Analysis of tumor mutation burden combined with immune infiltrates in endometrial cancer

Jun Zhang#, Lanfen An#, Xing Zhou, Rui Shi, Hongbo Wang

Department of Obstetrics and Gynecology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Contributions: (I) Conception and design: J Zhang, L An, H Wang; (II) Administrative support: H Wang; (III) Provision of study materials or patients: X Zhou, R Shi; (IV) Collection and assembly of data: J Zhang, L An; (V) Data analysis and interpretation: J Zhang, L An; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Hongbo Wang. Department of Obstetrics and Gynecology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China. Email: hb_wang1969@sina.com.

Background: Tumor mutational burden (TMB) is widely regarded as a predictor of response to immunotherapy. Few researchers have focused on the activity and prognosis of TMB in endometrial cancer (EC) and immune cells. Our study aimed to identify the prognostic role of TMB in EC.

Methods: We downloaded transcriptome data from The Cancer Genome Atlas (TCGA) database. Kaplan-Meier analysis with log-rank test was conducted to assess the difference in overall survival (OS) between the high and low TMB groups. The “CIBERSORT” scripts were performed to evaluate the immune compositions of EC patients. Cox regression analysis and survival analysis were used to verify the prognostic value prognosis of TMB.

Results: We obtained the single nucleotide mutation data for 529 EC patients. A missense mutation was the most common mutation type. TMB was associated with survival outcome, tumor grades, and pathological types. We identified 10 hub TMB-related signature and found that elevated T-cell subsets infiltrating density in the high TMB group revealed improved survival outcomes. According to Kaplan-Meier analysis, T cells gamma delta and T cells regulatory were prognostic immune cells in EC samples. Moreover, many top gene set enrichment analysis (GSEA) results, including amino sugar and nucleotide sugar metabolism, nucleotide excision repair, or p53 signaling pathway, were enriched significantly with TMB level as phenotype.

Conclusions: TMB is an important prognostic factor for EC, and TMB-related genes may be potential therapeutic targets for EC.

Keywords: Endometrial cancer (EC); immune cells; tumor mutational burden (TMB)

doi: 10.21037/atm-20-6049

View this article at: http://dx.doi.org/10.21037/atm-20-6049

Introduction

Endometrial cancer (EC) is the most common female reproductive tract malignancies in developed nations with an increasing incidence. It was estimated that 63,230 new cases and 11,350 deaths resulting from EC occurred in 2018 (1,2). EC has become a serious disease that threatens women’s health and quality of life, increasing the disease risk and medical burden on society. The majority of women with EC are in the early stage of the disease with a good prognosis after surgery therapy alone (2). However, about 15% of patients with EC are diagnosed with high-risk factors, and women with high-risk factors have an increased risk of distant metastasis and tumor-related death (3,4). As an independent organ, tumor formation means the complex evolution of neoplastic cells, immune cells, extracellular...
The clinical successes of chimeric antigen receptor T cell and immune checkpoint blockade therapies represent a milestone in cancer immunotherapy (6). Immunotherapy has become an essential treatment for advanced-stage cancers (2,7-9). The immune cells can play opposing roles in developing tumors: promotion of tumor progression versus clearance of neoplastic cells (6). And progressive results have been achieved. In many malignant tumors, immunotherapy is designed to induce type I inflammation and enhance cytotoxic T cells, type I T helper cells and type M1 macrophages to eliminate tumor cells (10,11). Indoleamine 2,3-dioxygenase (IDO) has been identified to interfere with tumor lymphatic endothelial cell infiltration and is associated with poor survival prognosis in EC (12). In recent years, the development of IDO inhibitors has shown preliminary therapeutic promise in cancer therapy (13). Since CD8+ T cells are the major direct effector of cytotoxic responses to cancer cells, after PD-1 blockade in lung cancer patients, PD-1 and CD8+ T cells expanded, these changes are considered to be the result of effective antitumor immunity, and tumor CD8+, T cell infiltration, correlates with positive clinical outcome (14,15).

The tumor immune microenvironment mainly refers to the microenvironment associated with immune cells (16). Pierini et al. has found that regulatory T cells (Treg) in the immune microenvironment have the ability to immune tolerance and suppress conventional T cells and other immune cells (such as natural killer cells and B lymphocytes) (17). Zhao indicated that Treg could secrete vascular endothelial growth factor A (VEGFA) and promote angiogenesis in a hypoxic environment, providing tumor cells with abundant nutrients (18). Morrison demonstrated that tumor stem cells are resistant to cytotoxic T lymphocytes’ effects and can escape immune recognition and cytotoxic T cell-mediated killing by low expression of major histocompatibility complex-I (19).

TMB, as a predictor of response to immunotherapy, refers to the total number of somatic coding mutations, gene insertion, base substitutions, and deletion errors detected per million bases (20). At present, the mechanism of the prediction of TMB tumor immune response is not fully understood. It is generally believed that the higher the TMB is, the more new antigens will be produced; that is, the body’s immunogenicity will increase, allowing tumor-specific T cells to recognize the new antigens and eventually produce an immune response (21-23). Previous reports have explored the role of TMB in immunotherapy (24). In metastatic colorectal carcinoma (mCRC), TMB is an important independent biomarker for mCRC with high microsatellite instability, which can stratify the possibility of patients responding to immune checkpoint inhibitors (ICPIs) (25). In melanoma patients, high TMB may be associated with the long-term clinical benefit of ICPIs (26). In patients with breast cancer, tumors with favorable immune-infiltrate and high TMB are correlated with prolonged survival (27). Besides, many tumor samples with mismatch repair (MMR) deficient respond more strongly to immunotherapy, which may be due to their high TMB (28).

In recent years, with the development of the sequencing technique, mutation data for different tumors can be downloaded from public databases such as The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. Many of these data resources were explored to identify the relationship between cancer and immune microenvironment and prognosis. However, few researchers have focused on the activity and prognosis of TMB in EC and immune cells; our study aimed to identify the prognostic role of TMB in EC.

We present the following article in accordance with the MDAR checklist (available at http://dx.doi.org/10.21037/atm-20-6049).

**Methods**

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

**Multi-omics data progress and acquisition**

We downloaded transcriptome data (529 EC cases, Workflow Type: HTseq-FPKM) from the TCGA database. In addition, through the Genomic Data Commons (GDC) tool, we can obtain the corresponding clinical characteristics of patients, including sex, age, tumor grade and clinical stage, and patients’ important status with follow-up.

**TMB assessment and prognosis analysis in EC**

In this study, TMB is defined as (total count of variants)/(the whole length of exons), in which base deletion, insertion, or substitution is considered as a variable (29). Using Practical Extraction and Report Language (Perl) scripts to run on JAVA software (30). The number of genomic changes in 529 EC patients was extracted, and the extracted TMB data were revealed in table online (https://cdn.amegroups.cn/static/
Differential methods were selected to analyze TMB groups and pathway analysis

We used “limma” package to conduct the differentially expressed genes (DEGs) analysis with |log (fold change)| >1.51 and false discovery rate (FDR) <0.05 (31). We selected the “org.Hs.eg.db” package to transferred gene symbols into Entrez ID. Using “clusterProfiler”, “ggplot2”, and “enrichplot” packages to conducted the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis of DEGs. Then, we downloaded the gene set enrichment analysis (GSEA) software and worked it on the JAVA platform. TMB data were selected as the phenotype, and then downloaded “c2.cp.kegg.v6.2.symbols.gmt gene sets” from the Molecular Signatures Database database (MSigDB).

Identifying hub TMB related signatures

We used the “survival” installation package and univariate Cox analysis to find prognostic genes DEGs with P<0.05. Then, the hub independent risk signature was identified using multivariate Cox analysis. The forest plot showed the hazard ratio (HR) of hub TMB-related genes.

CIBERSORT

First, using the “limma” package to normalized transcriptome profiles of EC patients. Next, we performed the “CIBERSORT” scripts to assess the immune compositions of EC patients (table online: https://cdn.amegroups.cn/static/public/atm-20-6049-2.xlsx) (32). We considered that the selected samples were usable when P<0.05.

Infiltrating immune cells prognostic analysis in EC

We integrated these data and calculated 22 important immune cells from 529 samples for analysis based on the immune cell survival data. Univariate Cox regression analysis was used to evaluate the prognostic value of immune cells. In high TMB and low TMB, Kaplan-Meier analysis was conducted to assess differential progression, recurrence, and survival outcomes.

Statistical analysis

Students’ t-test was conducted for continuous variables.
The Chi-square test was used to study the categorical variables. The comparison between the two groups uses the Wilcoxon rank-sum test. Our data analysis was performed in R software (version 3.6.1). To assess the difference in overall survival (OS) between the two levels, we performed a Kaplan-Meier analysis with a log-rank test. At the same time, we further analyzed the relationship between several clinical variables and TMB. The Wilcoxon rank-sum test was performed to compare the clinical characteristics between the two groups, and the Kruskal-Wallis test was preferred to compare the clinical characteristics among three or more groups. Statistical significance was defined as P value <0.05.

**Results**

**Genome-wide mutation maps in patients with EC**

We obtained the single nucleotide mutation data for 529 EC patients from the TCGA database. Different types of mutations were classified into different groups. Missense mutations were the most common mutation types (Figure 1A). Single nucleotide polymorphism (SNP) was more frequent than insertions and deletions (Figure 1B), and C>T was the most common type of single-nucleotide variant (SNV) (Figure 1C). We also used waterfall plots to show mutation information for each gene for each patient. Then, we counted the number of base mutations in each person, and mutation types were also shown in the box plot (Figure 1D,E). The top 10 mutated signatures in EC patients exhibited a descending order in a horizontal histogram. These mutated signatures include `PTEN` (64%), `PIK3CA` (48%), `ARID1A` (44%), `TTN` (38%), `TP53` (37%), `PIK3R1` (30%), `KMT2D` (26%), `CTCF` (26%), `MUC16` (25%), `CTNNB1` (24%) (Figure 1F). Other detailed mutation information was revealed in the waterfall plot (Figure 1G). Different colors represented different types of mutations. Clinical characteristic of 529 EC patients from TCGA cohort is showed in Table 1.

**TMB was associated with survival outcome, tumor grades, and pathological types**

The groups were classified into the high TMB group by the chosen value of TMB ≥10 and classified into the low TMB group by the selected value of TMB <10. High TMB correlated with better prognosis days with Kaplan-Meier analysis with log-rank test (P=0.0092) (Figure 2A). We also observed that TMB in different pathological types of EC has significant differences (Figure 2B). For the correlation between tumor grade and TMB, we revealed that high-grade tumor grade with high TMB (Figure 2C).

**Gene expression profiles in high and low TMB groups**

We selected the top 160 DEGs between two groups in the heatmap, where we could obtain that the in high TMB groups has decreased genome expression level (Figure 3A). Pattern specification process, regionalization, skeletal system development were enriched in the biological process. In the cellular component category, these TMB-related genes were mainly involved in the extracellular matrix. And in the molecular function group, these differential genes participated in receptor ligand activity, receptor regular activity, serine-type peptidase activity and serine hydrolase activity (Figure 3B). KEGG enriched results indicated that different genes might be correlated with other pathways (Figure 3C, Table 2). Moreover, many top GSEA results, including amino sugar and nucleotide sugar metabolism, nucleotide excision repair, or p53 signaling pathway were enriched significantly with TMB level as phenotype (Figure 3D,E,F).

**Identification and evaluation of 10 TMB-related signature**

Based on 160 DEGs (table online: https://cdn.amegroups.cn/static/public/atm-20-6049-3.pdf), we further selected ten prognostic signatures related to TMB. The ten independent risk signatures with P<0.05 using a multivariate regression model (Figure 4). The HR, with a 95% confidence interval, was shown in the forest plot. We also plotted time-dependent ROC curves of TMB and ten genes for predicting 1-, 3-, and 5-year survival of endometrial cancer (Figure S1).

**Inferred immune cell fractions between two TMB groups**

According to the CIBERSORT method, we deserved proportions of 22 immune cells in patients and showed the results in a bar plot, in which different colors represented different cell subsets (Figure 5). Moreover, we obtained that activated CD4+ T cells (P<0.001), plasma cells, and CD8+ T cells (P<0.001) exhibited higher infiltrating levels in the high TMB group by Wilcoxon rank-sum test. However, memory resting CD4+ T cell (P<0.001) and macrophages M0 (P<0.013) showed higher infiltrating levels in the low
Figure 1 Tumor mutation profile in EC samples. (A) A missense mutation is the most common mutations in variant classifications. (B) Invariant types, SNP account for a larger proportion than insertion and deletion. (C) In the SNV class, C>T exhibited more frequently than other types. (D,E,F) Tumor mutation burden in each sample and top ten mutated genes in endometrial cancer. (G) The landscape of mutation in EC samples. The waterfall plot was used to show the mutation information of each gene in each sample. The different color box below the waterfall plot represents various mutation types. The barplot was conducted to show the number of mutation burden. EC, endometrial cancer; SNP, single-nucleotide polymorphism; SNV, single-nucleotide variant.
Figure 2 Prognostic analysis of TMB and correlation with clinical characteristics. (A) Higher TMB level indicated a better OS with P=0.0092. (B) TMB level in different pathological types with P=9.54e−09. (C) Higher TMB level associated with advanced tumor grades with P=0.004. TMB, tumor mutation burden; OS, overall survival; EEA, endometrioid endometrial adenocarcinoma; mixed EA, mixed endometrioid adenocarcinoma; SEA, serous endometrial adenocarcinoma.

TMB group. No significant difference was found in naive B cells, resting NK cells, and activated NK cells between the high and low groups of TMB. However, T cells follicular helper is occupied more proportions (P<0.001) in the high TMB group, the total fraction of T cells follicular helper account for low component (Figure 6).

Elevated T-cell subsets infiltrating density in the high TMB group revealed improved survival outcomes

In view of the difference in the infiltration abundance of immune cells in different TMB groups, we analyzed the relationship between TMB-related immune cells and prognosis value. According to Kaplan-Meier analysis, T cells gamma delta and T cells regulatory were prognostic immune cells in EC samples. It is worth noting that T cells regulatory in the high TMB group had higher infiltrating abundance associated with improved survival rate (Figure 7).

Discussion

It is well known that growing tumor tissue contains tumor-infiltrating lymphocytes (TILs), and they have no effect on tumor elimination in vivo. Still, they can exert effector functions when removed from the immunosuppressive environment (33). The immune system plays a significant role in promoting and/or suppressing tumor progression, called cancer immunoediting. The immunoediting consists of three steps: elimination, equilibrium, and escape (34). In the past few decades, significant advances have been revealed in understanding how tumors escape the immune system, which led to new approaches to reduce tumor immune escape to eliminate tumors (33).

Promising results also have been obtained in immunotherapy (35). A comprehensive combination of adaptive and innate immunotherapeutic approaches can have a potent antitumor effect in a mouse melanoma model (36). Moreover, immunotherapy has been increasingly used in lung cancer patients. Programme death ligand 1 (PD-L1) as a newly found immunoregulatory molecule, interacting with its receptor, PD-1, inhibits cytotoxic immune response and exerts antitumor immune response (37-39). The emergence of ICPI has also changed some tumors’ treatment and has become a split-new standard of therapy (40,41).

Although immunotherapy has achieved encouraging results, only a small percentage of people have benefited (42).
Enrichment analysis of differential expression genes in high TMB and low TMB groups. (A) Heatmap shows the top 160 differentially expressed genes with log2(FC) > 1.51 and FDR < 0.05. (B,C) GO and KEGG enriched results displayed that these differentially expressed genes may be related to tuberculosis, regulation of lipolysis in adipocytes, pattern specification process, PI3K-Akt signaling pathway, and other functional pathways. (D,E,F) GSEA showed the top TMB-related signaling axis, including amino sugar and nucleotide sugar metabolism, nucleotide excision repair, or p53 signaling pathway with FDR < 0.25. TMB, tumor mutation burden; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology.
Table 2 KEGG pathway analysis for differential genes

<table>
<thead>
<tr>
<th>Description</th>
<th>BgRatio</th>
<th>P</th>
<th>P_adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroactive ligand-receptor interaction</td>
<td>9/278</td>
<td>2.39E−06</td>
<td>0.00032</td>
</tr>
<tr>
<td>Pathways in cancer</td>
<td>8/397</td>
<td>0.00023</td>
<td>0.01046</td>
</tr>
<tr>
<td>Melanoma</td>
<td>4/71</td>
<td>0.00023</td>
<td>0.01046</td>
</tr>
<tr>
<td>Oxytocin signaling pathway</td>
<td>5/158</td>
<td>0.00049</td>
<td>0.16899</td>
</tr>
<tr>
<td>Melanogenesis</td>
<td>4/100</td>
<td>0.0008</td>
<td>0.01946</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>5/179</td>
<td>0.00086</td>
<td>0.01946</td>
</tr>
<tr>
<td>cAMP signaling pathway</td>
<td>5/199</td>
<td>0.001358</td>
<td>0.02541</td>
</tr>
<tr>
<td>Basal cell carcinoma</td>
<td>3/55</td>
<td>0.0016</td>
<td>0.02541</td>
</tr>
<tr>
<td>Regulation of lipolysis in adipocytes</td>
<td>3/56</td>
<td>0.00168</td>
<td>0.02541</td>
</tr>
<tr>
<td>PI3K-Akt signaling pathway</td>
<td>6/342</td>
<td>0.00273</td>
<td>0.03518</td>
</tr>
<tr>
<td>Wnt signaling pathway</td>
<td>4/143</td>
<td>0.00285</td>
<td>0.03518</td>
</tr>
<tr>
<td>Hippo signaling pathway</td>
<td>4/154</td>
<td>0.00368</td>
<td>0.0365</td>
</tr>
<tr>
<td>mTOR signaling pathway</td>
<td>4/154</td>
<td>0.00368</td>
<td>0.03645</td>
</tr>
<tr>
<td>MAPK signaling pathway</td>
<td>5/255</td>
<td>0.00387</td>
<td>0.03645</td>
</tr>
<tr>
<td>JAK-STAT signaling pathway</td>
<td>4/158</td>
<td>0.00403</td>
<td>0.03645</td>
</tr>
<tr>
<td>ECM-receptor interaction</td>
<td>3/82</td>
<td>0.00474</td>
<td>0.03699</td>
</tr>
<tr>
<td>Vitamin digestion and absorption</td>
<td>2/24</td>
<td>0.00475</td>
<td>0.03699</td>
</tr>
<tr>
<td>Taste transduction</td>
<td>3/83</td>
<td>0.0049</td>
<td>0.03699</td>
</tr>
<tr>
<td>Salmonella infection</td>
<td>3/86</td>
<td>0.00539</td>
<td>0.0386</td>
</tr>
<tr>
<td>Protein digestion and absorption</td>
<td>3/90</td>
<td>0.00609</td>
<td>0.04059</td>
</tr>
</tbody>
</table>

KEGG, Kyoto Encyclopedia of Genes and Genomes.

Therefore, the research aims to search for predictive biomarkers for a better immune response (43). Densities of CD8^+^ T cells decreased along with the tumor progresses in colorectal cancer, indicating that intratumoral immune signatures can act as potential markers predicting OS (44).

Tumor mutational burden (TMB) is a new predictive marker of the immune response; in recent years, studies have shown that there is a linear relationship between the effectiveness of PD-1 inhibitors and TMB expression in a variety of solid tumors (45). Colorectal cancer patients treated with surgery followed by chemotherapy had a better prognosis with TMB-high (46). In our study, the high TMB group is correlated with OS, tumor grades, and pathological tumor types, which is in accordance with results in some other malignancies where high TMB was prone to induce local immune recognition and bring improved prognosis (47).

In recent years, researchers have an increasing interest in immune cell infiltration in tumor immunotherapy. TILs, which play an important role in immune surveillance and killing cancer cells in the lung cancer microenvironment, mainly include tumor-infiltrating T cells, natural killer cells, and tumor-infiltrating B lymphocytes (48). MMR deficiency patients with EC showed higher levels of CD8^+^ T cells, Tregs, and PD-L1 immune cells (49). Patients with EC with POLE mutations and microsatellite instability have higher TILs and PD-L1 expression levels than microsatellite stable TMB patients.
tumors (50).

We observed that, compared with the low TMB group, activated CD4⁺ T cell, plasma cells, and CD8⁺ T cell exhibited more abundant density in the high TMB group. Keane had observed that CD4⁺ TILs distinguished patients with different 5-year OS in diffuse large B-cell lymphoma (51). Fernández et al. found that changes in circulating inflammatory plasma-cells that can affect reality to the allergen (52). Donnem et al.
reported that CD8$^+$ TILs density has a prognostic impact on non-small cell lung cancer (53). In EC, regulatory T cells' expression might mediate T cell immune suppression within the cancer milieu and thus correlate with EC progression (54).

Based on KEGG and GO analysis, we selected DEGs, and we further selected ten TMB-related signatures using a multivariate regression model, including ACRL8, CLDN6, HIF3A, PDCL2, KRT71, L1CAM, ACTL8, SLC22A16, RYR1, PNMA3. HIF3A is a member of the transcription regulator family of hypoxia inducible factors (HIF), and many factors can affect target genes related to inflammation, cancer, and adipose tissue dysfunction (55). Recent research has shown that HIF3A can be a prognostic marker for prostate cancer (56). And Zhang et al. revealed that upregulated HIF3A expression promotes the progression of ovarian cancer (57). CLDN6 has a lower expression level in gastric cancer patients and is associated with age, distant metastasis, lymph node metastasis, and pathological staging showing as a potential biomarker (58). ACTL8 plays a key role in the metastasis, invasion, and poor prognosis of colorectal cancer (59). Similarly, in head and neck squamous cell carcinoma, ACTL8 may serve as a potential therapeutic target and prognostic marker (60).

In the present study, the proportion of immune cells in each patient was calculated according to the CIBERSORT method, which can be more convenient and save material resources in large-scale analysis of samples. Although we have obtained some results by analyzing biological information, a larger sample size of clinical research is needed to validate TMB potential correlation with immune infiltrates. Basic experiments verifying the association between ten immune genes signature and immune infiltrates should also be taken into consideration.

**Conclusions**

TMB is an important prognostic factor for EC, and TMB-related genes may be potential therapeutic targets for EC.

**Acknowledgments**

Funding: This work was supported by the National Natural Science Foundation of China (grant numbers 81974409).

**Footnote**

*Reporting Checklist:* The authors have completed the Materials Design Analysis Reporting (MDAR) checklist for authors. Available at [http://dx.doi.org/10.21037/atm-20-6049](http://dx.doi.org/10.21037/atm-20-6049)

*Peer Review File:* Available at [http://dx.doi.org/10.21037/atm-20-6049](http://dx.doi.org/10.21037/atm-20-6049)

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure (Available at [http://dx.doi.org/10.21037/atm-20-6049](http://dx.doi.org/10.21037/atm-20-6049)). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as...
revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


Figure S1 Time-dependent ROC curves of TMB and ten genes for predicting 1-, 3-, and 5-year survival of endometrial cancer. ROC, receiver operating characteristic; TMB, tumor mutation burden.