bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Loss of HMCES is synthetic lethal with APOBEC activity in cancer cells

Josep Biayna[1], Isabel Garcia-Cao[2], Miguel M. Álvarez[1], Marina Salvadores[1], Jose Espinosa-Carrasco[1], Marcel McCullough[1], Fran Supek[1,3,*] and Travis H. Stracker[2,4*]

[1]Genome Data Science, Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, C/ Baldiri Reixac 10, Barcelona 08028, Spain.
[2]Genomic Instability and Cancer, Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, C/ Baldiri Reixac 10, Barcelona 08028, Spain.
[3]Catalan Institution for Advanced Studies (ICREA), Passeig de Lluís Companys, 23, 08010 Barcelona, Spain.
[4]National Cancer Institute, Center for Cancer Research, Radiation Oncology Branch, 10 Center Dr, Building 10, Bethesda, MD 20814, USA

***Lead contacts:** fran.supek@irbbarcelona.org, travis.stracker@nih.gov

Running title: HMCES is synthetic lethal with APOBEC expression

Word count: abstract 199, introduction/results/discussion, figure legends, materials and methods 6101 words.

Key words: APOBEC, HMCES, synthetic lethality, lung cancer, CRISPR/Cas9

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Analysis of cancer mutagenic signatures provides information about the origin of mutations and can inform the use of clinical therapies, including immunotherapy. In particular, APOBEC3A (A3A) has emerged as a major driver of mutagenesis in cancer cells and its expression results in DNA damage and susceptibility to treatment with inhibitors of the ATR and CHK1 checkpoint kinases. Here we report the implementation of CRISPR/Cas9 genetic screening to identify susceptibilities of multiple A3A-expressing lung adenocarcinoma cell lines. We identify HMCES, a protein recently linked to the protection of abasic sites, as a central protein for the tolerance of A3A expression. HMCES depletion results in synthetic lethality with A3A expression specifically in a TP53-mutant background. Analysis of previous screening data reveals a strong association between A3A mutational signatures and sensitivity to HMCES loss and indicates that HMCES is specialized in protecting against a narrow spectrum of DNA damaging agents in addition to A3A. We experimentally show that both HMCES disruption and A3A expression increase susceptibility of cancer cells to ionizing radiation, oxidative stress and ATR inhibition; strategies that are often applied in tumor therapies. Overall, our results suggest that HMCES is an attractive target for selective treatment of A3A expressing tumors.

Introduction

The APOBEC3 (apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3) family of cytidine deaminases is a major source of mutagenesis in human cancers. Elevated mRNA levels of APOBEC3A (A3A) and APOBEC3B (A3B) enzymes, as

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

well as an activating germline polymorphism in the *A3A* and *A3B* genes, were associated with a particular mutational signature of C-to-T and C-to-G changes in a TCW trinucleotide context (where W is A or T)[1–4]. Both A3A and A3B have been implicated in localized hypermutation, which can occur in two different patterns: the focused *kataegis* ('mutation showers', likely occurring during repair of DNA double-strand (ds) breaks (DSB)[1,5]) and the diffuse *omikli* pattern ('mutation fog', proposed to occur during repair of mismatched or damaged nucleotides[6,7]. The A3s are a cause of intratumor genetic heterogeneity and generate driver mutations in tumors[7–10]. Consistently, A3 mutagenesis has prognostic value in cancers[1,11–13]. Recent genomics work suggests that A3 mutagenesis appears rare in various types of apparently non-cancerous somatic cells[14], and moreover A3 mutagenesis appears to increase in intensity in metastatic cancers[15]. This suggests that vulnerabilities of APOBEC-expressing cells would provide a window of opportunity to selectively target certain types of tumor cells while sparing their healthy counterparts.

Overexpressing A3 enzymes in yeast and human cell lines results in clustered mutation patterns resembling those seen in cancer genomes[13,16,17]. Therefore, such experimental models of A3 overexpression appear useful for recapitulating DNA damaging and mutagenic effects that occur in tumors due to APOBEC activity. The A3A mutagenesis signature is distinguishable from that of A3B and both signatures are present in varying proportions across cancer types. However, the A3A signature is predominant overall[7,18–21] consistent with experiments suggesting

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

that A3A induces high levels of DNA damage[2,22]. We therefore focused our attention on A3A.

A3s deaminate cytosine in DNA to generate uracil, which can be converted to an abasic (AP) site, following the action of uracil glycosylases[23]. Uracil is mutagenic, causing U:G mispairing during copying. Moreover, AP-sites cannot be directly copied by the replicative DNA polymerases during S-phase, necessitating the use of potentially mutagenic translesion synthesis (TLS) polymerases[24]. A3A induced damage occurs during S-phase and AP-sites can lead to replication fork stalling and replication stress[25–27]. Processing of AP-sites by AP-endonucleases can allow repair by the base excision repair (BER) pathway. This can promote further A3 mutagenesis, particularly if coupled with the activity of DNA mismatch repair that can 'hijack' BER intermediates[6]. Alternatively, the processing of AP-sites in ssDNA can convert them to DNA double-strand breaks (DSBs), a more cytotoxic lesion. Processing and repair of DSBs by the homologous recombination (HR) or break-induced replication (BIR) pathways generates additional ssDNA which may be targeted by APOBECs[5,28]. Thus, multiple DNA repair pathways are engaged as a consequence of A3-induced DNA damage, and activity of these pathways can promote further A3 DNA damage.

Increased reliance of some tumors on particular DNA repair pathways has long been exploited as a therapeutic avenue. For example, brain cancers that lose activity of the O-6-methylguanine-DNA methyltransferase (MGMT) enzyme, that can directly reverse O-6 adducts, are more sensitive to the DNA methylating drug temozolomide

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(TMZ)[29]. Ovarian and breast tumors with failures in HR repair pathways due to inactivated BRCA1 and BRCA2 genes are more sensitive to PARP inhibitors, such as Olaparib[30,31]. These examples of successful therapeutic applications encouraged us to search for targetable DNA repair pathways in cancer cells exposed to increased A3A activity.

Overexpression of A3A causes DNA damage and replication stress; the latter can be targeted by inhibitors of the ATR and CHK1 checkpoint kinases that respond to replication stress[22,32,33]. The observation that cell cycle checkpoint inhibitors enhanced the levels of DNA damage in A3A-expressing cells indicates that it is plausible that they have many additional inherent vulnerabilities that can be therapeutically exploited, apart from the replication stress response, which is a more general phenomenon not specific to A3A.

We performed a CRISPR/Cas9-based genome-wide screen for genes required to tolerate A3A-mediated DNA damage in a panel of cell lines from non-small cell lung cancer (NSCLC), where APOBEC activity has been shown to play an important role in tumor evolution[3,4,34–36]. Among other hits, we identified factors involved in multiple DSB repair pathways, including RAD9A, a component of the 9-1-1 alternative clamp loader and the recently characterized MCM8-MCM9-HROB complex[37–40]. Crucially, we found that different genetic backgrounds are consistently and strongly dependent on the gene encoding HMCES (5hmC binding, embryonic stem cell-specific-protein) for cell viability[8,15] under APOBEC stress, but not otherwise. Recently, HMCES (also known as SRAP Domain-containing

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Protein 1), as well as the related bacterial protein YedK, were shown to covalently bind to AP-sites in ssDNA, where they act as a suicide enzyme to protect them from TLS or AP-endonucleases[41–46]. In addition, HMCES has been proposed to function in the repair of DSBs in the canonical and the alternative non-homologous end-joining (NHEJ) pathways[42,47,48]. We validated HMCES depletion as a sensitizer to A3A in multiple cell lines, consistent with recent work[49], and show that HMCES limits DNA damage and prevents loss of cell viability resulting from *A3A* expression in a manner specific to TP53-mutant cells. Together, our results identify additional druggable targets to be considered in A3A expressing cancer cells and establish a central role for HMCES in preventing the toxicity of A3A expression.

Results

Generation of non-small cell lung cancer cell lines with inducible A3A expression

To examine the influence of A3A expression in NSCLC, we established a panel of cell lines with doxycycline (DOX) inducible expression of a haemagglutinin (HA)-tagged-A3A using the pSLIK-Neo vector system (Fig 1A-1B)[27]. This included NCI-H358, LXF-289, A549 and a TP53 null variant of A549, A549$^{TP53-/-}$, generated using CRISPR/Cas9 targeting (S1 Fig). Treatment of cells with DOX resulted in a dose-dependent increase in A3A mRNA expression and protein levels (Fig 1A-1B). Consistent with previous reports that A3A expression caused DNA damage and cell cycle checkpoint activation, we observed a slower growth rate in several of the NSCLC cell lines (Fig 1C)[27].

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

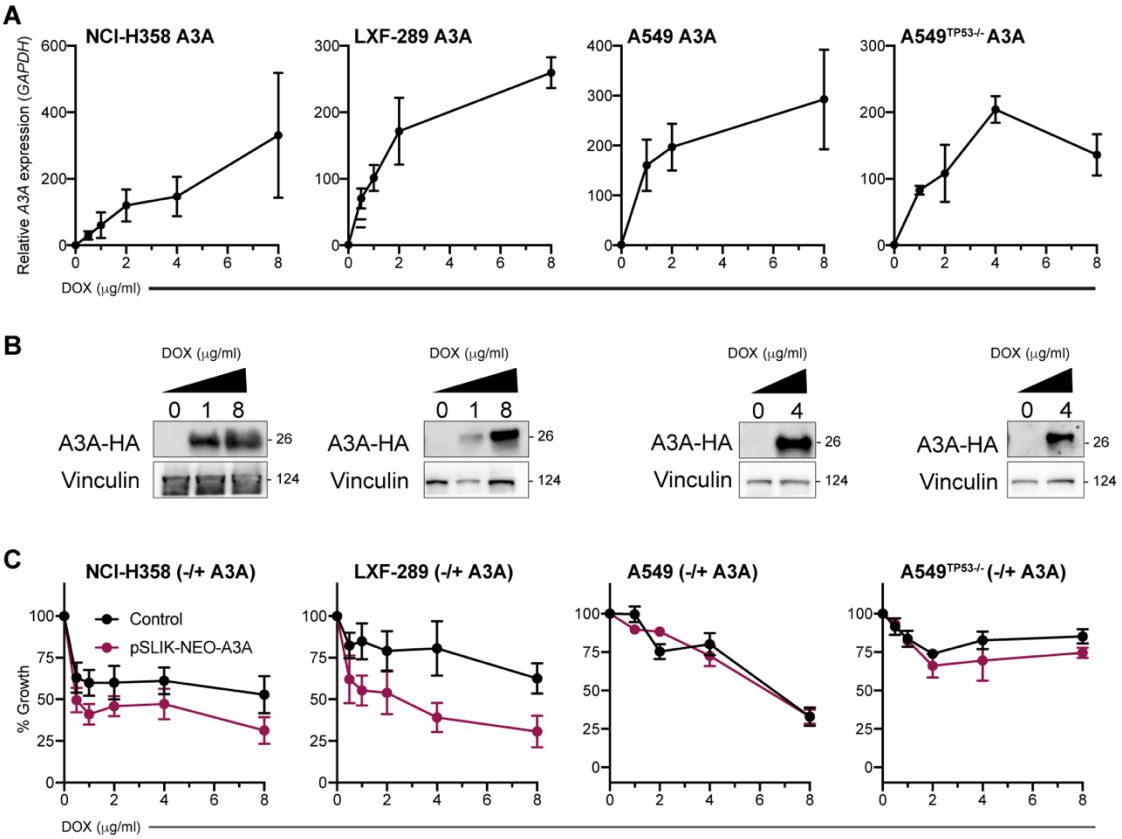


Fig 1. Inducible A3A expression reduces the fitness of lung adenocarcinoma cell lines. (**A**) A3A mRNA levels in NCI-H358, LXF-289, A549 and A549$^{TP53-/-}$ cell lines transduced with a doxycycline (DOX) inducible A3A cassette at various concentrations of DOX. Quantitative real-time PCR (qRT-PCR) and western blot analysis were repeated two times. (**B**) Western blot detection of HA-A3A upon DOX induction. LXF-289, NCI-H358, A549 and A549$^{TP53-/-}$ cells were collected and lysed 72h post-treatment. Vinculin serves as a loading control and molecular weight (MW) is indicated in kilodaltons. (**C**) Growth (percentage of growth rate relative to the cells without DOX) for the indicated cell lines after 72h of DOX treatment measured with alamarBlue. In red, cells transduced with the inducible A3A cassette and in black,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

the parental cell line (no A3A) exposed to the same concentration of DOX. Growth assays were repeated three times. For all graphs, mean and SEM are shown.

CRISPR/Cas9 genetic screen for A3A synthetic lethality

In order to identify vulnerabilities of A3A expressing cells, we performed genome-wide CRISPR/Cas9 screening in three lung adenocarcinoma cell lines: LXF-289, A549 and A549$^{TP53-/-}$. Screening was performed at an established IC_{25} dose of DOX for A549 and A549$^{TP53-/-}$ and IC_{25} and IC_{50} doses for LXF-289 (Fig 2A). The cell lines, with or without pSLIK-Neo A3A, were transduced with a single-vector lentiviral library expressing Cas9 (Brunello) and guide RNAs (gRNAs) for 19,114 genes (4 single gRNAs (sgRNAs) per gene) at a multiplicity of infection (MOI) ≤ 0.4 (Fig. 2A)[50]. Following puromycin selection, cells were lysed, genomic DNA extracted and preparation and analysis of initial gRNA representation (T0) was performed. Cells were subsequently expanded for 15 days and treated with DOX to induce A3A expression. At multiple time points following DOX treatment, we collected cells and amplified guide DNAs (gDNAs) using barcoded primers. DNA was sequenced and analyzed for changes in abundance of gDNAs targeting various genes, comparing the DOX-treated cells with the untreated (control) cell line at the same time point using the MAGeCK-RRA tool[51], thus revealing genes which have stronger fitness effects in A3A expressing cells. Additionally, we compared to T0 to determine overall essential genes. The control experiment showed that DOX itself (in a genetic background lacking the A3A plasmid) affected the essentiality of very few genes (S2 Fig).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

As TP53 status has been shown to influence CRISPR/Cas9 screening results, we first compared A549$^{TP53-/-}$ with the LXF-289 cell line that bears a TP53 mutation (c.742C>T; p.R248W per DepMap.org record ACH-000787) (Fig 2B)[52]. We prioritized genes by an overall APOBEC essentiality score: average $\log_2$ fold-change (LFC) over six measurements: three time points for the A549$^{TP53-/-}$ cell line (T9, T12 and T15) and three for the LXF-289 cell line (T5, T10 and T15). The top five hits by this score were the genes coding for the AP-site protecting protein HMCES[42], the RAD9A cell cycle checkpoint control protein, the MCM8 component of the MCM8-MCM9-HROB complex[37,38,53], ATXN7L3, a component of the SAGA chromatin modifying complex[54,55], and HGC6.3, an uncharacterized protein. For four of the five genes, the individual gRNAs, four per gene, consistently sensitized cells to A3A expression (Fig 2B and S3 Fig). However, HGC6.3 guides displayed an inconsistent temporal trend and the effect size for HGC6.3 was very different across the two cell lines (S3 Fig). An additional analysis by the MAGeCK-MLE method[51] (S4 Fig) suggested that all four gRNAs for *HGC6.3* had low knockout efficiency (all <=0.72; (S1 Table)) in contrast to other top hits, and we thus disregarded HGC6.3 in further analysis. The remaining four top hits did not show clear differences in effects between the two tested A3A dosages in the LXF-289 cell line (corresponding to IC_{25} and IC_{50}) (Panel D in S3 Fig). Next highest-ranking hits included the UBA6 ubiquitin activating enzyme, and a further five genes that were all related to DNA repair, DNA replication or cell cycle control (*DDX11, MCM9, CDC23, MAD2L2* (also known as REV7), and *HROB* (also known as C17orf53 or MCM8IP; S1 Table).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

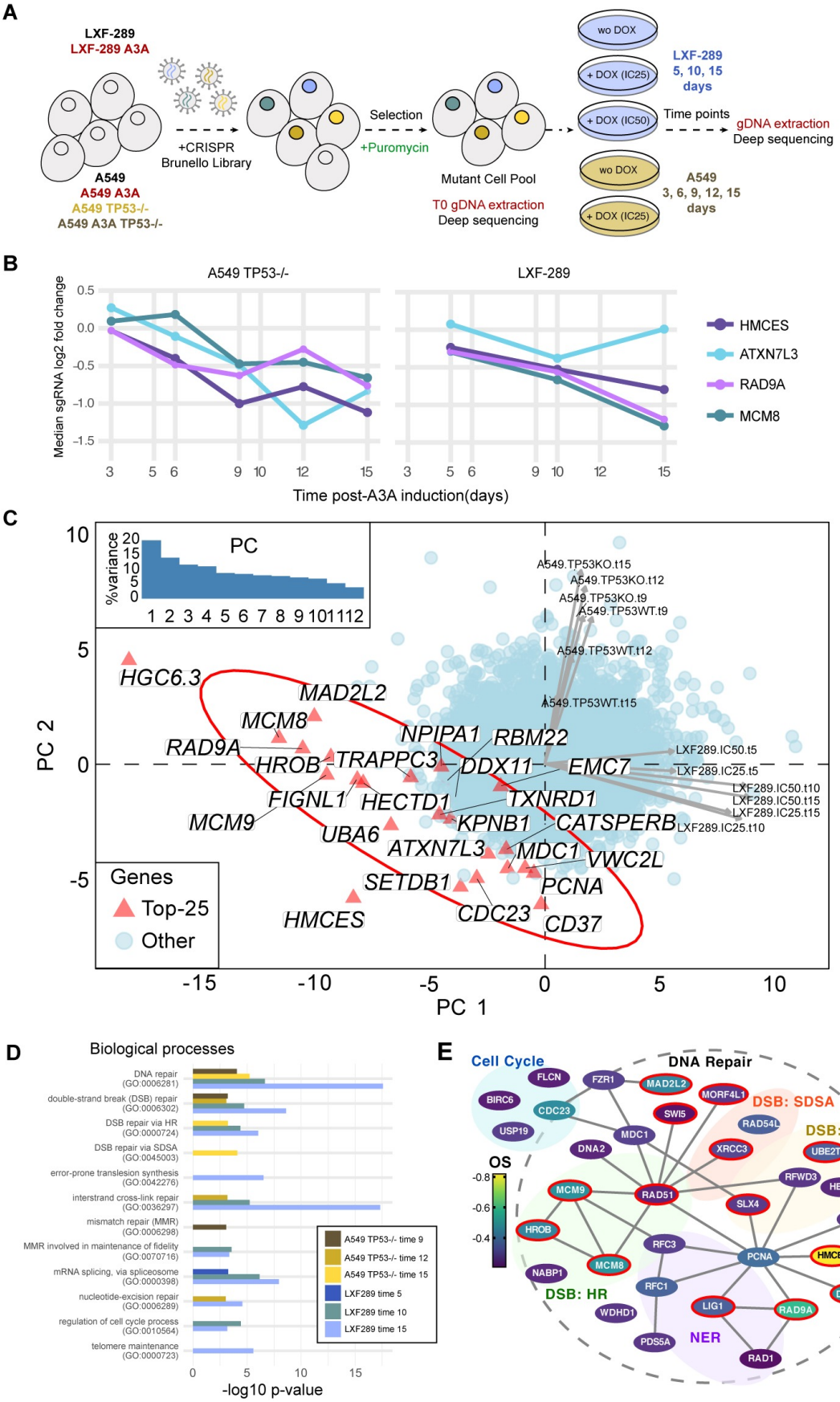
213 **Distinct genes but consistent pathways in A3A responses across genetic**
214 **backgrounds**

215 We further examined global trends in response to A3A expression across all ~19,000
216 genes and twelve different experimental conditions, using principal components
217 (PC) analysis (Fig 2C). This suggested that globally, the results differ considerably
218 between the A549 and LXF-289 cell lines, indicating that the genetic background
219 modulates conditional essentiality of many genes under A3A conditions (data for top
220 hits shown in S5 Fig). In particular, the first two PCs explained 34% variability in the
221 data and separated the LXF-289 from the A549 cell line data points, but they did not
222 appreciably separate (i) the three different time points within each cell line, nor (ii)
223 the TP53 *wild-type* versus $TP53^{-/-}$ background of the A549 cell line, nor (iii) the two
224 different A3A doses (IC_{25} and IC_{50}) in the LXF-289 cell line. The same PC analysis
225 highlighted two genes with an extremely strong signal in the A3A response: *HMCES*,
226 because it is consistently observed across both genetic backgrounds (Fig 2C), and
227 the LXF-298-specific *HGC6.3* gene, which we suspect is an artefact (see above).
228 We further substantiated these results using MAGeCK-MLE[51]; in this analysis,
229 *HMCES* was the only gene which was conditionally essential (>2 standard deviations
230 away from the mean of the beta coefficients, per MAGeCK-MLE recommendation)
231 in late time-point samples in both $A549^{TP53-/-}$ and LXF-289 cells (S4 Fig).
232
233 Despite the apparent differences in the effects of individual genes between $A549^{TP53-}$
234 $^{/-}$ and LXF-289 cells (S2 Table), Gene Ontology analysis yielded consistent results
235 (Fig 2D), identifying DNA repair-related pathways as strongly enriched. In both cell
236 lines, DSB repair was a major enriched biological process, with homologous

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

237 recombination (HR) representing the predominant pathway, and to a lesser extent,
238 interstrand crosslink (ICL) repair (at $p<10^{-3}$ using the GORILLA server; see S3 Table
239 for list of results). Furthermore, nucleotide excision repair (NER) and DNA mismatch
240 repair (MMR) were enriched in both cell lines, as well as the regulation of the cell
241 cycle (Fig 2D and S3 Table). In LXF-289 cells, the non-homologous end-joining
242 (NHEJ) and Fanconi anemia pathways were strongly represented among the top hits
243 enriched, as well as *MCM8, MCM9* and *HROB*, genes that have previously been
244 implicated in HR (Fig 2D and S3 Table)[37,38,53]. Further enriched pathways related
245 to DNA repair included error-prone TLS and telomere maintenance in LXF-289 cells.
246 Intriguingly, there was also a strong enrichment of mRNA splicing genes (S3
247 Table). In $A549^{TP53-/-}$ cells, there was also enrichment of DSB repair via synthesis-
248 dependent strand annealing (SDSA) (Fig 2D and S3 Table). Overall, we conclude
249 that A3A expression induces dependencies on a variety of DNA repair and related
250 pathways in cells, some of which may be specific to certain genetic backgrounds,
251 while others appear more universal.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Fig 2. CRISPR/Cas9 genetic screen indicates HMCES and other DNA repair genes as vulnerabilities of A3A expressing cells. (**A**) Experimental design using the Brunello genome-wide library[50]. (**B**) Depletion of the sgRNAs targeting top four genes upon A3A overexpression, as prioritized by the overall A3A conditional essentiality score: LFC across three time points of the LXF-289 cell line and the three latest time points of the $A549^{TP53-/-}$ cell line. y-axis shows the median of the four sgRNAs per gene. (**C**) Principal component (PC) analysis of A3A-conditional LFC scores for all genes across all 12 experimental conditions (see labels next to arrows, which show loadings of the conditions on PC1 and PC2; "KO" implies TP53-/- and "WT" TP53 *wild-type* A549 cell line; numbers in labels are the time points; "IC25" and "IC50" are two concentrations of DOX in the LXF-289 cell line). The top 25 genes, prioritized by the same A3A conditional score as in panel **B** are highlighted on the figure. Inlay shows a scree-plot, with the amount of variance explained by the 12 PCs. The LXF-289-specific *HGC6.3* hit is likely an artefact (see Results text). Scores for all genes/sgRNAs are included in S2 Table. (**D**) Gene Ontology enrichment analysis of the top hits in the six experiments considered for the overall A3A conditional score (as in panel **B).** Plot shows –log10 p-value (unadjusted) from the GORILLA server; $p > 10^{-3}$ are not shown. Additional information in S3 Table. (**E**) Network schematic of cell cycle and DNA repair-related genes from the top 300 hits in the screen (overall score: OS). Genes identified in the Gene Ontology analysis and appearing in the top 300 genes by OS are shown. Color denotes the OS, lines indicate physical interactions (thebiogrid.org)[56] and a red border indicates they were identified in the Gene Ontology analysis in both cell lines.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Analyses of large-scale genetic screening data suggest a unique role of HMCES

Our genetic screens performed in different cancer cell lines yielded many A3A conditionally essential genes that were specific to one of the two cell lines (Fig 2C). Quality control parameters of the screening data indicated the high quality of all the screens by gDNA representation, by the ability to discriminate common essential genes, and by the separation between APOBEC conditionally essential genes and non-targeting, control sgRNAs (S6 Fig and S4 Table). Therefore, a likely explanation for the differences between cell lines could be that the genetic background and/or epigenetic state of a cell line determines its complement of essential genes upon A3A activation.

This motivated us to seek further evidence that the top hits we observed across both cell lines would indeed be valid across a wider spectrum of genetic backgrounds. To this end, we analyzed data from 76 lung adenocarcinoma, lung squamous cell carcinoma and head and neck squamous cell carcinoma cell lines (thus approximately matching our experimental models by tissue or cell type) from the Project Achilles database[57,58]. In particular, we searched among the top 10 genes from our experiments for correlations between the burden of A3 context mutations in the cell line exomes and the essentiality of a gene. By this metric, the *HMCES* gene obtained the highest scores in the external Project Achilles data (Fig 3A and S5 Table) for APOBEC signature 13 and signature 2 (slope of fit -0.29 and -0.5, respectively; combined p=0.03, t-test on the regression coefficient, one-tailed). In contrast, the *MCM8, RAD9A* and *ATXN7L3* genes, even though observed in both

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

cell lines in our experiments, did not score highly in this analysis (Fig 3A; we note that *MCM8* does rank more highly than other top hits from the genetic screen, but is nonetheless not robustly supported; S5 Table). This provided additional confidence that the synthetic interaction between *HMCES* and A3 activity is likely to hold across very diverse genetic backgrounds, as it is observed across a large cell line panel. A caveat of this analysis is that the A3 mutational signature may reflect past activity or intermittent activity of A3, and thus the lack of correlation in this analysis does not necessarily rule out the validity of the hit.

In addition to HMCES, the analysis of our genetic screening data revealed many common hits participating in DSB repair (Fig 2E). While it is likely that AP-sites resulting downstream of APOBEC lesions may generate DSBs in need of repair, such hits in the screen would plausibly also result from other agents inducing DSBs. Because our screening effort is focused on finding potentially actionable vulnerabilities, we were less interested in finding hits that result from DNA damaging conditions in general, which may abundantly occur also in healthy cells and are not linked to a genetic marker, in contrast to APOBEC activity, which may be more common in tumors and is evident in mutational signatures. We therefore analyzed data from previous genetic screens performed in the RPE1-TP53$^{-/-}$ cell line under a variety of different genotoxic agents[59]. We found that some of the common hits for A3A are also sensitizers in these genetic screens. For example, *RAD9A* loss sensitizes to a variety of agents including gemcitabine, hydroxyurea, bleomycin, AZD6738 (ATR inhibitor) and others (Fig 3B). *MCM8, MCM9* or *HROB* loss sensitized to MNNG, cisplatin, MMS, trabectedin and camptothecin (Fig 3B). Loss

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of the *DDX11* helicase or *MAD2L2* (also known as REV7, component of the shieldin complex and accessory subunit of the error-prone DNA polymerase zeta) sensitized to a wide gamut of DNA damaging agents tested (Fig 3B). This suggests these hits may be generally critical to stalled forks, rather than specific to A3A-mediated damage[6,59–61].

In contrast, *HMCES, UBA6* and *ATXN7L3* appeared to have a more restricted pattern of sensitization to DNA damaging agents (Fig 3B), indicating that they may represent better targets for selective killing of APOBEC expressing cancer cells. Of those, *HMCES* exhibited a distinctive pattern that did not cluster with the other top-50 hits in our screen (Fig 3B). *HMCES* loss sensitized to exposure to $KBrO_3$ and H_2O_2 (oxidizing agents), and ionizing radiation (IR), that generates oxidative base damage and DSBs, in previous data[59]. This is consistent with the occurrence of AP-sites as repair intermediates of oxidatively damaged DNA and a role for HMCES in protecting such AP sites. Intriguingly, *HMCES* loss also sensitized to illudin-S and duocarmycin, alkylating drugs with incompletely understood mechanisms-of-action, but less so to other alkylators (Fig 3B and S7 Fig)[59]. Overall, this joint analysis of previous genetic screening data under DNA damaging conditions, together with our APOBEC screens, indicates that HMCES has a specialized, rather than a general role in protecting against DNA damage. Further, this suggests that inhibiting HMCES would be selective for treating tumors undergoing certain types of DNA damage, such as APOBEC-mediated cytosine deamination, or in combination with specific therapeutic strategies, such as radiotherapy that is widely used in cancer treatment.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

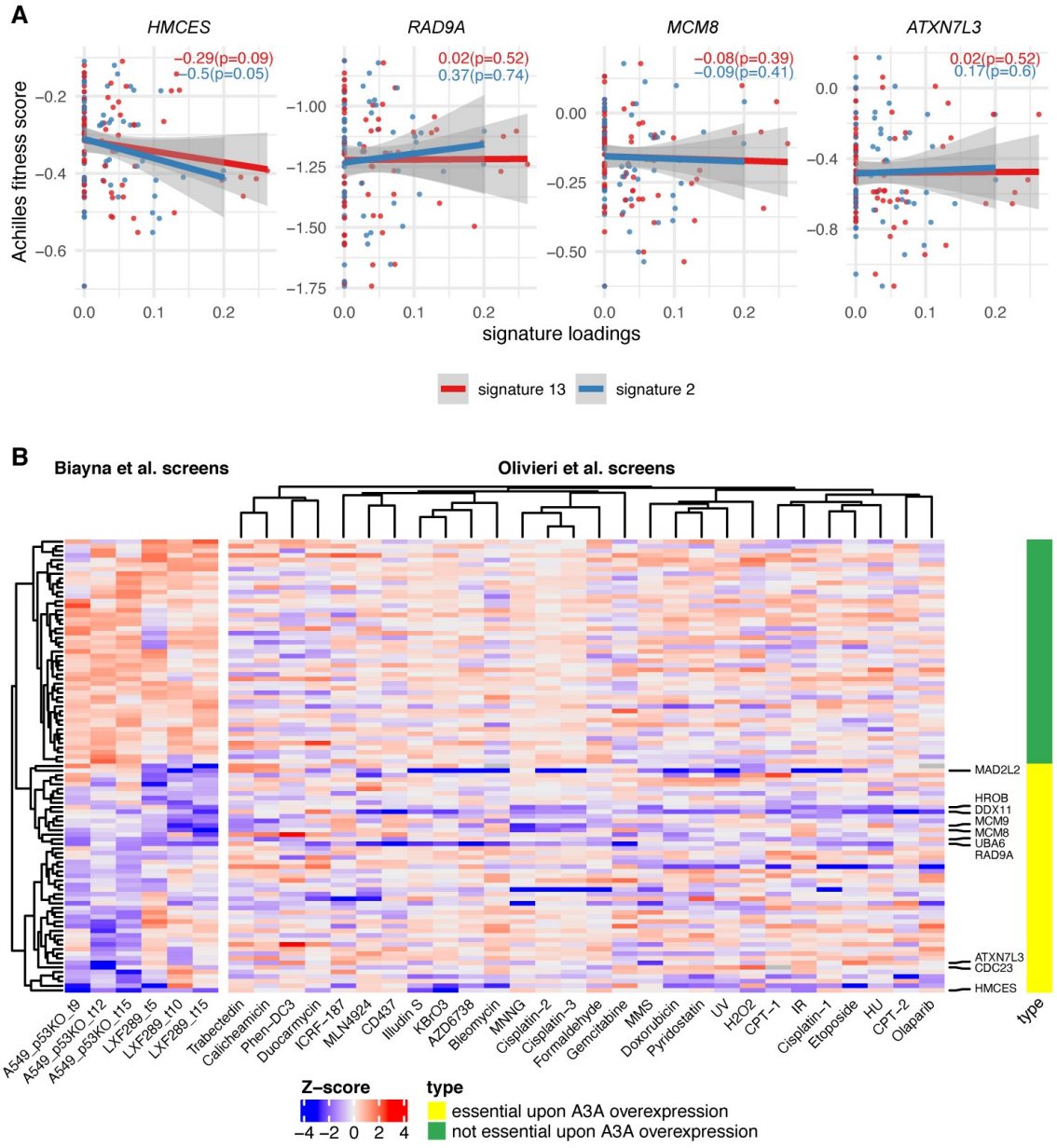


Fig 3. Dependency on HMCES is associated with mutational signatures of APOBEC across 76 lung and head-and-neck cancer cell lines. (**A**) Gene essentiality fitness score from Project Achilles versus APOBEC mutational signatures exposures, for cell lines from head-and-neck squamous cell carcinoma,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

lung adenocarcinoma, and lung squamous cell carcinoma, in four of the genes with the greatest overall score in our screens; see S5 Table and S8 Fig for associations with additional prominent genes. The slope and p-value (one-tailed, lower) for the regression model for both A3 signatures are shown within each panel. The more negative the slope the more sensitive the cell lines are to the depletion of the gene at a higher level of the APOBEC signature. (**B**) Heatmap shows a gene-level normalized log2 fold change (gene essentiality score) upon A3A overexpression for two cell lines and for three time points (Biayna et al. screens [59]); right panel shows Z-scores of gene essentiality after genotoxin exposure (Olivieri *et al.* screens). Data for 50 genes that are essential upon A3A overexpression in our screens (i.e. genes with the most negative mean log2 fold change across six data points) and 50 non-essential genes upon A3A overexpression in our screens. Labels on the right-hand side highlight the ten genes showing the highest overall A3A essentiality. An extended heatmap showing all genes from certain DNA repair pathways is included in S7 Fig.

HMCES depletion sensitizes A3A expressing cells

To test these possibilities, we depleted HMCES in multiple lung cancer cell line backgrounds by shRNA depletion or CRISPR/Cas9 knockout. Efficient depletion of HMCES mRNA and protein levels by shRNA in either LXF-289 or NCI-H358 (Fig. 4A), which was not used for the screening, enhanced sensitivity to A3A expression to different extents. Sensitivity in LXF-289 cells was apparent at early and late times (Fig. 4B), consistent with screening results, and accompanied by an arrest in G2/M phase (Fig. 4C) and increased levels of the γH2AX DNA damage marker (Fig. 4D).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

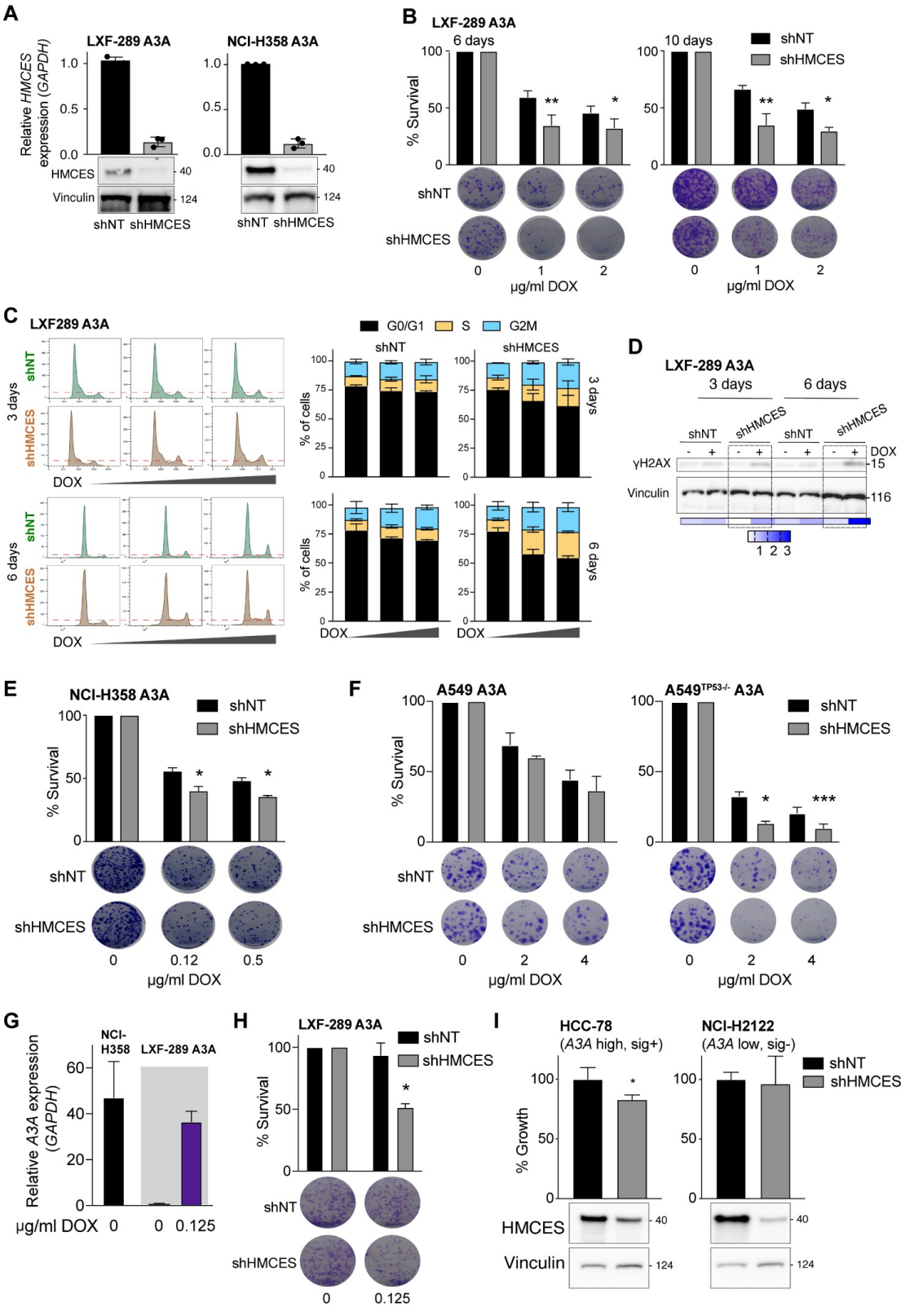
NCI-H358 cells also showed increased sensitivity to A3A expression (Fig. 4E). Together with the A549 screening data, these experiments further support that *HMCES* loss is more toxic to A3A-overexpressing cells in multiple genetic backgrounds.

We next asked whether the top hits in our A3A genetic screen were dependent on the activity of TP53, by comparing the derived A549$^{TP53-/-}$ cell line with its progenitor A549 that has a *wild-type TP53* status. Most of the top hits from the initial assay were not among the genes that differed depending on *TP53* status in A549. A prominent exception was *HMCES* (S5 Fig), which exhibits a stronger loss of fitness phenotype upon A3A expression in *TP53*$^{-/-}$ cells than in wild-type cells (we also noted some signal for *CDC23*; S5 Fig). As further support of this, in the statistical analysis of previous genetic screening data from Project Achilles (Fig 3A), we found that the association between the APOBEC mutational signatures and sensitivity to *HMCES* loss holds only for the *TP53* mutant cell lines, but not for the *TP53* wild-type cell lines (S9 Fig). To further test the epistatic interaction between *TP53* and *HMCES* under A3A expression, we directly assessed survival using colony forming assays in the A549 cell line pair following A3A expression and HMCES depletion. This showed that A549$^{TP53-/-}$ cells were more sensitive to A3A than A549 following HMCES depletion with shRNA (Fig 4F). This suggests that HMCES inhibition could be used to target A3A-expressing cells that have lost TP53, as is the case in many tumors, while sparing TP53 wild-type cells to a large extent.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Given that our experimental models rely on the inducible expression of exogenous *A3A*, we wanted to ascertain the relevance of these results to endogenous A3A expression levels in cancer cells. We examined endogenous *A3A* expression by qRT- PCR (Fig 4G) indicating that LXF-289 cells do not express detectable levels of *A3A*, while NCI-H358 cells do. We titrated our DOX levels down to achieve an induction of *A3A* mRNA levels in LXF289 cells similar to that of endogenous *A3A* that we observed in NCI-H358 cells (Fig 4G). We then examined the effect on cell growth and found that this impaired the growth of LFX-289 when HMCES was depleted (Fig 4H), consistent with experiments using higher DOX levels (Fig 4B). To further address the issue of applicability of HMCES inhibition to endogenous *A3A* levels, we examined public data for gene expression and mutational signatures in other NSCLC cell lines, highlighting two contrasting examples: HCC-78, that express *A3A* and exhibit APOBEC mutational signatures SBS2 and SBS13 (S10 Fig) [62,63], and NCI-H2122, that do not express detectable *A3A* nor exhibit the A3-mutational signatures (S10 Fig). We confirmed their relative *A3A* expression by qRT-PCR (S10 Fig) and depleted HMCES using shRNA. While both cell lines showed reduced levels of HMCES protein, only the naturally A3A-expressing HCC-78, but not the A3A non-expressing NCI-H2122, showed significant defects in cell growth upon HMCES depletion (Fig 4I). Together these data further support a role for HMCES in tolerating endogenous *A3A* expression in cancer cells.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



422

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Fig 4. Validation of the effects of HMCES depletion in multiple genetic backgrounds. (**A**) HMCES levels in LXF-289 A3A and NCI-H358 A3A cell lines by qRT-PCR and western blot following transduction with shHMCES or a non-targeting shRNA (shNT). The qRT-PCR validation was repeated at least three times and MW is indicated in kilodaltons. (**B**) Reduction of HMCES sensitizes LXF-289 A3A cells to *A3A* expression in a growth assay for 6 and 10 days. (**C**) Representative histograms of cell cycle progression (left panels) and quantitative analysis of LXF-289 A3A shHMCES and shNT (right panels). Cells were treated with DOX (0, 1 or 2 ug/ml) and harvested after 3 and 6 days. (**D**) Western blot of H2AX-S139 phosphorylation (γH2AX) in LXF-289 shHMCES and shNT cells after 3-6 days of A3A expression. Relative phosphorylation (0-3) was calculated normalizing the band densities of γH2AX to total Vinculin signal. MW is indicated in kilodaltons. (**E**) Reduction of HMCES sensitizes NCI-H358 A3A cells to *A3A* expression in a clonogenic survival assay after 15 days. (**F**) The effect of HMCES depletion is TP53-dependent. Clonogenic survival assays of A549 or A549$^{TP53-/-}$ cells are shown 10 days after treatment with the indicated dose of DOX. (**G**) *A3A* mRNA expression levels in LXF-289 A3A (0 and 0.125 ug/ml of doxycycline) and the parental NCI-H358 cell line relative to *GAPDH* measured by qRT-PCR (repeated two times, mean and SEM are shown). (**H**) Depletion of HMCES sensitizes LXF-289 A3A cells to *A3A* expression following low levels of DOX treatment in a clonogenic survival assay. (**I**) Growth inhibition (percentage of growth rate) measured with alamarBlue for HCC-78 (A3A expressing, A3 mutational signature positive) and NCI-2122 (A3A low, A3 mutational signature negative) cell lines transduced with shNT or shHMCES. HMCES levels are

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

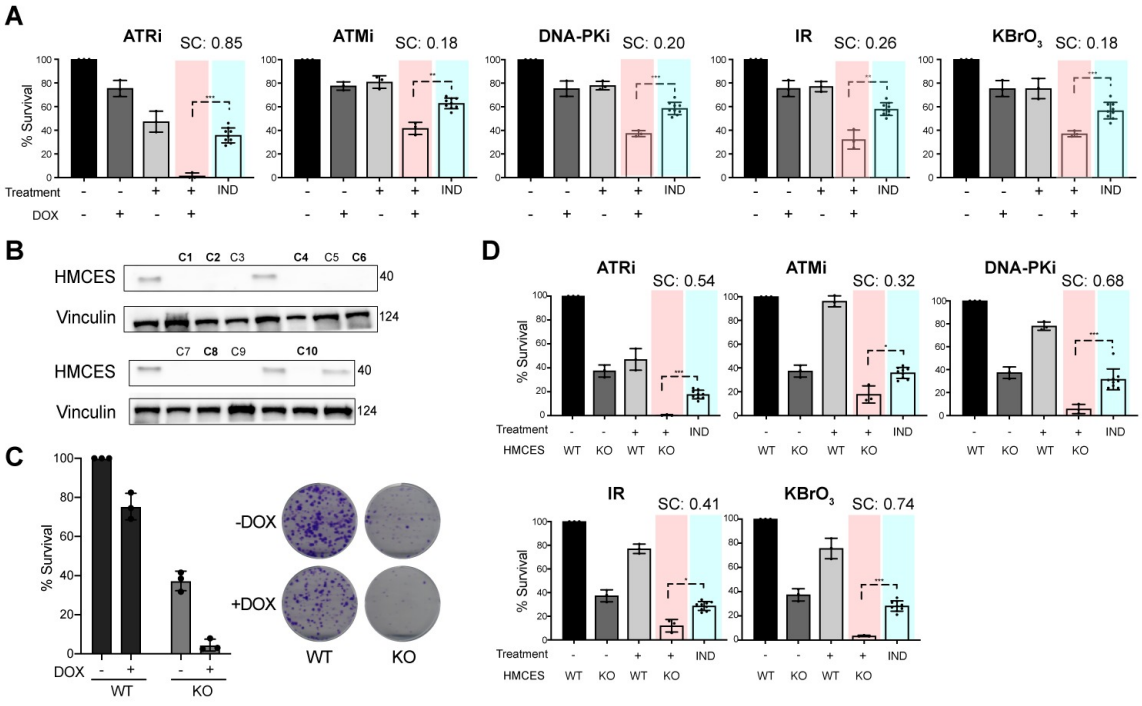
shown, Vinculin is used as a loading control. Statistical analysis of shHMCES versus shNT in all panels was performed using a one-tailed unpaired *t*-test; error bars indicate SD. *, $p \leq 0.05$, **, $p \leq 0.01$, ***, $p \leq 0.001$. For each time point, growth assays and clonogenic survival assays were repeated 3 times.

Sensitivity of A3A expressing cells to HMCES loss is enhanced by DNA damage

In addition to sensitizing to A3A, previous data implicated HMCES in the sensitivity to a limited number of DNA damaging agents that included ionizing radiation (IR) and $KBrO_3$[42,59]. Considering this, and that DNA repair factors involved in multiple DSB repair pathways were identified in our A3A screens (Fig 2E), we examined the effects of combinatorial treatments on A3A-mediated toxicity. Survival of LXF-289 A3A cells was analyzed with or without DOX in combination with IR or $KBrO_3$ treatment. Both agents showed increased toxicity in cells expressing A3A (Fig 5A; $p \leq 0.05$ for synergistic activity, by t-test against a Bliss independence baseline)[64]. We next examined the relative importance of the three PI3K-like kinases (PIKKs), ATM, ATR and DNA-PKcs, that regulate many of the individual proteins identified in the screens. LXF-289 cells were treated with DOX to induce A3A and each of the PIKKs was inhibited using small molecule inhibitors[65]. As previously reported, we saw that the toxicity of A3A overexpression was strongly enhanced following treatment with ATR inhibitors[32,33]. In addition, we saw that inhibitors of ATM and DNA-PKcs, that play a key role in DSB repair, led to synergistic cell killing with A3A overexpression ($p \leq 0.05$) albeit to a more modest extent than with the ATR inhibitor or with IR (Fig 5A).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

We next examined the influence of HMCES on combination treatments by generating knockout cell lines (HMCES KO) in the LXF-289 A3A background using CRISPR/Cas9 and single cell isolation (Fig 5B). Five clones with no detectable HMCES protein levels (Fig 5B; bold type) were pooled to generate an HMCES KO cell culture for subsequent analysis. HMCES KO impaired the colony-forming capacity of LXF-289 A3A cells, and this was further reduced upon expression of A3A following DOX treatment (Fig 5C). HMCES KO cells also showed hypersensitivity to the inhibition of any of the PIKKs, in the absence of exogenous DNA damaging agents, as well as to treatment with IR or $KBRO_3$ (Fig 5D). Together, our data point to HMCES as an important dependency of cancer cells for survival and suggest that this can be exploited using combination therapies.



bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Fig 5. A3A expression sensitizes to DNA damage and HMCES loss. (**A**) Clonogenic survival measured by the colony formation assay in LXF-289 A3A cells after exposure to the indicated small molecule inhibitor, IR (5 Gy), or treatment with 0.1 mM $KBrO_3$ with or without DOX (0.125 ug/ml) to induce *A3A* expression. (**B**) HMCES western blot of lysates from HMCES WT and KO LXF-289 A3A clones. Vinculin was used as a protein loading control and MW is indicated in kilodaltons. (**C**) Clonogenic survival assay of LXF-289 HMCES WT and KO cells upon over-expression of A3A by DOX. (**D**) Clonogenic survival assay comparing HMCES WT and KO cells after treatment with the indicated small molecule inhibitor, exposure to IR (5 Gy), or treatment with 0.1 mM $KBrO_3$ with or without DOX to induce A3A expression. For panels A and D, the "IND" column shows a Bliss independence model of additive activity of the two treatments, against which the combined treatment is tested (using *t*-test, two-tailed) to estimate synergistic activity[64]. SC, synergy score. *, $p \leq 0.1$, **, $p \leq 0.05$ ***, $p \leq 0.01$. Each experiment was repeated two times.

Screening HMCES deficient cells reveals additional modifiers of the A3A response

As HMCES KO cells could still tolerate some A3A expression, we performed a secondary CRISPR/Cas9 genome-wide screen to identify mediators of survival that were specific for LXF-289 A3A HMCES KO cells (Fig 6A). Positively selected genes in this assay are those whose deletion lessens the fitness penalty due to loss of *HMCES* upon *A3A* overexpression (alleviating epistasis), while negatively selected genes are those whose deletion increases this fitness penalty (aggravating

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

507 epistasis). Expectedly, this screen identified a strong positive selection for the loss
508 of *A3A*, presumably by targeting our inducible gene. Further, it identified *UNG*, the
509 gene encoding the primary uracil-DNA glycosylases UNG1 and UNG2, that localize
510 to the mitochondria and nucleus, respectively. This indicated that preventing the
511 generation of AP-sites by A3A conferred survival to HMCES KO cells, consistent
512 with previous work[22,49]. In addition, the loss of multiple genes encoding subunits
513 of the Mediator complex (*CCNC, MED24, MED25* and *MED16*), splicing regulators
514 (*SCAF1, SCAF8*), the *FBXW7* E3 Ubiquitin ligase (a common tumor suppressor
515 gene), the mismatch repair protein *MSH2*, the Protein phosphatase 4 subunit
516 *SMEK1* (PPP4R3A) that dephosphorylates γH2AX[66], the *ATF2* transcription factor
517 and *FADD* (FAS-associated death domain protein), were among high scoring hits
518 that were positively selected in HMCES KO cells (S6 Table).
519
520 Among negatively selected genes in HMCES KO cells, *TDP1* (tyrosyl-DNA
521 phosphodiesterase 1) was one of few genes implicated in DNA repair. TDP1 can
522 resolve 3'-AP-sites and is essential in cells lacking APE1, which promotes the repair
523 of AP-sites by BER[67–69]. *VCPIP1* was also a strong hit implicated in DNA repair.
524 VCPIP1 is a deubiquitinase that is activated by ATM/ATR and involved in the
525 removal of protein-DNA crosslinks through the regulation of SPRTN, that is critical
526 for damage prevention during DNA replication[70,71]. In addition, the gene encoding
527 the Protein phosphatase 2A subunit (*PPP2R2A*) was negatively selected. PPP2R2A
528 is a negative regulator of ATM-CHK2 and a candidate tumor suppressor gene
529 commonly deleted in ovarian, prostate, liver and bladder cancers[72–74].

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Together, these data indicate that the sensitivity of HMCES null cells can be mitigated by the loss of *UNG*, implicating AP-site generation in the toxicity, and that TDP1, and potentially VCPIP1/SPRTN, likely represent key backup activities for the resolution of APOBEC3A-mediated damage to promote cell survival.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

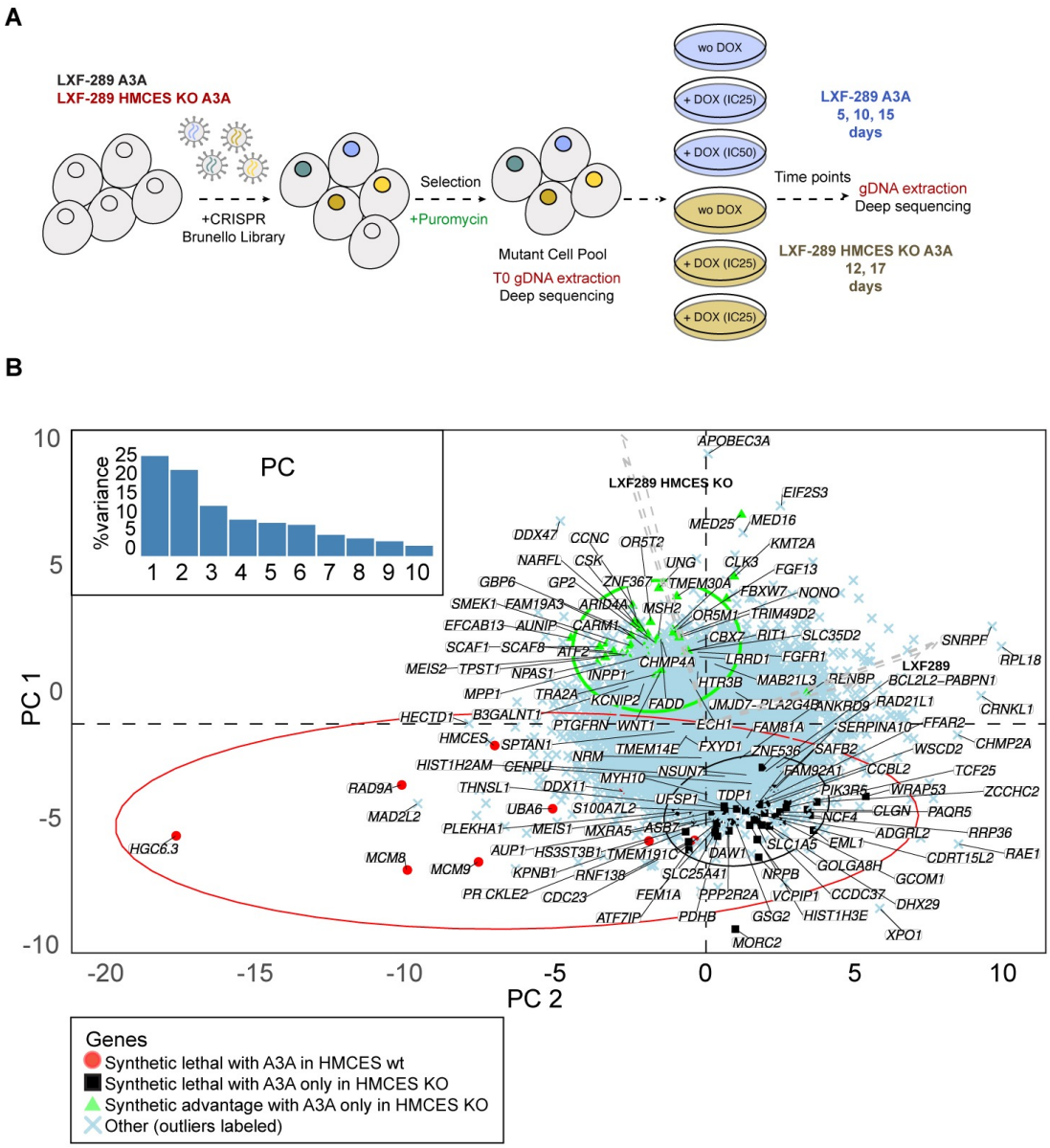


Fig 6: Secondary screening identifies modifiers of the response to A3A in HMCES KO LXF-289 cells. (**A**) Schematic of the secondary screen in HMCES KO LXF-289 cells. (**B**) PC analysis of A3A-conditional LFC scores for all genes across all 10 experimental conditions (see labels next to arrows, which show loadings of the conditions on PC1 and PC2; HMCES wt refers to the parental LXF-289 cell line. Top genes, prioritized by score are highlighted on the figure for each indicated

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

547 comparison: green, alleviating epistasis; black, aggravating epistasis; red synthetic
548 lethal in HMCES wt. Gene level scores and GO analysis are included in S6 Table
549 and S7 Table and additional epistasis analysis is included in S11 Fig.
550
551 **Discussion**
552 Our results establish that HMCES is a key mediator of A3A toxicity in cancer cells.
553 Given the increased levels of γH2AX DNA damage signaling observed in HMCES
554 knockdown cells, as well as the enrichment in DSB repair factors in our screens, our
555 results suggest that DSBs may be a major driver of A3A toxicity. Notably absent in
556 our primary screens were factors involved in the BER pathway that would normally
557 resolve AP-sites and promote the base changes that are evident in the APOBEC
558 mutational signature. We interpret this to mean that A3A lesions are likely not toxic
559 outside of S-phase, where they can be repaired by BER and likely additional
560 pathways.
561
562 This proposition is consistent with recently published work demonstrating that
563 HMCES depletion sensitizes both immortalized human cells and cancer cells to A3A
564 expression[49]. In addition, the study showed that HMCES loss reduced replication
565 fork elongation in a manner dependent on the UNG-mediated production of AP-sites,
566 using a uracil-DNA glycosylase inhibitor. UNG2 depletion was also previously shown
567 to suppress the accumulation of replication fork associated AP-sites and DSBs in
568 ATRi treated A3A-expressing cells[33]. Curiously, UNG2 depletion caused lethality
569 in A3B-expressing cells through a mechanism dependent on MMR and TP53[75].
570 We did not identify any glycosylases as sensitizers or suppressors of A3A-mediated

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

571 toxicity in our primary screens, regardless of TP53 status. However, our secondary
572 screen in HMCES KO cell lines identified UNG as a major positively selected hit in
573 cells expressing A3A, reinforcing the proposition that AP-site production underlies
574 the sensitivity of cells to A3A expression[42,49]. It also suggests a mechanistic
575 divergence between the toxicity of A3A and A3B expression. In addition, our
576 secondary screens identified the mismatch repair protein MSH2 as positively
577 selected, consistent with our recent data suggesting that APOBEC3A-mediated
578 mutagenesis is mediated by MMR activity, resulting in the characteristic *omikli*
579 pattern of clustered mutations[7]. In contrast, UNG, MSH2 or other MMR proteins,
580 were not identified as strong influencers of the survival of HMCES KO cells in the
581 absence of A3A or DNA damage in our screens, or in recent screens performed by
582 others in HMCES KO cells in a HEK293 background[76].
583
584 Replication fork slowing following A3A expression was shown to be due in part to
585 the recruitment of TLS polymerases, particularly POLζ, as well as increased
586 accessibility to APE1 endonuclease activity that is likely the source of DSBs[49].
587 While we did not identify APE1, potentially due to redundancy with other activities,
588 we did find the POLζ subunit MAD2L2 as an A3A sensitizer (Figs 2C-2E). Notably,
589 this was specific to the LXF-289 cancer cell line, suggesting that the genetic
590 background may have an appreciable impact on replication fork protection and
591 stability and how cells respond to AP-sites at replication forks.
592
593 Aside from HMCES, the overall agreement between cell lines at the individual gene
594 level was limited in our A3A screens. A previous analysis of essential genes using a

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

595 DNA repair library in HEK293 HMCES KO cells identified numerous proteins
596 involved in HR and TLS, suggesting that loss of HMCES was synthetically lethal with
597 the attenuation of these repair pathways[76]. In the absence of A3A expression, we
598 observed very limited overlap with that study when compared to our data in LXF-289
599 HMCES KO cells, with the ribonucleotide reductase subunit RRM1, clamp loader
600 subunit RFC3 and HR factor XRCC3 being the only notable overlapping hits among
601 negatively selected genes. Together, this further highlights that the individual genetic
602 or epigenetic status of particular cell lines may play a significant role in the response
603 to A3A expression, as well as to HMCES loss. This is supported by the fact that A3A
604 expression caused variable levels of cell cycle arrest that was cell line dependent
605 (Fig 1). This may reflect the status of individual DNA repair pathways or DNA
606 replication fork rates in the different genetic backgrounds. Despite the different
607 phenotypic outcomes between cell lines, HMCES emerged as a common and
608 prominent hit between the cell lines screened in our work (using an experimental
609 system for A3A overexpression), as well as many other cell lines examined in the
610 Project Achilles screens (using observational analysis of A3A mutational signatures
611 in the cell line exomes)[77].
612
613 TP53 status clearly has a major influence on the genetic interaction between
614 HMCES loss and A3A expression, implicating G1/S checkpoint status in the
615 tolerance of A3A expression. TP53 status was shown to have a major influence on
616 CRISPR/Cas9 survival screens and it is intuitive that checkpoint status would limit
617 toxic DNA damage generated during S-phase and prevent mitotic catastrophe
618 resulting from under-replicated DNA entering mitosis[78]. As A3A-mediated damage

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of ssDNA is S-phase specific, our results would again be consistent with the recently proposed model that HMCES shields AP-sites in ssDNA from processing by BER endonucleases that would generate AP-sites and toxic DSBs, or from replication by TLS polymerases that would result in increased mutagenesis[27,42,43,49,76]. As the enzymatic activities of HMCES have been implicated in this function, accumulating data suggests that targeting HMCES would be an attractive strategy for the specific sensitization of A3A-expressing, TP53-deficient cancers.

Inhibition of the ATR-CHK1 kinases, that activate cell cycle checkpoints in S and G2 in response to replication stress, was shown to enhance A3A-mediated toxicity in AML, lung, ovarian and osteosarcoma cell lines[32,33]. Our results further extend the robustness of this observation to additional lung cancer cell lines. In contrast to previous work, we also found increased sensitivity to ATM and DNA-PKcs inhibitors in LXF-289 NSCLCs lacking HMCES, potentially reflecting differences in the genetic backgrounds analyzed[75]. As inhibitors for the PIKKs are in multiple clinical trials, targeting HMCES may represent a strategy to enhance their efficacy, particularly in combination with radiotherapy. This may be particularly potent in cancers with *PPP2R2A* deletions[74]. Loss of PPP2R2A enhanced toxicity of A3A in HMCES KO cells, and depletion or PPP2R2A was shown to enhance the toxicity of ATR-CHK1 inhibitors in NSCLC[79]. Our finding that TDP1 was the primary DNA repair protein negatively selected specifically in HMCES KO cells expressing A3A indicate that TDP1 inhibitors could be also used in conjunction with HMCES depletion/inhibition to sensitize APOBEC3-expressing cancer cells. Numerous TDP1 inhibitors are

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

currently being explored in clinical trials to sensitize cells to topoisomerase inhibitors[80].

In addition to sensitizing to A3A-expression, HMCES deficiency also sensitized cells to treatment with IR and $KBrO_3$ treatment (Figs 3 and 5)[42,59]. As previously discussed, HMCES depletion sensitizes to a very limited spectrum of damaging agents compared to other hits in our screen that play more general roles in damage tolerance. The observation that additional hits, namely the poorly characterized SAGA complex component ATXN7L3 and ubiquitin-activating enzyme UBA6, shared overlapping profiles of DNA damage sensitivity as HMCES, suggests that a more definitive characterization of their roles is warranted in future work.

Collectively, existing data suggests that inhibition of HMCES is a promising strategy to suppress the APOBEC overexpressing, hypermutating tumor cell population, thereby slowing down the accumulation of genetic heterogeneity and preventing acquisition of new driver mutations or drug resistance mutations. Moreover, HMCES inhibition could augment the use of radiotherapy, which is widely used in the treatment of many cancer types and enhance the effectiveness of small molecule inhibitors for DNA damage signaling kinases and repair enzymes that are currently being developed and tested in clinical trials.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Methods

Cell culture and generation of doxycycline-induced lung adenocarcinoma (LUAD) cell lines

LXF-289, NCI-H358, HCC-78, NCI-H2122 and A549 cell lines were purchased from the American Type Culture Collection (ATCC) and the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ) and maintained with RPMI-1640 or DMEM medium and supplemented with 10% fetal bovine serum and 5% penicillin-streptomycin. The doxycycline-inducible HA-tagged A3A plasmid (pSLIK-Neo A3A) was a kind gift from the Weitzman lab[27]. Lentiviral particles were generated by transfection of HEK-293T cells. After transduction with the pSLIK-A3A lentivirus, LUAD cells were selected in 1 mg/mL Geneticin (Ibian Technologies).

Growth arrest and colony formation assays

For growth assays**,** cells were plated at density of 1,000 cells/well in a 96 well plate. After 24 hrs, cells were treated with increasing doses of doxycycline (0-8 ug/ml) and cultured for 72 hrs. AlamarBlue reagent (ThermoFisher Scientific) was added to cell culture media 4 hrs prior to reading fluorescence with a SYNERGY H1M fluorescence plate reader**.** For the colony formation assays, between 250-1000 cells per well were plated in a 12 well plate. Colonies were fixed with formalin (Sigma) and stained with a 0.01% crystal violet (Sigma) solution in 20% methanol. For some cell lines, quantification was performed by reading absorbance at 590 nm after the addition of 10% acetic acid.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Drug sensitivity assays

Colony-forming assays for drug sensitivity testing were performed by plating the cells at a density of 500 cells/well in a 6 well-plate, in triplicate. 24 hrs after plating, the following drug treatments were used: DNA-PKi (KU57788) 1uM (MedChemExpress), ATMi (KU55933) 5uM (Sigma), ATRi (AZD6738) 0.5uM (MedChemExpress), $KBrO_3$ 0.1mM (Sigma). IR (5 Gy) was administered using a Maxishot.200 X-Ray cabinet (Krautkramer Forster). For the induction of A3A expression, doxycycline was added at a concentration of 0.125 ug/ml (IC_{25}). The drug treatment was maintained in the growth media for the duration of the experiment (10 days), after which cells were fixed and stained with crystal violet. The number of colonies was quantified with Fiji (ImageJ). Colony-forming capacity is presented as a percentage of the vehicle-treated (DMSO 0.025%) control.

Cell cycle analysis

Cells were fixed in 70% ethanol for at least 2 hrs at -20ºC and resuspended in a PBS solution containing 35 ug/ml propidium iodide (Sigma) and 100 ug/ml RNAse A (Roche). Between 5000-10000 cells were analyzed per sample. Data was acquired on a Gallios A94303 Flow Cytometer (Beckman Coulter) in the Cytometry core facility of the University of Barcelona and analyzed by FlowJo software.

RNA extraction, cDNA synthesis and qRT-PCR

RNA extraction was performed using the Maxwell 16 LEV simplyRNA cell Kit (Promega) according to manufacturer's instructions. cDNA synthesis was performed

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

with the high-capacity cDNA reverse transcription kit (Life Technologies). Quantitative PCR (qRT-PCR) was performed with SYBR Select Master Mix for CFX (Applied Biosystems) or TaqMan universal PCR Master Mix II (Applied Biosystems) on a StepOnePlus Real-time PCR System (Applied Biosystems). Probes and primers are shown at S8 Table.

Generation of stable KO and knockdown cells

For A549 ± A3A and LXF-289 ± A3A cell lines, the NickaseNinja (ATUM) vector co-expressing two gRNAs (pD1401-AD: CMV-Cas9N-2A-GFP, Cas9-ElecD) was used to generate the TP53 KO and the HMCES KO cells. TP53 gRNA sequences (GCAGTCACAGCACATGACGG) (GATGGCCATGGCGCGGACGC) and HMCES gRNA sequences (CAGTGAATGGATCTCTACAA) (GAGCTTGCGCCTACCAGGAT) were designed using the ATUM gRNA Design Tool. 48 hrs post-transduction, positive GFP cells were sorted by FACS (BD FACSAriaTM Fusion) and plated into 96-well plates. After 15 days, clones were collected and validated by western blot using the following primary antibodies: p53 (sc-47698, Santa Cruz); vinculin (V9264, Sigma); HMCES (HPA044968, Atlas Antibodies). The HMCES knockdown stable cell lines (HCC-78, NCI-H2122, LXF-289 A3A, NCI-H358 A3A, A549 A3A, and A549$^{TP53-/-}$ A3A) were made using the Mission shRNA lentiviral vector NM_020187.1-133s1c1 (Sigma). Lentiviral particles were produced in HEK293T cells using a pLKO.1-shRNA plasmid. The cell lines were transduced and selected with puromycin for 72 hrs. As a control, we transduced LXF-289 (A3A) cells with the non-mammalian shRNA Control Plasmid DNA shC002 (Sigma).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

CRISPR/Cas9 screening

For sgRNA screening of the A549 ± A3A, A549$^{TP53-/-}$ ± A3A, LXF-289 ± A3A, cells were infected with the Brunello CRISPR Knockout Pooled Library (73179-LV, Addgene). Infection with lentiviruses was performed at a MOI≤0.4 for all cell lines. At 24 hrs post-infection, the medium was replaced with a selection medium containing puromycin (2 ug/mL). After 5-6 days of selection, cells were split into the different experimental conditions: For LXF-289 cell line, without and with doxycycline (0.125 and 2 ug/ml corresponding to IC_{25} and IC_{50} respectively). For LXF-289 HMCES KO secondary screening, without and with doxycycline (0.03 and 0,125 ug/ml corresponding to IC_{25} and IC_{50} respectively). For A549 cell line, without and with doxycycline (3.9 ug/ml). All cell lines were passaged every 3 days (up to 15 days) and for each time point, the number of cells needed to maintain the predetermined coverage of 400-500 fold was taken. DNA extraction was performed using the DNA genomic Kit (Puregene Cell and Tissue Kit).

NGS Library Preparation and sequencing

NGS libraries were prepared by two-step PCR, for the first one a total of 20 ug of DNA per a 12X reaction was used, for the second PCR a set of primers harboring Illumina TruSeq adapters as well as the barcodes for multiplexing were used (for all primers used see S8 Table). Sequencing was carried out in the CNAG sequencing unit using 6 lanes of a 1x50 HiSeq.

Statistics

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

751 The statistical analyses were performed using Prism software version 8.0. Each
752 functional experiment was repeated two times or three times (as specified in the
753 figure or legend). Differences between groups were analyzed by the Student t-test
754 assuming unequal variances.

755 **Independent validation**

756 We downloaded mutational signatures for the cell lines from Petljak et al[63] and
757 gene essentiality fitness score from Project Achilles[57]. We selected the cell lines
758 from HNSC, LUAD and LUSC. For the top 10 scoring genes in our analysis, we fitted
759 a linear regression model between the cell lines fitness score and the signature
760 loadings for signatures SBS2, SBS13 and SBS2+SBS13 (APOBEC signatures). We
761 compared the slope and p-values obtained. The p-value is obtained from a t-test
762 (one-tailed lower).

763 ***In silico* analysis of DNA damage sensitivity to DNA damaging agents**

764 We downloaded the previously published data for the z-scores after genotoxin
765 exposure screens[59]. We compared how our top 50 genes (essential upon A3A
766 overexpression) versus 50 genes that are not essential in our screens behaved after
767 the genotoxin exposure.

768 ***In silico* analysis of CRISPR/Cas9 screening results**

769 For alignment of the generated reads to the library, read counting, read count
770 normalization, quality control (QC) analysis of the samples, and calculation of the
771 sgRNA counts log2 fold change, we used MAGeCK-VISPR[51]. For pair wise

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

772 comparisons, we employed the robust rank aggregation (RRA) algorithm, using as
773 treatment the DOX-induced A3A sample, for each cell line (A549, A549$^{TP53-/-}$, and
774 LXF-289) and time point.
775
776 Estimation of gene essentiality and sgRNA efficiency was achieved using the
777 maximum-likelihood estimation (MLE) algorithm provided by MAGeCK-VISPR.
778 Namely, gene essentiality was estimated by comparison of the normalized sgRNA
779 counts between each sample (A549 and A549$^{TP53-/-}$ time 9, 12, and 15, and LXF-289
780 time 5, 10, and 15) and its corresponding time 0 sample, which yielded a beta score
781 per gene and sample. The beta score distribution for each sample was standardized
782 by subtraction of the mean and division by the standard deviation (SD), and a final
783 gene essentiality score was obtained by averaging the resulting *Z*-scores across
784 samples.
785
786 Finally, we used the FluteMLE function from the R package “MAGeCKFlute”[81] for
787 i) normalization of the beta scores yielded by MAGeCK-VISPR MLE using a built-in
788 set of 622 essential genes as a reference, and ii) comparison of the essentialities
789 between conditions (DOX-induced A3A vs. control) within each cell line and time
790 point, applying a significance cutoff of two SD (S4 Fig). This allowed us to identify
791 genes that were negatively selected in the A3A-expressing samples, but not selected
792 in the control samples.
793
794 **Data Availability Statement**

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

795 All data is included as Supplementary material or is available upon request. Analyses
796 that used data from Olivieri *et al.* [59] are available at Mendeley data:
797 https://data.mendeley.com/datasets/gfcn2wmrpf/1.
798

799 ## Acknowledgements

800 We thank members of the Stracker and Supek labs for input and discussions, M.
801 Weitzman and A. Green for sharing reagents and unpublished data.

802 ## Funding

803 J.B., M.M. and F.S. were supported by the ERC Starting Grant (757700 “HYPER-
804 INSIGHT”, to F.S). I.G.C. was funded by an AECC fellowship. M.M.A. and J.E.C.
805 were supported by the Spanish Ministry of Science, Innovation and Universities
806 (BFU2017-89833-P "RegioMut", to F.S.). T.H.S. was funded by the Spanish Ministry
807 of Science, Innovation and Universities (MCIU: PGC2018-095616-B-I00/GINDATA
808 and FEDER) and the Intramural Research Program of the National Institutes of
809 Health, National Cancer Institute. The T.H.S. and F.S. labs are supported by the
810 Centres of Excellence Severo Ochoa award and the CERCA Programme. F.S. is
811 funded by the ICREA Research Professor programme. The funders had no role in
812 study design, data collection and analysis, decision to publish, or preparation of the
813 manuscript.

814 **Author contributions**

815 J.B. performed the majority of the experiments, designed experiments, analyzed
816 data and edited the manuscript, I.G.C. performed experiments and analyzed data,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

817 M.M.A., M.S. and J.E.C. performed computational analyses, M.M performed
818 experiments, T.H.S. and F.S. supervised the project, designed experiments,
819 analyzed data, and wrote and edited the manuscript.
820
821 **Competing interests statement:**
822 The authors declare no competing interests
823

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

References

1. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, et al. Mutational processes molding the genomes of 21 breast cancers. Cell. 2012;149: 979–993. doi:10.1016/j.cell.2012.04.024
2. Nik-Zainal S, Wedge DC, Alexandrov LB, Petljak M, Butler AP, Bolli N, et al. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. Nat Genet. 2014;46: 487–491. doi:10.1038/ng.2955
3. Burns MB, Temiz NA, Harris RS. Evidence for APOBEC3B mutagenesis in multiple human cancers. Nat Genet. 2013;45: 977–983. doi:10.1038/ng.2701
4. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet. 2013;45: 970–976. doi:10.1038/ng.2702
5. Roberts SA, Sterling J, Thompson C, Harris S, Mav D, Shah R, et al. Clustered mutations in yeast and in human cancers can arise from damaged long single-strand DNA regions. Mol Cell. 2012;46: 424–435. doi:10.1016/j.molcel.2012.03.030
6. Chen J, Miller BF, Furano AV. Repair of naturally occurring mismatches can induce mutations in flanking DNA. Elife. 2014;3: e02001. doi:10.7554/eLife.02001
7. Mas-Ponte D, Supek F. DNA mismatch repair promotes APOBEC3-mediated diffuse hypermutation in human cancers. Nat Genet. 2020;52: 958–968. doi:10.1038/s41588-020-0674-6
8. Henderson S, Chakravarthy A, Su X, Boshoff C, Fenton TR. APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. Cell Rep. 2014;7: 1833–1841. doi:10.1016/j.celrep.2014.05.012
9. McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. Sci Transl Med. 2015;7: 283ra54. doi:10.1126/scitranslmed.aaa1408
10. Cannataro VL, Gaffney SG, Sasaki T, Issaeva N, Grewal NKS, Grandis JR, et al. APOBEC-induced mutations and their cancer effect size in head and neck squamous cell carcinoma. Oncogene. 2019;38: 3475–3487. doi:10.1038/s41388-018-0657-6
11. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013;500: 415–421. doi:10.1038/nature12477
12. Lada AG, Dhar A, Boissy RJ, Hirano M, Rubel AA, Rogozin IB, et al. AID/APOBEC cytosine deaminase induces genome-wide kataegis. Biol Direct. 2012;7: 47; discussion 47. doi:10.1186/1745-6150-7-47
13. Taylor BJ, Nik-Zainal S, Wu YL, Stebbings LA, Raine K, Campbell PJ, et al. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. Elife. 2013;2: e00534. doi:10.7554/eLife.00534

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

14. Franco I, Helgadottir HT, Moggio A, Larsson M, Vrtačnik P, Johansson A, et al. Whole genome DNA sequencing provides an atlas of somatic mutagenesis in healthy human cells and identifies a tumor-prone cell type. Genome Biol. 2019;20: 285. doi:10.1186/s13059-019-1892-z
15. Ullah I, Karthik G-M, Alkodsi A, Kjällquist U, Stålhammar G, Lövrot J, et al. Evolutionary history of metastatic breast cancer reveals minimal seeding from axillary lymph nodes. J Clin Invest. 2018;128: 1355–1370. doi:10.1172/JCI96149
16. Akre MK, Starrett GJ, Quist JS, Temiz NA, Carpenter MA, Tutt ANJ, et al. Mutation Processes in 293-Based Clones Overexpressing the DNA Cytosine Deaminase APOBEC3B. PLoS One. 2016;11: e0155391. doi:10.1371/journal.pone.0155391
17. Nikkilä J, Kumar R, Campbell J, Brandsma I, Pemberton HN, Wallberg F, et al. Elevated APOBEC3B expression drives a kataegic-like mutation signature and replication stress-related therapeutic vulnerabilities in p53-defective cells. Br J Cancer. 2017;117: 113–123. doi:10.1038/bjc.2017.133
18. Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc EP, et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet. 2015;47: 1067–1072. doi:10.1038/ng.3378
19. Supek F, Lehner B. Clustered Mutation Signatures Reveal that Error-Prone DNA Repair Targets Mutations to Active Genes. Cell. 2017;170: 534–547.e23. doi:10.1016/j.cell.2017.07.003
20. Cortez LM, Brown AL, Dennis MA, Collins CD, Brown AJ, Mitchell D, et al. APOBEC3A is a prominent cytidine deaminase in breast cancer. PLoS Genet. 2019;15: e1008545. doi:10.1371/journal.pgen.1008545
21. Law EK, Levin-Klein R, Jarvis MC, Kim H, Argyris PP, Carpenter MA, et al. APOBEC3A catalyzes mutation and drives carcinogenesis in vivo. J Exp Med. 2020;217. doi:10.1084/jem.20200261
22. Landry S, Narvaiza I, Linfesty DC, Weitzman MD. APOBEC3A can activate the DNA damage response and cause cell-cycle arrest. EMBO Rep. 2011;12: 444–450. doi:10.1038/embor.2011.46
23. Kavli B, Otterlei M, Slupphaug G, Krokan HE. Uracil in DNA--general mutagen, but normal intermediate in acquired immunity. DNA Repair (Amst). 2007;6: 505–516. doi:10.1016/j.dnarep.2006.10.014
24. Sale JE. Translesion DNA synthesis and mutagenesis in eukaryotes. Cold Spring Harb Perspect Biol. 2013;5: a012708. doi:10.1101/cshperspect.a012708
25. Marians KJ. Lesion bypass and the reactivation of stalled replication forks. Annu Rev Biochem. 2018;87: 217–238. doi:10.1146/annurev-biochem-062917-011921
26. Kidane D, Murphy DL, Sweasy JB. Accumulation of abasic sites induces genomic instability in normal human gastric epithelial cells during Helicobacter pylori infection. Oncogenesis. 2014;3: e128. doi:10.1038/oncsis.2014.42
27. Green AM, Landry S, Budagyan K, Avgousti DC, Shalhout S, Bhagwat AS, et al. APOBEC3A damages the cellular genome during DNA replication. Cell

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cycle. 2016;15: 998–1008. doi:10.1080/15384101.2016.1152426

28. Sakofsky CJ, Roberts SA, Malc E, Mieczkowski PA, Resnick MA, Gordenin DA, et al. Break-induced replication is a source of mutation clusters underlying kataegis. Cell Rep. 2014;7: 1640–1648. doi:10.1016/j.celrep.2014.04.053
29. Cahill DP, Levine KK, Betensky RA, Codd PJ, Romany CA, Reavie LB, et al. Loss of the mismatch repair protein MSH6 in human glioblastomas is associated with tumor progression during temozolomide treatment. Clin Cancer Res. 2007;13: 2038–2045. doi:10.1158/1078-0432.CCR-06-2149
30. Keung MYT, Wu Y, Vadgama JV. PARP inhibitors as a therapeutic agent for homologous recombination deficiency in breast cancers. J Clin Med. 2019;8. doi:10.3390/jcm8040435
31. Cleary JM, Aguirre AJ, Shapiro GI, D'Andrea AD. Biomarker-Guided Development of DNA Repair Inhibitors. Mol Cell. 2020;78: 1070–1085. doi:10.1016/j.molcel.2020.04.035
32. Green AM, Budagyan K, Hayer KE, Reed MA, Savani MR, Wertheim GB, et al. Cytosine deaminase APOBEC3A sensitizes leukemia cells to inhibition of the DNA replication checkpoint. Cancer Res. 2017;77: 4579–4588. doi:10.1158/0008-5472.CAN-16-3394
33. Buisson R, Lawrence MS, Benes CH, Zou L. APOBEC3A and APOBEC3B activities render cancer cells susceptible to ATR inhibition. Cancer Res. 2017;77: 4567–4578. doi:10.1158/0008-5472.CAN-16-3389
34. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science. 2014;346: 251–256. doi:10.1126/science.1253462
35. Chen Z, Wen W, Bao J, Kuhs KL, Cai Q, Long J, et al. Integrative genomic analyses of APOBEC-mutational signature, expression and germline deletion of APOBEC3 genes, and immunogenicity in multiple cancer types. BMC Med Genomics. 2019;12: 131. doi:10.1186/s12920-019-0579-3
36. Wang S, Jia M, He Z, Liu X-S. APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. Oncogene. 2018;37: 3924–3936. doi:10.1038/s41388-018-0245-9
37. Hustedt N, Saito Y, Zimmermann M, Álvarez-Quilón A, Setiaputra D, Adam S, et al. Control of homologous recombination by the HROB-MCM8-MCM9 pathway. Genes Dev. 2019;33: 1397–1415. doi:10.1101/gad.329508.119
38. Huang J-W, Acharya A, Taglialatela A, Nambiar TS, Cuella-Martin R, Leuzzi G, et al. MCM8IP activates the MCM8-9 helicase to promote DNA synthesis and homologous recombination upon DNA damage. Nat Commun. 2020;11: 2948. doi:10.1038/s41467-020-16718-3
39. Griffin WC, Trakselis MA. The MCM8/9 complex: A recent recruit to the roster of helicases involved in genome maintenance. DNA Repair (Amst). 2019;76: 1–10. doi:10.1016/j.dnarep.2019.02.003
40. Ohashi E, Tsurimoto T. Functions of Multiple Clamp and Clamp-Loader Complexes in Eukaryotic DNA Replication. Adv Exp Med Biol. 2017;1042: 135–162. doi:10.1007/978-981-10-6955-0_7
41. Halabelian L, Ravichandran M, Li Y, Zeng H, Rao A, Aravind L, et al.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

964 Structural basis of HMCES interactions with abasic DNA and multivalent
965 substrate recognition. Nat Struct Mol Biol. 2019;26: 607–612.
966 doi:10.1038/s41594-019-0246-6
967 42. Mohni KN, Wessel SR, Zhao R, Wojciechowski AC, Luzwick JW, Layden H,
968 et al. HMCES Maintains Genome Integrity by Shielding Abasic Sites in
969 Single-Strand DNA. Cell. 2019;176: 144–153.e13.
970 doi:10.1016/j.cell.2018.10.055
971 43. Thompson PS, Amidon KM, Mohni KN, Cortez D, Eichman BF. Protection of
972 abasic sites during DNA replication by a stable thiazolidine protein-DNA
973 cross-link. Nat Struct Mol Biol. 2019;26: 613–618. doi:10.1038/s41594-019-
974 0255-5
975 44. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, et
976 al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized
977 derivatives. Cell. 2013;152: 1146–1159. doi:10.1016/j.cell.2013.02.004
978 45. Aravind L, Anand S, Iyer LM. Novel autoproteolytic and DNA-damage sensing
979 components in the bacterial SOS response and oxidized methylcytosine-
980 induced eukaryotic DNA demethylation systems. Biol Direct. 2013;8: 20.
981 doi:10.1186/1745-6150-8-20
982 46. Wang N, Bao H, Chen L, Liu Y, Li Y, Wu B, et al. Molecular basis of abasic
983 site sensing in single-stranded DNA by the SRAP domain of E. coli yedK.
984 Nucleic Acids Res. 2019;47: 10388–10399. doi:10.1093/nar/gkz744
985 47. Colic M, Wang G, Zimmermann M, Mascall K, McLaughlin M, Bertolet L, et al.
986 Identifying chemogenetic interactions from CRISPR screens with drugZ.
987 Genome Med. 2019;11: 52. doi:10.1186/s13073-019-0665-3
988 48. Shukla V, Halabelian L, Balagere S, Samaniego-Castruita D, Feldman DE,
989 Arrowsmith CH, et al. HMCES Functions in the Alternative End-Joining
990 Pathway of the DNA DSB Repair during Class Switch Recombination in B
991 Cells. Mol Cell. 2020;77: 384–394.e4. doi:10.1016/j.molcel.2019.10.031
992 49. Mehta KPM, Lovejoy CA, Zhao R, Heintzman DR, Cortez D. HMCES
993 Maintains Replication Fork Progression and Prevents Double-Strand Breaks
994 in Response to APOBEC Deamination and Abasic Site Formation. Cell Rep.
995 2020;31: 107705. doi:10.1016/j.celrep.2020.107705
996 50. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et
997 al. Optimized sgRNA design to maximize activity and minimize off-target
998 effects of CRISPR-Cas9. Nat Biotechnol. 2016;34: 184–191.
999 doi:10.1038/nbt.3437
1000 51. Li W, Köster J, Xu H, Chen C-H, Xiao T, Liu JS, et al. Quality control,
1001 modeling, and visualization of CRISPR screens with MAGeCK-VISPR.
1002 Genome Biol. 2015;16: 281. doi:10.1186/s13059-015-0843-6
1003 52. Brown KR, Mair B, Soste M, Moffat J. CRISPR screens are feasible in TP53
1004 wild-type cells. Mol Syst Biol. 2019;15: e8679. doi:10.15252/msb.20188679
1005 53. Park J, Long DT, Lee KY, Abbas T, Shibata E, Negishi M, et al. The MCM8-
1006 MCM9 complex promotes RAD51 recruitment at DNA damage sites to
1007 facilitate homologous recombination. Mol Cell Biol. 2013;33: 1632–1644.
1008 doi:10.1128/MCB.01503-12
1009 54. Zhao Y, Lang G, Ito S, Bonnet J, Metzger E, Sawatsubashi S, et al. A
1010 TFTC/STAGA module mediates histone H2A and H2B deubiquitination,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

coactivates nuclear receptors, and counteracts heterochromatin silencing. Mol Cell. 2008;29: 92–101. doi:10.1016/j.molcel.2007.12.011

55. Lang G, Bonnet J, Umlauf D, Karmodiya K, Koffler J, Stierle M, et al. The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements. Mol Cell Biol. 2011;31: 3734–3744. doi:10.1128/MCB.05231-11
56. Oughtred R, Stark C, Breitkreutz B-J, Rust J, Boucher L, Chang C, et al. The BioGRID interaction database: 2019 update. Nucleic Acids Res. 2019;47: D529–D541. doi:10.1093/nar/gky1079
57. Meyers RM, Bryan JG, McFarland JM, Weir BA, Sizemore AE, Xu H, et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. Nat Genet. 2017;49: 1779–1784. doi:10.1038/ng.3984
58. Dempster JM, Pacini C, Pantel S, Behan FM, Green T, Krill-Burger J, et al. Agreement between two large pan-cancer CRISPR-Cas9 gene dependency data sets. Nat Commun. 2019;10: 5817. doi:10.1038/s41467-019-13805-y
59. Olivieri M, Cho T, Álvarez-Quilón A, Li K, Schellenberg MJ, Zimmermann M, et al. A genetic map of the response to DNA damage in human cells. Cell. 2020;182: 481–496.e21. doi:10.1016/j.cell.2020.05.040
60. Dempster JM, Rossen J, Kazachkova M, Pan J, Kugener G, Root DE, et al. Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines. BioRxiv. 2019; doi:10.1101/720243
61. Ardeljan D, Steranka JP, Liu C, Li Z, Taylor MS, Payer LM, et al. Cell fitness screens reveal a conflict between LINE-1 retrotransposition and DNA replication. Nat Struct Mol Biol. 2020;27: 168–178. doi:10.1038/s41594-020-0372-1
62. Jarvis MC, Ebrahimi D, Temiz NA, Harris RS. Mutation signatures including APOBEC in cancer cell lines. JNCI Cancer Spectr. 2018;2. doi:10.1093/jncics/pky002
63. Petljak M, Alexandrov LB, Brammeld JS, Price S, Wedge DC, Grossmann S, et al. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. Cell. 2019;176: 1282–1294.e20. doi:10.1016/j.cell.2019.02.012
64. Liu Q, Yin X, Languino LR, Altieri DC. Evaluation of drug combination effect using a Bliss independence dose-response surface model. Stat Biopharm Res. 2018;10: 112–122. doi:10.1080/19466315.2018.1437071
65. Blackford AN, Jackson SP. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. Mol Cell. 2017;66: 801–817. doi:10.1016/j.molcel.2017.05.015
66. Chowdhury D, Xu X, Zhong X, Ahmed F, Zhong J, Liao J, et al. A PP4-phosphatase complex dephosphorylates gamma-H2AX generated during DNA replication. Mol Cell. 2008;31: 33–46. doi:10.1016/j.molcel.2008.05.016
67. Lebedeva NA, Rechkunova NI, El-Khamisy SF, Lavrik OI. Tyrosyl-DNA phosphodiesterase 1 initiates repair of apurinic/apyrimidinic sites. Biochimie. 2012;94: 1749–1753. doi:10.1016/j.biochi.2012.04.004
68. Álvarez-Quilón A, Wojtaszek JL, Mathieu M-C, Patel T, Appel CD, Hustedt N, et al. Endogenous DNA 3’ blocks are vulnerabilities for BRCA1 and BRCA2

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

deficiency and are reversed by the APE2 nuclease. Mol Cell. 2020;78: 1152–1165.e8. doi:10.1016/j.molcel.2020.05.021

69. McNeill DR, Whitaker AM, Stark WJ, Illuzzi JL, McKinnon PJ, Freudenthal BD, et al. Functions of the major abasic endonuclease (APE1) in cell viability and genotoxin resistance. Mutagenesis. 2020;35: 27–38. doi:10.1093/mutage/gez046
70. Huang J, Zhou Q, Gao M, Nowsheen S, Zhao F, Kim W, et al. Tandem Deubiquitination and Acetylation of SPRTN Promotes DNA-Protein Crosslink Repair and Protects against Aging. Mol Cell. 2020;79: 824–835.e5. doi:10.1016/j.molcel.2020.06.027
71. Halder S, Torrecilla I, Burkhalter MD, Popović M, Fielden J, Vaz B, et al. SPRTN protease and checkpoint kinase 1 cross-activation loop safeguards DNA replication. Nat Commun. 2019;10: 3142. doi:10.1038/s41467-019-11095-y
72. Kalev P, Simicek M, Vazquez I, Munck S, Chen L, Soin T, et al. Loss of PPP2R2A inhibits homologous recombination DNA repair and predicts tumor sensitivity to PARP inhibition. Cancer Res. 2012;72: 6414–6424. doi:10.1158/0008-5472.CAN-12-1667
73. Goodarzi AA, Jonnalagadda JC, Douglas P, Young D, Ye R, Moorhead GBG, et al. Autophosphorylation of ataxia-telangiectasia mutated is regulated by protein phosphatase 2A. EMBO J. 2004;23: 4451–4461. doi:10.1038/sj.emboj.7600455
74. Ruvolo PP. The broken “Off” switch in cancer signaling: PP2A as a regulator of tumorigenesis, drug resistance, and immune surveillance. BBA Clin. 2016;6: 87–99. doi:10.1016/j.bbacli.2016.08.002
75. Serebrenik AA, Starrett GJ, Leenen S, Jarvis MC, Shaban NM, Salamango DJ, et al. The deaminase APOBEC3B triggers the death of cells lacking uracil DNA glycosylase. Proc Natl Acad Sci USA. 2019;116: 22158–22163. doi:10.1073/pnas.1904024116
76. Srivastava M, Su D, Zhang H, Chen Z, Tang M, Nie L, et al. HMCES safeguards replication from oxidative stress and ensures error-free repair. EMBO Rep. 2020;21: e49123. doi:10.15252/embr.201949123
77. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, et al. Defining a cancer dependency map. Cell. 2017;170: 564–576.e16. doi:10.1016/j.cell.2017.06.010
78. Sherr CJ, Bartek J. Cell cycle?targeted cancer therapies. Annu Rev Cancer Biol. 2017;1: 41–57. doi:10.1146/annurev-cancerbio-040716-075628
79. Qiu Z, Fa P, Liu T, Prasad CB, Ma S, Hong Z, et al. A Genome-Wide Pooled shRNA Screen Identifies PPP2R2A as a Predictive Biomarker for the Response to ATR and CHK1 Inhibitors. Cancer Res. 2020;80: 3305–3318. doi:10.1158/0008-5472.CAN-20-0057
80. Kawale AS, Povirk LF. Tyrosyl-DNA phosphodiesterases: rescuing the genome from the risks of relaxation. Nucleic Acids Res. 2018;46: 520–537. doi:10.1093/nar/gkx1219
81. Wang B, Wang M, Zhang W, Xiao T, Chen C-H, Wu A, et al. Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. Nat Protoc. 2019;14: 756–780. doi:10.1038/s41596-018-0113-7

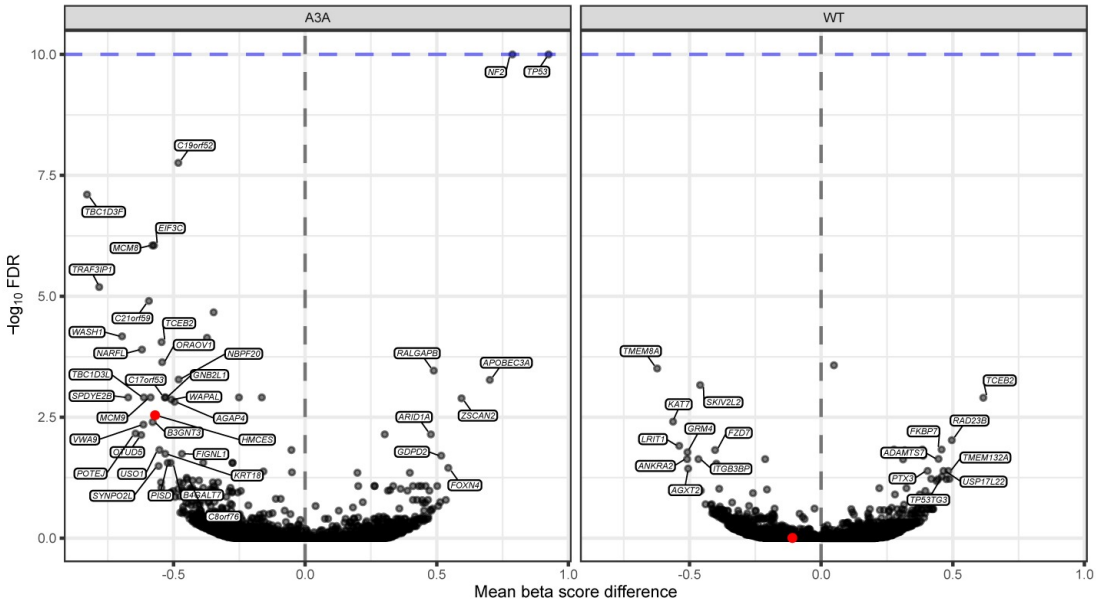
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

82. Hart T, Brown KR, Sircoulomb F, Rottapel R, Moffat J. Measuring error rates in genomic perturbation screens: gold standards for human functional genomics. Mol Syst Biol. 2014;10: 733. doi:10.15252/msb.20145216
83. Hart T, Tong AHY, Chan K, Van Leeuwen J, Seetharaman A, Aregger M, et al. Evaluation and Design of Genome-Wide CRISPR/SpCas9 Knockout Screens. G3 (Bethesda). 2017;7: 2719–2727. doi:10.1534/g3.117.041277
84. Spandidos A, Wang X, Wang H, Seed B. PrimerBank: a resource of human and mouse PCR primer pairs for gene expression detection and quantification. Nucleic Acids Res. 2010;38: D792-9. doi:10.1093/nar/gkp1005

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supporting Information

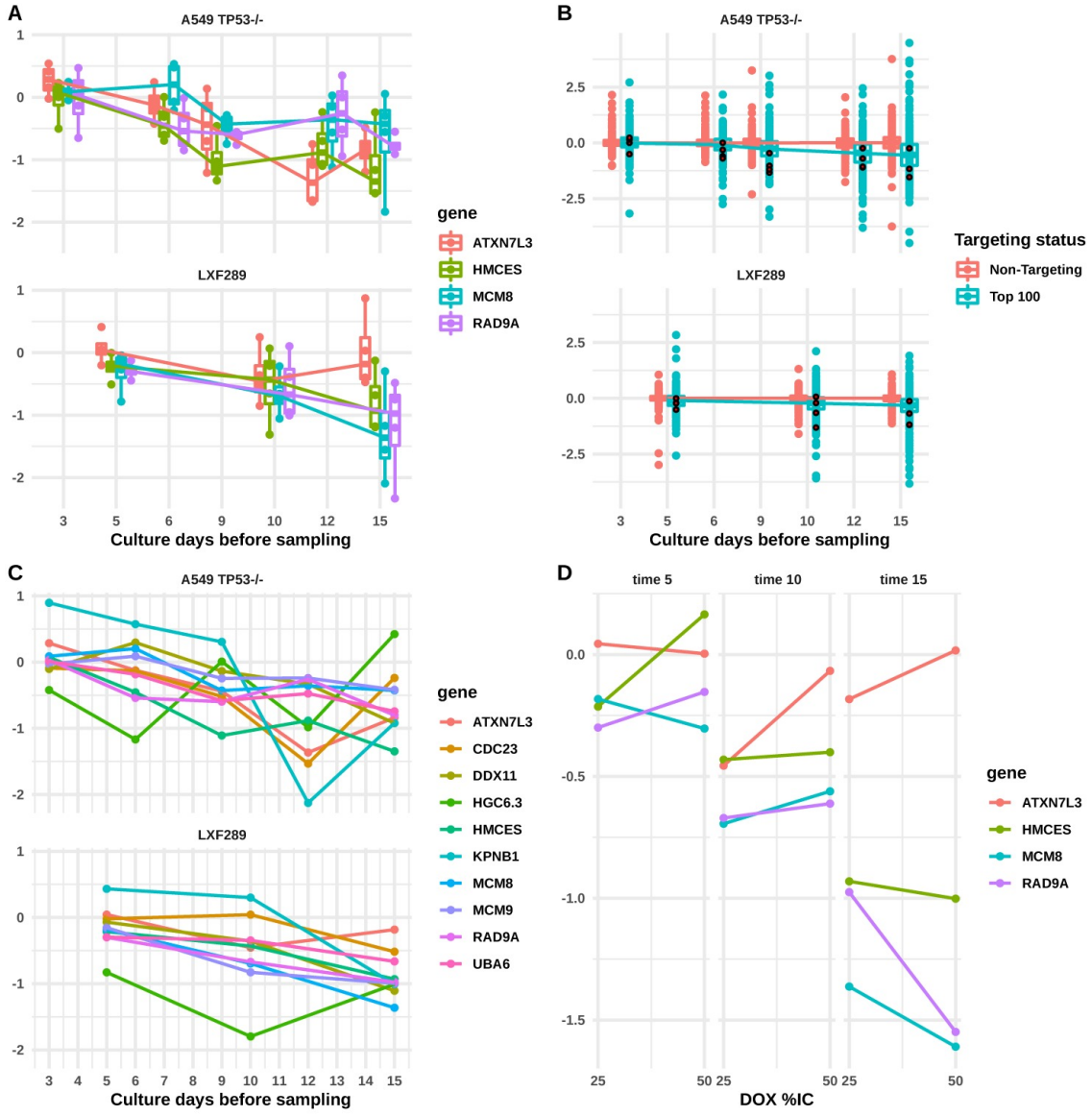
S1 Fig. Validation of the A549 ± A3A TP53-/- clones. Western blot of the A549 p53-/- (left panel) and the A549 A3A transduced p53-/- clones (right panel) generated using CRISPR/Cas9 targeting.



S2 Fig. Effects of DOX treatment. Volcano plot highlighting genes for which there could exist interaction with DOX *per se*. The x-axis represents the difference of the

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

(normalized) MAGeCK-MLE's beta score between treating a sample with DOX (IC_{25}) and the corresponding control sample, averaged across the A3A plasmid-free version of all cell lines sampled after 15 days of cell culture. Intuitively, genes with significant negative beta score differences suggest conditional essentiality with DOX. The y-axis represents the -$\log_{10}$ FDR of the Fisher's combined p-value (either lower or upper tail) across samples.



S3 Fig. Change in sgRNA count LFC dependent on days of cell culture before

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

sampling or DOX dose. LFC (y-axes) represents the cell count differences between a sample treated with DOX (IC_{25} in plots a, b, and c) and the corresponding control (untreated) sample. (**A**) The top four genes are shown after sorting based on the overall score. The four sgRNAs targeting each gene are shown separately, and their count distribution is represented as a boxplot. Lines join the median sgRNA counts for each gene. One of the top-5 genes, HGC6.3, was excluded from the plot due to low data quality (see S1 Table). (**B**) The sgRNAs shown are the 1000 non-targeting control sgRNAs in the Brunello library[50], and the top 100 genes after sorting genes based on the overall score. Lines join the median sgRNA counts for each distribution. Red dots indicate the LFC of sgRNAs for the HMCES gene. (**C**) The top ten genes are shown after sorting by the overall score. Here, the HGC6.3 gene is included, while according to MAGeCK-MLE it had low sgRNA efficiency (S1 Table), possibly causing the rather erratic trends across time points that it exhibits. (**D**) Difference in LFC dependent on DOX dose, either IC_{25} or IC_{50}, in the LXF289 cell line. Columns show LFCs at different sampling times. The top four genes by overall score are shown.

Gene	sgRNA	Cell line	Efficiency		Gene	sgRNA	Cell line	Efficiency
	s_75152	A549 TP53-/-	1			s_42031	A549 TP53-/-	1
		LXF289	0.727				LXF289	1
	s_75153	A549 TP53-/-	1			s_42032	A549 TP53-/-	1
HGC6.3		LXF289	0.596		*UBA6*		LXF289	1
	s_75154	A549 TP53-/-	0.977			s_42033	A549 TP53-/-	1
		LXF289	0.602				LXF289	1
	s_75155	A549 TP53-/-	1			s_42034	A549 TP53-/-	1

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

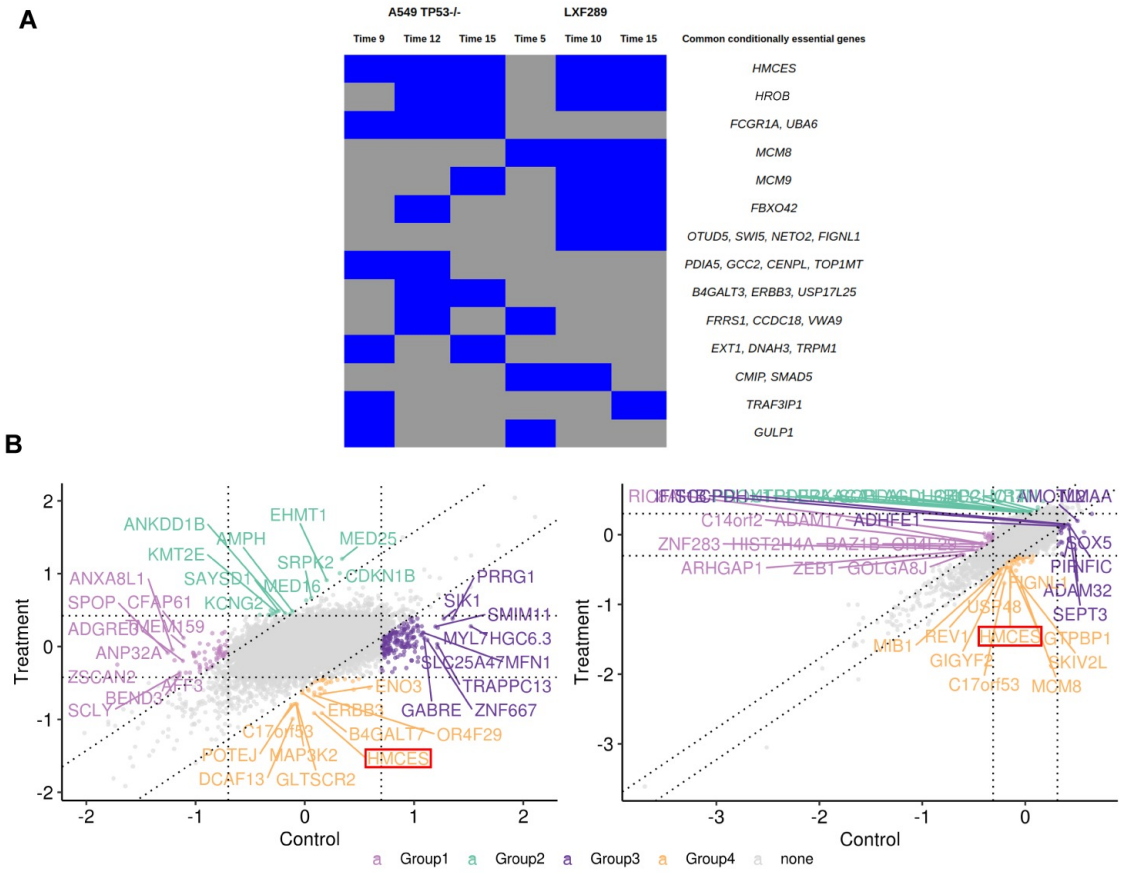
		LXF289	0.644				LXF289	1
	s_44647	A549 TP53-/-	1			s_4665	A549 TP53-/-	1
		LXF289	1				LXF289	1
	s_44648	A549 TP53-/-	1			s_4666	A549 TP53-/-	0.975
HMCES		LXF289	1		*DDX11*		LXF289	1
	s_44649	A549 TP53-/-	1			s_4667	A549 TP53-/-	0.713
		LXF289	0.992				LXF289	0.999
	s_44650	A549 TP53-/-	1			s_4668	A549 TP53-/-	0.968
		LXF289	1				LXF289	1
	s_15717	A549 TP53-/-	1			s_66465	A549 TP53-/-	0.996
		LXF289	1				LXF289	0.973
	s_15718	A549 TP53-/-	1			s_66466	A549 TP53-/-	0.998
RAD9A		LXF289	1		*MCM9*		LXF289	1
	s_15719	A549 TP53-/-	1			s_66467	A549 TP53-/-	0.79
		LXF289	1				LXF289	1
	s_15720	A549 TP53-/-	1			s_66468	A549 TP53-/-	1
		LXF289	1				LXF289	1
	s_53711	A549 TP53-/-	1			s_22709	A549 TP53-/-	1
		LXF289	1				LXF289	1
	s_53712	A549 TP53-/-	0.956			s_22710	A549 TP53-/-	1
MCM8		LXF289	0.999		*CDC23*		LXF289	1
	s_53713	A549 TP53-/-	0.937			s_22711	A549 TP53-/-	1
		LXF289	1				LXF289	1.22E-13
	s_53714	A549 TP53-/-	0.976			s_22712	A549 TP53-/-	1
		LXF289	1				LXF289	1
	s_44735	A549 TP53-/-	1			s_28149	A549 TP53-/-	1
		LXF289	1				LXF289	1
	s_44736	A549 TP53-/-	1			s_28150	A549 TP53-/-	1
ATXN7L3		LXF289	1		*MAD2L2*		LXF289	1
	s_44737	A549 TP53-/-	1			s_28151	A549 TP53-/-	1

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

		LXF289	1				LXF289	1
	s_44738	A549 TP53-/-	1			s_28152	A549 TP53-/-	1
		LXF289	1				LXF289	1
	s_10269	A549 TP53-/-	1					
		LXF289	1					
	s_10270	A549 TP53-/-	1					
KPNB1		LXF289	1					
	s_10271	A549 TP53-/-	1					
		LXF289	1					
	s_10272	A549 TP53-/-	1					
		LXF289	1					

S1 Table. Estimated sgRNA efficiencies for the top 11 genes. Efficiencies of sgRNAs were estimated by the MAGeCK-MLE algorithm in A549$^{TP53-/-}$ and LXF-289 cell lines. Informally, the efficiencies estimate the probability that a given sgRNA is able to generate an inactivating double-strand DNA break in the targeted gene. Only the *HGC6.3* gene shows overall low sgRNA efficiencies, particularly in the LXF-289 cell line, suggesting that the high log-fold change scores therein are an artefact (S3 Fig). *HMCES* has near-perfect sgRNA efficiencies.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

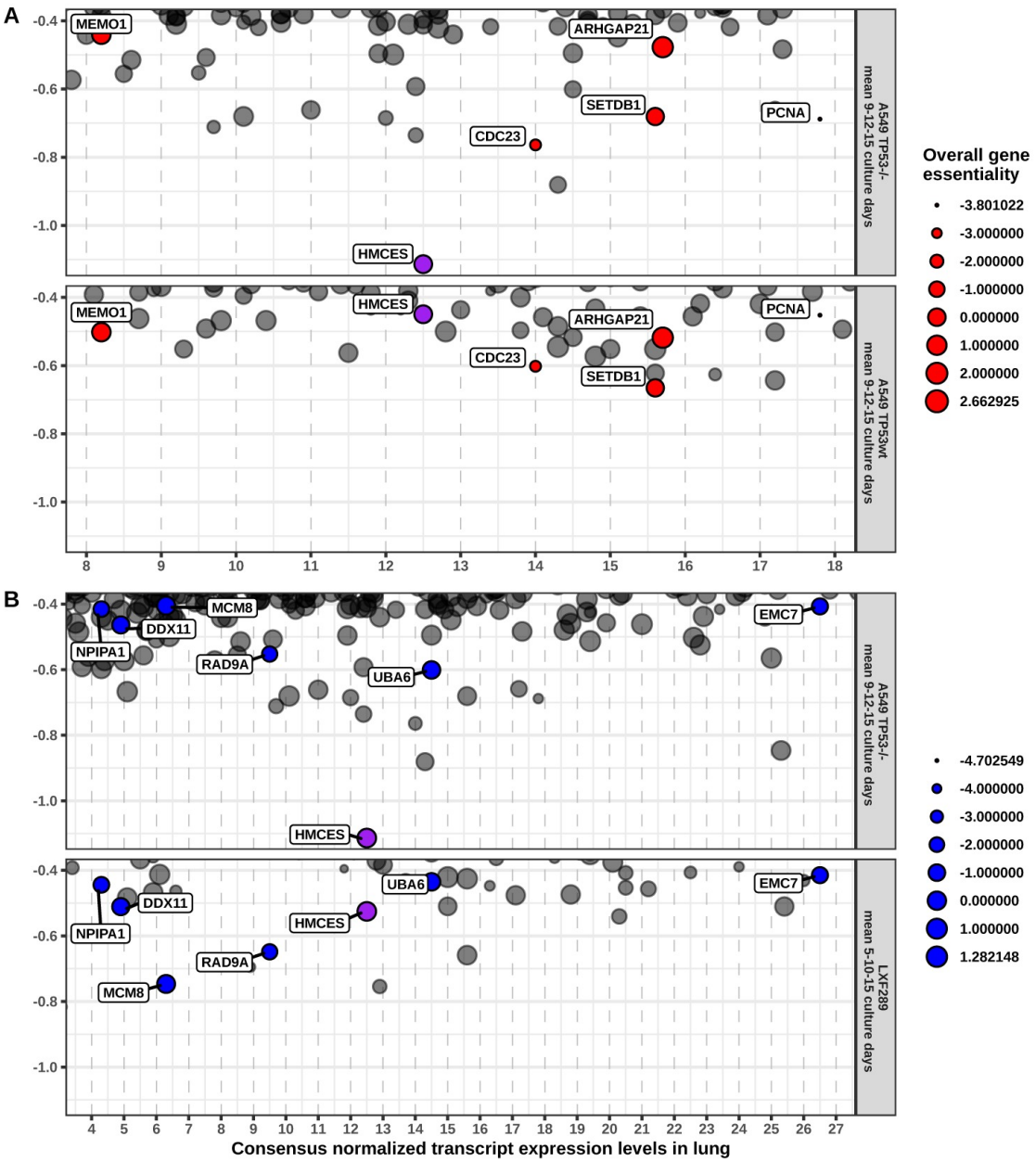


S4 Fig. Analysis of genetic screening data using an additional statistical methodology (MAGeCK-MLE). (**A**) Out of a total of 339 genes identified as conditionally essential by MAGeCK-MLE in either of the two cell lines examined (see next point), this figure shows those genes that were significant in more than one sample (time point / cell line combination). A blue box indicates that the genes in that row are conditionally essential (under A3A overexpression) in the corresponding sample, while a gray box indicates that there is no significant essentiality. HMCES is the gene found to be conditionally essential in the highest number samples -- all samples except the earliest time point of LXF-289, time 5. (**B**) MAGeCK-MLE visualizations ('nine-square plots' as in[81]) based on (left) A549 TP53-/- and (right) LXF-289 cell lines, both sampled at day 15. Points represent genes distributed

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

1184 according to the between-samples normalized beta scores (enrichments) for the
1185 control sample (untreated, x-axis) and DOX-treated sample (at IC_{25} concentration,
1186 y-axis). Vertical and horizontal dotted lines indicate two standard deviations of the
1187 beta score distribution away from zero to each side. Analogously, diagonal dotted
1188 lines represent two standard deviations of the distribution of between-treatment beta
1189 score differences away from zero to each side. Therefore, genes located in the
1190 bottom center square ("Group4" genes) have MAGeCK beta scores different
1191 between the control and treated sample, being not different from zero in the control
1192 (i.e. no evidence of selection) but significantly negative in the treatment (i.e.
1193 negatively selected); in other words, these genes are conditionally selected under
1194 A3A overexpression (DOX-induced). The top 10 genes are labeled in each square.
1195 HMCES is a top hit in both cell lines.
1196

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



S5 Fig. Contrasts of A3A-conditionally essential genes between different genetic backgrounds. (**A**) Contrast between *TP53* backgrounds of the A549 cell line. Red circles are genes with a consistently strong negative LFC (below -0.4) in both *TP53* backgrounds (-/- above, wild-type below), considering the mean LFC after 9, 12, and 15 culture days (y-axes); LFC represents the cell count differences

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

between a sample treated with DOX (IC_{25}) and the corresponding control (untreated) sample. The circle area shows the beta score (enrichment) calculated with MAGeCK-MLE and averaged across the time points: a more negative beta score indicates stronger gene essentiality irrespective of treatment. X-axes represent the Human Protein Atlas consensus normalized (across cell lines) transcript expression levels (NX) in lung tissue for each gene. Among the hits, *HMCES* is prominent in the TP53-/- background but not in the *wild-type* background, has moderate expression levels in lung tissue, and does not appear to be generally strongly essential. (**B**) In an analogous manner, this plot shows the contrast of A3A-conditionally essential genes between the $A549^{TP53-/-}$ and the LXF-289 genetic backgrounds; here, blue circles are genes with a consistently strong negative LFC (below -0.4) in both cell lines. Among the hits, *HMCES*, *RAD9A* and, to some extent, *MCM8* appear consistent in both backgrounds; of these three genes, *HMCES* has somewhat higher expression levels in lung tissue, and is the least essential in these cell lines. *HMCES* is the only hit that is consistent in both comparisons, and this is noted by using a purple color to highlight it.

See excel file.

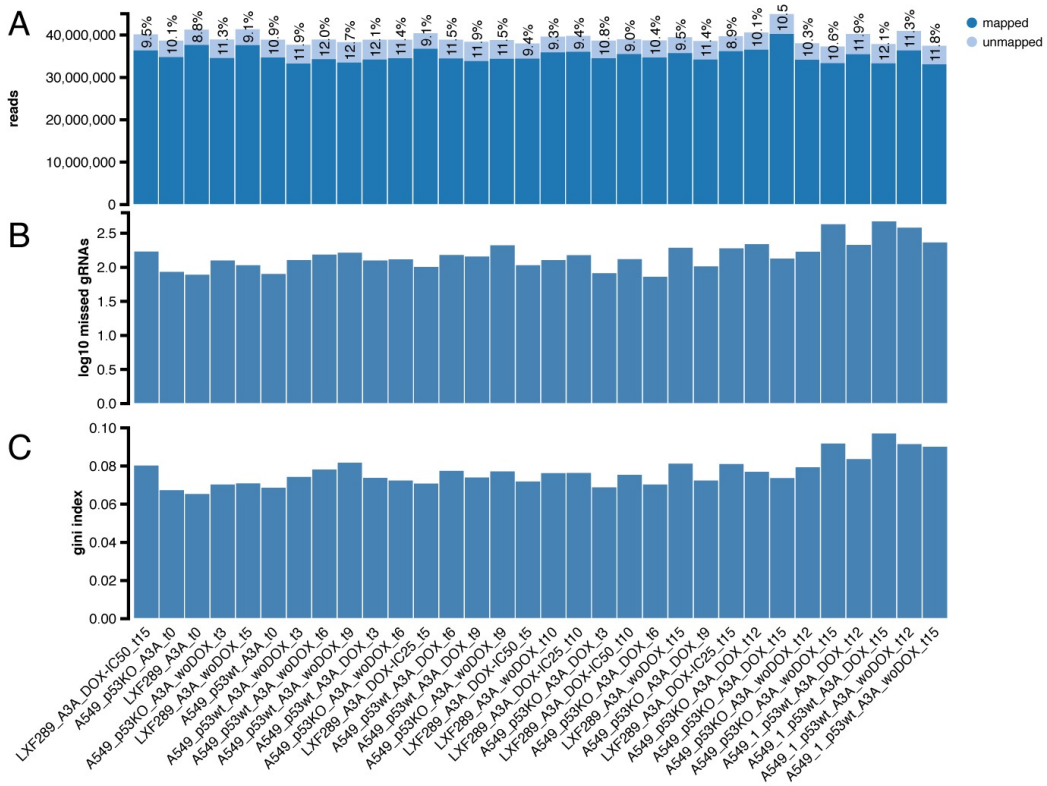
S2 Table. Gene-level data for all primary screens.

See excel file.

S3 Table. Enriched GO Biological Process terms obtained from GOrilla. GOrilla (http://cbl-gorilla.cs.technion.ac.il/) was used to analyze gene ontology using a P-

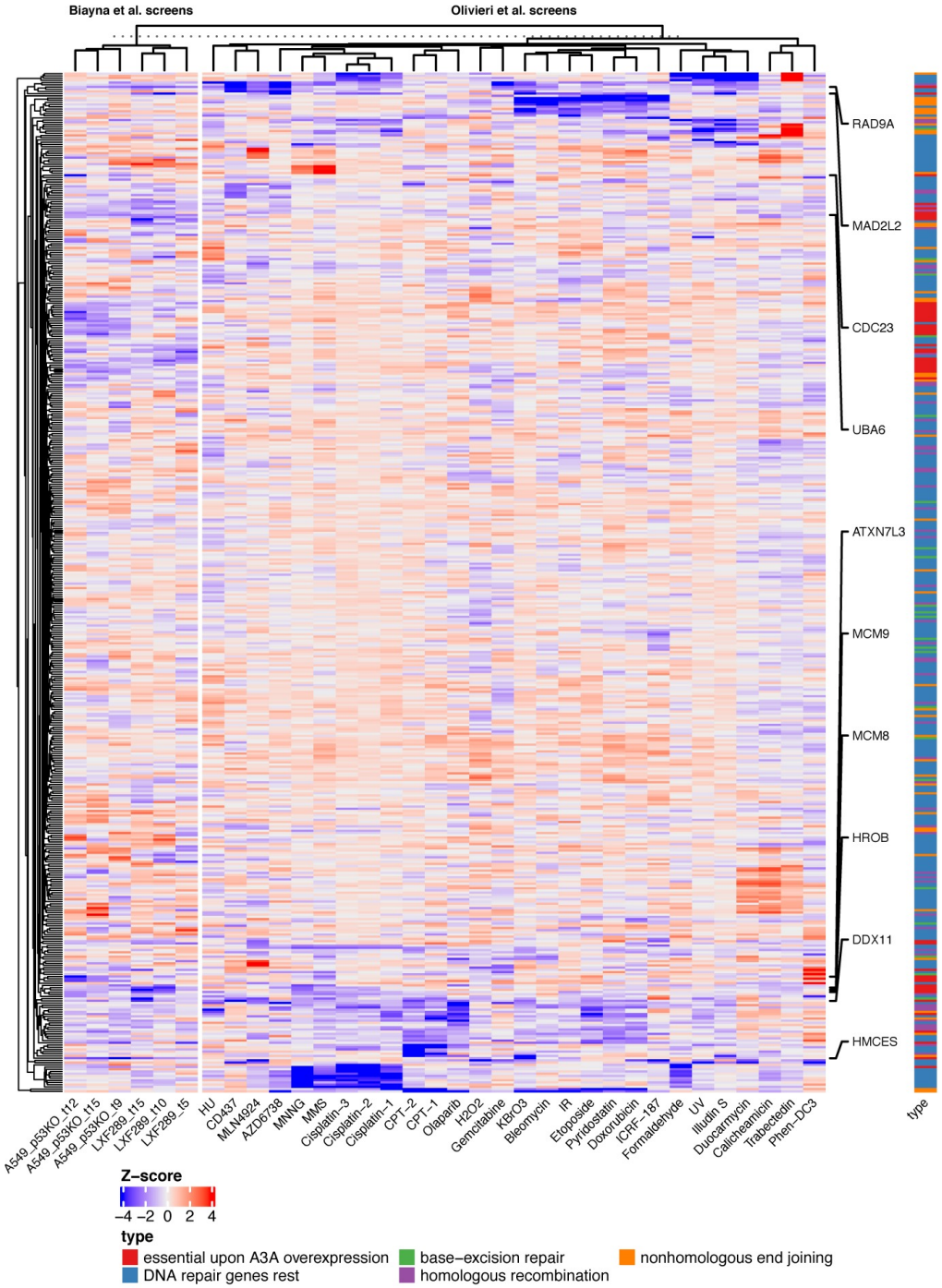
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

value threshold set at 10^{-3}. Samples included are A549$^{TP53-/-}$ at time points 9, 12, and 15 days, and LXF-289 at time points 5, 10, and 15 days.



S6 Fig. Quality control of sequencing reads. (**A**) Number of total sequenced reads per sample. Light blue fraction represents the percentage of reads that are unequivocally unmapped to the library, which is below the recommended maximum of 35% in all samples[51]. (**B**) Number of library sgRNAs that have zero counts per sample. Figures are higher in late samples, but this is to be expected due to negative selection. Overall, sgRNAs with zero counts are <1% of total sgRNAs in Brunello library (~77K)[50,51]. Namely, the maximum number of sgRNAs is 461. (**C**) Gini index of log-scaled read count distributions. This measure of the evenness across all sgRNA counts is below the recommended maximum of 0.2 in all samples[51]. Also, the Gini index is expected to increase in later time points.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



1239

1240 **S7 Fig. Many APOBEC-sensitizing genes, but not HMCES, also sensitize to a**

1241 **variety of other DNA damaging agents**. Left panel of heatmap shows a gene-level

1242 normalized log2 fold change (gene essentiality score) upon A3A overexpression for

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

two cell lines and for three time points (Biayna et al. screens); right panel shows Z-scores of gene essentiality after genotoxin exposure (Olivieri et al. screens)[59]. Data for 50 genes that are essential upon A3A overexpression in our screens (i.e. genes with the most negative mean log2 fold change across six data points) (labelled "top"), and 521 DNA repair genes. Labels on the right-hand side highlight the ten genes showing the highest overall A3A essentiality.

Cell line	**TP53**	**treatment**	**t9**	**t12**	**t15**	
	KO	**Control**	0.799	0.831	0.835	
		DOX-IC25	0.793	0.810	0.829	
A549 A3A						
	wt	**Control**	0.831	0.847	0.859	
		DOX-IC25	0.727	0.832	0.837	
			t5	**t10**	**t15**	
		Control	0.792	0.857	0.887	
LXF289 A3A	**mut**	**DOX-IC25**	0.789	0.854	0.889	
		DOX-IC50	0.779	0.847	0.885	

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

			t9	t12	t15	t18
	KO					0.881
RPE1[52]						
	wt		0.850	0.877	0.881	0.895

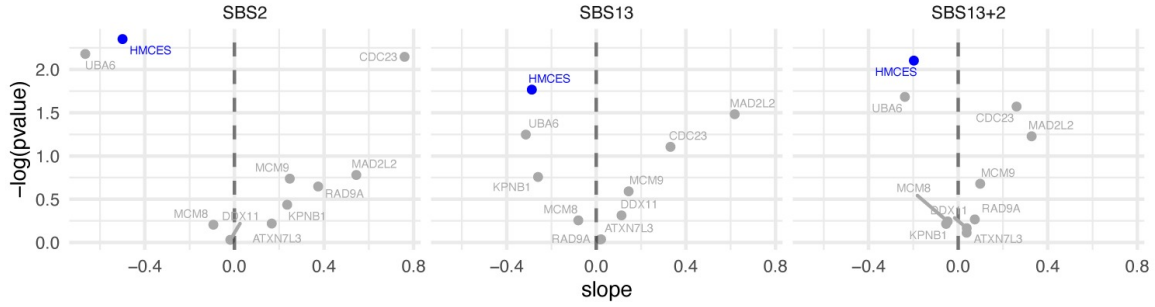
S4 Table. Area under the receiving operating characteristic curves (AUC) of each sample. AUC per sample, based on the capacity of the CRISPR screening to discriminate between known sets of essential[81–83] and non-essential[82] genes by their normalized read counts. For comparison, the AUCs for the same overall sets of genes in the genetic screens (RPE1 cell line) from Brown et al. 2019[52] have been included: note that, while our screening was based on the Brunello library[50], Brown et al. employed the TKO library, so the gene overlap is not total.

gene	signature	Regression coefficient	p-value (two-tailed)	pvalue (one-tailed, lower i.e. signature sensitizes to gene k.o.)
HMCES	SBS2	-0.499	0.095	0.048
UBA6	SBS2	-0.666	0.113	0.057
HMCES	SBS13+2	-0.197	0.122	0.061
HMCES	SBS13	-0.288	0.171	0.085
UBA6	SBS13+2	-0.238	0.186	0.093
UBA6	SBS13	-0.314	0.287	0.144
KPNB1	SBS13	-0.260	0.469	0.234
MCM8	SBS13	-0.080	0.774	0.387
MCM8	SBS13+2	-0.047	0.783	0.392
KPNB1	SBS13+2	-0.054	0.806	0.403
MCM8	SBS2	-0.094	0.814	0.407
DDX11	SBS2	-0.018	0.969	0.485
RAD9A	SBS13	0.018	0.965	0.517

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

ATXN7L3	SBS13	0.021	0.963	0.518
ATXN7L3	SBS13+2	0.038	0.893	0.553
DDX11	SBS13+2	0.038	0.847	0.576
ATXN7L3	SBS2	0.166	0.803	0.598
RAD9A	SBS13+2	0.075	0.766	0.617
DDX11	SBS13	0.113	0.731	0.634
KPNB1	SBS2	0.236	0.646	0.677
MCM9	SBS13	0.144	0.554	0.723
RAD9A	SBS2	0.374	0.524	0.738
MCM9	SBS13+2	0.099	0.507	0.746
MCM9	SBS2	0.247	0.478	0.761
MAD2L2	SBS2	0.544	0.458	0.771
CDC23	SBS13	0.331	0.331	0.834
MAD2L2	SBS13+2	0.328	0.293	0.854
MAD2L2	SBS13	0.618	0.227	0.886
CDC23	SBS13+2	0.261	0.208	0.896
CDC23	SBS2	0.760	0.117	0.942

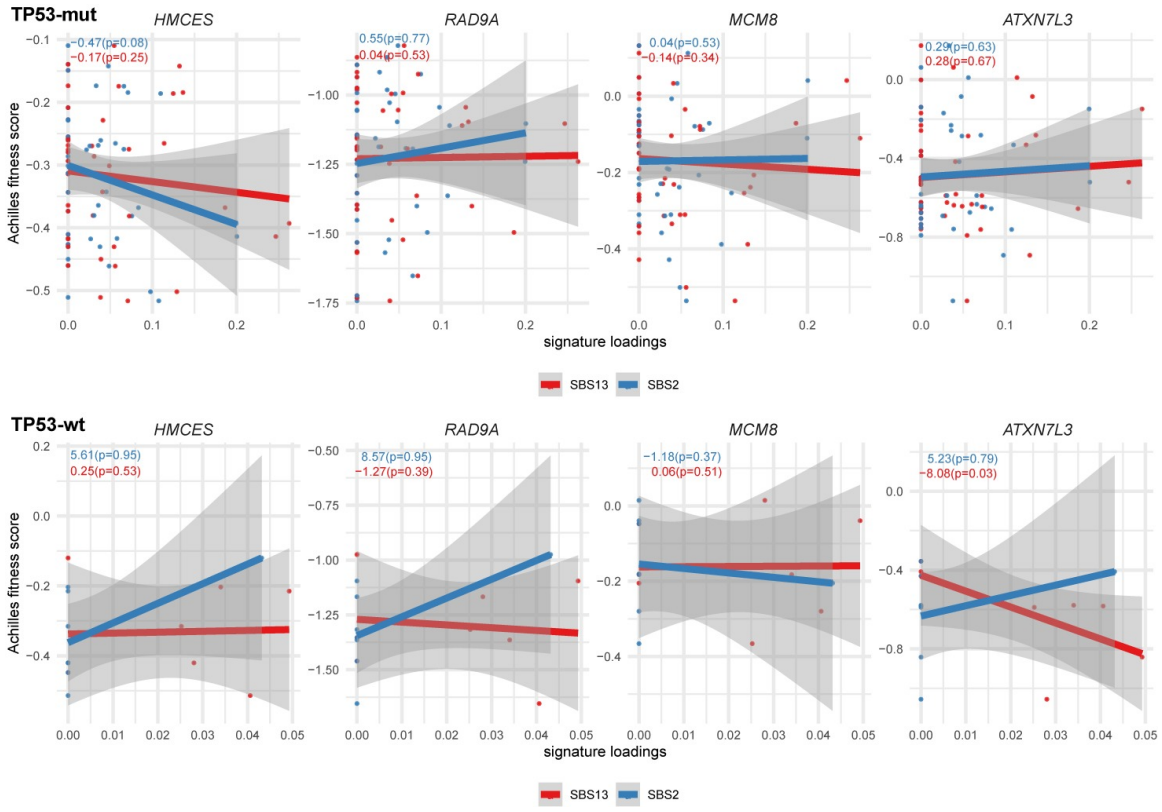
S5 Table. Table of differential fitness scores. Differential fitness scores (from project Achilles) upon APOBEC mutational signatures (SBS2, SBS13 and SBS13+2) burden for the top 10 genes that are essential upon A3A overexpression in our screens (i.e. genes with the most negative mean log2 fold change across six data points).



S8 Fig. Differential fitness scores. Differential fitness score (from project Achilles) upon APOBEC mutational signatures burden for the top 10 genes that are essential

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

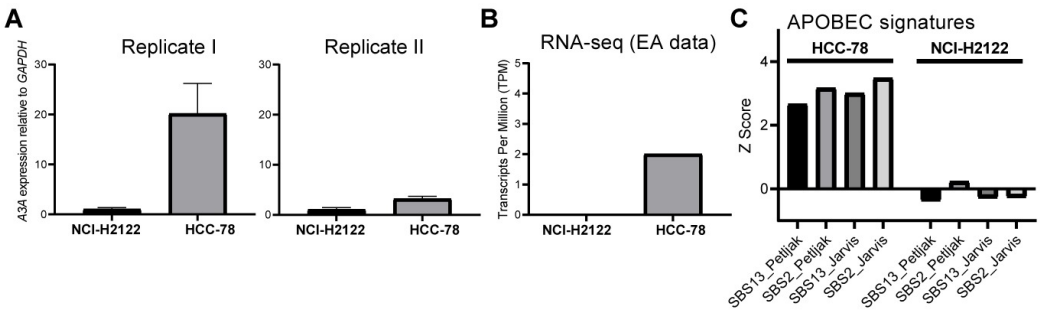
upon A3A overexpression in our screens (i.e. genes with the most negative mean log2 fold change across six data points).



S9 Fig. Gene essentiality fitness score from project Achilles versus APOBEC mutational signatures exposures. Cell lines originating from head-and-neck squamous cell carcinoma, lung adenocarcinoma, and lung squamous cell carcinoma were analyzed for the four genes with the greatest overall score in our genetic screens, while examining TP53 mutated (mut) and TP53 wild-type (wt) cell lines separately. The slope and p-value (one-tailed, lower) for the regression model for both APOBEC mutational signatures are shown within each panel. The more negative the slope, the more sensitive the cell lines are to the depletion of the particular gene at a higher level of the APOBEC mutational signature.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

1286

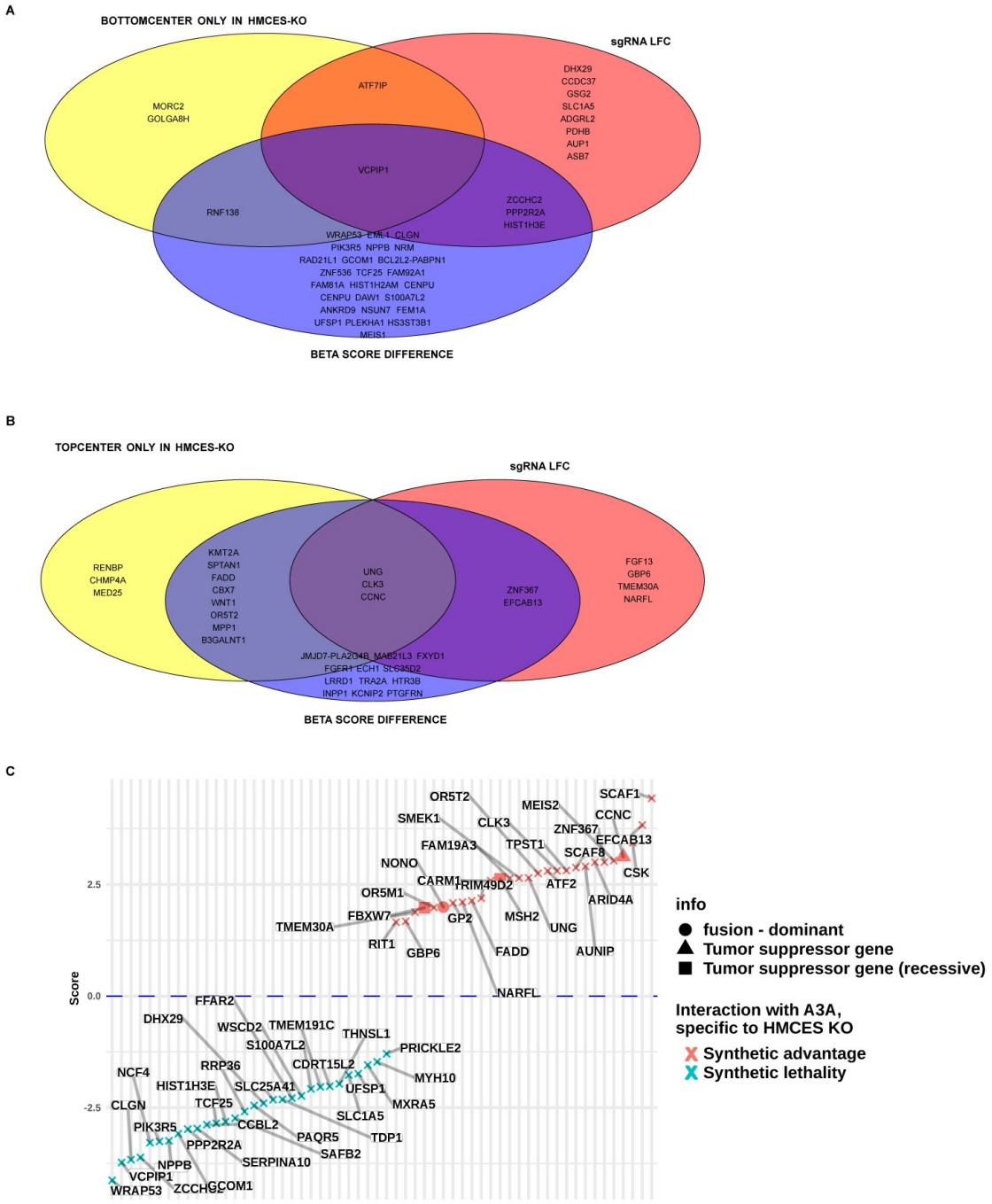


1287

1288 **S10 Fig. Endogenous expression and A3-mutational signature status of cell**
1289 **lines.** (**A**) Endogenous *A3A* mRNA expression levels in HCC-78 and NCI-H2122
1290 cells relative to *GAPDH* measured by quantitative real-time PCR (two independent
1291 biological replicates). (**B**) A3A gene expression (TPMs) downloaded from expression
1292 atlas (EA; https://www.ebi.ac.uk/gxa/home) and (**C**) APOBEC mutational
1293 signatures (SBS2 and SBS13) burden downloaded from Petljak et al. and Jarvis et.
1294 al and normalized across cell lines (z-score)[62,63].

1295

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



S11 Fig. Genes in epistasis with A3A expression in HMCES KO cells. (**A**, **B**): Venn diagrams containing genes that are in epistasis with A3A expression (panel **A**, synthetic sickness/lethality, panel **B**, synthetic advantage) exclusively within an HMCES KO background, when applying three complementary statistical methodologies. Genes in the

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

red circle have a standardized sgRNA LFC <-2 (**A**) or >2 (**B**) in the four DOX *versus* control comparisons (IC_{25}-t12, IC_{50}-t12, IC_{25}-t17, and IC_{50}-t17) exclusively in HMCES KO samples. Genes in the blue circle fulfill the same criteria but using the MAGeCK-MLE standardized beta score difference, instead of the LFC. Lastly, the yellow circle contains genes whose normalized beta scores are not different from 0 in the control sample while they are significantly different from 0 (**A**, lower; **B**, higher) in the A3A-expressing sample, exclusively in an HMCES KO background: specifically, this corresponds to the "bottom-center" (A) or "top-center" (B) square of MAGeCK-FLUTE's nine-square scatterplot visualization (see panel B of S4 Fig). (**C**) Genes shown in panels **A** and **B**, sorted by a score calculated as the mean of the standardized sgRNA LFCs and standardized MLE beta score differences from the four DOX vs. control comparisons of HMCES KO samples, minus the mean obtained in the same way for the HMCES wt samples. Therefore, a negative score (in blue) suggests that the gene could be synthetic lethal with A3A expression in HMCES KO but not in HMCES wt (e.g. *PPP2R2A*), and consequently a positive score (red) suggests that the gene could have synthetic advantage with A3A expression in HMCES KO but not in HMCES wt (e.g. *UNG*). Triangles indicate known tumor supressor genes in Cancer Gene Census, while a circles indicate a known oncogene.

See excel file.

S6 Table. Gene-level data for the secondary genetic screen in HMCES-/- cells.

See excel file.

S7 Table. GO enrichment analysis of the secondary screening data.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.05.429803; this version posted February 6, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Taqman Probes	
Gen Name	Probe ID
HMCES	Hs99999905_m1
GAPDH	Hs99999905_m1
Oligonucleotides	
Gene Name	Sequence (5'-3') / Sets
APOBEC3A Forward	TGGCATTGGAAGGCATAAGAC*
APOBEC3A Reverse	TTAGCCTGGTTGTGTAGAAAGC*
GAPDH Forward	AGCCACATCGCTCAGACAC
GAPDH Reverse	GCCCAATACGACCAAATCC
PCR1 Forward	AATGATACGGCGACCACCGAGATCTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACG
PCR1 Reverse	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCAATTCCCACTCCTTTCAAGACCT
Illumina Forward (PCR2)	AATGATACGGCGACCACCGAGATCT
NEBNext Multiplex Oligos (Reverse) (PCR2)	E7335S (Set 1), E7500S (Set 2), E7710S (Set 3) i E7730S (Set 4)
Custom Sequencing Primer	CGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCG

S8 Table. List of qRT-PCR primers (Taqman/Oligonucleotides) and PCR primers for library amplification and NGS. *Primer sequence obtained from PrimerBank[84].