



Article

# Primary Founder Mutations in the *PRKDC* Gene Increase Tumor Mutation Load in Colorectal Cancer

Hajnalka Laura Pálincás<sup>1,2</sup>, Lőrinc Pongor<sup>3,4</sup>, Máté Balajti<sup>3</sup>, Ádám Nagy<sup>3,4</sup>, Kinga Nagy<sup>1,2</sup>, Angéla Békési<sup>1,2</sup>, Giampaolo Bianchini<sup>5</sup>, Beáta G. Vértessy<sup>1,2,\*</sup> and Balázs Gyórfy<sup>3,4,\*</sup>

- <sup>1</sup> Genome Metabolism Research Group, Institute of Enzymology, Research Centre for Natural Sciences, Magyar Tudósok Körútja 2, H-1117 Budapest, Hungary; palinkas.hajnalka@ttk.hu (H.L.P.); nagy.kinga@ttk.hu (K.N.); bekesi.angela@ttk.hu (A.B.)
  - <sup>2</sup> Department of Applied Biotechnology and Food Sciences, BME Budapest University of Technology and Economics, Szt Gellért tér 4, H-1111 Budapest, Hungary
  - <sup>3</sup> TTK Lendület Cancer Biomarker Research Group, Institute of Enzymology, Research Centre for Natural Sciences, Magyar Tudósok Körútja 2, H-1117 Budapest, Hungary; pongorlorinc@gmail.com (L.P.); balajtimate@gmail.com (M.B.); nagy.adam@ttk.hu (Á.N.)
  - <sup>4</sup> Department of Bioinformatics and 2nd Department of Pediatrics, Semmelweis University, Tűzoltó u. 7-9, H-1094 Budapest, Hungary
  - <sup>5</sup> Department of Medical Oncology, San Raffaele Scientific Institute, via Olgettina 60, 20132 Milan, Italy; bianchini.giampaolo@hsr.it
- \* Correspondence: vertessy.beata@ttk.hu (B.G.V.); gyorffy.balazs@ttk.hu (B.G.)

**Abstract:** The clonal composition of a malignant tumor strongly depends on cellular dynamics influenced by the asynchronized loss of DNA repair mechanisms. Here, our aim was to identify founder mutations leading to subsequent boosts in mutation load. The overall mutation burden in 591 colorectal cancer tumors was analyzed, including the mutation status of DNA-repair genes. The number of mutations was first determined across all patients and the proportion of genes having mutation in each percentile was ranked. Early mutations in DNA repair genes preceding a mutational expansion were designated as founder mutations. Survival analysis for gene expression was performed using microarray data with available relapse-free survival. Of the 180 genes involved in DNA repair, the top five founder mutations were in *PRKDC* ( $n = 31$ ), *ATM* ( $n = 26$ ), *POLE* ( $n = 18$ ), *SRCAP* ( $n = 18$ ), and *BRCA2* ( $n = 15$ ). *PRKDC* expression was 6.4-fold higher in tumors compared to normal samples, and higher expression led to longer relapse-free survival in 1211 patients ( $HR = 0.72$ ,  $p = 4.4 \times 10^{-3}$ ). In an experimental setting, the mutational load resulting from UV radiation combined with inhibition of *PRKDC* was analyzed. Upon treatments, the mutational load exposed a significant two-fold increase. Our results suggest *PRKDC* as a new key gene driving tumor heterogeneity.

**Keywords:** cancer; DNA repair; mutation burden; non-homologous end joining; survival; next generation sequencing



**Citation:** Pálincás, H.L.; Pongor, L.; Balajti, M.; Nagy, Á.; Nagy, K.; Békési, A.; Bianchini, G.; Vértessy, B.G.; Gyórfy, B. Primary Founder Mutations in the *PRKDC* Gene Increase Tumor Mutation Load in Colorectal Cancer. *Int. J. Mol. Sci.* **2022**, *23*, 633. <https://doi.org/10.3390/ijms23020633>

Academic Editor: Paola Ghiorzo

Received: 19 October 2021

Accepted: 28 December 2021

Published: 6 January 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Following DNA damage, healthy human cells activate signaling cascades to prevent cell-cycle progression, as well as to initiate repair mechanisms through DNA damage response. Once the accumulated damage is beyond repair, apoptosis is induced [1]. The main mechanisms of DNA repair include base-excision repair (BER), mismatch repair (MMR), nucleotide-excision repair (NER), homologous recombination (HR), and non-homologous end joining (NHEJ).

NER is utilized after UV radiation-induced injury, where the resulting DNA adducts are recognized, and the damaged, short, single strand is removed and resynthesized by a DNA polymerase using the unharmed strand as the template. BER is involved in repairing non-helix distorting base lesions. These lesions are recognized by DNA glycosylases, which remove damaged and inappropriate bases, forming abasic AP sites, cleaved by AP

endonucleases. The single-strand break is then repaired by either short- or long-patch BER. MMR recognizes errors in the course of cell replication, where mismatch repair complexes repair errors occurring on the daughter strand using the parental strand as the template. After double-strand breaks, the damage is either repaired by homologous recombination during S and G2 phases, which enables the repair of the damage, or by NHEJ, where ends with microhomologies are joined. A comprehensive review of DNA repair mechanisms was presented previously [2].

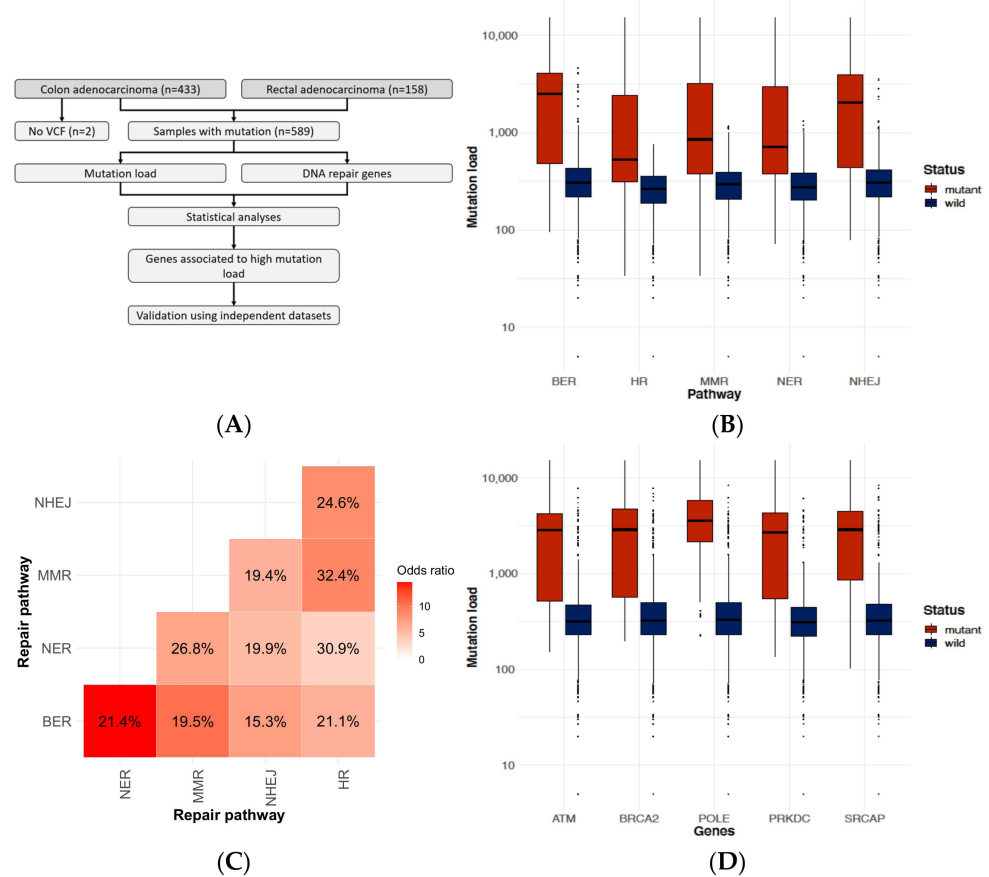
One of the major differences between normal and cancer cells is their response to DNA damage and repair. In normal cells, blocking DNA replication at lesions during double-strand repair leads to the collapse of the replication fork. The collapsed fork is then recognized by ATM and ATR proteins, which, by signaling through TP53, induce apoptosis by provoking BAX translocation and interaction to the mitochondrial voltage-dependent anion channel. In tumor cells, apoptotic pathways are suppressed by mutations in oncogenes (most commonly TP53) [3]. Cancer cells frequently have somatic mutations in DNA-repair genes (*ATM*, *BRCA1/2*, and *FANC* genes) that decrease repair capabilities, leading to the accumulation of mutations. The effect is further enhanced in the case of TP53 double mutants, showing a relationship between DNA repair defects and TP53-induced apoptosis [4].

Our aim was to identify founder mutation events by evaluating a large panel of colon cancer samples analyzed via next generation sequencing. Here, a founder mutation is designated as a genetic variation occurring when a new population is established—in terms of cancer etiology, such a novel population is always based on a clonal expansion of the cancer cells. By comparing mutation frequency, proportion, and estimated mutation time, we aimed to identify founder events leading to increased tumor heterogeneity. In the second part of the study, by utilizing a cell culture framework, we set the goal to validate the presence of a higher mutation load once a founder mutation occurred. Finally, by evaluating a large panel of independent clinical samples, we assessed the clinical relevance of our findings.

## 2. Results

### 2.1. DNA-Repair Gene Defects Cause an Increased Accumulation Rate of Mutations

We investigated mutation distribution in 591 colorectal cancer cases from the TCGA repository (see details in Section 4.2). Mutation burden (total mutation count) was calculated for each sample (for experimental workflow, see Figure 1A). In the case of mutation in any of the DNA repair mechanisms, a significant increase ( $p < 10^{-16}$ ) was observed in the overall mutation burden compared to the wild-type samples (Figure 1B). We also investigated whether a mutation in a DNA repair pathway will result in increased mutation prevalence in another pathway as well. Multiple mutations were not observed in 49.4% of the samples (25.8% had no mutation, while 23.6% had one mutation in a DNA repair pathway). However, once a mutation in one of the pathways occurred, another hit in an alternate pathway was more common (15.3–32.4% pairwise for all cases,  $7.42 \times 10^{-13} > p > 3.58 \times 10^{-34}$ ) (Figure 1C).



**Figure 1.** Characteristics of DNA repair pathway mutations in colorectal cancer. Summary of the analysis pipeline of selecting DNA-repair genes linked to increased mutation burden in colorectal cancer samples of the TCGA repository (A). A higher mutation load (median and interquartile ranges of the total mutation count) was observed when mutations were present in any of the DNA repair pathways (B). Simultaneous mutations in DNA repair pathways show enrichment for each signature as demonstrated by the high odds ratios of co-mutation when comparing DNA repair mechanisms summarized in a pairwise heatmap format. Percentage of patients with co-mutation in pathways is displayed within the tiles of the heatmap (C). Mutation load (boxplots showing median values with lower and upper quartile values and whiskers marking the range) related to mutation in most commonly mutated genes associated with DNA repair (D). BER = base excision repair, HR = homologous recombination repair, MMR = mismatch repair, NER = nucleotide excision repair, and NHEJ = non-homologous end joining repair.

### 2.2. Identifying Genes Where the Mutation Status Is Associated with High Mutation Burden

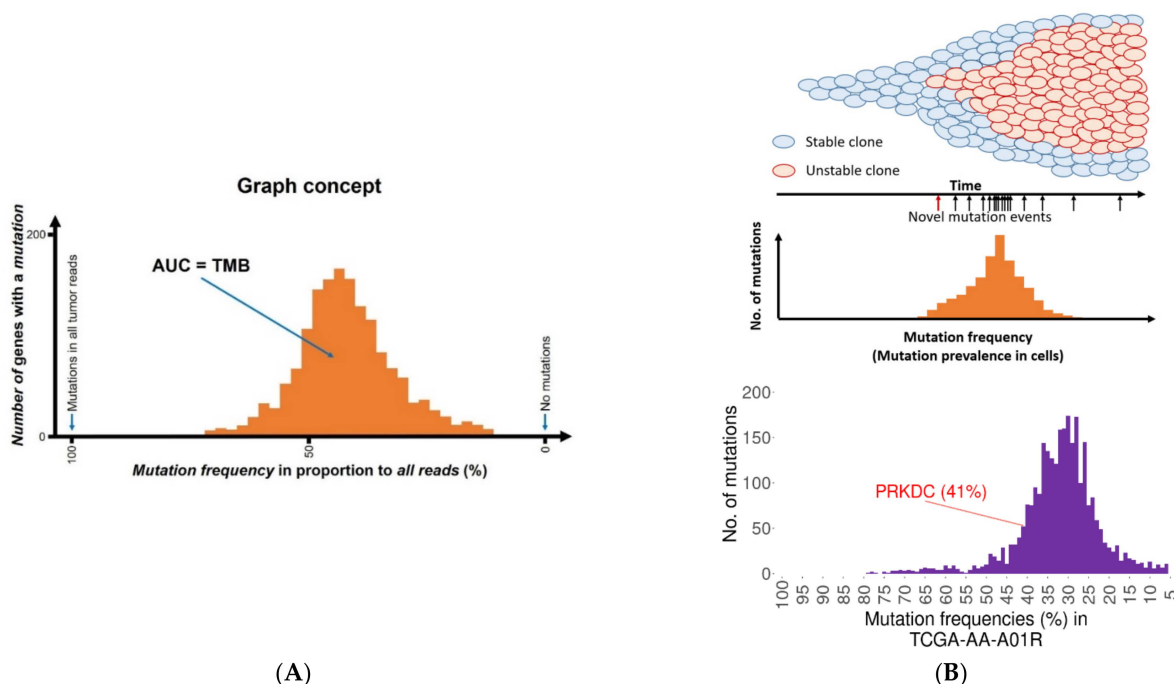
Calculation of mutation burden was performed to evaluate the effect of mutation in each DNA-repair gene and across three independent datasets (namely TCGA, DFCI, and Genentech repositories, see details in Section 4.2). When looking at the top genes in all three datasets (Table 1), three genes (*ATM*, *BRCA2*, and *PRKDC*) were recurrently identified. Samples with DNA-repair mutations in the top genes had a 6–12× increase in median mutation burden compared to wild-type samples (Figure 1D).

**Table 1.** Top genes associated with high mutation burden in colorectal cancer.

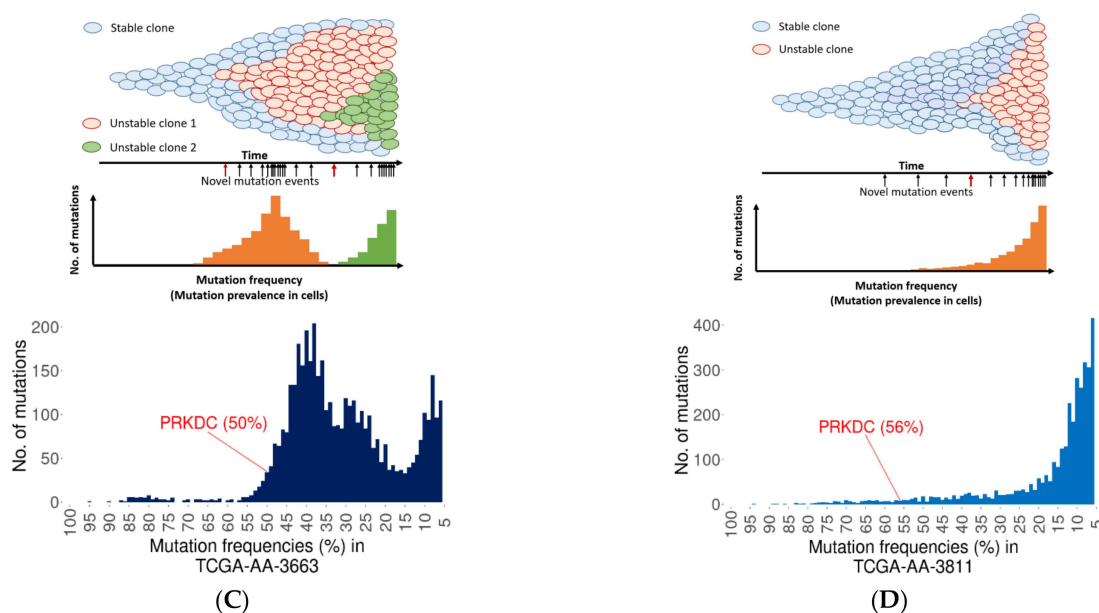
Cohort	Gene	p-Value	Mutation Burden (Mutant Samples)	Mutation Burden (Wild Samples)	Fold Increase	Samples with Mutation (%)	Total Samples
TCGA COAD (mutect2)	<i>PRKDC</i>	$9.5 \times 10^{-28}$	2853	329	8.7	22.0	431
	<i>ATM</i>	$3.7 \times 10^{-24}$	2926	333	8.8	20.0	431
	<i>POLE</i>	$1.3 \times 10^{-22}$	4002	336	11.9	13.9	431
	<i>BRCA2</i>	$1.2 \times 10^{-20}$	2951	335	8.8	15.1	431
	<i>POLD1</i>	$5.1 \times 10^{-18}$	3538	348	10.2	10.2	431
DFCI	<i>POLE</i>	$2.2 \times 10^{-14}$	919	123	7.5	7.3	619
	<i>BRCA2</i>	$1.2 \times 10^{-13}$	948	123	7.7	6.3	619
	<i>ATM</i>	$1.4 \times 10^{-12}$	819	123	6.7	7.4	619
	<i>PRKDC</i>	$1.3 \times 10^{-9}$	765	124	6.2	6.8	619
	<i>MLH3</i>	$1.1 \times 10^{-8}$	1122	126	8.9	3.2	619
GenenTech	<i>ATM</i>	$4.4 \times 10^{-4}$	909	82	11.1	18.1	72
	<i>RAD50</i>	$1.2 \times 10^{-3}$	2393	86	27.8	5.6	72
	<i>BRCA2</i>	$1.5 \times 10^{-3}$	1218	84	14.5	6.9	72
	<i>PRKDC</i>	$2.2 \times 10^{-3}$	1671	86	19.4	5.6	72
	<i>LIG1</i>	$3.2 \times 10^{-3}$	1195	84	14.2	6.9	72

2.3. Founder Mutation and Increased Mutational Load

An important aspect of mutations is not only the sole presence of a mutation, but also its ability to lead to a “mutator” phenotype. To investigate this issue, we identified DNA-repair gene mutations where the mutation developed early (this feature is proven by the high proportion of mutant reads when compared to other mutations) and the bulk of the accumulated mutations occurred later (manifested in lower mutation proportions). Figure 2A summarizes the concept of the analysis. Mutations with high prevalence preceding these lower mutation frequency peaks of mutations were defined as potential founder gene mutations. A total of 125 patients from the TCGA dataset (see details in Materials and Methods) had such a potential founder mutation—35.2% of these resulted in a single accumulation peak (Figure 2B), 21.6% led to double accumulation peaks (Figure 2C), and 14.4% delivered an exponential mutation accumulation (Figure 2D).



**Figure 2.** Cont.

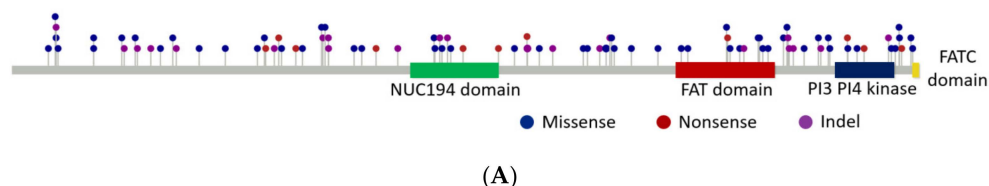


**Figure 2.** Founder mutations lead to an increased mutation load. Theoretical histogram showing the number of genes with a mutation as a function of the mutation frequency in a sample (A). The area under the curve (AUC) equals the total mutation burden (TMB) in a sample. Theoretical and actual representation using a representative example (with TCGA sample identifier) of founder mutation and different mutation load patterns in colorectal cancer patients with a single (B), double (C), and exponential (D) clonal expansion. The red arrows show the incidence of the founder mutation. Mutations in the *PRKDC* gene are indicated in the TCGA samples.

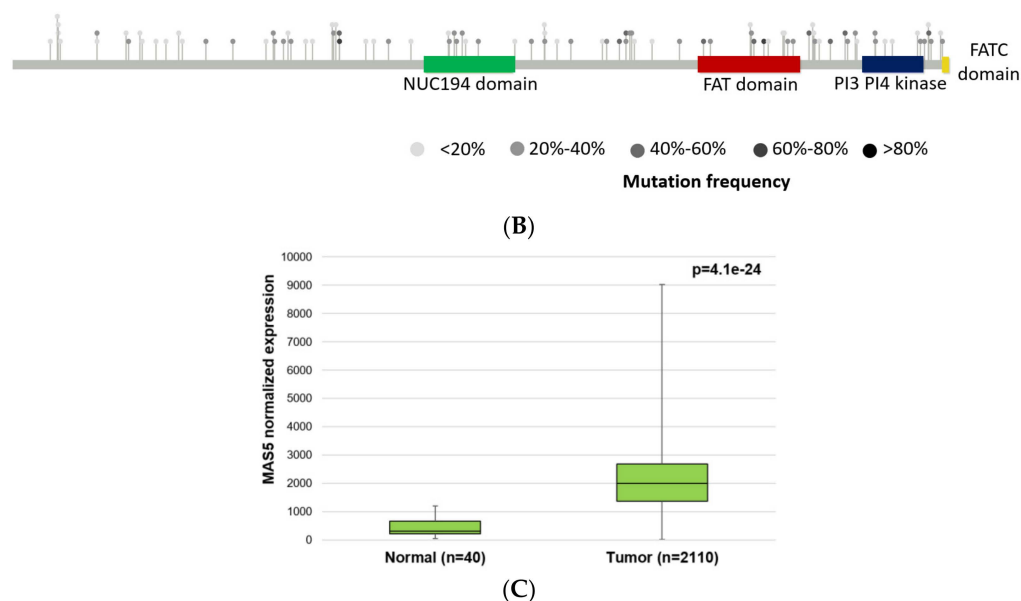
Following a DNA-repair gene mutation, one would expect that mutation rates would increase at similar (proximal) frequencies. Interestingly, DNA-repair mutations did not always lead to this effect, as in many cases the bulk of “randomly” accumulated mutations had a mean frequency difference over 40% compared to the DNA-repair gene defect. However, we have to note that these were patients with very few mutations.

#### 2.4. *PRKDC* Acting as a Mutator Phenotype

We restricted the founder genes to cases where at least half of the randomly accumulated mutations were in a mutation frequency range of 30% after the DNA-repair gene defect. Since one sample might have several DNA-repair genes that passed this criterion, we selected the top three genes for each sample. In this analysis, the top three most common genes with founder mutations were *PRKDC* ( $n = 31$  colorectal cancer samples from the TCGA dataset), *ATM* ( $n = 18$  samples), and *BRCA2* ( $n = 17$  samples). The median mutation frequency of *PRKDC* founder mutations was 69.5% in these patients (range 66.6%–98.9%). Most of the mutations were classified as missense variants, frame shift insertions, and deletions (Figure 3).



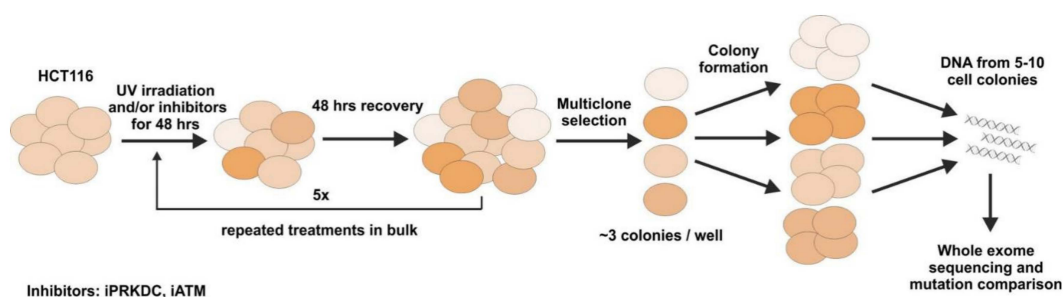
**Figure 3.** Cont.



**Figure 3.** Localization of *PRKDC* mutations show uniform distribution in the entire gene using lollipop plot figures representing mutations in an information-dense manner indicating the positions in amino acid coordinates within the domain organization scheme of the gene (A,B). Some mutations are localized in annotated domains of *PRKDC* (NUC194, FAT, and PI3–PI4 kinase domains), while other mutations are localized in structurally/functionally unknown protein segments (grey line). Distribution of different types of mutations color coded in *PRKDC* (A) and mutation frequency for each sample using different shades of grey (B). Expression of *PRKDC* is significantly higher in colorectal cancer samples (microarray gene expression data from the NCBI Gene Expression Omnibus (GEO) database, see Section 4.8) (C).

### 2.5. Mutagenesis Experiment

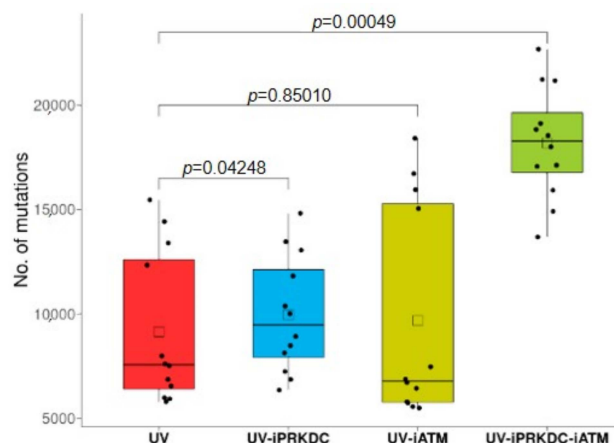
The experimental workflow of the mutagenesis experiments is summarized in Figure 4. For the details of sample collection, please see the Materials and Methods. Altogether, 16 samples were whole-exome sequenced following the mutagenic and inhibitory treatments. These include each treatment (iPRKDC, iATM, and the combination of the two) in triplicates, and the normal control in quadruplicate. The mean number of sequenced reads ranged between 30 and 50 million. All samples fulfilled the minimal quality requirement of at least 25 million reads, with a minimum coverage of 75×. To improve the quality of reads, a trimming was performed, which resulted in the increase of mean quality scores to 33. The alignment of sequences to the reference genome resulted in the placement of 99% of sequences.



**Figure 4.** Experimental overview of the in vitro experiments. Cell lines were treated with UV to induce DNA damage and inhibitors against *PRKDC* and *ATM*. Five rounds of treatment with recovery periods were applied. Multiclonal cell colonies (approx. three colonies/well) were cultured in 96-well plates. DNA was extracted from approximately 10 colonies, creating a polyclonal mixture, where drug-induced mutations were amplified over detection thresholds. Samples were subjected to whole-exome sequencing for mutation analysis.

### 2.6. Effect of PRKDC and ATM1 Inhibition on Mutation Rate in UV-Treated Cell Lines

The average number of mutations in the samples after UV exposure was 9151. This number showed an increase to 9960 in the iPRKDC-treated cell lines ( $p = 0.042$ ) and to a mean of 9681 in the iATM cell lines ( $p = 0.85$ ). The combined treatment of both PRKDC and ATM inhibitors resulted in almost double the number of mutations, to an average of 18,193 ( $p = 0.00049$ ) compared to the samples treated with UV radiation only (Figure 5).



**Figure 5.** The number of mutations show a significant increase in samples treated with UV and either PRKDC or both inhibitors. Boxplot showing the number of mutations including SNVs and indels in the samples treated with UV, iPRKDC, iATM, or both. The mean numbers of mutations in each category are represented by squares.

Of note, high impact mutations of either PRKDC, ATM, or both genes were also found to be present in essentially all of the samples treated with both PRKDC and ATM inhibitors. Around half of these mutations included the acquisition of a stop codon, resulting in a truncated protein. Mutations in other genes with a potentially mutator phenotype, such as POLE, were also present in the samples treated with both inhibitors.

### 2.7. Expression of DNA-Repair Genes and Survival Differences

To address the correlation between survival data and the expression levels of the DNA repair factors (especially those that were selected as top candidates for founder mutators—see Table 1), the microarray gene expression data of 2110 colorectal cancer patients were collected from NCBI Gene Expression Omnibus (GEO) database and analyzed, as described in detail in Section 4.8. Higher expression of the signatures representing each DNA repair pathway in colorectal cancer patients led to longer relapse-free survival (Supplementary Figure S1A). In the case of PRKDC and BRCA2, higher expression resulted in significantly better survival (HR = 0.72,  $p = 0.0044$ ; and HR = 0.73 and  $p = 0.0056$  for PRKDC and BRCA2, respectively, see Supplementary Figure S1B,C). The survival analysis using the mean expression of all DNA-repair genes resulted in the strongest correlation to improved survival (Supplementary Figure S1D). When comparing expression levels of tumor and normal samples, expression of PRKDC, BRCA2, and ATM genes was higher in the tumors. Fold-change increase of expression was over six-fold for PRKDC (Figure 3C), four-fold for BRCA2, and over 1.4 $\times$  for ATM. The survival analysis results are listed in Table 2.

**Table 2.** Results of the survival analysis for the expression of the top genes associated with high mutation burden in colorectal cancer patients. RFS = relapse-free survival, CI = 95% confidence interval.

Gene	HR	CI	p-Value	Expression Fold Change (Tumor vs. Normal)
<i>PRKDC</i>	0.72	0.58–0.9	$4.40 \times 10^{-3}$	6.44
<i>BRCA2</i>	0.73	0.58–0.91	$5.60 \times 10^{-3}$	4.30
<i>RAD50</i>	0.81	0.64–1.02	$6.90 \times 10^{-2}$	2.74
<i>ATM</i>	1.18	0.93–1.49	$1.68 \times 10^{-1}$	1.48
<i>LIG1</i>	1.17	0.94–1.46	$1.70 \times 10^{-1}$	1.40
<i>MLH3</i>	1.48	1.13–1.95	$4.70 \times 10^{-3}$	1.38
<i>POLD1</i>	0.67	0.53–0.85	$8.0 \times 10^{-4}$	1.09
<i>POLE</i>	0.61	0.47–0.77	$5.30 \times 10^{-5}$	0.95

### 3. Discussion

In this study, by combining *in silico* analysis and *in vitro* experiments, we identified mutations as founder events in hypermutating colorectal tumors. The recurrent genes affected by somatic mutations were *PRKDC*, *ATM*, *SCRAP*, *BRCA2*, and *POLE*. Patients with the highest mutation burden had mutations affecting multiple repair pathways. When examining mutation frequencies in patients, we identified four distinct accumulation patterns: neutral, single expansion, double/multiple expansions, and exponential expansion. In the case of non-neutral patterns, the most prevalent mutation harbored at the beginning of expansion was in the *PRKDC* gene.

The *PRKDC* protein (DNA-dependent protein kinase catalytic subunit, also abbreviated as DNA-PKcs) is involved in the DNA repair of double-strand DNA breaks [5,6] and V(D)J recombination [7,8] by non-homologous end joining. As the catalytic subunit of DNA-dependent protein kinase (DNA-PK), it is essential for the NHEJ process. Subsequent to DNA damage, a heterodimer comprised of Ku70 and Ku80 bind to the free DNA end, where *PRKDC* is recruited. In the case of a double-strand break, initial repairs are attempted by NHEJ, followed by HR [9,10]. During the repair process, DNA-PK functions as a “gatekeeper” by regulating DNA access through an autophosphorylation event [11]. By blocking autophosphorylation, DNA becomes inaccessible to the double-strand break repair complex, resulting in impaired HR [12,13].

Inhibition of *PRKDC* leads to enhanced cytotoxicity of radiotherapy treatment [14] and alkylating agents in cancer patients [15]. Mutations of *PRKDC* were significantly associated with a higher mutation load and response to immunotherapy in multiple cancer types [16]. In most responder cases, either truncating mutations or mutations in functional domains were identified. In addition, knockout of *PRKDC* enhanced the efficacy of the anti-programmed cell death protein one in the CT26 animal model, suggesting it as a drug target for immune checkpoint inhibitors [17]. Knockdown of *PRKDC* using small interfering RNA increased the sensitivity of malignant melanoma cells to cisplatin treatment [18]. Inhibition of *PRKDC* in osteosarcoma cell lines increased radiosensitivity, while co-treatment with the *PRKDC* inhibitor KU60648 resulted in enhanced DNA damage [19]. *PRKDC* mutant mice were unable to repair double-strand DNA breaks induced by ionizing radiation [20], resulting in shorter survival times.

On the other hand, the effects of elevated *PRKDC* expression are contradictory. In advanced prostate tumors, *PRKDC* is highly activated, promoting progression and metastasis [21]. In the case of breast cancer, higher expression levels of *PRKDC* were significantly associated with shorter overall survival, higher tumor grade, and positive lymph node status in patients, while downregulation sensitized MCF-7 cell lines to chemotherapeutics *in vitro* and in xenograft models [22]. In contrast, low protein expression of *PRKDC* was associated with a higher tumor grade and mitotic index, as well as survival, in breast cancer [23]. Here, we found that high expression of *PRKDC* resulted in better survival in colorectal cancer. The elevated *PRKDC* expression in tumors compared to normal samples may be a response to either DNA damage or increased replication. Tumors with higher



PRKDC expression may have more efficient DNA damage repair compared to low expressing tumors, while mutations in PRKDC will probably decrease or inhibit its DNA repair activity.

Recent studies have shown that inhibition or deletion and loss of function of PRKDC is compensated by hyperactivation of ATM [24,25], showing an interplay between pathways [26]. The three-dimensional structure of the PRKDC protein suggests competition of binding between Ku80 and BRCA1, and activation of NHEJ and homologous recombination [27], resulting in enhanced targeting of patients with BRCA1 and BER deficiency [28]. By performing mutational load analysis in HCT116 cells, we found that combined inhibition of both PRKDC and ATM proteins, together with UV radiation, led to a significantly increased number of mutations as compared to the radiation effect on its own.

We have to mention some limitations of our study: first, the number of colorectal cancer patients with next generation sequencing data was a restrictive factor. While PRKDC was the top hit with the highest frequency, the actual number of patients was still low. Thus, a future independent study must be performed with a larger patient number to validate our findings. Second, for the in cell validation of the impact of PRKDC and ATM, we used HCT116 cells that were deficient in mismatch repair that might affect the results. However, mismatch repair deficiency frequently occurs among colorectal cancers [29–31] and, in this sense, the HCT116 cell line is a well-established and relevant colorectal cancer model. Still, comparative studies on the MMR-proficient version of HCT116 cells or other cell lines is among our future goals. Third, the overall effect of DNA-repair gene defects was, to a certain degree, limited—this suggests that other factors could also have a significant influence on the mutation load, in addition to DNA repair. Here, we considered the possibility of increased DNA editing activity that might significantly contribute to the increased mutational burden during cancer progression [32–35], even if such events are often episodic, and cells tend to suppress such activities in long term [36]. However, even under such increased DNA editing conditions, the DNA repair capacity of the cells must have an impact. With intact DNA repair, the edited bases often can efficiently be repaired, and not cause increased mutational burden. In this sense, even if a DNA-repair gene mutation appears as a result of a temporary increased DNA editing activity (or due to other DNA damaging effect like oxidative stress or chemotherapy), such mutation can significantly contribute to the following expansion of the mutational load, and therefore might be considered as a founder mutation.

In summary, the most common founder mutation in DNA-repair genes leading to higher subsequent mutation load in colorectal cancer was in PRKDC, a gene involved in non-homologous end joining repair. Our results suggest PRKDC as a new key gene driving tumor heterogeneity.

## 4. Materials and Methods

### 4.1. Repair Gene Database Setup

The genes involved in different routes of DNA damage repair were identified through the KEGG pathway finder [37]. The following maps related to DNA repair and recombination proteins (ko03400) were used: base-excision repair (map03410), nucleotide-excision repair (map03420), mismatch repair (map03430), DSB repair homologous recombination (map03440), and DSB repair non-homologous end joining (map03450). The selected DNA-repair genes used in this study are listed in Supplementary Table S1.

### 4.2. Mutation Database Setup

A total of 433 colon adenocarcinoma and 158 rectal adenocarcinoma patients were included in the analysis. Mutations identified with Mutect2 were downloaded from the TCGA repository. Mutations were filtered based on the judgement system implemented in Mutect2, as well as additional filters, including a minimum of 50× coverage with at least 5× coverage of the mutant reads. Mutation frequency was calculated based on the number of mutant reads and the total coverage of the locus. Overall mutation burden for each

sample was the sum of mutations accepted by the applied filters. To validate the results, we performed cross analysis using three additional datasets obtained from the cBioPortal [38] repository. The analysis was performed using the TCGA provisional ( $n = 431$ ), DFCI [39] ( $n = 619$ ), and the Genentech [40] ( $n = 72$ ) colorectal adenocarcinoma datasets.

#### 4.3. Calculating Correlation between Patient Mutation Burden and Repair Gene Statuses

To identify repair genes linked to higher mutation burden, we calculated the correlation between mutation status of each gene and the overall mutation burden in all samples using the Mann–Whitney U test. The test was performed in the R environment using the *wilcox.test()* function. By this, we identified DNA-repair genes for which a mutation was linked to an increased overall mutation burden in these samples (Figure 1A). We performed the analysis using the mutation status of single genes, as well as the combination of multiple genes involved in a DNA repair mechanism. In this, a mutation could be present in any of the genes involved in the given repair mechanism. Samples were split based on the mutation status of a DNA-repair gene and mutation burden was compared between the two groups using a non-parametric test. To summarize the mutation accumulation timeline for each patient, mutations were represented using histograms, where the y-axis denotes the number of mutations in a given bin and the x-axis represents the mutation frequency based on percentile bins. In the case of selected repair gene mutations, the gene symbol of the repair gene is shown above the histogram.

#### 4.4. Identifying DNA-Repair Associated Founder Mutations

Since mutations in DNA-repair genes occurred in high rates with variable mutation frequencies, we aimed to identify genes where a high mutation prevalence was present before the start of accumulating genetic alterations. In principle, the mutation proportion of the investigated gene is higher than the subsequent mutations because this particular mutation will be present in all descendent cells. We termed these as founder mutations in case at least half the subsequent mutations were within a 30% mutation frequency range following the given founder gene mutation.

#### 4.5. Cell Culture Setup

The HCT116 cell line used in this study was purchased from the European Collection of Cell Cultures (ECACC, Salisbury, UK). Cells were cultured in McCoy's 5A medium (Gibco, Life Technologies, Carlsbad, CA, USA) supplemented with 50  $\mu\text{g}/\text{mL}$  penicillin-streptomycin (Gibco) and 10% fetal bovine serum (Gibco). Cells were maintained at 37 °C in a humidified incubator with 5%  $\text{CO}_2$  atmosphere.

#### 4.6. Mutagenesis Experiments

Cells were subjected to 20  $\text{J}/\text{m}^2$  ultraviolet-C (UV-C) light irradiation, then were grown for 48 h either in the absence or presence of 2  $\mu\text{M}$  NU7441 PRKDC inhibitor (iPRKDC) (also known as KU-57788) (Selleck Chemicals, Munich, Germany) [41]; or 20  $\mu\text{M}$  KU-55933 (Selleck Chemicals) [42], inhibitor of ATM (iATM); or both compounds. The PRKDC inhibitor NU7441 is routinely used in the literature and is a well-characterized specific small molecular compound to counteract PRKDC action [43–45]. After drug treatment, the medium was changed to fresh medium without drugs, allowing recovery for an additional 48 h. Altogether, five rounds of treatments (UV with or without the drugs) were performed in combination with recovery periods. Non-treated cells were handled in parallel without either UV irradiation or the addition of any drug.

Multi-cell clones were isolated by limiting dilution and grown first on 96-well, then on 24-well plates to obtain colonies originating from approximately three individual cell clones. There were no phenotypic differences observed between the colonies grown from either treated or non-treated cells with regard to colony size and proliferation. Finally, cells were harvested by trypsinization and cell pellets were combined in order to have pooled samples grown from approximately ten individual cell clones. For sample collection, four

samples were collected in the case of non-treated cells and three samples were collected from each of the following treatments: UV; UV + iPRKDC; UV + iATM; and UV + iPRKDC + iATM. Genomic DNA was subjected to whole-exome sequencing.

#### 4.7. Analysis of Whole Exome Sequencing Data

Data analysis was performed in the Galaxy platform [46]. Raw sequencing data were checked using FASTQC quality control tool (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>, accessed on 19 October 2021), and sequences with average base quality of less than 20 and a minimum length shorter than 20 were trimmed using Trimmomatic [47].

The alignment of paired-end reads to the reference genome GRCh38 was done using Bowtie2 [48] with default settings. The aligned reads with a MAPQ score of less than 20 and PCR-duplicate reads were filtered using the Picard toolkit (<https://broadinstitute.github.io/picard/>, accessed on 19 October 2021).

Mutations were identified with Mutect2 in the Genome Analysis Toolkit [49] with default settings applied. The number of mutations in normal and UV-treated samples were compared using the Mann–Whitney U test. The test was performed in the R environment using the *wilcox.test()* function.

#### 4.8. Construction of Gene Expression Database

Microarray gene expression data of colorectal cancer patients were obtained from the NCBI Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo/>, accessed on 14 January 2018). Array data files were processed in the R environment (<http://www.r-project.org>, accessed on 19 October 2021). The entire database contains 2110 samples from 15 datasets measured with the Affymetrix Human Genome U133A (GPL96) or the Human Genome U133 Plus 2.0 (GPL570) microarrays. Array quality control was performed for all samples using the “yaqcaffy” (<https://www.bioconductor.org/packages/release/bioc/html/yaqcaffy.html>, accessed on 19 October 2021) library. In this step, we checked the background, raw Q, percentage of present calls, presence of BioB-/C-/D- spikes, GAPDH 3' to 5' ratio, and the beta-actin 3' to 5' ratio. Gene chips were normalized with the MAS5 algorithm using the “affy” (<http://bioconductor.org/packages/release/bioc/html/affy.html>, accessed on 19 October 2021) library [50]. After the normalization, we retained only probes measured on both GPL96 and GPL570 platforms ( $n = 22,277$ ). Then, a second scaling normalization was applied to set the mean expression on each chip to 1000. For genes measured by several probe sets, we used JetSet to select the most reliable probe set [51].

#### 4.9. Analysis of Clinical Samples

Expression in normal and tumor tissues was compared using the Mann–Whitney test. We examined the correlation between the expression of the selected repair genes and relapse-free survival (RFS) using Cox proportional hazard regression and by plotting Kaplan–Meier survival plots. Cox regression analysis was performed using the “survival” R package v2.38 downloaded from CRAN (<https://cran.r-project.org/web/packages/survival/index.html>, accessed on 19 October 2021). Kaplan–Meier plots were generated for each gene separately and for the mean expression of genes involved in functional groups of DNA-repair genes by applying the “survplot” R package v0.0.7 (<http://www.cbs.dtu.dk/~eklund/survplot/>, accessed on 19 October 2021).

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijms23020633/s1>.

**Author Contributions:** Concept and design, B.G. and B.G.V.; cell culture experiments, H.L.P., K.N. and A.B.; survival analysis, G.B. and B.G.; whole-exome sequencing, Á.N.; TCGA data analysis, L.P.; processing of WES data, M.B.; writing of the draft manuscript, B.G., H.L.P. and B.G.V. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in part by the K135231, VEKOP-2.3.2-16-2017-00013, 2018-2.1.17-TET-KR-00001, 2020-1.1.6-JÖVŐ-2021-00013, FIEK\_16-1-2016-0005, and NKP-2018-1.2.1-NKP-2018-00005 grants of the NKFIH and by the Higher Education Institutional Excellence Programme (2020-4.1.1.-TKP2020) of the Ministry for Innovation and Technology in Hungary, and the BME-Biotechnology FIKP grant of EMMI (BME FIKP-BIO).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Singh, R.; Letai, A.; Sarosiek, K. Regulation of Apoptosis in Health and Disease: The Balancing Act of BCL-2 Family Proteins. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 175–193. [[CrossRef](#)]
2. Chatterjee, N.; Walker, G.C. Mechanisms of DNA Damage, Repair and Mutagenesis. *Environ. Mol. Mutagen.* **2017**, *58*, 235–263. [[CrossRef](#)] [[PubMed](#)]
3. Stambolic, V.; MacPherson, D.; Sas, D.; Lin, Y.; Snow, B.; Jang, Y.; Benchimol, S.; Mak, T.W. Regulation of PTEN Transcription by P53. *Mol. Cell* **2001**, *8*, 317–325. [[CrossRef](#)]
4. Geske, F.J.; Nelson, A.C.; Lieberman, R.; Strange, R.; Sun, T.; Gerschenson, L.E. DNA Repair Is Activated in Early Stages of P53-Induced Apoptosis. *Cell Death Differ.* **2000**, *7*, 393–401. [[CrossRef](#)]
5. Falck, J.; Coates, J.; Jackson, S.P. Conserved Modes of Recruitment of ATM, ATR and DNA-PKcs to Sites of DNA Damage. *Nature* **2005**, *434*, 605–611. [[CrossRef](#)]
6. Meek, K.; Dang, V.; Lees-Miller, S.P. Chapter 2 DNA-PK: The Means to Justify the Ends? In *Advances in Immunology*; Academic Press: New York, NY, USA, 2008; Volume 99, pp. 33–58.
7. Björkman, A.; Du, L.; Felgentreff, K.; Rosner, C.; Kamdar, R.P.; Kokaraki, G.; Matsumoto, Y.; Davies, E.G.; Burg, M.; van der Notarangelo, L.D.; et al. DNA-PKcs Is Involved in Ig Class Switch Recombination in Human B Cells. *J. Immunol.* **2015**, *195*, 5608–5615. [[CrossRef](#)]
8. Oksenyich, V.; Kumar, V.; Liu, X.; Guo, C.; Schwer, B.; Zha, S.; Alt, F.W. Functional Redundancy between the XLF and DNA-PKcs DNA Repair Factors in V(D)J Recombination and Nonhomologous DNA End Joining. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2234–2239. [[CrossRef](#)] [[PubMed](#)]
9. Shibata, A.; Conrad, S.; Birraux, J.; Geuting, V.; Barton, O.; Ismail, A.; Kakarougkas, A.; Meek, K.; Taucher-Scholz, G.; Löbrich, M.; et al. Factors Determining DNA Double-Strand Break Repair Pathway Choice in G2 Phase. *EMBO J.* **2011**, *30*, 1079–1092. [[CrossRef](#)]
10. Allen, C.; Halbrook, J.; Nickoloff, J.A. Interactive Competition Between Homologous Recombination and Non-Homologous End Joining11NIH Grant CA77693 to J.A.N. *Mol. Cancer Res.* **2003**, *1*, 913–920.
11. Uematsu, N.; Weterings, E.; Yano, K.; Morotomi-Yano, K.; Jakob, B.; Taucher-Scholz, G.; Mari, P.-O.; van Gent, D.C.; Chen, B.P.C.; Chen, D.J. Autophosphorylation of DNA-PKCS Regulates Its Dynamics at DNA Double-Strand Breaks. *J. Cell Biol.* **2007**, *177*, 219–229. [[CrossRef](#)]
12. Cui, X.; Yu, Y.; Gupta, S.; Cho, Y.-M.; Lees-Miller, S.P.; Meek, K. Autophosphorylation of DNA-Dependent Protein Kinase Regulates DNA End Processing and May Also Alter Double-Strand Break Repair Pathway Choice. *Mol. Cell. Biol.* **2005**, *25*, 10842–10852. [[CrossRef](#)] [[PubMed](#)]
13. Neal, J.A.; Dang, V.; Douglas, P.; Wold, M.S.; Lees-Miller, S.P.; Meek, K. Inhibition of Homologous Recombination by DNA-Dependent Protein Kinase Requires Kinase Activity, Is Titratable, and Is Modulated by Autophosphorylation. *Mol. Cell. Biol.* **2011**, *31*, 1719–1733. [[CrossRef](#)] [[PubMed](#)]
14. Van Oorschot, B.; Granata, G.; Di Franco, S.; Ten Cate, R.; Rodermond, H.M.; Todaro, M.; Medema, J.P.; Franken, N.A.P. Targeting DNA Double Strand Break Repair with Hyperthermia and DNA-PKcs Inhibition to Enhance the Effect of Radiation Treatment. *Oncotarget* **2016**, *7*, 65504–65513. [[CrossRef](#)]
15. Pinel, B.; Duchesne, M.; Godet, J.; Milin, S.; Berger, A.; Wager, M.; Karayan-Tapon, L. Mesenchymal Subtype of Glioblastomas with High DNA-PKcs Expression Is Associated with Better Response to Radiotherapy and Temozolomide. *J. Neurooncol.* **2017**, *132*, 287–294. [[CrossRef](#)] [[PubMed](#)]
16. Chen, M.H.; Tan, K.T.; Cheng, J.H.; Fang, W.-L.; Yeh, Y.C.; Yeh, C.-N.; Chang, Y.C.; Chen, S.-J.; Hsiao, M.; Chao, Y. PRKDC: A New Candidate for Checkpoint Blockade Immunotherapy? *J. Clin. Oncol.* **2017**, *35*, 3022. [[CrossRef](#)]
17. Tan, K.T.; Yeh, C.-N.; Chang, Y.-C.; Cheng, J.-H.; Fang, W.-L.; Yeh, Y.-C.; Wang, Y.-C.; Hsu, D.S.-S.; Wu, C.-E.; Lai, J.-I.; et al. PRKDC: New Biomarker and Drug Target for Checkpoint Blockade Immunotherapy. *J. Immunother. Cancer* **2020**, *8*, e000485. [[CrossRef](#)]

18. Li, N.; Ma, Y.; Ma, L.; Guan, Y.; Ma, L.; Yang, D. MicroRNA-488-3p Sensitizes Malignant Melanoma Cells to Cisplatin by Targeting PRKDC. *Cell Biol. Int.* **2017**, *41*, 622–629. [[CrossRef](#)]
19. Mamo, T.; Mladek, A.C.; Shogren, K.L.; Gustafson, C.; Gupta, S.K.; Riester, S.M.; Maran, A.; Galindo, M.; van Wijnen, A.J.; Sarkaria, J.N.; et al. Inhibiting DNA-PKcs Radiosensitizes Human Osteosarcoma Cells. *Biochem. Biophys. Res. Commun.* **2017**, *486*, 307–313. [[CrossRef](#)]
20. Gurley, K.E.; Ashley, A.K.; Moser, R.D.; Kemp, C.J. Synergy between Prkdc and Trp53 Regulates Stem Cell Proliferation and GI-ARS after Irradiation. *Cell Death Differ.* **2017**, *24*, 1853–1860. [[CrossRef](#)] [[PubMed](#)]
21. Goodwin, J.F.; Kothari, V.; Drake, J.M.; Zhao, S.; Dylgjeri, E.; Dean, J.L.; Schiewer, M.J.; McNair, C.; Jones, J.K.; Aytes, A.; et al. DNA-PKcs-Mediated Transcriptional Regulation Drives Prostate Cancer Progression and Metastasis. *Cancer Cell* **2015**, *28*, 97–113. [[CrossRef](#)]
22. Sun, G.; Yang, L.; Dong, C.; Ma, B.; Shan, M.; Ma, B. PRKDC Regulates Chemosensitivity and Is a Potential Prognostic and Predictive Marker of Response to Adjuvant Chemotherapy in Breast Cancer Patients. *Oncol. Rep.* **2017**, *37*, 3536–3542. [[CrossRef](#)] [[PubMed](#)]
23. Abdel-Fatah, T.; Arora, A.; Agarwal, D.; Moseley, P.; Perry, C.; Thompson, N.; Green, A.R.; Rakha, E.; Chan, S.; Ball, G.; et al. Adverse Prognostic and Predictive Significance of Low DNA-Dependent Protein Kinase Catalytic Subunit (DNA-PKcs) Expression in Early-Stage Breast Cancers. *Breast Cancer Res. Treat.* **2014**, *146*, 309–320. [[CrossRef](#)]
24. Zhou, Y.; Lee, J.-H.; Jiang, W.; Crowe, J.L.; Zha, S.; Paull, T.T. Regulation of the DNA Damage Response by DNA-PKcs Inhibitory Phosphorylation of ATM. *Mol. Cell* **2017**, *65*, 91–104. [[CrossRef](#)]
25. Finzel, A.; Grybowski, A.; Strasen, J.; Cristiano, E.; Loewer, A. Hyperactivation of ATM upon DNA-PKcs Inhibition Modulates P53 Dynamics and Cell Fate in Response to DNA Damage. *Mol. Biol. Cell* **2016**, *27*, 2360–2367. [[CrossRef](#)] [[PubMed](#)]
26. Jiang, W.; Crowe, J.L.; Liu, X.; Nakajima, S.; Wang, Y.; Li, C.; Lee, B.J.; Dubois, R.L.; Liu, C.; Yu, X.; et al. Differential Phosphorylation of DNA-PKcs Regulates the Interplay between End-Processing and End-Ligation during Nonhomologous End-Joining. *Mol. Cell* **2015**, *58*, 172–185. [[CrossRef](#)]
27. Sibanda, B.L.; Chirgadze, D.Y.; Ascher, D.B.; Blundell, T.L. DNA-PKcs Structure Suggests an Allosteric Mechanism Modulating DNA Double-Strand Break Repair. *Science* **2017**, *355*, 520–524. [[CrossRef](#)]
28. Albarakati, N.; Abdel-Fatah, T.M.A.; Doherty, R.; Russell, R.; Agarwal, D.; Moseley, P.; Perry, C.; Arora, A.; Alsubhi, N.; Seedhouse, C.; et al. Targeting BRCA1-BER Deficient Breast Cancer by ATM or DNA-PKcs Blockade Either Alone or in Combination with Cisplatin for Personalized Therapy. *Mol. Oncol.* **2015**, *9*, 204–217. [[CrossRef](#)]
29. Gupta, D.; Heinen, C.D. The Mismatch Repair-Dependent DNA Damage Response: Mechanisms and Implications. *DNA Repair* **2019**, *78*, 60–69. [[CrossRef](#)]
30. Germano, G.; Amirouchene-Angelozzi, N.; Rospo, G.; Bardelli, A. The Clinical Impact of the Genomic Landscape of Mismatch Repair-Deficient Cancers. *Cancer Discov.* **2018**, *8*, 1518–1528. [[CrossRef](#)] [[PubMed](#)]
31. Sekine, S.; Mori, T.; Ogawa, R.; Tanaka, M.; Yoshida, H.; Taniguchi, H.; Nakajima, T.; Sugano, K.; Yoshida, T.; Kato, M.; et al. Mismatch Repair Deficiency Commonly Precedes Adenoma Formation in Lynch Syndrome-Associated Colorectal Tumorigenesis. *Mod. Pathol.* **2017**, *30*, 1144–1151. [[CrossRef](#)] [[PubMed](#)]
32. Langenbucher, A.; Bowen, D.; Sakhtemani, R.; Bournique, E.; Wise, J.F.; Zou, L.; Bhagwat, A.S.; Buisson, R.; Lawrence, M.S. An Extended APOBEC3A Mutation Signature in Cancer. *Nat. Commun.* **2021**, *12*, 1602. [[CrossRef](#)]
33. Shi, M.-J.; Meng, X.-Y.; Fontugne, J.; Chen, C.-L.; Radvanyi, F.; Bernard-Pierrot, I. Identification of New Driver and Passenger Mutations within APOBEC-Induced Hotspot Mutations in Bladder Cancer. *Genome Med.* **2020**, *12*, 85. [[CrossRef](#)]
34. Buisson, R.; Langenbucher, A.; Bowen, D.; Kwan, E.E.; Benes, C.H.; Zou, L.; Lawrence, M.S. Passenger Hotspot Mutations in Cancer Driven by APOBEC3A and Mesoscale Genomic Features. *Science* **2019**, *364*, eaaw2872. [[CrossRef](#)] [[PubMed](#)]
35. Cannataro, V.L.; Gaffney, S.G.; Sasaki, T.; Issaeva, N.; Grewal, N.K.S.; Grandis, J.R.; Yarbrough, W.G.; Burtneis, B.; Anderson, K.S.; Townsend, J.P. APOBEC-Induced Mutations and Their Cancer Effect Size in Head and Neck Squamous Cell Carcinoma. *Oncogene* **2019**, *38*, 3475–3487. [[CrossRef](#)]
36. Petljak, M.; Alexandrov, L.B.; Brammell, J.S.; Price, S.; Wedge, D.C.; Grossmann, S.; Dawson, K.J.; Ju, Y.S.; Iorio, F.; Tubio, J.M.C.; et al. Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. *Cell* **2019**, *176*, 1282–1294.e20. [[CrossRef](#)]
37. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a Reference Resource for Gene and Protein Annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462. [[CrossRef](#)] [[PubMed](#)]
38. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the CBioPortal. *Sci. Signal.* **2013**, *6*, pl1. [[CrossRef](#)] [[PubMed](#)]
39. Giannakis, M.; Mu, X.J.; Shukla, S.A.; Qian, Z.R.; Cohen, O.; Nishihara, R.; Bahl, S.; Cao, Y.; Amin-Mansour, A.; Yamauchi, M.; et al. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep.* **2016**, *15*, 857–865. [[CrossRef](#)]
40. Seshagiri, S.; Stawiski, E.W.; Durinck, S.; Modrusan, Z.; Storm, E.E.; Conboy, C.B.; Chaudhuri, S.; Guan, Y.; Janakiraman, V.; Jaiswal, B.S.; et al. Recurrent R-Spondin Fusions in Colon Cancer. *Nature* **2012**, *488*, 660–664. [[CrossRef](#)]
41. Sun, X.; Liu, T.; Zhao, J.; Xia, H.; Xie, J.; Guo, Y.; Zhong, L.; Li, M.; Yang, Q.; Peng, C.; et al. DNA-PK Deficiency Potentiates CGAS-Mediated Antiviral Innate Immunity. *Nat. Commun.* **2020**, *11*, 6182. [[CrossRef](#)]

42. Zhao, X.; Ren, Y.; Lawlor, M.; Shah, B.D.; Park, P.M.C.; Lwin, T.; Wang, X.; Liu, K.; Wang, M.; Gao, J.; et al. BCL2 Amplicon Loss and Transcriptional Remodeling Drives ABT-199 Resistance in B Cell Lymphoma Models. *Cancer Cell* **2019**, *35*, 752–766.e9. [[CrossRef](#)]
43. Affandi, T.; Ohm, A.M.; Gaillard, D.; Haas, A.; Reyland, M.E. Tyrosine Kinase Inhibitors Protect the Salivary Gland from Radiation Damage by Increasing DNA Double Strand Break Repair. *J. Biol. Chem.* **2021**, *296*, 100401. [[CrossRef](#)] [[PubMed](#)]
44. Wang, G.; Guo, S.; Zhang, W.; Li, Z.; Xu, J.; Li, D.; Wang, Y.; Zhan, Q. A Comprehensive Analysis of Alterations in DNA Damage Repair Pathways Reveals a Potential Way to Enhance the Radio-Sensitivity of Esophageal Squamous Cell Cancer. *Front. Oncol.* **2020**, *10*, 575711. [[CrossRef](#)]
45. Wang, C.; Tang, H.; Geng, A.; Dai, B.; Zhang, H.; Sun, X.; Chen, Y.; Qiao, Z.; Zhu, H.; Yang, J.; et al. Rational Combination Therapy for Hepatocellular Carcinoma with PARP1 and DNA-PK Inhibitors. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 26356–26365. [[CrossRef](#)] [[PubMed](#)]
46. Giardine, B.; Riemer, C.; Hardison, R.C.; Burhans, R.; Elnitski, L.; Shah, P.; Zhang, Y.; Blankenberg, D.; Albert, I.; Taylor, J.; et al. Galaxy: A Platform for Interactive Large-Scale Genome Analysis. *Genome Res.* **2005**, *15*, 1451–1455. [[CrossRef](#)] [[PubMed](#)]
47. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
48. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
49. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytzky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce Framework for Analyzing next-Generation DNA Sequencing Data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)]
50. Gyorffy, B.; Molnar, B.; Lage, H.; Szallasi, Z.; Eklund, A.C. Evaluation of Microarray Preprocessing Algorithms Based on Concordance with RT-PCR in Clinical Samples. *PLoS ONE* **2009**, *4*, e5645. [[CrossRef](#)] [[PubMed](#)]
51. Li, Q.; Birkbak, N.J.; Gyorffy, B.; Szallasi, Z.; Eklund, A.C. Jetset: Selecting the Optimal Microarray Probe Set to Represent a Gene. *BMC Bioinform.* **2011**, *12*, 474. [[CrossRef](#)]