SUPPLEMENTARY MATERIALS

Tumor samples and peripheral blood collection

If the tumor content was estimated to be more than 40% after thorough pathological review, tumor DNA was extracted from freshly obtained tissues using a QIAamp Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. For DNA analysis, we used RNaseA (cat. #19101; Qiagen). We measured concentrations and 260/280- and 260/230-nm ratios with an ND1000 spectrophotometer (Nanodrop Technologies, Thermo Fisher Scientific, MA, USA) and then further quantified DNA using a Qubit fluorometer (Life Technologies, CA, USA).

Sequencing of Exome and Transcriptome

Genomic DNA from tumor tissue and matched blood samples was extracted using a QIAamp DNA Mini Kit (Qiagen). Library preparation was performed by using the SureSelect XT Human All Exon V6 kit (Agilent). Briefly, 900-1,000 ng of genomic DNA from tumor tissue and matched blood samples was sheared by Covaris Adaptive Focused Acoustics (AFA) sonication device (S2, Covaris Inc.), and 150-200 bp of the DNA fragments were processed for end-repairing, 3'-end adenylation, and ligation to adaptors. Sequencing libraries were performed on the HiSeq 2500 platform (Illumina) in 100-bp paired-end mode of TruSeq Rapid PE Cluster kit and TruSeq Rapid SBS kit (Illumina). RNA sequencing libraries were prepared using TruSeq RNA Exome Library Prep kit (Illumina) according to the manufacturer's protocol. Isolated total RNA was used in a reverse transcription reaction with random primers using SuperScript II reverse transcriptase (Invitrogen) according to manufacturer's protocols. RNA sequencing libraries were prepared via end-repair, 3'-end adenylation, adapter ligation, and amplification and those were sequenced 100-bp paired-end

mode of the TruSeq Rapid PE Cluster Kit and the TruSeq Rapid SBS Kit in Illuminia HiSeq 2500 (Illumina).

Variant calling and filtering

Exome sequencing reads were aligned to the human reference genome (GRCh37) using the bwa-mem algorithm from BWA (version 0.7.17). Further pre-processing of read alignment was performed using Genome Analysis Tool Kit (GATK, version 4.1.1.0)[1]. In brief, reads that were marked with duplicates were realigned to indels to remove alignment artifacts, and systematic errors in base quality scores were detected and recalibrated with known polymorphic sites of dbSNP (version 138)[2], 1000G (phase 1)[3] and HapMap (phase 3)[4] data using BaseRecalibrator and ApplyBQSR modules with default option parameters. We sought to identify pathogenic germline variants by scanning BAM files of each tumor and matched normal sample with the list for genomic regions of oncogenic germline variants found across pan-cancer[5]. The extraction of somatic mutations in the tumor samples was carried out with their matched normal samples using Mutect2[6] and Strelka2[7] to establish the highly sensitive somatic variant union sets. Variants with minimum depth ≥ 5 with at least 2 alternative alleles were used for further analysis, and annotated using Ensemble Variant Effect Predictor (VEP)[8] with the GRCh37 database. Allele-specific copy-number was quantified using FACETS (version0.6.0)[9] in default option parameters, and the resulting copy-ratio profiles were then used to estimate the fraction of cancer nuclei, average cancer genome ploidy, and somatic copy-number alterations by running ABSOLUTE (version 1.0.6)[10].

Mutational signature analysis

To assign single base substitutions (SBS) to the mutational signatures as defined by Alexandrov et al. [11], we utilized the deconstructSigs package (version 1.6.0) [12]. We

considered the characteristic patterns of mutations in trinucleotide context per sample with the reference of 'signatures.exome.cosmic.v3.may2019', and the contribution of each biological phenotype was represented as follows: age (SBS1 and SBS5), APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like; SBS2 and SBS13), UV (ultraviolet; SBS7a, SBS7b, SBS7c, and SBS7d), smoking (SBS4), immunoglobulin gene hypermutation (SBS9), HRD (homologous recombination deficiency; SBS3), MMRD (mismatch repair deficiency; SBS6, SBS15, SBS20 and SBS26), NERD (nucleotide excision repair deficiency; SBS8) [13], DPD (DNA proofreading deficiency; SBS10a and SBS10b) and BERD (base excision repair deficiency; SBS18).

Detection of LOH, TAI and LST

In search of additional genetic variants that are associated with HRD, we examined three independent measures of genomic instability: loss-of-heterozygosity (LOH) [14], telomeric allelic imbalance (TAI) [15], and large-scale state transitions (LST) [16]. From allele-specific copy number profiles, we determined LOH when absolute minor allelic copy number was zero and the other allele had absolute copy number > 0 spanning genomic regions over 15 megabases (Mb). LST was regarded in each chromosome when chromosome breaks (translocations, inversions or deletions) in adjacent segments of DNA was observed in larger than 10 Mb. TAI was defined when the absolute copy numbers between minor and major alleles were differentially observed in extending to the telomeric end, and not in crossing the centromere.

Identification of putative neoantigens

We identified putative neo-peptides using Mupexi[17] with NetMHCpan (version 4.0)[18] binding strength predictor between peptides and MHC molecules. The prediction depended on somatic mutations and HLA types. Transcript expression file was used to consider

expression of mutated peptides, which was optional. We performed HLA typing to determine HLA alleles and corresponding MHC complexes using Optitype[19]. We considered the neo-peptides with mutRANK < 2% for further analysis. We regarded a somatic mutation generating at least one neo-peptide as a neoantigen. Based on the cancer cell fraction (CCF) estimated by ABSOLUTE[10], newly emerged neoantigens were defined when the CCF of pre-treatment was zero and the CCF of on-treatment was not equal to zero.

Transcriptome sequencing analysis

Transcriptome reads were aligned to the human reference genome (GRCh38) using the 2-pass default mode of STAR (version 2.6.1)[20] with the annotation of ENSEMBL (version 98). Gene expression abundance as the unit of TPM (transcript per million) was estimated using the default option parameters in running RSEM (version 1.3.1)[21]. We estimated the gene-set enrichment scores of representative pathways involved in the TME using single-sample gene set enrichment analysis (ssGSEA) algorithm for each sample[22, 23] . By integrating the transcriptomic data, we classified each tumor sample into four distinct microenvironment subtypes: immune-depleted, fibrotic, immune-enriched, and immune-enriched/fibrotic[23]. We performed differentially expressed gene (DEG) analysis using the edgeR package[24]. Following normalization using trimmed mean of M values, tagwise dispersions were estimated and subjected to an exact test. DEGs were filtered according to the following criteria: expression fold change > 1.5; and *P* value from Wilcoxon rank sum test < 0.05. We calculated single-sample GSEA (ssGSEA) scores for MSigDB[25, 26] curated canonical genesets (C2.CP) using GSVA software package[27]. The change in ssGSEA scores after treatment was defined as the scores of pre-treatment subtracted from the scores of on-treatment. Intratumoral cell populations were estimated with MCP-Counter[28].

Sequencing of Single-cell RNA and T-cell receptor in peripheral blood mononuclear cells

Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll (GE Healthcare, Little Chalfont 17-5442-02 UK). After processing, PBMCs were resuspended in freezing media (RecoveryTM cell culture freezing medium, Gibco) and stored in liquid nitrogen. The cells were then cryopreserved in liquid nitrogen until use. All samples showed a viability of around 90% on average after thawing. Single-cell RNA-seq libraries were prepared by using the Chromium Next GEM Single Cell 5' Kit v1.1 (10× Genomics, Pleasanton, CA, USA) following the manufacturer's instructions. Briefly, the Chromium Controller instrument was utilized to encapsulate single cells into droplets and the single cells were barcoded reverse transcription (RT) of mRNA, followed by amplification, shearing and Illumina library construction. cDNA library quality was determined using Agilent Bioanalyzer and Next generation sequencing was performed by using the Novaseq6000 (200 cycles) cartridges from Illumina. TCR V(D)J segments were enriched from amplified cDNA from 5′ libraries via PCR amplification using a Chromium Single-Cell V(D)J Enrichment kit according to the manufacturer's protocol (10X Genomics).

Single cell RNA and TCR sequencing analysis

Single cell RNA sequencing (scRNA-seq) reads were aligned to the GRCh38 reference genome and quantified using Cell-Ranger (10X Genomics, version 3.1.0). We filtered out cells which met either of the following conditions: 1) putative doublets predicted by Scrublet[29], 2) low number of detected genes (< 200), and 3) high fraction of mitochondrial contamination (>10%). The remaining cells of raw UMI counts were then log-normalized with the scale factor of 10,000 and scaled across the given samples using Seurat package[30]. We performed the principal component analysis (PCA) on the integrated gene expression

profiles with the most variable 2500 genes, and the top principal components determined by 'elbow' heuristics were applied to remove batch effects across samples using Harmony algorithm[31]. For exploratory visualization, cells were projected into two-dimensional Uniform Manifold Approximation and Projection (UMAP) space[32]. Cells were clustered using FindNeighbors and FindCluster functions in Seurat package, and annotated by canonical marker genes.

To process single-cell TCR-sequencing data, we ran CellRanger 'vdj' pipeline (10X genomics, version.3.1.0) with the GRCh38 reference for demultiplexing, gene quantification and TCR clonotype assignment. TCR clonality was estimated as follows: [33]

$$Clonality = 1 - \frac{Shannon - Weiner\ index\ (H')}{\ln(\#\ of\ productive\ unique\ sequences)}$$

where Shannon-Wiener index (H') is the measure of diversity. Samples with monoclonal T cell population have a clonality of 0, while the clonality converges to 1 in samples with extremely diverse T cell population. To investigate how T cell clonotypes changed after the treatment, we grouped each T cell clone by comparing the clonal frequency of pre-treatment and that of on-treatment. T cell clones were defined as expanded clones when the clonal frequencies increased significantly compared to those of pre-treatment according to Fisher's exact test. T cell clones detected only in on-treatment samples were defined as novel clones. We excluded the clones with clone size 1 from the novel clones. We defined the T cell clones exhibiting decreased clonal frequencies as contracted clones based on Fisher's exact test. The rest of clones are defined as persistent clones.

Prediction of binding affinity of pMHC and TCR

To investigate the binding affinity of novel or expanded $CD8^+$ T cell receptors and newly emerged neo-peptides presented by MHC (pMHC) , we utilized ERGOII[34], a deep-learning

based tool for prediction of TCR-peptide binding. Inside the tool, long short-term memory (LSTM) neural network model was trained on McPAS-TCR[35], a database of pathology-associated T cell receptor sequences. We calculated the ERGOII binding score for all possible combinations of novel or expanded *CD8*+ TCRs and pMHCs derived from newly emerged neoantigens, and the highest affinity score for each clone is taken as a representative value for each clone.

1. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M *et al*: **A framework for variation discovery and genotyping using next-generation DNA sequencing data**. *Nature Genetics* 2011, **43**(5):491-498.
2. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: **dbSNP: the NCBI database of genetic variation**. *Nucleic acids research* 2001, **29**(1):308-311.
3. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA *et al*: **A global reference for human genetic variation**. *Nature* 2015, **526**(7571):68-74.
4. International HapMap C, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F *et al*: **Integrating common and rare genetic variation in diverse human populations**. *Nature* 2010, **467**(7311):52-58.
5. Huang K-l, Mashl RJ, Wu Y, Ritter DI, Wang J, Oh C, Paczkowska M, Reynolds S, Wyczalkowski MA, Oak N *et al*: **Pathogenic Germline Variants in 10,389 Adult Cancers**. *Cell* 2018, **173**(2):355-370.e314.
6. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G: **Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples**. *Nature Biotechnology* 2013, **31**(3):213-219.
7. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Källberg M, Chen X, Kim Y, Beyter D, Krusche P *et al*: **Strelka2: fast and accurate calling of germline and somatic variants**. *Nature Methods* 2018, **15**(8):591-594.
8. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F: **The Ensembl Variant Effect Predictor**. *Genome Biology* 2016, **17**(1):122.
9. Shen R, Seshan VE: **FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing**. *Nucleic Acids Research* 2016, **44**(16):e131-e131.

10. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, Laird PW, Onofrio RC, Winckler W, Weir BA *et al*: **Absolute quantification of somatic DNA alterations in human cancer**. *Nature Biotechnology* 2012, **30**(5):413-421.
11. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL *et al*: **Signatures of mutational processes in human cancer**. *Nature* 2013, **500**(7463):415-421.
12. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C: **DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution**. *Genome Biol*, **17**:31.
13. Jager M, Blokzijl F, Kuijk E, Bertl J, Vougioukalaki M, Janssen R, Besselink N, Boymans S, de Ligt J, Pedersen JS *et al*: **Deficiency of nucleotide excision repair is associated with mutational signature observed in cancer**. *Genome Res* 2019, **29**(7):1067-1077.
14. Abkevich V, Timms KM, Hennessy BT, Potter J, Carey MS, Meyer LA, Smith-McCune K, Broaddus R, Lu KH, Chen J *et al*: **Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer**. *British Journal of Cancer* 2012, **107**(10):1776-1782.
15. Birkbak NJ, Wang ZC, Kim J-Y, Eklund AC, Li Q, Tian R, Bowman-Colin C, Li Y, Greene-Colozzi A, Iglehart JD *et al*: **Telomeric Allelic Imbalance Indicates Defective DNA Repair and Sensitivity to DNA-Damaging Agents**. *Cancer Discovery* 2012, **2**(4):366.
16. Popova T, Manié E, Rieunier G, Caux-Moncoutier V, Tirapo C, Dubois T, Delattre O, Sigal-Zafrani B, Bollet M, Longy M *et al*: **Ploidy and Large-Scale Genomic Instability Consistently Identify Basal-like Breast Carcinomas with BRCA1/2; Inactivation**. *Cancer Research* 2012, **72**(21):5454.
17. Bjerregaard A-M, Nielsen M, Hadrup SR, Szallasi Z, Eklund AC: **MuPeXI: prediction of neo-epitopes from tumor sequencing data**. *Cancer Immunology, Immunotherapy* 2017, **66**(9):1123-1130.
18. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M: **NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data**. *The Journal of Immunology* 2017, **199**(9):3360.
19. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O: **OptiType: precision HLA typing from next-generation sequencing data**. *Bioinformatics* 2014, **30**(23):3310-3316.
20. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner**. *Bioinformatics* 2013, **29**(1):15-21.
21. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome**. *BMC Bioinformatics* 2011, **12**(1):323.
22. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C *et al*: **Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1**. *Nature* 2009, **462**(7269):108-112.

23. Bagaev A, Kotlov N, Nomie K, Svekolkin V, Gafurov A, Isaeva O, Osokin N, Kozlov I, Frenkel F, Gancharova O *et al*: **Conserved pan-cancer microenvironment subtypes predict response to immunotherapy**. *Cancer Cell* 2021, **39**(6):845-865.e847.
24. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data**. *Bioinformatics* 2010, **26**(1):139-140.
25. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: **Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles**. *Proceedings of the National Academy of Sciences* 2005, **102**(43):15545.
26. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP: **Molecular signatures database (MSigDB) 3.0**. *Bioinformatics* 2011, **27**(12):1739-1740.
27. Hänzelmann S, Castelo R, Guinney J: **GSVA: gene set variation analysis for microarray and RNA-Seq data**. *BMC Bioinformatics* 2013, **14**(1):7.
28. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH *et al*: **Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression**. *Genome Biology* 2016, **17**(1):218.
29. Wolock SL, Lopez R, Klein AM: **Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data**. (2405-4720 (Electronic)).
30. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zagar M *et al*: **Integrated analysis of multimodal single-cell data**. *bioRxiv* 2020:2020.2010.2012.335331.
31. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh P-r, Raychaudhuri S: **Fast, sensitive and accurate integration of single-cell data with Harmony**. *Nature Methods* 2019, **16**(12):1289-1296.
32. Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, Newell EW: **Dimensionality reduction for visualizing single-cell data using UMAP**. *Nature Biotechnology* 2019, **37**(1):38-44.
33. Cooper ZA, Frederick DT, Juneja VR, Sullivan RJ, Lawrence DP, Piris A, Sharpe AH, Fisher DE, Flaherty KT, Wargo JA: **BRAF inhibition is associated with increased clonality in tumor-infiltrating lymphocytes**. *OncoImmunology* 2013, **2**(10):e26615.
34. Springer I, Besser H, Tickotsky-Moskovitz N, Dvorkin S, Louzoun Y: **Prediction of Specific TCR-Peptide Binding From Large Dictionaries of TCR-Peptide Pairs**. *Front Immunol* 2020, **11**:1803-1803.
35. Tickotsky N, Sagiv T, Prilusky J, Shifrut E, Friedman N: **McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences**. *Bioinformatics* 2017, **33**(18):2924-2929.

Supplementary Figure S1. Study design and sample status (**a**) Overview of study design and sample collection. (**b**) Status of samples by data type. Blue filled squares indicate the presence of pre-treatment sample, and blue filled squares with "o" mark indicate the presence of both pre-treatment samples and on-treatment samples.

Supplementary Figure S2. Representative patients with distinct genetic profiles and their responses to ceralasertib in combination with durvalumab. (**a**) (Upper) The representative microscopic findings represent ATM complete loss and PD-L1 CPS score 1. (lower) Allele frequency plot demonstrates frequent loss of heterozygosity (LOH, red color), telomeric allelic imbalance (TAI, blue color), and large-scale transition (green checks). Areas colored purple represents both LOH and TAI. Plot showing absolute copy number marks copy number of minor allele (grey line) and total copy number (both major and minor allele). Amplified copy number lined in red and deleted copy number lined in blue. (**b**) Pre-treatment and on-treatment images of positron emission tomography–computed tomography. (**c**) (Upper) The representative microscopic findings represent intact ATM expression in tumor nucleus and PD-L1 CPS score 1. (lower) Allele frequency plot demonstrates no apparent LOH. (**d**) Pre-treatment and on-treatment coronal reconstructed images of computed tomography.

Supplementary Figure S3. Transcriptomic changes in TME signatures and cell types during treatment (**a**) Heatmap illustrating original GSVA scores of the gene sets shown in Figure 3b. The left dark grey bars represent log-transformed *P* value calculated by comparing GSVA score of responders (PR) and that of progressors (SD,

PD) in pre-treatment samples. (**b**) Changes in cytotoxic lymphocyte abundance during the study treatment. The cytotoxic lymphocyte score was estimated by deconvoluting the WTS data from pre- and on-treatment samples of eight patients.

Supplementary Figure S4. Cell type identification and comparison of the proportion of each cell type (**a**) UMAP plot of 32,787 total cells derived from pre-treatment (n=8) and on-treatment (n=7) peripheral blood samples and color-coded by global cell types, sample origin, timepoint, and best response. (**b**) (Left) UMAP plot of 19,621 T/NK cells color-coded by more subdivided cell type. (right) Comparison of the proportion of each subtype between pre-treatment samples from PR patients (green) and those from PD patients (ocher). (**c**) (Left) UMAP plot of 10,744 myeloid cells color-coded by more subdivided cell type. (right) Comparison of the proportion of each subtype between pre-treatment samples from PR patients (green) and those from PD patients (ocher). (**d**) (Left) UMAP plot of 1,143 B cells color-coded by more subdivided cell type. (right) Comparison of the proportion of each subtype between pre-treatment samples from PR patients (green) and those from PD patients (ocher). Only significant *P* values from Wilcoxon rank sum test are shown.

Supplementary Figure S5. TCR detection rate and distribution of CD8$^+$ T cell clones (**a**) UMAP plot of 19,621 T / NK cells from pre- (n=8) and on-treatment (n=7) peripheral blood samples. Cells are color-coded according to the timepoint. (**b**) (Left) UMAP plot of T / NK cells color-coded by TCR detection. (right) Bar plot showing the percentage of TCR-detected cells in each T cell subtype. (**c**) (Upper) Number of

each categorized clone type per patient. (lower) Proportion of each categorized clone type among all CD8$^+$T cell clones.

Supplementary Figure S6. Expression of canonical markers to assign cell types (**a**) Dot plot showing expression of marker genes used for global cell type annotation. (**b**) Dot plot showing expression of marker genes used for T / NK cell subtype annotation. (**c**) Dot plot showing expression of marker genes used for myeloid cell subtype annotation. (**d**) Dot plot showing expression of marker genes used for B cell subtype annotation.