**SUPPLEMENTARY APPENDIX**

**Pembrolizumab in Resectable Human-Papillomavirus-Unrelated Head and Neck Cancer**

Ravindra Uppaluri1,2,\*,#, Katie M. Campbell3,4,\*, Ann Marie Egloff1, Paul Zolkind5, Zachary L. Skidmore4, Brian Nussenbaum5,6, Randal C. Paniello5,6, Jason T. Rich5,6, Ryan Jackson5,6, Patrik Pipkorn5,6, Loren P. Michel6,7, Jessica Ley7, Peter Oppelt6,7, Gavin P. Dunn6,8, Erica K. Barnell3,4, Nicholas C. Spies4, Tianxiang Lin5, Tiantian Li9, David T. Mulder9, Youstina Hanna9, Iulia Cirlan9, Trevor J. Pugh9,10,11, Tenny Mudianto2, Rachel Riley2, Liye Zhou2, Vickie Jo12, Matthew Stachler12, Glenn J. Hanna2, Jason Kass1,2, Robert Haddad1,2, Jonathan D. Schoenfeld2,13, Evisa Gjini12, Ana Lako12, Wade Thorstad6,14, Hiram A. Gay6,14, Mackenzie Daly6,14, Scott J. Rodig112,15, Ian S. Hagemann16, Dorina Kallogjeri5, Jay F. Piccirillo5,6,Rebecca D. Chernock16, Malachi Griffith3,4,6,7, Obi L. Griffith3,4,6,7,#, and Douglas R. Adkins6,7,#

**Affiliations:**

1Department of Surgery, Brigham and Women’s Hospital; Boston, MA, USA

2Department of Medical Oncology, Dana-Farber Cancer Institute; Boston, MA, USA

3Department of Genetics, Washington University School of Medicine; St. Louis, MO, USA

4McDonnell Genome Institute, Washington University School of Medicine; St. Louis, MO, USA

5Department of Otolaryngology, Washington University School of Medicine; St. Louis, MO, USA

6Alvin J. Siteman Cancer Center, Washington University School of Medicine; St. Louis, MO, USA

7Department of Medicine, Division of Medical Oncology, Washington University School of Medicine; St. Louis, MO, USA

8Department of Neurological Surgery, Washington University School of Medicine; St. Louis, MO, USA

9Princess Margaret Cancer Center, University Health Network; Toronoto, Ontario, Canada

10Department of Medical Biophysics, University of Toronto; Toronto, Ontario, Canada

11Ontario Institute for Cancer Research; Toronto, Ontario, Canada

12Department of Pathology, Brigham and Women’s Hospital; Boston, MA, USA

13Department of Radiation-Oncology, Brigham and Women’s Hospital; Boston, MA, USA

14Department of Radiation-Oncology, Washington University School of Medicine; St. Louis, MO, USA

15Center for Immuno-Oncology, Brigham and Women’s Hospital; Boston, MA, USA

16Department of Pathology and Immunology, Washington University School of Medicine; St. Louis, MO, USA

\*Equal contribution, #Co-senior authors

**To whom correspondence should be addressed**:

Dr. Ravindra Uppaluri

Address: 450 Brookline Avenue, Boston, MA, 02215

Telephone: 617-632-3091

E-mail: Ravindra\_Uppaluri@DFCI.Harvard.edu

**Current Affiliations:**

Katie M. Campbell, Department of Medicine, Division of Hematology-Oncology, University of California, Los Angeles; Los Angeles, CA, USA

# 

# 

# SUPPLEMENTARY METHODS

Procedures

Statistical Analysis

Sample acquisition

Sequencing Analyses

Whole Exome Sequencing

RNA Sequencing

T Cell Receptor Sequencing

Variant/neoantigen predictions

Molecular Human papilloma virus assessment

Differential expression analyses

Pathology

PD-L1 expression

Multiplex Immunofluorescence

**SUPPLEMENTARY** **FIGURES**

Supplementary Figure 1. pTR related clinical outcomes

Supplementary Figure 2. Overall and progression-free survival and relapse-free interval by pathology risk category

Supplementary Figure 3. Baseline multiplex IHC tumor immune correlates of pathologic tumor response

Supplementary Figure 4. Immune cell expression patterns in post-treatment tumor RNA

Supplementary Figure 5. Significant changes in T cell receptor diversity between response groups

Supplementary Figure 6. Global summary of the peripheral T cell repertoire

Supplementary Figure 7. T cell clones detected in bulk tumor RNA and peripheral blood

Supplementary Figure 8. Baseline mutational profile does not correlate to pathological response to anti-PD-1 therapy

Supplementary Figure 9. Clonality plots comparing variants in matched baseline and post-treatment tumor DNA

Supplementary Figure 10. Representative tumor regions evaluated in spatial heterogeneity study.

Supplementary Figure 11. Additional patients with multiple post-treatment samples.

**SUPPLEMENTARY TABLES**

Supplementary Table 1. Correlative assays performed by subject

Supplementary Table 2. Adjuvant treatment by pathology risk category

Supplementary Table 3. Perisurgical adverse events by Clavien-Dindo Classification

Supplementary Table 4. Patient clinical stage, pathologic stage and pathologic tumor response

Supplementary Table 5. Patient characteristics by pathologic tumor response status

Supplementary Table 6. Prior institutional HNSCC high risk pathology and downstaging

Supplementary Table 7. Immune cell populations detected by deconvolution

Supplementary Table 8. Differential gene expression analysis

Supplementary Table 9. SNVs and indels detected across all patients

**CLINICAL PROTOCOL**

# 

# SUPPLEMENTARY METHODS

**Procedures**

Patients were monitored for 30 days after surgery for AEs and surgical/wound healing complications. During the administration of adjuvant pembrolizumab, patients were assessed every three weeks beginning with the first dose of pembrolizumab by physical examination, complete blood count, and metabolic panel, and every six weeks by thyroid panel and urinalysis. In these patients, AEs were monitored for 90 days after the last dose of adjuvant pembrolizumab. Patients were monitored for relapse every three months after surgery by physical examination and CT scans.

Two pathologists with head and neck expertise (RDC and IH) independently evaluated all slides from baseline and surgery specimens and quantified pTR in increments of 10%. Primary tumor and lymph node metastases were scored separately. Overall pTR was classified based on the best pTR observed in either primary tumor or lymph node. Joint review consensus was reached when discrepancies occurred.

Secondary endpoints included safety of administration of neoadjuvant pembrolizumab, and clinical tumor response to neoadjuvant pembrolizumab assessed by physical examination and, in some patients, by RECISTv1·1. Correlative endpoints assessed on matched tumor specimens obtained before and after (on day of surgery) neoadjuvant pembrolizumab included PD-L1 expression, histologic, immunologic, genomic, and tumor clonal dynamic changes. T-cell clonality was performed on peripheral blood obtained before and after neoadjuvant pembrolizumab (**Supplementary** **Figure 1**, **Supplementary** **Table 1**).

No major protocol deviations occurred. The protocol was amended six times over the course of the study. Amendments 1-3 included updates of the risk profile and dose modifications of study drug, clarified eligibility criteria, and added Dana-Farber/Brigham and Women’s Cancer Center, Boston, MA as a secondary site. Amendment 4 added correlative studies. Amendment 5 added an unplanned interim analysis after the first 20 patients enrolled into group 1 due to the lower than expected rate of patients with high-risk pathology, and added CT scan of the neck to be performed after neoadjuvant pembrolizumab and prior to surgery. Amendment 6 closed accrual to group 1, and added group 2.

**Statistical Analysis**

Distribution of demographic and clinical characteristics was defined and compared between patients in high-risk and other (low/intermediate-risk) pathology groups. Percent difference and 95% confidence intervals (CI) were calculated for categorical variables; median difference and 95% CI were calculated for continuous variables. Spearman’s rank correlation coefficient was used to estimate correlations between tumor PD-L1 staining and numbers of tumor infiltrating T cells (CD8+ and CD4+) and extent of pTR. Molecular correlates were evaluated for changes across pTR categories using the non-parametric test of trend. Kaplan-Meier estimates of OS, progression-free survival (PFS) and relapse-free survival (RFS) rates and 95% CI by pathology risk category or pTR category were determined and differences between categories assessed using the log rank test. OS was defined as time (months) from day of surgery to death; PFS was defined as time from day of surgery to first disease progression event (new primary, recurrence, distant metastasis or death from disease) or death from any cause; RFS was defined as time from day of surgery to first relapse event (recurrence or distant metastasis). In gene expression analysis, unpaired Mann-Whitney/Wilcoxon rank-sum tests were used to compare groups of patients, and paired Wilcoxon signed-rank tests were used to compare matched baseline and post-treatment tumor samples.

## Sample acquisition

Primary tumor biopsies and blood collection: Each patient underwent an incisional biopsy of the primary tumor site and venipuncture to collect peripheral blood (30 mL). Tissue and blood specimens were immediately processed and used for correlative studies. After neoadjuvant treatment and on the day of surgery, post-treatment blood and tumor was also collected.

Collection for heterogeneity study 10 micron sections of FFPE blocks containing tumor were mounted onto slides, and the first and tenth slides H&E stained. Regions containing tumor were marked by a pathologist (MS), and corresponding regions from unstained slides were collected by scraping under RNAse-free conditions. Samples were treated with deparaffinization solution (Qiagen Cat#939018) and DNA isolated using Reliaprep FFPE gDNA kit (Promega Cat#A2352) according to manufacturer instructions.

#### 

#### Whole Exome Sequencing

Genomic DNA was isolated from tumor and blood samples using the DNeasy Blood and Tissue Kit (Qiagen, Venlo, Netherlands). Whole exome sequencing (WES) library construction and sequencing were performed as previously described, with a few exceptions.[(Griffith, Miller, et al., 2015)](https://paperpile.com/c/D0dMrm/G1K5J) Genomic DNA was fragmented using the Covaris E210 DNA Sonicator (Covaris, WoBurn, MA). Dual indexed whole exome sequencing (WES) libraries were constructed and pooled according to the manufacturer recommendations using one of the following approaches:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Work Order | Reagent | Samples | Captures | Pools | Lanes | Platform | PE Reads |
| 2850994 | IDT xGen Exome Research Panel v1.0 | 10 | 2 | 2 | 2 | HiSeq4000 | 2x150 |
| 2855132 | IDT xGen Exome Research Panel v1.0 | 32 | 3 | 3 | 4 | HiSeq4000 | 2x150 |
| 2854092 | Nimblegen VCRome | 23 | 3 | 3 | 3 | HiSeq4000 | 2x150 |

There was median 84.27X coverage (64.91-105.92X) spanning the exome in blood samples, and 75.56X coverage (26.34-635.52X) in tumor samples. See **eTable 15** for further details.

#### RNA sequencing

Total RNA was isolated from tumor samples using the RNeasy Kit (Qiagen, Venlo, Netherlands). RNA sequencing (RNAseq) libraries (n=38) were prepared using the Illumina TruSeq Stranded Total RNA kit with 500 ng of starting material according to the manufacturer’s recommendations. Libraries were pooled and sequenced using one of the following approaches:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Work Order | Preparation | Index | Samples | Pools | Lanes | Platform | PE Reads |
| 2850995 | Illumina TruSeq | Single | 18 | 1 | 8 | HiSeq4000 | 2x150 |
| 2854121 | Illumina TruSeq | Single | 13 | 1 | 6 | HiSeq4000 V4 1Tb | 2x125 |
| 2855138 | Illumina TruSeq | Dual | 7 | 1 | 5 | HiSeq4000 | 2x150 |

Samples that had low amounts or poor quality RNA (n=7) were evaluated using cDNA hybridized capture sequencing using one of the following approaches:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Work Order | Preparation | Reagent | Index | Samples | Captures | Pools | Lanes | Platform | PE Reads |
| 2854122 | Illumina Truseq | IDT xGen Exome Research Panel v1.0 | Dual | 5 | 1 | 1 | 1 | HiSeq4000 V4 1Tb | 2x125 |
| 2855140 | Illumina Truseq | Nimblegen VCRome | Dual | 2 | 1 | 1 | 5 | HiSeq4000 | 2x150 |

RNA seq libraries were sequenced to a median depth of 347,311,447 reads (276,725,868-469,020,578); cDNA capture sequencing libraries were sequenced to a median depth of 165,001,122 reads (17,058,178-882,394,752). See **eTable 16** for further details.

#### 

#### T cell receptor sequencing

Genomic DNA was isolated from blood samples (n=50) using the DNeasy Blood and Tissue Kit (Qiagen, Venlo, Netherlands). Libraries were prepared using a hybridized capture designed to enrich for rearranged T cell receptor gene loci (*TRA, TRB*) using CapTCR sequencing, as previously described.[(Mulder et al., 2018)](https://paperpile.com/c/D0dMrm/t5Cu0) Single indexed libraries were pooled and sequenced on a single lane of the Illumina HiSeq 4000 platform with 2x150 bp reads to a depth of 2,423-28,107,860 (median 6,379,349 reads). See **eTable 17** for further details

### 

### Somatic variant detection

The Genome Modeling System (GMS) was used for all analysis, including the somatic variant detection and RNAseq analysis.[(Griffith, Griffith, et al., 2015)](https://paperpile.com/c/D0dMrm/zM3t2) Briefly, WES data was processed through SpeedSeq v0.1.0[(Chiang et al., 2015; Griffith, Griffith, et al., 2015)](https://paperpile.com/c/D0dMrm/zM3t2+beoBu) which uses BWA-MEM v0.7.10[(Heng Li, 2013)](https://paperpile.com/c/D0dMrm/tsZqD) alignment to the human reference genome (NCBI build 38, GRCh38) and marks duplicates using SAMBLASTER v0·1·22.[(Chiang et al., 2015; Faust & Hall, 2014; Griffith, Griffith, et al., 2015)](https://paperpile.com/c/D0dMrm/zM3t2+beoBu+2O2Sc) Single nucleotide variants (SNVs) were detected by comparing tumors to matched blood WES by SomaticSniper v1.0.4,[(Chiang et al., 2015; Faust & Hall, 2014; Griffith, Griffith, et al., 2015; Larson et al., 2014)](https://paperpile.com/c/D0dMrm/zM3t2+beoBu+2O2Sc+RdWuC) VarScan2 v2.3.6,[(Chiang et al., 2015; Faust & Hall, 2014; Griffith, Griffith, et al., 2015; Larson et al., 2014; Reble et al., 2017)](https://paperpile.com/c/D0dMrm/zM3t2+beoBu+2O2Sc+RdWuC+V2AT9) Strelka v1.0.11,[(Saunders et al., 2012)](https://paperpile.com/c/D0dMrm/Rstne) Samtools r982,[(H. Li et al., 2009)](https://paperpile.com/c/D0dMrm/SmfgI) and Mutect v1.1.4.[(Cibulskis et al., 2013)](https://paperpile.com/c/D0dMrm/m4eU2) Small insertions and deletions (indels) were detected by GATK v5336,[(Huang, 2018)](https://paperpile.com/c/D0dMrm/pFwR4) VarScan2,[(Reble et al., 2017)](https://paperpile.com/c/D0dMrm/V2AT9) Strelka,[(Saunders et al., 2012)](https://paperpile.com/c/D0dMrm/Rstne) and Mutect. Variants were annotated using the GMS transcript variant annotator against Ensembl v79 and VEP. False positives were further removed by manual reviewing all filtered, nonsynonymous SNVs and indels, according to previously defined guidelines.[(Barnell et al., 2018)](https://paperpile.com/c/D0dMrm/Xw0VP) Somatic mutations and alterations were detected in individual tumor samples; the union of variants in matched baseline/post-treatment samples were interrogated to recover low-frequency or low-coverage variants. Somatic copy number alterations (CNAs) were detected by CopyCat v0.1 [<https://github.com/chrisamiller/copyCat>], and structural variations (SVs) were detected by Manta v0.29.6.[(Chen et al., 2016; Saunders et al., 2012)](https://paperpile.com/c/D0dMrm/Rstne+fGw1j)

#### 

#### Targeted Capture Sequencing

A custom capture reagent (NimbleGen SeqCap EZ Choice) was designed, targeting all SNVs and Indels detected by WES across pre- and post-treatment tumor tissue. WES libraries (including 25 normal blood, 24 pre-treatment tumor tissue, and 23 post-treatment tumor tissue samples) were pooled and captured across one lane on the Illumina Novaseq platform (S4 Flow Cell). Additional FFPE samples acquired from post-treatment surgical specimen slides were pooled and captured across one lane on the Illumina NovaSeq platform (S4 Flow Cell). There was median 658.02X coverage (224.94-1139.76X) spanning the targeted capture region in blood samples, and 651.88X coverage (289.38-1108.58X) in tumor samples. See e**Table 13** for further details.

HLA typing for MHC Class I and II alleles were performed by xHLA [(Xie et al., 2017)](https://paperpile.com/c/D0dMrm/xH1t7). Class I and II peptide predictions were performed using all peptide lengths (8-11 for Class I and 13-17 for Class II) and all prediction algorithms available in the pVACtools pipeline, with each patient’s corresponding HLA types and somatic mutations.[(Hundal et al., 2016)](https://paperpile.com/c/D0dMrm/dHuYG) Mutant peptides were included in subsequent analysis as “neoantigens” if they had a predicted binding affinity to at least one corresponding HLA allele of less than 500nM. Neoantigens were further filtered by DNA variant allele fraction (VAF>5%) and expression. If matched RNA was available, neoantigens were filtered by variant expression (RNA VAF>5%) and overall gene expression (RNA FPKM>1). If matched RNA was not available, neoantigens were filtered by predicted gene expression by the tumor, where the average expression of the gene was summarized across all baseline RNAseq data (n=19; average FPKM>1).

### 

### Gene expression

RNA reads were aligned to GRCh38 by TopHat v2.0.8. Gene and transcript expression levels were quantified using Cufflinks v2.1.1 (FPKM)[(Chen et al., 2016)](https://paperpile.com/c/D0dMrm/fGw1j) and HTSeq-count v0.5.4p1 (raw counts).[(Anders et al., 2015)](https://paperpile.com/c/D0dMrm/HgtRz) Samples sequenced using cDNA capture sequencing were not used in gene expression analysis due to their lower input quality. Differential expression analysis was performed using the DESeq2 R package[(Love et al., 2014)](https://paperpile.com/c/D0dMrm/j0NWp) on gene raw counts, and gene set enrichment (GSEA) analysis was performed by the fgsea R package.[(Sergushichev, 2016)](https://paperpile.com/c/D0dMrm/kuEKP) Deconvolution of infiltrating immune populations was estimated on each tumor RNA sample by CIBERSORT[(Newman et al., 2015)](https://paperpile.com/c/D0dMrm/QokOp). CIBERSORT was run in Relative and Absolute modes, with 1000 permutations against the reference LM22 mixtures file. Quantile normalization was disabled. Statistics were performed on the Absolute abundances. Unpaired Mann-Whitney/Wilcoxon rank-sum tests were used to compare groups of patients, and paired Wilcoxon signed-rank tests were used to compare matched baseline and post-treatment tumor samples.

## 

## T cell receptor (TCR) analysis

Reads were aligned to the TCR gene loci and assembled using MixCr [2.1.11], with the following parameters: mixcr align --force --species HomoSapiens -OvParameters.geneFeatureToAlign=VGene --chains TCR. The number of successfully aligned reads ranged from 57-935,780 (median 467,569), for an overall alignment rate of 1.35-11.32% (median 7·43%). MixCr was used to assemble and export clones, filtered to exclude out-of-frame and truncating clones (containing stop codons). Clonality and diversity calculations were generated, as previously described, on reads associated with functional T cell clonotypes.[(Zhang et al., 2017)](https://paperpile.com/c/D0dMrm/nHFWQ) These calculations were individually performed on alpha and beta clones.

## 

## PD-L1 Immunohistochemistry

IHC using clone 9A11 mouse anti-human PD-L1 monoclonal antibody (IgG1, 10.4 μg/ml final concentration) was performed on a Bond III automated staining system (Leica Biosystems) following the manufacturer’s protocol as previously described.[(Howitt et al., 2016)](https://paperpile.com/c/D0dMrm/74fd) Heat-induced antigen retrieval was performed using Epitope Retrieval Solution 2 (ER2) (pH8) (Leica) for 30 minutes. Primary antibody was incubated for total of 2 hours with two separate applications, follow by 8 minutes of post-primary blocking reagent, 12 minutes of horseradish peroxidase-labeled polymer, 5 minutes of peroxidase block, 15 minutes of 3,3’-diaminobenzidine (DAB) developing, and 10 minutes of hematoxylin staining. Stained slides were scored by a pathologist (SR) for percentage of malignant cells with positive staining over background (0-100%) and the average intensity of staining (0= no staining, 1=weak staining above background, 2=moderate staining, 3=strong staining) without knowledge of the genetic status of the tumors. A modified H-score was generated by multiplying percentage of malignant cells with positive staining and the average intensity of positive staining (0-300).

**Multiplex Immunofluorescence Staining**  
Multiplex Immunofluorescent staining was performed overnight on BOND RX fully automated stainers (Leica Biosystems) as previously described.[(Carey et al., 2017)](https://paperpile.com/c/D0dMrm/HXigH) Briefly, tissue sections of 5-μm thick FFPE were baked for 3 hours at 60°C before loading into the BOND RX (Leica Biosystems). Slides were deparaffinized using BOND DeWax Solution (Leica Biosystems) and then rehydrated with series of graded ethanol to deionized water. Antigen retrieval was performed in BOND Epitope Retrieval Solution 1 (ER1, Leica Biosystems) at pH 6 for 10 minutes at 98°C. Deparaffinization, rehydration and antigen retrieval were all preprogrammed and executed by the BOND RX (Leica Biosystems). Next, slides were serially stained with primary antibodies, such as anti-CD8 (clone C8/144B; DAKO, dilution 1:5000). Incubation time per primary antibody was 40 minutes. Subsequently, anti-rabbit Polymeric Horseradish Peroxidase (Poly-HRP, BOND Polymer Refine Detection Kit, Leica Biosystems) was applied as a secondary label with an incubation time of 10 minutes. Signal for antibody complexes was labeled and visualized by their corresponding Opal Fluorophore Reagents (PerkinElmer) by incubating the slides for 5 minutes. The same process was repeated for the following antibodies / fluorescent dyes. Slides were air dried, mounted with Prolong Diamond Anti-fade mounting medium (#P36965, Life Technologies) and stored in a light-proof box at 4 ̊C prior to imaging. Image acquisition and cell identification was performed at 20x using the Mantra multispectral imaging platform (Vectra 3, PerkinElmer, Hopkinton, MA). The target antigens, antibody clones, and dilutions for markers included in this report and details of controls are listed in the Table below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Primary Antibody | Clone ID/Company | Dilution | Opal Kit Fluor | Opal Fluor Dilution |
| Anti-CD8 | C8/144B, DAKO | 1:5,000 | Opal 520 | 1:100 |
| Anti-human CD274 (PD-L1) | 405·9A11, Cell Signaling Technologies | 1:500 | Opal 540 | 1:200 |
| Anti-CD4 | 4B12, DAKO | 1:250 | Opal 540 | 1:200 |
| Anti-PD-1 | EH33, Cell Signaling Technologies | 1:11,000 | Opal 620 | 1:200 |
| Cytokeratin | AEI/AE3/ DAKO | 1:2000 | Opal 690 | 1:50 |

**Multiplex Image Acquisition and Analysis**  
Image acquisition was performed using the Mantra multispectral imaging platform (Vectra 3, PerkinElmer, Hopkinton, MA) as previously described.[(Carey et al., 2017)](https://paperpile.com/c/D0dMrm/HXigH) Areas with non-tumor or residual normal tissue (i.e. residual lymph node) were excluded from the analysis. Representative regions of interest were chosen by the pathologist, and 3-5 fields of view (FOVs) were acquired at 20x resolution as multispectral images. Image Analysis was performed using the Inform 2.3 Image Analysis Software (PerkinElmer, Hopkinton, MA).   
  
**Multiplex Immunofluorescence Cell Identification**  
Cell identification was performed as described.[(Carey et al., 2017)](https://paperpile.com/c/D0dMrm/HXigH) In short, after image capture, the FOVs were spectrally unmixed and then analyzed using supervised machine learning algorithms within Inform 2·3 (PerkinElmer). The Inform 2·3 image analysis software assigns phenotypes to all cells in the image, based on a combination of immunofluorescence characteristics associated with segmented nuclei (DAPI signal). Each cell-phenotype specific algorithm is based upon an iterative training / test process, whereby a small number of cells (training phase, typically 20-30 cells) are manually selected as being most representative of each phenotype of interest and the algorithm then predicts the phenotype for all remaining cells (testing phase). The pathologist can over-rule the decisions made by the software to improve accuracy, until phenotyping is optimized. Thresholds for "positive" staining and the accuracy of phenotypic algorithms were optimized and confirmed by the pathologist (SR) for each case.

## 

## Statistical analysis

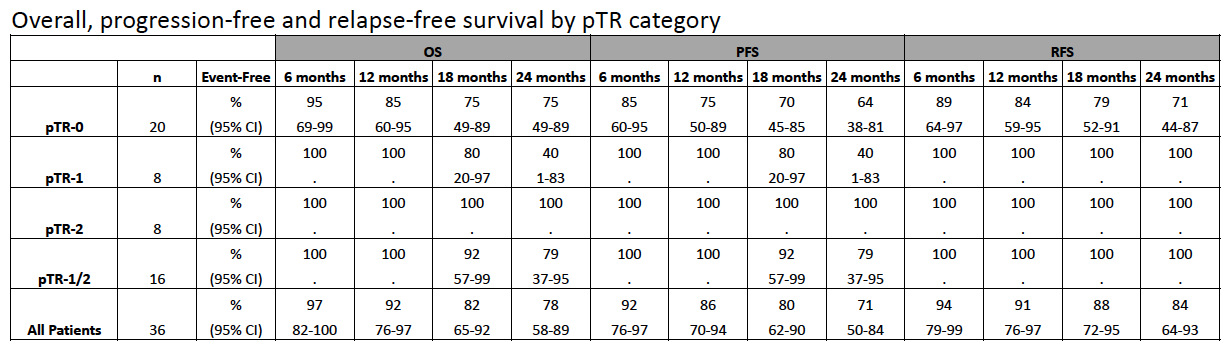
Sequencing data analysis and visualization were performed in R v.3.3.2 using the ggplot2 R package v2.2.1[(Wickham, 2016)](https://paperpile.com/c/D0dMrm/DHLYm) and GenVisR v1.8.0.[(Skidmore et al., 2016)](https://paperpile.com/c/D0dMrm/ikPhO) For box-plots, center line indicates median, box limits are upper and lower quartiles, whiskers indicate 1.5 x interquartile range and outliers are presented.

# Supplementary Figures

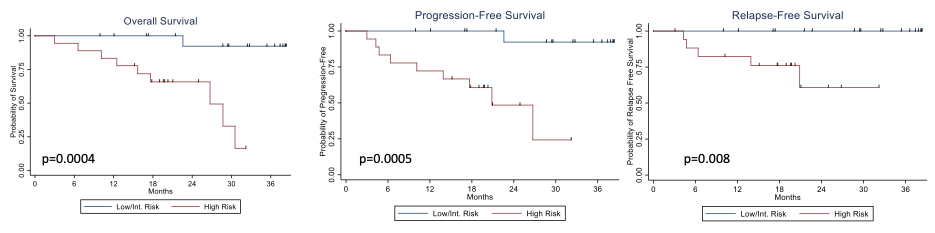
# 

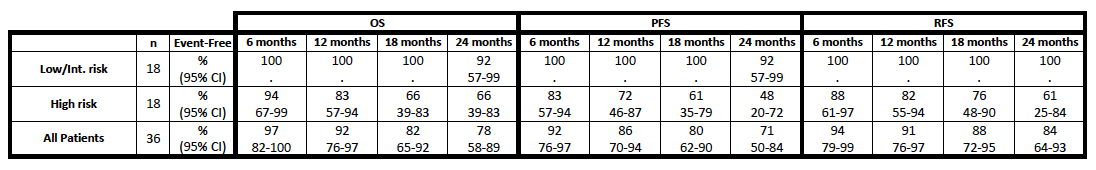
# 

# 



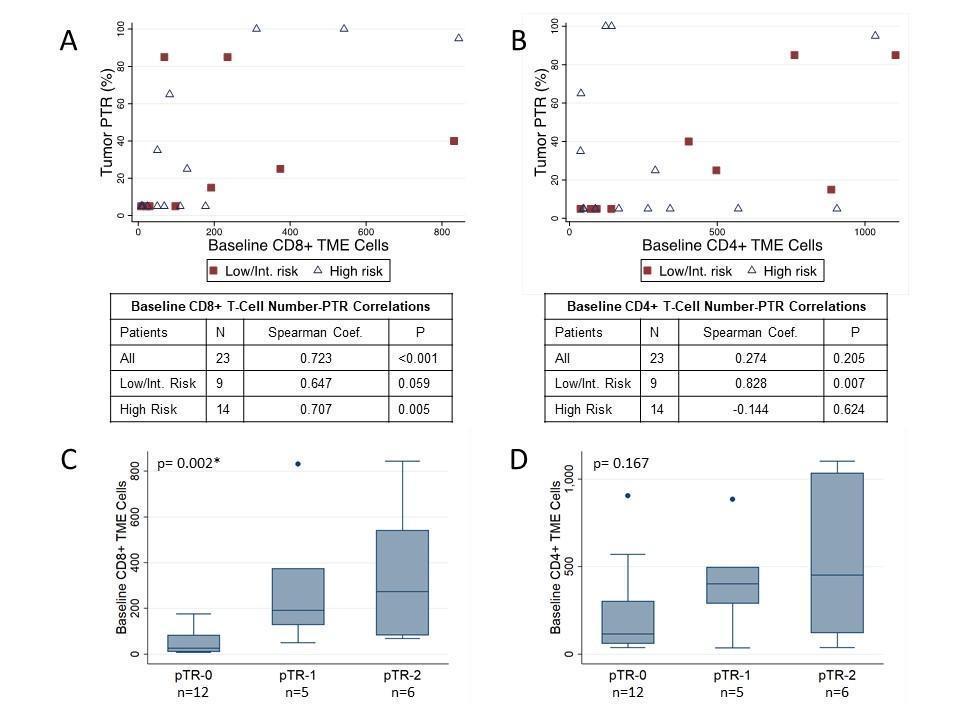
# Supplementary Figure 1. pTR related clinical outcomes (A) Kaplan-Meier (KM) curves for overall survival (OS), progression-free survival (PFS) and relapse-free survival (RFS) by pTR category (pTR-0, pTR-1, and pTR-2) suggested better survival among patients with pTR-2 but did not reach statistical significance (Log rank p-values comparing 3 groups is shown). (B) K-M curves for comparing pTR-0 to pTR-1/2 shows association of any PTR with improved RFS.





## Supplementary Figure 2. Overall, progression-free and relapse-free survival by pathology risk category

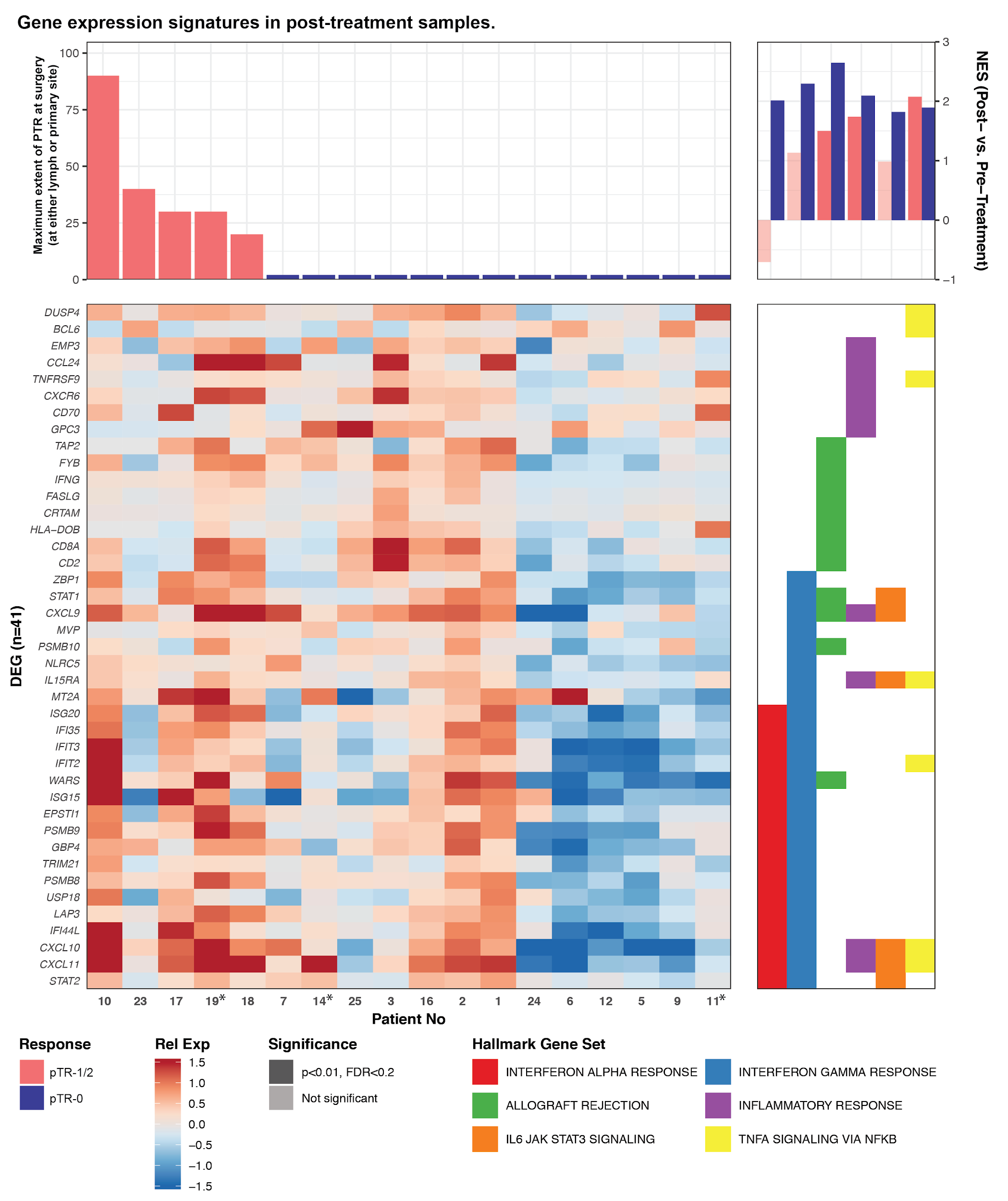
High-risk pathology defined as extranodal extension (ENE) and/or positive margins. Low/Intermediate-risk cases lacked ENE and had negative margins. Log rank p-values are provided.



## Supplementary Figure 3. Baseline immune multiplex IF correlates of pathologic tumor response.

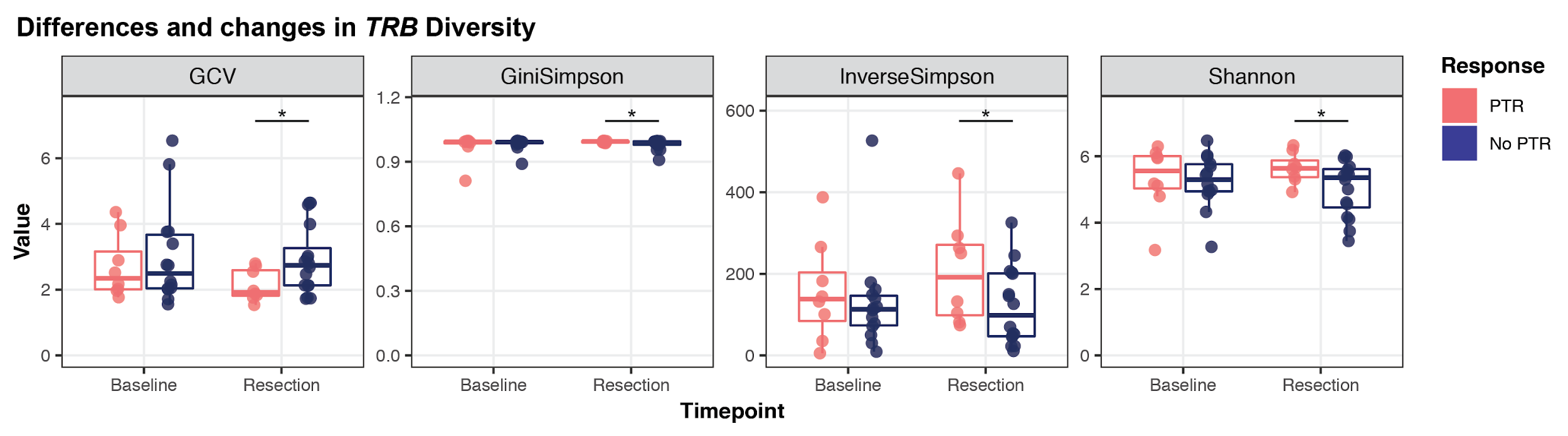
(A) Numbers of CD8+ T cells and (B) CD4+ T cells per mm2 in the baseline biopsy tumor microenvironment (TME) were evaluated separately for correlation with pTR. The maximum pTR, whether detected in the primary tumor, lymph node or both, was used for the analyses, which were performed using for subjects as well as stratified by pathology risk category. Spearman correlation coefficient and associated p values are provided. (C) Numbers of CD8+ and (D) CD4+ T cells in the baseline biopsy TME were evaluated by pTR category. Nonparametric test of trend p values are provided.

## 

****

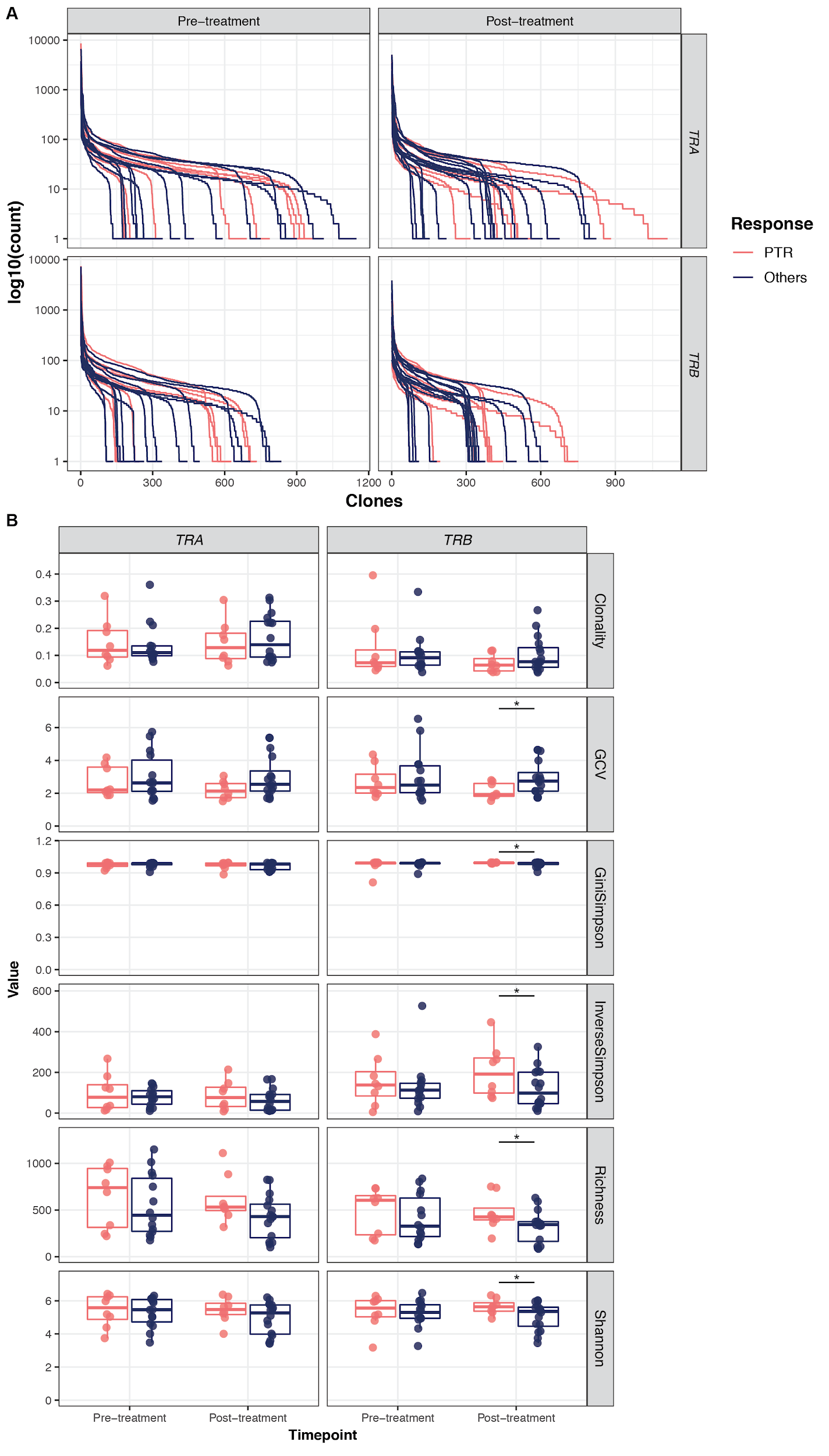
**Supplementary** **Figure 4. Immune cell expression patterns in post-treatment tumor RNA**

This heatmap shows genes (n=41) associated with hallmark gene sets (right panel) that were differentially expressed at baseline (p<0·01, FDR<0·25) between patients with pTR-1/2 and those with pTR-0. Genes are sorted in the order shown in **Figure 3**. Patients are sorted by decreasing maximum pTR (at either the tumor or lymph node site), then by Ward’s hierarchical clustering. Patients who did not have matched baseline samples are indicated by an asterisk by the Patient No (\*, x-axis). Expression is displayed as the gene-normalized expression across all RNA sequencing samples. The Normalized Enrichment Score (NES) depicted in the top right panel indicates the enrichment of hallmark gene sets by comparing matched Post-treatment to Pre-treatment RNA in either patients with pTR (pink) or no pTR (blue). The opacity of the bar indicates that the comparison was statistically significant (p<0·01, FDR<0·2). Positive values indicate that expression of the corresponding gene sets was enriched post-treatment; negative values indicate that expression of the corresponding gene set was enriched at baseline.



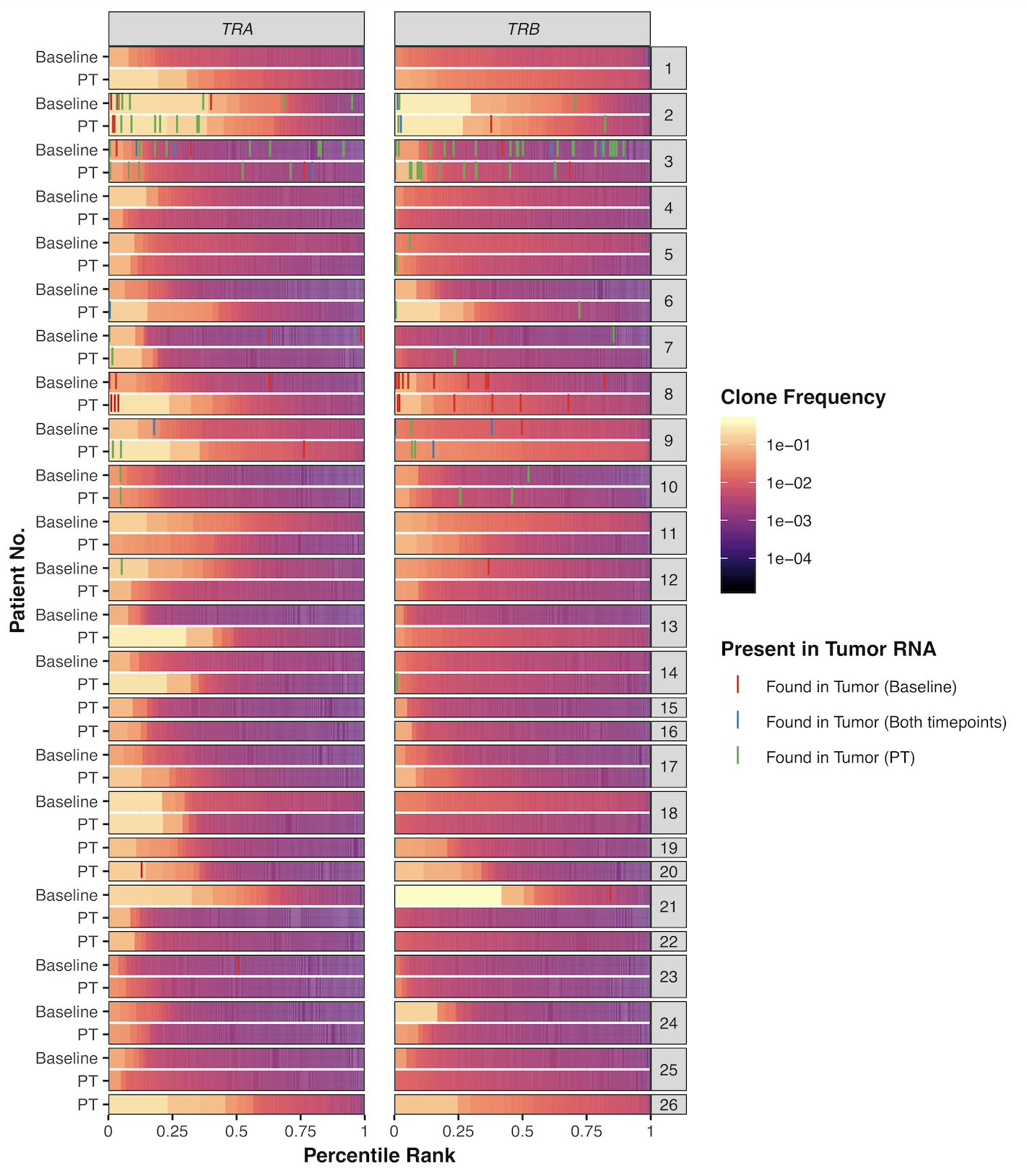
## Supplementary Figure 5. Significant changes in T cell receptor diversity between response groups

The geometric coefficient of variance (GCV) and the Gini Simpson, Inverse Simpson, and Shannon Diversity indices were used to quantify the TRB clones detected by TCR sequencing in the peripheral blood DNA. Wilcoxon tests were used to evaluate statistical significance across responder groups and timepoints; \* indicates p<0.05.

****

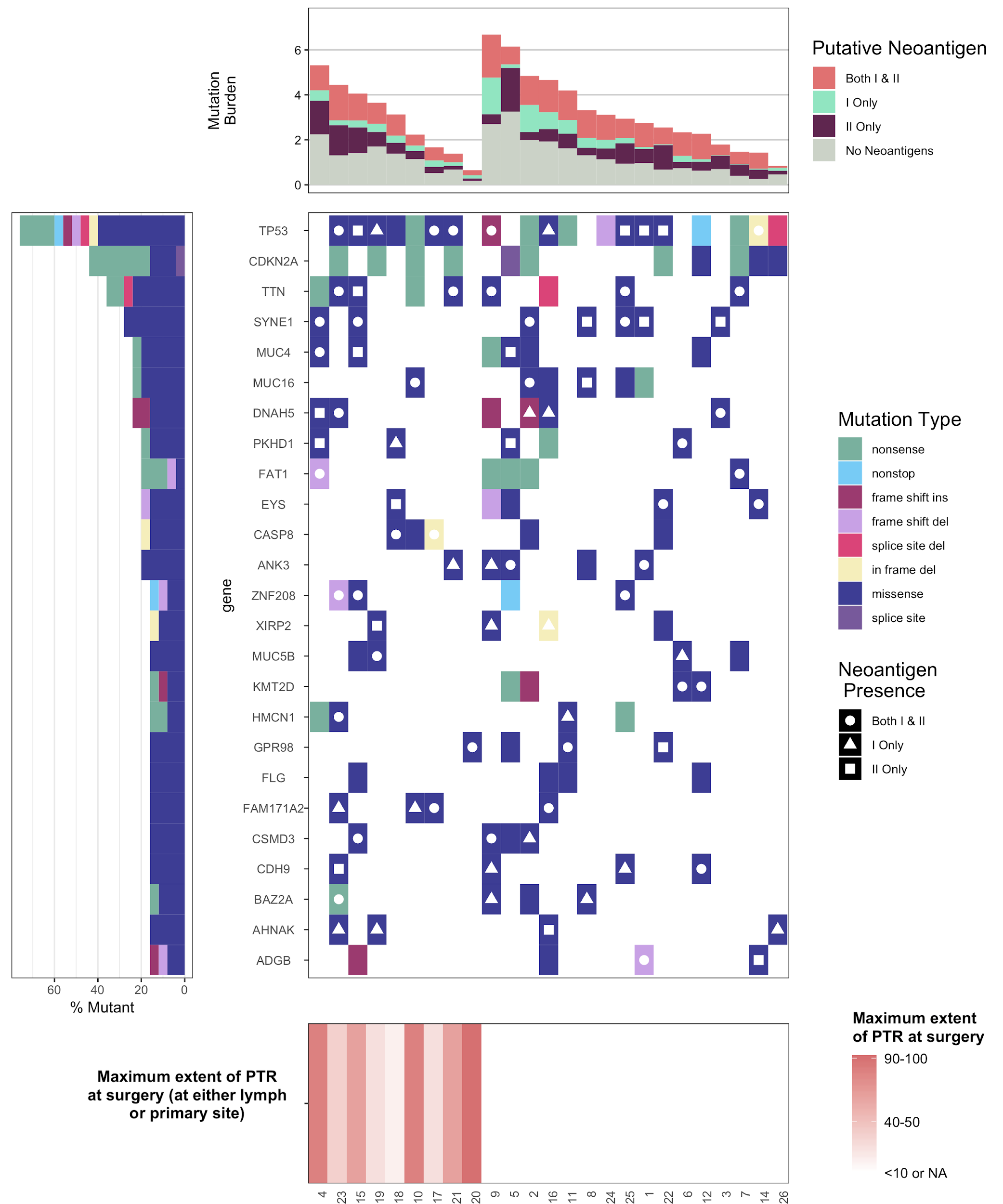
## Supplementary Figure 6. Global summary of the peripheral T cell repertoire

The T cell repertoire was assessed in DNA derived from the peripheral blood using CapTCR sequencing. A. Count distribution of unique TCR clones detected in each peripheral blood sample, corresponding to either the *TRA* (top panel) or *TRB* (bottom panel) gene. Line color indicates whether the patient showed pTR following neoadjuvant pembrolizumab (pink) or not (dark blue). B. Reads were divided into those corresponding to either *TRA* (left panel) or *TRB* (right panel) to quantify characteristics of the peripheral T cell repertoire. The calculations for clonality, geometric coefficient of variance (GCV), Gini Simpson Index, Inverse Simpson Index, richness (total number of clones detected), and Shannon Diversity Index were performed as previously described, and are provided in the Supplementary Methods. Wilcoxon signed rank tests were used to compare metrics across (paired) timepoints, within each response group. Mann-Whitney tests were used to compare the metrics across response groups, within each timepoint. \* indicates comparisons with p<0·05.

****

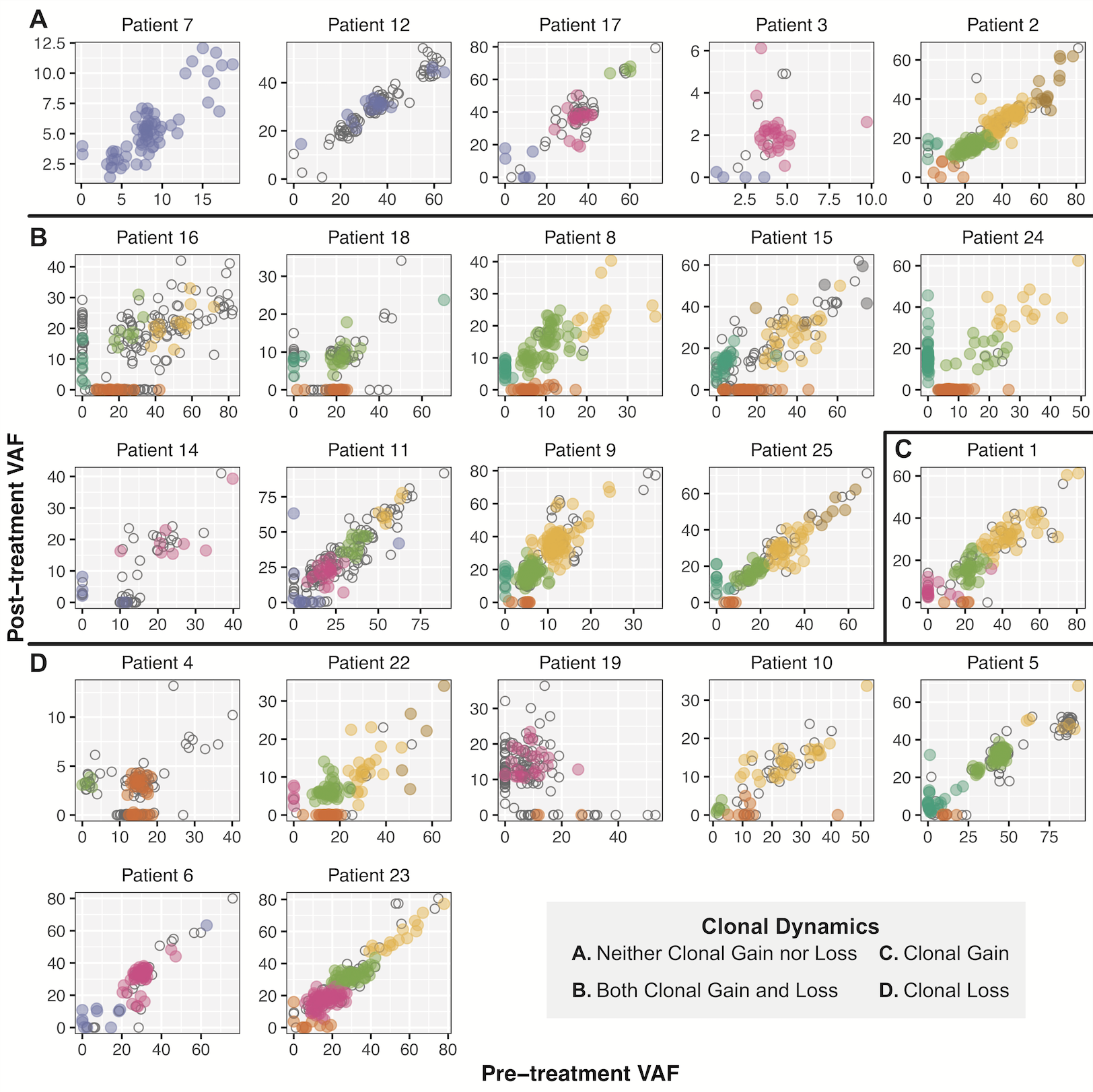
**Supplementary Figure 7. T cell clones detected in bulk tumor RNA and peripheral blood**

T cell clones detected in the peripheral blood DNA were ordered by decreasing rank based upon their frequency, with respect to each sample and gene (TRA, TRB). They are indicated by their percentile rank (x-axis) and colored by their frequency. Samples are indicated by patient number (y-axis, right-side labels) and their corresponding timepoint (y-axis, left-side labels). Clones that were detected in bulk tumor RNA and also detected in the peripheral blood DNA, are shown by lines at their corresponding percentile rank. The color of these lines designate whether the clone was detected at only one or both timepoints, in the tumor RNA.



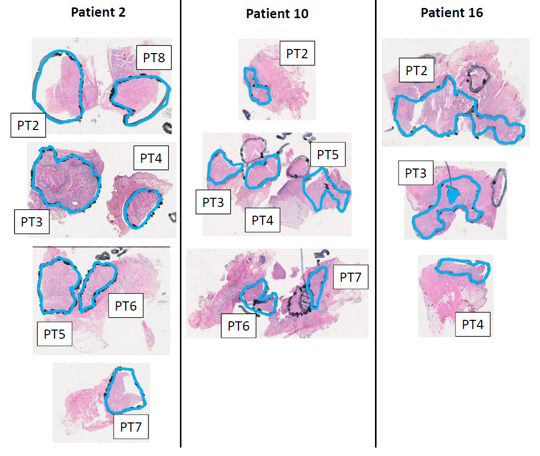
**Supplementary Figure 8. Baseline mutational profile does not correlate to pathological response to anti-PD-1 therapy**

Bar plot showing pathologic tumor response (lower panel), bar plot showing tumor mutational burden (upper panel), waterfall showing recurrently mutated genes (middle panel). Genes were filtered to those that were altered in at least 16% (4/25) of patients. Of note, none of these genes were significantly mutated across the cohort, due to the low number of patients. Mutations that have a putative neoantigen <500 nM (either a Class I, Class II, or both) are indicated. Mutations are defined whether they have a Class I, Class II, or at least one of each Class of neoantigen. Individuals are sorted by decreasing pathologic response (at the primary site, followed by the lymph node), and by decreasing mutational burden in the remainder of the cohort.



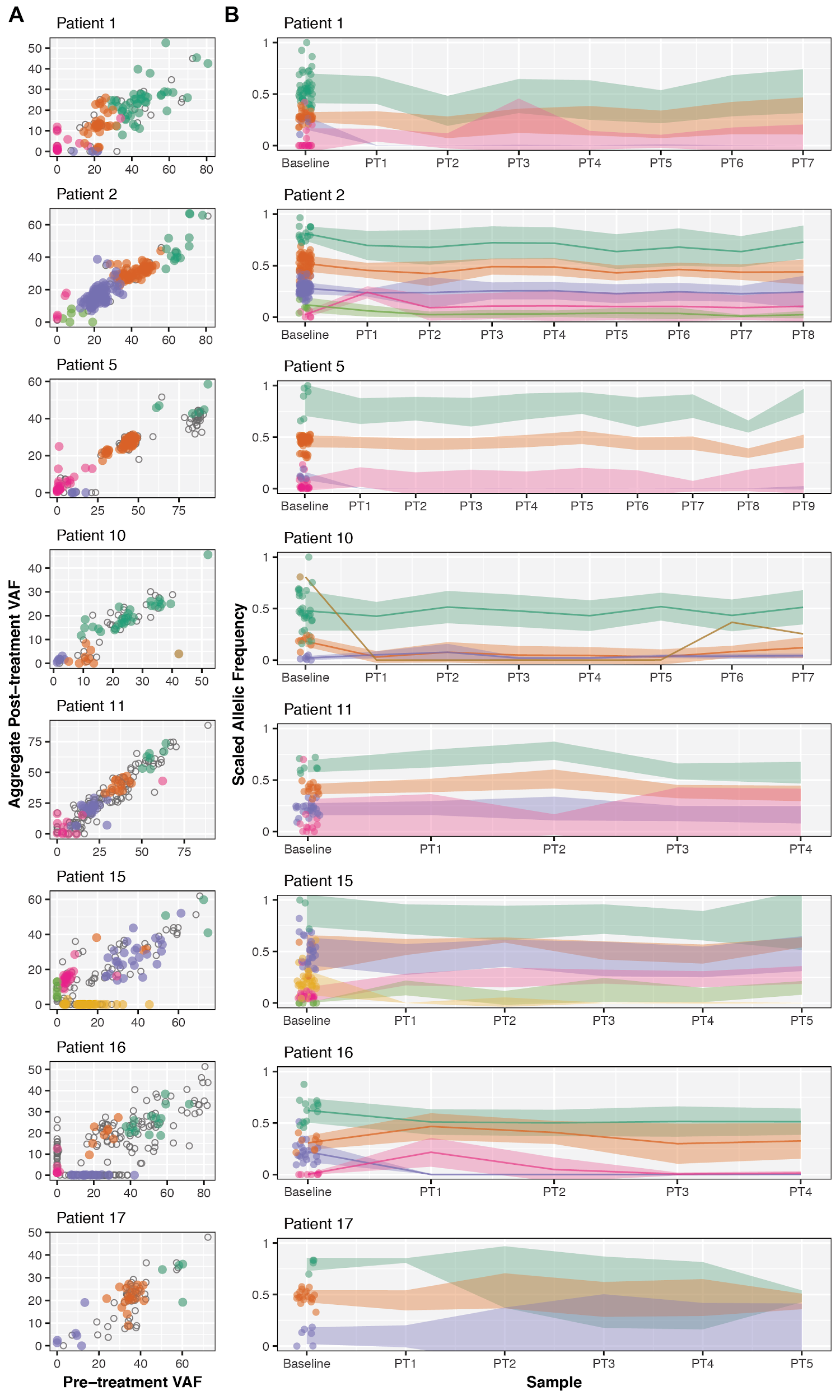
**Supplementary Figure 9. Clonality plots comparing variants in matched baseline and post-treatment tumor DNA**

Matched DNA samples were available for 22 patients. The variant allele fraction (VAF) reflects the aggregate metric determined by the combination of whole exome and high-depth (targeted) sequencing data in pre-treatment tumor biopsies (x-axis) and post-treatment surgical resections (y-axis). Variant clustering was performed using the combination of exome and high-depth (targeted) sequencing data by SciClone, and each color indicates a cluster of variants within each patient (i.e. similar colors across patients are not related). Clonal gain (variants commonly colored in turquoise) was defined by at least five variants that were detected in post-treatment, but not pre-treatment tumor DNA. Clonal loss (variants commonly colored in orange) was defined by at least five variants that were detected in pre-treatment, but not post-treatment tumor DNA. Variants in regions containing copy number alterations or loss-of-heterozygosity were excluded from clustering (open circles). Patients are organized by observed clonal dynamics: A. Neither clonal gain nor loss, B. Both clonal gain and loss, C. Clonal gain only, and D. Clonal loss only. Patients 20, 21, and 26 are not included because only one timepoint was assessed by DNA sequencing.



**Supplementary Figure 10. Representative tumor regions evaluated in spatial heterogeneity study.**

Non-overlapping regions, demarcated in blue, from surgically resected tumor from Patients 2,10 and 16 were individually evaluated in spatial heterogeneity sequencing analyses.



**Supplementary Figure 11. Additional patients with multiple post-treatment samples.**

A) Clonality plots comparing VAF of SNVs and Indels at baseline (x-axis) and resection (y-axis). These values represent aggregate metrics of all post-treatment samples available, and colors represent variants that clustered together by comparing these aggregate metrics per individual. Open circles represent variants that were either unassigned or were assigned to clusters with less than 5 variants. B) The scaled allelic frequency (AF) is depicted on the y-axis of each variant (point) that was assigned a cluster in (A). The solid line represents the average scaled AF of the cluster. Shaded ribbons represent the standard error from the mean of the scaled AF of the cluster.

# 

# SUPPLEMENTARY TABLES

## 

## Supplementary Table 1. Correlative assays performed by subject

WES, Whole exome sequencing. RNAseq indicates either RNA sequencing or cDNA capture sequencing. (X = assay performed, NA = assay not performed)

## Supplementary Table 2. Adjuvant treatment by pathology risk category

## 

## Supplementary Table 3. Perisurgical adverse events by Clavien-Dindo Classification

## 

## Supplementary Table 4. Patient clinical and pathologic staging and pathologic tumor response

Patients were clinically staged (TNM) prior to treatment and pathologically staged at the time of surgical resection. Quantification of treatment effect is defined by the percentage of sample containing keratinous debris, necrotic tumor, and histiocytic reactions. Surgical specimens defined as having PTR had pathologic findings of tumor cell necrosis and keratinous debris with giant cell/histiocytic reaction present in the tumor bed of surgery specimens but not in pretreatment biopsy specimen. PTR was scored for the primary tumor and lymph node resection at the first surgery (13-22 days post-treatment). Downstaging is defined by reduced TNM staging over the course of treatment, comparing baseline clinical to post-treatment pathologic staging.

## 

## Supplementary Table 5. Patient characteristics by pathologic tumor response status

## 

## Supplementary Table 6. Comparison of trial patients with prior institutional HNSCC cohort

## Supplementary Table 7. Immune cell populations detected by deconvolution (Absolute scores)

CIBERSORT was used to perform immune cell deconvolution on RNA sequencing data. CIBERSORT was run in Relative and Absolute modes, with 1000 permutations against the reference LM22 mixtures file. Quantile normalization was disabled. This table includes the Absolute scores associated with each immune cell population. Absolute values were also used for statistical analysis and visual display in the main text.

## 

## Supplementary Table 8. Differential gene expression analysis

Differential gene expression analysis was performed using the R package DESeq2 across the following comparisons: (Columns B-D) Pre-treatment tumor biopsies in patients with pTR (n=6) vs. those without pTR (n=10), (Columns E-G) Matched post-treatment vs. pre-treatment tumor samples in patients with pTR (n=4), and (Columns H-J) Matched post-treatment vs. pre-treatment tumor samples in patients without pTR (n=10). Genes (Column A) that are shown are those that were significantly differentially expressed (p<0·01, padj<0·2) in at least one of the comparisons.

## 

## Supplementary Table 9. SNVs and indels detected across all patients (Exome sequencing)

This table includes all nonsilent SNVs and Indels detected within each patient (Columns A-B), the coordinates and annotation (Columns C-N), and the presence of the variant across samples (Columns O-Z). The last column, ‘category,’ indicates whether there was at least one Class I and/or Class II neoantigen predicted for the corresponding patient, based upon their predicted HLA type.

## 

# REFERENCES (SUPPLEMENTARY APPENDIX)

1. [Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* , *31*(2), 166–169.](http://paperpile.com/b/D0dMrm/HgtRz)
2. [Barnell, E. K., Ronning, P., Campbell, K. M., Krysiak, K., Ainscough, B. J., Ramirez, C., Spies, N., Kunisaki, J., Hundal, J., Skidmore, Z. L., Gomez, F., Trani, L., Matlock, M., Wagner, A. H., Joshua Swamidass, S., Griffith, M., & Griffith, O. L. (2018). Standard operating procedure for somatic variant refinement of tumor sequencing data. In *bioRxiv* (p. 266262). https://doi.org/](http://paperpile.com/b/D0dMrm/Xw0VP)[10.1101/266262](http://dx.doi.org/10.1101/266262)
3. [Carey, C. D., Gusenleitner, D., Lipschitz, M., Roemer, M. G. M., Stack, E. C., Gjini, E., Hu, X., Redd, R., Freeman, G. J., Neuberg, D., Hodi, F. S., Liu, X. S., Shipp, M. A., & Rodig, S. J. (2017). Topological analysis reveals a PD-L1-associated microenvironmental niche for Reed-Sternberg cells in Hodgkin lymphoma. *Blood*, *130*(22), 2420–2430.](http://paperpile.com/b/D0dMrm/HXigH)
4. [Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A. J., Kruglyak, S., & Saunders, C. T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* , *32*(8), 1220–1222.](http://paperpile.com/b/D0dMrm/fGw1j)
5. [Chiang, C., Layer, R. M., Faust, G. G., Lindberg, M. R., Rose, D. B., Garrison, E. P., Marth, G. T., Quinlan, A. R., & Hall, I. M. (2015). SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nature Methods*, *12*(10), 966–968.](http://paperpile.com/b/D0dMrm/beoBu)
6. [Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E. S., & Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnology*, *31*(3), 213–219.](http://paperpile.com/b/D0dMrm/m4eU2)
7. [Faust, G. G., & Hall, I. M. (2014). SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics* , *30*(17), 2503–2505.](http://paperpile.com/b/D0dMrm/2O2Sc)
8. [Griffith, M., Griffith, O. L., Smith, S. M., Ramu, A., Callaway, M. B., Brummett, A. M., Kiwala, M. J., Coffman, A. C., Regier, A. A., Oberkfell, B. J., Sanderson, G. E., Mooney, T. P., Nutter, N. G., Belter, E. A., Du, F., Long, R. L., Abbott, T. E., Ferguson, I. T., Morton, D. L., … Wilson, R. K. (2015). Genome Modeling System: A Knowledge Management Platform for Genomics. *PLoS Computational Biology*, *11*(7), e1004274.](http://paperpile.com/b/D0dMrm/zM3t2)
9. [Griffith, M., Miller, C. A., Griffith, O. L., Krysiak, K., Skidmore, Z. L., Ramu, A., Walker, J. R., Dang, H. X., Trani, L., Larson, D. E., Demeter, R. T., Wendl, M. C., McMichael, J. F., Austin, R. E., Magrini, V., McGrath, S. D., Ly, A., Kulkarni, S., Cordes, M. G., … Wilson, R. K. (2015). Optimizing Cancer Genome Sequencing and Analysis. *Cell Systems*, *1*(3), 210–223.](http://paperpile.com/b/D0dMrm/G1K5J)
10. [Howitt, B. E., Sun, H. H., Roemer, M. G. M., Kelley, A., Chapuy, B., Aviki, E., Pak, C., Connelly, C., Gjini, E., Shi, Y., Lee, L., Viswanathan, A., Horowitz, N., Neuberg, D., Crum, C. P., Lindeman, N. L., Kuo, F., Ligon, A. H., Freeman, G. J., … Rodig, S. J. (2016). Genetic Basis for PD-L1 Expression in Squamous Cell Carcinomas of the Cervix and Vulva. *JAMA Oncology*, *2*(4), 518–522.](http://paperpile.com/b/D0dMrm/74fd)
11. [Huang, Z. (2018). GATK Test Protocol v1 (protocols.io.mhdc326). In *protocols.io*. https://doi.org/](http://paperpile.com/b/D0dMrm/pFwR4)[10.17504/protocols.io.mhdc326](http://dx.doi.org/10.17504/protocols.io.mhdc326)
12. [Hundal, J., Carreno, B. M., Petti, A. A., Linette, G. P., Griffith, O. L., Mardis, E. R., & Griffith, M. (2016). pVAC-Seq: A genome-guided in silico approach to identifying tumor neoantigens. *Genome Medicine*, *8*(1), 11.](http://paperpile.com/b/D0dMrm/dHuYG)
13. [Larson, D. E., Abbott, T. E., & Wilson, R. K. (2014). Using SomaticSniper to Detect Somatic Single Nucleotide Variants. *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]*, *45*, 15.5.1–8.](http://paperpile.com/b/D0dMrm/RdWuC)
14. [Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. In *arXiv [q-bio.GN]*. arXiv.](http://paperpile.com/b/D0dMrm/tsZqD) <http://arxiv.org/abs/1303.3997>
15. [Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* , *25*(16), 2078–2079.](http://paperpile.com/b/D0dMrm/SmfgI)
16. [Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550.](http://paperpile.com/b/D0dMrm/j0NWp)
17. [Mulder, D. T., Mahé, E. R., Dowar, M., Hanna, Y., Li, T., Nguyen, L. T., Butler, M. O., Hirano, N., Delabie, J., Ohashi, P. S., & Pugh, T. J. (2018). CapTCR-seq: hybrid capture for T-cell receptor repertoire profiling. *Blood Advances*, *2*(23), 3506–3514.](http://paperpile.com/b/D0dMrm/t5Cu0)
18. [Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., Hoang, C. D., Diehn, M., & Alizadeh, A. A. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*, *12*, 453.](http://paperpile.com/b/D0dMrm/QokOp)
19. [Reble, E., Castellani, C. A., Melka, M. G., O’Reilly, R., & Singh, S. M. (2017). VarScan2 analysis of de novo variants in monozygotic twins discordant for schizophrenia. *Psychiatric Genetics*, *27*(2), 62–70.](http://paperpile.com/b/D0dMrm/V2AT9)
20. [Saunders, C. T., Wong, W. S. W., Swamy, S., Becq, J., Murray, L. J., & Cheetham, R. K. (2012). Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* , *28*(14), 1811–1817.](http://paperpile.com/b/D0dMrm/Rstne)
21. [Sergushichev, A. A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. In *bioRxiv* (p. 060012). https://doi.org/](http://paperpile.com/b/D0dMrm/kuEKP)[10.1101/060012](http://dx.doi.org/10.1101/060012)
22. [Skidmore, Z. L., Wagner, A. H., Lesurf, R., Campbell, K. M., Kunisaki, J., Griffith, O. L., & Griffith, M. (2016). GenVisR: Genomic Visualizations in R. *Bioinformatics* , *32*(19), 3012–3014.](http://paperpile.com/b/D0dMrm/ikPhO)
23. [Wickham, H. (2016). Programming with ggplot2. In *Use R!* (pp. 241–253).](http://paperpile.com/b/D0dMrm/DHLYm)
24. [Xie, C., Yeo, Z. X., Wong, M., Piper, J., Long, T., Kirkness, E. F., Biggs, W. H., Bloom, K., Spellman, S., Vierra-Green, C., Brady, C., Scheuermann, R. H., Telenti, A., Howard, S., Brewerton, S., Turpaz, Y., & Venter, J. C. (2017). Fast and accurate HLA typing from short-read next-generation sequence data with xHLA. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 8059–8064.](http://paperpile.com/b/D0dMrm/xH1t7)
25. [Zhang, L., Cham, J., Paciorek, A., Trager, J., Sheikh, N., & Fong, L. (2017). 3D: diversity, dynamics, differential testing - a proposed pipeline for analysis of next-generation sequencing T cell repertoire data. *BMC Bioinformatics*, *18*(1), 129.](http://paperpile.com/b/D0dMrm/nHFWQ)