Am J Cancer Res 2022;12(2):744-762
www.ajcr.us /ISSN:2156-6976/ajcr0139477

Original Article

APOBEC3F expression in triple-negative breast cancer is associated with tumor microenvironment infiltration and activation of cancer immunity and improved survival

Rongrong Wu[1,2*], Masanori Oshi[1,3*], Mariko Asaoka[1,2*], Michelle R Huyser[1], Yoshihisa Tokumaru[1,4], Akimitsu Yamada[3], Li Yan[5], Itaru Endo[3], Takashi Ishikawa[2], Kazuaki Takabe[1,2,3,6,7,8]

*[1]Department of Surgical Oncology, Roswell Park Cancer Institute, Buffalo, NY, USA; [2]Department of Breast Surgery and Oncology, Tokyo Medical University, Tokyo, Japan; [3]Department of Gastroenterological Surgery, Yokohama City University Graduate School of Medicine, Yokohama, Kanagawa, Japan; [4]Department of Surgical Oncology, Graduate School of Medicine, Gifu University, Gifu, Japan; [5]Department of Biostatistics & Bioinformatics, Roswell Park Cancer Institute, Buffalo, NY, USA; [6]Department of Surgery, Jacobs School of Medicine and Biomedical Sciences, State University of New York, Buffalo, NY, USA; [7]Department of Surgery, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan; [8]Department of Breast Surgery, Fukushima Medical University, Fukushima, Japan. *Equal contributors.*

Received September 28, 2021; Accepted November 16, 2021; Epub February 15, 2022; Published February 28, 2022

Abstract: The apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like (APOBEC) causes a point mutation from cytidine to uracil in DNA and/or RNA. The role of APOBEC3A and APOBEC3B in breast cancer has been well described, whereas that of APOBEC3F remains unknown. To investigate the clinical relevance of APOBEC3F expression, we analyzed a total of 3000 breast cancer cases from multiple independent large patient cohorts including METABRIC, TCGA, GSE75688, and GSE114725. High expression of *APOBEC3F* was associated with improved disease-specific and overall survival in triple negative breast cancer (TNBC). *APOBEC3F* is not usually a reflection of cancer cell biology in TNBC or luminal breast cancer, except for homologous recombination deficiency in TNBC. In the TNBC homologous recombination deficiency group, *APOBEC3F* expression was not consistently associated with intratumor heterogeneity, mutation rates, or neoantigens. *APOBEC3F* expression did not correlate with response to any of the drugs tested in breast cancer cell lines in vitro. However, high *APOBEC3F* expression was associated with enrichment of several immune-related gene sets and immune activity. High *APOBEC3F* expression also accompanied higher infiltration of anti-cancer immune cell infiltration in TNBC. However, in luminal breast cancer, high *APOBEC3F* tumor significantly enriched not only immune-related gene sets, but also cell proliferation-, metastasis-, and apoptosis-related gene sets. Analysis of single-cell transcriptomes showed *APOBEC3F* exclusively expressed in immune cells and significantly associated with cytolytic activity of the immune cells, immune response, and immune cell proliferation. Expression of immune checkpoint genes was uniformly elevated in *APOBEC3F*-high tumors. We conclude that *APOBEC3F* is exclusively expressed in immune cells and this expression is associated with enhanced anti-cancer immune response as well as improved survival in TNBC.

Keywords: APOBEC, APOBEC3F, GSEA, immune cell, immune check point, single-cell sequencing, triple negative breast cancer, tumor microenvironment, xCell

Introduction

The apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3) subfamily of cytidine deaminases has been highlighted as a cause of somatic mutations in many types of cancer including breast. Human APOBEC3s are encoded by seven structurally related genes on chromosome 22 (3A, 3B, 3C, 3DE, 3F, 3G, and 3H) [1]. APOEBC3s preferentially deaminate the central cytidine (C) between thymine (T) and adenine (A) or T (referred to as TCW motif), resulting in many C to T and C to G substitutions within the TCW motif. This character-

istic APOBEC signature has been reported in approximately 75% of cancer types and >50% of all tumors analyzed [2]. This overall mutational load particularly contributes to breast, lung, esophagus, bladder, head and neck, and cervix cancers [3-5]. Nik-Zainal et al. reported the presence of clustered mutations called "kataegis" identified through genome sequencing of 21 breast cancer tumors in 2012, which are believed to be caused by APOBEC3s [6].

Increased somatic mutations in tumors may induce a more aggressive tumor phenotype as well as a greater opportunity for neoantigen presentation on cancer cells to evoke immune responses against the cancer cell [7-9]. Triple negative breast cancer (TNBC), characterized by estrogen receptor (ER)-negative, progesterone receptor (PgR)-negative, and epidermal growth factor receptor 2 (HER2)-negative, is known to have a high mutational burden. Among various breast cancer subtypes, TNBC is the most proliferative with a generally poor prognosis. However, TNBC patients with high tumor neoantigen presentation are known to have better survival [10].

In vitro cell line experiments and even syngeneic animal experiments may not fully model the interaction between cancer cells and immune cells in the tumor microenvironment (TME) of human TNBC patients. Our group has been utilizing an in silico translational research approach to overcome this challenge by analyzing clinical and molecular data from large breast cancer patient cohorts to objectively quantify immune cell infiltration and function in tumors, thereby elucidating cancer biology [11-19].

In breast cancer, APOBEC3A and APOBEC3B have been considered the key mutagenic drivers promoting mutation associated with poor prognosis [20, 21]. However, APOBEC3-related mutational signatures have been positively correlated with neoantigen production and immune cell infiltration [22, 23]. Although all APOBEC3 family members have deaminating activity causing nucleotide changes in DNA or RNA [24], their relevance in breast cancer remains largely unknown except for APOBEC3A and APOBEC3B.

Human APOBEC3F was discovered with APOBEC3D and APOBEC3G as a host factor that inhibits human immunodeficiency virus type 1 (HIV-1) replication by inducing hypermutation of HIV-1 DNA [25-27]. Since the viral infectivity factor (Vif) eliminates the activity of APOBEC3 and renders HIV-1 resistant to these antiviral effects, the interaction of HIV-1 with Vif has become a new HIV-1 drug target [28]. APOBEC3F was shown to partially inhibit Vif degradation [25]. The only reported role of APOBEC3F in cancer has been its high expression in hepatocellular carcinoma (HCC) tumors and association with poor recurrence-free survival [29]. To the best of our knowledge, no study demonstrates APOBEC3F's clinical importance in breast cancer.

Given the above-mentioned roles of APOBEC3 in breast cancer, we hypothesized that *APOBEC3F* gene expression may increase breast cancer tumor mutation by affecting immune cell infiltration, thus impacting anticancer immune response and patient survival.

Materials and methods

Breast cancer patient cohorts

Overall, 3000 breast cancer patients' information was obtained for analyses from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) and The Cancer Genome Atlas (TCGA) projects. We obtained annotated transcriptome and clinicopathological outcome data from cBioPortal, utilizing the same methodology as previously reported [30-32]. The discovery cohort included 1904 patients of METABRIC, and the validation cohort included 1077 patients of TCGA. Within each cohort, the top and bottom tertiles of patients as per tumor gene expression were assigned to high and low gene expressor groups respectively. 21253 single-cell transcriptomes of eight patients from GSE114725 [33] and 515 single-cell transcriptomics of 11 patients from GSE75688 [34] were used as primary breast cancer single-cell sequencing cohorts. Institutional review board approval was not required for this study because the four noted datasets are de-identified and publicly available. Gene expression data from all four cohorts we used were normalized by the time of the original publication [30, 33-35].

Gene set enrichment analysis

For investigation of gene function, we used preranked gene set enrichment analysis (GSEA) as

previously described [15, 36, 37]. False discovery rate (FDR) was calculated using multiple hypothesis tests comparing gene expression with Hallmark collection gene sets from the Molecular Signatures Database (MSigDB). GSEA recommended a false discovery rate (FDR) of 0.25 for statistical significance [38].

Composition of immune cells and cancer immunity score estimation

xCell algorithm [39] was used to estimate tumor composition of infiltrating immune cells from whole tumor transcriptomic data [16, 18, 40]. Expression levels of Perforin 1 (*PRF1*) and Granzyme A (*GZMA*) were used to measure immune cytolytic activity as the CYT score [41, 42], which we have previously associated with reduced intratumor heterogeneity and better clinical outcomes in TNBC [43]. Tumor transcriptome-based estimates of tumor-infiltrating lymphocyte (TIL) fraction, leukocyte fraction, Interferon (IFN)-γ Response, Lymphocyte Infiltration Signature, and B and T cell receptor (BCR/TCR) richness were all obtained from the published study by Thorsson et al. [44].

Others

For statistical analysis and data plotting, R (version 4.0.1) with Bioconductor (version 3.13) and Microsoft Excel (version 16 for Windows) were used. All *P*-value cut-offs for statistical significance were set at 0.05. Overall survival (OS) was defined as the time from treatment completion to all-cause related death, and disease-specific survival (DSS) was defined as the time from treatment completion to disease-specific death. Kaplan-Meier method with log rank test was used for survival analysis. Statistical comparisons between groups were made using Mann-Whitney U test for two groups or Kruskal-Wallis test for multiple groups. Median and interquartile level values are shown in Tukey boxplots. Transcriptomics and drug sensitivity data for each breast cancer cell line were from the cancer cell line encyclopedia (https://depmap.org/portal/).

Results

High tumor APOBEC3F expression indicates better prognosis in triple negative breast cancer (TNBC)

First, we investigated the relationship between tumor *APOBEC3F* expression and clinical features in different breast cancer subtypes-luminal (estrogen receptor [ER]-positive/human epidermal growth factor receptor 2 [HER2]-negative), HER2-positive, and TNBC. *APOBEC3F* expression was highest in TNBC compared to other subtypes in both METABRIC and TCGA cohorts (all $P<0.001$, **Figure 1A**). Compared to normal tissues in TCGA, APOBEC3F expression was only higher in TNBC than in normal tissues, and conversely lower in luminal and HER2-positive types (All $P<0.001$, **Figure 1A**, right). Expression of the gene, however, did not show any consistent association with Nottingham histological grade (grade I-III) or American Joint Committee on Cancer (AJCC) pathological stage (stage I-III) of tumors in the two cohorts (Figure S1). We defined the top two-thirds of cohorts by their tumor *APOBEC3F* expression as high expressors and the rest as low expressors. This cohort grouping led to the observation that *APOBEC3F* expression was significantly associated with better disease-specific survival (DSS) and overall survival (OS) consistently in both METABRIC and TCGA TNBC cohorts (all $P<0.05$), but not for other cancer subtypes (**Figure 1B**). Thus TNBC, which has the highest *APOBEC3F* expression among breast cancer subtypes, is associated with better outcomes.

Tumor APOBEC3F expression is associated with homologous recombination deficiency (HRD) in TNBC and mutation rates in luminal subtype

Because mutations can generate neoantigens that prime an anti-tumor T cell immune attack [45], we investigated the relationship between tumor *APOBEC3F* expression and tumor homologous recombination deficiency (HRD), intratumor heterogeneity, mutation- and neoantigen-related scores, previously established by Thorsson et al. [44] for the TCGA cohort. As shown in **Figure 2A**, HRD alone was associated with high *APOBEC3F* in TNBC. Conversely, in luminal subtype, low *APOBEC3F* was associated with a fraction altered score and both silent and non-silent mutation rate. These results suggest that although the APOBEC3 family, especially APOBEC3B, has been identified as enzymes that generate mutations, high expression of *APOBEC3F* is not associated with tumor mutation or neoantigen burden.

We also examined the association of drug sensitivity of various human breast cancer cell lines and their *APOBEC3F* expression to investi-

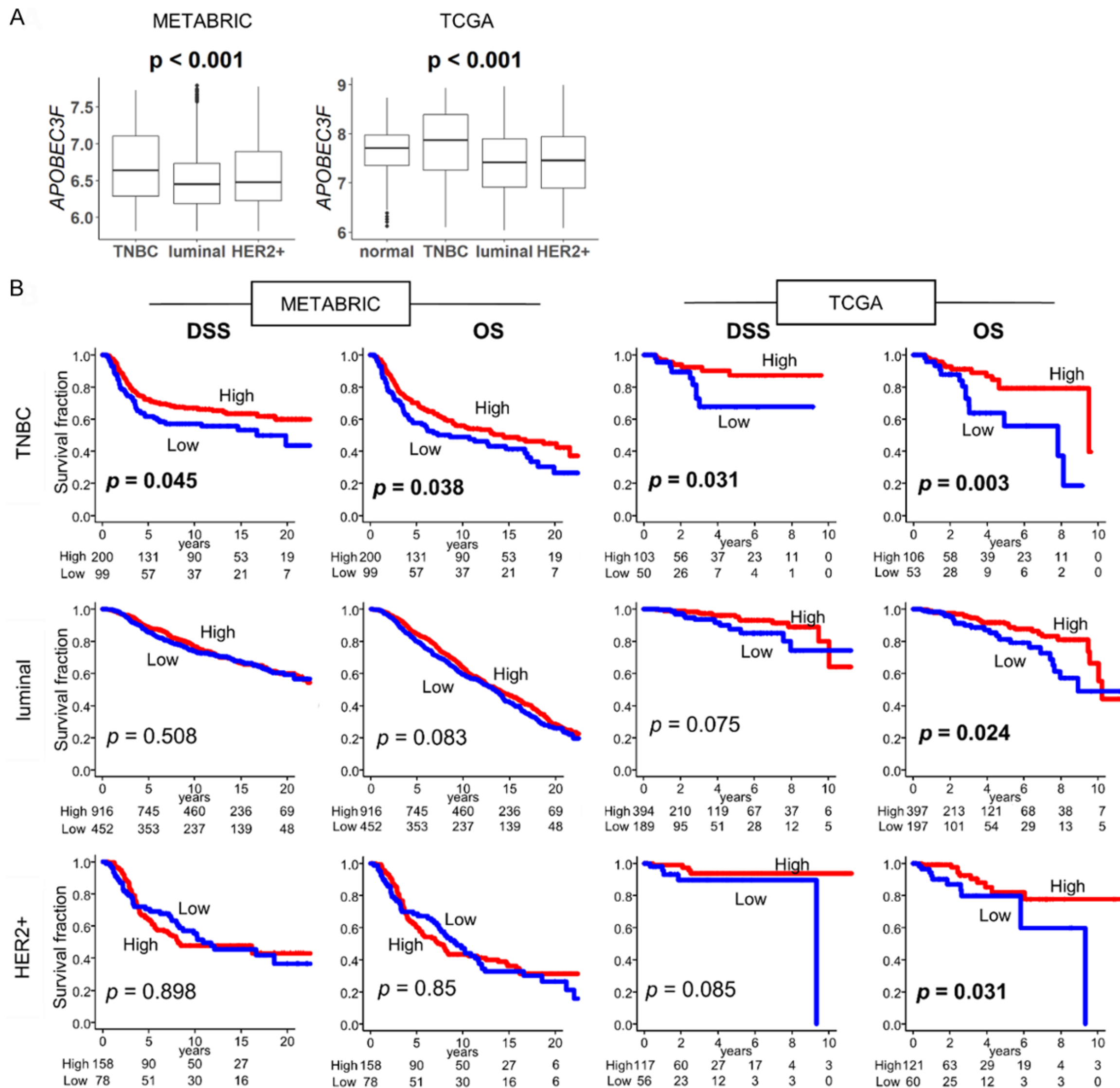
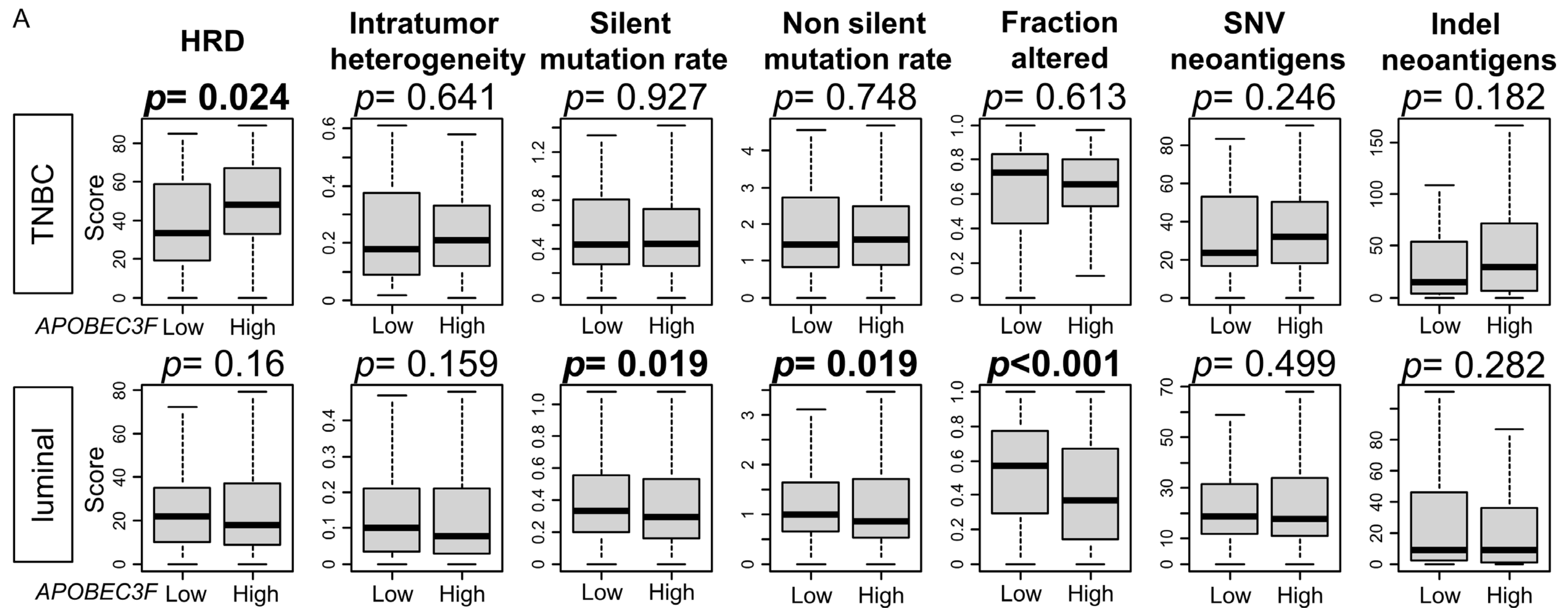


Figure 1. Tumor *APOBEC3F* expression by breast cancer subtype is shown in the (A) Boxplots show APOBEC3F expression between TNBC, ER or PR positive/HER2 negative (luminal), and HER2 positive (HER2+) subtypes in METABRIC, and normal tissue in TCGA. The *P*-values were derived by Kruskal-Wallis test for multi-group comparison and Mann-Whitney U test for two-group comparison. (B) Kaplan-Meier curves for DSS and OS by *APOBEC3F* gene expression in METABRIC and TCGA cohorts are shown for the luminal, TNBC, and HER2+ subtypes. High and low APOBEC3F expressors are defined using the intra-cohort bottom tertile as a cut-off. A log-rank test was used to calculate *P*-values. DSS; disease-specific survival, ER; estrogen receptor, HER2; human epidermal growth factor receptor 2, OS; overall survival, PR; progesterone receptor, TNBC; triple negative breast cancer.

gate the role of *APOBEC3F* and treatment response in cancer cells. We did not find a clear association for cell-lines of either TNBC or luminal subtypes (**Figures 2B**, S2).

High tumor APOBEC3F expression is associated with elevated immune response, immune cytolytic activity (CYT), and anti-cancer immune cell infiltration

Since *APOBEC3F* expression was not associated with tumor mutation rate, it was of interest to investigate whether it associates with infiltration of non-cancer immune cells in the breast cancer tumor microenvironment (TME). For this, we utilized xCell algorithm to estimate the composition of immune cell infiltrations from bulk tumor transcriptomes of the METABRIC and TCGA cohorts. Anti-cancer and pro-cancer immune cell infiltration were positively associated with high *APOBEC3F* expression in TNBC, including $CD8^{+}$ T cells, B cells, dendritic cells, and M1 and M2 macrophages,



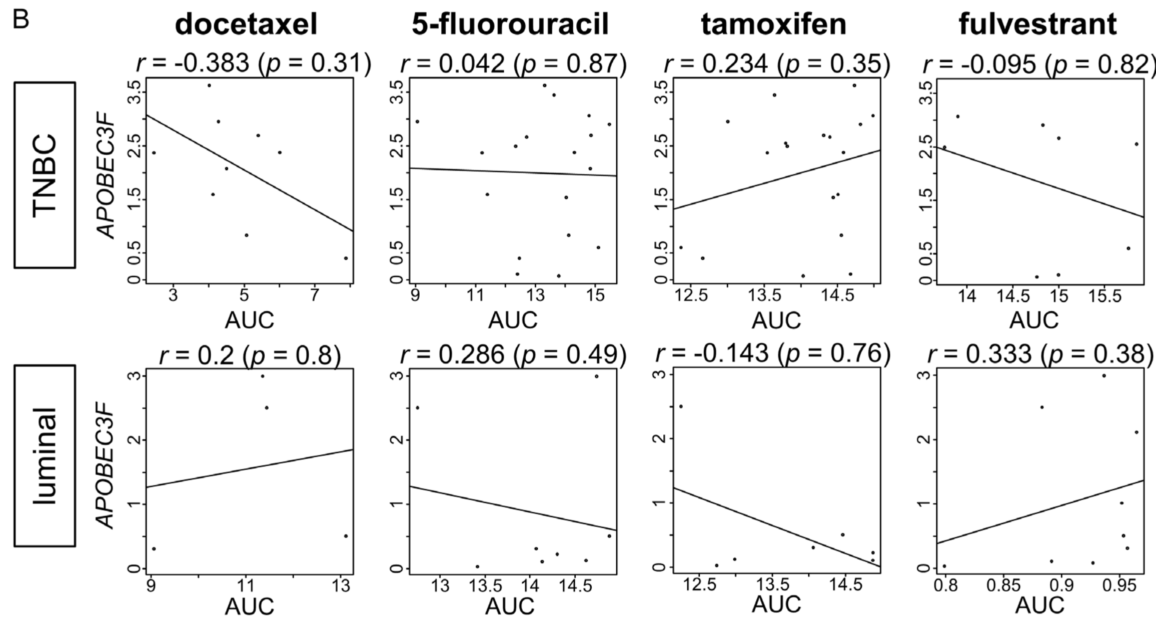


Figure 2. Association of *APOBEC3F* expression with tumor HRD, intratumor heterogeneity, mutation rates, and neoantigens in the TCGA cohort, and with drug sensitivity of human breast cancer cell lines. A. Boxplots according to low and high *APOBEC3F* breast cancer in the TCGA cohort of mutation scores; HRD, intratumor heterogeneity, silent mutation, non-silent mutation rate, altered fraction, SNV neoantigen, and indel neoantigen. High and low APOBEC3F expressors are defined using the intra-cohort bottom tertile as cut-off. Mann-Whitney U test was used to calculate all *P*-values. B. Correlation plots of *APOBEC3F* expression and drug sensitivity AUC; docetaxel, 5-fluorouracil, tamoxifen, fulvestrant in TNBC and luminal breast cancer cell lines. Cell line data was downloaded from the Depmap portal. The PRISM Repurposing 19Q4 was used for each drug sensitivity. Correlation coefficients were derived from the Spearman rank test. HRD; homologous recombination deficiency, SNV; single nucleotide variant, AUC; area under the curve.

consistently in both TCGA and METABRIC cohorts (**Figure 3A**, **3B**; all *P*<0.05). Immune cell killing activity, calculated as Cytolytic Activity score (CYT), was consistently and significantly elevated in both cohorts for high *APOBEC3F* expressors (**Figure 3C**, both *P*< 0.001). Similarly, high *APOBEC3F* expressors among patients with luminal subtype had greater tumor infiltration by both anti-cancer and pro-cancer immune cells (Figure S3).

High APOBEC3F expressors of TNBC subtype have enriched immune-related gene expression

Given that infiltrations of immune cells were very similar among breast cancer subtypes, to explain TNBC significantly better survival we hypothesized that high *APOBEC3F* expression is associated with enhanced immune response. To this end, gene set enrichment analysis (GSEA) using Hallmark gene set collection was performed to compare high and low *APOBEC3F* expressors of TNBC subtype in the METABRIC and TCGA cohorts. As expected, high expressors had enriched expression of a few immune-related gene sets consistently in both cohorts, including inflammatory response, IL-2/STAT5 signaling, interferon (IFN)-γ response, IFN-α response, IL-6/Jak/SSTAT3 signaling, and complement (**Figure 4A**, all FDR<0.25).

A high level of cancer immunity scores provided by Thorsson et al., such as IFN-γ response, lymphocyte infiltration signature, leukocyte fraction, tumor infiltrating lymphocytes (TIL) regional fraction, and T cell receptor (TCR) richness in TCGA cohort were significantly associated with high *APOBEC3F* TNBC (**Figure 4B**, all *P*<0.05). Hence, activation of multiple immune responses was associated with high *APOBEC3F* expression in TNBC.

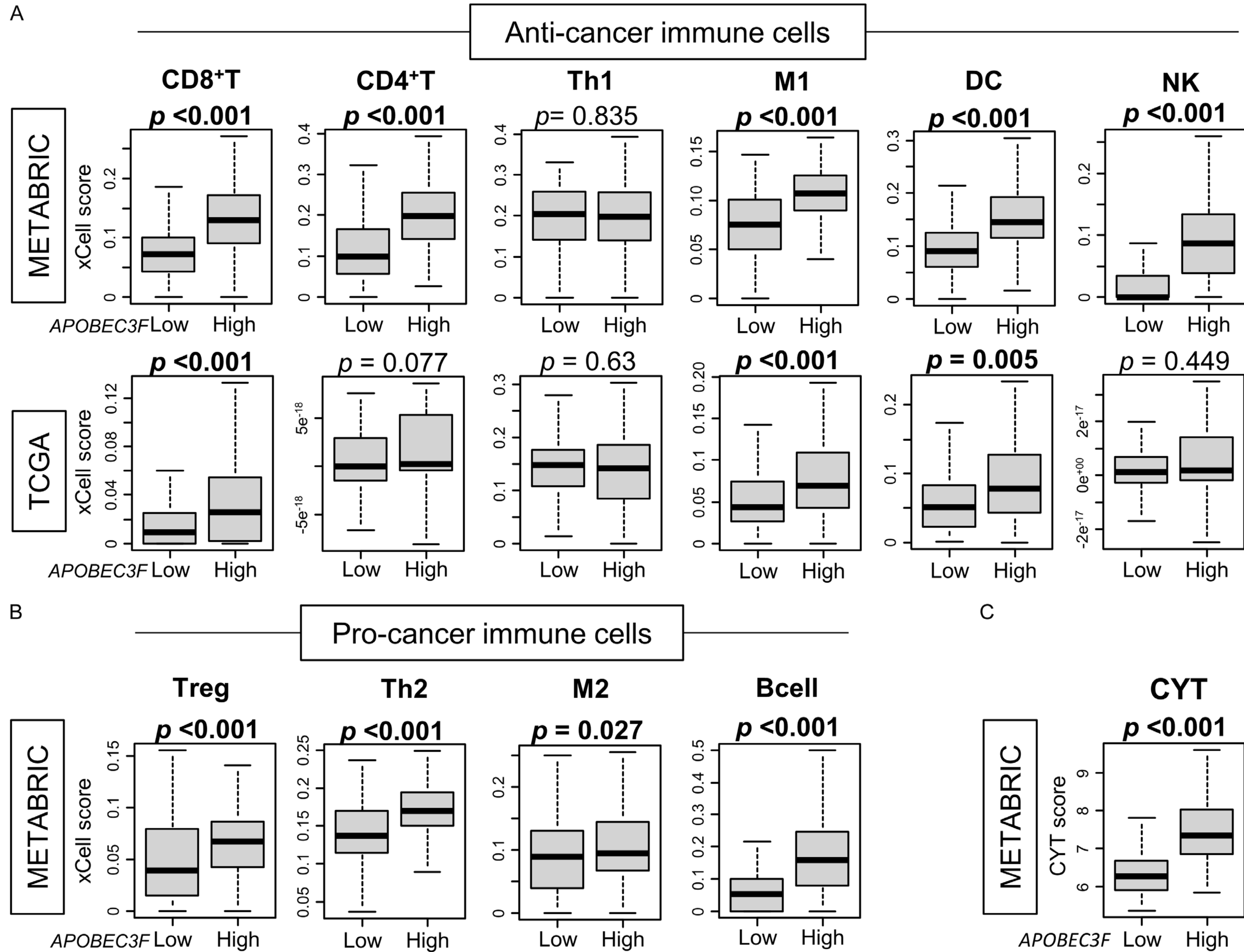
High APOBEC3F expression in luminal subtype is associated with enriched expression of cancer aggravating pathways in addition to immune-related gene sets

We also examined tumors of luminal subtype of the METABRIC and TCGA cohorts to compare high and low *APOBEC3F* expressors for enrichment of gene pathways. Like TNBC, high *APOBEC3F* expressors of luminal subtype also showed enrichment for several immune-related gene sets consistently in both cohorts, including inflammatory response, IL-2/STAT5 IFN-γ response, IFN-α response, IL6/Jak/STAT3 signaling, TNF-α signaling via NFkB, coagulation, and allograft rejection (**Figure 5A**). However, unlike TNBC, for the luminal subtype several other gene sets were also consistently enriched in both cohorts, such as apoptosis-related gene sets, tumor p53 pathway, hypoxia, reactive oxygen species (ROS) pathway, Kirsten rat sarcoma (KRAS) signaling up, phosphatidylinositol-4,5-bisphosphate 3-kinase (PI3K)/threonine protein kinase B (AKT)/mammalian target of rapamycin (mTOR) signaling, angiogenesis, fatty acid metabolism, bile acid metabolism, heme metabolism, apical surface, and UV response up (**Figure 5B-E**).

Regarding cancer immunity scores provided by Thorsson et al. for the TCGA cohort, high-*APOBEC3F* luminal cancer, like TNBC, had higher scores for IFN-γ response, lymphocyte infiltration signature, leukocyte fraction, tumor infiltrating lymphocytes (TIL) regional fraction, B cell receptor (BCR) and T cell receptor (TCR) richness (all *P*<0.05, **Figure 5F**). Taken together, high tumor *APOBEC3F* expression is significantly linked with cancer immunity, regardless of cancer subtype, while simultaneously promoting cancer aggressiveness in luminal breast cancer.

In both TNBC and luminal cancers, APOBEC3F is exclusively expressed in tumor immune cells at levels associated with heightened immune activity

Given the strong association of *APOBEC3F* expression with immune cell infiltration and immune response in the breast cancer TME, it was of interest whether *APOBEC3F* is a mere marker of immune cells or has function within immune cells. In the GSE75688 breast cancer tumor single-cell transcriptome dataset, we observed that *APOBEC3F* gene was exclusively expressed in immune cells compared with myeloid, stromal, and cancer cells (**Figure 6A**, *P*<0.001). Furthermore, among immune cells in the GSE114725 breast cancer tumor single-cell transcriptome dataset, *APOBEC3F* was expressed highly in $CD4^+$ T cells, $CD8^+$ T cells, natural killer cells, neutrophils, and B-cells, and at a relatively lower level in regulatory T cells (T-regs), dendritic cells (DC), macrophages (MΦ), and monocytes (Mono) (**Figure 6A**, *P*<0.001).


A
Anti-cancer immune cells
CD8+T
CD4+T
Th1
M1
DC
NK
METABRIC
TCGA
xCell score
APOBEC3F
Low
High
p <0.001
p= 0.835
p = 0.077
p = 0.63
p = 0.005
p = 0.449
B
Pro-cancer immune cells
Treg
Th2
M2
Bcell
p = 0.027
C
CYT
CYT score

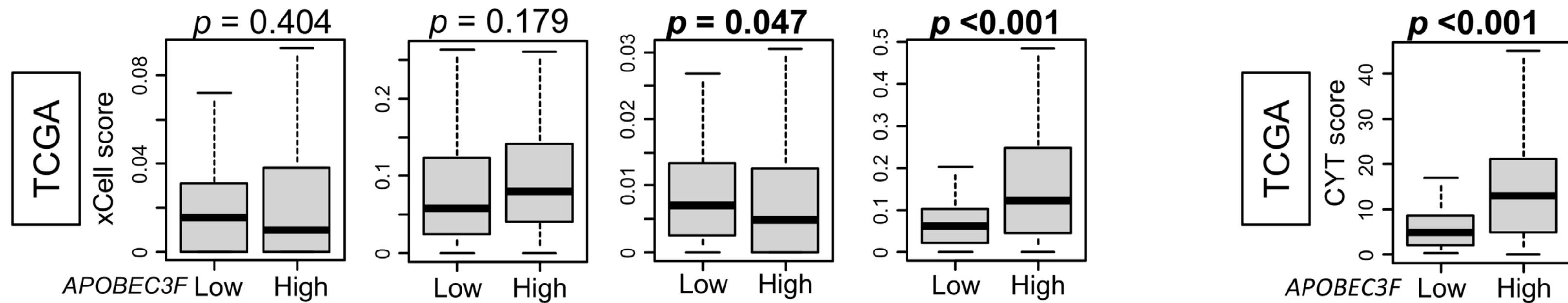
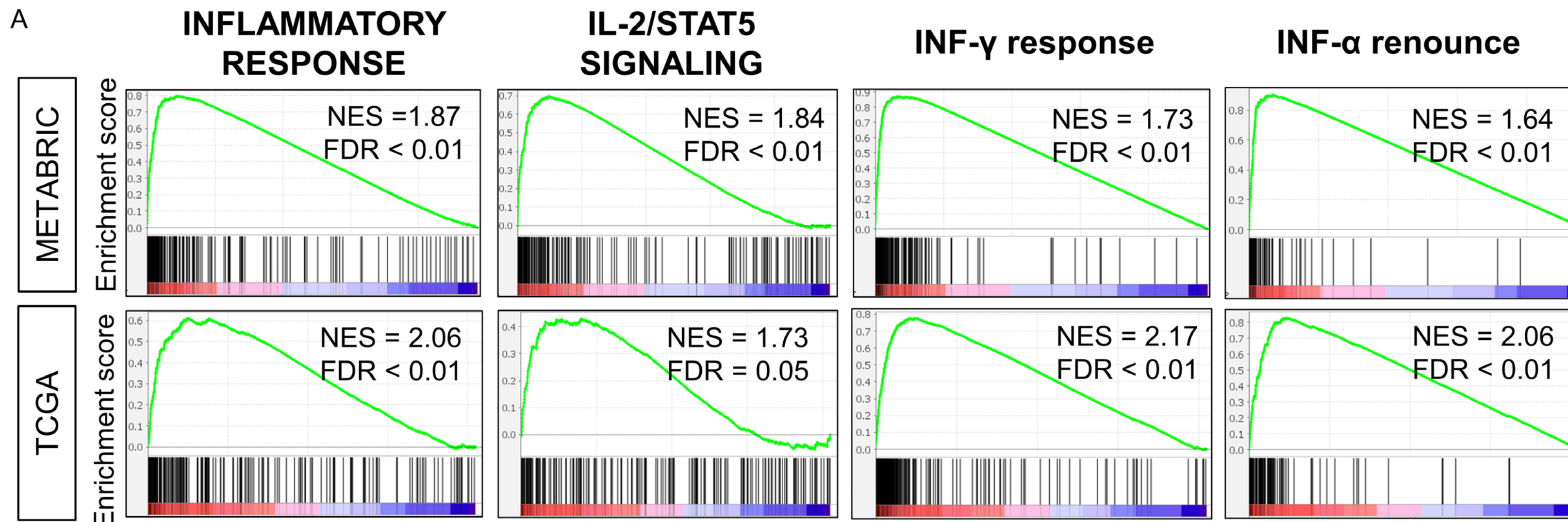


Figure 3. Association of tumor *APOBEC3F* expression with tumor immune composition and cytolytic activity (CYT). Boxplots for low and high *APOBEC3F* expressors of TNBC subtype in METABRIC and TCGA cohorts are shown for (A) anti-cancer immune cells: $CD8^+$ T cells, $CD4^+$ T cells, Th1, M1 macrophages, DC, NK; (B) pro-cancer immune cells: Treg, Th2, M2 macrophages, and B cells; (C) CYT score. High and low APOBEC3F expressors are defined using the intra-cohort bottom tertile as cut-off. Mann-Whitney U test was used to calculate all *P*-values. DC; dendritic cells, NK; natural killer T-cells, Th1; T helper type 1 cells, Th2; T helper type 2 cells, Treg; regulatory T cells.



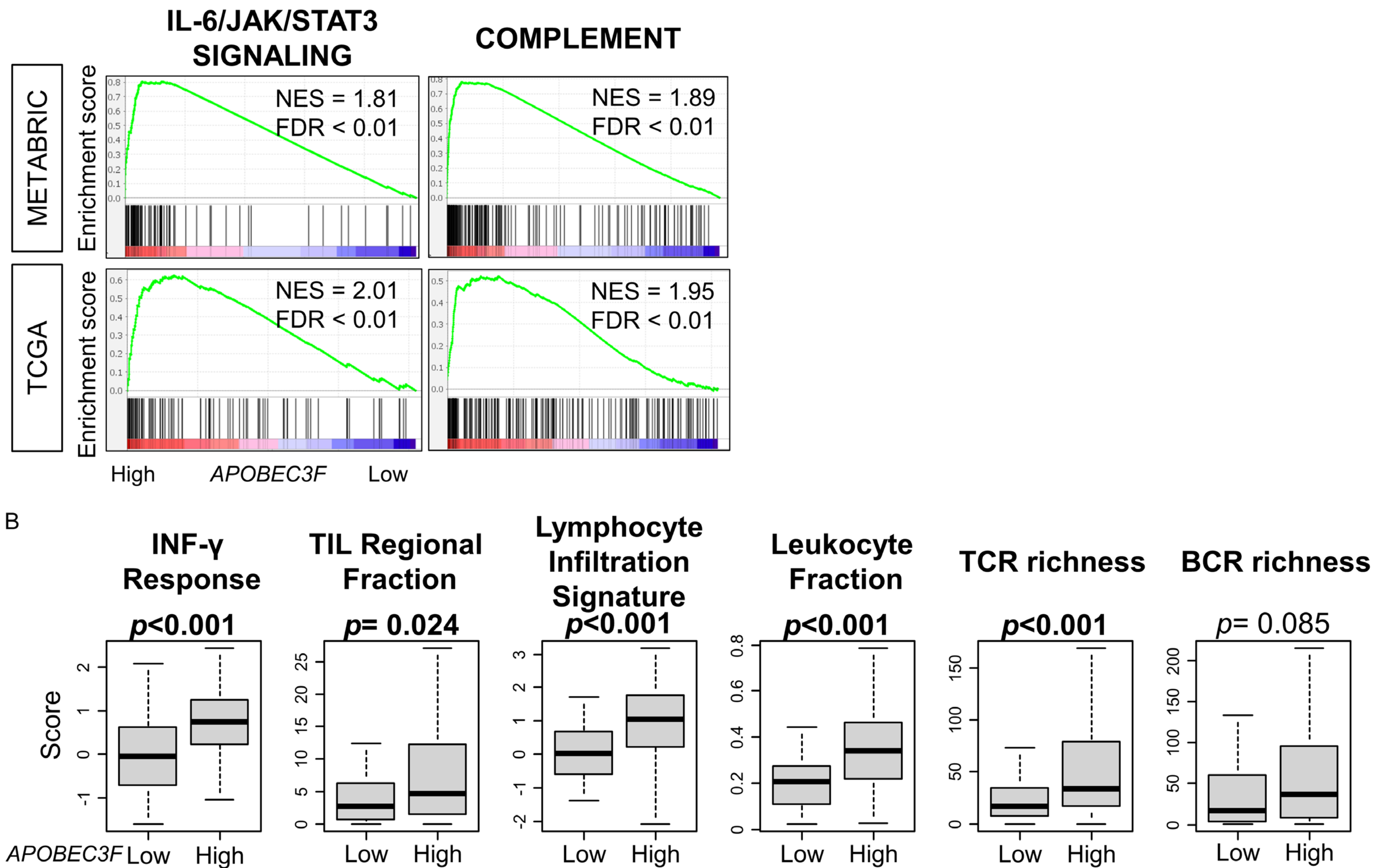
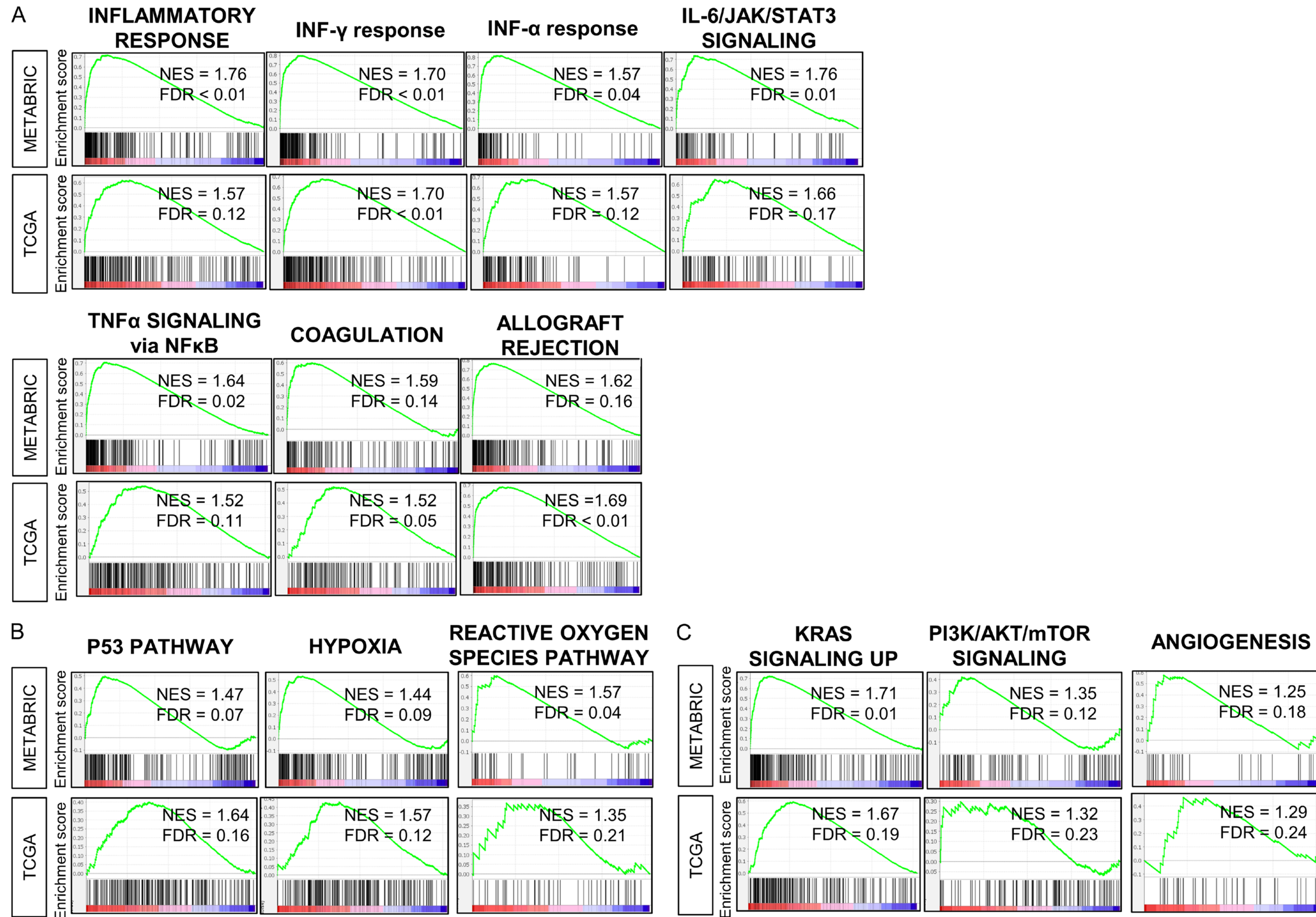


Figure 4. Enriched expression of immune response gene sets among high *APOBEC3F* expressors of TNBC subtype. A. Enrichment plots for Hallmark gene sets with significantly enriched expression in both METABRIC and TCGA cohorts are shown. FDR<0.25 was set as a significance threshold. B. Boxplots for various immune response-related scores among high and low *APOBEC3F* expressors of TNBC subtype in the TCGA cohort. High and low APOBEC3F expressors are defined using the intra-cohort bottom tertile as cut-off. BCR; B cell receptor, FDR; false discovery rate, GSEA; gene set enrichment analysis, IFN; interferon, NES; normalized enrichment score, TCR; T cell receptor, TIL; tumor infiltrating lymphocytes.


A
INFLAMMATORY RESPONSE
INF-γ response
INF-α response
IL-6/JAK/STAT3 SIGNALING
METABRIC
TCGA
Enrichment score
NES = 1.76
FDR < 0.01
NES = 1.70
FDR < 0.01
NES = 1.57
FDR = 0.04
NES = 1.76
FDR = 0.01
NES = 1.57
FDR = 0.12
NES = 1.70
FDR < 0.01
NES = 1.57
FDR = 0.12
NES = 1.66
FDR = 0.17
TNFα SIGNALING via NFκB
COAGULATION
ALLOGRAFT REJECTION
NES = 1.64
FDR = 0.02
NES = 1.59
FDR = 0.14
NES = 1.62
FDR = 0.16
NES = 1.52
FDR = 0.11
NES = 1.52
FDR = 0.05
NES =1.69
FDR < 0.01
B
P53 PATHWAY
HYPOXIA
REACTIVE OXYGEN SPECIES PATHWAY
NES = 1.47
FDR = 0.07
NES = 1.44
FDR = 0.09
NES = 1.57
FDR = 0.04
NES = 1.64
FDR = 0.16
NES = 1.57
FDR = 0.12
NES = 1.35
FDR = 0.21
C
KRAS SIGNALING UP
PI3K/AKT/mTOR SIGNALING
ANGIOGENESIS
NES = 1.71
FDR = 0.01
NES = 1.35
FDR = 0.12
NES = 1.25
FDR = 0.18
NES = 1.67
FDR = 0.19
NES = 1.32
FDR = 0.23
NES = 1.29
FDR = 0.24

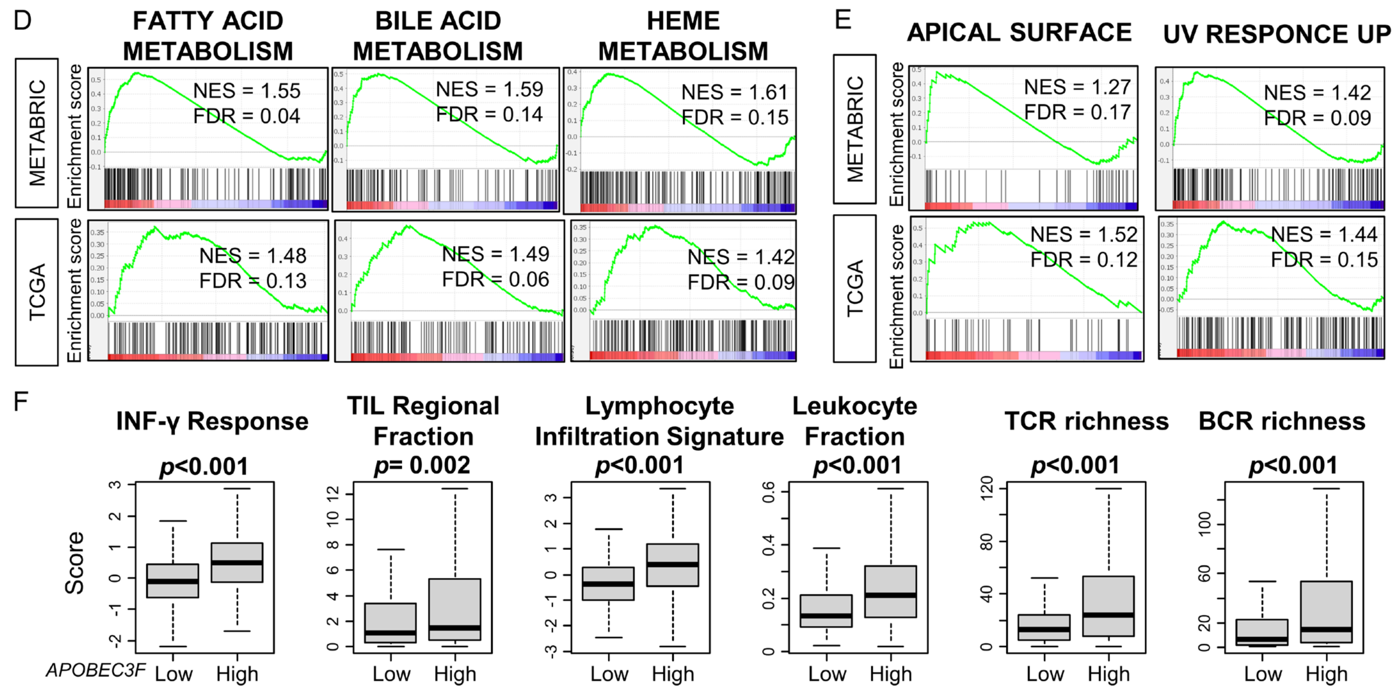


Figure 5. Enriched expression of immune response gene sets among high *APOBEC3F* expressors of luminal subtype. Enrichment plots for Hallmark gene sets with significantly enriched expression in both METABRIC and TCGA cohorts are shown for (A) immune-related gene sets, (B) apoptosis-related gene sets, (C) cancer progression-related gene sets, (D) metabolism-related gene sets, and (E) others. FDR<0.25 was set as significance threshold. (F) Boxplots for various immune response-related scores among high and low *APOBEC3F* expressors of the luminal subtype in the TCGA cohort. High and low APOBEC3F expressors are defined using the intra-cohort bottom tertile as cut-off. BCR; B cell receptor, FDR; false discovery rate, IFN; interferon, NES; normalized enrichment score, TCR; T cell receptor, TIL; tumor infiltrating lymphocytes.

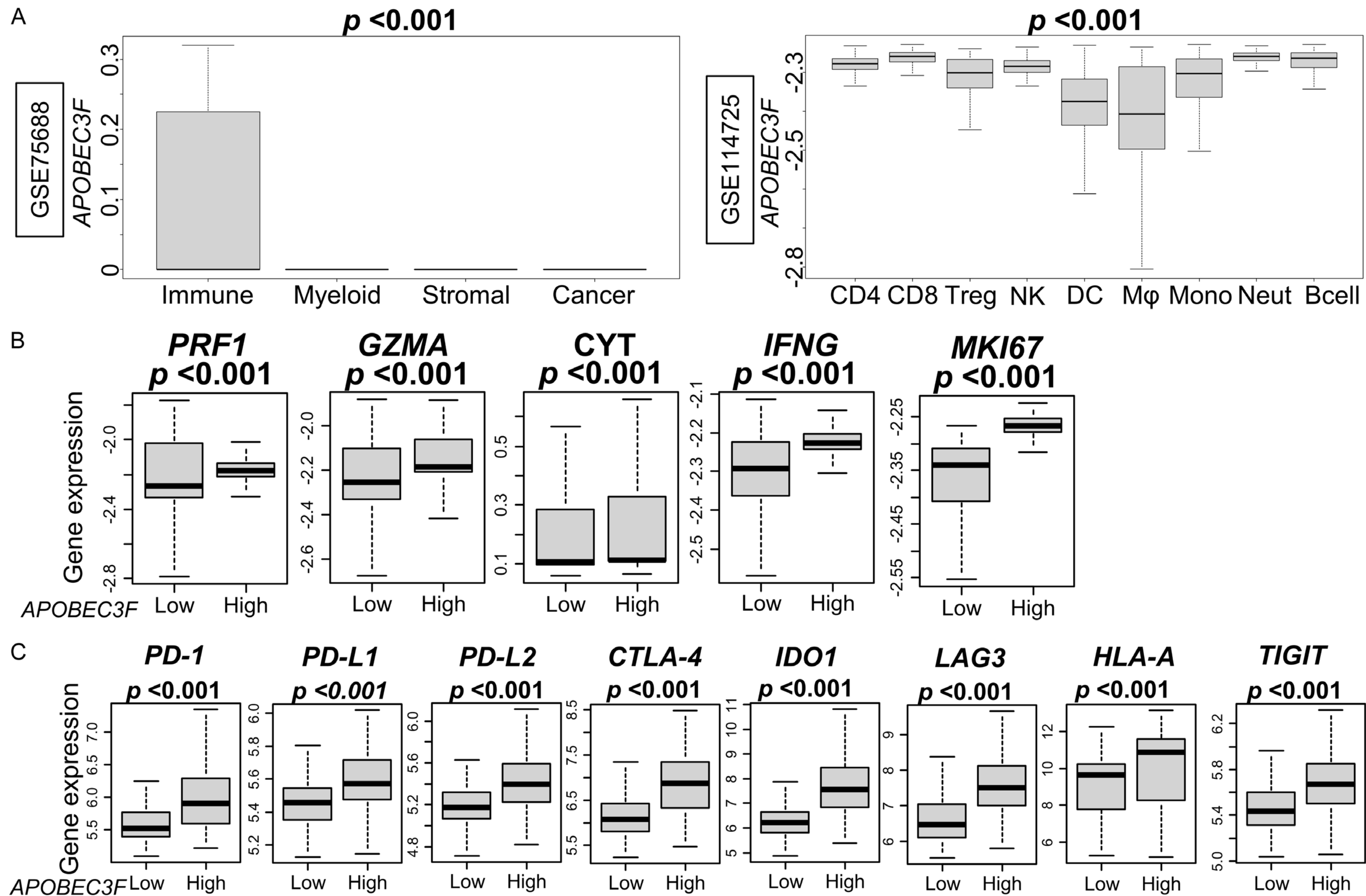


Figure 6. *APOBEC3F* expression in different cells of breast cancer tumors and immune function. Single-cell tumor transcriptome datasets of two breast cancer cohorts were analyzed. (A) Boxplots of *APOBEC3F* gene expression levels of immune, myeloid, stromal, and cancer cells in the GSE75688 cohort, and of immune cell subsets of the GSE114725 dataset. (B) Boxplots showing gene expression in immune cells from the GSE114725 cohort and (C) immune checkpoint gene expressions in TNBC with the low and high *APOBEC3F* groups of the METABRIC cohort. Kruskal-Wallis test and Mann-Whitney U test were used to compare multiple and two factors to calculate *P*-values. *CTLA4*; cytotoxic T-lymphocyte-associated protein 4, DC; dendritic cells, *GZMB*; Granzyme B, *HLA-A;* human leukocyte antigens-A, *IDO1*; indoleamine 2,3-dioxygenase 1, *INFG*; Interferon γ, *LAG3*; lymphocyte-activation gene 3, Mono; monocytes, Mφ; macrophage, Neutro; Neutrophils, NK; Natural killer cells, *PD-1;* programmed death protein-1, *PD-L1*; programmed death-ligand 1, *PD-L2*; programmed death-ligand 2, *PRF1*; Perforin 1, *TIGIT;* T cell immunoreceptor with Ig and ITIM domains, Treg; regulatory T-cells.

To investigate the association of *APOBEC3F* expression with immune function in immune cells, we analyzed the GSE75688 dataset and found that immune cells with high *APOBEC3F* expression also had higher expression of Perforin 1 (*PRF1*) and Granzyme A (*GZMA*), the two genes that determine cytolytic activity score (CYT). Interferon-γ (IFN-γ), reflecting immune response, and MKI67, reflecting cell proliferation, were also elevated in high-*APOBEC3F* expressing cells (**Figure 6B**, $P<0.001$). Our results suggest that *APOBEC3F* is expressed in immune cells, especially effector cells. Furthermore, its expression is also associated with immune function, including cytolytic activity, as well as proliferation in the human breast cancer TME.

Given the increased infiltration of immune cells and its strong correlation with immune response, we expected *APOBEC3F* expression to correspond with immune checkpoint genes. Indeed, various immune checkpoint genes that we investigated were highly expressed in high-*APOBEC3F* groups for both TNBC and luminal subtypes in the METABRIC cohort, including programmed death protein-1 (*PD-1*), cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*), programmed death-ligand 1 (*PD-L1*), *PD-L2*, lymphocyte-activation gene 3 (*LAG3*), indoleamine 2,3-dioxygenase 1 (*IDO1*), human leukocyte antigens-A (*HLA-A*), T cell immunoreceptor with Ig and ITIM domains (*TIGIT*) (**Figure 6C**). The TCGA cohort mirrored this observation (Figure S4). Thus, these results imply an association between global expression of immune check point molecules and high *APOBEC3F* expression in breast cancer.

Discussion

In this study, we demonstrate that in the human breast cancer tumor microenvironment, *APOBEC3F* is exclusively expressed in immune cells that enhance the anti-cancer immune response, including immune cell-mediated cytolysis of cancer cells. High expression of *APOBEC3F* is significantly associated with immune cell infiltration in tumor and anti-cancer immune activity, as well as improved patient survival in TNBC. These observations were obtained using two large independent primary breast cancer cohorts, METABRIC and TCGA. *APOBEC3F* expression was significantly high in TNBC, which among all breast cancer subtypes is the subtype with the most abundant immune cell infiltration. Patient survival (DSS and OS) was consistently better in high-*APOBEC3F* TNBC, which led us to hypothesize that *APOBEC3F* may enhance cancer immunity. In both TNBC and luminal breast cancer, *APOBEC3F* expression did not show association with cancer cell parameters such as tumor clonal heterogeneity, mutation burden, or neoantigen load (except for HRD in TNBC). *APOBEC3F* expression did not correlate with sensitivity to any of the anticancer drugs in any of the cancer cell lines that we investigated. However, *APOBEC3F* expression was associated with tumor infiltration by multiple types of immune cells and with high immune cytolytic activity regardless of cancer subtype. High *APOBEC3F* tumor significantly enriched immune-related gene sets including IL-2/STAT5 signaling, IFN γ response, inflammatory response, TNF-α signaling, and was associated with high levels of immune activity scores in TNBC. However, in luminal breast cancer, not only immune-related gene sets but also a variety of gene sets were enriched including apoptosis-related, metabolism-related, cell proliferation-related, apical surface, and UV response. Using the breast cancer single-cell transcriptome cohorts we found that compared to cancer and stromal cells, *APOBEC3F* was exclusively expressed in immune cells. High-*APOBEC3F* immune cells had significantly higher expression of immune response, including tumoricidal genes, as well as MKI67. Finally, high-*APOBEC3F* tumors expressed significantly higher immune checkpoint genes reflecting T cell infiltration in both TNBC and luminal subtypes.

Although members of the APOBEC3 family have high homology in their structures, their functions differ individually and may even be opposite [46]. These enzymes may activate oncogenes or inactivate tumor suppressors by DNA/RNA deamination to promote tumor heterogeneity and cancer cell growth [47]. In breast cancer hallmark trinucleotide contexts, excessive expression of *APOBEC3A* and *APOBEC3B* induces chromosomal mutations and DNA damage responses leading to those enzymes being regarded as the key mutagenic drivers [48]. *APOBEC3B* is associated with specific somatic mutations such as *TP53* and *PIK3CA* [49, 50] as well as other subclonal mutations that cause immune evasion [51]. Thus, *APOBEC3B* has been suggested to play a role

in the accumulation of somatic genomic changes during carcinogenesis. *APOBEC3B* expression has been reported to be associated with Nottingham histological grade and worse survival [52]. *APOBEC3A*, the most catalytically active of the APOBEC3 family, has also been suggested to be a powerful mutation promoter [53].

Conversely, mutations caused by APOBECs and other factors are also known to induce high immunogenicity and activate cancer immunity [24, 54]. Recent studies have reported that tumor gene expression of *APOBEC3B* correlates with that of *PD-1* in multiple cancer types [55-57]. Kataegis, a specific mutation pattern generated by APOBEC3s, is also associated with *PD-1/PD-2* expression [58]. In non-small cell lung cancer (NSCLC), high expression of *APOBEC3B* has been reported in patients with persistent response to immunotherapy [55]. There is a trade-off between these two effects since each APOBEC3 is linked to immunological and cell proliferative processes in cancer to varying degrees. Indeed, we have previously observed that increased mutational load in breast cancer counterbalances the heightened anti-cancer immunity and aggressive phenotype [59]. Our group has also shown that breast cancers with elevated levels of RNA editing by APOBEC3 have increased aneuploidy and copy number abnormalities, and enriched expression of immune-related, but not cell-related gene sets [60]. We have also found higher immune cells than epithelial cells in higher gene expressing *APOBEC3*. In the current study, we found that *APOBEC3F* activates cancer immunity strong enough to affect survival differences in TNBC as evident from single cell sequence dataset, enrichment of immune-related gene sets, immune cell infiltrations, enhanced immune response, and immune cell killing. Among breast cancer subtypes TNBC is known to be the most aggressive, however, some TNBC with high immunogenicity and abundant TILs reportedly respond well to chemotherapy and have a good prognosis [15-17, 61-64]. Our result that high tumor *APOBEC3F* expression level is associated with better survival may imply that these are the immune-activated TNBC patients, which is consistent with previous reports.

Interestingly, while more immune-related gene sets were enriched in TNBC, immune cell infiltration according to APOBEC3F expression levels did not differ significantly between TNBC and luminal breast cancer. In TNBC, high *APOBEC3F* tumor enriched immune-related gene sets, whereas it enriched a wide range of gene sets, along with cancer proliferation gene sets, in luminal subtype. To this end, *APOBEC3F* expression seems to be heavily tilted toward killing cancer cells in TNBC. Using the single-cell transcriptome datasets, we found that *APOBEC3F* was exclusively expressed in immune cells and significantly related to its immune activity, as well as immune cell proliferation. This suggests that the improved survival prognosis in TNBC patients with high tumor expression of *APOBEC3F* does not simply reflect the neoantigen activity of cancer accompanied by abundant immune cell infiltration, but *APOBEC3F* may have the ability to activate immune cells.

APOBEC3F gene expression was higher in TNBC, but lower in ER+HER2 and HER2+ subtypes compared to normal tissue in TCGA. However, all three subtypes had significantly less APOBEC3F protein when compared to normal tissue in the Clinical Proteomic Tumor Analysis Consortium (CPTAC) Confirmatory/Discovery dataset (personal communication). This further emphasizes the necessity for protein analysis at the tissue and/or cell level in the laboratory.

However, *APOBEC3F* expression was completely unrelated to drug sensitivity in breast cancer cell lines, which agrees with the notion that *APOBEC3F* functions in immune cells and not in cancer cells. Our results further suggest that studies of genes that work on non-cancer cells in the TME require analyses of human patient samples, since cancer cell culture cannot replicate the interactions that occur in vivo, and animal models may not replicate human tumors. We found that immune checkpoint molecules were uniformly elevated in high APOBEC3F breast cancers, which further demonstrates that APOBEC3F is involved in cancer immunity given that immune checkpoint molecules are known to be markers of lymphocyte exhaustion. Whether APOBEC3F expression has the utility as a biomarker for the response to immune checkpoint inhibitors is yet to be determined through future study.

Naturally, we are aware of several limitations in this study. First, this is a retrospective study using a previously collected cohort of patients.

Despite utilizing large dataset cohorts, selection bias exists in patient background and clinicopathological information. Secondly, entire analyses of our study were based on mRNA expression data, which may not be in sync with the APOBEC3F protein expression. Additionally, immune cell infiltration analyzed in this study were all fractions calculated using an algorithm from bulk tumor transcriptome, which were not validated by histology nor single cell isolation of immune cells. Reasons for the opposite or invalidated results for some immune cells between 2 cohorts may be due to differences in the platform used to measure mRNA expression, with METABRIC using gene expression microarray and TCGA using RNA-sequence, as well as TCGA having approximately half the sample size as METABRIC. Moreover, the RNA sequencing in the two single-cell transcriptome datasets that we utilized is not free from coverage bias with low overall gene expression. Finally, since we are incapable of conducting in vitro experiments at this time, we do not know how *APOBEC3F* specifically modifies DNA or RNA in the immune cells, and further studies are needed to determine its specific functions.

In conclusion, our study demonstrates that in breast cancer tumors, the *APOBEC3F* gene is exclusively expressed in immune cells. *APOBEC3F* expression in these tumors is positively associated with enhanced anti-cancer immune response, immune cell proliferation, and immune cell-mediated cytolytic activity in the tumor microenvironment, and with favorable survival in triple-negative breast cancer.

Acknowledgements

The authors would like to thank Dr. Santosh K. Patnaik for his scientific input and editing of this manuscript. This research was supported by National Institutes of Health, USA grant number R37CA248018, R01CA250412, 1R01CA-251545, as well as US Department of Defense BCRP grant number W81XWH-19-1-0674, W81XWH-19-1-0111, as well as Avelas Biosciences grant number AVB620-C-002 to K.T.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Kazuaki Takabe, Roswell Park Comprehensive Cancer Center, Elm & Carlton Streets, Buffalo, New York, NY, USA. Tel: 716-845-5128; E-mail: kazuaki.takabe@roswellpark.org

References

[1] Macduff DA and Harris RS. Directed DNA deamination by AID/APOBEC3 in immunity. Curr Biol 2006; 16: R186-189.

[2] Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, Islam SMA, Lopez-Bigas N, Klimczak LJ, McPherson JR, Morganella S, Sabarinathan R, Wheeler DA, Mustonen V, Group PMSW, Getz G, Rozen SG and Stratton MR; PCAWG Consortium. The repertoire of mutational signatures in human cancer. Nature 2020; 578: 94-101.

[3] Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyfjord JE, Foekens JA, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jager N, Jones DT, Jones D, Knappskog S, Kool M, Lakhani SR, Lopez-Otin C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt AN, Valdes-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain, Zucman-Rossi J, Futreal PA, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo E, Shibata T, Pfister SM, Campbell PJ and Stratton MR. Signatures of mutational processes in human cancer. Nature 2013; 500: 415-421.

[4] Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G, Harris S, Shah RR, Resnick MA, Getz G and Gordenin DA. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet 2013; 45: 970-976.

[5] Leonard B, Hart SN, Burns MB, Carpenter MA, Temiz NA, Rathore A, Vogel RI, Nikas JB, Law EK, Brown WL, Li Y, Zhang Y, Maurer MJ, Oberg AL, Cunningham JM, Shridhar V, Bell DA, April C, Bentley D, Bibikova M, Cheetham RK, Fan JB, Grocock R, Humphray S, Kingsbury Z, Peden J, Chien J, Swisher EM, Hartmann LC, Kalli KR, Goode EL, Sicotte H, Kaufmann SH and Harris RS. APOBEC3B upregulation and genomic mutation patterns in serous ovarian carcinoma. Cancer Res 2013; 73: 7222-7231.

[6] Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, McBride DJ, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jonsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerod A, Tutt A, Martens JW, Aparicio SA, Borg A, Salomon AV, Thomas G, Borresen-Dale AL, Richardson AL, Neuberger MS, Futreal PA, Campbell PJ and Stratton MR; Breast Cancer Working Group of the International Cancer Genome Consortium. Mutational processes molding the genomes of 21 breast cancers. Cell 2012; 149: 979-993.

[7] Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, Miller ML, Rekhtman N, Moreira AL, Ibrahim F, Bruggeman C, Gasmi B, Zappasodi R, Maeda Y, Sander C, Garon EB, Merghoub T, Wolchok JD, Schumacher TN and Chan TA. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 2015; 348: 124-128.

[8] Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, Walsh LA, Postow MA, Wong P, Ho TS, Hollmann TJ, Bruggeman C, Kannan K, Li Y, Elipenahli C, Liu C, Harbison CT, Wang L, Ribas A, Wolchok JD and Chan TA. Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med 2014; 371: 2189-2199.

[9] Takahashi H, Asaoka M, Yan L, Rashid OM, Oshi M, Ishikawa T, Nagahashi M and Takabe K. Biologically aggressive phenotype and anti-cancer immunity counterbalance in breast cancer with high mutation rate. Sci Rep 2020; 10: 1852.

[10] Liedtke C, Mazouni C, Hess KR, Andre F, Tordai A, Mejia JA, Symmans WF, Gonzalez-Angulo AM, Hennessy B, Green M, Cristofanilli M, Hortobagyi GN and Pusztai L. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol 2008; 26: 1275-1281.

[11] Chouliaras K, Tokumaru Y, Asaoka M, Oshi M, Attwood KM, Yoshida K, Ishikawa T and Takabe K. Prevalence and clinical relevance of tumor-associated tissue eosinophilia (TATE) in breast cancer. Surgery 2021; 169: 1234-1239.

[12] Gandhi S, Elkhanany A, Oshi M, Dai T, Opyrchal M, Mohammadpour H, Repasky EA and Takabe K. Contribution of immune cells to glucocorticoid receptor expression in breast cancer. Int J Mol Sci 2020; 21: 4635.

[13] Katsuta E, Rashid OM and Takabe K. Clinical relevance of tumor microenvironment: immune cells, vessels, and mouse models. Hum Cell 2020; 33: 930-937.

[14] Le L, Tokumaru Y, Oshi M, Asaoka M, Yan L, Endo I, Ishikawa T, Futamura M, Yoshida K and Takabe K. Th2 cell infiltrations predict neoadjuvant chemotherapy response of estrogen receptor-positive breast cancer. Gland Surg 2021; 10: 154-165.

[15] Oshi M, Asaoka M, Tokumaru Y, Angarita FA, Yan L, Matsuyama R, Zsiros E, Ishikawa T, Endo I and Takabe K. Abundance of regulatory T cell (Treg) as a predictive biomarker for neoadjuvant chemotherapy in triple-negative breast cancer. Cancers (Basel) 2020; 12: 3038.

[16] Oshi M, Asaoka M, Tokumaru Y, Yan L, Matsuyama R, Ishikawa T, Endo I and Takabe K. CD8 T cell score as a prognostic biomarker for triple negative breast cancer. Int J Mol Sci 2020; 21: 6968.

[17] Oshi M, Newman S, Tokumaru Y, Yan L, Matsuyama R, Kalinski P, Endo I and Takabe K. Plasmacytoid dendritic cell (pDC) infiltration correlate with tumor infiltrating lymphocytes, cancer immunity, and better survival in triple negative breast cancer (TNBC) more strongly than conventional dendritic cell (cDC). Cancers (Basel) 2020; 12: 3342.

[18] Oshi M, Tokumaru Y, Asaoka M, Yan L, Satyananda V, Matsuyama R, Matsuhashi N, Futamura M, Ishikawa T, Yoshida K, Endo I and Takabe K. M1 Macrophage and M1/M2 ratio defined by transcriptomic signatures resemble only part of their conventional clinical characteristics in breast cancer. Sci Rep 2020; 10: 16554.

[19] Yao S, Cheng TD, Elkhanany A, Yan L, Omilian A, Abrams SI, Evans S, Hong CC, Qi Q, Davis W, Liu S, Bandera EV, Odunsi K, Takabe K, Khoury T and Ambrosone CB. Breast tumor microenvironment in black women: a distinct signature of CD8+ T-cell exhaustion. J Natl Cancer Inst 2021; 113: 1036-1043.

[20] Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc EP, Kim J, Kwiatkowski DJ, Fargo DC, Mieczkowski PA, Getz G and Gordenin DA. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet 2015; 47: 1067-1072.

[21] Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, Refsland EW, Kotandeniya D, Tretyakova N, Nikas JB, Yee D, Temiz NA, Donohue DE, McDougle RM, Brown WL, Law EK and Harris RS. APOBEC3B is an enzymatic source of mutation in breast cancer. Nature 2013; 494: 366-370.

[22] Chen Z, Wen W, Bao J, Kuhs KL, Cai Q, Long J, Shu XO, Zheng W and Guo X. Integrative genomic analyses of APOBEC-mutational signature, expression and germline deletion of APOBEC3 genes, and immunogenicity in multiple cancer types. BMC Med Genomics 2019; 12: 131.

[23] Glaser AP, Fantini D, Wang Y, Yu Y, Rimar KJ, Podojil JR, Miller SD and Meeks JJ. APOBEC-mediated mutagenesis in urothelial carcinoma is associated with improved survival, mutations in DNA damage response genes, and immune response. Oncotarget 2018; 9: 4537-4548.

[24] Granadillo Rodriguez M, Flath B and Chelico L. The interesting relationship between APOBEC3 deoxycytidine deaminases and cancer: a long road ahead. Open Biol 2020; 10: 200188.

[25] Zheng YH, Irwin D, Kurosu T, Tokunaga K, Sata T and Peterlin BM. Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. J Virol 2004; 78: 6073-6076.

[26] Richards C, Albin JS, Demir O, Shaban NM, Luengas EM, Land AM, Anderson BD, Holten JR, Anderson JS, Harki DA, Amaro RE and Harris RS. The binding interface between human APOBEC3F and HIV-1 Vif elucidated by genetic and computational approaches. Cell Rep 2015; 13: 1781-1788.

[27] Hu Y, Desimmie BA, Nguyen HC, Ziegler SJ, Cheng TC, Chen J, Wang J, Wang H, Zhang K, Pathak VK and Xiong Y. Structural basis of antagonism of human APOBEC3F by HIV-1 Vif. Nat Struct Mol Biol 2019; 26: 1176-1183.

[28] An P, Penugonda S, Thorball CW, Bartha I, Goedert JJ, Donfield S, Buchbinder S, Binns-Roemer E, Kirk GD, Zhang W, Fellay J, Yu XF and Winkler CA. Role of APOBEC3F gene variation in HIV-1 disease progression and pneumocystis pneumonia. PLoS Genet 2016; 12: e1005921.

[29] Yang Z, Zhuang L, Yu Y, Zhou W, Lu Y, Xu Q, Tang B and Chen X. Overexpression of APOBEC3F in tumor tissues is potentially predictive for poor recurrence-free survival from HBV-related hepatocellular carcinoma. Discov Med 2015; 20: 349-356.

[30] Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R and McKinney S; MATABRIC Group; Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C and Aparicio S. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346-352.

[31] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C and Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6: pl1.

[32] Kim SY, Kawaguchi T, Yan L, Young J, Qi Q and Takabe K. Clinical relevance of microRNA expressions in breast cancer validated using the cancer genome atlas (TCGA). Ann Surg Oncol 2017; 24: 2943-2949.

[33] Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, Nainys J, Wu K, Kiseliovas V, Setty M, Choi K, Fromme RM, Dao P, McKenney PT, Wasti RC, Kadaveru K, Mazutis L, Rudensky AY and Pe'er D. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. Cell 2018; 174: 1293-1308, e1236.

[34] Chung W, Eum HH, Lee HO, Lee KM, Lee HB, Kim KT, Ryu HS, Kim S, Lee JE, Park YH, Kan Z, Han W and Park WY. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat Commun 2017; 8: 15081.

[35] From the American Association of Neurological Surgeons (AANS), American Society of Neuroradiology (ASNR), Cardiovascular and Interventional Radiology Society of Europe (CIRSE), Canadian Interventional Radiology Association (CIRA), Congress of Neurological Surgeons (CNS), European Society of Minimally Invasive Neurological Therapy (ESMINT), European Society of Neuroradiology (ESNR), European Stroke Organization (ESO), Society for Cardiovascular Angiography and Interventions (SCAI), Society of Interventional Radiology (SIR), Society of NeuroInterventional Surgery (SNIS), and World Stroke Organization (WSO), Sacks D, Baxter B, Campbell BCV, Carpenter JS, Cognard C, Dippel D, Eesa M, Fischer U, Hausegger K, Hirsch JA, Shazam Hussain M, Jansen O, Jayaraman MV, Khalessi AA, Kluck BW, Lavine S, Meyers PM, Ramee S, Rufenacht DA, Schirmer CM and Vorwerk D. Multisociety consensus quality improvement revised consensus statement for endovascular therapy of acute ischemic stroke. Int J Stroke 2018; 13: 612-632.

[36] Okano M, Oshi M, Butash AL, Asaoka M, Katsuta E, Peng X, Qi Q, Yan L and Takabe K. Estrogen receptor positive breast cancer with high expression of androgen receptor has less cytolytic activity and worse response to neoadjuvant chemotherapy but better survival. Int J Mol Sci 2019; 20: 2655.

[37] Okano M, Oshi M, Butash AL, Katsuta E, Tachibana K, Saito K, Okayama H, Peng X, Yan L, Kono K, Ohtake T and Takabe K. Triple-negative breast cancer with high levels of an-

nexin A1 expression is associated with mast cell infiltration, inflammation, and angiogenesis. Int J Mol Sci 2019; 20: 4197.
[38] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.
[39] Aran D, Hu Z and Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017; 18: 220.
[40] Tokumaru Y, Oshi M, Katsuta E, Yan L, Huang JL, Nagahashi M, Matsuhashi N, Futamura M, Yoshida K and Takabe K. Intratumoral adipocyte-high breast cancer enrich for metastatic and inflammation-related pathways but associated with less cancer cell proliferation. Int J Mol Sci 2020; 21: 5744.
[41] Takeshita T, Yan L, Asaoka M, Rashid O and Takabe K. Late recurrence of breast cancer is associated with pro-cancerous immune microenvironment in the primary tumor. Sci Rep 2019; 9: 16942.
[42] Narayanan S, Kawaguchi T, Yan L, Peng X, Qi Q and Takabe K. Cytolytic activity score to assess anticancer immunity in colorectal cancer. Ann Surg Oncol 2018; 25: 2323-2331.
[43] Oshi M, Kawaguchi T, Yan L, Peng X, Qi Q, Tian W, Schulze A, McDonald KA, Narayanan S, Young J, Liu S, Morris LG, Chan TA, Kalinski P, Matsuyama R, Otsuji E, Endo I and Takabe K. Immune cytolytic activity is associated with reduced intra-tumoral genetic heterogeneity and with better clinical outcomes in triple negative breast cancer. Am J Cancer Res 2021; 11: 3628-3644.
[44] Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, Sivakumar IKA, Gentles AJ, Malhotra R, Farshidfar F, Colaprico A, Parker JS, Mose LE, Vo NS, Liu J, Liu Y, Rader J, Dhankani V, Reynolds SM, Bowlby R, Califano A, Cherniack AD, Anastassiou D, Bedognetti D, Mokrab Y, Newman AM, Rao A, Chen K, Krasnitz A, Hu H, Malta TM, Noushmehr H, Pedamallu CS, Bullman S, Ojesina AI, Lamb A, Zhou W, Shen H, Choueiri TK, Weinstein JN, Guinney J, Saltz J, Holt RA, Rabkin CS; Cancer Genome Atlas Research Network, Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG and Shmulevich I. The immune landscape of cancer. Immunity 2019; 51: 411-412.
[45] Schumacher TN and Schreiber RD. Neoantigens in cancer immunotherapy. Science 2015; 348: 69-74.
[46] Ng JCF, Quist J, Grigoriadis A, Malim MH and Fraternali F. Pan-cancer transcriptomic analysis dissects immune and proliferative functions of APOBEC3 cytidine deaminases. Nucleic Acids Res 2019; 47: 1178-1194.
[47] Petljak M and Maciejowski J. Molecular origins of APOBEC-associated mutations in cancer. DNA Repair (Amst) 2020; 94: 102905.
[48] Nik-Zainal S, Wedge DC, Alexandrov LB, Petljak M, Butler AP, Bolli N, Davies HR, Knappskog S, Martin S, Papaemmanuil E, Ramakrishna M, Shlien A, Simonic I, Xue Y, Tyler-Smith C, Campbell PJ and Stratton MR. Association of a germline copy number polymorphism of APOBEC3A and APOBEC3B with burden of putative APOBEC-dependent mutations in breast cancer. Nat Genet 2014; 46: 487-491.
[49] Silwal-Pandit L, Vollan HK, Chin SF, Rueda OM, McKinney S, Osako T, Quigley DA, Kristensen VN, Aparicio S, Borresen-Dale AL, Caldas C and Langerod A. TP53 mutation spectrum in breast cancer is subtype specific and has distinct prognostic relevance. Clin Cancer Res 2014; 20: 3569-3580.
[50] Henderson S, Chakravarthy A, Su X, Boshoff C and Fenton TR. APOBEC-mediated cytosine deamination links PIK3CA helical domain mutations to human papillomavirus-driven tumor development. Cell Rep 2014; 7: 1833-1841.
[51] Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, Seth S, Chow CW, Cao Y, Gumbs C, Gold KA, Kalhor N, Little L, Mahadeshwar H, Moran C, Protopopov A, Sun H, Tang J, Wu X, Ye Y, William WN, Lee JJ, Heymach JV, Hong WK, Swisher S, Wistuba II and Futreal PA. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science 2014; 346: 256-259.
[52] Sieuwerts AM, Willis S, Burns MB, Look MP, Meijer-Van Gelder ME, Schlicker A, Heideman MR, Jacobs H, Wessels L, Leyland-Jones B, Gray KP, Foekens JA, Harris RS and Martens JW. Elevated APOBEC3B correlates with poor outcomes for estrogen-receptor-positive breast cancers. Horm Cancer 2014; 5: 405-413.
[53] Langenbucher A, Bowen D, Sakhtemani R, Bournique E, Wise JF, Zou L, Bhagwat AS, Buisson R and Lawrence MS. An extended APOBEC3A mutation signature in cancer. Nat Commun 2021; 12: 1602.
[54] Boichard A, Tsigelny IF and Kurzrock R. High expression of PD-1 ligands is associated with kataegis mutational signature and APOBEC3 alterations. Oncoimmunology 2017; 6: e1284719.
[55] Wang S, Jia M, He Z and Liu XS. APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. Oncogene 2018; 37: 3924-3936.
[56] Cescon DW, Haibe-Kains B and Mak TW. APOBEC3B expression in breast cancer re-

flects cellular proliferation, while a deletion polymorphism is associated with immune activation. Proc Natl Acad Sci U S A 2015; 112: 2841-2846.
[57] Wen WX, Soo JS, Kwan PY, Hong E, Khang TF, Mariapun S, Lee CS, Hasan SN, Rajadurai P, Yip CH, Mohd Taib NA and Teo SH. Germline APOBEC3B deletion is associated with breast cancer risk in an Asian multi-ethnic cohort and with immune cell presentation. Breast Cancer Res 2016; 18: 56.
[58] Purwanto I, Heriyanto DS, Ghozali A, Widodo I, Dwiprahasto I, Aryandono T and Haryana SM. Overexpression of programmed death-ligand 1 receptor mRNA as an independent negative prognostic factor for triple negative breast cancer. World J Oncol 2020; 11: 216-222.
[59] Schmid P, Rugo HS, Adams S, Schneeweiss A, Barrios CH, Iwata H, Dieras V, Henschel V, Molinero L, Chui SY, Maiya V, Husain A, Winer EP, Loi S and Emens LA; IMpassion130 Investigators. Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2020; 21: 44-59.
[60] Asaoka M, Ishikawa T, Takabe K and Patnaik SK. APOBEC3-mediated RNA editing in breast cancer is associated with heightened immune activity and improved survival. Int J Mol Sci 2019; 20: 5621.
[61] Oshi M, Newman S, Murthy V, Tokumaru Y, Yan L, Matsuyama R, Endo I and Takabe K. ITPKC as a prognostic and predictive biomarker of neoadjuvant chemotherapy for triple negative breast cancer. Cancers (Basel) 2020; 12: 2758.
[62] Oshi M, Newman S, Tokumaru Y, Yan L, Matsuyama R, Endo I and Takabe K. Inflammation is associated with worse outcome in the whole cohort but with better outcome in triple-negative subtype of breast cancer patients. J Immunol Res 2020; 2020: 5618786.
[63] Katsuta E, Yan L, Opyrchal M, Kalinski P and Takabe K. Cytotoxic T-lymphocyte infiltration and chemokine predict long-term patient survival independently of tumor mutational burden in triple-negative breast cancer. Ther Adv Med Oncol 2021; 13: 17588359211006680.
[64] Oshi M, Angarita FA, Tokumaru Y, Yan L, Matsuyama R, Endo I and Takabe K. A novel three-gene score as a predictive biomarker for pathologically complete response after neoadjuvant chemotherapy in triple-negative breast cancer. Cancers (Basel) 2021; 13: 2401.

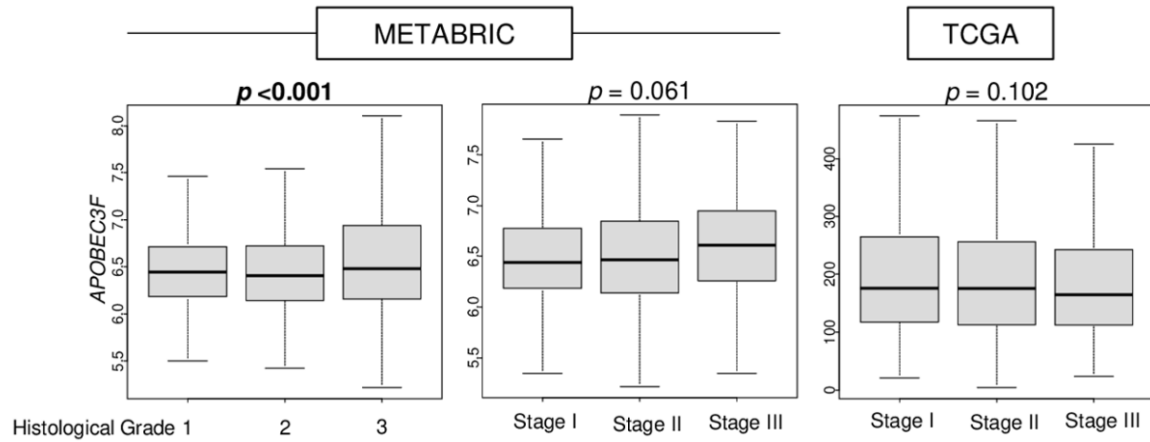


Figure S1. Association between clinical features in the groups with low and high *APOBEC3F* expression. Boxplots show *APOBEC3F* expression by Nottingham histological grade (Grade I, II and III) and the American Joint Committee on Cancer (AJCC) cancer staging (Stage I-III) in the METABRIC and TCGA cancer cohort. Kruskal-Wallis test was used to calculate *P*-values.

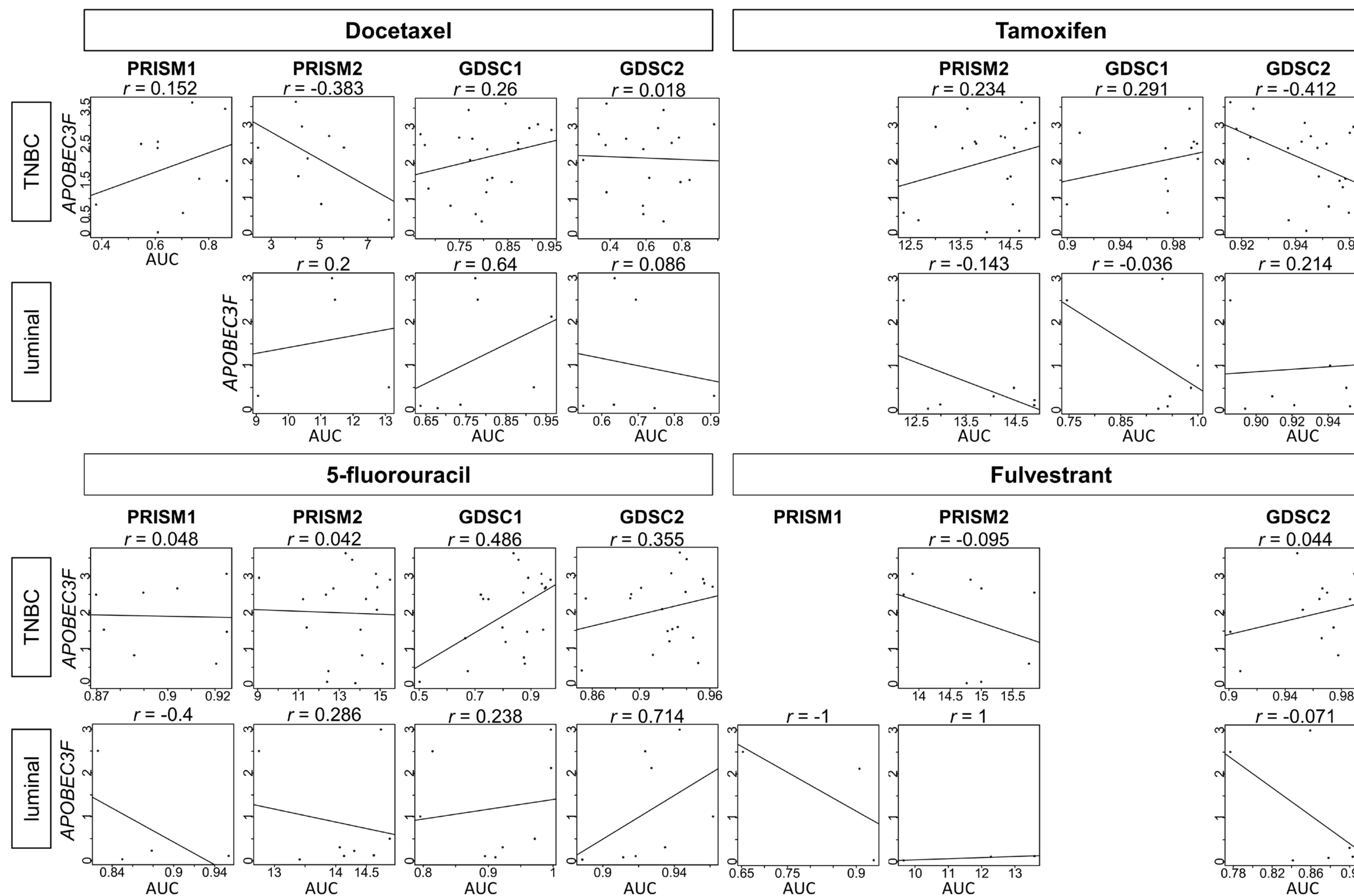


Figure S2. Correlation plots of *APOBEC3F* expression in cell lines and drug sensitivity AUC from different experimental results; docetaxel, 5-fluorouracil, tamoxifen, fulvestrant in TNBC and luminal breast cancer cell lines. Cell line data are the same as in **Figure 2** and were downloaded from the Depmap portal. The following experimental data for each drug were all downloaded from the Depmap portal, PRISM1; PRISM Repurposing 19Q3, PRISM2; PRISM Repurposing 19Q4, GDSC1; Sanger GDSC1, GDSC2; Sanger GDSC2. Correlation coefficients were derived from the Spearman rank test.

APOBEC3F is associated with cancer immunity and better survival in TNBC

A

Anti-cancer immune cells

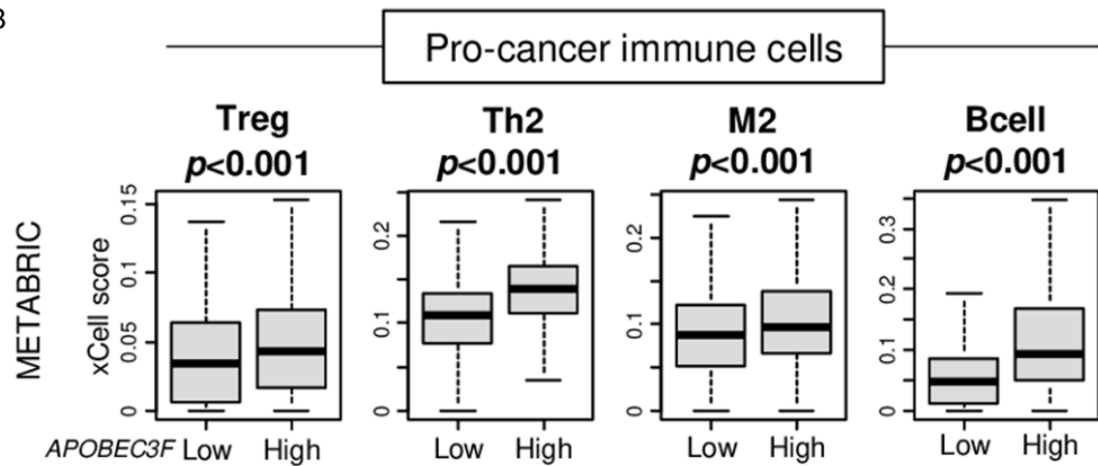
CD8+T $p<0.001$ | CD4+T $p<0.001$ | Th1 $p<0.001$ | M1 $p<0.001$ | DC $p<0.001$ | NK $p<0.001$

METABRIC, xCell score, APOBEC3F Low / High

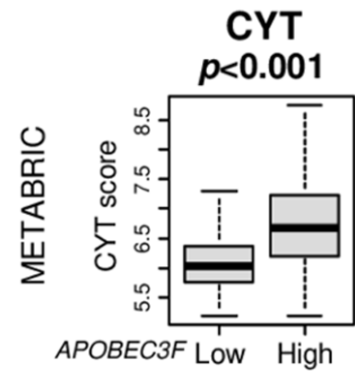
TCGA: $p<0.001$ | $p=0.007$ | $p=0.004$ | $p<0.001$ | $p<0.001$ | $p=0.235$

TCGA, xCell score, APOBEC3F Low / High

B



C



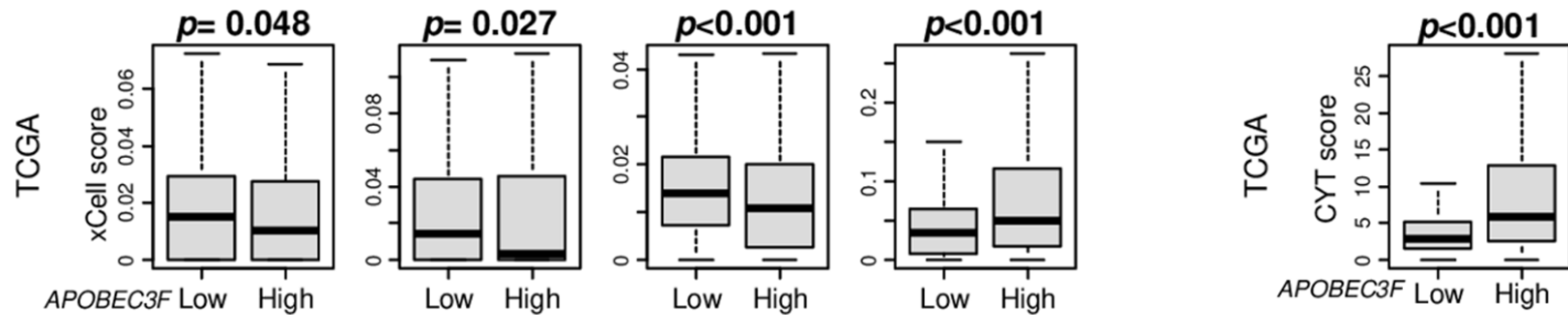


Figure S3. Association of tumor *APOBEC3F* expression with tumor immune composition and cytolytic activity (CYT). Boxplots for low and high APOBEC3F expressors of luminal breast cancer subtype in METABRIC and TCGA cohorts are shown for (A) anti-cancer immune cells: $CD8^+$ T cells, $CD4^+$ T cells, Th1, M1 macrophages, DC, NK; (B) pro-cancer immune cells: Treg, Th2, M2 macrophages, and B cells; (C) CYT score. High and low *APOBEC3F* expressors are defined using the intra-cohort bottom tertile as cut-off. Mann-Whitney U test was used to calculate all *P*-values. DC; dendritic cells, NK; natural killer T-cells, Th1; T helper type 1 cells, Th2; T helper type 2 cells, Treg; regulatory T cells.

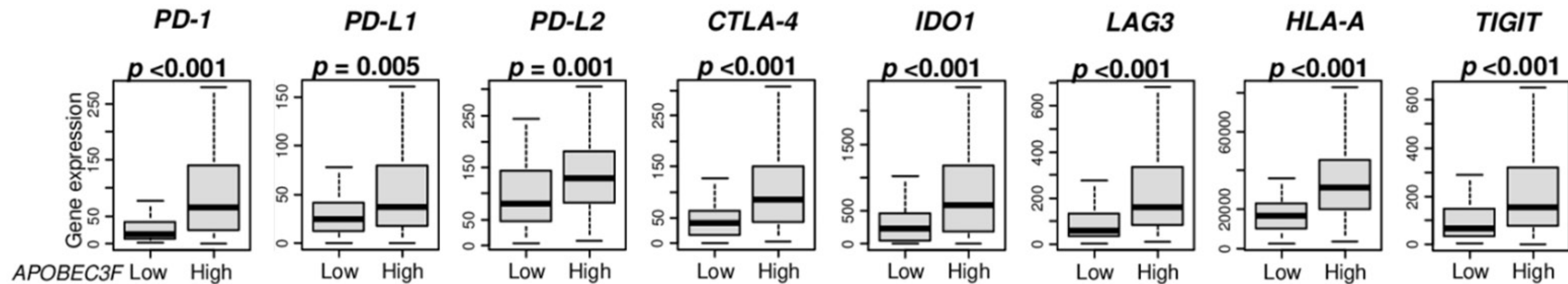


Figure S4. Immune checkpoint gene expressions in TNBC with the low and high *APOBEC3F* groups of the TCGA cohort. Mann-Whitney U test was used to calculate *P*-values. *CTLA4*; cytotoxic T-lymphocyte-associated protein 4, *HLA-A;* human leukocyte antigens-A, *IDO1*; indoleamine 2,3-dioxygenase 1, *LAG3*; lymphocyte-activation gene 3, *PD-1;* programmed death protein-1, *PD-L1*; programmed death-ligand 1, *PD-L2*; programmed death-ligand 2, *TIGIT;* T cell immunoreceptor with Ig and ITIM domains.