



Pan-Cancer Survey of Tumor Mass Dormancy and Underlying Mutational Processes

Anna Julia Wiecek^{1†}, Daniel Hadar Jacobson^{1,2†}, Wojciech Lason^{1†} and Maria Secrier^{1*}

¹ Department of Genetics, Evolution and Environment, UCL Genetics Institute, University College London, London, United Kingdom, ² UCL Cancer Institute, Paul O’Gorman Building, University College London, London, United Kingdom

Tumor 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 characterize. Little is understood of its prevalence across cancer types and of the mutational background that may favor 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 tumor mass dormancy are immunological and angiogenic. To understand the genomic and cellular context in which tumor 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 tumor mutagenesis. We find evidence for immunological and angiogenic dormancy-like signals in 16.5% of bulk sequenced tumors, with a frequency of up to 33% in certain tissues. Mutations in the *CASP8* and *HRAS* oncogenes were positively selected in dormant tumors, 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 aging-induced mutations were relatively depleted in these tumors, while patterns of smoking and defective base excision repair were linked with increased tumor 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. Tumor 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, tumor 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.

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

INTRODUCTION

Tumor 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 (Gerlinger et al., 2014; McGranahan and Swanton, 2017; Temko et al., 2018). Within this rapidly changing context, adaptation of cancer cells is paramount for survival and

OPEN ACCESS

Edited by:

Guang Yao,
The University of Arizona,
United States

Reviewed by:

Joseph William Landry,
Virginia Commonwealth University,
United States
Hai Hu,
Sun Yat-sen Memorial Hospital, China

***Correspondence:**

Maria Secrier
m.secrier@ucl.ac.uk

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

Specialty section:

This article was submitted to
Cell Growth and Division,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 21 April 2021

Accepted: 17 June 2021

Published: 09 July 2021

Citation:

Wiecek AJ, Jacobson DH,
Lason W and Secrier M (2021)
Pan-Cancer Survey of Tumor Mass
Dormancy and Underlying Mutational
Processes.
Front. Cell Dev. Biol. 9:698659.
doi: 10.3389/fcell.2021.698659

much of it is achieved through cellular plasticity (Yuan et al., 2019). As a manifestation of this plasticity, cancer dormancy has emerged as an important contributor to the early stages of tumor development, as well as to cancer progression and metastatic dissemination (Jahanban-Esfahlan et al., 2019; Phan and Croucher, 2020; 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 tumor mass dormancy (TMD), described as an equilibrium between cell proliferation and cell death shaped by the microenvironment that constrains tumor growth (Holmgren et al., 1995; Aguirre-Ghiso, 2007; Wang et al., 2019), are complementary but distinct mechanisms that contribute to the plasticity of cancer expansion (Shen and Clairambault, 2020; Huang, 2021). The former concept had already been coined by 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; Barr et al., 2017; Heldt et al., 2018). The latter, initially termed “population dormancy” by Gimbrone et al. (1972) and then renamed to tumor 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 et al., 2004b), arising as a temporary equilibrium between tumor elimination and immune escape (Koebel et al., 2007; Teng et al., 2008). 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 tumor cells arise (Dunn et al., 2004a). This keeps the tumor in a dormant state (Aguirre-Ghiso, 2007; Wang et al., 2019), hypothesized 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 et al., 2020). During this period of dormancy, the continued cytotoxic response triggered by pro-inflammatory signaling cytokines like interferon γ , as well as prolonged exposure to pro-inflammatory signaling in general, causes the cytotoxic cells to become inactive, a phenomenon known as T cell exhaustion (Dunn et al., 2006; Yi et al., 2010; Wherry and Kurachi, 2015). Finally, the angiogenic switch or immune escape shifts the balance in favor 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 tumor growth by controlling $\beta 1$ integrin signaling which drives a cascade of Ras/ERK mitogenic activation via the focal adhesion kinase (FAK) and the EGF receptor (EGFR) (Aguirre Ghiso et al., 1999; Aguirre Ghiso, 2002; Liu et al., 2002; White et al., 2004). Downregulation of any of these components has been shown to lead to tumor 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 tumor dormancy (Aguirre-Ghiso et al., 2003). This state is further corroborated by a limited ability in recruiting new blood vessels and vasculature remodeling, resulting in angiogenic dormancy (Moserle et al., 2009). This condition is often characterized by VEGF inhibition in the presence of anti-angiogenic factors such as angiostatin, endostatin, thrombospondin, etc., or chemokines (CXCL9 and CXCL10) (Aguirre-Ghiso, 2007; Lyu et al., 2013; Senft and Ronai, 2016; 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 vascularization 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 tumors 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 tumors from 31 tissues available from the Cancer Genome Atlas (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.

RESULTS

Pan-Cancer Characterization of Immunological, Angiogenic, and Tumor 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 programs 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 program 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 section “Materials and Methods,” **Supplementary Figures 2–4**). Both immunological and angiogenic dormancy represent potential mechanisms for restricting the expansion of primary tumor 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 program 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 tumor mass dormancy programs, we next profiled these cellular programs across 9,631 samples of solid primary tumors from TCGA. We observed a spectrum of TMD across tumors that ranges from highly dormant to highly expanding (Figures 1A–C and Supplementary Figures 6, 7). TMD is thought to emerge when tumor cell proliferation is balanced by apoptosis due to factors such as limitations in blood supply or an active immune system (Holmgren et al., 1995). Consequently, primary tumor samples with high TMD program scores (upper quartile of the score range) which also showed a proliferation/apoptosis ratio below 1 (see section “Materials and Methods”), indicative of limited primary tumor lesion expansion, were classed as exhibiting TMD (Figure 1D). Overall, 16.5% of samples across different tissues exhibited substantial evidence for a TMD-like phenotype (with up to 33% prevalence in certain cancers) and they were further subdivided into those indicating angiogenic dormancy (4.4%), immunological dormancy (5%), or both (7%) (Figure 1E and Supplementary Table 2). Head and neck (32%), sarcoma (32%) and breast (26%) cancer were among the cancers with highest rates of TMD. Angiogenic dormancy alone was most widespread in breast (14%) and head and neck cancers (13%). Some rarer cancers like thymoma or pheochromocytoma and paraganglioma also showed a remarkably high prevalence of TMD (~33% each) with the latter also exhibiting the highest immune-mediated dormancy (21%). The systematic differences in TMD program scores across different tissues (Figures 1E,F) suggest that the tissue environment may impact the ability of the tumor to enter a TMD state. Across the board, immune-mediated dormancy levels appeared higher than those of angiogenic dormancy, suggesting a dominant role for immune surveillance in determining TMD.

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

The Genomic Background of TMD

To gain a better understanding of the genomic context in which TMD can develop, we asked whether samples with high and low TMD differed in their association with known drivers of tumorigenesis. We found 15 genes with either a statistically significant enrichment or a depletion of mutations

within samples classed as showing TMD across the 31 solid cancer tissues, with several depletion signals characterizing both early- and late-stage tumors (Fisher’s exact test adjusted $p < 0.05$, Figure 2A, Supplementary Figures 9A, 10A, and Supplementary Table 3). Furthermore, mutations in *MUC4*, a gene involved in angiogenesis and metastasis (Zhi et al., 2014), were specifically enriched in stomach adenocarcinoma, while *EGFR*, *FAT3/4*, *LRP1B*, *KAT6B* mutations, mainly linked with cell proliferation and immune responses (Kato, 2012; Simó-Riudalbas et al., 2015; Sigismund et al., 2018; Chen et al., 2019), were depleted in colon cancer (Fisher’s exact test adjusted $p < 0.05$, Supplementary Figure 11). Many of these findings were reaffirmed using a random forest classification approach (Supplementary Figure 12).

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

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

The genes presenting a depletion of point mutations within dormant samples identified by our analysis have key functions in regulating tumor growth, including *TP53*, the master regulator that coordinates signals of stress such as DNA damage and aberrant growth signaling and can induce cell cycle arrest or apoptosis (Polyak et al., 1997; Vogelstein et al., 2000), or *APC*, which suppresses tumor growth through repression of the Wnt signaling pathway (Boman and Fields, 2013). Both the *APC* and *DCC* genes promote

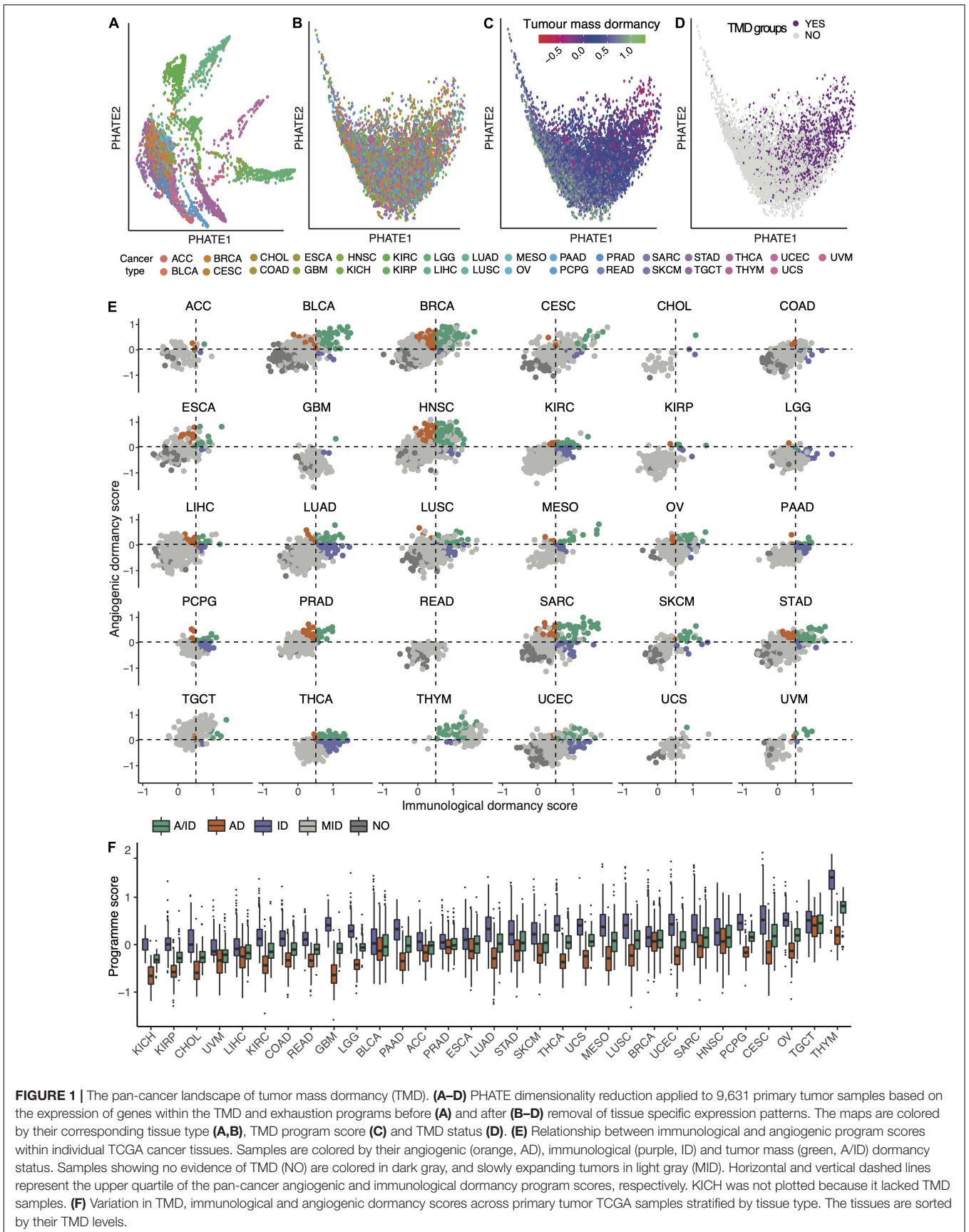


FIGURE 1 | The pan-cancer landscape of tumor mass dormancy (TMD). **(A–D)** PHATE dimensionality reduction applied to 9,631 primary tumor samples based on the expression of the genes within the TMD and exhaustion programs before **(A)** and after **(B–D)** removal of tissue specific expression patterns. The maps are colored by their corresponding tissue type **(A,B)**, TMD program score **(C)** and TMD status **(D)**. **(E)** Relationship between immunological and angiogenic program scores within individual TCGA cancer tissues. Samples are colored by their angiogenic (orange, AD), immunological (purple, ID) and tumor mass (green, A/ID) dormancy status. Samples showing no evidence of TMD (NO) are colored in dark gray, and slowly expanding tumors in light gray (MID). Horizontal and vertical dashed lines represent the upper quartile of the pan-cancer angiogenic and immunological dormancy program 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.

apoptosis by downregulation of survivin gene expression (Zhang et al., 2001) and caspase-9 cleavage (Forcet et al., 2001), respectively. Control of tumor vascularization was reflected in depletion of mutations in *PRDM16*, which inhibits angiogenesis by suppressing the expression of a HIF target semaphorin 5B (Kundu et al., 2020), and in *NCO1A*, a transcriptional coactivator that upregulates the expression of the VEGF α pro-angiogenic factor (Qin et al., 2015). Mutations in the *IDH1* gene, shown to be depleted in samples with TMD, result in the production of the 2-hydroxyglutarate metabolite which regulates the activity of α -ketoglutarate dependent dioxygenases and causes the ubiquitination and proteasomal degradation of *HIF1A*, a key sensor of hypoxia and initiator of angiogenesis (Ye et al., 2013). Moreover, 2-hydroxyglutarate can also result in the downregulation of leukocyte chemotaxis factors (Turcan et al., 2012; Tommasini-Ghelfi et al., 2019), which could contribute to the ability of tumor cells to escape immunological dormancy through immune system evasion.

In addition to the enrichment analysis, using a maximum-likelihood dN/dS method (Martincorena et al., 2017) we detected signals of positive selection for mutations within *CASP8* and *HRAS* in samples with TMD, but not in expanding tumor samples (Figure 2B). In contrast, *IDH1* and *NRAS*, both of which showed a depletion of mutations within TMD samples, showed signals of positive selection only in expanding tumors (Figure 2B). Interestingly, *CASP8* showed similarly increased mutation rate both in early- and late-stage cancers with TMD, while *HRAS* mutations appeared linked with TMD only in late-stage tumors – pointing toward a timing specificity of TMD-linked evolutionary pressures (Supplementary Figure 14). The associations between mutations in these genes and TMD were also validated in an independent dataset from the International Cancer Genome Consortium (ICGC), along with several others including *KRAS* and *TP53* (Supplementary Figure 15). While signals of positive selection are more difficult to validate in the sparser independent cohorts available due to the limited sample size, nevertheless *CASP8* and *TP53* selection signals were robustly recovered both

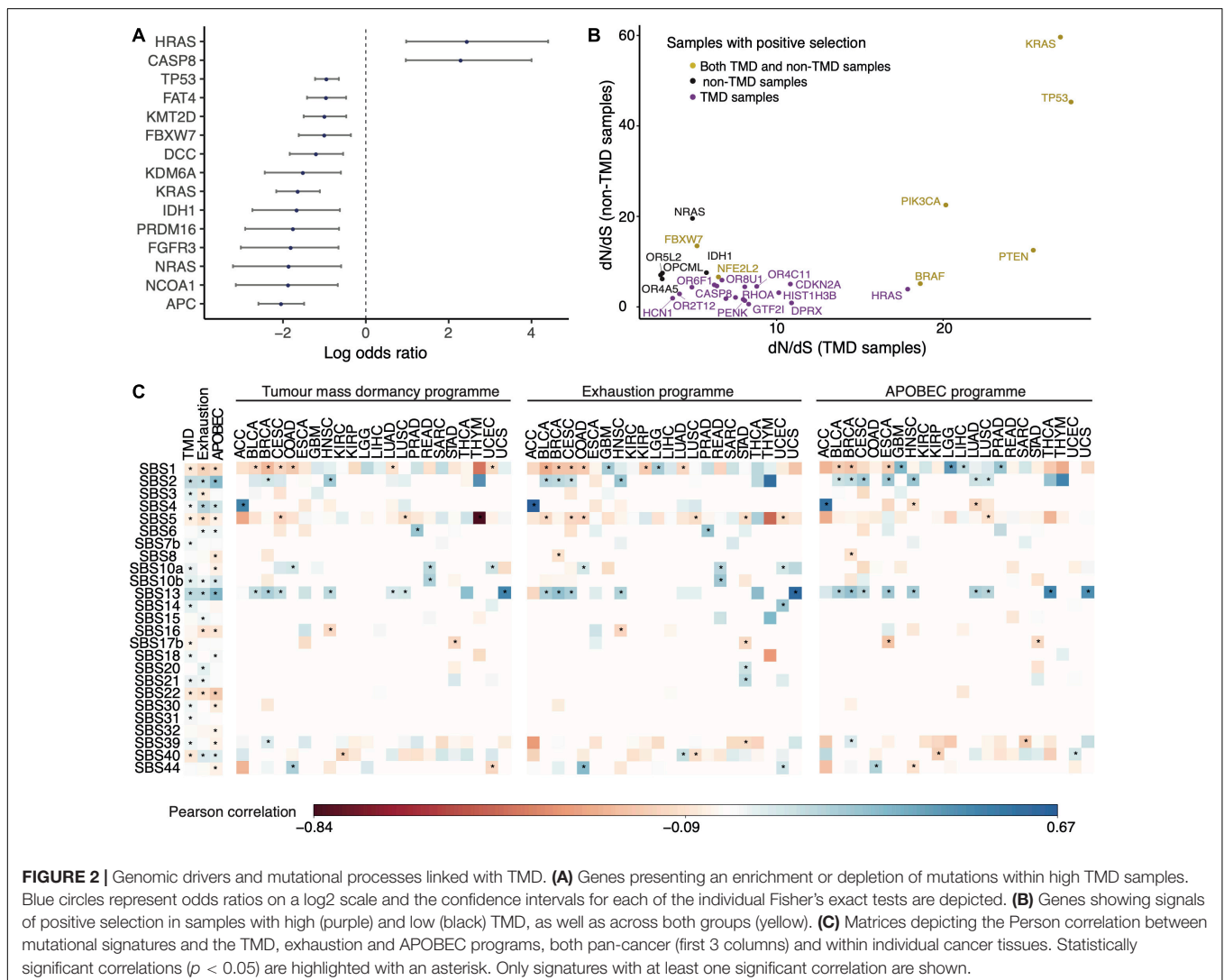


FIGURE 2 | Genomic drivers and mutational processes linked with TMD. **(A)** Genes presenting an enrichment or depletion of mutations within high TMD samples. Blue circles represent odds ratios on a log2 scale and the confidence intervals for each of the individual Fisher’s exact tests are depicted. **(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 Pearson correlation between mutational signatures and the TMD, exhaustion and APOBEC programs, 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.

in oral cancers from ICGC as well as in breast cancers from the METABRIC cohort (**Supplementary Figure 16**). These findings support the importance of *CASP8* and *RAS* mutational status in the context of TMD.

In terms of broader structural variation in the genome, we found no copy number alteration events (amplifications and deletions) specifically enriched in tumors with TMD (data not shown), potentially because such events would be preferentially selected for in fast growing tumors. Moreover, tumors with angiogenic and immunological dormancy showed a modest, but significant decrease in mutational burden when compared to tumors without TMD (**Supplementary Figure 17**).

Mutational Processes Linked With Dormancy and Exhaustion

In addition to investigating the links with specific driver events, we also set out to characterize broader mutational processes associated with TMD. Different risk factors of cancer induce DNA damage in the cells in a context dependent manner, such that nucleotide substitution patterns associated with such mutational processes can be observed within cancer genomes. These patterns of trinucleotide substitutions are termed “mutational signatures” and have been widely characterized across cancers (Alexandrov et al., 2013a, 2020). We carried out a mutational signature analysis to survey the contribution of known mutagenic processes and risk factors to the genomes of dormant tumors. The mutational signature prevalence was correlated with dormancy and exhaustion program scores (**Figure 2C**, left and center panels).

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

Dormancy and Exhaustion Programs Correlate With APOBEC Enzymatic Activity

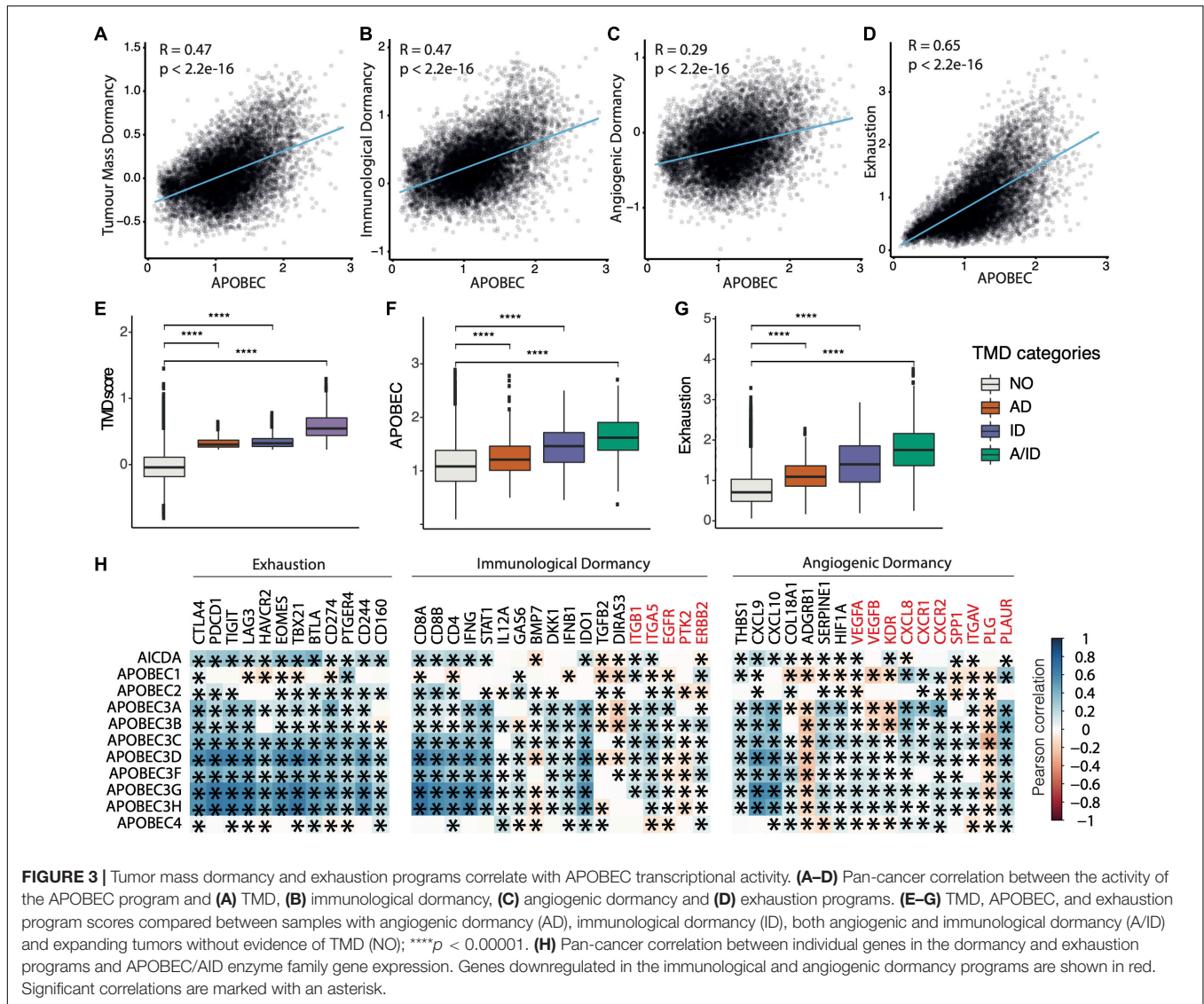
To further explore the association between TMD and exhaustion with APOBEC mutagenesis, we calculated the correlation of our program scores and the mean expression of genes belonging to the APOBEC family (**Figures 3A–F**). Across all cancers and regardless of clinical stage, we observed significant correlations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

We also observed positive correlations between most but not all genes included in the immunological, angiogenic and exhaustion programs, on the one side, and the expression of the AID/APOBEC enzyme family members, on the other (**Figure 3H**). This may be because of downstream regulation of expression, e.g., through mRNA degradation, compartmentalization, or inhibition of transcription. Notably, among all genes, the correlation was highest for the mRNAs of enzymatically active APOBEC3 and AID members of the APOBEC family. The correlation scores were similar for *APOBEC3A-H*, which is to be expected as these genes form a cluster on chromosome 22. Within the TMD program, type II interferon γ and the interferon-upregulated genes such as *STAT1* showed a positive correlation with APOBEC. This is interesting as interferon γ signaling is associated with anti-tumor response by favoring rejection of highly immunogenic tumors (Mittal et al., 2014; Benci et al., 2016). However, prolonged interferon γ signaling promotes epigenetic changes to *STAT1* and the expression of ligands for multiple T cell inhibitory receptors (Benci et al., 2016). Thus, we would expect a spectrum of exhaustion signaling to be present within the tumor, ranging from cells signaling danger to cells with anti-cytotoxic response. Anti-angiogenic markers, such as the urokinase receptor *PLAUR*, which downregulate angiogenic processes and thus restrict the supply of T cells (Oh et al., 2003), were also positively correlated with APOBEC expression. Similar associations were observed for several immune checkpoints (*CD244*, *CD166*, *TIGIT*, *CTLA4*, and *PDCD1*) and exhaustion markers (*EOMES* and *TBX21*).

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

TMD Activity Differences in the Context of APOBEC Mutagenesis

To further elucidate the association between TMD and APOBEC activity, we sought to investigate the specific TMD signals



that might have the strongest links with APOBEC-attributable mutagenesis, as quantified by the single-base substitution mutational signatures SBS2 and SBS13. In order to identify APOBEC-enriched tumors, we employed t-distributed stochastic neighbor embedding (tSNE) to cluster samples based on their overall mutational signature profiles (Figures 4A,B). Unsurprisingly, the clustering was impacted by the tumor tissue of origin, as mutational signatures are often tissue-specific (Degasperi et al., 2020) (Figure 4A). However, samples enriched for APOBEC-associated mutations, defined as the total enrichment of signatures SBS2 and SBS13, covered a fairly heterogeneous set of tissues (Figure 4B). Groups with distinct mutagenesis patterns were defined using expectation-maximization clustering (Supplementary Figure 29), and the procedure was repeated 100 times to obtain robust clusters. Samples were defined as APOBEC-enriched if they fell into the APOBEC-associated cluster more than 50 times (see section “Materials and Methods”), which broadly overlaps with the

APOBEC enrichment score approach developed by Roberts et al. (2013) (Supplementary Figure 30).

To determine whether APOBEC mutagenesis was specifically associated with particular aspects of the TMD program, we used random forest classifiers to rank TMD genes based on how much their expression helps to distinguish between the APOBEC-enriched and depleted groups (see section “Materials and Methods”). The accuracies for predicting APOBEC enrichment across 100 built models were remarkably high and broadly replicable (mean AUC = 0.8354, range: 0.8203–0.8534) (Figures 4C,D). Additionally, the gene ranking in the models was also highly reproducible (Figure 4E). *PLG*, which encodes for the plasminogen protein and triggers angiostatin release to inhibit angiogenesis (Oh et al., 2003), displayed the highest importance by far, defined as the mean associated decrease in Gini index following its removal from a model (mean = 119.2). Upon further investigation of *PLG* expression, we found that its high predictive importance could

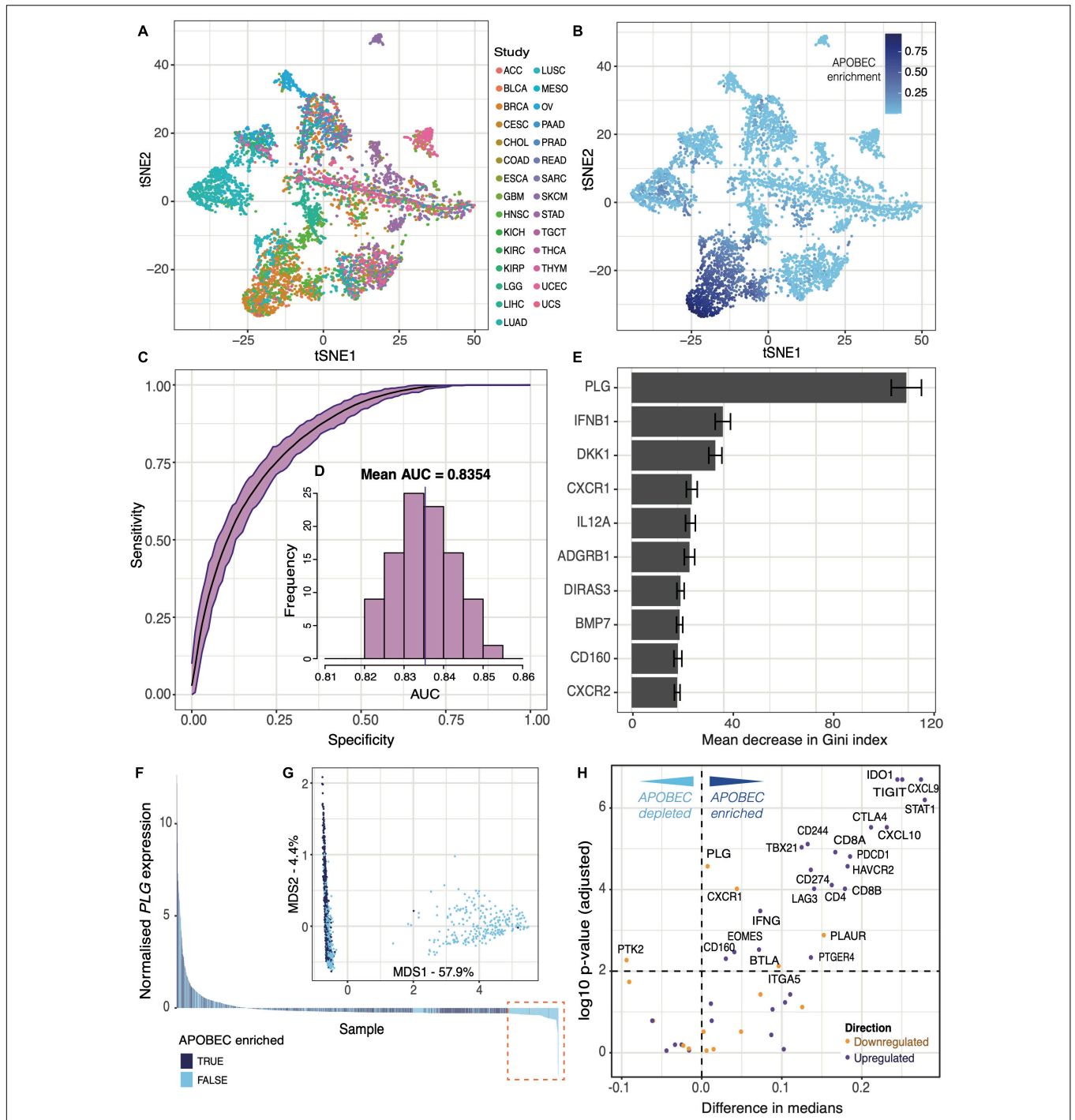


FIGURE 4 | Tumor mass dormancy and exhaustion programs distinguish an APOBEC mutagenesis cluster. tSNE dimensionality reduction of 6,410 primary tumor samples based on single-base substitution signature profiles labeled by (A) TCGA study and (B) the total contribution of mutations from APOBEC signatures SBS2 and SBS13. (C) Receiver operating characteristic (ROC) curves of 100 random forest classifiers of APOBEC signature enrichment based on expression of genes involved in the TMD and exhaustion programs. (D) Distribution of Area Under Curve (AUC) values across all 100 random forest classifiers. (E) The top 10 ranked genes with highest importance for APOBEC signature enrichment classification, ranked by mean associated decrease in Gini index across all 100 classifiers. (F) Waterfall plot displaying samples included in an exemplary classifier ranked by normalized *PLG* expression, colored by APOBEC enrichment labels. The dashed orange box highlights the subset of samples with low *PLG* expression and marked depletion of APOBEC signatures. (G) Multidimensional scaling plot displaying samples included in the exemplary classifier. (H) Volcano plot displaying the difference in median normalized expression for each gene involved in the TMD and exhaustion programs between APOBEC-enriched and non-enriched groups, colored by anticipated direction of regulation. Differences in medians above 0 correspond to an increase in expression in the APOBEC-enriched group. Genes with a difference in medians < 0 would instead be higher expressed in the non-enriched cluster (APOBEC depleted). A significance cut-off of $p < 0.01$ was applied after Benjamini–Hochberg multiple testing correction.

largely be attributed to a substantial group of low-*PLG*-expressing samples which were not labeled as APOBEC-enriched (Figure 4F). This becomes more apparent when considering the multidimensional scaling plot attributable to one of the random forest classifiers, which clearly displays a segregated cluster of non-APOBEC-enriched samples defined by a specific *PLG* expression threshold derived from the classifier (Figure 4G). Therefore, APOBEC mutagenesis appears to be completely lacking in the cases when *PLG* is not expressed, suggesting a link between APOBEC enzymatic activity and the uPA/uPAR system of angiogenic modulation.

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

Heterogeneity of TMD Maintenance in Hypoxic Conditions

Owing to the reported association between hypoxic environments and angiogenesis (Chen et al., 2009), we next sought to investigate whether hypoxia might differentially impact TMD. Hypoxia was quantified on a per-sample basis using established transcriptional signatures (see section “Materials and Methods”), and the calculation procedure was validated using three separate hypoxia signatures (Supplementary Figure 31).

From a pan-cancer perspective, we observed that hypoxia scores were significantly lower in samples displaying any type of dormancy, compared to samples with no evidence of TMD (Figure 5A). This was further corroborated by grouping tumors into distinct subsets (bins) based on their hypoxia levels: the proportion of samples with no evidence of TMD increased significantly as hypoxia scores increased (Figure 5B). In particular, we noted the lowest hypoxia levels appearing in tumors with angiogenic dormancy. These results align with the established consensus that the hypoxia-inducible factor (HIF) pathway is a key regulator of angiogenesis (Pugh and Ratcliffe, 2003; Krock et al., 2011), with angiogenic dormancy expected to develop in normoxic conditions. Similar patterns were also observed on a tissue-specific basis, particularly in the cases of sarcoma, stomach, lung squamous and esophageal carcinomas (Figure 5C). Cancer studies which tended to present low hypoxia, such as pheochromocytoma and paraganglioma (PCPG), thymoma (THYM) and thyroid carcinoma (THCA) often presented some evidence of TMD, whereas studies presenting high hypoxia, such as colon and rectum adenocarcinomas (COAD/READ) and glioblastoma (GBM), were abundant in expanding tumors (Supplementary Figure 32). Notable exceptions included bladder cancer and uveal melanoma, both of which presented samples with evidence

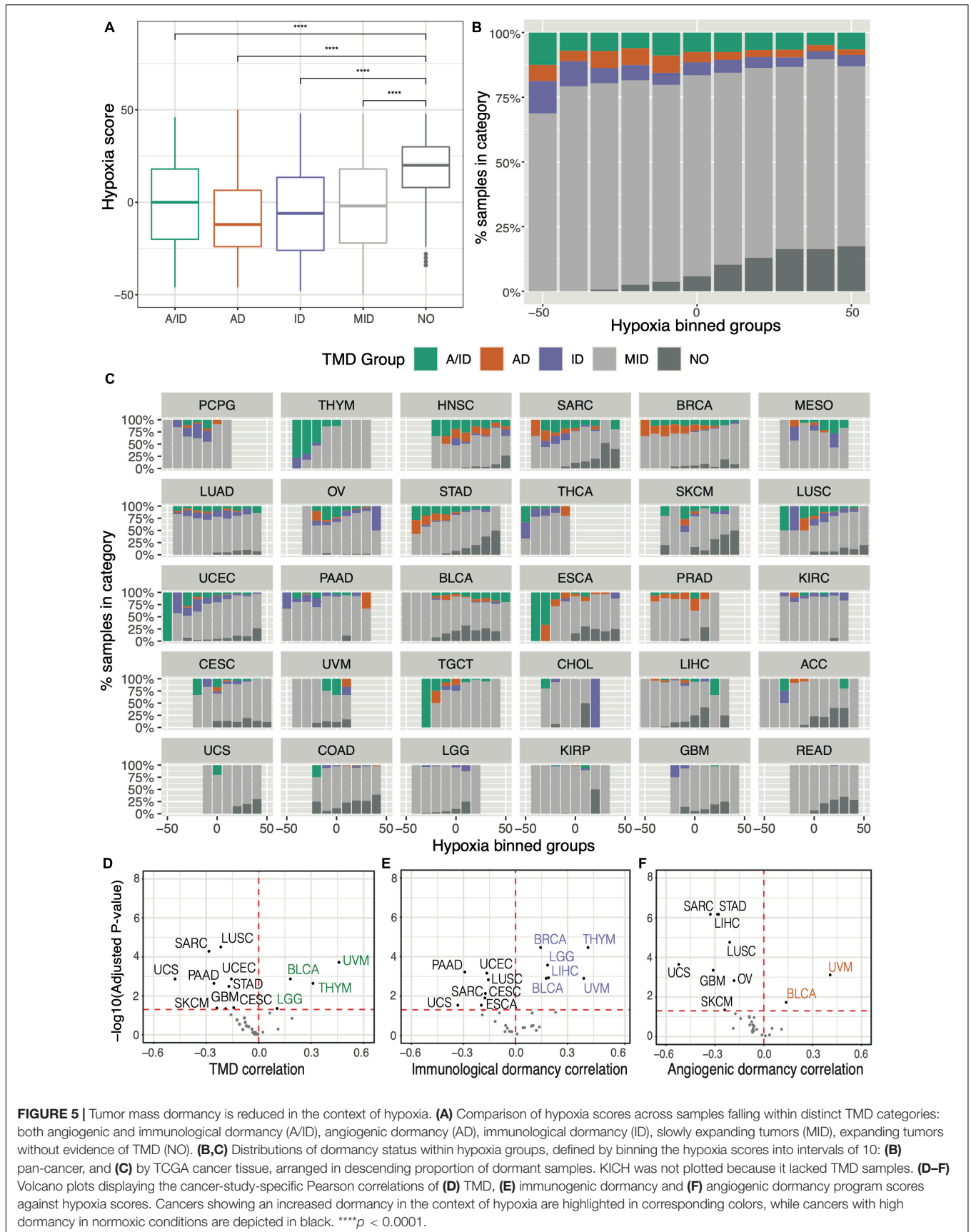
of TMD at the higher end of their respective hypoxia spectra (Figures 5C–F). While prior research has highlighted HIF-1 α expression as a promoter of angiogenesis and tumor invasion in bladder carcinoma (Theodoropoulos et al., 2004) and metastasis in uveal melanoma (Asnaghi et al., 2014), our analysis suggests that a period of TMD-induced latency may also be compatible with hypoxia in these cancers.

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

The Microenvironmental Context of TMD

Since TMD is by definition dependent on immune surveillance, we also set out to confirm this and potentially identify new components of the cancer microenvironment which are permissive to this type of dormancy. Specifically, because the balance in the populations of immune and stromal cells is central to forming anti-tumor responses, we investigated how immune composition might be shaped in the context of TMD, exhaustion and APOBEC programs. Using cell type-specific transcriptional markers (see section “Materials and Methods”), we observed strong positive correlations between TMD and various types of T cells (Figure 6A). Antitumoral cytotoxic cells such as CD8+ and CD4+ T cells, Th1/2 cells, as well as tumor promoting regulatory cells and macrophages were correlated with the APOBEC, exhaustion and dormancy programs. This observation is in line with the mechanism proposed in the literature whereby an initial immune response caused by neoantigens presented by the dormant tumors is followed by exhaustion of such signals (Ghorani et al., 2020). In addition to their cytotoxic activity, CD8+ and CD4+ T cells have been previously shown to limit tumor growth through the secretion of anti-proliferative cytokines, such as interferon γ which can stimulate the expression of p21 and p27 cell cycle inhibitors (Chin et al., 1996; Wall et al., 2003), and antiangiogenic chemokines, such as CXCL9 and CXCL10 (Ikeda et al., 2002).

Compared to Th1 and Th2 cells, natural killer (NK) cells showed a weaker correlation with dormancy programs, consistent with reports of tumor growth control by the immune system being mostly associated with adaptive T cell responses (Finn, 2006). The microenvironment of samples with high TMD was also depleted in inflammatory Th17 cells, which secrete the angiogenesis inducer IL-17A (Bailey et al., 2014). Interestingly, samples displaying signals of angiogenic dormancy had higher enrichment of Th2 instead of Th1 cells. While the Th1 activity is linked with interferon γ production and tumor suppression, Th2 activity has been implicated in cancer progression (Zhao et al., 2019). Despite the broader heterogeneity of T helper cell signals, they showed a similar overall trend as that of cytotoxic and dendritic cells when projected across the TMD landscape (Figures 6B–D). All these classes except for the Th2 cells appeared most enriched in tumors with evidence of both immunological and angiogenic dormancy (Figure 6E). Similar



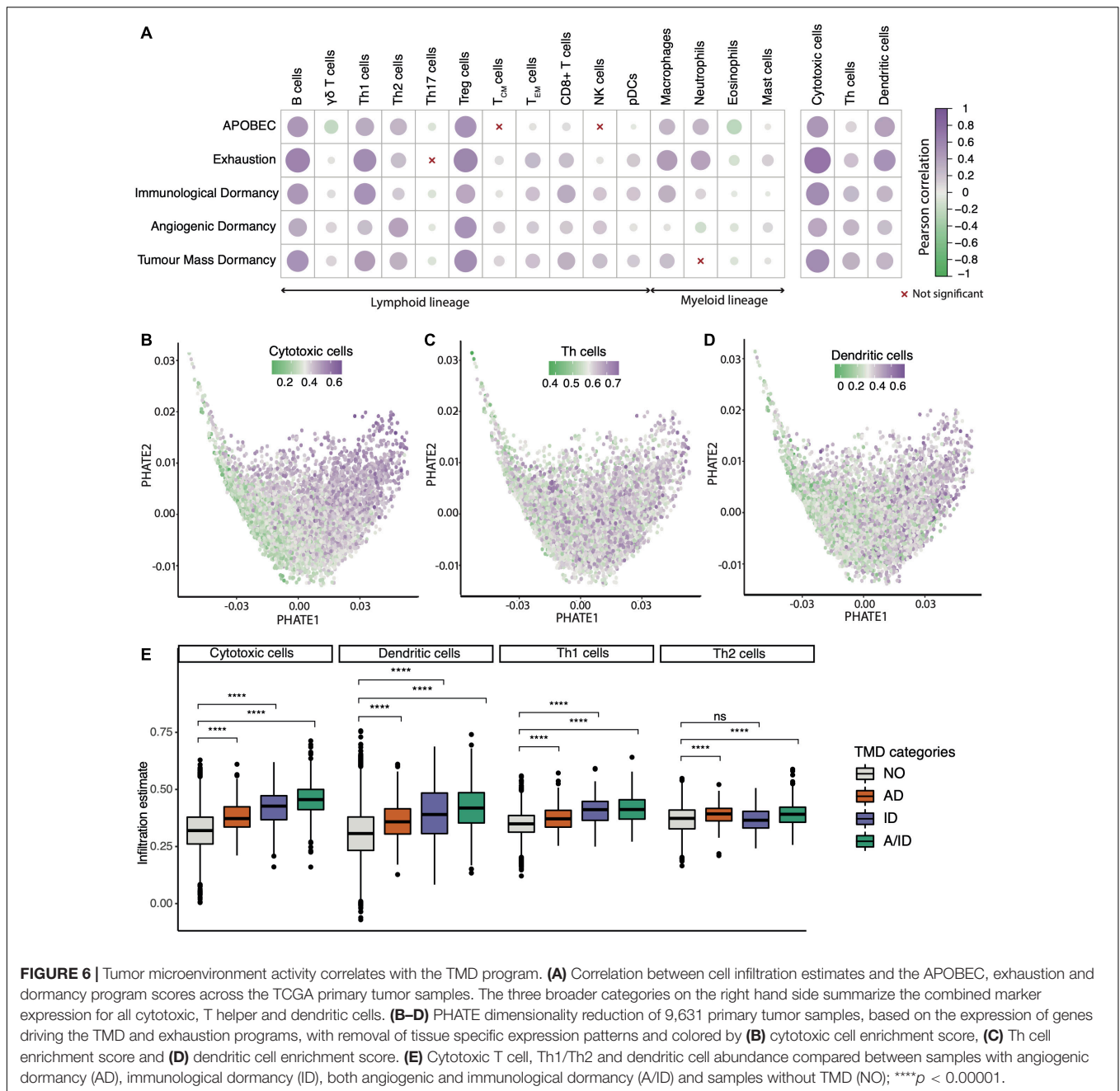


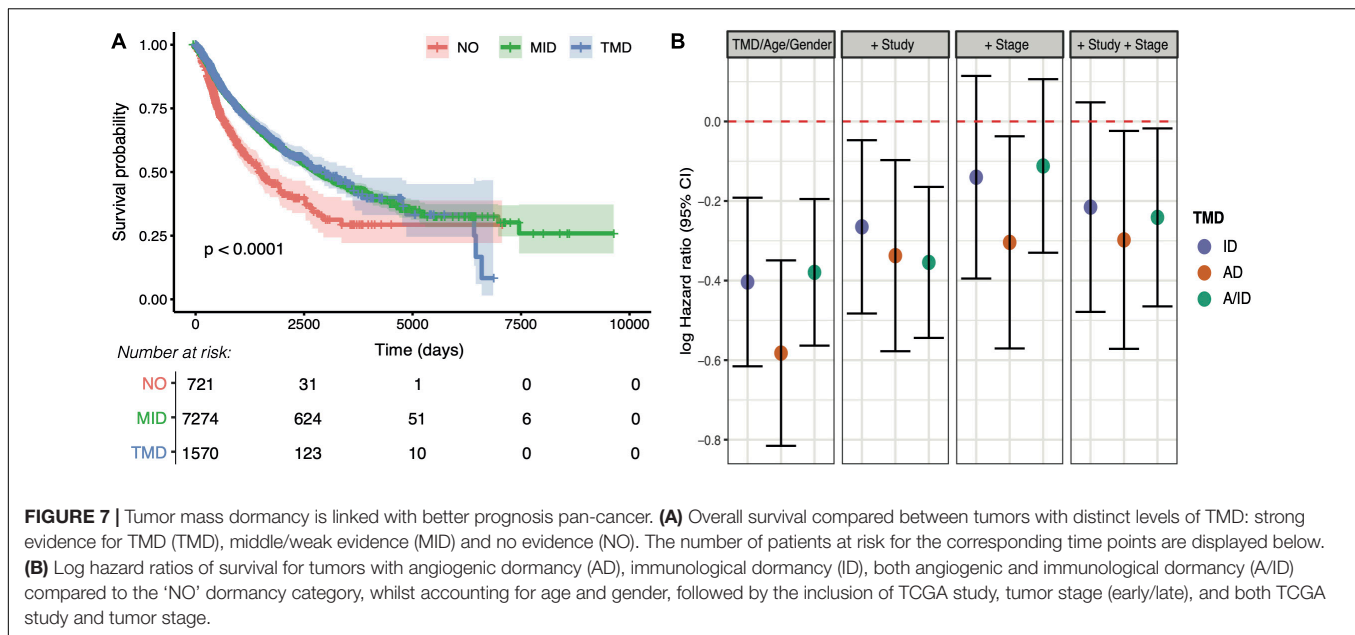
FIGURE 6 | Tumor microenvironment activity correlates with the TMD program. **(A)** Correlation between cell infiltration estimates and the APOBEC, exhaustion and dormancy program scores across the TCGA primary tumor samples. The three broader categories on the right hand side summarize the combined 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 programs, with removal of tissue specific expression patterns and colored 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); **** $p < 0.00001$.

associations were observed for early- and late-stage cancers (**Supplementary Figures 33–36**).

TMD Is Prognostic in Cancer

To understand the clinical relevance of TMD, we carried out survival analysis and found that patients with expanding tumors that presented no evidence of TMD had a significantly reduced prognosis (**Figure 7A**), consistent with the expectation that fast-proliferating tumors should be more aggressive than their dormant counterparts. To further dissect the potential sources of this variation along the immunity/angiogenesis axis, we employed Cox proportional hazards models to determine

the effects of different types of dormancy on survival whilst accounting for potential confounders such as patient age, gender, cancer type, and tumor stage. While all three categories were associated with decreased risk when compared to patients with no evidence of TMD, much of this variation was accounted for by tumor stage, with only angiogenic and tumor mass dormancy still marginally significant after correction (**Figure 7B** and **Supplementary Figures 37–41**). While the tumor stage effect may not be fully uncoupled from the TMD effect, TMD did present a marginal protective role within the early-stage tumors (stages I/II), as well as a larger protective role in late-stage tumors (stages III/IV) (**Supplementary Figure 42**).



DISCUSSION

The TMD is a key state in tumor development and progression, but its evolutionary constraints are largely unknown. In this study, we have presented an overview of TMD across 31 cancer tissues and highlighted potential novel genomic hallmarks of this state. While TMD remains a poorly characterized program whose current known markers are undoubtedly incomplete, our approach to evaluating TMD in bulk tumors through an equilibrium of proliferation, apoptosis, cytotoxicity, exhaustion, and angiogenesis signals has enabled us to systematically capture this temporary state across a variety of cancers. We show that TMD-like signals are present across a multitude of cancer tissues, and that immunological and angiogenic dormancy are not always concomitant. We confirm many of the expected associations reported in the literature, including lower mutational burden, correlations with CD8+, regulatory and helper T cells (Teng et al., 2008), improved prognosis (Park and Nam, 2020) and a strong link with immune exhaustion (Mittal et al., 2014). Exhaustion is a phenotype that likely immediately follows TMD, but in our analysis the two states are largely overlapping. This is most likely due to the bulk nature of the samples analyzed, where we may be capturing parts of the tumor that starting to expand in addition to areas still displaying TMD. Interestingly, we find TMD tends to occur in normoxic rather than hypoxic environments in the majority of cases, which is different than what is observed for cellular quiescence (Qiu et al., 2017; Butturini et al., 2019).

In addition to expected associations, our methodology has also enabled us to capture previously unreported links between TMD and clinical or genomic characteristics. Firstly, while TMD would be expected to dominate in early-stage tumors, we also notice TMD-like signals in a comparable fraction of late-stage tumors (18%) (Supplementary Figure 8). In the absence of longitudinal data that could shed further light onto the causes, we

can only hypothesize this could be due to a mechanism of slowly advancing cancers that are maintained in a semi-TMD state, potentially enabled by a strong, consistent immune response.

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

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

A remarkably consistent signal came from the mutagenesis trace left by the APOBEC enzyme family. APOBEC activity

appeared strongly linked with the exhaustion signals that presumably follow TMD, and with immunological dormancy. Several studies (Burns et al., 2013; Swanton et al., 2015; Yamazaki et al., 2019) have proposed that APOBEC expression may be one of the drivers of tumorigenesis, suggesting a role for APOBECs in evading the immune response. APOBEC3 signatures are prevalent in leukaemias (Rebhandl et al., 2014), while AID was demonstrated to induce hypermutation and chromosomal translocation in B cell lymphomas (Robbiani and Nussenzweig, 2013). On the other hand, APOBEC may promote the formation of neoantigens and activate cytotoxic T cell responses. It was shown that the APOBEC3 class of enzymes is upregulated in immune cells (monocytes, macrophages and pDCs) upon treatment with interferon γ (Wang et al., 2008; Stenglein et al., 2010). In ovarian carcinoma, the expression of APOBEC3G was positively correlated with T cell infiltration, expression of cytotoxic granzyme and perforin (*GZMB* and *PRF1*), and improved clinical outcome (Leonard et al., 2016). Moreover, APOBEC mutagenesis has previously been linked with responses to immunotherapy in lung, melanoma and head and neck cancers (Wang et al., 2018; Faden et al., 2019; Driscoll et al., 2020) through the induction of heteroclitic neoepitopes. We have demonstrated increased APOBEC mutagenesis in a TMD setting across multiple tumor types, and have shown that this process appears linked with the regulation of angiogenesis by the uPAR system. It is possible that the mutational footprint left in the genome by such activity enables immune recognition and maintenance of TMD in the time span until exhaustion and immune escape occurs. This could suggest a window of opportunity for the employment of checkpoint inhibitors, possibly in combination with APOBEC inhibitors, in early-stage tumors with TMD.

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

An important limitation of this analysis consists in the fact that there are no currently available datasets describing transcriptional changes specifically in the context of tumor mass dormancy. This phenotype is difficult to study due to its temporary nature and few model systems for this state exist. We have therefore not been able to unequivocally determine whether the chosen scoring method and even the cut-offs employed for assessing TMD capture this phenomenon in the most accurate manner. However, we are taking a conservative approach focusing only on the most extreme values, which are most likely

to reflect this state. This approach appears methodologically robust to small fluctuations in expression/gene signatures, and the associations identified have been validated in independent datasets. Moreover, the key findings of *CASP8* positive selection in TMD, associations with APOBEC mutagenesis and cytotoxic microenvironment were also reproducible using the PCA-based scoring methodology (**Supplementary Figures 43–48**). Further experimental evidence is needed to verify the biological nature of these associations and the most appropriate way of evaluating TMD, once experimental model systems for this state will have become more widely available and easier to study.

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

Furthermore, our study focused mainly on pan-cancer markers of TMD, due to the scarcity of this phenotype. From our analyses, we observe TMD is frequent in certain tissues such as head and neck, breast, bone, and has weak evidence or complete absence of signals in others, e.g., chromophobe renal 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.

MATERIALS AND METHODS

Datasets

FPKM normalized 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 tumor 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 (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 and 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 Adenocarcinoma (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 analyzed in the same manner as the TCGA data.

Quantifying the Dormancy, Exhaustion and APOBEC Programs

Mean log-transformed expression values of genes deemed to be associated with a given program by manual curation of the literature (**Supplementary Table 1**) were used to produce per-sample program scores for TMD, immunological/angiogenic dormancy, exhaustion and APOBEC activity. We have specifically selected genes that have been associated with immunological and angiogenic dormancy, 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 program scores were calculated by taking the mean expression of genes within the respective programs, as all genes in the programs were expected to be upregulated in the respective states.

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

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 program from the sum of expression values of upregulated genes (E_u) and dividing by the total number of genes within the program ($N_u + N_d$, where N_u is the total number of upregulated genes and N_d is the total number of downregulated genes):

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

Principal Component Analysis (PCA)

Principal component analysis was performed across the pan-cancer dataset using the up- and downregulated gene signature of the respective program. 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 program 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) program score range, respectively. Samples with no evidence for TMD (expanding tumors, marked as 'NO') were identified as being in the lower quartile of the TMD program 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.

Assessing the Robustness of the AD/ID Program 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 program 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 and 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.

Assessing the Robustness of APOBEC Associations

Correlations between the mean expression of APOBEC-related genes and random gene expression programs were calculated by randomly selecting an equal number of genes to that found in the TMD signatures (35) from the genome and assessing the mean expression correlation with APOBEC activity. 1,000 iterations were performed. Associations varied as shown in **Supplementary Figure 24**, and were consistently

lower than the one calculated for TMD. *P*-values were frequently significant, but this is likely due to the large number of samples.

PHATE Dimensionality Reduction

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

Association Between Cancer Driver Mutations and TMD

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

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

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

Mutational Signature Analysis

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

SigProfiler was used to infer mutational signatures from TCGA whole-exome sequencing data. For each TCGA study

of interest, input mutational matrices were generated using the *SigProfilerMatrixGeneratorR* function using all samples containing at least 50 mutations and aligning the MAF files to the hg38 genome build. *SigProfilerExtractorR* was used to extract *de novo* mutational signatures for each cancer type. For each study, the solution with the greatest number of mutational signatures was chosen, for which also the sum of the solution stability (calculated average silhouette coefficient) and minimum stability exceeded 1, and the minimum stability value did not fall below 0.4. For the selected solutions, the identity of the mutational signatures was determined by calculating cosine similarities with COSMIC v3.1 mutational signatures.

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

Tumor Microenvironment Deconvolution From Bulk RNA-Seq Data

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

Identification of APOBEC Mutagenesis Clusters

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

times and samples which appeared in the APOBEC-enriched cluster on more than 50 occasions were labeled as 'APOBEC enriched.'

Random Forest Modeling of APOBEC Mutagenesis-Gene Expression Associations

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

Genes whose expression contributed to APOBEC enrichment classification were identified using machine learning via a random forest approach, conducted using the *randomForest* R package. In total, 5,850 samples were included for which Z-scores and mutational signature profiles were available. Each random forest model required all samples labeled 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.

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. (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. (2019). A binning approach with a window size of 10 was employed to further subdivide tumors into discrete subgroups with increasing levels of hypoxia.

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.

Statistics

Pairwise correlations were calculated using the Pearson correlation statistics. The *corrplot* R package was used to carry out statistical analysis and visualize 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 $p < 0.05$.

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>.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://portal.gdc.cancer.gov/> (TCGA Genomics Data Commons Data Portal).

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

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

FUNDING

AW and DJ were supported by MRCDTP 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).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.698659/full#supplementary-material>

REFERENCES

- Aguirre Ghiso, J. A. (2002). Inhibition of FAK signaling activated by urokinase receptor induces dormancy in human carcinoma cells in vivo. *Oncogene* 21, 2513–2524. doi: 10.1038/sj.onc.1205342
- Aguirre Ghiso, J. A., Kovalski, K., and Ossowski, L. (1999). Tumor dormancy induced by downregulation of urokinase receptor in human carcinoma involves integrin and MAPK signaling. *J. Cell Biol.* 147, 89–104. doi: 10.1083/jcb.147.1.89
- Aguirre-Ghiso, J. A. (2007). Models, mechanisms and clinical evidence for cancer dormancy. *Nat. Rev. Cancer* 7, 834–846. doi: 10.1038/nrc2256
- Aguirre-Ghiso, J. A., Estrada, Y., Liu, D., and Ossowski, L. (2003). ERK(MAPK) activity as a determinant of tumor growth and dormancy; regulation by p38(SAPK). *Cancer Res.* 63, 1684–1695.
- Alexandrov, L. B., Jones, P. H., Wedge, D. C., Sale, J. E., Campbell, P. J., Nik-Zainal, S., et al. (2015). Clock-like mutational processes in human somatic cells. *Nat. Genet.* 47, 1402–1407. doi: 10.1038/ng.3441
- Alexandrov, L. B., Kim, J., Haradhvala, N. J., Huang, M. N., Tian Ng, A. W., Wu, Y., et al. (2020). The repertoire of mutational signatures in human cancer. *Nature* 578, 94–101. doi: 10.1038/s41586-020-1943-3
- Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., et al. (2013a). Signatures of mutational processes in human cancer. *Nature* 500, 415–421. doi: 10.1038/nature12477
- Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J., and Stratton, M. R. (2013b). Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* 3, 246–259. doi: 10.1016/j.celrep.2012.12.008
- Asnagli, L., Lin, M. H., Lim, K. S., Lim, K. J., Tripathy, A., Wendeborn, M., et al. (2014). Hypoxia promotes uveal melanoma invasion through enhanced Notch and MAPK activation. *PLoS One* 9:e105372. doi: 10.1371/journal.pone.0105372
- Bailey, S. R., Nelson, M. H., Himes, R. A., Li, Z., Mehrotra, S., and Paulos, C. M. (2014). Th17 cells in cancer: the ultimate identity crisis. *Front. Immunol.* 5:276. doi: 10.3389/fimmu.2014.00276
- Barr, A. R., Cooper, S., Heldt, F. S., Butera, F., Stoy, H., Mansfeld, J., et al. (2017). DNA damage during S-phase mediates the proliferation-quiescence decision in the subsequent G1 via p21 expression. *Nat. Commun.* 8:14728. doi: 10.1038/ncomms14728
- Benci, J. L., Xu, B., Qiu, Y., Wu, T. J., Dada, H., Twyman-Saint Victor, C., et al. (2016). Tumor interferon signaling regulates a multigenic resistance program to immune checkpoint blockade. *Cell* 167, 1540–1554.e12. doi: 10.1016/j.cell.2016.11.022
- Bhandari, V., Hoey, C., Liu, L. Y., Lalonde, E., Ray, J., Livingstone, J., et al. (2019). Molecular landmarks of tumor hypoxia across cancer types. *Nat. Genet.* 51, 308–318. doi: 10.1038/s41588-018-0318-2
- Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795. doi: 10.1016/j.immuni.2013.10.003
- Boire, A., Coffelt, S. B., Quezada, S. A., Vander Heiden, M. G., and Weeraratna, A. T. (2019). Tumour dormancy and reawakening: opportunities and challenges. *Trends Cancer* 5, 762–765. doi: 10.1016/j.trecan.2019.10.010
- Boman, B. M., and Fields, J. Z. (2013). An APC:WNT counter-current-like mechanism regulates cell division along the human colonic crypt axis: a mechanism that explains how APC mutations induce proliferative abnormalities that drive colon cancer development. *Front. Oncol.* 3:244. doi: 10.3389/fonc.2013.00244
- Buffa, F. M., Harris, A. L., West, C. M., and Miller, C. J. (2010). Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br. J. Cancer* 102, 428–435. doi: 10.1038/sj.bjc.6605450
- Burns, M. B., Temiz, N. A., and Harris, R. S. (2013). Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat. Genet.* 45, 977–983. doi: 10.1038/ng.2701
- Butturini, E., Carcereri de Prati, A., Boriero, D., and Mariotto, S. (2019). Tumor dormancy and interplay with hypoxic tumor microenvironment. *Int. J. Mol. Sci.* 20:4305. doi: 10.3390/ijms20174305
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2, 401–404. doi: 10.1158/2159-8290.CD-12-0095
- Chang, M. T., Asthana, S., Gao, S. P., Lee, B. H., Chapman, J. S., Kandoth, C., et al. (2016). Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. *Nat. Biotechnol.* 34, 155–163. doi: 10.1038/nbt.3391
- Chen, H., Chong, W., Wu, Q., Yao, Y., Mao, M., and Wang, X. (2019). Association of *LRP1B* mutation with tumor mutation burden and outcomes in melanoma and non-small cell lung cancer patients treated with immune check-point blockades. *Front. Immunol.* 10:1113. doi: 10.3389/fimmu.2019.01113
- Chen, L., Endler, A., and Shibasaki, F. (2009). Hypoxia and angiogenesis: regulation of hypoxia-inducible factors via novel binding factors. *Exp. Mol. Med.* 41, 849–857. doi: 10.3858/emmm.2009.41.12.103
- Chin, Y. E., Kitagawa, M., Su, W. C., You, Z. H., Iwamoto, Y., and Fu, X. Y. (1996). Cell growth arrest and induction of cyclin-dependent kinase inhibitor p21 WAF1/CIP1 mediated by STAT1. *Science* 272, 719–722. doi: 10.1126/science.272.5262.719
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44:e71. doi: 10.1093/nar/gkv1507
- Damen, M. P. F., van Rheenen, J., and Scheele, C. L. G. J. (2020). Targeting dormant tumor cells to prevent cancer recurrence. *FEBS J.* doi: 10.1111/febs.15626
- De Blasio, A., Di Fiore, R., Morreale, M., Carlisi, D., Drago-Ferrante, R., Montalbano, M., et al. (2016). Unusual roles of caspase-8 in triple-negative breast cancer cell line MDA-MB-231. *Int. J. Oncol.* 48, 2339–2348. doi: 10.3892/ijo.2016.3474
- Degasperi, A., Amarante, T. D., Czarnecki, J., Shooter, S., Zou, X., Glodzik, D., et al. (2020). A practical framework and online tool for mutational signature analyses show inter-tissue variation and driver dependencies. *Nat. Cancer* 1, 249–263. doi: 10.1038/s43018-020-0027-5
- Driscoll, C. B., Schuelke, M. R., Kottke, T., Thompson, J. M., Wongthida, P., Tonne, J. M., et al. (2020). APOBEC3B-mediated corruption of the tumor cell immunopeptidome induces heteroclitic neoepitopes for cancer immunotherapy. *Nat. Commun.* 11:790. doi: 10.1038/s41467-020-14568-7
- Dunn, G. P., Koebel, C. M., and Schreiber, R. D. (2006). Interferons, immunity and cancer immunoeediting. *Nat. Rev. Immunol.* 6, 836–848. doi: 10.1038/nri1961
- Dunn, G. P., Old, L. J., and Schreiber, R. D. (2004a). The immunobiology of cancer immunosurveillance and immunoeediting. *Immunity* 21, 137–148. doi: 10.1016/j.immuni.2004.07.017
- Dunn, G. P., Old, L. J., and Schreiber, R. D. (2004b). The three Es of cancer immunoeediting. *Annu. Rev. Immunol.* 22, 329–360. doi: 10.1146/annurev.immunol.22.012703.104803
- Eustace, A., Mani, N., Span, P. N., Irlam, J. J., Taylor, J., Betts, G. N., et al. (2013). A 26-gene hypoxia signature predicts benefit from hypoxia-modifying therapy in laryngeal cancer but not bladder cancer. *Clin. Cancer Res.* 19, 4879–4888. doi: 10.1158/1078-0432.CCR-13-0542
- Faden, D. L., Ding, F., Lin, Y., Zhai, S., Kuo, F., Chan, T. A., et al. (2019). APOBEC mutagenesis is tightly linked to the immune landscape and immunotherapy biomarkers in head and neck squamous cell carcinoma. *Oral. Oncol.* 96, 140–147. doi: 10.1016/j.oraloncology.2019.07.020
- Finn, O. J. (2006). Human tumor antigens, immunosurveillance, and cancer vaccines. *Immunol. Res.* 36, 73–82. doi: 10.1385/IR:36:1:73
- Forcet, C., Ye, X., Granger, L., Corset, V., Shin, H., Bredesen, D. E., et al. (2001). The dependence receptor DCC (deleted in colorectal cancer) defines an alternative mechanism for caspase activation. *Proc. Natl. Acad. Sci. U.S.A.* 98, 3416–3421. doi: 10.1073/pnas.051378298
- Gao, J., Chang, M. T., Johnsen, H. C., Gao, S. P., Sylvester, B. E., Sumer, S. O., et al. (2017). 3D clusters of somatic mutations in cancer reveal numerous rare mutations as functional targets. *Genome Med.* 9:4. doi: 10.1186/s13073-016-0393-x
- Gerlinger, M., McGranahan, N., Dewhurst, S. M., Burrell, R. A., Tomlinson, I., and Swanton, C. (2014). Cancer: evolution within a lifetime. *Annu. Rev. Genet.* 48, 215–236. doi: 10.1146/annurev-genet-120213-092314
- Ghorani, E., Reading, J. L., Henry, J. Y., de Massy, M. R., Rosenthal, R., Turati, V., et al. (2020). The T cell differentiation landscape is shaped by tumour mutations in lung cancer. *Nat. Cancer* 1, 546–561. doi: 10.1038/s43018-020-0066-y
- Gimbrone, M. A., Leapman, S. B., Cotran, R. S., and Folkman, J. (1972). Tumor dormancy *in vivo* by prevention of neovascularization. *J. Exp. Med.* 136, 261–276. doi: 10.1084/jem.136.2.261

- Green, A. M., and Weitzman, M. D. (2019). The spectrum of APOBEC3 activity: from anti-viral agents to anti-cancer opportunities. *DNA Repair. (Amst)* 83:102700. doi: 10.1016/j.dnarep.2019.102700
- Gužvić, M., and Klein, C. A. (2013). Cancer dormancy: time to explore its clinical relevance. *Breast Cancer Res.* 15:321. doi: 10.1186/bcr3590
- Hadfield, G. (1954). The dormant cancer cell. *Br. Med. J.* 2, 607–610. doi: 10.1136/bmj.2.4888.607
- Heldt, F. S., Barr, A. R., Cooper, S., Bakal, C., and Novák, B. (2018). A comprehensive model for the proliferation-quiescence decision in response to endogenous DNA damage in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 115, 2532–2537. doi: 10.1073/pnas.1715345115
- Holmgren, L., O'Reilly, M. S., and Folkman, J. (1995). Dormancy of micrometastases: balanced proliferation and apoptosis in the presence of angiogenesis suppression. *Nat. Med.* 1, 149–153. doi: 10.1038/nm0295-149
- Huang, S. (2021). Reconciling non-genetic plasticity with somatic evolution in cancer. *Trends Cancer* 7, 309–322. doi: 10.1016/j.trecan.2020.12.007
- Ikeda, H., Old, L. J., and Schreiber, R. D. (2002). The roles of IFN gamma in protection against tumor development and cancer immunoeediting. *Cytokine Growth Factor Rev.* 13, 95–109. doi: 10.1016/s1359-6101(01)00038-7
- Itahana, K., Dimri, G. P., Hara, E., Itahana, Y., Zou, Y., Desprez, P. Y., et al. (2002). A role for p53 in maintaining and establishing the quiescence growth arrest in human cells. *J. Biol. Chem.* 277, 18206–18214. doi: 10.1074/jbc.M201028200
- Jager, M., Blokzijl, F., Kuijk, E., Bertl, J., Vougioukalaki, M., Janssen, R., et al. (2019). Deficiency of nucleotide excision repair is associated with mutational signature observed in cancer. *Genome Res.* 29, 1067–1077. doi: 10.1101/gr.246223.118
- Jahanban-Esfahlan, R., Seidi, K., Manjili, M. H., Jahanban-Esfahlan, A., Javaheri, T., and Zare, P. (2019). Tumor cell dormancy: threat or opportunity in the fight against cancer. *Cancers (Basel)* 11:1207. doi: 10.3390/cancers11081207
- Jelaković, B., Castells, X., Tomić, K., Ardin, M., Karanović, S., and Zavadil, J. (2015). Renal cell carcinomas of chronic kidney disease patients harbor the mutational signature of carcinogenic aristolochic acid. *Int. J. Cancer* 136, 2967–2972. doi: 10.1002/ijc.29338
- Jiménez-Sánchez, A., Cast, O., and Miller, M. L. (2019). Comprehensive benchmarking and integration of tumor microenvironment cell estimation methods. *Cancer Res.* 79, 6238–6246. doi: 10.1158/0008-5472.CAN-18-3560
- Kato, M. (2012). Function and cancer genomics of FAT family genes (review). *Int. J. Oncol.* 41, 1913–1918. doi: 10.3892/ijo.2012.1669
- Kim, M. J., Cervantes, C., Jung, Y. S., Zhang, X., Zhang, J., Lee, S. H., et al. (2021). PAF remodels the DREAM complex to bypass cell quiescence and promote lung tumorigenesis. *Mol. Cell* 81, 1698–1714.e6. doi: 10.1016/j.molcel.2021.02.001
- Koebel, C. M., Vermi, W., Swann, J. B., Zerafa, N., Rodig, S. J., Old, L. J., et al. (2007). Adaptive immunity maintains occult cancer in an equilibrium state. *Nature* 450, 903–907. doi: 10.1038/nature06309
- Krock, B. L., Skuli, N., and Simon, M. C. (2011). Hypoxia-induced angiogenesis: good and evil. *Genes Cancer* 2, 1117–1133. doi: 10.1177/1947601911423654
- Kundu, A., Nam, H., Shelar, S., Chandrashekar, D. S., Brinkley, G., Karki, S., et al. (2020). PRDM16 suppresses HIF-targeted gene expression in kidney cancer. *J. Exp. Med.* 217:e20191005. doi: 10.1084/jem.20191005
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883. doi: 10.1093/bioinformatics/bts034
- Leonard, B., Starrett, G. J., Maurer, M. J., Oberg, A. L., Van Bockstal, M., Van Dorpe, J., et al. (2016). APOBEC3G Expression Correlates with T-Cell infiltration and improved clinical outcomes in high-grade serous ovarian carcinoma. *Clin. Cancer Res.* 22, 4746–4755. doi: 10.1158/1078-0432.CCR-15-2910
- Liu, D., Aguirre Ghiso, J., Estrada, Y., and Ossowski, L. (2002). EGFR is a transducer of the urokinase receptor initiated signal that is required for in vivo growth of a human carcinoma. *Cancer Cell* 1, 445–457. doi: 10.1016/s1535-6108(02)00072-7
- Lyu, T., Jia, N., Wang, J., Yan, X., Yu, Y., Lu, Z., et al. (2013). Expression and epigenetic regulation of angiogenesis-related factors during dormancy and recurrent growth of ovarian carcinoma. *Epigenetics* 8, 1330–1346. doi: 10.4161/epi.26675
- MacDonald, J., Ramos-Valdes, Y., Perampalam, P., Litovchick, L., DiMattia, G. E., and Dick, F. A. (2017). A systematic analysis of negative growth control implicates the DREAM complex in cancer cell dormancy. *Mol. Cancer Res.* 15, 371–381. doi: 10.1158/1541-7786.MCR-16-0323-T
- Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Van Loo, P., et al. (2017). Universal patterns of selection in cancer and somatic tissues. *Cell* 171, 1029–1041.e21. doi: 10.1016/j.cell.2017.09.042
- McGranahan, N., and Swanton, C. (2017). Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* 168, 613–628. doi: 10.1016/j.cell.2017.01.018
- Miller, I., Min, M., Yang, C., Tian, C., Gookin, S., Carter, D., et al. (2018). Ki67 is a graded rather than a binary marker of proliferation versus quiescence. *Cell Rep.* 24, 1105–1112.e5. doi: 10.1016/j.celrep.2018.06.110
- Mittal, D., Gubin, M. M., Schreiber, R. D., and Smyth, M. J. (2014). New insights into cancer immunoeediting and its three component phases—elimination, equilibrium and escape. *Curr. Opin. Immunol.* 27, 16–25. doi: 10.1016/j.coi.2014.01.004
- Moon, K. R., van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D. B., Chen, W. S., et al. (2019). Visualizing structure and transitions in high-dimensional biological data. *Nat. Biotechnol.* 37, 1482–1492. doi: 10.1038/s41587-019-0336-3
- Moserle, L., Amadori, A., and Indraco, S. (2009). The angiogenic switch: implications in the regulation of tumor dormancy. *Curr. Mol. Med.* 9, 935–941. doi: 10.2174/156652409789712800
- Naumov, G. N., Bender, E., Zurakowski, D., Kang, S. Y., Sampson, D., Flynn, E., et al. (2006). A model of human tumor dormancy: an angiogenic switch from the nonangiogenic phenotype. *J. Natl. Cancer Inst.* 98, 316–325. doi: 10.1093/jnci/dij068
- Oh, C. W., Hoover-Plow, J., and Plow, E. F. (2003). The role of plasminogen in angiogenesis *in vivo*. *J. Thromb. Haemost.* 1, 1683–1687. doi: 10.1046/j.1538-7836.2003.00182.x
- Park, S. Y., and Nam, J. S. (2020). The force awakens: metastatic dormant cancer cells. *Exp. Mol. Med.* 52, 569–581. doi: 10.1038/s12276-020-0423-z
- Petrova, V., Annicchiarico-Petruzzelli, M., Melino, G., and Amelio, I. (2018). The hypoxic tumour microenvironment. *Oncogenesis* 7:10. doi: 10.1038/s41389-017-0011-9
- Phan, T. G., and Croucher, P. I. (2020). The dormant cancer cell life cycle. *Nat. Rev. Cancer* 20, 398–411. doi: 10.1038/s41568-020-0263-0
- Polyak, K., Xia, Y., Zweier, J. L., Kinzler, K. W., and Vogelstein, B. (1997). A model for p53-induced apoptosis. *Nature* 389, 300–305. doi: 10.1038/38525
- Pugh, C. W., and Ratcliffe, P. J. (2003). Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat. Med.* 9, 677–684. doi: 10.1038/nm0603-677
- Qin, L., Xu, Y., Ma, G., Liao, L., Wu, Y., Li, Y., et al. (2015). NCOA1 promotes angiogenesis in breast tumors by simultaneously enhancing both HIF1 α - and AP-1-mediated VEGF α transcription. *Oncotarget* 6, 23890–23904. doi: 10.18632/oncotarget.4341
- Qiu, G. Z., Jin, M. Z., Dai, J. X., Sun, W., Feng, J. H., and Jin, W. L. (2017). Reprogramming of the tumor in the hypoxic niche: the emerging concept and associated therapeutic strategies. *Trends Pharmacol. Sci.* 38, 669–686. doi: 10.1016/j.tips.2017.05.002
- Ranganathan, A. C., Zhang, L., Adam, A. P., and Aguirre-Ghiso, J. A. (2006). Functional coupling of p38-induced up-regulation of BiP and activation of RNA-dependent protein kinase-like endoplasmic reticulum kinase to drug resistance of dormant carcinoma cells. *Cancer Res.* 66, 1702–1711. doi: 10.1158/0008-5472.CAN-05-3092
- Rebhandl, S., Huemer, M., Gassner, F. J., Zaborsky, N., Hebenstreit, D., Catakovic, K., et al. (2014). APOBEC3 signature mutations in chronic lymphocytic leukemia. *Leukemia* 28, 1929–1932. doi: 10.1038/leu.2014.160
- Robbiani, D. F., and Nussenzweig, M. C. (2013). Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. *Annu. Rev. Pathol.* 8, 79–103. doi: 10.1146/annurev-pathol-020712-164004
- Roberts, S. A., Lawrence, M. S., Klimczak, L. J., Grimm, S. A., Fargo, D., Stojanov, P., et al. (2013). An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.* 45, 970–976. doi: 10.1038/ng.2702
- Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B. S., and Swanton, C. (2016). DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol.* 17:31. doi: 10.1186/s13059-016-0893-4
- Sadasivam, S., and DeCaprio, J. A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nat. Rev. Cancer* 13, 585–595. doi: 10.1038/nrc3556
- Semenza, G. L. (2003). Targeting HIF-1 for cancer therapy. *Nat. Rev. Cancer* 3, 721–732. doi: 10.1038/nrc1187

- Senft, D., and Ronai, Z. A. (2016). Immunogenic, cellular, and angiogenic drivers of tumor dormancy—a melanoma view. *Pigment Cell Melanoma Res.* 29, 27–42. doi: 10.1111/pcmr.12432
- Shen, S., and Clairambault, J. (2020). Cell plasticity in cancer cell populations. *F1000Res* 9:F1000FacultyRev-635. doi: 10.12688/f1000research.24803.1
- Shivji, M. K., Podust, V. N., Hübscher, U., and Wood, R. D. (1995). Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry* 34, 5011–5017. doi: 10.1021/bi00015a012
- Sigmund, S., Avanzato, D., and Lanzetti, L. (2018). Emerging functions of the EGFR in cancer. *Mol. Oncol.* 12, 3–20. doi: 10.1002/1878-0261.12155
- Simó-Riudalbas, L., érez-Salvia, M. P., Setien, F., Villanueva, A., Moutinho, C., Martínez-Cardús, A., et al. (2015). KAT6B is a tumor suppressor histone H3 Lysine 23 acetyltransferase undergoing genomic loss in small cell lung cancer. *Cancer Res.* 75, 3936–3945. doi: 10.1158/0008-5472.CAN-14-3702
- Stenglein, M. D., Burns, M. B., Li, M., Lengyel, J., and Harris, R. S. (2010). APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nat. Struct. Mol. Biol.* 17, 222–229. doi: 10.1038/nsmb.1744
- Swanton, C., McGranahan, N., Starrett, G. J., and Harris, R. S. (2015). APOBEC enzymes: mutagenic fuel for cancer evolution and heterogeneity. *Cancer Discov.* 5, 704–712. doi: 10.1158/2159-8290.CD-15-0344
- Tate, J. G., Bamford, S., Jubb, H. C., Sondka, Z., Beare, D. M., Bindal, N., et al. (2019). COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* 47 (D1), D941–D947. doi: 10.1093/nar/gky1015
- Temko, D. I., Tomlinson, P. M., Severini, S., Schuster-Böckler, B., and Graham, T. A. (2018). The effects of mutational processes and selection on driver mutations across cancer types. *Nat. Commun.* 9:1857. doi: 10.1038/s41467-018-04208-6
- Teng, M. W., Swann, J. B., Koebel, C. M., Schreiber, R. D., and Smyth, M. J. (2008). Immune-mediated dormancy: an equilibrium with cancer. *J. Leukoc. Biol.* 84, 988–993. doi: 10.1189/jlb.1107774
- Theodoropoulos, V. E., Lazaris, A. C., Sofras, F., Gerzelis, I., Tsoukala, V., Ghikonti, I., et al. (2004). Hypoxia-inducible factor 1 alpha expression correlates with angiogenesis and unfavorable prognosis in bladder cancer. *Eur. Urol.* 46, 200–208. doi: 10.1016/j.eururo.2004.04.008
- Tommasini-Ghelfi, S., Murnan, K., Kouri, F. M., Mahajan, A. S., May, J. L., and Stegh, A. H. (2019). Cancer-associated mutation and beyond: the emerging biology of isocitrate dehydrogenases in human disease. *Sci. Adv.* 5:eaaw4543. doi: 10.1126/sciadv.aaw4543
- Tummers, B., and Green, D. R. (2017). Caspase-8: regulating life and death. *Immunol. Rev.* 277, 76–89. doi: 10.1111/imr.12541
- Turcan, S., Rohle, D., Goenka, A., Walsh, L. A., Fang, F., Yilmaz, E., et al. (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* 483, 479–483. doi: 10.1038/nature10866
- Vogelstein, B., Lane, D., and Levine, A. J. (2000). Surfing the p53 network. *Nature* 408, 307–310. doi: 10.1038/35042675
- Wall, L., Burke, F., Smyth, J. F., and Balkwill, F. (2003). The anti-proliferative activity of interferon-gamma on ovarian cancer: in vitro and in vivo. *Gynecol. Oncol.* 88 (1 Pt 2), S149–S151. doi: 10.1006/gyno.2002.6707
- Wang, F. X., Huang, J., Zhang, H., and Ma, X. (2008). APOBEC3G upregulation by alpha interferon restricts human immunodeficiency virus type 1 infection in human peripheral plasmacytoid dendritic cells. *J. Gen. Virol.* 89 (Pt 3), 722–730. doi: 10.1099/vir.0.83530-0
- Wang, H. F., Wang, S. S., Huang, M. C., Liang, X. H., Tang, Y. J., and Tang, Y. L. (2019). Targeting immune-mediated dormancy: a promising treatment of cancer. *Front. Oncol.* 9:498. doi: 10.3389/fonc.2019.00498
- Wang, S., Huimin, L., Minfang, S., Zaoke, H., Tao, W., Xuan, W., et al. (2020). Copy number signature analyses in prostate cancer reveal distinct etiologies and clinical outcomes. *medRxiv [Preprint]* doi: 10.1101/2020.04.27.20082404
- Wang, S., Jia, M., He, Z., and Liu, X. S. (2018). APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. *Oncogene* 37, 3924–3936. doi: 10.1038/s41388-018-0245-9
- Watnick, R. S., Rodriguez, R. K., Wang, S., Blois, A. L., Rangarajan, A., Ince, T., et al. (2015). Thrombospondin-1 repression is mediated via distinct mechanisms in fibroblasts and epithelial cells. *Oncogene* 34, 2823–2835. doi: 10.1038/onc.2014.228
- Wherry, E. J., and Kurachi, M. (2015). Molecular and cellular insights into T cell exhaustion. *Nat. Rev. Immunol.* 15, 486–499. doi: 10.1038/nri3862
- White, D. E., Kurpios, N. A., Zuo, D., Hassell, J. A., Blaess, S., Mueller, U., et al. (2004). Targeted disruption of beta1-integrin in a transgenic mouse model of human breast cancer reveals an essential role in mammary tumor induction. *Cancer Cell* 6, 159–170. doi: 10.1016/j.ccr.2004.06.025
- Winter, S. C., Buffa, F. M., Silva, P., Miller, C., Valentine, H. R., Turley, H., et al. (2007). Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res.* 67, 3441–3449. doi: 10.1158/0008-5472.CAN-06-3322
- Yamazaki, H., Shirakawa, K., Matsumoto, T., Hirabayashi, S., Murakawa, Y., Kobayashi, M., et al. (2019). Endogenous APOBEC3B overexpression constitutively generates DNA substitutions and deletions in myeloma cells. *Sci. Rep.* 9:7122. doi: 10.1038/s41598-019-43575-y
- Ye, D., Ma, S., Xiong, Y., and Guan, K. L. (2013). R-2-hydroxyglutarate as the key effector of IDH mutations promoting oncogenesis. *Cancer Cell.* 23, 274–276. doi: 10.1016/j.ccr.2013.03.005
- Yi, J. S., Cox, M. A., and Zajac, A. J. (2010). T-cell exhaustion: characteristics, causes and conversion. *Immunology* 129, 474–481. doi: 10.1111/j.1365-2567.2010.03255.x
- Yuan, S., Norgard, R. J., and Stanger, B. Z. (2019). Cellular plasticity in cancer. *Cancer Discov.* 9, 837–851. doi: 10.1158/2159-8290.CD-19-0015
- Zhang, J., Bajari, R., Andric, D., Gerthoffert, F., Lepsa, A., Nahal-Bose, H., et al. (2019). The international cancer genome consortium data portal. *Nat. Biotechnol.* 37, 367–369. doi: 10.1038/s41587-019-0055-9
- Zhang, T., Otevrel, T., Gao, Z., Ehrlich, S. M., Fields, J. Z., and Boman, B. M. (2001). Evidence that APC regulates survivin expression: a possible mechanism contributing to the stem cell origin of colon cancer. *Cancer Res.* 61, 8664–8667.
- Zhao, X., Liu, J., Ge, S., Chen, C., Li, S., Wu, X., et al. (2019). Saikosaponin a inhibits breast cancer by regulating Th1/Th2 Balance. *Front. Pharmacol.* 10:624. doi: 10.3389/fphar.2019.00624
- Zhi, X., Tao, J., Xie, K., Zhu, Y., Li, Z., Tang, J., et al. (2014). MUC4-induced nuclear translocation of β -catenin: a novel mechanism for growth, metastasis and angiogenesis in pancreatic cancer. *Cancer Lett.* 346, 104–113. doi: 10.1016/j.canlet.2013.12.021

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.

Copyright © 2021 Wiecek, Jacobson, Lason and Secrier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.