# scientific reports



## **OPEN** Globally shared TCR repertoires within the tumor-infiltrating lymphocytes of patients with metastatic gynecologic cancer

Taehoon Kim<sup>1,5</sup>, Hyeonseob Lim<sup>1,5</sup>, Soyeong Jun<sup>1,5</sup>, Junsik Park<sup>2</sup>, Dongin Lee<sup>1</sup>, Ji Hyun Lee<sup>3,4</sup>, Jung-Yun Lee<sup>2⊠</sup> & Duhee Bang<sup>1⊠</sup>

Gynecologic cancer, including ovarian cancer and endometrial cancer, is characterized by morphological and molecular heterogeneity. Germline and somatic testing are available for patients to screen for pathogenic variants in genes such as BRCA1/2. Tissue expression levels of immunogenomic markers such as PD-L1 are also being used in clinical research. The basic therapeutic approach to gynecologic cancer combines surgery with chemotherapy. Immunotherapy, while not yet a mainstream treatment for gynecologic cancers, is advancing, with Dostarlimab recently receiving approval as a treatment for endometrial cancer. The goal remains to harness stimulated immune cells in the bloodstream to eradicate multiple metastases, a feat currently deemed challenging in a typical clinical setting. For the discovery of novel immunotherapy-based tumor targets, tumor-infiltrating lymphocytes (TILs) give a key insight on tumor-related immune activities by providing T cell receptor (TCR) sequences. Understanding the TCR repertoires of TILs in metastatic tissues and the circulation is important from an immunotherapy standpoint, as a subset of T cells in the blood have the potential to help kill tumor cells. To explore the relationship between distant tissue biopsy regions and blood circulation, we investigated the TCR beta chain (TCR $\beta$ ) in bulk tumor and matched blood samples from 39 patients with gynecologic cancer. We found that the TCR clones of TILs at different tumor sites were globally shared within patients and had high overlap with the TCR clones in peripheral blood.

Gynecologic cancers, including ovarian cancer and endometrial cancer, cause widespread mortality among women<sup>1,2</sup>. Morphological and molecular heterogeneity in gynecologic cancers has been shown to affect patient survival<sup>3,4</sup>. Pathologic assessment of genomic (e.g. TP53, PTEN, and BRCA1/2) and immunogenomic (e.g. PD-1 and PD-L1) features to distinguish diverse subtypes of gynecologic cancer is important to determine optimal treatment strategies<sup>5-13</sup>. In this context, the approval of Dostarlimab is noteworthy, as it specifically targets patients with mismatch repair deficiency (dMMR) or microsatellite instability-high (MSI-H) markers<sup>14</sup>. The US Food and Drug Administration (FDA) granted accelerated approval to Dostarlimab for monotherapy use in patients with mismatch repair deficiency (dMMR) in 2021<sup>15</sup>. The approval was further extended in 2023, allowing the use of Dostarlimab for recurrent or advanced endometrial cancer patients with either mismatch repair deficiency (dMMR) or microsatellite instability-high (MSI-H), both in combination with chemotherapy and as monotherapy<sup>16,17</sup>. Additionally, the FDA granted approval to Pembrolizumab for patients with persistent, recurrent or metastatic cervical cancer whose tissue expresses PD-L1 (CPS  $\ge$  1) in 2021<sup>18,19</sup>. This regulatory endorsement underscores the importance of genomic and immunogenomic profiling in tailoring treatment approaches for gynecologic cancers.

The success of immune checkpoint inhibitors has led to heightened interest in the tumor immune microenvironment as a factor in the diagnosis and treatment of gynecologic cancers<sup>20-22</sup>. Investigation of TCR repertoire

<sup>1</sup>Department of Chemistry, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea. <sup>2</sup>Department of Obstetrics and Gynecology, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea. <sup>3</sup>Department of Clinical Pharmacology and Therapeutics, College of Medicine, Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Korea. <sup>4</sup>Department of Biomedical Science and Technology, Kyung Hee Medical Science Research Institute, Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Korea. <sup>5</sup>These authors contributed equally: Taehoon Kim, Hyeonseob Lim and Soyeong Jun. <sup>22</sup>email: yodrum682@ gmail.com; duheebang@yonsei.ac.kr

which is expressed by heterogeneous lymphocyte populations is important for understanding the immune activities surrounding tumors. Previous studies have shown evidence of T cell expansion with TCR repertoire or flow cytometry data of TILs in gynecologic cancers depending on the subtype of the tumor tissue<sup>23,24</sup> and T cells<sup>25</sup>. However, the number of tissue samples per patient was not high enough in these studies to assert that the TCRs found in tissues are highly individual-specific. Also, an explanation of the TCR found in multiple tissues and its overlap with blood was insufficient.

In this study, we sought to find patterns in inter-sample TCR repertoire overlap which consists of overlap between tissue-tissue and tissue-blood. We performed sequencing of the TCR $\beta$  from both tumor and peripheral blood samples of 31 patients with ovarian cancer and 8 patients with endometrial cancer. We also examined multiple metastatic tumor sites in 8 patients and compared to the primary tumor. Experiments were performed in duplicate to focus on clones that were twice-observed in both replicates of each sample. The overall scheme is shown in Fig. 1A,B. Our work highlights the possibility of using patient peripheral blood samples to discover TCR clones that are present among TILs distributed across multiple tissue sites.

### Methods

#### Subjects

Thirty-nine patients who underwent surgery in 2020 or 2021 in the Department of Obstetrics and Gynecology at Yonsei University, Korea, participated in the study. Institutional Review Board (IRB) approval was obtained from Yonsei University Health System (IRB number 4-2018-0342). All methods were performed in accordance with relevant guidelines and regulations. All patients gave informed consent to participate in the study. Tissue samples were collected from 31 patients with ovarian cancer and 8 patients with endometrial cancer. The 33 cases from 31 ovarian cancer patients included borderline ovarian tumor and peritoneal cancer (ICD-O codes: D39.1, C48.1, C48.2). Whole blood samples were collected from each patient prior to tumor resection. The patient labeled "OV30" and "OV31" had serial samples. The first series of samples were obtained for the purpose of pathologic confirmation and the second series of samples were obtained while reducing the size of the tumor.

#### **Clinical sample processing**

Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood samples by a density gradient method using Ficoll-Paque PLUS<sup>\*</sup> (GE Healthcare). DNA was extracted from tissue and PBMC samples using DNeasy Blood & Tissue Kits<sup>\*</sup> (Qiagen). The concentration of DNA was measured using a Qubit dsDNA BR Assay Kit<sup>\*</sup> (Invitrogen). Twenty microliters of extracted DNA were used for each experimental replicate, with input DNA ranging from 94 to 4240 ng.

#### TCRβ library preparation

Genomic DNA from tissues and PBMCs was amplified by three steps of multiplex PCR. Before library preparation, primer rebalancing with an oligo pool containing primer binding sites and random barcodes was conducted to prevent amplification bias. In the first PCR, the forward primer targeted the V region upstream of complementarity-determining region 3 (CDR3) in TCR $\beta$ , and the reverse primer targeted the J region of TCR $\beta$ . The primer sequences used in the first PCR contained shared flanking sequences, which enabled the second PCR to attach Illumina sequencing adapters (Supplementary Table S4). The first PCR was performed with 20 µL template DNA, 5 µL each of forward and reverse primer mix (100 µM), 10 µL 5X Q solution, and 25 µL 2X Qiagen Multiplex PCR Master Mix<sup>\*</sup> (Qiagen). The first PCR was performed with the following conditions: 95 °C for 10 min, 25 cycles of 95 °C for 30 s, 58 °C for 90 s, and 72 °C for 90 s, and a final extension at 72 °C for 3 min. The PCR products were purified using 1.5X AMPure XP beads (Beckman Coulter) and eluted in 12 µL nuclease-free water.

The second PCR was performed with 12  $\mu$ L of the first PCR product, 2.5  $\mu$ L each of the forward and reverse primer mix (10  $\mu$ M), 10  $\mu$ L 5X Q solution, and 25  $\mu$ L 2X Qiagen Multiplex PCR Master Mix<sup>\*</sup> (Qiagen). The second PCR was performed with the following conditions: 95 °C for 15 min, 6 cycles of 95 °C for 30 s, 60 °C for 40 s, and 72 °C for 1 min, and a final extension at 72 °C for 10 min. The PCR products were purified using 1.5X AMPure XP beads and eluted in 22  $\mu$ L nuclease-free water.

The third PCR was performed with 22  $\mu$ L of the second PCR product, 2.5  $\mu$ L each of the forward and reverse primer mix (10  $\mu$ M), and 25  $\mu$ L 2X KAPA Hifi PCR<sup>\*</sup> solution (KAPA Biosystems). The third PCR was performed with the following conditions: 98 °C for 3 min, 6 cycles of 98 °C for 10 s, 65 °C for 30 s, and 72 °C for 30 s, and a final extension at 72 °C for 5 min. The PCR products were purified using 1.2X AMPure XP beads and eluted in 22  $\mu$ L nuclease-free water. The concentration of the final PCR product was measured using a Qubit dsDNA BR Assay Kit<sup>\*</sup> (Invitrogen).

#### Sequencing and data processing

The library was sequenced as  $2 \times 150$  bp paired-end readout with NovaSeq 6000 System<sup>\*</sup> (illumina). The fastq files were randomly subsampled to a size of 1 Gb using Seqtk (version 1.3-r106). CDR3 sequences were called by MiXCR (version 3.0.13)<sup>26</sup> using the following options: -s Homosapiens –starting-material dna –adapters adapters-present –impute-germline-on-export –5-end v-primers –3-end j-primers –receptor-type trb –region-of-interest CDR3 –only-productive –align "-OvParameters.geneFeatureToAlign = VRegion" –assemble "-Oad-dReadsCountOnClustering = true" –verbose.

Clones with non-human epitopes were filtered out by a screen against 7,276,705 CDR3 amino acid sequences downloaded from TCRdb<sup>27</sup>, McPAS-TCR<sup>28</sup>, PIRD<sup>29</sup>, TCR3d<sup>30</sup>, and VDJdb<sup>31,32</sup>. Some clones were excluded if junction sequence of the clone does not contain proper conserved residues (e.g. V-Cystein at 104th or J-Phenylalanine at 118th to 129th). The average number of unique CDR3 amino acid sequences for each replicate in patients was 12,917.4, ranging from 2984.3 to 82,681.7.



**Figure 1.** TCR clonal similarity between patient samples. (**A**) Scheme of the workflow in the study. PBMCs were separated from patients' whole blood collected prior to tumor resection. Extracted DNA from paired PBMC and tissue samples was subjected to TCR $\beta$  library preparation. FASTQ files were subsampled to a size of 1 Gb. CDR3 sequences were excluded if they matched with downloaded public CDR3 sequences (see Methods). (**B**) A brief summary of the TCR $\beta$  library preparation procedures (see Methods). (**C**) Pairwise overlap was calculated and applied to hierarchical clustering (see Methods). Cases labeled as "OV30-1st" and "OV30-2nd" were from a single patient that underwent surgery twice. Six representative insets of patients with metastasis are followed. CDR3s observed in both replicates were termed twice-observed CDR3 segments, while those observed in only one replicate were referred to as once-observed CDR3 segments. In (**C**), only the twice-observed CDR3 segments were used for the pair-wise overlap analysis.

#### TCR repertoire analysis

The term "segment" was used to refer to functionally annotated sequences in TCR $\beta$ . Unique CDR3 segments were defined as unique CDR3 amino acid sequences. The frequency of each clone was calculated as the number of reads spanning a unique CDR3 segment over the total number of reads spanning all CDR3 segments. The abundance of each group was calculated as the total number of reads within a group divided by the total number of reads in the sample, which is equal to the sum of the frequencies of all clones in the group. Unique CDR3 segments observed in both replicates were considered as twice-observed CDR3 segments, while those observed in only one replicate were considered once-observed CDR3 segments. Only twice-observed CDR3 segments were used for the calculation of inter-sample overlap (Figs. 1C and 2A,B). Merged tissue replicates, which also contained once-observed CDR3 segments, were used to observe clones overlapping multiple tissues (Fig. 3A,B).

The sample-wise TCR repertoire overlap was determined by the "repOverlap" function in the R immunarch package using the "jaccard" and "morisita" arguments<sup>33</sup>. Graphs were generated using base R (https://www.r-project.org/) and the ggplot2 package<sup>34</sup>. The ggsignif package was used to annotate the plots with statistical significance levels. The reshape2 package was used to format data to generate heatmaps<sup>35</sup>. The viridis package was used to assign colors in the plots.

The term "motif" was used to refer to a unit of local or global similarity in CDR3 amino acid sequences observed as a k-mer. TCR $\beta$  CDR3 clustering was performed with GLIPH2 (version 0.01)<sup>36</sup> using CD8 reference sequences and the default parameters. Replicates were merged to capture more clusters with CDR3 sequence similarity. Because each patient's HLA type is unknown, tissue samples from different patients were clustered separately. If separate clusters of CDR3 sequences within a patient contained overlapping motifs, the cluster with the lowest Fisher score was retained, and the other clusters were removed. The motif frequency was calculated by dividing the total number of reads spanning an individual motif by the total number of reads in the sample.

Shannon's entropy was gauged by following equation:

Shannon' sentropy = 
$$-\sum_{i=1}^{n} p_i log_2 p_i$$

where n is the total number of unique clones in the TCR repertoire, and  $p_i$  is the frequency of one unique clone in the TCR repertoire. Shannon's entropy was used as a measure for repertoire diversity.

#### **Statistical analysis**

All statistical tests were performed in R. Wilcoxon rank-sum test and Wilcoxon signed-rank test were performed to assess significance in comparisons between groups. The Bonferroni method was applied to correct errors due to multiple comparisons. The Chi-squared test was used to calculate the *p*-value in Supplementary Table S2.

#### Target-enriched library preparation and variant calling

Genomic DNA from tumor tissues and PBMCs were sheared using covaris (Covaris), and 50–100 ng sheared genomic DNA was used for tumor variant analysis. End repair and A-tailing of sheared genomic DNA were performed using  $5 \times \text{ER/A-Tailing Enzyme Mix}$  (Enzymatics), and adaptor ligation was performed using WGS Ligase (Enzymatics). Adaptor-ligated genomic DNA was purified with  $1.2 \times \text{AMPure XP}$  beads and eluted in 20 µL nuclease-free water. PCR amplification was performed with the following steps: 3 min at 98 °C; 10 cycles of 15 s at 98 °C, 30 s at 60 °C, and 30 s at 72 °C; followed by 10 min at 72 °C. The PCR amplicons were purified using  $1.2 \times \text{AMPure XP}$  beads and analyzed using the 4150 Tapestation system (Agilent). Target enrichment was performed using the AlphaLiquid<sup>\*</sup> 100 target-capture panel, which consists of 106 cancer-related genes<sup>37</sup>. Target-enriched DNA libraries were sequenced using an Illumina NovaSeq 6000 System<sup>\*</sup> to create 150 bp paired-end reads.

Adaptor sequences and reads with low quality (<Q20) were trimmed using FASTP<sup>38</sup>. The trimmed reads were aligned to the human reference genome (hg38) using the BWA "mem" algorithm<sup>39</sup>. Then, duplicate reads were removed, and variants were called using VarDict<sup>40</sup>. Variants of tumor tissues from the same patient were combined, and germline variants were removed. For each patient, variants with allele frequency > 1% in at least one tumor tissue were retained, and other variants were removed.

#### Results

#### Similarity landscape of TCR repertoire across all gynecologic cancer patients

The median age of the study population was 57 years (range 17–86 years). Nine patients (23.1%) had a history of hypertension, and four patients (10.3%) had diabetes mellitus. One of the patients with endometrial cancer had a history of colon cancer, and another had a history of thyroid cancer. Among the patients with epithelial ovarian cancer (n = 27), most presented with high-grade serous carcinoma (n = 20, 74.1%) and stage III–IV disease (n = 25, 92.6%), and eight (29.6%) had BRCA1/2 mutation.

To analyze the TCR repertoire, we generated Illumina sequencing libraries from genomic DNA isolates of the samples by PCR reactions using primers targeting TRBV and TRBJ genes (see Methods). We assumed that unique CDR3 segment represents TCR clone since CDR3 of TCR $\beta$  is highly diverse due to VDJ recombination. In the process of filtering, clones associated with non-human epitopes were screened against a database consisting of 7,276,705 CDR3 amino acid sequences sourced from TCRdb<sup>27</sup>, McPAS-TCR<sup>28</sup>, PIRD<sup>29</sup>, TCR3d<sup>30</sup>, and VDJdb<sup>31,32</sup>. Precisely, clones were excluded based on an exact match with sequences from these databases. Non-human epitope CDR3 was removed to rule out the TCR repertoire that are unrelated to tumor antigen, such as TCR repertoires that are expanded by pathogens. Consequently, an average of 6,130.6 clonotypes, which cumulatively constituted about 29.6% frequency, were removed per sample. Among the CDR3 segments that



**Figure 2.** Overlap between TIL and PBMC TCR repertoires is patient specific. (**A**) Overlap in TCR repertoires between PBMCs and ovarian cancer tissue samples. Two patients who each underwent surgery at two different time points are included ("OV30-1st" and "OV30-2nd", "OV31-1st" and "OV31-2nd"). (**B**) Overlap in TCR repertoires between PBMCs and endometrial cancer tissue samples. For both (**A**) and (**B**), only the twice-observed CDR3 segments were utilized for the overlap analysis. (**C**) Comparison of overlap between patients and within patients. "Between patients" refers to sample pairs from different patients. "Within patients" refers to sample pairs from different time points. Wilcoxon rank-sum test, \*\*\**p*<0.005.

Scientific Reports | (2023) 13:20485 |



**Figure 3.** The PBMC TCR repertoire has high overlap with clones found in multiple tissues with metastasis. (**A**) In patients with metastatic cancer, unique tissue CDR3 segments were grouped based on the number of tissues in which they were detected. The proportion of these unique tissue CDR3 segments also found in the patients' PBMCs is represented as a percentage. (**B**) The proportions displayed in (**B**) are based on the cumulative counts of observed tissue CDR3 segments and those TIL CDR3s overlapping with PBMCs for all 9 cases, rather than averaging individual data. Merged tissue replicates refer to the union of CDR3 segments from both replicates, including once-observed CDR3 segments. In analyses (**A**) and (**B**), the TIL TCR repertoire was derived from these replicates to evaluate clonal overlap across multiple tissues. Conversely, the PBMC TCR repertoire was confined to twice-observed CDR3 segments, enhancing the reliability of comparisons with TIL TCR repertoires. This approach accounts for the diversity and clonal expansion within the PBMC repertoire, providing a comprehensive quantitative assessment.

passed the filtering process, twice-observed CDR3s showed higher frequencies compared to once-observed CDR3s (Methods, Fig. S1). The frequency of twice-observed CDR3 segment was consistent in both replicates (Fig. S2), so we thought that twice-observed CDR3 segments possibly represent expanded clones.

We assessed the pairwise similarity in TCR repertoire between all possible pairs of samples by hierarchical clustering and Jaccard index calculation<sup>41</sup> to see if there were shared clones. The Jaccard indices were zero in most cases which means no overlap is observed between different patients (Fig. 1C). This trend is consistent with previous articles<sup>42,43</sup>. We assumed that many factors, including TCR recombination<sup>44-46</sup> and variant MHC alleles<sup>47</sup>, contributed to the patient-specific nature of TCR repertoire.

#### Intra-patient similarity of TCR repertoire within multiple biopsies

While we were looking at sample-wise similarity, we found that similarity within multiple biopsies (e.g. tissues of multiple regions or PBMCs) from same patient were consistently high in many patients (Fig. 1C). Specifically, this intra-patient similarity was observed by the formation of patient-specific cluster, which was found in 6 out of the 9 PBMC samples (excluding OV12-PBMC, OV18-PBMC, EM5-PBMC) and all 31 tissue samples. To focus on intra-patient similarity, we re-organized similarity data based on each patient with metastatic cancer (Fig. 1C). The six representative insets show marked similarity constituting clusters of three or more samples. The maximum similarity of TCR repertoires occurred between tissues from the same organ within a single patient (Omentum1 and Omentum2 from patient "OV30", Jaccard index = 0.273) but similarity across metastatic organs were also high compared to inter-patient similarity.

Interestingly, a particular patient labeled "OV30", who underwent surgery twice so that two timepoints were investigated, showed similarity between samples at two different timepoints (Fig. 1C). Specifically, within the 1st timepoint samples, the average similarity was 0.153, and within the 2nd timepoint samples, the average similarity was 0.128. Importantly, between the samples of the 1st and 2nd time points, the average similarity was 0.114,

indicating a notable consistency in the tumor TCR repertoire across these timepoints. In patient labeled "OV31", who also had surgery twice showed a maintenance of previous clones from tissue and PBMC samples, which supports our observation on patient "OV30" (Fig. S3). We thought that this result can be presumably explained in two ways. One possibility is that tissue-resident memory T cells were expanded across adjacent tumor regions. The other possibility is that circulating memory lymphocytes infiltrated into tumor tissue.

In patients with ovarian cancer, we compared the number of unique CDR3 segments and Shannon's entropy to assess the level of diversity within TCR repertoires. A high value of Shannon's entropy corresponds to a high diversity in the distribution of TCR clones, which means that the TCR repertoire is likely to contain high proportion of low frequency TCR clones<sup>48</sup>. Both metrics in PBMC samples tended to be greater than that in tumor samples (Fig. S4). This is in accordance with previous results showing a relatively low number of unique CDR3 sequences in tissues compared with PBMCs<sup>49–51</sup>.

The data in Fig. 1C suggest that the overlap in TCR repertoires between PBMCs and TILs is higher within individual donors than between donors. We plotted the overlap in TCR repertoires between PBMCs and TILs (Fig. 2A,B) and observed that samples from the same patient tended to overlap more with each other than with samples from different patients (Fig. 2C). Furthermore, the results from patients "OV30" and "OV31" showed that overlap of TCR clones between PBMCs and TILs could be detected at different times within the same patients (Fig. 2A).

#### The PBMC TCR repertoire has high overlap with clones found in multiple tissue sites

We next focused on TCR clones overlapping in multiple organs. Since the similarity between samples (Figs. 1C and 2A,B) is a value calculated between two samples, there is a limitation in that it cannot explain the overlap with multiple organs as well. Thus, we pooled all of the TIL TCR clones within the same surgical case and grouped the unique TIL TCR clones by the number of tissues they are observed (Fig. 3A,B). Among the TCR clones found in multiple tissues, relatively high proportion of PBMC TCR clones were found compared to TCR clones found in single tissue. We then performed a Chi-square test to determine whether the increased proportion of PBMC TCR clones was statistically significant (Supplementary Tables S2 and S3). The resulting P values for nine cases converged to near zero, indicating that TCR clones found in multiple tissue samples are jointly detected in the PBMC TCR repertoire.

#### Multiple tissue sites share clonally abundant TCR clusters with CDR3 sequence similarity

Because TCR clones can have extensive diversity due to the complexity of combinations in their variable domains, we next focused on motifs or decomposed structural units of amino acid sequence. We used GLIPH2<sup>36</sup> to cluster the CDR3 motifs by frequency and sequence similarity within eight patients with metastatic cancer. All eight patients had unique CDR3 motifs that were shared among multiple tissues with metastasis (Fig. 4). The sample clustering based on individual motif frequencies showed that some TCR clusters were conserved among multiple tissues with metastasis, which might be the result of a tumor antigen response (Fig. S5).

To indirectly determine whether tumor antigens from multiple tissue sites within the same patient were shared, we profiled variants in tumor tissues using a target-enrichment method (Fig. S6). We observed a total



**Figure 4.** TCR motifs are shared among multiple metastatic tissues. Unique TCR motifs within eight patients with metastatic cancer were grouped. The number of tissues in which each motif occurred was then counted. For example, a total of five tissue samples in "OV30-1st" and "OV30-2nd" shared the same TCR motifs ("OV30-1st" and "OV30-2nd" were from a single patient that underwent surgery twice).

of 34 unique variants among the eight patients and found that 23 variants were shared among more than two tissues within the same patient. Although the total number of variants was not large, the proportion of variants that were shared among tissues was high. Therefore, we can infer that there was a high probability of shared neoantigens between metastasized tissues within the same patient.

#### Discussion

Gynecologic cancer is difficult to treat, because most patients are diagnosed at an advanced stage and experience recurrence despite surgery and chemotherapy<sup>52,53</sup>. Recently, various types of cancer immunotherapy have been developed, improving treatment outcomes<sup>54–56</sup>. However, the current response rate for immunotherapy in patients with gynecologic cancer is only ~  $20\%^{57,58}$ , and predictive biomarkers have been explored to improve immunotherapy efficacy and enable personalized targeted therapy. The presence of TILs in the tumor microenvironment is associated with improved patient survival. In this study, we observed TCR clones that were highly conserved among multiple sites of metastasis, suggesting the presence of tumor-specific T cells that might be harnessed for immunotherapy in patients with gynecologic cancer.

We profiled the CDR3 sequences of patients with gynecologic cancer to measure inter-sample similarity in TCR repertoires. The elucidation of the TIL CDR3 sequences and the extent to which it is shared in multiple metastatic tissues may lead to finding the novel TCR-based therapeutics. High clonal similarity between TIL TCR repertoires is supported by other studies that showed homogeneity of TCR repertoires within organs<sup>59</sup> and tissues<sup>60</sup>. Our data suggest that tumor samples from distant tissues share TCR clones, which might have expanded from similarly structured antigens.

Neoantigens shared among tumor tissues with high immunogenic potential may drive the clonal evolution of immune cells. Studies have shown that metastatic clones with the same origin share antigen variants<sup>61,62</sup>, and computationally predicted neoantigens were highly shared across metastatic sites<sup>59,63</sup>. Similarity of neoantigen pools between sites that are located closely together has not been explicitly shown, however. Hence, further study is needed to elucidate the neoantigens that are shared in tumor tissues.

We found that the TCR repertoire of PBMCs is relatively rich with clones that are shared by multiple metastases rather than clones that appear only at a single tumor site (Fig. 3A,B). The presence of tissue-shared clones in PBMCs was reported in a previous study<sup>64</sup>; however, that study contrasts with ours in that it compared TCR repertoires of tumor and non-tumor tissues with inflammation. In our understanding, recirculation of T cells is responsible for the high proportion of tissue-shared clones in blood<sup>65</sup>. However, we cannot exclude the possibility that the high proportion of tissue-shared clones might include bystander T cells<sup>66</sup>, which could potentially lessen the anti-tumorigenic potential, as we did not confirm whether the TCR sequences originated from tumor-specific T cells. Whether these clones function in tumor surveillance has not yet been studied.

We did not attempt to identify the sites where TCR repertoire expansion occurred. There remains a possibility that high-frequency TCR clones might have moved from the blood to tissues, which opposes our assumption that TIL TCR clones are generated from neoantigens in tumor tissues. In the work of Liangtao et al.<sup>67</sup>, tumor-specific TCR clones were verified by the RNA expression of proliferation-related markers. Furthermore, the fact that we did not perform experiments from Liangtao et al.<sup>67</sup> is a weakness of our study. We believe that our analysis of the TCR repertoire overlap between tissues can be supplemented with sequencing data obtained through other experimental methods.

In recent years, several studies have aimed to understand the TCR repertoire similarities among patients with different types of cancers, with some delving into multi-omics approaches to provide a broader perspective, though not placing a primary focus on TCR similarity as our study did<sup>68,69</sup>. Our research distinctly stands on the data derived from Korean patients, shedding light on the TCR repertoire in gynecologic cancers such as ovarian and endometrial cancer. This unique ethnic dimension adds a novel premise to the existing body of knowledge, as different genetic backgrounds can significantly impact the immune response and, consequently, the TCR repertoire. Moreover, unlike some studies that did not utilize replicates<sup>70–72</sup>, our approach of employing only the clones common between replicates aimed to enhance the reliability of our findings. This approach brings a higher degree of confidence in the observed TCR repertoire similarities (Fig. S2), thereby providing a more robust basis for our conclusions. While there's an overlap in the principal goal of exploring TCR repertoire similarity with other studies, our study diverges through the adoption of a methodological approach and the utilization of a unique dataset, which allowed us to focus more intensively on this aspect. Our work, therefore, contributes a unique lens through which the TCR repertoire in gynecologic cancers can be understood, and presents a stepping stone for more generalized or diversified studies in the future.

In summary, we found TCR clones that were highly conserved between tissue sites, which may be a result of shared neoantigens expressed in the tissue environment. These tissue-shared clones were enriched in the TCR repertoire of PBMCs. Further collection of data is necessary to uncover the underlying biology of the immune microenvironment of gynecologic cancer.

#### Data availability

The sequencing data for tissue and PBMC TCR repertoire samples from gynecologic cancer patients have been deposited at the Sequence Read Archive under accession number PRJNA939934.

Received: 15 April 2023; Accepted: 17 November 2023 Published online: 22 November 2023

#### References

1. Matulonis, U. et al. Ovarian cancer. Nat. Rev. Dis. Prim. 2, 1-22. https://doi.org/10.1038/nrdp.2016.61 (2016).

- 2. Siegel, R., Miller, K. & Jemal, A. Cancer statistics, 2019. CA Cancer J. Clin. 69, 7-34. https://doi.org/10.3322/caac.21551 (2019).
- Friedrich, M., Friedrich, D., Kraft, C. & Rogmans, C. Multimodal treatment of primary advanced ovarian cancer. Anticancer Res. 41, 3253–3260. https://doi.org/10.21873/anticanres.15111 (2021).
- 4. Makker, V. et al. Endometrial cancer. Nat. Rev. Dis. Prim. https://doi.org/10.1038/s41572-021-00324-8 (2021).
- Green, A., Feinberg, J. & Makker, V. A review of immune checkpoint blockade therapy in endometrial cancer. Am. Soc. Clin. Oncol. Educ. Book Am. Soc. Clin. Oncol. Annu. Meet. 40, 238–244. https://doi.org/10.1200/EDBK\_280503 (2020).
- Lheureux, S., Gourley, C., Vergote, I. & Oza, A. Epithelial ovarian cancer. Lancet (London, England) 393, 1240–1253. https://doi. org/10.1016/S0140-6736(18)32552-2 (2019).
- Testa, U., Petrucci, E., Pasquini, L., Castelli, G. & Pelosi, E. Ovarian cancers: Genetic abnormalities, tumor heterogeneity and progression, clonal evolution and cancer stem cells. *Medicines (Basel, Switzerland)* 5, 16. https://doi.org/10.3390/medicines5010016 (2018).
- Spurdle, A., Bowman, M., Shamsani, J. & Kirk, J. Endometrial cancer gene panels: Clinical diagnostic vs research germline DNA testing. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **30**, 1048–1068. https://doi.org/10.1038/modpathol.2017.20 (2017).
- Bell, D. Origins and molecular pathology of ovarian cancer. Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc 18, S19–S32. https:// doi.org/10.1038/modpathol.3800306 (2005).
- O'Hara, A. & Bell, D. The genomics and genetics of endometrial cancer. Adv. Genom. Genet. 2012, 33–47. https://doi.org/10.2147/ AGG.S28953 (2012).
- Zorn, K. et al. Gene expression profiles of serous, endometrioid, and clear cell subtypes of ovarian and endometrial cancer. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 11, 6422–6430. https://doi.org/10.1158/1078-0432.CCR-05-0508 (2005).
- Clifford, C. et al. Multi-omics in high-grade serous ovarian cancer: Biomarkers from genome to the immunome. Am. J. Reprod. Immunol. (New York, N.Y.: 1989) 80, e12975. https://doi.org/10.1111/aji.12975 (2018).
- Disis, M. et al. Society for immunotherapy of cancer (SITČ) clinical practice guideline on immunotherapy for the treatment of gynecologic cancer. J. Immunother. Cancer 11, e006624. https://doi.org/10.1136/jitc-2022-006624 (2023).
- Oaknin, A. *et al.* Clinical activity and safety of the anti-programmed death 1 monoclonal antibody dostarlimab for patients with recurrent or advanced mismatch repair-deficient endometrial cancer: A nonrandomized phase 1 clinical trial. *JAMA Oncol.* 6, 1766–1772. https://doi.org/10.1001/jamaoncol.2020.4515 (2020).
- 15. FDA. FDA Approves Immunotherapy for Endometrial Cancer with Specific Biomarker. https://www.fda.gov/news-events/press-annou ncements/fda-approves-immunotherapy-endometrial-cancer-specific-biomarker (2021).
- FDA. FDA Approves Dostarlimab-Gxly with Chemotherapy for Endometrial Cancer. https://www.fda.gov/drugs/drug-approvalsand-databases/fda-approves-dostarlimab-gxly-chemotherapy-endometrial-cancer (2023).
- Mirza, M. et al. Dostarlimab for primary advanced or recurrent endometrial cancer. N. Engl. J. Med. 388, 2145–2158. https://doi. org/10.1056/NEJMoa2216334 (2023).
- FDA. FDA Approves Pembrolizumab Combination for the First-Line Treatment of Cervical Cancer. https://www.fda.gov/drugs/resou rces-information-approved-drugs/fda-approves-pembrolizumab-combination-first-line-treatment-cervical-cancer (2021).
- Pothuri, B. et al. Practice changing cervical cancer clinical trials. Gynecol. Oncol. 165, 410–412. https://doi.org/10.1016/j.ygyno. 2022.03.004 (2022).
- 20. Bagchi, S., Yuan, R. & Engleman, E. Immune checkpoint inhibitors for the treatment of cancer: Clinical impact and mechanisms of response and resistance. *Annu. Rev. Pathol.* 16, 223–249. https://doi.org/10.1146/annurev-pathol-042020-042741 (2021).
- Wan, C. et al. Enhanced efficacy of simultaneous PD-1 and PD-L1 immune checkpoint blockade in high-grade serous ovarian cancer. Cancer Res. 81, 158–173. https://doi.org/10.1158/0008-5472.CAN-20-1674 (2021).
- Shen, S. *et al.* Development and validation of an immune gene-set based prognostic signature in ovarian cancer. *EBioMedicine* 40, 318–326. https://doi.org/10.1016/j.ebiom.2018.12.054 (2019).
- Lee, S. et al. Distinct T cell receptor repertoire diversity of clinically defined high-grade serous ovarian cancer treatment subgroups. iScience 24, 102053. https://doi.org/10.1016/j.isci.2021.102053 (2021).
- Gotoh, O. et al. Immunogenomic landscape of gynecologic carcinosarcoma. Gynecol. Oncol. 160, 547–556. https://doi.org/10. 1016/j.ygyno.2020.11.030 (2021).
- Jang, M. et al. Characterization of T cell repertoire of blood, tumor, and ascites in ovarian cancer patients using next generation sequencing. Oncoimmunology 4, e1030561. https://doi.org/10.1080/2162402X.2015.1030561 (2015).
- Bolotin, D. et al. MiXCR: Software for comprehensive adaptive immunity profiling. Nat. Methods 12, 380–381. https://doi.org/10. 1038/nmeth.3364 (2015).
- Armstrong, D. K. et al. Ovarian cancer, version 2.2020, NCCN clinical practice guidelines in oncology. J. Natl. Compr. Cancer Netw. JNCCN 19, 191–226. https://doi.org/10.6004/jnccn.2021.0007 (2021).
- 28. Tickotsky, N., Sagiv, T., Prilusky, J., Shifrut, E. & Friedman, N. McPAS-TCR: A manually curated catalogue of pathology-associated
- T cell receptor sequences. *Bioinformatics (Oxford, England)* 33, 2924–2929. https://doi.org/10.1093/bioinformatics/btx286 (2017).
  29. Zhang, W. *et al.* PIRD: Pan immune repertoire database. *Bioinformatics (Oxford, England)* 36, 897–903. https://doi.org/10.1093/bioinformatics/btz614 (2020).
- Gowthaman, R. & Pierce, B. TCR3d: The T cell receptor structural repertoire database. *Bioinformatics (Oxford, England)* 35, 5323–5325. https://doi.org/10.1093/bioinformatics/btz517 (2019).
- Shugay, M. et al. VDJdb: A curated database of T-cell receptor sequences with known antigen specificity. Nucleic Acids Res. 46, D419–D427. https://doi.org/10.1093/nar/gkx760 (2018).
- Bagaev, D. et al. VDJdb in 2019: Database extension, new analysis infrastructure and a T-cell receptor motif compendium. Nucleic Acids Res. 48, D1057–D1062. https://doi.org/10.1093/nar/gkz874 (2020).
- Vadim, N., immunarch.bot & Eugene, R. immunomind/immunarch: 0.6.5: Basic single-cell support. https://doi.org/10.5281/zenodo. 3893991 (2022).
- 34. Wickham, H. ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag, 2016).
- 35. Wickham, H. Reshaping data with the reshape package. J. Stat. Softw. 21, 1–20. https://doi.org/10.18637/jss.v021.i12 (2007).
- Huang, H., Wang, C., Rubelt, F., Scriba, T. & Davis, M. Analyzing the mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat. Biotechnol.* 38, 1194–1202. https://doi.org/10.1038/ s41587-020-0505-4 (2020).
- Kim, S. *et al.* Dynamic changes in longitudinal circulating tumour DNA profile during metastatic colorectal cancer treatment. *Br. J. Cancer* 127, 898–907. https://doi.org/10.1038/s41416-022-01837-z (2022).
- Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)* 34, i884–i890. https://doi.org/10.1093/bioinformatics/bty560 (2018).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 (2009).
- Lai, Z. et al. VarDict: A novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res. 44, e108–e108. https://doi.org/10.1093/nar/gkw227 (2016).
- Chung, N., Miasojedow, B., Startek, M. & Gambin, A. Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. *BMC Bioinformatics* 20, 1–11. https://doi.org/10.1186/s12859-019-3118-5 (2019).
- Teng, Y. et al. Analysis of T cell receptor clonotypes in tumor microenvironment identifies shared cancer-type-specific signatures. Cancer Immunol. Immunother.: CII 71, 1–10. https://doi.org/10.1007/s00262-021-03047-7 (2022).

- Conde, C. et al. Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science (New York, N.Y.) 376, eabl5197. https://doi.org/10.1126/science.abl5197 (2022).
- Chien, Y., Gascoigne, N., Kavaler, J., Lee, N. & Davis, M. Somatic recombination in a murine T-cell receptor gene. Nature 309, 322–326. https://doi.org/10.1038/309322a0 (1984).
- Burtrum, D., Kim, S., Dudley, E., Hayday, A. & Petrie, H. TCR gene recombination and alpha beta-gamma delta lineage divergence: Productive TCR-beta rearrangement is neither exclusive nor preclusive of gamma delta cell development. J. Immunol. (Baltimore, Md.: 1950) 157, 4293–4296 (1996).
- Dupic, T., Marcou, Q., Walczak, A. & Mora, T. Genesis of the αβ T-cell receptor. PLoS Comput. Biol. 15, e1006874. https://doi.org/ 10.1371/journal.pcbi.1006874 (2019).
- Logunova, N. et al. MHC-II alleles shape the CDR3 repertoires of conventional and regulatory naïve CD4+T cells. Proc. Natl. Acad. Sci. U. S. A. 117, 13659–13669. https://doi.org/10.1073/pnas.2003170117 (2020).
- Li, N., Yuan, J., Tian, W., Meng, L. & Liu, Y. T-cell receptor repertoire analysis for the diagnosis and treatment of solid tumor: A methodology and clinical applications. *Cancer Commun. (London, England)* 40, 473–483. https://doi.org/10.1002/cac2.12074 (2020).
- 49. Wang, Y. *et al.* T cell receptor beta-chain profiling of tumor tissue, peripheral blood and regional lymph nodes from patients with papillary thyroid carcinoma. *Front. Immunol.* **12**, 595355. https://doi.org/10.3389/fimmu.2021.595355 (2021).
- Reuben, A. et al. Comprehensive T cell repertoire characterization of non-small cell lung cancer. Nat. Commun. 11, 603. https:// doi.org/10.1038/s41467-019-14273-0 (2020).
- Tsuji, T. et al. Clonality and antigen-specific responses shape the prognostic effects of tumor-infiltrating T cells in ovarian cancer. Oncotarget 11, 2669. https://doi.org/10.18632/oncotarget.27666 (2020).
- Bowtell, D. et al. Rethinking ovarian cancer II: Reducing mortality from high-grade serous ovarian cancer. Nat. Rev. Cancer 15, 668–679. https://doi.org/10.1038/nrc4019 (2015).
- Allgar, V. & Neal, R. Delays in the diagnosis of six cancers: analysis of data from the national survey of NHS patients: Cancer. Br. J. Cancer 92, 1959–1970. https://doi.org/10.1038/sj.bjc.6602587 (2005).
- Rusquec, P., Calbiac, O., Robert, M., Campone, M. & Frenel, J. Clinical utility of pembrolizumab in the management of advanced solid tumors: An evidence-based review on the emerging new data. *Cancer Manag. Res.* 11, 4297–4312. https://doi.org/10.2147/ CMAR.S151023 (2019).
- 55. Santin, A. *et al.* Regression of chemotherapy-resistant polymerase ε (POLE) ultra-mutated and MSH6 hyper-mutated endometrial tumors with nivolumab. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 22, 5682–5687. https://doi.org/10.1158/1078-0432.CCR-16-1031 (2016).
- Lynam, S., Lugade, A. & Odunsi, K. Immunotherapy for gynecologic cancer: current applications and future directions. *Clin. Obstet. Gynecol.* 63, 48. https://doi.org/10.1097/GRE.00000000000513 (2020).
- Chardin, L. & Leary, A. Immunotherapy in ovarian cancer: Thinking beyond PD-1/PD-L1. Front. Oncol. 11, 795547. https://doi. org/10.3389/fonc.2021.795547 (2021).
- Nishio, H., Iwata, T. & Aoki, D. Current status of cancer immunotherapy for gynecologic malignancies. *Jpn. J. Clin. Oncol.* 51, 167–172. https://doi.org/10.1093/jjco/hyaa214 (2021).
- Mattos-Arruda, L. et al. The genomic and immune landscapes of lethal metastatic breast cancer. Cell Rep. 27, 2690–2708. https:// doi.org/10.1016/j.celrep.2019.04.098 (2019).
- Emerson, R. et al. High-throughput sequencing of T-cell receptors reveals a homogeneous repertoire of tumour-infiltrating lymphocytes in ovarian cancer. J. Pathol. 231, 433–440. https://doi.org/10.1002/path.4260 (2013).
- Avigdor, B. *et al.* Mutational profiles of breast cancer metastases from a rapid autopsy series reveal multiple evolutionary trajectories. *JCI Insight* 2, e96896. https://doi.org/10.1172/jci.insight.96896 (2017).
- Brown, D. *et al.* Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations. *Nat. Commun.* 8, 14944. https://doi.org/10.1038/ncomms14944 (2017).
- Jiang, T. et al. Heterogeneity of neoantigen landscape between primary lesions and their matched metastases in lung cancer. Transl. Lung Cancer Res. 9, 246. https://doi.org/10.21037/tlcr.2020.03.03 (2020).
- Läubli, H. et al. The T cell repertoire in tumors overlaps with pulmonary inflammatory lesions in patients treated with checkpoint inhibitors. Oncoimmunology 7, e1386362. https://doi.org/10.1080/2162402X.2017.1386362 (2017).
- Masopust, D. & Schenkel, J. The integration of T cell migration, differentiation and function. Nat. Rev. Immunol. 13, 309–320. https://doi.org/10.1038/nri3442 (2013).
- Simoni, Y. et al. Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. Nature 557, 575–579. https://doi.org/10.1038/s41586-018-0130-2 (2018).
- Zheng, L. et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. Science (New York, N.Y.) 374, abe6474. https://doi. org/10.1126/science.abe6474 (2021).
- Yang, B. et al. Spatial heterogeneity of infiltrating T cells in high-grade serous ovarian cancer revealed by multi-omics analysis. Cell Rep. Med. https://doi.org/10.1016/j.xcrm.2022.100856 (2022).
- Zhang, A. *et al.* Interfaces of malignant and immunologic clonal dynamics in ovarian cancer. *Cell* 173, 1755–1769. https://doi.org/ 10.1016/j.cell.2018.03.073 (2018).
- Shi, L. *et al.* Multi-omics study revealing the complexity and spatial heterogeneity of tumor-infiltrating lymphocytes in primary liver carcinoma. *Oncotarget* 8, 34844. https://doi.org/10.18632/oncotarget.16758 (2017).
- Reuben, A. *et al.* TCR repertoire intratumor heterogeneity in localized lung adenocarcinomas: An association with predicted neoantigen heterogeneity and postsurgical recurrence. *Cancer Discov.* 7, 1088–1097. https://doi.org/10.1158/2159-8290.CD-17-0256 (2017).
- Chen, R. et al. Evolution of genomic and T-cell repertoire heterogeneity of malignant pleural mesothelioma under dasatinib treatment. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 26, 5477–5486. https://doi.org/10.1158/1078-0432.CCR-20-1767 (2020).

#### Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2021R1A2C2008490); the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare, Republic of Korea (HI18C2282, HI14C1277); "Hankookilbo Myung-Ho Seung" Faculty Research Assistance Program of Yonsei University College of Medicine (6-2021-0102).

#### Author contributions

J.P., J.H.L, J.-Y.L. and D.B. conceived and designed the study. T.K., H.L., S.J. and D.L. performed the experiments and analyzed the data. T.K. wrote the paper.

### **Competing interests**

The authors declare no competing interests.

#### Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/ 10.1038/s41598-023-47740-2.

Correspondence and requests for materials should be addressed to J.-Y.L. or D.B.

Reprints and permissions information is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2023