Lo et al. Supplementary Figures:

Supplementary Figure 1: Tissue processing, MR-seq and data analysis workflow. H&E and unstained slides were generated to annotate the tumor region and perform tumor enrichment using the AVENIO Millisect prior to nucleic acid extraction, sequencing and analysis as shown to identify all somatic mutations and subsequently predict neoantigens.

Supplementary Figure 2: Tumor enrichment using AVENIO Millisect semi-automated tissue dissection. Annotated digital masks (A, B) were created from H&E slides to mask tumor regions (B, C) and applied to serially sectioned unstained slides for milling of selected areas for tumor enrichment. H&E stained post-dissection slides (E, F) demonstrates tumor regions that were successfully dissected for entry into nucleic acid extraction and sequencing.

Supplementary Table 1. Region types, captured metrics and rationale for region exclusion, when relevant.

Supplementary Table 2. Numbers of total versus global mutations, expressed mutations, and neoantigens in the five multi-region cases.

Supplementary Figure 3: Mutation class frequencies and somatic mutation signatures across tumor regions. (A) Somatic mutation counts and SNV mutation flanking nucleotide contexts for all tumor regions of all patients. (B) Relative contribution of established somatic signatures (Alexandrov et al 2013) to the SNV mutational spectra of each region across all patients.

Supplementary Figure 4: Copy number analysis of the five multi-region cases. (A) Copy number analysis of the CRC1 tumor samples. IGV plot of median logR values for copy number segments from TitanCNA. Specific CNAs with known association with CRC are indicated. (B) Copy number analysis of the CRC2 tumor samples. IGV plot of median logR values for copy number segments from TitanCNA. Specific CNAs with known association with CRC are indicated. (C) Copy number analysis of the RCC tumor samples. IGV plot of median logR values for copy number segments from TitanCNA. Specific CNAs with known association with RCC are indicated. (D) Copy number analysis of the NSCLC tumor samples. IGV plot of median logR values for copy number segments from TitanCNA. Specific CNAs with known association with NSCLC are indicated. (E) Copy number analysis of the UBC tumor samples. IGV plot of median logR values for copy number segments from TitanCNA.

Supplementary Figure 5: UpSetR analyses to address how and when sequencing a second biopsy would help to identify clonal mutations. A second biopsy can eliminate singleton mutations, enriching for clonal mutations. At the same time, sample quality is key to ensure sensitive mutation detection. (A) The CRC1 and UBC cases would benefit from two biopsies, eliminating the singleton mutations highlighted by the blue arrows (the CRC1 case is shown). (B) In the CRC2 and RCC cases, a second biopsy would enrich for clonal mutations, but cannot distinguish clade-specific from clonal mutations (the CRC2 case is shown). (C) The NSCLC would have minimal benefit from a second biopsy due to few singletons or lesion-specific mutations.

Supplementary Table 3: CNA based neoantigen loss in the five multi-region cases, without regard for RNAseq support of mutations. (A) Summary table showing numbers of mutations lost via genomic deletion or LOH events in each case, along with the proportion of them that are putatively neoantigenic. As a comparator, the proportion of all mutations that are putatively neoantigenic is included. (B-F) Enumeration of mutations that exhibited mutual exclusivity with LOH events in individual cases - CRC-1, CRC-2, UBC, RCC, and NSCLC respectively.

Supplementary Figure 6. (A) Analysis of mutant allele expression versus neoantigen status for all somatic mutations across UBC tumor regions, and (B) analysis of alt-allele read counts versus neoantigen status for all somatic mutations across UBC tumor regions.

Supplementary Table 4. Candidate neoantigens mediating the expression depletion effect across UBC tumor regions.

Supplementary Figure 7. Analysis of mutant allele expression versus neoantigen status after removing the 14 candidate neoantigens mediating the expression depletion effect across UBC tumor regions.

Supplementary Figure 8. (A) Analysis of mutant allele expression versus neoantigen status for all somatic mutations across CRC1 tumor regions, and (B) analysis of alt-allele read counts versus neoantigen status for all somatic mutations across CRC1 tumor regions. (C) Analysis of mutant allele expression versus neoantigen status for all somatic mutations across CRC2 tumor regions, and (D) analysis of alt-allele read counts versus neoantigen status for all somatic mutations across CRC2 tumor regions. (E) Analysis of mutant allele expression versus neoantigen status for all somatic mutations across NSCLC tumor regions, and (F) analysis of alt-allele read counts versus neoantigen status for all somatic mutations across NSCLC tumor regions. (G) Analysis of mutant allele expression versus neoantigen status for all somatic mutations across RCC tumor regions, and (H) analysis of alt-allele read counts versus neoantigen status for all somatic mutations across RCC tumor regions.

Supplementary Figure 9: Evidence for neoantigen expression depletion in the TCGA BLCA and IMvigor210 bladder cancer cohorts. (A) The relationship between neoantigen status and mutant allele expression was examined for UBC patients in the IMvigor210 cohort (n=238). The difference in median expression for neoantigenic (neo) mutations versus median expression for non-neoantigenic (non-neo) mutations was determined for each patient. Barplot shows the number of patients with a negative sign or a positive sign from this median analysis, with a Chi-square p-value from a test of the proportions. Patients with only neo or non-neo mutations and these were excluded from the analysis. (B) The table shows the specific numbers of patients with each median sign along with the numbers of patients that individually showed significant differences in their neo versus non-neo expression levels (by a Wilcoxon rank-sum test). (C) Scatter plot of tumor neoantigen burden (TNB) versus median for each patient. Note the leftward trend of median, indicative of samples with neoantigen expression depletion. (D) and (E) The relationship between neoantigen status and mutant allele expression

was examined for UBC patients in the TCGA BLCA cohort (n=202). Similar analyses are shown as in (A) and (B).

Supplementary Figure 10: Allele-specific copy number alterations in HLA-I genes across all tumor regions. (A) Single allele HLA loss was detected for the HLA-A/B/C genes in all NSCLC samples, and in the liver metastasis samples from the RCC case. (*) indicates there was a signal of HLA loss that fell just below the significance threshold. (B) Tumor/normal log ratios across chromosome 6p and the HLA region. Purple lines indicate segment means for the non-HLA exons as determined by the copynumber::pcf() function in R. Labeled points at right indicate the allele-specific tumor/normal log ratios for both alleles of all HLA-I and HLA-II genes. Note that the HLA-A/B/C loss event occurs only in the liver metastasis sample (LV1), and not in the primary (P1) or IVC tumor thrombus (IVC1) sample.

Supplementary Table 5. RNA-seq alignment statistics.

Supplementary Table 6. P-values and within-patient adjusted P-values for region-level neoantigen vs non-neoantigen expression comparisons using A) alt RPKM or B) alt-allele read counts.

Supplementary Table 7. CNA based neoantigen loss in the five multi-region cases, requiring that included mutations have at least two alt-allele-supporting RNAseq reads in at least one sample. (A) Summary table showing numbers of mutations lost via genomic deletion or LOH events in each case, along with the proportion of them that are putatively neoantigenic. As a comparator, the proportion of all mutations that are putatively neoantigenic is included. (B-F) Enumeration of mutations with RNAseq support that exhibited mutual exclusivity with LOH events in individual cases - CRC-1, CRC-2, UBC, RCC, and NSCLC respectively.

Supplementary Table 8. Neoantigens whose presentation was predicted to be lost due to HLA allele loss (A) in the NSCLC case, and (B) in the RCC case.

Supplementary Figure 11. The numbers of expressed neoepitopes predicted to be presented by each HLA allele is plotted for (A) all regions of the NSCLC case, and (B) the liver metastasis regions of the RCC case. Lost alleles are indicated in pink, and retained alleles are indicated in blue. A neoepitope is defined as a peptide-HLA pair with an EL-mut score <= 2 as determined by NetMHCpan-4.0. Expressed neoepitopes are those whose underlying mutation was supported by >= 2 RNA-seq reads.