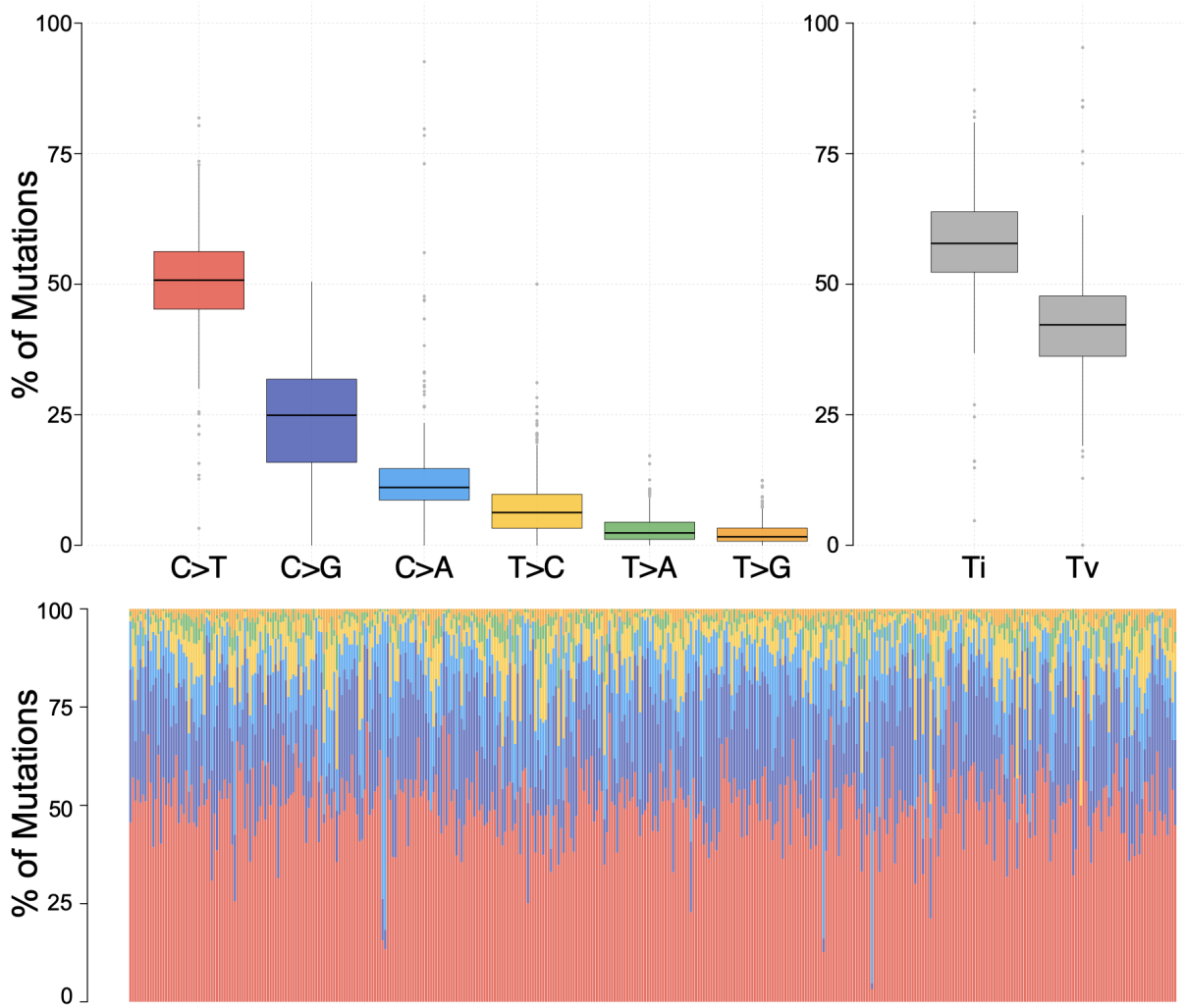

Figure S1. Pairwise nucleotide substitution patterns in the urothelial cancer cohort of the Cancer Genome Atlas (TCGA). The frequency of transitions (Ti) and transversions (Tv) is also shown. The number of samples is 411.

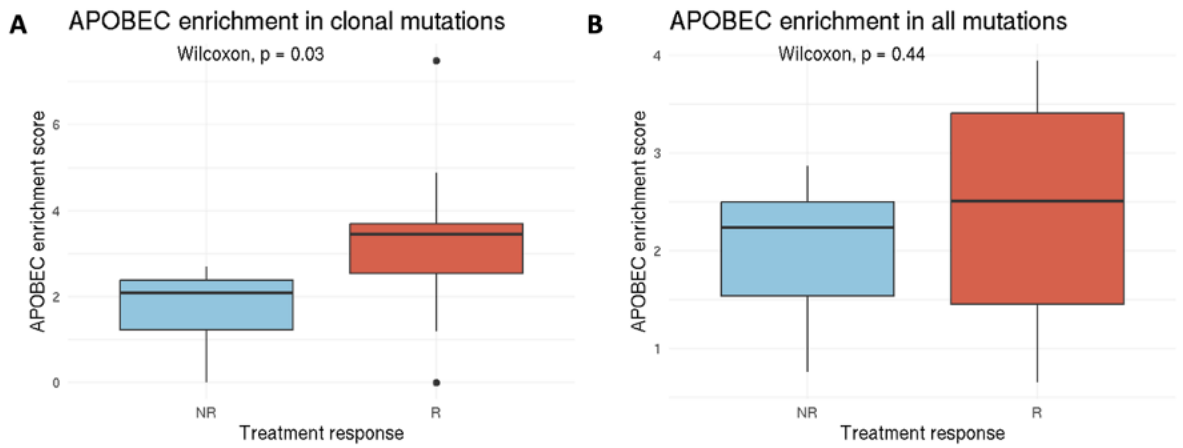

Figure S2. Relationship between APOBEC mutations and response to treatment. A. APOBEC enrichment in clonal mutations. There is a significant positive relationship between response to treatment and the APOBEC enrichment in clonal mutations (p-value = 0.03, Wilcoxon rank sum test). **B. APOBEC enrichment in all mutations.** The differences between responders and non-responders are not statistically significant. NR: no responders, R: responders.

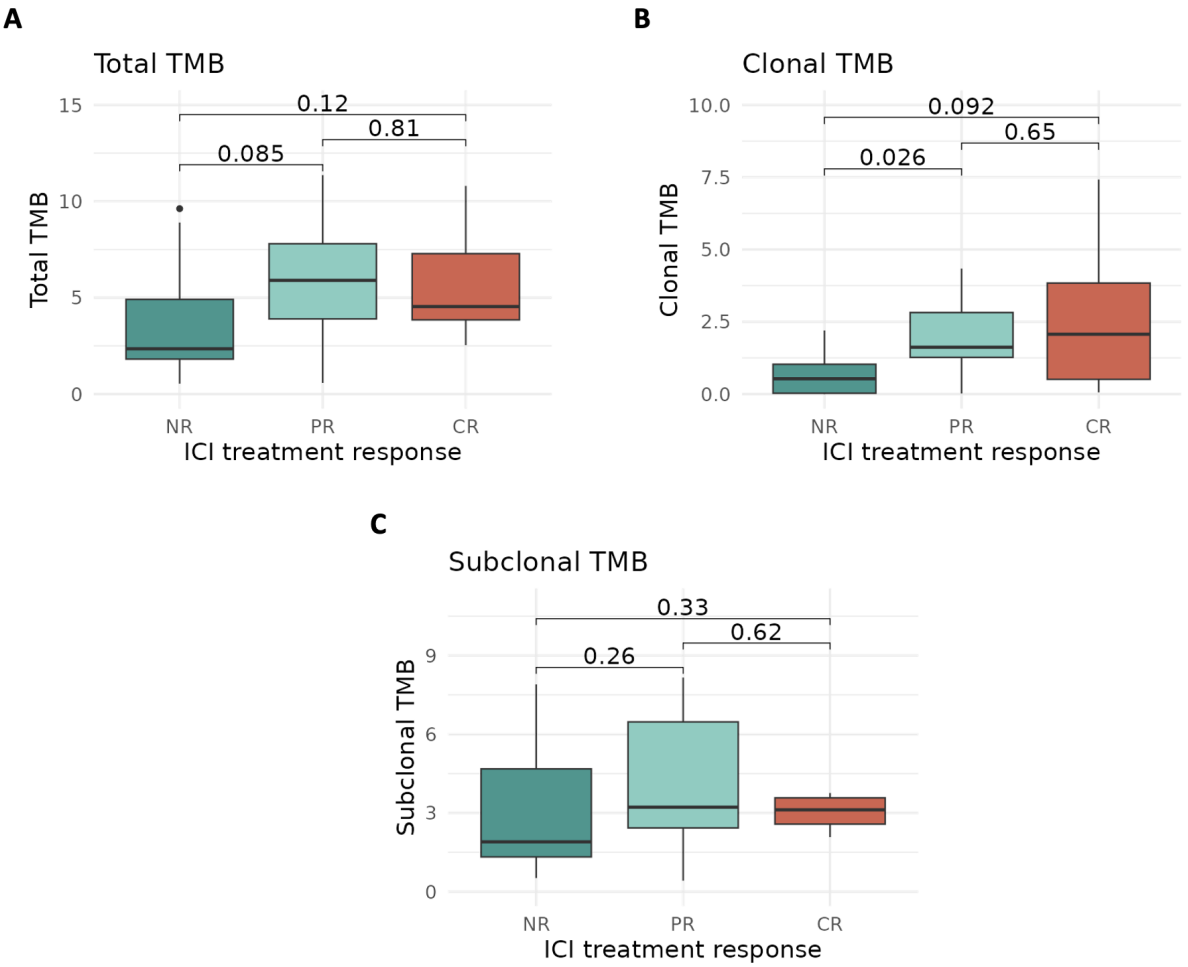

Figure S3. Relationship between TMB and the three different response groups to ICI therapy. A. Relationship between TMB and response to ICI treatment. The differences in TMB values between the three response groups were not significant. **B. Relationship between clonalTMB and response to ICI treatment.** The clonal TMB for partial responders is significantly higher compared to non-responders (p-value = 0.026). The differences in clonal TMB between non-responders and complete responders and partial responders compared to complete responders did not reach statistical significance. **C. Relationship between subclonal TMB and response to ICI treatment.** The differences in subclonal TMB values between the three response groups were not significant. NR: no responders, PR: partial responders, CR: complete responders.

NetMHCpan 4.0

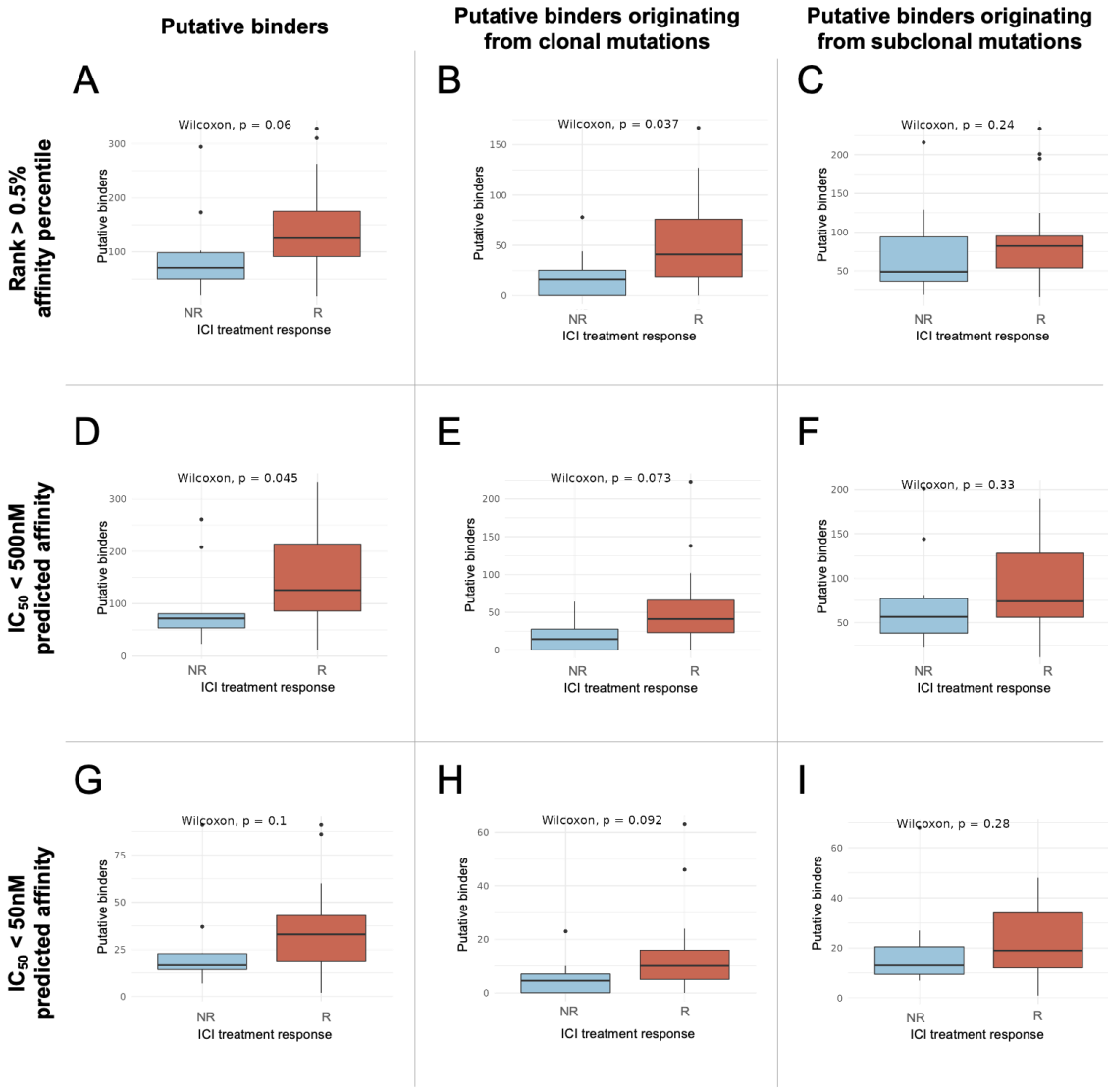

Figure S4. Relationship between ICI treatment response and the number of putative binders predicted with NetMHCpan 4.0. The number of predicted binders tends to be higher in responders than in non-responders. **A-C:** Number of putative binders with predicted affinity rank < 0.5%. **D-F:** Number of putative binders with predicted affinity IC_{50} < 500nM. **G-I:** Number of putative binders with predicted affinity IC_{50} < 50nM. P-values were calculated using the Wilcoxon rank sum test. NR: no responders, R: responders.

MHCflurry 2.0

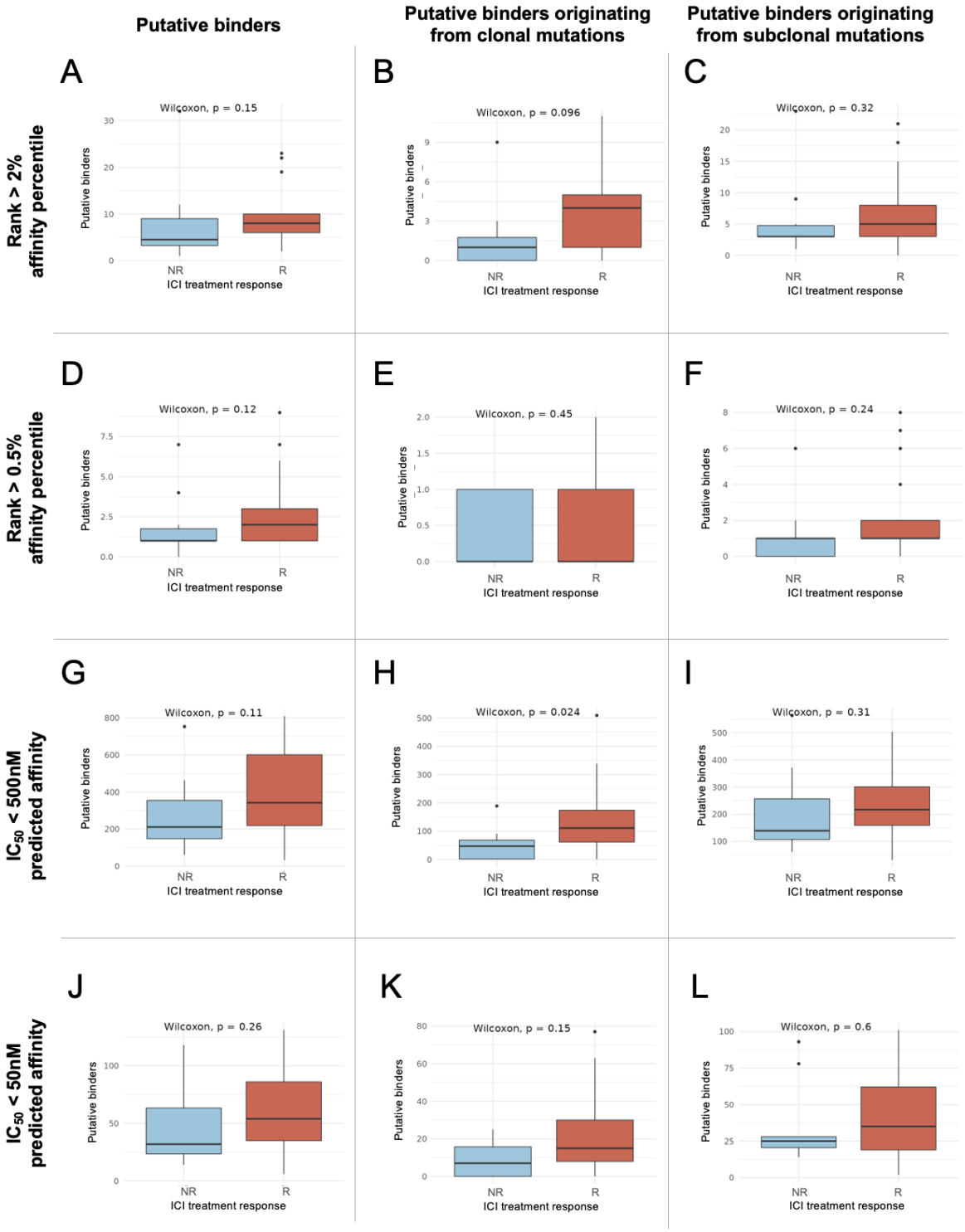

Figure S5. Relationship between ICI treatment response and the number of putative binders predicted with MHCflurry 2.0. The number of putative binders tends to be higher in responders than non-responders. **A-C:** Number of putative binders with predicted affinity rank < 2%. **D-F:** Number of putative binders with predicted affinity rank < 0.5%. **G-I:** Number of putative binders with predicted affinity IC_{50} < 500nM. **J-L:** Number of putative binders with predicted affinity IC_{50} < 50nM. P-values were calculated using the Wilcoxon rank sum test. NR: no responders, R: responders.

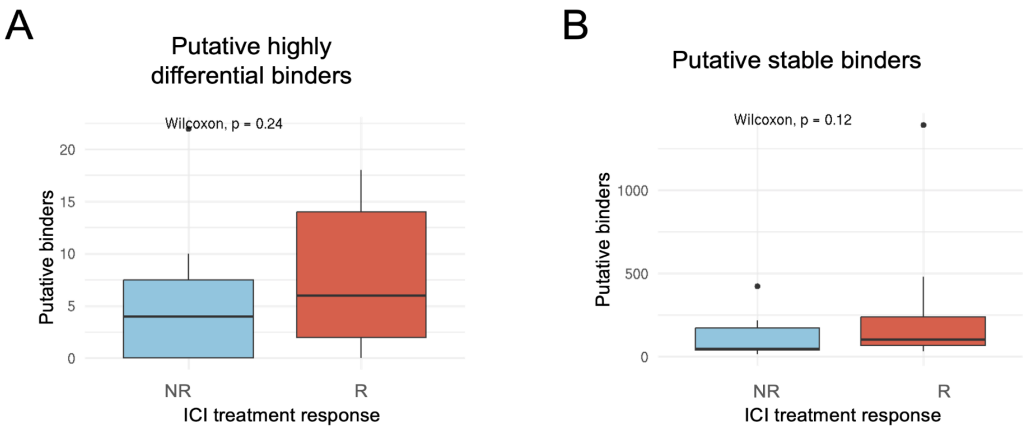

Figure S6. A. Relationship of differential agretopicity index (DAI) and treatment response. No significant difference in highly differential peptides (DAI>9) was seen between response groups. **B. Relationship between the number of predicted stable binders and treatment response.** No significant difference in the number of putative stable binders (binding stability<1.4h) was seen between response groups. P-values were calculated using Wilcoxon rank sum test. NR: no responders, R: responders.

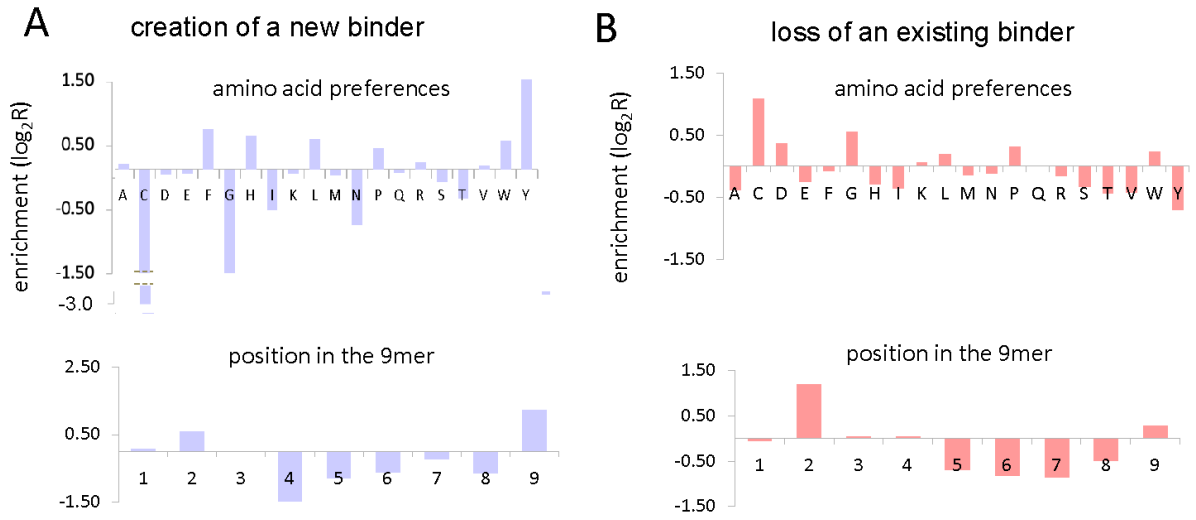

Figure S7. Non-synonymous mutations affect the binding affinity of the peptide. A. Formation of new MHC I binders. Enrichment for different amino acid substitutions (above) or positions in the peptide (below) are measured as the log2 ratio of the substitution frequency in the set of new binders versus the frequency in the set of peptides that do not change binding status. Enriched amino acids using a chi-square test: tyrosine (Y), phenylalanine (F), leucine (L) and histidine (H) at p-value < 10^{-5}, tryptophan (W) at p-value = 0.002872. **B. Loss of MHC I binding capacity**. Enrichment for different amino acid substitutions (above) or positions in the peptide (below) are measured as the log2 ratio of the substitution frequency in the set of peptides associated with loss of MHC I binding *versus* the frequency in the set of peptides that do not change their binding status. Enriched amino acids using a chi-square test: cysteine (C) (p-value < 10^{-5}), glycine (G) (p-value =6.55x10^{-5}).

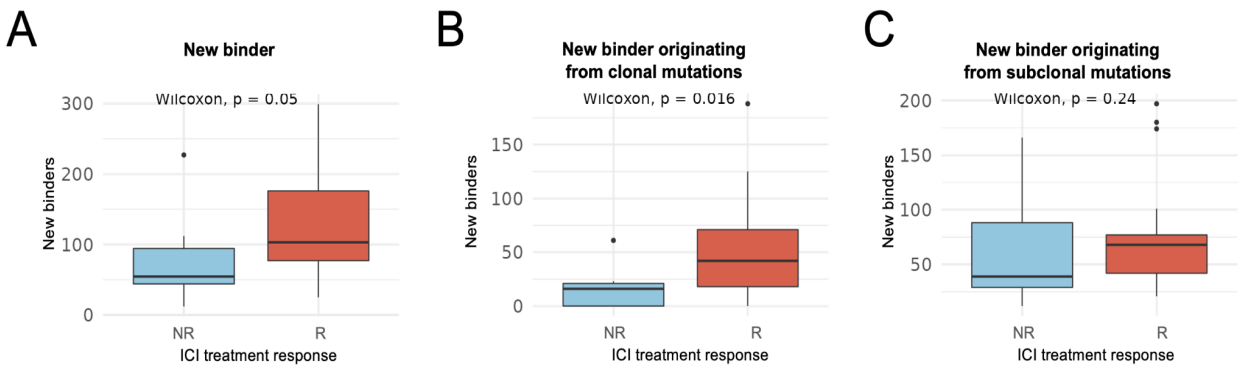

Figure S8. Relationship of number of predicted new binders and treatment response. A. Binders derived from total mutations. Responders have a higher number of new putative binders than non-responders (Wilcoxon test, p-value = 0.05). **B. Binders derived from clonal mutations.** Number of putative binders originating from clonal mutations is significantly higher in responders than non-responders (Wilcoxon test, p-value = 0.016). **C. Binders derived from subclonal mutations.** No significant difference can be observed for the number of putative binders originating from subclonal mutations between responders and non-responders. NR: no responders, R: responders.

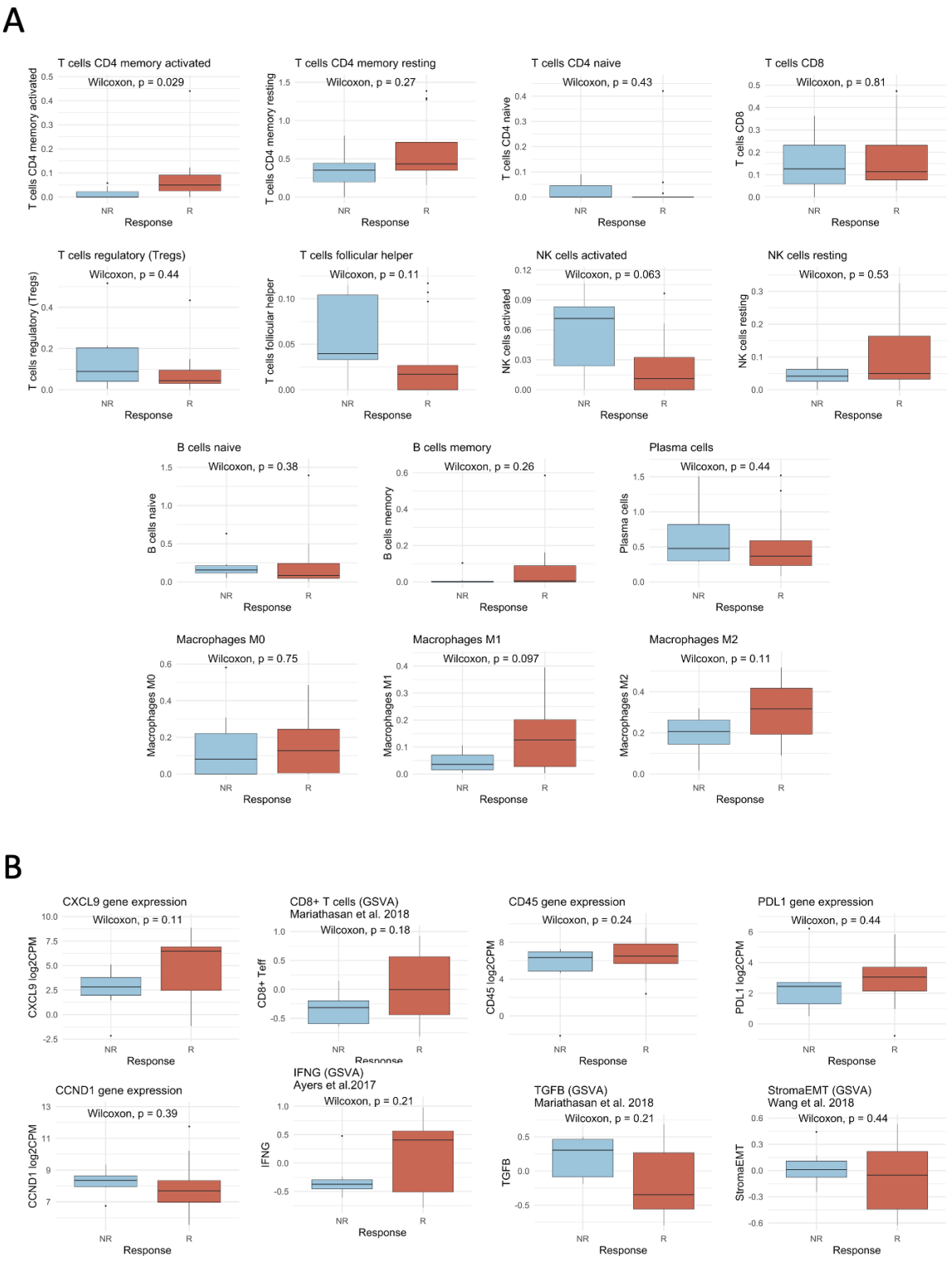

Figure S9. Relation between tumor-infiltrating immune cells and immune biomarkers and treatment response. Boxplots comparing **A.** fraction of immune cells and treatment response and **B.** gene expression of different immune biomarkers. Immune cell infiltration was estimated with CIBERSORT using gene expression profiles. P-values were calculated using Wilcoxon rank sum test. Details on immune biomarkers and immune cell profiles can be found in the supplementary data file 2. NR: no responders, R: responders.

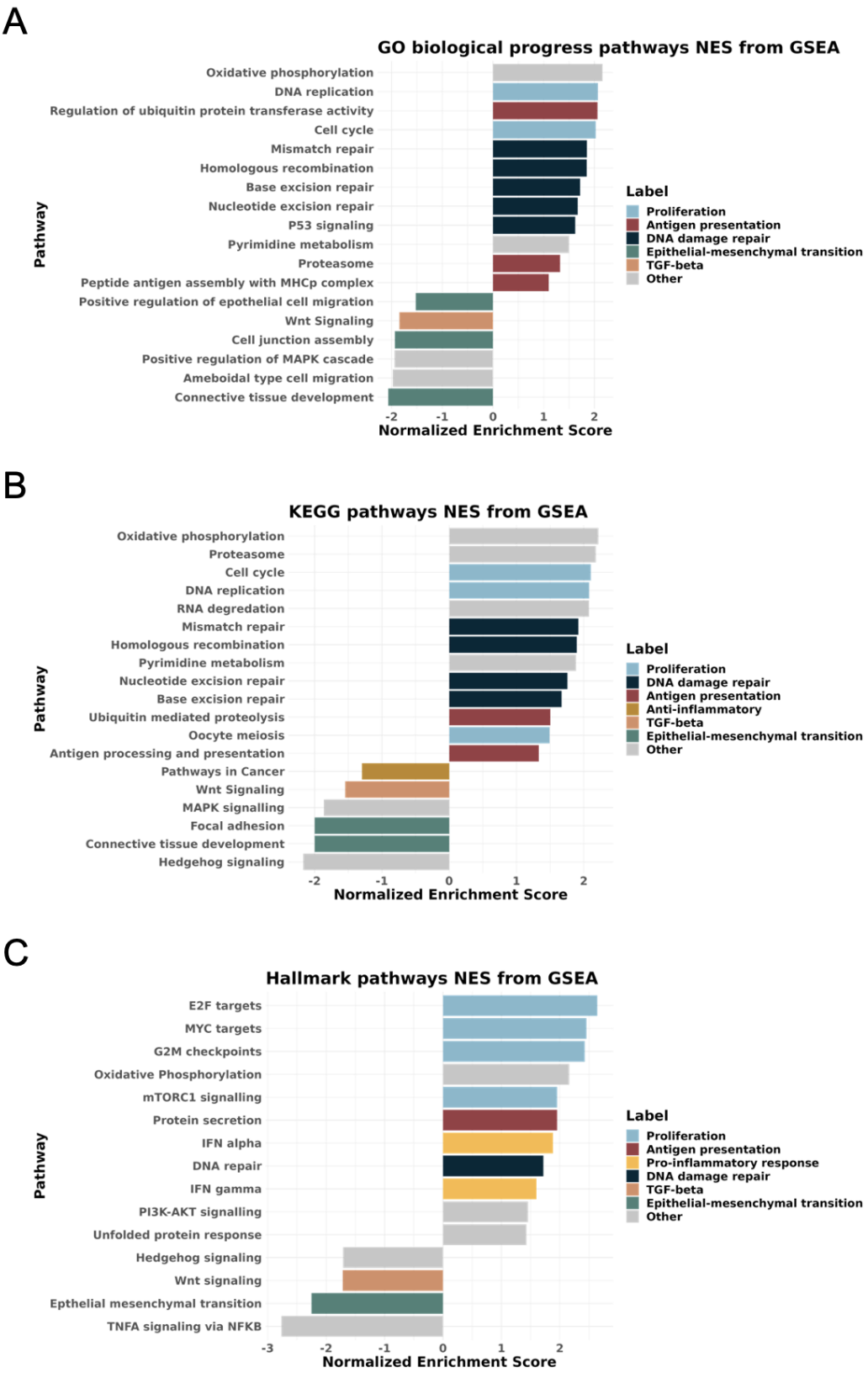

Figure S10. Pathways of gene set enrichment analysis significantly related to ICI response (adjusted *P* < 0.05, comparing 13 responders and 7 non-responders). Selected pathways are shown. The complete list of pathways with their adjusted p-value, normalized enrichment score and the included genes is provided in the Additional data file 3.

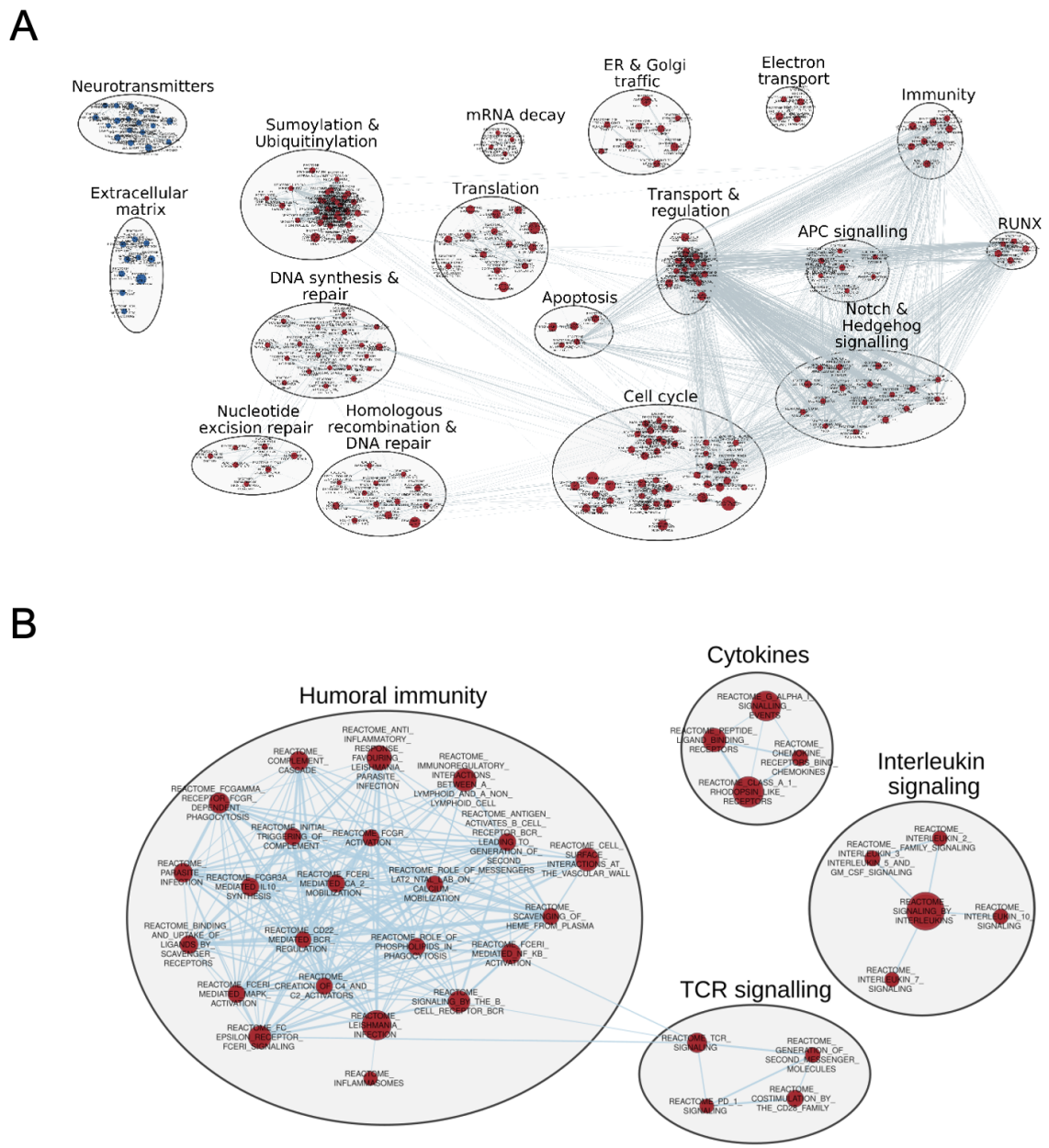

Figure S11. Network of significantly enriched pathways. Enrichment map of GSEA results (adjusted p-value < 0.05) comparing **A.** responders *versus* non-responders and **B.** complete responders *versus* non-responders. Nodes represent genesets and edges represent the connectivity between genesets (combined metric Jaccard+Overlap >0.375). Red and blue represent positive and negative enrichment scores, respectively.