nature biotechnology

ARTICLES

Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes

Amalie Kai Bentzen[1], Andrea Marion Marquard[2], Rikke Lyngaa[1], Sunil Kumar Saini[1], Sofie Ramskov[1], Marco Donia[3,4], Lina Such[1], Andrew J S Furness[5,6], Nicholas McGranahan[5,7], Rachel Rosenthal[5,7], Per thor Straten[3,8], Zoltan Szallasi[2], Inge Marie Svane[3,4], Charles Swanton[5,7], Sergio A Quezada[5,6], Søren Nyboe Jakobsen[1,9], Aron Charles Eklund[2] & Sine Reker Hadrup[1]

Identification of the peptides recognized by individual T cells is important for understanding and treating immune-related diseases. Current cytometry-based approaches are limited to the simultaneous screening of 10–100 distinct T-cell specificities in one sample. Here we use peptide–major histocompatibility complex (MHC) multimers labeled with individual DNA barcodes to screen >1,000 peptide specificities in a single sample, and detect low-frequency CD8 T cells specific for virus- or cancer-restricted antigens. When analyzing T-cell recognition of shared melanoma antigens before and after adoptive cell therapy in melanoma patients, we observe a greater number of melanoma-specific T-cell populations compared with cytometry-based approaches. Furthermore, we detect neoepitope-specific T cells in tumor-infiltrating lymphocytes and peripheral blood from patients with non-small cell lung cancer. Barcode-labeled pMHC multimers enable the combination of functional T-cell analysis with large-scale epitope recognition profiling for the characterization of T-cell recognition in various diseases, including in small clinical samples.

CD8 T cells are activated when their T-cell receptor (TCR) recognizes a peptide presented in the context of the cognate MHC-I molecule on an antigen-presenting cell. This interaction is a key element of adaptive immunity for defense against pathogens, for the recognition and elimination of cancer cells, and for the pathogenesis of autoimmune diseases. The ability to identify the peptide-MHC (pMHC) molecules recognized by CD8 T cells is crucial to understanding disease development and to developing specific therapeutic interventions. Characterization of T cells that recognize a given antigen can be assessed through functional analyses of antigen stimulation using cytokine-secretion assays, but these methods require large samples, depend on the secretory capacity of T cells, and are usually of low sensitivity[1].

Antigen-specific T cells have been detected using fluorescently labeled pMHC multimers, avoiding a dependence on secretory capacity and enabling parallel analyses of additional characteristics of T cells by flow cytometry[2]. 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[3,4]. Combinatorial encoding based on the application of a specific combination of tags to a given pMHC multimer has been employed to enhance the complexity of fluorescence-based analyses[5–7], and more recently of mass-cytometry-based analyses[8]. These advances enabled the parallel screening of up to 109 different pMHC-responsive T cells in single samples. Although the combinatorial encoding principle has enhanced our ability to describe immune interactions, the vast number of peptides potentially presented and the diversity of possible T-cell clones require a broader analytic range. An estimated 100,000–750,000 pMHC class I complexes are expressed for each allelic product (for human leukocyte antigen (HLA)-A and HLA-B loci)[9,10], and each individual carries ~10^7 different TCRs[11,12], each detecting up to 10^6 variations of a given peptide sequence[13]. Thus, a substantial increase in analytic range is required to comprehensively characterize T-cell recognition of cancer cells for personalized immunotherapeutic strategies, to search for T-cell targets involved in autoimmune disease development, and to identify TCR recognition profiles before clinical application. Here we use pMHC multimers tagged with unique DNA barcodes for parallel detection of >1,000 T-cell specificities in a single sample, and identify rare CD8 T-cell specificities in small clinical samples, including peripheral blood. By replacing fluorescence-based detection of pMHC-TCR interactions with DNA-barcoded reagents, which have the potential for ~10^{10} different markers, we remove the limitations in T-cell specificity analyses inherent to combinatorial encoding using fluorescent or metal tags.

RESULTS

Large-scale detection of T-cell specificities

We generated DNA barcode-labeled MHC-I multimer reagents on a dextran backbone that carried the common fluorescent label

[1]Section for Immunology and Vaccinology, National Veterinary Institute, Technical University of Denmark, Copenhagen, Denmark. [2]Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark. [3]Center for Cancer Immune Therapy, Department of Hematology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark. [4]Department of Oncology, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark. [5]CRUK Lung Cancer Center of Excellence, UCL Cancer Institute, London, UK. [6]Cancer Immunology Unit, UCL Cancer Institute, University College London, London, UK. [7]Translational Cancer Therapeutics Laboratory, The Francis Crick Institute, London, UK. [8]Department of Immunology and Microbiology, University of Copenhagen, Copenhagen, Denmark. [9]Immudex, Copenhagen, Denmark. Correspondence should be addressed to S.R.H. (sirha@vet.dtu.dk).

Received 11 March; accepted 4 August; published online 29 August 2016; doi:10.1038/nbt.3662

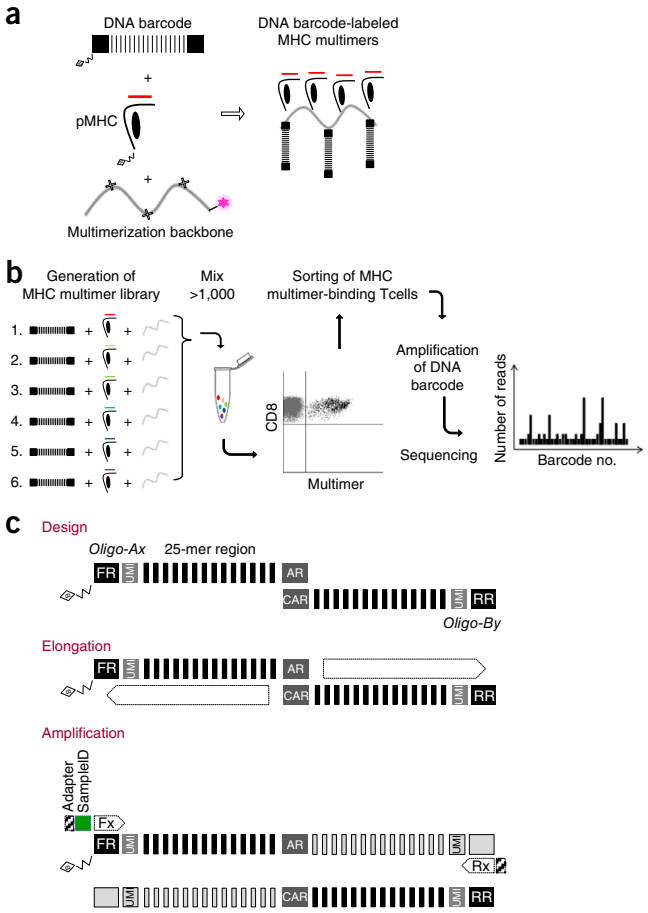

Figure 1 Preparation and use of DNA barcode-labeled MHC multimers. (**a**) Schematic overview showing the strategy for using DNA barcode-labeled MHC multimers for detection of antigen-specific T cells in complex cellular suspensions. Biotinylated DNA barcodes and pMHC molecules are attached to a PE-labeled dextran backbone carrying streptavidin; B, biotin. (**b**) Each MHC multimer is assembled with a given DNA barcode, forming a tag for the corresponding specificity (1 to >1,000). MHC multimer-binding T cells are sorted based on the PE label. DNA barcodes are amplified and sequenced, and the relative number *s* of DNA barcode reads is used to determine the composition of antigen-responsive T cells in the sample. (**c**) Schematic overview of the DNA barcode design. B, biotin; FR, forward region; UMI, unique molecular identifier; coding region, 25-mer barcode sequence assigning pMHC specificity; AR, annealing region; CAR, complementary annealing region; and RR, reverse region. The biotinylated oligo Ax comprises a 16-nucleotide region partially complementary to oligo By. Oligos Ax and By both contain an individual 25-mer oligonucleotide barcode sequence (determined by the 'x or y') and six randomly incorporated nucleotides, providing a UMI for each synthesized oligo. Oligo A contains a forward primer region (FR), and oligo B contains a reverse primer region (RR) (the oligonucleotide sequences are listed in **Supplementary Tables 1** and **2**). Following annealing of Oligo Ax and Oligo By and before 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. The forward primer is flanked by a 5′ sample identifier sequence (sampleID), and both the forward and reverse primers encode a 5′ sequencing adaptor sequence (IonTorrent, A-Key and P1-key, respectively).

phycoerythrin (PE) (**Fig. 1a**,**b**). The DNA barcodes were designed from a set of 240,000 unique 25-mer oligonucleotides with sequences described that have similar amplification properties while maintaining maximum diversity of their identification motifs[14]. To provide a system adaptable to large library screenings, we applied a combinatorial design of DNA barcodes (**Fig. 1c** and **Supplementary Tables 1–3**). Through this design we could generate at least 5.8×10^6 different barcodes. To determine the number of unique DNA sequences originating from each pMHC-associated barcode (clonally reduced barcode reads), we incorporated unique molecular identifiers (UMIs)[15] next to the barcode sequence (**Fig. 1c**), allowing the clonal reduction of sequences from the amplified product. MHC multimers were generated by addition of average two biotinylated DNA barcodes and 12 pMHC complexes per dextramer backbone. pMHC complexes were generated through UV-induced conditional ligand exchange to synthetic peptides of predicted relevance using methods described previously[16,17].

After sorting of MHC multimer-binding T cells using fluorescence-activated cell sorting (FACS) based on PE fluorescence intensity, the composition of the associated DNA barcodes can be retrieved by DNA amplification and high-throughput sequencing. Sequencing data were processed with the software package 'Barracoda' (http://www.cbs.dtu.dk/services/Barracoda, see Online Methods), to obtain the number of clonally reduced barcode reads assigned to a given sample and pMHC specificity (**Fig. 1b**). We used the number of clonally reduced reads for a given pMHC specificity to estimate the frequency of T cells specific for a given epitope, based on the average number of TCR–MHC multimer interactions detected in the total MHC multimer-binding T-cell pool in a given cell sample. This estimate was calculated using the following equation: [(number of DNA barcode reads associated with a specific pMHC)/(number of total barcode reads derived from the same sample)] × % MHC multimer-binding cells among total CD8 T cells (**Supplementary Table 4**).

Feasibility and detection limit for large-scale T-cell analysis

Using DNA barcodes to identify pMHC–TCR interactions enables, in principle, distinction of $>10^{10}$ different pMHC complexes in a single sample. However, several factors such as efficient and specific pMHC–TCR interactions, protein concentration and production, and peptide synthesis capabilities limit library sizes. To prove the feasibility of staining antigen-specific T cells in mixtures of >1,000 different pMHC multimers, we compared DNA-barcoded MHC multimers and combinatorial fluorescently labeled MHC multimers to detect T-cell populations responsive to virus or cancer-derived peptides[5,6]. We tested specific T-cell interactions using PE-labeled MHC multimers with and without an attached DNA barcode. The DNA barcode did not interfere with pMHC multimer binding, and MHC multimers with and without DNA barcodes showed identical staining capacities when staining healthy donor peripheral blood mononuclear cells (PBMCs) or tumor-infiltrating lymphocytes (TILs) from a melanoma patient (detection with HLA-A*0201 cytomegalovirus (CMV) pp65 $_{NLV}$ multimers or HLA-A*0201 hTERT p988 (YLQVNSLQTV) multimers, respectively) (**Supplementary Fig. 1a**,**b**). Moreover, specific DNA barcodes could be amplified after the isolation of HLA-B*0702 CMV pp65$_{TPR}$-multimer-binding T cells from healthy donor PBMCs, either following single MHC multimer staining or after mixture with 999 irrelevant MHC multimers. DNA barcodes associated with irrelevant MHC multimers were not identified (**Supplementary Fig. 1c**,**d**).

We generated a panel of 1,031 pMHC DNA-barcoded multimers with peptides from diverse antigens, including virus-derived, melanoma-associated[18], Merkel cell polyomavirus-derived[19], renal cell carcinoma-associated, and breast cancer-associated antigens (**Supplementary Table 5**). The pMHC-multimer library includes HLA-A*0101, HLA-A*0201, HLA-A*0301, HLA-A*1101, HLA-A*2402,

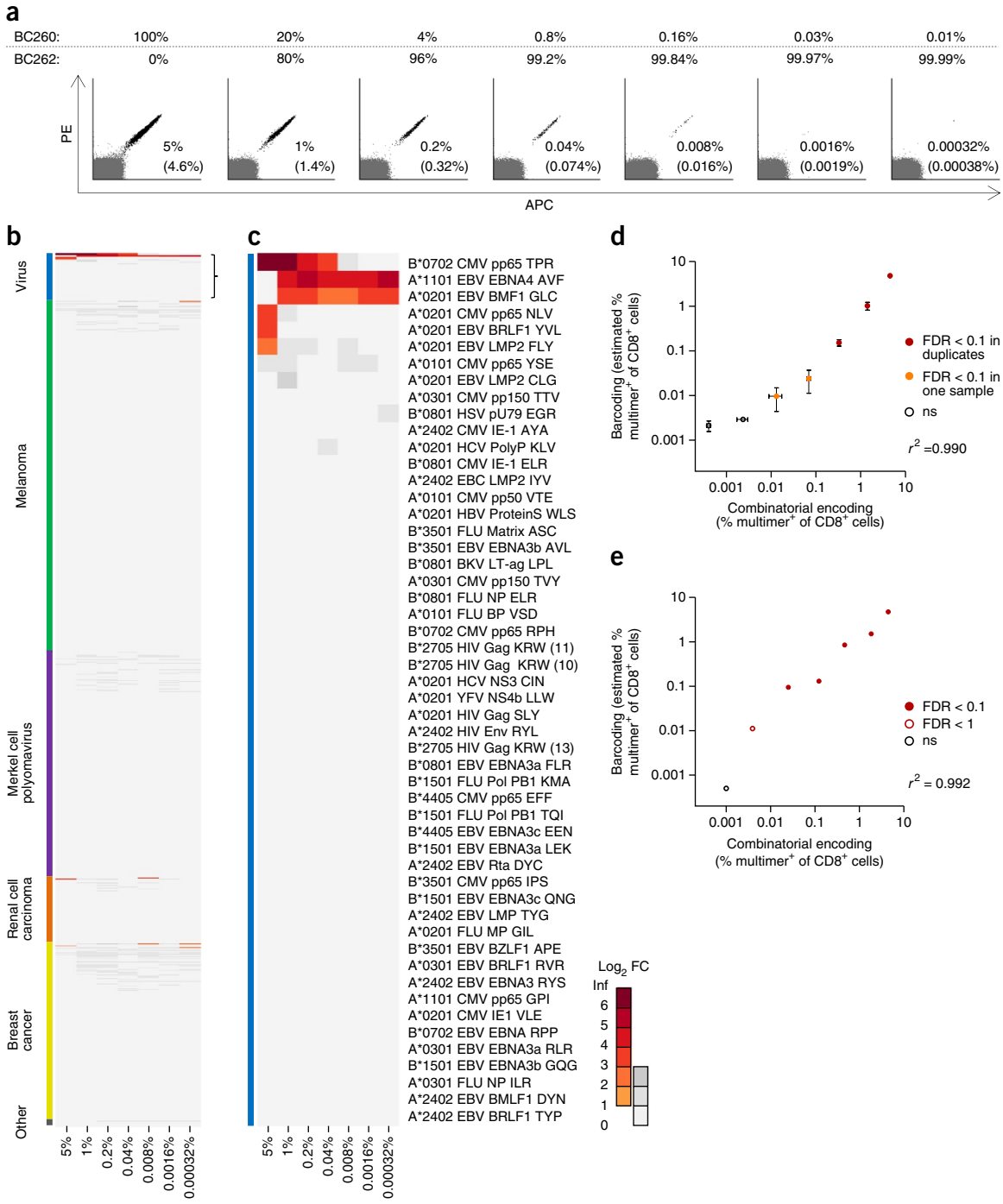

Figure 2 Dynamic range and limit of detection of DNA-barcoded MHC multimers. (**a**) Fluorescent multimer-based assessment of seven samples with various proportions of B*0702 CMV_{TPR}-specific T cells, theoretically: 5%, 1%, 0.2%, 0.04%, 0.008%, 0.0016% and 0.00032% of CD8 T cells. Samples are generated 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-positive 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 1,031 pMHC multimers. The heatmap shows changes in read proportions compared to background levels (log_2FC). 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 log_2FC, highest to lowest, compared to baseline samples. Orange-red or gray scaling, respectively, indicates a statistically significant (FDR < 0.1%) or insignificant (FDR ≥ 0.1%) number of reads. (**c**) Magnification focused on the virus-derived peptides. First row represents the target specificity B*0702 CMV_{TPR} present in BC260. Second and third rows below are T-cell responses from the HLA-B*0702 negative donor BC262, followed by three lower-frequency 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 1,031 DNA-barcoded pMHC multimers, when staining 2×10^6 PBMCs per sample (**d**) or 10×10^6 PBMCs per sample (**e**). Error bars represent the range of duplicates (s.d.). All pMHC multimers are 'dextramers'.

HLA-B*0702, HLA-B*0801, HLAB*1501, HLA-B*2705, HLA-B*3501, and HLA-B*4402 peptide-associated complexes. To investigate the linearity of detection and the detection limit of the method, we generated a titration curve with antigen-specific T cells. PBMCs from an HLA-A*0201, A*0301, B*0702, B*4001 donor (BC260) holding ~5% HLA-B*0702_CMV_pp65$_{TPR}$-specific T cells (as determined by combinatorial encoding of fluorescently labeled pMHC multimers) were mixed with varying amounts of cells from an HLA-B*0702-negative donor (BC262; **Fig. 2a** and **Supplementary Table 6**). We detected few T cell populations specific for peptides other than those derived from viruses (**Fig. 2b** and **Supplementary Fig. 2a**). Among the virus-derived specificities, we detected T cells specific for three epitopes in a sample of 100% PBMCs from the BC260 donor: 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 a mixed sample of 20% BC260 PBMCs and 80% BC262 PBMCs, we additionally 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) from the BC262 donor. Although the two low-frequency BC260-associated responses, HLA-A*0201 CMV$_{NLV}$ and EBV$_{YVL}$, were no longer identified, the HLA-B*0702 CMV$_{TPR}$ response remained detectable. Stepwise fivefold 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**).

We then compared T-cell detection by DNA barcode-labeled MHC multimers with flow cytometry–based detection to correlate the outcome of the two detection methods. The estimated frequencies of pMHC-responsive T cells correlated with the frequencies determined by combinatorial fluorescently labeled MHC multimers (**Fig. 2d,e**). When the HLA-B*0702 CMV$_{TPR}$-responsive T cells were assessed in an identical set of samples using a smaller library of DNA-barcoded MHC multimers (n = 110; **Supplementary Table 7**), we observed a comparable detection limit (0.001% of CD8 T cells) and correlation with detection by fluorescently labeled MHC multimers (r^2 = 0.999; **Supplementary Fig. 3a–d**). Moreover, results from the 1,031 DNA-barcoded MHC multimer library correlated with those from the 110 DNA-barcoded MHC multimer library, both in terms of 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) (**Supplementary Fig. 3d,e**). Improved identification of rare specific T cells using higher cell numbers (10×10^6 cells per sample), suggested that larger sample sizes may improve the limit of detection (**Fig. 2e**).

Detection of T-cell specificities across different donors

We used the 1,031-MHC multimer library to screen for T-cell reactivity in PBMCs from ten healthy donors with various HLA types (**Supplementary Table 6**). Virus-specific T cells were dominant (**Fig. 3a**) and, as expected, T-cell specificity was restricted by the donor's HLA haplotype (**Fig. 3b**): donor BC171 expressed HLA-A*1101, A*0301, B*0702, and B*1501, and the T-cell reactivities detected corresponded to these HLA molecules; samples from HLA-A*0201-positive individuals BC260, 259, 261, 262, 268, 254, and 251 showed variable T-cell reactivity toward HLA-A*0201-restricted virus-derived epitopes; donor BC268 expressed HLA-A*0101 and HLA-B*0801, and showed both CMV- and FLU-specific reactivity restricted to these HLA molecules. Among the ten healthy donors screened, we identified only one response not matching the HLA type of the given individual (**Fig. 3b;** HLA-B*0801, HSV pU79$_{EGR}$, from donor BC251).

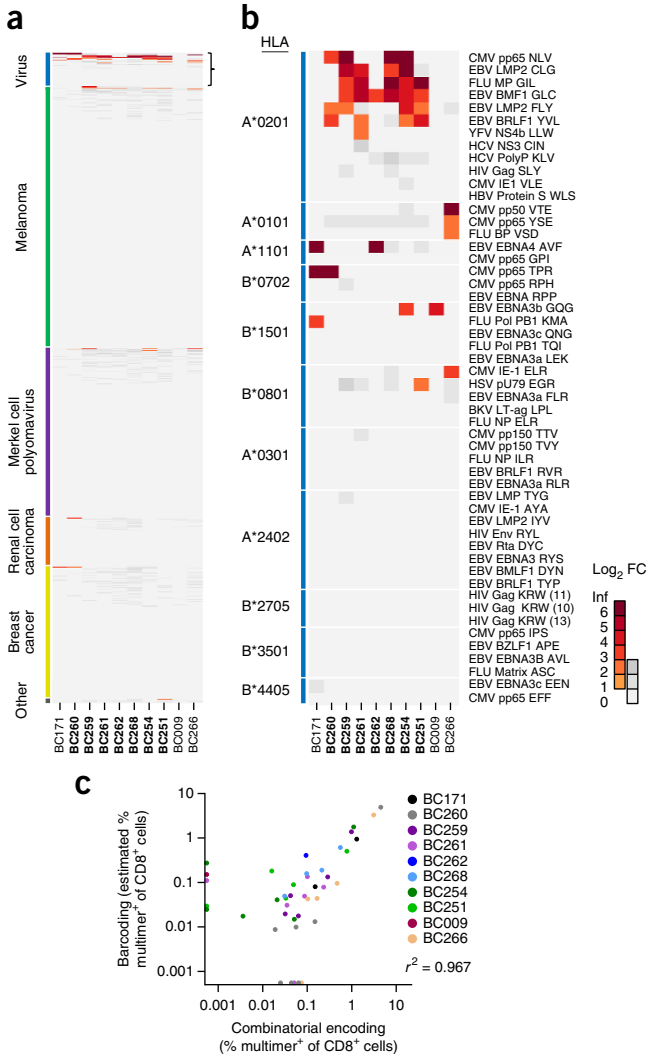

Figure 3 High-throughput assessment of T-cell reactivity using large peptide libraries. (**a**) Analysis of T-cell reactivity in ten different healthy donor PBMC samples (1–2 × 10^6 PBMCs) using 1,031 pMHC multimers, each carrying individual barcodes. The heatmap is organized as in **Figure 2b**, each column represents one donor. Donors marked in bold are all HLA-A*0201 positive. (**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 $\log_2$FC. Each sample shown here was analyzed once, but replicates with six of the same donors using another pMHC library are shown in **Supplementary Figure 4**. (**c**) Correlations between antigen-specific T-cell frequencies from analyses of T-cell responses in ten healthy donors using either combinatorial fluorescently labeled pMHC multimers (x axes) or 1,031 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.

We compared these results to fluorescently labeled pMHC multimers and found that the frequencies of antigen-responsive T cells determined using the two methods correlated well (r^2 = 0.967; **Fig. 3c**). Of the 42 potential responses identified, 88% were detected either by fluorescently labeled MHC multimers or by DNA-barcoded

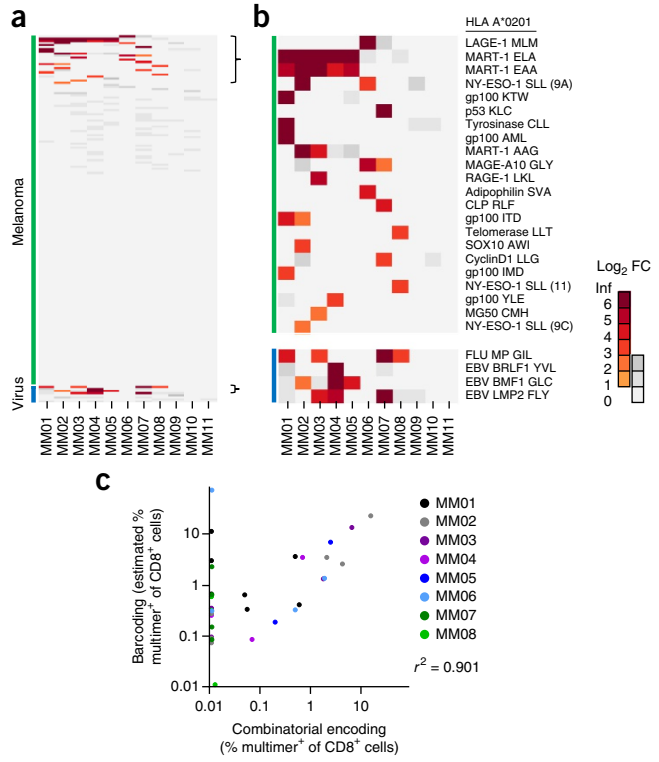

Figure 4 High-throughput assessment of tumor-reactive T cells. (**a**) Heatmap representing the barcode-based analysis of tumor-reactive T cells in 11 different samples of *in vitro*-expanded tumor-infiltrating lymphocytes (TILs) using a library of 167 HLA-A*0201 melanoma-associated pMHC multimers and 8 virus-derived pMHC multimers (175 pMHC library). The heatmap is organized as in **Figure 2b**. Patient MM10 and MM11 carries the HLA-A*0205 subtype. Each TIL culture was 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. 0.5–2 × 10^6 TILs were analyzed per sample in both methods.

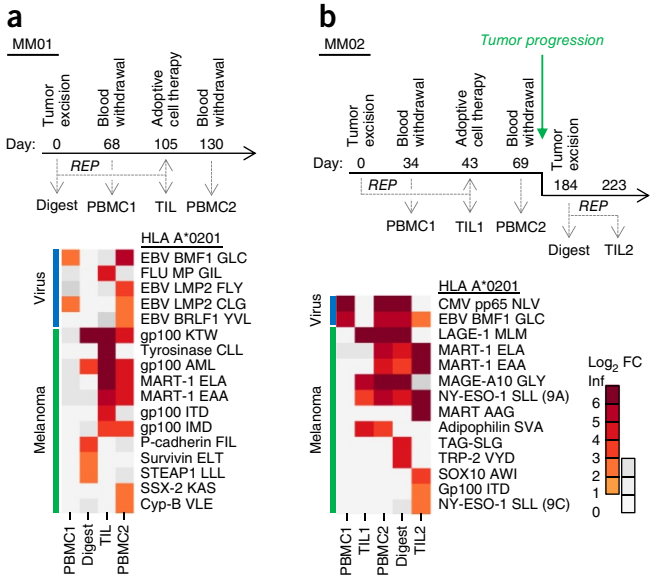

Figure 5 T-cell assessment in limited biological samples. (**a**,**b**) Analysis of dynamic changes in T-cell response to melanoma-associated antigens (175 pMHC library) before and after TIL adoptive cell transfer in two patients with metastatic melanoma. A timeline of sample collection and TIL adoptive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

MHC multimers. The five responses detected by flow cytometry but not with DNA-barcoded MHC multimers were 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 or 1/1,000 input reads in the DNA-barcode based analysis. The five T-cell specificities detected with DNA-barcoded MHC multimers and not with fluorescently labeled MHC multimers are likely to be specific given that they all match the HLA-expression and virus-response profiles of the respective donors.

When PBMCs from a subset of six of the same donors were analyzed using a library of 110 DNA-barcoded MHC multimers (**Supplementary Fig. 4** and **Supplementary Table 7**), we again observed strong correlations between T-cell reactivities detected using DNA-barcoded MHC multimers and those detected using combinatorial fluorescently labeled MHC multimers (r^2 = 0.985; **Supplementary Fig. 4c**), and between results from the 110- versus the 1,031-pMHC libraries (r^2 = 0.951) (**Supplementary Fig. 4d**,**e**). These results suggest that library size does not affect the capacity for detection of specific T cells using DNA-barcoded MHC-multimers considerably.

Detection of tumor-reactive T cells

To demonstrate the feasibility of using DNA-barcoded MHC multimers for the detection of tumor-restricted pMHC-specific T cells, we analyzed the T-cell reactivities in expanded TILs from 11 melanoma patients using an HLA-A*0201-restricted DNA-barcoded pMHC multimer library of 167 melanoma-associated peptides and 8 virus-derived peptides (**Supplementary Table 8**), described previously[18,20]. We detected numerous T-cell populations responsive to melanoma-associated antigens in 8 of the 11 TIL products (**Fig. 4a**,**b**). When the same library of melanoma-associated epitopes was used to analyze T cells in the healthy donor cohort, very few responses were detected (**Fig. 3a**). Notably, two patients that did not have T cells specific for HLA-A*0201-restricted peptides were found to be HLA-A*0205-positive upon detailed HLA typing (MM10 and MM11, **Supplementary Table 6**), supporting previous observations that HLA micropolymorphisms may affect T-cell detection by MHC multimers[21,22]. T-cell reactivity against several well-known melanoma-associated antigens, such as, MART-1, NY-ESO-1, gp100, and tyrosinase, was also detected. 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 (**Fig. 4b**), possibly reflecting partial cross-recognition of all three peptides from the same T-cell population, a phenomenon that has been described previously[23–26]. 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 frequency of antigen-responsive T cells estimated by 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 more melanoma-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**). The detection of an increased number of tumor-specific T-cell populations suggests that T-cell detection based on DNA barcode-labeled MHC multimers may be less dependent on stringent fluorescence-based separation (**Supplementary Fig. 5a**,**b**).

Detection of tumor-reactive T cells in small 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 the need for lymphocyte expansion. We used 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 2 trial of adoptive cell therapy using *in vitro* expanded TILs[27]. We screened for T-cell specificity against a library of virus- and shared melanoma-derived epitopes (328 barcoded pMHC multimers, **Supplementary Table 9**)[13] in enzymatically digested uncultured tumor fragments (i.e., unexpanded TILs), TILs expanded *in vitro* from small tumor fragments, dissected from the same metastatic lesion, and peripheral blood obtained before and ~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 several melanoma-associated T-cell specificities (**Fig. 5a**,**b**). In MM01, two of these specificities were also detected in the expanded TILs and in peripheral blood after transfer (**Fig. 5a**). The composition of T-cell specificities in peripheral blood after adoptive transfer was broadly similar to the composition of the TIL infusion product. In MM02, T-cell specificities detected in peripheral blood after TIL transfer reflected the composition of the TIL infusion product, but additional specificities, which matched those detected in a digested sample taken from a tumor lesion after tumor progression, were also detected (**Fig. 5b**). Overall, we detected an increased CD8 T-cell response to melanoma-associated antigens in peripheral blood after adoptive transfer of TILs that reflects T-cell populations detected in both tumor digest samples and expanded TIL products.

Detection of neoepitope-specific T cells

Mutation-derived neoepitopes have been suggested as important targets for tumor rejection mediated by T cells, and the predicted neoepitopes (based on the HLA profile of the patients in those studies) correlates positively with clinical outcome following treatment with immune checkpoint inhibitors[28,29]. Consequently, efficient tools for identification of neoepitope-restricted T cells are of potential prognostic and therapeutic value. We analyzed samples from two non-small cell lung carcinoma (NSCLC) patients for T-cell recognition of predicted HLA-binding peptide sequences containing cancer-specific mutations. For each patient, a large, personal neoepitope peptide library had been generated[25,30] ($n = 288$ and $n = 417$ for patient L011 and L013, respectively, as described in Online Methods). T-cell reactivity of *in vitro*–expanded TILs was assessed using both combinatorial fluorescently labeled MHC multimers (by parallel analysis of 30–36 different pMHC specificities per sample) and DNA-barcoded pMHC multimers. The analysis of the full neoepitope library using the first approach required splitting of the

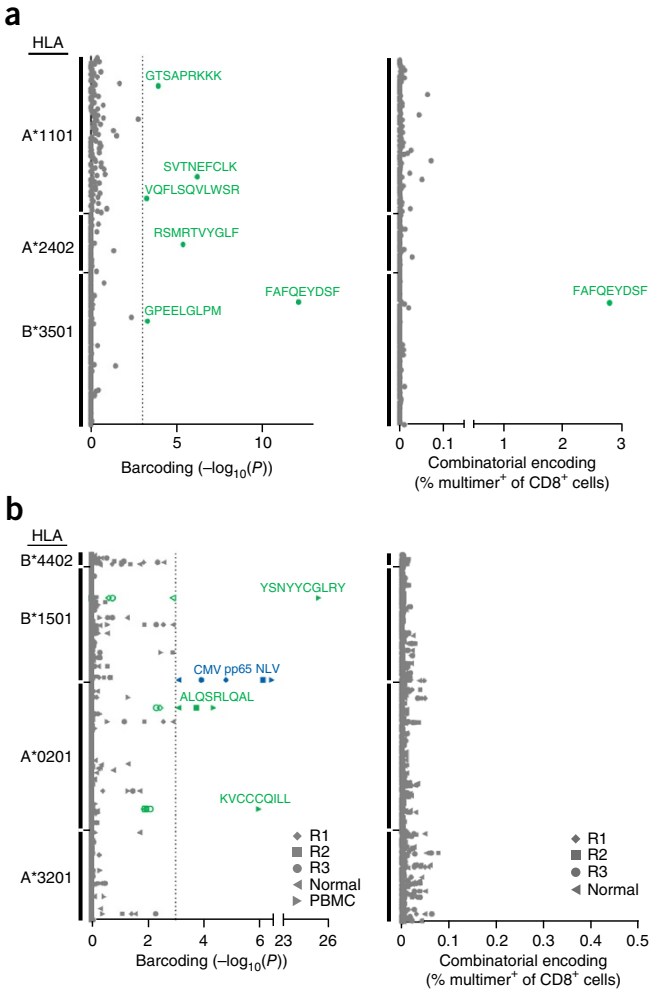

Figure 6 Detection of neoepitope responsive T cells in lung cancer. (**a**) Screening for T-cell recognition of 288 predicted neoepitopes and 10 HLA-matched virus-derived peptides in T cells expanded from a resected lesion in patient L011 using either DNA-barcoded MHC multimers (5×10^6 viable cells per tube in 1 tube) or combinatorial fluorescently labeled MHC multimers (1×10^6 viable cells per tube in nine tubes). Results from all 288 pMHC multimers included in the screening are plotted (*y* axis). Data plotted on the *x*-axis are the $-\log_{10}(P)$ (in respect to the pMHC-associated DNA barcode) for the DNA-barcoded MHC multimer analyses (average of duplicates), or percentage of MHC multimer positive T cells of total CD8 T cell for combinatorial fluorescently labeled MHC multimers. Dotted line at $x = 3$ ($-\log_{10}(0.001)$) 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 DNA-barcoded MHC multimers ($4–8 \times 10^6$ viable cells per tube in one tube) or combinatorial fluorescently labeled MHC multimers ($1–2 \times 10^6$ viable cells per tube in 13 tubes). Samples were derived from *in vitro*-expanded T cells from three resected tumor regions (R1–3) and from one normal lung region. PBMCs were only analyzed with DNA-barcoded MHC multimers (2×10^6 viable cells per tube). Results from all 423 pMHC multimers included in the screening are plotted (*y* axis). *x*-axis data are plotted as in **a**. Each L013-derived sample was analyzed once. Open symbols represent an epitope recognized in at least one of the remaining four samples with FDR < 0.1.

samples into 9 or 13 tubes for L011 and L013, respectively. Results from analyses using combinatorial fluorescently labeled MHC multimers were reported in detail previously[30], and resulted in detection of one neoepitope-responsive T-cell population in patient L011. Using DNA-barcoded MHC multimers, we observed nine neoepitope-specific T-cell population and one virus-specific T-cell population in these two patients (**Fig. 6a**,**b**). The use of DNA barcode-labeled MHC multimers enabled the screening of the whole library in one tube (4–8 × 10^6 TILs) rather than 9–13 parallel tubes (1 × 10^6 TILs per tube) for the fluorescent-based analysis. The larger number of cells in the initial binding reaction may account for the detection of a greater number of T-cell populations detected 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 specific for three neoepitopes were most prominently detected in peripheral blood (**Fig. 6b**), but could be detected, either as significantly (peptide 'ALQ', P = 0.004 (R1), 0.0002 (R2), 0.005 (R3), 0.0008 (Normal)) or as less enriched populations (peptide 'YSN', $P > 0.1$ (R1, R2, R3), P = 0.0012 (Normal); and peptide 'KVC', P = 0.01 (R1, R2, R3, Normal)), in T-cell cultures from the different tumor regions and normal lung tissue. CMV-specific T-cell populations were detected in all tumor regions, normal lung tissue, and in PBMCs. Detection of neoepitopes in peripheral blood was not feasible using combinatorial fluorescently labeled MHC multimers, due to the limited sample size, as this approach would require a minimum of 26 × 10^6 viable cells. Detection of such responses directly from peripheral

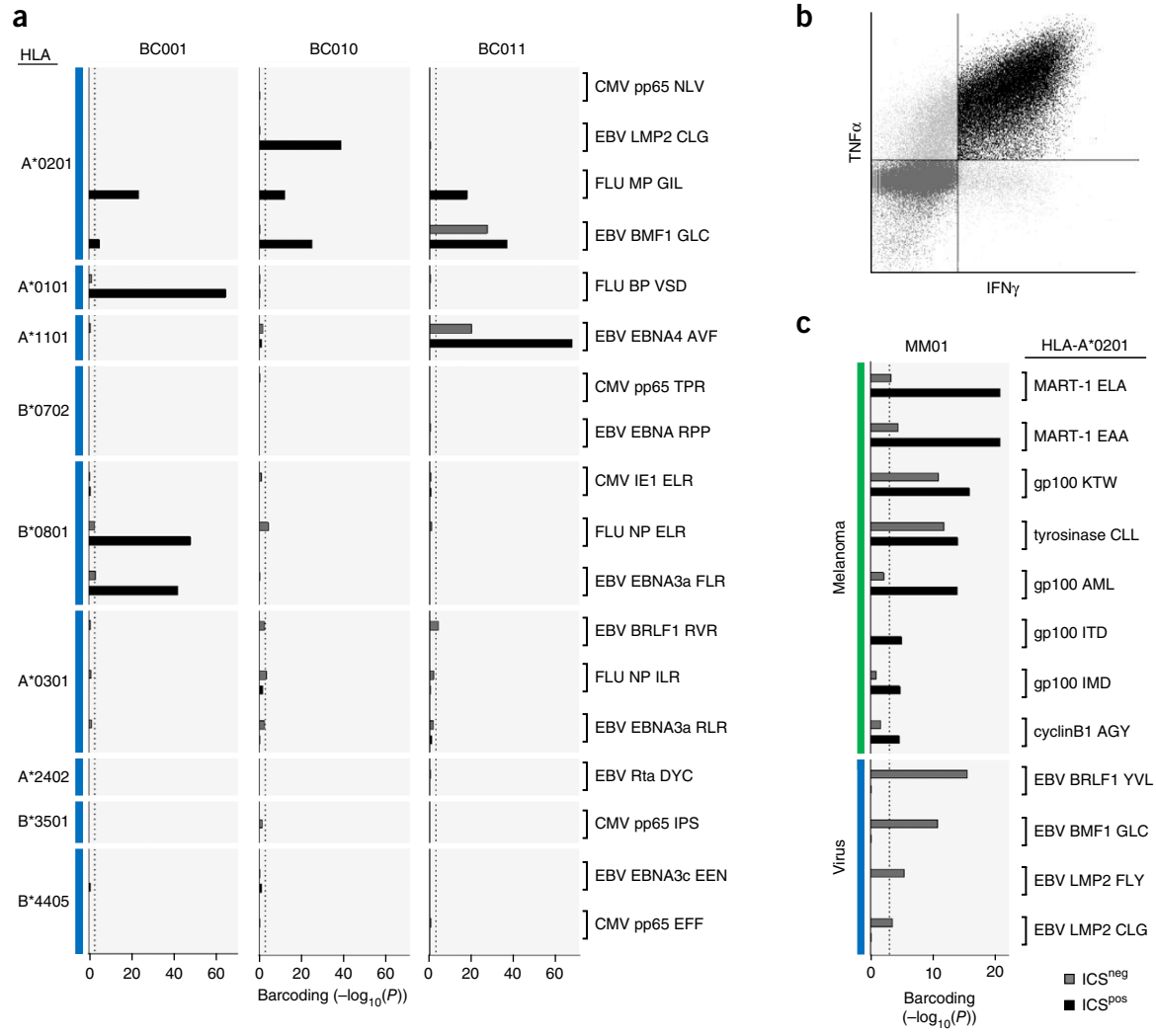

Figure 7 Functional assessment of pMHC-responsive T cells. (**a**) Bar plot representing the parallel assessment of T-cell specificity profiling and functional responsiveness of healthy donor PBMCs. PBMCs were stimulated with a CEF peptide pool, stained with IFNγ and TNFα antibodies and with a library of 328 DNA-barcoded MHC multimers. CD8 T cells were isolated based on production of IFNγ and TNFα (ICS^{pos}) versus no production of these cytokines (ICS^{neg}). MHC multimer binding was not included as an isolation criterion. Peptides listed were present both in the CEF pool and in the pMHC multimer library. Each bar represents the $-\log_{10}(P)$ value in respect to the pMHC associated DNA barcode. Dotted line at $y = 3$ ($-\log_{10}(0.001)$) represents 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 to 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 h stimulation with an autologous tumor cell line TILs were stained with IFNγ and TNFα antibodies and with a library of 328 DNA-barcoded MHC multimers. ICS^{pos} and ICS^{neg} TILs from patient MM01 were isolated based on **b**. MHC multimer binding was not included as an isolation criterion. 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.

blood has the potential to ease personalized T-cell therapy approaches and immunomonitoring.

Combining functionality and specificity profiling of T cells

We corroborated the functional capacity of T-cell specificities identified by DNA-barcoded pMHC multimers using cytokine production as a readout. In three healthy donor PBMC samples, we analyzed functional activity toward a pool of 32 different CMV, EBV, and FLU-derived peptides (CEF pool) by intracellular cytokine staining (ICS) and subsequently stained with a library of 328 DNA-barcoded pMHC multimers (**Supplementary Table 9** and Online Methods). We sorted T cells based on their cytokine secretion profile (IFNγ and TNFα secretion (ICS^{pos}) or no cytokine secretion (ICS^{neg})) and found that they matched the T-cell specificity assigned by pMHC multimer screening. Functional reactivity was observed for all virus-specific T cells present in these PBMCs (as detected in parallel by using barcoded pMHC multimers and sorting all multimer-positive T cells). Of 11 virus-specific T-cell populations, eight displayed full functionality, as they were only detected in ICS^{pos} fractions; the remaining three responses were present in both the ICS^{neg} and the ICS^{pos} pools, indicating partial functionality (A1101-EBV EBNA4 AVF and A0201-EBV BMF1 GLC in BC011; and B0801-EBV EBNA3a FLR in BC001) (**Fig. 7a**). Using the same principle, we analyzed the functional reactivity of TILs toward autologous tumor cells in a melanoma patient (**Fig. 7b**,**c**). We detected eight T-cell populations specific for melanoma-associated HLA-A*0201-restricted T-cell epitopes. These T-cell populations were either exclusively or predominantly detected in the ICS^{pos} population, indicating antigen processing and presentation by tumor cells and the functional capacity of epitope-specific TILs. As expected, the four virus-restricted T-cell populations present in this TIL product were detected only in the ICS^{neg} population. 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 enables multi-parallel screening of >1,000 pMHC molecules using DNA-barcoded pMHC multimers. The theoretical complexity of DNA barcode tags exceeds 10^{10} (ref. 14), and the complexity of pMHC-based T-cell analyses may consequently extend greatly beyond the 1,000 pMHC molecules used here. The limitations in scalability are based on the ability to achieve specific TCR-pMHC interactions within a very large pool of different pMHC multimers. The number (or concentration) of other MHC multimers that hinders specific TCR-pMHC multimer interactions might be >10,000–100,000 MHC multimers, but current libraries have not yet reached these numbers. Possible bottlenecks for increased library sizes are the automation of pMHC multimer production, high-throughput peptide synthesis, and microfluidics systems to ensure efficient pMHC interaction with all T cells in a given sample and specific interaction of the TCR with its cognate pMHC recognition motif (which can potentially be outmatched by CD8:MHC-constant-region unspecific interactions).

The detection of antigen-specific CD8 T cells using DNA-barcoded MHC multimers provides a detection limit comparable to that of fluorescence-based readouts and correlates tightly with the frequency of antigen-specific T cells found using combinatorial fluorescently labeled MHC multimers. We detected several T-cell populations within the same sample in the frequency range of 20–0.01% of CD8 T cells (**Figs. 2**–**4**), and showed that the limit of detection may be improved by increasing the number of T cells analyzed (**Figs. 2e** and **6a**,**b**). Experiments in our cohort of tumor-derived material in particular showed a greater number of T-cell populations detected with DNA barcode-labeled MHC multimers, suggesting that this method is less dependent on stringent fluorescence-based separation owing to sequencing of the pMHC multimer-associated DNA barcode to determine T-cell specificity instead of a fluorescent signal as a single parameter (**Supplementary Fig. 5a**,**b**). This may enable detection of low-avidity T cells, as the DNA barcode label is readily recovered from those sorted T cells that bind specifically to the MHC multimer even while obtaining only a marginally higher fluorescent intensity than the MHC multimer negative population. As tumor-reactive T cells, especially those recognizing shared self-peptide epitopes, often have lower affinity for the pMHC recognition motif compared to virus-specific T cells[31], an improvement in the level of resolution is a considerable advantage, even though our approach may not differentiate between high and low affinity T-cell specificities. Additionally, the dextran backbone used provides a higher-order multimer (~12) as opposed to the streptavidin-based tetramerization conventionally used for combinatorial fluorescently labeled MHC multimers, and may further assist the detection of low-avidity T cells[32,33]. The ability to assess T-cell antigen recognition at the relevant disease site is important for monitoring and potential intervention in many human disorders. Previous reports demonstrated loss of antigen recognition after *in vitro* expansion of T cells[18], which suggests that direct *ex vivo* assessments of T-cell reactivity will provide insights into the T-cell specificity at disease sites. For diagnostic purposes and for response-to-therapy monitoring, peripheral blood samples are preferable due to ease of access. However, the limited complexity and reduced screening breadth of MHC multimer-based analysis has 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 of patients with large libraries of pMHC multimers (**Fig. 6a**,**b**).

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 in parallel 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[34–36] to gather further information on relevant tumor epitopes. The analyses presented herein, based on bulk assessments of T-cell populations, largely reflect the information that can be obtained using fluorescently labeled pMHC multimers (**Figs. 2**–**7**), but represents a quantitative leap in the number of potential epitopes that can be assessed in parallel. Such large-scale analyses are particularly relevant when analyzing T-cell reactivity against tumor antigens based on personalized cancer genomic analyses. In general, this technology enables the assessment of T-cell reactivity at genome-wide levels for various immune-mediated diseases, and may enable T-cell specificity assessments that cover TCR variability.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We would like to thank U.K. Hansen, A. Burkal and T. Seremet for technical assistance; T. Schumacher, Netherlands Cancer Institute, for scientific discussions and sharing of MHC expression plasmids; and Dr. Altman, NIH Tetramer Core Facility, for sharing expression plasmids HLA-B*1501 and HLA-B*3501. The work was funded by The Danish Cancer Society (ID:R72-A4531-13-S2), The Lundbeck Foundation Fellowship (ID: R190-2014-4178), The Danish Research Council (FSS-ID: 1331-00283 and DFF-ID:4004-00422), Familien Erichsens Mindefond, Cancer Research UK (FC001169), the UK Medical Research Council (FC001169), the Wellcome Trust (FC001169), the UK Medical Research Council (MR/FC001169/1) and the Novo Nordisk Foundation (ID: 16584).

AUTHOR CONTRIBUTIONS

A.K.B. conceived the concept, designed and performed experiments, analyzed data, made figures and wrote the manuscript; A.M.M. designed the in silicon analysis, analyzed data, and made figures; R.L., S.K.S., and S.R. produced MHC monomers, designed and performed experiments, analyzed data and revised the manuscript; L.S. performed experiments, M.D. and I.M.S. provided patients material and generated tumor cell lines, discussed data; P.t.S. provided administrative support, flow facility and production of MHC monomers; A.J.S.F., S.A.Q. provided patients material and peptides from NSCLC; N.M.G., R.R., C.S. identified tumor mutagenome and predicted NSCLC-associated neoepitopes; Z.S. and A.C.E., designed the *in silico* analysis, and guided data analyses; S.N.J. designed DNA barcodes, conceived the concept, guided data analyses and revised the manuscript; S.R.H. conceived the concept, designed experiments, analyzed data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. McCutcheon, M. *et al.* A sensitive ELISPOT assay to detect low-frequency human T lymphocytes. *J. Immunol. Methods* **210**, 149–166 (1997).
2. Altman, J.D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996).
3. Chattopadhyay, P.K. & Roederer, M. Cytometry: today's technology and tomorrow's horizons. *Methods* **57**, 251–258 (2012).
4. Bendall, S.C. *et al.* Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **332**, 687–696 (2011).
5. 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).
6. 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).
7. 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).
8. Newell, E.W. *et al.* Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. *Nat. Biotechnol.* **31**, 623–629 (2013).
9. Rammensee, H.G., Falk, K. & Rötzschke, O. Peptides naturally presented by MHC class I molecules. *Annu. Rev. Immunol.* **11**, 213–244 (1993).
10. Stevanović, S. & Schild, H. Quantitative aspects of T cell activation--peptide generation and editing by MHC class I molecules. *Semin. Immunol.* **11**, 375–384 (1999).
11. Robins, H.S. *et al.* Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114**, 4099–4107 (2009).
12. Davis, M.M. & Bjorkman, P.J. T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395–402 (1988).
13. Wooldridge, L. *et al.* A single autoimmune T cell receptor recognizes more than a million different peptides. *J. Biol. Chem.* **287**, 1168–1177 (2012).
14. Xu, Q., Schlabach, M.R., Hannon, G.J. & Elledge, S.J. Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc. Natl. Acad. Sci. USA* **106**, 2289–2294 (2009).
15. Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* **9**, 72–74 (2011).
16. Toebes, M. *et al.* Design and use of conditional MHC class I ligands. *Nat. Med.* **12**, 246–251 (2006).
17. Rodenko, B. *et al.* Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nat. Protoc.* **1**, 1120–1132 (2006).
18. Andersen, R.S. *et al.* Dissection of T-cell antigen specificity in human melanoma. *Cancer Res.* **72**, 1642–1650 (2012).
19. 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).
20. Frøsig, T.M. *et al.* Broadening the repertoire of melanoma-associated T-cell epitopes. *Cancer Immunol. Immunother.* **64**, 609–620 (2015).
21. 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**, 641–648 (2014).
22. 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**, 967–975 (2015).
23. 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**, 1750–1758 (1998).
24. 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**, 7635–7645 (2007).
25. Wieckowski, S. *et al.* Fine structural variations of alphabetaTCRs selected by vaccination with natural versus altered self-antigen in melanoma patients. *J. Immunol.* **183**, 5397–5406 (2009).
26. Speiser, D.E. *et al.* Unmodified self antigen triggers human CD8 T cells with stronger tumor reactivity than altered antigen. *Proc. Natl. Acad. Sci. USA* **105**, 3849–3854 (2008).
27. Andersen, R. *et al.* Long-lasting complete responses in patients with metastatic melanoma after adoptive cell therapy with tumor-infiltrating lymphocytes and an attenuated IL2 regimen. *Clin. Cancer Res.* **22**, 3734–3745 (2016).
28. Snyder, A. *et al.* Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
29. Rizvi, N.A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).
30. McGranahan, N. *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469 (2016).
31. 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).
32. 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).
33. 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).
34. 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).
35. 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).
36. Klein, A.M. *et al.* Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* **161**, 1187–1201 (2015).

ONLINE METHODS

Ethical approval. All healthy donor and patient material was collected under approval by the local Scientific Ethics Committee, and written informed consent was obtained according to the Declaration of Helsinki. Related to healthy donor material and melanoma patient material, approval was obtained from the Scientific Ethics Committee of the Capital Region of Denmark. For NSCLC patients, informed consent allowing for genome sequencing had been obtained and samples were collected under approval from University College London Hospital, London (UCLHRTB 10/H1306/42).

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 FCS (FCS; Gibco) + 10% DMSO.

Material from metastatic melanoma patients. TILs were obtained from resected tumor lesions from individuals with stage IV melanoma (American Joint Committee on Cancer, AJCC). 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 6,000 U/mL IL-2) at 37 °C and 5% CO_2, allowing TILs to migrate into the medium. TILs were expanded to reach >50 × 10^6 total cells originated from ~48 individual fragments, which had expanded to confluent growth in 2-mL wells and eliminated adherent tumor cells (average of ~2 × 10^6 cells per well from each tumor fragment). TIL cultures were further expanded using a standard rapid expansion protocol (REP) as previously described[37]. Briefly, TILs were stimulated with 30 ng/mL anti-CD3 antibody (OKT-3; Ortho Biotech) and 6,000 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 6,000 U/mL IL-2), but after 7 d complete medium and serum were removed stepwise from the culture by adding REP medium without serum to maintain cell densities around 1 × 10^6–2 × 10^6 cells/mL. TIL cultures and tumor cell lines were cryopreserved at −150 °C in human serum + 10% DMSO. Tumor cell lines were tested for mycoplasma contamination, and visually inspected for morphological characteristics prior to cryopreservation.

Uncultured TILs were established as described previously[38]. 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 2 trial (ClinicalTrials.gov identifier: NCT00937625)[27]. Clinical outcome of patient MM01 and patient MM02 is reported elsewhere (the two patients are identified as M37 and M43, respectively)[27]. Tissue typing was conducted by HLA amplification and sequencing (IMGM Laboratories GmbH) or by PCR and flow cytometry (results listed in **Supplementary Table 6**).

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 before 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 (ref. 30). *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 5 × 5 × 5 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 L013 was conducted as described previously[39,40]. Patients were serotyped and simultaneously genotyped using Optitype[41], 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 seven out of the eight 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 rapid expansion protocol, all are described in detail in McGranahan *et al.*, 2016 (ref. 30).

Generation of DNA barcodes and dextran conjugation. Oligonucleotides containing distinct 25mer nucleotide sequences (from Xu *et al.* 2009)[14] 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 >1,000 unique double-stranded AxBy DNA barcodes (sequences are listed in **Supplementary Tables 1** and **2**). Combinations of A and B oligos (one of each) were mixed with 5× 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 As and Bs 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 mean fluorescence intensity 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 backbone), 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 1 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[16,17,42]. Recombinant HLA-A*0101, HLA-A*0201, HLA-A*0301, HLA-A*1101, HLA-A*2402, HLA-A*3201, HLA-B*0702, HLA-B*0801, HLA-B*1501, HLA-B*2705, HLA-B*3501, HLA-B*4402, and HLA-B*4405 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 (ref. 43). Specific peptide-MHC complexes were generated by UV-mediated peptide MHC exchange[16,17,22,44]. Nine of 1,031 pMHC monomers were refolded directly with the antigenic peptide as described in Garboczi, Hung, & Wiley, 1992 (ref. 45).

Identification of putative neoantigens related to NSCLC patients, L011 and L013. Identified non-silent mutations were used to generate a list of 9- to 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 (ref. 46). 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

doi:10.1038/nbt.3662

without further selection. Previous experimental validation of peptide-HLA binding has shown that ~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)[19,20].

Other peptide libraries included. In the 1,031 peptides library we included peptide libraries associated with Merkel cell carcinoma and breast cancer. Both of these libraries were selected as previously described[19], using a combination of SYFPEITHI[47] and netMHC[46]. Thresholds applied: > 19 in SYFPEITHI and < 1,000 in netMHC. All predicted peptides were experimentally validated for MHC binding as described in Rodenko *et al.*[17]. Peptide libraries associated with melanoma and RCC included previous published T-cell epitopes, identified as described in ref. 18.

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 pmol pMHC monomer was mixed with 160×10^{-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 909 nM and incubated 20 min on ice. DNA-barcoded MHC multimers were either stored for up to 5 d at 4 °C (PBS supplemented with NaN_3, final concentration 0.02% (w/v)) or for up to 2 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 3,300*g*, and pooled (2.3 pmol 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 50× by ultrafiltration to obtain a final volume of ~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, 3,300*g*, 4 °C, to reach a final concentration of 23 nM for each pMHC in the staining reaction. Any aggregates in the MHC multimer reagent pool were sedimented by centrifugation, 5 min at 3,300*g* before addition to the cell sample. An aliquot of ~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[5,6] 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 + 100 μg/mL herring DNA + 2 mM EDTA) and incubated 30 min, 37 °C in the presence of 50 nM dasatinib. 2×10^5–3×10^6 cells (except when specifically stated) were incubated, 15 min, 37 °C, with pooled DNA-barcoded multimers in a total volume of 100 μL (final concentration of each distinct pMHC, 23 nM). Next a 5× 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 (PFA). 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 0.2×10^6 cells mL^{-1}. Consecutive fivefold 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[5,6].

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 resuspended in X-vivo + 5% human serum. 3×10^5 TILs were stimulated with autologous tumor at a 3:1 ratio (TIL/tumor) and 1 μL/mL GolgiPlug (BD, 555029) was added before incubation, 5 h, 37 °C and 5% CO_2. 1×10^6 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 2 h incubation, 37 °C and 5% CO_2, 1 μL/mL GolgiPlug (BD, 555029) was added and cells were incubated 4 more hours at 37 °C and 5% CO_2. 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, 23 nM). 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 1/1,000) and incubated for 30 min, 4 °C. Cells were washed twice in barcode-cytometry buffer, incubated 30 min-overnight, 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 the MHC-multimer-stained cells were not acquired immediately, they were resuspended in 1% PFA and washed twice more in barcode-cytometry buffer after 1–24 h. 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 'dumped' (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, 5,000*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[5,6]. 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 50,000× 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 (USA) or GeneDx (USA) using Ion Torrent PGM 314 or 316 chip (Life Technologies).

Processing of sequencing data. To process the sequencing data and automatically identify the barcode sequences we designed a specific software package, 'Barracoda' (http://www.cbs.dtu.dk/services/barracoda). This software tool identifies the barcodes used in a given experiment, assigns sampleID and pMHC specificity to each barcode, and calculates the total number of reads and clonally reduced reads for each pMHC-associated DNA barcode. Furthermore, it includes a statistical processing of the data. Details are given below.

Raw sequence reads were aligned to the constant primer regions and the 16 nucleotide AxBy annealing region using Bowtie2 (ref. 48). Reads successfully mapped to at least two of these three features were aligned to the primer regions, the 25-mer 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 $n \times m$ 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[49]. 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[50]. *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 ten 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 1/1,000 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 were ≤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: CTGTGACTATG TGAGGCTTTC or ATGCAACCAAGAGCTTAAGT) and 0.1 µM sequence specific fluorescent reporter probes (probeB1: 6-FAM/GCCTGTAGTCCCA CGCGATCTAACA/BHQ and probeB2: HEX/CAACCATTGATTGGG GACAACTGGG/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 **Supplementary Figure 1c** 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-GAGATACGTTGACCTCGTTGAAN NNNNNTCTTGAACTATGAATCGTCTCACTTAAGCTCTTGGTTGCAT and ss-barcode2: Biotin-GAGATACGTTGACCTCGTTGAANNNNNNTC TATAGGTGTCTACTACCTCACTTAAGCTCTTGGTTGCAT (blue, green and black coloring indicates primer regions, UMI and the 25mer barcode region, respectively).

37. 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).
38. 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**, 3747–3759 (2015).
39. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
40. Gerlinger, M. *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat. Genet.* **46**, 225–233 (2014).
41. Szolek, A. *et al.* OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* **30**, 3310–3316 (2014).
42. 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. USA* **105**, 3825–3830 (2008).
43. Hadrup, S.R. *et al.* High-throughput T-cell epitope discovery through MHC peptide exchange. *Methods Mol. Biol.* **524**, 383–405 (2009).
44. 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**, 1109–1120 (2013).
45. 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. USA* **89**, 3429–3433 (1992).
46. 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).
47. Rammensee, H., Bachmann, J., Emmerich, N.P., Bachor, O.A. & Stevanović, S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50**, 213–219 (1999).
48. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
49. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
50. Robinson, M.D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).

doi:10.1038/nbt.3662