Immune-Related Gene Signatures to Predict the Effectiveness of Chemoimmunotherapy in Triple-Negative Breast Cancer Using Exploratory Subgroup Discovery

Olha Kholod 1, William I. Basket 1, Jonathan B. Mitchem 1,2,3,4, Jussuf T. Kaifi 1,2,3,4, Richard D. Hammer 1,4,5,6, Christos N. Papageorgiou 4,6 and Chi-Ren Shyu 1,7, *  

1 MU Institute for Data Science and Informatics, University of Missouri, Columbia, MO 65212, USA  
2 Department of Surgery, School of Medicine, University of Missouri, Columbia, MO 65212, USA  
3 Harry S. Truman Memorial Veterans’ Hospital, Columbia, MO 65201, USA  
4 Ellis Fischel Cancer Center, University of Missouri, Columbia, MO 65211, USA  
5 Department of Pathology & Anatomical Sciences, School of Medicine, University of Missouri, Columbia, MO 65212, USA  
6 Department of Medicine, School of Medicine, University of Missouri, Columbia, MO 65212, USA  
7 Department of Electrical Engineering & Computer Science, University of Missouri, Columbia, MO 65212, USA  
* Correspondence: shyuc@missouri.edu

Simple Summary: Chemoimmunotherapy combinations have transformed the treatment landscape for patients with triple-negative breast cancer (TNBC). However, the discovery of immune-related biomarkers is needed to optimally identify patients requiring the addition of immune-checkpoint inhibitors (ICIs) to chemotherapy. In this study, we identified immune-related gene signatures via exploratory subgroup discovery algorithm that substantially increase the odds of partial remission for TNBC patients on anti-PD-L1+chemotherapy regimen. We have also uncovered distinct cell populations for TNBC patients with various treatment outcomes. Our framework may result in better risk stratification for TNBC patients that undergo chemoimmunotherapy and lead to overall improvement of their health outcomes in the future.

Abstract: Triple-negative breast cancer (TNBC) is an aggressive subtype of breast cancer with limited therapeutic options. Although immunotherapy has shown potential in TNBC patients, clinical studies have only demonstrated a modest response. Therefore, the exploration of immunotherapy in combination with chemotherapy is warranted. In this project we identified immune-related gene signatures for TNBC patients that may explain differences in patients’ outcomes after anti-PD-L1+chemotherapy treatment. First, we ran the exploratory subgroup discovery algorithm on the TNBC dataset comprised of 422 patients across 24 studies. Secondly, we narrowed down the search to twelve homogenous subgroups based on tumor mutational burden (TMB, low or high), relapse status (disease-free or recurred), tumor cellularity (high, low and moderate), menopausal status (pre- or post) and tumor stage (I, II and III). For each subgroup we identified a union of the top 10% of genotypic patterns. Furthermore, we employed a multinomial regression model to predict significant genotypic patterns that would be linked to partial remission after anti-PD-L1+chemotherapy treatment. Finally, we uncovered distinct immune cell populations (T-cells, B-cells, Myeloid, NK-cells) for TNBC patients with various treatment outcomes. CD4-Tn-LEF1 and CD4-CXCL13 T-cells were linked to partial remission on anti-PD-L1+chemotherapy treatment. Our informatics pipeline may help to select better responders to chemoimmunotherapy, as well as pinpoint the underlying mechanisms of drug resistance in TNBC patients at single-cell resolution.

Keywords: triple-negative breast cancer; exploratory subgroup discovery; chemoimmunotherapy
1. Introduction

Triple-negative breast cancer (TNBC) occurs in about 10 to 20% of diagnosed breast cancers and is defined by the absence or minimal expression of estrogen receptor (ER), progesterone receptor (PR) and epidermal growth factor receptor 2 (HER2) [1,2]. Due to its aggressive clinical phenotype and limited response to hormonal therapy, one in three TNBC patients will likely relapse within the first three years of primary diagnosis [3]. Although numerous therapeutic agents have been evaluated for the treatment of early TNBC [4], only Olaparib has been approved for the treatment of the small group of patients with high-risk TNBC harboring germline BRCA1 or BRCA2 pathogenic variants in the adjuvant setting [5]. The emergence of cancer immunotherapy, however, is altering the paradigm in TNBC treatment.

TNBC, unlike other breast cancer subtypes, has high tumor mutational burden (TMB), which has been correlated with responsiveness to immune checkpoint inhibitors (ICIs) [6]. Indeed, checkpoint inhibition with the anti-PD1 antibody Pembrolizumab has been approved for advanced-stage, PD-L1 positive TNBC due to improved outcomes when combined with frontline chemotherapy [7]. Interestingly, ICIs are more effective in treating TNBC when given early in the course of the disease, which may be a result of immune escape mechanisms emerging as the condition progresses [8]. More recently, results from the KEYNOTE-522 trial indicated that adding checkpoint inhibition in the early stage setting does in fact improve long-term outcomes [9]. However, subgroup analyses did not pinpoint any strongly predictive biomarkers. For example, PD-L1 expression did not distinguish responders from non-responders in the early setting, with both PD-L1-negative and PD-L1-positive patients obtaining a benefit from Pembrolizumab. Moreover, the addition of immunotherapy increased adverse effects (AEs) [10]. In another study—IMPASSION131—the combination of Paclitaxel with the PD-L1 inhibitor Atezolizumab failed to improve progression-free survival (PFS) or overall survival (OS) in TNBC patients [11]. These findings could be due to imbalances in prognostic features or accidental discoveries in a relatively small trial. Therefore, the exploration of immune-related biomarkers is needed to optimally identify patients requiring the addition of ICIs to chemotherapy [12,13].

In this work we determined homogenous TNBC subgroups based on both phenotypic and genotypic parameters using exploratory subgroup mining. We have also identified significant predictors that increase chances of partial remission in TNBC patients on chemoimmunotherapy treatment using multinomial regression model on TNBC scRNA-seq dataset. Lastly, we uncovered distinct immune cell populations (T-cells, B-cells, Myeloid, NK-cells) for TNBC patients with various treatment outcomes. We interpreted our results using biomedical knowledge, including findings from existing clinical trials, immunohistochemistry experiments and functional characterization of specific genes. The proposed informatics pipeline may assist health care professionals in the selection of chemoimmunotherapy responders, as well as determine the underlying causes of drug resistance in TNBC patients at a single-cell level and resolution.

2. Materials and Methods

2.1. Data Mapping

In this study we employed two datasets. Each dataset consisted of multiple phenotypic (either categorical or continuous) and genotypic (continuous only) variables. Each categorical variable was labeled based on the National Comprehensive Cancer Network (NCCN) Guidelines in Oncology [14]. For example, relapse-free status was categorized as (1) disease-free or (2) recurred. Continuous variables were converted into categoric variables by grouping values into several categories. For example, normalized gene expression values were categorized as (1) downregulated, (2) upregulated, or (3) non-differentially expressed.

The first TNBC dataset comprised of 422 patients. These patients were selected from 24 breast cancer studies available at the cBioPortal platform [15]. The final dataset included breast cancer patients based on the following immunohistochemical profile: ER-negative, PR-negative and HER2-negative. This dataset consisted of 12 phenotypic variables, in-
cluding clinical-pathologic data (age at diagnosis, menopausal status, tumor type, tumor stage, tumor cellularity, histologic grade, TMB), treatment regimen (chemotherapy, radiotherapy, hormone therapy) and survival status (overall survival status, relapse-free status). There were 1067 genotypic variables in the form of normalized gene expression values derived from human immunome (immune-related genes) and human kinome (protein kinase genes).

The second TNBC dataset consisted of scRNA-seq profiles for 22 TNBC patients that underwent chemotherapy (Paclitaxel) or chemoimmunotherapy treatment (Paclitaxel with Atezolizumab) [16]. For this study we selected six phenotypic variables, including information about treatment timeline (pre-, post-treatment, progression), tissue type (tumor or blood), tumor site (brain, breast, chest wall, liver, lymph nodes), treatment type (anti-PD-L1+chemotherapy or chemotherapy only), treatment response (partial response (PR), stable disease (SD), progressive disease (PD) and cell cluster (T-cells, B-cells, Myeloid, NK-cells)). We used the same genotypic variables as in the TNBC subgroup discovery dataset.

2.2. The Informatics Pipeline

Our informatics pipeline has three modules: (1) exploratory subgroup discovery, (2) inference module based on multinomial regression model and (3) immune cell populations discovery. Our goal was two-fold: (1) to identify significant genes from exploratory subgroup discovery that increase odds of having partial remission after anti-PD-L1+chemotherapy treatment and (2) to uncover unique immune cell populations for TNBC subgroups with various treatment outcomes.

The main goal of exploratory subgroup discovery module was to determine homogenous patient subgroups based on expanatory phenotypic characteristics (Module A on Figure 1), where prevailing number of patients in that subgroup exemplify distinctive genotypic patterns (Module B on Figure 1). Each genotypic pattern had been represented as a combination of differentially expressed genes. For example, the genotypic pattern may consist of three genes: upregulated EGFR, downregulated MTOR and upregulated MAPK1 genes. On the first step, the algorithm determines the base subgroup (e.g., Chemotherapy = Yes) contingent on the most significant contrast against the rest of the population. On the next inclusion step it adds a new phenotypic variable, e.g., TMB = High, to the previous subgroup to generate a more focused subgroup (Chemotherapy = Yes and TMB = High). Subsequently, the exclusion step is employed to remove a less relevant inclusion move after each inclusion step. The exploratory search selects multiple paths that form multiple subgroups and have equally relevant genotypic patterns within each subgroup. When the algorithm reaches the most focused subgroup with the highest contrast score that cannot be further increased, the search would be terminated. Support [17] and growth rate [18] were used to measure the frequency for a specific genotypic pattern in the homogenous subgroup. We then applied a J-value [19] to prioritize each subgroup based on the relevance (contrasts) for all patterns in each subgroup [20].

To find significant predictors of partial remission on anti-PD-L1+chemotherapy regimen, we employed multinomial regression model (Module C on Figure 1) on the scRNA-seq TNBC dataset. The outcome variable was categorical and represented as a combination of treatment response, treatment timeline, and treatment type. For example, the level of outcome variable can be encoded as SD-Post_treatment-Chemo meaning that a fraction of TNBC patients achieved stable disease after treatment with chemotherapy only. Overall, there were ten levels of outcome variable. We set PD-Post_treatment-Chemo—progressive disease after chemotherapy—as a baseline for the model. The continuous covariates were encoded as genes with normalized gene expression values identified as a top 10% of genotypic patterns in the exploratory subgroup discovery stage. We used the multinom function from the nnet package [21] to estimate a multinomial logistic regression model. We computed p-values via two-tailed z-test to identify significant predictors of response to anti-PD-L1+chemotherapy treatment.
we generated heatmap plots for each condition in every TNBC subgroup of interest. For each subgroup, we compared immune cell populations in these two conditions: (1) anti-PD-L1+chemotherapy, post treatment, partial remission and (2) chemotherapy, post treatment, progression. The immune cell populations discovery module (Module D on Figure 1) determined distinct immune cell populations (T-cells, B-cells, Myeloid, NK-cells) for TNBC patients with various treatment outcomes. Each TNBC subgroup had two conditions: (1) anti-PD-L1+chemotherapy, post treatment, partial remission and (2) chemotherapy, post treatment, PD. Using the top 10% of genotypic patterns from exploratory mining stage as an input, we generated heatmap plots for each condition in every TNBC subgroup of interest. For example, NME3 gene was represented as a geometric mean of NME3 expression values in CD4-Tcm-LMNA cells [22]. Finally, we compared immune cell populations in these two conditions to identify mutually exclusive cell populations that were associated with either partial remission after anti-PD-L1+chemotherapy treatment or progressive disease after chemotherapy treatment.

3. Results

3.1. The Identification of Homogenous TNBC Subgroups

First, we ran the exploratory subgroup discovery algorithm on the TNBC dataset described in Section 2.1. The algorithm revealed 11,944 subgroups. We focused our analysis of the 460 subgroups where TNBC patients had undergone chemotherapy. On the next step, we narrowed down the search to homogenous subgroups based on TMB (low or high), relapse status (disease-free or recurrent), tumor cellularity (high, low and moderate), menopausal status (pre- or post) and tumor stage (I, II and III). Since the lengths of genotypic patterns vary (up to 5 genes), we decided to make a union of top 10% of genotypic patterns for each subgroup of interest. Let us assume that each genotypic pattern is a set of elements, where each element is a unique differentially expressed gene (e.g., upregulated MTOR gene). The union would represent a set of a collection of genotypic patterns, where each element would not be repetitive. These genotypic patterns were used as covariates for the multinomial regression model in the next section.

3.2. Significant Predictors of Partial Remission after Anti-PD-L1+Chemotherapy

The multinomial regression model on scRNA-seq TNBC dataset was able to identify significant predictors from exploratory subgroup discovery results that increase odds of
having partial remission after anti-PD-L1+chemotherapy treatment versus progressive disease after chemotherapy (Table 1).

Table 1. Significant predictors that increase odds of partial remission after anti-PD-L1+chemotherapy identified by our informatics pipeline.

| Subgroups                              | Predictors | Coefficients | p-Values | Odds Ratio |
|----------------------------------------|------------|--------------|----------|------------|
| Chemotherapy (Yes) TMB (High)          | ACVR1B     | 0.4506       | 0.0136   | 1.5692     |
|                                        | PDPK1      | 0.1271       | 0.0169   | 1.1355     |
| Chemotherapy (Yes) TMB (Low)           | CLK3       | 0.1402       | 0.0017   | 1.1505     |
|                                        | TAOK2      | 0.6603       | <0.0001  | 1.9354     |
| Chemotherapy (Yes) Relapse Status (Disease Free) | CDK9        | 0.4152       | <0.0001  | 1.5146     |
|                                        | CFP        | 0.2368       | 0.0438   | 1.2673     |
|                                        | VRK3       | 0.1591       | 0.0011   | 1.1725     |
| Chemotherapy (Yes) Relapse Status (Recurred) | BUB1       | 0.6563       | 0.0183   | 1.9278     |
|                                        | BAZ1B      | 0.3529       | <0.0001  | 1.4232     |
| Chemotherapy (Yes) Menopausal Status (Pre) | PD1K1L     | 0.6405       | 1.3243×10⁻⁵ | 1.8974     |
|                                        | KIR2DL4    | 1.8931       | 0.0094   | 6.6403     |
|                                        | MAPK3      | 0.8886       | <0.0001  | 2.4319     |
|                                        | STK24      | 0.2842       | <0.0001  | 1.3287     |
| Chemotherapy (Yes) Menopausal Status (Post) | IFI16      | 0.0665       | 0.0205   | 1.0688     |
|                                        | CSK        | 0.2397       | <0.0001  | 1.2709     |
|                                        | TAP2       | 0.2165       | 0.0185   | 1.2417     |
|                                        | TIGIT      | 0.5160       | <0.0001  | 1.6794     |
| Chemotherapy (Yes) Tumor Stage (I)      | CCR6       | 0.1307       | 0.0023   | 1.1396     |
|                                        | BCL10      | 0.1222       | 0.0064   | 1.1300     |
|                                        | PRKCA      | 0.9982       | <0.0001  | 2.7135     |
|                                        | EPB16      | 0.8559       | <0.0001  | 2.3536     |
|                                        | IFNAR2     | 0.4546       | 6.4298×10⁻⁶ | 1.5756     |
| Chemotherapy (Yes) Tumor Stage (II)     | PD1K1L     | 0.4952       | 0.0003   | 1.6409     |
|                                        | RPS6KA5    | 0.2836       | <0.0001  | 1.3279     |
|                                        | IKZF2      | 0.4714       | <0.0001  | 1.6023     |
| Chemotherapy (Yes) Tumor Stage (III)    | RIOK3      | 0.1683       | 2.4615×10⁻⁵ | 1.1833     |
|                                        | PRKCA      | 1.0213       | <0.0001  | 2.7770     |
| Chemotherapy (Yes) Tumor Stage (III)    | IFIH1      | 0.3366       | 1.2977×10⁻⁵ | 0.1773     |
|                                        | CDKL5      | 0.7786       | 0.0066   | 2.1875     |

Next, we highlight the importance of identified phenotypic features from Table 1 for TNBC patient outcomes. Using literature, high-TMB TNBC status may benefit specifically from ICIs in combination with chemotherapy [23] or ICIs alone [24]. TNBC patients have high TMB due to accumulation of genomic instability, which leads to the production neoantigens, thereby resulting in strong effector cell responses [25]. TNBC tumors have a "hot tumor phenotype", which characterized by a high degree of immune infiltration and associated with improved survival outcomes regardless of tumor stage, molecular subtype, PD-L1 status, age and treatment schedule [26]. The IMpassion130 trial tested immunotherapy agent Durvalumab in combination with chemotherapy or chemotherapy alone on 149 early stage TNBC patients. Median TMB was significantly higher in patients with pathologic complete response (pCR) (median 1.87 versus 1.39, \( p = 0.005 \)), and odds ratios for pCR per mut/MB were 2.06 (95% CI 1.33–3.20) among all patients, 1.72 (95% CI 1.00–3.13) in the Durvalumab arm, and 2.82 (95% CI 1.21–6.54) in the chemotherapy arm. Interestingly, the association between pCR and TMB was more pronounced in patients treated with chemotherapy alone. The KEYNOTE-119 trial evaluated metastatic TNBC patients treated with Pembrolizumab monotherapy versus chemotherapy. The positive association was observed between TMB and clinical response to Pembrolizumab (ORR
The proliferative MKI67+ T-cells (Tprf-MKI67) were exclusively present in TNBC patients achieving progressive disease after chemotherapy. Based on literature findings, the expression of MKI67 gene was significantly correlated with lymph node metastases, tumor invasion and adverse survival outcome in TNBC [29]. In addition, more unfavourable survival outcomes in breast cancer patients with recurrent lesions were significantly correlated
with high Ki-67 immunohistochemical expression levels (hazard ratio 2.307; 95% confidence interval 1.207–4.407, \( p \)-value = 0.011) \[30\]. Therefore, MKI67 may be an important biomarker of predictive and prognostic value in TNBC.

CD4-Tn-LEF1 and CD4-CXCL13 T-cells were linked to partial remission after anti-PD-L1+chemotherapy treatment. Importantly, these CD4+ T-cells express very high amounts of PD-1 and other co-stimulatory and inhibitory receptors. Therefore, they instrumental to B-cells for efficient antibody responses and their presence in tumor samples is often correlated with a better outcome in patients with solid tumors [31]. Based on biomedical literature, the presence of CD4-CXCL13 T-cells in TNBC tumors responsive to chemoimmunotherapy was detected through immunohistochemistry staining \[16,32\]. In addition to CXCL13+ T-cells, naïve LEF1+ T-cells (Tn-LEF1) were also linked to a favorable response to both anti-PD-L1+chemotherapy and chemotherapy. In a recent study, the magnitude of lymphocytic infiltration was assessed by a four-gene signature—HLF, CXCL13, SULT1E1 and GBP1, which was indicative of favourable outcome in TNBC after neoadjuvant therapy. This signature may help to identify early stage TNBC patients and being a novel prognostic biomarker of this aggressive disease [33].

The activated IFI6+ T-cells (Tact-IFI6) were linked to progressive disease after chemotherapy. The poor metastasis-free survival in breast cancer patients was linked to upregulation of mitochondrial antiapoptotic protein IFI6 that might be involved in regulation of mitochondrial ROS production [34]. Therefore, to improve clinical outcomes in breast cancer patients, the deactivation of mitochondrial functions of IFI6 is paramount.

3.3.2. B-Cells Global Cluster

The MKI67+ follicular B-cells (Bfoc-MKI67), NEIL1+ follicular B-cells (Bfoc-NEIL1) and MKI67+ memory B-cells (Bmem-MKI67) were exclusively present in TNBC patients with partial remission after anti-PD-L1+chemotherapy treatment. Based on biomedical literature, follicular B-cells was associated with favorable outcomes for TCGA patients with breast cancer [35]. The naïve B-cells, memory B-cells and follicular B-cells were present primarily in patients responsive to chemoimmunotherapy but not in patients responsive to chemotherapy treatment [16]. In regard with Bfoc-NEIL1 cell population, NEIL1 implicated in repair of oxidative damage associated with DNA replication or transcription [36]. Reduction in NEIL1 expression was associated with a poorer outcome in patients with breast invasive carcinoma [37]. Hence, NEIL1 could be a promising biomarker for TNBC patients that consider chemoimmunotherapy treatment.

Plasma IGHG1+ B-cells (pB-IGHG1) were linked to progressive disease after chemotherapy treatment. In TNBC, the expression of IGHG1 indicated the most significant prognostic value compared to trivial clinicopathological parameters [38]. Intriguingly, IGHG1 expression in B-cells and plasma cells could be associated with immune evasion and tumor cell proliferation in breast malignancies [39]. These data may imply that B cells or plasma cells could have pro-tumoral roles under particular conditions; however, the factors influencing the emergence of this pathologic phenotype and the roles played by B cells and plasma cells in these contexts remains unclear.

3.3.3. Myeloid Cells Global Cluster

The MMP9+ macrophages (macro-MMP9) were exclusively present in TNBC patients with partial remission after anti-PD-L1+chemotherapy treatment. The literature search revealed that MMPs have a intricate role in cancer progression and may exert both pro- and antitumorigenic activities [40]. Although MMP expression has been linked to tumor progression in various cancer types including breast cancer [41], clinical trials investigating the effect of broad-spectrum MMP inhibitors have failed, and in some cases, patients treated with these inhibitors even progressed after treatment comparing to control placebo group [42]. Indeed, the overexpression of MMP9 results in increased production of antiangiogenic fragments, decreased angiogenesis, and therapeutic effects of established breast cancer [43]. In another study, gene transfer of MMP-9 to ex vivo breast cancer
tumors caused tumor regression via increased neutrophil infiltration and an activation of tumor-associated macrophages (TAMs) [44]. Therefore, MMP9 can serve as a biomarker for predicting tumor regression in TNBC.

The macro-CCL2, macro-CX3CR1, macro-IFI27, macro-IGFBP7, macro-IL1B9, macro-MGP and macro-SLC40A1 cells were exclusively present in TNBC patients achieving progressive disease after chemotherapy. Based on biomedical findings, CCL2 expression in breast carcinomas was highly associated with macrophage infiltration, and its expression was correlated with poor prognosis in breast cancer patients [45]. In another study, chemokine receptor CX3CR1 showed a role in angiogenic macrophage survival in the tumor microenvironment contributing to tumor metastasis [46]. In a similar fashion, IFI27 overexpression was shown to impair the tamoxifen-induced apoptosis in breast cancer cells [47]. Finally, IL1B signalling contributed to breast cancer metastasis by enhancing tumor cell motility and inhibiting cell proliferation [48]. These findings highlight the importance of CCL2, CX3CR1, IFI27 and IL1B expressed in macrophages in progression of TNBC.

The CLEC9A⁺ dendritic cells (cDC1-CLEC9A), macro-CFD, macro-FOLR2, macro-MKI67, macro-SPP1, macro-TUBA1B, FCN1⁺ monocytes (mono-FCN1), mono-S100A89 and mono-SMIM25 cells were linked to progressive disease after chemotherapy treatment. Notably, CFD functioned as an enhancer of tumor proliferation and cancer stem cell properties in breast cancers [49]. In another study, SPP1-associated macrophages in the tumor-adipose microenvironment facilitate breast cancer progression [50]. Interestingly, S100A8/A9, which are calcium-binding proteins that are secreted primarily by granulocytes and monocytes, may be associated with the loss of estrogen receptor and may be involved in the poor prognosis of Her2⁺/basal-like subtypes of breast cancer. Therefore, myeloid cell populations expressing CFD, SPP1 and S100A89 might be crucial biomarkers of poor treatment response in TNBC.

3.3.4. NK-Cells Global Cluster

The CNOT2⁺ group 2 innate lymphoid cells (ILC2-CNOT2) were exclusively present in TNBC patients achieving partial remission after anti-PD-L1+chemotherapy treatment. Indeed, ILC2s involved in both anti-tumor and pro-tumoral immunity in a variety of human cancers [51]. In terms of pro-tumoral immunity, the promotion of tumor growth and metastasis is achieved by crosstalk between ILC2s and tumor microenvironment (TME) [52]. In addition, the ILC2s trigger the apoptosis of tumor cells by recruiting and activating eosinophils [53], CXCL1L/CXCL2L molecules and macrophages with M1 profile [54].

The ZNF683⁺ group 1 innate lymphoid cells (ILC1-ZNF683) cells were exclusively present in TNBC patients achieving progressive disease after chemotherapy. The biomedical literature demonstrates that ILC1 cells involved in inhibiting the antitumoral immune response, enabling the differential tumor infiltration of ILC1 cells in patients to improve the leveraging of immunity in cancer therapies [55]. However, the role of ZNF683 gene in particular remains elusive.

ILC3-AREG and ILC3-IL7R cells were linked to partial remission after anti-PD-L1+chemotherapy treatment. It had been shown that ILC3-IL7R could predict a favorable response to both treatment regimens, indicating its potential role in effective antitumor immunity [16]. In contrary, ILC1-VCAM1 cells were linked to progressive disease after chemotherapy treatment. Recent studies have shown that vascular cell adhesion molecule-1 (VCAM1) is aberrantly expressed in breast cancer cells and mediates prometastatic tumor-stromal interactions [56]. Therefore, AREG⁺, IL7R⁺ and VCAM1⁺ innate lymphoid cells can help determine prognosis for breast cancer patients.

4. Discussion

The analysis of the TNBC scRNA-seq data revealed distinct immune cell populations that are linked to either partial remission after anti-PD-L1+chemotherapy or progressive disease after chemotherapy only. In terms of T-cells, CD4-Tn-LEF1 and CD4-CXCL13 T-cells
were linked to partial remission after anti-PD-L1+chemotherapy treatment, while Tact-IFI6 T-cells were linked to progressive disease after chemotherapy. The naïve B-cells, memory B-cells and follicular B-cells were mainly enriched in tumors responsive to chemoimmunotherapy but not in tumors responsive to chemotherapy treatment. The MMP9+ macrophages (macro-MMP9) were exclusively present in TNBC patients with partial remission after anti-PD-L1+chemotherapy treatment, while heterogenous population of macrophages, including macro-CCL2, macro-CX3CR1, macro-IFI27, macro-IGFBP7, macro-IL1B9, macro-MGP and macro-SLC40A1 cells were exclusively present in TNBC patients achieving progressive disease after chemotherapy. Finally, group 3 innate lymphoid cells (ILC3-AREG and ILC3-IL7R) were linked to partial remission after anti-PD-L1+chemotherapy treatment, while ZNF683+ group 1 innate lymphoid cells (ILC1-ZNF683) cells were exclusively present in TNBC patients achieving progressive disease after chemotherapy. Each of these cell populations have distinctive genetic markers that could be useful therapeutic targets for chemoimmunotherapy.

The role of T follicular helper and B-cell crosstalk in tumor immunity has been extensively studied over the last decade. Accumulating evidence suggests that tumor infiltrated lymphocyte (TIL) subpopulations (CD4, CD8, and CD19/20) constitute of both suppressive (pro-tumor) or effector (anti-tumor) phenotypes whose functions are influenced by the surrounding TME [57]. Natural or treatment-induced immune activation or suppression may determine the balance between pro- or anti-tumor immune cell crosstalk within a given tumor. Key anti-tumor effector activities include antibody-dependent cell cytotoxicity, complement activation, antibody-mediated tumor cell phagocytosis, antigen presentation, T cell activation, cytokine secretion, and direct tumor killing by TIL, including CD8, NK, B cells, and/or macrophages [58].

Despite of significant survival advantages that could be achieved after treatment with chemoimmunotherapy, most TNBC patients would not benefit. Therefore, more and more attention has been paid to the identification and development of biomarkers for the response of chemoimmunotherapy in recent years. Our informatics pipeline identified novel phenotypic and genotypic predictors in unsupervised manner that indicative of favorable outcome after chemoimmunotherapy. These predictors could be important biomarkers in the design and analysis of intervention studies and ultimately could help to optimize therapy decisions for TNBC patients. In addition, it may help to select better responders to chemoimmunotherapy, as well as pinpoint the underlying mechanisms of drug resistance in TNBC patients at single-cell resolution.

5. Conclusions

To tackle patient heterogeneity, chemoimmunotherapy combinations represent a feasible alternative for TNBC patients. However, matching patient subgroups to effective treatments that increase their chance of survival remains a challenging endeavor. In this work, we augmented our exploratory subgroup discovery algorithm to identify TNBC subpopulations that may benefit from chemoimmunotherapy. Specifically, we identified immune-related gene signatures that increased the likelihood of partial remission after anti-PD-L1+chemotherapy regimen versus progressive disease after chemotherapy in TNBC patients. Our novel informatics pipeline identified immune cell populations that associated with various treatment outcomes in TNBC. We also showed the importance of TMB and menopausal status among the investigated TNBC subgroups. The potential limitations include the usage of two disjoint datasets and the absence of outcome variable for immunotherapy outcomes in TCGA datasets. Further validation of our computational results in wet-lab studies would be a significant step toward improving survival outcomes for TNBC patients.
Author Contributions: Conceptualization, O.K. and C.-R.S.; methodology, O.K.; software, W.I.B.; validation, O.K. and C.-R.S.; formal analysis, O.K., W.I.B. and C.-R.S.; data curation, O.K.; writing—original draft preparation, O.K. and C.-R.S.; writing—review and editing, C.-R.S., J.B.M., J.T.K., R.D.H. and C.N.P.; visualization, O.K.; supervision, C.-R.S. and J.B.M.; project administration, C.-R.S.; funding acquisition, C.-R.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research study was funded by University of Missouri Institute for Data Science and Informatics to O.K and W.I.B. and Data-Driven and Artificial Intelligence Initiatives to C.-R.S., O.K. and C.-R.S. were funded by Shumaker Endowment for Bioinformatics. J.B.M. received funding from the Department of Veteran’s Affairs (K2BX004346-01A1). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Veterans’ Affairs. The funding bodies had no role in the study design; in the collection, analysis, interpretation of data; or in the writing of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The TCGA dataset is available online at https://www.cbioportal.org (accessed on 22 August 2022). The scRNA-seq TNBC dataset is available online at http://tnbc_pd-l1.cancer-pku.cn (accessed on 22 August 2022).

Acknowledgments: We appreciate the critical input of members from Interdisciplinary Data Analytics and Search (iDAS) Laboratory at University of Missouri, Columbia.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Moss, J.L.; Tatalovich, Z.; Zhu, L.; Morgan, C.; Cronin, K.A. Triple-negative breast cancer incidence in the United States: Ecological correlations with area-level sociodemographics, healthcare, and health behaviors. Breast Cancer 2020, 28, 82–91. [CrossRef]
2. Lehmann, B.D.; Pietenpol, J.A. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. J. Pathol. 2014, 232, 142–150. [CrossRef]
3. Gupta, G.K.; Collier, A.L.; Lee, D.; Hoefer, R.A.; Zheleva, V.; Van Reesema, L.L.S.; Tang-Tan, A.M.; Guye, M.L.; Chang, D.Z.; Winston, J.S.; et al. Perspectives on Triple-Negative Breast Cancer: Current Treatment Strategies, Unmet Needs, and Potential Targets for Future Therapies. Cancers 2020, 12, 2392. [CrossRef]
4. Tarantino, P.; Corti, C.; Schmid, P.; Cortes, J.; Mittendorf, E.A.; Rugo, H.; Tolaney, S.M.; Bianchini, G.; Andrè, F.; Curigliano, G. Immunotherapy for early triple negative breast cancer: Research agenda for the next decade. NPJ Breast Cancer 2022, 8, 23. [CrossRef]
5. Tutt, A.N.; Garber, J.E.; Kaufman, B.; Viale, G.; Fumagalli, D.; Rastogi, P.; Gelber, R.D.; de Azambuja, E.; Fielding, A.; Balmaña, J.; et al. Adjuvant Olaparib for Patients with BRCA1- or BRCA2-Mutated Breast Cancer. N. Engl. J. Med. 2021, 384, 2394–2405. [CrossRef]
6. O’Meara, T.A.; Tolaney, S.M. Tumor mutational burden as a predictor of immunotherapy response in breast cancer. Oncotarget 2021, 12, 394–400. [CrossRef]
7. Cortes, J.; Cescon, D.W.; Rugo, H.S.; Nowecki, Z.; Im, S.-A.; Yusof, M.M.; Gallardo, C.; Lipatov, O.; Barrios, C.H.; Holgado, E.; et al. Pembrolizumab plus chemotherapy versus placebo plus chemotherapy for previously untreated locally recurrent inoperable or metastatic triple-negative breast cancer (KEYNOTE-355): A randomised, placebo-controlled, double-blind, phase 3 clinical trial. Lancet 2020, 396, 1817–1828. [CrossRef]
8. Hutchinson, K.E.; Yost, S.E.; Chang, C.-W.; Johnson, R.M.; Carr, A.R.; McAdam, P.R.; Halligan, D.L.; Chang, C.-C.; Schmolze, D.; Liang, J.; et al. Comprehensive Profiling of Poor-Risk Fined Primary and Recurrent Triple-Negative Breast Cancers Reveals Immune Phenotype Shifts. Clin. Cancer Res. 2020, 26, 657–668. [CrossRef]
9. Schmid, P.; Cortes, J.; Dent, R.; Fusi, L.; McArthur, H.; Kümmel, S.; Bergh, J.; Denkert, C.; Park, Y.H.; Hui, R.; et al. VP7-2021: KEYNOTE-522: Phase III study of neoadjuvant pembrolizumab + chemotherapy vs. placebo + chemotherapy, followed by adjuvant pembrolizumab vs. placebo for early-stage TNBC. Ann. Oncol. 2021, 32, 1198–1200. [CrossRef]
10. Criscitiello, C.; Corti, C.; Pravettoni, G.; Curigliano, G. Managing side effects of immune checkpoint inhibitors in breast cancer. Crit. Rev. Oncol. Hematol. 2021, 162, 103354. [CrossRef]
11. Miles, D.; Gilgorov, J.; André, F.; Cameron, D.; Schneeweiss, A.; Barrios, C.; Xu, B.; Wardley, A.; Kaen, D.; Andrade, L.; et al. Primary results from IMpassion131, a double-blind, placebo-controlled, randomised phase III trial of first-line pachitaxel with or without atezolizumab for unresectable locally advanced/metastatic triple-negative breast cancer. Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. 2021, 32, 994–1004. [CrossRef]
12. Davis, A.A.; Patel, V.G. The role of PD-L1 expression as a predictive biomarker: An analysis of all US Food and Drug Administration (FDA) approvals of immune checkpoint inhibitors. J. Immunother. Cancer 2019, 7, 278. [CrossRef]
13. Cristescu, R.; Aurora-Garg, D.; Albright, A.; Xu, L.; Liu, X.Q.; Loboda, A.; Lang, L.; Jin, F.; Rubin, E.H.; Snyder, A.; et al. Tumor mutational burden predicts the efficacy of pembrolizumab monotherapy: A pan-tumor retrospective analysis of participants with advanced solid tumors. J. Immunother. Cancer 2022, 10, e003091. [CrossRef]

14. Gradishar, W.J.; Anderson, B.O.; Abraham, J.; Aft, R.; Agnese, D.; Allison, K.H.; Blair, S.L.; Burstein, H.J.; Dang, C.; Elias, A.D.; et al. Breast Cancer: Version 3.2020, NCCN Clinical Practice Guidelines in Oncology. J. Natl. Compr. Cancer Netw. JNCCN 2020, 18, 452–478. [CrossRef]

15. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.E.; Sumner, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Genomics and Clinical Profiles Using the eBiologicalPortal. Sci. Signal. 2013, 6, pl1. [CrossRef]

16. Zhang, Y.; Chen, H.; Hu, X.; Gao, R.; Zhao, Y.; Liu, B.; Niu, L.; Sun, X.; Yu, X.; et al. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. Cancer Cell 2021, 39, 1578–1593.e8. [CrossRef]

17. Agrawal, R.; Srikan, R. Fast Algorithms for Mining Association Rules in Large Databases. In Proceedings of the 20th International Conference on Very Large Data Bases, VLDB ’94, Santiago de Chile, Chile, 12–15 September 1994; Morgan Kaufmann Publishers Inc.: Burlington, MA, USA, 1994; pp. 487–499.

18. Dong, G.; Li, J. Efficient mining of emerging patterns: Discovering trends and differences. In Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’99, San Diego, CA, USA, 15–18 August 1999; Association for Computing Machinery: New York, NY, USA, 1999; pp. 43–52. [CrossRef]

19. Egghe, L. Theory and practise of the g-index. Scientometrics 2006, 69, 131–152. [CrossRef]

20. Liu, D.; Baskett, W.; Beversdorf, D.; Shyu, C.-R. Exploratory Data Mining for Subgroup Cohort Discoveries and Prioritization. IEEE J. Biomed. Health Inform. 2020, 24, 1456–1468. [CrossRef]

21. Baumjohann, D.; Brossart, P. T follicular helper cells: Linking cancer immunotherapy and immune-related adverse events. J. Immunother. Cancer 2021, 9, e002588. [CrossRef]

22. Ma, Q.; Chen, Y.; Qin, Q.; Guo, F.; Wang, Y.-S.; Li, D. CXCL13 expression in mouse 4T1 breast cancer microenvironment elicits antitumor immune response by regulating immune cell infiltration. Precis. Clin. Med. 2021, 4, 155–167. [CrossRef]

23. Criscitiello, C.; Bayar, M.; Curigliano, G.; Symmans, F.; Desmedt, C.; Bonnefoi, H.; Sinn, B.; Pruneri, G.; Vicier, C.; Pierga, J.; et al. A gene signature to predict high tumor-infiltrating lymphocytes after neoadjuvant chemotherapy and outcome in patients with triple-negative breast cancer. Ann. Oncol. 2018, 29, 162–169. [CrossRef]

24. Cheryiath, V.; Kaur, J.; Davenport, A.; Khaled, A.; Chowdhury, N.; Gaddipati, L. GIP3 (IF16), a mitochondrial localized antiapoptotic protein, promotes metastatic potential of breast cancer cells through mtROS. Br. J. Cancer 2018, 119, 52–64. [CrossRef]

25. Hollern, D.P.; Xu, N.; Thennavan, A.; Godowski, C.; Garcia-Recio, S.; Mott, K.R.; He, X.; Garay, J.P.; Carey-Evend, K.; Marron, D.; et al. B Cells and T Follicular Helper Cells Mediate Response to Checkpoint Inhibitors in Higher Mutation Burden Mouse Models of Breast Cancer. Cell 2019, 179, 1191–1206.e21. [CrossRef]

26. Mori, H.; Ouchida, R.; Hijiikata, A.; Kitamura, H.; Ohara, O.; Li, Y.; Gao, X.; Yasui, A.; Lloyd, R.S.; Wang, J.-Y. Deficiency of the oxidative damage-specific DNA glycosylase NEIL1 leads to reduced germinal center B cell expansion. DNA Repair. 2009, 8, 1328–1332. [CrossRef]
