Pan-cancer survey of tumour mass dormancy and underlying mutational processes

Anna Julia Wiecek[1,†], Daniel Hadar Jacobson[1,†], Wojciech Lason[1,†], Maria Secrier[1*]

[1]UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, UK

[†]These authors have contributed equally to this work and share first authorship

*** Correspondence:**
Maria Secrier
m.secrier@ucl.ac.uk

Keywords: tumour mass dormancy, immunity, angiogenesis, mutational signatures, APOBEC, hypoxia

Abstract

Tumour mass dormancy is the key intermediate step between immune surveillance and cancer progression, yet due to its transitory nature it has been difficult to capture and characterise. Little is understood of its prevalence across cancer types and of the mutational background that may favour such a state. While this balance is finely tuned internally by the equilibrium between cell proliferation and cell death, the main external factors contributing to tumour mass dormancy are immunological and angiogenic. To understand the genomic and cellular context in which tumour mass dormancy may develop, we comprehensively profiled signals of immune and angiogenic dormancy in 9,631 cancers from the Cancer Genome Atlas and linked them to tumour mutagenesis. We find evidence for immunological and angiogenic dormancy-like signals in 16.5% of bulk sequenced tumours, with a frequency of up to 33% in certain tissues. Mutations in the *CASP8* and *HRAS* oncogenes were positively selected in dormant tumours, suggesting an evolutionary pressure for controlling cell growth/apoptosis signals. By surveying the mutational damage patterns left in the genome by known cancer risk factors, we found that ageing-induced mutations were relatively depleted in these tumours, while patterns of smoking and defective base excision repair were linked with increased tumour mass dormancy. Furthermore, we identified a link between APOBEC mutagenesis and dormancy, which comes in conjunction with immune exhaustion and may partly depend on the expression of the angiogenesis regulator *PLG* as well as interferon and chemokine signals. Tumour mass dormancy also appeared to be impaired in hypoxic conditions in the majority of cancers. The microenvironment of dormant cancers was enriched in cytotoxic and regulatory T cells, as expected, but also in macrophages and showed a reduction in inflammatory Th17 signals. Finally, tumour mass dormancy was linked with improved patient survival outcomes. Our analysis sheds light onto the complex interplay between dormancy, exhaustion, APOBEC activity and hypoxia, and sets directions for future mechanistic explorations.

1 INTRODUCTION

Tumour evolution is shaped by a variety of internal and external forces that act at different stages during cancer development, sometimes in an antagonistic manner, and drive distinct trajectories to advanced disease (McGranahan and Swanton 2017; Gerlinger et al. 2014; Temko et al. 2018). Within this rapidly changing context, adaptation of cancer cells is paramount for survival and much of it is achieved through cellular plasticity (Yuan, Norgard, and Stanger 2019). As a manifestation of this plasticity, cancer dormancy has emerged as an important contributor to the early stages of tumour development, as well as to cancer progression and metastatic dissemination (Phan and Croucher 2020; Jahanban-Esfahlan et al. 2019; Park and Nam 2020). Its two facets, cellular dormancy driven by arrest in the G0 state of the cell cycle (Phan and Croucher 2020), and tumour mass dormancy (TMD), described as an equilibrium between cell proliferation and cell death shaped by the microenvironment that constrains tumour growth (Aguirre-Ghiso 2007; H.F. Wang et al. 2019; Holmgren, O'Reilly, and Folkman 1995), are complementary but distinct mechanisms that contribute to the plasticity of cancer expansion (Huang 2021; Shen and Clairambault 2020). The former concept was coined already in 1954 by Geoffrey Hadfield (HADFIELD 1954) and has since led to the generation of more detailed mechanistic insights explaining its regulation by the DREAM complex (Sadasivam and DeCaprio 2013; MacDonald et al. 2017; Kim et al. 2021), with key dependencies on the p53/p21 activation axis (Itahana et al. 2002; Heldt et al. 2018; Barr et al. 2017). The latter, initially termed "population dormancy" by Judah Folkman in 1972 (Gimbrone et al. 1972) and then renamed to tumour mass dormancy, remains poorly understood due to the lack of suitable data and models (Boire et al. 2019).

The theoretical model of TMD is embedded into the "3Es of immunoediting" paradigm (Dunn, Old, and Schreiber 2004b), arising as a temporary equilibrium between tumour elimination and immune escape (Teng et al. 2008; Koebel et al. 2007). As the cancer lesion develops, a period of immunoediting follows when the immune system interacts with the malignant cells establishing a dynamic equilibrium: the immunogenic cells are eliminated, and non-immunogenic tumour cells arise (Dunn, Old, and Schreiber 2004a). This keeps the tumour in a dormant state (H.F. Wang et al. 2019; Aguirre-Ghiso 2007), hypothesised to account for the "latency" between the initiation of the first mutator phenotype and symptom manifestation in early disease, or for the disease-free period preceding cancer recurrence (Gužvić and Klein 2013; Damen, van Rheenen, and Scheele 2020). During this period of dormancy, the continued cytotoxic response triggered by pro-inflammatory signalling cytokines like interferon ɣ, as well as prolonged exposure to pro-inflammatory signalling in general, causes the cytotoxic cells to become inactive, a phenomenon known as T cell exhaustion (Dunn, Koebel, and Schreiber 2006; Wherry and Kurachi 2015; Yi, Cox, and Zajac 2010). Finally, the angiogenic switch or immune escape shifts the balance in favour of cancer progression (Jahanban-Esfahlan et al. 2019).

A variety of molecular mechanisms have been proposed to mediate these switches. The urokinase receptor (uPAR) regulates tumour growth by controlling β1 integrin signalling which drives a cascade of Ras/ERK mitogenic activation via the focal adhesion kinase (FAK) and the EGF receptor (EGFR) (Aguirre Ghiso, Kovalski, and Ossowski 1999; White et al. 2004; Liu et al. 2002; Aguirre Ghiso 2002). Downregulation of any of these components has been shown to lead to tumour growth arrest and dormancy (Aguirre-Ghiso 2007). Additionally, blocking uPAR activates p38, which induces dormancy under p53 upregulation and downregulation of JUN (Ranganathan et al. 2006). Metastatic lesions appear to manifest reduced p38 activity under sustained ERK activation. As a result, a low ERK:p38 protein expression ratio is often employed to assess tumour dormancy (Aguirre-Ghiso et al. 2003). This state is further corroborated by a limited ability in recruiting new blood vessels and vasculature remodelling, resulting in angiogenic dormancy (Moserle, Amadori, and Indraccolo 2009). This condition is often characterised by VEGF inhibition in the presence of anti-angiogenic factors such as

angiostatin, endostatin, thrombospondin, etc., or chemokines (CXCL9, CXCL10) (Senft and Ronai 2016; Lyu et al. 2013; Aguirre-Ghiso 2007; Park and Nam 2020).

Overall, it is clear that TMD results from an interplay between immunological and angiogenic dormancy, where cell growth is counterbalanced by apoptosis due to poor vascularisation and immune cytotoxicity, followed by exhaustion. It becomes evident that the states of T cell exhaustion and dormancy are not a simple dysfunction, but a purposeful homeostatic mechanism enabling the prevention of pathological immune responses. Hence, it is of crucial importance to be able to identify and target the dormant tumours early. This temporary equilibrium state provides a unique clinical opportunity, but its prevalence in cancer and the genetic determinants that may sustain it are currently unknown.

In this study, we surveyed the landscape of TMD along the two axes that shape it, immunological and angiogenic dormancy, in a cohort of 9,631 tumours from 31 tissues available from TCGA. We show that TMD can be captured from bulk sequencing datasets, that it is pervasive across a variety of cancer types and that its emergence is linked with key somatic alterations. We also investigate the environmental context of TMD and its relevance in the clinic.

2 RESULTS

2.1 Pan-cancer characterisation of immunological, angiogenic and tumour mass dormancy

To evaluate the levels of immunological and angiogenic dormancy across multiple cancer tissues, we manually curated lists of genes associated with the two programmes from the literature (Supplementary Table 1). The lists included receptor molecules, as well as soluble mediators, surface and structural proteins, enzymes and transcription factors, several of which are observed to be frequently mutated during cancer development (Supplementary Figure 1). Immunological and angiogenic dormancy programme scores were assigned on a per-sample basis using two distinct methodologies that exploited the expression of genes associated with either process, focusing either on differences between up/downregulated gene activity ('scaled difference of means') or on the largest variation explained by principal component analysis ('PCA') (see Materials and Methods, Supplementary Figures 2-4). Both immunological and angiogenic dormancy represent potential mechanisms for restricting the expansion of primary tumour cell populations due to either impaired vascularization or immunosurveillance (Aguirre-Ghiso 2007), therefore both processes can contribute to the development of TMD. As such, an overall per-patient TMD programme score was also derived using the expression of genes associated with both types of dormancy. We assessed the robustness of the two scoring methods to changes in the gene signature employed or small variations in gene expression, and showed that the 'scaled difference of means' approach was comparatively more stable to such fluctuations (Supplementary Figure 5). Therefore, we chose this method for the downstream analysis.

Having established a framework for quantifying immunological, angiogenic and tumour mass dormancy programmes, we next profiled these cellular programmes across 9,631 samples of solid primary tumours from the Cancer Genome Atlas (TCGA). We observed a spectrum of TMD across tumours that ranges from highly dormant to highly expanding (Figure 1A-C, Supplementary Figures 6-7). TMD is thought to emerge when tumour cell proliferation is balanced by apoptosis due to factors such as limitations in blood supply or an active immune system (Holmgren, O'Reilly, and Folkman 1995). Consequently, primary tumour samples with high TMD programme scores (upper quartile of the score range) which also showed a proliferation/apoptosis ratio below 1 (see Materials and Methods), indicative of limited primary tumour lesion expansion, were classed as exhibiting TMD (Figure 1D). Overall, 16.5% of samples across different tissues exhibited substantial evidence for a

127 TMD-like phenotype (with up to 33% prevalence in certain cancers) and they were further subdivided
128 into those indicating angiogenic dormancy (4.4%), immunological dormancy (5%), or both (7%)
129 (Figure 1E and Supplementary Table 2). Head and neck (32%), sarcoma (32%) and breast (26%) cancer
130 were among the cancers with highest rates of TMD. Angiogenic dormancy alone was most widespread
131 in breast (14%) and head and neck cancers (13%). Some rarer cancers like thymoma or
132 pheochromocytoma and paraganglioma also showed remarkably high prevalence of TMD (~33% each)
133 with the latter also exhibiting the highest immune-mediated dormancy (21%). The systematic
134 differences in TMD programme scores across different tissues (Figure 1E-F) suggest that the tissue
135 environment may impact the ability of the tumour to enter a TMD state. Across the board, immune-
136 mediated dormancy levels appeared higher than those of angiogenic dormancy, suggesting a dominant
137 role for immune surveillance in determining TMD.

138 Even though TMD is expected to manifest primarily in early forming tumours (Dunn, Old, and
139 Schreiber 2004b), we surprisingly found that a sizeable proportion of late-stage tumours also exhibited
140 similarly high TMD-linked activity levels (Supplementary Figure 8). When comparing the prevalence
141 of TMD between early- and late-stage cancers, expanding tumours appeared enriched in the later stages
142 (Fisher's exact test $p < 0.0001$, 1.5-fold enrichment), as expected. Remarkably, samples with
143 angiogenic dormancy were also marginally enriched (1.3-fold) in late-stage cancers (Fisher's exact test
144 $p < 0.05$).

145 **2.2 The genomic background of TMD**

146 To gain a better understanding of the genomic context in which TMD can develop, we asked whether
147 samples with high and low TMD differed in their association with known drivers of tumorigenesis. We
148 found 15 genes with either a statistically significant enrichment or a depletion of mutations within
149 samples classed as showing TMD across the 31 solid cancer tissues, with several depletion signals
150 characterizing both early- and late-stage tumours (Fisher's exact test adjusted $p<0.05$, Figure 2A,
151 Supplementary Figures 9A, 10A, Table 3). Furthermore, mutations in *MUC4*, a gene involved in
152 angiogenesis and metastasis (Zhi et al. 2014), were specifically enriched in stomach adenocarcinoma,
153 while *EGFR*, *FAT3/4*, *LRP1B*, *KAT6B* mutations, mainly linked with cell proliferation and immune
154 responses (Katoh 2012; H. Chen et al. 2019; Simó-Riudalbas et al. 2015; Sigismund, Avanzato, and
155 Lanzetti 2018), were depleted in colon cancer (Fisher's exact test adjusted $p<0.05$, Supplementary
156 Figure 11). Many of these findings were reaffirmed using a random forest classification approach
157 (Supplementary Figure 12).

158 The results of the pan-cancer analysis largely reflected the balance of proliferation, cell death and
159 pro/anti-angiogenic signals (Naumov et al. 2006; Semenza 2003; Aguirre-Ghiso 2007) one would
160 expect in the context of dormancy. Two genes stood out as having a ~5-fold enrichment of mutations
161 in patients displaying TMD: *HRAS* and *CASP8*. Interestingly, while *HRAS* alterations were positively
162 associated with dormancy, other members of the same oncogenic family, *KRAS* and *NRAS* showed a
163 depletion of mutations in the context of this phenotype. Specifically, we found an enrichment of *HRAS*
164 hotspot mutations at the Q61 and G13 positions across all solid primary tumour samples, but a
165 depletion of *KRAS* G13 and G12, as well as *NRAS* Q61 hotspot alterations (Supplementary Figure 13).
166 Two of these hotspots presented cancer stage specificity: *HRAS* G13 mutations were enriched in TMD
167 in late-stage cancers, while *KRAS* G12 mutations were depleted in early-stage tumours (Supplementary
168 Figures 9B, 10B). The oncogenic activation of the Ras protein has been associated with pro-angiogenic
169 signaling through repression of thrombospondin-1 (Watnick et al. 2015). It has been suggested that
170 while distinct oncogenic Ras alterations might have similar ability to promote cell cycle progression,

171 they might have different abilities to induce the pro-angiogenic programme (Aguirre-Ghiso 2007),
172 which may explain the discrepancy of mutation signals in the *RAS* genes in relation to TMD.

173 *CASP8*, a gene encoding a cysteine-aspartic acid protease involved in the execution of apoptosis by
174 cleaving and thereby activating caspase-3 and caspase-7 (Tummers and Green 2017), was also
175 preferentially mutated in dormancy. While *CASP8* loss of function would be predicted to impair the
176 ability of cancer cells to initiate apoptosis, silencing of *CASP8* in breast cancer cell lines has also been
177 shown to decrease cancer cell growth by delaying G0/G1- to S-phase transition and increasing the
178 expression of CDK inhibitors p21 and p27 (De Blasio et al. 2016).

179 The genes presenting a depletion of point mutations within dormant samples identified by our analysis
180 have key functions in regulating tumour growth, including *TP53*, the master regulator that coordinates
181 signals of stress such as DNA damage and aberrant growth signaling and can induce cell cycle arrest
182 or apoptosis (Vogelstein, Lane, and Levine 2000; Polyak et al. 1997), or *APC*, which suppresses tumour
183 growth through repression of the Wnt signaling pathway (Boman and Fields 2013). Both the *APC* and
184 *DCC* genes promote apoptosis by downregulation of survivin gene expression (T. Zhang et al. 2001)
185 and caspase-9 cleavage (Forcet et al. 2001) respectively. Control of tumour vascularization was
186 reflected in depletion of mutations in *PRDM16*, which inhibits angiogenesis by suppressing the
187 expression of a HIF target semaphorin 5B (Kundu et al. 2020), and in *NCO1A*, a transcriptional
188 coactivator that upregulates the expression of the VEGFα pro-angiogenic factor (Qin et al. 2015).
189 Mutations in the *IDH1* gene, shown to be depleted in samples with TMD, result in the production of
190 the 2-hydroxyglutarate metabolite which regulates the activity of α-ketoglutarate dependent
191 dioxygenases and causes the ubiquitination and proteasomal degradation of *HIF1A*, a key sensor of
192 hypoxia and initiator of angiogenesis (Ye et al. 2013). Moreover, 2-hydroxyglutarate can also result in
193 the downregulation of leukocyte chemotaxis factors (Turcan et al. 2012; Tommasini-Ghelfi et al.
194 2019), which could contribute to the ability of tumour cells to escape immunological dormancy through
195 immune system evasion.

196 In addition to the enrichment analysis, using a maximum-likelihood dN/dS method (Martincorena et
197 al. 2017) we detected signals of positive selection for mutations within *CASP8* and *HRAS* in samples
198 with TMD, but not in expanding tumour samples (Figure 2B). In contrast, *IDH1* and *NRAS*, both of
199 which showed a depletion of mutations within TMD samples, showed signals of positive selection only
200 in expanding tumours (Figure 2B). Interestingly, *CASP8* showed similarly increased mutation rate both
201 in early- and late-stage cancers with TMD, while *HRAS* mutations appeared linked with TMD only in
202 late-stage tumours – pointing towards a timing specificity of TMD-linked evolutionary pressures
203 (Supplementary Figure 14). The associations between mutations in these genes and TMD were also
204 validated in an independent dataset from the International Cancer Genome Consortium (ICGC), along
205 with several others including *KRAS* and *TP53* (Supplementary Figure 15). While signals of positive
206 selection are more difficult to validate in the sparser independent cohorts available due to the limited
207 sample size, nevertheless *CASP8* and *TP53* selection signals were robustly recovered both in oral
208 cancers from ICGC as well as in breast cancers from the METABRIC cohort (Supplementary Figure
209 16). These findings support the importance of *CASP8* and *RAS* mutational status in the context of TMD.

210 In terms of broader structural variation in the genome, we found no copy number alteration events
211 (amplifications and deletions) specifically enriched in tumours with TMD (data not shown), potentially
212 because such events would be preferentially selected for in fast growing tumours. Moreover, tumours

213 with angiogenic and immunological dormancy showed a modest, but significant decrease in mutational
214 burden when compared to tumours without TMD (Supplementary Figure 17).

215 2.3 Mutational processes linked with dormancy and exhaustion

216 In addition to investigating the links with specific driver events, we also set out to characterise broader
217 mutational processes associated with TMD. Different risk factors of cancer induce DNA damage in the
218 cells in a context dependent manner, such that nucleotide substitution patterns associated with such
219 mutational processes can be observed within cancer genomes. These patterns of trinucleotide
220 substitutions are termed "mutational signatures" and have been widely characterised across cancers
221 (Alexandrov, Nik-Zainal, Wedge, Aparicio, et al. 2013; Alexandrov et al. 2020). We carried out
222 mutational signature analysis to survey the contribution of known mutagenic processes and risk factors
223 to the genomes of dormant tumours. The mutational signature prevalence was correlated with
224 dormancy and exhaustion programme scores (Figure 2C, left and centre panels).

225 We observed that exposure to signatures SBS1, originating from aging-induced deamination of 5-
226 methylcytosines (Alexandrov et al. 2015), SBS5, also ageing-linked, and SBS22, linked with
227 aristolochic acid exposure (Jelaković et al. 2015), decreased as TMD and exhaustion increased. In
228 contrast, smoking (SBS4) and defective base excision repair linked with polymerase epsilon or *NTHL1*
229 mutations (SBS10a/b, SBS30) (Jager et al. 2019; Shivji et al. 1995) were associated with an increase
230 in TMD. Finally, we noted a consistently strong correlation between TMD/exhaustion and mutational
231 signatures SBS2 and SBS13, associated with mutagenesis induced by a class of cytidine deaminases
232 called APOBEC (apolipoprotein B mRNA editing catalytic polypeptide-like) (Alexandrov et al. 2020).
233 These enzymes induce mutagenesis in viral genetic material and are part of anti-viral defence, but can
234 also act on and damage host DNA (Green and Weitzman 2019). The correlation of these mutational
235 footprints of APOBEC with TMD was observed both pan-cancer and across individual cancer types,
236 including bladder, breast, cervical and head and neck cancers (Figure 2C). Reassuringly, there was a
237 similarly strong correlation between SBS2 and SBS13 and the mean expression of the AID/APOBEC
238 cytidine deaminases (Figure 2C, right panel). All the correlations highlighted were consistent between
239 early- and late-stage cancers (Supplementary Figure 18).

240 2.4 Dormancy and exhaustion programmes correlate with APOBEC enzymatic activity

241 To further explore the association between TMD and exhaustion with APOBEC mutagenesis, we
242 calculated the correlation of our programme scores and the mean expression of genes belonging to the
243 APOBEC family (Figure 3A-F). Across all cancers and regardless of clinical stage, we observed
244 significant correlations between the activity of the APOBEC programme as a whole, TMD and
245 exhaustion programmes (Figure 3 A-G, Supplementary Figures 19-23). These associations were
246 stronger than would be expected by chance (Supplementary Figure 24) and consistently validated in
247 independent datasets from ICGC and cBioPortal (Supplementary Figures 25-28). In particular, we
248 observed a marked association between APOBEC mutagenesis and immunological dormancy, which
249 we stipulate may precede or come in conjunction with immune exhaustion, the latter showing the
250 strongest correlation.

251 We also observed positive correlations between most but not all genes included in the immunological,
252 angiogenic and exhaustion programmes, on the one side, and the expression of the AID/APOBEC
253 enzyme family members, on the other (Figure 3H). This may be because of downstream regulation of
254 expression, e.g. through mRNA degradation, compartmentalization, or inhibition of transcription.
255 Notably, among all genes, the correlation was highest for the mRNAs of enzymatically active
256 APOBEC3 and AID members of the APOBEC family. The correlation scores were similar for

257 *APOBEC3A-H*, which is to be expected as these genes form a cluster on chromosome 22. Within the
258 TMD programme, type II interferon ɣ gene and the interferon-upregulated genes such as *STAT1*
259 showed a positive correlation with APOBEC. This is interesting as interferon ɣ signaling is associated
260 with anti-tumour response by favoring rejection of highly immunogenic tumours (Mittal et al. 2014;
261 Benci et al. 2016). However, prolonged interferon ɣ signaling promotes epigenetic changes to *STAT1*
262 and promotes expression of ligands for multiple T cell inhibitory receptors (Benci et al. 2016). Thus,
263 we would expect a spectrum of exhaustion signaling to be present within the tumour, ranging from
264 cells signaling danger to cells with anti-cytotoxic response. Anti-angiogenic markers, such as the
265 urokinase receptor *PLAUR*, which downregulate angiogenic processes and thus restrict the supply of
266 T cells (Oh, Hoover-Plow, and Plow 2003), were also positively correlated with APOBEC expression.
267 Similar associations were observed for several immune checkpoints (*CD244*, *CD166*, *TIGIT*, *CTLA4*,
268 *PDCD1*) and exhaustion markers (*EOMES*, *TBX21*).

269 Overall, the multiple correlations between APOBEC expression and the dormancy/exhaustion levels
270 in tumours suggest a complex interplay between the activity of APOBEC deaminases and the immune
271 microenvironment, which we set out to explore in greater depth.

272 **2.5 TMD activity differences in the context of APOBEC mutagenesis**

273 To further elucidate the association between TMD and APOBEC activity, we sought to investigate the
274 specific TMD signals that might have the strongest links with APOBEC-attributable mutagenesis, as
275 quantified by the single-base substitution mutational signatures SBS2 and SBS13. In order to identify
276 APOBEC-enriched tumours, we employed t-distributed stochastic neighbour embedding (tSNE) to
277 cluster samples based on their overall mutational signature profiles (Figure 4A-B). Unsurprisingly, the
278 clustering was impacted by the tumour tissue of origin, as mutational signatures are often tissue-
279 specific (Degasperi et al. 2020) (Figure 4A). However, samples enriched for APOBEC-associated
280 mutations, defined as the total enrichment of signatures SBS2 and SBS13, covered a fairly
281 heterogenous set of tissues (Figure 4B). Groups with distinct mutagenesis patterns were defined using
282 expectation-maximization clustering (Supplementary Figure 29), and the procedure was repeated 100
283 times to obtain robust clusters. Samples were defined as APOBEC-enriched if they fell into the
284 APOBEC-associated cluster more than 50 times (see Materials and Methods), which broadly overlaps
285 with the APOBEC enrichment score approach developed by Roberts et al (Roberts et al. 2013)
286 (Supplementary Figure 30).

287 To determine whether APOBEC mutagenesis was specifically associated with particular aspects of the
288 TMD programme, we used random forest classifiers to rank TMD genes based on how much their
289 expression helps to distinguish between the APOBEC-enriched and depleted groups (see Materials and
290 Methods). The accuracies for predicting APOBEC enrichment across 100 built models were
291 remarkably high and broadly replicable (mean AUC = 0.8354, range: 0.8203-0.8534) (Figure 4C-D).
292 Additionally, the gene ranking in the models was also highly reproducible (Figure 4E). *PLG*, which
293 encodes for the plasminogen protein and triggers angiostatin release to inhibit angiogenesis (Oh,
294 Hoover-Plow, and Plow 2003), displayed the highest importance by far, defined as the mean associated
295 decrease in Gini index following its removal from a model (mean = 119.2). Upon further investigation
296 of *PLG* expression, we found that its high predictive importance could largely be attributed to a
297 substantial group of low-*PLG*-expressing samples which were not labelled as APOBEC-enriched
298 (Figure 4F). This becomes more apparent when considering the multidimensional scaling plot
299 attributable to one of the random forest classifiers, which clearly displays a segregated cluster of non-
300 APOBEC-enriched samples defined by a specific *PLG* expression threshold derived from the classifier
301 (Figure 4G). Therefore, APOBEC mutagenesis appears to be completely lacking in the cases when

302 *PLG* is not expressed, suggesting a link between APOBEC enzymatic activity and the uPA/uPAR
303 system of angiogenic modulation.

304 Finally, we compared the expression of genes associated with TMD between APOBEC mutagenesis
305 enriched and depleted groups. The large majority of genes whose upregulation drives TMD displayed
306 a higher median expression in APOBEC-enriched samples compared with non-enriched samples
307 (Figure 4H). It is worth noting that some genes were found to be significantly changed whilst
308 displaying negligible difference in medians, likely reflecting the effects of a large sample size rather
309 than a real biological difference. Overall, these results reaffirm the positive correlation between TMD
310 and APOBEC enzyme activity, apparent both from an expression and mutagenesis perspective.

311 **2.6 Heterogeneity of TMD maintenance in hypoxic conditions**

312 Owing to the reported association between hypoxic environments and angiogenesis (L. Chen, Endler,
313 and Shibasaki 2009), we next sought to investigate whether hypoxia might differentially impact TMD.
314 Hypoxia was quantified on a per-sample basis using established transcriptional signatures (see
315 Materials and Methods), and the calculation procedure was validated using three separate hypoxia
316 signatures (Supplementary Figure 31).

317 From a pan-cancer perspective, we observed that hypoxia scores were significantly lower in samples
318 displaying any type of dormancy, compared to samples with no evidence of TMD (Figure 5A). This
319 was further corroborated by grouping tumours into distinct subsets (bins) based on their hypoxia levels:
320 the proportion of samples with no evidence of TMD increased significantly as hypoxia scores increased
321 (Figure 5B). In particular, we noted the lowest hypoxia levels appearing in tumours with angiogenic
322 dormancy. These results align with the established consensus that the hypoxia-inducible factor (HIF)
323 pathway is a key regulator of angiogenesis (Pugh and Ratcliffe 2003; Krock, Skuli, and Simon 2011),
324 with angiogenic dormancy expected to develop in normoxic conditions. Similar patterns were also
325 observed on a tissue-specific basis, particularly in the cases of sarcoma, stomach, lung squamous and
326 esophageal carcinomas (Figure 5C). Cancer studies which tended to present low hypoxia, such as
327 pheochromocytoma and paraganglioma (PCPG), thymoma (THYM) and thyroid carcinoma (THCA)
328 often presented some evidence of TMD, whereas studies presenting high hypoxia, such as colon and
329 rectum adenocarcinomas (COAD/READ) and glioblastoma (GBM), were abundant in expanding
330 tumours (Supplementary Figure 32). Notable exceptions included bladder cancer and uveal melanoma,
331 both of which presented samples with evidence of TMD at the higher end of their respective hypoxia
332 spectra (Figure 5C-F). While prior research has highlighted HIF-1α expression as a promoter of
333 angiogenesis and tumour invasion in bladder carcinoma (Theodoropoulos et al. 2004) and metastasis
334 in uveal melanoma (Asnaghi et al. 2014), our analysis suggests that a period of TMD-induced latency
335 may also be compatible with hypoxia in these cancers.

336 While a considerable heterogeneity across cancer tissues was observed, patterns of decreased hypoxia
337 in the context of TMD were noted in the majority of cancers (Figure 5D-F). This would imply that
338 hypoxia may be an impediment to TMD maintenance across the majority of tumours, consistent with
339 its increased prevalence in the later stages of tumour development (Petrova et al. 2018).

340 **2.7 The microenvironmental context of TMD**

341 Since TMD is by definition dependent on immune surveillance, we also set out to confirm this and
342 potentially identify new components of the cancer microenvironment which are permissive to this type
343 of dormancy. Specifically, because the balance in the populations of immune and stromal cells is
344 central to forming anti-tumour responses, we investigated how immune composition might be shaped

345 in the context of TMD, exhaustion and APOBEC programmes. Using cell type-specific transcriptional
346 markers (see Materials and Methods), we observed strong positive correlations between TMD and
347 various types of T cells (Figure 6A). Antitumoral cytotoxic cells such as CD8+ and CD4+ T cells,
348 Th1/2 cells, as well as tumour promoting regulatory cells and macrophages were correlated with the
349 APOBEC, exhaustion and dormancy programmes. This observation is in line with the mechanism
350 proposed in the literature whereby an initial immune response caused by neoantigens presented by the
351 dormant tumours is followed by exhaustion of such signals (Ghorani et al. 2020). In addition to their
352 cytotoxic activity, CD8+ and CD4+ T cells have been previously shown to limit tumour growth through
353 the secretion of anti-proliferative cytokines, such as IFN-γ which can stimulate the expression of p21
354 and p27 cell cycle inhibitors (Wall et al. 2003; Chin et al. 1996), and antiangiogenic chemokines, such
355 as CXCL9 and CXCL10 (Ikeda, Old, and Schreiber 2002).

356 Compared to Th1 and Th2 cells, natural killer (NK) cells showed a weaker correlation with dormancy
357 programmes, consistent with reports of tumour growth control by the immune system being mostly
358 associated with adaptive T cell responses (Finn 2006). The microenvironment of samples with high
359 TMD was also depleted in inflammatory Th17 cells, which secrete the angiogenesis inducer IL-17A
360 (Bailey et al. 2014). Interestingly, samples displaying signals of angiogenic dormancy had higher
361 enrichment of Th2 instead of Th1 cells. While the Th1 activity is linked with IFN-ɣ production and
362 tumour suppression, Th2 activity has been implicated in cancer progression (Zhao et al. 2019). Despite
363 the broader heterogeneity of T helper cell signals, they showed a similar overall trend as that of
364 cytotoxic and dendritic cells when projected across the TMD landscape (Figure 6B-D). All these
365 classes except for the Th2 cells appeared most enriched in tumours with evidence of both
366 immunological and angiogenic dormancy (Figure 6E). Similar associations were observed for early-
367 and late-stage cancers (Supplementary Figures 33-36).

368 **2.8 TMD is prognostic in cancer**

369 To understand the clinical relevance of TMD, we carried out survival analysis and found that patients
370 with expanding tumours that presented no evidence of TMD had a significantly reduced prognosis
371 (Figure 7A), consistent with the expectation that fast-proliferating tumours should be more aggressive
372 than their dormant counterparts. To further dissect the potential sources of this variation along the
373 immunity/angiogenesis axis, we employed Cox proportional hazards models to determine the effects
374 of different types of dormancy on survival whilst accounting for potential confounders such as patient
375 age, gender, cancer type, and tumour stage. While all three categories were associated with decreased
376 risk when compared to patients with no evidence of TMD, much of this variation was accounted for
377 by tumour stage, with only angiogenic and tumour mass dormancy still marginally significant after
378 correction (Figure 7B, Supplementary Figures 37-41). While the tumour stage effect may not be fully
379 uncoupled from the TMD effect, TMD did present a marginal protective role within the early-stage
380 tumours (stages I/II), as well as a larger protective role in late-stage tumours (stages III/IV)
381 (Supplementary Figure 42).

382 **3 DISCUSSION**

383 TMD is a key state in tumour development and progression, but its evolutionary constraints are largely
384 unknown. In this study, we have presented an overview of TMD across 31 cancer tissues and
385 highlighted potential novel genomic hallmarks of this state. While TMD remains a poorly characterised
386 programme whose current known markers are undoubtedly incomplete, our approach to evaluating
387 TMD in bulk tumours through an equilibrium of proliferation, apoptosis, cytotoxicity, exhaustion and
388 angiogenesis signals has enabled us to systematically capture this temporary state across a variety of

389 cancers. We show that TMD-like signals are present across a multitude of cancer tissues, and that
390 immunological and angiogenic dormancy are not always concomitant. We confirm many of the
391 expected associations reported in the literature, including lower mutational burden, correlations with
392 CD8+, regulatory and helper T cells (Teng et al. 2008), improved prognosis (Park and Nam 2020) and
393 a strong link with immune exhaustion (Mittal et al. 2014). Exhaustion is a phenotype that likely
394 immediately follows TMD, but in our analysis the two states are largely overlapping. This is most
395 likely due to the bulk nature of the samples analysed, where we may be capturing parts of the tumour
396 that starting to expand in addition to areas still displaying TMD. Interestingly, we find TMD tends to
397 occur in normoxic rather than hypoxic environments in the majority of cases, which is different than
398 what is observed for cellular quiescence (Qiu et al. 2017; Butturini et al. 2019).

399 In addition to expected associations, our methodology has also enabled us to capture previously
400 unreported links between TMD and clinical or genomic characteristics. Firstly, while TMD would be
401 expected to dominate in early-stage tumours, we also notice TMD-like signals in a comparable fraction
402 of late-stage tumours (18%) (Supplementary Figure 8). In the absence of longitudinal data that could
403 shed further light onto the causes, we can only hypothesise this could be due to a mechanism of slowly
404 advancing cancers that are maintained in a semi-TMD state, potentially enabled by a strong, consistent
405 immune response.

406 Secondly, we find that oncogenic mutations in *CASP8* (inactivating) and *HRAS* (activating) are
407 positively selected for during the development of TMD. This is either because they might confer an
408 advantage in terms of maintaining TMD, or because they might enable cells to escape TMD and expand
409 further. Interestingly, the selection pressure on *CASP8* in the context of TMD was seen across the entire
410 tumour progression spectrum, while *HRAS* mutations appeared positively selected specifically in late-
411 stage tumours. Counterbalancing these selective forces are *KRAS* hotspots that are preferentially
412 depleted in the context of TMD. These findings could suggest distinct adaptive mechanisms in the
413 context of early forming versus disseminated cancer in a temporary dormant state, as well as a complex
414 "tug-of-war"-like regulation of tumour growth/arrest via Ras oncogenesis. Beyond these mutational
415 associations, most of the classical cancer drivers were amplified or deleted preferentially in non-
416 dormant tumours, while TMD appeared largely devoid of specific copy number marks. These
417 observations may be partly explained by an association with early stage cancers, where these changes
418 have not yet occurred as a means of expansion and escape from immune surveillance.

419 In concordance with the reduced rate of cell proliferation in TMD, we found that ageing-induced
420 mutations, which are believed to accumulate at a relatively constant rate in cancer, were lesser
421 represented in dormant tumours. Correlations between TMD and several other mutational processes
422 (smoking, defective DNA repair) were noted, which might suggest that multiple neoplastic
423 mechanisms that are tissue-specific could favour this state.

424 A remarkably consistent signal came from the mutagenesis trace left by the APOBEC enzyme family.
425 APOBEC activity appeared strongly linked with the exhaustion signals that presumably follow TMD,
426 and with immunological dormancy. Several studies (Yamazaki et al. 2019; Swanton et al. 2015; Burns,
427 Temiz, and Harris 2013) have proposed that APOBEC expression may be one of the drivers of
428 tumorigenesis, suggesting a role for APOBECs in evading immune response. APOBEC3 signatures
429 are prevalent in leukaemias (Rebhandl et al. 2014), while AID was demonstrated to induce
430 hypermutation and chromosomal translocation in B cell lymphomas (Robbiani and Nussenzweig
431 2013). On the other hand, APOBEC may promote the formation of neoantigens and activate cytotoxic
432 T cell responses. It was shown that the APOBEC3 class of enzymes is upregulated in immune cells
433 (monocytes, macrophages and pDCs) upon treatment with interferon ɣ (F.X. Wang et al. 2008;

434 Stenglein et al. 2010). In ovarian carcinoma, the expression of APOBEC3G was positively correlated
435 with T cell infiltration, expression of cytotoxic granzyme and perforin (*GZMB*, *PRF1*), and improved
436 clinical outcome (Leonard et al. 2016). Moreover, APOBEC mutagenesis has previously been linked
437 with responses to immunotherapy in lung, melanoma and head and neck cancers (S. Wang et al. 2018;
438 Faden et al. 2019; Driscoll et al. 2020) through the induction of heteroclitic neoepitopes. We have
439 demonstrated increased APOBEC mutagenesis in a TMD setting across multiple tumour types, and
440 have shown that this process appears linked with the regulation of angiogenesis by the uPAR system.
441 It is possible that the mutational footprint left in the genome by such activity enables immune
442 recognition and maintenance of TMD in the time span until exhaustion and immune escape occurs.
443 This could suggest a window of opportunity for the employment of checkpoint inhibitors, possibly in
444 combination with APOBEC inhibitors, in early-stage tumours with TMD.

445 Our pan-cancer analysis unveils a complex interplay between TMD, exhaustion and APOBEC activity,
446 and a mutational context that may enable these phenotypes to develop. However, it is possible that
447 some of the associations uncovered may be due to the direct association between APOBEC activity
448 and immune responses/exhaustion or hypoxia, which have been previously reported in the literature as
449 highlighted above, rather than with tumour mass dormancy itself. Future experimental studies will be
450 needed to verify and further understand the implications of increased APOBEC mutagenesis in a TMD-
451 specific context. Moreover, changes in gene expression may not accurately reflect changes in
452 abundance of protein products in our defined dormancy and exhaustion programmes. Biochemical
453 assays and studies in model organisms will give a more definite answer to the nature of these
454 relationships.

455 An important limitation of this analysis consists in the fact that there are no currently available datasets
456 describing transcriptional changes specifically in the context of tumour mass dormancy. This
457 phenotype is difficult to study due to its temporary nature and few model systems for this state exist.
458 We have therefore not been able to unequivocally determine whether the chosen scoring method and
459 even the cut-offs employed for assessing TMD capture this phenomenon in the most accurate manner.
460 However, we are taking a conservative approach focusing only on the most extreme values, which are
461 most likely to reflect this state. This approach appears methodologically robust to small fluctuations in
462 expression/gene signatures, and the associations identified have been validated in independent datasets.
463 Moreover, the key findings of *CASP8* positive selection in TMD, associations with APOBEC
464 mutagenesis and cytotoxic microenvironment were also reproducible using the PCA-based scoring
465 methodology (Supplementary Figures 43-48). Further experimental evidence is needed to verify the
466 biological nature of these associations and the most appropriate way of evaluating TMD, once
467 experimental model systems for this state will have become more widely available and easier to study.

468 Another limitation refers to the fact that the samples analysed have been profiled using bulk sequencing
469 methods, which means that we are only capturing an average cellular signal across the entire tumour
470 which may not reflect its true heterogeneity. Furthermore, since TMD is a temporary phenotype, a
471 single static snapshot of tumours, even in this large dataset, may not be sufficient to capture the entire
472 array of TMD phenotypes that might develop in different tissues. Therefore, multi-regional sequencing
473 and longitudinal studies should be conducted in the future to shed further light on the diversity and
474 dynamics of TMD phenotypes.

475 Furthermore, our study focused mainly on pan-cancer markers of TMD, due to the scarcity of this
476 phenotype. From our analyses, we observe TMD is frequent in certain tissues such as head and neck,
477 breast, bone, and has weak evidence or complete absence of signals in others, e.g. chromophobe renal
478 cell carcinoma. Some of the smaller studies from TCGA may be insufficiently powered to capture such

signals. Therefore, future studies should focus on an in-depth characterization of TMD in distinct tissues and should be able to pinpoint tissue-specific genomic markers linked to TMD. Finally, the potential genomic biomarkers of dormancy that we identified need to be confirmed through experimental validation in suitable in vitro/in vivo models of TMD.

Overall, our study provides evidence for TMD, immunological and angiogenic dormancy signals across a variety of cancer types, and highlights key associated intrinsic and extrinsic hallmarks, including *CASP8*/*RAS* dependencies, APOBEC mutagenesis and hypoxia. These findings pave the way for further exploration of the mechanisms underlying TMD emergence, maintenance and exit.

4 MATERIALS AND METHODS

4.1 Datasets

FPKM normalised RNA-sequencing expression data as well as mutation annotation files aligned against the hg38 genome from the Muse pipeline, were downloaded using the *TCGAbiolinks* R package (Colaprico et al. 2016) for 9,712 treatment-naïve primary tumour TCGA samples across 31 solid cancer types. For patients with multiple samples available, one RNA-seq barcode entry was randomly selected for each individual patient resulting in 9,631 total entries. All expression data were log-transformed for downstream analysis.

For validation, RNA-seq and matched whole-genome sequencing data were downloaded for 12 cancer studies from the ICGC Data Portal (J. Zhang et al. 2019), accompanied by RNA-seq and matched targeted sequencing from cBioPortal (Cerami et al. 2012) for the following datasets: blca_mskcc_solit_2012 = Bladder Cancer (MSKCC, J Clin Onco 2013); brca_metabric = Breast Cancer (METABRIC, Nature 2012 & Nat Commun 2016); prad_mskcc = Prostate Adenocarcinoma (MSKCC, Cancer Cell 2010); sarc_mskcc = Sarcoma (MSKCC/Broad, Nat Genet 2010); whole-genome sequencing from rt_target_2018_pub = Pediatric Rhabdoid Tumor (TARGET, 2018) and prostate_dkfz_2018 = Prostate Cancer (DKFZ, Cancer Cell 2018); and whole-exome sequencing from: brca_smc_2018 = Breast Cancer (SMC 2018); GIS031 = Lung adenocarcinoma (GIS, Nat Genet 2019); paad_qcmg_uq_2016 = Pancreatic Adenocarcinoma (QCMG, Nature 2016); prad_broad = Prostate Adenocarinoma (Broad/Cornell, Nat Genet 2012); utuc_cornell_baylor_mdacc_2019 = Upper Tract Urothelial Carcinoma (Cornell/Naylor/MDACC, Nat Commun 2019); wt_target_2018_pub = Pediatric Wilm's Tumor (TARGET, 2018). The data were processed and analysed in the same manner as the TCGA data.

4.2 Quantifying the dormancy, exhaustion and APOBEC programmes

Mean log-transformed expression values of genes deemed to be associated with a given programme by manual curation of the literature (Supplementary Table 1) were used to produce per-sample programme scores for TMD, immunological/angiogenic dormancy, exhaustion and APOBEC activity. We have specifically selected genes that have been associated with immunological and angiogenic <u>dormancy</u>, rather than generic immunity or angiogenesis processes, in order to ensure that any associations identified downstream are likely to be TMD-related.

The APOBEC and exhaustion programme scores were calculated by taking the mean expression of genes within the respective programmes, as all genes in the programmes were expected to be upregulated in the respective states.

The angiogenic and immunological dormancy (AD/ID) programme scores, as well as the TMD scores, were calculated using two different approaches:

(a) Scaled difference of means

For every sample in the cohort, a score S was derived by subtracting the sum of expression values of downregulated genes (E_d) in the respective programme from the sum of expression values of upregulated genes (E_u) and dividing by the total number of genes within the programme (N_u+N_d, where N_u is the total number of upregulated genes and N_d is the total number of downregulated genes):

$$S = \frac{\sum_i^{N_u} E_{u,i} - \sum_j^{N_d} E_{d,j}}{N_u + N_d}$$

(b) Principal component analysis (PCA)

PCA was performed across the pan-cancer dataset using the up- and downregulated gene signature of the respective programme. The dormancy score was extracted as the coordinate of the first principal component (PC1).

The proliferation/apoptosis ratio was calculated by dividing the average expression of E2F target genes (including MKI67, a classical proliferation marker (Miller et al. 2018)), obtained from the "HALLMARK_E2F_TARGETS" gene list, by the average expression of genes involved in apoptosis from the "HALLMARK_APOPTOSIS" gene list. Both gene lists are part of the "H: hallmark gene sets" collection deposited at MSigDB.

Samples with high TMD were identified as being within the upper quartile of the TMD programme score range and presenting a ratio of proliferation/apoptosis < 1. Dormant samples were further classified as having immunological and/or angiogenic dormancy based on whether they were in the upper quartile of the angiogenic (AD) and/or immunological dormancy (ID) programme score range, respectively. Samples with no evidence for TMD (expanding tumours, marked as 'NO') were identified as being in the lower quartile of the TMD programme score range and showing a ratio of proliferation/apoptosis > 1. The remaining samples ('MID' category) were classed as having middle levels of expansion with some weak potential evidence of dormancy, but biologically unlikely to be dormant.

4.3 Assessing the robustness of the AD/ID programme scores

To assess the relative robustness of the three methodologies employed to calculate the TMD and AD/ID scores, we took two approaches:

(1) We systematically removed one gene from the respective programme at a time and recalculated the scores.
(2) A small amount of random noise was added to the expression of each gene in the signature, measured in each sample and scores were recalculated. The jitter R function was used to introduce variable amounts of noise in the data, by changing the factor variable level between 1-200). For each level of noise, the approach was repeated 100 times and the scores were recalculated.

The overall variability of the fold changes in the new score compared to the original scores was compared between the 'scaled difference of means' and PCA-based methods.

557 **4.4 Assessing the robustness of APOBEC associations**

558 Correlations between the mean expression of APOBEC-related genes and random gene expression
559 programmes were calculated by randomly selecting an equal number of genes to that found in the TMD
560 signatures (35) from the genome and assessing the mean expression correlation with APOBEC activity.
561 1000 iterations were performed. Associations varied as shown in Supplementary Figure 24, and were
562 consistently lower than the one calculated for TMD. P-values were frequently significant, but this is
563 likely due to the large number of samples.

564 **4.5 PHATE dimensionality reduction**

565 The *phateR* R package (Moon et al. 2019) was used to perform the dimensionality reduction based on
566 the expression of genes associated with the TMD and exhaustion programmes. A constant seed was
567 used for reproducibility. The *ComBat* function from the *sva* R package (Leek et al. 2012) was used to
568 remove tissue-specific expression patterns from the TCGA RNA-seq data.

569 **4.6 Association between cancer driver mutations and TMD**

570 The COSMIC database (Tate et al. 2019) was used to source a list of 723 known drivers of
571 tumorigenesis (Tiers 1+2). COSMIC genes mutated in at least 1% of samples across all solid primary
572 tumour samples were tested for enrichment or depletion of mutations between samples with high and
573 low TMD using a Fisher's exact test. Only missense, nonsense, nonstop, frameshift deletion/insertion
574 and inframe insertion/deletion mutations were considered in the analysis. For *HRAS*, *KRAS* and *NRAS*,
575 a Fisher's exact test was also performed to test for enrichment or depletion of specific recurrent hotspot
576 mutations, reported by the cBioPortal data hub and based in part on methodology from Chang et al
577 (Chang et al. 2016) and Gao et al (Gao et al. 2017). The analysis was also repeated on a cancer-by-
578 cancer basis, where COSMIC genes mutated in at least 5% of samples within the cancer-specific study
579 sample were tested for enrichment or depletion of mutations with TMD samples. The p-values were
580 corrected using Benjamini-Hochberg (BH) procedure to accommodate for multiple testing.

581 Enrichment analysis was reaffirmed using a random forest classification approach. This was conducted
582 using the *randomForest* R package. A balanced dataset of samples with high and low TMD was
583 generated by utilising all samples with low TMD (n = 657), and randomly sampling an equal number
584 of samples with high TMD (n = 1,154). The resulting random forest consisted of 500 trees, and
585 mutation enrichment was determined by calculating the mean decrease in Gini index between the two
586 groups following the removal of a gene.

587 To identify genes that are positively selected in the context of TMD, we carried out dN/dS analysis
588 separately for high and low TMD samples using the *dNdScv* R package (Martincorena et al. 2017), run
589 with default parameters.

590 **4.7 Mutational signature analysis**

591 Mutational signature contributions were inferred using *deconstructSigs* (Rosenthal et al. 2016) and the
592 choice of signatures was further informed using results from *SigProfiler* (Alexandrov, Nik-Zainal,
593 Wedge, Campbell, et al. 2013). Only samples with at least 50 mutations were employed in the analysis,
594 for a total of 6,410 samples.

595 *SigProfiler* was used to infer mutational signatures from TCGA whole-exome sequencing data. For
596 each TCGA study of interest, input mutational matrices were generated using the
597 *SigProfilerMatrixGeneratorR* function using all samples containing at least 50 mutations and aligning

598 the MAF files to the hg38 genome build. *SigProfilerExtractorR* was used to extract de novo mutational
599 signatures for each cancer type. For each study, the solution with the greatest number of mutational
600 signatures was chosen, for which also the sum of the solution stability (calculated average silhouette
601 coefficient) and minimum stability exceeded 1, and the minimum stability value did not fall below 0.4.
602 For the selected solutions, the identity of the mutational signatures was determined by calculating
603 cosine similarities with COSMIC v3.1 mutational signatures.

604 Mutational signatures from whole-exome sequencing data were also inferred using the *deconstructSigs*
605 R package. MAF files were aligned against the hg38 genome build and the COSMIC v3.1 mutational
606 signatures were employed in the analysis. For each cancer type, the contribution of all signatures within
607 the *deconstructSigs* solution was set to 0, apart from SBS1, SBS5, signatures identified in the
608 corresponding *SigProfiler* solution, as well as signatures which contribute on average at least 5% of
609 mutations across all samples within the *deconstructSigs* solution. For cancer types where *SigProfiler*
610 did not result in a stable solution, the *deconstructSigs* solution was used with the contribution of all
611 signatures set to 0, apart from SBS1, SBS5 as well as signatures which contribute at least 5% of
612 mutations across samples within the *deconstructSigs* solution.

613 **4.8 Tumour microenvironment deconvolution from bulk RNA-seq data**

614 The tumour microenvironment cell infiltration scores were calculated using the *ConsensusTME* R
615 package (Jiménez-Sánchez, Cast, and Miller 2019) based on gene sets from Bindea et al (Bindea et al.
616 2013). Cell abundance was estimated from the TCGA bulk RNA-seq data using singe sample Gene Set
617 Enrichment Analysis (ssGSEA). Broader categories (cytotoxic cells, Th cells, dendritic cells) were
618 scored by combining the expression across subtype-specific markers.

619 **4.9 Identification of APOBEC mutagenesis clusters**

620 First, APOBEC enrichment scores were calculated using the *sigminer* R package (Shixiang Wang et
621 al. 2020). Next, the mutational signature profiles of 6,410 samples underwent dimensionality reduction
622 with tSNE using the *Rtsne* R package, followed by expectation-maximization clustering using the
623 *EMCluster* R package. The optimal number of clusters was determined by considering the associated
624 increase in log-likelihood pertaining to additional clusters and was assumed to be equivalent for all
625 applications of tSNE (Supplementary Figure 9A). Whilst substantial increases were associated with
626 both 10 and 17 clusters, we selected 10 clusters in order to prevent the cluster of APOBEC-enriched
627 samples from being segmented. The APOBEC-enriched cluster was identified as the cluster with the
628 highest median collective enrichment of APOBEC signatures SBS2 and SBS13. The procedure was
629 conducted 100 times and samples which appeared in the APOBEC-enriched cluster on more than 50
630 occasions were labelled as 'APOBEC enriched'.

631 **4.10 Random forest modelling of APOBEC mutagenesis-gene expression associations**

632 Z-score normalisation of RNA-sequencing expression data was applied across each cancer type
633 individually. The FPKM-normalised values were log-transformed, and the results for each gene were
634 transformed into a $Z \sim N(0,1)$ distribution using the scale function in R. Only genes involved in the
635 dormancy and exhaustion programmes were considered as input in the classifier.

636 Genes whose expression contributed to APOBEC enrichment classification were identified using
637 machine learning via a random forest approach, conducted using the *randomForest* R package. In total,
638 5,850 samples were included for which Z-scores and mutational signature profiles were available. Each
639 random forest model required all samples labelled as APOBEC-enriched (n = 952), and an equal

number of samples taken randomly without replacement from the non-APOBEC-enriched samples to generate balanced groups. Each random forest consisted of 1,000 trees. The importance of each gene in the model, measured as the associated mean decrease in Gini index following its removal, was also calculated. This procedure was run 100 times.

Changes in median gene expression between the APOBEC enriched and non-enriched groups (Figure 4H) were determined using a two-sample Wilcoxon rank-sum test. P-values were corrected for multiple testing using the Benjamini-Hochberg procedure.

4.11 Calculation of hypoxia scores

Hypoxia scores were calculated using three transcription-based hypoxia signatures referred to as Buffa (Buffa et al. 2010), West (Eustace et al. 2013) and Winter (Winter et al. 2007), following the calculation procedure detailed by Bhandari et al (Bhandari et al. 2019). Genes which were included in the curated dormancy or exhaustion signatures were subsequently removed from the respective hypoxia signature. Scores were calculated based on the pan-cancer cohort to enable variation in hypoxia between cancer types to be identified. Since we observed a high degree of agreement between the three signatures (Supplementary Figure 11), the Buffa signature was employed in the downstream analysis on account of its use as a reference by Bhandari et al. A binning approach with a window size of 10 was employed to further subdivide tumours into discrete subgroups with increasing levels of hypoxia.

4.12 Survival Analysis

Survival analysis was conducted via the *survival* R package. Hazard ratios were calculated using Cox proportional hazards models and were displayed using the *finalfit* R package. Survival curves were calculated using the Kaplan-Meier formula, and plotted using the *survminer* R package.

4.13 Statistics

Pairwise correlations were calculated using the Pearson correlation statistics. The *corrplot* R package was used to carry out statistical analysis and visualise the correlation matrices. Groups were compared using the Student's t test, Wilcoxon rank-sum test or Kruskal-Wallis test as appropriate. Multiple testing correction was applied using the Benjamini-Hochberg method. The significance threshold was taken as $p < 0.05$.

4.14 Code

All code developed for the purpose of this analysis can be found at the following repository: https://github.com/secrierlab/tumourMassDormancy/tree/v1.0

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

MS designed and supervised the study. AJW, DHJ and WL performed the analyses. All authors wrote and approved the manuscript.

7 Funding

AJW and DHJ were supported by MRC DTP grants (MR/N013867/1). MS was supported by a UKRI Future Leaders Fellowship (MR/T042184/1), an Academy of Medical Sciences Springboard Award (SBF004\1042) and a Wellcome Trust Seed Award in Science (215296/Z/19/Z).

8 Data Availability Statement

The results published here are based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

9 REFERENCES

Aguirre Ghiso, J. A. 2002. "Inhibition of FAK signaling activated by urokinase receptor induces dormancy in human carcinoma cells in vivo." *Oncogene* 21 (16): 2513-24. https://doi.org/10.1038/sj.onc.1205342. https://www.ncbi.nlm.nih.gov/pubmed/11971186.

Aguirre Ghiso, J. A., K. Kovalski, and L. Ossowski. 1999. "Tumor dormancy induced by downregulation of urokinase receptor in human carcinoma involves integrin and MAPK signaling." *J Cell Biol* 147 (1): 89-104. https://doi.org/10.1083/jcb.147.1.89. https://www.ncbi.nlm.nih.gov/pubmed/10508858.

Aguirre-Ghiso, J. A. 2007. "Models, mechanisms and clinical evidence for cancer dormancy." *Nat Rev Cancer* 7 (11): 834-46. https://doi.org/10.1038/nrc2256. https://www.ncbi.nlm.nih.gov/pubmed/17957189.

Aguirre-Ghiso, J. A., Y. Estrada, D. Liu, and L. Ossowski. 2003. "ERK(MAPK) activity as a determinant of tumor growth and dormancy; regulation by p38(SAPK)." *Cancer Res* 63 (7): 1684-95. https://www.ncbi.nlm.nih.gov/pubmed/12670923.

Alexandrov, L. B., P. H. Jones, D. C. Wedge, J. E. Sale, P. J. Campbell, S. Nik-Zainal, and M. R. Stratton. 2015. "Clock-like mutational processes in human somatic cells." *Nat Genet* 47 (12): 1402-7. https://doi.org/10.1038/ng.3441. https://www.ncbi.nlm.nih.gov/pubmed/26551669.

Alexandrov, L. B., J. Kim, N. J. Haradhvala, M. N. Huang, A. W. Tian Ng, Y. Wu, A. Boot, K. R. Covington, D. A. Gordenin, E. N. Bergstrom, S. M. A. Islam, N. Lopez-Bigas, L. J. Klimczak, J. R. McPherson, S. Morganella, R. Sabarinathan, D. A. Wheeler, V. Mustonen, G. Getz, S. G. Rozen, M. R. Stratton, PCAWG Mutational Signatures Working Group, and PCAWG Consortium. 2020. "The repertoire of mutational signatures in human cancer." *Nature* 578 (7793): 94-101. https://doi.org/10.1038/s41586-020-1943-3. https://www.ncbi.nlm.nih.gov/pubmed/32025018.

Alexandrov, L. B., S. Nik-Zainal, D. C. Wedge, S. A. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A. L. Børresen-Dale, S. Boyault, B. Burkhardt, A. P. Butler, C. Caldas, H. R. Davies, C. Desmedt, R. Eils, J. E. Eyfjörd, J. A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinski, M. Imielinsk, N. Jäger, D. T. Jones, D. Jones, S. Knappskog, M. Kool, S. R. Lakhani, C. López-Otín, S. Martin, N. C. Munshi, H. Nakamura, P. A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J. V. Pearson, X. S. Puente, K. Raine, M. Ramakrishna, A. L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T. N. Schumacher, P. N. Span, J. W. Teague, Y. Totoki, A. N. Tutt, R. Valdés-Mas, M. M. van Buuren, L. van 't Veer, A. Vincent-Salomon, N. Waddell, L. R. Yates, J. Zucman-Rossi, P. A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S. M. Grimmond, R. Siebert, E. Campo, T. Shibata, S. M. Pfister, P. J. Campbell, M. R. Stratton, Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, and ICGC PedBrain. 2013. "Signatures of mutational processes in human cancer." *Nature* 500

720 (7463): 415-21. https://doi.org/10.1038/nature12477.
721 https://www.ncbi.nlm.nih.gov/pubmed/23945592.

722 Alexandrov, L. B., S. Nik-Zainal, D. C. Wedge, P. J. Campbell, and M. R. Stratton. 2013.
723 "Deciphering signatures of mutational processes operative in human cancer." *Cell Rep* 3 (1):
724 246-59. https://doi.org/10.1016/j.celrep.2012.12.008.
725 https://www.ncbi.nlm.nih.gov/pubmed/23318258.

726 Asnaghi, L., M. H. Lin, K. S. Lim, K. J. Lim, A. Tripathy, M. Wendeborn, S. L. Merbs, J. T. Handa,
727 A. Sodhi, E. E. Bar, and C. G. Eberhart. 2014. "Hypoxia promotes uveal melanoma invasion
728 through enhanced Notch and MAPK activation." *PLoS One* 9 (8): e105372.
729 https://doi.org/10.1371/journal.pone.0105372.
730 https://www.ncbi.nlm.nih.gov/pubmed/25166211.

731 Bailey, S. R., M. H. Nelson, R. A. Himes, Z. Li, S. Mehrotra, and C. M. Paulos. 2014. "Th17 cells in
732 cancer: the ultimate identity crisis." *Front Immunol* 5: 276.
733 https://doi.org/10.3389/fimmu.2014.00276. https://www.ncbi.nlm.nih.gov/pubmed/24987392.

734 Barr, A. R., S. Cooper, F. S. Heldt, F. Butera, H. Stoy, J. Mansfeld, B. Novák, and C. Bakal. 2017.
735 "DNA damage during S-phase mediates the proliferation-quiescence decision in the
736 subsequent G1 via p21 expression." *Nat Commun* 8: 14728.
737 https://doi.org/10.1038/ncomms14728. https://www.ncbi.nlm.nih.gov/pubmed/28317845.

738 Benci, J. L., B. Xu, Y. Qiu, T. J. Wu, H. Dada, C. Twyman-Saint Victor, L. Cucolo, D. S. M. Lee, K.
739 E. Pauken, A. C. Huang, T. C. Gangadhar, R. K. Amaravadi, L. M. Schuchter, M. D.
740 Feldman, H. Ishwaran, R. H. Vonderheide, A. Maity, E. J. Wherry, and A. J. Minn. 2016.
741 "Tumor Interferon Signaling Regulates a Multigenic Resistance Program to Immune
742 Checkpoint Blockade." *Cell* 167 (6): 1540-1554.e12.
743 https://doi.org/10.1016/j.cell.2016.11.022. https://www.ncbi.nlm.nih.gov/pubmed/27912061.

744 Bhandari, V., C. Hoey, L. Y. Liu, E. Lalonde, J. Ray, J. Livingstone, R. Lesurf, Y. J. Shiah, T.
745 Vujcic, X. Huang, S. M. G. Espiritu, L. E. Heisler, F. Yousif, V. Huang, T. N. Yamaguchi, C.
746 Q. Yao, V. Y. Sabelnykova, M. Fraser, M. L. K. Chua, T. van der Kwast, S. K. Liu, P. C.
747 Boutros, and R. G. Bristow. 2019. "Molecular landmarks of tumor hypoxia across cancer
748 types." *Nat Genet* 51 (2): 308-318. https://doi.org/10.1038/s41588-018-0318-2.
749 https://www.ncbi.nlm.nih.gov/pubmed/30643250.

750 Bindea, G., B. Mlecnik, M. Tosolini, A. Kirilovsky, M. Waldner, A. C. Obenauf, H. Angell, T.
751 Fredriksen, L. Lafontaine, A. Berger, P. Bruneval, W. H. Fridman, C. Becker, F. Pagès, M. R.
752 Speicher, Z. Trajanoski, and J. Galon. 2013. "Spatiotemporal dynamics of intratumoral
753 immune cells reveal the immune landscape in human cancer." *Immunity* 39 (4): 782-95.
754 https://doi.org/10.1016/j.immuni.2013.10.003.
755 https://www.ncbi.nlm.nih.gov/pubmed/24138885.

756 Boire, A., S. B. Coffelt, S. A. Quezada, M. G. Vander Heiden, and A. T. Weeraratna. 2019. "Tumour
757 Dormancy and Reawakening: Opportunities and Challenges." *Trends Cancer* 5 (12): 762-765.
758 https://doi.org/10.1016/j.trecan.2019.10.010.
759 https://www.ncbi.nlm.nih.gov/pubmed/31813453.

760 Boman, B. M., and J. Z. Fields. 2013. "An APC:WNT Counter-Current-Like Mechanism Regulates
761 Cell Division Along the Human Colonic Crypt Axis: A Mechanism That Explains How APC
762 Mutations Induce Proliferative Abnormalities That Drive Colon Cancer Development." *Front*
763 *Oncol* 3: 244. https://doi.org/10.3389/fonc.2013.00244.
764 https://www.ncbi.nlm.nih.gov/pubmed/24224156.

765 Buffa, F. M., A. L. Harris, C. M. West, and C. J. Miller. 2010. "Large meta-analysis of multiple
766 cancers reveals a common, compact and highly prognostic hypoxia metagene." *Br J Cancer*
767 102 (2): 428-35. https://doi.org/10.1038/sj.bjc.6605450.
768 https://www.ncbi.nlm.nih.gov/pubmed/20087356.

769 Burns, M. B., N. A. Temiz, and R. S. Harris. 2013. "Evidence for APOBEC3B mutagenesis in
770 multiple human cancers." *Nat Genet* 45 (9): 977-83. https://doi.org/10.1038/ng.2701.
771 https://www.ncbi.nlm.nih.gov/pubmed/23852168.

772 Butturini, E., A. Carcereri de Prati, D. Boriero, and S. Mariotto. 2019. "Tumor Dormancy and
773 Interplay with Hypoxic Tumor Microenvironment." *Int J Mol Sci* 20 (17).
774 https://doi.org/10.3390/ijms20174305. https://www.ncbi.nlm.nih.gov/pubmed/31484342.

775 Cerami, E., J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne,
776 M. L. Heuer, E. Larsson, Y. Antipin, B. Reva, A. P. Goldberg, C. Sander, and N. Schultz.
777 2012. "The cBio cancer genomics portal: an open platform for exploring multidimensional
778 cancer genomics data." *Cancer Discov* 2 (5): 401-4. https://doi.org/10.1158/2159-8290.CD-
779 12-0095. https://www.ncbi.nlm.nih.gov/pubmed/22588877.

780 Chang, M. T., S. Asthana, S. P. Gao, B. H. Lee, J. S. Chapman, C. Kandoth, J. Gao, N. D. Socci, D.
781 B. Solit, A. B. Olshen, N. Schultz, and B. S. Taylor. 2016. "Identifying recurrent mutations in
782 cancer reveals widespread lineage diversity and mutational specificity." *Nat Biotechnol* 34
783 (2): 155-63. https://doi.org/10.1038/nbt.3391.
784 https://www.ncbi.nlm.nih.gov/pubmed/26619011.

785 Chen, H., W. Chong, Q. Wu, Y. Yao, M. Mao, and X. Wang. 2019. "Association of." *Front Immunol*
786 10: 1113. https://doi.org/10.3389/fimmu.2019.01113.
787 https://www.ncbi.nlm.nih.gov/pubmed/31164891.

788 Chen, L., A. Endler, and F. Shibasaki. 2009. "Hypoxia and angiogenesis: regulation of hypoxia-
789 inducible factors via novel binding factors." *Exp Mol Med* 41 (12): 849-57.
790 https://doi.org/10.3858/emm.2009.41.12.103.
791 https://www.ncbi.nlm.nih.gov/pubmed/19942820.

792 Chin, Y. E., M. Kitagawa, W. C. Su, Z. H. You, Y. Iwamoto, and X. Y. Fu. 1996. "Cell growth arrest
793 and induction of cyclin-dependent kinase inhibitor p21 WAF1/CIP1 mediated by STAT1."
794 *Science* 272 (5262): 719-22. https://doi.org/10.1126/science.272.5262.719.
795 https://www.ncbi.nlm.nih.gov/pubmed/8614832.

796 Colaprico, A., T. C. Silva, C. Olsen, L. Garofano, C. Cava, D. Garolini, T. S. Sabedot, T. M. Malta,
797 S. M. Pagnotta, I. Castiglioni, M. Ceccarelli, G. Bontempi, and H. Noushmehr. 2016.
798 "TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data." *Nucleic*
799 *Acids Res* 44 (8): e71. https://doi.org/10.1093/nar/gkv1507.
800 https://www.ncbi.nlm.nih.gov/pubmed/26704973.

801 Damen, M. P. F., J. van Rheenen, and C. L. G.J Scheele. 2020. "Targeting dormant tumor cells to
802 prevent cancer recurrence." *FEBS J*. https://doi.org/10.1111/febs.15626.
803 https://www.ncbi.nlm.nih.gov/pubmed/33190412.

804 De Blasio, A., R. Di Fiore, M. Morreale, D. Carlisi, R. Drago-Ferrante, M. Montalbano, C. Scerri, G.
805 Tesoriere, and R. Vento. 2016. "Unusual roles of caspase-8 in triple-negative breast cancer
806 cell line MDA-MB-231." *Int J Oncol* 48 (6): 2339-48. https://doi.org/10.3892/ijo.2016.3474.
807 https://www.ncbi.nlm.nih.gov/pubmed/27082853.

808 Degasperi, A., T. D. Amarante, J. Czarnecki, S. Shooter, X. Zou, D. Glodzik, S. Morganella, A. S.
809 Nanda, C. Badja, G. Koh, S. E. Momen, I. Georgakopoulos-Soares, J. M. L. Dias, J. Young,
810 Y. Memari, H. Davies, and S. Nik-Zainal. 2020. "A practical framework and online tool for
811 mutational signature analyses show inter-tissue variation and driver dependencies." *Nat*
812 *Cancer* 1 (2): 249-263. https://doi.org/10.1038/s43018-020-0027-5.
813 https://www.ncbi.nlm.nih.gov/pubmed/32118208.

814 Driscoll, C. B., M. R. Schuelke, T. Kottke, J. M. Thompson, P. Wongthida, J. M. Tonne, A. L. Huff,
815 A. Miller, K. G. Shim, A. Molan, C. Wetmore, P. Selby, A. Samson, K. Harrington, H.
816 Pandha, A. Melcher, J. S. Pulido, R. Harris, L. Evgin, and R. G. Vile. 2020. "APOBEC3B-
817 mediated corruption of the tumor cell immunopeptidome induces heteroclitic neoepitopes for
818 cancer immunotherapy." *Nat Commun* 11 (1): 790. https://doi.org/10.1038/s41467-020-
819 14568-7. https://www.ncbi.nlm.nih.gov/pubmed/32034147.

820 Dunn, G. P., C. M. Koebel, and R. D. Schreiber. 2006. "Interferons, immunity and cancer
821 immunoediting." *Nat Rev Immunol* 6 (11): 836-48. https://doi.org/10.1038/nri1961.
822 https://www.ncbi.nlm.nih.gov/pubmed/17063185.

823 Dunn, G. P., L. J. Old, and R. D. Schreiber. 2004a. "The immunobiology of cancer
824 immunosurveillance and immunoediting." *Immunity* 21 (2): 137-48.
825 https://doi.org/10.1016/j.immuni.2004.07.017.
826 https://www.ncbi.nlm.nih.gov/pubmed/15308095.

827 ---. 2004b. "The three Es of cancer immunoediting." *Annu Rev Immunol* 22: 329-60.
828 https://doi.org/10.1146/annurev.immunol.22.012703.104803.
829 https://www.ncbi.nlm.nih.gov/pubmed/15032581.

830 Eustace, A., N. Mani, P. N. Span, J. J. Irlam, J. Taylor, G. N. Betts, H. Denley, C. J. Miller, J. J.
831 Homer, A. M. Rojas, P. J. Hoskin, F. M. Buffa, A. L. Harris, J. H. Kaanders, and C. M. West.
832 2013. "A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in
833 laryngeal cancer but not bladder cancer." *Clin Cancer Res* 19 (17): 4879-88.
834 https://doi.org/10.1158/1078-0432.CCR-13-0542.
835 https://www.ncbi.nlm.nih.gov/pubmed/23820108.

836 Faden, D. L., F. Ding, Y. Lin, S. Zhai, F. Kuo, T. A. Chan, L. G. Morris, and R. L. Ferris. 2019.
837 "APOBEC mutagenesis is tightly linked to the immune landscape and immunotherapy
838 biomarkers in head and neck squamous cell carcinoma." *Oral Oncol* 96: 140-147.
839 https://doi.org/10.1016/j.oraloncology.2019.07.020.
840 https://www.ncbi.nlm.nih.gov/pubmed/31422205.

841 Finn, O. J. 2006. "Human tumor antigens, immunosurveillance, and cancer vaccines." *Immunol Res*
842 36 (1-3): 73-82. https://doi.org/10.1385/IR:36:1:73.
843 https://www.ncbi.nlm.nih.gov/pubmed/17337768.

844 Forcet, C., X. Ye, L. Granger, V. Corset, H. Shin, D. E. Bredesen, and P. Mehlen. 2001. "The
845 dependence receptor DCC (deleted in colorectal cancer) defines an alternative mechanism for
846 caspase activation." *Proc Natl Acad Sci U S A* 98 (6): 3416-21.
847 https://doi.org/10.1073/pnas.051378298. https://www.ncbi.nlm.nih.gov/pubmed/11248093.

848 Gao, J., M. T. Chang, H. C. Johnsen, S. P. Gao, B. E. Sylvester, S. O. Sumer, H. Zhang, D. B. Solit,
849 B. S. Taylor, N. Schultz, and C. Sander. 2017. "3D clusters of somatic mutations in cancer
850 reveal numerous rare mutations as functional targets." *Genome Med* 9 (1): 4.
851 https://doi.org/10.1186/s13073-016-0393-x.
852 https://www.ncbi.nlm.nih.gov/pubmed/28115009.

853 Gerlinger, M., N. McGranahan, S. M. Dewhurst, R. A. Burrell, I. Tomlinson, and C. Swanton. 2014.
854 "Cancer: evolution within a lifetime." *Annu Rev Genet* 48: 215-36.
855 https://doi.org/10.1146/annurev-genet-120213-092314.
856 https://www.ncbi.nlm.nih.gov/pubmed/25292359.

857 Ghorani, E., J. L. Reading, J. Y. Henry, M. R. de Massy, R. Rosenthal, V. Turati, K. Joshi, A. J. S.
858 Furness, A. B. Aissa, S. K. Saini, S. Ramskov, A. Georgiou, M. W. Sunderland, Y. N. S.
859 Wong, M. V. De Mucha, W. Day, F. Galvez-Cancino, P. D. Becker, I. Uddin, M. Ismail, T.
860 Ronel, A. Woolston, M. Jamal-Hanjani, S. Veeriah, N. J. Birkbak, G. A. Wilson, K.
861 Litchfield, L. Conde, J. A. Guerra-Assunção, K. Blighe, D. Biswas, R. Salgado, T. Lund, M.
862 Al Bakir, D. A. Moore, C. T. Hiley, S. Loi, Y. Sun, Y. Yuan, K. AbdulJabbar, S. Turajilic, J.
863 Herrero, T. Enver, S. R. Hadrup, A. Hackshaw, K. S. Peggs, N. McGranahan, B. Chain, C.
864 Swanton, and S. A. Quezada. 2020. "The T cell differentiation landscape is shaped by tumour
865 mutations in lung cancer." *Nat Cancer* 1 (5): 546-561. https://doi.org/10.1038/s43018-020-
866 0066-y. https://www.ncbi.nlm.nih.gov/pubmed/32803172.

867 Gimbrone, M. A., S. B. Leapman, R. S. Cotran, and J. Folkman. 1972. "Tumor dormancy in vivo by
868 prevention of neovascularization." *J Exp Med* 136 (2): 261-76.
869 https://doi.org/10.1084/jem.136.2.261. https://www.ncbi.nlm.nih.gov/pubmed/5043412.

870 Green, A. M., and M. D. Weitzman. 2019. "The spectrum of APOBEC3 activity: From anti-viral
871 agents to anti-cancer opportunities." *DNA Repair (Amst)* 83: 102700.
872 https://doi.org/10.1016/j.dnarep.2019.102700.
873 https://www.ncbi.nlm.nih.gov/pubmed/31563041.

874 Gužvić, M., and C. A. Klein. 2013. "Cancer dormancy: time to explore its clinical relevance." *Breast*
875 *Cancer Res* 15 (6): 321. https://doi.org/10.1186/bcr3590.
876 https://www.ncbi.nlm.nih.gov/pubmed/24359585.

877 HADFIELD, G. 1954. "The dormant cancer cell." *Br Med J* 2 (4888): 607-10.
878 https://doi.org/10.1136/bmj.2.4888.607. https://www.ncbi.nlm.nih.gov/pubmed/13190204.

879 Heldt, F. S., A. R. Barr, S. Cooper, C. Bakal, and B. Novák. 2018. "A comprehensive model for the
880 proliferation-quiescence decision in response to endogenous DNA damage in human cells."
881 *Proc Natl Acad Sci U S A* 115 (10): 2532-2537. https://doi.org/10.1073/pnas.1715345115.
882 https://www.ncbi.nlm.nih.gov/pubmed/29463760.

883 Holmgren, L., M. S. O'Reilly, and J. Folkman. 1995. "Dormancy of micrometastases: balanced
884 proliferation and apoptosis in the presence of angiogenesis suppression." *Nat Med* 1 (2): 149-
885 53. https://doi.org/10.1038/nm0295-149. https://www.ncbi.nlm.nih.gov/pubmed/7585012.

886 Huang, S. 2021. "Reconciling Non-Genetic Plasticity with Somatic Evolution in Cancer." *Trends*
887 *Cancer* 7 (4): 309-322. https://doi.org/10.1016/j.trecan.2020.12.007.
888 https://www.ncbi.nlm.nih.gov/pubmed/33536158.

889 Ikeda, H., L. J. Old, and R. D. Schreiber. 2002. "The roles of IFN gamma in protection against tumor
890 development and cancer immunoediting." *Cytokine Growth Factor Rev* 13 (2): 95-109.
891 https://doi.org/10.1016/s1359-6101(01)00038-7.
892 https://www.ncbi.nlm.nih.gov/pubmed/11900986.

893 Itahana, K., G. P. Dimri, E. Hara, Y. Itahana, Y. Zou, P. Y. Desprez, and J. Campisi. 2002. "A role
894 for p53 in maintaining and establishing the quiescence growth arrest in human cells." *J Biol*
895 *Chem* 277 (20): 18206-14. https://doi.org/10.1074/jbc.M201028200.
896 https://www.ncbi.nlm.nih.gov/pubmed/11880381.

897 Jager, M., F. Blokzijl, E. Kuijk, J. Bertl, M. Vougioukalaki, R. Janssen, N. Besselink, S. Boymans, J.
898 de Ligt, J. S. Pedersen, J. Hoeijmakers, J. Pothof, R. van Boxtel, and E. Cuppen. 2019.
899 "Deficiency of nucleotide excision repair is associated with mutational signature observed in
900 cancer." *Genome Res* 29 (7): 1067-1077. https://doi.org/10.1101/gr.246223.118.
901 https://www.ncbi.nlm.nih.gov/pubmed/31221724.

902 Jahanban-Esfahlan, R., K. Seidi, M. H. Manjili, A. Jahanban-Esfahlan, T. Javaheri, and P. Zare.
903 2019. "Tumor Cell Dormancy: Threat or Opportunity in the Fight against Cancer." *Cancers*
904 *(Basel)* 11 (8). https://doi.org/10.3390/cancers11081207.
905 https://www.ncbi.nlm.nih.gov/pubmed/31430951.

906 Jelaković, B., X. Castells, K. Tomić, M. Ardin, S. Karanović, and J. Zavadil. 2015. "Renal cell
907 carcinomas of chronic kidney disease patients harbor the mutational signature of carcinogenic
908 aristolochic acid." *Int J Cancer* 136 (12): 2967-72. https://doi.org/10.1002/ijc.29338.
909 https://www.ncbi.nlm.nih.gov/pubmed/25403517.

910 Jiménez-Sánchez, A., O. Cast, and M. L. Miller. 2019. "Comprehensive Benchmarking and
911 Integration of Tumor Microenvironment Cell Estimation Methods." *Cancer Res* 79 (24):
912 6238-6246. https://doi.org/10.1158/0008-5472.CAN-18-3560.
913 https://www.ncbi.nlm.nih.gov/pubmed/31641033.

914 Katoh, M. 2012. "Function and cancer genomics of FAT family genes (review)." *Int J Oncol* 41 (6):
915 1913-8. https://doi.org/10.3892/ijo.2012.1669.
916 https://www.ncbi.nlm.nih.gov/pubmed/23076869.

917 Kim, M. J., C. Cervantes, Y. S. Jung, X. Zhang, J. Zhang, S. H. Lee, S. Jun, L. Litovchick, W. Wang,
918 J. Chen, B. Fang, and J. I. Park. 2021. "PAF remodels the DREAM complex to bypass cell
919 quiescence and promote lung tumorigenesis." *Mol Cell*.
920 https://doi.org/10.1016/j.molcel.2021.02.001.
921 https://www.ncbi.nlm.nih.gov/pubmed/33626321.

922 Koebel, C. M., W. Vermi, J. B. Swann, N. Zerafa, S. J. Rodig, L. J. Old, M. J. Smyth, and R. D.
923 Schreiber. 2007. "Adaptive immunity maintains occult cancer in an equilibrium state." *Nature*
924 450 (7171): 903-7. https://doi.org/10.1038/nature06309.
925 https://www.ncbi.nlm.nih.gov/pubmed/18026089.

926 Krock, B. L., N. Skuli, and M. C. Simon. 2011. "Hypoxia-induced angiogenesis: good and evil."
927 *Genes Cancer* 2 (12): 1117-33. https://doi.org/10.1177/1947601911423654.
928 https://www.ncbi.nlm.nih.gov/pubmed/22866203.

929 Kundu, A., H. Nam, S. Shelar, D. S. Chandrashekar, G. Brinkley, S. Karki, T. Mitchell, C. B. Livi, P.
930 Buckhaults, R. Kirkman, Y. Tang, G. C. Rowe, S. Wei, S. Varambally, and S. Sudarshan.
931 2020. "PRDM16 suppresses HIF-targeted gene expression in kidney cancer." *J Exp Med* 217
932 (6). https://doi.org/10.1084/jem.20191005. https://www.ncbi.nlm.nih.gov/pubmed/32251515.

933 Leek, J. T., W. E. Johnson, H. S. Parker, A. E. Jaffe, and J. D. Storey. 2012. "The sva package for
934 removing batch effects and other unwanted variation in high-throughput experiments."
935 *Bioinformatics* 28 (6): 882-3. https://doi.org/10.1093/bioinformatics/bts034.
936 https://www.ncbi.nlm.nih.gov/pubmed/22257669.

937 Leonard, B., G. J. Starrett, M. J. Maurer, A. L. Oberg, M. Van Bockstal, J. Van Dorpe, O. De Wever,
938 J. Helleman, A. M. Sieuwerts, E. M. Berns, J. W. Martens, B. D. Anderson, W. L. Brown, K.
939 R. Kalli, S. H. Kaufmann, and R. S. Harris. 2016. "APOBEC3G Expression Correlates with
940 T-Cell Infiltration and Improved Clinical Outcomes in High-grade Serous Ovarian

941 Carcinoma." *Clin Cancer Res* 22 (18): 4746-55. https://doi.org/10.1158/1078-0432.CCR-15-
942 2910. https://www.ncbi.nlm.nih.gov/pubmed/27016308.

943 Liu, D., J. Aguirre Ghiso, Y. Estrada, and L. Ossowski. 2002. "EGFR is a transducer of the urokinase
944 receptor initiated signal that is required for in vivo growth of a human carcinoma." *Cancer*
945 *Cell* 1 (5): 445-57. https://doi.org/10.1016/s1535-6108(02)00072-7.
946 https://www.ncbi.nlm.nih.gov/pubmed/12124174.

947 Lyu, T., N. Jia, J. Wang, X. Yan, Y. Yu, Z. Lu, R. C. Bast, K. Hua, and W. Feng. 2013. "Expression
948 and epigenetic regulation of angiogenesis-related factors during dormancy and recurrent
949 growth of ovarian carcinoma." *Epigenetics* 8 (12): 1330-46.
950 https://doi.org/10.4161/epi.26675. https://www.ncbi.nlm.nih.gov/pubmed/24135786.

951 MacDonald, J., Y. Ramos-Valdes, P. Perampalam, L. Litovchick, G. E. DiMattia, and F. A. Dick.
952 2017. "A Systematic Analysis of Negative Growth Control Implicates the DREAM Complex
953 in Cancer Cell Dormancy." *Mol Cancer Res* 15 (4): 371-381. https://doi.org/10.1158/1541-
954 7786.MCR-16-0323-T. https://www.ncbi.nlm.nih.gov/pubmed/28031411.

955 Martincorena, I., K. M. Raine, M. Gerstung, K. J. Dawson, K. Haase, P. Van Loo, H. Davies, M. R.
956 Stratton, and P. J. Campbell. 2017. "Universal Patterns of Selection in Cancer and Somatic
957 Tissues." *Cell* 171 (5): 1029-1041.e21. https://doi.org/10.1016/j.cell.2017.09.042.
958 https://www.ncbi.nlm.nih.gov/pubmed/29056346.

959 McGranahan, N., and C. Swanton. 2017. "Clonal Heterogeneity and Tumor Evolution: Past, Present,
960 and the Future." *Cell* 168 (4): 613-628. https://doi.org/10.1016/j.cell.2017.01.018.
961 https://www.ncbi.nlm.nih.gov/pubmed/28187284.

962 Miller, I., M. Min, C. Yang, C. Tian, S. Gookin, D. Carter, and S. L. Spencer. 2018. "Ki67 is a
963 Graded Rather than a Binary Marker of Proliferation versus Quiescence." *Cell Rep* 24 (5):
964 1105-1112.e5. https://doi.org/10.1016/j.celrep.2018.06.110.
965 https://www.ncbi.nlm.nih.gov/pubmed/30067968.

966 Mittal, D., M. M. Gubin, R. D. Schreiber, and M. J. Smyth. 2014. "New insights into cancer
967 immunoediting and its three component phases--elimination, equilibrium and escape." *Curr*
968 *Opin Immunol* 27: 16-25. https://doi.org/10.1016/j.coi.2014.01.004.
969 https://www.ncbi.nlm.nih.gov/pubmed/24531241.

970 Moon, K. R., D. van Dijk, Z. Wang, S. Gigante, D. B. Burkhardt, W. S. Chen, K. Yim, A. V. D.
971 Elzen, M. J. Hirn, R. R. Coifman, N. B. Ivanova, G. Wolf, and S. Krishnaswamy. 2019.
972 "Visualizing structure and transitions in high-dimensional biological data." *Nat Biotechnol* 37
973 (12): 1482-1492. https://doi.org/10.1038/s41587-019-0336-3.
974 https://www.ncbi.nlm.nih.gov/pubmed/31796933.

975 Moserle, L., A. Amadori, and S. Indraccolo. 2009. "The angiogenic switch: implications in the
976 regulation of tumor dormancy." *Curr Mol Med* 9 (8): 935-41.
977 https://doi.org/10.2174/156652409789712800.
978 https://www.ncbi.nlm.nih.gov/pubmed/19925406.

979 Naumov, G. N., E. Bender, D. Zurakowski, S. Y. Kang, D. Sampson, E. Flynn, R. S. Watnick, O.
980 Straume, L. A. Akslen, J. Folkman, and N. Almog. 2006. "A model of human tumor
981 dormancy: an angiogenic switch from the nonangiogenic phenotype." *J Natl Cancer Inst* 98
982 (5): 316-25. https://doi.org/10.1093/jnci/djj068.
983 https://www.ncbi.nlm.nih.gov/pubmed/16507828.

984 Oh, C. W., J. Hoover-Plow, and E. F. Plow. 2003. "The role of plasminogen in angiogenesis in vivo."
985 *J Thromb Haemost* 1 (8): 1683-7. https://doi.org/10.1046/j.1538-7836.2003.00182.x.
986 https://www.ncbi.nlm.nih.gov/pubmed/12911578.

987 Park, S. Y., and J. S. Nam. 2020. "The force awakens: metastatic dormant cancer cells." *Exp Mol*
988 *Med* 52 (4): 569-581. https://doi.org/10.1038/s12276-020-0423-z.
989 https://www.ncbi.nlm.nih.gov/pubmed/32300189.

990 Petrova, V., M. Annicchiarico-Petruzzelli, G. Melino, and I. Amelio. 2018. "The hypoxic tumour
991 microenvironment." *Oncogenesis* 7 (1): 10. https://doi.org/10.1038/s41389-017-0011-9.
992 https://www.ncbi.nlm.nih.gov/pubmed/29362402.

993 Phan, T. G., and P. I. Croucher. 2020. "The dormant cancer cell life cycle." *Nat Rev Cancer* 20 (7):
994 398-411. https://doi.org/10.1038/s41568-020-0263-0.
995 https://www.ncbi.nlm.nih.gov/pubmed/32488200.

996 Polyak, K., Y. Xia, J. L. Zweier, K. W. Kinzler, and B. Vogelstein. 1997. "A model for p53-induced
997 apoptosis." *Nature* 389 (6648): 300-5. https://doi.org/10.1038/38525.
998 https://www.ncbi.nlm.nih.gov/pubmed/9305847.

999 Pugh, C. W., and P. J. Ratcliffe. 2003. "Regulation of angiogenesis by hypoxia: role of the HIF
1000 system." *Nat Med* 9 (6): 677-84. https://doi.org/10.1038/nm0603-677.
1001 https://www.ncbi.nlm.nih.gov/pubmed/12778166.

1002 Qin, L., Y. Xu, G. Ma, L. Liao, Y. Wu, Y. Li, X. Wang, J. Jiang, J. Wang, and J. Xu. 2015. "NCOA1
1003 promotes angiogenesis in breast tumors by simultaneously enhancing both HIF1α- and AP-1-
1004 mediated VEGFa transcription." *Oncotarget* 6 (27): 23890-904.
1005 https://doi.org/10.18632/oncotarget.4341. https://www.ncbi.nlm.nih.gov/pubmed/26287601.

1006 Qiu, G. Z., M. Z. Jin, J. X. Dai, W. Sun, J. H. Feng, and W. L. Jin. 2017. "Reprogramming of the
1007 Tumor in the Hypoxic Niche: The Emerging Concept and Associated Therapeutic Strategies."
1008 *Trends Pharmacol Sci* 38 (8): 669-686. https://doi.org/10.1016/j.tips.2017.05.002.
1009 https://www.ncbi.nlm.nih.gov/pubmed/28602395.

1010 Ranganathan, A. C., L. Zhang, A. P. Adam, and J. A. Aguirre-Ghiso. 2006. "Functional coupling of
1011 p38-induced up-regulation of BiP and activation of RNA-dependent protein kinase-like
1012 endoplasmic reticulum kinase to drug resistance of dormant carcinoma cells." *Cancer Res* 66
1013 (3): 1702-11. https://doi.org/10.1158/0008-5472.CAN-05-3092.
1014 https://www.ncbi.nlm.nih.gov/pubmed/16452230.

1015 Rebhandl, S., M. Huemer, F. J. Gassner, N. Zaborsky, D. Hebenstreit, K. Catakovic, E. M.
1016 Grössinger, R. Greil, and R. Geisberger. 2014. "APOBEC3 signature mutations in chronic
1017 lymphocytic leukemia." *Leukemia* 28 (9): 1929-32. https://doi.org/10.1038/leu.2014.160.
1018 https://www.ncbi.nlm.nih.gov/pubmed/24840555.

1019 Robbiani, D. F., and M. C. Nussenzweig. 2013. "Chromosome translocation, B cell lymphoma, and
1020 activation-induced cytidine deaminase." *Annu Rev Pathol* 8: 79-103.
1021 https://doi.org/10.1146/annurev-pathol-020712-164004.
1022 https://www.ncbi.nlm.nih.gov/pubmed/22974238.

1023 Roberts, S. A., M. S. Lawrence, L. J. Klimczak, S. A. Grimm, D. Fargo, P. Stojanov, A. Kiezun, G.
1024 V. Kryukov, S. L. Carter, G. Saksena, S. Harris, R. R. Shah, M. A. Resnick, G. Getz, and D.
1025 A. Gordenin. 2013. "An APOBEC cytidine deaminase mutagenesis pattern is widespread in
1026 human cancers." *Nat Genet* 45 (9): 970-6. https://doi.org/10.1038/ng.2702.
1027 https://www.ncbi.nlm.nih.gov/pubmed/23852170.

1028 Rosenthal, R., N. McGranahan, J. Herrero, B. S. Taylor, and C. Swanton. 2016. "DeconstructSigs:
1029 delineating mutational processes in single tumors distinguishes DNA repair deficiencies and
1030 patterns of carcinoma evolution." *Genome Biol* 17: 31. https://doi.org/10.1186/s13059-016-
1031 0893-4. https://www.ncbi.nlm.nih.gov/pubmed/26899170.

1032 Sadasivam, S., and J. A. DeCaprio. 2013. "The DREAM complex: master coordinator of cell cycle-
1033 dependent gene expression." *Nat Rev Cancer* 13 (8): 585-95. https://doi.org/10.1038/nrc3556.
1034 https://www.ncbi.nlm.nih.gov/pubmed/23842645.

1035 Semenza, G. L. 2003. "Targeting HIF-1 for cancer therapy." *Nat Rev Cancer* 3 (10): 721-32.
1036 https://doi.org/10.1038/nrc1187. https://www.ncbi.nlm.nih.gov/pubmed/13130303.

1037 Senft, D., and Z. A. Ronai. 2016. "Immunogenic, cellular, and angiogenic drivers of tumor
1038 dormancy--a melanoma view." *Pigment Cell Melanoma Res* 29 (1): 27-42.
1039 https://doi.org/10.1111/pcmr.12432. https://www.ncbi.nlm.nih.gov/pubmed/26514653.

1040 Shen, S., and J. Clairambault. 2020. "Cell plasticity in cancer cell populations." *F1000Res* 9.
1041 https://doi.org/10.12688/f1000research.24803.1.
1042 https://www.ncbi.nlm.nih.gov/pubmed/32595946.

1043 Shivji, M. K., V. N. Podust, U. Hübscher, and R. D. Wood. 1995. "Nucleotide excision repair DNA
1044 synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA."
1045 *Biochemistry* 34 (15): 5011-7. https://doi.org/10.1021/bi00015a012.
1046 https://www.ncbi.nlm.nih.gov/pubmed/7711023.

1047 Sigismund, S., D. Avanzato, and L. Lanzetti. 2018. "Emerging functions of the EGFR in cancer."
1048 *Mol Oncol* 12 (1): 3-20. https://doi.org/10.1002/1878-0261.12155.
1049 https://www.ncbi.nlm.nih.gov/pubmed/29124875.

1050 Simó-Riudalbas, L., M. Pérez-Salvia, F. Setien, A. Villanueva, C. Moutinho, A. Martínez-Cardús, S.
1051 Moran, M. Berdasco, A. Gomez, E. Vidal, M. Soler, H. Heyn, A. Vaquero, C. de la Torre, S.
1052 Barceló-Batllori, A. Vidal, L. Roz, U. Pastorino, K. Szakszon, G. Borck, C. S. Moura, F.
1053 Carneiro, I. Zondervan, S. Savola, R. Iwakawa, T. Kohno, J. Yokota, and M. Esteller. 2015.
1054 "KAT6B Is a Tumor Suppressor Histone H3 Lysine 23 Acetyltransferase Undergoing
1055 Genomic Loss in Small Cell Lung Cancer." *Cancer Res* 75 (18): 3936-45.
1056 https://doi.org/10.1158/0008-5472.CAN-14-3702.
1057 https://www.ncbi.nlm.nih.gov/pubmed/26208904.

1058 Stenglein, M. D., M. B. Burns, M. Li, J. Lengyel, and R. S. Harris. 2010. "APOBEC3 proteins
1059 mediate the clearance of foreign DNA from human cells." *Nat Struct Mol Biol* 17 (2): 222-9.
1060 https://doi.org/10.1038/nsmb.1744. https://www.ncbi.nlm.nih.gov/pubmed/20062055.

1061 Swanton, C., N. McGranahan, G. J. Starrett, and R. S. Harris. 2015. "APOBEC Enzymes: Mutagenic
1062 Fuel for Cancer Evolution and Heterogeneity." *Cancer Discov* 5 (7): 704-12.
1063 https://doi.org/10.1158/2159-8290.CD-15-0344.
1064 https://www.ncbi.nlm.nih.gov/pubmed/26091828.

1065 Tate, J. G., S. Bamford, H. C. Jubb, Z. Sondka, D. M. Beare, N. Bindal, H. Boutselakis, C. G. Cole,
1066 C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S. C. Jupe, C. Y. Kok, K. Noble, L.
1067 Ponting, C. C. Ramshaw, C. E. Rye, H. E. Speedy, R. Stefancsik, S. L. Thompson, S. Wang,
1068 S. Ward, P. J. Campbell, and S. A. Forbes. 2019. "COSMIC: the Catalogue Of Somatic
1069 Mutations In Cancer." *Nucleic Acids Res* 47 (D1): D941-D947.
1070 https://doi.org/10.1093/nar/gky1015. https://www.ncbi.nlm.nih.gov/pubmed/30371878.

1071 Temko, D., I. P. M. Tomlinson, S. Severini, B. Schuster-Böckler, and T. A. Graham. 2018. "The
1072 effects of mutational processes and selection on driver mutations across cancer types." *Nat*
1073 *Commun* 9 (1): 1857. https://doi.org/10.1038/s41467-018-04208-6.
1074 https://www.ncbi.nlm.nih.gov/pubmed/29748584.

1075 Teng, M. W., J. B. Swann, C. M. Koebel, R. D. Schreiber, and M. J. Smyth. 2008. "Immune-
1076 mediated dormancy: an equilibrium with cancer." *J Leukoc Biol* 84 (4): 988-93.
1077 https://doi.org/10.1189/jlb.1107774. https://www.ncbi.nlm.nih.gov/pubmed/18515327.

1078 Theodoropoulos, V. E., A. C.h Lazaris, F. Sofras, I. Gerzelis, V. Tsoukala, I. Ghikonti, K. Manikas,
1079 and I. Kastriotis. 2004. "Hypoxia-inducible factor 1 alpha expression correlates with
1080 angiogenesis and unfavorable prognosis in bladder cancer." *Eur Urol* 46 (2): 200-8.
1081 https://doi.org/10.1016/j.eururo.2004.04.008.
1082 https://www.ncbi.nlm.nih.gov/pubmed/15245814.

1083 Tommasini-Ghelfi, S., K. Murnan, F. M. Kouri, A. S. Mahajan, J. L. May, and A. H. Stegh. 2019.
1084 "Cancer-associated mutation and beyond: The emerging biology of isocitrate dehydrogenases
1085 in human disease." *Sci Adv* 5 (5): eaaw4543. https://doi.org/10.1126/sciadv.aaw4543.
1086 https://www.ncbi.nlm.nih.gov/pubmed/31131326.

1087 Tummers, B., and D. R. Green. 2017. "Caspase-8: regulating life and death." *Immunol Rev* 277 (1):
1088 76-89. https://doi.org/10.1111/imr.12541. https://www.ncbi.nlm.nih.gov/pubmed/28462525.

1089 Turcan, S., D. Rohle, A. Goenka, L. A. Walsh, F. Fang, E. Yilmaz, C. Campos, A. W. Fabius, C. Lu,
1090 P. S. Ward, C. B. Thompson, A. Kaufman, O. Guryanova, R. Levine, A. Heguy, A. Viale, L.
1091 G. Morris, J. T. Huse, I. K. Mellinghoff, and T. A. Chan. 2012. "IDH1 mutation is sufficient
1092 to establish the glioma hypermethylator phenotype." *Nature* 483 (7390): 479-83.
1093 https://doi.org/10.1038/nature10866. https://www.ncbi.nlm.nih.gov/pubmed/22343889.

1094 Vogelstein, B., D. Lane, and A. J. Levine. 2000. "Surfing the p53 network." *Nature* 408 (6810): 307-
1095 10. https://doi.org/10.1038/35042675. https://www.ncbi.nlm.nih.gov/pubmed/11099028.

1096 Wall, L., F. Burke, J. F. Smyth, and F. Balkwill. 2003. "The anti-proliferative activity of interferon-
1097 gamma on ovarian cancer: in vitro and in vivo." *Gynecol Oncol* 88 (1 Pt 2): S149-51.
1098 https://doi.org/10.1006/gyno.2002.6707. https://www.ncbi.nlm.nih.gov/pubmed/12586108.

1099 Wang, F. X., J. Huang, H. Zhang, and X. Ma. 2008. "APOBEC3G upregulation by alpha interferon
1100 restricts human immunodeficiency virus type 1 infection in human peripheral plasmacytoid
1101 dendritic cells." *J Gen Virol* 89 (Pt 3): 722-730. https://doi.org/10.1099/vir.0.83530-0.
1102 https://www.ncbi.nlm.nih.gov/pubmed/18272764.

1103 Wang, H. F., S. S. Wang, M. C. Huang, X. H. Liang, Y. J. Tang, and Y. L. Tang. 2019. "Targeting
1104 Immune-Mediated Dormancy: A Promising Treatment of Cancer." *Front Oncol* 9: 498.
1105 https://doi.org/10.3389/fonc.2019.00498. https://www.ncbi.nlm.nih.gov/pubmed/31297335.

1106 Wang, S., M. Jia, Z. He, and X. S. Liu. 2018. "APOBEC3B and APOBEC mutational signature as
1107 potential predictive markers for immunotherapy response in non-small cell lung cancer."
1108 *Oncogene* 37 (29): 3924-3936. https://doi.org/10.1038/s41388-018-0245-9.
1109 https://www.ncbi.nlm.nih.gov/pubmed/29695832.

1110 Wang, Shixiang, Huimin Li, Minfang Song, Zaoke He, Tao Wu, Xuan Wang, Ziyu Tao, Kai Wu, and
1111 Xue-Song Liu. 2020. "Copy number signature analyses in prostate cancer reveal distinct
1112 etiologies and clinical outcomes." *medRxiv*: 2020.04.27.20082404.
1113 https://doi.org/10.1101/2020.04.27.20082404.
1114 http://medrxiv.org/content/early/2020/04/29/2020.04.27.20082404.abstract.

1115 Watnick, R. S., R. K. Rodriguez, S. Wang, A. L. Blois, A. Rangarajan, T. Ince, and R. A. Weinberg.
1116 2015. "Thrombospondin-1 repression is mediated via distinct mechanisms in fibroblasts and
1117 epithelial cells." *Oncogene* 34 (22): 2823-35. https://doi.org/10.1038/onc.2014.228.
1118 https://www.ncbi.nlm.nih.gov/pubmed/25109329.

1119 Wherry, E. J., and M. Kurachi. 2015. "Molecular and cellular insights into T cell exhaustion." *Nat*
1120 *Rev Immunol* 15 (8): 486-99. https://doi.org/10.1038/nri3862.
1121 https://www.ncbi.nlm.nih.gov/pubmed/26205583.

1122 White, D. E., N. A. Kurpios, D. Zuo, J. A. Hassell, S. Blaess, U. Mueller, and W. J. Muller. 2004.
1123 "Targeted disruption of beta1-integrin in a transgenic mouse model of human breast cancer
1124 reveals an essential role in mammary tumor induction." *Cancer Cell* 6 (2): 159-70.
1125 https://doi.org/10.1016/j.ccr.2004.06.025. https://www.ncbi.nlm.nih.gov/pubmed/15324699.

1126 Winter, S. C., F. M. Buffa, P. Silva, C. Miller, H. R. Valentine, H. Turley, K. A. Shah, G. J. Cox, R.
1127 J. Corbridge, J. J. Homer, B. Musgrove, N. Slevin, P. Sloan, P. Price, C. M. West, and A. L.
1128 Harris. 2007. "Relation of a hypoxia metagene derived from head and neck cancer to
1129 prognosis of multiple cancers." *Cancer Res* 67 (7): 3441-9. https://doi.org/10.1158/0008-
1130 5472.CAN-06-3322. https://www.ncbi.nlm.nih.gov/pubmed/17409455.

1131 Yamazaki, H., K. Shirakawa, T. Matsumoto, S. Hirabayashi, Y. Murakawa, M. Kobayashi, A. D.
1132 Sarca, Y. Kazuma, H. Matsui, W. Maruyama, H. Fukuda, R. Shirakawa, K. Shindo, M. Ri, S.
1133 Iida, and A. Takaori-Kondo. 2019. "Endogenous APOBEC3B Overexpression Constitutively
1134 Generates DNA Substitutions and Deletions in Myeloma Cells." *Sci Rep* 9 (1): 7122.
1135 https://doi.org/10.1038/s41598-019-43575-y.
1136 https://www.ncbi.nlm.nih.gov/pubmed/31073151.

1137 Ye, D., S. Ma, Y. Xiong, and K. L. Guan. 2013. "R-2-hydroxyglutarate as the key effector of IDH
1138 mutations promoting oncogenesis." *Cancer Cell* 23 (3): 274-6.
1139 https://doi.org/10.1016/j.ccr.2013.03.005. https://www.ncbi.nlm.nih.gov/pubmed/23518346.

1140 Yi, J. S., M. A. Cox, and A. J. Zajac. 2010. "T-cell exhaustion: characteristics, causes and
1141 conversion." *Immunology* 129 (4): 474-81. https://doi.org/10.1111/j.1365-2567.2010.03255.x.
1142 https://www.ncbi.nlm.nih.gov/pubmed/20201977.

1143 Yuan, S., R. J. Norgard, and B. Z. Stanger. 2019. "Cellular Plasticity in Cancer." *Cancer Discov* 9
1144 (7): 837-851. https://doi.org/10.1158/2159-8290.CD-19-0015.
1145 https://www.ncbi.nlm.nih.gov/pubmed/30992279.

1146 Zhang, J., R. Bajari, D. Andric, F. Gerthoffert, A. Lepsa, H. Nahal-Bose, L. D. Stein, and V. Ferretti.
1147 2019. "The International Cancer Genome Consortium Data Portal." *Nat Biotechnol* 37 (4):
1148 367-369. https://doi.org/10.1038/s41587-019-0055-9.
1149 https://www.ncbi.nlm.nih.gov/pubmed/30877282.

1150 Zhang, T., T. Otevrel, Z. Gao, S. M. Ehrlich, J. Z. Fields, and B. M. Boman. 2001. "Evidence that
1151 APC regulates survivin expression: a possible mechanism contributing to the stem cell origin
1152 of colon cancer." *Cancer Res* 61 (24): 8664-7.
1153 https://www.ncbi.nlm.nih.gov/pubmed/11751382.

1154 Zhao, X., J. Liu, S. Ge, C. Chen, S. Li, X. Wu, X. Feng, Y. Wang, and D. Cai. 2019. "Saikosaponin
1155 A Inhibits Breast Cancer by Regulating Th1/Th2 Balance." *Front Pharmacol* 10: 624.
1156 https://doi.org/10.3389/fphar.2019.00624. https://www.ncbi.nlm.nih.gov/pubmed/31214035.

1157 Zhi, X., J. Tao, K. Xie, Y. Zhu, Z. Li, J. Tang, W. Wang, H. Xu, J. Zhang, and Z. Xu. 2014. "MUC4-
1158 induced nuclear translocation of β-catenin: a novel mechanism for growth, metastasis and

1159 angiogenesis in pancreatic cancer." *Cancer Lett* 346 (1): 104-13.
1160 https://doi.org/10.1016/j.canlet.2013.12.021.
1161 https://www.ncbi.nlm.nih.gov/pubmed/24374017.

1162

1163

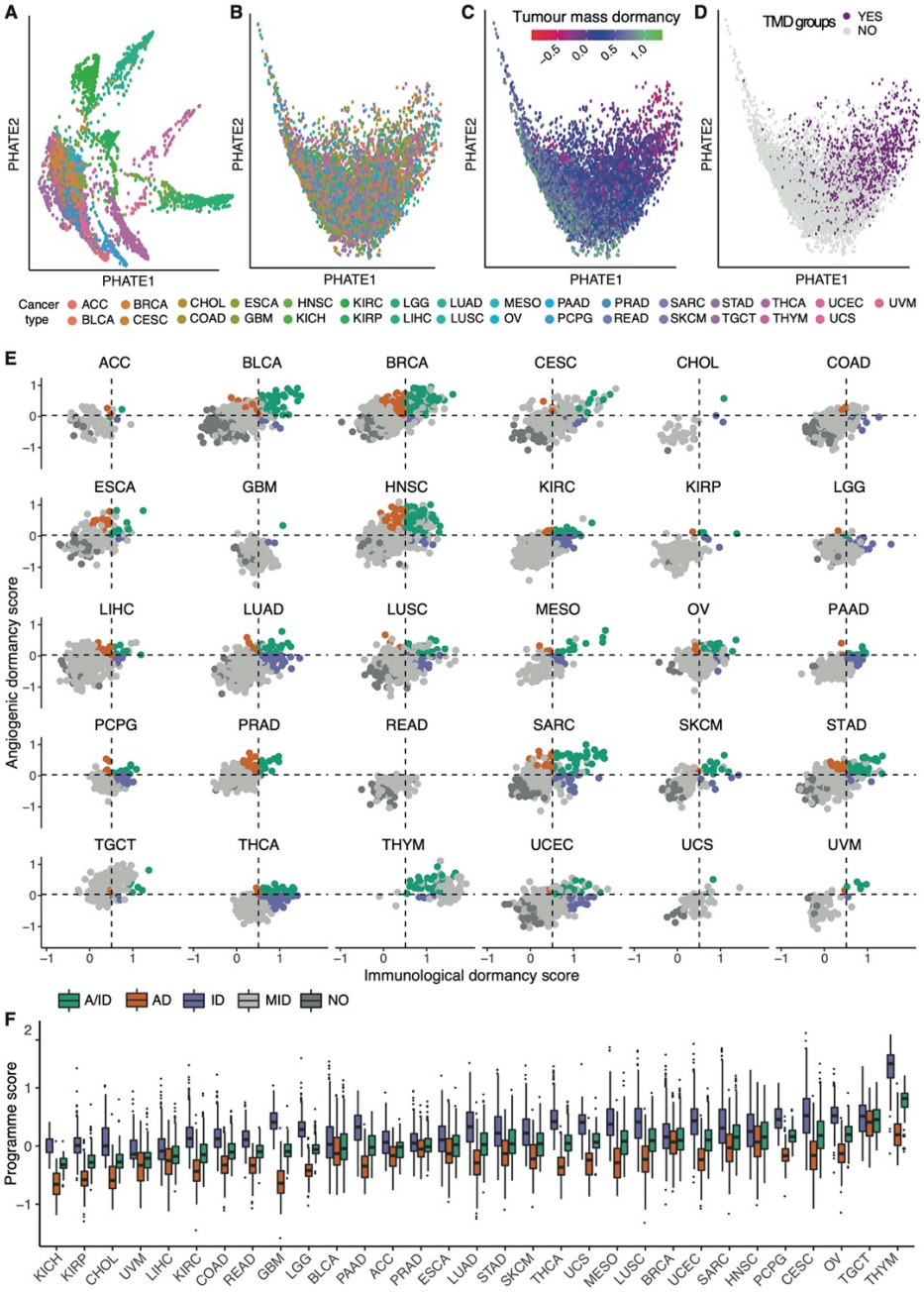

Figure 1. The pan-cancer landscape of TMD. (A-D) PHATE dimensionality reduction applied to 9,631 primary tumor samples based on the expression of genes within the TMD and exhaustion programmes before **(A)** and after **(B-D)** removal of tissue specific expression patterns. The maps are coloured by their corresponding tissue type **(A-B)**, TMD programme score **(C)** and TMD status **(D)**. **(E)** Relationship between immunological and angiogenic programme scores within individual TCGA cancer tissues. Samples are colored by their angiogenic (orange, AD), immunological (purple, ID) and tumour mass (green, A/ID) dormancy status. Samples showing no evidence of TMD (NO) are coloured in dark grey, and slowly expanding tumours in light grey (MID). Horizontal and vertical dashed lines represent the upper quartile of the pan-cancer angiogenic and immunological dormancy programme scores, respectively. KICH was not plotted because it lacked TMD samples. **(F)** Variation in TMD, immunological and angiogenic dormancy scores across primary tumor TCGA samples stratified by tissue type. The tissues are sorted by their TMD levels.

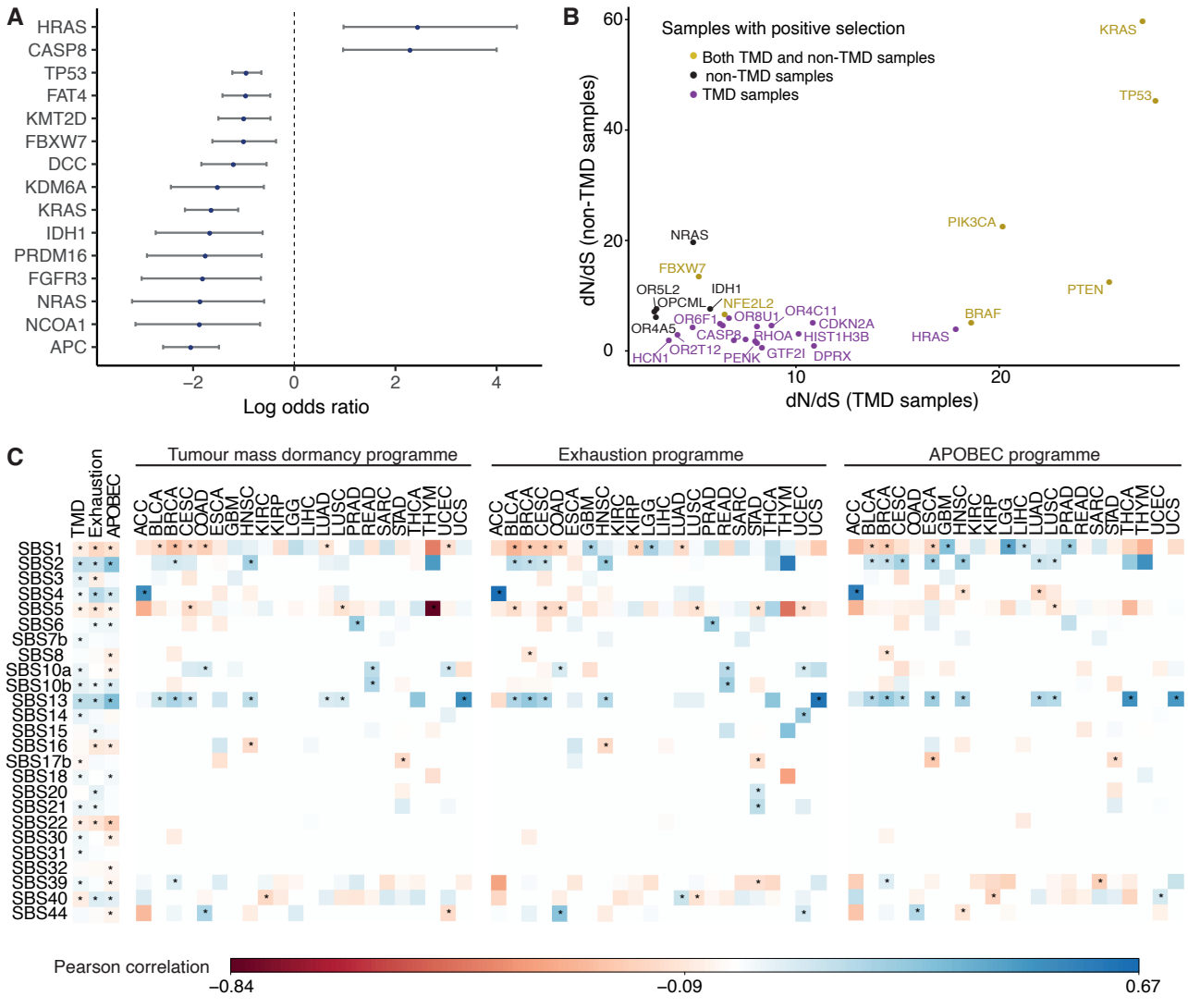

Figure 2. Genomic drivers and mutational processes linked with TMD. (A) Genes presenting an enrichment or depletion of mutations within high TMD samples. Blue circles represent odds ratios on a log2 scale and the confidence intervals for each of the individual Fisher's exact tests are depicted. **(B)** Genes showing signals of positive selection in samples with high (purple) and low (black) TMD, as well as across both groups (yellow). **(C)** Matrices depicting the Person correlation between mutational signatures and the TMD, exhaustion and APOBEC programmes, both pan-cancer (first 3 columns) and within individual cancer tissues. Statistically significant correlations ($p<0.05$) are highlighted with an asterisk. Only signatures with at least one significant correlation are shown.

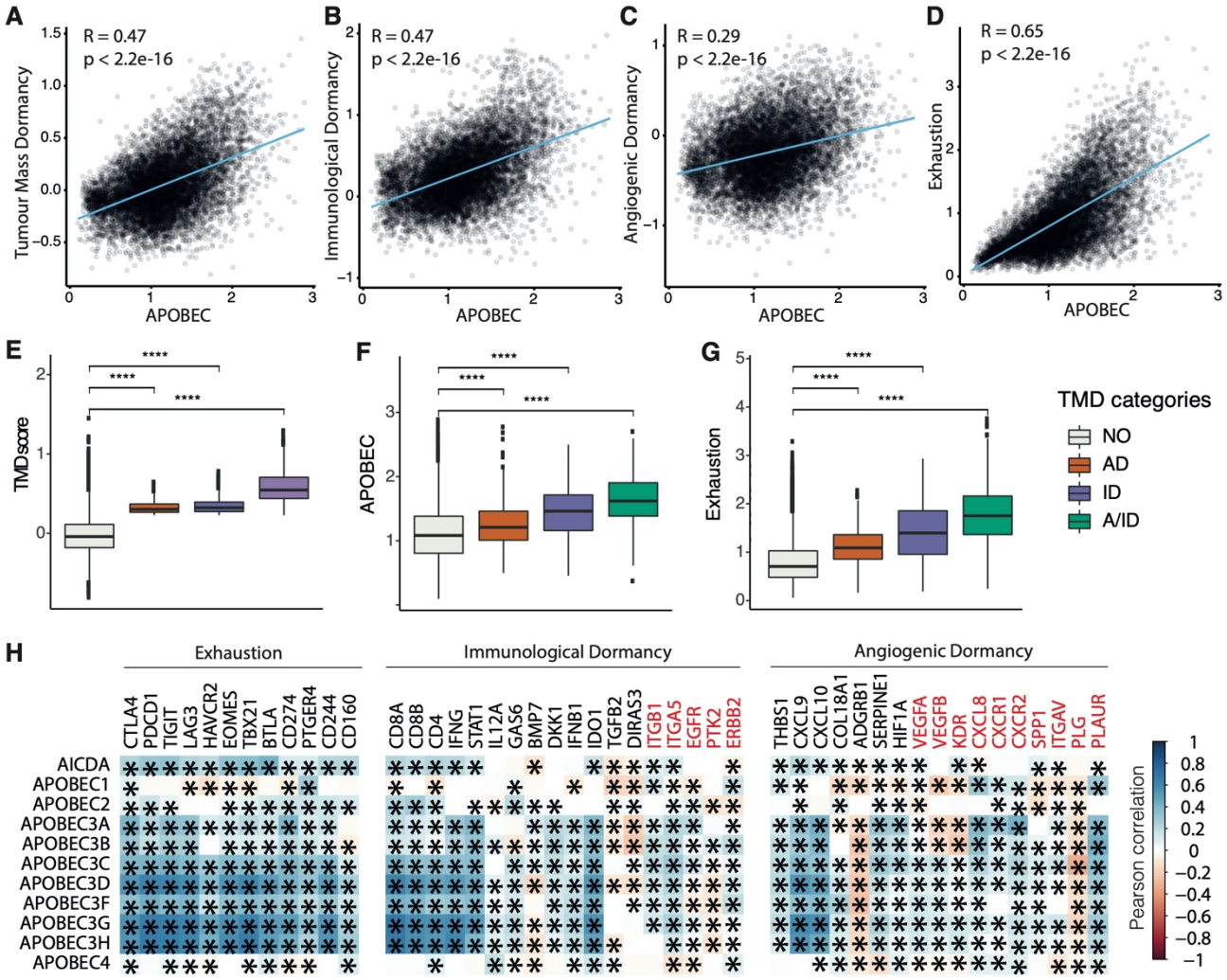

Figure 3. TMD and exhaustion programmes correlate with APOBEC transcriptional activity. (A-D) Pan-cancer correlation between the activity of the APOBEC programme and **(A)** TMD, **(B)** immunological dormancy, **(C)** angiogenic dormancy and **(D)** exhaustion programmes. **(E-G)** TMD, APOBEC and exhaustion programme scores compared between samples with angiogenic dormancy (AD), immunological dormancy (ID), both angiogenic and immunological dormancy (A/ID) and expanding tumours without evidence of TMD (NO); **** p<0.0001. **(H)** Pan-cancer correlation between individual genes in the dormancy and exhaustion programmes and APOBEC/AID enzyme family gene expression. Genes downregulated in the immunological and angiogenic dormancy programmes are shown in red. Significant correlations are marked with an asterisk.

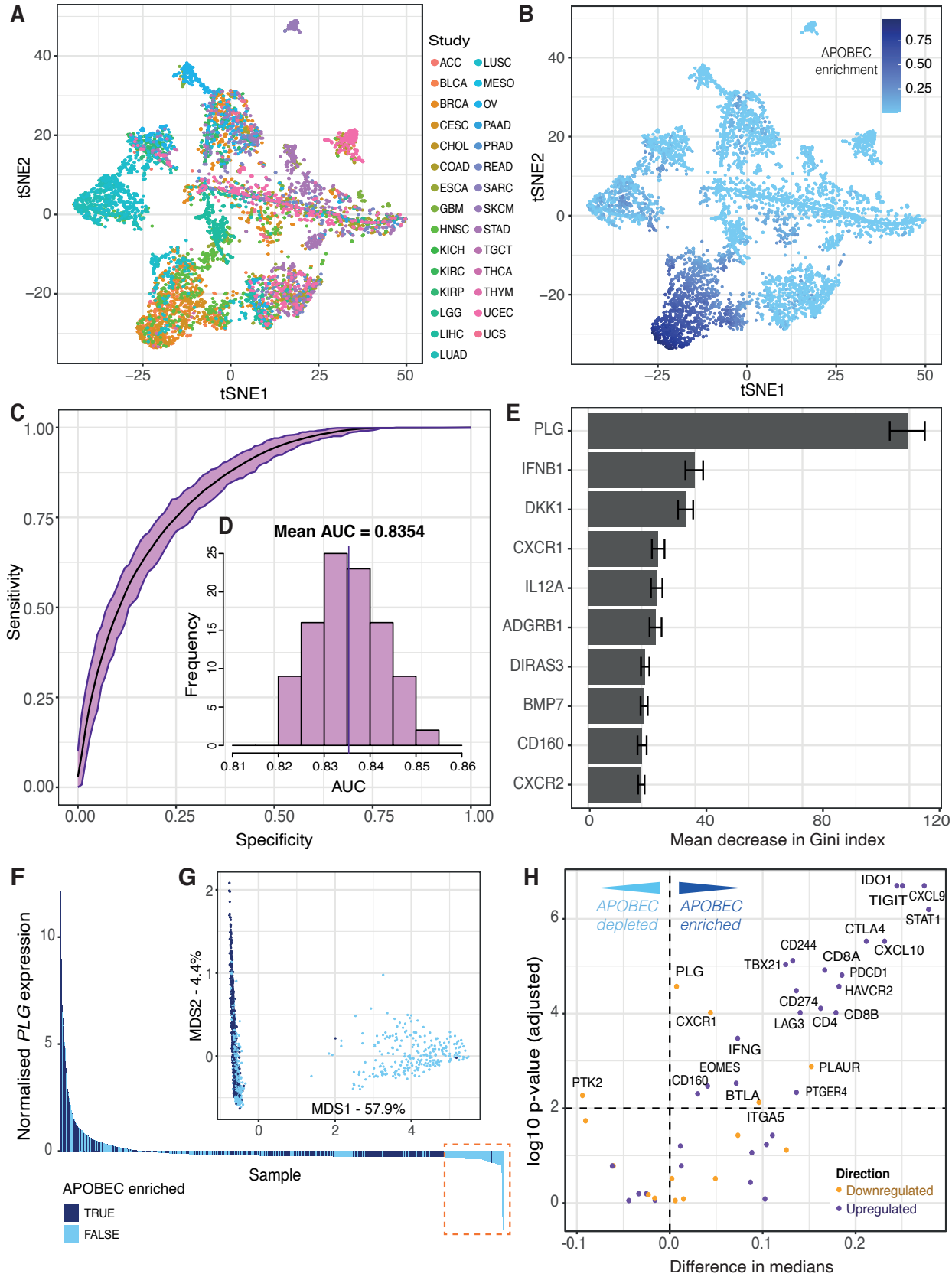

Figure 4. TMD and exhaustion programmes distinguish an APOBEC mutagenesis cluster. tSNE dimensionality reduction of 6,410 primary tumour samples based on single-base substitution signature

1201 profiles labelled by **(A)** TCGA study and **(B)** the total contribution of mutations from APOBEC
1202 signatures SBS2 and SBS13. **(C)** Receiver operating characteristic (ROC) curves of 100 random forest
1203 classifiers of APOBEC signature enrichment based on expression of genes involved in the TMD and
1204 exhaustion programmes. **(D)** Distribution of Area Under Curve (AUC) values across all 100 random
1205 forest classifiers. **(E)** The top 10 ranked genes with highest importance for APOBEC signature
1206 enrichment classification, ranked by mean associated decrease in Gini index across all 100 classifiers.
1207 **(F)** Waterfall plot displaying samples included in an exemplary classifier ranked by normalised *PLG*
1208 expression, coloured by APOBEC enrichment labels. The dashed orange box highlights the subset of
1209 samples with low PLG expression and marked depletion of APOBEC signatures. **(G)** Multidimensional
1210 scaling plot displaying samples included in the exemplary classifier. **(H)** Volcano plot displaying the
1211 difference in median normalised expression for each gene involved in the TMD and exhaustion
1212 programmes between APOBEC-enriched and non-enriched groups, coloured by anticipated direction
1213 of regulation. Differences in medians above 0 correspond to an increase in expression in the APOBEC-
1214 enriched group. Genes with a difference in median <0 would instead be higher expressed in the non-
1215 enriched cluster (APOBEC depleted). A significance cut-off of $p<0.01$ was applied after Benjamini-
1216 Hochberg multiple testing correction.

1217

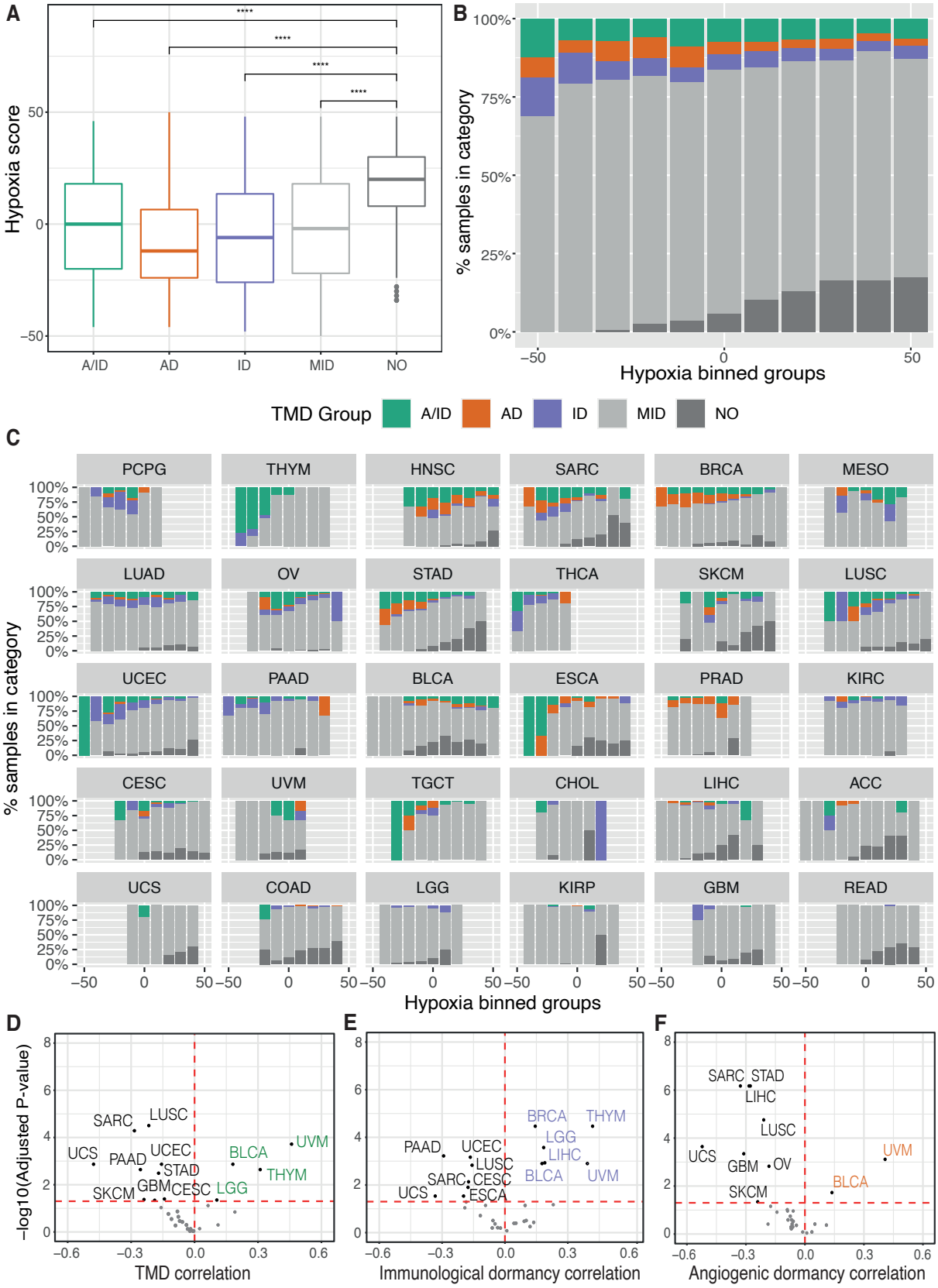

Figure 5. TMD is reduced in the context of hypoxia. (A) Comparison of hypoxia scores across samples falling within distinct TMD categories: both angiogenic and immunological dormancy (A/ID),

1221 angiogenic dormancy (AD), immunological dormancy (ID), slowly expanding tumours (MID),
1222 expanding tumours without evidence of TMD (NO). **** $p<0.0001$. **(B-C)** Distributions of dormancy
1223 status within hypoxia groups, defined by binning the hypoxia scores into intervals of 10: **(B)** pan-
1224 cancer, and **(C)** by TCGA cancer tissue, arranged in descending proportion of dormant samples. KICH
1225 was not plotted because it lacked TMD samples. **(D-F)** Volcano plots displaying the cancer-study-
1226 specific Pearson correlations of **(D)** TMD, **(E)** immunogenic dormancy and **(F)** angiogenic dormancy
1227 programme scores against hypoxia scores. Cancers showing an increased dormancy in the context of
1228 hypoxia are highlighted in corresponding colours, while cancers with high dormancy in normoxic
1229 conditions are depicted in black.

1230

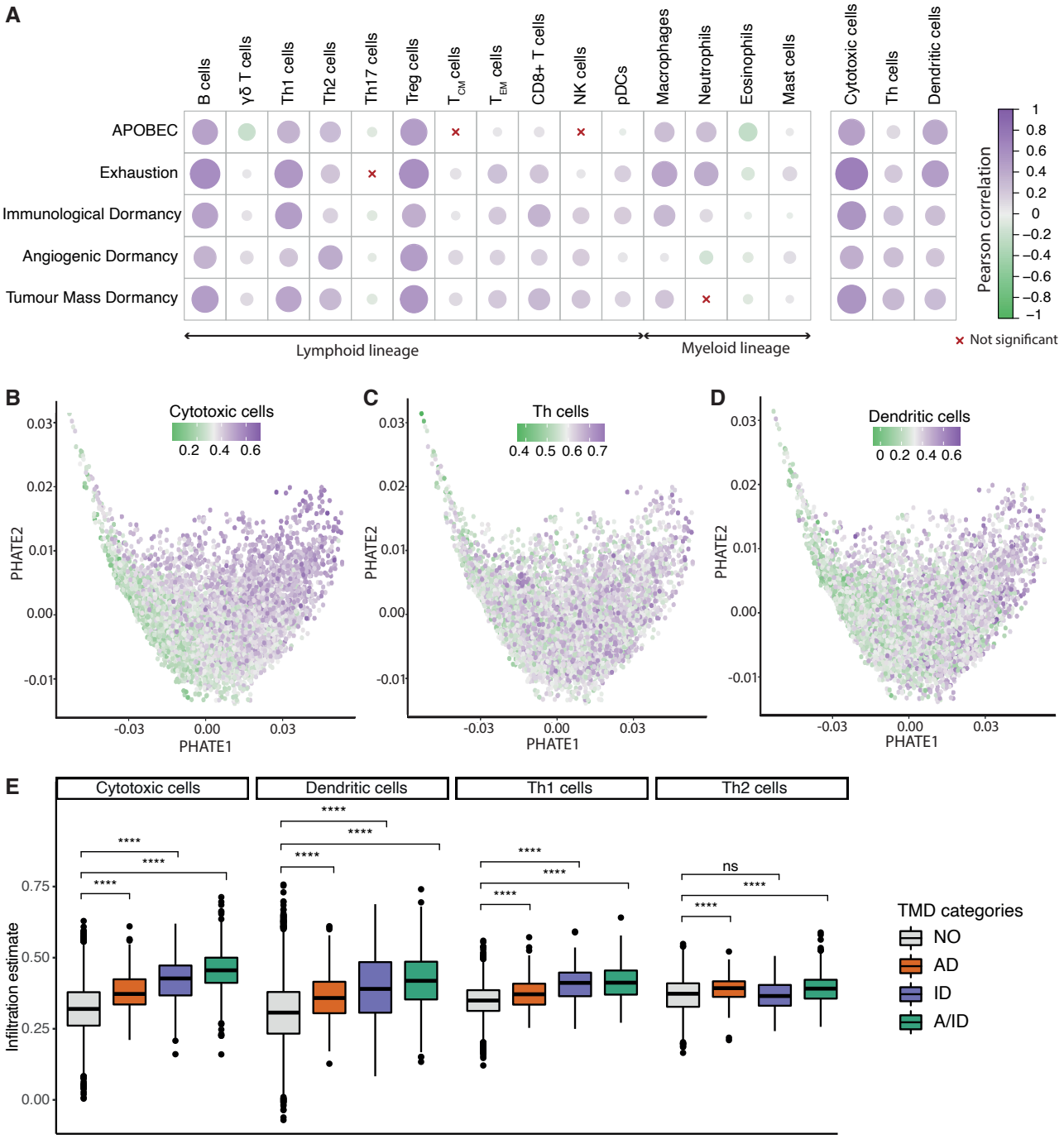

Figure 6. Tumour microenvironment activity correlates with the TMD programme. (A) Correlation between cell infiltration estimates and the APOBEC, exhaustion and dormancy programme scores across the TCGA primary tumor samples. The three broader categories on the right hand side summarise the combined marker expression for all cytotoxic, T helper and dendritic cells. **(B-D)** PHATE dimensionality reduction of 9,631 primary tumor samples, based on the expression of genes driving the TMD and exhaustion programmes, with removal of tissue specific expression patterns and coloured by **(B)** cytotoxic cell enrichment score, **(C)** Th cell enrichment score and **(D)** dendritic cell enrichment score. **(E)** Cytotoxic T cell, Th1/Th2 and dendritic cell abundance compared between samples with angiogenic dormancy (AD), immunological dormancy (ID), both angiogenic and immunological dormancy (A/ID) and samples without TMD (NO); **** $p<0.0001$.

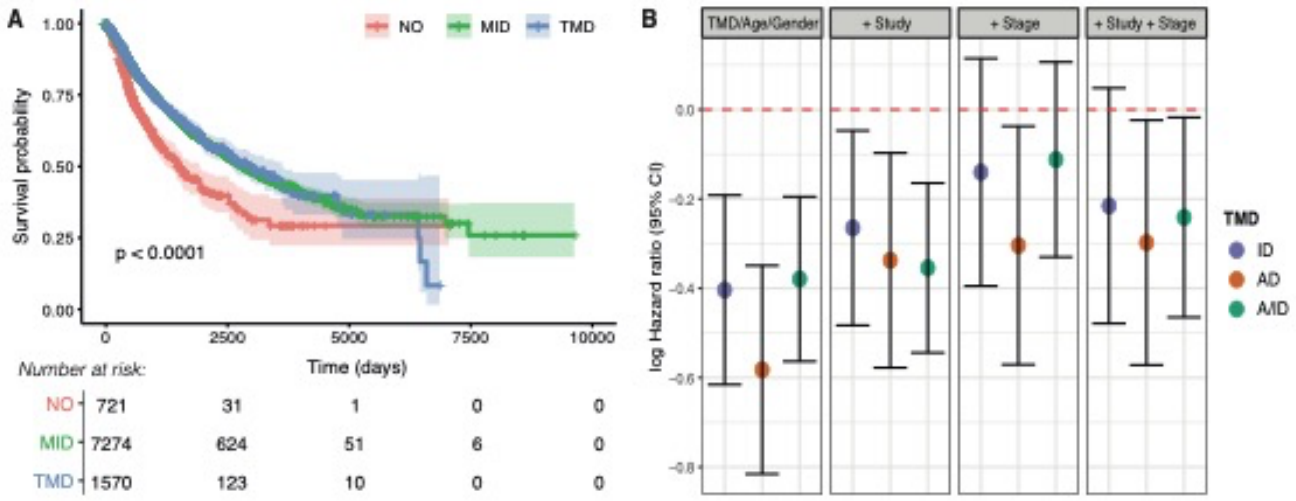

Figure 7. TMD is linked with better prognosis pan-cancer. (A) Overall survival compared between tumours with distinct levels of TMD: strong evidence for TMD (TMD), middle/weak evidence (MID) and no evidence (NO). The number of patients at risk for the corresponding time points are displayed below. **(B)** Log hazard ratios of survival for tumours with angiogenic dormancy (AD), immunological dormancy (ID), both angiogenic and immunological dormancy (A/ID) compared to the 'NO' dormancy category, whilst accounting for age and gender, followed by the inclusion of TCGA study, tumour stage (early/late), and both TCGA study and tumour stage.