Applications of TCR repertoire analysis for biomarker discovery and beyond

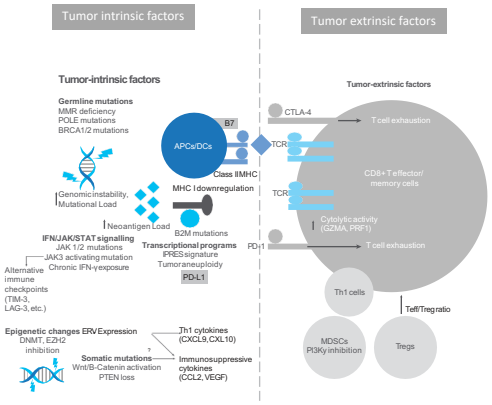
Ankita Das[1], Vasumathi Kode[1], Kayla Lee[1], Priyanka Shah[2], Xiaoshan "Shirley" Shi[1], Nitin Mandloi, Ravi Gupta[2], Amit Chaudhuri[1,2] and Papia Chakraborty[1]

1) MedGenome Inc, Foster City, CA, 2) MedGenome Labs, Bangalore, India; corresponding author : ankita.d@medgenome.com

Abstract

T cell immunity provides significant therapeutic benefit to cancer patients treated with checkpoint inhibitors, however a very small fraction of patients typically respond to checkpoint inhibitors, and a smaller fraction of them have any long-term benefit. This can be attributed to the lack of prognostic and predictive biomarkers. The infiltration levels of CD8 T cell in tumors is often used as a characteristic biomarker and can be correlated with response, but recent studies have shown that examining the functional state of the T cells and other immune cell types in the tumor, and the immunogenic neoantigen burden and the TCR repertoire clonality, might give a more appropriate representation of what might be going on in the tumors and hence can be used as predictive biomarkers for response. At MedGenome we have built a suite of tools that can be utilized to: a) Study the tumor microenvironment using the OncoPept*TUME* analysis pipeline, b) Predict and validate the immunogenic neoantigens using OncoPept*VAC* & OncoPept*SCRN*, and c) A suite of workflows to analyze the TCR repertoire from a wide range of sample types using bulk and single cell approaches. Here we present the data highlighting the applications of TCR repertoire as a biomarker for immunotherapy and also present our capabilities of providing these solutions as a service.

Overview of the tumor extrinsic and intrinsic factors that determine immunotherapy success



(modified from Ann. Rev. Med. 2018 69 : 33:47)

Figure 1: Shows an overview of the complex interplay between the tumor cell and the immune cell types that can be in the tumor, and the cell extrinsic and intrinsic factors in a patient's tumor that can be important in determining success to immunotherapy or not. On the left here are some of the tumor intrinsic molecular features : for instance if a tumor carries certain germline mutations that result in genomic instability, leading to high tumor mutational load and some of those mutations leading to expression of neoantigens in the tumor, if those neoantigens are then presented on the surface of the cell by HLA presentation, and presented via an APC, then cells in the microenvironment such as a CD8 T cell can be attracted to that target and induce tumor killing. However, they need to have the appropriate TCR repertoire to recognize the HLA bound peptide antigens to induce downstream signaling and mount an anti-tumor response. There are other immune cell types in the tumor such as MDSCs and T-regs that can be suppressive for inducing tumor killing mechanisms. Therefore, by studying the immune infiltration in the tumor, and identifying potential immunogenic neoantigens, and looking at the TCR repertoire can all be informative biomarkers for immunotherapy.

MedGenome's tools to enable to biomarker discovery for Immunotherapy

Table 1: Summary of the biomarker discovery services at MedGenome

Biomarker/ therapeutic	MedGenome's solution	Deliverables
Tumor microenvironment analysis	OncoPept*TUME*	Gene expression analysis Immune score cytokines & pathways
TCR repertoire analysis	OncoPept*TCR*	V(D)J genes CDR3 clonotypes TCR α/β pairing
Neoantigen prediction and validation	OncoPept*VAC* & OncoPept*SCRN*	Predicted & validated neoantigen vaccine candidates

Overview of the workflow for OncoPept*VAC* pipeline and TCR repertoire analysis

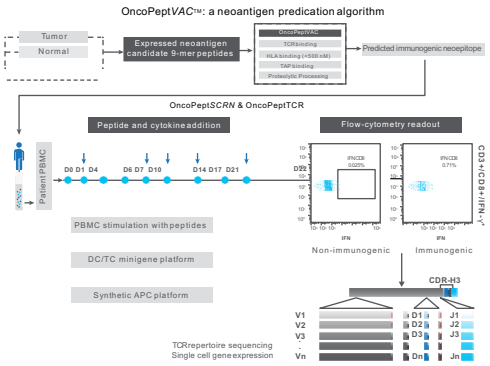


Figure 2: Workflow of neoantigen prediction and screening : The OncoPept*VAC* pipeline takes whole exome sequencing and RNA-sequencing data from are performed using the genomic DNA from blood and tumor and RNA from the patient's tumor. All mutant peptides are passed through the OncoPept*VAC* pipeline. The potential immunogenic peptides are predicted based on a series of prioritization steps including peptide processing, peptide TAP binding, HLA binding and TCR binding. Finalized peptides are then subjected to OncoPept*SCRN* wherein peptides are tested for CD8 T cell activation assay in presence of antigen loaded APCs. Functional TCR cell response is assessed by IFNγ production by flow cytometry, TCR repertoire sequencing and single cell gene expression analysis. The TCR repertoire and single cell gene expression analysis are performed to reveal the clonality and the functional T-cell states of the amplified T-cells.

TCR repertoire sequencing combined with single cell gene expression analysis reveals the clonal amplification and functional T-cell states

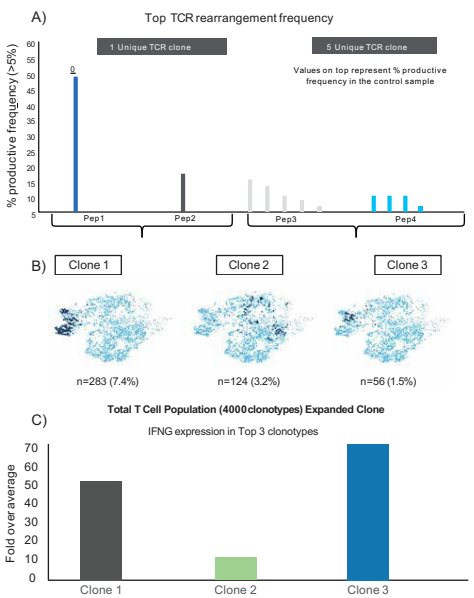
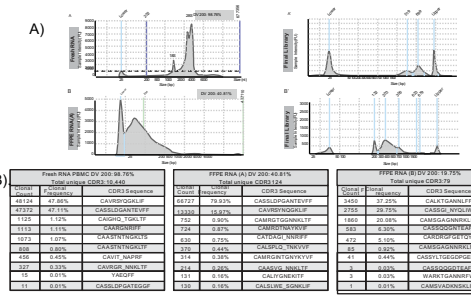


Figure 3: TCR repertoire and single cell gene expression analysis to identify antigen specific TCR repertoire clonal expansion: A) TCR repertoire analysis of PBMCs screened using the neoantigen vaccine candidates shows that compared to their wildtype counterparts, mutant forms of the peptide candidates show TCR α/β clonotype frequency changes between 20%-50%. We also observed that while some vaccine candidates mount a monoclonal response, others induce a polyclonal response. The TCR repertoire libraries were generated using the SMARTer TCR α/β Profiling Kit (Takara Bio USA), and sequencing was performed using the Illumina MiSeq platform, and repertoire analysis was performed using the MixCR pipeline. B&C) Single cell TCR repertoire analysis of the PBMCs (with positive responses to the peptide, assayed by IFNγ production) identified over 4000 TCR clonotypes, with the t-SNE plots showing the top 3 clonotypes and the frequencies are listed below. Single cell gene expression of the amplified T cells reveal that there is varying levels of IFNγ in T-cells that are clonally expanded, and that this information can provide insights on the functional state of the expanded T-cell clones. The single cell TCR and gene expression experiment was performed using the 10x Genomics Immune profiling solution & sequencing was performed using the Illumina Novaseq 6000, and 10x Cell Ranger analysis was performed downstream.

γ/δ TCR repertoire Profiling as a biomarker & potential therapeutic



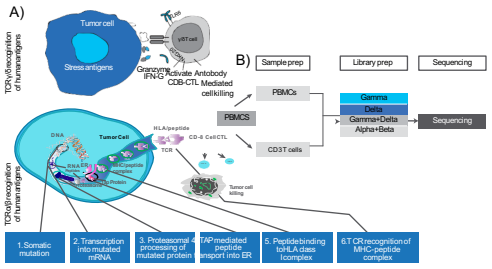
Fresh RNA PBMC DV 200: 98.76% Total unique CDR3: 10,440		
Clonal Counts	Clonal Frequency	CDR3 Sequence
48124	47.86%	CAVRSFQGKLIF
47372	47.11%	CASSLDGANTEVFF
1125	1.12%	CASQHQ_TGKLTF
1113	1.11%	CAWSRGNRIF
1073	1.07%	CAASTNTNQGLTS
808	0.80%	CAASTNTNQGKLTF
456	0.45%	CAVIT_NAPFF
327	0.33%	CAVRGR_NNKLTF
15	0.01%	TYEQFF
11	0.01%	CASSLGPGATEGGF

FFPE RNA (A) DV 200: 40.81% Total unique CDR3:124		
Clonal Counts	Clonal Frequency	CDR3 Sequence
79337	79.93%	CASSLGPGANTEVFF
19338	19.97%	CAVRSYNQGKLIF
752	0.90%	CAMRGTGGNNKLTF
724	0.87%	CAMRGTNNYKYVF
639	0.75%	CATDAGL_NNRFF
370	0.44%	CALSPLG_TNKVYF
314	0.38%	CAMRGNTGNYKYVF
214	0.26%	CAASVG_NNKLTF
131	0.16%	CALYGNEKITF
130	0.16%	CALSLWE_SGNKLF

FFPE RNA (B) DV 200: 19.75% Total unique CDR3:79		
Clonal Counts	Clonal Frequency	CDR3 Sequence
3450	37.25%	CALKTGANNLFF
2755	29.75%	CASSGI_NYQLIW
1869	20.08%	CAMSGAGNNRKLIW
583	6.30%	CASSPGQGNTEAFF
472	5.10%	CARDRGPGETQTYF
85	0.92%	CAMSGAGNNRKLIW
41	0.44%	CASSYLTGEGPGELFF
3	0.03%	CASSPGQGETEAFF
3	0.03%	WARKTGANNRFW
1	0.01%	CAMSVADNSKLIW

Figure 4: TCR repertoire analysis from FFPE tissue : A) Tapestatton profiles of the RNA from PBMC and an FFPE sample and the corresponding TCR α/β libraries generated from each of the samples. The expected TCR libraries from good quality RNA is 500-700bp, and with FFPE a skewing of the peaks between 200-500bp is observed. B) Table shows the top clonotypes identified from the three different FFPE samples. We show that for all the samples regardless of DV200, we identify over a 100 clonotypes. In addition to that, we observed that multiple samples from patients treated with the same vaccine show a shared clonotype among the top clonotypes observed. The FFPE TCR assay was performed using 10ng of total RNA, and a workflow developed by MedGenome using the SMARTer TCR α/β Profiling Kit (Takara Bio USA) , and sequencing was performed using the Illumina MiSeq and analysis was performed using the MixCR pipeline (MiLabs).

γ/δ TCR repertoire Profiling as a biomarker & potential therapeutic



C)

% of reads mapped	Gamma PBMC	Gamma CD3	Delta PBMC	Delta CD3	Gamma Delta PBMC	Gamma Delta CD3	Alpha Beta PBMC	Alpha Beta CD3
TRA	0.07	0.03	0.02%	0.06%	0.02	0.08	76.68	73.8
TRB	0	0.01	0.1	0.01	0.2	0.02	23.03	26.1
TRD	0	0	99.85	99.91	67.8	49.84	0	0
TRG	99.92	99.96	0.01	0	31.97	50.06	0	0
Total # clonotypes	1355	1932	203	270	1358	2171	24725	41115

Figure 5: TCR repertoire analysis from γ/δ **Repertoire**: A) Cartoon representation of the γ/δ TCR compared to the α/β TCR, where in contrast to MHC-presentation of antigens, γ/δ T cells are able to recognize stress antigens without MHC-presentation, making them attractive for cancer immunotherapy. They can have a CTL phenotype via engagement with CD8 T cells. B) Shows workflow for testing specificity of γ/δ TCR assay using 10ng of total RNA from PBMCs and CD3 enriched T cells. C) Table shows proof of concept data showing that in TCR libraries where γ/δ primers were added we see specific enrichment of γ/δ chains versus alpha beta and similarly in TCR α/β libraries we see specific amplification of TCR α/β chains. This assay is performed using 10ng of total RNA from PBMCs and CD3 sorted cells, and γ/δ primers designed and optimized by MedGenome using the SMARTer Takara TCR α/β Profiling Kit, and sequencing was performed using the Illumina MiSeq , and analysis was done using MixCR..

Summary of the TCR sequencing services at MedGenome

Table 4 : Summary of the for TCR repertoire sequencing at MedGenome

Technology	Source & Input Type	Required Amount	Sequencing Method	Analysis Method	Information Obtained
Single Cell TCR Immune Profiling (10X Genomics)	Isolated Cells (Fresh, Frozen, Fixed) (Human, Mouse)	Single cell suspension 500-80,000 cells	Illumina NovaSeq PE150	Loupe-browser	CDR3, α/β pairing and clonotypes, V(D)J-sequences
Bulk SMARTer α/β TCR Profiling kit (Takara Bio USA)	Isolated Cells RNA (Human, Mouse)	10 ng-3 ng RNA 50-10,000 cells	Illumina MiSeq PE300	MiXCR VDJTools	CDR3, V(D)J sequences and clonal frequency α/β pairing from single cell optional
Bulk Gamma delta TCR Profiling (MedGenome)	Isolated Cells RNA	10 ng-3 µg RNA 1000-10,000 cells	Illumina MiSeq PE300	MiXCR VDJTools	Full length V(D)J Region sequences with hits, clone count and frequencies
FFPE Bulk TCR Profiling (MedGenome)	FFPE, RNA (DV200>20)	> 10 ng total RNA	Illumina MiSeq PE300	MiXCR VDJTools	Full length V(D)J Region sequences with hits, clone count and frequencies
Single Cell BCR Immune Profiling (10X Genomics)	Isolated Cells (Fresh, Frozen, Fixed) (Human, Mouse)	Single cell suspension 500-80,000 cells	Illumina NovaSeq PE150	Loupe-browser	IgG pairing, clonotypes and V(D) J sequences
Bulk SMARTer BCR IgG Profiling Kit (Takara Bio USA)	Isolated cells RNA (Mouse)	10 ng-3 ng RNA 50-10,000cells	Illumina MiSeq PE300	MiXCR VDJTools	IgG H, K and L gene segments, V(D)J sequences, diversity of receptors, clonotypes, assignment of isotype

Table 5: Summary of the sample types for TCR

Source of Input Material	Approximate Starting number of Cells	Input Amount(s) of RNA	Total TCR alpha/beta clonotypes identified
PBMC Fresh/ frozen (healthy& patients)	500,000	10ng, 100ng	>20,000 clonotypes
White Blood Cells (healthy& patient)	500,000	10ng	>20,000 clonotypes
Sorted CD8+ from lesion~ lymphomas	15,000	1ng	>10,000 clonotypes
Tumor infiltrating lymphocytes	2,000	10 ng	>3,000 clonotypes
FFPE tissue	million	10ng, 50ng	> 100 clonotypes
Mouse T Cell sorted (tumor and normal)	34,000	10 ng	>40,000 clonotypes