Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/

RJLBPCS

ISSN 2454-6348

Original Research Article **DOI: 10.26479/2018.0404.05**

COMPARATIVE IN-SILICO ANALYSIS OF *PLASMODIUM VIVAX* SAL-I AND HUMAN MCM PROTEIN FAMILY

Neelam Sehrawat[1]* Hitesh Singh[1], Surendra Gakhar[2]

1. Deptt. of Genetics, Maharshi Dayanand University, Rohtak, Haryana, India.

2. Centre for Medical Biotechnology, Maharshi Dayanand University, Rohtak, Haryana, India.

ABSTRACT: Mini-chromosome maintenance (MCM) protein 2-9, a subgroup of AAA+ ATPase family. MCM proteins act as replicative helicase. These proteins are functionally responsible for the DNA unwinding at the time of DNA replication. MCM proteins are not responsible for DNA replication, also have role in transcription, translation and regulation of cell cycle and also inter-related with the human carcinogenesis. All the MCMs contain a conserved regions approx of ~200-250 amino acids, which is site for nucleotide binding. They have a nearly conserved Walker A motif, Walker B motif and Zinc finger motif. All MCM are localized to nucleus only. In this article we focus on members of this important family of MCM proteins from the malaria parasite *Plasmodium vivax* and their comparison with the human host.

KEYWORDS: DNA helicase, malaria, *Plasmodium vivax,* replication, helicase etc.

Corresponding Author: Dr. Neelam Sehrawat* Ph.D.

Deptt. of Genetics, Maharshi Dayanand University, Rohtak, Harayana, India.

Email Address: neelamsehrawat@gmail.com

1. INTRODUCTION

MCM (minichromosome maintenance) were first identified, to be involved in DNA replication as the result of a genetic screening for *Saccharomyces* cerevisiae mutant that are defective in minichromosome maintenance [1]. The best known among these family are the MCM2–7 proteins,

a family of six conserved proteins that are the key components of the replication initiation complex that initiates DNA synthesis in all eukaryotes [1]. MCMs are evolutionary conserved in all eukaryotes and archaea. MCM1 is the universal transcription factor which involve in the regulation of some other MCM and other transcriptional factors [3]. MCMs are highly and significantly up-regulated in human meningioma tumor, a significant increase in MCM2 (8 fold) and MCM3 (5 fold), MCM4 (4 fold), MCM5 (4 fold), MCM6 (3 fold), MCM7 (5 fold) expressions in meningiomas [4]. They are activated by cyclin dependent kinases, such as Cdc6, Cdt1 and Dbf4/Cdc7 in the early G1 phase of the cell cycle to form the origin complex called the pre-replication complex (pre-RC) [4]. Approx 40 MCMs are expressed in abundance in all phases of the cell cycle and degraded in quiescence, senescence and differentiation steps thus they can be used as a specific markers of the cell cycle state in tissues.MCM are categorized in to two groups, one group by MCM (2-7) and other byMCM8 and MCM9, which are found only in higher eukaryotes. Beside their role in DNA replication, MCM plays role in other cellular activities. This speculation was supported by two observation referred to as the ´MCM paradox': at least in the yeast, Xenopus, Drosophila, reduced concentration of MCM doesn't impair DNA replication [5] and majority of the MCM do not localize at the DNA syntheses in mammalians cells[6]. MCM subunits play essential role in DNA replication, deletion or default in any subunit causes the cell death or apoptosis.MCM 2-7 contains some evolutionary conserved sequences, 250- amino acid sequence that encoded for the ATPase active site (AAA Domains). They contain ring-shaped P-loop NTPases, which exert their activity through energy dependent remodeling or translocation of macromolecules [7, 8]. As with all P-loop NTPases, members of this group possess a core αβα nucleotide-binding domain which contains two major nucleotide-binding and hydrolysis motifs referred to as Walker A (P-loop), Walker B and arginine finger [18]. DNA synthesis in eukaryotes is a complex, multistep process that requires the participation of a number of MCM proteins. This process involves the binding of ORC to replication origins and the recruitment of Cdc6 and MCM2–7 to form the pre-replicative complex (pre-RC) and the activation of pre-RC by Cdc7 and Cdc28 protein kinases to initiate DNA synthesis [10]. Mcm1 in the regulation of replication initiation remains to be investigated. Mcm10 is another replication initiation factor that intimately interacts with the MCM2–7 proteins in replication initiation [9]. It should be pointed out here that Mcm1 and Mcm10 bear no sequence homology to the MCM2–7 family. Cdc23/ Mcm10 function is conserved between fission yeast and Xenopus [10], where in vitro analysis has indicated a similar requirement for Cdc45 binding, but apparently not compared with *Saccharomyces cerevisiae*, where Mcm10 is needed forMcm2 chromatin binding [10]. However,

unlike the situationin Xenopus, where Mcm 10 chromatin binding is dependenton Mcm2–7 [10],it showed that the fission yeast protein is bound to chromatin throughout the cell cycle in growing cells, and only displaced from chromatin during quiescence [10]. MCM proteins are isolated from various plants, yet their role is to be established [11]. Beside their role in DNA replication, MCM plays various roles, they most likely role in plant stress tolerance in some plants. Recently reported that in pea plant MCM6 overproduce in response of high salinity and cold stress [12]. The overexpression of pea MCM6 single subunit in tobacco plant promotes salinity stress tolerance without affecting its yield [12]. Basic understanding of the molecular basis of life cycle, cell growth and differentiation of malaria parasite is the essential key for the development of novel chemotherapeutic agents for the control of malaria. Plasmodium has complex life cycle in mosquito and human host [13].There are five point in plasmodium in which DNA replication occurs [13, 14]. MCM proteins are required for the pre-intiation complex in DNA synthesis. Although plasmodium has tight regulation on all processes, however little is known about the regulatory mechanism of replication process. ORC1 and MCM4 have role in pre-replication formation and both expressed only in gametocytes [15, 16]. However all six subunits of the MCM complex has been reported in *P. falciparum* [17]. There are 8 MCM (2-9) present in *P. falciparum* [18]. *Plasmodium vivax* is the next important strain of plasmodium causing high mortality than *P. falciparum*. Instead of *P. vivax* importance, there is no prior study of MCMs of P. *vivax*. Hence, in the present study we have made an effort to compare in-silico MCM family of proteins from *P. vivax* and their comparison with human host.

2. MATERIALS AND METHODS

a) Identification of putative MCM family genes in *Plasmodium vivax*

The complete set of predicted genomic, CDs, transcript and protein sequences from the open reading frame (ORFs) of the *Plasmodium vivax* genome has been obtained from PlasmoDB [http://plasmodb.org/] version 9.2. Gene text search was primarily used to collect putative MCM family genes from PlasmoDB.

b) Analysis of MCM family genes

All the information about MCM family genes regarding gene ID's, chromosomal location, genomic position, number of introns and exons, nucleiotide sequence length, molecular weight, amino acid sequence length, no. of MCM gene domains and isoelectric points was extracted from plasma DB and compiled in a table.

c) Domain Architecture Analysis

To identify the potential domains of MCM proteins (Walker motif A, Walker motif B and arginine finger motif) encoded by putative MCM family, we opted for SMART 7(Simple Modular Architecture Research Tool) [19], Pfam database of protein families that includes there annotations and multiple sequence alignment generated using hidden markov model and Interpro [20]. The conserved domain of the protein is known as fingerprint of that protein family, which provides distinctive signature and structural/functional domain.

d) Prediction of sub-cellular Localization MCM family proteins

Various online software like PlasMit, PlasmoAP and Euk-mploc 2.0 Server were used to predict Subcellular localization of MCM proteins in *Plasmodium vivax*. Euk-mploc 2.0 server database is use for predicting subcellular of eukaryotic proteins (www.csdio.sjtu.edu.cn/bioinf/euk-multi 2.0) version 2.0 [21]. Mitoprot (ihg.gsf.de/ihg/mitoprot.html) was used for prediction of mitochondrial proteins [22] and PATS (www.patshow.co.uk) was used to identify apicoplast targeting proteins [23]. Protein sequences in FASTA format were submitted to the server and localization results were displayed.

e) Human homolog search

The downloaded sequences of MCM were used as query and then matched with human homolog using BLAST search (www.ncbi.nlm.nih.gov). The corresponding human sequences were retrieved and their conserved domains were searched by using online CD search tool available on www.ncbi.nlm.nih.gov/structural/cdd/wrpsb.cgi. Human homologs were also identified by using BLASTp of *P. vivax* MCMs as query with Human genome database.

f) Prediction of Phosphorylation sites

Phosphorylation sites in Hs and PvxMCMs were find out by submitting each MCM sequence to www.cbs.dtu.dk/services/NetPhosK for prediction of kinase specific eukaryotic protein Phosphorylation sites.

g) Phylogenetic analysis

In order to identify the closely related homologs of the new family of MCM gene of *P. vivax* in other organisms, PSI-BLAST search on the non redundant database [NRDB available at the National Centre for Biotechnology Information (NCBI)] of proteins have been carried out using the MCM gene family as a query. The nucleotide sequences matching the query with highest sequence identity and with a reliable E-value (<0.0001) have been further extracted from each of these searches to collect the close homologs of the new MCM gene. Multiple sequence alignment of the MCM gene

and their most closely related homologs has been further carried out using CLUSTALW [http://clustalw.org/] version 2.1.A rooted neighbor joining phylogeny tree for *Plasmodium vivax* MCM genes and homologous genes was constructed by using MEGA 6.06.

h) 3D Structure prediction

The *P. vivax* MCMs genes showing potential homology with human homolog were used for the construction of 3D structure by submitting *P. vivax* MCM sequence as target and human MCM structure as a template by using SWISS-Model server.

3. RESULTS AND DISCUSSION

The genome of *P. vivax*, available at www.plasmodb.org was investigated using 'MCM' as query. The results of this search presented in Table -1. All the MCM members from this family were used to BLAST with *H. sapiens* and the comparative analysis is carried out. MCM2-9 proteins show sequence conservation particularly in a 200-250 amino acid residue domain, which is located almost in the center of these large proteins. Walker A motif contains the consensus sequence [(G)xxxxGK[S/T], where x is any residue and lysine residue is present which is characteristic of all the ATP-binding proteins. It has been observed that the Walker A motif sequence in all the MCMs is slightly deviated from the consensus and the glycines in the motif GK(S/T) are substituted by serine or alanine [24, 18]. The Walker B motif [hhhh(D/E)] has conserved nucleotide phosphate-binding motif (where h is a hydrophobic residue). These are the characteristic feature of members of the P-loop NTPase domain superfamily. The Walker A motif binds the β-γ phosphate moiety of the bound nucleotide (typically ATP or GTP) and Walker B motif binds with the Mg2+ cation [25]. All the MCMs contain a Zn finger motifs and another short motif SRFD, which is present approximately 70 residues after the Walker B motif and defines an arginine finger motif.

MCM1:-MCM1 Protein- A sequence-specific DNA-binding protein that plays an essential role as a global regulator of yeast cell cycle control [1]. It contains a MADS-box domain within the N-terminal 56 amino acids. It is one of the four founder proteins that structurally define the superfamily of MADS domain proteins. MCM1 found only in the Humans not in the *Plasmodium vivax* and *P. falciparum*.

Table 1- Comparison of MCM gene family of *Plasmodium vivax Sal-1* and Human homolog

Sr. No	Protein Name	Gene Id (Pvx/Hs)	Location (chrom. No/Position)	Size in Kda (Pvx/Hs)	Exon/intron in Pvx	% Identity with HsMCM
1	MCM2	PVX_085565/ENSP000000265056	13/1,2820411 to 285,468	111/91	2/1	29
2	MCM3	PVX_079890/ENSP000000229854	10/181,777 to 184,812	110/91	2/1	29
3	MCM4	PVX_122675/ENSP000000430329	14/ 772,315 to 775,182	108/80	1/0	33
4	MCM5	PPVX_084615/ENSP000000412847	13/482,362 to 486,858	85/82	5/4	43
5	MCM6	PVX_114735/ENSP000000264156	11/679,166 to 682,690	107/92	1/0	39
6	MCM7	PVX_087810/ENSP000000344006	1/185,096 to 188,086	94/81	1/0	40
7	MCM8	PVX_084595/ENSP000000478141	13/458,048 to 462,740	149/81	4/3	47
8	MCM9	PVX_089905/ENSP000000314505	5/926,346 to 930,278	145/127	1/0	49

MCM2: The protein encoded by this gene is one of the highly conserved mini-chromosome maintenance proteins (MCM) that are involved in the initiation of eukaryotic genome replication. This family is also present in the archea bacteria in 1 to 4 copies. *Methano caldococcus jannaschii* (*Methanococcus jannaschii*) has four members- MJ0363, MJ0961, MJ1489 and MJECL13. This protein forms a complex with MCM4, 6 and 7 and has been shown to regulate the helicase activity of the complex [26]. This protein is phosphorylated and thus regulated by protein kinases CDC2 and CDC7.It has been well established that HsMCM2 contributes to a variety of nuclear functions in addition to DNA replication. The detailed biochemical characterization of HsMCM2 showed that the C-terminal region of HsMCM2 contains ssDNA- binding activity that inhibits the DNA helicase activity [27]. On the other hand using pulldown analysis it was reported that two fragments from the central region were mainly responsible for the interaction between HsMCM2 and HsMCM4 [27]. The gene with Plasmo DB number PVX_085565 is a homolog of human MCM2 and is located on chromosome 14. The MCM2 protein presents in the Plasmodium both are contains same weight while the human MCM2 is slightly less in weight than PVxMCM. It

contains additional 163 amino acid than human homologs as shown in table-1. As eukaryotic MCMs and PvX MCM also contains putative C4-type zinc finger domain at its N-terminal region. The zinc finger domain may have a role in the binding of PvxMCMs to chromatin because these domains are known to be responsible for protein-DNA and protein-protein interactions and therefore contribute to complex assembly similar to PfMCM2.

MCM3:- The study about the phosphorylation of HsMCM3 reported that cyclin B–CDK1 catalyzes phosphorylation of HsMCM3 at Ser-112, hence regulating HsMCM3 association with other HsMCM2–7 subunits and loading of HsMCM3 onto chromatin [28]. MCM3 protein is a subunit of the protein complex that consists of MCM2-9. It has been shown to interact directly with MCM5/CDC46. This protein also interacts with and is acetylated by MCM3AP, a chromatin-associated acetyltransferase [50]. The acetylation of this protein inhibits the initiation of DNA replication and cell cycle progression. Human MCM acts as component of the MCM2-7 complex (MCM complex) which is the putative replicative helicase essential for once per cell cycle DNA replication initiation and elongation in eukaryotic cells. The active ATPase sites in the MCM2-7 ring are formed through the interaction surfaces of two neighboring subunits such that a critical structure of a conserved arginine finger motif is provided in trans relative to the ATP-binding site of the Walker A box of the adjacent subunit. The six ATPase active sites, however, are likely to contribute differentially to the complex helicase activity.The gene with of human MCM3 homologs in *Pf* and *Pvx* are located in the chromosome 5 and 10 respectively. The protein of human MCM3 is slightly smaller in size than PfMCM3 and Pvx MCM3. MCM3 with Ki-67 used as marker in diagnosis of salivary gland tumours.

MCM4:- The Phosphorylation of this protein by CDC2 kinase reduces the DNA helicase activity and chromatin binding of the MCM complex [29]. This gene is mapped to a region on the chromosome 8 head-to-head next to the PRKDC/DNA-PK, a DNA-activated protein kinase involved in the repair of DNA double-strand breaks. The phosphorylation at sites 3 and 32 of HsMCM4 required CDK2 in HeLa cells and this phosphorylated HsMCM4 had several distinct and site-specific roles in MCM function [29]. It was reported that the central region of HsMCM2, which contains zinc finger and ATPase motif interacts with HsMCM4 [29]. The active ATPase sites in the MCM2-7 ring are formed through the interaction surfaces of two neighboring subunits such that a critical structure of a conserved arginine finger motif is provided in trans relative to the ATP-binding site of the Walker A box of the adjacent subunit[51]. MCM4 is an important gene associated with Natural Killer Cell and Glucocorticoid Deficiency with DNA repair Defect [30].The gene with

PlasmoDB number PVX_122675 is a homolog of human MCM4 and is located on chromosome 14 and, is larger in size as compared with its human homolog. It has high molecular weight to its human counterpart. PvxMCM4 contains some unique features such as it is the largest of all the MCM4 and contains insertions at few places within its entire sequence including the zinc finger domain. It was also reported that PfMCM4 is expressed specifically in the sexual erythrocytic stages indicating that PfMCM4 may be involved in gametogenesis where DNA is replicated [31].

MCM5:- The protein encoded by this gene is structurally very similar to the CDC46 protein from *S. cerevisiae* [32]. The human MCM5 gene was shown to be expressed widely in many normal tissues, but its mRNA levels vary in different condition. Cyclin-E is shown to directly interact with and co-localize on centrosomes with the DNA replication factor MCM5 in a CLS-dependent but Cdk2-independent manner [33]. The domain in MCM5 that is responsible for interaction with cyclin E is distinct from any previously described for MCM5 function and is highly conserved in MCM5 proteins from yeast to mammals. Expression of MCM5 or its cyclin E-interacting domain, but not MCM2, significantly inhibits over-duplication of centrosomes in CHO cells arrested in S-phase. The highest levels of MCM5 mRNA transcripts were detected in A-431 epidermoid carcinoma cells, U-2 OS osteosarcoma cells and U-251 MG astrocytoma cells [34]. Expression of all human gene of the MCM family is induced by growth stimulation and their mRNA levels peak at G1/S transition. The growth-regulation expression of MCM5 is primarily regulated by members of the E2F family through binding to multiple sites of the MCM5 gene promoter. The gene with PlasmoDB number PVX_084615 is a homolog of HsMCM5. PvxMCM5 is almost similar in size and is located on chromosome number 13.

MCM6:-It is well established that Cdt1 physically interacts with the MCM complex and this interaction mainly occurs between Cdt1 and MCM6 in human cells. The detailed analysis indicated that the C-terminal 79 residues of hCdt1 interact with the C-terminal 113 residues of HsMCM6 while the large N-terminal Orc6-binding domain recruits Cdt1/MCM2-7 to ORC complex [35]. HsMCM6 expression correlated with the tumor in craniopharyngiomas [36]. The gene with PlasmoDB number PVX_114735 is a homolog of HsMCM6. PVXMCM6 is having more15Kda molecular weight than human homolog. MCM6 is located on chromosome number 11 in *P. vivax*. The comparative analysis of the conserved motifs showed slight differences in their amino acid sequences in walker A motif whereas residues are conserved in remaining two motifs.

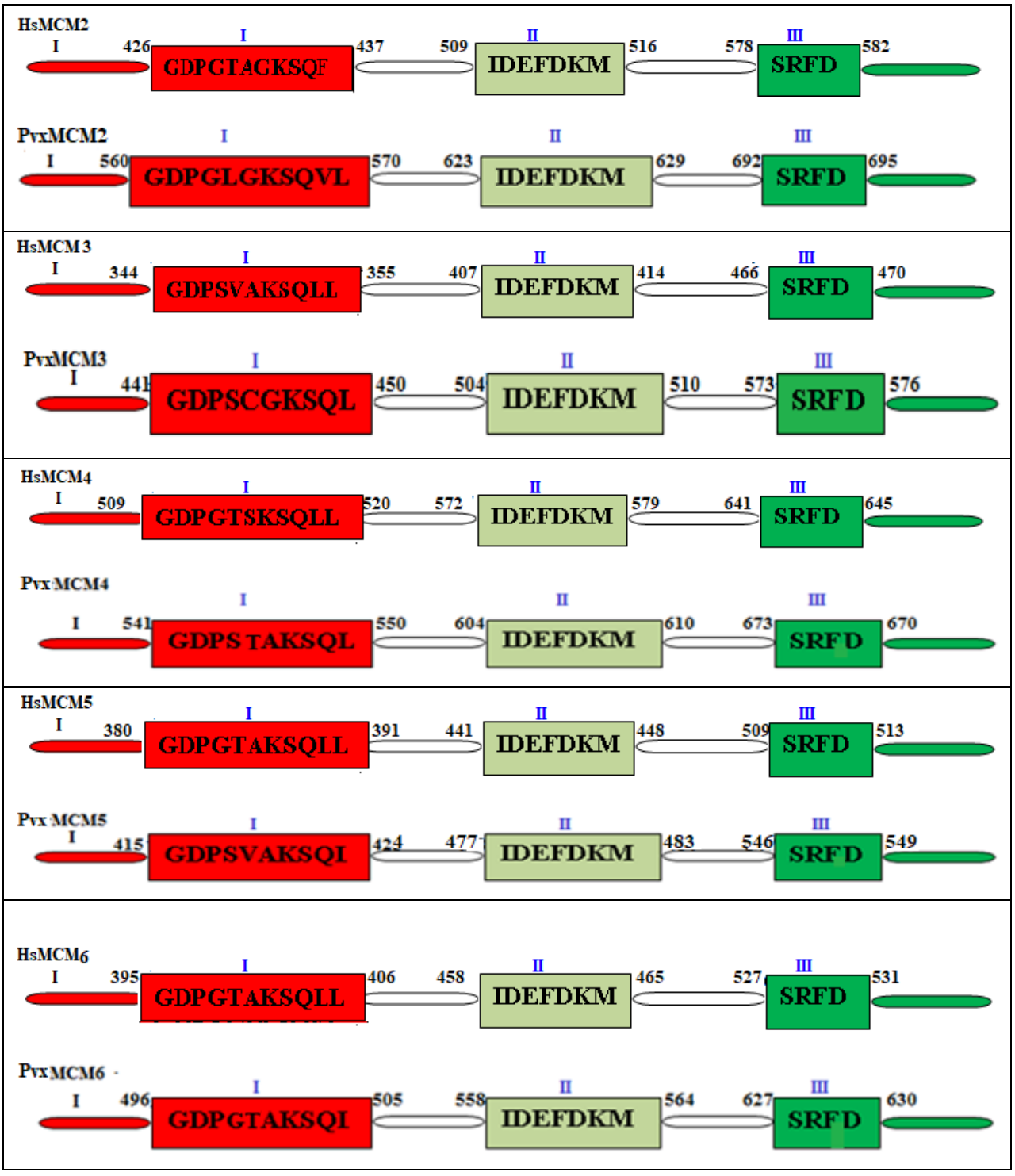
MCM7:-Recent study shows that MCM7 may be a useful proliferation marker in prostatic neoplasia [37]. The *MYCN* oncogene is amplified in ~25% of neuro blastoma tumors, Induction of *MYCN* in

conditional cell lines results in increased expression of endogenous MCM7 mRNA and a 3-fold [52]. MCM7 and MCM3 were identified as cyclin D1-binding proteins. Cyclin D1/CDK4 kinase binds to components of the MCM complex. Although the cyclin D1-dependent kinase did not phosphorylate MCM7, active cyclin D1/CDK4, but not cyclin E/CDK2, did catalyze the dissociation of an RB·MCM7 complex. The gene with plasmodb number PVX_087810 is a homolog of HsMCM7 and it is located on chromosome 1. PVX contains 14 additional amino acid in its MCM domain. The expression of PfMCM7 polypeptides was predominantly observed in late trophozoites and during schizont maturation [38] and decreased in the ring stages, which is in agreement with DNA replication in Plasmodium.

MCM8:- MCM8 is a new evolutionarily conserved family member but its homolog is not present in yeast [39]. MCM8 not take part in MCM2-7 pre-initiation complex. MCM8 mRNA is expressed in placenta, lung and liver, but is also significantly expressed in adult heart, a tissue with a low percentage of proliferating cells [53]. We report that the accumulation on chromatin of another member of the MCM protein family, human MCM8 (hMCM8), occurs during early G_1 phase, before the HMCM2-hMCM7 complex binds. HMCM8 interacts in vivo with two components of the pre-RC, namely, hcdc6 and hORC2 [54]. Depletion of endogenous HMCM8 protein by RNA interference leads to a delay of entry into S phase, suggesting a role for HMCM8 in G_1 progression. Furthermore, down-regulation of hMCM8 also leads to a reduced loading of hcdc6 and the hMCM2-hMCM7 complex on chromatin [40, 54]. The gene with PlasmoDB number PVX_084595 is a homolog of HsMCM8 and PfMCM8 is located on chromosome 13 (Table 1). There is a large difference in their size, as PvxMCM 8 is the largest MCM. PVXMCM8 contains additional 502 amino acid in comparison toHsMCM8.

MCM9:-In Xenopus egg extracts, MCM9 interacts with CDT1 to load MCM2-7 onto replication origins, and also counteracts the inhibitory effects of Geminin upon CDT1 for replication licensing.HsMCM9 is a novel member of MCM family [41, 42]. Similar to HsMCM8, HsMCM9 is only present in the genome of higher eukaryotes. It showed 24–31% total amino acid identity with HsMCM2–MCM8 proteins and contains a unique C-terminal domain which has only weak homology to MCM2-7 and MCM8 but is conserved within MCM9 homologs. Evolutionary history show that MCM9 closely related with MCM8.PlasmoDB number PVX_089905 is the homolog of HsMCM9 (Table 1). PVXMCM9 contains 171 additional amino acids as compared with HsMCM9. MCM9 are located on chromosome number 5 in *P. vivax*. The conserved motifs of PvxMCM and HsMCM were analyzed by using MEME Suite-GLAM2 version 4.8.0. The motif analysis indicated

that in *P. vivax*, walker A motif (motif I), 5 out of 11 residues are highly conserved and others are variable to some extent as shown in figure-1. The Walker B motif (Motif II) is highly conserved in MCM (2-8), whereas MCM 9 has 3 variable residues out of 7 residues. On the other hand Arginine motif (Motif III) has T instead of S in SRFD residues in MCM 7 and MCM 8 and others have conserved SRFD residues as shown in figure-1.



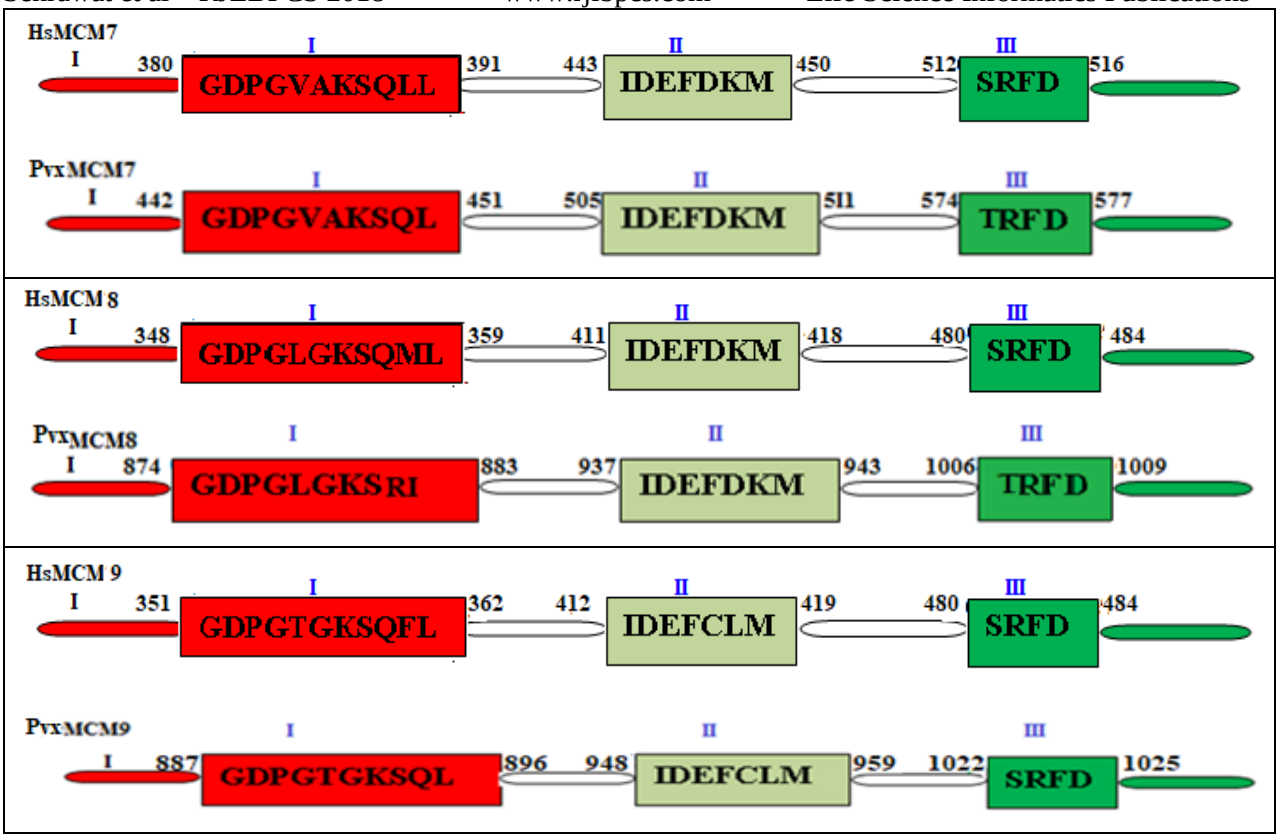


Figure1: Comparison of MCM (2-9) of *Plasmodium vivax* and Human homolog showing the aminoacid residues of Walker A motif (Motif-I), Walker B motif (Motif-I) and Zinc Finger Motif (Motif-III).

Post translational modification especially phosphorylation plays a very important role in regulating the function of various MCMs. The phosphorylation of MCM2, MCM3, MCM4, MCM6 and MCM7 has been observed in vivo and in vitro in different eukaryotic cells [43, 44, 45].Phosphorylation pattern of PfMCMs has also been discussed [18]. However the same is not reported in case of *P. vivax*. Therefore, the Phosphorylation potential of all the HsMCM and PvxMCM was analyzed using NetphosK at www.cbs.dtu.dk/services/NetPhosK. The results suggested that all members of MCM family are prone to Phosphorylation and contain recognition sites for PKC followed by PKA or CKII as shown in figure-2. Despite of the difference in size of MCM proteins in *P. vivax* and human, the Phosphorylation sites are almost equal in Hs and PvxMCM3, MCM5, MCM6, MCM8 and MCM9. Furthermore it is interesting to note that PvxMCM2 contain 8 sites as opposed to 12 sites in HsMCM2, PvxMCM4 contains 20 phosphorylation sites whereas HsMCM4 has 14 sites as shown in figure-2. The Subcellular localization was find out for PvxMCM proteins by using protein sequence of MCM 2-9 as a query for input to plasMit, PlasmoAP and Euk-mploc 2.0 Server www.csdio.sjtu.edu.cn/bioinf/euk-multi 2.0. The results shows that all MCM 2-9 are localized to the

nucleus only, therefore this analysis also strengthen the role of MCMs in DNA replication and transcription.

PvXMCM2

Site	Scores	Kinases
170-S	0.784	PKC
377-S	0.865	PKC
504-T	0.835	PKC
656-T	0.701	PKC
881-T	0.785	PKA
895-S	0.728	PKC
907-T	0.703	CKII
929-T	0.706	PKC

PvxMCM4

Site	Scores	Kinases
31-S	0.789	PKC
160-T	0.830	PKA
228-S	0.726	PKC
243-S	0.758	PKC
305-T	0.877	PKC
523-T	0.840	PKC
524-T	0.804	PKC
545-T	0.741	PKC
565-S	0.858	PKC
627-T	0.794	PKC
636-T	0.806	PKC
694-T	0.873	PKC
708-S	0.715	CKII
710-S	0.736	CKII
716-S	0.718	CKII
755-T	0.794	PKC
759-T	0.828	PKC
795-T	0.796	PKC
821-S	0.716	PKA
831-S	0.716	PKC

PvxMCM5

Site	Scores	Kinases
158-T	0.766	PKC
187-S	0.819	PKC
197-T	0.834	PKA
228-S	0.715	PKC
229-T	0.907	PKC
267-S	0.753	PKC
395-T	0.861	PKC
443-S	0.733	PKA

HsMCM2

Site	Scores	Kinases
43-S	0.77	CKII
133-S	0.75	PKA
190-S	0.73	PKA
213-S	0.74	PKA
431-T	0.73	PKC
529-S	0.76	PKC
586-S	0.71	PKC
709-T	0.71	PKC
713-S	0.80	PKC
717-S	0.80	PKC
721-T	0.84	PKC
727-S	0.71	PKC

HsMCM4

Site	Scores	Kinases
7-T	0.74	CDK5
13-S	0.84	PKA
34-S	0.77	PKC
96-S	0.75	PKC
268-T	0.85	PKC
317-S	0.85	PKC
424-T	0.86	PKC
442-S	0.00	PKC
488-T	0.76	PKC
495-T	0.79	PKC
514-T	0.83	PKC
538-S	0.70	PKC

PvxMCM6

Site	Score	Kinases
26-S	0.877	PKC
86-T	0.804	PKC
92 S	0.723	CKII
223-S	0.850	PKC
410-S	0.708	PKC
519-T	0.707	PKC
520-S	0.742	PKC
524-S	0.708	PKA
615-T	0.781	PKC
646-T	0.748	PKC
674-T	0.822	PKC
684-S	0.830	PKC
705-S	0.831	PKC

PVXMCM3

Site	Score	Kinases
76T	0.753	PKC
93T	0.768	PKC
219T	0.750	PKC
461T	0.806	PKC
465T	0.760	PKC
528T	0.892	PKC
602T	0.716	PKC
681T	0.879	PKC
756S	0.704	PKC
757S	0.822	PKC
845T	0.928	PKC
849T	0.934	PKC
850T	0.937	PKC
860T	0.771	PKC

HsMCM3

Site	Score	Kinases
90-T	0.792	PKC
108-S	0.74	PKC
151-T	0.73	PKC
160-S	0.84	PKC
164-T	0.76	PKC
191-S	0.72	PKC
240-T	0.84	PKC
255-T	0.75	PKC
288-T	0.90	PKC
340-S	0.75	PKC
361-T	0.71	PKC
369-T	0.75	PKC
373-S	0.77	PKC
374-S	0.79	PKC
555-T	0.91	PKC
672-S	0.77	CKII
674-T	0.79	CKII
711-S	0.74	CKII

HsMCM5

Site	Score	Kinases
52-T	0.90	PKC
116-S	0.77	PKC
139-S	0.72	PKC
166-T	0.80	PKA
278-S	0.73	PKC
292-T	0.74	PKA

612-S	0.735	PKA
630-S	0.823	PKC
632-T	0.791	PKC
662-T	0.711	PKC
680-T	0.842	PKC
716-S	0.822	PKA
754-S	0.780	PKC

PvxMCM7

Site	Score	Kinases
136-T	0.843	PKC
286-S	0.819	PKC
423-T	0.919	PKC
465-T	0.706	PKC
466-T	0.894	PKC
529-S	0.730	PKC
544-S	0.711	PKC
718-S	0.760	PKC
723-T	0.776	PKC
731-S	0.758	PKA
741-S	0.839	PKC
768-T	0.769	PKC

PVXMCM9

Site	Score	Kinases
20-S	0.734	CKII
63-S	0.875	PKC
265-S	0.756	PKC
288-S	0.712	PKC
300-T	0.764	PKC
393-T	0.863	PKC
478-S	0.753	PKC
522-T	0.735	PKC
657-T	0.747	PKC
739-S	0.734	PKA
742-S	0.726	PKA
827-S	0.787	PKC
891-T	0.810	PKC
921-T	0.801	PKC
999-T	0.746	PKC
1150-S	0.851	PKC
1172-S	0.737	PKA
1181-S	0.707	PKC
1193-T	0.780	PKC
1274-S	0.714	PKA

722-T	0.835	PKC
760-S	0.909	PKC
804-S	0.803	PKC
809-T	0.798	PKC
920-T	0.839	PKC

HsMCM6

Site	Score	Kinases
77-T	0.82	PKC
96-T	0.75	PKC
119-T	0.74	PKC
135-T	0.77	PKC
140-S	0.70	PKA
163-T	0.71	PKC
195-T	0.78	PKC
247-T	0.74	PKC
306-T	0.72	PKC
420-S	0.84	PKC
424-T	0.70	PKA
492-T	0.82	PKC
513-S	0.71	PKC
613-S	0.72	PKC
712-S	0.72	PKA
762-S	0.72	CKII

HsMCM7

Site	Score	Kinases
113-S	0.73	PKC
156-S	0.75	PKA
162-T	0.85	PKC
224-T	0.72	PKC
392-S	0.83	PKC
405-T	0.88	PKC
409-S	0.76	PKA
410-S	0.81	PKA
477-T	0.77	PKC
500-S	0.73	PKA

HsMCM8

Site	Score	Kinases
15-S	0.78	PKA
131-T	0.71	PKC
157-T	0.74	PKC
168-S	0.82	PKA
260-S	0.86	PKC
388-T	0.78	PKC

315-S	0.71	CKII
385-T	0.75	PKC
405-S	0.76	PKA
417-S	0.74	PKC
600-S	0.74	PKA
605-S	0.81	PKA
699-T	0.79	PKC

PVXMCM8

Site	Score	Kinases
346-S	0.705	PKC
420-S	0.728	PKA
436-S	0.707	PKC
459-S	0.742	PKA
539-T	0.733	PKC
582-S	0.730	PKC
659-S	0.728	PKA
839-S	0.869	PKC
981-S	0.726	PKC
015-T	0.806	PKC
015-T	0.806	PKC
1055-T	0.784	PKB
1129-S	0.723	CKII
1199-S	0.749	PKC

HSMCM9

Site	Score	Kinases
76-S	0.71	PKA
79-T	0.77	PKC
101-S	0.70	PKA
129-T	0.91	PKC
134-S	0.80	PKA
170-S	0.74	PKA
187-S	0.70	PKC
212-S	0.79	PKC
346-S	0.74	PKA
356-T	0.80	PKC
376-T	0.82	PKC
396-T	0.78	PKC
437-S	0.90	PKC
659-S	0.79	PKC
726-T	0.76	PKC
817-S	0.79	PKC
871-T	0.76	PKC
890-S	0.93	PKA
924-T	0.73	PKC

	387-S 0.73 PKC 444-S 0.72 PKC 516-S 0.75 PKA 519-T 0.81 PKC 595-S 0.71 PKA 610-S 0.77 PKC 670-T 0.89 PKC 676-S 0.76 PKA	1044-S 0.78 PKC 1087-S 0.91 PKA 1099-S 0.75 PKC 1143-S 0.84 PKA

Figure 2: Comparison of Phosphorylation sites of MCM (2-9) of *Plasmodium vivax* and Human homolog. PKA – Phospho kinase A, PKC-Phospho kinase C, CK-II –Casein Kinase -II.

PvxMCM 9 is showing 49% sequence similarity to human homolog, hence PvxMCM9 sequence was submitted to the Swissmodel homology-modeling server (swissmodel.expasy.org/). The primary sequence of PvxMCM 9 showed 38% identity to a MCM 3 protein homolog of *Saccharomyces cerevisiae* S288c, the eukaryotic replicative helicase [46]. The structural modeling of the PvxMCM9 was therefore done using the known crystal structure of this homolog as the template (PDB 3ja8_3at www.rcsb.org). The ribbon diagram of the predicted structure of PvxMCM9 is shown in Figure-3.When the modeled structure of PfMCM9 was superimposed, it is clear that these structures superimpose partially. Molecular graphic images were produced using theUCSF Chimera package(www.cgl.ucsf.edu/chimera) from the Resource for Bio-computing, Visualization, and Informatics at the University of California, San Francisco (supported by NIHP41 RR-01081) [47].

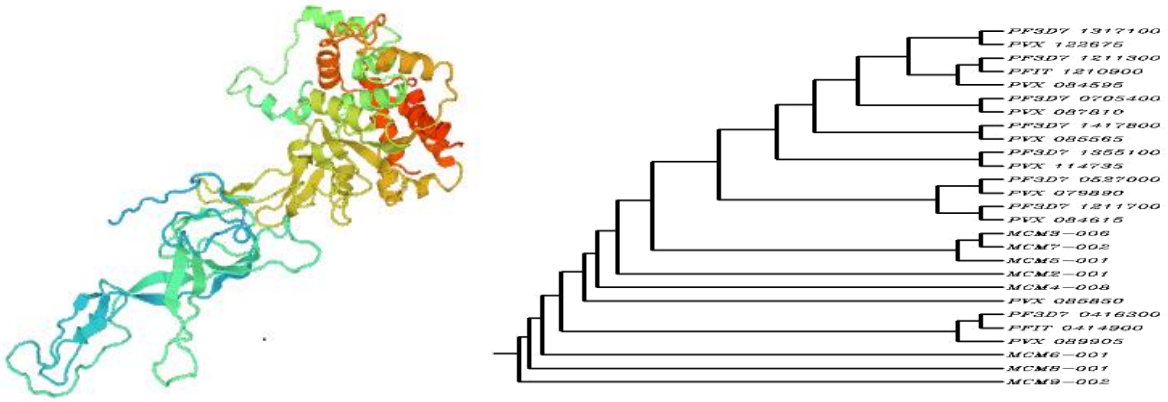


Figure- 3a & 3b. a) 3-D modeled structure of PvxMCM-9, which is showing maximum homology with human homolg. b) Phylogenetic tree of MCM proteins of *P. vivax, P. falciparium* & Human homologs.

The evolutionary pattern of MCM protein family of *P. vivax, P. falciparum* and human was also predicted. The downloaded MCM sequences were subjected to phylogenetic analysis by using MEGA 6.06 software (www.mega6.com) [48]. A neighbor joining rooted tree is generated as a result of phylogenetic analysis. In the phylogeny tree MCM family gene are divided into groups and further sub-divided into sub group as shown in figure- 3b, that clearly reflect differences in their protein sequences.

4. CONCLUSION

Overall in the manuscript, we have attempted to present comparative in-silico analysis of Mini chromosome maintenance MCM protein family of neglected malaria parasite, *Plasmodium vivax* Sal-1 and compare with human. All the MCMs contain a conserved regions approx of ~200-250 amino acids, which is site for nucleotide binding. They have a nearly conserved Walker A motif, Walker B motif and Zinc finger motif. All MCM are localized to nucleus only. Despite of the difference in size of MCM proteins in *P. vivax* and human, the Phosphorylation sites are almost equal in Hs and PvxMCM3, MCM5, MCM6, MCM8 and MCM9. Furthermore it is interesting to note that PvxMCM2 contain 8 sites as opposed to 12 sites in HsMCM2, PvxMCM4 contains 20 phosphorylation sites whereas HsMCM4 has 14 sites. MCM 9 shows 49% sequence similarity to its Human counterpart. This study could provide a way for further analysis/validation of this important family of proteins. MCMs have their role in parasite replication, so these proteins could be used in the formulation of antiparasitic drugs for the control of malaria.

5. ACKNOWLEDGEMENTS

The authors acknowledge the financial support provided by UGC and DBT-IPLS to carry out this work.

6. CONFLICT OF INTEREST

The authors have no conflict of interest.

REFERENCES

1. Maine GT, Sinha P, Tye BK- Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics*. 1984; 106:365-85.
2. Hennessy KM, Lee A, Chen E, Botstein DA-Group of interacting yeast DNA replication genes. *Genes Dev*. 1991; 5:958-69.
3. Takahasi K,Yamada H, Yanagida M- Fission yeast minichromesome loss mutant miscause lethal aneuploidy and replication abnormality. *Mol Biol Cell*. 1994; 15:1145-58.
4. Saydam O, Seno O, Schaaij-Visser TB, Pham TV, Piersma SR, Stemmer-Rachamimov AO, et al- Comparative protein profiling reveals minichromosomemaintenance (MCM) proteins as novel potential tumor markers for meningiomas. *J. Proteome Res*. 2010; 9: 485-94.

5. Crevel I, Crevel G, Gostan T, de Renty C, Coulon V, Cotterill S- Decreased MCM2–6 in Drosophila S2 cells does not generate significant DNA damage or cause a marked increase in sensitivity to replication interference. *PLoS One*. 2011; e27101.
6. Ibarra A, Schwob E, Mendez J - Excess MCM proteins protect human cells fromreplicative stress by licensing backup origins of replication. Proc. Natl. Acad. Sci. 2008; 105:8956–8961.
7. Iyer LM, Leipe DD, Koonin EV, Aravind L- Evolutionary history and higher orderclassification of AAA+ ATPases. *J Struct Biol*. 2008; 146:11-31.
8. Frickey T, Lupas AN- Phylogeneic analysis of AAA proteins. *J Structure Biol*. 2004; 146:2-10.
9. Bell SP and DuttaA - DNA replication in eukaryotic cells: *Annual Review of Biochemistry*. 2002; 71: 333–374.
10. Gregan J, Linder K, Brimage L, Franklin R, Namdar M, Hart EA- Fission yeastCdc23/MCM10 function after pre-replicative complex formation to promote Cdc 45chromatine binding. *Mol Biol Cell*. 2003; 14:3876-87.
11. Tuteja N, Tran NQ, Dang HQ, Tuteja R - Plant MCM proteins ; role in Dna replication and beyond. *Plant Mol Biol*. 2011; 77: 537-47.
12. Dang HQ, Tran NQ, Gill SS, Tuteja R, Tuteja NA- Single subunit MCM6 from pea promote salinity stress tolerance without affecting yield. *Plant Mol Biol*. 2011; 76: 19-34.
13. Amot DE, Gull K- The plasmodium cell-cycle facts and questions. *Ann Trop Med Parasitol*. 1998; 92:361-5.
14. Tuteja R- Malaria: an overview. *The FEBS Journal*. 2007; 274: 4670-9.
15. Bannister L, Mitchell G- The in outs and roundabout of malaria: *Trends Parasitol*. 2001; 19:209-13.
16. Li JL, Cox LS- Identification of an MCM4 homologue expressed specifically in the sexual stageof *Plasmodium Falciparum*. *Int J Parasitol*. 2001; 31:1246-52.
17. Patterson S, Robert C, Whittle C, Chakarbarti R, Doerig C, Chakrabarti D-Pre –replication complex organiszation in the typical DNA replication cycle of *Plasmodium falciparum*: characterisation of the mini-chromosome maintenance (MCM) complexes formation. *Mol Biochem Parasitol*. 2006; 145:50-9.
18. Ansari A, Tuteja R- Genome wide comparative comprehensive analysis of *Plasmodium falciparum* MCM family with human host: *Communicative & Integrativa Biology*. 2006; 5: 607-615.

19. Letunic I, Doerks T- SMART: recent updates, new developments and status in 2015. *Nucleic acid res.* 2015; 43:257-260.

20. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, DasU, Daugherty L, Duquenne L, Finn RD, Gough J- InterPro: the integrative protein signature database. *Nucleic Acids Res*. 2009; 37: 211-215.

21. Chou K, Shen H, A New Method for Predicting the Subcellular Localization of Eukaryotic Proteins with Both Single and Multiple Sites. *Euk-mPLoc*. 2010; 20.

22. Claros MG. MitoProt, a Macintosh application for studying mitochondrial proteins. *Comput Appl Biosci*. 1995; 11(4):441-7.

23. Foth BJ, Ralph SA, Tonkin TC, Nicole S, Fraunholz SK, Roos DS, Cowman, AF, Fadden M GI- Dissecting Apicoplast Targeting in the Malaria Parasite *Plasmodium falciparum*. *Science*. 2003; 299:5606.

24. Koonin EV-A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. *Nucleic Acids Res*.1993; 21:2541-7.

25. Ramakrishna C, DaniVS, Ramasarma TA- A conformational analysis of walker motif A[GXXXXGKT(S)] in nucleotide- binding and other proteins. *Protein Eng* .2002; 15:783-98.

26. Sigrist CJA, Castro DE, Cerutti L, Cuche BA, Hulo N, Bridge A, Bougueleret L Xenarios I-New and continuing developments at PROSITE. *Nucleic Acids Res*.2013; D344–D347.

27. Komamura KY, Tanaka R, Omori A, Kohno T, Ishimi Y- Biochemical characterization of fragmented human MCM2. *FEBS J*. 2008; 275: 727-38.

28. Lin DI, Aggarwal P, Diehl JA- Phosphorylation of MCM3 on Ser-112 regulates it incorporation into the MCM2-7 complex. *Proc Natl Acad Sci*. 2008; 105: 8079-84.

29. Komamura KY, Karasawa SK, Saitoh T, Sato M, Hanaoka F, Tanaka S - Site-specific phosphorylation of MCM4 during the cell cycle in mammalian cells. *FEBS J*. 2006; 273:1224.

30. Gineau L, Cognet C, Kara L, Lach FP, Dunn J, Veturi U -Partial MCM4deficiency in patients with growth retardation, adrenal in sufficiency, and natural killer cell deficiency. *J Clin Invest*. 2012; 122:821-32.

31. Banniter L, Mitchell G - The ins outs and rounds abouts of malaria. *Trends Parasitol*. 2003; 19:209-13.

32.Maine GT, Sinha P, Tye BK- Mutants of *S. cerevisiae* defective in the maintenance of minichromosomes. *Genetics*. 1984; 106:365-85.

33. Ferguson, R.L., Mallaer, J.L., 2008.Cyclin E- dependant localization of MCM5 regulate centrosome duplication. J Cell Sci.121: 3224-32.

34. Afanasyeva EA, MestdaghP, Kumps C, Vandesompele J, Ehemann V, Theissen J, Fischer M, Zapatka M, Brors B, Savelyeva L, Sagulenko V, Speleman F, Schwab M, Westermann F-MicroRNA miR-885-5p targets CDK2 and MCM5, activatesp53 and inhibits proliferation and survival. *Cell Death Differ*. 2011; 18(6):974-84.

35. Zhang J, Yu L, Wu X, Zou L, Sou KK, Wei Z-The interaction domains ofhCdt1 and hMCM6 involve in the chromatin loading of MCM complex in human cells. *Cell Cycle*. 2010; 9:4848-57.

36. Xu J, Zhang S, You C, Huang S, Cai B, Wang X - Expression of humanMCM6and DNA Topo II alpha in crainophayngiomas and its correlation with recurrence of the tumor. *J Neuro oncol*. 2007; 83:183-9.

37. Toyokawa G, Masuda K, Daigo Y, Cho HS, Yoshimatsu M, Takawa M- Minichromosome MaintenanceProtein 7 is a potential therapeutic target in human cancer and a novel prognostic marker of nonsmallcell lung cancer. *Mol Cancer*.2007; 10:65.

38. Doerig C, Chakrabarti D, Kappes B, Matthews K- The cell cycle in protozoan Parasites. *Prog Cell Cycle*. 2007; 4:163-83.

39. Gozuacik D, Chami M, Lagorce D, Faiyre J, MurakamiY, Poch O-Identification and functional characterisation of new member of the human MCM proteins family hMcm8. *Nucleic Acid Res*. 2003; 31:570-9.

40. Volkening M, Hoffmann I - Involvement of human MCM8 in prereplication complex assembly by recruiting hcdc6 to chromatin. *Mol Cell Biol*. 2005; 25:1560- 8.

41. Yoshida, K. Identification of novel cell- cycle - induced MCM family proteinsMCM9. *Biochem Biophys Res Commun*. 2005; 331:669-74.

42. Lutzmann M, Maiorano D, Mechali M- Identification of full gene and proteins ofMCM9, a novel vertebrate- specific member of MCM2-8 proteins family. *Gene.* 2005; 362:51-6.

43. Pereverzeva, Whitmire E, Khan B, Coue M -District phosphoisoform of the xenopus MCM4 proteins regulate the function of the Mcm complex. *Mol Cell Biol*. 2002; 20: 3667-76.

44. Ishimi Y, Komamura KY- Phosphorylation of MCM4, 6, 7 helicase activity. *J Bio Chem*. 2001; 276:34428-33.

45. Brunak, S- Prediction of post- Translation glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics*. 2004; 4:1633-49.

46. Ali FA, Douglas ME, Locke J, Pye VE, Nans A, Diffley JX., Costa A- Cryo-EM structure of a licensed DNA replication origin. *Nature Commun*. 2017; 8: 2241.

47. Pettersen EF,Goddard T, Huang CC, Couch GS, Greenblatt DM, Meng EC. UCSF Chimera – A visualization system for exploratory research and analysis. *J Comput Chem*. 2004; 25:1605:12.

48. Tamura K et al -MEGA6: Molecular Evolutionary Genetics Analysis version 6.0: *Mol Bio and Evol*. 2013; 30: 2725-2729.

49. Chen Y, Tye BK, -The yeast Mcm1 protein is regulated post transcriptionally by the flux of glycolysis. *Mol Cell Biol*. 1995; 15(8): 4631–4639.

50. Takei Y, Assenberg M, Tsujimoto G, Laskey R- The MCM3 acetylase MCM3AP inhibits initiation, but not elongation of DNA replication via interaction with MCM3. *JBC*. 2002; 277: 43121-43125.

51. Todorov IT, Pepperkok R, Philipova RN, Kearsey SE, Ansorge W-A human nuclear protein with sequence homology to a family of early S phase proteins is required for entry into S phase and for cell division. *J. Cell Sci*. 1994; 107:253-265.

52. Shohet M, John H, Sharon E, Plon SM, Burlingame SS, Si-YiC, Malcolm K, Jed G, Nuchtern- Minichromosome maintenance protein MCM7 is a direct target of the MYCN transcription factor in neuroblastoma. *Cancer Res.* 2002; 62: 1123–1128.

53. Johnson EM, Kinoshita Y, Daniel, DC -A new member of the MCM proteinfamily encoded by the human MCM8 gene, located contrapodal to GCD10 at chromosomeband. *NucleicAcid Reseaech.* 2003; 31(11): 2915–2925.

54. Volkening M, Hoffmann I- Involvement of human MCM8 in prereplication complex soluble in vitro plex assembly by recruiting hcdc6 to chromatin. *Mol. Cell. Biol*. 2005; 25:1560–68.