Identification and Characterization of Thermostable Y-Family DNA Polymerases η, ι, κ and Rev1 From a Lower Eukaryote, Thermomyces lanuginosus

Alexandra Vaisman, John P. McDonald, Mallory R. Smith, Sender L. Aspelund, Thomas C. Evans Jr, and Roger Woodgate

Y-family DNA polymerases (pols) consist of six phylogenetically separate subfamilies: two UmuC (polV) branches, DinB (pol IV, Dpo4, polκ), Rad30A/POLH (polη), and Rad30B/POLI (polι) and Rev1. Of these subfamilies, DinB orthologs are found in all three domains of life; eubacteria, archaea, and eukarya. UmuC orthologs are identified only in bacteria, whilst Rev1 and Rad30A/B pols are only detected in eukaryotes. Within eukaryotes, a wide array of evolutionary diversity exists. Humans possess all four Y-family pols (polη, polι, polκ, and Rev1), Schizosaccharomyces pombe has three Y-family pols (polη, polκ, and Rev1), and Saccharomyces cerevisiae only has polη and Rev1. Here, we report the cloning, expression, and biochemical characterization of the four Y-family pols from the lower eukaryotic thermophilic fungi, Thermomyces lanuginosus. Apart from the expected increased thermostability of the T. lanuginosus Y-family pols, their major biochemical properties are very similar to properties of their human counterparts. In particular, both Rad30B homologs (T. lanuginosus and human polι) exhibit remarkably low fidelity during DNA synthesis that is template sequence dependent. It was previously hypothesized that higher organisms had acquired this property during eukaryotic evolution, but these observations imply that polι originated earlier than previously known, suggesting a critical cellular function in both lower and higher eukaryotes.

Keywords: thermostable fungi, Y-family DNA polymerases, phylogenetic analysis, translesion DNA synthesis, DNA polymerase eta (polη), DNA polymerase iota (polι), DNA polymerase kappa (polκ), Rev1

INTRODUCTION

The Y-family DNA polymerases are responsible for copying damaged DNA during DNA replication in a process called translesion synthesis (TLS) (Sale et al., 2012). These enzymes are highly specialized in order to accommodate different structural DNA distortions caused by a wide variety of DNA lesions. The Y-family is divided into six phylogenetically distinct subfamilies: two UmuC (polV)
branches; Rad30A/POLH (polη); Rad30B/POL1 (pol); and DinB (pol IV, Dpo4, polκ); and Rev1 (Ohmori et al., 2001). Across the different domains of life, Y-family polymerase subfamilies are found in various combinations. For example, Umuc orthologs are only detected in Gram-positive and Gram-negative bacteria, whereas Rev1 and Rad30A/B orthologs are only detected in eukaryotes. The DinB subfamily is the most evolutionarily conserved, having members scattered throughout all three domains of life from unicellular bacteria to humans. However, differences in the distribution of Y-family DNA pols are present within each kingdom. For example, the eukaryote Saccharomyces cerevisiae (S. cerevisiae) contains neither a POLK nor a POLI gene. Indeed, it was originally assumed that poli was expressed only in higher eukaryotes. However, next generation whole genome sequencing has revealed that poli orthologs are actually distributed throughout the whole Eukaryota domain. One example is the thermophilic fungus, Thermomyces lanuginosus (T. lanuginosus) which possesses all four eukaryotic Y-family subfamilies much like humans, in contrast to its fungal relatives, S. cerevisiae and Schizosaccharomyces pombe (S. pombe). Is there logic in such seemingly random distribution of pol? Using phylogenetic analysis and comparing the biochemical characterization of Y-family pols from different species, we hoped to shed some light on this question.

The Y-family DNA pols are classically described as specializing in TLS activity that arises from their capacious active sites that accommodate DNA lesions which would otherwise obstruct the processive confined active sites of A- and B-family replicative pols (Sale et al., 2012). Each Y-family polymerase is tailored to process different lesions, leaving behind unique errors after gaining access to the replication fork (Yang and Gao, 2018). Polymerases belonging to the same subfamily often specialize in targeting the same type of DNA lesions. For example, DinB/polk orthologs are very adept at bypassing minor groove DNA adducts (e.g., N2-dG adducts) (Liu et al., 2014; Basu et al., 2017; Jha and Ling, 2018; Stern et al., 2019). Although, the archaeal ortholog, Dpo4, is notably less efficient at bypassing bulky aromatic lesions than its eukaryotic ortholog polκ (Avkin et al., 2004; Ling et al., 2004b; Choi et al., 2006). Along with a common preference for bulky aromatics, DinB pols also share a propensity for template slippage that increases deletion events in the mutation spectra (Kokoska et al., 2002).

Eukaryotic Rad30 orthologs polη and poli are similar in sequence but exhibit very different TLS properties. Human and S. cerevisiae polη are exceptionally efficient at bypassing a thymine-thymine cyclobutane pyrimidine dimer (CPD) (Johnson et al., 2000c; Masutani et al., 2006; McCulloch et al., 2004). Although poli can insert nucleotides opposite CPDs, the efficiency is substantially lower than polη-catalyzed TLS (Tissier et al., 2000a; Vaisman et al., 2003). Poli also has a unique feature; it misincorporates dG opposite dT, 3- to 10-fold more frequently than the correct base dA (Johnson et al., 2000b; Tissier et al., 2000b; Jha et al., 1996). This is achieved using a deoxycytidyl transferase mechanism that limits Rev1 to exclusively incorporate dC (Johnson et al., 2000b; Tissier et al., 2000b; Zhang et al., 2000c). The unique feature of eukaryotic Rev1 orthologs is their efficiency at bypassing both damaged guanines and abasic sites using a deoxycytidyl transferase mechanism that limits Rev1 to exclusively incorporate dC (Nelson et al., 1996). This is achieved by displacing the DNA lesion from the active site entirely and instead using a protein sidechain (R324 in S. cerevisiae and R357 in H. sapiens) as the “template” which base pairs solely with dC (Nair et al., 2005, 2011; Weaver et al., 2020).

In this manuscript, we describe the identification, purification and characterization of thermostable eukaryotic orthologs of polη, poli, polκ, and Rev1 from a thermophilic, multicellular fungal species, T. lanuginosus. Biochemical characterization of TLS DNA pols η, ι, κ, and Rev1 include determination of the enzyme’s fidelity, processivity, thermostability, metal ion requirements, and TLS specificity during bypass of CPDs, abasic sites, and benzo[a] pyrene diol epoxide (BPDE) adducts. Our findings serve as basis for comparative analysis of the properties of proteins from different species, thus providing an important insight into the functional evolution of the Y-family polymerases.

**MATERIALS AND METHODS**

**Bacterial Plasmids**

Plasmids used in this study are described in Table 1. Where noted, bacteria were grown on LB agar plates containing 20 μg/ml chloramphenicol; 25 μg/ml zeocin; 30 μg/ml kanamycin; 20 μg/ml spectinomycin; or 100 μg/ml ampicillin.

**Identification and Cloning of Y-Family Orthologs From T. lanuginosus**

At the onset of this investigation, the complete genomic sequence of T. lanuginosus was not yet available. Therefore, in order to identify and clone the various Y-family polymerase orthologs from T. lanuginosus, we used several PCR-based procedures. Initially, we employed a degenerate PCR approach by first generating protein homology comparisons for known fungal polη, poli, polκ, and Rev1 proteins. Regions of conserved amino acid stretches within these four polymerase families were identified and degenerate PCR primers were designed based on selected conserved amino acid regions. Purified T. lanuginosus genomic DNA was purchased from DSMZ (Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH) (DSM 10635). PCR and nested-PCR were performed utilizing this T. lanuginosus genomic DNA and amplicons of gene segments encoding these four Y-family polymerases were cloned and sequenced. Based on the cloned DNA sequences that encoded regions of the Y-family polymerase genes, gene-specific PCR primers were designed. Additional PCR reactions were performed either using...
genomic DNA or DNA from a *T. lanuginosus* cDNA library, kindly provided by the Fungal Genomics Project at Concordia University (Adrian Tsang; Director of the Center for Functional and Structural Genomics). PCR techniques employed included RACE PCR (cDNA library), Flanking-sequence PCR (Sorensen et al., 1993) utilizing degenerate biotinylated primers and overlapping exon PCR (genomic DNA). Full-length genes were amplified with gene-specific primers with the addition of an *NdeI* site at the 5’ end and either a *BamHI* or *BglII* (POLK) site at the 3’ end and cloned into pGEM vectors for sequencing. The *POLH, POLK* and *REV1* genes were subcloned into the *NdeI* to *BamHI* sites of pET22b+ for protein expression. An *E. coli* codon optimized version of the *POLI* gene was subsequently synthesized by Genscript (Piscataway, NJ) and cloned into the pJM871 low expression vector (Frank et al., 2012) from *NdeI* to *BamHI*.

Subsequently, the full *T. lanuginosus* SSBP genomic sequence was published and released (https://mycocosm.jgi.doe.gov/Thelan1/Thelan1.home.html) (McHunu et al., 2013) (GenBank assembly accession GCA_000315315.1), and we were able to confirm that our clones do indeed encode the *T. lanuginosus* polι, polθ, polκ, and Rev1 proteins. The *POLH* gene is encoded on contig00146 (GenBank accession ANHP01000280) from nucleotide 84213–86129. The *POLI* gene is encoded on contig00250 (GenBank accession ANHP01000232) from nucleotide 126087–127864 with two introns. The *POLK* gene is encoded on contig00025 (GenBank accession ANHP01000258) from nucleotide 58089–60067 with one intron. The Rev1 gene is encoded on contig00203 (GenBank accession ANHP01000194) from nucleotide 112065–115607 with one intron. The Joint Genome Institute MycoCosm database designations for these proteins is as follows: polι – jgi|Thelan1|4409|TLAN_03021-R0; polθ – jgi|Thelan1|3308|TLAN_02476-R0; polκ – jgi|Thelan1|3900|TLAN_01802-R0; Rev1 – jgi|Thelan1|3900|TLAN_01802-R0.

### Phylogenetic Tree Construction

A multiple sequence alignment of the newly described *T. lanuginosus* polι protein and polθ proteins from various other organisms was performed using the ClustalW algorithm in MacVector (version 15.5.4). Known polθ protein sequences were obtained from Genbank protein records or identified by BLAST homology searches of Genbank genomic sequence records. Genbank accession numbers are indicated within the legend of Figure 1. The alignment was exported from MacVector in the Nexus sequence file format. This Nexus file was imported into the SplitsTree4 (version 4.14.4) and an rooted phylogenetic tree was generated by setting the distance method to BioNJ.

### Purification of the *T. lanuginosus* DNA Polι

*E.coli* strain RW644 [F− dcm ompT hsdS(rB2− mB−) gal λ(DE3) ΔmutDC596:ermGT ΔdinB1:ble ΔaraD-polBΩ Δ(gpt-proA)62] harboring pJM596 was grown overnight in 20 ml LB media containing 100 μg/ml ampicillin at 37°C. The overnight culture was then transferred into 1L fresh LB-ampicillin and grown at 37°C until an OD600 ~0.5. At this point, IPTG was added to a final concentration of 1 mM to induce expression of *T. lanuginosus* DNA polι and grown for an additional 2 h before cells were harvested by centrifugation. The cell pellet was resuspended in 15 ml lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10 mM β-mercaptoethanol, 20% v/v glycerol) and dialyzed against buffer A (50 mM Tris pH 8.0, 10 mM NaCl, 10 mM β-mercaptoethanol, 20% v/v glycerol) and dialyzed overnight at 4°C against 1L of the same buffer. The dialyzed protein suspension was applied to a 5 ml Ni-NTA HiTrap DEAE FF column (Cytiva, cat#17515401) and bound proteins were eluted with a 20–500 mM linear gradient of NaCl. Fractions containing polι were pooled and dialyzed against buffer B (10 mM sodium phosphate, pH 7.0,10 mM β-mercaptoethanol, 20% v/v glycerol) and applied to a 5 ml Bio-Scale™ Mini CHT™ Type II cartridge (BioRad, cat#7324332) and eluted with a 10–300 mM linear gradient of sodium phosphate in buffer B. Peak polι-containing fractions were pooled, aliquoted and stored at −80°C.

### Purification of the *T. lanuginosus* DNA Polκ

*E.coli* strain RW644 harboring pJM682 and purified using the same protocol as described above for polι.
Our attempts to purify untagged *T. lanuginosus* pol using the above protocol for polι and polκ proved to be unsuccessful. We therefore decided to purify an N-terminal His-tagged version of *T. lanuginosus* polι. To do so, plasmid pJM966 was introduced into the *E.coli* SHuf® T7 Express strain [F lac, pro, lacIq/Δ ara-leu]7697 araD139 fhuA2 lacZ: T7 gene1Δ (phoA), Actinia tenebrosa (Sea anemone), Flustrellidae echiura (Segmented worm), Agaricus bisporus (Mushroom), Lepidocharysella muscorum (Green alga), Neurospora crassa (Fungi), Nematostella vectensis (Cnidaria) and *Xenopus laevis* (XP_018100015). Purification of the *T. lanuginosus* DNA Pol ι

Our attempts to purify untagged *T. lanuginosus* pol and human pol, was performed using the ClustalW algorithm in MacVector (version 15.5.4). This alignment was exported from MacVector as a Nexus file which was then imported into SplitsTree4 (version 4.14.4) to generate an unrooted phylogenetic tree by setting the distance method to BioNJ. Most of the pol proteins cluster into several groups comprised of chordates (blue), lower animals (red), lower plants (green), fungi (orange) and diatoms (pink). Pol protein sequences for each organism were obtained from the following Genbank sequence files: *Acanthaster planci* (XP_022107640), *Actinia tenebrosa* (XP_035668094), *Catenaria anguillulae* (ORZ32752), *Chara braunii* (GBG75489), *Crassostrea virginica* (XP_023314304), *Danio rerio* (NP_001017834), *Dimorphus gyrociilatus* (CAD512353), *Fistulifera solaris* (GAX13589), *Haliaeetus leucocephalus* (XP_010581340), *Homo sapiens* (NP_001338561), *Jimgerdemannia flammicorona* (RUP43481), *Klebsormidium nitens* (GAQ87748), *Lipomyces starkeyi* (ODQ70434), *Mus musculus* (NP_006102), *Neurospora crassa* (CAD70388), *Phaeodactylum tricornutum* (XP_00178064), *Raphidocelis subcapitata* (GBF90329), *Salpingoeca rosetta* (XP_004998634), *Selaginella moellendorffii* (CEL97102) and *Xenopus laevis* (XP_018100015).
(BioRad, cat#7324332) and eluted with a 10–300 mM linear gradient of sodium phosphate. Peak pol-containing fractions were pooled, aliquoted and stored at −80°C.

**Purification of *T. lanuginosus* Rev1**

RW644 harboring pRARE2 (expressing tRNAs for seven rare codons: AGA, AGG, AUU, CUA, GGA, CCC, and CGG) was transformed with pJM863. The strain was grown overnight in 60 ml LB media containing 100 μg/ml ampicillin at 37°C. The overnight culture was then transferred into 3L fresh LB-ampicillin (1.5 L each in 2 x 4L flasks) and grown at 37°C until an OD\textsubscript{600} ~0.5. At this point, IPTG was added to a final concentration of 1 mM to induce expression of *T. lanuginosus* Rev1 and grown for an additional 2 h before cells were harvested by centrifugation. The cell pellet was resuspended in 60 ml lysis buffer (50 mM Tris pH 7.5, 0.5 mM EDTA, 1 mM DTT, 20% glycerol), sonicated, and cleared by ultracentrifugation at 45,000 rpm in a Beckman 50.2 Ti rotor for 45 min. Ammonium sulfate (35% saturation, 0.2 g/ml) was slowly added to the cleared lysate and stirred at 4°C for 30 min. Precipitated proteins were harvested by centrifugation and the resulting pellet resuspended in a 200 μl of the same buffer. The dialysed lysate was applied to a 20 ml HiPrep 200 HR size exclusion column (Cytiva, Cat#17116601) and dialyzed overnight against 2L of the same buffer equilibrated with buffer (20 mM KPO4 pH 7.0, 100 mM NaCl) and dialyzed overnight against 2L of the same buffer at 4°C. The dialysed lysate was applied to a 20 ml HiPrep Heparin FF 16/100 column (Cytiva, cat#28936549) and eluted with a 25–400 mM NaCl linear gradient. Fractions containing Rev1 were pooled and applied to a HiPrep 16/60 Sephacryl S-200 HR size exclusion column (Cytiva, Cat#17116601) equilibrated in Phosphate Buffered Saline (PBS, pH6.8). Rev1 containing fractions were pooled and applied to a 5 ml Phosphocellulose P11 column (Whatman), equilibrated with buffer "F" (20 mM KPO4 pH 7.0, 100 μM EDTA, 10% glycerol, 1 mM DTT, and 100 mM KCl) and eluted with a 200–800 mM KCl linear gradient. Rev1-containing fractions were pooled, concentrated, and aliquoted for storage at −80°C.

**DNA Templates**

Undamaged synthetic oligonucleotide primers and an 48 bp abasic site-containing template were synthesized by Lofstrand Laboratories (Gaithersburg, MD) using standard technique and PAGE purified. A synthetic abasic site, dSpacer, was purchased from Glen Research (Sterling, VA) and incorporated into the oligonucleotide template using standard techniques by Lofstrand. The 7.2 kb M13mp18 circular template was purchased from New England Biolabs (Ipswich, MA). The *cis*-syn CPD containing oligonucleotide was synthesized by TriLink BioTechnologies (San Diego, CA) and has been described previously (Boudsocq et al., 2004). The synthesis of a template with BPDE-da adduct was described previously (Frank et al., 2002). The sequence of each primer-template pair is given in the legend of the respective figures.

**RESULTS AND DISCUSSION**

**Identification of Y-Family Polymerases in *T. lanuginosus***

Based on previous investigations and observations, it has been noted that the genomes of higher eukaryotes, such as humans and
mice, possess genes for the four known eukaryotic Y-family pols, whereas, the model organism *S. cerevisiae* only possesses genes for two of the eukaryotic Y-family pols, RAD30 and REV1. Based on this observation, it has been suggested the genes encoding the closely related polh and poli proteins are paralogs of each other that diverged during evolution sometime after the origination of fungi. However, BLAST homology searches revealed that the genomes of filamentous fungi such as *Neurospora crassa* and *Aspergillus nidulans* and even the corn smut *Ustilago maydis* harbor a POLI gene, indicating to us that the genetic and biochemical characteristics of translesion synthesis in some "higher" fungi could be quite akin to that in higher eukaryotes such as humans. This notion of conserved functionality provided the impetus to embark on our efforts to identify and clone the four eukaryotic Y-family pols from the fungi, *T. lanuginosus*. Since *T. lanuginosus* is thermophilic, we reasoned the Y-family pols from this organism would exhibit enhanced stability during purification and biochemical analysis (see below).

In order to identify and clone the *T. lanuginosus* Y-family pols, protein sequence similarity comparisons were performed using known polη, polκ, and Rev1 protein sequences, including other fungal sequences. Initially, degenerate PCR primers were designed based on highly conserved regions of protein homology between the various proteins within the four Y-family polymerase groups. Degenerate PCR was performed using *T. lanuginosus* genomic DNA and amplicons were cloned and sequenced. Subsequently, sequence of cloned regions of the Y-family polymerase genes were used to design gene-specific primers. A combination of various PCR techniques was then employed, for example degenerate PCR, nested PCR, RACE PCR, over-lapping exon PCR and flanking-sequence PCR using either genomic DNA or cDNA library DNA, to generate full-length sequences of each of the *T. lanuginosus* POLH, POLI, POLK and REV1 genes. Full-length genes were then subcloned into protein expression vectors to facilitate purification of the four Y-family DNA pols. After we completed cloning the Y-family pol genes, the *T. lanuginosus* genome was sequenced and published allowing us to confirm that our cloned sequences do indeed encode the correct *T. lanuginosus* proteins. Sequence comparisons between *T. lanuginosus* and human Y-family pols reveal modest similarities that are higher than those observed between *S. cerevisiae* and human homologs. *T. lanuginosus* and human polη have 22% identical and 15% similar residues (compare to *S. cerevisiae* versus human at 16% identical and 15% similar residues). *T. lanuginosus* and human Rev1 have 24% identical and 15% similar residues (compare to *S. cerevisiae* versus human at 16% identical and 16% similar residues). Additionally, *T. lanuginosus* and human polκ have 20% identical and 14% similar residues, and *T. lanuginosus* and human polη have 24% identical and 16% similar residues.

**Phylogenetic analysis of the Y-Family Polymerase poli**

The observation that the four eukaryotic Y-family pols are present in many filamentous fungi suggested to us that the divergence of POLI and POLH is even more ancient than the divergence of the fungi and metazoan (animals) groups. This supposition prompted us to perform a more in-depth phylogenetic analysis on the origins of the poli protein. We executed an exhaustive search for poli protein sequences within the GenBank database, including the fungi and metazoan groups, all other groups with clades containing various "lower" eukaryotes, and the plant kingdom. Although the POLI gene was absent from many "lower" or "simpler" eukaryotic organisms than fungi, as well as higher plants, we did find several examples of lower eukaryotes and simpler plants that do in fact encode a POLI gene.

A ClustalW alignment was performed using a group of sequenced POLI genes from organisms that include simple eukaryotes, lower plants, fungi and metazoans, all of which also possess genes for POLH, POLK and REV1. The unrooted phylogenetic tree of this alignment (Figure 1) reveals that the poli proteins cluster, for the most part, within several groups that include a chordate group, a lower animal grouping, including arthropods, annelids, echinoderms and mollusks, a fungi group, and a lower plant group. Poli proteins from other organisms that did not fit into these groups include a protein from a sea anemone (*N. solaris*), two proteins from diatom organisms and proteins from an Alveolate and a Choanoflagellate, which are examples of very simple single-celled eukaryotes. Interestingly, our analysis further supports the perception that the POLI gene is quite commonly found in fungal species, including some yeasts, and even other lower eukaryotic organisms and therefore substantiates our supposition that the evolutionary “split” between POLH and POLI may have preceded the emergence of fungi.

During this phylogenetic analysis of poli, one particularly noteworthy and unexpected observation was that POLI genes were quite frequently identified in some organisms within a particular group and absent in other organisms within that group. This finding holds true from the simplest eukaryotes to higher eukaryotes. For example, many protists do not possess a POLI gene such as Metamonads (*e.g.*, diplomonads, *i.e.*, *Giardia lamblia* and parabasalids, *i.e.*, *Trichomonas vaginalis*), Euglenozoans (*e.g.*, *Trypanosomatids, i.e.* *Trypanosoma brucei* and *Leishmania donovani*) and Amoebozoans (*e.g.*, *Entamoeba histolytica*). While other groups of protists do have a POLI gene such as Heteroloboseas (*e.g.*, *Naegleria gruberi* and *Alveolates* (*e.g.*, *Vitrella brassicaformis* (*Colpodellida clade*))). In direct contrast, other species of Alveolates in the Apicomplexans clade (*e.g.*, *Cryptosporidium parvum* and the Ciliophorans clade (*cliteta* (*e.g.*, *Paramecium tetraurelia*) do not have a POLI gene. In addition, some Stramenopiles have a POLI gene (*e.g.*, diatoms, *i.e.*, *Phaeodactylum tricornutum* and *Fistulifera solaris*) and some do not (*e.g.*, diatoms, *i.e.*, *Thalassiosira oceanica* and *e.g.*, *water mould, i.e.*, *Saprolegnia parasitica*). Thus, it seems the presence, or absence, of a POLI gene in the genome of any specific organism is quite indiscriminate.

This unpredictable “hit or miss” trait of POLI holds true in plants, fungi, and metazoans as well. For example, we found POLI genes in the genomes of lower plant forms such as microalgae,
green algae, and spikemoss but were unable to find POLI genes in any higher vascular plants. A similar situation exists in fungi. No POLI genes were found in lower fungal forms such as Microsporidia and Glomeromycota. Likewise, in Saccharomycotina (true yeasts) (e.g., S. cerevisiae) no POLI genes were found. Some fungi in the Taphrinomycotina subphylum, such as Saitoella complicate, have POLI genes, while the closely related fission yeasts, Schizosaccharomyces species, do not. Furthermore, POLI genes were found in other yeast forms (e.g., Lipomyces starkeyi), filamentous fungi, parasitic, and symbiotic fungi, and corn smut (Figure 1). However, we found no POLI genes in Basidiomycota which include club fungi and mushrooms.

Although most organisms within the various groups of animals, from the Choanoflagellates (considered to be the closest living relatives of the animals) up to the mammals, possess a POLI gene, we found it remarkable that there are numerous examples of higher eukaryotes organisms that do not possess a POLI gene. For example, we found no POLI genes in most Platyzoa (i.e., Platyhelminthes, e.g., flat worms), Rotifera (i.e., wheel animals), and Annelids (i.e., segmented worms), with the notable exception of Dimorphilus gyrociliatus which does harbor a POLI gene (Figure 1). While, many types of Mollusks do possess POLI genes, we found no POLI gene in octopi. Apparently, Arthropods such as spiders, scorpions, sea spiders, centipedes, millipedes and tardigrades (water bears) do not have POLI, but horseshoe crabs do. With a few exceptions, the majority of crustaceans and insects have POLI. Lastly, we found that the vast majority of chordates from the lancelet to humans do indeed possess a POLI gene (Figure 1). However, there are some extraordinary exceptions. We found that the tunicates (sea squirts), which are lower chordates, lack a POLI gene. Thus far, only a single fish species, Esox lucius (northern pike), was identified as lacking POLI. While alligators possess a POLI gene, the Australian saltwater crocodile (Crocodylus porosus) and the gharial (Gavialis gangeticus) both lack POLI. Most astounding, the majority of bird species, including Galliformes (e.g., chickens, turkeys, grouse, quail, partridges and pheasants), Bucerotiformes (hornbills), Upupiformes (hoopoes) and the vast majority of Passeriformes (e.g., sparrow, finch, tit, crow, canary, rifleman, and starlings) which includes more than half of all bird species, lack a POLI gene.

This phylogenetic analysis provides a “snapshot” of the evolution of POLI at this given moment in time. At some time in the past, the polη and polη paralogs diverged from one another due to a need to have two biochemically distinct translesion synthesis polymerases. The biochemical properties of both human polη and polη are well characterized. However, unlike polη which is defective in the human Xeroderma Pigmentosum Variant syndrome (Masutani et al., 1999), the cellular function of polη remains enigmatic (McIntyre, 2020; Vaisman and Woodgate, 2020). Therefore, the evolutionary pressure that led to this division of polη and polη remains a mystery. Perhaps this seemingly arbitrary “hit or miss” distribution of POLI genes in eukaryotic organisms can provide clues as to the cellular functioning of polη. Evaluation of the biological and biochemical translesion synthesis properties of polη and polη from closely related lower eukaryotes, one of which possesses a POLI gene and the other that does not, may provide hints as to when, and why, these two paralogs diverged and the cellular functions of polη.

### Expression and Purification of T. lanuginosus Polymerases η, ι, κ, and Rev1

All four Y-family pols from T. lanuginosus were expressed in E. coli, but their expression levels, solubility, and ease of purification varied. Both pols η and κ were well-expressed in E. coli and readily purified by conventional protocols including ammonium sulfate precipitation and ion-exchange chromatography (Figure 2). Rev1 was initially poorly expressed (unpublished observations), but expression increased dramatically in the presence of the pRARE2 plasmid coding for rare E.coli tRNAs. T. lanuginosus Rev1 was also purified by conventional methods including ammonium sulfate precipitation, size-exclusion and ion-exchange chromatography (Figure 2). T. lanuginosus polη was the most problematic of the four Y-family polymerases to purify, requiring codon optimization for E.coli expression and the addition of an N-terminal Histidine tag. Subsequently, the protein was purified by affinity- and ion exchange-chromatography (Figure 2). Once the polymerases were purified to >95% homogeneity, they were characterized for metal ion requirement, ability to bypass lesions, fidelity, thermostability, and processivity.
Activation of T. lanuginosus Polymerases η, κ, ι by Mn$^{2+}$ and Mg$^{2+}$

DNA polymerases are known to require divalent cations to catalyze the nucleotidyl transfer reaction (Bessman et al., 1958; Mildvan and Loeb, 1979; Loeb and Kunkel, 1982; Vashishtha et al., 2016). Although it is generally believed that Mg$^{2+}$ is the activating co-factor in vivo due to its cellular abundance, other divalent metal cations such as Mn$^{2+}$, Co$^{2+}$, Ni$^{2+}$, Zn$^{2+}$ have the ability to substitute for Mg$^{2+}$ under certain conditions (Mildvan and Loeb, 1979; Loeb and Kunkel, 1982). Different metal co-factors can affect the activity and fidelity of DNA polymerases, promote TLS, and in the case of X-family enzymes, increase polymerization efficiency (Martin et al., 2013). In rare cases, Mn$^{2+}$ appears to serve as a natural metal activator of DNA polymerases. This has been shown for X-family polymerases λ and μ and Y-family pol which exhibit great preference for Mn$^{2+}$ over Mg$^{2+}$ as the activating metal ion (Dominguez et al., 2000; Frank and Woodgate, 2007; Garcia-Diaz et al., 2007). In particular, human pol was inhibited by MgCl$_2$ present at physiological concentrations and exhibited peak activity at low MnCl$_2$ levels (0.05–0.25 mM). Furthermore, at optimal concentrations, Mn$^{2+}$ ions improved the fidelity of pol-catalyzed nucleotide incorporation opposite template dT.

We were interested in determining the optimal concentration and type of divalent metal ion that is required for activation of T. lanuginosus pols η, κ, and ι. We assessed the primer extension activity of these polymerases in the presence of Mg$^{2+}$ and Mn$^{2+}$ at concentrations ranging from 62.5 μM to 8 mM (Figure 3).
Polη exhibited greatest activity in reactions with 4–8 mM Mg\(^{2+}\) and 0.5–8 mM Mn\(^{2+}\) (Figure 3A). Reactions catalyzed by polκ were most efficient at 1–8 mM Mg\(^{2+}\) and 0.25–8 mM Mn\(^{2+}\) (Figure 3C). Polη was most active in the presence of ∼4 mM Mn\(^{2+}\) (Figure 3B).

As shown in Figure 3, Mn\(^{2+}\) appears to be a more conducive co-factor than Mg\(^{2+}\) in reactions with all three T. lanuginosus pols, as they all were active over a wider range of Mn\(^{2+}\) and exhibited greater activity at the same concentration of Mn\(^{2+}\) vs. Mg\(^{2+}\). However, there are some interesting differences seen between the reported metal ion requirements for human and T. lanuginosus pols. Human polη exhibits highest activity at 0.05–0.2 mM MnCl\(_2\) (Frank and Woodgate, 2007), whereas T. lanuginosus enzyme required much higher concentrations of Mn\(^{2+}\) with a sharp peak of activity at ∼4 mM MnCl\(_2\). In contrast, while human polκ was inhibited by MgCl\(_2\) at concentrations >1 mM, T. lanuginosus polκ required at least 0.5 mM Mg\(^{2+}\) and its activity gradually increased with concentrations up to 8 mM.

Because Mn\(^{2+}\) is generally known to decrease replication fidelity, we used 4 mM Mg\(^{2+}\) for the subsequent studies with polη and polκ. In contrast, further characterization of polη was performed in the presence of 4 mM Mn\(^{2+}\).

**Processivity of T. lanuginosus Polymerases η, ι and κ**

We next examined the processivity of T. lanuginosus pols η, ι, and κ (Figure 4A). As with the human enzymes, T. lanuginosus polη and polκ were much more processive than pol and incorporated ~6 and ~35 bases, respectively, in a single binding event after only 1 min. In contrast, pol incorporated ~10–11 bases in a distributive manner. It then exhibited a strong pause after encountering two adjacent template dTs located 11 and 12 nucleotides from the 5′ end of the primer. Presumably this is due to misincorporation of dG opposite the dTs and its subsequent poor extension (Vaisman et al., 2001). These results indicate that the processivity of T. lanuginosus pols η, ι, and κ are therefore similar to their human counterparts (Washington et al., 1999; Ohashi et al., 2000a; Tissier et al., 2000b).

**Thermostability of T. lanuginosus Polymerases**

*Thermomyces lanuginosus* is a thermophilic fungus with a natural habitat of growing in organic soils, e.g., decomposing vegetable matter, composts, and animal excrements, at an optimal temperature between 48 and 52°C (Tsiklinsky, 1899; Maheshwari et al., 2000; Singh et al., 2003). The ability of this fungus to live in such an environment suggests that its proteins possess enhanced thermostability (Maheshwari et al., 2000). We therefore examined the optimum temperature for the *Thermomyces lanuginosus* pols η, κ, and ι by incubating primer extension reactions at temperatures ranging between 25 and 75°C. All three polymerases exhibited robust activity between 25 and 50°C, but there was sharp decline in activity at temperatures greater than 50°C (Figure 4B).

![Polη, Polι, and polκ-dependent Lesion Bypass Past Abasic Sites, CPDs, and benzo[a]Pyrene Adducts](Image)

The Y-family polymerases are best characterized by their ability to catalyze translesion replication past a variety of DNA lesions, in a manner distinct from the cell’s conventional replicases (Sale et al., 2012; Vaisman and Woodgate, 2017). We were therefore...
interested in assessing the ability and fidelity of *T. lanuginosus* pols η, κ, ι and to carry out translesion synthesis across a *cis*-syn CPD, an abasic site, and a benzo[α]pyrene adduct.

Apurinic/apyrimidinic (or abasic) sites are generated as a result of spontaneous or enzymatically-induced base loss and are considered the most frequent form of DNA damage (Lindahl, 1993; Boiteux and Guillet, 2004). An estimated 10,000 abasic sites occur per human cell per day under physiological conditions (Loeb, 1989). It is therefore not surprising that the ability of polη to copy DNA templates containing this noninstructional lesion has been extensively investigated. What is surprising, is that studies performed by various groups with human and *S. cerevisiae* polη produced conflicting results, ranging from those stating that “hPolη has the highest abasic lesion bypass efficiency” among human Y-family DNA pols and “is a major pol involved in abasic site bypass” (Choi et al., 2010; Sherrer et al., 2011; Patra et al., 2015; Thompson and Cortez, 2020) to those which dismiss human, or *S. cerevisiae* polη as contributing to *in vivo* abasic site bypass because of “negligible nucleotide insertion opposite the abasic site and primer extension past the lesion” (Haracska et al., 2001; Pages et al., 2008). Somewhere in between these polar extremes are studies suggesting that polη catalyzes insertion and/or extension steps of abasic site bypass efficiently enough to ensure its involvement in TLS *in vivo*, but not as a major player (Zhang et al., 2000a; Masutani et al., 2000; Zhao et al., 2004; Sherrer et al., 2011).

We show here that the efficiency and fidelity of *T. lanuginosus* polη during TLS of abasic sites is reminiscent of properties of human and yeast polη uncovered in the third group of studies described above (Zhang et al., 2000a; Masutani et al., 2000; Zhao et al., 2004). In particular, in standing-start primer extensions *T. lanuginosus* polη inserted dA and dG opposite the abasic site with moderate efficiency, but further primer elongation was largely obstructed (Figure 5A). However, minuscule extension of the primer with terminal dA opposite the abasic site occurred by incorporation of an additional dA when the base 5′ to the abasic site was dT (Figure 5A, incorporation of dA). It should be noted that base selectivity of polη on the template with an abasic site is similar to its moderate fidelity on the undamaged DNA. As shown in Figure 5A, polη preferentially incorporates a correct dA opposite the undamaged dT although substantial misinsertion of dG and low level of dC and dT misincorporation can also be seen.

As with polη, there is no consensus regarding the ability of the DinB orthologs (polκ) to utilize DNA templates containing abasic sites (Strauss, 1991; Ohashi et al., 2000b; Zhang et al., 2000b; Boudsocq et al., 2001; Gerlach et al., 2001; Choi et al., 2010; Sherrer et al., 2011), which is at least partly explained by the strong effect of the surrounding sequence context on the bypass mechanism, base selectivity, and efficiency of insertion and extension steps of TLS. Thus, it has been shown that human polκ tolerates an abasic site much better when the next (5′) template base is dT, and the nucleotide preferentially incorporated opposite the lesion is dA. Furthermore, when primer/template slippage is possible, bypass often occurs through a stabilized misalignment mechanism leading to 1- or 2-base deletions (Boudsocq et al., 2001; Ling et al., 2004a).

Similarly, *T. lanuginosus* polκ-catalyzed TLS of the abasic site was quite efficient when the 5′ template was dT and dA was preferentially incorporated opposite the lesion, although the other three nucleotides were also inserted at substantial levels (Figure 5B). The low fidelity of polκ during abasic site bypass is very similar to its highly error-prone nucleotide incorporation on undamaged DNA. The fact that the major product of *T. lanuginosus* polκ primer extension reactions in the presence of dT was generated by addition of four nucleotides is consistent with the aforementioned human polκ template-slippage mechanism (Mukherjee et al., 2013).

Therefore, the resulting main product would contain two mismatches and total primer elongation by four bases because the DNA template used in the current study has two dA’s next to dT. When the DNA template was not susceptible to primer/template slippage (reactions with dC, dG, and dA) only one wrong nucleotide was incorporated opposite the abasic site.

Human polη has the capacity to insert deoxynucleotides opposite an abasic site, however further extension is limited (Zhang et al., 2000c; Vaisman et al., 2002; Nair et al., 2009; Choi et al., 2010; Sherrer et al., 2011). *T. lanuginosus* polη behaved in a similar manner, i.e., a single nucleotide was readily inserted opposite the abasic site with no indication of further extension (Figure 5C). Furthermore, akin to human polη, *T. lanuginosus* polη incorporates dG, dA, and dT with similar efficiency. Interestingly, when undamaged templates were used, *T. lanuginosus* polη was even more inaccurate than the human polymerase, *i.e.*, not only dG, but also dT, was inserted opposite template dT more often than the correct dA (Figure 5C).

The most common ultraviolet light photoproduce is a *cis*-syn cyclobutane pyrimidine dimer (CPD). Unlike the ambiguity of TLS of an abasic site, there is no question as to which TLS polymerase bypasses a CPD most efficiently. Indeed, it is well established that polη can bypass a CPD with the same efficiency and accuracy as undamaged template dT’s (Johnson et al., 2000c; Masutani et al., 2000; McCulloch et al., 2004). Most other pols halt synthesis either immediately before the lesion, or opposite the first template dT of the CPD, as their active site is unable to accommodate the covalently joined dT’s of the CPD. Similar to human and *S. cerevisiae* polη, *T. lanuginosus* polη efficiently bypassed the CPD lesion (Figure 5A). Like polymerases purified from other species, *T. lanuginosus* polη was even more accurate while inserting nucleotides opposite the damaged bases than on undamaged DNA templates (Figure 5A).

The pols in the DinB branch of the Y-family (polκ) are unable to replicate past UV-induced *cis*-syn CPD dimers (Johnson et al., 2000a; Ohashi et al., 2000b; Zhang et al., 2000b) with the one reported exception of archaeal *Sulfolobus solfataricus* Dp04, whose properties resemble eukaryotic polη (Boudsocq et al., 2001). Interestingly, the ability of *T. lanuginosus* polκ to copy a CPD-containing template resembles the behavior of its human orthologs rather than the thermophilic archaeal enzyme, *i.e.*, *T. lanuginosus* polκ was virtually blocked by the CPD lesion in reactions with all four dNTPs although a very low level of the
correct incorporation of two dAs opposite the damaged bases was detected in reactions where only dATP was present (Figure 5B).

It has been reported that human polι can bypass a CPD. Depending on the sequence context, it incorporates anywhere between 1–5 nucleotides, albeit with low efficiency (Tissier et al., 2000a; Vaisman et al., 2003). T. lanuginosus polι was much less efficient compared to the human enzyme. In fact, CPD bypass was barely detectible (Figure 5C). As with undamaged template dT, the preferred nucleotide inserted opposite the 3′-dT of the CPD was dG (Figure 5C).

Benzopyrene diol epoxides (BPDEs) are the highly carcinogenic metabolites of the environmental pollutant benzo[a]pyrene, found in tobacco smoke and automobile exhaust. In mammalian cells, various BPDE stereoisomers covalently bind to DNA forming adducts at the N2 and N6 position of guanine and adenine, respectively. Previous studies demonstrated that efficiency and fidelity of TLS past BPDE by different pols not only depends on the DNA base to which the adduct is linked, but is also often modulated by the stereochemistry of the adduct (Zhang et al., 2002a; Chiapperino et al., 2002; Frank et al., 2002; Rechkoblit et al., 2002; Suzuki et al., 2002; Jia et al., 2008). For example, in general, human polη is almost completely stalled by BPDE-dA adducts. Under similar reaction conditions, it accurately and efficiently bypasses most of the BPDE-dG stereoisomers (Zhang et al., 2002a; Rechkoblit et al., 2002; Suzuki et al., 2002; Jia et al., 2008), although the cis-S-BPDE-dG represents a significant block for the polymerase (Suzuki et al., 2002). In contrast, polι is much more efficient and accurate while incorporating nucleotides opposite BPDE-dA than opposite BPDE-dG (Frank et al., 2002). Furthermore, independent of adduct stereochemistry, the correct dT is inserted by polι equally efficiently opposite various BPDE-dA isomers. On the other hand, the ability of human polη to replicate past the BPDE is strongly affected by the adduct conformation and damaged base, i.e., it readily traverses through the trans-S-BPDE-dA, but not through the isomeric trans-R- adduct, or through the BPDE-
dG isomers (Chiapperino et al., 2002; Rechkoblit et al., 2002). In the current study with T. lanuginosus pols, we used DNA templates with a trans-S-BPDE-dA adduct. Reactions with polη and polκ contained Mg\(^{2+}\) or Mn\(^{2+}\), whereas polι was tested only in the presence of Mn\(^{2+}\). Similar to the human enzyme, T. lanuginosus polη was quite efficient, but rather inaccurate, while incorporating nucleotides opposite the trans-S-BPDE-dA (Figures 6A,C) especially when reactions were carried out in the presence of Mn\(^{2+}\) which significantly boosted TLS efficiency without noticeably reducing fidelity (Figure 6C). Nucleotide incorporation opposite the lesion by polι was at least as efficient as by polη (Figure 6E), but it was less accurate and further primer extension was inhibited (Figure 6E). We have previously found that fidelity of human polι was influenced by the local sequence context of the BPDE-dA lesion (Frank et al., 2002), and when the template bases 5' to the lesion were GAC, uncharacteristically accurate insertion of dT opposite the lesion occurred with similar efficiency, making synthesis by polι highly error-prone (Figure 6D).

While polι itself was incapable of extending primers by more than one nucleotide after insertion opposite the BPDE-dA even in the presence of Mn\(^{2+}\), it was able to assist polκ to bypass the lesion and elongate primers to the end of the template (Figure 6F). Nevertheless, TLS catalyzed by the combined action of polι and polκ was still less efficient than by polη alone (Figure 6F).

Characterization of T. lanuginosus Rev1
S. cerevisiae and human Rev1 function as a dCMP transferase, incorporating strictly dCMP opposite abasic sites and dG-lesions (Nelson et al., 1996; Lin et al., 1999; Masuda et al., 2001). dCMP transferase activity is instructed by Rev1’s conserved active site side chain arginine residue via a protein-templated mechanism (Nair et al., 2005; Weaver et al., 2020). We hypothesized that the conservation of the functional arginine and Rev1 consensus sequence (Figure 7A)
would make *T. lanuginosis* Rev1 a bonafide dCMP transferase. Therefore, the DNA synthesis activity of *T. lanuginosus* Rev1 was characterized in vitro opposite damaged and undamaged DNA. We found that *T. lanuginosus* Rev1 could utilize both Mn\(^{2+}\) and Mg\(^{2+}\), but reasoned that Mg\(^{2+}\) was a more likely candidate for physiological utilization given its wider efficacy in comparison to Mn\(^{2+}\) (Figure 7B). Similar to the activities of *S. cerevisiae* and human Rev1, DNA synthesis by *T. lanuginosus* Rev1 was not processive, whereby it inserted only once opposite a templating dG and stalled at the following dT base (Figure 7C). As expected, *T. lanuginosus* Rev1 had a similar thermostability to the other *T. lanuginosus* Y-family pols (Figure 7D). The well-characterized dCMP transferase activity of Rev1 ortholog was evaluated opposite undamaged, as well as abasic site- and CPD-containing templates (Figure 7E). The ability for *T. lanuginosus* Rev1 to perform DNA synthesis was efficient only in the presence of dCMP opposite undamaged, or abasic site templates, akin to *S. cerevisiae* and human REV1 (Zhang et al., 2002b; Nair et al., 2011; Weerasooriya et al., 2014). In summary, we found that the DNA synthesis properties of *T. lanuginosus* Rev1 are comparable with its higher eukaryotic ortholog.

**CONCLUDING REMARKS**

In this paper, we describe the cloning, purification, and biochemical characterization of Y-family pols from the thermophilic fungus *T. lanuginosus*. *T. lanuginosus* contains all four eukaryotic Y-family orthologs that each generally demonstrate biochemical properties similar to the properties of higher eukaryotic orthologs with the exception of the expected thermostability of the *T. lanuginosus* Y-family polymerases. Interestingly, while *T. lanuginosus* poli exhibited biochemical properties akin to the human enzyme, its metal ion requirement was significantly different. Whereas human poli is activated by very low levels of Mn\(^{2+}\) (250 \(\mu\)M) and inhibited by high levels of Mg\(^{2+}\) (>1 mM) (Frank and Woodgate, 2007), *T. lanuginosus* poli required 4–8 mM Mn\(^{2+}\) or Mg\(^{2+}\) respectively, for maximal activity in vitro.

**DATA AVAILABILITY STATEMENT**

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

**AUTHOR CONTRIBUTIONS**

Conceptualization: JM and RW. Data curation: AV, JM, MS, and SA. Formal analysis: AV, JM, and MS. Funding acquisition: RW. Investigation: AV, JM, MS, and SA. Visualization: AV, JM, and MS. Writing—original draft, Review and editing: AV, JM, MS, SA, TE, and RW.

**FUNDING**

Funding for this work was provided by the National Institute of Child Health and Human Development/National Institutes of Health Intramural Research Program (RW).

**ACKNOWLEDGMENTS**

We would like to thank Dominic Quiros (NIH) for assistance with the early exploratory phase of Rev1 purification. We also thank Francine B. Perler for her work on purifying *T. lanuginosus* poli.

The conservation of Y-family polymerase activities between lower and higher eukaryotic organisms suggests that TLS activity may be more essential to cellular viability than previously appreciated. Further structure-function studies of these polymerases, as well as identification and characterization of other Y-family orthologs from yet unsequenced species would help clarify the seemingly random ortholog dispersion among eukaryotes and how their presence/absence might provide evolutionary advantages in the face of different environmental stressors.

**REFERENCES**

Avkin, S., Goldsmith, M., Velasco-Miguel, S., Geacintov, N., Friedberg, E. C., and Livneh, Z. (2004). Quantitative Analysis of Translesion DNA Synthesis across a Benzo[α]pyrene-Guanine Adduct in Mammalian Cells. *J. Biol. Chem.* 279, 53298–53305. doi:10.1074/jbc.M409155200

Basu, A. K., Pande, P., and Bose, A. (2017). Translesion Synthesis of 2′-Deoxyguanosine Lesions by Eukaryotic DNA Polymerases. *Chem. Res. Toxicol.* 30, 61–72. doi:10.1021/acs.chemrestox.6b00285

Bessman, M. J., Lehman, I. R., Simms, E. S., and Kornberg, A. (1958). Enzymatic Synthesis of Deoxyribonucleic Acid. II. General Properties of the Reaction. *J. Biol. Chem.* 233, 171–177. doi:10.1016/s0021-9258(19)68049-x

Boiteux, S., and Guillet, M. (2004). Abasic Sites in DNA: Repair and Biological Consequences in *Saccharomyces cerevisiae*. *DNA Repair* 3, 1–12. doi:10.1016/j.dnarep.2003.10.002

Boudsocq, F., Iwai, S., Hanaoka, F., and Woodgate, R. (2001). *Sulfolobus solfataricus* P2 DNA Polymerase IV (Dpo4): an Archaeal DinB-like DNA Polymerase with Lesion-Bypass Properties Akin to Eukaryotic Pol. *Nucleic Acids Res.* 29, 4607–4616. doi:10.1093/nar/29.22.4607

Boudsocq, F., Kokoska, R. J., Plosky, B. S., Vaisman, A., Ling, H., Kunkel, T. A., et al. (2004). Investigating the Role of the Little Finger Domain of Y-Family DNA Polymerases in Low Fidelity Synthesis and Translesion Replication. *J. Biol. Chem.* 279, 32932–32940. doi:10.1074/jbc.m405249200

Chiaparino, D., Kroth, H., Kramarczuk, I. H., Sayer, J. M., Masutani, C., Hanaoka, F., et al. (2002). Preferential Misincorporation of Purine Nucleotides by Human DNA Polymerase η Opposite Benzo[α]pyrene 7,8-Diol 9,10-Epoxide Deoxyguanosine Adducts. *J. Biol. Chem.* 277, 11765–11771. doi:10.1074/jbc.m112139200

Choi, J.-Y., Angel, K. C., and Guengerich, F. P. (2006). Translesion Synthesis across Bulky N2-Alkyl Guanine DNA Adducts by Human DNA Polymerase δ. *J. Biol. Chem.* 281, 21062–21072. doi:10.1074/jbc.M602246200

Choi, J.-Y., Lim, S., Kim, E.-J., Jo, A., and Guengerich, F. P. (2010). Translesion Synthesis across Abasic Lesions by Human B-Family and Y-Family DNA Polymerases α, δ, η, λ, κ, and REV1. *J. Mol. Biol.* 404, 34–44. doi:10.1016/j.jmb.2010.09.015
Sale, J. E., Lehmann, A. R., and Woodgate, R. (2012). Y-family DNA Pol
276, 30615. doi:10.1016/j.dnarep.2020.102914
274, 7099–7108. doi:10.1128/mcb.20.19.7099-7108.2000
Yang, Z., Yuan, F., Wu, X., Rechkoblit, O., Taylor, J. S., Geacintov, N. E., et al.
274, 20869–20875. doi:10.1016/j.rbc.2016.02.049
286, 7099–7108. doi:10.1128/mcb.20.19.7099-7108.2000
Yang, W., and Gao, Y. (2018). Translesion and Repair DNA Polynucleases: Diverse Structure and Mechanism. Annu. Rev. Biochem. 87, 239–261. doi:10.1146/annurev-biochem-062917-012405
Yang, Z., Yuan, F., Wu, X., Rechkoblit, O., Taylor, J. S., Geacintov, N. E., et al.
286, 7099–7108. doi:10.1128/mcb.20.19.7099-7108.2000
Yang, Z., Yuan, F., Wu, X., and Wang, Z. (2000c). Preferential Incorporation of G
274, 20869–20875. doi:10.1016/j.rbc.2016.02.049
286, 20869–20875. doi:10.1016/j.rbc.2016.02.049
Washington, M. T., Johnson, R. E., Prakash, S., and Prakash, L. (1999). Fidelity and Processivity of Saccharomyces cerevisiae DNA Polymerase η. J. Biol. Chem. 274, 36835–36838. doi:10.1074/jbc.274.52.36835
Washington, M. T., Johnson, R. E., Prakash, S., and Prakash, L. (1999). Fidelity and Processivity of Saccharomyces cerevisiae DNA Polymerase η. J. Biol. Chem. 274, 36835–36838. doi:10.1074/jbc.274.52.36835
Washington, M. T., Johnson, R. E., Prakash, S., and Prakash, L. (1999). Fidelity and Processivity of Saccharomyces cerevisiae DNA Polymerase η. J. Biol. Chem. 274, 36835–36838. doi:10.1074/jbc.274.52.36835
Weaver, T. M., Cortez, L. M., Khoang, T. H., Washington, M. T., Agarwal, P. K., and Freudenthal, B. D. (2020). Visualizing Rev1 Catalyze Protein-Template DNA Synthesis. Proc. Natl. Acad. Sci. USA 117, 25494–25504. doi:10.1073/pnas.2010484117
Weerasooriya, S., Jasti, V. P., and Basu, A. K. (2014). Replicative Bypass of Abasic Site in Escherichia coli and Human Cells: Similarities and Differences. PLoS One 9, e107915. doi:10.1371/journal.pone.0107915
Weaver, T. M., Cortez, L. M., Khoang, T. H., Washington, M. T., Agarwal, P. K., and Freudenthal, B. D. (2020). Visualizing Rev1 Catalyze Protein-Template DNA Synthesis. Proc. Natl. Acad. Sci. USA 117, 25494–25504. doi:10.1073/pnas.2010484117
Washington, M. T., Johnson, R. E., Prakash, S., and Prakash, L. (1999). Fidelity and Processivity of Saccharomyces cerevisiae DNA Polymerase η. J. Biol. Chem. 274, 36835–36838. doi:10.1074/jbc.274.52.36835