Supplemental material for "Evolutionary conservation of residues in vertebrate POLN conferring low fidelity and bypass activity"

Kei-ichi Takata, Mercedes E. Arana, Mineaki Seki, Thomas A. Kunkel, and Richard D. Wood

Table S1. Base substitution errors generated by POLN and POLN derivatives

			WT			K679T			K679A		
Base	Mutation From → To	Mispair Template ·dNMP	Total # of template nucleotides[a]	Observed	Error rate (10^{-3})[b]	Total # of template nucleotides	Observed	Error rate (10^{-3})	Total # of template nucleotides	Observed	Error rate (10^{-3})
A	A → G	A · dCMP	99 (4653) [a]	1	0.2	99 (4455)	4	0.9	99 (4356)	2	0.5
	A → T	A · dAMP		0	≤0.2		0	≤0.2		0	≤0.2
	A → C	A · dGMP		0	≤0.2		0	≤0.2		0	≤0.2
T	T → C	T · dGMP	91 (4277)	8	1.9	91 (4095)	4	1.0	91 (4004)	2	0.5
	T → A	T · dTMP		0	≤0.2		0	≤0.2		0	≤0.25
	T → G	T · dCMP		0	≤0.2		0	≤0.2		0	≤0.25
G	G → A	G · dTMP	95 (4465)	132	30.0	95 (4275)	25	6.0	95 (4180)	46	11.0
	G → C	G · dGMP		5	1.1		3	0.5		4	1.0
	G → T	G · dAMP		1	0.2		0	≤0.2		1	0.2
C	C → T	C · dAMP	122 (5734)	0	≤0.2	122 (5490)	1	0.2	122 (5368)	4	0.7
	C → G	C · dCMP		0	≤0.2		0	≤0.2		0	≤0.2
	C → A	C · dTMP		0	≤0.2		0	≤0.2		0	≤0.2

[a] Listed are the numbers of template A, T, G and C nucleotides in the 407 nucleotide gap. The numbers in parentheses are that number times the number of *lacZ* clones sequenced, which were 47, 44 and 45, respectively, for WT, K679T and K679A.

[b] Error rates are the number of observed mutations of a particular type divided by the total number of nucleotides sequenced, a calculation used previously for particularly inaccurate polymerases (T. Matsuda, K. Bebenek, C. Masutani, I.B. Rogozin, F. Hanaoka and T.A. Kunkel Error rate and specificity of human and murine DNA polymerase eta, J Mol Biol 312 (2001) 335-346).

Supplemental Figure legends

Supplemental Figure 1. Sequence alignment of DNA polymerase domain of POLN homologs and prokaryotic A-family DNA polymerases. Numbers (1-6) show DNA polymerase motifs, alphabets (a-c) show insertions in POLN, 2 open arrow heads show the residues (K679 and Y682 of POLN) substituted in this study. Three closed arrowheads show residues that are important for strand displacement activity in *E. coli* pol I (S769, F771, and R841). The alignment was created using ClustalX. Similarity groups for colored residues are: {K, R, H}, {D, E}, {I, L, V, M}, {F, Y, W}, {Q, N}, {G, A, S, T, P, C}. Degrees of conservation scored using the Gonnet Pam250 matrix in ClustalX are shown by *, double dots, or single dots. Hs, *Homo sapiens*; Pt, *Pan troglodytes*; Ec, *Equus caballus*; Cf, *Canis familiaris*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Dr, *Danio rerio*; Ec, *Escherichia coli*; Pa, *Pseudomonas aeruginosa*; Re, *Rhodococcus erythropolis*; Taqpol, *Thermus aquaticus* DNA polymerase; BspolI, *Bacillus subtilis*, RhpolI, *Rickettsia helvetica*. PtPOLN, EcPOLN, CfPOLN, and RnPOLN are predicted from their genomic DNA sequences.

Supplemental Figure 2. Mutation error spectrum by POLN and derivatives. The 407 template nucleotides within the single-stranded gap region of the M13mp2 DNA substrate are shown as five lines of sequence. Nucleotide +1 is the first transcribed nucleotide of the *LacZ* gene. Base substitutions are represented as letters above the sequence. Single base deletions are represented by an open triangle below the sequence. Single base additions are depicted as a filled inverted triangle above the sequence. *Red* characters represent phenotypically detectable changes, and *gray* characters represent phenotypically silent 'hitchhikers' found in association with detectable changes. The Y682F POLN had weak processivity (Fig. 6) and could not fill the gap in M13mp2.

Supplemental Figure 3. Extension by POLN derivatives after a nucleotide insertion opposite a 5*S*-Tg. Increasing amounts of delP, WT, K679R, K679T, K679A and K679Q (6, 12 and 23 nM), Y682F (6, 12, 23, 29, 58 and 115 nM), POLQ (3, 6 and 12 nM), and RB69 gp43 (2.5, 5 and 10 pM) were incubated with the 5'-^{32}P-labeled primer-templates indicated beside the panel in the presence of four all nucleotides at 37°C for 10 min in the each reaction mixtures. The first lane contained no enzyme. Panels A, B, C, and D show extension after insertion of A, C, G, or T opposite 5*S*-Tg. The percentage (%) extension of the primer is shown below each lane.
Only POLQ could extend after incorrect nucleotide incorporations opposite the lesion (panel D). The short products observed in panel B were not bypass products. They might be products produced by slippage of the 3' end sequence of the primer, as "GATGC" can perfectly anneal to the template sequence "CTACG" located 2 bases towards the 5' end from 5*S*-Tg.

Supplemental Figure 4. Extension from T:G mismatch by POLN derivatives. Increasing amounts of delP, WT, K679R, K679T, K679A and K679Q (6, 12 and 23 nM), Y682F (6, 12, 23, 29, 58 and 115 nM), and RB69 gp43 (2.5, 5 and 10 pM) were incubated with the 5'-^{32}P-labeled primer-templates indicated beside the panel in the presence of four all nucleotides at 37°C for 10 min. The first lane contained no enzyme. The percentage (%) extension of the primer is shown below each lane.

Supplemental Figure 5. 5*S*-Tg bypass from 1 or 2 base behind the lesion by WT, K679T, K679A, and RB69 gp43. (A) DNA synthesis on a DNA template containing an undamaged thymine from the 14-mer primer (lanes 1-16) or the 15-mer primer (lanes 17-32). (B) DNA synthesis on a DNA template containing a 5*S*-Tg from the 14-mer primer (lanes 1-16) or the 15-mer primer (lanes 17-32). All reaction mixtures contained 100 nM substrate in the presence of all four nucleotides. Incubation time of each reaction was shown in bottom. Locations of unreacted end-labeled primer (N_0), each template base position (from N_1 to N_{16}), full-length product (N_{16}

for the 14-mer primer, N_{15} for the 15-mer primer), positions of 5*S*-Tg are shown as Tg. (C) Termination probabilities at positions along the 30-mer template containing an undamaged T or a 5*S*-Tg during the DNA synthesis from the 14-mer primer by WT (white), K679T (light gray), K679A (dark gray), and RB69 gp43 (black). Termination probabilities were defined as described in the materials and methods. Values are averages from 2 data points at reaction intervals from 2, 4 and 6 min with error bars representing standard deviations. Bypass, insertion, and extension probabilities and bypass efficiency were defined as described in the materials and methods. (D) As described for C, using the 30-mer template and the 15 primer.

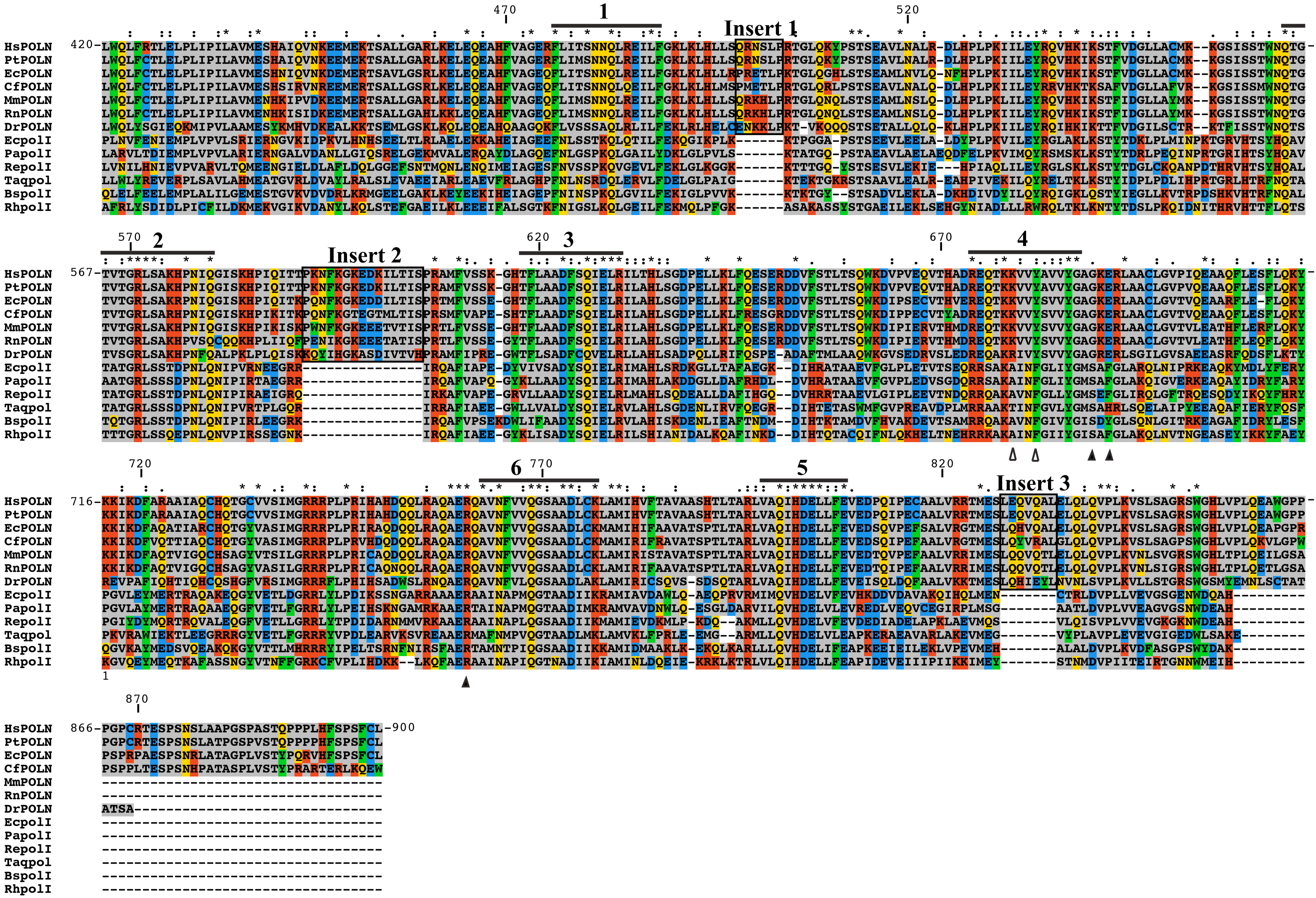
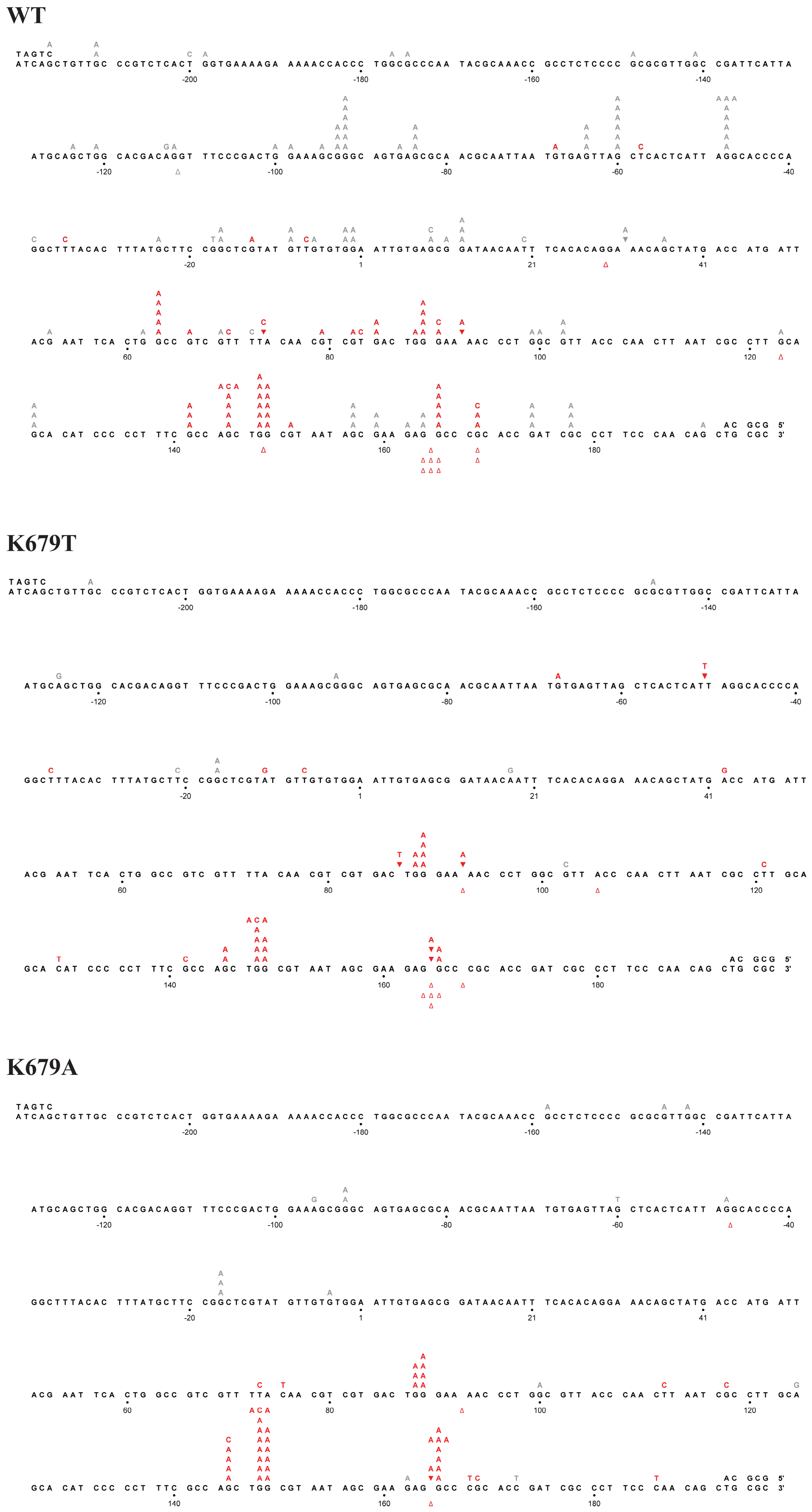

HsPOLN
PtPOLN
EcPOLN
CfPOLN
MmPOLN
RnPOLN
DrPOLN
EcpolI
PapolI
RepolI
Taqpol
BspolI
RhpolI
420
566
470
520
1
Insert 1
567
715
570
2
Insert 2
620
3
670
4
716
865
720
6
770
5
820
Insert 3
866
900
870

Fig. S1



WT

K679T

K679A

Fig. S2

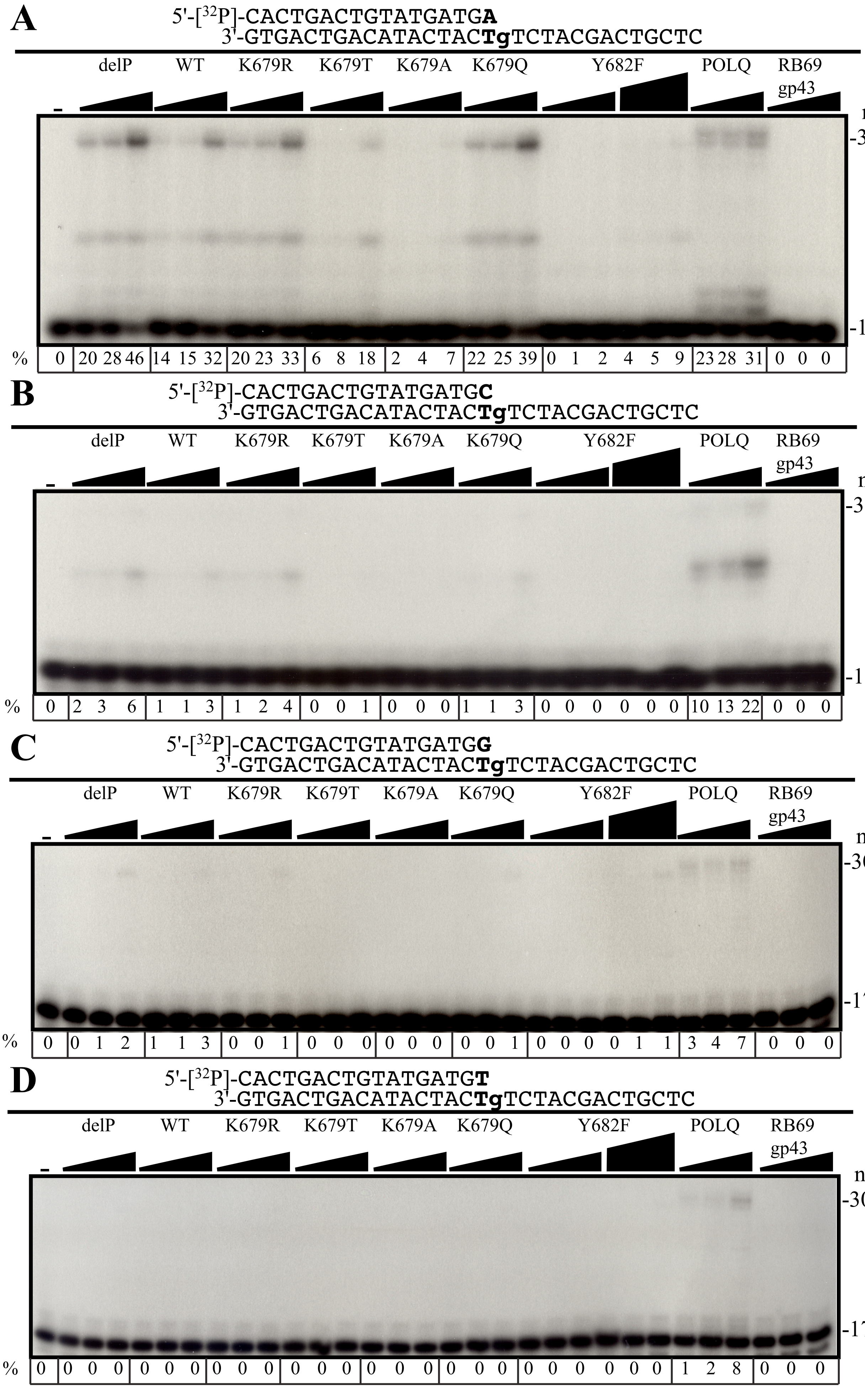

A
5'-[^{32}P]-CACTGACTGTATGATGA
3'-GTGACTGACATACTACTgTCTACGACTGCTC
delP
WT
K679R
K679T
K679A
K679Q
Y682F
POLQ
RB69 gp43
–
nt
-30
-17
%
0 | 20 28 46 | 14 15 32 | 20 23 33 | 6 8 18 | 2 4 7 | 22 25 39 | 0 1 2 4 5 9 | 23 28 31 | 0 0 0
B
5'-[^{32}P]-CACTGACTGTATGATGC
3'-GTGACTGACATACTACTgTCTACGACTGCTC
delP
WT
K679R
K679T
K679A
K679Q
Y682F
POLQ
RB69 gp43
–
nt
-30
-17
%
0 | 2 3 6 | 1 1 3 | 1 2 4 | 0 0 1 | 0 0 0 | 1 1 3 | 0 0 0 0 0 0 | 10 13 22 | 0 0 0
C
5'-[^{32}P]-CACTGACTGTATGATGG
3'-GTGACTGACATACTACTgTCTACGACTGCTC
delP
WT
K679R
K679T
K679A
K679Q
Y682F
POLQ
RB69 gp43
–
nt
-30
-17
%
0 | 0 1 2 | 1 1 3 | 0 0 1 | 0 0 0 | 0 0 0 | 0 0 1 | 0 0 0 0 1 1 | 3 4 7 | 0 0 0
D
5'-[^{32}P]-CACTGACTGTATGATGT
3'-GTGACTGACATACTACTgTCTACGACTGCTC
delP
WT
K679R
K679T
K679A
K679Q
Y682F
POLQ
RB69 gp43
–
nt
-30
-17
%
0 | 0 0 0 | 0 0 0 | 0 0 0 | 0 0 0 | 0 0 0 | 0 0 0 | 0 0 0 0 0 0 | 1 2 8 | 0 0 0

Fig. S3

5'-[^{32}P]-CACTGACTGTATGATG**T**
3'-GTGACTGACATACTAC**G**TCTACGACTGCTC

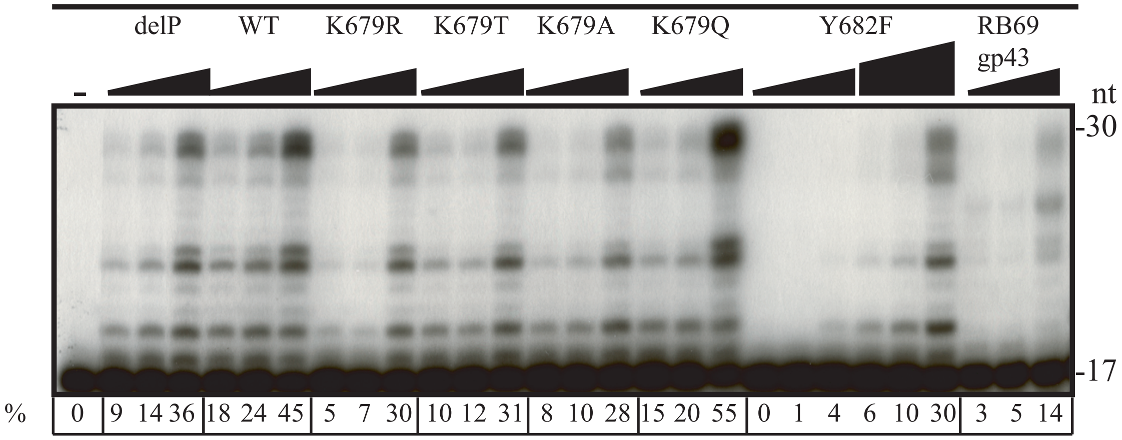


Fig. S4

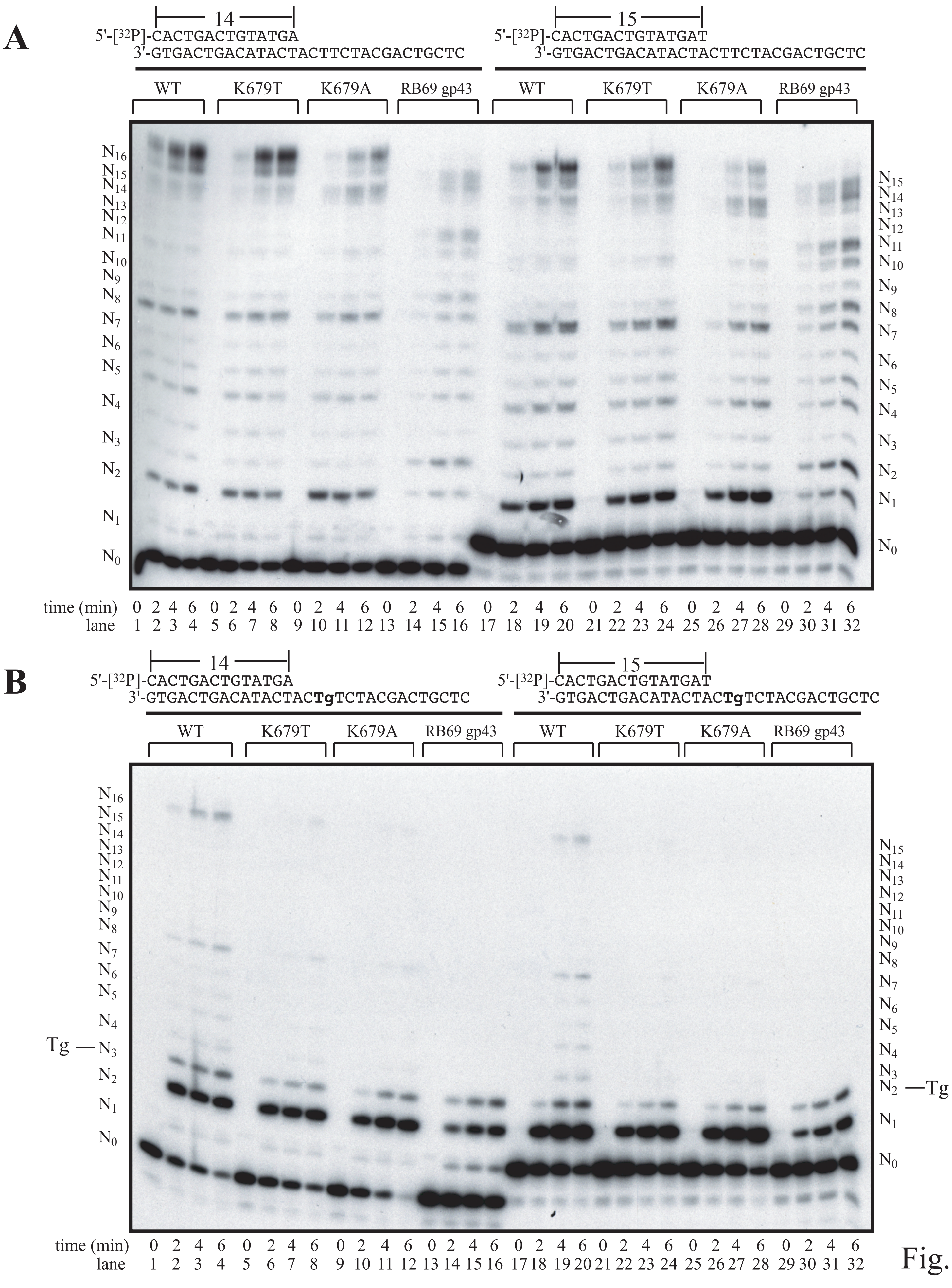
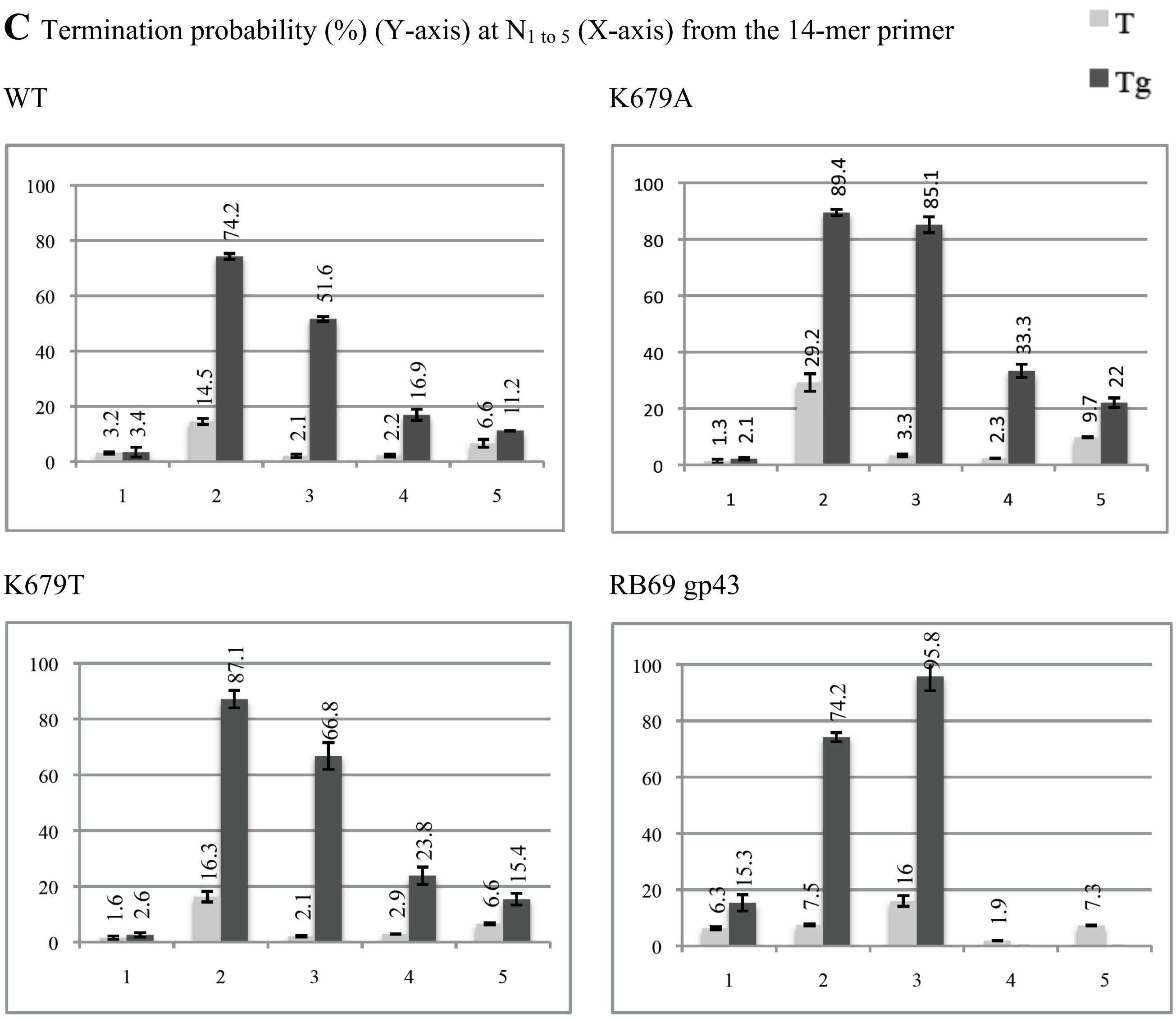

A
14
5'-[32P]-CACTGACTGTATGA
3'-GTGACTGACATACTACTTCTACGACTGCTC
15
5'-[32P]-CACTGACTGTATGAT
3'-GTGACTGACATACTACTTCTACGACTGCTC
WT
K679T
K679A
RB69 gp43
WT
K679T
K679A
RB69 gp43
N16 N15 N14 N13 N12 N11 N10 N9 N8 N7 N6 N5 N4 N3 N2 N1 N0
N15 N14 N13 N12 N11 N10 N9 N8 N7 N6 N5 N4 N3 N2 N1 N0
time (min) 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6
lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
B
14
5'-[32P]-CACTGACTGTATGA
3'-GTGACTGACATACTACTgTCTACGACTGCTC
15
5'-[32P]-CACTGACTGTATGAT
3'-GTGACTGACATACTACTgTCTACGACTGCTC
WT
K679T
K679A
RB69 gp43
WT
K679T
K679A
RB69 gp43
N16 N15 N14 N13 N12 N11 N10 N9 N8 N7 N6 N5 N4 N3 N2 N1 N0
Tg — N3
N15 N14 N13 N12 N11 N10 N9 N8 N7 N6 N5 N4 N3 N2 N1 N0
N2 — Tg
time (min) 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6 0 2 4 6
lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32

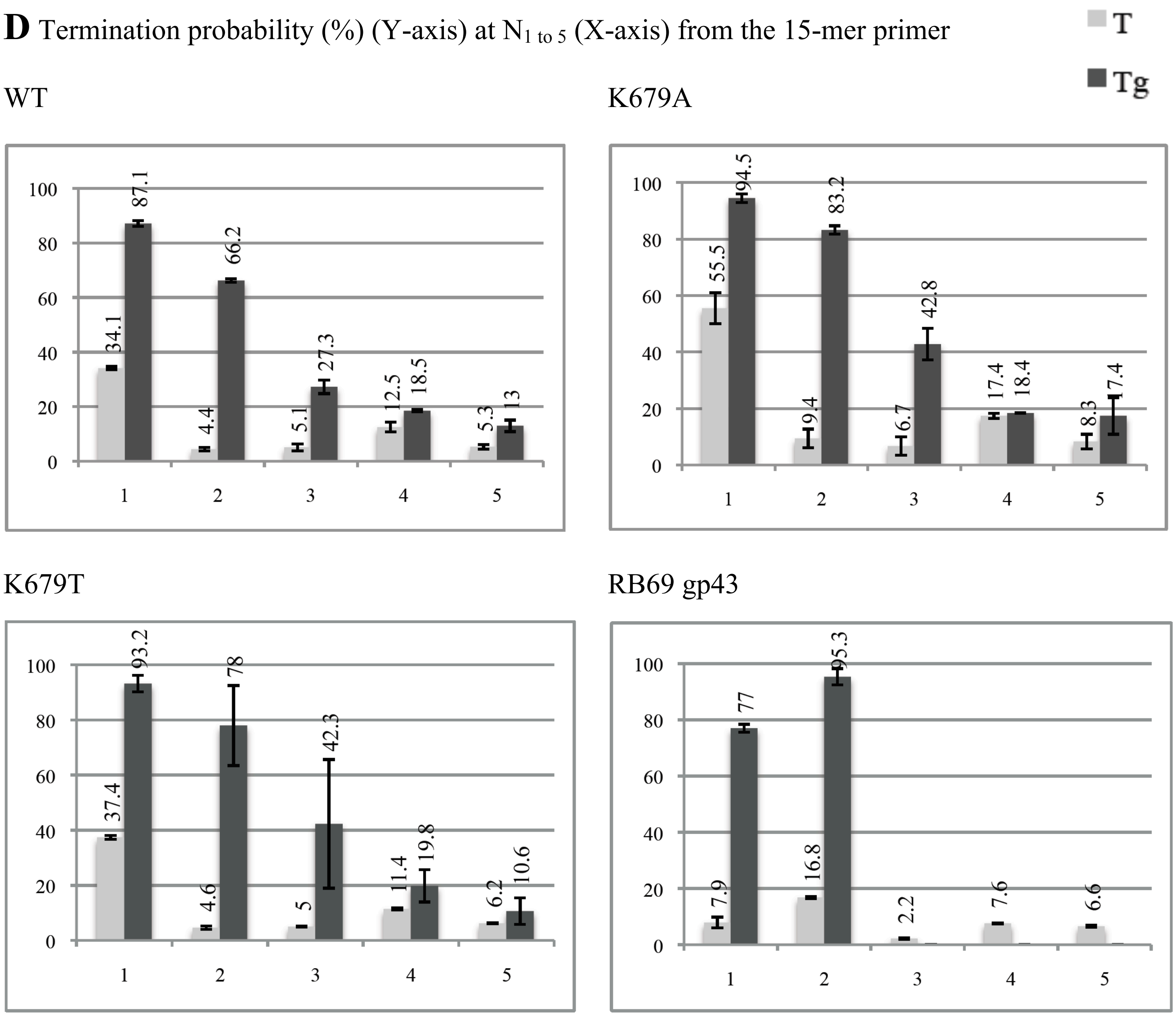
Fig. S5

C Termination probability (%) (Y-axis) at $N_{1\ \mathrm{to}\ 5}$ (X-axis) from the 14-mer primer



Template	Enzyme	Bypass probability	Insertion probability	Extension probability	Bypass efficiency
Undamaged T	WT	81.0±2.6	85.6±1.5	97.8±0.7	-
	K679T	84.4±0.7	86.8±1.4	98.3±0.2	-
	K679A	71.0±3.1	73.7±3.1	97.4±0.5	-
	RB69 gp43	74.7±1.1	93.1±0.1	85.5±0.7	-
5*S*-Tg	WT	12.8±1.0	26.5±0.9	50.1±1.0	15.8±1.0
	K679T	4.6±1.7	13.3±3.1	34.9±5.2	5.4±1.9
	K679A	2.2±0.2	11.8±1.2	18.8±0.5	3.1±0.1
	RB69 gp43	1.0±0.9	26.1±1.6	4.2±4.8	1.3±1.5

D Termination probability (%) (Y-axis) at $N_{1\text{ to }5}$ (X-axis) from the 15-mer primer


WT

K679A

K679T

RB69 gp43

Template	Enzyme	Bypass probability	Insertion probability	Extension probability	Bypass efficiency
Undamaged T	WT	65.5±0.7	68.3±0.2	96.0±0.7	-
	K679T	62.0±0.1	64.7±0.4	95.9±0.4	-
	K679A	43.2±1.2	46.8±2.6	92.4±2.4	-
	RB69 gp43	78.0±0.8	92.6±1.2	84.1±0.2	-
5*S*-Tg	WT	5.1±0.1	14.3±1.0	35.7±2.3	7.8±0.1
	K679T	1.8±1.4	7.1±2.7	23.1±11.4	2.9±0.6
	K679A	1.0±0.3	5.8±1.5	16.9±0.5	2.3±0.6
	RB69 gp43	1.1±0.8	23.5±1.7	4.7±3.2	1.5±1.1

Fig. S5