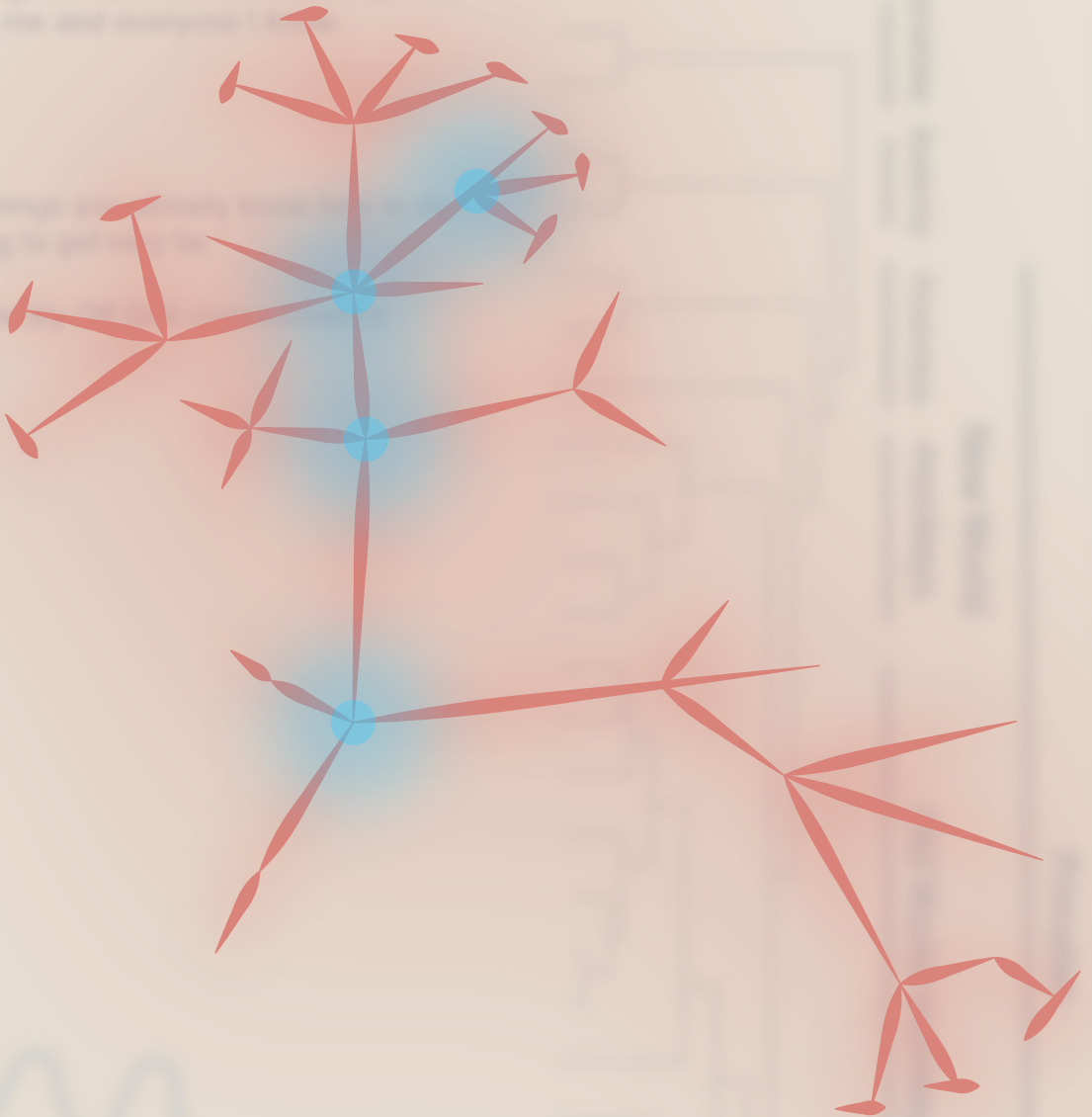


# APOBEC3B:

Regulation of an Oncogene in Sickness and in Health



Pieter Roelofs

**RADBOD  
UNIVERSITY  
PRESS**

Radboud  
Dissertation  
Series

# **APOBEC3B:**

Regulation of an Oncogene in Sickness and in Health

Pieter Roelofs

This publication has been made possible by KWF Kankerbestrijding (grant KWF10270) and René Vogels Stichting (travel grant)

**APOBEC3B: Regulation of an Oncogene in Sickness and in Health**

Pieter Roelofs

**Radboud Dissertation Series**

ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS

Postbus 9100, 6500 HA Nijmegen, The Netherlands

[www.radbouduniversitypress.nl](http://www.radbouduniversitypress.nl)

Design: Proefschrift AIO | Annelies Lips

Cover: Pieter Roelofs, derived from a drawing by Charles Darwin

Printing: DPN Rikken/Pumbo

ISBN: 9789465150628

DOI: 10.54195/9789465150628

Free download at: <https://doi.org/10.54195/9789465150628>

© 2025 Petrus Adriaan Roelofs

**RADBOUD  
UNIVERSITY  
PRESS**

This is an Open Access book published under the terms of Creative Commons Attribution-Noncommercial-NoDerivatives International license (CC BY-NC-ND 4.0). This license allows reusers to copy and distribute the material in any medium or format in unadapted form only, for noncommercial purposes only, and only so long as attribution is given to the creator, see <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

# **APOBEC3B:**

Regulation of an Oncogene in Sickness and in Health

Proefschrift

ter verkrijging van de graad van doctor  
aan de Radboud Universiteit Nijmegen  
op gezag van de rector magnificus prof. dr. J.M. Sanders,  
volgens besluit van het college voor promoties  
in het openbaar te verdedigen op  
maandag 24 maart 2025  
om 14:30 uur precies

door

**Petrus Adriaan Roelofs**

geboren op 26 december 1990  
te Alphen aan den Rijn

**Promotoren**

Dr. P.N. Span

Prof. dr. R.S. Harris (University of Minnesota, Verenigde Staten)

Prof. dr. J.W.M. Martens (Erasmus Universiteit Rotterdam)

**Manuscriptcommissie**

Prof. dr. M.J.L. Ligtenberg

Prof. dr. M.A.T.M. van Vugt (Universitair Medisch Centrum Groningen)

Dr. G.M.H.E. Dackus

# **APOBEC3B:**

Regulation of an Oncogene in Sickness and in Health

Dissertation

to obtain the degree of doctor

from Radboud University Nijmegen

on the authority of the Rector Magnificus prof. dr. J.M. Sanders,

according to the decision of the Doctorate Board

to be defended in public on

Monday, March 24, 2025

at 2:30 pm

by

**Petrus Adriaan Roelofs**

born on December 26, 1990

in Alphen aan den Rijn (the Netherlands)

**Supervisors**

Dr. P.N. Span

Prof. dr. R.S. Harris (University of Minnesota, United States)

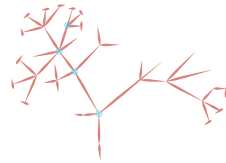
Prof. dr. J.W.M. Martens (Erasmus University Rotterdam)

**Manuscript Committee**

Prof. dr. M.J.L. Ligtenberg

Prof. dr. M.A.T.M. van Vugt (University Medical Center Groningen)

Dr. G.M.H.E. Dackus



This thesis is dedicated to my parents,  
whose flair, persistence, and delightful humbleness have inspired,  
and continue to inspire, me and everyone I know.

*"If you're only ever going to do things you already know how to do,  
you're not going to get very far."*

Nick Fouch, Fouch Family Off Grid (non-verbatim)





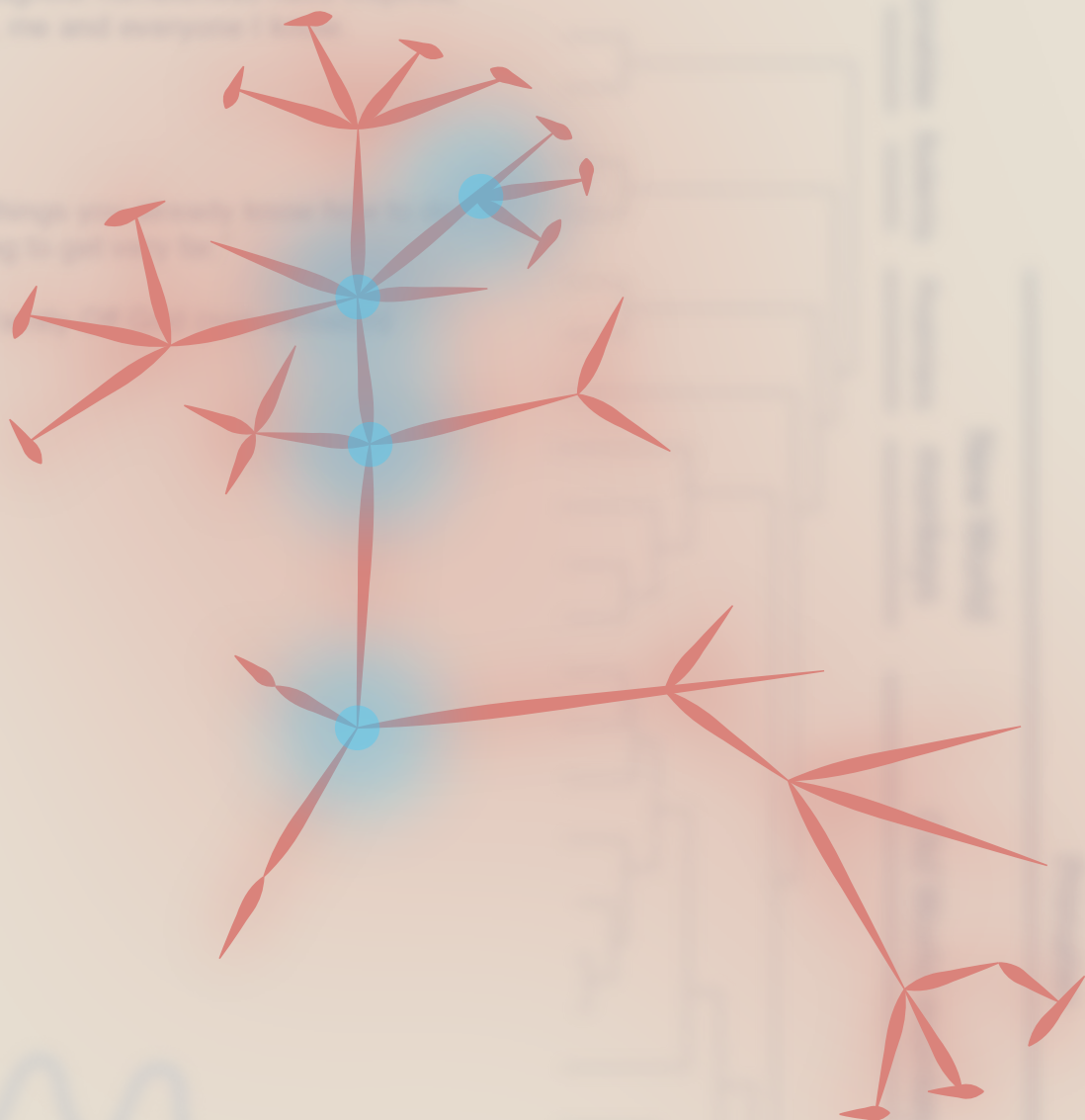
# Contents

<b>Chapter 1</b>	A Short Introduction and Thesis Outline	11
<b>Chapter 2</b>	Polyomavirus T Antigen Induces <i>APOBEC3B</i> Expression Using an LXCXE-Dependent and TP53-Independent Mechanism (published, <i>mBio</i> , 2019)	17
<b>Chapter 3</b>	Characterization of the Mechanism by which the RB/E2F Pathway Controls Expression of the Cancer Genomic DNA Deaminase <i>APOBEC3B</i> (published, <i>eLife</i> , 2020)	41
<b>Chapter 4</b>	Aberrant <i>APOBEC3B</i> Expression in Breast Cancer Is Linked to Proliferation and Cell Cycle Phase (published, <i>Cells</i> , 2023)	85
<b>Chapter 5</b>	Discussion: Clinical Implications of <i>APOBEC3</i> -Mediated Mutagenesis in Breast Cancer (published, <i>Clin Cancer Res</i> , 2023)	115
<b>Chapter 6</b>	English Summary	142
	Nederlandse Samenvatting	145
	Research Data Management Plan	148
	List of publications	149
	PhD Portfolio	150
	Acknowledgements	152
	Curriculum Vitae	156

A dedication to my parents,  
for delightful memories have inspired  
me, me and everyone.

For the things we have done together  
and going to get it done.

With love and respect.



# 1

## A Short Introduction and Thesis Outline

---

The human genome is a testament to the intricate relationship between host and pathogen. The gradual evolution of the defense mechanisms that underpin this relationship, can be seen as an evolutionary arms race: two invested parties holding one another in an apparent deathlock, while one-sided counters may detrimentally affect the competing party. Such an evolutionary arms race is not unlike the survival dynamics between prey animals and their predators which, through a series of stepwise evolutionary processes, developed formidable defense mechanisms (such as the bony osteoderms and club-like tails of Ankylosauria) and equally advanced predatory armaments and/or behavioral patterns, respectively (1). Likewise, past host-pathogen dynamics ultimately gave rise to our current collective of defense mechanisms against well-known microbes (including bacteria and fungi) and more exotic genomic invaders (such as retroviruses and endogenous LINE-1 genomic elements). One of these defense mechanisms involves the neutralizing activity of the APOBEC3 (A3) sub-family of deaminases. These evolutionary distinct enzymes are, amongst several other functions, well-known for their ability to neutralize foreign DNA.

The human repertoire of A3 genes is present on a single locus on chromosome 22 and is likely the product of past gene duplication events and subsequent functional divergence. Humans and other apes express up to seven distinct A3 proteins, namely A3A, A3B, A3C, A3D, A3F, A3G, and A3H. Amongst other functions, these A3 enzymes target cytosine nucleobases and catalyze the hydrolytic removal of an amine group and replacement with a carbonyl group. This process, called cytosine deamination, turns cytosine nucleotides into non-coding uracils, which restricts the propagation of foreign DNA. Specifically, A3A, A3D, A3F, A3G and A3H can restrict the replicative potential of retroviruses, while other A3 family members, including A3B and A3C, may restrict the DNA-based hepatitis B virus. Moreover, several members, including A3A, A3B, A3C and A3F aid in the restriction of LINE-1 retrotransposons, contributing to the maintenance of host genome integrity (2-4). Altogether, a range of A3 proteins has evolved to aid in the clearance of foreign DNA, contributing to overall host homeostasis.

However, like many other seemingly optimized adaptations, this host defense mechanism presents itself with an evolutionary trade-off. A case in point, A3 proteins may access and mutate the host genome, since the A3 deamination mechanism is agnostic to the provenance of nucleobases, be it foreign DNA or host genomic code. When not repaired in a timely fashion, uracils resulting from this off-target activity are read as thymines by the host replicative mechanism and eventually give rise to a C-to-T transition (5). Additionally, when uracils are excised by base excision

repair, error-prone translesion synthesis facilitates additional C-to-T transitions and C-to-G transversions. Interestingly, A3 proteins preferably target cytosines within a specific tri-nucleotide context. For example, A3A and A3B prefer a 5'-TCW-3' (W = A or T) context, which allows for tracing their activity in a *post-hoc* fashion using next-generation sequencing of human tissues. Collectively referred to as "APOBEC mutational signatures", these off-target mutations have been shown to significantly contribute to the overall mutational load of several cancer subtypes (6, 7). In fact, when placed alongside other mutational signatures (including those arising from UV-induced damage, cigarette smoke and several DNA damage repair deficiencies) A3-mediated deamination infamously makes podium as one of the most prominent sources of cancer mutation (7, 8). Of the seven A3 family members, A3A and A3B are thought to be most universally responsible for the APOBEC-mediated deamination in cancer (8-10).

Interestingly, APOBEC signature mutations are especially prominent in cancers caused by tumorigenic viruses, such as head-and-neck and cervical cancer (7). In these cancers certain viral oncoproteins can stimulate the proliferative signaling pathways and are thought to concomitantly activate transcription of A3B (11-14). Given the anti-viral functions of A3 proteins one might come to hypothesize that A3-related mutations almost exclusively occur in cancer types with a clear viral etiology. However, A3-related mutations are in no way unique to cancers with a viral component and, as such, are commonly found in, amongst many others, breast cancer (6, 7, 15). In fact, in breast cancer APOBEC signature mutations are detected in primary as well as metastatic disease, ranging from modest to extremely high levels and accounting for almost two-thirds of total mutations in hypermutated breast tumors (those with  $\geq 10$  mutations per megabase) (6, 15-19). Interestingly, this aspect may provide a useful insight in the regulation of A3 transcriptional activation in breast cancer, as the signaling pathways through which tumorigenic viruses induce A3 expression may overlap with A3 transcriptional activation in non-viral cancers.

This thesis reports on several recent advances in the investigation of A3B transcriptional (dys)regulation in breast cancer. In **Chapter 2** we describe the discovery that A3B expression can be induced by the polyomavirus large T antigen (PyV-LTA<sub>g</sub>), which likely stimulates the proliferative RB/E2F signaling pathway. Building on this work, findings in **Chapter 3** converge on the precise mechanism through which the RB/E2F signaling pathway represses A3B expression in breast cancer and, when dysregulated, allows for overexpression of the A3B oncoprotein. We also report that *in vitro* A3B expression increases to dramatic levels when loss

of repression coincides with activation of inflammatory PKC/ $\kappa$ B signaling pathway. Thus, these two (functionally separate) signaling pathways may be responsible for the lion's share of *A3B* induction in breast cancer. **Chapter 4** further deepens our understanding of *A3B* gene regulation within the context of the cell cycle, revealing that activation of *A3B* through the PKC/ $\kappa$ B signaling pathway is strongest in actively cycling cells and virtually absent in growth-arrested cells. Finally, **Chapter 5** provides a comprehensive overview of recent advances in the field, and provides further insight in clinical detection methods, regulatory pathways, and clinical relevance of the A3 proteins relevant to breast cancer. Specifically, these findings provide future directions to better understand and solidify the relevance of A3 proteins in breast cancer.

## References

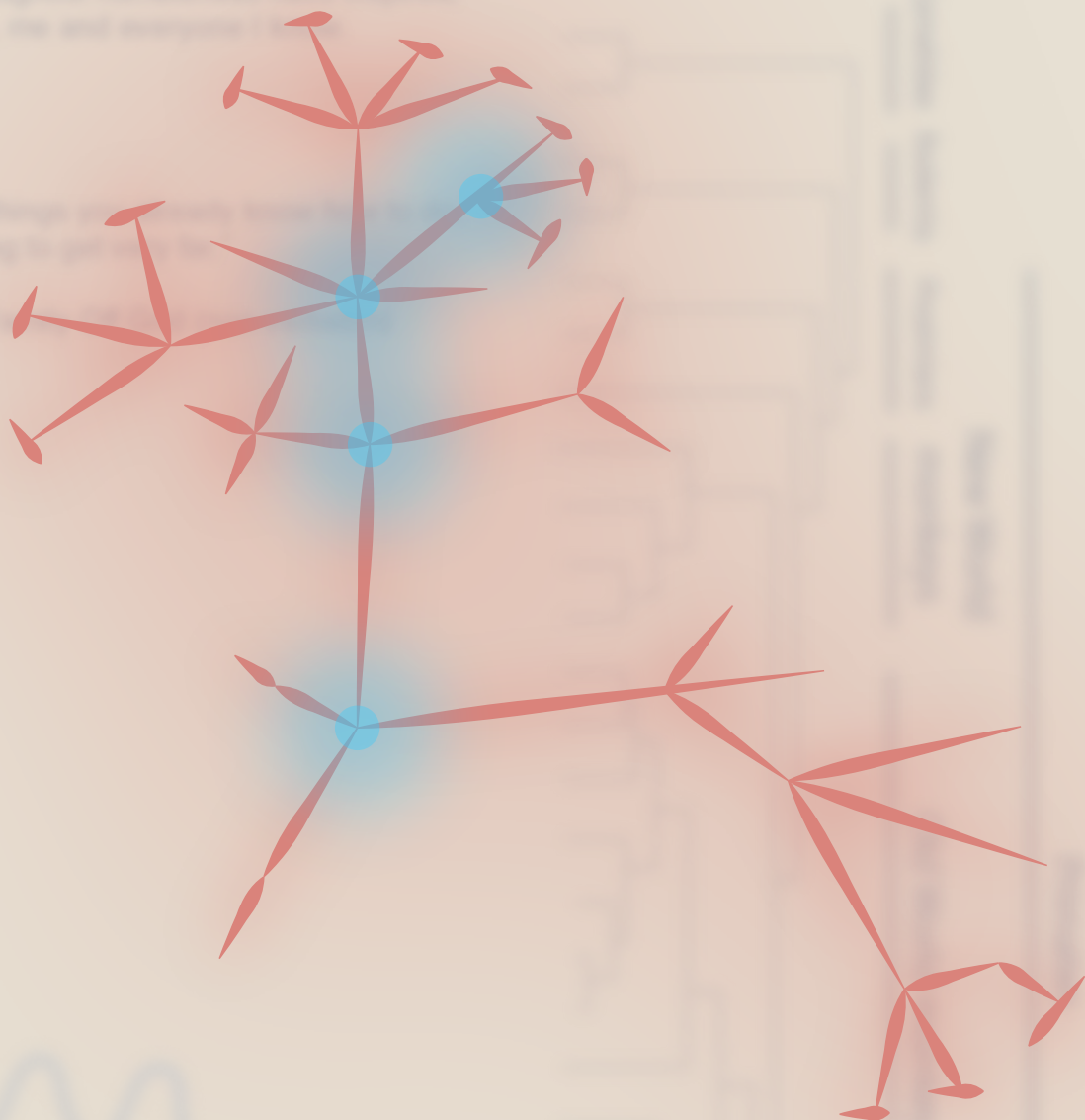
1. Arbour VM, Currie PJ. Ankylosaurid dinosaur tail clubs evolved through stepwise acquisition of key features. *Journal of Anatomy*. 2015;227(4).
2. Muckenfuss H, Hamdorf M, Held U, Perković M, Löwer J, Cichutek K, et al. APOBEC3 Proteins Inhibit Human LINE-1 Retrotransposition. *Journal of Biological Chemistry*. 2006;281(31).
3. Wissing S, Montano M, Garcia-Perez JL, Moran JV, Greene WC. Endogenous APOBEC3B Restricts LINE-1 Retrotransposition in Transformed Cells and Human Embryonic Stem Cells. *Journal of Biological Chemistry*. 2011;286(42).
4. Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV. APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. *eLife*. 2014;3.
5. Krokan HE, Sætrom P, Aas PA, Pettersen HS, Kavli B, Slupphaug G. Error-free versus mutagenic processing of genomic uracil—Relevance to cancer. *DNA Repair*. 2014;19.
6. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534(7605).
7. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. *Nature*. 2020;578(7793).
8. Jarvis MC, Carpenter MA, Temiz NA, Brown MR, Argyris PP, Brown WL, et al. Mutational impact of APOBEC3B and APOBEC3A in a human cell line. *bioRxiv*. 2022.
9. Jalili P, Bowen D, Langenbucher A, Park S, Aguirre K, Corcoran RB, et al. Quantification of ongoing APOBEC3A activity in tumor cells by monitoring RNA editing at hotspots. *Nat Commun*. 2020;11(1):2971.
10. 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(12).
11. Vieira VC, Leonard B, White EA, Starrett GJ, Temiz NA, Lorenz LD, et al. Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *mBio*. 2014;5(6).
12. Warren CJ, Van Doorslaer K, Pandey A, Espinosa JM, Pyeon D. Role of the host restriction factor APOBEC3 on papillomavirus evolution. *Virus Evol*. 2015;1(1).
13. Habraken Y, Piette J. NF-kappaB activation by double-strand breaks. *Biochem Pharmacol*. 2006;72(9):1132-41.
14. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT). *The Journal of Molecular Diagnostics*. 2015;17(3):251-64.
15. Brady SW, McQuerry JA, Qiao Y, Piccolo SR, Shrestha G, Jenkins DF, et al. Combating subclonal evolution of resistant cancer phenotypes. *Nat Commun*. 2017;8(1):1231.
16. Barroso-Sousa R, Jain E, Cohen O, Kim D, Buendia-Buendia J, Winer E, et al. Prevalence and mutational determinants of high tumor mutation burden in breast cancer. *Ann Oncol*. 2020;31(3):387-94.
17. Jarvis MC, Ebrahimi D, Temiz NA, Harris RS. Mutation Signatures Including APOBEC in Cancer Cell Lines. *JNCI Cancer Spectrum*. 2018;2(1).
18. Petljak M, Alexandrov LB, Brummel JS, Price S, Wedge DC, Grossmann S, et al. Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. *Cell*. 2019;176(6).
19. Angus L, Smid M, Wilting SM, van Riet J, Van Hoeck A, Nguyen L, et al. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nat Genet*. 2019;51(10):1450-8.



A dedication to my parents,  
for delightful memories have inspired  
me, me and everyone I love.

For the things we have done together  
and going to get it done.

With love and respect.



# 2

## Polyomavirus T Antigen Induces *APOBEC3B* Expression Using an LXCXE-Dependent and TP53-Independent Mechanism

---

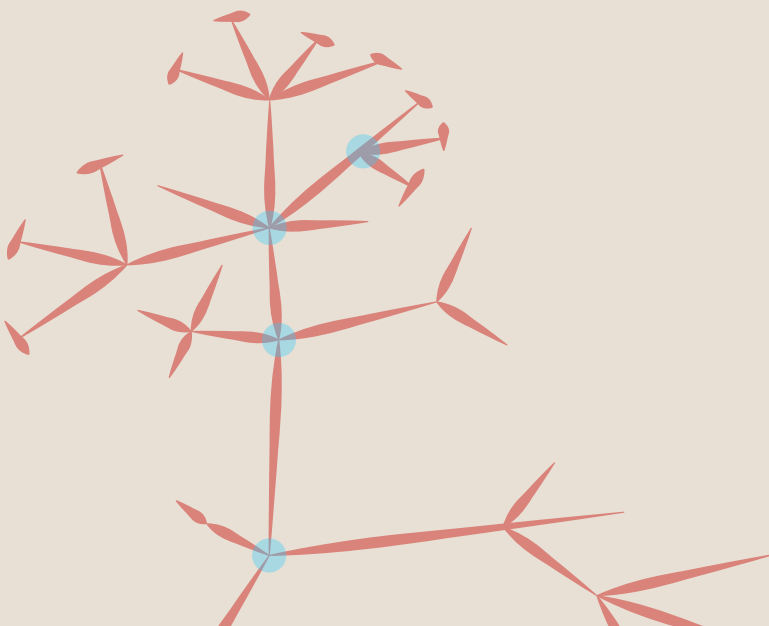
Gabriel J. Starrett<sup>1</sup>, Artur A. Serebrenik<sup>1</sup>, Pieter A. Roelofs<sup>2</sup>, Jennifer L. McCann<sup>1</sup>, Brandy Verhalen<sup>3</sup>, Matthew C. Jarvis<sup>1</sup>, Teneale A. Stewart<sup>1</sup>, Emily K. Law<sup>1</sup>, Annabel Krupp<sup>4</sup>, Mengxi Jiang<sup>3</sup>, John W. M. Martens<sup>5</sup>, Ellen Cahir-McFarland<sup>4</sup>, Paul N. Span<sup>2</sup>, Reuben S. Harris<sup>1,6</sup>

1. Department of Biochemistry, Molecular Biology and Biophysics, Masonic Cancer Center, Institute for Molecular Virology, University of Minnesota, Minneapolis, Minnesota, USA
2. Department of Radiation Oncology, Radiotherapy and OncoImmunology Laboratory, Radboud University Medical Center, Nijmegen, The Netherlands
3. Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama, USA
4. Department of Neuroimmunology, Biogen, Cambridge, Massachusetts, USA
5. Department of Medical Oncology, Cancer Genomics Netherlands, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands
6. Howard Hughes Medical Institute, University of Minnesota, Minneapolis, Minnesota, USA

Published in: mBio. 2019;10(1):e02690-18. doi: 10.1128/mBio.02690-18.

## Abstract

APOBEC3B is a single-stranded DNA cytosine deaminase with beneficial innate antiviral functions. However, misregulated APOBEC3B can also be detrimental by inflicting APOBEC signature C-to-T and C-to-G mutations in genomic DNA of multiple cancer types. Polyomavirus and papillomavirus oncoproteins induce APOBEC3B overexpression, perhaps to their own benefit, but little is known about the cellular mechanisms hijacked by these viruses to do so. Here we investigate the molecular mechanism of APOBEC3B upregulation by the polyomavirus large T antigen. First, we demonstrate that the upregulated APOBEC3B enzyme is strongly nuclear and partially localized to virus replication centers. Second, truncated T antigen (truncT) is sufficient for APOBEC3B upregulation, and the RB-interacting motif (LXCXE), but not the p53-binding domain, is required. Third, genetic knockdown of RB1 alone or in combination with RBL1 and/or RBL2 is insufficient to suppress truncT-mediated induction of *APOBEC3B*. Fourth, CDK4/6 inhibition by palbociclib is also insufficient to suppress truncT-mediated induction of *APOBEC3B*. Last, global gene expression analyses in a wide range of human cancers show significant associations between expression of *APOBEC3B* and other genes known to be regulated by the RB/E2F axis. These experiments combine to implicate the RB/E2F axis in promoting *APOBEC3B* transcription, yet they also suggest that the polyomavirus RB-binding motif has at least one additional function in addition to RB inactivation for triggering *APOBEC3B* upregulation in virus-infected cells.



## Introduction

Genetic diversity is key to virus replication, pathogenesis, and transmission, and particularly for escape from adaptive immune responses in vertebrate species (1–3). Each virus has evolved to maintain an optimized level of genetic diversity for its own unique life cycle through various mechanisms, with some viruses having high mutation rates and others much lower mutation rates, notably the dsDNA viruses (4, 5). Recently, it has been concluded that the genome compositions of multiple DNA tumor viruses, including high-risk human papillomavirus (HPV) types and BK polyomavirus (BKPv), have been shaped by long-term interactions with the innate, antiviral APOBEC deaminases (6–9). It has also been suggested that acutely occurring mutations by these enzymes in the major capsid gene of polyomaviruses promote antibody escape during polyomavirus-associated nephropathy and progressive multifocal leukoencephalopathy (10). The overlap between these disease variants and oncogenic enzymes is striking especially in light of growing evidence linking BKPv infection and a subset of urothelial carcinomas with high levels of APOBEC-signature mutations (11).

Several APOBEC enzymes, including APOBEC3B (A3B), bind 5'-TC dinucleotide motifs in single-stranded DNA and catalyze the hydrolytic conversion of cytosine to uracil (12, 13). Left unrepaired, uracil lesions can serve as the templates for new DNA synthesis and directly result in C-to-T mutations. Alternatively, if the uracil base is excised by cellular uracil DNA glycosylase 2 (UNG2), then the resulting abasic site becomes noninstructional and may trigger cellular DNA polymerases to insert an adenine opposite the lesion, except for REV1, which tends to incorporate either adenine or cytosine. Thus, APOBEC-catalyzed DNA deamination of 5'-TC motifs results in both C-to-T and C-to-G mutations (a signature frequently expanded to include the 3'-nucleobases A and T and referred to in the context of trinucleotide motifs 5'-TCA and 5'-TCT). An additional hallmark of virus mutagenesis by APOBEC enzymes is a bias toward occurring on the template of lagging-strand DNA replication (14–16). A likely mechanistic relationship with single-stranded DNA replication intermediates is supported by similar correlations in model yeast and *Escherichia coli* experiments (17, 18).

Human cells have the potential to express up to nine active DNA cytosine deaminases (AID, APOBEC1, and A3A/B/C/D/F/G/H) (19–22). Seven of these enzymes prefer 5'-TC motifs in single-stranded DNA, whereas AID uniquely prefers 5'-RC and APOBEC3G (A3G) prefers 5'-CC. A3B is the most likely APOBEC family member to contribute to the mutagenesis and evolution of small DNA tumor viruses because

it is specifically upregulated by viral oncoproteins. For high-risk HPV types, the oncoproteins E6 and E7 have been implicated through various pathways (23–26). For polyomaviruses, including JC, BK, and Merkel cell (JCPyV, BKPyV, and MCPyV, respectively), the large T antigen (TAg) is sufficient for A3B upregulation through a yet-to-be determined mechanism (6). However, the considerable functional overlap of these proteins, RB inactivation by E7 and TAg and p53 inactivation by E6 and TAg, may indicate limited pathways for A3B modulation by viruses (27, 28). Here we investigate the molecular mechanism by which polyomaviruses promote the transcriptional upregulation of *A3B* with results converging on the cellular RB/E2F pathway, which is often deregulated in cancer.

## Materials and Methods

**Cell lines, culture conditions, and lentivirus production.** Primary renal proximal tubule epithelial cells (RPTECs; Lonza) were grown in REGM (Lonza). MCF10A cells were grown in MEGM (Lonza) containing penicillin (100 U/ml) and streptomycin (100 µg/ml). HuK(i)G10 cells were grown in RenaLife epithelial medium (Lifeline Cell Technologies) with 5% FBS. MCF7 and derivative cell lines were grown in Richter's modification medium containing 5% fetal bovine serum, penicillin (100 U/ml), streptomycin (100 µg/ml), and 11.25 nM recombinant human insulin. All cell lines were grown at 37°C in a 5% CO<sub>2</sub> incubator. Lentiviruses expressing TAg and mutant derivatives were produced in 293T cells and transduced into RPTECs as described previously (23).

**Antibodies.** Large T and truncT forms of BKPyV T antigen were detected using pAb416 (30). JCPyV large T antigen was detected using PAB2000 (68). A Harris lab custom rabbit anti-human A3B monoclonal antibody, 5210.87.13, was used in immunoblotting assays and in high-resolution immunofluorescent microscopy experiments with JCPyV-infected HuK(i)G10 cells (69). Santa Cruz sc-130688 was used for A3B quantification and lower-resolution microscopy of BKPyV- and JCPyV-infected HuK(i)G10 cells (69). UNG2 was detected using the Abcam antibody ab23926. RB1 was detected using Santa Cruz sc-102, and RBL1 was detected using Cell Signaling 89798, whereas RBL2 could not be detected in immunoblots with Abcam antibody ab71143. TP53 (p53) was detected using clone DU-1 (Santa Cruz SC-126), p21 using CST clone 1201 (CST no. 2947), and beta-actin using CST clone 13E5 (CST no. 4970).

**RNA isolation, RT-qPCR, and immunoblots.** Total RNA was harvested by removal of medium and resuspension in TRIzol (Thermo Fisher), and purification was

done per the manufacturer's protocol. RT qPCR was used to quantify *A3B* and *UNG2* transcripts in siRNA experiments as described previously (11, 23) and these methods were adapted for *CCNE2*. Protein lysates from virus and siRNA experiments were harvested at 3 or 7 dpi or posttransduction, quantified, and immunoblotted as described previously (70). Data were plotted and *t* tests were calculated using GraphPad Prism 7.

**Immunofluorescent microscopy experiments.** HuK(i)G10 kidney cells were seeded at 6,000 cells/ well in a 96-well plate. Twenty-four hours later, infection with JCPyV was performed as described, and then cells were collected 7 dpi. Infected cells were incubated with EdU (Click-iT Plus EdU Alexa Fluor 647 imaging kit; Thermo Fisher Scientific) for 15 min and incubated with CSK buffer (10 mM HEPES-KOH, pH 7.4; 300 mM sucrose; 100 mM NaCl; 3mM KCl; 0.5% Triton X-100) (71) for 2 min on ice. Cells were then fixed in 4% PFA for 10 min followed by permeabilization with 0.5% Triton X-100 for 20 min. For EdU detection, the Click-iT reagent was added for 30 min in the dark according to the manufacturer's protocol and washed three times with PBS. Samples were incubated with BlockAid blocking solution (Thermo Fisher) for 1 h at room temperature. T antigen, VP1, and A3B staining was performed using the aforementioned antibodies at 1:1,000, 1:1,000, and 1:100 (1:50 for sc-130688) dilutions in BlockAid, respectively, overnight at 4°C followed by staining with the secondary antibodies for 1 h at room temperature. Images were acquired on the Opera Phenix (PerkinElmer) with the confocal 63 water objective. Immunofluorescence in RPTE cells was performed as described using the above-mentioned antibodies and imaged on the Invitrogen EVOS FL Imaging System (72).

**siRNA and expression construct transfection.** siRNAs targeting RB1 (J-003296-23; Dharmacon), RBL1 (SI02629921; Qiagen), and RBL2 (sc-29425; Santa Cruz) and fluorescein-conjugated nontargeting control siRNA (sc-36869; Santa Cruz) were purchased and diluted to a working concentration of 20 M. A final concentration of 20 nM was used for all targets in RPTECs, and 40 nM was used in MCF10A cells. siRNAs were delivered to cells using Lipofectamine RNAiMax (Thermo Fisher) as described previously (73).

**Drug treatment.** Cells were treated with 5  $\mu$ M nutlin (Sigma) for 24 h, and after 18 h of treatment, 25 ng/ml phorbol myristate acetate (PMA; Sigma) was added for the final 6 h prior to RNA extraction. MCF10A cells were treated only with PMA or DMSO, and RNA was isolated 6 h after treatment. For the palbociclib experiments, MCF10A and MCF7 (p53 WT and low *A3B* expression); HCC1937 and HCC1395 (low *A3B* expression); T47D, HCC1954, and Hs578T (intermediate *A3B* expression); and

HCC1599, HCC1143, SUM-225-CWN, and HCC202 (high *A3B* expression) cells were cultured in 6-well plates (Costar 3516; Corning Incorporated) until 70% confluence. Palbociclib (S1116; Selleckchem) was stored as a 5 mM solution in H<sub>2</sub>O and added to cells at concentrations of 0  $\mu$ M (H<sub>2</sub>O control), 0.5  $\mu$ M, and 2.5  $\mu$ M. No palbociclib was added to cells of the 0-h time point, which instead was transferred to ice prior to RNA isolation. RNA was also isolated at 0, 1, and 4 h post-palbociclib administration (total RNA purification kit 37500; Norgen), and cDNA was synthesized with 500 ng RNA (iScript 170-8891; Bio-Rad). RT-qPCR assays for *A3B* and *CCNE2* were performed using the C1000 Thermal Cycler (Bio-Rad). For pretreatment with PMA (tlrl-pma; InvivoGen; 1-mg/ml stock in DMSO), cells were treated with 0 ng/ml (DMSO) or 25 ng/ml PMA for 3 h, followed by treatment with 0  $\mu$ M, 0.5  $\mu$ M, and 2.5  $\mu$ M palbociclib. RNA was isolated 0, 1, and 3 h after addition of palbociclib and processed as described above. For pretreatment with palbociclib, cells were treated with 0  $\mu$ M, 0.5  $\mu$ M, and 2.5  $\mu$ M palbociclib for 30 min, followed by treatment with 0 ng/ml (DMSO) or 25 ng/ml PMA. RNA was isolated 0, 1, and 3 h after addition of PMA and processed as described above.

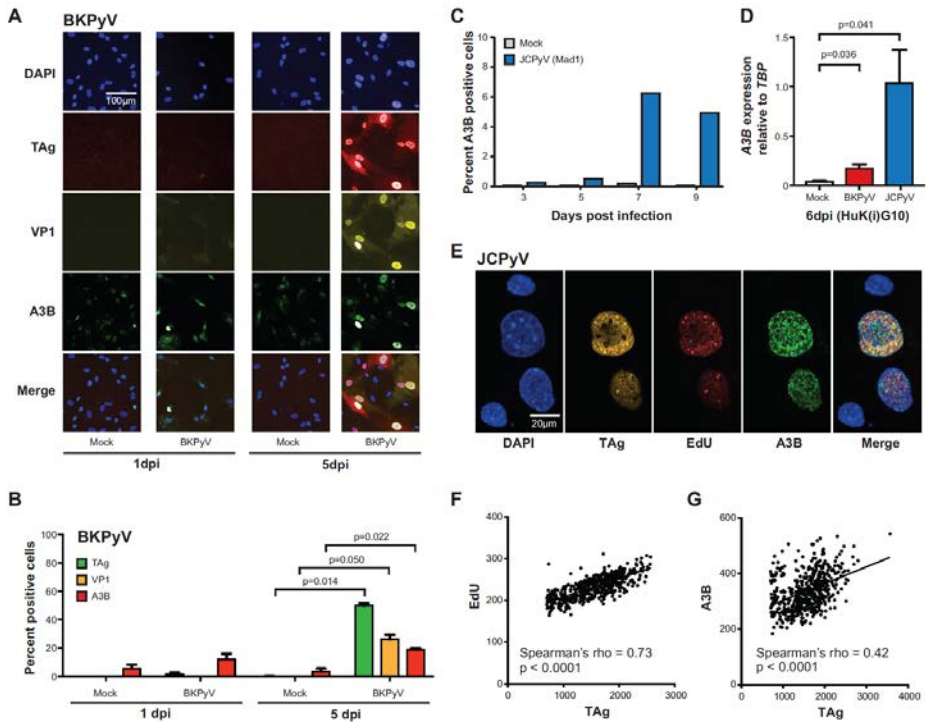
**Bioinformatics.** TCGA expression data were downloaded from the Broad GDAC Firehose as of January 2016. Expression correlations against *A3B* by all other genes in the breast cancer cohort were calculated, and significant correlates were used to determine significant activation or inhibition of upstream regulators using Ingenuity Pathway Analysis (Qiagen). All other Spearman correlations and P values were calculated, and heat maps were plotted using the R statistical environment. P values were adjusted for multiple comparisons using the Bonferroni correction, and resulting q values were reported. Cell cycle data were acquired from <http://www.targetgenereg.org/> in May 2018.

## Results

### Visualization of endogenous APOBEC3B protein in polyomavirus-infected cells

A3B induction by polyomaviruses has been shown at the mRNA level by RT-qPCR and at the protein level by immunoblotting in primary renal proximal epithelial cells (RPTECs) (6). To extend these results to other relevant cell types, RT-qPCR and immunofluorescent microscopy were used to ask whether polyomavirus infection causes a general pan-nuclear upregulation of A3B enzyme and/or localization to discrete subnuclear regions such as virus replication centers. Immortalized human kidney (HuK(i)G10) cells were infected with BKPyV (Dunlop strain) and

JCPyV (MAD1 strain) and subjected to analyses at various days postinfection (dpi). Infected cells have enlarged nuclei and robust expression of TAg and VP1 at 3 to 5 dpi (Fig. 1A). A3B expression was more variable but still clearly and significantly increased after infection with either virus compared to mock-infected controls (Fig. 1A to D). Generally, JCPyV is regarded to have slower replication dynamics than BKPyV (Dunlop), so initial JCPyV infections were run out in a time course showing peak A3B expression at 7 dpi (Fig. 1C). Across these experiments, JCPyV-infected HuK(i)G10 cells showed a greater differential expression of A3B mRNA and protein compared to mock-treated cells (Fig. 1B to D).



**Figure 1.** Visualization and quantification of A3B expression in PyV-infected cells. **(A and B)** Immunofluorescent images and quantification of TAg, VP1, and A3B in BKPyV-infected HuK(i)G10 cells at 1 and 5 dpi (significance determined using Welch's two-tailed *t* test; *n* = 2 biological replicates). **(C)** Time course of A3B mRNA levels in JCPyV (Mad1 strain) versus mock-infected HuK(i)G10 cells. **(D)** RT-qPCR quantification of A3B transcripts in mock-, BKPyV-, and JCPyV (Mad1)-infected HuK(i)G10 cells at 6dpi (significance determined by Welch's two-tailed *t* test; *n* = 3 technical replicates). **(E)** High-resolution immunofluorescent microscopy images of DAPI, A3B, EdU, and TAg in HuK(i)G10 cells infected with JCPyV (Mad1 strain). **(F and G)** Correlation coefficients and *P* values for EdU and A3B levels versus T antigen intensity in >100 cell images from a single experiment similar to that in panel E.



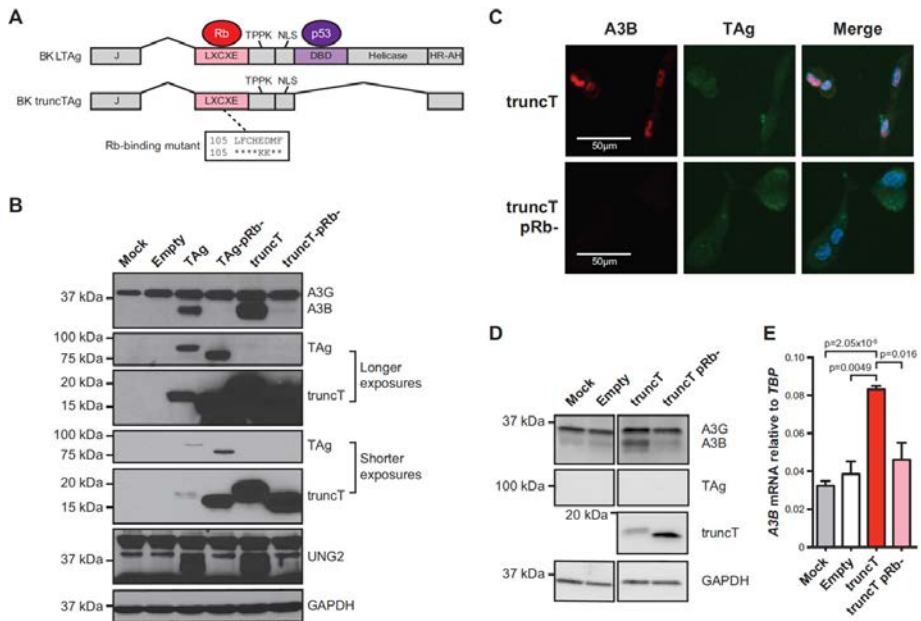
JCPyV-infected cells were also analyzed 7 dpi by high-resolution immunofluorescent microscopy for expression of A3B and viral proteins and for formation of virus replication foci. Cells were stained for DAPI, TAg, A3B, and EdU with virus replication centers appearing as brightly stained puncta positive for both TAg and EdU (representative images in Fig. 1A and E) (29). In infected cells, A3B is strongly induced with a pan-nuclear staining pattern that is sometimes coincident with EdU-positive virus replication foci. Incorporation of EdU into active replication foci is highlighted by strong positive correlations with TAg stain intensity, as expected, whereas A3B showed weaker but still significantly positive correlations (Fig. 1F and G). These data indicate that A3B upregulation may be a general property of polyomavirus infection and that A3B may access at least a subset of virus replication centers.

### **APOBEC3B upregulation by polyomavirus large T antigen requires the canonical RB-interacting motif LXCXE**

Based on the results above and our previous studies (6), multiple polyomaviruses have the conserved capacity to upregulate A3B expression in primary and immortalized kidney epithelial cells through the functions of large (L) TAg. To investigate the LTag domains responsible for A3B induction, and thus also implicate associated cellular factors, we tested a naturally occurring splice variant of BKPyV LTag, known as truncT, which lacks the DNA-binding and helicase domains essential for p53 neutralization (30–32) (schematic in Fig. 2A). In parallel, we also assessed derivatives of LTag and truncT with a disrupted LXCXE motif, which is required for inhibiting the tumor suppressor protein RB1 as originally shown for SV40 TAg (33). RPTECs were transduced with lentiviruses expressing an empty multiple cloning site as a negative control, BKPyV LTag as a positive control, BKPyV truncT, and RB-binding site mutant derivatives; incubated 3 days; and assessed by immunoblotting and fluorescence microscopy. Mock-transduced cells express low levels of A3B, and transduction with empty lentivirus causes a modest increase in this protein and also raises A3B levels to faintly detectable levels (Fig. 2B). In contrast, both LTag and truncT induce expression of A3B and UNG2, a known target of the RB-E2F axis (34), and all induction for both of these proteins is eliminated by two amino acid substitutions shown to abrogate RB binding in SV40 TAg (LFCHED to LFCHKK) (33) (Fig. 2B). The LTag and truncT mutants invariably migrate faster than the corresponding wild-type proteins during SDS-PAGE, which is likely due to a charge differential caused by the two amino acid substitutions. Immunofluorescent microscopy images also show truncT-mediated induction of nuclear A3B but not by the RB-binding mutant derivative (Fig. 2C).

To date, many aspects of A3B regulation and function have been determined using normal-like and cancerous mammary epithelial cell lines due to higher capacities

for genetic manipulation over primary cells and greater relevance to cancer (35, 36). To ask whether TAG induction of A3B also occurs in one of these more tractable systems, the normal-like mammary epithelial cell line MCF10A was transduced with constructs expressing BKPyV truncT or the RB-binding mutant and analyzed as described above. Both immunoblotting and RT-qPCR yielded similar results with truncT but not the RB-binding mutant causing A3B induction (Fig. 2D and E). Thus, polyomavirus T antigen appears to possess a conserved, LXCXE-dependent capacity to induce A3B in different epithelial cell types.

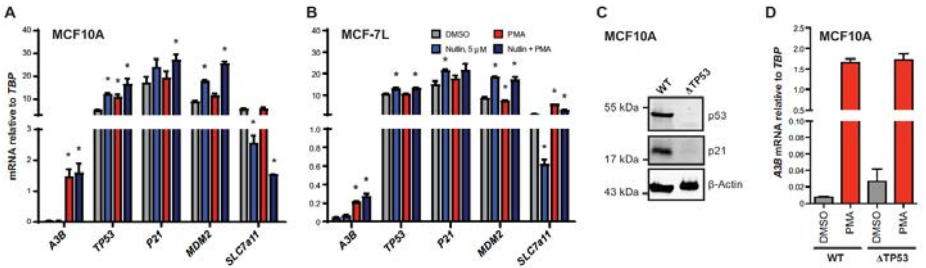


**Figure 2.** RB-binding domain is necessary for A3B induction by polyomavirus T antigen. **(A)** Diagram of BKPyV T antigen isoforms and the LXCXE mutant used here. **(B)** Immunoblots for the indicated proteins in RPTeCs transduced with a lentiviral vector expressing LTag, truncT, the indicated mutants, or nothing (empty). Mock-transduced cells were analyzed in parallel as an additional control. **(C)** Immunofluorescent microscopy images for truncT, truncT LXCXE mutant, and A3B in transduced RPTeCs. **(D and E)** Immunoblots and RT-qPCR results for MCF10A cells transduced with the indicated constructs as in panel B (mean and SEM shown in panel E;  $n=3$  biological replicates;  $P$  value determined by Welch's two-tailed  $t$  test).

### TP53 inactivation is dispensable for APOBEC3B induction

The aforementioned data comparing LTag and truncT simultaneously implicate RB1 and demonstrate that p53 inhibition is not required for A3B induction because truncT completely lacks the p53 binding domain (Fig. 2). To further ask whether p53 inactivation might influence A3B gene expression, we quantified A3B mRNA levels in

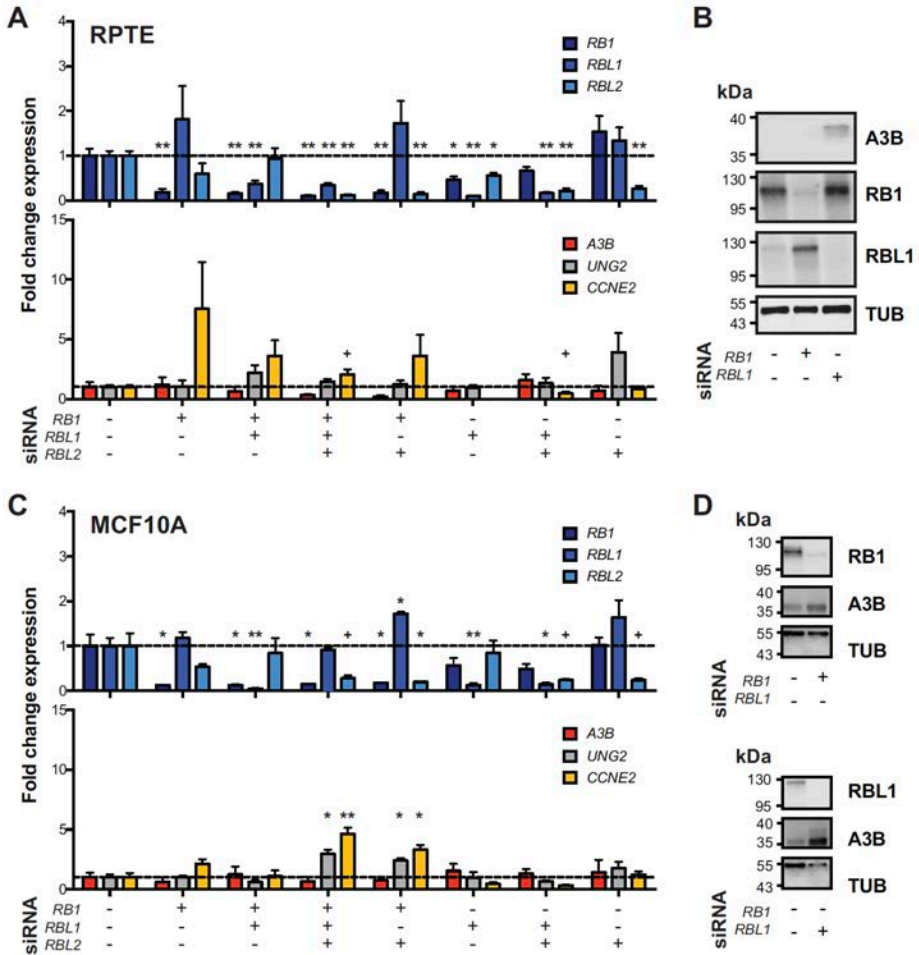
two cell lines that have been used to study A3B regulation, MCF10A and the human estrogen receptor-positive breast cancer cell line MCF-7L (above and references 35 to 38). Each cell line was treated with either DMSO or 5  $\mu$ M nutlin, which is a drug that protects p53 from MDM2-mediated degradation (39). As controls, mRNA levels were quantified for two genes activated by p53 (*P21*, *MDM2*) (40–43) and one gene repressed by p53 (*SLC7A11*) (44, 45). Respectively, the expression of these genes was derepressed or repressed by nutlin treatment (Fig. 3A and B). In comparison, neither steady-state nor PMA-induced A3B mRNA levels were changed by nutlin (Fig. 3A and B). Moreover, Cas9-mediated knockout of *TP53* in MCF10A cells also caused no significant effect on basal or PMA-induced A3B expression levels (Fig. 3C and D). These data combine to indicate that p53 by itself has no significant role in the either the PMA-induced pathway or basal-state transcriptional regulation of A3B, discouraging our original hypothesis (23) and conflicting with published data (38) (see Discussion).



**Figure 3.** Inactivation of p53 does not affect A3B expression. **(A and B)** Bar plots of RT-qPCR measurements of relevant genes in MCF10A (A) and MCF7L (B) cells treated with DMSO, 5  $\mu$ M nutlin, PMA, or nutlin + PMA. Statistically significant changes by Student's *t* test ( $P < 0.05$ ) are noted by an asterisk (mean and SEM;  $n = 3$  technical replicates). **(C)** Immunoblots of WT and *TP53* KO MCF10A cell lines. **(D)** RT-qPCR results showing the effects of PMA treatment on A3B mRNA levels in WT and *TP53* KO MCF10A cell lines.

### RB-family knockdown is insufficient to induce APOBEC3B expression

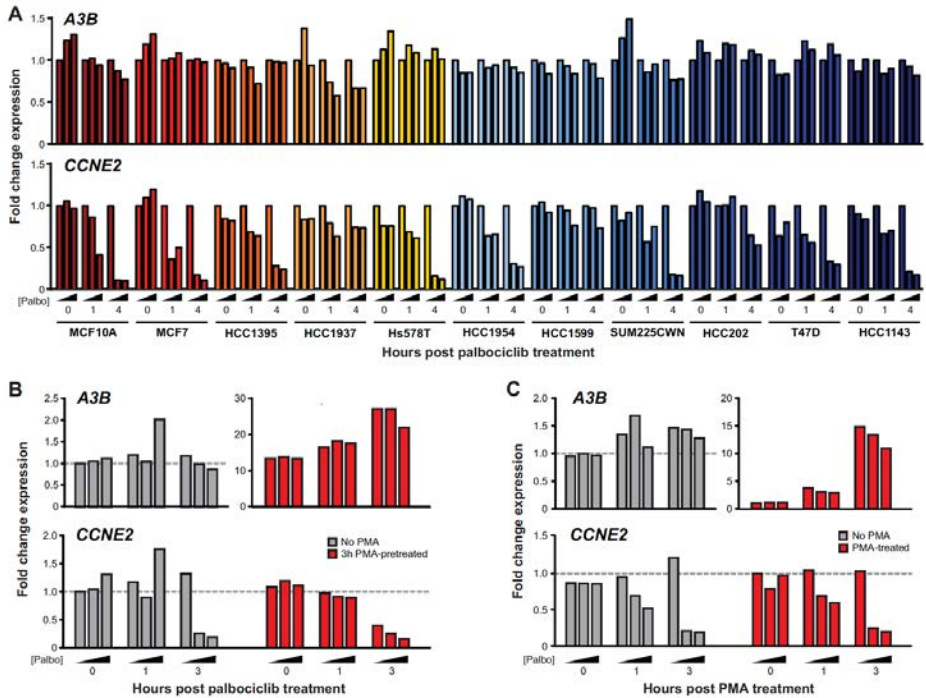
RB1 is arguably the most widely studied target of the LXCXE motif of viral proteins such as HPV E7, adenovirus E1A, and polyomavirus LTA $\gamma$  (27, 31, 32, 46). However, the related pocket proteins RBL1 (p107) and RBL2 (p130) also have an LXCXE-binding motif, are similarly targeted and inactivated by LTA $\gamma$  and truncT, and may be involved in A3B regulation (47–50). The aforementioned viral proteins bind to the hypophosphorylated forms of RB1, RBL1, and RBL2, which inhibits phosphorylation by cyclin-dependent kinases (CDKs) and leads to an accelerated cell cycle in part by deregulation of E2F transcriptional activities. To investigate the roles of RB1, RBL1, and RBL2 in A3B transcriptional regulation, a series of knockdown experiments was done with siRNAs targeting each of these factors in RPTECs and MCF10A cells



**Figure 4.** Modulation of RB family genes and A3B regulation. (**A** and **C**) Bar plots of RT-qPCR quantification of *RB*-family mRNAs (top) and predicted responsive genes, *A3B*, *UNG2*, and *CCNE2* (bottom), in RPTEC and MCF10A cells with siRNA-mediated KD of *RB*-family genes. *P* values for each siRNA combination compared to control were calculated using Welch's two-tailed *t* test and were indicated using the following symbols: +,  $P < 0.1$ ; \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ;  $n = 3$  biological replicates. (**B** and **D**) Immunoblots for A3B, RB1, and RBL1 in RPTEC and MCF10A cells following treatment with the indicated siRNA.

(Fig. 4A to D). RT-qPCR showed that 75% knockdown was achieved for each targeted gene (upper panels in Fig. 4A and C). As controls, *CCNE2* was upregulated upon *RB1* knockdown and *UNG2* was moderately upregulated by *RBL2* knockdown (lower panels in Fig. 4A and C). However, no combination of siRNAs resulted in significant upregulation of *A3B* mRNA levels (lower panels in Fig. 4A and C). Knockdown

of RB1 and RBL1 was validated at the protein level, but RBL2 could not be clearly discerned with available commercial antibodies (Fig. 4B and D). In contrast to the RT-qPCR results, protein-level A3B expression did appear to be elevated upon RBL1 knockdown. These results suggest that depletion of each RB family member alone or in combination is insufficient to significantly upregulate *A3B* mRNA levels, at least in these two different normal-like cell types where *A3B* is induced by TAG and truncT. However, the observed upregulation at the protein level in RBL1-depleted cells raises the possibility of an additional layer of regulation that may be posttranscriptional. These results also suggest that truncT may have at least one additional activity mediated by its LXCXE motif that contributes to *A3B* upregulation.



**Figure 5.** Palbociclib treatment of cancer cell lines and A3B expression. **(A)** RT-qPCR quantification of *A3B* and *CCNE2* mRNA expression in the indicated cell lines treated with 0, 0.5, or 2.5  $\mu$ M palbociclib for 0, 1, or 4 h. **(B)** RT-qPCR quantification of *A3B* and *CCNE2* expression in MCF10A cells pretreated with 0 ng/ml or 25 ng/ml PMA for 3 h prior to treatment with 0, 0.5, or 2.5  $\mu$ M palbociclib for 0, 1, and 3 h. **(C)** RT-qPCR quantification of *A3B* and *CCNE2* expression in MCF10A cells pretreated with 0, 0.5, or 2.5  $\mu$ M palbociclib for 30 min and then treated with 0 ng/ml or 25 ng/ml PMA for 0, 1, and 3 h.

## Pharmacological inhibition of CDK4/6 does not alter APOBEC3B expression

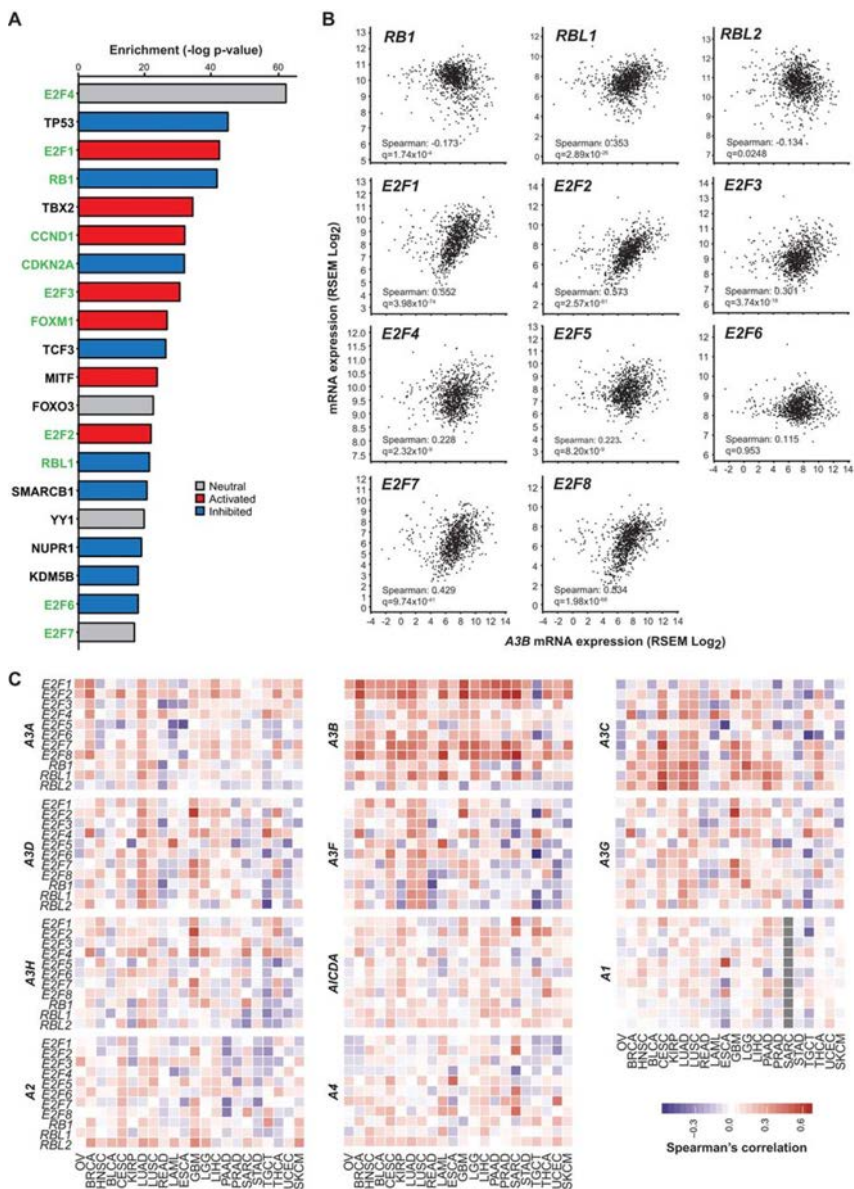
Palbociclib is a selective inhibitor of CDK4 and CDK6, which are kinases that function normally to phosphorylate pRB, prevent binding to E2F transcription factors, and stimulate the expression of many genes involved in cell cycle progression (51–53). To corroborate the knockdown experiments above, we treated a panel of transformed cell lines with palbociclib and quantified mRNA expression levels over time. This panel of cell lines was constructed based on *A3B* expression, ranging from low to high (54, 55), TP53 status, and ability to phosphorylate RB. As a positive control for palbociclib efficacy, we analyzed expression of *CCNE2*, which encodes cyclin E2, promotes entry into S phase, and is a known CDK4/6-RB-regulated gene (56, 57). The majority of cell lines showed a dose- and time-responsive decrease in *CCNE2* mRNA expression (Fig. 5A). This effect was minimal in HCC1937 and HCC1599 cells, which are known to display decreased RB phosphorylation (58, 59). In contrast, none of the palbociclib-treated cell lines showed a reproducible or significant change in *A3B* mRNA expression. In addition, MCF10A cells were treated with PMA to induce *A3B* mRNA expression by the PKC/ $\kappa$ NF- $\kappa$ B pathway, and again, palbociclib had little effect (palbociclib added post- or pre-PMA addition in Fig. 5B and C). These results combine to indicate that the kinase activity of CDK4 and CDK6 is dispensable for *A3B* expression in multiple different cell lines.

## Tumor transcriptome analyses support involvement of the RB pathway in APOBEC3B regulation

We next used bioinformatics approaches to mine TCGA data and assess global correlates with *A3B* mRNA expression in human tumors. First, we conducted pathway analysis using all genes with significant positive correlations between *A3B* expression in the TCGA breast tumor cohort. This analysis revealed that 11 of the top 20 significantly enriched upstream transcription factors contributing to this expression pattern are part of the CDK4/6-cyclin D-RB-E2F axis (green-labeled genes in Fig. 6A). These regulatory factors were either significantly activated or inhibited, generally corresponding with known functions, with the net outcome being accelerated cell cycling (respectively, red and blue bars in Fig. 6A). Upon closer pairwise examination of effectors in this signal transduction pathway, *A3B* mRNA expression has the strongest positive correlations with expression of *RBL1*, *E2F1*, *E2F2*, *E2F7*, and *E2F8* (Fig. 6B).

Last, we expanded this expression correlation analysis to include 22 different tumor types in TCGA and all 11 *APOBEC* family members. This global approach further highlighted strong correlations between *A3B* and expression of *E2F1*,





**Figure 6.** Evidence for *A3B* regulation by the RB/E2F pathway in tumors. **(A)** Top 20 hits from IPA enrichment analysis of upstream transcriptional regulators of *A3B* in TCGA breast cancer with RB-pathway/cell cycle-related genes highlighted in green. Negative logs of enrichment *P* values are shown in the bar graph on the right colored by predicted activation (red) or inhibition (blue) of the specific transcription factor. **(B)** Scatter plots showing the correlation between *A3B* mRNA levels and transcription factors in the RB pathway with Spearman's correlation coefficient and *q* value reported in the lower left corner of each subpanel. **(C)** Spearman's correlation coefficient values for all APOBEC-family members against RB pathway transcription factors across 22 cancers ordered by hierarchal clustering. Negative correlations are shown in blue, positive correlations are shown in red, and no data is represented by grey.

*E2F2*, *E2F7*, and *E2F8* and indicated that the association between *A3B*, this signal transduction pathway, and the cell cycle is evident in many cancer types (Fig. 6C). Heat map intensities also indicated that *A3B* is the only *APOBEC* family member that positively and globally correlates with activation of the RB-E2F axis. To further reconcile our experimental and bioinformatics data sets, we used the meta-analysis regulatory data from TargetGeneReg (<http://www.targetgenereg.org/>) to compare the regulation of *A3B* with known cell cycle-related genes (60). These results, summarized in Table 1, further indicate that the *A3B* mRNA expression profile is consistent with that of a cell cycle-regulated gene that becomes upregulated during the G<sub>2</sub>/M phase, which is similar to *FOXN1* and distinct from *UNG2*, *CCNE2*, and *A3C* (the last being an *APOBEC3* family member likely to be regulated by p53).

## Discussion

In this study, we investigate the mechanism of *A3B* upregulation by polyomavirus T antigen through analyses of separation-of-function mutants, genetic knockdowns, pharmacologic treatments, and transcriptomic data. We use high-resolution fluorescence microscopy to show that polyomavirus infection causes *A3B* upregulation and protein accumulation in the nuclear compartment with the potential to access viral replication foci. Second, we show that the LXCXE motif of LTA<sub>g</sub> and truncT, which is well known to inhibit the tumor suppressor RB, is essential for *A3B* upregulation, whereas the p53-binding domain is dispensable. Further investigation into this pathway using genetic and pharmacologic treatments indicate that solely perturbing RB family members (RB, RBL1, and RBL2) or kinases responsible for their phosphorylation (CDK4 and CDK6) is insufficient to cause *A3B* mRNA upregulation. However, bioinformatics analyses of tumor expression data show strong global correlations between *A3B* mRNA expression and expression of other genes regulated by the RB-E2F signaling pathway, including several members of the E2F family of transcription factors. Additional analysis of cell cycle-regulatory networks using the TargetGeneReg database suggest that *A3B* might be a late cell cycle gene repressed by the RB/E2F family members associated with the DREAM complex in quiescence and activated by other transcription factors. Suppression by the DREAM complex is supported by the mild upregulation of *A3B* protein observed upon RBL1 knockdown in this study.

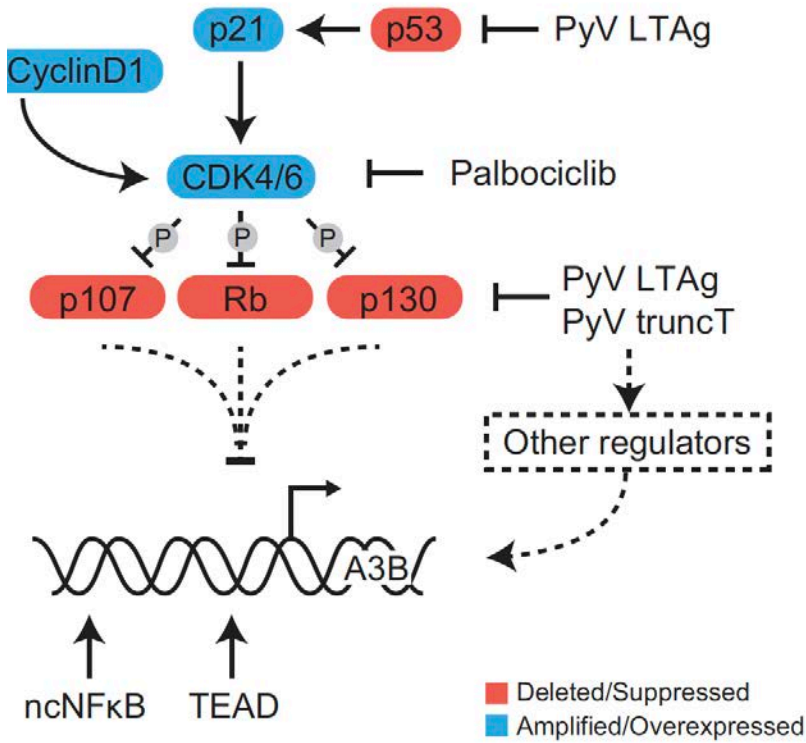


**Table 1.** Cell cycle analysis of A3B and other cell cycle-regulated genes

Type of value	Value for gene:				
	<i>APOBEC3B</i>	<i>APOBEC3C</i>	<i>CCNE2</i>	<i>FOXM1</i>	<i>UNG2</i>
Chromosome	22	22	8	12	12
p53 expression score	−4	14	−14	−17	−14
No. of cell cycle data sets	2	1	5	2	5
G <sub>1</sub> /S or G <sub>2</sub> /M	G <sub>2</sub> /M	0	G <sub>1</sub> /S	G <sub>2</sub> /M	G <sub>1</sub> /S
p53 target	No	Yes	No	No	No
Cell cycle gene	Yes	No	Yes	Yes	Yes
DREAM target	Yes	No	No	Yes	Yes
MMB-FoxM1 target	No	No	No	No	No
RB-E2F target	No	No	Yes	No	Yes

One putative activating transcription factor is FOXM1, which was significantly enriched as an upstream regulator of A3B and is known to activate genes in late G<sub>2</sub>/M (61). Interestingly, in lymphoblastoid B cells, FOXM1 has been reported to frequently co-occupy NF-κB binding sites and form protein complexes with NF-κB transcription factors, which have been implicated in A3B regulation (36, 62). Taken together, these results indicate that the RB/E2F pathway, which is commonly modified in cancer, likely contributes to A3B overexpression observed in virus infections and in different tumor types, but additional unknown signals are also likely to be required for full induction (model in Fig. 7).

Our original studies with HPV and A3B led us to propose a model in which p53 represses A3B transcription and that p53 inactivation by the viral oncoprotein E6 or by genetic mutation results in derepression of A3B transcription (23). This transcription repression model is consistent with strong correlations in tumors and cell lines between A3B overexpression and TP53 inactivation (54). A recent study confirmed these correlations and used pharmacologic and genetic approaches to provide further support for such a model (38). However, three different results in the cell-based systems presented here do not support a dominant role for p53 in A3B repression. Specifically, the p53 binding domain of BKPvV is dispensable for A3B upregulation, nutlin treatment has no effect on basal or induced A3B expression, and p53 knockout alone fails to induce A3B expression (Fig. 2 and 3). Moreover, another recent study showed that p53 inactivation renders cells more permissive for A3B mutagenesis (55). Therefore, we disfavor a transcriptional role and now favor a “tolerance model” in which p53 inactivation (genetic, epigenetic, or viral) is required for cells to be able to tolerate the increased levels of DNA damage caused by A3B overexpression. This model also explains why somatic TP53 mutations were identified as a significant global correlate with A3B overexpression in cancer (11).



**Figure 7.** Model for *A3B* transcriptional regulation. Schematic of the cell cycle-related proteins affected by T antigen and drug treatments used in this study with implications for *A3B* transcriptional regulation. Solid lines represent established processes, and dashed lines represent regulatory interactions/pathways implicated by our studies.

The RB-E2F signaling axis is one of the most frequently mutated pathways in cancer, which contributes to several hallmarks of cancer by deregulating the cell cycle (63). RB inactivation is also a common target for viral genes in order to promote the survival of infected cells. The integration and continued expression of viral oncogenes in the host genome are common characteristics of virus-associated tumors. For example, HPV-positive tumors frequently express the viral E7 oncoprotein, which also has an LXCXE motif, and this is thought to be critical for tumor development (64, 65). These tumors also tend to have a high burden of APOBEC-associated mutations on the DNA strand that serves as the template for lagging-strand replication, which is synthesized during the S phase of the cell cycle (11, 14, 17). Although some of these effects have been explained by perturbations to the p53 pathway, other pathways affected by E6 have also been shown to alter *A3B* gene expression, such as those regulated by the TEAD and ZNF384 transcription factors (23, 25, 36, 38).

It is therefore not surprising in hindsight that expression of the antiviral enzyme A3B is also induced upon disruption of the RB-E2F pathway. This idea is supported by a study showing that high-risk HPV E7 is capable of upregulating A3B (26). An additional study also found that elevated A3B expression is significantly correlated with proliferative features in breast cancer (66). Last, the LXCXE motif may be acting through another cellular signaling pathway. For instance, the LXCXE of various viral proteins triggers activation of the antiviral cGAS-STING pathway, which could also contribute to A3B transcriptional regulation (67).

All of these results combine to indicate that A3B transcriptional regulation is complex and governed by multiple pathways and different transcription factors. Perhaps linkage to the cell cycle evolved to prevent potentially oncogenic mutations of the host genome during normal cellular DNA replication or, alternatively, to maximize antiviral responses during particularly susceptible cell cycle stages. Further experiments using mutant viral oncogenes as molecular probes, such as T antigen, E6, and E7, are likely to continue to provide valuable insights into the regulation of A3B and lead to a greater understanding of its roles in tumorigenesis, virus evolution, and antiviral immunity.

## Acknowledgments

We thank James A. DeCaprio, Diako Ebrahimi, and N. Alpay Temiz for valuable discussions and comments.

## Funding

This work was supported in part by Biogen and by grants from the National Institutes of Health (NIAID R01 AI123162 to M.J., NCI R21 CA206309 to R.S.H., and NIAID R37 to R.S.H.). J.W.M.M., P.N.S., and R.S.H. received funding from the Dutch Cancer Society (KWF grant no. EMCR-2016-10270). Partial salary support for T.A.S. was provided by the Susan G. Komen Foundation, support for G.J.S. and J.L.M. was provided by a Graduate Research Fellowship from the National Science Foundation, and support for M.C.J. was provided by a training grant from the NCI (T32 CA009138). R.S.H. is the Margaret Harvey Schering Land Grant Chair for Cancer Research, a Distinguished McKnight University Professor, and an Investigator of the Howard Hughes Medical Institute.

## Conflicts of Interest

R.S.H. is a cofounder, shareholder, and consultant for ApoGen Biotechnologies Inc. Biogen sponsored a portion of the work in the Harris laboratory.

## References

1. Escalera-Zamudio M, Rojas-Anaya E, Kolokotronis SO, Taboada B, Loza-Rubio E, Méndez-Ojeda ML, Arias CF, Osterrieder N, Greenwood AD. 2016. Bats, primates, and the evolutionary origins and diversification of mammalian gammaherpesviruses. *mBio* 7:e01425-16. <https://doi.org/10.1128/mBio.01425-16>.
2. Cuevas JM, Geller R, Garijo R, López-Aldeguer J, Sanjuán R. 2015. Extremely high mutation rate of HIV-1 *in vivo*. *PLoS Biol* 13:e1002251. <https://doi.org/10.1371/journal.pbio.1002251>.
3. Buck CB, Van Doorslaer K, Peretti A, Geoghegan EM, Tisza MJ, An P, Katz JP, Pipas JM, McBride AA, Camus AC, McDermott AJ, Dill JA, Delwart E, Ng TFF, Farkas K, Austin C, Kraberger S, Davison W, Pastrana DV, Varsani A. 2016. The ancient evolutionary history of polyomaviruses. *PLoS Pathog* 12:e1005574. <https://doi.org/10.1371/journal.ppat.1005574>.
4. Duffy S, Shackelton LA, Holmes EC. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet* 9:267–276. <https://doi.org/10.1038/nrg2323>.
5. Sanjuán R, Domingo-Calap P. 2016. Mechanisms of viral mutation. *Cell Mol Life Sci* 73:4433–4448. <https://doi.org/10.1007/s00018-016-2299-6>.
6. Verhalen B, Starrett GJ, Harris RS, Jiang M. 2016. Functional upregulation of the DNA cytosine deaminase APOBEC3B by polyomaviruses. *J Virol* 90:6379–6386.
7. Warren CJ, Van Doorslaer K, Pandey A, Espinosa JM, Pyeon D. 2015. Role of the host restriction factor APOBEC3 on papillomavirus evolution. *Virus Evol* 1:vev015.
8. Wang Z, Wakae K, Kitamura K, Aoyama S, Liu G, Koura M, Monjurul AM, Kukimoto I, Muramatsu M. 2014. APOBEC3 deaminases induce hypermutation in human papillomavirus 16 DNA upon beta interferon stimulation. *J Virol* 88:1308–1317. <https://doi.org/10.1128/JVI.03091-13>.
9. Vartanian J, Guétard D, Henry M, Wain-Hobson S. 2008. Evidence for editing of human papillomavirus DNA by APOBEC3 in benign and precancerous lesions. *Science* 320:230–233. <https://doi.org/10.1126/science.1153201>.
10. Peretti A, Geoghegan EM, Pastrana DV, Smola S, Feld P, Sauter M, Lohse S, Ramesh M, Lim ES, Wang D, Borgogna C, FitzGerald PC, Bliskovsky V, Starrett GJ, Law EK, Harris RS, Killian JK, Zhu J, Pineda M, Meltzer PS, Boldorini R, Gariglio M, Buck CB. 2018. Characterization of BK polyomaviruses from kidney transplant recipients suggests a role for APOBEC3 in driving in-host virus evolution. *Cell Host Microbe* 23:628–635.e7. <https://doi.org/10.1016/j.chom.2018.04.005>.
11. Burns MB, Temiz NA, Harris RS. 2013. Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat Genet* 45:977–983. <https://doi.org/10.1038/ng.2701>.
12. Shi K, Carpenter MA, Banerjee S, Shaban NM, Kurahashi K, Salamango DJ, McCann JL, Starrett GJ, Duffy JV, Demir Ö, Amaro RE, Harki DA, Harris RS, Aihara H. 2017. Structural basis for targeted DNA cytosine deamination and mutagenesis by APOBEC3A and APOBEC3B. *Nat Struct Mol Biol* 24:131–139. <https://doi.org/10.1038/nsmb.3344>.
13. Kouno T, Silvas TV, Hilbert BJ, Shandilya SMD, Bohn MF, Kelch BA, Royer WE, Somasundaran M, Kurt Yilmaz N, Matsuo H, Schiffer CA. 2017. Crystal structure of APOBEC3A bound to single-stranded DNA reveals structural basis for cytidine deamination and specificity. *Nat Commun* 8:15024. <https://doi.org/10.1038/ncomms15024>.
14. Seplyarskiy VB, Soldatov RA, Popadin KY, Antonarakis SE, Bazykin GA, Nikolaev SI. 2016. APOBEC-induced mutations in human cancers are strongly enriched on the lagging DNA strand during replication. *Genome Res* 26:174–182. <https://doi.org/10.1101/gr.197046.115>.

15. Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, Hess JM, Rheinbay E, Kim J, Maruvka YE, Braunstein LZ, Kamburov A, Hanawalt PC, Wheeler DA, Koren A, Lawrence MS, Getz G. 2016. Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell* 164:538–549. <https://doi.org/10.1016/j.cell.2015.12.050>.
16. Morganella S, Alexandrov LB, Glodzik D, Zou X, Davies H, Staaf J, Sieuwerts AM, Brinkman AB, Martin S, Ramakrishna M, Butler A, Kim H-Y, Borg Å, Sotiriou C, Futreal PA, Campbell PJ, Span PN, Van Laere S, Lakhani SR, Eyfjord JE, Thompson AM, Stunnenberg HG, van de Vijver MJ, Martens JWM, Børresen-Dale A L, Richardson AL, Kong G, Thomas G, Sale J, Rada C, Stratton MR, Birney E, Nik-Zainal S. 2016. The topography of mutational processes in breast cancer genomes. *Nat Commun* 7:11383. <https://doi.org/10.1038/ncomms11383>.
17. Hoopes JI, Cortez LM, Mertz TM, Malc EP, Mieczkowski PA, Roberts SA. 2016. APOBEC3A and APOBEC3B preferentially deaminate the lagging strand template during DNA replication. *Cell Rep* 14:1273–1282. <https://doi.org/10.1016/j.celrep.2016.01.021>.
18. Bhagwat AS, Hao W, Townes JP, Lee H, Tang H, Foster PL. 2016. Strand-biased cytosine deamination at the replication fork causes cytosine to thymine mutations in *Escherichia coli*. *Proc Natl Acad Sci U S A* 113: 2176–2181. <https://doi.org/10.1073/pnas.1522325113>.
19. Harris RS, Dudley JP. 2015. APOBECs and virus restriction. *Virology* 479-480:131–145. <https://doi.org/10.1016/j.virol.2015.03.012>.
20. Conticello SG. 2008. The AID/APOBEC family of nucleic acid mutators. *Genome Biol* 9:229. <https://doi.org/10.1186/gb-2008-9-6-229>.
21. Vieira VC, Soares MA. 2013. The role of cytidine deaminases on innate immune responses against human viral infections. *Biomed Res Int* 2013:1. <https://doi.org/10.1155/2013/683095>.
22. Venkatesan S, Rosenthal R, Kanu N, McGranahan N, Bartek J, Quezada SA, Hare J, Harris RS, Swanton C. 2018. Perspective: APOBEC mutagenesis in drug resistance and immune escape in HIV and cancer evolution. *Ann Oncol* 29:563–572. <https://doi.org/10.1093/annonc/mdy003>.
23. Vieira VC, Leonard B, White EA, Starrett GJ, Temiz NA, Lorenz LD, Lee D, Soares MA, Lambert PF, Howley PM, Harris RS, Harris S. 2014. Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *mBio* 5:e02234-14. <https://doi.org/10.1128/mBio.02234-14>.
24. Mori S, Takeuchi T, Ishii Y, Kukimoto I. 2015. Identification of APOBEC3B promoter elements responsible for activation by human papillomavirus type 16 E6. *Biochem Biophys Res Commun* 460:555–560. <https://doi.org/10.1016/j.bbrc.2015.03.068>.
25. Mori S, Takeuchi T, Ishii Y, Yugawa T, Kiyono T, Nishina H, Kukimoto I. 2017. Human papillomavirus 16 E6 upregulates APOBEC3B via the TEAD transcription factor. *J Virol* 91:e02413-16. <https://doi.org/10.1128/JVI.02413-16>.
26. Warren CJ, Xu T, Guo K, Griffin LM, Westrich JA, Lee D, Lambert PF, Santiago ML, Pyeon D. 2015. APOBEC3A functions as a restriction factor of human papillomavirus. *J Virol* 89:688–702. <https://doi.org/10.1128/JVI.02383-14>.
27. Barbosa MS, Edmonds C, Fisher C, Schiller JT, Lowy DR, Vousden KH. 1990. The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. *EMBO J* 9:153–160.
28. Mietz JA, Unger T, Huibregtse JM, Howley PM. 1992. The transcriptional transactivation function of wild-type p53 is inhibited by SV40 large T-antigen and by HPV-16 E6 oncoprotein. *EMBO J* 11:5013–5020.
29. Low J, Humes HD, Szczypka M, Imperiale M. 2004. BKV and SV40 infection of human kidney tubular epithelial cells *in vitro*. *Virology* 323:182–188. <https://doi.org/10.1016/j.virol.2004.03.027>.

30. Abend JR, Joseph AE, Das D, Campbell-Cecen DB, Imperiale MJ. 2009. A truncated T antigen expressed from an alternatively spliced BK virus early mRNA. *J Gen Virol* 90:1238–1245. <https://doi.org/10.1099/vir.0.009159-0>.
31. DeCaprio JA, Garcea RL. 2013. A cornucopia of human polyomaviruses. *Nat Rev Microbiol* 11:264–276. <https://doi.org/10.1038/nrmicro2992>.
32. Ludlow JW, DeCaprio JA, Huang CM, Lee WH, Paucha E, Livingston DM. 1989. SV40 large T antigen binds preferentially to an underphosphorylated member of the retinoblastoma susceptibility gene product family. *Cell* 56:57–65. [https://doi.org/10.1016/0092-8674\(89\)90983-5](https://doi.org/10.1016/0092-8674(89)90983-5).
33. Christensen JB, Imperiale MJ. 1995. Inactivation of the retinoblastoma susceptibility protein is not sufficient for the transforming function of the conserved region 2-like domain of simian virus 40 large T antigen. *J Virol* 69:3945–3948.
34. Haug T, Skorpen F, Aas PA, Malm V, Skjelbred C, Krokan HE. 1998. Regulation of expression of nuclear and mitochondrial forms of human uracil-DNA glycosylase. *Nucleic Acids Res* 26:1449–1457.
35. Law EK, Sieuwerts AM, LaPara K, Leonard B, Starrett GJ, Molan AM, Temiz NA, Vogel RI, Meijer-van Gelder ME, Sweep FCGJ, Span PN, Foekens JA, Martens JWM, Yee D, Harris RS. 2016. The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. *Sci Adv* 2:e1601737. <https://doi.org/10.1126/sciadv.1601737>.
36. Leonard B, McCann JL, Starrett GJ, Kosyakovskiy L, Luengas EM, Molan AM, Burns MB, McDougale RM, Parker PJ, Brown WL, Harris RS. 2015. The PKC/NF- $\kappa$ B signaling pathway induces APOBEC3B expression in multiple human cancers. *Cancer Res* 75:4538–4547. <https://doi.org/10.1158/0008-5472.CAN-15-2171-T>.
37. Sieuwerts AM, Willis S, Burns MB, Look MP, Meijer-Van Gelder ME, Schlicker A, Heideman MR, Jacobs H, Wessels L, Leyland-Jones B, Gray KP, Foekens JA, Harris RS, Martens JWM. 2014. Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. *Horm Cancer* 5:405–413. <https://doi.org/10.1007/s12672-014-0196-8>.
38. Periyasamy M, Singh AK, Gemma C, Kranjec C, Farzan R, Leach DA, Navaratnam N, Pálinkás HL, Vertessy BG, Fenton TR, Doorbar J, Fuller-Pace F, Meek DW, Coombes RC, Buluwela L, Ali S. 2017. P53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. *Nucleic Acids Res* 45:11056–11069. <https://doi.org/10.1093/nar/gkx721>.
39. Vassilev LT. 2004. Small-molecule antagonists of p53-MDM2 binding: research tools and potential therapeutics. *Cell Cycle* 3:417–419. <https://doi.org/10.4161/cc.3.4.801>.
40. Lahav G, Rosenfeld N, Sigal A, Geva-Zatorsky N, Levine AJ, Elowitz MB, Alon U. 2004. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat Genet* 36:147–150. <https://doi.org/10.1038/ng1293>.
41. Barak Y, Juven T, Haffner R, Oren M. 1993. Mdm2 expression is induced by wild type p53 activity. *EMBO J* 12:461–468.
42. El-Deiry WS, Harper JW, O'Connor PM, Velculescu VE, Canman CE, Jackman J, Pietenpol JA, Burrell M, Hill DE, Wang Y, Wiman KG, Mercer WE, Kastan MB, Kohn KW, Elledge SJ, Kinzler KW, Vogelstein B. 1994. WAF1/CIP1 is induced in p53-mediated G<sub>1</sub> arrest and apoptosis. *Cancer Res* 54:1169–1174.
43. Seoane J, Le H-V, Massagué J. 2002. Cell culture Myc suppression of the p21 Cip1 Cdk inhibitor influences the outcome of the p53 response to DNA damage. *Nature* 419:729–734. <https://doi.org/10.1038/nature01119>.
44. Chen D, Tavana O, Chu B, Erber L, Chen Y, Baer R, Gu W. 2017. NRF2 is a major target of ARF in p53-independent tumor suppression. *Mol Cell* 68:224–232.e4. <https://doi.org/10.1016/j.molcel.2017.09.009>.
45. Jiang L, Kon N, Li T, Wang SJ, Su T, Hibshoosh H, Baer R, Gu W. 2015. Ferroptosis as a p53-mediated activity during tumour suppression. *Nature* 520:57–62. <https://doi.org/10.1038/nature14344>.

46. Whyte P, Buchkovich KJ, Horowitz JM, Friend SH, Raybuck M, Weinberg RA, Harlow E. 1988. Association between an oncogene and an antioncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. *Nature* 336:124–129. <https://doi.org/10.1038/334124a0>.
47. Caracciolo V, Reiss K, Khalili K, De Falco G, Giordano A. 2006. Role of the interaction between large T antigen and Rb family members in the oncogenicity of JC virus. *Oncogene* 25:5294–5301. <https://doi.org/10.1038/sj.onc.1209681>.
48. Henley SA, Dick FA. 2012. The retinoblastoma family of proteins and their regulatory functions in the mammalian cell division cycle. *Cell Div* 7:10. <https://doi.org/10.1186/1747-1028-7-10>.
49. Classon M, Dyson N. 2001. P107 and P130: versatile proteins with interesting pockets. *Exp Cell Res* 264:135–147. <https://doi.org/10.1006/excr.2000.5135>.
50. Harris KF, Christensen JB, Imperiale MJ. 1996. BK virus large T antigen: interactions with the retinoblastoma family of tumor suppressor proteins and effects on cellular growth control. *J Virol* 70:2378–2386.
51. Whittaker SR, Mallinger A, Workman P, Clarke PA. 2017. Inhibitors of cyclin-dependent kinases as cancer therapeutics. *Pharmacol Ther* 173:83–105. <https://doi.org/10.1016/j.pharmthera.2017.02.008>.
52. Asghar U, Witkiewicz AK, Turner NC, Knudsen ES. 2015. The history and future of targeting cyclin dependent kinases in cancer therapy. *Nat Rev Drug Discov* 14:130–146. <https://doi.org/10.1038/nrd4504>.
53. O'Leary B, Finn RS, Turner NC. 2016. Treating cancer with selective CDK4/6 inhibitors. *Nat Rev Clin Oncol* 13:417–430. <https://doi.org/10.1038/nrclinonc.2016.26>.
54. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, Yee D, Temiz NA, Donohue DE, McDougale RM, Brown WL, Law EK, Harris RS. 2013. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 494:366–370. <https://doi.org/10.1038/nature11881>.
55. Nikkilä J, Kumar R, Campbell J, Brandsma I, Pemberton HN, Wallberg F, Nagy K, Scheer I, Vertessy BG, Serebrenik AA, Monni V, Harris RS, Pettitt SJ, Ashworth A, Lord CJ. 2017. Elevated APOBEC3B expression drives a kataegic-like mutation signature and replication stress-related therapeutic vulnerabilities in p53-defective cells. *Br J Cancer* 117:113–123. <https://doi.org/10.1038/bjc.2017.133>.
56. Otto T, Sicinski P. 2017. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer* 17:93–115. <https://doi.org/10.1038/nrc.2016.138>.
57. Fry DW, Harvey PJ, Keller PR, Elliott WL, Meade M, Trachet E, Albassam M, Zheng X, Leopold WR, Pryer NK, Toogood PL. 2004. Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. *Mol Cancer Ther* 3:1427–1438.
58. Raspé E, Coulonval K, Pita JM, Paternot S, Rothé F, Twyffels L, Brohée S, Craciun L, Larsimont D, Kruys V, Sandras F, Salmon I, Van Laere S, Piccart M, Ignatiadis M, Sotiriou C, Roger PP. 2017. CDK4 phosphorylation status and a linked gene expression profile predict sensitivity to palbociclib. *EMBO Mol Med* 9:1052–1066. <https://doi.org/10.15252/emmm.201607084>.
59. Stoeck A, Lejnine S, Truong A, Pan L, Wang H, Zang C, Yuan J, Ware C, MacLean J, Garrett-Engle PW, Kluk M, Laskey J, Haines BB, Moskaluk C, Zawel L, Fawell S, Gilliland G, Zhang T, Kremer BE, Knoechel B, Bernstein BE, Pear WS, Liu XS, Aster JC, Sathyanarayanan S. 2014. Discovery of biomarkers predictive of GSI response in triple-negative breast cancer and adenoid cystic carcinoma. *Cancer Discov* 4:1154–1167. <https://doi.org/10.1158/2159-8290.CD-13-0830>.
60. Fischer M, Grossmann P, Padi M, DeCaprio JA. 2016. Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F target gene analyses identifies cell cycle gene regulatory networks. *Nucleic Acids Res* 44:6070–6086. <https://doi.org/10.1093/nar/gkw523>.



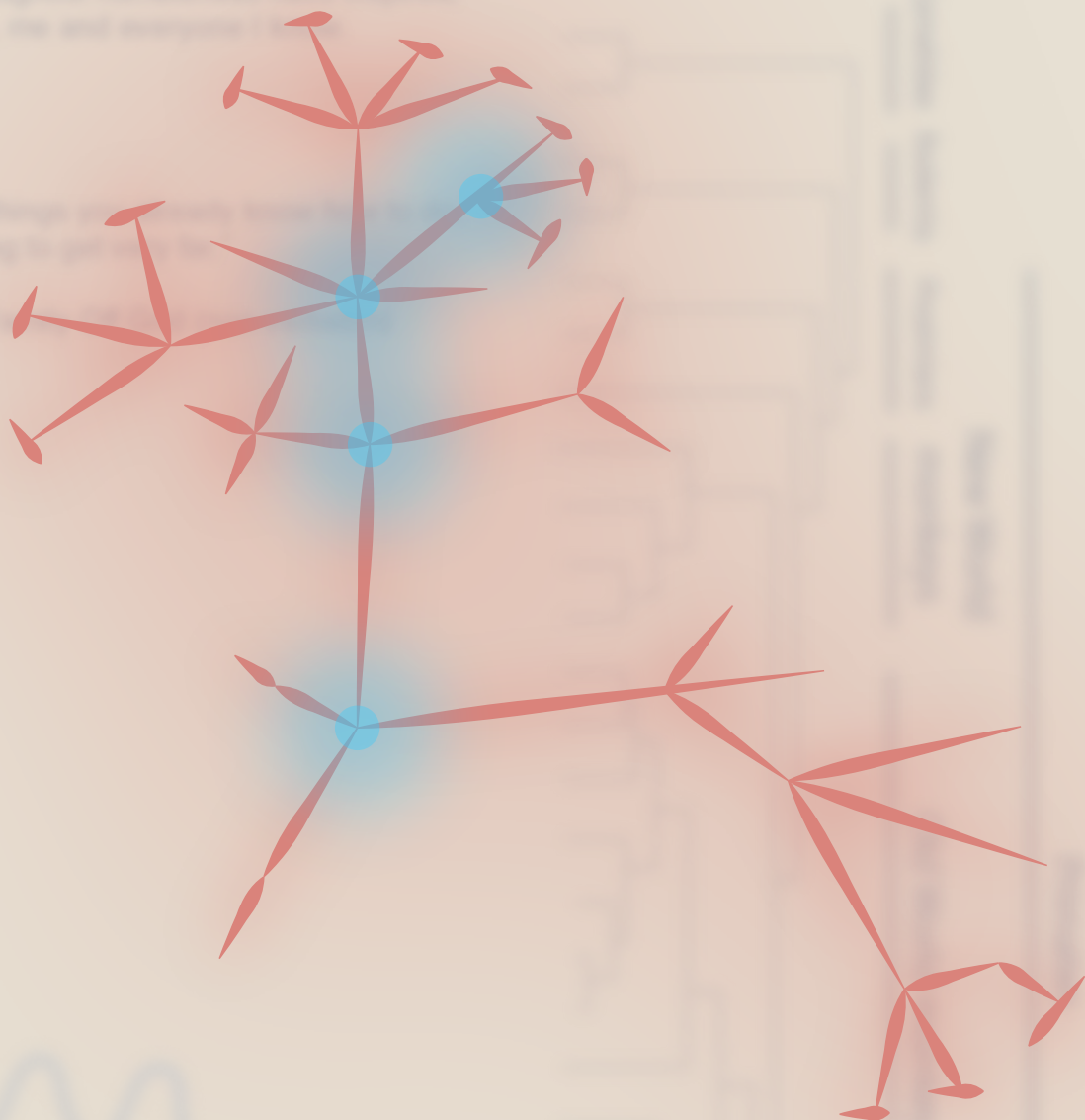
61. Laoukili J, Kooistra MRH, Brás A, Kauw J, Kerkhoven RM, Morrison A, Clevers H, Medema RH. 2005. FoxM1 is required for execution of the mitotic programme and chromosome stability. *Nat Cell Biol* 7:126–136. <https://doi.org/10.1038/ncb1217>.
62. Zhao B, Barrera LA, Ersing I, Willox B, Schmidt SCS, Greenfeld H, Zhou H, Mollo SB, Shi TT, Takasaki K, Jiang S, Cahir-McFarland E, Kellis M, Bulyk ML, Kieff E, Gewurz BE. 2014. The NF- B genomic landscape in lymphoblastoid B cells. *Cell Rep* 8:1595–1606. <https://doi.org/10.1016/j.celrep.2014.07.037>.
63. Sherr CJ, Beach D, Shapiro GI. 2016. Targeting CDK4 and CDK6: from discovery to therapy. *Cancer Discov* 6:353–367. <https://doi.org/10.1158/2159-8290.CD-15-0894>.
64. Mirabello L, Yeager M, Yu K, Clifford GM, Xiao Y, Zhu B, Cullen M, Boland JF, Wentzensen N, Nelson CW, Raine-Bennett T, Chen Z, Bass S, Song L, Yang Q, Steinberg M, Burdett L, Dean M, Roberson D, Mitchell J, Lorey T, Franceschi S, Castle PE, Walker J, Zuna R, Kreimer AR, Beachler DC, Hildesheim A, Gonzalez P, Porras C, Burk RD, Schiffman M. 2017. HPV16 E7 genetic conservation is critical to carcinogenesis. *Cell* 170:1164–1174.e6. <https://doi.org/10.1016/j.cell.2017.08.001>.
65. Zhen S, Hua L, Takahashi Y, Narita S, Liu YH, Li Y. 2014. *In vitro* and *in vivo* growth suppression of human papillomavirus 16-positive cervical cancer cells by CRISPR/Cas9. *Biochem Biophys Res Commun* 450:1422–1426. <https://doi.org/10.1016/j.bbrc.2014.07.014>.
66. Cescon DW, Haibe-Kains B, Mak TW. 2015. APOBEC3B expression in breast cancer reflects cellular proliferation, while a deletion polymorphism is associated with immune activation. *Proc Natl Acad Sci U S A* 112:2841–2846. <https://doi.org/10.1073/pnas.1424869112>.
67. Lau L, Gray EE, Brunette RL, Stetson DB. 2015. DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. *Science* 350:568–571. <https://doi.org/10.1126/science.aab3291>.
68. Knowles WA, Sharp IR, Efstratiou L, Hand JF, Gardner SD. 1991. Preparation of monoclonal antibodies to JC virus and their use in the diagnosis of progressive multifocal leucoencephalopathy. *J Med Virol* 34:127–131. <https://doi.org/10.1002/jmv.1890340211>.
69. Leonard B, Starrett GJ, Maurer MJ, Oberg A, Van Bockstal M, Van Dorpe J, De Wever O, Helleman J, Sieuwerts AM, Berns EMJJ, Martens JWM, Anderson B, Brown WL, Kalli KR, Kaufmann SH, Harris RS. 2016. APOBEC3G expression correlates with T cell infiltration and improved clinical outcomes in high-grade serous ovarian carcinoma. *Clin Cancer Res* 22:4746–4755. <https://doi.org/10.1158/1078-0432.CCR-15-2910>.
70. Jiang M, Abend JR, Tsai B, Imperiale MJ. 2009. Early events during BK virus entry and disassembly. *J Virol* 83:1350–1358. <https://doi.org/10.1128/JVI.02169-08>.
71. Kadaja M, Isok-Paas H, Laos T, Ustav E, Ustav M. 2009. Mechanism of genomic instability in cells infected with the high-risk human papillomaviruses. *PLoS Pathog* 5:e1000397. <https://doi.org/10.1371/journal.ppat.1000397>.
72. Jiang M, Entezami P, Gamez M, Stamminger T, Imperiale MJ. 2011. Functional reorganization of promyelocytic leukemia nuclear bodies during BK virus infection. *mBio* 2:e00281-10.



A dedication to my parents,  
for delightful memories have inspired  
me, me and everyone I love.

For the things we have done together  
and going to get it done.

With love and respect.



# 3

## Characterization of the mechanism by which the RB/E2F pathway controls expression of the cancer genomic DNA deaminase APOBEC3B

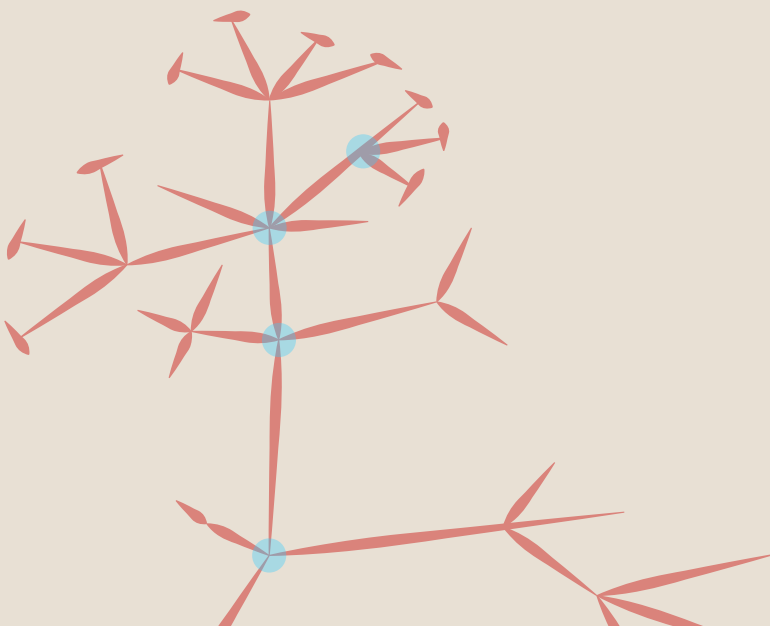
---

Pieter A Roelofs<sup>1,2</sup>, Chai Yeen Goh<sup>3†</sup>, Boon Haow Chua<sup>3,4†</sup>, Matthew C Jarvis<sup>1</sup>, Teneale A Stewart<sup>1,5</sup>, Jennifer L McCann<sup>1,6</sup>, Rebecca M McDougale<sup>1,7</sup>, Michael A Carpenter<sup>1,6</sup>, John WM Martens<sup>8</sup>, Paul N Span<sup>2</sup>, Dennis Kappej<sup>3,4</sup>, Reuben S Harris<sup>1,6\*</sup>

1. Department of Biochemistry, Molecular Biology and Biophysics, Masonic Cancer Center, Institute for Molecular Virology, Center for Genome Engineering, University of Minnesota, Minneapolis, United States
2. Department of Radiation Oncology, Radboud University Medical Center, Nijmegen, Netherlands
3. Cancer Science Institute of Singapore, National University of Singapore, Singapore, Singapore
4. Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore
5. Mater Research Institute, The University of Queensland, Faculty of Medicine, Brisbane, Australia
6. Howard Hughes Medical Institute, University of Minnesota, Minneapolis, United States;
7. Hennepin Healthcare, Minneapolis, United States
8. Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, Netherlands

## Abstract

APOBEC3B (A3B)-catalyzed DNA cytosine deamination contributes to the overall mutational landscape in breast cancer. Molecular mechanisms responsible for *A3B* upregulation in cancer are poorly understood. Here we show that a single E2F cis-element mediates repression in normal cells and that expression is activated by its mutational disruption in a reporter construct or the endogenous *A3B* gene. The same E2F site is required for *A3B* induction by polyomavirus T antigen indicating a shared molecular mechanism. Proteomic and biochemical experiments demonstrate the binding of wildtype but not mutant E2F promoters by repressive PRC1.6/E2F6 and DREAM/E2F4 complexes. Knockdown and overexpression studies confirm the involvement of these repressive complexes in regulating *A3B* expression. Altogether, these studies demonstrate that *A3B* expression is suppressed in normal cells by repressive E2F complexes and that viral or mutational disruption of this regulatory network triggers overexpression in breast cancer and provides fuel for tumor evolution.



## Introduction

Cancer is a collection of diseases characterized by a complex array of mutations ranging from gross chromosomal abnormalities to single-base substitution (SBS) mutations. Over the last decade, analyses of thousands of tumor genome sequences have confirmed this complexity and also, importantly, revealed common patterns or signatures indicative of the sources of DNA damage that led to these observed mutations (most recent pan-cancer analysis by Alexandrov et al., 2020; reviewed by Helleday et al., 2014; Roberts and Gordenin, 2014; Swanton et al., 2015; Venkatesan et al., 2018). One of the most prominent SBS mutation signatures to emerge is attributable to members of the APOBEC family of single-stranded (ss)DNA cytosine deaminases (Alexandrov et al., 2013; Burns et al., 2013a; Burns et al., 2013b; Nik-Zainal et al., 2012; Roberts et al., 2013). Breast, lung, head/neck, cervical, and bladder cancers often have strong APOBEC signatures and subsets of other cancer types have weaker APOBEC contributions. The APOBEC mutation signature consists of C-to-T transitions and C-to-G transversions occurring at cytosine nucleobases in 5'-TCW motifs (W = A or T; SBS2 and SBS13), respectively (Alexandrov et al., 2020; Alexandrov et al., 2013; Nik-Zainal et al., 2016).

The human APOBEC family has nine active family members: APOBEC1, AID, and APOBEC3A/B/ C/D/F/G/H (reviewed by Green and Weitzman, 2019; Harris and Dudley, 2015; Ito et al., 2020; Olson et al., 2018; Silvas and Schiffer, 2019; Simon et al., 2015; Siriwardena et al., 2016). Although several APOBEC3s have been implicated in cancer mutagenesis including APOBEC3A (A3A) and APOBEC3H (A3H) (Chan et al., 2015; Nik-Zainal et al., 2014; Starrett et al., 2016; Taylor et al., 2013), a particularly strong case can be made for APOBEC3B (A3B). First, A3B is overexpressed in a large fraction of tumors (Burns et al., 2013a; Burns et al., 2013b; Ng et al., 2019; Roberts et al., 2013). Second, A3B is the only deaminase family member localizing to the nuclear compartment (Bogerd et al., 2006; Burns et al., 2013a; Lackey et al., 2012; Lackey et al., 2013; Pak et al., 2011; Salamango et al., 2018; Stenglein et al., 2008). Third, A3B overexpression triggers strong DNA damage responses and overt cytotoxicity (Burns et al., 2013a; Nikkilä et al., 2017; Serebrenik et al., 2019; Taylor et al., 2013; Yamazaki et al., 2020). Fourth, A3B expression correlates positively with APOBEC signature mutation loads in breast cancer (Burns et al., 2013a), and its overexpression associates with branched evolution in breast and lung cancer (de Bruin et al., 2014; Lee et al., 2019; Roper et al., 2019). Fifth, A3B expression is induced by human papillomavirus (HPV) and polyomavirus (PyV) infections, which relates to the fact that cervical, head/neck, and bladder cancers have high proportions of APOBEC signature mutations (Gillison et al., 2019; Henderson et al., 2014; Starrett

et al., 2019; Verhalen et al., 2016; Vieira et al., 2014). Last, A3B overexpression associates with poor clinical outcomes including drug resistance and metastasis (Glaser et al., 2018; Law et al., 2016; Serebrenik et al., 2020; Sieuwerts et al., 2017; Sieuwerts et al., 2014; Walker et al., 2015; Xu et al., 2015; Yamazaki et al., 2019; Yan et al., 2016). However, in a different subset of cancer types, A3B has been shown to exert genotoxic stress that sensitizes tumor cells to DNA damaging chemotherapies (Glaser et al., 2018; Serebrenik et al., 2020).

The importance of A3B in cancer mutagenesis has stimulated interest in understanding the mechanisms by which this DNA mutator becomes overexpressed in tumors. A variety of stimuli have been shown to trigger transcriptional upregulation of endogenous *A3B* including small molecules, DNA damaging agents, and viral infections. Phorbol myristic acid (PMA) and lymphotoxin-b induce *A3B* by activating the protein kinase C (PKC) and non-canonical (nc)NF- $\kappa$ B signal transduction pathways (Leonard et al., 2015; Lucifora et al., 2014). Canonical NF- $\kappa$ B activation also leads to *A3B* upregulation (Maruyama et al., 2016) suggesting a mechanistic linkage between inflammatory responses and cancer mutagenesis. Various DNA damaging agents also stimulate *A3B* expression including hydroxyurea, gemcitabine, aphidicolin, and camptothecin (Kanu et al., 2016; Yamazaki et al., 2020). Interestingly, as alluded above, HPV infection induces *A3B* expression by mechanisms requiring the viral E6 and E7 oncoproteins (Mori et al., 2015; Mori et al., 2017; Starrett et al., 2019; Verhalen et al., 2016; Vieira et al., 2014; Warren et al., 2015; Westrich et al., 2018). E6 appears to induce *A3B* in part by recruiting the transcription factor TEAD4 to promoter sequences (Mori et al., 2015; Mori et al., 2017). JC and BK PyV upregulate *A3B* transcription by a mechanism requiring the LxCxE motif of the viral large T antigen (TA<sub>g</sub>; Starrett et al., 2019; Verhalen et al., 2016). HPV E7 also has a LxCxE motif suggesting a shared mechanism in which these viral oncoproteins may activate *A3B* transcription by antagonizing the canonical retinoblastoma tumor suppressor protein RB1 and the related pocket proteins RB-like 1 (RBL1) and RBL2 (reviewed by An et al., 2012; Bellacchio and Paggi, 2013; DeCaprio, 2014; DeCaprio and Garcea, 2013; Rashid et al., 2015). Viral inactivation of RB1 and RBL1/2 alters interactions with cellular E2F transcription factors and contributes to an accelerated cell cycle with dampened checkpoints. The RB/E2F axis is also frequently disrupted in non-viral cancers such as breast cancer, HPV-negative head/neck cancer, and lung cancer (Cancer Genome Atlas Network, 2012; Ertel et al., 2010; Nik-Zainal et al., 2016; Cancer Genome Atlas Network, 2015).

Central to the human RB/E2F axis are eight distinct E2F transcription factors (reviewed by Cao et al., 2010; Fischer and Müller, 2017; Sadasivam and DeCaprio, 2013). E2F1, E2F2, and E2F3 bind target promoters and recruit additional activating proteins to stimulate the expression of cell cycle genes during G<sub>1</sub>/S. RB1 binds the transactivation domain of these E2Fs and thereby prevents the recruitment of transcription activating factors. E2F4 and E2F5 form complexes with RBL1 or RBL2 and further associate with the MuvB complex, which includes LIN9, LIN37, LIN52, LIN54, and RBBP4. This bipartite assembly, known as the DREAM complex, represses transcription during the G<sub>0</sub> and early G<sub>1</sub> phases of the cell cycle (Litovchick et al., 2011; Litovchick et al., 2007; Pilkinton et al., 2007). Endogenous Cyclin/CDK complexes, as well as HPV E7 and PyV TAG through LxCxE motifs, dissociate RBL1 and RBL2 from E2F4 and E2F5 and thereby activate transcription (reviewed by An et al., 2012; Bellacchio and Paggi, 2013; DeCaprio, 2014; DeCaprio and Garcea, 2013; Rashid et al., 2015). E2F6, E2F7, and E2F8 exert their repressive function independent of RB1, RBL1, and RBL2 (Christensen et al., 2005; de Bruin et al., 2003; Trimarchi et al., 1998). E2F6 functions in the Polycomb Repressive Complex (PRC)1.6 complex to repress gene expression during G<sub>1</sub>-S (Qin et al., 2012; Scelfo et al., 2019; Stielow et al., 2018). The PRC1.6 complex consists of MGA, L3MBTL2, PCGF6, WDR5, E2F6, and TFDP1 (among other proteins), and directly binds DNA through MGA, L3MBTL2, and E2F6 (Stielow et al., 2018). Finally, E2F7 and E2F8 repress genes through the S-phase and prevent gene reactivation during the next cell cycle (Cuitin˜o et al., 2019).

Our previous studies showed that *A3B* expression is low in normal tissues (Burns et al., 2013a; Refsland et al., 2010) and inducible upon PyV TAG expression (Starrett et al., 2019; Verhalen et al., 2016). *A3B* induction by TAG may occur through the RB/E2F axis, as alluded above, or through a different LxCxE-dependent mechanism. The feasibility of such an alternative mechanism is supported by evidence that LxCxE is a common motif for protein-protein interactions and that HPV E7 uses this motif to bind >100 cellular proteins in addition to RB1, RBL1, and RBL2 (White et al., 2012). Here a series of molecular, biochemical, proteomic, and genomic approaches are used to distinguish between these molecular mechanisms. The combined results demonstrate the functionality of a single E2F binding site in the *A3B* promoter and reveal overlapping roles for both E2F4-based DREAM and E2F6-based PRC1.6 complexes in repressing *A3B* transcription in non-tumorigenic cells. Loss of this *A3B* repression mechanism in tumor cells is likely to promote cancer mutagenesis.

# Materials and Methods

## Cell lines and culture conditions

All cell lines were cultured at 37°C under 5% CO<sub>2</sub>. MCF10A cells and derivative cell lines were grown in advanced DMEM/F-12 (Invitrogen) with HEPES and L-Glutamine, supplemented with 5% horse serum (Invitrogen), 20 ng/mL EGF (Peprotech), 0.5 mg/mL hydrocortisone (Sigma), 100 ng/mL cholera toxin (Sigma), 10 mg/mL recombinant human insulin (Sigma), penicillin (100 U/mL), and streptomycin (100 mg/mL). MCF7 cells were cultured as described (Law et al., 2016) for A3B-luciferase reporter assays and for proteomics in DMEM containing 10% dialyzed FBS (PAN-Biotech), penicillin (100 U/mL), and streptomycin (100 mg/mL). BT-474 and Hs578T cells were cultured in DMEM supplemented with 10% FBS (Invitrogen), penicillin (100 U/mL), and streptomycin (100 mg/mL). MDA-MB-453 and 293 T cells were cultured in RPMI supplemented with 10% FBS, penicillin (100 U/mL), and streptomycin (100 mg/mL). All cell lines tested negative for Mycoplasma using a PCR-based assay (Uphoff and Drexler, 2011). PMA (ThermoFisher) was used at 25 ng/mL for 6 hr.

Key Resources Table				
Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
cell line (homo sapiens, female)	MCF10A	ATCC	Cat#:CRL-10317 RRID:CVCL_0598	
cell line (homo sapiens, female)	MCF10A-4C10	This paper		Hemizygous for A3B.
cell line (homo sapiens, female)	MCF7	ATCC	Cat#:HTB-22 RRID:CVCL_0031	
cell line (homo sapiens, female)	BT474	ATCC	Cat#:CLR-7913 RRID:CVCL_0179	
cell line (homo sapiens, female)	Hs578T	ATCC	Cat#:HTB-126 RRID:CVCL_0332	
cell line (homo sapiens, female)	MDA-MB-453	ATCC	Cat#:HTB-131 RRID:CVCL_0418	
cell line (homo sapiens, female)	HEK 293T	ATCC	Cat#:CRL-3216 RRID:CVCL_0063	
antibody	anti-RAD51 (Rabbit monoclonal)	Abcam	Cat#:ab133534 RRID:AB_2722613	WB (1:10,000)
antibody	anti-E2F4 (Mouse monoclonal)	Santa Cruz	Cat#:sc-398543	WB (1:250) ChIP: 5 ug per 20 ug Dynabeads

**Key Resources Table (continued)**

antibody	anti-E2F6 (Rabbit polyclonal)	Abcam	Cat#:ab53061 RRID:AB_2097254	WB (1:500) ChIP: 5 ug per 20 ug Dynabeads
antibody	anti-HA (Rabbit monoclonal)	Cell Signaling	Cat#:3724 RRID:AB_1549585	WB (1:5000)
antibody	anti-Rb (Mouse monoclonal)	Santa Cruz	Cat#:sc-102 RRID:AB_628209	WB (1:300)
antibody	anti-E2F1 (Mouse monoclonal)	Santa Cruz	Cat#:sc-251 RRID:AB_627476	WB (1:1,000)
antibody	anti-E2F3 (Mouse monoclonal)	Santa Cruz	Cat#:sc-56665 RRID:AB_1122397	WB (1:800)
antibody	anti-E2F5 (Mouse monoclonal)	Santa Cruz	Cat#:sc-374268 RRID:AB_10988935	WB (1:800)
antibody	anti-E2F6 (Mouse monoclonal)	Santa Cruz	Cat#:sc-53273 RRID:AB_783163	WB (1:300)
antibody	anti-LIN9 (Mouse monoclonal)	Santa Cruz	Cat#:sc-398234	WB (1:300)
antibody	anti-Tubulin (Mouse monoclonal)	Sigma-Aldrich	Cat#:T5168 RRID:AB_477579	WB (1:20,000)
antibody	anti-A3B (Rabbit monoclonal)	NIH AIDS Reagent Program	Cat#:12397 RRID:AB_2721202	WB (1:1,000)
antibody	anti-L3MBTL2 (Rabbit polyclonal)	Active Motif	Cat#:39569 RRID:AB_2615062	ChIP: 5 ug per 20 ug Dynabeads
recombinant DNA reagent	pGL4.74 TK-RL renilla control (plasmid)	Promega	Cat#:E692A	Internal control for luciferase assays
recombinant DNA reagent	pGL3 Basic (plasmid)	Promega	Cat#:E1751	Base vector for luciferase assays
recombinant DNA reagent	pA3B-luciferase (plasmid)	pRH7889		A3B promoter (wildtype) + luciferase
recombinant DNA reagent	pLenti-lox-empty vector (plasmid)	PMID: 30723127		
recombinant DNA reagent	pLenti-lox-BKPyV tTA <sub>g</sub> (plasmid)	PMID: 30723127		
recombinant DNA reagent	pLenti-lox-BKPyV tTA <sub>g</sub> LxCxE mutant (plasmid)	PMID: 30723127		



**Key Resources Table (continued)**

recombinant DNA reagent	pLenti4/TO-mCherry-T2A-MCS (plasmid)	pRH9693	Base vector for E2F expression
recombinant DNA reagent	pLenti4/TO-mCherry-T2A-HA-E2F4 (plasmid)	pRH9739	Lentiviral vector for expression of E2F4
recombinant DNA reagent	pLenti4/TO-mCherry-T2A-HA-E2F5 (plasmid)	pRH9740	Lentiviral vector for expression of E2F5
recombinant DNA reagent	pLenti4/TO-mCherry-T2A-HA-E2F6 (plasmid)	pRH9741	Lentiviral vector for expression of E2F6
recombinant DNA reagent	pLentiCRISPR-LoxP-A3B-gRNA#1 (plasmid)	pRH9696	Lentiviral vector for expression of gRNA targeting E2F site E
recombinant DNA reagent	pLentiCRISPR-LoxP-A3B-gRNA#3 (plasmid)	pRH9698	Lentiviral vector for expression of gRNA targeting E2F site E
sequence-based reagent	Cas9-encoding modified RNA	TriLink Biotech	Cat#:L7206-100
commercial assay or kit	Dual Luciferase Reporter Assay	Promega	Cat#:E1960
commercial assay or kit	Neon Transfection System 100μL Kit	ThermoFisher	Cat#:MPK10025
software, algorithm	MaxQuant version 1.5.2.8	MaxQuant	RRID:SCR_014485
software, algorithm	Fiji	Fiji	RRID:SCR_002285
software, algorithm	GraphPad Prism 6	GraphPad	RRID:SCR_002798
software, algorithm	Image Studio	LI-COR Biosciences	RRID:SCR_015795
other	Spark Multimode Microplate Reader	Tecan	
other	Neon Transfection System	ThermoFisher	Cat#:MPK5000
other	LI-COR Odyssey FC	LI-COR	Cat#:2800
other	LightCycler 480 Instrument	Roche	Cat#:04640268001
other	EASY-nLC 1200 system	ThermoFisher	Cat#:LC140
other	Q Exactive HF mass spectrometer	ThermoFisher	

## Plasmids and site-directed mutagenesis

The integrity of all plasmids was confirmed by Sanger sequencing. Oligos used for cloning, sequencing, and site-directed mutagenesis are listed in Supplementary file 4. The pLenti-lox constructs encoding BKPyV tTag or the LxCxE mutant were described (Starrett et al., 2019). The A3B promoter sequence (-900 to +50 corresponding to chr22:39,377,504–39,378,453 of the GRCH37/hg19 assembly) was ordered as a gBlock (IDT), subjected to overhang extension PCR to add 5' KpnI and 3' NheI restriction sites, and then cut and ligated into compatibly digested pGL3-basic (Promega). Site-directed mutagenesis was done following standard procedures (Quickchange, Agilent). E2F overexpression in BT-474 was done using pLent4/TO/V5-DEST (ThermoFisher), modified to lack the V5 tag through XhoI and AgeI digestion followed by insertion of a stuffer with compatible overhangs. 5' EcoRI and 3' AgeI sites were then added to a mCherry-T2A-MCS (multiple cloning site) cassette through overhang extension PCR and ligated into the base vector using compatible overhangs, resulting in the parental pLenti4/TO-mCherry-T2A-MCS vector. Then, coding regions of E2F4 (NM\_001950.3), E2F5 (NM\_001951.3 var 1), and E2F6 (NM\_198256.3 var A) were cloned into pcDNA3.1, and a N-terminal HA-tag was inserted by site-directed mutagenesis. Finally, 5' NheI and 3' AgeI sites were added to the HA-tagged E2F sequences by overhang extension PCR, followed by ligation into compatibly digested pLenti4/TO-mCherry-T2A-MCS parental vector. Transduction of BT-474, MDA-MB-453, and Hs578T, plated at 300,000 cells per well of a six-well plate, was then performed with lentiviral particles produced in 293 T cells as described (Burns et al., 2013a; Carpenter et al., 2019; Vieira et al., 2014).

## Dual luciferase assays

MCF10A cells were plated at 50,000 cells per well and MCF7 at 100,000 cells per well in a 24-well plate, grown as described above, and transfected 24 hr later with a 1:2 ratio of plasmid cocktail and TransIT-2020 following vendor instructions (Mirus). Each transfection reaction was comprised of 250 ng luciferase reporter construct (pGL3-basic, pA3B-luciferase, or mutant derivatives), 10 ng pGL4.74 TK-RL renilla control plasmid, and 50 ng of pLenti-lox vector expressing BKPyV tTag, tTag LxCxE mutant, or empty control (Starrett et al., 2019). Lysates were prepared 48 hr later using the Dual Luciferase Reporter Assay according to manufacturer's instructions (Promega). Luminescence was detected using a Spark Multimode Microplate Reader (Tecan).

## CRISPR/Cas9-mediated editing of the A3B promoter

All sequences of oligos used during CRISPR/Cas9-mediated editing are listed in Supplementary file 4. pLenti-based CRISPR/Cas9 gene disruption was used initially

to interrogate the *A3B* promoter using established protocols (Carpenter et al., 2019). Constructs targeting the *A3B* promoter or *lacZ* as a control were made using Golden Gate ligation and lentiviral particles were produced using 293T cells (Burns et al., 2013a; Carpenter et al., 2019; Vieira et al., 2014). Transduction of 300,000 MCF10A cells per well of a six-well plate was followed 48 hr later by selection with puromycin. Individual clones were obtained by limited dilution and multi-week outgrowth. The promoter region from >6 clones per condition was amplified, cloned into pJET1.2 (ThermoFisher), and subjected to Sanger sequencing. CRISPR/Cas9-mediated HDR was used to generate MCF10A clones with precise base substitutions in the *A3B* promoter. The MCF10A *A3B* hemizygous cell line was engineered by transducing MCF10A wildtype cells with pLentiCRISPR lentiviral particles expressing a single gRNA targeting the homologous 3'UTR of *A3A* and *A3B*, treating 48 hr with puromycin, and deriving single cell clones by limiting dilution. Clones were PCR screened for alleles mimicking the natural *A3B* deletion (Kidd et al., 2007). A clone hemizygous for *A3B* was selected for precision editing of the +21 to +28 region of the *A3B* promoter. In short, 50,000 cells were transfected (Neon Transfection, Invitrogen) with 1 ng modified gRNA targeting the *A3B* promoter or *lacZ* (Synthego), 1.5 mg Cas9-encoding modified RNA (TriLink Biotech), and 6.25 pmol HDR targeting ssDNA oligo based on prior literature (Prykhodzhiy et al., 2018). The 5' and 3' terminal nucleotides of the ssDNA oligo were protected with phosphorothioates (Richardson et al., 2016). Clones were retrieved by limiting dilution 72 hr post transfection, outgrown for several weeks, and subjected to *A3B* promoter region DNA sequencing. Primers used for screening can be found in Supplementary file 4.

## Immunoblotting

For all immunoblot experiments, cells were harvested and counted using an automated cell counter (Countess, ThermoFisher). Pelleted cells were resuspended in PBS, and whole-cell protein extracts prepared by adding Laemmli reducing sample buffer followed by incubation at 98°C for 15 min. Protein expression was analyzed by immunoblot using standard laboratory techniques. Antibodies were rabbit anti-RAD51, 1:10,000 (Abcam, ab133534), mouse anti-E2F4, 1:250 (Santa Cruz, sc-398543), rabbit anti-E2F6, 1:500 (Abcam, ab53061), rabbit anti-HA, 1:5000 (Cell Signaling, C29F4), mouse anti-Rb, 1:300 (Santa Cruz, sc-102), mouse anti-E2F1, 1:1000 (Santa Cruz, sc-251), mouse anti-E2F3, 1:800 (Santa Cruz, sc-56665), mouse anti-E2F5, 1:800 (Santa Cruz, sc-374268), mouse anti-E2F6, 1:300 (Santa Cruz, sc-53273), mouse anti-LIN9, 1:300 (Santa Cruz, sc-398234) mouse anti-tubulin, 1:20,000 (Sigma-Aldrich, T5168), and rabbit anti-A3B, 1:1,000 [5210-87-13] (Brown et al., 2019). mRNA quantification: mRNA was extracted (GenElute, Sigma-Aldrich)

and cDNA was synthesized using SuperScript First-Strand RT (ThermoFisher). mRNA expression of all APOBEC3 family members and TBP was quantified by RT-qPCR with specific primers (Refsland et al., 2010) in Ssofast Supermix (Bio-Rad) using a Lightcycler (Roche). Primer sequences are listed in Supplementary file 4.

### ChIP experiments

ChIP experiments were done as described (Leonard et al., 2015) with minor modifications. A 15 cm plate with approximately 107 sub-confluent cells was used as input for each immunoprecipitation. Chromatin was crosslinked for 10 min in 1% formaldehyde and then the crosslinking reaction was quenched using 125 mM glycine. Cells were washed in PBS, concentrated by centrifugation, and lysed in 1 mL Farnham lysis buffer (5 mM PIPES pH 8.0, 85 mM KCl, 1% Igepal CA-630, supplemented with protease inhibitors). After a 15 min incubation on ice, the cell nuclei were collected by 4°C centrifugation and then incubated 30 min on ice in nuclear lysis buffer (50 mM Tris-HCl pH8.1, 10 mM EDTA, 1% SDS, supplemented with protease inhibitors). Chromatin was sheared into 200–300 bp fragments using a Misonix sonicator for 13 cycles (30' on and 45' off) at an amplitude setting of 2. Chromatin was cleared of debris by centrifugation, diluted 5x with IP dilution buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1% Igepal CA-630, 0.25% deoxycholic acid, 1 mM EDTA), and incubated with 5 mg of each mAb coupled to 20 mg Dynabeads Protein G magnetic beads (Invitrogen). Input controls of 1% were frozen down for later analysis. ChIP antibodies were mouse anti-E2F4 (Santa Cruz, sc-398543), rabbit anti-E2F6 (Abcam, ab53061), and rabbit anti-L3MBTL2 (Active Motif, 39569). After overnight incubation beads were washed twice with IP wash buffer 1 (50 mM Tris-HCl pH7.4, 150 mM NaCl, 1% Igepal CA-630, 0.25% deoxycholic acid, 1 mM EDTA), three times with IP wash buffer 2 (100 mM Tris-HCl pH 9.0, 500 mM LiCl, 1% Igepal CA-630) and once with IP wash buffer 3 (100 mM Tris-HCl pH 9.0, 500 mM LiCl, 1% Igepal CA-630, 1% deoxycholic acid, 150 mM NaCl). Chromatin was eluted in elution buffer (50 mM NaHCO<sub>3</sub> 1% SDS) for 30 min at 65°C, and reverse-crosslinked in an overnight reaction at 62°C in 500 mM NaCl, 50 mM EDTA, 100 mM Tris-HCl pH 6.8 and 2 mg proteinase K (Roche). DNA was cleaned up and concentrated using a ChIP DNA Clean and Concentrator kit (Zymo Research). Quantitative PCR reactions were done using specific primer sets (Supplementary file 4).

### RNAi-mediated knockdown

MCF10A cells were plated at 175,000 cells per well of a six-well plate and transfected the next day using 30 pmol siRNAs targeting *E2F4* (SI02654694) and/or *E2F6* (SI00375445; Qiagen) using the RNAiMAX protocol (Invitrogen). Samples were harvested 24 hr post transfection.

### SILAC labeling and DNA pull-down experiments

For SILAC labeling, MCF7 cells were incubated in DMEM (-Arg, -Lys) medium containing 10% dialyzed FBS (PAN-Biotech) supplemented with 42 mg/L  $^{13}\text{C}_6^{15}\text{N}_4$  L-arginine and 73 mg/L  $^{13}\text{C}_6^{15}\text{N}_2$  L-lysine (Cambridge Isotope) or the corresponding non-labeled amino acids, respectively. SILAC incorporation was verified by in-gel trypsin digestion and MS analysis of 'heavy' input samples to ensure an incorporation rate of >98%. Cells were harvested and nuclear extracts were prepared as described (Kappei et al., 2017). Cells were harvested and incubated in hypotonic buffer (10 mM HEPES, pH 7.9, 1.5 mM  $\text{MgCl}_2$ , 10 mM KCl) on ice for 10 min. Cells were transferred to a Dounce homogenizer in hypotonic buffer supplemented with 0.1% Igepal CA630 (Sigma) and 0.5 mM DTT by 40 strokes. Nuclei were washed once in 1 PBS and extracted in hypertonic buffer (420 mM NaCl, 20 mM HEPES, pH 7.9, 20% glycerol, 2 mM  $\text{MgCl}_2$ , 0.2 mM EDTA, 0.1% Igepal CA630 (Sigma), 0.5 mM DTT) for 1 hr at 4°C on a rotating wheel. Samples were centrifuged at 4°C and >16,000 g for 1 hr and supernatants were used as nuclear protein extracts in the *in vitro* reconstitution DNA pull-down assays. DNA pull-downs were performed as described (Kappei et al., 2017). Briefly, 25 mg of forward and reverse sequence oligonucleotides (Supplementary file 4) were diluted in annealing buffer (20 mM Tris-HCl, pH 7.5, 10 mM  $\text{MgCl}_2$ , 100 mM KCl), denatured at 95°C and annealed by cooling. Annealed double-strand oligonucleotides were incubated with 100 units T4 kinase (ThermoFisher) for 2 hr at 37°C followed by incubation with 20 units T4 ligase overnight. Concatenated DNA strands were purified using phenol-chloroform extraction. Following biotinylation with desthiobiotin-dATP (Jena Bioscience) and 60 units DNA polymerase (ThermoFisher) the biotinylated probes were purified using MicroSpin G-50 columns (GE Healthcare). DNA baits were immobilized on 500 mg paramagnetic streptavidin beads (Dynabeads MyOne C1, ThermoFisher) on a rotation wheel for 30 min at room temperature. Subsequently, baits were incubated with 400 mg of nuclear extracts from MCF7 cells in PBB buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 5 mM  $\text{MgCl}_2$ , 0.5% Igepal CA-630 [Sigma]) while rotating for 2 hr at 4°C. 10 mg sheared salmon sperm DNA (Ambion) was added as a DNA binding competitor. After three PBS washes (450/500/600 mL), bound proteins were eluted in 2 Laemmli buffer and boiled for 5 min at 95°C.

### Mass spectrometry data acquisition and analysis

DNA pull-down samples were separated on a 12% NuPAGE Bis-Tris gel (ThermoFisher) for 30 min at 170 V in 1 MOPS buffer (ThermoFisher). The gel was fixed using the Colloidal Blue Staining Kit (ThermoFisher) and each lane was divided into four equal fractions. For in-gel digestion, samples were destained in destaining buffer (25 mM ammonium bicarbonate; 50% ethanol), reduced in 10 mM

DTT for 1 hr at 56°C followed by alkylation with 55 mM iodoacetamide (Sigma) for 45 min in the dark. Tryptic digest was performed in 50 mM ammonium bicarbonate buffer with 2 mg trypsin (Promega) at 37°C overnight. Peptides were desalted on StageTips and analyzed by nanoflow liquid chromatography on an EASY-nLC 1200 system coupled to a Q Exactive HF mass spectrometer (ThermoFisher). Peptides were separated on a C18 reversed-phase PicoFrit column (25 cm long, 75 mm inner diameter; New Objective) packed in-house with ReproSil-Pur C18-AQ 1.9 mm resin (Dr. Maisch). The column was mounted on an Easy Flex Nano Source and temperature controlled by a column oven (Sonation) at 40°C. A 105 min gradient from 2% to 40% acetonitrile in 0.5% formic acid at a flow of 225 nL/min was used. The spray voltage was set to 2.2 kV. The Q Exactive HF was operated with a TOP20 MS/MS spectra acquisition method per MS full scan. MS scans were conducted with 60,000 at a maximum injection time of 20 ms and MS/MS scans with 15,000 resolution at a maximum injection time of 50 ms. The raw files were processed with MaxQuant version 1.5.2.8 (Cox and Mann, 2008) with preset standard settings for SILAC labeled samples and the re-quantify option was activated. Carbamidomethylation was set as a fixed modification while methionine oxidation and protein N-acetylation were considered as variable modifications. Search results were filtered with a false discovery rate of 0.01. Known contaminants, proteins groups only identified by site, and reverse hits of the MaxQuant results were removed and only proteins were kept that were quantified by SILAC ratios in both 'forward' and 'reverse' samples. Raw mass spectrometry data will be accessible through the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD020473.

## Bioinformatics analyses

Primate genomes were accessed through Ensembl using the Compara application program interface via Bio::EnsEMBL::DBSQL::MethodLinkSpeciesSetAdaptor using the method\_link\_type 'EPO\_LOW\_COVERAGE' and the species\_set\_name 'primates'. The human *A3B* gene, including the -900 to +50 region, was compared to other sequences and further analyzed using Geneious Prime software (version 2019.1.3). *A3B* promoter transcription factor binding sites were predicted using the JASPAR database (Fornes et al., 2020) with score threshold set to 80%. TCGA primary breast tumors represented by both RNA-seq and whole-exome sequencing (Cancer Genome Atlas Network, 2012) were downloaded from the Firehose GDAC resource through the Broad Institute pipeline (<http://gdac.broadinstitute.org/>; n = 716). Genes correlated with *A3B* mRNA expression across the entire primary breast tumor data set were obtained using the USCS Xena Browser (bioRxiv 326470; doi: <https://doi.org/10.1101/326470>) and the top 20 genes in this list with the

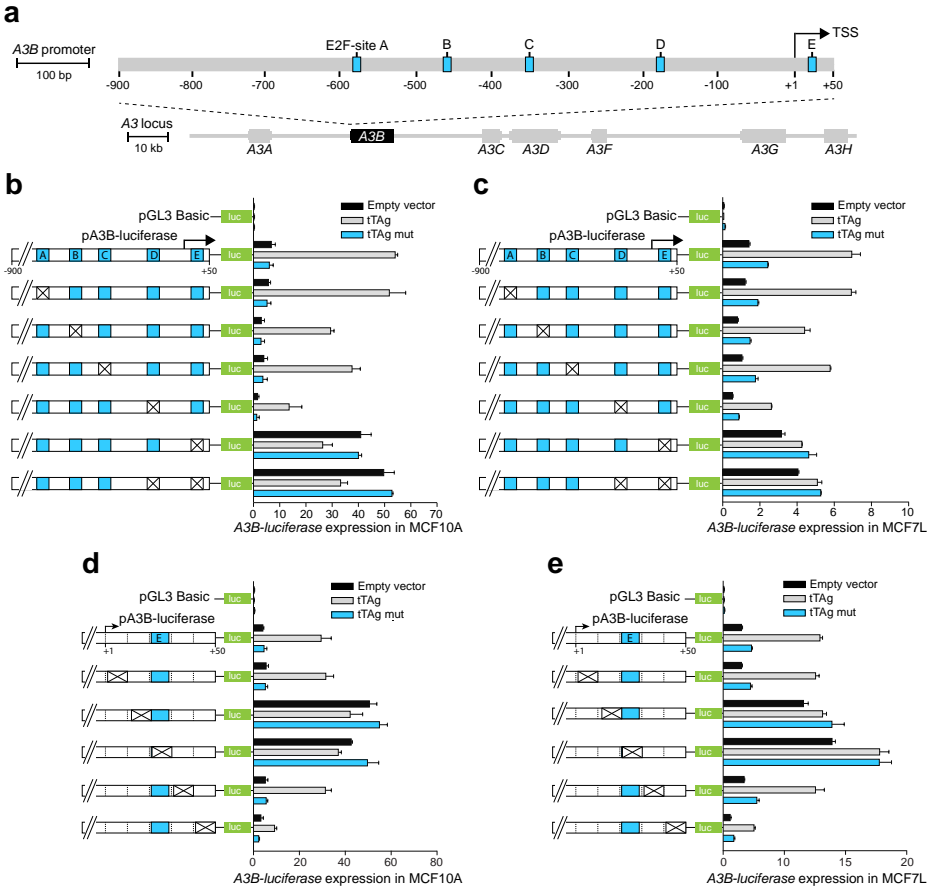
highest positive Spearman's correlations were used in this analysis. Quartiles for RSEM gene expression values relative to the housekeeping gene *TBP* were obtained for this gene list, and top (>75%) and bottom (<25%) quartiles of expression were calculated for every gene. Samples were then sorted into groups based on whether the expression of that gene fell into the top (n = 53 samples) or bottom (n = 111 samples) quartile for every gene in the list. RNA-seq data for *A3B* were also downloaded for every primary breast tumor, normalized to *TBP*, and used to establish expression correlations. The same methodology was used to calculate the top (>75%) and bottom (<25%) quartiles of expression for *A3B* mRNA to examine the mean *A3B* expression, APOBEC mutation signature, and APOBEC enrichment score. APOBEC mutation signatures were determined as described (Alexandrov et al., 2013; Jarvis et al., 2018) using the deconstructSigs R package (Rosenthal et al., 2016). APOBEC mutation enrichment scores were calculated using the hg19 reference genome and published methods (Chan et al., 2015). Enrichment score significance was assessed using a Fisher exact test with Benjamini-Hochberg false discovery rate (FDR) correction. All dataset analyses and visualizations were conducted using R and the ggplot2 package (<https://www.R-project.org/>).

## Results

### The *A3B* promoter contains a repressive transcriptional element

To study the mechanism of *A3B* transcriptional regulation, a 950 bp region spanning the *A3B* transcription start site (TSS) was cloned upstream of a firefly luciferase reporter (i.e. -900 to +50 relative to the +1 of the *A3B* TSS; Figure 1A). In MCF10A normal-like breast epithelial cells and MCF7 breast cancer cells, which both express low levels of *A3B* (Burns et al., 2013a), this construct supported modest levels of transcription activity above those of a promoter-less vector (compare black bars of pGL3-basic versus pA3B-luciferase in Figure 1B). Interestingly, similar to upregulation of the endogenous *A3B* gene in our previous studies (Starrett et al., 2019), transcription of the *A3B-luciferase* reporter was induced strongly in cells co-expressing the BK PyV truncated T antigen (tTag) but not in cells co-expressing a LxCxE mutant tTag (Figure 1B).

The JASPAR database (Fornes et al., 2020) was then used to predict transcription factor binding sites within the -900 to +50 *A3B* promoter region. This analysis yielded dozens of candidate sites including five putative E2F binding sites (labeled A-E in Figure 1A). The functionality of each E2F binding site was assessed by constructing site-directed mutant clusters and comparing *A3B-luciferase*



**Figure 1.** The A3B promoter harbors a repressive cis-element in the +1 to +50 region. (A) Schematic of the 7-gene human *APOBEC3* locus with the A3B promoter magnified to depict five predicted E2F binding sites (A-E in blue) relative to the TSS at +1 (scales indicated). (B-E) Relative luciferase activity of MCF10A or MCF7 cells expressing the indicated firefly luciferase construct (pGL3-basic, pA3B-luciferase, or mutant pA3B-luciferase), a renilla luciferase internal control plasmid (not shown), and a tTag plasmid (empty, wildtype, or LxCxE mutant). Mutation clusters are depicted by X's (mutant sequences in Supplementary file 4). Experiments report mean  $\pm$  SD of  $n \geq 2$  technical replicates and are representative of  $n = 3$  biologically independent replicates.

reporter activity in MCF10A and MCF7 (Figure 1B–C). Clustered base substitution mutations in sites A, B, and C had negligible effects on basal or tTag-induced levels of luciferase reporter expression. Clustered mutations in site D caused a two- to three-fold reduction in both basal and tTag-induced levels of luciferase reporter expression. However, clustered mutations in site E, located at +21 to +28 relative to the TSS, caused a strong five-fold induction of A3B-luciferase reporter activity that could not be further increased by tTag co-expression. Mutations in site E were also



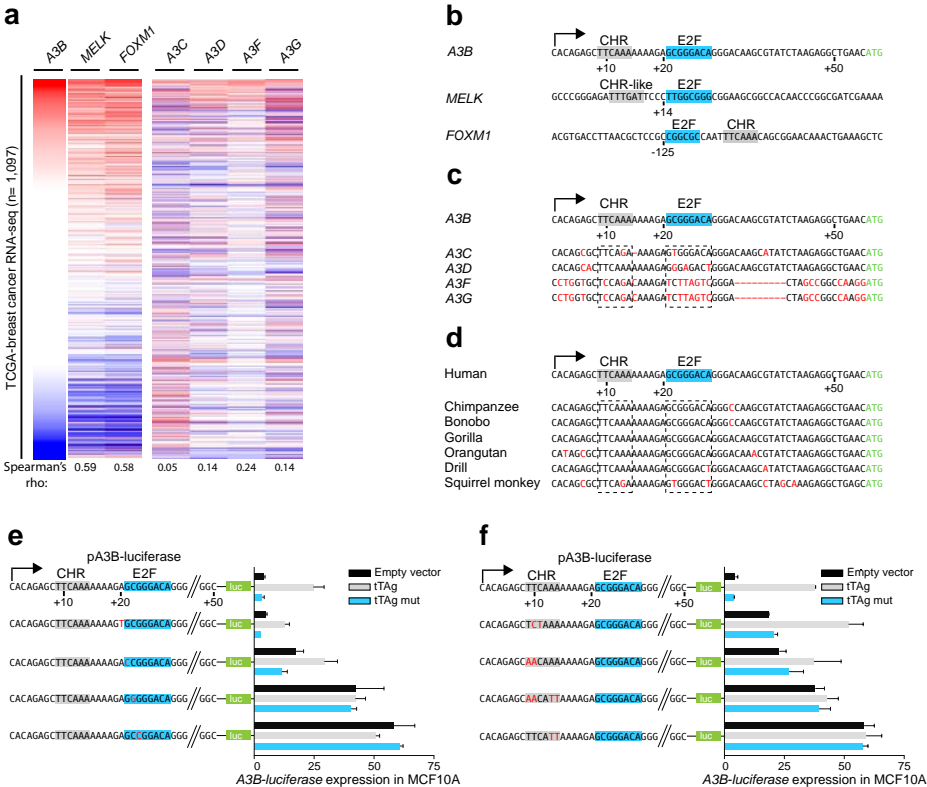
epistatic to those in site D, suggesting that site E may be the dominant regulatory site. The importance of site E was confirmed by analyzing additional mutation clusters, which partly or fully spanned site E and resulted in complete derepression of *A3B-luciferase* expression (Figure 1D–E). Mutation clusters +12 to +20 and +22 to +30 guided additional analyses including proteomics experiments below. Taken together, these results suggested that the +12 to +30 region of the *A3B* promoter including site E is normally bound by a repressive factor and different mutations prevent repression and allow high levels of transcription.

### ***A3B* promoter phylogenetic analyses delineate conserved CHR and E2F sites**

To gain additional insights into the possible involvement of an E2F complex in *A3B* transcriptional repression, TCGA breast cancer RNA-seq data sets were used to identify 114 genes with expression profiles positively associating with *A3B* (Spearman's  $\rho \geq 0.5$ ;  $n = 1,097$  RNA-seq data sets; Supplementary file 1, see online version of this publication). Remarkably, 87% of these genes were shown to be bound by repressive E2F complexes suggesting a common regulatory mechanism (Litovchick et al., 2007; Müller et al., 2014; Supplementary file 1, see online version of this publication). For instance, *A3B* mRNA levels across primary breast cancer associated strongly with expression levels of MELK and FOXM1 (Figure 2A), which both have well-described E2F-dependent repression mechanisms (Litovchick et al., 2007; Müller et al., 2017; Müller et al., 2014; Verlinden et al., 2005). A subset of these coordinately expressed genes also has a predicted consensus (or near-consensus) cell cycle gene homology region (CHR) element adjacent to the predicted E2F binding site (Figure 2B and Supplementary file 1, see online version of this publication). When juxtaposed, these two elements cooperatively facilitate the binding of repressive E2F complexes and suppress gene expression (Müller et al., 2012; Müller et al., 2017; Müller et al., 2014) and reviewed by Fischer and Müller, 2017; Sadasivam and DeCaprio, 2013. Interestingly, in the *A3B* promoter, both the predicted CHR (+9 to +14) and E2F (+21 to +28) elements occur within the +12 to +30 region defined above in mutagenesis experiments (Figure 1D–E).

The global profile of *A3B* mRNA expression in primary breast cancer is distinct from related *A3* genes except for *A3A* (Figure 2A). This is explained by differences at potentially critical nucleobase positions in both the CHR and E2F sites in the individual *A3* gene promoters including the most closely related *A3C* promoter region (Figure 2C and see below). The *A3A* promoter shares no obvious homology and the associated expression profiles cannot be explained mechanistically at this time. Sequence comparisons with other primates demonstrate that this region of

the *A3B* promoter, including juxtaposed CHR and E2F elements, is conserved in hominids and Old World monkeys (Figure 2D). Thus, adjacent CHR and E2F sites in the *A3B* promoter are unique amongst *A3* genes, specific to humans and other higher primates, and likely linked to the aforementioned expression patterns.



**Figure 2.** *A3B* repression requires both CHR and E2F cis-elements. (A) Heatmap depicting high-to-low *A3B* expression levels in TCGA breast cancer specimens (n = 1,097) and correlations with two known RB/E2F target genes, *MELK* and *FOXM1*, and related *APOBEC3* genes (Spearman's rho indicated). (B) Comparison of the *A3B* promoter and analogous regions of *MELK* and *FOXM1*. Known and predicted E2F and CHR elements are indicated in blue and light gray, respectively. (C–D) Alignments of the *A3B* promoter sequence and corresponding promoter sequences of related human *APOBEC3* genes and representative non-human primate *A3B* genes. (E–F) Relative luciferase activity of MCF10A cells expressing the indicated firefly luciferase construct (pA3B-luciferase or mutant pA3B-luciferase), a renilla luciferase internal control plasmid (not shown), and a tTag plasmid (empty, wildtype, or LxCxEmutant). Panel (E) reports data for E2F mutants and (F) for CHR mutants. Experiments report mean  $\pm$  SD of n  $\geq$  2 technical replicates and are representative of n = 3 biologically independent replicates.

To interrogate the functionality of the E2F and CHR elements, the *A3B-luciferase* reporter was subjected to additional rounds of site-directed mutagenesis and

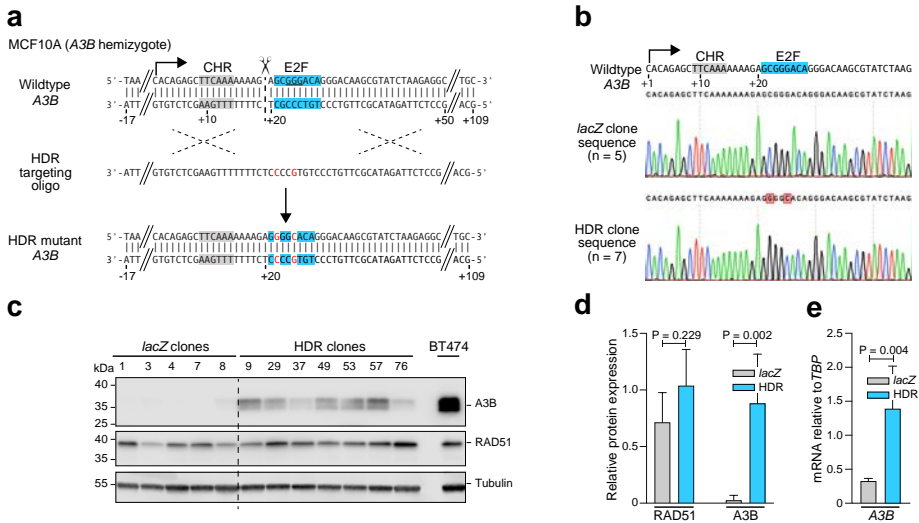
analysis in MCF10A. Altering the nucleobase immediately 5' of the predicted E2F binding site (+20 A-to-T) had no effect, and changing the first nucleobase of the predicted E2F binding site (+21 G-to-C) caused slight reporter activation but did not affect tTag inducibility (Figure 2E). In contrast, single nucleobase changes in the core of the predicted E2F binding site (+22 C-to-G or +23 G-to-C) caused full de-repression of the *A3B-luciferase* reporter that could not be further increased by tTag (Figure 2E). Single and combinatorial base substitution mutations in the CHR element also resulted in partial or full de-repression of the *A3B-luciferase* reporter (Figure 2F). For instance, mutation of +10 TC-to-CT or +9 TT-to-AA caused partial reporter de-repression, which could still be further enhanced by tTag. In contrast, mutating the two adenine nucleobases at the 3' end of the CHR element (+13 AA-to-TT) resulted in full reporter de-repression, which could not be increased by tTag. These fine-mapping results showed that both the putative E2F binding site and the adjacent CHR element are essential for repressing *A3B* transcription.

### **Targeted mutagenesis demonstrates a repressive role for the +21 to +28 E2F element in regulating endogenous *A3B* transcription independent of activation by the PKC/ncNF- $\kappa$ B pathway**

The abovementioned work indicated recruitment of a repressive complex to a putative E2F binding site in the *A3B-luciferase* reporter, which was necessarily episomal and may not be subject to the same regulatory mechanisms as the chromosomal *A3B* gene. To directly ask whether the endogenous +21 to +28 E2F site is involved in *A3B* repression, CRISPR/Cas9 technology was used to disrupt this region in diploid MCF10A cells. Four independent targeted clones showed elevated *A3B* protein levels in comparison to control *lacZ* clones, consistent with a repressive function for the putative E2F binding site (Figure 3—figure supplement 1A). DNA sequencing revealed allelic differences between the four clones, which could explain at least part of the variability observed in *A3B* elevation (Figure 3—figure supplement 1B).

To confirm and extend these results, homology-directed repair (HDR) was used to introduce precise base substitution mutations into the +21 to +28 E2F site in the endogenous *A3B* promoter of an MCF10A derivative engineered to be hemizygous for the entire *A3B* gene (Materials and methods). Tandem base substitution mutations, C22G and G25C, were chosen to disrupt the E2F site and simultaneously preserve the locus by maintaining the overall G:C content and spatial relationships between promoter elements (Figure 3A). Seven independent clones were obtained with the desired two base substitution mutations (Figure 3B). All seven showed robust increases in both *A3B* protein and mRNA levels with differences potentially due to clonal variation (Figure 3C–E). The mRNA levels of related *A3* family members

were unaffected, which further confirmed specificity of the targeted genomic changes (Figure 3—figure supplement 1C). Immunoblots were also performed for RAD51, an established RB/E2F-target (Dean et al., 2012; Müller et al., 2017), to show that global E2F regulation is unperturbed (Figure 3C–D). These results demonstrated that the endogenous E2F site at base pairs +21 to +28 of the *A3B* promoter contributes to transcriptional repression in MCF10A cells.



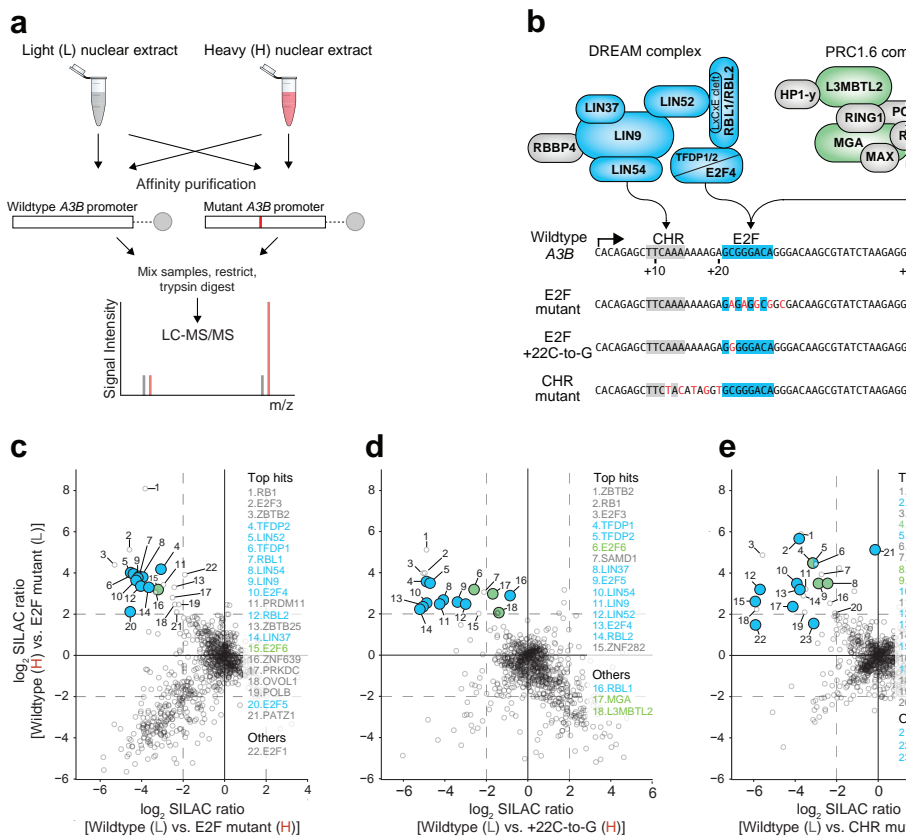
**Figure 3.** Single-base substitutions in the endogenous predicted E2F binding site induce *A3B* expression independent of activation by the PKC/ncNF- $\kappa$ B pathway. Complementary supporting data are in Figure 3—figure supplement 1. (A) Schematic of CRISPR/Cas9-mediated HDR of the predicted E2F binding site in *A3B* hemizygous MCF10A cells. Top: CRISPR/Cas9 (scissors) introduces a DNA break (dashed line) adjacent to the predicted E2F binding site (blue). Middle: The ssDNA oligo used for HDR has two point mutations in the predicted E2F binding site including one that disrupts the PAM (underlined). Bottom: *A3B* promoter sequence of properly targeted clones. (B) Sanger DNA sequencing chromatograms of the E2F promoter region of a representative control clone (*lacZ* clones, n = 5) and a representative clone with the targeted E2F point mutations (HDR clones, n = 7). (C–D) *A3B* and RAD51 protein levels in control *lacZ* and HDR clones with tubulin as a loading control (representative immunoblots and quantification from n = 3 experiments). *A3B*-overexpressing BT-474 cells were used as a positive control. P-values from unpaired t-test. (E) *A3B* mRNA expression levels in control *lacZ* and HDR clones quantified by RT-qPCR (mean  $\pm$  SD; p-value from unpaired t-test). (F) RT-qPCR (top) and immunoblot (bottom) results showing the effects of wildtype and LxCxE mutant tTag on the *A3B* gene (top) and protein (bottom) expression in two representative *lacZ* and HDR49 clones. Cyclin E2 was used as a positive immunoblot control for tTag-mediated induction of an RB/E2F-repressed gene. Tubulin was used as an immunoblot loading control. (G) Expression of *A3B* mRNA (top) and protein (bottom) upon PMA-treatment of the indicated *lacZ* control and HDR mutant clones. The magnitude of mRNA induction is indicated for each DMSO control and PMA-treated pair. Tubulin was used as an immunoblot loading control.

To determine whether this *cis*-element is solely responsible for endogenous *A3B* upregulation by tTag or whether multiple tTag-responsive mechanisms may combine to exert the observed phenotype, tTag was expressed in two representative HDR targeted MCF10A clones and two *lacZ* controls and *A3B* levels were analyzed by RT-qPCR and immunoblotting. Expression of tTag resulted in two-to three-fold higher *A3B* levels in *lacZ* control clones (Figure 3F), similar to results above with the episomal reporter. In contrast, neither expression of an LxCxE mutant nor an empty mCherry control vector induced *A3B*. Importantly, tTag had no effect on *A3B* mRNA or protein levels in the HDR targeted MCF10A clones (Figure 3F). This result was clear despite the fact that the LxCxE mutant was expressed more strongly than wildtype tTag (likely due to loss of an autoregulatory mechanism yet-to-be-defined) and that some variability in endogenous *A3B* expression was observed from experiment-to-experiment (even using the same HDR-targeted clone). Nevertheless, these results combined to demonstrate that all of the observed *A3B* induction by tTag is mediated by this single endogenous E2F site.

In parallel, representative HDR-targeted clones and *lacZ* controls were used to ask how this endogenous E2F site might impact *A3B* induction by PMA through the PKC/ $\kappa$ B signal transduction pathway (Leonard et al., 2015). This was done by treating cells with PMA and then quantifying *A3B* levels by RT-qPCR and immunoblotting. Interestingly, PMA caused similar induction of *A3B* mRNA and protein levels from both the wildtype endogenous promoter (*lacZ* controls) as well as the HDR-engineered promoter with tandem base substitution mutations C22G and G25C (Figure 3G). Overall, simultaneous de-repression through HDR-targeted mutation of the single E2F site and activation by PMA caused a thirty-fold increase in *A3B* levels above the uninduced basal level in the *lacZ* controls. Together with the above data, these results demonstrated that *A3B* expression is impacted independently by tTag/E2F and PKC/ $\kappa$ B and signal transduction mechanisms.

### **Repressive E2F4/DREAM and E2F6/PRC1.6 complexes bind to the *A3B* promoter**

Collectively, the data so far indicate that the putative E2F binding site is functionally relevant in repressing both *A3B-luciferase* reporter activity and endogenous *A3B* expression. However, the identity of the repressive complex(es) bound to this *cis*-element was unclear because multiple E2F family members are capable of transcriptional repression (Introduction). To address this problem in an unbiased manner, a series of proteomic experiments was conducted to identify MCF7 nuclear proteins capable of binding to the wildtype *A3B* +1 to +50 promoter sequence but not to repression-defective mutants (see Figure 4A for



**Figure 4.** The DREAM and PRC1.6 repressive complexes bind to the CHR-E2F region of the *A3B* promoter. Immunoblot validations of representative binding proteins from proteomic experiments are in Figure 4—figure supplement 1. (A) Schematic of the SILAC DNA pull-down strategy used to identify proteins from MCF7 cells capable of interacting with *A3B* promoter sequences. (B) Illustration of DREAM and PRC1.6 complexes positioned over the indicated *A3B* promoter elements (proteomic hits shaded blue and green, respectively). (C–E) Log<sub>2</sub>-transformed SILAC ratios of proteins purified using the indicated promoter sequences and identified through LC-MS/MS (dashed line, SILAC ratio threshold >2.0 [log<sub>2</sub>]). ‘Top hits’ are proteins surpassing the >2.0 log<sub>2</sub> SILAC ratio threshold in both datasets (rank based on heavy versus light SILAC ratio). ‘Others’ are proteins of interest surpassing the >2.0 log<sub>2</sub> SILAC ratio threshold in at least one dataset. Data for DREAM and PRC1.6 components are shaded blue and green, respectively.

a schematic of the proteomics workflow and Materials and methods for details). This approach was facilitated by stable isotope labeling with amino acids in cell culture (SILAC) to create heavy (H) and light (L) nuclear extracts for H versus L and L versus H comparisons with the different promoter substrates. Interestingly, in an experiment comparing proteins bound to the wildtype *A3B* promoter sequence versus a promoter sequence with mutations spanning the predicted E2F binding

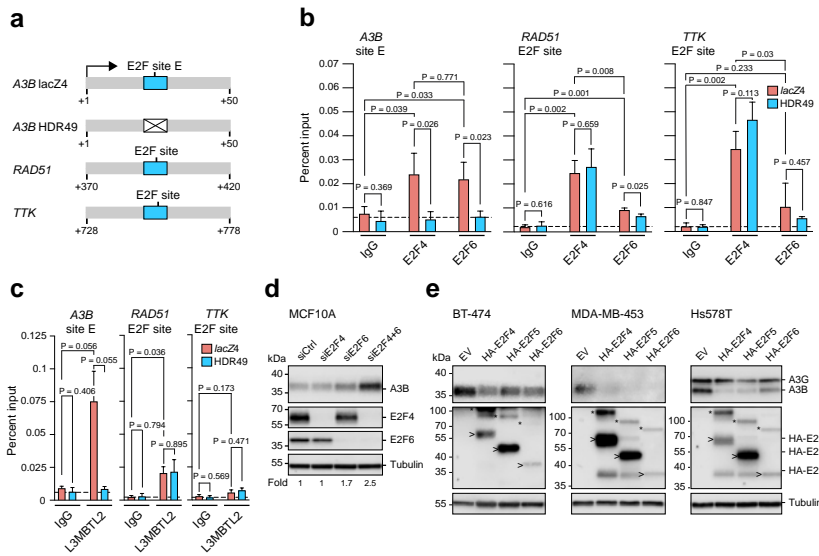
site (matching the +22-to-30 mutant in Figure 1D–E), a greater than four-fold enrichment was observed for almost all proteins in the repressive DREAM complex, including TFDP1, TFDP2, RBL1, RBL2, E2F4, E2F5, and the MuvB components LIN9, LIN37, LIN52, and LIN54 (Figure 4B–C and Supplementary file 2, see online version of this publication; confirmatory immunoblots for representative enriched proteins in Figure 4—figure supplement 1). Given that a single-base substitution +22 C-to-G was sufficient for full de-repression in reporter assays (Figure 2E), we repeated the SILAC DNA pull-downs comparing the wildtype promoter sequence and this mutant. Importantly, again, most members of the DREAM complex preferentially bound to the wildtype but not to the *A3B* mutant promoter sequence (Figure 4B,D, Supplementary file 2, see online version of this publication). Similar enrichments for DREAM complex components were also evident in a separate proteomics experiment comparing MCF7 nuclear proteins bound to the wildtype *A3B* promoter versus a promoter sequence with mutations spanning the CHR element (matching the +12-to-20 mutant in Figure 1D–E; Figure 4B, E, Supplementary file 2, see online version of this publication). These additional results indicated that the CHR site is also required for *A3B* promoter binding by the DREAM complex and that the E2F site alone is insufficient.

Interestingly, the proteomics data sets also implicated components of the PRC1.6 complex in binding to wildtype but not to E2F or CHR mutant *A3B* promoter sequences. In particular, E2F6, MGA, and L3MTBL2 were found enriched repeatedly (Figure 4B–E, Supplementary file 2, see online version of this publication, and confirmatory immunoblots for representative enriched proteins in Figure 4—figure supplement 1). Two additional PRC1.6 components, PCGF6 and WDR5, also approached the four-fold cut-off in one dataset (Supplementary file 2, see online version of this publication). These results indicated that the repressive PRC1.6 complex is also capable of binding to the wildtype *A3B* promoter sequence and may therefore also play a role in suppressing expression.

### **E2F4 and E2F6 complexes participate in *A3B* transcriptional repression**

A series of chromatin immunoprecipitation (ChIP) experiments was done to determine whether *A3B* repression in non-tumorigenic MCF10A cells is mediated by one or both of the identified E2F complexes. Although prior work has implicated the E2F4/DREAM complex (Periyasamy et al., 2017), the potential involvement of E2F6/PRC1.6 is novel. Anti-E2F4 and anti-E2F6 antibodies were used to immunoprecipitate cross-linked transcriptional regulatory complexes from MCF10A

*lacZ4* (control) and HDR49 (E2F site E mutant) cells described above and promoter occupancy was determined by quantitative PCR (Figure 5A–B). The wildtype *A3B* promoter in *lacZ4* cells showed similarly strong enrichment for binding by both E2F4 and E2F6, and single-base substitutions in E2F site E in HDR49 cells reduced binding of both proteins to background levels. In parallel analyses, significant E2F4 enrichment was evident in the promoter regions of two established E2F4/DREAM-repressed genes, *RAD51* and *TTK* (Dean et al., 2012; Engeland, 2018; Müller et al., 2017). E2F6 was enriched only at the *RAD51* promoter and not the *TTK* promoter.



**Figure 5.** Endogenous *A3B* regulation by both E2F4/DREAM and E2F6/PRC1.6 complexes. (A) Schematics of the promoter regions interrogated by ChIP experiments. Wildtype E2F sites are depicted by blue boxes and the mutant E2F site in the *A3B* promoter by a white X-box. (B–C) E2F4, E2F6, and L3MBTL2 occupancy at the indicated E2F sites in *A3B*, *RAD51*, and *TTK*, as analyzed by ChIP-qPCR using *lacZ4* and HDR49 cells. Experiments in (B) report mean  $\pm$  SD of  $n = 3$  biologically independent replicates and in (C) of  $n = 2$  biologically independent replicates (p values from unpaired t-test). Dashed lines indicate the average IgG background. (D) Immunoblots of A3B, E2F4, and E2F6 in MCF10A cells treated 24 hr with the indicated siRNAs. Tubulin was used as a loading control. Representative blots are shown and fold-changes below are based on the average values from  $n = 3$  biologically independent replicates. (E) Immunoblots of A3B and the indicated HA-tagged E2F proteins in BT-474, MDA-MB-453, and Hs578T cells transduced with lentiviral constructs encoding mCherry-T2A-HA-E2F4, -HA-E2F5, -HA-E2F6, or -EV (empty vector control). Representative blots of  $n = 2$  experiments are shown. Asterisks indicate E2Fs still fused with mCherry due to incomplete ribosome skipping at the T2A site, and arrowheads indicate bands for free E2F. Tubulin was used as a loading control.

The higher E2F6 signal at the *A3B* promoter compared to the *RAD51* and *TTK* promoters prompted us to ask whether other PRC1.6 components may also



bind preferentially. ChIP experiments for L3MBTL2 revealed strong binding of this PRC1.6 component to the *A3B* promoter, intermediate levels to the *RAD51* promoter, and insignificant levels to the *TTK* promoter (Figure 5C). These ChIP experiments indicated that the *A3B* promoter can be occupied by both E2F4 and E2F6 complexes, that the binding of either complex requires an intact +21 to +28 E2F site, and that the binding of the same proteins to other established E2F sites can vary significantly within the same cell population.

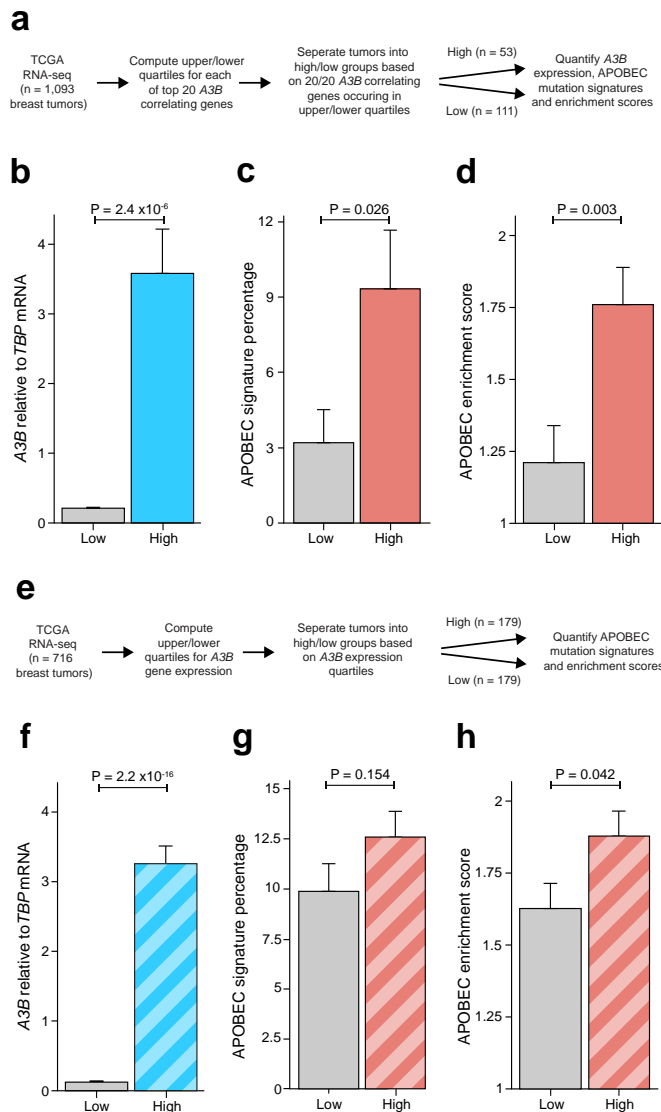
Next, we used small interfering (si)RNAs to interrogate the repressive function of each complex in MCF10A cells. Surprisingly, E2F4 depletion alone did not alter *A3B* expression, whereas E2F6 depletion caused an increase in *A3B* protein levels by immunoblotting (Figure 5D). We also observed that combined E2F4/E2F6 depletion increases *A3B* protein levels more than E2F6 alone, indicating that both complexes contribute to repression with the latter potentially being more dominant. Analogous knockdown attempts in MCF7 cells caused overt distress and inviability (data not shown). Conversely, overexpression of either E2F4 or E2F6, as well as E2F5, which also forms a DREAM complex (Litovchick et al., 2007), was able to repress *A3B* expression to varying extents in multiple different breast cancer cell lines (Figure 5E). Taken together, the ChIP, knockdown, and overexpression studies indicate that both E2F4/DREAM and E2F6/PRC1.6 complexes can occupy the *A3B* promoter and repress transcription. Moreover, the significant *A3B* upregulation observed upon E2F6 knockdown but not E2F4 knockdown suggests that the PRC1.6 complex repression mechanism may predominate.

### **Breast tumors with overexpression of an E2F-repressed gene set elicit higher levels of APOBEC signature mutations**

Breast tumors frequently display *A3B* overexpression and APOBEC signature mutations (Alexandrov et al., 2013; Angus et al., 2019; Bertucci et al., 2019; Burns et al., 2013a; Burns et al., 2013b; Nik-Zainal et al., 2012; Nik-Zainal et al., 2016; Roberts et al., 2013). However, association studies with large breast cancer cohorts have shown only weak positive or negligible associations between *A3B* expression levels and APOBEC signature mutation loads, and clear outliers exist including tumors with high *A3B* and few APOBEC signature mutations and low *A3B* and many APOBEC signature mutations (Buisson et al., 2019; Burns et al., 2013a; Burns et al., 2013b; Nik-Zainal et al., 2014; Roberts et al., 2013). This variability may be due to a number of factors including different durations of mutagenesis (i.e. tumor age is unknown and distinct from a patient's biological age) and mutagenic contributions from other APOBEC3 enzymes governed by distinct regulatory mechanisms (Buisson et al., 2019; Cortez et al., 2019; Nik-Zainal et al., 2014; Starrett et al., 2016). However,

given our results implicating both E2F4 and E2F6 complexes in *A3B* repression, we reasoned that effects from these and other potentially confounding variables may be overcome by asking whether the APOBEC mutation signature is enriched in breast tumors with functional overexpression of an E2F-repressed 20 gene set.

This was done by analyzing TCGA breast cancer RNA seq and whole-exome sequencing data (Cancer Genome Atlas Network, 2012) for gene expression levels and base substitution mutation signatures (workflow in Figure 6A). The top 20 genes associating positively with *A3B* and also showing evidence for E2F repression (Litovchick et al., 2007; Müller et al., 2014) were used to rank tumors based on highest to lowest expression levels of each gene (Figure 2A–B and Supplementary files 1 and 3). Tumors ranking in the top or bottom quartiles for expression of all 20 genes were considered for additional analyses ( $n = 53$  and  $n = 111$  tumors in the common high and common low groups, respectively). Once common high and low groups were delineated, pairwise comparisons were made for *A3B* expression levels, percentage of APOBEC signature mutations, and APOBEC signature enrichment values. As expected from the analysis work-flow and the likelihood of a shared transcriptional regulation mechanism, tumors with common high-expressing genes showed an average of twenty-fold higher *A3B* mRNA levels than tumors with common low-expressing genes ( $p < 2.4 \times 10^{-6}$  by Welch's test; Figure 6B). More interestingly, tumors with the common high-expressing genes showed an average of 9.3% APOBEC signature mutations versus 3.2% in the common low group ( $p = 0.026$  by Welch's test; Figure 6C). As an independent metric, significantly higher APOBEC mutation signature enrichment values were evident in tumors defined by the common set of high-expressed genes in comparison to tumors with the same genes expressed at low levels ( $p = 0.003$  by Welch's test; Figure 6D). A pairwise analysis of the mean expression value of the top 20 *A3B*-associating/E2F-repressed genes yielded similar positive associations with *A3B* mRNA expression levels, in tumors defined by the common set of APOBEC mutation percentages, and enrichment scores (Figure 6—figure supplement 1). In tumors defined by the common set of high-expressed genes in comparison to tumors with the same genes expressed at low levels ( $p = 0.003$  by Welch's test; Figure 6D). A pairwise analysis of the mean expression value of the top 20 *A3B*-associating/E2F-repressed genes yielded similar positive associations with *A3B* mRNA expression levels, APOBEC mutation percentages, and enrichment scores (Figure 6—figure supplement 1).



**Figure 6.** Elevated levels of APOBEC signature mutations in breast tumors with coordinated overexpression of an E2F-repressed gene set. Complementary analyses are presented in Figure 6—figure supplement 1. (A) Schematic depicting the bioinformatics workflow of TCGA breast tumor data sets based on the 20 genes most strongly associated with *A3B* expression (Figure 2A and Supplementary file 1, see online version of this publication). (B–D) The mean *A3B* mRNA levels, mean APOBEC mutation percentages, and mean APOBEC enrichment scores in breast tumors with coordinated overexpression (high) or repression (low) of the 20 gene set (mean ± SD; n = 53 tumors in the high group and n = 111 in the low group; p values from Welch’s t-test). (E) Schematic depicting the bioinformatics workflow of TCGA breast tumor data sets based solely on *A3B* mRNA expression levels. (F–H) The mean *A3B* mRNA levels, mean APOBEC mutation percentages, and mean APOBEC enrichment scores in breast tumors with high or low *A3B* mRNA levels (mean ± SD of top and bottom quartiles; n = 179 tumors in each group; p values from Welch’s t-test).

Although associations between *A3B* mRNA levels and APOBEC mutation signature have been analyzed previously (references above), we wanted to apply an *A3B*-focused quartile-binning approach to be able to compare results with those from the 20-gene set above (work-flow in Figure 6E). Therefore, TCGA breast tumor RNA-seq data were used to identify the top 25% and bottom 25% of *A3B* expressing tumors ( $n = 179$  per group). As mentioned above and expected from the work-flow, average *A3B* mRNA levels were much higher in the *A3B*-high group in comparison to the *A3B*-low group (Figure 6F). Also similar to the analysis above, both the average APOBEC mutation signature percentages and average APOBEC enrichment scores trended upward in *A3B*-high tumors (Figure 6G–H). However, in contrast to the analysis above, the difference in APOBEC mutation signature percentages was not significant and the difference in APOBEC enrichment scores was barely significant ( $p=0.154$  and  $0.042$  by Welch's test, respectively). Altogether, these results indicate that coordinated overexpression of an RB/E2F-repressed gene set may be a better indicator for APOBEC mutation susceptibility than expression of *A3B* itself. Potential explanations for the different results from each analysis approach are discussed below.

## Discussion

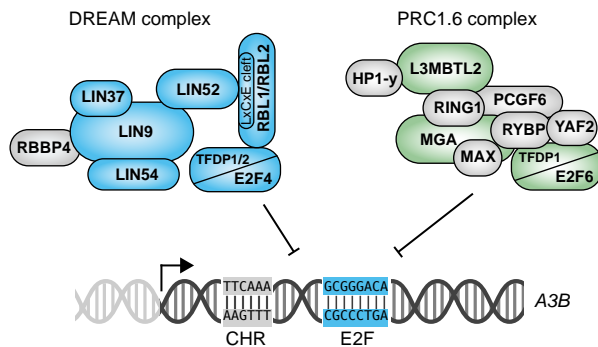
The studies here are the first to demonstrate that two repressive E2F complexes, E2F4/DREAM and E2F6/PRC1.6, combine to suppress *A3B* transcription and thereby protect genomic integrity in normal cells. The construction of a novel *A3B-luciferase* reporter enabled the delineation of a repressive cis-element comprised of juxtaposed E2F and CHR sites. Site-directed mutation of either site caused full de-repression that could not be further enhanced by co-expression of BK-PyV tAg. These results indicated that TAg-mediated upregulation of *A3B* reported previously (Starrett et al., 2019; Verhalen et al., 2016) is occurring exclusively through the RB/E2F axis and not through an alternative LxCxE-dependent mechanism. The importance of this E2F binding site in the endogenous *A3B* promoter was demonstrated definitively by CRISPR/Cas9-mediated base substitution mutation and experimentation with a panel of independent knock-in clones. Proteomics experiments revealed that two distinct repressive regulatory complexes, specifically E2F4/DREAM and E2F6/PRC1.6, are capable of binding to the wildtype *A3B* promoter but not to E2F or CHR mutant derivatives. Repressive roles for both E2F complexes were demonstrated by ChIP, knockdown, and overexpression studies. Finally, the potential pathological significance of E2F-mediated de-repression of *A3B* in breast cancer was supported by TCGA data analyses showing significant positive associations between elevated

expression of a set of 20 coordinately expressed E2F-regulated genes and higher levels of APOBEC signature mutations.

There is a broad interest in understanding the molecular mechanisms that govern *A3B* transcriptional regulation due to its physiological functions in antiviral immunity and pathological roles in cancer mutagenesis. Although prior studies implicated the E2F4/DREAM complex and generally the RB/E2F axis in repressing *A3B* transcription (Periyasamy et al., 2017; Starrett et al., 2019), the work here is the first to define the responsible cis-elements (juxtaposed CHR and E2F sites), show that all PyV tAg-mediated activation occurs through this single bipartite sequence, and demonstrate coordinated repression not only by the E2F4/DREAM complex but, surprisingly, also by the E2F6/PRC1.6 complex. Moreover, *A3B* induction by E2F4/6 de-repression occurs independently of *A3B* activation by PKC/ $\kappa$ B signal transduction. This additional result suggests that upregulation of *A3B* expression through genetic or viral perturbation of the RB/E2F cell cycle pathway has the potential to combine synergistically with inflammatory responses and trigger even greater levels of genomic DNA damage and mutagenesis. The role of p53 in *A3B* transcriptional regulation is less clear with some studies indicating that p53 inactivation leads to *A3B* upregulation (Menendez et al., 2017; Periyasamy et al., 2017) and others demonstrating that *TP53* knockout has no effect on *A3B* transcription (Nikkilä et al., 2017; Starrett et al., 2019). This may be due to differences in cell types and growth conditions. Alternatively, rather than playing an upstream role in *A3B* transcriptional regulation, p53 may function to help activate a downstream DNA damage response to prevent the accumulation of mutations by *A3B*, which also explains why genetic inactivation of *TP53* associates positively with elevated *A3B* mRNA levels (Burns et al., 2013a).

Our results support a model in which E2F4/DREAM and E2F6/PRC1.6 complexes combine to repress *A3B* transcription (Figure 7). These two complexes are likely to compete for binding to the same conserved E2F site located at +21 to +28 of the *A3B* promoter because tandem base substitution mutations (C22G and G25C) de-repress expression of endogenous *A3B* and render the locus non-responsive to further activation by tAg (Figure 3F). Similar results were obtained using E2F site E mutants of the episomal *A3B-luciferase* reporter (Figure 2E). Base substitution mutations in the adjacent CHR site in the episomal *A3B-luciferase* reporter also caused *A3B* de-repression to levels that could not be further increased by tAg (Figure 2E). These genetic results were corroborated by proteomics data sets indicating that base substitution mutations in either the E2F site or the CHR site fully abrogate promoter sequence binding by both the DREAM and PRC1.6

complexes (Figure 4). However, unlike E2F4, E2F6 is not known to be regulated through a TAg/LxCxE-dependent mechanism nor has its function been shown to require a CHR site. Future work will be required to bridge this knowledge gap. For instance, it may be possible that a subset of E2F6/PRC1.6 complex leverages an as-yet-unknown CHR binding factor to repress genes such as *A3B*. Alternatively, it may be possible that LxCxE-dependent interactions with PRC1.6 components other than E2F6 might interfere with the repressive function of PRC1.6. It is unlikely, however, that the E2F6/PRC1.6 complex requires the E2F4/DREAM complex as a cofactor for binding because E2F4-depleted cells maintain near-complete repression of *A3B* expression (Figure 5D).



**Figure 7.** Model for coordinated repression of *A3B* transcription by both E2F4/DREAM and E2F6/PRC1.6 complexes. Transcriptional repression of *A3B* through the combined activities of E2F4/DREAM and E2F6/PRC1.6 complexes. Other regulatory mechanisms including *A3B* transcriptional activation by PKC/ncNF-κB signal transduction are not shown. See text for details and discussion.

The E2F-governed regulatory mechanism described here provides an attractive explanation for a large proportion of reported *A3B* overexpression in both viral and non-viral cancer types. For instance, the HPV E7 and PyV TAg oncoproteins may trigger *A3B* upregulation directly by dissociating repressive E2F complexes. Accordingly, cervical cancers are almost invariably HPV-positive, *A3B*-overexpressing, and enriched for APOBEC signature mutations (Burns et al., 2013b; Cancer Genome Atlas Research Network, 2017; Roberts et al., 2013; Zapatka et al., 2020). HPV-positive head/neck cancers also show *A3B*-overexpression and APOBEC mutation signature enrichment (Burns et al., 2013b; Cancer Genome Atlas Network, 2015; Cannataro et al., 2019; Faden et al., 2017; Roberts et al., 2013; Vieira et al., 2014; Zapatka et al., 2020). Importantly, many HPV-negative cancers elicit similarly high *A3B* expression levels and APOBEC mutation burdens (Burns et al., 2013b; Cancer Genome Atlas Network, 2015; Cannataro et al., 2019; Gillison et al., 2019). Moreover, HPV status in head/neck cancer appears mutually exclusive with alterations of RB/

E2F axis genes, such that HPV-negative cancers often display copy number loss of *CDKN2A* (encoding p16) and overexpression of *CCND1* (encoding Cyclin D1; Cancer Genome Atlas Network, 2015; Gillison et al., 2019; Zapatka et al., 2020), which effectively mimics a subset of the oncogenic effects of E7. This indicates that both virus-dependent and independent tumors may exploit the same pathway to derepress *A3B* and gain an evolutionary advantage. This possibility is also supported by frequent lesions in the RB/E2F pathway in breast cancer, including loss of *RB1*, *CDKN1B* (encoding p27), and *CDKN2A* as well as amplification of *CCND1* (Angus et al., 2019; Bertucci et al., 2019; Cancer Genome Atlas Network, 2015; Ertel et al., 2010; Nik-Zainal et al., 2016; Cancer Genome Atlas Network, 2012).

Our studies also raise the possibility that high levels of expression of a set of 20 normally E2F-repressed genes may be used to identify tumors with elevated levels of APOBEC signature mutations (Figure 6A–D and Figure 6—figure supplement 1). Such information could be useful, for instance, to help identify the subset of patients with hypermutated tumors that may be most responsive to immunotherapy. It is also interesting that *A3B* mRNA levels do not associate as strongly with APOBEC signature mutation loads or enrichment values (Figure 6E–H). This discordance is unexpected and may be due to a combination of factors including cell cycle dysregulation (magnitude and mechanism), DNA damage response and DNA repair capabilities (including p53 functionality), tumor microenvironment (including inflammation and infection status), and possible contributions from related APOBEC3 family members including A3A and A3H. For instance, a more rapid cell cycle, dampened or disabled cell cycle checkpoints, downregulated (or saturated) DNA repair mechanisms, and potential coordination with A3A (Figure 2A) may combine to create conditions favoring an overall accumulation of APOBEC signature mutations. The overall APOBEC signature may also be influenced by an *A3A-B* fusion allele but its low frequency in TCGA data sets precluded analysis here. We are particularly intrigued by the potential for synergistic *A3B* induction by simultaneous E2F de-repression as part of cell cycle dysregulation and inflammation (modeled here by PyV tTag expression and PMA treatment, respectively, in Figure 3F–G). These perturbations, especially in combination with others such as viral or mutational inactivation of p53, may both activate the APOBEC mutation program and create an optimal environment for DNA damage tolerance, mutation accumulation, and tumor evolution.

Author Contributions

Pieter A Roelofs, Formal analysis, Investigation, Methodology, Writing - original draft, Writing -review and editing; Chai Yeen Goh, Data curation, Investigation, Methodology, Writing - review and editing; Boon Haow Chua, Matthew C Jarvis, Formal analysis, Investigation, Methodology, Writing - review and editing; Teneale A Stewart, Jennifer L McCann, Investigation, Methodology, Writing - review and editing; Rebecca M McDougale, Michael A Carpenter, Investigation, Writing - review and editing; John WM Martens, Paul N Span, Supervision, Funding acquisition, Methodology, Writing - review and editing; Dennis Kappei, Formal analysis, Supervision, Funding acquisition, Methodology, Writing - original draft, Writing - review and editing; Reuben S Harris, Conceptualization, Formal analysis, Supervision, Writing - original draft, Project administration, Writing - review and editing

Funding

This work was supported, in part, by the KWF Dutch Cancer Society (KWF10270, to J.W.M.M., P.N.S., and R.S.H.) and by the National Cancer Institute (P01-CA234228, to R.S.H.). R.S.H. is the Ewing Halsell President’s Council Distinguished Chair at University of Texas San Antonio, a CPRIT Scholar, and an Investigator of the Howard Hughes Medical Institute. None of the funding agencies had any role in the conceptualization, study design, data collection, interpretation of results, or the decision to submit this work for publication.

Data Availability

Raw mass spectrometry data is accessible through the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD020473. Additional data generated or analysed during this study are included in the manuscript and supporting files.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Kappei D	2020	APOBEC3B promoter interactors	<a href="https://www.ebi.ac.uk/pride/archive/projects/PXD020473">https://www.ebi.ac.uk/pride/archive/projects/PXD020473</a>	PRIDE, PXD020473

Acknowledgments

We thank Yanjun Chen and Bojana Stefanovska for helpful comments, James DeCaprio for constructive feedback, Diako Ebrahimi for help analyzing promoter conservation, Walker Lahr and Brandon Moriarity for advice on HDR, Shuyu Meng for early ChIP contributions, and Jesenia Perez and Daniel Salamango for sharing



preliminary data. The results presented here are in part based upon data generated by the TCGA Research Network: <http://www.cancer.gov/tcga>.

### **Conflicts of Interest**

Reuben S Harris: RSH is a co-founder, shareholder, and consultant of ApoGen Biotechnologies Inc. The other authors declare that no competing interests exist.

## References

- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjörd JE, Foekens JA, Greaves M, et al. 2013. Signatures of mutational processes in human cancer. *Nature* 500:415–421. DOI: <https://doi.org/10.1038/nature12477>, PMID: 23945592
- Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, Islam SMA, Lopez-Bigas N, Klimczak LJ, McPherson JR, Morganella S, Sabarinathan R, Wheeler DA, Mustonen V, Getz G, Rozen SG, et al. 2020. The repertoire of mutational signatures in human cancer. *Nature* 578:94–101. DOI: <https://doi.org/10.1038/s41586-020-1943-3>, PMID: 32025018
- An P, Saenz Robles MT, Pipas JM. 2012. Large T antigens of polyomaviruses: amazing molecular machines. *Annual Review of Microbiology* 66:213–236. DOI: <https://doi.org/10.1146/annurev-micro-092611-150154>, PMID: 22994493
- Angus L, Smid M, Wilting SM, van Riet J, Van Hoeck A, Nguyen L, Nik-Zainal S, Steenbruggen TG, Tjan-Heijnen VCG, Labots M, van Riel J, Bloemendal HJ, Steeghs N, Lolkema MP, Voest EE, van de Werken HJG, Jager A, Cuppen E, Sleijfer S, Martens JWM. 2019. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nature Genetics* 51:1450–1458. DOI: <https://doi.org/10.1038/s41588-019-0507-7>, PMID: 31570896
- Bellacchio E, Paggi MG. 2013. Understanding the targeting of the RB family proteins by viral oncoproteins to defeat their oncogenic machinery. *Journal of Cellular Physiology* 228:285–291. DOI: <https://doi.org/10.1002/jcp.24137>, PMID: 22718244
- Bertucci F, Ng KY, Patsouris A, Droin N, Piscuoglio S, Carbuccia N, Soria JC, Dien AT, Adnani Y, Kamal M, Garnier S, Meurice G, Jimenez M, Dogan S, Verret B, Chaffanet M, Bachelot T, Campone M, Lefeuvre C, Bonnefoi H, et al. 2019. Genomic characterization of metastatic breast cancers. *Nature* 569:560–564. DOI: <https://doi.org/10.1038/s41586-019-1056-z>, PMID: 31118521
- Boger HP, Wiegand HL, Hulme AE, Garcia-Perez JL, O'Shea KS, Moran JV, Cullen BR. 2006. Cellular inhibitors of long interspersed element 1 and Alu retrotransposition. *PNAS* 103:8780–8785. DOI: <https://doi.org/10.1073/pnas.0603313103>, PMID: 16728505
- Brown WL, Law EK, Argyris PP, Carpenter MA, Levin-Klein R, Ranum AN, Molan AM, Forster CL, Anderson BD, Lackey L, Harris RS. 2019. A rabbit monoclonal antibody against the antiviral and cancer genomic DNA mutating enzyme APOBEC3B. *Antibodies* 8:47. DOI: <https://doi.org/10.3390/antib8030047>
- Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, Lawrence MS. 2019. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* 364:eaaw2872. DOI: <https://doi.org/10.1126/science.aaw2872>, PMID: 31249028
- Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, Yee D, Temiz NA, Donohue DE, McDougale RM, Brown WL, Law EK, Harris RS. 2013a. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 494:366–370. DOI: <https://doi.org/10.1038/nature11881>, PMID: 23389445
- Burns MB, Temiz NA, Harris RS. 2013b. Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nature Genetics* 45:977–983. DOI: <https://doi.org/10.1038/ng.2701>, PMID: 23852168
- Cancer Genome Atlas Network. 2012. Comprehensive molecular portraits of human breast tumours. *Nature* 490:61–70. DOI: <https://doi.org/10.1038/nature11412>, PMID: 23000897
- Cancer Genome Atlas Network. 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 517:576–582. DOI: <https://doi.org/10.1038/nature14129>, PMID: 25631445
- Cancer Genome Atlas Research Network. 2017. Integrated genomic and molecular characterization of cervical cancer. *Nature* 543:378–384. DOI: <https://doi.org/10.1038/nature21386>, PMID: 28112728

- Cannataro VL, Gaffney SG, Sasaki T, Issaeva N, Grewal NKS, Grandis JR, Yarbrough WG, Burtneess B, Anderson KS, Townsend JP. 2019. APOBEC-induced mutations and their cancer effect size in head and neck squamous cell carcinoma. *Oncogene* 38:3475–3487. DOI: <https://doi.org/10.1038/s41388-018-0657-6>, PMID: 30647454
- Cao L, Peng B, Yao L, Zhang X, Sun K, Yang X, Yu L. 2010. The ancient function of RB-E2F pathway: insights from its evolutionary history. *Biology Direct* 5:55. DOI: <https://doi.org/10.1186/1745-6150-5-55>, PMID: 20849664
- Carpenter MA, Law EK, Serebrenik A, Brown WL, Harris RS. 2019. A lentivirus-based system for Cas9/gRNA expression and subsequent removal by Cre-mediated recombination. *Methods* 156:79–84. DOI: <https://doi.org/10.1016/j.jymeth.2018.12.006>, PMID: 30578845
- Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc EP, Kim J, Kwiatkowski DJ, Fargo DC, Mieczkowski PA, Getz G, Gordenin DA. 2015. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. *Nature Genetics* 47:1067–1072. DOI: <https://doi.org/10.1038/ng.3378>, PMID: 26258849
- Christensen J, Cloos P, Toftegaard U, Klinkenberg D, Bracken AP, Trinh E, Heeran M, Di Stefano L, Helin K. 2005. Characterization of E2F8, a novel E2F-like cell-cycle regulated repressor of E2F-activated transcription. *Nucleic Acids Research* 33:5458–5470. DOI: <https://doi.org/10.1093/nar/gki855>, PMID: 16179649
- Cortez LM, Brown AL, Dennis MA, Collins CD, Brown AJ, Mitchell D, Mertz TM, Roberts SA. 2019. APOBEC3A is a prominent cytidine deaminase in breast cancer. *PLOS Genetics* 15:e1008545. DOI: <https://doi.org/10.1371/journal.pgen.1008545>, PMID: 31841499
- Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* 26:1367–1372. DOI: <https://doi.org/10.1038/nbt.1511>, PMID: 19029910
- Cuitiño MC, Pécot T, Sun D, Kladney R, Okano-Uchida T, Shinde N, Saeed R, Perez-Castro AJ, Webb A, Liu T, Bae SI, Clijsters L, Selner N, Coppola V, Timmers C, Ostrowski MC, Pagano M, Leone G. 2019. Two distinct E2F transcriptional modules drive cell cycles and differentiation. *Cell Reports* 27:3547–3560. DOI: <https://doi.org/10.1016/j.celrep.2019.05.004>, PMID: 31130414
- de Bruin A, Maiti B, Jakoi L, Timmers C, Buerki R, Leone G. 2003. Identification and characterization of E2F7, a novel mammalian E2F family member capable of blocking cellular proliferation. *Journal of Biological Chemistry* 278:42041–42049. DOI: <https://doi.org/10.1074/jbc.M308105200>, PMID: 12893818
- de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, Jamal-Hanjani M, Shafi S, Murugaesu N, Rowan AJ, Grönroos E, Muhammad MA, Horswell S, Gerlinger M, Varela I, Jones D, Marshall J, Voet T, Van Loo P, Rasmussen DM, et al. 2014. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 346:251–256. DOI: <https://doi.org/10.1126/science.1253462>, PMID: 25301630
- Dean JL, McClendon AK, Knudsen ES. 2012. Modification of the DNA damage response by therapeutic CDK4/6 inhibition. *Journal of Biological Chemistry* 287:29075–29087. DOI: <https://doi.org/10.1074/jbc.M112.365494>, PMID: 22733811
- DeCaprio JA. 2014. Human papillomavirus type 16 E7 perturbs DREAM to promote cellular proliferation and mitotic gene expression. *Oncogene* 33:4036–4038. DOI: <https://doi.org/10.1038/onc.2013.449>, PMID: 24166507
- DeCaprio JA, Garcea RL. 2013. A cornucopia of human polyomaviruses. *Nature Reviews Microbiology* 11:264–276. DOI: <https://doi.org/10.1038/nrmicro2992>, PMID: 23474680
- Engeland K. 2018. Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. *Cell Death & Differentiation* 25:114–132. DOI: <https://doi.org/10.1038/cdd.2017.172>, PMID: 29125603

- Ertel A, Dean JL, Rui H, Liu C, Witkiewicz AK, Knudsen KE, Knudsen ES. 2010. RB-pathway disruption in breast cancer: differential association with disease subtypes, disease-specific prognosis and therapeutic response. *Cell Cycle* 9:4153–4163. DOI: <https://doi.org/10.4161/cc.9.20.13454>, PMID: 20948315
- Faden DL, Thomas S, Cantalupo PG, Agrawal N, Myers J, DeRisi J. 2017. Multi-modality analysis supports APOBEC as a major source of mutations in head and neck squamous cell carcinoma. *Oral Oncology* 74:8–14. DOI: <https://doi.org/10.1016/j.oraloncology.2017.09.002>, PMID: 29103756
- Fischer M, Müller GA. 2017. Cell cycle transcription control: dream/MuvB and RB-E2F complexes. *Critical Reviews in Biochemistry and Molecular Biology* 52:638–662. DOI: <https://doi.org/10.1080/10409238.2017.1360836>, PMID: 28799433
- Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, Modi BP, Correard S, Gheorghe M, Baranasic D, Santana-Garcia W, Tan G, Cheneby J, Ballester B, Parcy F, Sandelin A, Lenhard B, Wasserman WW, Mathelier A. 2020. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research* 48: gkz1001. DOI: <https://doi.org/10.1093/nar/gkz1001>
- Gillison ML, Akagi K, Xiao W, Jiang B, Pickard RKL, Li J, Swanson BJ, Agrawal AD, Zucker M, Stache-Crain B, Emde AK, Geiger HM, Robine N, Coombes KR, Symer DE. 2019. Human papillomavirus and the landscape of secondary genetic alterations in oral cancers. *Genome Research* 29:1–17. DOI: <https://doi.org/10.1101/gr.241141.118>, PMID: 30563911
- Glaser AP, Fantini D, Wang Y, Yu Y, Rimar KJ, Podojil JR, Miller SD, Meeks JJ. 2018. APOBEC-mediated mutagenesis in urothelial carcinoma is associated with improved survival, mutations in DNA damage response genes, and immune response. *Oncotarget* 9:4537–4548. DOI: <https://doi.org/10.18632/oncotarget.23344>, PMID: 29435122
- Green AM, Weitzman MD. 2019. The spectrum of APOBEC3 activity: from anti-viral agents to anti-cancer opportunities. *DNA Repair* 83:102700. DOI: <https://doi.org/10.1016/j.dnarep.2019.102700>, PMID: 31563041
- Harris RS, Dudley JP. 2015. APOBECs and virus restriction. *Virology* 479-480:131–145. DOI: <https://doi.org/10.1016/j.virol.2015.03.012>
- Helleday T, Shtad S, Nik-Zainal S. 2014. Mechanisms underlying mutational signatures in human cancers. *Nature Reviews Genetics* 15:585–598. DOI: <https://doi.org/10.1038/nrg3729>, PMID: 24981601
- Henderson S, Chakravarthy A, Su X, Boshoff C, Fenton TR. 2014. APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. *Cell Reports* 7:1833–1841. DOI: <https://doi.org/10.1016/j.celrep.2014.05.012>, PMID: 24910434
- Ito J, Gifford RJ, Sato K. 2020. Retroviruses drive the rapid evolution of mammalian APOBEC3 genes. *PNAS* 117:610–618. DOI: <https://doi.org/10.1073/pnas.1914183116>, PMID: 31843890
- Jarvis MC, Ebrahimi D, Temiz NA, Harris RS. 2018. Mutation signatures including APOBEC in cancer cell lines. *JNCI Cancer Spectrum* 2:pk002. DOI: <https://doi.org/10.1093/jncics/pky002>, PMID: 29888758
- Kanu N, Cerone MA, Goh G, Zalmas LP, Bartkova J, Dietzen M, McGranahan N, Rogers R, Law EK, Gromova I, Kschischo M, Walton MI, Rossanese OW, Bartek J, Harris RS, Venkatesan S, Swanton C. 2016. DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. *Genome Biology* 17:185. DOI: <https://doi.org/10.1186/s13059-016-1042-9>, PMID: 27634334
- Kappei D, Scheibe M, Paszkowski-Rogacz M, Bluhm A, Gossmann TI, Dietz S, Dejung M, Herlyn H, Buchholz F, Mann M, Butter F. 2017. Phylointeractomics reconstructs functional evolution of protein binding. *Nature Communications* 8:14334. DOI: <https://doi.org/10.1038/ncomms14334>, PMID: 28176777
- Kidd JM, Newman TL, Tuzun E, Kaul R, Eichler EE. 2007. Population stratification of a common APOBEC gene deletion polymorphism. *PLOS Genetics* 3:e63. DOI: <https://doi.org/10.1371/journal.pgen.0030063>, PMID: 17447845

- Lackey L, Demorest ZL, Land AM, Hultquist JF, Brown WL, Harris RS. 2012. APOBEC3B and AID have similar nuclear import mechanisms. *Journal of Molecular Biology* 419:301–314. DOI: <https://doi.org/10.1016/j.jmb.2012.03.011>, PMID: 22446380
- Lackey L, Law EK, Brown WL, Harris RS. 2013. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell Cycle* 12:762–772. DOI: <https://doi.org/10.4161/cc.23713>, PMID: 23388464
- Law EK, Sieuwerts AM, LaPara K, Leonard B, Starrett GJ, Molan AM, Temiz NA, Vogel RI, Meijer-van Gelder ME, Sweep FC, Span PN, Foekens JA, Martens JW, Yee D, Harris RS. 2016. The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. *Science Advances* 2:e1601737. DOI: <https://doi.org/10.1126/sciadv.1601737>, PMID: 27730215
- Lee JY, Schizas M, Geyer FC, Selenica P, Piscuoglio S, Sakr RA, Ng CKY, Carniello JVS, Towers R, Giri DD, de Andrade VP, Papanastasiou AD, Viale A, Harris RS, Solit DB, Weigelt B, Reis-Filho JS, King TA. 2019. Lobular carcinomas in situ display intralesion genetic heterogeneity and clonal evolution in the progression to invasive lobular carcinoma. *Clinical Cancer Research* 25:674–686. DOI: <https://doi.org/10.1158/1078-0432.CCR-18-1103>, PMID: 30185420
- Leonard B, McCann JL, Starrett GJ, Kosyakovsky L, Luengas EM, Molan AM, Burns MB, McDougale RM, Parker PJ, Brown WL, Harris RS. 2015. The PKC/NF- $\kappa$ B signaling pathway induces APOBEC3B expression in multiple human cancers. *Cancer Research* 75:4538–4547. DOI: <https://doi.org/10.1158/0008-5472.CAN-15-2171-T>, PMID: 26420215
- Litovchick L, Sadasivam S, Florens L, Zhu X, Swanson SK, Velmurugan S, Chen R, Washburn MP, Liu XS, DeCaprio JA. 2007. Evolutionarily conserved multisubunit RBL2/p130 and E2F4 protein complex represses human cell cycle-dependent genes in quiescence. *Molecular Cell* 26:539–551. DOI: <https://doi.org/10.1016/j.molcel.2007.04.015>, PMID: 17531812
- Litovchick L, Florens LA, Swanson SK, Washburn MP, DeCaprio JA. 2011. DYRK1A protein kinase promotes quiescence and senescence through DREAM complex assembly. *Genes & Development* 25:801–813. DOI: <https://doi.org/10.1101/gad.203421>, PMID: 21498570
- Lucifora J, Xia Y, Reisinger F, Zhang K, Stadler D, Cheng X, Sprinzl MF, Koppensteiner H, Makowska Z, Volz T, Remouchamps C, Chou WM, Thasler WE, Hüser N, Durantel D, Liang TJ, Münk C, Heim MH, Browning JL, DeJardin E, et al. 2014. Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA. *Science* 343:1221–1228. DOI: <https://doi.org/10.1126/science.1243462>, PMID: 24557838
- Maruyama W, Shirakawa K, Matsui H, Matsumoto T, Yamazaki H, Sarca AD, Kazuma Y, Kobayashi M, Shindo K, Takaori-Kondo A. 2016. Classical NF- $\kappa$ B pathway is responsible for APOBEC3B expression in cancer cells. *Biochemical and Biophysical Research Communications* 478:1466–1471. DOI: <https://doi.org/10.1016/j.bbrc.2016.08.148>, PMID: 27577680
- Menendez D, Nguyen TA, Snipe J, Resnick MA. 2017. The cytidine deaminase APOBEC3 family is subject to transcriptional regulation by p53. *Molecular Cancer Research* 15:735–743. DOI: <https://doi.org/10.1158/1541-7786.MCR-17-0019>, PMID: 28232385
- Mori S, Takeuchi T, Ishii Y, Kukimoto I. 2015. Identification of APOBEC3B promoter elements responsible for activation by human papillomavirus type 16 E6. *Biochemical and Biophysical Research Communications* 460:555–560. DOI: <https://doi.org/10.1016/j.bbrc.2015.03.068>, PMID: 25800874
- Mori S, Takeuchi T, Ishii Y, Yugawa T, Kiyono T, Nishina H, Kukimoto I. 2017. Human papillomavirus 16 E6 upregulates APOBEC3B via the TEAD transcription factor. *Journal of Virology* 91:16. DOI: <https://doi.org/10.1128/JVI.02413-16>
- Müller GA, Quaas M, Schümann M, Krause E, Padi M, Fischer M, Litovchick L, DeCaprio JA, Engeland K. 2012. The CHR promoter element controls cell cycle-dependent gene transcription and binds the DREAM and MMB complexes. *Nucleic Acids Research* 40:1561–1578. DOI: <https://doi.org/10.1093/nar/gkr793>, PMID: 22064854

- Müller GA, Wintsche A, Stangner K, Prohaska SJ, Stadler PF, Engeland K. 2014. The CHR site: definition and genome-wide identification of a cell cycle transcriptional element. *Nucleic Acids Research* 42:10331–10350. DOI: <https://doi.org/10.1093/nar/gku696>, PMID: 25106871
- Müller GA, Stangner K, Schmitt T, Wintsche A, Engeland K. 2017. Timing of transcription during the cell cycle: protein complexes binding to E2F, E2F/CLE, CDE/CHR, or CHR promoter elements define early and late cell cycle gene expression. *Oncotarget* 8:97736–97748. DOI: <https://doi.org/10.18632/oncotarget.10888>
- Ng JCF, Quist J, Grigoriadis A, Malim MH, Fraternali F. 2019. Pan-cancer transcriptomic analysis dissects immune and proliferative functions of APOBEC3 cytidine deaminases. *Nucleic Acids Research* 47:1178–1194. DOI: <https://doi.org/10.1093/nar/gky1316>, PMID: 30624727
- Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, et al. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell* 149:979–993. DOI: <https://doi.org/10.1016/j.cell.2012.04.024>, PMID: 22608084
- Nik-Zainal S, Wedge DC, Alexandrov LB, Petljak M, Butler AP, Bolli N, Davies HR, Knappskog S, Martin S, Papaemmanuil E, Ramakrishna M, Shlien A, Simoncic I, Xue Y, Tyler-Smith C, Campbell PJ, Stratton MR. 2014. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. *Nature Genetics* 46:487–491. DOI: <https://doi.org/10.1038/ng.2955>, PMID: 24728294
- Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC, Van Loo P, Ju YS, Smid M, Brinkman AB, Morganella S, Aure MR, Lingjærde OC, Langerød A, Ringnér M, Ahn SM, et al. 2016. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 534:47–54. DOI: <https://doi.org/10.1038/nature17676>, PMID: 27135926
- Nikkilä J, Kumar R, Campbell J, Brandsma I, Pemberton HN, Wallberg F, Nagy K, Scheer I, Vertessy BG, Serebrenik AA, Monni V, Harris RS, Pettitt SJ, Ashworth A, Lord CJ. 2017. Elevated APOBEC3B expression drives a kataegic-like mutation signature and replication stress-related therapeutic vulnerabilities in p53-defective cells. *British Journal of Cancer* 117:113–123. DOI: <https://doi.org/10.1038/bjc.2017.133>, PMID: 28535155
- Olson ME, Harris RS, Harki DA. 2018. APOBEC enzymes as targets for virus and cancer therapy. *Cell Chemical Biology* 25:36–49. DOI: <https://doi.org/10.1016/j.chembiol.2017.10.007>, PMID: 29153851
- Pak V, Heidecker G, Pathak VK, Derse D. 2011. The role of amino-terminal sequences in cellular localization and antiviral activity of APOBEC3B. *Journal of Virology* 85:8538–8547. DOI: <https://doi.org/10.1128/JVI.02645-10>, PMID: 21715505
- Periyasamy M, Singh AK, Gemma C, Kranjec C, Farzan R, Leach DA, Navaratnam N, Pa'linka's HL, Vértessy BG, Fenton TR, Doorbar J, Fuller-Pace F, Meek DW, Coombes RC, Buluwela L, Ali S. 2017. p53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. *Nucleic Acids Research* 45:11056–11069. DOI: <https://doi.org/10.1093/nar/gkx721>, PMID: 28977491
- Pilkinton M, Sandoval R, Colamonici OR. 2007. Mammalian mip/LIN-9 interacts with either the p107, p130/E2F4 repressor complex or B-Myb in a cell cycle-phase-dependent context distinct from the *Drosophila* dREAM complex. *Oncogene* 26:7535–7543. DOI: <https://doi.org/10.1038/sj.onc.1210562>, PMID: 17563750
- Prykhodzhiy SV, Fuller C, Steele SL, Veinotte CJ, Razaghi B, Robitaille JM, McMaster CR, Shlien A, Malkin D, Berman JN. 2018. Optimized knock-in of point mutations in zebrafish using CRISPR/Cas9. *Nucleic Acids Research* 46:e102. DOI: <https://doi.org/10.1093/nar/gky512>, PMID: 29905858

- Qin J, Whyte WA, Anderssen E, Apostolou E, Chen HH, Akbarian S, Bronson RT, Hochedlinger K, Ramaswamy S, Young RA, Hock H. 2012. The polycomb group protein L3mbtl2 assembles an atypical PRC1-family complex that is essential in pluripotent stem cells and early development. *Cell Stem Cell* 11:319–332. DOI: <https://doi.org/10.1016/j.stem.2012.06.002>, PMID: 22770845
- Rashid NN, Rothan HA, Yusoff MS. 2015. The association of mammalian dream complex and HPV16 E7 proteins. *American Journal of Cancer Research* 5:3525–3533. PMID: 26885443
- Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS. 2010. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Research* 38:4274–4284. DOI: <https://doi.org/10.1093/nar/gkq174>, PMID: 20308164
- Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. 2016. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature Biotechnology* 34:339–344. DOI: <https://doi.org/10.1038/nbt.3481>, PMID: 26789497
- Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G, Harris S, Shah RR, Resnick MA, Getz G, Gordenin DA. 2013. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nature Genetics* 45:970–976. DOI: <https://doi.org/10.1038/ng.2702>, PMID: 23852170
- Roberts SA, Gordenin DA. 2014. Hypermutation in human cancer genomes: footprints and mechanisms. *Nature Reviews Cancer* 14:786–800. DOI: <https://doi.org/10.1038/nrc3816>, PMID: 25568919
- Roper N, Gao S, Maity TK, Banday AR, Zhang X, Venugopalan A, Cultraro CM, Patidar R, Sindiri S, Brown AL, Goncarencu A, Panchenko AR, Biswas R, Thomas A, Rajan A, Carter CA, Kleiner DE, Hewitt SM, Khan J, Prokunina-Olsson L, et al. 2019. APOBEC mutagenesis and copy-number alterations are drivers of proteogenomic tumor evolution and heterogeneity in metastatic thoracic tumors. *Cell Reports* 26:2651–2666. DOI: <https://doi.org/10.1016/j.celrep.2019.02.028>, PMID: 30840888
- Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. 2016. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biology* 17:31. DOI: <https://doi.org/10.1186/s13059-016-0893-4>, PMID: 26899170
- Sadasivam S, DeCaprio JA. 2013. The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nature Reviews Cancer* 13:585–595. DOI: <https://doi.org/10.1038/nrc3556>, PMID: 23842645
- Salamango DJ, McCann JL, Demirö, Brown WL, Amaro RE, Harris RS. 2018. APOBEC3B nuclear localization requires two distinct N-terminal domain surfaces. *Journal of Molecular Biology* 430:2695–2708. DOI: <https://doi.org/10.1016/j.jmb.2018.04.044>, PMID: 29787764
- Scelfo A, Fernández-Pérez D, Tamburri S, Zanotti M, Lavarone E, Soldi M, Bonaldi T, Ferrari KJ, Pasini D. 2019. Functional landscape of PCGF proteins reveals both RING1A/B-dependent and RING1A/B-independent specific activities. *Molecular Cell* 74:1037–1052. DOI: <https://doi.org/10.1016/j.molcel.2019.04.002>, PMID: 31029542
- Serebrenik AA, Starrett GJ, Leenen S, Jarvis MC, Shaban NM, Salamango DJ, Nilsen H, Brown WL, Harris RS. 2019. The deaminase APOBEC3B triggers the death of cells lacking uracil DNA glycosylase. *PNAS* 116:22158–22163. DOI: <https://doi.org/10.1073/pnas.1904024116>, PMID: 31611371
- Serebrenik AA, Argyris PP, Jarvis MC, Brown WL, Bazzaro M, Vogel RI, Erickson BK, Lee SH, Goergen KM, Maurer MJ, Heinzen EP, Oberg AL, Huang Y, Hou X, Werooha SJ, Kaufmann SH, Harris RS. 2020. The DNA cytosine deaminase APOBEC3B is a molecular determinant of platinum responsiveness in clear cell ovarian cancer. *Clinical Cancer Research* 26:3397–3407. DOI: <https://doi.org/10.1158/1078-0432.CCR-19-2786>, PMID: 32060098



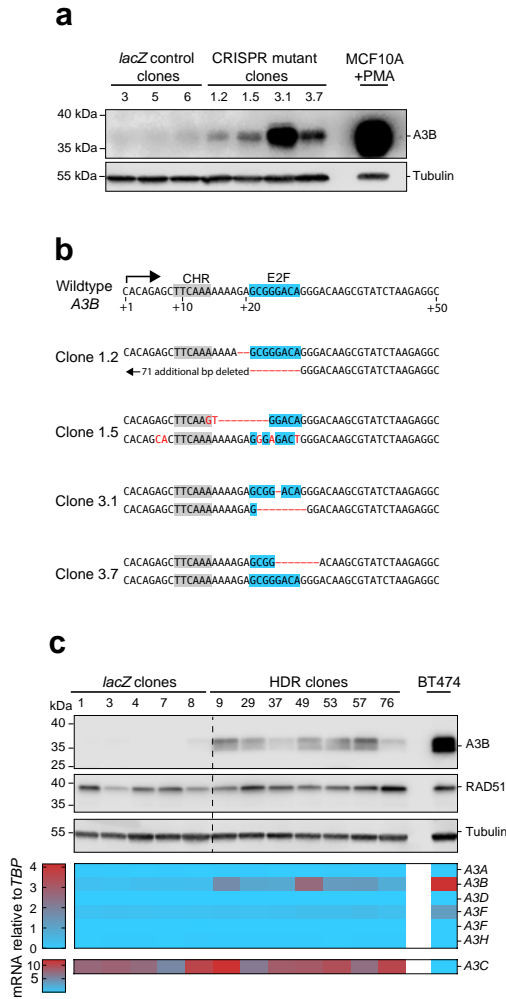
- Sieuwerts AM, Willis S, Burns MB, Look MP, Meijer-Van Gelder ME, Schlicker A, Heideman MR, Jacobs H, Wessels L, Leyland-Jones B, Gray KP, Foekens JA, Harris RS, Martens JW. 2014. Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. *Hormones and Cancer* 5:405–413. DOI: <https://doi.org/10.1007/s12672-014-0196-8>, PMID: 25123150
- Sieuwerts AM, Schrijver WA, Dalm SU, de Weerd V, Moelans CB, Ter Hoeve N, van Diest PJ, Martens JW, van Deurzen CH. 2017. Progressive APOBEC3B mRNA expression in distant breast cancer metastases. *PLOS ONE* 12:e0171343. DOI: <https://doi.org/10.1371/journal.pone.0171343>, PMID: 28141868
- Silvas TV, Schiffer CA. 2019. APOBEC3s: DNA-editing human cytidine deaminases. *Protein Science* 28:1552–1566. DOI: <https://doi.org/10.1002/pro.3670>, PMID: 31241202
- Simon V, Bloch N, Landau NR. 2015. Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. *Nature Immunology* 16:546–553. DOI: <https://doi.org/10.1038/ni.3156>, PMID: 25988886
- Siriwardena SU, Chen K, Bhagwat AS. 2016. Functions and malfunctions of mammalian DNA-cytosine deaminases. *Chemical Reviews* 116:12688–12710. DOI: <https://doi.org/10.1021/acs.chemrev.6b00296>, PMID: 27585283
- Starrett GJ, Luengas EM, McCann JL, Ebrahimi D, Temiz NA, Love RP, Feng Y, Adolph MB, Chelico L, Law EK, Carpenter MA, Harris RS. 2016. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *Nature Communications* 7:12918. DOI: <https://doi.org/10.1038/ncomms12918>, PMID: 27650891
- Starrett GJ, Serebrenik AA, Roelofs PA, McCann JL, Verhalen B, Jarvis MC, Stewart TA, Law EK, Krupp A, Jiang M, Martens JWM, Cahir-McFarland E, Span PN, Harris RS. 2019. Polyomavirus T antigen induces APOBEC3B expression using an LXCXE-dependent and TP53-independent mechanism. *mBio* 10:e02690-18. DOI: <https://doi.org/10.1128/mBio.02690-18>, PMID: 30723127
- Stenglein MD, Matsuo H, Harris RS. 2008. Two regions within the amino-terminal half of APOBEC3G cooperate to determine cytoplasmic localization. *Journal of Virology* 82:9591–9599. DOI: <https://doi.org/10.1128/JVI.02471-07>, PMID: 18667511
- Stielow B, Finkernagel F, Stiewe T, Nist A, Suske G. 2018. MGA, L3MBTL2 and E2F6 determine genomic binding of the non-canonical polycomb repressive complex PRC1.6. *PLOS Genetics* 14:e1007193. DOI: <https://doi.org/10.1371/journal.pgen.1007193>, PMID: 29381691
- Swanton C, McGranahan N, Starrett GJ, Harris RS. 2015. APOBEC enzymes: mutagenic fuel for cancer evolution and heterogeneity. *Cancer Discovery* 5:704–712. DOI: <https://doi.org/10.1158/2159-8290.CD-15-0344>, PMID: 26091828
- Taylor BJ, Nik-Zainal S, Wu YL, Stebbings LA, Raine K, Campbell PJ, Rada C, Stratton MR, Neuberger MS. 2013. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *eLife* 2:e00534. DOI: <https://doi.org/10.7554/eLife.00534>, PMID: 23599896
- Trimarchi JM, Fairchild B, Verona R, Moberg K, Andon N, Lees JA. 1998. E2F-6, a member of the E2F family that can behave as a transcriptional repressor. *PNAS* 95:2850–2855. DOI: <https://doi.org/10.1073/pnas.95.6.2850>, PMID: 9501179
- Uphoff CC, Drexler HG. 2011. Detecting Mycoplasma contamination in cell cultures by polymerase chain reaction. *Methods in Molecular Biology* 731:93–103. DOI: [https://doi.org/10.1007/978-1-61779-080-5\\_8](https://doi.org/10.1007/978-1-61779-080-5_8), PMID: 21516400
- Venkatesan S, Rosenthal R, Kanu N, McGranahan N, Bartek J, Quezada SA, Hare J, Harris RS, Swanton C. 2018. Perspective: apobec mutagenesis in drug resistance and immune escape in HIV and Cancer evolution. *Annals of Oncology* 29:563–572. DOI: <https://doi.org/10.1093/annonc/mdy003>, PMID: 29324969
- Verhalen B, Starrett GJ, Harris RS, Jiang M. 2016. Functional upregulation of the DNA cytosine deaminase APOBEC3B by polyomaviruses. *Journal of Virology* 90:6379–6386. DOI: <https://doi.org/10.1128/JVI.00771-16>, PMID: 27147740



- Verlinden L, Eelen G, Beullens I, Van Camp M, Van Hummelen P, Engelen K, Van Hellemont R, Marchal K, De Moor B, Fojier F, Te Riele H, Beullens M, Bollen M, Mathieu C, Bouillon R, Verstuyf A. 2005. Characterization of the condensin component Cnap1 and protein kinase Melk as novel E2F target genes down-regulated by 1,25-dihydroxyvitamin D3. *Journal of Biological Chemistry* 280:37319–37330. DOI: <https://doi.org/10.1074/jbc.M503587200>, PMID: 16144839
- Vieira VC, Leonard B, White EA, Starrett GJ, Temiz NA, Lorenz LD, Lee D, Soares MA, Lambert PF, Howley PM, Harris RS. 2014. Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *mBio* 5:e02234-14. DOI: <https://doi.org/10.1128/mBio.02234-14>, PMID: 25538195
- Vizcaino JA, Csordas A, Del-Toro N, Dianas JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H. 2016. 2016 update of the PRIDE database and its related tools. *Nucleic Acids Research* 44:11033. DOI: <https://doi.org/10.1093/nar/gkw880>, PMID: 27683222
- Walker BA, Wardell CP, Murison A, Boyle EM, Begum DB, Dahir NM, Proszek PZ, Melchor L, Pawlyn C, Kaiser MF, Johnson DC, Qiang YW, Jones JR, Cairns DA, Gregory WM, Owen RG, Cook G, Drayson MT, Jackson GH, Davies FE, et al. 2015. APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. *Nature Communications* 6:6997. DOI: <https://doi.org/10.1038/ncomms7997>, PMID: 25904160
- Warren CJ, Xu T, Guo K, Griffin LM, Westrich JA, Lee D, Lambert PF, Santiago ML, Pyeon D. 2015. APOBEC3A functions as a restriction factor of human papillomavirus. *Journal of Virology* 89:688–702. DOI: <https://doi.org/10.1128/JVI.02383-14>, PMID: 25355878
- Westrich JA, Warren CJ, Klausner MJ, Guo K, Liu C-W, Santiago ML, Pyeon D. 2018. Human papillomavirus 16 E7 stabilizes APOBEC3A protein by inhibiting Cullin 2-dependent protein degradation. *Journal of Virology* 92: JVI.01318-17. DOI: <https://doi.org/10.1128/JVI.01318-17>
- White EA, Sowa ME, Tan MJ, Jeudy S, Hayes SD, Santha S, Münger K, Harper JW, Howley PM. 2012. Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses. *PNAS* 109:E260–E267. DOI: <https://doi.org/10.1073/pnas.1116776109>, PMID: 22232672
- Xu L, Chang Y, An H, Zhu Y, Yang Y, Xu J. 2015. High APOBEC3B expression is a predictor of recurrence in patients with low-risk clear cell renal cell carcinoma. *Urologic Oncology: Seminars and Original Investigations* 33:340.e1–34340. DOI: <https://doi.org/10.1016/j.urolonc.2015.05.009>
- Yamazaki H, Shirakawa K, Matsumoto T, Hirabayashi S, Murakawa Y, Kobayashi M, Sarca AD, Kazuma Y, Matsui H, Maruyama W, Fukuda H, Shirakawa R, Shindo K, Ri M, Iida S, Takaori-Kondo A. 2019. Endogenous APOBEC3B overexpression constitutively generates DNA substitutions and deletions in myeloma cells. *Scientific Reports* 9:7122. DOI: <https://doi.org/10.1038/s41598-019-43575-y>, PMID: 31073151
- Yamazaki H, Shirakawa K, Matsumoto T, Kazuma Y, Matsui H, Horisawa Y, Stanford E, Sarca AD, Shirakawa R, Shindo K, Takaori-Kondo A. 2020. APOBEC3B reporter myeloma cell lines identify DNA damage response pathways leading to APOBEC3B expression. *PLOS ONE* 15:e0223463. DOI: <https://doi.org/10.1371/journal.pone.0223463>, PMID: 31914134
- Yan S, He F, Gao B, Wu H, Li M, Huang L, Liang J, Wu Q, Li Y. 2016. Increased APOBEC3B predicts worse outcomes in lung cancer: a comprehensive retrospective study. *Journal of Cancer* 7:618–625. DOI: <https://doi.org/10.7150/jca.14030>, PMID: 27076842
- Zapatka M, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M, Desai N, Sülthmann H, Moch H, Cooper CS, Eils R, Ferretti V, Lichter P, PCAWG Pathogens, PCAWG Consortium. 2020. The landscape of viral associations in human cancers. *Nature Genetics* 52:320–330. DOI: <https://doi.org/10.1038/s41588-019-0558-9>, PMID: 32025001

# Supplementary Information

## Supplementary Figure S1

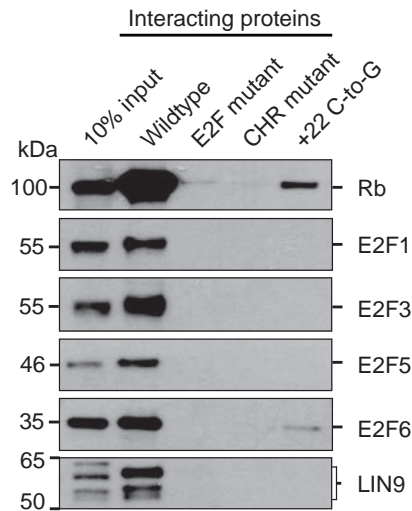


**Figure Supplement to Figure 3.** Exploratory clones showing disruption of A3B expression upon CRISPR/Cas9 targeting. **A** Immunoblot showing increased A3B expression in four independent clones after targeting the +21 to +28 E2F site using CRISPR/Cas9 technology. Sample on the far right indicates A3B expression upon PMA-treatment for scale.

**B** DNA sequences of the E2F region of the *A3B* promoter of the four CRISPR/Cas9-targeted clones in panel (A). Clones 1.2 and 1.5 were obtained using gRNA\_50UTR#1, and clones 3.1 and 3.7 were obtained using gRNA\_50UTR#3 (gRNA sequences in Supplementary file 4).

**C** Immunoblots from Figure 3C are shown again here for comparison with a heatmap summary of *APOBEC3* family member mRNA levels in the same control and HDR-targeted clones. *APOBEC3* mRNA levels were determined by RT-qPCR, normalized to *TBP*, and presented as a heatmap for comparison. A3C is shown on a separate scale due to relatively high expression in MCF10A.

Supplementary Figure S2



**Figure Supplement to Figure 4.** Confirmatory immunoblots for SILAC pull-down assay. Immunoblots for Rb, E2F1, E2F3, E2F5, E2F6, and LIN9 indicating binding of these proteins to the +1 to +50 *A3B* promoter sequence *in vitro*. Note that binding is vastly decreased in all mutated sequences. Specific mutations are identical to those mentioned in Figure 4B.



Supplementary Table S3

TCGA Sample ID	E2F Group	BUB1	CCNA2	CCNB2	CDC20	CDC45	CDCA5	CDCA8	CENPA	CEP55
TCGA-A1-A0SO	high	14,753	5,953	5,322	8,935	4,073	8,432	8,686	5,087	8,746
TCGA-A2-A04T	high	6,135	6,123	5,052	6,814	2,132	6,626	8,947	1,432	3,770
TCGA-A2-A0CX	high	6,675	6,307	5,008	5,162	3,494	5,683	3,223	1,585	6,117
TCGA-A2-A0D2	high	5,066	6,723	14,403	16,293	5,079	16,916	11,823	1,962	5,757
TCGA-A2-A25B	high	9,976	7,367	6,048	6,125	1,315	4,286	5,435	1,653	7,716
TCGA-A7-A13D	high	8,279	3,840	4,733	12,426	1,882	4,328	6,882	2,511	5,452
TCGA-A7-A13F	high	4,799	5,288	6,223	6,786	4,616	7,858	5,406	1,119	6,582
TCGA-A7-A6VY	high	4,847	5,248	3,063	11,165	1,545	6,252	5,116	1,755	3,103
TCGA-A8-A079	high	5,497	6,905	6,160	15,044	3,323	9,184	5,063	1,452	3,174
TCGA-A8-A07R	high	7,357	17,462	7,737	10,799	4,842	12,133	6,520	1,701	4,337
TCGA-A8-A07S	high	4,755	3,472	4,060	4,809	1,510	4,152	4,104	0,952	4,510
TCGA-A8-A08F	high	6,132	7,966	5,497	6,726	5,483	9,497	5,896	1,014	5,938
TCGA-A8-A08R	high	8,634	3,544	3,991	13,965	2,036	5,683	10,880	1,405	4,735
TCGA-AN-A04C	high	6,032	9,658	6,625	6,452	3,053	7,899	5,886	3,577	5,660
TCGA-AN-A0G0	high	4,653	4,002	5,459	7,232	4,851	5,042	3,870	1,094	4,541
TCGA-AN-A0XU	high	6,868	5,262	4,484	11,930	3,019	3,562	3,458	2,402	6,304
TCGA-AO-A03R	high	5,909	5,837	12,365	16,308	1,447	8,841	8,043	1,904	10,466
TCGA-AO-A0J2	high	11,225	7,502	20,958	25,120	6,639	15,141	10,702	3,921	8,791
TCGA-AO-A0J6	high	3,257	8,809	3,745	12,617	1,283	5,933	6,832	3,118	3,837
TCGA-AO-A0JM	high	3,146	4,939	5,440	6,620	2,226	6,640	4,557	0,889	3,631
TCGA-AO-A1KR	high	3,357	4,101	4,875	19,667	3,202	4,515	6,869	3,380	9,195
TCGA-AR-A0TP	high	7,293	3,241	3,128	7,436	2,457	6,630	4,712	1,217	5,673
TCGA-AR-A0U4	high	5,978	7,962	6,055	8,553	1,377	6,061	6,508	2,176	6,381
TCGA-AR-A1AY	high	3,769	3,423	3,187	9,803	2,234	5,566	5,640	0,854	4,640
TCGA-AR-A5QQ	high	2,803	4,824	5,367	13,207	2,194	6,171	3,810	1,404	4,599
TCGA-B6-A0RE	high	3,790	3,161	3,520	6,904	1,203	4,649	4,050	1,758	5,428
TCGA-B6-A0RU	high	3,775	4,720	5,558	9,180	2,220	4,358	7,146	1,488	2,757
TCGA-B6-A0X1	high	9,759	4,828	7,034	14,963	1,553	8,823	9,115	4,015	5,764
TCGA-B6-A0X5	high	3,363	6,187	4,687	6,335	1,532	5,534	3,873	1,393	4,725
TCGA-BH-A0B3	high	7,105	4,357	7,486	18,181	1,373	4,935	4,852	2,511	5,967
TCGA-BH-A18R	high	5,895	9,980	5,895	5,725	2,628	6,199	4,534	1,094	8,638
TCGA-BH-A18V	high	9,806	6,198	10,343	7,755	1,995	10,832	7,095	2,165	13,217
TCGA-C8-A12L	high	6,919	9,118	8,122	5,010	1,883	5,932	3,582	1,589	6,276
TCGA-C8-A134	high	7,240	5,411	7,098	8,256	1,204	6,959	6,804	5,005	4,241
TCGA-C8-A137	high	7,289	5,769	6,403	6,128	1,518	5,430	4,544	1,148	4,872
TCGA-C8-A275	high	4,387	5,554	3,986	4,118	1,516	3,714	3,550	1,469	4,411

	DEPDC1B	FOXM1	GTSE1	HJURP	KIF2C	KIFC1	MELK	PLK1	RAD51	TPX2	UBE2C
	2,438	25,258	4,388	3,206	9,662	10,916	5,608	11,345	2,518	62,278	5,204
	1,012	29,542	2,552	3,193	7,833	9,217	2,728	4,473	1,385	18,991	5,361
	0,955	11,706	2,215	3,215	4,642	5,004	3,238	10,502	2,302	20,223	8,540
	2,695	30,844	2,884	4,095	14,464	14,316	4,551	11,662	2,233	24,410	11,892
	2,123	7,307	1,911	4,819	5,788	4,452	2,550	7,234	1,760	24,423	6,800
	0,817	9,764	2,330	2,880	9,006	4,313	8,025	6,964	3,221	20,063	15,958
	1,613	22,903	2,487	2,664	7,513	9,160	4,063	3,500	5,223	16,035	14,214
	0,696	11,520	1,440	2,679	6,134	6,942	4,821	7,979	1,373	12,846	7,086
	1,740	15,630	3,358	2,339	11,942	6,135	2,565	8,516	2,387	12,908	15,210
	1,210	19,176	4,565	7,729	6,733	9,066	7,274	17,303	3,376	25,636	10,480
	1,027	10,571	2,275	2,296	4,725	6,821	2,552	5,176	1,412	11,600	3,955
	1,163	12,909	3,420	5,465	8,646	13,125	5,319	10,323	4,045	29,382	9,774
	0,985	15,896	3,428	3,697	6,911	8,906	3,427	4,612	2,923	15,514	6,346
	1,191	12,389	2,543	5,452	6,662	9,434	5,008	6,141	4,396	18,689	9,814
	1,002	10,753	4,819	3,563	4,722	4,016	4,009	6,458	1,674	11,160	4,287
	1,397	16,124	3,738	4,506	6,301	6,950	3,287	9,681	1,103	17,917	6,822
	2,490	13,538	4,279	7,163	12,687	14,519	5,043	5,692	3,058	25,269	22,837
	5,110	38,054	4,770	7,050	12,312	14,505	8,647	11,840	3,497	36,385	16,262
	0,981	10,272	2,222	2,808	6,905	6,123	5,773	6,270	1,368	17,239	5,298
	1,089	8,621	2,154	2,669	4,777	6,729	4,274	4,154	2,777	16,326	6,809
	1,394	19,895	2,956	5,613	9,660	6,593	3,236	4,387	1,374	18,478	18,300
	1,582	16,321	3,189	2,301	4,870	12,179	4,574	4,441	1,408	21,676	5,097
	1,162	20,976	3,061	2,555	4,860	3,872	2,919	7,551	1,611	16,101	7,982
	1,300	12,572	2,043	2,352	6,015	5,464	3,109	9,370	1,592	16,364	7,025
	1,431	9,618	2,110	3,577	6,277	6,847	4,341	6,762	2,221	10,465	9,331
	2,340	14,979	1,704	3,611	4,431	5,156	3,119	5,327	1,574	13,754	8,203
	1	8,224	3,723	2,454	6,043	7,020	4,094	5,263	1,001	18,047	5,403
	1,693	41,142	3,074	3,020	9,560	7,990	5,644	11,462	1,629	16,258	5,238
	2,421	12,486	3,212	2,867	4,509	5,708	2,878	4,251	2,333	11,064	5,436
	0,771	11,691	1,370	2,591	8,684	9,025	4,120	9,385	2,187	29,634	9,062
	1,450	21,642	3,028	3,975	6,947	5,528	4,160	7,966	1,810	17,438	5,596
	1,577	30,076	4,603	5,641	6,436	15,432	9,661	12,805	5,385	28,867	8,736
	1,557	20,617	6,645	2,518	4,914	4,274	3,484	10,782	2,181	21,369	8,134
	0,946	9,629	2,782	2,212	9,123	11,306	8,053	7,360	1,118	17,700	6,854
	1,285	9,268	3,082	1,793	6,974	3,987	3,230	5,702	1,400	17,010	6,692
	1,769	6,810	2,693	3,513	4,490	4,466	4,287	5,032	1,624	10,621	5,503

**Supplementary Table S3:** Continued

TCGA-C8-A27B	high	5,473	5,342	3,722	12,637	2,085	3,943	3,362	2,303	4,284
TCGA-D8-A1JD	high	3,795	5,538	3,222	5,120	1,594	4,465	2,272	0,727	6,507
TCGA-E2-A109	high	4,841	4,388	4,869	4,324	1,544	7,498	3,042	1,601	3,604
TCGA-E2-A14N	high	5,198	6,022	7,660	12,158	1,58	5,176	10,695	4,387	8,358
TCGA-E2-A14V	high	4,983	5,638	4,807	4,736	1,561	3,461	2,755	1,192	7,653
TCGA-E2-A1LH	high	2,446	2,855	3,586	7,054	2,144	4,775	4,180	1,158	4,117
TCGA-E2-A1LI	high	8,066	7,224	3,709	9,457	2,453	5,903	6,783	2,543	6,075
TCGA-E9-A1RH	high	6,697	11,832	4,319	8,116	2,454	6,171	3,869	5,492	5,154
TCGA-E9-A22G	high	3,669	3,451	4,444	8,783	3,138	6,689	8,867	3,881	5,648
TCGA-EW-A1P4	high	6,622	7,515	6,701	9,256	1,721	3,586	3,460	2,050	4,015
TCGA-EW-A1PC	high	5,839	3,421	3,092	8,818	1,645	4,269	3,523	2,251	6,060
TCGA-GM-A2DF	high	12,481	5,884	13,582	19,989	3,942	8,810	12,460	4,937	9,640
TCGA-GM-A3XL	high	4,136	7,957	8,049	11,030	4,326	9,197	4,379	2,041	3,473
TCGA-HN-A2NL	high	3,360	3,891	6,310	10,614	2,033	9,632	7,108	2,791	4,676
TCGA-LL-A8F5	high	4,339	4,521	4,945	16,517	4,920	6,787	3,881	1,570	3,279
TCGA-S3-AA10	high	4,735	8,999	3,140	9,098	1,554	6,452	5,832	2,155	7,353
TCGA-UU-A93S	high	3,935	5,045	5,659	10,705	5,004	11,831	2,720	2,322	6,651
TCGA-5L-AAT0	low	0,371	0,381	0,460	0,738	0,272	0,861	0,510	0,113	0,296
TCGA-5L-AAT1	low	0,493	0,220	0,351	1,067	0,067	0,483	0,301	0,057	0,142
TCGA-A2-A0CO	low	0,224	0,301	0,253	0,325	0,144	0,373	0,323	0,049	0,235
TCGA-A2-A0CP	low	0,104	0,154	0,287	0,501	0,028	0,392	0,318	0,068	0,194
TCGA-A2-A0CV	low	0,191	0,155	0,480	0,408	0,019	0,327	0,424	0,019	0,090
TCGA-A2-A0CZ	low	0,462	0,470	0,622	1,026	0,084	0,556	0,580	0,150	0,462
TCGA-A2-A0D3	low	0,235	0,227	0,248	0,313	0,046	0,293	0,556	0,068	0,163
TCGA-A2-A0EN	low	0,664	0,719	0,343	0,592	0,113	0,562	0,386	0,182	0,565
TCGA-A2-A0ES	low	0,661	0,549	0,550	0,453	0,200	0,357	0,525	0,102	0,573
TCGA-A2-A0EW	low	0,586	0,520	0,529	0,561	0,129	0,533	0,433	0,150	0,414
TCGA-A2-A0EX	low	0,346	0,462	0,362	0,534	0,148	0,516	0,531	0,089	0,290
TCGA-A2-A0T6	low	0,428	0,317	0,303	0,388	0,038	0,432	0,375	0,030	0,267
TCGA-A2-A0YI	low	0,402	0,392	0,382	0,500	0,106	0,382	0,322	0,038	0,347
TCGA-A2-A1G0	low	0,503	0,400	0,302	0,587	0,129	0,406	0,586	0,051	0,386
TCGA-A2-A1G6	low	0,078	0,132	0,095	0,106	0,022	0,114	0,130	0,013	0,072
TCGA-A2-A259	low	0,500	0,412	0,289	0,694	0,143	0,631	0,531	0,136	0,389
TCGA-A2-A4RY	low	0,289	0,364	0,369	0,527	0,129	0,506	0,408	0,053	0,229
TCGA-A2-A4S0	low	0,260	0,323	0,255	0,360	0,043	0,367	0,401	0,075	0,178
TCGA-A7-A0CH	low	0,776	0,742	0,598	0,733	0,183	0,657	0,653	0,080	0,571
TCGA-A7-A0D9	low	0,460	0,391	0,475	0,369	0,204	0,456	0,542	0,092	0,425
TCGA-A7-A13G	low	0,582	0,596	0,511	0,794	0,277	0,516	0,624	0,147	0,570
TCGA-A7-A13H	low	0,400	0,411	0,309	0,568	0,146	0,525	0,443	0,080	0,346

1,415	8,315	2,789	5,323	10,773	8,719	3,917	3,425	1,675	17,734	4,039
1,226	8,421	1,238	2,050	4,633	3,797	3,874	4,973	1,794	13,845	4,051
1,524	8,527	2,379	2,839	4,097	5,378	2,457	5,339	1,889	11,385	8,493
1,053	19,532	3,002	3,542	7,038	3,535	8,324	5,013	1,927	18,235	9,284
1,344	12,260	2,058	2,529	4,337	6,282	4,199	8,174	1,359	22,684	11,098
0,937	9,138	3,387	8,239	3,365	5,788	2,856	5,036	1,707	10,369	19,068
1,334	12,923	2,324	5,831	6,528	12,633	10,361	7,885	1,832	23,877	11,001
0,955	5,708	3,771	2,733	5,013	4,341	3,701	8,412	1,624	15,686	9,287
1,921	25,381	3,232	3,648	14,045	7,710	7,726	4,293	1,275	21,737	5,293
1,148	14,763	2,608	3,662	6,887	6,436	3,639	8,888	1,194	19,254	6,986
2,136	9,123	3,347	3,839	6,212	8,111	7,249	4,357	0,881	41,075	10,136
2,481	17,680	7,132	6,095	10,365	18,106	12,550	11,910	2,534	19,534	21,286
0,967	15,093	2,842	2,673	6,439	5,284	5,310	7,257	1,755	16,386	9,718
2,408	30,158	3,112	3,531	12,120	11,739	2,219	5,384	1,020	22,057	11,926
0,963	14,127	4,053	2,725	8,153	8,554	5,446	8,162	2,378	16,142	14,506
1,553	20,174	2,668	4,973	6,522	9,152	4,380	6,350	1,146	19,335	10,286
1,713	9,056	2,215	3,674	9,567	14,130	5,061	6,291	2,207	11,854	9,452
0,081	0,444	0,270	0,359	0,433	0,869	0,351	0,333	0,292	1,202	0,587
0,027	0,421	0,132	0,124	0,184	0,321	0,274	0,249	0,167	0,617	0,231
0,081	0,593	0,196	0,102	0,222	0,742	0,114	0,443	0,107	0,583	0,247
0,031	0,564	0,149	0,127	0,299	0,608	0,090	0,358	0,152	0,873	0,414
0,031	0,249	0,056	0,047	0,074	0,165	0,467	0,168	0,105	1,084	0,189
0,087	0,985	0,176	0,228	0,635	0,743	0,281	0,627	0,184	1,575	0,612
0,039	0,444	0,097	0,111	0,162	0,351	0,131	0,263	0,073	0,941	0,155
0,164	1,081	0,284	0,437	0,397	0,729	0,415	0,580	0,150	1,5	0,516
0,159	0,733	0,242	0,257	0,512	0,710	0,408	0,431	0,193	1,422	0,621
0,156	1,050	0,225	0,289	0,529	0,760	0,337	0,471	0,233	1,294	0,433
0,097	0,537	0,141	0,143	0,439	0,497	0,292	0,269	0,207	1,284	0,279
0,051	0,463	0,114	0,179	0,278	0,373	0,247	0,247	0,122	1,138	0,179
0,078	0,612	0,198	0,195	0,402	0,372	0,202	0,295	0,178	0,789	0,293
0,100	0,534	0,272	0,197	0,538	0,582	0,157	0,342	0,164	1,312	0,284
0,013	0,124	0,046	0,042	0,075	0,154	0,052	0,106	0,066	0,179	0,058
0,103	0,552	0,167	0,248	0,517	0,584	0,280	0,796	0,334	1,285	0,398
0,104	0,901	0,160	0,209	0,320	0,816	0,209	0,357	0,148	0,930	0,324
0,091	0,980	0,232	0,191	0,228	0,672	0,125	0,294	0,100	0,916	0,264
0,136	0,986	0,285	0,449	0,618	0,745	0,364	0,654	0,268	1,829	0,523
0,167	0,628	0,264	0,301	0,451	0,451	0,424	0,317	0,120	1,294	0,367
0,144	1,066	0,291	0,427	0,496	0,809	0,331	0,814	0,260	1,825	0,466
0,116	0,772	0,195	0,267	0,419	0,677	0,234	0,660	0,211	0,869	0,395



**Supplementary Table S3:** Continued

TCGA-A7-A5ZW	low	0,243	0,489	0,312	0,414	0,093	0,447	0,274	0,075	0,467
TCGA-A7-A5ZX	low	0,137	0,231	0,149	0,206	0,040	0,223	0,195	0,027	0,072
TCGA-A8-A06Y	low	0,390	0,513	0,354	0,407	0,097	0,539	0,393	0,108	0,425
TCGA-A8-A07G	low	0,812	0,647	0,675	0,711	0,231	0,838	0,656	0,141	0,555
TCGA-A8-A083	low	0,463	0,146	0,225	0,597	0,088	0,290	0,334	0,021	0,128
TCGA-A8-A08C	low	0,568	0,825	0,551	0,653	0,128	0,500	0,574	0,140	0,571
TCGA-A8-A0A2	low	0,495	0,646	0,466	0,527	0,143	0,491	0,484	0,110	0,385
TCGA-A8-A0A6	low	0,573	0,706	0,612	0,573	0,174	0,407	0,566	0,137	0,635
TCGA-AC-A2FG	low	0,279	0,358	0,325	0,369	0,227	0,479	0,478	0,050	0,325
TCGA-AC-A2FK	low	0,186	0,194	0,143	0,132	0,019	0,246	0,252	0,024	0,130
TCGA-AC-A2FO	low	0,588	0,625	0,489	0,620	0,186	0,518	0,717	0,089	0,419
TCGA-AC-A2QI	low	0,276	0,328	0,301	0,494	0,154	0,598	0,395	0,065	0,350
TCGA-AC-A3BB	low	0,509	0,464	0,713	0,640	0,175	0,790	0,484	0,158	0,453
TCGA-AC-A3QP	low	0,159	0,299	0,354	0,738	0,113	0,591	0,329	0,079	0,243
TCGA-AC-A3TN	low	0,187	0,355	0,348	0,768	0,306	0,720	0,396	0,088	0,258
TCGA-AC-A3YI	low	0,215	0,265	0,226	0,321	0,249	0,404	0,260	0,045	0,261
TCGA-AC-A3YJ	low	0,270	0,676	0,567	1,034	0,236	0,579	0,491	0,160	0,392
TCGA-AN-A0FS	low	0,291	0,570	0,522	0,708	0,258	0,664	0,604	0,098	0,440
TCGA-AN-A0XT	low	0,343	0,158	0,461	0,667	0,054	0,763	0,437	0,054	0,532
TCGA-AO-A0JG	low	0,462	0,518	0,356	0,312	0,086	0,324	0,501	0,067	0,302
TCGA-AO-A12C	low	0,688	0,658	0,632	0,687	0,083	0,542	0,470	0,145	0,409
TCGA-AO-A1KO	low	0,186	0,228	0,166	0,522	0,033	0,213	0,257	0,050	0,094
TCGA-AQ-A1H2	low	0,362	0,225	0,400	0,737	0,291	0,907	0,627	0,132	0,406
TCGA-AR-A0TR	low	0,611	0,291	0,408	0,652	0,078	0,408	0,580	0,077	0,270
TCGA-AR-A1AM	low	0,209	0,203	0,198	0,411	0,040	0,375	0,331	0,041	0,151
TCGA-AR-A24W	low	0,252	0,161	0,193	0,275	0,066	0,359	0,344	0,045	0,212
TCGA-AR-A2LM	low	0,447	0,369	0,284	0,395	0,152	0,476	0,363	0,074	0,312
TCGA-AR-A2LN	low	0,143	0,182	0,069	0,194	0,033	0,212	0,241	0,082	0,116
TCGA-AR-A2LQ	low	0,749	0,782	0,714	0,558	0,197	0,621	0,591	0,067	0,623
TCGA-B6-A0IH	low	0,278	0,114	0,191	0,205	0,050	0,553	0,465	0,008	0,089
TCGA-B6-A0RN	low	0,174	0,272	0,297	0,313	0,027	0,405	0,468	0,053	0,187
TCGA-B6-A0RP	low	0,407	0,647	0,419	0,601	0,094	0,422	0,487	0,135	0,503
TCGA-B6-A0RQ	low	0,101	0,200	0,148	0,199	0,048	0,256	0,180	0,043	0,125
TCGA-B6-A0X4	low	0,199	0,353	0,467	0,411	0,098	0,404	0,489	0,109	0,324
TCGA-B6-A0X7	low	0,393	0,566	0,393	0,714	0,098	0,456	0,600	0,143	0,406
TCGA-B6-A1KI	low	0,391	0,315	0,228	0,402	0,181	0,541	0,372	0,080	0,296
TCGA-BH-A0BM	low	0,144	0,364	0,191	0,440	0,035	0,218	0,463	0,054	0,268
TCGA-BH-A0BO	low	0,236	0,236	0,197	0,217	0,082	0,254	0,231	0,023	0,306
TCGA-BH-A0DO	low	0,328	0,287	0,250	0,287	0,120	0,352	0,334	0,038	0,183

0,094	0,385	0,161	0,183	0,280	0,562	0,216	0,308	0,177	0,824	0,288
0,054	0,258	0,077	0,078	0,135	0,319	0,067	0,157	0,086	0,379	0,188
0,144	1,449	0,181	0,292	0,392	0,514	0,536	0,727	0,171	1,917	0,744
0,161	1,181	0,333	0,300	0,600	0,754	0,261	0,632	0,310	2,093	0,529
0,041	0,431	0,168	0,118	0,180	0,404	0,199	0,266	0,121	1,447	0,266
0,097	0,990	0,294	0,318	0,528	1,169	0,384	0,800	0,200	1,787	0,476
0,163	1,161	0,126	0,387	0,450	0,790	0,259	0,485	0,313	1,641	0,485
0,143	0,846	0,234	0,278	0,452	0,591	0,315	0,409	0,282	1,562	0,467
0,099	0,739	0,161	0,231	0,369	0,548	0,196	0,361	0,246	0,820	0,258
0,030	0,287	0,085	0,075	0,113	0,241	0,058	0,111	0,096	0,376	0,105
0,138	0,718	0,219	0,212	0,576	0,639	0,314	0,436	0,124	1,460	0,655
0,097	0,320	0,243	0,182	0,391	0,623	0,182	0,489	0,149	0,783	0,246
0,093	1,000	0,178	0,233	0,511	0,845	0,323	0,601	0,301	1,418	0,754
0,038	0,542	0,107	0,128	0,337	0,581	0,150	0,508	0,157	0,662	0,443
0,044	0,682	0,159	0,208	0,394	0,865	0,217	0,380	0,271	0,872	0,645
0,057	0,523	0,184	0,143	0,230	0,488	0,260	0,247	0,214	0,661	0,323
0,153	1,066	0,334	0,336	0,540	0,869	0,250	0,793	0,187	1,267	0,891
0,143	1,030	0,242	0,324	0,487	0,947	0,249	0,437	0,357	1,487	0,425
0,083	1,014	0,157	0,194	0,335	0,960	0,249	0,510	0,261	1,737	0,523
0,100	0,981	0,149	0,252	0,389	0,690	0,307	0,355	0,171	1,147	0,245
0,078	1,234	0,204	0,355	0,687	0,672	0,260	0,781	0,186	1,610	0,647
0,028	0,290	0,088	0,089	0,105	0,104	0,072	0,100	0,079	0,294	0,082
0,131	1,004	0,311	0,251	0,658	0,272	0,276	0,540	0,320	1,875	0,747
0,077	0,483	0,129	0,300	0,462	0,942	0,165	0,562	0,077	1,371	0,362
0,049	0,432	0,089	0,122	0,298	0,572	0,141	0,376	0,133	0,649	0,256
0,141	0,773	0,092	0,116	0,287	0,359	0,082	0,247	0,104	1,410	0,190
0,086	0,350	0,160	0,173	0,342	0,461	0,288	0,291	0,196	1,066	0,257
0,020	0,263	0,065	0,069	0,196	0,275	0,043	0,137	0,069	0,353	0,125
0,175	1,515	0,348	0,548	0,693	0,829	0,571	0,900	0,271	1,775	0,537
0,026	0,307	0,082	0,085	0,165	0,125	0,123	0,109	0,070	1,042	0,123
0,056	0,430	0,105	0,146	0,336	0,463	0,149	0,237	0,182	0,950	0,271
0,084	0,963	0,229	0,302	0,463	0,755	0,249	0,720	0,206	1,432	0,522
0,025	0,247	0,045	0,132	0,160	0,284	0,047	0,233	0,089	0,389	0,141
0,076	0,901	0,261	0,232	0,321	0,571	0,170	0,429	0,116	1,194	0,382
0,125	0,729	0,232	0,351	0,522	0,827	0,225	0,558	0,138	1,680	0,574
0,075	0,944	0,252	0,218	0,353	0,605	0,323	0,350	0,264	1,351	0,477
0,066	0,375	0,152	0,132	0,171	0,206	0,078	0,078	0,082	0,638	0,304
0,080	0,558	0,128	0,090	0,206	0,467	0,193	0,127	0,151	0,622	0,159
0,064	0,550	0,144	0,147	0,238	0,411	0,168	0,277	0,182	0,626	0,233

**Supplementary Table S3:** Continued

TCGA-BH-A0DV	low	0,493	0,388	0,470	0,644	0,102	0,466	0,362	0,054	0,419
TCGA-BH-A0E7	low	0,750	0,503	0,486	0,707	0,146	0,676	0,599	0,148	0,652
TCGA-BH-A0EA	low	0,182	0,190	0,147	0,097	0,044	0,266	0,349	0,022	0,114
TCGA-BH-A0H3	low	0,562	0,493	0,349	0,434	0,154	0,669	0,602	0,073	0,538
TCGA-BH-A1ET	low	0,481	0,365	0,328	0,262	0,078	0,340	0,418	0,096	0,394
TCGA-BH-A1FB	low	0,649	0,677	0,542	0,619	0,179	0,609	0,514	0,141	0,559
TCGA-BH-A1FH	low	0,800	0,367	0,545	0,320	0,134	0,383	0,482	0,103	0,240
TCGA-BH-A28O	low	0,171	0,198	0,165	0,297	0,071	0,252	0,177	0,027	0,136
TCGA-BH-A28Q	low	0,547	0,374	0,370	0,553	0,104	0,341	0,447	0,070	0,301
TCGA-BH-A42U	low	0,174	0,217	0,270	0,383	0,092	0,439	0,191	0,047	0,151
TCGA-BH-A42V	low	0,238	0,400	0,274	0,549	0,150	0,429	0,300	0,093	0,275
TCGA-BH-A8G0	low	0,168	0,274	0,207	0,215	0,114	0,284	0,280	0,043	0,210
TCGA-C8-A12N	low	0,574	0,575	0,678	0,960	0,140	0,547	0,534	0,206	0,613
TCGA-D8-A1JB	low	0,638	0,493	0,380	0,385	0,217	0,540	0,458	0,086	0,402
TCGA-D8-A1JH	low	0,563	0,676	0,576	0,586	0,183	0,450	0,554	0,119	0,487
TCGA-D8-A1JN	low	0,216	0,348	0,413	0,356	0,099	0,372	0,520	0,046	0,215
TCGA-D8-A1JS	low	0,120	0,349	0,458	0,669	0,209	0,480	0,566	0,149	0,309
TCGA-D8-A1JU	low	0,099	0,183	0,101	0,166	0,025	0,178	0,219	0,036	0,113
TCGA-D8-A1XY	low	0,212	0,224	0,156	0,122	0,127	0,374	0,179	0,034	0,257
TCGA-D8-A27E	low	0,257	0,339	0,233	0,262	0,050	0,314	0,340	0,049	0,270
TCGA-D8-A27V	low	0,541	0,510	0,253	0,439	0,212	0,438	0,492	0,050	0,325
TCGA-D8-A3Z5	low	0,179	0,265	0,256	0,465	0,102	0,425	0,451	0,048	0,156
TCGA-D8-A4Z1	low	0,209	0,408	0,242	0,422	0,096	0,379	0,419	0,054	0,260
TCGA-D8-A73U	low	0,379	0,234	0,216	0,207	0,120	0,359	0,234	0,064	0,315
TCGA-D8-A73X	low	0,501	0,531	0,298	0,877	0,128	0,491	0,728	0,146	0,480
TCGA-E2-A14U	low	0,059	0,106	0,079	0,192	0,068	0,185	0,139	0,017	0,046
TCGA-E2-A15C	low	0,446	0,425	0,385	0,467	0,060	0,619	0,630	0,136	0,362
TCGA-E2-A15D	low	0,451	0,490	0,642	0,773	0,110	0,493	0,506	0,078	0,368
TCGA-E2-A15G	low	0,362	0,336	0,371	0,470	0,095	0,440	0,566	0,084	0,424
TCGA-E2-A1B4	low	0,170	0,180	0,245	0,538	0,046	0,183	0,404	0,075	0,187
TCGA-E2-A1BC	low	0,144	0,263	0,157	0,201	0,057	0,223	0,384	0,026	0,135
TCGA-E2-A3DX	low	0,483	0,511	0,358	0,502	0,249	0,535	0,490	0,094	0,409
TCGA-E9-A1R3	low	0,044	0,028	0,031	0,037	0,004	0,093	0,144	0,009	0,027
TCGA-E9-A3Q9	low	0,195	0,254	0,192	0,258	0,059	0,209	0,292	0,043	0,225
TCGA-E9-A5UP	low	0,164	0,342	0,193	0,209	0,116	0,333	0,210	0,071	0,273
TCGA-EW-A1J2	low	0,364	0,330	0,307	0,292	0,060	0,309	0,399	0,073	0,336
TCGA-EW-A1PG	low	0,185	0,187	0,130	0,250	0,025	0,185	0,219	0,036	0,140
TCGA-GI-A2C8	low	0,144	0,167	0,135	0,150	0,038	0,180	0,188	0,019	0,161

0,044	0,821	0,170	0,168	0,410	0,704	0,255	0,404	0,306	1,461	0,366
0,139	1,210	0,249	0,356	0,548	0,718	0,321	0,854	0,258	1,873	0,337
0,018	0,497	0,110	0,066	0,147	0,204	0,026	0,171	0,110	0,607	0,116
0,108	0,853	0,225	0,235	0,516	0,663	0,390	0,415	0,252	1,300	0,331
0,108	0,607	0,121	0,188	0,337	0,328	0,272	0,255	0,162	0,898	0,235
0,122	0,864	0,322	0,280	0,619	0,596	0,360	0,723	0,278	1,568	0,376
0,066	0,630	0,198	0,109	0,465	0,379	0,205	0,415	0,093	0,737	0,272
0,051	0,379	0,119	0,066	0,157	0,374	0,084	0,188	0,120	0,390	0,172
0,077	0,537	0,176	0,158	0,312	0,396	0,318	0,333	0,172	1,428	0,414
0,038	0,478	0,106	0,144	0,204	0,589	0,135	0,240	0,109	0,531	0,351
0,070	0,594	0,216	0,178	0,389	0,725	0,157	0,366	0,155	0,785	0,389
0,050	0,442	0,157	0,125	0,175	0,370	0,075	0,220	0,112	0,614	0,189
0,121	0,878	0,306	0,390	0,776	0,852	0,401	0,735	0,280	1,763	0,739
0,111	0,714	0,211	0,174	0,388	0,445	0,271	0,531	0,142	1,189	0,602
0,147	0,786	0,300	0,323	0,612	0,718	0,291	0,462	0,215	1,528	0,367
0,069	0,754	0,126	0,277	0,254	0,431	0,190	0,238	0,220	0,823	0,344
0,071	0,924	0,213	0,275	0,527	0,731	0,200	0,441	0,180	0,902	0,573
0,027	0,199	0,087	0,070	0,179	0,166	0,067	0,154	0,091	0,332	0,087
0,059	0,180	0,112	0,077	0,160	0,337	0,172	0,267	0,144	0,830	0,211
0,074	0,791	0,155	0,150	0,242	0,872	0,103	0,233	0,093	0,771	0,228
0,151	0,673	0,156	0,280	0,525	0,686	0,382	0,418	0,241	1,104	0,446
0,040	0,609	0,119	0,145	0,271	0,608	0,114	0,489	0,113	0,584	0,306
0,059	0,569	0,237	0,161	0,261	0,694	0,145	0,291	0,089	0,788	0,299
0,047	0,704	0,164	0,184	0,191	0,427	0,160	0,398	0,088	0,761	0,162
0,154	0,754	0,193	0,398	0,587	0,966	0,302	0,677	0,150	1,617	0,536
0,011	0,234	0,050	0,035	0,077	0,408	0,035	0,204	0,061	0,203	0,059
0,085	0,750	0,128	0,219	0,493	0,584	0,238	0,570	0,281	1,949	0,513
0,070	0,789	0,198	0,141	0,328	0,538	0,483	0,550	0,221	1,149	0,278
0,056	1,067	0,142	0,264	0,402	0,792	0,242	0,461	0,176	1,095	0,270
0,040	0,387	0,175	0,204	0,311	0,575	0,115	0,153	0,110	0,852	0,392
0,031	0,291	0,057	0,122	0,389	0,236	0,127	0,148	0,118	1,179	0,122
0,126	0,777	0,261	0,232	0,503	0,678	0,268	0,411	0,250	1,090	0,407
0,014	0,039	0,015	0,012	0,040	0,048	0,025	0,022	0,048	0,101	0,026
0,054	0,371	0,115	0,107	0,144	0,373	0,099	0,202	0,072	0,616	0,199
0,094	0,473	0,239	0,154	0,157	0,338	0,099	0,741	0,083	1,046	0,349
0,077	0,486	0,125	0,137	0,283	0,398	0,174	0,263	0,141	0,725	0,206
0,049	0,295	0,079	0,060	0,161	0,176	0,151	0,163	0,076	0,497	0,164
0,024	0,268	0,095	0,057	0,141	0,131	0,086	0,169	0,059	0,407	0,102

**Supplementary Table S3:** Continued

TCGA-GM-A2DM	low	0,313	0,387	0,389	0,441	0,127	0,398	0,370	0,101	0,364
TCGA-GM-A3XG	low	0,109	0,202	0,215	0,403	0,100	0,297	0,148	0,035	0,089
TCGA-GM-A4E0	low	0,128	0,316	0,209	0,239	0,049	0,337	0,118	0,123	0,239
TCGA-GM-A5PV	low	0,188	0,365	0,448	0,709	0,101	0,416	0,484	0,094	0,364
TCGA-GM-A5PX	low	0,129	0,153	0,082	0,111	0,050	0,231	0,142	0,038	0,099
TCGA-HN-A2OB	low	0,235	0,236	0,193	0,350	0,066	0,321	0,294	0,030	0,207
TCGA-OL-A5RU	low	0,279	0,416	0,361	0,569	0,201	0,433	0,393	0,057	0,302
TCGA-OL-A5RX	low	0,273	0,277	0,304	0,343	0,156	0,456	0,282	0,083	0,260
TCGA-OL-A66N	low	0,169	0,213	0,286	0,358	0,256	0,640	0,332	0,095	0,199
TCGA-PL-A8LX	low	0,214	0,224	0,358	0,331	0,109	0,459	0,266	0,051	0,209
TCGA-WT-AB44	low	0,041	0,217	0,031	0,125	0,017	0,130	0,204	0,019	0,077
TCGA-Z7-A8R5	low	0,093	0,324	0,310	0,459	0,120	0,409	0,244	0,052	0,219

0,062	0,748	0,191	0,259	0,345	0,509	0,229	0,370	0,200	1,039	0,377
0,039	0,276	0,080	0,075	0,155	0,456	0,066	0,206	0,117	0,416	0,203
0,067	0,568	0,175	0,148	0,182	0,485	0,158	0,241	0,059	0,567	0,372
0,091	0,479	0,196	0,209	0,440	0,746	0,221	0,310	0,313	0,996	0,701
0,041	0,218	0,043	0,088	0,090	0,129	0,049	0,157	0,082	0,316	0,100
0,064	0,402	0,197	0,108	0,226	0,462	0,124	0,214	0,099	0,605	0,232
0,102	0,903	0,230	0,176	0,369	0,805	0,228	0,492	0,136	0,908	0,326
0,084	0,602	0,169	0,115	0,321	0,576	0,266	0,305	0,214	0,865	0,271
0,072	0,697	0,139	0,143	0,338	0,427	0,139	0,477	0,165	0,848	0,395
0,060	1,092	0,115	0,193	0,202	0,715	0,164	0,311	0,305	1,291	0,481
0,005	0,198	0,019	0,084	0,038	0,070	0,038	0,154	0,043	0,154	0,161
0,044	0,370	0,120	0,136	0,159	0,469	0,106	0,341	0,136	0,433	0,372

**Supplementary Table S4**

<b>Oligos used for pLenti-CRISPR</b>	RSH16684	pLenti-CRISPR_gRNA_A3B_1-50_#1_top
	RSH16685	pLenti-CRISPR_gRNA_A3B_1-50_#1_bottom
	RSH16688	pLenti-CRISPR_gRNA_A3B_1-50_#3_top
	RSH16689	pLenti-CRISPR_gRNA_A3B_1-50_#3_bottom
	RSH16741	pLenti-CRISPR_gRNA_LacZ_top
	RSH16742	pLenti-CRISPR_gRNA_LacZ_bottom
	RSH10020	pLenti-CRISPR_gRNA_A3A/B_top
	RSH10021	pLenti-CRISPR_gRNA_A3A/B_bottom
<b>Synthetic gRNAs used</b>	gRNA_50UTR#1	Introducing DSB during precision base substitution
	gRNA_LacZ	Negative control during precision base substitution
<b>ssoDN</b>	RSH16912	HDR-template for precision genome engineering (* = PS)
<b>Genotyping primers for MCF10A hemizygote</b>	RSH3037	Validating hemizygote MCF10A cel line
	RSH3038	Validating hemizygote MCF10A cel line
	RSH3039	Validating hemizygote MCF10A cel line
	RSH3040	Validating hemizygote MCF10A cel line
	RSH14061	Screening for A3B promoter mutations
	RSH14062	Screening for A3B promoter mutations
<b>Cloning and sequencing</b>	RSH13836	Adding KpnI site to -900 +50 FWD
	RSH13831	Adding NheI site to -900 +50 REV
	RSH4467	Validation after SDM of pA3B-luciferase FWD
	RSH4468	Validation after SDM of pA3B-luciferase REV
	RSH2313	pJET1.2 cloning, sequencing FWD
	RSH2314	pJET1.2 cloning, sequencing REV
	RSH1360	Adding 5'EcoRI to mCherry-T2A-MCS
	RSH16769	Adding 3'BamHI to mCherry-T2A-MCS
	RSH14678	Adding N-terminal HA-tag to E2Fs FWD
	RSH14679	Adding N-terminal HA-tag to E2Fs REV
	RSH17006	Adding 5'NheI to HA-E2Fs, to place into pLenti4/TO-mCherry-T2A-MCS
	RSH12014	Adding 3'Agel to HA-E2Fs, to place into pLenti4/TO-mCherry-T2A-MCS
	RSH14902	Sequencing HA-E2F in pLenti4/TO-mCherry-T2A-MCS
	RSH15218	Sequencing HA-E2F in pLenti4/TO-mCherry-T2A-MCS
	RSH1263	Sequencing HA-E2F in pLenti4/TO-mCherry-T2A-MCS

caccgCAGAGCTTCAAAAAAGAGC gaacGCTCTTTTTTGAAGCTCTGc caccggTTCAAAAAAGAGCGGGACA gaacTGTCCTCCGCTCTTTTTTGAAcc caccgTGCGAATACGCCACGCGAT gaacATCGCGTGGGCGTATTGCGAc CACcagcaggggggtgaatggagc aaacGCTCCATTCAACCCCCCTGCT	
CAGAGCTTCAAAAAAGAGC CCCGAAUCUCUAUCGUGCGG	
G*CAGGAGGGGCCAGCAGGCAGAGTGGGCAG- CGGTACCTGATCTGTGGATTGATGTCAGCCTCT- TAGATACGCTTGTCCTGTGCCCTCTTTTTTT- GAAGCTCTGTGGTTTCACTTCCTGCTT*A	Design based on: Richardson, 2016, PMID: 26789497
TAGGTGCCACCCCGAT TTGAGCATAATCTTACTCTTGATC TTGGTGCTGCCCCCTC TAGAGACTGAGGCCCAT TTGCAATCCGTTGGGTCTT CCTGACTGGGATTCTGTGTGG	Kidd et al., 2007, PMID:17447845 Kidd et al., 2007, PMID:17447845 Kidd et al., 2007, PMID:17447845 Kidd et al., 2007, PMID:17447845
NNNNGGTACCCACGCATCCTCTCTCC NNNNGCTAGCGCCTCTTAGATACGCTTGTCC CTAGCAAATAGGCTGTCCC CTTTATGTTTTTGGCGTCTTCCA CGACTACTATAGGGAGAGCGGC AAGAACATCGATTTTCCATGGCAG NNNNGAATTCATGGTGAGCAAGGGCGAGGA NNNNGgatccAGGGCCGGGATTCTCTCTCC CTAGCGTTTACCACCATGGCCTACCCCTAC- GACGTGCCCGACTACGCCTCCCTCC TTAAGGAGGGAGGCGTAGTCGGGCACGTC- GTAGGGGTAGGCCATGGTGGTAAACG NNNNGCTAGCTACCCCTACGACGTGCCCGAC  AGGCCTACCGGTCCATAGAGCCACCGCATC  GGTGGGAGGTCTATATAAGCAGAGC ACCCTAACTGACACACATTCC CACCATCGTGAACAGTACGAAC	



**Supplementary Table S4:** Continued

	Primer name	Function
<b>qPCR primers</b>	RSH3220	A3B qPCR primer, FWD
	RSH3221	A3B qPCR primer, REV
	RSH3231	TBP qPCR primer, FWD
	RSH3232	TBP qPCR primer, REV
	RSH17341	A3B ChIP qPCR primer, FWD
	RSH17342	A3B ChIP qPCR primer, REV
	RSH17152	RAD51 ChIP qPCR primer, FWD
	RSH17153	RAD51 ChIP qPCR primer, REV
	RSH17154	TTK ChIP qPCR primer, FWD
	RSH17155	TTK ChIP qPCR primer, REV
	RSH2742	A3A qPCR primer, FWD
	RSH2743	A3A qPCR primer, REV
	RSH3085	A3C qPCR primer, FWD
	RSH3086	A3C qPCR primer, REV
	RSH2749	A3D qPCR primer, FWD
	RSH2750	A3D qPCR primer, REV
	RSH2751	A3F qPCR primer, FWD
	RSH2752	A3F qPCR primer, REV
	RSH2753	A3G qPCR primer, FWD
	RSH2754	A3G qPCR primer, REV
	RSH2757	A3H qPCR primer, FWD
	RSH2758	A3H qPCR primer, REV
<b>Oligos used for site-directed mutagenesis</b>	RSH14411	SDM A3B 5'UTR(-578) E2F SDM CC>TA (site A) FWD
	RSH14412	SDM A3B 5'UTR(-578) E2F SDM CC>TA (site A) REV
	RSH14100	SDM A3B 5'UTR(-460) E2F SDM CC>TA (site B) FWD
	RSH14101	SDM A3B 5'UTR(-460) E2F SDM CC>TA (site B) REV
	RSH14098	SDM A3B 5'UTR(-355) E2F SDM CC>TA (site C) FWD
	RSH14099	SDM A3B 5'UTR(-355) E2F SDM CC>TA (site C) REV
	RSH14096	SDM A3B 5'UTR(-180) E2F SDM CG>AT (site D) FWD
	RSH14097	SDM A3B 5'UTR(-180) E2F SDM CG>AT (site D) REV
	RSH14092	SDM A3B 5'UTR(+22) E2F SDM CG>AT (site E) FWD
	RSH14093	SDM A3B 5'UTR(+22) E2F SDM CG>AT (site E) REV
	RSH13821	SDM A3B 5'UTR(+1/+10) E2F SDM FWD
	RSH13822	SDM A3B 5'UTR(+1/+10) E2F SDM REV

Sequence (5'-3')	Reference, if applicable
GACCCCTTGGTCCTTCGAC	
GCACAGCCCCAGGAGAAG	
cccatgactcccatgacc	
tttacaaccaagattcactgtgg	
GGTCACTTTAAGGAGGGCTGT	Periyasamy et al. 2017. PMID: 28977491
TAGATACGCTGTGCCCTGTCC	Periyasamy et al. 2017. PMID: 28977491
GAGGCGGGGATACGTTAC	Müller, et al. 2017. PMID: 29228647
CTCCTTAGGGCTCGGTCTCT	Müller, et al. 2017. PMID: 29228647
TCACGACTGTCGCTTCTACG	Müller, et al. 2017. PMID: 29228647
GCTGGGGAAATGAAAAGATG	Müller, et al. 2017. PMID: 29228647
gagaagggacaagcacatgg	Refsland, et al. 2010, PMID: 20308164
tgatccatcaagtgtctgg	Refsland, et al. 2010, PMID: 20308164
AGCGCTTCAGAAAAGAGTGG	Refsland, et al. 2010, PMID: 20308164
AAGTTTCGTTCCGATCGTTG	Refsland, et al. 2010, PMID: 20308164
acccaaacgtcagtcgaatc	Refsland, et al. 2010, PMID: 20308164
cacatttctgctggttctc	Refsland, et al. 2010, PMID: 20308164
ccgtttggacgcaaagat	Refsland, et al. 2010, PMID: 20308164
ccagtgatctggaacactt	Refsland, et al. 2010, PMID: 20308164
ccgaggaccgaaggttac	Refsland, et al. 2010, PMID: 20308164
tccaacagtgtgaaattcg	Refsland, et al. 2010, PMID: 20308164
agctgtggccagaagcac	Refsland, et al. 2010, PMID: 20308164
cggaatgttcggctgtt	Refsland, et al. 2010, PMID: 20308164
GGCTGACCTGCTACCCACCTTCTGG	
CCAGAAGGTGGGTAGCAGGTCAGCC	
GAGTCCACACTTTATTCCCCACTGG	
CCAGTGGGGAATAAAGTGTGGACTC	
GAGAGGGGGCTTAAGCCTGCCCAAC	
GTTGGGCAGGCTTAAGCCCCCTCTC	
GTTCTCTGCCAGATGGAAGGGTCCGG	
CCGGACCCCTTCATCTGGCAGAGGAAC	
CTTCAAAAAAAGAGATGGACAGGGACAAG	
CTTGTCCTGTCCATCTCTTTTTTTGAAG	
GCAGGAAGTGAAACCTCTGCGA-	
TACAAAAAAGAGCGGG	
CCCGCTCTTTTTTGTATCGCA-	
GAGGTTTCACTTCTCTGC	

**Supplementary Table S4:** Continued

	Primer name	Function
<b>Oligos used for site-directed mutagenesis</b>	RSH13823	SDM A3B 5'UTR(+11/+20) E2F SDM FWD
	RSH13824	SDM A3B 5'UTR(+11/+20) E2F SDM REV
	RSH13825	SDM A3B 5'UTR(+21/+30) E2F SDM FWD
	RSH13826	SDM A3B 5'UTR(+21/+30) E2F SDM REV
	RSH13827	SDM A3B 5'UTR(+31/+40) E2F SDM FWD
	RSH13828	SDM A3B 5'UTR(+31/+40) E2F SDM REV
	RSH13829	SDM A3B 5'UTR(+41/+50) E2F SDM FWD
	RSH13830	SDM A3B 5'UTR(+41/+50) E2F SDM REV
	RSH16981	SDM A3B 5'UTR(+20) E2F SDM A>T FWD
	RSH16982	SDM A3B 5'UTR(+20) E2F SDM A>T REV
	RSH16801	SDM A3B 5'UTR(+21) E2F SDM G>C FWD
	RSH16802	SDM A3B 5'UTR(+21) E2F SDM G>C REV
	RSH16613	SDM A3B 5'UTR(+22) E2F SDM C>G FWD
	RSH16614	SDM A3B 5'UTR(+22) E2F SDM C>G REV
	RSH16805	SDM A3B 5'UTR(+23) E2F SDM G>C FWD
	RSH16806	SDM A3B 5'UTR(+23) E2F SDM G>C REV
	RSH16795	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>TCTAAA
	RSH16796	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>TCTAAA
	RSH16987	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>AACAAA
	RSH16988	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>AACAAA
	RSH16989	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>AACATT
	RSH16990	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>AACATT
	RSH16993	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>TTCATT
	RSH16994	SDM A3B 5'UTR (CHR) E2F SDM TTCAAA>TTCATT

Sequence (5'-3')	Reference, if applicable
GAAACCACAGAGCTTCTACATAGGT- GCGGGACAGGGACAAG	
CTTGTCCTGTCCGCACCTATG- TAGAAGCTCTGTGGTTTC	
GCTTCAAAAAAGAGAGAGGCG- GCGACAAGCGTATCTAAG	
CTTAGATACGCTTGTGCGCG- CCTCTCTCTTTTTTGAAGC	
GAGCGGGACAGGGTCTAACCTGTCTAAGAGGCGC	
GCGCCTCTTAGACAGTTAGACCTGTCCGCTC	
GGACAAGCGTATGTGACACGAGCTAGCCCGGC	
GCCCGGGCTAGCTCGTGTACATACGCTTGTC	
GAAACCACAGAGCTTCAAAAAAGT- GCGGGACAGGGACAAG	
CTTGTCCTGTCCGCACCTTTTTT- GAAGCTCTGTGGTTTC	
CACAGAGCTTCAAAAAAGAC- CGGGACAGGGACAAG	
CTTGTCCTGTCCGGTCTTTTTTTGAAGCTCTGTG	
CAGAGCTTCAAAAAAGAGGG- GGACAGGGACAAGCG	
CGCTTGTCCTGTCCCCCTCTTTTTTTGAAGCTCTG	
CAGAGCTTCAAAAAAGAGC- CGGACAGGGACAAGCG	
CGCTTGTCCTGTCCGGCTCTTTTTTTGAAGCTCTG	
GAAGTGAAACCACAGAGCTCTA- AAAAAGAGCGGGACAG	
CTGTCCCGCTCTTTTTTTA- GAGCTCTGTGGTTTCACTTC	
GTAAGCAGGAAGTGAAACCACAGAGCAA- CAAAAAAGAGCGGGACAGGGAC	
GTCCCTGTCCGCTCTTTTTTTGTGCTCT- GTGGTTTCACTTCCTGCTTAC	
GTAAGCAGGAAGTGAAACCACAGAGCAA- CATTAAAGAGCGGGACAGGGAC	
GTCCCTGTCCGCTCTTTTAATGTTGCTCT- GTGGTTTCACTTCCTGCTTAC	
GCAGGAAGTGAAACCACAGAGCTTCAT- TAAAAGAGCGGGACAGGGAC	
GTCCCTGTCCGCTCTTTTAAT- GAAGCTCTGTGGTTTCACTTCCTGC	

**Supplementary Table S4:** Continued

	Primer name	Function
Oligos used for SILAC	APOBEC3B promoter wt forward	WT bait for SILAC experiments
	APOBEC3B promoter wt reverse	WT bait for SILAC experiments
	APOBEC3B promoter CHR mut (SDM2) forward	CHR mutant bait for SILAC experiments
	APOBEC3B promoter CHR mut (SMD2) reverse	CHR mutant bait for SILAC experiments
	APOBEC3B promoter E2F mut (SDM3) forward	E2F mutant bait for SILAC experiments
	APOBEC3B promoter E2F mut (SDM3) reverse	E2F mutant bait for SILAC experiments
	APOBEC3B promoter C22G forward	22CtoG mutant bait for SILAC experiments
	APOBEC3B promoter C22G reverse	22CtoG mutant bait for SILAC experiments

Roelofs et al. 2020. All values relative to TBP

Sequence (5'-3')	Reference, if applicable
GTGAAACCACAGAGCTTCAAAAAAAGAGC- GGGACAGGGACAAGCGTATCTAAGAGGC	
ACGCCTCTTAGATACGCTTGTCCTGTC- CCGCTCTTTTTTTGAAGCTCTGTGGTTTC	
GTGAAACCACAGAGCTTCTACATAGGTGCG- GGACAGGGACAAGCGTATCTAAGAGGC	
ACGCCTCTTAGATACGCTTGTCCTGTCCCG- CACCTATGTAGAAGCTCTGTGGTTTC	
GTGAAACCACAGAGCTTCAAAAAAAGAGA- GAGGCGGGACAAGCGTATCTAAGAGGC	
ACGCCTCTTAGATACGCTTGTCGCCG- CCTCTCTCTTTTTTTGAAGCTCTGTGGTTTC	
GTGAAACCACAGAGCTTCAAAAAAAGAGG- GGGACAGGGACAAGCGTATCTAAGAGGC	
ACGCCTCTTAGATACGCTTGTCCTGTC- CCCCTCTTTTTTTGAAGCTCTGTGGTTTC	

A dedication to my parents,  
for delightful memories have inspired  
me and everyone I know.

For the things we have done together  
and for the things we are going to do together.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

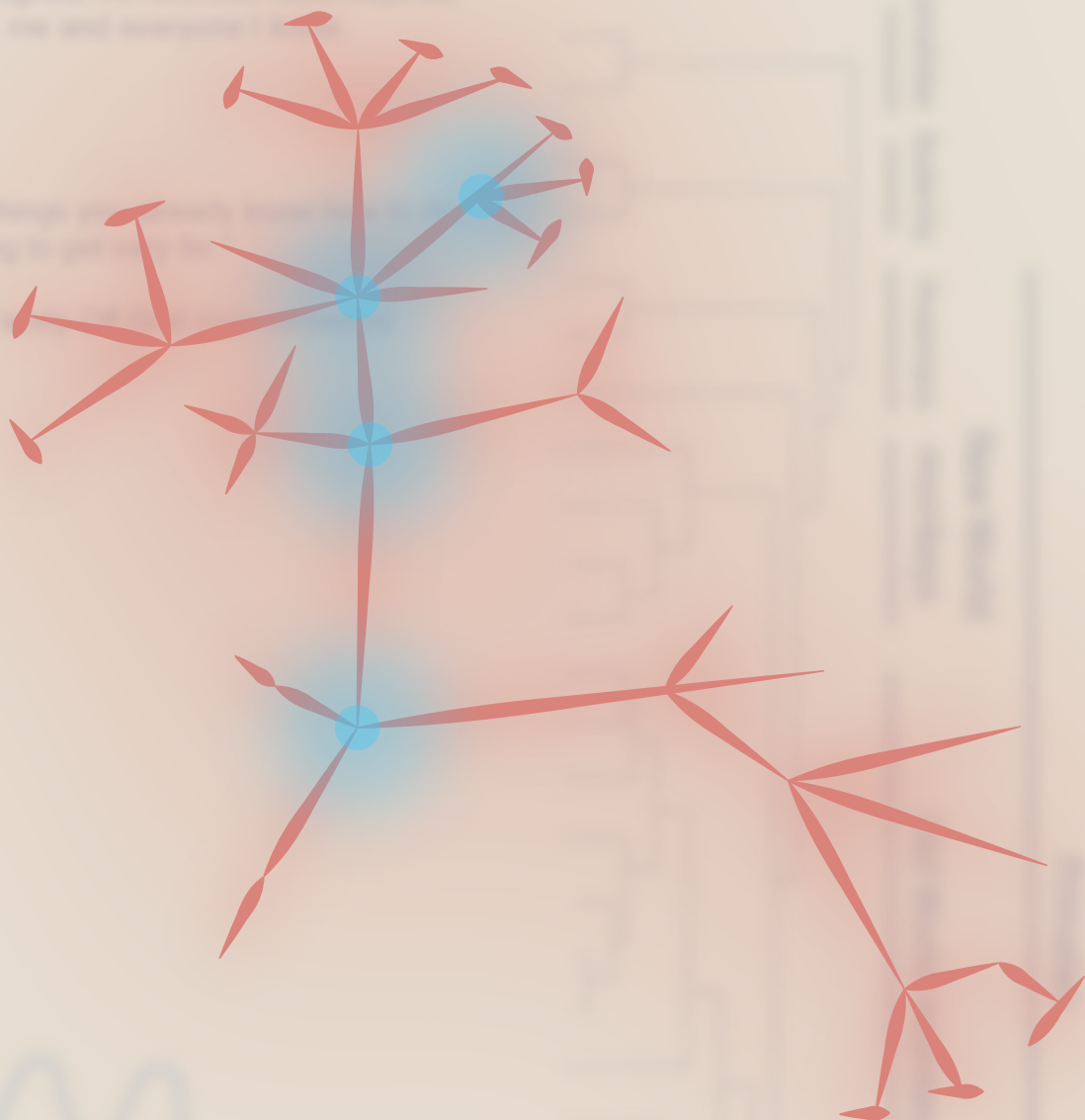
With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.

With love and affection,  
I dedicate this book to you.



## Aberrant APOBEC3B Expression in Breast Cancer Is Linked to Proliferation and Cell Cycle Phase

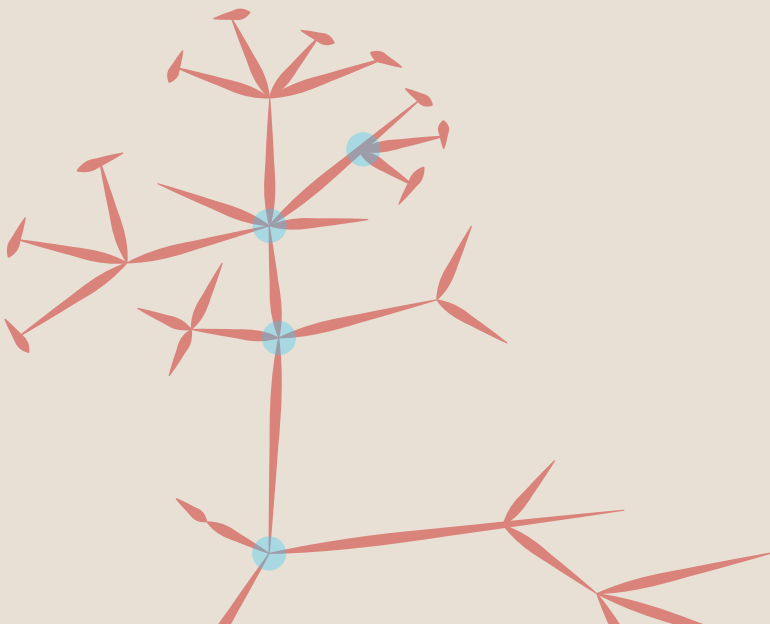
Pieter A Roelofs<sup>1,2</sup>, Mieke A. M. Timmermans<sup>3</sup>, Bojana Stefanovska<sup>2,4,5</sup>, Myrthe A. den Boestert<sup>3</sup>, Amber W. M. van den Borne<sup>1</sup>, Hayri E. Balcioglu<sup>3</sup>, Anita M. Trapman<sup>3</sup>, Reuben S. Harris<sup>2,4,5</sup>, John W. M. Martens<sup>3</sup> and Paul N. Span<sup>1,\*</sup>

1. Radiotherapy & Oncolimmunology Laboratory, Department of Radiation Oncology, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands
2. Department of Biochemistry, Molecular Biology and Biophysics, Masonic Cancer Center, Institute for Molecular Virology, and Center for Genome Engineering, University of Minnesota, Minneapolis, MN 55455, USA
3. Department of Medical Oncology, Erasmus MC Cancer Institute, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands
4. Department of Biochemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, TX 78229, USA
5. Howard Hughes Medical Institute, University of Texas Health San Antonio, San Antonio, TX 78229, USA



## Abstract

APOBEC3B (A3B) is aberrantly overexpressed in a subset of breast cancers, where it associates with advanced disease, poor prognosis, and treatment resistance, yet the causes of A3B dysregulation in breast cancer remain unclear. Here, A3B mRNA and protein expression levels were quantified in different cell lines and breast tumors and related to cell cycle markers using RT-qPCR and multiplex immunofluorescence imaging. The inducibility of A3B expression during the cell cycle was additionally addressed after cell cycle synchronization with multiple methods. First, we found that A3B protein levels within cell lines and tumors are heterogeneous and associate strongly with the proliferation marker Cyclin B1 characteristic of the  $G_2/M$  phase of the cell cycle. Second, in multiple breast cancer cell lines with high A3B, expression levels were observed to oscillate throughout the cell cycle and again associate with Cyclin B1. Third, induction of A3B expression is potently repressed throughout  $G_0$ /early  $G_1$ , likely by RB/E2F pathway effector proteins. Fourth, in cells with low A3B, induction of A3B through the PKC/ $\kappa$ B pathway occurs predominantly in actively proliferating cells and is largely absent in cells arrested in  $G_0$ . Altogether, these results support a model in which dysregulated A3B overexpression in breast cancer is the cumulative result of proliferation-associated relief from repression with concomitant pathway activation during the  $G_2/M$  phase of the cell cycle.



## Introduction

Normal and cancerous mammary epithelial cells are subject to a multitude of endogenous and exogenous sources of mutagenesis, scarring their genome with a barrage of genetic alterations [1–3]. Well-known sources that inflict these mutations are reactive oxygen species, inadequate repair of DNA lesions due to repair deficiencies, and carcinogenic compounds present in tobacco smoke [3]. Additionally, the mutational landscape of certain cancer subtypes, including breast cancer, is drastically shaped by members of the APOBEC family of deaminases, some of which catalyze the deamination of cytosines in genomic single-stranded DNA substrates [4–8]. Mutagenesis mediated by these DNA cytosine deaminases, referred to as APOBEC mutagenesis, actively contributes to inter- and intra-tumoral heterogeneity and increases in strength during the progression of breast cancer [9–11], as reviewed in [12]. In fact, together with spontaneous deamination of 5-methylcytosines and homologous recombination deficiency, APOBEC mutagenesis is one of the main sources of genomic scarring in breast cancer genomes and is a major mechanism of endogenous mutations [3,8,9,11]. For instance, mutations in known drivers of breast cancer such as *PIK3CA*, *KMT2C*, *ARID1A*, *NF1*, and *CDH1* have been attributed to APOBEC enzymatic activity and are enriched in metastatic breast cancer, supporting a model in which APOBEC proteins are promoting tumor evolution [9,11]. The activity of this family of enzymes also contributes to poor outcomes in response to breast cancer therapies [13,14] and is reviewed in [12].

Humans express up to seven different APOBEC3 (A3) enzymes, namely A3A, A3B, A3C, A3D, A3F, A3G, and A3H. To date, the most likely family members to contribute to APOBEC mutagenesis are A3A and A3B, with abundant evidence supporting a role for A3B in breast cancer [5,13,15–18], as reviewed in [19]. A3B, which is expressed at a low level in normal mammary tissue, is overexpressed in a subset of cancerous mammary tissues [5,15,20], as reviewed in [19]. In addition, high A3B expression in breast cancer is associated with aggressive clinicopathological characteristics [15–18] resulting in increased metastatic potential, as well as worse prognosis in treatment-naïve disease [13,16,17]. Therefore, considering the significance of A3B in breast cancer, determining what drives this protein to abnormally high expression levels is critically important.

Earlier work by our group and others has pointed out that both A3B mRNA and protein levels associate with proliferation markers in multiple tumor types [21–28]. Moreover, negative regulation of A3B transcription is facilitated through two

distinct repressive E2F complexes, at least one of which associates with the cell cycle [24,25,29]. However, it remains unknown whether aberrant A3B expression in breast cancer is caused by derepression and/or enhanced induction. We therefore set out to investigate the proliferation dependency of A3B expression and induction throughout the cell cycle of breast cancer cells. The findings presented here support a model in which A3B expression is universally repressed during cellular dormancy (i.e., G<sub>0</sub> and G<sub>1</sub> phases), whereas pathway activation during the proliferative stage (G<sub>2</sub> and M phases) contributes to the overexpression of A3B found in a subset of breast cancer cell lines and tumors. Together, these results provide a molecular explanation for the association of A3B expression and cancer cell proliferation and provide further insights into the transcriptional (dys)regulation of this cancer genomic DNA deaminase.

## Materials and Methods

### Cell lines and culture conditions

All cell lines used in this study were cultured at 37 °C under 5% CO<sub>2</sub>. MCF10A cells and the derived HDR and control clones were previously described and cultured using an identical medium recipe [24]. MCF7 cells were cultured in DMEM (Gibco) supplemented with 10% FCS, Glutamax (Thermo Fisher Scientific, Waltham, MA, USA), 6 µg/L recombinant human insulin (Sigma Aldrich, St. Louis, MO, USA), penicillin (100 U/mL), and streptomycin (100 µg/mL). HCC1599, HCC1954, SUM225CWN, and HCC1143 cells were cultured in RPMI supplemented with 10% FCS, penicillin (100 U/mL), and streptomycin (100 µg/mL). SKBR3, MDA-MB-415, BT474, and MDA-MB-468 cells were cultured in DMEM supplemented with 10% FCS, Glutamax, penicillin (100 U/mL), and streptomycin (100 µg/mL). All cell lines routinely tested negative for Mycoplasma using the MycoAlert Mycoplasma Detection Kit (Lonza, Basel, Switzerland). All cell counting took place using automated cell counters (Countess, Thermofisher and Luna, Logos Biosystems). PMA was always used at 25 ng/mL for 6 h. Knockdown of A3B was performed using shRNA technology, as described in [5]. An overview of the most relevant cell lines and their characteristics relating to this study is provided in Supplementary Table S1.

### Tissue Collection and (Multiplex) IHC

Cell lines used in conventional and multiplex IHC assays were part of an in-house cell line microarray (CMA), which contains 54 formalin-fixed paraffin-embedded (FFPE) breast cancer cell lines. Primary breast cancer specimens were obtained under the protocol to study biological markers associated with disease outcome

and in accordance with the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands (<https://www.coreon.org/wp-content/uploads/2020/04/coreon-code-of-conduct-english.pdf>, accessed on 14 April 2023) and the new European General Data Protection Regulation (GDPR). This protocol states that the use of coded left-over material for scientific purposes and, therefore, for the greater good, does not require informed consent according to Dutch law. This waiver of informed consent was acknowledged by the medical ethics committee of the Erasmus Medical Centre Rotterdam, the Netherlands, in MEC 02.953.

Detection of A3B by conventional IHC on these cell lines was performed as described in [26–28]. For multiplex IHC, slides containing cell lines or primary breast cancer specimens were rehydrated by three xylene submersions of 10 min, followed by three 100% ethanol washes, one 70% ethanol wash, one 50% ethanol wash, and a final wash in ddH<sub>2</sub>O. Tissue material was then fixed to the glass slides by a 10-minute incubation in neutral buffered formalin, followed by a wash in ddH<sub>2</sub>O. For antigen retrieval, a 15-minute heating cycle at 95 °C in 1X AR6 buffer (Akoya Biosciences, AR600250) was performed. Subsequently, multiplex staining was performed for A3B, Cyclin B1, Cyclin E2, and pan-Cytokeratin. In short, non-specific epitopes were blocked in blocking solution (Akoya Biosciences Marlborough, MA, USA, ARD 1001EA) for 10 min at RT, and rabbit anti-A3B (5210-87-13 [26]) was added at a 1:640 dilution in blocking solution. After a combined incubation of 1 h at RT and 4 °C O/N, slides were washed thrice in PBS-T, followed by another 10-minute blocking step. Opal HRP Polymer Ms + Rb (Akoya Biosciences, ARH1001EA) was added for 10 min, followed by PBS-T wash. Opal 650 (Akoya Biosciences, FP1496001KT) was added at a 1:200 dilution in amplification diluent (Akoya Biosciences, FP1498) and incubated for 10 min at RT, followed by three PBS-T washes. This procedure was then repeated using rabbit anti-Cyclin B1, 1:25 (Santa Cruz, Dallas, TX, USA, clone GNS1, sc-247, 1 h RT + 4 °C O/N), using the Opal 520 fluorophore, 1:200 (Akoya Biosciences, FP1487001KT); mouse anti-Cyclin E2, 1:100 (Santa Cruz, sc-245, 1 h RT), using the Opal 690 fluorophore, 1:200 (Akoya Biosciences, FP1497001KT); and mouse anti-pan-Cytokeratin (pan CK), 1:400 (Novus Biologicals, Englewood, CO, USA, clone AE-1/AE-3, 30 min RT), using the Opal 620 fluorophore, 1:200 (Akoya Biosciences, FP1495001KT). The slides were mounted after DAPI counterstaining.

Spectral library mapping was performed using single-stained sample slides of BT474 cells using the Vectra 3.0 system. This library was used for spectral overlap compensation and to determine the set exposure time per filter channel of the Vectra microscope. Greyscale images of samples were taken using inform advanced image analysis software package (PerkinElmer) and exported to a TME-Analyzer

version 2.4 beta (Balcioglu et al., manuscript in preparation), which was utilized to segment cells based on cytokeratin (i.e., negative, borderline, or positive) staining and appearance (i.e., grouped, individual, or spindle-shaped cells). A Python code was written (available upon request) to process data provided by the TME-Analyzer and determine co-localization at a cellular level. To determine the likelihood of the found co-localization occurring naturally, binomial distribution was performed using the following Formula (1):

$$P_{bin} = \binom{n}{k} p^k (1 - p)^{n-k}$$

where:

n = the total cells per cell line;

k = the number of double positives (i.e., a cell staining positive for A3B and Cyclin B1/Cyclin E2);

p = the chance of a double-positive cell based on the occurrence of the individual Cyclin per cell line, as calculated using the following Formula (2):

$$\left( \frac{\#A3B}{\#Total} \cdot \frac{\#Cyclin}{\#Total} \right)$$

The null hypothesis of the binomial test stated that the number of double-positive cells (e.g., A3B and Cyclin B1) was caused by chance, and the alternative hypothesis stated that co-localization was caused by a yet-to-be-determined underlying biological mechanism. The chosen p-value of 0.05 was corrected with a Bonferroni correction. The null hypothesis was rejected when p-values were lower than 0.05.

## Cell Cycle Synchronization

MCF10A cells were synchronized in  $G_0$  by a combined method involving contact inhibition and growth factor withdrawal. In short, cells were plated in T75 or T175 flasks (according to experiment size), and growth medium was refreshed after 2 days. Once cells had reached full confluence, growth medium was once again replaced and contact inhibition was initiated for 24 h. Cells were then harvested, washed in HBSS twice, counted, and split at a 1:5 ratio into separate T25 or T75 flasks (according to experiment size) into mitogen-free medium (MCF10A medium without horse serum and EGF). After 24 h of growth inhibition, mitogen-free medium of all flasks except  $t = 0$  was replaced with normal MCF10A growth medium. For each time point, cells were harvested from individual flasks. Double-thymidine block experiments were performed by plating cells in individual T25 flasks and adding thymidine at a concentration of 2.0 mM after 24 h (48 h for HCC1143). After 18 h of incubation, all flasks were washed with pre-warmed,  $CO_2$ -equilibrated HBSS and allowed to progress in pre-warmed,  $CO_2$ -equilibrated culture

medium. After 8 h, thymidine was added again at a concentration of 2.0 mM and was removed by washing twice with pre-warmed, CO<sub>2</sub>-equilibrated HBSS after 18 h. Thymidine was not removed for the t = 0 time point, which was harvested directly. Of note, due to excessive cytotoxicity, MCF10A cells could not be reproducibly synchronized using double-thymidine. The double-thymidine block was also unsuitable for the synchronization of BT474 cells, which is likely due to their low proliferation rate [30]. Serum withdrawal experiments with MCF7 and MDA-MB-468 cells were performed by culturing cells for 24 h in normal growth medium, followed by washing twice with PBS, and further culturing in serum-deprived medium for up to 24 h. When applicable, medium was removed and replaced with normal growth medium for further sampling. For experiments involving palbociclib, MCF10A and BT474 were plated in separate flasks or plates for 24 h, followed by treatment with 1 μM palbociclib for up to 24 h. Palbociclib was then washed out by washing cells thrice with PBS, after which periodic sampling commenced. For all experiments, cells were harvested using Trypsin-EDTA (ThermoFisher), neutralized in an equal volume of Trypsin Neutralization Solution (Genlantis, San Diego, CA, USA), counted, and divided up in separate tubes according to the number of analyses. All centrifugation steps with live cells were performed at 450 x g for 10 min, or 800 x g for 15 min for final sample collection.

### PI Staining, Flow Cytometry, and Fluorescence Microscopy

Single-cell suspensions were fixed in ice-cold 70% ethanol while vortexing and incubated on ice for at least 30 min. Cells were pelleted (450 x g, 10 min), and ethanol was removed. Cells were washed twice in PBS before staining for 45 min with staining solution (PBS supplemented with 0.1% (w/v) Triton X-100, 20 μg/mL propidium iodide, and 200 μg/mL RNase A) at 1–2 × 10<sup>6</sup> cells per mL at 37 °C. Cells were then analyzed immediately using the PE channel of a FACS Canto analyzer (BD Biosciences, Franklin Lakes, NJ, USA). To observe mitosis, live cells were stained with 0.3 g/mL Hoechst in PBS for 5 min at room temperature, enclosed, and analyzed using conventional fluorescence microscopy. Fluorescent immunocytochemistry was performed using rabbit anti-53BP1, 1:1000 (ThermoFisher, PA1-16565).

### Immunoblotting and mRNA Quantification

Immunoblotting and RT-qPCR were performed as described before [24]. The antibodies used for immunoblotting were rabbit anti-A3B, 1:1000 (5210-87-13 [26]); rabbit anti-Cyclin E2, 1:1000 (Cell Signaling, Danvers, MA, USA, #4132); rabbit anti-Cyclin B1, 1:1000 (Cell Signaling, D5C10); mouse anti-Cyclin A2, 1:2000 (Cell Signaling, BF683); rabbit anti-FoxM1, 1:1000 (Cell Signaling, D12D5); and mouse anti-Tubulin, 1:20,000 (Sigma-Aldrich, T5168). Primers used for RT-

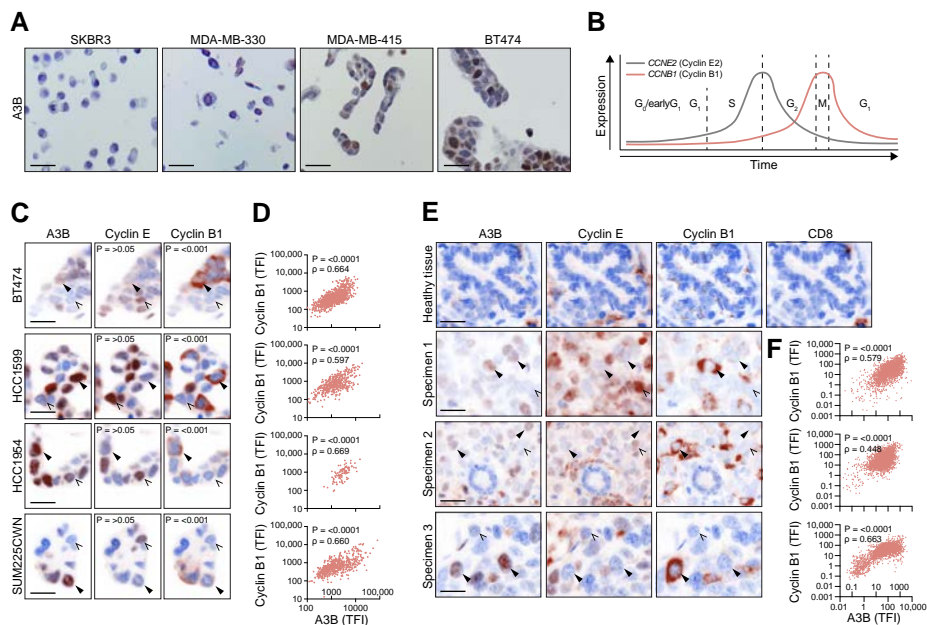
qPCR were as follows: *A3A*: FWD 5'-GAGAAGGGACAAGCACATGG-3' and REV 5'-TGGATCCATCAAGTGTCTGG-3'; *A3B*: FWD 5'-GCCACAGAGAAGATTCTTAGCC-3' and REV 5'-CGCCAGACCTACTTGTGCTA-3'; *CCNE2*: FWD 5'-TCCAAGAGTTTGCTTACGTCAC-3' and REV 5'-GCCAGGAGATGATTGTTACAGG-3'; *CCNB1*: FWD 5'-TCTTGACAGTAAATGATGTGG-3' and REV 5'-CAGTCAATTAGGATGGCTCT-3'; *B2M*: FWD 5'-CTTTGTCCAGCCCAAGATAG-3' and REV 5'-CAATCCAAATGCGGCATCTTC-3'; *HPRT*: FWD 5'-TGACACTGGCAAAACAATGCA-3' and REV 5'-GGTCCTTTTACCAGCAAGCT-3'; *TBP*: FWD 5'-CCCATGACTCCCATGACC-3' and REV 5'-TTTACAACCAAGATTCACTGTGG-3'.

## Results

### A3B Is Co-Expressed with Cell Cycle Proteins

To obtain a global overview of the *A3B* expression pattern in breast cancer cells, paraffin-embedded sections of an exploratory panel of breast cancer cell lines were subjected to immunohistochemistry (IHC) using a validated antibody against *A3B* [26–28]. While this antibody recognizes *A3A*, *A3B*, and *A3G*, detection of *A3B* is easily distinguished from *A3A* and *A3G* during microscopy and immunoblotting due to *A3B*'s distinct nuclear localization and size, respectively ([26–28]; discussed in [12]). We initially chose to investigate MDAMB-415 and BT474, two cell lines known to express *A3B* at relatively high levels [5,31]. Two cell lines, *A3B*-null SKBR3 and *A3B*-low MDA-MB-330 [5,31–33], served as independent negative controls. While undetectable in these negative controls, *A3B* protein was detectable in the remaining cell lines and almost exclusively localized to the nuclear compartment, which is typical of this enzyme (Figure 1A) [4,34]. Intriguingly, the overall staining pattern in both *A3B*-high cell lines was far from uniform, with approximately 50% of cells almost completely lacking *A3B* signal.

Considering the major influence cell cycle progression exerts on gene expression genome-wide and the association of *A3B* with markers of proliferation in breast cancer (see Section 1), these initial observations prompted us to further consider the cell cycle as a major determinant of *A3B* expression. Thus, to differentiate between various phases of the cell cycle, a multiplex IHC assay was developed, which included the immunomicroscopic detection of *A3B*, Cyclin E2, Cyclin B1, and pan-Cytokeratin (pCK) (Supplementary Figure S1, also see Section 2). Of note, for all analyses, subcellular localization was taken into account to ensure additional specificity; thus, nuclear staining of *A3B* and Cyclin E2 and cytoplasmatic staining of Cyclin B1 were quantified. Classically, Cyclin E2 expression peaks during the late G<sub>1</sub>/S phase and drops dramatically once G<sub>2</sub> starts. Conversely, Cyclin B1 peaks during



**Figure 1.** Protein expression of A3B in breast cancer is heterogeneous and correlates with cell cycle proteins. **(A)** Photomicrographs of A3B using conventional IHC on a diverse set of breast cancer cell lines, including the A3B-null (SKBR3), A3B-low (MDA-MB-330), and A3B-high (MDA-MB-415 and BT474) cell lines. Note the heterogeneous expression pattern of A3B in MDA-MB-415 and BT474. **(B)** General overview of the various cell cycle phases, overlaid with proteins known to peak during each respective phase.  $G_0/G_1$  is marked by low expression of both cyclins, while *CCNE2* (Cyclin E2) expression rises throughout  $G_1$  and into  $G_1/S$ . Expression of *CCNB1* (Cyclin B1) increases during  $G_2$  and peaks at mitosis. **(C)** Photomicrographs showing the subcellular localization of A3B, Cyclin E2, and Cyclin B1 in various A3B-high breast cancer cell lines. Images acquired by multiplex IHC staining. Solid arrowheads indicate cells expressing A3B; open arrows point out cells negative for A3B. P-values are calculated by binomial testing, determining the probability of A3B being co-expressed with individual cell cycle markers. **(D)** Quantification of (subcellular specific) total fluorescent intensity (TFI) of individual cells as in **(C)**, using Spearman correlation. **(E,F)** Healthy tissue and three primary breast cancer tumors stained for A3B, Cyclin E2, and Cyclin B1, using the same workflow and analysis presented in **(C,D)**. CD8 staining of healthy tissue is included to account for the sporadic cytoplasmic staining encountered in isolated cells, attributable to A3G expressing tumor-resident leukocytes. Note that detection of A3B protein through regular and multiplex IHC requires separate optimization and antibody concentrations, accounting for the slight difference in staining pattern between these two procedures. Scale bars denote 20  $\mu\text{m}$ . TFI: total fluorescent intensity.

$G_2$  to only enter the nuclear compartment during the relatively short mitotic phase (Figure 1B) [35,36]. Thus, these two cyclins are distinctly regulated and can be used as cell cycle markers. The multiplex assay was applied on asynchronous cultures (i.e., grown under normal tissue culture conditions) of various A3B-expressing breast cancer cell lines (BT474, HCC1599, HCC1954, and SUM225CWN; see [5,25]).



As indicated above, Cyclin E2 is readily detectable in the nuclear compartment, and in separate cells, Cyclin B1 predominantly localizes to the cytoplasm. Furthermore, in agreement with our expectations, nuclear A3B is detectable in all expressing cell lines, but not in all individual cells (Figure 1C). Interestingly, individual breast cancer cells that lack A3B are often positive for Cyclin E2.

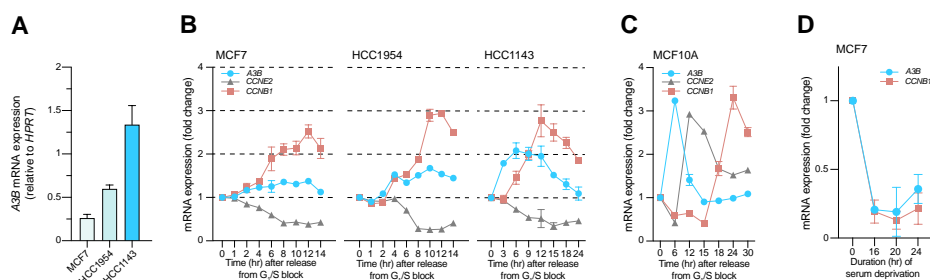
Binomial testing, determining the probability of A3B being co-expressed with either Cyclin, was applied on single-cell images and further determined that A3B is significantly co-expressed ( $p < 0.001$ ) with Cyclin B1, but not Cyclin E2 ( $p > 0.05$ ; see Figure 1C). Additionally, nonparametric Spearman rank correlation revealed a statistically significant positive correlation between A3B and Cyclin B1 in all four cell lines ( $p = 0.664, 0.597, 0.669, 0.660$ , respectively; Figure 1D). Most cell lines also showed cells negative for both Cyclin E2 and Cyclin B1 (and A3B), which might represent quiescent cells [37]. Subsequently, specimens of treatment-naïve primary invasive ductal carcinomas were analyzed. In concordance with the observations in cell lines, most A3B-positive cells scored negative for Cyclin E2 and positive for cytoplasmic Cyclin B1 (Figure 1E,F). Combined, these results show that the expression of endogenous A3B in cancer cells is coordinated with proliferation and predominantly with the G<sub>2</sub> and M phases of the cell cycle.

### **Endogenous A3B Expression Can Fluctuate throughout the Cell Cycle**

Next, multiple cell synchronization methods were used to investigate the degree to which A3B expression changes throughout the cell cycle. Given the consistent observations amongst a diverse set of sample types thus far (*i.e.*, multiple independent cell lines and tumor tissues), we opted to further diversify the cell line panel. Two additional breast cancer cell lines, MCF7 and HCC1143, as well as the previously featured HCC1954 cell line, were therefore included in cell cycle synchronization experiments. These cell lines represent the range of A3B expression found in clinical samples, with MCF7 expressing low levels, HCC1954 expressing moderate levels, and HCC1143 expressing high levels of A3B (Figure 2A; also see [5]). Additionally, these three cell lines were receptive to the same cell cycle synchronization method (Supplementary Table S1). Each cell line was synchronized with a double-thymidine block (see Section 2), and periodic samples were taken for RT-qPCR analysis upon release from the block. To complement the aforementioned multiplex IHC assay, gene expression of *CCNE2* (Cyclin E2) and *CCNB1* (Cyclin B1) was included to monitor cell cycle progression.

Interestingly, after release from the double-thymidine block, A3B mRNA levels rise in all three cell lines, in a manner that associates positively with *CCNB1* expression

and inversely with *CCNE2* expression (Figure 2B). However, it is notable that these trends are modest with the *A3B*-low cell line, MCF7A, intermediate in the *A3B*-medium cell line, HCC1954, and strongest with the *A3B*-high cell line HCC1143. In all instances, before the cultures reach peak *CCNB1* expression levels, a discernable increase in *A3B* expression is evident that is approximately 2-fold higher than expression levels at the  $G_1/S$  border. Notably, for HCC1143, where protein levels are sufficiently high to be tracked by immunoblot, both *A3B* protein and mRNA levels closely associate and often peak slightly sooner in comparison to those of *CCNB1* throughout the entire cell cycle (Supplementary Figure S2A).



**Figure 2.** *A3B* expression in relation to cell cycle markers *CCNE2* and *CCNB1*. **(A)** Levels of endogenous *A3B* mRNA expression during normal culture of *A3B*-low (MCF7), *A3B*-intermediate (HCC1954), and *A3B*-high (HCC1143) breast cancer cells ( $n = 3$ ). **(B)** Expression of *A3B*, *CCNE2*, and *CCNB1* in MCF7, HCC1954, and HCC1143 released from  $G_1/S$  arrest, induced by a double-thymidine block. The peak in *CCNB1* expression denotes the  $G_2/M$  phase, whereas *CCNE2* drops soon after release since *CCNE2* expression is high throughout  $G_1$ . The different sampling times are chosen to accommodate the differences in the growth speed of the used cell lines. Error bar represents SD. MCF7 and HCC1954 data represent  $n = 2$  independent experiments; HCC1143 data represent  $n = 6$  independent experiments. **(C)** Cell cycle progression and *A3B* expression in the normal-like breast epithelial cell line MCF10A, which expresses very low levels of *A3B*. Data represent  $n = 2$  independent experiments. **(D)** Gene expression levels of *A3B* and *CCNB1* as measured in MCF7 breast cancer cells after varying lengths of serum deprivation. Values relative to normal growth conditions. Error bar represents SD. Data represent  $n = 2$  independent experiments.

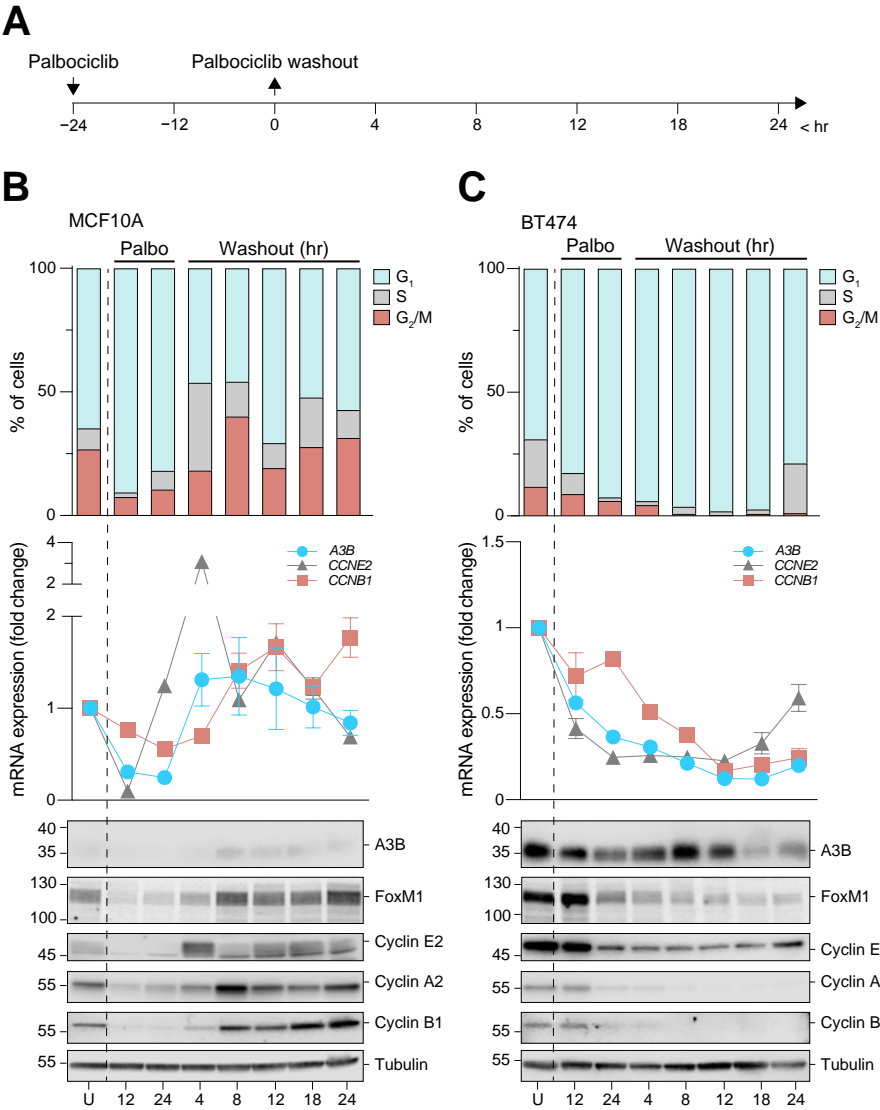
In additional experiments, the normal-like mammary epithelial cell line MCF10A, which has very low baseline levels of endogenous *A3B* expression [38], was synchronized using contact inhibition and serum starvation (see Section 2). When released in a normal medium, MCF10A cells consistently lag in  $G_0$ /early  $G_1$  for 6-9 h and reach the  $G_1/S$  phase in 12 h, S phase in 18 h, and  $G_2/M$  in 24 h, as judged by *CCNE2* and *CCNB1* expression (Supplementary Figure S2B). Other than an initial expression peak upon release in a normal medium (which likely is due to an effect of serum addition including many growth and signaling factors), the MCF10A cell line shows no appreciable changes in *A3B* during the bulk of its cell cycle (Figure 2C).

Protein levels of A3B in these experiments tended to fluctuate, likely due to low signal-to-noise ratio and minor sampling differences between samples. These findings indicate that A3B expression in cancer cells may be coordinated with the cell cycle but is unable to do so in the normal-like cell line MCF10A, where expression of A3B is typically very low.

To obtain a view of A3B expression in the dormant  $G_0$  state, proliferation was halted in MCF7 cells by serum withdrawal, which is classically known to induce proliferative arrest [39,40]. Compared to the double-thymidine block, A3B expression is even more influenced by serum withdrawal, showing an ~80% decrease after 16 h (Figure 2D). Comparable observations were made in the cell line MDA-MB-468 which, after serum starvation, shows dramatically decreased levels of A3B as compared to published expression values [5]. These values steadily increase when MDA-MB-468 cells are released back into the normal medium (Supplementary Figure S2C). Thus far, the observations indicate that A3B expression in breast cancer cells requires proliferation and associates strongly with Cyclin B (*CCNB1* expression), indicative of the  $G_2/M$  phase of the cell cycle. Additionally, and as defined previously [24], a repressive mechanism exists in normal-like MCF10A cells to prevent A3B from being expressed, even during active cell cycling and proliferation (discussed further below).

### **A3B Expression in Cancer Cell Lines Is Strongly Cell Cycle-Dependent**

RNA sequencing expression profiles of A3B were associated previously with cell proliferation [22]. Additionally, earlier work showed that A3B transcription is repressed by the RB/E2F pathway [24,25], which itself is heavily influenced by cell cycle stage [41]. However, these earlier studies did not establish a mechanistic link between the cell cycle and A3B expression. To investigate whether the RB/E2F pathway controls A3B expression, MCF10A and BT474 cells were subjected to prolonged treatment with the CDK4/6 inhibitor palbociclib. Inhibition of the RB/E2F pathway through inhibiting CDK4/6 for 24 h has been shown to induce growth arrest in early  $G_1$  (Supplementary Figure S3A; also see [37]). The BT474 cell line was chosen for its considerable expression of A3B, which actively shapes the genome of this cell line *in vitro* [31,42,43], and MCF10A cells were included as a normal like control. Moreover, in our hands, palbociclib treatment is the only suitable synchronization method for these two cell lines (see Supplementary Table S1), allowing for the investigation of cell cycle dynamics in an RB/E2F context. In our experiments, each prolonged palbociclib treatment was followed by washout, release into normal medium, and periodic sampling for mRNA and protein analysis (Figure 3A).



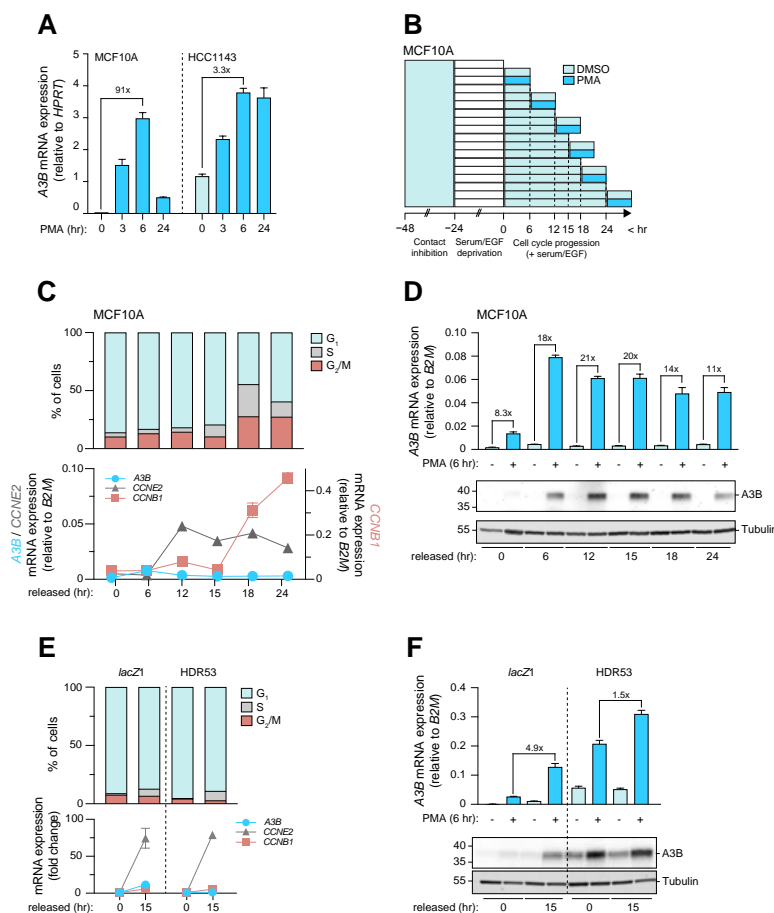
**Figure 3.** A3B is repressed during G<sub>0</sub>/early G<sub>1</sub> arrest in normal-like MCF10A cells and A3B-high BT474 cells. **(A)** Palbociclib arrests MCF10A and BT474 cells in G<sub>0</sub>/early G<sub>1</sub> by direct interference with the RB/E2F pathway. Cells were treated for 24 h to induce growth arrest, followed by washout and subsequent sampling. **(B,C)** Cell cycle progression (top panels) and RT-qPCR and immunoblot analysis (center and bottom panels, respectively) of A3B, Cyclin E2, and Cyclin B1 during and after palbociclib-induced growth arrest of MCF10A and B474 cells. For immunoblots, samples of both cell lines were loaded on the same gel, but data are depicted separately here due to image intensity adjustments required for visualization of A3B in each cell line. U = unsynchronized. Error bar represents SD. Data represent n = 4 independent experiments.

Synchronization in early  $G_1$  was successful in both cell lines, as measured by PI staining and flow cytometry (Figure 3B,C, upper panels). MCF10A cells swiftly resume proliferation after drug washout. Conversely, BT474 cells remain arrested for up to 18 h, during which near-complete early  $G_1$  arrest is obtained before cells begin to trickle into the S phase. Immunoblot analysis of cell cycle proteins, including Cyclin E2 and Cyclin B1, corresponded with the appropriate cell cycle phases (Figure 3B,C, lower panels). In MCF10A cells, A3B expression levels are relatively low, and protein levels drop below detection limits exclusively during growth arrest. However, following palbociclib washout, A3B expression levels increase and appear to track most closely with those of Cyclin B1 (Figure 3B, center panel). In comparison, A3B mRNA expression in BT474 is readily detectable and steadily declines during palbociclib treatment and subsequent early  $G_1$  arrest (Figure 3C, center panel). Importantly, although the decrease in A3B protein levels trails behind mRNA expression levels, both measures become significantly lower once a near complete  $G_1$  arrest is obtained (Figure 3C). Given the observed prolonged growth arrest in BT474 even after palbociclib washout, an additional time course was set up that involved an extended outgrowth period. This experiment shows comparable results and additionally that A3B expression can revert to normal (i.e., high) levels at 72 h after palbociclib washout (Supplementary Figure S3B). Notably, in both cell lines, expression levels of A3B closely mirror those of the  $G_2/M$  regulator FOXM1, which, like A3B, is repressed through the RB/E2F axis [24,44].

### A3B Induction Dynamics throughout the Cell Cycle

Several mechanisms that activate A3B transcription depend on the activation of the inflammatory PKC/ $\kappa$ B pathway [38,45–49]. This pathway, in turn, is activated by various intra- and extracellular stimuli, such as DNA damage from ionizing radiation (IR), chemical interference with pathway components, and/or ligand-based receptor activation (reviewed in [12,50]; also see [38]). It is therefore mechanistically insightful to know in which cell cycle phase the PKC/ $\kappa$ B pathway is most capable of inducing A3B. For instance, because A3B preferentially targets ssDNA [51,52], which is more abundant in actively dividing cells as compared to growth arrested cells, induction during proliferation may be particularly mutagenic.

In order to efficiently investigate the cell cycle dynamics of A3B expression, we first set out to select cell lines with differential responses to induction. For this purpose, the chemical compound phorbol myristic acid (PMA) was used to induce PKC/ $\kappa$ B-dependent expression of A3B in MCF10A and HCC1143. MCF10A cells show high A3B inducibility, whereas HCC1143 breast cancer cells, which have much higher baseline levels of endogenous A3B, show appreciably lower A3B



**Figure 4.** Induction of A3B expression is strongest in actively proliferating cells, and repression during G<sub>0</sub> is RB/E2F-dependent. **(A)** Analysis used to determine the most suitable model for the investigation of PKC/ncNF- $\kappa$ B-induced expression of A3B within the context of cell cycle progression. Unsynchronized MCF10A and HCC1143 cells were exposed to PMA for up to 24 h, during which samples were taken for RT-qPCR analysis. Error bar represents SD. Data represent  $n = 2$  independent experiments. **(B)** Method used to induce A3B expression using PMA at defined cell cycle stages of MCF10A cells. Teal denotes contact inhibition, followed by growth factor withdrawal (white), cell cycle progression (light blue), and treatment with PMA or vehicle. Samples taken at the start of each PMA or vehicle treatment served as independent controls for cell cycle progression. **(C)** Cell cycle progression of MCF10A as analyzed by PI staining followed by flow cytometry, and RT-qPCR analysis of A3B, *CCNE2*, and *CCNB1*. Error bar represents SD. Data represent  $n = 2$  independent experiments. **(D)** A3B mRNA (top) and protein (bottom) expression upon PMA treatment during indicated time points following synchronization. Error bar represents SD. Data represent  $n = 2$  independent experiments. **(E)** Synchronization of *lacZ* control and MCF10A-HDR53 cells, as analyzed by PI stain and RT-qPCR of A3B, *CCNE2*, and *CCNB1*. Gene expression is given as fold change over baseline relative to the cell cycle housekeeping gene *B2M*. Error bar represents SD. Data represent  $n = 3$  independent experiments. **(F)** Induction of A3B upon PMA treatment in G<sub>0</sub>-arrested and actively cycling *lacZ* control and MCF10A-HDR53 cell lines. Fold changes above the brackets indicate the post-induction levels of A3B in arrested cells with those in actively cycling cells. Error bar represents SD. Data represent  $n = 3$  independent experiments.

induction potential (Figure 4A, Supplementary Table S1). We propose that the PKC/ $\kappa$ B pathway and the *A3B* promoter may already be partially activated in HCC1143 cells. Indeed, earlier work has shown that the PKC/ $\kappa$ B pathway in many cancer cell lines is constitutively activated during culture and is actively driving *A3B* expression [38]. MCF10A was synchronized and cells in various cell cycle stages were treated with PMA (Figure 4B). Cell cycle progression was analyzed by flow cytometry and RT-qPCR, and kinetics comparable to earlier experiments were obtained (compare Figure 4C to Supplementary Figure S2B). As a control, the expression of Cyclin E2 was analyzed since PKC/ $\kappa$ B activation is known to repress *CCNE2* [53] (Supplementary Figure S4A). Importantly, *A3B* was induced minimally in  $G_0$ -arrested cells, indicating that potent repression is present in this cell cycle phase even when PKC/ $\kappa$ B pathway activation is strong (Figure 4D).

Interestingly, *A3B* was induced to higher levels in cells released from  $G_0$  arrest (Figure 4D). An independent time course yielded comparable results (Supplementary Figure S4B,D). To control for the possibility that the relative lack of *A3B* inducibility during  $G_0$  is due to a starvation-induced decrease in viability and signal responsiveness, *A3A* expression was analyzed in parallel. Like *A3B*, *A3A* has been found to be inducible by PMA and therefore served as a potent control for induction [45]. Importantly, the immediate induction of *A3A* by PMA in  $G_0$ -arrested cells indicates that the relative lack of *A3B* induction throughout  $G_0$  is not an artifact of our synchronization method (Supplementary Figure S4D).

We also analyzed whether the strong induction of *A3B* during proliferation could lead to DNA damage, especially considering the association of *A3B* with DNA strand breaks and the activation of the DNA damage response [5]. We therefore treated growth-arrested and actively cycling cells with PMA and analyzed the nuclear accumulation of the DNA damage marker 53BP1 by immunofluorescence microscopy. Contrary to our expectations, treatment with PMA did not lead to the induction of 53BP1 foci under any condition tested (Supplementary Figure S5). An explanation for this observation is discussed below.

So far, the results show that *A3B* transcriptional activation through the PKC/ $\kappa$ B pathway is influenced by cell cycle progression and that induction is impaired in  $G_0$ . Repression of cell cycle-associated genes is often facilitated by E2F-based transcriptional complexes, including the DREAM complex, which predominantly represses genes throughout the dormant  $G_0$  phase (reviewed in [41]). We have previously shown that E2F-based complexes also repress *A3B* transcription in normal mammary epithelial cells by binding to an *A3B* resident E2F site [24].

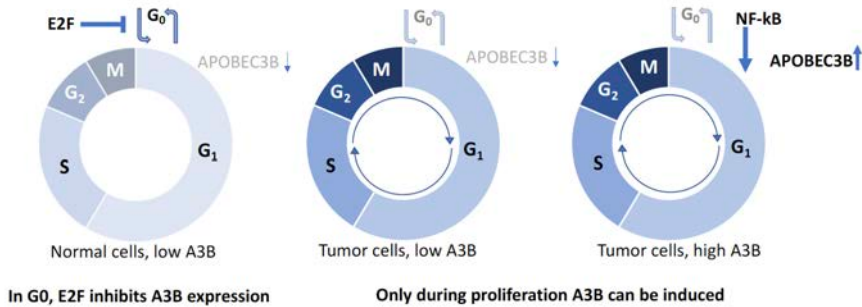
To investigate whether this site facilitates repression of A3B specifically during  $G_0$ , MCF10A cells engineered to lack the E2F site through CRISPR/Cas9-mediated base substitution (HDR clones, published in [24]) were synchronized and treated with PMA. The cell line MCF10A-HDR53 was used; similar to other HDR clones, it expresses increased levels of A3B as compared to controls (Supplementary Figure S4E). Both control and HDR cell lines arrested effectively at  $G_0$  and reached the  $G_1/S$  phase in 15 h, as indicated by a sharp increase in *CCNE2* (Figure 4E). Intriguingly, while the A3B induction pattern in the control cell line corresponds to earlier results (i.e., low induction in  $G_0$  and high induction after, compare Figure 4D,F), induction of A3B in  $G_1/S$ -phase cells is only marginally higher than that in  $G_0$  in the MCF10A-HDR53 cell line. Thus, the transcriptional activation of A3B by the inflammatory PKC/ $\kappa$ NF- $\kappa$ B pathway is most efficient when cells are actively proliferating and is actively repressed by E2F repressive complexes when cells are arrested in  $G_0$  and/or early  $G_1$ . Additionally, a comparison of data from the 0 and 15 h timepoints suggests that A3B expression levels may remain stable throughout the cell cycle in cells lacking this single E2F site (Figure 4F, light blue bars of HDR53). Combined with the findings mentioned above, these data show that A3B is universally repressed by the RB/E2F pathway throughout  $G_0$  in normal and cancer cells alike. The gain-of-function experiments in our engineered MCF10A cells also suggest that there may be no other repressive mechanisms during all other cell cycle phases and that aberrant pathway activation during proliferation likely explains A3B overexpression observed in a significant proportion of breast cancers.

## Discussion

This study constitutes the first detailed insights into the cell cycle dynamics of A3B expression in breast cancer. Starting with the observation that A3B expression is heterogeneous in cultured breast cancer cells, we show that A3B expression associates strongly with Cyclin B1 and not Cyclin E2 on a single-cell level in cell lines and clinical breast cancer specimens. Furthermore, breast cancer cell line synchronization studies indicate that A3B expression levels oscillate throughout the cell cycle in a manner that associates positively with Cyclin B1 ( $G_2/M$  phase) and negatively with Cyclin E ( $G_0/G_1$  phase). Surprisingly, expression and induction of A3B during growth arrest are tightly repressed by the RB/E2F pathway in all cell lines tested. Conversely, A3B induction is a prominent feature during cell cycle progression of A3B-overexpressing breast cancer cells, and only possible during proliferation in cells with low constitutive expression. Thus, our study offers new insights into the mechanism of A3B regulation during the cell cycle. Building on



our previous work [24,25], we propose that a general lack of RB/E2F-mediated repression does not fully explain the observed dysregulation of A3B expression in breast cancers. Instead, the combined effects of RB/E2F de-repression and PKC/ $\kappa$ B pathway activation appear to be required to achieve the high A3B expression levels observed in many cancer cell lines and tumors (model in Figure 5). This model is also supported by work from other groups, which indicated that A3B expression is associated with proliferation [21–23,54].



**Figure 5.** Model explaining current results. In normal, non-proliferating cells in G<sub>0</sub>, E2F inhibits A3B gene expression (left). During proliferation, such as in rapidly dividing tumor cells, E2F no longer inhibits A3B expression (middle), yet for A3B to be overexpressed in a subset of tumors, stimulation during proliferation is required through the NF- $\kappa$ B pathway (right).

Over the recent years, various publications have reported on the transcriptional regulation of A3B. Since its first association, the PKC/ $\kappa$ B pathway has continued to take a central role in A3B regulation [38,47,55]. The PKC/ $\kappa$ B pathway conveys intracellular and extracellular signals such as DNA damage and TNF $\alpha$ , respectively, and thus relays cellular stress and inflammatory signals to the A3B promoter [38,47,49]. As seen in this study and others, the PKC/ $\kappa$ B pathway and, by extension, the expression of A3B can also be stimulated directly using PMA. Conversely, transcriptional *repression* is facilitated by the RB/E2F pathway and is enforced by the binding of E2F-based repressive complexes to an E2F site within the A3B promoter [24,25]. Disruption of this pathway raises A3B expression in a way that mirrors the action of viral oncogenes E7 and polyomavirus large T antigen, which are well known for their ability to induce A3B during viral infections [25,29]. The data presented in this work dovetail with this prior knowledge and indicate that when these regulatory mechanisms are no longer in homeostasis, a “perfect storm” scenario might unfold. In this scenario, the combination of DNA damage, inflammation, and a generally pro-proliferative cellular state raise A3B protein levels to a degree where they can actively influence disease trajectory.

Three findings in particular were unexpected in this study. First, whereas in the current study a 24 h exposure to palbociclib yielded a potent downregulation of A3B, earlier work showed that inhibition of CDK4/6 using palbociclib does not directly influence A3B expression [25]. At the time, it was proposed that a mechanism secondary to the RB/E2F pathway might also regulate A3B expression. We now show that a decrease in A3B expression coincides with arrest in G<sub>0</sub>/early G<sub>1</sub>, strongly implicating proliferative status with the expression dynamics of A3B. In follow-up research, it would be informative to include a transcription factor screen of the A3B promoter during palbociclib treatment and during promoter activation in response to PMA, which would further elucidate the transcriptional dynamics of A3B. Secondly, repression of A3B during G<sub>0</sub> appears to be largely intact in all cell lines included in this study, including cells with high expression of A3B. We previously proposed that overexpression of A3B, in part, may be attributable to the failure of some breast cancer cells to repress A3B through the RB/E2F pathway. Although this is still supported by multiple lines of evidence, the current study offers more detailed insights and indicates that cancer cells with high A3B expression reside disproportionately in cell cycle phases where endogenous and exogenous stimuli are free to activate the promoter. Thus, we propose that overexpression of A3B is the culmination of untimed proliferation and activation of (inflammatory) pathways, including the PKC/ $\kappa$ B pathway (Figure 5). Finally, it was expected that increased expression of A3B in response to PMA would mount a DNA damage response. The absence of this response may be explained by the possibility that PMA interferes directly with the cell's ability to mount a DNA damage response. As also seen in our studies, acute administration of PMA severely slows cell cycle progression [39,53]. Thus, by reducing ssDNA present at replication forks, cytosine deamination and the associated DNA damage response are avoided [45]. Alternatively, the levels of A3B reached with PMA may not be high enough to inflict serious DNA damage in our model. In either scenario, we propose the use of other techniques, such as CRISPR activation or lentiviral transduction at high MOI, to investigate the effects of cell cycle-specific APOBEC activity on the DNA damage response. Combining these efforts will provide valuable insights into how A3B-stimulating pathways and the cell cycle of tumor cells combine to induce A3B, providing the mutagenic fuel that drives tumor evolution.

The clinical relevance of this work is at least two-fold. First, it explains the heterogeneity observed in tumor A3B protein expression by IHC, with strong positivity associating with active cell cycling and specifically with the G<sub>2</sub>/M phase of the cell cycle (coincident with Cyclin B1 positivity). Such tumors are not only actively dividing but are also therefore likely to be undergoing active DNA deamination, damage, and mutagenesis caused by A3B (*i.e.*, ongoing evolution). This finding is corroborated by other work that shows that A3B is preferentially expressed during G<sub>2</sub>/M in myeloid cells [23]. Since

APOBEC can synergize with DNA damage-inducing drugs across multiple cancer types, including breast and ovarian cancer [56,57], it may be beneficial to combine such treatments with G<sub>2</sub>-arresting drugs. Inducing G<sub>2</sub> arrest in A3B-positive tumors to be treated with IR may also provide clinical benefit since tumor cells are most sensitive to DNA damage in G<sub>2</sub>, during which A3B expression may be induced further to possibly toxic levels. Future cell-based assays should determine the dosage, timing, and feasibility of such an approach. Second, this work shows the applicability of multiplex IHC with an A3B antibody and two cell cycle proteins. In essence, any other clinically relevant marker may be co-stained using this approach, which may then inform the tumor microenvironment of A3B-expressing tumors. For example, the infiltration of CD4<sup>+</sup>/CD8<sup>+</sup> T-cells into A3B-expressing tumor mass could be directly observed in clinical biopsies. This is exceptionally relevant since recent evidence indicates that APOBEC-dominated tumors might be receptive to immune-based anticancer therapies [58–60].

In conclusion, the data presented here consolidate the proposed association between cellular proliferation and A3B expression and offer further insights into the cell cycle dynamics of the regulation of this cancer genomic DNA deaminase. In breast cancers, A3B overexpression appears to be the cumulative result of proliferation-associated relief from repression with concomitant pathway activation (Figure 5).

### Supplementary Materials

The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cells12081185/s1>, Table S1: Relevant characteristics of the main cell lines used in this study. Figure S1: Additional information regarding multiplex IHC; Figure S2: Detailed time-courses and MDA-MB-468 serum deprivation; Figure S3: Additional cell cycle synchronization: 28 courses using palbociclib; Figure S4: Additional cell cycle synchronization: 38 courses using palbociclib; Figure S5: PMA induction of A3B and DNA damage response.

### Author Contributions

Conceptualization and methodology: P.A.R., R.S.H., J.W.M.M. and P.N.S. Data collection and experimentation: P.A.R., A.W.M.v.d.B., B.S., M.A.M.T., M.A.d.B., H.E.B. and A.M.T. Writing: P.A.R. Revision: all authors. Supervision: R.S.H., J.W.M.M. and P.N.S. Funding acquisition: R.S.H., J.W.M.M. and P.N.S. All authors have read and agreed to the published version of the manuscript.

### Funding

This work was supported, in part, by the KWF Dutch Cancer Society (KWF10270, to J.W.M.M., P.N.S., and R.S.H.) and by the National Cancer Institute (P01-CA234228,

to R.S.H.). R.S.H. is the Ewing Halsell President's Council Distinguished Chair at University of Texas San Antonio, a CPRIT Scholar, and an Investigator of the Howard Hughes Medical Institute. None of the funding agencies had any role in the conceptualization, study design, data collection, interpretation of results, or the decision to submit this work for publication.

### **Institutional Review Board Statement**

Primary breast cancer specimens were obtained under the protocol to study biological markers associated with disease outcome and in accordance with the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands (<https://www.coreon.org/wp-content/uploads/2020/04/coreon-code-of-conduct-english.pdf>, accessed on 17 April 2023) and the new European General Data Protection Regulation (GDPR). This protocol states that the use of coded left-over material for scientific purposes and, therefore, for the greater good, does not require informed consent according to Dutch law. This waiver for informed consent was acknowledged by the medical ethics committee of the Erasmus Medical Centre Rotterdam, the Netherlands, in MEC 02.953. Informed Consent Statement: Primary breast cancer specimens were obtained under the protocol to study biological markers associated with disease outcome and in accordance with the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands (<https://www.coreon.org/wp-content/uploads/2020/04/coreon-code-of-conduct-english.pdf>, accessed on 17 April 2023) and the new European General Data Protection Regulation (GDPR). This protocol states that the use of coded left-over material for scientific purposes and, therefore, for the greater good, does not require informed consent according to Dutch law. This waiver for informed consent was acknowledged by the medical ethics committee of the Erasmus Medical Centre Rotterdam, the Netherlands, in MEC 02.953.

### **Data Availability Statement**

Data and materials are available upon request to P.N.S.

### **Acknowledgments**

We thank Prokopios P. Argyris, for providing assistance during the interpretation of A3B staining in breast (cancer) tissues, Jennifer L McCann, for helpful discussions and contributions during the early conceptualization phase of this study, and Anya Normandeau for careful reading of the manuscript.

### **Conflicts of Interest**

R.S.H. is cofounder of ApoGen Biotechnologies, Inc., which closed operations in April 2021. All other authors declare no competing or financial interests.

## References

1. Yizhak, K.; Aguet, F.; Kim, J.; Hess, J.M.; Kubler, K.; Grimsby, J.; Frazer, R.; Zhang, H.; Haradhvala, N.J.; Rosebrock, D.; et al. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science* 2019, 364, eaaw0726.
2. Martincorena, I.; Fowler, J.C.; Wabik, A.; Lawson, A.R.J.; Abascal, F.; Hall, M.W.J.; Cagan, A.; Murai, K.; Mahbubani, K.; Stratton, M.R.; et al. Somatic mutant clones colonize the human esophagus with age. *Science* 2018, 362, 911–917.
3. Alexandrov, L.B.; Kim, J.; Haradhvala, N.J.; Huang, M.N.; Tian Ng, A.W.; Wu, Y.; Boot, A.; Covington, K.R.; Gordenin, D.A.; Bergstrom, E.N.; et al. The repertoire of mutational signatures in human cancer. *Nature* 2020, 578, 94–101.
4. Burns, M.B.; Temiz, N.A.; Harris, R.S. Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat. Genet.* 2013, 45, 977–983.
5. Burns, M.B.; Lackey, L.; Carpenter, M.A.; Rathore, A.; Land, A.M.; Leonard, B.; Refsland, E.W.; Kotandeniya, D.; Tretyakova, N.; Nikas, J.B.; et al. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 2013, 494, 366–370.
6. Nik-Zainal, S.; Alexandrov, L.B.; Wedge, D.C.; Van Loo, P.; Greenman, C.D.; Raine, K.; Jones, D.; Hinton, J.; Marshall, J.; Stebbings, L.A.; et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* 2012, 149, 979–993.
7. Nik-Zainal, S.; Davies, H.; Staaf, J.; Ramakrishna, M.; Glodzik, D.; Zou, X.; Martincorena, I.; Alexandrov, L.B.; Martin, S.; Wedge, D.C.; et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 2016, 534, 47–54.
8. Roberts, S.A.; Lawrence, M.S.; Klimczak, L.J.; Grimm, S.A.; Fargo, D.; Stojanov, P.; Kiezun, A.; Kryukov, G.V.; Carter, S.L.; Saksena, G.; et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.* 2013, 45, 970–976.
9. Angus, L.; Smid, M.; Wilting, S.M.; van Riet, J.; Van Hoeck, A.; Nguyen, L.; Nik-Zainal, S.; Steenbruggen, T.G.; Tjan-Heijnen, V.C.G.; Labots, M.; et al. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nat. Genet.* 2019, 51, 1450–1458.
10. Brady, S.W.; McQuerry, J.A.; Qiao, Y.; Piccolo, S.R.; Shrestha, G.; Jenkins, D.F.; Layer, R.M.; Pedersen, B.S.; Miller, R.H.; Esch, A.; et al. Combating subclonal evolution of resistant cancer phenotypes. *Nat. Commun.* 2017, 8, 1231.
11. Bertucci, F.; Ng, C.K.Y.; Patsouris, A.; Droin, N.; Piscuoglio, S.; Carbuccion, N.; Soria, J.C.; Dien, A.T.; Adnani, Y.; Kamal, M.; et al. Genomic characterization of metastatic breast cancers. *Nature* 2019, 569, 560–564.
12. Roelofs, P.A.; Martens, J.W.M.; Harris, R.S.; Span, P.N. Clinical Implications of APOBEC3-mediated Mutagenesis in Breast Cancer. *Clin. Cancer Res.* 2022, OF1–OF12.
13. Law, E.K.; Sieuwerts, A.M.; LaPara, K.; Leonard, B.; Starrett, G.J.; Molan, A.M.; Temiz, N.A.; Vogel, R.I.; Meijer-van Gelder, M.E.; Sweep, F.C.; et al. The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. *Sci. Adv.* 2016, 2, e1601737.
14. Sammons, S.; Raskina, K.; Danziger, N.; Alder, L.; Schrock, A.B.; Venstrom, J.M.; Knutson, K.L.; Thompson, E.A.; McGregor, K.; Sokol, E.; et al. APOBEC Mutational Signatures in Hormone Receptor-Positive Human Epidermal Growth Factor Receptor 2-Negative Breast Cancers Are Associated With Poor Outcomes on CDK4/6 Inhibitors and Endocrine Therapy. *JCO Precis. Oncol.* 2022, 6, e2200149.
15. Sieuwerts, A.M.; Doebar, S.C.; de Weerd, V.; Verhoef, E.I.; Beauford, C.M.; Agahozo, M.C.; Martens, J.W.M.; van Deurzen, C.H.M. APOBEC3B Gene Expression in Ductal Carcinoma In Situ and Synchronous Invasive Breast Cancer. *Cancers* 2019, 11, 1062.

16. Sieuwerts, A.M.; Schrijver, W.A.; Dalm, S.U.; de Weerd, V.; Moelans, C.B.; Ter Hoeve, N.; van Diest, P.J.; Martens, J.W.; van Deurzen, C.H. Progressive APOBEC3B mRNA expression in distant breast cancer metastases. *PLoS ONE* 2017, 12, e0171343.
17. Sieuwerts, A.M.; Willis, S.; Burns, M.B.; Look, M.P.; Meijer-Van Gelder, M.E.; Schlicker, A.; Heideman, M.R.; Jacobs, H.; Wessels, L.; Leyland-Jones, B.; et al. Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. *Horm. Cancer* 2014, 5, 405–413.
18. Periyasamy, M.; Patel, H.; Lai, C.F.; Nguyen, V.T.M.; Nevedomskaya, E.; Harrod, A.; Russell, R.; Remenyi, J.; Ochocka, A.M.; Thomas, R.S.; et al. APOBEC3B-Mediated Cytidine Deamination Is Required for Estrogen Receptor Action in Breast Cancer. *Cell Rep.* 2015, 13, 108–121.
19. Swanton, C.; McGranahan, N.; Starrett, G.J.; Harris, R.S. APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. *Cancer Discov.* 2015, 5, 704–712.
20. Tsuboi, M.; Yamane, A.; Horiguchi, J.; Yokobori, T.; Kawabata-Iwakawa, R.; Yoshiyama, S.; Rokudai, S.; Odawara, H.; Tokiniwa, H.; Oyama, T.; et al. APOBEC3B high expression status is associated with aggressive phenotype in Japanese breast cancers. *Breast Cancer* 2016, 23, 780–788.
21. Barbosa-Desongles, A.; Hernandez, C.; Simo, R.; Selva, D.M. Testosterone induces cell proliferation and cell cycle gene overexpression in human visceral preadipocytes. *Am. J. Physiol. Cell Physiol.* 2013, 305, C355–C359.
22. Cescon, D.W.; Haibe-Kains, B.; Mak, T.W. APOBEC3B expression in breast cancer reflects cellular proliferation, while a deletion polymorphism is associated with immune activation. *Proc. Natl. Acad. Sci. USA* 2015, 112, 2841–2846.
23. Hirabayashi, S.; Shirakawa, K.; Horisawa, Y.; Matsumoto, T.; Matsui, H.; Yamazaki, H.; Sarca, A.D.; Kazuma, Y.; Nomura, R.; Konishi, Y.; et al. APOBEC3B is preferentially expressed at the G<sub>2</sub>/M phase of cell cycle. *Biochem. Biophys. Res. Commun.* 2021, 546, 178–184.
24. Roelofs, P.A.; Goh, C.Y.; Chua, B.H.; Jarvis, M.C.; Stewart, T.A.; McCann, J.L.; McDougale, R.M.; Carpenter, M.A.; Martens, J.W.; Span, P.N.; et al. Characterization of the mechanism by which the RB/E2F pathway controls expression of the cancer genomic DNA deaminase APOBEC3B. *eLife* 2020, 9, e61287.
25. Starrett, G.J.; Serebrenik, A.A.; Roelofs, P.A.; McCann, J.L.; Verhalen, B.; Jarvis, M.C.; Stewart, T.A.; Law, E.K.; Krupp, A.; Jiang, M.; et al. Polyomavirus T Antigen Induces APOBEC3B Expression Using an LXCXE-Dependent and TP53-Independent Mechanism. *mBio* 2019, 10, e02690-18.
26. Brown, W.L.; Law, E.K.; Argyris, P.P.; Carpenter, M.A.; Levin-Klein, R.; Ranum, A.N.; Molan, A.M.; Forster, C.L.; Anderson, B.D.; Lackey, L.; et al. A Rabbit Monoclonal Antibody against the Antiviral and Cancer Genomic DNA Mutating Enzyme APOBEC3B. *Antibodies* 2019, 8, 47.
27. Argyris, P.P.; Naumann, J.; Jarvis, M.C.; Wilkinson, P.E.; Ho, D.P.; Islam, M.N.; Bhattacharyya, I.; Gopalakrishnan, R.; Li, F.; Koutlas, I.G.; et al. Primary mucosal melanomas of the head and neck are characterised by overexpression of the DNA mutating enzyme APOBEC3B. *Histopathology* 2023, 82, 608–621.
28. Argyris, P.P.; Wilkinson, P.E.; Jarvis, M.C.; Magliocca, K.R.; Patel, M.R.; Vogel, R.I.; Gopalakrishnan, R.; Koutlas, I.G.; Harris, R.S. Endogenous APOBEC3B overexpression characterizes HPV-positive and HPV-negative oral epithelial dysplasias and head and neck cancers. *Mod. Pathol.* 2021, 34, 280–290.
29. Periyasamy, M.; Singh, A.K.; Gemma, C.; Kranjec, C.; Farzan, R.; Leach, D.A.; Navaratnam, N.; Palinkas, H.L.; Vertessy, B.G.; Fenton, T.R.; et al. p53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. *Nucleic Acids Res.* 2017, 45, 11056–11069.

30. Sweeney, K.J.; Swarbrick, A.; Sutherland, R.L.; Musgrove, E.A. Lack of relationship between CDK activity and G<sub>1</sub> cyclin expression in breast cancer cells. *Oncogene* 1998, 16, 2865–2878.
31. Jarvis, M.C.; Ebrahimi, D.; Temiz, N.A.; Harris, R.S. Mutation Signatures Including APOBEC in Cancer Cell Lines. *JNCI Cancer Spectr.* 2018, 2, pky002.
32. Goldman, M.J.; Craft, B.; Hastie, M.; Repecka, K.; McDade, F.; Kamath, A.; Banerjee, A.; Luo, Y.; Rogers, D.; Brooks, A.N.; et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* 2020, 38, 675–678.
33. Kanu, N.; Cerone, M.A.; Goh, G.; Zalmas, L.P.; Bartkova, J.; Dietzen, M.; McGranahan, N.; Rogers, R.; Law, E.K.; Gromova, I.; et al. DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. *Genome Biol.* 2016, 17, 185.
34. Lackey, L.; Law, E.K.; Brown, W.L.; Harris, R.S. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell Cycle* 2013, 12, 762–772.
35. Gong, D.; Ferrell, J.E., Jr. The roles of cyclin A2, B1, and B2 in early and late mitotic events. *Mol. Biol. Cell* 2010, 21, 3149–3161.
36. Grant, G.D.; Brooks, L., 3rd; Zhang, X.; Mahoney, J.M.; Martyanov, V.; Wood, T.A.; Sherlock, G.; Cheng, C.; Whitfield, M.L. Identification of cell cycle-regulated genes periodically expressed in U2OS cells and their regulation by FOXM1 and E2F transcription factors. *Mol. Biol. Cell* 2013, 24, 3634–3650.
37. Huang, X.; Di Liberto, M.; Jayabalan, D.; Liang, J.; Ely, S.; Bretz, J.; Shaffer, A.L., 3rd; Louie, T.; Chen, I.; Randolph, S.; et al. Prolonged early G(1) arrest by selective CDK4/CDK6 inhibition sensitizes myeloma cells to cytotoxic killing through cell cycle-coupled loss of IRF4. *Blood* 2012, 120, 1095–1106.
38. Leonard, B.; McCann, J.L.; Starrett, G.J.; Kosyakovsky, L.; Luengas, E.M.; Molan, A.M.; Burns, M.B.; McDougale, R.M.; Parker, P.J.; Brown, W.L.; et al. The PKC/NF-kappaB signaling pathway induces APOBEC3B expression in multiple human cancers. *Cancer Res.* 2015, 75, 4538–4547.
39. Dufourny, B.; Alblas, J.; van Teeffelen, H.A.; van Schaik, F.M.; van der Burg, B.; Steenbergh, P.H.; Sussenbach, J.S. Mitogenic signaling of insulin-like growth factor I in MCF-7 human breast cancer cells requires phosphatidylinositol 3-kinase and is independent of mitogen-activated protein kinase. *J. Biol. Chem.* 1997, 272, 31163–31171.
40. Johnson, N.; Bentley, J.; Wang, L.Z.; Newell, D.R.; Robson, C.N.; Shapiro, G.I.; Curtin, N.J. Pre-clinical evaluation of cyclindependent kinase 2 and 1 inhibition in anti-estrogen-sensitive and resistant breast cancer cells. *Br. J. Cancer* 2010, 102, 342–350.
41. Sadasivam, S.; DeCaprio, J.A. The DREAM complex: Master coordinator of cell cycle-dependent gene expression. *Nat. Rev. Cancer* 2013, 13, 585–595.
42. Petljak, M.; Alexandrov, L.B.; Brummel, J.S.; Price, S.; Wedge, D.C.; Grossmann, S.; Dawson, K.J.; Ju, Y.S.; Iorio, F.; Tubio, J.M.C.; et al. Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. *Cell* 2019, 176, 1282–1294.e20.
43. Jarvis, M.C.; Carpenter, M.A.; Temiz, N.A.; Brown, M.R.; Richards, K.A.; Argyris, P.P.; Brown, W.L.; Yee, D.; Harris, R.S. Mutational impact of APOBEC3B and APOBEC3A in a human cell line. *bioRxiv* 2022, 26, 539–551.
44. Litovchick, L.; Sadasivam, S.; Florens, L.; Zhu, X.; Swanson, S.K.; Velmurugan, S.; Chen, R.; Washburn, M.P.; Liu, X.S.; DeCaprio, J.A. Evolutionarily conserved multisubunit RBL2/p130 and E2F4 protein complex represses human cell cycle-dependent genes in quiescence. *Mol. Cell* 2007, 26, 539–551.
45. Siriwardena, S.U.; Perera, M.L.W.; Senevirathne, V.; Stewart, J.; Bhagwat, A.S. A Tumor-Promoting Phorbol Ester Causes a Large Increase in APOBEC3A Expression and a Moderate Increase in APOBEC3B Expression in a Normal Human Keratinocyte Cell Line without Increasing Genomic Uracils. *Mol. Cell. Biol.* 2019, 39, e00238-18.

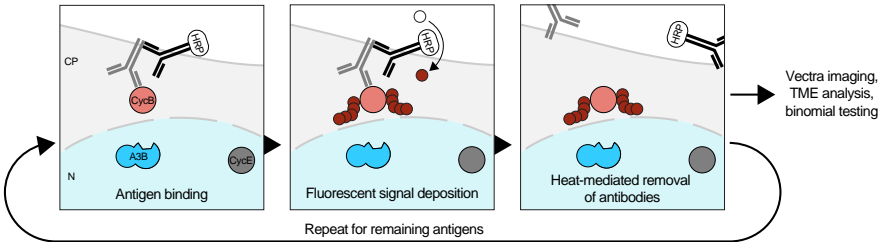


46. Yamazaki, H.; Shirakawa, K.; Matsumoto, T.; Kazuma, Y.; Matsui, H.; Horisawa, Y.; Stanford, E.; Sarca, A.D.; Shirakawa, R.; Shindo, K.; et al. APOBEC3B reporter myeloma cell lines identify DNA damage response pathways leading to APOBEC3B expression. *PLoS ONE* 2020, 15, e0223463.
47. Lin, L.; Holmes, B.; Shen, M.W.; Kammeron, D.; Geijsen, N.; Gifford, D.K.; Sherwood, R.I. Comprehensive Mapping of Key Regulatory Networks that Drive Oncogene Expression. *Cell Rep.* 2020, 33, 108426.
48. Saito, Y.; Miura, H.; Takahashi, N.; Kuwahara, Y.; Yamamoto, Y.; Fukumoto, M.; Yamamoto, F. Involvement of APOBEC3B in mutation induction by irradiation. *J. Radiat. Res.* 2020, 61, 819–827.
49. Periyasamy, M.; Singh, A.K.; Gemma, C.; Farzan, R.; Allsopp, R.C.; Shaw, J.A.; Charmsaz, S.; Young, L.S.; Cunnea, P.; Coombes, R.C.; et al. Induction of APOBEC3B expression by chemotherapy drugs is mediated by DNA-PK-directed activation of NF-kappaB. *Oncogene* 2021, 40, 1077–1090.
50. Wang, W.; Mani, A.M.; Wu, Z.H. DNA damage-induced nuclear factor-kappa B activation and its roles in cancer progression. *J. Cancer Metastasis Treat.* 2017, 3, 45–59.
51. Hoopes, J.I.; Cortez, L.M.; Mertz, T.M.; Malc, E.P.; Mieczkowski, P.A.; Roberts, S.A. APOBEC3A and APOBEC3B Preferentially Deaminate the Lagging Strand Template during DNA Replication. *Cell Rep.* 2016, 14, 1273–1282.
52. McDaniel, Y.Z.; Wang, D.; Love, R.P.; Adolph, M.B.; Mohammadzadeh, N.; Chelico, L.; Mansky, L.M. Deamination hotspots among APOBEC3 family members are defined by both target site sequence context and ssDNA secondary structure. *Nucleic Acids Res.* 2020, 48, 1353–1371.
53. Gazova, I.; Lefevre, L.; Bush, S.J.; Clohisey, S.; Arner, E.; de Hoon, M.; Severin, J.; van Duin, L.; Andersson, R.; Lengeling, A.; et al. The Transcriptional Network That Controls Growth Arrest and Macrophage Differentiation in the Human Myeloid Leukemia Cell Line THP-1. *Front. Cell Dev. Biol.* 2020, 8, 498.
54. Kim, Y.S.; Sun, S.; Yoon, J.S.; Ko, Y.H.; Won, H.S.; Kim, J.S. Clinical implications of APOBEC3A and 3B expression in patients with breast cancer. *PLoS ONE* 2020, 15, e0230261.
55. Madsen, P.; Anant, S.; Rasmussen, H.H.; Gromov, P.; Vorum, H.; Dumanski, J.P.; Tommerup, N.; Collins, J.E.; Wright, C.L.; Dunham, I.; et al. Psoriasis upregulated phorbol-1 shares structural but not functional similarity to the mRNA-editing protein apobec-1. *J. Investig. Dermatol.* 1999, 113, 162–169.
56. Serebrenik, A.A.; Argyris, P.P.; Jarvis, M.C.; Brown, W.L.; Bazzaro, M.; Vogel, R.I.; Erickson, B.K.; Lee, S.H.; Goergen, K.M.; Maurer, M.J.; et al. The DNA Cytosine Deaminase APOBEC3B is a Molecular Determinant of Platinum Responsiveness in Clear Cell Ovarian Cancer. *Clin. Cancer Res.* 2020, 26, 3397–3407.
57. Denkert, C.; Untch, M.; Benz, S.; Schneeweiss, A.; Weber, K.E.; Schmatloch, S.; Jackisch, C.; Sinn, H.P.; Golovato, J.; Karn, T.; et al. Reconstructing tumor history in breast cancer: Signatures of mutational processes and response to neoadjuvant chemotherapy. *Ann. Oncol.* 2021, 32, 500–511.
58. Barroso-Sousa, R.; Keenan, T.E.; Pernas, S.; Exman, P.; Jain, E.; Garrido-Castro, A.C.; Hughes, M.; Bychkovsky, B.; Umeton, R.; Files, J.L.; et al. Tumor Mutational Burden and PTEN Alterations as Molecular Correlates of Response to PD-1/L1 Blockade in Metastatic Triple-Negative Breast Cancer. *Clin. Cancer Res.* 2020, 26, 2565–2572.
59. Litchfield, K.; Reading, J.L.; Puttick, C.; Thakkar, K.; Abbosh, C.; Bentham, R.; Watkins, T.B.K.; Rosenthal, R.; Biswas, D.; Rowan, A.; et al. Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* 2021, 184, 596–614.e14.
60. Boichard, A.; Pham, T.V.; Yeerna, H.; Goodman, A.; Tamayo, P.; Lippman, S.; Frampton, G.M.; Tsigelny, I.F.; Kurzrock, R. APOBEC-related mutagenesis and neo-peptide hydrophobicity: Implications for response to immunotherapy. *Oncoimmunology* 2019, 8, 1550341.

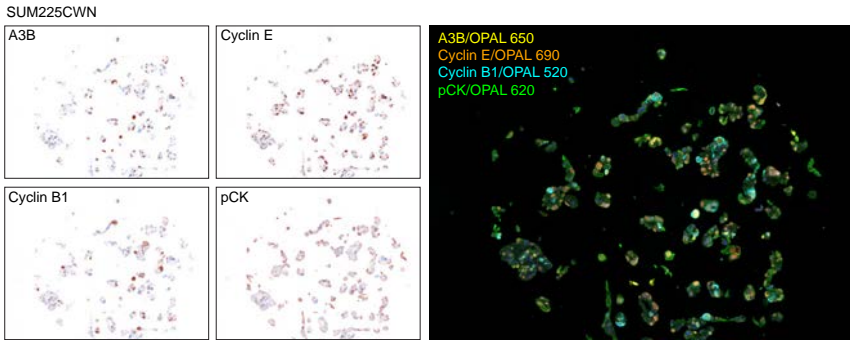


# Supplementary Information

**A**

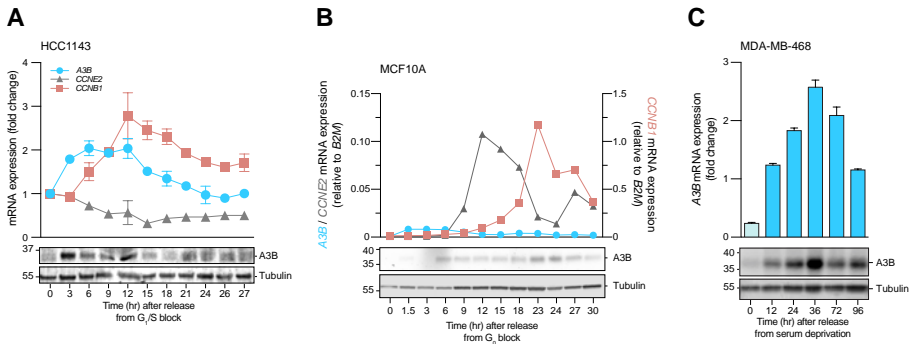


**B**



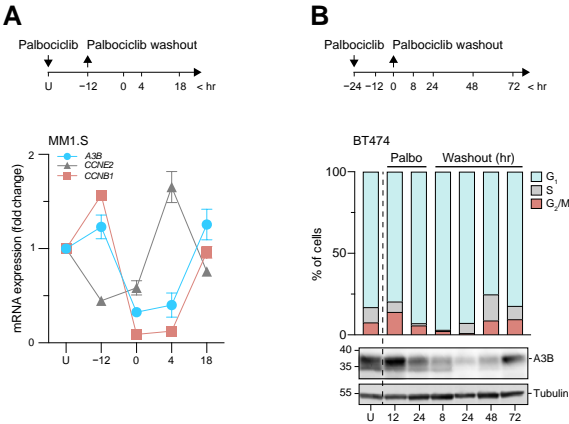
## Supplementary Figure 1. Additional information regarding multiplex IHC

**A** Schematic overview of the workflow used during multiplex IHC, which involves the sequential binding of antigens by monoclonal primary antibodies, the development and deposition of fluorescent signal through HRP conjugated secondary antibodies, and the heat-mediated removal of both primary and secondary antibodies. After completion, slides were analyzed (see Methods) and data was subjected to binomial testing using an in-house Python script. **B** Representative microscope field-of-view during image acquisition using multiplex IHC for A3B, Cyclin E2, Cyclin B1, and pan CK. An example of the breast cancer cell line SUM225CWN is shown. Note that while fluorescent markers are used for signal deposition, images used for display of individual markers are pseudo-colored to emulate chromogenic staining patterns. Pan cytokeratin was used to determine cell boundaries after which pair-wise comparisons were made between A3B and individual Cyclins. The image on the right represents a composite view of all four channels pseudo-colored in yellow (A3B), orange (Cyclin E), cyan (Cyclin B1), and green (pan-CK)



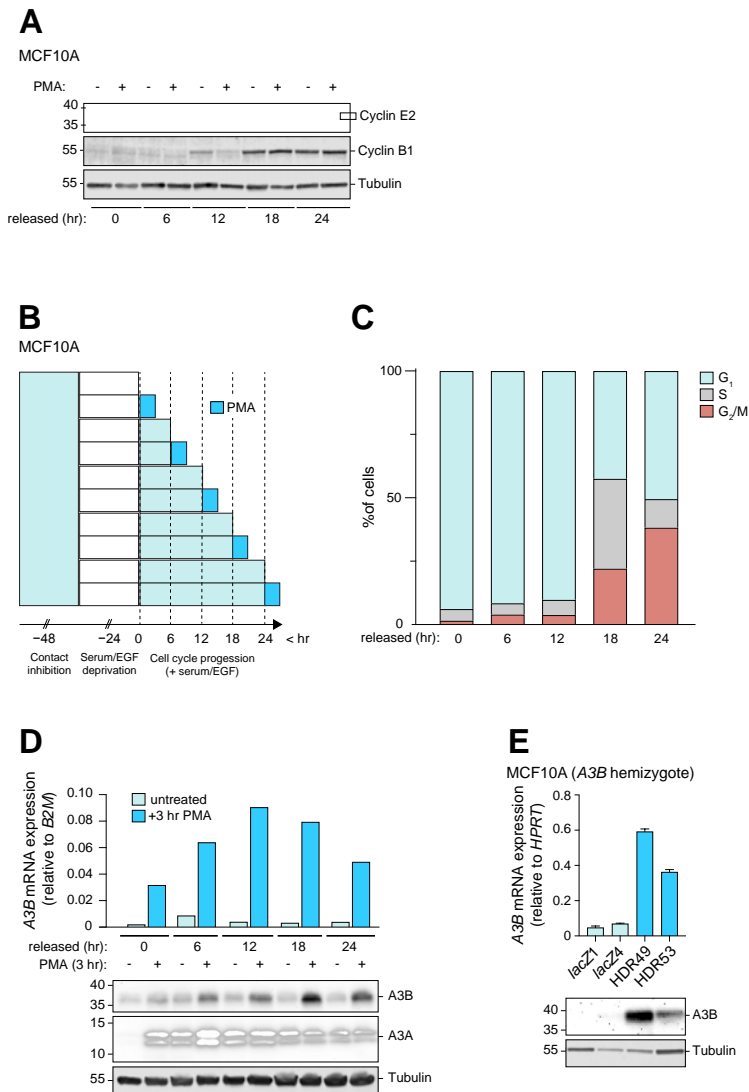
**Supplementary Figure 2. Detailed time-courses and MDA-MB-468 serum deprivation**

**A-B** RT-qPCR for *A3B*, *CCNE2*, and *CCNB1* (top) and *A3B* immunoblot (bottom) in HCC1143 and MCF10A, taken at the indicated times following a double thymidine block or serum starvation, respectively. These data represent the full extent of the HCC1143 and MCF10A samples used in the experiment described in Figure 2B-C and additionally show *A3B* protein expression. **C** RTqPCR (top) and immunoblot (bottom) for *A3B* expression performed on MDA-MB-468 cells at various hours following serum deprivation.



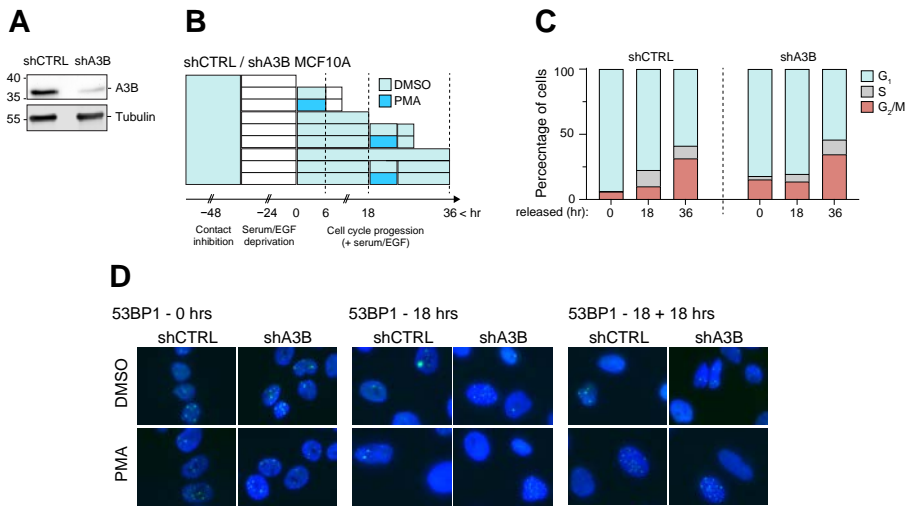
**Supplementary Figure 3. Additional cell cycle synchronization courses using palbociclib**

**A** Reanalyzed *A3B*, *CCNE2*, and *CCNB1* expression data in the multiple myeloma cell line MM1.S as sampled before, during, and after synchronization using palbociclib. Publicly available data (GSE35728) was used. **B** Additional time course of BT474 synchronized using palbociclib followed by an extended outgrowth period (see top panel). Cell cycle progression, as analyzed by PI stain, and an *A3B* immunoblot are shown.



**Supplementary Figure 4. Additional cell cycle synchronization courses using palbociclib**

**A** Supporting immunoblots showing the expression of Cyclin E2 and Cyclin B1 before and after PMA treatment throughout the experiments described in Figure 4B-D. Note that Cyclin E2 expression is consistently impacted by PMA treatment. **B-D** Additional cell cycle experiment, with minor adaptations as compared to Figure 4B-D. Note that in this experiment, PMA treated cells were compared to control samples taken at the time of PMA addition (B). Cell cycle progression was measured by flowcytometry (C). RT-qPCR and immunoblot assays were performed to analyze A3B expression in response to PMA treatment (D). A3A expression was included as a control for the general inducibility of genes under control of the PKC/ncNF- $\kappa$ B pathway. **E** A3B mRNA (top) and protein (bottom) expression of two *lacZ* control (light blue) and HDR clones (blue). The HDR clones in this study possess a deactivating mutation in the repressive E2F element within the *A3B* promoter.



**Supplementary Figure 5. PMA induction of A3B and DNA damage response**

**A** Immunoblots showing A3B expression in MCF10A cells after transduction with shCTRL and shA3B lentiviral particles and puromycin selection. **B** Treatment regimen used to investigate the activation of the DNA damage response upon PMA treatment in growth-arrested cells, cells released for 18 hrs followed by a 3 hr outgrowth period (early response), and cells released for 18 hours followed by a 12 hr outgrowth period (delayed response). **C** Flowcytometry plots showing cell cycle progression of both cell lines upon release in normal growth medium. **D** Fluorescent microscopy images showing 53BP1 foci after PMA treatment in growth-arrested cells and cells released for 18 hrs (followed by a 3 or 12 hr outgrowth period).

**Supplementary Table 1. Relevant characteristics of the main cell lines used in this study**

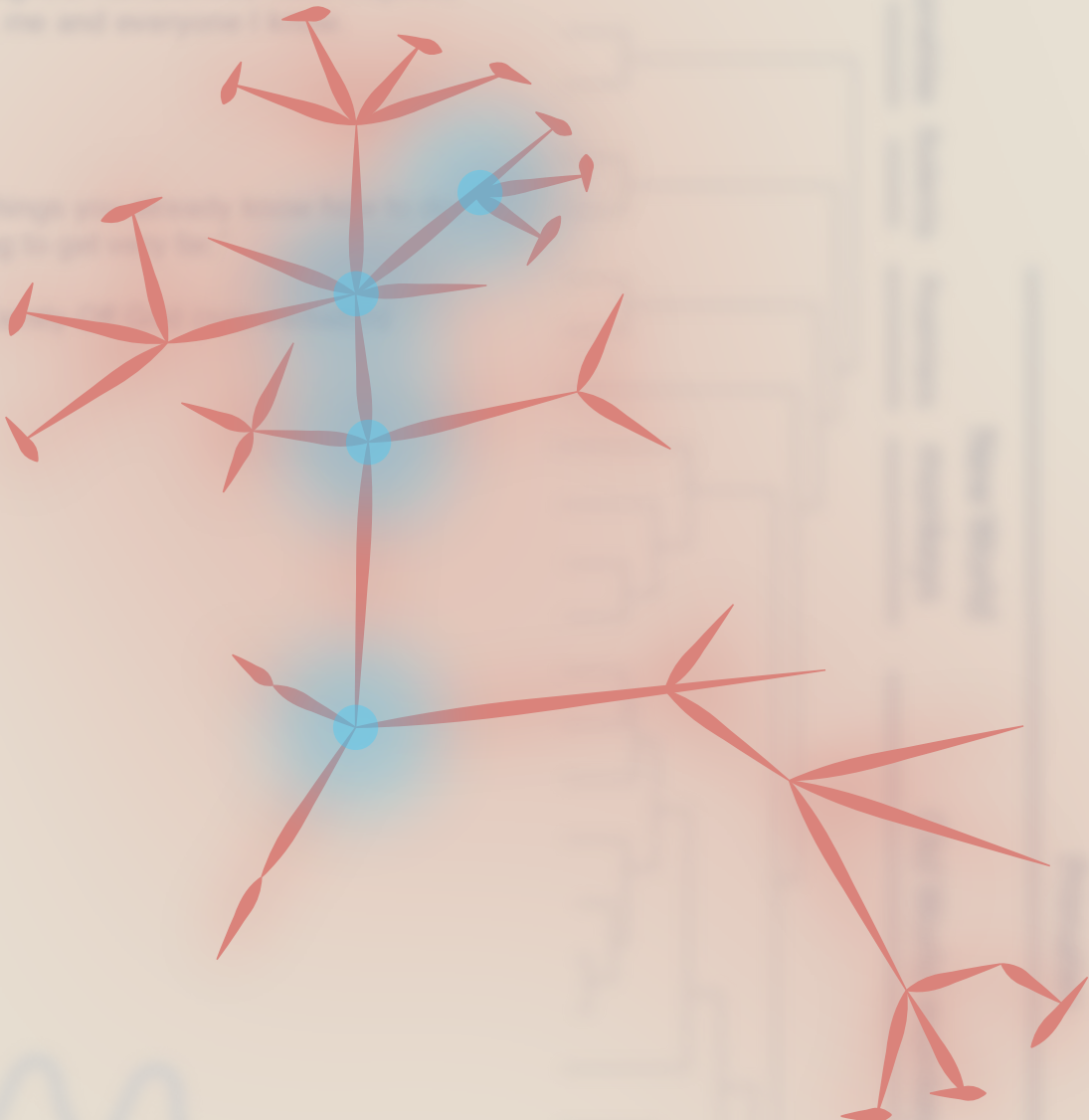
Cell line	Type of cell line	Level of A3B expression (relative to normal)	Synchronizable by double thymidine block?	Alternative cell cycle synchronization option	A3B inducibility by PMA
MCF10A	Immortalized (normal-like)	Very low	No (results in cell death)	Contact inhibition followed by mitogen deprivation	High during proliferation
MCF7	Breast cancer	Low	Yes	Not tested	Limited, during proliferation
HCC1954	Breast cancer	Low/intermediate	Yes	Not tested	Not tested
BT474	Breast cancer	Intermediate	No (proliferation rate too low)	Palbociclib	Not tested
HCC1143	Breast cancer	High	Yes	Not tested	Limited, during proliferation

A dedication to my parents,  
for delightful memories have inspired  
me and everyone I know.

For the things we have done together  
and for the things we are going to do together.

With love and affection,

With love and affection,



## Clinical Implications of APOBEC3-Mediated Mutagenesis in Breast Cancer

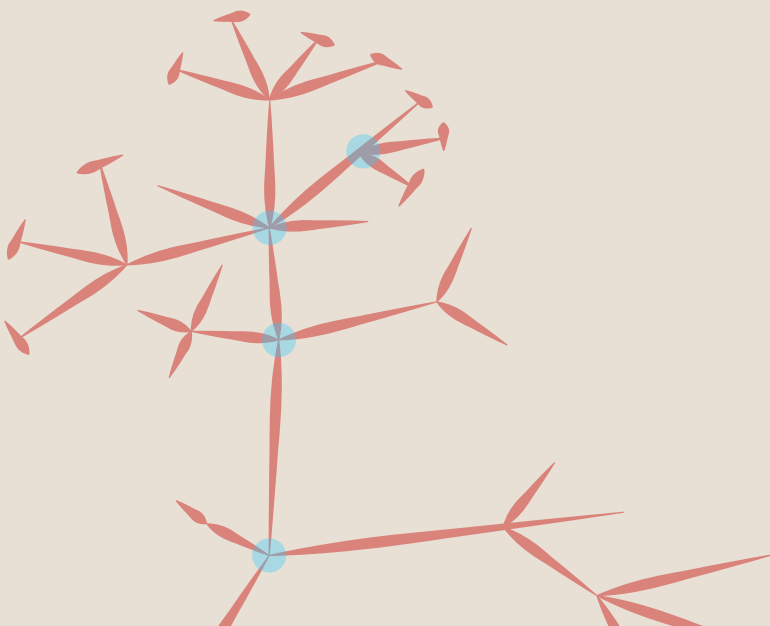
Pieter A Roelofs<sup>1,2</sup>, John W. M. Martens<sup>3</sup>, Reuben S. Harris<sup>2,4,5</sup>, and Paul N. Span<sup>1,\*</sup>

1. Department of Radiation Oncology, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands
2. Department of Biochemistry, Molecular Biology and Biophysics, Masonic Cancer Center, Institute for Molecular Virology, and Center for Genome Engineering, University of Minnesota, Minneapolis, MN 55455, USA
3. Department of Medical Oncology, Erasmus MC Cancer Institute, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands
4. Department of Biochemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, TX 78229, USA
5. Howard Hughes Medical Institute, University of Texas Health San Antonio, San Antonio, TX 78229, USA

Published in: Clinical Cancer Research. 2023;29(9):1658. doi: 10.1158/1078-0432.CCR-22-2861.

## Introduction

The genomic landscape of breast cancer is shaped by many mutagenic processes, which promote intratumor and intertumor genomic heterogeneity and contribute to tumor evolution and thereby treatment resistance (1, 2). These mutational processes are computationally distinguishable as signatures with different etiologic causes (1). For example, a mutational signature attributable to deamination of cytosine bases in DNA catalyzed by apolipoprotein B mRNA editing enzyme, catalytic polypeptide-3 (APOBEC3) enzymes is evident in a large proportion of breast cancers (1). Two of the seven APOBEC3 enzymes found in humans, APOBEC3A (A3A) and APOBEC3B (A3B), have been causally linked to the observed APOBEC mutation signature in breast cancer. This review focuses first on how APOBEC3-positive tumors can be diagnosed, second on how the proteins responsible may become dysregulated in breast cancer, and finally on the clinical impact and implications of APOBEC3-mediated mutagenesis for novel and patient-specific treatment opportunities. Of note, APOBEC3 enzymes may also have roles in cancer that are independent of mutagenesis, for example, estrogen-dependent gene expression (3), R-loop homeostasis (bioRxiv 2021.08.30.458235v1), and RNA editing (bioRxiv 2022.06.01.494353), which are beyond the scope of this review.



## The APOBEC3 ABC's and How to Detect Them

### Family member profiles

APOBEC3 proteins catalyze the deamination of cytosines, thereby converting them into premutagenic uracils (reviewed in ref. 4). Human cells can express up to seven APOBEC3 proteins, A3A, A3B, A3C, A3D, A3F, A3G, and A3H, which can be further distinguished by single amino acid variants, of which A3H is the most variable in the human population with over a dozen phylogenetically distinct haplotypes (refs. 5, 6; Fig. 1A). In addition, although all A3 members are structurally similar, differences in amino acid sequence and functionality allow subclassification into different domain groupings (called Z1, Z2, and Z3, respectively color-coded green, orange, and blue in Fig. 1A). The composition of these domains is evolutionarily conserved amongst higher primates and most of these three domains are also expressed in other mammalian orders, including even- and odd-toed ungulates, bats, and afrotheres (5). In humans, the Z1 domain provides the catalytically active pocket in A3A, A3B, and A3G, while the catalytically active pockets of A3C-F and A3H are encoded by Z2 and Z3 domains, respectively (Fig. 1A). In addition, a catalytically inactive form of Z2 is present as the N-terminal domain of A3B, A3D, A3F, and A3G, possibly serving to regulate catalytic activity, subcellular localization, and the packaging into viral particles (7–9). As such, several A3 members, including A3D, A3F, A3G, and A3H, are capable of restricting HIV-1, whereas other virus types may be restricted by A3A, A3B, and A3H [(10–13) and reviewed in refs. 4, 14]. For instance, the DNA-based hepatitis B virus may be restricted through the editing capabilities of A3B, A3C, A3F, and A3G (15), and the large DNA herpesviruses by A3B and potentially also by A3A (16).

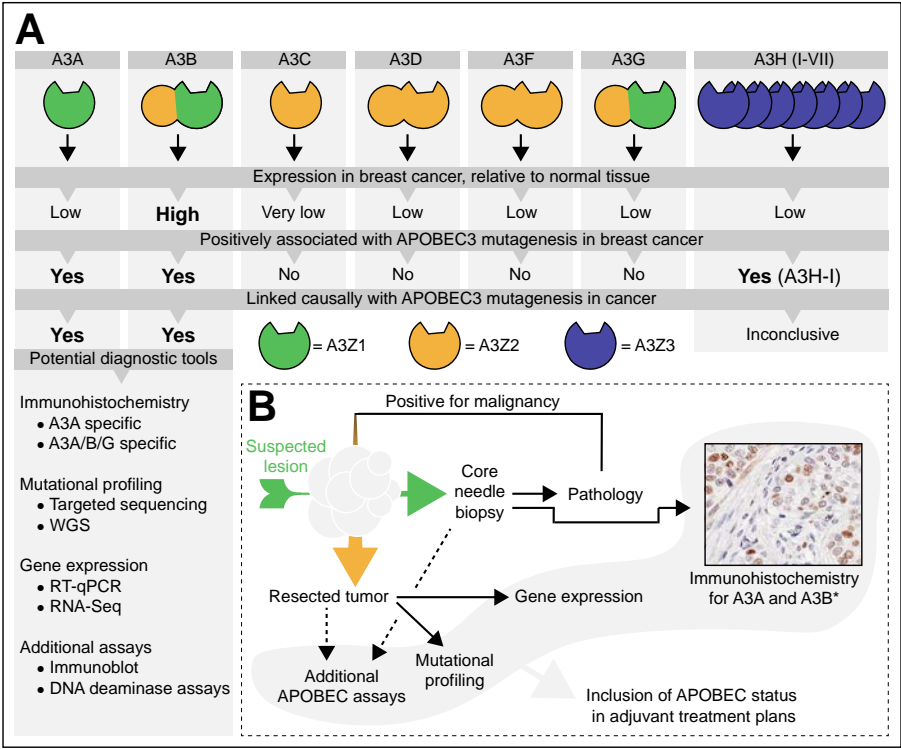
### A3A and A3B are major contributors to cancer

Whole-genome sequencing (WGS) found that off-target activity of APOBEC3 to chromosomal DNA constitutes a major source of somatically acquired mutations in a variety of malignancies, including breast cancer (17–20). Mutational activity by APOBEC3 proteins can be computationally identified in sequencing data as mutations occurring at cytosines within a 5'-TCW (W = A or T) trinucleotide context (21). The mutational process within this context starts with the deamination of cytosine into uracil, which then templates for thymine during replication and base pairs with adenosine. After a round of replication, a C-to-T transition is then immortalized into the genome. Alternatively, uracil excision by DNA glycosylases and subsequent error-prone repair by translesion synthesis polymerases can generate C-to-G transversions. These two distinguishable single base substitutions (SBS) mutations are included in the over 40 etiologically distinct mutation signatures found in pan cancer datasets (referred to as SBS2 and SBS13,



respectively; see ref. 1). Importantly, these mutation signatures are consistent with the mutagenic characteristics of A3A, A3B, and A3H haplotype I (A3H-I). Although previously proposed as a likely contributor, the role of A3H-I in cancer-related mutagenesis has recently been questioned [bioRxiv 2022.04.26.489523v2 (22, 23)]. In comparison, A3A and A3B are currently seen as major contenders to the origin of APOBEC3 deaminase activity in cancer, which is further described below. Of all APOBEC3 proteins possibly involved in breast cancer mutagenesis, A3B is the only deaminase overexpressed and steadily present in the nuclear compartment (7, 17, 24, 25). Various studies have directly and indirectly connected A3B activity with APOBEC3-mediated mutagenesis in several cancer types, including breast cancer (Fig. 1A; bioRxiv 2022.04.26.489523v2; refs. 17, 26). However, APOBEC3-mediated mutagenesis can still be detected in breast cancers of patients carrying loss of A3B. Loss of A3B presents as a chimeric allele, where the A3A coding region is fused to the A3B 3' untranslated region, which is rare in European and African populations, but present in approximately 37%, approximately 58%, and approximately 93% of East Asians, native American, and Oceanic populations, respectively (27, 28). This A3A-B fusion allele indicates that additional APOBEC3 members, such as A3A, also contribute to the overall level of APOBEC signature SBS mutations in breast cancer.

Like A3B, the potent deamination activity of A3A has also been causally linked to increased levels of APOBEC3-mediated mutagenesis (Fig. 1A; bioRxiv 2022.04.26.489523v2; refs. 23, 29–32). In fact, A3A has been proposed as the dominant deaminase over A3B in breast cancer, predominantly based on the reported computationally distinguishable mutational signatures of A3A and A3B as established in yeast, cell lines and subsequently tumors (30, 33, 34). However, although potentially useful in the detection of A3A-mediated mutagenesis specifically, more recent analyses in HAP1 cells have shown that this approach may not provide the necessary resolution between these two deaminases and highlights the appreciable contribution of A3A as well as A3B (bioRxiv 2022.04.26.489523v2). Furthermore, while A3A was recently proposed as a major contributor to SBS2 and SBS13 in breast cancer cells, A3B still has an appreciable contribution to APOBEC3-mediated mutagenesis (23). Therefore, because both A3A and A3B are directly implicated with the accumulation of APOBEC signatures, and a detailed correlation of the most relevant deaminase in relation to breast cancer development is yet to be fully established (and they may also contribute combinatorially), APOBEC3-mediated mutagenesis in this review will not specifically be attributed to either enzyme (Fig. 1A). Reproducible and clinically implementable detection methods, discussed below, can further consolidate the clinical relevance of APOBEC3 proteins and their mutagenic activity.



**Figure 1.** The APOBEC3 enzymes, their association with breast cancer, and the diagnostic methods available. **(A)** Breakdown of individual APOBEC3 family members, their conserved domain composition (5, 6), their expression levels in cancer (17, 22,30), and their causal involvement in the observed APOBEC mutagenesis pattern observed in cancer (13, 17, 22, 33). The list of potential diagnostic tools denotes published methods suitable for the detection of APOBEC3 enzymes, their deaminase activity, or the APOBEC SBS signatures (bioRxiv 2022.04.26.489523v2; refs. 12, 17, 35, 38, 44–49). **(B)** Proposed flow chart for the inclusion of the APOBEC status in the consideration of suitable adjuvant treatment plans. An initial core needle biopsy is taken from the suspected lesion (green arrow) and IHC for A3A, and/or A3B is performed in parallel to conventional clinical pathology. If malignant and operable, the freshly resected tumor (orange arrow) is subjected to additional assays, including mutational profiling and gene expression analyses. The resultant APOBEC status may then be included in the adjuvant treatment plans. The example shown here is considered A3B specific because of its nuclear localization.

### Options for clinical detection of APOBEC3-mediated mutagenesis

As part of the initial histopathologic assessment of malignancy, expression of APOBEC3 proteins can be detected through IHC in, for instance, diagnostic core needle biopsies (Fig. 1B). A rabbit mAb has recently proven to be suitable for the specific detection of A3A protein using immunofluorescence (bioRxiv 2022.04.26.489523v2). For detection of A3B, the most frequently used antibody is 5210-87-13, a rabbit

monoclonal that detects A3A, A3B, and A3G due to a shared epitope (35). As the only APOBEC3 protein with dominant nuclear localization, A3B expressed by tumor cells can be clearly distinguished from other APOBEC3 proteins, including A3A, which are present as cell-wide or cytoplasmic proteins (14). The IHC detection of A3B has been demonstrated in tumor tissue from several cancer types, including head/neck and ovarian cancer (36–38). Given the low expression of APOBEC3 proteins in most healthy cells, staining of A3A and A3B in tumor cells can be readily detected. It is currently unknown which molecular breast cancer subtype is most likely to score positive for the IHC detection of A3A and A3B protein. However, protein expression can be seen at early stages of tumor development (39) and, therefore, IHC detection of A3A and A3B may conveniently be included in histopathologic analysis and stained in parallel to molecular markers such as estrogen receptor (ER) and HER2. We therefore recommend all samples be subjected to the A3A/B IHC analysis. Furthermore, to establish mutational contributions, the resected primary tumor may also be analyzed by DNA sequencing (Fig. 1B). To the best of our knowledge, targeted sequencing approaches have yet to be optimized to detect APOBEC SBS signatures. Therefore, whole-exome sequencing or WGS is necessary to gain a comprehensive overview of APOBEC3-mediated mutagenesis. On the basis of available sequencing data, HER2-amplified breast cancers are most likely to display pan-genomic APOBEC SBS signatures, and ER<sup>+</sup> disease may additionally contain APOBEC induced mutations with clinical relevance (40, 41). These samples may be prioritized in sequencing efforts. Moreover, because tumors with homologous recombination (HR) deficiencies such as loss of BRCA1 or BRCA2 rarely show APOBEC SBS signatures (42, 43), testing labs may be prudent to focus APOBEC diagnostic efforts on HR-proficient tumors. Importantly, by combining genome sequencing with IHC, historical APOBEC3 activity (i.e., presence of APOBEC SBS signatures, but absence of A3A and A3B protein), may be distinguished from ongoing APOBEC3 activity (i.e., presence of both APOBEC SBS signatures and IHC positivity). Other techniques that can consolidate the expression and/or activity of A3A and A3B include quantification by RNA-based methods (i.e., qRT-PCR, RNA sequencing, and/or RNA scope), immunoblotting, A3A-dependent RNA editing (44), and DNA deaminase assays (Fig. 1B; refs. 1, 12, 17, 38, 45–49). To understand the biological context surrounding APOBEC3-mediated mutagenesis, it is important to know how these enzymes can be expressed in breast cancer (discussed below).

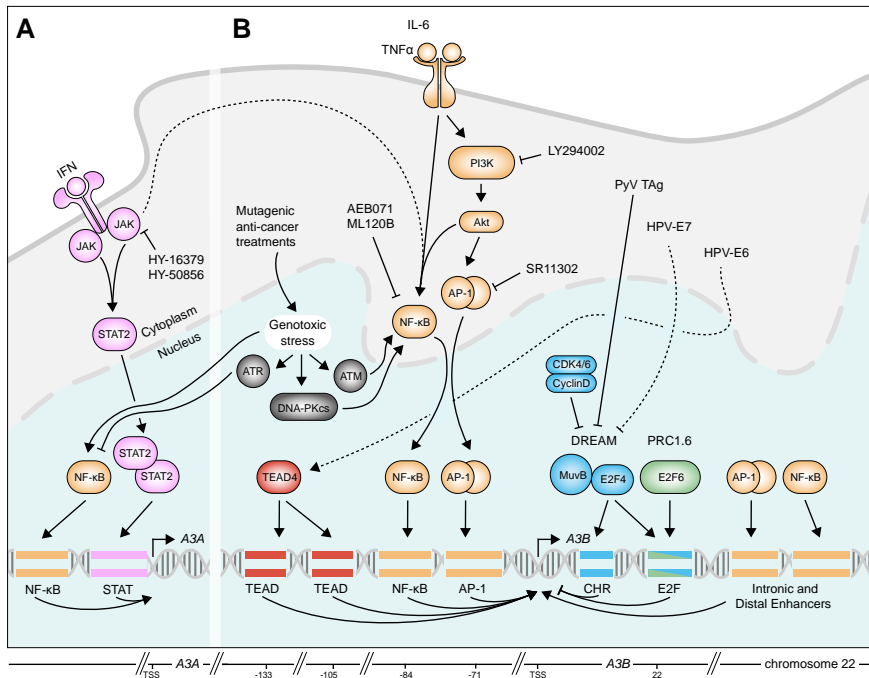
## **(Dys)regulation of APOBEC3 Enzymes in Breast Cancer**

Because of their prominent role in the innate immune system, much of the initial data on APOBEC3 regulation stem from virology research. These observations have

proven insightful in the identification of mechanisms underlying APOBEC3 (dys) regulation, even though viruses are unlikely to be directly involved in breast cancer (reviewed in ref. 50).

## Regulation of A3A expression

One of the most prominent factors that induce A3A expression are IFNs. The pleiotropic group of IFNs, most commonly type-I and type-II IFN, are produced as first-responder inflammatory cytokines by, among others, tumor-resident immune cells. Type-I IFNs potently induce A3A, while type-II IFNs only activate A3A marginally (51). In breast epithelial cells A3A transcriptional activation through type-I IFNs requires the transcription factor complex STAT2 and its upstream regulator JAK (Fig. 2A; ref. 52). Type-I IFN also induces A3A in tumor cell lines, including those from lung, bladder, and breast cancer (44, 45, 53–55).



**Figure 2.** Transcriptional regulation of A3A and A3B. **(A)** Activation of IFN signaling facilitates the recruitment of STAT2 to the promoter of A3A, while genotoxic stress promotes the recruitment of the NFκB transcription factor complex. Transcriptional activation of A3A through genotoxic stress is inhibited by ATR. **(B)** The PKC/ncNFκB pathway dominates A3B transcription and recruits transcription factor complexes to the A3B promoter, intronic, and distal enhancers. In turn, the PKC/ncNFκB pathway is activated by genotoxic stress, TNFα, IL6, and possibly IFN signaling. Transcriptional repression is facilitated by the DREAM complex (and the upstream RB/E2F pathway) and the PRC1.6 complex. Viral oncogenes, including HPV-E6, HPV-E7, and PyV Tag, can also activate A3B, although this is unlikely to contribute to A3B expression in breast cancer.

A3A transcription can also be activated by the canonical PKC/ncNfκB pathway, which itself is activated by a myriad of inflammatory and genotoxic stresses (Fig. 2A and discussed below; refs. 52, 56). However, the activation of A3A upon genotoxic insults is closely guarded by the protein ataxia telangiectasia and rad3 related (ATR), which generally serves as a protective protein during DNA replication stress (Fig. 2A; ref. 57). Interestingly, ATR also prevents direct induction of A3A through commonly used cancer treatments, particularly those that cause replication stress (52). However, whether anticancer treatments can stimulate inflammatory pathways and subsequently induce A3A expression, and whether this impacts disease trajectory, remains unclear. Moreover, the current body of knowledge on the transcriptional regulation of A3A strongly indicates that chemical inhibition of the IFN and/or PKC/ncNfκB pathways may be a useful approach to limit A3A expression. Currently, the best-defined inhibitor that has been directly investigated within this context targets JAK2 and effectively prevents A3A induction by IFN (Fig. 2; ref. 52).

### Regulation of A3B expression

The PKC/ncNfκB signaling pathway and its associated proteins PI3K and AKT are at the center of A3B transcriptional activation (38, 58, 59). PKC/ncNfκB and related AP-1 complexes are recruited to sites within the A3B promoter, intronic regions, and a distant enhancer, thus activating A3B transcription (Fig. 2B; ref. 60). Various upstream stimuli converge to activate the PKC/ncNfκB signaling pathway, thereby eliciting an increase in A3B expression in breast cancer. For example, TNFα, a pro-inflammatory cytokine, activates the PKC/ncNfκB signaling pathway and stimulates A3B expression (Fig. 2B; ref. 60). Furthermore, the proinflammatory cytokine IL6, produced by both leukocytes and several solid tumor cell lines, can activate PKC/ncNfκB (reviewed in ref. 61) and thereby A3B (Fig. 2B; refs. 60, 62).

Induction of A3B by type I and type II IFNs by PKC/ncNfκB cross-activation is also observed in (oropharyngeal and lung) cancer cell lines (53, 55). However, relative to A3A, this induction of A3B by IFNs is less consistent between different tissue types, indicating the presence of currently unknown regulatory mechanisms (Fig. 2B). Finally, DNA double-strand breaks, which commonly occur in response to ionizing radiation, various chemotherapeutic drugs, or advanced genomic instability (reviewed in ref. 63) can increase A3B expression in, among others, breast cancer cell lines (54, 59, 60, 64–66). Induction of A3B through the PKC/ncNfκB pathway is also dependent on several main responders to DNA double-strand breaks, specifically DNA-PKcs and ataxia telangiectasia mutated; refs. Fig.2B; refs. 59, 60). Interestingly, and further emphasizing the central role of the PKC/ncNfκB pathway in A3B induction, several preclinical and clinically approved PKC inhibitors have been

shown to effectively, and dose-responsively, inhibit expression of A3B in various (breast) cancer cell lines (Fig. 2; refs. 38, 60). Future studies could further explore the usefulness of these compounds in restricting the mutagenic activities of A3B.

In addition, several viral oncoproteins, including human papillomavirus (HPV) E6, E7, and polyomavirus T-antigen, are strongly implicated with A3B transcriptional dysregulation and the accumulation of APOBEC SBS signatures in several virally induced cancers (Fig. 2B; refs. 51, 67–72). Specifically, HPV-E6 may drive expression of A3B through recruitment of the transcription factor TEAD to two distinct binding sites at the A3B promoter (67). In addition, both HPV-E7 and polyomavirus T-antigen target the transcriptional repressor DREAM which, as an integral component of the RB/E2F pathway, facilitates the timely expression of cell cycle-associated genes during proliferation (Fig. 2B). A3B is repressed by the DREAM and the PRC1.6 complex, which are recruited to an E2F binding site within the A3B promoter in normal-like breast epithelial cells (Fig. 2B; refs. 26, 71). Importantly, disruption of the RB/E2F pathway is common in breast cancer and associates with increased APOBEC SBS signatures (26, 73). Moreover, given the functional implication of the RB/E2F pathway with proliferation, it is likely that A3B expression is regulated in a fashion similar to many cell-cycle genes. Multiple lines of evidence have indeed classified A3B as a gene that associates strongly with cell-cycle progression and proliferation in cancer cells (26, 36, 74).

Thus, in contrast to A3A, A3B is readily induced by therapeutic agents. In fact, the occurrence of treatment-induced mutations is relatively well documented and predominantly showcases the direct induction of mutations by cancer drugs, including cis-platin (43), and induction of APOBEC mutagenesis has been documented comparing gliomas before and after irradiation (75). However, it would be insightful, although challenging to control, to investigate the contribution of treatment-induced A3B on the total APOBEC signature load of patients with breast cancer. A minor contribution of treatment-induced A3B activity is to be expected and might, on a background level, amplify the impact of A3B on disease progression (discussed below).

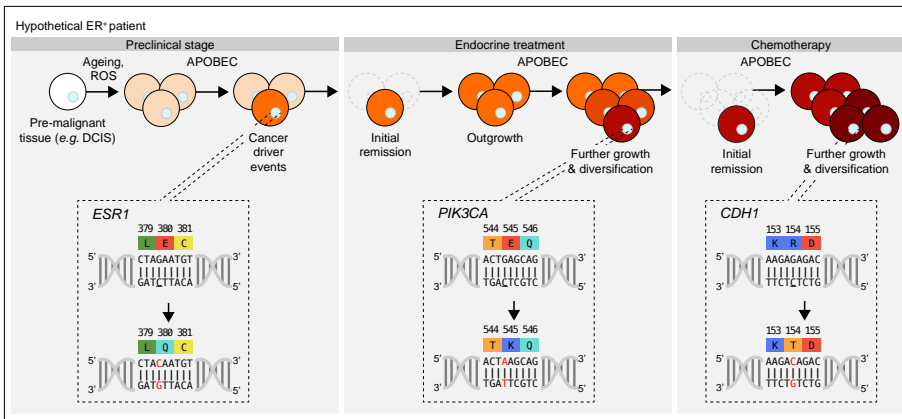
## **APOBEC3-Mediated Mutagenesis and Disease Trajectory**

Recent tumor sequencing efforts have revealed the genetically heterogeneous and evolving nature of tumor cells, changing their genetic makeup during different cancer stages or when facing anticancer treatments (1, 76, 77). Specifically,

APOBEC3-mediated mutagenesis has been shown to, in varying degrees, influence the mutational landscape of premalignant breast lesions, primary disease, and metastatic breast cancer.

### APOBEC3 activity during premalignancy

Although samples sizes remain limited, APOBEC3 expression and/or APOBEC3-mediated mutagenesis is found in about 8% of ductal carcinoma *in situ* (DCIS) samples and is detectable in approximately 16% of specimens when DCIS is associated with invasive disease (49, 78–80). However, although APOBEC3-mediated mutagenesis is appreciable in DCIS, no clear evidence of APOBEC3-mediated mutagenesis toward driver genes has been found (79, 80). This indicates that although APOBEC3-mediated mutagenesis can influence the cellular genome at the precursor stage, this frequently takes place before clonal selection overtakes the overall genomic makeup (Fig. 3). Therefore, the overall impact on tumor evolution is yet to be fully determined.



**Figure 3.** APOBEC3-mediated mutagenesis and disease trajectory. Clinical progression of a hypothetical ER<sup>+</sup> patient from preclinical stage (initial outgrowth and initial cancer driver events), to first- and second-line anticancer treatments. In this example, APOBEC3 mediated mutation of ESR1 is the first driver mutation. Mutations in PIK3CA provide the tumor with resistance to adjuvant endocrine treatment. Although remission is obtained, further growth and APOBEC3-mediated diversification occurs. Second-line chemotherapeutics provided temporary remission, but further APOBEC3-mediated mutagenesis affected genes involved in metastatic behavior, here exemplified with CDH1, leading to treatment failure. The cytosine targeted by APOBEC3 proteins is underlined.

### APOBEC3 activity throughout malignant disease

APOBEC3-mediated deamination is actively involved in tumor evolution in early and advanced breast cancer (2, 43, 55, 77, 81, 82), adding novel “branches” to the cancer evolutionary “tree” that may unfavorably impact disease trajectory. For instance, in

hypermutated breast tumors, APOBEC3 activity can account for almost two-thirds of the total mutational burden (43, 83). Indeed, mutations in approximately 25% of cancer driver genes occur within the preferred 50-TCW APOBEC context (see Fig. 3; Table 1 for examples).

**Table 1.** Nonexhaustive list of APOBEC-associated mutations and clinical associations in breast cancer driver genes.

Gene	Mutations	Trinucleotide context	Pathogenicity score (FATHMM)	Hotspot (Yes/No)	Clinical associations	References
<i>PIK3CA</i>	E453K	TCT > TTT	0.99	No	Loss of normal <i>PIK3CA</i> function has been recognized as a contributing factor to the acquired resistance to chemotherapy and endocrine therapy. Multi-hit mutations in <i>PIK3CA</i> are also associated with APOBEC activity.	(77, 84, 85, 93, 121)
	E542K	TCA > TTA	0.97	Yes		
	E545K	TCA > TTA	0.97	Yes		
	E545Q	TCA > TGA	0.98	Yes		
	E726K	TCA > TTA	0.99	No		
<i>SPEN</i>	E970K	TCT > TTT	0.99	No	Associated with acquired resistance to endocrine treatment	(93, 94)
	E2151K	TCT > TTT	N/A	No		
<i>ESR1</i>	E380Q	TCT > TGT	1.00	Yes	Associated with acquired resistance to endocrine treatment	(93, 95)
<i>NF1</i>	Q554*	TCA > TTA	0.99	No	Confers resistance to endocrine treatment, possibly by enabling cell-cycle progression overdrive	(90, 122)
	Q1218*	TCA > TTA	0.98	No		
	Q1399*	TCA > TTA	0.98	No		
	Q2234*	TCA > TTA	1.00	No		
<i>CDH1</i>	R154T	TCT > TGT	0.98	No	Associated with cellular discohesion and hyperplasia	(77, 92)
<i>ARID1A</i>	SNNN*	Not reported	Various	No	Associated with acquired resistance to endocrine treatment and lowered PFS	(81)
<i>KMT2C</i>	QNNN*	Not reported	Various	No		

Abbreviation: FATHMM, Functional Analysis Through Hidden Markov Models (123).

As driver mutations occur predominantly in the early stages of tumor evolution, this indicates that APOBEC3-mediated mutagenesis provides mutagenic fuel during the early stages of breast cancer (77, 81). However, APOBEC3-mediated alterations of driver genes can also occur as late events (Fig. 3; refs. 43, 77, 81). Notable APOBEC3-associated driver mutations are the E542K and E545K hotspot mutations in *PIK3CA*, the second most frequently altered breast cancer driver gene (2). These mutations account for  $\geq 1/3$ rd of the *PIK3CA* single-point mutations in breast cancer and are thought to predominantly occur as early events (29, 41, 81). Tumors can also carry multiple APOBEC-associated *PIK3CA* mutations which are enriched in metastatic breast cancer as compared with primary breast cancer (81). This *cis-PIK3CA* mutational genotype provides enhanced downstream signaling, associates with lower progression-free survival, and has been recognized as a contributing factor to acquired treatment resistance (41, 84–86).

Other genes affected by APOBEC3-mediated mutagenesis include, among others, *KMT2C* and *ARID1A* (77), which exhibit widespread, non-hotspot-truncating S>X and Q>X mutations in an APOBEC context (81). Importantly, loss of *KMT2C* is associated with resistance to endocrine therapy, while *ARID1A* mutations may



confer resistance to both endocrine therapy and chemotherapy (87–89). Moreover, the emergence of truncating *NF1* mutations bearing APOBEC SBS signatures may also occur during endocrine therapy (90). Loss of *NF1* has been shown to confer resistance to endocrine treatment, possibly by enabling cell-cycle progression overdrive (91). An APOBEC-associated mutation in the tumor suppressor *CDH1*, unique to the metastatic tumor, grew to dominance during chemotherapy (77). Pathogenic mutations in *CDH1* have been associated with cellular decohesion and hyperplasia and contribute to lobular breast cancer (reviewed in ref. 92). Finally, postmortem sequencing of metastases of endocrine-resistant ER. breast cancer revealed novel acquired APOBEC-associated mutations in *SPEN* and *ESR1* (93), genes that have been associated with acquired resistance to endocrine treatment (94, 95). Importantly, APOBEC-associated mutations in almost all the aforementioned driver genes (*PIK3CA*, *KMT2C*, *ARID1A*, *NF1*, and *CDH1*) have recently been shown to be enriched in metastatic breast cancer, strongly emphasizing their relevance in cancer development (81). Other genes carrying APOBEC-associated mutations in these studies were, amongst others, the tumor suppressors *MAP3K1*, *TP53*, and *ZFHX3*.

## Leveraging APOBEC3 Activity for Clinical Benefit

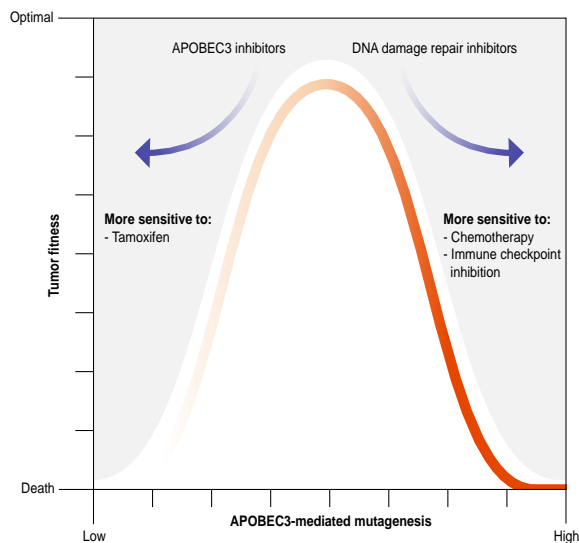
### APOBEC3 as a prognostic biomarker

Considering the active contribution of APOBEC3-mediated mutagenesis to disease trajectory, it can be expected that APOBEC3 expression and/or APOBEC SBS signatures can serve as prognostic biomarkers in breast cancer. Indeed, in ER. breast cancer, high expression of *A3B* correlates with unfavorable clinical parameters, including disease-free survival, metastasis-free survival, and overall survival (47). Although prognostic studies are rare as these have to be performed in the absence of systemic treatment to distinguish them from predictive biomarkers, the independent nature of these findings emphasizes the suitability of *A3B* as a prognostic marker in ER<sup>+</sup> breast cancer.

### APOBEC3 as a predictive biomarker

The role of APOBEC3 expression and/or APOBEC SBS signatures as predictive biomarkers has become increasingly established over recent years. Increased *A3B* expression in breast cancer has been strongly associated with treatment failure of adjuvant endocrine drugs (48), implying that commonly used endocrine drugs (such as tamoxifen) are less suitable for APOBEC3-positive tumors (Fig. 4). Conversely, in breast cancer as well as ovarian carcinoma and bladder cancer, APOBEC3-mediated mutagenesis was found to predict beneficial treatment response (37, 54, 96, 97).

Although diverse treatment regimens were used in these studies, all applied DNA-intercalating agents. These observations suggest that APOBEC3-mediated mutagenesis and DNA intercalators combine to exert synergistic levels of DNA damage during breast cancer treatment (Fig. 4). Overall, APOBEC3-mediated mutagenesis is a predictive biomarker for response to both endocrine therapy and chemotherapy. Of note, recent work in lung cancer suggests A3A and A3B may also contribute to resistance to targeted therapies (bioRxiv 2021.01.20.426852v1; bioRxiv 2020.12.18.423280v2).



**Figure 4.** Clinical implications of APOBEC3-mediated mutagenesis. Overview depicting the relationship between APOBEC3-mediated mutagenesis (X-axis) and tumor fitness (Y-axis) as a function of suitable anticancer treatments. Standard-of-care therapies, such as tamoxifen and radiotherapy, are suitable when APOBEC3 activity is low, driving down tumor fitness. On the other hand, ICI, synergistic drugs (such as DNA-intercalating agents), and synthetically lethal combinations (such as ATRi) can exploit the weaknesses brought about by APOBEC3-mediated mutagenesis. Inhibitors of A3A and/or A3B, currently in development, may also be used to limit APOBEC3-mediated genomic diversification and sensitize to other treatments.

### **APOBEC3 as a predictor for immune checkpoint inhibition response**

Cancer growth relies on a disturbance in the balance between detection and subsequent elimination of cancer cells by immune cells and, conversely, the escape of cancer cells from immune cells. At the core of this interaction are antigen-presenting dendritic cells (DC) and CD4<sup>+</sup> or CD8<sup>+</sup> T cells. DCs are innate immune cells specialized in recognizing neoantigens, which are proteins released by tumor cells that contain nonautologous antigens as a result from somatic mutations (98).

Upon recognition, these neoantigens are used to prime naïve CD4<sup>+</sup> and CD8<sup>+</sup> T cells, which infiltrate tumor tissue and eliminate tumor cells displaying these neoantigens (reviewed in ref. 99). Because of the highly plastic nature of tumors, cancer cells need to tip the scale in their favor to avoid cytotoxic elimination. Proteins expressed by tumor cells, such as PD-1, can interact with inhibitory ligands expressed by T cells, such as PD-L1, initiating a shift toward immune tolerance. This interaction, called an immune checkpoint, forms the basis for immune checkpoint inhibition (ICI), which seeks to enhance immunogenic tumor cell killing by using antibodies against key immune checkpoint proteins (Fig. 4, and reviewed in refs. 100, 101).

Given the involvement of APOBEC activity with shaping the tumor genome, and thereby the antigen repertoire, its suitability to predict ICI response in breast cancer has become subject of investigation. In a recent study, murine breast cancer cells that normally do not possess the *A3B* gene were engineered to express *A3B* and orthotopically injected (102). Interestingly, expression of *A3B* alone already significantly slowed tumor growth as compared with cells devoid of *A3B*. This partial inhibition of tumor growth was dependent on CD4<sup>+</sup> and CD8<sup>+</sup> immune cells, suggesting that at least some cytotoxic tumor cell killing was achieved. *A3B* expression also promoted tumor infiltration by T cells that were likely primed with APOBEC3-induced neoantigens. Strikingly, when combined with ICI, potent and sustained growth inhibition was achieved in *A3B*-expressing cells, but not control cells.

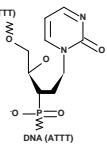
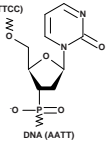
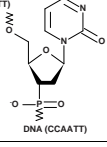
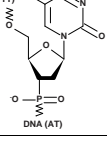
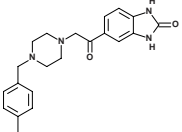
Increased APOBEC3-mediated mutagenesis has also been associated with immunoactivation in human breast cancer. For instance, first, indicators of APOBEC3-mediated mutagenesis such as increased APOBEC SBS signatures and expression of *A3A* or *A3B* often correlate positively with infiltration of tumor tissue by immune cells, including DCs and CD8<sup>+</sup> T cells (102–105). Second, the same indicators of APOBEC3-mediated mutagenesis correlate positively with expression of PD-1 and PD-L1 (106). Third, in several studies—although with a limited number of patients with breast cancer—the presence of APOBEC SBS signatures significantly improved the chance of ICI response (83, 106–108). Combined, these findings suggest that APOBEC3 activity can function as a powerful predictor of ICI responsiveness (Fig. 4), which merits further investigation with larger cohorts of patients with breast cancer.

## Synthetically lethal interactions with APOBEC3-mediated mutagenesis

In addition to somatic C-to-T and C-to-G mutations, the activity of APOBEC3 enzymes can also induce DNA double-strand breaks and replication stress (17, 40, 109–111). To counteract these pressures, and to ensure survival, cancer cells that display APOBEC3 activity increasingly rely on DNA damage repair (109–111). This vulnerability has led to a number of synthetic lethality approaches that target DNA damage repair in A3A- or A3B-expressing cancer cells (Fig. 4). Cells with high APOBEC3 activity are exceptionally vulnerable to inhibition of ATR, an important DNA damage checkpoint (29, 111). There are currently multiple ATR inhibitors being evaluated for clinical use (112). Furthermore, inhibition of at least three major repair factors involved in the resolution or neutralization of deaminated lesions, including UNG, HMCES, and APE1, show similar synthetic lethal phenotypes (109, 113–115). Altogether, these synthetic lethality models represent a rational design to systematically attack the vulnerabilities of cancers that show ongoing APOBEC3 mutagenesis and warrant further research into their suitability in breast cancer.

## Dampening APOBEC3 activity using inhibitors

Efforts to inhibit A3A and A3B enzymes have chiefly relied on the design of chemical inhibitors and, so far, revolve around two molecular classes (Fig. 5). One such approach exploits the trinucleotide preference of the Z1 domains of A3A and A3B, and features a chemically modified cytosine, called dZ, in an oligo-based substrate. These efforts yielded promising substrate-like inhibitors within the low micromolar range in *in vitro* assays (116–119). In addition, another recent approach identified several candidate small molecules (as opposed to substrate-like inhibitors), with comparable *in vitro* effectiveness within the low micromolar range (120). After further characterization *in vitro*, APOBEC3 inhibitors can subsequently be investigated in clinically relevant preclinical platforms, as relevant genetically engineered mouse models for A3A and A3B have become available recently (31, 102). Future studies should aim to determine possible systemic toxicity of candidate inhibitors and their ability to lower the accumulation of APOBEC3 SBS signatures in murine cancers. Moreover, considering the established role of both these enzymes in the development of (breast) cancer, further research into APOBEC3 inhibition should stay focused on the dual inhibition of both A3A and A3B. Ultimately, APOBEC3 inhibitors should be investigated as synergistic treatments in conjunction with existing anti-breast cancer therapies, including surgery and targeted treatments based on genetic markers (Fig. 4).

Abbreviation	Inhibitor class	Molecular structure	Binding, inhibitory and dissociative properties	References
TdZA-Oligo	Oligo-based substrate	<p>DNA (ATTT)</p>  <p>DNA (ATTT)</p>	<p>vs. A3A <math>K_i = 7.5 \pm 1.7 \mu\text{M}</math>, as per NMR assay</p> <p>vs. A3B <math>K_i = 11.4 \pm 2.6 \mu\text{M}</math>, as per NMR assay</p> <p>vs. A3B <math>K_i = 5.5 \pm 0.6 \mu\text{M}</math>, as per ITC assay</p>	(117)
CCdZ-Oligo	Oligo-based substrate	<p>DNA (ATTCC)</p>  <p>DNA (AATT)</p>	<p>vs. A3A <math>K_i = 12.3 \pm 1.8 \mu\text{M}</math>, as per NMR assay</p> <p>vs. A3A <math>K_i = 0.97 \pm 0.15 \mu\text{M}</math>, as per ITC assay</p> <p>vs. A3B, not an inhibitor</p>	(116) (117)
dZCC-Oligo	Oligo-based substrate	<p>DNA (ATT)</p>  <p>DNA (CCAATT)</p>	<p>vs. A3A, not an inhibitor</p> <p>vs. A3B <math>K_i = 18.5 \pm 5.5 \mu\text{M}</math>, as per NMR assay</p>	(116)
5FdZ-Oligo	Oligo-based substrate	<p>DNA (TTTT)</p>  <p>DNA (AT)</p>	<p>vs. A3A: <math>\text{IC}_{50} = 0.16 \pm 0.01 \mu\text{M}</math>, as per oligo cleavage assay</p>	(118)
C8.5	Small molecule inhibitor		<p>vs. A3A: <math>\text{IC}_{50} = 9 \mu\text{M}</math>, as per oligo cleavage assay</p> <p>vs. A3B: <math>\text{IC}_{50} = 40 \mu\text{M}</math>, as per oligo cleavage assay</p>	(120)

**Figure 5.** APOBEC3A and APOBEC3B inhibitors. Published inhibitors of APOBEC3A and APOBEC3B, their molecular structures and pharmacologic properties.

## Conclusions

Over the past decade, the perspective on APOBEC3 enzymes, specifically A3A and A3B, has drastically shifted from beneficial members of the innate system to direct influencers of cancer development and disease trajectory most notably in breast cancer. A collection of cellular pathways, including the PKC/ $\text{ncNF}\kappa\text{B}$ , the RB/ E2F pathway, and IFN signaling relay proliferative and inflammatory signals that stimulate expression of A3A and/or A3B. The mutagenic activities of APOBEC3 proteins can now be traced using high throughput sequencing approaches, implicating them with genomic alterations that stand in direct association with treatment response. They also provide prognostic and predictive value and reveal potential cancer weak spots. Furthermore, the promising characteristics of potential APOBEC3 inhibitors merit further investigation and may be instrumental in restricting the mutagenic arsenal of cancer cells.

5

## Authors' Disclosures

J.W.M. Martens reports grants from The Dutch Cancer Society during the conduct of the study as well as personal fees from Novartis; other support from Roche; and grants from The Dutch Cancer Society, The Dutch Organisation of Scientific Research (NWO), Breast Cancer Now, Pfizer, GSK, MLDS, Oncode, Eurostar, and ERC Advanced outside the submitted work. No disclosures were reported by the other authors.

## Acknowledgments

We want to extend our gratitude to Wilke Castelijns, MSc, for helpful suggestions during the early conceptualization phase of this review. We also thank Bojana Stefanovska, PhD, and Anya Normandeau, BA, for helpful feedback. We thank Dr. Prokopios Argyris for the IHC image of the breast cancer specimen in Figure 1.

This work was supported, in part, by the KWF Dutch Cancer Society (KWF10270, to J.W.M. Martens, P.N. Span, and R.S. Harris) and by the NCI (P01-CA234228, to R.S. Harris). R.S. Harris is the Ewing Halsell President's Council Distinguished Chair, a CPRIT Scholar, and an Investigator of the Howard Hughes Medical Institute at University of Texas Health San Antonio. None of the funding agencies had any role in conceptualization, study design, data collection, interpretation of results, or the decision to submit this work for publication.

## References

1. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. *Nature* 2020;578:94–101.
2. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* 2016;534:47–54.
3. Periyasamy M, Patel H, Lai CF, Nguyen VTM, Nevedomskaya E, Harrod A, et al. APOBEC3B-mediated cytidine deamination is required for estrogen receptor action in breast cancer. *Cell Rep* 2015;13:108–21.
4. Harris RS, Dudley JP. APOBECs and virus restriction. *Virology* 2015;479–480:131–45.
5. Ito J, Gifford RJ, Sato K. Retroviruses drive the rapid evolution of mammalian APOBEC3 genes. *Proc Natl Acad Sci U S A* 2020;117:610–8.
6. Wang X, AbuduA, Son S, Dang Y, Venta PJ, Zheng YH, et al. Analysis of human APOBEC3H haplotypes and anti-human immunodeficiency virus type 1 activity. *J Virol* 2011;85:3142–52.
7. Salamango DJ, McCann JL, Demir O, Brown WL, Amaro RE, Harris RS, et al. APOBEC3B nuclear localization requires two distinct N-terminal domain surfaces. *J Mol Biol* 2018;430:2695–708.
8. Navarro F, Bollman B, Chen H, Konig R, Yu Q, Chiles K, et al. Complementary function of the two catalytic domains of APOBEC3G. *Virology* 2005;333:374–86.
9. Shi K, Demir O, Carpenter MA, Wagner J, Kurahashi K, Harris RS, et al. Conformational switch regulates the DNA cytosine deaminase activity of human APOBEC3B. *Sci Rep* 2017;7:17415.
10. Anderson BD, Ikeda T, Moghadasi SA, Martin AS, Brown WL, Harris RS, et al. Natural APOBEC3C variants can elicit differential HIV-1 restriction activity. *Retrovirology* 2018;15:78.
11. Weisblum Y, Oiknine-Djian E, Zakay-Rones Z, Vorontsov O, Haimov-Kochman R, Nevo Y, et al. APOBEC3A is upregulated by human cytomegalovirus (HCMV) in the maternal-fetal interface, acting as an innate anti-HCMV effector. *J Virol* 2017;91:e01296 17.
12. Refsland EW, Stenglein MD, Shindo K, Albin JS, Brown WL, Harris RS, et al. Quantitative profiling of the full APOBEC3 mRNA repertoire in lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res* 2010;38:4274–84.
13. Bishop KN, Holmes RK, Sheehy AM, Davidson NO, Cho SJ, Malim MH, et al. Cytidine deamination of retroviral DNA by diverse APOBEC proteins. *Curr Biol* 2004;14:1392–6.
14. Hultquist JF, Harris RS. Leveraging APOBEC3 proteins to alter the HIV mutation rate and combat AIDS. *Future Virol* 2009;4:605.
15. Suspene R, Guetard D, Henry M, Sommer P, Wain-Hobson S, Vartanian JP, et al. Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases *in vitro* and *in vivo*. *Proc Natl Acad Sci U S A* 2005;102:8321–6.
16. Cheng AZ, Yockteng-Melgar J, Jarvis MC, Malik-Soni N, Borozaan I, Carpenter MA, et al. Epstein-Barr virus BORF2 inhibits cellular APOBEC3B to preserve viral genome integrity. *Nat Microbiol* 2019;4:78–88.
17. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, et al. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 2013;494:366–70.
18. 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;149:979–93.
19. 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;45:970–6.

20. Petljak M, Alexandrov LB, Brummendorf JS, Price S, Wedge DC, Grossmann S, et al. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. *Cell* 2019;176:1282–94.
21. Nik-Zainal S, Morganella S. Mutational signatures in breast cancer: the problem at the DNA level. *Clin Cancer Res* 2017;23:2617–29.
22. Starrett GJ, Luengas EM, McCann JL, Ebrahimi D, Temiz NA, Love RP, et al. The DNA cytosine deaminase APOBEC3H haplotype I likely contributes to breast and lung cancer mutagenesis. *Nat Commun* 2016;7:12918.
23. Petljak M, Dananberg A, Chu K, Bergstrom EN, Striemen J, von Morgen P, et al. Mechanisms of APOBEC3 mutagenesis in human cancer cells. *Nature* 2022;607:799–807.
24. Lackey L, Law EK, Brown WL, Harris RS. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell Cycle* 2013;12:762–72.
25. McCann JL, Klein MM, Leland EM, Law EK, Brown WL, Salamango DJ, et al. The DNA deaminase APOBEC3B interacts with the cell-cycle protein CDK4 and disrupts CDK4-mediated nuclear import of cyclin D1. *J Biol Chem* 2019;294:12099–111.
26. Roelofs PA, Goh CY, Chua BH, Jarvis MC, Stewart TA, McCann JL, et al. Characterization of the mechanism by which the RB/E2F pathway controls expression of the cancer genomic DNA deaminase APOBEC3B. *Elife* 2020;9:e61287.
27. Kidd JM, Newman TL, Tuzun E, Kaul R, Eichler EE. Population stratification of a common APOBEC gene deletion polymorphism. *PLoS Genet* 2007;3:e63.
28. 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;46:487–91.
29. Buisson R, Langenbucher A, Bowen D, Kwan EE, Benes CH, Zou L, et al. Passenger hotspot mutations in cancer driven by APOBEC3A and mesoscale genomic features. *Science* 2019;364:eaaw2872.
30. 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.
31. 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:e20200261.
32. Hoopes JJ, Cortez LM, Mertz TM, Malc EP, Mieczkowski PA, Roberts SA, et al. APOBEC3A and APOBEC3B preferentially deaminate the lagging strand template during DNA replication. *Cell Rep* 2016;14:1273–82.
33. 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;47:1067–72.
34. DeWeerd RA, Nemeth E, Poti A, Petryk N, Chen CL, Hyrien O, et al. Prospectively defined patterns of APOBEC3A mutagenesis are prevalent in human cancers. *Cell Rep* 2022;38:110555.
35. Brown WL, Law EK, Argyris PP, Carpenter MA, Levin-Klein R, Ranum AN, et al. A rabbit monoclonal antibody against the antiviral and cancer genomic DNA mutating enzyme APOBEC3B. *Antibodies* 2019;8:47.
36. Argyris PP, Wilkinson PE, Jarvis MC, Magliocca KR, Patel MR, Vogel RI, et al. Endogenous APOBEC3B overexpression characterizes HPV-positive and HPV-negative oral epithelial dysplasias and head and neck cancers. *Mod Pathol* 2021;34:280–90.



37. Serebrenik AA, Argyris PP, Jarvis MC, Brown WL, Bazzaro M, Vogel RI, et al. The DNA cytosine deaminase APOBEC3B is a molecular determinant of platinum responsiveness in clear cell ovarian cancer. *Clin Cancer Res* 2020;26:3397–407.
38. Leonard B, McCann JL, Starrett GJ, Kosyakovsky L, Luengas EM, Molan AM, et al. The PKC/NF-kappaB signaling pathway induces APOBEC3B expression in multiple human cancers. *Cancer Res* 2015;75:4538–47.
39. Lee JY, Schizas M, Geyer FC, Selenica P, Piscuoglio S, Sakr RA, et al. Lobular carcinomas in situ display intralesion genetic heterogeneity and clonal evolution in the progression to invasive lobular carcinoma. *Clin Cancer Res* 2019;25:674–86.
40. Venkatesan S, Angelova M, Puttick C, Zhai H, Caswell DR, Lu W-T, et al. Induction of APOBEC3 exacerbates DNA replication stress and chromosomal instability in early breast and lung cancer evolution. *Cancer Discov* 2021;11:2456–73.
41. Kingston B, Cutts RJ, Bye H, Beaney M, Walsh-Crestani G, Hrebien S, et al. Genomic profile of advanced breast cancer in circulating tumour DNA. *Nat Commun* 2021;12:2423.
42. Pitt JJ, Riester M, Zheng Y, Yoshimatsu TF, Sanni A, Oluwasola O, et al. Characterization of Nigerian breast cancer reveals prevalent homologous recombination deficiency and aggressive molecular features. *Nat Commun* 2018;9:4181.
43. Angus L, Smid M, Wilting SM, van Riet J, Van Hoeck A, Nguyen L, et al. The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nat Genet* 2019;51:1450–8.
44. Jalili P, Bowen D, Langenbucher A, Park S, Aguirre K, Corcoran RB, et al. Quantification of ongoing APOBEC3A activity in tumor cells by monitoring RNA editing at hotspots. *Nat Commun* 2020;11:2971.
45. Carpenter MA, Li M, Rathore A, Lackey L, Law EK, Land AM, et al. Methylcytosine and normal cytosine deamination by the foreign DNA restriction enzyme APOBEC3A. *J Biol Chem* 2012;287:34801–8.
46. Sieuwerts AM, Schrijver WA, Dalm SU, deWeerd V, Moelans CB, TerHoeve N, et al. Progressive APOBEC3B mRNA expression in distant breast cancer metastases. *PLoS One* 2017;12:e0171343.
47. Sieuwerts AM, Willis S, Burns MB, Look MP, Meijer-Van Gelder ME, Schlicker A, et al. Elevated APOBEC3B correlates with poor outcomes for estrogenreceptor-positive breast cancers. *Horm Cancer* 2014;5:405–13.
48. Law EK, Sieuwerts AM, LaPara K, Leonard B, Starrett GJ, Molan AM, et al. The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. *Sci Adv* 2016;2:e1601737.
49. Sieuwerts AM, Doebar SC, de Weerd V, Verhoef EI, Beauford CM, Agahozo MC, et al. APOBEC3B gene expression in ductal carcinoma in situ and synchronous invasive breast cancer. *Cancers* 2019;11:1062.
50. Lawson JS, Glenn WK. Catching viral breast cancer. *Infect Agent Cancer* 2021; 16:37.
51. Baker SC, Mason AS, Slip RG, Skinner KT, Macdonald A, Masood O, et al. Induction of APOBEC3-mediated genomic damage in urothelium implicates BK polyomavirus (BKPyV) as a hit-and-run driver for bladder cancer. *Oncogene* 2022;41:2139–51.
52. Oh S, Bournique E, Bowen D, Jalili P, Sanchez A, Ward I, et al. Genotoxic stress and viral infection induce transient expression of APOBEC3A and proinflammatory genes through two distinct pathways. *Nat Commun* 2021;12:4917.
53. Kondo S, Wakae K, Wakisaka N, Nakanishi Y, Ishikawa K, Komori T, et al. APOBEC3A associates with human papillomavirus genome integration in oropharyngeal cancers. *Oncogene* 2017;36:1687–97.

54. Middlebrooks CD, Banday AR, Matsuda K, Udquim KI, Onabajo OO, Paquin A, et al. Association of germline variants in the APOBEC3 region with cancer risk and enrichment with APOBEC-signature mutations in tumors. *Nat Genet* 2016;48:1330–8.
55. Roper N, Gao S, Maity TK, Banday AR, Zhang X, Venugopalan A, et al. APOBEC mutagenesis and copy-number alterations are drivers of proteogenomic tumor evolution and heterogeneity in metastatic thoracic tumors. *Cell Rep* 2019;26:2651–66.
56. Siriwardena SU, Perera MLW, Senevirathne V, Stewart J, Bhagwat AS. A tumor-promoting phorbol ester causes a large increase in APOBEC3A expression and a moderate increase in APOBEC3B expression in a normal human keratinocyte cell line without increasing genomic uracils. *Mol Cell Biol* 2019;39:e00238–18.
57. Saldivar JC, Cortez D, Cimprich KA. The essential kinase ATR: ensuring faithful duplication of a challenging genome. *Nat Rev Mol Cell Biol* 2017;18:622–36.
58. Madsen P, Anant S, Rasmussen HH, Gromov P, Vorum H, Dumanski JP, et al. Psoriasis upregulated phorbolin-1 shares structural but not functional similarity to the mRNA-editing protein apobec-1. *J Invest Dermatol* 1999;113:162–9.
59. Periyasamy M, Singh AK, Gemma C, Farzan R, Allsopp RC, Shaw JA, et al. Induction of APOBEC3B expression by chemotherapy drugs is mediated by DNA-PK-directed activation of NF- $\kappa$ B. *Oncogene* 2021;40:1077–90.
60. Lin L, Holmes B, Shen MW, Kammeron D, Geijsen N, Gifford DK, et al. Comprehensive mapping of key regulatory networks that drive oncogene expression. *Cell Rep* 2020;33:108426.
61. Liu T, Zhang L, Joo D, Sun SC. NF- $\kappa$ B signaling in inflammation. *Signal Transduct Target Ther* 2017;2:17023.
62. Liu W, Wu J, Yang F, Ma L, Ni C, Hou X, et al. Genetic polymorphisms predisposing the interleukin 6-induced APOBEC3B-UNG imbalance increase HCC risk via promoting the generation of APOBEC-signature HBV mutations. *Clin Cancer Res* 2019;25:5525–36.
63. Habraken Y, Piette J. NF-kappaB activation by double-strand breaks. *Biochem Pharmacol* 2006;72:1132–41.
64. Kanu N, Cerone MA, Goh G, Zalmas LP, Bartkova J, Dietzen M, et al. DNA replication stress mediates APOBEC3 family mutagenesis in breast cancer. *Genome Biol* 2016;17:185.
65. Saito Y, Miura H, Takahashi N, Kuwahara Y, Yamamoto Y, Fukumoto M, et al. Involvement of APOBEC3B in mutation induction by irradiation. *J Radiat Res* 2020;61:819–27.
66. Yamazaki H, Shirakawa K, Matsumoto T, Kazuma Y, Matsui H, Horisawa Y, et al. APOBEC3B reporter myeloma cell lines identify DNA damage response pathways leading to APOBEC3B expression. *PLoS One* 2020;15:e0223463.
67. Mori S, Takeuchi T, Ishii Y, Yugawa T, Kiyono T, Nishina H, et al. Human papillomavirus 16 E6 upregulates APOBEC3B via the TEAD transcription factor. *J Virol* 2017;91:e02413–16.
68. PeriyasamyM, Singh AK, Gemma C, Kranjec C, Farzan R, Leach DA, et al. p53 controls expression of the DNA deaminase APOBEC3B to limit its potential mutagenic activity in cancer cells. *Nucleic Acids Res* 2017;45:11056–69.
69. Vieira VC, Leonard B, White EA, Starrett GJ, Temiz NA, Lorenz LD, et al. Human papillomavirus E6 triggers upregulation of the antiviral and cancer genomic DNA deaminase APOBEC3B. *mBio* 2014;5:e02234–14.
70. Warren CJ, Van Doorslaer K, Pandey A, Espinosa JM, Pyeon D. Role of the host restriction factor APOBEC3 on papillomavirus evolution. *Virus Evol* 2015;1:vev015.

71. Starrett GJ, Serebrenik AA, Roelofs PA, McCann JL, Verhalen B, Jarvis MC, et al. Polyomavirus T antigen induces APOBEC3B expression using an LXCXE-Dependent and TP53-independent mechanism. *mBio* 2019;10:e02690–18.
72. Zapatka M, Borozan I, Brewer DS, Iskar M, Grundhoff A, Alawi M, et al. The landscape of viral associations in human cancers. *Nat Genet* 2020;52:320–30.
73. Cancer GENOME Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* 2012;490:61–70.
74. Hirabayashi S, Shirakawa K, Horisawa Y, Matsumoto T, Matsui H, Yamazaki H, et al. APOBEC3B is preferentially expressed at the G<sub>2</sub>-M phase of cell cycle. *Biochem Biophys Res Commun* 2021;546:178–84.
75. Kocakavuk E, Anderson KJ, Varn FS, Johnson KC, Amin SB, Sulman EP, et al. Radiotherapy is associated with a deletion signature that contributes to poor outcomes in patients with cancer. *Nat Genet* 2021;53:1088–96.
76. Brady SW, McQuerry JA, Qiao Y, Piccolo SR, Shrestha G, Jenkins DF, et al. Combating subclonal evolution of resistant cancer phenotypes. *Nat Commun* 2017;8:1231.
77. McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C, et al. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci Transl Med* 2015;7:283ra54.
78. Nachmansohn D, Officer A, Mori H, Gordon J, Evans MF, Steward J, et al. The breast pre-cancer atlas illustrates the molecular and micro-environmental diversity of ductal carcinoma in situ. *NPJ Breast Cancer* 2022;8:6.
79. Pareja F, Brown DN, Lee JY, Da Cruz Paula A, Selenica P, Bi R, et al. Whole exome sequencing analysis of the progression from non-low-grade ductal carcinoma in situ to invasive ductal carcinoma. *Clin Cancer Res* 2020;26: 3682–93.
80. Pang JB, Savas P, Fellowes AP, Mir Arnau G, Kader T, Vedururu R, et al. Breast ductal carcinoma in situ carry mutational driver events representative of invasive breast cancer. *Mod Pathol* 2017;30:952–63.
81. Bos MK, Smid M, Sleijfer S, Martens JWM. Apolipoprotein B mRNA-editing catalytic polypeptide-like-induced protein changes in estrogen receptor-positive, human epidermal growth factor receptor 2-negative breast cancer throughout disease progression. *JCO Precis Oncol* 2022;6:e2100190.
82. Yates LR, Knappskog S, Wedge D, Farmery JHR, Gonzalez S, Martincorena I, et al. Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell* 2017;32:169–84.
83. Barroso-Sousa R, Jain E, Cohen O, Kim D, Buendia-Buendia J, Winer E, et al. Prevalence and mutational determinants of high tumor mutation burden in breast cancer. *Ann Oncol* 2020;31:387–94.
84. Miller TW, Balko JM, Arteaga CL. Phosphatidylinositol 3-kinase and antiestrogen resistance in breast cancer. *J Clin Oncol* 2011;29:4452–61.
85. Mosele F, Stefanovska B, Lusque A, Tran Dien A, Garberis I, Droin N, et al. Outcome and molecular landscape of patients with PIK3CA-mutated metastatic breast cancer. *Ann Oncol* 2020;31:377–86.
86. Vasan N, Razavi P, Johnson JL, Shao H, Shah H, Antoine A, et al. Double PIK3CA mutations in cis increase oncogenicity and sensitivity to PI3Kalpha inhibitors. *Science* 2019;366:714–23.
87. Gala K, Li Q, Sinha A, Razavi P, Dorso M, Sanchez-Vega F, et al. KMT2C mediates the estrogen dependence of breast cancer through regulation of ERalpha enhancer function. *Oncogene* 2018;37:4692–710.

88. Hanka AB, Sudhan DR, Arteaga CL. Overcoming endocrine resistance in breast cancer. *Cancer Cell* 2020;37:496–513.
89. Xu G, Chhangawala S, Cocco E, Razavi P, Cai Y, Otto JE, et al. ARID1A determines luminal identity and therapeutic response in estrogen-receptor-positive breast cancer. *Nat Genet* 2020;52:198–207.
90. Sokol ES, Feng YX, Jin DX, Basudan A, Lee AV, Atkinson JM, et al. Loss of function of NF1 is a mechanism of acquired resistance to endocrine therapy in lobular breast cancer. *Ann Oncol* 2019;30:115–23.
91. Pearson A, Proszek P, Pascual J, Fribbens C, Shamsheer MK, Kingston B, et al. Inactivating NF1 mutations are enriched in advanced breast cancer and contribute to endocrine therapy resistance. *Clin Cancer Res* 2020;26:608–22.
92. Bruner HC, Derksen PWB. Loss of E-cadherin-dependent cell-cell adhesion and the development and progression of cancer. *Cold SpringHarb Perspect Biol* 2018;10:a029330.
93. Savas P, Teo ZL, Lefevre C, Flensburg C, Caramia F, Alsop K, et al. The subclonal architecture of metastatic breast cancer: results from a prospective community-based rapid autopsy program "CASCADE. *PLoS Med* 2016;13:e1002204.
94. Legare S, Cavallone L, Mamo A, Chabot C, Sirois I, Magliocco A, et al. The estrogen receptor cofactor SPEN functions as a tumor suppressor and candidate biomarker of drug responsiveness in hormone-dependent breast cancers. *Cancer Res* 2015;75:4351–63.
95. Zundeleovich A, Dadiani M, Kahana-Edwin S, Itay A, Sella T, Gadot M, et al. ESR1 mutations are frequent in newly diagnosed metastatic and loco-regional recurrence of endocrine-treated breast cancer and carry worse prognosis. *Breast Cancer Res* 2020;22:16.
96. Mullane SA, Werner L, Rosenberg J, Signoretti S, Callea M, Choueiri TK, et al. Correlation of apobec mRNA expression with overall survival and pd-1 expression in urothelial carcinoma. *Sci Rep* 2016;6:27702.
97. Denkert C, Untch M, Benz S, Schneeweiss A, Weber KE, Schmatloch S, et al. Reconstructing tumor history in breast cancer: signatures of mutational processes and response to neoadjuvant chemotherapy (small star, filled). *Ann Oncol* 2021;32:500–11.
98. Blass E, Ott PA. Advances in the development of personalized neoantigen- based therapeutic cancer vaccines. *Nat Rev Clin Oncol* 2021;18:215–29.
99. Dougan M, Pietropaolo M. Time to dissect the autoimmune etiology of cancer antibody immunotherapy. *J Clin Invest* 2020;130:51–61.
100. Blank CU, Enk A. Therapeutic use of anti-CTLA-4 antibodies. *Int Immunol* 2015;27:3–10.
101. Philips GK, Atkins M. Therapeutic uses of anti-PD-1 and anti-PD-L1 antibodies. *Int Immunol* 2015;27:39–46.
102. DiMarco AV, Qin X, McKinney BJ, Garcia NM, Van Alsten SC, Mendes EA, et al. APOBECmutagenesis inhibits breast cancer growth through induction of T cell-mediated antitumor immune responses. *Cancer Immunol Res* 2022;10:70–86.
103. Cescon DW, Haibe-Kains B, Mak TW. APOBEC3B expression in breast cancer reflects cellular proliferation, while a deletion polymorphism is associated with immune activation. *Proc Natl Acad Sci U S A* 2015;112:2841–6.
104. 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.
105. Smid M, Rodriguez-Gonzalez FG, Sieuwerts AM, Salgado R, Prager-Van der Smissen WJ, Vlugt-Daane MV, et al. Breast cancer genome and transcriptome integration implicates specific mutational signatures with immune cell infiltration. *Nat Commun* 2016;7:12910.

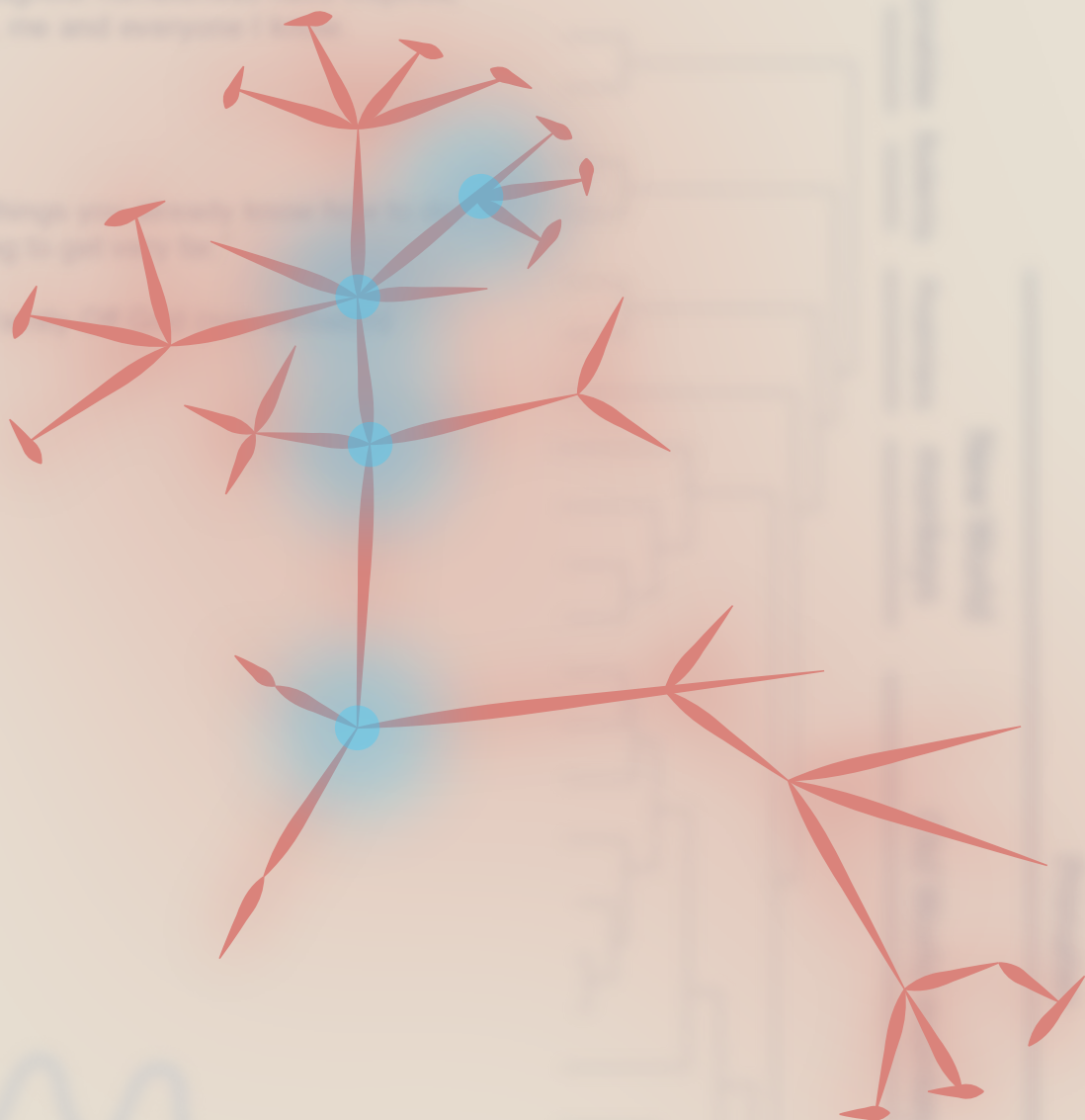
106. Boichard A, Pham TV, Yeerna H, Goodman A, Tamayo P, Lippman S, et al. APOBEC-related mutagenesis and neo-peptide hydrophobicity: implications for response to immunotherapy. *Oncoimmunology* 2019;8:1550341.
107. Litchfield K, Reading JL, Puttick C, Thakkar K, Abbosh C, Bentham R, et al. Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* 2021;184:596–614.
108. Guo H, Zhu L, Huang L, Sun Z, Zhang H, Nong B, et al. APOBEC alteration contributes to tumor growth and immune escape in pan-cancer. *Cancers* 2022;14:2827.
109. Buisson R, Lawrence MS, Benes CH, Zou L. APOBEC3A and APOBEC3B activities render cancer cells susceptible to ATR inhibition. *Cancer Res* 2017;77:4567–78.
110. Green AM, Landry S, Budagyan K, Avgousti DC, Shalhout S, Bhagwat AS, et al. APOBEC3A damages the cellular genome during DNA replication. *Cell Cycle* 2016;15:998–1008.
111. 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;77:4579–88.
112. Barnieh FM, Loadman PM, Falconer RA. Progress towards a clinically successful ATR inhibitor for cancer therapy. *Curr Res Pharmacol Drug Discov* 2021;2:100017.
113. Biayna J, Garcia-Cao I, Alvarez MM, Salvadores M, Espinosa-Carrasco J, McCullough M, et al. Loss of the abasic site sensor HMCES is synthetic lethal with the activity of the APOBEC3A cytosine deaminase in cancer cells. *PLoS Biol* 2021;19:e3001176.
114. 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 U S A* 2019;116:22158–63.
115. Mehta KPM, Lovejoy CA, Zhao R, Heintzman DR, Cortez D. HMCES maintains replication fork progression and prevents double-strand breaks in response to APOBEC deamination and abasic site formation. *Cell Rep* 2020;31:107705.
116. Barzak FM, Harjes S, Kvach MV, Kurup HM, Jameson GB, Filichev VV, et al. Selective inhibition of APOBEC3 enzymes by single-stranded DNAs containing 2'-deoxyzebularine. *Org Biomol Chem* 2019;17:9435–41.
117. Kvach MV, Barzak FM, Harjes S, Schares HAM, Jameson GB, Ayoub AM, et al. Inhibiting APOBEC3 activity with single-stranded dna containing 2'-deoxyzebularine analogues. *Biochemistry* 2019;58:391–400.
118. Kvach MV, Barzak FM, Harjes S, Schares HAM, Kurup HM, Jones KF, et al. Differential inhibition of APOBEC3 DNA-mutator isozymes by fluoro- and non-fluoro-substituted 2'-deoxyzebularine embedded in single-stranded DNA. *Chembiochem* 2020;21:1028–35.
119. Barzak FM, Ryan TM, Kvach MV, Kurup HM, Aihara H, Harris RS, et al. Small-angle X-ray scattering models of APOBEC3B catalytic domain in a complex with a single-stranded DNA inhibitor. *Viruses* 2021;13:290.
120. King JJ, Borzooee F, Im J, Asgharpour M, Ghorbani A, Diamond CP, et al. Structure-based design of first-generation small molecule inhibitors targeting the catalytic pockets of AID, APOBEC3A, and APOBEC3B. *ACS Pharmacol Transl Sci* 2021;4:1390–407.
121. De Mattos-Arruda L. PIK3CA mutation inhibition in hormone receptor positive breast cancer: time has come. *ESMO Open* 2020;5:e000890.
122. Su J, Ruan S, Dai S, Mi J, Chen W, Jiang S. NF1 regulates apoptosis in ovarian cancer cells by targeting MCL1 via miR-142-5p. *Pharmacogenomics* 2019;20:155–65.
123. Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR, Campbell C. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. *Bioinformatics* 2018;34:511–3



A dedication to my parents,  
for delightful memories have inspired  
me, me and everyone I love.

For the things we have done together  
and going to get it done.

With love and respect.



# 6

English Summary

Nederlandse Samenvatting

Research Data Management Plan

List of publications

PhD Portfolio

Acknowledgements

Curriculum Vitae



## English Summary

APOBEC3 (A3) enzymes are powerful and sophisticated custodians of the mammalian genome, potently restricting propagation of viruses and retrotransposons. However, their efficiency and agnosticism towards their substrate's origin leaves the host vulnerable to incursions into their own genomic integrity. This "design flaw" (phrasing used in the loosest of ways) has now been laid bare by an exciting example of interdisciplinary research, namely that between virology and cancer biology.

The causal link between A3 enzymatic activity and the accumulation of mutations in viral genomes was established over two decades ago and focused predominantly on the anti-retroviral activity of A3D, A3F, A3G, and A3H. Additionally, several viral proteins that counteract and avoid A3-mediated restriction, including the proteins Vif and BORF2, have since been identified. An increasing number of reports have since been published that demonstrate the contribution of A3-mediated mutagenesis to the accumulation of sporadic mutations in many different cancer types. Later research homed in on the most important candidates (A3A, A3B, and A3H), of which now A3A and A3B are regarded as the most ubiquitous causes for APOBEC signature mutations in cancer. The high mutational burden associated with A3-activity throws a grim shadow at traditional anti-cancer treatments, highlighting the importance of a better understanding of A3-associated mutagenesis and the innovation of existing treatment regimens. As part of this endeavor, this thesis has focused on elucidating the transcriptional activation of A3B, as well as reviewing several clinical aspects surrounding A3-mediated mutagenesis.

Although APOBEC signature mutations are found in a diverse collection of cancer subtypes its presence in breast cancer is remarkable, especially considering the unlikely presence of a significant viral component in the etiology of breast cancer. However, the pathways that trigger A3A and A3B in cancers *with* a viral background do leave important clues as to why and how the same enzymes are upregulated in breast cancer. Therefore, understanding the mechanisms through which A3A and A3B enzymes are upregulated is a crucial element in the overall effort to combat the accumulation of genomic mutations in breast cancer.

The journey described in this thesis started out with the family of polyomaviruses (PyV), a distinct class of viruses that has been associated with rare tumor types, including Merkel cell carcinoma and kidney transplant-associated cancer. Specifically, in human kidney cells A3B expression increases sharply upon PyV-infection (see **Chapter 2**). For this induction a viral protein called large-T antigen

(LTA<sub>g</sub>) is required, which, importantly, contains a domain that strongly implied the involvement of the inhibitory RB/E2F pathway. This finding provided a clue that disruption of *A3B* inhibition (creating a double negative) by LTA<sub>g</sub> is responsible for *A3B* upregulation in this system. Further analysis of tumor transcriptome data showed that in breast cancer the expression of *A3B* tightly correlates with activation of the RB/E2F pathway. This paper, although quite exploratory in nature, laid much of the groundwork for the mechanistic papers that followed. Furthermore, this paper highlighted the position of *A3B* in the cell's "rapid response system" that aids in maintaining host genome integrity and overall health.

Expanding on this work, **Chapter 3** provided insight into the mechanism that controls *A3B* expression in healthy cells and illustrated how failure to maintain control can jeopardize genomic integrity. Using LTA<sub>g</sub> merely as a means to induce *A3B* expression, a region on the *A3B* promoter was identified that is indispensable for the upregulation of *A3B* by LTA<sub>g</sub>. This promoter region contains two binding sites (the E2F-binding site and CHR-element) that can recruit the inhibitory protein complex DREAM. As a crucial element of the RB/E2F-pathway, the DREAM complex inhibits a host of genes involved in cell cycle regulation. Site-specific gene editing and pull-down assays later confirmed that the DREAM complex indeed binds this promoter region and is responsible for the transcriptional inhibition of *A3B*. Moreover, by analyzing transcriptional and genomic data of a large cohort of breast cancers, it was shown that APOBEC signature mutations are more likely to occur in breast cancer where the RB/E2F pathway is dysregulated.

**Chapter 4** provides further insight into the regulation of *A3B* throughout the cell cycle in both healthy and cancer cells. This work displayed one of few applications of an *A3B*-antibody in immunohistochemical staining of breast cancer cell lines and tissues. Interestingly, staining patterns of *A3B* appeared spotty, even in cell lines known for their high *A3B* expression. This finding strongly indicated that the induction of *A3B* is a specific feature of certain cell cycle phases. Further analyses confirmed that *A3B* expression fluctuates in a cell cycle-dependent manner and chemical induction of *A3B* is most potent during proliferative cell cycle phases. Together, these findings postulate that *A3B* expression in breast cancer is a function of the *presence* of an inducer (such as the PKC/ncNF-κB signaling pathway) and the *absence* of an inhibitor (as is the case with DREAM during dormant cell cycle phases). Thus, the work presented in **Chapters 2, 3 and 4** contribute to the greater collection of data that describes the regulatory mechanisms of *A3* proteins.

Finally, **Chapter 5** aimed to provide a comprehensive overview of A3A/A3B regulation, as well as to outline several clinical aspects of APOBEC-related cancer research. This chapter highlights some of the most promising clinical detection methods for A3A and A3B, provides clinical examples of A3-associated mutations in cancer and explores several options for leveraging (as well as inhibiting) A3-mediated mutagenesis in the clinic. Taken together, the work presented in this thesis contributes to a better understanding of the regulation of A3 in normal and cancer cells. It also aptly emphasizes the need to develop strategies that mitigate the detrimental effects of A3-mediated mutagenesis, which will greatly impact the clinical outlook of breast cancer patients worldwide.

## Nederlandse Samenvatting

Het genoom van zoogdieren, waaronder die van de mens, wordt beschermd en beïnvloed door een groot aantal eiwitten, waaronder de APOBEC3 (A3) enzymen. Het menselijk genoom codeert voor 7 verschillende A3-varianten, genaamd A3A, A3B, A3C, A3D, A3F, A3G en A3H. Deze enzymen hebben als meest bekende functie het bestrijden van virussen, hetgeen wordt bereikt door het introduceren van mutaties in het virale genoom. Het begrip dat deze eiwitten viraal DNA niet afdoende kunnen onderscheiden van gastheer DNA levert echter een nadeel op voor de mens, wiens DNA ook kan worden aangevallen door deze efficiënte A3 enzymen. Deze “ontwerpfout” (in de meest losse zin van het woord) is inmiddels blootgelegd dankzij een klinkend voorbeeld van interdisciplinair wetenschappelijk onderzoek, namelijk dat tussen virologie en kankerbiologie.

Het verband tussen de enzymatische activiteit van A3 enzymen en de opeenstapeling van mutaties in virale genomen werd al meer dan 20 jaar geleden gelegd. Dit onderzoek heeft zich voornamelijk gericht op de A3 varianten A3D, A3F, A3G en A3H. Wederzijds zijn er virale eiwitten beschreven die de werking van A3 eiwitten juist kunnen belemmeren, zoals de eiwitten Vif en BORF2. Sinds die tijd is er een sterke toename zichtbaar van publicaties die laten zien dat de mutagene activiteiten van A3 eiwitten ook bijdragen aan het ontstaan van mutaties in veel verschillende kankersoorten. Onderzoek heeft zich voornamelijk gericht op de meest prominente kandidaten (A3A, A3B en A3H), waarvan nu A3A en A3B als voornaamste boosdoeners naar voren zijn gekomen. De hoge mutatielast die wordt geassocieerd met A3-activiteit werpt een grimmige schaduw over traditionele kankerbehandelingen. Het benadrukt enerzijds het belang van een dieper begrip van A3-geassocieerde mutagenese en anderzijds het belang van innovatie binnen bestaande behandelplannen. Om hier een deel aan bij te dragen is dit proefschrift toegewijd aan het ophelderen van de transcriptionele activatie van A3B en presenteert het een diepgaande review over de verscheidene klinische aspecten rondom de mutagenese die wordt veroorzaakt door A3 eiwitten.

Hoewel de typerende APOBEC-mutaties in een erg diverse verzameling kanker subtypes worden gevonden is de aanwezigheid ervan in borstkanker opmerkelijk, aangezien het onwaarschijnlijk is dat borstkanker een virale achtergrond heeft. De wijzen waarop A3A en A3B in kankers met een virale achtergrond worden geactiveerd geven echter wel belangrijke indicaties over hoe (en waarom) deze enzymen worden geüpereguleerd in borstkanker. Het begrijpen van de mechanismen waarmee A3A- en A3B-enzymen worden geüpereguleerd is daarom van cruciaal belang bij het bestrijden van de accumulatie van genomische mutaties in borstkanker.

De reis die in dit proefschrift staat beschreven begon met de virale familie van polyomavirussen (PyV), een aparte klasse van virussen die is geassocieerd met zeldzame tumortypen, zoals Merkel-celcarcinoom en kanker geassocieerd met niertransplantatie. Specifiek neemt de expressie van A3B in menselijke niercellen scherp toe na PyV-infectie (zie **Hoofdstuk 2**). Voor deze inductie is een viraal eiwit genaamd "large-T-antigen" (LTA<sub>g</sub>) vereist. Dit eiwit bevat een domein dat in sterk verband staat met de "RB/E2F signaling pathway"; welke onder andere een sterke remmende functie heeft op de celdeling. Deze bevinding gaf een aanwijzing dat transcriptie van A3B normaliter wordt geremd door de RB/E2F signaling pathway en dat verstoring hiervan door LTA<sub>g</sub> (wat een dubbel negatief zou creëren) leidt tot upregulatie van A3B in dit systeem. Verder onderzoek van transcriptomische gegevens van tumoren toonde aan dat in borstkanker de expressie van A3B sterk correleert met de activatie van de RB/E2F signaling pathway. Dit artikel, hoewel vrij verkennend van aard, legde de basis voor de mechanistische studies die volgden. Bovendien benadrukte dit artikel de positie van A3B in het "snelle respons systeem" van de cel dat helpt bij het behoud van de integriteit van het gastheergenoom en de algehele gezondheid.

**Hoofdstuk 3** bouwt verder op dit werk en geeft meer inzicht in het mechanisme dat de expressie van A3B reguleert in gezonde cellen. Bevindingen in dit hoofdstuk beschrijven tevens dat het verliezen van controle op A3B de genomische integriteit van de gastheer in gevaar kan brengen. Door LTA<sub>g</sub> louter te gebruiken als middel om de expressie van A3B op te wekken, werd een regio op de A3B-promotor geïdentificeerd die onmisbaar is voor de upregulatie van A3B door LTA<sub>g</sub>. Deze promotorregio bevat twee bindingsplaatsen (de E2F-bindingsplaats en CHR-element) die het proteïnecomplex DREAM kunnen rekruteren. Als cruciaal element van de RB/E2F signaling pathway remt het DREAM-complex een groot aantal genen die betrokken zijn bij de regulatie van de celcyclus. Door het *in vitro* aanbrengen van systematische mutaties en het door het uitvoeren van pull-down assays bevestigden wij hierna dat het DREAM-complex inderdaad deze promotorregio bindt en verantwoordelijk is voor de transcriptionele remming van A3B. Bovendien werd door de analyse van transcriptionele en genomische gegevens van een grote cohort van borstkankers aangetoond dat mutaties die veroorzaakt worden door A3 enzymen vaker voorkomen in borstkanker waar de RB/E2F signaling pathway is verstoord.

**Hoofdstuk 4** biedt meer inzicht in de regulatie van A3B gedurende de celcyclus in zowel gezonde cellen als kankercellen. Dit werk toonde een zeldzame toepassing van een A3B-antilichaam in immunohistochemische kleuring van borstkankercellijnen en -weefsels. Interessant was dat de aankleuring van A3B

veel verschilde tussen cellen onderling, zelfs in cellijnen die bekend staan om hun hoge A3B-expressie. Deze bevinding gaf sterk aan dat de inductie van A3B met name plaatsvindt binnen bepaalde fasen van de celdeling. Nadere analyses bevestigden dat de expressie van A3B fluctueert op een celcyclusafhankelijke wijze en dat chemische inductie van A3B het krachtigst is tijdens proliferatieve fasen van de celcyclus. Samengevat onderbouwen deze bevindingen het model dat de expressie van A3B in borstkanker een uitkomst is van de aanwezigheid van een inductor (zoals de PKC/ $\kappa$ B signaling pathway) en de afwezigheid van een remmer (zoals DREAM tijdens de rustfasen van de celcyclus). Zodoende draagt het werk gepresenteerd in **hoofdstukken 2, 3 en 4** bij aan een betrouwbaardere modellering van de regulatie van A3B.

Tot slot had **hoofdstuk 5** als doel een uitgebreid overzicht te geven van de regulatie van A3A en A3B en om de verschillende klinische aspecten van APOBEC-gerelateerd kankeronderzoek uiteen te zetten. Dit hoofdstuk belicht enkele van de meest veelbelovende klinische detectiemethoden voor A3A en A3B, geeft klinische voorbeelden van A3-geassocieerde mutaties bij kanker en verkent verschillende opties om de mutagene activiteit van A3 enzymen klinisch te benutten of juist te remmen. Samengevat draagt het werk gepresenteerd in deze scriptie bij aan een beter begrip van de regulatie van A3 in normale cellen en kankercellen. Het benadrukt ook de noodzaak om strategieën te ontwikkelen die de schadelijke mutagene activiteit van A3 enzymen tegengaan; iets wat van grote invloed zal zijn op de klinische vooruitzichten van borstkankerpatiënten wereldwijd.

## Research Data Management Plan

Raw data and preliminary interpretations generated at the Radboudumc between May 2017 and February 2019 have been recorded in a digital lab notebook (Labguru), which is backed up using a hard drive (WD Elements Time Machine) at the Department of Radiation Oncology (ROI Lab) at the Radboudumc. Raw data obtained at the Radboudumc between July 2020 and November 2021 were similarly backed up at the Department of Radiation Oncology. Data generated over two separate visits to the University of Minnesota (February 2019 - July 2020 and November 2021 – June 2022) were duplicated with permission and also backed up using a hard drive (WD Elements Time Machine) at the Department of Radiation Oncology (ROI Lab) at the Radboudumc. Backed-up data was organized in four separate folders (Radboudumc 2017-2019, UMN 2019-2020, Radboudumc 202-2021 and UMN 2021-2022), and experiments were given unique identifiers, contributing to sustainable Findability. Where applicable these folders include a subfolder with lab-specific standard operating procedures. All data files on Labguru and back-ups on the local server are exclusively accessible by relevant research employees and scientific staff members, ensuring proper Access management. Software required to properly access and interpret data are widely used, although some systems are available as commercial versions only, such as FlowJo, Snapgene and GraphPad Prism. Data obtained at the Radboudumc are conditionally obtainable through Paul Span, PhD. Data obtained at the University of Minnesota are conditionally obtainable through Professor Reuben Harris (currently a resident professor at UT Health San Antonio, TX, USA). The supplementary files (Excel sheets) mentioned in Chapter 3 are available as online downloads using the appropriate DOI: <https://doi.org/10.7554/eLife.61287>, or by contacting Paul Span, PhD.

## List of publications

### Published

- Dhekne HS, Hsiao NH, **Roelofs P**, Kumari M, Slim CL, Rings EH, van Ijzendoorn SC. Myosin Vb and Rab11a regulate phosphorylation of ezrin in enterocytes. *J Cell Sci.* 2014 Mar 1;127(Pt 5):1007-17. doi: 10.1242/jcs.137273. Epub 2014 Jan 10. PMID: 24413175.
- Starrett GJ, Serebrenik AA, **Roelofs PA**, McCann JL, Verhalen B, Jarvis MC, Stewart TA, Law EK, Krupp A, Jiang M, Martens JWM, Cahir-McFarland E, Span PN, Harris RS. Polyomavirus T Antigen Induces *APOBEC3B* Expression Using an LXCXE-Dependent and TP53-Independent Mechanism. *mBio.* 2019 Feb 5;10(1):e02690-18. doi: 10.1128/mBio.02690-18. PMID: 30723127; PMCID: PMC6428753.
- **Roelofs PA**, Martens JWM, Harris RS, Span PN. Clinical Implications of APOBEC3-Mediated Mutagenesis in Breast Cancer. *Clin Cancer Res.* 2023 May 1;29(9):1658-1669. doi: 10.1158/1078-0432.CCR-22-2861. PMID: 36478188; PMCID: PMC10159886.
- **Roelofs PA**, Goh CY, Chua BH, Jarvis MC, Stewart TA, McCann JL, McDougale RM, Carpenter MA, Martens JW, Span PN, Kappei D, Harris RS. Characterization of the mechanism by which the RB/E2F pathway controls expression of the cancer genomic DNA deaminase APOBEC3B. *Elife.* 2020 Sep 28;9:e61287. doi: 10.7554/eLife.61287. PMID: 32985974; PMCID: PMC7553775.
- **Roelofs PA**, Timmermans MAM, Stefanovska B, den Boestert MA, van den Borne AWM, Balcioglu HE, Trapman AM, Harris RS, Martens JWM, Span PN. Aberrant APOBEC3B Expression in Breast Cancer Is Linked to Proliferation and Cell Cycle Phase. *Cells.* 2023 Apr 18;12(8):1185. doi: 10.3390/cells12081185. PMID: 37190094; PMCID: PMC10136826.



## PhD Portfolio of Petrus Adriaan Roelofs

**Department:** Department of Radiation Oncology  
**PhD period:** 01/05/2017 – 01/09/2021  
**PhD Supervisor(s):** dr. Paul Span, prof. dr. R.S. Harris, prof. dr. J.W.M. Martens  
**PhD Co-supervisor(s):** None

Training activities	Hours
<b>Courses</b>	
• Radboudumc - Introduction day (2017)	6.00
• RIMLS - Introduction course "In the lead of my PhD" (2017)	15.00
• Statistiek voor promovendi met Excel (2017)	42.00
• Working with radionuclides level 5b (2017)	56.00
• The Art of Presenting Science (2018)	42.00
• German language course in preparation of post-PhD career (2020)	56.00
• RU - Mindfulness Based Stress Reduction (2020)	45.00
• Radboudumc - Scientific integrity (2020)	20.00
<b>Seminars</b>	
• PhD retreat (with poster) (2017)	7.00
• Radboud New Frontiers in Microbiome, attended (2017)	18.00
• RIMLS Meet the Expert CRISPR/Cas9, attended (2017)	3.00
• Meet the Expert Simon Dovedi, attended (2017)	5.00
• Women's Cancer PhD lunch Meetings, attended (2018)	4.00
• Seminar Inge Verbruggen, attended (2018)	2.00
• Seminar "De nieuwe mens komt uit het lab", attended (2018)	2.00
• Radboud Reflects Seminars, attended (2018)	6.00
• PhD Lunchmeetings Theme Women's cancers, attended, presented once (oral) (2019)	11.00
• MinnCResT seminar by Michael Emerman (University of Minnesota), attended (2019)	2.00
• Synthego CRISPR seminar, attended (2019)	1.50
• Chemical Biology Seminar, Prof. Geoff Jameson, attended (2020)	2.00
• Radboud Research Rounds, attended (3x) (2020)	8.50
• Masonic Cancer Center Talks (University of Minnesota), attended (2019)	5.00
• Masonic Cancer Center Talks (University of Minnesota), attended (2020)	6.00
<b>Conferences</b>	
• 26th meeting KWF Society for Cancer Biology (CB2017), attended (2017)	14.00
• Cancer Immunology Immunotherapy (CII) meeting January 2018, attended (2018)	3.00
• 13th Annual EORTC PathoBiology Group Meeting, oral presentation (2018)	28.00
• 27th meeting KWF Society for Cancer Biology, attended (CB2017) (2018)	14.00
• American Society for Virology Conference 2019, attended (2019)	35.00

**Other**

• Radboudumc - General Radboudumc introduction for research personnel (2017)	9.00
• Laptop presentation Research Theme Women's Cancers (2017)	1.00
• Radiotherapy meets Immunology symposium (2017)	8.50
• Masonic Cancer Center's 9th Annual Cancer Research Symposium (2018)	7.00
• Journal Clubs Dept. of Radiation Oncology (2018)	8.50
• Cancer Immunology Immunotherapy (CII) meetings (2018)	6.00
• Journal Clubs Dept. of Radiation Oncology & ROI lab (2019)	14.00
• Radboud Oncologie Fonds Task Force meetings (2019)	7.00
• Annual Institute for Molecular Virology symposium (2019)	7.00
• Annual Genetic Mechanisms Program Retreat (2019)	3.50
• Biochemistry, Molecular Biology & Biophysics Departmental Retreat (Itasca Biological Station) (2019)	14.00
• CRISPR seminar by Thermo: A Workflow for Knock-in Genome Editing: Simplified (2020)	2.00
• Weekly Journal Clubs Dept. of Biochemistry, Molecular Biology & Biophysics (University of Minnesota) (2020)	77.00

**Teaching activities****Supervision of internships / other**

• Supervising Honour's Students program 2018 block 2 (2018)	42.00
• Supervising Thesis Student Molecular Sciences (2018)	8.00
• Supervising: Introduction to Lab- Lieke ter Steeg, Honour's student. (2018)	22.50
• Assisting Onderwijspracticum MIN21 (2018)	28.00
• Honours Programme: preparation Internship Abroad- Lieke ter Steeg (2019)	56.00
• Student (2020)	18.00
• Student (2021)	44.50
• Student (2022)	41.50

**Total****884.00**

## Acknowledgements

Long, long years- it has truly been a long time coming. There were times I thought I would never be writing these acknowledgements; yet, here we are. Now that this journey is reaching its close, I would like to extend my gratefulness to all those involved, either directly or indirectly. In a (somewhat) chronological order:

**Paul**, thank you so much for including me in the APOBEC3B project. It has been a while ago, but I still remember how we almost immediately bonded over a combination of Twin Peaks and our passion for science. You managed to lay a very powerful collaborative foundation based on shared interests, mutual trust, empathy and our shared tendency for long-winded discussions. Thank you for having created such a supportive and inspiring environment and for teaching me to think critically whilst staying open-minded. Above all other things I want to thank you for your exemplary patience. It still baffles me how you managed to make me stand up, be pragmatic and finish what I started.

**Reuben**, you were crucial during my formative years as a scientist. I did not tell you back then, but my first day at the CCRB has since become a core memory; how it felt to suddenly run with the wolves! I still remember the lab parties, early morning meetings with bagels, and writing together in the sun outside of CCRB. Your ability to rapidly assemble a list of possible explanations for any observation is truly commendable. Thanks to you I have learned to be more pragmatic, to think big, and to keep a helicopter-view. I also learned so much from your excellent writing skills; you opened up a whole new world to me. Your invite for a second visit to Minneapolis was highly appreciated, thank you again for that opportunity. I wish you all the best in San Antonio.

**John**, I would like to thank you for the great guidance throughout my PhD research. I have a deep admiration for your ability to be critical yet supportive, highly perceptive yet approachable. You have an excellent skill in connecting people with different expertise, leading to great scientific collaborations. The (breast) cancer field is lucky to have you, and I am grateful that I got to be part of the APOBEC team.

I want to thank all the past and present members of the Department of Radiation Oncology (formerly ROI laboratory), especially **Annemarie Post, Eva-Leonne Göttgens, Fokko Huizinga, Jan Bussink, Gosse Adema, Marleen Ansems, Wenny Peeters, Maaïke Looman, Jasper Lok, Hans Peters, Lenneke Cornelissen, Daan**

**Boreel, Fabian Schuurmans, Flavia Reinema, Vera Kho, Vera Mekers, Eline van Houtum, Katrin Rabold, Lisa Huis in 't Veld, Natasja Balneger, and Renske van den Bijgaart.** Thank you for the warm welcome, fruitful discussions, support and watchful eyes.

A very warm and heart-felt thank you to all my BSc and Msc students; **Lieke ter Steeg, Wilke Castelijns, Shuyu Meng, David Tissen, and Amber van den Borne.** Thank you for your help in the lab, the interesting discussions and, especially, for making me realize how incredible rewarding it is to see people learn. I wish all of you all the best in your future endeavors.

All the good friends I have made in the USA: thank you, thank you, THANK YOU! It is with tears in my eyes that I am writing this: you have truly carried me through the toughest of times. You have showed me what friendship truly means and for that I am forever grateful. **Jake Brindley**, I miss you dearly and I am looking forward to the day our paths cross again. We will talk soon, my dear friend. Until that time, I will treasure the time we have spent together. **Cameron Durfee**, you have been a great support during both my stays in Minneapolis. Thank you for all the laughs, deep discussions, bouldering sessions, and late (and early!) hours in the lab. **Miranda Olson**, you were always there for me, even when I was on my crazy kicks and even when I was a bad friend. You were a rock and a great house-mate. I wish you all the best for the future; I know you will get anything you put your mind to! Say hi to Leo for me.

I want to express my gratitude to all my former co-workers of the Harris Lab that I got to work with at the Cancer and Cardiovascular Research Building. Especially, **Jenny McCann, Artur Serebrenik, Matthew Jarvis, Bojana Stefanovska, Emily Law, Daniel Salamango, Mike Carpenter, and Bill Brown.** You all formed the rock-solid scientific foundation that kept me going when push came to shove. All the best, be it in Texas, North Carolina, Minnesota or elsewhere. May your plasmid database always remain pristine.

A great shout out to **Teneale Stewart**; it was your meticulous and well-organized work that laid the foundation for this thesis. Thank you for the discussions we had, and the best of luck to you!

I want to thank professor **Arie van Winkelhoff** for the very first time that someone told me I should reach higher. Thank you kindly for making me believe in myself and for introducing me to the wonderful world of science.

Dear **Kirti**. Where to even start! We met by happenstance on a PON picnic close to five years ago. So many things have changed, yet other things never will! Thank you for being such a supporting friend during my PhD and for being an amazing friend throughout. I will never forget (or will I?) our sometimes-heated discussions, laughs and overall good times. It has been inspiring to see you grow into a great scientist, a strong voice amongst the crowd. I wish you all the best in your future career, wherever this may take place!

Dear **Imre**, you are a great friend; I feel so lucky to have met you in Nijmegen. The whisky tastings and the times we went bouldering together are some of my fondest memories. It has been too long already, let us meet again soon. All the best with moving soon, I will come and find you.

Dear **Tommy**, ever since we met back in 2020 you were there for me. Your empathy, optimism, and genuine interest in how I was doing supported me throughout my PhD and after. Also, thank you for fueling my bouldering craze, it was a great way to release stress after work! Good luck in Wuppertal, and I hope to see you soon.

**Steven**, thanks for having been my home for almost 1.5 years. It sure was a rocky road, but looking back I would not have wanted it any different way. Enjoy your new engine and turbocharger; from here on out you will get a timely oil change. Here is to a bright future!

Although I have only seen in person some of you, I also want to thank a number of artists. **James Taylor**, your music carried me through some of the toughest moments and taught me to appreciate the small things. The things I can still understand. Thank you, and I hope to meet you once more. Also, **Mark Knopfler** and **Paul Simon**, your music inspired me during the long days at the lab. I am thankful for your music. Another "Thank you!" goes out to podcasters **Steven Novella**, **Cara Santa Maria**, **Noah Lugeons**, **Heath Enwright**, and **Eli Bosnick**. Through your podcasts all of you introduced me to scientific skepticism, critical thinking, humanism, atheism, and flat-out good humor. Your voices of reason do not go unheard in this time of unrest. It is hard to put into words how critical of a role you played in getting me where I am today, and I thank you for that!

**Eefke**, wat was het ontzettend fijn om je te ontmoeten in Venlo, alweer bijna 2,5 jaar geleden! Het was voor mij erg prettig om te kunnen praten met iemand die ook actief was binnen de academische wereld. Bedankt voor de fijne en diepgaande

gesprekken die we hebben gehad, ik ga ons flauwe grappen en grollen nooit vergeten. Ik wens je het alle beste toe in Maastricht, ik twijfel er geen moment aan dat je je PhD gaat halen. Voor deze ene keer zal ik je met rust laten en niet afsluiten met een flauw grapje.

Ik wil ook de rest van mijn bouldervriendengroepje bedanken, in het bijzonder **Vita, Bram, Kristie, Yvonne** en **Emma**. Deze sport (en jullie aanwezigheid!) is zowat therapeutisch; ik ben blij dat ik jullie heb mogen vinden. We gaan nog veel uren op de mat (kom van de mat af!) en aan de muur beleven, in welke hal dat dan ook moge zijn. Ook een bedankje richting de jeugdtraining in Venlo! Door jullie enthousiasme hebben jullie me stiekem erg gemotiveerd gehouden, ik heb veel van jullie mogen leren.

Lieve **mama, papa, oma, (schoon)broers** en **(schoon)zussen**. Verreweg mijn beste herinnering uit mijn PhD-onderzoek is de dag dat ik papa een verrassingsbezoek mocht brengen op zijn 60e verjaardag. Ik ga dat nooit vergeten! Wat was het fijn om jullie toen allemaal weer te zien. Bedankt voor jullie voelbare steun tijdens mijn PhD-onderzoek. Het moet voor jullie soms bevreemdend zijn geweest dat ik me zo verloor in zaken waarvan jullie niet goed wisten wat ze betekenden. Ik waardeer het enorm dat jullie mij bleven steunen en vertrouwen in de zaak bleven hebben.

**Mama van Fenna**, lieve Anita, bedankt voor ook jouw steun in het afgelopen jaar. Je helpt ons bijna dagelijks uit de brand en ik ben je nog steeds dankbaar dat ik samen met Fenna zowat twee maanden bij je mocht neerstrijken. Ik waardeer het enorm dat ik aan het einde van een lange dag even bij je mag uitrazen als ik Rups kom halen. We gaan nog veel leuke en mooie dingen beleven met zijn allen!

**Rups**, je weet het zelf al, maar je bent simpelweg de beste.

Lieve **Fenna**, toen wij elkaar ontmoetten lag mijn PhD-traject alweer even achter me. Ik ben ontzettend blij dat wij elkaar hebben ontmoet in een tijd waarop onze beide toekomst en eindelijk overzichtelijk begonnen te worden. Je steun was voelbaar tijdens het afronden van mijn PhD, bedankt dat je deze rit tot het einde met me hebt uitgezet. Ik houd van je en ik kijk uit naar de komende 60 jaar!

## Curriculum Vitae

Petrus Adriaan (Pieter) Roelofs was born on December 26<sup>th</sup> 1990 in Alphen aan den Rijn, The Netherlands. Being a dreamy kid with an interest in “how things work”, a general passion for biology began to develop in high school. This led to Pieter enrolling in the Biology and Medical Laboratory Research (BASc) program at the Saxion University of Applied Sciences in Deventer, The Netherlands. As part of this program Pieter interned at two laboratories at the University Medical Center Groningen. Throughout his first internship Pieter investigated the possible role of *Porphyromonas gingivalis* infection in the etiology of rheumatoid arthritis, under the supervision of professor Arie van Winkelhoff. His second internship focused on the effect of *MYO5* defects on the localization of Rab11 proteins in Microvillus Inclusion Disease, which resulted in his first co-authorship together with Herschel Dhekne, PhD, and professor Sven van IJendoorn. After graduating in 2013, Pieter initially aimed to work as a laboratory technician, but opted to continue his studies at a MSc level. After having completed a Pre-MSc program in 2013, he enrolled in the Biomedical Sciences (MSc) program at the University of Groningen in 2014. During this program Pieter developed a passion for stem cell biology and cancer research and devoted two internships to pursue these new interests. His first Master’s internship was supervised by Arend Overeem, PhD, and focused on reprogramming kidney-resident cells into hepatocyte-like cells. For his second Master’s internship Pieter moved to Utrecht in 2015 to join the lab of professor Hans Clevers (Hubrecht Institute), and investigated p53 stabilization therapy in APC-deficient intestinal organoids under the supervision of Jarno Drost, PhD. After obtaining his MSc degree he joined the Department of Radiation Oncology (formerly known as the Radiotherapy & Oncolmunology Laboratory) as a PhD student, under direct advisory of Paul Span, PhD. Pieter’s two other advisors were professor John Martens (Erasmus Universiteit Rotterdam) and professor Reuben Harris (then at the University of Minnesota, now UT Health San Antonio). Pieter moved to Minneapolis for a 16-month research visit at the Reuben Harris research group in February 2019. He continued his research at the Department of Radiation Oncology in 2020, and revisited the Reuben Harris research group for six months in November 2021. During his PhD research Pieter focused on the regulation of *APOBEC3B* expression in breast cancer through the repressive DREAM protein complex (this work). He also put substantial effort into the investigation of a synthetic lethality model involving APOBEC3A/B expression in homologous recombination deficient cancer cells (not included in this work). During and after his time in the lab Pieter developed a keen interest in education, mentoring and coaching. Currently, Pieter is working as a general coach and Biology teacher at Buurtcollege Agora Maas en Peel, in Helden (The Netherlands).

