Feb 20, 2025

MEMORI-Methods for computational analyses

Nature Cancer

DOI
https://dx.doi.org/10.17504/protocols.io.j8nlk9xpwv5r/v1

Melissa Barroux[1], Jacob Househam[2], Eszter Lakatos[2], Tahel Ronel[2], Trevor Graham[2]

[1]Medical Clinic and Polyclinic II, Klinikum rechts der Isar, Technical University of Munich, München, Germany; [2]Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK

Melissa Barroux
Klinikum rechts der Isar

OPEN ACCESS

DOI: https://dx.doi.org/10.17504/protocols.io.j8nlk9xpwv5r/v1

External link: https://doi.org/10.1038/s43018-025-00955-w

Protocol Citation: Melissa Barroux, Jacob Househam, Eszter Lakatos, Tahel Ronel, Trevor Graham 2025. MEMORI-Methods for computational analyses. **protocols.io** https://dx.doi.org/10.17504/protocols.io.j8nlk9xpwv5r/v1

Manuscript citation:
Barroux, M., Househam, J., Lakatos, E. *et al.* Evolutionary and immune microenvironment dynamics during neoadjuvant treatment of esophageal adenocarcinoma. *Nat Cancer* **6**, 820–837 (2025). https://doi.org/10.1038/s43018-025-00955-w

Protocol status: Working

Created: February 20, 2025

Last Modified: February 20, 2025

Protocol Integer ID: 123125

Keywords: oesophageal adenocarcinoma patients in the memori trial, oesophageal adenocarcinoma patient, immune microenvironment dynamics during neoadjuvant treatment, oesophageal adenocarcinoma, immune cell dynamics during treatment, immune cell dynamic, immune microenvironment dynamic, methods for computational analysis, neoadjuvant treatment, image mass cytometry analysis, computational analysis, mass cytometry, memori trial, computational method

Funders Acknowledgements:
German Cancer Aid
DKTK
CRUK

Abstract

This protocol describes computational methods used for the manuscript "Evolutionary and immune microenvironment dynamics during neoadjuvant treatment of oesophageal adenocarcinoma" by Barroux et al. In this study we performed longitudinal multi-omic analysis (WES, RNA-Seq, TCR-Seq and image mass cytometry analyses) of oesophageal adenocarcinoma patients in the MEMORI trial. Multi-timepoint multi-omic analyses was performed on 27 patients (before, during and after neoadjuvant treatment) to study evolutionary and immune cell dynamics during treatment.

1.1 Whole-exome sequencing – alignment:

1 Contaminating adapter sequences were removed using Skewer v0.2.2[1], with a maximum error rate of 0.1, minimum mean quality value of 10 and a minimum read length of 50 after trimming using options "-l 50 -r 0.1 -d 0.03 -Q 10 -n". The trimmed and filtered reads from each sequencing run and library where separately aligned to the GRCh38 reference assembly of the human genome[2] using the BWA-MEM algorithm v0.7.17[3]. Following the GATK best practices and the associated set of tools v4.1.4.1[4-6] reads were sorted by coordinates (GATK SortSam), independent sequencing runs from the same tissue sample were merged and duplicated reads were marked using GATKs MarkDuplicates. The coverage statistics of the final bam files were verified with Qualimap v2.2.1[7].

1.2 Copy number analysis:

2 Copy number analysis and cellularity estimates were performed with Sequenza[8]. Pre-processing of bam files to seqz files was done using the sequenza-utils bam2seqz algorithm and binned using the sequenza-utils seqz_binning -w50 command. The per patient set of break points, binned depth-ratio and B-allele frequencies data from binned seqz files were then inputted into the sequenza algorithm (version 2.1.2) to determine allele specific copy-numbers, ploidy Ψ and purity ρ estimates[8]. The threshold segment length for fitting cellularity and ploidy parameters was set to 10^6 bp to filter out short segments, that can cause noise for cellularity and ploidy estimates. The initial parameter space searched was restricted to $[\rho \mid 0.1 \leq \rho \leq 1]$ and $[\Psi \mid 1 \leq \Psi \leq 5]$. Upon manual review of the results, we identified several samples with unreasonable fits (cases where calls suggested extremely variable ploidy values across samples from the same patient). For these samples, we manually identified alternative solutions consistent with the other samples and somatic variant calls. If samples showed no EAC-like CNAs or EAC-specific SNVs, samples were excluded from genetic analyses. Clonal CNAs were defined as CNAs present in all samples of a patient, subclonal CNAs were defined as CNAs present in > 1/3 of samples, but not all samples of a patient and private CNAs as CNAs present in ≤ 1/3 of samples of a patient. For patients with only 2 available samples subclonal/private CNAs were defined as CNAs present in 1 sample. To calculate the fraction of exome, that changes its CNS between timepoint A and B (Fig 3D), 15 patients with available samples from timepoint A and B and no multi-region samples at timepoint A or B were included.

1.3 SNV detection:

3 Somatic mutations were first called for each tumour sample separately against matched blood samples with Mutect2 (version 4.1.2.0)[9]. Variants detected were annotated using

ANNOVAR[10]. Variants detected in any tumour sample (marked PASS, coverage AD 5 in both normal and tumour, at least 3 variant reads in the tumour and a maximum variant read of 0 in the normal reference genotype) were merged into a single list of "candidate mutations". For later MOBSTER-analysis[11] for multi-region/multi-timepoint samples (see below), SNV caller bam-readcount[12] was used to call nucleobases and coverage at each candidate mutation position in all samples of a respective patient. Candidate driver genes of EAC were annotated using 108 cancer driver genes for oesophageal and gastric cancer reported by IntOGen© [13],[14].

1.4 Identification of subclonal and clonal mutations per sample:

4 For each sample "candidate mutations" were classified as clonal or subclonal mutation using MOBSTER[15]. Based on the calculated variant allele frequency (VAF), i.e., the ratio of read counts mapping to the mutant allele, over the total coverage at the variant locus, which was previously normalized by the tumour purity and tumour copy number segments, MOBSTER clusters input mutations as clonal peak mutations (=clonal mutations) or tail mutations for neutral mutations (= subclonal mutations).

1.5 Construction of phylogenetic trees:

5 For phylogenetic tree analysis we analysed SNVs from multi-region/multi-timepoint samples of each patient using MOBSTER for multiple samples. To avoid previously described biases in clonality analysis of multi-region/multi-timepoint analysis[11], MOBSTER excludes mutations belonging to the evolutionary neutral-tail for clonality analysis in multi-region/multi-timepoint samples. From non-tail mutations maximum-parsimony trees were reconstructed with the Parsimony Ratchet method[16] using the phangorn package for R[17]. Included non-tail mutations within a sample were considered to be mutated (state 1) and others considered to be non-mutated (state 0) in a given sample. The Ratchet was run for a minimum of 10 and a maximum of 10^3 iterations and terminated after 10 rounds without improvement. The acctran algorithm[17-19] was utilized to approximate ancestral character states. From these a set of shared mutations between samples and mutations that were uniquely mutated on each clade of the phylogeny were obtained.

1.6 Extraction of mutational signatures and dN/dS analyses:

6 The extraction of mutational signatures in tumour samples was carried out with deconstructSigs[20], using COSMIC signatures as reference signatures[21],[22] and the default settings. Unfiltered signatures for each tumour sample are shown in the extended data (Extended Data fig 2E). In order to reduce the number of called mutation signatures to more prevalent ones, we rerun the analysis excluding signatures that had a COSMIC signature weight of <5% in any of the responsiveness groups at a given timepoint. Next, we evaluated mutational footprints of chemo- and radiochemotherapy in our cohort.

COSMIC signatures were called separately for treatment-naïve mutations, mutations occurring under chemotherapy (= "CTx-SNVs") and mutations occurring under RCTx (="RCTx-SNVs "). Treatment-naive mutations included mutations at timepoint A in REPs and NRPs, CTx-SNVs included newly occurring mutations at timepoint B in REPs and NRPs and newly occurring mutations at timepoint C in REPs and RCTx-SNVs harboured newly occurring mutations at timepoint C in NRPs. dN/dS ratios were calculated for treatment-naïve mutations, CTx-SNVs and RCTx-SNVs using the dNdScv R package[23].

1.7 Genetic alterations and RNA expression in EAC drivers:

7 SNVs, indels and CNA were assessed in 108 IntOGen© drivers[13] of oesophageal and gastric cancer. To assess the number of altered driver genes in each sample, driver genes were counted as altered if they harboured one or more SNVs, indels or CNAs. To assess the correlation between CNS and RNA expression of 15 high frequency driver genes, samples with matching WES and RNA-Sequencing data were analysed (n=64). RNA-Sequencing data was pre-processed as described in 2.1 and 2.2. Changes in expression levels of driver genes between treatment-naïve samples and post-treatment samples were calculated by Wilcoxon tests. For correlation analyses between the expression level and the CNS of driver genes, the expression level was adjusted for the estimated tumour cellularity. Correlation between expression levels and CNS were calculated for each driver genes by Pearson correlation. For expression analyses of driver genes in samples with mutated and non-mutated driver status, only CDKN2A and TP53 displayed a sufficiently powered analysis. The expression level of the respective driver was adjusted for the estimated tumour cellularity. Expression levels between mutated and non-mutated samples were compared by the Wilcoxon test.

1.8 HLA haplotyping:

8 HLA haplotyping for each patient was carried out with the raw paired fastqs of blood-derived normal samples using Optitype[24]. The pair of A, B and C alleles with the highest score were taken as a patient's haplotype.

1.9 Calling immune escape:

9 We predicted four types of immune escape mechanisms: (1) somatic mutations in one of the HLA alleles or (2) in the *B2M* gene[25]; (3) loss of an HLA haplotype through LOH[26]; and (4) PD-L1 overexpression[27].
Somatic mutations in the HLA locus were called using polysolver[25]. First, alleles were converted into a polysolver-compatible format (lower case, digits separated by underscore) and outputted into a patient-specific winners.hla.txt file, following the standard output of polysolver. Next, the mutation detection script of polysolver (shell_call_hla_mutations_from_type) was run on matched tumour- normal pairs

to call tumour-specific alterations in HLA-aligned sequencing reads using MuTect (v1.16). In addition, Strelka2 (v2.9.9)[28] was run independently to detect short insertions and deletions in HLA-aligned reads as this software offers increased sensitivity over polysolver's default caller. Finally, both single nucleotide mutations and insertions/deletions passing quality control were annotated by polysolver's built-in annotation script, shell_annotate_hla_mutations. LOH at the HLA locus was assessed using the software LOHHLA[26] and sequenza[8]. First, the allele-specific copy number as predicted by sequenza at the HLA-A, -B, -C loci was evaluated. Samples with a predicted minor allele copy number of 0 (e.g. 2:0, 3:0) were labelled as candidate LOH. Then, we ran LOHHLA with the previously generated winners.hla.txt as input. A type-I allele of a patient was annotated as "allelic imbalance" (AI) if the p-value testing the difference in evidence for the two alleles was lower than 0.01. Alleles with AI were labelled as LOH if the following criteria held: (i) the predicted copy number of the lost allele was below 0.5 with confidence interval strictly below 0.7; (ii) the copy number of the kept allele was above 0.75; (iii) the number of mismatched sites between alleles was above 10. All LOHHLA-based LOH called were in candidate LOH samples, but some candidate LOH could not be confirmed most probably due to low purity and read density.

Mutations in *B2M* were called if the sample contained a nonsynonymous somatic, frameshift (FS), stop-loss or stop-gain mutation located inside the exonic regions of the exons of the *B2M* gene, as called by ANNOVAR[10]. PD-L1 overexpression was assessed using matched RNA-seq data from the samples. RNA-Seq raw data was pre-processed and then normalized as described in 2.1 and 2.2. The mean expression value of normalized gene counts (in transcripts per million mapped reads (TPM)) of PD-L1) in treatment-naïve EAC samples was calculated and considered as PDL-1 treatment-naïve baseline. PDL-1 overexpression was defined as ≥ PD-L1 treatment-naïve baseline + 1 standard deviation. When assessing immune escape mechanisms (transcriptomic versus genetic immune escape) samples with matching WES and RNA-Sequencing data were included (n=64).

1.10 Calling neoantigens:

10 Neoantigens were predicted from variant call tables and HLA types using the NeoPredPipe[29], python-based pipeline combining Annovar and netMHCpan 4.0. Briefly, all quality-controlled somatic mutations were annotated using Annovar[10], and for all non-synonymous exonic mutations the mutated peptide sequence was predicted. We took any 9- and 10-mer spanning the mutated amino acid(s), resulting in either (i) a 19-aa window for SNVs or (ii) a peptide until the next predicted stop codon for FSs. These peptides were evaluated according to their novelty and predicted binding strength to the patient's six-allele HLA set. Peptides novel compared to the healthy human proteome with binding rank 2 and below (amongst the best 2% of binders compared to a large set of random peptides) were reported as neoantigens. All patient-specific HLA alleles were used for neoantigen predictions, regardless of mutation or LOH status of the

HLA locus. We considered a mutation neoantigenic if at least one of its downstream mutated peptides were a neoantigen with respect to any of the patient's six HLA alleles. The clonality of neoantigentic mutations was assessed based on MOBSTER[15] analyses (more details in 1.4). To calculate the enrichment/depletion of neoantigens in regions of different copy number states, we calculated for each sample, the ratio between the fraction of copy-number adjusted NeoSNVs present in a certain CN category (loss, cnLOH, high LOH, diploid, gain and amplification) and the fraction of genome showing the respective CN category.

1.11 Expression of neoantigens:

11 To assess the expression levels of neoantigenic peptides, we analyzed gene expression levels of genes harbouring neoantigenic mutations. RNA-Seq raw data was pre-processed and then normalized as described in 2.1 and 2.2. As a proxy for neoantigen expression, expression levels (CPM) of protein-coding genesharbouring neoantigenic mutations were assessed. Mean expression levels of neoantigens were calculated for each sample and compared between samples from different timepoints via Wilcoxon tests. For analyses of neoantigen expression levels in low and high CD4 and CD8 cell groups, CD4 and CD8 T-cell levels were deconvoluted from normalized RNA-sequencing data using CIBERSORT[30] (more details in methods part 2.5). Samples were assigned to high and low CD4 and CD8 cell groups based on the median CD4 and CD8 score of the study sample.

2.1 RNA sequencing – alignment and filtering of RNA samples:

12 After initial quality control with FastQC[31] and default adaptor trimming with Skewer[1], paired-end reads were aligned to GRCh38 reference genome and version 28 of the Gencode GTF annotation using the STAR 2-pass method[32]. Read groups were added with Picard v.2.5.0[33]. Per gene read counts were produced with htseq-count that is incorporated into the STAR pipeline[32]. Raw gene counts were first filtered for reads uniquely assigned to non-ribosomal protein-coding genes located on canonical chromosomes (chr1-22, X and Y). If samples had less than 1.5M of these 'usable' reads they were re-sequenced to improve coverage. Where possible, the same library preparation pool was sent again for sequencing. These `top-ups' proved to be true technical replicates, since the resulting gene expression of the re-sequenced samples clustered very closely to their original samples on a principal component analysis (PCA). It was therefore determined that the fastqs of these samples could simply be merged at the start of the pipeline. In cases where resequencing was required but insufficient library remained, a new library was prepared and the sequencing run that produced the highest read was used in subsequent analysis. If sequencing depth (adding `top-ups' of the second library where necessary) contained $<1.5\times10^6$ reads, samples were excluded for further downstream analysis.

2.2 Gene expression normalisation and filtering of expressed genes:

13 A master table including raw read counts for each samples of 58,243 genes were converted to a DGEList object via DESeq2[34]. Next, a list of expressed genes (n=11,900) was determined by edgeR[35] default filtering, to determine genes with sufficiently large counts to be retained in a statistical analysis. Normalization and conversion into counts per million (CPM) was performed using edgeR[35]. Samples from different sequencing batches and post-treatment samples with different fixation methods clustered mixed on a PCA, indicating no notable batch effects (Extended Data fig. 9A+B).

2.3 Pathway enrichment clustering:

14 Hallmark pathways were download from MSigDB[36] and unrelated pathways (SPERMATOGENSIS, MYOGENESIS and ANDROGEN_RESPONSE) were excluded, leading to a final list of 47 pathways. Based on their biological function pathways were categorized into 5 groups: oncogenic, immune, cellular stress, stromal and other. For each tumour sample, the CPM expression of expressed genes converted to entrez gene IDs (n= 10,463) was used as input for single sample gene set enrichment analysis using the GSVA R package[37]. The mean and standard deviation of enrichment was then recorded for each sample. Unsupervised clustering of samples based on gene set enrichment scores for each pathway was used to cluster tumour samples via principal component analysis. Loadings of defined hallmark pathways and hallmark classes (oncogenic, immune, stromal, cellular stress) on the PCA dimensions of interest was visualized separately. Pathways categorised in the class of “others” and regarded as cancer unrelated were removed from the loading plot for clearer visualisation purposes. As PCA clustering showed most samples from timepoint A and B clustering separately from timepoint C on PC1, differentially expressed pathways between samples from timepoint A/B and timepoint C were determined via FDR adjusted p-values determined via Wilcoxon signed-rank tests. Supervised clustering with significantly differentially expressed hallmark pathways between samples from timepoint A/B and timepoint C were visualised in a heatmap created with the pheatmap R package.

2.4 KEGG pathway enrichment analysis:

15 For each tumour sample DESeq normalized counts (or mean normalized counts across samples for multi-region sampled tumours) were used for KEGG pathway enrichment analysis. For each gene, log2 fold changes between patient groups were calculated via DESeq analysis to determine differential expression. Conversion to entrez gene IDs and gene symbols was performed using the Biological Id Translator function from ClusterProfiler (v4.2.2)[38]. The enrichment of KEGG pathways for differentially expressed genes between patient groups was determined with enrichKEGG from ClusterProfiler[38]. Adjusted p-values were calculated using the false discovery rate.

Enriched pathways with adjusted p-value <0.1 were plotted in Fig 4D and significantly enriched pathways (p<0.05) were demarcated by a dotted line in the graph.

2.5 Immune deconvolution from RNA-Sequencing data

16 Immune cells were deconvoluted with CIBERSORT[30] a computational method for quantifying cell fractions from bulk tissue gene expression profiles based on support vector regression with prior knowledge of expression profiles from purified leukocyte subsets. For each tumour sample, the CPM expression of genes converted to Gene Symbols was used as input for CIBERSORT single sample immune deconvolution and was uploaded to the online available CIBERSORT tool (https://cibersort.stanford.edu/index.php). A LM22 signature file compromising 22 immune cell types and included in the deconvolution tool was used for deconvolution. Deconvoluted immune cell data from our samples was outputted as absolute and relative scores. For further analysis CIBERSORT output tables were processed in R. For each sample CIBERSORT scores of "Macrophages.M0", "Macrophages.M1" and "Macrophages.M2", were summed to a joint score for “Macrophages”, scores of activated and resting cells of the same cell types were summed (e.g. "NK.cells.resting" and "NK.cells.activated" were summed to a “NK cell” score), scores for "B.cells.naive", "B.cells.memory" and "Plasma.cells" were summed to “Other B cells”, scores for "T.cells.follicular.helper", "T.cells.regulatory..Tregs.", "T.cells.gamma.delta" were joint to “Other T cells” and scores for “T.cells.CD4.naive”, “T.cells.CD4.memory.activated” and “T.cells.CD4.memory resting” were added to common score for “CD4 T-cells”. For validation purposes further immune cell deconvolution methods were used. For each tumour sample, the CPM expression of protein-coding genes and expressed genes converted to entrez gene IDs (n= 10,463) was used as input for single sample gene set enrichment analysis using the Consensus^{TME} gene sets[39] and the GSVA R package[37] in R. As a third immune deconvolution tool Syllogist[40] was used. For each tumour sample, the CPM expression of protein-coding genes converted to HUGO nomenclature (n= 39,921) was used as input for the provided Syllogist script[40], which was run in R. Odds ratios of immune cells of interests were visualized in bar graphs generated in R.

3.1 IMC raw data processing:

17 Raw data were exported as MCD files. MCD files were converted into tiff files and segmented into single cells using an unbiased, supervised analysis pipeline adapted from https://github.com/BodenmillerGroup/ImcSegmentationPipeline. In more detail, nucleated cells were segmented using Ilastik v1.3.3[41] and CellProfiler v4.1.3[42]. Using Ilastik, pixels were classified into nuclei, cytoplasm and background. A probability map for the three classifications was generated and exported to create a cell mask with CellProfiler. Data folders containing tiff images of the 21 markers and the combined cell mask were loaded into histoCAT v1.76. In histocat and Cellprofiler mean marker intensity

of pixels and spatial features of segmented cells were analyzed and exported as a table containing single cell data for each marker. Subsequent processing of single-cell level data was performed in OMIQ.

3.2 Cluster analysis

18 IMC single-cell level data and clinical data for ROIs from the exploratory IMC set were uploaded in OMIQ. Arcsinh transformation (factor 0.5) was performed. To exclude batch effects based on PFPE and FFPE fixation methods, mean expression of each marker between FFPE and PFPE post-treatment samples from the exploratory set was assessed (Extended Data fig 9C). 21 markers with specific staining in the mcd images and no significant difference in staining intensity between post-treatment FFPE and PFPE samples (Extended Data fig 9C) were further used for principal component analysis. PCA showed no major batch effect between samples from different fixation methods (Extended Data fig 9D, E). The three FFPE ROIs clustering slightly apart in the PCAs (circled in the PCAs) were from the same patient (RE23) (Extended Data fig 9E), indicating most likely a different biology in this patient. Thus all 21 markers were eligible for further analyses. Visualization of the global single-cell landscape was performed in OMIQ after arcsinh transformation (factor 0.5) using Opt-SNE based on channels (CD163, CD204, CD3, CD45RO, CD45, CD4, CD8a, Ecadherin, Eomes, GzmB, Ki67, PD-1, PD-L1, PanCK, SMA, Tim-3) (random seed 3367, max iterations 1000, opt-SNE end 5000, perplexity 30, theta 0.5 and verbosity 25). Phenograph clusters were visualized using the Phenograph algorithm based on the cell markers (CD163, CD204, CD3, CD45RO, CD4, CD8a, Ecadherin, Eomes, GzmB, Ki67, PD-1, PD-L1, PanCK, SMA, Tim-3) (k nearest neighbours 60, Louvain seed 8216, distance metric euclidean, Louvain runs 1, number of results 1) and Z-score of median marker expression of 28 phenograph clusters was visualized with a heatmap in OMIQ. For initially ambiguous clusters expressing markers of distinct cell types (C1, C4, C6, C11, C14, C18, C23 and C24), individual cluster cells were manually reviewed on the original image file in order to define the cluster's cell type. Exemplary IMC stainings of cells from those clusters are shown in Extended Data fig 10. Based on their marker profiles, all clusters were classified into different cluster types (tumour cell, immune cell and clusters of other cells). For immune cell clustering, $CD45^+$ cells were gated based on their CD45 expression. Marker positivity was determined based on thresholds obtained by sampling background areas and positive cells of the image in MCD viewer. Visualization of the immune single-cell landscape was performed in OMIQ after arcsinh transformation (factor 0.5) using Opt-SNE based on channels (CD163, CD204, CD38, CD3, CD45RO, CD4, CD8a, Eomes, FoxP3, GzmB, HLA-DR, Ki67, PD-1, TCF1, Tbet, Tim-3) (random seed 8836, max iterations 1000, opt-SNE end 5000, perplexity 30, theta 0.5 and verbosity 25). Phenograph clusters were visualized using the Phenograph algorithm based on the same markers (k nearest neighbours 50, Louvain seed 7583, distance metric euclidean, Louvain runs 1, number of results 1). Differences in cell counts of identified clusters between REPs and NRPs at

different timepoints were analysed by a two-sided anova tests (e.g. Fig 6E) and * indicates $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

3.3 Phenotype analyses of CD4+ and CD8+ T cells:

19 $CD4^+$ and $CD8^+$ cells were gated from each sample based on their CD4 and CD8a expression respectively. Marker positivity was determined based on thresholds obtained by sampling background areas and positive cells of the image in MCD viewer (supplemental table 4). The imaging mass cytometry from the single region dataset consisted of n= 25,973 $CD4^+$ cells and n= 7,830 $CD8^+$ cells. Quantification of distinct T-cells with different phenotypes, was assessed by further subsampling $CD4^+$ and $CD8^+$ cells based on their expression of markers for differentiation, transcriptional programming, activation/exhaustion, and cytotoxicity (CD45RO, Granzyme B, HLA-DR, PD-1, Eomes, Tim-3, CD38, Ki67). Marker positivity was determined based on thresholds obtained by sampling background areas and positive cells of the image in MCD viewer (supplemental table 4). Gated T-cell composition including counts per sample and fractions of respective parent cell, were exported from OMIQ. Different T-cell phenotypes between patient groups were visualized using violinplots generated with the ggplot2 function in R studio and compared between patient groups by Wilcoxon tests. Our multiregion dataset incorporated 47 additional ROIs from resection specimens only (supplemental table 7) and consisted of n= 45,940 $CD4^+$ cells and n= 27,018 $CD8^+$ cells. Marker positivity needed to be redetermined for these validation ROIs due to the deterioration of signal between staining and image acquisition and revised thresholds are listed in supplemental table 4.

4. TCR sequencing:

20 Full details for both the experimental TCRseq library preparation and the subsequent computational analysis (V, J, and CDR3 annotation) using Decombinator are published in our previous work[43],[44].

4.1 TCR statistical analysis:

21 Statistical analyses were performed in R. Wilcoxon rank tests were used to compare the abundance of detected TCRs between timepoints in NRPs and REP. For correlation analysis between deconvoluted T-cells from CIBERSORT and total TCRs, CIBERSORT scores for 'CD4 T-cells naive', 'CD4 T-cells rest', 'CD4 T-cells active' and 'CD8 T-cells' were summed to a ‘Total T-cell score’ for each sample. Correlation between calculated ‘Total T-cell scores’ and abundance of detected TCRs was calculated using Pearson correlation analysis. To calculate expansions of TCRs during treatment, individual TCRs were normalized in each sample to TCRs/million. The sets of TCRs that had expanded in proportion by at least 4 times, and at least 8 times, between

timepoints A and B, A and C, and B and C were identified. The proportions of the TCRs that had expanded between timepoint A and C were tracked across all timepoints and plotted for patients with TCR data from all 3 timepoints. The number of at least 4-fold and 8-fold expanded TCRs between any 2 timepoints was plotted for patients with different clinical treatment response. As TCRseq data are relatively sparse with large differences in data quality, large fold changes were selected to highlight major changes in the T-cell repertoire. Statistical comparisons between the number of expanded TCRs and regression scores were calculated by Mann-Whitney U tests.

References:

22

1. Jiang H, Lei R, Ding SW, Zhu S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics. 2014; 15: 182. doi: 10.1186/1471-2105-15-182.
2. Schneider VA, Graves-Lindsay T, Howe K, Bouk N, Chen HC, Kitts PA, Murphy TD, Pruitt KD, Thibaud-Nissen F, Albracht D, Fulton RS, Kremitzki M, Magrini V, et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. Genome Res. 2017; 27: 849-64. doi: 10.1101/gr.213611.116.
3. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 1754-60. doi: 10.1093/bioinformatics/btp324.
4. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013; 43: 11 0 1- 0 33. doi: 10.1002/0471250953.bi1110s43.
5. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20: 1297-303. doi: 10.1101/gr.107524.110.
6. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43: 491-8. doi: 10.1038/ng.806.
7. Okonechnikov K, Conesa A, Garcia-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics. 2016; 32: 292-4. doi: 10.1093/bioinformatics/btv566.
8. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, Szallasi Z, Eklund AC. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. Ann Oncol. 2015; 26: 64-70. doi: 10.1093/annonc/mdu479.
9. 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. Nat Biotechnol. 2013; 31: 213-9. doi: 10.1038/nbt.2514.
10. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010; 38: e164. doi: 10.1093/nar/gkq603.
11. Caravagna G, Heide T, Williams MJ, Zapata L, Nichol D, Chkhaidze K, Cross W, Cresswell GD, Werner B, Acar A, Chesler L, Barnes CP, Sanguinetti G, et al. Subclonal reconstruction of tumors by using machine learning and population genetics. Nat Genet. 2020; 52: 898-907. doi: 10.1038/s41588-020-0675-5.
12. Khanna A, Larson DE, Srivatsan SN, Mosior M, Abbott TE, Kiwala S, Ley TJ, Duncavage EJ, Walter MJ, Walker JR, Griffith OL, Griffith M, Miller CA. Bam-readcount -- rapid generation of basepair-resolution sequence metrics. ArXiv. 2021. doi:
13. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, Santos A, Lopez-Bigas N. IntOGen-mutations identifies cancer drivers across tumor types. Nat Methods. 2013; 10: 1081-2. doi: 10.1038/nmeth.2642.
14. Martinez-Jimenez F, Muinos F, Sentis I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, Mularoni L, Pich O, Bonet J, Kranas H, Gonzalez-Perez A, Lopez-Bigas N. A compendium of mutational cancer driver genes. Nat Rev Cancer. 2020; 20: 555-72. doi: 10.1038/s41568-020-0290-x.
15. Caravagna G, Sanguinetti G, Graham TA, Sottoriva A. The MOBSTER R package for tumour subclonal deconvolution from bulk DNA whole-genome sequencing data. BMC Bioinformatics. 2020; 21: 531. doi: 10.1186/s12859-020-03863-1.
16. Nixon KC. The Parsimony Ratchet, a New Method for Rapid Parsimony Analysis. Cladistics. 1999; 15: 407-14. doi: 10.1111/j.1096-0031.1999.tb00277.x.
17. Schliep KP. phangorn: phylogenetic analysis in R. Bioinformatics. 2011; 27: 592-3. doi: 10.1093/bioinformatics/btq706.
18. Swofford DLM, W. P. Reconstructing ancestral character states under Wagner parsimony. Math Biosci. 1987; 87, 199–229. doi:
19. Farris JS. Methods for Computing Wagner Trees. Syst Zool. 1970; 19, 83. doi:
20. 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. 2016; 17: 31. doi: 10.1186/s13059-016-0893-4.
21. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale A-L, Boyault S, Burkhardt B, Butler AP, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500: 415-21. doi: papers2://publication/doi/10.1038/nature12477.
22. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. Cell Rep. 2013; 3: 246-59. doi: 10.1016/j.celrep.2012.12.008.
23. Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, Davies H, Stratton MR, Campbell PJ. Universal Patterns of Selection in Cancer and Somatic Tissues. Cell. 2017; 171: 1029-41 e21. doi: 10.1016/j.cell.2017.09.042.

24. Szolek A, Schubert B, Mohr C, Sturm M, Feldhahn M, Kohlbacher O. OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics. 2014; 30: 3310-6. doi: 10.1093/bioinformatics/btu548.
25. Shukla SA, Rooney MS, Rajasagi M, Tiao G, Dixon PM, Lawrence MS, Stevens J, Lane WJ, Dellagatta JL, Steelman S, Sougnez C, Cibulskis K, Kiezun A, et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nat Biotechnol. 2015; 33: 1152-8. doi: 10.1038/nbt.3344.
26. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, Birkbak NJ, Veeriah S, Van Loo P, Herrero J, Swanton C, Consortium TR. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell. 2017; 171: 1259-71 e11. doi: 10.1016/j.cell.2017.10.001.
27. Alsaab HO, Sau S, Alzhrani R, Tatiparti K, Bhise K, Kashaw SK, Iyer AK. PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome. Front Pharmacol. 2017; 8: 561. doi: 10.3389/fphar.2017.00561.
28. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Kallberg M, Chen X, Kim Y, Beyter D, Krusche P, Saunders CT. Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods. 2018; 15: 591-4. doi: 10.1038/s41592-018-0051-x.
29. Schenck RO, Lakatos E, Gatenbee C, Graham TA, Anderson ARA. NeoPredPipe: high-throughput neoantigen prediction and recognition potential pipeline. BMC Bioinformatics. 2019; 20: 264. doi: 10.1186/s12859-019-2876-4.
30. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods Mol Biol. 2018; 1711: 243-59. doi: 10.1007/978-1-4939-7493-1_12.
31. s-andrews. FastQC: A quality control analysis tool for high throughput sequencing data. https. (GitHub.
32. 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 (Oxford, England). 2013; 29: 15-21. doi: papers2://publication/doi/10.1093/bioinformatics/bts635.
33. Picard Tools.
34. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15: 550. doi: 10.1186/s13059-014-0550-8.
35. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26: 139-40. doi: 10.1093/bioinformatics/btp616.
36. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015; 1: 417-25. doi: 10.1016/j.cels.2015.12.004.
37. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14: 7. doi: 10.1186/1471-2105-14-7.

38. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16: 284-7. doi: 10.1089/omi.2011.0118.
39. Jimenez-Sanchez A, Cast O, Miller ML. Comprehensive Benchmarking and Integration of Tumor Microenvironment Cell Estimation Methods. Cancer Res. 2019; 79: 6238-46. doi: 10.1158/0008-5472.CAN-18-3560.
40. Lu IN, Dobersalske C, Rauschenbach L, Teuber-Hanselmann S, Steinbach A, Ullrich V, Prasad S, Blau T, Kebir S, Siveke JT, Becker JC, Sure U, Glas M, et al. Tumor-associated hematopoietic stem and progenitor cells positively linked to glioblastoma progression. Nat Commun. 2021; 12: 3895. doi: 10.1038/s41467-021-23995-z.
41. Berg S, Kutra D, Kroeger T, Straehle CN, Kausler BX, Haubold C, Schiegg M, Ales J, Beier T, Rudy M, Eren K, Cervantes JI, Xu B, et al. ilastik: interactive machine learning for (bio)image analysis. Nat Methods. 2019; 16: 1226-32. doi: 10.1038/s41592-019-0582-9.
42. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A. CellProfiler 4: improvements in speed, utility and usability. BMC Bioinformatics. 2021; 22: 433. doi: 10.1186/s12859-021-04344-9.
43. Uddin I, Woolston A, Peacock T, Joshi K, Ismail M, Ronel T, Husovsky C, Chain B. Quantitative analysis of the T cell receptor repertoire. Methods Enzymol. 2019; 629: 465-92. doi: 10.1016/bs.mie.2019.05.054.
44. Oakes T, Popple AL, Williams J, Best K, Heather JM, Ismail M, Maxwell G, Gellatly N, Dearman RJ, Kimber I, Chain B. The T Cell Response to the Contact Sensitizer Paraphenylenediamine Is Characterized by a Polyclonal Diverse Repertoire of Antigen-Specific Receptors. Front Immunol. 2017; 8: 162. doi: 10.3389/fimmu.2017.00162.