

eISSN: 2581-9615 CODEN (USA): WJARAI Cross Ref DOI: 10.30574/wjarr Journal homepage: https://wjarr.com/

	Ware World Journal of Advanced Research and Reviews	UNDER JOHN HAUH
		World Journal Series INDIA
Che	ck for up	dates

(RESEARCH ARTICLE)

# Polymerase and Proofreading Exonuclease Domains of the Nuclear-encoded DNAdependent RNA Polymerase of Plant Mitochondria

Peramachi Palanivelu \*

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

World Journal of Advanced Research and Reviews, 2023, 19(02), 989-1004

Publication history: Received on 08 July 2023; revised on 18 August 2023; accepted on 21 August 2023

Article DOI: https://doi.org/10.30574/wjarr.2023.19.2.1670

# Abstract

Mitochondria, found in all eukaryotic cells, play a crucial role in generating much needed biological energy for the cells in the form of adenosine triphosphates (ATPs). It is a semi-autonomous organelle and is partly controlled by its own genome and mostly by the nuclear imports. To replicate its own genome, it uses two DNA polymerases, viz. polymerases IA and IB which are essentially similar to the *E. coli* DNA polymerase I. The nuclear-encoded RNA polymerase (NEP) (EC 2.7.7.6) is imported from the nucleus and involves in the transcription of all mitochondrial genes. In *Arabidopsis thaliana*, the mitochondrial NEP showed 59.05% identity to the NEP of the chloroplasts, but only 28.24% identity to the T7 RNA polymerase, suggesting the NEPs of mitochondria and chloroplasts are distinctly different. However, in both the plant NEPs, the polymerase catalytic core and proofreading (PR) exonuclease domains are completely conserved. The mitochondrial NEP's catalytic core from different plant sources is remarkably conserved and is in close agreement with other RNA polymerases reported already and possesses a typical template-binding pair (-YG-), a basic catalytic metal-binding motifs are identified based on sequence similarity and site-directed mutagenesis (SDM) experiments. The PR exonuclease of the mitochondrial NEP belongs to the DEDD-superfamily of exonucleases.

**Keywords:** Mitochondrial transcription; Nuclear-encoded RNA polymerase; RNA polymerase active site; Proofreading exonuclease; Exonuclease active site; Mechanism of action

# 1 Introduction

Mitochondria are double-membrane structures found in all eukaryotic cells and harbour the mitochondrial genomic DNAs (mtDNAs) and mitochondrial plasmids. The mitochondria are of sizes ranging from 0.5 to 4  $\mu$ m and usually about dozens of mtDNA copies are found in a single mitochondrion and furthermore, as a single eukaryotic cell harbours 100s of mitochondria, the total number of copies of mtDNA exceeds > 1000 copies per cell. The copy number and shape of mitochondria vary considerably in different cell types and under different physiological conditions. They generate the much needed chemical energy for cellular activities in the form of ATPs by the process of aerobic respiration and hence, known as the power house of the cells. Thus, they are essential organelles found in all eukaryotes from yeasts to animals and plants.

Plant mitochondrial genomes are large and complex as compared to animal mitochondrial genomes. For example, nearly all animal mitochondrial genomes are ~16.5 kb in length with no or a few introns and possess little non-coding DNA, whereas plant mitochondrial genomes are quite large and their sizes range from 200 to 2500 kb. For example, the model plant *Arabidopsis thaliana*'s mitochondrial genome is of 3, 66,924 bp and codes for 33 proteins, 3 rRNAs and 21 tRNAs [1]. It is interesting to note that even though their genomes are much bigger in size, they do not contain significantly more genes than their animal counterparts. The larger size of plant mitochondrial genomes is mainly due to large numbers of introns, repeats and non-coding regions in their genomes. Using mostly nuclear-encoded proteins,

Copyright © 2023 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution Liscense 4.0.

<sup>\*</sup> Corresponding author: Peramachi Palanivelu

mitochondria contain all of the machinery for transcribing and translating genes encoded by the mtDNA. Unlike the nucleus, a single mitochondrion possesses many copies of its own genome. For example, depending on the tissue type and age of the plant, the number of copies per mitochondrion can vary between 50 and 500 [2].

Replication of mtDNA utilizes a bacterial-like, single subunit DNA polymerases, viz. IA and IB. Thus, most plants possess two nuclear-encoded copies of the DNA polymerases, IA and IB. It is interesting to note that in all plants where localization has been examined, these enzymes are dual targeted to both mitochondria and chloroplasts [3]. Structural, phylogenetic and multiple sequence alignment (MSA) analysis of plant mitochondrial and chloroplast DNA polymerases reveal a similarity to bacterial DNA polymerase I [4,5]. mtDNAs play an important role in cytoplasmic male sterility and this property is attributed mainly to mtDNA rearrangements, causing plants to lose their ability to make pollen, a trait which is exploited by crop breeders for hybrid generation.

# 1.1 Transcription by Plant Mitochondrial RNA Polymerase

RNA polymerases are crucial enzymes of life as they control the gene expression at the transcription level. Therefore, understanding the structure, function, mechanism and regulation of RNA 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 and plants. [6]. Whereas the DNA polymerases essentially involve in the replication and repair of genomes, the RNA polymerases are involved in transcription of genomes, i.e., involved in the flow of genetic information from  $DNA \rightarrow RNA$  in both prokaryotes and eukaryotes.

Different types of RNA polymerases perform transcription in prokaryotes and eukaryotes. For example, viruses contain two types of RNA polymerases (RNAPs), viz. DNA-dependent RNAPs (DdRps) and RNA-dependent RNAPs (RdRps), which are single-subunit (SSU) types. Eubacteria and archaebacteria employ a single type of RNA polymerase for all their transcription needs, but it is of a multi-subunit (MSU) type. However, eukaryotes use 5 different types of RNA polymerases (I-V), which are also of MSU type [6]. The chloroplast transcription in higher plants is performed by two types of RNAPs, 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. However, the mitochondria use only one type of RNA polymerase that is also nuclear-encoded (NEP). Both the plant NEPs are SSU enzymes which are of T7 bacteriophage-type. [7]. Three different subtypes of NEPs are reported from mitochondria and chloroplasts in plants, viz. RPOT-1, RPOT-2 and RPOT-3. The RPOT-1 and RPOT-3 are imported into mitochondria and chloroplasts, respectively, whereas the RPOT-2 has dual targeting properties and transcribes genes in both the organelles [8].

# 1.2 PR Functions in RNA Polymerases

Both the DNA and RNA polymerases do make mistakes during the replication/transcription processes, but rarely [9]. However, the mistakes are corrected promptly by the PR enzymes associated with or part of these polymerases. 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 in biological systems. There are two types of DEDD-superfamilies: One type of DEDD-superfamily is intrinsic type and is found in the same DNA/RNA polymerase polypeptide and functions as a multifunctional enzyme (e.g.), *E. coli* DNA pol I. In the second type of DEDD-superfamily is extrinsic type and the exonuclease function is performed by a tightly associated subunit along with the polymerase subunits in a multienzyme complex system (e.g.),  $\varepsilon$ -subunits of bacterial DNA pols III, ExoNs in the SARS-Coronaviruses, in the PA subunits of the polymerase in human influenza viruses and T4 DNA polymerase [7,10-13]. Again, the DEDD-superfamily of PR exonuclease 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 [14].

These two superfamilies, DEDD and PHP are invariably found/associated with the DNA/RNA replicases and transcriptases to repair any error that might occur during the replication/transcription processes [15]. However, the PHP-superfamily is not that common and has been reported only from the bacterial kingdom, but also has been reported from the viral kingdom by Palanivelu, recently [11]. The PHP-superfamily is found in the bacterial replicative DNA polymerases III (intrinsic) (the 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. Whereas the A-family polymerases possess two exonuclease domains  $(3\rightarrow5' \text{ and } 5'\rightarrow3' \text{ exonucleases})$ , the B-family polymerases are mainly involved in genome replication and repair in biological systems.

The error rates of RNAPs are very minimal and generally in the range of 10<sup>-4</sup> to 10<sup>-6</sup> [7] and are corrected promptly during the transcription process itself. For error corrections, the RNAPs have evolved essentially two different mechanisms (Table 1). In the MSU-RNAPs of prokaryotes, eukaryotes, and chloroplasts, which are involved mainly in the transcription of mRNAs, the PR active site is embedded within the polymerase active site itself [6,9]. However, in other SSU RNAPs, 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) as discussed earlier. The mitochondrial mRNAs have no 5' caps but have poly–A tails. Table 1 shows the types of PR activities and their localization in viruses, prokaryotes, eukaryotes and organelles.

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

RNA/DNA P	ols	Source(s)	PR active site	Reference(s)
Dd <b>R</b> ps- <mark>MSU</mark> - Dd <b>R</b> ps- <mark>MSU</mark> -	-RNAP -RNAPs	Prokaryotes Eukaryotes	Intrinsic within the RNAP active site itself	[6,9]
Dd <b>R</b> ps- <mark>MSU</mark>	-RNAPs	All eukaryotic (pols I-III)	Intrinsic within the RNAP active site itself	[6,9]
Dd <b>R</b> ps- <mark>MSU</mark>	-RNAPs	Plants (pols IV & V)	Intrinsic within the RNAP active site itself	[6,9]
Rd <b>R</b> ps- SSU Types	(+) Str (e.g., S SARS Humai	and RNA Virus SARS-CoVs, -related CoVs, 1-CoVs)	DEDD( <b>H</b> )-superfamily# (In the associated ExoN subunit)	[10]
Rd <b>R</b> ps- SSU Types	<mark>(-) Str</mark> (e.g., H Viruse	<mark>and RNA Viruso</mark> Human Influenza s A, B & C)	DEDD( <b>H</b> )- superfamily# (In the PA subunit of the RNA p	[11] olymerase)
Dd <b>D</b> ps- SSU Type	<b>T4 DN</b> ( <i>E. col</i>	l <b>A pol</b> <i>i</i> Phage)	DEDD(Y) superfamily^ (In the same polypeptide)	[12]
Dd <b>D</b> ps- SSU Type	E. coli	DNA pol I	DEDD(Y) superfamily^ (In the same polypeptide)	[13]
	<u>Plant</u>	(Chloroplasts)		
Dd <b>D</b> ps- SSU Types	<b>DNA</b> J (Plant	pols IA & IB chloroplasts)	DEDD(Y) superfamily^ (In the same polypeptide)	[5]
Dd <b>R</b> ps- SSU Type	( <b>NEP</b> ) (Plant	chloroplasts)	DEDD( <b>H</b> ) superfamily# (In the same polypeptide)	[5]
	<u>Plant</u>	(Mitochondria	<u>)</u>	
Dd <b>R</b> ps- SSU Type	RNA I (Plant	Pol ( <b>NEP</b> ) Mitochondria)	DEDD( <b>H</b> ) superfamily# (In the same polypeptide, MFE)	This work

Pols, Polymerases. <sup>^</sup>E. coli DNA pol I and pol II types; # Similar to the ε-subunit of bacterial DNA pols III (DNA replicases).

The polymerase and PR exonuclease domains in the DNA polymerases IA, IB and Nuclear-Encoded RNA Polymerase (NEP) of the plant chloroplasts and the MSU eubacterial type PEP are already discussed in detail by Palanivelu [5 and references therein], and, therefore, the polymerase and PR exonuclease domains of the NEP of plant mitochondria are analyzed and reported in this communication.

# 2 Material and methods

The protein sequence data of the NEPs from various plant mitochondria 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 sequence similarities, site-directed mutagenesis (SDM) and X-ray crystallographic data from other DNA and RNA polymerases, already reported.

#### 3 Results and discussion

#### **3.1 Analysis of Mitochondrial NEPs**

Figure 1 shows the MSA of the mitochondrial NEPs from various plant sources. (Only the regions required for the discussions are shown here). The A. thaliana sequence is used as the reference and highlighted in vellow. The N-terminal region of  $\sim 150$  amino acids is not conserved and showed many gaps in the alignment (not shown). 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-superfamily active site amino acids and is highlighted in light blue. The first conserved triad DV/I/TE of the DEDD-superfamily of exonucleases is seen after  $\sim$ 500 amino acids from the Nterminal. Again, after  $\sim$ 745 amino acids, a second demarcation in the sequences is seen and that contains the polymerase domain and the active site amino acids are highlighted in vellow. The polymerase core region is completely conserved in all NEPs from different plant sources and contains the template-binding -YG- pair, the catalytic amino acid -746VDR-K and the nucleotide discriminating amino acid R at -4 from the catalytic K. The active site, <sup>4</sup>KLVK<sup>752</sup>Q<sup>1</sup>TVMTSVY<sup>8</sup>GVT-, is found to be very similar to the confirmed active site of *E. coli* DNA pol I, -OR-<sup>4</sup>RSAK<sup>758</sup>A<sup>1</sup>INFGLIY<sup>8</sup>GM- [13] and in close agreement to the active sites of the other DNA/RNA polymerases already reported (Table 2) [7]. The catalytic metal-binding Ds in -HQD<sup>677</sup>- and -HD<sup>909</sup>S- motifs are highlighted in dark green. Site-directed mutagenesis data have shown that the substitutions of the Ds in  $-AFD^{537}G$ - and  $-HD^{812}S$  motifs to  $D^{537} \rightarrow N$ and  $D^{812} \rightarrow N$  resulted in the complete loss of T7 RNAP activity [17.18]. An additional -DxD- type metal-binding motif is found in this domain and highlighted in light green. All the RNA polymerases invariably end in the tetrapeptide, -YFFNand its significance is not clear now. Interestingly, a highly conserved decapeptide, -IPYL/I/VPML/I/VPP-, is found in all the mitochondrial NEPS at the N-terminal region (highlighted) and its significance is not clear now.

#### 3.1.1 CLUSTAL O (1.2.4) MSA of plant mitochondrial RNA polymerases (NEPs)

tr A0A816ZBA3 A0A816ZBA3 BRANA	QSSKNVTLENT-KLSRRYGCIECDPLIRKGLDKSARHMV <mark>IPYLPMLIPP</mark> RNWTGYDQGAH	405
sp P92969 RPOT1 ARATH	QNFRTVTLENT-KTSRRYGCIECDPLVLKGLDKSARHMVIPYLPMLIPPQNWTGYDQGAH	422
tr S8EG02 S8EG02_9LAMI	HTSRTVCREKRQTRRYGVVQCDPLVHKGLEKAARHMV <mark>IPYMPMLVPP</mark> VNWTGYDKGGY	412
tr A0A0Q3EKS1 A0A0Q3EKS1 BRADI	HEMRTVTREQQ-KSSRRYGVIKCDPLIRQGLDRTAKHMVIPYMPMLIPPISWTGYDKGAH	438
tr A0A8S0QEE5 A0A8S0QEE5_OLEEU	HSLKTVETQ-RGSRRYGVIECDPLVRKGLERTARHMV <mark>IPYMPMLVPPLNWTGYDRGAY</mark>	424
tr A0A0V0IXS1 A0A0V0IXS1 SOLCH	HTLKTVETV-KGSRRYGVIQCDPLVRKGLDKTARHMV <mark>I</mark> PYMPMLVPPQNWSGYDRGAY	439
NP 001312318.1	HTLKTVETM-KGSRRYGVIQCDPLVRKGLDKTARHMV <mark>IPYMPMLVPP</mark> SWLGYDKGGY	448
NP 001289502.1	HTLKTVETM-KGSRRYGVIQCDPLVRKGLDKTARHMV <mark>I</mark> PYMPMLVPPQSWLGYDKGAY	448
tr A0A565CQ27 A0A565CQ27 9BRAS	HKFKTVTKYPGSRMVRRYGVIECDSLLLAGLDKSAKHMLIPYVPMLVPPKRWKGYDKGGY	445
tr A0A087GS25 A0A087GS25_ARAAL	HKFKTVTKYPGSKLVRRYGVIECDSLLLAGLDKSAKHMLIPYVPMLVPPKRWKGYDKGGY	426
tr ROHH96 ROHH96 9BRAS	HKFKTVTKYPGSKLVRRYGVIECDSLLLAGLDKSAKHMIIPYVPMLVPPKRWKGYDKGGY	437
tr A0A3P6GE78 A0A3P6GE78 BRAOL	HKFKTVTKYPGSKLVRRYGVIECDSLLLAGLDKSAKYMLIPYVPMLVPPKRWRGYDKGGY	421
tr A0A078HME9 A0A078HME9_BRANA	HKFKTVTKYPGSKMVRRYGVIECDSLLLAGLDKSAKHMLIPYVPMLVPPKRWKGYDKGGY	419
tr A0A8S9IIZ9 A0A8S9IIZ9 BRACR	HKFKTVTKYPGSKLVRRYGVIECDSLLLAGLDKSAKHMLIPYVPMLVPPKRWKGYDKGGY	422
tr A0A2G2WAZ7 A0A2G2WAZ7 CAPBA	HVFRIATRDPRKNIVKKYGVIECDPLVIVGVDRTVKQMMIPYVPMLVPPKKWRGYDKGGY	425
tr A0A6P6SN94 A0A6P6SN94 <sup>-</sup> COFAR	HTFKIATNEAGQNPVKRYGVIVCDPVIHSGLDRSVNHMMVPYVPMLVPPKKWKGYDKGGY	476
tr A0A445AVH3 A0A445AVH3_ARAHY	HRFKAISKNPGEKILKNYGIIECDPLVLAGLDKSAHHML <mark>M</mark> PYMPMLIPPKKWKGYEKGGH	440
tr A0A445END8 A0A445END8 ARAHY	HRFKAISKNPGEKILKNYGIIECDPLVLTGLDKSAHHML <mark>M</mark> PYMPMLIPPKKWKGYEKGGH	441
tr F6HYL3 F6HYL3 VITVI -	HTFKTVTKESGQKDWRKYGVIECDPLVLIGLDRTAKHMV <mark>I</mark> PYVPMLIPPKKWKGYDKGGY	425
tr A0A5N6RVV2 A0A5N6RVV2 9ROSI	HTFKPVT-NPGQKIVKKYGVIECDPLVLMGLDKTAKHMLIPYVPMLVPPKKWKGYDKGGH	426
tr A0A6J1KHL7 A0A6J1KHL7 <sup>-</sup> CUCMA	HSFKPVAKTPGQKLVKKYGVIECDPIVLAGLDRTAKQMLIPYIPMLVPPKKWKGYDKGGH	422
tr A0A6J1E061 A0A6J1E061 MOMCH	HRFKPVTKTPGQKIVKTYGVIECDPIVLAGLDRTAKHVLIPYIPMLVPPKKWKGYDKGGH	418
tr A0A0A0LH38 A0A0A0LH38 <sup>-</sup> CUCSA	HGFKPVAKTPGQNFVKKYGVVECDPIVLAGLDRTAKHML <mark>I</mark> PYIPMLVPPKKWKGYDKGGH	422
tr A0A1S3BU26 A0A1S3BU26 CUCME	HGFKPVAKTPGQNFVKKYGVIECDPIVLAGLDQTAKHMLIPYIPMLVPPKKWKGYDKGGH	422
tr A0A8T3B2K3 A0A8T3B2K3 DENNO	HILKAVSVEDG-KLKNKYGVIECHPLVHMGLGSTARLMV <mark>I</mark> PYLPMLIPPKKWKGYNKGGH	431
tr A0A3B6RRH2 A0A3B6RRH2 WHEAT	HVLRKPIVENG-RLKKKHFVIECDPLVHEGFESTARHVE <mark>I</mark> PYLPMLVPPTKWKGYDKGGH	395
tr A0A8I6Z6W4 A0A8I6Z6W4 HORVV	HVLRKPIVENG-RLKKKHFVIECDPLVHEGFESTARHVE <mark>IPYLPMLVPP</mark> TKWKGYDKGGH	394
tr A0A3L6DFV0 A0A3L6DFV0 MAIZE	HVLRQPIIENG-RLKKKHWVIECDHLVHEGFESTARHVD <mark>I</mark> PYLPMLVPPKKWKGYDKGGH	398
tr A0A921U2C4 A0A921U2C4_SORBI	HVLRQPIIENG-RLKKKHWVIECDHLVHDGFESTARHVD <mark>IPYLPMLVPP</mark> KKWKGYNKGGH	405
tr A0A3L6PDV3 A0A3L6PDV3 PANMI	HVLRQPIIENG-RLKKKHWVIECDHLVHEGFESTARHVD <mark>IPYLPMLVPP</mark> KKWKGYDKGGH	396
tr A0A2S3HHG5 A0A2S3HHG5_9POAL	HVLRQPIIENG-RLKKKHWVIECDHLVHEGFESTARHVD <mark>IPYLPMLVPP</mark> KKWKGYDKGGH	400
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	HVLRQPIIENG-RLKKKHWVIECDHLVHEGFESTARHVD <mark>IPYLPMLVPP</mark> KKWKGYDKGGH	396
tr A0A6G1CRT5 A0A6G1CRT5_9ORYZ	HILRQPIIENG-RLKKKHWVIECDPLVHEGFESTARHVE <mark>I</mark> PYLPMLVPPKKWKGYDTGGY	256
tr A0A8J5SZQ4 A0A8J5SZQ4_ZIZPA	HVLRQPIVENG-RLKKKHWVIECDHLVHEGFESTARHVE <mark>IPYLPMLIPP</mark> KKWKSYDTGGY	394
_	: : : : *. :: *. : : :**:*********	

		PR exonuclease	
tr A0A816ZBA3 A0A816ZBA3_BRANA	FFLP	SYLMRTNGSKQQRIAVNKTPKAQLEPVFKALDTLGNTKWRINKKVLSLVDRIWANG	465
sp P92969 RPOT1_ARATH	FFLP	SYVMRTHGAKQQRTVMKRTPKEQLEPVYEALDTLGNTKWKINKKVLSLVDRIWANG	482
tr S8EG02 S8EG02_9LAMI	LFLP	SQVMRTHGAKQQREAVKRVPREQLYPVYEALDTLGTTGWRVNKRVFSIVERIWGSG	472
tr A0A0Q3EKS1 A0A0Q3EKS1_BRADI	LFLP	SYVMRTHGARQQREAVKKAPKEQMQLIFEALDTLGSTKWRINKKVLSIVDRIWSSG	498
tr A0A8S0QEE5 A0A8S0QEE5_OLEEU	LFLP	SYIMRTHGAKQQREAVKRVPRKQLEPVFEALNTLGTTKWKVNKRILAVIDRIWASG	484
tr A0A0V0IXS1 A0A0V0IXS1_SOLCH	LFLP	SYIMRTHGAKQQREAVKRVPKKQLEPVFQALDTLGNTKWRVNKRVLGILDRIWASG	499
NP_001312318.1	LFLP	SYIMRTHGAKQQREAVKRVPKKQLEPVFQALDTLGNTKWRVNRKVLGIVNRIWASG	508
NP_001289502.1	LFLP	SYIMRTHGAKQQREAVKRVPKKQLEPVFQALDTLGNTKWRLNRKVLGIVDRIWASG	508
tr A0A565CQ27 A0A565CQ27 9BRAS	LFLP	SYIMRTHGSKKQQDALRDVSHKTAHRVFEALDTLGNTKWRVNRKILDVVERLWADG	505
tr A0A087GS25 A0A087GS25_ARAAL	LFLP	SYIMRTHGSKKQQDALRDVSHKTAHRVFEALDTLGNTKWRVNRNILDVVERLWADG	486
tr ROHH96 ROHH96 9BRAS	LFLP	SYIMRTHGSKKQQDALRDISSKTAHRVFEALDTLGNTKWRVNRKILDVVERLWADG	497
tr A0A3P6GE78 A0A3P6GE78_BRAOL	LFLP	SYIMRTHGSKKQQDALKDISSKTAHRVFEALDTLGNTKWRVNRKILDVVERLWADG	481
tr A0A078HME9 A0A078HME9 BRANA	LFLP	SYIMRTHGSKKQQDALKDISSKTAHRVFEALDTLGNTKWRVNRKILDVVERLWADG	479
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	LFLP	SYIMRTHGSKKQQDALKDISSKTAHRVFEALDTLGNTKWRVNRKILDVVERLWADG	482
tr A0A2G2WAZ7 A0A2G2WAZ7 CAPBA	LFLP	SYLMRTHGSRRQQDAVRSVSGKQMQQVYEALDTLGSTKWRVNKRILSVVESIWSGG	485
tr A0A6P6SN94 A0A6P6SN94 COFAR	LFLP	SYLMRTHGSRQQQDAIKCAPVKQMQKVYEALDTLGNTKWRVNRRILNVVETVWGGG	536
tr A0A445AVH3 A0A445AVH3_ARAHY	LFLP	SYIMRTHGSKKQQDTMKNVKRTQMQKVFEALDVLGSTKWRINRRILAVVEAVWAGG	500
tr A0A445END8 A0A445END8 ARAHY	LFLP	SYIMRTHGSKKQQDTLKNVKRTQMQKVFEALDVLGSTKWRINRRILAVVEAVWAGG	501
tr F6HYL3 F6HYL3_VITVI	LFLP	SYVMRTHGSRKQQDAVKSVPRRQLQKVFEALDTLGNTKWRINRRVLSVVESIWARG	485
tr A0A5N6RVV2 A0A5N6RVV2 9ROSI	LFLP	SYVMRTHGSRKQQDAVKIVPAKQMQKVFEALDVLGNTKWRVNRRVLNIVECIWARG	486
tr A0A6J1KHL7 A0A6J1KHL7 CUCMA	LFLP	SYVMRTHGSIRQQDAMKNISGKQMQKVFEALDMLGSTKWRVNRRVLSVVESIWSQG	482
tr A0A6J1E061 A0A6J1E061 MOMCH	LFLP	SYVMRTHGSSRQQDAMKNISGKQMQKVFEALDMLGSTKWRVNRRVLSVVESIWAQG	478
tr A0A0A0LH38 A0A0A0LH38 CUCSA	FFLP	SYVMRTHGSSRQQDAMKNISGKQMQKVFEALDMLGSTKWRVNRSVLSVVESLWAQG	482
tr A0A1S3BU26 A0A1S3BU26 CUCME	FFLP	SYVMRTHGSSRQQDAMKNISGKQMQKVFEALDMLGSTKWRVNRRVLSVVESLWAQG	482
tr A0A8T3B2K3 A0A8T3B2K3 DENNO	LFLP	SYVMRTHGAKDQQNAIKSVPKRQLNKVFEALDTLGSTKWRVNRKILQVVETIWSEG	491
tr A0A3B6RRH2 A0A3B6RRH2 WHEAT	LFLP	SYVMRTHGVKDQKEAIKSVPRKQLRKVFEALDILGGTKWRVNRRVHDVVETIWSRG	455
tr A0A81626W4 A0A81626W4_HORVV	LFLP	SYVMRTHGVKDQKEAIKSVPRKQLRKVFEALDILGGTKWRVNRRVHDVVETIWSRG	454
tr A0A3L6DFV0 A0A3L6DFV0 MAIZE	LFLP	SYIMRTHGVKDQKDAINSVPRKQLRKVFEALDILGSTKWRVNRRVHDVVETIWSQG	458
tr A0A921U2C4 A0A921U2C4 SORBI	LFLP	SYIMRTHGVKDQKDAINSVPRKQLRKVFEALDILGSTKWRVNRRVHDVVETIWSRG	465
tr A0A3L6PDV3 A0A3L6PDV3 PANMI	LFLP	SYIMRTHGVKDQKDAIKSVPRKQLRKVFEALDILGSTKWRVNRRVHDVVETIWSRG	456
tr A0A2S3HHG5 A0A2S3HHG5 9POAL	LFLP	SYIMRTHGVKDQKDAIKSVPRKQLRKVFEALDILGSTKWRVNRRVHDVVETIWSRG	460
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	LFLP	SYIMRTHGVKDQKDAIKSVPRKQLRKVFEALDILGSTKWRVNRRVHDVVETIWSRG	456
tr A0A6G1CRT5 A0A6G1CRT5 90RYZ	LFLP	SYIMRTHGVKDQKEAIKSVPRKQLRKVFEALDTLGSTKWRVNRRVHDAVETVWSRG	316
tr A0A8J5SZQ4 A0A8J5SZQ4 ZIZPA	LFLP	SYIMRTHGVKDQKDAIKSVPRKQLRKVFEALDILGSTKWRVNRRVHDVVETIWSRG	454
_	:***	* :***:* *: .:. :::**: ** * *::*: : :::*. *	

+		5.0.5
tr AUA8162BA3 AUA8162BA3 BRANA	GRLGGLVDRDDVPIPEEPDGEDQEELKQWKWKFKEANKENSERHSQRC <mark>DVE</mark> LKLEVARKM	525
spip92969 RPOTI_ARATH		542
trisseguzisseguz_9LAMI		33Z
triauaugseksijauaugseksi_BRADI		558
tr AUA8SUQEES AUA8SUQEES_OLEEU	GQLADLVDREDVPLPEEPDTEDEAEIRKWKWKVKAAKKENSERHSQRCDIELKLAVARKM	544
tr AUAUVUIXSI AUAUVUIXSI_SOLCH	GRLADLVDREDVPLPEEPDTEDEEEIKKWKWKVKAAKKENCERHSQRCDIELKLAVARKM	559
NP_001312318.1	GRLADLVDREDVPLPEAPDTEDEAEIRKWKWKVKGVKKENCERHSQRCDIELKLAVARKM	568
NP_001289502.1	GRLADLVDREDVPLPEEPDAEDEAQIRKWKWKVKGVKKENCERHSQRCDIELKLAVARKM	568
tr A0A565CQ27 A0A565CQ27_9BRAS	GDIAGLVNREDVPIPEKPSSEDPEEIQSWKWSVRKANKINRERHSLRC <mark>DVE</mark> LKLSVARKM	565
tr A0A087GS25 A0A087GS25_ARAAL	GNIAGLVNREDVPIPEKPSSEDPEEIQSWKWSARKANKINRERHSLRO <mark>DVE</mark> LKLSVARKM	546
tr R0HH96 R0HH96_9BRAS	GNIAGLVNREDVPIPEKPSSEDPEELQSWKWSVRKANKINRERHSLRC <mark>DVE</mark> LKLSVARKM	557
tr A0A3P6GE78 A0A3P6GE78_BRAOL	GNIAGLVNREDVPIPEKPSSEDPEEIQSWKWSVRKAKKTNRERHSLRC <mark>DVE</mark> LKLSVARKM	541
tr A0A078HME9 A0A078HME9_BRANA	GNIAGLVNREDVPIPEKPLSEDPEEIQTWKWSVRKAKKINRERHSLRC <mark>DVE</mark> LKLSVARKM	539
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	GNIAGLVNREDVPIPEKPLSEDPEEIQLWKWSVRKAKKINRERHSLRC <mark>DVE</mark> LKLSVARKM	542
tr A0A2G2WAZ7 A0A2G2WAZ7_CAPBA	GNIAGLVDRNDVPIPDLHS-DDIMEVKRWKWRVRKAKKINQELHSQRC <mark>D<mark>T</mark>ELKLSVARKL</mark>	544
tr A0A6P6SN94 A0A6P6SN94_COFAR	GNIAGLVNRNDVHIPELHS-DAAEEIKKWKWNMRKAKKINRERHSQRC <mark>DIE</mark> LKLSVAHKM	595
tr A0A445AVH3 A0A445AVH3_ARAHY	GNTAGLIDRKDVPIPARPPLDDPKQIQEWKWNVMKAKKINAERHSLRC <mark>DTE</mark> LKLSVARKM	560
tr A0A445END8 A0A445END8_ARAHY	GNTAGLIDRKDVPIPARPPVDDPKQIQEWKWNVMKAKKINAERHSLRC <mark>DTE</mark> LKLSVARKM	561
tr F6HYL3 F6HYL3_VITVI	GNLGGLVDRENVPLPEKPSTEDLTEIQNWKWNARKAKKINQERHSLRC <mark>DTE</mark> IKLSVARKM	545
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	GNIAGLVNREDVPIPEKPSSEDLTEIQEWKWNLRKANKINRERHSQRC <mark>DTE</mark> LKLSVAQKM	546
tr A0A6J1KHL7 A0A6J1KHL7_CUCMA	GNAAGLVDRKDVPIPEKPL-GDLTKWKRSMWKAKKINQELHSQRC <mark>DVE</mark> LKLSVARKM	538
tr A0A6J1E061 A0A6J1E061 MOMCH	GNTAGLVDRKDVPIPEQPL-GDLTEMQEWKWSVRKAKKINQELHSQRC <mark>DVE</mark> LKLSVARKM	537
tr A0A0A0LH38 A0A0A0LH38 CUCSA	GNTAGLVDRKDVPIPEKPL-GDLTEMQEWRWSMKKAKKINQELHSQRC <mark>DVE</mark> LKLSVARKM	541
tr A0A1S3BU26 A0A1S3BU26 CUCME	GNTAGLVDRKDVPIPEKQL-GDLTETQEWRWSMKKAKKINQELHSQRC <mark>DVE</mark> LKLSVARKM	541
tr A0A8T3B2K3 A0A8T3B2K3 DENNO	GGVAGLVDREDIPIQEKPGLEDSMEIKKWRWNARKAKKANSEMHAQRC <mark>DTE</mark> LKLSVARKM	551
tr A0A3B6RRH2 A0A3B6RRH2 WHEAT	GGIAGLVDKGNIPLPEQPETEDPDEIQKWKWSVKKTKKANRELHAERC <mark>DTE</mark> LKLSVARKM	515
tr A0A816Z6W4 A0A816Z6W4 HORVV	GGIAGLVDKENIPLPERPETEDPDEIQKWRWSVKKTKKTNRELHAERC <mark>DTE</mark> LKLSVARKM	514
tr A0A3L6DFV0 A0A3L6DFV0 MAIZE	GGIAGLVDKANIPLPERPESEDPDEMQKWKWSLKKAKKTNRELHAERC <mark>DTE</mark> LKLSVARKM	518
tr A0A921U2C4 A0A921U2C4 SORBI	GGIAGLVDKANIPLPERPESEDPDEMQKWKWSLKKAKKTNRELHAERC <mark>DTE</mark> LKLSVARKM	525
tr A0A3L6PDV3 A0A3L6PDV3 PANMI	GGIAGLVDKTNIPLPERPELEDPDEVQKWKWSLKKAKKTNRELHAERC <mark>DTE</mark> LKLSVARKM	516
tr A0A2S3HHG5 A0A2S3HHG5 9POAL	GGIAGLVDKTNIPLPERPELEDPDEVQKWKWSLKKAKKTNRELHAERC <mark>DTE</mark> LKLSVARKM	520
tr A0A2T7DVW7 A0A2T7DVW7 9POAL	GGIAGLVDKTNIPLPERPELEDPDEVQKWKWSLKKAKKTNRELHAERCDTELKLSVARKM	516
tr   A0A6G1CRT5   A0A6G1CRT5 90RYZ	GGIAGLVDKGNIPLPERPESDDPDEIQIWKWSLKKAKKTNRELHAERCDTELKLSVARKM	376
tr A0A8J5SZQ4 A0A8J5SZQ4 ZIZPA	GGIAGLVDKGNIPLPERPESEDPDEIQKWKWSLKKAKKTNRELHAERODTELKLSVARKM	514
	* *: . * * * *: ** <mark>* *</mark> :** **:*:	
1		

tr A0A816ZBA3 A0A816ZBA3 BRANA	KEEE <mark>GFYFPHNVDFRGRAYPMHPYLNHLGSDLCR</mark> GILEFCEGKPLGESGLRWLKVHIANL	585
sp P92969 RPOT1 ARATH	KDEEGFYYPHNVDFRGRAYPIHPYLNHLGSDLCRGILEFCEGKPLGKSGLRWLKIHIANL	602
tr S8EG02 S8EG02 9LAMI	KDEDGFFYPHNLDFRGRAYPMHPHLNHLGSDICEGILEFSEGRPLGKSGLHWLKIHLANL	592
tr A0A0Q3EKS1 A0A0Q3EKS1 BRADI	KDEVGFYYPHN <mark>LD</mark> FRGRAYPMHPYLNHLGS <mark>D</mark> LCEGVLEFSEGRPLGDSGLCWLKIHLANL	618
tr A0A8S0QEE5 A0A8S0QEE5_OLEEU	KDEDSFYYPHN <mark>LD</mark> FRGRAYPMHPYLN <mark>HLGSD</mark> LCEGILEFAEGRPLGKSGLRWLKIHLANL	604
tr A0A0V0IXS1 A0A0V0IXS1_SOLCH	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPYLN <mark>HL</mark> GS <mark>D</mark> LCEGILEFSEGRPLGKSGLRWLKIHLANV	619
NP_001312318.1	KDEDGFYYPHN <mark>LD</mark> FRGRAYPMHPYLN <mark>HL</mark> GS <mark>D</mark> LCEGILEFAEGRPLGTSGLRWLKIHLANV	628
NP_001289502.1	KDEDGFYYPHN <mark>LD</mark> FRGRAYPMHPYLN <mark>HL</mark> GS <mark>D</mark> LCEGILEFAEGRPLGKSGLRWLKIHLANV	628
tr A0A565CQ27 A0A565CQ27_9BRAS	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLNHLSS <mark>D</mark> LCEGTLEFAEGRPLGKSGLHWLKIHLANL	625
tr A0A087GS25 A0A087GS25_ARAAL	KDEEGFYYPHN <mark>LD</mark> FRGRAYPM PHLN <mark>HL</mark> SS <mark>D</mark> LCEGTLEFAEGRPLGKSGLHWLKIHLANL	606
tr R0HH96 R0HH96_9BRAS	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGTLEFAEGRPLGKSGLHWLKIHLANL	617
tr A0A3P6GE78 A0A3P6GE78_BRAOL	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGTLEFAEGRPLGKSGLYWLKIHLANL	601
tr   A0A078HME9   A0A078HME9 BRANA	KDEEGFYYPHN <mark>LD</mark> FRGRAYPM PHLN <mark>HLSSD</mark> LCEGTLEFAEGRPLGKSGLYWLKIHLANL	599
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGTLEFAEGRPLGKSGLYWLKIHLANL	602
tr A0A2G2WAZ7 A0A2G2WAZ7_CAPBA	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGILEFSEGRPLGKSGLRWLKIHLANL	604
tr A0A6P6SN94 A0A6P6SN94_COFAR	KDEDGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGLLEFAEGRPLGKSGVYWLKIHLANL	655
tr A0A445AVH3 A0A445AVH3_ARAHY	KDEEGFYYPHN <mark>LD</mark> FRGRAYPM SHLN <mark>HL</mark> SO <mark>D</mark> LCEGLLEFAEGKPLGKSGLQWLKIHLANL	620
tr A0A445END8 A0A445END8_ARAHY	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMISHLN <mark>HL</mark> SC <mark>D</mark> LCEGLLEFAEGKPLGKSGLQWLKIHLANL	621
tr F6HYL3 F6HYL3_VITVI	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGVLEFAEGRPLGKSGLRWLKIQLANL	605
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	RDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGILEFAEGRPLGKSGLHWLKIHLANL	606
tr A0A6J1KHL7 A0A6J1KHL7_CUCMA	KDEEGFYYPHN <mark>LD</mark> FRGRAYPM PHLN <mark>HL</mark> SS <mark>D</mark> LCEGVLEFAEGRPLGKSGLHWLKIHLANL	598
tr A0A6J1E061 A0A6J1E061_MOMCH	RDEEGFYYPHN <mark>LD</mark> FRGRAYPM PHLN <mark>HL</mark> SS <mark>D</mark> LCEGVLEFAEGRPLGKSGLHWLKIHLANL	597
tr A0A0A0LH38 A0A0A0LH38_CUCSA	KDEEGFYYPHN <mark>LD</mark> FRGRAYPMHPHLN <mark>HL</mark> SS <mark>D</mark> LCEGVLEFAEGRPLGKSGLHWLKIHLANL	601
tr A0A1S3BU26 A0A1S3BU26_CUCME	KDEEGFYYPHN <mark>VD</mark> FRGRAYPMHPHLN <mark>HLSSD</mark> LCEGVLEFAEGRPLGKSGLHWLKIHLANL	601
tr A0A8T3B2K3 A0A8T3B2K3_DENNO	QDEEGFYYPHN <mark>LD</mark> FRGRAYPM SHLN <mark>HL</mark> SS <mark>D</mark> LCEGILEFAEGRQLGKSGLCWLKIHLANI	611
tr A0A3B6RRH2 A0A3B6RRH2_WHEAT	REEDGFYYPHN <mark>LD</mark> FRGRAYPMPHLS <mark>HLGSD</mark> LCEGVLEYAEGRPLGKSGLRWLKIHLANK	575
tr A0A8I6Z6W4 A0A8I6Z6W4_HORVV	REEDGFYYPHN <mark>LD</mark> FRGRAYPMHPHLS <mark>HLGSD</mark> LCEGVLEYGEGRPLGKSGLRWLKIHLANK	574
tr A0A3L6DFV0 A0A3L6DFV0_MAIZE	REEDGFYYPHN <mark>LD</mark> FRGRAYPMPHLS <mark>HLGSD</mark> LCRGVLEYAEGRPLGKSGLCWLKIHLANK	578
tr A0A921U2C4 A0A921U2C4_SORBI	REEDGFYYPHN <mark>LD</mark> FRGRAYPMHPHLS <mark>HLGSD</mark> LCEGVLEYAEGRPLGKSGLCWLKIHLANK	585
tr A0A3L6PDV3 A0A3L6PDV3_PANMI	REEDGFYYPHN <mark>LD</mark> FRGRAYPM PHLS <mark>HLGSD</mark> LCEGVLEYAEGRPLGKSGLCWLKIHLANK	576
tr A0A2S3HHG5 A0A2S3HHG5_9POAL	REEDGFYYPHN <mark>LD</mark> FRGRAYPMHPHLS <mark>HLGSD</mark> LCEGVLEYAEGRPLGKSGLCWLKIHLANK	580
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	REEDGFYYPHN <mark>LD</mark> FRGRAYPMHPHLS <mark>HLGSD</mark> LCEGVLEYAEGRPLGKSGLCWLKIHLANK	576
tr A0A6G1CRT5 A0A6G1CRT5_90RYZ	REEDGFYYPHN <mark>LD</mark> FRGRAYPMHPHLS <mark>HLGSD</mark> LCEGVLEYAEGRPLGKSGLRWLKIHLANK	436
tr A0A8J5SZQ4 A0A8J5SZQ4_ZIZPA	REEDGFYYPHN <mark>IDFRGRAYPM</mark> HPHLS <mark>HLGSDLCE</mark> GVLEYAEGRSLGKSGLRWLKIHLANK	574
	* <mark>.****<mark>.*</mark>***********************</mark>	

	PR Exonuclease	Polymerase	
tr A0A816ZBA3 A0A816ZBA3_BRANA	YAGGVDKFANEDRVAFTESHLEDIFDSSDRPLEGKRWWLNAEDP	FQCLAACMNLSKALRS	645
sp P92969 RPOT1 ARATH	YAGGVDKLAYEDRIAFTESHLEDIFDSSDRPLEGKRWWLNAEDP	FQCLAACINLSEALRS	662
tr S8EG02 S8EG02_9LAMI	YANGIDKLSHEGRLAFTDNHLDDICDSADRPLEGKRWWLSAEDP	FQCLAACIDLAEALRC	652
tr A0A0Q3EKS1 A0A0Q3EKS1_BRADI	YAGGVDKLSYDGRIAFTENHLEDIFDSANRPLEGKRWWLGAEDP	FQCLAVCMDLTEALRS	678
tr A0A8S0QEE5 A0A8S0QEE5 OLEEU	YGGGVDKLSYEGRIAFAENHLEDIFDSADRPLEGKRWWLGAEDP	FQCLATCMNLSESLRS	664
tr A0A0V0IXS1 A0A0V0IXS1_SOLCH	YGGGVDKLSYEGRAAFSENHVEDIFDSADRPLEGRRWWLGAEDP	FQCLATCINISEALRS	679
NF_001312318.1	YGGGVDKLSYEGRVAFSENHLEDIFDSAERPLEGKRWWLGAEDP	FQCLATCINIAEALRS	688
NF_001289502.1	YGGGVDKLSYEGRVAFSENHVEDIFDSAERPLEGKRWWLGAEDP	FQCLATCINIAEALRS	688
tr A0A565C027 A0A565C027_9BRAS	YAGGVEKLSHDGRLAFVENHLDVIMDSAENPIHGKRWWLKAEDP	FQCLAACVILTQALKS	685
tr A0A087GS25 A0A087GS25_ARAAL	YAGGVEKLSHDGRLAFVENHLDVIIDSAENPIHGKRWWLKAEDP	FQCLAACVILTQALKS	666
tr ROHH96 ROHH96 9BRAS	YAGGVEKLSHDARLAFVETHLDDVMDSAENPIHGKRWWLKAEDP	FQCLAACVILTQALKS	677
tr A0A3P6GE78 A0A3P6GE78_BRAOL	YAGGVEKLSHEGRLAFVENHLDDIMDSAENAIHGRRWWLKAEDP	FQCLAACVVLAQALKS	661
tr A0A078HME9 A0A078HME9_BRANA	YAGGVEKLSHDGRLAFVENHLDDIIDSAENAVHGKRWWLKAEDP	FQCLAACVILAQALKS	659
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	YAGGVEKLSHDGRLAFVENHLDDIIESAENAVHGKRWWLKAEDP	FQCLAACIILAQALKS	662
tr A0A2G2WAZ7 A0A2G2WAZ7 CAPBA	YAGGIEKRCYDARIAFIENHLDDILDSAHNPLNGNKWWLNAEDP	FQCLAACINLSEALKS	664
tr A0A6P6SN94 A0A6P6SN94 COFAR	YGGVVKKLSYDQRLAFVENHLNEVLDSADNPLNGNRWWLKAEDP	FQCLAACMNLSEALKS	715
tr A0A445AVH3 A0A445AVH3_ARAHY	YAGGVEKLSYDGRLTFVENHLHDIFDSADNPVNGNRWWLMAEDP	FQCLATCINLSEALRS	680
tr A0A445END8 A0A445END8_ARAHY	YAGGVEKLSYDGRLTFVENHLHDIFDSADNPVNGNRWWLMAEDP	FQCLATCINLSEALRS	681
tr F6HYL3 F6HYL3_VITVI	YAGGVEKLSYDGRLAFVDNHLDDVFDSADNPLNGNRWWLTAEDP	FQCLAACINLSEALRS	665
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	YSGGVEKLSHDGRLAFVDNHLRDIFDSADNPINGNRWWLTAEDP	FQFLAACINLSEALRS	666
tr A0A6J1KHL7 A0A6J1KHL7_CUCMA	YAGGVEKLSYDRRLAFVDDHLGSIFDSASNPVHGDRWWLAAEDP	FQCLAACINLSEALRS	658
tr A0A6J1E061 A0A6J1E061_MOMCH	YAGGVEKLSYDRRLAFVDDHLDSIFDSASNPVNGNRWWLTAEDP	FQCLAACINLSEALKS	657
tr A0A0A0LH38 A0A0A0LH38_CUCSA	YAGGVEKLSYDERLAFVDDHLDNIFDSASNPVNGDRWWLTAEDP	FQCLAACMNLSEALKS	661
tr A0A1S3BU26 A0A1S3BU26_CUCME	YAGGVEKLSYDRRLAFVDDHLDNIFDSASNPVNGDRWWLTAEDP	FQCLAACMNLSEALKS	661
tr A0A8T3B2K3 A0A8T3B2K3_DENNO	YGGGVEKLSYDARLTFVENHLNNIFDSAMNPVDGKRWWIHAEDP	FQCLAACINLSDALTS	671
tr A0A3B6RRH2 A0A3B6RRH2_WHEAT	YGGGIEKLSHESKLAFVEDHLPDIFDSASNPVDGNCWWINAEDP	FQCLAACMDLSKALES	635
tr A0A81626W4 A0A81626W4_HORVV	YGGGIEKLTHESKLAFVEDRLPAIFDSAANPVDGNCWWINAEDP	FQCLAACIDLSDALKS	634
tr A0A3L6DFV0 A0A3L6DFV0_MAIZE	YGGGIEKLSHEGKLAFVENQLFDIFDSAANPVDGNCWWTNAEDP	FQCLAACMDLSDALRS	638
tr A0A921U2C4 A0A921U2C4_SORBI	YGGGVEKLSHEGKLAFVENQLLDIFDSAANPVDGNCWWTNAEDP	FQCLAACMDLSDALNS	645
tr A0A3L6PDV3 A0A3L6PDV3 PANMI	YGGGVEKLSHEGKLAFVENQLLDIFDSAANPVDGNCWWINAEDP	FQCLAACMDLSDALKS	636
tr A0A2S3HHG5 A0A2S3HHG5_9POAL	YGGGVEKLSHEGKLAFVENQLLDIFDSAANPVDGNCWWINAEDP	FQCLAACMDLSDALKS	640
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	YGGGVEKLSHEGKLAFVENQLLDIFDSAANPVDGNCWWINAEDP	FQCLAACMDLSDALKS	636
tr A0A6G1CRT5 A0A6G1CRT5_90RYZ	YGGGIEKLSHEGKVAFVENQLPDVFDSAATNPDGNCWWMNAEDP	FQCLAACMDLSDALKS	496
tr A0A8J5SZQ4 A0A8J5SZQ4 ZIZPA	YGGGIEKLSHEGKLAFVENQLPDIFDSAANPTDGNCWWMNAEDP	FQCLAASMDLFDALTS	634
	* :.* : : * : : : : * * * * * *****	* **: :* .	

tr A0A816ZBA3 A0A816ZBA3_BRANA	SVPEAAISHIPI	HQD	3SCNGLQHYAALGROKTGAAAVNLFTGEKPADIYADIAARVLKIM	705
sp P92969 RPOT1_ARATH	PFPEAAISHIPI	HQD	GSCNGLQHYAALGRDKLGADAVNLVTGEKPADVYTEIAARVLKIM	722
tr S8EG02 S8EG02_9LAMI	PSPEMSIS <mark>HIPV</mark>	HQD	GSCNGLQHYAALGROTLGAAAVNLVAGEKPADVYTGIANRVLDIM	712
tr A0A0Q3EKS1 A0A0Q3EKS1_BRADI	SSPEAMISHIPV	HQD	GSCNGLQHYAALGROKLGAIAVNLVSGEKPADVYSGIATRVVEIM	738
tr A0A8S0QEE5 A0A8S0QEE5_OLEEU	PSPETTISHMPV	HQD	GSCNGLQHYAALGROKLGAAAVNLVVGDKPADVYSGIASRVLDIM	724
tr A0A0V0IXS1 A0A0V0IXS1_SOLCH	PSPETSISHMPI	HQD	GSCNGLQHYAALGROKLGAAAVNLVAGDKPADVYSGIAARVLDIM	739
NP_001312318.1	PSPETAISYMPI	HQD	GSCNGLQHYAALGROKLGAAAVNLVAGDKPADVYSGIAARVLDIM	748
NP_001289502.1	PSPETAIS <mark>YMPI</mark>	HQD	GSCNGLQHYAALGROTLGAAAVNLVAGDKPADVYSGIAARVLDIM	748
tr A0A565CQ27 A0A565CQ27_9BRAS	SSPYSVISHLPI	HQD	GSCNGLQHYAALGROSFEAAAVNLVAGEKPADVYSEISLRVHEIM	745
tr A0A087GS25 A0A087GS25_ARAAL	PSPYSVISHLPI	HQD	GSCNGLQHYAALGROSFEAAAVNLVAGEKPADVYSEISLRVHEIM	726
tr R0HH96 R0HH96_9BRAS	PSPYSVISHLPI	HQD	GSCNGLQHYAALGROSFEAAAVNLVAGEKPADVYSEISRRVHEIM	737
tr A0A3P6GE78 A0A3P6GE78_BRAOL	PSPYSVISHLPI	HQD	GSCNGLQHYAALGROSFEAAAVNLVAGEKPADVYSEISFRVHEIM	721
tr A0A078HME9 A0A078HME9_BRANA	PSPHSVISNLPI	HQD	GSCNGLQHYAALGROSFEAAAVNLVAGDKPADVYSEISLRVHEIM	719
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	PSPYSVISNLPI	HQD	GSCNGLQHYAALGROSFEAAAVNLVAGEKPADVYSEISLRVHEIM	722
tr A0A2G2WAZ7 A0A2G2WAZ7_CAPBA	SSPHTVISHLPI	HQD	GSCNGLQHYAALGROSMEAAAVNLVAGEKPADVYTEIAMRVDHII	724
tr A0A6P6SN94 A0A6P6SN94_COFAR	SSPHTVISHLPI	HQD	GSCNGLQHYAALGKOSMEAAAVNLIAGDKPADVYSEIAARVHDII	775
tr A0A445AVH3 A0A445AVH3_ARAHY	SSPSSFISHLPI	HQD	GSCNGLQHYAALGROTMEAAAVNLVAKEKPADVYSEIAVRVHDIM	740
tr A0A445END8 A0A445END8_ARAHY	SSPSSFISHLPI	HQD	GSCNGLQHYAALGROTMEAAAVNLVAKEKPADVYSEIAVRVHDIM	741
tr F6HYL3 F6HYL3_VITVI	SSPHTVISHLPI	HQD	GSCNGLQHYAALGRNSLEASAVNLVAGEKPADVYSEIAARVHEIM	725
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	SSPLTIISHLPI	HQD	GSCNGLQHYAALGROSMEAAAVNLTAGDKPADVYSEIAMRVHDIM	726
tr A0A6J1KHL7 A0A6J1KHL7_CUCMA	SSPHTVISHLPI	HQD	GSCNGLQHYAALGROTFEAAAVNLVAGDKPADVYSEIAARVHNIM	718
tr A0A6J1E061 A0A6J1E061_MOMCH	SSPHTVISHLPI	HQD	GSCNGLQHYAALGROTFEAAAVNLVAGEKPADVYTEIAARVHNIM	717
tr A0A0A0LH38 A0A0A0LH38_CUCSA	SAPHTVISYLPI	HQD	GSCNGLQHYAALGROTLEAAAVNLVAGEKPADVYSEIAARVHTIM	721
tr A0A1S3BU26 A0A1S3BU26_CUCME	SAPHTVISYLPI	HQD	GSCNGLQHYAALGROTLEAAAVNLVAGEKPADVYSEIAARVHTIM	721
tr A0A8T3B2K3 A0A8T3B2K3_DENNO	SSPHSMISYLPI	HQD	GSCNGLQHYAALGROSLEAAAVNLVAQEKPADVYSGIAARVYEIV	731
tr A0A3B6RRH2 A0A3B6RRH2_WHEAT	PSPHDAVSHLPI	HQD	GSCNGLQHYAALGROYMGAAAVNLVPGEKPADIYSEIAARVLDVV	695
tr A0A8I6Z6W4 A0A8I6Z6W4_HORVV	SSPHAAVSHLPI	HQD	GSCNGLQHYAALGROYMGAAAVNLVPGEKPADIYSEIAARVLNVV	694
tr A0A3L6DFV0 A0A3L6DFV0_MAIZE	PSPYHAVSHLPI	HQD	GSCNGLQHYAALGROYMGAVAVNLVPGEKPADIYSEIASRVLNVV	698
tr A0A921U2C4 A0A921U2C4_SORBI	PSPYRTVSHLPI	HQD	GSCNGLQHYAALGROYMGAVAVNLVPGEKPADIYSEIAARVLNVV	705
tr A0A3L6PDV3 A0A3L6PDV3_PANMI	SSPYRAVSHLPI	HQD	GSCNGLQHYAALGROYMGAVAVNLVPGEKPADIYSEIAARVLDVV	696
tr A0A2S3HHG5 A0A2S3HHG5_9POAL	SSPYRAVSHLPI	HQD	GSCNGLQHYAALGROYMGAVAVNLVPGEKPADIYSEIAARVHDVV	700
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	SSPYRAVSHLPI	HQD	GSCNGLQHYAALGROYMGAVAVNLVPGEKPADIYSEIAARVHDVV	696
tr A0A6G1CRT5 A0A6G1CRT5_90RYZ	SSPQCAVSHLPI	HQD	GSCNGLQHYAALGROYMGAAAVNLVPGEKPADIYSEIAARVLDVV	556
tr A0A8J5SZQ4 A0A8J5SZQ4_ZIZPA	SSPYCAVSHLPI	HQD	GSCNGLQHYAALGROYMGAAAVNLVPGEKPADIYSEIAARVLIVV	694
	* :*:	* * *	***********	

tr A0A816ZBA3 A0A816ZBA3_BRANA	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTTV <mark>YG</mark> VT <mark>YSGARNQIKKRLKERGAFADDSQNFHAAYYAARVTTNALE</mark>	783
sp P92969 RPOT1_ARATH	VDRKLVKOTVMTSVYGVTYSGARDOIKKRLKERGTFEDDSLTFHASCYAAKITLKALE	803
tr S8EG02 S8EG02_9LAMI	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTIVGARDQITKRLKERDAIADDVELFGAACYAAKVILKAIG	793
tr A0A0Q3EKS1 A0A0Q3EKS1_BRADI	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLKERGVIGDDSELFAASCYAAKVTLTALG	819
tr A0A8S0QEE5 A0A8S0QEE5_OLEEU	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARDQIKKRLKERGAIEDDTELFSAACYAAKTTLIALG	805
tr A0A0V0IXS1 A0A0V0IXS1_SOLCH	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARDQIKRRLKERGAIEDDNELFSAACYAAKTTLTALG	820
NP_001312318.1	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARDQIKKRLKERGVIEDDNELFAAACYAAKTTLTALG	829
NP_001289502.1	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARDQIKRRLKERGVIEDDNELFAAACYAAKTTLTALG	829
tr A0A565CQ27 A0A565CQ27_9BRAS	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTIVGAREQIKRRLEEKGVITDERVLFAAACYAAKVTLAALG	826
tr A0A087GS25 A0A087GS25_ARAAL	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLEEKGVITDERMLFAAACYSAKVTLAALG	807
tr R0HH96 R0HH96_9BRAS	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLEEKGVITDERMLFAAACYSAKVTLAALG	818
tr A0A3P6GE78 A0A3P6GE78_BRAOL	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VT <mark>YVGAREQIKRRLEEKGVITDERMLFAAACYSAKVTLAALG</mark>	802
tr A0A078HME9 A0A078HME9_BRANA	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTIVGAREQIKRRLEEKGVITDERMLFAAACYSAKVTLAALG	800
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLEEKGVITDERMLFAAACYSAKVTLAALG	803
tr A0A2G2WAZ7 A0A2G2WAZ7_CAPBA	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLEEKGLIDDDRLRFTASCYAAKVTLAALG	805
tr A0A6P6SN94 A0A6P6SN94_COFAR	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLQEKGLITDDRLLFSAACYGAKVTMAALG	856
tr A0A445AVH3 A0A445AVH3_ARAHY		821
tr A0A445END8 A0A445END8_ARAHY		822
tr F6HYL3 F6HYL3_VITVI	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAREQIKRRLAEKGLITDERLLFTAACYAAKVTLAALG	806
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTIVGAREQIKRRLEEKGLITDDRLLFTAACYAAKVTLAALG	807
tr A0A6J1KHL7 A0A6J1KHL7_CUCMA	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTFIGAREQIKKRLGEKGLISDDRLLFQAACYAAKVTLAALG	799
tr A0A6J1E061 A0A6J1E061_MOMCH	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTIVGAREQIKKRLEEKGLISDDRLLFRAACYAAKVTLAALG	798
tr A0A0A0LH38 A0A0A0LH38_CUCSA	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAREQIKRRLEEKGLISDDRLRFRASCYAAKVTLSALG	802
tr A0A1S3BU26 A0A1S3BU26_CUCME	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTIVGAREQIKKRLEEKGLISDDRLLFRASCYAAKVTLAALG	802
tr A0A8T3B2K3 A0A8T3B2K3_DENNO	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLIEKGKITDDRLLFRVSCYAAKVTLNALG	812
tr A0A3B6RRH2 A0A3B6RRH2_WHEAT	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTFIGARQQIMKRLQEKGHITDDKLLYDVSCYATRVTLDALG	776
tr A0A8I6Z6W4 A0A8I6Z6W4_HORVV	-VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTFIGARQQIMKRLQEKGHITDEKLLYDVSCYAARVTLDALG	775
tr A0A3L6DFV0 A0A3L6DFV0_MAIZE	IQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARQQITKRLQEKGLITDDKLLYDVSCYATRVTLDALG	818
tr A0A921U2C4 A0A921U2C4_SORBI	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARQQITKRLQEKGLIADDKLLYDVSCYATRVTLDALG	786
tr A0A3L6PDV3 A0A3L6PDV3_PANMI	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARQQITKRLQEKGLITDDKLLYDVSCYATRVTLDALG	777
tr A0A2S3HHG5 A0A2S3HHG5_9POAL	-VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARQQITKRLQEKGLITDDKLLYDVSCYATRVTLDALG	781
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	-VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARQQITKRLQEKGLITDDKLLYDVSCYATRVTLDALG	777
tr A0A6G1CRT5 A0A6G1CRT5_90RYZ	VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARQQITKRLQEKGLITDDKLLYEVSCYATRVTLDALG	637
tr A0A8J5SZQ4 A0A8J5SZQ4_ZIZPA	-VI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGARQQITKRLQEKGLITDDKLLYDVSCYATRVTLDALG	775
	:* <mark>**</mark> *** <mark>**</mark> ****:* <mark>**</mark> ** <mark>:</mark> ****:** :** *:. : *: : .: *: *:	
1		

A0A816ZBA3 A0A816ZBA3_BRANA	TTLQVLTIKRDTEKVKARKQMTAFAPNFVHSLDASHMMMTAVACNRAGLSF <mark>AGVHD</mark> SFWFB93	
sp P92969 RPOT1_ARATH	TTLQVLTLSRETDKVMARRQMTAFAPNFIHSLDGSHMMMTAVACNRAGLSFAGVHDSFWT	913
tr S8EG02 S8EG02_9LAMI	TSLQVLTLQRESEKVLARRQRTAFPPNFVHSLDGSHMMMTAVACKKTGL <mark>NFAGV<mark>HDS</mark>YWT</mark>	903
tr A0A0Q3EKS1 A0A0Q3EKS1_BRADI	TSLQVLTLQRETDKVMVKRQRTAFPPNFVHSLDGSHMMMTAVACNKQGL <mark>Y</mark> FAGV <mark>HDS</mark> YVT	929
tr A0A8S0QEE5 A0A8S0QEE5_OLEEU	TSLQILTLQRETDKVMVKRQRTAFPPNFVHSLDGSHMMMTAIACKEAGL <mark>S</mark> FAGV <mark>HDS</mark> KVT	915
tr A0A0V0IXS1 A0A0V0IXS1_SOLCH	TSLQILTLQRETDKVMVKRQRTAFPPNFVHSLDGSHMMMTAITCKEAGL <mark>S</mark> FAGV <mark>HDS</mark> VVT	930
NP_001312318.1	TSLQILTLQRETDKVMVKRQRTAFPPNFVHSLDGSHMMMTAIACKESGL <mark>SFAGVHDS</mark> KVT	939
NP_001289502.1	TSLQILTLQRETDKVMVKRQRTAFPPNFVHSLDGSHMMMTAIACKESGL <mark>S</mark> FAGV <mark>HDS</mark> KVT	939
tr A0A565CQ27 A0A565CQ27_9BRAS	TSLQVLALQREGNTVDVRKQRTAFPPNFVHSLDGTHMMMTAVACREAGL <mark>NFAGVHDS</mark> KVT	936
tr A0A087GS25 A0A087GS25_ARAAL	TSLQVLALQREGNTVDVRKQRTAFPPNFVHSLDGTHMMMTAVACREAGL <mark>NFAGV<mark>HDS</mark>VVT</mark>	917
tr R0HH96 R0HH96_9BRAS	TSLQVLALQREGNTVDVRKQRTAFPPNFVHSLDGTHMMMTAVACREAGL <mark>NFAGVHDS</mark> KVT	928
tr A0A3P6GE78 A0A3P6GE78_BRAOL	TSLQVLALQREGNTVDVRKQRTAFPPNFVHSLDGTHMMMTAVACREAGL <mark>NFAGV<mark>HDS</mark>VWT</mark>	912
tr A0A078HME9 A0A078HME9_BRANA	TSLQVLALQREGNTVDVRKQRTAFPPNFVHSLDGTHMMMTAVACREAGL <mark>NFAGVHDS</mark> KVT	910
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	TSLQVLALQREGNTVDVRKQRTAFPPNFVHSLDGTHMMMTAVACREAGL <mark>NFAGV<mark>HDS</mark>VVT</mark>	923
tr A0A2G2WAZ7 A0A2G2WAZ7_CAPBA	TSLQVLALQREGDVVEVRKQRTAFPPNFVHSLDGSHMMMTAVACRDAGLHFAGV <mark>HDS</mark> FVT	915
tr A0A6P6SN94 A0A6P6SN94_COFAR	TSLQVLALQREGDSVEARKQRTAFPPNFVHSLDSTHMMMTAIACRDSGL <mark>Q</mark> FAGV <mark>HDS</mark> FVT	966
tr A0A445AVH3 A0A445AVH3_ARAHY	TSLQILALKREGNTVDAKKQRTAFPPNFVHSLDGSHMMMTALACRDAGLCFAGV <mark>HDS</mark> FVT	931
tr A0A445END8 A0A445END8_ARAHY	TSLQILALKREGNTVDAKKQRTAFPPNFVHSLDGSHMMMTALACRDAGL <mark>FAGVHDS</mark> FVT	932
tr F6HYL3 F6HYL3_VITVI	TSLQVLALQREGSVVVIRKQRTAFPPNFVHSLDGSHMMMTAVACRDAGL¢FAGV <mark>HDS</mark> FVT	916
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	TSLQVLALQREGKSVDVRKQRTAFPPNFVHSLDGTHMMMTAIACRDAGLRFAGV <mark>HDS</mark> FVT	917
tr A0A6J1KHL7 A0A6J1KHL7_CUCMA	TSLQVLALQREGNSVDVRKQRTAFPPNFVHSLDGSHMMLTALACRDAGLRFAGV <mark>HDS</mark> FVT	909
tr A0A6J1E061 A0A6J1E061_MOMCH	TSLQVLALRREGNSVDVRKHRTAFPPNFVHSLDGSHMMLTALACRDAGL <mark>R</mark> FAGV <mark>HDS</mark> FVT	908
tr A0A0A0LH38 A0A0A0LH38_CUCSA	TSLQVLALQREGNLVDVRKQRTAFPPNFVHSLDGSHMMLTALACRDAGL <mark>R</mark> FAGV <mark>HDS</mark> FVT	912
tr A0A1S3BU26 A0A1S3BU26_CUCME	TSLQVLALRREGNSVDVRKQRTAFPPNFVHSLDGSHMMLTALACRDAGL <mark>R</mark> FAGV <mark>HDS</mark> FVT	912
tr A0A8T3B2K3 A0A8T3B2K3_DENNO	TSLQVLALKRDGCLVAAKQQKTAFPPNFVHSLDGSHMMMTAISCKNSGLKFAGV <mark>HDS</mark> FVT	922
tr A0A3B6RRH2 A0A3B6RRH2_WHEAT	T SLQCLALRREGDAIATQRQKAAFPPN FVH SLDS SHMMMTAITCKEAGL <mark>H</mark> FAGV <mark>HDS</mark> FWV	886
tr A0A8I6Z6W4 A0A8I6Z6W4_HORVV	TSLQCLALRREGDAIATQRQKAAFPPNFVHSLDSSHMMMTAITCKEAGLHFAGV <mark>HDS</mark> FVV	885
tr A0A3L6DFV0 A0A3L6DFV0_MAIZE	T SLQCLALRREGDAIAIQRQKAAFPPN FVH SLDSSHMMMTAIACKESGL <mark>H</mark> FAGV <mark>HDS</mark> FWV	928
tr A0A921U2C4 A0A921U2C4_SORBI	TSLQCLALRREGDAIAIQRQKAAFPPNFVHSLDSSHMMMTAIACKEAGL <mark>H</mark> FAGV <mark>HDS</mark> FVV	896
tr A0A3L6PDV3 A0A3L6PDV3_PANMI	TSLQCLALRREGDAIAIQRQKAAFPPNFVHSLDSSHMMMTAIACKEAGLHFAGV <mark>HDS</mark> FVV	887
tr A0A2S3HHG5 A0A2S3HHG5_9POAL	TSLQCLALRREGDAIAIHRQKAAFPPNFVHSLDSSHMMMTAIACKEAGL <mark>H</mark> FAGV <mark>HDS</mark> FVV	891
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	TSLQCLALRREGDVIAIQRQKAAFPPNFVHSLDSSHMMMTAIACKEAGL <mark>H</mark> FAGV <mark>HDS</mark> FVV	887
tr A0A6G1CRT5 A0A6G1CRT5_9ORYZ	TSLQCLALRREGDAIALQRQKAAFPPNFVHSLDSSHMMMTAIACKKAGL <mark>H</mark> FAGV <mark>HDS</mark> FVV	747
tr A0A8J5SZQ4 A0A8J5SZQ4_ZIZPA	TSLQCLALRREGDTIAIRRQKAAFPPNLVHSLDSSHMMMTAIACKEADLHFAGVHDSFWV	885
_	*:** *:: *: : ::::** **::***::**::**	

		Polymerase	СТD	
tr A0A816ZBA3 A0A816ZBA3_BRANA	HASDVI	VMNRILRESLLSSTTSQ	S	917
sp P92969 RPOT1_ARATH	HACDVI	DVMNTILREKFVELYEKP	ILENLLESFQKSFPDISFPPLPERGDFDLRKVLESTY	973
tr S8EG02 S8EG02_9LAMI	HACDVI	DELNRILREKFVELYELP	ILENLLESFQVSFPTLSFPPLPPRGDFDLREVMQSPY	963
tr A0A0Q3EKS1 A0A0Q3EKS1_BRADI	HACDVI	D <mark>EMNKILREKFVELYDAP</mark>	ILENLLDSFETSFPTLRFPPLPERGDFDLNDVLQSPY	989
tr A0A8S0QEE5 A0A8S0QEE5_OLEEU	HACDVI	EDMNKILREKFVELYDAP	ILENLLESFQQSFPKLDFPPLPERGDFDLREVLASPY	975
tr A0A0V0IXS1 A0A0V0IXS1_SOLCH	HASDVI	DOMNKILREKFVELYDAP	ILENLLESFQQSFPDLQFPPLPERGDFDLREVLESPY	990
NP_001312318.1	HACDVI	DOMNKILREKFVELYDAP	ILENLLESFQQSFPDLQFPPLPERGDFDLREVLESPY	999
NP_001289502.1	HASDVI	DOMNKILREKFVELYDAP	ILENLLESFQQSFPDLQFPPLPERGDFDLREVLESPY	999
tr A0A565CQ27 A0A565CQ27_9BRAS	HACDVI	DTMNRILREKFVELYNIP	ILEDLLQSFQESYPNLEFPPVPQRGDFDLKEVLKSQY	996
tr A0A087GS25 A0A087GS25_ARAAL	HACDVI	DTMNRILREKFVELYNTP	ILENLLQSFEESFPNLEFPPVPKRGDFDLKEVLKSQY	977
tr R0HH96 R0HH96_9BRAS	HACDVI	DTMNRILREKFVELYNTP	ILEDLLQSFQESYPNLVFPPVPKRGDFDLKEVLKSQY	988
tr A0A3P6GE78 A0A3P6GE78_BRAOL	HACOVI	DTMNRILREKFVELYSTP	ILEDLLQSFQESYPDLVFPSVPKRGDFDLKEVLRSQY	972
tr A0A078HME9 A0A078HME9_BRANA	HACDVI	DTMNRILREKFVELYSSP	VLEDLLQSFQESYPTLVFPPVPKRGEFDLKEVLKSQY	970
tr A0A8S9IIZ9 A0A8S9IIZ9_BRACR	HACDVI	DTMNRILREKFVELYSSP	ILEDLLQSFQESYPTLVFPPVPKRGDFDLEEVLKSQY	983
tr A0A2G2WAZ7 A0A2G2WAZ7 CAPBA	HACNVI	DOMNRILREKFVELYSMP	ILEDLLQSFQESYPALTFPPLPKRGDFDLREVLESPY	975
tr A0A6P6SN94 A0A6P6SN94_COFAR	HACDVI	DKMNQILREKFVELYSMP	ILENLLENFHTSYPTLTFPALPKRGNFDLREVLKSPY	1026
tr A0A445AVH3 A0A445AVH3_ARAHY	HACOVI	ENMSQILREKFVELYDMP	ILENLLEDFQTTYPGLVFPPLPERGDFDLRKVLDSPY	991
tr A0A445END8 A0A445END8_ARAHY	HACDVI	ENMSQILREKFVELYDMP	ILENLLEDFQTSYPGLVFPPLPERGDFDLRKVLDSPY	992
tr F6HYL3 F6HYL3_VITVI	HACDVI	DKMNQILREKFVELYSMP	ILENLLESFQTSHPTLTFPPLPDRGNFDLREVLESPY	976
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	HACDVI	EKMNQILRDKFVELYSKP	ILENLLVSFQTSHPVLVFPPLPERGDFDLRKVLESPY	977
tr A0A6J1KHL7 A0A6J1KHL7_CUCMA	HACDVI	DOMNLILREKFVELYSMP	VLDSLLEEFETTYPGLTFPPLPERGDFDLREVLKSPY	969
tr A0A6J1E061 A0A6J1E061_MOMCH	HACDVI	DOMNVILRETFVELYNMP	VLESLLEGFETTYPGLTFPSLPDRGDFDLREVLNSPY	968
tr A0A0A0LH38 A0A0A0LH38_CUCSA	HACDVI	DRMNLILREKFVELYSMP	VLESLLEGFETTYPGLTFPSLPGRGDFDLQEVLKSPY	972
tr A0A1S3BU26 A0A1S3BU26_CUCME	HACDVI	DRMNLILREKFVELYSMP	VLESLLEGFETTYPGLTFPSLPERGDFNLQEVLRSPY	972
tr A0A8T3B2K3 A0A8T3B2K3_DENNO	HACDVI	NRMNQILREQFVELYNIP	ILENLLESFESSFPTLTFPPLPDRGDFDLRKVLESPY	982
tr A0A3B6RRH2 A0A3B6RRH2_WHEAT	HACDVI	DKMNQILREQFVELYSMP	ILENLLEEFQTLFPTVEFPPCPAQGNFDVREVLKSTY	946
tr A0A816Z6W4 A0A816Z6W4_HORVV	HACDVI	DKMNQILREQFVELYSMP	ILENLLEEFQTLFPTVEFPPCPAQGNFDVREVLTSTY	945
tr A0A3L6DFV0 A0A3L6DFV0_MAIZE	HACDVI	DRMNQILRQQFVELYSMP	ILENLLEEFQTSFPTLEFPPCPPQGNFDVREVLTSTY	988
tr A0A921U2C4 A0A921U2C4_SORBI	HACDVI	DOMNQILRQQFVELYSMP	ILDNLLEEFQMAFPTLEFPPCPPQGNFDVREVLTSTY	956
tr A0A3L6PDV3 A0A3L6PDV3_PANMI	HACDVI	DOMNQILREQFVELYSMP	ILENLLEEFQKSFPTLEFPPCPPQGNFDVREVLTSTY	947
tr A0A2S3HHG5 A0A2S3HHG5_9POAL	HACDVI	DOMNQILREQFVELYSMP	ILENLLEEFQKSFPTLEFPPCPPQGNFDVREVLTSTY	951
tr A0A2T7DVW7 A0A2T7DVW7_9POAL	HACDVI	DOMNQILREQFVELYSMP	ILENLLEEFQKSFPTLEFPPCPPQGNFDVREVLTSTY	947
tr A0A6G1CRT5 A0A6G1CRT5_90RYZ	HACOVI	DKMNQILREQFVELYSMP	ILENLLEEFQASFPPLEFPPCPPQGDFDVREVLASTY	807
tr A0A8J5SZQ4 A0A8J5SZQ4_ZIZPA	HACDVI	MNQILREQFVELYSMP	ILENLLKELQTSFPTLEFPPCPPQGDFDVREVLAST	945
	** *	:. ***: ::.		

/End of Mitochondrial NEP sequences		
TIA0A816ZBA3IA0A816ZBA3 BRANA		917
splP929691RPOT1 ARATH	FFN	976
r S8EG02 S8EG02 9LAMI	FFN	966
r A0A003EKS1 A0A003EKS1 BRADI	FFN	992
r A0A8S0QEE5 A0A8S0QEE5 OLEEU	FFN	978
r A0A0V0IXS1 A0A0V0IXS1 SOLCH	FFN	993
IP 001312318.1	FFN	1002
P 001289502.1	FFN	1002
r A0A565CQ27 A0A565CQ27 9BRAS	FFN	999
r A0A087GS25 A0A087GS25 ARAAL	FFN	980
r R0HH96 R0HH96_9BRAS	FFN	991
r A0A3P6GE78 A0A3P6GE78 BRAOL	FFN	975
r A0A078HME9 A0A078HME9_BRANA	FFN	973
r A0A8S9IIZ9 A0A8S9IIZ9_BRACR	FFN	986
r A0A2G2WAZ7 A0A2G2WAZ7_CAPBA	FFN	978
r  A0A6P6SN94  A0A6P6SN94 COFAR	FFN	1029
r A0A445AVH3 A0A445AVH3_ARAHY	FFN	994
r A0A445END8 A0A445END8_ARAHY	FFN	995
r F6HYL3 F6HYL3_VITVI	FFN	979
r A0A5N6RVV2 A0A5N6RVV2_9ROSI	FFN	980
r A0A6J1KHL7 A0A6J1KHL7_CUCMA	FFN	972
r A0A6J1E061 A0A6J1E061_MOMCH	FFN	971
r A0A0A0LH38 A0A0A0LH38_CUCSA	FFN	975
r A0A1S3BU26 A0A1S3BU26_CUCME	FFN	975
r A0A8T3B2K3 A0A8T3B2K3_DENNO	FFN	985
r A0A3B6RRH2 A0A3B6RRH2_WHEAT	FFN	949
r A0A8I6Z6W4 A0A8I6Z6W4_HORVV	FFN	948
r A0A3L6DFV0 A0A3L6DFV0_MAIZE	FFN	991
r A0A921U2C4 A0A921U2C4_SORBI	FFN	959
r A0A3L6PDV3 A0A3L6PDV3_PANMI	FFN	950
r A0A2S3HHG5 A0A2S3HHG5_9POAL	FFN	954
r A0A2T7DVW7 A0A2T7DVW7_9POAL	FFN	950
r A0A6G1CRT5 A0A6G1CRT5_90RYZ	FFN	810
r A0A8J5SZQ4 A0A8J5SZQ4_ZIZPA	FFN	948

A0A816ZBA3 BRANA Brassica napus, S8EG02 9LAMI, Genlisea aurea, A0A8S0OEE5 OLEEU Olea europaea. 001312318.1 RNA pol 1B, Nicotiana tabacum, A0A565CQ27 9BRAS Arabis nemorensis, R0HH96 9BRAS Capsella rubella, A0A078HME9\_BRANA Brassica napus, A0A2G2WAZ7 CAPBA Capsicum baccatum, A0A445AVH3 ARAHY Arachis hypogaea, F6HYL3 VITVI Vitis vinifera A0A6J1KHL7\_CUCMA Cucurbita maxima, A0A0A0LH38 CUCSA Cucumis sativus, A0A8T3B2K3 DENNO Dendrobium nobile, A0A8I6Z6W4 HORVV Hordeum vulgare. A0A921U2C4 SORBI Sorghum bicolor, A0A2S3HHG5\_9POAL Panicum hallii, A0A6G1CRT5 9ORYZ Orvza meveriana,

P92969 RPOT1 ARATH Arabidopsis thaliana A0A0Q3EKS1 BRADI Brachypodium distachyon A0A0V0IXS1 SOLCH Solanum chacoense 001289502.1 Pol1, Nicotiana sylvestris A0A087GS25 ARAAL Arabis alpine A0A3P6GE78 BRAOL Brassica oleracea A0A8S9IIZ9 BRACR Brassica cretica A0A6P6SN94 COFAR Coffea Arabica A0A445END8 ARAHY Arachis hypogaea A0A5N6RVV2 9ROSI Carpinus fangiana A0A6J1E061 MOMCH Momordica charantia A0A1S3BU26 CUCME Cucumis melo A0A3B6RRH2 WHEAT Triticum aestivum A0A3L6DFV0 MAIZE Zea mavs A0A3L6PDV3 PANMI Panicum miliaceum A0A2T7DVW7\_9POAL Panicum hallii A0A8J5SZO4 ZIZPA Zizania palustris

Figure 1 MSA of mitochondrial RNAPs (NEP) from various plant sources

#### 3.2 'Mix and Match' MSA analysis of Plant Mitochondrial and Chloroplast NEPs

Figure 2 shows the 'mix and match' MSA of the plant mitochondrial and chloroplast NEPs. The N-terminal regions do not show much conservation, up to ~150 amino acids. Even though both show many consecutive Ss in the N-terminal regions, but the chloroplast N-terminal region shows a large number of consecutive Ss up to 9 of them and in addition to that, it also has a –DxD- type metal-binding motif (highlighted in green). BLASTp analysis of both the NEPs of *Arabidopsis thaliana* was performed. The mitochondrial NEP of *Arabidopsis thaliana* shows 59.05% identity to the chloroplast NEP, suggesting the NEPs of mitochondria and chloroplasts are distinctly different, but showed complete conservation of their catalytic regions (Fig. 2). Even though these NEPs are classified as T3/T7 RNA polymerase type, but they showed very little identity (the mitochondrial NEP of *Arabidopsis thaliana* showed only 28.24% identity to the T7 RNAP) suggesting a drastic divergence during evolution, but maintaining the catalytic cores still intact. The decapeptides in the NTD regions are highly conserved in both NEPs and highlighted. It is interesting to note that the metal-binding motifs, -HQD-, -HDS- and –DvD-, and the catalytic core regions, viz. -QVDRKLVKQTVMTSVYGVTY- are completely conserved in both the NEPs and highlighted. The CTD contains an invariant tetrad, –YFFN- motif (marked in red) in both the NEPs suggesting a possible common role.



sp P92969 RPOT1_ARATH(Mt) sp 024600 RPOT3_ARATH(Cp)	VKFLSDS MASAAAASPSLSLNPTSHFQHQTSLVTWLKPPPSSALFRRKTLPFFERHSLPISA * : ::* : *: *: *: *: *:	20 60
sp P92969 RPOT1_ARATH(Mt) sp O24600 RPOT3_ARATH(Cp)	SSS       GTHYPVNRVRGILSSVNLSGVRNGLSINPVNEMGGLSSFRHGQCYVFEGYATAAQAI         SSSS       TSLSVH-EKPISNSVHFHGNLIESFENQDSSYAGTIKGASLI         ***.*       *:       *.*	80 105
sp P92969 RPOT1_ARATH(Mt) sp O24600 RPOT3_ARATH(Cp)	DSTDPEDESSGSDEVNELITEMEKETERIRKKARLAAIPPKRVIAGMGAQKFYMLKQR EELENPVERNGLSGRRRLFMQDPPWISALFLKGLSKMVDQTLKIERK <mark>DID</mark> KRKFDSLRRR :. : * .**: : :: :: * **** *::*	138 165
	NTD 🔺 PR exonucl	ease
sp P92969 RPOT1_ARATH(Mt) sp 024600 RPOT3_ARATH(Cp)	TSRRYGCIECDPLVLKGLDKSARHMYIPYLPMLIPPONWTGYDQGAHFFLP LVRRYGVIECDSLLLAGLDKSAKHMIIPYVPMLVPPKRWKGYDKGGYLFLP **** **** *:* ******:*****************	435 452
sp P92969 RPOT1_ARATH(Mt) sp O24600 RPOT3_ARATH(Cp)	KQQRTVMKRTPKEQLEPVYEALDTLGNTKWKINKKVLSLVDRIWANGGRIGGLVDREDVP KKQQDALKDISHKTAHRVFEALDTLGNTKWRVNRNILDVVERLWADGGNIAGLVNREDVP *:*: .:* :: .**:***********************	495 512
sp P92969 RPOT1_ARATH(MT) sp O24600 RPOT3_ARATH(Cp)	IPEEPEREDQEKFKNWRWESKKAIKQNNERHSQRC <b>DIE</b> LKLEVARKMKDEEGFYYPHN <mark>VD</mark> IPEKPSSEDPEELQSWKWSARKANKINRERHSLRC <b>DVE</b> LKLSVARKMKDEEGFYYPHN <mark>LD</mark> ***:* .** *:::.*:*:*:* * *.**** ***:****.********	555 572
sp P92969 RPOT1_ARATH(Mt) sp 024600 RPOT3_ARATH(Cp)	FRGRAYPIHPYLNHLGSDLCRGILEFCEGKPLGKSGLRWLKIHIANLYAGGVDKLAYEDR FRGRAYPMHPHLNHLSSDLCRGTLEFAEGRPLGKSGLHWLKIHLANLYAGGVEKLSHDAR *******	615 632
sp P92969 RPOT1_ARATH(Mt) sp O24600 RPOT3_ARATH(Cp)	Prexo     Polymerase       IAFTESHLEDIFDSSDRPLEGKRWWLNAEDP     FQCLAACINLSEALRSPFPEAAISHIPIE       LAFVENHLDDIMDSAENPIHGKRWWLKAEDP     FQCLAACVILTQALKSPSPYSVISHLPIE       :**.*.**:**:**:***     *******: *::*****	675 692
sp P92969 RPOT1_ARATH(Mt) sp O24600 RPOT3_ARATH(Cp)	QD       GSCNGLQHYAALGRDKLGADAVNLVTGEKPADVYTEIAARVLKIMQQDAEEDPETFPN         QD       GSCNGLQHYAALGRDSFEAAAVNLVAGEKPADVYSEISRRVHEIMKKDSSKDPESNPT         ************************************	735 752
sp P92969 RPOT1_ARATH(Mt) sp O24600 RPOT3_ARATH(Cp)	ATYAKLMLIQVD <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYSGARDQIKKRLKERGTFEDDSLTFHASCYAA AALAKILIIQVD <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAREQIKRRLEEKGVITDERMLFAAACYSA *: **::: ******	795 812
sp P92969 RPOT1_ARATH(Mt) sp O24600 RPOT3_ARATH(Cp)	KITLKALEEMFEAARAIKSWFGDCAKIIASENNAVCWTTPLGLPVVQPYRKPGRHLVKTT KVTLAALGEIFEAARAIMSWLGDCAKIIASDNHPVRWITPLGLPVVQPYCRSERHLIRTS *:** ** *:******* **:********:*: * * ******	855 872
sp P92969 RPOT1_ARATH(Mt) sp 024600 RPOT3_ARATH(Cp)	LQVLTLSRETDKVMARRQMTAFAPNFIHSLDGSHMMMTAVACNRAGLSFAGVHDS LQVLALQREGNTVDVRKQRTAFPPNFVHSLDGTHMMMTAVACREAGLNFAGVHDS ****:*.**.:*.*************************	915 932
sp P92969 RPOT1_ARATH(Mt) sp 024600 RPOT3_ARATH(Cp)	CDVDVMNTILREKFVELYEKP ILENLLESFQKSFPDISFPPLPERGDFDLRKVLESTYFFN CDVDTMNRILREKFVELYNTP ILEDLLQSFQESYPNLVFPPVPKRGDFDLKEVLKSCYFFN ****.** ***********:.* ****************	976 993

Figure 2 'Mix and Match' MSA analysis of the NEPs of mitochondria and chloroplasts from various plant sources

Figure 3 shows the organization of the NEPs of the mitochondria and chloroplast of *A. thaliana*. The NTDs are not conserved in both the cases, whereas the other three domains are highly conserved and organized in the same order.



Figure 3 Domain-organization of the NEPs of mitochondria and chloroplasts (Numberings from the *A. thaliana* sequence)

### 3.3 Analysis of Active Site Amino Acids of the NEP from Plant Mitochondria

#### 3.1.3 Active Site Amino Acids of the RNA Polymerase Catalytic Core

The catalytic core region essentially contains three components in DNA/RNA polymerases, 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 of the catalytic core are found in the NEPs of the plant mitochondria also. Interestingly, the plant mitochondrial NEP's catalytic core, -<sup>748</sup>D**R**-<sup>4</sup>KLV**K**<sup>753</sup>**Q**<sup>1</sup>TVMTSV**Y**<sup>8</sup>**G**V-, is identical to the catalytic core of the chloroplast NEP of *A. thaliana*, suggesting their common origin. Furthermore, the mitochondrial NEP's 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 -	<sup>820</sup> WLA <mark>Y<sup>-8</sup>G</mark> VT <mark>R<sup>-4</sup>SVT<mark>M</mark>R<sup>1</sup>SVMTLA<mark>Y<sup>8</sup>G</mark>S-</mark>
Viral SP6 SSU RNA Pol	- <sup>612</sup> WDS <mark>I-<sup>8</sup>G</mark> IT <mark>R-4</mark> SLT <mark>KK</mark> 1PVMTLP <mark>Y<sup>8</sup>G</mark> S-
Mitochondrial SSU RNA pol (Sc)	<mark>-<sup>1009</sup>TR<sup>-4</sup>KVVKQ<sup>1</sup>TVMTNVY<sup>8</sup>GV-</mark>
Mitochondrial SSU RNA pol (H. sapie	ens) - <sup>986</sup> TR <sup>-4</sup> KVVKQ <sup>1</sup> TVMTVVY <sup>8</sup> GV-
E. coli DNA pol I (SSU)	- <sup>753</sup> Q <mark>R <sup>4</sup></mark> RSA <mark>K <sup>738</sup>A<sup>1</sup>INFGLI</mark> V <sup>8</sup> GM-
Chloroplast SSU DNA pol IA (ARATH	<sup>1</sup> ) - <sup>873</sup> E <b>R⁻⁴</b> RKA <mark>K<sup>878</sup>M</mark> ¹LNFSIA <b>Y<sup>8</sup>G</b> K-
Chloroplast SSU DNA pol IB (ARATH	/) - <sup>857</sup> E <b>R⁻⁴</b> RKA <mark>K<sup>862</sup>M¹</mark> LNFSIA <b>Y<sup>8</sup>G</b> K-
Chloroplast SSU RNA pol (NEP) (AR	ATH) - <sup>765</sup> D <b>R<sup>-4</sup>KLVK<sup>770</sup>Q</b> <sup>1</sup> TVMTSV <b>Y<sup>8</sup>G</b> V-
Mitochondrial SSU RNA pol (NEP) (A	\ <i>RATH</i> ) - <sup>748</sup> DR <sup>-4</sup> KLVK <sup>753</sup> Q <sup>1</sup> TVMTSVY <sup>8</sup> GV-

Adapted from Palanivelu [5]; *Sc, Saccharomyces cerevisiae; ARATH, Arabidopsis thaliana;* The active site amino acids, highlighted in dark blue, are confirmed by SDM; and other techniques.

#### 3.1.4 Polymerase Metal-Binding Sites

The second important site for the polymerase function is the catalytic metal-binding site which involves in the addition of the incoming nucleotides to the 3'-OH of the growing primer. The catalytic metal-binding amino acids of the mitochondrial NEP are arrived at sequence similarity. The two **Ds** in the  $-HQD^{677}$ - and  $-HD^{909}S$ - motifs (highlighted in dark green in Figs. 1 and 2) are proposed as the catalytic metal-binding amino acids and they could bind a  $Mg^{2+}$  during catalysis. Interestingly, these two motifs are completely conserved in the NEPs of both mitochondria and chloroplasts from various plant sources (Fig. 2). The  $-HD^{812}S$ - motif is also found in T7 RNAP and its involvement was confirmed by an SDM experiment. It was found that the  $-D^{812} \rightarrow N$  in the -HDS- motif exhibited no detectable activity [17]. The 2Ds, proposed in the metal-binding sites are also reported to be highly conserved among most of the SSU RNA polymerases. Usually the D in -QD- and the D in HDS are found to be involved in  $Mg^{2+}$ -binding and 'NTP charge shielding'. They are widely reported in the SSU RNA polymerases like T3/T7 phage RNA polymerases and SSU RNA polymerases of plant and fungal mitochondria and plant chloroplasts [7]. Based on these data, Fig. 4 shows the proposed active site amino acids of the polymerase domain of the plant mitochondrial NEP. For a detailed description of the polymerase's mechanism, see Palanivelu [19].

#### 3.1.5 PR Exonuclease Active Site

The PR exonuclease active site amino acids of NEPs from various plant mitochondria are arrived at from MSA analysis and their sequence similarities to other well-established DEDD-exonuclease superfamilies. The NEPs from various plant mitochondria belong to DEDD(H) subfamily as discussed below. The DEDD-superfamily of PR exonucleases uses either a Y or a H as the catalytic proton acceptor. An invariant H is found in the active site of mitochondrial NEPs (Table 3), and thus, this one belongs to the DEDD(H)-subfamily. Fijalkowska and Schaaper [20] have found DEDD(H)-subfamily of

PR exonuclease in the  $\varepsilon$ -subunits of the bacterial replicase multienzyme complexes (DNA pols III) which belongs to the DnaQ-H-family with the four active site carboxylates (D<sup>12</sup>, E<sup>14</sup>, D<sup>103</sup>, and D<sup>167</sup>) with the invariant **H<sup>162</sup>**, which acts as the general base in catalysis. They also found that modification of the two conserved amino acid residues, viz. **D**<sup>12</sup> $\rightarrow$ Ala and **E**<sup>14</sup> $\rightarrow$ Ala, in the  $\varepsilon$ -subunit by SDM experiments, resulted in the loss of the exonuclease function. These observations were further corroborated by X-ray crystallographic analysis of the  $\varepsilon$ -subunit,  $\varepsilon$ -186, by Hamdan et al. [21]. An interesting DEDD-superfamily of PR 3' $\rightarrow$ 5' exonuclease was found not in polymerases, but in the tRNA processing enzyme, RNase T. At least five amino acid residues, viz. D<sup>23</sup>, E<sup>25</sup>, D<sup>125</sup>, **H<sup>181</sup>** and D<sup>186</sup> were found on the active site of *E. coli* RNase T. The crystal structures of RNase T from *Pseudomonas aeruginosa* and *E. coli* have been reported by Zuo et al. [22] and they found it also belongs to DEDD(H)-subfamily.

Tables 3 and 4 show the summary of the identified/confirmed DEDD(Y/H)-exonuclease active site amino acids in DNA/RNA polymerases from different viral, bacterial, fungal, plant and animal sources.

Table 3 DEDD-superfa	mily of PR	exonuclease	active	site a	mino	acids	from	viral,	bacterial,	fungal,	plant	and	animal
sources													

Exo-Family	Consensus A-site Pattern			Proton Acceptor	Catalytic Metal ion*	ZBSs		
DEDD(Y/H)-superfamily								
Prokaryotic PR Enzymes								
T4 DNA pol ( <i>E. coli</i> Phage)	- <b>1</b> 12x 114	<sup>219</sup> _Υ <sup>320</sup> _	324_	Tyr	Zn <sup>2+</sup>	1		
DNA pol I (E. coli)	-D <sup>355</sup> x = <sup>357</sup> -	<sup>424</sup> -Y <sup>497</sup> -	501_	Tyr	Zn <sup>2+</sup>	1		
DNA pol II (E. coli)	-D <sup>156</sup> xE <sup>158</sup> -	229_γ <sup>331</sup> _	335_	Tyr	Zn <sup>2+</sup>	1		
RNase D (E. coli)	-D <sup>28</sup> xE <sup>30</sup> C	) <sup>85</sup> Y <sup>151</sup> -[	0 <sup>155</sup> -	Tyr	Zn <sup>2+</sup>	1		
DNA pol III, ε-subunit <i>(E. coli</i> )	-D <sup>12</sup> x= <sup>14</sup> -D <sup>1</sup>	<sup>03</sup> -H <sup>162</sup> -D <sup>1</sup>	67_	His	Zn <sup>2+</sup>	1		
RNase T (E. coli)	- <mark>D</mark> <sup>23</sup> X <mark>E</mark> <sup>25</sup> -D <sup>1</sup>	25_ <mark> </mark> 181_	86_	His	Zn <sup>2+</sup>	1		
PR Enzymes in Eukaryotic DN	IA Replicas	es						
DNA pol ε cat. subunit ( <i>Sc</i> )	- <mark>D</mark> 290∨ <mark>=</mark> 292	-D <sup>383</sup> -Y <sup>473</sup>	-D <sup>477</sup> -	His	Zn <sup>2+</sup>	1		
DNA pol δ cat. subunit ( <i>Hs</i> )	-D <sup>316</sup> IE <sup>318</sup>	-D <sup>402</sup> -Y <sup>511</sup>	-D <sup>515</sup> -	His	Zn <sup>2+</sup>	1		
PR Enzymes in DNA Polymera	ases 1A and	1B from	Plant C	Chloroplas	ts			
DNA polymerase 1A ( <i>ARATH</i> )	-DTE <sup>296</sup>	D <sup>368</sup>	-Y <sup>470</sup> —C	0 <sup>474</sup> - Tyr	Zn <sup>2+</sup>	1		
DNA polymerase 1B (ARATH)	-DTE <sup>272</sup>	D <sup>346</sup>	-Y <sup>448</sup> —[	0 <sup>452</sup> - Tyr	Zn <sup>2+</sup>	1		
PR Enzymes in RNA Polymer	ase, NEP fro	m Plant	Chloro	olasts				
Arabidopsis Thaliana	-DVE550	D <sup>572</sup>	-H <sup>586</sup> —[	D <sup>590</sup> -His	Zn <sup>2+</sup>	1		
Arachis hypogaea	-DVE	D	H	D- His	Zn <sup>2+</sup>	1		
Oryza rufipogon	-DVE	D	HI	D- His	Zn <sup>2+</sup>	1		
Nelumbo nucifera	-DIE	D	H(	D- His	Zn <sup>2+</sup>	1		
Nicotiana tabacum	-DTE	D	H(	D- His	Zn <sup>2+</sup>	1		
Capsicum annuum	-DTE	D	H[	D- His	Zn <sup>2+</sup>	1		
Carpinus fangiana	-DTE	D	H[	D- His	Zn <sup>2+</sup>	1		
PR Enzymes in RNA Polymerase, NEP from Plant Mitochondria								
Arabidopsis Thaliana	-DIE <sup>533</sup>	D <sup>555</sup>	H <sup>569</sup> —D	<sup>573</sup> - His	Zn <sup>2+</sup>	1		
Brassica napus	-DVE	D	HD	)- His	Zn <sup>2+</sup>	1		
Coffea Arabica	-DTE	Dl	1D	)- His	Zn <sup>2+</sup>	1		
Triticum aestivum	-DTE	Dl	1D	)- His	Zn <sup>2+</sup>	1		
Zea mays	-DTE	Dl	HD	)- His	Zn <sup>2+</sup>	1		
Orvza meveriana	-DTE	Dl	HD	- His	Zn <sup>2+</sup>	1		

Adapted from Palanivelu [5]. A-site, Active site; \*Water-bound Zn<sup>2+</sup>; ZBSs, Number of zinc-binding sites; Active site amino acids, confirmed by SDM are highlighted in dark blue and X-ray crystallography, in light blue. *Sc, Saccharomyces cerevisiae; Hs, Homo sapiens;* ARATH, *Arabidopsis thaliana*.

Table 4 DEDD(Y/H)-superfamily exonuclease active site amino acids and their distance conservation

DEDD-Superfamily of PR Exonucleases (-DxE-E-H*/Y*D*-)							
Phage DNA Polymera	ses						
T4 DNA pol ( <i>E. coli</i> Pha	age) - <mark>D</mark>	<sup>114</sup> F	219	S <b>Y<sup>320</sup>N→3</b> aa	← <mark>□<sup>324</sup>∨F-</mark> [23]		
T7 DNA pol ( <i>E. coli</i> Pha	age) -DIE	7F	D <sup>235</sup>	- DY <sup>170</sup> N→3 aa	←D <sup>174</sup> VV-		
Prokanyotic DNA Pen	icase (DNA r	ol III- s-sub					
				<b>162</b>	<b>167</b>		
E. COll			105	L <mark>H</mark> <sup>™</sup> G→4 aa←			
	/s -D <sup>12</sup>		D <sup>102</sup>	LH <sup>™</sup> G→4 aa↔	-D <sup>167</sup> AQ		
Snigella dysenteriae	-D		D	LH G→4 aa↔	-D AQ		
Samonella typhilmunun		<b>C</b> L	U	L⊓ G→4 ad←			
Eukaryotic DNA Repli	cases						
DNA pol ε cat. subunit (	(Sc) - <mark>0</mark> 290	⁰I <mark>E</mark> FC	) <sup>383</sup>	E <b>Y<sup>473</sup>S→3 aa</b> ←	D <sup>477</sup> AV- [24,25]		
DNA pol δ cat. subunit	(Hs) - <mark>D</mark> <sup>316</sup>	ا <b>E</b> F	<sup>402</sup> \	/ <mark>Y<sup>511</sup>C→3 aa</mark> ←	D <sup>515</sup> AY- [26 ]		
Plant DNA Polymeras	es IA (Plant (	Chloroplasts	s <sup>#</sup> )				
Arabidonsis Thaliana	-D <sup>294</sup> TE	FD <sup>368</sup> S	SY <sup>470</sup>	$S \rightarrow 3 aa \rightarrow D^{474}A$			
Chlorella desiccate	-D <sup>294</sup> TE	FD <sup>394</sup> R	SY <sup>495</sup>	S→3 aa→D <sup>499</sup> /	K-		
Nelumbo nucifera	-D <sup>447</sup> TE	FD <sup>521</sup> S	FY <sup>634</sup>	$S \rightarrow 3 aa \rightarrow D^{638}$	SI-		
Raphanus sativus	-D <sup>284</sup> TE	FD <sup>358</sup> S	SY <sup>460</sup>	S→3 aa→ <b>D</b> <sup>464</sup> A	J		
Plant DNA Polymeras	es IR (Plant (	Chloroplaste	<mark>.#)</mark>		-		
		- <b>346</b>		<b></b> 452 <i>c</i>	N1-		
Arabidopsis Thaliana	-D <sup></sup> IE	FD <sup>466</sup> N	SY CV <sup>572</sup>	$S \rightarrow 3 aa \rightarrow D^{-5}$	51- 51		
Nicoliaria labacum Sesamum indicum	-D 1E	FD N FD <sup>445</sup> N	01 SV <sup>502</sup>	$5 \rightarrow 3$ aa $\rightarrow D$ $3$			
Raphanus sativus	-D <sup>258</sup> TE	FD <sup>332</sup> N	S <b>Y</b> <sup>434</sup>	$S \rightarrow 3 aa \rightarrow D^{438}S$	SI-		
	D <sup>355</sup> TE	VD424	DV497				
E. COILDINA POLIEXO			RT	A→3 aa→D F	<b>ND-</b>		
SARS-CoVs <sup>^</sup>							
SARS-CoV-1 ExoN/AC	F2 - <sup>90</sup>	V <b>=</b>	243	AH <sup>268</sup> V, →4 aa	273 A I-		
MERS-CoV ExoN/DPI	P4 _ <sup>90</sup>	v V <b>E</b>	243	AH <sup>268</sup> V→4 aa	273		
SARS-CoV-2 ExoN//AC	E2 - <sup>90</sup>	V <b>E</b>	243	AH <sup>268</sup> V→4 aa	- <sup>273</sup> AI-		
Nuclear-Encoded RN/	Polymerase	e (NEP) from	n Plant Ch	loroplasts^			
Arabidopsis Thaliana	-D <sup>548</sup> VE	LD <sup>572</sup> F	N <b>H</b> <sup>586</sup> L	→3 aa→ <mark>D</mark> <sup>590</sup> LC	-		
Arachis hypogaea	-D <sup>575</sup> VE	V <b>D</b> <sup>599</sup> F	N <b>H</b> <sup>613</sup> L	.→3 aa→ <mark>D<sup>617</sup>L</mark> C	;-		
Oryza rufipogon	-D <sup>538</sup> VE	LD <sup>562</sup> F	N <b>H</b> <sup>576</sup> L	→3 aa→D <sup>580</sup> LC	-		
Nelumbo nucifera	-D <sup>586</sup> VE	LD <sup>610</sup> F	N <b>H<sup>624</sup>L</b>	→3 aa→ <mark>D<sup>628</sup>LC</mark>	-		
Nuclear Encoded BN/	Dolymoracy		Diant Mit	ochondria			
					_		
Arabidopsis Thaliana	-DIE <sup>533</sup>	VD <sup>555</sup> F	N <b>H</b> <sup>565</sup> L	$\rightarrow$ 3 aa $\rightarrow$ <b>D</b> <sup>3/3</sup> LC			
Brassica napus	-DVE	LD <sup>322</sup> F	NH <sup>333</sup> L	$\rightarrow 3 \text{ aa} \rightarrow \text{D}^{\text{Str}}LC$			
Coffea Arabica	-DIE <sup>506</sup>	LD***F	NH <sup>642</sup> L	$\rightarrow$ 3 aa $\rightarrow$ <b>D</b> <sup>44</sup> LC			
Triticum aestivum	-DIE	LD	SH L- SH <sup>545</sup>	$\rightarrow$ 3 aa $\rightarrow$ <b>D</b> LC			
Cea mays	-DTE <sup>367</sup>	LD F	SH L- SH <sup>403</sup> I	→3 aa→ <mark>D</mark> LC →3 aa→ <mark>D<sup>407</sup>⊢C</mark>			

Adapted from Palanivelu [5]. *Sc, Saccharomyces cerevisiae; Hs, Homo sapiens* \*The distance between the proton acceptor (H/Y) and the last D are 3 to 4 amino acids Active site amino acids confirmed by SDM analysis are highlighted in dark blue and by X-ray are highlighted in light blue #Similar active site amino acids are found in *E. coli* DNA pol I ^Interestingly, similar active site amino acids are found in SARS-CoVs and other human CoVs ACE2, Angiotensin-Converting Enzyme 2; DPP4, Dipeptidyl peptidase 4.

Figure 4 shows the proposed polymerase and PR exonuclease active sites. In the NEP of mitochondria, the active site  $H^{569}$  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, resulting in the removal of a misincorporated nucleotide. Thus, the 3' $\rightarrow$ 5' exonuclease $\leftrightarrow$ polymerase activities switch between excision and incorporation modes without dissociation of the enzyme-substrate complex [27]. For a detailed mechanism of polymerase, see Palanivelu [19].



Mitochondrial NEP (*ARATH*) (PR exo) – RC**D**<sup>531</sup>I**E**<sup>533</sup>-----NV**D**<sup>555</sup>F-----N**H**<sup>569</sup>L—S**D**<sup>573</sup>LC-



# 4 Conclusion

Mitochondria are semi-autonomous organelles and are found in all eukaryotic cells. They are partly controlled by their own genome and mostly by the nuclear imports. The nuclear-encoded RNA polymerase (NEP) is imported from the nucleus to the mitochondria and is involved in the transcription of all mitochondrial genes. The mitochondrial NEP showed only 59.05% identity to the NEP of chloroplasts from *Arabidopsis thaliana*, suggesting that the NEPs of mitochondria and chloroplasts are distinctly different. However, in both the NEPs, the polymerase catalytic core and PR exonuclease domains are completely conserved. The PR exonuclease of the mitochondrial NEP belongs to the DEDD-superfamily.

# **Compliance with ethical standards**

#### Acknowledgments

The author wishes to thank Dr. N. Srinivasan, Former Professor, Department of Endocrinology, Post Graduate Institute of Basic Medical Sciences, University of Madras, Chennai and Dr. P. Sathiyamoorthy, Chief Scientist, GEM Research Foundation, Chennai, for their useful suggestions on the manuscript.

#### References

- [1] Unseld M, Marienfeld JR, Brandt R, Brennicke A. The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. Nat Genet. 1997; 15:57-61.
- [2] Preuten T, Cincu E, Fuchs J, Zoschke R, Liere K, Borner T. Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. Plant J. 2010; 64:948-959.
- [3] Cupp JD and Nielsen BL. Minireview: DNA replication in plant mitochondria. Mitochondrion 2014; 19:231-237.
- [4] Moriyama T, Sato N. Enzymes involved in organellar DNA replication in photosynthetic eukaryotes. Front Plant Sci. 2014; 5. DOI: 10.3389/fpls.2014.00480.
- [5] Palanivelu P. Identification of Polymerase and Proofreading Exonuclease Domains in the DNA Polymerases IA, IB and Nuclear-Encoded RNA Polymerase of the Plant Chloroplasts. World J Adv Res Rev. 2023; 17:706–727.
- [6] 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
- [7] 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.
- [8] Hedtke B, Borner T, Weihe A. One RNA polymerase serving two genomes. EMBO Reports. 2000; 1:435-440.
- [9] 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.
- [10] Palanivelu P. An overview of the proofreading functions in bacteria and SARS-Coronaviruses. Int J Biochem Res Rev. 2021; 30:33-62.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] Zuo Y, Deutscher MP. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res. 2001; 29:1017–1026.
- [15] 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.
- [16] 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.
- [17] Osumi-Davis PA, Sreerama N, Volkin DB, Middaugh RC, Woody RW, Woody AYM. Bacteriophage T7 RNA Polymerase and its Active-site Mutants: Kinetic, Spectroscopic and Calorimetric Characterization. J Mol Biol. 1994; 237:5-19.
- [18] Kochetkov SN, Rusakova EE, V.L. Tunitskaya VL. Recent studies of T7 RNA polymerase mechanism. FEBS Letters. 1998; 440:264-267.
- [19] 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. 2022; 17:625–565.
- [20] 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.
- [21] 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.

- [22] 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.
- [23] Elisseeva E, Mandal SS, Reha-Krantz LJ. Mutational and pH Studies of the 3' → 5' Exonuclease Activity of Bacteriophage T4 DNA Polymerase. J Biol Chem. 1999; 274:25151-25158.
- [24] 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. 2021; 143:16147-16153.
- [25] 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.
- [26] 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.
- [27] 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.