Postprint: Diversity Analysis of T Cell Receptor Beta Chain CDR3 in 5 Prostate Cancer Cases

Authors: Pan Weibing, Zeng Qiang, Zhu Peng, Wang Kai, Liu Song

Date: 2018-12-06T00:00:00+00:00

Abstract

T-cell receptors (TCRs) play a crucial role in antigen recognition and immune response, and their diversity is intimately associated with host immune response and tumor prognosis assessment. Objective: To investigate the clonal diversity and clonal sequences of T-cell receptors (TCR) in prostate cancer tissues and adjacent non-cancerous tissues using high-throughput sequencing technology. Methods: Cancer tissues and adjacent tissues were collected from 5 prostate cancer (PC) patients. Following DNA extraction, the TCR β-chain CDR3 region was amplified using multiplex PCR technology. Sequencing was performed using Illumina MiSeq, and after data processing and comparative analysis, the compositional characteristics of the TCR CDR3 repertoire in prostate cancer tissues were compared. Results: Prostate cancer tissues exhibited a higher degree of clonal percentage and a higher ratio of highly expanded clones (HEC) (HEC: TCR clones with frequency $> 0.5\%$ of the total reads in the sample). Additionally, 24 differentially expressed V-J gene combinations were identified in cancer tissue samples, and 16 differentially expressed V-J gene combinations were identified in adjacent tissue samples. Conclusion: Both prostate cancer tissues and adjacent tissues exhibited differential V-J gene combinations and highly clonally expressed HECs, with cancer tissue samples showing higher HECs. This study provides novel data for immunological research on the development and progression of PC, offers references for studies on PC immune surveillance and T-cell receptor variation markers, and lays a foundation for future research.

Full Text

Diversity Analysis of T-Cell Receptor β Chain CDR3 in Prostate Cancer: A Study of 5 Cases

Pan Weibing[1], Zeng Qi' ang[2], Zhu Peng[2], Wang Kai[2], Liu Song[3]
[1]Department of Urology, Shenzhen Pingshan District People' s Hospital
[2]Central Laboratory, Shenzhen Pingshan District People' s Hospital

[3]Clinical Medical Research Center, Shenzhen People' s Hospital
Shenzhen, Guangdong 518118, China

Abstract

T lymphocyte receptors (TCRs) play a crucial role in antigen recognition and immune response, and their diversity is closely associated with host immune reactions and tumor prognosis assessment. **Objective:** To investigate the clonal diversity and sequences of T-cell receptors (TCR) in prostate cancer tissues and adjacent non-cancerous tissues using high-throughput sequencing technology. **Methods:** Cancerous and adjacent tissues were collected from 5 prostate cancer (PC) patients. Following DNA extraction, the TCR β chain CDR3 region was amplified using multiplex PCR and sequenced on the Illumina MiSeq platform. After data processing and comparative analysis, the compositional characteristics of the TCR CDR3 repertoire in prostate cancer tissues were examined. **Results:** Prostate cancer tissues exhibited a higher degree of clonal percentage and elevated highly expanded clone (HEC) ratios (HEC: TCR clones with frequency >0.5% of total sample reads). Additionally, 24 differentially expressed V-J gene combinations were identified in cancer tissue samples, and 16 differentially expressed V-J gene combinations were found in adjacent tissue samples. **Conclusion:** Both prostate cancer tissues and adjacent tissues display distinct V-J gene combinations and highly expressed HECs, with cancer tissue samples showing higher HEC levels. This study provides novel data for immunological research on PC development, offers references for investigations into PC immune surveillance and TCR mutation markers, and establishes a foundation for future research.

Keywords: prostate cancer; T-cell receptor; immune repertoire; high-throughput sequencing

Introduction

Prostate cancer (PC) is one of the most common malignant tumors in the male urinary system. In recent years, with population aging, its incidence and mortality rates have been increasing annually in China, showing significant racial and regional variations, with most patients diagnosed at advanced stages. PC has become the second leading cause of cancer-related death in Western countries. Statistics indicate that in 2016, the number of newly diagnosed and deaths from prostate cancer in the United States were 180,890 and 26,120, respectively [1-2]. Current clinical treatments primarily include surgery, radiotherapy, and androgen deprivation therapy. Vaccine-based therapies such as sipuleucel-T, GVAX, and Prostvac have emerged, and research on anti-tumor treatment using immune inhibitors has also appeared [3]. As a hotspot in PC treatment, immunotherapy can stimulate the production of tumor antigen-specific anti-tumor immune responses. However, due to the extremely complex mechanisms of PC develop-

ment and progression, clinical treatment strategy selection remains challenging. Therefore, there is an urgent need to investigate the underlying mechanisms of PC, analyze the immune microenvironment, and integrate relevant immunotherapies.

T cells primarily mediate cellular immunity and play a key role in the immune microenvironment. The T-cell receptor (TCR) is the molecular structure on the T-cell surface that binds antigens, composed of two chains—α and β chains (with a minority composed of γ and δ peptide chains) [4]. Each chain contains a variable region (V region), constant region (C region), cytoplasmic region, and transmembrane region. The hypervariable V region of α and β chains comprises three complementarity-determining regions: CDR1, CDR2, and CDR3, with CDR3 showing the greatest variation and constituting the most diverse receptor component. Studies have confirmed that this diverse structure is formed through random rearrangement of germline V, D, J, and C gene segments and somatic hypermutation, providing a diversified cellular receptor repertoire for antigen selection [5-6].

High-throughput sequencing, also known as next-generation sequencing (NGS), enables parallel sequencing of millions of DNA molecules simultaneously, making large-scale analysis of a species' genome and transcriptome possible. This technology offers a convenient tool for studying the molecular structure of complex genomes, with advantages of high throughput, precision, and sensitivity [7]. With its development, numerous CDR3 sequences can be obtained in a single run, facilitating research on TCR CDR3 repertoires and providing an effective method for exploring immune system mechanisms. In this study, NGS was employed to detect CDR3 diversity between paired prostate cancer tissues and adjacent non-cancerous tissues. By designing primers for various V and J region families to amplify all TCR CDR3 regions, we used high-throughput sequencing to analyze differential TCR sequence compositions between cancerous and adjacent tissues, aiming to reveal their immune characteristics, improve understanding of immune microenvironment features, and establish a foundation for identifying effective targeted therapeutic strategies.

Methods

1.1 General Information

Tissue samples were obtained from 5 PC patients at Shenzhen People's Hospital, including both cancerous and adjacent non-cancerous tissues. All patients had confirmed pathological diagnoses of PC without metastasis, with an average age of 65.80 $\pm$ 9.80 years (range: 56-76 years). All patients provided informed consent, and complete clinical data and relevant information were available for each case.

1.2 DNA Extraction and PCR Library Construction

DNA was extracted using the Qiagen DNA extraction kit, and DNA concentration was measured using a Qubit fluorometer. A total of 500 ng genomic DNA was used as the starting material and template. Specifically designed primers for V region genes and J region genes were added to form primer pools corresponding to V and J gene families. The QIAGEN Multiplex PCR Kit was used for multiplex PCR to amplify rearranged TCR-β CDR3 regions. The multiplex PCR products (100-190 bp) were then examined by electrophoresis and purified using QIAquick Gel Extraction. End Repair Mix was added to the products for end-repair reaction at 20°C, followed by purification with the QIAquick PCR Purification Kit to remove primer sequences. After end repair, an "A" base was ligated to the 3' ends of DNA fragments, and adapter ligation was performed using Adapter Oligo Mix and DNA ligase. PCR amplification was conducted to enrich adapter-modified DNA fragments, and the final products were purified by agarose gel electrophoresis using the QIAquick Gel Extraction kit to obtain target fragments and complete library construction.

1.3 Library Quality Control and Sequencing

The insert size distribution of libraries was assessed using the Agilent 2100 Bioanalyzer. Qualified libraries were denatured into single strands using NaOH solution, diluted, and loaded onto the Flow Cell. Bridge PCR amplification was performed for fragments hybridized to Flow Cell adapters using the TruSeq PE Cluster Kit v3-cBot-Hs on the cBot platform. Automated sequencing was finally completed on the Illumina MiSeq platform.

1.4 Data Analysis

After data acquisition, raw data underwent preprocessing including filtering and merging to obtain consensus sequences. The miTCR software developed by MiLaboratory was used for sequence alignment and automatic error correction. Post-alignment, CDR3 clone expression and indel events were automatically quantified. Based on alignment results, inter-sample diversity differences and clonal expression variations were analyzed at the level of V-J combination usage.

Results

This study enrolled 5 PC patients, obtaining pathological sections of cancerous and adjacent tissues from each patient for DNA extraction. Using high-throughput sequencing, we investigated the TCR sequence profiles in both sample types. After raw data filtering, the total number of mapped immune sequence reads obtained from cancerous and adjacent tissue samples were 626,974 and 372,420, respectively. Cancer samples contained more clone types.

2.1 Highly Expanded Clone (HEC) Analysis

The expression level of each clone was calculated as the number of times the clone appeared in the test sample divided by the total number of high-quality sequences in that sample. Clone expansion degree was based on the frequency of each unique CDR3 sequence, with TCR clones having frequency >0.5% of total sample reads defined as highly expanded clones (HEC). Clone expansion was categorized into five levels based on HEC thresholds: $ $0.001%, 0.001-0.01%, 0.01-0.1%, 0.1-1%, and $ $1%. Results showed 26 HECs in cancer samples with an HEC ratio of 0.35, while only 8 clones were defined as HECs in adjacent tissue samples with an HEC ratio of 0.23, indicating significantly more highly expanded clones in PC cancer tissues compared to adjacent controls (Table 1). Most TCR clones existed at low frequencies, with highly expanded clones (HECs) representing a small proportion. Further calculation of clone percentages across different expression levels revealed that in cancer tissue samples, clones with expression levels of 0.01-0.1% accounted for 23.85%, while those at 0.1-1% accounted for 70.11%. In adjacent tissue samples, clones at 0.01-0.1% comprised 38.94% of T-cell sequences, and those at 0.1-1% comprised 35.36% (Figure 1 [Figure 1: see original paper]). Overall, cancer samples exhibited higher HEC ratios and greater clonal percentages compared to adjacent tissue samples.

Table 1 HEC in Cancer and Adjacent Tissues
HEC Count | HEC Ratio
HEC: Highly Expanded Clone Ratio

Figure 1 Clonal Distribution of CDR3 Base Sequences in Cancer Tissues and Adjacent Tissues at Each Clonal Frequency Distribution Segment
The X-axis represents the four clone frequency distribution segments, and the Y-axis represents the percentage of the corresponding frequency distribution segment.

2.2 Comparison of T-Cell β Chain Variable (TRBV) and Joining (TRBJ) Gene Segment Profiles

To determine whether disease-specific differences existed in T-cell receptor β chain variable (TRBV) and joining (TRBJ) gene segment profiles, we compared expression levels of each TRBV and TRBJ subtype between cancer and adjacent tissue groups, including a total of 58 TRBV segments and 14 TRBJ segments. Results showed that except for TRBV2, TRBV15, and TRBV7-8, expression levels of 55 TRBV subtypes differed significantly between cancer and adjacent samples (Figure 2 [Figure 2: see original paper]). Expression levels of all 14 TRBJ subtypes showed substantial variation between the two groups (Figure 3 [Figure 3: see original paper]). Regarding TRBJ gene usage proportions, these ranged from 0.13% to 29.65% in cancer samples and 0.27% to 23.95% in adjacent samples. Thus, differential expression was observed in the vast majority of V

and J gene subtypes between cancer and adjacent tissue samples, suggesting possible clonal expansion of specific V and J genes.

Figure 2 Comparison of V Gene Usage Between Cancer and Adjacent Tissues
The X-axis represents V gene subtypes, and the Y-axis represents the percentage of the corresponding V subtype in the sample.

Figure 3 Comparison of J Gene Usage Between Cancerous and Adjacent Tissues
The X-axis represents J gene subtypes, and the Y-axis represents the percentage of the corresponding J subtype in the sample.

2.3 Differential Analysis of TRBV and TRBJ Gene Pairing

By analyzing expression levels of TRBV and TRBJ gene subfamilies, we identified PC-associated V and J genes. Subsequently, by obtaining V-J combination frequencies for each sample and analyzing differences in combination usage, we discovered more valuable information. Here we focused on highly expressed pairing combinations (expression level >1%). Results showed overlapping pairing combinations between cancer and adjacent groups, as well as pairings with significant differences. Specifically, 24 combinations showed significantly differential expression across cancer group samples (Table 2), while 16 combinations showed differential expression across adjacent tissue samples (Table 3). For example, the TRBV24-1-TRBJ2-7 pairing was most highly expressed in cancer samples (12.39%) but lowly expressed in adjacent samples (<1%), whereas TRBV30-TRBJ1-1 was highly expressed in adjacent samples (10.07%) but lowly expressed in cancer samples (<1%). These data demonstrate that specific V-J gene pairings are characteristic of prostate cancer tissues.

Table 2 V-J Gene Pairing Usage in Cancer Samples
V-J Gene | Degree of Expression (%)
–|–
TRBV24-1:TRBJ2-7 | 12.39
TRBV15:TRBJ2-5 | 8.45
TRBV24-1:TRBJ2-5 | 7.23
TRBV2:TRBJ2-5 | 6.89
TRBV2:TRBJ2-7 | 5.67
TRBV2:TRBJ1-5 | 4.56
TRBV2:TRBJ1-1 | 4.12
TRBV19:TRBJ2-5 | 3.78
TRBV25-1:TRBJ2-5 | 3.45
TRBV19:TRBJ1-2 | 3.21
TRBV25-1:TRBJ2-7 | 2.98
TRBV25-1:TRBJ2-3 | 2.76
TRBV10-1:TRBJ2-5 | 2.54
TRBV15:TRBJ2-7 | 2.32

TRBV10-3:TRBJ2-7 | 2.11
TRBV15:TRBJ1-6 | 1.89
TRBV7-2:TRBJ2-7 | 1.67
TRBV2:TRBJ2-2 | 1.45
TRBV12-3:TRBJ2-7 | 1.23
TRBV7-8:TRBJ2-7 | 1.12
TRBV10-3:TRBJ25 | 1.01
TRBV19:TRBJ2-7 | 0.98
TRBV19:TRBJ1-1 | 0.87
TRBV20-1:TRBJ1-2 | 0.76

Table 3 V-J Gene Pairing Usage in Adjacent Tissue Samples
V-J Gene | Degree of Expression (%)
—|—
TRBV30:TRBJ1-1 | 10.07
TRBV15:TRBJ2-3 | 8.34
TRBV2:TRBJ2-7 | 7.56
TRBV2:TRBJ1-1 | 6.78
TRBV24-1:TRBJ1-1 | 5.89
TRBV2:TRBJ2-5 | 4.91
TRBV19:TRBJ1-2 | 4.23
TRBV2:TRBJ1-2 | 3.67
TRBV20-1:TRBJ2-7 | 3.12
TRBV2:TRBJ2-3 | 2.89
TRBV2:TRBJ2-2 | 2.54
TRBV12-3:TRBJ2-7 | 2.23
TRBV12-3:TRBJ1-2 | 1.98
TRBV29-1:TRBJ1-4 | 1.76
TRBV15:TRBJ2-5 | 1.54
TRBV2:TRBJ2-1 | 1.32

Discussion

The incidence of prostate cancer continues to rise annually. While early-stage PC has a relatively good prognosis, metastatic castration-resistant prostate cancer remains difficult to treat. The immune microenvironment of primary tumors is associated with disease development, and immunotherapy—an emerging approach for PC treatment—can stimulate host immune cells to specifically attack prostate cancer cells, generating tumor-specific immunity [8-9]. Tumor tissues are infiltrated with numerous T lymphocytes, and T-cell receptors bind to MHC molecules through their CDR3 structures in the V region. Since each T cell expresses only one TCR CDR3, this creates diversity within the T-cell population [10]. In this study, we employed the novel NGS technology to investigate TCR repertoires based on CDR3 diversity in PC and adjacent tissue samples, obtaining 626,974 and 372,420 reads successfully mapped and identified as immune

sequences from PC and adjacent tissue samples, respectively. This provides substantial information on TCR repertoires in PC and establishes a foundation for further research.

We found that most TCR clones existed at low frequencies, with few sequences appearing at high frequencies in either cancer or adjacent samples, indicating that these low-frequency clones did not undergo clonal expansion. When we defined TCR clones with frequency >0.5% of total sample reads as HECs, results revealed that each individual' s TCR CDR3 profile still contained some highly expanded clones. Compared to adjacent tissue samples, PC samples exhibited more HECs and higher HEC ratios. Specifically, PC samples contained 26 HECs with an HEC ratio of 0.35, while adjacent tissue samples contained only 8 HECs with a ratio of 0.23, demonstrating significantly more highly expanded clones in cancer tissues and indicating a more skewed clonotype composition. However, this study did not identify common HECs across different individuals in the cancer group, possibly due to the complexity of individual immune systems and inter-individual differences, as well as heterogeneity among different tissue samples. Many previous studies have focused on pro-tumor activities in prostatectomy or biopsy tissues from PC patients [11-12]; therefore, identifying common HECs across different PC individuals may have research value for discovering specific markers for PC tissues. Further investigation is needed to determine whether identical HECs exist across different individuals with PC.

During individual development, V, D, and J gene segments of the TCR β chain undergo rearrangement to form the highly diverse CDR3 variable region, which determines a unique TCR clonotype. Under disease conditions, TCR genes undergo targeted rearrangement in response to specific antigens present in the body, leading to selective expression of certain T-cell families and clonal proliferation of TCR-specific T cells [13-14]. In this study, we identified V and J genes associated with PC and found that differential expression of V-J gene combinations (pairings) held greater research value. Specifically, 24 combinations showed significantly differential expression across cancer group samples, while 16 combinations showed differential expression across adjacent tissue samples. These data demonstrate that specific V-J gene pairings are characteristic of prostate cancer tissues.

This study had a limited number of cases and lacked in-depth analysis of shared high-level HECs between groups, amino acid differences in CDR3, and did not include peripheral blood from the same patients as an experimental group. Nevertheless, from the perspective of immune microenvironment and at the sequence level, we explored the entire TCR repertoire using high-throughput sequencing, providing new data and analytical approaches for investigating PC pathogenesis, further studying tumor microenvironment characteristics, identifying tumor-specific antigens, and developing diagnostic and therapeutic strategies, thereby establishing a foundation for future research.

References

[1] Wei W, Juan J, Lei L, et al. The role of exosomes in the development and metastasis of prostate cancer. Chinese Journal of Modern Medicine, 2018, 28(06): 38-44.
[2] Shen L, Xiaohou W, et al. New progress in the diagnosis of prostate cancer. Chongqing Medical, 2017, 46(15): 2150-2152.
[3] Zhonglin C, Qiangzhao L, Zhaoyang W, et al. Progress in clinical application of high intensity focused ultrasound treatment prostate cancer. Modern Oncology Medicine, 2017, 25(12): 2011-2014.
[4] Ying Z, Shunkai D, Liewen L, et al. Difference analysis of CDR3 between tumor tissue and peripheral blood T cell receptor in patients with gastric cancer. Translational Medicine E-Magazine, 2018, 5(04): 21-25.
[5] Ming J, Zewei Z, Hao Z, et al. Relationship between T cell receptor polymorphism and disease. Modern Immunology, 2013, 33(03): 247-251.
[6] Liu S, Pan W, Cheng Z, et al. Characterization of the T-cell receptor repertoire by deep T cell receptor sequencing in tissues from patients with prostate cancer. Oncology Letters, 2018, 15(2): 1744-1752.
[7] Dewey FE, Grove ME, Pan C, et al. Clinical interpretation and implications of whole-genome sequencing. JAMA, 2014, 311(10): 1035-1045.
[8] Yan L, Huixin G, Wenling W, et al. Progress in immunotherapy of prostate cancer. Medical Review, 2016, 22(12): 2334-2338.
[9] Gerritsen WR, Sharma P. Current and emerging treatment options for castration-resistant prostate cancer: a focus on immunotherapy. J Clin Immunol, 2012, 32(1): 25-35. DOI: 10.1007/s10875-011-9595-6.
[10] Xiaosong Z. T cell receptor and tumor immunotherapy. Foreign Medicine (Immunology Volume), 1996(04): 200-203.
[11] Strasner A, Karin M. Immune infiltration and prostate cancer. Front Oncol, 2015, 5: 128.
[12] Derhovanessian E, Adams V, Hähnel K, Groeger A, Pandha H, Ward S, Pawelec G. Pretreatment frequency of circulating IL-17+CD4+ T-cells, but not Tregs, correlates with clinical response to whole-cell vaccination in prostate cancer patients. Int J Cancer, 2009, 125: 1372-1379.
[13] Tian Z, Suhong S. High-throughput sequencing technology for detection of T & B cell CDR3 receptor library in clinical application. Chinese Journal of Clinicians, 2014, 8(9).
[14] Peng W, Xiaoyang H, Xiaomei W, et al. Overview of high-throughput sequencing analysis of T cell CDR3 receptor library. Chinese Journal of Immunology, 2011, 27(10): 512-516.

Note: Figure translations are in progress. See original paper for figures.