Prediction of homologous recombination deficiency from cancer gene expression data

Jun Kang1, Jieun Lee2, Ahwon Lee1,3 and Youn Soo Lee1

Abstract
Objective: Homologous recombination deficiency (HRD) is the main mechanism of tumorigenesis in some cancers. HRD causes abnormal double-strand break repair, resulting in genomic scars. Some scoring HRD tests have been approved as companion diagnostics of polyadenosine diphosphate-ribose polymerase (PARP) inhibitor treatment. This study aimed to build an HRD prediction model using gene expression data from various cancer types.

Methods: The cancer genome atlas data were used for HRD prediction modeling. A total of 10,567 cases of 33 cancer types were included, and expression data from 5128 out of 20,502 genes were included as predictors. A penalized logistic regression model was chosen as a modeling technique.

Results: The area under the curve of the receiver operating characteristic curve of HRD status prediction was 0.98 for the training set and 0.93 for the test set. The accuracy of HRD status prediction was 0.93 for the training set and 0.88 for the test set.

Conclusions: Our study suggests that the HRD prediction model based on penalized logistic regression using gene expression data can be used to select patients for treatment with PARP inhibitors.

1Department of Hospital Pathology, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea
2Division of Medical Oncology, Department of Internal Medicine, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea
3Cancer Research Institute, The Catholic University of Korea, Seoul, Korea

Corresponding author:
Youn Soo Lee, Department of Hospital Pathology, Seoul St. Mary’s Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-daero, Seocho-gu, Seoul 06591, Republic of Korea.
Email: lys9908@catholic.ac.kr
Introduction

Homologous recombination repair (HRR) is a DNA-damage repair process that restores DNA double-strand breaks without causing mutations. Double-strand breaks invoke chromosome instability and can cause cancers, and impairments in homologous recombination are seen in some cancers. Homologous recombination deficiency (HRD) is caused by a mutation, deletion, or epigenetic silencing in homologous recombination-related genes including BRCA1/2, CHEK1/2, PALB2, and RAD50. HRD is the main mechanism of tumorigenesis in breast, ovarian, and prostate cancers.1,2

HRD causes abnormal double-strand break repair and results in genomic scars, which are aberrant chromosome integrations including large-scale loss of heterogeneity,3 telomere allelic imbalance,4 and large-scale state transitions.5 These three large-scale aberrant genomic alterations can be quantified by counting,6,7 resulting in the HRD score.6 Some genomic scar-based HRD tests such as Myriad myChoice CDx or FoundationOne CDx, which involve HRD score calculations, have been approved as companion diagnostics of polyadenosine diphosphate-ribose polymerase (PARP) inhibitor treatment for patients with ovarian cancer8,9 and prostate cancer.10,11

It should be noted that HRD tests based on genomic scars have limitations.12 As systemic treatment proceeds, various resistance mechanisms are developed and the restoration of HRR by reversion or secondary mutations of HRD-related genes may be possible.13,14 The HRD test cannot reflect HRR restoration because genomic scars make an irreversible footprint and persist even after HRR restoration.12 Therefore, discordance between the HRD status and sensitivity to platinum or PARP inhibitors may be observed.15

There is a consensus that HRD assays are necessary to predict PARP inhibitor efficacy, and that an assessment of dynamic tumor evolution other than HRD is important to guide future treatment.16 In this study, we built an HRD prediction model using gene expression data in various cancer types and applied a previously developed machine learning-based model to predict PARP inhibitor response using gene expression17 to assess the efficacy of PARP inhibitors in solid tumors.

Materials and Methods

Dataset

The cancer genome atlas (TCGA) data were used for HRD prediction modeling. HRD scores of various TCGA cancer types have previously been published,18 and were downloaded. The three components of the HRD score are: large (>15 Mb) non-arm-level regions with loss of heterozygosity (LOH), large-scale state transitions (breaks between adjacent segments of >10 Mb), and subtelomeric regions with allelic imbalance. Genome-wide segmentation data for allele-specific copy number and LOH, and tumor
purity were estimated using the ABSOLUTE computational method.\textsuperscript{19}

Normalized gene expression data of the pan-cancer TCGA dataset were estimated from mRNA expression data that were generated using RNAseq and normalized. TCGA gene expression data were downloaded from the National Cancer Institute (NCI) Genomic Data Commons (GDC) website.\textsuperscript{20} Ethical approval was not required because public databases were used according to TCGA publication guidelines.\textsuperscript{21}

**Prediction modeling**

We constructed an HRD status prediction model using TCGA gene expression data. The target variable was the HRD status, and predictor variables were cancer type and gene expression. Genes with a large median absolute deviation value (greater than third-quartiles) were selected to reduce computational load. The HRD status was defined as positive if the HRD score was \( \geq 42 \) and negative if it was \(<42.6\). A penalized logistic regression model was selected as the modeling technique.

Three-quarters of the dataset was designated as the training set and the remainder as the test set according to HRD status stratification. Data preprocessing was performed on the training set using a Yeo–Johnson transformation to resolve skewness, followed by normalization, standardization, and oversampling to resolve the case imbalance with the Synthetic Minority Over-sampling Technique (SMOTE).\textsuperscript{22} Data preprocessing was performed using the recipes R package\textsuperscript{23} and themis R package.\textsuperscript{24}

Model selection was performed with a 10-fold validation on a grid of hyperparameters: \( \lambda \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}\} \), \( \alpha \{0.0, 0.25, 0.5, 0.75, 1.0\} \). Ten-fold validation was stratified with the combined HRD status and cancer type to evenly divide the assessment sets according to HRD status and cancer type.

**Assessing model performance**

The performance of the prediction was estimated using an area under the curve of the receiver operating characteristic curve (AUROC) for HRD status. The modeling process and assessment of model performance were performed using tidymodels\textsuperscript{25} and glmnet rpackages.\textsuperscript{26}

**Important predictor selection and gene ontology analysis**

Important predictors were selected by coefficients. Genes were selected with a large absolute coefficient value (\( >0.02 \)). A gene ontology test was undertaken using PANTHER\textsuperscript{27} to determine which biological processes are important in HRD status prediction. The following detailed parameters were used: analysis type: PANTHER overrepresentation test (released 2022-07-12); annotation version and release date: PANTHER version 17.0 released 2022-02-22; test type: Fisher’s exact; correction: false discovery rate.

**Results**

**Dataset summary**

A total of 10,567 cases of 24 cancer types were included, ranging in number from 45 (cholangiocarcinoma) to 1171 (breast invasive carcinoma). Cancer type abbreviations and the number of cases are listed in the S1 Appendix. Gene expression of 5128 of the 20,502 genes was included as predictors. The number of HRD status-positive cases was 1169 (11.1%). Ovarian cancer showed the highest HRD-positive rate (0.56).

**Selecting model and performance estimation**

The model with \( \lambda \) (penalty) = 0.1 and \( \alpha \) (mixture) = 0 (ridge regression) showed the best performance in terms of AUROC in
10-fold-validation (Figure 1). We fitted the final model with the selected hyperparameters and the entire training set. The AUROC of HRD status prediction was 0.98 for the training set and 0.93 for the test set. The accuracy of HRD status prediction was 0.93 for the training set and 0.88 for the test set.

**Performance by cancer type**

All cancer types showed good prediction in terms of AUROC and accuracy in the training set. In training set prediction, AUROC was above 0.92 for every cancer type and the accuracy was above 0.72 (Table 1 and Figure 2). Most cancer types maintained good performance in the test set prediction, except for uterine carcinosarcoma, lung squamous cell carcinoma, and esophageal carcinoma. In the test set prediction, AUROC was above 0.76 and the accuracy was above 0.74 in every cancer type except uterine carcinosarcoma and pancreatic adenocarcinoma (Table 1 and Figure 3). The cancer types with a low HRD positive rate (<0.10) showed very good accuracy (>0.90) in the test set. Table 1 summarizes the HRD positive rate, accuracy, and AUROC of the training set and test set.

![Figure 1](image_url). Hyperparameter tuning and performance assessment in 10-fold cross-validation resampling. The x-axis is a penalty scaling parameter: $\lambda \{ 10^{-3}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 10^0 \}$, color is mixture hyperparameter of penalty function: $\alpha \{ 0.0, 0.25, 0.5, 0.75, 1.0 \}$. The y-axis is a mean estimate of the area under the curve of the receiver operating characteristic curve (AUROC) using 10-fold cross-validation resampling.
Table 1. Accuracy and area under the curve of the receiver operating characteristic curve for each cancer type.

| Cancer type | HRD positive rate | Training set | Test set |
|-------------|------------------|--------------|----------|
|             |                  | Accuracy | AUROC | Accuracy | AUROC |
| OV          | 0.56             | 0.75     | 0.97  | 0.65     | 0.87  |
| UCS         | 0.34             | 0.86     | 1.00  | 0.46     | 0.56  |
| LUSC        | 0.32             | 0.74     | 0.92  | 0.55     | 0.71  |
| SARC        | 0.30             | 0.84     | 0.96  | 0.76     | 0.87  |
| ESCA        | 0.29             | 0.72     | 0.93  | 0.71     | 0.81  |
| STAD        | 0.23             | 0.85     | 0.96  | 0.79     | 0.91  |
| BLCA        | 0.21             | 0.79     | 0.94  | 0.76     | 0.82  |
| BRCA        | 0.19             | 0.88     | 0.97  | 0.82     | 0.94  |
| LUAD        | 0.14             | 0.89     | 0.97  | 0.82     | 0.84  |
| HNSC        | 0.11             | 0.92     | 0.97  | 0.79     | 0.82  |
| UCEC        | 0.10             | 0.94     | 0.99  | 0.91     | 0.96  |
| CESC        | 0.06             | 0.94     | 0.99  | 0.99     | 0.99  |
| SKCM        | 0.04             | 0.98     | 0.99  | 0.95     | NA    |
| PAAD        | 0.03             | 0.98     | 1.00  | 0.98     | 0.53  |
| LIHC        | 0.03             | 0.97     | 0.98  | 0.93     | 0.72  |
| MESO        | 0.02             | 1.00     | 1.00  | 0.90     | 0.95  |
| CHOL        | 0.02             | 1.00     | 1.00  | 1.00     | NA    |
| DLBC        | 0.02             | 0.97     | 1.00  | 1.00     | NA    |
| COAD        | 0.01             | 0.99     | 0.98  | 0.98     | 0.98  |
| ACC         | 0.01             | 1.00     | 1.00  | 1.00     | NA    |
| READ        | 0.01             | 1.00     | NA    | 0.95     | 0.92  |
| PRAD        | 0.01             | 1.00     | 1.00  | 1.00     | NA    |
| KIRC        | 0.00             | 1.00     | 1.00  | 1.00     | NA    |
| KIRP        | 0.00             | 1.00     | 1.00  | 1.00     | NA    |
| LGG         | 0.00             | 1.00     | NA    | 0.99     | 0.91  |
| GBM         | 0.00             | 1.00     | NA    | 1.00     | NA    |
| KICH        | 0.00             | 0.98     | NA    | 1.00     | NA    |
| LAML        | 0.00             | 1.00     | NA    | 1.00     | NA    |
| PCPG        | 0.00             | 1.00     | NA    | 1.00     | NA    |
| TGCT        | 0.00             | 1.00     | NA    | 1.00     | NA    |
| THCA        | 0.00             | 1.00     | NA    | 1.00     | NA    |
| THYM        | 0.00             | 1.00     | NA    | 1.00     | NA    |
| UVM         | 0.00             | 1.00     | NA    | 1.00     | NA    |

Note: Study abbreviations are listed in S1 Appendix.
AUROC, area under the curve of the receiver operating characteristic curve.

**Important predictors and gene ontology test**

Genes with large absolute coefficient values can be considered useful in predicting the HRD status. These genes with positive coefficients are described in the S2 Appendix, and those with negative coefficients are described in the S3 Appendix. Genes with large positive coefficients were overrepresented in actin cytoskeleton organization, protein-containing complex subunit organization, the cellular metabolic process, and the organic substance metabolic process. Genes with large negative coefficients were
Figure 2. Receiver operating characteristic (ROC) curve of homologous recombination deficiency status prediction of the training set.
overrepresented in negative regulation of the apoptotic process and the cellular protein metabolic process. Gene ontology analyses of genes with large negative and positive coefficients are described in S4 Appendix.

Figure 3. Receiver operating characteristic (ROC) curve of homologous recombination deficiency status prediction of the test set.
Discussion

An accumulation of genomic alterations or aberrations based on HRD ultimately leads to the carcinogenesis of solid cancers such as ovary, breast, and prostate cancer. PARP inhibitors such as olaparib and talazoparib inhibit and trap poly(ADP-ribose) polymerase 1, which repairs single-strand DNA breaks (SSB). The prevention of SSB restoration by PARP inhibitors results in double-strand breaks and eventually leads to cellular death in \( BRCA1/2 \) mutant or HRD cancers. Based on this mechanism, PARP inhibitors have shown clinical efficacy in \( BRCA1/2 \) mutant ovarian, breast, and prostate cancer, and also in HRD-positive ovarian and prostate cancer.

The tumor environment undergoes dynamic evolution as treatment proceeds, so an assessment of the overall tumor biology is warranted. Evaluating genomic characteristics based on gene expression profile can overcome the limitation of not reflecting a change in HRD because the gene expression profile represents the current functional cellular state. Although genomic scars do not disappear once they are created, the HRD gene expression signature changes to the HRR signature if HRR restoration occurs.\(^{28}\)

Our penalized logistic regression model predicted the HRD status with gene expression data in various cancer types with a 0.93 test set AUROC. This prediction performance is similar to the HRD status prediction previously achieved using deleterious \( BRCA1/2 \),\(^{29}\) suggesting that an HRD prediction model based on penalized logistic regression using gene expression data can be applied to patient selection for treatment with PARP inhibitors. This model is preliminary and should be validated in a clinic with prospectively collected cancer tissues, but this study confirms the possibility of HRD prediction using gene expression data. This offers the advantage of evaluating the current HRD status of cancer using serially-collected tumor tissues when disease shows resistance to previous treatment; in contrast, HRD tests based on genomic scars do not reflect dynamic changes in HRD status.

Our prediction model showed good performance in cancer types associated with high HRD-expressing malignancies including ovarian cancer and breast cancer, but poorly predicted other cancer types, including uterine carcinosarcoma and pancreatic adenocarcinoma. This poor prediction performance might reflect underfitting by a small dataset, so a large dataset showing poor prediction may be needed for these cancer types.

Some gene expression-based HRD prediction tests have previously been studied.\(^{28,30-32}\) However, the HRD status in these studies was determined by the presence of pathogenic variants in HRD-related genes such as \( BRCA1/2 \). Conversely, our prediction model used an HRD score to determine the HRD status. An HRD test based on genomic scars has the advantage of detecting more patients who can be treated with PARP inhibitors.\(^{9}\)

The primary purpose of this study was to determine whether gene expression data can predict HRD status. However, by investigating important predictors of our model, we also found that genes with large negative coefficients were overrepresented in negative regulation of the apoptotic process. These findings are consistent with previous studies.\(^{33}\) Apoptosis is suppressed in the response to DNA damage.\(^{34}\)

Our prediction model has some advantages and limitations. The main advantage of HRD prediction with gene expression is that it reflects the current HRD status. To understand how well this predicts the current HRD status, our model should be tested with gene expression profile data collected from patients who showed resistance to PARP inhibitors or platinum. The
restoration of HRR is a major mechanism of such resistance, and gene expression datasets from patients showing resistance may reflect current changes in HRD. Additionally, our prediction model was limited with respect to generalization. The dataset used for the modeling process only included 24 common cancer types, so the performance of other cancer types is unknown. Moreover, TCGA RNAseq normalization is cancer set-dependent. Therefore, the standardization of gene expression data processing should be resolved for the practical use of this HRD prediction model using gene expression data. External validation is also needed for practical use.

In summary, our study shows that the HRD prediction model based on penalized logistic regression using gene expression data can be applied to patient selection for treatment with PARP inhibitors. However, gene expression data standardization and validation in cancer with HRD restoration are required to apply these models in clinical testing.

Data availability
RNAseq data used in this study are publicly available at: https://api.gdc.cancer.gov/data/3586c0da-64d0-4b74-a449-5ff4d9136611. HRD data are available at: https://api.gdc.cancer.gov/data/66dd07d7-6366-4774-83c3-5ad1e22b177e. The code used for data processing is also available in this repository.

Declaration of conflicting interests
The Catholic University of Korea, Industry-Academic Cooperation Foundation has been filed a patent for “Modeling method for homologous recombination deficiency predictive model” (Application No. 10-2022-0014718). Jun Kang, Jieun Lee, Ahwon Lee, and Youn Soo Lee are listed as inventors of the patent.

Funding
The authors disclosed receipt (pending publication) of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1I1A1A01043754).

ORCID iD
Jun Kang https://orcid.org/0000-0002-7967-0917

Supplemental material
Supplemental material for this article is available online.

References
1. Nguyen L, Martens J, Van Hoeck A, et al. Pan-cancer landscape of homologous recombination deficiency. Nat Commun 2020; 11: 5584.
2. Von Werdt A, Brandt L, Schärer OD, et al. PARP inhibition in prostate cancer with homologous recombination repair alterations. JCO Precis Oncol 2021: 1639–1649. DOI: 10.1200/PO.21.00152.
3. Abkevich V, Timms KM, Hennessy BT, et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br J Cancer 2012; 107: 1776–1782. DOI: 10.1038/bjc.2012.451.
4. Birkbak NJ, Wang ZC, Kim JY, et al. Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. Cancer Discov 2012; 2: 366–375. DOI: 10.1158/2159-8290.CD-11-0206.
5. Popova T, Manié E, Rieunier G, et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. Cancer Res 2012; 72: 5454–5462. DOI: 10.1158/0008-5472.CAN-12-1470.
6. Telli ML, Timms KM, Reid J, et al. Homologous recombination deficiency (HRD) Score predicts response to platinum-containing neoadjuvant...
7. Watkins JA, Irshad S, Grigoriadis A, et al. Genomic scars as biomarkers of homologous recombination deficiency and drug response in breast and ovarian cancers. *Breast Cancer Res* 2014; 16: 211. DOI: 10.1186/bcr3670.

8. Moore K, Colombo N, Scambia G, et al. Maintenance olaparib in patients with newly diagnosed advanced ovarian cancer. *N Engl J Med* 2018; 379: 2495–2505. DOI: 10.1056/NEJMoa1810858.

9. Ray-Coquard I, Pautier P, Pignata S, et al. Olaparib plus bevacizumab as first-line maintenance in ovarian cancer. *N Engl J Med* 2019; 381: 2416–2428. DOI: 10.1056/NEJMoa1911361.

10. Hussain M, Mateo J, Fizazi K, et al. Survival with olaparib in metastatic castration-resistant prostate cancer. *N Engl J Med* 2020; 383: 2345–2357. DOI: 10.1056/NEJMoa2022485.

11. De Bono J, Mateo J, Fizazi K, et al. Olaparib for metastatic castration-resistant prostate cancer. *N Engl J Med* 2020; 382: 2091–2102. DOI: 10.1056/NEJMoa1911440.

12. Hoppe MM, Sundar R, Tan DSP, et al. Biomarkers for homologous recombination deficiency in cancer. *J Natl Cancer Inst* 2018; 110: 704–713. DOI: 10.1093/jnci/djy085.

13. Edwards SL, Brough R, Lord CJ, et al. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* 2008; 451: 1111–1115. DOI: 10.1038/nature06548.

14. Pishvaian MJ, Biankin AV, Bailey P, et al. BRCA2 secondary mutation-mediated resistance to platinum and PARP inhibitor-based therapy in pancreatic cancer. *Br J Cancer* 2017; 116: 1021–1026. DOI: 10.1038/bjc.2017.40.

15. Ngoi NYL and Tan DSP. The role of homologous recombination deficiency testing in ovarian cancer and its clinical implications: do we need it? *ESMO Open* 2021; 6: 100144. DOI: 10.1016/j.esmoop.2021.100144.

16. Miller RE, Leary A, Scott CL, et al. *ESMO* recommendations on predictive biomarker testing for homologous recombination deficiency and PARP inhibitor benefit in ovarian cancer. *Ann Oncol* 2020; 31: 1606–1622. DOI: 10.1016/j.annonc.2020.08.2102.

17. Hill KE, Rattani A, Lietz CE, et al. Machine learning-based prediction of response to PARP inhibition across cancer types. *medRxiv* 2019; 19007757.

18. Thorsson V, Gibbs DL, Brown SD, et al. The immune landscape of cancer. *Immunity* 2018; 48: 812–830.e814. DOI: 10.1016/j.immuni.2018.03.023.

19. Carter SL, Cibulsksis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol* 2012; 30: 413–421. DOI: 10.1038/nbt.2203.

20. Grossman RL, Heath AP, Ferretti V, et al. Toward a shared vision for cancer genomic data. *New Eng J Med* 2016; 375: 1109–1112.

21. National Cancer Institute. The Cancer Genome Atlas Program, https://cancergenome.nih.gov/publications/guidelines (accessed 28 October 2022).

22. Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* 2002; 16: 321–357. DOI: 10.1613/jair.953.

23. Kuhn M and Wickham H. recipes: Preprocessing and Feature Engineering Steps for Modeling. https://github.com/tidymodels/recipes, https://recipes.tidymodels.org/, 2022.

24. Hvitfeldt E. themis: Extra Recipes Steps for Dealing with Unbalanced Data. https://github.com/tidymodels/themis, https://themis.tidymodels.org/, 2022.

25. Kuhn M and Wickham H. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. https://www.tidymodels.org, 2020.

26. Friedman J, Hastie T and Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010; 33: 1–22.

27. Mi H, Muruganujan A and Thomas PD. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* 2013; 41: D377–D386. DOI: 10.1093/nar/gks1118.

28. Peng G, Chun-Jen Lin C, Mo W, et al. Genome-wide transcriptome profiling of
29. Rempel E, Kluck K, Beck S, et al. Pan-cancer analysis of genomic scar patterns caused by homologous repair deficiency (HRD). *NPJ Precis Oncol* 2022; 6: 36. 2022/06/11. DOI: 10.1038/s41698-022-00276-6.

30. Jazaeri AA, Yee CJ, Sotiriou C, et al. Gene expression profiles of BRCA1-linked, BRCA2-linked, and sporadic ovarian cancers. *J Natl Cancer Inst* 2002; 94: 990–1000. DOI: 10.1093/jnci/94.13.990.

31. Konstantinopoulos PA, Spentzos D, Karlan BY, et al. Gene expression profile of BRCAAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. *J Clin Oncol* 2010; 28: 3555–3561. DOI: 10.1200/JCO.2009.27.5719.

32. Leibowitz BD, Dougherty BV, Bell JSK, et al. Validation of genomic and transcriptomic models of homologous recombination deficiency in a real-world pan-cancer cohort. *BMC Cancer* 2022; 22: 587. 2022/06/02. DOI: 10.1186/s12885-022-09669-z.

33. Yao R and Cooper GM. Requirement for phosphatidylinositol-3 kinase in the prevention of apoptosis by nerve growth factor. *Science* 1995; 267: 2003–2006. DOI: 10.1126/science.7701324.

34. Aubrey BJ, Kelly GL, Janic A, et al. How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death Differ* 2018; 25: 104–113. DOI: 10.1038/cdd.2017.169.