Immunogenomics: A Negative Prostate Cancer Outcome Associated with TcR- γ/δ Recombinations

Yaping N. Tu¹ · Wei Lue Tong¹ · John M. Yavorski¹ · George Blanck¹

Received: 6 November 2017 / Accepted: 2 January 2018 / Published online: 22 January 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract

We developed a scripted algorithm, based on previous, earlier editions of the algorithm, to mine prostate cancer exome files for T-cell receptor (TcR) recombination reads: Reads representing TcR gene recombinations were identified in 497 prostate cancer exome files from the cancer genome atlas (TCGA). As has been reported for melanoma, co-detection of productive TcR- α and TcR- β recombination reads correlated with an RNA expression signature representing T-cell exhaustion, particularly with high RNA levels for PD-1 and PD-L1, in comparison to several different control sets of samples. Co-detection of TcR- α and TcR- β recombination reads also correlated with high level expression of genes representing antigen presenting functions, further supporting the conclusion that co-detection of TcR- α and TcR- β recombination reads represents an immunologically relevant microenvironment. Finally, detection of unproductive TcR- δ recombinations, and unproductive and productive TcR- γ recombinations, strongly correlated with, and may represent a convenient biomarker for a poor clinical outcome. These results underscore the value of the genomics-based assessment of unproductive TcR recombinations and raise questions about the impact of tumor microenvironment lymphocytes in the absence of antigenicity.

Keywords T-cell receptor recombinations · Prostate adenocarcinoma · Tumor specimen exomes · Immune checkpoint proteins · Antigen processing · The cancer genome atlas

Introduction

Much of cancer biology and cancer therapy has moved in the direction of the anti-tumor immune response. Additionally, cancer biology has advanced with big data based immune characterizations of cancer specimens. For example, tumor specimen RNASeq, exome (WXS), and whole genome sequence files have been assessed for the presence of immune receptor recombinations, and in the case of RNASeq files in particular, for expression signatures of immune activity or lack of immune activity. However, little if any of the genomics-based approaches to immunoscore have been integrated into

clinical trials or standard of care. Several genomics-based approaches have indicated potential for future patient-oriented applications. For example, high mutation rates have, in the case of some cancers, been associated with T-cell infiltrates and with responses to immune checkpoint inhibitor therapies [1–4], a conclusion that has recently been reproduced in a mouse model, possibly allowing for a more precise understanding of the high mutation rate, T-cell infiltrate connection, via empirical approaches. Recovery of TcR- β VDJ recombination reads from bladder cancer WXS files is associated with better outcomes [5]. And, co-detection of TcR- α and TcR- β recombination reads in single melanoma WXS files is associated with PD-1 expression in corresponding RNASeq files [6]. While both survival outcome-TcR VDJ associations, and PD-1 expression, can be assessed with more sophisticated and more precise applications, maximizing the information available from WXS files offers the hope of increasing efficiencies and cost-effectiveness, particularly because WXS generation from tumor specimens is becoming widespread. Cost-effectiveness may also be an important consideration in rapidly moving certain therapies from internationally developed to undeveloped settings.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12307-018-0204-6>) contains supplementary material, which is available to authorized users.

✉ George Blanck
gblanck@health.usf.edu

¹ Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, 12901 Bruce B. Downs Blvd., Tampa, FL 33612, USA

To advance the opportunity to understand the value of obtaining TcR-recombination reads from WXS files, we processed WXS files from TCGA primary prostate cancer (PRAD) specimens and from an independent set of metastatic prostate samples for the TcR-recombination reads. The processing approach reported here is similar to previously reported approaches [6, 7], but with certain modifications that led to a more robust recovery of TcR-recombination reads. The results further confirm the opportunity to obtain significant immunoscore information from tumor specimen exome files. The results identify a very small population with a highly negative outcome, particularly important in the prostate cancer setting where “watchful waiting” is a commonly employed approach, given the large percentage of cases that will not have significant clinical impact.

Materials and Methods

Processing Exome Files for Reads Representing V(D)J Recombinations

The nucleotide sequences for both V regions and J regions for the human TcR genes were obtained from NCBI, and additional allelic variants were obtained from IMGT/GENE-DB. The V and J sequences for TcR- β , - γ , and - δ are listed in the supporting online material (SOM) of ref. [5]. The TcR- α sequences are in the SOM for ref. [7]. Whole exome sequences (WXS) of prostate adenocarcinoma (PRAD) were downloaded from the Genome Data Commons, using a token file provided for dbGaP approved project number 6300. An example of the download manifest for the PRAD WXS files is in Table S1 of the SOM. A `Module_Search_IgTcR` shell script was used to execute individual search scripts which identify candidate V region and J region sequences specific for each TcR gene (Table S2). The search scripts contain a list of unduplicated 10 nucleotide long sequences close to the 3' end of the V regions. The V regions sequences used in the script were 3, 5, 7, 9, and 11 nucleotides away from the 3' end of the canonical V sequences referred to above, to avoid losses in the search due to N-region diversity. Similarly, 10 nucleotide long sequences from 3, 5, 7, 9, and 11 nucleotides away from the 5' end of the J sequence were used. The search script compared each read in the vicinity of the TcR gene regions with the set of V and J regions, and the reads containing a match for both V and J regions were deposited into a TSV file and were then the object of the script, `IMGTSearchTCR.php`. This php process used the read containing the VJ combination to query the Immune GeneTics (IMGT) tool for determining whether the read is productive or unproductive; how many nucleotides are an exact match to V and J regions recognized by IMGT; as well as the quality of the match. For this project, only recombination reads containing greater than 20

nucleotide matches in both V and J regions were accepted as a productive or unproductive read. This relatively high standard was applied to ensure a reliable identification of the recombination reads. To process the `Module_Search_IgTcR` script, a `Module_Search_IgTcR_Header` execution script containing the environment paths and settings for the `Module_Search_IgTcR` script was created (Table S3). A summary output of the above software for each of the TcR receptor genes, providing results for various parameters, such as the V and J region nucleotide matches, is provided in the SOM (Table S4–7). See also ref. [8], and references therein; and the following link (<https://github.com/martinda/gnu-parallel/blob/master/CITATION>).

RNASeq Data

The RNASeq V2 RSEM (RNASeq Expectation Maximization) scores for the genes indicated in Results were downloaded directly from cBioPortal.org [9–12] (Table S8A). [TY1] The PRAD barcodes in the output, from the above processing steps for recovering V(D)J recombination reads, were grouped (using Microsoft Excel) into the following categories: (i) all barcodes, (ii) barcodes with unproductive TcR- β recombination reads, (iii) barcodes with co-detection of productive TcR- α/β recombination reads. This latter category (iii) was further divided into: (iv) barcodes with co-detection of three or more productive TcR- α/β recombination reads; and (v) barcodes with co-detection of only one productive TcR- α/β recombination read. The RNASeq V2 RSEM scores for each gene were then matched to the barcodes in each category mentioned above using the `VLOOKUP` function in Excel. The CD274 gene is provided as an example in the SOM for the construction of the box and whisker plots from the RNASeq data (Table S8B). A t-test was performed to compare the RNA expression levels between each category for each gene to provide the *p*-values listed in Table S8C and in Results.

Survival Data

Clinical information was downloaded from cBioPortal.org (Table S9), and then separated into categories. The categories were determined based on the detection of TcR recombination reads in the barcodes. The following groups of barcodes were established: (a) productive TcR- α , productive TcR- β , (b) co-detection of productive TcR- α and TcR- β , (c) unproductive TcR- β , (d) unproductive TcR- δ . After matching each barcode category to its survival data using the `VLOOKUP` function in Microsoft Excel, the categories were compared to each other using the average function and t-test function (Results). The IBM program, Statistical Package for the Social Sciences (SPSS) version 23 was then used to construct Kaplan-Meier curves and generate Log Rank *p*-values representing comparisons of the barcode groups.

PRAD Metastatic Sample Download from dbGaP

An additional 50 metastatic prostate adenocarcinoma WXS samples were obtained from the database of Genotypes and Phenotypes (dbGaP), study accession: phs000915.v1.p1, approval of request number # 54707-1. SRA (Sequence Read Archive) Toolkit was used to download the files. The utility in SRA Toolkit, “prefetch”, was used in conjunction with HTTP transfer to download the SRA files. The raw SRA files were then converted to BAM files to be compatible with Samtools before being subjected to the same processing steps used for TCGA-PRAD WXS files.

Results

Recovery of TcR-Recombination Reads from PRAD WXS Files

Results of processing the PRAD WXS files for TcR recombination reads indicate a robust detection of T-cell infiltration in the primary PRAD tumor barcodes (Table 1). Both productive and unproductive recombination reads were detected for TcR- α and TcR- β . As in several past cases [5], the ratio of barcodes with productive TcR- β recombination reads, to barcodes with unproductive TcR- β recombination reads is relatively high, suggesting that T-cells which undergo unproductive TcR- β recombinations do not survive long enough to be detected or maintained in the tumor microenvironment. This is in sharp contrast to TcR- α , where in numerous cancers, unproductive recombination reads are commonly detected, although, this represents the first report of detection of unproductive TcR- α recombination reads in 100% of the barcodes. The very high level of TcR- α recombination is consistent with the understanding that many T-cells undergo the recombination of both TcR- α alleles [13] and consistent with prostate cancer specimens often having infiltrating lymphocytes [14, 15].

There was an overlap in barcodes that represented both productive TcR- α and productive TcR- β recombination reads: 7.3% of all the barcodes, with an average of 3.6 reads per sample.

Two examples of a productive TcR- α and TcR- β recombination read, with the N-region nucleotides and the VDJ usage, are shown in Fig. 1.

There was detection of both productive and unproductive TcR- γ recombination reads, and a small detection of unproductive TcR- δ recombination reads. Productive TcR- γ recombination reads were detected in nearly 12% of the barcodes, unproductive TcR- γ was detected in about 23% of the barcodes (Table 1).

A collection of productive and unproductive recombination reads, and the V and J usage for TcR- β recombination is presented in Table S5; productive TcR- α and unproductive TcR- δ recombination reads, and the respective V and J usage, are in Tables S4 and S6; productive and unproductive TcR- γ recombination reads and the V and J usage are in Table S7.

T-Cell Exhaustion RNASeq Markers Correlated with Co-Detection of TcR- α and TcR- β Recombination Reads in WXS Files

To verify and extend the understanding of T-cell markers in prostate cancer genomics files, RNASeq V2 RSEM scores were obtained from cBioPortal.org for the following genes: IFNG, CD3D, CD4, and CD8A. The average RNASeq values for each gene were then compared across barcode groups, as follows: (i) all barcodes, as a control set; (ii) barcodes with unproductive TcR- β reads only, as a control set; and (iii) barcodes with co-detection of productive TcR- α / β reads, as a test set. Results indicated that the above genes had a statistically significant, higher level of RNASeq values for the group of barcodes representing co-detection of TcR- α / β reads in comparison to the two above indicated control barcode groups (Fig. 2a, Table S8C), except in the case of CD3D when the TcR- α / β barcode group was compared to the control

Table 1 TCGA barcode and read counts for productive and unproductive recombination representing TcR- α , - β , - γ , - δ in prostate adenocarcinoma WXS files

Human T cell receptors	Detection of recombination reads in 479 samples from TCGA ^a			
	Productive		Unproductive	
	Read count/ Barcode count	Percentage of barcodes	Read count/ Barcode count	Percentage of barcodes
α	190/121	25.3%	6554/479	100%
β	165/101	21.1%	27/24	5.01%
γ	89/57	11.9%	154/112	23.4%
δ	0	0%	8/8	1%
Co-detection of α & β	126/35	7.31%	369/24	5.01%

^a TCGA, the cancer genome atlas

Fig. 1 **a** Examples of TcR- α productive reads from WXS PRAD files. Barcode TCGA-CH-5738 (top); TCGA-CH-5743 (bottom). **b** Examples of TcR- β productive reads from WXS PRAD files. Barcode TCGA-VI-A8WW (top); TCGA-EJ-A8FP (bottom)

group represented by barcodes where there was the detection of unproductive TcR- β only. In this latter case, results for CD3D did show that the average RNASeq value for the TcR- α/β co-detection group was higher than in the case of the unproductive TcR- β group, however, this comparison did not yield a significant p -value ($p = 0.08$).

Overall, the above results indicate a strong association of T-cell markers with the barcode group that represents co-detection of TcR- α/β recombination reads. To further verify the association of T-cell markers with this barcode group, the barcode group that represents co-detection of TcR- α/β was subdivided into barcodes with three or more TcR- α/β

Fig. 2 **a** Box and whisker plot of RNA Expression (RNA Seq V2 RSEM) for all barcodes (dark) vs barcodes with detection of TcR- β unproductive recombination reads (medium) vs co-detection of productive TcR- α/β recombination reads (light). For IFNG, the y-axis ranges from 0 to 9; CD3D, y-axis ranges from 0 to 300; CD4, y-axis ranges from 0 to 2500; CD8A, y-axis ranges from 0 to 700. **b** Box and whisker plot of RNA expression (RNA Seq V2 RSEM) for all barcodes (dark) vs barcodes with detections of TcR- β unproductive recombination reads (medium) vs co-detection of productive TcR- α/β recombination reads (light). For CD33, the y-axis ranges from 0 to 100; CTSS, y-axis ranges

from 0 to 350; ITGAX, y-axis ranges from 0 to 700; CD68, y-axis ranges from 0 to 3000; CD83, y-axis ranges from 0 to 900; CIITA, y-axis ranges from 0 to 700; CTSL, y-axis ranges from 0 to 4000. **c** Box and whisker plot of RNA Expression (RNA Seq V2 RSEM) for all barcodes (dark) vs barcodes with detections of TcR- β unproductive recombination reads (medium) vs co-detection of productive TcR- α/β recombination reads (light). For C10ORF54 (VISTA), the y-axis ranges from 0 to 2000; PDCD1, y-axis ranges from 0 to 140; CD274, y-axis ranges from 0 to 45; CTLA4, y-axis ranges from 0 to 120. A detailed version of all of the box and whisker plots in this figure is provided in the SOM, Fig. S1

recombination reads and barcodes with only one TcR- α/β recombination read. Average RNASeq values were obtained for the four T-cell markers studied above (IFNG, CD3D, CD4, and CD8A), for the two groups of barcodes representing the two different read quantification categories. Results indicated that, for each of the four T-cell markers, there was a statistically significant higher level of expression in the barcode group representing the high number of TcR- α/β recombination reads (Table S8C).

To determine the potential level of antigen-presentation activity in the tumor microenvironment, where there was co-detection of TcR- α/β recombination reads, RNASeq V2 RSEM scores were obtained for the following genes: CD33, CD68, CIITA, CTSS, ITGAX, CD83, and CTSL. We then compared the average RNASeq values for the barcodes representing the co-detection of TcR- α/β and the two barcode groups indicated as control groups above, i.e., the barcode groups represented by all barcodes and barcodes with only detection of unproductive TcR- β . There was statistically significant, higher level expression for the above indicated antigen-presentation related genes in the barcode group representing co-detection TcR- α/β recombination reads, except in the cases of CIITA and CTSL (Fig. 2b, Table S8C). In the case of CIITA, the co-detection barcode group was found to have significantly higher average RNASeq value only when compared to the average RNASeq value for all barcodes and not when compared to the average RNASeq value for the barcode group representing detection of unproductive TcR- β recombination reads. In the case of CTSL, although the co-detection barcode group on average had a higher RNASeq value, compared to both control groups, neither control group represented a statistically significant comparison. When comparing the group of barcodes representing co-detection of three or more TcR- α/β recombination reads, in each WXS file, to the group of barcodes representing co-detection of only one such recombination read in the WXS file, the barcode group with three or more reads had a statistically significant, higher average RNASeq value than the barcode group with only one read, for all of the genes representing antigen presentation activity, except in the cases of CD68 ($p < 0.06$) and CTSL ($p < 0.06$) (Table S8C).

In a previous study of recovery of TcR recombination reads from melanoma WXS files, it was determined that co-detection TcR- α/β recombination reads correlated with a high level of PD-1 expression [6]. To assess the relationship between recovery of TcR recombination reads and T-cell exhaustion markers in PRAD, we obtained the RNASeq V2 RSEM scores for the following genes: PDCD1, CD274, CTLA4, and VISTA [16]. When comparing the barcode group representing co-detection of TcR- α/β recombination reads to the two control barcode groups indicated above (all barcodes and barcodes representing unproductive TcR- β , respectively), the average RNASeq values were higher, to a statistically

significant degree, for the barcodes representing simultaneous recovery of TcR- α and TcR- β recombination reads (i.e., the co-detection barcode group), in comparison with both control sets (Fig. 2c, Table S8C), for PDCD1, CD274, and CTLA4. In the case of VISTA, the barcode group representing co-detection of TcR- α/β had a statistically significant, higher average RNASeq value when compared with the average RNASeq value for all barcodes. However, while the average RNASeq value for VISTA of the co-detection barcode group was higher than the average VISTA RNASeq value for barcode group representing detection of unproductive TcR- β only, the difference was not statistically significant. Overall, these results provided an indication that barcodes with co-detection of productive TcR- α/β recombination reads have a higher level of immune checkpoint gene expression.

When examining the difference in RNASeq expression levels of these genes for barcodes representing co-detection of TcR- α/β recombination reads with three or more reads versus co-detection of TcR- α/β recombination with only one read, the barcodes representing only one read had a higher level of mRNA expression than barcodes containing three or more reads for genes CD274, CTLA4, and VISTA (Table S8C). This is an indication of the reverse trend observed when comparing barcodes containing co-detection of TcR- α/β recombination reads to the control groups (all barcodes and barcodes representing unproductive TcR- β) conducted earlier. This reverse trend was not observed for PDCD1 during the read count comparison; barcodes containing three or more co-detection of TcR- α/β recombination reads had a higher level of PDCD1 expression compared to barcodes containing only one read, however, this result was not statistically significant ($p < 0.06$).

Poor survival rates associated with detection of TcR- δ and $-\gamma$ recombination reads

To determine whether there was an association between survival rates and detection of particular TcR recombinations in the PRAD WXS files, barcode categories, representing recoveries of different types of recombination reads, were plotted on a Kaplan-Meier (KM) curve against the remaining population, using a cBioPortal.org web tool. The barcode group representing detection of unproductive TcR- δ recombination reads had statistically significantly worse disease-free survival than the remaining population $p < 0.029$ (Fig. S2), although the sample size for that barcode group is small. The KM curve obtained from cBioPortal.org was replicated using the program, IBM SPSS, to confirm survival difference between the two barcode groups (Fig. 3a; $p < 0.032$).

To further assess the survival rates for WXS files containing TcR- δ recombination reads, particularly because of the small TcR- δ sample size, disease-free survival and overall survival data were obtained from cBioPortal.org and further

Fig. 3 **a** Kaplan-Meier Survival Curve for barcodes representing recovery of TcR- δ unproductive recombinations (TCGA-CH-5748-01, TCGA-CH-5751-01, TCGA-EJ-A6RC-01, TCGA-HC-7080-01, TCGA-HC-7081-01, TCGA-HC-7209-01, TCGA-HC-8259-01, TCGA-J4-A83M-01) compared to the survival curve for the remaining barcodes. Log Rank comparison p -value <0.032 . **b** Kaplan-Meier Survival Curve for barcodes representing co-detection of TcR- γ

productive and unproductive recombinations compared to the survival curve for the remaining barcodes. Log Rank comparison p -value <0.008 . **c** Kaplan-Meier Survival Curve for barcodes representing co-detection of TcR- γ productive and unproductive recombinations plus barcodes representing TcR- δ unproductive recombinations compared to the survival curve for the remaining barcodes. Log Rank comparison p -value <0.002

analyzed. Specific barcode categories were established as control groups to determine whether a survival difference, compared to the TcR- δ , recombination-read barcode (test) group, could be verified. The control groups were as follows: (i) barcodes with co-detection of productive TcR- α and TcR- β recombination reads, (ii) barcodes with productive TcR- β recombination reads, (ii) barcodes with productive TcR- β or unproductive TcR- β recombination reads, and (iv) barcodes with unproductive TcR- β recombination. Results indicated a statistically significant difference in both disease-free and overall survival rates for nearly all comparisons of the control groups and the test group, where the control groups had higher rates of survival, compared to the test group, i.e., compared to the barcodes with unproductive TcR- δ recombination reads (Table 2). The only instance where statistical significance was not obtained was in comparing the test group to the

barcodes with the detection of productive TcR- β recombination reads, for disease-free survival. Among the control groups tested, the barcode group that yielded the highest survival averages was the group of barcodes representing unproductive TcR- β recombination reads, with $p < 0.002$ for both disease-free survival and overall survival, in comparison to barcodes containing TcR- δ recombination reads (Table 2). The difference in the average disease-free survival between the two categories was 24 months, and the difference in the average overall survival between the two categories was 22 months.

Because the above data representing the barcode group where there was detection of unproductive, TcR- δ recombination reads presumably represented γ/δ T-cells in the tumor microenvironment, we determined whether barcodes representing WXS detection of TcR- γ recombination reads would also be associated with worse disease survival,

Table 2 Disease-free survival and overall survival for barcodes with unproductive TcR- δ recombination reads detectable in the WXS files. (p -values in bold)

Barcode groups and matches for statistical analyses	PRAD barcode groups	Disease-free survival (months)	p value (less than)	Overall survival (months)	p value (less than)
1	Co-detection of Productive TcR- α and Productive TcR- β	29.99		35.46	
2	Unproductive δ	21.29		25.12	
1 vs 2	p-values for Co-detection of Productive TcR- α & TcR- β vs Unproductive TcR- δ		0.028*		0.021*
3	Productive TcR- β	28.92		32.79	
3 vs 2	p-values for Productive TcR- β vs Unproductive TcR- δ		0.074		0.011*
4	Productive TcR- β or Unproductive TcR- β	30.87		34.68	
4 vs 2	p-values for Both Productive or Unproductive TcR- β vs Unproductive TcR- δ		0.007*		0.011*
5	Unproductive TcR- β	45.25		47.78	
5 vs 2	p-values for Unproductive TcR- β vs Unproductive TcR- δ		0.0005*		0.002*

compared to the remaining population. The comparison of barcodes with productive TcR- γ recombination reads to the remaining population, or barcodes with unproductive TcR- γ recombination reads to the remaining population, did not indicate statistically significant different survival rates. However, barcodes representing co-detection of TcR- γ productive and unproductive barcodes did demonstrate a significantly lower survival rate, compared to all remaining barcodes, using both the IBM SPSS software (Fig. 3b) and the KM webtool at cBioPortal.org (p -value <0.033 ; SOM Fig. S3).

When comparing barcodes with co-detection of productive and unproductive TcR- γ recombination reads, plus barcodes with unproductive TcR- δ recombination reads, to the remaining barcode population on the KM curve, the survival difference was even more significant $p < 0.002$ (Fig. 3c) (Fig. S4).

KM curves and Log Rank Tests were then generated for the three new γ/δ T-cell, barcode test categories against a control group. The new test categories were (i) barcodes representing unproductive TcR- δ recombination reads, (ii) barcodes representing co-detection of productive and unproductive TcR- γ recombination reads, and (iii) the barcodes representing unproductive TcR- δ recombination reads combined with the barcodes representing co-detection of productive and unproductive TcR- γ recombination reads. The control group used was barcodes representing unproductive TcR- β recombination reads. Statistical significance, representing worse survival for the barcode groups representing the γ/δ T-cells, was obtained for all three comparisons (Fig. 4).

Additional, metastatic prostate files were downloaded and processed to extract TcR- recombination reads. For 50 metastatic samples, there was an increase in the ratio of TcR- δ recombination reads per barcode: 1 % in primary tumor samples from TCGA compared to 10 % in metastatic tumor samples from dbGaP.

Discussion

The above results support the validity and value of the opportunity to detect immune receptor recombination reads in WXS files, particularly important given the likelihood of the WXS approach as being the mainstay of cancer genomics in future, standard of care settings. That is, it is important to obtain as much information from the WXS file as possible, considering its likely preeminence in cancer genomics. This immunoscore opportunity is not only important in view of the fact that the WXS approach is likely to be a minimal standard for cancer genomics, but may also be important in settings where costs for additional tests are simply prohibitive. Thus, in developed countries, where many follow up tests are available, maximizing the information from a WXS file could support or make unnecessary additional tests; in other venues, a WXS file may be all that is available to guide therapy.

The software used in this study to identify the TcR recombination reads in the WXS files is similar in approach to software used in our previous, closely related work [5–7, 17]. However, the current algorithm has been modified to increase detection sensitivity, detection speed, as well as quality of the extracted reads. In particular, the previous algorithms involved a single search for V sequences, 5 nucleotides away from the 3' end, to avoid overlaps in the area of N-region diversity. In the current algorithm, the V-sequence and J-sequence searches both involve a step-wise search for V- and J-sequences away from the joining region. Again, this approach minimizes loss due to overlap with N-region nucleotides, but increases recovery by varying the distance, in particular for the V-sequences, from the joining site.

In this report, and in other reports, in general, barcodes that represent recovery of TcR recombination reads also have higher levels of RNASeq T-cell biomarkers, supporting the association of the recovery of the TcR reads with T-cell

Fig. 4 **a** Kaplan-Meier Survival Curve for barcodes representing recovery of TcR- δ unproductive recombinations compared to the survival curve for barcodes representing TcR- β unproductive. Log Rank comparison p -value <0.007 . **b** Kaplan-Meier Survival Curve for barcodes representing co-detection of TcR- γ productive and unproductive recombinations compared to the survival curve for

barcodes representing TcR- β unproductive. Log Rank comparison p -value <0.005 . **c** Kaplan-Meier Survival Curve for barcodes representing co-detection of TcR- γ productive and unproductive recombinations plus barcodes representing TcR- δ unproductive recombinations compared to the survival curve for barcodes representing TcR- β unproductive reads. Log Rank comparison p -value <0.003

presence in the tumor specimen microenvironment [6, 18]. Barcodes associated with co-detection of productive TcR- α and TcR- β recombinations had the highest RNASeq expression levels for T-cell marker genes, T-cell exhaustion genes, as well as genes associated with antigen-presentation when compared to barcodes associated with detection of unproductive TcR rearrangements or the total barcode collection. The most important aspect of these results would appear to be the indication that co-detection of TcR- α/β recombination reads in a WXS file is a strong indicator of T-cell exhaustion. These results raise the question of whether patients with co-detection of TcR- α/β recombination reads in a WXS file would benefit most from anti-immune checkpoint therapy approaches? However, it should be noted that, in the current study, there was an indication that barcodes with the highest number of reads, among the TcR- α/β co-detection barcodes, had a reduced level of immune checkpoint biomarker expression. While clearly additional study is needed, this result could indicate that the highest level of T-cell exhaustion may correlate with an intermediate level of TcR recombination read detection.

To date, there has been only one report whereby detection of TcR recombination reads in WXS files has been associated with patient outcome, namely in the case of bladder cancer, where a combined TCGA-Moffitt Cancer Center cohort of bladder cancer patients, with detection of TcR- β recombination reads in the WXS files, represented a favorable outcome [5]. In this current study, several different approaches indicated that detection of TcR- δ and $-\gamma$ recombination reads was associated with a poor outcome, with the greatest statistical significance represented by co-detection of productive TcR- γ , unproductive $-\gamma$ and unproductive $-\delta$. Clearly these results suggest a role for γ/δ T-cells in a poor outcome, but the heavy connection to unproductive recombinations of the TcR- γ and TcR- δ genes opens up numerous possible roles for these types of T-cells, for example, secretion of specific cytokines supporting tumor growth. In addition, the fact that the most significant associations of the TcR- δ and TcR- γ recombination reads represented extensive co-detection, the impact on prostate cancer development is likely to be greater with a greater level of γ/δ T-cell infiltrate. In particular, this result represents a striking example of the value of assessing the presence of unproductive immune receptor recombinations in tumor specimen WXS files. Without the unproductive reads, the genomics-based detection of γ/δ T-cell presence in the immune infiltrate, or the genomics-assessment of γ/δ T-cell association with a poor outcome, would not have been possible.

γ/δ T-cells ordinarily represent a very small proportion of either tissue resident or blood-born T-cells, in comparison to α/β T-cells. However, there have been reports that γ/δ T-cell infiltrates can suppress antigen presentation, in a tumor microenvironment, to α/β T-cells, for example, ref. [19]. Also, prostate tumor samples in particular have a higher proportion of

γ/δ T-cells, compared to tumors not derived from an epithelial origin, such as melanoma [19]. Reports have indicated how γ/δ T-cells, cytokine elaboration could play a role in immunosuppression, with the following examples IL10 [20–22]; IFNG and IL4 [23]; IL2, IL3, and GM-CSF [24]; and TGF- β [25]. Among that list of cytokines produced by γ/δ T-cells, IL-10 and TGF- β , are most widely regarded as being immunosuppressive in the tumor microenvironment [26, 27].

However, even without a clear understanding of mechanism, the above data may provide the basis for sequencing patient tumor samples prior to giving prognosis of disease, particularly for the development of an era of “watchful waiting” for prostate cancer.

Acknowledgements Authors thank USF research computing and the taxpayers of the State of Florida. This article is dedicated to William.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Bouffet E, Larouche V, Campbell BB, Merico D, de Borja R, Aronson M et al (2016) Immune checkpoint inhibition for hypermutant glioblastoma multiforme resulting from Germline Biallelic mismatch repair deficiency *J Clin Oncol: Off J Am Soc Clin Oncol*. <https://doi.org/10.1200/JCO.2016.66.6552>
2. Li SK, Martin A (2016) Mismatch repair and colon cancer: mechanisms and therapies explored. *Trends Mol Med* 22(4):274–289. <https://doi.org/10.1016/j.molmed.2016.02.003>
3. Westdorp H, Fennemann FL, Weren RD, Bisseling TM, Ligtenberg MJ, Figdor CG et al (2016) Opportunities for immunotherapy in microsatellite instable colorectal cancer. *Cancer Immunol Immunother: CII* 65(10):1249–1259. <https://doi.org/10.1007/s00262-016-1832-7>
4. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, Walsh LA, Postow MA, Wong P, Ho TS, Hollmann TJ, Bruggeman C, Kannan K, Li Y, Elipenahli C, Liu C, Harbison CT, Wang L, Ribas A, Wolchok JD, Chan TA (2014) Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* 371(23):2189–2199. <https://doi.org/10.1056/NEJMoa1406498>
5. Samy MD, Tong WL, Yavorski JM, Sexton WJ, Blanck G (2017) T cell receptor gene recombinations in human tumor specimen exome files: detection of T cell receptor-beta VDJ recombinations associates with a favorable oncologic outcome for bladder cancer. *Cancer Immunol Immunother: CII* 66(3):403–410. <https://doi.org/10.1007/s00262-016-1943-1>
6. Tu YN, Tong WL, Samy MD, Yavorski JM, Kim M, Blanck G (2017) Assessing microenvironment immunogenicity using tumor specimen exomes: co-detection of TcR-alpha/beta V(D)J recombinations correlates with PD-1 expression. *Int J Cancer* 140(11):2568–2576. <https://doi.org/10.1002/ijc.30675>
7. Gill TR, Samy MD, Butler SN, Mauro JA, Sexton WJ, Blanck G (2016) Detection of productively rearranged TcR-alpha V-J sequences in TCGA exome files: implications for tumor

- Immunoscore and recovery of antitumor T-cells. *Cancer Inform* 15:23–28. <https://doi.org/10.4137/CIN.S35784>
8. Malaspinas AS, Tange O, Moreno-Mayar JV, Rasmussen M, DeGiorgio M, Wang Y, Valdiosera CE, Politis G, Willerslev E, Nielsen R (2014) Bammds: a tool for assessing the ancestry of low-depth whole-genome data using multidimensional scaling (MDS). *Bioinformatics* 30(20):2962–2964. <https://doi.org/10.1093/bioinformatics/btu410>
 9. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6(269):pl1. <https://doi.org/10.1126/scisignal.2004088>
 10. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2(5):401–404. <https://doi.org/10.1158/2159-8290.CD-12-0095>
 11. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12(1):323. <https://doi.org/10.1186/1471-2105-12-323>
 12. Guo Y, Sheng Q, Li J, Ye F, Samuels DC, Shyr Y (2013) Large scale comparison of gene expression levels by microarrays and RNAseq using TCGA data. *PLoS One* 8(8):e71462. <https://doi.org/10.1371/journal.pone.0071462>
 13. Padovan E, Casorati G, Dellabona P, Meyer S, Brockhaus M, Lanzavecchia A (1993) Expression of two T cell receptor alpha chains: dual receptor T cells. *Science* 262(5132):422–424. <https://doi.org/10.1126/science.8211163>
 14. Karja V, Aaltomaa S, Lipponen P, Isotalo T, Talja M, Mokka R (2005) Tumour-infiltrating lymphocytes: a prognostic factor of PSA-free survival in patients with local prostate carcinoma treated by radical prostatectomy. *Anticancer Res* 25(6C):4435–4438
 15. Vesalainen S, Lipponen P, Talja M, Syrjänen K (1994) Histological grade, perineural infiltration, tumour-infiltrating lymphocytes and apoptosis as determinants of long-term prognosis in prostatic adenocarcinoma. *Eur J Cancer* 30A(12):1797–1803
 16. Gao J, Ward JF, Pettaway CA, Shi LZ, Subudhi SK, Vence LM, Zhao H, Chen J, Chen H, Efsthathiou E, Troncso P, Allison JP, Logothetis CJ, Wistuba II, Sepulveda MA, Sun J, Wargo J, Blando J, Sharma P VISTA is an inhibitory immune checkpoint that is increased after ipilimumab therapy in patients with prostate cancer. *Nat Med* 2017;23(5):551–5. <https://doi.org/10.1038/nm.4308>, 555.
 17. Tong WL, Tu YN, Samy MD, Sexton WJ, Blanck G (2017) Identification of immunoglobulin V(D)J recombinations in solid tumor specimen exome files: evidence for high level B-cell infiltrates in breast cancer. *Hum Vaccin Immunother* 13(3):1–6. <https://doi.org/10.1080/21645515.2016.1246095>
 18. Butler SN, Blanck G (2016) Immunoscore by correlating MHC class II and TCR expression: high level immune functions represented by the KIRP dataset of TCGA. *Cell Tissue Res* 363(2):491–496. <https://doi.org/10.1007/s00441-015-2261-1>
 19. Peng G, Wang HY, Peng W, Kiniwa Y, Seo KH, Wang RF (2007) Tumor-infiltrating gammadelta T cells suppress T and dendritic cell function via mechanisms controlled by a unique toll-like receptor signaling pathway. *Immunity* 27(2):334–348. <https://doi.org/10.1016/j.immuni.2007.05.020>
 20. Fournie JJ, Sicard H, Poupot M, Bezombes C, Blanc A, Romagne F et al (2013) What lessons can be learned from gammadelta T cell-based cancer immunotherapy trials? *Cell Mol Immunol* 10(1):35–41. <https://doi.org/10.1038/cmi.2012.39>
 21. Azuara V, Levraud JP, Lembezat MP, Pereira P (1997) A novel subset of adult gamma delta thymocytes that secretes a distinct pattern of cytokines and expresses a very restricted T cell receptor repertoire. *Eur J Immunol* 27(2):544–553. <https://doi.org/10.1002/eji.1830270228>
 22. Wen L, Barber DF, Pao W, Wong FS, Owen MJ, Hayday A (1998) Primary gamma delta cell clones can be defined phenotypically and functionally as Th1/Th2 cells and illustrate the association of CD4 with Th2 differentiation. *J Immunol* 160(4):1965–1974
 23. Ferrick DA, Schrenzel MD, Mulvania T, Hsieh B, Ferlin WG, Lepper H (1995) Differential production of interferon-gamma and interleukin-4 in response to Th1- and Th2-stimulating pathogens by gamma delta T cells in vivo. *Nature* 373(6511):255–257. <https://doi.org/10.1038/373255a0>
 24. Matsue H, Cruz PD Jr, Bergstresser PR, Takashima A (1993) Profiles of cytokine mRNA expressed by dendritic epidermal T cells in mice. *J Invest Dermatol* 101(4):537–542. <https://doi.org/10.1111/1523-1747.ep12365917>
 25. Seo N, Tokura Y, Takigawa M, Egawa K (1999) Depletion of IL-10- and TGF-beta-producing regulatory gamma delta T cells by administering a daunomycin-conjugated specific monoclonal antibody in early tumor lesions augments the activity of CTLs and NK cells. *J Immunol* 163(1):242–249
 26. Ke Y, Kapp LM, Kapp JA (2003) Inhibition of tumor rejection by gammadelta T cells and IL-10. *Cell Immunol* 221(2):107–114. [https://doi.org/10.1016/S0008-8749\(03\)00066-2](https://doi.org/10.1016/S0008-8749(03)00066-2)
 27. Kazemfar K, Chen R, Nicholson K, Coppola D, Zhou JM, Chen X, Wei S, Blanck G (2009) Combined IL-8 and TGF-beta blockade efficiently prevents neutrophil infiltrates into an A549-cell tumor. *Immunol Lett* 122(1):26–29. <https://doi.org/10.1016/j.imlet.2008.10.010>