Title: The role of APOBEC3B in lung tumour evolution and targeted therapy resistance

Authors: Deborah R. Caswell[a*†], Manasi K. Mayekar[b†], Philippe Gui[b†], Emily K. Law[e-h], Natalie I. Vokes[c-d], Carlos Martinez Ruiz[k,r], Michelle Dietzen[k,r], Mihaela Angelova[a], , Chris Bailey[a], Oriol Pich[a], Bjorn Bakker[a], Wei Wu[b], Timothy J. Humpton[u], D. Lucas Kerr[b], William Hill[a], Wei-Ting Lu[a], Franziska Haderk[b], Collin M. Blakely[b], Maise Al Bakir[a], Ai Nagano[a], Francisco Gimeno-Valiente[r], Clare Weeden[a], Andrew Rowan[a], Caroline E. McCoach[b], Bruna Almeida[l], Mary Green[l], Carlos Gomez[b], Beatrice Gini[b], Shigeki Nanjo[b], Dora Barbosa[b], Chris Moore[m], Joanna Przewrocka[a], Eva Grönroos[a], Simon Priestnall[t], Alejandro Suarez-Bonnet[t], Caroline Zverev[s], Scott Lighterness[s], James Cormack[s], Victor Olivas[b], Lauren Cech[b], Brandon Rule[a], Paul Ashford[w], Cameron Durfee[e-f], Nuri Alpay Temiz[i], Lisa Tan[b], Lindsay K. Larson[e-g], Prokopios P. Argyris[e-h,n], William L. Brown[e-g], Elizabeth A. Yu[a], Julia K. Rotow[j], Nitin Roper, Johnny Yu[o], Rachel I. Vogel[f,p], Nicholas J. Thomas[b], Subramanian Venkatesan[a,r], Samuel F. Bakhoum[x], Su Kit Chew[r], Mariam Jamal-Hanjani[r], Karen Vousden[u], Nicholas McGranahan[k,r], Eliezer M. Van Allen[q], Nnennaya Kanu[r], Reuben S. Harris[e-h], Julian Downward[m], Trever G. Bivona[b,v*], Charles Swanton[a,r]

Affiliations:

[a]Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK
[b]Department of Medicine, University of California, San Francisco, San Francisco, CA 94158, USA
[c]Department of Thoracic and Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas
[d]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas
[e]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN, USA
[f]Department of Biochemistry and Structural Biology, University of Texas Health San Antonio, San Antonio, TX 78229, USA
[g]Howard Hughes Medical Institute, University of Texas Health San Antonio, San Antonio, TX 78229, USA
[h]Howard Hughes Medical Institute, University of Minnesota, Minneapolis, MN, USA
[i]Health informatics, University of Minnesota, MN, USA
[j]Lowe Center for Thoracic Oncology, Dana-Farber Cancer Institute, Boston, MA, USA
[k]Cancer Genome Evolution Research Group, University College London, Cancer Institute, London UK
[l]Experimental Histopathology, The Francis Crick Institute, London, UK
[m]Oncogene Biology Laboratory, The Francis Crick Institute, London, UK
[n]Division of Oral and Maxillofacial Pathology, School of Dentistry, University of Minnesota, Minneapolis, MN, USA
[o]Biomedical Sciences Program, University of California, San Francisco, San Francisco, CA 94158, USA

[p]Department of Obstetrics, Gynecology, and Women's Health, University of Minnesota, Minneapolis, MN, USA
[q]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA
[r]Cancer Research UK Lung Cancer Centre of Excellence, UCL Cancer Institute, London, UK
[s]Biological Research Facility, The Francis Crick Institute, London, UK
[t]Department of Pathobiology & Population Sciences, The Royal Veterinary College, London, UK
[u] p53 and Metabolism Laboratory, The Francis Crick Institute, London, UK
[v]Chan-Zuckerber Biohub
[w] Institute of Structural and Molecular Biology, University of College London, Gower Street, London, UK
[x]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, 10065
[y]Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, 10065

[†]Authors contributed equally
[*]Correspondence: trever.bivona@ucsf.edu, deborah.caswell@crick.ac.uk

Summary:

Apolipoprotein B mRNA-editing catalytic subunit-like (APOBEC) enzymes are cytidine deaminases that play an important role in intrinsic responses to infection by viruses where they inhibit infection by deaminating deoxycytidine residues in viral single-stranded DNA[1,2]. APOBEC3 enzymes can act as potent host genome mutagens in multiple cancer types including non-small cell lung cancer (NSCLC)[3,4]. There are seven human APOBEC3 family members (APOBEC3A-H) and all but one (G) preferentially deaminate TC dinucleotide substrates[3]. In patients, both APOBEC3A (A3A)[5] and APOBEC3B (A3B)[6] have been implicated to play a major role in NSCLC[3]. In this study, the impact of A3B on epidermal growth factor receptor (EGFR)-driven lung cancer evolution was assessed. EGFR mutant NSCLC mouse models containing a human A3B transgene[7] were generated, and the effect of enforced A3B expression was evaluated at multiple timepoints. Enforced A3B expression was detrimental to tumour initiation and progression in the absence of targeted therapy due to elevated chromosomal instability (CIN) and p53 activation, which increased significantly in tumours with A3B expression. This finding was substantiated in the TRACERx421 study demonstrating a significant correlation between CIN and A3B in patients with NSCLC. In contrast, A3B expression with targeted therapy resulted in a significant increase in treatment resistance in mouse models and human preclinical models of EGFR mutant lung cancer. In specimens collected from patients with lung adenocarcinoma, targeted therapy upregulated A3B expression, and induced significantly more APOBEC-associated mutations including variants associated with therapy resistance. Analyses of human NSCLC models revealed that targeted therapy induced A3B upregulation in part via treatment-mediated activation of nuclear factor (NF)-κB signalling, and downregulation of the DNA repair enzyme uracil-DNA glycosylase (UNG). This study uncovers the multifaceted roles of A3B on tumour evolution and identifies APOBEC deaminases as future targets for eliciting more durable responses to targeted cancer therapy.

Main Text:

The mechanisms underlying cancer genome heterogeneity and evolution prior to and during therapy remain an active and important area of investigation. Recent work has established the adaptive mutability of cancer cells in response to targeted therapy through suppression of the mismatch repair (MMR) pathway and up-regulation of error-prone polymerases in drug-tolerant

94 cells[8]. This emerging paradigm holds important implications for the mechanistic understanding of
95 tumourigenesis as well as the evolution of tumours during therapy.
96
97 Prior tumour genome sequencing studies, including work from the multi-region sequencing
98 NSCLC TRACERx study, revealed subclonal enrichment for mutations in an APOBEC-preferred
99 substrate context, suggesting a possible role for this enzyme family in adaptive mutability later in
100 tumour evolution[9-12]. Gene expression analysis of the entire APOBEC3 family across multiple
101 stages of lung adenocarcinoma progression revealed the significantly elevated expression of A3B
102 at multiple timepoints (adenocarcinoma in situ and invasive lung adenocarcinoma) compared to
103 normal tissue[4].
104
105 Mouse models are one key system for the advancement of our understanding of cancer evolution
106 and drug response[13-16]. However, recent findings demonstrate that these models lack the mutational
107 heterogeneity present in human tumours[17-19]. This may be due, in part, to the fact that mice encode
108 only a single, cytoplasmic and non-genotoxic APOBEC3 enzyme[20,21]. To understand the role of
109 A3B in tumour evolution and therapy response, several new mouse strains incorporating a human
110 A3B transgene were engineered to mimic clonal and subclonal induction of A3B and to evaluate
111 the potential role of A3B in tumour evolution and therapy resistance in oncogene-driven NSCLC.
112 Finally, patient-derived NSCLC models and clinical specimens were also examined to further
113 establish the role of A3B in tumour evolution and therapy resistance.
114
115 **A3B is detrimental to tumour initiation in an EGFR mutant lung cancer mouse model.**
116 The role of A3B in tumour initiation was first investigated in a mouse strain that we engineered
117 by combining a novel Cre-inducible model for human *A3B* expression (*Rosa26::LSL-A3Bi*)[7] with
118 an established $EGFR^{L858R}$-driven lung cancer mouse model (*TetO-EGFR*^{L858R}) [13,14,22] and then
119 crossing with a Cre-inducible tetracycline controlled transactivator strain (*R26*$^{LNL\text{-}tTA}$). Tumours
120 were induced in *TetO-EGFR*^{L858R}*;R26*$^{LNL\text{-}tTA/+}$ *(E)* and *TetO-EGFR*^{L858R}*;R26*$^{LNL\text{-}tTA/LSL\text{-}APOBEC3B}$
121 (*EA3B)* mice using viral vector mediated Cre recombinase (Fig. 1a) and measured at 3 months
122 post-induction. Unexpectedly, the tumour number per mouse, total tumour volume per mouse, and
123 the fraction of mice with tumours was significantly lower in *EA3B* mice compared to *E* control
124 mice as measured by microCT imaging (Fig. 1b-c and Extended Data Fig. 1a). To verify this result,

125 and to assess tumour burden at a higher resolution, the experiment was repeated but animals were
126 culled at 3 months and tumour burden was assessed by immunohistochemistry (IHC) staining of
127 $EGFR^{L858R}$. This revealed a significantly decreased number of $EGFR^{L858R}$-positive cells per lung
128 area per mouse in *EA3B* mice compared to *E* control mice, substantiating the detrimental effect of
129 continuous *A3B* expression at tumour initiation (Fig. 1d). This was further confirmed by staining
130 for the programmed cell death marker caspase-3, which showed a significantly higher level of
131 tumour cell death in *EA3B* mice (Fig. 1e-f).
132
133 We hypothesized that enforced *A3B* expression at tumour initiation in *EA3B* mouse models may
134 induce increased CIN, p53 pathway activation, and cell death. This hypothesis is based on recent
135 work from our group revealing in a p53-deficient $EGFR^{L858R}$ mouse model that expression of the
136 human *A3B* transgene resulted in a significant increase in CIN[4]. In our models, a significantly
137 higher fraction of lagging chromosomes and chromatin bridges were observed in anaphase tumour
138 cells of *EA3B* mice compared with *E* mice at 3 months post-initiation[4] (Fig. 1g). At this same early
139 timepoint, there was also a significant increase in p53 nuclear positivity in tumours of *EA3B* mice
140 compared with *E* mice (Fig. 1h, and Extended Data Fig. 1b). No difference was observed in
141 proliferation (Ki67+ cells) or in the level of the DNA damage marker γH2AX+ between *E* and
142 *EA3B* mice (Extended Data Fig. 1c-d). Taken together, these data suggest that enforced *A3B*
143 expression impedes tumourigenesis through increased CIN, p53 pathway activation, and cell
144 death.
145
146 We hypothesized that enforced *A3B* expression at tumour intiation could also drive the detrimental
147 effect on tumour growth through enhanced immune surveillance in response to ongoing
148 mutagenesis[23]. To explore this, CD4 and CD8 T-cell infiltrates were examined at early (3 months
149 post induction) and late timepoints (at termination, ~6-17 months). At the early timepoint, a
150 significant increase in both CD4 (Extended Data Fig. 1e) and CD8 T-cells (Extended Data Fig. 1f-
151 g) in *EA3B* mice compared with *E* control mice was observed. In contrast, at the late timepoint a
152 lower level of both CD4 and CD8 T-cells was observed (Extended Data Fig. 1h-i) compared to the
153 early timepoint, and a significant increase in CD4 (Extended Data Fig. 1h), but not CD8 T-cells
154 (Extended Data Fig. 1i) in *EA3B* mice compared with *E* control mice was observed. These data
155 suggest immune evasion mechanisms to enforced *A3B* expression as tumours progress[24].

156

157 To examine if the connection between *A3B* and CIN observed in our mouse models was also
158 present in patients with lung adenocarcinoma from the TRACERx421 (TX421) cohort, the
159 correlation between *A3B* expression and multiple measures of CIN was assessed, expanding on
160 previously published analysis of the TRACERx100 (TX100) dataset. First the CIN70 score in
161 TX421 patient samples was examined. CIN70 score is an estimate of the extent of CIN in a sample
162 based on a 70-gene CIN signature[4,25]. Confirming and expanding on previous findings from
163 TX100[4], a significant correlation between *A3B* expression and CIN70 score was observed in the
164 TX421 cohort at patient tumour level (Extended Data Fig. 2a) and tumour region level (Extended
165 Data Fig. 2b). Although CIN70 can sometimes be an imperfect measure of CIN that also reflects
166 tumour proliferation status[4,25], we gained confidence in its utilisation after further confirming a
167 positive correlation between the percent missegregation errors quantified using H&E slides in the
168 anaphases of TRACERx tumours (TRACERx300, 300 H&E patient samples) and matched CIN70
169 score (Spearman $p = 0.0038$) (Extended Data Fig. 2c). A correlation between *A3B* expression and
170 subclonal CIN fraction was also observed in TX421 patients (Spearman $p = 0.00087$) (Extended
171 Data Fig. 2d).
172

173 When subsetting for patients with EGFR mutant lung adenocarcinoma in this expanded TX421
174 dataset, significant positive correlations were observed between *A3B* expression and multiple
175 measures of CIN at both the patient tumour level (Extended Data Fig. 2e) and tumour region level
176 (Extended Data Fig. 2f). A trend towards significance between the percent missegregation errors
177 quantified using haematoxylin and eosin (H&E) slides in the anaphases of TRACERx tumours
178 (TRACERx300 cohort) and matched CIN70 score (Extended Data Fig. 2g) was found, and a
179 significant correlation between *A3B* expression and subclonal CIN fraction was also observed
180 (Extended Data Fig. 2h). Finally, a significant correlation between the percentage of the genome
181 altered subclonally and APOBEC-driven mutations (TCN/TCW C>T/G) in patients with EGFR
182 mutant lung adenocarcinoma was shown (Extended Data Fig. 2i). These data together suggest that
183 the increased CIN observed with *A3B* expression in EGFR mutant mouse models is reflected in
184 NSCLC patients in the TX421 dataset.
185

186 Despite the detrimental effect observed at tumour initiation in *EA3B* mice, overall survival was
187 not significantly longer in the *EA3B* model compared with the *E* model (Fig. 1i). At a late
188 timepoint (termination), no significant difference in the fraction of chromosome missegregation
189 errors (Extended Data Fig. 3a) or p53 activation between *E* and *EA3B* mice was observed (Fig. 1j
190 and Extended Data Fig. 3b-g). These data suggest that tumours can overcome initial growth
191 disruption by *A3B* and tolerate its continuous expression, through selection for a lower fraction of
192 chromosome missegregations in anaphase tumours of *EA3B* mice (Extended Data Fig. 3a). This
193 kind of adaptation to excessive CIN through alterations that optimize cancer cell fitness in tumour
194 evolution, was shown in a previous publication from our lab illustrating that disruption of APC/C
195 activity occurs as a result of excessive CIN, to reduce CIN and restore equilibrium[26].
196
197 Loss of *A3B* expression was ruled out as an explanation for this lack of difference in survival, as
198 no significant difference was observed in *A3B* expression levels by IHC analyses between tumours
199 at 3 months post-induction and tumours at termination (~6-17 months) (Extended Data Fig. 4a-b).
200 Additionally, assessment of the recombination efficiency of the *LSL-A3Bi* transgene revealed a
201 similar fraction (~80%) of recombination in *EA3B* tumours harvested from mice at endpoint
202 compared to the fraction of tumours expressing *A3B* by IHC assessment at early timepoints
203 (Extended Data Fig. 4c), supporting that *A3B* expression is not lost with tumour progression in this
204 model.
205
206 To assess if inactivation of the p53 pathway is sufficient to decrease the detrimental effect observed
207 at tumour initiation with continuous *A3B* expression, tumours were induced in an $EGFR^{L858R}$ p53-
208 deficient mouse model either with or without *A3B* expression (*EP* and *EPA3B*) (Extended Data
209 Fig. 5a). No difference in either the number of tumours at 3 months post-induction (Extended Data
210 Fig. 5b), or in overall survival (Extended Data Fig. 5c) between *EP* and *EPA3B* mice was observed.
211 These data suggest that in a p53-deficient model of EGFR-driven lung cancer, continuous *A3B*
212 expression is tolerated. Thus, p53 in this model limits the tolerance of cancer cells to the
213 detrimental effects of early and continuous A3B expression during tumour initiation.
214
215 **Subclonal *A3B* inhibits tumour progression and increases overall survival in an EGFR**
216 **mutant lung cancer mouse model.**

217 Data from analysis of the TRACERx421 (TX421) cohort revealed that APOBEC-mediated
218 mutagenesis is enriched subclonally during tumour evolution in *EGFR* mutant disease (Extended
219 Data Fig. 6a-b) and the wider cohort[11]. Based on these findings from TX421 and on our findings
220 that enforced clonal expression of *A3B* was detrimental at tumour initiation, we generated mice in
221 which *A3B* expression could be temporally separated from $EGFR^{L858R}$ expression to mirror
222 subclonal APOBEC induction (Extended Data Fig. 7a). A tetracycline-inducible pneumocyte
223 specific mouse model of $EGFR^{L858R}$-dependent lung cancer[13,14,22] was combined with a Cre
224 recombinase-steroid receptor fusion ($Rosa26CreER^{T2}$)[27], with or without the conditional human
225 *APOBEC3B* minigene used above to generate $CCSP\text{-}rtTA;TetO\text{-}EGFR^{L858R};Rosa26^{CreER(T2)/LSL\text{-}}$
226 ^{APOBEC3B} (*EA3Bi* or *Ei*) mice (Extended Data Fig. 7a). As expected, subclonal A3B-positive tumour
227 cells were observed in *EA3Bi* mice but not in *Ei* control mice, reproducing the subclonal expression
228 of *A3B* in human NSCLC (Extended data Fig. 7b).
229
230 In this model, $EGFR^{L858R}$-positive hyperplasias pervade the lung tissue, forcing the termination of
231 mice prior to progression of the majority of $EGFR^{L858R+}$ lesions. Despite this, a fraction of
232 $EGFR^{L858R}$-positive hyperplasia progress to form tumour nodules. Tumour nodules per lung
233 section and tumour area per lung area were measured and found to be significantly lower in *EA3Bi*
234 mice compared with *E* control mice at termination (all culled within 2 weeks of one another)
235 (Extended Fig. 7c-d). Lung weight and $EGFR^{L858R}$ positivity per lung area, measures that include
236 the $EGFR^{L858R}$-positive hyperplasia, were quantified at termination in *Ei* and *EA3Bi* mice and no
237 significant difference was found (Extended Data Fig. 7e-f). These data suggest that *Ei* and *EA3Bi*
238 mice have a similar tumour burden at termination, but that a higher fraction of the $EGFR^{L858R}$-
239 positive hyperplasia in *Ei* control mice had progressed to form tumour nodules. Overall survival
240 in *Ei* and *EA3Bi* mice was also measured, and a significant increase in survival of *EA3Bi* mice
241 compared with *E* control mice was observed (Extended Fig. 7g). These data suggest that subclonal
242 *A3B* in p53 wildtype mice also has a detrimental effect on tumourigenesis substantiating the
243 unfavourable effects of continuous *A3B* expression at tumour initiation in a p53 wild-type
244 background.
245
246 **Subclonal A3B promotes tyrosine kinase inhibitor (TKI) resistance**

247 Having established that continuous clonal or subclonal expression of *A3B* is detrimental to tumour
248 initiation in p53 wildtype; *EGFR*^{L858R} mouse models, the effect of *A3B* on tumour evolution under
249 the selective pressure of targeted therapy was examined next. *A3B* was expressed subclonally in
250 the tumours of mice treated with TKI therapy in our *EA3Bi* mouse model (Fig. 2a). A significant
251 increase in tumour grade (Fig. 2b), tumour nodules per lung section (Fig. 2c), and tumour area per
252 tissue area (Fig. 2d) was observed in TKI-treated *EA3Bi* mice compared with TKI-treated *Ei*
253 control mice. Heterogeneous A3B tumour positivity was observed, consistent with subclonal
254 induction of A3B (Fig. 2e), and a significant increase in A3B positivity with TKI therapy compared
255 to untreated *EA3Bi* mice was also observed (Fig. 2f), suggesting a selection for *A3B* expression
256 with TKI therapy.
257
258 Tumour growth and progression was also compared in *Ei* versus *EA3Bi* mice in an additional
259 experiment with continuous administration of TKI therapy (erlotinib) to more closely model the
260 treatment regimens of patients with lung cancer (Fig. 2g). The mice were culled once the maximum
261 oral gavage number had been reached at 10 months, and a significant increase in tumour nodules
262 (Fig. 2h) and an increase in tumour grade (Fig. 2i) were observed in *EA3Bi* mice compared with
263 *Ei* control mice.
264
265 Based on previous work illustrating an important role for uracil DNA glycosylase (UNG) in
266 repairing APOBEC-induced uracil lesions[28], we evaluated UNG expression in *A3B* expressing
267 *EA3Bi* tumours. IHC staining for UNG revealed a significant decrease in UNG-positive cells per
268 tumour in *EA3Bi* mice compared with *Ei* mice treated with TKI therapy (Extended Data Fig. 8a-
269 b). Taken together, these findings suggest that subclonal *A3B* expression with TKI therapy in
270 conjunction with *UNG* downregulation contributes to increased tumour growth and TKI resistance
271 in EGFR mutant mouse models.
272
273 To validate our findings from the mouse models in human systems, cell viability under targeted
274 therapy was assessed in established patient-derived cell line models of oncogenic *EGFR*-driven
275 lung adenocarcinoma in which *A3B* was depleted in both clonal and polyclonal populations using
276 RNAi or CRISPR-Cas9 technologies. Depletion of *A3B* by shRNA treatment resulted in a
277 significant decrease in cell viability upon treatment with the EGFR TKI osimertinib compared to

278 an empty vector control in two EGFR mutant cell lines, HCC827 (*EGFR*^{exon19del}) (Extended Data
279 Fig. 9a-b) and H1975 (*EGFR*$^{L858R+T790M}$) (Extended Data Fig. 9c-d), suggesting that *A3B* confers
280 enhanced survivability during targeted therapy. To further verify this, multiple single cell clones
281 were generated from PC9 cells (*EGFR*^{exon19del}) with CRISPR depletion of *A3B*, which also showed
282 a significant decrease in cell viability with *A3B* loss upon treatment (Extended Data Fig. 9e-f).
283 Additionally, a similar result was also observed in the *EML4-ALK* cancer cell line H3122 treated
284 with crizotinib (Extended Data Fig. 9g-h) suggesting that these observations are not limited to
285 EGFR-mutant lung adenocarcinoma models. These data in human cell line models support our *in*
286 *vivo* mouse findings that *A3B* expression is advantageous to tumour growth and resistance during
287 TKI therapy.
288
289 **Targeted therapy induces A3B upregulation and UNG downregulation.**
290 Our mouse lung cancer models demonstrated that *UNG* expression is downregulated with targeted
291 therapy, and that *A3B* expression is selected for and promotes targeted therapy resistance. We
292 hypothesized that targeted therapy may also induce adaptations that increase *A3B* expression and
293 decrease *UNG* expression in human models. Targeted therapy-induced transcriptional changes of
294 both *A3B* and *UNG* were assessed in established human lung cancer cell line models present in
295 publicly available datasets (GEO2R). Treatment of multiple EGFR-mutant cell lines (HCC827 and
296 PC9 harbouring *EGFR*^{exon19del} and HCC4006 harbouring EGFR$^{L747\text{-}E749del,\ A750P}$) with the EGFR
297 TKI erlotinib was associated with transcriptional upregulation of *A3B* and downregulation of the
298 repair factor *UNG*, both acutely (6-hour to 1-day treatment) and at later timepoints (8-day
299 treatment) (Extended Data Fig. 10a). To confirm these transcriptional changes, an independent
300 RNA-seq dataset was examined[29], generated from EGFR-mutant PC9 cells treated with the EGFR
301 TKI osimertinib for 9 days. These EGFR TKI-treated cells also displayed a significant
302 upregulation of *A3B* and downregulation of *UNG* transcripts (Extended Data Fig. 10b). The
303 transcriptional effects identified were therefore a conserved consequence of EGFR pharmacologic
304 inhibition in *EGFR*-driven lung adenocarcinomas and not specific to the class of EGFR inhibitor,
305 as similar effects were observed with a first generation TKI (erlotinib) (Extended Data Fig. 10a)
306 and a third generation TKI (osimertinib) (Extended Data Fig. 10b). A3B induction was observed
307 with either lower physiological doses of erlotnib or osimertinib (data not shown), or higher doses
308 previously used to study clinically relevant mechanisms of acquired resistance[30-35]. These findings

309 were subsequently validated in multiple oncogenic *EGFR*-driven (Fig. 3a-b) and *EML4-ALK*-
310 driven (Fig. 3c) cellular models of lung adenocarcinoma at both the RNA (Fig. 3a-c) and protein
311 level (Fig. 3d, Extended Data Fig. 11g). To rule out the possibility that these observations could
312 be due to off-target pharmacological effects of EGFR TKIs, genetic inhibition of EGFR through
313 siRNA-mediated silencing in an oncogenic EGFR-driven model was performed, confirming
314 significant *A3B* upregulation and *UNG* downregulation (Fig. 3e). Furthermore, induction of A3B
315 was also observed upon treatment with an inhibitor of mitogen-activated protein kinase enzymes
316 MEK1, selumetinib, (Extended Data Fig. 10a). The induction of A3B by different inhibitors of
317 oncogenic RTKs and their downstream signalling components, such as MEK1, indicates that
318 upregulation of A3B is likely a consequence of inhibition of oncogenic signalling, and not specific
319 to EGFR TKIs. TKI-induced changes at the functional level were validated using established
320 assays to quantify APOBEC activity[36] and uracil excision capacity[28]. Consistent with RNA and
321 protein level changes, TKI treatment resulted in a significant increase in nuclear APOBEC activity
322 and a decrease in nuclear uracil excision capacity of UNG in multiple *EGFR*-driven cell line
323 models of lung adenocarcinoma, including $EGFR^{exon19del}$ cells (PC9 and HCC827), and
324 $EGFR^{L858R+T790M}$ cells (H1975) (Fig. 3f-g, Extended Data Fig. 10c-f). Increased *A3B* expression
325 and activity and decreased *UNG* expression and activity were also observed in *EML4-ALK*-driven
326 cellular models (H3122 and H2228) (Extended Data Fig. 10g-j).
327
328 To exclude an indirect effect of targeted therapy on cell cycle arrest which might alter APOBEC
329 and DNA repair enzyme expression, PC9 cells were treated with the CDK4/6 cell cycle inhibitor
330 palbociclib[37] to induce a G0/G1 cell cycle arrest and *A3B* and *UNG* expression levels were
331 measured. Palbociclib treatment induced G0/G1 cell cycle arrest as expected (Extended Data Fig.
332 11a). While *UNG* expression decreased upon palbociclib treatment, there was a substantial decline
333 in *A3B* expression, in contrast to the increased expression observed upon TKI therapy (Extended
334 Data Fig. 11b). These data suggest that TKI-mediated induction of *A3B* is unlikely to be a
335 consequence of TKI treatment-induced cell cycle inhibition.
336
337 To confirm our findings, *A3B* and *UNG* expression levels were examined in multiple human
338 tumour xenograft models: three cell line-derived tumour xenografts (CDXs) and one patient-
339 derived tumour xenograft (PDX). An increase in A3B protein levels and a decrease in UNG protein

340 levels were detected by western blot and immunohistochemistry (IHC) analyses in EGFR TKI-
341 treated tumour tissues from three distinct oncogenic *EGFR*-driven CDX models of human lung
342 adenocarcinoma: osimertinib-treated H1975 xenografts ($EGFR^{L858R+T790M}$) (Fig. 3h), osimertinib-
343 treated PC9 xenografts ($EGFR^{exon19del}$) (Extended Data Fig. 12a) and erlotinib-treated 11-18
344 xenografts ($EGFR^{L858R}$) (Extended Data Fig. 12b-e). Additionally, RNA-seq analyses from an
345 $EGFR^{L858R}$-harbouring PDX model of lung adenocarcinoma[38] treated with the EGFR inhibitor
346 erlotinib (Extended Data Fig. 12f) or the EGFR inhibitor osimertinib[29] (Extended Data Fig. 12g)
347 also revealed an increase in *A3B* mRNA levels and a trend towards decrease in *UNG* mRNA levels
348 upon treatment. The collective findings from these multiple independent human tumour xenograft
349 cohorts support a model whereby EGFR oncoprotein inhibition induces *A3B* upregulation and
350 *UNG* downregulation.
351
352 To examine if the increase in *A3B* expression observed during TKI therapy is directly linked to
353 A3B activity, nuclear APOBEC activity was measured during both acute and long term TKI
354 therapy. Osimertinib treatment for up to 9 days resulted in a periodic induction of APOBEC
355 activity, with peaks corresponding to over 4-fold increase compared to non-treated cells at day 1,
356 and an almost 5-fold increase at the 8-day timepoint (Extended Data Fig. 13a-b). The pattern of
357 *A3B* expression upon TKI treatment was assessed and was found to be significantly correlated
358 with the dynamics of nuclear APOBEC activity (Extended Data Fig. 13c). Accordingly, both
359 baseline and TKI-induced nuclear APOBEC activity were abrogated by RNAi depletion of *A3B* in
360 the $EGFR^{exon19del}$ human lung cancer cell line PC9, which was confirmed with both acute siRNA
361 (Extended Data Fig. 14d-f) and stable shRNA methods (Extended Data Fig. 13g-h). To further
362 confirm the predominant role of A3B in APOBEC activity, following the depletion of *A3B*, either
363 wild-type A3B (A3B WT-HA tagged) or a catalytically inactive A3B containing a mutation in the
364 C-terminal deaminase domain (A3B E225A-HA tagged) were reconstituted[39]. Accordingly,
365 expression of the wild-type catalytically active, but not of the mutant catalytically inactive A3B,
366 rescued the decline in nuclear APOBEC activity caused by *A3B* depletion (Extended Data Fig.13i-
367 j). The collective findings support a model whereby EGFR targeted therapy induces conditions in
368 cancer cells that are favourable for *A3B* expression and activity.
369
370 **NF-κB signalling is critical for TKI-induced A3B upregulation.**

371 Previous work from us and others revealed that NF-κB signalling is activated upon EGFR
372 oncogene inhibition in human lung cancer as a stress and survival response[38], and that NF-κB
373 signaling may be one prominent inducer of *A3B* gene expression[40,41]. To test the hypothesis that
374 NF-κB signaling that is activated upon targeted therapy and promotes *A3B* upregulation, an
375 established RNA-seq dataset generated from *EGFR*-driven human lung adenocarcinoma cells
376 treated acutely with either erlotinib or an established NF-κB inhibitor, PBS-1086, or both in
377 combination was examined[38]. TKI treatment-induced transcriptional upregulation of *A3B* was
378 attenuated by co-treatment with the NF-κB inhibitor[38] (Extended Data Fig. 14a). Upon TKI
379 treatment, an increase in nuclear levels of NF-κB subunit RELB tracked in synchrony and
380 episodically with *A3B* expression, peaking at 1 and 7 days (Extended Data Fig. 14b). These data
381 suggest that the NF-κB pathway may induce *A3B* expression. To confirm this, first the NF-κB
382 pathway was activated with increasing concentrations of TNFα, which elevated nuclear RELA and
383 RELB as well as A3B protein levels (Extended Data Fig. 14c) and cellular *A3B* mRNA expression
384 (Extended Data Fig. 14d). Furthermore, inhibition of the NF-κB pathway by simultaneous
385 depletion of both RELA and RELB (Extended Data Fig. 14e) reduced TKI-induced *A3B* mRNA
386 expression (Extended Data Fig. 14f) and protein levels (Extended Data Fig. 14g). Consistent with
387 these data from *in vitro* studies, co-inhibition of EGFR and NF-κB pathways blocked EGFR
388 inhibition-induced *A3B* upregulation in oncogenic EGFR-driven NSCLC xenografts (Extended
389 Data Fig. 12c). To determine if NF-κB could promote TKI-induced A3B-driven nuclear APOBEC
390 activity, we examined the effects of co-depletion of RELA and RELB transcription factors on TKI-
391 induced nuclear APOBEC activity. Consistent with our hypothesis, co-depletion of both NF-κB
392 transcription factors impaired TKI-induced nuclear APOBEC activity (Extended Data Fig. 14h).
393 These data support NF-κB activation resulting from EGFR TKI treatment as an inducer of *A3B*
394 upregulation in response to therapy. The intermittent nature of NF-κB activation is a classic feature
395 of NF-κB signalling[42-46], and likely leads to the downstream episodic A3B upregulation observed
396 during targeted therapy.
397
398 To investigate the clinical relevance of these findings, we examined single-cell RNA-seq data in
399 an established dataset obtained from clinical specimens of NSCLC procured from patients at 3
400 timepoints: 1) prior to initiation of systemic targeted therapy (classified as treatment naïve, TN),
401 2) while on targeted therapy when the tumour was regressing or at stable state as evaluated by

402 standard clinical imaging (classified as residual disease, RD), and 3) at clear progressive disease
403 (acquired resistance) as determined by standard clinical imaging (classified as progressive disease,
404 PD). The classification of response was based on RECIST criteria[47]. In total, 49 samples obtained
405 from patients pre- or post-TKI therapy (erlotinib, osimertinib and crizotinib being the most
406 frequent targeted therapies) from 30 lung cancer patients were analysed (Supplementary Table 1a).
407 We observed that mRNA expression of A3B and of the NF-κB components and targets RELA and
408 RELB, as well as an NF-κB gene signature[48], were significantly increased in tumours exposed to
409 EGFR TKI treatment, in particular at tumour progression during therapy (Extended Data Fig. 15a-
410 d).
411
412 ***UNG* downregulation is associated with *c-JUN* suppression during TKI treatment.**
413 Having noted that a key DNA repair enzyme UNG, that counteracts APOBEC-mediated
414 mutagenesis was downregulated during targeted therapy in both our mouse and pre-clinical
415 models, we investigated the potential mechanism by which this occurs. *UNG* gene promoter
416 analysis (using PROMO)[49] revealed the presence of predicted c-Jun consensus binding sites. Our
417 RNA-seq data from EGFR TKI-treated PC9 EGFR-mutant NSCLC cells indicated that like *UNG*,
418 *c-JUN* was also transcriptionally downregulated upon TKI treatment, which we validated using
419 RT-qPCR analysis (Extended Data Fig. 16a). These data are consistent with the expectation that
420 *c-JUN* expression should decrease upon inhibition of the MAPK pathway that occurs as a result
421 of EGFR inhibition by TKI treatment[50]. Thus, the possibility that TKI treatment-induced *UNG*
422 downregulation might be caused by *c-JUN* downregulation was investigated. Consistent with this
423 hypothesis, *c-JUN* depletion alone was sufficient to suppress *UNG* expression, suggesting that
424 *UNG* downregulation may be a consequence, in part, of the *c-JUN* suppression that occurs during
425 MAPK signalling suppression that is a classic effect of TKI treatment (Extended Data Fig. 16b).
426
427 **A3B generates genetic diversity that helps to fuel resistance to oncogene-targeted therapy.**
428 We next tested the hypothesis that *A3B* upregulation and activity (and *UNG* downregulation) could
429 create conditions conducive to cancer genome mutagenesis, genetic diversity, and targeted therapy
430 resistance. We sought to identify evidence for this and putative TKI resistance mutations induced
431 by *A3B* in both pre-clinical human and mouse models together with human NSCLC specimens.

432 Whole exome and whole genome sequencing (WES and WGS) was performed in our cell line
433 models and patient datasets collected from current and prior studies[31,51].
434
435 To examine the effects of TKI-induced A3B expression and activity on mutagenesis, doxycycline-
436 inducible CRISPR-Cas9 depletion of A3B was introduced in multiple clones of the human PC9
437 *EGFR*^{exon19del} cell line (Fig. 4a, and Extended Data Fig. 17a). WES analysis of these cells revealed
438 that TKI treatment of A3B-proficient PC9 cells for close to 3.5 months induced a significantly
439 higher mutation burden and a significantly higher number APOBEC-associated mutations in TKI-
440 treated samples (Fig. 4a). WGS sequencing analysis was also performed on an additional A3B-
441 proficient and A3B-deficient PC9 clone, treated with DMSO or with Osimertinib, again illustrating
442 a higher mutational burden and higher number of APOBEC-associated mutations specifically in
443 A3B-proficient osimertinib-treated sample and not in the A3B-proficient DMSO treated sample
444 (Fig. 4b, and Extended Data Fig. 17b-c). This data suggests that the accumulation of mutations in
445 an APOBEC context in the A3B-proficient osimertinib-treated sample is likely due to the induction
446 of A3B expression upon treatment with TKI therapy, rather than intermittent mutagenesis that in
447 some cases might naturally occur[52].
448
449 In the EGFR mutant human lung cancer cell line PC9, an A3B-proficient PC9 TKI resistant clone
450 acquired an early stop codon mutation in the tumour suppressor gene NRXN3 (Q54*)[53,54] in the
451 APOBEC-preferred context (T[C>T]A). The impact of this potential loss-of-function change was
452 validated by depleting NRXN3 (given the early stop codon mutation detected, which is likely a
453 loss-of-function event) in a naïve PC9 lung cancer cell line, which increased levels of
454 phosphorylated AKT and conferred resistance to EGFR TKI treatment (Extended Data Fig. 18a-
455 c).
456
457 The increase in mutational burden and APOBEC-associated mutations with targeted therapy was
458 confirmed again in mouse tumour cell lines derived from *EGFR*^{L858R} p53-deficient *A3B*-expressing
459 mouse models (Extended Fig. 6a, *EP* and *EPA3B* naïve mouse tumour cell lines), treated with
460 targeted therapy (erlotinib or osimertinib). A significantly higher number of mutations overall were
461 detected in TKI resistant A3B expressing EGFR mutant tumour cell lines (*EPA3B)* compared with
462 both treatment naïve *EPA3B* tumour cell lines and TKI resistant cell lines lacking A3B expression

463 (EP) (Extended Data Fig. 19a). A significantly higher number of APOBEC-associated mutations
464 were also detected in TKI resistant *A3B*-expressing *EPA3B* cell lines compared with TKI resistant
465 *EP* cell lines (Extended Data Fig. 19b). No EGFR T790M mutations were detected in any of the
466 naïve or TKI resistant mouse tumour cell lines. These data in mouse preclinical models strengthen
467 the hypothesis that A3B can promote genetic diversity during targeted therapy treatment, as these
468 models lack all other human APOBEC3 family enzymes.
469
470 In two of the TKI-resistant *A3B*-expressing mouse lung cancer cell lines, two unique A3B-
471 mediated de novo putative loss-of-function mutations in the *Ptprs* gene were identified (Extended
472 Data Fig 19c). Loss of PTPRS function through mutation or deletion has been shown to
473 significantly increase levels of active phosphorylated EGFR and is linked to resistance to targeted
474 therapy in human preclinical cancer models[55-57]. Additionally, loss of PTPRS expression has been
475 shown to drive significantly more rapid disease progression and worse overall survival in patients
476 with EGFR-driven lung adenocarcinoma[55-57]. Both *Ptprs* mutations were in the second Ig domain
477 that is conserved between the mouse and human PTPRS[58]. Loss of Ig domains in PTPRS has been
478 shown to reduce the phosphatase activity of PTPRS resulting in increased phospho-ERK levels[59].
479 Ig domains are conserved across all three human type IIa PTPs, including PTPRS, PTPRD, and
480 PTPRF[58]. One of the A3B-driven mutations in the mouse *Ptprs* gene, D138N, (Ptprs_mut1,
481 Extended Fig 19c) was also identified in patients with lung, colorectal, and bladder cancer,
482 including one lung cancer TRACERx patient with an $EGFR^{L858R}$ driver mutation. Subclonal
483 enrichment of APOBEC-associated mutation signatures 2 and 13 was observed in this TRACERx
484 patient compared with all other patients with EGFR mutant lung adenocarcinoma in the TX421
485 cohort dataset (Extended Data Fig. 19d). These findings suggest that the A3B-driven D138N
486 mutation could be a novel APOBEC hotspot mutation.
487
488 To verify the clinical relevance of our findings, *A3B* expression was also examined in the NSCLC
489 clinical dataset at TN, RD and PD timepoints (Supplementary Table 1b)[47,60]. Over 50% of human
490 tumours that were exposed to targeted therapy showed increased *A3B* expression with a significant
491 increase at PD ($P = 0.0318$) (Fig. 4c). No significant difference was observed between PD tumours
492 treated with single agent TKI therapy compared with PD tumours treated with single agent TKI
493 therapy combined with chemotherapy (Extended Data Fig. 20a). IHC analyses performed on these

494 clinical samples revealed a significant increase in A3B protein levels in EGFR TKI-treated tumour
495 samples both at RD and PD timepoints (P = 0.0113) (Fig. 4d-e). Thus, A3B was elevated not only
496 in human and mouse cellular models of oncogene-driven lung cancer but also in human clinical
497 tumours exposed to targeted therapy.
498
499 To further test whether *A3B* induction during targeted therapy fuels resistance development, the
500 overall mutation burden (SNV count) (Fig. 4f), and number of APOBEC-associated mutations
501 (C>T or C>G mutations in a TCN context) (Fig. 4g) examined in patients prior to and post-TKI.
502 Consistent with the increase of *A3B* expression, a significant increase was observed in both overall
503 mutation burden and APOBEC-associated mutations post-treatment (Fig. 4f-g). Moreover, known
504 APOBEC signature-associated mutations were identified in genes previously associated with TKI
505 resistance in tumours of patients who had progressed on or shown incomplete response to EGFR
506 inhibitor therapy (Fig. 4h). These mutations include an activating mutation in *PIK3CA* (E545K)[61],
507 WNT signalling-activating mutations in β-catenin at a GSK-3β phosphorylation site[62], MAPK
508 pathway reactivating-mutations through inactivation of *PP2A*, a negative regulator of MAPK
509 signalling[63,64], as well as an ALK inhibitor desensitizing mutation in *ALK* (E1210K)[65] in the
510 tumours of some patients who had progressed on or shown incomplete response to EGFR or ALK
511 inhibitor therapy (Fig. 4h). AKT, WNT and MAPK pathway activation have previously been
512 shown to cause EGFR and ALK inhibitor resistance[66-71].
513
514 All but one of these APOBEC-associated putative resistance mutations were detected selectively
515 post-treatment in tumours that have developed resistance, suggesting not only that these mutations
516 are induced by APOBEC during targeted therapy but also that they could in fact be promoting
517 resistance. Furthermore, all samples containing these APOBEC-associated mutations except for
518 one did not also harbour the EGFR T790M mutation, which has been reported to be present in
519 approximately 50 to 60% of first- and second-generation EGFR TKI-resistant cases[72,73] and arising
520 from a non-APOBEC clock-like mutation signature (Signature 1[74]) (Fig. 4h). The observation that
521 APOBEC-mediated mutations in resistance genes detected in post-treatment samples and the
522 EGFR T790M mutation appear to be mutually exclusive suggests that these APOBEC-mediated
523 mutations could be the principal mechanisms of resistance to targeted therapy in these patients.
524

525 Taken together, these data illustrate the diverse effects of A3B at different stages of tumour
526 evolution with or without the selective pressure of therapy, demonstrating the dual role of A3B,
527 both as an inhibitor of tumour progression at initiation, and a contributor to genetic diversity that
528 helps to promote resistance during the selective pressure of targeted therapy (Extended Data Fig.
529 21a-b)
530
531 **Discussion**:
532 Our collective findings shed light on the important, context-specific roles of APOBEC on lung
533 cancer pathogenesis and tumour evolution, from initiation to resistance development. Our mouse
534 models and human datasets suggest that A3B-driven adaptive mutability[8] is specific to a TKI
535 therapeutic context; for instance in the mouse models, enforced A3B expression in the absence of
536 TKI therapy at initiation promotes increased T-cell infiltrate, as well as increased CIN, p53
537 activation and cell death, resulting in reduced tumour growth. Later in tumour evolution in the
538 mouse models, we no longer observe a significant difference in p53 activation or CIN between *E*
539 and *EA3B* mice (Fig. 1j and Extended Data Fig. 3a-g), which could be due to selection for a lower
540 chromosome missegregation rate to optimize cell fitness enabling survival of A3B expressing
541 cells. As mentioned previously, attenuation of the chromosome segregation error rate has been
542 shown to occur over time in the presence of persistent spindle assembly checkpoint dysfunction
543 due to selection of APC/C mutations in cell models[26]. We also see a reduced immune infiltrate
544 overall and no significant difference in cytotoxic CD8 T cell infiltrate between *E* and *EA3B* mice
545 later in tumour evolution, in contrast to significant increase in both CD4 and CD8 T-cells in EA3B
546 at tumour initiation (Extended Data Fig. 1e-i). These data may reflect selection for adaptive
547 immune evasive mechanisms over time with enforced A3B expression[75].
548
549 Our data reinforce the idea that targeted therapies can induce changes that promote genome
550 mutagenesis, including that which is APOBEC-mediated, which was not previously appreciated.
551 This A3B induction during therapy helps to fuel the development of treatment resistance and
552 appears to be clinically relevant based on our clinical specimen datasets obtained from targeted
553 therapy-treated patient datasets, which are challenging to obtain and analyse. Based on our findings
554 here, additional clinical cohort analysis will be important to conduct as further human tumours
555 obtained from patients on targeted therapy become available in the future.

556

557 While APOBEC has been implicated in drug resistance previously[39,76], our study reveals a distinct
558 mechanism by which targeted cancer therapy is actively responsible for the upregulation of
559 APOBEC via NF-κB-mediated transcriptional induction that occurs in response to therapy. This
560 process then helps to promote tumour genetic evolution to facilitate the development of resistance.
561 Examination of single-cell RNA-seq data of cancer cells isolated from clinical samples of patients
562 with NSCLC before or whilst on targeted therapy[47] revealed, similar to A3B, significant increases
563 in the expression of the NF-κB transcriptional subunits RELA and RELB, and an established NF-
564 κB gene signature, in tumours exposed to EGFR TKI treatment (Extended Data Fig.15a-d),
565 reinforcing the relevance of NF-κB-mediated transcriptional upregulation of A3B as a mechanistic
566 feature of this evolutionary process. Our study increases the understanding of why co-treatment
567 with an NF-κB pathway inhibitor is more effective than EGFR inhibitor treatment alone at
568 preventing the emergence of resistance, as previously observed in multiple *in vitro* and *in vivo*
569 models of oncogenic-EGFR driven lung adenocarcinomas[38].

570

571 We demonstrated that expression of *A3B* helps to promote resistance in preclinical human and
572 mouse models of lung adenocarcinoma. Although we focused on oncogenic *EGFR*-driven lung
573 adenocarcinomas, our findings appear to extend to other molecular subsets as well such as *EML4-*
574 *ALK*-driven lung cancer (Fig. 3c and Extended Data Fig. 10g-j) and likely reflect a more general
575 principle of therapy-induced genetic evolution. Our analysis of patient data suggests an increase
576 in *A3B* expression in the post-TKI setting. We also identified specific APOBEC-associated
577 mutations that can promote resistance to targeted therapy. Our work is consistent with, and
578 expands upon, prior studies suggesting a potential association between APOBEC-mediated
579 mutagenesis and acquisition of putative resistance mutations in the APOBEC-preferred context
580 during treatment of EGFR-driven lung cancers[77,78]. The collective data suggest that inhibiting
581 APOBEC3B-driven tumour evolution could suppress the emergence of one key class of resistance
582 mutations and thereby improve response to targeted therapy. It is also possible and likely that other
583 APOBEC3 family members beyond A3B contribute to genetic evolution of resistance[5,52], although
584 A3B has been shown to be a contributor to mutagenesis in a human cell line model system[6]. We
585 must also not discount the potential contribution of chromosomal instability induced by A3B to
586 TKI resistance. We decided in this manuscript to focus specifically on TKI resistance mechanisms

587 driven by APOBEC-associated mutations but believe that disentangling the different contributions
588 of CIN and mutagenesis driven by APOBEC enzymes to TKI resistance is an important area of
589 future research. Finally, it is certain that APOBEC-mediated resistance is one of many paths to
590 resistance development.
591
592 Recent work has also demonstrated that decreased mismatch repair capacity and increased
593 expression of error-prone DNA polymerases may contribute to increased adaptability of colorectal
594 cancer to EGFR inhibition therapy[8]. While we observed downregulation of *UNG*, we did not
595 observe upregulation of these error-prone polymerases by gene expression analysis of EGFR TKI-
596 treated PC9 cells. Thus, our findings describe a distinct mode and mechanism of therapy-induced
597 genetic adaptability in cancer. Our evidence here and these emerging collective findings[5,8,39,79]
598 suggest that endogenous drivers of mutagenesis play diverse roles that are both detrimental and
599 beneficial to tumour evolution depending on context of tumour pathogenesis, and that treatment-
600 induced genetic mutability may be a general principle by which tumours evolve in response to
601 treatment.
602
603 **Author contributions**:
604 **Conception and design of the study**: D.R.C., M.K.M., P.G., E.K.L., N.K., R.S.H., J.D. T.G.B,
605 C.S.; **Data acquisition (cell line and animal studies)**: D.R.C., M.K.M., P.G., F.H., B.G., S.N.,
606 C.G., E.G., M. A. B., A.N., F.G.V., W.H., W.T.L., B.A., M.G., C.M., J.P., E.G., C.Z., S.L., J.C.,
607 B.R., W.B., A.R., B.A., R.I.V., M.M., N.J.T., T.J.H., C.W., N.K., S.V., K.V., S.H., V.O., D.B.;
608 **Data acquisition for clinical studies**: C.M.R., M.D., M.A., C.B., O.P., B.B., C.E.M., C.M.B.,
609 D.L.K., J.K.R., M.J.H., P.A., E.A.Y., L.T.; **Mutational Signature Analysis and/or other**
610 **computational analysis**: C.B., O.P., B.B., M.D., M.A., N.I.V., N.A.T., W.W., L.C., E.M.V.A.,
611 J.Y.; **Analysis and interpretation of data**: D.R.C,. M.K.M., P.G., N.I.V., T.G.B., C.S., E.K.L,
612 R.S.H., W.L.B, L.K.L., C.D., P.P.A., J.P., M.A.B., A.N., M.D., C.M.R., S.F.B., S.K.C., S.P.,
613 A.S.B., N.M., C.M., B.R., B.B., W.W., K.V., D.L.K, F.H., C.B., O.P., B.B., N.K., N.A.T., N.R.;
614 **Manuscript writing and revision**: D.R.C., M.K.M., P.G., M.D, M.A., C.B., O.P., B.B., K.V.,
615 N.M., E.M.V.A, N.K., R.S.H. J.D., T.G.B., C.S.
616
617 **Acknowledgements:**

618 This project is supported by the NIH/NCI U54CA224081, R01CA169338, R01CA211052,
619 R01CA204302, U01CA217882 and the Chan-Zuckerberg Biohub (to T.G.B). Pfizer, as well as
620 the University of California Cancer League (United States) (to C.E.M), AstraZeneca (United
621 Kingdom), The Damon Runyon Cancer Research Foundation P0528804 (United States), Doris
622 Duke Charitable Foundation P2018110 (United States), V Foundation P0530519 (United States),
623 and NIH/NCI R01CA227807 (to C.M.B.), F.H. was supported by the Mildred Scheel postdoctoral
624 fellowship from the German Cancer Aid. E.A.Y is supported by T32 HL007185 from the NHLBI.
625 Cancer studies in the Harris lab are supported in part by the National Cancer Institute P01-
626 CA234228. R.S.H. is the Ewing Halsell President's Council Distinguished Chair at University of
627 Texas San Antonio and an Investigator of the Howard Hughes Medical Institute.
628
629 D.R.C was supported by the Francis Crick Institute which receives its core funding form Cancer
630 Research UK (FC001169), the UK Medical Research Council (FC002269), and the Wellcome
631 Trust (FC001169), as well as an NC3Rs training fellowship (NC/S001832/1). C.S. acknowledges
632 grant support from Pfizer, AstraZeneca, Bristol Myers Squibb, Roche-Ventana, Boehringer-
633 Ingelheim, Archer Dx Inc (collaboration in minimal residual disease sequencing
634 technologies) and Ono Pharmaceutical, is an AstraZeneca Advisory Board member and Chief
635 Investigator for the MeRmaiD1 clinical trial, has consulted for Pfizer, Novartis,
636 GlaxoSmithKline, MSD, Bristol Myers Squibb, Celgene, AstraZeneca, Illumina, Genentech,
637 Roche-Ventana, GRAIL, Medicxi, Bicycle Therapeutics, and the Sarah Cannon Research
638 Institute, has stock options in Epic Bioscience, GRAIL, and has stock options and is co-founder
639 of Achilles Therapeutics. The other University of Minnesota authors have no competing interests
640 to declare. T.G.B. is an advisor to Novartis, Astrazeneca, Revolution Medicines, Array/Pfizer,
641 Springworks, Strategia, Relay, Jazz, Rain, EcoR1 and receives research funding from Novartis and
642 Revolution Medicines and Strategia.
643 N.I.V. served on an Advisory Board for Sanofi Genzyme. E.M.V.A. is a consultant for Tango
644 Therapeutics, Genome Medical, Invitae, Enara Bio, Janssen, Manifold Bio, Monte Rosa; receives
645 research funding from Novartis, BMS; has equity in Tango Therapeutics, Genome Medical,
646 Syapse, Enara Bio, Manifold Bio, Microsoft, Monte Rosa; has received travel reimbursement from
647 Roche/Genentech and own institutional patents filed on chromatin mutations and immunotherapy
648 response, and methods for clinical interpretation. C.E.M. is on advisory board of Genentech;

649 receives honoraria from Novartis, Guardant, Research and receives funding from Novartis,
650 Revolution Medicines. C.M.B. is a consultant for Amgen, Foundation Medicine, Blueprint
651 Medicines, Revolution Medicines; receives research funding from Novartis, AstraZeneca, Takeda;
652 and institutional research funding from Mirati, Spectrum, MedImmune and Roche.
653
654 Journal Editorial Board membership: C.S. is editorial board member for *Cell* and *PLOS*
655 *medicine*, associate editor and editorial board member for *Annals of Oncology*, scientific editor
656 for *Cancer Discovery* and advisory board member for *Nature Reviews Clinical*
657 *Oncology* and *Cancer Cell*. T.G.B is editor-in-chief of *npj Precision Oncology* and an editorial
658 board member of *Molecular and Cellular Biology*.
659
660 Patents: C.S. holds European patents relating to assay technology to detect tumour recurrence
661 (PCT/GB2017/053289); to targeting neoantigens (PCT/EP2016/059401), identifying patent
662 response to immune checkpoint blockade (PCT/EP2016/071471), determining HLA LOH
663 (PCT/GB2018/052004), predicting survival rates of patients with cancer
664 (PCT/GB2020/050221), identifying patients who respond to cancer treatment
665 (PCT/GB2018/051912), a US patent relating to detecting tumour mutations
666 (PCT/US2017/28013) and both a European and US patent related to identifying
667 insertion/deletion mutation targets (PCT/GB2018/051892).
668
669 C.S. is Royal Society Napier Research Professor (RP150154). His work is supported by the
670 Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001169),
671 the UK Medical Research Council (FC001169), and the Wellcome Trust (FC001169). C.S. is
672 funded by Cancer Research UK (TRACERx, PEACE and CRUK Cancer Immunotherapy
673 Catalyst Network), Cancer Research UK Lung Cancer Centre of Excellence, the Rosetrees
674 Trust, Butterfield and Stoneygate Trusts, NovoNordisk Foundation (ID16584), Royal Society
675 Research Professorship Enhancement Award (RP/EA/180007), the NIHR BRC at University
676 College London Hospitals, the CRUK-UCL Centre, Experimental Cancer Medicine Centre and
677 the Breast Cancer Research Foundation, USA (BCRF). His research is supported by a Stand Up
678 To Cancer-LUNGevity-American Lung Association Lung Cancer Interception Dream Team
679 Translational Research Grant (SU2C-AACR-DT23-17). Stand Up To Cancer is a program of

680 the Entertainment Industry Foundation. Research grants are administered by the American
681 Association for Cancer Research, the Scientific Partner of SU2C. C.S. also receives funding
682 from the European Research Council (ERC) under the European Union's Seventh Framework
683 Programme (FP7/2007-2013) Consolidator Grant (FP7-THESEUS-617844), and ERC
684 Advanced Grant (PROTEUS) from the European Research Council under the European Union's
685 Horizon 2020 research and innovation program (835297) and Chromavision from the European
686 Union's Horizon 2020 research and innovation programme (665233).
687
688 Special thanks to the Biological Research Facility at the Francis Crick Institute, specifically to Ade
689 Adekoya, James Cormack, Antony Horwood, and Scott Lighterness for their hard work and
690 support. Special thanks also to the Experimental Histopathology Laboratory at the Francis Crick
691 Institute, specifically to Emma Nye, Bruna Almeida, Mary Green, and Richard Stone for their help
692 and support. Special thanks to all the members of the Bivona laboratory (former and current),
693 Dmitry Gordenin, Alejandro Sweet-Cordero, Sourav Bandyopadhyay, Marcus Breese, Swati
694 Kaushik, Brandon Leonard and Sharath Raju for their insights and support and Sarah Elmes,
695 Ashley Maynard, David V. Allegakoen and Abhik Tambe for their technical support.
696
697 **Methods:**
698 **Cell line and growth assays**: Cell lines were grown in RPMI-1640 medium supplemented with
699 1% penicillin-streptomycin solution (10,000 Units/ml) and 10% FBS or in Iscove's Modified
700 Dulbecco's Medium (IMDM) supplemented with 1% penicillin-streptomycin solution (10,000
701 Units/ml), L-Glutamine (200 mM) and 10% FBS in a humidified incubator with 5% CO2
702 maintained at 37°C and fresh media was added to cells every 3-4 days. Drugs used for treatment
703 except PBS-1086[38] were purchased from Selleck Chemicals or from MedKoo Biosciences. For
704 growth assays, cells were exposed to the indicated drugs for few days-2.5 months (long-term
705 growth assays) or DMSO for 3-4 days (controls used for long-term growth assays) in 6 well plates
706 or 96 well plates (details in the figure legends) and assayed using crystal violet staining or Celltiter-
707 Glo luminescent viability assay (Promega) according to manufacturer's instructions.
708
709 **Deriving clonal populations and generating APOBEC3B knockout cells**: Clonal cells were
710 derived by sorting single cells into 96-well plates and expanding them over a few weeks. We then

711 derived pools of one of the clones expressing either a non-targeting guide or APOBEC3B targeting
712 guide along and puromycin marker and CRISPR/Cas9 by lentiviral transduction as done in a
713 previously published study[80]. A couple of gRNA target sequences, which were designed by the
714 Zhang lab to specifically target APOBEC3B[81] were first subcloned into the all-in-one lentiCRISPR
715 v2 plasmid (Addgene plasmid # 52961- a gift from Feng Zhang)[81]. They were then lentivirally
716 transduced into PC9 clonal cells, tested and the one that showed better APOBEC3B-depletion in
717 western blot analysis was selected for further analysis. Additionally, we also generated
718 APOBEC3B knockout clones from clonal derivatives of PC9 cells (C2 and E4). For this we used
719 the Dox-inducible version of LentiCRISPR v2 (Addgene plasmid # 87360 - a gift from Adam
720 Karpf) and subcloned the APOBEC3B-targeting sgRNA that worked well, into it[82]. The plasmid
721 was then lentivirally transduced into PC9 clonal cells, subjected to puromycin selection and then
722 half the cells were treated with 1 ug/ml doxycycline (dox) to induce the expression of Cas9 and
723 thereby generate APOBEC3B knockouts. Induction of APOBEC3B deletion was confirmed by
724 western blot analysis. Clonal cells were then derived from these cells using single cell sorting as
725 indicated earlier and the clonal populations that deleted APOBEC3B were identified by western
726 blot analysis and tested further. The cells not treated with dox served as APOBEC3B-proficient
727 control for the APOBEC3B-knockout clones.
728
729 **Transductions and transfections**: Hek293T cells were co-transfected with lentiviral packaging
730 plasmids pCMVdr8 and pMD2.G plasmid, along with the plasmid of interest using Fugene 6
731 transfection reagent (Promega). All shRNA plasmids used were purchased from Sigma.
732 APOBEC3B shRNA (TRCN0000142875). Cells were transduced by incubation with 1:1 diluted
733 lentivirus for 1- 2 days and then selected with the antibiotic marker (puromycin or hygromycin)
734 until untransduced control plate was clear. Negative control siRNA and other siRNA cocktails
735 used were purchased from GE Dharmacon. siRNAs were transfected using Lipofectamine RNAi
736 Max according to manufacturer's protocol and the cells were harvested 48 hr of transfection for
737 subsequent assays. or western blot analysis were used to confirm the shRNA, siRNA of CRISPR-
738 mediated depletions.
739
740 **RT-qPCR**: Total RNA was extracted using GeneJet RNA purification kit (Thermo Fisher) or
741 RNeasy Mini kit (Qiagen) and cDNA was synthesized from it using sensiFast cDNA synthesis kit

742 or Applied Biosystems High-Capacity cDNA Reverse Transcription Kit in accordance with the
743 manufacturer's instructions. qPCR reactions were performed using PowerUP SYBR Green Master
744 Mix (Applied Biosystems) or TaqMan Universal PCR Master Mix (Applied Biosystems) and
745 previously validated primers[83] (PrimerBank) on a QuantStudio. GAPDH, Actin, 18S RNA or
746 β2Microglobulin were used as reference genes. The following primers were used for p53 pathway
747 activation: Actin: Mm02619580_g1, Bax: Mm00432051_m1, Cdkn1a/p21: Mm04205640_g1,
748 Mdm2: Mm01233138_m1, Pmaip1/Noxa: Mm00451763_m1, Sesn2: Mm00460679_m1. Data
749 was analyzed using QuantStudio 12K Flex Software V1.3 and GraphPad Prism 7.
750
751 **Western Blot Assay**: Cells were treated with the indicated drugs the day after plating on a 6 well
752 plate or 10 cm plates for the indicated duration. Whole cell extracts were made with the former
753 were harvested by first washing with ice cold PBS and then lysing using ice cold RIPA buffer
754 containing protease and phosphatase inhibitors followed by sonication and centrifugation for
755 clarification of extracts. Cells grown on 10 cm plates were used for making nuclear-cytoplasmic
756 extracts as described previously using ice-cold 0.1% NP-40 in PBS[84]. Equal amounts of extracts
757 (quantified using Lowry assay) were subjected to SDS-PAGE analysis on 4-15% Criterion TX
758 gels (Bio-Rad), transferred to nitrocellulose membranes using semi-dry transfer apparatus and
759 Tans-Blot Turbo RTA Midi Nitrocellulose Transfer kit (Bio-Rad). The membranes were then
760 blocked with 3% milk in TBST, probed with primary antibody overnight at 4°C and then with the
761 corresponding secondary antibody- either HRP-conjugated or fluorescently labelled for 1-2 hr at
762 room temperature. They were then imaged on a Li-Cor imager or ImageQuant LAS 4000 (GE
763 Healthcare). anti-APOBEC3B (5210-87-13)[85] and anti-UNG[28] antibodies were kindly provided by
764 Reuben Harris and anti-GAPDH (sc-59540) was purchased from Santa Cruz Biotechnology. All
765 other primary antibodies including anti-EGFR (#4267), anti-phospho-EGFR (Y1068, #3777 or
766 #2236), anti-STAT3 (#9139), anti-phospho-STAT3 (Y705, #9145), anti-AKT (#2920), anti-
767 phospho-AKT (S473, #4060), anti-phospho-ERK (T202, Y204; #4370 or #9106), anti-ERK
768 (#9102), anti-RELA (#8242), anti-RELB (#4922), anti-Hsp90 (#4874), anti-TUBB (#2146) and
769 anti-histone H3 (#9715) were purchased from Cell Signaling Technology.
770
771 **Enzymatic assays**: APOBEC assays or Uracil Excision assays were performed by incubating
772 nuclear extracts made using REAP method [59] with either of the following DNA oligo substrates

773 (IDT): 5' – ATTATTATTATTCAAATGGATTTATTTATTTATTTATTTATTT-FAM-3' or 5'-
774 AGCAGTATTUGTTGTCACGA-FAM-3', respectively using established protocols[28,36]. Upon
775 completion of the reactions, they were heated at 95 degrees for 5 minutes after addition of TBE
776 Urea buffer (Novex) and immediately run on a 15% TBE-Urea gel (Bio-Rad). The gels were then
777 imaged using Cy2 filter on ImageQuant LAS 4000 (GE Healthcare) and quantified using ImageJ
778 and Microsoft Excel or GraphPad Prism 7.
779
780 **Subcutaneous Tumour Xenografts and PDX Studies:** All animal experiments were conducted
781 under UCSF IACUC-approved animal protocols. PC-9 tumour xenografts were generated by
782 injection of one million cells in a 50/50 mixture for matrigel and PBS into 6- to 8-wk-old female
783 NOD/SCID mice. Once the tumours grew to ~200 mm^3, the mice were then treated with vehicle
784 or 5 mg/kg osimertinib once daily for 4 days and the tumours were harvested on day 4 and
785 subsequently analyzed by western blot analysis. For H1975 xenografts, 1 million cells were
786 injected into SCID mice and the indicated treatment were initiated (once daily, oral gavage) when
787 the tumours were about 100 mm^3. The tumour tissues were harvested on day 4 of treatment. PDX
788 was generated as indicated in a previous study[38]. Tumours were passaged in SCID mice and
789 treatment was initiated after the tumours were about 400 mm^3. Tumours were harvested after 2
790 days of treatment with 25 mg/kg erlotinib, administered once daily by oral gavage.
791
792 **Mouse strains and tumour induction**: The Cre-inducible *Rosa26::LSL-APOBEC3Bi* mice are
793 described[7]. The *TetO-EGFR*^{L858R}*; Rosa26*$^{LNL\text{-}tTA}$ *(E)* and *CCSP-rtTA;TetO-*
794 *EGFR*^{L858R}*;Rosa26*$^{CreER(T2)}$ mice have been described[13,14,86,87]. Tumours were initiated in *E*, *EA3B,*
795 *EP, and EPA3B* mice by intratracheal infection with adenoviral vectors expressing Cre
796 recombinase as described[88]. Adenoviral-Cre (Ad-Cre-GFP) was from the University of Iowa Gene
797 Transfer Core. Tumours were initiated in *EA3Bi* mice using chow containing doxycycline (625
798 ppm) obtained from Harlan-Tekland. All animal regulated procedures were approved by the
799 Francis Crick Institute BRF Strategic Oversight Committee that incorporates the Animal Welfare
800 and Ethical Review Body and conformed with the UK Home Office guidelines and regulations
801 under the Animals (Scientific Procedures) Act 1986 including Amendment Regulations 2012. To
802 assess recombination efficiency of the LSL allele upstream of APOBEC3B, PCR primers
803 targeting the R26 site, the LSL cassette and the APOBEC3B transgene were used as described[7].

804

805 ***In vivo* treatment with Erlotinib**: Erlotinib was purchased from Selleckchem (Erlotinib Osi-744),
806 dissolved in 0.3% methylcellulose and administered intraperitoneally at 25 mg/kg, 5 days a week.
807

808 ***In vivo* treatment with tamoxifen**: Tamoxifen was administered by oral gavage 3 times in one
809 week with at a 2-4 day intervals (3 injections total). Mice received tamoxifen at 150 mg/kg
810 dissolved in sunflower oil.
811

812 **Assessment of recombination efficiency**: PCR was performed to assess recombination of the LSL
813 cassette upstream of the *A3B* allele in six tumours harvested at progression. Five out of six (5/6)
814 of the tumours had a recombination efficiency above 90% and one tumour out of six was
815 unrecombined (Extended Fig. 2c). This rate of recombination aligns with the rate of recombination
816 observed by IHC staining at 3 months and at termination (Extended Fig. 2a) and suggests that a
817 lack of recombination of the LSL cassette upstream of the A3B transgene explains A3B negative
818 tumours.
819

820 **MicroCT Imaging**: Mice were anaesthetized with isoflurane/oxygen for no more than an hour
821 each. They were minimally restrained whilst anaesthetized during the actual imaging (usually 8-
822 10 minutes). After scanning mice were observed and if necessary, placed in cages in a recovery
823 chamber/rack until they regained consciousness and start to eat/drink. Tumour burden in each
824 animal was quantified by calculating the tumour volume of visible tumours using AnalyzeDirect.
825

826 **Histological preparation and immunohistochemical staining**: Tissues were fixed in 10%
827 formalin overnight and transferred to 70% ethanol until paraffin embedding. IHC was performed
828 using the following primary antibodies: $EGFR^{L858R}$ mutant specific (Cell Signaling: 3197, 43B2),
829 APOBEC3B (5210-87-13)[85], Ki67 (Abcam: Ab15580), Caspase 3 (R&D (Bio-Techne): AF835),
830 p-Histone H2AX (Sigma-Aldrich, Ref. no. 05-636), Phospho-Histone H3 (Ser10) (Cell Signaling
831 Technology, 9706), CD4 (Abcam: ab183685, EPR19514), CD8 (Thermo Fisher Scientific, cod14-
832 0808-82, 4SM15), and UNG (NB600-1031, Novus Biologicals). Sections were developed with
833 DAB and counterstained with hematoxylin. The number of $EGFR^{L858R}$, APOBEC3B, Ki67,
834 Caspase 3, and gH2AX positive cells were quantified using QuPath.

835

836 **P53 Immunohistochemistry (IHC) staining:** Staining for p53 was performed on a Dako
837 Autostainer Link 48 (Agilent, UK) as previously described[89]. Briefly, formalin-fixed paraffin
838 embedded (FFPE) sections were dewaxed and subjected to heat induced epitope retrieval (HIER)
839 on-board using a Dako PT module. Once loaded, sections were heated to 97ºC for 20 minutes in
840 high pH Target Retrieval Solution (Agilent, UK). Sections then underwent peroxidase blocking
841 (Agilent, UK) before application of primary antibody (Leica, cat # NCL-L-p53-CM5p) for 40
842 minutes at a previously optimised dilution (1:750). Sections were then incubated with rabbit
843 envision secondary antibody (Agilent, UK) for 30 minutes. Signal was detected with liquid DAB
844 reagent (Agilent, UK). After detection, sections were counterstained with haematoxylin,
845 coverslipped using DPX mountant (CellPath, UK), and allowed to dry overnight. Slides were then
846 scanned at 20X magnification using an Aperio AT2 slide scanner (Leica).

847

848 **Evaluation of Chromosome Missegregation Errors in H&E and/or pH3 Stained Samples:**
849 Lung sections were evaluated for anaphases with chromosome missegregation events using a 100x
850 objective light microscope. For *E* and *EA3B* mice at early and late timepoints, the percentage of
851 missegregation errors was calculated and averaged across all mice using the harmonic mean. For
852 *EA3B* mice the percent error was normalised to an A3B recombination efficiency of 82% based
853 on observed recombination efficiency observed (Extended Fig. 4). For E and EA3Bi mice with
854 subclonal A3B expression, normalisation for the recombination efficiency was not possible so the
855 percentage of missegregation errors was calculated based on the number of errors verses normal
856 anaphases observed.

857

858 **Tumour processing Mouse**: Tumour tissue was collected form -80° freezer and placed on dry
859 ice. Tissue was placed on a petri dish and cut into pieces. RLT Buffer with β-Mercaptoethanol
860 was dispensed into 15 ml reaction vessels. A small part of the tumour was added to the lysis
861 buffer, and remaining parts were refrozen. TissueRuptor was used for disruption and
862 homogenization of tissue. The tip of the disposable probe was placed into the vessel and the
863 TissueRuptor was turned to speed 1 until the lysate was homogeneous (usually about 10
864 seconds). Lysate was added to a previously prepared QIAshredder tube and centrifuged at full

865 speed for 1 minute. The homogenised solution was then added to AllPrep DNA spin columns.
866 The Qiagen AllPrep DNA/RNA Mini Kit (80204).
867
868 **Histopathological Examination Mouse:** Four µm-thick, formalin-fixed, paraffin-embedded
869 (FFPE) sections from lung lobes were stained with haematoxylin & eosin (HE) and examined by
870 two board-certified Veterinary Pathologists (ASB & SLP). Sections were histopathologically
871 assessed and graded for the presence and type of proliferative epithelial lung lesions using the
872 INHAND guide for non-proliferative and proliferative lesions of the respiratory tract of the
873 mouse[90]. Histopathological assessment was performed blind to experimental grouping using a
874 conventional light microscope (Olympus BX43). Tissue sections were examined individually by
875 ASB & SLP and in case of discordance in diagnosis a consensus was reached using a double-head
876 microscope.
877
878 Proliferative lesions were diagnosed when the following histopathological characteristics were
879 observed involving the alveolar lining epithelium:
880
881 - Alveolar hyperplasia: Isolated plump pneumocytes which frequently involved locally extensive-
882 flat areas of lung parenchyma, mostly located subpleurally. Cytologically hyperplastic
883 pneumocytes lacked cellular atypia and had a single regular ovoid nucleus with homogeneous
884 chromatin. Cellular atypia was absent and mitotic figures were not observed.
885
886 - Bronchiolo-alveolar adenoma: A well-circumscribed expansile focus of neoplastic cells, ranging
887 from 100 µm to 500 µm in diameter. Neoplastic cells arranged in either a lepidic or arboriform
888 pattern. Cellular atypia was minimal and mitoses were absent.
889
890 - Well differentiated bronchiolo-alveolar adenocarcinoma: Similar histological features and size
891 as in adenomas but with increased cellular atypia and occasional mitotic figures.
892
893 - Moderately and poorly-differentiated bronchiolo-alveolar adenocarcinoma: A moderately well-
894 demarcated neoplastic focus, ranging from 100 µm to 3 mm or more in diameter. Poorly-
895 differentiated neoplasms normally exhibited densely packed sheets of neoplastic cells without

896 evidence of arboriform or trabecular growth. Frequent cellular and nuclear atypia and increased
897 mitotic activity.
898
899 Neoplastic foci where manually quantified in all lung lobe sections. The total area of neoplastic
900 infiltration was semi-quantitatively assessed and expressed as a percentage of the lung parenchyma
901 affected, avoiding areas of alveolar hyperplasia.

902 **Whole exome sequencing – mouse data**: WES was performed by the Advanced Sequencing
903 Facility at The Francis Crick Institute using the Twist BioScience Human Core Exon Kit for library
904 preparation and Agilent SureSelectXT Mouse All Exon, 16, Kit for library preparation
905 respectively. Sequencing was performed on HiSeq 4000 platforms.

906 **RNA sequencing – mouse data:** RNA-seq was performed by the Advanced Sequencing Facility
907 at the Francis Crick Institute using the KAPA mRNA HyperPrep Kit (KK8581 – 96 Libraries)
908 and KAPA Dual-Indexed Adapters (Roche-KK8720). Sequencing was performed on HiSeq 4000
909 platforms. The processed fastq files were mapped to mm10 reference genome using STAR
910 (version 2.4) algorithm and transcript expressions were quantified using RSEM (version 1.2.29)
911 algorithm with the default parameters. The read counts were used for downstream analysis.

912 **Alignment – mouse:** All samples were de-multiplexed and the resultant FASTQ files aligned to
913 the mm10 mouse genome, using bwa mem (bwa v0.7.15). De-duplication was performed using
914 Picard (v2.1.1) (http://broadinstitute.github.io/picard). Quality control metrics were collated using
915 FASTQC (v0.10.1 - http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), Picard and
916 GATK (v3.6). SAMtools (v1.3.1) was used to generate mpileup files from the resultant BAM files.
917 Thresholds for base phred score and mapping quality were set at 20. A threshold of 50 was set for
918 the coefficient of downgrading mapping quality, with the argument for base alignment quality
919 calculation being deactivated. The median depth of coverage for all samples was 92x (range: 58-
920 169x).

921 **Variant detection & annotation – mouse:** Variant calling was performed using
922 VarScan2(v2.4.1), MuTect(v1.1.7) and Scalpel(v0.5.4)[91-93].

923 The following argument settings were used for variant detection using VarScan2:

924 --min-coverage 8 --min-coverage-normal 10 --min-coverage-tumor 6 --min-var-freq 0.01 --min-
925 freq-for-hom 0.75 --normal-purity 1 --p-value 0.99 --somatic-p-value 0.05 --tumor-purity 0.5 --
926 strand-filter 0

927 For MuTect, only "PASS" variants were used for further analyses. With the exception of allowing
928 variants to be detected down to a VAF of 0.001, default settings were used for Scalpel
929 insertion/deletion detection.

930 To minimise false positives, additional filtering was performed. For single-nucleotide variants
931 (SNVs) or dinucleotides detected by VarScan2, a minimum tumor sequencing depth of 30, variant
932 allele frequency (VAF) of 5%, variant read count of 5, and a somatic p-value <0.01 were required
933 to pass a variant. For variants detected by VarScan2 between 2 and 5% VAF, the mutation also
934 needs to be detected by MuTect.

935 As for insertions/deletions (INDELs), variants need to be passed by both Scalpel ("PASS") and
936 VarScan2 (somatic p value <0.001). A minimum depth of 50x, 10 alt reads and VAF of 2% was
937 required.

938 For all SNVs, INDELs and dinucleotides, any variant also detected in the paired germline sample
939 with more than 5 alternative reads or a VAF greater than 1% was filtered out.

940 The detected variants were annotated using Annovar[94].

941 **Functional annotation of SNVs – mouse:** Murine gene mutation callings from whole exome
942 sequencing were parsed with some modification including genomic coordinates (removing 'chr'
943 before chromosomal numbers, only 'SNV' was selected). The modified files were fed into
944 PROVEAN (Protein Variation Effect Analyzer)[95-97] software tool
945 (http://provean.jcvi.org/index.php) to predict whether an amino acid substitution has an impact
946 on the biological function of a protein (SIFT score). The predict files were merged with original
947 files at gene level annotation using R program.
948

949 **Human *EGFR* transgene amplicon sequencing of mouse:** FASTQ files were aligned to hg19
950 obtained from the GATK bundle (v2.8) using bwa mem (bwa v0.7.15)[98,99]. Analyses were

951 performed using R (version 3.3.1) and deepSNV (v1.18.1)[100]. The median depth of coverage of
952 sequenced EGFR exons (19,20,21) was 5290x (range: 2238-8040). Variants associated with
953 resistance to EGFR tyrosine kinase inhibitors were queried using deepSNV's bam2R function,
954 with the arguments q=20 & s=2. The variants explored include: T790M, D761Y, L861Q, G796X,
955 G797X, L792X, L747S. L858R was identified in every sequenced sample.

956

957 **Generation of EGFR^{L858R} mutant mouse tumour cell lines**: To generate mouse tumour cell
958 lines, a tumour was micro-dissected from mouse lungs. A portion of the tumour was then cut from
959 dissected (1/3 to 1/2 of the original tumour depending on size) and placed into a 15 ml tube and
960 cut into small pieces with scissors. Tumours chunks were then digested for 30 minutes in an oven
961 at 37 degrees while rotating at full speed in digestion media (1400 ul HBSS-free (w/o Ca2+), 200
962 ul Collagenase IV, and 40 U/ml DNase). Next tumour cells were pelleted down in a centrifuge
963 (1100 rpm for 4 minutes) and resuspended in Iscove's Modified Dulbecco's Medium (IMDM)
964 supplemented with 1% penicillin-streptomycin solution (10,000 Units/ml), L-Glutamine (200
965 mM) and 10% FBS. This cell suspension was then plated in a 10 cm plate and passaged over a
966 period of 1 to 3 months until consistent growth was observed.
967

968 **Generation of tyrosine kinase inhibitor (TKI) resistant mouse tumour cell lines**: TKI naïve
969 cell lines were cultured in increasing levels of the TKIs Erlotinib or Osimertinib until growing
970 with minimal cell death in 1500 nM TKI therapy, which was considered fully resistant.
971

972 **Cell line Whole Genome Mutational Signature Analysis:** Sequences were aligned to the human
973 genome (hg38) using the Burrows-Wheeler Aligner (version 0.7.17). PCR duplicates were
974 removed using Picard (version 2.18.16). Reads were locally realigned around InDels using
975 GATK3 (version 3.6.0) tools RealignerTargetCreator to create intervals, followed by
976 IndelRealigner on the aligned bam files. MuTect2 from GATK3 (version 3.6.0) was used in
977 tumour/normal mode to call mutations in test vs control cell lines. Single nucleotide variants
978 (SNVs) that passed the internal GATK3 filter with read depths over 30 reads at called positions,
979 at least 4 reads in the alternate mutation call and an allele frequency greater than 0.05 were used
980 for downstream analysis. Figures were plotted using the deconstructSigs R package[101].
981

982 **DNA and RNA isolation from cell line models for sequencing**: DNA or RNA were extracted
983 using from frozen cell pellets using Qiagen's DNeasy Blood and tissue kit or Qiagen's RNeasy
984 mini kit, respectively, as per the manufacturer's instructions. The isolated DNA or RNA was
985 quantified and qualitatively assessed using Qubit Fluorometer (Thermo Fisher) and a Bioanalyzer
986 (Agilent) as per the manufacturer's instructions. The DNA or RNA were then sent to BGI for
987 whole genome sequencing (30X) or Novogene for mRNA or whole exome sequencing.
988
989 **Cell cycle analysis**: Cell cycle analysis was performed by propidium iodide (PI) staining. Briefly,
990 PC9 cells were treated for 24h with DMSO, 2 μM osimertinib or 1 μM palbociblib, and then fixed
991 in ice cold 70% ethanol and stained with a 50 ug/mL PI (Millipore Sigma, P4864) + 0.1% Triton
992 X-100 (Millipore Sigma, X100) solution. PI fluorescence was then measured on a flow cytometer
993 (BD FACSAria II).
994 **Human Subjects**: All patients gave informed consent for collection of clinical correlates, tissue
995 collection, research testing under Institutional Review Board (IRB)-approved protocols (CC13-
996 6512 and CC17-658, NCT03433469). Patient demographics are listed in Supplemental Tables.
997 Patient studies were conducted according to the Declaration of Helsinki, the Belmont Report, and
998 the U.S. Common Rule."
999
1000 **Studies with specimens from lung cancer patients**: Frozen or formalin-fixed paraffin embedded
1001 (FFPE) tissues from lung cancer patients for DNA or RNA Sequencing (Bulk and Single Cell)
1002 studies were, processed and sequenced as described previously[47,66]. We obtained tumor tissues
1003 from advanced stage NSCLC patients prior to initiation of systemic targeted therapy (classified as
1004 treatment naïve, TN), while on targeted therapy while the tumor was regressing or at stable state
1005 as evaluated by standard clinical imaging (classified as residual disease, RD) and upon clear
1006 progressive disease (acquired resistance) as determined by standard clinical imaging (classified as
1007 progressive disease, PD). The classification of response was based on RECIST (Response
1008 Evaluation Criteria in Solid Tumors) criteria. Some of these biopsies were subjected to whole
1009 exome sequencing at the QB3-Berkley Genomics for which library preparation was performed
1010 using IDT's xGen exome panel. For additional specimens, tumour DNA from formalin-fixed
1011 paraffin embedded (FFPE) tissues and matched non-tumour from blood aliquots or stored buffy
1012 coats were collected as part of University of California, San Francisco's (UCSF) biospecimen

1013 resource program (BIOS) in accordance with UCSF's institutional review board approved
1014 protocol. DNA from blood aliquots were isolated at the BIOS. Other non-tumour samples and
1015 FFPE tumour tissues were sent for extraction, assessment of quality and quantity to Novogene and
1016 those meeting the required samples standards were subjected to whole exome sequencing at
1017 Novogene's sequencing facility.
1018
1019 **Mutation analysis**:
1020 Paired-end reads were aligned to the hg19 human genome using the Picard pipeline
1021 (https://gatk.broadinstitute.org/). A modified version of the Broad Institute Getz Lab CGA WES
1022 Characterization pipeline (https://docs-google-com.ezp-
1023 prod1.hul.harvard.edu/document/d/1VO2kX_fgfUd0x3mBS9NjLUWGZu794WbTepBel3cBg08
1024) was used to call, filter and annotate somatic mutations. Specifically, single-nucleotide variants
1025 (SNVs) and other substitutions were called with MuTect (v1.1.6)[92]. Mutations were annotated
1026 using Oncotator[102]. MuTect mutation calls were filtered for 8-OxoG artifacts, and artifacts
1027 introduced through the formalin fixation process (FFPE) of tumour tissues [67]. Indels were called
1028 with Strelka (v1.0.11). MuTect calls and Strelka calls were further filtered through a panel of
1029 normal samples (PoN) to remove artifacts generated by rare error modes and miscalled germline
1030 alterations[92]. To pass quality control, samples were required to have <5% cross-sample
1031 contamination as assessed with ContEst[92]; mean target coverage of at least 25x in the tumour
1032 sample and 20x in the corresponding normal as assessed using GATK3.7 DepthOfCoverage; and
1033 a percentage of tumour-in-normal of < 30% as determined by deTiN[103]. This pipeline was modified
1034 for analysis of cell lines rather than tumour-normal pairs as follows: indels were called through
1035 MuTect2 alone rather than Strelka; deTiN was not performed; and a common variant filter was
1036 applied to exclude variants present in The Exome Aggregation Consortium (ExAC) if at least 10
1037 alleles containing the variant were present across any subpopulation, unless they appeared in a list
1038 of known somatic sites[104,105].
1039
1040 **Mutational signature analysis**
1041 Active mutational processes[106] were determined using the *deconstructSigs* R package[64], with a
1042 signature contribution cutoff of 6%. This cutoff was chosen because it was the minimum
1043 contribution value required to obtain a false-positive rate of 0.1% and false-negative rate of 1.4%

1044 per the authors' in-silico analysis, and is the recommended cutoff[101]. Samples with < 10 mutations
1045 were excluded from analysis due to poor signature discrimination with so few mutations and a
1046 sample with less than 15 days of exposure to TKI therapy was excluded because it is too short a
1047 time to accumulate detectable mutations due to therapy. For TRACERx data analysis data
1048 processing were performed in the R statistical environment version > = 3.3.1.
1049
1050 **RNA-Seq analyses**: PDX-tissue and mouse tumour cell line RNA extractions were carried using
1051 RNeasy micro kit (Qiagen). RNA-Seq was performed on PDX-tissue using replicate samples on
1052 the Illumina HiSeq4000, paired-end 100-bp reads at the Center for Advanced Technology (UCSF).
1053 For the differential gene expression analysis, DESeq program was used to compare controls to
1054 erlotinib samples as previously described[107].
1055
1056 RNAseq samples from patients and cell lines were sequenced by Novogene
1057 (https://en.novogene.com/) with paired-end sequencing (150bp in length). There were ~20 million
1058 reads for each sample. The processed fastq files were mapped to hg19 reference genome using
1059 STAR (version 2.4) algorithm and transcript expressions were quantified using RSEM (version
1060 1.2.29) algorithm. The default parameters in the algorithms were used. The normalized transcript
1061 reads (TPM) were used for downstream analysis. Gene set enrichment analysis was performed
1062 using GSEA software[108].
1063
1064 For single cell RNA Seq analyses the data from a previously published study (all cancer cells from
1065 advanced lung cancer patients) was used and analyzed in a similar manner[47]. All cells used are
1066 identified as malignant by marker expression and CNV inference and originated in from various
1067 biopsy sites (adrenal, liver, lymph node, lung, pleura/pleural fluid). Nonparametric, pairwise
1068 comparisons (Wilcoxon Rank Sum Test) was used to determine the statistical significance of the
1069 pairwise comparisons of different timepoints for their average scaled expression.
1070
1071 **Statistical Analysis**
1072 One-way or two-way ANOVA test with Holm's Sidak correction for multiple comparisons (>2
1073 groups) or two-tailed t-test (2 groups) were used to determine the statistical significance of the
1074 differences between groups for RT-qPCR, growth and enzymatic assays and Bulk RNA Seq

analysis. Normality of immunohistochemical and microCT data was determined using multiple testing methods (Anderson-Darling test, D'Agostino & Pearson test, Shapiro-Wilk test, and Kolmogorov-Smirnov test). A two-sided t-test or two-sided Mann-Whitney test was used for immunohistochemical and microCT data depending on the normality tests to determine the statistical significance of the differences between groups. Analysis for these assays was done using Graphpad Prism 7.

Data availability

The Whole Exome Sequencing (WES) data and RNAseq data (from the TRACERx study) generated, used or analysed during this study are available through the Cancer Research UK and University College London Cancer Trials Centre (ctc.tracerx@ucl.ac.uk) for academic non-commercial research purposes only, and subject to review of a project proposal that will be evaluated by a TRACERx data access committee, entering into an appropriate data access agreement and subject to any applicable ethical approvals. For the single cell RNA-Seq analyses shown in Ext. Fig. 15a-d the data from a previously published study (all advanced lung cancer cell data) were used and analyzed in a similar manner[47]. This data is available as an NCBI Bioproject # PRJNA591860. The RNA-Seq data for Ext. Fig. 14a was from a previously published study[38]. These data are available at NCBI GEO under accession number GSE65420. Related RNA-Seq and WES sequencing data will be deposited into NCBI GEO and SRA database.

References:

1 Cheng, A. Z. *et al.* APOBECs and Herpesviruses. *Viruses* **13**, doi:10.3390/v13030390 (2021).

2 Harris, R. S. & Liddament, M. T. Retroviral restriction by APOBEC proteins. *Nat Rev Immunol* **4**, 868-877, doi:10.1038/nri1489 (2004).

3 Swanton, C., McGranahan, N., Starrett, G. J. & Harris, R. S. APOBEC Enzymes: Mutagenic Fuel for Cancer Evolution and Heterogeneity. *Cancer discovery* **5**, 704-712, doi:10.1158/2159-8290.CD-15-0344 (2015).

1105 4 Venkatesan, S. *et al.* Induction of APOBEC3 Exacerbates DNA Replication Stress and
1106 Chromosomal Instability in Early Breast and Lung Cancer Evolution. *Cancer Discov* **11**,
1107 2456-2473, doi:10.1158/2159-8290.CD-20-0725 (2021).
1108 5 Isozaki, H. *et al.* APOBEC3A drives acquired resistance to targeted therapies in non-
1109 small cell lung cancer. *biorxiv.org*, doi:10.1101/2021.01.20.426852 (2021).
1110 6 Jarvis, M. C. *et al.* Mutational impact of APOBEC3B and APOBEC3A in a human cell
1111 line. *biorxiv.org*, doi:10.1101/2022.04.26.489523 (2022).
1112 7 Boumelha, J. *et al.* An immunogenic model of KRAS-mutant lung cancer for study of
1113 targeted therapy and immunotherapy combinations. *biorxiv.org*,
1114 doi:10.1101/2020.12.22.423126 (2020).
1115 8 Russo, M. *et al.* Adaptive mutability of colorectal cancers in response to targeted
1116 therapies. *Science* **366**, 1473-1480, doi:10.1126/science.aav4474 (2019).
1117 9 de Bruin, E. C. *et al.* Spatial and temporal diversity in genomic instability processes
1118 defines lung cancer evolution. *Science* **346**, 251-256, doi:10.1126/science.1253462
1119 (2014).
1120 10 Chen, J. *et al.* Genomic landscape of lung adenocarcinoma in East Asians. *Nature*
1121 *Genetics* **52**, 177-186, doi:10.1038/s41588-019-0569-6 (2020).
1122 11 Jamal-Hanjani, M. *et al.* Tracking the Evolution of Non-Small-Cell Lung Cancer. *The*
1123 *New England journal of medicine*, doi:10.1056/NEJMoa1616288 (2017).
1124 12 Henderson, S., Chakravarthy, A., Su, X., Boshoff, C. & Fenton, T. R. APOBEC-mediated
1125 cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-
1126 driven tumor development. *CellReports* **7**, 1833-1841, doi:10.1016/j.celrep.2014.05.012
1127 (2014).
1128 13 Politi, K. *et al.* Lung adenocarcinomas induced in mice by mutant EGF receptors foundin
1129 human lung cancers respondto a tyrosine kinase inhibitor orto down-regulation of the
1130 receptors. *Genes & Development* **20**, 1496-1510, doi:10.1101/gad.1417406 (2006).
1131 14 Politi, K., Fan, P.-D., Shen, R., Zakowski, M. & Varmus, H. Erlotinib resistance in
1132 mouse models of epidermal growth factor receptor-induced lung adenocarcinoma.
1133 *Disease Models & Mechanisms* **3**, 111-119, doi:10.1242/dmm.003681 (2010).

1134 15 Starrett, J. H. *et al.* Drug Sensitivity and Allele Specificity of First-Line Osimertinib
1135 Resistance EGFR Mutations. *Cancer Research* **80**, 2017-2030, doi:10.1158/0008-
1136 5472.CAN-19-3819 (2020).
1137 16 de Bruin, E. C., McGranahan, N. & Swanton, C. Analysis of intratumor heterogeneity
1138 unravels lung cancer evolution. *Molecular & Cellular Oncology* **2**, e985549,
1139 doi:10.4161/23723556.2014.985549 (2015).
1140 17 Westcott, P. M. K. *et al.* The mutational landscapes of genetic and chemical models of
1141 Kras-driven lung cancer. *Nature Publishing Group* **517**, 489-492,
1142 doi:10.1038/nature13898 (2015).
1143 18 McFadden, D. G. *et al.* Mutational landscape of EGFR-, MYC-, and Kras-driven
1144 genetically engineered mouse models of lung adenocarcinoma. **113**, E6409-E6417,
1145 doi:10.1073/pnas.1613601113 (2016).
1146 19 Chung, W.-J. *et al.* Kras mutant genetically engineered mouse models of human cancers
1147 are genomically heterogeneous. *Proceedings of the National Academy of Sciences of the*
1148 *United States of America*, 201708391, doi:10.1073/pnas.1708391114 (2017).
1149 20 Jónsson, S. R. *et al.* Evolutionarily conserved and non-conserved retrovirus restriction
1150 activities of artiodactyl APOBEC3F proteins. *Nucleic acids research* **34**, 5683-5694,
1151 doi:10.1093/nar/gkl721 (2006).
1152 21 MacMillan, A. L., Kohli, R. M. & Ross, S. R. APOBEC3 inhibition of mouse mammary
1153 tumor virus infection: the role of cytidine deamination versus inhibition of reverse
1154 transcription. *Journal of Virology* **87**, 4808-4817, doi:10.1128/JVI.00112-13 (2013).
1155 22 Ji, H. *et al.* The impact of human EGFR kinase domain mutations on lung tumorigenesis
1156 and in vivo sensitivity to EGFR-targeted therapies. *Cancer cell* **9**, 485-495,
1157 doi:10.1016/j.ccr.2006.04.022 (2006).
1158 23 DiMarco, A. V. *et al.* APOBEC mutagenesis inhibits breast cancer growth through
1159 induction of a T cell mediated antitumor immune response. *biorxiv.org*,
1160 doi:10.1101/2021.02.13.431068 (2021).
1161 24 Facciabene, A., Motz, G. T. & Coukos, G. T-regulatory cells: key players in tumor
1162 immune escape and angiogenesis. *Cancer Res* **72**, 2162-2171, doi:10.1158/0008-
1163 5472.CAN-11-3687 (2012).

1164 25 Carter, S. L., Eklund, A. C., Kohane, I. S., Harris, L. N. & Szallasi, Z. A signature of
1165 chromosomal instability inferred from gene expression profiles predicts clinical outcome
1166 in multiple human cancers. *Nat Genet* **38**, 1043-1048, doi:10.1038/ng1861 (2006).
1167 26 Sansregret, L. *et al.* APC/C Dysfunction Limits Excessive Cancer Chromosomal
1168 Instability. *Cancer Discov* **7**, 218-233, doi:10.1158/2159-8290.CD-16-0645 (2017).
1169 27 Seibler, J. Rapid generation of inducible mouse mutants. *Nucleic acids research* **31**, 12e-
1170 12, doi:10.1093/nar/gng012 (2003).
1171 28 Serebrenik, A. A. *et al.* The deaminase APOBEC3B triggers the death of cells lacking
1172 uracil DNA glycosylase. *Proceedings of the National Academy of Sciences of the United*
1173 *States of America* **116**, 22158-22163, doi:10.1073/pnas.1904024116 (2019).
1174 29 Haderk, F. *et al.* A focal adhesion kinase-YAP signaling axis drives drug tolerant
1175 persister cells and residual disease in lung cancer. *bioRxiv*, 2021.2010.2023.465573,
1176 doi:10.1101/2021.10.23.465573 (2021).
1177 30 Finn, K. J., Martin, S. E. & Settleman, J. A Single-Step, High-Dose Selection Scheme
1178 Reveals Distinct Mechanisms of Acquired Resistance to Oncogenic Kinase Inhibition in
1179 Cancer Cells. *Cancer Res* **80**, 79-90, doi:10.1158/0008-5472.CAN-19-0729 (2020).
1180 31 Ramirez, M. *et al.* Diverse drug-resistance mechanisms can emerge from drug-tolerant
1181 cancer persister cells. *Nature Communications* **7**, ncomms10690,
1182 doi:doi:10.1038/ncomms10690 (2016).
1183 32 Hashida, S. *et al.* Acquisition of cancer stem cell-like properties in non-small cell lung
1184 cancer with acquired resistance to afatinib. *Cancer Sci* **106**, 1377-1384,
1185 doi:10.1111/cas.12749 (2015).
1186 33 Shah, K. N. *et al.* Aurora kinase A drives the evolution of resistance to third-generation
1187 EGFR inhibitors in lung cancer. *Nat Med* **25**, 111-118, doi:10.1038/s41591-018-0264-7
1188 (2019).
1189 34 Shien, K. *et al.* Acquired resistance to EGFR inhibitors is associated with a manifestation
1190 of stem cell-like properties in cancer cells. *Cancer Res* **73**, 3051-3061, doi:10.1158/0008-
1191 5472.CAN-12-4136 (2013).
1192 35 Poulikakos, P. I. *et al.* RAF inhibitor resistance is mediated by dimerization of aberrantly
1193 spliced BRAF(V600E). *Nature* **480**, 387-390, doi:10.1038/nature10662 (2011).

1194 36 Vieira, V. C. *et al.* Human papillomavirus E6 triggers upregulation of the antiviral and
1195 cancer genomic DNA deaminase APOBEC3B. *mBio* **5**, doi:10.1128/mBio.02234-14
1196 (2014).
1197 37 Bollard, J. *et al.* Palbociclib (PD-0332991), a selective CDK4/6 inhibitor, restricts tumour
1198 growth in preclinical models of hepatocellular carcinoma. *Gut* **66**, 1286-1296,
1199 doi:10.1136/gutjnl-2016-312268 (2017).
1200 38 Blakely, C. M. *et al.* NF-κB-activating complex engaged in response to EGFR oncogene
1201 inhibition drives tumor cell survival and residual disease in lung cancer. *CellReports* **11**,
1202 98-110, doi:10.1016/j.celrep.2015.03.012 (2015).
1203 39 Law, E. K. *et al.* The DNA cytosine deaminase APOBEC3B promotes tamoxifen
1204 resistance in ER-positive breast cancer. *Science Advances* **2**, e1601737-e1601737,
1205 doi:10.1126/sciadv.1601737 (2016).
1206 40 Leonard, B. *et al.* The PKC/NF-κB signaling pathway induces APOBEC3B expression in
1207 multiple human cancers. *Cancer Research* **75**, 4538-4547, doi:10.1158/0008-5472.CAN-
1208 15-2171-T (2015).
1209 41 Maruyama, W. *et al.* Classical NF-κB pathway is responsible for APOBEC3B expression
1210 in cancer cells. *Biochemical and biophysical research communications* **478**, 1466-1471,
1211 doi:10.1016/j.bbrc.2016.08.148 (2016).
1212 42 Nelson, D. E. *et al.* Oscillations in NF-kappaB signaling control the dynamics of gene
1213 expression. *Science* **306**, 704-708, doi:10.1126/science.1099962 (2004).
1214 43 Tay, S. *et al.* Single-cell NF-kappaB dynamics reveal digital activation and analogue
1215 information processing. *Nature* **466**, 267-271, doi:10.1038/nature09145 (2010).
1216 44 Cheng, Q. J. *et al.* NF-kappaB dynamics determine the stimulus specificity of epigenomic
1217 reprogramming in macrophages. *Science* **372**, 1349-1353, doi:10.1126/science.abc0269
1218 (2021).
1219 45 Zambrano, S., Bianchi, M. E. & Agresti, A. High-throughput analysis of NF-kappaB
1220 dynamics in single cells reveals basal nuclear localization of NF-kappaB and spontaneous
1221 activation of oscillations. *PLoS One* **9**, e90104, doi:10.1371/journal.pone.0090104
1222 (2014).
1223 46 Inoue, K. *et al.* Oscillation dynamics underlie functional switching of NF-kappaB for B-
1224 cell activation. *NPJ Syst Biol Appl* **2**, 16024, doi:10.1038/npjsba.2016.24 (2016).

1225 47 Maynard, A. *et al.* Therapy-Induced Evolution of Human Lung Cancer Revealed by
1226 Single-Cell RNA Sequencing. *Cell* **182**, 1232-1251.e1222,
1227 doi:10.1016/j.cell.2020.07.017 (2020).
1228 48 Gilmore, T. D. Introduction to NF-kappaB: players, pathways, perspectives. *Oncogene*
1229 **25**, 6680-6684, doi:10.1038/sj.onc.1209954 (2006).
1230 49 Farré, D. *et al.* Identification of patterns in biological sequences at the ALGGEN server:
1231 PROMO and MALGEN. *Nucleic acids research* **31**, 3651-3653, doi:10.1093/nar/gkg605
1232 (2003).
1233 50 Hollenhorst, P. C. RAS/ERK pathway transcriptional regulation through ETS/AP-1
1234 binding sites. *Small GTPases* **3**, 154-158, doi:10.4161/sgtp.19630 (2012).
1235 51 Jia, P. *et al.* Next-generation sequencing of paired tyrosine kinase inhibitor-sensitive and
1236 -resistant EGFR mutant lung cancer cell lines identifies spectrum of DNA changes
1237 associated with drug resistance. *Genome Research* **23**, 1434-1445,
1238 doi:10.1101/gr.152322.112 (2013).
1239 52 Petljak, M. *et al.* Characterizing Mutational Signatures in Human Cancer Cell Lines
1240 Reveals Episodic APOBEC Mutagenesis. *Cell* **176**, 1282-1294 e1220,
1241 doi:10.1016/j.cell.2019.02.012 (2019).
1242 53 Sun, H. T., Cheng, S. X., Tu, Y., Li, X. H. & Zhang, S. FoxQ1 promotes glioma cells
1243 proliferation and migration by regulating NRXN3 expression. *PLoS One* **8**, e55693,
1244 doi:10.1371/journal.pone.0055693 (2013).
1245 54 Kusinska, R. *et al.* Influence of genomic variation in FTO at 16q12.2, MC4R at 18q22
1246 and NRXN3 at 14q31 genes on breast cancer risk. *Mol Biol Rep* **39**, 2915-2919,
1247 doi:10.1007/s11033-011-1053-2 (2012).
1248 55 Veeriah, S. *et al.* The tyrosine phosphatase PTPRD is a tumor suppressor that is
1249 frequently inactivated and mutated in glioblastoma and other human cancers. *Proc Natl*
1250 *Acad Sci U S A* **106**, 9435-9440, doi:10.1073/pnas.0900571106 (2009).
1251 56 Ortiz, B. *et al.* Loss of the tyrosine phosphatase PTPRD leads to aberrant STAT3
1252 activation and promotes gliomagenesis. *Proc Natl Acad Sci U S A* **111**, 8149-8154,
1253 doi:10.1073/pnas.1401952111 (2014).
1254 57 Morris, L. G. *et al.* Genomic dissection of the epidermal growth factor receptor
1255 (EGFR)/PI3K pathway reveals frequent deletion of the EGFR phosphatase PTPRS in

1256 head and neck cancers. *Proc Natl Acad Sci U S A* **108**, 19024-19029,
1257 doi:10.1073/pnas.1111963108 (2011).
1258 58 Nikolaienko, R. M., Agyekum, B. & Bouyain, S. Receptor protein tyrosine phosphatases
1259 and cancer: new insights from structural biology. *Cell Adh Migr* **6**, 356-364,
1260 doi:10.4161/cam.21242 (2012).
1261 59 Davis, T. B. *et al.* PTPRS Regulates Colorectal Cancer RAS Pathway Activity by
1262 Inactivating Erk and Preventing Its Nuclear Translocation. *Sci Rep* **8**, 9296,
1263 doi:10.1038/s41598-018-27584-x (2018).
1264 60 Roper, N. *et al.* Clonal Evolution and Heterogeneity of Osimertinib Acquired Resistance
1265 Mechanisms in EGFR Mutant Lung Cancer. *Cell Rep Med* **1**,
1266 doi:10.1016/j.xcrm.2020.100007 (2020).
1267 61 Samuels, Y. *et al.* Mutant PIK3CA promotes cell growth and invasion of human cancer
1268 cells. *Cancer cell* **7**, 561-573, doi:10.1016/j.ccr.2005.05.014 (2005).
1269 62 Kim, S. & Jeong, S. Mutation Hotspots in the beta-Catenin Gene: Lessons from the
1270 Human Cancer Genome Databases. *Mol Cells* **42**, 8-16, doi:10.14348/molcells.2018.0436
1271 (2019).
1272 63 Ruvolo, P. P. The broken "Off" switch in cancer signaling: PP2A as a regulator of
1273 tumorigenesis, drug resistance, and immune surveillance. *BBA Clin* **6**, 87-99,
1274 doi:10.1016/j.bbacli.2016.08.002 (2016).
1275 64 Silverstein, A. M., Barrow, C. A., Davis, A. J. & Mumby, M. C. Actions of PP2A on the
1276 MAP kinase pathway and apoptosis are mediated by distinct regulatory subunits. *Proc*
1277 *Natl Acad Sci U S A* **99**, 4221-4226, doi:10.1073/pnas.072071699 (2002).
1278 65 Gainor, J. F. *et al.* Molecular Mechanisms of Resistance to First- and Second-Generation
1279 ALK Inhibitors in ALK-Rearranged Lung Cancer. *Cancer discovery* **6**, 1118-1133,
1280 doi:10.1158/2159-8290.CD-16-0596 (2016).
1281 66 Blakely, C. M. *et al.* Evolution and clinical impact of co-occurring genetic alterations in
1282 advanced-stage EGFR-mutant lung cancers. *Nature Genetics* **49**, 1693-1704,
1283 doi:10.1038/ng.3990 (2017).
1284 67 Shen, S. *et al.* Upregulation of miR-136 in human non-small cell lung cancer cells
1285 promotes Erk1/2 activation by targeting PPP2R2A. *Tumour biology : the journal of the*

1286 *International Society for Oncodevelopmental Biology and Medicine* **35**, 631-640,
1287 doi:10.1007/s13277-013-1087-2 (2014).
1288 68 Engelman, J. A. *et al.* Effective use of PI3K and MEK inhibitors to treat mutant Kras
1289 G12D and PIK3CA H1047R murine lung cancers. *Nature Medicine* **14**, 1351-1356,
1290 doi:10.1038/nm.1890 (2008).
1291 69 Ercan, D. *et al.* Reactivation of ERK signaling causes resistance to EGFR kinase
1292 inhibitors. *Cancer discovery* **2**, 934-947, doi:10.1158/2159-8290.CD-12-0103 (2012).
1293 70 Hrustanovic, G. *et al.* RAS-MAPK dependence underlies a rational polytherapy strategy
1294 in EML4-ALK-positive lung cancer. *Nature Medicine* **21**, 1038-1047,
1295 doi:10.1038/nm.3930 (2015).
1296 71 Papadakis, A. I. *et al.* SMARCE1 suppresses EGFR expression and controls responses to
1297 MET and ALK inhibitors in lung cancer. *Cell research* **25**, 445-458,
1298 doi:10.1038/cr.2015.16 (2015).
1299 72 Yu, H. A. *et al.* Analysis of tumor specimens at the time of acquired resistance to EGFR-
1300 TKI therapy in 155 patients with EGFR-mutant lung cancers. *Clin Cancer Res* **19**, 2240-
1301 2247, doi:10.1158/1078-0432.CCR-12-2246 (2013).
1302 73 Sequist, L. V. *et al.* Genotypic and histological evolution of lung cancers acquiring
1303 resistance to EGFR inhibitors. *Sci Transl Med* **3**, 75ra26,
1304 doi:10.1126/scitranslmed.3002003 (2011).
1305 74 Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell*
1306 **149**, 979-993, doi:10.1016/j.cell.2012.04.024 (2012).
1307 75 Ribas, A. Adaptive Immune Resistance: How Cancer Protects from Immune Attack.
1308 *Cancer Discov* **5**, 915-919, doi:10.1158/2159-8290.CD-15-0563 (2015).
1309 76 Sieuwerts, A. M. *et al.* Elevated APOBEC3B correlates with poor outcomes for estrogen-
1310 receptor-positive breast cancers. *Hormones & cancer* **5**, 405-413,
1311 doi:10.1007/s12672-014-0196-8 (2014).
1312 77 Leshchiner, I. *et al.* Comprehensive analysis of tumour initiation, spatial and temporal
1313 progression under multiple lines of treatment. *biorxiv.org*, doi:10.1101/508127.
1314 78 Lee, J.-K. *et al.* Clonal History and Genetic Predictors of Transformation Into Small-Cell
1315 Carcinomas From Lung Adenocarcinomas. *Journal of Clinical Oncology*,
1316 JCO.2016.2071.2909, doi:10.1200/JCO.2016.71.9096 (2017).

1317 79 Ma, W. *et al.* APOBEC3B promotes hepatocarcinogenesis and metastasis through novel
1318 deaminase-independent activity. *Mol Carcinog* **58**, 643-653, doi:10.1002/mc.22956
1319 (2019).
1320 80 Xie, S., Cooley, A., Armendariz, D., Zhou, P. & Hon, G. C. Frequent sgRNA-barcode
1321 recombination in single-cell perturbation assays. *PLoS ONE* **13**, e0198635,
1322 doi:10.1371/journal.pone.0198635 (2018).
1323 81 Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for
1324 CRISPR screening. *Nature Methods* **11**, 783-784, doi:10.1038/nmeth.3047 (2014).
1325 82 Barger, C. J., Branick, C., Chee, L. & Karpf, A. R. Pan-Cancer Analyses Reveal
1326 Genomic Features of FOXM1 Overexpression in Cancer. *Cancers* **11**,
1327 doi:10.3390/cancers11020251 (2019).
1328 83 Burns, M. B. *et al.* APOBEC3B is an enzymatic source of mutation in breast cancer.
1329 **494**, 366-370, doi:10.1038/nature11881 (2013).
1330 84 Suzuki, K., Bose, P., Leong-Quong, R. Y., Fujita, D. J. & Riabowol, K. REAP: A two
1331 minute cell fractionation method. *BMC research notes* **3**, 294, doi:10.1186/1756-0500-3-
1332 294 (2010).
1333 85 Brown, W. L. *et al.* A Rabbit Monoclonal Antibody against the Antiviral and Cancer
1334 Genomic DNA Mutating Enzyme APOBEC3B. *Antibodies (Basel, Switzerland)* **8**,
1335 doi:10.3390/antib8030047 (2019).
1336 86 Murillo, M. M. *et al.* Disruption of the Interaction of RAS with PI 3-Kinase Induces
1337 Regression of EGFR-Mutant-Driven Lung Cancer. *CellReports* **25**, 3545-3553.e3542,
1338 doi:10.1016/j.celrep.2018.12.003 (2018).
1339 87 Tichelaar, J. W., Lu, W. & Whitsett, J. A. Conditional expression of fibroblast growth
1340 factor-7 in the developing and mature lung. *The Journal of biological chemistry* **275**,
1341 11858-11864, doi:10.1074/jbc.275.16.11858 (2000).
1342 88 Dupage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using
1343 adenoviral or lentiviral delivery of Cre recombinase. *Nature Protocols* **4**, 1064-1072,
1344 doi:10.1038/nprot.2009.95 (2009).
1345 89 Humpton, T. J. *et al.* A noninvasive iRFP713 p53 reporter reveals dynamic p53 activity
1346 in response to irradiation and liver regeneration in vivo. *Sci Signal* **15**, eabd9099,
1347 doi:10.1126/scisignal.abd9099 (2022).

1348 90 Renne, R. *et al.* Proliferative and nonproliferative lesions of the rat and mouse respiratory
1349 tract. *Toxicol Pathol* **37**, 5S-73S, doi:10.1177/0192623309353423 (2009).
1350 91 Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery
1351 in cancer by exome sequencing. *Genome Res* **22**, 568-576, doi:10.1101/gr.129684.111
1352 (2012).
1353 92 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and
1354 heterogeneous cancer samples. *Nature biotechnology* **31**, 213-219, doi:10.1038/nbt.2514
1355 (2013).
1356 93 Fang, H. *et al.* Indel variant analysis of short-read sequencing data with Scalpel. *Nat*
1357 *Protoc* **11**, 2529-2548, doi:10.1038/nprot.2016.150 (2016).
1358 94 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic
1359 variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164,
1360 doi:10.1093/nar/gkq603 (2010).
1361 95 Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. Predicting the functional
1362 effect of amino acid substitutions and indels. *PLoS One* **7**, e46688,
1363 doi:10.1371/journal.pone.0046688 (2012).
1364 96 Choi, Y. in *Proceedings of the ACM Conference on Bioinformatics, Computational*
1365 *Biology and Biomedicine - BCB '12* 414-417 (2012).
1366 97 Choi, Y. & Chan, A. P. PROVEAN web server: a tool to predict the functional effect of
1367 amino acid substitutions and indels. *Bioinformatics* **31**, 2745-2747,
1368 doi:10.1093/bioinformatics/btv195 (2015).
1369 98 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for
1370 analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303,
1371 doi:10.1101/gr.107524.110 (2010).
1372 99 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**,
1373 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).
1374 100 Gerstung, M. *et al.* Reliable detection of subclonal single-nucleotide variants in tumour
1375 cell populations. *Nat Commun* **3**, 811, doi:10.1038/ncomms1814 (2012).
1376 101 Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S. & Swanton, C.
1377 DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA

1378 repair deficiencies and patterns of carcinoma evolution. *Genome Biology* **17**, 31-11,
1379 doi:10.1186/s13059-016-0893-4 (2016).
1380 102 Ramos, A. H. *et al.* Oncotator: cancer variant annotation tool. *Human mutation* **36**,
1381 E2423-2429, doi:10.1002/humu.22771 (2015).
1382 103 Taylor-Weiner, A. *et al.* DeTiN: overcoming tumor-in-normal contamination. *Nature*
1383 *Methods* **15**, 531-534, doi:10.1038/s41592-018-0036-9 (2018).
1384 104 Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature*
1385 **536**, 285-291, doi:10.1038/nature19057 (2016).
1386 105 Consortium, A. P. G. AACR Project GENIE: Powering Precision Medicine through an
1387 International Consortium. *Cancer discovery* **7**, 818-831, doi:10.1158/2159-8290.CD-17-
1388 0151 (2017).
1389 106 Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. **500**, 415-
1390 421, doi:10.1038/nature12477 (2013).
1391 107 Anders, S. & Huber, W. Differential expression analysis for sequence count data.
1392 *Genome Biology* **11**, R106-112, doi:10.1186/gb-2010-11-10-r106 (2010).
1393 108 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for
1394 interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-
1395 15550, doi:10.1073/pnas.0506580102 (2005).

1396

1397 **Supplementary Tables**
1398
1399 **Supplementary Table 1: a,** Metadata of patient tumour samples analyzed using single-cell RNA-
1400 Seq analysis. **b,** Metadata of patient tumour samples analyzed using bulk RNA-Seq analysis. **c**,
1401 Metadata of human tumour samples stained for APOBEC3B via immunohistochemistry.
1402
1403 **Supplementary Table 2:** GSEA analysis of long-term osimertinib-treated APOBEC3B-proficient
1404 PC9 E4 derivatives compared to APOBEC3B-deficient derivatives.
1405
1406 **Supplementary Table 3:** Metadata of patient tumor samples processed for whole exome
1407 sequencing and subsequent mutational signature analysis.
1408
1409 **Supplementary Table 4: a**, Mutations observed in EGFR- and ALK-driven lung cancer patients.
1410 **b**, Selective putative resistance mutations observed in *EGFR*- and *ALK*-driven lung cancer
1411 patients.

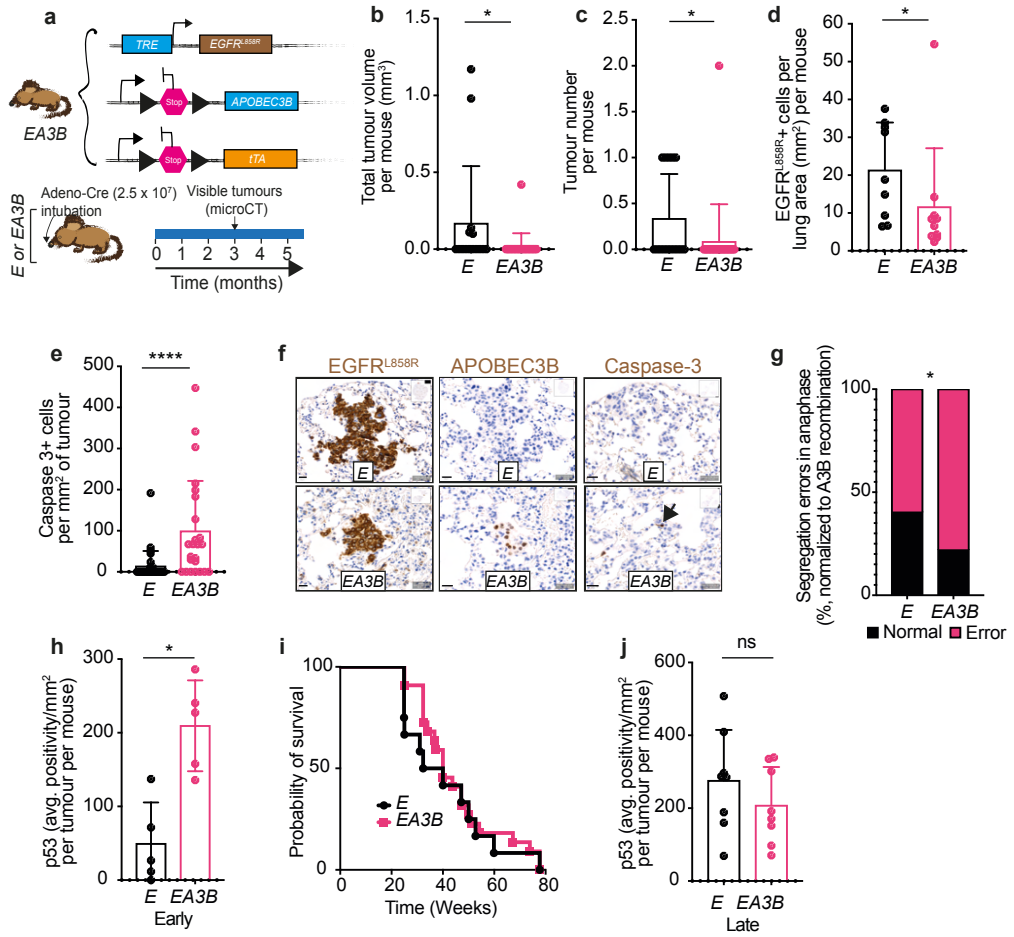


Fig. 1| APOBEC3B is detrimental for tumourigenesis in an *EGFR*^{L858R} mouse model of lung cancer. a, Tumourigenesis in *E* and *EA3B* mice was induced using the indicated viral titer (2.5×10^7 viral particles/mouse). Tumour growth was assessed by microCT analysis. **b,** Total tumour volume per mouse at 3 months post-induction quantified by microCT analysis (*E* = 15, *EA3B* = 24, mean with SD, two-sided Mann-Whitney Test, P = 0.0163, each dot represents a mouse). **c**, Total tumour number per mouse at 3 months post-induction quantified by microCT analysis (*E* = 15, *EA3B* = 24, mean with SD, two-sided Mann-Whitney Test, P = 0.0236, each dot represents a mouse). **d,** Quantification of EGFR^{L858R}-positive cells per lung area (mm^2) by immunohistochemical (IHC) staining at 3 months post-induction (*E* = 9, *EA3B* = 10, mean with SD, two-sided Mann-Whitney Test, P =0.0435, each dot represents a mouse). **e,** Quantification of caspase-3-positive cells per mm^2 of tumour at 3 months post-induction (*E* = 9, *EA3B* = 10, mean with SD, two-sided Mann-Whitney Test, P<0.0001, each dot represents a tumour). **f,** Representative IHC stainings of EGFR^{L858R}, APOBEC3B, and caspase-3 (scale bar = 20 µm, arrow indicates positive cell). **g**, Percent chromosome missegregation errors at 3 months post-induction (two-sided Fisher's exact test, p = 0.016, *E* = 9, *EA3B* =10). **h**,Quantification of p53 positive cells per lung area by IHC staining at 3 months post-induction (*E* = 5, *EA3B* = 5, mean with SD, two-sided Mann-Whitney Test, p = 0.0159). **i**, Survival curve of *E* versus *EA3B* mice (*E* =15, *EA3B* = 24, each dot represents a mouse). **j**, Quantification of p53 positive cells per lung area by IHC staining at late timepoint (termination) (*E* = 8, *EA3B* = 8, mean with SD).

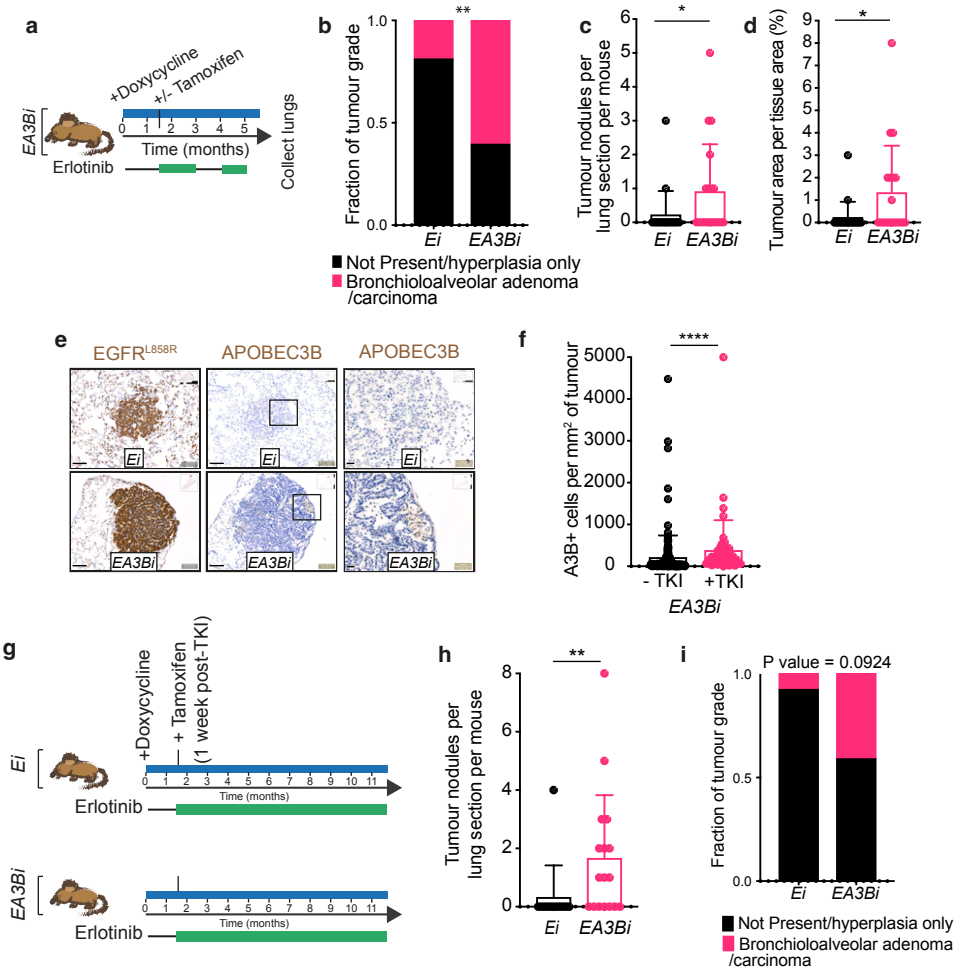


Fig. 2| Expression of hA3B in combination with TKI therapy induces adaptations favorable for A3B expression and TKI resistance. **a,** Experimental set up of A3B induction post-initiation of TKI therapy using *TetO-EGFR^{L858R};CCSP-rtTA;R26$^{LSL-APOBEC3B/Cre-ER(T2)}$(EA3Bi)* mice. **b,** Fraction of tumour grade, divided into no tumours present (not present) or hyperplasia only and bronchioloalveolar adenoma or carcinoma, in *Ei* and *EA3Bi* mice at 5 months post-induction post-TKI treatment (*Ei* = 19, *EA3Bi* = 19, two sided Fisher's exact test, **P= 0.0044). **c,** Tumour nodules per lung section per mouse at 5 months post induction post-TKI treatment (*Ei* =19, *EA3Bi* = 19,median = dashed line, 1st and 3rd quartiles = dotted lines, two-sided Mann-Whitney Test, *P =0.0443). **d,**Tumour area per lung area per mouse at 5 months post-induction post-TKI treatment (*Ei*=19, *EA3Bi* = 19, median = dashed line, 1st and 3rd quartiles = dotted lines, two-sided Mann-Whitney Test, *P =0.0212). **e,** Representative immunohistochemical staining of EGFR^{L858R} and A3B. Scale bar = 100 µm and 20 µm. **f,** A3B positive cells/mm^2 of tumour per mouse (*EA3Bi* - TKI = 151, *EA3Bi* + TKI = 52, median = dashed line, 1st and 3rd quartiles = dotted lines, two sided Mann Whitney Test, **** P = <0.0001) **g**, Experimental set up of induction of subclonal APOBEC3B using *TetO-EGFR^{L858R};CCSP-rtTA;R26$^{LSL-APOBEC3B/Cre-ER(T2)}$(EA3Bi)* or *TetO-EGFR^{L858R};CCSP-rtTA;R26$^{Cre-ER(T2)/+}$(Ei)* mice with continuous TKI therapy (erlotinib). **h**, Tumour nodules per lung section per mouse (*Ei* = 13, *EA3Bi* = 17, median = dashed line, 1st and 3rd quartiles = dotted lines, two sided Mann Whitney Test, ** P = 0.0086). **i**, Fraction of tumour grade in *Ei* and *EA3Bi* mice (*E* =13, *EA3Bi* = 17, NP = not present, H = hyperplasia, A =bronchioloalveolar adenoma, WDC = bronchioloalveolar carcinoma, well differentiated, Fisher's exact test).

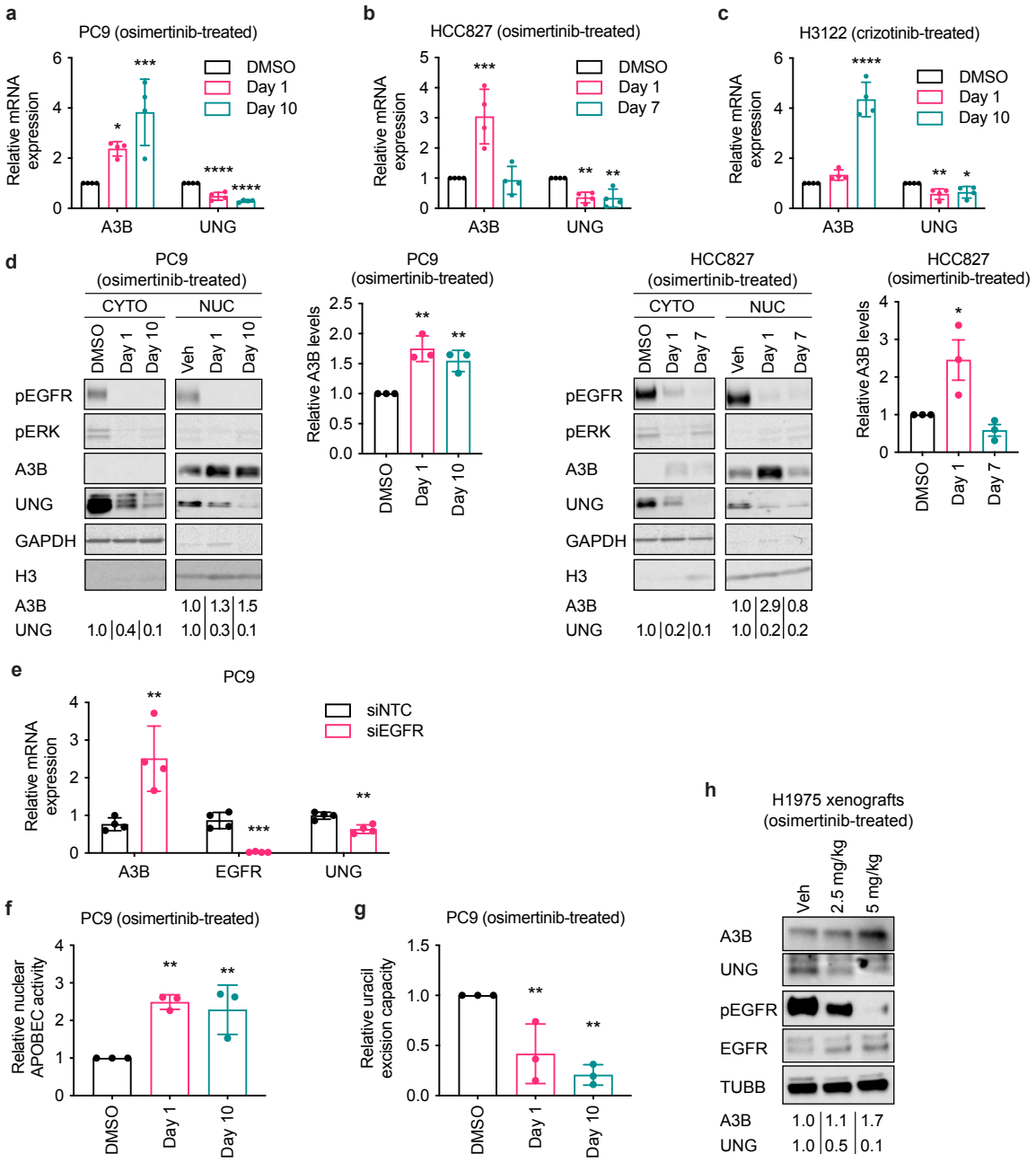


Fig. 3| Treatment with tyrosine kinase inhibitors (TKIs) induces A3B upregulation and UNG downregulation. **a,** RT-qPCR analysis of PC9 cells treated with DMSO or 2 μM osimertinib (osi) for 18 hours or 10 days (n = 4 biological replicates, mean ± SEM, ANOVA test, *P = 0.0349, ***P = 0.0006, ****P < 0.0001). **b**, RT-qPCR analysis of HCC827 cells treated with DMSO or 0.4 μM osi for 18 hours or 7 days (n = 4 biological replicates, mean ± SEM, ANOVA test, ***P = 0.0008, **P = 0.0014, **P = 0.0011). **c**, RT-qPCR analysis of H3122 cells treated with DMSO or 0.5 μM crizotinib (criz) for 18 hours or 10 days (n = 4 biological replicates, mean ± SEM, ANOVA test, ****P < 0.0001, **P = 0.0086, *P = 0.0219). **d**, Western blot analysis of cells treated in a and b (CYTO: cytoplasmic extracts, NUC: nuclear extracts, H3: Histone H3) with quantification of A3B levels in PC9 cells (n = 3 biological replicates, mean ± SEM, ANOVA test, **P = 0.0012, **P = 0.0058) and HCC827 cells (n = 3 biological replicates, mean ± SEM, ANOVA test, *P = 0.0186). **e,** RT-qPCR analysis of PC9 cells treated with non-targeting siRNA (siNTC) or EGFR siRNA (siEGFR) for 18 hours and grown for 2 days (n = 4 biological replicates, mean ± SEM, t-test, **P = 0.0075, ***P = 0.0002, **P = 0.0027). **f,** APOBEC activity assay performed using nuclear extracts of PC9 cells treated as in panel a (n = 3 biological replicates, mean ± SD, ANOVA test, **P = 0.0036, **P = 0.0073). **g,** Uracil excision capacity assay using nuclear extracts of PC9 cells treated as in panel a (n = 3 biological replicates, mean ± SD, ANOVA test, **P = 0.0077, **P = 0.0017). **h**, Western blot analysis using extracts of *EGFR*-mutant H1975 human NSCLC xenografts harvested after 4 days of treatment with vehicle or the indicated doses of osimertinib (TUBB: Tubulin Beta Class I).

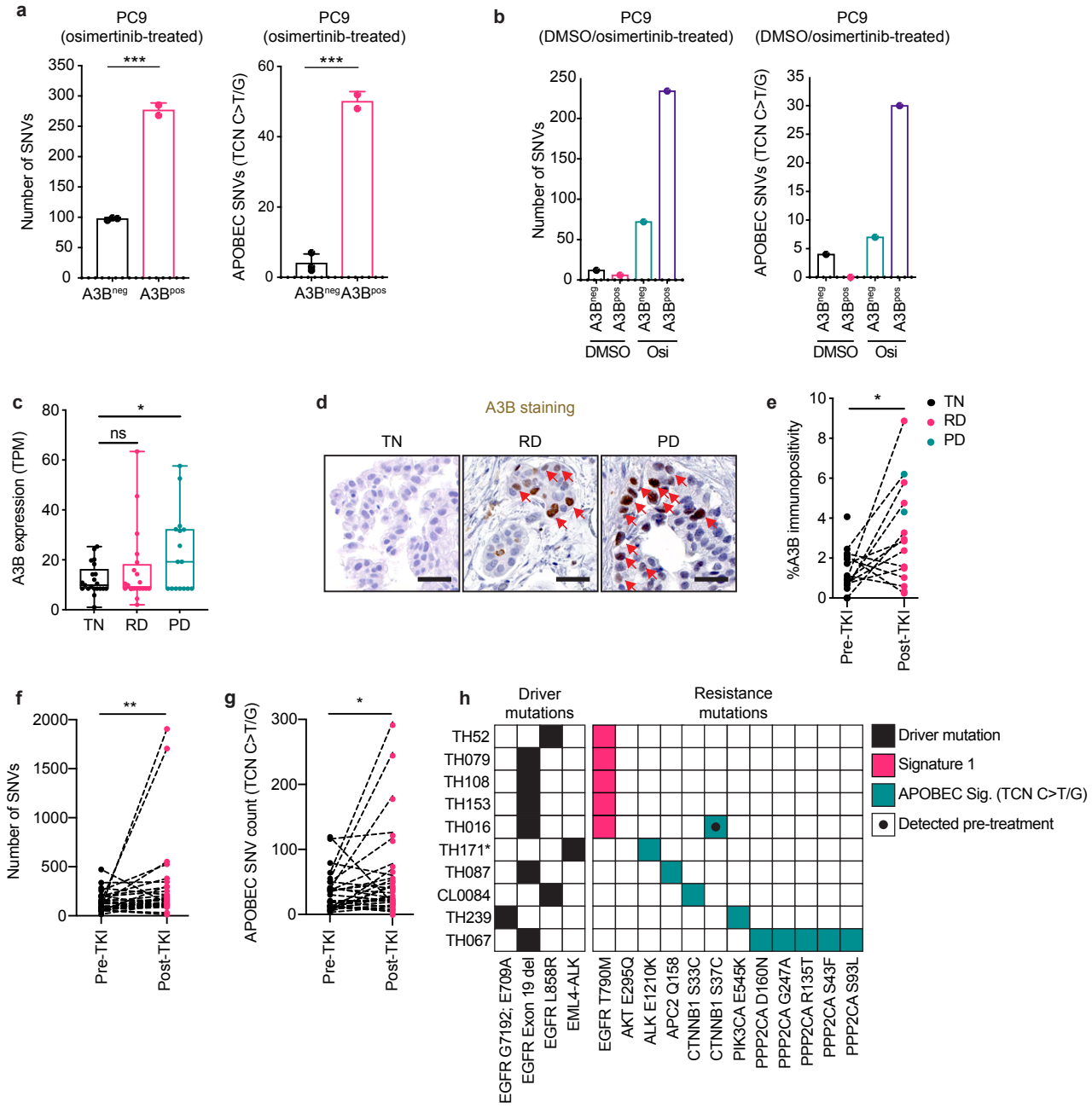
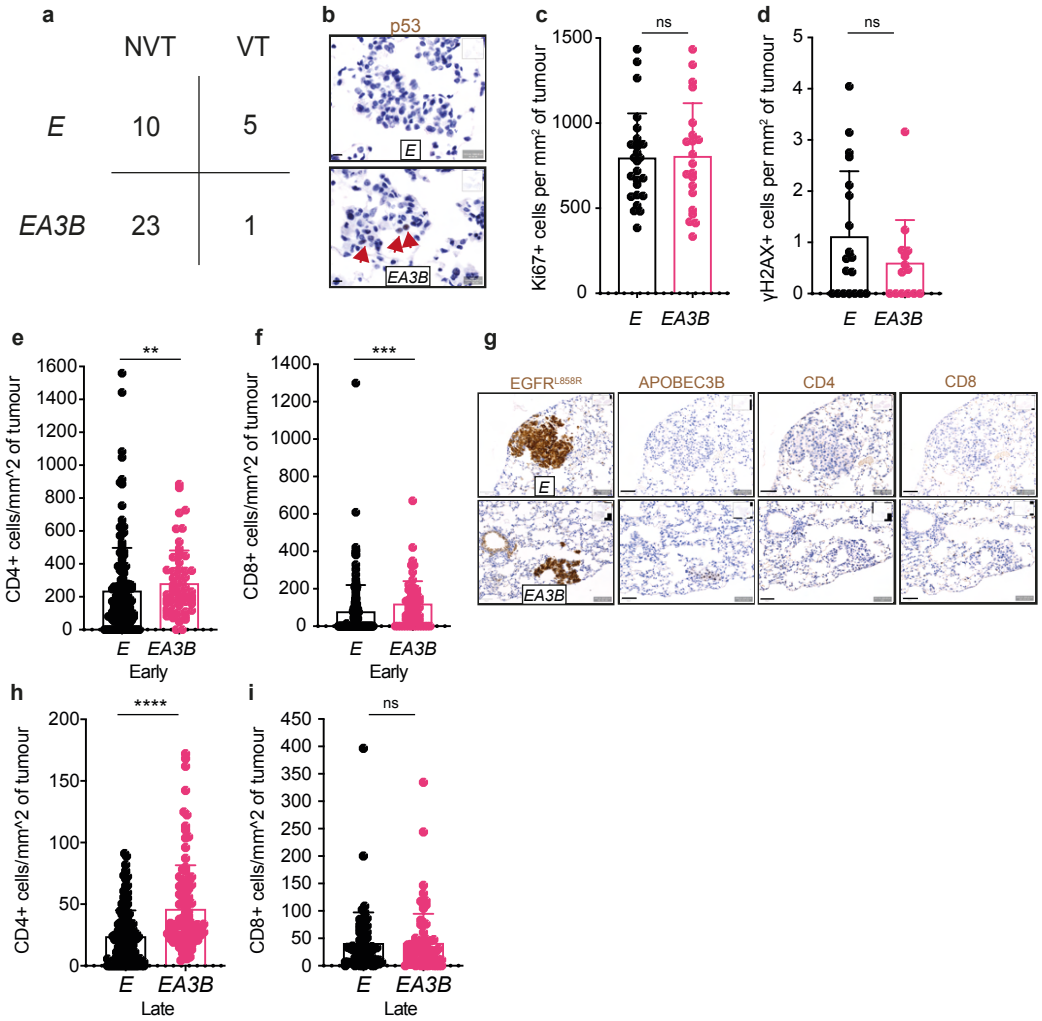
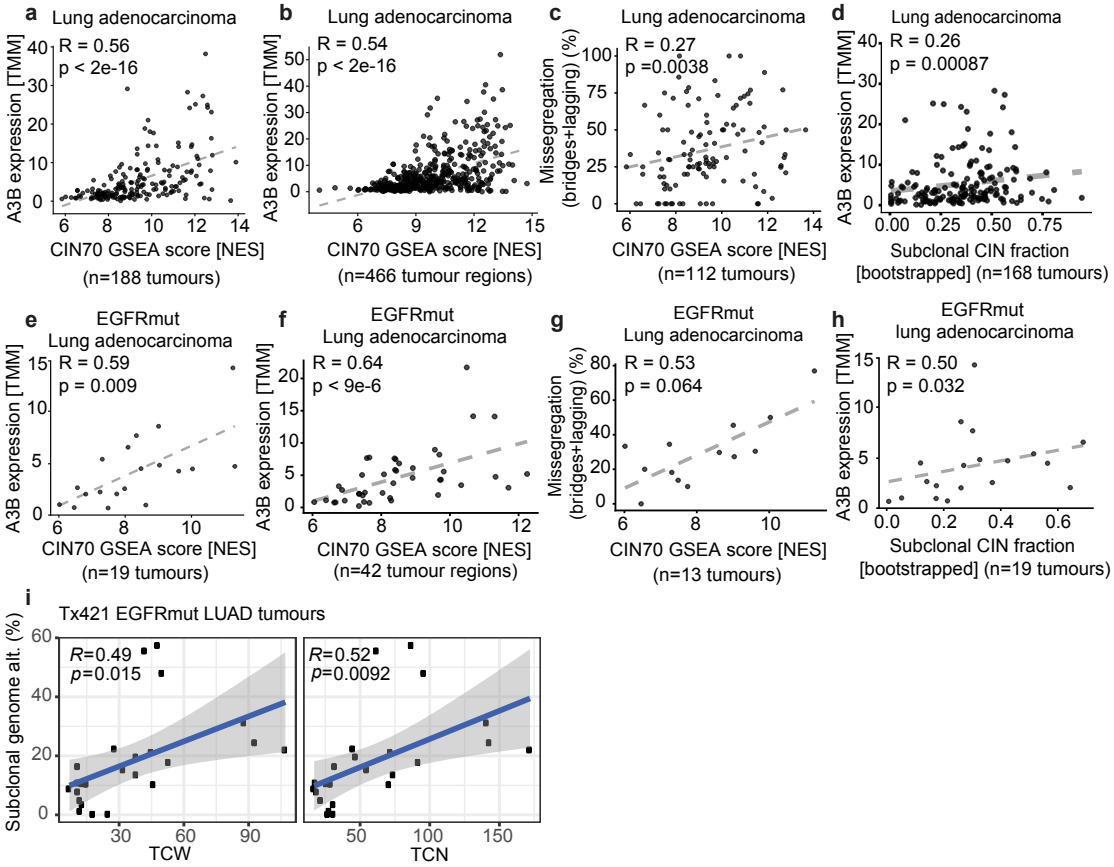


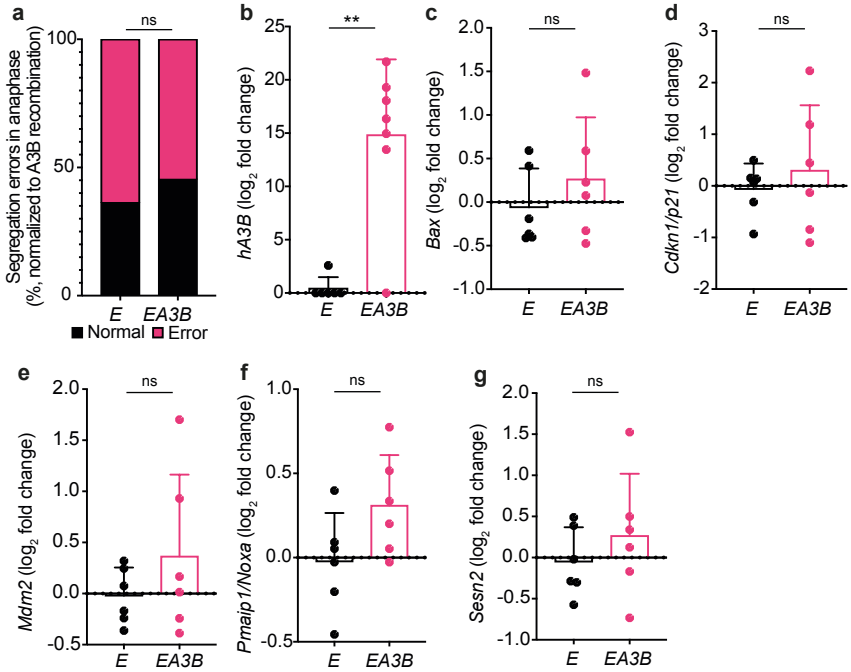
Fig. 4| APOBEC3B generates genetic diversity that fuels resistance to oncogene-targeted therapy. **a**, Whole exome sequencing analysis of mutation burden and APOBEC-associated mutations (TCN, C>T/C>G) acquired during osimertinib treatment of A3B proficient or deficient PC9 cells (Knockout verified in Ext. Data. Fig18a, unpaired t-test, ***P < 0.0005). **b,** Whole genome sequencing analysis of mutation burden and APOBEC-associated mutations (TCN, C>T/C>G) acquired with DMSO or osimertinib treatment of A3B proficient or deficient PC9 cells (Knockout verified in Ext. Data. Fig. 18b-c, DMSO 1 month, osimertinib 2.5 months). **c**, Comparison of APOBEC3B expression levels (Exp: batch corrected TPM) measured using RNA-Seq analysis in human NSCLC specimens obtained before treatment (treatment naïve, TN), or on-treatment at residual disease (RD) or at progressive disease (PD) stages from lung cancer patients undergoing treatment with tyrosine kinase inhibitors (all data points shown; ANOVA test, mixed effects model, *P = 0.0318). **d**, Representative images of IHC analysis of A3B protein levels in patients with NSCLC at TN, RD and PD stages, red arrows indicate positive stained cells (scale bar: 30 μM). **e**, IHC quantification of human NSCLC samples obtained before treatment (Pre-TKI) or after treatment (Post-TKI) (all data points shown, unpaired t-test, *P = 0.0113, patients treated with single agent TKI). **f,** Quantification of total mutation burden (SNV count) in human NSCLC samples obtained before (Pre-TKI) or after (Post-TKI) treatment (Wilcoxon matched-pairs signed rank test, two tailed, ** P = 0.0023). **g,** Quantification of APOBEC-associated mutation count in human NSCLC samples obtained before treatment (Pre-TKI) or after treatment (Post-TKI) (Wilcoxon matched-pairs signed rank test, two tailed, ** P = 0.0152). **h**, Mutation signature associated with each putative de novo TKI resistance mutation detected in clinical patient samples analyzed post-TKI at progressive disease. *Denotes sample from patient having received prior chemotherapy.



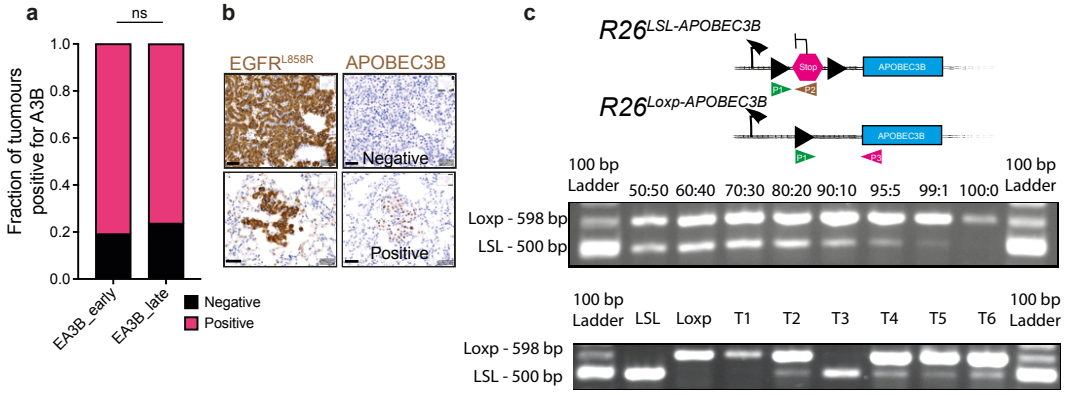
Extended Data Fig. 1| APOBEC3B is detrimental for tumourigenesis in an *EA3B* mouse model of lung cancer. a, Two by two contingency table of the number of mice with visible tumours (VT) or no visible tumours (NVT) by microCT at 3 months post-induction (two-sided Fisher's exact test, P<0.05). **b**, Representative images of p53 nuclear immunohistochemical staining. Scale bar = 10 µm, arrows indicates positive cells. **c**, Quantification of Ki67-positive cells per mm^2 of tumour at 3 months post-induction (*E* = 9, *EA3B* = 10, each dot represents a tumour). **d**, Quantification of γH2AX-positive cells per mm^2 of tumour at 3 months post-induction (*E* = 9, *EA3B* = 10, each dot represents a tumour). **e**, Quantification of CD4+ cells per mm^2 of tumour at 3 months post-induction (Mice *E* =8 *EA3B* = 7, each dot represents a tumour, mean with SD, two-sided Mann-Whitney Test, P = 0.0086). **f**, Quantification of CD8+ cells per mm^2 of tumour at 3 months post-induction (Mice *E* =8 *EA3B* = 8, each dot represents a tumour, mean with SD, two-sided Mann-Whitney Test, P = 0.0003). **g**, Representative IHC stainings of EGFR^{L858R}, APOBEC3B, and CD4 and CD8 T cells (scale bar = 50 µm). **h**, Quantification of CD4+ cells per mm^2 of tumour at late timepoint (termination) (Mice *E* =8 *EA3B* = 7, each dot represents a tumour, mean with SD, two-sided Mann-Whitney Test, P < 0.0001). **i,** Quantification of CD8+ cells per mm^2 of tumour at late timepoint (termination)(Mice *E* =6 *EA3B* = 6, each dot represents a tumour, mean with SD).



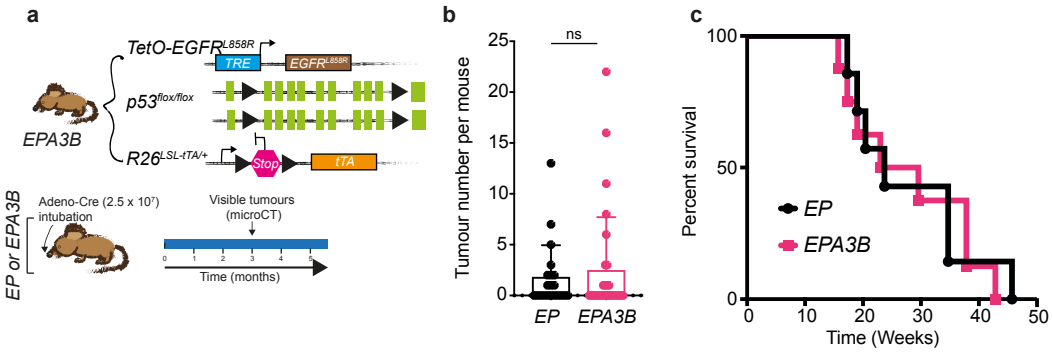
Extended Data Fig. 2| APOBEC3B expression correlates multiple measures of chromosomal instability in TRACERx 421. **a,** Significant correlation between A3B expression and CIN70 GSEA score calculated using tumours from patients with lung adenocarcinoma in the TRACERx421 cohort (n = 188 tumours, Spearman, rho = 0.56, p <2e-16). **b,** Significant correlation between A3B expression and CIN70 GSEA score calculated using tumour regions in patients with lung adenocarinoma in the TRACERx421 cohort (n=466 tumour regions, Spearman, rho =0.54 p <2e-16). **c,** Significant correlation between percent missegregation errors (anaphase bridges (bridges) and lagging chromosomes (lagging)) and CIN70 score calculated using tuours from patients in the TRACERx300 cohort (n = 112 tumours, Spearman, rho =0.27, p=0.0038). **d,** Correlation between A3B expression and subclonal CIN fraction calculated patients with lung adenocarcinoma in the TRACERx421 cohort (n= 168 tumours, Bootstrapped Spearman, rho = 0.26 p=0.00087). **e**, Significant correlation between A3B expression and CIN70 GSEA score calculated using EGFR mutant (mut) tumours from patients with lung adenocarcinoma in the TRACERx421 cohort (n = 19 tumours, Spearman, rho = 0.59, p =0.009). **f**, Significant correlation between A3B expression and CIN70 GSEA score calculated using EGFRmut tumour regions in patients with lung adenocarinoma in the TRACERx421 cohort (n=42 tumour regions, Spearman, rho =0.64 p <9e-6). **g**, Correlation between percent missegregation errors and CIN70 score calculated using EGFRmut tumours from patients in the TRACERx300 cohort (n = 13 tumours, Spearman, rho =0.53, p=0.064). **h**, Correlation between A3B expression and subclonal CIN fraction calculated in EGFRmut patients with lung adenocarcinoma in the TRACERx421 cohort (n= 168 tumours, Bootstrapped Spearman, rho = 0.5 p=0.032). **i**, Comparisons between C>T and C>G mutation counts at TCN and TCW trinucleotide context and % genome altered subclonally (Pearson, TCW rho = 049, p =0.015, TCN rho = 0.52, p =0.0092).



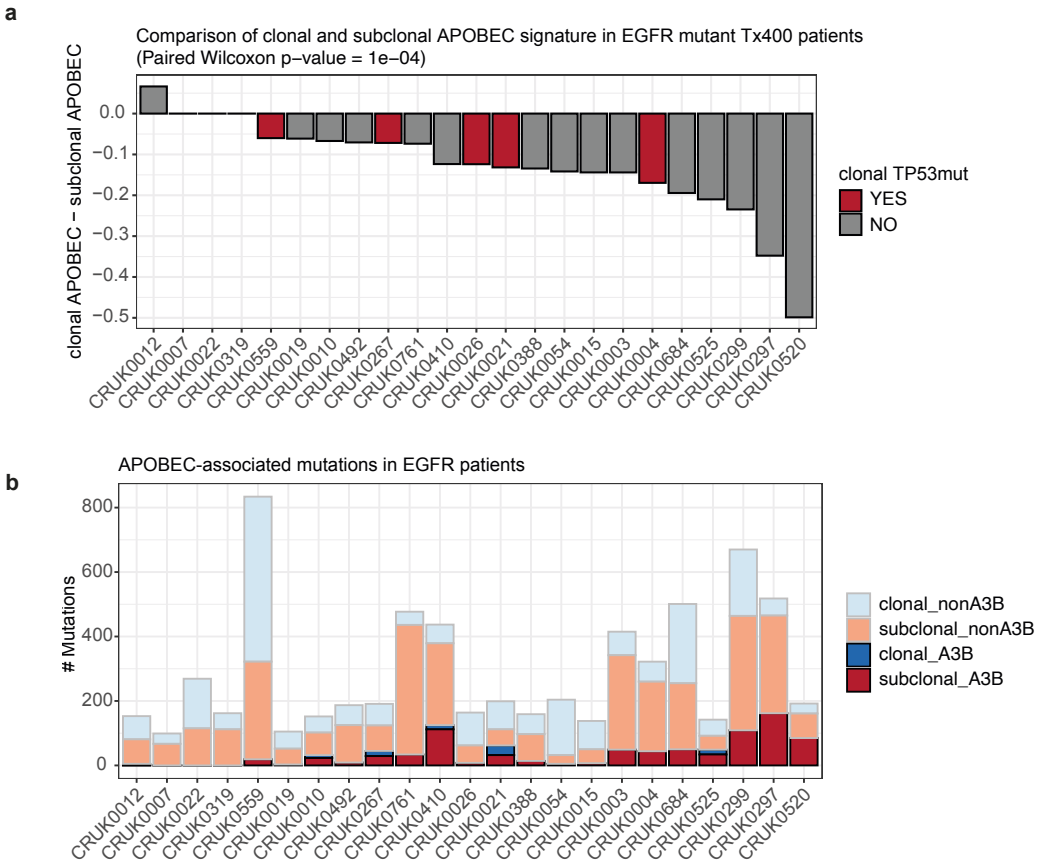
Extended Data Fig. 3| No increase in CIN or p53 activation is observed with continous A3B expression late in tumour evolution . a, Percent chromosome missegregation errors at late timepoint (termination) (*E* = 9, *EA3B* =10)**. b,** *A3B* expression in *E* and *EA3B* tumours at termination (*E* =6, *EA3B* = 7, $\log_2$ fold change). **c**, *Bax* expression in *E* and *EA3B* tumours at termination (*E* =6, *EA3B* = 6, $\log_2$ fold change, A3B negative tumour in EA3B mouse identified in b removed from subsequent analysis of p53 activation). **d**, *Cdkn1/p21* expression in *E* and *EA3B* tumours at termination (*E* =6, *EA3B* = 6, $\log_2$ fold change, A3B negative tumour in EA3B mouse identified in b removed from analysis). **e**, *Mdm2* expression in *E* and *EA3B* tumours at termination (*E* =6, *EA3B* = 6, $\log_2$ fold change, A3B negative tumour in EA3B mouse identified in b removed from analysis). **f**, *Pmaip1/Noxa* expression in *E* and *EA3B* tumours at termination (*E* =6, *EA3B* = 6, $\log_2$ fold change, A3B negative tumour in EA3B mouse identified in b removed from analysis). **g**, *Sesn2* expression in *E* and *EA3B* tumours at termination (*E* =6, *EA3B* = 6, $\log_2$ fold change, A3B negative tumour in EA3B mouse identified in b removed from analysis).



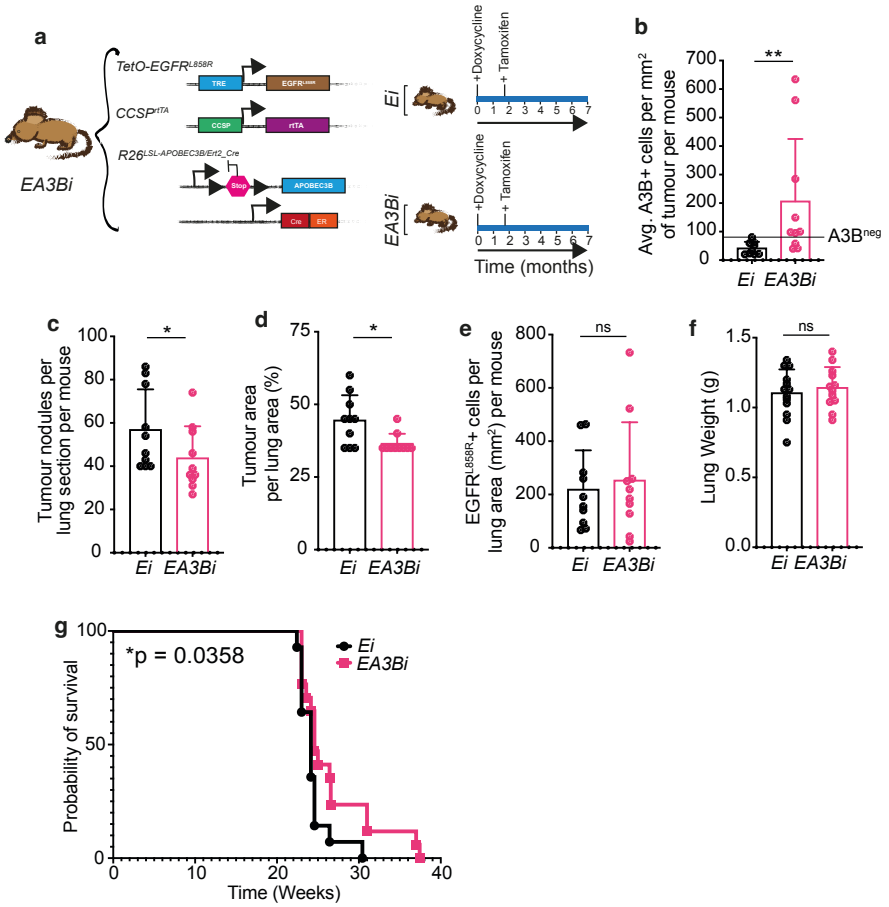
Extended Data Fig. 4| Recombination efficiency of LSL-APOBEC3B. a, Fraction of tumours positive for A3B at early and late timepoints measured by IHC staining (Early E = 3, EA3B = 3, Late E = 8, EA3B = 8). **b**, Representative IHC staining of EGFR^{L858R}, and APOBEC3B. Scale bar = 50 µm. **c**, PCR analysis of recombination efficiency of LoxP-Stop-LoxP site upstream of human APOBEC3B transgene. Primer 1 (P1) in LoxP, Primer 2 (P2) in stop codon, Primer 3 (P3) in human APOBEC3B transgene. P1+P2 = LSL 500 bp, P1+P3 = 598 bp. *EA3B* tumours at a late timepoint (termination). LSL = unrecombined tail DNA, Loxp = fully recombined cell line DNA. Ratios represent fraction of unrecombined tail DNA to fully recombined cell line DNA.



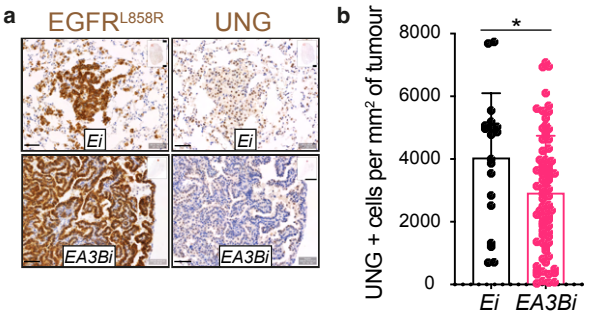
Extended Data Fig. 5| TP53 pathway inactivation rescues tumour growth with continous A3B expression. a, Tumours in *EP* and *EPA3B* mice were induced using the indicated viral titer (2.5×10^7 viral particles/mouse). Tumour growth was assessed by microCT analysis. **b,** Total tumour number per mouse at 3 months post-induction quantified by microCT analysis (*EP* = 21, *EPA3B* = 30, median = dashed line, 1st and 3rd quartiles = dotted lines. Combined from two separate experiments). **c,** Survival curve of *EP* versus *EPA3B* mice (*EP* =8, *EPA3B* = 7, each dot represents a mouse).



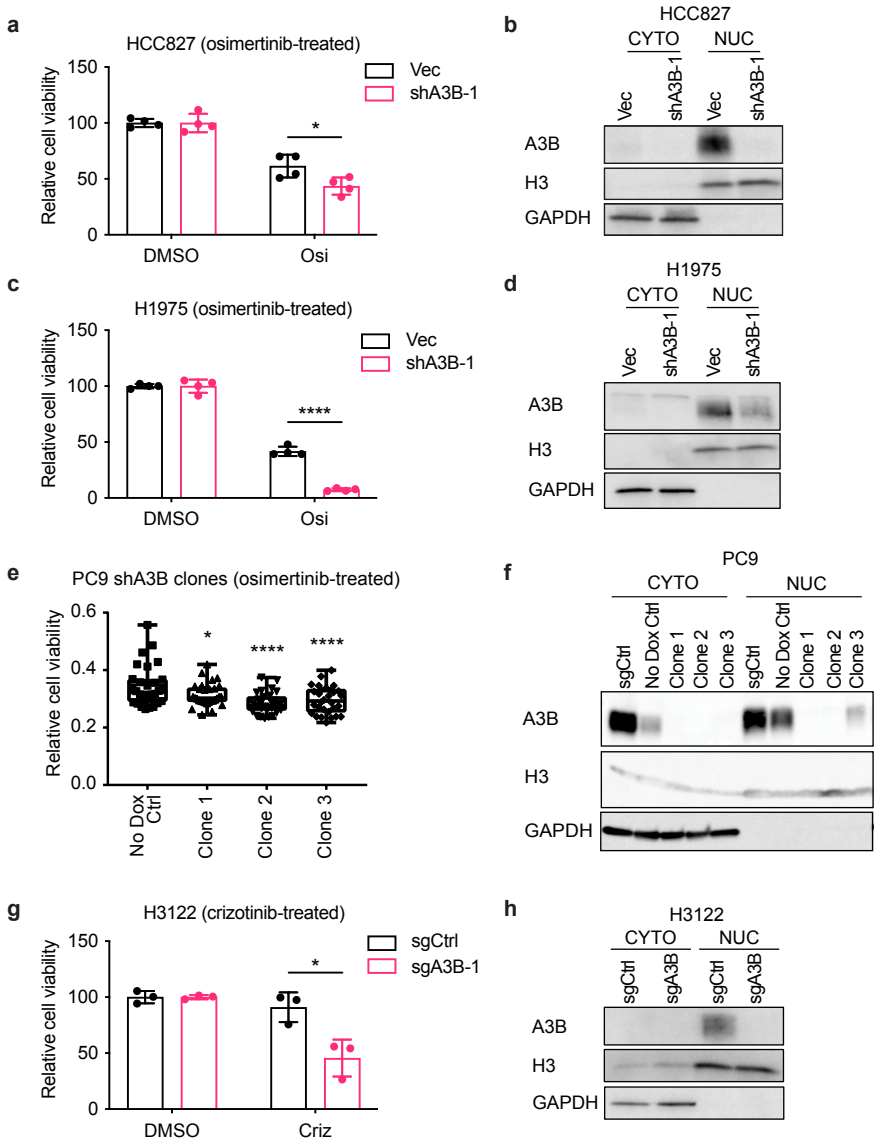
Extended Data Fig. 6| Subclonal APOBEC mutagenesis in EGFR mutant patients from TRACERx 400 dataset. a, Comparison of clonal and subclonal APOBEC-associated mutation signature (clonal APOBEC - subclonal APOBEC) in patients with EGFR driver mutations (1, 1a, exon 19 deletion). Grey bars indicate the patient is TP53 wildtype or has a subclonal TP53 mutation. Red bars indicate that the patient has a clonal TP53 mutation. (P = 1e^-4). **b,** Number of APOBEC-associated mutations in patients with EGFR driver mutations (1, 1a, exon 19 deletion). Colors indicate clonal or subclonal APOBEC or non APOBEC-associated mutations.



Extended Data Fig. 7| APOBEC3B expression without TKI therapy inhibits tumour growth. ***a***, Experimental set up of induction of subclonal APOBEC3B using *TetO-EGFR*^{L858R}*;CCSP-rtTA;R26*$^{LSL\text{-}APOBEC3B/Cre\text{-}ER(T2)}$*(EA3Bi)* or *TetO-EGFR*^{L858R}*;CCSP-rtTA;R26*$^{Cre\text{-}ER(T2)/+}$*(CCSP_Ei)* mice. **b**, Average A3B positive cells/mm^2 of tumour per mouse (*Ei* = 9, *CCSP_EA3Bi* = 10, median = dashed line, 1st and 3rd quartiles = dotted lines, two sided Mann-Whitney Test, **P = 0.0041, each dot represents a mouse, dots below A3B^{neg} line are A3B negative). **c**,Tumour nodules per lung section per mouse at termination (*Ei* =10, *CCSP_EA3Bi* = 10, median = dashed line, 1st and 3rd quartiles = dotted lines, two-sided Mann-Whitney Test, *P = 0.0494). **d**, Tumour area per lung area at termination (*Ei* =10, *EA3Bi* = 10, median = dashed line, 1st and 3rd quartiles = dotted lines, two-sided Mann-Whitney Test, *P = 0.0216). **e**, EGFR^{L858R} positive cells per lung area per mouse in *Ei* and *EA3Bi* mice at termination. **f,** Lung weight in *Ei* vs *EA3Bi* mice at termination *Ei* = 13, *EA3Bi* = 12, median = dashed line, 1st and 3rd quartiles = dotted lines, each dot represents a mouse). **g,** Survival curve of *Ei* versus *EA3Bi* mice (*Ei* = 14, *EA3Bi* = 17, each dot represents a mouse, Log-rank (Mantel-Cox) test, *P = 0.0358).



Extended Data Fig. 8| Significant decrease in UNG expression in A3B expressing tumours with TKI treatment. **a,** Representative immuno-histochemical staining of EGFR^{L858R} and UNG. Scale bar = 50 µm. **b,** Quantification of UNG+ cells per mm^2 of tumour at 5 months post-induction (*E* n=10, *EA3Bi* n=10, median = dashed line, 1st and 3rd quartiles = dotted lines, two-sided t-test, P = 0.0226, each dot represents a tumour).



Extended Data Fig. 9| Depletion of A3B impairs survival of lung adenocarcinoma cells subjected to TKI treatment. a, CellTiter-Glo (CTG) viability assays performed on A3B-deficient or A3B-proficient HCC827 cells treated with osimertinib (osi; day 1 to 5: 2 μM; day 6 to 22: 100 nM) (n = 4 biological replicates, data normalized to respective DMSO conditions, mean ± SD, t-test, *P = 0.0306). **b**, Immunoblot confirming A3B knock-down in HCC827 (CYTO: cytoplasmic extracts, NUC: nuclear extracts, H3: Histone H3). **c,** CTG viability assays performed on A3B-deficient or A3B-proficient H1975 cells treated with osi (day 1 to 8: 4 μM) (n = 4 biological replicates, data normalized to respective DMSO conditions, mean ± SD, t-test, ****P < 0.0001). **d,** Immunoblot confirming A3B knockdown in H1975. **e**, CTG viability assays performed on derivatives of PC9 single cell clones transduced with pTLCV2 plasmid derivative expressing sgRNA targeting A3B with or without Cas9 induction and treated with osi (day 1 to 4: 100 nM; day 5 to 42: 500 nM-2 μM, day 42 to 2.5 months: 100 nM) (n = 30, whiskers: min to max, ANOVA test). **f**, Immunoblot confirming knockout of A3B in PC9 clones. **g**, Crystal violet growth assay performed on A3B-deficient or A3B-proficient H3122 cells treated with crizotinib (day 1 to 8: 100 nM; day 8 to 11: 50 nM; day 11 to 15: 100 nM; day 15 to 32: 1-2 μM) (n = 3 biological replicates, data normalized to respective DMSO conditions, mean ± SD, t-test, *P = 0.0206). **h,** Immunoblot confirming A3B knockout in H3122.

a HCC827 6 hr erlotinib vs DMSO (GSE51212)

Gene symbol	logFC	p-value	adj. p-val.
APOBEC3B	1.97	5.37E-10	2.98E-06
UNG	-0.656	5.67E-06	0.000337

HCC4006 1 day erlotinib vs DMSO (GSE57156)

Gene symbol	logFC	p-value	adj. p-val.
APOBEC3B	1.29	0.000226	0.009784
UNG	-0.907	0.000888	0.022115

PC9 8 days erlotinib vs DMSO (GSE67051)

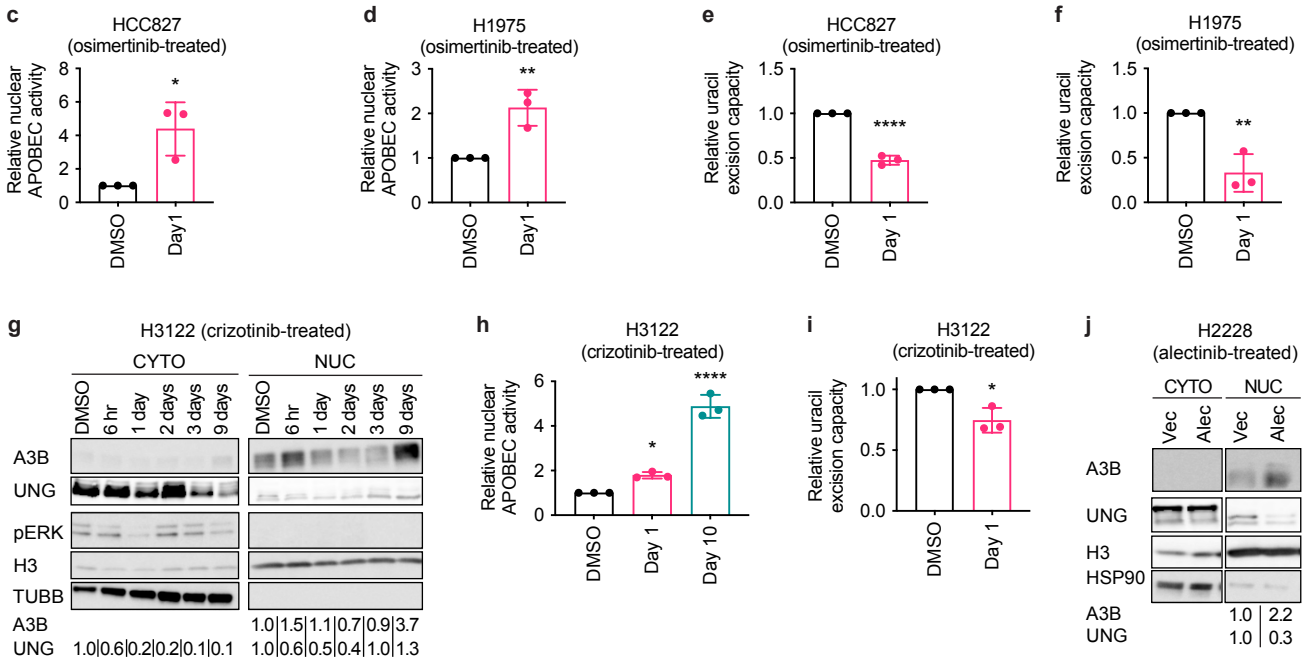
Gene symbol	logFC	p-value	adj. p-val.
APOBEC3B	0.494	7.64E-03	0.039366
UNG	-0.204	0.0673	0.176893

HCC827 6 hr MEK inhibitor (AZD6244) vs DMSO (GSE51212)

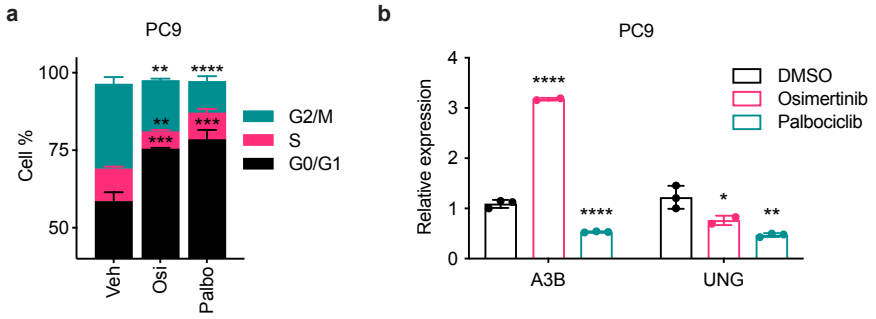
Gene symbol	logFC	p-value	adj. p-val.
APOBEC3B	0.572	1.08E-05	4.69E-03
UNG	-0.135	5.73E-02	3.62E-01

b PC9 9 days osimertinib vs DMSO

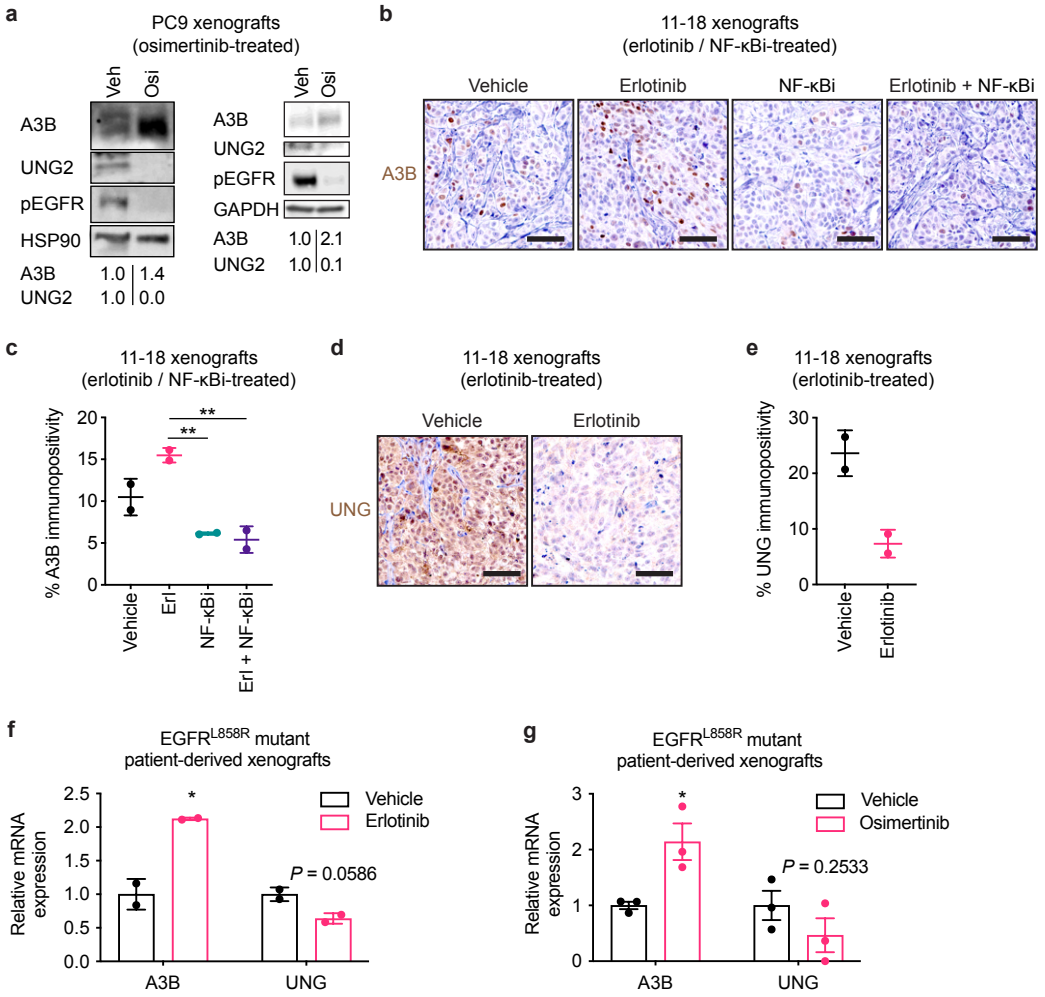
Gene symbol	logFC	p-value	adj. p-val.
APOBEC3B	2.25	5.69E-71	4.28E-69
UNG	-0.797	2.26E-09	1.68E-08



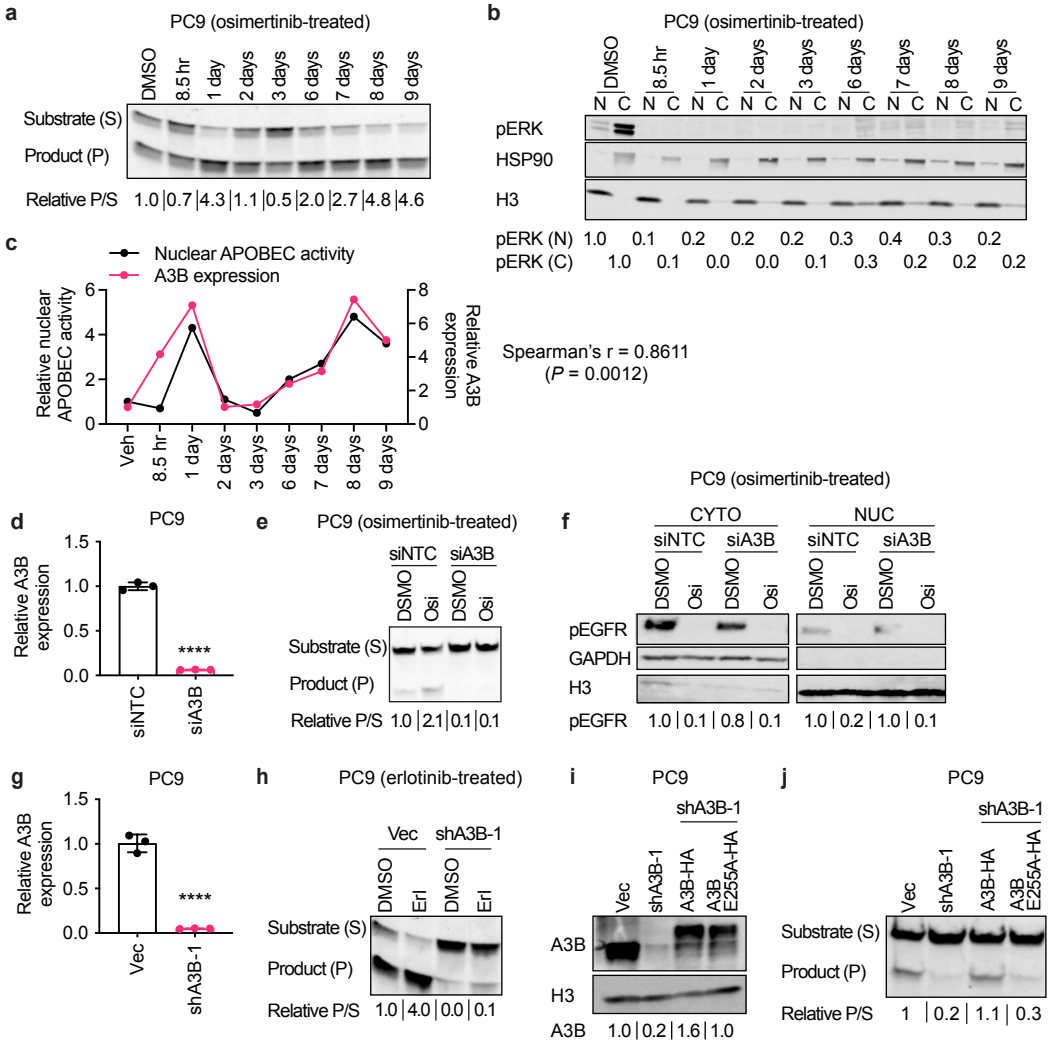
Extended Data Fig. 10| Treatment with tyrosine kinase inhibitors (TKIs) induces increased A3B expression and activity and decreased UNG expression and activity in multiple pre-clinical models of lung adenocarcinoma. **a,** GEO2R analysis of the indicated GSEA datasets of *EGFR*-driven cellular models of human lung adenocarcinoma treated with erlotinib or a MEK inhibitor (AZD6244). **b,** RNA-Seq analysis of gene expression changes in PC9 cells treated with 2 μM osimertinib (osi) for 9 days relative to DMSO-treated cells identified using (n = 3 biological replicates, mean ± SEM, ANOVA test). **c**, APOBEC activity assay using nuclear extracts of HCC827 cells treated with DMSO or 0.4 μM osi for 18 hours (n = 3 biological replicates, mean ± SD, t-test, *P = 0.0213). **d**, APOBEC activity assay using nuclear extracts of H1975 cells treated with DMSO or 1 uM osi for 18 hours (n = 3 biological replicates, mean ± SD, t-test, **P = 0.0084). **e**, Uracil excision capacity assay using extracts from panel c (n = 3 biological replicates, mean ± SD, t-test, ****P < 0.0001). **f,** Uracil excision capacity assay using extracts from part d (n = 3 biological replicates, mean ± SD, t-test, **P = 0.0054). **g**, Western blot analysis of extracts from H3122 cells treated with DMSO or 1 μM crizotinib (criz) for the indicated durations. **h,** APOBEC activity assay performed using nuclear extracts of H3122 cells treated with DMSO or 0.5 μM crizotinib for 18 hours or 10 days (n = 3 biological replicates, mean ± SD, ANOVA, *P = 0.0204, ****P < 0.0001). **i**, Uracil excision capacity assay using extracts from part h (n = 3 biological replicates, mean ± SD, t-test, *P = 0.0123). **j**, Western blot analysis of EML4-ALK positive H2228 human NSCLC cells treated with DMSO or 0.5 μM alectinib (alec) for 18 hours (CYTO: cytoplasmic extracts; NUC: nuclear extracts; H3: Histone H3).



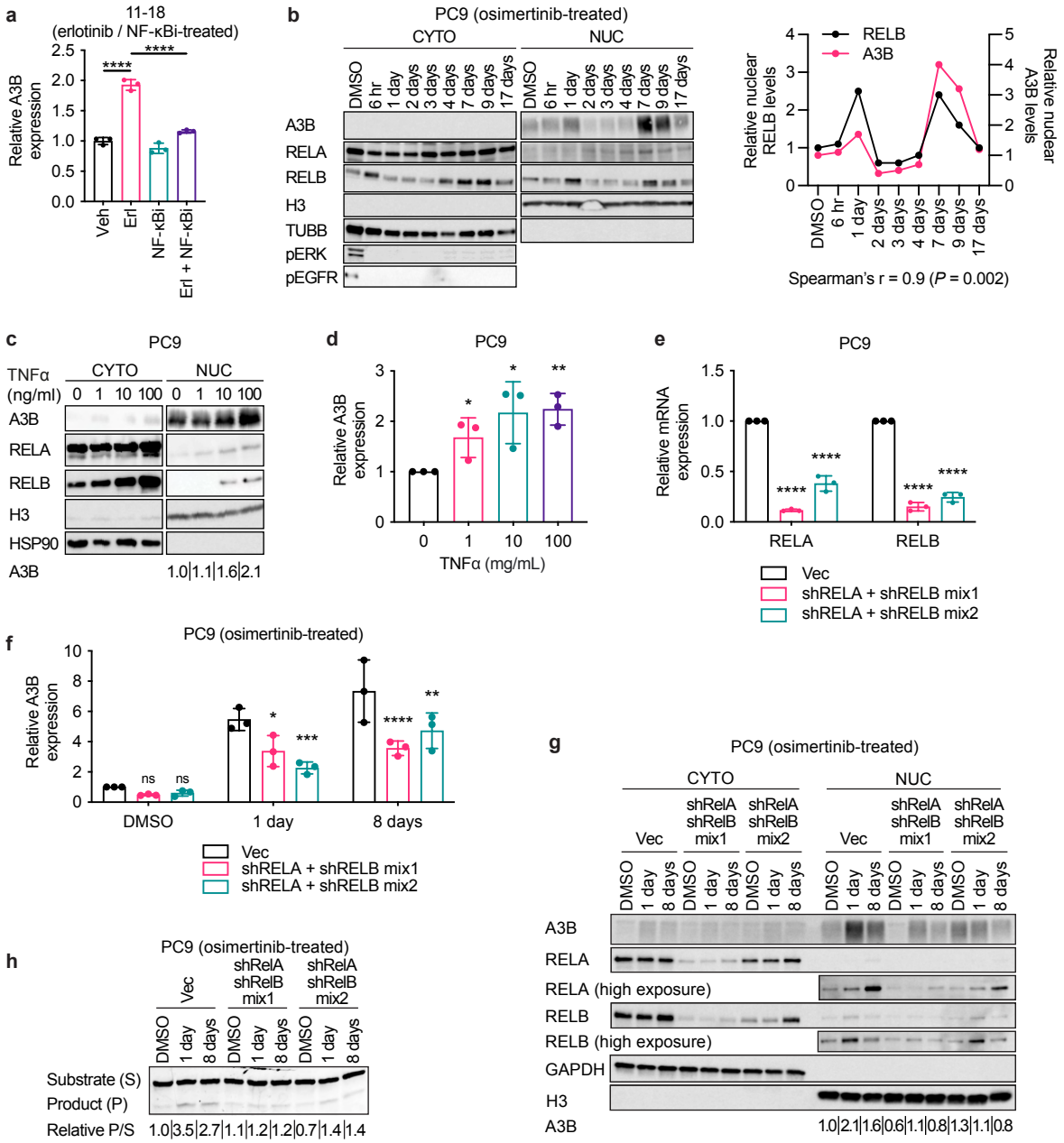
Extended Data Fig. 11| Cell cycle inhibition with palbociclib significantly decreases A3B expression. a, Cell cycle analysis of PC9 cells treated with DMSO, 2 μM osimertinib (osi) or 1 μM palbociclib (palbo) for 18 hours. Teal: G2/M phase; Pink: S phase; Black: G0/G1 phase (n = 4 biological replicates, mean ± SEM, ANOVA test, **P = 0.0010, **P = 0.0049, ***P = 0.0009, ****P < 0.0001, ***P = 0.0003). **b,** RT-qPCR analysis of PC9 cells treated as in panel a, (n = 2 or 3 biological replicates, mean ± SD, ANOVA test, ****P < 0.0001, *P = 0.0215, **P = 0.0018).



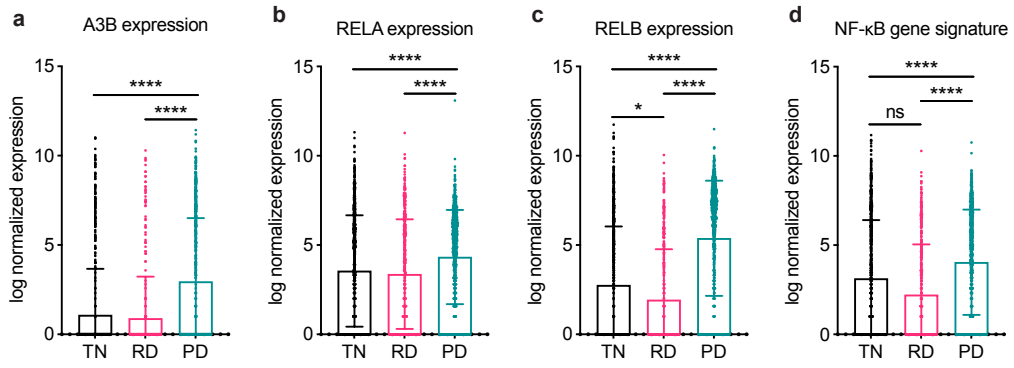
Extended Data Fig. 12| EGFR inhibition induces A3B upregulation and UNG2 downregulation in xenograft models. a, Western blot analyses of extracts of PC9 tumour xenografts treated with vehicle or 5 mg/kg osimertinib (n = 2 biological replicates). **b,** Representative images of IHC analysis of A3B protein levels in 11-18 xenografts treated with vehicle, 12.5 mg/kg/day erlotinib, 7.5 mg/kg/day NF-κB inhibitor (NF-κBi, PBS-1086) or combination (Erlotinib + NF-κBi) for 2 months (scale: 60 μM)[17]. **c**, Quantification of immunohistochemical staining for A3B in 11-18 xenografts treated with vehicle, erlotinib (Erl), NF-κB inhibitor (NF-κBi, PBS-1086) or combination (Erl + NF-κBi) for 2 months (n = 2 biological replicates, mean, ANOVA test, erlotinib vs. NF-κBi **P = 0.0095, erlotinib vs. erl + NF-κBi **P = 0.0072). **d**, Representative images of IHC analysis of UNG protein levels in 11-18 xenografts treated with vehicle or 12.5 mg/kg/day erlotinib for 2 months. **e**, Quantification of immunohistochemical staining for UNG in 11-18 xenografts treated with vehicle or erlotinib for 2 months (n = 2 biological replicates). **f,** RNA-Seq analysis upon treatment of a PDX model of human *EGFR*-driven lung adenocarcinoma with vehicle or erlotinib (2 days, 25 mg/kg) (n = 2 biological replicates, mean ± SD, t-test, *P = 0.0203). **g,** RNA-Seq analysis upon treatment of a PDX model of human *EGFR*-driven lung adenocarcinoma with vehicle or osimertinib (6 days, 10 mg/kg) (n = 3 biological replicates, mean ± SD, t-test, *P = 0.0267).



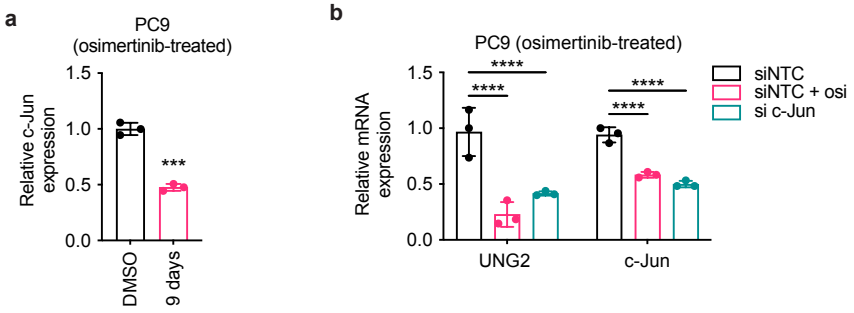
Extended Data Fig. 13| A3B drives a TKI-induced increase in nuclear APOBEC activity. **a,** APOBEC activity assay performed using nuclear extracts of PC9 cells treated with increasing doses of osimertinib (osi; day 1 to 3: 200 nM; day 3 to 6: 1 μM; day 6 to 14: 2 μM). **b**, Western blot analysis of pERK from extracts in panel a (C: cytoplasmic extracts; N: nuclear extracts). **c,** Comparison of nuclear APOBEC activity and A3B expression of cells in a and b (assayed in triplicate, Spearman's correlation). **d,** RT-qPCR validation of A3B siRNA knockdown (n = 3 technical replicates, t-test performed on ΔCt values shown, ****P < 0.0001). **e,** APOBEC activity assay performed using nuclear extracts of PC9 cells treated with non-targeting siRNA (siNTC) or A3B siRNA (siA3B), DMSO or 500 nM osi for 18 hours. **f,** Western blot analysis of pEGFR using treated control and A3B knockdown cells in d-e. **g,** RT-qPCR validation of A3B shRNA knockdown (n = 3 technical replicates, t-test performed on ΔCt values shown, ****P < 0.0001). **h,** APOBEC activity assay performed using nuclear extracts of PC9 cells transduced with pLKO vector or its derivatives encoding an shRNA against A3B and treated with DMSO or 1 μM erlotinib (erl) for 18 hours. **i,** Western blot analysis performed using nuclear extracts of PC9 cells transduced with pLKO vector or A3B knockdown vector alone or together with wild-type HA-tagged A3B or HA-tagged catalytic-ly-inactive A3B mutant (E255A) expression plasmid. **j,** APOBEC activity assay performed using PC9 cells as in panel i, in the absence of RNase A.



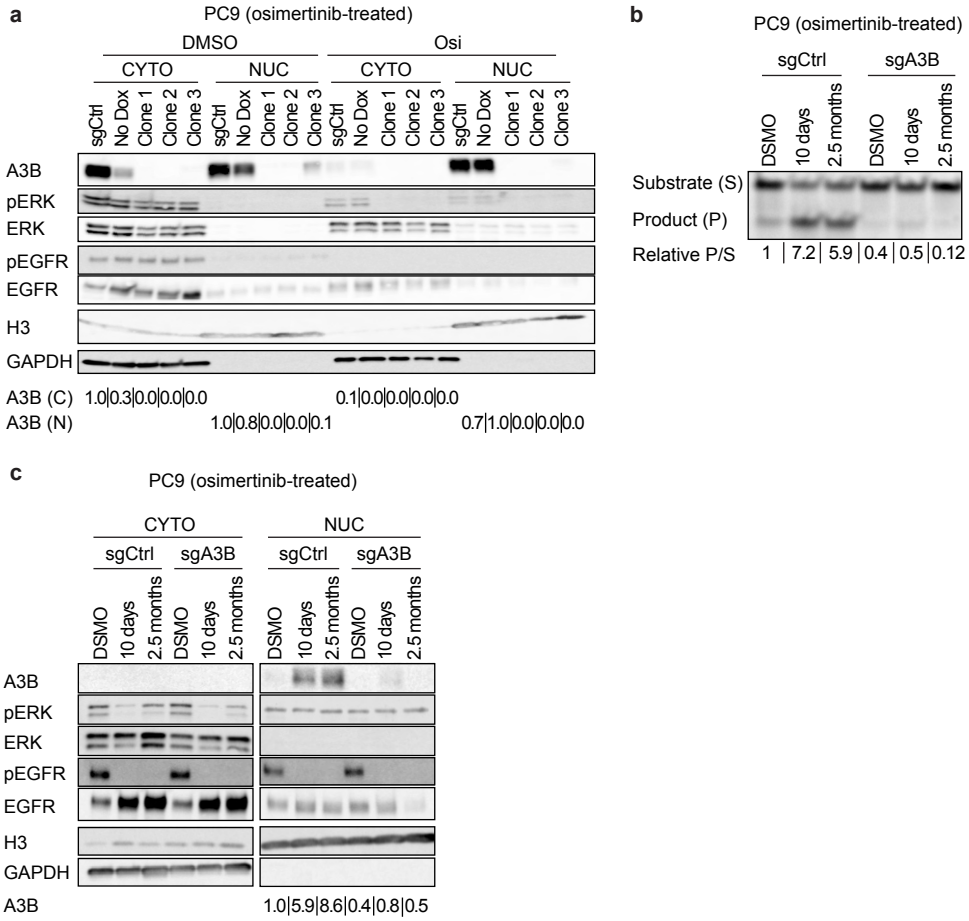
Extended Data Fig. 14| NF-κB signalling is critical for TKI-induced A3B upregulation. a, RNA-Seq analysis of *EGFR*-mutant 11-18 NSCLC cells treated with DMSO, 100 μM erlotinib (erl), 5 μM NF-κB inhibitor (NF-κBi, PBS-1086) or combination (Erl + NF-κBi) (n = 3 biological replicates, mean ± SEM, ANOVA test, ****P < 0.0001). **b,** Western blot analysis of extracts from PC9 cells treated with DMSO or 4 μM osimertinib (osi) for the indicated durations (CYTO: cytoplasmic extracts; NUC: nuclear extracts), with quantifications of relative nuclear A3B and RELB levels (Spearman's r = 0.9, P = 0.002). **c**, Western blot analysis of extracts from PC9 cells treated with DMSO or with TNFα for 8.5 hours. **d,** RT-qPCR analysis of TNFα-treated PC9 cells (n = 3 biological replicates, t-test, *P = 0.0406, *P = 0.0299, **P = 0.0024). **e,** RT-qPCR validation of RELA and RELB knockdown in PC9 cells transfected with non-targeting vector plasmid or combination of either shRELA-1 + shRELB-1 (mix1) or shRELA-2 + shRELB-2 (mix2) (n = 3 biological replicates; mean ± SD, ANOVA test, ****P < 0.0001). **f**, RT-qPCR analysis of A3B expression in PC9 cells transfected with non-targeting vector plasmid or combination of either shRELA-1 + shRELB-1 (mix1) or shRELA-2 + shRELB-2 (mix2), treated with DMSO or 500 nM osi for 1 or 8 days (n = 3 biological replicates; mean ± SD, ANOVA test, *P = 0.0130, ***P = 0.0005, ****P < 0.0001, **P = 0.0028). **g**, Western blot analysis of A3B, RELA, and RELB protein expression of PC9 cells used in f. **h**, APOBEC activity assay performed using nuclear extracts of PC9 cells used in f.



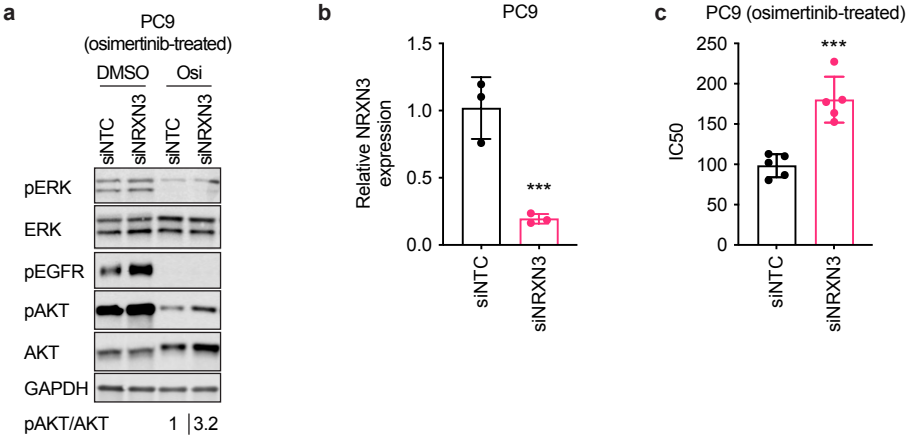
Extended Data Fig. 15| A3B expression and NF-κB pathway activation in patients with NSCLC. **a,** Box plots showing single cell RNA-Seq analysis of A3B expression in lung cancer cells from patient tumors at treatment naive (TN), residual disease (RD), and progressive disease (PD) timepoints (Wilcoxon Rank Sum test, ****P < 0.0001). **b**, Box plots showing single cell RNA-Seq analysis of RELA expression in lung cancer cells from panel a (Wilcoxon Rank Sum test, ****P < 0.0001). **c,** Box plots showing single cell RNA-Seq analysis of RELB expression in lung cancer cells from panel a (Wilcoxon Rank Sum test, ****P < 0.0001). **d,** Box plots showing single cell RNA-Seq analysis of NF-κB signature (from Gilmore_Core_NFκB_Pathway, GSEA, C2) in lung cancer cells from panel a (Wilcoxon Rank Sum test, ****P < 0.0001).



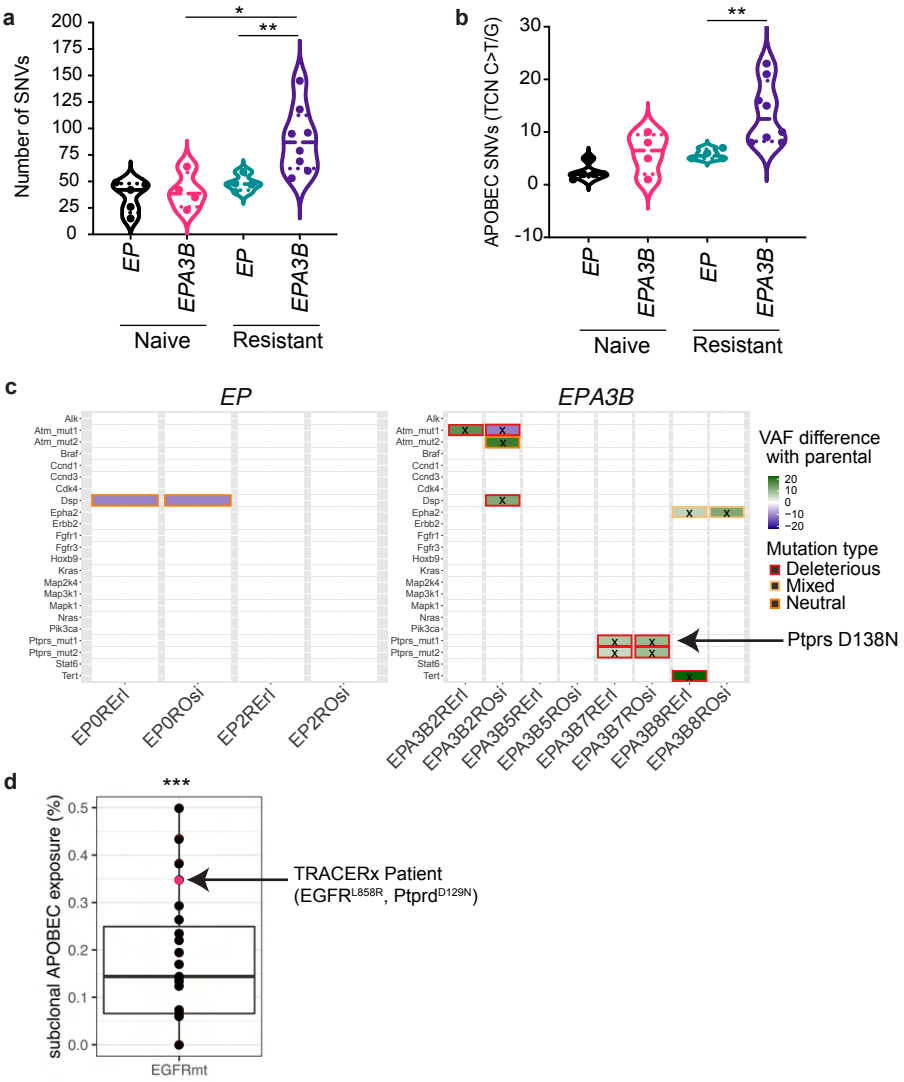
Extended Data Fig. 16| Mechanisms inducing adaptations favoring APOBEC-mutagenesis after TKI treatment. **a**, RT-qPCR analysis of c-Jun expression of PC9 cells treated with 2 μM osimertinib for 9 days relative to DMSO-treated cells using (3 biological replicates from Ext. Data Fig.11b, mean ± SEM, t-test, ***P = 0.0009). **b**, RT-qPCR analysis of PC9 cells transfected with non-targeting siRNA (siNTC) or c-Jun siRNA, and treated with DMSO or 2 μM osimertinib for 18 hours (n = 3 biological replicates; mean ± SD, ANOVA test, ****P < 0.0001).



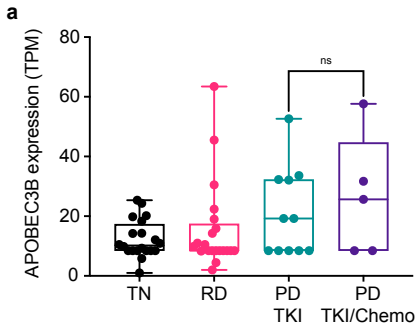
Extended Data Fig. 17| Confirmation of APOBEC3B deficiency in A3B knockout cancer cell clones. a, Western blot analysis for A3B, ERK, and EGFR of A3B proficient or A3B deficient PC9 cells treated with DMSO or osimertinib (Treatment: 113 days; dosing: 100 nM day 1-4, 500 nM-2 μM day 5 - 42, 50-100 nM day 43-76) (CYTO [C] - cytoplasmic extracts, NUC [N] - nuclear extracts). **b,** Nuclear APOBEC activity assay of A3B proficient control or A3B knockout PC9 clones treated with DMSO or osimertinib for the indicated duration. **c**, Western blot analysis for A3B, ERK and EGFR of clonal PC9 cells (Treatment: 2.5 months; dosing: 100 nM day 1-4, 1 μM - 2 μM day 5 - 52, 100 nM day 52 - 70, 25 nM day 70 -71) (CYTO - cytoplasmic extracts, NUC - nuclear extracts).



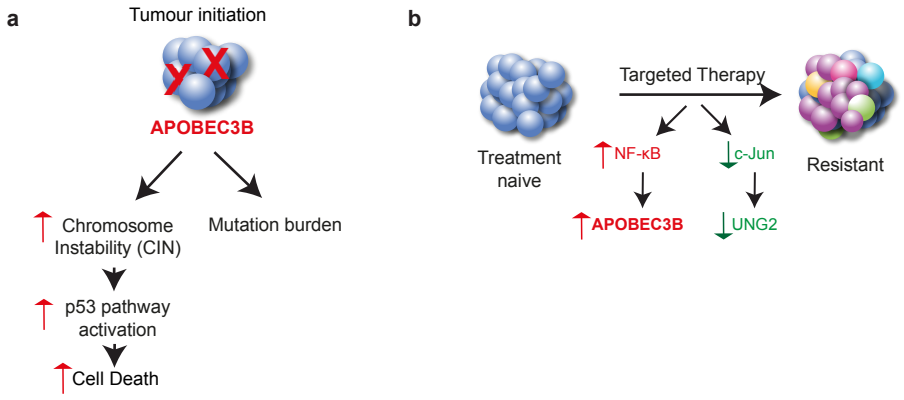
Extended Data Fig. 18| Loss of NRXN3 reduces sensitivity to TKI therapy. a, Western blot analysis of PC9 cells treated with non-targeting (siNTC) or NRXN3-targeting (siNRXN3) siRNA and treated with DMSO or 500 nM osimertinib for 2 days. **b,** RT-qPCR-based validation of NRXN3 knockdown in cells shown in a (n = 3 technical replicates, t-test performed on ΔCt values shown, ***P = 0.0007). **c,** IC50 analysis of PC9 siNTC or siNRXN3 after 3 day treatment (5 biological replicates for each of the following doses of osimertinib- 0 nM, 5 nM, 50 nM, 100 nM, 500 nM and 5000 nM, t-test, ***P = 0.0004).



Extended Data Fig. 19| An increase in the number of mutations in an APOBEC context in genes previously associated with TKI resistance in EPA3B mouse tumour cell lines. a, Comparison of EP and EPA3B mutation burdens in TKI naive and TKI resistant mouse lung cancer cell lines (median = dashed line, 1st and 3rd quartiles = dotted lines, two-sided Mann-Whitney Test, *P = 0.0162 **P = 0.008). **b,** Comparison of EP and EPA3B APOBEC driven mutations (TCN, C>T or C>G SNVs) in TKI naive and TKI resistant mouse lung cancer cell lines (median = dashed line, 1st and 3rd quartiles = dotted lines, two-sided Mann-Whitney Test, **P = 0.004). **c,** Functional annotation of TCN mutations in TKI resistance genes with change in variant allele frequency shown (x = TCN, Red square = deleterious mutation, yellow square = mixed (neutral and deleterious), orange square = neutral). **d**, Significant subclonal enrichment of the APOBEC-associated mutation signature in the TRACERx patient with A3B driven D129N mutation in the type IIa PTP PTPRD (equivalent to D138N mutation in PTPRS ***P = 0.0002, one-sample Wilcoxon test).



Extended Data Fig. 20| A3B expression in patients prior to and post therapy . a, Comparison of APOBEC3B expression levels (Exp: Batch corrected TPM) measured using RNA-Seq analysis in human NSCLC specimens obtained before treatment (TN), or on-treatment at residual disease (RD) or at progressive disease (PD) stages from lung cancer patients undergoing treatment with single agent tyrosine kinase inhibitors (TKI) or a combination of chemo- and TKI therapy (all data points shown).



Extended Data Fig. 21| The role of APOBEC3B in EGFR driven tumour evolution. a, At tumour initiation continous APOBEC3B expression induces, chromosomal instability and p53 pathway activation, resulting in cell death. **b**, With targeted therapy, APOBEC3B expression is induced in an NF-κB dependent manner, and UNG2 expression is lost driving genetic diversity and TKI resistance.