Published online 14 January 2015 *Nucleic Acids Research, 2015, Vol. 43, No. 3* ***1637–1645***
doi: 10.1093/nar/gku1398

DNA lesion identity drives choice of damage tolerance pathway in murine cell chromosomes

Isadora S. Cohen[1], Carmit Bar[1], Tamar Paz-Elizur[1], Elena Ainbinder[2], Karoline Leopold[1], Niels de Wind[3], Nicholas Geacintov[4] and Zvi Livneh[1,*]

[1]Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel, [2]Stem Cell Core Unit, Department of Biological Services, Weizmann Institute of Science, Rehovot 76100, Israel, [3]Department of Human Genetics, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands and [4]Chemistry Department, New York University, New York, NY 10003-5180, USA

Received November 23, 2014; Revised December 24, 2014; Accepted December 27, 2014

ABSTRACT

DNA-damage tolerance (DDT) via translesion DNA synthesis (TLS) or homology-dependent repair (HDR) functions to bypass DNA lesions encountered during replication, and is critical for maintaining genome stability. Here, we present piggyBlock, a new chromosomal assay that, using piggyBac transposition of DNA containing a known lesion, measures the division of labor between the two DDT pathways. We show that in the absence of DNA damage response, tolerance of the most common sunlight-induced DNA lesion, TT-CPD, is achieved by TLS in mouse embryo fibroblasts. Meanwhile, BP-G, a major smoke-induced DNA lesion, is bypassed primarily by HDR, providing the first evidence for this mechanism being the main tolerance pathway for a biologically important lesion in a mammalian genome. We also show that, far from being a last-resort strategy as it is sometimes portrayed, TLS operates alongside nucleotide excision repair, handling 40% of TT-CPDs in repair-proficient cells. Finally, DDT acts in mouse embryonic stem cells, exhibiting the same pattern—mutagenic TLS included—despite the risk of propagating mutations along all cell lineages. The new method highlights the importance of HDR, and provides an effective tool for studying DDT in mammalian cells.

INTRODUCTION

DNA repair mechanisms, though highly efficient, cannot completely eliminate DNA damage, that is estimated to occur at a rate of tens of thousands of lesions in each mammalian cell, every day (1). This has particular implications for DNA replication during S phase, as constant lesion formation renders the encounter of the replication machinery with damaged bases inevitable. When this happens, the completion of chromosome replication depends upon processes collectively labeled DNA damage tolerance (DDT) (2–4). Two classes of damage tolerance mechanisms are known: translesion DNA synthesis (TLS) and homology-dependent repair (HDR) (5). In TLS, the lesion is bypassed via synthesis of DNA across it by specialized DNA polymerases, while in HDR the missing sequence information opposite the lesion is obtained from the intact nested sister chromatid. Not much is known about the division of labor between the two pathways in mammals.

Much of the study of DNA damage repair and tolerance is carried out by treating cells with DNA damaging agents and quantifying their effect on aspects of the cell's life such as viability, mutation load, genome integrity or replication progression. To obtain a quantifiable population-level effect, treatment must exceed a certain threshold, that often lies beyond common real-life exposure levels, and that triggers activation of DNA damage response signaling. Such approaches are therefore ill suited to the study of low level, sporadic DNA damage. This challenge can be addressed by functional assays in which sequencing the bypass outcome of individual known lesions integrated into chromosomal DNA helps identify the DDT mechanism involved. Recent work in *Escherichia coli* (6) and human cells (5) demonstrated the feasibility of this approach. Here we present piggyBlock, a piggyBac transposition-based system for the chromosomal integration of replication-blocking lesions. This new assay system has the advantages of highly efficient integration and of a broad, hot spot-free integration locus spectrum (7–9). Its flexible integration cassette design is another improvement from a phage-derived system (5,10) that promotes whole plasmid loop-in. We use piggyBlock

*To whom correspondence should be addressed. Tel: +972 8 934 3203; Fax: +972 8 934 4169; Email: zvi.livneh@weizmann.ac.il
Present addresses:
Carmit Bar, Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029-6574, USA.
Karoline Leopold, Department of Molecular Biology, University of Geneva, 30 Quai Ernest Ansermet, Geneva 1211, Switzerland.

to transpose DNA containing known replication-blocking lesions into cultured cells' chromosomes and isolate individual DDT events via clonal selection. Using this single cell–single event assay system, we show that in murine cells tolerance of different lesions is achieved by distinct DDT pathways, and that this occurs in the absence of exogenous stress and DDR signaling. We investigate damage tolerance of two representative DNA lesions, cyclobutane pyrimidine dimer (CPD) and benzo[*a*]pyrene-guanine (BP-G), each implicated in one of the best-known environmentally caused cancers. CPD is the most common lesion caused by ultraviolet (UV) radiation from sunlight or other sources, while BP-G is a product of smoke. We show that these lesions are subject to very different DDT treatments in chromosomes. While a single thymine–thymine cis-syn cyclobutane pyrimidine dimer (TT-CPD) in a mouse genome is bypassed by TLS, tolerance of a single BP-G is achieved in most cases by HDR.

Embryonic stem (ES) cells are the starting material of every multi-cellular organism, and as such will transmit any mutations and chromosomal aberration to all tissues. Therefore, extremely stringent management of DNA damage may seem advantageous for these cells and indeed, ES cells exhibit lower survival rates following treatment with several genotoxic agents than that observed in mouse embryonic fibroblasts (MEFs) (11–14). Genomic stability of the surviving cells depends upon DNA damage repair and tolerance mechanisms. *Rad54*, a possible HDR participant, was implicated in replication fork recovery in murine ES (mES) cells following UV irradiation, in particular in nucleotide excision repair (NER)-deficient cells, but the precise identity of the UV-induced lesion that is its substrate is not known (15). Here, we show that DDT by both TLS and HDR occurs in mES cells, and exhibits characteristics very similar to those observed in MEFs, including mutagenic TLS.

MATERIALS AND METHODS

Lesion plasmid construction

To construct piggyBlock lesion plasmids, lesion cores were ligated into piggyBlock (Supplementary Data S1), a vector derived from 5′-PTK-3′ piggyBac (7). In the construction of this vector, restriction sites for BpiI and SfiI, ∼1 kb apart, were introduced, which are used for lesion core introduction. Lesions were (+)-*trans*-BPDE-N^2-dG (BP-G) and TT-CPD. To generate each lesion-containing duplex DNA insert, one lesion-containing oligo and five supporting oligos, or two lesion-containing oligos and four supporting oligos were used. To assemble the lesion core, 100 pmol of each oligo were phosphorylated using polynucleotide kinase and adenosine triphosphate. The six oligos were then combined at final 1 μM each and annealed in the presence of 125 mM NaCl by heating to 75°C for 10 min and cooling to 4°C at 1°C/min in a Biometra thermal cycler. The annealed lesion core was combined with 100 μg (∼30 pmol) of restriction-digested piggyBlock vector at 1:2 vector–insert ratio and ligated in final 5 ml using 2×10^4 units T4 DNA ligase New England Biolabs (NEB), overnight at 4°C. The ligation product was ethanol precipitated and the supercoiled fraction excised from agarose gel.

To assemble the lesion core for the piggyBlockBP-G1 single lesion plasmid, we used lesion oligo BP-G1: CAT[BP-G]CGTCCTAC. Supporting oligos were oIC67: AATGCGATCTGAC, oIC68: CAGTGGAATATC-TAGTGTAGGACGtATGCTCCTTGAACGCGGT, oIC70: GCGTTCAAGGAG, oIC71: ACTAGATATTC-CACTGCACGTGACGAACGTCAGATCG and oIC82: GTTCGTGACGTG. The lesion core for piggyBlockBP-G3 was created using lesion oligo BP-G3: GTTCGT[BP-G]ACGTG. Supporting oligos were oIC67, oIC68, oIC70, oIC71 and oIC83: CATGCGTCCTAC. Lesion oligos used in the construction of the piggyBlockBP-G1–3 dual-lesion plasmid were BP-G1 and BP-G3. Supporting oligos were oIC67, oIC68, oIC70 and oIC71. To assemble the single lesion duplex for piggyBlockTT-CPD1, we used lesion oligo TT-CPD1: TGCGA[TT]GCACG and supporting oligos oIC132: AATGCGAGGACTG, oIC133: GAGATG-GAATATCTAGTCGTGCGGTCGCATCCCAAG-GATGCGGT, oIC135: GCATCCTTGGGA, oIC136: ACTAGATATTCCATCTCTGCGCCCTCACGCAGTC-CTCG and oIC151: CGTGATTGCGCA. Assembly of the single lesion duplex for piggyBlockTT-CPD2 used lesion oligo TT-CPD2 CGTGA[TT]GCGCA and supporting oligos oIC132, oIC133, oIC135, oIC136 and oIC150: TGCGATTGCACG. To generate lesion cores for piggyBlockTT-CPD dual plasmids, lesion oligos CPD1 and CPD2 were used. Supporting oligos in the lesion core for piggyBlockTT-CPDCF were oIC132, oIC133, oIC135 and oIC136. piggyBlockTT-CPDFC supporting oligos were oIC138: AATGCCATGACAT, oIC139: CTCTGGTTTATCTAAGGCGTGCGGTCG-CACGGGTTGAACGCGGT, oIC141: GCGTTCAAC-CCG and oIC142: CCTTAGATAAACCAGAGTGCGC-CCTCACGATGTCATGG.

Transposase-encoding helper plasmids were mPB (7) and HyPB (16).

Cell culture and transfection

Xpa$^{-/-}$ MEFs were cultured in Dulbecco's modified Eagle's medium (DMEM; GIBCO/BRL) supplemented with 10% fetal bovine serum (FBS; HyClone), 100 units/ml penicillin and 100 μg/ml streptomycin (Biological Industries). DR-4 irradiated, puromycin-resistant mouse embryonic fibroblasts (iMEFs) prepared by the WIS stem cell unit served as feeder layer for cultivating mESC. Feeder layers were cultivated on 0.1% gelatin- (Sigma) coated plates in DMEM supplemented with 10% FBS, 2 mM L-alanyl L-Gln (Biological Industries), sodium pyruvate (Biological Industries) and 100 units/ml penicillin and 100 μg/ml streptomycin. Neomycin- and hygromycin-resistant *Xpa*$^{-/-}$ mES cells were cultivated in DMEM supplemented with FBS 15%, non-essential amino acid solution (Biological Industries), 2 mM L-alanyl L-Gln, β-mercaptoethanol (GIBCO/BRL), 10ng/ml Leukemia inhibitory factor (LIF; Peprotech), CHIR99021 (3 μM, GSK3βi, Axon Medchem) and PD0325901 (1 μM, ERK1/2i, Axon Medchem). The cells were incubated at 37°C in a 5% CO_2 atmosphere and periodically examined for mycoplasma contaminations by EZ-PCR test kit (Biological Industries).

Single lesion piggyBlock constructed plasmids were transfected into MEFs in 10 cm culture dishes using Jet PEI (Polyplus). Each dish was transfected with 10 ng of piggyBlock constructed lesion plasmid and 1 μg HyPB helper plasmid (16). Puromycin selection (1 μg/ml) was administered 24 h post-transfection. Transfection of dual piggyBlock plasmids was performed in six-well format. Each well was transfected with 50 ng of constructed piggyBlock lesion plasmid and 200 ng mPB helper plasmid (7,16). After 48 h, the cells from each well were sub-cultured in puromycin (Sigma, final 1 μg/ml) medium in four 10 cm culture dishes. After 7–10 days of selection, individual colonies were each manually removed into a well of a 96-well culture plate and further cultivated to full confluence.

Mouse ES cells were transfected in gelatinized 12-well plates without feeder layers, in 1:1 iMEF- and mESC-conditioned medium with LIF, CHIR99021 and PD0325901. Cells were transfected using Xfect (Clontech) with 50 ng piggyBlock constructed lesion plasmid and 1 μg mPB helper plasmid. Cells from each well were sub-cultured 24 h post-transfection into two 10 cm culture dishes containing puromycin resistant feeder layers and selection was performed with 1.8 μg/ml puromycin for 4–6 days. Mouse ES colonies were isolated using AVISO Mechatronic Systems' CellCelector. Clones were cultivated in 96-well plates, in 1:1 iMEF- and mESC-conditioned medium without feeder layer, for 2 days prior to harvesting.

DNA isolation, amplification and sequencing

Genomic DNA were isolated in 96-well format from frozen cell pellets, using Epicentre's (Illumina) QuickExtract. Nested Genomic polymerase chain reaction (PCR) in 96-well format was performed using Phire hot start II DNA polymerase from Thermo Scientific (Finnzymes). Primer sequences for the first amplification were oIC129: GTCGCTGTGCATTTAGGACA and oIC131: CAGAAAGCGAAGGAGCAAAG. Second amplification primer sequences were oIC14: CTTCCATTTGTCACGTCCTG and oIC130: GCTTGTCAATGCGGTAAGTG.

To produce inverse PCR (iPCR) templates, 20 μl of genomic DNA were digested with restriction enzyme Csp6I in a final volume of 50 μl, then self-ligated for 2 h at room temperature. Primers for the first iPCR amplification were oIC130 and oIC 269: CGACTGAGATGTCCTAAATGCAC. Second iPCR amplification primer sequences were oIC265: TGTCCTAAATGCACAGCGAC and oIC270: GACCAATTGGAAGACCCAAT. Products were precipitated with polyethylene glycol prior to sequencing as previously described (17). Fluorescent labeling was performed using BigDye terminator 1.1 (Invitrogen) with sequencing primer oIC106: GGGGAACTTCCTGACTAGGG for regular PCR and oIC263: TTAGAAAGAGAGAGCAATATTTCAAGAATG for iPCR. Sequences were read using Applied Biosystems' AB3130XL Genetic Analyzer. iPCR products were analyzed using the iMapper software (18) and the chromosomal map (Figure 1D) was produced using Ensembl (19).

RESULTS

piggyBlock, a chromosomal replication block assay system

piggyBlock is a new assay system that introduces DNA carrying specific lesions in a known sequence context into the chromosomes of mammalian cells in culture, via piggyBac transposition. It was developed as a tool for investigating DDT at the sequence level. piggyBlock's plasmid-borne chromosomal integration cassette consists of piggyBac terminal repeats flanking a single known DNA lesion, or pair of lesions positioned 30 base pair apart on opposite strands (Figure 1A and B). Also in the transposition cassette is the gene encoding the puromycin resistance protein for selection of stably transfected clones (Figure 1A). These were transfected, along with a helper plasmid encoding the piggyBac transposase protein, into murine cells. Transposase expressed from the helper plasmid then integrated the lesion-containing DNA into a chromosomal locus (Figure 1A) in a quasi-random manner (8,9). After transfection, cells were subjected to puromycin selection and clones containing individual lesion bypass events propagated and isolated, from which genomic DNA was extracted. This DNA was used as template for regular PCR or iPCR (20), and subsequent sequence analysis (Figure 1C). To exclude the possibility of lesions being removed by the NER pathway, experiments were conducted using cells that lack a functional *Xpa* gene, known to be indispensable for NER activity (21).

To discriminate between TLS and HDR, the nucleotide placed opposite the lesion was one which was previously shown to be rarely inserted by TLS (22,23). TLS most frequently inserts the correct nucleotide—dCMP—opposite BP-G (accurate TLS) and, at a lower incidence, dAMP (mutagenic TLS). We therefore engineered the rarely occurring T or G opposite BP-G (Figure 1B). Thus, synthesis by TLS through a BP-G:T pair is expected to most frequently restore the original G:C base pair (Figure 2, TLS branch), or introduce a T:A mutation in case of the less frequent mutagenic TLS. In contrast, HDR acquires the opposite strand nucleotide, leading to the formation of an A:T base pair at the lesion site (Figure 2, HDR branch). Each lesion was assayed using two different piggyBlock plasmids, each containing the damaged base in a different sequence context and with a different mismatch (Figure 1B). This ensures that the opposite strand signature representing HDR is indeed that, rather than a default signature unrelated to the identity of the base opposite the lesion. Cells were propagated in the transfection dish under puromycin selection to form mosaic colonies, each containing replication products of both the lesion-containing strand and the intact strand. TLS proceeds through the lesion strand and as such it preserves, when accurate, the identity of the lesion-carrying base. Therefore, by virtue of a damaged G being present opposite a T (or G), TLS events give rise to a double peak signature in the sequence file, as progeny of both the lesion strand and the intact, opposite strand are present at the same sequence position in the mosaic colony (Figure 2; sequence peaks of both A and G are seen at the location corresponding to the lesion). HDR, on the other hand, proceeds through the base opposite the lesion, so its signature

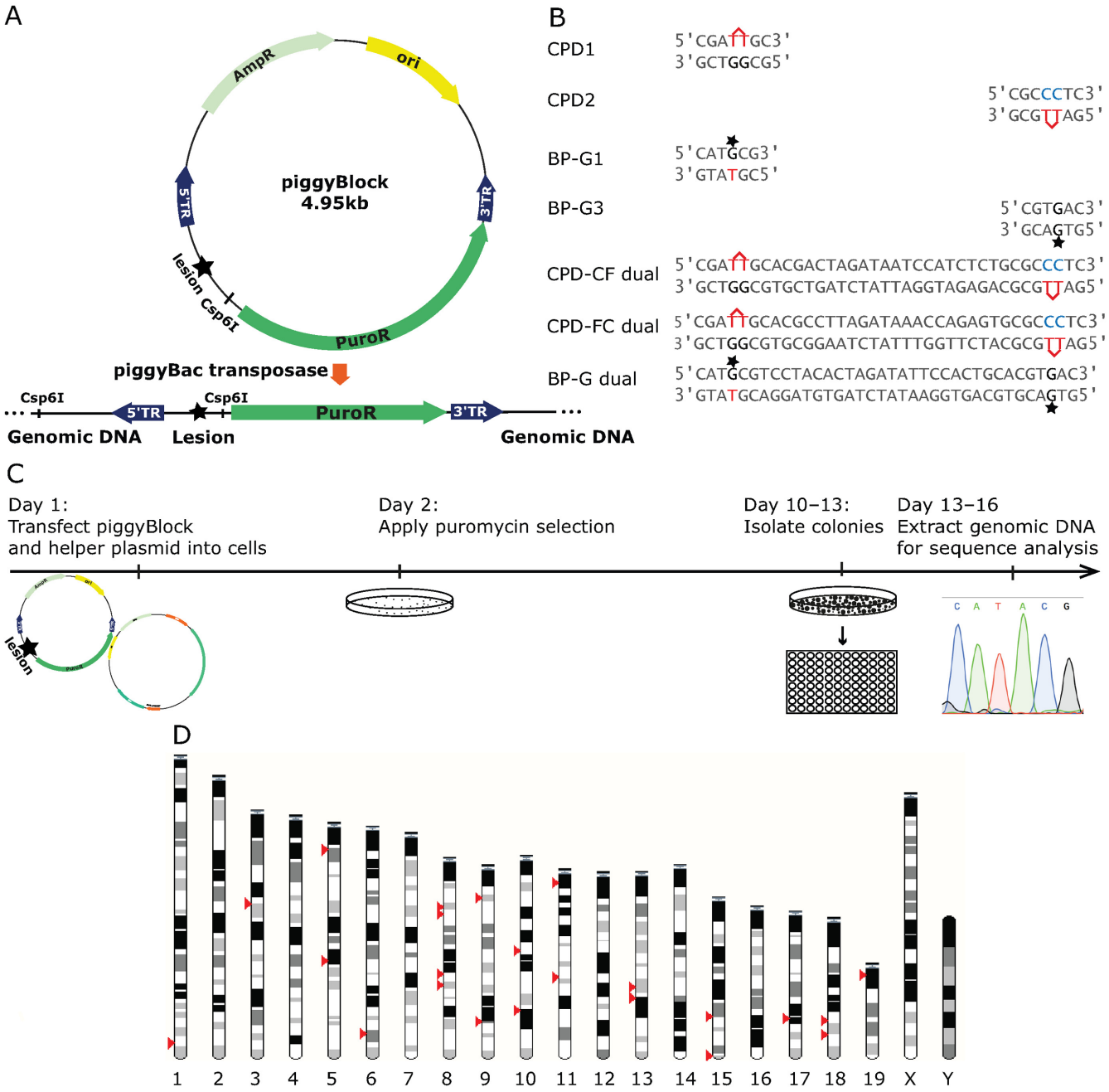

Figure 1. The piggyBlock assay system. (**A**) A cassette consisting of DNA lesion(s) and puromycin selection marker is transposed from the piggyBlock constructed plasmid into a chromosomal locus by piggyBac transposase expressed from a co-transfected helper plasmid. (**B**) Local sequences of double-stranded lesion oligos ligated into the piggyBlock vector to form the constructed lesion plasmid. A star represents a benzo[*a*]pyrene adduct and a right angle above two pyrimidines represents dimerization. (**C**) Experiment timeline: constructed piggyBlock lesion plasmid and helper plasmid were co-transfected into cells. 24–48 h later, puromycin selection was administered. Cells were maintained under selection for 4–10 days, until colonies formed. Individual colonies were then transferred into wells of 96-well culture dishes and cultivated to full confluence (2–5 days). Colonies were harvested, chromosomal DNA was isolated and sequence analysis was performed. (**D**) Chromosomal integration loci identified by iPCR.

consists of a single peak, the one complementary to the intact base opposite the lesion (Figure 2; a single sequence peak of A is present at the location corresponding to the lesion). When piggyBlock plasmids containing a lesion in each strand were transfected, cells were sub-cultured 48 h post-transfection, to allow the progeny of each strand to propagate separately into an individual colony. In this setup, sequences are predicted to always contain single peaks. Opposite strand peaks in the case of HDR and lesion strand peaks for TLS (5).

We performed iPCR (20) on a subset of our clones, to examine whether the lesion cassette is indeed present in a chromosomal location. The restriction site used in template preparation, Csp6I, was positioned in the piggyBlock

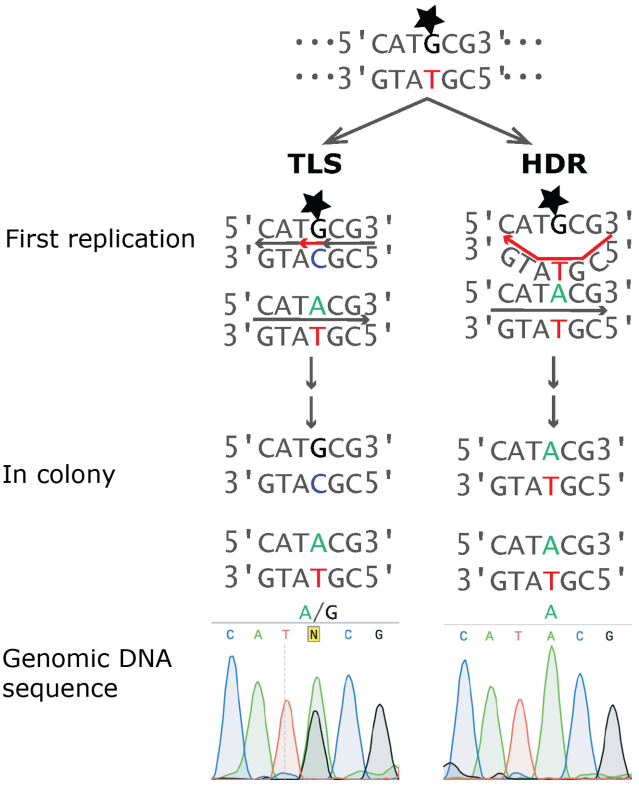

Figure 2. Lesion bypass pathways and corresponding sequence signatures. A non-complementary nucleotide that is rarely inserted by TLS was engineered opposite the damaged base(s). During the first-replication post-transfection, bypass by accurate TLS inserts the correct complementary base (C) opposite the damaged template base (BP-G; marked by a star), while replication on the opposite strand in the same cell places A opposite the template T. Further rounds of replication result in a mosaic colony, and the lesion position is visualized as a double peak in the sequence file. Bypass by HDR in the first replication places T in the same sequence position, that is inserted opposite the A in the nascent sister chromatid that serves as alternative template. This is visualized in the sequence file as a single peak, corresponding to the intact strand sequence.

integration cassette such that the amplicon spans the lesion core (Figure 1A), in addition to the chromosomal context in which lesion bypass took place (Figure 1D). This ensures that the sequenced bypass signatures are, indeed, present in a chromosomal context, rather than on a plasmid. We obtained genomic sequences from 27 clones. Two loci were observed twice, while the remaining 23 sequences were unique (Figure 1D and Table 1). The 3′ ends of 18 sequences contained the lesion core sequence, as predicted. In other clones, the genomic sequence was too long for the sequencing to reach the integration cassette at the end of the amplicon. In contrast to the ΦC31-mediated integration loci that we previously reported (5), no integration hotspot was observed in the piggyBlock-derived clones.

Choice of damage tolerance pathway for a single DNA lesion is governed by the identity of the lesion

CPDs are the main DNA lesions produced by UV light. While they disrupt aromaticity, they only confer minor distortion on the DNA strand (24). This lesion's excellent suitability as a TLS substrate (25,26) raises the question

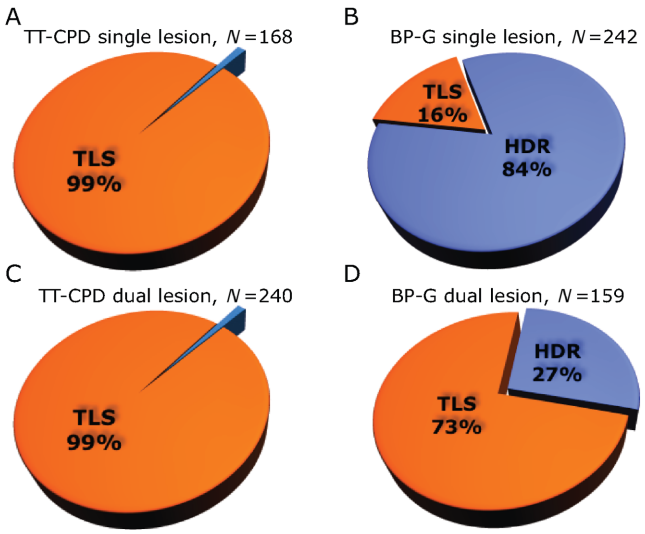

Figure 3. DNA sequence outcome of piggyBlock TLS/HDR experiments in MEF. (**A**) TT-CPD single lesion, $n = 168$. (**B**) BP-G single lesion, $n = 242$. (**C**) TT-CPD dual lesion, $n = 240$. (**D**) BP-G dual lesion, $n = 159$.

whether it also undergoes HDR, thus making it an intriguing candidate for the piggyBlock system. We scored 82 accurate TLS signatures across TT-CPD in the CPD1 data set and 86 in the CPD2 data set, while only 1 (<1%) signature consistent with HDR was observed in the CPD2 data set (Figure 3A and Supplementary Table S1). Thus, not only is TLS an excellent mechanism for CPD tolerance, our results show it to be essentially the only DDT mechanism acting on this lesion, confirming at the molecular level the phenotypic observation that tolerance of CPD crucially depends upon TLS.

When we assayed BP-G in the same manner, 84% of the sequences exhibited HDR signatures (96/109 in BPG1 and 107/133 in BPG3, total 203/242, Figure 3B and Supplementary Table S2). Of the remaining 39 sequences (13 BP-G1 and 26 BP-G3), that represented TLS, five (13%) were mutagenic. This corresponds to about half the BP-G mutagenicity level observed in MEFs when assayed with an episomal lesion plasmid, that is a substrate for TLS but not for HDR (22,27). All five mutagenic TLS events contained the A signature characteristic of mutagenic TLS across BP-G (Supplementary Table S2), recapitulating previous observations both in mouse and in human (22,27). Our results show that the DDT pattern across BP-G is nearly diametrically opposed to that employed for CPD. Surprisingly, once assayed alongside HDR, TLS across BP-G by DNA polymerases κ and ζ (22) turns out to be a secondary tolerance pathway, not the main one for this lesion.

In interpreting sequences containing mixed peaks, we considered the fact that by virtue of the assay being conducted in NER-deficient cells, lesions that are integrated into the cells' chromosomes persist through the first replication and must be bypassed again during subsequent replications. Since lesions are not propagated, the contribution of iterative lesion bypass to the colony's population is diluted by a factor of two with each generation. If TLS occurred in the first replication, subsequent events will not affect the

Table 1. Loci of lesion cassette piggyBac-mediated chromosomal integration, identified by iPCR

Chromosome	Position	Gene within 10 kb	Occurrences
1	185655149	None	1
3	60422572	None	1
5	17598155	Sema3c	1
5	89035474	Slc4a4	2
6	51758609	None	1
6	133597533	None	1
8	31845795	Nrg1	1
8	36613434	Dlc1	1
8	75083660	Hmox1	2
8	82161734	U7	1
9	21330794	Slc44a2	1
9	100660922	Stag1	1
10	41822649	Sesn1	1
10	61285408	None	1
10	99599438	None	1
11	8951120	Pkd1l1	1
11	69934251	Ybx2,Slc2a4	1
13	74390171	None	1
13	81355882	Gpr98	1
15	77182826	Rbfox2	1
15	102105528	Eif4b,Tenc1	1
17	69416583	C030034I22Rik	1
18	66237552	Ccbe1	1
18	75386903	Smad7	1
19	7484022	Rtn3	1

bypass signature. Whether an additional TLS event recapitulates the original TLS signature or a HDR event copies it, the resulting sequence will exhibit the same double peak (Figure 2, TLS branch). However, a TLS event occurring during the second replication will leave a mark on a first-replication HDR sequence, adding a smaller but still visible lesion-strand peak to the opposite-strand peak indicative of HDR (Supplementary Figure S3). Therefore, the determination of the extents of TLS and HDR relies on the ability to discern whether a mixed-peak sequence should be taken as a TLS event or a 'contaminated' HDR event. To this end, we performed a simulation of the sequence signatures expected from various TLS/HDR ratios (Supplementary Figures S3 and S4 and Supplementary Data S2). The detection threshold for a sequence variant by the capillary sequencing technology that we used is 10% of a mixed-sequence DNA sample (28). This corresponds in the simulation to the third generation post-transposition. We analyzed our data using the QSV Analyser software (28), that calculates allele ratio from Sanger sequencing output peak heights. We counted QSV Analyser peak ratios of 0.7:0.3 or closer to equal as representing TLS events. Opposite-strand to lesion-strand peak ratios from 0.8:0.2 to 1:0 were interpreted as HDR. Of the sequences reported above, 79% of the total (86 BP-G1 sequences and 106 BP-G3 sequences) contained pure opposite strand peaks (peak ratio 1:0 in QSV Analyser output), indicating that in these clones the lesion was bypassed by HDR in all three replications detectable by the system (28). Of the remaining, mixed-peak sequences, 10 in the BP-G1 set and one in the BP-G3 set (4% of the total) were additional HDR events that contained residual lesion-strand peaks. Quantification by QSV Analyser confirms our direct observation, based on peak appearance, that the effect of repeated lesion bypass on the results is small.

TLS is the preferred tolerance pathway for clustered lesions in opposite strands

DNA damage can occur as isolated events, as in the case of the single synthetic lesions described above, or as clusters of two or more lesions (29). The latter configuration may affect the availability of HDR in particular, as this pathway must involve some form of interaction between the nascent sister chromatids, whether through the formation of a Holliday junction, replication fork regression or another, uncharacterized process. Close proximity between clustered lesions in the two strands may therefore pose a steric constraint on the availability of HDR. To address this question, we used dual-lesion TT-CPD and BP-G piggyBlock plasmids, that contain a lesion in each strand, 30 bp apart (Figure 1B). Two TT-CPD piggyBlock constructs of slightly different local sequence contexts were used, with the thought that this may affect DDT pathway choice, but as it turned out, it did not. In accordance with the single TT-CPD results, HDR signatures were virtually absent from the dual TT-CPD data (Figure 3C). Of 128 events in the CPD-CF data set, 126 bore accurate TLS signatures, while the remaining two were consistent with a HDR signature (Supplementary Table S3). Of the 112 events in the CPD-FC set, 111 gave the signature of accurate TLS, and only one was consistent with HDR (Supplementary Table S3). Overall, 237/240 events were accurate TLS, and only three may represent HDR. However, at this low incidence (~1%), it is impossible to distinguish between the HDR and the rare mutagenic TLS signatures. These data show that no inter-lesion proximity effect is at play in DDT across CPD, as the dual- and single-lesion constructs give the same results (Figure 3A and C).

When BP-G was assayed in the dual-lesion configuration, again HDR events were observed, but at a much lower rate than with the single BP-G piggyBlock plasmids: Now, only 27% (43/159) of the sequences contained the HDR signa-

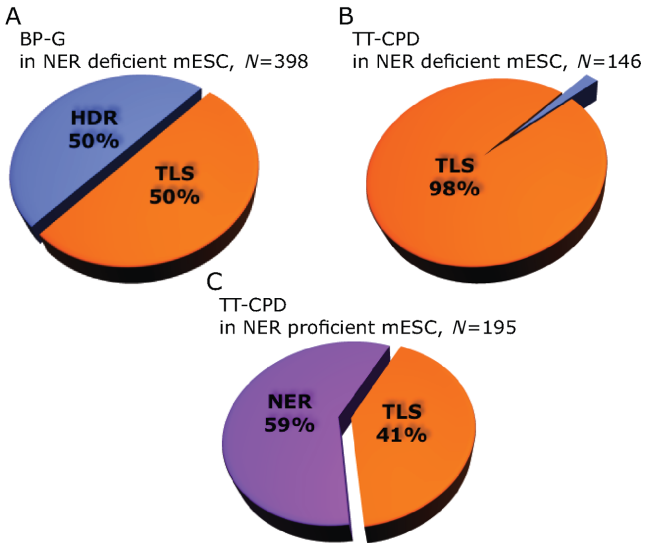

Figure 4. DNA sequence outcome of piggyBlock TLS/HDR and TLS/NER experiments in mES cells. (**A**) BP-G in NER-deficient mESC, $n = 398$. (**B**) TT-CPD in NER-deficient mESC, $n = 146$. (**C**) TT-CPD in NER-proficient mESC, $n = 195$.

ture (Figure 3D). This result follows the same trend we previously reported for human skin fibroblasts, using a ΦC31-based (10) chromosomal integration system (5). Of the 116 TLS signatures, 16 (14%) were mutagenic. Nearly all mutagenic signatures (14/16) were, again, insertions of A opposite BP-G (Supplementary Table S4).

DNA damage tolerance and repair in embryonic stem cells

To directly assay DDT in mES cells, we piggyBlock-transposed DNA segments containing dual BP-G or TT-CPD lesions into *Xpa* mES cell chromosomes. Expression of pluripotency markers was verified by quantitative reverse transcriptase-PCR (Supplementary Figure S1). HDR accounted for half the BP-G events in these cells (199/398), and TLS for the other half (Figure 4A and Supplementary Table S5). Mutation frequency of TLS across BP-G was 13% (25/199; Supplementary Table S5), similar to the 14% mutagenicity we observed in MEFs using the same piggyBlock plasmid. Of the 25 mutagenic signatures, 23 were A and the remaining two were G opposite BP-G insertions (Supplementary Table S5). Thus, we show that both DDT mechanisms are active in mES cells. Of note, no evidence was found for safeguards against point mutations beyond those available in differentiated cells. Of the 149 CPD sequences, three (2%) were consistent with HDR and the remaining 146 with accurate TLS (Figure 4B and Supplementary Table S6). These results are essentially identical to the ones obtained in MEFs.

To assess the contribution of DDT to genome maintenance, we assayed our piggyBlockTT-CPD dual-lesion plasmids in mES cells of the NER-proficient IB10 line, from which our *Xpa* cells were derived. Like HDR, NER uses the opposite strand sequence as template. Therefore, in cells in which this pathway is active, NER signatures are indistinguishable from those of HDR. In contrast to the *Xpa* data sets, this time we observed intact strand signatures in 59% of the events (47/78 in CPD-CF and 69/117 in CPD-FC, total 116/195; Figure 4C and Supplementary Table S7), most likely due to the activity of NER. Remarkably, a significant fraction—41% (79/195)—were TLS events, of which five (6%) were mutagenic (Supplementary Table S7).

DISCUSSION

The piggyBlock assay system for analyzing the fate of replication-blocking lesions in chromosomes provides higher chromosomal integration efficiency than our previously reported ΦC31-based system (5). It therefore requires smaller amounts of the limited-resource constructed lesion plasmid. piggyBlock also offers higher construction flexibility, as only a DNA fragment of choice, rather than the entire plasmid, is integrated into the genome. Unlike the chromosomal integration loci that we obtained using the ΦC31-baed system, a third of which mapped to the same locus (5), the piggyBlock loci that we identified in this work revealed no integration hotspot. This is consistent with previous large-scale analyses' results (8,9). Lesion strand identification in this system is achieved via a mismatched lesion position, which raises the question whether mismatch repair (MMR) is implicated in producing the bypass signature. MMR is known to follow behind replication and to possess nascent strand recognition ability (30). In our system, the lesion-mismatch is integrated into the chromosome *en bloc*, in a transposition event that is not coupled to replication, so neither strand is 'old' or 'new'. Indeed, our TT-CPD results, where each strand maintains its distinct original sequence, due to TLS that proceeds through the lesion strand, demonstrate that MMR does not act on our substrates. Combined with genome-editing techniques that became available in recent years (31), the piggyBlock system can be used to genetically dissect DDT mechanisms in NER-deficient cell cultures. We have previously shown that in xeroderma pigmentosum variant (*XPV*) cells, that lack a functional polη, an episomal TT-CPD in a structure that can undergo TLS but not HDR is bypassed by either polκ or polι, followed by polζ, in a mutagenic manner (32). The piggyBlock system can be used in cells that are deficient in both Xpa and polη, to examine whether this mutagenic TLS pathway or HDR is the main backup in the XPV syndrome.

It has been recognized for some time that DNA lesions have cognate TLS polymerases that specialize in their (relatively) efficient and accurate bypass. We have shown in previous work using an episomal plasmid system that is a substrate of TLS but not of HDR, that polη performs efficient and accurate bypass of TT-CPD, apparently unaided by any other TLS polymerase. Meanwhile, TLS across BP-G has a slower kinetics and requires the activity of two polymerases: polκ for insertion of a base opposite the damaged base and polζ for extension of the nascent strand from the lesion-mismatch base pair (23). Our present findings place these earlier ones in a broader context. They show that polη performs this task in all DDT events across CPD, while polκ and polζ only operate in a minority of DDT across BP-G events. Thus, the concept 'cognate polymerase' must be extended to 'cognate DDT pathway'. Accordingly, the question of how the right polymerase reaches the right lesion at

the right time must be modified, to encompass the possibility of none of the TLS polymerases, but rather the HDR mechanism reaching the lesion site when it is the best tool for the job. Nevertheless, when HDR-related strand negotiation is made difficult by the presence of another lesion nearby in the opposite strand, as in the case of lesion clusters, tolerance of BP-G can still be achieved. In those cases, the bulk of the work is performed by TLS.

The high abundance of DNA damage and its frequent encounters with replication forks suggest that DDT mechanisms must be constitutively available, to facilitate the timely progression of DNA replication and cell division. Indeed, both TLS and HDR were observed in our experiments without induction of the DNA damage response. Cells were not treated with DNA damaging agents, and transfection/transposition did not induce the DNA damage response, as indicated by the lack of Ser345 Chk1 phosphorylation (33) (Supplementary Figure S2). Thus, DDR is not essential for TLS and HDR operation, nor for orchestrating the different patterns of division of labor between the two observed for each of the lesions assayed. It would be interesting, however, to perform the piggyBlock assay in cells stressed by UV or another DDR-triggering agent, to investigate whether DDR alters the division of labor between the two pathways.

The use of piggyBlock to gauge TLS/HDR ratio requires that cells be NER-deficient, to prevent elimination of the lesions prior to replication. Yet the experiment performed in NER-proficient cells using piggyBlockTT-CPD revealed a significant fraction (41%) of events that proved to be TLS, despite the availability of the repair pathway. This can be attributed to TT-CPDs present in small numbers escaping detection by NER, by virtue of their structure that only mildly distorts the DNA double helix. Thus, TLS is not merely a last-resort mechanism for dealing with DNA damage that escaped removal by accurate repair systems, as it is portrayed at times. Rather, in the absence of DDR it is one of the mechanisms operating at the first line of defense against lesions blocking DNA replication.

Several studies have shown mES cells to be more sensitive to DNA damaging agents than differentiated cells (11). This is often explained as a means for eliminating cells with damaged or mutated genomes that can endanger entire cell lineages. Yet, the abundance of DNA damage and the frequent encounters of replication forks with DNA lesions make it unlikely that every such encounter leads to cell death, as this would severely undermine embryo development. Indeed, we show that under conditions of low DNA damage abundance, mES cells are proficient at DDT, carrying out both TLS and HDR. Remarkably, their behavior is largely similar to that of MEFs, bypassing TT-CPD by accurate TLS and BP-G by either HDR or TLS. Of note, although TLS across BP-G was mostly accurate, a significant fraction (13%), similar to that observed in MEFs, was mutagenic. Epidemiological studies in humans have shown that BP-G DNA lesions form in embryos as a result of maternal exposure during gestation (34,35). Our results suggest that point mutations resulting from TLS across these lesions may emerge at the earliest stages of embryo development, with the implication of genomic sequence alterations being propagated throughout the adult organism, as well as its own germ line.

The contrast between DDT mechanisms employed in the bypass of BP-G and CPD sheds new light on the unique status of polη among TLS polymerases. Deficiency in this polymerase in human patients results in the XPV syndrome, the only form of xeroderma pigmentosum that is not caused by disruption of the NER pathway. XPV is characterized by extreme UV sensitivity and increased skin cancer predisposition (36,37), that are recapitulated in polη knockout mice (38). No equivalent phenotype directly attributed to faulty DNA damage handling has been reported in relation to any other TLS polymerase. Polη performs TLS across CPD with efficiency and accuracy unparalleled by TLS of any other lesion tested and, unlike other TLS polymerases tested, does not require the aid of any other TLS polymerase for lesion bypass (23,25,39,40). Based on the findings presented in this work, we propose that it is the unavailability of an alternative DDT pathway for CPD that placed polη under the selective pressure that drove it to evolve to this degree of specialization. Meanwhile, the parallel operation of HDR relaxed selective pressure driving polymerase specialization in TLS across BP-G, and possibly other lesions, resulting in their lesser performance.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We thank Allan Bradley (Wellcome Trust Sanger Institute, UK) for the generous gift of piggyBac plasmids, Yael Fried from WIS Stem Cell Core Unit for technical help and Jacob Hanna (Weizmann Institute of Science, Rehovot) for his help with the mES cells.

FUNDING

Flight Attendant Medical Research Institute, Florida, USA; Leona M. and Harry B. Helmsley Charitable Trust, NY, USA [to Z.L.]; Israel Science Foundation [1136/08 and 684/12 to Z.L.]; U.S. National Institutes of Health/National Cancer Institute [CA099194 to N.G.]. Funding for open access charge: Flight Attendant Medical Research Institute, Florida, USA; Leona M. and Harry B. Helmsley Charitable Trust, NY, USA [to Z.L.]; Israel Science Foundation [1136/08 and 684/12 to Z.L.]; U.S. National Institutes of Health/National Cancer Institute [CA099194 to N.G.].
Conflict of interest statement. None declared.

REFERENCES

1. Friedberg,E.C. (2006) *DNA Repair And Mutagenesis*. Amer Society for Microbiology, Herndon, VA.
2. Friedberg,E.C. (2005) Suffering in silence: the tolerance of DNA damage. *Nat. Rev. Mol. Cell. Biol.*, **6**, 943–953.
3. Lehmann,A.R. and Fuchs,R.P. (2006) Gaps and forks in DNA replication: rediscovering old models. *DNA Repair*, **5**, 1495–1498.
4. Boiteux,S. and Jinks-Robertson,S. (2013) DNA repair mechanisms and the bypass of DNA damage in *Saccharomyces cerevisiae*. *Genetics*, **193**, 1025–1064.

5. Izhar,L., Ziv,O., Cohen,I.S., Geacintov,N.E. and Livneh,Z. (2013) Genomic assay reveals tolerance of DNA damage by both translesion DNA synthesis and homology-dependent repair in mammalian cells. *Proc. Natl. Acad. Sci.*, **110**, E1462–E1469.
6. Pagès,V., Mazón,G., Naiman,K., Philippin,G. and Fuchs,R.P. (2012) Monitoring bypass of single replication-blocking lesions by damage avoidance in the *Escherichia coli* chromosome. *Nucleic Acids Res.*, **40**, 9036–9043.
7. Cadiñanos,J. and Bradley,A. (2007) Generation of an inducible and optimized piggyBac transposon system. *Nucleic Acids Res.*, **35**, e87.
8. Akhtar,W., de Jong,J., Pindyurin,A.V., Pagie,L., Meuleman,W., de Ridder,J., Berns,A., Wessels,L.F.A., van Lohuizen,M. and van Steensel,B. (2013) Chromatin position effects assayed by thousands of reporters integrated in parallel. *Cell*, **154**, 914–927.
9. Liang,Q., Kong,J., Stalker,J. and Bradley,A. (2009) Chromosomal mobilization and reintegration of sleeping beauty and PiggyBac transposons. *Genesis*, **47**, 404–408.
10. Calos,M.P. (2006) The phiC31 integrase system for gene therapy. *Curr. Gene Ther.*, **6**, 633–645.
11. Liu,J.C., Lerou,P.H. and Lahav,G. (2014) Stem cells: balancing resistance and sensitivity to DNA damage. *Trends Cell Biol.*, **24**, 268–274.
12. Nouspikel,T. (2013) Genetic instability in human embryonic stem cells: prospects and caveats. *Future Oncol.*, **9**, 867–877.
13. Rocha,C.R.R., Lerner,L.K., Okamoto,O.K., Marchetto,M.C. and Menck,C.F.M. (2013) The role of DNA repair in the pluripotency and differentiation of human stem cells. *Mutat. Res./Rev. Mutat. Res.*, **752**, 25–35.
14. Tichy,E.D. and Stambrook,P.J. (2008) DNA repair in murine embryonic stem cells and differentiated cells. *Exp. Cell Res.*, **314**, 1929–1936.
15. Eppink,B., Tafel,A.A., Hanada,K., van Drunen,E., Hickson,I.D., Essers,J. and Kanaar,R. (2011) The response of mammalian cells to UV-light reveals Rad54-dependent and independent pathways of homologous recombination. *DNA Repair*, **10**, 1095–1105.
16. Yusa,K., Zhou,L., Li,M.A., Bradley,A. and Craig,N.L. (2011) A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 1531–1536.
17. Lis,J.T. and Schleif,R. (1975) Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Res.*, **2**, 383–389.
18. Kong,J., Zhu,F., Stalker,J. and Adams,D.J. (2008) iMapper: a web application for the automated analysis and mapping of insertional mutagenesis sequence data against Ensembl genomes. *Bioinformatics*, **24**, 2923–2925.
19. Flicek,P., Amode,M.R., Barrell,D., Beal,K., Billis,K., Brent,S., Carvalho-Silva,D., Clapham,P., Coates,G., Fitzgerald,S. *et al.* (2014) Ensembl 2014. *Nucleic Acids Res.*, **42**, D749–D755.
20. Ochman,H., Gerber,A.S. and Hartl,D.L. (1988) Genetic applications of an inverse polymerase chain reaction. *Genetics*, **120**, 621–623.
21. de Vries,A., van Oostrom,C.T., Hofhuis,F.M., Dortant,P.M., Berg,R.J., de Gruijl,F.R., Wester,P.W., van Kreijl,C.F., Capel,P.J., van Steeg,H. *et al.* (1995) Increased susceptibility to ultraviolet-B and carcinogens of mice lacking the DNA excision repair gene XPA. *Nature*, **377**, 169–173.
22. Shachar,S., Ziv,O., Avkin,S., Adar,S., Wittschieben,J., Reißner,T., Chaney,S., Friedberg,E.C., Wang,Z., Carell,T. *et al.* (2009) Two-polymerase mechanisms dictate error-free and error-prone translesion DNA synthesis in mammals. *EMBO J.*, **28**, 383–393.
23. Livneh,Z., Ziv,O. and Shachar,S. (2010) Multiple two-polymerase mechanisms in mammalian translesion DNA synthesis. *Cell Cycle*, **9**, 729–735.
24. Park,H., Zhang,K., Ren,Y., Nadji,S., Sinha,N., Taylor,J.-S. and Kang,C. (2002) Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc. Natl. Acad. Sci.*, **99**, 15965–15970.
25. Masutani,C., Kusumoto,R., Yamada,A., Dohmae,N., Yokoi,M., Yuasa,M., Araki,M., Iwai,S., Takio,K. and Hanaoka,F. (1999) The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase [eta]. *Nature*, **399**, 700–704.
26. Johnson,R.E., Prakash,S. and Prakash,L. (1999) Efficient bypass of a thymine–thymine dimer by yeast DNA polymerase, Polη. *Science*, **283**, 1001–1004.
27. Avkin,S., Goldsmith,M., Velasco-Miguel,S., Geacintov,N., Friedberg,E.C. and Livneh,Z. (2004) Quantitative analysis of translesion DNA synthesis across a benzo[a]pyrene-guanine adduct in mammalian cells: the role of DNA polymerase κ. *J. Biol. Chem.*, **279**, 53298–53305.
28. Carr,I.M., Robinson,J.I., Dimitriou,R., Markham,A.F., Morgan,A.W. and Bonthron,D.T. (2009) Inferring relative proportions of DNA variants from sequencing electropherograms. *Bioinformatics*, **25**, 3244–3250.
29. Skosareva,L.V., Lebedeva,N.A., Rechkunova,N.I., Kolbanovskiy,A., Geacintov,N.E. and Lavrik,O.I. (2012) Human DNA polymerase λ catalyzes lesion bypass across benzo[a]pyrene-derived DNA adduct during base excision repair. *DNA Repair*, **11**, 367–373.
30. Iyer,R.R., Pluciennik,A., Burdett,V. and Modrich,P.L. (2005) DNA mismatch repair: functions and mechanisms. *Chem. Rev.*, **106**, 302–323.
31. Mali,P., Esvelt,K.M. and Church,G.M. (2013) Cas9 as a versatile tool for engineering biology. *Nat. Methods*, **10**, 957–963.
32. Ziv,O., Geacintov,N., Nakajima,S., Yasui,A. and Livneh,Z. (2009) DNA polymerase ζ cooperates with polymerases κ and ι in translesion DNA synthesis across pyrimidine photodimers in cells from XPV patients. *Proc. Natl. Acad. Sci.*, **106**, 11552–11557.
33. Walker,M., Black,E.J., Oehler,V., Gillespie,D.A. and Scott,M.T. (2009) Chk1 C-terminal regulatory phosphorylation mediates checkpoint activation by de-repression of Chk1 catalytic activity. *Oncogene*, **28**, 2314–2323.
34. Jedrychowski,W.A., Perera,F.P., Tang,D., Rauh,V., Majewska,R., Mroz,E., Flak,E., Stigter,L., Spengler,J., Camann,D. *et al.* (2013) The relationship between prenatal exposure to airborne polycyclic aromatic hydrocarbons (PAHs) and PAH-DNA adducts in cord blood. *J. Expos. Sci. Environ. Epidemiol.*, **23**, 371–377.
35. Perera,F.P., Tang,D., Tu,Y.H., Cruz,L.A., Borjas,M., Bernert,T. and Whyatt,R.M. (2004) Biomarkers in maternal and newborn blood indicate heightened fetal susceptibility to procarcinogenic DNA damage. *Environ. Health Perspect.*, **112**, 1133–1136.
36. Maher,V.M., Ouellette,L.M., Curren,R.D. and McCormick,J.J. (1976) Frequency of ultraviolet light-induced mutations is higher in xeroderma pigmentosum variant cells than in normal human cells. *Nature*, **261**, 593–595.
37. Lehmann,A., McGibbon,D. and Stefanini,M. (2011) Xeroderma pigmentosum. *Orphanet J. Rare Dis.*, **6**, 70.
38. Lin,Q., Clark,A.B., McCulloch,S.D., Yuan,T., Bronson,R.T., Kunkel,T.A. and Kucherlapati,R. (2006) Increased susceptibility to UV-induced skin carcinogenesis in polymerase η-deficient mice. *Cancer Res.*, **66**, 87–94.
39. Johnson,R.E., Washington,M.T., Prakash,S. and Prakash,L. (2000) Fidelity of human DNA polymerase η. *J. Biol. Chem.*, **275**, 7447–7450.
40. Hendel,A., Ziv,O., Gueranger,Q., Geacintov,N. and Livneh,Z. (2008) Reduced efficiency and increased mutagenicity of translesion DNA synthesis across a TT cyclobutane pyrimidine dimer, but not a TT 6–4 photoproduct, in human cells lacking DNA polymerase eta. *DNA Repair (Amst.)*, **7**, 1636–1646.