APOBEC3G expression correlates with T cell infiltration and improved clinical outcomes in high-grade serous ovarian carcinoma

Brandon Leonard,1,2,* Gabriel J. Starrett,1,2,* Matthew J. Maurer,3 Ann L. Oberg,3 Mieke Van Bockstal,4 Jo Van Dorpe,4 Olivier De Wever,5 Jozien Helleman,6 Anieta M. Sieuwerts,6 Els M.J.J. Berns,6 John W.M. Martens,6 Brett D. Anderson,1,2 William L. Brown,1,2 Kimberly R. Kalli,7 Scott H. Kaufmann,8,9 Reuben S. Harris1,2,10

1Biochemistry, Molecular Biology and Biophysics Department, University of Minnesota, Minneapolis, MN 55455, USA. 2Masonic Cancer Center, University of Minnesota, Minneapolis, MN 55455, USA. 3Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA. 4Department of Pathology, Ghent University Hospital, Ghent, Belgium. 5Department of Radiation Oncology and Experimental Cancer Research, Ghent University and Ghent University Hospital, Ghent, Belgium. 6Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, Netherlands. 7Women’s Cancer Program, Mayo Clinic Cancer Center, Rochester, MN 55905 USA. 8Division of Oncology Research, Department of Oncology, Mayo Clinic, Rochester, MN 55905 USA. 9Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN 55905, USA. 10Howard Hughes Medical Institute, University of Minnesota, Minneapolis, MN 55455, USA.

* These authors contributed equally to this work

Running Title: High APOBEC3G expression in TILs in ovarian cancer

Keywords: APOBEC3G; immune cell infiltration; ovarian cancer; ovarian tumor heterogeneity; prognostic biomarker
Address correspondence to: Reuben S. Harris, Howard Hughes Medical Institute, University of Minnesota, 2231 6th St. S.E., Minneapolis, MN 55455, USA.
Phone: 612.624.0457; Fax: 612.625.2163; E-mail: rsh@umn.edu

Conflict of Interest Statement: R.S.H. is a co-founder of ApoGen Biotechnologies Inc. The other authors have no conflicts of interest to disclose.

Words (excluding references): 4,303/5,000; Tables: 2; Figures: 4; References: 50/50; Supplementary Items: 5

Translational Relevance

Ovarian cancer is the deadliest cancer of the female reproductive tract, partly due to a lack of biomarkers for early clinical detection and assessment of prognosis. Several members of the APOBEC family of antiviral DNA cytosine deaminases, including APOBEC3G, are expressed broadly in human tissues, and APOBEC3B is overexpressed in many tumor types including ovarian cancer. Here, we show that APOBEC3G is expressed at surprisingly high levels in T lymphocytes within high-grade serous ovarian tumors. APOBEC3G expression correlates positively with improved patient outcomes. The latter result is consistent with prior work demonstrating that T cell infiltration correlates with improved outcomes. Our results are the first to demonstrate that high APOBEC3G levels in ovarian tumors are attributable to T cell infiltration and not, for instance, expression in other normal cells or tumor cells. Thus, APOBEC3G is a new candidate biomarker for T cell infiltration and positive anti-tumor immune responses.
Abstract

**Purpose:** APOBEC3 DNA cytosine deaminase family members normally defend against viruses and transposons. However, deregulated APOBEC3 activity causes mutations in cancer. Due to broad expression profiles and varying mixtures of normal and cancer cells in tumors, including immune cell infiltration, it is difficult to determine where different APOBEC3s are expressed. Here, we ask whether correlations exist between APOBEC3 expression and T cell infiltration in high-grade serous ovarian cancer (HGSOC), and assess whether these correlations have prognostic value.

**Experimental Design:** Transcripts for APOBEC3G, APOBEC3B, and the T cell markers, CD3D, CD4, CD8A, GZMB, PRF1, and RNF128 were quantified by RT-qPCR for a cohort of 354 HGSOC patients. Expression values were correlated with each other and clinical parameters. Two additional cohorts were used to extend HGSOC clinical results. Immuno-imaging was used to colocalize APOBEC3G and the T cell marker CD3. TCGA data extended expression analyses to additional cancer types.

**Results:** A surprising positive correlation was found for expression of APOBEC3G and several T cell genes in HGSOC. Immunohistochemistry and immunofluorescent imaging showed protein colocalization in tumor-infiltrating T lymphocytes. High APOBEC3G expression correlated with improved outcomes in multiple HGSOC cohorts. TCGA data analyses revealed that expression of APOBEC3D and APOBEC3H also correlates with CD3D across multiple cancer types.

**Conclusions:** Our results identify APOBEC3G as a new candidate biomarker for tumor-infiltrating T lymphocytes and favorable prognoses for HGSOC. Our data also highlight the complexity of the tumor environment with respect to differential APOBEC family gene expression in both tumor and surrounding normal cell types.
Introduction

Ovarian cancer is the deadliest gynecological malignancy worldwide (1). The most common type of ovarian cancer, high-grade serous ovarian carcinoma (HGSOC), accounts for over 60% of cases, and is the most aggressive reproductive tract malignancy (2). Due to the lack of efficient detection methods, HGSOC generally presents at advanced stages and is associated with high rates of recurrence and mortality. Interestingly, several studies have identified T cell infiltration as a favorable prognostic factor for HGSOC (3-7). Additional markers of T cell infiltration and particularly those reflecting high-quality anti-tumor responses are needed to fully realize the clinical impact of this finding.

The APOBEC enzymes are an 11-member family of zinc-coordinating enzymes that convert cytosines to uracils (C-to-U) in ssDNA (8). The enzymatic activity of specific family members is essential for both adaptive (AID) and innate immune responses [APOBEC3 subfamily members; (9) and references therein]. AID is well-studied due to its known roles in antibody diversification through somatic hypermutation and class switch recombination in B cells (10). Other well-studied family members include APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H because of demonstrated anti-HIV-1 activity (11). These enzymes are capable of introducing C-to-U lesions in viral cDNA intermediates that manifest as G-to-A mutations in proviral genomes (12-14). APOBEC3 subfamily members have also been implicated in restricting the replication of many other DNA-based parasites including transposable elements [(9) and references therein]. It is important to note that most APOBEC family members are expressed broadly and constitutively (12,15-17), but several also can be further upregulated by specific conditions such as APOBEC3A by interferon-α (18,19), APOBEC3B by non-canonical NF-κB activation and HPV infection (20-22), and APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H by the combination of T cell activation and HIV-1 infection (13,14).
While the APOBEC3 enzymes have been shown to have important physiologic roles in protecting cells from endogenous and exogenous DNA-based pathogens, their dysregulation has also been linked to cancer. For instance, AID has been linked to various B cell malignancies [reviewed in (23)]. A second prominent example is the recent finding that APOBEC3B is overexpressed and a significant source of mutation in breast, ovarian, and several other cancers [(17,24-27) and reviewed in (28-30)]. APOBEC3B deaminates cytosines in genomic DNA to produce promutagenic uracil lesions, which can result in mutations if they are not repaired faithfully. APOBEC3B is overexpressed and its mutation signature is overrepresented in genomic sequences of HGSOC and other cancers. Elevated APOBEC3B expression levels have also been linked to poor prognosis in multiple cancer types (31-36).

In addition to APOBEC3B, related subfamily members have also been implicated in cancer mutagenesis to varying degrees (37-40). Because most APOBEC family members are expressed in many normal cell types and tissues, a major confounding factor in quantification of APOBEC expression levels in tumors is cellular heterogeneity due to surrounding normal tissue and immune cell infiltration. To address these issues, we quantified APOBEC expression in a large cohort of HGSOC patients and asked whether expression levels correlated with immune cell infiltration and could be distinguished from surrounding normal tissues. A strong positive correlation was identified between APOBEC3G expression and markers of T cell infiltration, and co-expression was confirmed by immunohistochemical and immunofluorescent staining of primary HGSOC specimens. Moreover, high APOBEC3G expression levels correlated significantly with improved prognosis. These findings were extended to multiple cancer types through an analysis of publically available RNA sequencing (RNAseq) data from The Cancer Genome Atlas (TCGA). Collectively, these results highlight the complexity of APOBEC3 family member expression in HGSOC specimens comprised of tumor cells, surrounding normal tissues, and...
in many instances infiltrating immune cells. Our studies are the first to identify APOBEC3G as a new candidate biomarker for effective T cell responses and potentially for immunotherapies against HGSOC.

Materials and Methods

Analysis of expression correlations in Mayo Clinic cohort

Primary tumor samples from 354 HGSOC cases at the Mayo Clinic were selected based on histology, grade, stage, and availability of clinical outcome data (IRB #13-002487). Following cryosectioning of each snap frozen specimen, TRIzol based RNA extractions were performed. cDNA was synthesized in triplicate using Transcriptor Reverse Transcriptase (Roche) and reverse transcription quantitative PCR (RT-qPCR) for APOBEC3G, APOBEC3B, CD3D, CD4, CD8A, GZMB, PRF1, RNF128, and TBP was performed using the primer and probe combinations listed in Table S1 (validation in Fig. S1). Correlations between APOBEC3G, APOBEC3B, and the various T cell markers were determined using Spearman’s correlation. Spearman’s correlation coefficient (r$_S$) and p-values are reported.

Immunohistochemistry

IHC for CD3, CD4, CD8, and APOBEC3G was performed on 7 paraffin embedded primary HGSOC specimens obtained from patients who underwent debulking surgery at Ghent University Hospital. This study was approved by the ethics committee of the Ghent University Hospital. The tissues were fixed in a 4% formaldehyde solution for 12-48 hours and embedded in paraffin. IHC was performed on 3.5 μm tissue sections on Superfrost slides (Menzel-Gläser) using a Benchmark XT automated slide stainer (Roche), according to the manufacturer’s instructions. The following antibodies were used: mouse monoclonal anti-CD3 (clone F7.2.38, dilution 1/10, Dako), rabbit monoclonal anti-CD4 (clone SP35,
dilution 1/25, Cell Marque), mouse monoclonal anti-CD8 (clone C8/144B, no dilution, NeoMarkers), and a locally developed rabbit monoclonal anti-APOBEC3G [clone 5211-110-19, dilution 1/50]. Although this rabbit monoclonal antibody recognizes APOBEC3G, APOBEC3A, and APOBEC3B due to unavoidable homology, we are confident that it is only detecting APOBEC3G in the immunohistochemistry studies described here because APOBEC3A is not expressed in T lymphocytes and, for reasons still under investigation, this monoclonal antibody does not recognize endogenous APOBEC3B by these procedures. It is important to emphasize that APOBEC3A is not expressed in most normal cell types, and it is only induced to detectable levels in interferon-activated myeloid lineage cells (12,15,17,18). Moreover, comparatively few CD68-positive macrophages were detected in the HGSOC specimens described here and positive signals from APOBEC3A or APOBEC3G expression in this cell type are minor (data not shown). Heat-induced epitope retrieval was done using Cell Conditioning 1 (Roche) for CD3, CD8, and APOBEC3G, and using Cell Conditioning 2 (Roche) for CD4. Visualization of all primary antibodies was achieved with the ultraViewTM Universal DAB Detection Kit (Roche). Counterstaining with hematoxylin, dehydration of the tissue sections, and application of coverslips were carried out using an automated coverslipper (Tissue-Tek).

Immunofluorescent imaging

Fluorescence-based co-localization experiments were done using a subset of the same HGSOC specimens used above for IHC following published procedures (41) After sample preparation, permeabilization, and blocking with 10% goat serum in PBS at room temperature (DAKO X0907), each slide was stained first with rabbit polyclonal anti-APOBEC3G (Sigma Atlas Antibody HPA001812) diluted 1/50 in PBS (RT, 1 hr), second with mouse monoclonal anti-CD3 (described above) diluted 1/10 in PBS (4°C, overnight), and finally with a combination of secondary antibodies diluted 1/500 in PBS [RT, 1 hr; anti-rabbit
IgG-AlexaFluor 594 (Invitrogen A11012) and anti-mouse IgG1-AlexaFluor 488 (Invitrogen A21121)]. Finally, slides were stained with DAPI diluted 1/5000 in methanol (RT, 5 min), mounted with a cover slip, and imaged using a fluorescence microscope equipped with appropriate filters (Olympus BX40, Tokyo, Japan). Multiple PBS washes were done between each step of the procedures.

Survival analysis in Mayo Clinic cohort

Kaplan-Meier curves were constructed by dividing specimens at the median expression level for each gene. OS data was available for all 354 patients, while PFS data was only available for 348 patients. P-values, hazard ratios and 95% confidence intervals were determined using Cox regression models on the continuous log2-transformed expression, adjusting for stage and debulking status.

Survival analysis in additional cohorts

Ovarian cancer data from TCGA and Gene Expression Omnibus (GEO) were compiled by the 2015 version of the Kaplan-Meier plotter database on a PostgreSQL server (42). GEO accession numbers were GSE14764, GSE15622, GSE18520, GSE19829, GSE23554, GSE26193, GSE26712, GSE27651, GSE30161, GSE3149, GSE51373, and GSE9891. All gene expression data were determined using only the HG-U133A, HG-U133 Plus 2.0, and HG-U133A 2.0 Affymetrix microarray platforms so that comparisons could be made between datasets. Specifically, APOBEC3G expression was determined using the 204205_at probe. Grade 3 serous ovarian cancers were the only samples used in this analysis. OS data were available for 856 patients, while PFS data was available for 753 patients. Kaplan-Meier plots were constructed and p-values, hazard ratios and 95% confidence intervals were calculated using the Mantel-Cox log-rank test.
Medical ethics approval for the Dutch cohort was obtained in part previously (43) (n=37) and in part more recently (n=36 and n=15; MEC-2008-183). APOBEC3G mRNA levels were measured by RTqPCR using an assay on demand (Hs00222415_m1, Applied Biosystems) and three reference genes were measured with the primers listed in Table S1 and quantification using SYBR green. Relative APOBEC3G expression levels were determined by normalization to the average of 3 reference genes as described (31). Kaplan-Meier curves were constructed by dividing specimens at the median expression level for each gene. P-values, hazard ratios and 95% confidence intervals were determined using the Mantel-Cox log-rank test.

**TCGA analysis**

Normalized RNAseqV2 data were downloaded from TCGA in July 2015. APOBEC3, CD3D, and CD20 mRNA levels were quantified using normalized read counts. rs and p-values for linear models of APOBEC versus immune-marker genes were calculated using Spearman’s rank correlation coefficient with the R statistical environment. Cancer types were grouped by hierarchical clustering (hclust) of the rs values for each APOBEC family member in the R statistical environment, and these results were used to generate a dendrogram of these relationships. All data were graphed using the ggplot2 R package (44). P-values were calculated from the rs values and adjusted for multiple comparisons using the Benjamini-Hochberg method and significance was defined as a p-value less than 0.05.

**Results**

**APOBEC3G expression correlates with activated T lymphocyte infiltration in HGSOC**

APOBEC3G is expressed constitutively in many different cell lines and tissue types and is also known to be upregulated by HIV-1 infection of primary T lymphocytes, which is
one of many distinct mechanisms of immune cell activation [e.g., (12-14)]. Although virus infection is unlikely to be part of the etiology of ovarian cancer (45), the presence of activated immune cells (infiltrate) in HGSOC is known to correlate with better overall outcomes, most likely due to strong anti-tumor immune responses (3-7). We therefore asked whether APOBEC3G expression correlated with T cell infiltration and clinical outcomes using a cohort of 354 primary HGSOC samples procured at the Mayo Clinic (clinical characteristics in Table 1).

RNA was prepared from fresh frozen HGSOC tissues, and a previously validated, highly specific RT-qPCR assay was used to assess APOBEC3G mRNA levels (12). In parallel, expression of the related gene APOBEC3B, which has been implicated in ovarian cancer genome mutagenesis (24), was assayed. In addition, the mRNA levels of several established T cell markers were quantified, including CD3D (total T cells), CD4 (helper T cells), CD8A (cytotoxic T cells), GZMB (activated cytotoxic T cells), PRF1 (activated cytotoxic T cells), and RNF128 (anergic T cells) (primers and probes in Table S1) (46,47).

Strong positive correlations were observed between APOBEC3G mRNA expression levels and CD3D (p<0.0001, r_s=0.6159), CD4 (p<0.0001, r_s=0.5825), CD8A (p<0.0001, r_s=0.6168), GZMB (p<0.0001, r_s=0.6591), and PRF1 (p<0.0001, r_s=0.6422) (Fig. 1A-E). The lone exception was RNF128 (p=0.7665, r_s=0.0161), which is a marker for T cell anergy (48), suggesting that the positive signals emanate from bona fide activated T lymphocytes (Fig. 1F). Interestingly, APOBEC3G expression showed a similar positive correlation with CD8A and CD4, suggesting that APOBEC3G is expressed in both the cytotoxic and helper T cell subsets (Fig. 1B vs. 1C). This result was corroborated by similarly strong positive correlations between APOBEC3G and two markers of CTL activation, GZMB and PRF1 (Fig. 1D and 1E). In contrast, weaker and less significant correlations were found between APOBEC3B and the expression of any of these T cell genes (Fig. 1G-L). This result was expected, however, because prior studies have indicated that APOBEC3B is only expressed
at very low levels in normal tissues and upregulated in tumor cells (12,17,24,25). Taken together, our data confirm prior studies documenting T cell infiltration in HGSOC (3-7) and, importantly, reveal an unanticipated association between high levels of APOBEC3G expression and CTL activation in HGSOC. This result for APOBEC3G was especially unexpected given its broad expression profile documented previously by several groups including our own (12,15-17).

**APOBEC3G protein visualization in infiltrating T lymphocytes in HGSOC**

To determine whether APOBEC3G protein expression co-localizes with the same T cell markers analyzed above at the mRNA level by RT-qPCR, we performed hematoxylin staining and immunohistochemistry for CD3, CD4, CD8, and APOBEC3G on seven unrelated HGSOC samples (representative image sets in Fig. 2A-T with hematoxylin staining in Fig. 2A, F, K, and P). As expected from the RT-qPCR analysis above, several of these additional tumor specimens showed clear evidence for T lymphocyte infiltration. Two HGSOC lesions contained low expression of all examined T cell markers (representative images for CD3, CD4, and CD8 in Fig. 2B-D) and correspondingly low levels of APOBEC3G expression (representative image in Fig. 2E). In contrast, two tumors showed extensive T lymphocyte infiltration (representative images for CD3, CD4 and CD8 in Fig. 2G-I, L-N, and Q-S). Interestingly, there was a strong colocalization between these markers and the expression of APOBEC3G in these tumors (representative image in Fig. 2J, O, and T ). Moreover, different regions of a single HGSOC can be heterogenous, with one region showing dense clusters of APOBEC3G-high infiltrating T cells and another showing a more dispersed distribution (compare Fig. 2F-J and K-O). The remaining three samples showed moderate expression of both infiltrating T cell markers and APOBEC3G, and colocalization was still observed (data not shown). These immunohistochemical experiments confirm the above correlations and show that APOBEC3G is indeed expressed at the protein level.
within tumor infiltrating T lymphocytes.

To directly test whether APOBEC3G is expressed within tumor infiltrating T lymphocytes, we performed a series of fluorescence-based co-localization experiments with CD3 and APOBEC3G. Here, we were only able to image well-isolated infiltrating T lymphocytes because groups of cells caused excessive background fluorescence. Nevertheless, these experiments enabled us to demonstrate for the first time that CD3 and APOBEC3G are indeed co-expressed in the same T lymphocyte (representative image in Fig. 2U-X). Moreover, the DAPI co-stain and 1000x total magnification combine to show that both proteins are excluded from the nuclear compartment and predominantly cytoplasmic (the additional cell surface localization of CD3 is more difficult to visualize in tissue cross sections).

**APOBEC3G is a candidate biomarker for improved HGSOC patient outcomes**

Long-term clinical follow-up data were available for all of the Mayo Clinic HGSOC patients. The aforementioned gene expression data were therefore correlated with clinical information to determine whether APOBEC3G expression levels predict the length of progression free survival (PFS) and/or overall survival (OS) in HGSOC. As a positive control, the T cell markers above were also analyzed with respect to clinical information. Kaplan-Meier plots were constructed by splitting each gene expression data set at the median to create high and low expression groups (Fig. 3A and B). P-values, hazard ratios (HR), and 95% confidence intervals (CI) were determined using Cox regression models on the log2-transformed expression that were corrected for stage and debulking status (Table 2).

Higher expression levels of CD3D (p=0.020, HR=0.94 [95% confidence interval 0.90,0.99]), CD4 (p=0.0046, HR=0.90 [0.84, 0.97]), CD8A (p=0.0053, HR=0.93 [0.88, 0.98]), GZMB (p=0.011, HR=0.94 [0.90, 0.99]), and PRF1 (p=0.0049, HR=0.91 [0.86, 0.97]) were all associated with improved PFS (Fig. 3A and Table 2). As expected, RNF128 (p=0.43,
HR=0.97 [0.91, 1.04]) did not correlate with PFS (Fig. 3A and Table 2). Interestingly, APOBEC3G (p<0.0001, HR=0.81 [0.73, 0.89]) surpassed all of these genes as the most indicative marker of improved PFS in HGSOC (Fig. 3A). The results compiled from an analysis of OS largely mirrored those of PFS (Fig. 3B and Table 2).

Next, we used Kaplan-Meier plotter (kmplot.com) to generate a validation cohort using HGSOC data from TCGA and GEO. In this composite analysis of HGSOC patients, high levels of APOBEC3G correlated with improved durations of PFS (p=0.0057, HR=0.78 [0.65, 0.93]) and OS (p=0.063, HR=0.84 [0.70, 1.01]) (Fig. S2). Although only the former correlation reached statistical significance, the latter trended toward significance and also supported the observations above with the HGSOC Mayo Clinic cohort. A larger degree of variation is to be expected in this composite cohort because clinical variables are more difficult to take into account.

Finally, we analyzed HGSOC specimens from a Dutch cohort consisting of 88 patient samples (clinical information in Table S2). APOBEC3G expression was quantified using an independent RT-qPCR strategy (housekeeping gene primer sequences in Table S1 and see Material and Methods for details). Again, APOBEC3G expression levels associated with improved OS and more weakly with PFS (Fig. S2).

**APOBEC3B expression levels do not associate with HGSOC outcomes**

APOBEC3B has been implicated recently as an endogenous mutagen in several cancers [(28-30) and references therein], including ovarian cancer (24). Moreover, its overexpression has been linked to poor patient outcomes in multiple cancer types (31-36). Using the Mayo cohort gene expression data and clinical information from above, we asked whether APOBEC3B affects patient outcomes in HGSOC. These analyses revealed a trend toward high APOBEC3B and improved, rather than worsened, PFS and OS outcomes, although these relationships were less significant statistically than those for APOBEC3G (Fig.
**APOBEC expression correlates with immune cell markers in multiple human cancers**

To extend our findings from HGSOC to additional human cancers, publicly available RNAseq data from TCGA were analyzed to determine if correlations exist between expression of APOBEC genes and immune cell markers. At the time of these analyses, the TCGA had RNAseq data available for 7,861 samples spanning 22 different tumor types (details in Table S3). For each tumor type, expression of each APOBEC family member was quantified and correlated with the T cell marker CD3D (Fig. 4; top heatmap). Hierarchical clustering was also performed to elucidate similar correlation patterns between cancer types (Fig. 4A top). These analyses revealed that, in addition to APOBEC3G, APOBEC3D and APOBEC3H also correlated significantly with CD3D across multiple cancer types. APOBEC3F, also known to be expressed broadly and in T cells (12,13,15-17), did not correlate as strongly. The same analysis was also performed with CD20, a well-known marker for B cells (Fig. 4 bottom heatmap). The expression of the antibody diversification gene, AID, was the only APOBEC family member that significantly correlated with CD20 in a majority of cancer types (Fig. 4B). These analyses indicate that much of the expression of several APOBEC family members in cancer is likely due to T and B cell infiltration.

**Discussion**

Our studies have led to the surprising identification of APOBEC3G as a new candidate biomarker for activated T lymphocyte infiltration in HGSOC and improved patient outcomes. This result was unexpected because prior work has shown that APOBEC3G is expressed broadly, constitutively, and in some instances inducibly (12,13,15-17). The analysis of a cohort of 354 HGSOC patients also identified a strong correlation between
APOBEC3G and several markers of T cell infiltration (Fig. 1A-E). These results were validated at the protein level by immunohistochemistry and immunofluorescent staining of independent HGSOC tumor samples (representative image sets in Fig. 2). Clinical data revealed that APOBEC3G also associates with improved outcomes in a large HGSOC cohort as well as in two additional independent ovarian cancer cohorts (Table 2, Fig. 3, and Fig. S2). Finally, a global analysis of gene expression in 22 cancer types identified a similar correlation for two additional APOBEC3 genes, APOBEC3D and APOBEC3H, and a marker of T cells, CD3D (Fig. 4). Together, these data suggest that APOBEC3G expression levels in tumor infiltrating T lymphocytes may be a predictive biomarker for strong anti-cancer T cell responses and improved HGSOC outcomes.

Over the past several years, substantial evidence has accumulated to indicate that APOBEC3B and AID contribute to cancer genome mutagenesis. APOBEC3B is thought to mutate the genome of several different cancer types, including breast, lung, bladder, cervical, head/neck, and ovarian (17,24-27). The carcinogenic effect of AID appears more limited, as its characteristic deamination signature has only been found in certain types of B cell leukemias and lymphomas (23,27). The idea that AID expression is constrained to B lineage cell types is consistent with our data showing that AID mRNA expression correlates with CD20 mRNA expression in several solid tumor types (Fig. 4 bottom heatmap). In addition to APOBEC3B and AID, several other APOBEC family members have also been implicated in carcinogenesis (37-40). For example, it has been proposed that APOBEC3G drives hepatocellular carcinoma tumorigenesis (37). This idea, however, is difficult to reconcile with our observation that APOBEC3G mRNA expression in HGSOC and other tumor types correlates with the expression of activated T lymphocyte markers and, at least for HGSOC, this strong correlation could be validated visually at the protein level (Figs. 1, 2, and 4). Additional studies are clearly warranted to extend the APOBEC3G research.
immunohistochemistry results described here to other tumor types including hepatocellular carcinoma.

We were also surprised by the analysis of APOBEC3B in HGSOC, which revealed no significant clinical correlations and even a slight trend in the Mayo Clinic cohort toward better overall survival (Fig. 3). This finding differs significantly from estrogen receptor (ER)-positive breast cancer, where high APOBEC3B expression associates with shorter periods of disease-free survival and poorer rates of overall survival (31-34). A previous deep-sequencing study highlighted the similarities between HGSOC and another type of breast cancer, triple negative breast cancer (TNBC) (49), and, also found no correlation between APOBEC3B expression and survival in TNBC (32). One major difference between these two cancer types and ER-positive breast cancer is therapeutic options. There are multiple targeted therapeutics available for the treatment of ER-positive breast cancer that are administered based on molecular markers. In contrast, nearly all HGSOC and many TNBC patients are treated with platinum-based therapies. Because platinum-based therapies induce DNA damage, it is possible that these drugs become synergistic with APOBEC3B-catalyzed cytosine deamination and create a synthetic lethal state in cancer cells. This idea is reasonable as increased mutation loads have been shown to correlate with improved clinical outcomes in HGSOC patients treated with cisplatin (50). Furthermore, a synergistic effect created by these two forms of DNA damage could explain the slight trend found here toward a positive correlation between increased APOBEC3B expression and improved outcomes. Another possible explanation is that the levels of APOBEC3B mutagenesis in ovarian cancer may be not high enough to manifest clinically. Indeed, the APOBEC3B mutation signature is not as strong in ovarian cancer as it is in many other cancer types despite similar mRNA and protein expression levels (17,21,24-27). The underlying causes for this apparent discordancy are unknown, but several factors could be involved, including altered DNA repair capacities, differential protein regulation, and mutational contributions.
from other sources. More work is needed to determine the threshold of APOBEC3B mutagenesis needed to have a clinical impact.

Our global analysis of TCGA data revealed that the correlation between APOBEC3F and CD3D was substantially diminished as compared to APOBEC3D, APOBEC3G, and APOBEC3H (Fig. 4). This is surprising because APOBEC3F is thought to be broadly expressed and play an equally important role as APOBEC3D, APOBEC3G, and APOBEC3H in HIV-1 restriction in CD4-positive T lymphocytes, and prior studies have found that APOBEC3F is expressed at comparable levels to APOBEC3D and higher than APOBEC3H in human primary CD4+ T cells (12-14). Our data suggest that APOBEC3F may be under- and/or heterogeneously-expressed in T cells associated with the tumor microenvironment. These differences are potentially interesting as the overall APOBEC family member expression profile may be a useful property for distinguishing anti-viral from anti-tumor immune responses. It will be interesting in future studies to determine whether APOBEC3G, the overall APOBEC family expression profile, APOBEC mutation signature, and/or classical markers for T cell activation are the best predictors of successful clinical responses to immunotherapies for HGSOC as well as for other tumor types.

Acknowledgements

We thank Emily Law and David Masopust for thoughtful discussions, and Xuan Bich Trinh, Peter van Dam, Peter Vermeulen, Steven van Laere, and Luc Dirix from the Center for Oncological Research (CORE) at the University of Antwerp (Belgium) for contributing clinical specimens to the Dutch cohort.

Grant Support

Salary support for BL was provided by a Cancer Biology Training Grant (NIH T32
CA009138), and GJS by a National Science Foundation Graduate Research Fellowship (DGE 13488264). MJM, ALO, KRK, and SHK were supported in part by the Mayo Clinic SPORE in Ovarian Cancer (P50 CA136393). Ovarian cancer research in the Harris laboratory was supported by grants from the Minnesota Partnership for Biotechnology and Medical Genomics and the Minnesota Ovarian Cancer Alliance. RSH is an Investigator of the Howard Hughes Medical Institute.

References


25. Burns MB, Temiz NA, Harris RS. Evidence for APOBEC3B mutagenesis in multiple


35. Xu L, Chang Y, An H, Zhu Y, Yang Y, Xu J. High APOBEC3B expression is a predictor


Table 1. Clinical information for Mayo cohort (n=354)

<table>
<thead>
<tr>
<th>Tumor characteristic</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Morphology</strong></td>
<td></td>
</tr>
<tr>
<td>Serous</td>
<td>354 (100%)</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>11 (3.1%)</td>
</tr>
<tr>
<td>3</td>
<td>343 (96.9%)</td>
</tr>
<tr>
<td><strong>Stage</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14 (3.9%)</td>
</tr>
<tr>
<td>2</td>
<td>8 (2.3%)</td>
</tr>
<tr>
<td>3</td>
<td>253 (71.5%)</td>
</tr>
<tr>
<td>4</td>
<td>79 (22.3%)</td>
</tr>
<tr>
<td><strong>Debulking status</strong></td>
<td></td>
</tr>
<tr>
<td>No residual disease</td>
<td>162 (45.8%)</td>
</tr>
<tr>
<td>&lt;=1 cm remaining</td>
<td>144 (40.7%)</td>
</tr>
<tr>
<td>&lt;=1 cm remaining, possibly</td>
<td>48 (13.5%)</td>
</tr>
</tbody>
</table>


### Table 2. Mayo cohort Cox regression analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Parameter</th>
<th>p-value</th>
<th>HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD3D</td>
<td>PFS</td>
<td>0.020</td>
<td>0.94</td>
<td>[0.90, 0.99]</td>
</tr>
<tr>
<td>CD3D</td>
<td>OS</td>
<td>0.087</td>
<td>0.95</td>
<td>[0.90, 1.01]</td>
</tr>
<tr>
<td>CD4</td>
<td>PFS</td>
<td>0.0046</td>
<td>0.90</td>
<td>[0.84, 0.97]</td>
</tr>
<tr>
<td>CD4</td>
<td>OS</td>
<td>0.018</td>
<td>0.91</td>
<td>[0.84, 0.98]</td>
</tr>
<tr>
<td>CD8A</td>
<td>PFS</td>
<td>0.0053</td>
<td>0.93</td>
<td>[0.88, 0.98]</td>
</tr>
<tr>
<td>CD8A</td>
<td>OS</td>
<td>0.015</td>
<td>0.93</td>
<td>[0.87, 0.99]</td>
</tr>
<tr>
<td>GZMB</td>
<td>PFS</td>
<td>0.011</td>
<td>0.94</td>
<td>[0.90, 0.99]</td>
</tr>
<tr>
<td>GZMB</td>
<td>OS</td>
<td>0.047</td>
<td>0.95</td>
<td>[0.90, 1.00]</td>
</tr>
<tr>
<td>PRF1</td>
<td>PFS</td>
<td>0.0049</td>
<td>0.91</td>
<td>[0.86, 0.97]</td>
</tr>
<tr>
<td>PRF1</td>
<td>OS</td>
<td>0.018</td>
<td>0.92</td>
<td>[0.86, 0.99]</td>
</tr>
<tr>
<td>RNF128</td>
<td>PFS</td>
<td>0.43</td>
<td>0.97</td>
<td>[0.91, 1.04]</td>
</tr>
<tr>
<td>RNF128</td>
<td>OS</td>
<td>0.44</td>
<td>0.97</td>
<td>[0.91, 1.04]</td>
</tr>
<tr>
<td>APOBEC3G</td>
<td>PFS</td>
<td>&lt;0.0001</td>
<td>0.81</td>
<td>[0.73, 0.89]</td>
</tr>
<tr>
<td>APOBEC3G</td>
<td>OS</td>
<td>0.0003</td>
<td>0.82</td>
<td>[0.73, 0.91]</td>
</tr>
<tr>
<td>APOBEC3B</td>
<td>PFS</td>
<td>0.034</td>
<td>0.92</td>
<td>[0.85, 0.99]</td>
</tr>
<tr>
<td>APOBEC3B</td>
<td>OS</td>
<td>0.06</td>
<td>0.93</td>
<td>[0.87, 1.00]</td>
</tr>
</tbody>
</table>

1. OS= overall survival (n= 354); PFS= progression free survival (n= 348)
2. HR= hazard ratio
3. CI= confidence interval
Figure Legends

**Figure 1.** Correlations between APOBEC3 expression and T cell markers in HGSOC.

Dot plots illustrating correlations between APOBEC3G (A-F) or APOBEC3B (G-L) expression and the indicated T cell marker (n=354). mRNA expression was determined using RT-qPCR and normalized to the housekeeping gene TBP. Spearman’s correlation coefficients ($r_s$) and p-values are shown. Best-fit lines are shown for qualitative comparison, and were calculated using linear regression models.

**Figure 2.** Immunohistochemistry and immunofluorescence of T cell markers in HGSOC.

Photomicrographs of immunohistochemistry and immunofluorescence staining performed on HGSOC specimens illustrating the association between levels of T lymphocyte infiltration and the intensity of APOBEC3G expression. Representative staining of one HGSOC specimen with low (patient 6) and three staining sets from two HGSOC specimens with high (patients 3 and 2) levels of T cell infiltration are shown. The images depict hematoxylin (A, F, K, and P), CD3 (B, G, L, and Q), CD4 (C, H, M, and R), CD8 (D, I, N, and S), and APOBEC3G (E, J, O, and T). The dotted box in the 40x hematoxylin images indicates the approximate location of the subsequent panels at 100x magnification. The bottom row shows representative 1000x magnification images of colocalization of CD3 and APOBEC3G by immunofluorescent staining. DAPI-stained nuclei are blue.

**Figure 3.** Clinical correlates of T cell marker expression in HGSOC.

Kaplan-Meier plots illustrating associations between progression free survival (PFS) (A, n=354) or overall survival (OS) (B, n=348) and either one of the conventional T cell markers or APOBEC3G or APOBEC3B expression in the Mayo cohort of patients. Samples were
split at the median expression level for each gene with red representing tumors with high and blue representing tumors with low mRNA levels.

**Figure 4.** Correlations between *APOBEC* expression and immune cell markers across 22 cancer types.

Heatmap of Spearman’s correlation coefficients calculated from the comparison of the T cell marker *CD3D* (A) or the B cell marker *CD20* (B) with the indicated *APOBEC* family member. Expression levels were determined using TCGA RNAseq data (see Table S3 for the long form for each tumor abbreviation). Dark red squares indicate strong positive correlations, dark blue squares indicate strong negative correlations and white squares indicate a lack of correlation.
Fig. 1
Fig. 3
Fig. 4
Clinical Cancer Research

APOBEC3G expression correlates with T cell infiltration and improved clinical outcomes in high-grade serous ovarian carcinoma

Brandon Leonard, Gabriel J Starrett, Matthew J. Maurer, et al.

Clin Cancer Res  Published OnlineFirst March 25, 2016.