

# Compilation, alignment, and phylogenetic relationships of DNA polymerases

Dan K. Braithwaite and Junetsu Ito\*

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

Received January 14, 1993; Accepted January 22, 1993

## INTRODUCTION

This is an update of an earlier compilation and alignment of DNA polymerase sequences (Ito and Braithwaite, 1991). As in the previous compilation, we attempted to compile complete sequences, to facilitate the identification of conserved and viable regions of the DNA polymerases (1). This update includes, for the first time, three DNA polymerase sequences from *Archaea* (2); two new members of the Family A DNA polymerases; and 19 new members of the Family B DNA polymerases. In addition, we included nucleases that have related amino acid sequences to *E. coli* DNA polymerase I, and the sequence of *E. coli* DNA polymerase III ( $\epsilon$ -subunit) was aligned to Family C due to its homology to *Bacillus subtilis* DNA polymerase III.

As in the previous compilation (1), Family A DNA polymerases are named for their homology to the product of the *polA* gene specifying *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* encoding *E. coli* DNA polymerase III alpha subunit.

Table 1 summarizes the molecular weights and isoelectric points of each DNA polymerase and nuclease. Table 1 also serves as a reference guide to the sequences shown in Figures 1A, 1B, and 1C. Since no new sequences were published for the Family X DNA polymerases ( $\beta$ -like), we have excluded them from this compilation.

## SEQUENCE ALIGNMENT

The multiple alignments of the amino acid sequences for this update were performed in most cases by merely adding on to our original alignments (1) where possible. Due to the large number of sequences added to the alignment for Family B we have changed the original alignment in some areas between obvious blocks of conserved sequences. The newer sequences were added by aligning each to the closest related sequence already aligned, or in many cases to the closest related group of sequences already aligned. A more recent addition to the UWGCG (University of Wisconsin Genetic Computer Group) program package, PILEUP, a multiple alignment program, was used extensively to try and locate significant homology in groups

of closely related sequences. These newly formed groups of highly related sequences were then regapped to conform with the entire alignment based upon the previous alignment of those sequences in the new group from the original alignment. As in the previous paper, all the final adjustments had to be made by eye and, as stated above, in Family B the added sequences led to some improvements to the original alignment that became evident to the eye when they were being combined with the entire alignment by hand.

## GENERATION OF PHYLOGENETIC TREES FOR THE DNA POLYMERASE DOMAINS

Using Felsenstein's PHYLIP program package (71), specifically the programs named in the outline below, we generated phylogenetic trees for the 9 Family A DNA polymerases (Figures 2A and 2B) and for the 47 Family B DNA polymerases (Figures 3A and 3B). The trees for Family A were created from the alignment in Figure 1A using the most conserved regions found at the following positions: 798 to 814, 877 to 998, 1047 to 1090, 1104 to 1123, 1131 to 1158, 1175 to 1206, 1236 to 1251, 1284 to 1305, 1322 to 1340, and 1365 to 1379. These conserved regions were recombined and 100 bootstrap samples were generated using SEQBOOT program. Using the DNADIST program, we turned the samples into distance matrices using the Kimura-2 parameter method. The resulting matrices were then input to the NEIGHBOR program using the UPGMA method to produce approximately 100 trees. Finally those trees were reduced to a single tree using the CONSENSE program. This final tree was then plotted for publication using two different methods. The trees in Figures 2A and 3A were created by the DRAWGRAM program setup to produce a phenogram type tree and the trees in Figures 2B and 3B were created by the DRAWTREE program. The trees for Family B were created from the alignment in Figure 1B, according to the same procedure, using the most conserved regions found at the following positions: 1407 to 1760, 1885 to 1901, 1956 to 1990, 2081 to 2100, 2181 to 2210, and 2280 to 2320. The Family B DNA polymerases can be subdivided into two subfamilies, the protein-primed DNA polymerase subfamily and the RNA-primed DNA polymerase subfamily.

\* To whom correspondence should be addressed







Table with 10 columns (625-800) and 1000 rows of nucleotide sequences. Each row contains a sequence of nucleotides aligned across the columns. The table is organized into two main sections, with the top section ending at row 499 and the bottom section starting at row 500. The columns are labeled with their respective sequence positions: 625, 650, 675, 700, 725, 750, 775, and 800. The rows are labeled with various identifiers such as Adeno-1, Adeno-2, kalilo, and others, representing different genetic constructs or sequences. The sequences consist of the four nucleotides: Adenine (A), Guanine (G), Cytosine (C), and Thymine (T), with some variations indicated by lowercase letters or specific symbols.



	1425	1450	1475	1500	1525	1550	1575	1600
PRD1	191	1425	1450	1500	1525	1550	1575	1600
Adeno-1	214	1425	1450	1500	1525	1550	1575	1600
Adeno-2	513	1425	1450	1500	1525	1550	1575	1600
Adeno-3	511	1425	1450	1500	1525	1550	1575	1600
Adeno-7	289	1425	1450	1500	1525	1550	1575	1600
Adeno-10	607	1425	1450	1500	1525	1550	1575	1600
ka110	794	1425	1450	1500	1525	1550	1575	1600
ka112	676	1425	1450	1500	1525	1550	1575	1600
mar	412	1425	1450	1500	1525	1550	1575	1600
prM	641	1425	1450	1500	1525	1550	1575	1600
pgK1	607	1425	1450	1500	1525	1550	1575	1600
pgK2	607	1425	1450	1500	1525	1550	1575	1600
HSV-1	605	1425	1450	1500	1525	1550	1575	1600
Echv-1	693	1425	1450	1500	1525	1550	1575	1600
ZVZ	627	1425	1450	1500	1525	1550	1575	1600
Human	579	1425	1450	1500	1525	1550	1575	1600
Bovine	606	1425	1450	1500	1525	1550	1575	1600
S.c. pol	606	1425	1450	1500	1525	1550	1575	1600
S.p. pol	574	1425	1450	1500	1525	1550	1575	1600
S.f. pol	574	1425	1450	1500	1525	1550	1575	1600
FPV	592	1425	1450	1500	1525	1550	1575	1600
Human	508	1425	1450	1500	1525	1550	1575	1600
Chimp	511	1425	1450	1500	1525	1550	1575	1600
Human	511	1425	1450	1500	1525	1550	1575	1600
Lary	511	1425	1450	1500	1525	1550	1575	1600
E.c. pol	471	1425	1450	1500	1525	1550	1575	1600
P. furiosus	382	1425	1450	1500	1525	1550	1575	1600
S.c. pol	519	1425	1450	1500	1525	1550	1575	1600
Human	830	1425	1450	1500	1525	1550	1575	1600
D.m. pol	830	1425	1450	1500	1525	1550	1575	1600
S.c. pol	830	1425	1450	1500	1525	1550	1575	1600
S.p. pol	830	1425	1450	1500	1525	1550	1575	1600
T.B. pol	723	1425	1450	1500	1525	1550	1575	1600
Consensus	1401	1425	1450	1500	1525	1550	1575	1600

	1625	1650	1675	1700	1725	1750	1775	1800
PRD1	283	1625	1650	1675	1700	1725	1775	1800
Adeno-1	429	1625	1650	1675	1700	1725	1775	1800
Adeno-2	604	1625	1650	1675	1700	1725	1775	1800
Adeno-3	604	1625	1650	1675	1700	1725	1775	1800
Adeno-7	672	1625	1650	1675	1700	1725	1775	1800
Adeno-10	289	1625	1650	1675	1700	1725	1775	1800
ka31	780	1625	1650	1675	1700	1725	1775	1800
ka32	780	1625	1650	1675	1700	1725	1775	1800
mar	412	1625	1650	1675	1700	1725	1775	1800
prM	713	1625	1650	1675	1700	1725	1775	1800
pgK1	700	1625	1650	1675	1700	1725	1775	1800
pgK2	700	1625	1650	1675	1700	1725	1775	1800
HSV-1	726	1625	1650	1675	1700	1725	1775	1800
Echv-1	726	1625	1650	1675	1700	1725	1775	1800
ZVZ	636	1625	1650	1675	1700	1725	1775	1800
Human	636	1625	1650	1675	1700	1725	1775	1800
Bovine	625	1625	1650	1675	1700	1725	1775	1800
S.c. pol	625	1625	1650	1675	1700	1725	1775	1800
S.p. pol	625	1625	1650	1675	1700	1725	1775	1800
S.f. pol	625	1625	1650	1675	1700	1725	1775	1800
FPV	621	1625	1650	1675	1700	1725	1775	1800
Human	552	1625	1650	1675	1700	1725	1775	1800
Bovine	660	1625	1650	1675	1700	1725	1775	1800
S.c. pol	660	1625	1650	1675	1700	1725	1775	1800
S.p. pol	660	1625	1650	1675	1700	1725	1775	1800
S.f. pol	660	1625	1650	1675	1700	1725	1775	1800
Human	591	1625	1650	1675	1700	1725	1775	1800
Chimp	591	1625	1650	1675	1700	1725	1775	1800
Human	591	1625	1650	1675	1700	1725	1775	1800
Lary	591	1625	1650	1675	1700	1725	1775	1800
E.c. pol	429	1625	1650	1675	1700	1725	1775	1800
P. furiosus	442	1625	1650	1675	1700	1725	1775	1800
S.c. pol	552	1625	1650	1675	1700	1725	1775	1800
Human	974	1625	1650	1675	1700	1725	1775	1800
S.c. pol	974	1625	1650	1675	1700	1725	1775	1800
S.p. pol	974	1625	1650	1675	1700	1725	1775	1800
S.f. pol	974	1625	1650	1675	1700	1725	1775	1800
T.B. pol	862	1625	1650	1675	1700	1725	1775	1800
Consensus	1601	1625	1650	1675	1700	1725	1775	1800

	1825	1850	1875	1900	1925	1950	1975	2000
PRO1	370	370	370	370	370	370	370	370
Adeno	400	400	400	400	400	400	400	400
Adeno	420	420	420	420	420	420	420	420
Adeno	440	440	440	440	440	440	440	440
Adeno	460	460	460	460	460	460	460	460
Adeno	480	480	480	480	480	480	480	480
Adeno	500	500	500	500	500	500	500	500
Adeno	520	520	520	520	520	520	520	520
Adeno	540	540	540	540	540	540	540	540
Adeno	560	560	560	560	560	560	560	560
Adeno	580	580	580	580	580	580	580	580
Adeno	600	600	600	600	600	600	600	600
Adeno	620	620	620	620	620	620	620	620
Adeno	640	640	640	640	640	640	640	640
Adeno	660	660	660	660	660	660	660	660
Adeno	680	680	680	680	680	680	680	680
Adeno	700	700	700	700	700	700	700	700
Adeno	720	720	720	720	720	720	720	720
Adeno	740	740	740	740	740	740	740	740
Adeno	760	760	760	760	760	760	760	760
Adeno	780	780	780	780	780	780	780	780
Adeno	800	800	800	800	800	800	800	800
Adeno	820	820	820	820	820	820	820	820
Adeno	840	840	840	840	840	840	840	840
Adeno	860	860	860	860	860	860	860	860
Adeno	880	880	880	880	880	880	880	880
Adeno	900	900	900	900	900	900	900	900
Adeno	920	920	920	920	920	920	920	920
Adeno	940	940	940	940	940	940	940	940
Adeno	960	960	960	960	960	960	960	960
Adeno	980	980	980	980	980	980	980	980
Adeno	1000	1000	1000	1000	1000	1000	1000	1000

Human pol I  
 Bovine pol I  
 S.c. pol I  
 P. furiosus  
 S.c. pol II  
 Human pol II  
 S.c. pol II  
 P. furiosus  
 Human pol III  
 S.c. pol III  
 P. furiosus  
 Consensus



	2225	2250	2275	2300	2325	2350	2375	2400
PRD1	547	199						533
Adeno-5	573	185						572
Adeno-7	1053	1033						1053
Adeno-7	1122	1122						1083
ka	1199							1132
ka	1212							917
mar	1085							1202
mar	1002							1097
BSKL	286							965
pkL1	286							964
pkL2	988							994
pkL1	1049							992
pkL2	1049							1200
EBV	935							1194
HV-1	1031							1015
HV-2	1061							1203
HV-3	972							1097
HV-6	920							1012
CC-IV	943							993
Human pol 6	916							1046
Human pol 6	919							1045
S.c. pol 6	911							1034
S.c. pol 6	911							1037
P.p. pol 6	905							989
FvV	898							937
Vac.V	898							937
PCMV	843							984
PCMV	862							1006
Lampy	859							998
E.c. pol 14	694							774
P. furiosus	694							775
S. solfata	809							892
S.c. pol 11	1083							1208
D.m.	1178							1208
S.c. pol 6	1162							1311
S.p. pol 6	1142							1303
S.p. pol 6	1142							1213
I.S.c. Rev 5	1301							1213
I.S.c. Rev 5	1301							1213
Consensus	2201							2400

	2425	2450	2475	2500	2525	2550	2575	2600
PRD1	547	199						533
Adeno-5	573	185						572
Adeno-7	1053	1033						1053
Adeno-7	1122	1122						1083
ka	1199							1132
ka	1212							917
mar	1085							1202
mar	1002							1097
BSKL	286							965
pkL1	286							964
pkL2	988							994
pkL1	1049							992
pkL2	1049							1200
EBV	935							1194
HV-1	1031							1015
HV-2	1061							1203
HV-3	972							1097
HV-6	920							1012
CC-IV	943							993
Human pol 6	916							1046
Human pol 6	919							1045
S.c. pol 6	911							1034
S.c. pol 6	911							1037
P.p. pol 6	905							989
FvV	898							937
Vac.V	898							937
PCMV	843							984
PCMV	862							1006
Lampy	859							998
E.c. pol 14	694							774
P. furiosus	694							775
S. solfata	809							892
S.c. pol 11	1083							1208
D.m.	1178							1208
S.c. pol 6	1162							1311
S.p. pol 6	1142							1303
S.p. pol 6	1142							1213
I.S.c. Rev 5	1301							1213
I.S.c. Rev 5	1301							1213
Consensus	2201							2400

	2425	2450	2475	2500	2525	2550	2575	2600
PRD1	547	199						533
Adeno-5	573	185						572
Adeno-7	1053	1033						1053
Adeno-7	1122	1122						1083
ka	1199							1132
ka	1212							917
mar	1085							1202
mar	1002							1097
BSKL	286							965
pkL1	286							964
pkL2	988							994
pkL1	1049							992
pkL2	1049							1200
EBV	935							1194
HV-1	1031							1015
HV-2	1061							1203
HV-3	972							1097
HV-6	920							1012
CC-IV	943							993
Human pol 6	916							1046
Human pol 6	919							1045
S.c. pol 6	911							1034
S.c. pol 6	911							1037
P.p. pol 6	905							989
FvV	898							937
Vac.V	898							937
PCMV	843							984
PCMV	862							1006
Lampy	859							998
E.c. pol 14	694							774
P. furiosus	694							775
S. solfata	809							892
S.c. pol 11	1083							1208
D.m.	1178							1208
S.c. pol 6	1162							1311
S.p. pol 6	1142							1303
S.p. pol 6	1142							1213
I.S.c. Rev 5	1301							1213
I.S.c. Rev 5	1301							1213
Consensus	2201							2400

	2425	2450	2475	2500	2525	2550	2575	2600
PRD1	547	199						533
Adeno-5	573	185						572
Adeno-7	1053	1033						1053
Adeno-7	1122	1122						1083
ka	1199							1132
ka	1212							917
mar	1085							1202
mar	1002							1097
BSKL	286							965
pkL1	286							964
pkL2	988							994
pkL1	1049							992
pkL2	1049							1200
EBV	935							1194
HV-1	1031							1015
HV-2	1061							1203
HV-3	972							1097
HV-6	920							1012
CC-IV	943							993
Human pol 6	916							1046
Human pol 6	919							1045
S.c. pol 6	911							1034
S.c. pol 6	911							1037
P.p. pol 6	905							989
FvV	898							937
Vac.V	898							937
PCMV	843							984
PCMV	862							1006
Lampy	859							998
E.c. pol 14	694							774
P. furiosus	694							775
S. solfata	809							892
S.c. pol 11	1083							1208
D.m.	1178							1208
S.c. pol 6	1162							1311
S.p. pol 6	1142							1303
S.p. pol 6	1142							1213
I.S.c. Rev 5	1301							1213
I.S.c. Rev 5	1301							1213
Consensus	2201							2400

	2425	2450	2475	2500	2525	2550	2575	2600
PRD1	547	199						533
Adeno-5	573	185						572
Adeno-7	1053	1033						1053
Adeno-7	1122	1122						1083
ka	1199							1132
ka	1212							917
mar	1085							1202
mar	1002							1097
BSKL	286							965
pkL1	286							964
pkL2	988							994
pkL1	1049							992
pkL2	1049							1200
EBV	935							1194
HV-1	1031							1015
HV-2	1061							1203
HV-3								



	3025	3050	3075	3100	3125	3150	3175	3200
PRD1								
#2								
#2								
Adeno-2								
Adeno-1								
Adeno-3								
Adeno-1								
Adeno-1								
kalilo								
pa12								
mar								
sch								
peh								
pac11								
pac12								
pac13								
pac14								
pac15								
pac16								
pac17								
pac18								
pac19								
pac20								
pac21								
pac22								
pac23								
pac24								
pac25								
pac26								
pac27								
pac28								
pac29								
pac30								
pac31								
pac32								
pac33								
pac34								
pac35								
pac36								
pac37								
pac38								
pac39								
pac40								
pac41								
pac42								
pac43								
pac44								
pac45								
pac46								
pac47								
pac48								
pac49								
pac50								
pac51								
pac52								
pac53								
pac54								
pac55								
pac56								
pac57								
pac58								
pac59								
pac60								
pac61								
pac62								
pac63								
pac64								
pac65								
pac66								
pac67								
pac68								
pac69								
pac70								
pac71								
pac72								
pac73								
pac74								
pac75								
pac76								
pac77								
pac78								
pac79								
pac80								
pac81								
pac82								
pac83								
pac84								
pac85								
pac86								
pac87								
pac88								
pac89								
pac90								
pac91								
pac92								
pac93								
pac94								
pac95								
pac96								
pac97								
pac98								
pac99								
pac100								
pac101								
pac102								
pac103								
pac104								
pac105								
pac106								
pac107								
pac108								
pac109								
pac110								
pac111								
pac112								
pac113								
pac114								
pac115								
pac116								
pac117								
pac118								
pac119								
pac120								
pac121								
pac122								
pac123								
pac124								
pac125								
pac126								
pac127								
pac128								
pac129								
pac130								
pac131								
pac132								
pac133								
pac134								
pac135								
pac136								
pac137								
pac138								
pac139								
pac140								
pac141								
pac142								
pac143								
pac144								
pac145								
pac146								
pac147								
pac148								
pac149								
pac150								
pac151								
pac152								
pac153								
pac154								
pac155								
pac156								
pac157								
pac158								
pac159								
pac160								
pac161								
pac162								
pac163								
pac164								
pac165								
pac166								
pac167								
pac168								
pac169								
pac170								
pac171								
pac172								
pac173								
pac174								
pac175								
pac176								
pac177								
pac178								
pac179								
pac180								
pac181								
pac182								
pac183								
pac184								
pac185								
pac186								
pac187								
pac188								
pac189								
pac190								
pac191								
pac192								
pac193								
pac194								
pac195								
pac196								
pac197								
pac198								
pac199								
pac200								
pac201								
pac202								
pac203								
pac204								
pac205								
pac206								
pac207								
pac208								
pac209								
pac210								
pac211								
pac212								
pac213								
pac214								
pac215								
pac216								
pac217								
pac218								
pac219								
pac220								
pac221								
pac222								
pac223								
pac224								
pac225								
pac226								
pac227								
pac228								
pac229								
pac230								
pac231								
pac232								
pac233								
pac234								
pac235								
pac236								
pac237								
pac238								
pac239								
pac240								
pac241								
pac242								
pac243								
pac244								
pac245								
pac246								
pac247								
pac248								
pac249								
pac250								
pac251								
pac252								
pac253								
pac254								
pac255								
pac256								
pac257								
pac258								
pac259								
pac260								
pac261								
pac262								
pac263								
pac264								
pac265								





**Table 1.** The main families and subclassifications of DNA polymerases. The number of amino acids, molecular weight and isoelectric point of each DNA polymerase or nuclease were computed using the Proteinsort program from the UWGCG. Those Family B DNA polymerases marked with a star (\*) are protein-primed DNA polymerases. The sequence marked with (+) pEM (62) is not a complete sequence missing some unknown number of amino acids on the N-terminus and so the numbers in the table only represent this short sequence.

	Amino acids (No.)	Mol. Wt.	Isoelectric pt.	Reference
<b>A. Family A DNA polymerases</b>				
<b>1. Bacterial DNA polymerases</b>				
a) <i>E.coli</i> DNA polymerase I	928	103,117	5.37	(3)
b) <i>Streptococcus pneumoniae</i> DNA polymerase I	877	99,078	4.78	(4)
c) <i>Thermus aquaticus</i> DNA polymerase I	832	93,909	6.38	(5)
d) <i>Thermus flavus</i> DNA polymerase I	831	93,783	6.00	(6)
<b>2. Bacteriophage DNA polymerases</b>				
a) T5 DNA polymerase	829	94,410	6.19	(7)
b) T7 DNA polymerase	704	79,691	6.45	(8)
c) Spo1 DNA polymerase	924	106,808	5.34	(9)
d) Spo2 DNA polymerase	648	72,561	8.50	(10)
<b>3. Mitochondrial DNA polymerase</b>				
Yeast mitochondrial DNA polymerase (MIP1)	1254	143,479	9.23	(11,12,13)
<b>4. 5' to 3' Exonucleases with homologous sequences to <i>E.coli</i> DNA polymerase I</b>				
a) T4 RNase H (gp 33.2)	305	35,558	9.00	(14,15)
b) T5 Exonuclease (gp D15)	291	33,448	5.12	(7,16)
c) T7 Exonuclease (gp 6)	348	40,126	4.54	(4,8)
<b>B. Family B DNA polymerases</b>				
<b>1. Bacterial DNA polymerase</b>				
<i>E.coli</i> DNA polymerase II	783	90,020	6.85	(17)
<b>2. Bacteriophage DNA polymerases</b>				
a) PRD1 DNA polymerase*	553	63,336	6.68	(18,19)
b) $\phi$ 29 DNA polymerase*	575	66,714	8.83	(20)
c) M2 DNA polymerase*	572	66,423	7.69	(21)
d) T4 DNA polymerase	898	103,609	6.20	(22)
<b>3. Archaeobacterial DNA polymerases</b>				
a) <i>Thermococcus litoralis</i> DNA polymerase (Vent)	774	89,913	8.29	(23)
b) <i>Pyrococcus furiosus</i> DNA polymerase	775	90,112	7.92	(24)
c) <i>Sulfolobus solfataricus</i> DNA polymerase	882	101,332	9.72	(25)
<b>4. Eukaryotic Cell DNA polymerases</b>				
(1) DNA polymerase alpha				
a) Human DNA polymerase (alpha)	1,462	165,859	5.71	(26)
b) <i>S.cerevisiae</i> DNA polymerase I (alpha)	1,468	166,776	6.14	(27)
c) <i>S.pombe</i> DNA polymerase I (alpha)	1,405	159,348	6.85	(28)
d) <i>Drosophila melanogaster</i> DNA polymerase (alpha)	1,505	171,167	8.22	(29)
e) <i>Trypanosoma brucei</i> DNA polymerase (alpha)	1,339	151,611	6.39	(30)
(2) DNA polymerase delta				
a) Human DNA polymerase (delta)	1,107	123,634	6.94	(31,32)
b) Bovine DNA polymerase (delta)	1,106	123,707	7.52	(33)
c) <i>S.cerevisiae</i> DNA polymerase III (delta)	1,097	124,618	7.96	(34)
d) <i>S.pombe</i> DNA polymerase III (delta)	1,084	123,211	7.63	(35)
e) <i>Plasmodium falciparum</i> DNA polymerase (delta)	1,094	126,883	8.76	(36)
(3) DNA polymerase epsilon				
<i>S.cerevisiae</i> DNA polymerase II (epsilon)	2,222	255,669	6.92	(37)
(4) Other eukaryotic DNA polymerases				
<i>S.cerevisiae</i> DNA polymerase Rev3	1,504	172,956	8.86	(38)
<b>5. Viral DNA polymerases</b>				
a) Herpes Simplex virus type 1 DNA polymerase	1,235	136,547	7.35	(39)
b) Equine herpes virus type 1 DNA polymerase	1,220	135,955	6.67	(40)
c) Varicella-Zoster virus DNA polymerase	1,194	134,047	7.80	(41)
d) Epstein-Barr virus DNA polymerase	1,015	113,417	7.38	(42)
e) <i>Herpesvirus saimiri</i> DNA polymerase	1,009	113,934	7.31	(43)
f) Human cytomegalovirus DNA polymerase	1,242	137,101	7.25	(44)
g) Murine cytomegalovirus DNA polymerase	1,097	123,573	6.68	(45)
h) Human herpes virus type 6 DNA polymerase	1,012	115,819	7.11	(46)
i) Channel Catfish virus DNA polymerase	985	113,468	7.98	(47)
j) Chlorella virus DNA polymerase	913	104,955	6.66	(48)
k) Fowlpox virus DNA polymerase	988	116,658	8.11	(49)
l) Vaccinia virus DNA polymerase	937	108,564	7.50	(50)
m) Choristoneura biennis DNA polymerase	964	114,818	7.95	(51)
n) Autographa californica nuclear polyhedrosis virus (AcMNPV) DNA polymerase	984	114,337	8.35	(52)
o) Lymantria dispar nuclear polyhedrosis virus DNA polymerase	1,013	115,921	9.08	(53)

p) Adenovirus-2 DNA polymerase*	1,056	120,431	6.65	(54)
q) Adenovirus-7 DNA polymerase*	1,122	128,648	6.73	(55)
r) Adenovirus-12 DNA polymerase*	1,053	120,863	6.86	(56)
<b>6. Eukaryotic linear DNA plasmid encoded DNA polymerases</b>				
a) S-1 maize DNA polymerase*	917	105,935	8.62	(57)
b) <i>kalilo neurospora intermedia</i> DNA polymerase*	970	112,902	9.71	(58)
c) pAI2 <i>Ascobolus immersus</i> DNA polymerase*	1,202	138,279	10.10	(59)
d) pCLK1 <i>Claviceps purpurea</i> DNA polymerase*	1,097	126,627	8.76	(60)
e) <i>maranhar neurospora crassa</i> DNA polymerase*	1,021	119,074	9.62	(61)
f) pEM <i>Agaricus bisporus</i> DNA polymerase*	<sup>†</sup> 797	<sup>†</sup> 91,922	<sup>†</sup> 8.24	(62)
g) pGKL1 <i>Kluyveromyces lactis</i> DNA polymerase*	995	116,345	8.04	(63)
h) pGKL2 <i>Kluyveromyces lactis</i> DNA polymerase*	994	117,560	8.33	(64)
i) pSKL <i>Saccharomyces kluyveri</i> DNA polymerase*	999	117,544	9.79	(65)
<b>C. Family C DNA polymerases</b>				
<b>1. Bacterial replicative DNA polymerases</b>				
a) <i>E.coli</i> DNA polymerase III $\alpha$ subunit	1160	129,903	5.04	(66)
b) <i>S. typhimurium</i> DNA polymerase III $\alpha$ subunit	1160	130,118	5.05	(67)
c) <i>Bacillus subtilis</i> DNA polymerase III	1437	162,648	5.23	(68)
<b>2. <i>E.coli dnaQ</i> (MutD)</b>				
<i>E.coli</i> DNA polymerase III $\epsilon$ subunit	243	27,099	5.68	(69,70)

4. Lopez, P., Martinez, S., Diaz, A., Espinosa, M. and Lacks, S. A. (1989) *J. Biol. Chem.* **264**, 4255–4263.
5. Lawyer, F. C., Stoffel, S., Saiki, R. K., Myambo, K., Drummond, R. and Gelfand, D. H. (1989) *J. Biol. Chem.* **264**, 6427–6437.
6. Akhmetzhanov, A.A. and Vakhitov, V.A. (1992) *Nucl. Acids Res.* **20**, 5839.
7. Leavitt, M. C. and Ito, J. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4465–4469.
8. Dunn, J. J. and Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477–535.
9. Scarlato, V. and Gargano, S. (1992) *Gene* **118**, 109–113.
10. Rádén, B. and Rutberg, L. (1984) *J. Virol.* **52**, 9–15.
11. Foury, F. (1989) *J. Biol. Chem.* **264**, 20552–20560.
12. Ito, J. and Braithwaite, D.K. (1990) *Nucl. Acids Res.* **18**, 6716.
13. Blanco, L., Bernad, A. and Salas, M. (1991) *Nucl. Acids Res.* **19**, 955.
14. Hahn, S. and Rüger, W. (1989) *Nucl. Acids Res.* **17**, 6729.
15. Hollingsworth, H.C. and Nossal, N.G. (1991) *J. Biol. Chem.* **266**, 1888–1897.
16. Kaliman, A.V., Krutilina, A.I., Kryukov, V.M. and Bayev, A.A. (1986) *FEBS Lett.* **195**, 61–64.
17. Iwasaki, H., Ishino, Y., Toh, H., Nakata, A. and Shinagawa, H. (1991) *Mol. Gen. Genet.* **226**, 24–33.
18. Jung, G., Leavitt, M. C., Hsieh, J.-C. and Ito, J. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8287–8291.
19. Savilahti, H. and Bamford D.H. (1987) *Gene* **57**, 121–130.
20. Yoshikawa, H. and Ito, J. (1982) *Gene* **17**, 323–335.
21. Matsumoto, K., Takano, H., Kim, C. I. and Hirokawa, H. (1989) *Gene* **84**, 247–255.
22. 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.
23. Perler, F.B., Comb, D.G., Jack, W.E., Moran, L.S., Qiang, B., Kucera, R.B., Benner, J., Slatko, B.E., Nwankwo, D.O., Hempstead, S.K., Carlow, C.K.S. and Jannasch, H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5577–5581.
24. Cline, J.M., Scott, B.R., Adams, M.W.W., Sorge, J.A. and Mathur, E.J. *Gene* (in press).
25. Pisani, F.M., De Martino, C. and Rossi, M. (1992) *Nucl. Acids Res.* **20**, 2711–2716.
26. 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.
27. Pizzagalli, A., Valsasini, P., Plevani, P. and Lucchini, G. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 3772–3776.
28. Damagnez, V., Tillit, J., de Recondo, A.-M. and Baldacci, G. (1991) *Mol. Gen. Genet.* **226**, 182–189.
29. Hirose, F., Yamaguchi, M., Nishida, Y., Masutani, M., Miyazawa, H., Hanaoka, F. and Matsukage, A. (1991) *Nucl. Acids Res.* **19**, 4991–4998.
30. Leegwater, P.A.J., Strating, M., Murphy, N.B., Kooy, R.F., van der Vliet, P.C. and Overdule, J.P. (1991) *Nucl. Acids Res.* **19**, 6441–6447.
31. Chung, D.W., Zhang, J., Tan C.-K., Davie, E.W., So, A.G. and Downey, K.M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11197–11201.
32. Yang, C.-L., Chang, L.-S., Zhang, P., Hao, H., Zhu, L., Tommey, N.L. and Lee, M.Y.W.T. (1992) *Nucl. Acids Res.* **20**, 735–745.
33. Zhang, J., Chung, D.W., Tan, C.-K., Downey, K.M., Davie, E.W. and So, A.G. (1991) *Biochemistry* **30**, 11742–11750.
34. Morrison, A. and Sugino, A. (1992) *Nucl. Acids Res.* **20**, 375.
35. Pignède, G., Bouvier, D., de Recondo, A.-M. and Baldacci, G. (1991) *J. Mol. Biol.* **222**, 209–218.
36. Ridley, R.G., White, J.H., McAleese, S.M., Goman, M., Alano, P., de Vries, E. and Kilbey, B.J. (1991) *Nucl. Acids Res.* **19**, 6731–6736.
37. Morrison, A., Araki, H., Clark, A.B., Hamatake, R.K. and Sugino, A. (1990) *Cell* **62**, 1143–1151.
38. 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.
39. 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.
40. Telford, E.A., Watson, M.S., McBride, K. and Davison, A.J. (1992) *Virology* **189**, 304–316.
41. Davison, A. J. and Scott, J. E. (1986) *J. Gen. Virol.* **67**, 1759–1816.
42. 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.
43. Albrecht, J.-C. and Fleckenstein, B. (1990) *Virology* **174**, 533–542.
44. Kouzarides, T., Bankier, A. T., Satchwell, S. C., Weston, K., Tomlinson, P. and Barrell, B. G. (1987) *J. Virol.* **61**, 125–133.
45. Elliott, R., Clark, C. Jaquish, D. and Spector, D.H. (1991) *Virology* **185**, 169–186.
46. Teo, I.A., Griffin, B.E. and Jones, M.D. (1991) *J. Virol.* **65**, 4670–4680.
47. Davison, A.J. (1992) *Virology* **186**, 9–14.
48. Grabherr, R., Strasser, P. and Van Etten, J.L. (1992) *Virology* **188**, 721–731.
49. Binns, M. M., Stenzler, L., Tomley, F. M., Campbell, J. and Bournsnel, M. E. G. (1987) *Nucl. Acids Res.* **15**, 6563–6573.
50. Earl, P. L., Jones, E. V. and Moss, B. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3659–3663.
51. Mustafa, A. and Yuen, L. (1991) *DNA Seq.* **2**, 39–45.
52. Tomalski, M. D., Wu, J. and Miller, L. K. (1988) *Virology* **167**, 591–600.
53. Bjornson, R.M. and Rohrmann, G.F. (1992) *J. Gen. Virol.* **73**, 1499–1504.
54. Gingeras, T. R., Sciaky, D., Gelinis, 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.
55. Engler, J.A., Hoppe, M.S. and van Bree, M.P. (1983) *Gene* **21**, 145–159.
56. Shu, L., Hong, J.S., Wei, Y.-f. and Engler, J.A. (1986) *Gene* **46**, 187–195.
57. Paillard, M., Sederoff, R. R. and Levings, C. S. III (1985) *EMBO J.* **4**, 1125–1128.
58. Chan, B.S.-S., Court, D.A., Vierula, P.J. and Bertrand, H. (1991) *Curr. Genet.* **20**, 225–237.

59. Kempken, F., Meinhardt, F. and Esser, K. (1989) *Mol. Gen. Genet.* **218**, 523–530.
60. Oeser, B. and Tudzynski, P. (1989) *Mol. Gen. Genet.* **217**, 132–140.
61. Court D.A. and Bertrand, H. (1992) *Curr. Genet.* **22**, 385–397.
62. Robison, M.M., Royer, J.C. and Horgen, P.A. (1991) *Curr. Genet.* **19**, 495–502.
63. Stark, M.J.R., Mileham, A.J., Romanos, M.A. and Boyd, A. (1984) *Nucl. Acids Res.* **12**, 6011–6030.
64. Tommasino, M., Ricci, S. and Galeotti, C. L. (1988) *Nucl. Acids Res.* **16**, 5863–5878.
65. Hishinuma, F. and Hirai, K. (1991) *J. Gen. Genet.* **226**, 97–106.
66. Tomasiewicz, H. G. and McHenry, C. S. (1987) *J. Bacteriol.* **169**, 5735–5744.
67. Lancy, E. D., Lifshits, M. R., Munson, P. and Maurer, R. (1989) *J. Bacteriol.* **171**, 5581–5586.
68. Hammond, R. A., Barnes, M. H., Mack, S. L., Mitchener, J. A. and Brown, N. C. (1991) *Gene* **98**, 29–36.
69. Maki, H., Horiuchi, T. and Sekiguchi, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 7137–7141.
70. Joyce, C.M., Ollis, D.L., Rush, J., Steitz, T.A., Konigsberg, W.H. and Grindley, N.D.F. (1986) in 'Protein structure, folding and design (UCLA Symposia on Molecular and Cellular Biology, Vol 32)', D. Oxender, Ed., pp 197–205, Alan R. Liss.
71. Felsenstein, J. (1989) *Cladistics* **5**, 164–166.