# The DNA repair genes FAM35A and SPIDR are prognostic biomarkers in Glioblastoma and other solid tumors

**May Moe Thu Han**

*Received September 06, 2023*
*Accepted December 10, 2023*
*Electronic access December 31, 2023*

Glioblastoma Multiforme (GBM) is a highly fatal form of brain cancer with meager 5-year survival rates. Despite this, there is a lack of good prognostic biomarkers for this tumor. Previous studies have suggested that DNA repair genes represent potential prognostic and predictive biomarkers in different cancer types. However, there is still limited evidence on the use of DNA repair genes as prognostic biomarkers for GBM. Here, we analyzed the prognostic value of the gene expression of DNA repair genes belonging to the Homologous Recombination (HR) pathway in GBM and other solid tumors. For this purpose, we retrieved data from The Cancer Genome Atlas and used GEPIA2 to associate the gene expression of these genes with patient survival. We discovered 2/37 HR genes, FAM35A and SPIDR, whose expression was significantly associated with patient survival in GBM. Higher expression of FAM35A was associated with a better prognosis, while SPIDR was associated with a worse survival rate. Both genes were also associated with patient survival in other solid tumors. These findings support the importance of DNA repair genes as prognostic markers, particularly from the HR pathway. Further studies should validate these bio-markers in different cohorts of patients and stages of the disease and elucidate the mechanisms behind these findings.

**Keywords:** DNA repair, Biomarker, Glioblastoma, Prognosis, GBM, Homologous Recombination, Tumor, Pathway

## Introduction

Cancer is an umbrella term for several diseases characterized by uncontrolled cell growth [1]. Every year, millions of people's lives are taken away by cancer. Cancer has been typically treated in many ways, including surgery, chemotherapy, and radiotherapy. Most recent approaches include immunotherapy, stem cell therapy, and targeted therapy [2]. Glioblastoma Multiforme (GBM) is one of the most fatal types of cancer, and it represents the most common malignant brain tumor in adults [3]. Seizures, loss of consciousness, headache, speech or visual disturbance, weakness, and confusion are typical symptoms, as with any other tumor site in the brain [4]. Its fatality also comes from not having uncommon clinically meaningful molecular markers that aid in prognosis and therapy response prediction despite the recent advancements in our understanding of the genetic changes in GBM [5]. Once GBM is diagnosed, treatment usually consists of maximal surgical resection, then medication and radiation [6]. However, due to the invasive nature of GBM, complete removal has proven extremely difficult [6]. Despite the available treatments, the 5-year survival rate for adults (>40 years) is 4.3'% [7]. Moreover, due to the metastatic nature of GBM, achieving complete eradication is improbable, leading to an elevated risk of recurrence and its spread across the body [8]. As a result, GBM becomes highly unpredictable and has a poor prognosis.

The DNA can be damaged in various ways, including single- or double-strand breaks (SSB and DSB, respectively) and mismatched DNA strands [9,10]. SSB occurs when one of the two strands of DNA double helix is affected, whereas DSB occurs when both strands are affected. This is when DNA repair genes would come into play. Homologous recombination (HR) is one of most important DNA repair pathways that repair SSB, DSB, and interstrand crosslinks [11] and it is known to be a highly precise DNA repair mechanism [12]. It has been shown that the HR pathway is relevant in the context of cancer biology and disease prognosis. In fact, Homologous recombination deficiency (HRD) is a tumor characteristic that is defined by the inability to accurately repair double-strand breaks (DSBs) in DNA via homologous recombination. It has been shown that HRD drives genomic instability in different cancer types like breast and ovarian and confers therapeutic vulnerability. In GBM, upregulation of some HR repair components correlates with prognosis [13], and this might be related to the fact that augmented HR appears to underpin chemotherapy resistance in GBM [14]. However, a comprehensive analysis on the potential role of HR genes as prognostic markers in GBM is still missing.

In this paper, a comprehensive list of HR DNA repair genes have been collected in order to analyse their potential role as biomarkers in GBM. For this purpose, patient data was retrieved from The Cancer Genome Atlas through GEPIA2 [15]. This tool

© The National High School Journal of Science 2023

was used to identify differentially expressed DNA repair genes in GBMs and determine whether they were associated with survival. Two genes were found to be associated with patient survival in GBM and other solid tumors.

## Results

First, we retrieved a curated list of DNA Repair Genes [16,17] focusing on the Homologous Recombination (HR) pathway and found a total of 37 genes. We analyzed patient data derived from The Cancer Genome Atlas using the GEPIA2 [15] tool to interrogate whether they were differentially expressed in GBM. We found that there were 18 genes differentially expressed, with most genes being upregulated in tumor cells when compared to normal cells (Fig. 1). Even though the trend of these genes is the same, the basal expression of some of these genes in normal cells is different, suggesting that the intrinsic importance of these genes might be different. This may imply that some genes are more relevant in the context of normal cells because of their higher expression.

Next, out of the 18 genes that were differentially expressed, we found that only FAM35A was significantly associated with patient survival among these genes (HR = 0.64, pHR = 0.016, Table 1). SPIDR, although not differentially expressed, was also associated with patient survival (HR = 1.7, pHR = 0.0062). From this, we can see that patients expressing higher levels of FAM35A have a better prognosis while patients expressing higher levels of SPIDR have a worse one, as the genes having a hazard ratio (HR)$>$1 indicate a worse prognosis, while HR$<$1 indicates better prognosis.

Given this information, we also used GEPIA2 [15] to explore whether these two genes have a prognostic value in other cancer types. The results show that FAM35A is also associated with better survival in three tumor types (Kidney cancer (KIRC), Low-Grade Glioma (LGG), and Rectum Adenocarcinoma (READ), but has the opposite effect in Thyroid Carcinoma (THCA) (Fig. 2). Interestingly, FAM35A only showed a significantly different expression in GBM and LGG. On the other hand, SPIDR also has a prognostic value in four other tumor types: LGG, LIHC, SARC, and SKCM. However, SPIDR has a worse prognosis value for all these tumor types except SKCM. Notably, SPIDR is not differentially expressed within any of these tumor types (Fig. 3).

## Discussion

In this paper, we screened 37 DNA repair genes from the Homologous Recombination pathway to validate their prognostic value in GBM. 18 genes were differentially expressed, with most of them being upregulated in tumor cells. Out of these 18 genes, only 2 were significantly associated with patient survival in GBM. Higher expression of FAM35A was associated with a better prognosis. On the other hand, higher expression of SPIDR was associated with a worse prognosis. After that, we interrogated the prognostic value of these two genes in all solid tumor types and expanded the potential prognostic value to these genes.

Indeed, we found the same prognostic value for FAM35A in LGG, KIRC and READ. Likewise, we found the same prognostic value for SPIDR in LGG, LIHC, and SARC.

Previous research has highlighted the potential use of DNA repair genes as prognostic biomarkers. For instance, Lv and colleagues constructed a prognostic DNA repair-related gene signature using 11 genes, and they were able to associate the signature with the response of colorectal cancer patients to immunotherapy [18]. In GBM, it has been shown that upregulation of HR repair components correlates with prognosis [13], and this might be related to the fact that augmented HR appears to underpin chemotherapy resistance in GBM [14]. In this context, resistance is thought to result from the restoration of precise DNA repair of lesions caused by chemotherapy and radiotherapy (i.e., double-strand breaks).

In our analysis, FAM35A and SPIDR were the only two genes that had a prognostic value for GBM. These genes also had different impacts on patient survival in different tumor types. Tomida et al. found that the knockdown of FAM35A caused sensitivity to DNA-damaging agents [19]. In their work, depletion of FAM35A increased resistance to camptothecin in a BRCA1-mutant cell line. This implies that FAM35A might be involved in the processing of DNA ends, allowing for more effective DNA repair. Fackerell et al. also discovered that FAM35A defective cell lines are sensitive to DNA double-strand break inducing agents [20]. It has been proposed that there is a negative correlation between DNA repair deficiency and the prognosis of cancer patients. However, there is still some controversy as some paper show that low activity of the DNA repair system in tumor tissues appears to be associated with a better prognosis [21]. Although we do not have enough evidence for a formal explanation, based on the current literature, we suggest that this association comes from the potentially higher response to DNA-damaging agents used in GBM.

On the other hand, SPIDR is one of the genes that does not show a statistically different gene expression among all solid tumor types but it is correlated with patient survival. In our work, higher expression of SPIDR is linked to a worse prognosis in GBM patients. Hazan et al. supported that the depletion of SPIDR causes an increase in the rate of sister chromatid exchange, HR abnormalities, hypersensitivity to DNA-damaging agents, and overall genome instability [22]. Moreover, it has been shown that SPIDR depletion leads to genome instability and causes hypersensitivity to DNA damaging agents [23]. Although there is still not enough information on the role of SPIDR in cancer, we can conclude from the information available that


© The National High School Journal of Science 2023

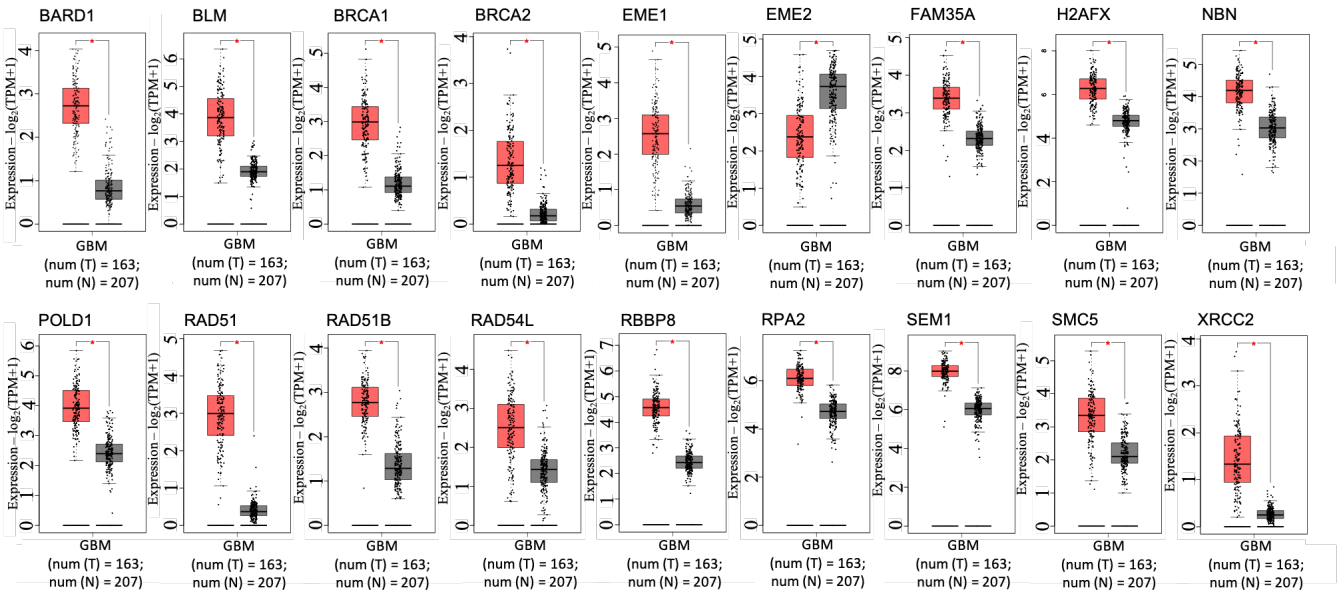


**Fig. 1** The expression level of the DNA repair genes belonging to the Homologous Recombination (HR) pathway in GBM vs. normal cells. The asterisk denotes statistical significance (P<0.01). Red indicates tumor cells, and gray indicates normal cells.

**Table 1** Changes in the gene expression level of Homologous recombination genes associated with patient survival in GBM. FC= Fold Change, HR= Hazard Ratio, p(HR)= p-value of hazard ratio. Statistically significant genes are denoted in bold and italics in the column of p(HR). Red means upregulation in tumor cells, and blue indicates downregulation.

| Gene type | FC | log2 | HR | p(HR) |
|---|---|---|---|---|
| RAD51 | 24 | 4.58 | 0.83 | 0.3 |
| RAD51B | 4.05 | 2.02 | 0.83 | 0.29 |
| RAD51D | 1.36 | 0.44 | 1.1 | 0.45 |
| HELQ | 2.18 | 1.12 | 0.83 | 0.31 |
| SWSAP1 | 2.45 | 1.29 | 1 | 0.92 |
| ZSWIM7 | 1.83 | 0.87 | 1.2 | 0.42 |
| SPIDR | 0.84 | -0.25 | 1.7 | ***0.0062*** |
| PDS5B | 1.12 | 0.16 | 1 | 0.99 |
| DMC1 | 0.6 | -0.74 | 1 | 0.88 |
| XRCC2 | 8 | 3 | 1.2 | 0.31 |
| XRCC3 | 1.66 | 0.73 | 0.96 | 0.82 |
| RAD52 | 0.89 | -0.17 | 0.98 | 0.92 |
| RAD54L | 2.75 | 1.46 | 1 | 0.91 |
| RAD54B | 1.88 | 0.91 | 0.92 | 0.64 |
| BRCA1 | 5.99 | 2.59 | 1.2 | 0.41 |
| BARD1 | 8.01 | 3 | 0.84 | 0.33 |
| PAXIP1 | 2.2 | 1.14 | 1.1 | 0.71 |
| SMC5 | 2.77 | 1.47 | 1.1 | 0.5 |
| SMC6 | 1.58 | 0.66 | 0.84 | 0.35 |

| Gene type | FC | log2 | HR | p(HR) |
|---|---|---|---|---|
| SHLD2 (FAM35A) | 2.38 | 1.25 | 0.64 | ***0.016*** |
| SEM1 | 3.84 | 1.94 | 1.4 | 0.068 |
| RAD50 | 1.73 | 0.79 | 1.1 | 0.56 |
| NBN | 2.42 | 1.28 | 0.88 | 0.5 |
| RBBP8 | 5.18 | 2.37 | 1.2 | 0.26 |
| MUS81 | 1.36 | 0.44 | 0.86 | 0.41 |
| EME1 | 10.74 | 3.42 | 0.95 | 0.78 |
| EME2 | 0.34 | -1.56 | 1.3 | 0.15 |
| SLX1B (GIYD2) | 0.92 | -0.12 | 1.2 | 0.36 |
| GEN1 | 2.43 | 1.28 | 1 | 0.86 |
| ATM | 1.49 | 0.58 | 0.83 | 0.31 |
| BLM | 4.97 | 2.31 | 1.2 | 0.4 |
| BRCA2 | 10.62 | 3.41 | 0.91 | 0.62 |
| H2AFX | 2.86 | 1.52 | 0.93 | 0.67 |
| POLD1 | 3.31 | 1.73 | 1.3 | 0.11 |
| RAD51C | 1.16 | 0.21 | 1.2 | 0.39 |
| RPA2 | 2.65 | 1.41 | 0.98 | 0.93 |
| TP53BP1 | 1.2 | 0.26 | 0.89 | 0.54 |

patients with high level of SPIDR might be more resistant to therapies damaging the DNA and that might account for their worse survival.

Further studies should validate the universality of these biomarkers in different cohorts of patients at different stages of the disease and elucidate the mechanisms behind our findings. Preclinical studies might use these findings to uncover the relationship between HR and DNA damaging therapies in GBM and other solid tumors. We hope that in the near future, these markers help predict patient survival and might help guide

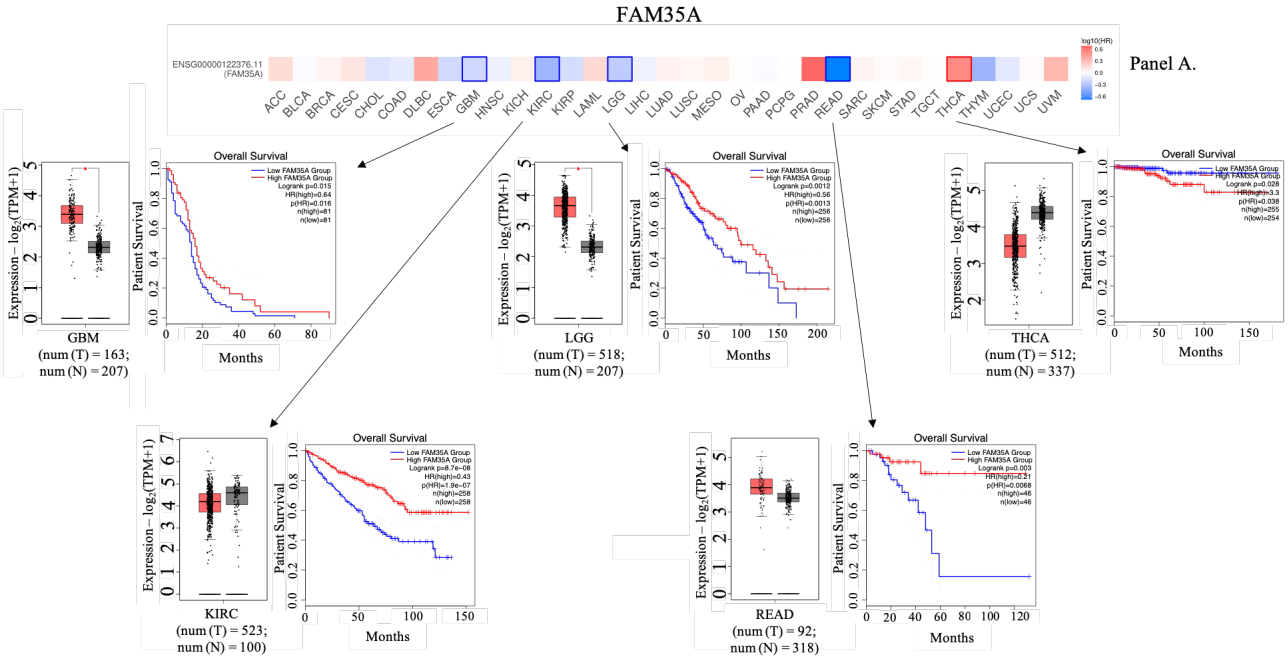


**Fig. 2** Cancer types associated with patient survival in relation to FAM35A. Panel A: Bold squares indicate statistical significance at a significance level of P<0.05. For TCGA abbreviations.

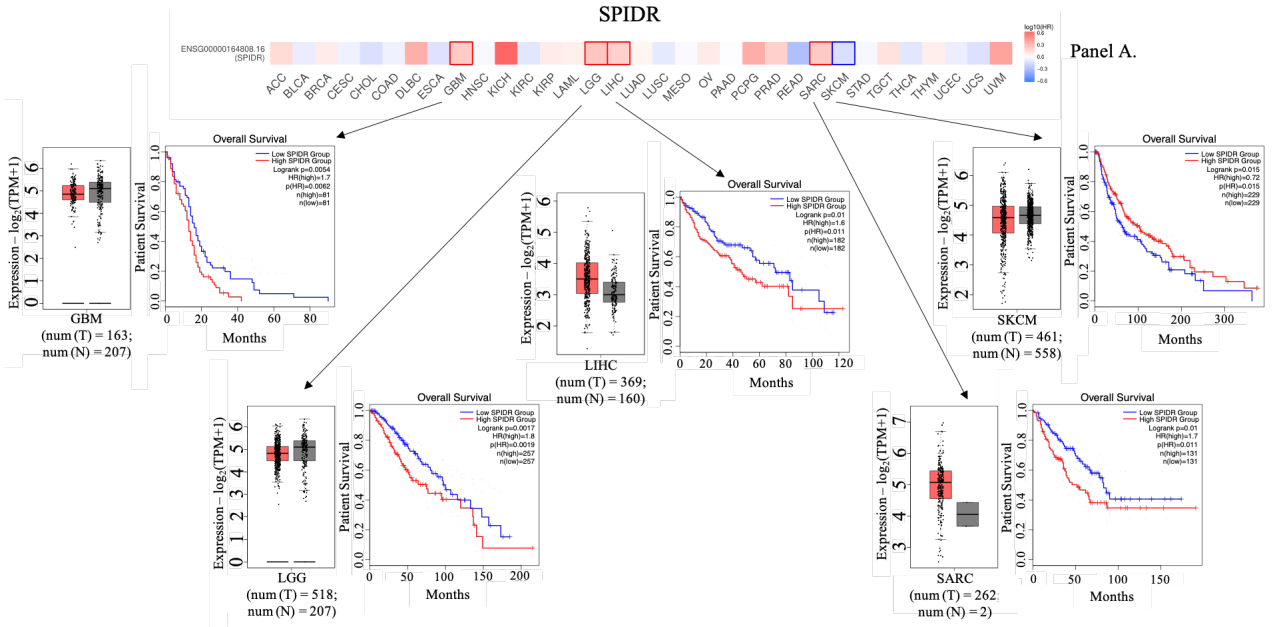


**Fig. 3** Cancer types associated with patient survival in relation to SPIDR. Panel A: Bold squares indicate statistical significance at a significance level of P<0.05. For TCGA abbreviations.

medical decisions on patient follow-up and treatment.

## Methodology

For transcriptomic analysis, we will be using the GEPIA2[15] (Gene Expression Profiling Interactive Analysis) webserver to


© The National High School Journal of Science 2023

identify the patterns of expression of the selected DNA repair genes in GBM. Only the log2 fold change (log2_FC) of those genes with statistically significant (Benjamini-Hochberg) adjusted p-values will be included for this task.

For the survival analysis, webserver The GEPIA2[15] (Gene Expression Profiling Interactive Analysis) webserver will be used to determine the correlation between gene expression and patient survival. GEPIA2[15] is based on The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) projects. The thresholds for high and low gene expression levels between cohorts will be adjusted to 50%. Log-rank P values, hazard ratios (HR), and 95% confidence intervals (CI) will be obtained for each gene.

The gene types in the figures are first listed according to Mdanderson[16] and Haricharan et al[17]. The gene types are in the homologous recombination pathway, and the supplementary information included when determining the figures includes the gene type, the expression of tumor and normal cells in transcript per million (TPM), fold change, log_2 of FC, hazard ratio (HR), and p(HR). All of these genes are analyzed and then chosen to fit the criteria for each figure.

To find out if the genes that are good at predicting GBM survival are also linked to patient survival in other types of cancer; first, the survival rates of the genes are looked at, and then the overall survival rate is plotted on a survival map with months as the unit of measurement. This is done on THE GEPIA2[15] (Gene Expression Profiling Interactive Analysis) webserver. The significance level is set to 0.05 and has no adjustment for the p-value. The group cutoff is at the median, and all the cancer types (TCGA abbreviation) are included in the map.

## Acknowledgements

I'd like to acknowledge my mentor, Luis Bermudez-Guzman, and my writing coach, Tsach Mackey for providing generously with knowledge and expertise. I would also like to thank my program manager, Nandini Pesaladinne, and the Lumiere Research Program for always being available to answer my questions and offer assistance when I needed it.

## References

1 A. Upadhyay, *Genes and Diseases*, **8**, 655–661.

2 D. Debela, S. Muzazu, K. Heraro, M. Ndalama, B. Mesele, D. Haile, S. Kitui and T. Manyazewal, *Sage Open Medicine*, **9**, 205031212110343.

3 E. Holland, *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 6242–6244.

4 P. Yuile, O. Dent, R. Cook, M. Biggs and N. Little, *Journal of Clinical Neuroscience*, **13**, 747–751.

5 D. Krex, B. Klink, C. Hartmann, A. Deimling, T. Pietsch, M. Simon, M. Sabel, J. Steinbach, O. Heese, G. Reifenberger, M. Weller and G. Schackert, *Brain*, **130**, 2596–2606.

6 A. Shergalis, A. Bankhead, U. Luesakul, N. Muangsin and N. Neamati, *Pharmacological Reviews*, **70**, 412–445.

7 Q. Ostrom, H. Gittleman, J. Xu, C. Kromer, Y. Wolinsky, C. Kruchko and J. Barnholtz-Sloan, *Neuro-oncology*, **18**, 1– 75.

8 S. Yalamarty, N. Filipczak, X. Li, A. Subhan, F. Parveen, J. Ataide, B. Rajmalani and V. Torchilin, *Cancers*, **15**, 2116.

9 J. Alhmoud, J. Woolley, A. Moustafa and M. Malki, *Cancers*, **12**, 1050.

10 R. Hakem, *The EMBO Journal*, **27**, 589–605.

11 L. Krejčí, V. Altmannová, M. ŠpıRek and X. Zhao, *Nucleic Acids Research*, **40**, 5795–5818.

12 M. Toh and J. Ngeow, *Oncologist*, **26**, 1526– 1537.

13 T. Knijnenburg, L. Wang, M. Zimmermann, N. Chambwe, G. Gao, A. Cherniack, H. Fan, H. Shen, G. Way, C. Greene, Y. Liu, R. Akbani, B. Feng, L. Donehower, C. Miller, I. Yang, M. Karimi, H. Chen, P. Kim and C. Wang, *Cell Reports*, **23**, 239–254 6.

14 C. Alcazar, P. Todorova, A. Habib, B. Mukherjee and S. Burma, *Molecular Cancer Research*, **14**, 928–940.

15 Z. Tang, B. Kang, C. Li, T. Chen and Z. Zhang, *Nucleic Acids Research*, **47**, 556– 560.

16 *Human DNA repair genes*, `https://www.mdanderson.org/documents/Labs/Wood-Laboratory/human-dna-repair-genes.html`, n.d.).

17 S. Haricharan, M. Bainbridge, P. Scheet and P. Brown, *Breast Cancer Research and Treatment*, **146**, 211–220.

18 M. Lv, W. Wang, M. Zhong, D. Cai, D. Fan, C. Li, W. Kou, Z. Huang, X. Duan, C. Hu, Q. Zhu, X. He and F. Gao, *Frontiers in Genetics*, **13**, year.

19 J. Tomida, K. Takata, S. Bhetawal, M. Person, H. Chao, D. Tang and R. Wood, *The EMBO Journal*, **37**, year.

20 K. Fackrell, L. Bobins, Parul and J. Tomida, *Environmental and Molecular Mutagenesis*, **61**, 709–715.

21 K. Kiwerska and K. Szyfter, *Journal of Applied Genetics*, **60**, 329–334.

22 I. Hazan, T. Hofmann and R. Aqeilan, *PLOS Genetics*, **12**, year.

23 L. Wan, J. Han, T. Liu, S. Dong, F. Xie, H. Chen and J. Huang, *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 10646–10651.

© The National High School Journal of Science 2023