

A New Way of Exploring Immunity - Linking Highly Multiplexed Antigen Recognition to Immune Repertoire and Phenotype

INTRODUCTION

T cells, key components of the adaptive immune system, are involved in the initiation of the immune response mediated by specific antigen recognition. T-cell receptors (TCRs), located on the surface of T cells, recognize and bind specific antigenic peptides presented by a major histocompatibility complex (MHC) on the surface of other cells (1).

TCRs are heterodimeric proteins, commonly comprising an alpha and a beta chain. The heterodimers recognize peptide-MHC (pMHC) through their complementarity-determining region (CDR) loops: CDR1, CDR2, and CDR3, such that each paired TCR can bind to a particular set of pMHCs. Since each T cell typically expresses a single paired TCR, sufficient TCR sequence diversity is essential across an individual's entire repertoire to bind the broad spectrum of antigens that may be encountered. Somatic V(D)J recombination of the TCR loci during T cell development gives rise to an enormous potential space of TCR sequences, resulting in TCR diversity (2).

The amino acid sequence of the paired TCR directly determines its binding specificity. However, we do not yet have a complete understanding of the factors underlying the recognition of pMHC complexes by their cognate TCRs. A generalized and predictive understanding of this interaction would enable entirely new approaches in basic immunological research and in clinical practice; this would encompass areas such as TCR-based diagnostic methods and the rational design of immunotherapies.

Although we do not yet have a full understanding of TCR recognition, recent work has shown that it is possible to identify shared motifs within TCRs that are specific to particular pMHCs (3, 4). These studies identified TCR sequences that bind a limited number of pMHCs and use the shared motifs to classify previously unseen TCRs according to their predicted ability to bind one of the pMHCs in the training data. These initial studies are extremely promising but are limited by the amount and diversity of training data. A generalizable understanding of TCR recognition will require more extensive data with orders of magnitude more pMHC diversity.

HIGHLIGHTS

We demonstrate the power of the Single Cell Immune Profiling Solution with Feature Barcoding technology to:

- Assess the binding specificities of over 150,000 CD8⁺ T cells from 4 human donors across a highly multiplexed panel of 44 distinct, specific peptide-MHC (pMHC) multimers
- Uncover data at single cell resolution for gene and cell surface protein expression, TCR-pMHC binding specificity, and paired $\alpha\beta$ -T-cell receptor (TCR) sequences
- Generate highly multiplexed binding data from many different pMHC multimers in a single experiment
- Enable the potential to develop novel experimental and analytical approaches
- Help facilitate a greater understanding of how the adaptive immune system responds to immune insults and contribute to the development of therapeutics

The 10x Genomics Chromium Single Cell Immune Profiling Solution with Feature Barcoding technology is particularly well suited to generating the data required to address this question. Single cell-based methods are the only way to directly observe antigen binding to T cells, while simultaneously sequencing the paired and recombined alpha and beta TCR genes. This is achieved by generating Single Cell 5' libraries and V(D)J enriched libraries that provide paired TCR sequences in combination with Feature Barcoding technology, which use highly multiplexed pMHC multimer reagents, each carrying a different pMHC and a distinct molecular barcode, to identify binding specificities.

Here, we demonstrate the use of a highly multiplexed panel of 44 distinct, specific pMHC multimers (dCODE™ Dextramer® reagents) to assess the binding specificities of over 150,000 CD8⁺ T cells from four human donors. This large dataset contains data at single cell resolution on gene expression, expression of eleven surface proteins (using TotalSeq™-C antibodies with Feature Barcoding technology), pMHC binding, and paired αβ TCR sequences (Figure 1).

From an initial analysis of the dataset, we have identified 62,858 cells and 14,320 distinct paired TCR sequences with apparent specificity for at least one pMHC within the panel. This single dataset has the potential to account for a significantly larger amount of knowledge regarding paired human TCR specificities than has been generated to date.

Within our data, we observed TCRs with cognate antigens that had been reported previously, while also identifying entirely new TCR–pMHC interactions. In addition, we observed specific expanded non-naïve T cell clones along with more diverse binding in the naïve compartment.

This rich and large dataset illustrates the power and scalability of the 10x Genomics Chromium Single Cell Immune Profiling Solution with Feature Barcoding technology and presents an exciting opportunity for researchers to explore and draw

further conclusions about the mechanisms of TCR–pMHC interaction. Furthermore, this experiment serves as the next step on the path toward the even larger-scale experiments that will be necessary to fully comprehend the rules of antigen recognition in the adaptive immune system. In addition, the experiment allowed us to enhance our understanding of experimental design and computational analysis, both essential for single cell–based immunology research.

METHODS

Cell samples

We obtained CD8⁺ T cells from four healthy donors from AllCells and Stem Cell Technologies (Reference Table 1). Donors were chosen to ensure each HLA allele in the pMHC panel was present in at least one donor. Cytomegalovirus (CMV) and Epstein Barr Virus (EBV) serostatus was known for Donors 3 and 4.

Peptide-MHC multimers with Feature Barcoding technology

We used a panel of 44 dCODE™ Dextramer® reagents (Immudex, Reference Table 2) with antigenic peptides derived from infectious diseases (CMV, EBV, influenza, HTLV, HPV and HIV) and known cancer antigens.

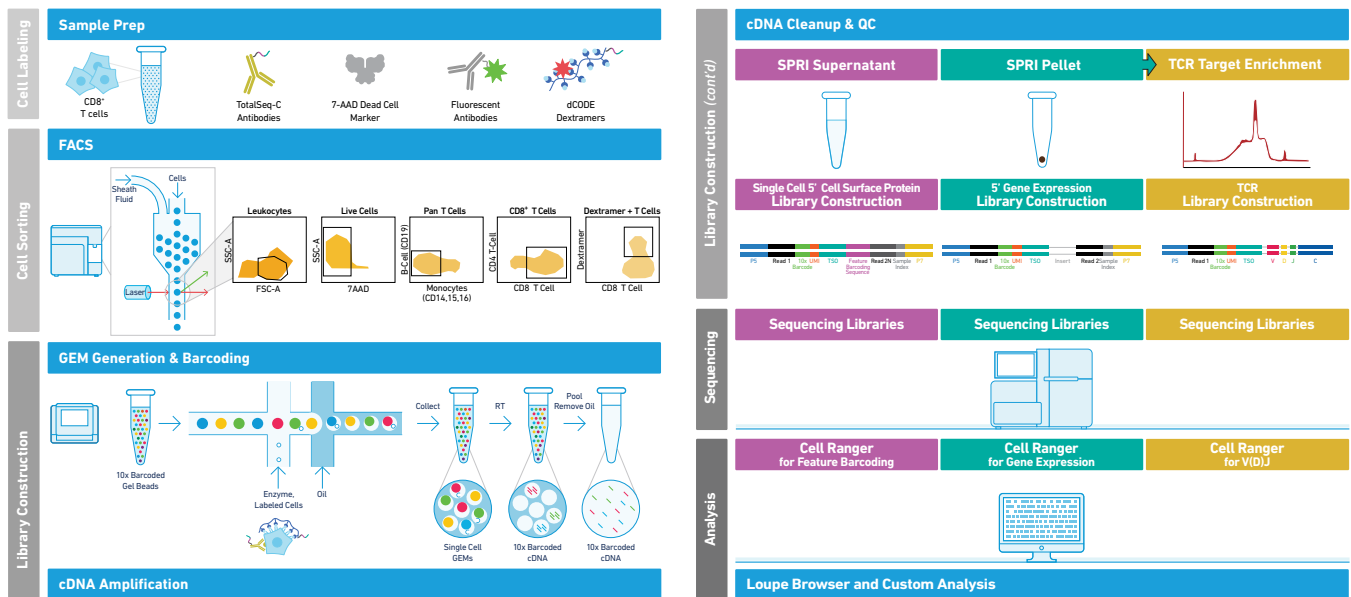


Figure 1. Experimental approach for generating high-throughput, highly multiplexed TCR–antigen binding data. CD8⁺ T cells from healthy human donors were labeled with fluorescent antibodies, TotalSeq™-C antibodies, and dCODE™ Dextramer® reagents. Dextramer® positive CD8⁺ T cells were sorted by flow cytometry and used as input for the 10x Genomics Chromium Single Cell Immune Profiling Solution with Feature Barcoding technology. Libraries were prepared to characterize gene expression, cell surface protein expression, paired TCR sequences, and TCR–pMHC binding in each single cell.

Each Dextramer® reagent included a distinct nucleic acid barcode, along with a phycoerythrin (PE) fluorophore.

The panel also contained 6 Dextramers® with irrelevant negative control peptides to assist in the detection of non-specific binding events.

Just prior to labeling, 2 µl of each Dextramer® reagent was combined for a total volume of 100 µl.

Surface marker antibodies with Feature Barcoding technology

We used a panel of eleven TotalSeq™-C antibodies (BioLegend, Reference Table 3) that were chosen to enable discrimination between CD8⁺ T cell (CD8A⁺/CD3⁺) subpopulations, such as naïve (CD45RA⁺/CD45RO⁻/CCR7⁺), effector (TEF, CD45RA⁺/CD45RO⁺/CCR7⁻), effector memory (TEM, CD45RA⁻/CD45RO⁺/CCR7⁻), exhausted (PD-1⁺), central memory (TCM, CD45RA⁻/CD45RO⁺/CCR7⁺), and activated (HLA-DR⁺/CD127⁺), and exclude any CD4⁺ T cells (CD4⁺), B cells (CD19⁺), and monocytes (CD14⁺, CD15⁺, CD16⁺). Prior to staining, the TotalSeq™-C antibodies were pooled by combining 0.5 µg of each antibody per reaction. CCR7 and HLA-DR antibodies were not included in the panel for staining cells from Donor 1.

Fluorescently labeled surface marker antibodies

We used a panel of fluorescently labeled antibodies (BioLegend, Reference Table 4) to enable sorting of pure CD8⁺ T cells. Fluorescent antibodies were used in staining reactions at a dilution of 1:100 and were different clones than the TotalSeq™-C antibodies to avoid binding competition.

Cell labeling

For each donor, 20 million CD8⁺ T cells were thawed in a 37°C water bath. 1 ml warm medium (RPMI + 10% FBS) was added to the cells in a 15 ml conical tube containing 9 ml warm medium. The cells were washed once in 5 ml medium and finally resuspended in 4–6 ml PBS + 2% FBS and then counted. For Dextramer® and antibody labeling, 20 million cells were centrifuged and resuspended in 105 µl PBS + 2% FBS. 12.5 µl Human TruStain FcX™ (Fc receptor blocking solution, BioLegend, Cat# 422301), 5 µl dextran sulfate (10 mg/ml), and 100 µl pooled Dextramer® panel were added to the cells and incubated for 10 minutes on ice. 7 µl TotalSeq™-C antibody panel and 15 µl fluorescent antibody panel were then added and the cells were incubated for 30 minutes on ice. After incubation, cells were washed by adding 1.2 ml PBS + 2% FBS, followed by centrifugation for 5 minutes at 350 Xg (swing bucket). Cells were washed 2 more times in 1.5 ml PBS + 2% FBS. After the final centrifugation, the cells were resuspended in 3–4 ml PBS + 2% FBS. 100 µl cell

suspension was reserved for a non-sorted cell population. Before cell sorting, 7-AAD viability staining solution (BioLegend, Cat# 420403) was added at 1:200 dilution. See our Demonstrated Protocol [CG000149](#) for full details.

Cell sorting

Cells were sorted using a MA900 Multi-Application Cell Sorter (Sony Biotechnology) with 100 µm flow cell chip. Cells were sorted into 55.9 µl Reaction Mix containing RT Reagent Mix and Poly dT RT primers in 8 wells of a chilled 96-well plate. We targeted a cell yield of 9000 cells per well. Cells were gated using the Sony MA900 system software to obtain single (FSH/BSC, FSW/BSC), live (7-AAD⁻), CD8⁺ cells whilst excluding CD14⁺CD15⁺CD16⁺CD19⁺CD4⁺ cells. We then selected all Dextramer®-positive cells within the CD8⁺ T cell population. After sorting, 5 µl water and a total 12.4 µl of Additive A (2.4 µl) and RT Enzyme Mix B (10 µl) were added to each well to complete the Reaction Mix, which was then directly loaded onto a Chromium chip.

Generation of the single cell immune profiling libraries

The Chromium Single Cell V(D)J workflow with Feature Barcoding technology generates Single Cell 5' Gene Expression, V(D)J, and Cell Surface Protein libraries. The Cell Surface Protein libraries are generated from amplified DNA derived from antibody or Dextramer® conjugated Feature Barcode oligonucleotide, which are bound to the cell surface proteins and TCRs, respectively. A total 180 libraries were generated (60 libraries each for Gene Expression, Cell Surface Protein, and V(D)J). The libraries were prepared following the User Guide for Chromium Single Cell V(D)J Reagent kits with Feature Barcoding technology for Cell Surface Protein ([CG000186](#)).

Sequencing libraries

Chromium Single Cell V(D)J enriched libraries, 5' Gene Expression libraries, and Cell Surface Protein libraries were quantified, normalized, and sequenced according to the User Guide for Chromium Single Cell V(D)J Reagent kits with Feature Barcoding technology for Cell Surface Protein ([CG000186](#)). The 5' Gene Expression libraries and Cell Surface Protein libraries from Donors 1 and 3 and from Donors 2 and 4 were combined into 2 library pools and sequenced on an Illumina NovaSeq sequencer with a NovaSeq 6000 S4 Reagent Kit (200 cycles) (Illumina). The V(D)J enriched libraries from all 4 donors were pooled and sequenced with an Illumina NovaSeq 6000 S2 Reagent Kit (300 cycles) (Illumina). The Gene Expression, Cell Surface Protein, and V(D)J libraries were targeted, respectively, for sequencing depths of at least 20,000, 20,000 and 10,000 read pairs per cell (rppc). V(D)J libraries were downsampled to 5,000 rppc prior to analysis.

Analysis

The sequencing data combined with the description of the sequences identifying TotalSeq™-C antibodies and dCODE™ Dextramer® reagents were analyzed with the Cell Ranger analysis pipeline (see [What is Cell Ranger?](#)).

RESULTS AND PRELIMINARY OBSERVATIONS

Generating high-throughput, highly multiplexed TCR binding data

We acquired CD8⁺ T cells from four healthy human donors (Reference Table 1) and used FACS to isolate those T cells that appeared to bind at least one pMHC multimer from a panel of 50 dCODE™ Dextramer® reagents. These T cells were used to generate gene expression, TCR sequence, cell surface protein expression, and pMHC binding data (44 specific pMHCs; Reference Table 2) in 160,914 single cells where at least one productive TCR chain was detected.

We used the annotated TCR sequences produced by Cell Ranger to identify putative mucosal associated invariant T (MAIT) cells, based on whether they expressed a TCR α chain, comprising the typical V and J segments (5). We identified 2,769 MAIT cells and excluded them from further analysis. No iNKT cells were identified by a similar approach.

Within the remaining cells, we were able to identify a productive TCR α and β chain in 118,007 (73%) and we assigned these sequences to 43,721 distinct clonotypes, where each clonotype had a unique set of recombined TCR nucleotide sequences.

We used graph-based clustering to cluster the cells into subpopulations based on their gene expression profiles. We classified the clusters as either naïve (CD45RA⁺CD45RO⁻CCR7⁺) or non-naïve (the remaining cells) by analyzing gene and cell surface protein expression in Loupe Cell Browser. As expected, the non-naïve T cell population was highly enriched for expanded clonotypes (Figure 2).

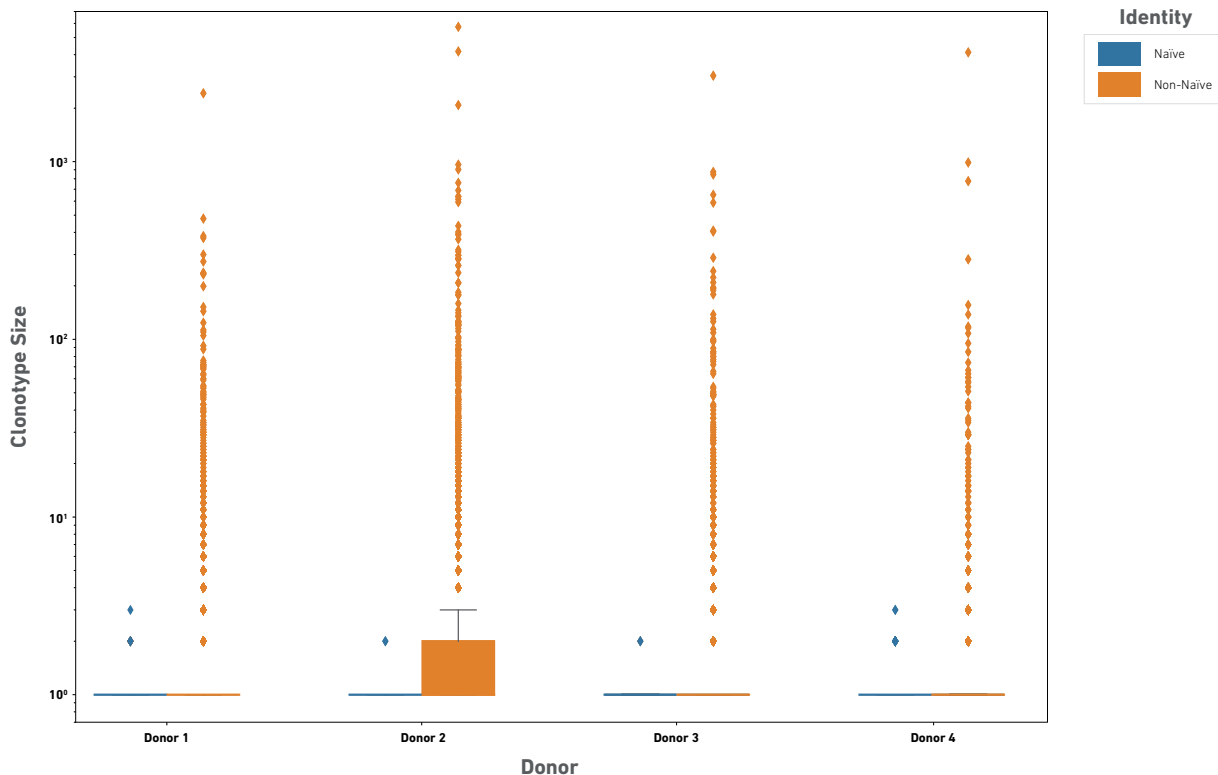


Figure 2. Expanded clonotypes are found in the non-naïve compartment. CD8⁺ T cells were classified as either naïve or non-naïve, based on cell surface marker and gene expression. Cells were assigned to clonotypes based on TCR sequences. Box plots indicate the distribution of the number of cells in each clonotype within the two compartments.

Preliminary analysis of T cells that bind to specific pMHC multimers

We classified the T cells according to whether they exhibited specific binding to at least one of the dCODE™ Dextramer® reagents used in the experiment. We focused on cells that had strong evidence of specific binding by setting a threshold such that a specific binding event required a UMI count greater than 10 that was also greater than five times the highest negative control UMI count for that cell. In cases where a cell was assigned more than one specificity, we considered it to be specific only for the pMHC with the highest UMI count. We also excluded any cells with apparent specificities for more than four Dextramers®. This gave 62,858 cells with pMHC specificity distributed across 14,320

unique clonotypes. Amongst the 37 dCODE™ Dextramer® reagents analyzed for specificity, the ones without a specific binding T cell were the following: HLA-A*01:01-VTEHDTLLY, HLA-A*02:01-KVAELVHFL, HLA-A*02:01-LLMGTGLGIVC, HLA-A*02:01-CLGGLTLMV, HLA-A*02:01-NLVPMTATV, HLA-A*02:01-KLQCVDLHV, HLA-B*07:02-TPRVTGGGAM.

pMHC-specific cells were not distributed evenly between the donors (Supplemental Figure 1). This was expected given that only Donors 1 and 2 possessed the HLA-A*02:01 allele that was the most common MHC in the antigen panel (Reference Tables 1 and 2). Additionally, combining the cell type classifications with binding data allowed us to determine the specific binding events between the naïve and memory T cell compartments for each donor (Figure 3 and Supplemental Figure 1).

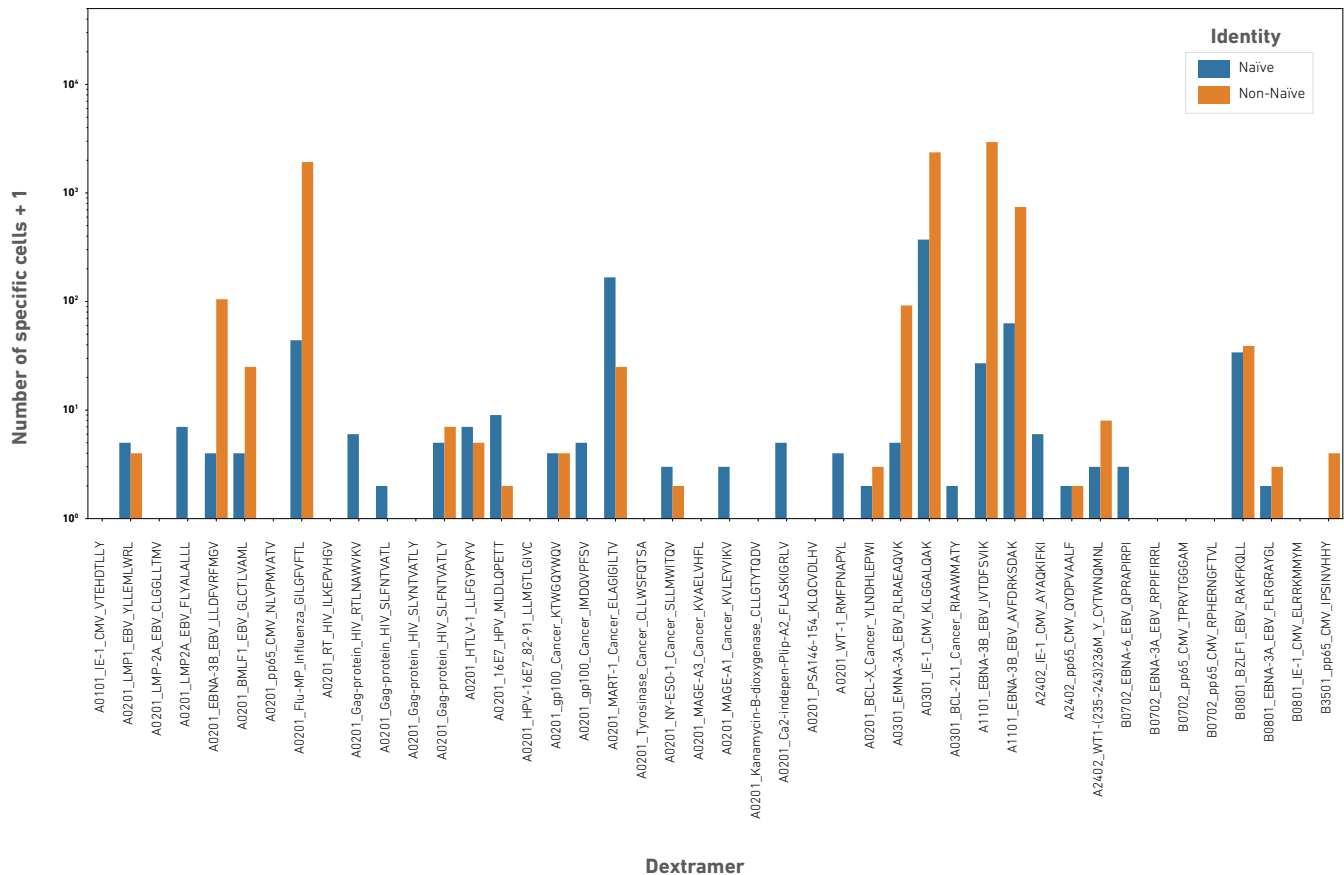


Figure 3. pMHC binding landscape for Donor 1. Bars indicate the number of cells classified as having specific binding to each pMHC multimer in the naïve and non-naïve compartments.

We observed multiple pMHC-specific naïve T cells with diverse TCR sequences. These naïve binders were often specific to endogenous or tumor-associated antigens (e.g. MART-1) or to antigens derived from viruses for which the donor was seronegative (e.g. HIV).

We looked for evidence of pMHC-specific expanded clonotypes within the CD8⁺ memory compartment. We observed binding of multiple clonally-related cells to the same pMHC, implying a specific clonal expansion driven by response to that particular antigen (Figure 4). These apparently specific expansions were often associated with pMHCs that were HLA-matched to a particular donor; in addition, expansions were consistent with donors being seropositive for particular common viral infections (e.g. EBV and Influenza in Donor 1 and Donor 2).

To further investigate binding within expanded clonotypes, we determined the binding concordance within each clonotype. For each expanded clonotype with a binding specificity, we calculated the proportion of the entire clonotype within the donor that had the particular binding specificity (Figure 4 and Supplementary Figure 2). A proportion of 1 indicates that all the cells within the clonotype have the same binding specificity. Clonotypes with high binding concordance were observed in instances of expected specific expansions, while lower binding concordances were more typical for pMHCs that were not HLA-matched.

Intriguingly, we observed clonotypes with apparently cross-reactive binding to A0301_EMNA-3A_EBV_RLRAEAQVK, A0301_IE-1_CMV_KLGGALQAK, A1101_EBNA-3B_EBV_IVTDFSVIK and A1101_EBNA-3B_EBV_AVFDRKSDAK.

Binding to these multimers was observed in all four donors irrespective of their HLA haplotype and serostatus. While some of the expanded clonotypes exhibited high levels of binding concordance, others did not. This suggests an avenue for further study to investigate the causes of these binding events and whether they represent specific TCR-mediated binding or another, less specific binding mode.

Comparison of pMHC-binding TCR sequences with previously reported observations

We compared the pMHC-specific TCRs that we observed in these data with those that have been previously reported to have the same binding specificities in VDJdb (6). VDJdb contains paired, human TCR sequences for 11 of the pMHCs for which we observed binding events. We calculated the TCRdist metric (3) between each TCR pair we observed and the closest entry in VDJdb. We found examples of exactly matching sequences (Table 1) alongside sequences with varying degrees of similarity to those already reported (Figure 5).

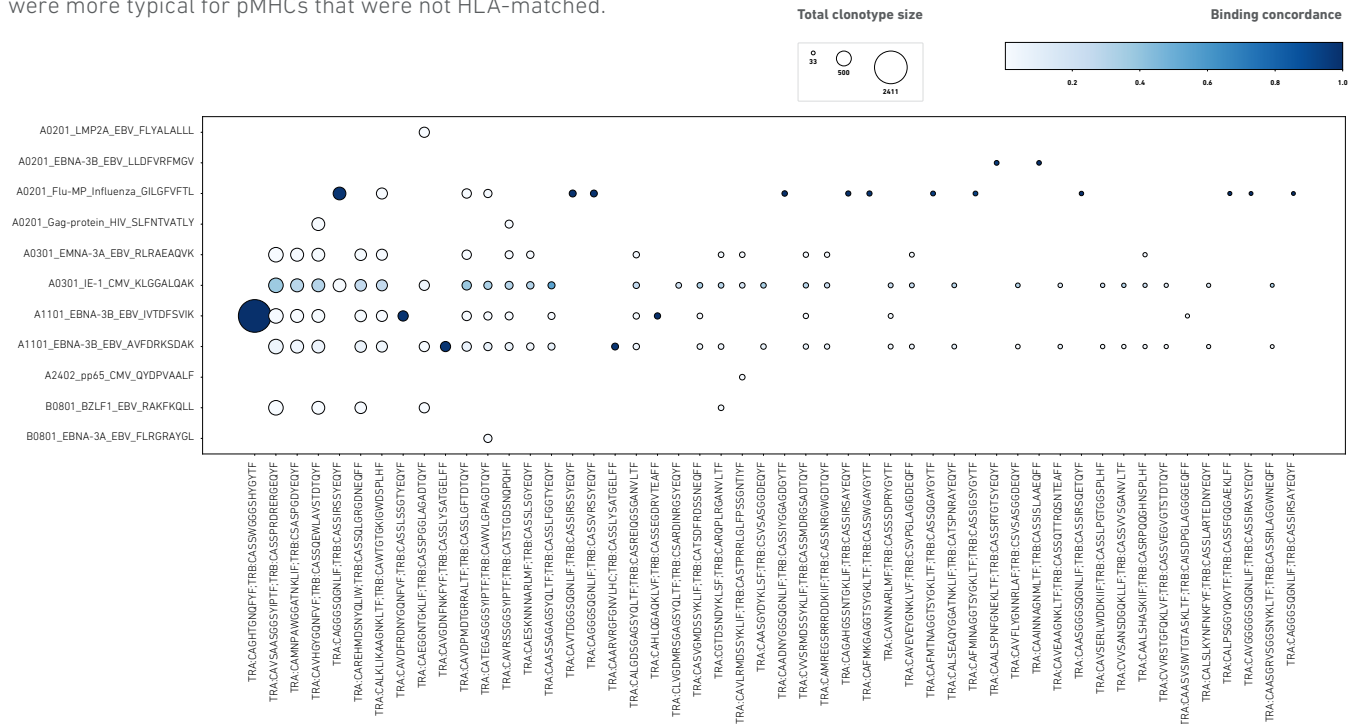


Figure 4. Binding specificities of expanded clonotypes from Donor 1. The 50 largest clonotypes are plotted along with their binding specificities and concordance. A circle indicates that at least one member of the clonotype was classified as specific for a particular pMHC. Circle size indicates the total within-donor clonotype size. Circle color indicates the proportion of cells within the clonotype that bind the Dextramer® (the 'binding concordance').

		Dataset	VDJdb	Overlap
donor1	A0201_ELAGIGILTV_MART-1_Cancer	177	44	0
donor1	A0201_SLLMWITQV_NY-ESO-1_Cancer	3	1	0
donor1	A0201_GILGFVFTL_Flu-MP_Influenza	339	577	44
donor1	A0201_GLCTLVAML_BMLF1_EBV	14	148	4
donor1	A0201_LLFGYPVYV_HTLV-1	9	2	0
donor1	A0201_RMFPNAPYL_WT-1	3	5	0
donor1	A1101_IVTDFSVIK_EBNA-3B_EBV	146	8	0
donor1	A1101_AVFDRKSDAK_EBNA-3B_EBV	276	7	0
donor1	B3501_IPSINVHHY_pp65_CMV	3	2	0
donor1	B0801_FLRGRAYGL_EBNA-3A_EBV	3	5	0
donor2	A0201_ELAGIGILTV_MART-1_Cancer	25	44	0
donor2	A0201_GILGFVFTL_Flu-MP_Influenza	650	577	54
donor2	A0201_GLCTLVAML_BMLF1_EBV	64	148	2
donor2	A0201_LLFGYPVYV_HTLV-1	6	2	0
donor2	A1101_IVTDFSVIK_EBNA-3B_EBV	58	8	0
donor2	A1101_AVFDRKSDAK_EBNA-3B_EBV	236	7	0
donor2	B3501_IPSINVHHY_pp65_CMV	1	2	0
donor2	B0801_FLRGRAYGL_EBNA-3A_EBV	16	5	0
donor3	A0201_ELAGIGILTV_MART-1_Cancer	33	44	0
donor3	A0201_SLLMWITQV_NY-ESO-1_Cancer	1	1	0
donor3	A0201_GILGFVFTL_Flu-MP_Influenza	9	577	0
donor3	A0201_LLFGYPVYV_HTLV-1	8	2	0
donor3	A0201_RMFPNAPYL_WT-1	1	5	0
donor3	A1101_IVTDFSVIK_EBNA-3B_EBV	66	8	0
donor3	A1101_AVFDRKSDAK_EBNA-3B_EBV	646	7	0
donor3	B3501_IPSINVHHY_pp65_CMV	2	2	0
donor3	B0801_ELRRKMMYM_IE-1_CMV	1	2	0
donor4	A0201_ELAGIGILTV_MART-1_Cancer	19	44	0
donor4	A0201_GILGFVFTL_Flu-MP_Influenza	2	577	0
donor4	A0201_GLCTLVAML_BMLF1_EBV	1	148	0
donor4	A0201_LLFGYPVYV_HTLV-1	2	2	0
donor4	A1101_IVTDFSVIK_EBNA-3B_EBV	88	8	0
donor4	A1101_AVFDRKSDAK_EBNA-3B_EBV	76	7	0
donor4	B0801_ELRRKMMYM_IE-1_CMV	1	2	0

Table 1. Comparison between pMHC-specific paired TCR sequences identified in this experiment and those reported in VDJdb. Overlapping sequences are those with a TCRdist separation of 0.

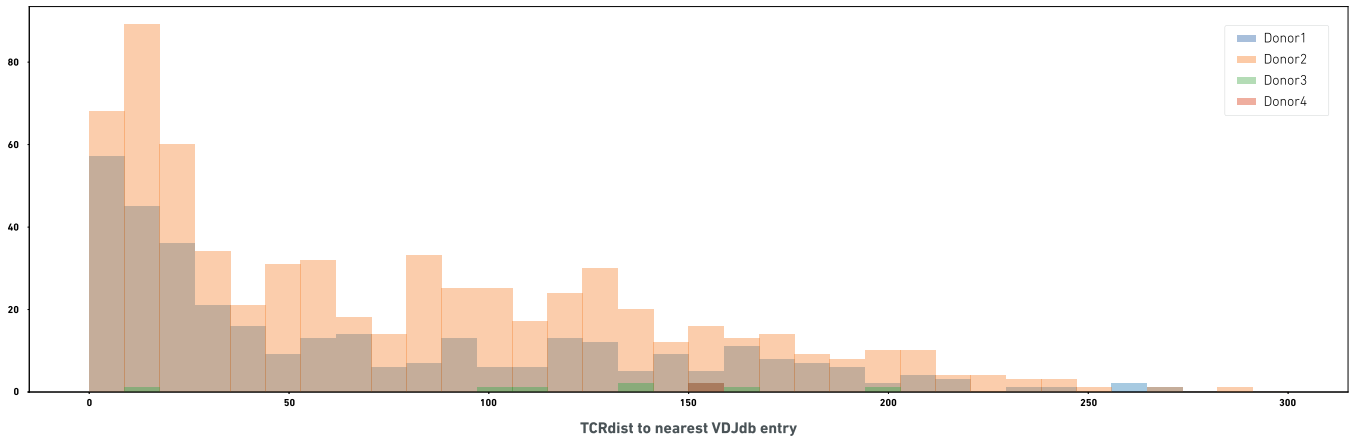
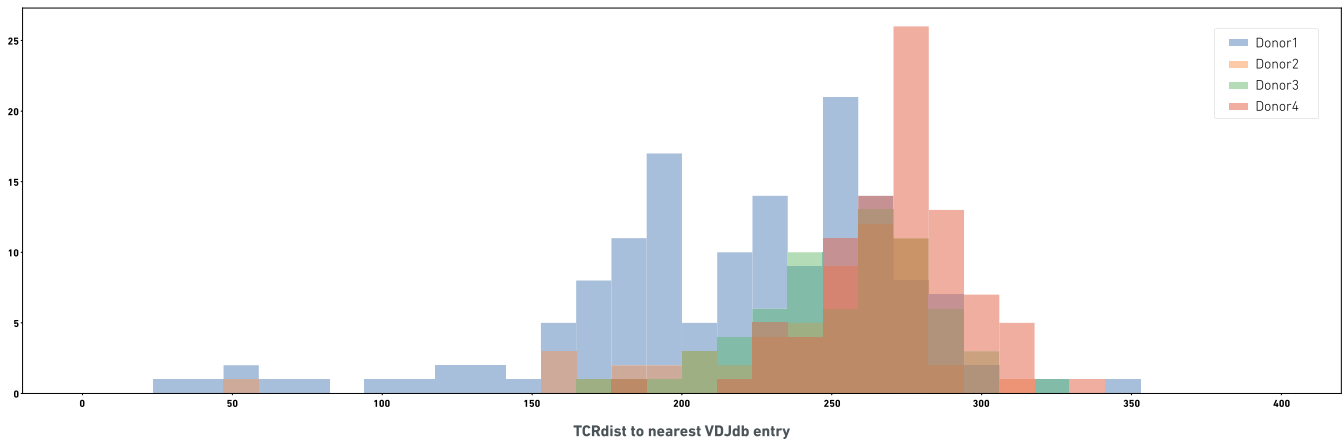
A0201_Flu-MP_Influenza_GILGFVFTL**A1101_EBNA-3B_EBV_IVTDFSVIK**

Figure 5. Similarities between paired TCR sequences binding to two pMHC-specific Dextramers® with a peptide previously reported in VDJdb. Histograms show the distribution of TCRdist values between pMHC-specific paired TCRs and the closest TCR paired chain entry in VDJdb with the same reported specificity.

Discussion

This experiment, the first of its kind at this scale, demonstrates that Feature Barcoding technology enables the generation of highly multiplexed binding data from many different pMHC multimers in a single experiment. The dataset also illustrates the complexity inherent in performing and analyzing these experiments. Even though the experimental design and the analyses presented here are not fully optimized for understanding TCR binding at such a large scale, the data will be useful in extracting detailed biological insights as well as developing future experimental and analytical approaches.

Key areas for experimental improvement include optimizing the design of pMHC multimer panels and the source of donor cells with respect to HLA types. For the analysis, the most crucial step is the classification of cells according to their binding specificities; there is a valuable opportunity to develop more sophisticated approaches that maximize sensitivity and specificity by moving beyond simple count-based thresholds.

Associating TCR sequences with pMHC specificity and cellular phenotypes at this scale and resolution will enable a new way of exploring adaptive immunity and understanding some of the most fundamental questions in immunology.

REFERENCE TABLES

Donor	Provider	Alleles				CMV status	EBV status	Other status
		HLA-A 1	HLA-A 2	HLA-B 1	HLA-B 2			
Donor 1	StemCell Technologies	02:01	11:01	35:01	ND	ND	ND	HIV-1, HIV-2, HBV, HCV negative
Donor 2	StemCell Technologies	02:01	01:01	08:01	ND	ND	ND	HIV-1, HIV-2, HBV, HCV negative
Donor 3	AllCells	24:02	29:02	35:02	44:03	+	-	HIV, HBV, HCV negative
Donor 4	AllCells	03:01	03:01	07:02	57:01	-	-	HIV, HBV, HCV negative

Reference Table 1. Information on the T cell donors used in this study. No data.

Cat#	Allele	Category	Peptide
WA2131	A*0101	IE-1/CMV	VTEHDTLLY
WB2660	A*0201	gp100/Cancer/	KTWGQYWQV
WB2162	A*0201	MART-1/Cancer/	ELAGIGILTV
WB3697	A*0201	Tyrosinase/Cancer/	CLLWSFQTSA
WB2158	A*0201	gp100/Cancer/	IMDQVPFSV
WB3247	A*0201	NY-ESO-1/Cancer/	SLLMWITQV
WB3497	A*0201	MAGE A3/Cancer/	KVAELVHFL
WB3474	A*0201	MAGE-A1/Cancer/	KVLEYVIKV
WB5066	A*0201	Kanamycin B dioxygenase	CLLGTYTQDV
WB2143	A*0201	EBNA 3B/EBV	LLDFVRFMGV
WB3307	A*0201	HPV 16E7, 82-91	LLMGTLGIVC
WB2144	A*0201	LMP-2A/EBV	CLGGLTMV
WB3531	A*0201	LMP1/EBV	YLLEMLWRL
WB3529	A*0201	LMP2A/EBV	FLYALALL
WB2161	A*0201	Flu MP/Influenza	GILGFVFTL
WB2130	A*0201	BMLF1/EBV	GLCTLVAML
WB2132	A*0201	pp65/CMV	NLVPMVATV
WB2139	A*0201	RT/HIV	ILKEPVHGV
WB5335	A*0201	Ca2+-indepn. Plip A2	FLASKIGRLV
WF2639	A*2402	WT1 (235-243)236M->Y	CYTWNQMNL
WB2646	A*0201	Gag protein/HIV	RTLNAWVKV
WB2157	A*0201	PSA 146-154	KLQCVDLHV
WB2141	A*0201	HTLV-1	LLFGYPVYV
WB3338	A*0201	Gag protein/HIV	SLFNTVATL
WB3339	A*0201	Gag protein/HIV	SLYNTVATLY

WB3340	A*0201	Gag protein/HIV	SLFNTVATLY
WB2177	A*0201	WT-1	RMFPNAPYL
WB2191	A*0201	BCL-X/Cancer/	YLNDHLEPWI
WB2652	A*0201	16E7/HPV	MLDLQPETT
WC2197	A*0301	IE-1/CMV	KLGGALQAK
WC2656	A*0301	EMNA 3A/EBV	RLRAEAQVK
WC2632	A*0301	BCL-2L1/Cancer/	RIAAWMATY
WD2175	A*1101	EBNA 3B/EBV	IVTDFSVIK
WD2149	A*1101	EBNA 3B/EBV	AVFDRKSDAK
WK2138	B*3501	pp65/CMV	IPSINVHHY
WF2196	A*2402	IE-1/CMV	AYAQKIFKI
WF2133	A*2402	pp65/CMV	QYDPVAALF
WH2165	B*0702	EBNA 6/EBV	QPRAPIRPI
WH2136	B*0702	pp65/CMV	TPRVTGGGAM
WH2166	B*0702	EBNA 3A/EBV	RPPIFIRRL
WH2135	B*0702	pp65/CMV	RPHERNGFTVL
WI2148	B*0801	BZLF1/EBV	RAKFKQLL
WI2137	B*0801	IE-1/CMV	ELRRKMMYM
WI2147	B*0801	EBNA 3A/EBV	FLRGRAYGL
WA3580	A*0101	NC	SLEGGGLGY
WA3579	A*0101	NC	STEGGGLAY
WB2666	A*0201	NC	ALIAPVHAV
WF3231	A*2402	NC	AYSSAGASI
WH3397	B*0702	NC	GPAESAAGL
NI3233	NR(B0801)	NC	AAKGRGAAL

Reference Table 2. List of the dCODE™ Dextramer® reagents used in the study.

Cat#	Clone	Description
300479	UCHT1	TotalSeq-C 0034 anti-human CD3 Antibody
302265	HIB19	TotalSeq-C 0050 anti-human CD19 Antibody
304163	HI100	TotalSeq-C 0062 anti-human CD45RA Antibody
300567	RPA-T4	TotalSeq-C 0072 anti-human CD4 Antibody
301071	RPA-T8	TotalSeq-C 0080 anti-human CD8A Antibody
301859	M5E2	TotalSeq-C 0081 anti-human CD14 Antibody
304259	UCLH1	TotalSeq-C 0087 anti-human CD45RO Antibody
329963	EH12.2H7	TotalSeq-C 0088 anti-human CD279 (PD-1) Antibody
351356	A019D5	TotalSeq-C 0390 anti-human CD127 (IL7Ra) Antibody
353251	G043H7	TotalSeq-C 0148 anti-human CD197 (CCR7) Antibody
307663	L243	TotalSeq-C 0159 anti-human HLA-DR Antibody
400187	MOPC-21	TotalSeq-C 0090 Mouse IgG1, K isotype control Antibody
400293	MOPC-173	TotalSeq-C 0091 Mouse IgG2a, K isotype control Antibody
400381	MPC-11	TotalSeq-C 0092 Mouse IgG2b, K isotype control Antibody

Reference Table 3. List of the TotalSeq™-C antibodies used in the study.

Cat#	Clone	Fluorochrome	Description
363011	SJ25C1	PE/Cy7	Anti-human CD19 antibody
325615	HCD14	Pacific Blue™	Anti-human CD14 antibody
323021	W6D3	Pacific Blue	Anti-human CD15 (SSEA-1) antibody
360723	B73.1	Brilliant Violet 421™	Anti-human CD16 antibody
344716	SK1	FITC	Anti-human CD8 antibody
344633	SK3	Brilliant Violet 510	Anti-human CD4 antibody

Reference Table 4. List of the antibodies used in the study for cell sorting.

RESOURCES

Datasets	go.10xgenomics.com/vdj/datasets
Seminars	go.10xgenomics.com/vdj/seminars
Application Notes	go.10xgenomics.com/vdj/app-notes
Support	go.10xgenomics.com/vdj/support
Publications	go.10xgenomics.com/vdj/pubs

SUPPORT

support@10xgenomics.com

10x Genomics
7068 Koll Center Parkway, Suite 401
Pleasanton, CA 94566 USA

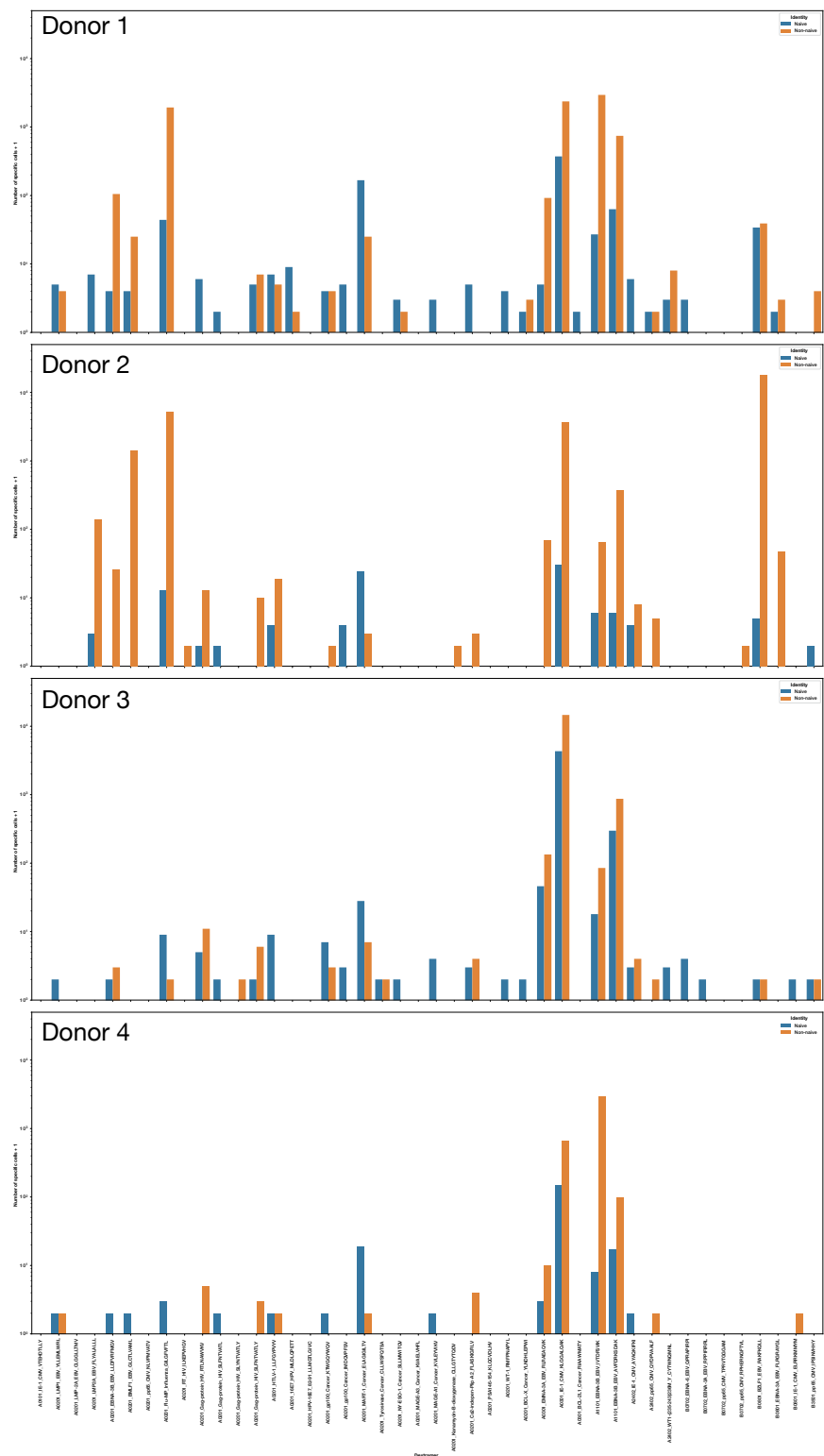
LEGAL NOTICES

For 10x Genomics legal notices visit:
10xgenomics.com/legal-notices

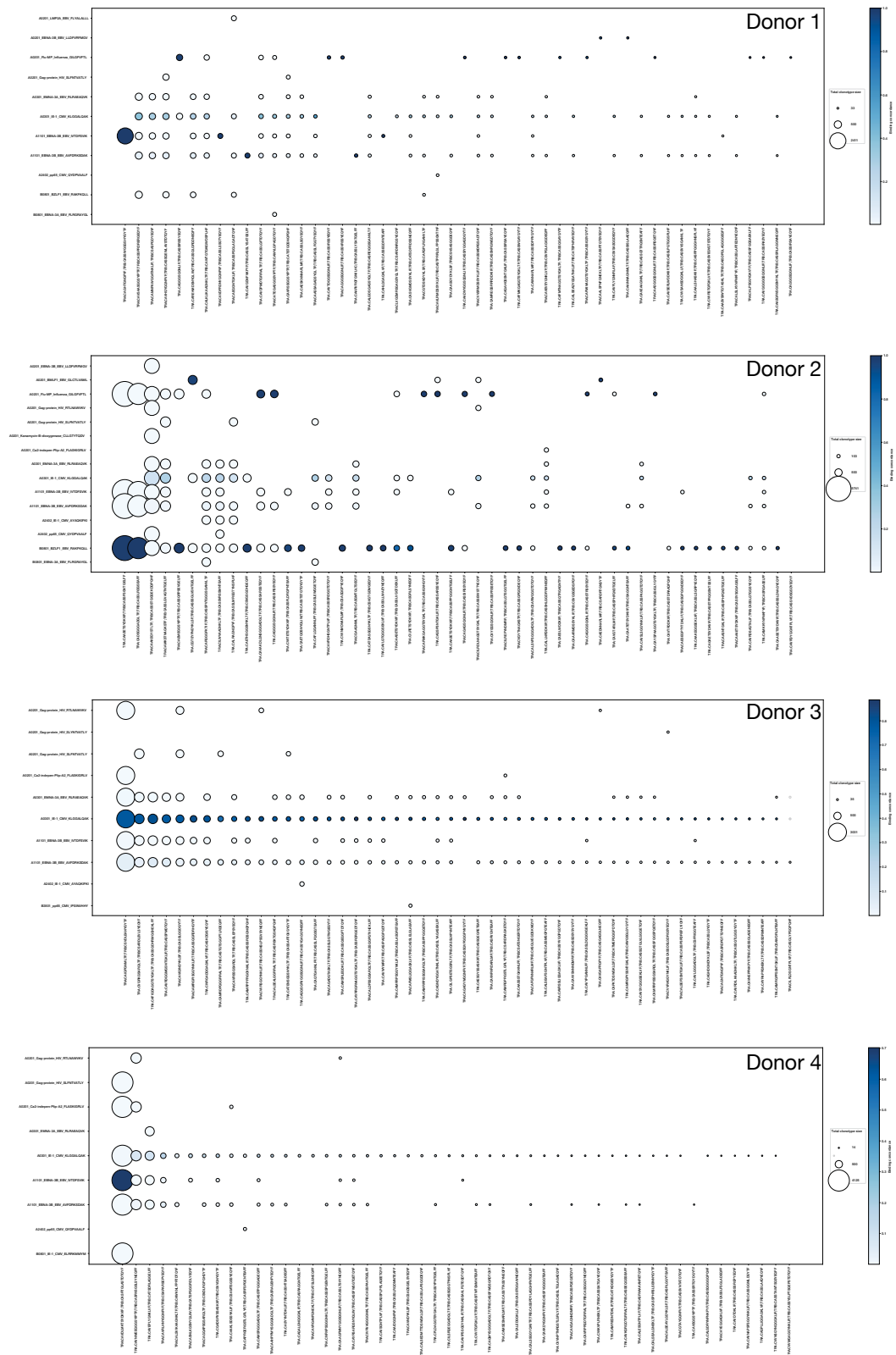
REFERENCES

1. N. L. La Gruta, S. Gras, S. R. Daley, P. G. Thomas, J. Rossjohn, Understanding the drivers of MHC restriction of T cell receptors. *Nat. Rev. Immunol.* **18**, 467-478 (2018).
2. A. Murugan, T. Mora, A. M. Walczak, C. G. Callan Jr, Statistical inference of the generation probability of T-cell receptors from sequence repertoires. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16161-16166 (2012).
3. P. Dash et al., Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature.* **547**, 89-93 (2017).
4. J. Glanville et al., Identifying specificity groups in the T cell receptor repertoire. *Nature.* **547**, 94-98 (2017).
5. D. I. Godfrey, A. P. Uldrich, J. McCluskey, J. Rossjohn, D. B. Moody, The burgeoning family of unconventional T cells. *Nat. Immunol.* **16**, 1114-1123 (2015).
6. M. Shugay et al., VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res.* **46**, D419-D427 (2018).

SUPPLEMENTAL FIGURES



Supplemental Figure 1. pMHC binding landscape for all donors. Bars indicate the number of cells classified as having specific binding to each pMHC multimer in the naive and non-naive compartments.



Supplemental Figure 2. Binding specificities of expanded clonotypes from all donors. The 50 largest clonotypes are plotted along with their binding specificities and concordance. A circle indicates that at least one member of the clonotype was classified as specific for a particular pMHC. Circle size indicates the total within-donor clonotype size. Circle color indicates the proportion of cells within the clonotype that bind the Dextramer® (the 'binding concordance').