

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

WJAR) R W	JARR	
Re	world Journal of Advanced search and Reviews		
		World Journal Series INDIA	
Check for undates			

(RESEARCH ARTICLE)

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 semiautonomous 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

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.

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 $(3\rightarrow 5' \rightarrow 3' \text{ exonucleases})$, but the B-family polymerase possesses only one exonuclease domain, i.e., $3'\rightarrow 5'$ 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 Δ PolIB 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 *pollb*, but not *polla* 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 ' σ 70 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.





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	Refere	nce
Dd R ps- MSU-R	NAP	Prokaryotes <u>Eukaryotes</u>	Intrinsic within the RN active site itself	IAP	[4]
Dd R ps- MSU-R	NAPs	All eukaryotic (pols I-III)	Intrinsic within the RM active site itself	IAP	[4]
Dd R ps- MSU-R	R <mark>NAPs</mark> (Plants pols IV & V)	Intrinsic within the RM active site itself	IAP	[4]
Rd R ps <mark>-</mark> SSU Types	(+) Str (e.g., SA SARS-r Human	<u>and RNA Virus</u> ARS-CoVs, elated CoVs,) i-CoVs	ess DEDD(H)-superfamily (In the associated Exol	# N subun	[6] it, MEC)
Rd R ps <mark>-</mark> SSU Types	<u>(-) Stra</u> (e.g., H Viruses	and RNA Viruse uman influenza s A, B & C)	<u>es</u> DEDD(H)- superfamily (In the PA subunit of t	7# he RNA j	[7] polymerase, MEC)
	Bacter	iophages			
<mark>DdD</mark> ps- SSU Types	T4 DNA (<i>E. coli</i>	A pol Phage)	DEDD(Y) superfamily (In the same polypepti	^ ide, MFE	[28])
	<u>Proka</u>	ryotes			
Dd <mark>D</mark> ps- SSU Types	E. coli I <u>Chloro</u>	DNA pol I plasts	DEDD(Y) superfamily' (In the same polypepti	^ ide, MFE	[1])
Dd R ps- SSU Types	DNA po (Plant o	ols IA & IB chloroplasts)	DEDD(Y) superfamily (In the same polypept	# ide, MFE	This work)
Dd <mark>R</mark> ps- SSU Types	RNAP (I (Plant o	NEP) chloroplasts)	DEDD(H) superfamily (In the same polypept	# ide, 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 -CX₆CX₂CX_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**⁻⁴RKA**K**⁸⁷⁸M¹LNFSIA**Y**⁸**G**⁸⁸⁷, is very similar to the well-established active site of *E. coli* DNA pol I, -**R**⁻⁴RSA**K**⁷⁵⁸A¹INFGLI**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

	▲ Exo domain ———	
tr A0A8T2W5E6 A0A8T2W5E6_9CHLO	PPPDVYVVNTIADARRIAARLASSELSSRLHACDTEVMDIDLKRHSPCCHGHVICFSIYC	320
tr A0A1U8QBT1 A0A1U8QBT1_NELNU	-YKKVLVVDNIAVAKEIVRML-TTRYKDLVHAC <mark>DTEV</mark> AKIDAKQETPVDHGELICFSIYS	473
tr A0A6J0P749 A0A6J0P749_RAPSA	-YDKVVVVDTVPAARNVVAKL-VDQYRNLVHSCDTEVSDIEVKDETPVDHGKLICFSIYC	310
sp F4I6M1 POLIA_ARATH	-YDKVLIVDNVQAAKDTVAKL-VNQFRNHVHSCDTEVSGIEVKEETPVDHGELICFSIYC	320
tr A0A6P4CP16 A0A6P4CP16 ARADU	-YEDILVVDNISLAEKVAKML-TVNYRHLIHACDTEVAKIDVKQETPVDHGEIICFSIYS	310
tr A0A834U2N1 A0A834U2N1_9FABA	-YKDILVVDNISLAKKVVRML-TNKYRHLVHACDTEVAKINVKEETPVDHGEITCFSIYS	318
tr A0A6P4DWB3 A0A6P4DWB3 ARADU	-YEDVLVVDSIPLAEEVAKML-TVKYRHLIHACDTEVAKIDVKDETPVDHGEIICFSIYS	357
tr A0A1S2XND1 A0A1S2XND1_CICAR	-YEDILVVDSIPLAEEVVKML-TVKYRHLIYACDTEVAKIDVKQETPVDHGEITCFSIYC	358
tr A0A8B8MN50 A0A8B8MN50 ABRPR	-YEDILVVDNIPLAEEVSKLL-TTKYRHHIFACDTEVAKIDVKQETPVDHGEIICFSIYC	355
tr A0A1S3UFS3 A0A1S3UFS3 VIGRR	-YEDILVVDNISLAEEVSKML-TTKYRHLIYACDTEVAKIDVKQETPVDHGEIICFSIYC	323
tr A0A371FAB7 A0A371FAB7 MUCPR	-YEDVLVVDSIPLAEEVSKML-TTKYRHLIYAQDTEVAKIDVKQETPVDHGEIICFSIYS	344
tr A0A445LVJ8 A0A445LVJ8_GLYSO	-YDDILVVDNIPLAEEVSKML-TTKYRHLIYACDTEVAKIDVKQETPVDHGEITCFSIYC	349
tr A0A445LVL1 A0A445LVL1 GLYSO	-YDDILVVDNIPLAEEVSKML-TTKYRHLIYACDTEVAKIDVKQETPVDHGEITCFSIYC	349
tr A0A8B8Q8R9 A0A8B8Q8R9_9MYRT	-YGKVLIVNDKSVAKEVVRKL-TNQYKHLVHACDTEVAKIDVKQATPVDHGEIICFSIYS	354
tr A0A6P5WN92 A0A6P5WN92 DURZI	-YDQVLVVDNISVAKEVVLML-TTQYRHLVHAQDTEVSNIDVKKETPVDHGEITCFSIYS	443
tr A0A6J1EVX1 A0A6J1EVX1 CUCMO	-YESVLVVDGVSAAKEVVSML-TTKYKNLVHACDTEVAKIDVKQETPVDHGEIICFSIYS	427
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA	VSAAKEVVSML-TTKYKNLVHAQDTEVAKIDVKQETPVDHGEIICFSIYS	401
tr A0A6J1D9Z5 A0A6J1D9Z5_MOMCH	SVSAAKEVVSML-TTKYKNLVHACDTEVAKIDVKQETPVDHGEIICFSIYS	394
tr A0A6J1IIP1 A0A6J1IIP1 CUCMA	SICAAREVVSML-TMKYRNLVHACDTEVAKIDVKQETPVDHGEIICFSIYS	393
tr A0A4U5R370 A0A4U5R370 POPAL	ERVLVVDNASMAREVVSKL-TNQYRHLIHAQDTEVAKIEVKDETPIDHGEITCFSIYS	270
tr A0A6P4ALP7 A0A6P4ALP7_ZIZJJ	GKVLVVNSISVAKKIVQKL-TSQYRHLVHAQDTEVAKIEVKQETPVDHGEIICFSIYS	413
-	*. * . :. <mark>:*<mark>****</mark> *: * :* **.: *****.</mark>	

tr A0A8T2W5E6 A0A8T2W5E6_9CHLO	SESIKKVWHNYS <mark>FD</mark> RHVMERLGVKMNGFDGDTMHMARLWDSSRVGRGGYSLEALSSDENL	440
tr A0A1U8QBT1 A0A1U8QBT1_NELNU	DPSIR <mark>KVWHNYS<mark>FD</mark>SHVIENYGLKISGFHADTMHMARLWDSSRRKEGGYSLEALTMDPKV</mark>	567
tr A0A6J0P749 A0A6J0P749_RAPSA	DSSIR <mark>KVWHNYS<mark>FD</mark>SHIIRNHGIELSGFHADTMHMARLWDSARRTEGGYSLEALTSDPKV</mark>	404
sp F4I6M1 POLIA ARATH	DSFIRKVWHNYSFDSHIIRNHGIEISGFHADTMHMARLWDSARRIKGGYSLEALTSDPKV	414
tr A0A6P4CP16 A0A6P4CP16 ARADU	DPSIKKVWHNYS <mark>FD</mark> CHVIENYGFKVSGFYADTMHMARLWDSSRRWVGGYSLEALTGDREV	404
tr A0A834U2N1 A0A834U2N1 9FABA	DQSIKKVWHNYS <mark>FD</mark> HHVIENYGLKVSGFHADTMHMARLWNSSRRLDGGYSLEALTSDKRV	412
tr A0A6P4DWB3 A0A6P4DWB3 ARADU	DSSIKKVWHNYS <mark>FD</mark> CHIIENYGFKVSGFHADTMHMARLWDSSRRLDGGYSLEGLTGDKRV	451
tr A0A1S2XND1 A0A1S2XND1 CICAR	DPSIKKVWHNYS <mark>FD</mark> CHVIENYGFKVSGFHADTMHMARLWDSSRQLDGGYSLEGLTGDRRV	452
tr A0A8B8MN50 A0A8B8MN50 ABRPR	DSSIEKVWHNYS <mark>FD</mark> CHVIENYGFKVSGFHADTMHMARLWDSSRHLDGGYSLEGLTGDRRV	449
tr A0A1S3UFS3 A0A1S3UFS3 VIGRR	DSSIRKVWHNYS <mark>FD</mark> CHVIENYGFKVSGFHADTMHMARLWDSSRLLDGGYSLEGLTGDRRV	417
tr A0A371FAB7 A0A371FAB7 MUCPR	DSSIKKVWHNYS <mark>FD</mark> CHVIENYGFKVSGFHADTMHMARLWDSSRHLDGGYSLEGLTGDRRV	438
tr A0A445LVJ8 A0A445LVJ8 GLYSO	DSSIKKVWHNYS <mark>FD</mark> CHVIENYGFKVSGFHADTMHMARLWDSSRHLDGGYSLEGLTGDRRV	443
tr A0A445LVL1 A0A445LVL1 GLYSO	DSSIKKVWHNYS <mark>FD</mark> CHVIENYGFKVSGFHADTMHMARLWDSSRHLDGGYSLEGLTGDRRV	443
tr A0A8B8Q8R9 A0A8B8Q8R9 ⁻ 9MYRT	DPDIRKVWHNYS <mark>FD</mark> WHVIENYGIKVSGFHADTMHMARLWNSSRRTEGGYSLEALTGDRRV	447
tr A0A6P5WN92 A0A6P5WN92 DURZI	DQSIKKVWHNYS <mark>FD</mark> WHVIKNYGLEVSGFHADTIHMARLWDSSRRTSGGYSLEALTGDRNI	537
tr A0A6J1EVX1 A0A6J1EVX1 CUCMO	DPLIRKVWHNYS <mark>FD</mark> WHIIENYGIKISGFHADTMHMARLWDSSRRMNGGYSLEALSCDTKV	521
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA	DPLIRKVWHNYS <mark>FD</mark> WHIIENYGIKISGFHADTMHMARLWDSSRRMNGGYSLEALSYDTKV	495
tr A0A6J1D9Z5 A0A6J1D9Z5 MOMCH	DPLIKKVWHNYS <mark>FD</mark> WHIIENYGIKVSGFHADTMHMARLWDSSRRANGGYSLEALSGDVKV	488
tr A0A6J1IIP1 A0A6J1IIP1 CUCMA	DPSIRKVWHNYS <mark>FD</mark> WHIIENYGIKVSGFHADTMHMARLWDSSRRINGGYSLEALSGDTKV	487
tr A0A4U5R370 A0A4U5R370 POPAL	SPDIKKVWHNYS <mark>FD</mark> WHVIENYGISVSGFHADTMHMARLWDSSRRINGGYSLEALTGDQKV	364
tr A0A6P4ALP7 A0A6P4ALP7 ⁻ ZIZJJ	DPSIKKVWHNYS <mark>FD</mark> WHVIENYGLKVSGFHADTMHLARLWDSSRRIKGGYSLEALTGDPKV	507
_	. *. <mark>*******</mark> ** <mark>**</mark> *:: *:.** .**:*:****:*:* ******.*: * .:	
1		

tr A0A8T2W5E6 A0A8T2W5E6 9CHLO	MGATASG	486
TTIA0A1U80BT1 A0A1U80BT1 NELNU	MSEVOOCTKGE PIKKKWCT FGET, IGKI SMKTI FGKKKIKKDGSEGKVVTVPPVFFI.OFF	627
		450
CT AUA6JUP/49 AUA6JUP/49 RAPSA	LGATQTKEEADFLGKISMKTIFGKRKLKKDGTEGKIIVIPPVEELQRED	403
sp F4I6M1 POLIA_ARATH	LGGTQTKEFLGKISMKTIFGKRKLKKDGSEGKIVVIPPVEELQRED	463
tr A0A6P4CP16 A0A6P4CP16 ARADU	MSRTMLTCEKDLIGKVSMKTIFGQKKVKKDGSEGKIVNIAPVEVLQRDE	453
tr A0A834U2N1 A0A834U2N1 9FABA	MCGAQLNHEKDLIGKVSMKTIFGRKKVKKDGSEGKIITIAPVEELQRDE	461
+ r 2 2 6 P 4 D W B 3 2 2 6 P 4 D W B 3 2 2 P 2 D U	MCAAOINHDKDIJGKVSMKTIEGKKKVKKDGSEGKTITIAPVEELORDE	50.0
+ x A 0 A 1 S 2 X ND 1 A 0 A 1 S 2 X ND 1 CT CAP		501
CI AUAISZANDI AUAISZANDI CICAR	MSRAFLDCERDLIGKVSMRAIFSRRALKDGIEGRI LIMAFVEELQKDE	301
tr AUA8B8MN50 AUA8B8MN50_ABRPR	MSRAQLNHEKDLIGKLSMKTIFSKKKVKKDGSEGKTVIIAPVEELQRDE	498
tr A0A1S3UFS3 A0A1S3UFS3_VIGRR	MSRTKLNH <mark>QKD</mark> LIGKVSMKTIFSKKKLKKDGSEGKTSVIAPVEELQRNE	466
tr A0A371FAB7 A0A371FAB7 MUCPR	MSRAQLNHEKDLIGKLSMKTIFSKKKVKKDGSEGKTTIIASVEELQRDE	487
tr A0A445LVJ8 A0A445LVJ8 GLYSO	MSRAOLN	492
+ = 1 = 0 = 44 5 T. VT. 1 = 0 = 44 5 T. VT. 1 GT. Y SO	MSDAOLNHEKDLTGKVSMKTTFSKKKLKKDGSFGKTSTTADVFFLOFF	492
		152
CI AUAOBOQOR9 AUAOBOQOR9_9MIRI	I SDAELGNE KELIGKASMKI IFGKNK VKRDGSEGKMI II PP VEELQKEE	496
tr AUA6P5WN92 AUA6P5WN92_DUR21	MQRTKWRKEEKELIGKTSMKTIFGKKKLKKDGSEGKIITIAPVEELQRQE	587
tr A0A6J1EVX1 A0A6J1EVX1_CUCMO	MSGAKLDQ <mark>EKE</mark> LIGKVSMKTIFGRKKMKKDGSEGKLIVIPPVEELQREE	570
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA	MSGAKLGQ EKE LIGKVSMKTIFGRKKMKKDGSEGKLIVIPPVEELQREE	544
trla0a6J1D9Z5la0a6J1D9Z5_MOMCH	MSGAKLGHEKELIGKVSMKTIFGRKKMKKDGYEGKLTVIPPVEELOREE	537
+ r IAOA6TITTPI IAOA6TITTPI CUCMA	MSGAKIGOFKRI, IGKI SMKTI FGRKKMKKDGSFGKI IVI PRVFFLOPFF	536
		41.2
CT AUA4U5R3/0 AUA4U5R3/0 POPAL	MRGAEPCIK-ELIGKVSMKNIFGKQKVKKDGSEGKLTTIAPVEELQREA	412
tr AUA6P4ALP7 AUA6P4ALP7_ZIZJJ	MSGAQLCDELELIGKVSMKTIFGKRKLKKDGKPGKIVTIPPVEVLQKEE	556
	· · · · · · · · · · · · · · · · · · ·	
tr A0A8T2W5E6 A0A8T2W5E6_9CHLO	DTHWRWVN <mark>Y</mark> FAY <mark>D</mark> AKSTWDLHAALESELREEPAKLDEAVAAAYARSGIVIETLWDLYRAT	546
tr A0A1U8QBT1 A0A1U8QBT1 NELNU	RIPWIF <mark>Y</mark> STI <mark>D</mark> SISTLKLFESLKIKLKNMEWVLDGVKRGTMYDFYEEY	675
tr A0A6J0P749 A0A6J0P749 RAPSA	REAWI SY SAIDAI STLKLYESMSKKLOOKEWRLDGKLLSGRTMLDEVHEE	503
en FATEMI DOLTA ADADU		510
hela0a(p4cpl(la0a(p4cpl(apapu		513
tr AUA6P4CPI6 AUA6P4CPI6_ARADU	RIPWICTBSIDERSTERLIESERNHEEDMPWRFDGVRVSNRTMIDFIREI	503
tr A0A834U2N1 A0A834U2N1_9FABA	RIPWIC <mark>Y</mark> SA <mark>ID</mark> ARSTLKLYESLKSHLSEMTWKVDGVQVRGPVSEKSMYDFYEEY	515
tr A0A6P4DWB3 A0A6P4DWB3 ARADU	RIPWIC <mark>YS</mark> AI D ARSTLNLYESLKSHLSAMPWKLDGVLVSGRTMYDFYNEF	550
tr A0A1S2XND1 A0A1S2XND1 CICAR	RIPWICYSAIDAKSTLKLYESLKSYLSDMPWKFNGVPVSGKTMYDFYNEY	551
+ - 1 A 0 A 9 B 9 M N 5 0 1 A 0 A 9 B 9 M N 5 0 A B D D D		549
		510
tr AUAIS3UF53 AUAIS3UF53_VIGRR	RIPWICY SALDASSTLKLYESLKGHLSDMPWKFDGVPVYGKTMFDFYNEY	516
tr A0A371FAB7 A0A371FAB7_MUCPR	RIPWIC <mark>Y</mark> SAI <mark>D</mark> ARSTLKLYESLKSHLSDMPWNLDGVPVLGKTMYDFYNKY	537
tr A0A445LVJ8 A0A445LVJ8 GLYSO	RIPWICYSAIDASSTLKLYESLKSHLSDMPWKFDGLPVYGKTMYDFYNEY	542
tr A0A445LVL1 A0A445LVL1 GLYSO	RIPWICYSALDASSTLKLYESLKSHLSDMPWKFDGLPVYGKTMYDFYNEY	542
		EAC
CI AUAOBOQOK9 AUAOBOQOK9 MIRI	REWIGI ALDAMSILREESLKKLUKMSWREDGRPCPQKSMLDEINEI	346
tr A0A6P5WN92 A0A6P5WN92_DURZI	RKLWICY SAIDAISTLRLYESLKNKLSSMSWVFNDKPVSGKSMYHFYEEY	637
tr A0A6J1EVX1 A0A6J1EVX1 CUCMO	RKPWVS <mark>YSAID</mark> SICTLKLYESLKKKLSDMPWERDGERIPDKTMFNFYEDY	620
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA	RKPWVS <mark>V</mark> SAIDSICTLKLYESLKKTLSDMPWERDGERIPDKTMFNFYEDY	594
+r A (A 6.T] D 975 A (A 6.T] D 975 MOMCH	RRPWVSYSAIDSICTLKLYFSLKNKLSNMPWFRDGFMIPDKTMFNFYFFY	587
		E0.C
CIAUA60IIIPI AUA60IIIPI CUCMA	KKLWVSI PALIDSICILKLIESLKNKLSGMPWERNGEMIPNKIMENFIEEI	200
tr A0A4U5R370 A0A4U5R370_POPAL	REPWIC <mark>Y</mark> SAI <mark>D</mark> AISTLQLYKSLESQLSKMPWNMDGKPVLKKSMFDFYLEY	462
tr A0A6P4ALP7 A0A6P4ALP7 ZIZJJ	REPWIC <mark>Y</mark> SAI <mark>D</mark> AISTLKLYESLKNHLSMKPWQIDGNLAPGKTMFNFYETY	606
	** ** ** ** ** ** ** ** ** **	
	▲ Pol →	
trlA0A8T2W5E6LA0A8T2W5E6_9CHLO	SPLVPDIYTPSGVPACSTPVLKTLAGKPGRARKALADIGAANEAVNGAAGLEALPDPDD	72.4
	NEWOTEWVTATCWDSUSMCAL WNLSCWUSAEVDETD_VCCOCCEC_NDT	020
L ADD CTOPRADID OF CTOPRADID	NEWGIEWIIAIGWESYSMGALMADSGAYSAEIDETD-ISSQSSES-NDIL	030
triauA6JUP749[AUA6J0P749_RAPSA	TPISTETYTASGWPSVSGVTLKTLAGKVSAEYDFMEDVTDTGAEEDDD	666
sp F4I6M1 POLIA_ARATH	SPLSTENFTASGWPSVGGDVLKELAGKVSAEYDFMDDVSDISLEEVVE	675
tr A0A6P4CP16 A0A6P4CP16 ARADU	YPLRTEMYTSSGWPSVSGDALKVLAGKVSAEYDFDD-FDDDCDLELDDDFR	668
+r 1 0 2 8 3 4 11 2 N 1 2 0 2 8 3 4 11 2 N 1 0 2 8 3 4 11 2 N 1		670
		215
LIAUA0P4DWD3IAUA0P4DWB3_AKADU	INDIIMIIADGWPDVDGDADKADAGKIDAEIDEVDENCDEDESEDEIPE	/15
tr AUA1S2XND1 A0A1S2XND1_CICAR	YKLETEMYTASGWPSVSGDALKILAGNISSDFDFTNNLDLDDE-QG	711
tr A0A8B8MN50 A0A8B8MN50 ABRPR	YNLKTEMYTATGWPSVSGDALKALAGNNSAEYDFFDEACDVD-LDDE-DE	711
tr A0A1S3UFS3 A0A1S3UFS3 VIGRR	YNLKTEMYTATGWPSVSGDALKAMAGKISADYDFFDEDCNLFL-GF-DG	67.8
+ x A 0 A 27 FAR7 A 0 A	VNLETENVTATCHESUSCENT, VALACHTSAEVERR	600
L LA CA AA FAMADA AA FAMADA LA CA AA FAMADA LA CA AA FAMADA LA CA AA FAMADA LA CA AA FAMADA AAA FAMADA AA FAMA		099
tr AUA445LVJ8 AUA445LVJ8_GLYSO	YNLETEMYTATGWPSVSGDALKALAGSISADYDFFDEDCNLDDLDDE-DE	706
tr A0A445LVL1 A0A445LVL1_GLYSO	YNLETEMYTATGWPSVSGDALKALAGSISADYDFFDEDCNLDDLDDE-DE	706
tr A0A8B8Q8R9 A0A8B8Q8R9 9MYRT	VQLPAEIFTPSGWPAVSGAALKNLAGKVSAEYEFTDDISNLESGDD	707
tr A0A6P5WN92 A0A6P5WN92 DUP7T	VKLPARIYTATGWPSVSGDALKTLAGKVSARYDRTD	80.0
La la Ca Calennal La Ca Calennal Concord		200
CT AUAGUTEVAL AUAGUTEVAL CUCMO	FWD21FD11H16M52A26DHDFT1H6KA2HELDDF1DD502DLEAAN	/83
tr AUA6J1J7E3 A0A6J1J7E3_CUCMA	EALSTELYTATGWPSVSRDALKILAGKVSAEFDDFTDNSQSDTEVVN	757
tr A0A6J1D9Z5 A0A6J1D9Z5 MOMCH	KVLSTEMYTASGWPSVSGDALKILAGKVSAEFDDFTDAHDDLQSDNEVDN	754
tr A0A6J1IIP1 A0A6J1IIP1 CUCMA	EDLSTEMYTATGWPSVSGDALKVLAGKVSAEYDYFTDDLOSDNFFGD	749
+ r A 0 A 4115 B 37 0 A 0 A 4115 B 37 0 B 0 B A 1	VDL DVFTYTASGWDSVSGDALKTLARKISSEVVVND	625
La la Ca Ca La Da Ca		020
CI AUA6P4ALP/ AUA6P4ALP/_ZIZJJ	VKL5TDLITATGWPSVSVDALKVLAGKVSADFDIVDDAE-ESQFDDDIGN	//0
	: : :* :* *:*: ::	

tr A0A8T2W5E6 A0A8T2W5E6 9CHLO	DGAT PWSTDDAE I AKLQIERQEALRERT PKEMADI GE KNGYGKLFQYF PTV PDGLAA <mark>CA</mark> A	784
tr A0A1U8QBT1 A0A1U8QBT1 NELNU	-PEQTVNEV-EKRKGTSVSEEETDISAYGTAYTAFGGGKEGREA <mark>CH</mark> A	883
tr A0A6J0P749 A0A6J0P749 RAPSA	aaaav-dqaseaqksktdvdtsaygtayaafgggergkea <mark>ch</mark> a	708
sp F4I6M1 POLIA ARATH	DDDVE-TSETQKSKTDDETDTSAYGTAYVAFGGGERGKEACHA	717
tr A0A6P4CP16 A0A6P4CP16_ARADU	NSROTEVKPVEIEKSAYGTALAAFPTMEEGREA <mark>CH</mark> A	704
tr A0A834U2N1 A0A834U2N1 9FABA	NPSQSEVAREIIDDTAYGTAFSAFTNPEEGREA <mark>CH</mark> A	714
tr A0A6P4DWB3 A0A6P4DWB3 ARADU	IPFAAFPTEEEGREA <mark>CH</mark> A	751
tr A0A1S2XND1 A0A1S2XND1 CICAR	NSGQSHVAVSKVDNSAYGTAFAAFPTEKEGREA <mark>CH</mark> A	747
tr A0A8B8MN50 A0A8B8MN50 ABRPR	NPFAAFPTEEEGREA <mark>CH</mark> A	747
tr A0A1S3UFS3 A0A1S3UFS3 VIGRR	NTPARTECTION NT	714
tr A0A371FAB7 A0A371FAB7_MUCPR	NPROSQVAPVTIDKSAYGTAFVAFPTEEEGREA <mark>CH</mark> A	735
tr A0A445LVJ8 A0A445LVJ8 GLYSO	NPRASART NP	742
tr A0A445LVL1 A0A445LVL1 GLYSO	NPAQSQVASVKIDKSAYGTAYAAFPTEEEGREA <mark>CH</mark> A	742
tr A0A8B8Q8R9 A0A8B8Q8R9 9MYRT	-GEDISESETSE-ASGVLGSTASRNIDTSAYGTAFSAFHNEDEGREA <mark>CH</mark> A	755
tr A0A6P5WN92 A0A6P5WN92 DURZI	SPEKRTDVDTSAYGTAFAAFEDEEKGREA <mark>CH</mark> A	832
tr A0A6J1EVX1 A0A6J1EVX1 CUCMO	DFETMPHEENRRRIVHECANMSDYGTALKAFKLKEEGMEA <mark>C</mark> HA	826
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA	DFETMPREENRRRIIHECANMSDYGTTLTAFKLKEEGMEA <mark>CH</mark> A	800
tr A0A6J1D9Z5 A0A6J1D9Z5 MOMCH	DSETMPHGENKKPVIHESANMSDYGTAFEAFASKEEGREA <mark>CH</mark> A	797
tr A0A6J1IIP1 A0A6J1IIP1 CUCMA	DSETVSHVENKKHIIHESANMSDYGTALKAFGSSEKGREA <mark>CH</mark> A	792
tr A0A4U5R370 A0A4U5R370 POPAL	DSETMTDEDLESKELSVVENENESGHVGNLRRFQTPEEGIEA	670
tr A0A6P4ALP7 A0A6P4ALP7 ZIZJJ	DYKAATDVSEKEEPEEPKAVDSSAYGTAFTAFADSKLEEEGREA	817
	* * <mark>*</mark> *	

tr A0A8T2W5E6 A0A8T2W5E6_9CHLO	VDSL <mark>I</mark> DA <mark>S</mark> AIDTLLSNFIIPLQSDDISKPDASGVY <mark>RVH<mark>C</mark>SLNINTETGRLSARRPNLQNQ</mark>	844
tr A0A1U8QBT1 A0A1U8QBT1 NELNU	IAA <mark>LCEVC</mark> SIDSLVSNFILPLQGSHILGKNGRVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	939
tr A0A6J0P749 A0A6J0P749 RAPSA	IAS <mark>LCEVC</mark> SIDSLISNFILPLOGSNVSGKDGRVHCSLNINTETGRLSARRPNLONO	764
sp F4I6M1 POLIA ARATH	IASLCEVCSIDSLISNFILPLQGSNVSGKDGRVHCSLNINTETGRLSARRPNLQNQ	773
tr A0A6P4CP16 A0A6P4CP16_ARADU	IAS <mark>LCEVC</mark> 5IDSLISNFILPLQGCNILGKDHRIH <mark>C</mark> 5ININTETGRLSARRP5LQNQ	760
tr A0A834U2N1 A0A834U2N1 9FABA	IAAL <mark>CEVC</mark> SIDSLISNFILPLOGSNVSGRNGRIH <mark>C</mark> SLNINTETGRLSARRPNLONO	770
tr A0A6P4DWB3 A0A6P4DWB3_ARADU	IAAL <mark>CEVC</mark> SINSLISNFILPLQGHNISGKDNRIH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	807
tr A0A1S2XND1 A0A1S2XND1_CICAR	IAAL <mark>CEVS</mark> SINSLISNFILPLOGHNISGKDNRVHCSLNINTETGRLSARRPNLONO	803
tr A0A8B8MN50 A0A8B8MN50_ABRPR	IAAL <mark>CEVC</mark> SINSLISNFILPLQGHNISGKDNRVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	803
tr A0A1S3UFS3 A0A1S3UFS3_VIGRR	IAAL <mark>COVC</mark> SINSLISNFILPLOGQNISGKDLRVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	770
tr A0A371FAB7 A0A371FAB7_MUCPR	IAAL <mark>CTVC</mark> SINSLISNFILPLQGHNISGKDLRVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	791
tr A0A445LVJ8 A0A445LVJ8_GLYSO	IAA <mark>LCQVC</mark> SINSLISNFILPLQGHNISGKDLRVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	798
tr A0A445LVL1 A0A445LVL1_GLYSO	IAAL <mark>COVC</mark> SINSLISNFILPLOGHNISGKDLRVHCSLNINTETGRLSARRPNLONO	798
tr A0A8B8Q8R9 A0A8B8Q8R9_9MYRT	IAAL <mark>CEVC</mark> SIDSLISNFILPLQGSNVSGKDGRVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	811
tr A0A6P5WN92 A0A6P5WN92_DURZI	IAS <mark>LCTVC</mark> SIDSLISNFILPLQGSNVSGKSERVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	888
tr A0A6J1EVX1 A0A6J1EVX1_CUCMO	IAAL <mark>CTIC</mark> SIDSLISNFILPLQGSNISGKNGRVH <mark>C</mark> SLNINTETGRLSARRPSLQNQ	882
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA	ISAL <mark>CTIC</mark> SIDSLISNFILPLOGSNISGKNGRVHCSLNINTETGRLSARRPSLONO	856
tr A0A6J1D9Z5 A0A6J1D9Z5_MOMCH	IAAL <mark>CEVC</mark> SIDSLISNFILPLQGSNISGKNGRVH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	853
tr A0A6J1IIP1 A0A6J1IIP1_CUCMA	IAAL <mark>CEVC</mark> SIDSLISNFILPLQGSNISGKNGRIH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	848
tr A0A4U5R370 A0A4U5R370_POPAL	ISS <mark>LCEIC</mark> 5IDSLISNFILPLQSSNLSGKSGRVH <mark>C</mark> 5LNINTETGRLSARRPNLQNQ	726
tr A0A6P4ALP7 A0A6P4ALP7_ZIZJJ	IAAI <mark>CEVC</mark> SIDSLISNFILPLQGSNISGKNGRIH <mark>C</mark> SLNINTETGRLSARRPNLQNQ	873
_	: : <mark>* : .</mark> *::*:******** : . *:* <mark>*</mark> : ******************************	

tr A0A8T2W5E6 A0A8T2W5E6 9CHLO	PALEKDRYRVRKAFTADRKASKTLIVADYGQLELRILAHMANCQSMLTAFKLGGDFHSRT	904
tr A0A1U8QBT1 A0A1U8QBT1 NELNU	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLANCKSMLDAFKAGGDFHSRT	997
tr A0A6J0P749 A0A6J0P749 RAPSA	PALEKDRYKIRKAFVAAPGNSLIVADYGQLELRILAHLAGCKSMMQAFKAGGDFHSRT	822
sp F4I6M1 POLIA_ARATH	PALEKDRYKIRKAFVASPGNTLVVADYGQLELRILAHLTGCKSMMEAFKAGGDFHSRT	831
tr A0A6P4CP16 A0A6P4CP16 ARADU	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLANCKSMLEAFEAGGDFHSRT	818
tr A0A834U2N1 A0A834U2N1 9FABA	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLASCKSMLDAFKAGGDFHSRT	828
tr A0A6P4DWB3 A0A6P4DWB3 ARADU	PALEKDRYKIRKAFIAAPGNSLIVADYGQLELRILAHLANCKSMLDAFKAGGDFHSRT	865
tr A0A1S2XND1 A0A1S2XND1_CICAR	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLANCKSMMEAFKAGGDFHSRT	861
tr A0A8B8MN50 A0A8B8MN50 ABRPR	PALEKDRYKIRQAFIASPGNSLIVADYGQLELRILAHLANCKSMLEAFEAGGDFHSRT	861
tr A0A1S3UFS3 A0A1S3UFS3_VIGRR	PALEKDRYKIRQAFIASPGNSLIVADYGQLELRILAHLADCKSMLEAFKAGGDFHSRT	828
tr A0A371FAB7 A0A371FAB7 MUCPR	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLADCKSMLEAFKAGGDFHSRT	849
tr A0A445LVJ8 A0A445LVJ8_GLYSO	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLADCKSMLEAFEAGGDFHSRT	856
tr A0A445LVL1 A0A445LVL1_GLYSO	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLADCKSMLEAFEAGGDFHSRT	856
tr A0A8B8Q8R9 A0A8B8Q8R9_9MYRT	PALEKDRYKIRQAFIAKPGHSLIVADYGQLELRILAHLADCKSMLDAFRAGGDFHSRT	869
tr A0A6P5WN92 A0A6P5WN92_DURZI	PALEKDRYKIRQAFIAAPGNSLVVADYGQLELRILAHLADCKSMLDAFKAGGDFHSRT	946
tr A0A6J1EVX1 A0A6J1EVX1_CUCMO	PALEKDRYKIRQAFIASPGNSLIVADYGQLELRILAHLANCQSMLDAFKAGGDFHSRT	940
tr A0A6J1J7E3 A0A6J1J7E3_CUCMA	PALEKDRYKIRQAFIASPGNSLIVADYGQLELRILAHLANCKSMLDAFKAGGDFHSRT	914
tr A0A6J1D9Z5 A0A6J1D9Z5_MOMCH	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLANCKSMLEAFKAGGDFHSRT	911
tr A0A6J1IIP1 A0A6J1IIP1_CUCMA	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLANCKSMLEAFKAGGDFHSRT	906
tr A0A4U5R370 A0A4U5R370_POPAL	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRVLAHLANCKSMLDAFKAGGDFHSRT	784
tr A0A6P4ALP7 A0A6P4ALP7_ZIZJJ	PALEKDRYKIRQAFIAAPGNSLIVADYGQLELRILAHLANCKSMLDAFKAGGDFHSRT	931
	*******::*:** * *::*******************	

tr A0A8T2W5E6 A0A8T2W5E6_9CHLO ALG	MYDHIQDAIAKGNCLLEWDGVAEDGSYTPPPVPLLKDLYG <mark>AE</mark> RRA <mark>KVL</mark> NFSIA <mark>YG</mark> K 96	54
tr A0A1U8QBT1 A0A1U8QBT1_NELNU AMN	MYPHIRKAVEQKQVLLEWHPQTGEVKPPVPLLKDAFA <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 10	54
tr A0A6J0P749 A0A6J0P749 RAPSA AMN	MYPHIRKAVDNGEVLLEWHPQPGQDKPPVPLLKDAFG <mark>5E</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 87	9
sp F4I6M1 POLIA_ARATH AMN	MYPHVREAVENGQVILEWHPEPGEDKPPVPLLKDAFGSERRKAKMLNFSIAYGK 88	8
tr A0A6P4CP16 A0A6P4CP16_ARADU AMN	MYPYICDAVNEKQVLLEWHPQPGEDKPPVPLLKDAFA <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 87	15
tr A0A834U2N1 A0A834U2N1 9FABA AMN	MYPYIREAVEKNQVLLEWYPQPGEEKPPVPLLKDAFA <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 88	5
tr A0A6P4DWB3 A0A6P4DWB3 ARADU AMN	MYPYIREAVDKKEVLLEWHPQPGEEKPPVPLLKDAFG <mark>5E</mark> RRA <mark>KML</mark> NFSIA <mark>YG</mark> K 92	22
tr A0A1S2XND1 A0A1S2XND1 CICAR AMN	MYPYIREAVEKKEVLLEWHPQPGEDKPPVPLLKDAFG <mark>5E</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 91	. 8
tr A0A8B8MN50 A0A8B8MN50 ABRPR AMN	MYPYIRQAVEKKEVLLEWHPRPGEDKPPVPLLKDAFGSE <mark>R</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 91	8
tr A0A1S3UFS3 A0A1S3UFS3 VIGRR AMN	MYPFIREAVEKKEVLLEWHPKPGEDKPPVPLLKDAFA <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 88	35
tr A0A371FAB7 A0A371FAB7 MUCPR AMN	MYPYIREAVEKKQVLLEWHPQPGEDKPPVPLLKDAFG <mark>5E</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 90	6
tr A0A445LVJ8 A0A445LVJ8 GLYSO AMN	MYPHIREAVEKKEVLLEWHPQPGEDKPPVPLLKDAFASE <mark>R</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 91	.3
tr A0A445LVL1 A0A445LVL1_GLYSO AMN	MYPHIREAVEKKEVLLEWHPQPGEDKPPVPLLKDAFA <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 91	. 3
tr A0A8B8Q8R9 A0A8B8Q8R9 9MYRT AMN	MYQHIREAVEKKEVLLEWHPQPGDDKPPVPLLKDAFA <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 92	26
tr A0A6P5WN92 A0A6P5WN92 DURZI AMN	MYSHIREAVEKRKVLLEWHPQPGQEKPPVPLLKDAFA <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 10	03
tr A0A6J1EVX1 A0A6J1EVX1 CUCMO AMN	MYPHIRKAVEEGSVLLEWDPQPGEDKPPVPLLKDAFASE <mark>R</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 99	7
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA AMN	MYPHIRKAVEEGSVLLEWDPQPGEDKPPVPLLKDAFASE <mark>R</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 97	1
tr A0A6J1D9Z5 A0A6J1D9Z5_MOMCH AMN	MYPHIRNAVEKGSVLLEWDPQPGEDKPPVPLLKDAFG <mark>5E</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 96	58
tr A0A6J1IIP1 A0A6J1IIP1 CUCMA AMN	MYPHIRKAVEDGSVLLEWDPQPGEDKPPVPLLKDAFGSE <mark>R</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 96	53
tr A0A4U5R370 A0A4U5R370 POPAL AVN	MYPHIREAIEKKRVLLEWYPOPGEDKPPVPLLKDAFASE <mark>R</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 84	1
tr A0A6P4ALP7 A0A6P4ALP7 ZIZJJ AMN	MYPHIREAIEKKQVLLEWDPQPGEEKPPVPLLKDAFG <mark>5ER</mark> RKA <mark>KML</mark> NFSIA <mark>YG</mark> K 98	8
- *:.	** .: .*: . :***	

	Pol 🛶 🛶	▲ — → СТО	
tr A0A8T2W5E6 A0A8T2W5E6_9CHL0	KSGNKAAKAHALRAAINTPI	QGSAADVATAAMLRITADERLREMGW	1070
tr A0A1U8QBT1 A0A1U8QBT1 NELNU	DHASNAQCGHIERAAINTPV		1133
tr A0A6J0P749 A0A6J0P749_RAPSA	QSRAQKNHIQRAAINTPV	QGSAADVAMCAMLEITTNQRLKELGW	983
sp F4I6M1 POLIA_ARATH	KSRAQRNHIQRAAINTPV	QGSAADVAMCAMLEISINQQLKKLGW	992
tr A0A6P4CP16 A0A6P4CP16_ARADU	DQATNYQKGHIERAAINTPV	QGSAADVAMCAMLQIWNNEQLKDLGW	981
tr A0A834U2N1 A0A834U2N1_9FABA	ARATPSQKNHIERAAINTPVQLFSSDDFLIYKMQ	QGSAADVAMCAMLEISKNTRLKELGW	1005
tr A0A6P4DWB3 A0A6P4DWB3_ARADU	GQANKYQKGHIERAAINTPV	QG SAADVAMCAMLE I SNNKQLKELGW	1028
tr A0A1S2XND1 A0A1S2XND1_CICAR	AQANTYQKGHIERAAINTPV	QGSAADVAMCAMIQISNNKQLKELGW	1024
tr A0A8B8MN50 A0A8B8MN50_ABRPR	AQANTYQKGHIERAAINTPV	QG SAADVAMCAMLQ I SNNKQLKELGW	1024
tr A0A1S3UFS3 A0A1S3UFS3_VIGRR	AQANTYQKGHIERAAINTPV	QGSAADVAMCAMLQISKNKQLKELGW	991
tr A0A371FAB7 A0A371FAB7_MUCPR	AQANTYQKGHIERAAINTPV	QG SAADVAMCAMLQ I SNNKQLKELGW	1012
tr A0A445LVJ8 A0A445LVJ8_GLYSO	AQANTYQKGHIERAAINTPV	QGSAADVAMCAMLQISKNKRLKELGW	1019
tr A0A445LVL1 A0A445LVL1_GLYSO	AQANTYQKGHIERAAINTPV	QGSAADVAMCAMLQISKNKRLKELGW	1019
tr A0A8B8Q8R9 A0A8B8Q8R9_9MYRT	AHASRAHRNHIERAAINTPV	QGSAADVAMCAMLRITNNKQLEDLGW	1032
tr A0A6P5WN92 A0A6P5WN92_DURZI	AQCTYAQKGHIERAAINTPV	QGSAADVAMCAMLQILKNEQLKELGW	1109
tr A0A6J1EVX1 A0A6J1EVX1_CUCMO	KQVTRAQKGHIERAAINTPV	QGSAADVAMCAMLEISKNSRLRELGW	1103
tr A0A6J1J7E3 A0A6J1J7E3_CUCMA	KQVTRAQKGHIERAAINTPV	QGSAADVAMCAMLEISKNSRLRELGW	1077
tr A0A6J1D9Z5 A0A6J1D9Z5_MOMCH	KHATRAQRGHIERAAINTPV	QGSAADVAMCAMLEISNNSGLRELGW	1074
tr A0A6J1IIP1 A0A6J1IIP1_CUCMA	KHATRAHKGHIERAAINTPV	QGSAADVAMCAMLEISKNSRLRELGW	1069
tr A0A4U5R370 A0A4U5R370_POPAL	TDASSSLRGHVERAAINTPV	QGSAADVAMCAMLEISKNNRLRELGW	947
tr A0A6P4ALP7 A0A6P4ALP7_ZIZJJ	RHATRSQRGHIERAAINTPV	QGSAADVAMCAMLEISNNETLKELGW	1094
	* ******	1	
tr A0A8T2W5E6 A0A8T2W5E6 9CHLO	RLLL-QVHDEVILEGPKETAEIAQAVVVECMRSF	PFGDAGGDPLRVELSVDSKYAD <mark>TWYDA</mark>	1129
tr A0A1U8QBT1 A0A1U8QBT1_NELNU	QVHDEVILEGPNESAEEARAIVVKCMSK	FYGTNFLKVDLSVDAKCAR <mark>NWYAA</mark>	1186
tr A0A6J0P749 A0A6J0P749 RAPSA	KLLL-QIHDEVILEGPMESAELAKDIVVDCMSKE	FNGKNILSVDLSVDAKCACNWYAA	1040
sp F4I6M1 POLIA_ARATH	RLLL-QIHDEVILEGPIESAEIAKDIVVDCMSKE	FNGRNILSVDLSVDAKCACNWYAA	1049
tr A0A6P4CP16 A0A6P4CP16_ARADU	KLLL-QVHDEVILEGPTESAEVAKAIVVECMSKE	FNGKNILKVGLSVDAKCACNWYAG	1038
tr A0A834U2N1 A0A834U2N1_9FABA	KLLL-QVHDEVILEGPTESAEEAKAIVVDCMSKE	FNGKNILQVDLSVDAKCALNWYSA	1062
+ - 1 A 0 A 6 P 4 D WB 3 1 A 0 A 6 P 4 D WB 3 A P A D U	KILL-OVHDEVILEGESESAEVAKATVVECMAK	FHCKNTLKVDLSVDAKCACNWVSA	1085

tr A0A6J0P749 A0A6J0P749_RAPSA	KLLL-QIHDEVILEGPMESAELAKDIVVDCMSKPFNGKNILSVDLSVDAKCAQNWYAA	1040
sp F4I6M1 POLIA_ARATH	RLLL-QIHDEVILEGPIESAEIAKDIVVDCMSKPFNGRNILSVDLSVDAKCAQNWYAA	1049
tr A0A6P4CP16 A0A6P4CP16_ARADU	KLLL-QVHDEVILEGPTESAEVAKAIVVECMSKPFNGKNILKVGLSVDAKCAQNWYAG	1038
tr A0A834U2N1 A0A834U2N1_9FABA	KLLL-QVHDEVILEGPTESAEEAKAIVVDCMSKPFNGKNILQVDLSVDAKCAINWYSA	1062
tr A0A6P4DWB3 A0A6P4DWB3_ARADU	KLLL-QVHDEVILEGPSESAEVAKAIVVECMAKPFHGKNILKVDLSVDAKCAQNWYSA	1085
tr A0A1S2XND1 A0A1S2XND1_CICAR	KLLL-QVHDEVILEGPTESAEVAKSIVVDCMSKPFYGKNILKVDLSVDAKCAQNWYSA	1081
tr A0A8B8MN50 A0A8B8MN50_ABRPR	KLLLQQVHDEVILEGPTESAEFAKSIVVECMSKPFNGKNILKVDLSVDAKCAQNWYSA	1082
tr A0A1S3UFS3 A0A1S3UFS3 VIGRR	KLLL-QVHDEVILEGPTESAEVAKSIVVECMSKPFNGKNILKVDLSVDAKCAQNWYAG	1048
tr A0A371FAB7 A0A371FAB7_MUCPR	KLLL-QVHDEVILEGPTESSEVAKSIVVECMSKPFNGKNILKVDLSVDAKCALNWYSG	1069
tr A0A445LVJ8 A0A445LVJ8_GLYSO	KLLL-QVHDEVILEGPTESAEVAKSIVVECMSKPFNGKNILKVDLSVDAKCAQNWYSG	1076
tr A0A445LVL1 A0A445LVL1_GLYSO	KLLL-QVLESTINSIHSFADWGNK	1047
tr A0A8B8Q8R9 A0A8B8Q8R9_9MYRT	RLLL-QVHDEVILEGPTESAEVAKAIVVECMSKPFFE-GVNILNVDLAVDAKSAQNWYAA	1090
tr A0A6P5WN92 A0A6P5WN92_DURZI	RLLL-QVHDEVILEGPSESAEIAKAIIIECMSKPFKG-GKNILKVDLAVDAKCAQNWYAA	1167
tr A0A6J1EVX1 A0A6J1EVX1_CUCMO	RLLL-QVHDEVILEGPTESAEVAKAIVVDCMSKPFNGKNILKVDLAVDAKCAQNWYSA	1160
tr A0A6J1J7E3 A0A6J1J7E3_CUCMA	RLLL-QVHDEVILEGPTESAEVAKAIVVDCMSKPFNGKNILKVDLAVDAKCAQNWYSA	1134
tr A0A6J1D9Z5 A0A6J1D9Z5_MOMCH	RLLL-QVHDEVILEGPTESAEVAKAIVVECMSKPFSGKNILNVDLAVDAKCAQNWYSA	1131
tr A0A6J1IIP1 A0A6J1IIP1_CUCMA	RLLL-QVHDEVILEGPTESAEVAKAIVVECMSKPFNGKNILNVDLAVDAKCAQNWYSA	1126
tr A0A4U5R370 A0A4U5R370_POPAL	KLLL-QVHDEVILEGPTESAEVAKAIVVDCMSKPFGGKNFLKVDLAVDAKCAQNWYSA	1004
tr A0A6P4ALP7 A0A6P4ALP7_ZIZJJ	RLLL-QVHDEVILEGPSESAEVARAIVIECMSKPFGGVNFLKVDLAVDAKCAQNWYSA	1151
	: : . : * *	

//End of DNA pol IA sequences		
tr A0A8T2W5E6 A0A8T2W5E6_9CHLO	K	1130
tr A0A1U8QBT1 A0A1U8QBT1_NELNU	ĸ	1187
tr A0A6J0P749 A0A6J0P749 RAPSA	K	1041
sp F4I6M1 POLIA ARATH	ĸ	1050
tr A0A6P4CP16 A0A6P4CP16 ARADU	ĸ	1039
tr A0A834U2N1 A0A834U2N1 9FABA	ĸ	1063
tr A0A6P4DWB3 A0A6P4DWB3 ARADU	ĸ	1086
tr A0A1S2XND1 A0A1S2XND1 CICAR	ĸ	1082
tr A0A8B8MN50 A0A8B8MN50 ABRPR	ĸ	1083
tr A0A1S3UFS3 A0A1S3UFS3 VIGRR	ĸ	1049
tr A0A371FAB7 A0A371FAB7 MUCPR	ĸ	1070
tr A0A445LVJ8 A0A445LVJ8 GLYSO	ĸ	1077
tr A0A445LVL1 A0A445LVL1 GLYSO	H	1048
tr A0A8B8Q8R9 A0A8B8Q8R9 ⁻ 9MYRT	ĸ	1091
tr A0A6P5WN92 A0A6P5WN92 DURZI	ĸ	1168
tr A0A6J1EVX1 A0A6J1EVX1 CUCMO	K	1161
tr A0A6J1J7E3 A0A6J1J7E3 CUCMA	ĸ	1135
tr A0A6J1D9Z5 A0A6J1D9Z5_MOMCH	ĸ	1132
tr A0A6J1IIP1 A0A6J1IIP1 CUCMA	ĸ	1127
tr A0A4U5R370 A0A4U5R370 POPAL	ĸ	1005
tr A0A6P4ALP7 A0A6P4ALP7 ZIZJJ	K	1152
-	:	

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 -CX₆CX₂C_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**⁻⁴RKA**K**⁸⁶²M¹LNFSIA**Y**⁸**G**⁸⁷¹⁻, is identical to DNA pol IA (-**R**⁻⁴RKA**K**^{M1}LNFSIA**Y**⁸**G**-) and is very similar to the well-established active site of *E. coli* DNA pol I, -**R**⁻⁴RKA**K**⁷⁵⁸A¹INFGLI**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

		Exo	
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC	EKNAIQSMATDVVNGTKTRIVSDEGSGVSQVSLRERLGAMYDKVHIV	ONLSAAKEIV	375
tr A0A1U7VL87 A0A1U7VL87_NICSY	EKNAIQSMETDVVNGTKTRIVSDEGTGVSQVSLRERLGAMYDKVHIV	ONLSAAKEVV	375
tr A0A0V0J0J1 A0A0V0J0J1 SOLCH	EKNAIKSVATDFVNGTETKIVSDEGTGLGQITLRERLGAMYEKVHIV	ONLSAAKEVV	325
tr A0A1U8FC39 A0A1U8FC39_CAPAN	EKNTIQSVATTVVNGTETKIVSDEGTGLGQVTLRERLGVMYEKVHIV	ONLSAAKEVV	339
tr A0A6I9UH37 A0A6I9UH37 sesin	K-KEAKPAAKKTVLSDTVSEPLSEKITASGGTELHERLSQVYDTVLVV	SIPAARQVV	311
tr A0A8S0VP65 A0A8S0VP65_OLEEU	VIDEIKNGAADRKCIVADIAYRTQIE-TITHESKNIFERLRTVYDKVLVV	SISVAREVV	331
tr A0A6J0P5F8 A0A6J0P5F8 RAPSA	DVRGRQRPLVASFDSARNESTVTISKVGKRTDLSRVRANLTKIYNKVRVV	ONVSTAKEIV	243
sp Q84ND9 POLIB_ARATH	-VTLKPLNSDTTLDNASYKKTATISKVEKCTNLSQVRANLKKIYNRVRVV	ONVSSAKETV	257
tr A0A6J1ALP4 A0A6J1ALP4_9ROSI	NQDTGHTNPNVTRRDRANESGVASTEEDNVVSQEDISKRLARIYDQVLVV	ONVSVAREVV	396
tr A0A1U8MKL7 A0A1U8MKL7 GOSHI	IQDRGHMDPNVTRRDQANENGVASSEENLPVYRNDIHKQLAKIYDQVLVV I	ONISVAKEVV	374
tr A0A0B0MAF8 A0A0B0MAF8_GOSAR	IQYRRHMDPNATRRDQANENGVASSEENLPVYRNDIHKQLAKIYNQVLVV I	ONISVAKEVV	401
tr A0A1U8KJG2 A0A1U8KJG2_GOSHI	IQYRGHMDPNATRRDQANENGVASSEENLPVYRNDIHKQLAKIYNQVLVV	ONISVAKEVV	401
tr A0A6J1FW48 A0A6J1FW48 CUCMO	NGLKRGAAVEEFSKMTINGGGTKITEAPATSHKPDIKERLNGVYDSVLVV	SIQAAREVV	352
tr A0A6P8EE16 A0A6P8EE16 PUNGR	VGTS-VLVSEPFEDEAVVAFGVDAAKDKASNTSEDVRKKLRGIYEEVIVV	ONISMAREIV	404
tr A0A1S2Y7M1 A0A1S2Y7M1_CICAR	NGNH-SLAT-TAKDKTQAKSAVAMIRSDEQLKLRDRLCSIYEDILVV	INLSHAEEVA	321
tr A0A1S3VL91 A0A1S3VL91_VIGRR	EDKC-NLET-IAKYEANATSVKKAR-SSEQLKLRGRLCSIYEDILVV	INISLAKEVA	285
tr A0A371ED47 A0A371ED47 MUCPR	NGDC-DMDT-TAKDATNATNATSVKKAR-STEQSKLRDRLCSIYEDILVV	INISLAEEVA	262
tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	NGNH-SSGT-TAKDATNATSVKKTRSREEQSKLCDRLCSIYEDVLVV	IDISLAEKVA	301
_	: .* :*: : :*	*	

tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A80VP65 A0A830VP65 _OLEEU tr A0A80VP65 A0A830VP65 _OLEEU tr A0A6J0P5F8 A0A6J0P5F8 RAPSA SD 084ND9 POLIE_ARATH tr A0A6J1ALP4 A0A6J1ALP4 _9ROSI tr A0A108MAF8 A0A0B0MAF8 _GOSAR tr A0A108KJG2 A0A1U8KJG2 _GOSHI tr A0A6J1FW48 A0A6J1FW48 _CUCMO tr A0A6J1FW48 A0A6J1FW48 _CUCMO tr A0A6FE16 A0A6F8EE16 _PUNGR tr A0A132Y7M1 A0A1S2Y7M1 _CICAR tr A0A371ED47 A0A371ED47 _MUCPR tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	RKLTSRYRHLV HAC DT VAR I DVKQQTPVDHGEIICFSIYSGPEADFGDGRSCIWVDVLD RKLTSQYRHLV HAC DT VAR I DVKQQTPVDHGEVICFSIYSGPEADFGDGRSCIWVDVLD SKLTSQYRHLV HAC DT VAN I DVKQQTPVDHGEVICFSIYSGPEADFGDGRSCIWVDVLD SKLTHQYRHFV HAC DT VAN I DVKQQTPVDHGEVICFSIYSGPEADFGDGRSCIWVDVLD SKLTNQYRHFV HAC DT VAR I DVKEETPVDHGEVICFSIYSGPEADFGDGRSCIWVDVLD GILTNQYRHLV HAC DT VAR I DVKEETPVDHGEVICFSIYSGPADFGDGRSCIWVDVLD ARLVNQYRDLV HAC DT VAR I DVKEETPVDHGEVICFSIYSGPADFGDGRSCIWVDVLD AKLVNQYRDLV HAC DT VAR I DVKEETPVDHGEVICFSIYSGPADFGDGRSCIWVDVLG ALLMNQYRNLV HAC DT VSR I DVKEETPVDHGEVICFSIYSGPADFGDGRSCIWVDVLG KMLTTQYSHLV HAC DT VSR I DVKQETPVDHGEVICFSIYSGADADFGNGRSCIWVDVLD LMLTTRFRHLV HAC DT VSN I DVKQETPVDHGEVICFSIYSGADADFGNGRSCIWVDVLD LMLTTRFRHLV HAC DT VSN I DVKQETPVDHGEVICFSIYSGADADFGNGRSCIWVDVLD SMLTMKYRNLV HAC DT VSN I DVKQETPVDHGEVICFSIYSGADADFGNGRSCIWVDVLD LMLTTRFRHLV HAC DT VSN I DVKQETPVDHGEVICFSIYSGADADFGNGRSCIWVDVLD SMLTMKYRNLV HAC DT VSN I DVKQETPVDHGEVICFSIYSGPADFGNGRSCIWVDVLD KMLTVNYRHLI HAC DT VSN I DVKQETPVDHGEVICFSIYSGPADFGNGRSCIWVDVLD KMLTVNYRHLI HAC DT VAR I DVKQETPVDHGEVICFSIYSGPADFGGRSCIWVDVLD KMLTVNYRHLI HAC DT VAR I DVKQETPVDHGEVICFSIYSGPADFGGRSCIWVDVLD	435 435 385 399 371 303 317 456 434 461 461 461 461 461 381 345 322 361
++1202153706412021537064 TOP2C	CCCKNILUFFADFFODDSTDKUMHNYSFDNHUTFNVCFKUSCFHADFMHMADIWDSSDD#	405
tr A0A107VL87 A0A107VL87_NICSY	GGGRNLLVEFAFFFQDFSIRKVWINTSFDNHVIENYGFKVSGFHADTMHMARLWDSSRT GGGRNLLVEFAFFFQDFSIRKVWHNYSFDNHVIENYGFKVSGFHADTMHMARLWDSSRT	495
tr A0A0V0J0J1 A0A0V0J0J1_SOLCH	GGGNDLLVEFAPFFQDPSIRKVWHNYSFDKHVIENYGFKVSGFHADTMHMARLWDSSRRI	445
tr AOA1U8FC39 AOA1U8FC39_CAPAN	GGGKDLLVEFAPFFQDSSIRKVWHNYSFDNHVIENYGFKLSGFHADTMHMARLWDSSRRT	459
tr A0A619UH37 A0A619UH37 _SESIN	GGGKDLLKE FAPFFEDPSIKKVWHNYSFDNHVIENYGLKVSGFYADTMHMARLWNSSRRT GGGKNLIAFFAPFFEDSSTKKWWHNYSFDDHIVENYGLKVSGFFADTMHMADLWNSSDDT	431
+ r 2026.10P5F8 2026.10P5F8 2225	GGGKN LIAE FARFFEDSSIAK VWRN ISFDDHIVEN IGLAV SGFFADIMRMARLWN SSRAI FNGKDVLAFFKPFFEDSSIKKVWHNYSFDNHIIRNYGIKLSGFHGDTMHMARLWD SSROT	363
sp 084ND9 POLIB ARATH	ENGRDILAEFKPFFEDSSIKKVWHNYSFDNHIIRNYGIKLSGFHGDTMHMARLWDSSRGI	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 AUA6P8EE16 AUA6P8EE16 PUNGR	GGGNDLLLEFAFFFEDFKIKKVWHNISFDNHVLENIGIKLSGFHADTMHMARLWNSSRKT GGGREMLNKENNEFODDSIKKVWHNISFDCHVIONYGEKUSGEHNDTMHMADLWDSSDDW	524 441
trla0als3VL91 a0als3VL91 VIGRB	GGGKEIFNKFARFFODSSIKKVWHNYSFDCHVIONYGFKVSGFHADTMHMARHWDSSRW	40.5
tr A0A371ED47 A0A371ED47 MUCPR	GGGKEILDKFADFFQDSSIKKVWHNYSFDCHVVQNYGFKVSGFHADTMHMARLWDSSRRW	382
tr A0A8B8KCR7 A0A8B8KCR7 ABRPR	GGGKEIMNKFARFFQDSSIKKVWHNYS <mark>FD</mark> CHVIQNYGFKVSGFHADTMHMARLWDSSRRW	421
_	.*. :: :* **:* .*:******* *:: ***::***********	
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI	555
tr A0AlS3YPG4 A0AlS3YPG4_TOBAC tr A0AlU7VL87 A0AlU7VL87_NICSY	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI	555 555
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTMI	555 555 505
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A108FC39 A0A108FC39_CAPAN	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTMI LGGYSLEALTGDSSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLFALTGDSOIMSDAKKGPGFKVIGKVSMKNIFGRKKLKKDGSEGKLITI	555 555 505 519 485
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A8S0VP65 A0A8S0VP65 OLEEU	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSVINAKRGFGEEVIGKVSMKNIFGRKKLKKDGSEGKLITI	555 555 505 519 485 505
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A8S0VP65 A0A8S0VP65 _OLEEU tr A0A6J0P5F8 A0A6J0P5F8 _RAPSA	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTMI EGGYSLEALTGDSSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQIMSDAKKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKLIVI AGGYSLEALTGDPSVIPNAKPRFGEELIGKVSMKTIFGRKKLKKDGSEGKLIVI	555 555 505 519 485 505 417
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A8S0VP65 A0A8S0VP65 _OLEEU tr A0A6J0P5F8 A0A6J0P5F8 _RAPSA sp Q84ND9 POLIB_ARATH	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQIMSDAKKGPGEKVIGKVSMKNIFGRKKLKKDGSEGKLITI EGGYSLEALTGDPSVIPNAKPRFGEEMIGKVSMKTIFGRKKLKKDGSEGKLIVI AGGYSLEALTSDPRVLGGTETKFEAELFGKISMKTIFGKGKLKKDGSEGKLVVI	555 555 505 519 485 505 417 431
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A8S0VP65 A0A8S0VP65 _OLEEU tr A0A6J0P5F8 A0A6J0P5F8 _RAPSA sp 084ND9 POLIB _ARATH tr A0A6J1ALP4 A0A6J1ALP4 _9ROSI	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSSVMRDARLVHAERLFDGGGLFGKISMKTIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSSVMRDARLVHAERLFDGGGLFGKISMKTIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSSVMRDARLVHAERLFDEGLFGKISMKTIFGRKKLKKDGSEGKLIVI AGGYSLEALTGDPSVIPNAKPR <u>EEELFGKISMKTIFGKGKLKKDGSEGKLVVI</u> SGGYSLEALTSDPKVLGGTETK <u>EEAELFGKISMKTIFGKGKLKKDGSEGKLVVI</u>	555 555 505 519 485 505 417 431 571
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A6J9UH37 A0A6J9UH37 _SESIN tr A0A630VP65 A0A830VP65 _OLEEU tr A0A6J0P5F8 A0A6J0P5F8 _RAPSA sp Q84ND9 POLIB _ARATH tr A0A6J1ALP4 A0A6J1ALP4 _ 9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _ GOSH1 tr A0AB0MMF8 A0A6J0AP6 _ COSD2	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTMI EGGYSLEALTGDSSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSSVMRDARLVHAERLFDGGGLFGKISMKTIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSSVMRDARLVHAERLFDGGLFGKISMKTIFGRKKLKKDGSEGKLIVI AGGYSLEALTGDSVIPNAKPR GEEVIGKVSMKTIFGRKKLKKDGSEGKLIVI AGGYSLEALTSDPRVLGGTETK EAELFGKISMKTIFGKGKLKKDGSEGKLVII AGGYSLEALTSDRVLGGTETK EAELFGKISMKTIFGKKKLKKDGSEGKLVII AGGYSLEALTGDRNVMNRTKWRK EKELIGKISMKTIFGKKKLKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRK EKELIGKTSMKTIFGKKKLKKDGSEGKMITI	555 555 505 519 485 505 417 431 571 571 549
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A107VL87 A0A107VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A108FC39 A0A108FC39 CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A8S0VP65 A0A8S0VP65 OLEEU tr A0A6J0P58 A0A6J0P578 RAPSA sp Q84ND9 POLIB_ARATH tr A0A6J1AP4 A0A6J1ALP4_9ROSI tr A0A6J1AP4 A0A108MKL7_GOSHI tr A0A80MKL7 A0A108MKL7_GOSHI tr A0A80MKR8 A0A6B0MAR8_GOSAR tr A0A108KK2 A0A108KL72 GOSHI	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSTVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGSEGKVTVI EGGYSLEALTGDSVIPDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSVIPDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSVIPDARFR GEELIGKVSMKTIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETK EARLFGKISMKTIFGRKKLKKDGSEGKLVII AGGYSLEALTSDPRVLGGTETK EARLFGKISMKTIFGKKKLKKDGSEGKLVII AGGYSLEALTGDRNVMQRSTWRK EKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRK EKELIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRK EKELIGKTSMKTIFGKKKVKKDGSEGKMITI	555 505 519 485 505 417 431 571 549 576
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A01U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A8S0VP65 A0A8S0VP65_OLEEU tr A0A6J0P5F8 A0A6J0P5F8_RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A1U8MKL7 A0A1U8MKL7_GOSHI tr A0A80B0MAF8 A0A6B0MAF8_GOSAR tr A0A1U8KJ2 A0A1U8KJ2_GOSHI tr A0A6J1FW48 A0A6J1FW48_CUCMO	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSTVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQIMSDAKKGFGEKISMKTIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSQIMSDAKKGFGEKISMKTIFGRKKLKKDGSEGKLITI AGGYSLEALTGDPSVIPNAKPRFGEKISMKTIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGGKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKLVI GGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKLVI AGGYSLEALTSDRVNQGSTWRKFEAFLFGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI	555 505 519 485 505 417 431 571 549 576 576 526
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A619UH37 A0A619UH37_SESIN tr A0A630P58 A0A630P58 RAPSA sp Q84ND9 POLIB_ARATH tr A0A6J1AD24 A0A6J1AL24_9ROSI tr A0A1U8MKL7 A0A1U8MKL7_GOSHI tr A0A108MKL7 A0A1U8KL62_GOSHI tr A0A108KL8 A0A6J1FW48_CUCM0 tr A0A6J1FW48 A0A6J8EE16_PUNGR	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQIMSDAKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKITI EGGYSLEALTGDSQIMSDAKKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTGDPRVLGGTETKFGEKJIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAFJFGKISMKTIFGGKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKIMVI GGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKKKDGSEGKIMTI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI	555 505 519 485 505 417 431 571 549 576 526 578
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A108FC39 A0A108FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A8S0VP65 A0A8S0VP65_OLEEU tr A0A6J0P5F8 A0A6J0P5F8 RAPSA sp Q84ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4 9ROSI tr A0A108MKL7 A0A108MKL7_GOSHI tr A0A108MKL7 A0A108MKL7_GOSHI tr A0A108MAF8 A0A60B0MAF8 GOSAR tr A0A108MJ62 A0A108KJ62_GOSHI tr A0A6J8E46 A0A6J1FW48 CUCMO tr A0A698E416 A0A6J1FW48 CUCMO tr A0A698E416 A0A698E416_PUNGR tr A0A12277M1 A0A12277M1_CICAR	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQIMSDAKKGFGEVIGKVSMKNIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSQIMSDAKKGFGEVIGKVSMKNIFGRKKLKKDGSEGKLITI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGRKKLKKDGSEGKLWVI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKLWI AGGYSLEALTSDRVLGGTETKFEAFLFGKISMKTIFGKKKKKDGSEGKLWI GGYSLEALTGDRNVMQRSTWRKFEAFLGKISMKTIFGKKKVKKDGSEGKLWI AGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALSGDRVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALSGDRVMQRSAKLGFEKFLIGKTSMKTIFGKKKVKDGSEGKITTL VGGYSLEALSGDRVMGSAKLGFEKFLIGKTSMTTIFGKKKVKDGSEGKMTTI	555 505 505 505 417 485 505 417 431 571 549 576 576 576 576 578 495
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A630P58 A0A630P55 _OLEEU tr A0A630P578 A0A630P578 _RAPSA sp 084ND9 POLIB _ARATH tr A0A631A1P4 A0A631A1P4 _ 9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A1U8KL7 A0A1U8KKJG2 _GOSHI tr A0A631F44 A0A631F448 _CUCMO tr A0A631F448 A0A631F448 _CUCMO tr A0A637F48 = [A0A631F448 _CUCMO tr A0A637F48 = [A0A631F448 _CUCMO tr A0A632Y7M1 A0A1S2Y7M1 _CICAR tr A0A1S3VL91 A0A1S3VL91 _VIGRR	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQIMSDAKKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSQIMSDAKKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKLIVI AGGYSLEALTSDPRVLGGTETKFEAFJFGKISMKTIFGRKKLKKDGSEGKLIVI AGGYSLEALTSDPRVLGGTETKFEAFJFGKISMKTIFGKKKLKKDGSEGKLWVI GGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKLWI AGGYSLEALTGDRVNUQRSTWRKFEXFLIGKISMKTIFGKKKLKKDGSEGKLWI AGGYSLEALTGDRNVMQRSTWRKFEKFLIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMQRSTWRKFEKFLIGKISMKTIFGKKKVKKDGSEGKITI VGGYSLEALTGDRVMAGAKLGFEKFLIGKISMKTIFGGKKWKKDGSEGKITTL VGGYSLEALSGDTKVMSRAKLDFEKFLIGKISMKTIFGQKKAKIDGSAGKMVTI VGGYSLEALSGDTKVMARAKLDFEKFLIGKISMTTIFGQKKAKIDGSAGKMVTI VGGYSLEALSGDTKVMARAKLDFEKFLIGKISMTTIFGQKKAKIDGSAGKMVTI VGGYSLEALTGDRVMARAKLDFEKFLIGKISMTTIFGQKKAKIDGSAGKMVTI VGGYSLEALTGDRVMARAKLDFEKFLIGKISMTTIFGQKKAKIDGSAGKMVTI VGGYSLEALTGDRVMARAKLDFEKFLIGKISMTTIFGQKKAKIDGSAGKMVTI VGGYSLEALTGDRVMARAKLDFEKFLIGKISMTTIFGQKKAKIDGSAGKMVTI	555 505 505 505 417 431 549 576 576 576 578 495 459
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A630P5F8 A0A6J0P5F8 _RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4 _ 9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A0B0MAF8 A0A0B0MAF8 _GOSAR tr A0A1U8KL62 A0A1U8KL52 _GOSHI tr A0A6J1FW48 A0A6J1FW48 _CUCMO tr A0A6J21FW48 A0A6J1FW48 _CUCMO tr A0A6J21FW48 A0A6J27M1_CICAR tr A0A1S2Y7M1 A0A1S2Y7M1_CICAR tr A0A1S3VL91 A0A1S3VL91 _ VIGRR tr A0A1S3VL91 A0A1S3VL91 _ VIGRR tr A0A371ED47 A0A371ED47 _MUCPR	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSJMSDAKKGPGEKYIGKVSMKNIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSQIMSDAKKGPGEKYIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTGDSVIPNAKPRFGEEMIGKVSMKNIFGRKKLKKDGSEGKLITI EGGYSLEALTGDPVLGGTETKFGEEMIGKVSMKNIFGRKKLKKDGSEGKLIVI AGGYSLEALTSDPRVLGGTETKFEAELFGKISMKKIFGKGKLKKDGSEGKLVI SGGYSLEALTGDRVNQGSTWRKFEXELIGKISMKKIFGKGKLKKDGSEGKLVI LGGYSLEALTGDRNVMQRSTWRKFEXELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMGSANLDFEKELIGKISMSTIFGKKKVKDGSEGKITL VGGYSLEALSGDTKVMARAKLDFEKELIGKISMSTIFGQKKAKIDGSAGKMVTI VGGYSLEALTGDRVMARAKLDFEKELIGKISMTIFGQKKAKIDGSEGKMITI	555 505 505 417 431 571 549 576 576 576 576 576 578 495 459 436 475
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A69UP55 A0A8S0VP65_OLEEU tr A0A60J0P578 A0A60J0P578_RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A60B0MAF8 A0A60B0MAF8_GOSAR tr A0A1U8KKL7 A0A1U8KKL7_GOSHI tr A0A108KK27 A0A1U8KL62_GOSHI tr A0A6J1FW48 A0A60J1FW48_CUCM0 tr A0A698EE16 A0A698EE16_PUNGR tr A0A182Y7M1 A0A1S2Y7M1_CICAR tr A0A13VL91 A0A1S3VL91_VIGRR tr A0A371E047 A0A371E047_MUCPR tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGSEGKVTVI EGGYSLEALTGDSQIMSDAKKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKLITI EGGYSLEALTGDSVIPNAKPRFGEKVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAELFGKISMKTIFGRKKLKKDGSEGKIMVI GGYSLEALTSDPRVLGGTETKFEAELFGKISMKTIFGKKKLKKDGSEGKLVII IGGYSLEALTGDRNVMQRSTWRKFENELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDTKVMSGAKLGFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDRVMGRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMSGAKLGFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMSGAKLDFEKELIGKTSMTTIFGCKKKVKDGSEGKMITI VGGYSLEALTGDRVMRSAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVSKAKLD	555 505 519 485 505 417 431 571 549 576 526 578 495 426 578 495 436 475
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A630P5F8 A0A630VP65_OLEEU tr A0A630P5F8 A0A630VP65_OLEEU tr A0A630P5F8 A0A631ALP4_9ROSI tr A0A108MK17 A0A1U8MK17_GOSHI tr A0A018MK17 A0A108MK17_GOSHI tr A0A018KJG2 A0A1U8KJ22_GOSHI tr A0A631FW48 A0A631FW48_CUCMO tr A0A631FW48 A0A631FW48_CUCMO tr A0A632Y7M1 A0A123Y7M1_CICAR tr A0A1S3VL91 A0A1S3VL91_VIGRR tr A0A371ED47 A0A371ED47_MUCPR tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSQIMSDAKKGFGKISMKTIFGRKKLKKDGSEGKITI EGGYSLEALTGDSQIMSDAKKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKITI AGGYSLEALTGDPSVIPNAKPRFGELFGKISMKKIFGRKKLKKDGSEGKIVVI SGGYSLEALTSDPRVLGGTETKFEAELFGKISMKKIFGKGKLKKDGSEGKIVVI AGGYSLEALTGDRVVMCKTWRKFEAELFGKISMKKIFGKGKLKKDGSEGKLVI GGYSLEALTGDRVVMQRSTWRKFEAELFGKISMKKIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRVVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRVVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRVVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKITTL VGGYSLEALTGDRVMGRALGFEKELIGKISMKTIFGKKKVKKDGSEGKITTL VGGYSLEALTGDRVMSGALGFEKELIGKISMKTIFGQKKVKKDGSEGKITTI VGGYSLEALTGDRVMSGAKLDFEKELIGKISMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKISMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLD	555 505 505 417 431 571 549 576 576 576 576 526 578 495 459 436 475
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A630P5F8 A0A630VP65_OLEEU tr A0A630P5F8 A0A630VP65_OLEEU tr A0A630P5F8 A0A630VP65_OLEEU tr A0A630P5F8 A0A630VP65_OLEEU tr A0A630P5F8 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A108MK17 A0A108MK17_GOSHI tr A0A630VP65 A0A600MAF8_GOSAR tr A0A108KJ62 A0A108KJ62_GOSHI tr A0A630VP65 A0A600MAF8_CUCMO tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A630VP65 A0A630VP65_OLEEU tr A0A6370VP65 A0A630VP65_OLEEU tr A0A8380VC7 A0A888VC7_ABRPR tr A0A1S3YP64 A0A1S3YP64_TOBAC	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSQIMSDAKKGPEKVIGKVSMKNIFGRKKLKKDGSEGKITI EGGYSLEALTGDSQIMSDAKKGFGEEVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTGDPSVIPNAKPRFGEEVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAELFGKISMKKIFGKGKLKKDGSEGKLVI SGGYSLEALTGDRNVLGGTETKFEAELFGKISMKKIFGKGKLKKDGSEGKLVI GGYSLEALTGDRNVMQRSTWRKFEAELFGKISMKKIFGKGKLKKDGSEGKLVI GGYSLEALTGDRNVMQRSTWRKFEAELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMSGALGFEKELIGKISMKTIFGKKKVKKDGSEGKITTL VGGYSLEALTGDREVMSGALGFEKELIGKISMKTIFGQKKVKKDGSEGKITTI VGGYSLEALTGDREVMSGAKLDFEKELIGKISMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGYSLEALTGDREVMSKAKLD	555 505 505 417 431 571 549 576 576 576 576 576 576 576 576 576 576
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A630P558 A0A6J0P558 _RAPSA sp Q84ND9 POLIB _ARATH tr A0A6J1ALP4 A0A6J1ALP4 _9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A0B0MAF8 A0A6D1ALP4 _9ROSI tr A0A0B0MAF8 A0A6D1ALP4 _9ROSI tr A0A0B0MAF8 A0A6D1ALP4 _9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A6J1F448 A0A6D1F448 _GOSAR tr A0A6J2 = COM tr A0A537 = COM tr A0A537 = COM tr A0A37 = COM tr A0A27 = COM tr	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSQIMSDAKKG PEKVIGKVSMKNIFGRKKLKKDGSEGKUTI EGGYSLEALTGDSVIPNAKPR FGELMIGKVSMKTIFGRKKLKKDGSEGKIMVI AGGYSLEALTGDPSVIPNAKPR FGELFGKISMKKIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETK EAELFGKISMKKIFGKKKLKKDGSEGKLVI SGGYSLEALTGDRNVMQRSTWRK EKELIGKISMKKIFGKKKKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRK EKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRK EKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRK EKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRK EKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRNVMQRSTWRK EKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDTKVMSGAKLG EKELIGKISMSTIFGKKKVKKDGSEGKITI VGGYSLEALSGDTKVMSGAKLG EKELIGKISMSTIFGKKKVKKDGSEGKITI VGGYSLEALTGDREVMSGALG EKELIGKISMSTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGAKLD EKELIGKISMSTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGAKLD EKELIGKISMSTIFGQKKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD EKELIGKISMSTIFGQKKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD EKELIGKISMSTIFGQKKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD EKELIGKISMSTIFGQKKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD EKELIGKISMSTIFGQKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD EKELIGKISMSTIFGQKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD EKELIGKISMSTIFGQKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD EKELIGKISMSTIFGQKVKNDGSEGKMTTI VGGYSLEALTGDREVMSKAKLD	555 505 519 485 505 417 431 571 549 576 576 526 578 495 459 436 475
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UB37 A0A619UH37 _SESIN tr A0A630VP65 A0A630VP65 _OLEEU tr A0A630VP65 A0A630VP65 _GOSAR tr A0A1U8KJ7 A0A1U8KJ7 _GOSHI tr A0A631FW48 A0A631FW48 _CUCMO tr A0A632V7M1 A0A12SY7M1 _CICAR tr A0A637VP1 A0A1S3VL91 _VIGRR tr A0A371ED47 A0A371ED47 _MUCPR tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A1S3YF64 A0A1S3YF64 _TOBAC tr A0A10V7U87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSQIMSDAKKGPGKVIGKVSMKNIFGRKKLKKDGSEGKITI EGGYSLEALTGDSQIMSDAKKGPGKVIGKVSMKNIFGRKKLKKDGSEGKLITI EGGYSLEALTGDPSVIPNAKPRFGEE IGKISMKTIFGRKKLKKDGSEGKLIVI AGGYSLEALTSDPRVLGGTETKFEAELFGKISMKKIFGKKKLKKDGSEGKLVI SGGYSLEALTGDRVNMRTKWRKFEAELFGKISMKKIFGKKKLKKDGSEGKLVI GGYSLEALTGDRNVMQRSTWRKFEAELIGKISMKTIFGKKKLKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMQRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDTKVMSGALGFEKELIGKISMKTIFGKKKVKKDGSEGKITTL VGGYSLEALSGDTKVMSGAKLGFEKELIGKISMKTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGALGFEKELIGKISMKTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGALLGFEKELIGKISMKTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGALLGFEKELIGKISMKTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGALLGFEKELIGKISMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGAKLDFEKELIGKISMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDFEKELIGKISMTFFGQKVKKDGSEGKMTTI VGGYSLEALTGD	555 505 519 485 505 417 431 571 549 576 576 526 578 495 459 436 475
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A01U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A630P578 A0A630P578 RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A6J0BMKL7 A0A1U8MKL7_GOSHI tr A0A6J0BMK8 A0A60BMAR8_GOSAR tr A0A1U8KJ62 A0A1U8KJ62_GOSHI tr A0A6J1FW48 A0A6J1FW48_CUCMO tr A0A6J8E16 A0A698EE16_PUNGR tr A0A133VL91 A0A1S3VL91_VIGRR tr A0A183VL91 A0A1S3VL91_VIGRR tr A0A8B8KCR7 A0A8B8KCR7_ABRPR tr A0A107VL87 A0A1U7VL87_NICSY tr A0A1087V539 A0A1U8FC39_CAPAN tr A0A108FC39 A0A51FW48_CCAPAN tr A0A108FC39 A0A5107VL87 SOLCH tr A0A108FC39 A0A5107VL87 SOLCH tr A0A108FC39 A0A108FC39_CAPAN	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGSEGKLTI EGGYSLEALTGDSVINDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGSEGKLTI EGGYSLEALTGDSVINDARVRFGEKIGKISMKTIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAELFGKISMKTIFGRKKLKKDGSEGKLVVI SGGYSLEALTSDPRVLGGTETKFEAELFGKISMKTIFGKKKLKKDGSEGKLVVI AGGYSLEALTSDPRVLGGTETKFEAELFGKISMKTIFGKKKLKKDGSEGKLVI GGYSLEALTGDRNVMQRSTWRKFENELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALSGDTKVMSGALGFEKELIGKISMKTIFGKKKVKKDGSEGKITI VGGYSLEALSGDRVMGRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKITI VGGYSLEALTGDRVMGRSTWRKFEKELIGKISMKTIFGKKKVKKDGSEGKITI VGGYSLEALTGDRVMSGALGFEKELIGKISMTTIFGQKKVKKDGSEGKITI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRVSKAKLDFEKELIGKVSMTFJGQVRKGSMYFFYEKY PSVEELQKTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFFYEKY PSVEELQKTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFFYEKY PSVEELQRTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFFYEKY PSVEELQRTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFFYEKY PSVEELQRTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFFYEKY PSVEELQRTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFFYEKY PSVEELQRTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFYEKY PSVEELQRTERELVICYSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYFYEKY P	555 505 519 485 505 417 431 571 549 576 576 526 578 495 436 475 436 475
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A630P578 A0A619UH37_SESIN tr A0A630P578 A0A630P578 RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A0B0MAF8 A0A60B0MAF8_GOSAR tr A0A108KL7 A0A1U8KL7_GOSHI tr A0A108KL7 A0A1U8KL62_GOSHI tr A0A108KL7 A0A108KL7_GOSHI tr A0A108KJ9 POLIB_ARATH tr A0A108KJ9 A0A108KL7_GOSHI tr A0A137FW48 A0A631FW48_CUCMO tr A0A631S1FW48 A0A631FW48_CUCMO tr A0A678EE16 A0A698EE16_PUNGR tr A0A1S37L947 A0A312S77M1_CICAR tr A0A1S37L947 A0A312D47_MUCPR tr A0A888KCR7 A0A888KCR7_ABRPR tr A0A107VL87 A0A1U7VL87_NICSY tr A0A108FC39 A0A108FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A8S0VF65 A0A8S0VF65_OLEEU	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSVMRDARLVHAERLFDGDGLFGKISMKTIFGRKKLKKDGSEGKVTVI EGGYSLEALTGDSVIPNAKPRFGEKVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTGDPSVIPNAKPRFGEKVIGKVSMKTIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEALFGKISMKTIFGKKKLKKDGSEGKLVII GGYSLEALTSDPRVLGGTETKFEALFGKISMKTIFGKKKLKKDGSEGKLVII LGGYSLEALTGDRNVMQRSTWRKFEALFGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRNVMQRSTWRKFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDTKVMSGALGFEKELIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMGRSTWRKFEKELIGKTSMKTIFGKKKVKDGSEGKMITI VGGYSLEALTGDRVMSGALGFEKELIGKTSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVMSGALGFEKELIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVMSGAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRVMSRAKLDFEKELIGKVSMTTIFGQKKVKKDGSEGKMITI ******.**: * * * * * * * * * * * * * * *	555 505 519 485 505 417 431 571 549 576 526 578 495 459 436 475 459 436 475
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0AU07VL87 A0A1U7VL87 _NICSY tr A0A0108FC39 A0A108FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A630P558 A0A630P558 _GOLEU tr A0A630P558 A0A630P558 _GOLEU tr A0A630P558 A0A630P558 _GOLEU tr A0A630P558 A0A631ALP4 _9ROSI tr A0A631ALP4 A0A631ALP4 _9ROSI tr A0A018MK17 A0A108MK17 _GOSHI tr A0A018KJG2 A0A108KJG2 _GOSHI tr A0A631FW48 A0A631FW48 _CUCMO tr A0A69E16 A0A69EE16 _PUNGR tr A0A632Y7M1 A0A123Y7M1 _CICAR tr A0A1S3YPG4 A0A1S3YL91 _VIGRR tr A0A371ED47 A0A371ED47 _MUCPR tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A107VL87 A0A1U7VL87 _NICSY tr A0A619UH37 A0A1U7VL87 _MICSY tr A0A619UH37 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A108FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A108C39 _GAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 _SESIN tr A0A619UH37 _SESIN tr A0A6	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSHVMCDARLVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSOMSDAKKGPEKVIGKVSMKNIFGRKKLKKDGSEGKUTI EGGYSLEALTGDSUNDAKKGPEKVIGKVSMKNIFGRKKLKKDGSEGKUTI AGGYSLEALTGDPVLGGTETKEAELFGKISMKTIFGKKKLKKDGSEGKIVI AGGYSLEALTGDPVLGGTETKEAELFGKISMKKIFGKGKLKKDGSEGKUVI GGYSLEALTGDRVNGGTETKEAELFGKISMKKIFGKGKLKKDGSEGKUVI GGYSLEALTGDRNVMQRSTWRKEKELIGKISMKKIFGKKKKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKEKELIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKEKELIGKISMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKEKELIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMGRAKLDHEKLIGKISMKTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGALGHEKLIGKISMSTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGALGHEKLIGKISMSTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSGAKLDHEKDLIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDHEKDLIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSKAKLDHEKDLIGKVSMTTIFGQKKVKKDGSEGKMTTI VGYSLEALTGDREVMSKAKLDHEKDLIGKVSMTTIFGQKKVKKDGSEGKMTTI VGYSLEALTGDREVMSKAKLD	555 505 519 485 505 417 431 571 549 576 526 578 495 459 436 475 613 613 563 577 543 563
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87 NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A630P578 A0A619UH37_SESIN tr A0A630P578 A0A630P578 RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A6080MAF8 A0A60B0MAF8_GOSAR tr A0A108KL7 A0A1U8KL7_GOSHI tr A0A108KL7 A0A1U8KL7_GOSHI tr A0A108KL7 A0A108KL7_GOSHI tr A0A617FW48 A0A631ALP4_9ROSI tr A0A619E16 A0A698E16_PUNGR tr A0A698E16 A0A698E16_PUNGR tr A0A133VL91 A0A1S3VL91_VIGRR tr A0A1S3VL91 A0A1S3VL91_VIGRR tr A0A888KCR7 A0A888KCR7_ABRPR tr A0A888KCR7 A0A888KCR7_ABRPR tr A0A107VL87 A0A107VL87_NICSY tr A0A108F39 A0A108F39_CAPAN tr A0A619H37 A0A619UH37_SESIN tr A0A630P558 A0A630P578_RAPSA ID 094 DD 19 COLIB_ARATH tr A0A631ALP4 A0A630L9578 ID 094 DD 19 COLIB_ARATH tr A0A631ALP4 A0A631ALP4 9ROSI	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSTVMRDARPVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSVMRDARLVHAERLFDEGLFGKISMKTIFGRKKLKKDGSEGKLTI EGGYSLEALTGDSVINDARLVHAERLFDEGLFGKISMKTIFGRKKLKKDGSEGKLTI EGGYSLEALTGDSVINDARPTTER AGGYSLEALTSDPRVLGGTETKFGELFGKISMKTIFGRKKLKKDGSEGKLVVI SGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKLVVI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKLVII LGGYSLEALTGDRNVMQRSTWRKFENFLIGKISMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDTKVMSGAKLGFEKFLIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDRVMQRSTWRKFEKFLIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMQRSTWRKFEKFLIGKISMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMRGAALGFEKFLIGKISMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVMSGAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVMSKAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVMSKAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRVMSKAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI ******.**:*:*:::::::::::::::::::::::::	555 505 519 485 505 417 431 571 549 576 576 526 578 495 459 436 475 613 613 563 577 543 563 577 491 631
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UH37 A0A619UH37_SESIN tr A0A630P5F8 A0A6J0P5F8_RAPSA sp Q84ND9 POLIB_ARATH tr A0A6J1ADF8 A0A6J1ALP4_9ROSI tr A0A0B0MAF8 A0A6D0MAF8_GOSAR tr A0A0B0MAF8 A0A6D0MAF8_GOSAR tr A0A1U8KL7 A0A1U8KL7_GOSHI tr A0A1U8KL7 A0A1U8KL62_GOSHI tr A0A6J1FW48 A0A6J1FW48_CUCMO tr A0A6J1FW48 A0A6J1FW48_CUCMO tr A0A6J1FW48 A0A6J1FW48_CUCMO tr A0A6J1FW48 A0A6J1FW48_CUCMO tr A0A6J1FW48 A0A6J1FW48_CUCMO tr A0A6J1FW48 A0A53YL91_VIGRR tr A0A1S3YL91 A0A1S3YL91_VIGRR tr A0A1S3YL91 A0A1S3YL91_VIGRR tr A0A8B8KCR7 A0A8B8KCR7_ABRPR tr A0A8B8KCR7 A0A8B8KCR7_ABRPR tr A0A6J0VJJ1 A0A1S3YL95_OLCH tr A0A6J0VJ58 A0A6J0P578_RAPSA sp Q8AD9 SDLB_ARATH tr A0A6J1ALP4 A0A6J1ALP4_9ROSI tr A0A1U8KL7 A0A1U8KL7_GOSHI	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQIMSDAKGFGEKVIGKISMKTIFGRKKLKKDGSEGKVTVI EGGYSLEALTGDSQIMSDAKGFGEKVIGKVSMKNIFGRKKLKKDGSEGKITI EGGYSLEALTGDSVIPNAKPRFGEKVIGKVSMKNIFGRKKLKKDGSEGKITI EGGYSLEALTGDPSVIPNAKPRFGEKVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGRKKLKKDGSEGKLVII AGGYSLEALTSDPRVLGGTETKFEAFLGKISMKTIFGKKKLKKDGSEGKLVII LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALSGDRVMGRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRNVMQRSTWRKFEKFLIGKSMKTIFGGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMSGALGFEKFLIGKISMKTIFGGKKKVKKDGSEGKMTTI VGGYSLEALTGDRGVNSRAKLDFEKFLIGKSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRGVNSRAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRGVNSRAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRGVNSRAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMTTI VGGYSLEALTGDRGVNSRAKLDFEKFLIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRGVNSRAKLDFEKFLIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRGVNSRAKLDFEKFLIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRGVSSAKLDFEKFLIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRGVSSAKLDFEKFLIGKVSMTTIFGQKVKKDGSEGKMTTI VGGYSLEALTGDRGVSSAKLDFEKFLIGKVSVTTIFGQKVKKDGVCVKKFGVSGVFFYEKY PSVEELQRTERELVICVSALDSISTIMLYESLKNKLSKRIWTFDGVRKGSMYEFYEKY PSVEELQRTERELVICVSALDSISTIMLYESLKNKLSKRVWTFDGVRKGSMYEFYEKY PSVEELQRTERELVICVSALDSISTIKLYESLKNKLSKRVWTFDGVVKSMYEFYEKY PFVEELQRUERELVICVSALDSISTIKLYESLKNKLSKNSVFDGKVFGVSGSKSMFFYEKY PFVEELQREFKLVIVSSALDSISTIKLYESLKNKLSKNSVFDGKVFDGVVYA PFVEELQREFKLVIVSSALDSISTIKLYESLKNKLSKNSVFDGKV	555 505 519 485 505 417 431 571 549 576 526 578 495 459 436 475 459 436 475 613 563 577 543 563 577 543 563 577 543 563 477
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC tr A0A1U7VL87 A0A1U7VL87_NICSY tr A0A0V0J0J1 A0A0V0J0J1_SOLCH tr A0A1U8FC39 A0A1U8FC39_CAPAN tr A0A619UB37 A0A619UB37_SESIN tr A0A619UB37 A0A619UB37_SESIN tr A0A8S0VP65 A0A8S0VP65_OLEEU tr A0A6J0P5F8 A0A6J0P5F8 RAPSA sp Q84ND9 POLIB_ARATH tr A0A108MKL7 A0A1U8MKL7_GOSHI tr A0A108MKL7 A0A108MKL7_GOSHI tr A0A108MKL7 A0A108KK27_GOSHI tr A0A108KJ62 A0A108KJ62_GOSHI tr A0A108KJ62 A0A108KJ62_GOSHI tr A0A108KJ91 A0A1S2Y7M1_CICAR tr A0A1S3VP64 A0A371ED47_MUCPR tr A0A8B8KCR7 A0A371ED47_MUCPR tr A0A8B8KCR7 A0A8B8KCR7_ABRPR tr A0A888KCR7 A0A8B8KCR7_ABRPR tr A0A8049 A0A108539_CAPAN tr A0A89564 A0A108579_CAPAN tr A0A8049578 A0A61908739_CAPAN tr A0A619087 A0A61908759_CAPAN tr A0A619578 A0A630755_OLEEU tr A0A619578 A0A630755_RAPSA BD D84ND9 POLIB_ARATH tr A0A808MF67 A0A108KL7_GOSHI tr A0A108ML7 A0A1081ALP4_PROSI tr A0A108ML7 A0A10841AP4_PROSI tr A0A108ML7 A0A10841AP4_PROSI tr A0A108ML7 A0A10841AP4_PROSI tr A0A108ML7 A0A10841AP4_PROSI tr A0A108ML7 A0A1080MAF5_GOSAR tr A0A080MAF8 A0A080MAF5_GOSAR tr A0A080MAF8 A0A080MAF5_GOSAR	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFHDEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSQIMSDAKKGFGEVIGKVSMKNIFGRKKLKKDGSEGKUTI EGGYSLEALTGDSQIMSDAKKGFGEVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTGDSVIPNAKPRFGEVIGKVSMKNIFGRKKLKKDGSEGKIMVI AGGYSLEALTSDPRVLGGTETKFGEVIGKVSMKNIFGRKKLKKDGSEGKUVI SGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKUVI AGGYSLEALTSDPRVLGGTETKFEAFLFGKISMKTIFGKKKLKKDGSEGKUVI GGYSLEALTSDPRVLGGTETKFEAFLGKISMKTIFGKKKKKDGSEGKUVI AGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI UGGYSLEALTGDRNVMQRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALGDRVMGRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALTGDRVMGRSTWRKFEKFLIGKTSMKTIFGKKKVKKDGSEGKMITI VGGYSLEALGDRVMSGAALGFEKFLIGKISMKTIFGQKKAKKDGSEGKMITI VGGYSLEALTGDRCVMSGAALGFEKFLIGKISMKTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRCVMSRAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRCVMSRAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRCVMSRAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI VGGYSLEALTGDRCVMSRAKLDFEKFLIGKVSMTTIFGQKKVKKDGSEGKMITI ******.**:*	555 505 505 505 505 505 417 431 571 549 576 576 576 576 576 576 576 576 576 576
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A630P58 A0A6J0P578 RAPSA sp Q84ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4 _ 9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A0B0MAF8 A0A60B0MAF8 GOSAR tr A0A0B0MAF8 A0A60B0MAF8 GOSAR tr A0A108KL7 A0A1U8KL62 _GOSHI tr A0A6J1FW48 A0A6J1FW48 _CUCMO tr A0A698E16 A0A6J1FW48 _CUCMO tr A0A698E16 A0A6371ED47 _MUCPR tr A0A888KCR7 A0A371ED47 _MUCPR tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A800J0J1 _A0A1090J1 _SOLCH tr A0A619UH37 A0A619UH37 _SESIN tr A0A630AE5 A0A800P55 _OLEEU tr A0A630AE5 A0A808VF65 _OLEEU tr A0A631AEP4 A0A631AEP4 _9ROSI tr A0A631AEP4 A0A631AEP4 _9ROSI tr A0A631AEP4 A0A631AEP4 _9ROSI tr A0A80MAF8 A0A0B0MAF8 _GOSAR tr A0A80MAF8 A0A0B0MAF8 _GOSAR tr A0A80MAF8 A0A0B0MAF8 _GOSAR tr A0A80MAF8 A0A0B0MAF8 _GOSAR tr A0A108KL7 A0A108KL7 _GOSHI tr A0A080MAF8 A0A080MAF8 _GOSAR tr A0A080MAF8 A0A080MAF	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGK I SMKT IFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGK I SMKT IFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSTVMRDARLVHAERLFHDEGLFGK I SMKT IFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSQMRDARLVHAERLFHDEGLFGK I SMKT IFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQMRDARLVHAERLFHDEGLFGK I SMKT IFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQMRDARLVHAERLFHDEGLFGK I SMKT IFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSQMRDARLVHAERLFHDEGLFGK I SMKT IFGRKKLKKDGTEGKVTVI AGGYSLEALTGDSQMRDARVHAKRF FGEL IGKVSMKT IFGRKKLKKDGSEGKLVI AGGYSLEALTSDPRVLGGTETK FERE IGKTSMKT IFGKKKLKKDGSEGKLVI AGGYSLEALTGDRVVLQGTETK FERE IGKTSMKT IFGKKKLKKDGSEGKLVI AGGYSLEALTGDRVVLQRSTWRK EKE IGKTSMKT IFGKKKLKKDGSEGKLVI AGGYSLEALTGDRVVLQRSTWRK EKE IGKTSMKT IFGKKKLKKDGSEGKMITI LGGYSLEALTGDRVVLQRSTWRK EKE IGKTSMKT IFGKKKLKKDGSEGKLTI UGGYSLEALTGDRVVLQRSTWRK EKE IGKTSMKT IFGKKKVKKDGSEGK IITI VGGYSLEALTGDRVVMQRSTWRK EKE IGKTSMKT IFGKKKVKKDGSEGK IITI VGGYSLEALTGDRVVMQRSTWRK EKE IGKTSMKT IFGGKKKVKDGSEGK IITI VGGYSLEALTGDRVVMRARKLD EKE IGKTSMKT IFGGKKVKKDGSEGK IITI VGGYSLEALTGDREVMSGAALG EKE IGKVSMTT IFGQKKVKKDGSEGKMTTI VGGYSLEALTGDREVMSRAKLD EKE IGKVSMTT IFGQKKVKNGSSEGKMTTI VGYSLEALTGDREVMSRAKLD EKE IGKVSMTT IFGQKKVKNGSSEGKMTTI VGYSLEALTGDREVMSKAKLD EKE IGKVSMTT IFGVSMYFDGVRKGSMYEFYEKY PSVEELQRTERE VI YSALDSISTLLYSSLKNLSKRVWTFDGVRKGSMYEFYEKY PSVEELQRTERE VI YSALDSISTLLYSSLKNLSKRVTFDGVRKGSMYE	555 505 505 505 505 417 431 571 549 576 576 576 576 576 578 495 459 436 475 436 475 613 563 577 543 563 577 543 563 577 543 563 577
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A619UH37 _SESIN tr A0A6050P58 A0A6J0P578 _RAPSA sp Q84ND9 POLIB _ARATH tr A0A6J1ALP4 A0A6J1ALP4 _ 9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A108MKL7 A0A1U8MKL7 _GOSHI tr A0A6J1FW48 A0A6J1FW48 _CUCMO tr A0A6J8EE16 A0A6548EE16 _PUNGR tr A0A1S3VPG4 A0A1S3VPG4 _TOBAC tr A0A1S3VPG4 A0A1S3VPG4 _TOBAC tr A0A8B8KCR7 A0A8B8KCR7 _ABRPR tr A0A1U8FC39 A0A1U7VL87 _NICSY tr A0A108FS A0A601PH45 _SIN tr A0A108FS A0A609UH37 _SESIN tr A0A108FS A0A609UH37 _SESIN tr A0A108FS A0A619UH37 _SESIN tr A0A609UH37 A0A619UH37 _SESIN tr A0A6019UH37 A0A619UH37 _SESIN tr A0A6019LF4 A0A601ALF4 _9ROSI tr A0A601ALF4 A0A601ALF4 _9ROSI tr A0A601ALF4 A0A601ALF4 _9ROSI tr A0A108KK7 A0A108KKT7 _GOSHI tr A0A108KK7 A0A108KKT7 _GOSHI tr A0A108KK7 A0A108KKT7 _GOSHI tr A0A108KK7 A0A108KKT7 _GOSHI tr A0A108KK7 A0A108KK77 _GOSHI tr A0A108KK76 A0A00B0MAF8 _GOSAR tr A0A108KK76 A0A601F448 _CUCMO tr A0A601ALF4 A0A601F448 _CUCMO tr A0A601ALF4 A0A601F448 _CUCMO tr A0A601ALF4 A0A601F448 _CUCMO tr A0A607 = EIA0A601F448 _CUCMO tr A0A607 = EIA0A607 = EIA0F tr A0A607 = EIA0A607 = EIA07 = EIA0F tr A0A607 = EIA0A607 = EIA07 = EIA0F tr A0A607 =	SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI SGGYSLEALTGDSTVMRDARPVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSTVMRDARLVHAERLFHGEGLFGKISMKTIFGRKKLKKDGTEGKVTVI LGGYSLEALTGDSSVMRDARLVHAERLFDGGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSSVMRDARLVHAERLFDGGLFGKISMKTIFGRKKLKKDGTEGKVTVI EGGYSLEALTGDSVIRDARLVHAERLFDGGLFGKISMKTIFGRKKLKKDGSEGKLVI AGGYSLEALTGDFVIJDAKRFFGELFIGKISMKTIFGRKKLKKDGSEGKLVI AGGYSLEALTSDPRVLGGTETKFGELFIGKISMKTIFGKKKLKKDGSEGKLVI AGGYSLEALTGDRNVMQRSTWRKFEALFGKISMKTIFGKKKLKKDGSEGKLVI LGGYSLEALTGDRNVMQRSTWRKFEKLJIGKTSMKTIFGKKKLKKDGSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKLJIGKTSMKTIFGKKKVKNGSSEGKMITI LGGYSLEALTGDRNVMQRSTWRKFEKLJIGKTSMKTIFGKKKVKNGSSEGKMITI VGGYSLEALTGDRNVMQRSTWRKFEKLJIGKISMKTIFGKKKVKNGSSEGKMITI VGGYSLEALTGDRNVMQRSTWRKFEKLJIGKISMKTIFGKKKVKNGSSEGKMITI VGGYSLEALTGDRNVMQRSTWRKFEKLJIGKISMKTIFGKKKVKNGSSEGKMITI VGGYSLEALTGDRVMARAKLDFEKLJIGKISMKTIFGKKKVKNGSSEGKMITI VGGYSLEALTGDRVMARAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDRVMARAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDRVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDRVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGYSLEALTGDREVMSRAKLDFEKLJIGKVSMTTIFGQKKVKNGSSEGKMTI VGYSLEALTGDREVMSRAKLD	555 505 505 505 417 431 549 576 576 576 576 576 576 578 495 459 436 475 459 436 475 613 613 613 563 577 543 563 477 491 631 609 636 638 638 6638
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UB37 A0A619UB37 _SESIN tr A0A630P58 A0A630P558 _RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1A1P4 A0A6J1A1P4 _ 9ROSI tr A0A1U8MKL7 A0A1U8MK17 _GOSHI tr A0A108MK17 A0A1U8MK17 _GOSHI tr A0A108MK17 A0A108MK17 _GOSHI tr A0A108MK17 A0A108MK17 _GOSHI tr A0A6J1FW48 A0A6J1FW48 _CUCMO tr A0A698EE16 A0A659EE16 _PUNGR tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A109 _ A0A100 _ A0A109 _ A0A100 _ A0A100 _ A0A1000 _ A0A1000 _ A0A100 _ A0A1000 _ A0A10000 _ A0A10000 _ A0A10000 _ A0A100000 _ A0A1000000 _ A0A10000000000	SGGY SLEALTGD STVMRDAR PVHAERLFHGE GLFGK I SMKT I FGRKKLKKDGTEGKVTVI SGGY SLEALTGD STVMRDAR PVHAERLFHGE GLFGK I SMKT I FGRKKLKKDGTEGKVTVI LGGY SLEALTGD STVMRDAR LVHAERLFHGE GLFGK I SMKT I FGRKKLKKDGTEGKVTVI LGGY SLEALTGD SVMRDAR LVHAERLFD GD CLFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDAR LVHAERLFD GD CLFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDAR LVHAERLFD GD CLFGK I SMKT I FGRKKLKKDGSEGK ITT I EGGY SLEALTGD SVMRDAR LVHAERLFD GD CLFGK I SMKT I FGRKKLKKDGSEGK ITT I SGGY SLEALTGD FVI FDNAKFR FERE LGK I SMKT I FGRKKLKKDGSEGK ITT I GGY SLEALTGD FVU GGTETK EAELFGK I SMKT I FGRKKLKKDGSEGK ITT I LGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I LGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I LGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I LGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I VGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I VGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I VGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I VGGY SLEALTGD RNVMORSTWRK EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I VGGY SLEALTGD RNVMORSTRKLD EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I VGGY SLEALTGD RGVMS GALG EKEL LGKT SMKT I FGRKKVKKDGSEGK ITT I VGGY SLEALTGD RGVMS RAKLD	555 505 505 505 417 431 549 576 576 576 576 576 578 495 459 436 475 613 613 613 563 577 543 563 577 543 563 477 491 631 609 636 638 6586 638 6584
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A0V0J0J1 A0A0V0J0J1 _SOLCH tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UB37 A0A619UH37 _SESIN tr A0A6300578 A0A6300578 RAPSA sp 084ND9 POLIB_ARATH tr A0A6J1ALP4 A0A6J1ALP4 _9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A108MK17 A0A1U8MK17 _GOSHI tr A0A108MK17 A0A1U8MK17 _GOSHI tr A0A108MK17 A0A108MK17 _GOSHI tr A0A698EE16 A0A651FW48 _CUCMO tr A0A698EE16 A0A652077M1 _CICAR tr A0A133YPG4 A0A1S3YPG4 _TOBAC tr A0A133YFG4 A0A1S3YPG4 _TOBAC tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A107VL87 A0A107VL87 _MICSY tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A613LP44 _9ROSI tr A0A613LP44 A0A613LP44 _9ROSI tr A0A108KL7 A0A108MK17 _GOSHI tr A0A108KJ62 A0A108KJ62 _GOSHI tr A0A1032Y7M1 A0A108XY7M1 _CICAR tr A0A132Y7M1 A0A123Y7M1 _CICAR	SGGY SLEALTGD STVMRDAR PVHAERLFHGE GLFGK I SMKT I FGRKKLKKDGTEGKVTVI SGGY SLEALTGD STVMRDAR PVHAERLFHGE GLFGK I SMKT I FGRKKLKKDGTEGKVTVI LGGY SLEALTGD STVMRDAR LVHAERLFHGE GLFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDAR LVHAERLFD GD LJFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDAR LVHAERLFD GD LJFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDAR LVHAERLFD GD LJFGK I SMKT I FGRKKLKKDGSEGK I TI EGGY SLEALTGD SVMRDAR LVHAERLFD GD LJFGK I SMKT I FGRKKLKKDGSEGK I TI EGGY SLEALTGD FVI LGTETK EGEL IGKI SMKT I FGRKKLKKDGSEGK I TI AGGY SLEALTGD RVUGGTETK EKEL IGKI SMKT I FGKKKLKKDGSEGK I TI LGGY SLEALTGD RVUGGTETK EKEL IGKI SMKT I FGKKKLKKDGSEGK I TI LGGY SLEALTGD RVVNQRSTWRK EKEL IGKI SMKT I FGKKKVKKDGSEGK I TI LGGY SLEALTGD RVVNQRSTWRK EKEL IGKI SMKT I FGKKKVKKDGSEGK I TI LGGY SLEALTGD RVVNQRSTWRK EKEL IGKI SMKT I FGKKKVKKDGSEGK I TI VGGY SLEALTGD RVVNQRSTWRK EKEL IGKI SMKT I FGKKKVKKDGSEGK I TI VGGY SLEALTGD RVVNQRSTWRK EKEL IGKI SMKT I FGKKKVK KDGSEGK I TI VGGY SLEALTGD RVVNGRALG EKEL IGKI SMKT I FGKKKVK KDGSEGK I TI VGGY SLEALTGD RVMSGALG EKEL IGKI SMKT I FGGKKVK KDG SEGK I TI VGGY SLEALTGD RVMSGALL EKEL IGKI SMKT I FGGKKVK KDG SEGK I TI VGGY SLEALTGD RVMSRARLD EKEL IGKI SMKT I FGGKKVK KDG SEGK I TI VGGY SLEALTGD RGVMSRARLD EKEL IGKI SMKT I FGGKKVK KDG SEGK I TI VGGY SLEALTGD RGVMSRARLD EKEL IGKI SMKT I FGGKKVK KDG SEGK I TI VGGY SLEALTGD RGVMSRARLD EKEL IGKI SMTTI FGGKKVK KDG SEGK I TI VGGY SLEALTGD RGVMSRARLD	555 505 505 505 417 431 571 549 576 576 576 576 578 495 459 436 475 613 613 613 563 577 543 563 563 577 543 563 563 577 543 563 563 577 543 563 563 577 543 563 563 555 576 576 576 576 576 576 576 576 576
tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A1U7VL87 A0A1U7VL87 _NICSY tr A0A1U8FC39 A0A1U8FC39 _CAPAN tr A0A619UB37 A0A619UB37 _SESIN tr A0A6300P58 A0A630P578 _RAPSA sp 084ND9 POLIB_ARATH tr A0A631ALP4 A0A631ALP4 _ 9ROSI tr A0A1U8MKL7 A0A1U8MKL7 _GOSHI tr A0A108MK17 A0A1U8MK17 _GOSHI tr A0A631FW48 A0A631FW48 _CUCMO tr A0A698EE16 A0A6327M1_CICAR tr A0A698EE16 A0A6327M1_CICAR tr A0A1S3YPG4 A0A1S3YPG4 _TOBAC tr A0A888KCR7 A0A888KCR7 _ABRPR tr A0A109 F A0A107VL87 _MUCPR tr A0A619UH37 A0A107VL87 _MUCPR tr A0A619UH37 A0A107VL87 _ABRPR tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A619UH37 _SESIN tr A0A619UH37 A0A613LP4 _ 9ROSI tr A0A613LP4 A0A631ALP4 _ 9ROSI tr A0A108KL7 A0A108KL7 _GOSHI tr A0A108KL7 A0A108KL7 _GOSHI tr A0A108KJ62 A0A108SYP64 _CUCMO tr A0A613LP4 A0A631ALP4 _ 9ROSI tr A0A108KJ62 A0A108KJ62 _GOSHI tr A0A108KJ62 A0A108KJ62 _GOSHI tr A0A118KJ62 A0A108KJ62 _GOSHI tr A0A118KJ62 A0A108KJ62 _GOSHI tr A0A118KJ62 A0A108SY7M1_CICAR tr A0A118KJ62 A0A108SY7M1_CICAR tr A0A118XJ64 A0A531LP4 _M0A53 tr A0A118KJ64 A0A531LP4 _M0A54 tr A0A631LP4 A0A631LP4 _M0A54 tr A0A631LP4 A0A5431LP4 _M0CPR tr A0A88KCR7 A0A88KCR7 ABRPR	SGGY SLEALTGD STVMRDAR PVHAERLFHGEGLFGK I SMKT I FGRKKLKKDGTEGKVTVI SGGY SLEALTGD STVMRDAR PVHAERLFHGEGLFGK I SMKT I FGRKKLKKDGTEGKVTVI LGGY SLEALTGD SVMRDARLVHAERLFHDEGLFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDARLVHAERLFHDEGLFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDARLVHAERLFHDEGLFGK I SMKT I FGRKKLKKDGTEGKVTVI EGGY SLEALTGD SVMRDARLVHAERLFHDEGLFGK I SMKT I FGRKKLKKDGTEGKLVVI AGGY SLEALTGD SVI PNAKPR FGE VIGKVSMKNI FGRKKLKKDGSEGK IIVI EGGY SLEALTGD PRVLGGTETK EALFGKI SMKTI FGRKKLKKDGSEGKLVI AGGY SLEALTGD FRVLGGTETK EALFGKI SMKTI FGRKKLKKDGSEGK IIVI LGGY SLEALTGD RNVMORSTWRK EKELIGKTSMKTI FGRKKVKKDGSEGK IITI LGGY SLEALTGD RNVMORSTWRK EKELIGKTSMKTI FGRKKVKKDGSEGK IITI LGGY SLEALTGD RNVMORSTWRK EKELIGKTSMKTI FGRKKVKKDGSEGK IITI LGGY SLEALTGD RNVMORSTWRK EKELIGKTSMKTI FGRKKVKKDGSEGK IITI UGGY SLEALTGD RNVMORSTWRK EKELIGKTSMKTI FGRKKVKKDGSEGK IITI VGGY SLEALTGD RNVMORSTWRK EKELIGKTSMKTI FGRKKVKKDGSEGK IITI VGGY SLEALTGD RNVMORSTWRK EKELIGKTSMKTI FGRKKVKKDGSEGK IITI VGGY SLEALTGD RVMSARALD FKELIGKTSMTTI FGGKKVKKDGSEGK IITI VGGY SLEALTGD REVMSARALD FKELIGKVSMTTI FGGKKVKKDGSEGK IITI VGGY SLEALTGD REVMSARALD FKELIGKVSMTTI FGGKKVKKDGSEGKMTTI VGGY SLEALTGD REVMSARALD HEKDLIGKVSMTTI FGGKKVKKDGSEGKMTTI VGGY SLEALTGD REVMSARALD HEKDLIGKVSMTTI FGGKKVKKDGSEGKMTTI VGGY SLEALTGD REVMSARALD HEKDLIGKVSMTTI FGGKKVKKDGSEGKMTTI VGGY SLEALTGD REVMSARALD HEKDLIGKVSMTTI FGGKKVKKDGSEGKMTTI VGGY SLEALTGD REVMSKAKLD HEKDLIGKVSMTTI FGGKKVKNDGSEGKMTTI VGGY SLEALTGD REVMSKAKLD HEKDLIGKVSMTTI FGGKKVKNDGSEGKMTTI VGGY SLEALTGD REVMSKAKLD HEKDLIGKVSMTTI FGGKKVKNDGSEGKMTTI VGGY SLEALTGD REVMSKAKLD FKELPSLKSKSKNMTT FGGKSKMTTFYEY PVVELQRE	555 505 519 485 505 417 431 571 549 576 526 578 495 436 475 459 436 475 459 436 475 613 613 563 577 543 563 577 543 563 577 543 563 577 499 631 609 636 636 586 584 554 554 554 554 554 554 554 554 554

	Exo 🔶 🚬 Pol	
tr A0A1S3YPG4 A0A1S3YPG4 TOBAC	TOLROLFFGGIONRKNSDESLPYEKEFKVPINVDKVIEFGKKAPTKFRKIRLHRICD-LID	732
tr A0A1U7VL87 A0A1U7VL87 NTCSY	TOLROLFFGGIONRKNSDESLPYEKEFKVP NVDKVIEEGKKAPTKFRKIRLHRICD-LID	732
tr A0A0V0J0J1 A0A0V0J0J1 SOLCH	TOLROLFFGGIONRRNVDESLPNEKEFKVP NVDKVIEEGKKAPTKFRKIHLHRICD-PIN	682
tr 202108FC39 202108FC39 C2P2N	TOLROLFFGGTONRKNIDECLPTEKEFKVP NVDKVIFEGKKAPTKFRKIOLHRICD-PID	696
tr 20261901037 20261901037 SESTN	LOLROLFFGGVONSKDONFFLEVER VERVER NVDNY I FFGKKNPTKVRKI TI.BKPDGVNIF	663
trla0a850VP651a0a850VP65_OLEFU	AOLOOFFGGVOND-NANFFLOLFKNEKTE NYDKITEFGKKADYKYDKITLHKDLOVNIF	682
tr A0A6J0P5F8 A0A6J0V105_00EE0	TOLDOL FEGGISSIST GEDL DEFEL FEVD NUDNUT FEGEVED AVERULATION DUD	596
ED LOS ANDS LOOL TR ADATH	TOLDOL PECCISNSON OLDALI PECKINKYI WYAYI PECKKANA KATALAN MATADA DA	610
tyla0afiate4la0afiate4 0post	TO DATE VOLTANSKA DEDETE FERTEN UT VERTENSKA AND STATE AND	750
LI AGAGGIADET AGAGGIADET SKOSI	FOLDOLING CALVER AND A CALVERY AND AND A CALVERY AND A CALVE	730
CT AUAIUSMKL/ AUAIUSMKL/ GOSHI	TOLEQUE GGENSADPNVSEPDEKTERTP NVDKVIEEGKKAPTKERNIKLOSIGV-KEP	720
tr AUAUBUMAF8 AUAUBUMAF8_GOSAR	TOLROLLIGGILNSKDPNVSLPEEKTFKIP NVDKVIEEGKKAPTKFKNIKLCSIGV-KLP	/ 3 3
tr AUAIU8KJG2 AUAIU8KJG2 GOSHI	TQLRQLLYGGLLNSKDPNVSLPGEKTFKIP NVDKVIEEGKKAPTKFKNIKLCSIGV-KLP	/ 5 5
tr AUA6J1FW48 AUA6J1FW48_CUCMO	AQLRQLLFGGTSNSKNPDESLPTERTFKVPNTENVIEEGKKTPSKFRNINLKRISVEDLS	706
tr A0A6P8EE16 A0A6P8EE16_PUNGR	TQLRQLFFGGILNSKDPNEALPEEKTFKIT NDDGFIEEGKKKASKYRNITLRSIGV-KFP	757
tr A0A1S2Y7M1 A0A1S2Y7M1_CICAR	AQLRVLLFGGTANRKNHNEAIPTERIFKVP NVDKVLEEGKKTPLKLRAIKLNSIGY-NLN	673
tr AUAIS3VL91 AUAIS3VL91_VIGRR	AQLRVLLFGGTINRKNHNDDIPTERIFKIP NVDKVIAEGKKAPTKFCDIKLNSIGY-KLN	638
tr A0A371ED47 A0A371ED47_MUCPR	AQLRVLLFGGTVNRKNPDEAIPTERIFKIP NVDKVIEKGKKAPTKFRDVKLNSIGY-NLN	615
tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	AQLRVLLFGGTVNRKNHNEVIPTERIFKIP NLDKVIEEGKKAPTKFRDIKMNSIGY-NLN	654
	·*** ····** * · · · · ** * * * * * ·	
tr A0A1S3YPG4 A0A1S3YPG4 TOBAC	NEVPSCEPEVSIYGSAYNAFGGG-QKGIE <mark>ACH</mark> AIAAI <mark>CEMCS</mark> IDSLISNFILPLQGDDVS	844
tr A0A1U7VL87 A0A1U7VL87 NICSY	NEIPSCEPEVSIYGSAYNAFGGG-QKGIEACHAIAAI <mark>CEMCS</mark> IDSLISNFILPLQGDVS	844
tr A0A0V0J0J1 A0A0V0J0J1 SOLCH	NEALGONPEISAYGTAYHAFGGG-QKGIEACHAIAAI <mark>CEVCS</mark> IDSLISNFILPLOGHDVS	794
tr A0A1U8FC39 A0A1U8FC39 CAPAN	NEALSCNPEVSAYGTAYHAFGGG-QKGIEACHAISAICEVCSIDSLISNFILPLOGDDVS	808
tr A0A619UH37 A0A619UH37 SESTM	KNTAGIGIDTSACGAAYSAFGGG-OAGIFACHATAATCEVCETDSLISNELLPLOCNETS	771
tr A0A8S0VP65 A0A8S0VP65 OLEFU	AWNDPDKSDYGAAYTAFGGG-OAGVEACHATAATCRVCSTDSLISNFILPLOGNHIS	784
+r 2026.T0P5F8 2026.T0P5F8 2026.T	TSWESDIGATIALOG WAS LEVELATION OF DELISITIED WORKED	712
sp Q84ND9 POLIB_ARATH	HVNTSVESDTSAYGTAFDAFGGG-ESGKEACHAIAALCEVCSIDSLISNFILPLOGSNVS	728
tr A0A6J1ALP4 A0A6J1ALP4 9ROSI	CPEMVIDVDTSAYGTAFAAFGDE-EKGRE <mark>ACHAIASICEVCS</mark> IDSLISNFILPLOGENVS	850
tr A0A1U8MKL7 A0A1U8MKL7 GOSHT	YPETMTAVDKSAYGTAFAAFEDE-EKGREACHAIASICEVCSIDSLISNFILPLOGSNVS	82.8
TT A0A0B0MAF8 A0A0B0MAF8 GOSAP	YPETMIAVDISAYGTAFAAFEDE-EKGREACHAIASICEVCEIDSLISNEILPLOCENVS	855
	Y DETMINUTS A GTA FAA FF DE - FKGERECKA TA SICEWC ST DSLT SNETL PLOGENVS	855
tr 1000611 FW48 1000611 FW48 CUCMO		910
tyla0A601FW40 A0A601FW40_COCMO	REFINESANTSDIGAALAAFGSSELKONDAALAALG VCSIDSHISNETHELQOPNIS	019
LI AUAGPOLLIG AUAGPOLLIG_FUNGK		770
LI AUAISZI MI AUAISZI MI_CICAR		779
tr AUAIS3VL91 AUAIS3VL91_VIGRR	SEVLPERIDKSAYGTAPSCFPTD-QEGREACHATAATCEVCSIDSLISNFLPLQGENIS	744
tr AUA3/IED4/AUA3/IED4/_MUCPR	SEVLPAETDKSAYGTAFSAFPTE-EEGREACHATAATCEVCSIDSLISNFILPLQGHNIS	721
tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	NEVAPVEIDNSAYGTALNAFPTY-EEGRE <mark>AC</mark> HAIAAH <mark>CE</mark> YCSIDSLISNFILPLQGHNIS	760
	* *:* .* : * * <mark>****::******************</mark> .:*	
trla0a1s3ypg4la0a1s3ypg4 TOBAC	GENGETH CSTINTNTETGELSAREPNI.ONOPALEKDEYKTROAFVAAOGNSLTVADYGOLE	904
+ - ΙΔΟΔ1117VT.87 ΙΔΟΔ1117VT.87 NTCSV	GENGE THE STATE TO BE DATE OF DE DATE OF DATE	904
	CENCE II CONTRACTOR SADDINI ONDAT PUDDVITI OVEVA SECNET IVADVOT P	054
LIAOAUVOODOI AOAUVOODOI _SOLCH		0.54
tr AUAIUSFC39 AUAIUSFC39 CAPAN	GENGRIHCELNINTETGRISARRPNLONOPALERDRYKIROAFVAAEGNSLIVADYGQLE	868
tr AUA619UH37 AUA619UH37_SESIN	GKNGRIHCSININTETGRLSARRPNLQNQPALEKDRYKIRQAFIAAPGNSLIVADYGQLE	831
tr A0A8S0VP65 A0A8S0VP65_OLEEU	GKDRRIH <mark>C\$</mark> LNINTETGRLSARRPNLQNQPALEKDRYKIRQAFIAAPDNSLIVADYGQLE	844
tr A0A6J0P5F8 A0A6J0P5F8_RAPSA	GKDGRVH <mark>C\$</mark> ININTETGRLSARRPNLQNQPALEKDRYKIRQAFIASPGNSLIVADYGQLE	772
sp Q84ND9 POLIB_ARATH	GKDGRVHCSININTETGRLSARRPNLONOPALEKDRYKIROAFIASPGNSLIVADYGOLE	788
tr A0A6J1ALP4 A0A6J1ALP4_9ROSI	GKSGRVH <mark>C\$</mark> ININTETGRLSARRPNLQNQPALEKDRYKIRQAFVAAPGNSLIVADYGQLE	910
tr A0A1U8MKL7 A0A1U8MKL7 GOSHI	GKGGRVHCSLNINTETGRLSARRPNLQNQPALEKDRYKIRQAFVAAPGNSLVVADYGQLE	888
tr A0A0B0MAF8 A0A0B0MAF8_GOSAR	GKGGRVH <mark>C\$</mark> LNINTETGRLSARRPNLQNQPALEKDRYKIRQAFVAAPGNSLVVADYGOLE	915
tr A0A1U8KJG2 A0A1U8KJG2_GOSHI	GKGGRVHCSLNINTETGRLSARRPNLONOPALEKDRYKIROAFVAAPGNSLVVADYGOLE	915
tr A0A6J1FW48 A0A6J1FW48 CUCMO	GENGRIHCSLNINTETGRLSARRPNLONOPALEKDRYKTROAFTAAPGNSLTVADYGOLF	879
trla0a6p8FF16la0a6p8FF16 pingp	GENGEVICSINTNEETGDI.SADDONI.ONODAI.FEDDYETDOFTAADGNSI.TVADVGOI.F	930
+r A0A192V7M1 A0A0F0EET0_F0NGR	GKDLDVHCSTNTNTFTGDLSADDDSLONODALEKDATATAACMSLTVADVCOTE	030
triacaiozi mijacaiozi mi_cicak	CADDD IN TWINE CONDAUCED ON DD OL ON DD OL DAD DAD ON DD OL DAD AND AND AND AND AND AND AND AND AND	0.0.0
	GRUDRINGS IN INTETGRUSARRPSLONOPALERDRINIR OAF HAAPGNSLIVADIGOLE	804
tr AUA3/IED4/AUA3/IED4/_MUCPR	GKDHRVHCSININTETGRLSARRPNLQNQPALEKDRYKIRQAFIAAPGNSLIVADYGQLE	/81
tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	GKDQRVHCSININTETGRLSARRPNLQNQPALEKDRYKIRQAFIAAPGNSLIVADYGQLE	820
	*:. *:***:*****************************	
trla0a1S3ypg4la0a1S3ypg4 mora	RKFFARFFRRUHTLIGRARWFPSUKNATGSUKGHTFDAATNTDUGGSAADUAMCAMTDTG	1094
T AVAIDSIFGT AVAIDSIFGT TOBAC	REFERENCES AND A CONDUCTION AND CONTRACT AND A CONTRACT	1004
CTIAUALU/VL8/IAUALU/VL8/_NICSY	KKFLAKEFKKVHTLLGKAKWFF5VKNATG5VKGH1ERAAINTPVQGSAADVAMCAMLEIS	1084
tr AUAUVUJUJI AUAUVUJUJI_SOLCH	KKFEAREFGRVHTLLGRARWFPSVKNATGSVKGHIERAAINTPVQGSAADVAMCAMLEIS	1034
tr AUA1U8FC39 A0A1U8FC39_CAPAN	RRFEAREFGRVHTLLGRARWFPSVKNATGSVKGHIERAAINTPVQGSAADVAMCAMIEIS	1048
tr A0A6I9UH37 A0A6I9UH37_SESIN	RKKEARKYRSVHTLLGRARHFPSLKNASSAHRAHIERAAINTPVQGSAADVAMCAMLQIS	1011
tr A0A8S0VP65 A0A8S0VP65_OLEEU	RKKEAVVDGCVHTLLGRARTFPSLRNATPWHKSHIERAAINTPVQGSAADVAMCAMLEIS	1024
tr A0A6J0P5F8 A0A6J0P5F8 RAPSA	RKKEAIQNGYVLTLLGRARKFPAYRSRAQKNHIERAAINTPVQGSAADVAMCAMLEIT	950
sp Q84ND9 POLIB ARATH	RKKEAIQKGYVLTLLGRARKFPEYRSRAQKNHIERAAINTPVQGSAADVAMCAMLEIS	966
tr A0A6J1ALP4 A0A6J1ALP4 9ROST	RKYEAOKLRRVKTLLGRARLFPSYAHATRAOKGHIERAAINTPVOGSAADVAMCAMLOIS	1090
tr A0A1U8MKL7 A0A1U8MKL7 GOSHT	RKFEAOKFGCVKTLLGRARRFPSFAHCTRAOKGHTFRAAINTPVOGSAADVAMCAMLOIS	1068
triananenare lananenare cocar	DREEY OR ECONRALT COVED BEGEN ROAD YOK CALED YY LMEDNOCS Y DRYNGYM O'L C	1005
LI AUAUDUMARO JAUAUDUMARO GOSAR	REEAVEGOVETHUGKAKKEFSFARGTKAVEGHEKAAINTPVQGSAADVAMCAMLQIS	T0 82
LI AUAIUSKUGZ AUAIUSKUGZ GOSHI	KKELAQKEGUVKTLLGKAKKEFSFAHUTRAQKGHLERAAINTPVQGSAADVAMCAMLQIS	1092
tr AUA6J1FW48 A0A6J1FW48_CUCMO	KKKEAEEKSCVRTLLGRARRFPSMTHATRAHKGHIERAAINTPVQGSAADVAMCAMLEIS	1059
tr A0A6P8EE16 A0A6P8EE16_PUNGR	RKQEAREHGQVYTLLGRARRFPSFSRASYAQKGHIERAAINTPVQGSAADVAMSAMLQIF	1110
tr A0A1S2Y7M1 A0A1S2Y7M1_CICAR	RKKEAFEFGCVYTLLGRARQFPEINQGQNYYKGHIERAAINTPVQGSAADVAMLAMLEIS	1019
tr A0A1S3VL91 A0A1S3VL91 VIGRR	RKKEACEFQRVHTLLGRARRFPKIDPDHSSHRGHIERAAINTPVQGSAADVAMCAMLQIS	984
tr A0A371ED47 A0A371ED47 MUCPR	RKKEACESQCVHTLLGRARRFPKIDYAHSYQKGHIERAAINTPVQGSAADVAMCAMLOIS	961
tr A0A8B8KCR7 A0A8B8KCR7 ABRPR	RKKEACESQCVHTLLGRARRFPKIDPAYSYOKGHIERAAINTPVOGSAADVAMCAMLOIS	1000
	4. 44 4 444444 44 . 44444444444 444	

//End of DNA pol IB sequences		
tr A0A1S3YPG4 A0A1S3YPG4_TOBAC	CARNWYSAR	1152
tr A0A1U7VL87 A0A1U7VL87 NICSY	CAKNWYSAK	1152
tr A0A0V0J0J1 A0A0V0J0J1_SOLCH	CAKNWYSAK	1102
tr A0A1U8FC39 A0A1U8FC39 CAPAN	CAKNWYAAK	1116
tr A0A6I9UH37 A0A6I9UH37 SESIN	CACNWYSAK	1079
tr A0A8S0VP65 A0A8S0VP65 OLEEU		1058
tr A0A6J0P5F8 A0A6J0P5F8 RAPSA	CAONWYAAK	1018
sp Q84ND9 POLIB ARATH	CAQNWYAGK	1034
tr A0A6J1ALP4 A0A6J1ALP4_9ROSI	CAONWYAAK	1159
tr A0A1U8MKL7 A0A1U8MKL7 GOSHI	CAQNWYAAK	1136
tr A0A0B0MAF8 A0A0B0MAF8_GOSAR	CAONWYAAK	1163
tr A0A1U8KJG2 A0A1U8KJG2 GOSHI	CACNWYAAK	1163
tr A0A6J1FW48 A0A6J1FW48_CUCMO	CARNWYSAK	1127
tr A0A6P8EE16 A0A6P8EE16 PUNGR	CAONWYAAK	1179
tr A0A1S2Y7M1 A0A1S2Y7M1_CICAR	CAONWYAAK	1087
tr A0A1S3VL91 A0A1S3VL91 VIGRR	CAQNWYAAK	1052
tr A0A371ED47 A0A371ED47_MUCPR	CAONWYAAK	1029
tr A0A8B8KCR7 A0A8B8KCR7_ABRPR	CACNWYTAK	1068
_		

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

		2.0.1
sp/P00282/DPOI_ECOLI	AWIAKLEKAPVFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHDYLD	391
sp F4I6M1 POLIA ARATH	DTVAKLVNQFRNHVH <mark>SCDTE</mark> VSGIEVKEETPVDHGELI <mark>C</mark> FSIY <mark>C</mark> GPEADFGNGKS <mark>C</mark> IWVD	336
sp/084ND9/POLIB ARATH	ETVALLMNÖYRNLVHA <mark>CDTE</mark> VSRIDVKTETPVDHGEMICFSIYCGSEADFGDGKSCIWVD	314
	** ************************************	
sp P00582 DP01 ECOLI	APDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAFDTMLESYILNS	451
sp F4I6M1 POLIA ARATH	VLGE-NGREVLAEFKPYFEDSFIRKVWHNYSEDSHIIRNHGIEISGFHADTMHMARLWDS	395
SDIO84ND9 POLTE ARATH	VLGE-NGRDILAEFKPFFEDSSIKKVWHNYSFDNHIIRNYGIKLSGFHGDTMHMARLWDS	373
spige mostronin_main		575
	Exo 🖌 👝 Pol	
		5.0.4
sp P00582 DP01_ECOLI	FNQIALEEAGRYAAEDADVTLQLHLKMWPDLQ KHKGPLNVFENIEM	531
sp F4I6M1 POLIA ARATH	VEELQREDREAWIS <mark>Y</mark> SAI <mark>D</mark> AISTLKLYESMTKKLQLMDWHLDGKPV LGRTMLDFYHEFWR	515
sp 084ND9 POLIB_ARATH	VKELOMEDREAWIS <mark>Y</mark> SAIDSISTLKLYESMKKOLOAKKWFLDGKLI SKKNMFDFYOEYWO	493
- · · · _	···~ *· * * · ** · ** · · · · · · · · ·	
	•••••••••••••••••••••••••••••••••••••••	

sp P00582 DP01 ECOLI	LEYR	GLAKLKSTYTD-KLPL	646
sp F4I6M1 POLIA ARATH	SETQKSKTDDETDTSAYGTAYVAFGGGERGKEACHA	AIASL <mark>C</mark> EV <mark>C</mark> SIDSLISNFILPLQGS	741
sp Q84ND9 POLIB_ARATH	VETQHVNTSVESDTSAYGTAFDAFGGGESGKEACH	AIAALCEVCSIDSLISNFILPLQGS	725
sp P00582 DP01 ECOLI	MINPKTGRVHTSYHQAVTATGRLSSTDPNLQNIPVH	RNEEGRRIRQAFIAPEDYVIVSADY	706
sp F4I6M1 POLIA_ARATH	NVSGKDGRVH <mark>C</mark> SLN-INTETGRLSARRPNLQNQPAI	LEKDRYKIRKAFVASPGNTLVV <mark>ADY</mark>	800
sp Q84ND9 POLIB_ARATH	NVSGKDGRVH <mark>C</mark> SLN-INTETGRLSARRPNLQNQPAI :. * ******** * * *****: ***** *. — — —	LEKDRYKIRQAFIASPGNSLIVADY ::: :**:**:* . :: ***	784
sp P00582 DP01_ECOLI	TVTSEQRRSAKAINFGLIYGMSAFGLAR	QLNIPRKEAQKYMDLYFERYPGVLE	<mark>800</mark>
sp F4I6M1 POLIA_ARATH	PPVPLLKDAFGSE <mark>R</mark> RKA <mark>K</mark> MLNFSIA <mark>YG</mark> KTAVGLSRI	DWKVSTKEAQETVDLWYNDRQEVRK	920
sp Q84ND9 POLIB_ARATH	PPVPLLKDAFASE <mark>RRKAK</mark> MLNFSIA <mark>YG</mark> KTAIGLSRI	DWKVSREEAQDTVNLWYNDRQEVRK : :: :***. ::*::: * :	904
sp P00582 DP01 ECOLI	YMERTRAQAKEQGYVETLDGRRLYLPDIKSSNGAR	RAAAERAAINAPMQGTAADIIKRAM	860 860
sp F4I6M1 POLIA_ARATH	WQEMRKKEAIEDGYVLTLLGRSRRFPASKSRAQH	RNHIQRAAINTPVQGSAADVAMCAM	978
sp Q84ND9 POLIB_ARATH	WQELRKKEAIQKGYVLTLLGRARKFPEYRSRAQF : * : :* :.*** ** ** :* :* :* *: 	<pre>KNHIERAAINTPVQGSAADVAMCAM : :*****:*:**:****: **</pre>	962
sp P00582 DP01_ECOLI	IAVDAWLQAEQPRVRMIMQVHDELVFEVHKDDVDAV	VAKQIHQLMENCTRLDVPLLV	916
sp F4I6M1 POLIA_ARATH	LEISINQQLKKLGWRLLLQI <mark>HDE</mark> VILEGPIESAEI#	AKDIVVDCMSKPFNGRNILSVDLSV	1038
sp Q84ND9 POLIB_ARATH	LEISNNQRLKELGWKLLLQV HDE VILEGPSESAENA : :. : :: :::*: <mark>***</mark> :::* ::	AKDIVVNCMSEPFNGKNILSVDLSV : : :: *.* * *	1022
sp P00582 DP01 ECOLI	EVGSGENWDQAH 928		
sp F4I6M1 POLIA ARATH	DAKCAQNWYAAK 1050		
sp Q84ND9 POLIB_ARATH	DAKCAQNWYAGK 1034 ::**		

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 metalbound 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 $3' \rightarrow 5'$ exonuclease \leftrightarrow 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 SS Y ⁴⁷⁰ S→3 aa→ D ⁴⁷⁴ AI-
A. Thaliana DNA pol IB	- D ²⁷² TE ²⁷⁴ F D ³⁴⁶ NS Y ⁴⁴⁸ S→3 aa→ D ⁴⁵² SI-
<i>E. coli</i> DNA pol I	$-D^{355}\mathbf{T}E^{357}YD^{424}RR\mathbf{Y}^{497}A\rightarrow 3 \text{ aa}\rightarrow D^{501}AD-$

(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].



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**⁻⁴KLV**K**Q¹TVMTSV**Y**⁸**G**-, is very similar to the active site of *E. coli* DNA pol I, -**R**⁻⁴RSA**K**⁷⁵⁸A¹INFGLI**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

triD7SML31D7SML3_VITVI

		PR Exo	
	4		FOC
TT AUA251TQF3 AUA251TQF3 HELAN			506
+~ 1303565c0271303565c027 00038			510
++ 1V41.7M1 V41.7M1 FUTSA		ALDILGNIKWRYNRRILDYYERLWADGOLAGLYNREDYF ALDTLGNTKWDVNDKILDYYERLWADGOLAGLYNDFDVD	512
tr DOHHGE DOHHGE GBDAS		ALDTLGNTKWRVNRKILDVVERLWADGGNIAGLVNREDVI	510
splo246001BPor3 ARATH		ALDTLGNTKWRVNRNII.DVVFRLWADGGNIAGLVNREDVI	512
tr D7LGB7 D7LGB7 ABALL		ALDTLGNTKWRVNRNILDVVERLWADGGNIAGLVNREDVP	509
tr A0A078G6J8 A0A078G6J8 BRANA		ALDTLGNTKWRVNRKILDVVERLWADGGNIAGLVNREDVP	495
tr A0A6J0K1N7 A0A6J0K1N7 RAPSA		ALDTLGNTKWRVNRKILDVVERLWADGGNIAGLVNREDVP	501
tr A0A397Y433 A0A397Y433 BRACM		ALDTLGNTKWRVNRKILDVVERLWADGGNIAGLVNREDVP	498
sp Q8L6J1 RPO3B TOBAC		ALDTLGSTKWRVNKRILNVVESIWAGGGNIAGLVDRKDVP	496
tr A0A1U8H8N9 A0A1U8H8N9 CAPAN		ALDTLGSTKWRVNKRILSVVESIWSGGGNIAGLVDRNDVP	498
tr A0A803QAW0 A0A803QAW0 CANSA		ALDMLGSVKWRVNKKVFSVVESIWDGGGNVAGLVNRKDVT	498
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI		ALDVLGNTKWRVNRRVLNIVECIWARGGNIAGLVNREDVP	499
tr A0A6J111C4 A0A6J111C4_CUCMA		ALDMLGSTKWRVNRRVLSVVESIWEQGGNTAGLVDRKDVP	498
tr A0A6J1E061 A0A6J1E061_MOMCH		ALDMLGSTKWRVNRRVLSVVESIWAQGGNTAGLVDRKDVP	491
tr A0A0A0LH38 A0A0A0LH38_CUCSA		ALDMLGSTKWRVNRSVLSVVESLWAQGGNTAGLVDRKDVP	495
tr A0A1S3BU26 A0A1S3BU26 CUCME		ALDMLGSTKWRVNRRVLSVVESLWAQGGNTAGLVDRKDVP	495
tr A0A835UM73 A0A835UM73 VANPL		ALDTLGATKWRINKRVLGIVDRLWSSGGRLAGLVDREDLP	544
tr A0A4S8JC56 A0A4S8JC56_MUSBA	YICLVIFFDHTYIDKVSANC	ALDTLGNTKWRVNKRVLSIVDRIWSSGGRLADLVDHEDIP	575
tr A0A445DB12 A0A445DB12_ARAHY		ALDTLGNTKWRVNKRVLSVIDRIWASGGRLADLVDRNDVP	539
tr A0A0E0QNF8 A0A0E0QNF8_ORYRU		ALNTLGSTKWRVNKRVLSIVDRIWSSGGRLADLVDRTDVA	502
tr A0A6P6BHY6 A0A6P6BHY6_DURZI		ALDTLGNTKWRINKRVLGVVDRIWANGGRIADLVDREDVP	515
tr A0A5E4F8X8 A0A5E4F8X8_PRUDU		ALDTLGSTKWRVNKRVLGVIDRIWASGGRLADLVDREDVP	525
tr M5XKH1 M5XKH1_PRUPE		ALDTLGSTKWRVNKRVLGVIDRIWASGGRLADLVDREDVP	525
tr A0A2C9WDN1 A0A2C9WDN1_MANES		ALDTLGRTKWRINKRILSIVDRIWNSGGRLADLVDRSDVP	549
tr A0A6P6TN25 A0A6P6TN25_COFAR		ALDTLGNTKWRVNKRVLSVVDRIWASGGRLGNLVDRNDIP	559
tr A0A1U7ZW01 A0A1U7ZW01_NELNU		ALDTLGNTKWRVNKRVLSVVDRIWASGGRLADLVDREDVP	550
tr D7SML3 D7SML3_VITVI		ALNTLGNTKWRINKRLFDVVDRIWAGGGCLADLVDRNDVP	554
	TRATH ORDORATION		
LEIAUA251TGF3IAUA251TGF3_HELAN	IFDLN-SEDSSEIREWSWSV	NKAKKINGELESCROOLEDELEYARRONDE ESPTTPHNO	363
LE LAGASUSMEES LAGASUSMEES_CREGI	LPERPLIED PARIDIWEWSV	NAKKINDEREAL COMPARING AGENCIE EUPTIPHIC	366
- UATTWI UATTWI PURCH	T DE ND CCED DED TOCHENCU	DESCRIPTION DECT SCHOOL ST STATES DE STATES DE STATES	872
LEIVINIALIVINIA_BUINA	T DE ED COEDDE EL CONTROV	DEASE THE PERSON AND AND THE PERSON PROPERTY OF THE PERSON	570
CE INGINE CONTRACTOR	THE POLOTELE OF THE T		310
triD7LGR71D7LGR7 ARALL	IFERPSSEDPERIOSWEWSV	BRANKINDERS LEG OVERSE VARENS DE OFTERRE	569
tria0a078c6j8 a0a078c6j8 BBANA	IPCKPLSEDPEEICLWKWSV	BRAKKINDERS SLACOVELELSVADENSDERGEYTEHNLO	555
tr AOA6JOE1N7 AOA6JOE1N7 RAPSA	IPERPSSEDPEEIQSWEWSV	REAKETNRERESLECOVELELTVARENEDERSFTTPHELO	361
tr1A0A397Y4331A0A397Y433 BBACM	IPERPSSEDPEELQSWKWSV	REAKKINGERESLECOVEDEL VARENEDERSPYTPHNLD	558
sp g8L6J1 RPO3B_TOBAC	I PELSNSDO IMEVERNENEV	RESERINGELES ORCOTELE VAREIER FYYPHILD	556
tr A0A1USNSN9 A0A1USESN9 _ CAPAN	IPDLN-SDOIMEVERNEWRV	REARKINGELES GROOTELELSVAREDER FYYPHSLO	557
EF AGASO 3GANG AGAGA CANSA	VFERFSSDOLTEINGWEWEV	REAREINSERESLEC <mark>OVELELEVARMEDED</mark> GFTFFHELD	558
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	IPERPSSEDLTEIGENRWML	PRASKINGERESCROTTERLEVACENEDESSTTPHELD	559
trlA0A6J1I1C4lA0A6J1I1C4_CUCMA	IPERPLOGDLTENGEWEWSM	BEAKKINGELESCHOILGELEVARENEDERSTYPHALD	558
tr A0A6J1E061 A0A6J1E061_MOMCH	IPEQPLG-DLTENGEWENSV	REARKINGELESGROUTLELEVARESEDEESFITFHILD	550
tr A0A0A0LH38 A0A0A0LH38_CUCSA	IPERPLG-DLTENGENRWSM	KEAKKINGELUSGROVELELEVAREMEDERSFTEPHNLD	554
tr A0A1S3BU26 A0A1S3BU26_CUCME	IPERQLG-DLTETQENRWSM	REARKINGELESGROPULALIVAREMEDERSPTTPHNUD	554
LE AUASSSUN73 AUASSSUN73 VANPL	LPDKPETEDENLISRNRWOV	RGIRKENIERESORGDVELELEVARSMEDERGFTTPHELD	604
LE IAUA458JC56 IAUA458JC56_MJ5BA	LPERPETDOEAELSENFUEL	S.SVEKENSERESORCOVELEEVARSISEEESFTTPHNVD	635
DETAVA1150B12TAVA1150B12_ABAHY	LPEEPDIDDEATIKKWWWKA	S SYRKEN SENISCING THE REAL AREAS DEED FYTEHNYD	399
EFINONVEDQNFGINONUEDQNFG_CRTRU	SPERFUTEDEDRIGRNRWTL	KAARAENGERESUKA PARAMIARANAN ANTIPINGS	362
LE INVACE CONTE INVACE CONTE	STREPUTEDELETRENKWKV	ANTALASEKSU ASTRONOMIAK MATERIALI	373
EFINANCETONO (NUNSEEFERS_FRUDU	L DE E DOTE DE LA EL DEN KWEL	SAARNENSERS VOLGEN AND ASKS SPECIFIED IN THE SECOND STATES OF THE SECOND	363
EFINANALINANAL PROFE	LPERPOTENCE INSTRUCT	RSVEEPNEEPSOND PROBLEM PROPERTY FILL	606
EVILOASPETNOSIAOSPETNOS COPAR	LEREPTENZA TLEREPT	REFERENCES OF THE REAL PARTY AND THE PARTY A	619
tr 1A0A1U72N01 1A0A1U72N01 SPLNU	I PEEPDTEDEAR I RENEWEV	RAAKENSERESCHOLENELEVARMADENSE PUPUL	610

LPERPOTEDEAQLBENEWEVESVEEANSERESQ

 $R \geq R$

:

_ *

::

:

614

ARES

PPTPHN

tr A0A251TQF3 A0A251TQF3_HELAN	FRGRAYPMEPMEN <mark>HE</mark> SS <mark>O</mark> LCRGILEFAEGRPLGKSGLR#LRIHLANLYGSGVEKLSYDGR	625
CTIAOAGOJNPEJIAOAGOJNPEJ_CHEQI	FRORATERINE SALCROVLEFAEGRPLGRSGLOWLRIELANVIAGGVDRLSIEAR	626
EFIAUA565CQ27IAUA565CQ27_9BRAS	FROMATPREFELMENTS SECRET LE FAEGRPLGKSGLEWIK INLANDYAGGVEKLSHOGR	638
Er V4L7M1 V4L7M1_EUTSA	FRGRAYPHEPHLNHLS SOLCRGT LE FAEGRPLGKSGLYWLK IHLANLYAGGVEKLSHDGR	632
trIROHH96IROHH96_9BRAS	FRORAYPHEPELN HIS SOLCROTLE PAEGRPLGKSGLEWLKIELANLYAGGVERLSEDAR	630
MA10336101 MD023 _ #H#23H	PROVED AND ADDRESS OF A DATA DATA DATA DATA DATA DATA DATA D	
triD7LGR7 D7LGR7_ARALL	FRGRAYPMHPHLNHLSSDLCRGTLEFAEGRPLGESGLHWLKIHLANLFAGGVEKLSHDAR	62.9
tr1a0a078g6j81a0a078g6j8_brana	FRGRAYPMNPHLNHLSS CRGTLEFAEGRPLGRSGLYWLKINLANLYAGGVEKLSHDGR	615
tr A0A6J0K1N7 A0A6J0K1N7_RAPSA	FRORAYDMHDHLMHLSSOLCRGTLEFAEGRPLGKSGLYWLKIHLANLYAGGVEKLSHDGR	621
tr A0A397Y433 A0A397Y433_BRACM	FRGRAYPMHPHLNHLSSOLCRGTLEFAEGRPLGKSGLYWLKIHLANLYAGGVEKLSHEGR	618
splQ8L6J1 RPO3B_TOBAC	FRGRATPMHPHLNHLSSOLCRGILEFAEGRPLGRSGLRWLKIHLASLYAGGVERLCYDAR	616
tr A0AlU8H8N9 A0AlU8H8N9 CAPAN	FROMAYPHHEBLNHLSS CRGILEFSEGRPLGESGLEWLKIHLANLYAGGIEERCYDAR	617
tria0a803gam0la0a803gam0 cansa	FRGRAYPHERELNESS CRGVLEFAEGRPLGESGLEWLEIGLANLYAGGVEELSYDGE	618
trla0a5n6rvv2la0a5n6rvv2 9rosi	FRGRAYPHNPHLNHLSS CRGILEFAEGRPLGKSGLHWLKIHLANLYSGGVEKLSHDGR	619
tr A0A6J1I1C4 A0A6J1I1C4 CUCMA	FRORAYPHEPHINELSS CROVLE FAEGRPLGESGLEWLEIHLANLYAGGVEELSYDRR	618
tr A0A6J1E061 A0A6J1E061 MONCH	FRGRAYPHHERLANDS STLCRGVLE FAEGRPLGKSGLHWLKIHLANLYAGGVERLSYDRR	610
tr IA0A0A0LH38 IA0A0A0LH38 CUCSA	FRGRAYPHHPELNHSSDLCRGVLEFAEGRPLGKSGLHWLKIHLANLYAGGVEKLSYDER	614
trlA0A1S3BU26lA0A1S3BU26 CUCME	FRORAYPHHEELNHESS C.CRGVLEFAEGRPLGRSGLHWLKIHLANLYAGGVEKLSYDRR	614
tr A0A835UM73 A0A835UM73 VANPL	FRGRAYPHEPELNILGSOLCRGVLEFAEGROLGESGLEWLEIHLANLYGNGVDEISYDGE	664
tr A0A458JC56 A0A458JC56 MUSBA	FRORATPHMPTLNMLGSDLCRGVLEFAEGRPLGRSGLRWLKINLANLYAGGVDKLSNDAR	695
trIA0A445DB12IA0A445DB12 ARAHY	FRORAYPHEPELNELOSOLCRGILE FAEGRPLGKSGLHWLKIHLANLYAGGVDKLSHEGR	65.9
triadadeoonfsiadadeoonfs oryru	FRGRAYPMHPYLNHLGSDLCRGVLEFAEGRPLGKSGLBWLKIHLANLYAGGVDKLSYDGR	622
tr A0A6P6BHY6 A0A6P6BHY6 DURZI	FRORAYPHEPTLNHLOSDLCRGVLEFAEGRPLGRSGLEWLKIHLANLYAGGVDKLPYEGR	635
trla0a5e4F8X8la0a5e4F8X8 PRUDU	FRORAYPHEPYLNHLOSOFCRGILEFSEGRHLGESGLEWLEIHLANLYAGGVDELSFDDE	645
tr M5XEH1 M5XEH1 PRUPE	FRORAYPHEPTLNHLGS CCRGILEFSEGRHLGRSGLEWLKIHLANLYAGGVDKLSFDDR	645
trlA0A2C9WDN11A0A2C9WDN1 MANES	FRORAYPMEPYLNHLOSOLCRGILEFAEGRPLGKSGLOWLKIHLANLPAGGVDKLSYDGR	669
tr IA0A6P6TN25 IA0A6P6TN25 COPAR	FRORAVPMEREINHOSDVCRGMLEFAEGRPLGKSGLEWLKIHLANLFANGVDKLSHEAR	679
tr[A0A1U72W01]A0A1U72W01 NELNU	PROBATEMENT AND A CREATE FACTOR AND A CREATE	670
triD7SML3 D7SML3 VITVI	FRORATPHHEPTINE COLCEGILE FAEGRPLGSSGLHWLKINLANL FAGGVDELSHEGR	67.4
		27.8

tr A0A251TQF3 A0A251TQF3 HELAN	SYPNALLARLLIGQVN <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTFVGAR	778
tr A0A803MFE3 A0A803MFE3 CHEQI	RklvkQTVMTSVYGVTFVGAR	779
tr A0A565CQ27 A0A565CQ27 9BRAS	RANDER SNPTAQLAKIVINQVERKLVKQTVMTSVYGAR	791
tr V4L7M1 V4L7M1 EUTSA	Reverse SNPTAALAKILINQVERKLVKQTVMTSVYGAR	785
tr R0HH96 R0HH96 9BRAS	RADING SNPTAALAKILITQVERKLVKQTVMTSVYGAR	783
sp 024600 RPOT3 ARATH	SNPTAALAKILITOVDRKLVKQTVMTSVYGVTYVGAR	785
tr D7LGR7 D7LGR7_ARALL	SNPTAALAKILITQVERKLVKQTVMTSVYGAR	782
tr A0A078G6J8 A0A078G6J8_BRANA	RADING SNPTAALAKILINQVERKLVKOTVMTSVYGAR	768
tr A0A6J0K1N7 A0A6J0K1N7_RAPSA	RADING SNPTAALAKILINQVIRKLVKQTVMTSVYGAR	774
tr A0A397Y433 A0A397Y433_BRACM	RADING SNPTAALAKILINQVERKLVKQTVMTSVYGAR	771
sp Q8L6J1 RPO3B_TOBAC	IDPNALLAKLLIDQVIRKLVKQTVMTSVYGAR	769
tr A0A1U8H8N9 A0A1U8H8N9_CAPAN	RDPNALLAKLLIDQVIRKLVKQTVMTSVYGAR	770
tr A0A803QAW0 A0A803QAW0_CANSA	RHPCASLAKILIDQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAR	771
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	TNPHAVLAKILIDQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAR	772
tr A0A6J1I1C4 A0A6J1I1C4_CUCMA	RNPNALLAKILIDQVIRKLVKQTVMTSVYGAR	771
tr A0A6J1E061 A0A6J1E061_MOMCH	TNPNALLAKILIDQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAR	763
tr A0A0A0LH38 A0A0A0LH38_CUCSA	TNPNASLAKLLIDQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAR	767
tr A0A1S3BU26 A0A1S3BU26_CUCME	TNPNASLAKLLIDQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAR	767
tr A0A835UM73 A0A835UM73_VANPL	YDGSEQRIFGLKEILLAEIKKLNKVEQRKSVKMFTLKVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAR	843
tr A0A4S8JC56 A0A4S8JC56_MUSBA	VNRDALRARLLVDQVI <mark>R</mark> KLVKQTVMTSVYGAR	848
tr A0A445DB12 A0A445DB12_ARAHY	IFPNALHARRLVNQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAR	812
tr A0A0E0QNF8 A0A0E0QNF8_ORYRU	TDPDAARARLLLDQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAR	775
tr A0A6P6BHY6 A0A6P6BHY6_DURZI	RNPNALHARLLINQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAR	788
tr A0A5E4F8X8 A0A5E4F8X8_PRUDU	TNPNALHARLLINQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYVGAR	798
tr M5XKH1 M5XKH1_PRUPE	TNPNALHARLLINQVI <mark>R</mark> KLVKQTVMTSV <mark>YG</mark> VTYVGAR	798
tr A0A2C9WDN1 A0A2C9WDN1_MANES	VFPDALHARTLINQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAR	822
tr A0A6P6TN25 A0A6P6TN25_COFAR	WFPDALRARVLVNQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAR	832
tr A0A1U7ZW01 A0A1U7ZW01_NELNU	THPNAMHARILINQVI<mark>R</mark>KLV<mark>KO</mark>TVMTSV<mark>YG</mark>VTYIGAR	823
tr D7SML3 D7SML3_VITVI	IFPDALRARILINQVI <mark>R</mark> KLV <mark>KQ</mark> TVMTSV <mark>YG</mark> VTYIGAR	827
	_ = <u>* * * * * * * * * * * * * * * * * </u>	

				_	. Pol – A	CTD	
tr A0A251TQF3 A0A251TQF3 HELAN	FVHSLDGTHMMMTAVACRDAGLRFAGV	HDS	FWTHAC	DVD	KMNVILREKFVELYRTPILE	N	950
tr A0A803MFE3 A0A803MFE3 CHEQI	YVHSLDGTHMMMTAVACRDAGLQFAGV	HDS	FWTHAC	DVE.	KMNQILREKFVELYSMPILE	N	951
tr A0A565CQ27 A0A565CQ27 9BRAS	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	TMNRILREKFVELYNIPILE	D	963
tr V4L7M1 V4L7M1 EUTSA	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	TMNRILREKFVELYNTPILE	D	957
tr R0HH96 R0HH96 9BRAS	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	TMNRILREKFVELYNTPILE	D	955
sp 024600 RPOT3_ARATH	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	TMNRILREKFVELYNTPILE	D	957
tr D7LGR7 D7LGR7_ARALL	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	FMNRILREKFVELYNTPILE	D	954
tr A0A078G6J8 A0A078G6J8_BRANA	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	[MNRILREKFVELYSSPILE	D	940
tr A0A6J0K1N7 A0A6J0K1N7_RAPSA	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	FMNRILREKFVELYSTPILE	D	946
tr A0A397Y433 A0A397Y433_BRACM	FVHSLDGTHMMMTAVACREAGLNFAGV	HDS	YWTHAC	DVD	[MNRILREKFVELYSTPILE	D	943
sp Q8L6J1 RP03B_TOBAC	FVHSLDGSHMMMTAVACRDAGLQFAGV	HDS	FWTHAC	DVD	OMNRILREKFVELYSLPILE	D	941
tr A0A1U8H8N9 A0A1U8H8N9_CAPAN	FVHSLDGSHMMMTAVACRDAGLHFAGV	HDS	FWTHAC	DVD	OMNRILREKFVELYSMPILE	D	942
tr A0A803QAW0 A0A803QAW0_CANSA	FVHSLDGSHMMMTALACRDAGLCFAGV	HDS	FWTHAC	DVD	QMNQILRQQFVELYSMPVLE	N	943
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	FVHSLDGTHMMMTAIACRDAGLRFAGV	HDS	FWTHAC	DVE.	KMNQILRDKFVELYSKPILE	N	944
tr A0A6J111C4 A0A6J111C4_CUCMA	FVHSLDGSHMMLTALACRDAGLRFAGV	HDS	FWTHAC	DVD	OMNLILREKFVELYNMPVLE	т	943
tr A0A6J1E061 A0A6J1E061_MOMCH	FVHSLDGSHMMLTALACRDAGLRFAGV	HDS	FWTHAC	DVD	OMNVILRETFVELYNMPVLE	s	935
tr A0A0A0LH38 A0A0A0LH38_CUCSA	FVHSLDGSHMMLTALACRDAGLRFAGV	HDS	FWTHAC	DVD.	RMNLILREKFVELYSMPVLE	s	939
tr A0A1S3BU26 A0A1S3BU26_CUCME	FVHSLDGSHMMLTALACRDAGLRFAGV	HDS	FWTHAC	DVD	RMNLILREKFVELYSMPVLE	s	939
tr A0A835UM73 A0A835UM73_VANPL	FIHSLDGSHMMMTAVACRDSGLNFAGV	HDS	YWTHAC	DVD	KINIILREKFVELYKQPILE	D	1015
tr A0A4S8JC56 A0A4S8JC56_MUSBA	FVHSLDGSHMMMTAVACKRAGLNFAGV	HDS	YWTHAC	DVD	OMNRILREKFVELYERPILE	E	1028
tr A0A445DB12 A0A445DB12_ARAHY	FVHSLDGTHMMMTAVACKKEGLNFAGV	HDS	YWTHAC	DVD.	RMNRILREKFVELYETPVLE	N	984
tr A0A0E0QNF8 A0A0E0QNF8_ORYRU	FVHSLDGSHMMMTAVACKRQGLNFAGV	HDS	YWTHAC	DVD	TMNKILREKFVELYDTPILE	N	947
tr A0A6P6BHY6 A0A6P6BHY6_DURZI	FVHSLDGSHMMMTAVACKKAGLNFAGV	HDS	YWTHAC	DVD	EMNRILREKFVELYEAPILE	N	960
tr A0A5E4F8X8 A0A5E4F8X8_PRUDU	FVHSLDGSHMMMTAVACKKAGLNFAGV	HDS	YWTHAC	DID	EMNRILREKFVELYEAPILE	N	970
tr M5XKH1 M5XKH1_PRUPE	FVHSLDGSHMMMTAVACKKAGLNFAGV	HDS	YWTHAC	DID	EMNRILREKFVELYEAPILE	N	970
tr A0A2C9WDN1 A0A2C9WDN1_MANES	FVHSLDGSHMMMTAVACKNAGLNFAGV	HDS	YWTHAC	DVD.	EMNRILREKFVELYEKPILE	S	994
tr A0A6P6TN25 A0A6P6TN25_COFAR	FVHSLDGSHMMMTAVACRKAGLNFAGV	HDS	YWTHAC	DVD.	EMNRILREKFVELYETPILE	N	1004
tr A0A1U7ZW01 A0A1U7ZW01_NELNU	FVHSLDGSHMMMTAVACRKAGLNFAGV	HDS	YWTHAC	DID	EMNKILREKFVELYEKPILE	N	995
tr D7SML3 D7SML3_VITVI	FVHSLDGSHMMMTAIACKKAGLNFAGV	HDS	YWTHAC	DVD	EMNRLLREKFVQLYETPILE	N	999
	· · * * * * * * * * * * * * * * * * * *	* * *	*****	*::	:* :**: **:** *:**	1	

//End of NEP sequences			
tr A0A251TQF3 A0A251TQF3_HELAN	LLESFQTTYPGLEFPPLPPRGDFDLNEVLKSF	Y F FN	986
tr A0A803MFE3 A0A803MFE3_CHEQI	LLQSFQESYPESTFPPLPERGNFELEKVLESP	Y F FN	987
tr A0A565CQ27 A0A565CQ27_9BRAS	LLQSFQESYPNLEFPPVPQRGDFDLKEVLKSQ	Y F FN	999
tr V4L7M1 V4L7M1 EUTSA	LLQSFQESYPNLVFPPVPKRGDFDLKEVLKSQ	Y F FN	993
tr R0HH96 R0HH96 9BRAS	LLQSFQESYPNLVFPPVPKRGDFDLKEVLKSQ	Y F FN	991
sp 024600 RPOT3_ARATH	LLOSFQESYPNLVFPPVPKRGDFDLKEVLKSQ	YFFN	993
tr D7LGR7 D7LGR7_ARALL	LLQSFQESYPNLVFPPVPKRGDFDLKEVLKSQ	Y F FN	990
tr A0A078G6J8 A0A078G6J8 BRANA	LLQSFQESYPTLVFPPVPKRGDFDLKEVLKSQ	Y F FN	976
tr A0A6J0K1N7 A0A6J0K1N7 RAPSA	LLQSFQESYPNLVFPPVPKRGDFDLKEVLKSH	Y F FN	982
tr A0A397Y433 A0A397Y433_BRACM	LLQSFQESYPNLVFPPVPKRGGFDLKEVLKSQ	Y F FN	979
sp Q8L6J1 RPO3B TOBAC	LLENFQKSYPALTFPPLPKRGDFNLREVLESP	YFFN	977
tr A0Alu8H8N9 A0Alu8H8N9 CAPAN	LLQSFQESYPALTFPPLPKRGDFDLREVLESP	Y F FN	978
tr A0A803QAW0 A0A803QAW0_CANSA	LLENFQTSYPELEFPPIPERGSFNLQEVLNSF	Y F FN	979
tr A0A5N6RVV2 A0A5N6RVV2_9ROSI	LLVSFQTSHPVLVFPPLPERGDFDLRKVLESF	Y F FN	980
tr A0A6J1I1C4 A0A6J1I1C4_CUCMA	LLEDFETTYPGLTFPPLPERGNFDLREVLKSR	Y F FN	979
tr A0A6J1E061 A0A6J1E061 MOMCH	LLEGFETTYPGLTFPSLPDRGDFDLREVLNSF	Y F FN	971
tr A0A0A0LH38 A0A0A0LH38_CUCSA	LLEGFETTYPGLTFPSLPGRGDFDLQEVLKSP	Y F FN	975
tr A0A1S3BU26 A0A1S3BU26_CUCME	LLEGFETTYPGLTFPSLPERGDFNLQEVLRSF	Y F FN	975
tr A0A835UM73 A0A835UM73_VANPL	LLESFRKSFPTLSFPELPERGDFDLKEVLDSF	Y F FN	1051
tr A0A4S8JC56 A0A4S8JC56_MUSBA	LLESFQQSFPTLNFPPLPERGDFDLKDVLESF	Y F FN	1064
tr A0A445DB12 A0A445DB12_ARAHY	LLESFQGSFSTLMFPPLPDR <mark>ETST</mark>		1008
tr A0A0E0QNF8 A0A0E0QNF8_ORYRU	LLESFEKSFPELKFPPLPERGDFDLTDVLGSP	Y F FN	983
tr A0A6P6BHY6 A0A6P6BHY6_DURZI	LLESFQKSFPSLNFPPLPERGDFDLREVLESF	Y F FN	996
tr A0A5E4F8X8 A0A5E4F8X8_PRUDU	LLEGFQQSFPTLTFPPLPDRGDFDLRDVLESF	Y F FN	1006
tr M5XKH1 M5XKH1_PRUPE	LLEGFQQSFPTLTFPPLPDRGDFDLRDVLESP	Y F FN	1006
tr A0A2C9WDN1 A0A2C9WDN1_MANES	LLESFEQSFPTLSFPSLPERGDFDLREVLESP	Y F FN	1030
tr A0A6P6TN25 A0A6P6TN25_COFAR	LCRSGVTSILKMFWIPRISLTDTI	Y <mark>PG</mark> N	1032
tr A0A1U7ZW01 A0A1U7ZW01_NELNU	LLESFQQSFPTLNFPPLPERGDFDLREVLDSF	Y F FN	1031
tr D7SML3 D7SML3_VITVI	LLESFQQSFPALEFPPLPERGDFDLREVLESF	Y F FN	1035
	* . : :		

A0A251TQF3_HELAN, Helianthus annuus A0A565CQ27_9BRAS Arabis nemorensis R0HH96_9BRAS, Capsella rubella D7LGR7_ARALL, Arabidopsis lyrata subsp. lyrata A0A6J0K1N7_RAPSA, Raphanus sativus Q8L6J1|RP03B_TOBAC, Nicotiana tabacum A0A803QAW0_CANSA, Cannabis sativa A0A6J111C4_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 -620	WLA <mark>Y⁻⁸G</mark> VT <mark>R-</mark> 4SVT <mark>K</mark> R1SVMTLA <mark>Y8G</mark> S-
Viral SP6 SSU RNA Pol -61	² WDS <mark>I[®]G</mark> IT <mark>R-</mark> 4SLT <mark>K</mark> K ¹ PVMTLP <mark>Y⁸G</mark> S-
Mitochondrial SSU RNA pol (<i>Sc</i>)	- ¹⁰⁰⁹ TR ⁻⁴ KVVKQ ¹ TVMTNVY ⁸ GV-
Mitochondrial SSU RNA pol (H. sapiens	s) - ⁹⁸⁶ TR⁻⁴KVVKQ¹TVMTVVY ⁸ GV-
<i>E. coli</i> DNA pol I (SSU)	- ⁷⁵³ Q <mark>R</mark> ⁴ RSA <mark>K⁰⁰³A</mark> 1INFGLI <mark>Y</mark> ⁸ GM-
Chloroplast SSU RNA pol (NEP) (ARA7	TH) - ⁷⁶⁵ DR ⁻⁴ KLVK ⁷⁷⁰ Q ¹ TVMTSVY ⁸ GV-
Chloroplast SSU DNA pol IA (ARATH)	- ⁸⁷³ E R - ⁴ RKA <mark>K⁸⁷⁸M</mark> ¹ LNFSIA Y⁸G K-
Chloroplast SSU DNA pol IB (ARATH)	- ⁸⁵⁷ E R - ⁴ RKAK ⁸⁶² M ¹ LNFSIAY ⁸ GK-

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 $3' \rightarrow 5'$ PR exonuclease domain and amino acids from 368 to 783 were involved in polymerase function. They found that the $D^{335} \rightarrow N$ 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 ($D^{156} \rightarrow N$, $D^{229} \rightarrow N$ and $D^{335} \rightarrow N$) 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 $D^{290} \rightarrow A/E^{292} \rightarrow A$ (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 (Asp¹², Glu¹⁴, Asp¹⁰³, and Asp¹⁶⁷) with the invariant His¹⁶², which acts as the general base in catalysis. They also found that modification of the two conserved amino acid residues, viz. Asp¹² \rightarrow Ala and Glu¹⁴ \rightarrow 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 ⁻545LRCD⁵⁴⁸VE⁵⁵⁰LKL-----NLD⁵⁷²F------LNH⁵⁸⁶→3aa→D⁵⁹⁰LC₅₉₂⁻



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 $3' \rightarrow 5'$ exonuclease \leftrightarrow 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 (<i>E. coli</i> Phage)	-0 ¹¹² x = ¹¹⁴ -0 ²¹⁹ -Y ³²⁰ -0 ³²⁴ -	Tyr	Zn ²⁺	1
DNA pol I <i>(E. coli</i>)	-0 ³⁵⁵ x= ³⁵⁷ -0 ⁴²⁴ -Y ⁴⁹⁷ -0 ⁵⁰¹ -	Tyr	Zn ²⁺	1
DNA pol II <i>(E. coli</i>)	- <mark>0¹⁵⁶xE¹⁵⁸-0²²⁹-Y³³¹-0³³⁵-</mark>	Tyr	Zn ²⁺	1
RNase D <i>(E. coli</i>)	-D ²⁸ xE ³⁰ D ⁸⁵ Y ¹⁵¹ -D ¹⁵⁵ -	Tyr	Zn ²⁺	1
DNA pol III, ε-subunit <i>(E. coli</i>)	-D ¹² x= ¹⁴ -D ¹⁰³ -H ¹⁶² -D ¹⁶⁷ -	His	Zn ²⁺	1
RNase T <i>(E. coli</i>)	- 23 x - 25- 125- 181- 186-	His	Zn ²⁺	1
PR Enzymes in Eukaryotic D	NA Replicases			
DNA pol ε cat. subunit (S.c)	- <mark>D²⁹⁰X=²⁹²-D³⁸³-Y⁴⁷³-D⁴⁷⁷-</mark>	- His	Zn ²⁺	1
DNA pol δ cat. subunit (<i>H.s</i>)	-D ³¹⁶ XE ³¹⁸ -D ⁴⁰² -Y ⁵¹¹ -D ⁵¹⁵	- His	Zn ²⁺	1
PR Enzymes in DNA Polyme	rases 1A and 1B from Plan	t Chloropla	<mark>sts</mark>	
DNA polymerase 1A (ARATH)	-DxE ²⁹⁶ D ³⁶⁸ Y ⁴⁷⁰ -	–D ⁴⁷⁴ - Tyr	Zn ²⁺	1
DNA polymerase 1B (<i>ARATH</i>)	-DxE ²⁷² D ³⁴⁶ Y ⁴⁴⁸ -	–D ⁴⁵² - Tyr	Zn ²⁺	1
PR Enzymes in RNA Polyme	rase, NEP from Plant Chlor	oplasts		
Arabidopsis Thaliana	-DxE ⁵⁵⁰ D ⁵⁷² H ⁵⁸⁶ -	–D ⁵⁹⁰ - His	Zn ²⁺	1
Arachis hypogaea	-DxEDH	D- His	Zn ²⁺	1
Oryza rufipogon	-DxEDH	D- His	Zn ²⁺	1
Nelumbo nucifera	-DxEDH	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

DEDD-superfamily of PR examinationses (-DyE-E-H+W*D*-)				
DNA Polymerases				
T4 DNA pol (E. col/ Pha	ge) -∎I∎ ¹¹⁴ F∎ ²¹⁹ SY ³²⁰ N→3 aa∎ ⁵²⁴ VE- [40]			
T7 DNA pol (E. col/ Pha	Je) -DIE ⁷ FD ²³⁵ DY ¹⁷⁰ N→3 aaD ¹⁷⁴ VV-			
Prokaryotic DNA Repli	cases (DNA pol III- c-subunits)			
E. coli	- 2 ¹² T 2 F 2 ¹⁰² L 2 ¹⁰² G→4 aa- 2 ⁰⁷ AQ-			
Citrobacter amalonaticu	s -D ¹⁵ TEFD ¹⁰⁵ LH ¹⁶⁵ G→4 aaD ¹⁷⁰ AQ			
Shigella dysenteriae	-D12TEFD192LH1952G4 aaD957AQ			
Salmonella typhimurium	-D ¹² TEFD ¹⁰² LH ¹⁶² G→4 aa←D ¹⁶⁷ AQ			
Eukaryotic DNA Replic	ases			
DNA policicat, subunit (;	Sc) = ED ³⁸³ = EY ⁴⁷³ S = 3 aa = D ⁴⁷⁷ AV - [35,36]			
DNA pol 5 cat subunit (He)			
Drivi por o car. Suburne (
DNA Polymerases IA fi	om Plant Chioroplasts"			
Arabidopsis Thallana	-D ^{2ad} TEFD ^{ad} SSY ^{ad} S→3 aa→D ^{ad} Al-			
Chlorella desiccate	-D2# TEFD##RSY##S -3 aa -D##AK-			
Nelumbo nucifera	-Der TEFDer SFY S3 aa -Der SI-			
Raphanus sativus	-D*** TEFD***SSY***S3 aaD***AI			
DNA Polymerases IB fr	om Plant Chloroplasts [*]			
Arabidopsis Thaliana	-D ²⁷² TEFD ¹⁴⁵ NSY ⁴⁴⁹ S→3 aa→D ⁴⁵² SI-			
Nicotiana tabacum	-0 ²⁰⁰ TEFD ⁴⁴⁴ NCY ⁵¹² S3 aaD ⁵⁴⁸ SF			
Sesamum indicum	-0140 TEFD441NSY502S3 aaD506 SI-			
Raphanus sativus	-0 ²⁵⁶ TEFD ¹⁰⁰ NSY ⁴³⁶ S3 aaD ⁴³⁶ Sh			
E. coli DNA pol I Exo*	-D ³⁶⁶ TEYD ¹³¹ RY ¹⁹⁷ A→3 aa→D ⁵⁹¹ AD-			
SARS-CoVs*				
SARS-CoV-1 ExoN/ACE	2 - ³⁰ ∎V∎			
MERS-CoV ExoN/DPP	4 - ⁵⁰ ∎V∎			
SARS-CoV-2 ExoN//AC	E2 - ⁹⁰ VH			
Nuclear Encoded DNA	Polymerana (NEO) from Diant ChloroplantsA			
HOGINAL-ENCODED RNA	Formerase mer internation riant Childrophasts.			
Arabidopsis Thallana	-D***VELD***FNH***L→3 aa→D***LC-			
Arachis hypogaea	-Orrove-VOrror-VOrror-NH***L-3 aa-OrroLC-			
Oryza ruhpogon	-DIVE			
Nelumbo nucifera	•Dam /EFOrmeNHFO3 sa→Dam FC			

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

 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 nucleusencoded 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.