## Compilation and alignment of DNA polymerase sequences

#### Junetsu Ito and Dan K.Braithwaite

Department of Microbiology and Immunology, College of Medicine, University of Arizona Health Sciences Center, Tucson, AZ 85724, USA

Received May 29, 1991; Revised and Accepted July 9, 1991

#### INTRODUCTION

More than 40 different DNA polymerases, including some putative DNA polymerase sequences deduced from nucleotide sequence data, have recently been reported (1-39). The amino acid sequences of these DNA polymerases have been aligned and partial homologous regions identified by many investigators (2-4,9,10,12-25,27-36,42-51). Based on the segmental amino acid sequence similarities, DNA polymerases have been classified into two major groups; E. coli DNA polymerase I-Type and eukaryotic DNA polymerase  $\alpha$ -Type (14,44,47,48,51), or family A DNA polymerases and family B DNA polymerases (4,9,50). As the number of DNA polymerase sequences increases, the classification of DNA polymerases becomes increasingly ambiguous. For example, DNA polymerase delta of yeast was shown to have amino acid sequence similarity to the  $\alpha$ -Type DNA polymerases (17). It has become necessary to establish a unified classification of DNA polymerases. Here we propose to classify DNA polymerases into families A, B, and C (Figure 1: A, B, and C), according to the amino acid sequence homologies with E. coli DNA polymerases I, II, and III, respectively. As new and different prokaryotic and eukaryotic DNA polymerases are identified, the number of families can easily be expanded by using additional letters of the alphabet (i.e., D, E, etc.).

The bacterium  $E.\ coli$  (strain K12) contains three distinct DNA polymerases I, II, and III (52).  $E.\ coli$  DNA polymerase I, the first DNA polymerase discovered, is specified by the polA gene (52).  $E.\ coli$  DNA polymerase II, encoded by the polB gene, was recently sequenced and found to be identical to the dinA gene, a DNA damage inducible gene whose expression is regulated by the SOS system in  $E.\ coli$  (8,53). Amino acid sequence alignment shows that  $E.\ coli$  DNA polymerase II has significant homology with family B ( $\alpha$ -Type) DNA polymerases (8,53,54).

E. coli DNA polymerase III is a multisubunit enzyme encoded by various dna genes (55); the DNA polymerizing  $\alpha$ -subunit encoded by the polC (dnaE) gene (56) and the  $3' \rightarrow 5'$  exonuclease performing  $\epsilon$ -subunit encoded by the dnaQ gene (57). The  $\alpha$ -subunit of E. coli DNA polymerase III exhibits an extensive homology with the corresponding  $\alpha$ -subunit of Salmonella typhimurium DNA polymerase III (35); and both show significant homology to Bacillus subtilis DNA polymerase III, a single-polypeptide encoded by the polC gene (36).

In summary, family A DNA polymerases are named for their homology to the product of the *polA* gene encoding *E. coli* DNA polymerase I; family B DNA polymerases are named for their

homology to the product of the *polB* gene encoding *E. coli* DNA polymerase II; and family C DNA polymerases are named for their homology to the product of the *polC* gene encoding *E. coli* DNA polymerase III.

The eukaryotic DNA polymerase  $\beta$ , the smallest known DNA polymerase, does not have homology with those of any of the DNA polymerase families described above. Instead, DNA polymerase  $\beta$  has homology with terminal transferases (37). This  $\beta$  group we will call family X (Figure 1D). The classification and original reference(s) for the amino acid sequences of each DNA polymerase are shown in Table 1.

All of the family A DNA polymerases, except for yeast mitochondrial DNA polymerase I, are prokaryotic and are very sensitive to dideoxynucleotide inhibitors, and therefore are useful enzymes for DNA sequencing by the chain-termination method (58). The family A DNA polymerases are resistant to aphidicolin. The family B DNA polymerases are quite extensive in number and variety. Most of the family B DNA polymerases, if not all, are sensitive to aphidicolin and relatively resistant to dideoxynucleotide inhibitors. Most of the family B DNA polymerases, except for pAI2 (33) and yeast DNA polymerase II (16), contain the highly conserved amino acid sequence motif YGDTD, which has been suggested to form part of the dNTP binding site. Amino acid substitutions in this conserved sequence resulted in defects in the DNA polymerase activity without affecting the  $3' \rightarrow 5'$  exonuclease activity (59,60,61). The family C DNA polymerases are major bacterial replicative DNA polymerases which do not have appreciable homology with those of family A and B DNA polymerases. B. subtilis DNA polymerase III is a single polypeptide that is highly sensitive to hydroxyphenylazouracil (62). It is anticipated that the number of sequenced family C DNA polymerases will increase rapidly, since all of the aerobic bacteria may contain a member of this family of DNA polymerases.

#### **SEQUENCE ALIGNMENT**

The 37 complete DNA polymerase sequences and 3 complete terminal deoxynucleotidyltransferase (TDT) sequences are listed in 4 groups; the far Columbia Ex. 2089 polymerases, the far polymerases (inclustrated for the alignment, are very similar to polymerases not City of New York



## Compilation and alignment of DNA polymerase sequences

#### Junetsu Ito and Dan K.Braithwaite

Department of Microbiology and Immunology, College of Medicine, University of Arizona Health Sciences Center, Tucson, AZ 85724, USA

Received May 29, 1991; Revised and Accepted July 9, 1991

#### INTRODUCTION

More than 40 different DNA polymerases, including some putative DNA polymerase sequences deduced from nucleotide sequence data, have recently been reported (1-39). The amino acid sequences of these DNA polymerases have been aligned and partial homologous regions identified by many investigators (2-4,9,10,12-25,27-36,42-51). Based on the segmental amino acid sequence similarities, DNA polymerases have been classified into two major groups; E. coli DNA polymerase I-Type and eukaryotic DNA polymerase  $\alpha$ -Type (14,44,47,48,51), or family A DNA polymerases and family B DNA polymerases (4,9,50). As the number of DNA polymerase sequences increases, the classification of DNA polymerases becomes increasingly ambiguous. For example, DNA polymerase delta of yeast was shown to have amino acid sequence similarity to the  $\alpha$ -Type DNA polymerases (17). It has become necessary to establish a unified classification of DNA polymerases. Here we propose to classify DNA polymerases into families A, B, and C (Figure 1: A, B, and C), according to the amino acid sequence homologies with E. coli DNA polymerases I, II, and III, respectively. As new and different prokaryotic and eukaryotic DNA polymerases are identified, the number of families can easily be expanded by using additional letters of the alphabet (i.e., D, E, etc.).

The bacterium *E. coli* (strain K12) contains three distinct DNA polymerases I, II, and III (52). *E. coli* DNA polymerase I, the first DNA polymerase discovered, is specified by the *pol*A gene (52). *E. coli* DNA polymerase II, encoded by the *pol*B gene, was recently sequenced and found to be identical to the *din*A gene, a DNA damage inducible gene whose expression is regulated by the SOS system in *E. coli* (8,53). Amino acid sequence alignment shows that *E. coli* DNA polymerase II has significant homology with family B ( $\alpha$ -Type) DNA polymerases (8,53,54).

E. coli DNA polymerase III is a multisubunit enzyme encoded by various dna genes (55); the DNA polymerizing  $\alpha$ -subunit encoded by the polC (dnaE) gene (56) and the  $3' \rightarrow 5'$  exonuclease performing  $\epsilon$ -subunit encoded by the dnaQ gene (57). The  $\alpha$ -subunit of E. coli DNA polymerase III exhibits an extensive homology with the corresponding  $\alpha$ -subunit of Salmonella typhimurium DNA polymerase III (35); and both show significant homology to Bacillus subtilis DNA polymerase III, a single-polypeptide encoded by the polC gene (36).

In summary, family A DNA polymerases are named for their homology to the product of the *polA* gene encoding *E. coli* DNA polymerase I; family B DNA polymerases are named for their

homology to the product of the *polB* gene encoding *E. coli* DNA polymerase II; and family C DNA polymerases are named for their homology to the product of the *polC* gene encoding *E. coli* DNA polymerase III.

The eukaryotic DNA polymerase  $\beta$ , the smallest known DNA polymerase, does not have homology with those of any of the DNA polymerase families described above. Instead, DNA polymerase  $\beta$  has homology with terminal transferases (37). This  $\beta$  group we will call family X (Figure 1D). The classification and original reference(s) for the amino acid sequences of each DNA polymerase are shown in Table 1.

All of the family A DNA polymerases, except for yeast mitochondrial DNA polymerase I, are prokaryotic and are very sensitive to dideoxynucleotide inhibitors, and therefore are useful enzymes for DNA sequencing by the chain-termination method (58). The family A DNA polymerases are resistant to aphidicolin. The family B DNA polymerases are quite extensive in number and variety. Most of the family B DNA polymerases, if not all, are sensitive to aphidicolin and relatively resistant to dideoxynucleotide inhibitors. Most of the family B DNA polymerases, except for pAI2 (33) and yeast DNA polymerase II (16), contain the highly conserved amino acid sequence motif **YGDTD**, which has been suggested to form part of the dNTP binding site. Amino acid substitutions in this conserved sequence resulted in defects in the DNA polymerase activity without affecting the  $3' \rightarrow 5'$  exonuclease activity (59,60,61). The family C DNA polymerases are major bacterial replicative DNA polymerases which do not have appreciable homology with those of family A and B DNA polymerases. B. subtilis DNA polymerase III is a single polypeptide that is highly sensitive to hydroxyphenylazouracil (62). It is anticipated that the number of sequenced family C DNA polymerases will increase rapidly, since all of the aerobic bacteria may contain a member of this family of DNA polymerases.

#### SEQUENCE ALIGNMENT

The 37 complete DNA polymerase sequences and 3 complete terminal deoxynucleotidyltransferase (TDT) sequences are listed in 4 groups; the family A DNA polymerases, the family B DNA polymerases, the family C DNA polymerases, and family X DNA polymerases (including TDTs). In order to limit the space needed for the alignment, we omitted DNA polymerase sequences that are very similar to the prototype DNA polymerase. The DNA polymerases not shown include: herpes virus type-2 (63),



adenovirus type-5 (64), bacteriophage T3 (65), and bacteriophage PZA (66).

#### **ACCURACY OF SEQUENCE DATA**

Whenever a sequence ambiguity existed in a published sequence, we contacted the authors to obtain the updated sequence information. We found that a few published amino acid sequences differ at one or more positions from their GenBank/EMBL entry. Again, we have communicated with the primary author to confirm the correct sequences.

The multiple alignment of the amino acid sequences was obtained by a series of pairwise alignments combined and adjusted by eye into larger and larger subsets of similar sequences. The process of combining and adjusting by eye was aided by modified versions of the MOTIF program (67) and the ALIGN program (68). The GAP and BESTFIT programs, from UWGCG (University of Wisconsin Genetic Computer Group) (69), initially generated the pairwise alignments, adjusted for maximum alignment that allowed for a considerable number of gaps. We then compressed these alignments by eye to give a more contiguous alignment. The alignment of the sequences for optimal similarity is straightforward in the areas of relatively conserved structure, but is much more arbitrary in the more varied sequence areas. The alignment of the varied areas should therefore be regarded as less than optimal in view of the difficulties concerned with multiple alignments in these areas.

Finally, we invite further correction from readers, and welcome suggested revisions and alternative alignments.

#### **ACKNOWLEDGEMENTS**

This work was supported by grant GM28013 from the National Institutes of Health and by grant NP-704 from the American Cancer Society.

#### REFERENCES

- Joyce, C.M., Kelley, W.S. and Grindley N.D.F. (1982) J. Biol. Chem. 257, 1958-1964.
- Lopez, P., Martinez, S., Diaz, A., Espinosa, M. and Lacks, S.A. (1989)
   J. Biol. Chem. 264, 4255-4263.
- Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R. and Gelfand, D.H. (1989) J. Biol. Chem. 264, 6427-6437.
- Leavitt, M.C. and Ito, J. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 4465-4469.
- 5. Dunn, J.J. and Studier, F.W. (1983) J. Mol. Biol. 166, 477-535.
- 6. Rådén, B. and Rutberg, L. (1984) J. Virol. 52, 9-15.
- 7. Foury, F. (1989) J. Biol. Chem. 264, 20552-20560.
- Iwasaki, H., Ishino, Y., Toh, H., Nakata, A. and Shinagawa, H. (1991)
   Mol. Gen. Genet. 226, 24-33.
- Jung, G., Leavitt, M.C., Hsieh, J.-C. and Ito, J. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 8287-8291.
- 10. Savilahti, H. and Bamford D.H. (1987) Gene 57, 121-130.
- 11. Yoshikawa, H. and Ito, J. (1982) Gene 17, 323-335.
- Matsumoto, K., Takano, H., Kim, C.I. and Hirokawa, H. (1989) Gene 84, 247-255.
- Spicer, E.K., Rush, J., Fung, C., Reha-Krantz, L.J., Karam, J.D. and Konigsberg, W.H. (1988) J. Biol. Chem. 263, 7478-7486.
- Wong, S.W., Wahl, A.F., Yuan, P.-M., Arai, N., Pearson, B.E., Arai, K.-i., Korn, D., Hunkapiller, M.W. and Wang, T.S.-F. (1988) EMBO J. 7, 37-47.
- Pizzagalli, A., Valsasnini, P., Plevani, P. and Lucchini, G. (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 3772-3776.
- 16. Morrison, A., Araki, H., Clark, A.B., Hamatake, R.K. and Sugino, A. (1990)

- Boulet, A., Simon, M., Faye, G., Bauer, G.A. and Burgers, P.M.J. (1989) *EMBO J.* 8, 1849-1854.
- Morrison, A., Christensen, R.B., Alley, J., Beck, A.K., Bernstine, E.G., Lemontt, J.F. and Lawrence, C.W. (1989) J. Bacteriol. 171, 5659-5667.
- Gibbs, J.S., Chiou, H.C., Hall, J.D., Mount, D.W., Retondo, M.J., Weller, S.K. and Coen, D.M. (1985) Proc. Natl. Acad. Sci. U.S.A. 82, 7969 – 7973.
- Kouzarides, T., Bankier, A.T., Satchwell, S.C., Weston, K., Tomlinson, P. and Barrell, B.G. (1987) J. Virol. 61, 125-133.
- Baer, R., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Séguin, C., Tuffnell, P.S. and Barrell, B.G. (1984) Nature 310, 207-211.
- 22. Davison, A.J. and Scott, J.E. (1986) J. Gen. Virol. 67, 1759-1816.
- Binns, M.M., Stenzler, L., Tomley, F.M., Campbell, J. and Boursnell, M.E.G. (1987) Nucl. Acids Res. 15, 6563-6573.
- Earl, P.L., Jones, E.V. and Moss, B. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 3659-3663.
- 25. Tomalski, M.D., Wu, J. and Miller, L.K. (1988) Virology 167, 591-600.
- Gingeras, T.R., Sciaky, D., Gelinas, R.E., Bing-Dong, J., Yen, C.E., Kelly, M.M., Bullock, P.A., Parsons, B.L., O'Neill, K.E. and Roberts, R.J. (1982)
   J. Biol. Chem. 257, 13475-13491.
- 27. Engler, J.A., Hoppe, M.S. and van Bree, M.P. (1983) Gene 21, 145-159.
- 28. Shu, L., Hong, J.S., Wei, Y.-f. and Engler, J.A. (1986) Gene 46, 187-195.
- Paillard, M., Sederoff, R.R. and Levings, C.S. III (1985) EMBO J. 4, 1125-1128.
- Stark, M.J.R., Mileham, A.J., Romanos, M.A. and Boyd, A. (1984) Nucl. Acids Res. 12, 6011-6030.
- Tommasino, M., Ricci, S. and Galeotti, C.L. (1988) Nucleic Acids Res. 16, 5863-5878.
- 32. Oeser, B. and Tudzynski, P. (1989) Mol. Gen. Genet. 217, 132-140.
- Kempken, F., Meinhardt, F. and Esser, K. (1989) Mol. Gen. Genet. 218, 523-530.
- Tomasiewicz, H.G. and McHenry, C.S. (1987) J. Bacteriol. 169, 5735-5744.
- Lancy, E.D., Lifsics, M.R., Munson, P. and Maurer, R. (1989) J. Bacteriol. 171, 5581-5586.
- Hammond, R.A., Barnes, M.H., Mack, S.L., Mitchener, J.A. and Brown, N.C. (1991) Gene 98, 29-36.
- Matsukage, A., Nishikawa, K., Ooi, T., Seto, Y. and Yamaguchi, M. (1987)
   J. Biol. Chem. 262, 8960-8962.
- 38. Abbotts, J., SenGupta, D.N., Zmudzka, B., Widen, S.G., Notario, V. and Wilson, S.H. (1988) *Biochemistry* 27, 901-909.
- SenGupta, D.N., Zmudzka, B.Z., Kumar, P., Cobianchi, F., Skowronski, J. and Wilson, S.H. (1986) Biochem. Biophys. Res. Comm. 136, 341-347.
- and Wilson, S.H. (1980) Biochem. Biophys. Res. Comm. 130, 341-347.
   Peterson, R.C., Cheung, L.C., Mattaliano, R.J., White, S.T., Chang, L.M.S. and Bollum F.J. (1985) J. Biol. Chem. 260, 10495-19502.
- Koiwai, O., Yokota, T., Kageyama, T., Hirose, T., Yoshida, S. and Arai, K.-i. (1986) Nucl. Acids Res. 14, 5777-5792.
- 42. Argos, P., Tucker, A.D. and Phillipson, L. (1986) Virology 149, 208-216.
- 43. Larder, B.A., Kemp, S.D. and Darby, G. (1987) EMBO J. 6, 160-175.
- 44. Hall, J.D. (1988) Trends Genet. 4, 42-46.
- 45. Reha-Krantz, L.J. (1988) J. Mol. Biol. 202, 711-724.
- Bernad, A., Zaballos, A., Salas, M. and Blanco, L. (1987) EMBO J. 6, 4219-4225.
- 47. Wang, T.S.-F., Wong, S.W. and Korn, D. (1989) FASEB 3, 14-21.
- Bernad, A., Blanco, L., Lazaro, J.M., Martin, G. and Salas, M. (1989) Cell 59, 219-228.
- Polesky, A.H., Steitz, T.A., Grindley, N.D.F. and Joyce, C.M. (1990) J. Biol. Chem. 24, 14579-14591.
- 50. Ito, J. and Braithwaite, D.K. (1990) Nucl. Acids Res. 18, 6716.
- 51. Blanco, L., Bernad, A. and Salas, M. (1991) Nucl. Acids Res. 19, 955.
- Kornberg, A. (1974) DNA replication. W.H. Freeman and Co., San Francisco.
- Bonner, C.A., Hays, S., McEntee, K. and Goodman M.F. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 7663-7667.
- Chen, H., Lawrence, C.B., Bryan, S.K. and Moses, R.E. (1990) Nucl. Acids Res. 18, 7185-7186.
- 55. McHenry, C.S.(1985) Mol. Cell. Biochem. 66, 71-85.
- Shepard, D., Oberfelder, R.W., Welch, M.W. and McHenry, C.S. (1984)
   J. Bacteriol. 158, 455-459.
- Scheuermann, R., Tam, S., Burgers, P.M.J., Lu, C. and Echols, H. (1983)
   Proc. Natl. Acad. Sci. U.S.A. 80, 7085-7089.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463 – 5467.
- 59. Dorsky, D.I. and Crumpacker, C.S. (1990) J. Virol. 64, 1394-1397.
- 60. Bernad, A., Lázaro, J.M., Salas, M. and Blanco, L. (1990) Proc. Natl. Acad.



- Jung, G., Leavitt, M.C., Schultz, M. and Ito, J. (1990) Biochem. Biophys. Res. Comm. 170, 1294-1300.
- 62. Neville, M.N. and Brown, N.C. (1972) Nature New Biol. 240, 80-82.
- 63. Tsurumi, T., Maeno, K. and Nishiyama, Y. (1987) Gene 52, 129-137.
- 64. Dekker, B.M.M. and Van Ormondt, H. (1984) Gene 27, 115-120.
- Beck, P.J., Gonzalez, S., Ward, C.L. and Molineux, I.J. (1989) J. Mol. Biol. 210, 687-701.
- Paces, V., Vlcek, C., Urbanek, P. and Hostomsky, Z. (1985) Gene 38, 45-56.
- Smith, H.O., Annau, T.M. and Chandrasegaran, S. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 826-830.
- Doolittle, R.F. and Feng, D.-F. (1990) In Doolittle, R.F. (ed.), Methods in Enzymol. – Molecular Evolution: Computer Analysis of Protein and Nucleic Acid Sequences. Academic Press, New York, Vol. 183, pp. 659-669.
- Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucl. Acids Res. 12, 387-395.

#### Classification of DNA polymerases

<b>A.</b>	Family A DNA polymerases								
	1.	References							
		a) E. coli DNA polymerase I	(1)						
		b) Streptococcus pneumoniae DNA polymerase I	(2)						
		c) Thermus aquaticus DNA polymerase I	(3)						
	2.	Bacteriophage DNA polymerases	(-)						
	2.		40						
		a) T5 DNA polymerase b) T7 DNA polymerase	(4) (5)						
		c) Spo2 DNA polymerase	(5) (6)						
	3.	Mitochondrial DNA polymerase	(4)						
		Yeast mitochondrial DNA polymerase (MIP1)	(7)						
	<b>1</b> 7	D TNA nekwanana							
В.	Family B DNA polymerases								
	1.	Bacterial DNA polymerases	401						
		E. coli DNA polymerase II	(8)						
	2.	Bacteriophage DNA polymerases							
		a) PRD1 DNA polymerase*	(9,10)						
		b) $\phi$ 29 DNA polymerase*	(11)						
		c) M2 DNA polymerase*	(12)						
		d) T4 DNA polymerase	(13)						
	3.	Eukaryotic DNA polymerases							
		a) Human DNA polymerase alpha	(14)						
		b) Yeast DNA polymerase I	(15)						
		c) Yeast DNA polymerase II d) Yeast DNA polymerase III (delta)	(16)						
		e) Yeast DNA polymerase Rev3	(17) (18)						
	4.	Viral DNA polymerases	(10)						
		a) Herpes-1 DNA polymerase	(19)						
		b) Human cytomegalovirus DNA polymerase	(20)						
		c) Epstein-Barr virus DNA polymerase	(21)						
		d) Varicella-Zoster virus DNA polymerase	(22)						
		e) Fowlpox virus DNA polymerase	(23)						
		f) Vaccinia virus DNA polymerase	(24)						
		g) Autographa californica nuclear	. (05)						
		polyhedrosis virus (AcMNPV) DNA polymeras h) Adenovirus-2 DNA polymerase*	e (25) (26)						
		i) Adenovirus-7 DNA polymerase*	(27)						
		j) Adenovirus-12 DNA polymerase*	(28)						
	5. Eukaryotic linear DNA plasmid encoded DNA polymerases								
		a) S-1 maize mitochondrial DNA polymerase*	(29)						
		b) Kluyveromyces lactis plasmid pGKL1 DNA polymerase							
		c) Kluyveromyces lactis plasmid pGKL2 DNA polymerase	* (31)						
		d) Claviceps purpurea plasmid pCLK1 DNA polymerase* e) Ascobolus immersus plasmid pAI2 DNA polymerase*	(32) (33)						
2	Family C DNA polymerases								
	Bacterial replicative DNA polymerases								
		a) E. coli DNA polymerase III α subunit	(34)						
		<ul> <li>b) Salmonella typhimurium DNA polymerase III α subun</li> <li>c) Bacillus subtilis DNA polymerase III</li> </ul>	it (35)						
D.	Fam	ily X DNA polymerases	(36)						
		a) Rat DNA polymerase β	(37)						
		b) Human DNA polymerase β	(38,39						
		c) Human terminal deoxynucleotidyltransferase (TdT)	(40)						
		d) Bovine terminal deoxynucleotidyltransferase (TdT)	(41)						
		e) Mouse terminal deoxynucleotidyltransferase (TdT)	(41)						

**Table 1.** The main families and subclassifications of DNA polymerases. Those DNA polymerases marked with a star (\*) are protein-primed DNA polymerases.



			50000000000000000000000000000000000000	20000-1000 2000-200-200 2000-200-200	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	
DANTYYDAKAPSFRIEATGGTKAGGAFTPED PROLIALIKELYOLLGLAR EYGTEADDYLASLAKKAEEGTEVRIL TADDU YOLLSDR.  275  275  275  275  275  370  375  400  1KTOVELEL TGGGLEVOOP AREEL GLFKKYEFKRITADVEAGKULOAKGAKPAAKOGTSYADGAEVTINTSYDNYT LIDEETLANTALEKAPVARDTETDSTDINSSANTYGLSFALEPGVANT PVAH  FYLLYOVELEL TGGGLEVOOP AREEL GLFKKYEFKRITADVEAGKULOAKGAKPAAKOGTSYADGAEVTINTSYDNYT LIDEETLANTALEKAPVARDTETDSTDINSSANTYGLSFALEPGVANT PVAH  FYLLYOVELEL TGGGLEVOOP AREEL GLFKKYEFKRITADVEAGKULOAKGAKPATATVEAL II NOTINGESTEAT TI TOTTAL TATATVEAT TI TOTTAL GAVYNAWI GAVARDSELLANTAGUS ALL GAVELY TOTTAL TOTTAL GAVYNAWI GAVARDSELTAL YGGSTEAT CHOOST LIDEET CONTACTOR TOTTAL GAVARDSELTAL GAVARDSELTAL GAVARDSELTAL YGGSTEAT CHOOST LIDEET CONTACTOR TOTTAL GAVARDSELTAL GAVARDSELTAL YGGSTEAT CHOOST LIDEET CONTACTOR TOTTAL GAVARDSELTAL GAVAR	ANTICATAR ANTICA	SOCIETA SOCIET	TOOD  RIMAN SDOKGLIAF  FYLAN SOCENIARY  FYLAN SOCENIARY  RECAN  RECAN  LAS YOUR STANKEST OF TARE	1200 APMOGTAAD SP 1063AAD MIVOGTAAD MICOSAGA 1108ASS TLLGSAGA 1108ASS TLLGSAGA 1108ASS MAIOSSGVO WAIOSSGVO WAIOSSGVO	1400	
LEFALEPGI LESFALEPGI LESFALMES		######################################	HE SECOND	FERNARIA FERNARIA FERNARIA FERNARIA FERNARIA FERNARIA FERNARIA	13.75 MOVESOKRÓVNILESEVLRECTSKEYARÓ 15.75	
WLWERTG	FSSVOLLESYTRIG FSSVOLLESTOR 575 FSVOLKAGUS 575 FSVOKRAGUS FSVCARAG	PAKELY PVEHV VAKKLTO				
375 375 WENGKFYDEMISANI VENGKFYDEMG	MANUAL PROPERTY AND	TTS LIAGEE LIAGEE CVAGLIAFT SVENSPVIVSKI	POYLYSDYSOIERRIA EDEVLISSOYSOIERRIA FORKINADITABLERVIA POTRIVANGIBERVIA FOREFRYDESELERVI FOREFRYDESELERVIA FOREFRYDESELERVIA FOREFRYDESELERVIA	1175 1175 1175 1175 1175 1175 1175 1175	1375 2000 1375 1575	
AKLEKAPYFAFDTE IGLEDLYYSGPOVE VYNNYIĞPVAFOSE	GWKYVEGUÞVEVAYPOEELVFBUETSBUETSBALLES GWKYVEGUÞVEVAYPOEELVFBUETSBUETSBUETSBUETSBUETSBUETSBUETSBUETS	T750 T750 T750 T750 T750 T750 T750 T750		######################################	FLAMOVE	
KAUJAKLE RPJAIGLE	AL		SECRETARY SECRET	1150 THE POLICE SHOTE STATE OF THE POLICE SHOTE SECURITY OF THE POLICE SHOTE	YTPEFLK)	
350 TILDEETLK TATIOTKA	MKYVPGA 550 550 CEARCL CEARCL CCR I CLK	750 ACLIENS CETLENS COCKARES COCCARES COCKARES C	950 GARINGO GA	1150 CGRR 11P CGRRETH	1350 KVCKGLGWYKRILGA KVKEFVKEFYKSYT	
ISYDNYV EGAFLSK	EVLRKPG GROODL ORFGVOLC TYTALDER FASTER FASTER FASTER		674 74 74 74 74 74 74 74	COCYVETLY RGTVETLY RGTVETLY RGFTXSH RG	KFTAVCE	
APEYTATI KENLINDI	TENVARP FRXVVIII WWA PELL MILIO TO MELL DELLO TO MELLO TO	APVLASE APVLASE SIO LANGE ESERCIA		A A A A A A A A A A A A A A A A A A A	NSKYKOD I DSKYSO	
325 325 GETSYADE	72 22 22 22 22 22 22 22 22 22 22 22 22 2	PER	925 IPKI VPHO	1125 1125 1125 1125 1125 1125 1125 1125	1325 KKLAMOKEHPO KPHSKLGGPSLG	
325 CAKPAAKPOETKISVA HKPVEATVENIHDI	######################################	2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	925 MAAF LINNEKTNOLATTERLYPHGTTTRRATE	AIR STANKER	GFEYKKU	
EAGKULOAK	CONTRACTOR	EPICIE KANICIE TATOVINE TESSATOR	A S S S S S S S S S S S S S S S S S S S	PEGENT PE	VKGVLLD.	
SOO SOO CKYEFKWIADVE	SOO SOO SOOLYATI GOOLYATI PPECAFVI PACHECIA	700 KLAVIVET AALSERIF ELLEKLES	NACON PAREE NACON PAREE NACON PAREE NACON PAREE NACON PAREE OVER PAREE OVER PAREE NACON PA	1100 FPGVLEYN FPGVLEYN FPGLKRUIT FFKFVGG FKFVGG	1300 LIDDOETRY LIDDOETRY 1500	
LFKKYEF	LITOPISFP PPLOGRSLOEHFOX IGRINSEPYKSFCOOKFTENARPA SOO S25 SOO S25 SELLEEAN-PPELA YORKELLOND FLOCKER LINYOFKYVUL PKALLEEAN-PPELA YORK SKEPHANDLALLA KOOKYVUL PKALLEEAN-PPELA YORK SKEPHANDLALLA KOOKYVUL PKALLEEAN-PPELA YORK SKEPHANDLANG SKEPHAND SKEPHANDLANG SKEPHAND SKEPHANDLANG SKEPHAND SKEPHAND SKEPHAND SKEPHAND SKEPHAND SKEPHAND SKEPHAND SKEPHAND SKEPHAND SKEP	75  MILAGERAGPFOTAL TEELVELAKTOPUS TOO  TEELSTEELTEELTEELTEELTEELTEELTEELTEELTEE	L SOLVERS	TOLYFER TOTYFER TETYFOR TETYFOR TELVKKFLE TELVKKFLE TETYFOR	13go Grakkampsi acidodettryvkävi lääfetykki jäjäkykönskykookt invakoldavkai i Ga* Svirkevamolitpsaki aipägeliä imai läkänsi Gapsi jäskysänä yyreetyi 15go 15go	
OPAAEELL GOOTGO	LUGKNTL	KAIEELY	11000 11000	PRKEADK SRKEADKA PYKEADKA TVADAKE TVADAKE TKLYEGEL KLYEGEL	C. S. C. C. S. C. C. S. C. C. S. C.	
275 CEGLEVA	SKKSLIOH 475 LLEFGSLLH RORPHFHK	675 ERNGFPFD	PKTGRVHTSTHOAVTATGR LAGGKHTRFWOOLTGTGR PRTGRLHTFWOOLTGTGR DAGGC RTGHEHNTTSGR AFGGKHGSWIPHGAVTGR OFFGGSTTGSSTLASARER	1075 AKTGEPTEN A	1275 GENLDOAN". GENLDOAN". SERFAGGSRD FOYTKU. FOYTKU.	
KTOVELEL	ÉSSOKLMÉLSKÁSLKÖNGLUGKATLÍ TÖP I SPPLPP LOGASLDENFOK LORFHSEPYKSFCOOK PTEMARÞAENLRPGUNKYVPGANÞVEVA POEELVPDVETT WYSDIFMALLESYKLIKES 473 550 575 575 570 570 575 575 575 575 576 575 575 575 575 575	RAAM.LAKOR VQIESKIVO	SAPERATE SAP	ALLEGAM	170 170 170 170 170 170 170 170 170 170	
AL KEDGE	CG C	LYOSKY	KANCE SELL	GLARO FRISHE RIVANSVIE KKIGO ALIGNGA.	AVELLAVE COLONIA COLON	
O CONKEVATI	OTLESSIGNOVESSGROEN 450 1874	Trysesu	THE SECTION AND ADDRESS OF THE SECTION ADDRESS OF THE SECTION ADDRESS OF THE SECTION AND ADDRESS OF THE SECTION ADDRESS	NAME OF THE PROPERTY OF THE PR	A MEN COLUMN COL	
KTHAAKL	TITANY TITANY TYPE TYPE TYPE TYPE TYPE TYPE TYPE TYP	6 DFTDVGY PTKLNDM	KYTTLEAK KYT	1050 NAKALINFOLI NAKALINFOLI NAKALINFOLI FAKILI FAKILINFOLI FAKILINFOLI FAKILI F	DENEMBER 128	
AGL SFRG	**************************************	KS.	MEDATE STATES OF THE STATES OF	7	1	
TARE TO THE TARE T	FERAL KON THE TREE TO THE TREE TO THE TREE TO THE TREE TREE TREE TREE TREE TREE TREE	于 学 子 子 子 子 子 子 子 子 子 子 子 子 子 子 子 子 子 子	LALO LAEDAPIC VANDOPPIC VIENCE EN POLIC VIENCE EN POLIC PERONE EN POLICE	PET REPON CCEPROVE FKVPVES	MAIN CONTRACTOR OF CONTRACTOR	
COUNTY TANK	YYGECARIFICATION CONTRACTOR CONTR	PEREKF GGENT PEREKF GGENT SOKNYF PVFLKKCP	825 TSEEVIER TSAAVIER	1025 AAEVFGEP AMMVFKGE AMMVFKCL O	1225 1225 117 116 1625 1425	
MEGICAL FERNANL FORMILATRI PARLIG.  25  25  25  26  26  27  29  20  20  20  20  20  20  20  20  20	HER ALE	GOCHE GROVEKA THENCARRA TALLER RAILER SO Y FUR SO Y FUR S	TEKTPO TEKTPO TEKTPO TOK TEKT TEKTPO	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG		
MEGENCI GENERAL FORMAN FRANCIA LE I SEGEPORAN GRASLIKAL. REG 25 25 20 25 00 100 CGLO T. VAPPEL AGI SFROAT MAALL CONKENAN LS YOLAT ON IPONYGENGENGEN SYOLAT ON IPONYGENGENGENGEN SYOLAT ON IPONYGENGENGENGENGENGENGENGENGENGENGENGENGENG	LWWFECHLWWRRPLRYGFCARWFSTRCMAAEAPR HPYGIGY LGESLGROWFGSCGGGGGE 425  FTV DOI SOON SEER ALELLEP LEBEKALEYGONI, KTORG LANYGI FTV DOI SOON SEER ALELLEP LEBEKALEYGONI, KTORG LANYGI FTV TORG LOS TORG LOS TORG TO THE TORG TO T	625 850 RYAEDAVTI, QLILLUM-DOLOGII TULODERIT PEGNARUGELI PYALODORI TEURIF FERENT VICTORI	625 61CP. IXTPGGAPSTSEEVLELLOYPPKY LETRGLAKLKSTRULLUM 1CG. PA CT. O. TKOOSTSEEVLELLOYPPKY LETRGLAKLKSTRULLUM 1CG. PA CT. O. TKOOSTSEEVLELLOYPPKK LOYPET KLKSTRULLUM 1CG. PA CT. O. TKOOSTSEAVLELLONDP PKT LETR HI TKOT TKOTSECTORS AGAPTET. TKO CAPPONOCH EGWOODECOAA DI TEST HI TGOT TGOSSECOKAALTY AGAPTET. TSC CAPPONOCH EGWOODECOAA DI TEST HI TGOT TGOSSECOKAALTY 1CLOSSOSCEEE FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. LOSSOSCEEE FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. LOSSOSCEEE FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. LOSSOSCEEE FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. LOSSOSCEEE FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. LOSSOSCEEE FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. LOSSOSCEEE FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. LOSSOSCEEP FINANT GGCTG. VL. KPAPPONOCPT MCTAL L'KSTMIFFER. GALESSEEV. 1VL. L'KSTMIFFER. CONTRACT L'KSTMIFFER. CONTRACT L'KSTMIFFER. GALESSEEV. 1VL. L'KSTMIFFER. CONTRACT L'KSTMIFFER. CONTRACT L'KSTMIFFER. GALESSEEV. 1VL. L'KSTMIFFER. CONTRACT L'KSTMIFFER. GALESSEE CONTRACT L'KSTMIFFER. GALESSEEV. 1VL. L'KSTMIFFER. CONTRACT L'KSTMIFFER. GALESSEE CONTRACT L'KSTMIFFER. GALESSEEV. 1VL. L'KSTMIFFER. CONTRACT	1050  AEGO INATALEVELLET  VTSEORREACA INFGLI 1'GOSAFGLARO  INANIEDOS OS ON INTERNACIONE OS	1250 1250 1250 1250 1250 1250 1250 1250	
805.0-		\$01.000000 \$01.000000	22000000 22000000 22000000 2000000	525226855 525226855	######################################	9
Consensus 7.75	S.p. pol 1	200 SE	0 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	Consensus 1	20 20 20 20 20 20 20 20 20 20 20 20 20 2	

\*\*Igure 14. Family A DNA polymerases—E. coli DNA pol. I (E.c. pol I)(1), Streptococcus pneumoniae DNA pol. I (S.p. pol I)(2), Thermus aquaticus DNA pol. I (Taq pol I)(3), T5 DNA pol. (4), T7 DNA pol. (6), and yeast mitochondrial DNA pol. (MIP1)(7).

# DOCKET

## Explore Litigation Insights



Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

## **Real-Time Litigation Alerts**



Keep your litigation team up-to-date with **real-time** alerts and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

## **Advanced Docket Research**



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

### **Analytics At Your Fingertips**



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

#### API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

#### **LAW FIRMS**

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

#### **FINANCIAL INSTITUTIONS**

Litigation and bankruptcy checks for companies and debtors.

### **E-DISCOVERY AND LEGAL VENDORS**

Sync your system to PACER to automate legal marketing.

