

## Review Article

# Diversity of the DNA Replication System in the Archaea Domain

Felipe Sarmiento,<sup>1</sup> Feng Long,<sup>1</sup> Isaac Cann,<sup>2</sup> and William B. Whitman<sup>1</sup>

<sup>1</sup> Department of Microbiology, University of Georgia, 541 Biological Science Building, Athens, GA 30602-2605, USA

<sup>2</sup> 3408 Institute for Genomic Biology, University of Illinois, 1206 W Gregory Drive, Urbana, IL 61801, USA

Correspondence should be addressed to William B. Whitman; whitman@uga.edu

Received 21 October 2013; Accepted 16 February 2014; Published 26 March 2014

Academic Editor: Yoshizumi Ishino

Copyright © 2014 Felipe Sarmiento et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The precise and timely duplication of the genome is essential for cellular life. It is achieved by DNA replication, a complex process that is conserved among the three domains of life. Even though the cellular structure of archaea closely resembles that of bacteria, the information processing machinery of archaea is evolutionarily more closely related to the eukaryotic system, especially for the proteins involved in the DNA replication process. While the general DNA replication mechanism is conserved among the different domains of life, modifications in functionality and in some of the specialized replication proteins are observed. Indeed, *Archaea* possess specific features unique to this domain. Moreover, even though the general pattern of the replicative system is the same in all archaea, a great deal of variation exists between specific groups.

## 1. Introduction

*Archaea* are a diverse group of prokaryotes which are united by a number of unique features. Many of these microorganisms normally thrive under extreme conditions of temperature, salinity, pH, or pressure. However, some also coexist in moderate environments with *Bacteria* and eukaryotes [1, 2]. Likewise, a wide range of lifestyles, including anaerobic and aerobic respiration, fermentation, photo- and chemoautotrophy, and heterotrophy, are common among this group. In spite of this diversity, members of the *Archaea* domain possess specific physiological characteristics that unite them, such as the composition and structure of their membrane lipids [3].

*Archaea*, for the most part, resemble *Bacteria* in size and shape, but they possess important differences at genetic, molecular, and metabolic levels. Remarkable differences are present in the machinery and functionality of the information processing systems. For example, the histones in some *Archaea* [4], the similarity of proteins involved in transcription and translation, and the structure of the ribosome are more closely related to those of eukaryotes than *Bacteria* [5]. The most striking similarities between *Archaea* and eukaryotes are observed in DNA replication, one of the most conserved processes in living organisms, where the

components of the replicative apparatus and the overall process closely resemble a simplified version of the eukaryotic DNA replication system [6, 7]. However, a few characteristics of the archaeal information processing system are shared only with *Bacteria*, such as the circular topology of the chromosome, the small size of the genome, and the presence of polycistronic transcription units and Shine-Dalgarno sequences in the mRNA instead of Kozak sequences [8, 9]. These observations suggest that archaeal DNA replication is a process catalyzed by eukaryotic-like proteins in a bacterial context [10]. In addition, there are specific features which belong only to the domain *Archaea*, such as the DNA polymerase D, which add another level of complexity and uniqueness to these microorganisms. While a number of evolutionary hypotheses have been proposed to account for the mosaic nature of this process [11, 12], comparative genomics and ultrastructural studies support two alternative scenarios. The archezoan scenario holds that an archaeal ancestor developed a nucleus and evolved into a primitive archezoan, which later acquired an  $\alpha$ -proteobacterium to form the mitochondrion and evolved into the eukaryotes. The symbiogenesis scenario holds that an archaeal ancestor plus an  $\alpha$ -proteobacterium formed a chimeric cell, which then evolved into the eukaryotic cell [12]. A major difference between these hypotheses is the time when the first

eukaryotes evolved. In the archezoan scenario, the ancestors of the early eukaryotes are as ancient as the two prokaryotic lineages, the Bacteria and Archaea. In the symbiogenesis scenario, the eukaryotes evolved relatively late after the diversification of the ancient archaeal and bacterial lineages.

The cultivated *Archaea* are taxonomically divided into five phyla. *Crenarchaeota* and *Euryarchaeota* are best characterized, and this taxonomic division is strongly supported by comparative genomics. A number of genes with key roles in chromosome structure and DNA replication are present in one phylum but absent in the other. *Crenarchaeotes* also exhibit a vast physiological diversity, including aerobes and anaerobes, fermenters, chemoheterotrophs, and chemolithotrophs [2]. The majority of the cultivated *crenarchaeotes* are also thermophiles [13]. *Euryarchaeotes* exhibit an even larger diversity, with several different extremophiles among their ranks in addition to large numbers of mesophilic microorganisms. Interestingly, all the known methanogens belong to this phylum. Additionally, environmental sequencing, 16S rRNA analysis, and cultivation efforts provide strong evidence for three additional phyla: *Thaumarchaeota*, *Korarchaeota*, and *Aigarchaeota* [14–16]. A sixth phylum has also been proposed previously, the *Nanoarchaeota*. This phylum includes the obligate symbiont *Nanoarchaeum equitans*, which has a greatly reduced genome [17]. However, a more complete analysis of its genome suggests that *N. equitans* corresponds to a rapidly evolving euryarchaeal lineage and not an early archaeal phylum [18].

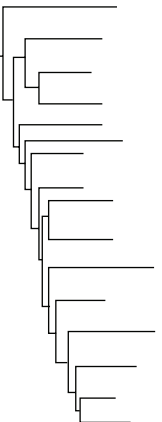
## 2. General Overview of the DNA Replication Process

The general process of DNA replication is conserved among the three domains of life, but each domain possesses some functional modifications and variations in key proteins. DNA replication may be divided into four main steps. Step 1 (Preinitiation) starts when specific proteins recognize and bind the origin of replication, forming a protein-DNA complex. This complex recruits an ATP-dependent helicase to unwind the double-stranded DNA (dsDNA). The single-stranded DNA (ssDNA) formed is then protected by ssDNA-binding proteins (SSB). Step 2 (Initiation) corresponds to the recruitment of the DNA primase and DNA polymerase. Step 3 (Elongation) corresponds to the duplication of both strands of DNA at the same time by the same unidirectional replication machine. Because of the antiparallel nature of the DNA, the leading strand is copied continuously, and the lagging strand is copied discontinuously as Okazaki fragments [10]. During replication, a sliding-clamp protein surrounds the DNA and binds to the polymerase to increase processivity or the number of nucleotides added to the growing strand before disassociation of the polymerase from the template. Step 4 (Maturation) corresponds to the completion of the discontinuous replication in the lagging strand by the actions of RNase, DNA ligase, and Flap endonuclease. The variability of the players involved in each of these processes between the three domains of life and more specifically within the *Archaea* domain is discussed in further detail below.

## 3. DNA Replication Components

**3.1. Preinitiation (Origin of Replication and Origin Recognition).** The number of origins of replication in a genome is correlated with phylogeny and varies among the different domains [19]. Members of the *Bacteria* domain usually possess only one replication origin, but eukaryotic chromosomes contain multiple replication origins. This difference was generally accepted as a clear divisor in DNA replication of eukaryotes and prokaryotes. However, members of the *Archaea* domain display either one or multiple origins of replication. The origins of replication in archaea are commonly AT-rich regions which possess conserved sequences called origin of recognition boxes (ORB) and are well conserved across many archaeal species. Also, smaller versions of the ORBs, called mini-ORBs, have been identified [20]. Members of the *Crenarchaeota* phylum display multiple origins of replication. For example, species belonging to the *Sulfolobales* order contain three origins of replication [20, 21]. Likewise, two and four origins of replication have been found in members of the *Desulfurococcales* (*Aeropyrum pernix*) and *Thermoproteales* (*Pyrobaculum calidifontis*), respectively [22, 23]. In addition to shortening the time required for replication, multiple origins may play a role in counteracting DNA damage during hyperthermophilic growth [24]. In contrast, most members of the *Euryarchaeota* phylum seem to possess only one origin of replication. For example, species belonging to the *Thermococcales* (*Pyrococcus abyssi*) and *Archaeoglobales* (*Archaeoglobus fulgidus*) orders contain only one origin of replication [25–27]. However, members of the order *Halobacteriales* (e.g., *Halobacterium* sp. NRC-1, *Haloferax volcanii*, and *Haloarcula hispanica*) possess numerous origins of replication that may be part of an intricate replication initiation system [28–30]. In fact, a recent report demonstrated that the different origins of replication in *Haloarcula hispanica* are controlled by diverse mechanisms, which suggest a high level of coordination between the multiple origins [31]. In contrast, the deletion of a single origin of replication in *Haloferax volcanii* affects replication dynamics and growth, but the simultaneous deletion of all four origins of replication accelerated growth compared to the wild type strain [32]. These results suggest that initiation of replication in *Haloferax volcanii* may be dependent of homologous recombination, as it is in some viruses, and generate questions about the purpose of the replication origins [32]. With the exception of *Methanothermobacter thermautotrophicus*, the DNA replication origin has not been experimentally demonstrated in the methanogens [33]. However, bioinformatics analysis by GC-skew in *Methanosarcina acetivorans* [34] and Z-curve analysis for *Methanocaldococcus jannaschii* and *Methanosarcina mazei* suggest a single origin of replication. A similar analysis of *M. maripaludis* S2 was inconclusive [35]. The fact that the order *Halobacteriales* is the only euryarchaeote that displays multiple origin of replication suggests that there may have been a lineage-specific duplication in this group. Moreover, the number of origins is not highly conserved even within a single archaeal phylum.

The origin of replication is recognized by specific proteins that bind and melt the dsDNA and assist in the loading of



Order	Phylum	Number of genomes	Number of homologs for replication genes									
			ORC/CDC6	GINS15	GINS23	RPA	MCM	RFCS	PCNA	PolB	DPI	DP2
Nitrosopumilales	<i>Thaumarchaeota</i>	2	1	0 or 1	0 or 1	1	1	1	1	1	1	1
<i>Thermoproteales</i>		14	1*	0	0	0*	1	1 to 3	1 to 3	2*	0	0
<i>Desulfurococcales</i>	<i>Crenarchaeota</i>	9	2	0	1*	0 or 1	1	1*	3*	2*	0	0
<i>Sulfolobales</i>		15	3	1	1	1	1	1	2*	3	0	0
Not named	<i>Korarchaeota</i>	1	2	0	0	0	1	1	1	2	1	1
<i>Nanoarchaeota</i>		1	1	0	0	1	1	1	1	1	1	1
<i>Thermococcales</i>		11	1	1	1*	3	1*	1	1*	1	1	1
<i>Methanopyrales</i>		1	0	0	0	0	1	1	1	1	1	1
<i>Methanobacteriales</i>		7	1 or 2	0 or 1	0	1*	1	1	1	2	1	1
<i>Methanococcales</i>	<i>Euryarchaeota</i>	15	0	0	0	2 to 4	2 to 8	1	1*	1	1	1
<i>Thermoplasmatales</i>		4	1*	0*	0	1	1	1	1	1	1	1
<i>Archaeoglobales</i>		4	2 or 3	0	0	0*	1 or 2	1	1	1*	1	1
<i>Halobacteriales</i>		15	3 to 15	0	0	1	1 or 2	2*	1	1*	1	1
<i>Methanomicrobiales</i>		6	2*	0	0	1	1	2	1	1 or 2	1	1
<i>Methanocellales</i>		1	2	0	0	1	1	1	1	3	1	1
<i>Methanosarcinales</i>		9	2*	0	0	3	1*	1*	1	1*	1*	1

FIGURE 1: Homologs for some key genes involved in archaeal DNA replication. Archaeal orders are phylogenetically organized following a rooted maximum likelihood tree of *Archaea* based on 53 concatenated ribosomal proteins [36]. The homology search was performed by RAST v4.0 (Rapid Annotation using Subsystem Technology), and the annotated data was viewed through the SEED viewer (<http://www.theseed.org/>). A total of 115 archaeal genome sequences were obtained from NCBI and uploaded into the RAST server. RAST annotation was performed using default parameters with the genetic code for *Bacteria* and *Archaea*. The replication proteins homologs were checked through the SEED subsystem DNA replication. Homology results of *Thermococcus kodakaraensis* KOD1, *Pyrococcus furiosus* DSM 3638, and *Thermofilum pendens* Hrk 5 were searched for replication protein homologs using BLAST (blastx v2.2.28+). Results with homology coverage of >80% and E-values less than 0.001 were considered as real homologs. The results were also supplemented with data from literature reviews and BLAST searches. When indicated in the figure, zero (0) homologs means that no homolog was found for a specific gene by using the described methodology. Asterisks indicate that one or two of the analyzed microorganisms possess an exception for that specific feature. Exceptions noted are (feature/order) ORC1/CDC6/*Thermoproteales*, *Thermofilum pendens* Hrk5 (3 homologs), and *Thermosphaera aggregans* DSM11486 (2 homologs); ORC1/CDC6/*Thermoplasmatales* and *Picrophilus torridus* DSM 9790 (1 homolog); ORC1/CDC6/*Methanomicrobiales* and *Methanoplanus petrolearius* DSM11571 (4 homologs); ORC1/CDC6/*Methanosarcinales* and *Methanosalsum chilinae* DSM4017 (4 homologs); GINS15/*Thermoplasmatales* and *Thermoplasma acidophilum* DSM1728 (1 homolog); GINS23/*Desulfurococcales*, *Staphylothermus hellenicus* DSM12710 (no homolog), and *Staphylothermus marinus* F1 (no homolog); GINS23/*Thermococcales*, *Pyrococcus horikoshii* OT3 (2 homologs), RPA/*Thermoproteales*, *Thermofilum pendens* Hrk5 (3 homologs), and *Thermosphaera aggregans* DSM11486 (1 homolog); RPA/*Methanobacteriales* and *Methanothermobacter feravidus* DSM2088 (no homolog); RPA/*Archaeoglobales* and *Archaeoglobus fulgidus* DSM4304 (1 homolog); MCM/*Thermococcales* and *Thermococcus kodakaraensis* KOD1 (3 homologs); MCM/*Methanosarcinales* and *Methanosarcina acetivorans* C2A (2 homologs); RFCS/*Desulfurococcales* and *Hyperthermus butylicus* DSM 5456 (2 homologs); RFCS/*Halobacteriales* and *Haloquadratum walsbyi* DSM16790 (1 homolog); RFCS/*Methanosarcinales*, *Methanosaeta concilii* CG6 (2 homologs), and *Methanosaeta thermophila* PT (2 homologs); PCNA/*Desulfurococcales* and *Ignisphaera aggregans* DSM17230 (1 homolog); PCNA/*Sulfolobales* and *Metallosphaera cuprina* Ar-4 (1 homolog); PCNA/*Thermococcales*, *Pyrococcus horikoshii* OT3 (2 homologs), and *Thermococcus kodakaraensis* KOD1 (2 homologs); PCNA/*Methanococcales*, *Methanococcus maripaludis* S2 (2 homologs), and *Methanoterris igneus* Kol5 (2 homologs). PolB/*Thermoproteales*, *Caldivirga maquilingensis* IC-167 (3 homologs), and *Pyrobaculum calidifontis* JCM 11548 (3 homologs); PolB/*Desulfurococcales*, *Staphylothermus hellenicus* DSM12710 (3 homologs), and *Staphylothermus marinus* F1 (3 homologs); PolB/*Archaeoglobales*, *Archaeoglobus fulgidus* DSM4304 (2 homologs); PolB/*Halobacteriales*, *Halorhabdus utahensis* DSM12940 (2 homologs); PolB/*Methanosarcinales*, *Methanosaeta concilii* GP6 (2 homologs); DPI/*Methanosarcinales*, *Methanococcoides burtonii* DSM6242 (2 homologs).

the replicative helicases. In *Bacteria* this protein is DnaA, which binds to the DnaA boxes [37]. In eukaryotes, a protein complex known as ORC (origin recognition complex) composed of 6 different proteins (Orcl-6) binds at the replication origin and recruits other proteins [38]. In *Archaea*, the candidate for replication initiation is a protein that shares homology with the eukaryotic Orcl and Cdc6 (cell division cycle 6) proteins and is encoded by almost all archaeal genomes [39]. Their genes are commonly located adjacent to the origin of replication. These proteins, termed Orcl/Cdc6 homologs, have been studied in detail in representatives of the *Euryarchaeota* (*Pyrococcus*) and *Crenarchaeota* (*Sulfolobus*) and have been shown to bind to the ORB region with high specificity [20, 26, 40]. In addition, purified Orcl/Cdc6 from

*S. solfataricus* binds to ORB elements from other crenarchaeotes and euryarchaeotes *in vitro*, suggesting that these proteins recognize sequences conserved across the archaeal domain [20]. The crystal structures of these proteins from *Pyrobaculum aerophilum* [41] and *Aeropyrum pernix* [42] have been solved, contributing to the overall understanding of their activity during the initiation of replication (reviewed in [43]). In general, the number of Orcl/Cdc6 homologs found in archaeal genomes varies between one and three (Figure 1). However, there are some interesting exceptions. Members of the order *Halobacteriales* possess between 3 and 15 Orcl/Cdc6 homologs, which may reflect the large number of origins of replication. Interestingly, like other members of the *Methanosarcinales*, the extremely halophilic methanogen

*Methanohalobium evestigatum* Z-7303 only possesses two Orc1/Cdc6 homologs. Thus, the ability to grow at extremely high concentrations of salt and to maintain high intracellular salt concentrations does not necessarily require a large number of Orc1/Cdc6 homologs. Although most members of the order *Methanomicrobiales* possess two copies of this gene, *Methanoplanus petrolearius* DSM 11571 possesses four copies. Thus, the number of Orc1/Cdc6 homologs varies considerably among the euryarchaeotes. In an extreme case, representatives of the orders *Methanococcales* and *Methanopyrales* do not possess any recognizable homologs for the Orc1/Cdc6 protein [7, 44], and the mechanics of replication initiation in these archaea remain an intriguing unknown of the archaeal replication process.

**3.2. Initiation (DNA Unwinding and Primer Synthesis).** After Orc1/Cdc6 proteins bind to the origin of replication, a helicase is recruited to unwind the dsDNA. In *Bacteria* the replicative helicase is the homohexamer DnaB. In eukaryotes the most likely candidate for the replicative DNA helicase is the MCM (minichromosome maintenance) complex, which is a heterohexamer that is activated when associated with other replicative proteins, forming the CMG (Cdc45-MCM-GINS) complex [45–48]. As in eukaryotes, in *Archaea* the most probable candidate for the replicative helicase is the MCM complex. Intriguingly, the MCM complexes in eukaryotes and archaea possess a 3′-5′ unwinding polarity, but the bacterial DnaB helicase unwinds duplex DNA on the opposite direction with a distinct 5′-3′ polarity [49–51].

Most archaeal genomes studied so far encode one MCM homolog. Its helicase activity has been demonstrated *in vitro* for several archaea, including *Methanothermobacter thermoautotrophicus* ΔH, where the MCM protein forms a dimer of hexamers [52, 53]. The *in vivo* interaction between the MCM complex and the Orc1/Cdc6 proteins has been demonstrated in other *Archaea* [54–56]. Indeed, through an *in vitro* recruiting assay, Orc1/CDC6 from *P. furiosus* has been demonstrated as the possible recruiter of the MCM complex to the origin of replication [57]. A recent study identified thirteen species of archaea with multiple *mcm* genes encoding the MCM complexes (Figure 1, [58]). The number of *mcm* homologs is especially high in the order *Methanococcales*, where different representatives possess from 2 to 8 copies [59, 60]. For instance, *Methanococcus maripaludis* S2 possesses four homologs of the MCM protein [44]. Through a shotgun proteomic study, peptides for three of them have been detected [61], suggesting that multiple MCMs are expressed and functional. However, a genome-wide survey of gene functionality in *M. maripaludis* demonstrated that only one of these genes, MMP0030, was likely to be essential for growth [62]. In addition, coexpression of recombinant MCMs from *M. maripaludis* S2 allowed copurification of all four proteins [59], suggesting that *M. maripaludis* may form a heterologous multimeric MCM complex. However, because MMP0030 protein is the only *mcm* gene essential for growth, a homologous multimeric complex may also be possible. Similar results have been found in *Thermococcus kodakaraensis*, in which three MCM homologs are present but only one is essential [63]. The *T. kodakaraensis*

MCM system is suggested to form homologous multimeric complexes [64]. A recent study proposed that two of the MCM homologs that are conserved among representatives of the order *Methanococcales* are a consequence of an ancient duplication that occurred prior to the divergence of this group [59]. It has also been proposed that the large number of MCM homologs in the order *Methanococcales* is a direct consequence of mobile elements, which may have taken advantage of the ancient duplication of the MCM genes to take over the replication system by forming cellular MCM heterocomplexes [58]. In any case, the large number of MCM homologs in the order *Methanococcales* may be a product of an intricate and complex evolutionary history. Whether or not it is related to the absence of the replication initiation protein Orc1/Cdc6 is an interesting possibility [58].

In eukaryotes the MCM complex is not active on its own and requires the association of two accessory factors, the heterotetrameric GINS complex (from the Japanese goishi-ni-san meaning 5-1-2-3, after the four subunits Sld5, Psf1, Psf2, and Psf3) and the Cdc45 protein. This complex, called CMG (Cdc45-MCM-GINS), is thought to be the active replicative helicase [47]. Homologs for GINS subunits have been identified in *Archaea* by bioinformatics analysis. One homolog is most similar to the eukaryotic proteins Psf2 and Psf3 (GINS23) and is common in the crenarchaeotes. Another homolog is most similar to the eukaryotic proteins Psf1 and Sld5 (GINS15) and is found largely in the euryarchaeotes [65]. The GINS complex is expected to be essential for DNA replication, but the distribution of these two different GINS homologs suggests that the presence of one of them is sufficient for archaeal DNA replication. Some *Crenarchaeota* (*S. solfataricus*) and some *Euryarchaeota* (*P. furiosus* and *T. kodakaraensis*) possess both homologs and form heterotetramers similar to the ones found in eukaryotes with a ratio 2:2 [66–68]. To find homologs of GINS15 and GINS23 in *Archaea*, two known GINS15 (from *Sulfolobus solfataricus* P2 and *Thermococcus gammatolerans* EJ3) and three known GINS23 (from *Sulfolobus solfataricus* P2, *Pyrococcus yayanosii* CH1, and *Thermococcus gammatolerans* EJ3) were used for homology searches. These analyses failed to identify GINS homologs in several archaeal groups (Figure 1). These results suggest that there are other unknown GINS homologs in archaea or that the sequences have diverged enough between genera not to be recognized by homology searches. In any case, a more comprehensive analysis of GINS proteins is necessary to fully understand their function and distribution in *Archaea*.

The detailed interactions of the GINS and MCM proteins in archaea appear to be highly variable. Although the *S. solfataricus* GINS and MCM proteins physically interact *in vitro*, MCM helicase activity was not stimulated in the complex [66]. In contrast, the MCM helicase activity of *T. kodakaraensis* and *P. furiosus* is clearly stimulated by their GINS complexes [64, 67]. The crystal structure of the GINS complex of *T. kodakaraensis* was recently determined. The backbone structure and the assembly are similar to the human complex with some notable differences [68]. Interestingly, many other euryarchaeotes possess only the homolog to GINS15. One of those, *Thermoplasma acidophilum*, forms

a homotetrameric complex [68, 69]. Moreover, *in vitro* the *T. acidophilum* MCM helicase activity was not affected by the GINS complex [69]. These results suggest that other proteins may be involved in the formation of a stable helicase in many Archaea. *M. maripaludis* S2 possesses a hypothetical protein which may correspond to the gene encoding for GINS15, which is probably essential for growth [60, 62]. No homologs of GINS23 are present in the genome of *M. maripaludis*.

The replication related Cdc45 protein is ubiquitous in eukaryotes, but its exact role has not yet been elucidated. Interestingly, homologs of the Cdc45 protein are not present in Archaea. However, a bioinformatics analysis revealed that the eukaryotic Cdc45 and the prokaryotic RecJ, which is a conserved 5'-3' exonuclease in most bacteria and archaea, possess a common ancestry and share a homologous DHH domain [70]. These results suggest that the archaeal RecJ may substitute for the eukaryotic Cdc45 during replication. Circumstantial evidence supports this conclusion. In *S. solfataricus*, a homolog of the DNA binding domain of RecJ was purified together with a GINS protein [66]. Similarly, in *T. kodakaraensis*, the RecJ homolog (TK1252p) copurified with proteins of the GINS complex [71] and formed a stable *in vitro* association with the GINS complex [72]. Two RecJ homologs are found in *M. jannaschii* (MJ0977 and 0831), and they partially complement a *recJ* mutation in *E. coli*. The recombinant MJ0977 also possesses high levels of thermostable single-stranded DNA degrading activity similar to RecJ [73]. In *M. maripaludis* S2, four proteins possess the DHH domain with similarity to the *M. jannaschii* RecJ homologs (MMP1682, 0547, 1078, and 1314). However, none of them was essential for growth, suggesting that the activities were redundant [62]. Alternatively, another protein, which has a low similarity with *M. jannaschii* RecJ homolog (MMP0285) but possesses the DHH domain, was possibly essential for growth [62]. However, this protein possesses other domains related to transport, which make it an unlikely candidate for RecJ. These results are not definitive, and more experimental data is needed to assign a function to this protein.

As soon as the DNA is unwound, ssDNA is protected from nucleases, chemical modification, and other disruptions by ssDNA binding proteins. These proteins are present in all three domains of life and are called SSB in *Bacteria*, where they form homotetramers or homodimers [74, 75], and RPA (replication protein A) in eukaryotes, where they form a stable heterotrimer composed of 70, 32, and 14 kDa proteins [76]. Although all ssDNA binding proteins contain different combinations of the oligonucleotide/oligosaccharide-binding fold or OB-fold [77], the sequence similarity is low among RPA and the bacterial SSBs [43]. In Archaea, different types of ssDNA binding proteins have been reported. In crenarchaeotes, an ssDNA binding protein is only well characterized in *S. solfataricus*. The *S. solfataricus* ssDNA binding protein contains only one OB-fold and can adopt a heterotetramer conformation [78]. However, the protein also appears to exist as a monomer [79]. This protein appears to be more similar to the bacterial proteins in sequence, but its structure is more similar to that of the eukaryotic RPAs [43, 80]. While homologs are recognizable in other members of the *Sulfolobales* and most other Archaea, they are absent from

the genomes of many members of the closely related order *Thermoproteales* with some few exceptions (Figure 1) [81].

RPAs have been studied from different representatives of the euryarchaeotes, revealing additional diversity. *M. jannaschii* and *M. thermautotrophicus* possess a single RPA subunit, which shares amino acid similarity with the eukaryotic RPA70 and possesses four and five OB-folds, respectively [82, 83]. *P. furiosus* possesses three different RPA subunits, which form a stable heterotrimer that is involved in homologous recombination *in vitro* [84]. The *P. furiosus* RPA, therefore, is similar in subunit organization to the eukaryotic RPA. The other members of the *Thermococcales* genera also possess three RPA homologs, which likely form the homotrimeric complex described for *P. furiosus* (Figure 1). *Methanosarcina acetivorans* also possesses three different RPA proteins (RPA1-3), with four, two, and two OB-folds, respectively. However, they do not interact with each other and appear to function as homodimers [85]. This arrangement is common among the *Methanosarcinales* but is not found in the closely related methanogen orders such as the *Methanocellales* (Figure 1). Members of the order *Methanococcales* possess two to four RPA homologs. The genome of *M. maripaludis* S2 contains three possible RPA homologs (MMP0122, 0616, and 1032) [60]. Only MMP0616 and 1032 were likely to be essential for growth [62], suggesting a possible different complex configuration for RPA proteins in this archaeon. It has been hypothesized that the diversity in OB-folds in the archaeal RPAs is a direct consequence of homologous recombination [86].

DNA synthesis starts with the production of a RNA primer, since the DNA polymerases that replicate genomes are incapable of *de novo* DNA synthesis. The RNA primer is synthesized by a DNA-dependent RNA polymerase or primase. DnaG, a single subunit protein, is the primase in *Bacteria*. In eukaryotes, a two-subunit primase, consisting of a small catalytic subunit (PriS) and a large subunit (PriL), is found in a complex with DNA polymerase  $\alpha$  and its accessory B subunit [87]. In Archaea, the first biochemically characterized primase was the eukaryotic primase-like small subunit PriS in *M. jannaschii* [88]. Later, homologs of the PriS and PriL subunits were described in several *Pyrococcus* species [89–93]. The eukaryotic-like DNA primase has also been characterized in the crenarchaeote *S. solfataricus* [94, 95], where it has been shown to interact with MCM through GINS23 [66]. A unique polymerization activity across discontinuous DNA templates has been characterized in the PriSL of *S. solfataricus*, suggesting that this primase may be involved in double-stranded break repair in Archaea [96]. So far the eukaryotic-like primases found in archaea exhibit similar properties. PriS functions as the catalytic subunit and PriL modulate its activity [87]. In addition, the archaeal eukaryotic-like primase has the unique ability to synthesize both DNA and RNA *in vitro*. Indeed, the *P. furiosus* PriS synthesizes long DNA fragments, but the addition of PriL regulates the process by decreasing the DNA polymerase activity, increasing the RNA polymerase activity, and decreasing the product length [90]. Additionally, studies of the *P. abyssi* enzyme suggest that DNA primase may also be involved in DNA repair because of the DNA polymerase, gap filling, and strand displacement activities

also present *in vitro* [93]. Interestingly, the archaeal primase small subunit resembles DNA polymerases from the Pol X family in sequence and structure, suggesting a similarity in their catalytic mechanism [96].

Homologs of the bacterial DnaG primase are also found in archaea. However, *in vitro* studies of the *P. furiosus* enzyme failed to detect primer synthesis activity, and the *T. kodakaraensis* DnaG was copurified with proteins of the exosome [71]. Also, in *S. solfataricus*, DnaG primase was found to be a core exosome subunit involved in RNA degradation [97, 98]. However, recent studies demonstrated that the *S. solfataricus* DnaG homolog has limited primer synthesis activity, and a dual primase system with PriSL has been proposed to function during DNA replication [99]. Hu et al. hypothesized an interesting theory which suggests that LUCA employed a dual-primase system comprising DnaG and PriSL, which also served roles in RNA degradation and the nonhomologous end joining (NHEJ) pathway. The system was inherited by all three domains. In the *Archaea* domain, it retained its original functions. However, in the *Bacteria* domain DnaG became the major replicative primase and PriSL evolved into the Pol domain of LigD involved in the NHEJ pathway. In contrast, DnaG was lost in the *Eukaryota* domain, and PriSL became the only replicative primase. In addition, it also evolved into the DNA Pol X family involved in NHEJ [96].

Like other *Archaea*, *M. maripaludis* S2 possesses both the eukaryotic- and bacterial-like primases. A new genome-wide survey of this methanoarchaea revealed that genes encoding both subunits of the eukaryotic like primase, PriS, and PriL (MMP0009 and MMP0071) were likely to be essential for growth, which is consistent with a role in replication. Similar observations have been demonstrated in other *Archaea*, such as *Halobacterium* sp. NRC-1 and *Haloferax volcanii* [93, 100]. In contrast the DnaG primase (MMPI286) was nonessential, which would be consistent with a role in the exosome [62]. These results suggest that in *M. maripaludis* and probably other euryarchaeotes, a dual primase system is not present.

**3.3. Elongation (Replicative DNA Polymerase).** The primer synthesized by the primase is further extended by a replicative DNA polymerase. DNA polymerases have been classified into seven families based on their amino acid sequence similarity (A, B, C, D, E, X, and Y) [101–105]. In *Bacteria*, *E. coli* possesses five DNA polymerases, Pol I–V, where the first three belong to families A, B, and C, respectively, and the last two belong to the Y family. The major replicative DNA polymerase in *E. coli* is Pol III, a member of the C family, which synthesizes DNA with high processivity when functioning within the holoenzyme [106]. In eukaryotes, a great diversity of DNA polymerases is also found, but DNA replication requires three replicative DNA polymerases which belong to family B (Pol $\alpha$ , Pol $\delta$ , and Pol $\epsilon$ ) [107]. In *Archaea*, fewer types of DNA polymerases are present, but they have an interesting evolutionary division. Representatives of the DNA polymerase B family are found in all *Archaea*. The D family DNA polymerase is present in every phyla studied so far with the exception of the *Crenarchaeota* (Figure 1). Lastly, a member of the Y family has been identified and biochemically

characterized in *S. solfataricus* [108] and the *Methanosarcinales*, in which several species harbor two homologs [109]. Even though the activity of DNA polymerases belonging to the Y family in archaea is not fully understood, they are suggested to play an important role in DNA repair [109].

Representatives of the DNA polymerase family B have been identified in all *Archaea*. The crenarchaeotes possess at least two family B DNA polymerases (PolB I and PolB II, and in some cases PolB III) [102, 110, 111]. The recombinant *S. solfataricus*, *Pyrodictium occultum*, and *A. pernix* enzymes have been characterized [102]. In contrast, euryarchaeotes only contain PolB I. The recombinant *P. furiosus* enzyme has been characterized [112]. The PolB I enzymes have similar amino acid sequence and overall structure, and they possess a potent 3'-5' exonuclease proofreading activity [102]. A unique property of the archaeal family B DNA polymerase is the ability to stall DNA polymerization in the presence of uracil or hypoxanthine deaminated bases [113, 114]; this property helps to prevent the copy of template-strand uracil and the transmission of fixed mutations to progeny [115]. Commonly, cytosine deamination converts G:C base pairs into the promutagenic G:U mismatches, which result in 50% of the offspring containing an A:T transition mutation after replication [116]. In addition, when A-U base pairs are formed, the loss of the 5-methyl moiety upon replacement of thymine can exert a detrimental effect on protein DNA interactions. Interestingly, in some crenarchaeotes and euryarchaeotes, one of the B family DNA polymerase paralogs possesses disrupted versions of the sequence motifs that are essential for catalysis. Possibly, these enzymes possess a structural role [117].

DNA polymerase family D (PolD) is a novel enzyme that was originally discovered in *P. furiosus* [118]. It has since then been identified in all euryarchaeotes [119]. For a long time, this enzyme was considered a euryarchaeote-specific polymerase, but the three newly discovered archaeal phyla were also found to possess genes for PolD [14–16]. PolD is a heterodimer with a small subunit (DP1) and a large subunit (DP2). It has been proposed that the large subunit harbors the polymerase activity, although its sequence is very different from other hitherto described DNA polymerases. The small subunit possesses high similarity with the noncatalytic B-subunits of the eukaryotic DNA polymerases  $\alpha$ ,  $\delta$ , and  $\epsilon$  [120]. Studies of the *M. jannaschii* enzyme demonstrated that the small subunit possesses a strong 3'-5' exonuclease activity, suggesting that it may be involved in proofreading activity [7, 121]. Efficient polymerase and proofreading activity have also been detected from a purified PolD of *P. furiosus*, and the residues Asp-1122 and Asp-1124 are essential for the polymerization reaction in *P. horikoshii* [122, 123]. Every euryarchaeote examined possessed one homolog for DP1 and DP2, except for *Methanococcoides burtonii* DSM 6242, which possesses two homologs for DP1 (Figure 1).

Because the crenarchaeotes only possess family B DNA polymerases, one of them must be the replicative polymerase. However, in euryarchaeotes it was not clear which enzyme was involved in replication. Studies in *Halobacterium* sp. NRC-1 showed that both PolB and PolD were essential for viability, and it was proposed that they could be working

together at the replication fork, synthesizing the leading and lagging strand, respectively [100, 124]. However, in other *Euryarchaeota*, PolB does not appear to be essential. For instance, a recent report demonstrated that both subunits of PolD were essential, but PolB was nonessential for the growth of *M. maripaludis* S2 [62]. Likewise, in *T. kodakaraensis* PolD can be coisolated with different proteins of the archaeal replication fork, but PolB was mainly coisolated with proteins of unknown function [71]. A deletion of the PolB in *T. kodakaraensis* also had no detectable effect on cell viability [125]. Interestingly, *T. kodakaraensis* PolB mutants have increased sensitivity to UV radiation. These results together suggest that PolD is the essential replicative DNA polymerase in these euryarchaeotes and not a PolB family polymerase. In this scenario, PolB may have some other crucial function in *Halobacterium* and possibly other closely related species, which it is not conserved throughout the phylum. Finally, it has been demonstrated that purified PolD from *P. abyssi* is able to perform strand displacement and RNA primer extension *in vitro*. Purified PolB, on its own, cannot perform these activities *in vitro*, but in the presence of PCNA strand displacement activity is stimulated [124, 126]. A recent *in vitro* study in *P. abyssi* demonstrated that RNase HIII is required to initiate strand displacement DNA synthesis by PolB [127].

**3.4. Elongation (Other Accessory Proteins).** Purified DNA polymerases possess low processivity; however, the addition of an accessory factor, the sliding clamp, gives DNA polymerases the required processivity to replicate genomes. This factor, whose structure resembles a doughnut, functions as a molecular platform which recruits several replication-associated enzymes and act together with DNA and the polymerase to stabilize their interactions during replication [128]. The structure of sliding clamps is conserved in the three domains of life. In *Bacteria*, the sliding clamp or  $\beta$ -subunit is a homodimeric ring [129]. In eukaryotes, the sliding clamp or proliferating cell nuclear antigen (PCNA) is a homotrimeric ring [128]. In *Archaea*, the majority of crenarchaeotes have multiple PCNA homologs which form heterotrimeric rings [130, 131] (Figure 1). In contrast, most euryarchaeotes possess a single PCNA homolog (Figure 1). Interactions of DNA polymerase with the sliding clamp are mediated through a common motif, called the PCNA Interacting Protein (PIP) box [132]. The three dimensional structures of PCNA-PolB and the PolB-PCNA-DNA complex from *P. furiosus* have been solved, shedding light on the interactions between these molecules [133, 134]. In addition, a study demonstrated that the PolD/PCNA and PolB/PCNA interactions require two and one PIP boxes, respectively [135]. Recently, through a protein-interaction network of *P. abyssi*, a previously unknown protein with nuclease activity (Pab0431) and a PIP canonical motif was discovered to interact with PCNA, suggesting a possible involvement in DNA replication [136]. *T. kodakaraensis* is the only well-documented example of an euryarchaeote with two homologs for PCNA, TK0535 (PCNA1), and TK0582 (PCNA2). Recent work demonstrated both proteins form stable homotrimeric rings that interact with *T. kodakaraensis* PolB *in vitro* [137]. A more detailed study of the two homologs of PCNA in *T.*

*kodakaraensis* demonstrated that both homologs stimulate *in vitro* the primer extension activity of PolB, but only PCNA1 and not PCNA2 stimulates the same activity of PolD [138]. Also, the same report showed that the *pcna2* gene can be disrupted without causing growth deficiencies, but it was not possible to isolate mutants of *pcna1*. These results suggest that PCNA1 but not PCNA2 is essential for DNA replication [138]. It has been proposed that one of the PCNA genes was acquired by lateral gene transfer [139]. Members of the order *Methanococcales* possess one PCNA homolog with two exceptions. *M. maripaludis* S2 possesses two PCNA homologs (MMPI126 and 1711) [60]. Both genes appeared to be essential for growth, suggesting that they both play an important role in replication [62]. The gene MMPI1711 possesses high similarity to both *T. kodakaraensis* genes and is likely the true methanococcal PCNA. In contrast, MMPI126 possesses only low similarity to the *T. kodakaraensis* genes and also contains an S-adenosylmethionine-dependent methyltransferases domain. Thus, it is likely to possess some alternative function. *Methanotorris igneus* Kol 5 also possesses two PCNA homologs.

For PCNA to assemble around DNA, a specific loading factor is required. In *Bacteria* the loading factor is known as  $\gamma$ -complex and comprises three different subunits in a  $\gamma_3$ - $\delta$ - $\delta'$  stoichiometry [140]. In eukaryotes, the loading factor is known as Replication Factor C (RFC) and is a heteropentameric complex comprising one large subunit and four different small subunits [141]. In contrast, in *Archaea* the RFC consists of two different proteins, a small subunit RFCS and a large subunit RFCL. They also form a heteropentameric complex in a 4:1 ratio (RFCS:RFCL). Interestingly, many crenarchaeotes possess one homolog of RFCL and two or three homologs of RFCS. Alternatively, most euryarchaeotes possess only one RFCL and RFCS homolog. The exceptions are some members of the orders *Halobacteriales*, *Methanomicrobiales*, and *Methanosarcinales*, which possess two RFCS homologs (Figure 1). Structural and biochemical studies of RFC have been conducted in several euryarchaeotes, such as *Archaeoglobus fulgidus* and *Pyrococcus* species [142–144]. An interesting case is observed in *Methanosarcina acetivorans*. This RFC possesses three subunits (RFC1, 2, L) found in a ratio 3:1:1 [145]. Homologs for these three subunits are also present in most of the genomes of the haloarchaea [145]. It is inferred that this type of RFC complex represents an intermediary form transitional between the canonical archaeal RFC complex of one small and one large subunit and the eukaryotic RFC complex of four different small and one large subunit [146]. The organization and spatial distribution exhibit a similarity to the *E. coli* minimal  $\gamma$ -complex, but the function of the subunits is probably not conserved [145].

**3.5. Maturation.** During DNA replication, the lagging strand is discontinuously synthesized by extending the RNA primers or Okazaki fragments. This discontinuously synthesized DNA strand requires maturation to form a single, covalently closed strand to end the replication process. In *Archaea*, Okazaki fragments were first demonstrated during replication in *P. abyssi* and *Sulfolobus acidocaldarius* [147]. During replication, these RNA primers are replaced with DNA. The

removal of the primers is performed by the enzyme RNase H, which is ubiquitous in the three domains of life. According to their sequence similarity, in prokaryotes the RNase H proteins are classified into three groups: RNase HI, HII, and HIII. In eukaryotes, they are classified as RNase H1 and H2 [148]. However, phylogenetic analyses suggest that the RNase H proteins can also be classified into two different groups: Type 1 (prokaryotic RNase HI, eukaryotic RNase H1, and viral RNase H) and Type 2 (prokaryotic RNase HII and HIII and eukaryotic RNase H2) [149]. These two types of RNase H possess different specific activities, metal ion preferences, and cleavage sites [150]. Originally it was thought that *Archaea* only possess type 2 RNase H [149], but several type 1 archaeal RNase H enzymes have since then been discovered [151, 152]. Interestingly, in eukaryotes, RNase H1 and H2 tend to coexist, and different combinations of the three prokaryotic RNases H (I, II, and III) are found, except for the combination of RNases HI and HIII. This mutually exclusive evolution of some of the prokaryotic RNases seems to be related to functional redundancy [148].

Another enzyme involved in primer replacement in DNA replication is the Flap endonuclease I (FEN-1), which recognizes double-stranded DNA with an unannealed 5'-flap, and cleaves it. A eukaryotic homolog of FEN-1 is found in every archaeal member analyzed, and one member of the *Thermoproteales*, *Thermofilum pendens* Hrk5, possesses two homologs.

*M. maripaludis* S2 possesses genes for the prokaryotic RNase HI and HII, and for FEN-1, but none are essential, suggesting that they may be redundant in their functions [44, 62]. Possibly, both RNase H proteins persist and are evolutionary stable in the genome. The gene products may perform the same function but one is less efficient than the other [153].

After the Okazaki fragments are replaced by DNA in the lagging strand, the nick between the newly synthesized and the elongated DNA is repaired by a DNA ligase. This enzyme uses a nucleotide cofactor to catalyze the formation of the phosphodiester bond in three well-characterized steps [154]. The enzyme is common to all three domains but can be grouped into two families based on cofactor specificity (ATP or NAD<sup>+</sup>). The DNA ligases from several crenarchaeotes and euryarchaeotes have been further characterized (reviewed in [43]). Many archaeal DNA ligases possess dual cofactor specificity (ATP/NAD<sup>+</sup> or ATP/ADP), but every archaeal DNA ligase characterized so far uses ATP. A thermophilic DNA ligase from *M. thermautotrophicus* which uses ATP as sole cofactor was characterized *in vitro* [155]. A characterized DNA ligase from the crenarchaeote *Sulfophobococcus zilligii* displayed specificity for three cofactors (ATP/NAD<sup>+</sup>/GTP) [156]. Major structural work in this enzyme has been achieved in *P. furiosus* and *S. solfataricus*, and the findings have been reviewed in detail [43]. In *M. maripaludis* S2, the DNA ligase is likely to be essential for growth [44, 62].

Understanding the maturation process in archaea has become a focus of increasing research in recent years. An *in vitro* study reconstituted the Okazaki fragment maturation using proteins derived from the crenarchaeote *S. solfataricus*. They demonstrated that only six proteins are necessary for

coupled DNA synthesis, RNA primer removal, and DNA ligation. In this model a single PCNA (heterotrimeric ring) coordinates the activities of PolBI, FEN-1, and DNA ligase into an efficient maturation complex [157]. A different *in vitro* study in *P. abyssi* has demonstrated two different models for Okazaki fragments maturation, which differ in the RNA primer removal process. In the first model RNA primer elimination is executed by continuous PolD strand displacement DNA synthesis and 5'RNA flap cleavage by FEN-1. The resulting nicks are closed by DNA ligase. Alternatively, the second model proposes that RNase HII cleaves the RNA primer, followed by the action of FEN-1 and strand displacement DNA synthesis by polB or polD. The nick is then sealed by DNA ligase [127]. Testing these models *in vivo* is of great interest to complete the picture of how these processes behave in the cell. In addition, other factors that can modulate these processes may be required for coordination of the *in vivo* maturation process.

#### 4. Conclusions

DNA replication in archaea possesses a dual nature, where the machinery is structurally and functionally similar to the eukaryotic replication system, but it is executed within a bacterial context [11, 25]. Additionally, unique archaeal features demonstrate the complexity of this process, and it does not appear to be just a simplified version of the eukaryotic system. For example, the archaeal specific DNA polymerase D is conserved across the *Archaea* domain with the exception of the *Crenarchaeota* phylum. The recent discovery that it is the essential replicative polymerase in two different euryarchaeotic species demonstrates an unanticipated variability in archaeal DNA replication and a fundamental difference in the replication mechanism between crenarchaeotes and euryarchaeotes. Interestingly, the absence of PolD in crenarchaeotes is not the only difference in DNA replication. These differences include the absence of histones in crenarchaeotes [4], the presence of multiple origins of replication in crenarchaeotes and a single origin of replication in most euryarchaeotes, the absence of GINS23 in most euryarchaeotes, the absence of RPA protein in many crenarchaeotes, the presence of multiple MCM homologs in *Methanococcales*, and the presence of one homolog of PCNA in euryarchaeotes compared to the multiple homologs in crenarchaeotes. These distinctive characteristics between the phyla highlight the complexity of archaeal DNA replication and suggest a complex evolutionary history (Table 1).

Among the *Euryarchaeota*, the orders *Halobacteriales* and *Methanococcales* possess differences in the DNA replication system that make them unique. For instance, *Halobacteriales* possess many more origin recognition proteins (Orcl/Cdc6) compared to the rest of the archaea. On the other hand, *Methanococcales* and *Methanopyrales* lack recognizable homologs for the Orcl/Cdc6 proteins, suggesting the presence of a very different mechanism for initiation of replication. In addition, the *Methanococcales* possess a large number of MCM protein homologs. It has been proposed that these distinctive features are connected, and because of the absence of the Orcl/Cdc6 proteins, MCM proteins may



TABLE 1: DNA replication proteins and features in the domains *Bacteria*, *Eukaryota*, and the two major phyla of the *Archaea* domain. Modified from [43].

DNA replication stage	Process	Archaea		
		Bacteria	Eukaryota	<i>Crenarchaeota</i>
Preinitiation	Origin of replication	Single DnaA	Multiple ORC complex (ORC 1-6)	Multiple OrcI/Cdc6 <sup>b</sup>
	Origin recognition	DnaB	MCM complex (MCM 2-7)	MCM complex
Initiation	DNA unwinding (Helicase)	DnaC	Cdc6 Cdt1	GIN523/GINS15
	DNA unwinding (Accessory proteins)		GIN5 complex (Sld5, Psf1-3) Cdc45	GIN515 <sup>c</sup> RecJ homolog? RecJ homolog?
	Primer synthesis	DnaG	Pol $\alpha$ /primase complex	DNA primase (PriSL) DNA primase (PriSL)/DnaG <sup>d</sup>
Elongation	DNA synthesis (polymerase)	Pol III (Family C DNA polymerase)	Pol $\delta$ , and Pol $\epsilon$ (Family B DNA polymerase)	Family B DNA polymerase polymerase <sup>e</sup>
	DNA synthesis (Processivity factors)	$\gamma$ -complex (clamp loader) $\beta$ -clamp (clamp)	RFC (clamp loader) PCNA (clamp)	RFC (clamp loader) PCNA (clamp)
	Maturation (Okazaki fragment processing)	Pol I (family A DNA polymerase) RNase H DNA ligase	Fen1/Dna2 RNase H DNA ligase	Fen1 RNase H DNA ligase

<sup>a</sup> Exception, the order *Halobacteriales*.

<sup>b</sup> Not known for members of the *Euryarchaeota* orders *Methanococcales* and *Methanopyrales*.

<sup>c</sup> GINS23 has been founded only in the order *Thermococcales* of the *Euryarchaeota*.

<sup>d</sup> *Sulfolobus solfataricus* did show primase activity *in vitro*.

<sup>e</sup> Family B DNA polymerase is also essential in *Halobacterium*. Because its function has not been clearly elucidated, it might also play a role in replication in this and closely related organisms.

interact with other unknown initiation enzymes, resulting in a complex phylogeny of the MCM homologs [49]. Presumably, these differences also account for the inability to recognize the origin of replication in these microorganisms. This unique scenario for initiation of DNA replication may be a direct consequence of various processes such as duplication, mobile genetic elements, and interactions with viruses [58].

The differences in DNA replication among archaeal higher taxa demonstrate an unexpected variability and suggest two alternative evolutionary models. In the first model, DNA replication evolved late after the diversification of the archaeal lineages. Because the replicative system was not fully formed, it was possible to develop differently between the lineages. Once formed, replication would then be highly conserved. However, this model is not consistent with differences observed between lineages that must have formed relatively late, such as those between the orders or even, in some cases, within certain orders. The alternative model is that DNA replication has changed throughout the evolution of the archaea. Thus, differences in the replicative systems may represent ancient as well as modern adaptations to changing environments. In this case, differences in the replicative systems may provide important insights into the evolutionary pressures in play during different episodes of archaeal evolution. From this perspective, the differences between the eukaryotic and archaeal replicative systems may not be evidence for an ancient origin of the eukaryotes. Instead, it is entirely plausible that the replicative system in eukaryotes could have evolved relatively late from a well-developed archaeal system.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] E. F. DeLong and N. R. Pace, "Environmental diversity of bacteria and archaea," *Systematic Biology*, vol. 50, no. 4, pp. 470–478, 2001.
- [2] B. Chaban, S. Y. M. Ng, and K. F. Jarrell, "Archaeal habitats—from the extreme to the ordinary," *Canadian Journal of Microbiology*, vol. 52, no. 2, pp. 73–116, 2006.
- [3] M. J. Hanford and T. L. Peeples, "Archaeal tetraether lipids: unique structures and applications," *Applied Biochemistry and Biotechnology A*, vol. 97, no. 1, pp. 45–62, 2002.
- [4] K. Sandman and J. N. Reeve, "Archaeal histones and the origin of the histone fold," *Current Opinion in Microbiology*, vol. 9, no. 5, pp. 520–525, 2006.
- [5] O. Lecompte, R. Ripp, J.-C. Thierry, D. Moras, and O. Poch, "Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale," *Nucleic Acids Research*, vol. 30, no. 24, pp. 5382–5390, 2002.
- [6] D. R. Edgell and W. F. Doolittle, "Archaea and the origin(s) of DNA replication proteins," *Cell*, vol. 89, no. 7, pp. 995–998, 1997.
- [7] E. R. Barry and S. D. Bell, "DNA replication in the archaea," *Microbiology and Molecular Biology Reviews*, vol. 70, no. 4, pp. 876–887, 2006.
- [8] S. Gribaldo and C. Brochier-Armanet, "The origin and evolution of archaea: a state of the art," *Philosophical Transactions of the Royal Society B*, vol. 361, no. 1470, pp. 1007–1022, 2006.
- [9] P. Londei, "Evolution of translational initiation: new insights from the archaea," *FEMS Microbiology Reviews*, vol. 29, no. 2, pp. 185–200, 2005.
- [10] B. Grabowski and Z. Kelman, "Archaeal DNA replication: eukaryal proteins in a bacterial context," *Annual Review of Microbiology*, vol. 57, pp. 487–516, 2003.
- [11] P. Forterre, J. Filée, and H. Myllykallio, "Origin and evolution of DNA and DNA replication machineries," in *The Genetic Code and the Origin of Life*, L. Ribas, Ed., Landes Bioscience, Georgetown, Tex, USA, 2007.
- [12] E. V. Koonin, "The origin and early evolution of eukaryotes in the light of phylogenomics," *Genome Biology*, vol. 11, no. 5, article 209, 2010.
- [13] M. Könneke, A. E. Bernhard, J. R. de la Torre, C. B. Walker, J. B. Waterbury, and D. A. Stahl, "Isolation of an autotrophic ammonia-oxidizing marine archaeon," *Nature*, vol. 437, no. 7058, pp. 543–546, 2005.
- [14] C. Brochier-Armanet, B. Boussau, S. Gribaldo, and P. Forterre, "Mesophilic crenarchaeota: proposal for a third archaeal phylum, the *Thaumarchaeota*," *Nature Reviews Microbiology*, vol. 6, no. 3, pp. 245–252, 2008.
- [15] J. G. Elkins, M. Podar, D. E. Graham et al., "A korarchaeal genome reveals insights into the evolution of the Archaea," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 23, pp. 8102–8107, 2008.
- [16] T. Nunoura, Y. Takaki, J. Kakuta et al., "Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group," *Nucleic Acids Research*, vol. 39, no. 8, pp. 3204–3223, 2011.
- [17] H. Huber, M. J. Hohn, R. Rachel, T. Fuchs, V. C. Wimmer, and K. O. Stetter, "A new phylum of *Archaea* represented by a nanosized hyperthermophilic symbiont," *Nature*, vol. 417, no. 6884, pp. 63–67, 2002.
- [18] C. Brochier, S. Gribaldo, Y. Zivanovic, F. Confalonieri, and P. Forterre, "Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to *Thermococcales*?" *Genome Biology*, vol. 6, no. 5, article R42, 2005.
- [19] S. J. Aves, "DNA replication initiation," *Methods in Molecular Biology*, vol. 521, pp. 3–17, 2009.
- [20] N. P. Robinson, I. Dionne, M. Lundgren, V. L. Marsh, R. Bernander, and S. D. Bell, "Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*," *Cell*, vol. 116, no. 1, pp. 25–38, 2004.
- [21] M. Lundgren, A. Andersson, L. Chen, P. Nilsson, and R. Bernander, "Three replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and asynchronous termination," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 18, pp. 7046–7051, 2004.
- [22] N. P. Robinson and S. D. Bell, "Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 14, pp. 5806–5811, 2007.
- [23] E. A. Pelve, A. C. Lindas, A. Knoppel, A. Mira, and R. Bernander, "Four chromosome replication origins in the archaeon *Pyrobaculum calidifontis*," *Molecular Microbiology*, vol. 85, no. 5, pp. 986–995, 2012.

- [24] A. F. Andersson, E. A. Pelve, S. Lindeberg, M. Lundgren, P. Nilsson, and R. Bernander, "Replication-biased genome organisation in the crenarchaeon *Sulfolobus*," *BMC Genomics*, vol. 11, no. 1, article 454, 2010.
- [25] H. Myllykallio, P. Lopez, P. López-García et al., "Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon," *Science*, vol. 288, no. 5474, pp. 2212–2215, 2000.
- [26] F. Matsunaga, P. Forterre, Y. Ishino, and H. Myllykallio, "In vivo interactions of archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 20, pp. 11152–11157, 2001.
- [27] S. Maisnier-Patin, L. Malandrin, N. K. Birkeland, and R. Bernander, "Chromosome replication patterns in the hyperthermophilic euryarchaea *Archaeoglobus fulgidus* and *Methanocaldococcus (Methanococcus) jannaschii*," *Molecular Microbiology*, vol. 45, no. 5, pp. 1443–1450, 2002.
- [28] R. Zhang and C.-T. Zhang, "Multiple replication origins of the archaeon *Halobacterium* species NRC-1," *Biochemical and Biophysical Research Communications*, vol. 302, no. 4, pp. 728–734, 2003.
- [29] C. Norais, M. Hawkins, A. L. Hartman, J. A. Eisen, H. Myllykallio, and T. Allers, "Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*," *PLoS Genetics*, vol. 3, no. 5, article e77, 2007.
- [30] J. A. Coker, P. DasSarma, M. Capes et al., "Multiple replication origins of *Halobacterium* sp. strain NRC-1: properties of the conserved *orc7*-dependent *oriC1*," *Journal of Bacteriology*, vol. 191, no. 16, pp. 5253–5261, 2009.
- [31] Z. Wu, J. Liu, H. Yang, H. Liu, and H. Xiang, "Multiple replication origins with diverse control mechanisms in *Haloarcula hispanica*," *Nucleic Acids Research*, 2013.
- [32] M. Hawkins, S. Malla, M. J. Blythe, C. A. Nieduszynski, and T. Allers, "Accelerated growth in the absence of DNA replication origins," *Nature*, vol. 503, no. 7477, pp. 544–547, 2013.
- [33] A. I. Majernik and J. P. J. Chong, "A conserved mechanism for replication origin recognition and binding in archaea," *The Biochemical Journal*, vol. 409, no. 2, pp. 511–518, 2008.
- [34] J. E. Galagan, C. Nusbaum, A. Roy et al., "The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity," *Genome Research*, vol. 12, no. 4, pp. 532–542, 2002.
- [35] R. Zhang and C.-T. Zhang, "Identification of replication origins in archaeal genomes based on the Z-curve method," *Archaea*, vol. 1, no. 5, pp. 335–346, 2005.
- [36] A. Spang, R. Hatzepichler, C. Brochier-Armanet et al., "Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota," *Trends in Microbiology*, vol. 18, no. 8, pp. 331–340, 2010.
- [37] K. J. Marians, "Prokaryotic DNA replication," *Annual Review of Biochemistry*, vol. 61, pp. 673–719, 1992.
- [38] S. P. Bell, "The origin recognition complex: from simple origins to complex functions," *Genes and Development*, vol. 16, no. 6, pp. 659–672, 2002.
- [39] V. L. Marsh and S. D. Bell, "DNA replication and the cell cycle," in *Archaea: Evolution, Physiology, and Molecular Biology*, R. A. Garrett and H. P. Klenk, Eds., Blackwell Publishing, Hoboken, NJ, USA, 2007.
- [40] F. Matsunaga, A. Glatigny, M.-H. Mucchielli-Giorgi et al., "Genomewide and biochemical analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in *Pyrococcus* sp," *Nucleic Acids Research*, vol. 35, no. 10, pp. 3214–3222, 2007.
- [41] J. Liu, C. L. Smith, D. DeRyckere, K. DeAngelis, G. S. Martin, and J. M. Berger, "Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control," *Molecular Cell*, vol. 6, no. 3, pp. 637–648, 2000.
- [42] M. R. Singleton, R. Morales, I. Grainge, N. Cook, M. N. Isupov, and D. B. Wigley, "Conformational changes induced by nucleotide binding in Cdc6/ORC from *Aeropyrum pernix*," *Journal of Molecular Biology*, vol. 343, no. 3, pp. 547–557, 2004.
- [43] Y. Ishino and S. Ishino, "Rapid progress of DNA replication studies in Archaea, the third domain of life," *Science China Life Sciences*, vol. 55, no. 5, pp. 386–403, 2012.
- [44] E. L. Hendrickson, R. Kaul, Y. Zhou et al., "Complete genome sequence of the genetically tractable hydrogenotrophic methanogen *Methanococcus maripaludis*," *Journal of Bacteriology*, vol. 186, no. 20, pp. 6956–6969, 2004.
- [45] B. K. Tye, "MCM proteins in DNA replication," *Annual Review of Biochemistry*, vol. 68, pp. 649–686, 1999.
- [46] K. Labib and J. F. X. Diffley, "Is the MCM2-7 complex the eukaryotic DNA replication fork helicase?" *Current Opinion in Genetics & Development*, vol. 11, no. 1, pp. 64–70, 2001.
- [47] I. Ilves, T. Petojevic, J. J. Pesavento, and M. R. Botchan, "Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins," *Molecular Cell*, vol. 37, no. 2, pp. 247–258, 2010.
- [48] M. L. Bochman and A. Schwacha, "The Mcm complex: unwinding the mechanism of a replicative helicase," *Microbiology and Molecular Biology Reviews*, vol. 73, no. 4, pp. 652–683, 2009.
- [49] Y. Ishimi, "A DNA helicase activity is associated with an MCM4, -6, and -7 protein complex," *The Journal of Biological Chemistry*, vol. 272, no. 39, pp. 24508–24513, 1997.
- [50] Z. Kelman and M. F. White, "Archaeal DNA replication and repair," *Current Opinion in Microbiology*, vol. 8, no. 6, pp. 669–676, 2005.
- [51] J. H. LeBowitz and R. McMacken, "The *Escherichia coli* dnaB replication protein is a DNA helicase," *The Journal of Biological Chemistry*, vol. 261, no. 10, pp. 4738–4748, 1986.
- [52] Z. Kelman, J.-K. Lee, and J. Hurwitz, "The single minichromosome maintenance protein of *Methanobacterium thermoautotrophicum* ΔH contains DNA helicase activity," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 26, pp. 14783–14788, 1999.
- [53] D. F. Shechter, C. Y. Ying, and J. Gautier, "The intrinsic DNA helicase activity of *Methanobacterium thermoautotrophicum* ΔH minichromosome maintenance protein," *The Journal of Biological Chemistry*, vol. 275, no. 20, pp. 15049–15059, 2000.
- [54] M. de Felice, L. Esposito, B. Pucci, M. de Falco, M. Rossi, and F. M. Pisani, "A CDC6-like factor from the archaea *Sulfolobus solfataricus* promotes binding of the mini-chromosome maintenance complex to DNA," *The Journal of Biological Chemistry*, vol. 279, no. 41, pp. 43008–43012, 2004.
- [55] R. Kasiviswanathan, J.-H. Shin, and Z. Kelman, "Interactions between the archaeal Cdc6 and MCM proteins modulate their biochemical properties," *Nucleic Acids Research*, vol. 33, no. 15, pp. 4940–4950, 2005.
- [56] J.-H. Shin, B. Grabowski, R. Kasiviswanathan, S. D. Bell, and Z. Kelman, "Regulation of minichromosome maintenance helicase activity by Cdc6," *The Journal of Biological Chemistry*, vol. 278, no. 39, pp. 38059–38067, 2003.
- [57] M. Akita, A. Adachi, K. Takemura, T. Yamagami, F. Matsunaga, and Y. Ishino, "Cdc6/Orc1 from *Pyrococcus furiosus* may act as the origin recognition protein and Mcm helicase recruiter," *Genes to Cells*, vol. 15, no. 5, pp. 537–552, 2010.

- [58] M. Krupović, S. Gribaldo, D. H. Bamford, and P. Forterre, "The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements," *Molecular Biology and Evolution*, vol. 27, no. 12, pp. 2716–2732, 2010.
- [59] A. D. Walters and J. P. J. Chong, "An archaeal order with multiple minichromosome maintenance genes," *Microbiology*, vol. 156, part 5, pp. 1405–1414, 2010.
- [60] A. D. Walters and J. P. J. Chong, "Methanococcus maripaludis: an archaeon with multiple functional MCM proteins?" *Biochemical Society Transactions*, vol. 37, no. 1, pp. 1–6, 2009.
- [61] Q. Xia, E. L. Hendrickson, Y. Zhang et al., "Quantitative proteomics of the archaeon *Methanococcus maripaludis* validated by microarray analysis and real time PCR," *Molecular & Cellular Proteomics*, vol. 5, no. 5, pp. 868–881, 2006.
- [62] F. Sarmiento, J. Mrázek, and W. B. Whitman, "Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon *Methanococcus maripaludis*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 12, pp. 4726–4731, 2013.
- [63] M. Pan, T. J. Santangelo, Z. Li, J. N. Reeve, and Z. Kelman, "Thermococcus kodakarensis encodes three MCM homologs but only one is essential," *Nucleic Acids Research*, vol. 39, no. 22, pp. 9671–9680, 2011.
- [64] S. Ishino, S. Fujino, H. Tomita et al., "Biochemical and genetical analyses of the three mcm genes from the hyperthermophilic archaeon, *Thermococcus kodakarensis*," *Genes to Cells*, vol. 16, no. 12, pp. 1176–1189, 2011.
- [65] K. S. Makarova, Y. I. Wolf, S. L. Mekhedov, B. G. Mirkin, and E. V. Koonin, "Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell," *Nucleic Acids Research*, vol. 33, no. 14, pp. 4626–4638, 2005.
- [66] N. Marinsek, E. R. Barry, K. S. Makarova, I. Dionne, E. V. Koonin, and S. D. Bell, "GINS, a central nexus in the archaeal DNA replication fork," *EMBO Reports*, vol. 7, no. 5, pp. 539–545, 2006.
- [67] T. Yoshimochi, R. Fujikane, M. Kawanami, F. Matsunaga, and Y. Ishino, "The GINS complex from *Pyrococcus furiosus* stimulates the MCM helicase activity," *The Journal of Biological Chemistry*, vol. 283, no. 3, pp. 1601–1609, 2008.
- [68] T. Oyama, S. Ishino, S. Fujino et al., "Architectures of archaeal GINS complexes, essential DNA replication initiation factors," *BMC Biology*, vol. 9, article 28, 2011.
- [69] H. Ogino, S. Ishino, K. Mayanagi et al., "The GINS complex from the thermophilic archaeon, *Thermoplasma acidophilum* may function as a homotetramer in DNA replication," *Extremophiles*, vol. 15, no. 4, pp. 529–539, 2011.
- [70] L. Sanchez-Pulido and C. P. Ponting, "Cdc45: the missing recJ ortholog in eukaryotes?" *Bioinformatics*, vol. 27, no. 14, pp. 1885–1888, 2011.
- [71] Z. Li, T. J. Santangelo, L. Čuboňová, J. N. Reeve, and Z. Kelman, "Affinity purification of an archaeal DNA replication protein network," *mBio*, vol. 1, no. 5, 2010.
- [72] Z. Li, M. Pan, T. J. Santangelo et al., "A novel DNA nuclease is stimulated by association with the GINS complex," *Nucleic Acids Research*, vol. 39, no. 14, pp. 6114–6123, 2011.
- [73] L. A. Rajman and S. T. Lovett, "A thermostable single-strand DNase from *Methanococcus jannaschii* related to the RecJ recombination and repair exonuclease from *Escherichia coli*," *Journal of Bacteriology*, vol. 182, no. 3, pp. 607–612, 2000.
- [74] K. R. Williams, J. B. Murphy, and J. W. Chase, "Characterization of the structural and functional defect in the *Escherichia coli* single-stranded DNA binding protein encoded by the *ssb-1* mutant gene. Expression of the *ssb-1* gene under  $\lambda$ p(L) regulation," *The Journal of Biological Chemistry*, vol. 259, no. 19, pp. 11804–11811, 1984.
- [75] S. Dąbrowski, M. Olszewski, R. Piątek, and J. Kur, "Novel thermostable ssDNA-binding proteins from *Thermus thermophilus* and *T. aquaticus*-expression and purification," *Protein Expression and Purification*, vol. 26, no. 1, pp. 131–138, 2002.
- [76] C. Iftode, Y. Daniely, and J. A. Borowiec, "Replication protein A (RPA): the eukaryotic SSB," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 34, no. 3, pp. 141–180, 1999.
- [77] A. G. Murzin, "OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences," *The EMBO Journal*, vol. 12, no. 3, pp. 861–867, 1993.
- [78] C. A. Haseltine and S. C. Kowalczykowski, "A distinctive single-stranded DNA-binding protein from the Archaeon *Sulfolobus solfataricus*," *Molecular Microbiology*, vol. 43, no. 6, pp. 1505–1515, 2002.
- [79] R. I. M. Wadsworth and M. F. White, "Identification and properties of the crenarchaeal single-stranded DNA binding protein from *Sulfolobus solfataricus*," *Nucleic Acids Research*, vol. 29, no. 4, pp. 914–920, 2001.
- [80] I. D. Kerr, R. I. M. Wadsworth, L. Cubeddu, W. Blankenfeldt, J. H. Naismith, and M. F. White, "Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of *Sulfolobus* SSB protein," *The EMBO Journal*, vol. 22, no. 11, pp. 2561–2570, 2003.
- [81] X. Luo, U. Schwarz-Linek, C. H. Botting, R. Hensel, B. Siebers, and M. F. White, "CC1, a novel crenarchaeal DNA binding protein," *Journal of Bacteriology*, vol. 189, no. 2, pp. 403–409, 2007.
- [82] T. J. Kelly, P. Simacek, and G. S. Brush, "Identification and characterization of a single-stranded DNA-binding protein from the archaeon *Methanococcus jannaschii*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14634–14639, 1998.
- [83] Z. Kelman, S. Pietrokovski, and J. Hurwitz, "Isolation and characterization of a split B-type DNA polymerase from the archaeon *Methanobacterium thermoautotrophicum*  $\Delta$ H," *The Journal of Biological Chemistry*, vol. 274, no. 40, pp. 28751–28761, 1999.
- [84] K. Komori and Y. Ishino, "Replication protein A in *Pyrococcus furiosus* is involved in homologous DNA recombination," *The Journal of Biological Chemistry*, vol. 276, no. 28, pp. 25654–25660, 2001.
- [85] J. B. Robbins, M. C. Murphy, B. A. White, R. I. Mackie, T. Ha, and I. K. O. Cann, "Functional analysis of multiple single-stranded DNA-binding proteins from *Methanosarcina acetivorans* and their effects on DNA synthesis by DNA polymerase B1," *The Journal of Biological Chemistry*, vol. 279, no. 8, pp. 6315–6326, 2004.
- [86] Y. Lin, L.-J. Lin, P. Sriratana et al., "Engineering of functional replication protein a homologs based on insights into the evolution of oligonucleotide/oligosaccharide-binding folds," *Journal of Bacteriology*, vol. 190, no. 17, pp. 5766–5780, 2008.
- [87] D. N. Frick and C. C. Richardson, "DNA primases," *Annual Review of Biochemistry*, vol. 70, pp. 39–80, 2001.
- [88] G. Desogus, S. Onesti, P. Brick, M. Rossi, and F. M. Pisani, "Identification and characterization of a DNA primase from the hyperthermophilic archaeon *Methanococcus jannaschii*," *Nucleic Acids Research*, vol. 27, no. 22, pp. 4444–4450, 1999.

- [89] A. A. Bocquier, L. Liu, I. K. O. Cann, K. Komori, D. Kohda, and Y. Ishino, "Archaeal primase: bridging the gap between RNA and DNA polymerases," *Current Biology*, vol. 11, no. 6, pp. 452–456, 2001.
- [90] L. Liu, K. Komori, S. Ishino et al., "The archaeal DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*," *The Journal of Biological Chemistry*, vol. 276, no. 48, pp. 45484–45490, 2001.
- [91] E. Matsui, M. Nishio, H. Yokoyama, K. Harata, S. Darnis, and I. Matsui, "Distinct domain functions regulating de novo DNA synthesis of thermostable DNA primase from hyperthermophile *Pyrococcus horikoshii*," *Biochemistry*, vol. 42, no. 50, pp. 14968–14976, 2003.
- [92] N. Ito, I. Matsui, and E. Matsui, "Molecular basis for the subunit assembly of the primase from an archaeon *Pyrococcus horikoshii*," *The FEBS Journal*, vol. 274, no. 5, pp. 1340–1351, 2007.
- [93] M. le Breton, G. Henneke, C. Norais et al., "The heterodimeric primase from the euryarchaeon *Pyrococcus abyssi*: a multifunctional enzyme for initiation and repair?" *Journal of Molecular Biology*, vol. 374, no. 5, pp. 1172–1185, 2007.
- [94] S.-H. Lao-Sirieix and S. D. Bell, "The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities," *Journal of Molecular Biology*, vol. 344, no. 5, pp. 1251–1263, 2004.
- [95] S.-H. Lao-Sirieix, L. Pellegrini, and S. D. Bell, "The promiscuous primase," *Trends in Genetics*, vol. 21, no. 10, pp. 568–572, 2005.
- [96] J. Hu, L. Guo, K. Wu, B. Liu, S. Lang, and L. Huang, "Template-dependent polymerization across discontinuous templates by the heterodimeric primase from the hyperthermophilic archaeon *Sulfolobus solfataricus*," *Nucleic Acids Research*, vol. 40, no. 8, pp. 3470–3483, 2012.
- [97] E. Evgueniev-Hackenberg, P. Walter, E. Hochleitner, F. Lottspeich, and G. Klug, "An exosome-like complex in *Sulfolobus solfataricus*," *EMBO Reports*, vol. 4, no. 9, pp. 889–893, 2003.
- [98] P. Walter, F. Klein, E. Lorentzen, A. Ilchmann, G. Klug, and E. Evgueniev-Hackenberg, "Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon *Sulfolobus solfataricus*," *Molecular Microbiology*, vol. 62, no. 4, pp. 1076–1089, 2006.
- [99] Z. Zuo, C. J. Rodgers, A. L. Mikheikin, and M. A. Trakselis, "Characterization of a functional DnaG-type primase in archaea: implications for a dual-primase system," *Journal of Molecular Biology*, vol. 397, no. 3, pp. 664–676, 2010.
- [100] B. R. Berquist, P. DasSarma, and S. DasSarma, "Essential and non-essential DNA replication genes in the model halophilic Archaeon, *Halobacterium* sp. NRC-1," *BMC Genetics*, vol. 8, article 31, 2007.
- [101] D. K. Braithwaite and J. Ito, "Compilation, alignment, and phylogenetic relationships of DNA polymerases," *Nucleic Acids Research*, vol. 21, no. 4, pp. 787–802, 1993.
- [102] I. K. O. Cann and Y. Ishino, "Archaeal DNA replication: identifying the pieces to solve a puzzle," *Genetics*, vol. 152, no. 4, pp. 1249–1267, 1999.
- [103] G. Lipps, S. Röther, C. Hart, and G. Krauss, "A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity," *The EMBO Journal*, vol. 22, no. 10, pp. 2516–2525, 2003.
- [104] J. Yamtich and J. B. Sweasy, "DNA polymerase family X: function, structure, and cellular roles," *Biochimica et Biophysica Acta*, vol. 1804, no. 5, pp. 1136–1150, 2010.
- [105] H. Ohmori, E. C. Friedberg, R. P. P. Fuchs et al., "The Y-family of DNA polymerases," *Molecular Cell*, vol. 8, no. 1, pp. 7–8, 2001.
- [106] N. Y. Yao, R. E. Georgescu, J. Finkelstein, and M. E. O'Donnell, "Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 32, pp. 13236–13241, 2009.
- [107] U. Hübscher, G. Maga, and S. Spadari, "Eukaryotic DNA polymerases," *Annual Review of Biochemistry*, vol. 71, pp. 133–163, 2002.
- [108] F. Boudsocq, S. Iwai, F. Hanaoka, and R. Woodgate, "*Sulfolobus solfataricus* P2 DNA polymerase IV (Dp04): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol $\eta$ ," *Nucleic Acids Research*, vol. 29, no. 22, pp. 4607–4616, 2001.
- [109] L.-J. Lin, A. Yoshinaga, Y. Lin et al., "Molecular analyses of an unusual translesion DNA polymerase from *Methanosarcina acetivorans* C2A," *Journal of Molecular Biology*, vol. 397, no. 1, pp. 13–30, 2010.
- [110] I. K. O. Cann, S. Ishino, N. Nomura, Y. Sako, and Y. Ishino, "Two family B DNA polymerases from *Aeropyrum pernix*, an aerobic hyperthermophilic crenarchaeote," *Journal of Bacteriology*, vol. 181, no. 19, pp. 5984–5992, 1999.
- [111] T. Uemori, Y. Ishino, H. Doi, and I. Kato, "The hyperthermophilic archaeon *Pyrodictium occultum* has two  $\alpha$ -like DNA polymerases," *Journal of Bacteriology*, vol. 177, no. 8, pp. 2164–2177, 1995.
- [112] T. Uemori, Y. Ishino, H. Toh, K. Asada, and I. Kato, "Organization and nucleotide sequence of the DNA polymerase gene from the archaeon *Pyrococcus furiosus*," *Nucleic Acids Research*, vol. 21, no. 2, pp. 259–265, 1993.
- [113] B. A. Connolly, "Recognition of deaminated bases by archaeal family-B DNA polymerases," *Biochemical Society Transactions*, vol. 37, no. 1, pp. 65–68, 2009.
- [114] B. A. Connolly, M. J. Fogg, G. Shuttleworth, and B. T. Wilson, "Uracil recognition by archaeal family B DNA polymerases," *Biochemical Society Transactions*, vol. 31, no. 3, pp. 699–702, 2003.
- [115] T. T. Richardson, X. Wu, B. J. Keith, P. Heslop, A. C. Jones, and B. A. Connolly, "Unwinding of primer-templates by archaeal family-B DNA polymerases in response to template-strand uracil," *Nucleic Acids Research*, vol. 41, no. 4, pp. 2466–2478, 2013.
- [116] T. Lindahl, "Instability and decay of the primary structure of DNA," *Nature*, vol. 362, no. 6422, pp. 709–715, 1993.
- [117] I. B. Rogozin, K. S. Makarova, Y. I. Pavlov, and E. V. Koonin, "A highly conserved family of inactivated archaeal B family DNA polymerases," *Biology Direct*, vol. 3, article 32, 2008.
- [118] M. Imamura, T. Uemori, I. Kato, and Y. Ishino, "A non- $\alpha$ -like DNA polymerase from the hyperthermophilic archaeon *Pyrococcus furiosus*," *Biological and Pharmaceutical Bulletin*, vol. 18, no. 12, pp. 1647–1652, 1995.
- [119] I. K. O. Cann, K. Komori, H. Toh, S. Kanai, and Y. Ishino, "A heterodimeric DNA polymerase: evidence that members of *Euryarchaeota* possess a distinct DNA polymerase," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 24, pp. 14250–14255, 1998.
- [120] L. Aravind, D. D. Leipe, and E. V. Koonin, "Toprim—a conserved catalytic domain in type IA and II topoisomerases,

- DnaG-type primases, OLD family nucleases and RecR proteins," *Nucleic Acids Research*, vol. 26, no. 18, pp. 4205–4213, 1998.
- [121] M. Jokela, A. Eskelinen, H. Pospiech, J. Rouvinen, and J. E. Syväoja, "Characterization of the 3' exonuclease subunit DPI of *Methanococcus jannaschii* replicative DNA polymerase D," *Nucleic Acids Research*, vol. 32, no. 8, pp. 2430–2440, 2004.
- [122] T. Uemori, Y. Sato, I. Kato, H. Doi, and Y. Ishino, "A novel DNA polymerase in the hyperthermophilic archaeon, *Pyrococcus furiosus*: gene cloning, expression, and characterization," *Genes to Cells*, vol. 2, no. 8, pp. 499–512, 1997.
- [123] Y. Shen, K. Musti, M. Hiramoto, H. Kikuchi, Y. Kawarabayashi, and I. Matsui, "Invariant Asp-1122 and Asp-1124 are essential residues for polymerization catalysis of family D DNA polymerase from *Pyrococcus horikoshii*," *The Journal of Biological Chemistry*, vol. 276, no. 29, pp. 27376–27383, 2001.
- [124] G. Henneke, D. Flament, U. Hübscher, J. Querellou, and J.-P. Raffin, "The hyperthermophilic euryarchaeota *Pyrococcus abyssi* likely requires the two DNA polymerases D and B for DNA replication," *Journal of Molecular Biology*, vol. 350, no. 1, pp. 53–64, 2005.
- [125] L. Cubonova, T. Richardson, B. W. Burkhart et al., "Archaeal DNA polymerase D but not DNA polymerase B is required for genome replication in *Thermococcus kodakarensis*," *Journal of Bacteriology*, vol. 195, no. 10, pp. 2322–2328, 2013.
- [126] C. Rouillon, G. Henneke, D. Flament, J. Querellou, and J.-P. Raffin, "DNA polymerase switching on homotrimeric PCNA at the replication fork of the euryarchaea *Pyrococcus abyssi*," *Journal of Molecular Biology*, vol. 369, no. 2, pp. 343–355, 2007.
- [127] G. Henneke, "In vitro reconstitution of RNA primer removal in Archaea reveals the existence of two pathways," *The Biochemical Journal*, vol. 447, no. 2, pp. 271–280, 2012.
- [128] G.-L. Moldovan, B. Pfander, and S. Jentsch, "PCNA, the maestro of the replication fork," *Cell*, vol. 129, no. 4, pp. 665–679, 2007.
- [129] X.-P. Kong, R. Onrust, M. O'Donnell, and J. Kuriyan, "Three-dimensional structure of the  $\beta$  subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp," *Cell*, vol. 69, no. 3, pp. 425–437, 1992.
- [130] K. Daimon, Y. Kawarabayashi, H. Kikuchi, Y. Sako, and Y. Ishino, "Three proliferating cell nuclear antigen-like proteins found in the hyperthermophilic archaeon *Aeropyrum pemix*: interactions with the two DNA polymerases," *Journal of Bacteriology*, vol. 184, no. 3, pp. 687–694, 2002.
- [131] I. Dionne, R. K. Nookala, S. P. Jackson, A. J. Doherty, and S. D. Bell, "A heterotrimeric PCNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*," *Molecular Cell*, vol. 11, no. 1, pp. 275–282, 2003.
- [132] E. Warbrick, "PCNA binding through a conserved motif," *BioEssays*, vol. 20, no. 3, pp. 195–199, 1998.
- [133] H. Nishida, K. Mayanagi, S. Kiyonari et al., "Structural determinant for switching between the polymerase and exonuclease modes in the PCNA-replicative DNA polymerase complex," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 49, pp. 20693–20698, 2009.
- [134] K. Mayanagi, S. Kiyonari, H. Nishida et al., "Architecture of the DNA polymerase B-proliferating cell nuclear antigen (PCNA)-DNA ternary complex," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 5, pp. 1845–1849, 2011.
- [135] B. Castrec, C. Rouillon, G. Henneke, D. Flament, J. Querellou, and J.-P. Raffin, "Binding to PCNA in Euryarchaeal DNA replication requires two PIP motifs for DNA polymerase D and one PIP motif for DNA polymerase B," *Journal of Molecular Biology*, vol. 394, no. 2, pp. 209–218, 2009.
- [136] P. F. Pluchon, T. Fouqueau, C. Creze et al., "An extended network of genomic maintenance in the archaeon *Pyrococcus abyssi* highlights unexpected associations between eucaryotic homologs," *PLoS One*, vol. 8, no. 11, Article ID e79707, 2013.
- [137] J. E. Ladner, M. Pan, J. Hurwitz, and Z. Kelman, "Crystal structures of two active proliferating cell nuclear antigens (PCNAs) encoded by *Thermococcus kodakarensis*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 7, pp. 2711–2716, 2011.
- [138] Y. Kuba, S. Ishino, T. Yamagami et al., "Comparative analyses of the two proliferating cell nuclear antigens from the hyperthermophilic archaeon, *Thermococcus kodakarensis*," *Genes to Cells*, vol. 17, no. 11, pp. 923–937, 2012.
- [139] T. Fukui, H. Atomi, T. Kanai, R. Matsumi, S. Fujiwara, and T. Imanaka, "Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakarensis* KOD1 and comparison with *Pyrococcus* genomes," *Genome Research*, vol. 15, no. 3, pp. 352–363, 2005.
- [140] D. Jeruzalmi, O. Yurieva, Y. Zhao et al., "Mechanism of processivity clamp opening by the delta subunit wrench of the clamp loader complex of *E. coli* DNA polymerase III," *Cell*, vol. 106, no. 4, pp. 417–428, 2001.
- [141] Y. Shiomi, J. Usukura, Y. Masamura et al., "ATP-dependent structural change of the eukaryotic clamp-loader protein, replication factor C," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 26, pp. 14127–14132, 2000.
- [142] T. Miyata, T. Oyama, K. Mayanagi, S. Ishino, Y. Ishino, and K. Morikawa, "The clamp-loading complex for processive DNA replication," *Nature Structural and Molecular Biology*, vol. 11, no. 7, pp. 632–636, 2004.
- [143] T. Oyama, Y. Ishino, I. K. O. Cann, S. Ishino, and K. Morikawa, "Atomic structure of the clamp loader small subunit from *Pyrococcus furiosus*," *Molecular Cell*, vol. 8, no. 2, pp. 455–463, 2001.
- [144] A. Seybert, D. J. Scott, S. Scaife, M. R. Singleton, and D. B. Wigley, "Biochemical characterisation of the clamp/clamp loader proteins from the euryarchaeon *Archaeoglobus fulgidus*," *Nucleic Acids Research*, vol. 30, no. 20, pp. 4329–4338, 2002.
- [145] Y.-H. Chen, Y. Lin, A. Yoshinaga et al., "Molecular analyses of a three-subunit euryarchaeal clamp loader complex from *Methanosarcina acetivorans*," *Journal of Bacteriology*, vol. 191, no. 21, pp. 6539–6549, 2009.
- [146] Y.-H. Chen, S. A. Kocherginskaya, Y. Lin et al., "Biochemical and mutational analyses of a unique clamp loader complex in the archaeon *Methanosarcina acetivorans*," *The Journal of Biological Chemistry*, vol. 280, no. 51, pp. 41852–41863, 2005.
- [147] F. Matsunaga, C. Norais, P. Forterre, and H. Myllykallio, "Identification of short "eukaryotic" Okazaki fragments synthesized from a prokaryotic replication origin," *EMBO Reports*, vol. 4, no. 2, pp. 154–158, 2003.
- [148] H. Kochiwa, M. Tomita, and A. Kanai, "Evolution of ribonuclease H genes in prokaryotes to avoid inheritance of redundant genes," *BMC Evolutionary Biology*, vol. 7, article 128, 2007.
- [149] N. Ohtani, M. Haruki, M. Morikawa, and S. Kanaya, "Molecular diversities of RNases H," *Journal of Bioscience and Bioengineering*, vol. 88, no. 1, pp. 12–19, 1999.

- [150] N. Ohtani, M. Haruki, M. Morikawa, R. J. Crouch, M. Itaya, and S. Kanaya, "Identification of the genes encoding  $Mn^{2+}$ -dependent RNase HII and  $Mg^{2+}$ -dependent RNase HIII from *Bacillus subtilis*: classification of RNases H into three families," *Biochemistry*, vol. 38, no. 2, pp. 605–618, 1999.
- [151] N. Ohtani, H. Yanagawa, M. Tomita, and M. Itaya, "Identification of the first archaeal type 1 RNase H gene from *Halobacterium* sp. NRC-1: archaeal RNase HI can cleave an RNA-DNA junction," *The Biochemical Journal*, vol. 381, no. 3, pp. 795–802, 2004.
- [152] N. Ohtani, H. Yanagawa, M. Tomita, and M. Itaya, "Cleavage of double-stranded RNA by RNase HI from a thermoacidophilic archaeon, *Sulfolobus tokodaii* 7," *Nucleic Acids Research*, vol. 32, no. 19, pp. 5809–5819, 2004.
- [153] M. A. Nowak, M. C. Boerlijst, J. Cooke, and J. M. Smith, "Evolution of genetic redundancy," *Nature*, vol. 388, no. 6638, pp. 167–170, 1997.
- [154] A. E. Tomkinson, S. Vijayakumar, J. M. Pascal, and T. Ellenberger, "DNA ligases: structure, reaction mechanism, and function," *Chemical Reviews*, vol. 106, no. 2, pp. 687–699, 2006.
- [155] M. Nakatani, S. Ezaki, H. Atomi, and T. Imanaka, "A DNA ligase from a hyperthermophilic archaeon with unique cofactor specificity," *Journal of Bacteriology*, vol. 182, no. 22, pp. 6424–6433, 2000.
- [156] Y. Sun, M. S. Seo, J. H. Kim et al., "Novel DNA ligase with broad nucleotide cofactor specificity from the hyperthermophilic crenarchaeon *Sulfolobus solfataricus*: influence of ancestral DNA ligase on cofactor utilization," *Environmental Microbiology*, vol. 10, no. 12, pp. 3212–3224, 2008.
- [157] T. R. Beattie and S. D. Bell, "Coordination of multiple enzyme activities by a single PCNA in archaeal Okazaki fragment maturation," *The EMBO Journal*, vol. 31, no. 6, pp. 1556–1567, 2012.