ISSN 0975-6299 Vol 3/Issue 2/April – June 2012

International Journal of Pharma and Bio Sciences

RESEARCH ARTICLE BIOINFORMATICS

GLUCOSE-6-PHOSPHATE DEHYDROGENASE DEFICIENCY BY METHYLATION MEDIATED GENE SILENCNG PROVIDES RESISTANCE AGAINST FALCIPARUM MALARIA

TABISH QIDWAI AND FARRUKH JAMAL*

Department of Biochemistry (DST-FIST & UGC-SAP Supported), Dr. Ram Manohar Lohia Avadh University, Faizabad-22400, U.P., India

FARRUKH JAMAL
Department of Biochemistry (DST-FIST & UGC-SAP Supported), Dr. Ram Manohar Lohia Avadh University, Faizabad-22400, U.P., India

ABSTRACT

Malaria is the third most obivious cause of mortality in the world and is an example of evolutionary selection. Several host genetic factors have been selected in response to *P. falciparum* infection, such as hemoglobin variants, glucose-6-phosphate dehydrogenase (G6PD) and pyruvate kinase deficiency. Among these, G6PD deficiency is an important factor that provides resistance against malaria. Therefore, it is important to evaluate the effect of factors causing the deregulation of gene contributing to the deficiency of enzyme. Epigenetic changes are the most important factor that causes the gene silencing. We have predicted the methylation of CpG Iceland in G6PD and compared it with other genes. Based on this prediction, our hypothesis is that high promoter methylation causes epigenetic silencing of *G6PD* gene, which provides resistance against the falciparum malaria.

KEYWORDS

G6PD deficiency, DNA methylation, Plasmodium falciparum, malaria, epigenetic

INTRODUCTION

Malaria is one of the most studied infectious disease affecting 500 million people each year and is a leading cause of death in the global human population[1]. Although epidemics are prevalent in many areas of the world, Africa is the most strongly affected by this disease[2]. In India, several endemic regions like Rourkela, Chhattisgarh and Madhya Pradesh have been reported. Several human genes show a strong association with malarial resistance[2, 25]. One major candidate gene encodes glucose-6-phosphate dehydrogenase (G6PD [MIM 305900]), an important "housekeeping" enzyme in the glycolytic pathway for glucose metabolism. G6PD also plays a critical role in maintaining the balance of NADPH, a necessary cofactor for cell detoxification. G6PD is the sole generator of NADPH in the red blood cells and alone may prevent oxidative damage and severe anemia. G6PD deficiency is the most common known enzymopathy and is estimated to affect 400 million people worldwide. More than 130 different G6PD variants that lead to reduced enzyme activity have been discovered at the DNA level[3]. The deficiency of this enzyme is not only associated with many clinical disorders, such as neonatal jaundice, hemolytic anemia and several cardiovascular diseases but also with the distribution of malarial endemicity; and many variants are found at rare-to-high frequencies in different populations [4,5]. A single amino acid replacement is responsible for the classic G6PD A/B polymorphism. The B variant, which possesses normal enzyme activity and dominates in frequency worldwide, is predicted to be the ancestral state by comparison with chimpanzee *G6PD* [6]. The A variant, which is due to a derived amino acid replacement in exon 5, possesses 85% enzyme activity and is found in sub-Saharan Africa at frequencies as high as 40% but rarely reaches frequencies 11% outside Africa and the Middle East[4,7]. Several other genetic factors also provide resistance against malaria such as hemoglobin variants and pyruvate kinase deficiency, α- and β-thalassaemias. High frequency ha been observed for sickle cell haemoglobinopathy (0–22.4%) and G6PD deficiency (4.3–17.4%), with beta-thalassemia trait (0–8.5%) taking almost an intermediate position. For G6PD deficiency, hemizygous males as well as female heterozygotes and female homozygotes were detected [8]. Pyruvate kinase deficiency is the second most common erythrocyte enzyme disorder in humans. Report showed that pyruvate kinase deficiency provides protection against infection and replication of *P. falciparum* in human erythrocytes. The pyruvate kinase deficiency has a protective effect against replication of the malaria parasite in human erythrocytes.

THE HYPOTHESIS

In mammalian genome, methylation takes place mostly at cytosine bases in the CpG dinucleotide. Half of the genes in mammalian genome in the proximal promoter regions (0.5-4 kb) have CpG Iceland. In normal cells, these are unmethylated whereas DNA methylation in the promoter of certain genes is associated with transcriptional silencing. Methylation affects gene expression directly by interfering with transcription factor binding and/or indirectly by recruiting histone deacetylases through methyl-DNA binding proteins. DNA methylation in promoter regions is strongly correlated with the absence of gene expression. We have predicted the methylation of CpG Iceland in G6PD and compared it with other genes. Based

on this prediction, our hypothesis is that high promoter methylation causes epigenetic silencing of *G6PD* gene, which provides resistance against the falciparum malaria.

EVALUATION OF HYPOTHESIS

Prediction of methylation status of CpG Icelands:

Bhasin *et al.* designed the software Methylator to predict the methylation status of a single CpG using a much smaller window[9]. It has a publicly available web server (http://bio.dfci.harvard.edu/Methylator). We have taken upstream sequences from the (TRED) transcriptional regulatory element database and run through the Methylator to predict the methylation status (**Table 1).**

Table 1
Percentage Methylation of different genes (-2000 to +1 with respect to TSS) predicted using Methylator software

S.N.	Gene	Pathway	%Methylation
1.	TNF-alpha	Inflammation	0.84
2.	**G6PD**	**Carbohydrate metabolism**	**2.24**
3.	IL-10	Inflammation	0.399
4.	GSTP1	Detoxification	1.54
5.	GSTM1	Detoxification	1.54
6.	Pklr	Carbohydrate metabolism	0.59
7.	PNMT3B	Signaling	1.54
8.	CDK3	Signaling	1.799

Prediction of conservation pattern in different organisms:

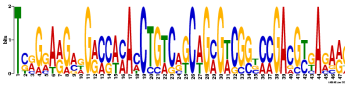
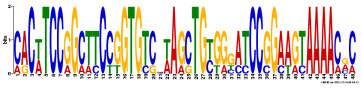
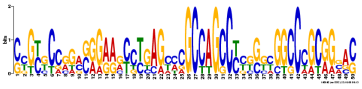
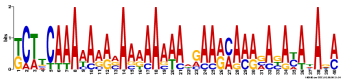
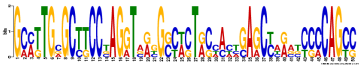
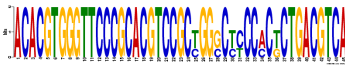
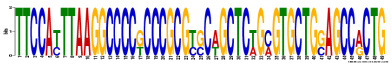
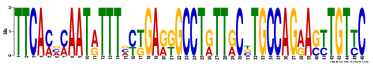
Ghosh D., (2004) developed MEME software that has been used for detection of conserved sequences in the promoter region[10]. We have access the promoter sequences of different organisms from Transcription Regulatory Element Database (TRED) and NCBI. A set of promoter sequences of different organisms run through MEME software. We have access some promoter sequences of the same gene in different organisms from NCBI and form a set of promoters for the same gene in different organisms. These promoter sequences were analyzed through MEME software **(Table 2 and Table 3)**.

Table 2
Training set for MEME software analysis

Sequence Name	Description	E-value	Length
ref|NW_879563.1|CfaX_W	C.lupus G6PD gene region,...	9.9e-66	1120
ref|NC_007331.2|NC_00	B.taurus G6PD gene region...	2.9e-56	1050

ref|NC_000023.9|NC_000	H.sapiens G6PD gene regio...	6e-53	1120
ref|NC_000086.6|NC_00	M.musculus G6pdx gene reg...	5.2e-49	1050
ref|NC_005783.4|:c235	E.gossypii AGOS_ABL206C g...	1.6e-33	1050
ref|NC_007134.3|NC_00	D.rerio LOC100148915 gene...	1.3e-11	1050
ref|NC_007134.3|NC_00	D.rerio wu: fj78b06 gene r...	4.6e-11	1120
ref|NW_047320.1|:20972	N.crassa NCU09111.1 gene ...	1.1e-09	1050
ref|NC_001146.3|:c1989	S.cerevisiae ZWF1 gene re...	2.8	1050
ref|NC_004354.3|:1956	D.melanogaster Zw gene re...	8.3	1120

Table 3
The MEME results of G6PD and different genes

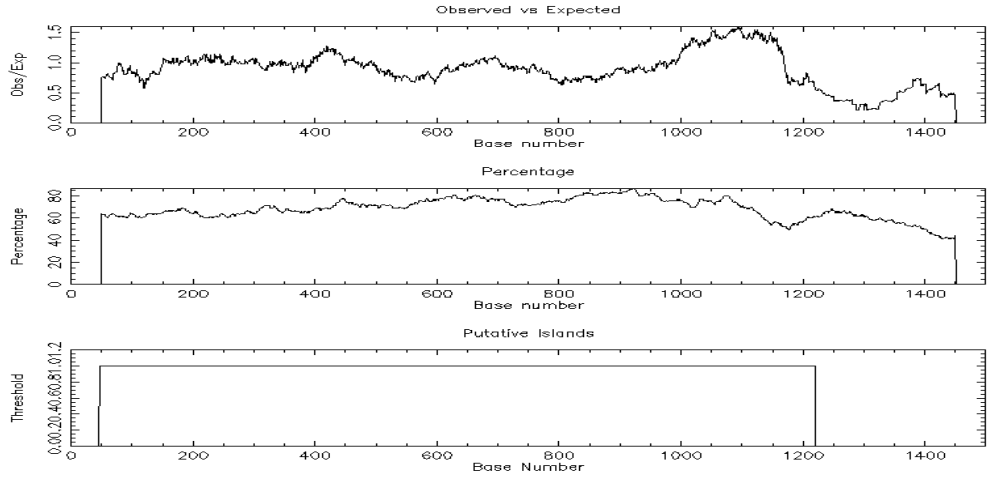
Gene	Logo1	Logo2	Logo3
G6PD genes in different organism as in table 2			
GSTP1 (1100 bp) In 3 organisms(rat, mouse,humans)			
GSTP1 (Varying length) In 7 organisms			
GSTP1 (700bp) In 7 organisms			
Ldha Human, rat, Mouse			

Prediction of CpG Icelands by using CpG Iceland searcher:

The upstream sequences were taken from the TRED and run through the CpG Iceland searcher which gives the CpG Iceland in the in the upstream sequences **(Figure 1 and Figure 2)**.

The CpG Iceland Searcher is available on the World Wide Web at http://www.cpgIcelands.com or http://www.uscnorris.com/cpgIcelands/cpg.cgi

.



CPGPLOT Icelands of unusual CG composition 151360414 from 1 to 1500
Observed/Expected ratio > 0.60, Percent C + Percent G > 50.00
Length > 200 Length 1174 (48..1221)

Figure1.
CpG Iceland profile of G6PD as predicted by EMBOSS software (-1000-500)BP

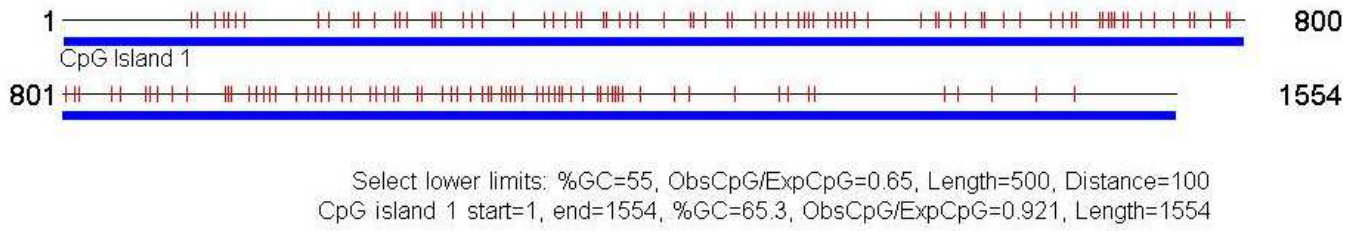


Figure 2
CpG Iceland profile of G6PD predicted by CpG Iceland searcher(-1000-500)BP

Prediction of CpG Icelands by using EMBOSS software:
The promoter sequences were taken from the TRED database. These promoter sequences were run in the EMBOSS software for prediction of CpG Iceland **(Figure 3a and 3b)**.

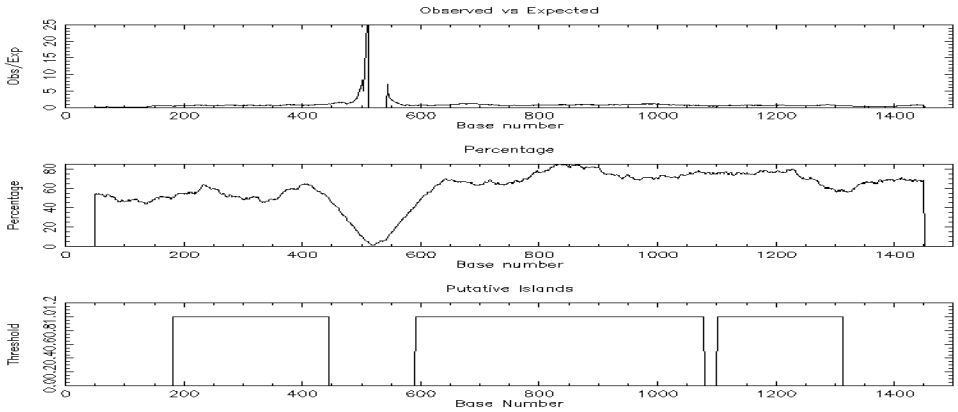


Figure3a.

CpG Iceland profile of GSTP1 as predicted by EMBOSS software(-1000-500)BP

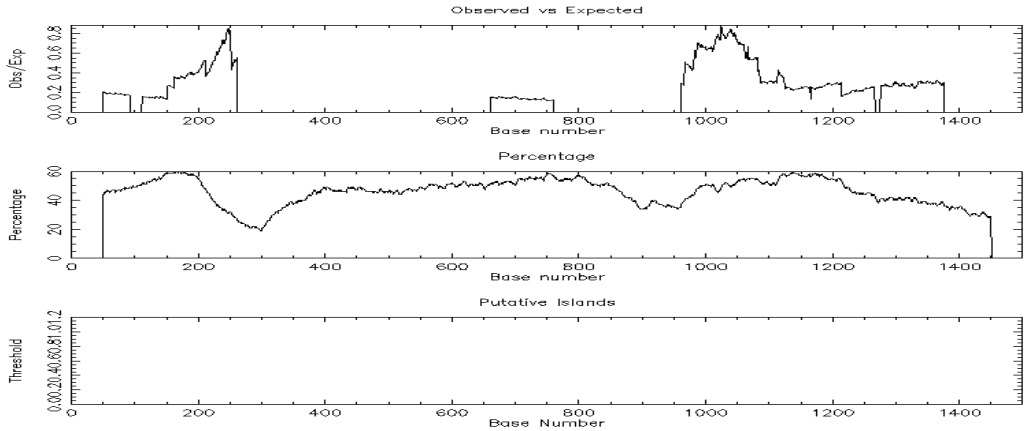


Figure3b.

CpG Iceland profile of NOS2A as predicted by EMBOSS software (-1000-500)BP

DISCUSSION

The prevalence of G6PD deficiency has been extensively studied in India but there is no clear evidence establishing the correlation of deficiency with *P.falciparum* malaria in the country[8, 11]. G6PD deficiency is a known red cell defect that confers resistance to malaria. It has been reported that the percentage deficiency of G6PD was higher in the malarial endemic regions as compared to non-endemic regions. The frequencies of low-activity alleles of glucose-6-phosphate dehydrogenase in humans are highly correlated with the prevalence of malaria. Despite the hemopathologies that they cause these deficiency alleles are thought to provide reduced risk from infection by the *Plasmodium* parasite and are maintained at high frequency [12]. We have predicted the methylation status of CpG Icelands in *G6PD* and compare it with those genes in which promoter hypermethylation causes gene silencing. Methylation affects gene expression directly by interfering with transcription factor binding and/or indirectly by recruiting histone deacetylases through methyl-DNA-binding proteins[13].

DNA methylation is involved in the regulation of many genes. At present, there are two major views regarding the causal relationship between DNA methylation and transcription. One popular view is that DNA methylation could repress

transcription by either or both of the following two mechanisms: (1) the methyl group may disrupt the binding sites of the transcription factors and result in the failure of transcription[14,15]; (2) methylated cytosines may attract methyl-CpG binding domain proteins (MBD) which would bring repressors to silence the chromatin [16]. In contrast, the alternative view argues that the transcriptionally inactive chromatin domains may be the true cause as they can serve as the targets for *de novo* methylation[17]. Regardless of which review is more accurate, it is clear that the methylation status of a CpG Iceland is highly correlated with transcription. Studying such epigenetic events is crucial for understanding of transcriptional regulation in development and differentiation. It has been reported that GSTP1, GSTM1, PNMT3B, CDK3 genes are transcriptionally repressed under some cancers because of high promoter methylation.

We calculated the GC content and CpG density of all the studied amplicons and compared it with DNA methylation. We found high promoter methylation in *G6PD.* Our hypothesis is that, due to high promoter methylation in *G6PD* as compared to *GSTP1* and *GSTM1,* it is more prone to epigenetic silencing leading to *G6PD* deficiency. Such epigenetic silencing of *G6PD* gene causes G6PD deficiency which provides resistance against malaria.

DNA methylation plays an important role in development and disease processes. For evaluation of its biological role, DNA methylation patterns of 190 gene promoter regions on chromosome 21 in five human cell types has been analyzed[18]. DNA methylation acts in a switch-like manner. Consistent with the well established role of DNA methylation in gene silencing, they found that DNA methylation in promoter regions strongly correlated with absence of gene expression and low levels of additional activating epigenetic marks. Inverse correlation of DNA Methylation with gene expression is known to lead to gene silencing[19]. The distribution of RNA polymerase II pre-initiation complex (PIC) is another indicator of gene expression. Kim *et al.* mapped the PIC binding sites across the genome by immunoprecipitation of TFIID-bound DNA from primary fibroblast cells[20]. They extracted all 92 PIC binding positions that are within 2.5 kb of the annotated TSS of genes on chromosome 21. These PIC binding sites are related with 69 gene promoters, out of which 67 genes are with DNA methylation data. 95% of these gene promoters (88 out of 93) exhibit low levels of DNA methylation (30%) in fibroblast. Comparing with the proportion of low methylated genes in the whole dataset of fibroblast (67%), the absence of DNA methylation is highly significant in the genes occupied with PIC. The absence of methylated genes in the PIC occupancy data set is consistent with the inhibitory role of DNA methylation on gene expression. Our data confirms that high levels of DNA methylation in CpG rich promoters are strongly associated with down-regulation of gene expression. However, the inverse relation does not hold, because there are many examples of unmethylated genes that are not expressed, possibly due to lack of expression caused by mechanisms other than DNA methylation like absence of the relevant activating transcription factors. The amplicons with high GC content and CpG density tend to be low methylated while those with low GC content and CpG density tend to be highly methylated, which is consistent with previous results [21, 22, 23]. The correlation between periodic distribution of CpG and DNA methylation has been reported recently [24]. Several reports observed a strong anticorrelation of DNA methylation with CpG density and GC content, indicating that DNA methylation declines, the more the sequence resembles a CpG Iceland [18, 21-23]. Thus, the GC content and CpG density are two critical properties for the biological effects of CpG Icelands.

ACKNOWLEDGEMENT

Facilities provided by the Department of Biochemistry, Dr. RML Avadh University(DST-

FIST and UGC-SAP supported) Faizabad and CIMAP, Lucknow for carrying out this work is gratefully acknowledged

REFERENCES

1. Greenwood B and Mutabingwa T, Malaria in 2002. Nature, 415(6872):670–672 (2002).
2. Miller LH, Impact of malaria on genetic polymorphism and genetic diseases in Africans and African Americans. Proc Natl Acad Sci USA, 91(7):2415–2419(1994).
3. Luzzatto L, Mehta A and Vulliamy T, Glucose 6-phosphate dehydrogenase deficiency. New York, McGraw Hill, pp.4517-4555(2001).
4. Beutler E, G6PD deficiency. Blood, 84(11):3613–3636(1994).
5. Vulliamy TJ, Mason P and Luzzatto L, The molecular basis of glucose-6-phosphate dehydrogenase deficiency. Trends Genet, 8(4):138–142(1992).
6. Kay AC, Kuhl W, Prchal JT and Beutler E, The origin of glucose-6-phosphate dehydrogenase (G6PD) polymorphisms in Afro-Americans. Am J Hum Genet, 50(2):394–398 (1992).
7. Ruwende C, Khoo SC, Snow RW and Yates SN, Kwiatkowski D, Gupta S, Warn P, Allsopp CE, Gilbert SC, Peschu N, and et al., Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. Nature, 376(6537):246–249(1995).
8. Balgir RS, Do tribal communities show an inverse relationship between sickle cell disorders and glucose-6-phosphate dehydrogenase deficiency in malaria endemic areas of Central-Eastern India? Homo, 57(2):163-176 (2006).
9. Bhasin M, Garg A and Raghava GP, PSLpred: prediction of subcellular localization of bacterial proteins. Bioinformatics, 21(10):2522-2524 (2005).
10. Ghosh D, Mixture models for assessing differential expression in complex tissues using microarray data. Bioinformatics, 20(11):1663-1669 (2004).
11. Sukumar S, Mukherjee MB, Colah RB and Mohanty D, Molecular basis of G6PD deficiency in India. Blood Cells Mol Dis, 33(2):141-145(2004).
12. Tishkoff SA, Varkonyi R, Cahinhinan N, Abbes S, Argyropoulos G, Destro-Bisol G, Drousiotou A, Dangerfield B, Lefranc G, Loiselet J, Piro A, Stoneking M,Tagarelli A, Tagarelli G, Touma EH, Williams SM and Clark AG, Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. Science, 293(5529):455–462(2001).
13. Zhu WG, Srinivasan K, Dai Z, Duan W, Druhan LJ, Ding H, Yee L, Villalona-Calero MA, Plass C and Otterson GA, Methylation of adjacent CpG sites affects Sp1/Sp3 binding and activity in the p21(Cip1) promoter. Mol Cell Biol, 23(12): 4056-4065(2003).
14. Kim J, Kollhoff A, Bergmann A and Stubbs L, Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene. Hum Mol Genet, 12(3): 233–245(2003).
15. Strunnikova M, Schagdarsurengin U, Kehlen A, Garbe JC, Stampfer MR and Dammann R, Chromatin inactivation precedes de novo DNA methylation during the progressive epigenetic silencing of the RASSF1A promoter. Mol Cell Biol, 25(10): 3923–3933(2005).
16. Bird AP and Wolffe AP, Methylation-induced repression-belts, braces, and chromatin. Cell, 99(5): 451–454(1999).
17. Bird A, DNA methylation patterns and epigenetic memory. Genes Dev, 16(1): 6-21 (2002).

18. Zhang Y, Rohde C, Tierling S, Jurkowski TP, Bock C, Santacruz D, Ragozin S, Reinhardt R, Groth M, Walter J and Jeltsch A, DNA Methylation Analysis of Chromosome 21 Gene Promoters at Single Base Pair and Single Allele Resolution. PLoS Genet. 2009 Mar; 5(3):e1000438. Epub 2009 Mar 27.

19. Klose RJ and Bird AP, Genomic DNA methylation: the mark and its mediators. Trends in Biochemical Sci, 31(2): 89–97(2006).

20. Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD and Ren B, A high resolution map of active promoters in the human genome. Nature, 436(7052): 876–880(2005).

21. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M and Schübeler D, Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet, 39(4): 457–466 (2007).

22. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R and Lander ES, Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature, 454(7205): 766–770 (2008).

23. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA, Haefliger C, Horton R, Howe K, Jackson DK, Kunde J, Koenig C, Liddle J, Niblett D, Otto T, Pettett R, Seemann S, Thompson C, West T, Rogers J, Olek A, Berlin K and Beck S, DNA methylation profiling of human chromosomes 6, 20 and 22. Nat Genet, 38(12): 1378–1385(2006).

24. Jia D, Jurkowska RZ, Zhang X, Jeltsch A and Cheng X, Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature, 449(7159): 248–251(2007).

25. Miller LH, Baruch DI, Marsh K and Doumbo OK, The pathogenic basis of malaria. Nature,415(6872):673–679(2002).