

OPEN

# Genomic alterations and abnormal expression of APE2 in multiple cancers

Katherine A. Jensen<sup>1,2</sup>, Xinghua Shi<sup>2,3\*</sup> & Shan Yan<sup>1\*</sup>

Although APE2 plays essential roles in base excision repair and ATR-Chk1 DNA damage response (DDR) pathways, it remains unknown how the APE2 gene is altered in the human genome and whether APE2 is differentially expressed in cancer patients. Here, we report multiple-cancer analyses of APE2 genomic alterations and mRNA expression from cancer patients using available data from The Cancer Genome Atlas (TCGA). We observe that APE2 genomic alterations occur at ~17% frequency in 14 cancer types ( $n = 21,769$ ). Most frequent somatic mutations of APE2 appear in uterus (2.89%) and skin (2.47%) tumor samples. Furthermore, APE2 expression is upregulated in tumor tissue compared with matched non-malignant tissue across 5 cancer types including kidney, breast, lung, liver, and uterine cancers, but not in prostate cancer. We also examine the mRNA expression of 13 other DNA repair and DDR genes from matched samples for 6 cancer types. We show that APE2 mRNA expression is positively correlated with PCNA, APE1, XRCC1, PARP1, Chk1, and Chk2 across these 6 tumor tissue types; however, groupings of other DNA repair and DDR genes are correlated with APE2 with different patterns in different cancer types. Taken together, this study demonstrates alterations and abnormal expression of APE2 from multiple cancers.

Genomic stability is constantly susceptible to irregularities that arise from endogenous and exogenous sources. One of the greatest threats to genomic integrity and implicated as a driving factor in numerous diseases is oxidative stress (OS), which is defined as imbalance between the production of reactive oxygen species (ROS) and antioxidant defense<sup>1-3</sup>. OS leads to various DNA damage such as DNA single-stranded breaks (SSBs) and DNA double-stranded breaks (DSBs). Whereas DSBs are among the most deleterious lesions, SSBs can occur upwards of 10,000 daily per cell, representing the most abundant type of DNA damage<sup>4</sup>. To better preserve integrity of the genome, various DNA repair and DNA damage response (DDR) pathways have evolved to mitigate the deleterious effects of OS, functioning effectively in healthy individuals<sup>4</sup>. It has been widely accepted that deficiencies in DNA repair and DDR pathways have been implicated in human diseases such as cancer and neurodegenerative disorders.

Previous studies have elucidated several critical proteins that regulate and coordinate various DNA repair and DDR pathways. Whereas MRN complex (Mre11, Rad50, and Nbs1) as well as BRCA1 and BRCA2 promote homology recombination (HR)-directed DSB repair<sup>5-7</sup>, Poly(ADP-ribose) polymerase 1 (PARP1) and X-ray repair cross-complementing protein 1 (XRCC1) have been implicated in SSB repair pathway<sup>8</sup>. Generally speaking, DSBs trigger the activation of ATM (ataxia-telangiectasia mutated) – Checkpoint kinase 2 (Chk2) DDR pathway<sup>2,4</sup>. On the other hand, SSBs and DNA replication stress trigger the activation of ATM and Rad3-related (ATR)- Checkpoint kinase 1 (Chk1) DDR pathway<sup>9-11</sup>. Proliferating cell nuclear antigen (PCNA) coordinates the faithful and processive DNA replication and repair of the genome when needed<sup>12,13</sup>. In addition to its critical role in ATR-Chk1 DDR pathway, TopBP1 participates in DSB repair directly<sup>14-17</sup>. Therefore, prior knowledge demonstrates that these proteins crosstalk among different DNA repair and DDR pathways.

Oxidative stress-induced DNA damage is primarily repaired by base excision repair (BER) pathway<sup>2</sup>. Oxidative DNA damage is first recognized and processed by various DNA glycosylases including but not limited to MUTYH, OGG1, NTH1, and NTHL1<sup>18</sup>. Defects in BER genes predispose to hereditary cancers. Whereas

<sup>1</sup>Department of Biological Sciences, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC, 28223, USA. <sup>2</sup>Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, NC, 28223, USA. <sup>3</sup>Present address: Department of Computer & Information Sciences, Temple University, 1925 N. 12th Street, Philadelphia, PA, 19122, USA. \*email: [shan.yan@unc.edu](mailto:shan.yan@unc.edu); [mindyshi@temple.edu](mailto:mindyshi@temple.edu)

MUTYH has been established in the etiology of a colorectal cancer predisposition syndrome<sup>19</sup>, a recent case study reports that MUTYH germline and somatic aberrations are implicated in pancreatic ductal adenocarcinoma and breast cancer oncogenesis<sup>20</sup>. OGG1 is also implicated in association with lung cancer<sup>21</sup>. Furthermore, NTHL1 mutations or deficiencies are associated with multi-cancer phenotype including colorectal cancer and breast cancer<sup>22</sup>. In addition, nucleotide excision repair (NER) protein XPC may regulate OGG1 expression and participates in BER pathway to repair oxidative DNA damage, suggesting crosstalk between different DNA repair pathways<sup>23,24</sup>.

AP endonuclease 1 (APE1, also known as APEX1, Apn1, or Ref-1) and AP endonuclease 2 (APE2, also known as APEX2 or Apn2) have been implicated in many regulatory mechanisms in the maintenance of genome stability, especially in BER pathway<sup>8,25,26</sup>. In addition to its transcriptional activity of redox regulation, APE1 is the main AP endonuclease with high endonuclease activity yet weak exonuclease activity to promote BER<sup>8,25,27</sup>. In contrast, APE2 has high exonuclease activity but weak endonuclease activity<sup>11,28</sup>. APE2 is composed of three functional domains: N-terminal EEP, PCNA-interacting (PIP) motif, and a highly-conserved C-terminal zinc finger Zf-GRF<sup>10</sup>. APE2 interacts with PCNA via two different modes, promoting 3'-5' SSB end resection for the activation of ATR-Chk1 DDR pathway<sup>9-11</sup>. Accumulating evidence has suggested that abnormal expression of APE1 is implicated in cancer such as gastric cancer and ovarian cancer<sup>29-31</sup>. Both APE1 and APE2 have previously been found to be upregulated in multiple myeloma cell lines<sup>32</sup>. However, it remains unknown whether the expression of APE2 in patient-derived tumor tissues is altered when compared with non-malignant tissues across multiple cancer types.

When genetic irregularities arise within DNA repair and DDR mechanisms, OS-induced damage may not be properly repaired, leading to genome instability and compromised protein production<sup>2,3,33</sup>. Copy number variations (CNVs), point mutations, and irregular gene expression patterns are examples of genomic alterations that have the potential to impact protein structure and function. Here, we conduct a multi-cancer bioinformatics analysis of APE2 at DNA, mRNA, and protein levels from publicly available data across multiple studies and cancer tissue types. In this study, we examine genomic alterations (somatic mutations and CNVs) occurring in APE2 *in vivo* from 14 cancer types and if APE2 is differentially expressed in 6 types of tumor tissue compared with non-malignant tissue. Furthermore, we analyze mRNA expression between APE2 and 13 critical DNA repair and DDR proteins. To the best of our knowledge, this work is the first to provide patient-derived evidence showing that abnormal expression of APE2 is implicated in multiple cancer types and that APE2 expression is correlated with the expression of various DNA repair and DDR proteins. We also discuss the function and biology of APE2 in genome integrity.

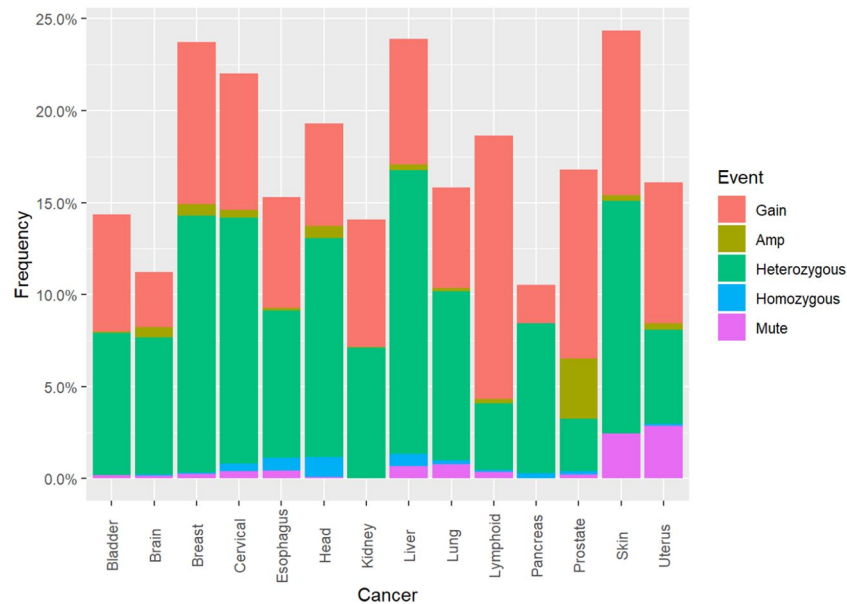
## Methods

**Data sets.** Multiple-study data for genomic alteration events in 14 different cancer types was retrieved from the cBioPortal for Cancer Genomics. Although the cBio repository contains data for 30 different primary sites, only data for these 14 cancer types adhered to the following criteria at the time of download: (1) sample size greater than 250, (2) alteration data available and/or (3) sufficient gene expression data for the same tissue available from TCGA. Alteration events include amplifications (high-level amplification and typically focal), gains (a few additional copies but usually broad), heterozygous deletions (shallow loss), homozygous deletions (deep loss), and protein-level somatic mutations (docs.cbioportal.org). Study-specific details on genomic alteration event data creation can be found within individual studies. A list of all studies - with links - from which datasets have been provided is available on the web (cbioportal.org/datasets).

Gene expression quantification data of sequenced mRNA from TCGA for 6 cancer types was downloaded from Genome Data Portal (GDC) v14.0 via GDC's transfer tool. Sample size  $n \geq 20$  where matched tumor and non-malignant tissue data was available for each individual was the determining factor by which cancer types were chosen for gene expression analysis. See Discussion for more information on an exception made to this criterion. Tissue samples for gene expression quantification data were originally obtained from patients at tissue source sites<sup>34</sup> such as Columbia University, Mayo Clinic, Duke University, and the University of Pittsburgh. Tissue samples were then sent to one of two biospecimen collection sites (BCRs) where RNA was isolated, clinical data standardized, and analyte distribution conducted. Samples and associated data were then sent to a genome characterization or sequencing center (GCC or GSC, respectively) where data was sequenced with Illumina HiSeq technologies. For gene expression quantification, mRNA transcripts were aligned and the .bam files were quantified with HTSeq. Read count data from HTSeq and from STAR are available, as well as fragments per kilobase of transcript per million mapped reads' (FPKM) and FPKM-UQ normalized data. Read count data is often used for global differential expression analysis of gene sets. FPKM normalization method is much more specific, taking into consideration gene length differences and is often a preferred method for gene-gene expression comparisons<sup>35</sup>. Since we focus on single gene and gene-gene pair analysis in this study, we use the FPKM normalized data where read count is divided by gene length.

**Data pre-processing.** Genomic alteration event files containing information on APE2 per cancer type were downloaded from cBioPortal website in TSV format. Files were then programmatically parsed for relevant information from tumor tissue samples, removing duplicate data and any individuals for whom alteration event data was either not available or not explored. To reduce redundancy, only the first record, for only tumor tissue samples which had more than one record, was retained for analyses.

File pre-processing for mRNA gene expression quantification included renaming files from UUID to the associated TCGA barcode. Two new sets of files were created after parsing original files: one set contained data for all tumor data for APE2 and 13 additional genes implicated in DDR pathways, one set contained only expression data from individuals for whom both tumor and matched non-malignant tissue samples were available. Table S3 shows sample sizes for tumor-only tissue and matched tissues per cancer type in our analysis.



**Figure 1.** Frequency of genomic alteration events in APE2 across 14 cancer types and its mRNA expression analysis between tumor tissue and matched non-malignant tissue from 6 different cancer types. cBioPortal analysis of APE2 genomic alterations including gains (Gain), amplifications (Amp), heterozygous deletions, homozygous deletions, and somatic mutations (Mute). Redundant samples were removed before analysis. Multiple somatic mutation events within a single sample and duplicate mutations across different individuals were each distinctly counted. The image was created with PowerPoint and R-package ggplot2 - 'ggplot' (AlterationEvents.Rmd).

**Genomic alteration event analysis.** To characterize the frequency of genomic alteration events in APE2, the frequency of each event type per cancer type was calculated in Excel and plotted in R. Genomic alteration plots were created in R and figures representing APE2 protein domains and mutations – along with R-generated plots – were created in PowerPoint.

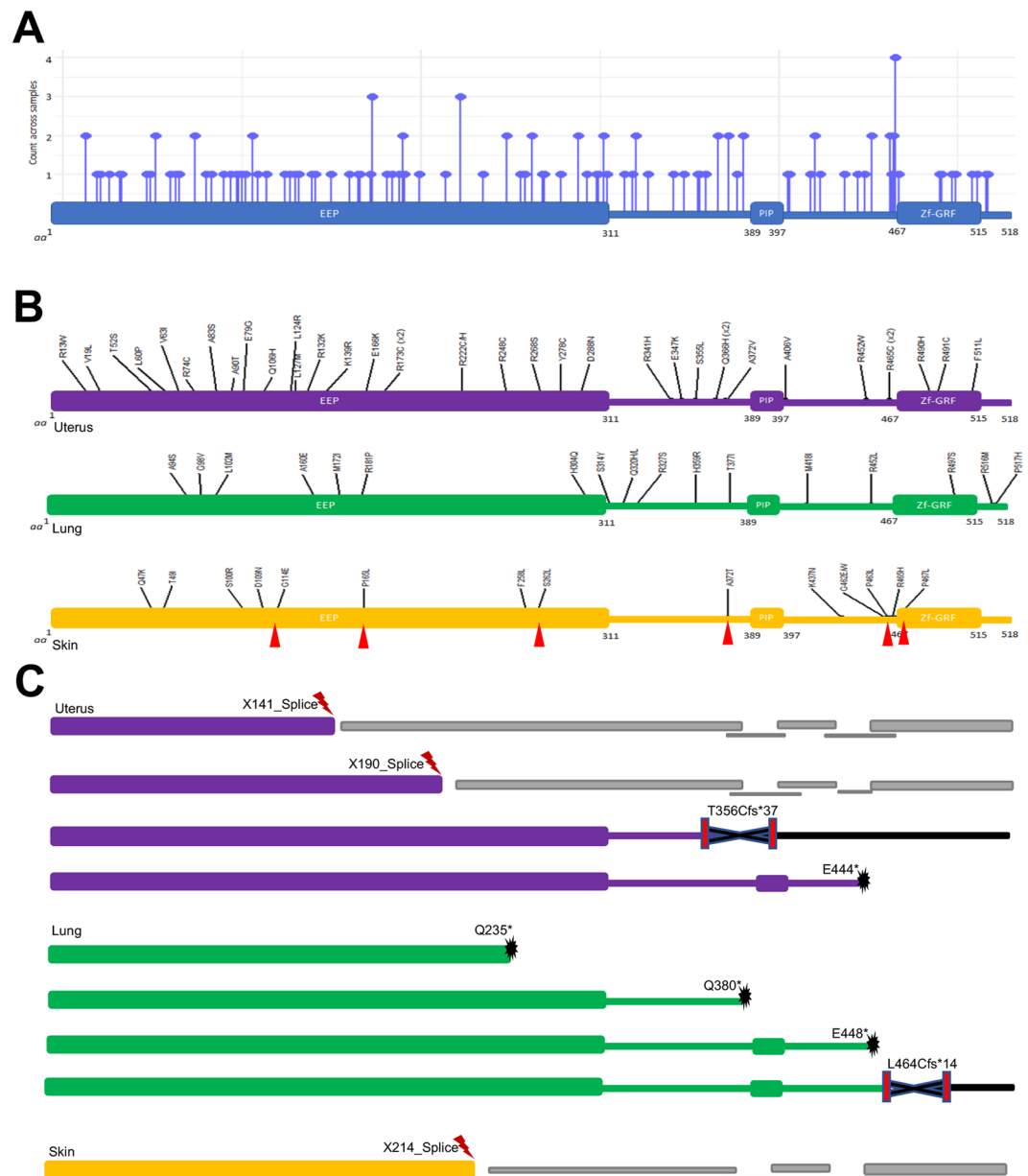
**mRNA analysis.** To find the difference in mRNA expression values of APE2 between tumor and matched non-malignant tissue, a paired two-sided t-test was conducted and boxplots were created in R. To find correlations between mRNA expression of APE2 to other genes implicated in DDR pathways, FPKM values were log-transformed, Pearson's correlation run and scatterplots created in R.

**Bioinformatics tools.** Data from the GDC portal was downloaded using the provided transfer tool. For pre-processing of data, scripts written in Python, utilizing the *os.walk* module, were used to parse, create, and rename files. To find matching TCGA barcode per UUID, *UUIDtoBarcode* from the R-package TCGAutils was used<sup>36</sup>. From R-package *ggplot2* and dependencies, genomic alteration frequency plots and somatic mutation plots were created using *ggplot* and scatterplots created with *ggscatter*<sup>37</sup>. All R-packages were implemented, and statistical analyses and plots were created, with RMarkdown in RStudio 1.1.456<sup>37–41</sup>. All Python scripts were run on Ubuntu 18.04LTS from Windows 64-bit OS. Figures created in PowerPoint were done so using Microsoft Office 2019 PowerPoint Version 1912 (Build 12325.20240 Click-to-Run).

## Results

**Genomic alterations in APE2 across all cancer types.** To verify the occurrence of APE2-situated genomic alterations in cancer patients, data from 14 cancer-types in the cBioPortal database was programmatically analyzed. From all samples for which data were available and had been evaluated for genomic alterations ( $n = 21,769$ ), genomic alterations in APE2 occurred at ~17% frequency and appeared in each cancer type (Fig. 1). Cancers with the highest frequency of total events were skin (24.35%), liver (23.88%), and breast (23.72%) (Fig. 1, Table S1). From the total alteration events observed, heterozygous deletions occurred most frequently (51.18%), gains next (40.77%), then amplifications (3.77%) and mutations (3.11%) with homozygous deletions occurring at the lowest frequency (1.17%). CNVs in APE2 were found at varying levels across each cancer type. Gains occurred most frequently in lymphoid cancer (14.30%), amplifications most often in prostate cancer (3.28%), heterozygous deletions observed most frequently in liver tumor tissue (15.42%), and head and neck cancerous samples revealed the highest frequency of homozygous deletions (1.07%) (Fig. 1, Table S1). Frequency of somatic mutations in APE2 ranged from 2.89% (uterus) to 0.10% (head and neck), although its somatic mutations did not appear in kidney nor pancreas tumor tissue.

**Occurrences of coding truncations and amino acid changes in APE2.** In addition to the analysis of APE2 genomic alterations, protein consequences of somatic mutations in APE2 were analyzed. Protein-level annotations of somatic mutations in APE2 were extracted from genomic alteration event data for 12 cancer types



**Figure 2.** Somatic mutations in APE2 at the protein level across multiple cancer types. **(A)** mutation counts of different mutations in APE2 protein in 12 different cancer types. **(B)** missense mutations in APE2 protein in uterus, skin, and lung cancers. Mutations from non-melanoma patients are highlighted with small red triangles. **(C)** representative splicing and truncations in APE2 protein in uterus, lung, and skin cancers. The images were created by PowerPoint and R-package ggplot2 - 'ggplot' (LolliPlotMutes.Rmd).

(Table S2). A total of 117 mutations were found in APE2 with uterine (40), lung (22), and skin (10 melanoma, and 6 non-melanoma) tumor tissue revealing the highest number of mutations (Fig. 2A, Table S2). Mutation events were comprised mainly of missense mutations with several nonsense mutations which create premature stop codons, frameshift deletions which are deletions of chunks of protein sequence that create a shift in remaining amino acids, and alternative splices denoted with an 'X' indicating a translation termination codon at the site (Fig. 2A–C). Tumor tissues with the highest frequency of somatic mutations in APE2 were uterus (2.89% = 40 out of 1,386), skin (2.47% = 16 out of 649) and lung (0.78% = 22 out of 2,831) tumor samples (Fig. 2B).

Notably, out of the 117 total mutations across multiple cancer types, 29 Arginine residues (~25%) were mutated to other residues, suggesting a distinct feature of APE2 missense mutations (Fig. 2A,B, Table S2). An Arginine to Cysteine mutation was found at position 173 (i.e., R173C) in one breast cancer and two uterine tumor samples. A R222 C and a R222H mutations were found in uterine cancer, and the R222 C mutation was also found in a brain cancer sample. The Arginine residue at position 465 was mutated to Histidine (i.e., R465H) in skin and cervical tumor samples, and to Cysteine (i.e., R465C) in more than one individual with uterine cancer. In particular, there are 8 Arginine mutations localized in the extreme C-terminus Zf-GRF motif of APE2 (i.e.,

R465C, R465H, R490H, R491C, R497S, R499W, R508Q, R516M). Our previous studies have shown that R473A, R473E, R502A, and R502E mutant APE2 in *Xenopus* (homologous to R479 and R508 mutants in human APE2) is deficient in ssDNA binding<sup>10</sup>, and that G483A-R484A double mutants of *Xenopus* APE2 (homologous to G489A-R490A double mutants in human APE2) are defective for ssDNA binding and PCNA interaction<sup>11</sup>. We speculate that the Arginine missense mutations in human APE2 identified from this study may compromise its ssDNA binding and/or PCNA interaction, leading to compromised exonuclease activity and associated genome instability.

Further distinction of skin cancer mutations into melanoma and non-melanoma tumor-types revealed three of the six non-melanoma mutations were Proline to Leucine residue changes (i.e., P165L, P463L and P467L) (Fig. 2B, Table S2). Interestingly, P463L and P467L flank the start of the Zf-GRF domain and P165L falls in the middle of the EEP domain.

Additionally, four nonsense mutations of APE2 were found in lung cancers (Q235\*, Q380\*, and E448\*) and uterine cancer (E444\*) creating premature stop codons (Fig. 2C). Of the two frameshift deletions found, one was in uterine tissue (T356Cfs\*37) and one in lung cancer (L464Cfs\*14) (Fig. 2C). Furthermore, three alternative splices were revealed in uterus (X141, X190) and skin cancers (X214) (Fig. 2C).

**Upregulation of APE2 mRNA expression in tumor tissue.** After establishing the occurrence of alterations in APE2 at the genomic level, APE2 expression at the mRNA level was analyzed. To find if APE2 is differentially expressed in cancer patients, gene expression quantification data was computationally compared between tumor and non-malignant tissue using a two-sided *t*-test in *R*. Samples referred to as ‘matched’ indicate individuals where non-malignant and tumor tissue were both available in the data (Tables S3 and S4). FPKM (fragments per kilobase of exon model per million reads mapped) values of mRNA sequencing data for APE2 in matched samples revealed significant upregulation ( $\alpha = 0.05$ ) in tumor tissue compared with non-malignant tissues across 5 cancer types including kidney ( $n = 126$ ), breast ( $n = 112$ ), lung ( $n = 106$ ), liver ( $n = 58$ ), and uterine ( $n = 23$ ) cancers (Fig. 3A–D,F). Matched samples from prostate cancer patients ( $n = 52$ ) did not present a significant difference in APE2 FPKM values (Fig. 3E).

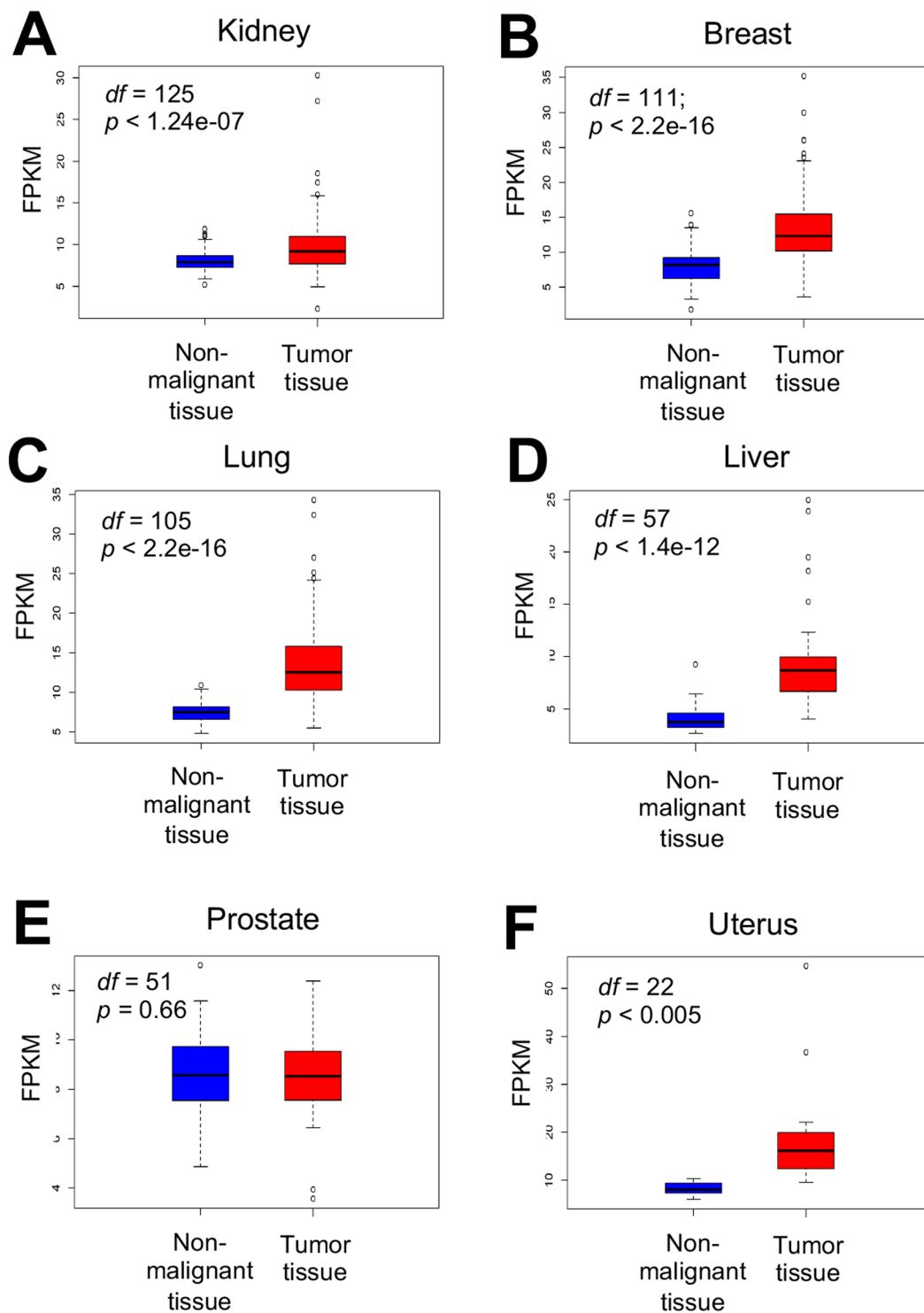
To establish a baseline for finding correlations - between expression patterns in APE2 and 13 other DNA repair and DDR pathway genes - in tumor tissue, matched samples were further analyzed for differential expression in these genes. Notably, in most, if not all, cancer types analyzed, mRNA expression of 11 out of the 13 DNA repair and DDR genes are upregulated in tumor tissue compared with that in non-malignant tissue. Analysis of these genes revealed significant upregulation in APE1, BRCA1, Chk1, Chk2, and TopBP1 for all cancer types (Figs. S1–S5). Like APE2, PCNA, BRCA2 and XRCC1 were also significantly upregulated in all cancer types except prostate (Figs. S6–S8). PARP1 was upregulated in all cancer types except kidney cancer (Fig. S9). mRNA expression of ATR in tumor tissue was increased compared with matched non-malignant tissue in liver, lung, uterus, and prostate cancers, but not in breast and kidney cancers (Fig. S10). Rad50 mRNA expression was upregulated in almost all cancer types except uterine cancer (Fig. S11).

Intriguingly, mRNA expression of ATM and Mre11 was upregulated in tumor tissue compared with matched samples in some cancer types but downregulated in other cancer types (Figs. S12–S13). ATM expression was upregulated in kidney and liver cancers, but not in lung and prostate cancers, while it was downregulated in breast and uterine cancers (Fig. S12). Mre11 expression was upregulated in lung and liver cancers, but not in kidney, prostate, and uterus cancers, and was significantly downregulated in breast cancer (Fig. S13).

**Patterns in mRNA expression between APE2 and other DNA repair and DDR genes.** Because co-expression and interactive partners of APE2 are not yet well known, it is significant to compare expression levels of APE2 to other DNA repair and DDR genes. After establishing differential expression patterns for each gene per cancer types, FPKM values of APE2 were correlated with 13 other DNA repair and DDR genes in tumor tissue only for breast ( $n = 1,105$ ), lung ( $n = 1,028$ ), kidney ( $n = 891$ ), uterine ( $n = 552$ ), prostate ( $n = 499$ ), and liver ( $n = 407$ ) cancers (Tables S3 and S5). While APE2 had a significant ( $\alpha = 0.05$ ) positive correlation with PCNA, APE1, XRCC1, PARP1, Chk1, and Chk2 across all 6 cancer types, groupings of DNA repair and DDR genes were found to be correlated with APE2 in different patterns in different cancer types (Figs. 4, and S14–S18). In liver cancer, correlations of APE2 with these 13 DNA repair and DDR genes were all positive (Fig. S14). Out of the 13 other genes, 11 had the significantly strongest relationship per gene with APE2 in liver cancer (determined by *R*-value and  $\alpha = 0.05$ ) (Fig. S14); however, direction, strength, and significance of correlations varied in other cancer types.

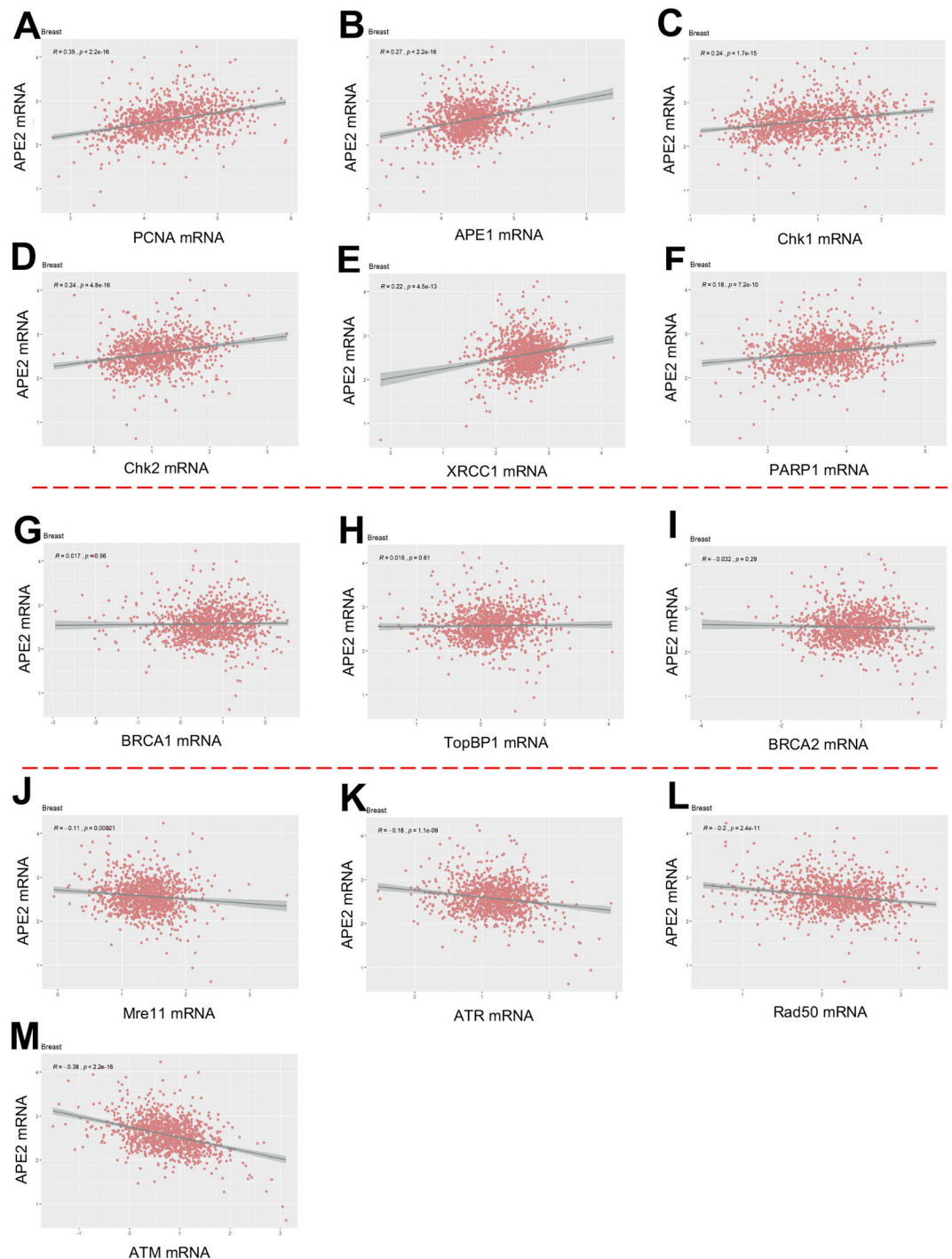
**PCNA, APE1, XRCC1, and PARP1.** Across all 6 cancer types, APE2 expression was significantly positively correlated with the expression of PCNA, APE1, PARP1 and XRCC1 yet expression ranges varied per cancer type and gene-gene pair. The positive correlation between the expression of APE2 and PCNA was most significant in prostate cancer. However, neither APE2 nor PCNA expression values was significantly different in matched prostate tumor (Figs. 3E and S6E), thus there is no baseline to strengthen the implications of this APE2-PCNA correlation. Weak APE2-APE1 expression correlations were observed in lung and kidney cancers, which is consistent with the significant role of APE1 and APE2 in SSB repair<sup>11,42,43</sup>. When APE2 and PARP1 values were correlated, liver and uterine cancer samples showed more variability whereas very little spread occurred in breast and kidney cancer samples. The highest positive correlation of APE2 and XRCC1 occurred in liver cancer.

**Chk1 and Chk2.** APE2 mRNA expression was positively and significantly correlated with the mRNA expression of Chk1 and Chk2 in all 6 cancer types (breast, liver, prostate, uterine, lung, and kidney) (Figs. 4, S14–S18). It is noted that the strongest positive correlation of APE2 with Chk1 and Chk2 is found in lung cancer (Fig. S17).



**Figure 3.** APE2 mRNA expression between tumor tissue and matched non-malignant tissue per individual from 6 different cancer types including kidney (A), breast (B), lung (C), liver (D), prostate (E), and uterus (F). A paired two-sided *t*-test was conducted to determine significance of difference. *df* (degree of freedom) and *p* values are listed in each panel. The images were created by PowerPoint and R-default 'boxplot' (matchedGeneExp.Rmd).

**BRCA1 and BRCA2.** When APE2 expression was correlated with expression of BRCA1 and BRCA2, the top 4 strongest positive correlations were, in order starting with strongest, liver, lung, prostate, and uterus. Whereas there were no significant correlations of APE2 to BRCA1 and BRCA2 in breast cancer (Fig. 4), kidney cancer samples exhibited a negative correlation of APE2-BRCA2 and a positive correlation of APE2-BRCA1 (Fig. S18, and Table S5).



**Figure 4.** Correlation between mRNA expression of APE2 and 13 other DNA repair and DDR proteins in tumor tissues of breast cancer. (A–F) are positive correlation of APE2 mRNA expression with that of PCNA (A), APE1 (B) Chk1 (C), Chk2 (D), XRCC1 (E), and PARP1 (F). (G–I) No correlation of APE2 mRNA expression with that of BRCA1 (G), TopBP1 (H), and BRCA2 (I). (J–M) show negative correlation of APE2 mRNA expression with that of Mre11 (J), ATR (K), Rad50 (L), and ATM (M). Pearson's  $R$  and  $p$  values are listed in each panel. The images were created by PowerPoint and R-package ggplot2 - 'ggscatter' (GeneExpressionAllCancer.Rmd).

**TopBP1 and ATR.** The APE2-ATR relationship was positive in liver and lung cancer but negative in breast cancer. Furthermore, the APE2-ATR correlation was not significant in prostate, uterine, nor kidney cancers. Although a positive correlation of APE2-TopBP1 was revealed in liver, prostate, uterus, and lung cancers, the APE2-TopBP1 correlation was not significant in kidney nor breast cancer.

**ATM, Mre11, and Rad50.** The APE2-ATM correlation was negative in all cancer types except liver. APE2 did not have a significant relationship with Mre11 or Rad50 in uterine nor prostate cancers. The correlation of APE2 with Mre11 was positive in liver and lung cancers, and negative in breast and kidney cancers. In addition, the APE2-Rad50 correlation was negative in breast, kidney, and lung cancers, yet positive in liver cancer.

## Discussion

**Function and biology of APE2.** The human APE2 gene was first cloned and briefly characterized in 2000<sup>44</sup>. However, the understanding of APE2 in DNA repair and DDR pathways has been derived from studies of model organisms such as *Xenopus* and yeast<sup>9–11,45–49</sup>, despite some biochemical and sub-cellular localization characterization of human APE2 protein<sup>28,50,51</sup>. Recent genetic screens identified APE2 as a synthetic lethal target in BRCA1- and BRCA2-deficient colonic and ovarian cancer cell lines<sup>52</sup>. Although the exact underlying mechanism remains unknown, this report suggests that APE2 may contribute to different DNA repair pathways other than BRCA1- and BRCA2-mediated DSB repair. Consistent with this, our series of studies using *Xenopus* egg extract system suggest that APE2 plays a direct role in SSB repair via the 3'-5' SSB end resection<sup>11,26,43</sup>. Of note, gene expression of APE2 and APE1 has been found up-regulated in multiple myeloma (MM) patients and MM cells, which may lead to dysregulation of HR via regulating Rad51 expression<sup>32</sup>. These new evidences suggest that APE2 contributes to genome integrity via different mechanisms.

A prior study has shown that APE2-knock out (KO) mice are viable but develop immune response defects and growth retardation<sup>53</sup>. Subsequent characterization of APE2-KO mice revealed the significance of APE2 in B cell development and immunoglobulin class switch recombination<sup>54,55</sup>. Interestingly, recent evidences suggest that ATR and BER pathways are involved in regulating the expression of Programmed death-ligand 1 (PD-L1), an important player for immunotherapy<sup>56–58</sup>. As APE2 has been shown in both BER and ATR pathways<sup>11,50</sup>, it is interesting to test whether APE2 is directly involved in cancer immunotherapy in future studies.

Due to the lack of in-dept knowledge about APE2 functions in human diseases, the purpose of this study has been to contribute knowledge from such an angle to investigate the genomic alterations and abnormal expression of APE2 in human cancers. Using data from samples across multiple cancer types has been advantageous in providing a snapshot of APE2 characteristics in cancer yet there are limitations which must also be acknowledged.

**Genomic alterations of APE2 in cancer samples.** Our data from 21,769 cancer patients revealed that genomic alterations of APE2 occur in all cancer types analyzed. This observation suggests the potential involvement of APE2 in cancer development, although future studies are needed to test this question directly. Frequencies of CNVs in APE2 for gains, amplifications, heterozygous deletions, and homozygous deletions vary significantly per cancer type. It is important to note that prior studies from which data was obtained did not provide CNV information for all patients<sup>32</sup> and frequencies observed in this study may be relatively lower than those in the actual disease population, warranting further investigation.

We note some unique features of APE2 somatic mutations. The relatively high rate of Arginine mutations dominating APE2 somatic mutations across several cancer types are a distinct feature. The identified Proline to Leucine mutations flanking the Zf-GRF domain (P463L and P467L) may affect its secondary structure and potential function to interact with ssDNA and PCNA<sup>10</sup>. Future investigations are needed to test whether these point mutations in APE2 affect its function in SSB repair and signaling via different capacity of interacting with ssDNA and PCNA or its nuclease activity *per se*.

**Abnormal expression of APE2 in cancer patients.** Our results on APE2 mRNA expression from matched tumor and non-malignant tissue demonstrate its overexpression in 5 out of 6 cancer types (Fig. 3). This pattern of APE2 overexpression is also found in PCNA, BRCA2, and XRCC1 (Figs. S6–S8). Consistent with our findings, Kumar *et al.* recently reported that APE2 is overexpressed at the levels of mRNA and protein in MM cell lines and 112 MM patients from two datasets<sup>32</sup>. Previous studies have demonstrated that PCNA interacts with APE2 to regulate its exonuclease activity<sup>11,28,45,47,50</sup>. A recent genetic screen has revealed that APE2 is a synthetic lethal target in BRCA2-deficient cells<sup>52</sup>. Although APE2 and XRCC1 are involved in BER pathway, it seems that they contribute to SSB repair via different mechanisms<sup>43</sup>.

A prior study revealed that mRNA expression of APE1 and PARP1 is upregulated in tumor tissue compared with that in non-malignant tissues from 53 paired colorectal cancer patients<sup>59</sup>. This is consistent with our observation of upregulation of APE1 mRNA (Fig. S1), and increased PARP1 mRNA expression (Fig. S9). Although upregulation or downregulation of individual genes in the BER pathway is often found in tumor tissue over non-malignant tissue, overall BER capacity has been proposed a determinant of prognosis and therapy response to DNA-damaging 5-fluorouracil in colon cancer patients<sup>60</sup>.

**Correlation of APE2 mRNA expression with other DNA repair and DDR proteins.** We have shown that APE2 expression is positively correlated with the expression of PCNA, APE1, PARP1, XRCC1, Chk1, and Chk2 across all six cancer types analyzed in this work (Figs. 4, S14–S18). This may be interpreted that APE2 is involved in BER and SSB repair pathways as same as PCNA, APE1, XRCC1, and PARP1<sup>11,28,42,43</sup>. Recent studies showing the role of APE2 in Chk1 phosphorylation in oxidative stress could provide feasible explanation for the positive correlation found between APE2 and Chk1<sup>9,10</sup>. Interestingly, the negative correlation of APE2 expression with ATM expression in all cancer types analyzed (excepting liver), suggests that APE2 and ATM may be in different DDR pathways.

Due to tumor heterogeneity, there is no 'one-size-fits-all' formula that can be applied to the prediction, prevention, or therapy of cancer or other human diseases. In this study we demonstrate the varying gene expression and genomic alteration profiles of APE2 by tissue site and even by individual in multiple cancers. Future studies that include more tissue sites, additional metadata, and consider other potential molecular partners of APE2 will be useful in establishing the genomic landscape in which certain APE2 characteristics can be implicated as predictive markers or therapeutic targets.



**Consideration of sample size and data availability.** The presence of matched samples in TCGA dataset provides clear indication of whether APE2 changes in tumor tissue per individual, yet many of the cancer type sample sizes of matched tissues in TCGA were too small to provide statistical value. Due to the prevalence of some cancer types over others, it is infeasible to obtain consistently large quantities of data for all cancer types. Furthermore, some tumor tissue samples are available in ‘study-worthy’ amounts, yet the collection of matched non-malignant tissues is not available due to the nature of the tissue.

Gene expression quantification data was useful in understanding the transcriptome surrounding APE2. In this study, data from multiple cancer types reveals not only APE2 characteristics in disease but suggests tissue-dependent or tissue-specific characteristics. All available data from TCGA has been subject to uniform protocol and strict quality control, indications of valid and consistent datasets. However, TCGA data from bladder and colon cancer studies contained identical gene expression quantification data and thus, due to lack of clarity, were not included in the current study. Protein expression data from The Cancer Proteome Atlas (TCPA) is available for many of the same individuals in the TCGA datasets we have analyzed here. An analysis of matched protein expression could be duly advantageous for examining APE2 transcription levels to the level of expressed protein and to compare translation of APE2 with other DNA repair and DDR genes. Regardless, APE2 is not included in the TCGA/TCPA protein quantification analyses, deeming the protein expression data inefficient for the current study.

Downloadable files of somatic mutation data from cBio contained conflicting information from that which could be visualized on the web interface of cBioPortal. For example, somatic mutations A372T and G462E found in skin cancer were not included in the files but do appear on the cBioPortal website. These missing mutations were found and manually added to analyses. There were a total of 633 melanoma patients whose available CNV information was analyzed and recorded (491 normal copies of APE2; 58 gains; 82 heterozygous deletions; 2 amplifications. Table S2). There were an additional 10 melanoma patients and 6 non-melanoma patients for whom APE2 somatic mutations were recorded. Therefore, the ~24% genomic alteration frequency from skin cancer patients was derived primarily from melanoma patients (Fig. 1).

Likewise, genomic alteration data from cBio is missing for a large number of samples, resulting in slightly less robust results. Furthermore, ploidy and purity of copy number calls could be different across studies. Due to a general lack of prior manual review that could ensure uniform quality of data available from cBioPortal, the CNV data has the potential to represent a number of false positives and false negatives (docs.cbioportal.org). Despite minor limitations with sample sizes and lack of availability for particular data-types, we believe the magnitude and overall quality of data analyzed here has been sufficient to capture certain patient-derived characteristics of APE2 in disease.

## Data availability

The gene expression quantification datasets analyzed during the current study are publicly available at <https://portal.gdc.cancer.gov/repository>. Genomic alteration data is publicly available at <https://www.cbioportal.org/>. GDC’s transfer tool can be downloaded from <https://gdc.cancer.gov/access-data/gdc-data-transfer-tool>. Code is available at <https://github.com/shilab/cancer-data>.

Received: 11 November 2019; Accepted: 13 February 2020;

Published online: 28 February 2020

## References

- Hoch, N. C. *et al.* XRCC1 mutation is associated with PARP1 hyperactivation and cerebellar ataxia. *Nat.* **541**, 87–91, <https://doi.org/10.1038/nature20790> (2017).
- Yan, S., Sorrell, M. & Berman, Z. Functional interplay between ATM/ATR-mediated DNA damage response and DNA repair pathways in oxidative stress. *Cell Mol. Life Sci.* **71**, 3951–3967, <https://doi.org/10.1007/s00018-014-1666-4> (2014).
- Betteridge, D. J. What is oxidative stress? *Metabolism: Clin. Exp.* **49**, 3–8, [https://doi.org/10.1016/s0026-0495\(00\)80077-3](https://doi.org/10.1016/s0026-0495(00)80077-3) (2000).
- Caldecott, K. W. Single-strand break repair and genetic disease. *Nat. Rev. Genet.* **9**, 619–631, <https://doi.org/10.1038/nrg2380> (2008).
- Chen, C. C., Feng, W., Lim, P. X., Kass, E. M. & Jasin, M. Homology-Directed Repair and the Role of BRCA1, BRCA2, and Related Proteins in Genome Integrity and Cancer. *Annu. Rev. Cancer Biol.* **2**, 313–336, <https://doi.org/10.1146/annurev-cancerbio-030617-050502> (2018).
- Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7**, 263–272, [https://doi.org/10.1016/s1097-2765\(01\)00174-5](https://doi.org/10.1016/s1097-2765(01)00174-5) (2001).
- Zhao, W. *et al.* BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. *Nat.* **550**, 360–365, <https://doi.org/10.1038/nature24060> (2017).
- Abbotts, R. & Wilson, D. M. 3rd Coordination of DNA single strand break repair. *Free. Radic. Biol. Med.* **107**, 228–244, <https://doi.org/10.1016/j.freeradbiomed.2016.11.039> (2017).
- Willis, J., Patel, Y., Lentz, B. L. & Yan, S. APE2 is required for ATR-Chk1 checkpoint activation in response to oxidative stress. *Proc. Natl Acad. Sci. USA* **110**, 10592–10597, <https://doi.org/10.1073/pnas.1301445110> (2013).
- Wallace, B. D. *et al.* APE2 ZF-GRF facilitates 3’-5’ resection of DNA damage following oxidative stress. *Proc. Natl Acad. Sci. USA* **114**, 304–309, <https://doi.org/10.1073/pnas.1610011114> (2017).
- Lin, Y. *et al.* APE2 promotes DNA damage response pathway from a single-strand break. *Nucleic Acids Res.* **46**, 2479–2494, <https://doi.org/10.1093/nar/gky020> (2018).
- Choe, K. N. & Moldovan, G. L. Forging Ahead through Darkness: PCNA, Still the Principal Conductor at the Replication Fork. *Mol. Cell* **65**, 380–392, <https://doi.org/10.1016/j.molcel.2016.12.020> (2017).
- Slade, D. Maneuvers on PCNA Rings during DNA Replication and Repair. *Genes (Basel)* **9**, <https://doi.org/10.3390/genes9080416> (2018).
- Liu, Y. *et al.* TOPBP1(Dpb11) plays a conserved role in homologous recombination DNA repair through the coordinated recruitment of 53BP1(Rad9). *J. Cell Biol.* **216**, 623–639, <https://doi.org/10.1083/jcb.201607031> (2017).
- Ogiwara, H. *et al.* Dpb11, the budding yeast homolog of TopBP1, functions with the checkpoint clamp in recombination repair. *Nucleic Acids Res.* **34**, 3389–3398, <https://doi.org/10.1093/nar/gkl411> (2006).

16. Kumagai, A., Lee, J., Yoo, H. Y. & Dunphy, W. G. TopBP1 activates the ATR-ATRIP complex. *Cell* **124**, 943–955, <https://doi.org/10.1016/j.cell.2005.12.041> (2006).
17. Yan, S. & Michael, W. M. TopBP1 and DNA polymerase alpha-mediated recruitment of the 9-1-1 complex to stalled replication forks: implications for a replication restart-based mechanism for ATR checkpoint activation. *Cell Cycle* **8**, 2877–2884, <https://doi.org/10.4161/cc.8.18.9485> (2009).
18. Krokan, H. E. & Bjoras, M. Base excision repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012583, <https://doi.org/10.1101/cshperspect.a012583> (2013).
19. Banda, D. M., Nunez, N. N., Burnside, M. A., Bradshaw, K. M. & David, S. S. Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: Mechanism, metals and medicine. *Free. Radic. Biol. Med.* **107**, 202–215, <https://doi.org/10.1016/j.freeradbiomed.2017.01.008> (2017).
20. Thibodeau, M. L. *et al.* Base excision repair deficiency signatures implicate germline and somatic MUTYH aberrations in pancreatic ductal adenocarcinoma and breast cancer oncogenesis. *Cold Spring Harb Mol Case Stud* **5**, <https://doi.org/10.1101/mcs.a003681> (2019).
21. Vlahopoulos, S., Adamaki, M., Khoury, N., Zoumpourlis, V. & Boldogh, I. Roles of DNA repair enzyme OGG1 in innate immunity and its significance for lung cancer. *Pharm. Ther.* **194**, 59–72, <https://doi.org/10.1016/j.pharmthera.2018.09.004> (2019).
22. Grolleman, J. E. *et al.* Mutational Signature Analysis Reveals NTHL1 Deficiency to Cause a Multi-tumor Phenotype. *Cancer Cell* **35**, 256–266 e255, <https://doi.org/10.1016/j.ccell.2018.12.011> (2019).
23. de Melo, J. T. *et al.* XPC deficiency is related to APE1 and OGG1 expression and function. *Mutat. Res.* **784–785**, 25–33, <https://doi.org/10.1016/j.mrfmmm.2016.01.004> (2016).
24. Melis, J. P. *et al.* Slow accumulation of mutations in Xpc(–/–) mice upon induction of oxidative stress. *DNA Repair.* **12**, 1081–1086, <https://doi.org/10.1016/j.dnarep.2013.08.019> (2013).
25. Tell, G., Quadrifoglio, F., Tiribelli, C. & Kelley, M. R. The many functions of APE1/Ref-1: not only a DNA repair enzyme. *Antioxid. Redox Signal.* **11**, 601–620, <https://doi.org/10.1089/ars.2008.2194> (2009).
26. Hossain, M. A., Lin, Y. & Yan, S. Single-Strand Break End Resection in Genome Integrity: Mechanism and Regulation by APE2. *Inter. J. Mol. Sci.* **19**, 2389, <https://doi.org/10.3390/ijms19082389> (2018).
27. Wilson, D. M. 3rd Properties of and substrate determinants for the exonuclease activity of human apurinic endonuclease Ape1. *J. Mol. Biol.* **330**, 1027–1037, [https://doi.org/10.1016/S0022-2836\(03\)00712-5](https://doi.org/10.1016/S0022-2836(03)00712-5) (2003).
28. Burkovics, P., Hajdu, I., Szukacsov, V., Unk, I. & Haracska, L. Role of PCNA-dependent stimulation of 3'-phosphodiesterase and 3'-5' exonuclease activities of human Ape2 in repair of oxidative DNA damage. *Nucleic Acids Res.* **37**, 4247–4255, <https://doi.org/10.1093/nar/gkp357> (2009).
29. Manoel-Caetano, F. S., Rossi, A. F. T., Calvet de Moraes, G., Severino, F. E. & Silva, A. E. Upregulation of the APE1 and H2AX genes and miRNAs involved in DNA damage response and repair in gastric cancer. *Genes. Dis.* **6**, 176–184, <https://doi.org/10.1016/j.gendis.2019.03.007> (2019).
30. Yuan, C. L. *et al.* APE1 overexpression is associated with poor survival in patients with solid tumors: a meta-analysis. *Oncotarget* **8**, 59720–59728, <https://doi.org/10.18632/oncotarget.19814> (2017).
31. Wen, X. *et al.* APE1 overexpression promotes the progression of ovarian cancer and serves as a potential therapeutic target. *Cancer Biomark* **17**, 313–322, <https://doi.org/10.3233/CBM-160643> (2016).
32. Kumar, S. *et al.* Role of apurinic/apyrimidinic nucleases in the regulation of homologous recombination in myeloma: mechanisms and translational significance. *Blood Cancer J.* **8**, 92, <https://doi.org/10.1038/s41408-018-0129-9> (2018).
33. Gafter-Gvili, A. *et al.* Oxidative stress-induced DNA damage and repair in human peripheral blood mononuclear cells: protective role of hemoglobin. *Plos One* **8**, e68341, <https://doi.org/10.1371/journal.pone.0068341> (2013).
34. Helgason, H. *et al.* Loss-of-function variants in ATM confer risk of gastric cancer. *Nat. Genet.* **47**, 906–910, <https://doi.org/10.1038/ng.3342> (2015).
35. Li, X. *et al.* A comparison of per sample global scaling and per gene normalization methods for differential expression analysis of RNA-seq data. *Plos One* **12**, e0176185, <https://doi.org/10.1371/journal.pone.0176185> (2017).
36. Ramos, M., Schiffer, L. & Waldron, L. TCGAutils: TCGA utility functions for data management. R package version 1.4.0. (2019).
37. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. (2016).
38. Xie, Y., Allaire, J. J. & Grolemund, G. R Markdown: The Definitive Guide. Chapman and Hall/CRC. (2018).
39. Allaire, J. J. *et al.* R: rmarkdown: Dynamic Documents for R. R package version 1.14. (2019).
40. RStudio Team. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. (2015).
41. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2013).
42. Lin, Y. *et al.* APE1 senses DNA single-strand breaks for repair and signaling. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkz1175> (2019).
43. Cupello, S., Lin, Y. & Yan, S. Distinct roles of XRCC1 in genome integrity in *Xenopus* egg extracts. *Biochem. J.* **476**, 3791–3804, <https://doi.org/10.1042/BCJ20190798> (2019).
44. Hadi, M. Z. & Wilson, D. M., 3rd. Second human protein with homology to the *Escherichia coli* abasic endonuclease exonuclease III. *Environ Mol Mutagen* **36**, 312–324, [10.1002/1098-2280\(2000\)36:4<312::AID-EM7>3.0.CO;2-K](https://doi.org/10.1002/1098-2280(2000)36:4<312::AID-EM7>3.0.CO;2-K) (2000).
45. Li, F. *et al.* Apn2 resolves blocked 3' ends and suppresses Top1-induced mutagenesis at genomic rNMP sites. *Nat. Struct. Mol. Biol.* **26**, 155–163, <https://doi.org/10.1038/s41594-019-0186-1> (2019).
46. Johnson, R. E. *et al.* Identification of APN2, the *Saccharomyces cerevisiae* homolog of the major human AP endonuclease HAP1, and its role in the repair of abasic sites. *Genes. Dev.* **12**, 3137–3143, <https://doi.org/10.1101/gad.12.19.3137> (1998).
47. Unk, I. *et al.* Stimulation of 3'→5' exonuclease and 3'-phosphodiesterase activities of yeast apn2 by proliferating cell nuclear antigen. *Mol. Cell Biol.* **22**, 6480–6486, <https://doi.org/10.1128/mcb.22.18.6480-6486.2002> (2002).
48. Ribar, B., Izumi, T. & Mitra, S. The major role of human AP-endonuclease homolog Apn2 in repair of abasic sites in *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **32**, 115–126, <https://doi.org/10.1093/nar/gkh151> (2004).
49. Ma, W., Resnick, M. A. & Gordenin, D. A. Apn1 and Apn2 endonucleases prevent accumulation of repair-associated DNA breaks in budding yeast as revealed by direct chromosomal analysis. *Nucleic Acids Res.* **36**, 1836–1846, <https://doi.org/10.1093/nar/gkm1148> (2008).
50. Tsuchimoto, D. *et al.* Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen. *Nucleic Acids Res.* **29**, 2349–2360, <https://doi.org/10.1093/nar/29.11.2349> (2001).
51. Burkovics, P., Szukacsov, V., Unk, I. & Haracska, L. Human Ape2 protein has a 3'-5' exonuclease activity that acts preferentially on mismatched base pairs. *Nucleic Acids Res.* **34**, 2508–2515, <https://doi.org/10.1093/nar/gkl259> (2006).
52. Mengwasser, K. E. *et al.* Genetic Screens Reveal FEN1 and APEX2 as BRCA2 Synthetic Lethal Targets. *Mol. Cell* **73**, 885–899 e886, <https://doi.org/10.1016/j.molcel.2018.12.008> (2019).
53. Ide, Y. *et al.* Growth retardation and dyslymphopoiesis accompanied by G2/M arrest in APEX2-null mice. *Blood* **104**, 4097–4103, <https://doi.org/10.1182/blood-2004-04-1476> (2004).
54. Guikema, J. E. *et al.* APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. *J. Exp. Med.* **204**, 3017–3026, <https://doi.org/10.1084/jem.20071289> (2007).

55. Guikema, J. E. *et al.* Apurinic/apyrimidinic endonuclease 2 is necessary for normal B cell development and recovery of lymphoid progenitors after chemotherapeutic challenge. *J. Immunol.* **186**, 1943–1950, <https://doi.org/10.4049/jimmunol.1002422> (2011).
56. Permata, T. B. M. *et al.* Base excision repair regulates PD-L1 expression in cancer cells. *Oncogene* **38**, 4452–4466, <https://doi.org/10.1038/s41388-019-0733-6> (2019).
57. Sun, L. L. *et al.* Inhibition of ATR downregulates PD-L1 and sensitizes tumor cells to T cell-mediated killing. *Am. J. Cancer Res.* **8**, 1307–1316 (2018).
58. Vendetti, F. P. *et al.* ATR kinase inhibitor AZD6738 potentiates CD8+ T cell-dependent antitumor activity following radiation. *J. Clin. Invest.* **128**, 3926–3940, <https://doi.org/10.1172/JCI96519> (2018).
59. Slyskova, J. *et al.* Functional, genetic, and epigenetic aspects of base and nucleotide excision repair in colorectal carcinomas. *Clin. Cancer Res.* **18**, 5878–5887, <https://doi.org/10.1158/1078-0432.CCR-12-1380> (2012).
60. Vodenkova, S. *et al.* Base excision repair capacity as a determinant of prognosis and therapy response in colon cancer patients. *DNA Repair.* **72**, 77–85, <https://doi.org/10.1016/j.dnarep.2018.09.006> (2018).

## Acknowledgements

We thank the members from the Yan lab and the Shi lab for critical comments during early stage of the study. The Yan lab is supported, in part, by funds from UNC Charlotte and grants from the NIH/NCI (R01CA225637) and NIH/NIGMS (R15GM114713). The Shi lab is supported by NIH/NHGRI (R15HG009565). K.J. was supported by the Charlotte Research Scholar Program in University of North Carolina at Charlotte.

## Author contributions

K.J., X.S., and S.Y. designed experiments. K.J. performed experiments. K.J., X.S., and S.Y. analyzed the data. K.J., X.S., and S.Y. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-60656-5>.

**Correspondence** and requests for materials should be addressed to X.S. or S.Y.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020