Abstract. Triple-negative breast cancer (TNBC) accounts for 15-30% of all breast cancer cases and is clinically difficult to treat due to the lack of hormone or human epidermal growth factor receptor 2 receptors, which are usually targeted by the most successful therapeutic approaches. Immune checkpoint inhibitors (ICIs) have offered long-term survival benefits in several types of solid tumors, however with low response rates. Thus, there is an urgent need to develop feasible biomarkers for identifying patients with TNBC, who are responsive. The present study demonstrated that the immune microenvironment of TNBC has the highest expression of immunoregulatory molecules among all pathologic types. The tumor mutation burden (TMB) of TNBC was not strongly correlated with cytolytic activity and showed no significant associations with different degrees of immune cell infiltration and TMB. The machine learning method divided patients with TNBC into two groups characterized by ‘hot’ and ‘cold’ tumors, according to whether immune-associated genes were highly expressed, and different responses to immunotherapy were seen between these two groups. Furthermore, patients with a TP53MutPIK3CAWild genotype demonstrated favorable immunotherapy-responsive signatures and may have improved outcomes with ICIs. In conclusion, the present study revealed that TP53 and PIK3CA may be appropriate biomarkers to screen for patients who would benefit most from ICIs, which could guide precise immunotherapy for patients with TNBC.

Introduction

Breast cancer (BC) is a highly heterogeneous cancer in both biological mechanisms and clinical treatment. Triple-negative breast cancer accounts for ~15-30% of invasive breast cancer, lacks the expression of estrogen receptors (ERs) and progesterone receptors (PRs), and does not overexpress human epidermal growth factor receptor 2 (HER2) (1). The lack of receptors on TNBC that can be targeted by drugs has made the development of treatments for TNBC challenging compared with other BC types. To date, to the best of our knowledge, not a single targeted therapy has been approved for the treatment of TNBC, and traditional chemotherapeutic reagents remain the standard treatment. Patients who do not show a pathologically complete response have a poor prognosis with a high incidence of recurrence (2-4); thus, there is an urgent need to explore new therapies and provide more patients with TNBC with other treatment options (5,6).

Immune checkpoint inhibitors (ICIs), such as antibodies targeting the checkpoints cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4; for example, ipilimumab) and programmed cell death protein 1 (PDCD1; for example, pembrolizumab and nivolumab), have shown remarkable achievements in a variety of cancer types, such as melanoma, Hodgkin’s lymphoma and non-small-cell lung cancer (7-10). A multicenter, nonrandomized phase I-b clinical trial named KEYNOTE-012 (Clinical Trials.gov identifier: NCT01848834) aims to investigate the safety, tolerability and antitumor activity of pembrolizumab (a programmed death-ligand 1, PDL1, inhibitor) in patients with TNBC and adenocarcinoma of the stomach or gastro-oesophageal junction (11). Among the 27 women with TNBC who were evaluable for antitumor activity, the overall response rate (ORR) was 18.5%, and the median time to response was 17.9 weeks (range, 7.3-32.4 weeks). Only a small fraction of patients with TNBC exhibited satisfactory clinical responses compared with patients with other solid tumors enrolled in KEYNOTE-012 (ORR: Advanced gastric cancer, 22%; advanced or metastatic urothelial cancer, 26%; and head and neck squamous cell carcinoma, 59%) (12-16). Therefore, it has become a primary priority to identify potential targetable biomarkers for ICI therapy and to investigate strategies to increase patient response rates.
Biomarkers, such as expression of the checkpoint PDL1 (17,18), tumor mutation burden (TMB) (19), tumor-infiltrating lymphocytes (TILs) (20), neoantigen load (21), and immune-regulatory mRNA expression signatures (22), are potentially applicable to predict the efficacy of ICIs. However, challenges in defining a validated cut-off value, intratumoral heterogeneity, test platform uniformities and dynamic changes have limited the clinical application (19,23,24). Furthermore, individual tumor clonal genotypes may require the screening of patients with TNBC to determine who will benefit from ICI therapy (25).

The present study aimed to verify the traditional ICI biomarkers, to explore novel biomarkers of TNBC and screen the patients who would benefit from ICI therapy.

Materials and methods

Data sources. The expression of the checkpoint molecules CTLA-4, indoleamine 2,3-dioxygenase 1 (IDO1), lymphocyte-activation gene 3 (LAG3), PDCD1, PDL1 and T cell immunoglobulin and mucin domain-containing protein 3 (TIM3) across four types of BC (HER2, luminal A, luminal B and TNBC) were obtained from a previous comprehensive study of human breast cancer (26) in The Cancer Genome Atlas (TCGA) database (cancer.gov/tcga) and log_2-transformed.

Materials and methods

Data sources. The expression of the checkpoint molecules CTLA-4, indoleamine 2,3-dioxygenase 1 (IDO1), lymphocyte-activation gene 3 (LAG3), PDCD1, PDL1 and T cell immunoglobulin and mucin domain-containing protein 3 (TIM3) across four types of BC (HER2, luminal A, luminal B and TNBC) were obtained from a previous comprehensive study of human breast cancer (26) in The Cancer Genome Atlas (TCGA) database (cancer.gov/tcga) and log_2-transformed. Normalized RNA-sequencing (RNA-seq) data were downloaded from the TCGA data portal. TMB data were also retrieved from the database. All data were analyzed using R (version 3.2.2) (27) and RStudio version 1.1.463 (28), unless otherwise stated.

Cytolytic activity (CYT). In order to study immune effector activity in TNBC, CYT was calculated as the log-average (geometric mean) of perforin-1 (PRF1) and granzyme-A (GZMA) expression in transcripts per million using RNA-seq data from TCGA.

Immune signature and single-sample gene set enrichment analysis (ssGSEA). To evaluate the immune signatures in the tumor microenvironment of TNBC, ssGSEA was used to identify gene sets from the Molecular Signatures Database (software.broadinstitute.org/gsea/msigdb/) that were enriched in TNBC; 28 heterogeneous immune cells were classified according to gene sets that share common biological function, chromosomal location, or regulation. Based on the immune signature spectrum, the degree of immune cell infiltration was determined by the ssGSEA scores, which were computed using R-package ‘Gene Set Variation Analysis’ version 1.34.0 (bioconductor.org/packages/GSVA/). The immune signature was clustered into high-, medium- and low-infiltration populations.

Random forests. To classify tumors of TNBC according to CYT, an accurate classification method named random forests was introduced. Random forests are comprised of a multitude of tree predictors such that each tree depends on a random vector independently, and all decision trees in the forests have the same distribution. A total of 782 parameters were used to assess CYT according to the importance score used to determine CYT. After dimensionality-reduced visualization by the multidimensional scaling (MDS) algorithm, the proximity of samples indicated their similarity in immunophenotype.

Ethical approval and informed consent. This study did not involve experiments on humans or animals performed by any of the authors.

Results

Suppressive tumor immune microenvironment in TNBC. Pre-existing expression of immune regulatory molecules is known as an indication of favorable responsiveness to ICIs in solid tumors. To investigate the immunophenotype of TNBC in comparison with other pathological types of BC, gene expression profiling and luminal categorization data were used from the publicly available TCGA database. Six immune regulatory molecules, including CTLA-4, IDO1, LAG3, PDCD1, PDL1 and TIM3, were examined in patients with each luminal type of BC. In this analysis, the highest expression intensity of CTLA-4, IDO1, LAG3, PDCD1 and PDL1 was identified in patients with TNBC among the four types. TIM3 demonstrated the second highest expression level in TNBC among the four types (Fig. 1). Moreover, higher expression of these immune regulatory molecules, including CTLA-4, IDO1, LAG3, PDCD1 and PDL1, was not associated with overall survival in patients with TNBC (Fig. S1). Thus, this enrichment of immune regulatory molecules suggested a suppressive antitumor response and poor prognosis in TNBC.

Cytolytic immune response is independent of TMB. A comprehensive evaluation of the tumor microenvironment of TNBC contributes to population selection. Effective antitumor cytotoxic responses rely on the recognition and presentation of newly generated and tumor-specific antigenic peptides on the cancer cell surface (30). Thus, whether the TMB could indicate the magnitude of the cytolytic immune response in the immune microenvironment of TNBC was first examined. To this end, CYT, a surrogate measurement of the magnitude of the cytolytic immune response was used (31). CYT was calculated as the geometric mean of GZMA and PRF1 expression, and has served as a highly specific marker in human glioblastoma (32). Based on this index, a very weak correlation was observed between the TMB and CYT in patients with TNBC (Fig. 2A), which strongly suggested that TMB may not represent the magnitude of the cytolytic immune response in TNBC.

In addition to CYT, immune cell infiltration in the tumor immune microenvironment is directly correlated with the magnitude of the cytolytic immune response (33). Furthermore,
systemic characterization of the pattern of tumor-infiltrating immune cell populations could lead to a deeper understanding of the TNBC immune microenvironment. Thus, whether a higher/lower TMB could predict the different infiltration levels of immune cells was tested. To estimate the relative infiltration of several intratumoral immune cell populations, the ssGSEA method was employed. This bioinformatics method translates the transcriptomic expression data as a normalized score to represent the relative abundance of specific cell types (34). Patients with TNBC were classified into three categories according to the degree of infiltration that was displayed by an unsupervised clustering algorithm (Fig. 2B). The color of each sample in the heatmap ranged from a cool color to a warm color, which represented scores from -3 to 3, respectively, and depicted infiltration of low, medium and high degrees. The clinicopathological characteristics, such as Tumor-Node-Metastasis stage, tissue type, and the expression of HER2, ER and PR, were also annotated.

Among these characteristics, it was found that clinical factors did not affect immune cell infiltration. TMB was not associated with the degree of infiltration (Fig. 2C). This suggests that TMB was not the only factor that determined the final immune response; further research is required to identify other biomarkers.

Machine learning identified dominant factors in determining CYT. Given that the TMB is not directly correlated with CYT in patients with TNBC, the optimization of the local immune response by a systematic approach was explored. To this end, the random forest method, a machine learning method based on multiple random-built decision trees, was employed. Hundreds of variants were employed as input parameters, including the relative infiltration of 28 types of immune cells, somatic mutation counts, 78 immune-related molecules, and 50 signaling pathways from the HALLMARK collection. By MDS algorithm, the close proximity of every two patients was illustrated to determine the similarity of immunological statuses between them. Two distinct immunological statuses were clearly categorized when guided by the density contour (Fig. 3A). To avoid artificial bias, out-of-bag samples supported a rational and acceptable error rate for the decision trees in the present analysis (Fig. 3B). Recently, Ayers et al (35) demonstrated that an expanded panel with 18 genes could distinguish different immunological statuses and predict a greater likelihood of response to immune checkpoint inhibitors. The panel includes genes involved in immune cell enrichment (CD3D, CD3E and NKG7), activation and function (CD2, IL2RG, TAGAP, GZMB, GZMK and STAT1), antigen presentation (CIITA, HLA-DRA, and HLA-E), chemokines and a chemokine receptor (CCL5, CCL10, CXCL13, and CXCR6) and immune checkpoint molecules (IDO1 and LAG3). Thus, these gene panels were introduced to further characterize the immunological statuses and ICI-responsive potential of patients with TNBC. Almost all 18 genes showed higher expression in the immunologically hot compared with the immunologically cold patients with TNBC, supporting an ICI-responsive potential for immunologically hot patients (Fig. 3C).

The most determining features in forming local CYT were identified using the importance score over all the input parameters (Fig. 3D). This score was derived from the machine learning process and represented the relative contribution of each factor to the resulting immune response. Several effector molecules, key transcription factors in the Th1/CTL response, were found in this analysis, which is consistent with established knowledge on antitumor immunity.

**TP53 Mut PIK3CA Wild** genotype defines patients with ICI-responsive potential. A previous study successfully demonstrated that specific genomic alterations are correlated with different efficacies of ICI treatment (36). The present
CHENG et al: BIOMARKER PREDICTS THE EFFICACY OF IMMUNOTHERAPY

study aimed to test whether oncogenic mutations in TNBC could cause different immunological statuses and sensitivities to immunotherapy. It is known that mutations in TP53 and PIK3CA occur frequently in patients with TNBC, and thus alterations in these genes may involve as many clinic cases as possible (Fig. 4A). The ability of TP53 / PIK3CA to define a subset of the tumor immune microenvironment was firstly examined. Four subsets were defined according to genomic alterations in TP53 and PIK3CA (Fig. 4B). Among the four different genotypes, the top 10 dominant factors in determining (Fig. 3D) CYT were examined (Fig. 4C). TNBC harboring wild-type PIK3CA had a significant abundance of all factors in comparison with the wild-type counterparts, suggesting a provoked antitumor response in patients with wild-type PIK3CA. Moreover, when focusing on PIK3CAWild patients, it was found that the TP53 mutation further indicated a subgroup with higher expression of the dominant factors. These findings implied that patients with genomic PIK3CAWild and TP53Mut alterations exhibited an elicited pre-existing CYT.

In recent years, new RNA-seq-derived biomarkers characterizing the inflamed tumor immune microenvironment represent one of the most exciting avenues for predicting the sensitivity of ICI immunotherapies in the treatment of solid tumors (35-37). In order to further evaluate the ICI-responsive potential, a scoring system using an immunophenotype score over four typically immunological factors was introduced: Effector cells (activated CD4+ T cells and CD8+ T cells, CD4+ Tem cells and CD8+ Tem cells), suppressor cells (regulatory T cells and MDSCs), MHC molecules (HLA-A, HLA-B, HLA-C, HLA-DPA1, HLA-E, HLA-F, HAVCR2, B2M, TAP1

Figure 2. TMB alone is not sufficient to assess CYT. (A) Correlation between CYT and TMB in patients with TNBC. The metric of CYT was quantified by mapping unmapped RNA-seq reads and normalizing to the count of mapped reads. (B) ssGSEA contributes to identifying the relative expression of immune cell populations and classifies patients with TNBC into three categories according to immune cell infiltration. The gradient of colors, from cool colors to warm colors, represents scores of immune cell infiltration (from -3 to 3, respectively). (C) There was no association between TMB and different degrees of immune cell infiltration (high, medium and low) in patients with TNBC. CYT, cytolytic activity; TNBC, triple negative breast cancer; TMB, tumor mutation burden; ssGSEA, single-sample gene set enrichment analysis; ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2.
and TAP2), and immunomodulators (immunostimulatory: CD27 and ICOS; and immune inhibitor molecules: LAG3, CTLA-4, PDCD1, PDCD1LG2, CD274, IDO1 and TIGIT; Fig. 4D). An elevated expression of the aforementioned immunological parameters was demonstrated in general, suggesting the TP53\(^{\text{Mut}}\)PIK3CA\(^{\text{Wild}}\) genotype as a potential biomarker for ICI treatment in patients with TNBC. This finding verified the hypothesis that patients with wild-type PIK3CA and TP53 mutations may have an improved response to immunotherapy.

**Discussion**

The present study aimed to identify biomarkers that would help to screen for patients with BC who were most likely to benefit from ICI therapy. A large gene analysis suggested that patients with TNBC might respond differently to ICIs based on their inclusion in one of the four subtypes of BC. The machine learning method facilitated the classification of TNBC according to its heterogenicity; a panel of 18 immune-associated molecules further divided TNBC into ‘hot’ and ‘cold’ tumors and indicated two different treatment outcomes. Based on these findings, specific gene mutations also affected the antitumor response, and patients with the TP53\(^{\text{Mut}}\)PIK3CA\(^{\text{Wild}}\) genotype may have an improved response to immunotherapy.

Gene analysis has proven to be a novel approach for judging the potential clinical benefit of immunotherapy. The expression of immune checkpoint genes, such as CTLA-4, IDO1, LAG3, PDCD1, PD-L1 or TIM3, which has traditionally been associated with responsiveness to ICIs, was analyzed as an indicator...
for screening tumors that were more suitable for immunotherapy (38). TNBC was identified as such a tumor, with genes that were highly expressed. A recent study has shown that PD-L1 expression is associated with the presence of TILs (39). The patients with TNBC were divided into three groups, according to the infiltration of TILs: High, medium and low infiltration.

Figure 4. Patients with the TP53<sup>Mut</sup>PIK3CA<sup>Wild</sup> genotype have a better response to immunotherapy. (A) Most common gene mutations in BC and TNBC were: PIK3CA, TP53, MLL3, RB1 and AFF2. (B) According to the mutation status of TP53 and PIK3CA, the patients with TNBC were divided into four groups. The pie chart shows the ratio of these four genotype-based groups: TP53<sup>Mut</sup>PIK3CA<sup>Mut</sup>, TP53<sup>Wild</sup>PIK3CA<sup>Mut</sup>, TP53<sup>Mut</sup>PIK3CA<sup>Wild</sup> and TP53<sup>Wild</sup>PIK3CA<sup>Wild</sup>. (C) The importance score of the top ten determinant signatures across four mutation statuses of TNBC. (D) The immunophenotype score across four typical immunological factors. *P<0.05 vs. with TP53<sup>Mut</sup>PIK3CA<sup>Wild</sup> group. TNBC, triple-negative breast cancer; ssGSEA, single-sample gene set enrichment analysis; BC, breast cancer.
Notably, there were no significant differences between TMB and TILs. This finding suggested that in addition to TMB, other factors, including the infiltration of immunosuppressive cells, the state of blood vessels, the size of the lesion, and the molecular typing of the lesion, also affected the TILs, which might have caused the significant differences. Advances in technological tools and high-throughput sequencing have increased the possibility of comprehensive characterization of somatic mutations (40). TMB has an important impact on the understanding of the efficacy of ICIs, as antigens that arise as a consequence of TMB are often targets of anti-PD-L1 and anti-CTLA-4 in mice (41). While TMB was not strongly associated with CYT in TNBC, specific biomarkers are needed to confirm these findings.

More than half of human tumors carry TP53 gene mutations, and these mutations are also frequently found in tumors but not in patients with luminal A and basal-like tumors (43). It was found that patients with TP53 mutations had an improved response to ICIs, in terms of higher expression of immune-associated molecules, which meant a more suitable antitumor microenvironment. PIK3CA is the second most frequently mutated gene, following the TP53 gene, and is associated with different types of BC. PIK3CA mutations may have favorable outcomes for patients with hormone receptor-positive BCs but may also constitute a major mechanism of resistance to trastuzumab treatment (44,45). The oncogene PIK3CA activates multipotent genetic programs and influences intratumoral heterogeneity (46).

In conclusion, the findings of the present study suggested that patients with TNBC may be the best suited for immunotherapy among the four subtypes of BC, and patients with the TP53<sup>mut</sup>PIK3CA<sup>Wild</sup> genotype will benefit most from ICI treatment. However, the clinical significance needs to be verified by further investigation.

Acknowledgements

The abstract was presented at the 17th International Congress of Immunology, 19th-23rd October 2019 in Beijing, China.

Funding

This study was funded by the National Natural Science Foundation of China (grant nos. 81472648, 81620108023 and 31700776).

Availability of data and materials

The datasets generated and/or analyzed during the present study are available in the TCGA database (https://gdc-portal.nci.nih.gov/).

Authors' contributions

JC and QJ analyzed the data, constructed figures and drafted the manuscript. JC, XD and SX were involved in data collection and analysis. BZ and QJ contributed to the conception and design of the study. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


