haematologica
Journal of the Ferrata Storti Foundation

Single-cell analysis of the T-cell receptor repertoire in untreated myeloma patients suggests potential myeloma-reactive CD8+ T-cells are shared between blood and marrow

by James Favaloro, Christian E. Bryant, Edward Abadir, Samuel Gardiner, Shihong Yang, Tracy King, Najah Nassif, Bronwyn A. O'Brien, Lisa M. Sedger, Richard Boyle, Douglas E. Joshua, and P. Joy Ho

Received: June 29, 2024.
Accepted: September 23, 2024.

Citation: James Favaloro, Christian E. Bryant, Edward Abadir, Samuel Gardiner, Shihong Yang, Tracy King, Najah Nassif, Bronwyn A. O'Brien, Lisa M. Sedger, Richard Boyle, Douglas E. Joshua, and P. Joy Ho. Single-cell analysis of the T-cell receptor repertoire in untreated myeloma patients suggests potential myeloma-reactive CD8+ T-cells are shared between blood and marrow. Haematologica. 2024 Oct 3. doi: 10.3324/haematol.2024.285952 [Epub ahead of print]

Single-cell analysis of the T-cell receptor repertoire in untreated myeloma patients suggests potential myeloma-reactive CD8$^+$ T-cells are shared between blood and marrow

*James Favaloro,[1,2] *Christian E. Bryant,[1,3] Edward Abadir,[1,3] Samuel Gardiner,[4] Shihong Yang,[1] Tracy King,[1,3] Najah Nassif,[2] Bronwyn A. O'Brien,[2] Lisa M. Sedger,[5,6] Richard Boyle,[7] Douglas E. Joshua[1,3] and #P. Joy Ho[1,2,3]

*J.F. and C.E.B. contributed equally to this study.

Affiliations:

[1]. Institute of Hematology, Multiple Myeloma Research Laboratory, New South Wales Health Pathology, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia

[2]. School of Life Sciences, University of Technology Sydney, Ultimo, New South Wales, Australia

[3]. Sydney Medical School, the University of Sydney, Sydney, New South Wales, Australia

[4]. Sydney Local Health District Clinical Research Institute, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia

[5]. Institute for Clinical Pathology and Medical Research (ICPMR), New South Wales Health Pathology, Westmead Hospital, Westmead, New South Wales, Australia

[6]. Center for Virus Research, Westmead Institute for Medical Research, Westmead, New South Wales, Australia

[7]. Orthopedics Department, Sydney Local Health District, Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia

Authorship: J.F. and C.E.B. designed and performed the research, analyzed the data, wrote the human ethics and wrote the paper; E.A. and S.Y. performed the research and analyzed the data; S.G. assisted with the bioinformatics and analyzed the data; N.N., B.A.O and L.M.S. assisted in research design and data analysis; C.E.B., E.A., T.K, and P.J.H. assisted in research design, reviewed patients, and assisted with the collection of patient samples and clinical information; R.B. reviewed patients undergoing hip arthroplasty and designed research; and D.E.J. and P.J.H. were involved in research design, analysis of data and writing the paper.

Short title: Single cell analysis of the TCR repertoire of CD8$^+$ T-cells in Multiple Myeloma

Correspondence: James Favaloro, Royal Prince Alfred Hospital, Level 5, Building 77, Missenden Rd, Camperdown, NSW 2050, Australia; e-mail: james.favaloro@health.nsw.gov.au; and Christian E. Bryant, Royal Prince Alfred Hospital, Level 5, Building 77, Missenden Rd, Camperdown, NSW 2050, ealth.nsw.gov.au.

Data sharing: Data may be made available by contacting the corresponding author.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Acknowledgments: The authors thank Dr Alberto Catalano for excellent research management support, RPAH Haematology Clinicians for collection of patient samples and informed consent, and patients and their families for donating samples for research. They would also like to thank Prof Joseph Powell and Dr Walter Muscovic from the Garvan-Weizmann Centre for Cellular Genomics as well as Dr Slavica Vuckovic for their contributions in the earlier stages of this project.

Sources of funding: This work is funded by a Brian D. Novis research grant from the International Myeloma Foundation (C.E.B.) and Sydney Blood Cancer Research.

Substantial evidence supports the concept that $CD8^+$ T-cells contribute to the control of Multiple Myeloma (MM),[1] a malignancy characterized by the uncontrolled proliferation of neoplastic plasma cells (PCs) in the bone marrow (BM), and an inflammatory tumor microenvironment (TME).[2] We previously demonstrated that T-cell receptor (TCR) Vβ restricted terminally differentiated $CD8^+$ T-cell (T_{TE}) clonal expansions, capable of exerting specific killing of autologous MM cells *in-vitro,* exist within BM and peripheral blood (PB) of MM patients.[3] Although these cells appear to possess a senescent secretory effector phenotype,[4] their presence in PB correlates with improved patient outcomes.[5] Recently, we presented an in-depth multiomic single-cell analysis of paired $CD8^+$ T-cell samples isolated from BM and PB of untreated, newly diagnosed (ND) MM patients,[6] demonstrating that potent cytotoxic effectors with low levels of co-inhibitory molecules reside within the tumor bed. Herein we present the results of a novel bioinformatics pipeline probing the effect of the TME on clonal $CD8^+$ T-cells and investigating the cognate antigens of these cells. We analyzed $CD8^+$ T-cells in the BM and PB of untreated, NDMM patients using single-cell RNA sequencing (scRNA-Seq), inclusive of paired TCR sequencing (TCR-Seq), via a novel bioinformatics workflow (detailed https://github.com/JFavaloro)). We demonstrate that dominant clones are evident in both BM and PB, expanded in both, and are composed of similar T-cell subsets. Dominant clones appear transcriptionally unaffected by the TME, suggesting exploitation of these cells in a disease, which remains incurable,[2] is of immunotherapeutic interest.

Purified $CD8^+$ T-cells isolated from BM and PB of four NDMM patients were subjected to the 10x workflow and processed as previously described (Supp. Table 1).[6] Analysis of scRNA-Seq data was performed with the R package 'Seurat' (v. 4.04) with TCR clonality and diversity assessed with the R packages 'scRepertoire' (v. 1.7.2) and 'Immunarch' (v. 0.7.0), for data with and without paired gene-expression data available, respectively. Clonal bins were established at log 10 intervals and clones were considered expanded if they accounted for >10% of a sample's repertoire, with additional categorisation into large (between 1% and 10%), medium (between 0.1% and 1%) and small (<0.1%). Sequence similarity was performed using the web-based algorithms for predictive TCR-epitope binding, TCRex[7] and TCRMatch.[8] Predictive analysis of antigen-specificity of TCR clonotypes was performed against a list of myeloma antigens using a web-based algorithm. This utilized extended TCR-Peptide Binding Predictor (ERGO-II)[9] in conjunction with a list of 197 immunogenic peptides over-expressed on malignant PCs.[10] Briefly, identified clonotypes were assessed for reactivity using the web-based tool and results with a confidence score of ≥0.9 were selected and assessed for frequency, tissue distribution and, where possible, transcriptome. The study was approved by the

Institutional Human Research Ethics Committee. All patients provided informed consent before sample collection, following the amended Declaration of Helsinki.

Assessment of clonal homeostasis demonstrated that small clonotypes dominated in most samples, with expanded clones present in two samples: PB43 and PB63 (Figure 1A, i). Transcriptional similarities with established clusters (as described in our recent publication)[6] were evident upon projecting data on Uniform Manifold Approximation and Projection (UMAP), with larger expansions demonstrating the expression of genes defining more mature cells and more diverse clonotypes clustering in the region of naïve-like clusters, matching the canonical understanding of T-cell ontogeny (Figure 1A, ii). To determine if clonal cells exist within both the BM and PB in similar ontological states, assessment of cells across clusters and tissue compartment was performed by means of Morisita index. This revealed a high degree of clonal sharing across tissue, primarily in the T_{TE} cluster, with the remaining overlap chiefly evident in the T_{EM} cluster of the PB and T_{EM}, Cyto-T_{EM} and P_{RE}-Ex clusters of the BM (Figure 1B). Focusing on the top 10 dominant clones within an individual (Supp. Table 2), analysis of distribution revealed these to be shared between BM and PB, expanded to a greater degree in the PB in three of four samples, and accounted for a highly variable proportion of the total TCR repertoire between patients (Figure 1C). To determine if differences in the distribution of cell subsets between shared dominant clones were affected by existing within the tumour microenvironment relative to the PB, a ranked pie chart of cluster distribution across the top 10 dominant clones was constructed. This revealed a high degree of compositional overlap; with shared clones appearing compositionally similar irrespective of tissue location (Figure 1D).

To query the potential reactivity of $CD8^+$ T-cells, several computational approaches were employed. Analysis of all clonotypes by TCRex for the MM-associated LLLGIGILV epitope of the HM21.4 antigen identified only a small number of potentially reactive clones, none of which were within the top 10 dominant clones (data not shown). Computational prediction of TCR specificity across the top 10 dominant clones in all samples, undertaken using TCRMatch,[8] suggested potential reactivity against several viral antigens, chiefly cytomegalovirus, Epstein–Barr virus and influenza A, and a minority against non-viral antigens (e.g., gliadin and insulin) (Figure 2A). However, as TCRMatch is limited to reference databases that primarily consist of data on viral-specific T-cells, we leveraged the deep-learning tool, ERGO-II[9] to detect cells potentially reactive to a published list of 197 peptides from 58 proteins known to be presented by class I MHC on malignant plasma cells.[10] After restricting results to "hits" with confidence scores ≥0.9 and clones comprising ≥0.1% of a sample's repertoire, a total of 14 clonotypes were identified against 6 unique peptides from 6 proteins with a degree of gene-

sharing evident in the overrepresentation of *TRBV27* (Table 1). Clones were generally observable within both BM and PB compartments but with greater representation in BM, with most clustered within the T_{EM} cluster, and some clustering in the P_{RE}-EX and Cyto-T_{EM} clusters (Figure 2B, i-vi). However, the most dominant clone in the dataset, accounting for 4.9% and 20.6% of the BM and PB repertoires of NDMM #43, respectively, and identified as potentially reactive against a peptide derived from Influenza A in the previous analysis, and one derived from Cap Methyltransferase 1 (CMTR1) by ERGO-II, demonstrated near exclusive clustering within the T_{TE} cluster (Figure 2B, vii). Differential expression testing of this (Figure 2C and Supp. Table 3), and other (data not shown) dominant clones shared across the BM and PB demonstrated few differences.

Of the 197 peptides assessed by deep learning, three stood out: RPRKAFLLLL, RPFHGWTSL and IHILDVLVL, derived from the proteins disulphide isomerase precursor A 4 (PDIA4), Mannosyl-Oligosaccharide Glucosidase (MOGS) and CMTR1 respectively. These were observed to be shared across multiple patients with evident restriction in TCR-Vβ gene usage including an over-representation of *TRBV27* (Table 1). While PDIA4 has an established role in MM,[11] both MOGS and CMTR1 remain unexploited. MOGS is a transmembrane protein found in the endoplasmic reticulum that catalyzes a reaction essential for immunoglobulin production, while CMTR1 is essential for mRNA stability and is vital for the propagation of Influenza A.[12] Although our analysis revealed that most potential myeloma-reactive $CD8^+$ T-cells reside primarily within the dominant BM-T_{EM} cluster, the most highly expanded clonotype in our dataset, which both TCRMatch and ERGO-II suggested was associated with influenza A, revealed a highly consistent cytotoxic effector transcriptional profile across both BM and PB. These findings highlight that shared public antigens may be implicated in T-cell immunity in MM, as has previously been suggested in melanoma,[13] however, further investigations, including utility of an age-matched control cohort are required.

While these *in-silico* results are not definitive evidence of specificity, our approach of using a disease-specific MHC-presented peptide library, a TCR-library derived from $CD8^+$ T-cells within the tumor bed and machine-learning to look for high probability hits is novel, has utility not only in MM, but also other diseases, fast-tracking peptide selection for use in validation assays. Although speculative, it is possible that identified clonal cells residing primarily in memory clusters either once contributed to disease control until such a time that myeloma clonal evolution rendered these targets mute or represent potential auto-reactive clones that may be activated in the presence of immunomodulatory drugs. Future work should endeavor to assess the transcriptome/proteome of the myeloma tumor itself; although not proof of specificity, this would provide supportive evidence

if it was determined that tumor does indeed express the epitope that dominant $CD8^+$ T-cell clones are reactive against. Further, analysis of paired tumor ligandome by mass spectrometry would grant further confidence in the *in-silico* results and guide functional studies, particularly were this method applied to a disease where mutation is more commonly observed (e.g., melanoma).[13] Clinical trials have demonstrated *ex-vivo* expanded T-cells harvested from the BM of MM patients are able to extend the autologous graft vs. myeloma effect attributable to transplant.[14] In support of this, our data suggests that most potential anti-myeloma $CD8^+$ T-cells likely reside primarily (but not exclusively) within *CD69* expressing clusters in the BM (i.e., the T_{EM}, Cyto T_{EM} and P_{RE}-Ex clusters). Taken together, these data suggest that the gene-expression profiles of dominant T-cell clones are not influenced to an appreciable degree by the TME. Nevertheless, this work highlights two peptides as currently unexploited immunotherapeutic targets in multiple myeloma, warranting further investigation.

References:

1. Vuckovic S, Minnie SA, Smith D, et al. Bone marrow transplantation generates T cell-dependent control of myeloma in mice. J Clin Invest. 2019;129(1):106-121.
2. Rajkumar SV. Multiple Myeloma: 2020 update on Diagnosis, Risk-stratification and Management. Am J Hematol. 2020; 95(5):548-567
3. Vuckovic S, Bryant CE, Lau KHA, et al. Inverse relationship between oligoclonal expanded CD69- TTE and CD69+ TTE cells in bone marrow of multiple myeloma patients. Blood Adv. 2020;4(19):4593-4604.
4. Suen H, Brown R, Yang S, et al. Multiple myeloma causes clonal T-cell immunosenescence: identification of potential novel targets for promoting tumour immunity and implications for checkpoint blockade. Leukemia. 2016;30(8):1716-1724.
5. Bryant C, Suen H, Brown R, et al. Long-term survival in multiple myeloma is associated with a distinct immunological profile, which includes proliferative cytotoxic T-cell clones and a favourable Treg/Th17 balance. Blood Cancer J. 2013;3(9):e148.
6. Favaloro J, Bryant CE, Abadir E, et al. Single-cell analysis of the CD8(+) T-cell compartment in multiple myeloma reveals disease specific changes are chiefly restricted to a CD69(-) subset suggesting potent cytotoxic effectors exist within the tumor bed. Haematologica. 2024;109(4):1220-1232.
7. Gielis S, Moris P, Bittremieux W, et al. Detection of Enriched T Cell Epitope Specificity in Full T Cell Receptor Sequence Repertoires. Front Immunol. 2019;10:2820.
8. Fleri W, Paul S, Dhanda SK, et al. The Immune Epitope Database and Analysis Resource in Epitope Discovery and Synthetic Vaccine Design. Front Immunol. 2017;8:278.
9. Springer I, Besser H, Tickotsky-Moskovitz N, Dvorkin S, Louzoun Y. Prediction of Specific TCR-Peptide Binding From Large Dictionaries of TCR-Peptide Pairs. Front Immunol. 2020;11:1803.
10. Walz S, Stickel JS, Kowalewski DJ, et al. The antigenic landscape of multiple myeloma: mass spectrometry (re)defines targets for T-cell-based immunotherapy. Blood. 2015;126(10):1203-1213.
11. Robinson RM, Reyes L, Duncan RM, et al. Inhibitors of the protein disulfide isomerase family for the treatment of multiple myeloma. Leukemia. 2019;33(4):1011-1022.
12. Li B, Clohisey SM, Chia BS, et al. Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. Nat Commun. 2020;11(1):164.
13. Snyder A, Makarov V, Merghoub T, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med. 2014;371(23):2189-2199.
14. Noonan KA, Huff CA, Davis J, et al. Adoptive transfer of activated marrow-infiltrating lymphocytes induces measurable antitumor immunity in the bone marrow in multiple myeloma. Sci Transl Med. 2015;7(288):288ra278.

Table 1:
Results of ERGO-II peptide prediction against peptides demonstrated to be overexpressed on malignant Plasma Cells

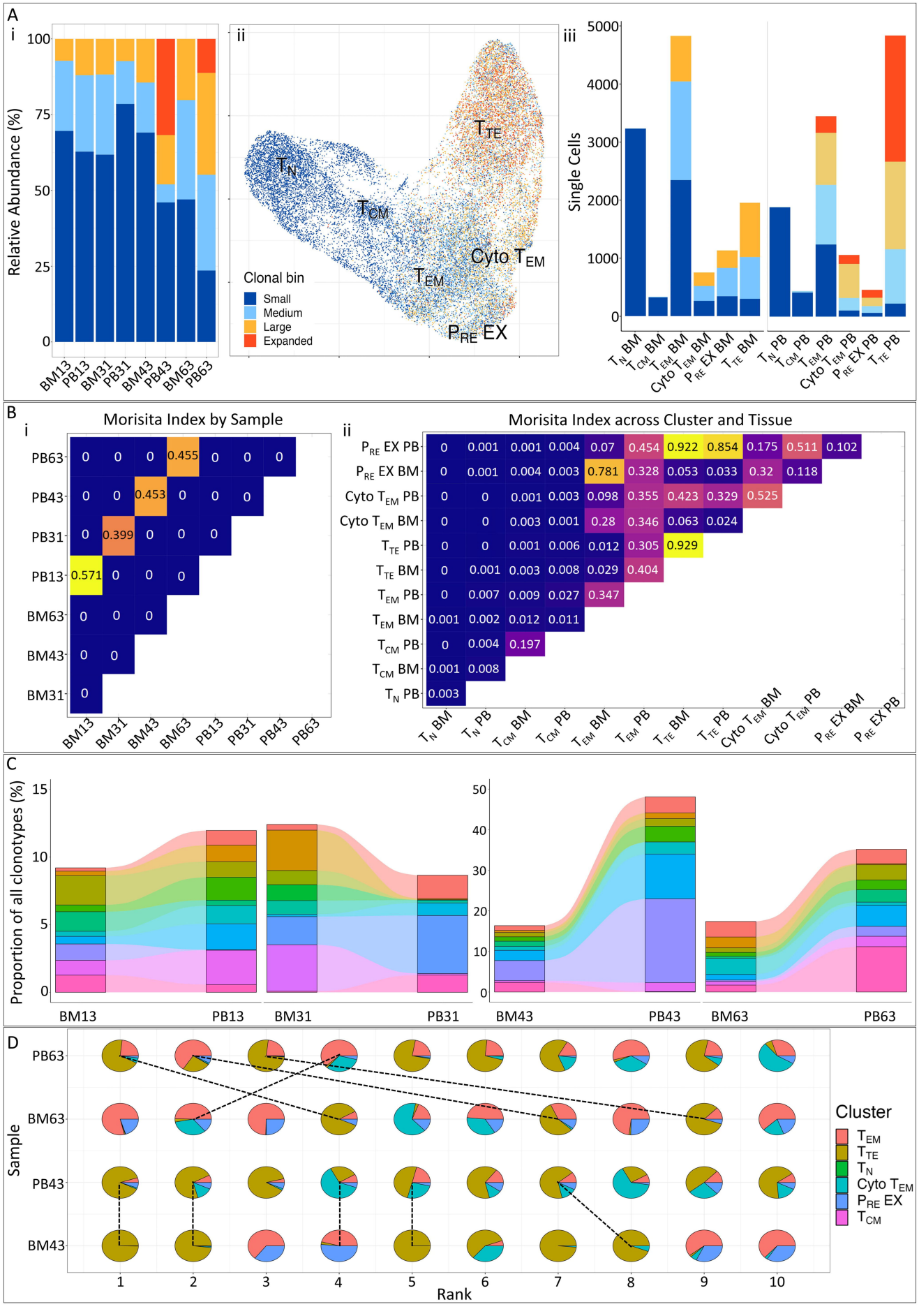
Patient ID	Tissue	[I]Size (%) repertoire	Clone	Confidence	Protein	Peptide	Dominant Cluster
NDMM #13	BM	0.396	*TRAV12-1;* CVVPWYSSASKIIF; *TRAJ3*	0.93	PDIA4	RPRKAFLLLL	N/A
	PB	0.012	***TRBV27***; CASGTGQNQPQHF; *TRBJ1-5*				N/A
NDMM #13	BM	0.247	*TRAV38-2/DV8;* CAYTSGTYKYIF; *TRAJ40*	0.91	PDIA4	RPRKAFLLLL	N/A
	PB	0.299	***TRBV27***; CASSLSPVNYGYTF; *TRBJ1-2*				N/A
NDMM #13	BM	0.04	*TRAV19*; CALHNAGKSTF; *TRAJ27*	0.91	MOGS	RPFHGWTSL	N/A
	PB	0.383	***TRBV27***; CASSLQRNTEAFF; *TRBJ1-1*				N/A
NDMM #13	BM	0.025	*TRAV27*; CAGPSGNTGKLIF; *TRAJ37*	0.92	MOGS	RPFHGWTSL	N/A
	PB	0.23	***TRBV27***; CASSLSRRVGSYGYTF; *TRBJ1-2*				N/A
NDMM #31	BM	0.128	*TRAV14/DV4;* CAMREPLNNAGNMLTF; *TRAJ39*	0.9	MOGS	RPFHGWTSL	N/A
	PB	N/A	***TRBV27***; CASSLGGGWTEAFF; *TRBJ1-1*				N/A
NDMM #31	BM	0.176	*TRAV22*; CALTDSWGKLQF; *TRAJ24*	0.94	CMTR1	IHILDVLVL	N/A
	PB	N/A	*TRBV5-4*; CASLPYSGANVLTF; *TRBJ2-6*				N/A
NDMM #31	BM	0.064	*TRAV8-1*; CAVIGFQKLVF; *TRAJ8*	0.91	LAP3	DVNNIGKYR	N/A
	PB	0.815	***TRBV27***; CASSLTASHYGYTF; *TRBJ1-2*				N/A
NDMM #43	BM	4.929	*TRAV26-2*; CILRDNGGKSTF; *TRAJ27*	0.94	CMTR1	IHILDVLVL	T_{TE}
	PB	20.567	*TRBV19*; CASSIWGTSNQPQHF; *TRBJ1-5*				T_{TE}
NDMM #43	BM	0.473	*TRAV1-2*; CAVRSLYNFNKFYF; *TRAJ21*	0.91	MOGS	RPFHGWTSL	T_{EM}
	PB	0.014	***TRBV27***; CASNDRGTDTQYF; *TRBJ2-3*				T_{EM}
NDMM #43	BM	0.3	*TRAV4*; CLVGDNTNAGKSTF; *TRAJ27*	0.97	CMTR1	IHILDVLVL	T_{EM}
	PB	0.014	*TRBV12-4*; CASRPSSGRGYNEQFF; *TRBJ2-1*				T_{EM}
NDMM #63	BM	0.674	*TRAV38-2/DV8*; CAYRSATHDMRF; *TRAJ43*	0.96	ASS1	NIGQKEDFEEA	T_{EM}
	PB	0.059	*TRBV12-4*; CASSPRDFWETQYF; *TRBJ2-5*				T_{EM}
NDMM #63	BM	0.196	*TRAV13-1*; CAASSPSNDYKLSF; *TRAJ20*	0.94	NDUFAF4	APRHPSTNSL	T_{EM}
	PB	0.03	*TRBV2*; CASIGGTYLL; *TRBJ2-1*				T_{EM}
NDMM #63	BM	0.174	*TRAV27*; CAGAFRGSNYKLTF; *TRAJ53*	0.93	CMTR1	IHILDVLVL	T_{EM}
	PB	0.015	*TRBV19*; CASRVAEGAYDPAFF; *TRBJ1-1*				T_{EM}
NDMM #63	BM	0.196	*TRAV9-2; CALSGIAGFQKLVF; TRAJ8*	0.9	PDIA4	RPRKAFLLLL	T_{EM}
	PB	0.03	*TRBV9; CASSATRPGQLNSPLHF; TRBJ1-6* ***TRBV27****; CASSSLDRPGDGYTF; TRBJ1-2*				T_{EM}

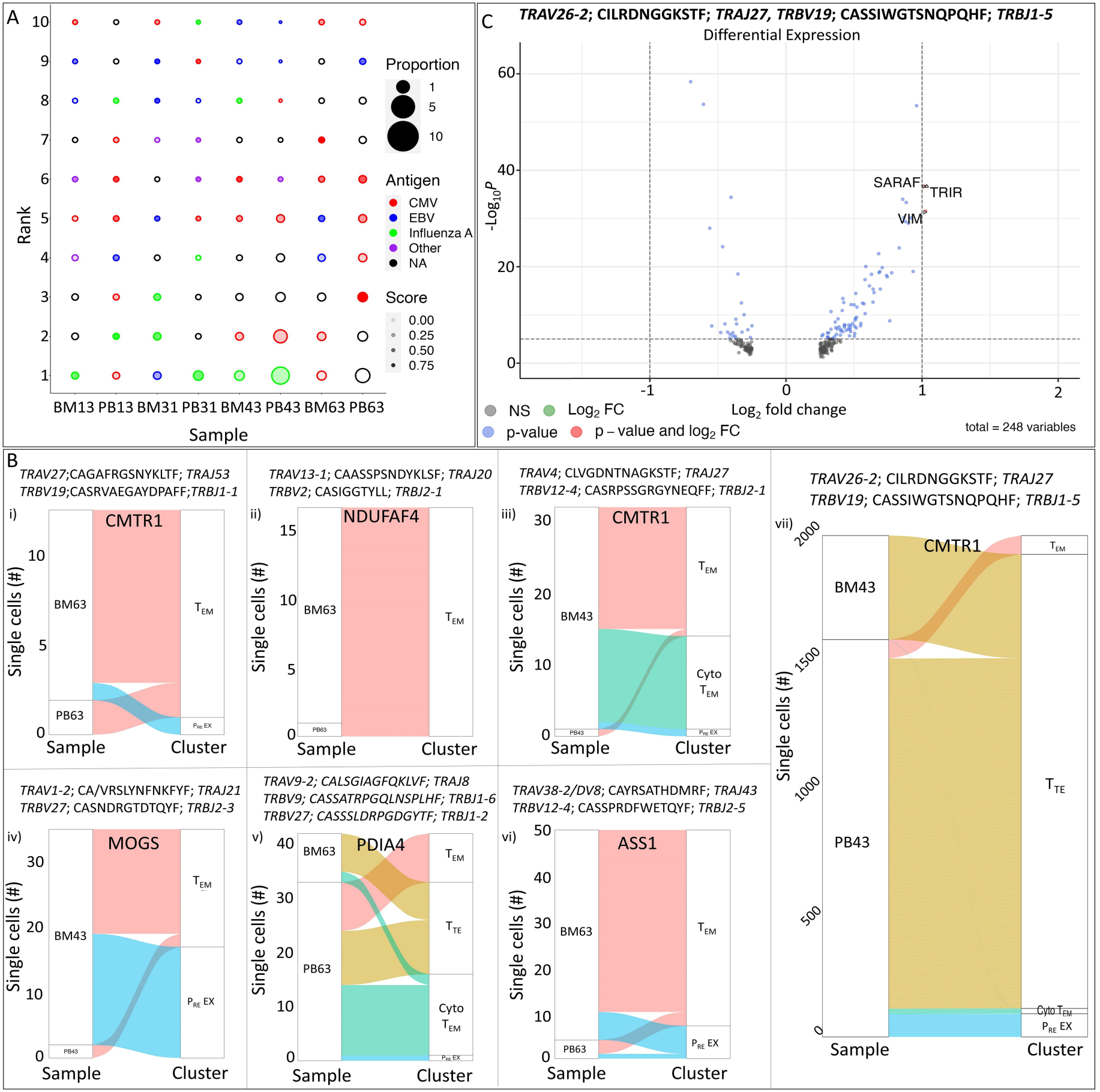
[I]Proportion of repertoire is derived from Immunarch analysis of raw data with non-productive clones removed. Clones with *TRBV27* are bolded.

Key: NDMM, Newly Diagnosed Multiple Myeloma; BM, Bone Marrow; PB, Peripheral Blood; T_{EM}, Effector memory T-cells; T_{TE}; Terminally differentiated T-cells; N/A, Not Applicable; ASS1, Argininosuccinate Synthase 1; CMTR1, Cap Methyltransferase 1; LAP3 Leucine Aminopeptidase 3; MOGS, Mannosyl-Oligosaccharide Glucosidase; NDUFAF4, Ubiquinone Oxidoreductase Complex Assembly Factor 4; PDIA4, Protein Disulphide Isomerase Family A Member 4.

Figure 1. Analysis of T-Cell Receptor clonal overlap of CD8$^+$ T-cells between Bone Marrow and Peripheral Blood of Newly Diagnosed Multiple Myeloma patients shows inconsistency in frequency but compositional similarity A) i) Clonal space homeostasis of T-cell receptor (TCR) clonotypes in paired bone marrow (BM) and peripheral blood (PB) CD8$^+$ T-cells across small (<0.1% of the sample's TCR repertoire); dark blue, medium (0.1%-1.0%); light blue, large (1.0%-10.0%); yellow and expanded (≥ 10%); red) clonal bins. **ii)** Uniform Manifold Approximation and Projection (UMAP) of the Combined single cell object coloured according to clonal bin. single cell (sc)RNA-Sequencing data without matching scTCR-Sequencing data is highlighted in grey (NA). **iii)** Stacked bar plots of clonal bin distribution across identified clusters in the BM (left) and PB (right) of two newly diagnosed multiple myeloma (NDMM) patients (#43 and #63). **B)** Clonal overlap across tissue and cluster as measured by the Morisita Index. Higher numbers (coloured yellow) indicate a greater level of clonal sharing, low numbers (coloured blue) indicate low/no overlap. **C)** Alluvial graphs demonstrating proportion of the top 10 observed TCR clonotypes between paired BM/PB samples of four NDMM patients (#13 far left, #31 left, #43 right and #63 far right). Each colour represents a TCR clonotype observable in both tissues. The size of the bar represents the proportion (y-axis) of the repertoire of either the BM (left) or PB (right). **D)** Pie chart of the top 10 dominant TCR clonotypes in paired BM and PB CD8$^+$T-cells from NDMM patients (#043 and #063). Colors represent the cluster to which each clone belongs, ranked 1-10 (x-axis) by level of expansion. The dashed line represents clones shared between BM and PB evident in the top 10 of each sample. TCR analysis of NDMM patients demonstrates enriched BM diversity and evident clonal sharing.

Figure 2. Results of computational predication of reactivity and transcriptional analysis of identified clones. A) Bubble plot illustrating potential reactivity of the top 10 dominant T-cell receptor (TCR) clonotypes in paired bone marrow (BM) and peripheral blood (PB) CD8$^+$ T-cells from newly diagnosed multiple myeloma (NDMM) patients (#013, #031, #043 and #063). The size of the dot represents the proportion (as a percentage) of all observed clonotypes within that individual sample, ranked 1-10 (y-axis) based on level of expansion. The opacity of the dot represents the confidence (darker implies greater confidence) to which a particular clonotype is reactive to a particular antigen (Cytomegalovirus (CMV), red; Epstein-Barr virus (EBV), blue; Influenza A, green; Other (e.g., yellow fever), purple). N/A refers to clonotypes to which no data was recovered. **B)** Alluvial graphs of clones identified by deep learning as being potentially myeloma reactive arranged by size (smallest to greatest; i-vii). Each line represents the tissue of origin of a single cell (left alluvium) colored by the cluster in which it resides (right alluvium). Each pair is listed with the identified TCR sequence along with the protein from which the peptide the clone is potentially reactive to is derived. **C)** Volcano plot illustrating differential expression results comparing the most dominant clone in NDMM#43 across the BM and PB. Red dots represent genes that are both significantly upregulated and expressed at a Log2 fold change (FC) value of >1 in the BM (right) relative to the PB or <1 in the PB (left) relative to the BM. Blue dots represent genes that are significantly different but not expressed at a great enough level (i.e., >1 or <1 Log2 FC). Green dots represent genes that are highly expressed (i.e., >1 or <1 Log2 FC) but found to be not significantly differentially expressed between the two tissues. Black dots represent tested genes that are neither highly expressed nor significantly different between the two tissues.


A
i
Relative Abundance (%)
100
75
50
25
0
BM13
PB13
BM31
PB31
BM43
PB43
BM63
PB63
ii
T_N
T_{CM}
T_{TE}
Cyto T_{EM}
T_{EM}
P_{RE} EX
Clonal bin
Small
Medium
Large
Expanded
iii
Single Cells
5000
4000
3000
2000
1000
0
T_N BM
T_{CM} BM
T_{EM} BM
Cyto T_{EM} BM
P_{RE} EX BM
T_{TE} BM
T_N PB
T_{CM} PB
T_{EM} PB
Cyto T_{EM} PB
P_{RE} EX PB
T_{TE} PB
B
i
Morisita Index by Sample
PB63 0 0 0 0.455 0 0 0
PB43 0 0 0.453 0 0 0
PB31 0 0.399 0 0 0
PB13 0.571 0 0 0
BM63 0 0 0
BM43 0 0
BM31 0
BM13 BM31 BM43 BM63 PB13 PB31 PB43 PB63
ii
Morisita Index across Cluster and Tissue
P_{RE} EX PB 0 0.001 0.001 0.004 0.07 0.454 0.922 0.854 0.175 0.511 0.102
P_{RE} EX BM 0 0.001 0.004 0.003 0.781 0.328 0.053 0.033 0.32 0.118
Cyto T_{EM} PB 0 0 0.001 0.003 0.098 0.355 0.423 0.329 0.525
Cyto T_{EM} BM 0 0 0.003 0.001 0.28 0.346 0.063 0.024
T_{TE} PB 0 0 0.001 0.006 0.012 0.305 0.929
T_{TE} BM 0 0.001 0.003 0.008 0.029 0.404
T_{EM} PB 0 0.007 0.009 0.027 0.347
T_{EM} BM 0.001 0.002 0.012 0.011
T_{CM} PB 0 0.004 0.197
T_{CM} BM 0.001 0.008
T_N PB 0.003
T_N BM T_N PB T_{CM} BM T_{CM} PB T_{EM} BM T_{EM} PB T_{TE} BM T_{TE} PB Cyto T_{EM} BM Cyto T_{EM} PB P_{RE} EX BM P_{RE} EX PB
C
Proportion of all clonotypes (%)
15
10
5
0
50
40
30
20
10
0
BM13
PB13
BM31
PB31
BM43
PB43
BM63
PB63
D
Sample
PB63
BM63
PB43
BM43
1 2 3 4 5 6 7 8 9 10
Rank
Cluster
T_{EM}
T_{TE}
T_N
Cyto T_{EM}
P_{RE} EX
T_{CM}


A
Rank
Sample
BM13 PB13 BM31 PB31 BM43 PB43 BM63 PB63
Proportion
1
5
10
Antigen
CMV
EBV
Influenza A
Other
NA
Score
0.00
0.25
0.50
0.75
C
TRAV26-2; CILRDNGGKSTF; TRAJ27, TRBV19; CASSIWGTSNQPQHF; TRBJ1-5
Differential Expression
-Log10P
Log2 fold change
SARAF
TRIR
VIM
NS
Log2 FC
p-value
p – value and log2 FC
total = 248 variables
B
TRAV27;CAGAFRGSNYKLTF; TRAJ53
TRBV19;CASRVAEGAYDPAFF;TRBJ1-1
i)
CMTR1
Single cells (#)
BM63
PB63
TEM
PRE EX
Sample
Cluster
TRAV13-1; CAASSPSNDYKLSF; TRAJ20
TRBV2; CASIGGTYLL; TRBJ2-1
ii)
NDUFAF4
TRAV4; CLVGDNTNAGKSTF; TRAJ27
TRBV12-4; CASRPSSGRGYNEQFF; TRBJ2-1
iii)
CMTR1
BM43
Cyto TEM
TRAV26-2; CILRDNGGKSTF; TRAJ27
TRBV19; CASSIWGTSNQPQHF; TRBJ1-5
vii)
CMTR1
BM43
PB43
TTE
Cyto TEM
PRE EX
TRAV1-2; CA/VRSLYNFNKFYF; TRAJ21
TRBV27; CASNDRGTDTQYF; TRBJ2-3
iv)
MOGS
TRAV9-2; CALSGIAGFQKLVF; TRAJ8
TRBV9; CASSATRPGQLNSPLHF; TRBJ1-6
TRBV27; CASSSLDRPGDGYTF; TRBJ1-2
v)
PDIA4
TRAV38-2/DV8; CAYRSATHDMRF; TRAJ43
TRBV12-4; CASSPRDFWETQYF; TRBJ2-5
vi)
ASS1

Supp. Table 1 Summary of recovered Chromium™ 10x data following QC

Sample ID	Sample	Number of cell barcodes recovered	Ave. number of detected genes per cell	Ave. number of recovered UMIs per cell	Mito genes (%)	Total TCR Clonotypes observed (n/%)	Cells with matching TCR data (n/%)
NDMM #13	[I]BM	7871	732	1625	1.9	4,455 (100)	6551 (83.2)
	PB	7822	1267	3966	4.1	3,876 (100)	7395 (94.5)
NDMM #31	[I]BM	2429	375	256	3.22	3,103 (100)	2317 (95.4)
	PB	7558	1098	3218	5.29	4,838 (100)	6559 (86.8)
NDMM #43	BM	7120	1005	3317	6.51	5,063 (100)	6869 (96.5)
	PB	6497	1097	3690	7.13	2,975 (100)	6207 (95.5)
NDMM #63	BM	7824	1251	3572	5.23	3,259 (100)	7472 (95.5)
	PB	6275	1374	3933	4.85	1,239 (100)	6147 (98.0)

[I]Gene expression library failed QC, excluded from further analysis.

Key: BM, Bone Marrow; PB, Peripheral Blood; NDMM, Newly Diagnosed Multiple Myeloma; ID, Identification; Mito, Mitochondrial; n, number; TCR, T-cell Receptor; UMIs, Unique Molecular Identifiers.

Supp. Table 2 Top 10 clones ranked by size of expansion (raw data)

BM13					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
1	182	2.25%	*TRAV12-2;TRBV7-8*	CAVKGNAGKSTF;CASSLESGGTEAFF	*TRAJ27;TRBJ1-1*
2	109	1.35%	*TRAV14/DV4;TRBV6-6*	CAMREGAEGAQKLVF;CASSYPSSGSLYNEQFF	*TRAJ54;TRBJ2-1*
3	99	1.22%	*TRAV29/DV5;TRBV7-9*	CAYGAGSYQLTF;CASSLIERETDTQYF	*TRAJ28;TRBJ2-3*
4	88	1.09%	*TRAV38-2/DV8;TRBV7-9*	CAYRRLGANNLFF;CASSPLDGRGTDTQYF	*TRAJ36;TRBJ2-3*
5	85	1.05%	*TRAV38-2/DV8;TRAV23/DV6;TRBV4-1*	CAYRGLFGNEKLTF;CAAKDPLMNRDDKIIF;CASSQDGTGEAKNIQYF	*TRAJ48;TRAJ30;TRBJ2-4*
6	63	0.78%	*TRAV12-2;TRBV2*	CAVMRDGGSQGNLIF;CASNRPQSGGSSYNSPLHF	*TRAJ42;TRBJ1-6*
7	49	0.61%	*TRBV10-3*	CAISEFGGTYEQYF	*TRBJ2-7*
8	48	0.59%	*TRAV38-2/DV8;TRBV4-1*	CAYRGLFGNEKLTF;CASSQDGTGEAKNIQYF	*TRAJ48;TRBJ2-4*
9	43	0.53%	*TRAV14/DV4;TRBV19*	CAMRDPSGGSYIPTF;CASSIDRAGAYNEQFF	*TRAJ6;TRBJ2-1*
10	39	0.48%	*TRAV13-1;TRBV3-1*	CAASPAQLTF;CASSQEGATGNEKLFF	*TRAJ22;TRBJ1-4*
PB13					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
1	209	2.50%	*TRAV38-2/DV8;TRBV7-9*	CAYRRLGANNLFF;CASSPLDGRGTDTQYF	*TRAJ36;TRBJ2-3*
2	141	1.69%	*TRAV13-1;TRBV3-1*	CAASPAQLTF;CASSQEGATGNEKLFF	*TRAJ22;TRBJ1-4*
3	111	1.33%	*TRAV24;TRBV7-6*	CAFIPNTNAGKSTF;CASSFLLAGADTQYF	*TRAJ27;TRBJ2-3*
4	110	1.32%	*TRAV29/DV5;TRBV2*	CAGSGGGNKLTF;CASSVGTDEQYF	*TRAJ10;TRBJ2-7*
5	101	1.21%	*TRAV5;TRBV6-5*	CAEMGNYGQNFVF;CASSPGALNYGYTF	*TRAJ26;TRBJ1-2*
6	101	1.21%	*TRAV12-1;TRBV5-1*	CVVRTDKLIF;CASSPDSQSSGNTIYF	*TRAJ34;TRBJ1-3*
7	93	1.11%	*TRAV12-2;TRBV7-8*	CAVKGNAGKSTF;CASSLESGGTEAFF	*TRAJ27;TRBJ1-1*
8	85	1.02%	*TRAV19;TRBV6-2*	CALSEADSGYSTLTF;CASSYWDGRVNTEAFF	*TRAJ11;TRBJ1-1*
9	71	0.85%	*TRAV12-2;TRBV2*	CAVMRDGGSQGNLIF;CASNRPQSGGSSYNSPLHF	*TRAJ42;TRBJ1-6*
10	69	0.83%	*TRAV20;TRBV9*	CAVRVGLTGGGNKLTF;CASSSLGDSNQPQHF	*TRAJ10;TRBJ1-5*

BM31					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
1	214	3.42%	*TRAV38-2/DV8;TRBV5-6*	CAYRRSNNDMRF;CASSLGAFTIYF	*TRAJ43;TRBJ1-3*
2	132	2.11%	*TRAV35;TRBV19*	CAGPVGDMRF;CASSVAGVRETQYF	*TRAJ43;TRBJ2-5*
3	130	2.08%	*TRAV12-1;TRBV9*	CVVNVLDMRF;CASSVGINEKLFF	*TRAJ43;TRBJ1-4*
4	72	1.15%	*TRAV20;TRBV2*	CAVQLDGQKLLF;CASRSRANVLTF	*TRAJ16;TRBJ2-6*
5	64	1.02%	*TRAV12-3;TRBV11-3*	CASWNSGYSTLTF;CASSLDGVTQFF	*TRAJ11;TRBJ2-1*
6	53	0.85%	*TRAV12-1;TRAV16;TRBV9*	CVVNVLDMRF;CSHNNAGNMLTF;CASSVGINEKLFF	*TRAJ43;TRAJ39;TRBJ1-4*
7	48	0.77%	*TRAV20;TRBV6-5*	CAVDMDSNYQLIW;CASSYGRAKGHFF	*TRAJ33;TRBJ2-1*
8	42	0.67%	*TRAV21;TRBV21-1;TRBV25-1*	CAAPNSGGGADGLTF;CASSKAAGQGSETQYF;CASSDSGANVLTF	*TRAJ45;TRBJ2-5;TRBJ2-6*
9	37	0.59%	*TRAV26-1;TRAV35;TRBV28*	CIVRGHGSSNTGKLIF;CAAPGTLTGNQFYF;CASSLGGPPAYGYTF	*TRAJ37;TRAJ49;TRBJ1-2*
10	35	0.56%	*TRAV14/DV4;TRBV19*	CAIENTDKLIF;CASSTRRAVGEKLFF	*TRAJ34;TRBJ1-4*
PB31					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
1	340	4.70%	*TRAV35;TRBV19*	CAGPVGDMRF;CASSVAGVRETQYF	*TRAJ43;TRBJ2-5*
2	95	1.31%	*TRBV19*	CASSVAGVRETQYF	*TRBJ2-5*
3	87	1.20%	*TRAV9-2;TRBV18*	CALSEIATGRRALTF;CAGSHPTGVLNGYTF	*TRAJ5;TRBJ1-2*
4	67	0.93%	*TRAV21;TRBV28*	CAAAGTGGFKTIF;CASSSSTSGSGLNEQFF	*TRAJ9;TRBJ2-1*
5	48	0.66%	*TRAV14/DV4;TRBV12-4*	CAMRERQAGTALIF;CASSGTANYGYTF	*TRAJ15;TRBJ1-2*
6	48	0.66%	*TRAV38-1;TRBV10-3*	CAFMKHSGFGNVLHC;CAISGGNTGELFF	*TRAJ35;TRBJ2-2*
7	45	0.62%	*TRAV13-2;TRBV7-9*	CAENTNAGKSTF;CASSLGGFEQYF	*TRAJ27;TRBJ2-7*
8	38	0.53%	*TRAV38-2/DV8;TRBV25-1*	CAYGLGGGADGLTF;CASTFGTGTDTQYF	*TRAJ45;TRBJ2-3*
9	33	0.46%	*TRAV12-2;TRBV4-3*	CAVNTGGFKTIF;CASSQDVASGAGELFF	*TRAJ9;TRBJ2-2*
10	33	0.46%	*TRAV38-2/DV8;TRBV7-9*	CAWWDNYGQNFVF;CASSHPDRVEKLFF	*TRAJ26;TRBJ1-4*

BM43					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
1	345	4.32%	*TRAV26-2;TRBV19*	CILRDNGGKSTF;CASSIWGTSNQPQHF	*TRAJ27;TRBJ1-5*
2	192	2.40%	*TRAV8-6;TRBV7-3*	CAVRVDYKLSF;CASSPNPGDNYGYTF	*TRAJ20;TRBJ1-2*
3	164	2.05%	*TRAV21;TRBV6-5*	CAEGTGNQFYF;CASRGTGPPYEQYF	*TRAJ49;TRBJ2-7*
4	100	1.25%	*TRAV13-1;TRBV19*	CAAYDYKLSF;CASTGGRTDTQYF	*TRAJ20;TRBJ2-3*
5	79	0.99%	*TRAV17;TRBV4-1*	CATDAGSSGNTPLVF;CASSQAGRGTTYNEQFF	*TRAJ29;TRBJ2-1*
6	76	0.95%	*TRAV38-2/DV8;TRBV2*	CAYRSADSGYALNF;CASSQAVSGGNSPLHF	*TRAJ41;TRBJ1-6*
7	69	0.86%	*TRBV6-5*	CASRGTGPPYEQYF	*TRBJ2-7*
8	67	0.84%	*TRAV26-2;TRAV29/DV5;TRBV19*	CILRDNGGKSTF;CTSVQRPGNTGKLIF;CASSIWGTSNQPQHF	*TRAJ27;TRAJ37;TRBJ1-5*
9	60	0.75%	*TRBV19*	CASSIWGTSNQPQHF	*TRBJ1-5*
10	52	0.65%	*TRAV12-2;TRAV9-2;TRBV28*	CAALSWGKLQF;CAPYNTNAGKSTF;CASRVDRMSNQPQHF	*TRAJ24;TRAJ27;TRBJ1-5*
PB43					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
1	1206	17.46%	*TRAV26-2;TRBV19*	CILRDNGGKSTF;CASSIWGTSNQPQHF	*TRAJ27;TRBJ1-5*
2	625	9.05%	*TRAV21;TRBV6-5*	CAEGTGNQFYF;CASRGTGPPYEQYF	*TRAJ49;TRBJ2-7*
3	253	3.66%	*TRAV26-2;TRAV29/DV5;TRBV19*	CILRDNGGKSTF;CTSVQRPGNTGKLIF;CASSIWGTSNQPQHF	*TRAJ27;TRAJ37;TRBJ1-5*
4	209	3.03%	*TRAV17;TRBV4-1*	CATDAGSSGNTPLVF;CASSQAGRGTTYNEQFF	*TRAJ29;TRBJ2-1*
5	151	2.19%	*TRBV19*	CASSIWGTSNQPQHF	*TRBJ1-5*
6	150	2.17%	*TRAV1-2;TRBV24-1*	CALRGDYKLSF;CATSDSPRTSGNNEQFF	*TRAJ20;TRBJ2-1*
7	141	2.04%	*TRDV1;TRAV21;TRBV6-5*	CSWGGQVMNYGGSQGNLIF;CAEGTGNQFYF;CASRGTGPPYEQYF	*TRAJ42;TRAJ49;TRBJ2-7*
8	110	1.59%	*TRBV6-5*	CASRGTGPPYEQYF	*TRBJ2-7*
9	104	1.51%	*TRAV1-2;TRAV10;TRBV24-1*	CALRGDYKLSF;CGEREGNARLMF;CATSDSPRTSGNNEQFF	*TRAJ20;TRAJ31;TRBJ2-1*
10	99	1.43%	*TRAV38-2/DV8;TRBV10-3*	CAYRSADSNYQLIW;CAISESVSGAGNTIYF	*TRAJ33;TRBJ1-3*

BM63					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
1	372	4.04%	*TRAV19;TRBV20-1*	CALRTPTNSNSGYALNF;CSASGPAEINEQFF	*TRAJ41;TRBJ2-1*
2	311	3.38%	*TRAV12-1;TRBV28*	CVVNWRSNDYKLSF;CASSFPSGGVSTDTQYF	*TRAJ20;TRBJ2-3*
3	219	2.38%	*TRAV12-1;TRBV20-1*	CVVGDWFGDMRF;CSALKPGTSSYNEQFF	*TRAJ43;TRBJ2-1*
4	152	1.65%	*TRAV14/DV4;TRBV7-6*	CAMREVESNMLTF;CASSTFSYEQYF	*TRAJ39;TRBJ2-7*
5	148	1.61%	*TRAV8-1;TRBV11-2*	CAVNHYNTDKLIF;CASSLDYFSGNTIYF	*TRAJ34;TRBJ1-3*
[!]6	129	1.40%	*TRAV23/DV6;TRBV4-3*	CAASIGNFGNEKLTF;CASSPSRNTEAFF	*TRAJ48;TRBJ1-1*
7	107	1.16%	*TRAV14/DV4;TRBV12-4*	CAMREAITQGGSEKLVF;CASRGGWGSPLHF	*TRAJ57;TRBJ1-6*
8	101	1.10%	*TRAV8-6;TRBV27*	CAVSDHNNARLMF;CASSWVRGDTGELFF	*TRAJ31;TRBJ2-2*
9	98	1.06%	*TRBV20-1*	CSASGPAEINEQFF	*TRBJ2-1*
10	92	1.00%	*TRAV13-1;TRBV28*	CAAPVDFGNEKLTF;CASSQISGGNTIYF	*TRAJ48;TRBJ1-3*
PB63					
Rank	**Clones**	**Proportion**	**Variable genes**	**CDR3 (aa)**	**Joining genes**
[!]1	359	5.32%	*TRAV23/DV6;TRBV4-3*	CAASIGNFGNEKLTF;CASSPSRNTEAFF	*TRAJ48;TRBJ1-1*
[@]2	224	3.32%	*TRAV8-6;TRAV8-6;TRBV28*	CAVPPTGGGNKLTF;CAVSDRSGGGADGLTF;CASSLGLHYEQYV	*TRAJ10;TRAJ45;TRBJ2-7*
[@]3	223	3.31%	*TRAV8-6;TRAV8-6;TRBV28*	CAVSDRSGGGADGLTF;CAVPPTGGGNKLTF;CASSLGLHYEQYV	*TRAJ45;TRAJ10;TRBJ2-7*
4	209	3.10%	*TRAV19;TRBV20-1*	CALWYGRDDKIIF;CSAKVNTEAFF	*TRAJ30;TRBJ1-1*
5	198	2.93%	*TRAV12-1;TRBV28*	CVVNWRSNDYKLSF;CASSFPSGGVSTDTQYF	*TRAJ20;TRBJ2-3*
6	189	2.80%	*TRAV13-1;TRBV28*	CAAPVDFGNEKLTF;CASSQISGGNTIYF	*TRAJ48;TRBJ1-3*
7	173	2.56%	*TRAV29/DV5;TRBV6-2*	CAARNTGNQFYF;CASSSRPGPSGYF	*TRAJ49;TRBJ2-7*
8	171	2.53%	*TRAV14/DV4;TRBV12-4*	CAMREAITQGGSEKLVF;CASRGGWGSPLHF	*TRAJ57;TRBJ1-6*
[@]9	156	2.31%	*TRAV8-6;TRAV8-6;TRBV28;TRBV3-1*	CAVPPTGGGNKLTF;CAVSDRSGGGADGLTF; CASSLGLHYEQYV;CISVPAAKSFPGTTSSYEQYF	*TRAJ10;TRAJ45;TRBJ2-7;TRBJ2-7*
[@]10	145	2.15%	*TRAV8-6;TRAV8-6;TRBV28;TRBV3-1*	CAVSDRSGGGADGLTF;CAVPPTGGGNKLTF; CASSLGLHYEQYV;CISVPAAKSFPGTTSSYEQYF	*TRAJ45;TRAJ10;TRBJ2-7;TRBJ2-7*

[!]Clone found to be specific (with 100% confidence) for the CMV derived epitope, RPHERNGFTVL by all tested methods. [@]Example of difficulties in reconstructing paired data; clone is identical differing only in order of the 2nd TCRα J gene and the inclusion of a second TCRβ.

Key: BM, Bone Marrow; PB, Peripheral Blood; CDR3, Complementarity-determining region 3

Supp. Table 3 Top 20 differentially expressed genes comparing the dominant clone in PT43 between the BM and PB

Gene	P value	Average Log_2FC	Proportion of cells expressing gene in the bone marrow (%)	Proportion of cells expressing gene in the peripheral blood (%)	Adjusted P value
HLA-A	4.41×10^{-59}	-0.70	99.1	100	7.31×10^{-55}
HLA-C	2.11×10^{-54}	-0.60	99.7	100	3.50×10^{-50}
PPDPF	4.11×10^{-54}	0.96	93.3	77.8	6.81×10^{-50}
SARAF	2.18×10^{-37}	1.01	72.9	47.4	3.6×10^{-33}
TRIR	2.19×10^{-37}	1.036	46.2	16.9	3.62×10^{-33}
HLA-B	4.06×10^{-35}	-0.40	100	100	6.72×10^{-31}
S100A10	1.06×10^{-34}	0.89	86.0	65.2	1.75×10^{-30}
C9orf16	4.91×10^{-34}	0.88	62.9	34.3	8.13×10^{-30}
VIM	3.19×10^{-32}	1.035	72.9	51.9	5.27×10^{-28}
NDUFB10	6.85×10^{-31}	0.87	51.4	24.9	1.13×10^{-26}
C12orf75	7.73×10^{-31}	0.921	59.6	33.3	1.28×10^{-26}
ISCU	4.40×10^{-30}	0.88	31.0	8.6	7.287×10^{-26}
PAXX	9.54×10^{-30}	0.90	62.0	35.6	1.58×10^{-25}
IL32	1.05×10^{-28}	-0.56	94.2	98.9	1.74×10^{-24}
CD52	7.25×10^{-25}	-0.466	97.6	99.5	1.20×10^{-20}
LIME1	1.27×10^{-24}	0.83	40.4	17.2	2.11×10^{-20}
CD8B	2.01×10^{-23}	0.68	73.3	54.3	3.33×10^{-19}
CALM1	9.24×10^{-21}	0.589	87.5	78.5	1.53×10^{-16}
GYPC	1.57×10^{-20}	0.71	26.4	8.8	2.60×10^{-16}

Average Log_2FC: log_2 fold-change of the average expression between the two groups. Positive values indicate that the feature is more highly expressed in the BM Adjusted P value is based on Bonferroni correction using all features in the dataset.

Key: PT, Patient; BM, Bone Marrow; PB, Peripheral Blood