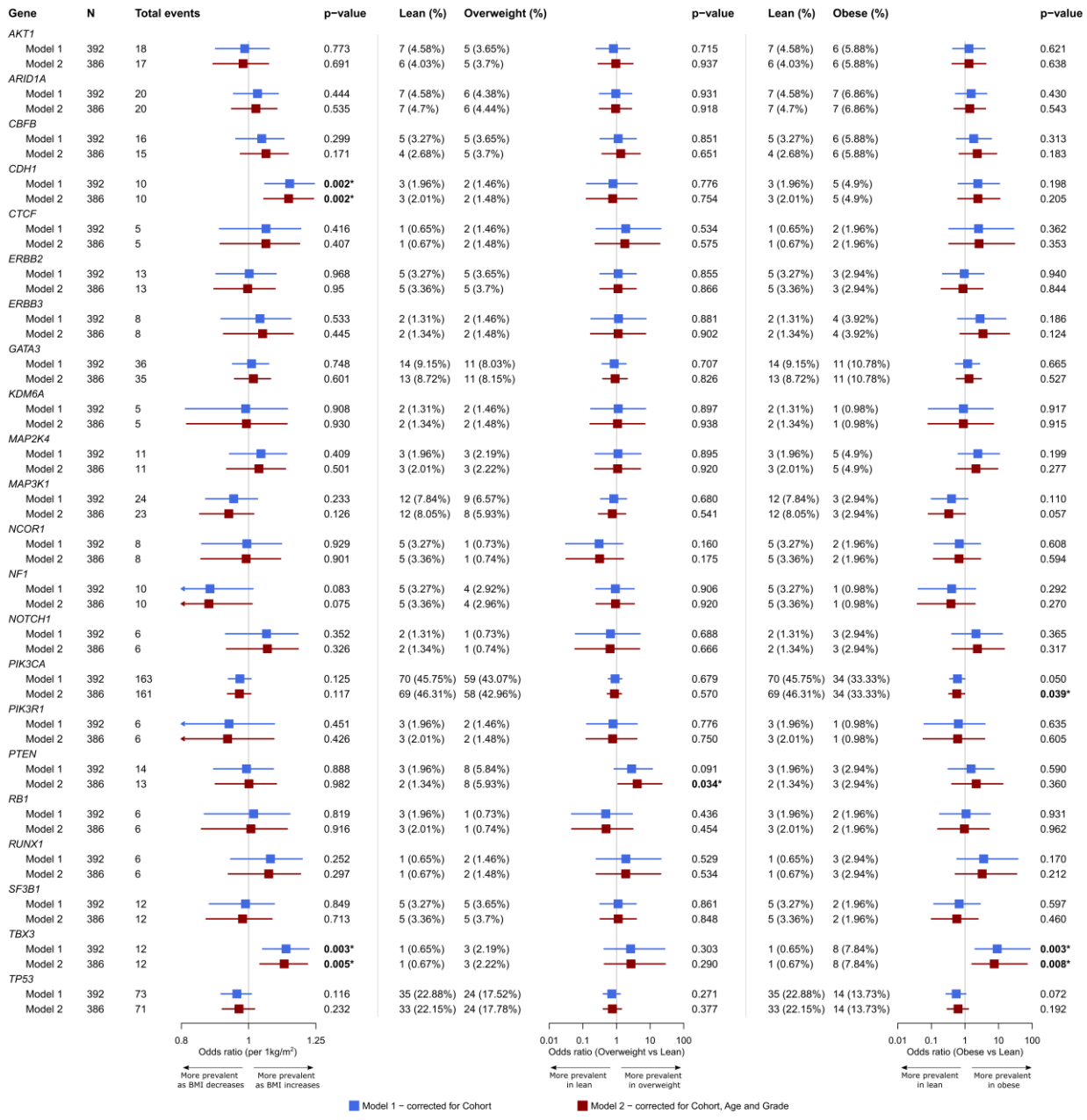
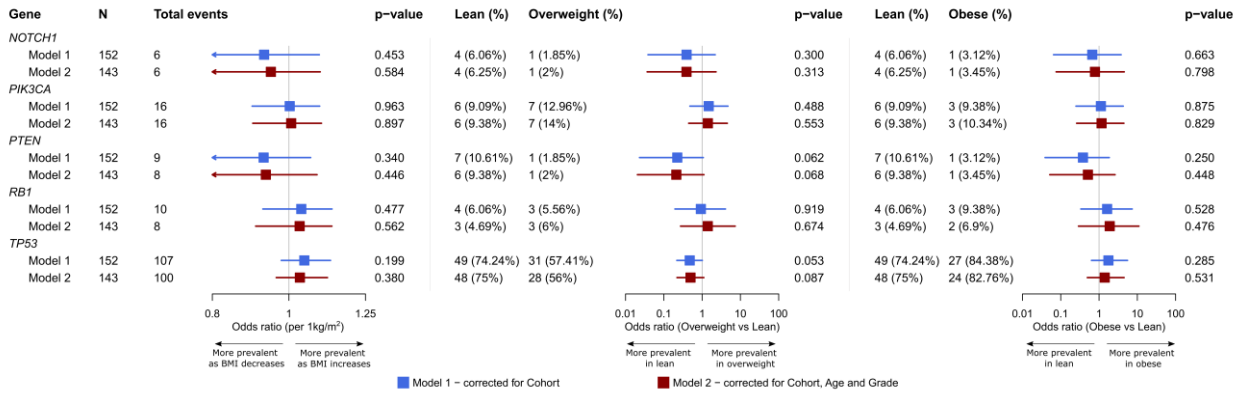
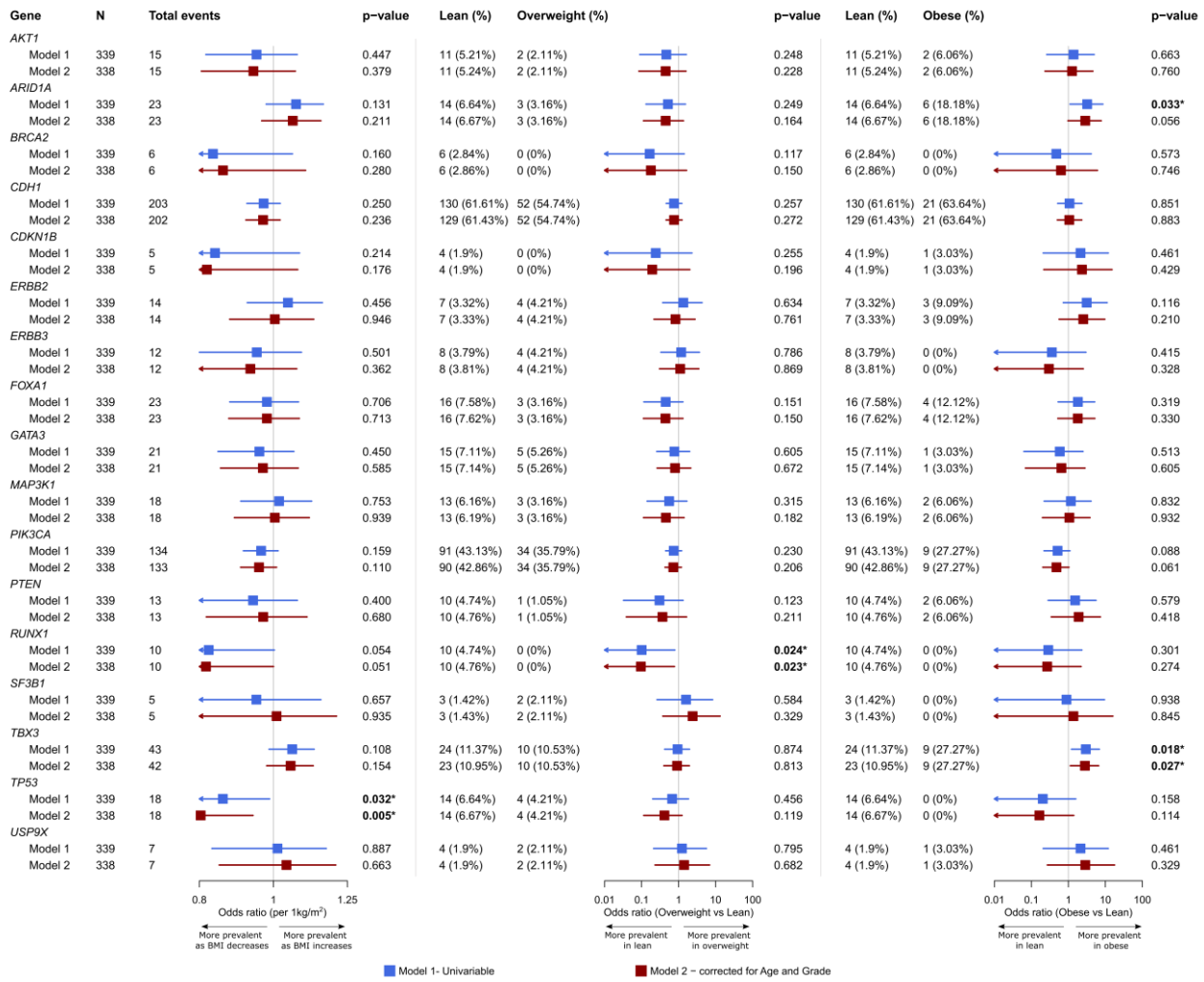
Extended Data



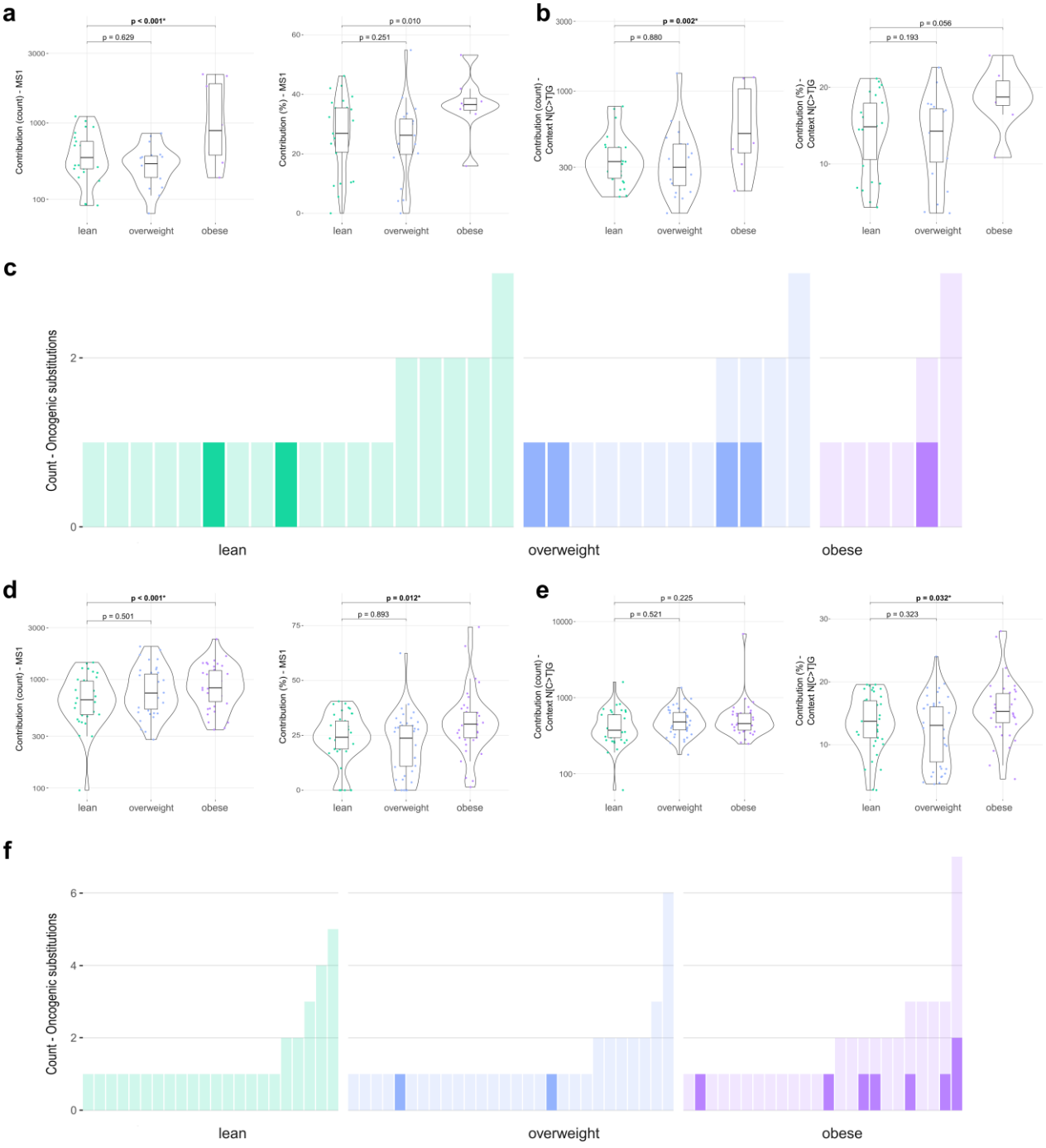
Extended Data Figure 1. Association of BMI as a categorical variable with oncogenic mutations of breast cancer-specific driver genes in NST ER+/HER2- patients. Forest plots showing the association between BMI, as a categorical variable (overweight vs lean, and obese vs lean), and driver gene mutations. Merged data from two cohorts, METABRIC and ICGC, was used. Gene mutations with at least 5 events detected in the respective cohorts were evaluated.



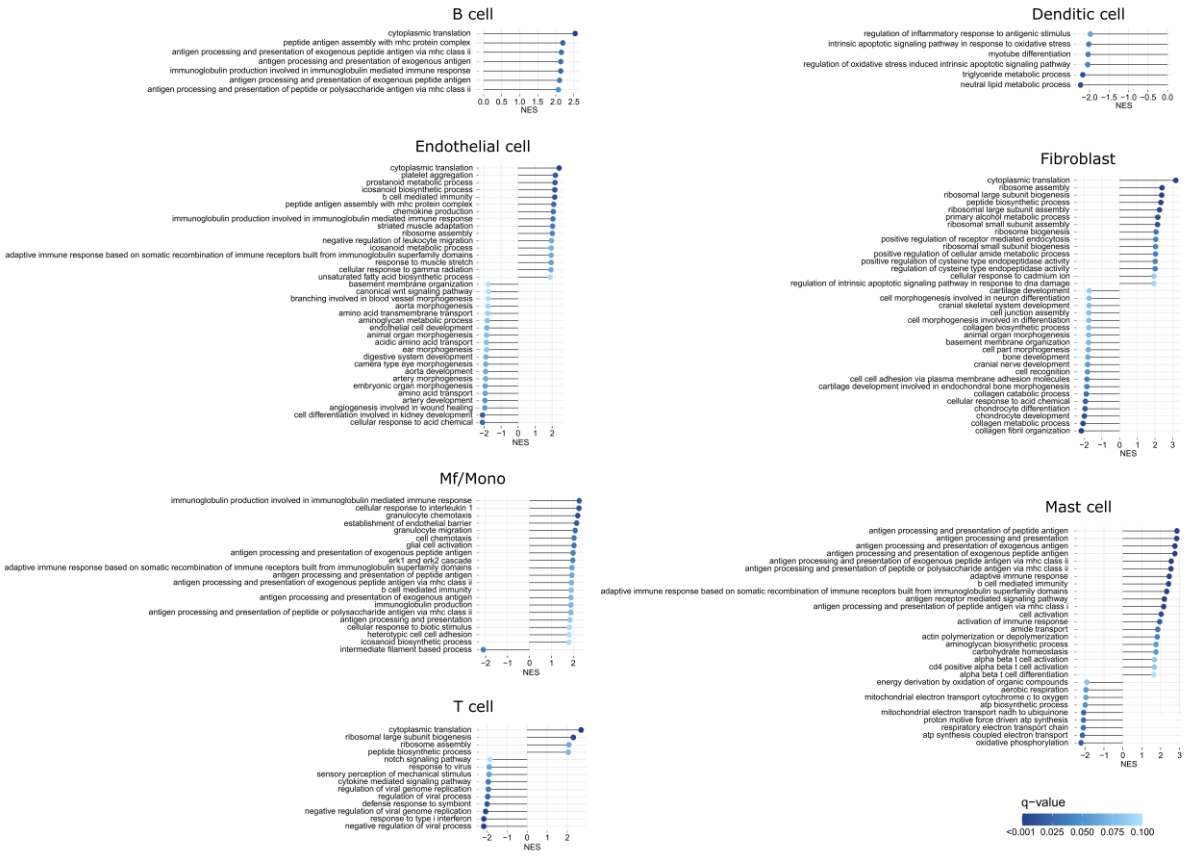
Extended Data Figure 2. Association of BMI with oncogenic mutations of breast cancer-specific driver genes in NST ER-/HER2- patients. Forest plots showing the association between BMI, as a categorical variable (overweight vs lean, and obese vs lean), and driver gene mutations. Merged data from two cohorts, METABRIC and ICGC, was used. Gene mutations with at least 5 events detected in the respective cohorts were evaluated.



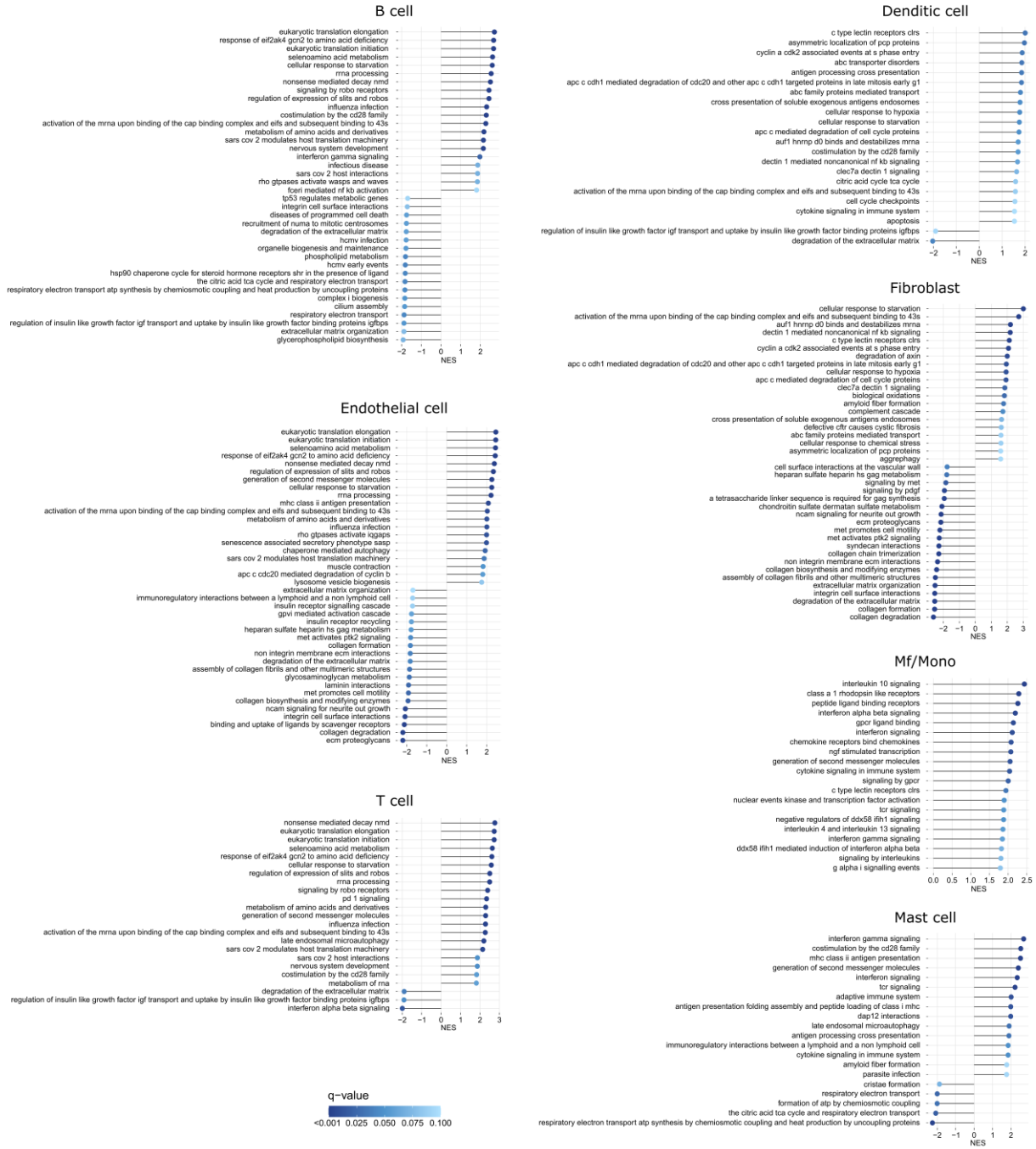
Extended Data Figure 3. Association of BMI with oncogenic mutations of breast cancer-specific driver genes in ILC ER+/HER2- patients. Forest plots showing the association between BMI, as a categorical variable (overweight vs lean, and obese vs lean), and driver gene mutations. Data from the ELBC cohort was used. Gene mutations with at least 5 events detected in the respective cohorts were evaluated.



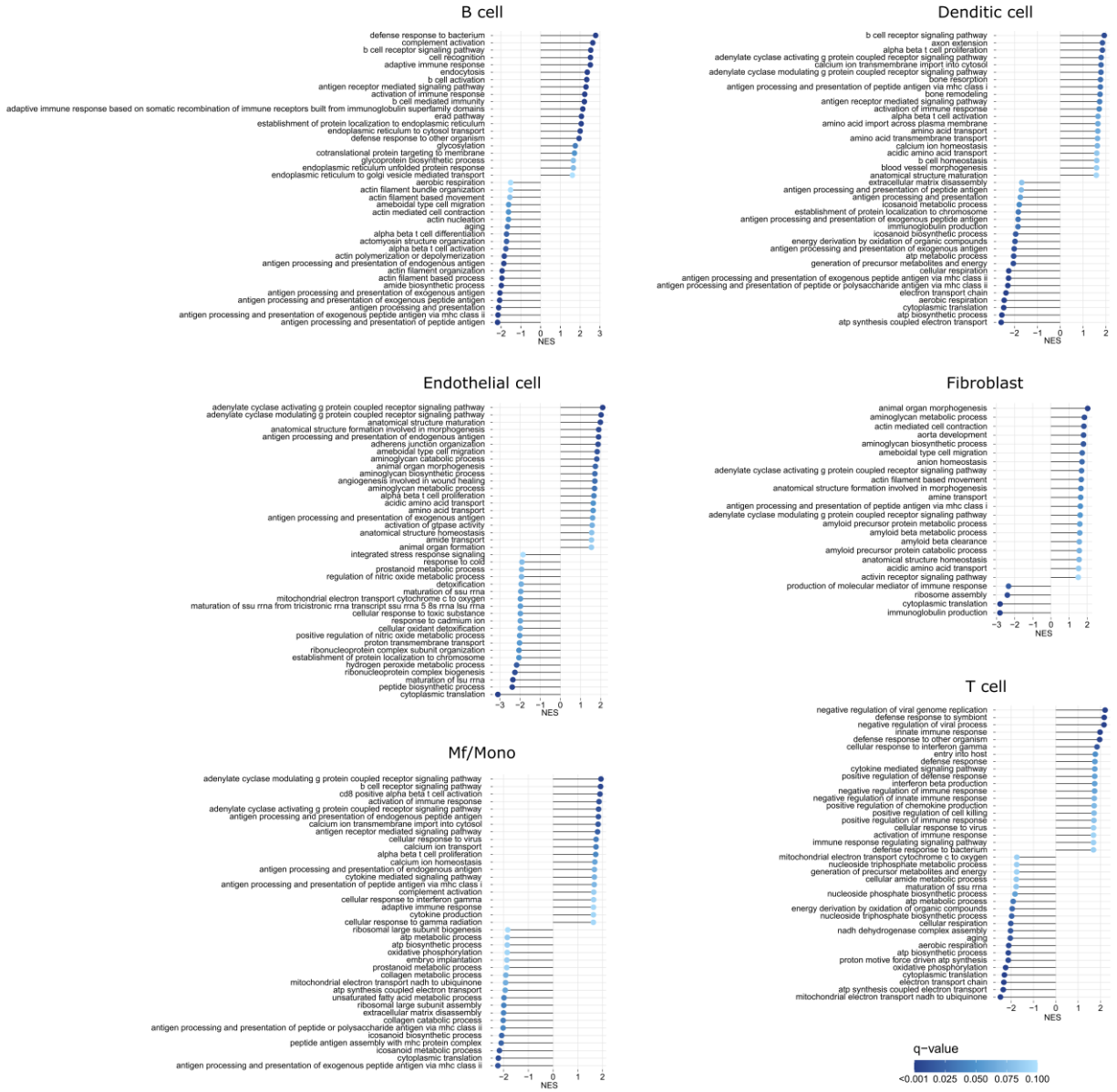
Extended Data Figure 4. Association of BMI with the age-associated mutational signature (Signature.1) in patients with NST ER+/HER2- in different age groups, >50 (a-c) and ≤50 years (d-f), from the ICGC cohort. **(a-b, d-e)** Violin plots showing the contribution in absolute count (left) and relative percentage (right) of the mutational signature 1 (MS1) (a, d) and the sequence context N[C>T]G (b, e) according to BMI categories. The p-values were derived from linear regressions adjusted for age (continuous) and tumor grade. **(c, f)** Bar plots presenting the number of oncogenic mutations occurring in BC-specific driver genes in each tumor. The count of mutations having the sequence context N[C>T]G is highlighted in bold colors. Each bar represents a tumor.



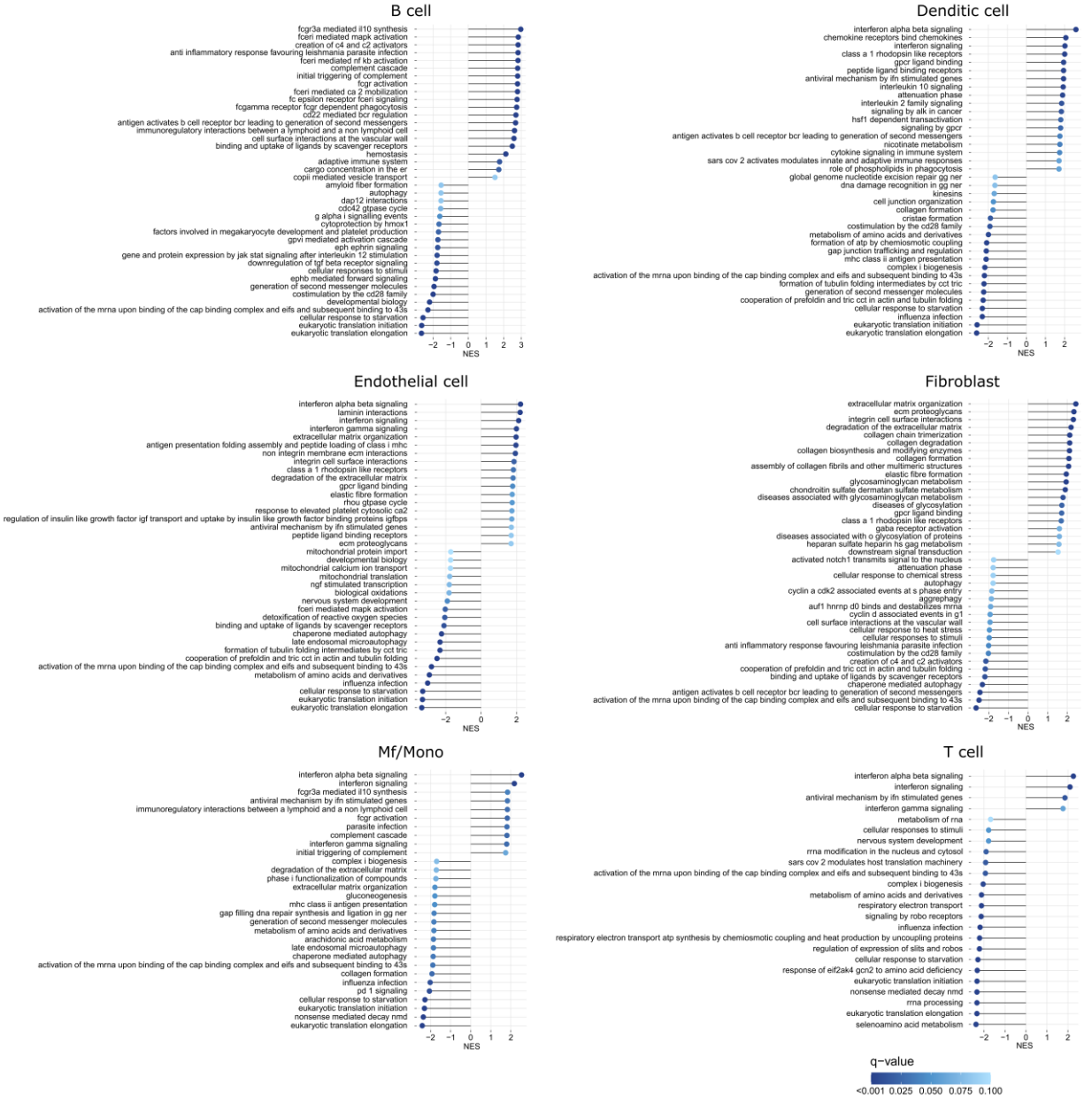
Extended Data Figure 5. Gene set enrichment analysis (GSEA) on Gene Ontology biological process (GOBP) gene sets of cancer cells from obese versus lean patients with NST ER+/HER2- from the BioKey cohort. Lollipop plots displaying top 20 upregulated (sorted by q-value, NES > 0) and top 20 downregulated (NES < 0) GOBP gene sets according to BMI category (obese vs lean) detected by GSEA (q-value < 0.1).



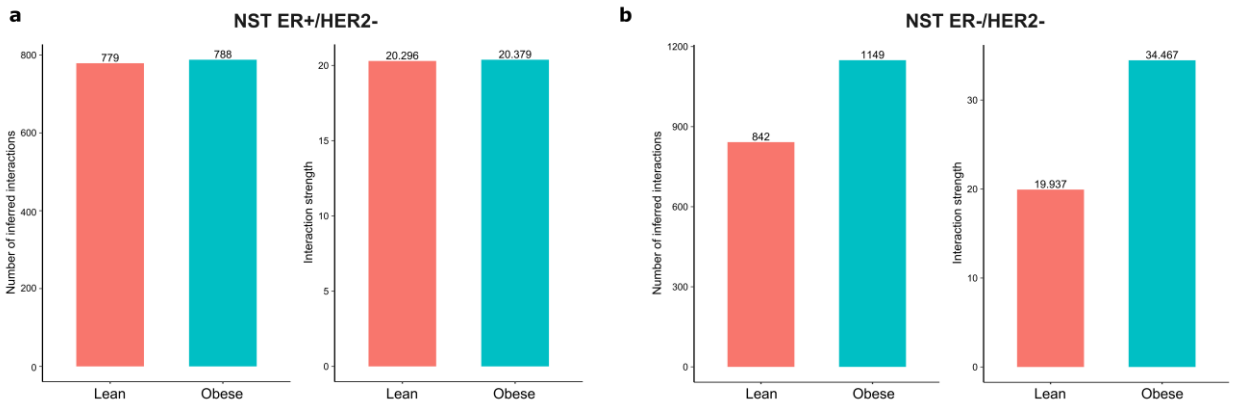
Extended Data Figure 6. GSEA on REACTOME gene sets of cancer cells from obese versus lean patients with NST ER+/HER2- from the BioKey cohort. Lollipop plots displaying top 20 upregulated (sorted by q-value, NES > 0) and top 20 downregulated (NES < 0) REACTOME gene sets according to BMI category (obese vs lean) detected by GSEA (q-value < 0.1).



Extended Data Figure 7. GSEA on GOBP gene sets of cancer cells from obese versus lean patients with NST ER-/HER2- from the BioKey cohort. Lollipop plots displaying top 20 upregulated (sorted by q-value, NES > 0) and top 20 downregulated (NES < 0) GOBP gene sets according to BMI category (obese vs lean) detected by GSEA (q-value < 0.1).



Extended Data Figure 8. Gene set enrichment analysis on REACTOME gene sets of cancer cells from obese versus lean patients with NST ER-/HER2- from the BioKey cohort. Lollipop plots displaying top 20 upregulated (sorted by q-value, NES > 0) and top 20 downregulated (NES < 0) REACTOME gene sets according to BMI category (obese vs lean) detected by GSEA (q-value < 0.1).



Extended Data Figure 9. Total intercellular interactions in tumors from lean and obese patients in the BioKey cohort. (a-b) Sum of signaling interactions in terms of number and strength inferred by CellChat in NST ER+/HER2- (a) and NST ER-/HER2- (b) tumors from lean and obese patients.