Immune repertoire and responses to neoadjuvant TCHP therapy in HER2-positive breast cancer

Ther Adv Med Oncol

2023, Vol. 15: 1–19

DOI: 10.1177/17588359231157654

Junyoung Shin, Baknoon Ham, Jeong-Han Seo, Sae Byul Lee, In Ah Park, Gyungyub Gong, Sung-Bae Kim and Hee Jin Lee

Abstract

Background: Despite the introduction of trastuzumab, pathologic complete response (pCR) is not attained in approximately 30–40% of Human epithelial growth factor receptor-2-positive breast cancer. Tumor-infiltrating lymphocytes (TIL) have been suggested as a predictive marker of treatment response, albeit not always effective. We investigated the relationship between trastuzumab, docetaxel, carboplatin, and pertuzumab (TCHP) treatment and immune repertoire as a treatment response predictor.

Design: In all, 35 cases were divided into two experimental groups: 10 and 25 cases in the preliminary and main experiments, respectively. In the preliminary experiment, the biopsy tissues before TCHP treatment and the surgical tissues after TCHP treatment were compared. In the main experiment, the biopsy tissues before TCHP treatment were compared according to the TCHP treatment response.

Methods: The T-cell repertoire for TRA, TRB, TRG, and TRD, and B-cell repertoire for immunoglobulin heavy, immunoglobulin kappa, and immunoglobulin lambda were evaluated. Whole transcriptome sequencing was also performed.

Results: In the preliminary experiment, the density and richness of the T-cell receptor (TCR) and B-cell receptor (BCR) repertoires decreased after treatment, regardless of TCHP response. In the main experiment, the Shannon's entropy index, density, and length of CDR3 of the TCR and BCR repertoires did not differ significantly in patients who did and did not achieve pCR. The pCR and non-pCR subgroups according to the level of TILs revealed that the non-pCR/lowTIL group had a higher proportion of low-frequency clones than the pCR/lowTIL group in TRA (non-pCR/lowTIL *versus* pCR/lowTIL, 0.01–0.1%, 63% *versus* 45.3%; <0.01%, 32.9% *versus* 51.8%, $p < 0.001$) and TRB (non-pCR/lowTIL *versus* pCR/lowTIL, 0.01–0.1%, 26.5% *versus* 14.7%; <0.01%, 72.0% *versus* 84.1%, $p < 0.001$).

Conclusions: The role of the diversity, richness, and density of the TCR and BCR repertoires as predictive markers for TCHP response was not identified. Compositions of low-frequency clones could be candidates for predictive factors of TCHP response; however, validation studies and further research are necessary.

Keywords: HER2-positive breast cancer, immune repertoire, neoadjuvant TCHP treatment, TCHP response prediction

Received: 17 August 2022; revised manuscript accepted: 30 January 2023.

Correspondence to:
Hee Jin Lee
Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro, 43-gil, Songpa-gu, Seoul 05505, Republic of Korea.

NeogenTC Corp., Seoul, Republic of Korea.
backlila@gmail.com

Sung-Bae Kim
Department of Oncology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro, 43-gil, Songpa-gu, Seoul 05505, Republic of Korea.
sbkim3@amc.seoul.kr

Junyoung Shin
Gyungyub Gong
Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

Baknoon Ham
Jeong-Han Seo
NeogenTC Corp., Seoul, Republic of Korea

Sae Byul Lee
Department of Breast Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

In Ah Park
Department of Pathology, Kangbuk Samsung Hospital, Seoul, Republic of Korea

Introduction

Human epithelial growth factor receptor-2 (HER2) is overexpressed in 15–30% of invasive breast cancers, with these tumors being more aggressive and having reduced overall survival (OS) than tumors not overexpressing HER2.[1,2] The inclusion of trastuzumab in combination treatment has improved pathologic complete

response (pCR) and 10-year OS rates in patients with early-stage, locally advanced HER2-positive breast cancer.[3] Although the combination of trastuzumab, docetaxel, carboplatin, and pertuzumab (TCHP) has become a favored neoadjuvant regimen, residual tumors remain in 30–40% of patients with HER2-positive breast cancer following neoadjuvant TCHP treatment.[1]

Several studies have endeavored to investigate predictors of response to trastuzumab-based chemotherapy in HER2-positive breast cancer, and the level of tumor-infiltrating lymphocytes (TILs) has been predictive of the benefit of trastuzumab in retrospective and prospective trials.[4–6] Accordingly, there have been steady efforts to subdivide and describe TILs in detail.

The T-cell receptor (TCR) and B-cell receptor (BCR), which are the antigen-recognizing parts of the T cell and B cell, respectively, are highly diverse proteins, potentially reaching 10^{20}–10^{61} in the TCR and 10^{18} in the BCR.[7,8] This high diversity is caused by the molecular structures of the TCR and BCR, which are composed of complexes of several amino acid (AA) chains and the V(D)J recombination process of each AA chain. The TCR is a heterodimer composed of an alpha and a beta chain consisting of TCR alpha (TRA) and TCR beta (TRB) in the case of alpha and beta T cells, and a gamma and delta chain consisting of TCR gamma (TRG) and TCR delta (TRD) in the case of gamma and delta T cells.[9] The BCR has a Y-shaped structure composed of a heavy chain produced by the immunoglobulin heavy (IGH) chain recombination in the pre-B cell and two light chains produced by immunoglobulin kappa (IGK) or immunoglobulin lambda (IGL) in the immature B cell.[10]

The recent development of high-throughput next-generation sequencing (NGS) has enabled studies on TCR and BCR repertoires, with immune repertoires reported to have a significant influence on the prognosis and pathophysiology of several diseases.[11–13] This study utilized high-throughput NGS to analyze the wide range of the TCR (TRA, TRB, TRG, and TRD) and BCR (IGH, IGL, and IGK) repertoires in patients with HER2-positive tumors treated with neoadjuvant TCHP, to determine the effects of immune repertoires on responses to TCHP treatment. Other factors that can help predict treatment responses were also analyzed.

Materials and methods

Study participants

Between 2017 and 2020, 35 patients with HER2 overexpressed breast cancer were recruited from the Asan Medical Center, Seoul, Republic of Korea. All patients underwent a core-needle biopsy before initiating neoadjuvant chemotherapy with TCHP. HER2 overexpression was determined in core-needle biopsy samples according to the HER2 testing guidelines.[14] After neoadjuvant treatment, all patients underwent surgical treatment, and disease response (pCR or non-pCR) was pathologically evaluated by examination of the surgical specimens.

The 35 patients were divided into two groups. The preliminary experiment group included 10 patients, 6 with a high level of TILs (highTIL $\geqslant 10\%$) who achieved pCR and 4 with a low level of TILs (lowTIL $< 10\%$) who did not achieve pCR (non-pCR). The main experiment included 25 consecutively enrolled patients, 13 who achieved pCR and 12 who did not (non-pCR) (Figure 1). The preliminary experiment assessed the effects of TCHP on TCR and BCR repertoires by comparing the pre-TCHP biopsy with the surgical samples. The main experiment analyzed the pre-TCHP biopsy samples alone, comparing the immune repertoire in the patients who achieved pCR and in patients with residual tumors. To minimize the effect of TILs on treatment response, patients in the pCR and non-pCR groups were each subdivided into those with high ($\geqslant 10\%$) and low ($< 10\%$) TILs, and their immune responses were compared separately (Figure 2).[15,16] The highTIL group included seven patients who achieved pCR (pCR/highTIL) and three who did not (non-pCR/highTIL), whereas the lowTIL group included six patients who achieved pCR (pCR/lowTIL) and seven who did not (non-pCR/lowTIL).

Methods

Immunohistochemistry and silver in situ *hybridization.* Formalin-fixed paraffin-embedded (FFPE) tissue sections, 4 μm thick, of core-needle biopsy samples from all patients, and the surgical specimens of the non-pCR group were subjected to immunohistochemical (IHC) staining for estrogen receptor (6F11, 1:200, Novo Castra, Newcastle, UK), progesterone receptor (PGR312, 1:200, Novo Castra), and HER2 (4B5, 1:8,

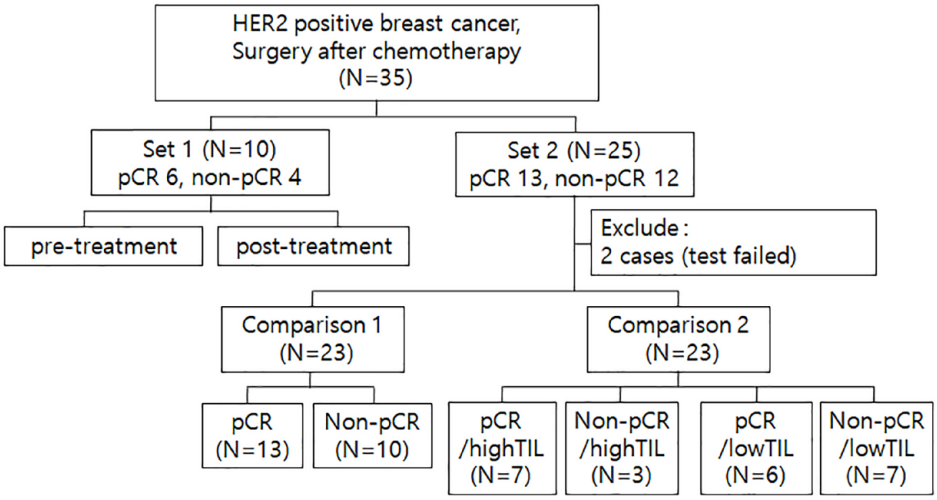

Figure 1. Schema of the study population.

Ventana, Tucson, AZ, USA). IHC labeling was performed using an autostainer (Benchmark XT; Ventana Medical Systems, AZ, USA) according to the manufacturer's protocol. Two pathologists reviewed these immune-labeled slides.

FFPE tissue sections, 4 μm thick, of samples scored as 2+ for HER2 were subjected to automated silver *in situ* hybridization (SISH) with INFORM HER2 DNA and chromosome 17 probes (Ventana Medical Systems, AZ, USA), using an ultraView SISH Detection Kit (Ventana Medical Systems, AZ, USA) according to the manufacturer's instructions.

Pathologic evaluation. All specimens were pathologically reviewed. Evaluation of the core-needle biopsy samples included assessments of stromal TILs and the nuclear and histologic grades of each breast carcinoma sample. The levels of TILs are expressed as percentages, as described by the International TIL Working Group.[17] The residual cancer burden (RCB) score, Miller-Payne grade, residual tumor size, and TILs of the surgical specimens were determined.[18,19] The tumor bed area in samples of patients who achieved pCR group was marked and collected for high-throughput NGS.

High-throughput sequencing for the immune repertoire. A microtome was used to obtain 5 μm thick specimens of selected areas of FFPE tissue from the 10 patients included in the preliminary experiment. RNA was extracted from these specimens using AllpreRNA FFPE kits (Qiagen, Hilden, Germany) according to the manufacturer's instructions. RNA was extracted from the 25 samples in the main experiment with Maxwell RSC RNA FFPE kits and a Maxwell RSC device (Promega, Madison, WI, USA), according to the manufacturer's instructions. All RNA samples were quantified using QuantiTdsDNA fluorescence assay kits (Invitrogen, Waltham, MA, USA).

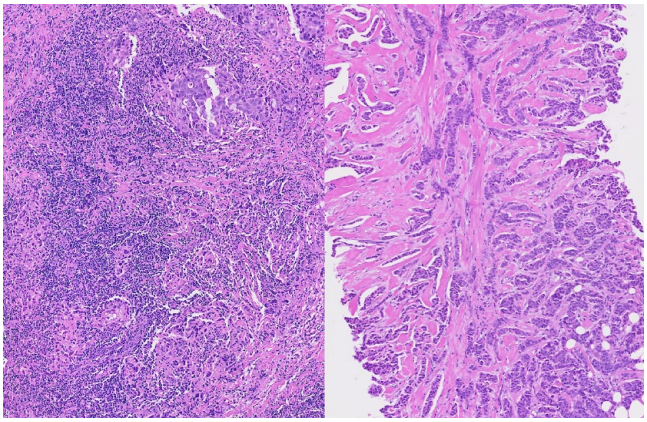

Figure 2. An example of a high TIL and low TIL in a tumor from the core-needle biopsy obtained before pre-neoadjuvant treatment. (Left) TILs were evaluated as 80% and classified as highTIL (H&E, ×10). (Right) Inflammatory cells are rarely observed, thus classifying this as lowTIL (H&E, ×10).
H&E, hematoxylin and eosin; TILs, tumor-infiltrating lymphocytes.

Libraries were constructed using 100 μg aliquots of RNA extracted from the 10 specimens included in the preliminary experiment, and 400 μg

aliquots of RNA extracted from the 25 specimens included in the main experiment of RNA was used as input of library generation with the immunoverse™-HS TCR alpha/delta/beta/gamma and BCR IGH/K/L kits (Archer, Zaragoza, Spain) according to the manufacturer's instructions. The intended concentrations of pooled libraries were confirmed using Collibri library quantification kits (Thermo Fisher Scientific, Waltham, MA, USA). Libraries were sequenced on an Illumina NextSeq 500 using NextSeq 500v2 reagents (Illumina, San Diego, CA, USA) for paired end, 150 base pair (bp), and dual index reads. Libraries were multiplexed, and the concentrations of sample libraries were calculated from a standard curve.

The Illumina sequencer was used to generate raw images utilizing sequencing control software for system control and base calling through an integrated primary analysis software program called Real-Time Analysis. The base calls binary was converted into FASTQ utilizing illumine package bcl2fastq. Adapters were not trimmed away from the reads.

Data were analyzed by Archer Analysis Immunoverse version 6.0. (ArcherDX, https://analysis.archerdx.com) according to the manufacturer's instructions. Briefly, Archer Analysis utilizes the AMP-specific molecular barcode adaptors for deduplication and PCR sequencing error correction. Deduplicated and error-corrected reads are used as inputs to MiXCR for comprehensive adaptive immunity profiling. T cell density, clonality, clonal fraction, CDR3 sequences, AAs, *V*, *D*, and *J* segments were obtained from the analyzer website. Only productive TCRs and BCRs were analyzed.

Whole transcriptome analysis. Libraries were prepared for 151 bp-paired end sequencing using the TrueSeq stranded mRNA Sample Preparation Kit (Illumina, San Diego, CA, USA), according to the manufacturer's instructions. Briefly, mRNA molecules were purified from 1 μg aliquots of total RNA and fragmented using oligo magnetic beads. After sequential end repair, A-tailing, and adapter ligation, cDNA libraries were amplified by PCR, and their quality was evaluated with an Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's protocol. The libraries were quantified using KAPA library quantification kits (Kapa Biosystems, Waltham, MA, USA) as described by the manufacturer. Following cluster amplification of denatured templates, the libraries were subjected to paired end (2×151 bp) sequencing using Illumina NovaSeq6000 (Illumina, San Diego, CA, USA). The adapter sequences and the ends of the reads with a Phred quality score < 20 were trimmed and reads shorter than 50 bp were simultaneously removed, using cutadapt v.2.8. The filtered reads were mapped to the reference genome using the aligner STAR v.2.7.1a, following ENCODE standard options with the 'quantMode TranscriptomeSMA' option to estimate transcriptome expression level. Gene expression estimation was done using RSEM v.1.3.1. Genes with at least 10 read counts were selected, and genes expressed to a lower extent were trimmed. To normalize sequencing depth among samples, DESeq2 package v.1.22.2 and FPKM were used. After comparing the gene expression level between the pCR and non-pCR groups, genes with thresholds of log2 fold change $\geqslant 1.5$ and p value < 0.05 were selected. A hierarchical clustering heatmap was constructed with these selected genes using the heatmap2 package in R. The filtered DEGs were sorted in the order of p value; hallmark pathway analysis using DAVID Bioinformatics Resources 6.8. A q value < 0.05 was considered to represent statistical significance.

Data from published papers. Databases of TCRdb,[20] TCR3d,[21] and vdjdb[22] were used to identify HER2 targeting sequences (*TRAV8-6*, 'CAVSVNTDKLIF' & *TRBV2-5*/*TRBJ2-5*, 'CASPPLAGDETQYF') (downloaded in 2019), and clustered TRB sequences showing significantly higher frequency in the pCR than in the non-pCR group were searched. The HER2 targeting sequences were compared with the current study results, and sequences with substitutions of one or two AAs were considered similar.[23,24]

Cell-type identification by estimating relative subtypes of RNA transcripts (CIBERSORT). RNA expression data were normalized with RMA and uploaded to the cell-type identification by estimating relative subtypes (CIBERSORT) web portal (https://cibersort.stanford.edu/) as a mixture file. The process of data manipulation, uploading, and analysis followed the instruction published on the homepage and article with the following options: relative and absolute modes together, LM22 signature gene file, 100 per mutation, and quantile normalization disabled. After the processing, the results were then downloaded.[25,26]

PAM50 subtypes and gene expression data. The levels of expression of the genes of prediction analysis of microarray 50 (PAM50) were normalized and standardized relative to five housekeepers, as described.[27] The samples were classified using the published PAM50 algorithm into breast cancer subtypes: luminal A, luminal B, basal-like, HER2-enriched, normal-like, and not applicable (NA).

Statistical analysis. The richness, density, Shannon's entropy index (SDI), number of AAs, and composition of each immune repertoire were evaluated. Richness was defined as the total number of separate clones with the same V, D, and J usage and CDR3 sequences. Density was defined as the sum of each clonal count. The SDI was calculated using the equation $H' = -\sum_{i=1}^{R} p_i \ln p_i$, where H' is the SDI, R is the number of separate clones, and p is the proportion of each clone. The AA length was determined by counting the number of AAs forming the CDR3 of each TCR/BCR. Composition was defined as the number of groups divided by the proportion of each clone occupying its entire clone count.

Groups were compared by Mann–Whitney U-tests, chi-square tests, or Fisher's exact tests, as applicable. Pre- and post-TCHP results in the preliminary experiment were compared using pairwise t-tests. The relationships between RCB scores and immune repertoires were determined by calculating Pearson correlation coefficients. In the main experiment, sequences with a significantly high frequency in the pCR group were selected, and principal component analysis (PCA) was performed by determining the relative Levenshtein distance. False discovery rate (FDR) correction was applied to correct for type 1 errors due to multiple comparisons. For all analyses, a p value or q value of <0.05 was considered to represent statistical significance. All data were analyzed using R Statistical Software (version 3.6.1).

Data availability. The data generated in this study are available within the article and its Supplemental Data files.

Results

Clinicopathologic characteristics of the patient population

Supplemental Table 1 summarizes the clinicopathologic characteristics of the 35 patients included in the preliminary and main experiments. A comparison of the 10 patients included in the preliminary experiment and the 25 included in the main experiment showed that the mean TIL was significantly higher (13% *versus* 7%, $p=0.008$) and the mean RCB score was significantly lower (0.471 *versus* 1.020, $p=0.045$) in the preliminary population. None of the other characteristics differed significantly.

Effect of TCHP treatment on immune repertoires

The 10 cases of the preliminary experiment were used to compare T-cell and B-cell repertoires before and after TCHP treatment. The immune repertoire characteristics of both groups are summarized in Supplemental Table 2. In TRA and TRB, most of the SDI, density, and richness were significantly decreased after treatment (the SDI of TRA, $p=0.037$ and TRB, $p=0.037$; density of TRA, $p=0.032$, and TRB, $p=0.022$; richness of TRA, $p=0.041$), but only the decrease in the richness of TRB was not statistically significant ($p=0.053$). In the TCR, the pre- and post-TCHP treatment lengths of CDR3 were similar. The clones occupying large proportions (1–10% and 0.1–1%) of both TRA and TRB were significantly increased after TCHP treatment ($p<0.001$) (Figure 3) (Supplemental Table 2). In the BCR, the SDI of IGH, IGL, and IGK were increased, while the density of IGH was decreased after treatment (the SDI of IGH, $p=0.014$, IGL, $p=0.024$, and IGK, $p=0.006$; density of IGH, $p=0.006$). However, the richness of IGH, IGL, and IGK, and density of IGL and IGK showed no significant change. In the aspect of the proportion of each class of immunoglobulin, IGHA significantly decreased after TCHP treatment ($p=0.004$). In the IGHG subclass, *IGHG2* was increased after TCHP treatment (pre-TCHP 20.4% and post-TCHP 27.9%; $p=0.023$), and *IGHG4* was decreased (pre-TCHP, 2.2% and post-TCHP 1.0%, $p=0.023$). The length of CDR3 of the light chains (IGL and IGK) was decreased significantly after TCHP treatment (IGL, $p=0.024$, and IGK, $p=0.006$); however, there was no difference in the heavy chain (IGH). After TCHP treatment, the composition of 1–10% and 0.1–1% clones was consistently significantly increased in IGH, IGK, and IGL ($p<0.001$) (Figure 3) (Supplemental Table 2). The pCR group and the non-pCR group showed similar changes in the SDI, density, and richness of TCR before and after TCHP treatment. In the

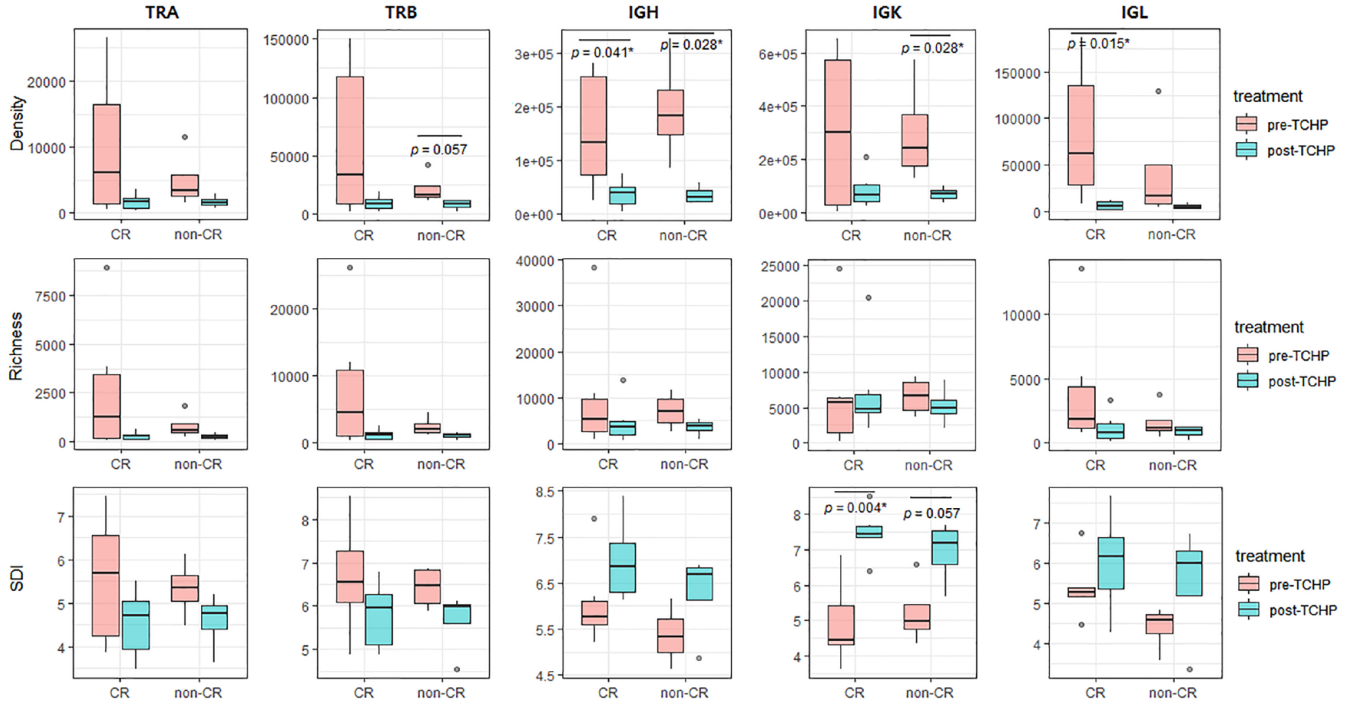

Figure 3. Boxplot of the immune repertoire according to the TCHP response. Pre- and post-TCHP immune repertoires in the 10 patients in the preliminary experiment who did and did not achieve pCR in response to TCHP. Regardless of the response, the density and richness of the TCR and BCR were decreased after treatment. The SDI tended to decrease in the TCR and increase in the BCR, with IGK showing a statistically significant increase.
*p value < 0.05.
BCR, B-cell receptor; IGK, immunoglobulin kappa; pCR, pathologic complete response; SDI, shannon's entropy index; TCHP, trastuzumab, docetaxel, carboplatin, and pertuzumab; TCR, T-cell receptor.

pCR group, there were significant differences before and after TCHP treatment in the density of IGH and IGL, and in the SDI of IGK (pCR group: density of IGH, $p = 0.041$; density of IGL, $p = 0.015$, and SDI of IGK, $p = 0.004$). In the non-pCR group, there were significant differences in the density of IGH and IGK, and the SDI of IGK before and after TCHP treatment (non-pCR group: density of TRB, $p = 0.057$; density of IGH, $p = 0.028$; density of IGK, $p = 0.028$; SDI of IGK, $p = 0.057$) (Figure 3) (Supplemental Table 2). In pre-TCHP, the richness of TRD and TRG was less than 10 in 40% of the cases, and in post-TCHP, 70%. These results were not statistically comparable because it was difficult to interpret whether the sequencing was successful. To supplement these results, the input RNA was increased to 400 μg in the main experiment.

Effect of response to TCHP on the immune repertoires

The main experiment included the pre-TCHP biopsy samples of 25 patients; however, two samples were excluded because amplification did not occur during the sequencing process. The remaining 23 patients included 13 who achieved pCR and 10 who did not (non-pCR); their clinicopathologic characteristics are summarized in Supplemental Table 3. None of these characteristics differed significantly in these two groups.

Evaluation of TCR showed that the SDI, density, richness, and length of CDR3 of TRA, TRB, TRD, and TRG did not differ significantly in the pCR and non-pCR groups [FIGURE4Figure 4(a)]. Although the clonal composition differed significantly between these groups ($p < 0.001$), the proportion of *TRGV9* and *TRDV2*, which represent Vγ9Vδ2 T cells, did not (pCR *versus* non-pCR, *TRGV9*, 22.5% *versus* 25.2%, $p = 0.614$; *TRDV2*, 37.3% *versus* 46.5%, $p = 0.107$) (Supplemental Table 4). None of the characteristics of the TRA and TRB repertoires correlated significantly with RCB scores [Figure 4(b)].

Evaluation of the BCR showed that the SDI, density, richness, and length of CDR3 of IGH, IGK,

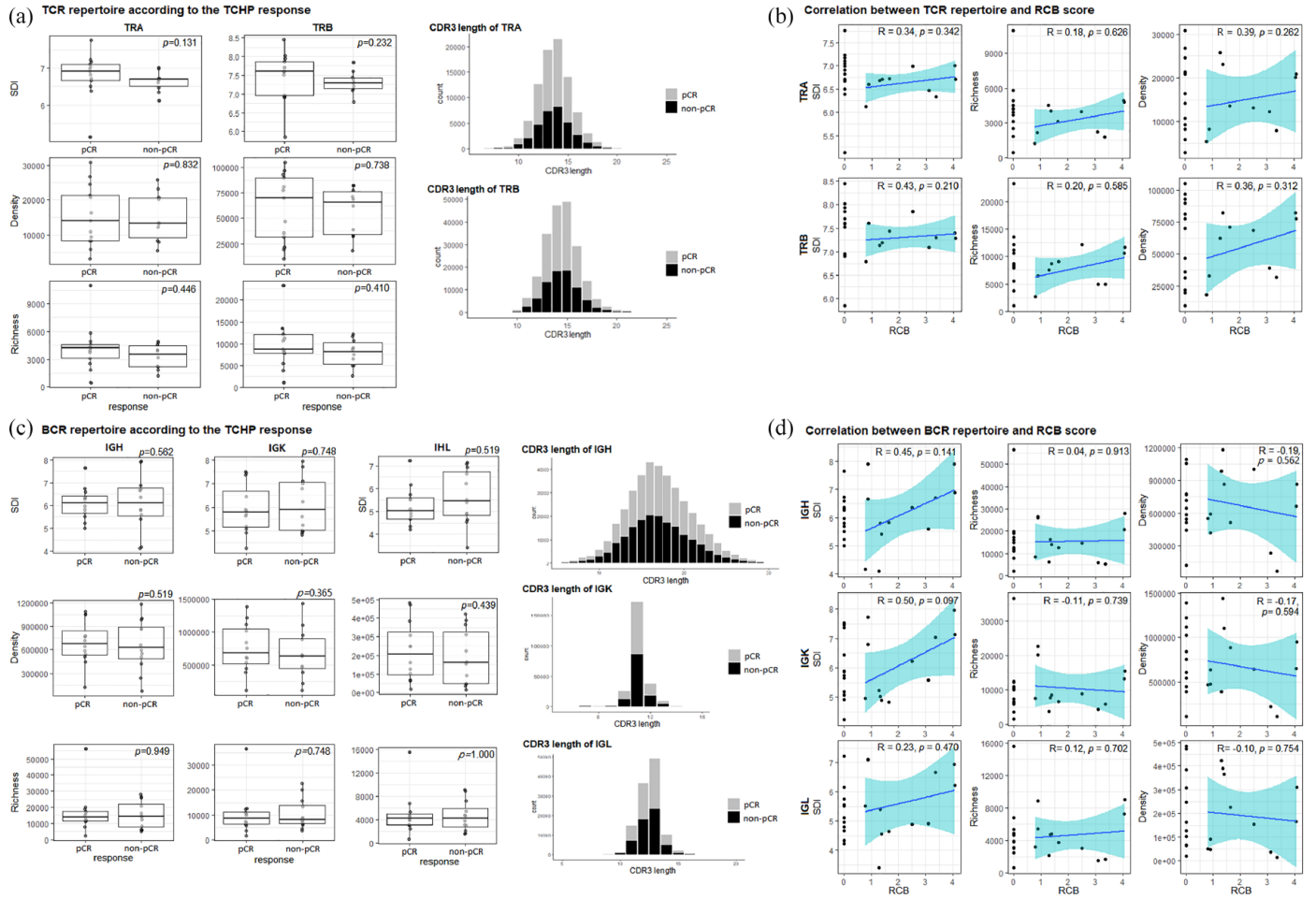

Figure 4. Immune repertoires in the 25 patients in the main experiment. (a) Boxplot of TCR repertoires according to the TCHP response (left) and bar plot of CDR3 length (right). The SDI, density, richness, and length of CDR3 of both TRA and TRB did not differ significantly in the pCR and non-pCR groups. (b) Scatter plot with correlation coefficient plot of the TCR repertoire and RCB score. The blue area represents a 95% confidence interval. None of the repertoires was significantly associated with RCB scores. (c) Boxplot of BCR repertoires according to the TCHP response (left) and bar plot of the CDR3 length (right). The SDI, density, richness, and length of CDR3 of IGH, IGK, and IGL did not differ significantly in the pCR and non-pCR groups. (d) Scatter plot with correlation coefficient plot of the BCR repertoire and RCB score. The blue area represents a 95% confidence interval. None of the BCR repertoires correlated with RCB scores.

BCR, B-cell receptor; IGK, immunoglobulin kappa; IGL, immunoglobulin lambda; pCR, pathologic complete response; RCB, residual cancer burden; SDI, shannon's entropy index; TCHP, trastuzumab, docetaxel, carboplatin, and pertuzumab; TCR, T-cell receptor.

and IGL did not differ significantly between the pCR and non-pCR groups. IgG was more abundant in the pCR (71.2%) than in the non-pCR (63.3%) group, albeit not statistically significant ($p = 0.052$). IgG subtypes showed similar distributions in both groups. The compositions of IGH, IGK, and IGL in the pCR and non-pCR groups differed significantly ($p < 0.001$). In all, 63 clones of hypermutated IGH were identified, with a mean of 2.5 clones per patient in the pCR group and 2.8 clones per patient in the non-pCR group. Conversely, the mean density of hypermutated IGH clones per patient was 87,794 in the pCR group and 101,330 in the non-pCR group, although the density, richness, and frequency of hypermutated IGH did not differ significantly in the pCR and non-pCR groups [Figure 4(c)] (Supplemental Table 4). None of the characteristics of the BCR repertoires correlated with RCB scores [Figure 4(d)].

Effect of response to TCHP on V, D, and J usage of TRA and TRB

The usage of *V*, *D*, and *J* genes of TRA and TRB was compared in patients who did and did not

achieve pCR in response to TCHP treatment. The most commonly used *TRAV* segment was *TRAV12-2*, followed by *TRAV13-1* and *TRAV9-2*. Only *TRAV34* differed in the pCR and non-pCR groups, albeit not statistically significant after correction for type 1 error ($p = 0.022$, $q = 0.735$). The most common TRAJ segment in the pCR group was *TRAJ23* followed by *TRAJ39*, whereas the most common TRAJ segments in the non-pCR group were *TRAJ43* and *TRAJ20*. None of the TRAJ segments differed significantly in the pCR and non-pCR groups after type 1 error correction [Figure 5(a)] (Supplemental Table 5). The most common TRB segments were *TRBV20-1* and *TRBV28* in the pCR group, and *TRBV18* and *TRBV6-1* in the non-pCR group. *TRBD1* was more common than *TRBD2* in both groups. *TRBJ2-1* was the most frequent segment in both groups, followed by *TRBJ2-7* and *TRBJ1-1*, with no significant differences between these two groups [Figure 5(b)] (Supplemental Table 5).

For TRA, the most common *V-J* gene pairs in the pCR group were *TRAV13-1*/*TRJ13*, *TRAV12-2*/*TRAJ3*, and *TRAV9-2*/*TRAJ54* in that order (mean of frequency, *TRAV13-1*/*TRAJ13*, 0.4%; *TRAV12-2*/*TRAJ3*, 0.4%; and *TRAV9-2*/*TRAJ54* 0.3%). In the non-pCR group, *TRAV4*/*TRAJ34* was the most frequent *V-J* gene pair, followed by *TRAV2*/*TRAJ10* and *TRAV1-2*/*TRAJ20* (mean of frequency, *TRAV4*/*TRAJ34*, 0.5%; *TRAV2*/*TRAJ10*, 0.5%; and *TRAV1-2*/*TRAJ20*, 0.4%). In all, 63 V-J gene pairs of TRA were differently expressed ($p < 0.05$) in the pCR and non-pCR groups, but none had a statistically significant q value. Sixteen *V-J* gene pairs of TRB differed significantly in the pCR and non-pCR groups, but none differed significantly after type 1 error correction (Supplemental Table 6).

Specific gene-targeting TCR sequences and PCA analysis of TCR sequences

A search of TCRdb for sequences targeting HER2 (Supplemental Table 7) identified one TRA and three TRB sequences, with all of these similar sequences found in the pCR group.

Evaluation of the samples in the main experiment identified 6738 TRA and 13,568 TRB sequences that were significantly ($p < 0.05$) more frequent in the pCR group. PCA analysis was performed by calculating relative Levenshtein distances. No clusters were observed in TRA [Figure 6(a)], whereas in TRB, a cluster was observed around a TCR sequence having the CDR3 AA sequence of 'CASSLLGSYNEQFF' with *TRBV5-1* and *TRBJ2-1* [Figure 6(b)]. A search of the 527 TRB sequences having 'CASSLLGSYNEQFF' in TCRdb showed that 3.6% had *TRAV5-1*/*TRBJ2-1*. These TRB sequences have been reported in various diseases, including type 1 diabetes (14.9%), non-small-cell lung cancer (NSCLC; 14.0%), and breast cancer (8.3%) [Figure 6(c)].

Characteristics of the immune repertoire according to TILs

The level of TILs was reported to be significantly associated with pCR in patients with HER2 positive breast cancer.[4,28] To exclude the effect of TILs, patients in both the pCR and non-pCR groups were subdivided into those with high and low levels of TILs. The clinicopathologic characteristics of these four groups, the pCR/highTIL, non-pCR/highTIL, pCR/lowTIL, and non-pCR/lowTIL groups, are summarized in Supplemental Table 8. None of these characteristics differed significantly in the pCR/highTIL and non-pCR/highTIL groups. Although most characteristics were similar in the pCR/lowTIL and non-pCR/lowTIL groups, lymph node metastasis was significantly more frequent in the non-pCR/lowTIL group ($p < 0.001$) (Supplemental Table 8).

Evaluation of TCRs showed no differences in the mean SDI, density, and richness of TRA, TRB, TRD, and TRG. The composition of the clones differed significantly in the pCR/highTIL and non-pCR/highTIL groups and in the pCR/lowTIL and non-pCR/lowTIL groups ($p < 0.001$, each) [Figure 7(a)]. The proportions of *TRGV9* and *TRDV2*, which represent Vγ9Vδ2 T cells, did not differ significantly in the pCR and non-pCR groups (Table 1).

Evaluation of BCRs showed no significant differences in the SDI, density, and richness of IGH, IGK, and IGL, and the isotypes of IGH among the four groups. The compositions of IGH, IGL, and IGK, however, showed statistically significant differences ($p < 0.001$) [Figure 7(b)]. The density, richness, and frequency of hypermutated clones did not differ significantly in comparisons of the pCR/highTIL and non-pCR/highTIL groups and the pCR/lowTIL and non-pCR/lowTIL groups [Figure 7(c), Table 1].

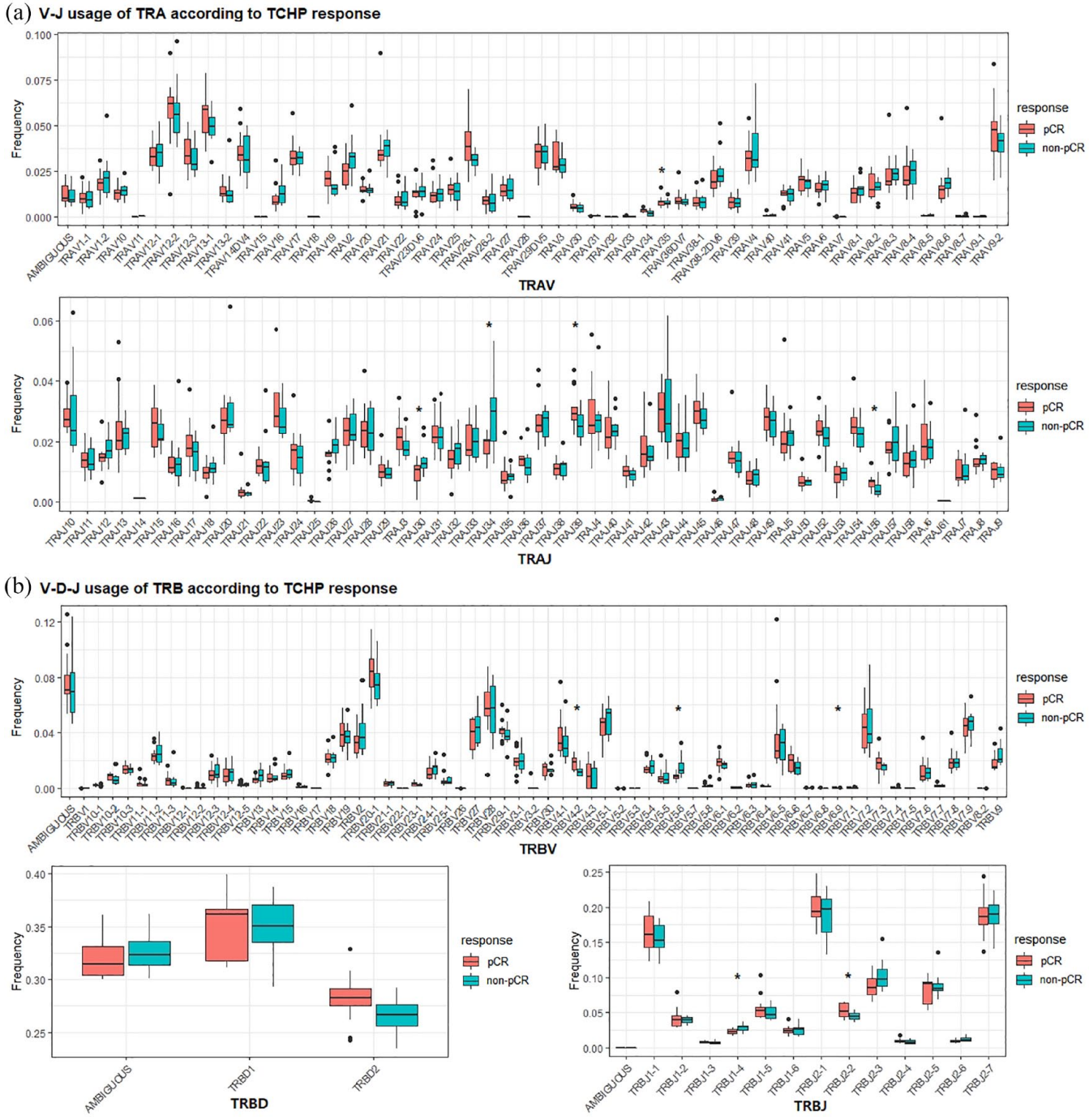

Figure 5. *V(D)J* usage of TCR. (a) Only *TRAV34*, *TRAJ30*, *TRAJ34*, *TRAJ39*, and *TRAJ56* showed p values <0.05, but their q values were not significant. (b) *TRBV4-2, TRBV5-6, TRBV6-9, and TRBJ1-5* also showed p values < 0.05, but their q values were not significant. *p value < 0.05.

Treatment response and gene expression analysis

The relationship of gene expression profile with treatment response was analyzed to better appreciate the differences between tumors that did and did not achieve pCR. A comparison of gene expression levels in the two groups showed that 185 genes were differentially expressed (Supplemental Table 9), including *ESR1* (estrogen receptor 1) (fold change 6.030, $p = 0.031$, $q = 0.185$) and *PDCD1* (fold change 2.6, $p = 0.049$, $q = 0.104$). However, none of these

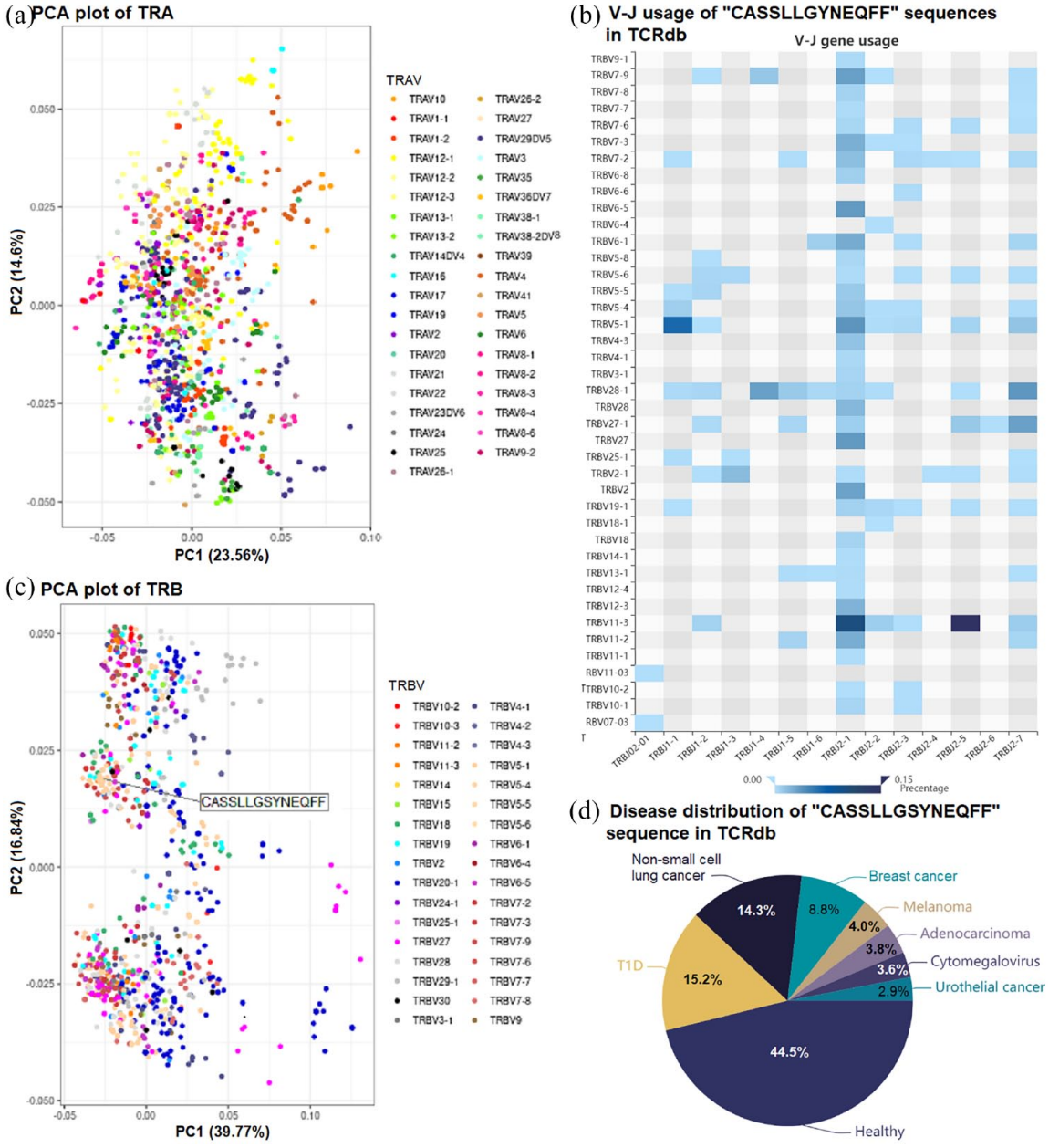

Figure 6. PCA plot of TCRs showing a significantly higher frequency in the pCR group, and of TCRdb search results. (a) TRA showed no cluster. (b) TRB showed a cluster around CDR3 AA of 'CASSLLGSYNEQFF' with *TRAV 5-1*. (c) A search of TCRdb showed that 3.6% of 'CASSLLGYNEQFF' CDR3 sequences had *TRBV5-1* and *TRBJ2-1*. (d) 'CASSLLGYNEQFF' CDR3 sequences were most frequently observed in healthy patients (44.5%), although 8.8% were present in patients with breast cancer.
PCA, principal component analysis; pCR, pathologic complete response; TCRs, T-cell receptors.

genes differed significantly after FDR adjustment. The top upregulated genes in the non-pCR group included those involved in responses to steroid hormones (GO: 0048545), the regulation of cell shape (GO: 0008360), the oxidation-reduction process (GO: 0055114), the regulation of *ERK1* and *ERK2* cascades (GO: 0070372), and the negative regulation of cell growth (GO: 0030308) (Figure 8).

Based on prediction analysis of microarray 50 (PAM50) classifications, nine patients (39%) were classified as being HER2-enriched, seven (30%) as luminal A, two (8%) as luminal B, three

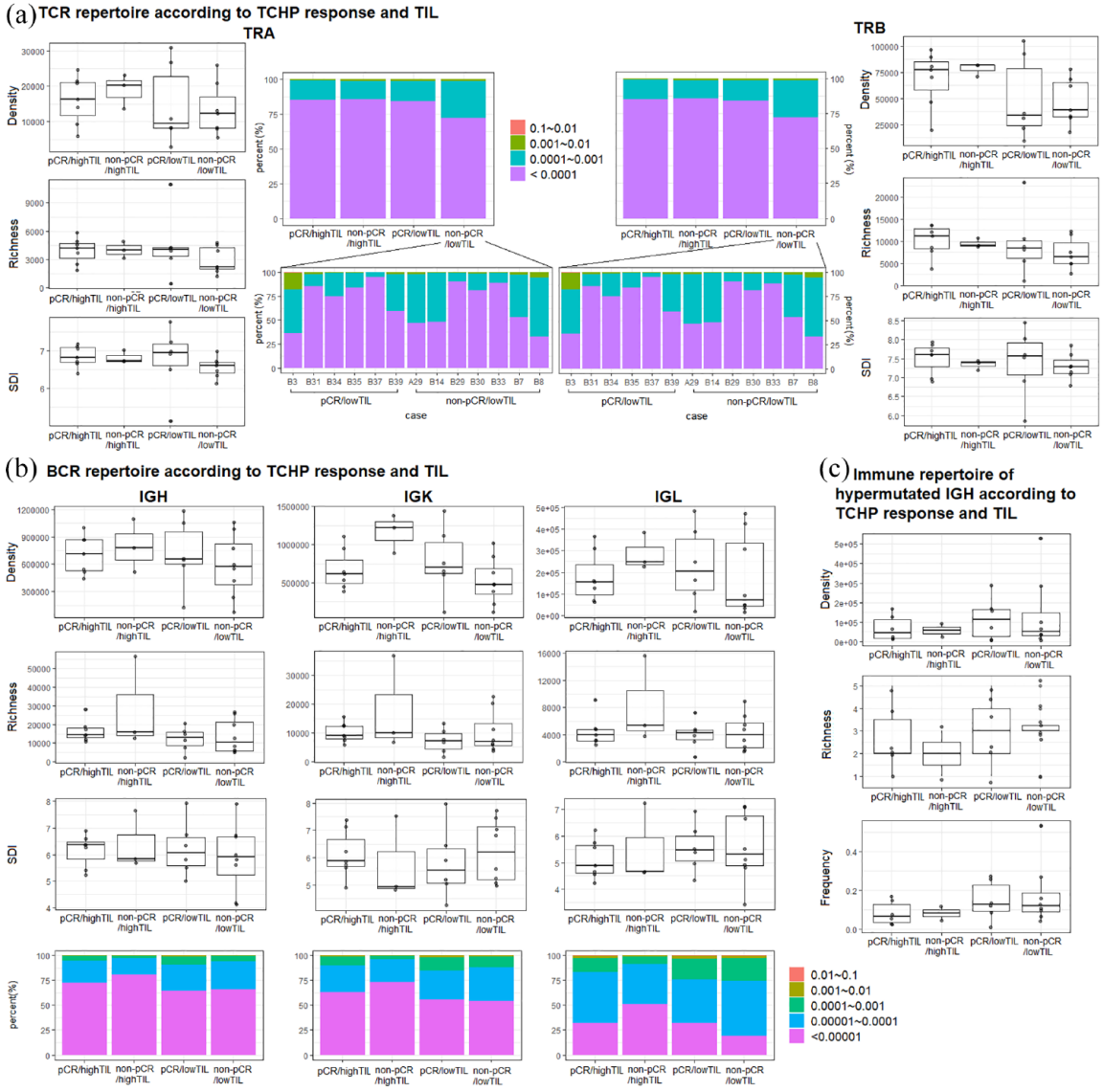

Figure 7. Immune repertoires as a function of both TCHP response and TILs. (a) The density, SDI, and richness of TRA and TRB did not differ significantly in comparisons of the pCR/highTIL and non-pCR/highTIL groups and the pCR/lowTIL and non-pCR/lowTIL groups. The composition of low-frequency sequences was lower in the non-pCR/lowTIL than in the pCR/lowTIL group. (b) There were no significant differences in the density, richness, and frequency of IGH, IGL, and IGK. Low-frequency clonal compositions of IGH, IGL, and IGK were higher in the non-pCR/highTIL than in the pCR/highTIL group. (c) There were no significant differences in hypermutated IGH.

IGH, immunoglobulin heavy; IGK, immunoglobulin kappa; IGL, immunoglobulin lambda; pCR, pathologic complete response; TCHP, trastuzumab, docetaxel, carboplatin, and pertuzumab; TILs, tumor-infiltrating lymphocytes.

(13%) as normal, and two (8%) as NA. HER2-enriched subtypes were more frequently observed in the pCR (seven patients, 54%) than in the non-pCR (two patients, 20%) group, but the difference was not statistically significant ($p = 0.099$) (Figure 8).

Immune cell sorting with CIBERSORT and treatment response

To examine the relationship between the composition of immune cells and treatment response, the decomposition of transcriptome sequencing data for immune cell sorting was performed using

Table 1. Immune repertoires of groups of patients as a function of both TCHP response and TILs.

	pCR/highTIL (*n* = 7)	Non-pCR/highTIL (*n* = 3)	*p* Value	pCR/lowTIL (*n* = 6)	Non-pCR/lowTIL (*n* = 7)	*p* Value
T-cell receptors						
SDI (mean)						
TRA	6.845	6.816	0.833	6.751	6.556	0.294
TRB	7.508	7.342	0.383	7.395	7.296	0.538
TRD	3.566	4.404	0.267	3.886	3.478	0.294
TRG	2.479	2.765	0.516	2.545	2.332	0.234
Density (mean)						
TRA	16,082	18,992	0.833	14,661	13,413	0.945
TRB	68,850	78,524	0.667	49,447	47,281	0.835
TRD	351	870	0.067	473	317	0.730
TRG	282	690	0.183	429	320	0.5338
Richness (mean)						
TRA	3,941	4,023	1	4,489	2,967	0.6282
TRB	10,097	9,472	0.833	9,589	7,229	0.6282
TRD	121	224	0.183	184	112	0.294
TRG	100	181	0.267	160	92	0.445
Composition (mean)						
TRA (%)						
1–10%	9 (0)	6 (0)	<0.001*	19 (1.0)	15 (1.0)	<0.001*
0.1–1%	793 (2.8)	362 (3.0)		761 (2.8)	844 (4.1)	
0.01–0.1%	15,423 (55.9)	6,121 (50.7)		12,202 (45.3)	13,090 (63.0)	
<0.01%	11,364 (41.2)	5,582 (46.2)		13,952 (51.8)	6,825 (32.9)	
TRB (%)						
1–10%	19 (0)	9 (0)	<0.001*	25 (0)	22 (0)	<0.001*
0.1–1%	739 (1.0)	385 (1.3)		706 (1.2)	735 (1.5)	
0.01–0.1%	9,710 (13.7)	3,724 (13.1)		8,441 (14.7)	13,396 (26.5)	
<0.01%	60,216 (85.1)	24,300 (85.5)		48,362 (84.1)	36,456 (72.0)	
TRG (%)						
TRGV9	60 (21.3)	243 (35.2)	0.116	89 (20.7)	67 (20.9)	0.616
Others	222 (78.7)	447 (64.8)	0.303	340 (79.3)	253 (79.1)	0.628

(Continued)

Table 1. (Continued)

	pCR/highTIL (*n* = 7)	Non-pCR/highTIL (*n* = 3)	*p* Value	pCR/lowTIL (*n* = 6)	Non-pCR/lowTIL (*n* = 7)	*p* Value
TRD (%)						
TRDV2	173 (49.4)	253 (29.1)	0.383	275 (58.1)	149 (47.2)	0.616
Others	177 (50.6)	617 (70.1)	0.067	198 (41.9)	167 (52.8)	0.628
B-cell receptors						
SDI (mean)						
IGH	6.161	6.397	1.000	6.219	6.708	0.755
IGK	6.121	5.760	0.667	5.797	7.121	0.573
IGL	5.124	5.513	1.000	5.552	6.371	1.000
Density (mean)						
IGH	707,776	794,986	0.667	709,185	668,712	0.491
IGK	668,675	1,164,079	0.067	780,198	595,456	0.228
IGL	178.790	286,026	0.183	234,046	205,732	0.490
Richness (mean)						
IGH	16,505	28,348	0.667	12,231	15,804	0.950
IGK	10,190	17,810	0.833	7,203	11,675	0.662
IGL	4,486	8,253	0.267	4,026	4,919	0.852
Richness, each isotype (mean)						
IGHM (%)	1,720 (10.4)	2,219 (7.8)	0.953	1,318 (10.7)	1,220 (7.7)	0.176
IGHD (%)	181 (1.1)	162 (0.6)	0.268	176 (1.3)	39 (0.2)	0.581
IGHA (%)	3,418 (20.7)	6,156 (21.7)	0.802	2,922 (23.9)	4,563 (28.9)	0.197
IGHA1	2,199 (64.3)	4,006 (65.1)	0.560	1,896 (64.9)	2,944 (64.5)	0.468
IGHA2	1,216 (35.6)	2,142 (34.8)	0.908	1,024 (35.1)	1,617 (35.5)	0.452
Indeterminate	3 (0.1)	8 (0.1)	0.576	1 (0.0)	2 (0.0)	0.896
IGHG (%)	11,181 (67.8)	19,786 (69.8)	0.915	7,806 (63.8)	9,974 (63.1)	0.075
IGHG1	6,972 (62.3)	12,687 (64.1)	0.385	4,490 (57.5)	6,712 (67.3)	0.840
IGHG2	2,523 (22.6)	4,186 (21.2)	0.352	2,154 (27.6)	2,007 (20.1)	0.582
IGHG3	1,001 (9.0)	1,529 (7.7)	0.087	732 (9.4)	696 (7.0)	0.873
IGHG4	255 (2.3)	466 (2.4)	0.223	141 (1.8)	162 (1.6)	0.320
Indeterminate	430 (3.8)	918 (4.6)	0.243	289 (2.7)	397 (4.0)	0.331
Indeterminate	5 (0.0)	25 (0.1)		9 (0.1)	8 (0.1)	0.813

(Continued)

Table 1. (Continued)

	pCR/highTIL (*n* = 7)	Non-pCR/highTIL (*n* = 3)	*p* Value	pCR/lowTIL (*n* = 6)	Non-pCR/lowTIL (*n* = 7)	*p* Value
Composition (mean)						
IGH (%)						
1–10%	12 (0.1)	10 (0.0)	<0.001*	10 (0.1)	13 (0.1)	<0.001*
0.1–1%	265 (0.7)	56 (0.4)		100 (1.0)	94 (0.6)	
0.01–0.1%	914 (4.7)	284 (2.3)		860 (8.2)	2121 (5.8)	
0.001–0.01%	3,638 (22.0)	2,029 (16.7)		6,402 (26.2)	5,698 (30.9)	
<0.001%	11,957 (72.5)	22,827 (80.6)		7,898 (64.5)	9,903 (62.6)	
IGK (%)						
1–10%	13 (0.1)	12 (0.1)	<0.001*	11 (0.2)	11 (0.1)	<0.001*
0.1–1%	119 (1.2)	100 (0.6)		112 (1.5)	101 (0.9)	
0.01–0.1%	899 (8.8)	634 (3.5)		962 (13.4)	1,326 (11.4)	
0.001–0.01%	2,753 (27.0)	4,083 (22.9)		2,109 (29.3)	3,925 (33.6)	
<0.001%	6,405 (62.9)	12,981 (72.9)		4,007 (55.6)	6,311 (54.1)	
IGL (%)						
1–10%	14 (0.3)	15 (0.2)	<0.001*	14 (0.3)	9 (0.2)	<0.001*
0.1–1%	222 (2.1)	45 (1.1)		103 (2.6)	88 (2.0)	
0.01–0.1%	764 (14.6)	249 (7.0)		853 (21.2)	1,005 (23.4)	
0.001–0.01%	2,297 (51.2)	1,439 (40.7)		1,769 (43.9)	2,383 (55.4)	
<0.001%	1,424 (31.8)	4,207 (51.0)		1,286 (32.0)	818 (19.0)	
Hypermutated IGH (mean)						
Density	50,414	30,793	0.560	125,174	131,561	0.943
Richness	2.2	1.3	0.321	2.8	3.4	0.495
Frequency	0.079	0.042	0.422	0.144	0.172	0.726

*p value < 0.05.
Non-pCR, non-pathologic complete response; pCR, pathologic complete response; SDI, shannon diversity index; TIL, tumor-infiltrating lymphocytes.

CIBERSORT. The most common cells in the pCR group were M2 macrophages and naïve B cells, whereas the most common cells in the non-pCR group were M2 macrophages, naïve B cells, and monocytes. Few eosinophils, neutrophils, naïve CD4 T cells, and γδ T cells were observed. There were no significant differences between the pCR and non-pCR groups (Supplemental Table 10, Supplemental Figure 1).

Discussion

T cells and B cells function as essential components of the humoral and cellular immune

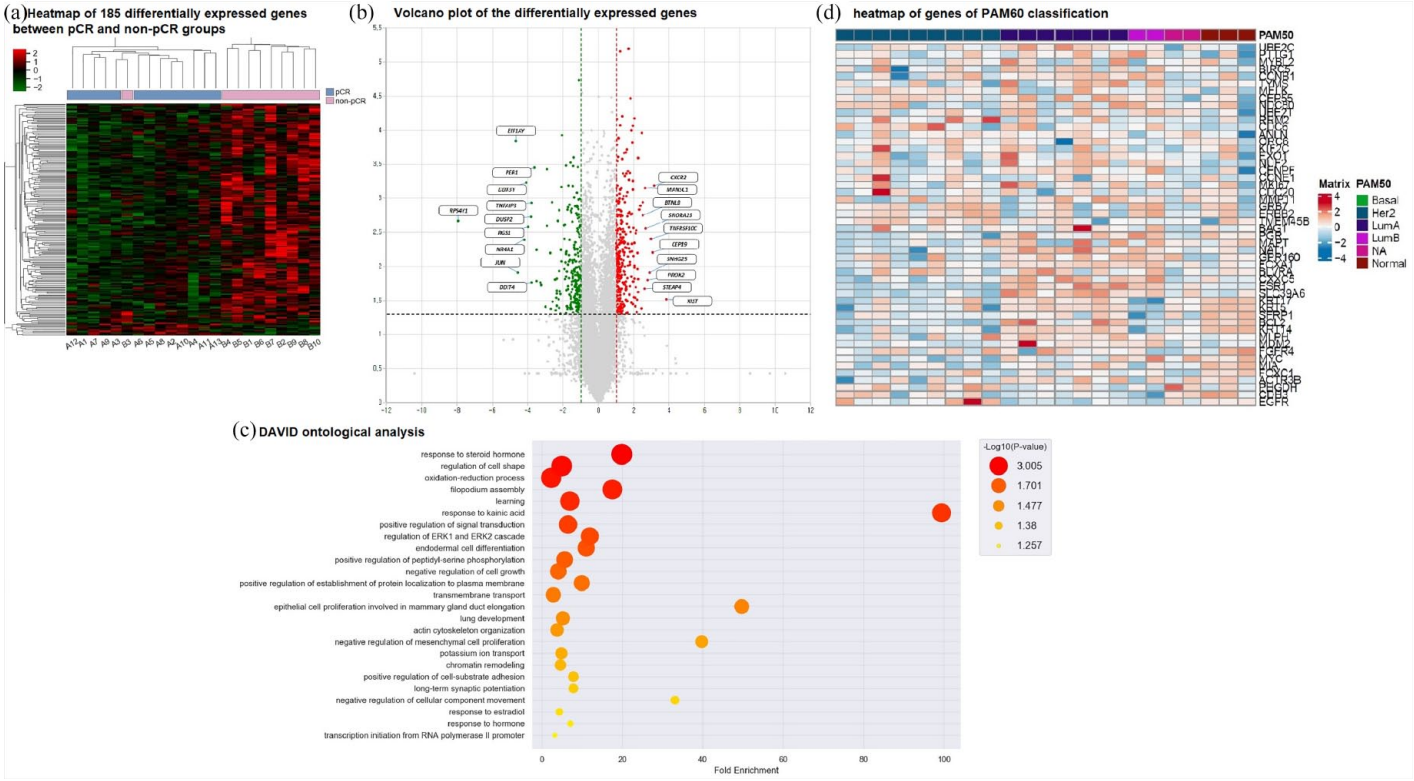

Figure 8. RNA expression in samples in the main experiment. (a) Data were normalized and scaled to give all genes equal variance. pCR and non-pCR groups differed in the expression of 185 genes. (b) The top 10 genes are displayed in the order of fold change. (c) Genes associated with responses to kainic acid and steroid hormones, filopodium assembly, regulation of cell shape, and regulation of *ERK1* and *ERK2* cascades were significantly more activated in the non-pCR group with fold enrichment >2 and p value <0.05. (d) Heatmap of genes of PAM50 classification. The expression data have been normalized.
pCR, pathologic complete response.

systems, with TCR and BCR repertoires having important effects on a wide range of diseases, including malignancies, autoimmune disorders, and infectious diseases.[11,13,29–32] To our knowledge, this study is the first to describe the immune repertoires of T cells (TRA, TRB, TRG, and TRD) and B cells (IGH, IGK, and IGL) in the tumor tissues of HER2-positive early-stage breast cancer and the first to observe changes in immune repertoires after TCHP treatment.

Several previous studies have described changes in immune repertoires after treatment. For example, the SDI of TCR repertoires has been reported to increase in NSCLC after several lines of treatment, and to be associated with durable clinical benefit and longer progression-free survival.[12] Conversely, adjuvant chemotherapy has been found to result in immune cell depletion and suppressed T-cell immunity.[33,34] In their recent breast cancer study, Uruena *et al.* found that the SDI and density of TCR decreased, and the Simpson index increased after neoadjuvant treatment.[35] This study, targeting HER2-positive tumors treated with TCHP, found similar results of decreased SDI, richness, and density of TCR repertoires after TCHP treatment.

The main experiment showed that the richness, density, and SDI of the TCR and BCR repertoires did not differ significantly between the pCR and non-pCR groups, regardless of subdivision according to TILs. Moreover, there was no significant difference between BCR isotypes and hypermutated BCR in the pCR and non-pCR groups. These results differ from those of a previous study of the T-cell repertoire in NSCLC, which found a low evenness of TCR in tumor tissue, with a high percentage of the top 1% of sequences being significantly associated with pCR.[11] In addition, a high SDI in patients with cervical cancer following concurrent chemotherapy was associated with a better prognosis.[34] However, although only TRB was examined, a

previous study on immune repertoires in HER2-positive tumors did not find significant differences in TRB immune repertoires in pretreatment biopsy samples, as in the current study.[1] This may be due to the characteristics of these tumors, as HER2-positive breast cancer has a lower tumor mutation burden than NSCLC, with fewer neoantigens.[36–38]

γδ-T cells are components of the innate immune system with multiple favorable antitumor characteristics.[39] Trastuzumab has been reported to enhance γδ-T cell-mediated antibody-dependent cytotoxicity against HER2-overexpressing breast cancer cell lines *in vitro*. Vγ9Vδ2 T cells are among the most abundant γδ-T cell subpopulations, accounting for approximately 5% of peripheral blood T cells.[39] Vγ9Vδ2 T cells have been reported to recognize various cancers and exert strong antitumor effects, by releasing pro-inflammatory cytokines, granzymes, and perforin and binding to apoptosis-inducing receptors. This study found that there was no significant difference in the immune repertoires of TRD and TRG and the distribution of TRGV9 and TRDD2 between the pCR and non-pCR groups.

One TRA and three TRBs, all in the pCR group, were found to have sequences targeting HER2. Because these sequences were observed in a small proportion of the pCR group, they alone do not explain the response to TCHP. PCA analysis of TRBs having a significantly higher frequency in the pCR group showed a cluster around the CDR3 AA of 'CASSLLGYNEQFF'. Although the encoded antigen has not been determined, it was found in TCRdb in patients with various cancers, including breast cancer.[20] In general, this study analyzed immune repertoire diversity on a bulk level, whereas TCR/BCR sequence identification using a particular antigen is more suitable for a single-cell approach. This approach may enable the identification of the specific TRA and specific T-cell function, as well as help to determine the antitumor effect of T cells with this sequence.

Based on the criteria used in the TCR3d, we attempted to identify similar CDR3 sequences with substitution of 1 or 2 AA using the Levenshtein distance.[21,40] The Levenshtein distance, however, is a mathematical method of calculating the similarity of two sentences, but it does not take into account the biological similarities of AA. PAM or BLOSUM are more biologically oriented methods of assessing the similarity of AA.[41] However, due to the lack of research results using these methods, it was difficult to determine objective criteria to evaluate the similarity of TCR.

RNA expression has been reported to affect treatment responses in patients with HER2 positive breast cancer. A high expression of FGFR4, ERBB2/HER2, low level of ESR1, molecular intrinsic subtype, and immune cell activation have been reported to be associated with pCR.[42–45] This study found that pathways associated with response to steroid hormones, *ERK1*/*ERK2* cascade regulation, and negative regulation of cell growth were upregulated in the non-pCR group. Similarly, pCR rates have been reported as lower in estrogen receptor-positive than estrogen receptor-negative patients, with activation of the ERK pathway contributing to resistance to trastuzumab.[46] However, in this study, there was no significant difference in the PAM50 classification, which was reported to be significantly associated with HER2 treatment response.[42]

This study had several limitations. First, the number of cases is relatively small, and the observation period was too short to assess the associations of breast cancer characteristics with survival and disease progression. Moreover, because this study was not a single-cell TCR repertoire analysis, it was not possible to determine the function of cells bearing each sequence. We unsuccessfully sought to find single-cell RNA and TCR sequencing results obtained in patients with HER2 positive tumors in the public database. These limitations may be overcome by single-cell-based research, which includes determinations of cell function. Finally, although various DNA mutations are known prognostic factors in HER2-positive breast cancer, we did not perform studies on DNA and tumor mutation burden in this study.[42,47] Further research on these areas will also be needed.

In summary, despite a broad investigation of the immune repertoire, the role of the diversity, richness, and density of the TCR and BCR repertoires as predictive markers for TCHP response was not identified. However, the inclusion of a large number of low-frequency sequences was associated with good treatment response. TRB sequences having significantly higher frequencies in the pCR group showed a cluster around 'CASSLLGYNEQFF' CDR3 AA. Further

research using single cells may be helpful in evaluating the antitumor effects of TCR sequences.

Declarations

Ethics approval and consent to participate

The Institutional Review Board of Asan Medical Center approved this study (approval number, 2019-0527). All participating patients provided signed written informed consent.

Consent for publication

The patients included in this study consented to have their study-related data published.

Author contribution(s)

Junyoung Shin: Data curation; Formal analysis; Investigation; Methodology; Validation; Visualization; Writing – original draft.

Baknoon Ham: Investigation; Software.

Jeong-Han Seo: Investigation; Software; Visualization.

Sae Byul Lee: Resources; Writing – review & editing.

In Ah Park: Methodology; Writing – review & editing.

Gyungyub Gong: Funding acquisition; Resources; Writing – review & editing.

Sung-Bae Kim: Project administration; Supervision; Writing – review & editing.

Hee Jin Lee: Methodology; Supervision; Writing – review & editing.

Acknowledgements

None.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was partly supported by a grant from the Asan Institute for Life Sciences, Asan Medical Center, Seoul, Korea (Grant No. 2019-144).

Competing interests

Sung-Bae Kim is a consultant on the advisory boards of Novartis, Astrazeneca, Lilly, Dae Hwa Pharmaceutical Co. Ltd., ISU Abxis, and Daiichi-Sankyo; received research funding from Novartis, Sanofi-Aventis, and DongKook Pharm. Co.; and owns stock in Genopeaks and NeogenTC. Hee Jin Lee is CEO and owns stock in NeogenTC. None of the other authors has any conflicts of interest to declare in relation to this study.

Availability of data and materials

All datasets generated for this study are included in the article/supplementary material.

Translational relevance

None.

ORCID iD

Junyoung Shin https://orcid.org/0000-0002-4833-9738

Supplemental material

Supplemental material for this article is available online.

References

1. Force J, Howie LJ, Abbott SE, *et al*. Early stage HER2-positive breast cancers not achieving a pCR from neoadjuvant trastuzumab- or pertuzumab-based regimens have an immunosuppressive phenotype. *Clin Breast Cancer* 2018; 18: 410–417.
2. Genuino AJ, Chaikledkaew U, The DO, *et al*. Adjuvant trastuzumab regimen for HER2-positive early-stage breast cancer: a systematic review and meta-analysis. *Expert Rev Clin Pharmacol* 2019; 12: 815–824.
3. Senkus E, Kyriakides S, Ohno S, *et al*. Primary breast cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2015; 26: v8–v30.
4. Ignatiadis M, Van den Eynden G, Roberto S, *et al*. Tumor-infiltrating lymphocytes in patients receiving trastuzumab/pertuzumab-based chemotherapy: A TRYPHAENA substudy. *J Natl Cancer Inst* 2019; 111: 69–77.
5. Salgado R, Denkert C, Campbell C, *et al*. Tumor-infiltrating lymphocytes and associations with pathological complete response and event-free survival in HER2-positive early-stage breast cancer treated with lapatinib and trastuzumab: a secondary analysis of the NeoALTTO trial. *JAMA Oncol* 2015; 1: 448–454.
6. Ingold Heppner B, Untch M and Denkert C, *et al*. Tumor-infiltrating lymphocytes: a predictive and prognostic biomarker in neoadjuvant-treated HER2-positive breast cancer. *Clin Cancer Res* 2016; 22: 5747–5754.

7. Hoehn KB, Fowler A, Lunter G, *et al.* The diversity and molecular evolution of B-Cell receptors during infection. *Mol Biol Evol* 2016; 33: 1147–1157.

8. Mora T and A W. *Systems immunology*. 1st ed. Boca Raton: CRC Press, 2019, p.354.

9. Janeway CA Jr, Travers P and Walport M *et al. Immunobiology: the immune system in health and disease*. 5th ed. New York: Garland Science, 2001.

10. Collins AM and Watson CT. Immunoglobulin light chain gene rearrangements, receptor editing and the development of a self-tolerant antibody repertoire. *Front Immunol* 2018; 9: 2249.

11. Casarrubios M, Cruz-Bermúdez A, Nadal E, *et al.* Pretreatment tissue TCR repertoire evenness is associated with complete pathologic response in patients with NSCLC receiving neoadjuvant chemoimmunotherapy. *Clin Cancer Res* 2021; 27: 5878–5890.

12. Liu YY, Yang QF, Yang JS, *et al.* Characteristics and prognostic significance of profiling the peripheral blood T-cell receptor repertoire in patients with advanced lung cancer. *Int J Cancer* 2019; 145: 1423–1431.

13. Cui JH, Lin KR, Yuan SH, *et al.* TCR repertoire as a novel indicator for immune monitoring and prognosis assessment of patients with cervical cancer. *Front Immunol* 2018; 9: 2729.

14. Farshid G, Bilous M, Morey A, *et al.* ASCO/CAP 2018 breast cancer HER2 testing guidelines: summary of pertinent recommendations for practice in Australia. *Pathology* 2019; 51: 345–348.

15. Jang N, Kwon HJ, Park MH, *et al.* Prognostic value of tumor-infiltrating lymphocyte density assessed using a standardized method based on molecular subtypes and adjuvant chemotherapy in invasive breast cancer. *Ann Surg Oncol* 2018; 25: 937–946.

16. Goto W, Kashiwagi S, Asano Y, *et al.* Predictive value of improvement in the immune tumour microenvironment in patients with breast cancer treated with neoadjuvant chemotherapy. *ESMO Open* 2018; 3: e000305.

17. Salgado R, Denkert C, Demaria S, *et al.* The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs working group 2014. *Ann Oncol* 2015; 26: 259–271.

18. Symmans WF, Peintinger F, Hatzis C, *et al.* Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J Clin Oncol* 2007; 25: 4414–4422.

19. Ogston KN, Miller ID, Payne S, *et al.* A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. *Breast* 2003; 12: 320–327.

20. Chen SY, Yue T, Lei Q, *et al.* TCRdb: a comprehensive database for T-cell receptor sequences with powerful search function. *Nucleic Acids Res* 2021; 49: D468–D474.

21. Gowthaman R and Pierce BG. TCR3d: the T cell receptor structural repertoire database. *Bioinformatics* 2019; 35: 5323–5325.

22. Shugay M, Bagaev DV, Zvyagin IV, *et al.* VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res* 2018; 46: D419–D427.

23. Souter MNT and Eckle SBG. Biased MAIT TCR usage poised for limited antigen diversity? *Front Immunol* 2020; 11: 1845.

24. Dash P, Fiore-Gartland AJ, Hertz T, *et al.* Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature* 2017; 547: 89–93.

25. Chen B, Khodadoust MS, Liu CL, *et al.* Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol* 2018; 1711: 243–259.

26. Newman AM, Steen CB, Liu CL, *et al.* Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* 2019; 37: 773–782.

27. Parker JS, Mullins M, Cheang MC, *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009; 27: 1160–1167.

28. Luen SJ, Salgado R, Fox S, *et al.* Tumour-infiltrating lymphocytes in advanced HER2-positive breast cancer treated with pertuzumab or placebo in addition to trastuzumab and docetaxel: a retrospective analysis of the CLEOPATRA study. *Lancet Oncol* 2017; 18: 52–62.

29. Liu X, Zhang W, Zhao M, *et al.* T cell receptor β repertoires as novel diagnostic markers for systemic lupus erythematosus and rheumatoid arthritis. *Ann Rheum Dis* 2019; 78: 1070–1078.

30. Li N, Yuan J, Tian W, *et al.* T-cell receptor repertoire analysis for the diagnosis and treatment of solid tumor: a methodology and clinical applications. *Cancer Commun (Lond)* 2020; 40: 473–483.

31. Niu X, Li S, Li P, *et al.* Longitudinal analysis of T and B cell receptor repertoire transcripts

reveal dynamic immune response in COVID-19 patients. *Front Immunol* 2020; 11: 582010.

32. Zhang W, Feng Q, Wang C, *et al.* Characterization of the B cell receptor repertoire in the intestinal mucosa and of tumor-infiltrating lymphocytes in colorectal adenoma and carcinoma. *J Immunol* 2017; 198: 3719–3728.
33. Gustafson CE, Jadhav R, Cao W, *et al.* Immune cell repertoires in breast cancer patients after adjuvant chemotherapy. *JCI Insight* 2020; 5: e134569.
34. Li R, Liu Y, Yin R, *et al.* The dynamic alternation of local and systemic tumor immune microenvironment during concurrent chemoradiotherapy of cervical cancer: a prospective clinical trial. *Int J Radiat Oncol Biol Phys* 2021; 110: 1432–1441.
35. Urueña C, Lasso P, Bernal-Estevez D, *et al.* The breast cancer immune microenvironment is modified by neoadjuvant chemotherapy. *Sci Rep* 2022; 12: 7981.
36. Wen Y, Ouyang D, Chen Q, *et al.* Prognostic value of tumor mutation burden and the relationship between tumor mutation burden and immune infiltration in HER2+ breast cancer: a gene expression-based study. *Gland Surg* 2022; 11: 100–114.
37. Castle JC, Uduman M, Pabla S, *et al.* Mutation-derived neoantigens for cancer immunotherapy. *Front Immunol* 2019; 10: 1856.
38. Reuben A, Zhang J, Chiou SH, *et al.* Comprehensive T cell repertoire characterization of non-small cell lung cancer. *Nat Commun* 2020; 11: 603.
39. Hoeres T, Smetak M, Pretscher D, *et al.* Improving the efficiency of Vγ9Vδ2 T-cell immunotherapy in cancer. *Front Immunol* 2018; 9: 800.
40. Gordin M, Philip H, Zilberberg A, *et al.* Breast cancer is marked by specific, public T-cell receptor CDR3 regions shared by mice and humans. *PLoS Comput Biol* 2021; 17: e1008486.
41. Mount DW. Comparison of the PAM and BLOSUM amino acid substitution matrices. *CSH Protoc* 2008; 2008: pdb.ip59.
42. Tanioka M, Fan C, Parker JS, *et al.* Integrated analysis of RNA and DNA from the phase III trial CALGB 40601 identifies predictors of response to trastuzumab-based neoadjuvant chemotherapy in HER2-positive breast cancer. *Clin Cancer Res* 2018; 24: 5292–5304.
43. Fumagalli D, Venet D, Ignatiadis M, *et al.* RNA sequencing to predict response to neoadjuvant anti-HER2 therapy: a secondary analysis of the NeoALTTO randomized clinical trial. *JAMA Oncol* 2017; 3: 227–234.
44. Carey LA, Berry DA, Cirrincione CT, *et al.* Molecular heterogeneity and response to neoadjuvant human epidermal growth factor receptor 2 targeting in CALGB 40601, a randomized phase III trial of paclitaxel plus trastuzumab with or without lapatinib. *J Clin Oncol* 2016; 34: 54–549.
45. Zou Y, Zheng S, Xie X, *et al.* N6-methyladenosine regulated FGFR4 attenuates ferroptotic cell death in recalcitrant HER2-positive breast cancer. *Nat Commun* 2022; 13: 2672.
46. Zazo S, González-Alonso P, Martín-Aparicio E, *et al.* Autocrine CCL5 effect mediates trastuzumab resistance by ERK pathway activation in HER2-positive breast cancer. *Mol Cancer Ther* 2020; 19: 1696–1707.
47. Guarneri V, Dieci MV, Frassoldati A, *et al.* Prospective biomarker analysis of the randomized CHER-LOB study evaluating the dual anti-HER2 treatment with trastuzumab and lapatinib plus chemotherapy as neoadjuvant therapy for HER2-positive breast cancer. *Oncologist* 2015; 20: 1001–1010.