THE EXPANDING POLYMERASE UNIVERSE

Myron F. Goodman and Brigette Tippin

Over the past year, the number of known prokaryotic and eukaryotic DNA polymerases has exploded. Many of these newly discovered enzymes copy aberrant bases in the DNA template over which 'respectable' polymerases fear to tread. The next step is to unravel their functions, which are thought to range from error-prone copying of DNA lesions, somatic hypermutation and avoidance of skin cancer, to restarting stalled replication forks and repairing double-stranded DNA breaks.

DCMP TRANSFERASE
The DNA template-directed reaction catalysed by the yeast Rev1 protein, where C is favoured for incorporation opposite an abasic template site, and to a much lesser extent opposite a normal G site.

PROOFREADING
Excision of a misincorporated nucleotide at a growing 3′-primer end by a 3′ exonuclease associated with the polymerase.

T–T 6–4 PHOTOPRODUCT
A form of damage occurring when DNA is exposed to UV radiation, in which a covalent bond is formed between the 6 and 4 pyrimidine ring positions, coupling adjacent thymines on the same DNA strand.

University of Southern California, Department of Biological Sciences and Chemistry, Stauffer Hall of Science 172, Los Angeles, California 90089-1340, USA.
e-mail: mgoodman@rcf.usc.edu

In the beginning there was just one DNA polymerase — *Escherichia coli* DNA polymerase I (pol I), discovered by Arthur Kornberg and colleagues[1,2] in 1956. Thirteen years later, Paula de Lucia and John Cairns, at Stony Brook, New York, isolated an *E. coli* mutant, *polA* (its designation being a play on de Lucia's first name, as proposed to Cairns by Julian Gross) that seemed to have less than 1% of the normal pol I activity[3]. From this strain, a new DNA-polymerizing enzyme, pol II, was isolated[4].

The *polA* strain was much more sensitive to ultraviolet (UV) radiation than wild-type cells, suggesting that pol I might be involved in DNA repair in addition to chromosomal replication. Shortly after, using this same *polA* strain, Thomas Kornberg and Malcolm Gefter[5], and Friedrich Bonhoeffer, Heinz Schaller and colleagues[6], independently discovered DNA polymerase III (pol III). Isolation of a conditionally lethal temperature-sensitive pol III mutant[7] showed that this enzyme is required for replicating the *E. coli* chromosome[6]. In contrast, pol II remained an enigma until last year, when it was shown[8] to be pivotal in restarting replication in UV-irradiated cells.

Last year also saw the identification of a new class of DNA polymerases — the UmuC/DinB/Rev1p/Rad30 superfamily (TABLE 1) — on the basis of five conserved sequence motifs present in all of these proteins (FIG. 1). The yeast Rev1 protein had been shown to contain DNA-template-dependent DCMP TRANSFERASE activity nearly three years earlier[9], but it was not until 1999 that the other family members were isolated and shown to be capable of replicating DNA using all four bases. Biological functions have been established for some members, including the *E. coli* UmuD′$_2$C complex (now known as pol V), the yeast Rev1 protein and human DNA polymerase eta (pol η/Rad30). However, the functions of the remaining members of the UmuC/DinB/Rev1p/Rad30 polymerase superfamily are less certain.

A feature common to many of these polymerases is their tendency to copy undamaged DNA with remarkably poor fidelity, whether or not they are involved in translesion synthesis. As its name suggests, translesion synthesis is the unimpaired copying of aberrant bases (see below) at which other cellular polymerases stall. With undamaged DNA, these low-fidelity polymerases incorporate an incorrect nucleotide once every 100–1,000 bases on average[10–12] (TABLE 1). For comparison, normal polymerases that do not PROOFREAD misincorporate nucleotides in the range of once every 10^4–10^6 bases[13]. Examples of low-fidelity polymerases include *E. coli* pol V, which preferentially misincorporates G opposite a 3′ T of a T–T 6–4 PHOTOPRODUCT; *E. coli* DNA polymerase IV (pol IV/DinB), which adds a nucleotide onto the end of a misaligned primer; Rev1p, which incorporates C opposite a non-coding ABASIC LESION; and human DNA polymerase iota (pol ι/Rad30B), which favours misincorporation of G opposite T on undamaged DNA. All of these events lead to mutation. There is also the remarkable case of pol η, which copies pyrimidine T–T DIMERS accurately, resulting in mutation avoidance at this type of DNA damage (FIG. 2).

The number of DNA polymerases has now grown from 3 to 5 in *E. coli*, and from 5 to at least 14 and count-

ABASIC LESION
A common form of DNA damage in which a base is lost from a strand of DNA, spontaneously or by the action of DNA repair enzymes such as apurinic endonucleases or uracil glycosylase, while leaving the phosphodiester bond intact.

T–T DIMER
A form of damage occurring when DNA is exposed to UV radiation, in which two covalent bonds are formed between both the 5 and 6 positions of the pyrimidine ring on adjacent thymines located on the same DNA strand.

NUCLEOTIDE-EXCISION REPAIR
The main pathway for removal of UV-damaged bases.

REPLICATION FORK
Site in double-stranded DNA at which the template strands are separated, allowing a newly formed copy of the DNA to be synthesized, with the fork moving in the direction of leading strand synthesis.

DISTRIBUTIVE POLYMERASE
A polymerase that dissociates from the primer–template DNA after incorporating one (or at most a few) nucleotides.

ing in eukaryotes (TABLE 1). Indeed, in a 'back to the future' moment during a recent conversation with Bob Lehman, Arthur Kornberg remarked, "In 1955, who would have imagined that there could be five DNA polymerases in *E. coli*?". So what were the events that led to the discovery of these polymerases, and what do we now know of their biochemical functions and cellular properties? And why are there so many of them in eukaryotic cells? Whereas prokaryotic cells have just one choice — replicate damaged DNA or die — eukaryotic cells can, in principle, use programmed cell death (apoptosis) as an 'escape hatch' to avoid a potential catastrophe.

A growing family

Genetic studies in *Saccharomyces cerevisiae* and *E. coli* have been instrumental in defining groups of proteins required for mutagenesis. For example, yeast lacking the *REV3* (REF. 14), *REV7* (REFS 15,16) or *REV1* (REF. 17) genes show significantly decreased spontaneous and UV-induced mutation rates. In *E. coli*, SOS mutagenesis (BOX 1) requires the *umuC* and *umuD* genes[18].

In 1968, Dean Rupp and Paul Howard-Flanders[19] observed discontinuities (daughter-strand gaps) in DNA synthesized in an excision-defective strain of *E. coli* after UV irradiation. Because these strains cannot carry out NUCLEOTIDE-EXCISION REPAIR[20], Rupp and Howard-Flanders suggested that a single pyrimidine dimer is enough to kill the cell, presumably by blocking DNA replication. But although it may be advantageous to copy a variety of template lesions as an alternative to cell death, there is no such thing as a free lunch. The associated cost of survival is an increased number of mutations, targeted at the lesion sites. In *E. coli*, this is referred to as UV-induced SOS error-prone repair (BOX 1).

Because both *E. coli* and yeast were known to have three DNA polymerases, there was no reason to suspect that there is a special class of polymerases to copy damaged DNA templates. Instead, proteins such as UmuC and UmuD′ (the mutagenically active form of UmuD; BOX 1) were thought to reduce the fidelity of *E. coli* pol III, enabling a blocked REPLICATION FORK to carry out error-prone translesion synthesis[21]. We now know that this is not the case — indeed, the cellular function of several new, errant DNA polymerases is translesion DNA synthesis[22–25] (TABLE 1). Further studies using these 'sloppier copier' DNA polymerases are now revealing a rich biochemical tapestry. For example, one member of this family — the *E. coli* UmuD′$_2$C polymerase (pol V)[26–28] — does not act alone, but requires three further proteins to catalyse translesion synthesis[12,28]. Moreover, yeast Rev1p requires DNA polymerase zeta (pol ζ) to copy past abasic sites[9] and T–T 6–4 photoproducts[29]. Other family members probably also use accessory proteins.

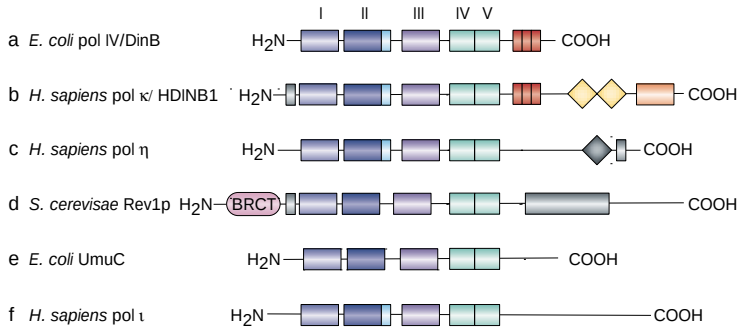


Figure 1 | **Representative members of the UmuC/DinB/Rev1p/Rad30 superfamily.** Five highly conserved domains (indicated by roman numerals I–V) are believed to contain the nucleotide binding and catalytic residues. The subgroups within the family can be easily distinguished by the presence or absence of unique domains. **a**, **b** | The DinB subgroup contains a further three small domains near the carboxyl terminus of the protein (red boxes), whereas zinc finger motifs are uniquely found in **b** | pol κ/HDINB1 (C_2HC type, yellow diamonds) and **c** | pol η (C_2H_2 type, dark grey diamond) that may be involved in DNA binding and selective targeting. **d** | Rev1p is the longest member of the family and contains two regions that are only conserved within the Rev1p subgroup (light grey boxes) as well as a BRCT (BRCA-1 carboxy-terminal) domain believed to mediate protein–protein interactions for cell cycle checkpoints and DNA repair. **e** | UmuC and **f** | pol ι are both characterized by unique carboxy-terminal ends in which no known functional domains have been identified. These unique regions could possibly mediate protein interactions that stimulate and target UmuC or pol ι to their cellular destinations.

The *E. coli* pol V mutasome

A good place to start any tour of the new DNA polymerases is with the *E. coli* pol V mutasome. The DNA-damage-inducible SOS response in *E. coli* was discovered more than 25 years ago (BOX 1). Many of the 30 or more SOS-regulated genes are involved in repairing DNA damage[30,31], but two genes, *umuC* and *umuD*, are instead required for SOS-induced mutagenesis[18,32–34]. Although SOS mutation rates are typically 100-fold higher than spontaneous rates[31], increased mutagenesis cannot occur unless UmuD is first converted (by cleavage) to the mutagenically active UmuD′ protein in a reaction that depends on another SOS-induced protein, RecA (REF. 35). The UmuC and UmuD′ proteins then interact to form a tight complex[36,37], UmuD′$_2$C (pol V)[27,28], with intrinsic DNA polymerase activity.

Working alone on an undamaged primed DNA template, pol V is a poor DISTRIBUTIVE POLYMERASE[26,27]. However, pol V cannot copy damaged DNA by itself — it requires RecA, single-stranded DNA binding protein (SSB) and β/γ complex[12,27] (where β is the PROCESSIVITY CLAMP and γ the CLAMP-LOADER component of the replicative pol III holoenzyme). This multiprotein system, consisting of pol V, RecA, SSB and β/γ, is called the pol V "mutasome" (FIG. 3), a term coined by Harrison Echols[38]. The specific activity of pol V is amplified by an extraordinary 15,000-fold in the presence of RecA-coated template, allowing it to copy past damaged DNA bases[12].

Although the SOS system typically introduces mutations at sites of DNA damage, there is also an increase in untargeted mutations in the absence of damage[39]. All three common forms of DNA damage (FIG. 2) are copied efficiently by the pol V mutasome, but synthesis by either the pol III holoenzyme or pol IV (DinB) is blocked[12]. The specificity of incorporation by the pol V mutasome opposite the three forms of lesion mimics the *in vivo* mutational data[12]. For example, the 3′ T of a T–T 6–4 photoproduct is a T→C mutational 'hotspot' caused by the misincorporation of G opposite T (FIG. 2b) — precisely the reaction favoured by the pol V mutasome[12]. In contrast, pols III and IV preferentially incorporate A, which agrees with the 'A-rule'[40], but not with the *in vivo* data.

What is the mechanism of translesion synthesis by

Table 1 | The expanding polymerase universe

DNA polymerase	Error rate	Properties	Possible function
UmuC/DinB/Rev1p/Rad30 superfamily			
E. coli pol V (*umuDC*)[12,27,28,99]	10^{-2}–10^{-3}	Translesion synthesis Low-fidelity synthesis	SOS lesion-targeted and untargeted mutagenesis
E. coli pol IV (*dinB*)[12,46] *H. sapiens* pol κ/θ (*HDINB1*)[49,50,100]	10^{-3}–10^{-4}	Mismatch extension	Untargeted and lesion-targeted mutagenesis Rescues stalled replication forks
H. sapiens pol η (*XPV*)[11,61,62] *S. cerevisiae* pol η (*RAD30*)[10]	10^{-2}–10^{-3}	Error-free synthesis of T–T UV photodimers	Prevents sunlight-induced skin cancer xeroderma pigmentosum
H. sapiens pol ι (*HRAD30B*)[59,79]	10^{-1}–10^{-4}	Low-fidelity synthesis	Somatic hypermutation
H. sapiens Rev1 (*HREV1*)[56,57] *S. cerevisiae* Rev1 (*REV1*)[9,29]		Incorporation of C opposite abasic sites	UV mutagenesis
Family B			
S. cerevisiae pol ζ (*REV3/REV7*)[58,59] *H. sapiens* pol ζ (*HREV3/HREV7*)[55,66,67,101]	10^{-4}–10^{-5}	Mismatch extension at lesions	UV mutagenesis
Family X			
H. sapiens pol λ/pol β2 (*POLL/POLβ2*)[82,83]			Meiosis-associated DNA repair
H.sapiens pol μ (*POLM*)[81]			Somatic hypermutation
S. cerevisiae pol κ (*TRF4*)[84]*			Sister-chromatid cohesion
Family A			
H. sapiens pol θ (*POLQ*)[85]* *D. melanogaster MUS308*[86]			Repair of DNA crosslinks

* The HDINB1 polymerase has also been designated pol κ/pol θ.

pol V? The key to arriving at quantitative, kinetic-based conclusions for the effects of RecA, SSB and β/γ on pol V-catalysed lesion bypass is to have the pol V mutasome bound in a confined region just before the lesion[41]. Two interactions occur: the first is between pol V and RecA; the second is between pol V and SSB (FIG. 3). Assembly of a RecA filament requires ATP binding (it proceeds 5′→3′ along a single-stranded DNA template in the presence of ATP or a poorly hydrolysable analogue, ATPγS). But disassembly of the filament in the same direction requires ATP hydrolysis[42]. So in the presence of ATPγS, RecA is bound stably to DNA as a helical filament. The other polymerases (pols II, III or IV) cannot copy DNA in the form of a RecA filament, even if the template is undamaged. Remarkably, however, pol V, along with SSB and β/γ, copies damaged and undamaged stabilized filaments with high processivity[12,41], perhaps providing the key to unlock the lesion-copying mechanism.

The RecA filament is 100 Å in diameter, whereas the β-clamp has an inner diameter of only 35 Å. But processive synthesis takes place on the filament. The obvious explanation is that pol V, acting in conjunction with SSB, strips RecA off the DNA in a 3′→5′ direction — a 100 Å RecA molecule cannot be threaded through the eye of a 35 Å β-dimer 'needle'[41]. The stripping process is akin to the action of a locomotive 'cowcatcher' (a pointed device attached to the front of trains to push obstacles off the track). In this case, the RecA 'cow' is pushed off the DNA template ahead of the advancing pol V–SSB 'locomotive'[41]. We have recently proposed that bidirectional disassembly of the RecA filament, driven in the 3′→5′ direction by pol V–SSB and in the opposite direction by ATP hydrolysis, confines SOS mutations to the sites of DNA damage[41], although SOS untargeted mutations do occur, albeit at a much lower frequency. After disassembly of the RecA filament and dissociation of pol V, the pol III holoenzyme presumably resumes replication on undamaged DNA downstream from the lesion[43].

The DinB subfamily

Escherichia coli pol IV (DinB) is believed to copy undamaged DNA at stalled replication forks[44], which arise *in vivo* from mismatched or misaligned primer ends that are not proofread. A function for pol IV in alleviating stalling of the pol III holoenzyme is potentially significant, given the estimate that *E. coli* replication forks probably stall at least once during each replication cycle[45]. Overexpression of pol IV results in increased frameshift mutagenesis[44], consistent with the ability of the enzyme to extend misaligned primer termini[46] (FIG. 2). Whereas DinB homologues are among the most conserved members of the UmuC/DinB/Rev1p/Rad30 superfamily, almost nothing is known about what they do in other organisms.

A second function for pol IV has been found, however, in adaptive mutation, a process in which non-proliferating microbial populations accumulate mutations when placed under non-lethal selective pressure[47]. In *E. coli*, pol IV is responsible for roughly half the *lacZ* adaptive frameshift mutations occurring on a plasmid in a wild-type background, and essentially all of the increased frameshifts in the absence of pol II[48]. So muta-

PROCESSIVITY CLAMP
A doughnut-shaped protein complex that threads the DNA through its hole while tethering the polymerase to DNA, typically increasing the processivity of the polymerase (the number of nucleotides incorporated into DNA per polymerase–template binding event).

CLAMP LOADER
A protein complex that binds and then assembles the processivity clamp onto the DNA at a 3′-OH primer end, in a reaction requiring ATP.

GENETIC COMPLEMENTATION GROUPS
A distinct group of genes coding for separate polypeptides (proteins) required in the same biological pathway.

tor polymerases provide flexibility in dealing with environmental stress, particularly in prokaryotic organisms. By investigating competition for survival using *E. coli* strains containing combinations of single, double and triple pol II, pol IV and pol V mutants, it should be possible to determine the contribution of each polymerase to the relative fitness of the organism.

Little is known about the *in vivo* function of human DINB (DNA polymerase kappa, pol κ). Purified pol κ, with a carboxy-terminal truncation (polκΔC) that deletes two zinc clusters (FIG. 1) found only in the higher eukaryote homologues[49], retains its polymerase activity. This implies that the carboxy-terminal region is dispensable for binding and catalysis, but that it mediates protein interactions that target or regulate the enzyme[50]. *In vitro*, polκΔC can bypass an abasic site by preferential insertion of A opposite the lesion, creating a –1 frameshift mutation by a template loop-out mechanism. This occurs when the abasic site is followed by a T in the template; if the abasic site is followed by a template A, only a simple base substitution is observed (FIG. 2). PolκΔC can also synthesize past an *N*-2-acetylaminofluorine (AAF)-modified G in an error-prone manner by preferential incorporation of T, without generating a frameshift mutation[50].

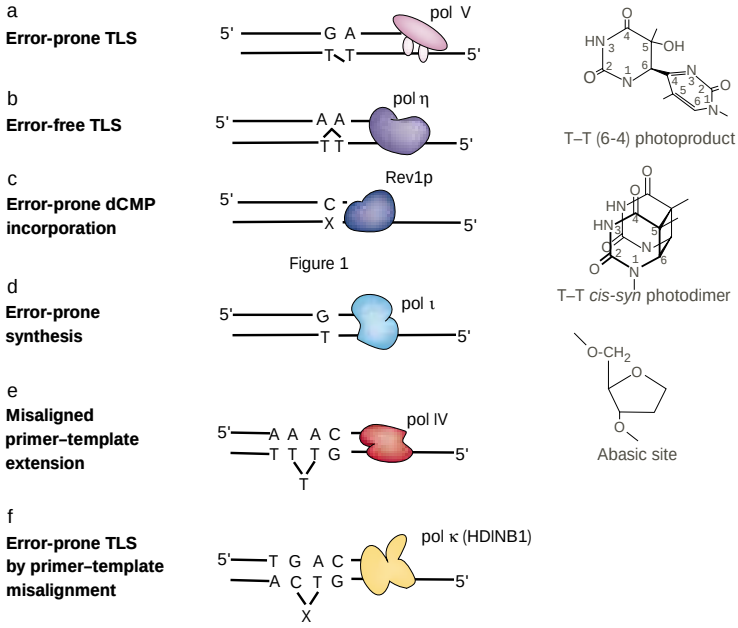


Figure 1

Figure 2 | **Biochemical properties of new polymerases. a** | Error-prone translesion synthesis (TLS) by *E. coli* pol V results in misincorporation of G opposite the 3′ T of a T–T 6–4 photoproduct. **b** | DNA polymerase η incorporates two A bases as it replicates across a T–T *cis–syn* photodimer, thereby avoiding mutation. **c** | The DNA-dependent dCMP transferase activity of Rev1p incorporates C opposite an abasic site. **d** | Misincorporation of G opposite T on an undamaged template is carried out by pol ι, in preference to correct incorporation of A, resulting in a high incidence of A to G transitions. **e** | Misaligned primer–template ends are extended efficiently by *E. coli* pol IV, leading to frameshift errors. Extension of a mismatched primer end (not shown) would lead to a base-substitution mutation. **f** | Bypass of an abasic site by pol κ (HDINB1) results in a –1 frameshift mutation when the lesion is followed by T.

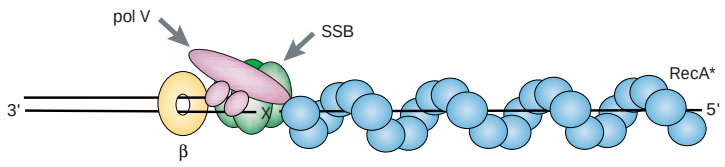


Figure 3 | **The pol V mutasome.** The pol V mutasome consists of pol V ($UmuD'_2C$), activated RecA (RecA*), β sliding clamp, γ clamp loading complex and the single-stranded DNA binding protein (SSB). Pol V associates at a 3′-primer end (vacated by the pol III core), while establishing direct contact with SSB, and the 3′ tip of a RecA filament.

Rev1p C transferase

When confronted with a missing DNA template base — that is, an abasic site (FIG. 2) — most DNA polymerases favour an 'A-rule' default mechanism in which A is strongly preferred (about tenfold) for incorporation opposite the non-templating lesion[40,51–53]. However, when Christopher Lawrence and colleagues[54] used plasmid DNA containing a site-directed abasic moiety to infect yeast cells, they observed preferential incorporation of C, not A (FIG. 2). This unexpected effect depends on the Rev1 protein, which is required for UV mutagenesis[17]. Mutations in *REV1*, *REV3* or *REV7* eliminate more than 95% of base-substitution mutations in yeast[55,56]. The human homologue of yeast Rev1p has since been found to be required for UV-induced mutagenesis[56], and it also behaves as a template-dependent dCMP transferase[57].

There is very little bypass of an abasic site *in vivo* in the absence of Rev1p, and what little bypass does occur obeys the A-rule[29]. Rev1p also shows weak incorporation of C opposite G, at about a tenfold lower rate compared with incorporation opposite an abasic site. Because it incorporates only C, Rev1p is perhaps better characterized as a template-based dCMP transferase rather than a bona fide DNA polymerase[9].

But Rev1p does not act alone in catalysing translesion replication — for this it requires pol ζ (the Rev3 and Rev7 proteins)[58] (FIG. 4). Pol ζ has the remarkable property of adding correct nucleotides onto mismatched 3′-primer ends with exceptionally high efficiencies, only 10–100-fold less than observed for correct primer extension[59]. So it is likely that pol ζ takes over from Rev1p, which incorporates C opposite an abasic lesion but cannot go further[58]. Lawrence and co-workers have also reported[29] that Rev1p is needed to copy past pyrimidine 6–4 photoproducts but, in contrast to bypass of abasic sites, C is not incorporated. Rev1p therefore seems to have two distinct functions in copying DNA damage. One requires its C transferase activity (FIG. 4a), whereas the other facilitates translesion synthesis by another polymerase, most probably pol ζ (FIG. 4b). However, a direct interaction between Rev1p and pol ζ has not been reported.

DNA polymerase η

DNA pol η, a human homologue of the yeast Rad30 protein[60], was identified as the product of the *XPV* gene[61,62] last year. Xeroderma pigmentosum (XP) is characterized by mutations in eight GENETIC COMPLEMENTATION GROUPS, seven of which code for enzymes involved in nucleotide-excision repair[31]. The eighth is the *XPV* gene. Although XPV cells can carry out nucleotide-excision

V(D)J RECOMBINATION
The site-specific recombination of immunoglobulin coding regions from multiple copies in the germ line to just one variable (V), one diversity (D) and one joining (J) region in the process of forming a functional immunoglobulin gene in B cells.

repair, they are deficient in copying UV-damaged DNA[63]. People who carry defects in *XP* genes show increased susceptibility to sunlight-induced skin cancer.

The prevalent form of UV damage to DNA is the T–T *cis–syn* photodimer (FIG. 2). T–T photodimers block replication by various polymerases *in vitro*, but they do not significantly impede human or yeast pol η, both of which copy these photodimers by correctly incorporating two A bases opposite each T site[64,65] (FIG. 2). This is consistent with a cellular role for pol η in the error-free replication of UV photodimers: error-free synthesis when copying T–T photodimers does not translate into error-free synthesis on undamaged DNA. Indeed, error rates for pol η on natural DNA templates can be as high as about 5% for T→C mutations (T•dGMP mispairs)[11], with most base-substitution errors in a range of around 0.5–1% (REFS 10,11). In comparison, most non-proofreading cellular polymerases (which do not have an associated 3′→5′ exonuclease activity for editing out misinserted nucleotides[13]) have error rates of about 10^{-3}–10^{-5}. So relaxed active-site specificity, enabling pol η to copy 'blocking' T–T dimers accurately, is probably responsible for its low fidelity on undamaged DNA.

In XPV cells, error-prone replication of T–T photodimers by some other polymerase could potentially cause an increase in mutations. One candidate is a human homologue of the yeast pol ζ. The yeast polymerase is composed of a complex of the Rev3 and Rev7 proteins, and, as already discussed, is required in an error-prone translesion-synthesis pathway[58]. Indeed, human *REV3* (REF. 66) and *REV7* (Ref. 67) homologues have been identified, but even if human pol ζ is not responsible for error-prone replication in XPV cells, there are plenty of other candidates to choose from (TABLE 1).

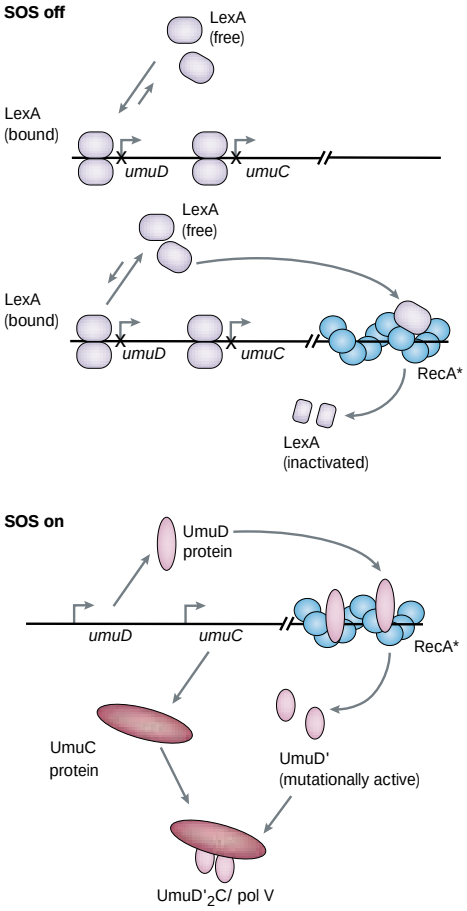
Candidates for somatic hypermutation

The kind of mutagenesis discussed so far is not the only process in which errors can be introduced into DNA. Take somatic hypermutation, for example, which is one of the processes responsible for generating the roughly one billion antibody variants in humans[68]. An initial repertoire of antibodies results from non-random V(D)J RECOMBINATION. After exposure to an antigen, activation of B cells expressing the correct antibody starts a second phase of diversity, termed 'affinity maturation', caused by somatic hypermutation in rapidly dividing GERMINAL CENTRE cells[69]. These mutations occur exclusively in the variable region of the immunoglobulin gene; they begin proximal to the promoter and diminish about 1–2 kilobases downstream[70]. The base-substitution error rate of around 3×10^{-4} per base pair per generation is about 10^6-fold above spontaneous background levels[68]. In other words, somatic hypermutation is exquisitely targeted, and is unaccompanied by a global alteration in the fidelity of B-cell replication.

Two *cis*-acting transcriptional enhancers located downstream of the variable region in light and heavy chains regulate somatic hypermutation[71] (FIG. 5). The intronic enhancer (E_i) and flanking matrix attachment region (MAR) of the κ light chain are both essential, eliminating somatic hypermutation completely when deleted[72]. Another κ light chain 3′ enhancer (E_3') affects mutations to a lesser extent[73]. A promoter sequence upstream of the immunoglobulin gene variable region is also essential, but any promoter can be used, and any DNA inserted into the variable region can act as the mutational target.

Before the discovery of error-prone polymerases, several models for somatic hypermutation were proposed. One suggests that hypermutation rates could arise from repetitive application of transcription-coupled repair, by which stalled replication forks in the

Box 1 | SOS mutagenesis



Escherichia coli keeps an arsenal of regulated pathways that help it to survive when under stress[31]. One of these is the 'SOS' regulon, which is thought to be induced in response to regions of single-stranded DNA — presumably a hallmark of large-scale DNA damage. Normally, the LexA repressor binds to the operators of more than 30 SOS genes and keeps them repressed. But in the presence of single-stranded DNA, the RecA protein forms a close-packed 'activated' RecA filament, RecA*, which acts as a co-protease to cleave any LexA released from low-affinity operators. Further cleavage of LexA frees up the more weakly bound operators, and the SOS genes are relieved from repression (see figure).

The SOS proteins are mainly involved in nucleotide-excision and recombination-repair pathways to remove the DNA damage. However, the two 'UV mutagenesis' (umu) genes, *umuC* and *umuD*, are instead required for replication past unrepaired lesions in the DNA template. They leave behind mutations targeted to sites of DNA damage. To be active, UmuD must be post-translationally cleaved to UmuD′ on the RecA* filament[31,35,97,98] (see figure). UmuC and UmuD′ then form a tight complex, $UmuD'_2C$, which has an intrinsic, low-fidelity DNA polymerase activity[43].

A replication fork blocked by DNA damage is dealt with by two SOS-induced DNA polymerases — pol II and pol V ($UmuD'_2C$). About two minutes after SOS induction, pol II reinitiates replication downstream from the lesion, leaving a gapped structure that is resolved by homologous recombination[43]. Replication restart is an error-free repair process. Pol V appears 30–45 minutes later. It binds at the 3′-OH adjacent to the lesion, then copies past the lesion, often inserting the wrong base opposite it. This process also requires RecA*, single-stranded DNA binding protein (SSB), and the β/γ processivity proteins (FIG. 3).

variable region recruit transcription-coupled repair proteins[74]. Repeated replication of the 'repair' region could increase the chance of mutation. But transcription-coupled repair shows a large bias for the actively transcribed strand, a feature that is uncharacteristic of somatic hypermutation[75]. Moreover, a huge amount of repetitive replication would be necessary to reach mutation rates of about 3×10^{-4} per base pair. In another model, reverse transcriptase is thought to synthesize a complementary DNA copy from an elongating messenger RNA that could replace the gene by homologous recombination[76]. However, this would be a convoluted way to obtain chromosomal mutations.

Two other models invoke a collision between stalled transcription forks and moving replication forks, causing either a reduction in fidelity of a normal polymerase[77], or a signal from a stalled fork for an error-prone polymerase to take over[78]. The idea that somatic hypermutation is caused by error-prone DNA polymerases has been given fresh impetus by the discovery of the errant polymerases (although the models invoking transcription-coupled repair, reverse transcripts, and collisions between transcription and replication machinery remain alive, albeit tenuously). At a meeting of The Royal Society on 'Hypermutation in antibody genes' (5–6 July 2000), which devoted one of its four sessions to the new polymerases, two favoured candidate polymerases emerged — pol ι (Rad30B) and DNA polymerase mu (pol μ) (TABLE 1).

Interest in human pol ι stems from its preference for incorporating G opposite a template T, making a G•T wobble base pair with a 3:1 preference over a Watson–Crick A•T pair[79] (FIG. 2). A T•T mispair is also easily formed, about 70% as efficiently as A•T (REF. 79). In another experiment[59], incorporation of both G and T opposite a template T is favoured by about tenfold and fivefold, respectively, relative to incorporation of A opposite T. Most of the other mispairs occur in the range of 10^{-2}–10^{-3}. The pol ι misincorporation preferences seem consistent with immunoglobulin mutational spectra in which TRANSITIONS are favoured 2:1 over TRANSVERSIONS, and A mutates more often than T (REF. 80).

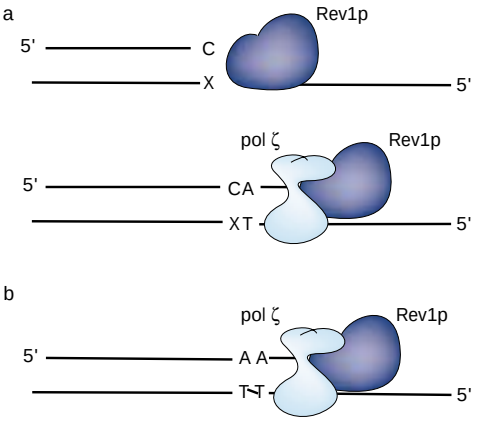


Figure 4 | **Rev1p/pol ζ lesion bypass. a** | Rev1p dCMP transferase activity incorporates C opposite a non-instructional abasic site in a DNA template in the absence of pol ζ, but it cannot extend the primer beyond the mismatch. Pol ζ can then take over for Rev1p and efficiently extend the mismatched primer terminus. However, it is not known whether Rev1p stays associated with the DNA or directly interacts with pol ζ during this process. **b** | Pol ζ can efficiently incorporate two A bases opposite a T–T 6–4 photoproduct *in vitro*, resulting in error-free bypass of the template lesion, but it can only do so in the presence of Rev1p, although the C transferase activity of Rev1p is not involved.

Human pol μ is most closely related (41% amino-acid identity) to terminal deoxynucleotidyltransferase (TdT)[81], a template-independent DNA-synthesizing enzyme. Some similiarity (23% identity) is also observed between pol μ and polymerase beta (pol β). Pol μ has weak intrinsic TERMINAL TRANSFERASE activity, and can also act as a DNA-dependent polymerase that shows poor base selection when manganese replaces magnesium as a cofactor in replication reactions *in vitro*[81]. This polymerase is expressed preferentially in peripheral lymphoid tissues and, based on analysis of the expressed sequence tag database, could be overrepresented in human B-cell germinal centres, which are critical for maturation of the immune response[81].

Should experiments using knockout mouse strains reveal a requirement for one (or perhaps several) of the errant polymerases in somatic hypermutation, it will be just the beginning of the story. Biochemical reconstitution of somatic hypermutation *in vitro* is likely to be a challenge. Any model for somatic hypermutation will have to account for the localization, polarity, magnitude and specificity of the point mutations — a tall order.

Although it is premature to speculate on specific mechanisms of somatic hypermutation *in vitro*, at least two models can be envisaged. In the first, interactions between transcription factors, somatic hypermutation-specific enhancer elements and co-activator proteins result in formation of a DNA secondary structure that recruits a mutator polymerase (FIG. 6a). In the second model, the errant polymerase is recruited to the site of a DNA nick, short gap, or perhaps even to a double-strand break (FIG. 6b). In both models, DNA synthesis by the mutator enzyme across a short gap — analogous to base-excision repair by pol β — generates mutations targeted to the variable-region gene. Moreover, nucleotide misincorporation and mismatch extension might require the action of separate polymerases. For example, pol ι often makes misincorporation errors, but seems to have difficulty in extending a mismatched primer end[59]. On the other hand, pol ζ synthesizes DNA with essentially nor-

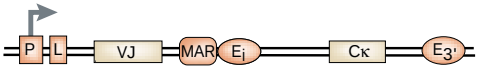


Figure 5 | **Genetic elements required for somatic hypermutation in the kappa light-chain immunoglobulin gene.** Both a promoter (P) and leader (L) sequence are required, but may be replaced with non-immunoglobulin counterparts from other genes. The intronic enhancer element (E_i) and associated nuclear matrix attachment region (MAR), as well as the 3′ enhancer (E_3'), must be present for hypermutation in the variable region (VJ). This region is flanked by the upstream promoter and downstream MAR/Ei sequences. The constant domain of the kappa light chain (Cκ) is not a target for somatic hypermutation.

GERMINAL CENTRE
A highly organized structure that develops around follicles in peripheral lymphoid organs, such as the spleen and lymph nodes, in which B cells undergo rapid proliferation and selection on formation of antigen–antibody complexes during the immune response.

TRANSITION
A point mutation in which a purine base (A or G) is substituted for a different purine base, and a pyrimidine base (C or T) is substituted for a different pyrimidine base, for example, an A•T→G•C transition.

TRANSVERSION
A point mutation in which a purine base is substituted for a pyrimidine base and vice versa, for example, an A•T→C•G transversion.

TERMINAL TRANSFERASE
An enzyme found primarily in the thymus gland that incorporates nucleotides randomly onto the 3′ end of single-stranded DNA (a non-templated reaction), in contrast to a polymerase, which incorporates nucleotides onto a 3′-primer-end in a double-stranded, template-directed reaction.

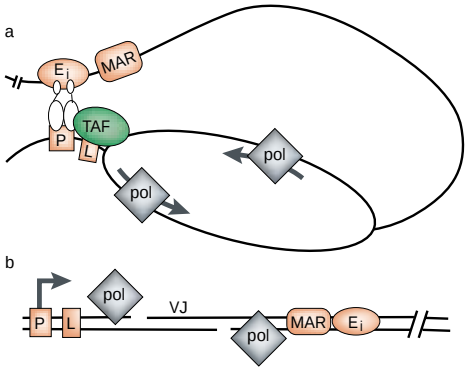


Figure 6 | **Models for somatic hypermutation by an error-prone polymerase.** **a** | Enhancer/primer-mediated targeting of somatic hypermutation. Strong interactions between enhancer-binding proteins (shown in white) and transcription-associated factors (TAFs) at the promoter (P) mediate the formation of a unique open DNA complex. Error-prone polymerases (pol), possibly pol ι or pol μ, may preferentially substitute for the normal replicative polymerases in this region of the DNA, resulting in mutations. Mutation is shown on both strands as there is no evidence for a strand bias. **b** | Nick- or gap-dependent targeting of somatic hypermutation. The variable region of the immunoglobulin genes may contain small nicks or gaps as a by-product of V(D)J recombination or very active transcription. The DNA breaks may be substrates for an error-prone polymerase to bind and generate mutations. (MAR, matrix attachment region; E_i, enhancer element.)

mal fidelity (10^{-4}–10^{-5}) but seems very efficient at mismatch extension (10^{-1}–10^{-2}). So the sequential action of both polymerases may be necessary for translesion synthesis in eukaryotic cells[59].

Other new polymerases

Pol λ/β2. Pol μ is not the only new polymerase related to pol β. Another DNA polymerase (pol λ) was identified in mouse[82], and later in humans (POL β2)[83], that shares a 32% amino-acid identity with pol β and contains the conserved family X (TABLE 1) residues critical for DNA and nucleotide binding, as well as catalysis[82]. Pol λ/β2 is expressed to significant levels only in the testes and ovaries, indicating that this enzyme may be involved in meiotic cell division[82], but this remains to be shown. Purified pol λ/β2 has polymerase activity, although nothing is known about its fidelity or preferred DNA substrates[82,83]. A BRCA1-containing carboxy-terminal (BRCT) domain located in the amino-terminal region can also be deleted from human pol λ/β2 without significant reduction in polymerase activity[83]. The *in vivo* function of BRCT domains in both pol λ/β2 and Rev1p pathways is an area that awaits further investigation.

Pol κ (Trf4). In *S. cerevisiae*, another β-like polymerase has been identified as the product of the *TRF4* gene. With its close homologue Trf5, theTrf4 protein is involved in maintaining sister-chromatid cohesion during S-phase replication. Fluorescence *in situ* hybridization (FISH) in *trf4* mutant cells[84] has revealed a marked increase in nuclei that fail to maintain cohesion of sister chromatids near centromeres and on chromosome arms. Moreover, *trf4ts/trf5* double mutant cells show delays in the G1 to S-phase transition and contain levels of DNA between those found in G1 and G2. These results imply that cells lacking *TRF* function do not completely replicate the genome. *In vitro*, Trf4 has a DNA polymerase activity with an elevated K_m (the substrate concentration that allows the reaction to proceed at half its maximum rate) for nucleotides, designated pol κ (this name has been applied to two other DNA polymerases; TABLE 1). One suggested function for Trf4 DNA polymerase is to facilitate progression of the replication fork through chromatid cohesion sites that might inhibit other DNA polymerases, such as pol δ or pol ε, causing the replication fork to collapse[84].

Why so many polymerases?

There are well established biological roles for pol V ($UmuD'_2C$), pol η (Rad30) and Rev1p. Although less certain, pol IV (DinB) is probably required to rescue stalled replication complexes. What about the other new polymerases? The DNA polymerase theta (pol θ/POLQ)[85] and MUS308 (REF. 86) proteins, both family A polymerases (TABLE 1), may be involved in repairing DNA crosslinks, but functions for pols ι (Rad 30B), μ, λ and κ (HDINB1) remain speculative (TABLE 1). It is reasonable to conclude that pol V and pol η usually copy damaged DNA templates, with pol V required for error-prone and pol η for error-free repair. Because pol V causes SOS untargeted mutations in the absence of DNA damage in *E. coli* [31], it is probably also important in natural selection and evolution[87].

In eukaryotes, the new polymerases might fill in short gaps emanating from non-homologous end-joining and from homologous recombination. Indeed, a marked increase in sister-chromatid exchange in transformed XPV cells led to the discovery[88], earlier this year, of a relationship between the S-phase checkpoint of the cell cycle and an X-ray-induced recombination pathway for repairing double-strand breaks[88]. Even *E. coli* seems to have a DNA-damage checkpoint in which UmuC and uncleaved UmuD coordinate progression through the cell cycle, signalling when it is safe to switch from stationary phase to exponential growth[89,90].

Recently, two groups have independently demonstrated the embryonic lethality of disrupting the mouse homologue of the *REV3* gene[91,92], the presumed catalytic subunit of mouse pol ζ. These studies emphasize the potential importance of specialized DNA polymerases in development and raise even more interesting questions regarding the extent of lesions that may occur during rapid cell proliferation, or rather if pol ζ might be closely linked to the mitotic checkpoints in the absence of DNA damage[91].

Macromolecular traffic control

The fact that DNA polymerases have become a 'growth industry' in the cell raises concerns about traffic control. To copy damaged DNA, rescue blocked replication forks, catalyse somatic hypermutation or fill in gaps during homologous and non-homologous recombination, the enzymes have to show up where and when they are needed and then depart when finished — and not a moment later. The basic idea, and it is not a new one, is that DNA repair proteins might always be

PRIMOSOME
A complex of proteins whose role is to initiate DNA synthesis by the *de novo* synthesis of an oligonucleotide RNA primer on a DNA template strand; a primosome is typically used to initiate synthesis at a replication origin or to re-initiate synthesis downstream of a stalled replication fork.

present at the replication fork, perhaps bound to the replication complex. A multiprotein complex, composed of two interconnected polymerase holoenzymes for coordinated leading- and lagging-strand synthesis, a lagging-strand PRIMOSOME, DNA helicase and SSB protein, would have the added baggage of other specialized polymerases, to be used sparingly when called for. This picture is easier to imagine if the replication complex is stationary, with the DNA moving through the 'replication factory'[93]. In contrast, however, the textbook version of events is that polymerase bound to accessory proteins traverses a DNA track. Nevertheless, an *in situ* assay using a *Bacillus subtilis* polymerase (PolC) tagged with green fluorescent protein identified the enzyme at discrete intracellular loci[94], indicating that the DNA may be moving through an anchored DNA polymerase.

Interactions between proteins of the replication complex and a superfamily polymerase have indeed been found in *E. coli*. Last year, Graham Walker and co-workers[95] reported differential binding between components of pol V (UmuD/UmuD′) and the α-, β- and ε-subunits of pol III. Further evidence[27] comes from the stabilization of a thermolabile pol III α-subunit, at non-permissive temperature, in the presence of pol V. These data hint that there could be a coordinated exchange between high- and low-fidelity polymerases, acting as partners in a macromolecular complex at sites of DNA damage.

However, the challenge, as Arthur Kornberg has often cautioned, is the need "to capture it alive" — that is, to reassemble an intact 'replisome–mutasome' macromolecular complex from individually purified protein components. The technology is available — *in situ* immunofluorescence, immunoprecipitation and multiple-hybrid screening can all be used to identify interactions between the new polymerases and other proteins. Even if such a fishing expedition bears fruit, to mix a metaphor, it still will not be easy to reconstitute a replisome–mutasome macromolecular complex *in vitro*. But there are successful precedents. For example, prokaryotic and eukaryotic replication, repair and recombination complexes have been built from the ground up with purified polymerases, primases, processivity factors, DNA-binding proteins, mismatch-binding proteins, recombinases, helicases and ligases. Indeed, this tried and tested approach is the first of Arthur Kornberg's[96] *Ten commandments: Lessons from the enzymology of DNA replication* — "rely on enzymology to clarify biological questions".

Update — added in proof

M. Goldsmith *et al.*[102] have shown that the MucB protein, a plasmid-encoded homologue of the *E. coli.* UmuC protein, is a DNA polymerase capable of translesion synthesis past an abasic site in the presence of MucA′ (the UmuD′ homologue), RecA and SSB protein.

Links

DATABASE LINKS *REV3* | *REV7* | *REV1* | *umuC* | *umuD* | pol κ | pol η | Rad30 | pol ι | pol μ | pol λ | *TRF4* | pol θ
ENCYCLOPEDIA OF LIFE SCIENCES DNA polymerase fidelity mechanisms | Eukaryotic replication fork

1. Kornberg, A., Lehman, I. R., Bessman, M. J. & Simms, E. S. Enzymatic synthesis of deoxyribonucleic acid. *Biochim. Biophys. Acta* **21**, 197–198 (1956).
2. Lehman, I. R., Bessman, M. J., Simms, E. S. & Kornberg, A. Enzymatic synthesis of deoxyribonucleic acid I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *J. Biol. Chem.* **233**, 163–170 (1958).
3. de Lucia, P. & Cairns, J. Isolation of an *E. coli* strain with a mutation affecting DNA polymerase. *Nature* **224**, 1164–1166 (1969).
4. Knippers, R. DNA polymerase II. *Nature* **228**, 1050–1053 (1970).
5. Kornberg, T. & Gefter, M. L. Purification and DNA synthesis in cell-free extracts: properties of DNA polymerase II. *Proc. Natl Acad. Sci. USA* **68**, 761–764 (1971).
6. Nusslein, V., Otto, B., Bonhoeffer, F. & Schaller, H. Function of DNA polymerase 3 in DNA replication. *Nature New Biol.* **234**, 285–286 (1971).
7. Gefter, M. L., Hirota, Y., Kornberg, T., Wechsler, J. A. & Barnoux, C. Analysis of DNA polymerases II and III in mutants of *Escherichia coli* thermosensitive for DNA synthesis. *Proc. Natl Acad. Sci. USA* **68**, 3150–3153 (1971).
8. Rangarajan, S., Woodgate, R. & Goodman, M. F. A phenotype for enigmatic DNA polymerase II: a pivotal role for pol II in replication restart in UV-irradiated *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **96**, 9224–9229 (1999).
9. Nelson, J. R., Lawrence, C. W. & Hinkle, D. C. Deoxycytidyl transferase activity of yeast REV1 protein. *Nature* **382**, 729–731 (1996).
   Shows that yeast Rev1 protein has template-dependent dCMP transferase activity, and is the first member of the UmuC/DinB /Rev1p/Rad30 superfamily found to be involved in DNA synthesis.
10. Washington, M. T., Johnson, R. E., Prakash, S. & Prakash, L. Fidelity and processivity of *Saccharomyces cerevisiae* DNA polymerase η. *J. Biol. Chem.* **274**, 36835–36838 (1999).
    Yeast Rad30 protein is shown to have low-fidelity DNA polymerase activity and is designated as pol η.
11. Matsuda, T., Bebenek, K., Masutani, C., Hanaoka, F. & Kunkel, T. A. Low fidelity DNA synthesis by human DNA polymerase η. *Nature* **404**, 1011–1013 (2000).
12. Tang, M. *et al.* Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted mutagenesis. *Nature* **404**, 1014–1018 (2000).
    The *E. coli* DNA pol V mutasome is shown to be responsible for SOS-induced mutagenesis through error-prone translesion synthesis.
13. Echols, H. & Goodman, M. F. Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.* **60**, 477–511 (1991).
14. Lawrence, C. W. & Christensen, R. B. Ultraviolet-induced reversion of *CYC1* alleles in radiation-sensitive strains of yeast. III. *REV3* mutant strains. *Genetics* **92**, 397–408 (1979).
15. Lawrence, C. W., Das, G. & Christensen, R. B. *REV7*, a new gene concerned with UV mutagenesis in yeast. *Mol. Gen. Genet.* **200**, 80–85 (1985).
16. Lawrence, C. W., Nisson, P. E. & Christensen, R. B. UV and chemical mutagenesis in *REV7* mutants of yeast. *Mol. Gen. Genet.* **200**, 86–91 (1985).
17. Lawrence, C. W. & Christensen, R. Ultraviolet-induced reversion of *cyc1* alleles in radiation sensitive strains of yeast. I. *rev1* mutant strains. *J. Mol. Biol.* **122**, 1–22 (1978).
18. Sommer, S., Knezevic, J., Bailone, A. & Devoret, R. Induction of only one SOS operon, *umuDC*, is required for SOS mutagenesis in *E. coli*. *Mol. Gen. Genet.* **239**, 137–144 (1993).
19. Rupp, W. D. & Howard-Flanders, P. Discontinuities in the DNA synthesized in an excision-defective strain of *Escherichia coli* following ultraviolet radiation. *J. Mol. Biol.* **31**, 291–304 (1968).
20. Sancar, A. & Sancar, G. B. DNA repair enzymes. *Annu. Rev. Biochem.* **57**, 29–67 (1988).
21. Bridges, B. A. & Woodgate, R. The two-step model of bacterial UV mutagenesis. *Mutat. Res.* **150**, 133–139 (1985).
22. Walker, G. Skiing the black diamond slope: Progress on the biochemistry of translesion DNA synthesis. *Proc. Natl Acad. Sci. USA* **95**, 10348–10350 (1998).
23. Woodgate, R. A plethora of lesion-replicating DNA polymerases. *Genes Dev.* **13**, 2191–2195 (1999).
24. Goodman, M. F. & Tippin, B. Sloppier copier DNA polymerases involved in genome repair. *Curr. Opin. Genet. Dev.* **10**, 162–168 (2000).
25. Friedberg, E., Feaver, W. J. & Gerlach, V. L. The many faces of DNA polymerases: Strategies for mutagenesis and for mutational avoidance. *Proc. Natl Acad. Sci. USA* **97**, 5681–5683 (2000).
26. Tang, M. *et al.* Biochemical basis of SOS mutagenesis in *Escherichia coli*: Reconstitution of *in vitro* lesion bypass dependent on the UmuD′$_2$C mutagenic complex and RecA protein. *Proc. Natl Acad. Sci. USA* **95**, 9755–9760 (1998).
    The first evidence for the intrinsic polymerase activity in UmuD′$_2$C (now pol V).
27. Tang, M. *et al.* UmuD′$_2$C is an error-prone DNA polymerase, *Escherichia coli* pol V. *Proc. Natl Acad. Sci. USA* **96**, 8919–8924 (1999).
    Purified UmuD′$_2$C from a temperature-sensitive DNA polymerase III mutant possesses intrinsic polymerase activity and can do translesion synthesis, demonstrating that UmuD′$_2$C is a bona fide DNA polymerase.
28. Reuven, N. B., Arad, G., Maor-Shoshani, A. & Livneh, Z. The mutagenic protein UmuC is a DNA polymerase activated by UmuD′, RecA, and SSB and is specialized for translesion replication. *J. Biol. Chem.* **274**, 31763–31766 (1999).
    The *E. coli* UmuC protein was shown to have DNA polymerase activity.
29. Nelson, J. R., Gibbs, P. E. M., Nowicka, A. M., Hinkle, D. C. & Lawrence, C. W. Evidence for a second function for *Saccharomyces cerevisiae* Rev1p. *Mol. Microbiol.* **37**, 549–553 (2000).
30. Walker, G. C. Inducible DNA repair systems. *Annu. Rev. Biochem.* **54**, 425–457 (1985).
31. Friedberg, E. C., Walker, G. C. & Siede, W. in *DNA Repair and Mutagenesis* Vol. 1, 407–522 (ASM Press, Washington DC, 1995).
32. Kato, T. & Shinoura, Y. Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutagenesis by ultraviolet light. *Mol. Gen. Genet.* **156**, 121–131 (1977).
33. Steinborn, G. Uvm mutants of *Escherichia coli* K12 deficient in UV mutagenesis. I. Isolation of *uvm* mutants

and their phenotypical characterization in DNA repair and mutagenesis. *Mol. Gen. Genet.* **165**, 87–93 (1978).
34. Bagg, A., Kenyon, C. J. & Walker, G. C. Inducibility of a gene product required for UV and chemical mutagenesis in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **78**, 5749–5753 (1981).
35. Burckhardt, S. E., Woodgate, R., Scheuremann, R. H. & Echols, H. UmuD mutagenesis protein of *Escherichia coli*: Overproduction, purification, and cleavage by RecA. *Proc. Natl Acad. Sci. USA* **85**, 1811–1815 (1988).
36. Woodgate, R., Rajagopalan, M., Lu, C. & Echols, H. UmuC mutagenesis protein of *Escherichia coli*: Purification and interaction with UmuD and UmuD′. *Proc. Natl Acad. Sci. USA* **86**, 7301–7305 (1989).
37. Bruck, I., Woodgate, R., McEntee, K. & Goodman, M. F. Purification of a soluble UmuD′C complex from *Escherichia coli*: Cooperative binding of UmuD′C to single-stranded DNA. *J. Biol. Chem.* **271**, 10767–10774 (1996).
38. Echols, H. & Goodman, M. F. Mutation induced by DNA damage: A many protein affair. *Mutat. Res.* **236**, 301–311 (1990).
39. Witkin, E. M. Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol. Rev.* **40**, 869–907 (1976).
40. Strauss, B. S. The 'A rule' of mutagen specificity: a consequence of DNA polymerase bypass of non-instructional lesions? *Bioessays* **13**, 79–84 (1991).
41. Pham, P., O'Donnell, M., Woodgate, R. & Goodman, M. F. A 'cowcatcher' model for SOS–lesion targeted mutations in *E. coli* involving pol V, RecA, SSB and β sliding clamp. (submitted).
42. Kuzminov, A. Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage λ. *Microbiol. Mol. Biol. Rev.* **63**, 751–813 (1999).
43. Goodman, M. F. Coping with replication 'train wrecks' in *Escherichia coli* using pol V, pol II, and RecA proteins. *Trends Biochem. Sci.* **25**, 189–195 (2000).
44. Kim, S.-R. *et al*. Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: An overexpression of *dinB*/*dinP* results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. *Proc. Natl Acad. Sci. USA* **94**, 13792–13797 (1997).
45. Cox, M. M. *et al*. The importance of repairing stalled replication forks. *Nature* **404**, 37–41 (2000).
46. Wagner, J. *et al*. The *dinB* gene encodes a novel *Escherichia coli* DNA polymerase (DNA pol IV). *Mol. Cell* **4**, 281–286 (1999).
E. coli DinB/pol IV protein is shown to have distributive polymerase activity that can also extend misaligned 3′-primer ends.
47. Foster, P. L. Adaptive mutation: The uses of adversity. *Annu. Rev. Microbiol.* **47**, 467–504 (1993).
48. Foster, P. L. Adaptive mutation in *Escherichia coli*. *Cold Spring Harb. Symp. Quant. Biol.* (in the press).
49. Gerlach, V. L. *et al*. Human and mouse homologs of *Escherichia coli* DinB (DNA polymerase IV), members of the UmuC/DinB superfamily. *Proc. Natl Acad. Sci. USA* **96**, 11922–11927 (1999).
50. Ohashi, E. *et al*. Error-prone bypass of certain DNA lesions by the human DNA polymerase κ. *Genes Dev.* **14**, 1589–1594 (2000).
51. Boiteux, S. & Laval, J. Coding properties of poly(deoxycycytidylic acid) templates containing uracil or apyrimidinic sites: *In vitro* modulation of mutagenesis by DNA repair enzymes. *Biochemistry* **21**, 6746–6751 (1982).
52. Schaaper, R. M., Kunkel, T. A. & Loeb, L. A. Infidelity of DNA synthesis associated with bypass of apurinic sites. *Proc. Natl Acad. Sci. USA* **80**, 487–491 (1983).
53. Sagher, D. & Strauss, B. S. Insertion of nucleotides opposite apurinic/apyrimidinic sites in deoxyribonucleic acid during *in vitro* synthesis: uniqueness of adenine nucleotides. *Biochemistry* **22**, 4518–4526 (1983).
54. Lawrence, C. W., Borden, A., Banerjee, S. K. & LeClerc, J. E. Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector. *Nucleic Acids Res.* **18**, 2153–2157 (1990).
55. Gibbs, P. E., McGregor, W. G., Maher, V. M., Nisson, P. & Lawrence, C. W. A human homolog of the *Saccharomyces cerevisiae* REV3 gene, which encodes the catalytic subunit of DNA polymerase ζ. *Proc. Natl Acad. Sci. USA* **95**, 6876–6880 (1998).
56. Gibbs, P. E. M. *et al*. The function of the human homolog of *Saccharomyces cerevisiae REV1* is required for mutagenesis induced by UV light. *Proc. Natl Acad. Sci. USA* **97**, 4186–4191 (2000).
57. Lin, W. *et al*. The human *REV1* gene codes for a DNA template-dependent dCMP transferase. *Nucleic Acids Res.* **27**, 4468–4475 (1999).
58. Nelson, J. R., Lawrence, C. W. & Hinkle, D. C. Thymine–thymine dimer bypass by yeast DNA polymerase ζ. *Science* **272**, 1646–1649 (1996).
59. Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S. & Prakash, L. Mutagenic bypass of DNA lesions by sequential action of eukaryotic DNA polymerases ι and ζ. *Nature* **406**, 1015–1019 (2000).
60. Johnson, R. E., Prakash, S. & Prakash, L. Requirement of DNA polymerase activity of yeast Rad30 protein for its biological function. *J. Biol. Chem.* **274**, 15975–15977 (1999).
61. Masutani, C. *et al*. The *XPV* (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. *Nature* **399**, 700–704 (1999).
Purification of human xeroderma pigmentosum variant (XPV) and demonstration that it is the human homologue of yeast DNA pol η.
62. Johnson, R. E., Kondratick, C. M., Prakash, S. & Prakash, L. *hRAD30* mutations in the variant form of Xeroderma pigmentosum. *Science* **285**, 263–265 (1999).
63. Cordeiro-Stone, M., Zaritskaya, L. S., Price, L. K. & Kaufmann, W. K. Replication fork bypass of a pyrimidine dimer blocking leading strand DNA synthesis. *J. Biol. Chem.* **272**, 13945–13954 (1997).
64. Masutani, C. *et al*. Xeroderma pigmentosum variant (XP-V) correcting protein from HeLa cells has a thymine dimer bypass DNA polymerase activity. *EMBO J.* **18**, 3491–3501 (1999).
65. Johnson, R. E., Prakash, S. & Praskah, L. Efficient bypass of a thymine–thymine dimer by yeast DNA polymerase, Pol η. *Science* **283**, 1001–1004 (1999).
66. Lin, W., Wu, X. & Wang, Z. A full-length cDNA of hREV3 is predicted to encode DNA polymerase ζ for damage-induced mutagenesis in humans. *Mutat. Res.* **433**, 89–98 (1999).
67. Murakumo, Y. *et al*. A human REV7 homolog that interacts with the polymerase ζ catalytic subunit hREV3 and the spindle assembly checkpoint protein hMAD2. *J. Biol. Chem.* **275**, 4391–4397 (2000).
68. Berek, C. & Milstein, C. The dynamic nature of the antibody repertoire. *Immunol. Rev.* **105**, 5–26 (1988).
69. Zhang, J., MacLennan, I. C., Liu, Y. J. & Lane, P. J. Is rapid proliferation in B centroblasts linked to somatic mutation and memory B cell clones? *Immunol. Lett.* **18**, 2393–2400 (1988).
70. Kim, S., Davis, M., Sinn, E., Patten, P. & Hood, L. Antibody diversity: somatic hypermutation of rearranged VH genes. *Cell* **27**, 573–581 (1981).
71. Neuberger, M. S. *et al*. Monitoring and interpreting the intrinsic features of somatic hypermutation. *Immunol. Rev.* **162**, 107–116 (1998).
72. Betz, A. G. *et al*. Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell* **77**, 239–248 (1994).
73. Winter, D. B. & Gearhart, P. J. Another piece in the hypermutation puzzle. *Curr. Biol.* **5**, 1345–1346 (1995).
74. Storb, U. The molecular basis of somatic hypermutation of immunoglobulin genes. *Curr. Opin. Immunol.* **8**, 206–214 (1996).
75. Milstein, C., Neuberger, M. S. & Staden, R. Both DNA strands of antibody genes are hypermutation targets. *Proc. Natl Acad. Sci. USA* **95**, 8791–8794 (1998).
76. Steel, E. J., Rothenfluh, H. S. & Blanden, R. V. Mechanism of antigen-driven somatic hypermutation of rearranged immunoglobulin *V(D)J* genes in the mouse. *Immunol. Cell. Biol.* **75**, 82–95 (1997).
77. Winter, D. B., Sattar, N. & Gearhart, P. J. in *Current topics in Microbiology and Immunology* (eds Kelsoe, G. & Flajnits, M. F.) 1–10 (Springer, Berlin, 1998).
78. Diaz, M. & Flajnik, M. F. Evolution of somatic hypermutation and gene conversion in adaptive immunity. *Immunol. Rev.* **162**, 13–24 (1998).
An excellent summary of the key features of somatic hypermutation spectra.
79. Tissier, A., McDonald, J. P., Frank, E. G. & Woodgate, R. pol ι, a remarkably error-prone human DNA polymerase. *Genes Dev.* **14**, 1642–1650 (2000).
Demonstration that human Rad30B/pol ι prefers to make the wobble G–T base pair over the correct A–T base pair, and overall has extremely poor fidelity.
80. Foster, S. J., Dorner, T. & Lipsky, P. E. Somatic hypermutation of VκJκ rearrangements: targeting of RGYW motifs on both DNA strands and preferential selection of mutated codons within RGYW motifs. *Eur. J. Immunol.* **29**, 4011–4021 (1999).
81. Dominguez, O. *et al*. DNA polymerase mu (Pol μ), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *EMBO J.* **19**, 1731–1742 (2000).
82. Garcia-Diaz, M. *et al*. DNA polymerase lambda (Pol lambda), a novel eukaryotic DNA polymerase with a potential role in meiosis. *J. Mol. Biol.* **301**, 851–867 (2000).
83. Nagasawa, K.-I. *et al*. Identification and characterization of human DNA polymerase β2, a DNA polymerase β–related enzyme. *J. Biol. Chem.* **275**, 31233–31238 (2000).
84. Wang, Z., Castano, I. B., de Las Penas, A., Adams, C. & Christman, M. F. Pol κ: A DNA polymerase required for sister chromatid cohesion. *Science* **289**, 774–779 (2000).
Discovery that the yeast Trf4 protein is a polymerase essential for sister-chromatid cohesion, indicating that possible polymerase exchange mechanisms may be needed to replicate chromosomes fully.
85. Sharief, F. S., Vojta, P. J., Ropp, P. A. & Copeland, W. C. Cloning and chromosomal mapping of the human DNA polymerase θ (POLQ), the eighth human DNA polymerase. *Genomics* **59**, 90–96 (1999).
86. Oshige, M., Aoyage, N., Harris, P. V., Burtis, K. C. & Sakaguchi, K. A new DNA polymerase species from *Drosophila melanogaster*: a probable *mus308* gene product. *Mutat. Res.* **433**, 183–192 (1999).
87. Radman, M. Enzymes of evolutionary change. *Nature* **401**, 866–869 (1999).
88. Limoli, C. L., Giedzinski, E., Morgan, W. F. & Cleaver, J. E. Polymerase η deficiency in the Xeroderma pigmentosum variant uncovers an overlap between the S phase checkpoint and double-strand break repair. *Proc. Natl Acad. Sci. USA* **97**, 7939–7946 (2000).
89. Opperman, T., Murli, S., Smith, B. T. & Walker, G. C. A model for *umuDC*-dependent prokaryotic DNA damage checkpoint. *Proc. Natl Acad. Sci. USA* **96**, 9218–9223 (1999).
90. Murli, S., Opperman, T., Smith, B. T. & Walker, G. C. A role for the *umuDC* gene products of *Escherichia coli* in increasing resistance to DNA damage in stationary phase by inhibiting the transition to exponential growth. *J. Bacteriol.* **182**, 1127–1135 (2000).
91. Bemark, M., Khamlichi, A. A., Davies, S. L. & Neuberger, M. S. Disruption of mouse polymerase ζ (*Rev3*) leads to embryonic lethality and impairs blastocyst development *in vitro*. *Curr. Biol.* **10**, 1213–1216 (2000).
92. Wittschieben, J. *et al*. Disruption of the developmentally regulated *Rev3l* gene causes embryonic lethality. *Curr. Biol.* **10**, 1217–1220 (2000).
93. Shapiro, L. & Losick, R. Dynamic spatial regulation in the bacterial cell. *Cell* **100**, 89–98 (2000).
94. Lemon, K. P. & Grossman, A. D. Localization of bacterial DNA polymerase: evidence for a factory model of replication. *Science* **282**, 1516–1519 (1998).
95. Sutton, M. D., Opperman, T. & Walker, G. C. The *Escherichia coli* SOS mutagenesis proteins UmuD and UmuD′ interact physically with the replicative DNA polymerase. *Proc. Natl Acad. Sci. USA* **96**, 12373–12378 (1999).
96. Kornberg, A. Ten commandments: Lessons from the enzymology of DNA replication. *J. Bacteriol.* **182**, 3613–3618 (2000).
97. Perry, K. L., Elledge, S. J., Mitchell, B. B., Marsh, L. & Walker, G. C. *umuDC* and *mucAB* operons whose products are required for UV light- and chemical-induced mutagenesis: UmuD, MucA, and LexA proteins share homology. *Proc. Natl Acad. Sci. USA* **82**, 4331–4335 (1985).
98. Shinagawa, H., Iwasaki, T., Kato, T. & Nakata, A. RecA protein-dependent cleavage of UmuD protein and SOS mutagenesis. *Proc. Natl Acad. Sci. USA* **85**, 1806–1810 (1988).
99. Maor-Shoshani, A., Reuven, N. B., Tomer, G. & Livneh, Z. Highly mutagenic replication by DNA polymerase V (UmuC) provides a mechanistic basis for SOS untargeted mutagenesis. *Proc. Natl Acad. Sci. USA* **97**, 565–570 (2000).
100. Johnson, R. E., Prakash, S. & Prakash, L. The human *DINB1* gene encodes the DNA polymerase pol θ. *Proc. Natl Acad. Sci. USA* **97**, 3838–3843 (2000).
101. Morelli, C., Mungall, A. J., Negrini, M., Barbanti–Brodano, G. & Croce, C. M. Alternative splicing, genomic structure, and fine chromosome localization of REV3L. *Cytogenet. Cell Genet.* **83**, 18–20 (1998).
102. Goldsmith, M., Sarov-Blat, L. & Livneh, Z. Plasmid-encoded MucB protein is a DNA polymerase (pol RI) specialized for lesion bypass in the presence of MucA′, RecA, and SSB. *Proc. Natl Acad. Sci. USA* **97**, 11227–11231 (2000).

Acknowledgements

This work was supported by grants from the National Institutes of Health. I want to express heartfelt gratitude to my collaborators, Roger Woodgate, Mike O'Donnell, John-Stephen Taylor, Kevin McEntee and especially to Hatch Echols. I also want to express my sincere appreciation to the students in my laboratory, Phuong Pham, Mengjia Tang, Xuan Shen, Irina Bruck and Jeffrey Bertram.