Next-generation detection of antigen-responsive T cells using DNA barcode-labeled MHC-I multimers

Authors:

Amalie Kai Bentzen[1], Andrea Marion Marquard[2], Rikke Lyngaa[1], Sunil Kumar Saini[1], Sofie Ramskov[1], Marco Donia[3], Lina Such[1], Andrew J.S. Furness[4,5], Nicholas McGranahan[4,6], Rachel Rosenthal[4,6], Per thor Straten[3], Zoltan Szallasi[2], Inge Marie Svane[3], Charles Swanton[4,6], Sergio A. Quezada[4,5], Søren Nyboe Jakobsen[1,7], Aron Charles Eklund[2], Sine Reker Hadrup[1,*]

Affiliations:

[1]Section for Immunology and Vaccinology, National Veterinary Institute, Technical University of Denmark, 1870 Frederiksberg C, Denmark

[2]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Lyngby, Denmark

[3]Center for Cancer Immune Therapy, Herlev University Hospital, 2730 Herlev, Denmark

[4]CRUK Lung Cancer Center of Excellence, UCL Cancer Institute, London, UK.

[5]Cancer Immunology Unit, UCL Cancer Institute, University College London, UK

[6]Translational Cancer Therapeutics Laboratory, The Francis Crick Institute 44 Lincoln's Inn Fields, London, UK.

[7]Immudex, 2100 Copenhagen, Denmark

[*]Correspondence: sirha@vet.dtu.dk (S.R.H.)

Contact Information:

Sine Reker Hadrup, Technical University of Denmark, Bülowsvej 27, 1870 Frederiksberg C, Denmark

ABSTRACT

The identification of specific peptides recognized by T cells is important for understanding and treating immune-related diseases. Current cytometry-based approaches are limited to the simultaneous screening of T cell reactivity towards 10–100 distinct peptide specificities in a single sample. We present a novel technology validated to screen for T cell recognition of multiple (>1000) peptide specificities in a single sample using peptide-MHC multimers labeled with individual DNA barcodes. Among MHC multimer-binding T cells the relative frequency of sequenced DNA barcodes originating from a given peptide-MHC motif is related to the size of the antigen-responsive T cell population. We demonstrate the application of DNA barcode-labeled MHC multimers for the detection of rare T cell populations of both virus- and cancer-restricted origins in various tissues. This technology enables true high-throughput detection of antigen-responsive T cells and opens the possibility of genome-wide immune assessments on a personalized basis.

Main text

CD8 T cells recognize peptide antigens presented by their cognate major histocompatibility complex I (MHC I) molecule. This is a key element for adaptive immunity in the control of intracellular pathogens, recognition and elimination of cancer and the pathogenesis of autoimmune diseases. To understand disease development and to foster specific therapeutic interventions, it is

crucial to characterize the specific peptide-MHC molecules (pMHC) recognized by CD8 T cells of relevance in a given disease. For the past two decades, antigen-specific T cells have been detected using fluorescently labeled pMHC multimers[1]. Currently, MHC multimer-based detection of antigen-responsive T cells is limited by the number of fluorescent or metal tags available for either flow- or mass cytometry[2,3]. Combinatorial encoding, applying a specific combination of several tags to a given pMHC multimer, has been employed to enhance the complexity of both analyses, which has recently enabled the parallel screening of 120 different pMHC-responsive T cells in single samples[4–7]. Although the combinatorial encoding principle has significantly enhanced our ability to describe immune interactions, it is dwarfed by the vast number of potentially presented peptides and the diversity of the T cell clones that recognize them[8,9].

DNA barcodes holds advantage as a novel type of tag for MHC multimers since oligonucleotides of suitable lengths can be assembled to generate $>10^8$ different DNA barcodes[10] and because the composition of DNA barcodes in a sample can easily be determined through high-throughput sequencing.

We provide proof-of-concept for the mixing >1000 different pMHC multimers tagged with unique DNA barcodes to enable the parallel detection of potentially >1000 different pMHC-responsive T cells in a single sample. The increased complexity of T cell detection enabled by use of DNA-barcoded MHC multimers advances our understanding of immune recognition, from model antigens to complex screenings that allow selection of candidate peptides on a genome-wide level, such as the mutagenome of human cancer.

RESULTS

Multi-parallel detection of antigen-responsive T cells in single samples

To allow complex assessments of T cell reactivity in limited biological samples, we developed a technology using DNA barcodes as tags for specific interactions between pMHC molecules and CD8 T cells. DNA barcoded MHC multimer reagents also carrying a common fluorescent label, phycoerythrin (PE), were generated on a dextran backbone as shown in Fig. 1a. This strategy enabled single-tube-based detection of pMHC responsive T cells using mixtures of >1000 distinct pMHC multimers where each specific pMHC molecule was associated with a given DNA barcode. MHC multimer-binding T cells were isolated using fluorescence-activated cell sorting (FACS) based on their PE fluorescence intensity, and the composition of the associated DNA barcodes were identified through amplification and high-throughput sequencing. Antigen-responsive T cells within the isolated T cell pool were identified based on the number of reads for a specific pMHC-associated barcode compared to the complete pMHC multimer library. The number of unique DNA sequences originating from each pMHC-associated barcode was assessed based on the incorporation of unique molecular identifiers (UMIs)[11], allowing the clonal reduction of sequences from the amplified product. This revealed the number of specific pMHC multimers that interacted with T cells in the given cell sample.

The DNA barcodes were designed from sets of unique 25mer oligonucleotides using previously published sequences described as having similar amplification properties while containing maximum diversity of their identification motifs[10]. To provide a system adaptable to large library screenings, we applied a combinatorial design of DNA barcodes as depicted in Fig. 1b and listed in Supplementary Table 1-3.

Feasibility and limit of detection using large libraries of DNA-barcoded MHC multimers for T cell analyses

To provide proof-of-feasibility for staining antigen-specific T cells in mixtures of >1000 different pMHC multimers, we compared detection of various T cell populations responsive to virus-derived peptides using DNA-barcoded MHC multimers or combinatorial fluorescently-labeled MHC multimers, respectively[4,12].We verified that PE labeled MHC multimers carrying DNA barcodes were able to bind specifically to their target T cell population even in excess of 999 irrelevant pMHC multimers, and that the DNA barcode associated with positive control reagents could be specifically recovered after isolation of MHC multimer binding cells (Supplementary Fig. 1a-d).

To investigate the linearity of detection and the detection limit of the method we prepared a titration curve of antigen-specific T cells by mixing peripheral blood mononuclear cells (PBMCs) from one donor (BC260) who carried ~5% HLA-B0702_CMV_pp65$_{TPR}$-specific T cells into an HLA-B0702-negative donor (BC262) (Fig. 2a). T cell recognition of a panel of 1031 pMHC molecules (Supplementary Table 4), each individually multimerized and encoded with a specific DNA barcode, was assessed. Peptides were divided into different categories based on the origin of their antigen. These categories included virus-derived antigens, melanoma-associated antigens[13], Merkel cell polyomavirus-derived antigens[14], renal cell carcinoma-associated antigens and breast cancer-associated antigens. Very few responses were detected, except within the virus-derived category (Fig. 2b and Supplementary Fig. 2A). Here, we observed T cell responses to three epitopes in the '5% sample' (100% of BC260 PBMCs), i.e., HLA-B*0702 CMV$_{TPR}$ (4.6% of CD8 T cells), HLA-A*0201 CMV$_{NLV}$ (0.15% of CD8 T cells), and HLA-A*0201 EBV$_{YVL}$ (0.06% of CD8 T cells). In the '1% sample' (20% of BC260 PBMCs and 80% of BC262 PBMCs), we observed T cell reactivity toward HLA-A*1101 EBV$_{AVF}$ (0.1% of CD8 T cells) and HLA-A*0201 EBV$_{GLC}$ (0.04% of CD8 T cells) originating from BC262. Although the two low-frequency BC260-associated responses, HLA-A*0201 CMV$_{NLV}$ and EBV$_{YVL}$ disappeared, the HLA-B*0702 CMV$_{TPR}$ response remained detectable. Stepwise 5-fold titrations led to a corresponding reduction in the detection of

the HLA-B*0702 CMV_{TPR}-associated DNA barcode and a loss of significance in terms of sequence-read numbers at 0.008% of CD8 T cells, corresponding to an average of 20 MHC multimer-positive T cells in a flow cytometry-based assessment (Fig. 2a-c). To compare T cell detection based on DNA barcode-labeled MHC multimers to state-of-the-art flow cytometry based assessment of pMHC specific T cell frequencies, the number of clonally reduced barcode reads was applied to estimate the frequency of antigen-responsive T cells, based on an average number of TCR-MHC multimer interactions in the total MHC multimer-binding T cell pool in a given cell sample. This estimate was obtained from clonally reduced barcode reads retrieved from a given sample and the frequency of all MHC multimer-positive T cells in that sample using the following equation: (number of DNA barcode reads associated with a specific pMHC/number of total barcode reads derived from the same sample) x percentage of MHC multimer-binding cells among total CD8 T cells (Supplementary Table 5). The estimated frequencies of pMHC-responsive T cells correlated strongly with the frequencies determined by combinatorial fluorescently-labeled MHC multimers (Fig. 2d). When the HLA-B*0702 CMV_{TPR}-responsive T cells were assessed in an identical set of samples using a smaller library of MHC multimers, n=110 (Supplementary Table 6), we observed comparable detection limit (0.001% of CD8 T cells) and strong correlation with detection via fluorescently labeled MHC multimers (r^2=0.999) (Supplementary Fig. 3a-d). Moreover, strong correlations were observed when results from the 1031 MHC multimer library were compared with the 110 MHC multimer library, both related to the estimated frequencies of HLA-B*0702 CMV_{TPR}-specific T cells (r^2=0.999) and the clonally reduced read counts (r^2=0.805), respectively (Supplementary Fig. 3d-e). The number of cells used for analysis could be increased at least to $10x10^6$ cells per sample while maintaining specific MHC multimer staining. Such 5-fold increase in the number a precursor cells seems to increase the limit of detection (Fig. 2e).

Detection of pMHC-responsive T cells across different donors

We continued using the 1031 member MHC multimer library to screen for T cell reactivity in PBMCs across 10 different healthy donors with various HLA types (Supplementary Table 7). We observed a T cell response signature dominated by reactivity toward virus-derived peptide sequences (Fig. 3a). Focusing on the virus-derived peptides, we observed a highly HLA-dependent signature of T cell reactivity (Fig. 3b). BC171 expressed HLA-A*1101, A*0301, B*0702, and B*1501, and the detected T cell reactivity corresponded to these tissue types. BC260, 259, 261, 261, 268, 254, and 251 expressed HLA-A*0201, and they showed variable T cell reactivity towards HLA-A*0201-restricted virus-derived epitopes. BC268 expressed HLA-A*0101 and HLA-B*0801 and showed both CMV- and FLU-specific reactivity restricted to these HLA molecules. Among the 10 healthy donors screened with 1031 different pMHC multimers, we identify only one response not matching the HLA type of the given individual (Fig. 3b, HLA-B*0801, HSV pU79$_{EGR}$, BC251). T cell reactivity against virus-derived peptides was determined by flow cytometry in parallel. The frequencies of antigen-responsive T cells determined using the two methods correlated strongly, r^2=0.967 (Fig. 3c). Of 42 potential responses, 88% were detected either by fluorescent labeled MHC multimers or by DNA-barcoded MHC multimers. The five responses detected by flow cytometry but not with DNA-barcoded MHC multimers, were all present at low frequencies in the analyzed cell samples (<0.01 % of CD8 T cells), and four of these were just below the threshold of 0.1% FDR in the DNA-barcode based analysis. The additional T cell specificities detected with DNA-barcoded MHC multimers and not with fluorescent labeled MHC multimers is likely to represent specific detection because they all match the HLA-expression and virus-response profiles of the given donors.

When PBMCs from six of the same donors were analyzed using a library of 110 DNA-barcoded MHC multimers (Supplementary Fig. 4 and Supplementary Table 6), we again observed strong

correlations when comparing T cell reactivity detected using DNA-barcoded MHC multimers and that detected using combinatorial fluorescently labeled MHC multimers, r^2=0.985 (Supplementary Fig. 4c), and when comparing results from the 110 versus the 1031 member pMHC library, r^2=0.951 (Supplementary Fig. 4d-e). Thus, indicating that library size does not considerably affect the capacity for detection of specific T cells using DNA-barcoded MHC-multimers.

Detection of tumor-reactive T cells

Recent clinical success in cancer immunotherapy has led to great interest in examining T cell reactivity against cancer. To demonstrate the feasibility of using DNA-barcoded MHC multimers for the detection of tumor-restricted pMHC-specific T cells, we analyzed T cell reactivity among tumor-infiltrating lymphocytes (TILs) from melanoma patients against a previously described library of shared melanoma-associated peptides[13,15]. Antigen-specificity was assessed within 11 expanded TIL products against an HLA-A*0201-restricted DNA-barcoded pMHC multimer library of 167 melanoma-associated peptides and 8 virus-derived peptides (Supplementary Table 8). We detected numerous T cell populations responsive to melanoma associated antigens in 8 of the 11 TIL products (Fig. 4A-B). In comparison the same library of melanoma-associated epitopes was included for T cell screening in the healthy donor cohort resulting in very few detected responses (Fig. 3a). Interestingly, two patients not responding to these HLA-A*0201 presented peptides, were evaluated as HLA-A*0205 positive upon detailed HLA typing (MM10 and MM11, Supplementary Table 7), adding to previous observations that HLA micropolymorphisms may strongly affect T cell detection using MHC multimers[16,17]. T cell reactivity was detected against several well-known melanoma-associated antigens, such as MART-1, NY-ESO-1, gp100 and tyrosinase. We observed reactivity against all three variants of the MART-1 peptide (ELAGIGILTV, EAAGIGILTV, and AAGIGILTV) within the same TIL samples in patients MM02 and MM03. This finding is likely to

reflect partial cross-recognition of all three peptides from the same T cell population, a phenomenon that has been described previously[18–21]. Additionally, T cell reactivity against two different variants of the gp100 peptide (ITDQVPFSV and IMDQVPFSV) and variants of the NY-ESO-1 peptide (SLLMWITQC and SLLMWITQA) was detected in patient MM01 and MM02, respectively (Fig. 4b). Similar to our observations in the healthy donor cohort, there was a tight correlation between the number of antigen-responsive T cells estimated based on association to DNA-barcoded MHC multimers, and that detected using combinatorial fluorescently labeled MHC multimers, r^2=0.901 (Fig. 4c). In the melanoma samples, DNA barcode-based screening resulted in detection of an increased numbers of melanoma-associated peptide-specific T cell populations (n=45) compared to screening with fluorescently labeled MHC multimers (n=16). The additional peptide-restricted populations detected by DNA-barcoded MHC multimers showed a patient-specific pattern that matched the HLA expression and antigen-response signature of the patient (Fig. 4a-b) and the detection of an increased number of tumor-specific T cell populations may reflect that T cell detection based on DNA barcode-labeled MHC multimers provides less dependence on stringent fluorescence-based separation (Supplementary Fig. 5a-b).

Detection of tumor-reactive T cells from small-size clinical samples

One major advantage of multiplex technologies for T cell detection is the possibility of determining the composition of antigen-specific T cells in small biological samples without a need for lymphocyte expansion. We utilized the high-throughput screening capacity of DNA-barcoded MHC multimers to study the dynamics of T cell responses in various samples from two patients with metastatic melanoma participating in a phase II trial, with adoptive cell therapy using in-vitro expanded TILs[22]. We screened for T cell recognition towards a large library of virus- and shared melanoma-derived epitopes (328 barcoded pMHC multimers, Supplementary Table 9) in: a)

uncultured tumor fragments following enzymatic digest, i.e. unexpanded TILs (digest), b) TILs in vitro expanded from small tumor fragments, dissected from the same metastatic lesion and c) peripheral blood obtained before and approximately 1 month after infusion of expanded TILs. Although very few lymphocytes were available in the melanoma samples directly after enzymatic digest (18,000 and 48,000, respectively for MM01 and MM02), we detected numerous melanoma-associated T cell responses in these tissue samples (Fig. 5a-b). In MM01 two of these responses were also detected in the expanded TILs and in peripheral blood after transfer (Fig. 5a). The composition of T cell responses in peripheral blood after adoptive transfer was to a large extent matching the composition of the TIL infusion product. In MM02, T cell responses detected in peripheral blood after TIL transfer reflected the composition of the TIL infusion product, but also additional responses were detected. These additional responses matched those detected in a digest sample taken from a tumor lesion after tumor progression (Fig. 5b). Overall, we detected an increased response to melanoma associated antigens in peripheral blood after adoptive transfer of TILs, reflecting T cell populations detected in both tumor digest samples and expanded TIL products.

Detection of neoepitope-responsive T cells in non-small cell lung carcinoma

Mutation-derived neoepitopes have been suggested as important targets for tumor rejection mediated by T cells, and the predicted presence of neoepitopes correlates positively with clinical outcome following treatment with immune checkpoint inhibitors[23,24]. Consequently, efficient means of identifying neoepitope-restricted T cells are of potential prognostic and therapeutic value. We analyzed two non-small cell lung carcinoma (NSCLC) patients for T cell recognition of predicted HLA-binding peptide sequences containing cancer-specific mutations, i.e., potential neoepitopes. For each patient, a large personal neoepitope peptide library was identified, n=288 and

n=417 for patient L011 and L013, respectively (described in Supplementary Methods). T cell reactivity of in vitro-expanded TILs was assessed using both combinatorial fluorescently labeled MHC multimers and DNA-barcoded MHC multimers. The results from the first approach are reported in detail in McGranahan et al, 2016[25], and resulted in detection of one neoepitope responsive T cell population in L011. Using DNA-barcoded MHC multimers, we observed nine neoepitope responsive T cell populations and one virus responsive T cell population in these two patients (Fig. 6a-b). The use of DNA barcode-labeled MHC multimers allowed us to screen the whole library in one tube (4-8 mill TILs) rather than using 10 parallel tubes (1 mill TILs per tube) for the fluorescent-based analysis. This increased number of cells may account for the detection of a greater number of T cell populations using the DNA-barcoded MHC multimers. In patient L013 we analyzed TILs from three different tumor regions, a fragment of adjacent normal lung tissue, and peripheral blood taken at the time of tumor removal. For this patient T cells responsive to the three neoepitopes were most prominently detected in peripheral blood (Fig. 6b), but could be detected, either significantly (ALQ), or as lower enriched populations (YSN and KVC) in T cell cultures from the different regions. CMV-responsive T cell populations were detected in all regions and in PBMCs. Detection of neoepitopes in peripheral blood was not feasible using combinatorial fluorescently labeled MHC multimers, due to the limited samples size. Detection of such responses directly from peripheral blood has the potential to significantly ease personalized T cell therapy approaches and immunomonitoring.

Assessing the functional reactivity among pMHC-responsive T cells

To understand the functional capacity of pMHC responsive T cells, we combined the assessment of cytokine production following target cell stimulation and the binding to pMHC multimers. In three healthy donor PBMC samples we analyzed for functional reactivity towards a pool of CMV, EBV,

and FLU-derived peptides (CEF pool) using intracellular cytokine staining (ICS) and subsequently stained with a library of 328 DNA-barcoded pMHC multimers (Supplementary Table 9 and Supplementary Methods). T cells were sorted solely based on their cytokine secretion profile, INFγ and TNFα secretion (ICS^{pos}) or no cytokine secretion (ICS^{neg}), but the pMHC associated DNA barcode could retrospectively reveal the antigen specific composition of these different cell-subsets. Functional reactivity was observed for all virus responsive T cells previously detected in these PBMCs. 8 in 11 virus specific T cell populations displayed full functionality, with none of these T cells detected in the ICS^{neg} fraction (Fig. 7a). Using the same principle we analyzed for functional reactivity of TILs towards autologous tumor cells in a melanoma patient (Fig. 7b-c). We detected eight T cell populations responsive to melanoma-associated HLA-A*0201-restricted T cell epitopes. These T cell populations were either exclusively or predominantly detected in the ICS^{pos} populations, indicating antigen processing and presentation on tumor cells and functional capabilities among the epitope-specific TILs. In contrast, the four virus-restricted T cell populations present in this TIL product were detected only in the ICS^{neg} population, as expected, since the virus responsive T cells would not have met their target on autologous tumor cells. These analyses prove the feasibility for combining T cell detection using large libraries of DNA-barcoded pMHC multimers with assessment of functional responsiveness upon target recognition.

DISCUSSION

We report herein a technology that allows multi-parallel detection of >1000 different antigen-specific CD8 T cell populations in a single sample using DNA-barcoded MHC multimers. The potential complexity of DNA barcode tags exceeds 10^{8} [ref10], and the functional limitations rely on the ability to achieve specific TCR-pMHC interactions within a very large pool of different pMHC multimers. Limitations in terms of maximal number or concentration of MHC multimers hindering

specific TCR-pMHC multimer interactions have not been reached. The complexity may likely be extended to libraries of 10,000-100,000 MHC multimers, but this will depend to some extent on the automation of MHC multimer production, high-throughput peptide synthesis and microfluidics systems ensuring both efficient pMHC interaction with all T cells in a given sample and that specific interaction of the TCR with its cognate pMHC recognition motif will not be outmatched by numerous CD8:MHC constant region interactions.

The detection of antigen-specific CD8 T cells using DNA-barcoded MHC multimers provides a detection limit similar to that of fluorescence-based readouts and correlates tightly with the frequency of antigen-responsive T cells found using combinatorial fluorescently labeled MHC multimers. We could detect several T cell populations within the same sample in the frequency range of 20%-0.01% of CD8 T cells (Figs. 2-4). We showed that this limit of detection may be increased by increasing the number of T cells analyzed (Fig. 2e and 6a-b). Especially in the cohort of tumor-derived material, we detected an increased number of pMHC-responsive T cell populations when the reactivity was assessed using the DNA-barcoded MHC multimers compared to the fluorescently labeled MHC multimers. This increased detection of specific T cells may reflect that T cell detection based on DNA barcode-labeled MHC multimers provides less dependence on stringent fluorescence-based separation, because the T cell specificities are determined by sequencing the pMHC multimer associated DNA barcode and not based on the fluorescent signal as a single parameter (Supplementary Fig. 5a-b). This characteristic may allow enhanced detection of low-avidity T cells, since the DNA barcode label is readily recovered from those T cells that bind specifically to the MHC multimer reagents even when they only obtain a marginally higher fluorescent intensity than the MHC multimer negative population, but are still included in the sorted fraction of T cells. Tumor-reactive T cells, especially those recognizing shared self-peptide epitopes, often demonstrate lower affinity for the pMHC recognition motif compared to virus-

specific T cells[26]. Additionally, the use of the dextran backbone, providing a higher-order multimer, as opposed to the streptavidin-based tetramerization conventionally used for combinatorial fluorescently labeled MHC multimers, may further assist the detection of low-avidity T cells[27,28]. Dextran reagents are however available carrying only three different fluorescent labels and are therefore not optional for multiplex fluorescently-based assays.

In numerous diseases, it is crucial to assess T cell recognition at the relevant disease site using, for example, tissue biopsies. Due to the complexity of the present technology, we allow assessment of T cell reactivity toward large libraries of MHC multimers even in small-size biological samples. Within two melanoma patients, we compared T cell reactivity directly from tissue digest with in vitro-expanded T cell product and peripheral blood before and after TIL adoptive transfer (Fig. 5a-b). We prove the feasibility for detection of tumor-reactive T cells in small biological samples, which allow interrogation of the dynamics in T cell recognition across large peptide libraries. Previous results showed loss of antigen recognition after in vitro expansion of T cells[13], which indicates that direct ex vivo assessments of T cell reactivity will provide novel insights into T cell recognition at the disease site. For diagnostic purposes and for response-to-therapy evaluations, peripheral blood samples would be the preferable material to analyze due to the ease of access. However, the inadequate complexity and the limit of detection of MHC multimer-based analysis previously prohibited direct detection of tumor antigen-responsive T cells from peripheral blood using large peptide libraries. Here, we showed the feasibility of detecting neoepitope-reactive T cells when screening peripheral blood with large libraries of pMHC multimers selected based on personalized tumor mutagenome analyses (Fig. 6a-b). Screening for T cell reactivity against large libraries of neoepitopes using the DNA-barcoded MHC multimer one-pot approach led to enhanced detection of neoepitope directed T cell reactivity compared to the use of combinatorial fluorescently labeled MHC multimers that required testing of numerous parallel samples.

We demonstrate that it is feasible to combine the use of DNA-barcoded MHC multimers with functional analysis. Upon peptide or tumor cell stimulation we selected cytokine-secreting T cells and showed that functional capacity and target recognition can be assessed across large peptide libraries (Fig. 7). Such functional assessments are valuable for understanding which antigens form potential tumor rejection epitopes. A future strategy could combine DNA-barcoded MHC multimer staining with single-cell sorting and gene expression profiling approaches [29–31]. The analyses presented herein, based on bulk assessments of T cell populations, largely reflect the information that can be obtained from bulk analyses using fluorescently labeled pMHC multimers (Fig. 2-7), but represents a quantitative leap in the number of potential epitopes that can be assessed in parallel. This in turn offers great potential for advancing T cell assessments to genome-wide levels. Moreover, it will allow such screens to be conducted not only on in vitro-expanded material but also directly on tissue biopsies and peripheral blood.

Author Contributions

AKB conceived the concept, designed and performed experiments, analyzed data, made figures and wrote the manuscript; AMM designed the in silicon analysis, analyzed data, and made figures; RL, SKS, and SRA, produced MHC monomers, designed and performed experiments, analyzed data and revised the manuscript; LS performed experiments, MD and IMS, provided patients material and generated tumor cell lines, discussed data; PtS provided administrative support, flow facility and production of MHC monomers; AF, SAQ, provided patients material and peptides from NSCLC; NMG, RR, CAS, identified tumor mutagenome and predicted NSCLC-associated neoepitopes; ZS and ACE, designed the in silicon analysis, and guided data analyses; SNJ designed DNA barcodes, conceived the concept, guided data analyses and revised the manuscript; SRH conceived the concept, designed experiments, analyzed data and wrote the manuscript.

Acknowledgments

We would like to thank Ulla Kring Hansen, Anna Burkal and Tina Seremet for technical assistance, and Ton Schumacher, Netherlands Cancer Institute for scientific discussions and sharing of MHC expression plasmids. The work was funded by The Danish Cancer Society, The Lundbeck Foundation and The Danish Research Council.

Competing financial interests: The technology is patented by the authors (SRH, AKB, and SNJ), the Capital Region of Copenhagen, University Hospital Herlev and licensed to Immudex.

References

1. Altman, J. D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274,** 94–96 (1996).

2. Chattopadhyay, P. K. & Roederer, M. Cytometry: today's technology and tomorrow's horizons. *Methods* **57,** 251–8 (2012).

3. Bendall, S. C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332,** 687–96 (2011).

4. Hadrup, S. R. *et al.* Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat.Methods* **6,** 520–526 (2009).

5. Newell, E. W., Klein, L. O., Yu, W. & Davis, M. M. Simultaneous detection of many T-cell specificities using combinatorial tetramer staining. *Nat.Methods.* **6,** 497–499 (2009).

6. Newell, E. W. *et al.* Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. *Nat. Biotechnol.* **31,** 623–9 (2013).

7. Newell, E. W., Sigal, N., Bendall, S. C., Nolan, G. P. & Davis, M. M. Cytometry by time-of-

flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. *Immunity* **36,** 142–52 (2012).

8. Robins, H. S. *et al.* Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114,** 4099–107 (2009).

9. Davis, M. M. & Bjorkman, P. J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334,** 395–402 (1988).

10. Xu, Q., Schlabach, M. R., Hannon, G. J. & Elledge, S. J. Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 2289–94 (2009).

11. Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* **9,** 72–74 (2011).

12. Andersen, R. S. *et al.* Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. *Nat. Protoc.* **7,** 891–902 (2012).

13. Andersen, R. S. *et al.* Dissection of T-cell antigen specificity in human melanoma. *Cancer Res.* **72,** 1642–50 (2012).

14. Lyngaa, R. *et al.* T-cell responses to oncogenic Merkel cell polyomavirus proteins distinguish patients with Merkel cell carcinoma from healthy donors. *Clin. Cancer Res.* **20,** 1768–1778 (2014).

15. Frøsig, T. M. *et al.* Broadening the repertoire of melanoma-associated T-cell epitopes. *Cancer Immunol. Immunother.* **64,** 609–20 (2015).

16. van Buuren, M. M. *et al.* HLA Micropolymorphisms Strongly Affect Peptide-MHC Multimer-Based Monitoring of Antigen-Specific CD8+ T Cell Responses. *J. Immunol.* **192,** 641–8 (2014).

17. Frøsig, T. M. *et al.* Design and validation of conditional ligands for HLA-B*08:01, HLA-B*15:01, HLA-B*35:01, and HLA-B*44:05. *Cytometry. A* **87,** 967–75 (2015).

18. Valmori, D. *et al.* Enhanced generation of specific tumor-reactive CTL in vitro by selected Melan-A/MART-1 immunodominant peptide analogues. *J. Immunol.* **160,** 1750–8 (1998).

19. Derré, L. *et al.* A novel population of human melanoma-specific CD8 T cells recognizes Melan-AMART-1 immunodominant nonapeptide but not the corresponding decapeptide. *J. Immunol.* **179,** 7635–45 (2007).

20. Wieckowski, S. *et al.* Fine structural variations of alphabetaTCRs selected by vaccination with natural versus altered self-antigen in melanoma patients. *J. Immunol.* **183,** 5397–406 (2009).

21. Speiser, D. E. *et al.* Unmodified self antigen triggers human CD8 T cells with stronger tumor reactivity than altered antigen. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3849–54 (2008).

22. Andersen, R. Long-lasting complete responses in patients with metastatic melanoma after adoptive cell therapy with tumor-infiltrating lymphocytes and an attenuated IL-2 regimen. *Clin. Cancer Res. Accept. Publ.* (2016).

23. Snyder, A. *et al.* Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371,** 2189–99 (2014).

24. Rizvi, N. A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348,** 124–8 (2015).

25. McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351,** 1463–9 (2016).

26. Aleksic, M. *et al.* Different affinity windows for virus and cancer-specific T-cell receptors:

Implications for therapeutic strategies. *Eur. J. Immunol.* **42,** 3174–3179 (2012).

27. Dolton, G. *et al.* Comparison of peptide-major histocompatibility complex tetramers and dextramers for the identification of antigen-specific T cells. *Clin. Exp. Immunol.* **177,** 47–63 (2014).

28. Massilamany, C. *et al.* Direct staining with major histocompatibility complex class II dextramers permits detection of antigen-specific, autoreactive CD4 T cells in situ. *PLoS One* **9,** e87519 (2014).

29. Han, A., Glanville, J., Hansmann, L. & Davis, M. M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* **32,** 684–692 (2014).

30. Dössinger, G. *et al.* MHC multimer-guided and cell culture-independent isolation of functional T cell receptors from single cells facilitates TCR identification for immunotherapy. *PLoS One* **8,** e61384 (2013).

31. Klein, A. M. *et al.* Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell* **161,** 1187–1201 (2015).

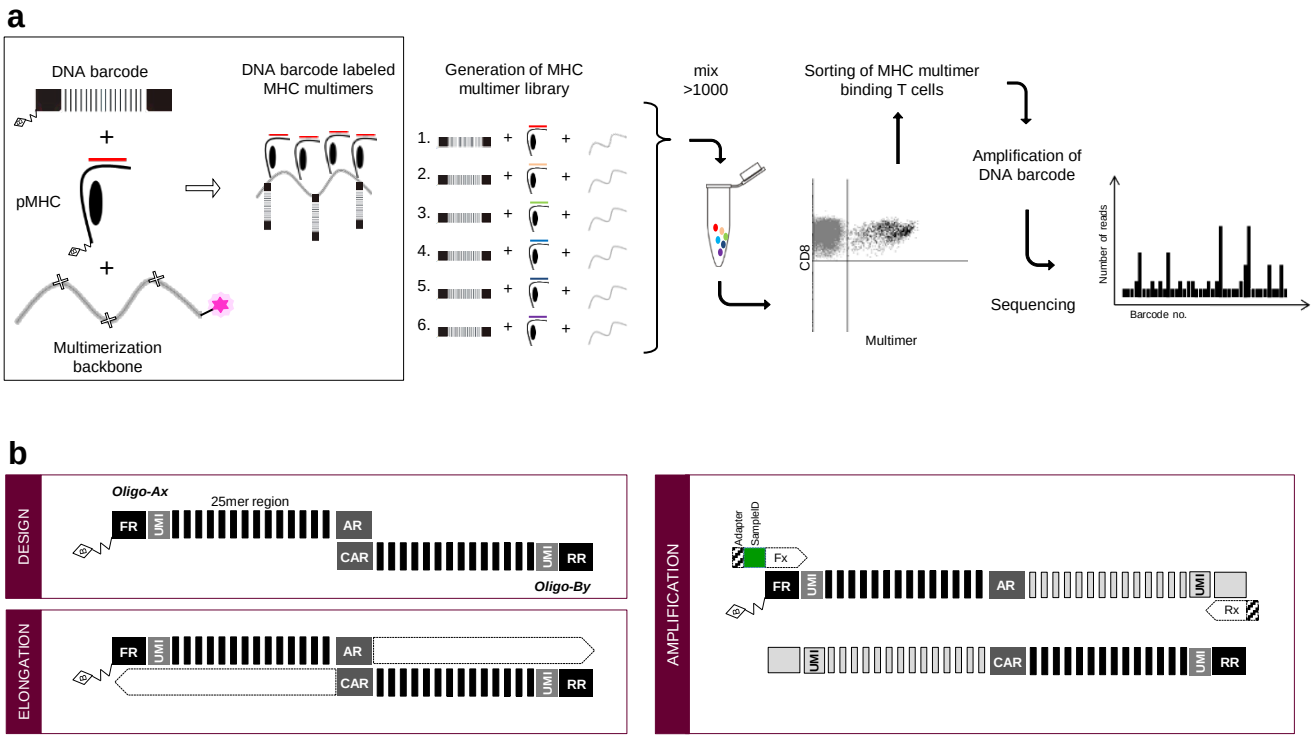


Figure 1. Preparation and use of DNA barcode-labeled MHC multimers

(**a**) Schematic overview showing the strategy for using DNA barcode-labeled MHC multimers for detection of antigen-specific T cells in complex cellular suspensions. Biotinylated DNA barcodes and pMHC molecules are co-attached to a PE-labeled dextran backbone carrying streptavidin; B=biotin. Each MHC multimer is assembled with a given DNA barcode, forming a tag for the corresponding specificity (1 to 1000). MHC multimer-binding T cells are isolated and sorted based on the PE label. DNA barcodes are amplified and sequenced, and the relative numbers of DNA barcode reads are used to determine the composition of antigen-responsive T cells in the sample. (**b**) Schematic overview of the DNA barcode design. B=biotin, FR= forward region, UMI=Unique Molecular Identifier, coding region= 25mer barcode sequence assigning pMHC specificity, AR=annealing region, CAR= complementary annealing region, and RR= reverse region. The biotinylated oligo Ax comprise a 15-nucleotide region partially complementary to oligo By. Oligos Ax and By both contain an individual 25mer oligonucleotide sequence (determined by the 'x or y') and six randomly incorporated nucleotides, providing a UMI for each synthesized oligo. Oligo A contain a forward primer region (FR), and Oligo B contain a reverse primer region (RR) (the oligonucleotide sequences are listed in Supplementary Table 1-2). Following annealing of Oligo Ax and Oligo By and prior to the attachment to the multimerization backbone, the oligos are elongated to obtain their double-stranded form. After the isolation of MHC multimer-binding T cells, the DNA barcodes are amplified by PCR (Supplementary Table 3). The forward primer is flanked by a 5' sample identifier sequence (sampleID), and both the forward and reverse primers encodes a 5' sequencing adaptor sequence (IonTorrent, A-Key and P1-key, respectively) (Supplementary Table 3).

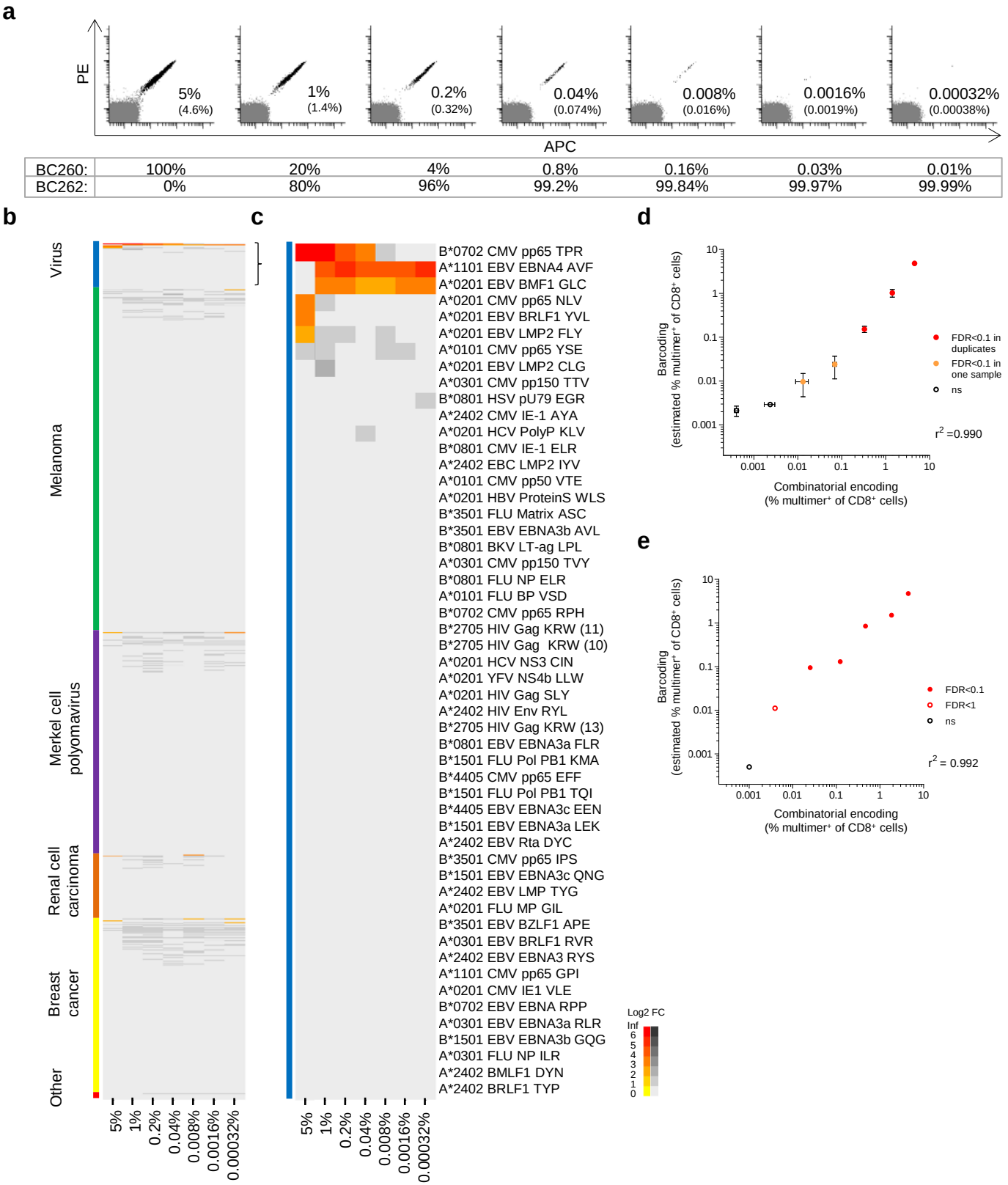


Figure 2. Dynamic range and limit of detection of DNA-barcoded MHC multimers

(**a**) Fluorescent multimer-based assessment of seven samples with various proportions of B*0702 CMV_{TPR}-specific T cells, theoretically: 5%, 1%, 0.2%, 0.04%, 0.008%, 0.0016% and 0.00032% of CD8 T cells. Samples are generated from BC260 PBMCs mixed with HLA-B7-negative BC262 at fivefold dilutions. Total percentage of cells derived from each donor is indicated below the corresponding dot plot. The flow-based percentages of MHC multimer$^+$ CD8 T cells are given in brackets (detection limit: ≥ 10 events and > 0.002% of CD8 T cells). (**b**) Heatmap representing the DNA barcode-based analysis of samples from (a). Each sample was screened with a panel of 1031 pMHC multimers. The heatmap shows changes in read proportions compared to background levels (Log2FC, see Supplementary Methods, Statistical analysis). Each column represents the reads associated with a given sample and each row a given antigen specificity, i.e. reads mapped to the DNA barcode associated with the corresponding pMHC. Epitopes are grouped based on their antigen origin. Within each antigen group rows are sorted based on Log2FC, highest to lowest, compared to baseline samples. Yellow-red or gray scaling, respectively, indicates statistically significant (FDR<0.1%) or insignificant (FDR≥0.1%) number of reads. A duplicate analysis is shown in Supplementary Figure 2. (**c**) Magnification focused on the virus-derived peptides. First row represents the target specificity B*0702 CMV_{TPR} present in BC260. 2nd and 3rd rows below are T cell responses from the HLA-B*0702 negative donor BC262, followed by three lower-frequent responses that are present in BC260. (**d, e**) Correlations between the frequency of HLA-B*0702 CMV_{TPR}-specific T cells determined by analyzing the same samples using combinatorial fluorescently labeled pMHC multimers or a panel of 1031 DNA-barcoded pMHC multimers, (**d**) when staining $2x10^6$ PBMCs per sample or (**e**) $10x10^6$ PBMCs per sample. Error bars represent range of duplicates. All pMHC multimers are 'dextramers'. A similar experiment using another pMHC multimer library is shown in Supplementary Figure 3.

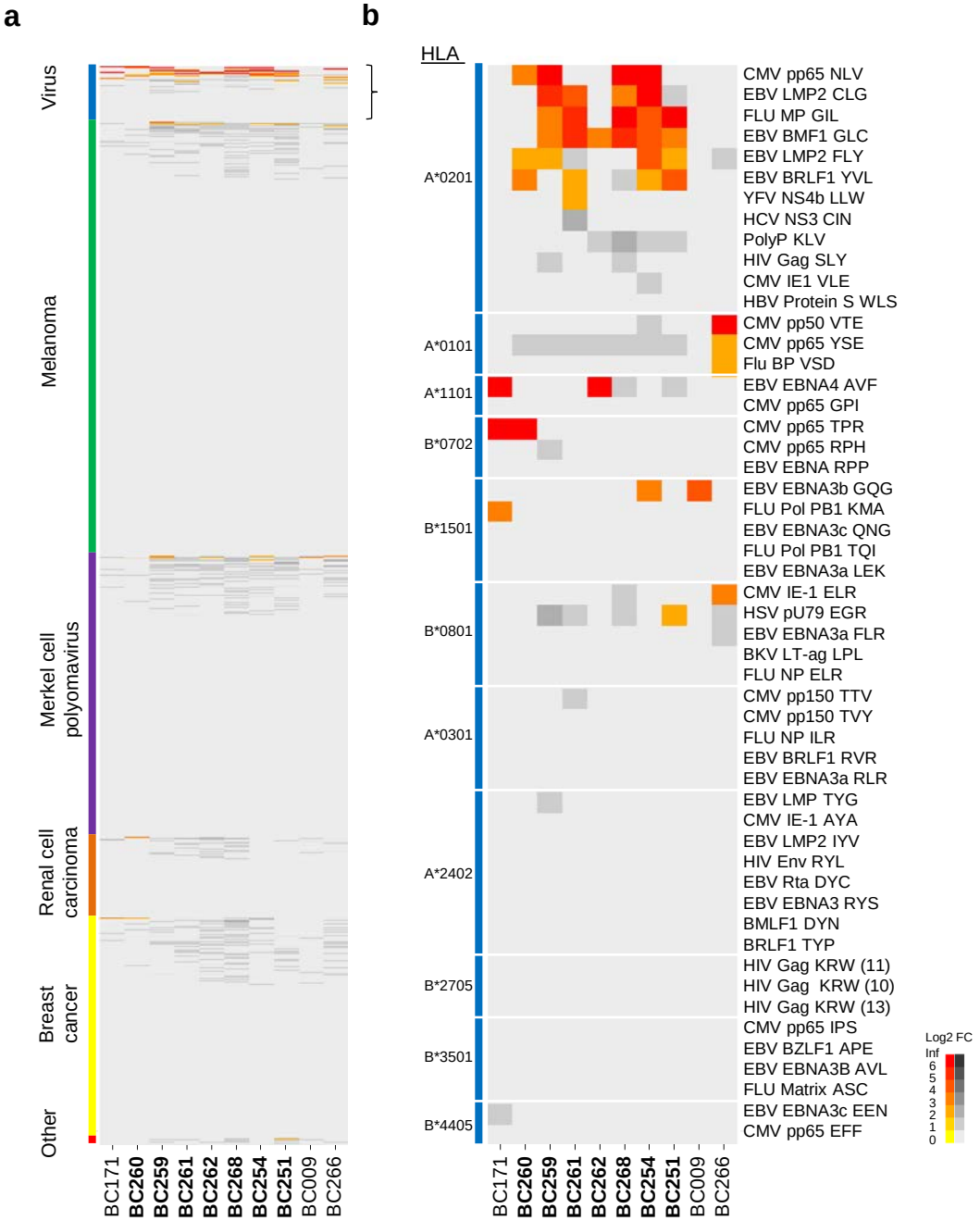
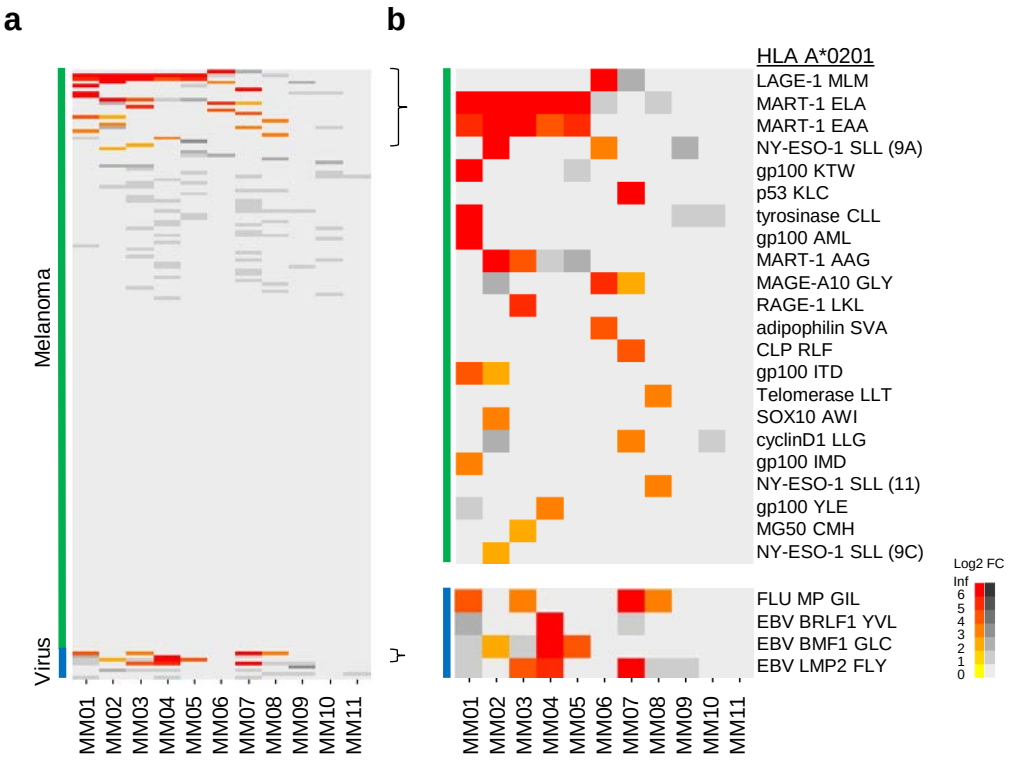


Figure 3. High-throughput assessment of T cell reactivity using large peptide libraries

(**a**) DNA barcode-based analysis of 10 different healthy donor PBMC samples (1-2x10^6 PBMCs) using 1031 pMHC multimers, each carrying individual barcodes. The heatmap is organized as in Figure 2b, each column represents one donor. Donors marked in bold are all HLA-A*0201. Further details on HLA-types of all donors are presented in Supplementary Table 7. (**b**) Magnification focused on the virus-derived peptides (52 epitopes). Rows representing antigen specificities are grouped according to HLA-type and sorted within each group based on Log2FC. Each sample shown here was analyzed once, but replicates with six of the same donors using another pMHC library, shown in Supplementary Figure 4. (**c**) Correlations between antigen-specific T cell frequencies, from analyses of T cell responses in 10 healthy donors using either combinatorial fluorescently labeled pMHC multimers (x-axes) or 1031 DNA-barcoded pMHC multimers (y-axes). Each dot represents one specificity. Only T cell populations that fulfilled the significance criteria for DNA barcode assessment (FDR<0.1%) or the threshold for fluorescent-based analysis (≥ 10 events and > 0.002% of CD8 T cells) are plotted. All specificities included in the plot were tested using both a combinatorial encoding analysis and DNA-barcoded MHC multimers. Dots plotted on the axes are nonsignificant for one of the methods.



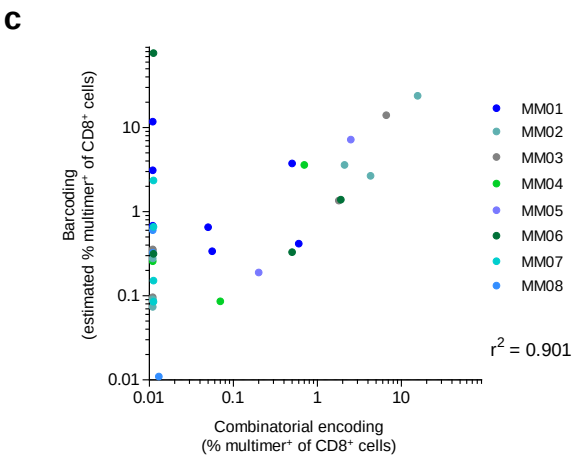


Figure 4. High-throughput assessment of tumor-reactive T cells (**a**) Heatmap representing the barcode-based analysis of 11 different samples of in vitro-expanded tumor-infiltrating lymphocytes (TILs) using a library of 167 HLA-A*0201 melanoma-associated pMHC multimers and 8 virus-derived pMHC multimers (175 pMHC library, Supplementary Table 8). The heatmap is organized as in Figure 2b. Donor MM9, MM10 and MM11 carries the HLA-A*0205 subtype (Supplementary Table 7). Each TIL culture were analyzed once. Three TIL cultures were reanalyzed with a larger pMHC library, with similar results (data not shown) (**b**) Magnification focusing on significant responses among both categories of peptides. (**c**) Correlations between antigen-specific T cell frequencies determined across the 11 different TIL samples from melanoma patients analyzed using combinatorial fluorescently labeled pMHC multimers (x-axes) or DNA-barcoded pMHC multimers (y-axes). Each dot represents one specificity. Only T cell populations that fulfilled the significance criteria described in Figure 3c were plotted. All specificities were tested using both a combinatorial fluorescently labeled pMHC multimers and DNA-barcoded pMHC multimers. 0.5-2x10^6 TILs were analyzed per sample in both methods.

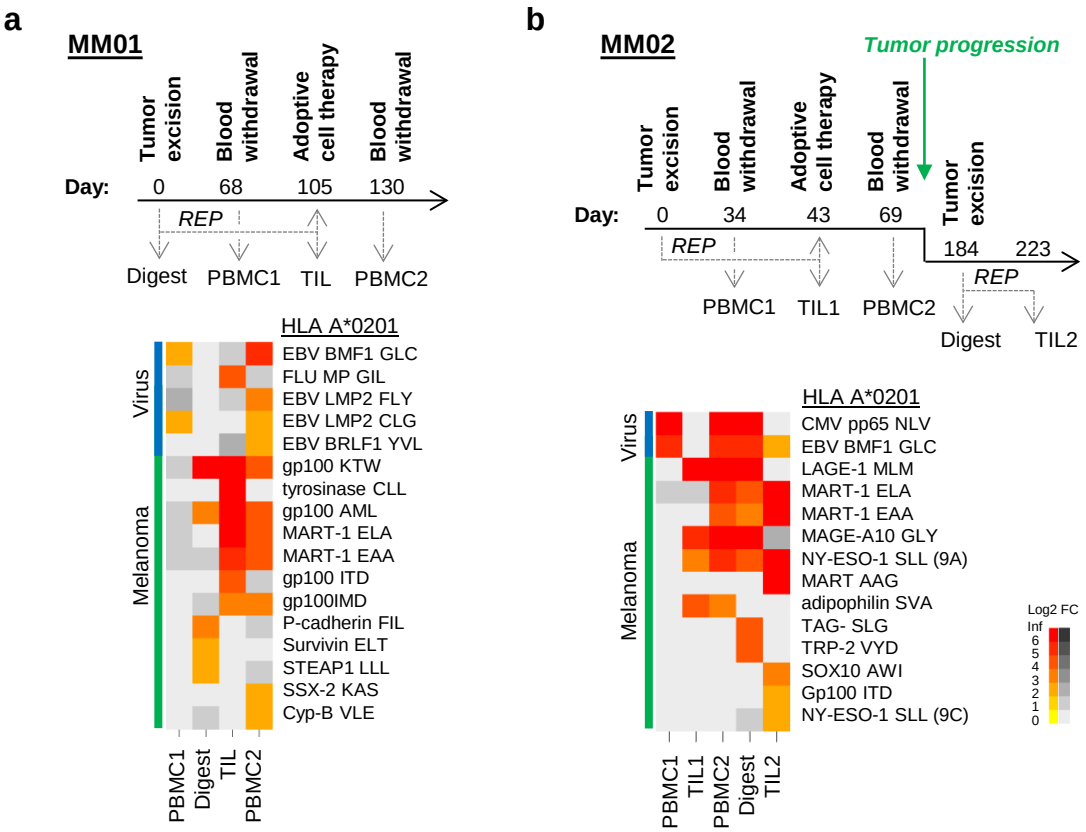
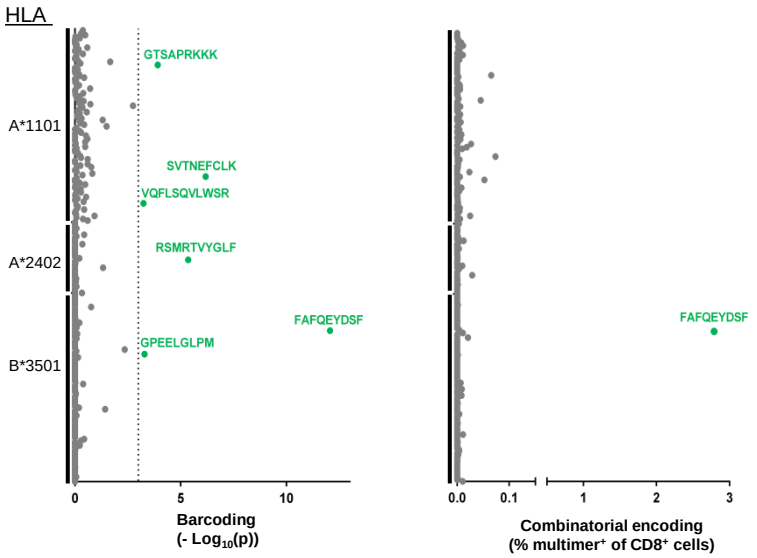


Figure 5. T cell assessment of limited biological samples
(**a**, **b**) Analysis of dynamic changes in T cell response to melanoma-associated antigens (175 pMHC library, Supplementary Table 8) before and after TIL adoptive cell transfer in two patients with metastatic melanoma. A timeline of sample collection and TIL adoptive transfer are presented for each patient together with a heatmap focusing on T cell specificities detected in any of the samples from (**a**) patient MM01 and (**b**) patient MM02. The heatmaps are organized as described in Figure 2b. All samples were analyzed once.

a



b

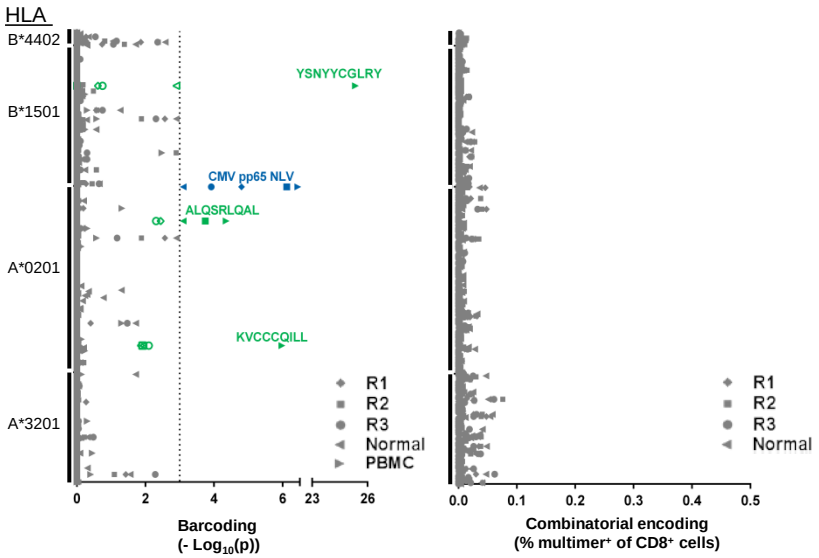


Figure 6. Detection of neoepitope responsive T cells in lung cancer
(**a**) Screening for T cell recognition of 288 predicted neoepitopes and 10 HLA-matched virus-derived peptides in T cells expanded from a resected lesion in patient L011 using either combinatorial fluorescently labeled MHC multimers ($1x10^6$ viable cells per tube in 9 tubes) or DNA-barcoded MHC multimers ($5x10^6$ viable cells per tube in 1 tube). Results from all 288 pMHC multimers included in the screening are plotted (x-axis). Data plotted on the y-axis is the percentage of MHC multimer positive T cells of total CD8 T cell for combinatorial fluorescently labeled MHC multimers, or -Log10(p) (in respect to the pMHC associated DNA barcode) for the DNA-barcoded MHC multimer analyses (average of duplicates). Dotted line at y=3 (-Log10(0.001)) represent the selected threshold of FDR < 0.1%. The FAFQEYDSF specific response was confirmed by multiple (n>3) additional fluorescent-based analyses [21]. (**b**) Screening for T cell recognition of 417 predicted neoepitopes and six HLA-matched virus derived epitopes in samples of different origin from patient L013 using either combinatorial fluorescently labeled MHC multimers ($1\text{-}2x10^6$ viable cells per tube in 13 tubes) or DNA-barcoded MHC multimers ($4\text{-}8x10^6$ viable cells per tube in 1 tube). Samples were derived from in vitro expanded T cells from three resected tumor regions and from one normal lung region. PBMCs were only analyzed with DNA-barcoded MHC multimers ($2x10^6$ viable cells per tube). Results from all 423 pMHC multimers included in the screenings are plotted (x-axis). Y-axis data is plotted as in (a). Each L013-derived sample were analyzed once. Open symbols represents an epitope recognized in at least one of the remaining four samples with FDR<0.1.

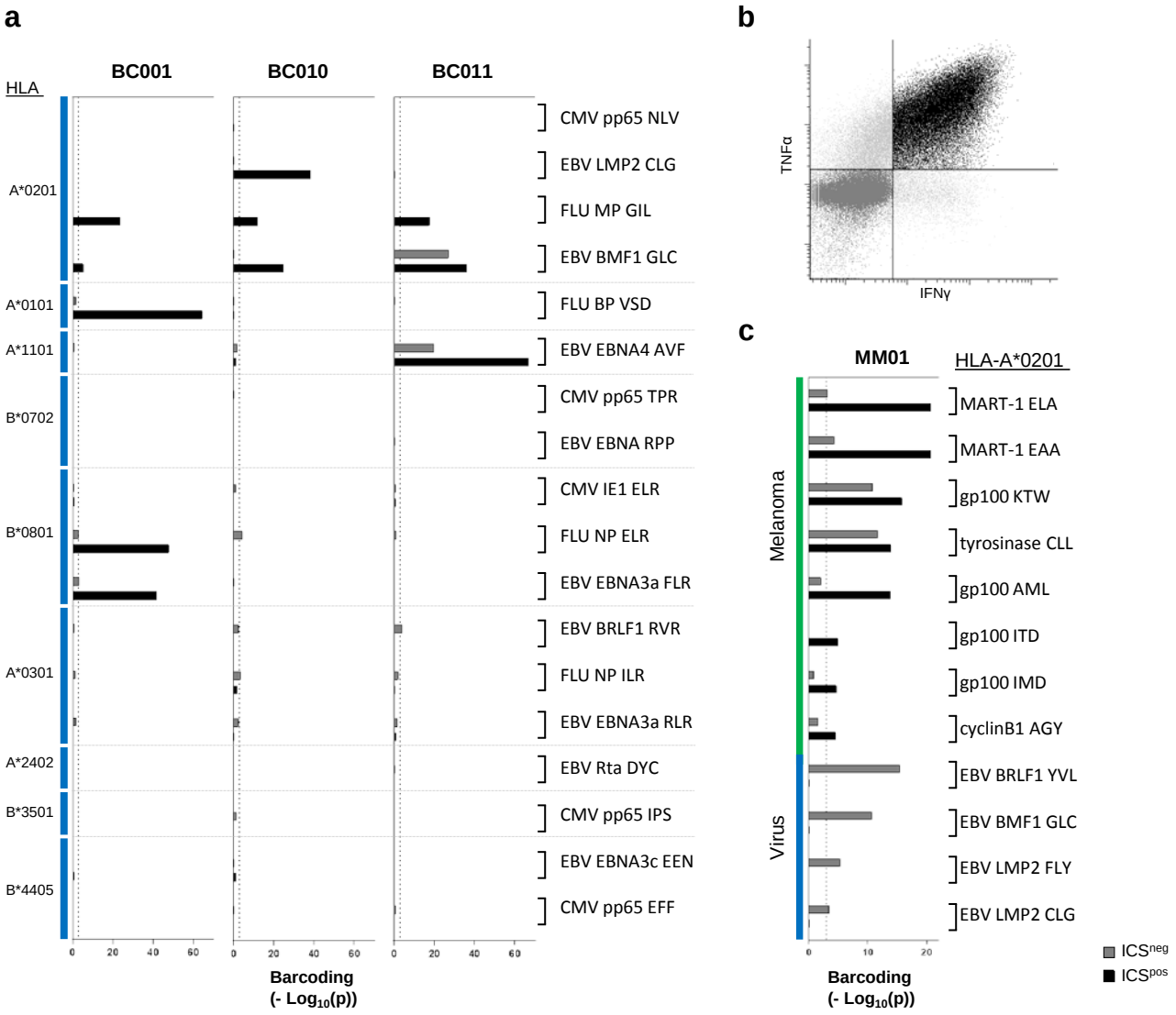


Figure 7. Functional assessment of pMHC-responsive T cells

(**a**) Bar plot representing the parallel assessment of T cell recognition and functional responsiveness of healthy donor PBMCs. PBMCs were stimulated with a CEF peptide pool, stained with INFγ and TNFα antibodies and with a library of 328 DNA-barcoded MHC multimers. CD8 T cells were isolated based on production of INFγ and TNFα (ICS^{pos}) versus no production of these cytokines (ICS^{neg}). MHC multimer binding was not included as isolation criteria. Peptides listed were present both in the CEF pool (Supplementary Methods) and in the pMHC multimer library (Supplementary Table 9). Each bar represents the -Log10(p) value in respect to the pMHC associated DNA barcode. Dotted line at y=3 (-Log10(0.001)) represent the threshold of FDR < 0.1%. Black bars represent T cell responses detected in the ICS^{pos} fraction and gray bars represent T cell responses detected in the ICS^{neg} fraction. T cell responses are grouped according the HLA restriction. (**b**) A representative dot plot used for FACS-based isolation of ICS^{pos} (black) or ICS^{neg} (gray) cell subset (here MM01). (**c**) After 5 hours stimulation with an autologous tumor cell line TILs were stained with INFγ and TNFα antibodies and with a library of 328 DNA-barcoded MHC multimers (Supplementary Table 9). ICS^{pos} and ICS^{neg} TILs from patient MM01 were isolated based on (b). MHC multimer binding was not included as isolation criteria. Data plotted as in (a). Only T cell responses with FDR<0.1% are included. T cell responses are sorted according to antigen category. All samples were analyzed once.

EXPERIMENTAL PROCEDURES

Cell samples

Peripheral blood mononuclear cells (PBMC) from healthy donors were isolated from whole blood by density centrifugation on Lymphoprep (Axis-Shield PoC), and cryopreserved at -150 °C in fetal calf serum (FCS; Gibco) + 10% DMSO.

Material from metastatic melanoma patients: TILs were obtained from resected tumor lesions from individuals with AJCC stage IV melanoma. Tumor lesions were resected following surgical removal. Tumor fragments from one metastatic lesion (1–3 mm^3) were either enzymatically digested overnight and immediately cryopreserved (uncultured TILs, see below), or cultured individually in complete medium (RPMI with 10% human serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 1.25 µg/mL fungizone (Bristol-Myers Squibb) and 6000 U/ml IL-2) at 37 °C and 5% CO_2, allowing TILs to migrate into the medium. TILs were expanded to reach >50x10^6 total cells originated from approximately 48 individual fragments, which had expanded to confluent growth in 2-mL wells and eliminated adherent tumor cells (average of approximately 2x10^6 cells per well from each tumor fragment). TIL cultures were further expanded using a standard rapid expansion protocol (REP) as previously described[32]. Briefly, TILs were stimulated with 30 ng/mL anti-CD3 antibody (OKT-3; Ortho Biotech) and 6000 U/mL IL-2 in the presence of irradiated (40 Gy) allogenic feeder cells (healthy donor PBMCs) at a feeder:TIL ratio of 200:1. Initially, TILs were rapidly expanded in a 1:1 mix of complete medium and REP medium (AIM-V [Invitrogen] + 10% human serum, 1.25 µg/mL fungizone and 6000 U/mL IL-2), but after seven days complete medium and serum were removed stepwise from the culture by adding REP medium without serum to maintain cell densities around 1x10^6–2x10^6 cells/mL. TIL cultures and tumor cell lines were cryopreserved at -150 °C in human serum + 10% DMSO.

Uncultured TILs were established as described previously[33]. Briefly, tumor fragments freshly

dissected from the same metastatic lesion as the expanded TILs were resuspended in RPMI media containing 1 mg/mL of collagenase type IV (Sigma-Aldrich) and 0.0125 mg/mL dornase alpha (Pulmozyme, Roche). After overnight digestion at room temperature under gentle sample agitation, single cell suspensions were cryopreserved at -150 °C in human serum + 10% DMSO.

Blood samples from patients with melanoma were collected at different time points, i.e. before and after infusion of rapidly expanded TILs in a phase II trial[22]. Clinical outcome of patient MM01 and patient MM02 is reported elsewhere (the two patients are identified as M37 and M43, respectively)[22]. Tissue typing was conducted by HLA amplification and sequencing (IMGM Laboratories GmbH) or by PCR and flow cytometry (results listed in Supplementary Table 7). All procedures were approved by the Scientific Ethics Committee for the Capital Region of Denmark. Written informed consent was obtained according to the Declaration of Helsinki.

Material from NSCLC patients:

Samples for sequencing were obtained from patients (L011 and L013) diagnosed with non-small cell lung cancer (NSCLC) who underwent definitive surgical resection prior to receiving any form of adjuvant therapy, such as chemotherapy or radiotherapy. Informed consent allowing for genome sequencing had been obtained. All samples were collected from University College London Hospital, London (UCLHRTB 10/H1306/42). L011 were classified with CK7+/TTF1+ adenocarcinoma (LUAD), L013 with squamous cell carcinoma (LUSC). Patient material was processed as described in detail in McGranahan et al, 2016[25].

Tumor Processing: Up to five regions from a single tumor mass, separated by 1 cm intervals, and adjacent normal tissue were resected, and snap-frozen. Peripheral blood was collected at the time of surgery. Approximately 5x5x5 mm snap-frozen tumor tissue and 500 μL blood were used for genomic DNA extraction. Multi-region Whole-Exome Sequencing and variant calling for each tumor region and matched germ-line from patient L011 and L013was conducted as described

previously[34,35]. Patients were serotyped and simultaneously genotyped using Optitype[36], which produced concordant results. Patients the following HLA-A and B genotype: L011; HLA-A*2402, A*1101, B*3501 and B*4901, and L013; HLA-A*0201, A*3201, B*4402 and B*1501. For our T cell analyses we were able to cover 7 out of the 8 HLA A and B molecules (HLA-B*4901 was not included).

Isolation and in-vitro expansion of tumor-infiltrating lymphocytes (TILs) for L011 and L013: Directly after tumor resection the sample was divided into regions. Samples were digested enzymatically, mechanically dissociated and TILs were expanded using a REP protocol, all are described in detail in McGranahan et al, 2016[25].

Generation of DNA barcodes and dextran conjugation

Oligonucleotides containing distinct 25mer nucleotide sequences (from Xu et al. 2009) were purchased from DNA Technology (Denmark). Oligonucleotides modified with a 5' biotin tag (oligo A) were joined to unmodified, partially complementary oligonucleotides (oligo B) to generate >1000 unique double-stranded AxBy DNA barcodes (sequences are listed in Supplementary Table 1 and 2). Combinations of A and B oligos (one of each) were mixed with 5x Sequenase Reaction Buffer mix (PN 70702, Affymetrix) to final concentrations of 26 µM (Oligo A) and 52 µM (Oligo B), respectively; heated to 65 °C for 2 min; and allowed to anneal by cooling slowly to <35 °C over 15-30 min. The annealed oligo A's and B's were elongated to create double-stranded AxBy DNA barcodes by adding Sequenase polymerase (70775Y, Affymetrix), 20 µM DTT and 800 µM or 72 µM dNTPs, followed by incubation for 5-10 min at RT. Elongated AxBy barcodes were diluted in nuclease-free water + 0.1% Tween to 2.17 µM (with respect to the A oligo) and stored at -20 °C. 5' biotinylated AxBy DNA barcodes were attached to PE- and streptavidin-conjugated dextran (from Immudex). The amount of DNA barcode was titrated for each lot of dextran conjugate to

ensure that the MFI was unaffected or affected only minimally when staining antigen-specific T cells with DNA-barcoded multimers compared to non-barcoded multimers generated from the same dextran-conjugate lot. For the dextran-conjugates applied during this study (6-7 SA per dextran), the amount of attached DNA barcode was 4-6 times less than the amount that would lead to a complete loss of antigen-specific interactions between barcoded multimers and T cells, i.e., an estimated two DNA barcodes per dextran backbone. DNA barcodes were attached by mixing with dextran-conjugate, followed by incubation, 30 min at 4 °C. DNA barcode-assembled dextran-conjugates were stored for up to one week at 4 °C.

Peptides and MHC monomer production

Peptides were purchased from Pepscan (Pepscan Presto) and dissolved to 10 mM in DMSO. UV-sensitive ligands were synthesized as previously described[37–39]. Recombinant HLA-A0101, HLA-A0201, HLA-A0301, HLA-A1101, HLA-A2402, HLA-A3201, HLA-B0702, HLA-B0801, HLA-B1501, HLA-B2705, HLA-B3501, HLA-B4402 and HLA-B4405 heavy chains and human β2 microglobulin light chain were produced in Escherichia coli. HLA heavy and light chains were refolded with UV-sensitive ligands and purified as described in Hadrup et al., 2009[41]. Specific peptide-MHC complexes were generated by UV mediated peptide MHC exchange[17,38,39,42]. Nine of 1031 pMHC monomers were refolded immediately with antigenic peptide as described in Garboczi, Hung, & Wiley, 1992[43].

<u>Identification of Putative Neoantigens related to NSCLC patients, L011 and L013:</u>

Identified non-silent mutations were used to generate a list of 9-11 amino acid long peptides with the mutated amino acid represented in each possible position. The binding affinity of every mutant peptide and its corresponding wild-type peptide to the patient's germline HLA alleles was predicted

using netMHCpan-2.8[44,45]. Candidate neo-antigens were identified as those with a predicted binding strength of <500 nM. We predicted 288 potential HLA class I binding peptides for L011, with the relative distribution of 45 HLA-A*2402, 121 A*1101 and 122 B*3501 binding peptides. For L013 we predicted a total of 417 potential HLA class I binding neoepitopes, including 173 HLA-A*0201, 104 A*3201, 10 B*4402 and 130 B*1501 binding peptides. All predicted neoepitopes were synthesized. Peptide binding was not validated experimentally, as the described technology allow screening of the full library without further selection. Previous experimental validation of peptide-HLA binding has shown that approximately 50% of the predicted peptides will bind to HLA molecules with medium to high affinity (defined as >60% of a comparable virus-derived T cell epitope)[14,15]

Other peptide libraries included:

In the 1031 peptides library we included peptide libraries associated to Merkel Cell Carcinoma and Breast cancer. Both of these libraries were selected as previously described[14]; using a combination of SYFPEITHI[46] and netMHC [44,45]. Thresholds applied: > 19 in SYFPEITHI and < 1000 in netMHC. All predicted peptides were experimentally validated for MHC binding as described in Rodenko et al.[38]. Peptide libraries associated to Melanoma and RCC included previous published T cell epitopes, identified as described in[13].

Generation of DNA barcode-labeled peptide-MHC multimer libraries

Unoccupied SA-binding sites on the DNA barcode-assembled dextran conjugates were used for the co-attachment of biotinylated pMHC molecules. 2.3 nmol pMHC monomer was mixed with $160x10^{-15}$ mol DNA-barcoded dextran-conjugate and incubated 30 min at RT. MHC multimers were diluted in PBS with 5.2 µM D-biotin (Avidity, Bio200) to 50 µg/mL and incubated 20 min on ice. DNA-barcoded MHC multimers were either stored for up to 5 days at 4 °C (PBS supplemented

with NaN_3, final concentration 0.02% (w/v)) or for up to two months at -80 °C (PBS supplemented with glycerol and BSA, final concentrations 5% and 0.5%, respectively).

Immediately before staining barcode-labeled MHC multimers were centrifuged, 5 min at 3300 g, and pooled (2.3 nmol of each pMHC/sample) to enable the detection of multiple T cell responses in parallel. If necessary, the volume of the reagent pool was reduced up to 50x by ultrafiltration to obtain a final volume of approx. 80 µL of MHC multimers: centrifugal concentrators (Vivaspin 20, 100,000 Dalton, Sartorius) were saturated in PBS + 5 % BSA by centrifugation until >10 ml had passed through the membrane. Concentrators were washed twice, each time by centrifuging 20 ml PBS through the membrane. The pooled multimers were loaded onto the concentrator and centrifuged, 3300 g, 4 °C, to reach a final concentration of 23 µM for each pMHC in the staining reaction. Any aggregates in the MHC multimer reagent pool were sedimented by centrifugation, 5 min 3300 g before addition to the cell sample. An aliquot of approx. 5 µl of the MHC multimer reagent pool was stored at -20 °C for baseline analysis.

Generation of fluorescently labeled pMHC tetramers

MHC tetramers were assembled as previously described[4,12] with nine fluorescent SA conjugates (SA-PE, SA-APC, SA-PE-Cy7, SA-BV421, SA-BV510, SA-BV605, SA-BV650, SA-PE-CF594, SA-BUV395) (BioLegend, Nordic Biosite, Denmark) to allow two-dimensional staining with fluorescent labels of up to 36 T cell receptor specificities in parallel.

Peptide-MHC multimer staining

Cryopreserved PBMCs and TILs were thawed and washed in RPMI + 10% FCS. Cells were washed in a barcode-cytometry buffer (PBS + 0.5% BSA + 100ug/ml herring DNA + 2mM EDTA) and

incubated 30 min, 37 °C in the presence of 50 nM dasatinib. $2x10^5$-$3x10^6$ cells (except when specifically stated) were incubated, 15 min, 37 °C, with pooled DNA-barcoded multimers in a total volume of 100 ul (final concentration of each distinct pMHC, 23 µM). Next a 5x antibody mix composed of CD8-PerCP (Invitrogen MHCD0831) (final dilution 1/50), dump channel antibodies: CD4-FITC (BD 345768) (final dilution 1/80), CD14-FITC (BD 345784) (final dilution 1/32), CD19-FITC (BD 345776) (final dilution 1/16), CD40-FITC (Serotech MCA1590F) (final dilution 1/40), CD16-FITC (BD 335035) (final dilution 1/64) and a dead cell marker (LIVE/DEAD Fixable Near-IR; Invitrogen L10119) (final dilution 1/1000) was added and incubated 30 min, 4 °C. Cells were washed twice in barcode-cytometry buffer and if not acquired immediately, fixed in 1% paraformaldehyde. If the cells were not acquired within 24 h they were washed twice more and resuspended in barcode-cytometry buffer. Cells were acquired within a week after multimer staining.

Generation of spiked cell samples: Initially, i.e. just after thawing and washing PBMCs in RPMI + 2% FCS, cells were incubated with 50 nM dasatinib, 30 min, 37 °C and resuspended in RPMI + 10% FCS to 0.2 $x10^6$ cells ml^{-1}. Consecutive five-fold dilutions of PBMCs from one healthy donor into PBMCs from another healthy donor generated seven samples with: 100%, 20%, 4%, 0.8%, 0.16%, 0.032% and 0.0064% cells derived from the first donor PBMCs. Cells were washed in barcode-cytometry buffer and incubated with multimers and antibodies as described above.

Cells applied for tetramer staining were cryopreserved, thawed, washed and incubated with dasatinib in the same way as cells stained with barcode-labeled multimers. Tetramer stainings were performed as described previously[4,12].

Intracellular cytokine staining

Healthy donor PBMCs and TILs were thawed and washed in RPMI + 10% FCS. TILs were rested overnight in X-vivo + 5% human serum at 37 °C and 5% CO_2, washed and re-suspended in X-vivo + 5% human serum. $3x10^5$ TILs were stimulated with autologous tumor at a ratio 3:1 (TIL:tumor) and 1µL/mL GolgiPlug (BD, 555029) was added prior to incubation, 5h, 37 °C and 5% CO2. $1x10^6$ healthy donor PBMCs were stimulated with an extended CEF peptide pool (jpt, PM-CEF-E) at a concentration of 1 µg/ml of each peptide. After 2h incubation, 37 °C and 5% CO2, 1 µL/mL GolgiPlug (BD, 555029) was added and cells were incubated 4 more hours at 37 °C and 5% CO2. TILs and healthy donor PBMCs were washed twice in barcode-cytometry buffer and incubated, 15 min, 37 °C, with pooled DNA-barcoded multimers in a total volume of 100 µL (final concentration of each distinct pMHC, 23µM). Cells were stained with surface antibodies: FITC-conjugated anti-CD3 antibody (BD 345763) and PerCP-conjugated anti-CD8 antibody (Invitrogen MHCD0831) (final dilution 1/40 of each antibody) and dead cell marker (LIVE/DEAD Fixable Near-IR; Invitrogen L10119) (final dilution 1/1000) and incubated for 30 min, 4 °C. Cells were washed twice in barcode-cytometry buffer, incubated 30 min-O.N., 4 °C, in fixation buffer (1:4, eBioscience 00-5123-43, to diluent eBioscience 00-5223-56), washed and resuspended in permeabilization buffer (1:10 buffer to water, eBioscience 00-8333-56). Cells were stained with intracellular antibodies: PE-Cy7-conjugated anti-TNFα antibody (BioLegend 502930) and APC-conjugated anti-IFNγ antibody (BD 341117) (final dilution 1/10 of each antibody) and incubated 30 min, 4 °C. Cells were washed in permeabilization buffer and resuspended in barcode-cytometry buffer. If cells were not acquired immediately, they were resuspended in 1% PFA and washed twice more in barcode-cytometry buffer after 1-24h. Fixed cells were acquired within a week after staining.

Flow cytometry and cell sorting

Cells stained with DNA-barcoded multimers were sorted on a FACSAria (Aria, Aria-II or AriaFusion) (Becton Dickinson) into tubes containing 200 µL barcode-cytometry buffer (tubes were saturated with PBS + 2% BSA in advance). Using FACSDiva software we gated on single, live $CD8^+$ positive and 'dump' (CD4, 14, 16, 19 and 40) negative lymphocytes and sorted all multimer (PE) positive cells within this population. For the samples stained with antibodies for intracellular activation-markers and DNA-barcoded-multimers, we gated on single, live $CD8^+/CD3^+$ lymphocytes. In all ICS experiments we sorted the IFN-γ/TNF-α double positive and the double negative population into separate tubes. The sorted cells were centrifuged 10 min, 5000 g, and the buffer was removed. The cell pellet was stored at -80 °C in a minimal amount of residual buffer (<20 µL).

Tetramer stained cells were acquired on a LSR-II or a LSR-Fortessa flow cytometer (Becton Dickinson). Antigen-specific T cells were identified as described in[4,12]. Briefly, we gated on single, live $CD8^+$ positive and 'dump' (CD4, 14, 16, 19 and 40) negative lymphocytes and selected cells positive in two tetramer channels and negative in the seven remaining tetramer channels.

DNA barcode amplification

Taq PCR Master Mix Kit (Qiagen, 201443) was used for amplification of DNA barcodes using 0.3 µM of appropriate forward and reverse primers comprising Ion Torrent PGM 5' and 3' adaptors (A-key and P1-key, respectively). PCR amplification was conducted on isolated cells (in <20 µl of buffer) or on the stored aliquot of the MHC multimer reagent pool (diluted 50,000x in the final PCR) used as baseline to determine the number of DNA barcode reads within a non-processed MHC multimer reagent library. Each sample was assigned a distinct sampleID embedded in the

forward primer. PCR was performed under the following conditions: 95 °C 10 min; 36 cycles: 95 °C 30 s, 60 °C 45 s, 72 °C 30 s; and 72 °C 4 min. The PCR products were purified with QIAquick PCR Purification kit (Qiagen, 28104) or QIAquick gel extraction kit (Qiagen, 28704). The amplified DNA barcodes were sequenced at Sequetech (U.S.A.) or GeneDx (U.S.A.) using Ion Torrent PGM 314 or 316 chip (Life Technologies).

Processing of sequencing data

Raw sequence reads were aligned to the constant primer regions and the 16 nucleotide AxBy annealing region using Bowtie2[47]. Reads successfully mapped to at least two of these three features were aligned to the primer regions, the 25mer barcode regions, and the annealing region using the Smith-Waterman algorithm. Based on these alignment positions, the sampleID and the UMI (Fig 1B) were located and extracted. Sequence reads that could not be unambiguously assigned to a DNA barcode were discarded. To avoid multiple counting of duplicate sequence reads generated by amplification, the two random UMI regions were concatenated and duplicate counts for each AxBy barcode combination were removed. The resulting $\boldsymbol{n} \times \boldsymbol{m}$ matrix of clonally unique sequence counts, where n is the number of samples and m is the number of AxBy barcodes, was used in all further analyses.

Statistical analyses

The analysis of barcode enrichment was based on methods designed for the analysis of RNA-seq data and was implemented in the R package edgeR[48]. Fold changes in read counts mapped to a given sample relative to mean read counts mapped to triplicate baseline samples were estimated using normalization factors determined by the trimmed mean of M-values method[49]. P values were

calculated by comparing each experiment individually to the mean baseline sample reads using a negative binomial distribution with a fixed dispersion parameter set to 0.1; this value was derived from the background distribution in the 10 healthy donor PBMC samples shown in Figure 3. False discovery rates (FDRs) were estimated using the Benjamini-Hochberg method. Specific barcodes with an FDR<0.1% were defined as significant. At least 1/1000 reads associated with a given DNA barcode relative to the total number of DNA barcode reads in that given sample was set as threshold avoid false positive detection of T cell populations due to low number of reads in the baseline samples.

QPCR-based recovery of enriched DNA barcodes

QPCR-based analyses were used to recover enriched barcodes when the number of distinct barcodes in a given experiment where ≤2. The Brilliant II QRT-PCR Low ROX Master Mix Kit (Agilent 600837) was applied with 0.3 µM 5' and 3' primers (forward primers: GAAGTTCCAGCCAGCGTC or GAGATACGTTGACCTCGTTG and reverse primers: CTGTGACTATGTGAGGCTTTC or ATGCAACCAAGAGCTTAAGT) and 0.1 µM sequence specific fluorescent reporter probes (probeB1: 6-FAM/GCCTGTAGTCCCACGCGATCTAACA/BHQ and probeB2: HEX/CAACCATTGATTGGGGACAACTGGG/BHQ or probe-ss1: 6-FAM/TCT[+T][+G][+A]AC[+T][+A]TG[+A][+A][+T]CGTC/BHQ-1-plus and probe-ss2: HEX/TCT[+A][+T][+A]GG[+T][+G]TC[+T][+A][+C]TACC/BHQ-1-plus) ([+X]= locked nucleic acid). Template was drawn from residual buffer containing the sorted cells (≤20 µL per sample). A 40-cycle qPCR (Mx3000P, Agilent) was performed with the following cycling conditions: 95 °C 10 min; 95 °C 30 s, 60 °C 60 s × 40 cycles.

Note on DNA-barcodes

For the initial experiments presented in figure S1C we used single-stranded DNA barcodes (ss-barcodes) that resembled the A-oligonucleotides, but where the annealing region at the 3'end on the A-oligonucleotides were replaced with a reverse primer region as follows: ss-barcode1: Biotin-GAGATACGTTGACCTCGTTGAANNNNNNTCTTGAACTATGAATCGTCTCACTTAAGCTCTTGGTTGCAT and ss-barcode2: Biotin-GAGATACGTTGACCTCGTTGAANNNNNNTCTATAGGTGTCTACTACCTCACTTAAGCTCTTGGTTGCAT (blue, green and black coloring indicates primer regions, UMI and the 25mer barcode region, respectively).

Supplementary References

32. Donia, M. *et al.* Characterization and comparison of 'standard' and 'young' tumour-infiltrating lymphocytes for adoptive cell therapy at a danish translational research institution. *Scand. J. Immunol.* **75,** 157–167 (2012).

33. Donia, M. *et al.* Aberrant Expression of MHC Class II in Melanoma Attracts Inflammatory Tumor-Specific CD4+ T- Cells, Which Dampen CD8+ T-cell Antitumor Reactivity. *Cancer Res.* **75,** 3747–59 (2015).

34. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366,** 883–92 (2012).

35. Gerlinger, M. *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat. Genet.* **46,** 225–33 (2014).

36. Szolek, A. *et al.* OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* **30,** 1–7 (2014).

37. Bakker, A. H. *et al.* Conditional MHC class I ligands and peptide exchange technology for the human MHC gene products HLA-A1, -A3, -A11, and -B7. *Proc.Natl.Acad.Sci.U.S.A* **105,** 3825–3830 (2008).

38. Rodenko, B. *et al.* Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nat.Protoc.* **1,** 1120–1132 (2006).

39. Toebes, M. *et al.* Design and use of conditional MHC class I ligands. *Nat.Med.* **12,** 246–251 (2006).

40. Hadrup, S. R. *et al.* High-throughput T-cell epitope discovery through MHC peptide exchange. *Methods Mol. Biol.* **524,** 383–405 (2009).

42. Chang, C. X. L. *et al.* Conditional ligands for Asian HLA variants facilitate the definition of CD8+ T-cell responses in acute and chronic viral diseases. *Eur. J. Immunol.* **43,** 1109–20 (2013).

43. Garboczi, D. N., Hung, D. T. & Wiley, D. C. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in Escherichia coli and complexed with single antigenic peptides. *Proc.Natl.Acad.Sci.U.S.A* **89,** 3429–3433 (1992).

44. Hoof, I. *et al.* NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* **61,** 1–13 (2009).

45. Nielsen, M. *et al.* NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS.One.* **2,** e796 (2007).

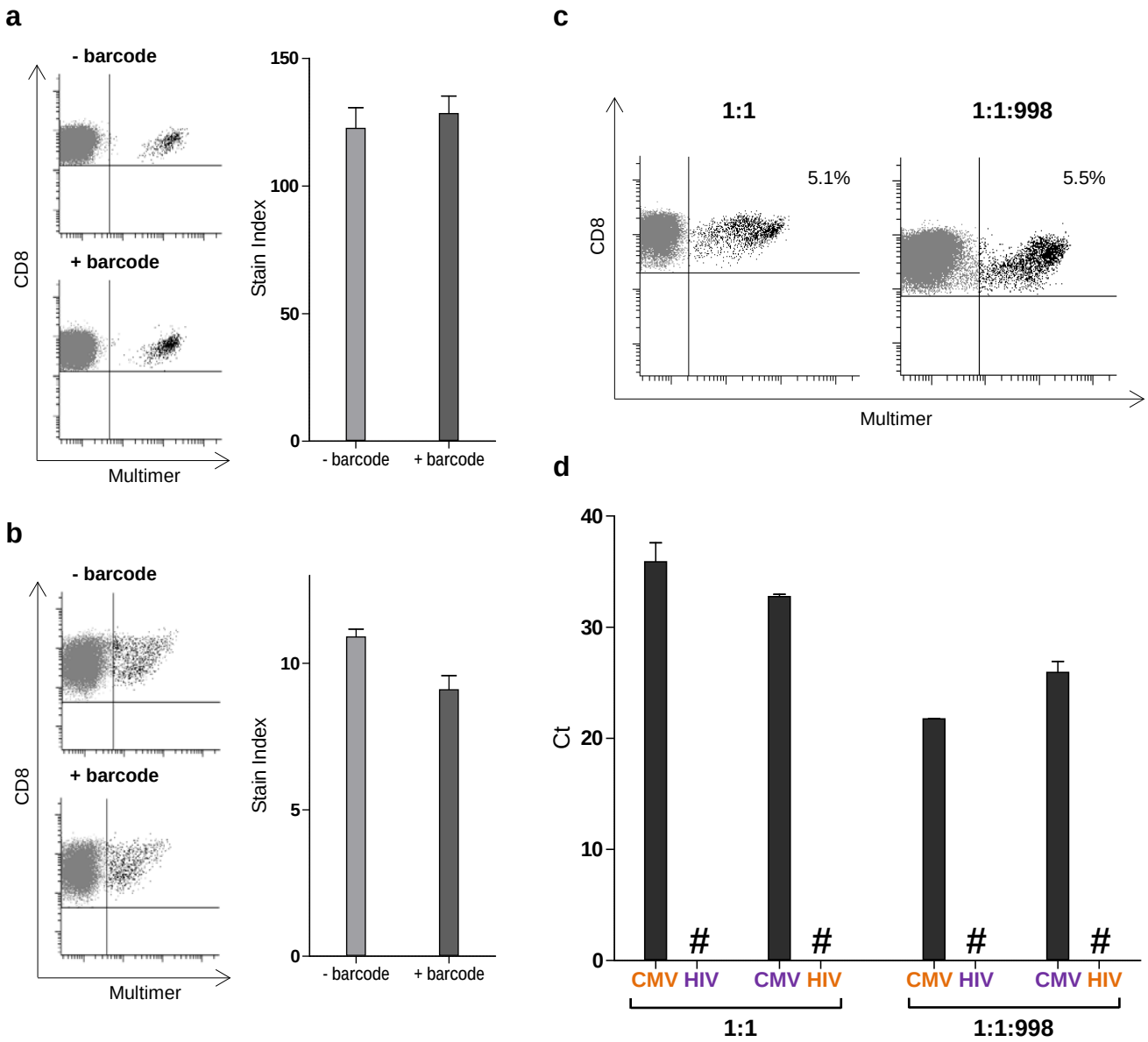
46. Rammensee, H. G., Bachmann, J., Emmerich, N. P., Bachor, O. A. & Stevanovic, S.

SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50,** 213–219 (1999).

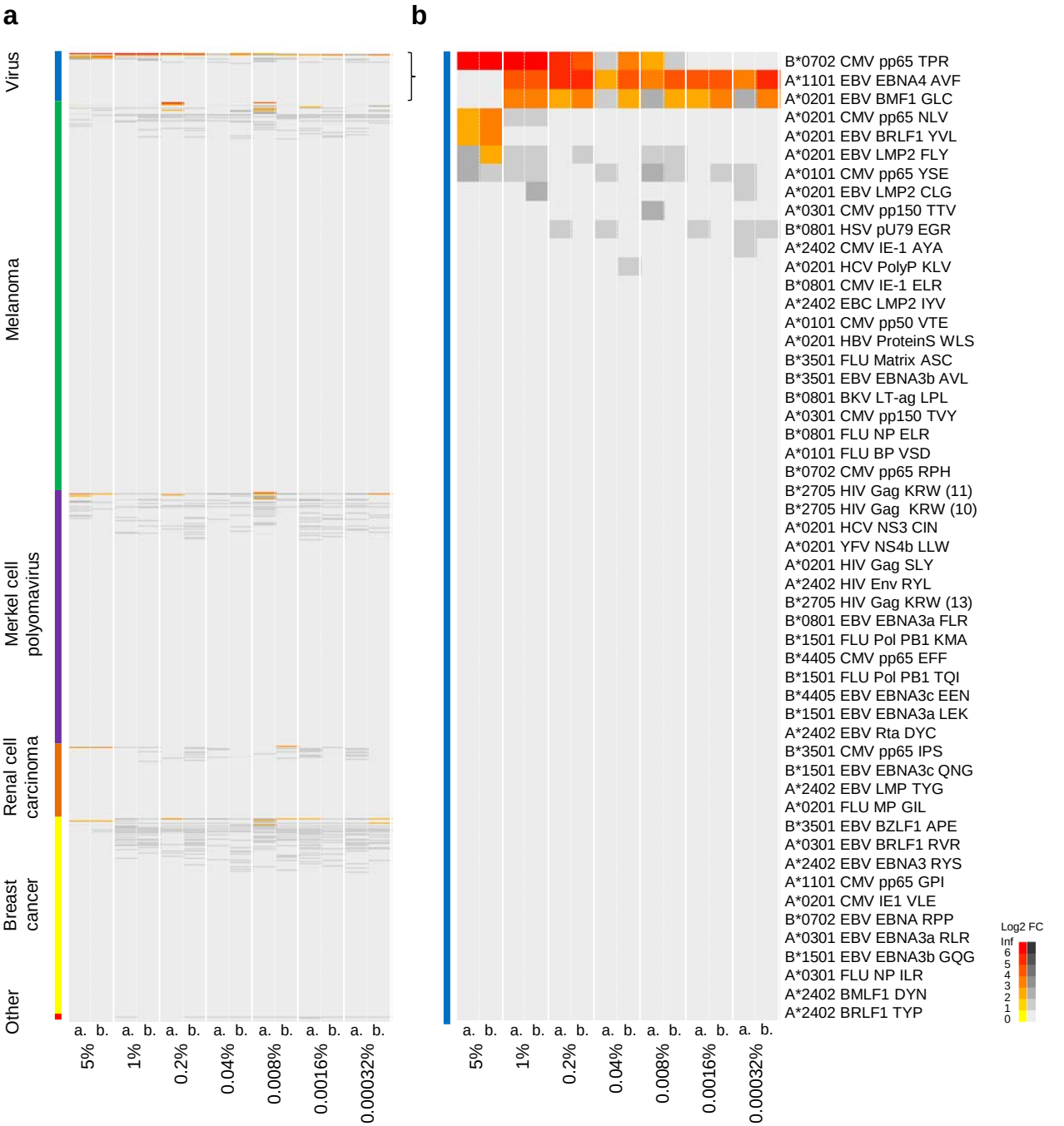
47. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–9 (2012).

48. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–40 (2010).

49. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11,** R25 (2010).



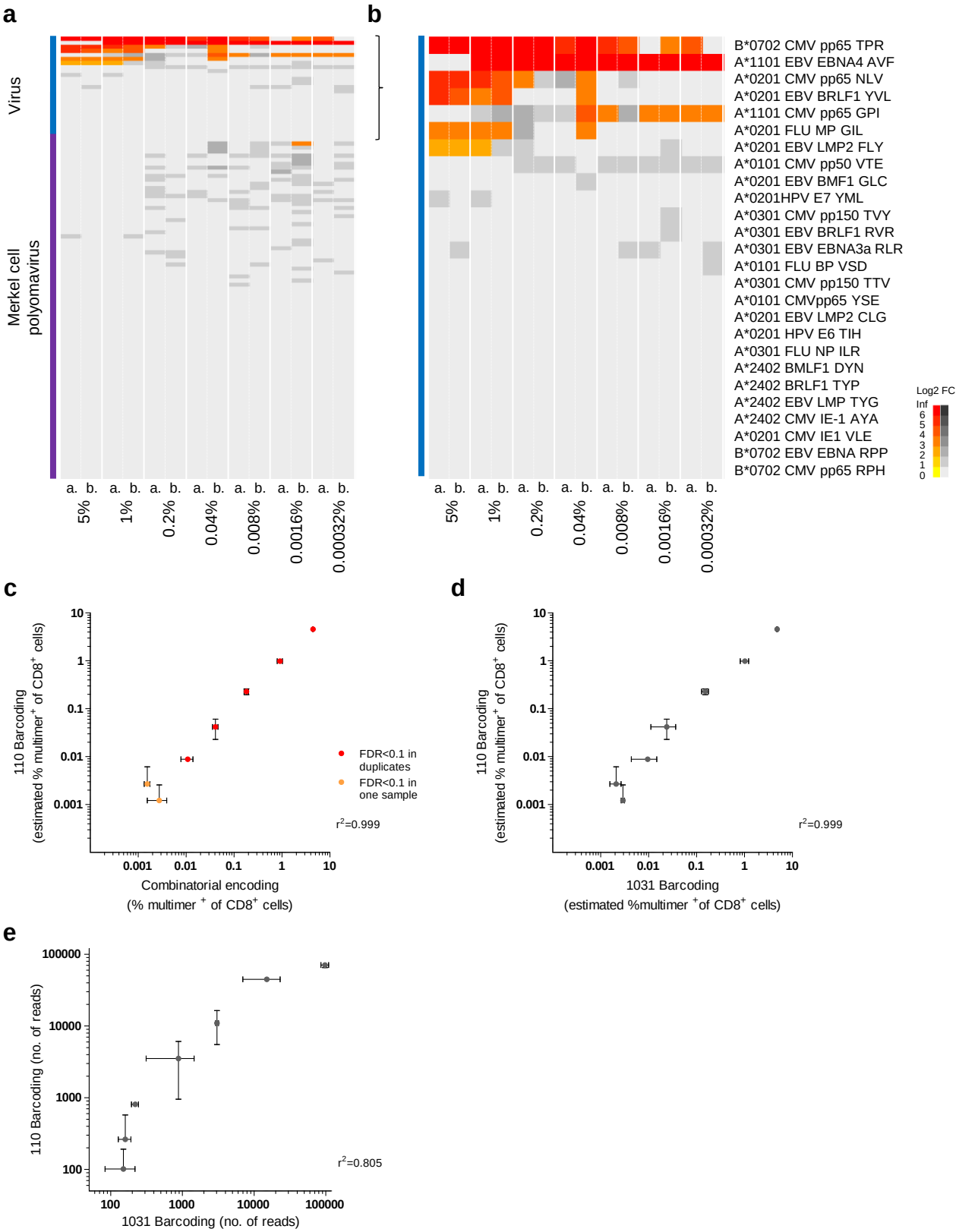
Supplementary Figure 1. Binding capacity of DNA-barcoded MHC multimers and recovery of antigen specificity. (**a, b**) Fluorescent-based determination of the binding capacity of DNA-barcoded MHC multimers (+barcode) compared to fluorescent labeled MHC multimers (-barcode). DNA-barcoded MHC multimer reagents were assembled with two identical DNA barcodes attached to each multimerization backbone and subsequent co-attachment of pMHC molecules. Non-barcoded multimers were generated similarly without prior attachment of DNA barcodes (i.e. '-barcode' and '+barcode' were both assembled on a dextramer backbone). Reagents assembled with HLA-A*0201 CMV pp65 $_{NLV}$ were applied for staining (**a**) of healthy donor PBMCs and reagents assembled with HLA-A*0201 hTERT p988 (YLQVNSLQTV) were applied for staining of (**b**) expanded TILs from a melanoma patient. The binding capacity is evaluated in terms of the stain index of the multimer fluorescent intensity of T cells stained with non-barcoded or DNA-barcoded MHC multimers respectively, along with the frequencies of the given multimer positive cell population of CD8 T cells, (**a**) 0.9-1.2% and (**b**) 3.8%-4.3%. Bar plots show mean stain index values of three stainings ± SD. (**c**) Fluorescent-based analysis of antigen specific T cells stained with relevant virus pMHC multimers and excess of irrelevant pMHC multimers. An equimolar (1:1) mixture of individually barcode labeled HLA-B*0702 CMV pp65$_{TPR}$-multimers and HLA-A*0201 HIV Pol$_{ILK}$-multimers, or a mixture with further 998 additional of fluorescent labeled pMHC multimers, were used for staining of healthy donor PBMCs. The 998 additional MHC multimers comprised equal amounts of irrelevant-peptide HLA-A*0201 and HLA-B*0702 multimers, i.e. multimers carrying MHC molecules refolded with UV-sensitive ligand (Supplementary Methods). Using either reagent mixture the concentrations of each equimolar pMHC were 23 µM in the final staining volume, i.e. for staining with the 1:1:998 equimolar reagents, the volume of the MHC multimer pool were reduced 50x. Percentages of the multimer positive population of CD8 T cells are given in dot plots. (**d**) The multimer positive populations from (**c**) were sorted by FACS and DNA barcodes associated with the sorted cell population were subjected to qPCR with fluorescent reporter probes targeting each individual DNA barcode. The experiments were performed with reagent mixtures with DNA barcodes inverted between the CMV and HIV pMHC multimers (indicated in orange and purple respectively). Cross threshold (Ct) values of DNA barcodes recovered from qPCR with approximately 200 and 600 cells in separate reactions (derived from staining with 1:1 and 1:1:998 reagent mixtures respectively) are shown in bar plots (mean±range of duplicate samples). Hashtag indicate that a given barcode was not detected.



Supplementary Figure 2. Duplicate assessment of the dynamic range and detection limit using a panel of >1000 DNA-barcoded pMHC multimers. Related to Figure 2

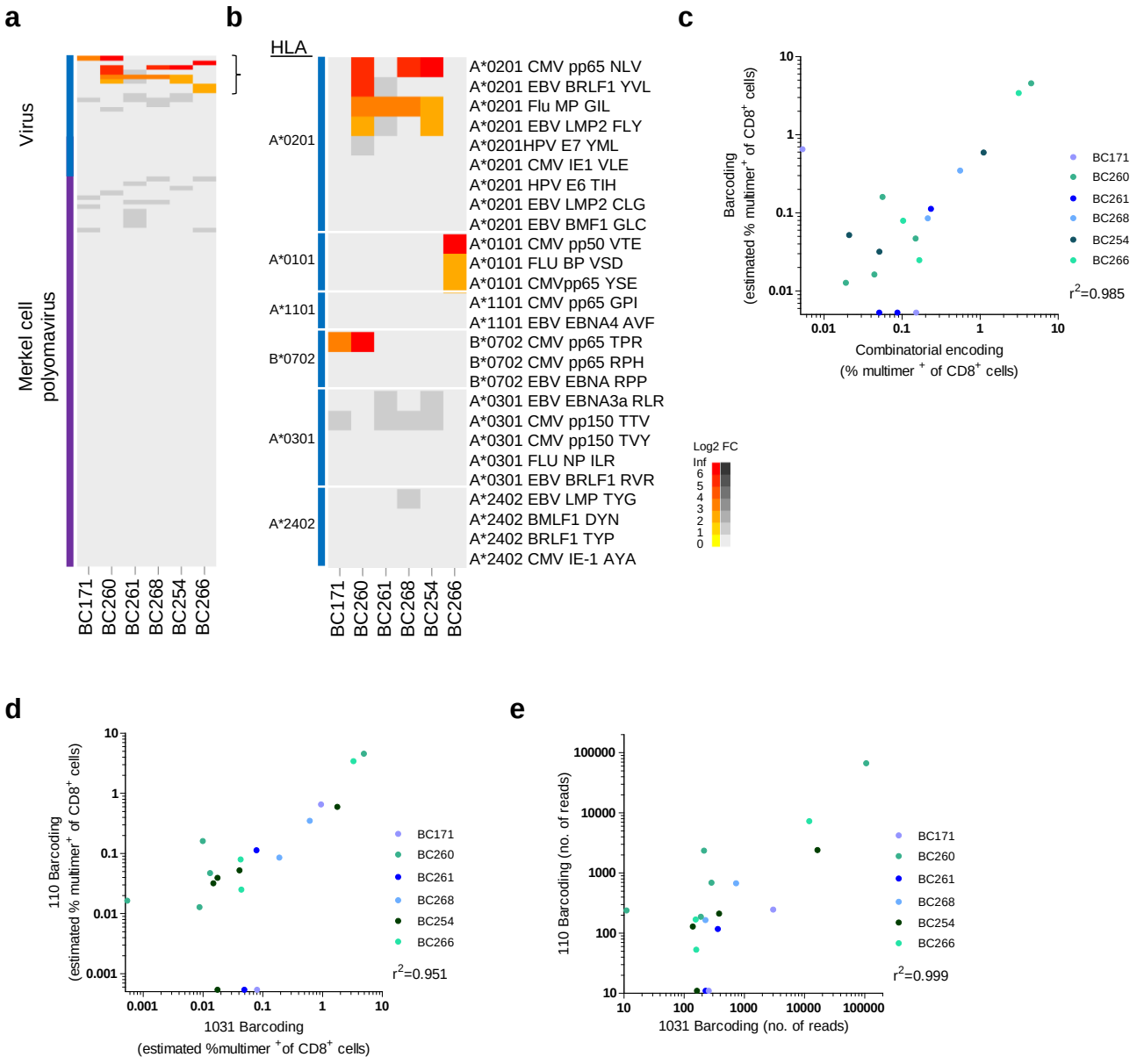
(**a**) Heatmap representing the pMHC multimer analysis of duplicates of seven samples with various proportions of HLA-B*0702 CMVTPR-specific T cells: 5%, 1%, 0.2%, 0.04%, 0.008%, 0.0016% and 0.00032% of CD8 T cells. The '5%' sample corresponds to 100% BC260 PBMCs. Each sample was screened with a panel of 1031 pMHC multimers, all carrying individual DNA barcodes. The heatmap shows changes in read proportions compared to background levels, as in Figure 2b. Each column represents the reads associated with a given sample. Epitopes are grouped based on their antigen origin. Within each antigen group, rows are sorted based on LogFC, highest to lowest, compared to baseline samples. Yellow-red coloring in the heatmap represents a statistically significant number of DNA barcode reads, FDR < 0.1%, defined as antigen-specific T cell responses. The gray scaling indicates non-significant number of barcode reads, i.e. any antigen-responsive T cells associated with such barcode would be present in numbers too low to discern from background. Duplicate samples are grouped side-by-side, indicated with a. and b. respectively. (**b**) Magnification of the panel of virus-derived peptides. Rows representing antigen specificities are sorted based on Log2FC and duplicate samples a grouped as in (a). The first row represents the target '5%' titrated specificity, B*0702 CMV_{TPR}. The rows below are two T cell responses present in the HLA-B*0702 negative PBMC sample (BC262), followed by at least two lower-frequent responses that are present in the '5%' donor (BC260). Dark gray scaling, i.e. barcode reads that that are non-significant but with Log2FC>1, may represent T cell responses just below the detection limit. All pMHC multimers are 'dextramers'.

.



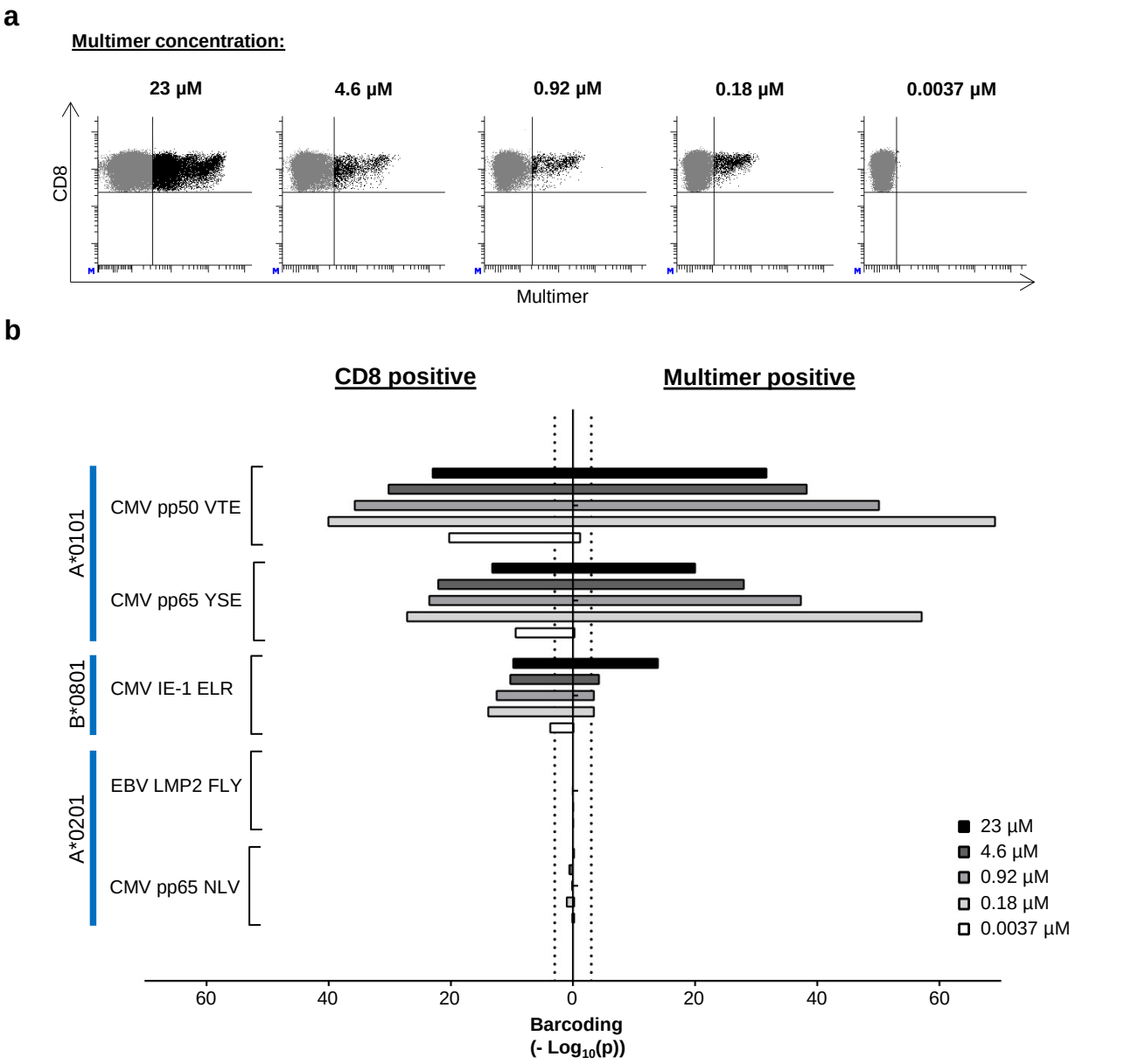
Supplementary Figure 3. Comparison of the dynamic range and detection limit using panels of 110 DNA-barcoded pMHC multimers or 1031 DNA-barcoded pMHC multimers. Related to Figure 2

(**a**) Heatmap representing the pMHC multimer analysis of duplicates of seven samples with various proportions of HLA-B*0702 CMV_{TPR}-specific T cells: 5%, 1%, 0.2%, 0.04%, 0.008%, 0.0016% and 0.00032% of CD8 T cells. The '5%' sample corresponds to 100% BC260 PBMCs. Each sample was screened with a panel of 110 pMHC multimers, all carrying individual DNA barcodes. The heatmap is organized as in Figure 2b, each column represents one donor. (**b**) Magnification of the panel of virus-derived peptides. Rows representing antigen specificities are sorted based on LogFC and duplicate samples a grouped as in (a). The first row represents the target '5%' titrated specificity, B*0702 CMV_{TPR}. The rows below are T cell responses present in the HLA-B*0702 negative PBMC sample (BC262) or lower-frequent responses present in the "5%" donor (BC260). Dark gray scaling, i.e. barcode reads that that are non-significant but with LogFC>1, may represent T cell responses just below the detection limit. **(c)** Correlations between the frequency of HLA-B*0702 CMV_{TPR}-specific T cells determined by analyzing the same samples using combinatorial fluorescent labeled pMHC multimers or a panel of 110 DNA barcode labeled pMHC multimers. (**d, e**) Correlation between results obtained from screening the same samples of $2x10^6$ PBMCs with the panel of 1031 or 110 DNA-barcoded pMHC multimers represented as (**d**) the estimated frequency of HLA-B*0702 CMV_{TPR}-specific T cells or (**e**) the number of clonally reduced barcode reads associated with this pMHC multimer. Error bars represent range of duplicates. The accumulated number of non-reduced read counts that mapped to any of the DNA barcodes among the 14 samples screened were $2.7x10^6$ and $1.4\ x10^6$ reads for the 110 and 1031 pMHC multimer library respectively. All pMHC multimers are 'dextramers'.



Supplementary Figure 4. Assessment of T cell reactivity across healthy donor samples using a panel of 110 DNA-barcoded pMHC multimers. Related to Figure 3

(**a**) Analysis of PBMCs (1-2x10^6) from six different healthy donors using 110 pMHC multimers each carrying individual barcodes. The heatmap is organized as in figure 2B, each column represents one donor. Epitopes are grouped based on their antigen origin. Significant responses are shown in yellow-red colors. Significance was defined as FDR<5% since the number of reads within a given sample were compared with only one baseline sample (opposed to three baseline samples in other experiments). (**b**) Magnification of the panel of virus-derived peptides (26 epitopes). Rows representing antigen specificities are grouped according to HLA-type and sorted within each group based on Log2FC. HLA types of donors can be seen in table S8. (**c**) Correlations between the frequencies of antigen-specific T cells determined by analyzing the same samples with combinatorial encoded fluorescently labeled pMHC multimers or with110 DNA-barcoded pMHC multimers. Each dot represents one specificity. T cell populations with FDR<5% in DNA-barcode MHC multimer analysis or ≥10 events and >0.002% of CD8 T cells in combinatorial encoding analysis were plotted. All specificities included in the plot were tested using both a combinatorial encoding analysis and DNA-barcoded MHC multimers. Dots plotted on the axes are nonsignificant for one of the methods. (**d, e**) Correlation between results obtained from screening the same samples with the panel of 1031 or 110 DNA-barcoded pMHC multimers represented as (**d**) the estimated frequency of antigen-specific T cells or (**e**) the number of clonally reduced barcode reads associated with the given pMHC multimers. Each dot represents one specificity. Only T cell populations that fulfilled the significance criteria for DNA barcode assessment (FDR<0.1% and 5% for the 1031 and 110 library, respectively) in at least one of the analyses were plotted. Dots plotted on the axes are nonsignificant for one of the library screenings. The accumulated number of non-reduced read counts that mapped to any of the DNA barcodes among the six screened samples were 2.6x10^5 and 4.3x10^5 reads for the 110 and 1031 pMHC multimer library respectively.



Supplementary Figure 5. T cell reactivity assessed independently from fluorescent-based separation of MHC multimer binding T cells. Related to Figure 4

(**a**) PBMCs ($2x10^6$) from one healthy donor was stained with varying amounts of DNA barcoded-pMHC multimers, i.e. a titration from 23 µM to 0.0037 µM in respect to each pMHC as indicated above each dot plot. This corresponds to 100%-0.016% of the amount used elsewhere in this study. Samples were stained in duplicates and either the full CD8 population (all cells in the dot plots) or only the MHC multimer positive population (cells indicated in black) were sorted. (**b**) Bar plot representing the distribution of DNA barcodes in the isolated cells from (a). Left side is based on the full CD8 population, right side is based of the MHC multimer positive population. Each bar represents the -Log10(p) value in respect to the pMHC associated DNA barcode. Dotted line at y=3 (-Log10(0.001)) represent the threshold of FDR < 0.1%. The donor BC035 is HLA-A*0101 and B*0801 positive and A*0201 negative (see full HLA-type in Supplementary Table 7). BC035 was previously found to carry antigen-responsive T cells restricted to HLA-A*0101, CMV $pp50_{VTE}$ (2.2% of CD8 T cells) and CMV $pp65_{YSE}$ (0.6% of CD8 T cells). HLA- A*0201 restricted epitopes functions as negative controls. Irrespectively of sorting the full CD8 or only the multimer positive population the same responses are detected after sequencing of the DNA barcodes. When the MHC multimer binding T cells could not be separated from the full CD8 T cell population based on there fluorescent intensity, i.e. when applying the lowest amount of MHC multimer reagent, the DNA barcodes associated with positive control reagents were still recovered after sorting the full CD8 population indicating that T cell reactivity can be assessed independent on fluorescent separation of the MHC multimer binding T cells.