nature biotechnology

Molecular breeding of polymerases for amplification of ancient DNA

Marc d'Abbadie[1], Michael Hofreiter[2], Alexandra Vaisman[3], David Loakes[1], Didier Gasparutto[4], Jean Cadet[4], Roger Woodgate[3], Svante Pääbo[2] & Philipp Holliger[1]

© 2007 Nature Publishing Group http://www.nature.com/naturebiotechnology

In the absence of repair, lesions accumulate in DNA. Thus, DNA persisting in specimens of paleontological, archaeological or forensic interest is inevitably damaged[1]. We describe a strategy for the recovery of genetic information from damaged DNA. By molecular breeding[2] of polymerase genes from the genus *Thermus* (*Taq* (*Thermus aquaticus*), *Tth* (*Thermus thermophilus*) and *Tfl* (*Thermus flavus*)) and compartmentalized self-replication[3,4] selection, we have evolved polymerases that can extend single, double and even quadruple mismatches, process non-canonical primer-template duplexes and bypass lesions found in ancient DNA, such as hydantoins and abasic sites. Applied to the PCR amplification of 47,000–60,000-year-old cave bear DNA, these outperformed *Taq* DNA polymerase by up to 150% and yielded amplification products at sample dilutions at which *Taq* did not. Our results demonstrate that engineered polymerases can expand the recovery of genetic information from Pleistocene specimens and may benefit genetic analysis in paleontology, archeology and forensic medicine.

Ancient DNA sequences have been isolated from a wide variety of sources[1] and have provided information about human migration[5], animal and crop domestication[6,7] and the genetic relationship between modern *Homo sapiens* and its closest extinct relative *H. neandertalensis*[8,9]. Even under optimal burial conditions, however, DNA damage is progressive and either limits the length of continuous sequence that can be recovered or renders even well-preserved specimens unproductive despite the demonstrable (by hybridization) presence of DNA[10]. We reasoned that genetic information encoded in such samples may not be lost but may simply be inaccessible because the DNA polymerases commonly used for PCR stall at sites of damage[11]. Polymerases capable of PCR amplification of damaged DNA would therefore have the potential to enhance the retrieval of ancient DNA sequences and, in combination with direct sequencing approaches[8,12], to expand paleogenomic data.

Engineering polymerases that combine the processivity and selectivity required for PCR with a high tolerance for template damage is challenging. Furthermore, damage tolerance should be generic, as detailed information about the forms of DNA damage in ancient

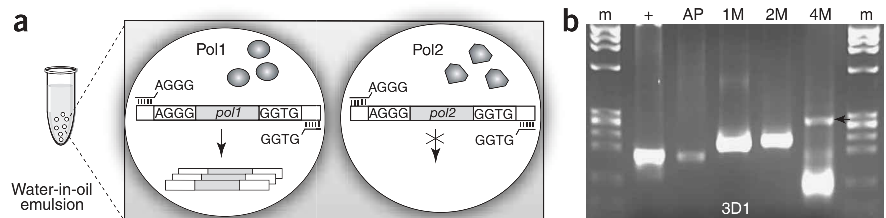

Figure 1 Selection scheme. (**a**) Principle of CSR selection. CSR is based on a simple feedback loop, in which a polymerase replicates only its own encoding gene. Compartmentalization in the aqueous compartments of a water-in-oil emulsion[30] serves to isolate individual self-replication reactions from each other. In such a system adaptive gains directly (and proportionally) translate into genetic amplification of the encoding gene. Two independent aqueous emulsion compartments are shown. Polymerases (such as Pol1, left compartment) that are capable of using quadruple-mismatch primers (AGGG•AGGG, GGTG•GGTG) to replicate their one encoding gene (*pol1*) produce 'offspring'—that is, increase their copy number in the post-selection population—whereas polymerases such as Pol2 (right compartment) that are unable to use quadruple-mismatch primers disappear from the gene pool (a more detailed step-by-step selection scheme is shown in **Supplementary Fig. 1**). (**b**) Polymerase activity in PCR of the selected polymerase 3D1 using matched (+), single-mismatch (1M: G•A, C•C), double-mismatch (2M: CC•CC, CC•CC), quadruple-mismatch (4M: AGGG•AGGG, GGTG•GGTG) and abasic site primers (AP). Under the same conditions, wild-type polymerases *Taq*, *Tth* and *Tfl* yielded only amplification products using matched (+) primers. There is substantial primer-dimer formation with the quadruple-mismatch primers. The correct-sized quadruple-mismatch amplification product is marked by an arrow. m, ϕx174 *Hae*III digest marker.

[1]Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK. [2]Max Planck Institute for Molecular Anthropology, Department of Evolutionary Genetics, Deutscher Platz 6, D-04103 Leipzig, Germany. [3]Section on DNA Replication, Repair and Mutagenesis, Building 6, Room 1A13, National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892-2725, USA. [4]Laboratoire Lésions des Acides Nucléiques, SCIB/UMR-E 3 (CEA/UJF), Département de Recherche Fondamentale sur la Matière Condensée, Commissariat à l'Énergie Atomique (CEA)/Grenoble, F-38054 Grenoble Cedex 9, France. Correspondence should be addressed to P.H. (ph1@mrc-lmb.cam.ac.uk).

Received 21 February; accepted 18 May; published online 15 July 2007; doi:10.1038/nbt1321

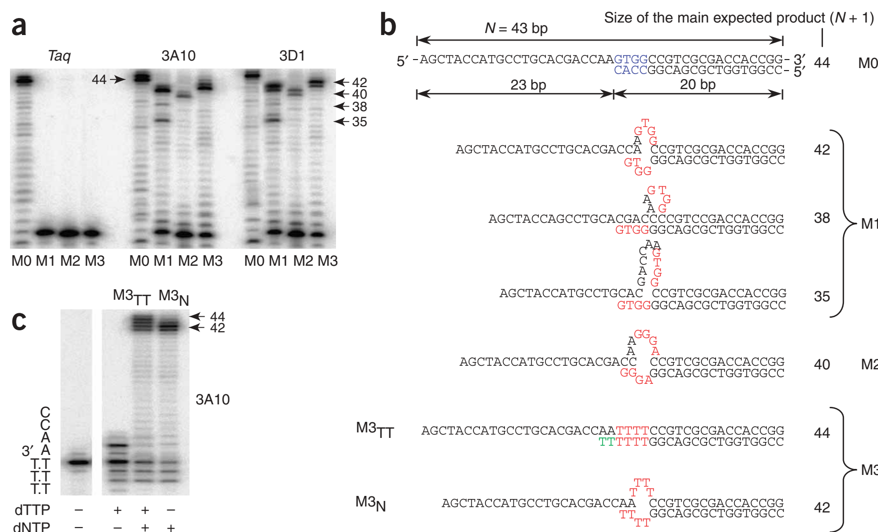

Figure 2 Quadruple-mismatch extension and template slippage. (**a**) Enzyme concentrations were normalized on an undamaged template and the polymerases were assayed for their activity in extending radiolabeled quadruple-mismatch primers (M1, GGTG•GGTG; M2, AGGG•AGGG; M3: TTTT•TTTT) as compared to matched primer (M0). Only selected polymerases 3A10 and 3D1 yield extension products with M1–3, but extension products are predominately shorter than M0. (**b**) Possible primer template configurations and expected main product lengths (N+1). Matched primer-template sequences (M0) at primer 3′ end are shown in blue; mismatched and misaligned structures are shown in red. (**c**) Quadruple-mismatch extension of M3 by 3A10, when extended with dTTP only (lane 2, $M3_{TT}$). Possible primer template configuration for M3 for quadruple-mismatch extension is shown with the two incorporated dTs in green ($M3_{TT}$). When initial dTTP extension is followed by addition of the remaining dNTPs (lane 3), 3A10 yields a significant proportion of full-length extension product, as evidenced by the appearance of bands at (+43), (+44), whereas addition of dNTPs from the start (lane 4, $M3_N$)) favors the misaligned structure that yields a shorter extension product centered around (+42) as in **a**.

© 2007 Nature Publishing Group http://www.nature.com/naturebiotechnology

samples is lacking (and damage may vary depending on burial conditions). Many lesions (with the exception of miscoding lesions such as uracil) abrogate base-pairing and yield distorted, non-cognate 3′ structures, similar to transversion mismatches. In the case of A-family polymerases, such mismatches cause significant stalling not just at the primer 3′ end[13] but up to four bases upstream[14]. To maximize tolerance of such distorted primer-template structures, we decided to select for polymerases capable of extending a primer 3′ terminus preceded by up to four mismatched bases.

Previous library designs for polymerase evolution were based on random mutagenesis of the polymerase gene[3,4] or defined regions thereof[15–17], but these proved unproductive for the selection of damage-tolerant polymerases (data not shown). We therefore prepared new libraries using molecular breeding[2]. We recombined three A-family polymerase (DNA pol I) genes from the genus *Thermus*—*Taq* (from *T. aquaticus*), *Tth* (from *T. thermophilus*) and *Tfl* (from *T. flavus*)—to create library 3T for selection by compartmentalized self-replication (CSR)[3] (**Fig. 1a** and **Supplementary Fig. 1** online). To test library performance, we first selected for CC•CC double-mismatch extension.

Single C•C mismatches are extended $>10^6$-fold less efficiently than matched termini by *Taq*[13]. Double-mismatch extension has been reported only for the Y-family polymerases polη and polι[18,19]. Nevertheless, a single round of CSR selection of 3T produced several clones with efficient double-mismatch extension. In particular, H10, a *Tth-Taq* chimera containing nine additional point mutations (F74I, F280L, P300S, T387A, A441V, A519V Q536R, R679G, F699L (*Tth* numbering)), could extend primers with double (CC•CC) as well as single (C•C; G•A) mismatches in PCR (data not shown). However, H10 showed no clear improvement in damage tolerance compared to the single-mismatch extension polymerase M1 previously described[4].

We therefore proceeded to select for extension of the more challenging quadruple mismatches GGTG•GGTG, AGGG•AGGG (primer (5′-3′)•template (3′-5′)) (**Fig. 1a**). After three rounds of CSR, we recovered a diverse set of polymerase chimeras with new properties (**Supplementary Table 1** online), including the generic ability to use single, double and quadruple mismatches (for example, 3D1, **Fig. 1b**) or to bypass template lesions such as abasic sites or hydantoins.

Despite their diverse properties, all the selected polymerases share a similar arrangement of gene segments. Though diverging in detail, in all polymerases the N-terminal region (forming part or all of the 5′-3′ exonuclease domain) and the C terminus derive from *Tth*, whereas the protein core derives mainly from *Taq*. Four point mutations (L33P, E76K, D145G and E822K) diverging from *Taq*, *Tth* or *Tfl* were present in several polymerases. One variant (3B10) had acquired a 16-amino-acid C-terminal extension through a frameshift mutation (see **Supplementary Fig. 2** online).

We chose two polymerases, 3D1 (a *Tth-Taq* chimera with six additional mutations: L33P, E76K, D145G, P552S, E775G, M777T) and 3A10 (a *Tth-Taq* chimera with eight additional mutations: E76K, E91Q, D145G, R336Q, A448T, I616M, V739M and E744G), for detailed investigation. We first examined extension of three different quadruple mismatches: GGTG•GGTG, AGGG•AGGG (used for selection) and the unrelated TTTT•TTTT. Although *Taq* was unable to extend any of the mismatches, both 3A10 and 3D1 could, but the majority of the reaction products were shorter than expected (**Fig. 2a,b**). Extension of matched termini generated by strand misalignment seemed to compete with direct extension of quadruple mismatches. To assess whether extension of quadruple mismatches was possible, we performed extension of the TTTT•TTTT quadruple mismatch with a single nucleotide (dTTP) that was complementary to the first two template bases (dA-dA) with 3A10. Misalignment was suppressed and the TTTT•TTTT quadruple mismatch was efficiently extended by correct incorporation of two dTs (**Fig. 2c**). Furthermore, addition of dNTPs resulted in the further extension of the quadruple mismatch and the formation of a significant amount of full-length product.

Both 3A10 and 3D1 were capable of efficiently bypassing DNA lesions relevant to ancient DNA, such as abasic sites, 5-hydroxyhydantoin (5Hyd) and 5-methyl-5-hydroxyhydantoin (5MeHyd), all of which are associated with PCR failure from ancient samples[20], whereas the same lesions blocked extension by the parent polymerases *Taq* (**Fig. 3a** and **Supplementary Fig. 3a** online), *Tth* and *Tfl* (data not shown). Both polymerases, particularly 3A10, were also proficient at bypassing abasic sites in PCR (see **Supplementary Fig. 3b** online). 3A10 achieved translesion synthesis of both abasic sites and 5Hyd lesions that was up to 50-fold superior to *Taq* as judged by polymerase ELISA[17] (see **Supplementary Fig. 3c** online). Both 3A10 and 3D1

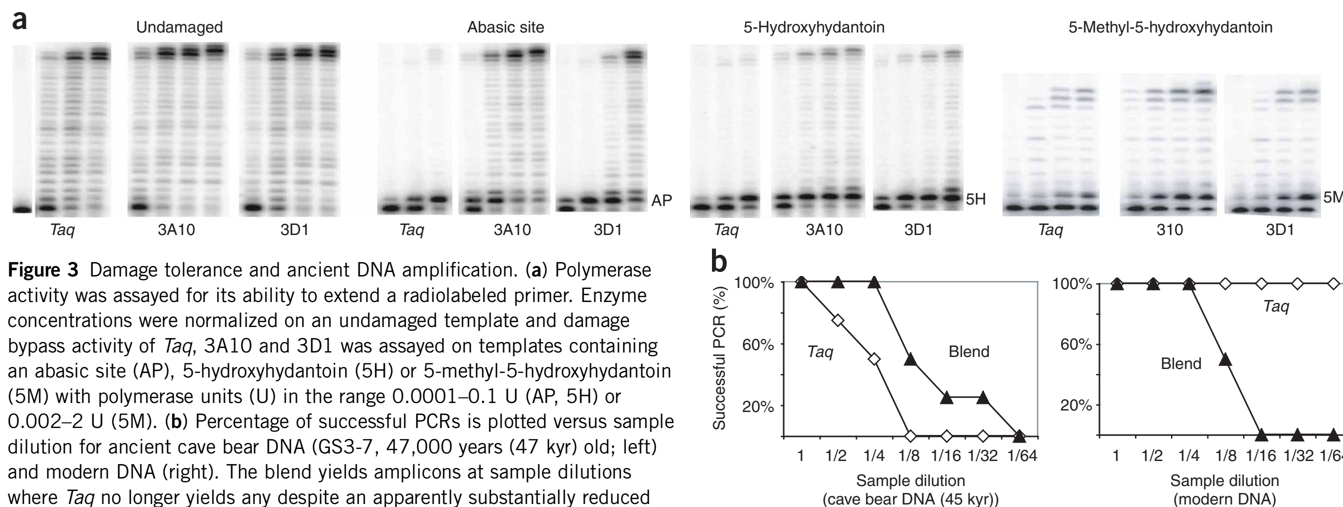

Figure 3 Damage tolerance and ancient DNA amplification. (**a**) Polymerase activity was assayed for its ability to extend a radiolabeled primer. Enzyme concentrations were normalized on an undamaged template and damage bypass activity of *Taq*, 3A10 and 3D1 was assayed on templates containing an abasic site (AP), 5-hydroxyhydantoin (5H) or 5-methyl-5-hydroxyhydantoin (5M) with polymerase units (U) in the range 0.0001–0.1 U (AP, 5H) or 0.002–2 U (5M). (**b**) Percentage of successful PCRs is plotted versus sample dilution for ancient cave bear DNA (GS3-7, 47,000 years (47 kyr) old; left) and modern DNA (right). The blend yields amplicons at sample dilutions where *Taq* no longer yields any despite an apparently substantially reduced affinity for (undamaged) template DNA.

© 2007 Nature Publishing Group http://www.nature.com/naturebiotechnology

showed a similar spectrum of dNTP insertion (A>G>>T>>>C) opposite abasic sites or 5Hyd/5MeHyd lesions. For abasic sites, which derive from depurination, bypass either is silent or leads to G-to-A transitions, whereas for the hydantoins, which originate from pyrimidines, bypass yields C/T-to-A transversion mutations. We estimated fidelity on an undamaged template to be decreased by two- to fourfold (3D1) and sevenfold (3A10) as compared with *Taq* (see **Supplementary Fig. 4a**,**b** online).

The ability of the selected polymerases to efficiently bypass template lesions in PCR encouraged us to test their activity for the recovery of ancient DNA. We carried out subsequent experiments using a blend of *Taq* with the most promising of the selected polymerases (3A10, 3B5, 3B6, 3B8, 3B10, 3C12 and 3D1), rather than testing individual combinations, to minimize wastage of precious ancient samples and maximize the chances of success. We first performed 56 PCR amplifications at limiting dilutions of ancient DNA (aDNA) derived from a 47,000-year-old cave bear (*Ursus spelaeus*) bone and scored successful amplifications for the blend and *Taq* alone. We found that the blend yielded amplification products at between two- and fivefold lower concentrations of aDNA than *Taq* did, and indeed yielded amplification products at DNA concentrations low enough that *Taq* no longer generated any (**Fig. 3b**).

Normalizing PCR activity on a dilution series of 'modern' DNA showed that this effect was not due to higher PCR efficiency of the blend. On the contrary, *Taq* appeared to be more than an order of magnitude more active at low template concentrations of modern DNA (**Fig. 3b**), suggesting that the blend requires more template than *Taq* does to produce an equivalent PCR signal. This suggests that the measured activity of the blend for the amplification of ancient DNA is likely to represent an underestimate of its true potential. Moreover, it implies that the blend can tap into a pool of DNA molecules that are inaccessible to *Taq*, presumably because they are damaged.

To stringently exclude sample heterogeneity and stochastic variation as the source of this effect, we performed a further 608 independent PCR amplifications from two different samples of cave bear bone (~47,000 and ~60,000 years old, respectively) and scored the number of PCR amplicons at limiting dilution (see **Supplementary Fig. 5** online). The blend yielded a larger number of amplicons (8–150%) than *Taq* in all but one experiment (**Table 1**), confirming our earlier results.

We cloned amplicons from experiment 5 (**Table 1**), sequenced independent clones (blend, 28; *Taq*, 26) and analyzed their sequences for systematic and sporadic errors. As ancient DNA PCR products may arise from a single template molecule, systematic errors (that is, deviations from the consensus sequence occurring in all clones from one PCR) mostly arise from lesions in the ancient DNA template. Sporadic errors, by contrast, occur at any point during the PCR amplification. Systematic errors therefore reflect lesion bypass, whereas sporadic errors largely reflect polymerase fidelity in PCR.

There seemed to be a higher incidence of systematic errors in the blend PCRs (5 blend/3 *Taq*). Although the numbers are small, we speculate that the difference may arise from the amplification products deriving from bypass of previously blocking lesions, broadly consistent with ~30% higher amplification success in experiment 5 (**Table 1**).

Blend amplicons also showed a 3.25-fold higher number of sporadic errors (14 blend/4 *Taq*). To better determine polymerase fidelity, we sequenced further amplicons (300 base pairs (bp); 10 blend/15 *Taq*) deriving from a different region of the cave bear mitochondrial genome. We found a 3.75-fold higher number of sporadic errors (15 blend/6 *Taq*), presumably reflecting the reduced fidelity of blend polymerases such as 3D1 and 3A10 (see **Supplementary Fig. 4a**,**b**).

Table 1 Ancient DNA amplification

Experiment	Sample[a]	Dilution	*Taq* PCR[b]	Blend PCR[b]	Improvement[c]
1	GS3-7	1/500	24/36	28/36	+16%
2	GS3-7	1/2,000	2/24	5/24	+150%
3	GS3-7	1/1,000	21/48	24/48	+14%
4	366	1/5	2/4	4/4	+100%
5	366	1/10	12/24	16/24	+33%
6	GS3-7	1/1,000	12/48	14/48	+16%
7	GS3-7	1/1,000	10/24	7/24	–30%
8	GS3-7	1/1,000	12/48	13/48	+8%
9	GS3-7	1/1,000	12/48	13/48	+8%
Total			107/304	124/304	+15%

[a]366 derives from the Herdengel cave (Austria; ~60,000 years old); GS3-7 derives from the Gamsulzen cave (Austria; ~47,000 years old). [b]Successful PCR amplifications/total PCR amplifications per experiment. [c]Improvement, blend versus *Taq* (%). Relative PCR success in the presence of limiting amounts of two different ancient cave bear DNA samples (GS3-7 and 366 (60,000 years old)). In eight out of nine experiments, the blend of mismatch polymerases yielded more amplicons than did *Taq*.

The polymerases described here are unique in combining the ability to bypass multiple lesions (for example, two abasic sites in PCR; **Supplementary Fig. 2a**) with robust PCR activity. Together with their ability to process non-cognate primer duplexes, this may contribute to the polymerases' ability to enhance the recovery of ancient DNA sequences.

Further improvements in polymerase performance appear possible but may require more insight into the bottlenecks that occur with ancient DNA recovery. We evolved bypass of just two classes of lesions that are known to occur in ancient DNA. Abasic sites are generated by spontaneous depurination or depyrimidination and as the end-product of various oxidation-induced DNA damage pathways[21,22]. High levels of oxidized pyrimidines, such as 5-hydroxy-5-methylhydantoin and 5-hydroxyhydantoin, have been found in ancient samples and are associated with PCR failure[20]. However, it is possible or even likely that other poorly understood forms of damage cause PCR failure. These may include intrastrand cross-links, which appear to be prevalent in older samples[23] and may be bypassed poorly, if at all, even by the selected polymerases with their tolerance for non-canonical primer-template duplex structures. In addition, PCR recovery may be limited by the presence of potent inhibitors, such as heme or polyphenolic acids, produced by the decomposition of organic matter. CSR should allow the selection of polymerases with additional improvements in their ability to bypass lesions and noncanonical DNA structures. These may be combined with polymerases selected for resistance to environmental inhibitors, as demonstrated previously for heparin[3].

It is difficult to infer from sequence data alone the molecular determinants of the notable abilities of 3A10 or 3D1 to extend double and quadruple mismatches and process misaligned primer template structures. The features that are most consistently shared among the selected polymerases (for example, mutations L33P, E76K and D145G; see **Supplementary Figs. 1** and **2**) all implicate the 5′-3′ exonuclease domain. In most selected polymerases, this domain derives from *Tth*. Although we found that the *Tth* exonuclease domain is more thermostable than its *Taq* counterpart (F.J. Ghadessy and P.H., unpublished results) and may therefore promote evolvability through increased tolerance of destabilizing mutations as described[24], there may be other factors contributing to its universal selection. At least in the case of the previously described polymerase M1 (ref. 4), we found that mutations in the 5′-3′ exonuclease domain contributed substantially to mismatch extension (M.d'A. and P.H., unpublished results), suggesting that this domain may contribute, in a manner that is not yet understood, to the processing of non-cognate 3′ ends.

In conclusion, molecular breeding and directed evolution by CSR have allowed us to isolate polymerases that enhance the recovery of genetic material from Pleistocene specimens, presumably as a result of their ability to amplify damaged DNA. Polymerases such as these should improve the recovery of ancient DNA and reduce bias towards modern DNA contamination. They should also be suited for direct sequencing approaches[8,12], as they are pre-adapted to emulsion PCR[3]. Polymerases capable of amplifying damaged DNA have applications and impact beyond paleobiology, for example, in archeology, historic and forensic medicine and the genetic analysis of clinical specimens damaged by preservatives, cancer drugs or ionizing radiation.

METHODS

DNA manipulation and protein expression. *Tth* and *Tfl* polymerase genes were cloned from *Thermus thermophilus* (*Tth*) and *Thermus flavus* (*Tfl*) genomic DNA (DSMZ) using gene-specific primers 1–4 (oligonucleotide sequences are provided in **Supplementary Methods** online), cloned into pASK75 and assayed for PCR activity as described[3]. Polymerase libraries were prepared by molecular breeding of polymerase genes. In molecular breeding, homologous genes from different organisms (orthologs) are recombined to yield a library of chimeras comprising segments of the different orthologs. Molecular breeding samples only functional diversity and therefore molecular breeding libraries often comprise a larger number of active clones than random mutant libraries. Genes encoding *Taq*, the *Taq* mutant T8 (ref. 3), *Tth* and *Tfl* were recombined using the staggered extension protocol (StEP) as described[25]. In StEP, genes to be recombined are PCR amplified with common flanking primers but using extension times that are too short to allow complete primer extension during each cycle of PCR. This promotes template switching between homologous regions, leading effectively to recombination between the genes. Tuning the extension time allows some control over the length of gene segments that are swapped. Here, equal concentrations of polymerase genes were cycled 40 times (94 °C, 30 s; 55 °C, 1 s) using primers 5 and 6. The product was gel-purified and reamplified with primers 7 and 8 and then cloned using *Xba*I and *Sal*I into pASK75 to create library 3T (1×10^9 c.f.u., 70% active clones). Expression of polymerases for characterization and ancient DNA PCR was as described[3,4] but using a 16/10 Hi-Prep heparin FF Column (Amersham Pharmacia Biotech) to purify heat-cleared (50 °C, 30 min) Bugbuster (Novagen) lysate with Complete EDTA-free protease inhibitor cocktail (Roche). Polymerase fractions eluted around 0.3 M NaCl and were concentrated and diafiltered into 50 mM Tris, pH 7.4, 1 mM DTT, 50% glycerol and stored at –20 °C.

Selection and screening. Emulsification and CSR selection were performed as described[3,26] using either matched primers 5 and 6 or single-mismatch primers 9 and 10 (ref. 4), double-mismatch primers 11 and 12, and quadruple-mismatch primers 13 and 14; cycled 20 times (94 °C, 1 min; 50 °C, 1 min; 72 °C, 8 min); reamplified with out-nested primers 5 and 6 or with gene-specific primers 1–4 and 15 or combinations thereof; and recloned as above. After selection rounds one and two, clones were screened by mismatch PCR (94 °C 30 s, 55 °C 30 s, 72 °C 1 min) with primers 5,6, (20 cycles (20×)) or 9,10 (30×) or 11,12 (30×) or 13,14 (50×) abasic site bypass PCR with primers 16,17 (25×) and by polymerase ELISA as described[17] but using hairpins 18,19. Promising clones from rounds one and two were StEP shuffled and backcrossed with parent polymerase genes. Clones analyzed in more detail in this report derive from selection round 3. Mutation rates were determined using the mutS ELISA assay[27] (Genecheck) according to manufacturer's instructions.

Primer extension. Synthesis of 5-hydroxyhydantoin phosphoramidite is described in **Supplementary Methods** online. DNA primers and templates substrates were prepared as described[28] by annealing single-stranded circular DNA with ^{32}P-labeled primers at a 1.5:1 molar ratio in annealing buffer (50 mM Tris-HCl (pH 8), 5 mM $MgCl_2$, 50 μg/ml BSA, 1.42 mM 2-mercaptoethanol) for 10 min at 100 °C followed by slow cooling to 20 °C over 2 h. Annealing efficiencies were >95%, as evidenced by the different mobility of the ^{32}P-labeled primers before and after hybridization to the template on non-denaturing polyacrylamide gels. Primer 18 was used to study abasic and 5-hydroxyhydantoin bypass on templates 21–23. To study bypass of 5-methyl-5-hydroxyhydantoin, primer 24 and templates 25 and 26 were used. To study quadruple-mismatch extension, primers 27–30 and templates 31–33 were used. Standard extension reactions contained 10 nM DNA templates (expressed as primer termini), 100 μM of either all four dNTPs or each dNTP individually, 40 mM Tris-HCl at pH 8.0, 5 mM $MgCl_2$, 10 mM dithiothreitol, 250 μg/ml bovine serum albumin; 2.5% glycerol and various amounts of DNA polymerases. Reactions were incubated for 10 min at 65 °C unless otherwise specified. The reactions were terminated by mixing with one volume of formamide loading dye solution containing 50 mM EDTA, 0.1% xylene cyanol and 0.1% bromophenol blue in 90% formamide. Before being loaded onto the gel, the reactions were denatured by heating at 100 °C for 10 min and immediately transferred into ice for 2 min. Products were resolved by denaturing PAGE (8 M urea, 15% acrylamide, 3 h at 2,000 V) and then visualized and quantified using a Fuji image analyzer FLA-3000 and Multi-Gauge software. The levels of primer extension and primer elongation past the damaged or corresponding undamaged sites (undamaged T, abasic site or hydantoin) are expressed as a percentage of the total primer termini.

© 2007 Nature Publishing Group http://www.nature.com/naturebiotechnology

Ancient DNA. aDNA was extracted as described[29]. Essentially, bone or teeth were ground with mortar and pestle. 10 ml extraction buffer containing 0.45 M EDTA (pH 8), 0.5% *N*-lauroylsarcosine, 1% polyvinylpyrrolidone, 50 mM DTT, 2.5 mM PTB and 0.25 mg/ml proteinase K were added to 200 mg–1 g of bone powder and incubated for 16 h at 37 °C under rotation. The remaining bone powder was collected by centrifugation and only the supernatant was used for further processing. aDNA was purified by binding to silica. 40 ml of L2 buffer (5.5 M guanidinium isothiocyanate, 25 mM NaCl, 100 mM Tris (pH 8)) and 50 µl of silica suspension were added to 10 ml supernatant and incubated for approximately 30 min. The pellet was collected by brief centrifugation, the supernatant discarded, and the silica pellet washed in buffer L2 and once with NewWash (Bio 101). After drying of the pellet, the DNA was eluted at 56 °C in aliquots of 100 µl TE (10 mM Tris pH 7.4, 1 mM EDTA) Mock extractions were performed alongside all extractions. The final volume of the extract was 100 µl. aDNA was amplified by two-step PCR. Briefly, 2 µl of ancient sample were added to a 20 µl PCR in SuperTaq buffer (HT Biotech) with 1 µM primers, 2 µM dNTPs and 0.5 U of SuperTaq or blend and amplified for 28 cycles using primers 34 and 35 (PCR1). This PCR was set up in a clean room following precautions appropriate for aDNA. PCR1 was then diluted 1/20 in a secondary clean room and reamplified for 32 cycles using in-nested primers 36 and 37 and standard PCR parameters. Blend was prepared by mixing polymerases (activity normalized on undamaged templates) in a ratio of 90% SuperTaq/10% mutant polymerases (equivalent amounts of 3A10, 3B5, 3B6, 3B8, 3B10, 3C12 and 3D1). Amplifications using SuperTaq alone were compared with amplifications using blend in PCR1 and SuperTaq in PCR2. No-template controls were always included to detect contamination.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS
M.d'A. was supported by a Medical Research Council studentship and a Junior Research Fellowship from Trinity College, Cambridge, UK. A.V. and R.W. were supported by funds from the NICHD/NIH Intramural Research Program.

AUTHOR CONTRIBUTIONS

M.d'A. contributed to library construction, CSR selection and, together with M.H. and S.P., to ancient DNA amplification. M.d'A., P.H., A.V. and R.W. contributed to polymerase characterization using template lesions synthesized by D.L., D.G. and J.C. P.H. contributed to the planning and design of the project and manuscript writing.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturebiotechnology/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Pääbo, S. *et al.* Genetic analyses from ancient DNA. *Annu. Rev. Genet.* **38**, 645–679 (2004).
2. Crameri, A., Raillard, S.A., Bermudez, E. & Stemmer, W.P. DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288–291 (1998).
3. Ghadessy, F.J., Ong, J.L. & Holliger, P. Directed evolution of polymerase function by compartmentalized self-replication. *Proc. Natl. Acad. Sci. USA* **98**, 4552–4557 (2001).
4. Ghadessy, F.J. *et al.* Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution. *Nat. Biotechnol.* **22**, 755–759 (2004).
5. Leonard, J.A. *et al.* Ancient DNA evidence for Old World origin of New World dogs. *Science* **298**, 1613–1616 (2002).
6. Jaenicke-Despres, V. *et al.* Early allelic selection in maize as revealed by ancient DNA. *Science* **302**, 1206–1208 (2003).
7. Anderung, C. *et al.* Prehistoric contacts over the Straits of Gibraltar indicated by genetic analysis of Iberian Bronze Age cattle. *Proc. Natl. Acad. Sci. USA* **102**, 8431–8435 (2005).
8. Green, R.E. *et al.* Analysis of one million base pairs of Neanderthal DNA. *Nature* **444**, 330–336 (2006).
9. Krings, M. *et al.* Neandertal DNA sequences and the origin of modern humans. *Cell* **90**, 19–30 (1997).
10. Geigl, E.M. Inadequate use of molecular hybridization to analyze DNA in Neanderthal fossils. *Am. J. Hum. Genet.* **68**, 287–291 (2001).
11. Plosky, B.S. & Woodgate, R. Switching from high-fidelity replicases to low-fidelity lesion-bypass polymerases. *Curr. Opin. Genet. Dev.* **14**, 113–119 (2004).
12. Poinar, H.N. *et al.* Metagenomics to paleogenomics: large-scale sequencing of mammoth DNA. *Science* **311**, 392–394 (2005).
13. Huang, M.-M., Arnheim, N. & Goodman, M.F. Extension of base mispairs by *Taq* polymerase: implications for single nucleotide discrimination in PCR. *Nucleic Acids Res.* **20**, 4567–4573 (1992).
14. Miller, H. & Grollman, A.P. Kinetics of DNA polymerase I (Klenow fragment exo-) activity on damaged DNA templates: effect of proximal and distal template damage on DNA synthesis. *Biochemistry* **36**, 15336–15342 (1997).
15. Suzuki, M., Baskin, D., Hood, L. & Loeb, L.A. Random mutagenesis of *Thermus aquaticus* DNA polymerase I: concordance of immutable sites in vivo with the crystal structure. *Proc. Natl. Acad. Sci. USA* **93**, 9670–9675 (1996).
16. Xia, G. *et al.* Directed evolution of novel polymerase activities: mutation of a DNA polymerase into an efficient RNA polymerase. *Proc. Natl. Acad. Sci. USA* **99**, 6597–6602 (2002).
17. Ong, J.L., Loakes, D., Jaroslawski, S., Too, K. & Holliger, P. Directed evolution of DNA polymerase, RNA polymerase and reverse transcriptase activity in a single polypeptide. *J. Mol. Biol.* **361**, 537–550 (2006).
18. Matsuda, T. *et al.* Error rate and specificity of human and murine DNA polymerase eta. *J. Mol. Biol.* **312**, 335–346 (2001).
19. Frank, E.G. *et al.* Altered nucleotide misinsertion fidelity associated with poliota-dependent replication at the end of a DNA template. *EMBO J.* **20**, 2914–2922 (2001).
20. Höss, M., Jaruga, P., Zastawny, T.H., Dizdaroglu, M. & Pääbo, S. DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Res.* **24**, 1304–1307 (1996).
21. Lindahl, T. & Karlstrom, O. Heat-induced depyrimidination of deoxyribonucleic acid in neutral solution. *Biochemistry* **12**, 5151–5154 (1973).
22. Lindahl, T. & Nyberg, B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11**, 3610–3618 (1972).
23. Hansen, A.J. *et al.* Crosslinks rather than strand breaks determine access to ancient DNA sequences from frozen sediments. *Genetics* **173**, 1175–1179 (2006).
24. Bloom, J.D., Labthavikul, S.T., Otey, C.R. & Arnold, F.H. Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. USA* **103**, 5869–5874 (2006).
25. Zhao, H., Giver, L., Shao, Z., Affholter, J.A. & Arnold, F.H. Molecular evolution by staggered extension process (StEP) *in vitro* recombination. *Nat. Biotechnol.* **16**, 258–261 (1998) [see comments].
26. Ghadessy, F.J. & Holliger, P. Compartmentalized self-replication: a novel method for the directed evolution of polymerases and other enzymes. *Methods Mol. Biol.* **352**, 237–248 (2007).
27. Debbie, P. *et al.* Allele identification using immobilized mismatch binding protein: detection and identification of antibiotic resistant bacteria and determination of sheep susceptibility to scrapie. *Nucleic Acids Res.* **25**, 4825–4829 (1997).
28. McDonald, J.P. *et al.* Novel thermostable Y-family polymerases: applications for the PCR amplification of damaged or ancient DNAs. *Nucleic Acids Res.* **34**, 1102–1111 (2006).
29. Hofreiter, M. *et al.* Evidence for reproductive isolation between cave bear populations. *Curr. Biol.* **14**, 40–43 (2004).
30. Tawfik, D.S. & Griffiths, A.D. Man-made cell-like compartments for molecular evolution. *Nat. Biotechnol.* **16**, 652–656 (1998).

© 2007 Nature Publishing Group http://www.nature.com/naturebiotechnology