Cell Reports, Volume *21*

Supplemental Information

Primer-Independent DNA Synthesis by a Family B DNA Polymerase from Self-Replicating Mobile Genetic Elements

Modesto Redrejo-Rodríguez, Carlos D. Ordóñez, Mónica Berjón-Otero, Juan Moreno-González, Cristian Aparicio-Maldonado, Patrick Forterre, Margarita Salas, and Mart Krupovic

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Bioinformatic analyses

Phylogenetic analysis. The non-redundant database of protein sequences at the NCBI was searched using the PSI-BLAST (Altschul et al., 1997). For phylogenetic analyses protein sequences were aligned with Promals3D (Pei and Grishin, 2014). Poorly aligned (low information content) positions were removed using the Gappyout function of Trimal (Capella-Gutierrez et al., 2009). The dataset of viral, plasmid and polintons pDNAP sequences was collected previously (Krupovic and Koonin, 2015). Maximum likelihood phylogenetic tree was constructed using the PhyML program (Guindon et al., 2010) the latest version of which (http://www.atgc-montpellier.fr/phyml-sms/) includes automatic selection of the best-fit substitution model for a given alignment. The best model identified by PhyML was LG +G6 +I +F (LG, Le-Gascuel matrix; G6, Gamma shape parameter: fixed, number of categories: 6; I, proportion of invariable sites: fixed; F, equilibrium frequencies: empirical).

Identification and annotation of integrated MGE. The pipolins' were identified by thorough analysis of genomic neighborhoods of the piPolB-encoding genes. The precise boarders of integration were defined based on the presence of direct repeats corresponding to attachment sites. The repeats were searched for using Unipro UGENE (Okonechnikov et al., 2012). Pipolin genes were annotated based on the PSI-BLAST searches (Altschul et al., 1997) against the non-redundant protein database at NCBI and HHpred searches (Soding et al., 2005). Pipolins were compared to each other and visualized using EasyFig (Sullivan et al., 2011).

Protein expression and purification

Primer-independent DNA polymerase (piPolB) from *E. coli* 3-373-03_S1_C2 Pipolin (NCBI GI:693097161) was obtained from GeneScript into NdeI-XhoI sites of pET23a (see Table S4). A stop codon was included to obtain the untagged recombinant protein. Polymerase (D368A) and exonuclease (D59A/E61A) deficient proteins, as well as wild type, K614A, H615A, K624A and R626A his-tagged variants, were obtained by site directed mutagenesis (Table S4).

All piPolB variants were expressed in Bl21(DE3) *E. coli* cells, using ZYM-5052 autoinduction medium (Studier, 2005) in the presence of 100 mg/L ampicillin. Cultures were grown for 20 h at 28 ºC. For purification of untagged piPolB variants, cells were disrupted by grinding with alumina and suspended in buffer A (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 7 mM β-mercaptoethanol, 5% (v/v) glycerol) containing 1 M NaCl. Alumina and cell debris were removed by centrifugation, and absorbance at 260 nm was adjusted to 120 units/ml prior to DNA precipitation with 0.3% (w/v) polyethyleneimine. After centrifugation at 20,000 x g for 20 min, ammonium sulfate was added to the supernatant to 69% saturation and centrifuged at 20,000 x g for 30 min. The piPolB (wild type and mutants) containing pellet was resuspended in buffer A and applied to serial Q Sepharose® fast flow (GE Healthcare) and phosphocellulose (P11, Whatman) columns, at an ionic strength about 0.2 M NaCl. After extensive wash with increasing concentrations of NaCl in buffer A, purified DNA polymerase was eluted with 0.35 M NaCl and applied to Heparin-Sepharose® CL-6B column (GE Healthcare), where, after washing with 0.35, 0.4 and 0.45 M NaCl, they were eluted at 1 M NaCl in buffer A.

Histidine-tagged variants were purified by standard method (Spriestersbach et al., 2015). Briefly, cells were resuspended in buffer C (50 mM phosphate buffer, pH 8, 7 mM β-mercaptoethanol, 5% (v/v) glycerol, 1 M NaCl, 5 mM imidazole) and incubated for 30 min at room temperature with 1 mg/mL lysozyme (Sigma) and 1 unit of benzonase (Sigma), prior to cell disruption by sonication. After centrifugation at 20,000 x g for 30 min, the soluble fraction was applied to a Ni-NTA column (Qiagen). After extensive wash with 5, 10, 25 and 50 mM imidazole, the protein was eluted with 200 mM imidazole and subsequently applied to Heparin-Sepharose® CL-6B column (GE Healthcare), where, after washing with 0.35, 0.4 and 0.45 M NaCl, it was eluted at 1 M NaCl in buffer A.

In all cases, pooled fractions containing pure piPolB variants were dyalized overnight against 500 volumes of buffer B (50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 7 mM β-mercaptoethanol, 0.25 M NaCl and 50% (v/v) glycerol) and kept at -20ºC, or at -70ºC for long storage. Final purity of the proteins was estimated to be >90 % by SDS–PAGE followed by Coomassie blue staining.

DNA substrates

Oligonucleotides (Table S3) were purchased from Sigma or IDT in PAGE purification grade. To form a primer/template substrate as indicated in the top of each figure, the P15 oligonucleotide (Table S3) was 5´-labeled with [γ-^{32}P]ATP using T4 Polynucleotide Kinase and hybridized to 1.2-fold molar excess of

complementary unlabeled template oligonucleotides (T33GTA, T33GTT or T33GFA, Table S3) in the presence of 50 mM NaCl and 50 mM Tris-HCl, pH 7.5.
Genomic M13mp18 single-stranded circular DNA (laboratory stock) was diluted up to 50 ng/μL in a buffer containing 0.2 M NaCl and 60 mM Tris-HCl, pH 7.5 with or without M13 UP primer (Table S3), heated for 5 min at 65 ºC and cooled slowly overnight to allow the annealing of the primer. Primed and non-primed M13 substrates were stored at -20 ºC in small aliquots to minimize DNA nicking due to repetitive cycles of freeze-thaw.

Replication of homopolymeric single stranded oligonucleotides
Assays were made in 20 μL final volume containing 50 mM Tris-HCl, pH 7.5, 1 mM DTT, 4% (v/v) glycerol, 0.1 mg/ml BSA, 0.05% (v/v) Tween 20, 1 μM oligonucleotide template, 100 μM dNTPs, 500 nM of the indicated piPolB variant and 0.5 μCi [α-^{32}P]dATP. After incubation for the indicated times at 30 ºC, the reactions were stopped by adding 10 μL of formamide loading buffer (98% formamide, 20 mM EDTA, 0.5% (w/v) bromophenol blue, and 0.5% (w/v) xylene cyanol). Samples were analyzed by 8 M urea-20% polyacrylamide gel electrophoresis (20x30x0.5 mm) in 1X TBE buffer. Gel bands were detected either by autoradiography or phosphorimages (Typhoon FLA 7000) and processed with ImageJ software.

De novo primer synthesis detection
To detect de novo primer synthesis we used [γ-^{32}P]ATP as the labeled nucleotide. M13 ssDNA (3.2 nM) was used as template. The reaction mixture contained, in a final volume of 25 μL, 50 mM Tris-HCl, pH 7.5, 1 mM DTT, 4% (v/v) glycerol, 0.1 mg/ml BSA, 0.05% (v/v) Tween 20, 100 μM dNTPs, 0.5 μCi [γ-^{32}P]ATP, the indicated template and DNA polymerase (500 nM). Reactions were triggered by addition of either 10 mM $MgCl_2$ or 1 mM $MnCl_2$ and incubated for indicated times at 30 ºC. Then, the reactions were stopped by adding 10 μL of formamide loading buffer (98% formamide, 20 mM EDTA, 0.5% (w/v) bromophenol blue, and 0.5% (w/v) xylene cyanol). Samples were analyzed by 8 M urea-20% polyacrylamide gel electrophoresis (20x30x0.5 mm) in 1X TBE buffer. When indicated, high resolution gels (40 cm long) were used. Gel bands were detected by phosphorimages (Typhoon FLA 7000) and processed with ImageJ software. The [γ^{32}P]ATP-(dGMP)n DNA ladder used as size marker, generated by human PrimPol (García-Gómez et al., 2013) was a gift from Dr. Luis Blanco (CBMSO, Madrid).

Survival of piPolB expressing bacteria upon DNA damaging agents challenges
Starter cultures of *E. coli* Bl21(DE3) harboring pET23a::piPolB or pET23a::piPolB(D368A) plasmids were inoculates in LB media in the presence of ampicillin (150 μg/mL) and glucose (40 mM) and grown overnight shaking at 37 ºC. Saturated cultures were diluted (1:100) in fresh LB media with ampicillin and grown 1-2 h at 28 ºC until DO600nm=0.4. Recombinant protein expression was then induced by 0.5 mM IPTG during one hour prior to genotoxic challenge. At this point an aliquot was withdrawn to verify recombinant protein expression by SDS-PAGE (not shown).
For MMC treatment, the indicated drug concentration was added directly into the cultures that were grown for an additional hour and then serial-diluted in fresh LB and plate onto LB-agar plates (without antibiotic). In the case of UV-exposure, the induced cultures were serial-diluted in sterile PBS and plated onto LB-agar plates prior to irradiation with the indicated UV-light intensity in a Spectrolinker™ XL-1000 (Spectronics Corporation).

Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Φ29DNAP	M.Salas laboratory stock	Lazaro et al. 1995
Experimental Models: Cell Lines		
E. coli Bl21(DE3)	Stratagene	C6000-03
Recombinant DNA		
pET23a(+)::piPolB	This paper (GeneScript)	See Table S4
pET23a(+)::piPolB derivatives	This paper	See Table S4
Sequence-Based Reagents		
M13mp18 ssDNA	M.Salas laboratory stock	Sambrok and Russel, 2001
Oligonucleotide substrates for DNA replication assays	Sigma/IDT	See Table S3
Software and Algorithms		
PROMALS3D	University of Texas Southwestern Medical Center	Pei et al.NAR 2008 and http://prodata.swmed.edu/promals3d/promals3d.php
Illustrator	Adobe	adobe.com/illustrator

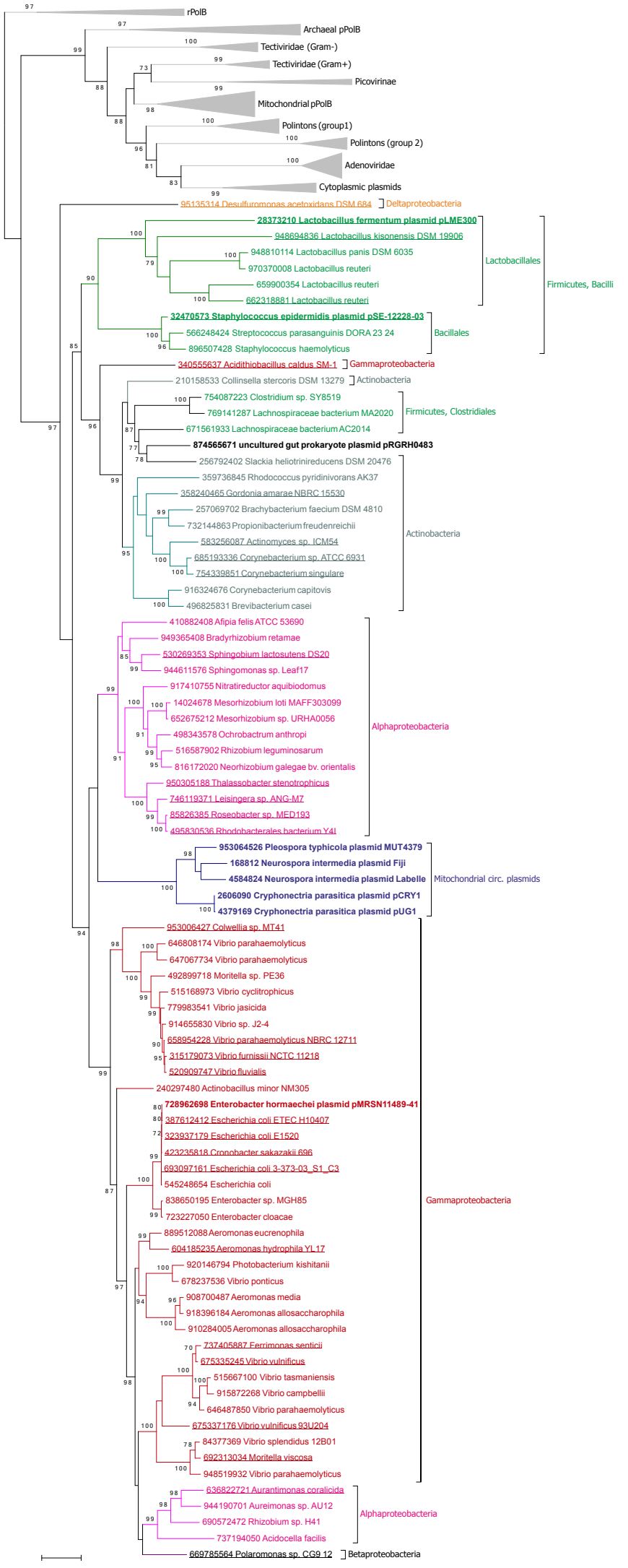
rPolB
Archaeal pPolB
Tectiviridae (Gram-)
Tectiviridae (Gram+)
Picovirinae
Mitochondrial pPolB
Polintons (group1)
Polintons (group 2)
Adenoviridae
Cytoplasmic plasmids
95135314 Desulfuromonas acetoxidans DSM 684
Deltaproteobacteria
28373210 Lactobacillus fermentum plasmid pLME300
948694836 Lactobacillus kisonensis DSM 19906
948810114 Lactobacillus panis DSM 6035
970370008 Lactobacillus reuteri
659900354 Lactobacillus reuteri
662318881 Lactobacillus reuteri
Lactobacillales
32470573 Staphylococcus epidermidis plasmid pSE-12228-03
566248424 Streptococcus parasanguinis DORA 23 24
896507428 Staphylococcus haemolyticus
Bacillales
Firmicutes, Bacilli
340555637 Acidithiobacillus caldus SM-1
Gammaproteobacteria
210158533 Collinsella stercoris DSM 13279
Actinobacteria
754087223 Clostridium sp. SY8519
769141287 Lachnospiraceae bacterium MA2020
671561933 Lachnospiraceae bacterium AC2014
Firmicutes, Clostridiales
874565671 uncultured gut prokaryote plasmid pRGRH0483
256792402 Slackia heliotrinireducens DSM 20476
359736845 Rhodococcus pyridinivorans AK37
358240465 Gordonia amarae NBRC 15530
257069702 Brachybacterium faecium DSM 4810
732144863 Propionibacterium freudenreichii
583256087 Actinomyces sp. ICM54
685193336 Corynebacterium sp. ATCC 6931
754339851 Corynebacterium singulare
916324676 Corynebacterium capitovis
496825831 Brevibacterium casei
Actinobacteria
410882408 Afipia felis ATCC 53690
949365408 Bradyrhizobium retamae
530269353 Sphingobium lactosutens DS20
944611576 Sphingomonas sp. Leaf17
917410755 Nitratireductor aquibiodomus
14024678 Mesorhizobium loti MAFF303099
652675212 Mesorhizobium sp. URHA0056
498343578 Ochrobactrum anthropi
516587902 Rhizobium leguminosarum
816172020 Neorhizobium galegae bv. orientalis
950305188 Thalassobacter stenotrophicus
746119371 Leisingera sp. ANG-M7
85826385 Roseobacter sp. MED193
495830536 Rhodobacterales bacterium Y4I
Alphaproteobacteria
953064526 Pleospora typhicola plasmid MUT4379
168812 Neurospora intermedia plasmid Fiji
4584824 Neurospora intermedia plasmid Labelle
2606090 Cryphonectria parasitica plasmid pCRY1
4379169 Cryphonectria parasitica plasmid pUG1
Mitochondrial circ. plasmids
953006427 Colwellia sp. MT41
646808174 Vibrio parahaemolyticus
647067734 Vibrio parahaemolyticus
492899718 Moritella sp. PE36
515168973 Vibrio cyclitrophicus
779983541 Vibrio jasicida
914655830 Vibrio sp. J2-4
658954228 Vibrio parahaemolyticus NBRC 12711
315179073 Vibrio furnissii NCTC 11218
520909747 Vibrio fluvialis
240297480 Actinobacillus minor NM305
728962698 Enterobacter hormaechei plasmid pMRSN11489-41
387612412 Escherichia coli ETEC H10407
323937179 Escherichia coli E1520
423235818 Cronobacter sakazakii 696
693097161 Escherichia coli 3-373-03_S1_C3
545248654 Escherichia coli
838650195 Enterobacter sp. MGH85
723227050 Enterobacter cloacae
889512088 Aeromonas eucrenophila
604185235 Aeromonas hydrophila YL17
920146794 Photobacterium kishitanii
678237536 Vibrio ponticus
908700487 Aeromonas media
918396184 Aeromonas allosaccharophila
910284005 Aeromonas allosaccharophila
737405887 Ferrimonas senticii
675335245 Vibrio vulnificus
515667100 Vibrio tasmaniensis
915872268 Vibrio campbellii
646487850 Vibrio parahaemolyticus
675337176 Vibrio vulnificus 93U204
84377369 Vibrio splendidus 12B01
692313034 Moritella viscosa
948519932 Vibrio parahaemolyticus
Gammaproteobacteria
636822721 Aurantimonas coralicida
944190701 Aureimonas sp. AU12
690572472 Rhizobium sp. H41
737194050 Acidocella facilis
Alphaproteobacteria
669785564 Polaromonas sp. CG9 12
Betaproteobacteria

Figure S1

Figure S1. Maximum likelihood phylogeny of pipolin-related PolBs. Related to Figure 1.
Clades that are only distantly related to the pipolin-encoded proteins were collapsed. The tree is rooted with cellular and viral RNA-primed PolBs. Branches are colored according to the classification of the corresponding taxa. Pipolins for which precise integration sites were identified (Table S1) are underlined. Numbers at the branch points represent the Bayesian-like transformation of aLRT (aBayes) local support values. The scale bar represents the number of substitutions per site.

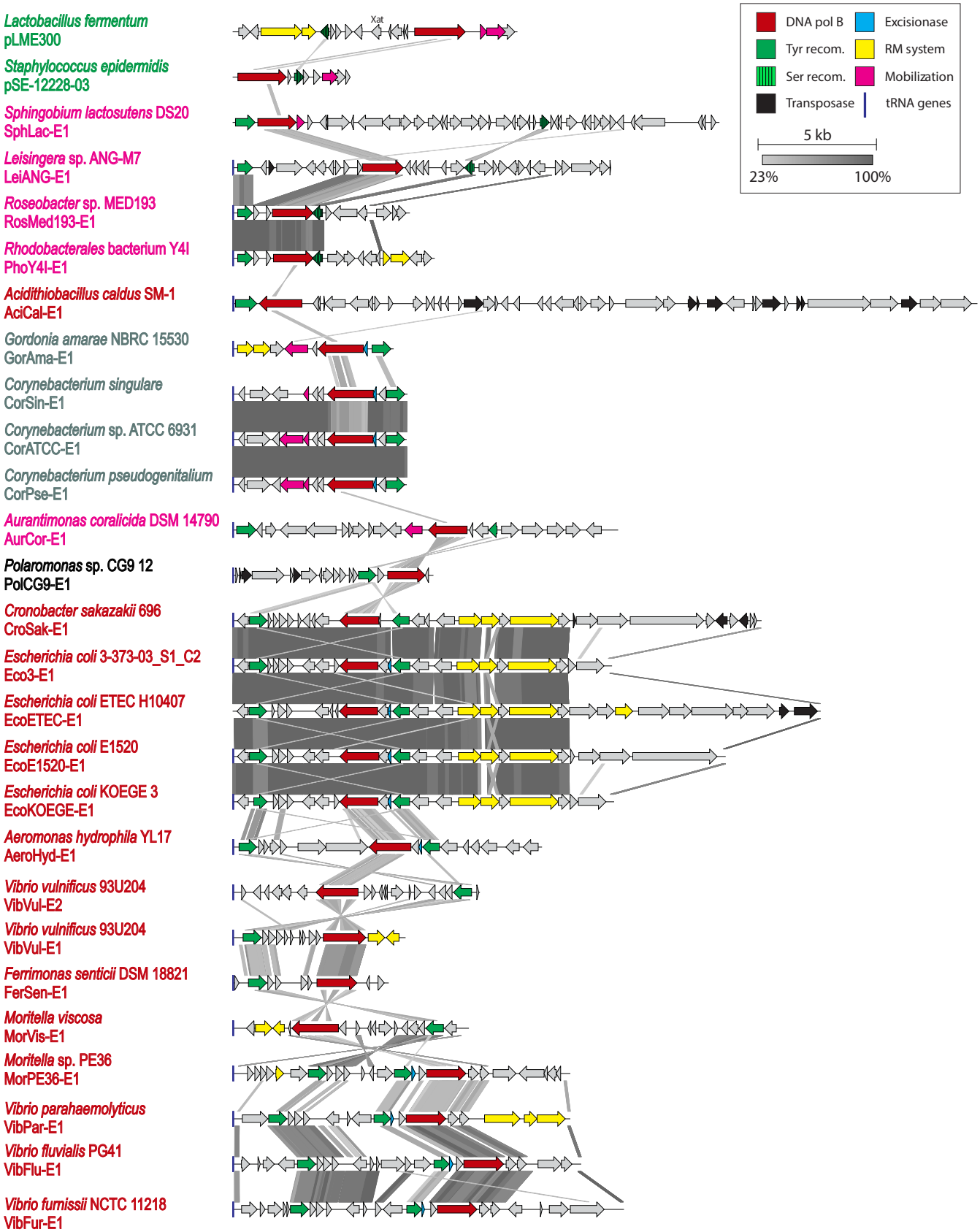

Figure S2. Comparative genomic analysis of bacterial Pipolins. Related to Figure 1.
Predicted protein-coding genes are indicated with arrows, indicating the direction of transcription. 'Xat' stands for xenobiotic acyltransferase. The color key for the designation of the common genes and the greyscale for homology degree are shown in the top right area of the figure. Names of pipolin carrying bacteria are colored as in Figure S1.

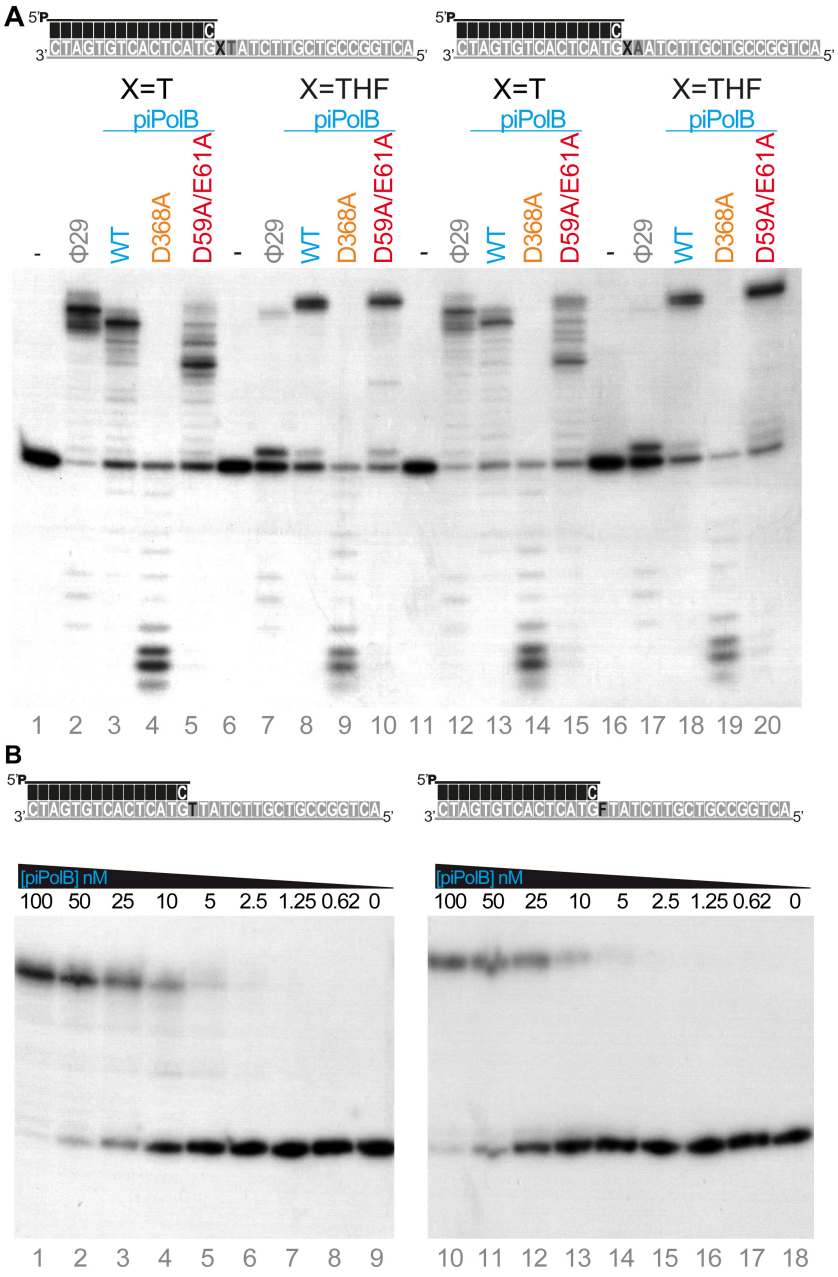

Figure S3. Effect of sequence context and enzyme concentration on TLS capacity by piPolB. Related to Figures 2 and 3.

(A) Denaturing PAGE analysis of primer extension products by Φ29 DNA polymerase (Φ29) and piPolB. As indicated, wild type, D368A (polymerase deficient) and D59A/E61A (exonuclease deficient) piPolB variants were tested using two different sequence contexts (lanes 1-10 vs. 11-20), in the absence (lanes 1-5 and 11-15) or presence (lanes 6-10 and 16-20) of a THF abasic site analog. Schematic representations of each template/primer substrate are depicted above. (B) Processive replication of primer/template substrates by decreasing concentrations of wild type piPolB. Reactions were performed in the presence of either 1 μM (lanes 1-9) or 100 μM (lanes 10-18) dNTPs for undamaged or damaged templates, respectively.

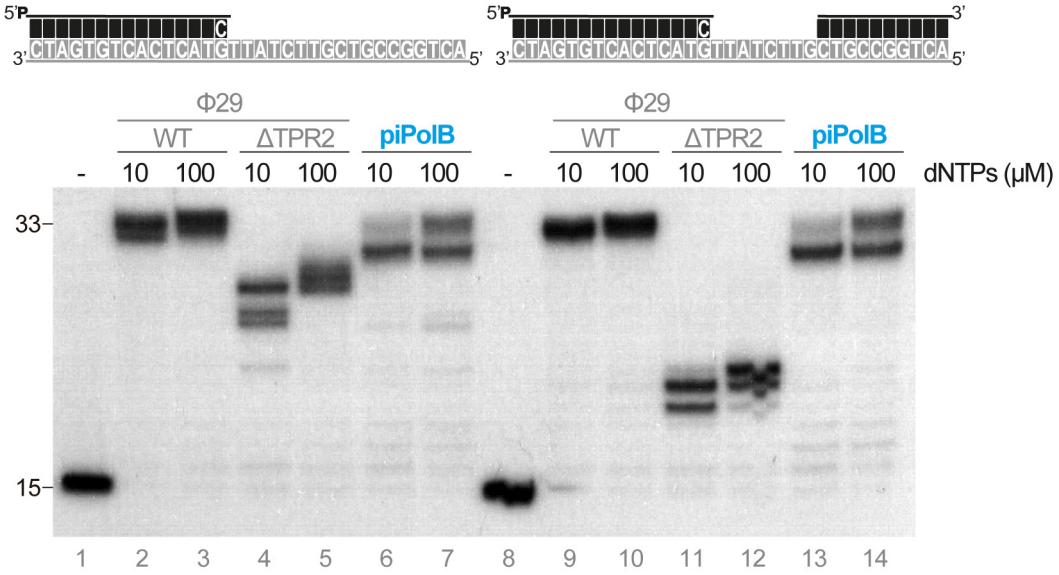

Figure S4. Strand displacement capacity of piPolB. Related to Figure 4.
Denaturing PAGE analysis of replication of a recessed primer/template duplex (lanes 1-7) or a 5 nucleotides gap (lanes 8-14) obtained with a downstream oligonucleotide (P20-33, Table S3) that must be displaced to resume replication by piPolB (lanes 6-7 and 13-14). Wild type and ΔTPR2 variants of Φ29 DNA polymerase were used as positive and negative controls, respectively.

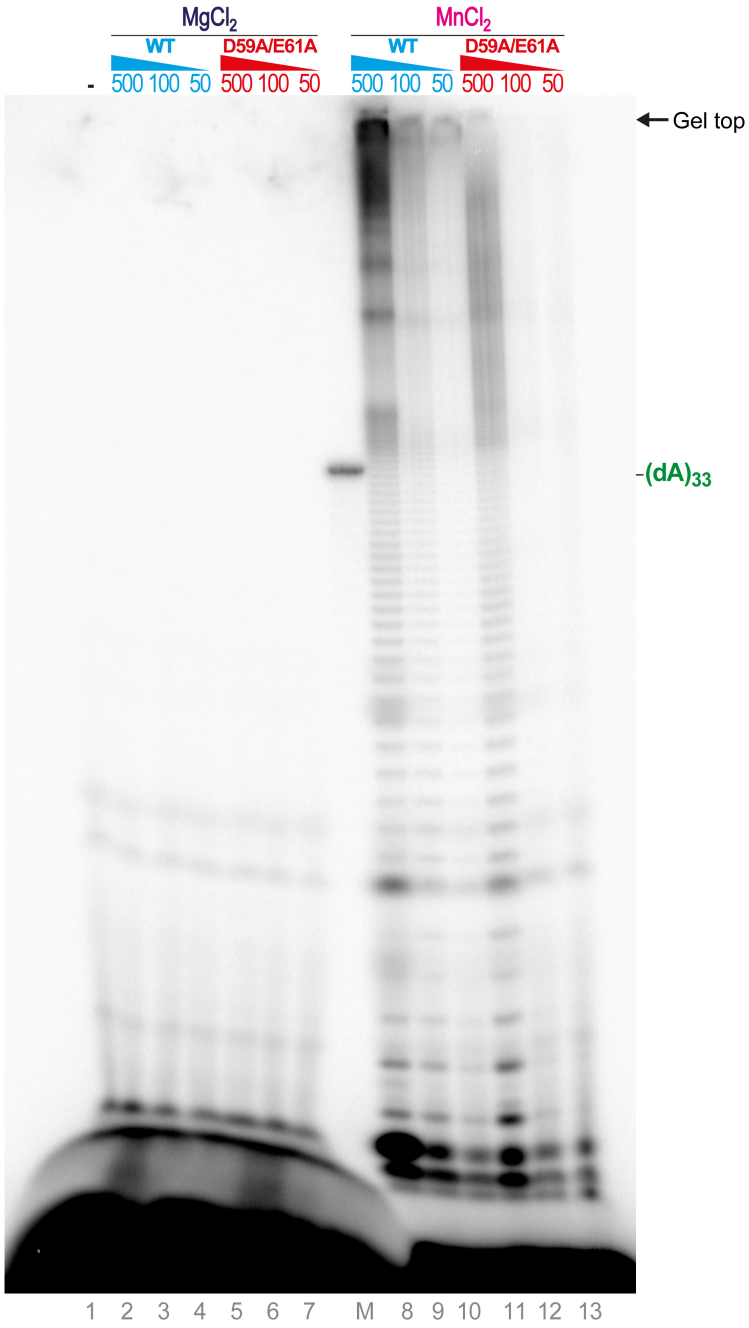

Figure S5. Effect of divalent cations on *de novo* replication of homopolymeric ssDNA substrate by piPolB. Related to Figure 4.

Primer synthesis and replication of homopolymeric poly-dT DNA template (1 μM) by wild type (WT) or exonuclease deficient (D59A/E61A) piPolB variants (500 nM). Reactions were triggered either with 10 mM $MgCl_2$ or 1 mM $MnCl_2$ and resolved in high resolution 8 M urea-20% PAGE. The P4, P10, P15 and 33A primers (see STAR Methods and Table S3) labeled with a 5'-^{32}P were loaded as size markers (lane M), and the size of the shorter products detected are indicated on the left and right, respectively.

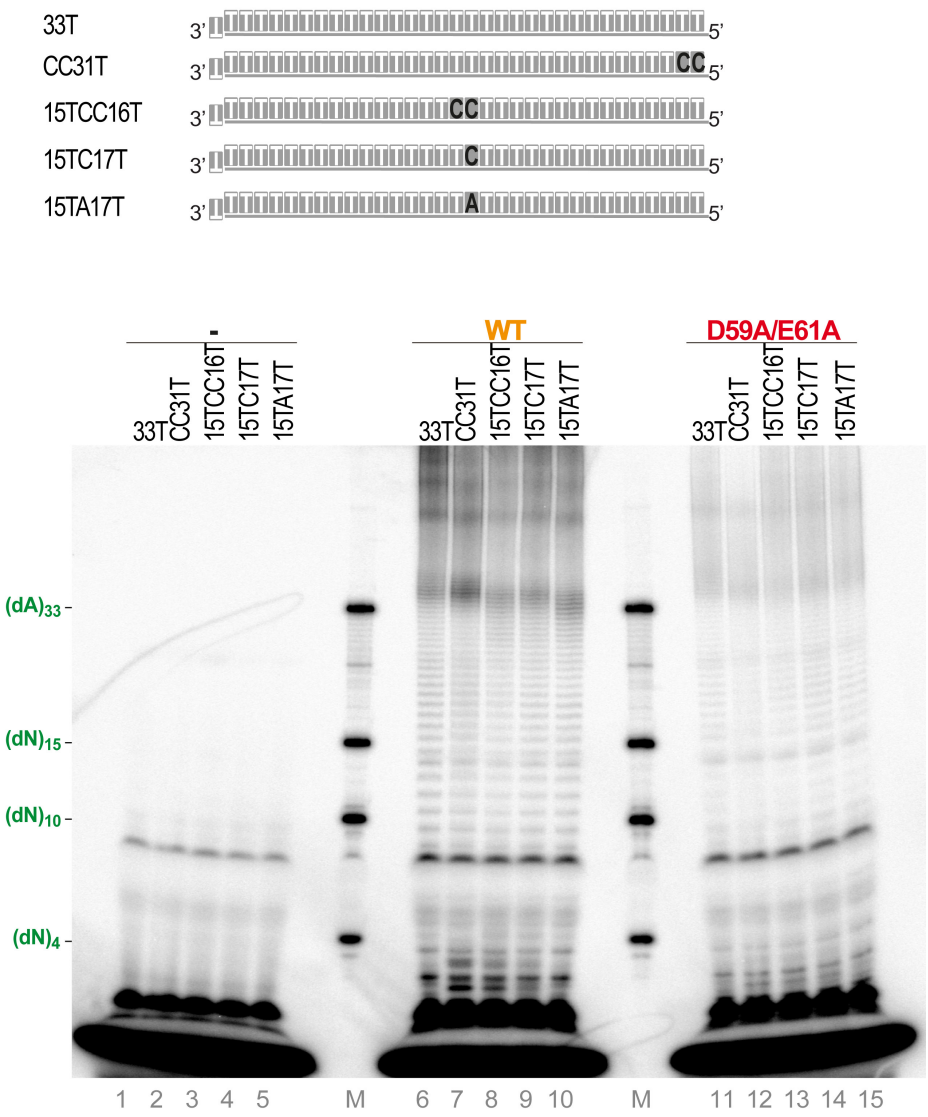

Figure S6. Small effect of single modifications of template sequence on *de novo* replication of homopolymeric ssDNA substrate by piPolB. Related to Figure 5.
Alternative ssDNA templates are depicted above the gel.

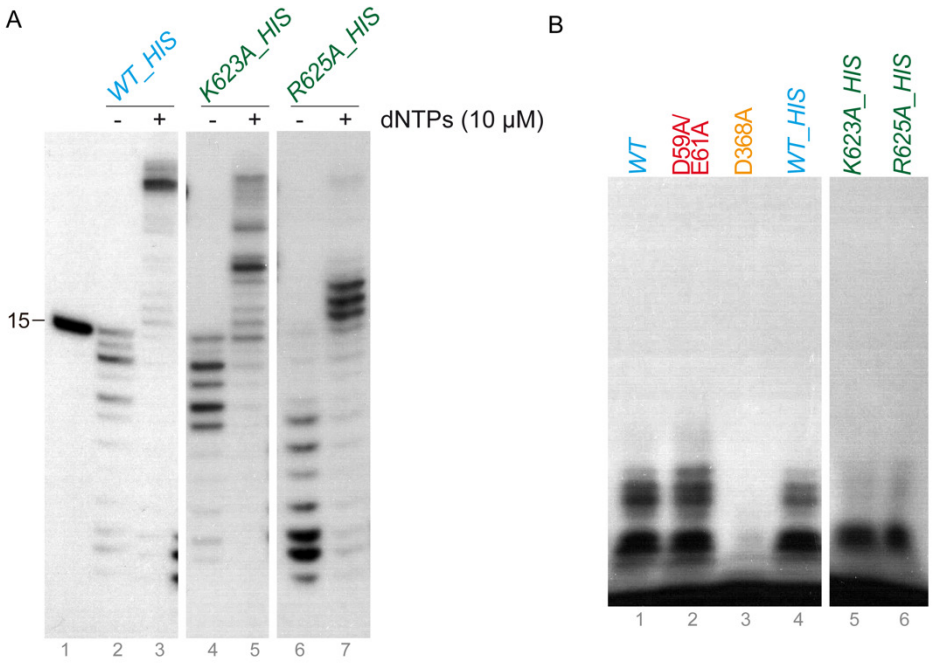

Figure S7. Polymerase and primase capacities of piPolB variants in the analog residues of PolB KxY motif. Related to Figure 6.

Primer extension (A) and primer synthesis (B) capacity of wild type, K624A and R626A His-tagged piPolBs. Assays were performed as in Figure 5. For reference, non-tagged wild type and exonuclease and polymerase deficient piPolB variants were also included in panel B.

Supplementary Table S1. Major characteristics of bacterial integrated pipolins. Related to Figure 1.

Name	Organism	Accession number	Coordinates	Size, bp	att, bp	Target
AciCal-E1	*Acidithiobacillus caldus* SM-1	CP002573	742597..793219	50,623	22	tRNA-Ser
ActICM54-E1	*Actinomyces* sp. ICM54	JDFI01000069	13259..25930	12,672	42	tRNA-Arg
AeroHyd-E1	*Aeromonas hydrophila* YL17	CP007518	1176890..1197918	21,029	57	tRNA-Leu
AurCor-E1	*Aurantimonas coralicida* DSM 14790	NZ_ATXK01000001	336290..362498	26,209	45	tRNA-Met
ColMT41	*Colwellia* sp. MT41	CP013145	3871337..3888198	16,862	70	tRNA-His
CorPse-E1	*Corynebacterium pseudogenitalium* ATCC 33035	GL542874	17142..28962	11,821	44	tRNA-Met
CorSin-E1	*Corynebacterium singular*	CP010827	813240..825107	11,868	77	tRNA-Met
CorATCC-E1	*Corynebacterium* sp. ATCC_6931	CP008913	649037..660848	11,812	35	tRNA-Met
CroSak-E1	*Cronobacter sakazakii* 696	CALF01000071	2319337..2355309	35,973	115	tRNA-Leu
DesAce-E1	*Desulfuromonas acetoxidans* DSM 684	AAEW02000002	57804..90620	32,817	12	tRNA-Glu*
Eco3-373-E1	*Escherichia coli* 3-373-03_S1_C2	NZ_JNMI01000006	64262..90063	25,802	113	tRNA-Leu
EcoE1520-E1	*Escherichia coli* E1520	AEHT01000035	186352..219882	33,531	116	tRNA-Leu
EcoETEC-E1	*Escherichia coli* ETEC H10407	NC_017633	2091087..2131115	40,029	126	tRNA-Leu#
EcoKOEGE-E1	*Escherichia coli* KOEGE_3	NZ_KE701180	343066..369002	25,934	122	tRNA-Leu
FerSen-E1	*Ferrimonas senticii* DSM 18821	NZ_AUGM01000004	223370..233973	10,604	118	tRNA-Trp
GorAma-E1	*Gordonia amarae* NBRC 15530	BAED01000024	21677..32602	10,926	36	tRNA-Lys
LacKis-E1	*Lactobacillus kisonensis* DSM 19906	AZEB01000002	15062..25687	10,626	53	intergenic
LacReu-E1	*Lactobacillus reuteri* TMW1.112	JOKX02000001	220481..239506	19,026	89	Intragenic (KEQ20280)
LeiANG-E1	*Leisingera* sp. ANG-M7	NZ_JWLI01000013	31462..57169	25,708	48	tRNA-Phe
MorPE36-E1	*Moritella* sp. PE36	NZ_ABCQ01000001	278975..301937	22,963	327	tRNA-Leu
MorVis-E1	*Moritella viscosa*	LN554852	4574306..4590347	16,042	187	tRNA-Leu*
PolCG9-E1	*Polaromonas* sp. CG9_12	CCJP01000005	2655959..2669601	13,643	179	tRNA-Phe
RhoY4I-E1	Rhodobacterales bacterium Y4I	NZ_DS995281	71969..85705	13,737	47	tRNA-Phe
RosMED193-E1	*Roseobacter* sp. MED193	AANB01000002	378959..391005	12,047	47	tRNA-Phe
SphLac-E1	*Sphingobium lactosutens* DS20	ATDP01000106	11226..44366	33,141	51	tRNA-Gly#
ThaSte-E1	*Thalassobacter stenotrophicus*	CYRX01000031	20244..36763	16,520	51	tRNA-Asn
VibFlu-E1	*Vibrio fluvialis* PG41	NZ_ASXS01000007	262170..285863	23,694	61	tRNA-Leu
VibFur-E1	*Vibrio furnissii* NCTC 11218	CP002377	779039..805644	26,606	95	tRNA-Leu
VibPar-E1	*Vibrio parahaemolyticus* NBRC 12711	NZ_BBQD01000016	8844..31803	22,960	57	tRNA-Leu
VibVul-E1	*Vibrio vulnificus* 93U204	CP009261	705355..717099	11,745	53	tRNA-Trp
VibVul-E2	*Vibrio vulnificus* 93U204	CP009261	2876933..2893717	16,785	174	tRNA-Leu*

* – integration into 5' end of the tRNA gene; # – tandem integration of two distinct elements.

Table S3. Sequences of oligonucleotides used in this work. Related to Experimental Procedures..

Name*	***Sequence (5'-3')***
P4	GATC
P10	GACTGCTTAC
P15	GATCACAGTGAGTAC
T33GTA	ACTGGCCGTCGTTCTATTGTACTCACTGTGATC
T33GTT	ACTGGCCGTCGTTCTAATGTACTCACTGTGATC
T33GFT	ACTGGCCGTCGTTCTAT**F**GTACTCACTGTGATC
P20-33	GAACGACGGCCAGT
33A	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
33T*	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT/invT/
CC31T*	CCTTTTTTTTTTTTTTTTTTTTTTTTTTTT/invT/
15TCC16T*	TTTTTTTTTTTTTCCTTTTTTTTTTTTTTT/invT/
15TC17T*	TTTTTTTTTTTTTCTTTTTTTTTTTTTTTT/invT/
15TA17T*	TTTTTTTTTTTTTATTTTTTTTTTTTTTTT/invT/
M13 UP	GTAAAACGACGGCCAGT

***F** represents the THF abasic site analog. */invT/ stands for a last dTMP nucleotide linked by an inverted 3'-3' bond.

Table S4. Gene Sequence Information and mutagenesis primers. Related to Experimental Procedures.

Codon-optimized sequence of B-family DNA polymerase from E. coli 3-373-03_S1_C2 pipolin (piPolB) cloned into NdeI-XhoI sites (underlined) of pET23a(+) plasmid	
CATATGAGCAATAACCTGCAAGACATCCTGGCTGCCGCTTCTGGCTACCAAAGTGTCACC TCGGAACCGGCCCTGAATCGCAAACGCCCGAAAACCCTGGATGACTATCCGGTTATTCCG CCGGCGAGCAAGAAAGTGAGCGTGATTAGCTCTGATCTGACCCTGCATATCGGTTTTGAC ACGGAATACGTGTTCAACCCGGAAACCCGCCAGAATGACATCCTGTCGTATCAAAGCTAC GTGGTTCTGCCGGATAACACGGGCATTTCCAATATTATCTATCCGCCGGACTCACAGAAA AAATCTCGTCTGAGTTTCAAAGATTTCCTGTGCCAAACCATTACGCCGCTGCTGGAAACC GGTGTTATCACGAAATGGCCGGGCATTATCAACATTTACGCCCACTTTATTCGCGCGGAC ATCGCCTCGTTTGCAAACTTCTGGAGCGATTACAAAATCCTGCTGAAAGGCATCCGTGGT ACCGTTAGTTCCTTTAAAAACCGCTACGGTATCGATTTCGACGAACAGCAAGAACGTCGC GTCAAAACCGAACAGATTATGTTTGATAAACGTACGTCTCCGCCGCGCTGCAGTAATGTG GCCTTCATTGATACCCTGCTGATCACGCCGGGCGGTATGGGTCTGGCAGAATGTGGCGAA CTGCTGGGTCTGCCGAAACTGACCATTCCGGCTCCGTATAGCATCACGAACATGCGCGAA TACCTGCTGGGTGACCGTGCAGGTTTTGAAGCGTATGCGCTGCGTGATGCTGAAATCGCG GTTCGCTACGCTCTGCAGGTCCGTAATTTCTGCGCGCGCGAACTGATGATTGATCGTGTGC CGGCAACCATTGGTGCCATGGCCGTTTCTCGTTTCACCAAAACGCTGAAAGAAAACAACA TGAGTCCGGAAGTGTGCCTGGGCACCCATATCAAAACGCGTGAACTGTGGCTGACCGAA AAACAGGCCTTTCGCACGATTAAAAACCCGGCATCCGTTCCGTCACGTGAACTGTTTGAA ACCTTCCCGATTAACTGCTATCATGGCGGTCGCAATGAATGTTTCATGATGGGTGTGACC CCGTCAGATCACTGGTATGATTACGACCTGGCAGGCGCTTATACCACGGGTCTGCTGGAT ATTCTGACCCCGGACTACGGCAACATCCGTCTGAGCAAAAATCCGGATGACTATTGCGGC CATGTGATGGGTTTTGCGCTGGTTACCTTTCGCTTCCCGGAATCCGTCCCGTATCCGTCAC TGCCGGTGCGTACGGATCAGTACGGTCTGTTTTTCCCGCTGAGCGGTGAAAGCTGGGCCA CCGCCCCGGAAATTGAACTGGCCCTGTCCCTGGGTGCAGAAATGACGATCCATAACGGCA TTATCGTGCCGTGGATTTGTGATACCAGCCCGCACAATAGTGAATCCACGTCAGTTTTTCT GCCGTTTGTGCAGCAAGTTCGCGAAAACCGTAATCGCCATATCAAAGGTTCCCTGGAAGA AAAATTCTGGAAAGAAATCGGCAACTCACTGTATGGTAAACTGGCTCAGGGCCTGCGTGC CAAAACCGCATTCGATACCGCGCGTGGCCTGAATCGCAGCCTGCCGCCGTCATCGGTCAC CCAACCGTTTTTCGCGGCCCACGTGACGGGTTTTATTCGCGCTGTCGTGGGCGAACTGAT GAATGCGCTGCCGTCTGATAGCTCTGTTGTCAGTGTGACCACGGACGGCTTTCTGACCAA CTGTCCGCTGGATAAAATCAATATGTCGGGTCCGCTGAGTTCCCGCTTCCAGAGCCTGTG CGATATTGTTGACCCGGGTTCATCGATGCTGACCTGTAAACATGAAGTCTCTCAACTGAT CGCCATGAAAACCCGTGGTCAGCTGACGTATAAAGCAATTCAAGGCAAACCGGTGGTTC ATGCACGCGCTGGTGTCAAACCGCCGGCCGACATTCCGCGTAGTGATTATAACGACTACA TGGTGGATCTGTACCTGAATCGTCTGCCGGGTCAGACCCTGTCGCGTAGCACCCTGATCT CGACGCGCGAAATGTGGCTGTCTGAAAGTGATCTGGTTAGCCGTGAACAGGACATTCGCC TGAACCTGGAATTTGATTTCAAACGTCAACCGGTGCGCCCGGCGATGAACGAAGGCCATC TGCTGATGTTTTCTCGTCCGTGGGATAATATGGAAGAAGCCCTGCAGCAACGTAGTCTGT TCGATGACTGGCGCCAGACCCACACGCTGAAAACCCTGGCCGATTGGGATGACTGGTGC GACTTTCTGTATTGTCGCACGGTTTTCTCTGATATGAAACTGAAAGTCGGCTCTAAACGTA GTGATGACATCCTGGTTCGTCTGTTTCTGCGCGCACTGACCCAGTGCCAATGGGGTCTGAT GCTGAAAGATAAAAAATCCTACTCATGTAAAGAAGTGGCGGAATGGCTGACCTCGGAAG GCTATAGCGTTACCGTCACGGATGTCAAAAATGCTGTGCGTGCGAAAATTCCGCAGATGA AATTTAGCTCTGTGACCCCGCGTATGAAATCCCTGATGGACATTATCGCCCGTAAATATC CGACCTTTTGCCTGCCGGTTTAACTCGAG	
Oligonucleotide pairs used for site-directed mutagenesis	
piPolB variant	*Sequence (5'-3')*
D368A	CCCCGTCAGATCACTGGTATGATTACGCCCTGGCAGGCGCTTATACCACG
	GCTGGTATAAGCGCCTGCCAGGGCGTAATCATACCAGTGATCTGACGGG G
D59A/E61A	CCCTGCATATCGGTTTTGCCACGGCATACGTGTTCAACCCGGAAACCC
	GGGTTTCCGGGTTGAACACGTATGCCGTGGCAAAACCGATATGCAGGG
WTHis	CCTTTTGCCTGCCGGTTTTACTCGAGCACCACCACCACCACCAC
	GTGGTGGTGGTGGTGGTGCTCGAGTAAAACCGGCAGGCAAAAGG
K614A	GGGTTCATCGATGCTGACCTGTGCACATGAAGTCTCTCAACTGATCGC
	GCGATCAGTTGAGAGACTTCATGTGCACAGGTCAGCATCGATGAACCC
H615A	GGGTTCATCGATGCTGACCTGTAAAGCTGAAGTCTCTCAACTGATCGC

	GCGATCAGTTGAGAGACTTCAGCTTTACAGGTCAGCATCGATGAACCC
K624A	GTCTCTCAACTGATCGCCATGGCAACCCGTGGTCAGCTGACG
	CGTCAGCTGACCACGGGTTGCCATGGCGATCAGTTGAGAGAC
R626A	GTCTCTCAACTGATCGCCATGAAAACCGCTGGTCAGCTGACGTATAAAG C
	GCTTTATACGTCAGCTGACCAGCGGTTTTCATGGCGATCAGTTGAGAGAC

Supplementary References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res *25*, 3389-3402.

Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972-1973.

García-Gómez, S., Reyes, A., Martínez-Jiménez, M.I., Chocron, E.S., Mourón, S., Terrados, G., Powell, C., Salido, E., Méndez, J., Holt, I.J.*, et al.* (2013). PrimPol, an archaic primase/polymerase operating in human cells. Mol Cell *52*, 541-553.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol *59*, 307-321.

Krupovic, M., and Koonin, E.V. (2015). Polintons: a hotbed of eukaryotic virus, transposon and plasmid evolution. Nat Rev Microbiol *13*, 105-115.

Okonechnikov, K., Golosova, O., Fursov, M., and team, U. (2012). Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics *28*, 1166-1167.

Pei, J., and Grishin, N.V. (2014). PROMALS3D: multiple protein sequence alignment enhanced with evolutionary and three-dimensional structural information. Methods Mol Biol *1079*, 263-271.

Soding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res *33*, W244-248.

Spriestersbach, A., Kubicek, J., Schafer, F., Block, H., and Maertens, B. (2015). Purification of His-Tagged Proteins. Methods Enzymol *559*, 1-15.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. Protein Expr Purif *41*, 207-234.

Sullivan, M.J., Petty, N.K., and Beatson, S.A. (2011). Easyfig: a genome comparison visualizer. Bioinformatics *27*, 1009-1010.