Cross-Platform in-silico Analyses Exploring Tumor Immune Microenvironment with Prognostic Value in Triple-Negative Breast Cancer

Victor C Kok 1,2, Charles CN Wang 2,3, Szu-Han Liao 2, De-Lun Chen 2

1Division of Medical Oncology, Kuang Tien General Hospital Cancer Center, Taichung, 43303, Taiwan; 2Department of Bioinformatics and Medical Engineering, Asia University, Taichung, 41354, Taiwan; 3Center for Artificial Intelligence and Precision Medicine Research, Asia University, Taichung, 41354, Taiwan

Correspondence: Victor C Kok; Charles CN Wang, Email: vkok@alumni.harvard.edu; cnwang@asia.edu.tw

Introduction: Only a proportion of triple-negative breast cancer (TNBC) is immunotherapy-responsive. We hypothesized that the tumor microenvironment (TME) influences the outcomes of TNBC and investigated the relevant signaling pathways.

Materials and Methods: Immune score (IS) and stromal score (SS) were calculated using the ESTIMATE and correlated with the overall survival (OS) in TNBC. RNA-seq data from 115 TNBC samples and 112 normal adjacent tissues were retrieved. Validations in the methylation levels in 10 TNBC and five non-TNBC cell lines were obtained. Cox model overall survival (OS) validated the derived transcription factor (TF) genes in cBioPortal breast cancer patients.

Results: SS-low predicts a higher OS compared with SS-high patients (P = 0.0081). IS-high/SS-low patients had better OS (P = 0.045) than IS-low/SS-high patients. More macrophages were polarized to the M2 state in patients with IS-low/SS-high patients (P < 0.001). Moreover, CIBERSORTx showed more CD8+ cytotoxic T-cells in IS-high/SS-low patients (P = 0.0286) and more resting NK cells in the IS-low/SS-high TME (P = 0.0108). KEGG pathway analysis revealed that overexpressed genes were enriched in the IL-17 and cytokine-cytokine receptor interaction pathways. The lncRNA DRAIC, a tumor suppressor, was consistently deactivated in the 10 TNBC cell lines. On the cBioPortal platform, we validated that 13% of ER-negative, HER2-unamplified BC harbored IL17RA deep deletion and 25% harbored TRAF3IP2 amplification. OncBioPortal datasets, then the nine altered TF genes derived from the X2K analysis showed significantly worse relapse-free survival in 2377 patients and OS in 4819 invasive BC patients than in the unaltered cohort.

Conclusion: Of note, the results of this integrated in silico study can only be generalized to approximately 17% of patients with TNBC, in which infiltrating stromal cells and immune cells play a determinant prognostic role.

Keywords: triple-negative breast cancer, immune cells, tumor microenvironment, IL-17 signaling pathway, immune evasion, in silico study

Introduction

In 2020, the global incidence of breast cancer was 2,261,419 diagnosed across 185 countries, accounting for 11.7% of all cancer types.1 Triple-negative breast cancer (TNBC) is characterized as estrogen receptor-negative, progesterone receptor-negative, and a lack of HER2 expression or amplification and accounts for approximately 15–20% of all breast cancers.2,3 TNBCs are heterogeneous (in terms of genomics, transcriptomics, and histopathology) and demonstrate the heterogeneity of response to anti-programmed death-1/ligand-1 (anti-PD-1/PD-L1) checkpoint inhibition immunotherapy.3–5 In fact, recent studies by Schmid et al revealed that PD-1 or PD-L1 checkpoint blockade immunotherapy combined with neoadjuvant chemotherapy (NAC) improved the pathological complete response (pCR) rate in the neoadjuvant setting; therefore, they proposed such an approach would probably improve the progression-free survival (PFS) of patients with advanced or metastatic TNBC staining positive for PD-L1 in tumor-infiltrating immune cells, if used as the first-line therapy.6–8 In contrast, other tumors are weakly immunogenic depending on the composition of the infiltrating immune cell populations and extrinsic factors (eg different metabolites or specific cytokines) enriched in the immune microenvironment.
tumor microenvironment (TME). Of note, recent bioinformatics research, based on bulk tumor gene expression data demonstrated that a low abundance of regulatory CD4+ T cells (Treg) was significantly associated with an increased pCR rate in TNBC patients after NAC.9

In a TNBC surgical specimen, untreated tumor cells typically represent approximately 60% of the cellular component, lymphoid and immune cells account for 20%, and stromal cells such as fibroblasts, histiocytes, endothelial cells, myofibroblasts, and adipocytes represent the remaining 20%.10 The primary function of stromal cells is to establish an immune response. The TME creates a chemokine-rich milieu inside, promoting the encounter between the boundary tumor cells and a variety of surrounding infiltrating immune cells in addition to cancer-associated fibroblasts, neo-vessels, neo-neurites, and other supportive tissues.11 Prior research has demonstrated that the prognosis of TNBC in terms of disease-free survival and disease-specific survival is worse in the basal-like immunosuppressed subtype and fare better in the basal-like immune-activated subtype, indicating that the immune TME plays a crucial role in the formation of either a tumor-permissive or tumor-expulsive milieu.4 A clinicopathological study also demonstrated that TNBC with a high number of tumor-infiltrating CD56-positive natural killer (NK) cells was associated with a more favorable disease-free survival.12 Additionally, research focusing on cancer-associated chemokines revealed that NK cells expressing abundant CXCR3 (also known as GPR9 and CD183) molecules on the cell surface are recruited by the chemokines CXCL9, CXCL10, or CXCL11 secreted by immune cells or stromal cells in the TME.13 However, it is still unclear what driving force explains the immunogenic or immunosuppressive phenotype in the TME in patients with TNBC.

Recently, there have been marked improvements in the research efficiency using cross-platform in-silico bioinformatics analyses of multi-omics, multi-layer cancer data.4,10,14–17 Hence, here, we leveraged multi-platform in-silico analyses of genomic data to investigate which infiltrating immune cell populations would be associated with the immune phenotype, and which signaling pathways could determine a subgroup of TNBC that is strongly immunogenic, thus, having a better prognosis. Lastly, we also aimed to determine whether transcription factor signatures derived from the RNA-seq data have prognostic value.

Materials and Methods
Ethics Statement and Study Design
This human data-based research study leveraged multiple publicly accessible RNA-seq datasets containing only mature, anonymous, and de-identified genetic and demographic data. The Institute Review Board of Kuang Tien General Hospital approved the study with a Certificate of Approval numbered KTGH-10458. This study was performed in accordance with the Declaration of Helsinki.

The schematic diagram (Figure 1) shows the study design and methodology adopted in this study. First, we downloaded from TCGA, a breast cancer gene expression RNA-seq dataset (TCGA-BRCA). Using this dataset, we sorted 115 patients with non-metaplastic TNBC, characterized by the lack of estrogen receptors, progesterone receptors, and HER2 non-overexpression or HER2-FISH un-amplification.

Estimation of the Immune/Stromal Scores
As mentioned above, the gene expression data and corresponding clinical information from a total of 1222 cases were retrieved from a publicly available dataset (TCGA-BRCA).18 After extracting the ER, PgR, and HER2 information of each sample, a total of 115 TNBC cases, 112 normal cases, and 994 non-TNBC cases were identified. To predict TNBC purity, we applied the Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data (ESTIMATE) algorithm to the normalized expression matrix to determine the immune/stromal scores of each TNBC patients. The immune score (IS; derived from the immune signature of 141 genes) and stromal score (SS; derived from the stromal signature of 141 genes) were calculated using the r-ESTIMATE package in R.19 Supplementary Table S1 presents a gene list of these immune-signature and stromal-signature genes. This algorithm rank-normalizes and rank-orders a set of gene expression values in a given sample and calculates the empirical cumulative distribution function from the signature gene set of the remaining genes; therefore, ESTIMATE can calculate the SS and IS from the RNA-seq data. The algorithm also allows the combination of these scores into an estimate score, used to infer DNA copy number-
based tumor purity, as per the following formula: Fraction of tumor cells in a clinical sample (namely, transcriptomic-based tumor purity) = cos (0.6049872018 + 0.0001467884 × [combined stromal-immune score]).

Correlation Analyses
The overall survival (primary prognosis endpoint-based) was estimated using Kaplan-Meier survival analysis. Log rank test was used to compare the survivals stratified by stage. Based on the calculated SS and IS, the corresponding patients were classified into two groups, and their prognoses were individually examined. A previously published method (maximally selected rank statistics), was employed for optimal cutoff identification using the survival in R to explore the relationship between the overall survival and the two groups of TNBC cases.

CIBERSORT Analysis
We used CIBERSORT X (https://cibersort.stanford.edu) to compare the proportion of infiltrating immune cells in the TME of IS-low/SS-high and IS-high/SS-low TNBC patients. The CIBERSORT accurately allows the evaluation of the relative levels of 22 immune cell phenotypes; such and analysis was performed using the immunedecov package in R. The immune cell fraction level divided by the cutoff value was 0 or 1 in the subsequent scoring formula.

Differential Expression Analysis
According to the ESTIMATE results, the intersection genes were selected based on the stromal/immune scores. The limma package in R was used to screen DEGs between normal and TNBC samples. Genes with a p-value ≤ 0.0001 and absolute log2 fold-change ≥ 4 were considered to be differentially expressed and extracted further for network construction. Heatmaps were generated using the pheatmap package in R. The generated heatmaps and volcano plots show the differentially expressed genes in either the IS-high/SS-low or IS-low/SS-high TNBC groups.

Gene Ontology and KEGG Pathway Enrichment Analyses
Using these DEGs defined as above for the two TNBC groups, we investigated the enriched signaling pathways based on GO terms and KEGG pathways. Functional enrichment analyses of GO terms, including cellular components, molecular functions, and biological processes, and KEGG pathways were performed using the clusterProfiler package in R. The functions among genes of interest were adjusted with a cutoff criterion of p < 0.05.
eXpression2Kinases Analysis

We utilized eXpression2Kinases (X2K), as reported elsewhere, to disclose the potential enrichment of transcription factors.26,27

Gene Set Variation Analysis

Gene set variation analysis (GSVA) is a bioinformatics framework that organizes gene expression data in the form of a pathway or signature summary.28 GSVA is a popular pathway-related immune infiltration and tumor mutational burden immune-related analysis. Here, we leveraged GSVA to provide an accurate definition of pathway enrichment between samples from different groups. We used the GSVA package in R to evaluate the t score and assigned the pathway activity conditions. We then used the pheatmap package in R to display the distinctions in pathway activation between normal tissues and those of IS-high/SS-low and IS-low/SS-high patients.

Validation in TNBC Cancer Cell Lines and Non-TNBC Cancer Cell Lines

In the Depmap platform (https://depmap.org/portal/), we used the methylation dataset (1 kb upstream transcription start sites) to compare the lncRNA methylation between the TNBC cell lines and non-TNBC cell lines (ER+/HER2+ or HER2-negative). We defined fractional methylation > 0.6 indicates “methylated” and < 0.6 “unmethylated”. Ten TNBC cell lines used were HCC2157, BT20, MDAMB231, HCC1395, HCC1937, HCC1599, MDAMB436, MDAMB468, HCC1806, and HCC1143. Five non-TNBC cell lines were BT474, UACC812, EFM192A, EFM19, and ZR751.

Validation in cBioPortal for Cancer Genomics

Finally, we validated our findings in a different, more extensive dataset of breast cancer via cBioPortal analysis.29,30

Statistical Analyses

We used the R package statistical software to implement the survival analysis with relevant R functions as survfit(), survdiff(), coxfit1. In addition, we used the StatsDirect version 3.3 to compute the Mann–Whitney U-test and the grouped linear regression. Finally, the built-in default statistical software performed all other statistical analyses on each platform.

Results

From the TCGA Breast Cancer dataset, gene expression data for 115 patients with triple-negative breast cancer and 112 normal adjacent tissues were retrieved. Figure 1 demonstrates the study design, flow, and methodology used in the current study. In this cohort of 115 TNBC patients, favorable overall survival as per the Kaplan-Meier curves was significantly correlated with an earlier disease stage (P = 0.0012; Figure 2).

The ESTIMATE algorithm computationally estimates the fraction of stromal and immune cells in the TME of all the 115 TNBC patients. The SS was high (SS-high) in 10 patients, and the IS was high (IS-high) in the ten other patients. There were no significant correlations between the SS or IS and the cancer stage in all the 115 patients (Supplementary Figure S1). However, as per the Kaplan-Meier analyses, SS-low patients showed a higher overall survival (OS) than that of SS-high patients (P = 0.0081; Supplementary Figure S2), while IS-high patients showed a higher OS than that of IS-low patients (P = 0.2; too few cases in the IS-high; Supplementary Figure S3). Of note, a strong correlation between both the SS and IS and the patients’ overall prognosis was observed. Expectedly, when compared with IS-low/SS-high patients, IS-high/SS-low patients showed a better OS (P = 0.045) (Figure 3). Cytoscape revealed the immune cell infiltration levels between samples grouped by IS-high/SS-low or IS-low/SS-high (Supplementary Figure S4). Interestingly, grouped linear regression showed a statistically significant increase in M2 macrophages in TNBC patients with the IS-high/SS-low phenotype. In line with these results, in the tumor microenvironment (TME), a higher proportion of M2 was also found in IS-low/SS-high patients, as estimated by CYBERSORTx (P < 0.001) (Supplementary Figure S5). Additionally, significantly more CD8+ cytotoxic T-cells, memory B cells (P = 0.0304), activated CD4+ memory T cells (P = 0.0056), follicular helper T cells (P = 0.0044), and activated NK cells (P = 0.0511) were also observed in IS-high/SS-low patients (14% vs 5.3%, P = 0.0143) (Table 1). On the other hand, more resting NK cells were observed in IS-low/SS-high patients (P = 0.0108).
Figure 2 Overall survival of 115 patients with triple-negative breast cancer stratified by cancer stage at diagnosis. Kaplan-Meier curves referring to different overall TNBC stages. Log rank test was used to compare the survivals stratified by stage. The higher stage correlates significantly with poorer survival (P = 0.0012).

Figure 3 Overall survival of 115 patients with triple-negative breast cancer stratified by the combination of the Stromal Score and Immune Score. Both scores were inferred by the ESTIMATE gene expression signatures. SS-low/IS-high patients were associated with excellent overall survival, whereas SS-high/IS-low patients showed the worst overall survival.
Moreover, DEG analysis showed 651 DEGs (284 upregulated, and 367 downregulated) in IS-high/SS-low patients, and 370 DEGs (187 upregulated, and 183 downregulated) in IS-low/SS-high patients (Supplementary Figure S6). Heatmaps and volcano plots of the DEGs in the context of these two groups are shown in Figure 4A–D. DESeq2 uses the Wald test to identify differentially expressed genes by comparing the immunogenic or immunotolerant classes and then the normal adjacent tissue. To put things in context, we displayed the top 30 DEGs ranked by the Wald statistic for the four categories by the combined immune and stromal scores. In IS-high/SS-low TNBC tumors, the top 30 upregulated DEGs were LAG3, CCNE1, GBP5, CDCA8, ETV7, CDCA3, CXCL10, KCNJ10, IDO1, PTG1, SKA1, CXCL11, IFNG, LOC105373098, TPX2, PLK1, WFDC21P, A2ML1, KIF2C, MMP11, IL21R, NDC80, IL4I1, AIM2, CXCL9, CDC20, CCL25, RTP3, NUF2, and CENPA; in contrast, the downregulated DEGs were MAMDC2, FAXDC2, VEGFD, TGFB3, NOV1, ADAMTS5, ECRG4, LMOD1, CAVIN2, ADAM33, TMTC1, DACH1, RERGL, MAB21L1, SLC7A2, AGTR1, PGR, SCN2B, PAMR1, MYH11, ATP1A2, ADAMTS9-AS2, CD300LG, CA4, CHL1, GPC3, ARHGAP20, CCL14, ABCA8, and ADGRL3. In IS-low/SS-high TNBC tumors, the top 30 upregulated DEGs were MMP11, CEMIP, COL11A1, COL10A1, CA9, LOC101929128, LOC157273.

### Table 1 Comparison of the Immune Cell Profiles in the Tumor Microenvironment (TME) of IS-High/SS-Low and IS-Low/SS-High Triple-Negative Breast Cancer Patients, Estimated Using CYBERSORTx

<table>
<thead>
<tr>
<th>Immune Cell Type in the TME</th>
<th>Immune Score-High/Stromal Score-Low TNBC (n = 10)</th>
<th>Immune Score-Low/Stromal Score-High TNBC (n = 10)</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive B cells</td>
<td>0.0363 ± 0.0532</td>
<td>0.0258 ± 0.0263</td>
<td>0.9288</td>
</tr>
<tr>
<td>Memory B cells</td>
<td>0.0168 ± 0.0121</td>
<td>0.0061 ± 0.0153</td>
<td>0.0304</td>
</tr>
<tr>
<td>Plasma cells</td>
<td>0.0162 ± 0.0376</td>
<td>0.0148 ± 0.0397</td>
<td>0.7599</td>
</tr>
<tr>
<td>CD8+ T cells</td>
<td>0.1403 ± 0.0934</td>
<td>0.0531 ± 0.0616</td>
<td>0.0143</td>
</tr>
<tr>
<td>Naive CD4+ T cells</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Resting CD4+ memory T cells</td>
<td>0.0956 ± 0.0557</td>
<td>0.1604 ± 0.1036</td>
<td>0.2176</td>
</tr>
<tr>
<td>Activated CD4+ memory T cells</td>
<td>0.0441 ± 0.0316</td>
<td>0.0176 ± 0.0159</td>
<td>0.0056</td>
</tr>
<tr>
<td>Follicular helper T cells</td>
<td>0.0700 ± 0.0235</td>
<td>0.0379 ± 0.0305</td>
<td>0.0044</td>
</tr>
<tr>
<td>Regulatory T cells</td>
<td>0.0517 ± 0.0409</td>
<td>0.0271 ± 0.0219</td>
<td>0.8767</td>
</tr>
<tr>
<td>Gamma delta T cells</td>
<td>0.0421 ± 0.0361</td>
<td>0.0259 ± 0.0268</td>
<td>0.2454</td>
</tr>
<tr>
<td>Resting NK cells</td>
<td>0</td>
<td>0.0199 ± 0.0395</td>
<td>0.0108</td>
</tr>
<tr>
<td>Activated NK cells</td>
<td>0.0360 ± 0.0257</td>
<td>0.0178 ± 0.0256</td>
<td>0.0511</td>
</tr>
<tr>
<td>Monocytes</td>
<td>0.0009 ± 0.0030</td>
<td>0.0164 ± 0.0428</td>
<td>0.2105</td>
</tr>
<tr>
<td>M0 macrophages</td>
<td>0.1196 ± 0.0540</td>
<td>0.1979 ± 0.1377</td>
<td>0.2799</td>
</tr>
<tr>
<td>M1 macrophages</td>
<td>0.1497 ± 0.0789</td>
<td>0.1357 ± 0.0670</td>
<td>0.8534</td>
</tr>
<tr>
<td>M2 macrophages</td>
<td>0.1043 ± 0.0628</td>
<td>0.1534 ± 0.0939</td>
<td>0.1431</td>
</tr>
<tr>
<td>Resting dendritic cells</td>
<td>0.0395 ± 0.0663</td>
<td>0.0211 ± 0.0178</td>
<td>0.5889</td>
</tr>
<tr>
<td>Activated dendritic cells</td>
<td>0.0001 ± 0.0004</td>
<td>0.0008 ± 0.0026</td>
<td>0.5</td>
</tr>
<tr>
<td>Resting mast cells</td>
<td>0.0344 ± 0.0214</td>
<td>0.0681 ± 0.0466</td>
<td>0.123</td>
</tr>
<tr>
<td>Activated mast cells</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Eosinophils</td>
<td>0.0008 ± 0.0028</td>
<td>0</td>
<td>&gt; 0.9999</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>0.0013 ± 0.0032</td>
<td>0.0002 ± 0.0006</td>
<td>0.582</td>
</tr>
</tbody>
</table>

Notes: Bold text denotes statistical significance. (%) Mean±SD; *Using Mann–Whitney U-test. CYBERSORTx: https://cibersortx.stanford.edu/index.php.

Moreover, DEG analysis showed 651 DEGs (284 upregulated, and 367 downregulated) in IS-high/SS-low patients, and 370 DEGs (187 upregulated, and 183 downregulated) in IS-low/SS-high patients (Supplementary Figure S6). Heatmaps and volcano plots of the DEGs in the context of these two groups are shown in Figure 4A–D. DESeq2 uses the Wald test to identify differentially expressed genes by comparing the immunogenic or immunotolerant classes and the normal adjacent tissue. To put things in context, we displayed the top 30 DEGs ranked by the Wald statistic for the four categories by the combined immune and stromal scores. In IS-high/SS-low TNBC tumors, the top 30 upregulated DEGs were LAG3, CCNE1, GBP5, CDCA8, ETV7, CDCA3, CXCL10, KCNJ10, IDO1, PTG1, SKA1, CXCL11, IFNG, LOC105373098, TPX2, PLK1, WFDC21P, A2ML1, KIF2C, MMP11, IL21R, NDC80, IL4I1, AIM2, CXCL9, CDC20, CCL25, RTP3, NUF2, and CENPA; in contrast, the downregulated DEGs were MAMDC2, FAXDC2, VEGFD, TGFB3, NOV1, ADAMTS5, ECRG4, LMOD1, CAVIN2, ADAM33, TMTC1, DACH1, RERGL, MAB21L1, SLC7A2, AGTR1, PGR, SCN2B, PAMR1, MYH11, ATP1A2, ADAMTS9-AS2, CD300LG, CA4, CHL1, GPC3, ARHGAP20, CCL14, ABCA8, and ADGRL3. In IS-low/SS-high TNBC tumors, the top 30 upregulated DEGs were MMP11, CEMIP, COL11A1, COL10A1, CA9, LOC101929128, LOC157273,
Figure 4 Heatmaps and volcano plots of the differentially expressed genes in SS-low/IS-high (A and B) and SS-high/IS-low (C and D) TNBC patients, and the corresponding KEGG pathway analysis (E). We labeled the DEGs satisfying the conditions of having an adjusted P-value < 0.0001 and |log2FoldChange| ≥ 5 in panels (B and D). The overexpressed DEGs in the context of both phenotypes are enriched in the IL-17 and cytokine-cytokine receptor interactions’ signaling pathways.
INHBA, CCL11, OLAH, CILP2, LYPD1, MMP1, HAPLN1, HOXB9, LINC02487, EPYC, GBP5, TDO2, CCKBR, CCN4, CCL11, LIPG, LINC01615, ANLN, GJB2, LINC00511, LINC00673, CXCL10, and SHISAL1; whereas, the downregulated DEGs were ABCA10, CAVIN2, ECRG4, SCN2B, ANGPTL7, VEGFD, MYH11, CD300LG, SCARA5, HPSE2, MYOC, HEPACAM2, LINC00993, HIF3A, SLC16A12, ATP1A2, PGM5-AS1, BTNL9, TNXB, ANKR3D30A, GDF10, PGR, PTPRQ, B3GALT1, NPY2R, PLPPR1, HSD17B13, DEFB132, LEP, and GLYAT.

KEGG pathway enrichment analysis showed that the overexpressed DEGs from both phenotypes were enriched in the IL-17 signaling pathway and viral protein interaction cytokine and cytokine receptor genes. Notably, the overexpressed DEGs in the context of the SS-low/IS-high phenotype were also enriched in other two cytokine-related pathways: cytokine-cytokine receptor interactions pathway and chemokine signaling pathway (Figure 4E).

Using X2K, we also inferred the transcription factors associated with the two immune phenotypes. In SS-high/IS-low TNBC patients, the inferred TFs were PPARG, HNF4A (also known as farnesoid X receptor, FXR), NR1H4, NR0B2, and MLXIPL, whereas the TF signature of SS-low/IS-high TNBC patients was composed of the PPARG, CEBPA, and MLXIPL TFs (Figure 5). Additionally, we performed a literature search on PubMed and discovered nine TF genes linked to IL-17-mediated signaling, including PPARG, CEBPA, MEOX1, KLF15, CD36, ZNF750, EZH2, HNF4A, and NR0B2 (Table 2).

In addition, GSVA also suggested the involvement of additional pathways, including the JAK-STAT signaling, T cell receptor signaling, B cell receptor signaling, cytokine-cytokine receptor interaction, TGF-β signaling, and PPAR signaling pathways (Supplementary Figure S7; Figure S7A shows the GSVA of the 370 DEGs of the SS-high/IS-low subgroup and Figure S7B shows the GSVA of the 651 DEGs of the SS-low/IS-high subgroup).

Validation of the Results in the TNBC Cell Lines

IL-17 signaling will kick off the NFκB signaling and induce M1 to M2-like transformation in TME macrophages. Long non-coding RNA, DRAIC, a tumor suppressor, would interact with the IκB kinase (IKK) to inhibit NFκB activation. In the Depmap platform, we used the methylation dataset (1 kb upstream transcription start sites) to compare the lncRNA DRAIC methylation between the ten TNBC cell lines and five non-TNBC cell lines (ER+/HER2+ or HER2-negative). We defined fractional methylation > 0.6 indicates “methylated” and < 0.6 “unmethylated”. Our study results showed that 10 TNBC cell lines had methylated lncRNA DRAIC, whereas all five non-TNBC cell lines were unmethylated (Figure 6). Therefore, the lncRNA DRAIC tumor suppressor was consistently deactivated in the 10 TNBC cell lines.

Validation of the Results in the cBioPortal

IL-17 is a proinflammatory cytokine that signals mainly via the TRAF3 Interacting protein 2 (TRAF3IP2), as reported previously. TRAF3IP2 is an inflammatory mediator and upstream regulator of several crucial transcription factors,
such as AP-1 and NF-κB. Act1, an essential component in IL-17 signaling complex, is encoded by the gene TRAF3IP2. Finally, we used the Oncoprint from cBioPortal to validate the frequency of IL-17 genes in TNBC. We discovered that 13% of ER-negative and HER2-FISH unamplified breast cancers harbored IL17RA deep deletions and 25% harbored TRAF3IP2 amplifications (Supplementary Figure S8).

Interestingly, we also discovered using the cBioPortal platform that aberrations in the nine TF genes mentioned above are associated with a worse prognosis, as per the relapse-free survival of 2377 patients (log-rank, P = 0.00007) and the overall survival of a larger group of 4819 patients (log-rank, P = 0.001697) (Figure 7). Finally, we interrogated the different expression of all the related genes of the IL-17 signaling pathway in the two polarized immune-stroma-scored TNBCs. The hierarchical clustering heatmaps show evidence in supporting that most IL-17 pathway genes had differentially expressed in the tumor specimens than the normal adjacent tissue counterparts (Supplementary Figure S9). In addition, it is notable that IL-17D and IL-17F were more differentially expressed than the normal adjacent tissues in the stromal-low immune-high tumors.

**Discussion**

Several recent studies have shed light on the pivotal role of the TME in the shaping of the tumor behavior in TNBC. Act1, an essential component in IL-17 signaling complex, is encoded by the gene TRAF3IP2. Finally, we used the Oncoprint from cBioPortal to validate the frequency of IL-17 genes in TNBC. We discovered that 13% of ER-negative and HER2-FISH unamplified breast cancers harbored IL17RA deep deletions and 25% harbored TRAF3IP2 amplifications (Supplementary Figure S8).

Interestingly, we also discovered using the cBioPortal platform that aberrations in the nine TF genes mentioned above are associated with a worse prognosis, as per the relapse-free survival of 2377 patients (log-rank, P = 0.00007) and the overall survival of a larger group of 4819 patients (log-rank, P = 0.001697) (Figure 7). Finally, we interrogated the different expression of all the related genes of the IL-17 signaling pathway in the two polarized immune-stroma-scored TNBCs. The hierarchical clustering heatmaps show evidence in supporting that most IL-17 pathway genes had differentially expressed in the tumor specimens than the normal adjacent tissue counterparts (Supplementary Figure S9). In addition, it is notable that IL-17D and IL-17F were more differentially expressed than the normal adjacent tissues in the stromal-low immune-high tumors.

**Discussion**

Several recent studies have shed light on the pivotal role of the TME in the shaping of the tumor behavior in TNBC. Act1, an essential component in IL-17 signaling complex, is encoded by the gene TRAF3IP2. Finally, we used the Oncoprint from cBioPortal to validate the frequency of IL-17 genes in TNBC. We discovered that 13% of ER-negative and HER2-FISH unamplified breast cancers harbored IL17RA deep deletions and 25% harbored TRAF3IP2 amplifications (Supplementary Figure S8).

Interestingly, we also discovered using the cBioPortal platform that aberrations in the nine TF genes mentioned above are associated with a worse prognosis, as per the relapse-free survival of 2377 patients (log-rank, P = 0.00007) and the overall survival of a larger group of 4819 patients (log-rank, P = 0.001697) (Figure 7). Finally, we interrogated the different expression of all the related genes of the IL-17 signaling pathway in the two polarized immune-stroma-scored TNBCs. The hierarchical clustering heatmaps show evidence in supporting that most IL-17 pathway genes had differentially expressed in the tumor specimens than the normal adjacent tissue counterparts (Supplementary Figure S9). In addition, it is notable that IL-17D and IL-17F were more differentially expressed than the normal adjacent tissues in the stromal-low immune-high tumors.

**Table 2 Selection of the Transcription Factors Linked to Interleukin-17-Mediated Signaling**

<table>
<thead>
<tr>
<th>Transcription Factor (TF)</th>
<th>Investigated Relationship</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPARγ</td>
<td>PPARγ-induced SOCS3 expression prevents IL-17-mediated cancer growth.</td>
<td>[31]</td>
</tr>
<tr>
<td>CEBPA (=C/EBPα)</td>
<td>Transcription factor that coordinates proliferation arrest and the differentiation of myeloid progenitors, adipocytes, hepatocytes, and cells of the lung and the placenta. IL-17 suppresses expression of several pro-adipogenic TFs, including PPARγ and CEBPA.</td>
<td>[32,33]</td>
</tr>
<tr>
<td>MEOX1</td>
<td>TGF-β1 transcriptionally regulates MEOX1 expression via Smad2/3 in adult human dermal fibroblasts, thus promoting cell migration.</td>
<td>[34]</td>
</tr>
<tr>
<td>KLF15</td>
<td>Specifically, IL-17 suppresses KLF15, a pro-adipogenic TF, and enhances expression of KLF2 and KLF3, which are anti-adipogenic. Thus, IL-17 suppresses adipogenesis at least in part through the combined effects of TFs that regulate adipocyte differentiation.</td>
<td>[32]</td>
</tr>
<tr>
<td>CD36</td>
<td>In the presence of palmitic acid (PA), IL-17a could directly increase the cellular uptake of PA, leading to the proliferation of ovarian cancer cells via the IL-17a/IL-17RA/p-STAT3/FABP4 axis rather than via CD36.</td>
<td>[35]</td>
</tr>
<tr>
<td>ZNF750</td>
<td>ZNF750 is the p63 target gene. The levels of inflammatory cytokines (IL17d and Tnfsf15) were significantly reduced by Rbm38 deficiency in senescence-resistant Rbm(38/-);TAp63(+/-) mouse livers and MEFs. Rbm38 and p63 function as intergenic suppressors in aging and tumorigenesis.</td>
<td>[36]</td>
</tr>
<tr>
<td>EZH2</td>
<td>EZH2 positively regulate the expression of IL-17a and IL-17f. The inducible binding of EZH2 at the IL-17a promoter was dependent on signaling pathways downstream of the TCR. IL-17f bears 50% homology to IL-17a and has recently been suggested to play a role in inflammation.</td>
<td>[37,38]</td>
</tr>
<tr>
<td>HNF4A (= farnesoid X receptor, FXR)</td>
<td>In addition to the classical Jak-Stat antiviral signaling pathway, IFN-λ1 inhibits hepatitis C virus replication through the suppression of miRNA-122 transcription via an inflammatory Stat 3-HNF4A feedback loop. Inflammatory feedback circuits activated by IFNs during chronic inflammation expose non-responders to the risk of hepatocellular carcinoma.</td>
<td>[39]</td>
</tr>
<tr>
<td>NR0B2 (= SHP)</td>
<td>This TF is a tumor suppressor. Both FXR (-/-) and NR0B2(-/-) mice develop spontaneous hepatocellular carcinoma. Upregulated NR0B2 will regulate the IL-6-dependent pathway.</td>
<td>[40,41]</td>
</tr>
</tbody>
</table>
Methylation fraction of the IncRNA DRAIC was compared between 10 TNBC cancer cell lines and 5 non-TNBC cancer cell lines. Long non-coding RNA, DRAIC, a tumor suppressor, would interact with IKK to inhibit NFκB activation. In the Depmap platform, we used the methylation dataset (1 kb upstream transcription start sites) to compare the PPARG methylation between the TNBC cell lines and non-TNBC cell lines (ER+/HER2+ or HER2-negative). We defined fractional methylation > 0.6 indicates “methylated” and < 0.6 “unmethylated”. Our study results showed that 10 TNBC cell lines (HCC2157, BT20, MDAMB231, HCC1395, HCC1937, HCC1599, MDAMB436, MDAMB468, HCC1806, and HCC1143) had methylated IncRNA DRAIC, whereas all five non-TNBC cell lines (BT474, UACC812, EFM192A, EFM19, and ZR751) were unmethylated.

Validation using the cBioPortal demonstrates the relapse-free survival in 2377 patients and the overall survival in 4819 patients with invasive breast carcinoma stratified based on the alteration of the transcription factor genes PPARG, CEBPA, MEOX1, CD36, ZNF750, KLF15, EZH2, HNF4A, and NR0B2. The altered group refers to any alteration in one of these transcription factors.
Given the paucity of studies investigating the mechanism behind the strikingly different prognoses of immunosuppressive or immunogenic TNBCs, our study aimed to explore the differences in stromal cells (particularly immune cells) within the TME, as well as in the enriched DEGs in the search for the driving force behind the formation of these two different phenotypes. Interestingly, we discovered that the composition of immune cells reacting to TNBC tumor cells was strikingly different in the highly immuno-weak stromal TME and the strong stromal-weak immunogenic TME. Additionally, we discovered that the immunosuppressive type (SS-high/IS-low) showed gene signatures enriched in the IL-17, and viral-cytokine and cytokine receptor interactions’ pathways. Of note, this study is unique considering the demonstration that altered TF genes derived from IL-17-mediated signaling showed strong significance with respect to both the relapse-free and overall survival, as per the analysis of another extensive dataset.

More than 1600 known (or likely) human TF genes represent approximately 8% of the human genes. Mutations in TF genes are often highly deleterious in humans. Nine TF genes, PPARG, CEBPA, MEOXI, CD36, ZNF750, KLF15, EZH2, HNF4A, and NR0B2, were enriched in our bioinformatics analysis. Importantly, this is the first time that the alteration in one or more of the nine TF in the context of TNBC was linked to the IL-17 signaling pathway and associated with significantly poorer prognosis in terms of relapse-free survival and overall survival in a large number of breast cancer patients, excluding ER-positive and HER2-amplified cases.

In cancer tissues, activated T cells secrete the proinflammatory cytokine IL-17, which through the regulation of the MAPKs and NF-κB activities, gives rise to the increased expression of IL-6 and cyclooxygenase-2. A score of research recapitulated these observations and revealed a more detailed mechanistic understanding. A plethora of studies have revealed increased IL-17A levels in ER-negative or triple-negative breast cancer. In fact, the upregulation of IL-17A signaling is associated with increased expression of programmed cell death protein 1 (PD-1) and programmed death-ligand 1 (PD-L1) in breast cancer with low ER expression, which may elevate the infiltration of CD8+ T cells in the tumor tissues. Of note, CD4+ T cells among tumor-infiltrating lymphocytes in the TME are the main source of IL-17. IL-17-mediated downstream signaling plays a critical role in the TME, inducing the expression of different genes either to switch on pro-tumor effector cytokines or to inhibit tumor growth in a context- and system-dependent manner. For example, IL-17 induces the production of IL-6, which in turn induces STAT3 (signal transducer and activator of transcription 3); in fact, IL-6 orchestrates both the increase in the recruitment of suppressive tumor-associated myeloid cells (MDSCs) and impacts their ability to inhibit anti-tumor T-cell responses. Early on, a recent study recovered how an IL-17–mediated paracrine network could recruit immature myeloid cells into the TME, causing tumor progression; this study demonstrated that through NF-κB and ERK signaling, tumor-infiltrating T helper type 17 (Th17) cells and IL-17 would induce granulocyte colony-stimulating factor expression, which is crucial for the mobilization of immature myeloid cells into the TME. A recent study of the role IL-17A plays in the TNBC’s TME identified that IL-17A stimulates the angiogenic and migratory activity of TNBC cells and modulates the TME’s immune landscape towards a pro-metastatic phenotype. In contrast, Bianchini et al demonstrated that a subset of highly proliferative, ER-negative breast cancers with high expression of a B-cell/plasma cell stromal metagene corresponding to immune functions and extracellular matrix components were associated with a favorable prognosis. These studies lend support to the composition of cellular constituents in the TME, and their association with the patient’s prognosis.

Using the methylation dataset, our in-silico cell line study in the DepMap platform demonstrates that 10 TNBC cell lines had methylated IncRNA DRAIC. In contrast, all five non-TNBC cell lines were unmethylated. This tumor suppressor, consistently deactivated in the 10 TNBC cell lines, shall require further research and investigation into its clinical significance.

Still, regarding the TF genes associated with TNBC, a previous study suggested that EZH2 plays decisive roles in immune cells (eg, T cells, NK cells, dendritic cells, and macrophages) in the tumor microenvironment. Liu et al demonstrated a novel function of PPARγ in lymphocyte trafficking and the cross-talk between Th17 and B cells. Additionally, investigators from the Institut Curie and INSEERM found that a high Th17 metagene was associated with a good prognosis in T cell non-inflamed type TNBC, suggesting Th17 is a novel prognostic composite biomarker. Altogether, these studies clearly support the notion that integrating immune cells and tumor molecular diversity is an efficient strategy for the prognostic stratification of cancer patients.

Several recently published studies provide some functional data to show how the IL-17-mediated signaling pathway drives immune response and modulates the therapeutic response to chemotherapy or immunotherapy in solid tumors,
including breast cancers. These data will inform future therapeutic strategy design both pre-clinically and clinically. Some exciting and relevant research findings are discussed. In the TME, adenosine is a well-known immunosuppressive molecule. Adenosine and adenosine triphosphate (ATP) are the most abundant metabolites within a cell and extracellular space, acting as an autocrine and paracrine messenger. While ATP acts as an accelerator to promote proinflammatory activities, adenosine, via the Gs-coupled A2a and A2b receptors, suppresses various immune cells. Thibaudin et al identified that in breast cancer tumors, Th17 cells would express the ectonucleotidases, CD73 and CD39, which are the enzymes that degrade extracellular ATP into adenosine. The investigators further observed that the adenosine molecule would suppress T cell immune responses. Wang et al investigated how IL-17A regulated natural killer (NK) cell activity and found that by restraining IL-15-driven NK cell terminal maturation, IL-17A constrained the NK cell's anti-cancer activity. Th17 cells also contribute to the acquired resistance to PD-L1 blockade immunotherapy and MEK inhibitor in KRAS/p53-mutant lung cancers. Th17 cells create resistance by secreting IL-17 and IL-22 cytokines. The investigators established an in vivo model to prove the principle that through antibody depletion of IL-17A, combining the MEK inhibition and PD-L1 blockade markedly reduce the therapy-induced acquired resistance. 

Finally, the purpose of the current research was neither to develop a predictive molecular panel biomarker, nor to identify innate resistance versus sensitive phenotypes to checkpoint inhibitor immunotherapy; we aimed to investigate the factors behind the TME’s immune-stromal state and to look for the systemic functions of the involved molecular signaling pathways. However, this study is not without limitations. One of them is the available low number of cases used to separate tumors into two extreme immune-stromal phenotypes.

**Conclusions**

Overall, our data suggest that the disclosure of the distinct molecular anatomy of the TME in patients with TNBC (not of the cancer cells per se) will assist in the determination of the ultimate tumor behavior. Of note, the results of this integrated in silico study can only be generalized to approximately 17% of patients with TNBC, in which infiltrating stromal cells and immune cells play a determinant prognostic role. Remarkably, our data suggest that the determining factor for the differences in the tumor immune microenvironment (immunogenic versus immunosuppressive) can be explained by IL-17 signaling and its paracrine network.

**Data Sharing Statement**

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

**Acknowledgments**

The results published here are based on data generated by the TCGA Research Network: https://www.cancer.gov/tcga. Part of the results in this study were presented as an e-poster at the American Association for Cancer Research 2021 Annual Meeting, held during April 10–15, 2021. The prior version without peer review has been posted as a preprint with DOI: 10.21203/rs.3.rs-495771/v1.

**Funding**

This study was partly supported by a research fund from the Ministry of Science and Technology, Taiwan, with the grant number as MOST 109-2321-B-468-001. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Disclosure**

The authors declare that they have no competing interests in this work.
References


