World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

Identification of Polymerase and Proofreading Exonuclease Domains in the DNA Polymerases IA, IB and Nuclear-Encoded RNA Polymerase of the Plant Chloroplasts

Peramachi Palanivelu *

Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai – 625 021, India.

World Journal of Advanced Research and Reviews, 2023, 17(03), 706–727

Publication history: Received on 09 February 2023; revised on 14 March 2023; accepted on 18 March 2023

Article DOI: https://doi.org/10.30574/wjarr.2023.17.3.0455

Abstract

Chloroplast plays a crucial role in all photosynthetic plants and converts the light energy to chemical energy. It is a semi-autonomous organelle and is mostly controlled by its own genome and partly by the nuclear imports. To replicate its own genome, it uses two DNA polymerases, viz. polymerases IA and IB. DNA polymerase IA showed 72.45% identity to polymerase IB, but only 35.35% identity to the *E. coli* DNA polymerase I. Multiple sequence alignment (MSA) analysis have shown that the DNA polymerases IA and IB and the *E. coli* DNA polymerase I possess almost identical active sites for polymerization and proofreading (PR) functions, suggesting their possible common evolutionary origin. The nuclear-encoded RNA polymerase (NEP) is imported from the nucleus and involves in the transcription of all the four subunits of the chloroplast RNA polymerase. The polymerase catalytic core of the DNA polymerases IA, IB and the NEP are remarkably conserved and is in close agreement with other DNA/RNA polymerases reported already, and possess a typical template-binding pair (-YG-), a basic catalytic amino acid (K) to initiate catalysis and a basic nucleotide selection amino acid R at -4 from K. The DNA polymerases IA and IB are very similar to prokaryotic DNA polymerases, except in possessing a zinc-binding motif (ZBM) in them, like the eukaryotic replicases. Interestingly, the PR exonucleases of all three polymerases belong to the DEDD-superfamily of exonucleases. The DNA polymerases IA and IB belong to the DEDD(Y)-subfamily, whereas the NEP belongs to the DEDD(H)-subfamily.

Keywords: Chloroplast replication; DNA polymerase IA; DNA polymerase IB; Polymerase active sites; Chloroplast transcription; Nuclear-encoded RNA polymerase; Proofreading exonucleases; Exonuclease active sites

1. Introduction

DNA and RNA polymerases are crucial enzymes of life and, therefore, they are found in all living cells. They play an important role in not only maintaining the blue-print of life in all living cells but also in the control of gene expression at the transcription level. Therefore, understanding the structure, function, mechanism and regulation of these polymerases has been the primary goal of molecular biologists since its discovery. Interestingly, these polymerases are highly conserved in all kingdoms of life, from viruses to animals. [1, 2]. DNA polymerases essentially involve in the replication and repair of genomes in both prokaryotes and eukaryotes whereas RNA polymerases involve in transcription of genomes, i.e., the flow of genetic information from DNA to RNA. The organelles such as mitochondria and chloroplasts also use these enzymes for replication and repair, and transcription processes. Prokaryotes use a single type of DNA polymerase for replication, viz. DNA pol III, whereas the replication in eukaryotes is a complex process and they use three types of DNA polymerases for genome replication and repair, viz. DNA pol α, pol ε and pol δ [3]. Organelles such as mitochondria and chloroplasts use mostly a single type of replicase, viz. DNA pol γ and DNA pols IA and IB, respectively.

* Corresponding author: Peramachi Palanivelu

Transcriptions are performed by different types of RNA polymerases in prokaryotes and eukaryotes. Viruses contain two types of RNA polymerases, viz. DNA-dependent RNAPs (DdRps) and RNA-dependent RNAPs (RdRps), which are of single-subunit (SSU) types. Eubacteria and archaebacteria employ a single type of RNA polymerase for all their transcription needs and it is a multi-subunit (MSU) type. However, eukaryotes use 5 different types of RNA polymerases (I-V), which are also of MSU type [4]. The chloroplast transcription in higher plants is performed by two types of RNA polymerases, i) plastid-encoded RNA polymerase (PEP), and ii) nuclear-encoded RNA polymerase (NEP). PEP is essentially a eubacterial-type MSU enzyme whose catalytic core subunits are encoded by the chloroplast genome itself, whereas NEP is a SSU enzyme which is of T7 bacteriophage-type. PEP is crucial for the biogenesis and maintenance of chloroplasts, but is controlled by the NEP and nuclear-encoded sigma factors (Figs. 1A and 1B) [5].

Both the DNA and RNA polymerases do make mistakes during the replication/transcription processes, but rarely. However, the mistakes are corrected promptly by the PR enzymes in both the cases. Based on the active site structures, the PR enzymes are classified into two major groups, viz. DEDD- and PHP– superfamilies. The DEDD-superfamily is the most common and there are mainly two types of DEDD-superfamily of PR exonucleases found in all biological systems which repair any error(s) that might occur during replication and transcription processes: One type of DEDD-superfamily of PR exonuclease (intrinsic type) is found in the same DNA/RNA polymerase polypeptide itself and functions as a multifunctional enzyme (e. g.), *E. coli* DNA pol I. The second type of PR exonuclease (extrinsic type) exists as a tightly associated subunit along with the polymerase subunits, in a multienzyme complex system (e.g.), bacterial DNA pols III, ExoNs of RdRps the SARS-Coronaviruses and the PA subunits of the RNA polymerases in human influenza viruses [1, 6, 7]. The DEDD-superfamily of PR exonuclease, the most common type of PR enzyme, consists of two subfamilies, viz. DEDD(Y) and DEDD(H), depending upon whether they employ an invariant Y or a H as the proton acceptor to initiate catalysis [8].

These two superfamilies are invariably found/associated with the DNA/RNA replicases and transcriptases to repair any error that might occur during the replication/transcription processes [8, 9]. The PHP-superfamily is not that common and has been reported mainly from the bacterial kingdom, but recently from the viral kingdom also by Palanivelu (7). For example, the PHP-superfamily is commonly found in the bacterial replicative DNA polymerases III (DNA pols III belong to the C-family polymerases), and in bacterial DNA polymerases X. (It is interesting to note, that the A- and B-families of the replicative polymerases are found both in prokaryotes and eukaryotes, but the C-family is found only in bacterial kingdom. All the three families are involved in genome replication and repair. Whereas the A–family polymerase possesses two exonuclease domains (3→5' and 5'→3' exonucleases), but the B-family polymerase possesses only one exonuclease domain, i.e., 3'→5' PR exonuclease domain) [10].

1.1. DNA Polymerases of Chloroplasts and their Roles in the Replication

Chloroplasts, an important organelle of all photosynthetic plants, house the photosynthetic enzyme systems that convert the light energy to chemical energy. Thus, they play a vital role not only in the primary carbon metabolism but also in the biosynthesis of fatty acids, amino acids, and tetrapyrroles. Therefore, chloroplast's function is indispensable throughout the life-cycle of plants and a compromised activity results in embryo lethality. Chloroplasts and mitochondria are semi-autonomous organelles and are considered descendants of endosymbiotic α-proteobacteria-like and cyanobacteria-like organisms, respectively. They possess their own genomes, and replication, transcription and translation machineries. Their genomes exist primarily as homogeneous circular DNA molecules and replicated by its own DNA polymerases. Although the replication mechanisms of these organellar genomes are poorly understood in photosynthetic plants, several enzymes related to its genome replication, such as DNA polymerases IA/IB, primase, replicative DNA helicase, DNA topoisomerase, single-strand DNA-binding proteins, RNase H and DNA ligase have been localized in chloroplasts.

Our understanding of plant organelle DNA replication is still very incomplete [11]. This is largely due to insufficient knowledge about their replication and repair enzymes, and their role(s). To date, two organellar DNA polymerases, IA and IB, resembling bacterial DNA Pol I, have been identified in chloroplasts [3,12]. The DNA polymerase and PR exonuclease active sites of 1A and 1B show similarity to the *E. coli* DNA pol I, but differ from the bacterial DNA polymerase by harbouring a typical zinc-binding motif (ZBM) like eukaryotic replicases. However, both DNA pol IA and IB have been shown to replicate the entire chloroplast genome with a greater efficiency than the microbial DNA pol I.

The importance and involvement of these two polymerases is proved by genetic analysis. The pol IB knockout plants were shown to have fewer genome copy numbers per organelle and grew slowly [13], and the ΔPolIB deletion mutant showed increased sensitivity to double-stranded DNA breaks, suggesting its predominant role in chloroplast DNA damage repair [14]. Parent et al [14] found that mutation of both genes was lethal, and thus, confirming an essential and redundant role for these two proteins. However, the mutation of a single gene was sufficient to cause a reduction

in the levels of DNA in both mitochondria and plastids. They also demonstrated that *polIb*, but not *polIa* mutant lines, are hypersensitive to ciprofloxacin, a small molecule that specifically induces DNA double-strand breaks in plant organelles, suggesting a function for PolIB in DNA repair. MSA analysis shows that pol IA and pol IB are very similar to each other, and possess identical polymerase and PR exonuclease active sites (Figs. 2 and 3). Unlike the *E. coli* DNA pol I, both DNA pols IA and IB are able to bypass DNA lesions and continue replicating DNA [15]. Therefore, it was suggested that a high degree of fidelity could be due to the presence of 3'-5' PR exonuclease domains, which are generally absent in DNA polymerases, involve in typical translesion synthesis.

1.2. RNA Polymerases of Plant Chloroplasts (PEP and NEP)

As discussed elsewhere, the PEP and NEP are the two enzymes that are involved in the transcription of the chloroplast genome and are extensively studied [5, 16]. Even though both the polymerases are high–fidelity enzymes, they can add a few (10^{-4} to 10^{-5}) mismatched nucleotides during the high-speed transcription processes (2). These mismatched nucleotides may have a null-effect on the growth and survival of an organism or they maybe deleterious, depending upon its location(s) and essentiality of the translated protein(s). Therefore, in order to overcome this problem, it is found that these polymerases do have an intrinsic PR exonuclease, which corrects these mistakes, resulting in error-free transcriptions.

1.3. Roles of PEP and NEP in Plant Chloroplasts

As discussed elsewhere, the chloroplast in higher plants use two different types of RNA polymerases to transcribe all its genes. One is the NEP, which is homologous to the SSU-RNAPs of bacteriophages and mitochondria [2, 17-19] and the other one is the PEP, which is structurally and functionally very similar to the eubacterial MSU-RNAPs. It is interesting to note that the NEP, encoded by the nucleus, is structurally unrelated to PEP, but it belongs to the "SSU-RNAP" protein family of bacteriophages T3, T7, SP6, etc. [5-9]. The NEP mainly involves in the transcription of ~20% of the genes which are essentially non-photosynthetic and housekeeping ones, whereas the PEP involves in transcription of the rest of ~80% of the genes, which are photosynthesis-related and tRNA genes [20,21]. This was evident from the knockout mutants of PEP that showed an albino phenotype and lacked photosynthesis [22,23].

Both NEP and PEP differ in their sensitivity to the antibiotic rifampicin. PEP is sensitive to rifampicin (similar to eubacterial MSU-RNAPs), whereas the NEP is insensitive to rifampicin [5,23,24]. Furthermore, both the enzymes differ in their promoter selections too, (i. e.), NEP and PEP use different promoters for transcriptions. While most of the NEP promoters have the conserved YRTA motif, the PEP promoters resemble the bacterial 'σ70 promoters type' which are typically characterized by −10 and −35 consensus sequence motifs [25]. However, it should also be noted that a number of PEP promoters lack the −10 or the −35 elements, a few, even both [26,27]. Although chloroplasts possess all the genes for the core subunits of a PEP, but this enzyme can correctly initiate transcription only with the nuclear-encoded σ factor. The NEP transcribes the housekeeping genes as well as the genes of the PEP subunits, viz. rpoA, rpoB, rpoC1 and rpoC2 from two different operons [21] as shown in Figs. 1a and 1b suggesting the primary controls are exercised by the nucleus. In other words, the PEP transcription is controlled by the nucleus by providing the σ factor and transcription of its own genes by the NEP. Therefore, that makes the chloroplast as only a semi-autonomous organelle. (This is true for mitochondria also where NEP transcribes all the mitochondrial genes and hence, it is also a semi-autonomous organelle like chloroplast). Figs. 1a and 1b. show the transcription of PEP genes by NEP and assembly of PEP.

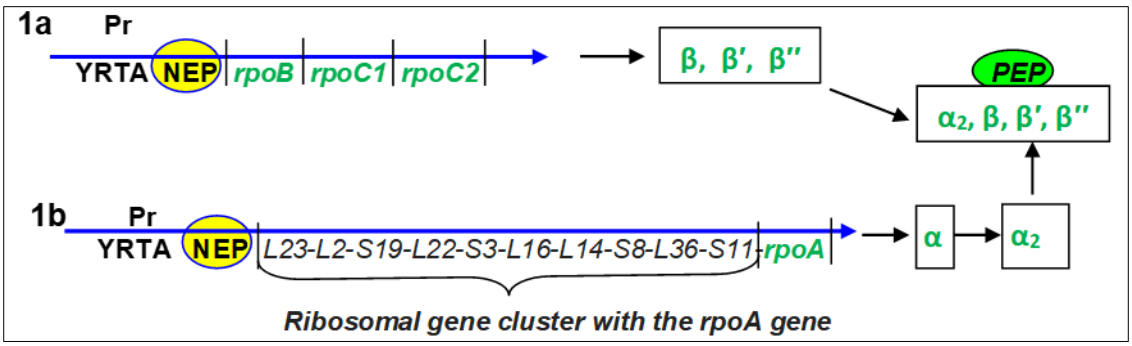


Figure 1a and 1b Schematic diagram showing transcription of PEP genes by NEP and assembly of PEP

Pr, Promoter region; L and S, Large and Small ribosomal protein subunits, *rpoA. rpoB, rpoC1, and rpoC2* are the genes for the different subunits of the PEP.

1.4. RNA Polymerases and Their PR Exonuclease Activities

The error rates of RNA polymerases are very minimal and generally in the range of 10^{-4} to 10^{-5} [2]. However, these errors are corrected during the transcription process itself. Because of this, RNA polymerases have evolved essentially with two different types of PR exonucleases (Table 1). In the MSU-RNAPs of prokaryotes, eukaryotes and chloroplasts, which are involved mainly in mRNA transcriptions, the RR active site is embedded within the polymerase active site itself [4]. However, in other RNA polymerases the PR activity is either found on the same polypeptide as a separate domain (multifunctional enzyme type, MFE) or with a closely associated subunit of the enzyme (multienzyme complex type, MEC). Table 1 shows the different types of PR activities of RNA polymerases known in viruses, prokaryotes, eukaryotes and organelles.

Table 1 PR activities and their localization in viruses, prokaryotes, eukaryotes and organellar RNA polymerases

RNAP/DNAP	Organism	PR active site	Reference
Dd**R**ps-**MSU-RNAP**	**Prokaryotes**	Intrinsic within the RNAP active site itself	[4]
	Eukaryotes		
Dd**R**ps-**MSU-RNAPs**	All eukaryotic (pols I-III)	Intrinsic within the RNAP active site itself	[4]
Dd**R**ps-**MSU-RNAPs**	Plants (pols IV & V)	Intrinsic within the RNAP active site itself	[4]
Rd**R**ps-SSU Types	**(+) Strand RNA Viruses**		
	(e.g., SARS-CoVs, SARS-related CoVs,) Human-CoVs	DEDD(**H**)-superfamily# (In the associated ExoN subunit, MEC)	[6]
Rd**R**ps-SSU Types	**(-) Strand RNA Viruses**		
	(e.g., Human influenza Viruses A, B & C)	DEDD(**H**)- superfamily# (In the PA subunit of the RNA polymerase, MEC)	[7]
	Bacteriophages		
Dd**D**ps-SSU Types	T4 DNA pol (*E. coli* Phage)	DEDD(**Y**) superfamily^ (In the same polypeptide, MFE)	[28]
	Prokaryotes		
Dd**D**ps-SSU Types	*E. coli* DNA pol I	DEDD(**Y**) superfamily^ (In the same polypeptide, MFE)	[1]
	Chloroplasts		
Dd**R**ps-SSU Types	DNA pols IA & IB (Plant chloroplasts)	DEDD(**Y**) superfamily# (In the same polypeptide, MFE)	This work
Dd**R**ps-SSU Types	RNAP (NEP) (Plant chloroplasts)	DEDD(**H**) superfamily# (In the same polypeptide, MFE)	This work

^*E. coli* DNA pol I and pol II types; DNAP, DNA polymerase; RNAP, RNA polymerase;
Similar to the ε-subunit of bacterial DNA pols III (replicases)

As the PR activity in the MSU eubacterial type PEP is already discussed in detail by Palanivelu [4], the PR active sites of the NEP and the DNA Pols IA and IB are analyzed and reported in this communication.

2. Material and methods

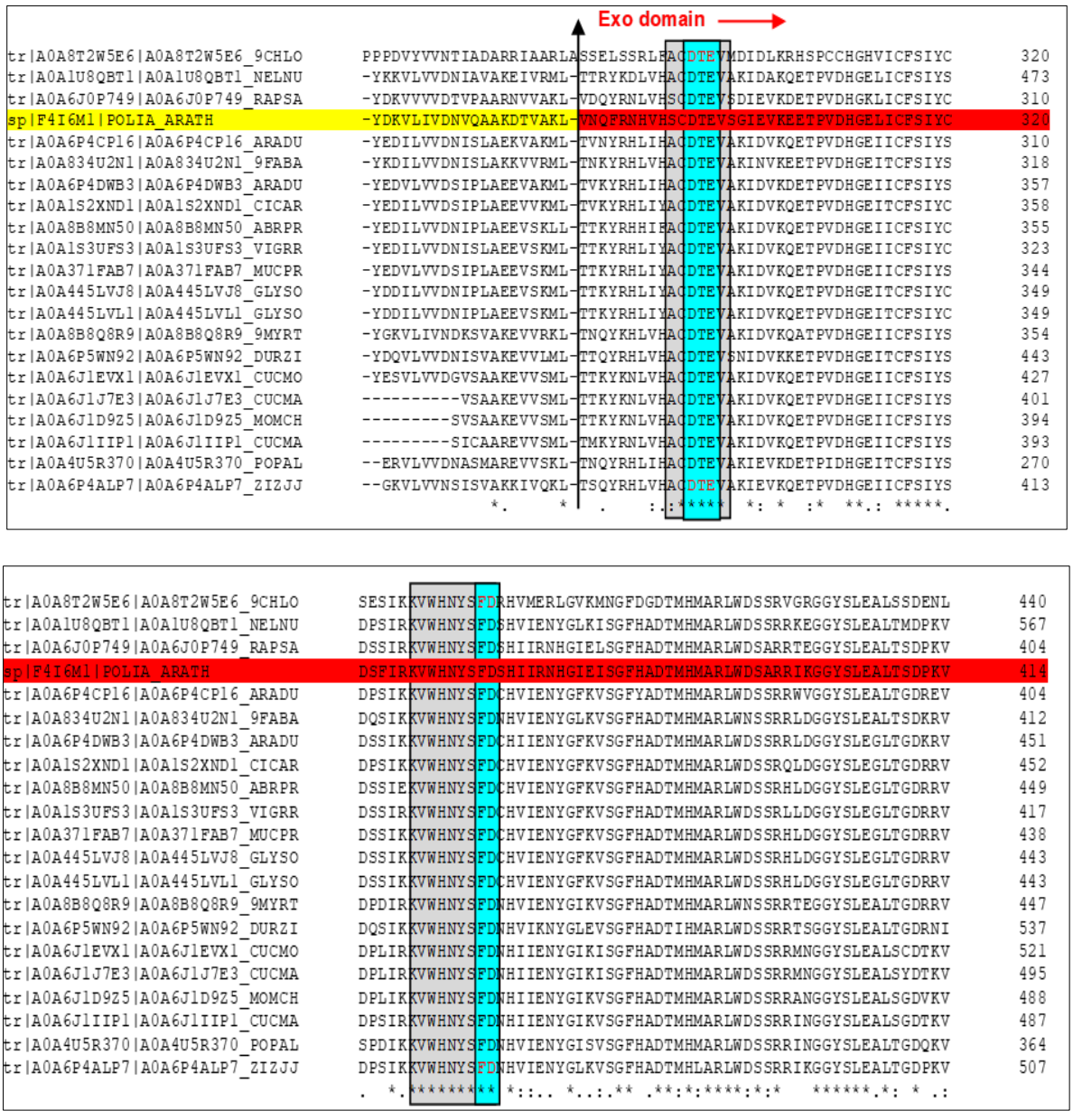
The protein sequence data of DNA polymerases IA, IB and NEPs from various plant chloroplasts were obtained from PUBMED and SWISS-PROT databases. The advanced version of Clustal Omega was used for protein sequence analysis. The polymerase and PR active sites are arrived at by the sequence similarity with other DNA and RNA polymerases already reported.

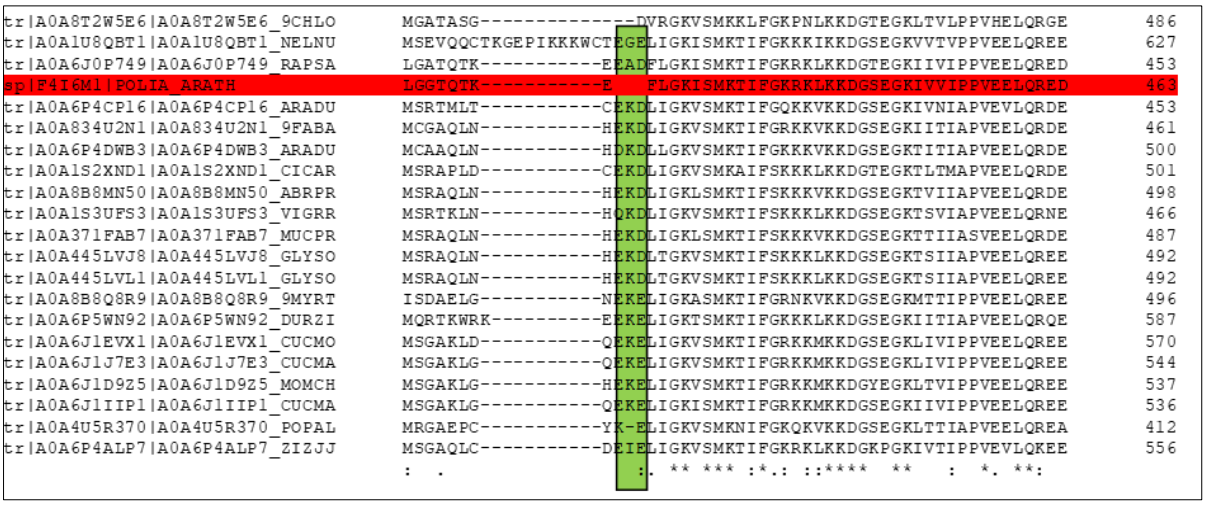
3. Results and discussion

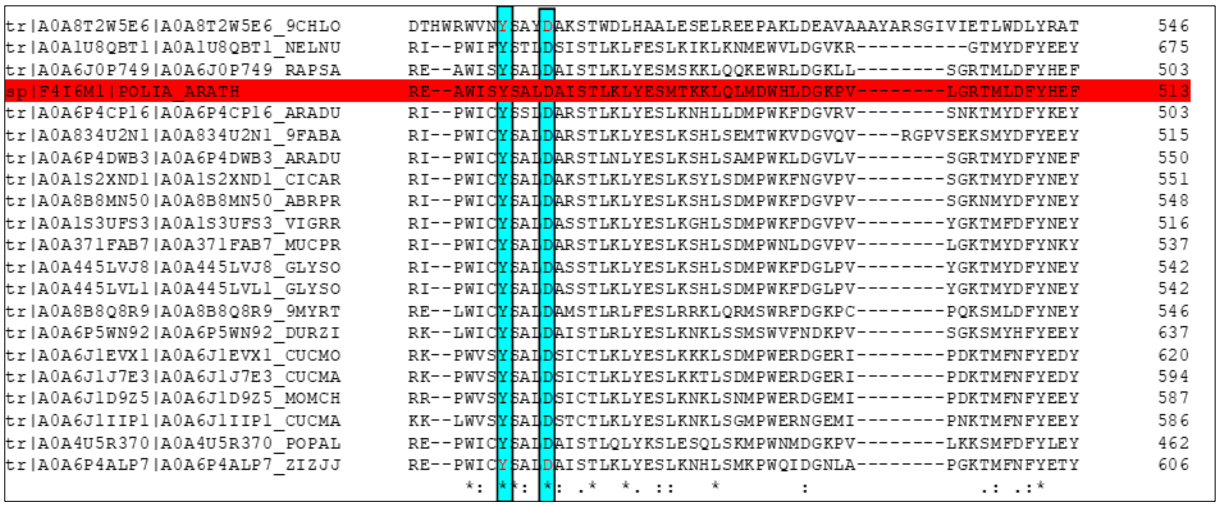
Figure 2 shows the MSA of the DNA polymerases IA of the chloroplasts from various plant sources. (Only the required regions for discussions are shown here). The *A. thaliana* sequence is used as the reference and highlighted in yellow.

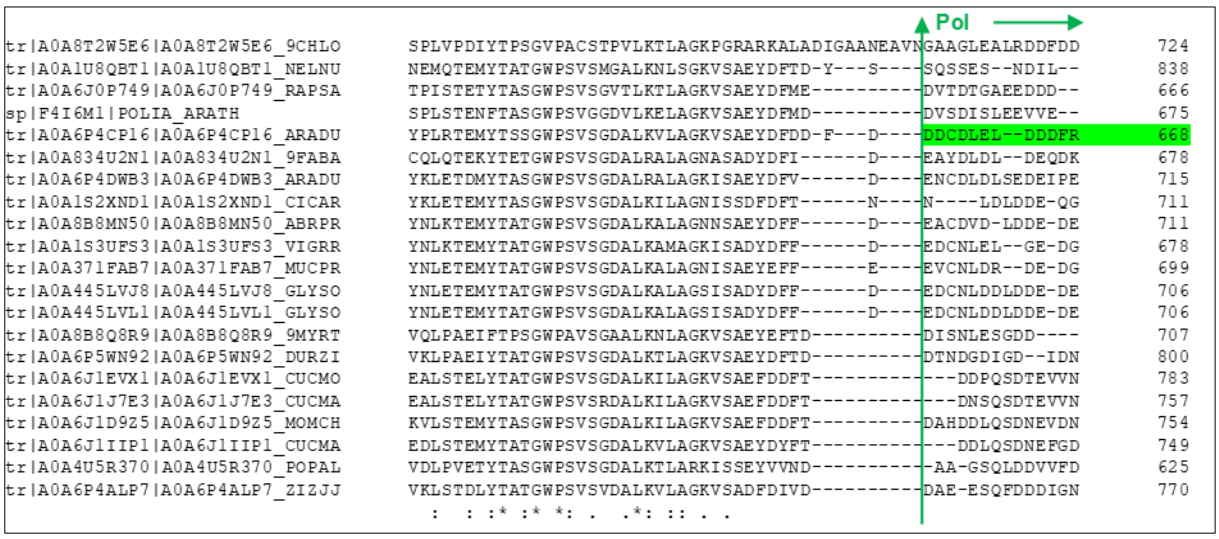
The N-terminal regions of ~300 amino acids are not conserved and showed many gaps in the alignment, after that, conservations are observed and a clear demarcation of the PR exonuclease domain is seen (highlighted in red). The PR exonuclease domain contains the typical DEDD(Y)-superfamily active site amino acids and is highlighted in light blue. There is a DxD type of metal-binding site within the PR exonuclease site (highlighted in light green). Again, after ~650 amino acids, a second demarcation in the sequences is seen and that contains the polymerase active site amino acids and are highlighted in green. The polymerase active site region contains the typical active site amino acids highlighted in yellow and a ZBM of -$CX_6CX_2CX_nC$- type is highlighted in orange. The polymerase region is completely conserved in all pol IA from different plant sources and contains the template-binding -YG- pair, the catalytic amino acid K and the nucleotide discriminating amino acid R at -4 position from the catalytic K. The active site, -**R**$^{-4}$RKA**K**^{878}M^{1}LNFSIA**Y**^{8}**G**^{887}-, is very similar to the well-established active site of *E. coli* DNA pol I, -**R**$^{-4}$RSA**K**^{758}A^{1}INFGLI**Y**^{8}**G**- [3] and in close agreement to the active sites of the other DNA/RNA polymerases already reported [1, 2]. The ZBM, within the polymerase region, is suggested to play a structural role. The C-terminal ends in a conserved hexapeptide –N**W**YSA/G**K**- in most all of the DNA pol IA, suggesting an important role.

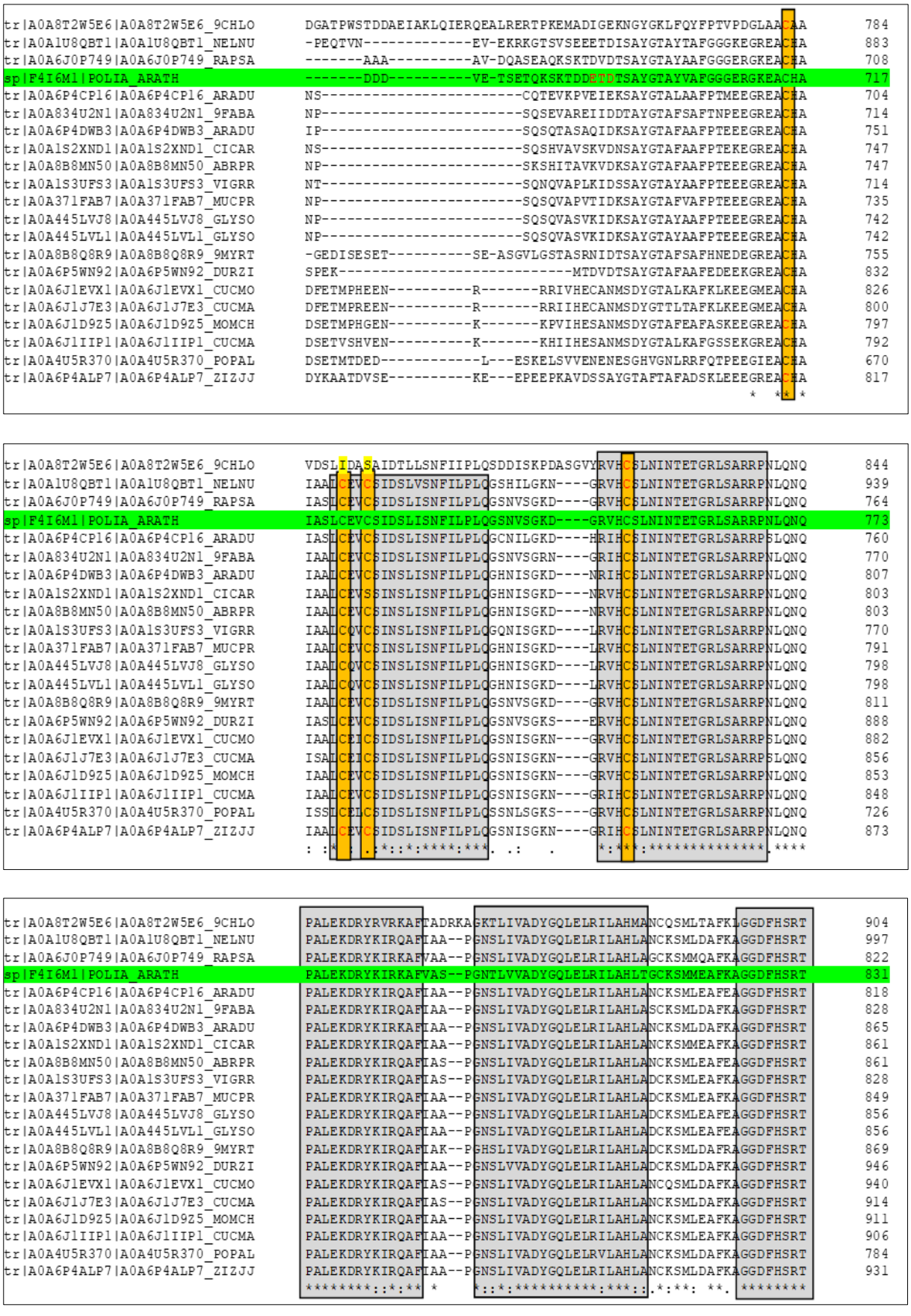
3.1. CLUSTAL O (1.2.4) MSA of the Chloroplast DNA polymerase IA

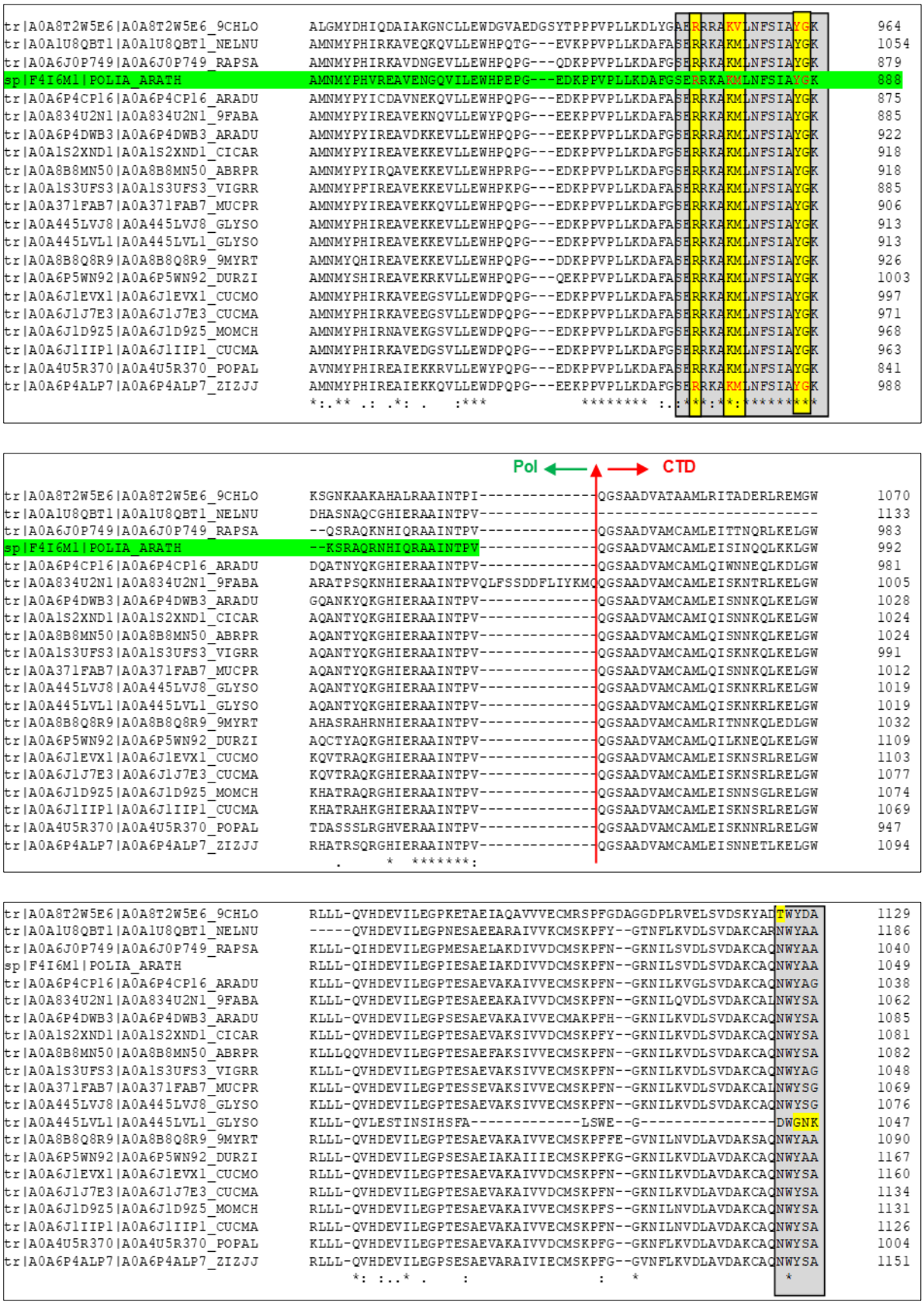


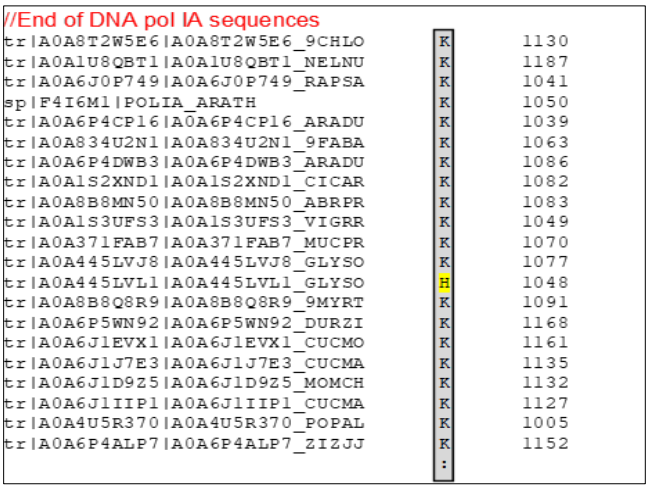

tr|A0A8T2W5E6|A0A8T2W5E6_9CHLO MGATASG----------DVRGKVSMKKLFGKPNLKKDGTEGKLTVLPPVHELQRGE 486
tr|A0A1U8QBT1|A0A1U8QBT1_NELNU MSEVQQCTKGEPIKKKWCTEGELIGKISMKTIFGKKKIKKDGSEGKVVTVPPVEELQREE 627
tr|A0A6J0P749|A0A6J0P749_RAPSA LGATQTK----------EEADFLGKISMKTIFGKRKLKKDGTEGKIIVIPPVEELQRED 453
sp|F4I6M1|POLIA_ARATH LGGTQTK----------E FLGKISMKTIFGKRKLKKDGSEGKIVVIPPVEELQRED 463
tr|A0A6P4CP16|A0A6P4CP16_ARADU MSRTMLT----------CEKDLIGKVSMKTIFGQKKVKKDGSEGKIVNIAPVEVLQRDE 453
tr|A0A834U2N1|A0A834U2N1_9FABA MCGAQLN----------HEKDLIGKVSMKTIFGRKKVKKDGSEGKIITIAPVEELQRDE 461
tr|A0A6P4DWB3|A0A6P4DWB3_ARADU MCAAQLN----------HDKDLLGKVSMKTIFGKKKVKKDGSEGKTITIAPVEELQRDE 500
tr|A0A1S2XND1|A0A1S2XND1_CICAR MSRAPLD----------CEKDLIGKVSMKAIFSKKKLKKDGTEGKTLTMAPVEELQRDE 501
tr|A0A8B8MN50|A0A8B8MN50_ABRPR MSRAQLN----------HEKDLIGKLSMKTIFSKKKVKKDGSEGKTVIIAPVEELQRDE 498
tr|A0A1S3UFS3|A0A1S3UFS3_VIGRR MSRTKLN----------HQKDLIGKVSMKTIFSKKKLKKDGSEGKTSVIAPVEELQRNE 466
tr|A0A371FAB7|A0A371FAB7_MUCPR MSRAQLN----------HEKDLIGKLSMKTIFGKKKVKKDGSEGKTTIIASVEELQRDE 487
tr|A0A445LVJ8|A0A445LVJ8_GLYSO MSRAQLN----------HEKDLTGKVSMKTIFSKKKLKKDGSEGKTSIIAPVEELQREE 492
tr|A0A445LVL1|A0A445LVL1_GLYSO MSRAQLN----------HEKDLTGKVSMKTIFSKKKLKKDGSEGKTSIIAPVEELQREE 492
tr|A0A8B8Q8R9|A0A8B8Q8R9_9MYRT ISDAELG----------NEKELIGKASMKTIFGRNKVKKDGSEGKMTTIPPVEELQREE 496
tr|A0A6P5WN92|A0A6P5WN92_DURZI MQRTKWRK----------EEKELIGKTSMKTIFGKKKLKKDGSEGKIITIAPVEELQRQE 587
tr|A0A6J1EVX1|A0A6J1EVX1_CUCMO MSGAKLD----------QEKELIGKVSMKTIFGRKKMKKDGSEGKLIVIPPVEELQREE 570
tr|A0A6J1J7E3|A0A6J1J7E3_CUCMA MSGAKLG----------QEKELIGKVSMKTIFGRKKMKKDGSEGKLIVIPPVEELQREE 544
tr|A0A6J1D9Z5|A0A6J1D9Z5_MOMCH MSGAKLG----------HEKELIGKVSMKTIFGRKKMKKDGYEGKLTVIPPVEELQREE 537
tr|A0A6J1IIP1|A0A6J1IIP1_CUCMA MSGAKLG----------QEKELIGKISMKTIFGRKKMKKDGSEGKIIVIPPVEELQREE 536
tr|A0A4U5R370|A0A4U5R370_POPAL MRGAEPC----------YK-ELIGKVSMKNIFGKQKVKKDGSEGKLTTIAPVEELQREA 412
tr|A0A6P4ALP7|A0A6P4ALP7_ZIZJJ MSGAQLC----------DEIELIGKVSMKTIFGKRKLKKDGKPGKIVTIPPVEVLQKEE 556


tr|A0A8T2W5E6|A0A8T2W5E6_9CHLO DTHWRWVNYSAYDAKSTWDLHAALESELREEPAKLDEAVAAAYARSGIVIETLWDLYRAT 546
tr|A0A1U8QBT1|A0A1U8QBT1_NELNU RI--PWIFYSTLDSISTLKLFESLKIKLKNMEWVLDGVKR----------GTMYDFYEEY 675
tr|A0A6J0P749|A0A6J0P749_RAPSA RE--AWISYSALDAISTLKLYESMSKKLQQKEWRLDGKLL--------SGRTMLDFYHEF 503
sp|F4I6M1|POLIA_ARATH RE--AWISYSALDAISTLKLYESMTKKLQLMDWHLDGKPV--------LGRTMLDFYHEF 513
tr|A0A6P4CP16|A0A6P4CP16_ARADU RI--PWICYSSLDARSTLKLYESLKNHLLDMPWKFDGVRV--------SNKTMYDFYKEY 503
tr|A0A834U2N1|A0A834U2N1_9FABA RI--PWICYSALDARSTLKLYESLKSHLSEMTWKVDGVQV----RGPVSEKSMYDFYEEY 515
tr|A0A6P4DWB3|A0A6P4DWB3_ARADU RI--PWICYSALDARSTLNLYESLKSHLSAMPWKLDGVLV--------SGRTMYDFYNEF 550
tr|A0A1S2XND1|A0A1S2XND1_CICAR RI--PWICYSALDAKSTLKLYESLKSYLSDMPWKFNGVPV--------SGKTMYDFYNEY 551
tr|A0A8B8MN50|A0A8B8MN50_ABRPR RI--PWICYSALDARSTLKLYESLKSHLSDMPWKFDGVPV--------SGKNMYDFYNEY 548
tr|A0A1S3UFS3|A0A1S3UFS3_VIGRR RI--PWICYSALDASSTLKLYESLKGHLSDMPWKFDGVPV--------YGKTMFDFYNEY 516
tr|A0A371FAB7|A0A371FAB7_MUCPR RI--PWICYSALDARSTLKLYESLKSHLSDMPWNLDGVPV--------LGKTMYDFYNKY 537
tr|A0A445LVJ8|A0A445LVJ8_GLYSO RI--PWICYSALDASSTLKLYESLKSHLSDMPWKFDGLPV--------YGKTMYDFYNEY 542
tr|A0A445LVL1|A0A445LVL1_GLYSO RI--PWICYSALDASSTLKLYESLKSHLSDMPWKFDGLPV--------YGKTMYDFYNEY 542
tr|A0A8B8Q8R9|A0A8B8Q8R9_9MYRT RE--LWICYSALDAMSTLRLFESLRRKLQRMSWRFDGKPC--------PQKSMLDFYNEY 546
tr|A0A6P5WN92|A0A6P5WN92_DURZI RK--LWICYSALDAISTLRLYESLKNKLSSMSWVFNDKPV--------SGKSMYHFYEEY 637
tr|A0A6J1EVX1|A0A6J1EVX1_CUCMO RK--PWVSYSALDSICTLKLYESLKKLSDMPWERDGERI--------PDKTMFNFYEDY 620
tr|A0A6J1J7E3|A0A6J1J7E3_CUCMA RK--PWVSYSALDSICTLKLYESLKKTLSDMPWERDGERI--------PDKTMFNFYEDY 594
tr|A0A6J1D9Z5|A0A6J1D9Z5_MOMCH RR--PWVSYSALDSICTLKLYESLKNKLSNMPWERDGEMI--------PDKTMFNFYEEY 587
tr|A0A6J1IIP1|A0A6J1IIP1_CUCMA KK--LWVSYSALDSTCTLKLYESLKNKLSGMPWERNGEMI--------PNKTMFNFYEEY 586
tr|A0A4U5R370|A0A4U5R370_POPAL RE--PWICYSALDAISTLQLYKSLESQLSKMPWNMDGKPV--------LKKSMFDFYLEY 462
tr|A0A6P4ALP7|A0A6P4ALP7_ZIZJJ RE--PWICYSALDAISTLKLYESLKNHLSMKPWQIDGNLA--------PGKTMFNFYETY 606


Pol
tr|A0A8T2W5E6|A0A8T2W5E6_9CHLO SPLVPDIYTPSGVPACSTPVLKTLAGKPGRARKALADIGAANEAVNGAAGLEALRDDFDD 724
tr|A0A1U8QBT1|A0A1U8QBT1_NELNU NEMQTEMYTATGWPSVSMGALKNLSGKVSAEYDFTD-Y---S----SQSSES--NDIL-- 838
tr|A0A6J0P749|A0A6J0P749_RAPSA TPISTETYTASGWPSVSGVTLKTLAGKVSAEYDFME----------DVTDTGAEEDDD-- 666
sp|F4I6M1|POLIA_ARATH SPLSTENFTASGWPSVGGDVLKELAGKVSAEYDFMD----------DVSDISLEEVVE-- 675
tr|A0A6P4CP16|A0A6P4CP16_ARADU YPLRTEMYTSSGWPSVSGDALKVLAGKVSAEYDFDD-F---D----DDCDLEL--DDDFR 668
tr|A0A834U2N1|A0A834U2N1_9FABA CQLQTEKYTETGWPSVSGDALRALAGNASADYDFI------D----EAYDLDL--DEQDK 678
tr|A0A6P4DWB3|A0A6P4DWB3_ARADU YKLETDMYTASGWPSVSGDALRALAGKISAEYDFV------D----ENCDLDLSEDEIPE 715
tr|A0A1S2XND1|A0A1S2XND1_CICAR YKLETEMYTASGWPSVSGDALKILAGNISSDFDFT------N----N----LDLDDE-QG 711
tr|A0A8B8MN50|A0A8B8MN50_ABRPR YNLKTEMYTATGWPSVSGDALKALAGNNSAEYDFF------D----EACDVD-LDDE-DE 711
tr|A0A1S3UFS3|A0A1S3UFS3_VIGRR YNLKTEMYTATGWPSVSGDALKAMAGKISADYDFF------D----EDCNLEL--GE-DG 678
tr|A0A371FAB7|A0A371FAB7_MUCPR YNLETEMYTATGWPSVSGDALKALAGNISAEYEFF------E----EVCNLDR--DE-DG 699
tr|A0A445LVJ8|A0A445LVJ8_GLYSO YNLETEMYTATGWPSVSGDALKALAGSISADYDFF------D----EDCNLDDLDDE-DE 706
tr|A0A445LVL1|A0A445LVL1_GLYSO YNLETEMYTATGWPSVSGDALKALAGSISADYDFF------D----EDCNLDDLDDE-DE 706
tr|A0A8B8Q8R9|A0A8B8Q8R9_9MYRT VQLPAEIFTPSGWPAVSGAALKNLAGKVSAEYEFTD----------DISNLESGDD---- 707
tr|A0A6P5WN92|A0A6P5WN92_DURZI VKLPAEIYTATGWPSVSGDALKTLAGKVSAEYDFTD----------DTNDGDIGD--IDN 800
tr|A0A6J1EVX1|A0A6J1EVX1_CUCMO EALSTELYTATGWPSVSGDALKILAGKVSAEFDDFT----------DDPQSDTEVVN 783
tr|A0A6J1J7E3|A0A6J1J7E3_CUCMA EALSTELYTATGWPSVSRDALKILAGKVSAEFDDFT----------DNSQSDTEVVN 757
tr|A0A6J1D9Z5|A0A6J1D9Z5_MOMCH KVLSTEMYTASGWPSVSGDALKILAGKVSAEFDDFT----------DAHDDLQSDNEVDN 754
tr|A0A6J1IIP1|A0A6J1IIP1_CUCMA EDLSTEMYTATGWPSVSGDALKVLAGKVSAEYDYFT----------DDLQSDNEFGD 749
tr|A0A4U5R370|A0A4U5R370_POPAL VDLPVETYTASGWPSVSGDALKTLARKISSEYVVND----------AA-GSQLDDVVFD 625
tr|A0A6P4ALP7|A0A6P4ALP7_ZIZJJ VKLSTDLYTATGWPSVSVDALKVLAGKVSADFDIVD----------DAE-ESQFDDDIGN 770




Pol
CTD

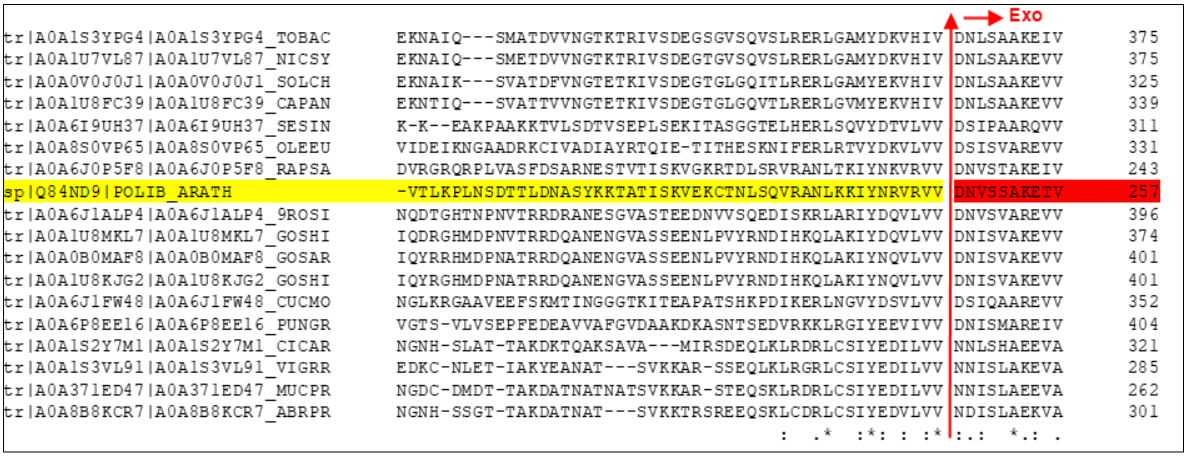


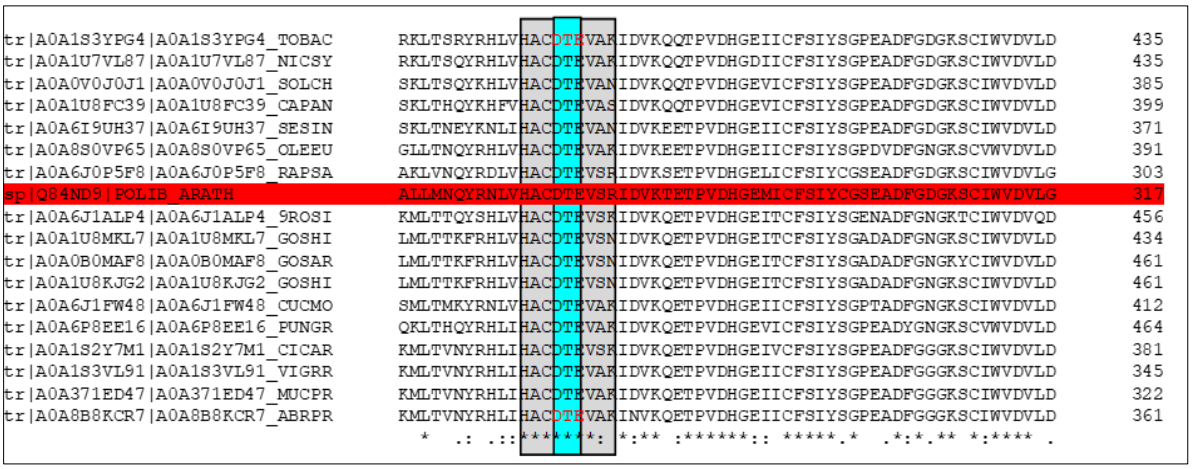
A0A8T2W5E6_9CHLO, *Chlorella desiccate*
A0A6J0P749_RAPSA, *Raphanus sativus*
A0A6P4CP16_ARADU, *Arachis duranensis*
A0A6P4DWB3_ARADU, *Arachis duranensis*
A0A8B8MN50_ABRPR, *Abrus precatorius*
A0A371FAB7_MUCPR, *Mucuna pruriens*
A0A445LVL1_GLYSO, *Glycine soja*
A0A6P5WN92_DURZI, *Durio zibethinus*
A0A6J1J7E3_CUCMA, *Cucurbita maxima*
A0A6J1IIP1_CUCMA, *Cucurbita maxima*
A0A6P4ALP7_ZIZJJ, *Ziziphus jujube*
A0A1U8QBT1_NELNU, *Nelumbo nucifera*
F4I6M1|POLIA_ARATH, *Arabidopsis thaliana*
A0A834U2N1_9FABA, *Senna tora*
A0A1S2XND1_CICAR, *Cicer arietinum*
A0A1S3UFS3_VIGRR, *Vigna radiata*
A0A445LVJ8_GLYSO, *Glycine soja*
A0A8B8Q8R9_9MYRT, *Rhodamnia argentea*
A0A6J1EVX1_CUCMO, *Cucurbita moschata*
A0A6J1D9Z5_MOMCH, *Momordica charantia*
A0A4U5R370_POPAL, *Populus alba*

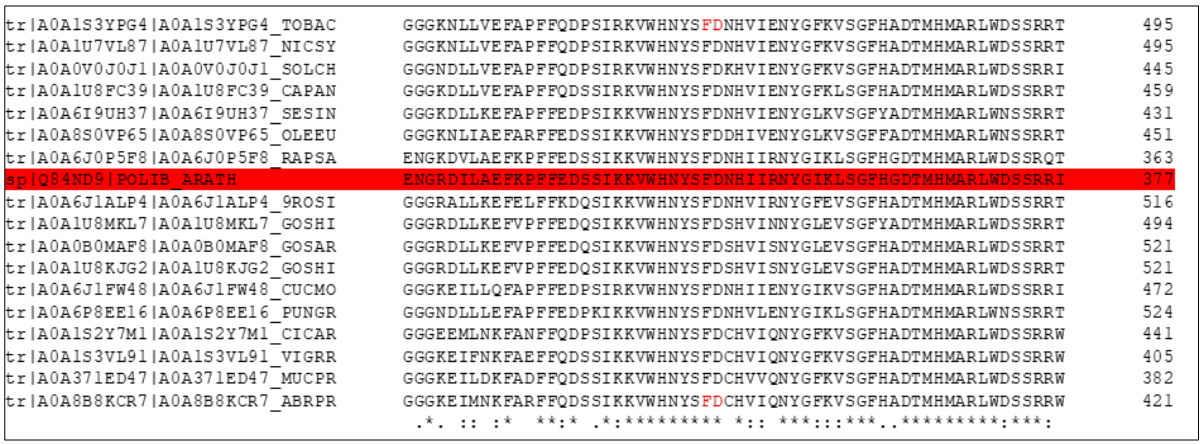
Figure 2 MSA of DNA polymerase IA of various plant chloroplasts

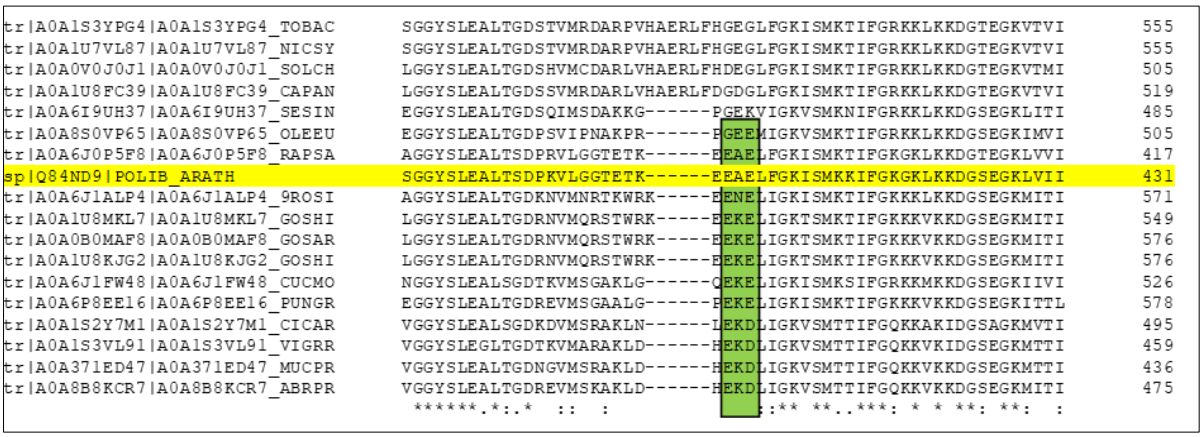
Figure 3 shows the MSA of the DNA polymerase 1B of the chloroplasts from various plant sources. (Only the required regions for discussions are shown here). The *A. thaliana* sequence is used as the reference and highlighted. The N-terminal region of ~250 amino acids is not conserved and shows many gaps in the alignment as in pol IA and, after that, conservations are observed and a clear demarcation of the PR exonuclease domain is seen (highlighted in red) as in pol IA. The PR exonuclease domain contains the typical DEDD(Y)-superfamily active site amino acids and are highlighted in light blue. There is a DxD type of metal-binding site within the PR exonuclease site in many of them (highlighted in light green). The polymerase active site region is highlighted in green. The polymerase active site region contains the typical active site amino acids highlighted in yellow and a ZBM of -$CX_6CX_2C_nC$- type is highlighted in orange and is identical to the DNA pol IA. The polymerase region is completely conserved in all and contains the template-binding -YG- pair, the catalytic amino acid K and the nucleotide discriminating amino acid R at -4 position from the catalytic K. The active site, **–R**$^{-4}$RKA**K**^{862}M^{1}LNFSIA**Y**^{8}**G**^{871}-, is identical to DNA pol IA (-**R**$^{-4}$RKA**K**M^{1}LNFSIA**Y**^{8}**G**-) and is very similar to the well-established active site of *E. coli* DNA pol I, -**R**$^{-4}$RSA**K**^{758}A^{1}INFGLI**Y**^{8}**G**- [3]. The polymerase active sites from DNA pol IA and IB are in close agreement with the active sites of the other DNA/RNA polymerases already reported [1, 2]. The ZBM, in the polymerase region, is suggested to play a structural role. The C-terminal ends in a conserved hexapeptide -NWYTAK- in all, suggesting an important role.

3.2. CLUSTAL O (1.2.4) MSA of the Chloroplast DNA Polymerase IB

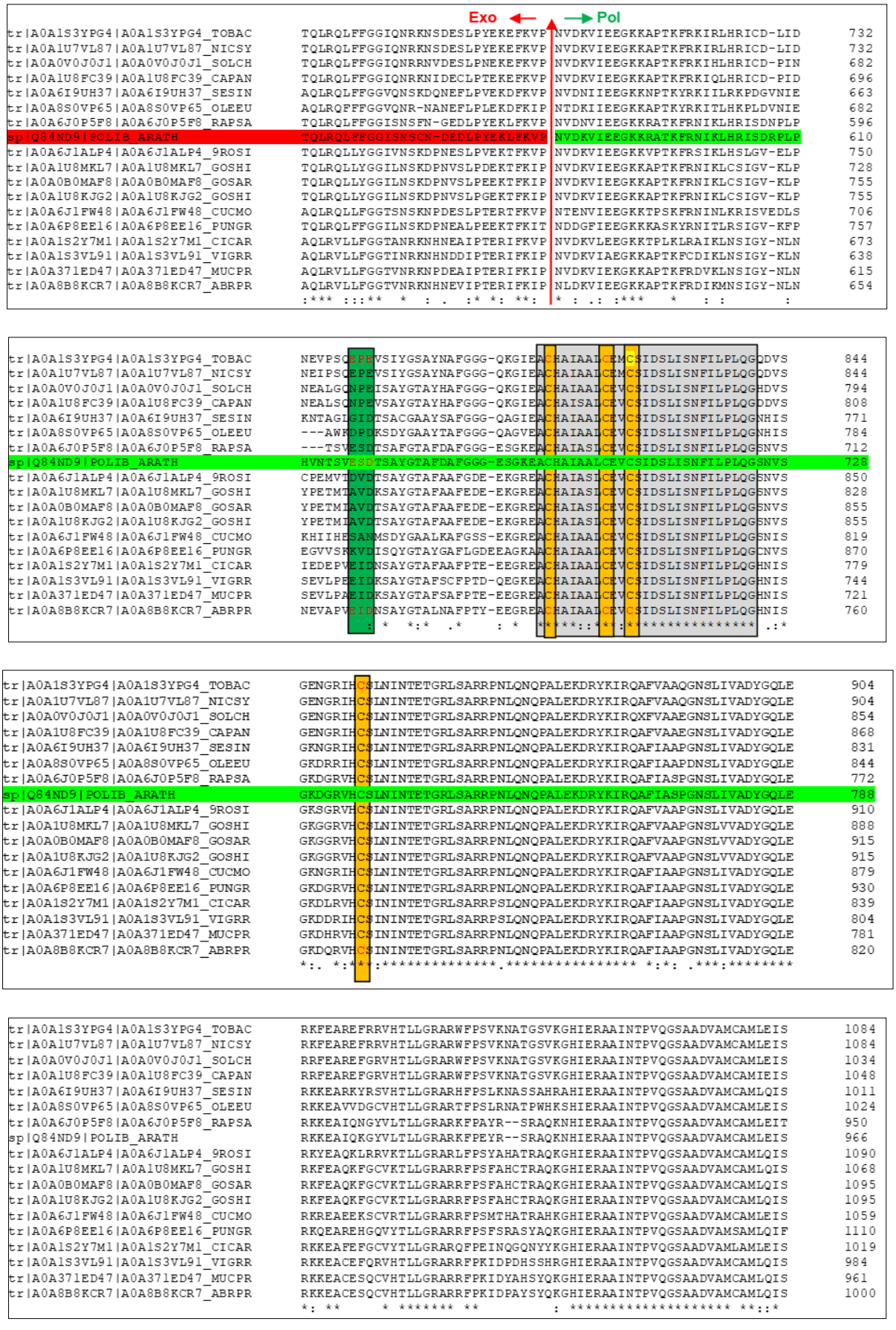


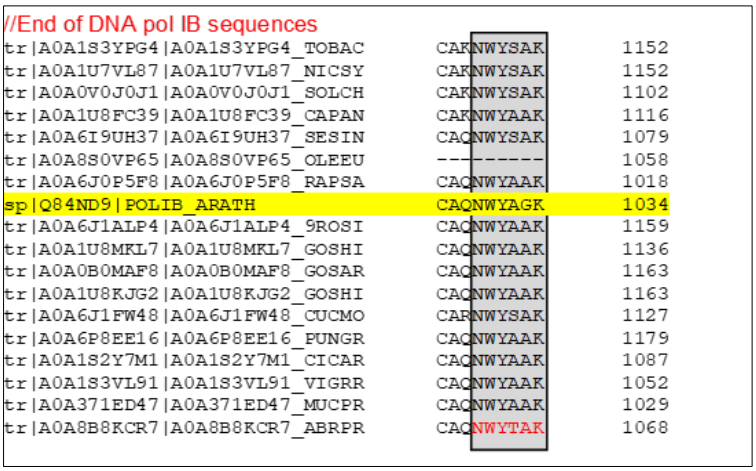

tr|A0A1S3YPG4|A0A1S3YPG4_TOBAC RKLTSRYRHLVHACDTEVARIDVKQQTPVDHGEIICFSIYSGPEADFGDGKSCIWVDVLD 435
tr|A0A1U7VL87|A0A1U7VL87_NICSY RKLTSQYRHLVHACDTEVARIDVKQQTPVDHGDIICFSIYSGPEADFGDGKSCIWVDVLD 435
tr|A0A0V0J0J1|A0A0V0J0J1_SOLCH SKLTSQYKHLVHACDTEVANIDVKQQTPVDHGEVICFSIYSGPEADFGDGKSCIWVDVLD 385
tr|A0A1U8FC39|A0A1U8FC39_CAPAN SKLTHQYKHFVHACDTEVASIDVKQQTPVDHGEVICFSIYSGPEADFGDGKSCIWVDVLD 399
tr|A0A6I9UH37|A0A6I9UH37_SESIN SKLTNEYKNLIHACDTEVANIDVKEETPVDHGEIICFSIYSGPEADFGDGKSCIWVDVLD 371
tr|A0A8S0VP65|A0A8S0VP65_OLEEU GLLTNQYRHLVHACDTEVARIDVKEETPVDHGEIICFSIYSGPDVDFGNGKSCVWVDVLD 391
tr|A0A6J0P5F8|A0A6J0P5F8_RAPSA AKLVNQYRDLVHACDTEVSHIDVKSETPVDHGELICFSIYCGSEADFGDGKSCIWVDVLG 303
sp|Q84ND9|POLIB_ARATH ALLMNQYRNLVHACDTEVSRIDVKTETPVDHGEMICFSIYCGSEADFGDGKSCIWVDVLG 317
tr|A0A6J1ALP4|A0A6J1ALP4_9ROSI KMLTTQYSHLVHACDTEVSKIDVKQETPVDHGEITCFSIYSGENADFGNGKTCIWVDVQD 456
tr|A0A1U8MKL7|A0A1U8MKL7_GOSHI LMLTTKFRHLVHACDTEVSNIDVKQETPVDHGEITCFSIYSGADADFGNGKSCIWVDVLD 434
tr|A0A0B0MAF8|A0A0B0MAF8_GOSAR LMLTTKFRHLVHACDTEVSNIDVKQETPVDHGEITCFSIYSGADADFGNGKYCIWVDVLD 461
tr|A0A1U8KJG2|A0A1U8KJG2_GOSHI LMLTTKFRHLVHACDTEVSNIDVKQETPVDHGEITCFSIYSGADADFGNGKSCIWVDVLD 461
tr|A0A6J1FW48|A0A6J1FW48_CUCMO SMLTMKYRNLVHACDTEVARIDVKQETPVDHGEIICFSIYSGPTADFGNGKSCIWVDVLD 412
tr|A0A6P8EE16|A0A6P8EE16_PUNGR QKLTHQYRHLIHACDTEVARIDVKQETPVDHGEVICFSIYSGPEADYGNGKSCVWVDVLD 464
tr|A0A1S2Y7M1|A0A1S2Y7M1_CICAR KMLTVNYRHLIHACDTEVSRIDVKQETPVDHGEIVCFSIYSGPEADFGGGKSCIWVDVLD 381
tr|A0A1S3VL91|A0A1S3VL91_VIGRR KMLTVNYRHLIHACDTEVARIDVKQETPVDHGEIICFSIYSGPEADFGGGKSCIWVDVLD 345
tr|A0A371ED47|A0A371ED47_MUCPR KMLTVNYRHLIHACDTEVARIDVKQETPVDHGEIICFSIYSGPEADFGGGKSCIWVDVLD 322
tr|A0A8B8KCR7|A0A8B8KCR7_ABRPR KMLTVNYRHLIHACDTEVARINVKQETPVDHGEIICFSIYSGPEADFGGGKSCIWVDVLD 361


tr|A0A1S3YPG4|A0A1S3YPG4_TOBAC GGGKNLLVEFAPFFQDPSIRKVWHNYSFDNHVIENYGFKVSGFHADTMHMARLWDSSRRT 495
tr|A0A1U7VL87|A0A1U7VL87_NICSY GGGKNLLVEFAPFFQDPSIRKVWHNYSFDNHVIENYGFKVSGFHADTMHMARLWDSSRRT 495
tr|A0A0V0J0J1|A0A0V0J0J1_SOLCH GGGNDLLVEFAPFFQDPSIRKVWHNYSFDKHVIENYGFKVSGFHADTMHMARLWDSSRRI 445
tr|A0A1U8FC39|A0A1U8FC39_CAPAN GGGKDLLVEFAPFFQDSSIRKVWHNYSFDNHVIENYGFKLSGFHADTMHMARLWDSSRRT 459
tr|A0A6I9UH37|A0A6I9UH37_SESIN GGGKDLLKEFAPFFEDPSIKKVWHNYSFDNHVIENYGLKVSGFYADTMHMARLWNSSRRT 431
tr|A0A8S0VP65|A0A8S0VP65_OLEEU GGGKNLIAEFARFFEDSSIKKVWHNYSFDDHIVENYGLKVSGFFADTMHMARLWNSSRRT 451
tr|A0A6J0P5F8|A0A6J0P5F8_RAPSA ENGKDVLAEFKPFFEDSSIKKVWHNYSFDNHIIRNYGIKLSGFHGDTMHMARLWDSSRQT 363
sp|Q84ND9|POLIB_ARATH ENGRDILAEFKPFFEDSSIKKVWHNYSFDNHIIRNYGIKLSGFHGDTMHMARLWDSSRRI 377
tr|A0A6J1ALP4|A0A6J1ALP4_9ROSI GGGRALLKEFELFFKDQSIKKVWHNYSFDNHVIRNYGFEVSGFHADTMHMARLWDSSRRT 516
tr|A0A1U8MKL7|A0A1U8MKL7_GOSHI GGGRDLLKEFVPFFEDQSIKKVWHNYSFDSHVINNYGLEVSGFYADTMHMARLWDSSRRT 494
tr|A0A0B0MAF8|A0A0B0MAF8_GOSAR GGGRDLLKEFVPFFEDQSIKKVWHNYSFDSHVISNYGLEVSGFHADTMHMARLWDSSRRT 521
tr|A0A1U8KJG2|A0A1U8KJG2_GOSHI GGGRDLLKEFVPFFEDQSIKKVWHNYSFDSHVISNYGLEVSGFHADTMHMARLWDSSRRT 521
tr|A0A6J1FW48|A0A6J1FW48_CUCMO GGGKEILLQFAPFFEDPSIRKVWHNYSFDNHIIENYGIKVSGFHADTMHMARLWDSSRRI 472
tr|A0A6P8EE16|A0A6P8EE16_PUNGR GGGNDLLLEFAPFFEDPKIKKVWHNYSFDNHVLENYGIKLSGFHADTMHMARLWNSSRRT 524
tr|A0A1S2Y7M1|A0A1S2Y7M1_CICAR GGGEEMLNKFANFFQDPSIKKVWHNYSFDCHVIQNYGFKVSGFHADTMHMARLWDSSRRW 441
tr|A0A1S3VL91|A0A1S3VL91_VIGRR GGGKEIFNKFAEFFQDSSIKKVWHNYSFDCHVIQNYGFKVSGFHADTMHMARLWDSSRRW 405
tr|A0A371ED47|A0A371ED47_MUCPR GGGKEILDKFADFFQDSSIKKVWHNYSFDCHVVQNYGFKVSGFHADTMHMARLWDSSRRW 382
tr|A0A8B8KCR7|A0A8B8KCR7_ABRPR GGGKEIMNKFARFFQDSSIKKVWHNYSFDCHVIQNYGFKVSGFHADTMHMARLWDSSRRW 421


tr|A0A1S3YPG4|A0A1S3YPG4_TOBAC SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI 555
tr|A0A1U7VL87|A0A1U7VL87_NICSY SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI 555
tr|A0A0V0J0J1|A0A0V0J0J1_SOLCH LGGYSLEALTGDSHVMCDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTMI 505
tr|A0A1U8FC39|A0A1U8FC39_CAPAN LGGYSLEALTGDSSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGTEGKVTVI 519
tr|A0A6I9UH37|A0A6I9UH37_SESIN EGGYSLEALTGDSQIMSDAKKG------PGEKVIGKVSMKNIFGRKKLKKDGSEGKLITI 485
tr|A0A8S0VP65|A0A8S0VP65_OLEEU EGGYSLEALTGDPSVIPNAKPR------PGEEMIGKVSMKTIFGRKKLKKDGSEGKIMVI 505
tr|A0A6J0P5F8|A0A6J0P5F8_RAPSA AGGYSLEALTSDPRVLGGTETK------PEAELFGKISMKTIFGKGKLKKDGTEGKLVVI 417
sp|Q84ND9|POLIB_ARATH SGGYSLEALTSDPKVLGGTETK------EEAELFGKISMKKIFGKGKLKKDGSEGKLVII 431
tr|A0A6J1ALP4|A0A6J1ALP4_9ROSI AGGYSLEALTGDKNVMNRTKWRK-----PENELIGKISMKTIFGKKKLKKDGSEGKMITI 571
tr|A0A1U8MKL7|A0A1U8MKL7_GOSHI LGGYSLEALTGDRNVMQRSTWRK-----PEKELIGKTSMKTIFGKKKVKKDGSEGKMITI 549
tr|A0A0B0MAF8|A0A0B0MAF8_GOSAR LGGYSLEALTGDRNVMQRSTWRK-----PEKELIGKTSMKTIFGKKKVKKDGSEGKMITI 576
tr|A0A1U8KJG2|A0A1U8KJG2_GOSHI LGGYSLEALTGDRNVMQRSTWRK-----PEKELIGKTSMKTIFGKKKVKKDGSEGKMITI 576
tr|A0A6J1FW48|A0A6J1FW48_CUCMO NGGYSLEALSGDTKVMSGAKLG------QEKELIGKISMKSIFGRKKMKKDGSEGKIIVI 526
tr|A0A6P8EE16|A0A6P8EE16_PUNGR EGGYSLEALTGDREVMSGAALG------PEKELIGKISMKTIFGKKKVKKDGSEGKITTL 578
tr|A0A1S2Y7M1|A0A1S2Y7M1_CICAR VGGYSLEALSGDKDVMSRAKLN------LEKDLIGKVSMTTIFGQKKAKIDGSAGKMVTI 495
tr|A0A1S3VL91|A0A1S3VL91_VIGRR VGGYSLEGLTGDTKVMARAKLD------HEKDLIGKVSMTTIFGQKKVKIDGSEGKMTTI 459
tr|A0A371ED47|A0A371ED47_MUCPR VGGYSLEALTGDNGVMSRAKLD------HEKDLIGKVSMTTIFGQKKVKKDGSEGKMTTI 436
tr|A0A8B8KCR7|A0A8B8KCR7_ABRPR VGGYSLEALTGDREVMSKAKLD------HEKDLIGKVSMTTIFGQKKVKKDGSEGKMITI 475


tr|A0A1S3YPG4|A0A1S3YPG4_TOBAC PSVEELQKTERELWICYSALDSISTLMLYESLKNKLSKRIWTFDGV--RKGSMYEFYEKY 613
tr|A0A1U7VL87|A0A1U7VL87_NICSY PSVEELQKTERELWICYSALDSISTLMLYESLKNKLSKRIWTFDGV--RKGSMYEFYEKY 613
tr|A0A0V0J0J1|A0A0V0J0J1_SOLCH PSVEELQRTERELWICYSALDSISTLMLYESLKKLSKRIWTFDGV--RKGSMYEFYEKY 563
tr|A0A1U8FC39|A0A1U8FC39_CAPAN PSVEELQRTERELWICYSALDSISTLMLYESLKNKLSKRVWTFDGV--RKGSMYEFYEKY 577
tr|A0A6I9UH37|A0A6I9UH37_SESIN PPVEELQRVERKLWISYSALDSISTLRLYESLEKKLRKTPWSVDGH--SKGSMFDNYVKY 543
tr|A0A8S0VP65|A0A8S0VP65_OLEEU PPVEELQRVERRLWICYSALDSISTLKLYDSLRNKLSKMPWNLDGV--VKGTMFNFYEKY 563
tr|A0A6J0P5F8|A0A6J0P5F8_RAPSA PPVEELQKDDREAWISYSALDSISTLKLYESMKKQLQVKKWFLDGKLVSGKNMFDFYQEY 477
sp|Q84ND9|POLIB_ARATH PPVKELQMEDREAWISYSALDSISTLKLYESMKKQLQAKKWFLDGKLISKKNMFDFYQEY 491
tr|A0A6J1ALP4|A0A6J1ALP4_9ROSI APVEELQREERKLWISYSALDAISTLRLYESLKSKLSSMSWVFDGKPVSGKSMYHFYKEY 631
tr|A0A1U8MKL7|A0A1U8MKL7_GOSHI APVEELQREERKLWICYSALDSISTLRLYESLKNKLSSMSWVFDGKTIPGKSMYHFYEEY 609
tr|A0A0B0MAF8|A0A0B0MAF8_GOSAR APVEELQREERKLWICYSALDSISTLRLYESLKSKLSSMSWVFDGKTIPGKSMYHFYEEY 636
tr|A0A1U8KJG2|A0A1U8KJG2_GOSHI APVEELQREERKLWICYSALDSISTLRLYESLKSKLSSMSWVFDGKTIPGKSMYHFYEEY 636
tr|A0A6J1FW48|A0A6J1FW48_CUCMO PPVEELQREEKKLWVSYSGLDSICTLKLYESLKNKLSDMPWERNGEMIPNKTMFNFYEEY 586
tr|A0A6P8EE16|A0A6P8EE16_PUNGR APVEELQREEQELWISYSSLDAISTLKLYESLRNKLLGMAWRFNGMPQAGKSMFDFYKEY 638
tr|A0A1S2Y7M1|A0A1S2Y7M1_CICAR APVEVLQREERIPWICYSALDARSTLKLYESLKSHLSDLPWKLDGE-LLHGSMFDFYEKY 554
tr|A0A1S3VL91|A0A1S3VL91_VIGRR VPVDVLQRDERIPWICYSVLDARSTLKLYESLKSHLSDMPWNLDSLPVLHKNMYDFYQEY 519
tr|A0A371ED47|A0A371ED47_MUCPR APVDVLQRDERIPWICYSVLDARSTLKLYESLKSHLSDMPWNLDGVSVLHKNMYDFYQEY 496
tr|A0A8B8KCR7|A0A8B8KCR7_ABRPR APVEVLQRDERIPWICYSVLDARSTLKLYESLKSHLSGMPWNLDGVRVLHKNMYDFYQEY 535


Exo
Pol



A0A1S3YPG4_TOBAC, *Nicotiana tabacum*
A0A0V0J0J1_SOLCH, *Solanum chacoense*
A0A6I9UH37_SESIN, *Sesamum indicum*
A0A0B0MAF8_GOSAR, *Gossypium arboretum*
Q84ND9|POLIB_ARATH, *Arabidopsis thaliana*
A0A1U8MKL7_GOSHI, *Gossypium hirsutum*
A0A6J1FW48_CUCMO, *Cucurbita moschata*
A0A1S2Y7M1_CICAR, *Cicer arietinum*
A0A1U7VL87_NICSY, *Nicotiana sylvestris*
A0A1U8FC39_CAPAN, *Capsicum annuum*
A0A8S0VP65_OLEEU, *Olea europaea*
A0A6J0P5F8_RAPSA, *Raphanus sativus*
A0A6J1ALP4_9ROSI, *Herrania umbratica*
A0A1U8KJG2_GOSHI, *Gossypium hirsutum*
A0A6P8EE16_PUNGR, *Punica granatum*
A0A1S3VL91_VIGRR, *Vigna radiate*

Figure 3 MSA of DNA polymerase IB from various plant chloroplasts

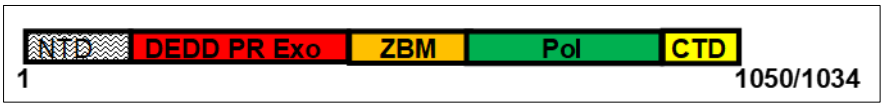
All three polymerases, viz. *E. coli* DNA pol I, DNA pol IA and IB from *A. thaliana* chloroplast are subjected to Mix and Match MSA analysis to find out the identities, similarities and conservation(s)between them. Figure 4 shows the 'Mix and Match' MSA of all three DNA pols. (only the required regions for discussions are shown). The *E. coli* DNA pol I is used as the reference and is highlighted in yellow. Interestingly, all three show identical polymerase and PR exonuclease catalytic core amino acids, suggesting their close evolutionary relatedness. However, in contrast to the *E. coli* DNA pol I, two possible ZBMs (highlighted in orange) are found in both the active site regions of *A. thaliana's* DNA pol IA and IB (highlighted in orange) (Fig. 4), suggesting structural roles.

3.3. CLUSTAL O (1.2.4) MSA of *E. coli* DNA pol I, and DNA pols IA and IB from *A. thaliana* chloroplasts

Figure 4 Mix and Match MSA of the *E. coli* DNA pol I, and the chloroplast DNA pols IA and IB from *A. thaliana*

3.4. PR exonuclease Active Site Structures of the DNA Polymerases IA and IB

Figure 5 shows the organization of the different domains on the DNA polymerases IA and IB from *A. thaliana* chloroplasts. The NTD domain of the polymerases is not conserved whereas the other four domains are highly conserved in both. A typical ZBM precedes the Pol domain and distinguishes these polymerases from the *E. coli* DNA pol I.



NTD, N-terminal domain; DEDD, PR exonuclease domain; ZBM, zinc-binding motif; Pol, polymerase domain; CTD, C-terminal domain.

Figure 5 A schematic diagramme showing the domain structure of the DNA polymerases IA and IB from plant chloroplasts (numbering from *A. thaliana*)

Based on the above data, amino acids at active sites of PR exonucleases of chloroplast DNA polymerases IA and IB are proposed (Fig. 6). The exonuclease active site is found to be very similar to the established active site of the *E. coli* DNA pol I and follows DEDD–superfamily (Fig. 4). In all three enzymes, the active site Tyr accepts the proton from the metal-bound water molecule to initiate catalysis which is followed by the formation of a highly reactive Zn-hydroxyl free radical, leading to the removal of a misincorporated nucleotide. Thus, the 3'→5' exonuclease↔polymerase activities switch between excision and incorporation modes without dissociation of the enzyme-substrate complex [29].

A. Thaliana DNA pol IA -**D**^{294}**TE**^{296}----------F**D**^{368}S--------S**Y**^{470}S→3 aa→**D**^{474}AI-

A. Thaliana DNA pol IB -**D**^{272}**TE**^{274}----------F**D**^{346}N--------S**Y**^{448}S→3 aa→**D**^{452}SI-

E. coli DNA pol I -D^{355}**T**E^{357}----------YD^{424}R--------R**Y**^{497}A→3 aa→D^{501}AD-

(Active site amino acids highlighted in dark blue are confirmed by site-directed mutagenesis (SDM) experiments analysis).

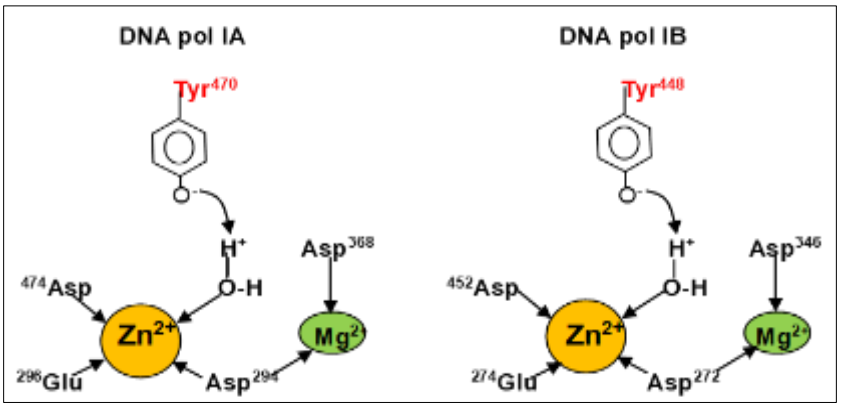


Figure 6 Proposed amino acids at the PR exonuclease active sites of the DNA polymerases IA and IB of plant chloroplasts. (the amino acid numberings are from the *A. thaliana*)

3.5. Active site Structures of the Polymerases of DNA Polymerases IA and IB

Based on the above data, amino acids at active sites of the DNA polymerase IA and IB of chloroplasts are proposed (Fig. 7). Interestingly, the polymerase active site amino acids are found to be identical to the well-established active site amino acids of the *E. coli* DNA pol I (Fig. 4). In all three enzymes, the active site K abstracts the proton from the 3'-OH of the growing primer to initiate catalysis, which is followed by a nucleophilic-electrophilic attack with the subsequent addition of the incoming base to the growing primer. For a detailed mechanism, see Palanivelu [30].

DNA pol IA **DNA pol IB**

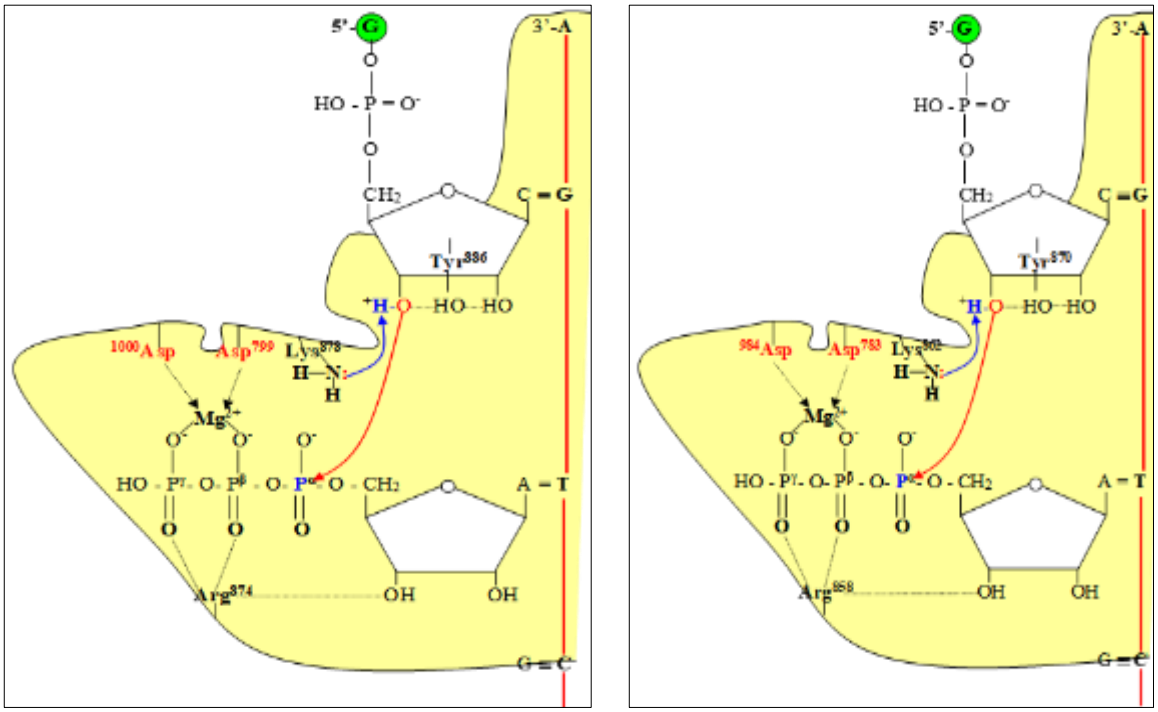
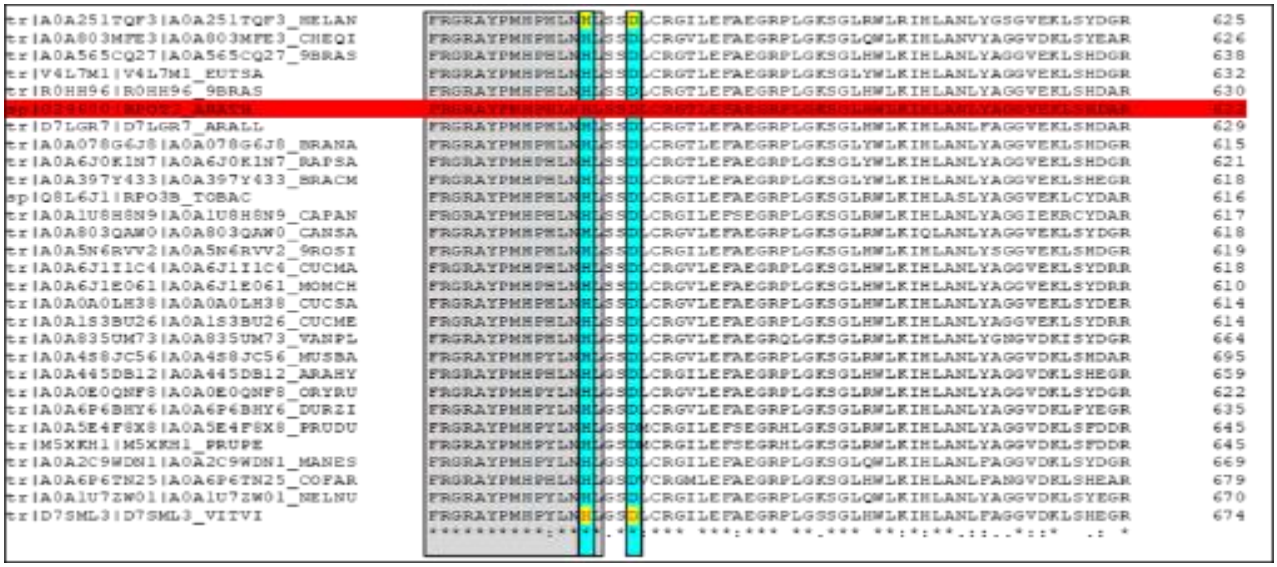


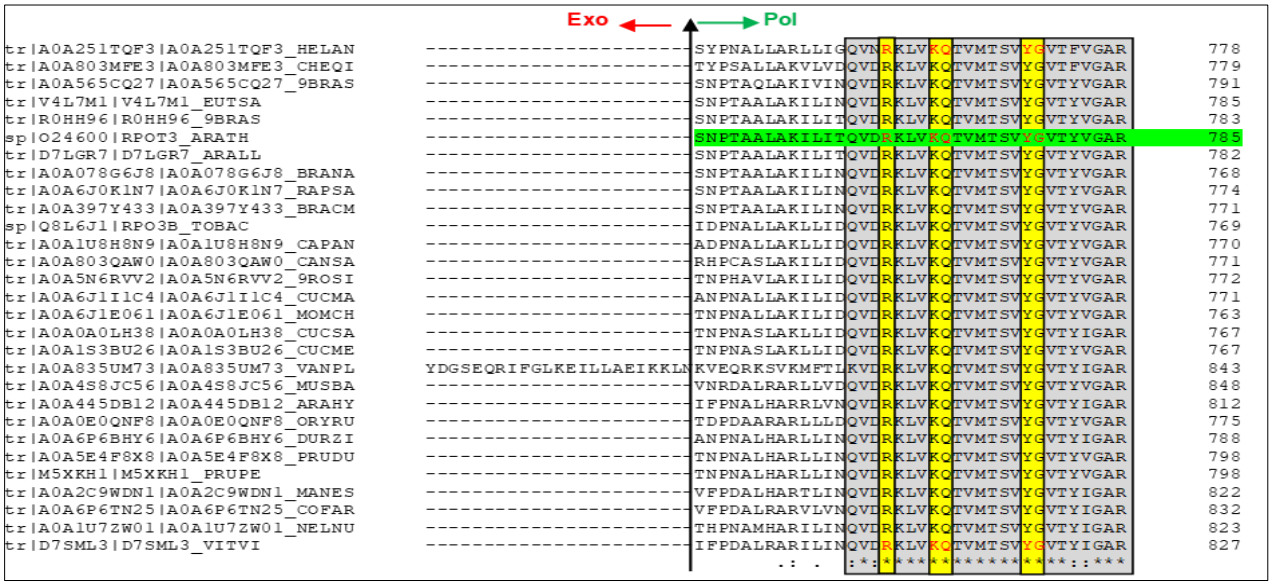
Figure 7 Proposed amino acids at the active sites of the DNA polymerases IA and IB of the plant chloroplasts. (the amino acid numberings are from the *A. thaliana*)

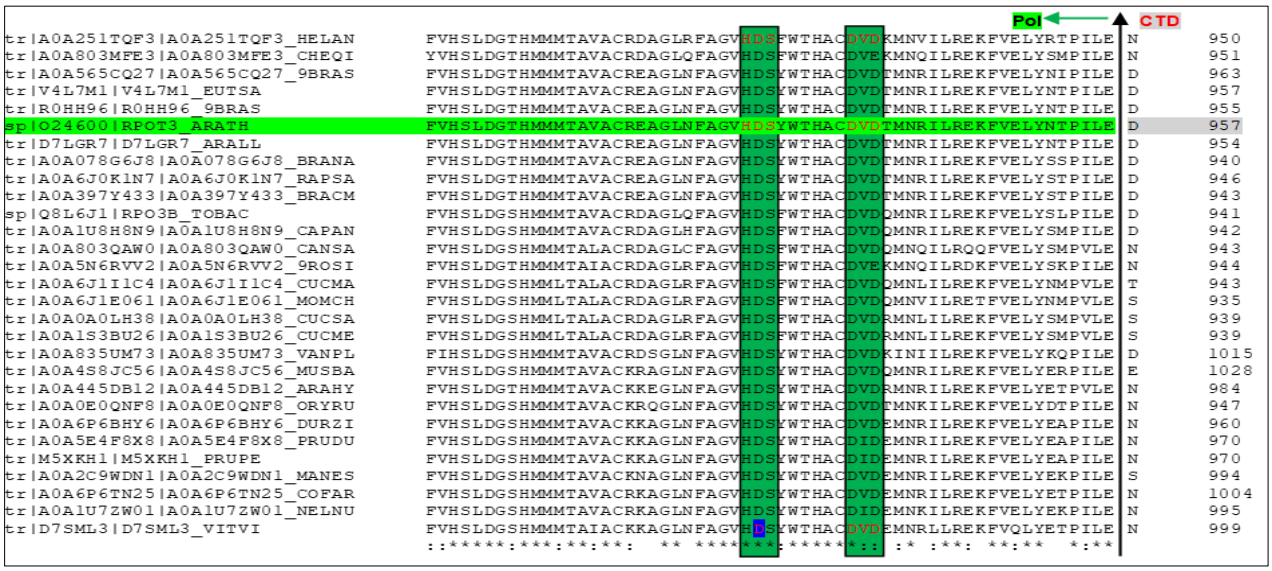
Figure 8 shows the MSA of the NEPs from the chloroplasts of various plant sources. (Only the required regions for discussions are shown here). The *A. thaliana* sequence is used as the reference and highlighted in yellow. The N-terminal region up to 400-500 amino acids is not conserved and shows many gaps in the alignment and, after that, conservations are observed and a clear demarcation of the PR exonuclease domain is also seen (highlighted in red). The PR exonuclease domain contains the typical DEDD(H) superfamily active site amino acids and is highlighted in light blue. A DxD type of metal-binding site within the PR exonuclease site is also found (highlighted in light green). Again, after ~750 amino acids a second demarcation in the sequences is observed and that contains the polymerase active site

region and highlighted in green. The polymerase region contains the typical active site amino acids highlighted in yellow and a DxD type metal-binding motif (highlighted in dark green). The polymerase region is completely conserved in all and contains the template-binding -YG- pair, the catalytic amino acid K and the nucleotide discriminating amino acid R at -4 position from the catalytic K. The active site–**R**$^{-4}$KLV**K**Q^{1}TVMTSV**Y**^{8}**G**-, is very similar to the active site of *E. coli* DNA pol I, -**R**$^{-4}$RSA**K**^{758}A^{1}INFGLI**Y**^{8}**G**- [3] and in close agreement to the active sites of the other DNA/RNA polymerases already reported [1,2]. The C-terminal end is remarkably conserved and ends in the tetrapeptide –YFFN- suggesting an important role. However, the ZBMs are not located in the RNA polymerase NEP, as in chloroplast DNA pols IA and IB.

3.6. CLUSTAL O (1.2.4) MSA of NEPs from the Chloroplasts of Various Plant Sources




Exo
Pol


Pol
CTD

A0A251TQF3_HELAN, *Helianthus annuus*
A0A565CQ27_9BRAS *Arabis nemorensis*
R0HH96_9BRAS, *Capsella rubella*
D7LGR7_ARALL, *Arabidopsis lyrata* subsp. *lyrata*
A0A6J0K1N7_RAPSA, *Raphanus sativus*
Q8L6J1|RPO3B_TOBAC, *Nicotiana tabacum*
A0A803QAW0_CANSA, *Cannabis sativa*
A0A6J1I1C4_CUCMA, *Cucurbita maxima*
A0A0A0LH38_CUCSA, *Cucumis sativus*
A0A835UM73_VANPL, *Vanilla planifolia*
A0A445DB12_ARAHY, *Arachis hypogaea*
A0A6P6BHY6_DURZI, Durio zibethinus
M5XKH1_PRUPE, Prunus persica
A0A6P6TN25_COFAR, Coffea Arabica
D7SML3_VITVI, Vitis vinifera
A0A803MFE3_CHEQI, *Chenopodium quinoa*
V4L7M1_EUTSA, *Eutrema salsugineum*
O24600|RPOT3_ARATH, *Arabidopsis thaliana*
A0A078G6J8_BRANA, *Brassica napus*
A0A397Y433_BRACM, *Brassica campestris*
A0A1U8H8N9_CAPAN, *Capsicum annuum*
A0A5N6RVV2_9ROSI, *Carpinus fangiana*
A0A6J1E061_MOMCH, *Momordica charantia*
A0A1S3BU26_CUCME, *Cucumis melo*
A0A4S8JC56_MUSBA, *Musa balbisiana*
A0A0E0QNF8_ORYRU, Oryza rufipogon
A0A5E4F8X8_PRUDU, Prunus dulcis
A0A2C9WDN1_MANES, Manihot esculenta
A0A1U7ZW01_NELNU, Nelumbo nucifera

Figure 8 MSA of NEPs from the chloroplasts of various plant sources

3.7. Active Site Amino Acid Analyses of the DNA Polymerases and NEP of Plant Chloroplasts

Figure 9 shows the domain organization of the NEP of *A. thaliana* chloroplast. The NTD is not conserved whereas the other three regions are highly conserved.

Figure 9 Different domains of the NEP of chloroplasts (Numbering from the *A. thaliana* sequence)

The polymerase catalytic core region essentially contains three components, viz. a template-binding pair –YG-, a basic catalytic amino acid -K/R- and a nucleotide discriminating amino acid –R-, placed at -4 to -5 from the catalytic K/R. These three highly conserved components are found located in the NEP of the chloroplasts and as found in the chloroplast DNA pols IA and IB also (Figs. 2,3) The catalytic core regions are in close agreement with those already reported from other DNA/RNA polymerases (Table 2).

Table 2 Catalytic core regions of various RNA and DNA polymerases

Polymerase type	Catalytic core
SSU RNA/DNA pols	
Viral T7 SSU RNA pol	-620WLAY-8GVTR-4SVTKR1SVMTLAY8GS-
Viral SP6 SSU RNA Pol	-612WDSI-8GITR-4SLTKK1PVMTLPY8GS-
Mitochondrial SSU RNA pol (*Sc*)	-1009TR-4KVVKQ1TVMTNVY8GV-
Mitochondrial SSU RNA pol (*H. sapiens*)	-986TR-4KVVKQ1TVMTVVY8GV-
E. coli DNA pol I (SSU)	-753QR-4RSAK758A1INFGLIY8GM-
Chloroplast SSU RNA pol (NEP) (*ARATH*)	-765DR-4KLVK770Q1TVMTSVY8GV-
Chloroplast SSU DNA pol IA (*ARATH*)	-873ER-4RKAK878M1LNFSIAY8GK-
Chloroplast SSU DNA pol IB (*ARATH*)	-857ER-4RKAK862M1LNFSIAY8GK-

Sc, *Saccharomyces cerevisiae*; *ARATH*, *Arabidopsis thaliana*;
The active site amino acids, highlighted in dark blue, are confirmed by SDM analysis

3.8. Active Site Amino Acids at the DEDD(Y/H)-superfamily of PR Exonucleases in DNA Polymerases IA, IB and NEP of Plant Chloroplasts

The PR exonuclease active sites of NEPs from various plant chloroplasts are arrived at from the sequence similarity to other well-established DEDD-exonuclease superfamily as substantiated below. The DNA pols IA and IB use the DEDD(Y)-superfamily of exonuclease to PR the errors during chloroplast genome replication. The *E. coli* DNA pol I also uses the DEDD(Y)-superfamily PR exonuclease and its active site amino acids were confirmed both by SDM and X-ray crystallographic analyses by different investigators [31-33]. The *E. coli* DNA pol II also uses the -DEDD(Y)-superfamily and its active site amino acids were again confirmed both by SDM and X-ray crystallographic analyses by Wang and Yang [34]. They have shown that the amino acids 147 to 367 comprised the 3'→5' PR exonuclease domain and amino acids from 368 to 783 were involved in polymerase function. They found that the D^{335}→N mutant of the active site lost its exonuclease activity and hence, suggested its involvement in the catalysis. The exonuclease active site amino acids of the pol II were further corroborated by the following (D^{156}→N, D^{229}→N and D^{335}→N) SDM exo$^-$ mutants [6] (Table 3). Furthermore, RNase D (EC 3.1.13.5), one of the seven exoribonucleases, which is involved in the 3'-maturation of several stable RNAs like tRNAs, 5S rRNA, and other small structured RNAs, was also shown to belong to the -DEDD(Y)-family [8]. Furthermore, the involvement of the DEDD(Y)-superfamily in PR exonuclease in the eukaryotic replicases, viz. DNA polymerases δ and ε from *Saccharomyces cerevisiae* have been reported recently [35,36]. They found that the double mutant D^{290}→A/E^{292}→A (from the invariant first triad –D^{290}xE^{292}- of the DEDD(Y)-superfamily), was exonuclease deficient (Table 3). The replicative DNA polymerases δ and ε exhibit both polymerase and exonuclease activities [3].

Many of the PR exonucleases use DEDD(H)-superfamily. Fijalkowska and Schaaper [37] have found DEDD(H)-superfamily of PR exonuclease in the ε-subunits of the bacterial replicase multienzyme complexes (DNA pols III) belongs to the DnaQ-H-family with the four active site carboxylates (Asp^{12}, Glu^{14}, Asp^{103}, and Asp^{167}) with the invariant His^{162}, which acts as the general base in catalysis. They also found that modification of the two conserved amino acid residues, viz. Asp^{12}→Ala and Glu^{14}→Ala, in the ε-subunit by SDM experiments, resulted in the loss of the exonuclease function and hence, suggested that they might play a role in the coordination of the catalytic metal ion. These observations were further corroborated by X–ray crystallographic analysis of the ε-186 by Hamdan et al. [38].

3.9. PR Exonuclease Active Site Structure of the NEPs from Plant Chloroplasts

Figure 10 shows the proposed amino acids at the active site of the PR exonuclease of chloroplast NEPs. The active site is found to be very similar to the established active site of the ε-subunit of *E. coli* DNA pol III (Table 3) as both are coming from nuclear encoded.

NEP-PR Exonuclease $^{-}{}_{545}$**L**RC**D**^{548}V**E**^{550}LKL----------NL**D**^{572}F-------LN**H**^{586}→3aa→**D**^{590}L$C_{592}{}^{-}$

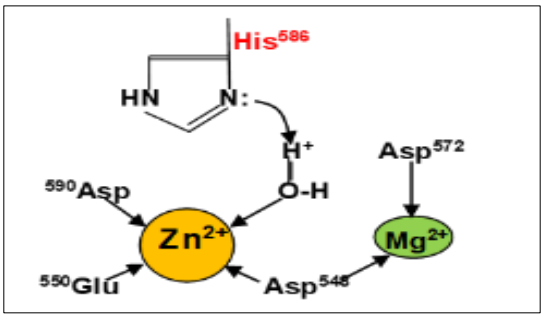


Figure 10 Proposed amino acids at the PR exonuclease active site of the NEP from chloroplasts (the amino acids numberings are from the *A. thaliana*)

In all three exonucleases from the chloroplasts, the active site Tyr/His accepts the proton from the metal-bound water molecule to initiate catalysis, which is followed by the formation of a highly reactive Zn-hydroxyl free radical, leading to the removal of a misincorporated nucleotide. Thus, the 3'→5' exonuclease↔polymerase activities switch between excision and incorporation modes without dissociation of the enzyme-substrate complex [29].

Table 3 DEDD-superfamily of exonuclease active site amino acids from viral, bacterial, fungal, plant and animal sources

Exo-Family	Consensus A-site Pattern	Proton Acceptor	Catalytic Metal ion*	No. of Zn-binding site(s)
DEDD(Y/H)-family				
Prokaryotic PR Enzymes				
T4 DNA pol (*E. coli* Phage)	-D^{112}xE^{114}-D^{219}-Y^{320}-D^{324}-	Tyr	Zn^{2+}	1
DNA pol I (*E. coli*)	-D^{355}xE^{357}-D^{424}-Y^{497}-D^{501}-	Tyr	Zn^{2+}	1
DNA pol II (*E. coli*)	-D^{156}xE^{158}-D^{229}-Y^{331}-D^{335}-	Tyr	Zn^{2+}	1
RNase D (*E. coli*)	-D^{28}xE^{30}---D^{85}--Y^{151}-D^{155}-	Tyr	Zn^{2+}	1
DNA pol III, ε-subunit (*E. coli*)	-D^{12}xE^{14}-D^{103}-H^{162}-D^{167}-	His	Zn^{2+}	1
RNase T (*E. coli*)	-D^{23}xE^{25}-D^{125}-H^{181}-D^{186}-	His	Zn^{2+}	1
PR Enzymes in Eukaryotic DNA Replicases				
DNA pol ε cat. subunit (*S.c*)	-D^{290}XE^{292}-D^{383}-Y^{473}-D^{477}-	His	Zn^{2+}	1
DNA pol δ cat. subunit (*H.s*)	-D^{316}XE^{318}-D^{402}-Y^{511}-D^{515}-	His	Zn^{2+}	1
PR Enzymes in DNA Polymerases 1A and 1B from Plant Chloroplasts				
DNA polymerase 1A (*ARATH*)	-DxE^{296}------D^{368}----Y^{470}—D^{474}-	Tyr	Zn^{2+}	1
DNA polymerase 1B (*ARATH*)	-DxE^{272}------D^{346}----Y^{448}—D^{452}-	Tyr	Zn^{2+}	1
PR Enzymes in RNA Polymerase, NEP from Plant Chloroplasts				
Arabidopsis Thaliana	-DxE^{550}------D^{572}----H^{586}—D^{590}-	His	Zn^{2+}	1
Arachis hypogaea	-DxE---------D--------H-----D-	His	Zn^{2+}	1
Oryza rufipogon	-DxE---------D--------H-----D-	His	Zn^{2+}	1
Nelumbo nucifera	-DxE---------D--------H-----D-	His	Zn^{2+}	1

Adapted from Palanivelu [7]

A-site, Active site; *Water-bound Zn2+; Active site amino acids confirmed by SDM are highlighted in dark blue and X-ray crystallography in light blue. *S.c, Saccharomyces cerevisiae; H.s, Homo sapiens*; *ARATH, Arabidopsis thaliana.*

Table 4 shows the DEDD-superfamily exonuclease active site amino acids and their distance conservations.

Table 4 DEDD-superfamily exonuclease active site amino acids and their distance conservations

Adapted from Palanivelu [7]. *Sc, Saccharomyces cerevisiae; Hs, Homo sapiens.*

*The distance between the proton acceptor (Y/H) and the last invariant D is 3 to 4 amino acids.

#Very similar active site amino acids are found in *E. coli* DNA pol I. ^Similar active site amino acids are found and confirmed by SDM in SARS-CoVs (6); ACE2, Angiotensin-Converting Enzyme 2; DPP4, Dipeptidyl peptidase 4. Active site amino acids confirmed by SDM analysis are highlighted in dark blue and by X-ray crystallography, are highlighted in light blue.

4. Conclusions

Chloroplast DNA polymerases IA and IB are highly homologous with 72.45% identity and possess very similar active sites for polymerization and proofreading functions. Their polymerization and proofreading active sites are very similar to the bacterial DNA polymerase I, suggesting their possible common evolutionary origin. However, the chloroplast DNA polymerases IA and IB differ from the bacterial DNA polymerase I by possessing a typical ZBM in them, as found in eukaryotic replicases. The ZBM is suggested to play a structural role. The polymerase catalytic cores of the DNA polymerases IA, IB and NEP are remarkably conserved. The proofreading exonucleases of all three polymerases (IA, IB and NEP) belong to the DEDD-superfamily. The DNA polymerases IA and IB belong to the DEDD(Y)- subfamily, whereas the NEP belongs to the DEDD(H)- subfamily.

Compliance with ethical standards

Acknowledgments

The author wishes to thank Dr. P. Sathiyamoorthy, Chief Scientist, GEM Research Foundation, Chennai, for useful suggestions on the manuscript.

References

[1] Palanivelu P. DNA polymerases – An insight into their active sites and mechanism of action, In: Recent Advances in Biological Research, Vol 1, Chapter 2, pp 1-39, SCIENCEDOMAIN International Book Publishers, UK. ISBN: 9788193422441, DOI: 10.9734/bpi/rabr/v1; 2019.

[2] Palanivelu P. Single Subunit RNA Polymerases: An Insight into their Active Sites and Catalytic Mechanism, In: Advances and Trends in Biotechnology and Genetics. Vol 1, Chapter 1, pp 1-38, SCIENCEDOMAIN International Book Publishers, UK. ISBN: 978-93-89246-59-9, DOI:10.9734/bpi/atbg/v1; 2019.

[3] Palanivelu P. An Insight into the Mechanism of Genome Duplication in Eukaryotes: Polymerase and Proofreading Functions by Eukaryotic DNA Replicases. Int J Rec Sci Res. 13:2076-2116, (2022).

[4] Palanivelu P. Multi-subunit RNA polymerases of bacteria - An insight into their active sites and catalytic mechanism. Indian J Sci Technol. 2018; 11: 1-37.

[5] Palanivelu P. Active Sites of the Multi-subunit RNA Polymerases of Eubacteria and Chloroplasts are Similar in Structure and Function: Recent Perspectives. In: Current Research Trends in Biological Science Vol. 2. Chapter-3 pp 26-61, SCIENCEDOMAIN International Book Publishers, UK. Print ISBN: 978-93-90149-66-7, eBook ISBN: 978-93-90149-14-8, DOI: 10.9734/bpi/crtbs/v2;2020

[6] Palanivelu P. An overview of the proofreading functions in bacteria and SARS-Coronaviruses. Int J Biochem Res Rev. 2021; 30: 33-62.

[7] Palanivelu P. Identification of DEDD- and PHP-Superfamilies of Proofreading Exonucleases in the Acidic Protein Subunit PA of RNA Polymerase of Human Influenza Viruses, World J Adv Res Rev. 2022; 16: 804–824.

[8] Zuo Y, Deutscher MP. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res. 2001; 29:1017–1026

[9] Nagpal S, Nair DT. The PHP domain of PolX from Staphylococcus aureus aids high fidelity DNA synthesis through the removal of misincorporated deoxyribo-, ribo- and oxidized nucleotides. Nat Sci Rep. 2021; 11:4178.

[10] Timinskas K, Balvociute M, Timinskas A, Venclovas C. Comprehensive analysis of DNA polymerase III a subunits and their homologs in bacterial genomes. Nucleic Acids Res. 2014; 42: 1393–1413.

[11] Nielsen BL, Cupp JD, Brammer J (2010) Mechanisms for maintenance, replication, and repair of the chloroplast genome in plants. J Exp Bot. 61: 2535–2537.

[12] Palanivelu P. Analyses of priming reactions and proofreading functions during initiation of replication of prokaryotic and eukaryotic genomes. Brit J Pharma Med Res. 2022; 7:3790-3828.

[13] Morley SA, Nielsen BL. Chloroplast DNA Copy Number Changes during Plant Development in Organelle DNA Polymerase Mutants. Front Plant Sci. 2016, 7.

[14] Parent JS, Lepage E, Brisson N. Divergent Roles for the Two PolI-Like Organelle DNA Polymerases of Arabidopsis. Plant Physiol. 2011, 156, 254–262.

[15] Baruch-Torres N, Brieba LG. Plant organellar DNA polymerases are replicative and translesion DNA synthesis polymerases. Nucleic Acids Res. 2017, 45, 10751–10763.

[16] Sugiura M. The chloroplast genome. Essays Biochem. 1995, 30, 49–57.

[17] Serino G, Maliga P. RNA polymerase subunits encoded by the plastid rpo genes are not shared with the nucleus-encoded plastid enzyme. Plant Physiol. 1998; 117:1165-1170.

[18] Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 2004; 5:123–135.

[19] Fukuzawa H, Kohchi T, Sano T, Nakahigashi K, Umesono K. Cold spring harbor symposium for quantitative biology. 1987; 52:791-804.

[20] Hedtke B, Börner T, Weihe A. Mitochondrial and chloroplast phage-type RNA polymerase in Arabidopsis. Science, 1997;277:809–811. 11.

[21] Hajdukiewicz PT, Allison LA, Maliga P. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J. 1997; 16:4041–4048. 12.

[22] Allison LA, Simon LD, Maliga P. Deletion of rpoB reveals a second distinct transcription system in plastids of higher plants. EMBO J. 1996; 15:2802–2809.

[23] Maciossek G, Kofer W, Bock A, Schoch S, Maier RM, Wanner G, Rudiger W, Koop HU, Herrmann RG. Targeted disruption of the plastid RNA polymerase genes rpoA, B and C1: molecular biology, biochemistry and ultrastructure. Plant J. 1999; 18:477–489.

[24] Severinov K, Mustaev A, Kukarini A, Muzzin O, Little MC, Hallick RB. Chloroplast rpoA, rpoB, and rpoC Genes Specify at least Three Components of a Chloroplast DNA-dependent RNA polymerase active in tRNA and mRNA Transcription. J Biol Chem. 1998; 263:14302-14307.

[25] Allison LA. The role of sigma factors in plastid transcription. Biochimie. 2000; 82: 537–548.

[26] Gruissem W, Zurawski G. Analysis of promoter regions for the spinach chloroplast rbcL, atpB and psbA genes. EMBO J. 1985; 4: 3375-3383.

[27] Gruissem W, Zurawski G. Identification and mutational analysis of the promoter for a spinach chloroplast transfer RNA gene. EMBO J. 1985; 4:1637–1644.

[28] Reha-Krantz LJ, R L Nonay RL. Genetic and biochemical studies of bacteriophage T4 DNA polymerase 3'-->5'-exonuclease activity. J Biol Chem. 1993; 268:27100-27108.

[29] Mizrahi V, Benkovic P, Benkovic SJ. Mechanism of DNA polymerase I: exonuclease/ polymerase activity switch and DNA sequence dependence of pyrophosphorolysis and misincorporation reactions. Proc Natl Acad Sci, (USA). 1986; 83:5769-5773.

[30] Palanivelu P. An Insight into the Active Sites of the Catalytic Basic Protein Subunit PB1 of the RNA Polymerase of Human Influenza Viruses, World J Adv Res Rev. 17: 625–565 (2022).

[31] Derbyshire V, Freemont PS, Sanderson MR, Beese L, Friedman JM, Joyce CM, Steitz TA. Genetic and crystallographic studies of the 3,5"-exonucleolyitc site of DNA polymerase I. Science. 1988;240: 199-201.

[32] Derbyshire V, Grindley ND, Joyce CM. The 3'-5' exonuclease of DNA polymerase I of Escherichia coli: contribution of each amino acid at the active site to the reaction. EMBO J. 1991; 10:17-24.

[33] Joyce CM, Steitz TA. DNA polymerase I: From crystal structure to function via genetics. Trends Biochem Sci. 1987;12: 288-292.

[34] Wang F, Yang W. Structural insight into translesion synthesis by DNA Pol II. Cell. 2009; 139:1279–1289.

[35] Pinto MN, ter Beek J, Ekanger LA, Erik Johansson E, Barton JK. The [4Fe4S] Cluster of Yeast DNA Polymerase ε Is Redox Active and Can Undergo DNA-Mediated Signaling. J Am Chem Soc. https://doi.org/10.1021/jacs.1c07150

[36] Jain R, Rajashankar KR. Buku A, Johnson RE, Prakash L, Prakash S, et al. Crystal structure of yeast DNA polymerase epsilon catalytic domain. PLoS ONE. 2014; 9:e94835. doi: 10.1371/journal.pone.0094835.

[37] Fijalkowska IJ, Schaaper RM. Mutants in the Exo I motif of Escherichia coli dnaQ: Defective proofreading and in viability due to error catastrophe. Proc Natl Acad Sci. (USA). 1996; 93:2856-2861.

[38] Hamdan S, Carr PD, Brown SE, Ollis DL, Dixon NE. Structural Basis for Proofreading during Replication of the Escherichia coli Chromosome. Structure. 2002; 10:535–546.

[39] Zuo Y, Zheng H, Wang Y, Chruszcz M, Cymborowski M, Skarina T, Savchenko A, Malhotra A, Minor W. Crystal Structure of RNase T, an exoribonuclease involved in tRNA maturation and end-turnover. Structure. 2007; 15:417–428.

[40] Maksimova TG, Mustayev AA, Zaychikov EF, Lyakhov DL, Tunitskaya VL, Akbarov AK, Luchin SV, et al. Lys631 residue in the active site of the bacteriophage T7 RNA polymerase. Affinity labeling and site-directed mutagenesis. Eur J Biochem. 1991; 195:841-847.

[41] Pavlov YI, Maki S, Maki H, Kunkel TA. Evidence for interplay among yeast replicative DNA polymerases alpha, delta and epsilon from studies of exonuclease and polymerase active site mutations. BMC Biol. 2004; 2:1-13.