

Evolution and origin of sliding clamp in bacteria, archaea and eukarya

Sandesh Acharya¹, Amol Dahal¹, Hitesh Kumar Bhattarai^{1*},

¹Department of Biotechnology, Kathmandu University, Dhulikhel, Nepal.

Author information.

Affiliation

Department of Biotechnology, Kathmandu University, Dhulikhel, Nepal.

***Corresponding author.**

Correspondence to **Hitesh K. Bhattarai**. (email: hitesh321@gmail.com)

Abstract

Replication of DNA is an essential process in all domains of life. A protein often involved without exception in replication is the sliding clamp. The sliding clamp encircles the DNA and helps replicative polymerase stay attached to the replication machinery increasing the processivity of the polymerase. In eukaryotes and archaea the sliding clamp is called the Proliferating Cell Nuclear Antigen (PCNA) and consists of two domains. This PCNA forms a trimer encircling the DNA as a hexamer. In bacteria, the structure of the sliding clamp is highly conserved, but the protein itself, called beta clamp, contains three domains, which dimerize to form a hexamer. The bulk of literature touts a conservation of the structure of the sliding clamp, but fails to recognize conservation of protein sequence among sliding clamps. In this paper we have used PSI blast to the second iteration in NCBI to show a statistically significant sequence homology between *Pyrococcus furiosus* PCNA and *Kallipyga gabonensis* beta clamp. The last two domains of beta clamp align with the two domains of PCNA. This homology data demonstrates that PCNA and beta clamp arose from a common ancestor. In this paper, we have further used beta clamp and PCNA sequences from diverse bacteria, archaea and eukarya to build maximum likelihood phylogenetic tree. Most, but not all, species in different domains of life harbor one sliding clamp from vertical inheritance. Some of these species that have two or more sliding clamps have acquired them from gene duplication or horizontal gene transfer events.

Introduction:

Replication, transcription and translation are fundamental events occurring in all living organisms. Homologous machineries exist for the transcription and translation pathways of all cells however replication machineries vary considerably among the different domains, mainly between bacteria and archaea/eukarya(Yao & O'Donnell 2016)². DNA polymerases, primases and helicases involved in replication process are quite distinct and non-homologous between prokaryotes and eukaryotes(Yao & O'Donnell 2016)². Interestingly two replication proteins, DNA sliding clamps and clamp loader share structural homology in the different domains of life.

DNA sliding clamp proteins, found in all organisms are essential components of DNA polymerase enzyme as they bind the enzyme and prevent it from dissociating from the template DNA strand. The sliding clamp proteins are called beta clamp in bacteria and proliferating cell nuclear antigen (PCNA) in eukaryotes(Bruck & O'Donnell 2001)². Since the replication machinery of archaea demonstrates similarity to eukaryotic replication machinery than to bacterial replication machinery, archaea contain PCNA and not beta clamp. While beta clamp is a homo-dimer consisting two sub-units of three domains(Kong et al. 1992)², PCNA is a homo-trimer made up of three subunits of two domains(Matsumiya 2001)². In some archaea of Crenarchaeal sub-domain, three PCNA homologs are present, which exist as monomers in solution and self-assemble in to a functional hetero-trimer(Dionne et al. 2003; Williams et al. 2006)². Despite the difference in the sub-units and domains, all sliding clamps assemble to form a ring-shaped protein, such that the ring is wide enough to accommodate double-stranded DNA(Kelman 1995; Jeruzalmi et al. 2002)². The architecture consists of outer negatively charged surface of continuous beta-sheets, with positively charged inner surface composed of alpha helices(Kelman 1995; Hedglin et al. 2013)². In spite of sub-unit compositions, the 3-D structure of the sliding clamp has been highly conserved throughout evolution.

Sliding clamps encircle DNA(Stukenberg et al. 1991)² and thereby, physically tether the DNA polymerase to the DNA template, hence increasing the processivity of the enzyme(Stukenberg et al. 1991)². Previously, the function of sliding clamps was thought to be limited only to maintain the processivity of DNA polymerase. However, with recent studies, various interactions of sliding clamps with other proteins have been recognized that shed light on different roles of sliding clamps in the cell. PCNA has been known to interact with several proteins like XPG, MSH3, MSH6, MCH1, PMS2, hMYH, Fen1 endonuclease, and DNA ligase I suggesting roles in nucleotide excision repair, mismatch repair, base excision repair, and maturation of Okazaki fragments(Parker et al. 2001)². Similarly, the prokaryotic homolog of PCNA, viz. beta-clamp, has been observed to interact with MutS, Poll, DNA ligase, PolII and PolV suggesting roles in mismatch repair, processing of Okazaki fragments, and DNA repair(Wilkins 2000; López de Saro & O'Donnell 2001)². Thus, not only the 3-D structure, but also interactions and functions are conserved between sliding clamps of prokaryotes and eukaryotes.

The existing body of information reveals that bacterial beta clamps and eukaryotic PCNA share structural homology. It has been reported that beta clamps and PCNA share no sequence homology at all although beta clamp proteins and PCNA are highly conserved in bacteria and in archaea/eukaryotes, respectively (Kong et al. 1992; Bruck & O'Donnell 2001; Hedglin et al. 2013). In this study, we analyzed sequence homology between bacterial beta clamp and archaeal/eukaryotic PCNA. Additionally, to understand their evolution, sliding clamp sequences from representative bacteria, archaea and eukarya were retrieved and phylogenetic tree was drawn for each domain of life. As expected the phylogenetic trees largely follow the vertical evolution pattern of organisms. In some instances they reveal gene duplication and horizontal gene transfer events.

Material and methods:

Determination of sequence homology between bacterial beta clamp and archaeal PCNA

Two hundred forty nine amino acid long *Pyrococcus furiosus* PCNA sequence was obtained from Uniprot website. To discover homologues of the sequence in bacteria, the sequence was protein blasted against bacteria in the NCBI database (taxID :2). The non-redundant protein sequence database was used, and position specific iterated blast (PSI-Blast) was carried out using the default BLOSUM 62 Matrix, gap cost of Existence:11 Extension:1, with conditional compositional score matrix adjustment at the threshold of 0.005. To discover more distant homologues, second iteration of PSI-blast was carried out.

From the second iteration of the PSI blast a number of bacterial beta clamps were found as homologues of *P. furiosus* PCNA. Specifically domain 2 of bacterial beta clamp was found homologous to *P. furiosus* PCNA domain 1 and domain 3 of bacterial beta clamp was found homologous to *P. furiosus* PCNA domain 2. *Kallipyga gabonensis* beta clamp was found to be the most homologous bacterial beta clamp. To demonstrate the homology between *K. gabonensis* beta clamp and *P. furiosus* PCNA, domain 2 and 3 of *K. gabonensis* beta clamp (starting amino acid 128) as discovered by pfam website was aligned with *P. furiosus* PCNA using clustal omega software. Similarly, human PCNA and *P. furiosus* PCNA, and domain 2 and 3 of *K. gabonensis* beta clamp and *E. coli* beta clamp (starting amino acid 129) were aligned using clustal omega. Earlier papers have reported that there is no homology between PCNA and beta clamp. To evaluate this point, human PCNA was aligned with domain 2 and 3 of *E. coli* beta clamp.

Sequence Retrieval and Multiple Sequence Alignment for Tree Construction:

In order to observe whether the sliding clamp proteins were conserved in their respective domain, a phylogenetic tree was constructed for different domains. The first step in this process was the sequence retrieval.

From the three domains of life, one organism each was chosen and their protein sequence for Beta-Clamp or PCNA was downloaded. The beta-clamp protein from *E.coli* (Uniport Accession Number : P0A988) was PSI-blasted against Non-Redundant Protein Sequences (NRPS) of 75 different species in

bacterial domain, and all homologs of the protein from different organisms were downloaded. Similarly, the PCNA from *Pyrococcus furiosus* (Uniport Accession Number: **073947**) was PSI-blasted against Non-Redundant Protein Sequences of 74 archeal species and PCNA protein from *Homo sapiens* (**Uniport Accession Number: P12004**) was PSI-blasted against Non-Redundant Protein Sequences of 31 eukaryal species(Postberg et al. 2010)² to obtain homologous proteins in each domain. The organisms in bacterial and archea domain were chosen from a list of organisms used by Moreira in 2014(Petitjean et al. 2014)³. The organisms were selected to represent all the classes, and also had their whole genome sequenced.

The obtained protein sequences were then aligned using Muscle algorithm (Edgar 2004)⁴ with default settings in MEGA X. In this way, different Multiple Sequence Alignments were generated for the three domains of life.

Tree Construction

The evolutionary history was inferred by using the Maximum Likelihood method and JTT matrix-based model(Jones et al. 1992)⁵. The bootstrap consensus tree inferred from 500 replicates was taken to represent the evolutionary history of the taxa analyzed(Felsenstein 1985)⁶. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) were shown next to the branches(Felsenstein 1985)⁶. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. All positions with less than 95% site coverage were eliminated, i.e., fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position (partial deletion option). Evolutionary analyses were conducted in MEGA X(Kumar et al. 2018)⁷.

A.

Description	Max score	Total score	Query cover	E value	Per. Ident	Accession
proliferating cell nuclear antigen (pcna) [Flavobacteriales bacterium]	222	222	96%	1e-68	23.74%	MAD50601.1
proliferating cell nuclear antigen (pcna) [Acidimicrobiaceae bacterium]	218	218	98%	2e-67	23.32%	MBG02288.1
proliferating cell nuclear antigen (pcna) [Dehalococcoidaceae bacterium]	217	217	99%	3e-67	30.92%	MBR74595.1
proliferating cell nuclear antigen (pcna) [Alphaproteobacteria bacterium TMED62]	217	217	96%	1e-66	21.76%	OUU62278.1
proliferating cell nuclear antigen (pcna) [Halobacteriovoraceae bacterium]	217	217	96%	1e-66	23.55%	MAW08577.1
proliferating cell nuclear antigen (pcna) [Flavobacteriales bacterium TMED191]	217	217	96%	1e-66	23.74%	RPG58329.1
proliferating cell nuclear antigen (pcna) [Candidatus Marinimicrobia bacterium]	216	216	97%	1e-66	21.12%	MBD23483.1
proliferating cell nuclear antigen (pcna) [Flavobacteriaceae bacterium]	217	217	96%	1e-66	23.14%	MAD12048.1
proliferating cell nuclear antigen (pcna) [Nitrospinae bacterium]	216	216	97%	2e-66	21.83%	MBE18271.1

B.

DNA polymerase III subunit beta [Kallipyga gabonensis]	53.8	53.8	94%	1e-04	19.34%	WP_053942720.1
hypothetical protein DRP84_07425 [Spirochaetes bacterium]	52.7	52.7	94%	2e-04	18.70%	RKX94156.1
DNA polymerase III subunit beta [Kallipyga massiliensis]	51.5	51.5	94%	6e-04	18.11%	WP_019133621.1
DNA polymerase III subunit beta [Cyanobacteria bacterium QH_2_48_84]	50.0	50.0	92%	0.002	16.67%	PSO65263.1
DNA polymerase III subunit beta [Cyanobacteria bacterium QH_7_48_89]	50.0	50.0	92%	0.002	16.67%	PSO58692.1
DNA polymerase III subunit beta [Cyanobacteria bacterium QH_1_48_107]	49.6	49.6	92%	0.003	16.67%	PSO55030.1
DNA polymerase III beta subunit [Edaphobacter modestus]	49.6	49.6	87%	0.003	13.87%	RZU41692.1

Fig 1. Screenshot of the results obtained when *P. furiosus* PCNA is PSI blasted against the NCBI bacterial database. (A) shows the results obtained when regular blast is carried out. PCNA sequences in undefined bacterial genomes appear. It is possible that these bacterial PCNAs emerged in bacterial genome from horizontal gene transfer from archaea and eukarya. (B) shows bacterial beta clamp from some defined and some undefined species that appear in the blast after second iteration. The yellow color represents hits that only appear in the second iteration of the blast. *Kallipyga gabonensis* is the top defined beta clamp hit in the bacteria. Other defined bacterial beta clamps that appear in the blast search are *Kallipyga massiliensis* and *Edaphobacter modestus* beta clamps. Query cover above 90 percent in these searches indicates that the entire PCNA was covered during the blast search.

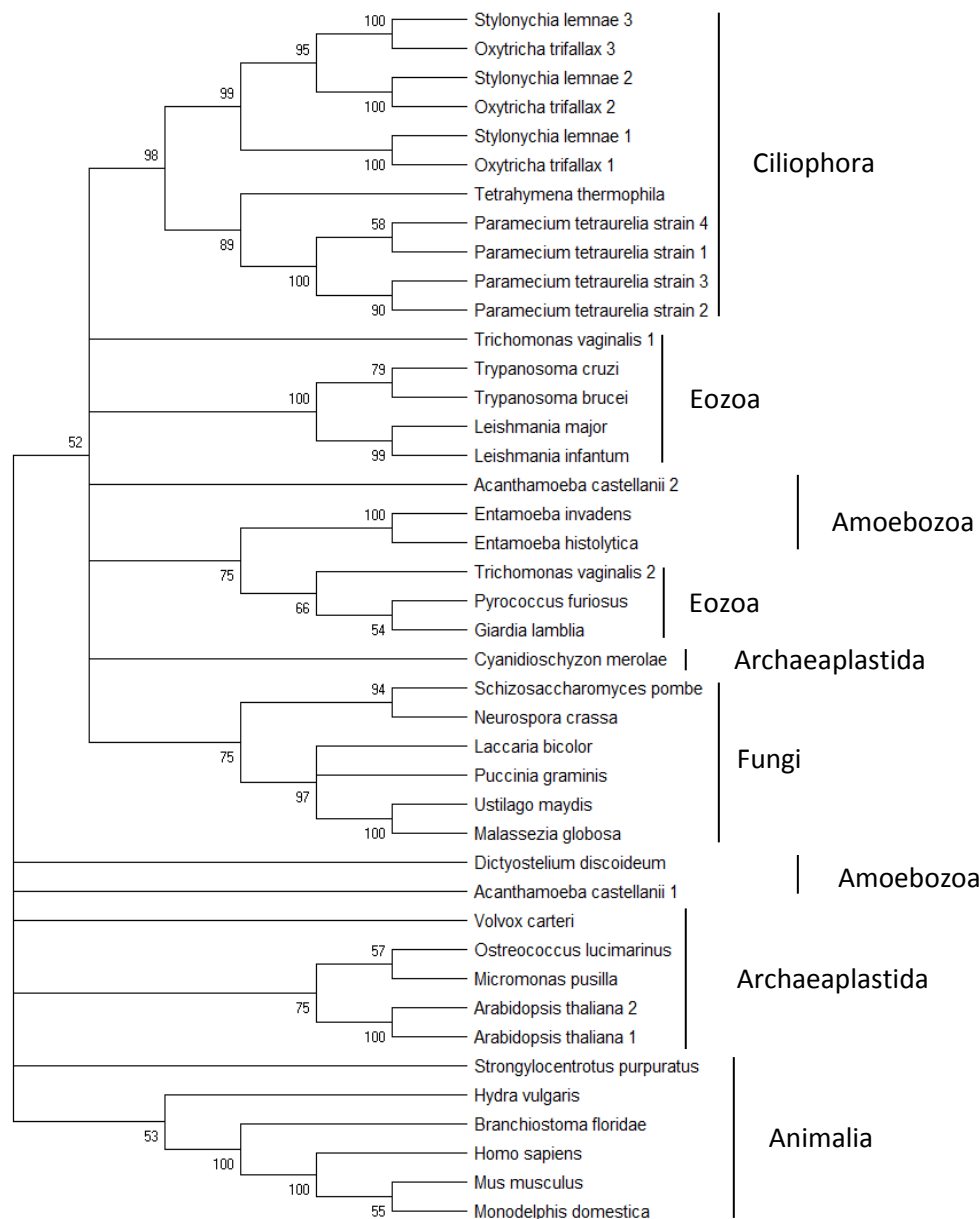


Fig 3. Phylogenetic Maximum Likelihood tree drawn for PCNA of 31 different eukaryotic species. Both domains of PCNA were used for alignment. Species from 6 kingdoms of eukaryotes—animalia, fungi, archaeplastida, amoebozoa, eozoa and ciliophora (inside the kingdom chromista)—were used to draw the tree. The purpose of using organisms from 6 different kingdoms is to represent the entire eukaryotic domain, not just the crown groups. This tree demonstrates that most species contain only one PCNA that is vertically inherited. The instances where there is gene duplication and horizontal gene transfer is indicated in the text.

Fig. 4

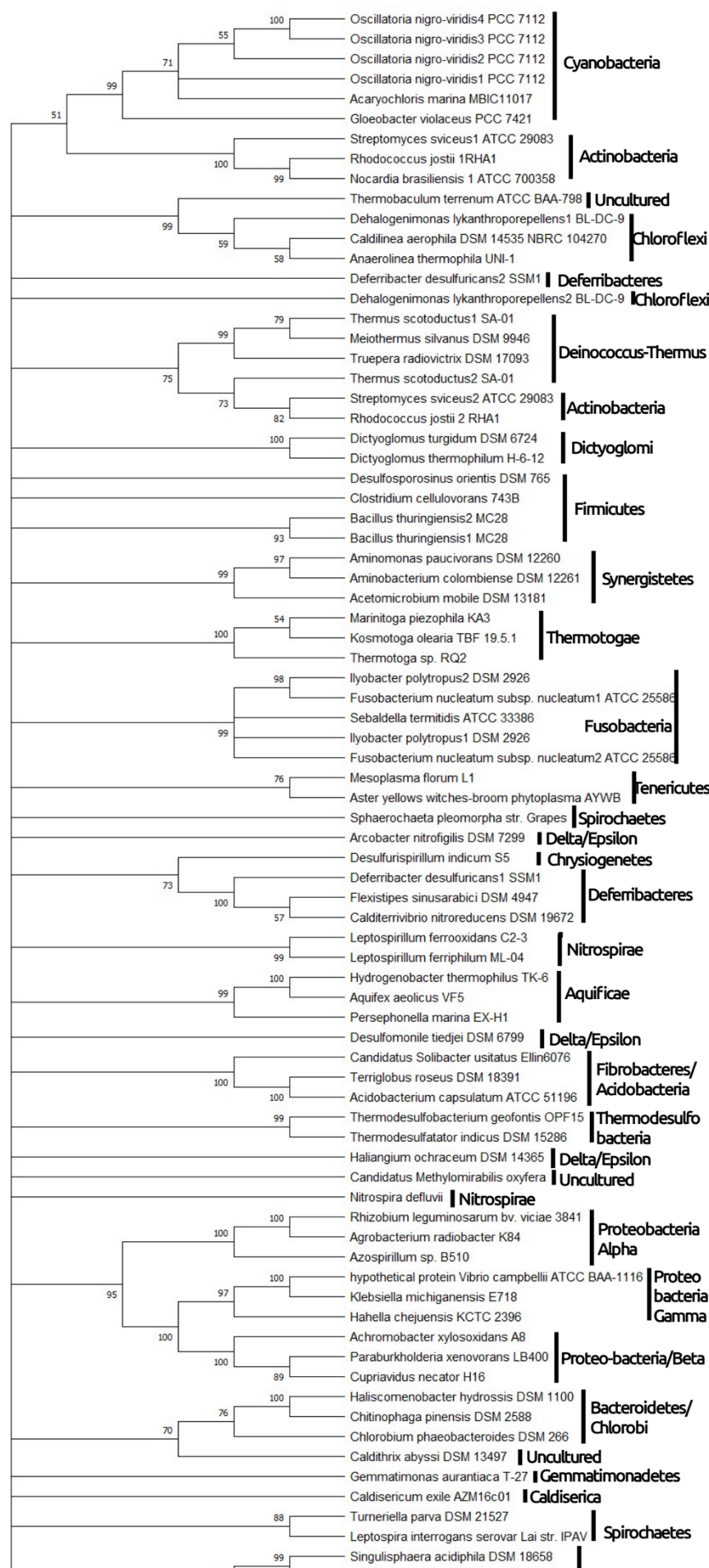
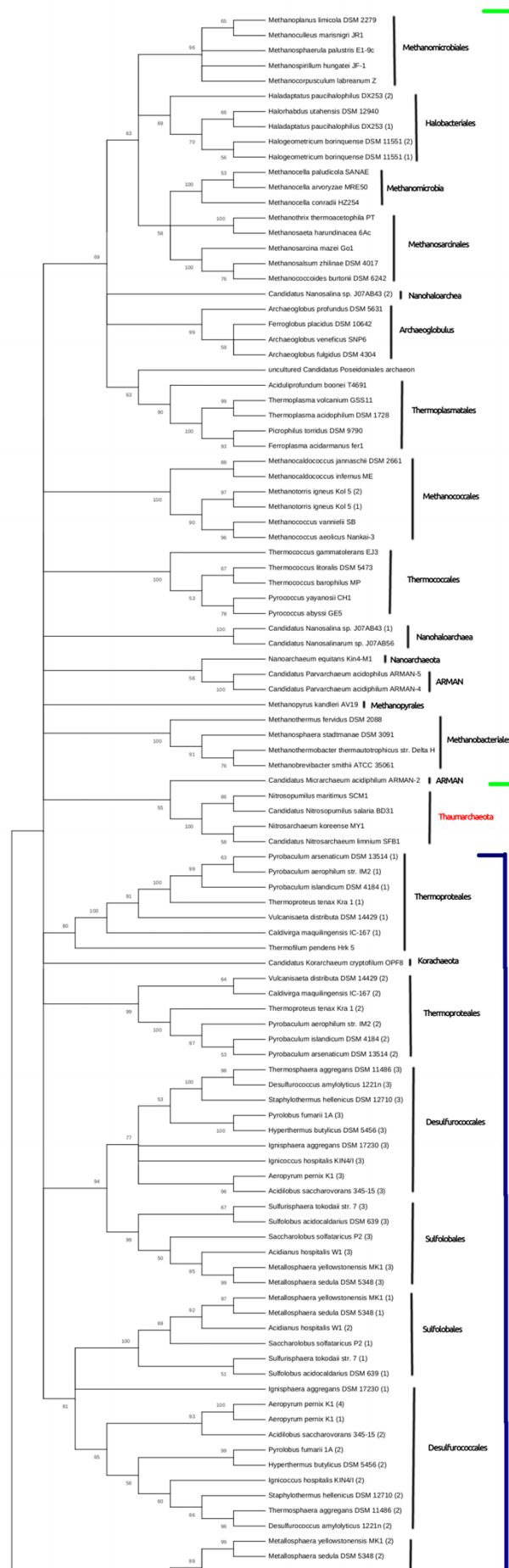


Fig.4 Maximum Likelihood tree drawn for beta-clamp of 75 different bacterial species representing different phylum and classes. All three domains of beta-clamp were used for alignment. The tree demonstrates that most bacteria have only a single homolog for beta-clamp protein, which is vertically inherited. Some instances of gene duplication and horizontal gene transfer are also observed. Bacteria belonging to the same class are clustered together with minor exceptions.

Fig.5 Maximum Likelihood tree drawn for PCNA of 75 different archaeal species representing three major phyla: Euryarchaeota, Thaumarchaeota and Crenarchaeota (Green Line: Euryarchaeota, Blue Line: Crenarchaeota). The tree demonstrates that the three phyla are positioned separately. Some archaea are observed to have more than one homolog for PCNA. It is observed that PCNA protein is highly conserved in archaea, with few exceptions of gene duplication and horizontal gene transfer.

Fig 5.



Result and discussion:

Detection of homology between PCNA and sliding clamp

With a threshold of 0.005, *P. furiosus* PCNA was PSI blasted against the NCBI non-redundant bacterial protein sequence database to the second iteration. Figure 1A shows a screenshot of the top hits from the blast. A number of PCNAs were discovered in the bacterial genome. Given that PCNAs are characteristic proteins of eukaryotes and archaea, these proteins may have arisen by horizontal gene transfer from eukaryotes or archaea to bacteria. Further analysis is necessary to determine their source of origin. The proteins not found in the first iteration, but found in the second iteration are highlighted in yellow as in Figure 1b. A slew of bacterial beta clamp were discovered as homologues of *P. furiosus* PCNA upon PSI blast at second iteration. Top among the hits was *Kallipyga gabonensis* beta clamp. The query cover for the blast was 94 percent, E score was 1×10^{-4} , and sequence identity was above 19%. Since a random blast of PCNA yielded beta clamp of various species, it can be claimed that the two are homologous.

A closer inspection of the alignment result shows that out of the three domains of beta clamp, domains 2 and 3 align with the full length of the PCNA. This suggests that either the Lowest Universal Common Ancestor (LUCA) had three domains in its sliding clamp, the first of which got lost in archaea, or LUCA had two domains in its sliding clamp and the third domain got added after duplication event. The alignment of *K. gabonensis* beta clamp and *P. furiosus* PCNA shows 20%(50/249) sequence identity, 44%(109/249) sequence similarity and 11%(27/249) gaps. On the other hand, *P. furiosus* PCNA and human PCNA show 25%(61/249) sequence identity, 61%(136/249) sequence similarity and 5%(13/249) gaps. *E. coli* beta clamp and human PCNA, which were earlier claimed to not have sequence homology, show 16%(42/261) sequence identity, 36%(94/261) sequence similarity and 16%(43/261) gap. These results demonstrate that although more common model organism (human and *E. coli*) PCNA and beta clamp do not demonstrate high degree of homology, non-model organism (*K. gabonensis* and *P. furiosus*) beta clamp and PCNA demonstrate statistically noteworthy sequence homology.

Given that domain 2 and 3 of beta clamp align with the two domains of PCNA, it can be hypothesized that domain 1 from beta clamp interacts with bacteria specific replication and repair protein. Most of the interaction assays carried out with PCNA and beta clamp have been conducted with the entire protein. It would be interesting to zero in on which domain of the sliding clamps is important for individual interactions.

Phylogenetic tree of eukaryotic PCNA

To study the origin and evolution of eukaryotic PCNA, PCNA from human was protein blasted in NCBI against the non-redundant protein sequences of selected eukaryotic organisms from different classes. Certain species of the kingdom/class eozoa, amoebzoa, ciliophora, archaeplastida, fungi and animalia were chosen. Where multiple PCNAs were discovered all sequences were downloaded. A phylogenetic maximum likelihood tree was drawn using these sequences. A tree thus constructed is shown in Figure 3. The proteins from different classes tended to segregate as labeled on the right. For example all PCNAs from ciliophora and animalia cluster together. This demonstrates that there is very little horizontal gene transfer. A largely vertical inheritance pattern can be expected for an essential protein like PCNA. However, a number of species in ciliophora class demonstrate gene duplication events. Three PCNA genes are found in both *Stylonychia lemnae* and *Oxytricha trifallax*. Each of the three proteins from one species lies next to a protein from another species. This suggests that a single PCNA gene underwent gene duplication twice in the progenitor organism, which gave rise to three PCNAs, each of which evolved separately when speciation occurred.

The figure also demonstrates that there are four closely related PCNAs in *Paramecium tetraurelia*. The tree pattern suggests that there was an initial duplication event, followed by duplications of the duplicated PCNA.

Amoebozoa and eozoa PCNAs tended to cluster at two locations each. *Trichomonas vaginalis* had two PCNAs, one clustering with other eozoa PCNAs and other clustering with *Pyrococcus furiosa* (archaea) PCNA. This demonstrates that one PCNA was vertically inherited, while other PCNA might have arisen from horizontal gene transfer from archaea. Other eozoa (*Leishmania* and *Trypanosoma*) PCNA clustered together, suggesting that they all came from a common ancestor. *Giardia lamblia* (an eozoa) PCNA, however segregated separately close to *Pyrococcus furiosus* PCNA, hinting at a possible archaeal origin.

As for amoebozoa PCNAs, three PCNAs from different species lie next to each other, whereas two other PCNAs, one from *Dictostylium discoideum* and another from *Acanthamoeba castellanii*, are sandwiched between archaeplastida and fungi. It has been observed from earlier research that *Dictostylium* molecules resemble archaeplastida and fungi molecules more than amoebozoa molecules (Baldauf & Doolittle 1997; Eichinger et al. 2005) □. So this placement in the tree does not come as a surprise. The second *Acanthamoeba castellanii* PCNA might have arisen from horizontal gene transfer from *Dictostylium* or archaeplastida or fungi.

All fungi and animalia PCNAs cluster together suggesting a common origin. There seems to be no gene duplication event in the species we have considered in the paper. All except one archaeplastida molecules cluster together also suggesting a common origin. *Cyanidioschyzon merolea* PCNA is found at a different location from other archaeplastida genome suggesting a different ancestry.

Phylogenetic tree of bacterial beta clamp:

The Beta-clamp from *E. coli* was PSI-blasted against non-redundant protein sequences from different organisms belonging to different classes in bacteria. A total of 85 homologs of the protein were identified from 75 species of bacteria. All the homologs were downloaded, and aligned together to build a phylogenetic tree as shown in figure 4. It was observed that most of the bacteria consisted of only one homolog for Beta-clamp. However, some bacteria had two or more homologs of the protein, suggesting ancient gene duplication or horizontal gene transfer. The bacteria belonging to same phylum, or sharing similar properties were found to be clustered together apart from some exceptions, which suggests that the beta-clamp protein is conserved among similar species in bacterial domain.

In proteobacteria, three classes viz. Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria are found to be clustered together, originating from a single node suggesting common origin. However, the Delta/Epsilon sub-divisions seem to scatter away from other Proteobacteria in the phylogenetic tree. It suggests that the the phylum Proteobacteria is not monophyletic which has also been previously suggested (Hug et al. 2016) □. The phyla Planctomycetes, Verrucomicrobia and Chlamydiae which are often grouped together as PVC superphylum (Hug et al. 2016) □ were observed to be clustered together, suggesting similar origin and properties.

Multiple homologs of the same protein in a species can be accounted for by ancient gene duplication or horizontal gene transfer events. In case of *Oscillatoria nigroviridis*, four homologs of sliding beta clamp were observed. It was a result of three ancient gene duplication at different stages of time, followed by vertical gene transfer. Similar was the case for *Bacillus thuringensis*. The resulting homologs didn't undergo much variation and thus were clustered together. The phylum Actinobacteria was positioned at two different locations in the phylogenetic tree, one clustered with Cyanobacteria and the other with Deinococcus-thermus. It suggests that the horizontal gene transfer occurred from the ancestor of Deinococcus-thermus to actinobacteria before speciation. This gene was maintained in some species,

which showed two homologs for beta-clamp protein, while disappeared in some species (*Nocardia brasiliensis*), which had only a single homolog.

In the phylum Fusobacteria, two homologs each of *Ilyobacter polytropus* and *Fusobacterium nucleatum subsp. nucleatum* were observed, which were clustered together. It might have resulted from ancient gene duplication, or lateral gene transfer among the species. Two homologs of the beta-clamp protein of *Deferribacter desulfuricans* were observed, where one homolog was separated from its counterpart and located between species of Chloroflexi suggesting lateral gene transfer from Chloroflexi. Other cases of lateral gene transfer can be claimed in *Nitrospira defluvii*, and *Dehalogenimonas lykanthroporepellens*.

Apart from some cases of horizontal gene transfer and gene duplication, the species from same class were clustered together in the phylogenetic tree. Thus, we can conclude that beta-clamp protein is fairly conserved in bacterial domain.

Phylogenetic Tree of Archeal PCNA

The PCNA from *P. furiosus* was PSI-blasted against non-redundant protein sequences from different archaea belonging to different classes in archaea. 115 homologs of PCNA were identified in 74 species of archaea. All of the homolog sequences were downloaded, and a multiple sequence alignment and phylogenetic tree was constructed similarly as discussed above.

The obtained phylogenetic tree is shown in Figure 5. It can be observed that the three main phyla of domain archaea, viz Thaumarchaeota, Euryarchaeota and Crenarchaeota were positioned separately in the phylogenetic tree, apart from some exceptions. The phylum Thaumarchaeota was sandwiched in between Euryarchaeota and Crenarchaeota. A single species of Korarchaeota was present within the phylum Crenarchaeota, which leads to questions the monophyly of Crenarchaeota.

In the phylum, Euryarchaeota, the species belonging to the same class were classified together. However, some cases of duplication and horizontal gene transfer were observed. Two homologs of PCNA were observed in *Halogeometricum borinquense* DSM 11551, which were separated from the same node, hence suggesting ancient gene duplication. Similar was the case for *Methanoterris igneus* K015. Also, two homologs of the protein were discovered for *Haladaptatus paucihalophilus*, and *Candidatus Nanosalina*. These two homologs were not positioned together, hence suggests horizontal gene transfer.

The phylum Crenarchaeota is divided into three classes, viz. Sulfolobales, Desulfurococcales, and Thermoproteales. The phylogenetic tree obtained still supports the close relationship of Sulfolobales and Desulfurococcales, also mentioned by Armanet(Brochier-Armanet et al. 2011)^{□□}. Both these classes had three homologs of PCNA. The two homologs appear to have originated by gene duplication in the ancestor of these two classes, followed by modification to form different proteins, while the third homolog might have been acquired by lateral gene transfer from other diverge species, possibly eukaryotes. As all three homologs are required for the assembly of a functional hetero-trimer, the homologs are observed to be vertically inherited in future generations. However, the third phylum of Crenarchaeota: Thermoproteales had two homologs of PCNA. Thus, it can be predicted that the functional PCNA in Thermoproteales is a homo-trimer (of any one of the homolog) and two homologs might have arisen due to ancient gene duplication.

The phylum Euryarchaeota can be divided into two groups: Sub-Phyla I and Sub-Phyla II. Sub-Phyla I consists of Thermococcales, Nanoarchaea, and class I methanogens, while Sub-Phyla II consists of Archeoglobulus, Halobacteriales, Thermoplasmatales and Class II methanogens.(Forterre 2015)[□] The phylogenetic tree in figure 5 suggests the monophyly of Sub-Phyla II, which is also supported by the updated tree of life by Forterre.(Forterre 2015)[□] However, the monophyly of Sub-Phyla I couldn't be supported statistically.

The phylogenetic tree further clarifies the position of Methanopyrales in the phylogenetic tree. As *Methanopyrales kandleri* was positioned along with Methanobacteriales, Methanopyrales can be classified as Class I Methanogen. However, the monophyly of Class I and Class II methanogen couldn't be supported. Methanobacteria, Methanobacteriales and Methanosarcinales were observed to share a common ancestor along with Halobacteriales. Methanococcus was positioned near to Thermococcales, while Nanohaloarchaea and Nanoarchaeota were sandwiched between Methanococcus and Methanobacteriales.

The phylum Nanoarchaeota was positioned near to Thermococcales, along with Nanohaloarchaea, and ARMAN group. The close association of Nanoarchaeota to Thermococcales was also observed by Brochier in 2005 (Brochier et al. 2005)[□]. Also, Thermococcales, Nanoarchaea, ARMAN and Nanohaloarchaea are placed together suggesting they are closely related. It has also been supported by phylogeny based on ribosomal proteins and DNA replication proteins (Raymann et al. 2014; Forterre 2015)[□]. The tree however, questions the position of ARMAN-2. It has been positioned as a sister group to Thaumarchaeota, forming a border-line between Crenarchaeota and Euryarchaeota.

Thus, the phylogenetic tree based on archaeal PCNA follows vertical pattern of evolution in most cases. Also, archaeal PCNA is conserved among closely related species of the Archaea domain.

To conclude, the sliding clamp proteins (Beta-Clamp in Bacteria, PCNA in Archaea and Eukarya) are highly conserved within their respective domain apart from some exceptions. Also, the observed sequence homology between some Beta-Clamps and PCNA generates the possibility of common origin for the sliding clamp proteins.

References

- Baldauf SL, Doolittle WF. 1997. Origin and evolution of the slime molds (Mycetozoa). *Proc Natl Acad Sci U S A*. 94(22):12007–12012.
- Brochier-Armanet C, Forterre P, Gribaldo S. 2011. Phylogeny and evolution of the Archaea: One hundred genomes later. *Curr Opin Microbiol*. 14(3):274–281.
- Brochier C, Gribaldo S, Zivanovic Y, Confalonieri F, Forterre P. 2005. Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol*. 6(5):1–10.
- Bruck I, O'Donnell M. 2001. The ring-type polymerase sliding clamp family. *Genome Biol*. 2(1):2–4.
- Dionne I, Nookala RK, Jackson SP, Doherty AJ, Bell SD. 2003. A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Cell*. 11(1):275–282.
- Edgar RC. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32(5):1792–1797.
- Eichinger I, Pachebat JA, Glöckner G, Rajandream MA, Sugrang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, et al. 2005. The genome of the social amoeba *Dictyostelium discoideum*. *Nature*. 435(7038):43–57.
- Felsenstein J. 1985. Confidence Limits on Phylogenies: an Approach Using the Bootstrap. *Evolution* (N Y). 39(4):783–791.

- Forterre P. 2015. The universal tree of life: An update. *Front Microbiol.* 6(JUN):1–18.
- Hedglin M, Kumar R, Benkovic SJ. 2013. Replication clamps and clamp loaders. *Cold Spring Harb Perspect Biol.* 5(4):1–19.
- Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hermsdorf AW, Amano Y, Ise K, et al. 2016. A new view of the tree of life. *Nat Microbiol.* 1(5):1–6.
- Jeruzalmi D, O'Donnell M, Kuriyan J. 2002. Clamp loaders and sliding clamps. *Curr Opin Struct Biol* [Internet]. 12(2):217–224.
<https://www.sciencedirect.com/science/article/abs/pii/S0959440X02003135?via%3Dihub>
- Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices. *Comput Appl Biosci.* 8(3):275–282.
- Kelman Z. 1995. DNA Polymerase III Holoenzyme: Structure and Function of a Chromosomal Replicating Machine. *Annu Rev Biochem.* 64(1):171–200.
- Kong XP, Onrust R, O'Donnell M, Kuriyan J. 1992. Three-dimensional structure of the β subunit of E. coli DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell.* 69(3):425–437.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 35(6):1547–1549.
- López de Saro FJ, O'Donnell M. 2001. Interaction of the β sliding clamp with MutS, ligase, and DNA polymerase I. *Proc Natl Acad Sci U S A.* 98(15):8376–8380.
- Matsumiya S. 2001. Crystal structure of an archaeal DNA sliding clamp: Proliferating cell nuclear antigen from *Pyrococcus furiosus*. *Protein Sci.* 10(1):17–23.
- Parker A, Gu Y, Mahoney W, Lee SH, Singh KK, Lu AL. 2001. Human Homolog of the MutY Repair Protein (hMYH) Physically Interacts with Proteins Involved in Long Patch DNA Base Excision Repair. *J Biol Chem.* 276(8):5547–5555.
- Petitjean C, Deschamps P, López-García P, Moreira D. 2014. Rooting the domain archaea by phylogenomic analysis supports the foundation of the new kingdom Proteoarchaeota. *Genome Biol Evol.* 7(1):191–204.
- Postberg J, Forcob S, Chang WJ, Lipps HJ. 2010. The evolutionary history of histone H3 suggests a deep eukaryotic root of chromatin modifying mechanisms. *BMC Evol Biol.* 10(1):1–13.
- Raymann K, Forterre P, Brochier-Armanet C, Gribaldo S. 2014. Global phylogenomic analysis disentangles the complex evolutionary history of DNA replication in Archaea. *Genome Biol Evol.* 6(1):192–212.
- Stukenberg PT, Studwell-Vaughan PS, O'Donnell M. 1991. Mechanism of the sliding β -clamp of DNA polymerase III holoenzyme. *J Biol Chem.* 266(17):11328–11334.
- Wilkins AS. 2000. The puzzle of PCNA's many partners. *BioEssays.* 22(11):997–1006.

Williams GJ, Johnson K, Rudolf J, McMahon SA, Carter L, Oke M, Liu H, Taylor GL, White MF, Naismith JH. 2006. Structure of the heterotrimeric PCNA from *Sulfolobus solfataricus*. *Acta Crystallogr Sect F Struct Biol Cryst Commun.* 62(10):944–948.

Yao NY, O'Donnell ME. 2016. Evolution of replication machines. *Crit Rev Biochem Mol Biol.* 51(3):135–149.