Supplementary Information

Genome-wide mapping of oxidative DNA damage via engineering of 8-oxoguanine DNA glycosylase

Yuxin Fang[1], Peng Zou[1,2,3,*]

[1] College of Chemistry and Molecular Engineering, Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University, Beijing, 100871, China

[2] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, 100871, China

[3] PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, 100871, China

[*] Correspondence: zoupeng@pku.edu.cn

Table of contents

Materials and Methods
Cell culture
Mouse embryonic fibroblast (MEF) cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM, Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco). Cells were cultured at 37℃ under 5% CO_2. For genomic DNA extraction, cells were harvested at 80% confluence.

Preparation of model DNA sequences
Fluorescently labeled OG-DNA was generated by annealing of a 5'-6-FAM-modified 30 nt oligonucleotide containing an OG modification (Sangon) and its complementary strand. The 35 bp OG-DNA for ITC experiments was generated by annealing of a 35 nt oligonucleotide containing one OG modification (Sangon) and its complementary strand. Model sequences used in enrichment procedure were generated by PCR amplification with an OG-modified forward primer (Sangon) and its corresponding reverse primer. For the preparation of AP site sequence, a 35 bp double-strand DNA containing a U:C mispair was ligated with two 250 bp DNA duplex by T4 ligase (NEB, M0202). The ~500 bp ligation product was purified by TIANgel Midi Purification Kit (TIANGEN, DP209). Uracil excision was completed by incubation with UDG (NEB, M0280) at 37℃ for 30 min to generate the AP site. Model DNA sequences were stored at -20℃. Detailed sequences are listed in **Table S2**.

Recombinant expression and purification of DNA glycosylases
Glycosylase genes were inserted to vector plasmid pET21a, followed with a His_6 tag at the C terminal. For purification of Fpg, T7 tag on the vector plasmid was removed to avoid interference with glycosylase activity. *E. coli* strain BL21-DE3 cells were transformed with the plasmids and grown on LB agar with 70 μg/mL ampicillin at 37℃ overnight. Individual colony was proliferated in 5 mL LB medium with 70 μg/mL ampicillin for 8 h, then 1:100 diluted to 500 mL LB medium and grown at 37℃ to an OD_{600} of 0.5-1.0. Protein expression was induced with 1 mM IPTG (isopropyl β-d-1-thiogalactopyranoside, INALCO, 1758-1400) at 20℃ for 12-14 hours. Cells were harvested by centrifugation and then lysed at 4℃ by ultra-sonication in 20 mL binding buffer (50 mM Tris-HCl, pH7.5, 300 mM NaCl). The suspension was centrifuged at 20,000 rpm for 30 min and the clear supernatant was incubated with 2 mL Ni-NTA agarose slurry (Qiagen, 30210) at 4℃ for 30 min. Cell lysate with unbound proteins flowed out under gravity. The resin was then washed with 20 mL wash buffer (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 50 mM imidazole). Bound proteins were eluted with 10 mL elute buffer (200 mM imidazole in 50 mM Tris-HCl, pH 7.5, 300 mM NaCl). Purity of target proteins was assessed by SDS-PAGE (**Figure S14**). The fractions with eluted glycosylases were dialyzed against 1×PBS (pH 7.4, Solarbio, P1010) with 1 mM DTT (dithiothreitol, INALCO, 1758-9030) and concentrated using Amicon Ultra-4 filter with Ultracel-10 membrane (Millipore). Protein concentration was determined by absorption at 280nm for general uses or by Pierce BCA Protein Assay Kit (Thermo, 23227) for ITC experiments. Aliquots of purified proteins were stored at -80℃.

Gel-shift assay
For borohydride trapping assays, 5 μM DNA was incubated with 10 μM DNA glycosylases in 25 mM Tris-HCl pH 7.6, 2 mM EDTA, 50 mM $NaBH_3CN$ at 37℃ for 20 min. The reaction was quenched by 2×formamide-SDS loading buffer (40% SDS, 0.02% bromophenol blue, 90% formamide, 100 mM Tris-HCl pH 6.8, 2 mM DTT) and heated at 90℃ for 10 min. Samples were separated by SDS-PAGE (12% polyacrylamide) at 220 V for 30min and imaged by Typhoon FLA 9500.

For electrophoretic mobility shift assay (EMSA), 5 μM DNA was incubated with 10 μM DNA glycosylases in 25 mM Tris-HCl pH 7.6, 2 mM EDTA at 37℃ for one hour. After incubation, a final concentration of 5% glycerol and 0.02% bromophenol blue was added. Samples were separated by native PAGE (12% polyacrylamide, running in 0.5×TBE) at 80 V for 80 min and imaged by Typhoon FLA 9500.

Isothermal titration calorimetry (ITC)

Calorimetric binding experiments were conducted on MicroCal iTC$_{200}$ system (Malven). The reaction buffer used in ITC experiments was comprised of 10 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM DTT. Both the protein and DNA were diluted with reaction buffer to ensure a uniform buffer environment during titration. 8 μM OG-DNA was titrated with 80 μM ^{K249Q}hOGG1 in 2 μL increments (20 injections in total) at 150 s intervals. Titration experiment was conducted at 25℃. A control experiment where proteins were titrated into blank solution without DNA was subtracted from protein-DNA binding isotherm. The quantity of heat absorbed upon binding was measured by integrating the area of each endotherm. Results were analyzed by MicroCal AddOn for Origin 7.0. The association constant (K_a) for the binding reaction was obtained by fitting of the data to a single-site binding model. Data points with significant deviations were excluded from fitting analysis.

Enrichment procedure of model DNA sequence

1.5 ng of each model sequence was spiked into 50 μg salmon sperm DNA (Invitrogen, 15632-011). DNA mixture was incubated with 4.5 μM glycosylase in 20 μL system containing 10 mM Tris-HCl pH 7.5, 150 mM NaCl and 1 mM DTT. Reaction system was incubated at 37℃ for 40 min under the condition with or without 50 mM sodium cyanoborohydride. The excessive salmon sperm DNA was added to simulate non-OG-containing DNA from genomic samples, and also to improve enrichment fold by blocking the non-specific binding of glycosylases to control DNA (**Figure S7**). Before enrichment, 5% of the reaction system was set aside as input. Complex of glycosylase and DNA was enriched by MagneHis Ni-Particles (Promega, V8560). 10 μL beads were wash twice with binding buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl) and mixed with each 20 μL reaction system. Then the mixture was vortexed at 1,400 rpm on the shaker at room temperature for 30 min, before being placed on a magnetic rack. After removing the supernatant, beads were washed four times with 500 μL binding buffer. Complex of glycosylase and OG-DNA was eluted by 500 mM imidazole in binding buffer. After DNA purification, Power Up SYBR Green Master Mix (Applied Biosystems, A25777) was used for qPCR analysis, which was conducted on Step One Plus Real-Time PCR system (Applied Biosystems). The enrichment fold was calculated by the delta-delta Ct method, in which the difference of Ct values (ΔCt) between pull-down sample and input sample was calculated first, then the difference of ΔCt (ΔΔCt) between modified DNA and unmodified DNA was calculated and transformed into folds of enrichment. Primers used in qPCR analysis are listed in **Table S3.**

Enrichment procedure of genomic DNA

Cells were harvested and genomic DNA was extracted with TIANamp Genomic DNA Kit (TIANGEN, DP304). 0.1 mM desferal (TargetMol, T1637) was added at each step to prevent further DNA oxidation during genomic DNA extraction. Genomic DNA was digested to 100-1000 bp with NEBNext dsDNA Fragmentase (NEB, M0348). DNA fragments were purified with 2×CMPure beads (CWbiotech, CW2508). Sonication was avoided during the fragmentation process as it has been shown to introduce

artificial DNA damages[1]. 2 pg model sequences were spiked into each 10 μg genomic DNA fragments. Half of the DNA mixture was incubated with Fpg (NEB, M0240) at 37℃ for one hour for damage excision. DNA was purified with 2×CMPure beads. 100 ng DNA was separated from each sample as input. For each enrichment, 4 μg genomic DNA was incubated with 4.5 μM ^{K249Q}hOGG1 in 100 μL system with 1×NEBuffer1(NEB, B7001). To improve the enrichment fold, 20 μg empty vector plasmid pcDNA3.1 was added as blocking DNA in each enrichment system. The mixture was incubated at 37℃ for 40 min and enriched as procedures described above. After elution, DNA was purified with 2× CMPure beads.

NGS library construction was performed with VAHTS Universal DNA Library Prep Kit for Illumina (Vazyme, NP604) following the manufacturer's guidelines. Size selection was conducted after PCR amplification by using 0.55× and 0.2× VAHTS DNA clean beads (Vazyme, N411) to obtain 300-500 bp fragments. Size distribution of libraries was verified with capillary electrophoresis. Sequencing was performed by Illumina HiSeq X Ten platform.

NGS Data analysis

All sequencing data and analysis results have been deposited into the Gene Expression Omnibus (GEO), with the accession number GSE136862. Raw reads were first performed with adaptor and quality trimming using cutadapt[2]. Reads with quality values below 25 and reads shorter than 45 nt after trimming were excluded. The processed reads were mapped to mouse reference genome (mm9) using bowtie2[3]. A bowtie2 index was built from combination of mm9 reference genome, spike-in DNA sequences and blocking DNA sequences. After alignment, peaks were called by model-based analysis of ChIP-Seq (MACS2[4]) using non-redundant reads. For peaks called by enrich/input, p<0.001 was used as cutoff. While for peaks called by enrich/control, a cutoff p<0.05, was used. Only those peaks called by two methods simultaneously were retained. Overlap of two biological replicates was defined as high-confidence peaks and was applied to further analysis. Reads visualization was performed with Integrative Genomics Viewer (IGV[5]). Visualization of peaks over chromosomes was performed with ChIPseeker[6].

Genomic features were extracted from TxDb.Mmusculus.UCSC.mm9.knownGene (R package version 3.2.2.) using GenomicRanges[7]. Regions between 500 bp upstream and 100 bp downstream the transcription start site (TSS) were extracted as promoters. Annotation of CpG islands and DNase I highly sensitive regions (DNaseI_HS) were downloaded from AnnotationHub (AH6117; AH6129). ChIP-Seq data of RNA polII (encodeproject, ENCFF291FCH), H3K27ac (encodeproject, ENCFF467TWY) and H3K9me3 (NCBI GEO, GSM1621023) were also downloaded. Intersection between peaks and genomic features was analyzed using GenomicRanges. Enrichment in genomic elements was calculated by dividing fraction of peaks by fraction of genome. Detailed results are listed in **Table S1.**

For potential G-quadruplex sequence (PQS) analysis, the DNA sequences of peaks were extracted by BEDTools[8]. PQSs were predicted with pqsfinder[9] with requirements of loops between 1-12 nt and a minimal score of 20.

Pearson's correlation test was performed with deepTools[10]. For calculations based on raw reads, reads mapped to the genome were counted by bins of 10 kb in length. For calculations based on peak regions, FPKM (fragments per kilobase per million mapped fragments) of peak regions were used.

To identify mutations within the peak regions, we performed SNP calling on the alignment results of the input and the enriched samples. SNPs were called by bcftools over the mm9 genome and filtered

by the following criteria: 1) QUAL values greater than 20; 2) raw read depth larger than 30; 3) occurrence in both replicates. SNPs from the enriched sample within the peak regions were extracted, and those that overlap with SNPs in the input sample were excluded from further analysis. A total of 20 mutations were observed among 1,470 OG peaks (Table S4).

Supplementary Tables
Table S1. Distribution of peaks in genomic elements

Features	Fraction of genome[*]	Fraction of peaks#1[**]	Fraction of peaks#2[**]	Enrichment #1[***]	Enrichment #2[***]	Average	SD
Intergenic	0.6103	0.5788	0.5785	0.9484	0.9480	0.9482	0.0003
Promoter	0.0078	0.0225	0.0220	2.8963	2.8281	2.8622	0.0482
Intron	0.3755	0.3672	0.3691	0.9779	0.9829	0.9804	0.0036
5'UTR	0.0025	0.0071	0.0084	2.8255	3.3565	3.0910	0.3755
3'UTR	0.0107	0.0084	0.0084	0.7840	0.7881	0.7860	0.0029
CDS	0.0129	0.0161	0.0136	1.2503	1.0557	1.1530	0.1376
CpG islands	0.0039	0.0199	0.0194	5.1771	5.0360	5.1065	0.0998
DNase I_HS	0.0232	0.0489	0.0524	2.1104	2.2609	2.1857	0.1064
PolII	0.0094	0.0264	0.0246	2.7985	2.6072	2.7028	0.1353
H3K27ac	0.0189	0.0264	0.0252	1.3953	1.3341	1.3647	0.0433
H3K9me3	0.3515	0.2990	0.2961	0.8509	0.8424	0.8466	0.0060

[*]Fraction of genome = length of element (bp) / length of whole genome (bp)

[**]Fraction of peaks = number of peaks in element / number of peaks in whole genome

[***]Enrichment = fraction of peaks / fraction of genome

Table S2. Model DNA sequences in this study

5'-6-FAM-OG	TGTTCATCATGGGTC/**OG**/TCGGTATATCAGCT
5'-6-FAM-control	TGTTCATCATGGGTCGTCGGTATATCAGCT
ITC-OG	AGCTTGTTCATCATGGGTC/**OG**/TCGGTATATCAGCTG
ITC-control	AGCTTGTTCATCATGGGTCGTCGGTATATCAGCTG
OG sequence 1	TGTTCATCATGGGTC/**OG**/TCGGTATATCAGCTGGATCGAATTCCCTACTTCATACATTTTCAATTAAGATGCCACCAAAAAAAAAAAAGAAAAGTTAAGGACAATAGCTCGACGATTGAAGGTAGATACCCATACGACGTTCCAGACTACGCTCTGCAGCAAGGTCTGCAGGCTAGTGGTGGAGGAGGCTCTGGTGCTAGCGGAAAGTCTTACCCAACTGTGAGTGCTGATTACCAGGACGCCGTTGAGAAGG
OG sequence 2	TGTTCATCATGGGTC/**OG**/TCGGTATATCAGCTATGCCTGAATTACCCGAAGTTGAAACCAGCCGCCGCGGCATAGAACCGCATCTCGTTGGTGCAACCATTCTTCATGCAGTGGTGCGCAACGGACGCTTGCGCTGGCCGGTTTCAGAAGAGATCTACCGTTTAAGCGACCAACCAGTGCTTAGCGTGCAGCGGCGGGCTAAATATCTGCTGCTGGAGCTGCCTGAGGGCTGGATTATCATTCATTTAGGGA
Control sequence 1	GGTGAGCCGTGTCGGGTGTGCGGTACGCCGATTGTGGCGACTAAACATGCGCAGCGGGCAACGTTTTATTGTCGGCAGTGCCAGAAGAAGCTTGACTACAAAGACCACGACGGTGACTACAAAGACCACGACATCGACTACAAAGACGACGACAAACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAG

Control sequence 2	TGGAACCTTTGACAAGGGCACGAAGACCGGTGGACCCTTCGGAACC ATCAAGCACCCTGCCGAACTGGCTCACAGCGCTAACAACGGTCTTGA CATCGCTGTTAGGCTTTTGGAGCCACTCAAGGCGGAGTTCCCTATTTT GAGCTACGCCGATTTCTACCAGTTGGCTGGCGTTGTTGCCGTTGAGG TCACGGGTGGACCTAAGGTTCCATTCCACCCTGGAAGAGAGGACAA GCCTGAGCCACCACCA
35bp U:C mispair forward	5'P-AGCTTGTTCATCATGGGTC**U**TCGGTATATCAGCTG
35bp U:C mispair reverse	5'P-AATTCAGCTGATATACCGA**C**GACCCATGATGAACA
Full-length AP site	CATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTG AAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGC GGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCG CTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTT CCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGG GGCTCCCTTTAGGGAAGCTTGTTCATCATGGGTC/**AP**/TCGGTATATCA GCTGAATTCTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAA ACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGAC GGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACT CTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTT TGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGA GCTGATTTAACAAAAATTTA
500bp control sequence	TGTTCATCATGGGTCGTCGGTATATCAGCTATGCCTGAATTACCCGAA GTTGAAACCAGCCGCCGCGGCATAGAACCGCATCTCGTTGGTGCAAC CATTCTTCATGCAGTGGTGCGCAACGGACGCTTGCGCTGGCCGGTTT CAGAAGAGATCTACCGTTTAAGCGACCAACCAGTGCTTAGCGTGCAG CGGCGGGCTAAATATCTGCTGCTGGAGCTGCCTGAGGGCTGGATTATC ATTCATTTAGGGATGTCTGGCAGCCTGCGCATCCTTCCAGAAGAACTT CCCCCTGAAAAGCATGACCATGTGGATTTGGTGATGAGCAACGGCAA AGTGCTGCGCTACACCGATCCGCGCCGCTTTGGTGCCTGGCTGTGGA CCAAAGAGCTGGAAGGGCATAATGTGCTGACCCATCTTGGACCGGAG CCGCTTAGCGACGATTTCAATGGTGAGTATCTGCATCAGAAGTGCGC GAAGAAAAAAACGGCGATTAAACCGTG

Table S3. qPCR primers in this study

OG sequence 1 forward	ACGCTCTGCAGCAAGGTC
OG sequence 1 reverse	CAACGGCGTCCTGGTAATCAG
OG sequence 2 forward	GTTGGTGCAACCATTCTTCATGCA
OG sequence 2 reverse	CGCTAAGCACTGGTTGGTCG
Control sequence 1 forward	GTCGGCAGTGCCAGAAGAAG
Control sequence 1 reverse	TGGTGGTGCTCGAGTTTGTCG
Control sequence 2 forward	GCCGAACTGGCTCACAGC
Control sequence 2 reverse	AGAAATCGGCGTAGCTCAAAATAGG
AP site forward	GATTAGGGTGATGGTTCACGTAGTGG
AP site reverse	GGAACAAGAGTCCACTATTAAAGAACGTGG

Table S4. Occurrence of mutations within OG peaks

Mutation	Occurrence
A-T	3
A-G	1
A-C	1
T-G	1
T-C	5
G-A	2
G-C	5
C-T	2

Supplementary Figures
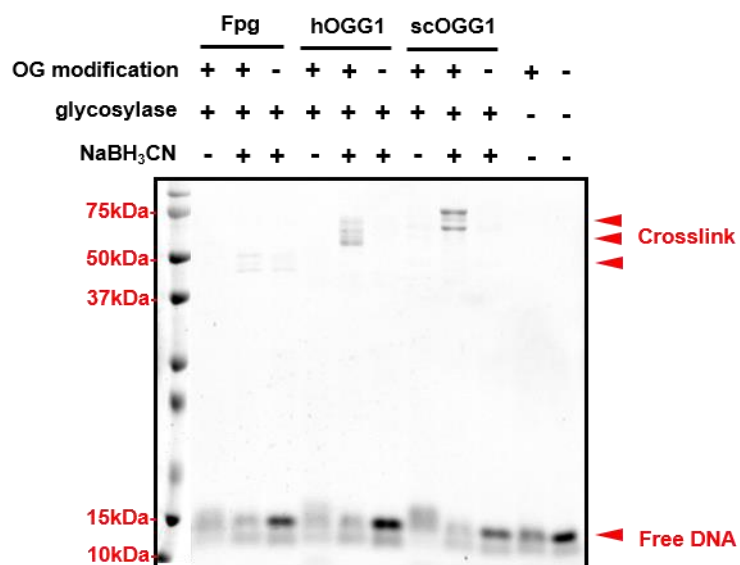
Figure S1. Borohydride trapping assay of DNA glycosylases. Reactions were analyzed by SDS-PAGE and in-gel fluorescence. The molecular weight of pre-stained protein marker (1st lane) was indicated on the left. Upper bands with large molecular weights indicated the formation of glycosylase-DNA crosslink.
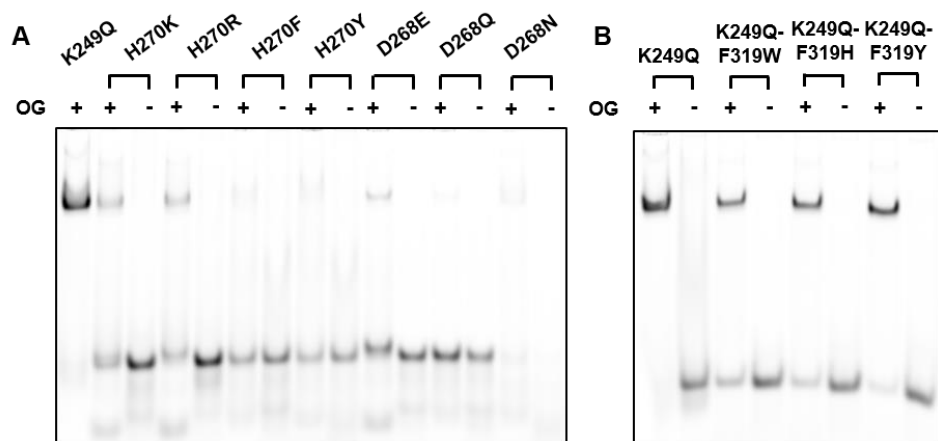
Figure S2. Binding of DNA glycosylases to OG-DNA detected by EMSA. Reactions were analyzed by native PAGE and in-gel fluorescence. Upper bands with larger molecular weights indicated the formation of glycosylase-DNA complex. (**A**) Performance of hOGG1 variants. (**B**) Performance of hOGG1 variants containing double mutations of Lys249 and Phe319.
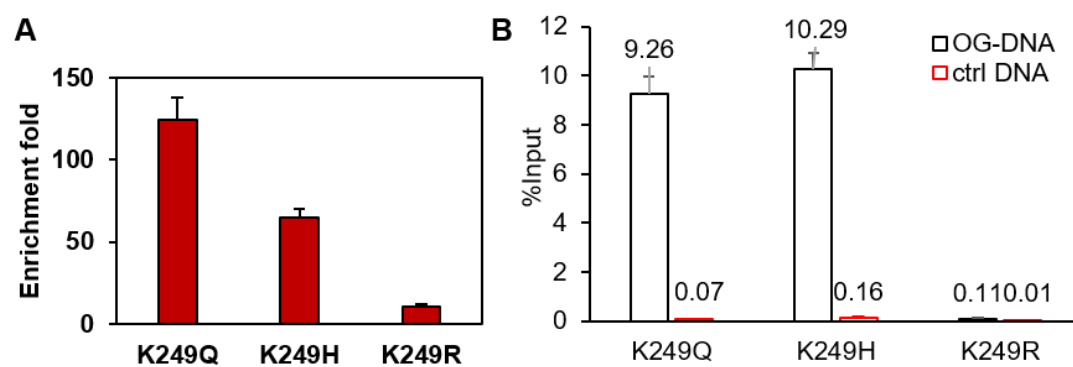
Figure S3. Comparison of Lys249 mutants in OG-DNA enrichment. (**A**) Relative enrichment fold when using Lys249 mutants. Values represent fold of enrichment over the input, normalized to a control sequence consisting of canonical bases. (**B**) Recovery rate of model DNA sequences after enrichment.
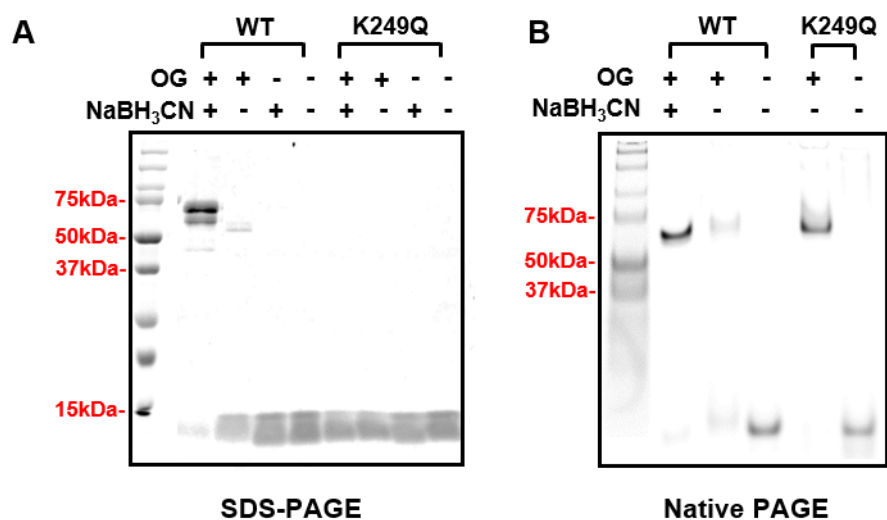
Figure S4. Non-covalent interaction between K249Q and OG-DNA. Reactions were separated by SDS-PAGE (**A**) or native PAGE (**B**), followed by in-gel fluorescence analysis. The molecular weight of pre-stained protein marker (1st lane) was indicated on the left. Upper bands with larger molecular weights indicated the formation of glycosylase-DNA complex. WT refers to wild type hOGG1.
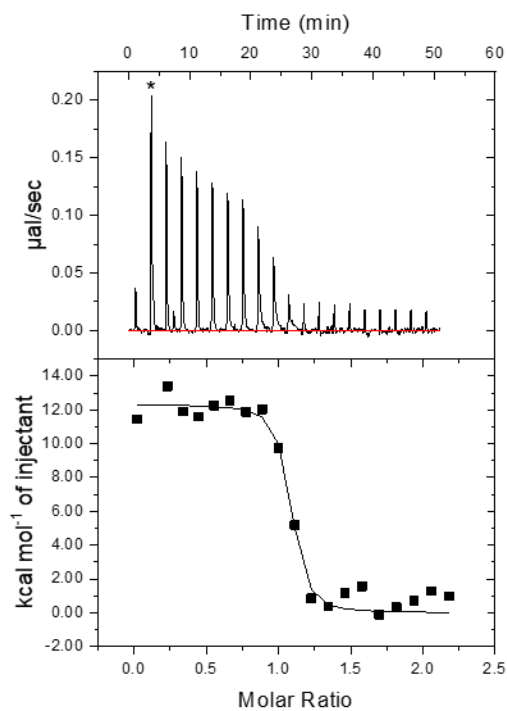
Figure S5. ITC binding isotherm for the association of OG-DNA and [K249Q]**hOGG1.** Upper, total heat absorbed per injection. Lower, resultant-binding isotherm (dots) obtained by integrating the peak areas of each injection. The solid line represents the fitting to a single-site-binding model. The second injection (denoted by asterisk) was marked as an outlier and excluded from fitting.
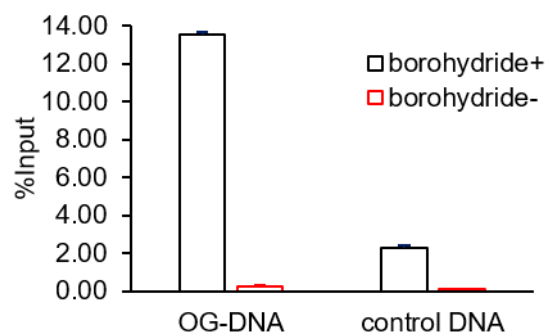
Figure S6. Borohydride-induced crosslink between hOGG1 and control DNA. Recovery of control DNA in the reaction omitting borohydride indicated the basal level of background. Addition of borohydride increased the recovery of control DNA, which led to poor enrichment specificity.
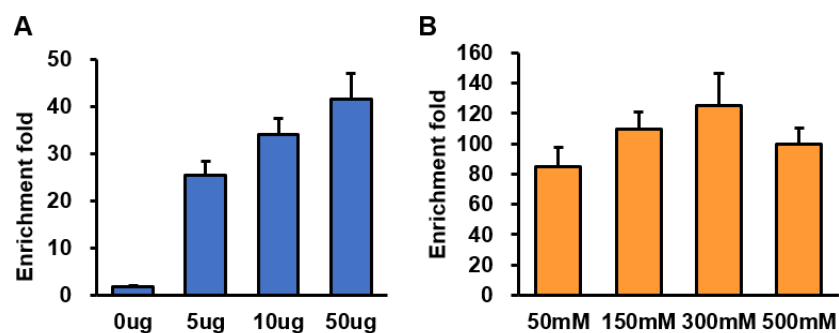
Figure S7. Optimization of enrichment procedures. (**A**) Enrichment of OG-DNA with different amounts of blocking DNA. (**B**) Enrichment of OG-DNA in reaction buffers with different NaCl concentrations. Results were analyzed by qPCR using the ΔΔCt method for calculation. Values represent fold enrichment relative to the input, normalized to a control sequence consisting of canonical bases (i.e. without OG).
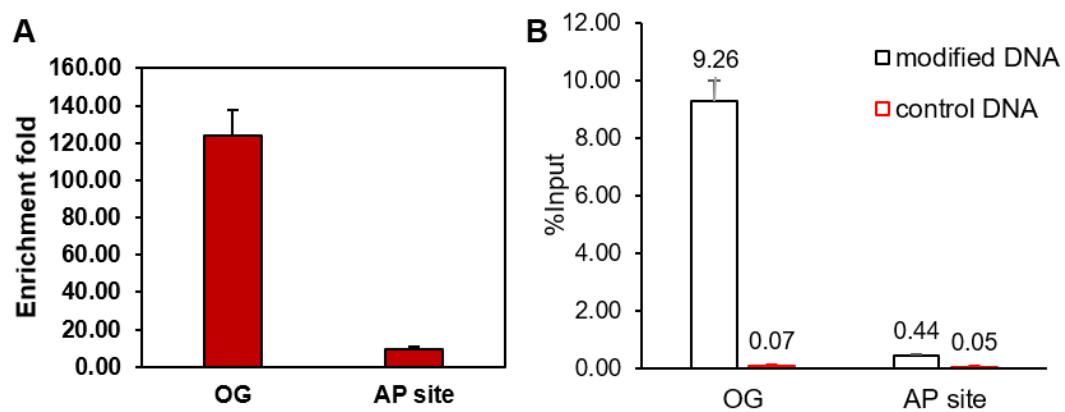
Figure S8. Enrichment of AP site using ^{K249Q}**hOGG1.** (**A**) Enrichment fold of OG-DNA and AP site sequence. Values represent fold of enrichment over the input, normalized to a control sequence consisting of canonical bases. (**B**) Recovery rate of model DNA sequences after enrichment.
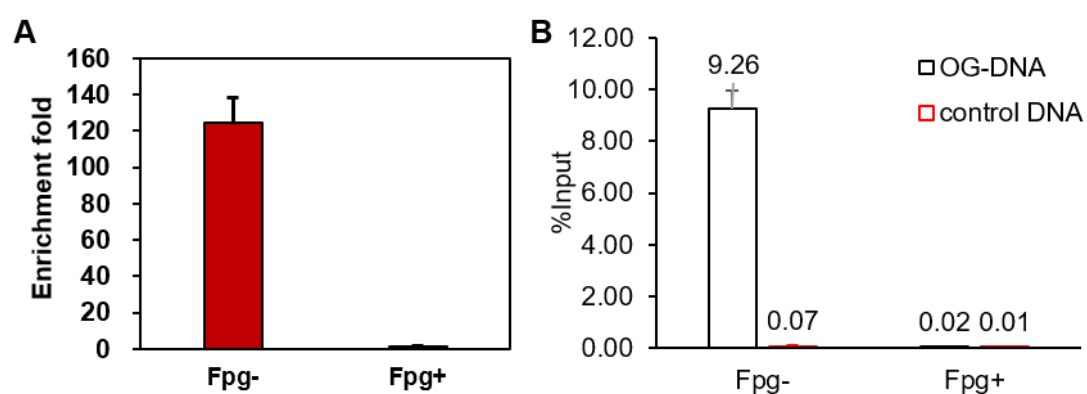
Figure S9. Performance of damage excision on model DNA sequences. (**A**) Enrichment fold of OG-DNA with and without damage excision by Fpg. Values represent fold of enrichment over the input, normalized to a control sequence consisting of canonical bases. (**B**) Recovery rate of model DNA sequences after enrichment.
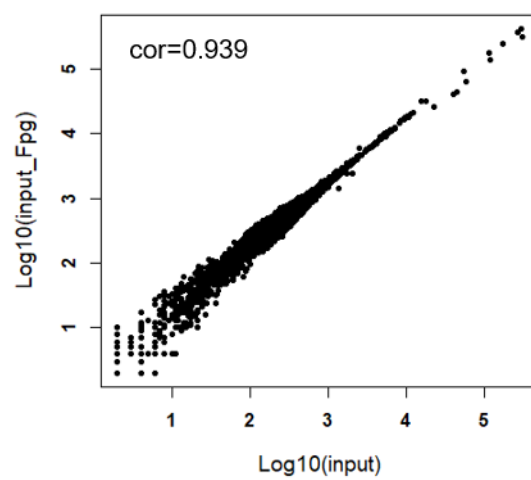
Figure S10. Evaluation of interference in sequencing caused by Fpg treatment. Pearson's correlation coefficient (cor=0.939) between input samples with and without Fpg treatment. The coefficient was calculated from raw reads.
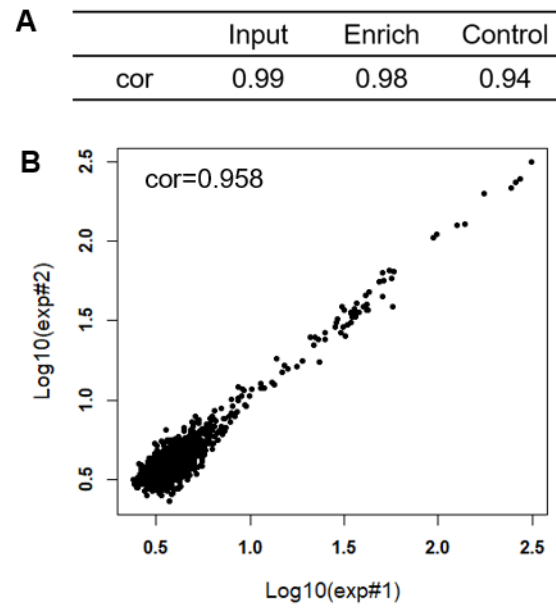
Figure S11. Correlation coefficient between two biological replicates. Pearson's correlation coefficient (cor=0.958) between two biological replicates. The coefficients were calculated from raw reads (**A**) or peak regions (**B**).
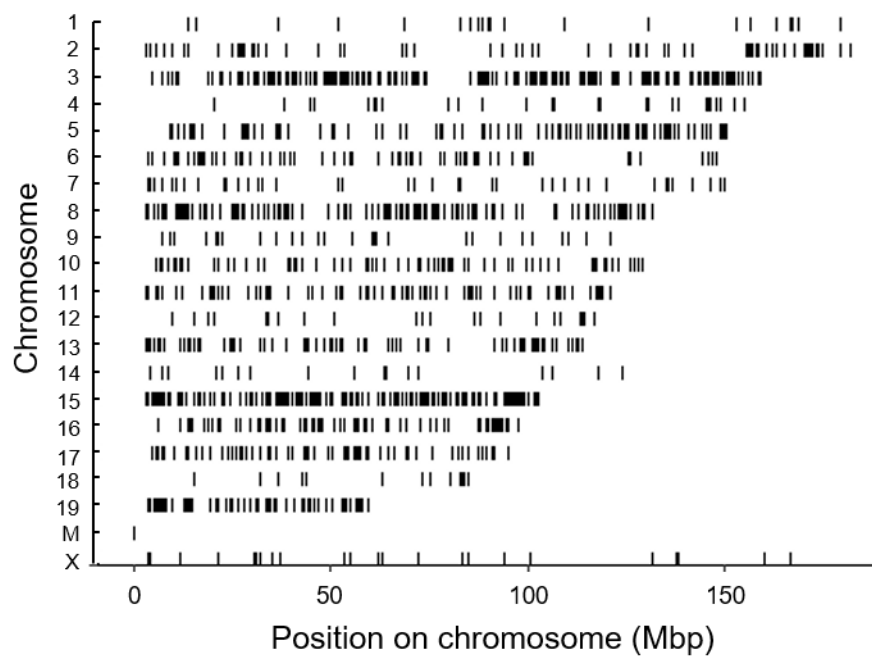
Figure S12. Distribution of peaks over chromosomes. The horizontal axis indicates the position along the chromosomes (bp) and the vertical axis indicates chromosome number. No peak was identified in the Y chromosome.
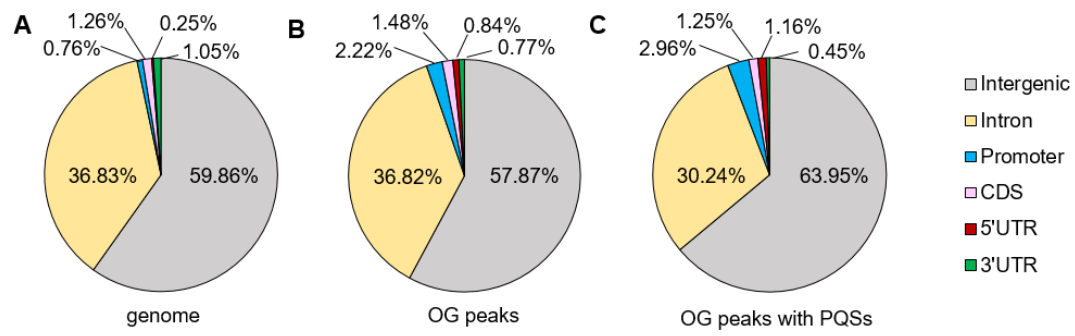
Figure S13. Distribution of OG peaks in genomic elements. (**A**) Content of each element in mouse genome. (**B**) Distribution of total peaks. (**C**) Distribution of peaks possessing PQSs
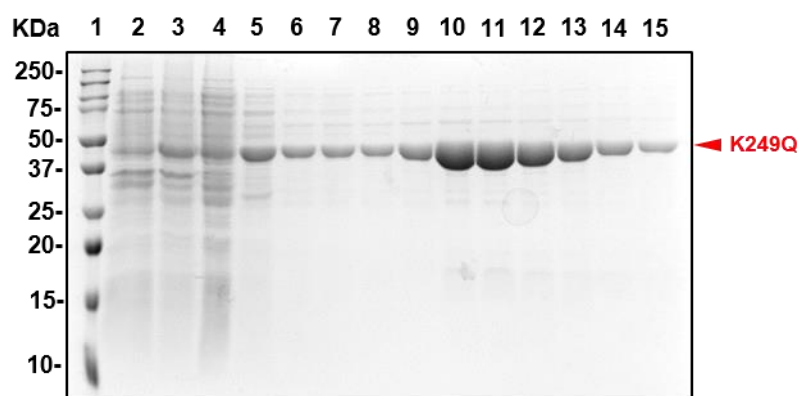
Figure S14. Purity of recombinant ^{K249Q}hOGG1. Samples were separated by SDS-PAGE (12% polyacrylamide) and stained with Commassie brilliant blue. The molecular weight of pre-stained protein marker (1st lane) was indicated on the left. Lane 1: pre-stained protein marker; lane 2: whole cell lysate from cells without IPTG induction; lane 3: whole cell lysate from cells induced by IPTG for 12 h; lane 4: flow through of cell lysate with unbound proteins; lane 5: flowing out of wash buffer; lane 6-15: fractions with the eluted proteins. Fractions indicated by lane 9-14 were collected for dialysis and concentration.

Reference

(1) Ali, M. H. M., Al-Saad, K. A., and Ali, C. M. (2014) Biophysical studies of the effect of high power ultrasound on the DNA solution. *Phys. Medica 30*, 221–227.

(2) Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal 17*, 10-12.

(3) Langmead, B., and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods 9*, 357–359.

(4) Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nussbaum, C., Myers, R. M., Brown, M., Li, W., and Liu, X. S. (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biol. 9*, R137.1-R137.9.

(5) Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform. 14*, 178–192.

(6) Yu, G., Wang, L. G., and He, Q. Y. (2015) ChIP seeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics 31*, 2382–2383.

(7) Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., and Carey, V. J. (2013) Software for Computing and Annotating Genomic Ranges. *PLoS Comput. Biol. 9*, 1–10.

(8) Quinlan, A. R., and Hall, I. M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics 26*, 841–842.

(9) Hon, J., Martínek, T., Zendulka, J., and Lexa, M. (2017) pqsfinder: an exhaustive and imperfection-tolerant search tool for potential quadruplex-forming sequences in R. *Bioinformatics 33*, 3373–3379.

(10) Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., and Manke, T. (2014) DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res. 42*, 187–191.