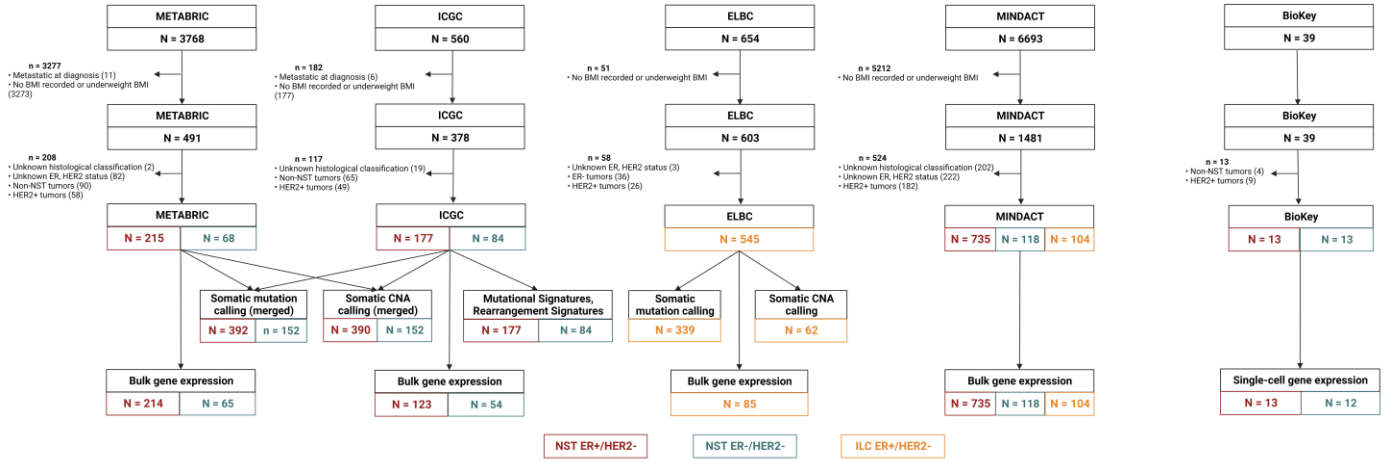
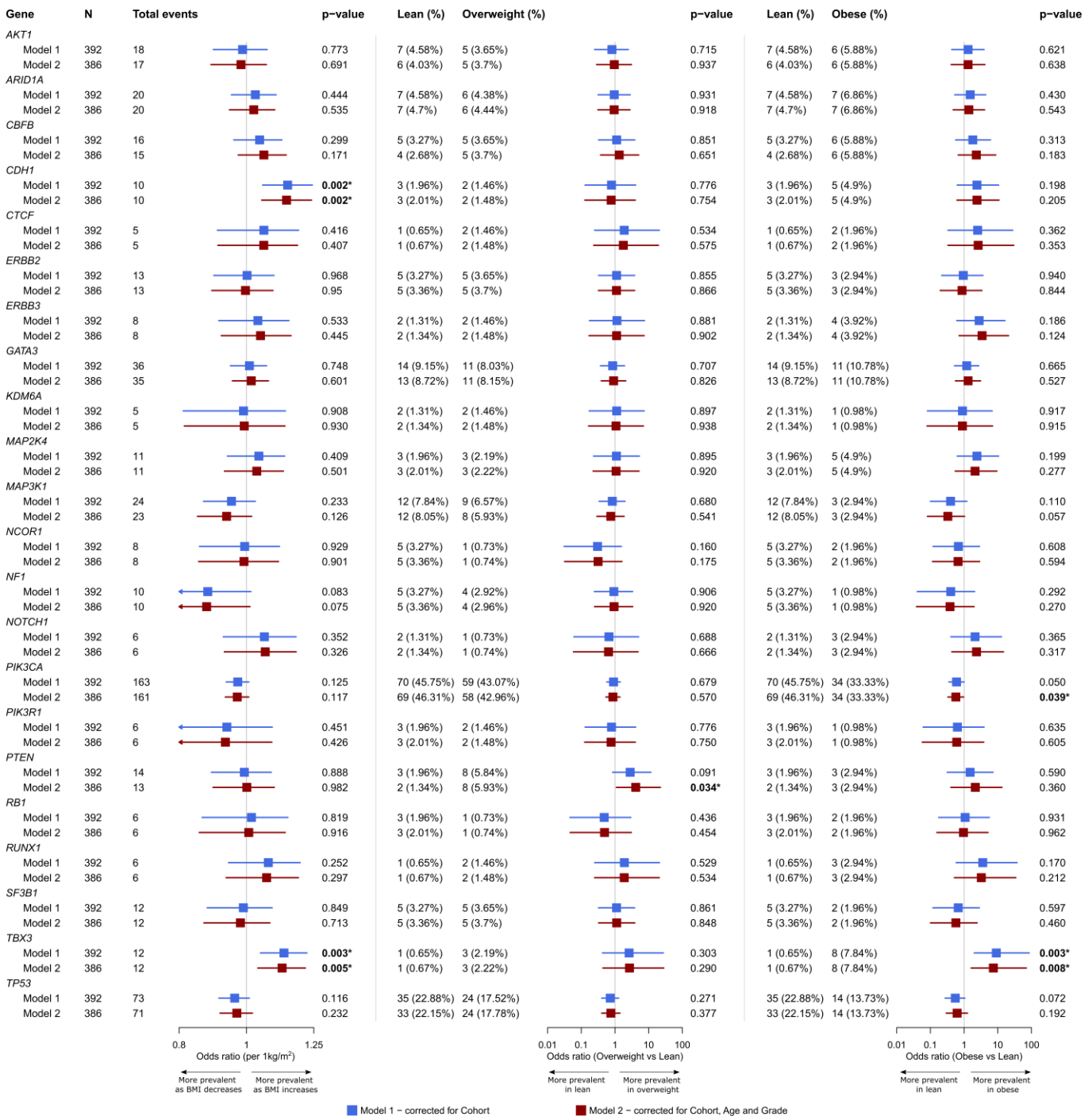
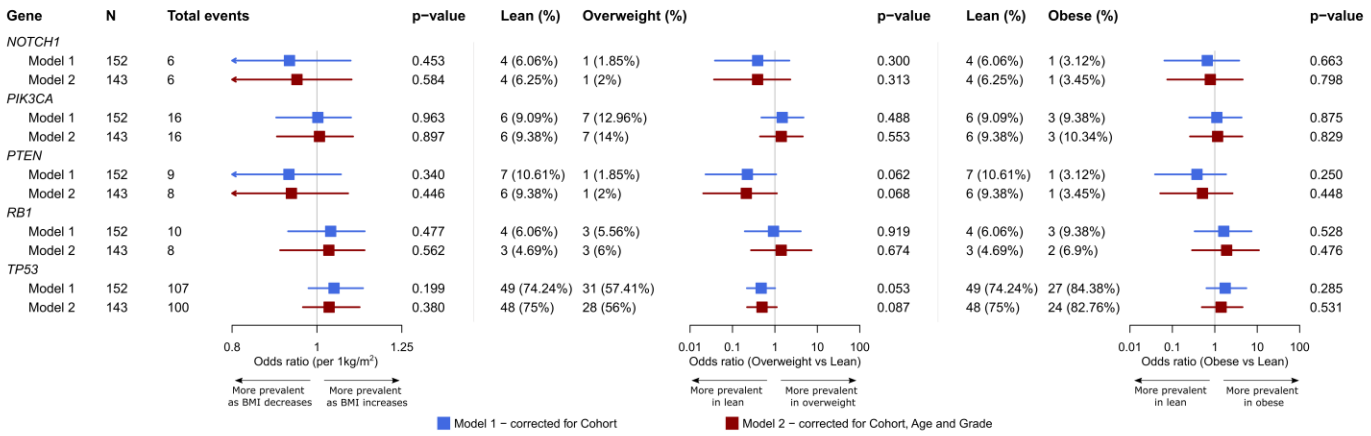
Supplementary Figures



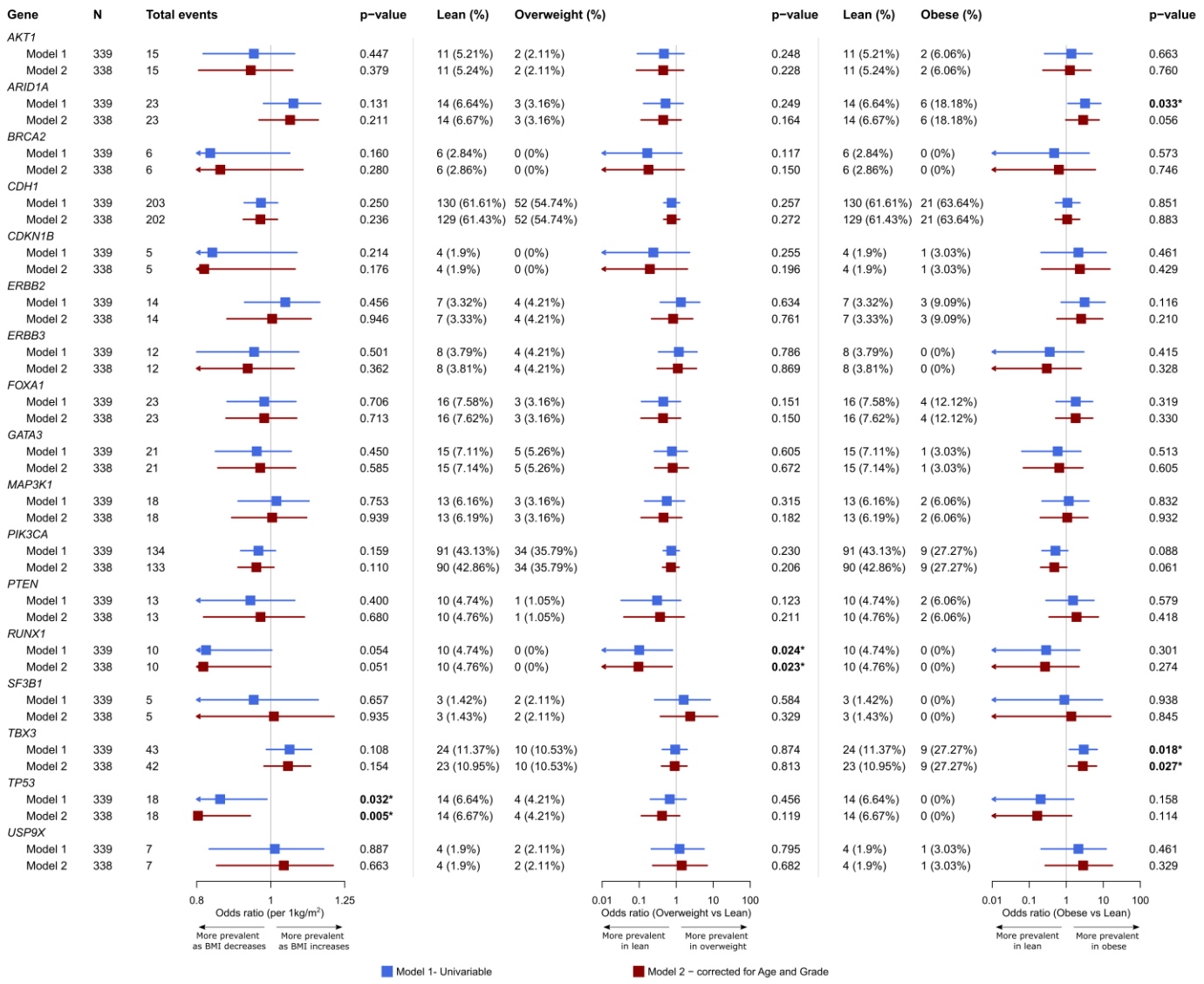
Supplementary Figure 1. Overview of the data flow in the study. Flowchart displaying the data selection and stratification process and showing the final numbers of cases analyzed in the study.



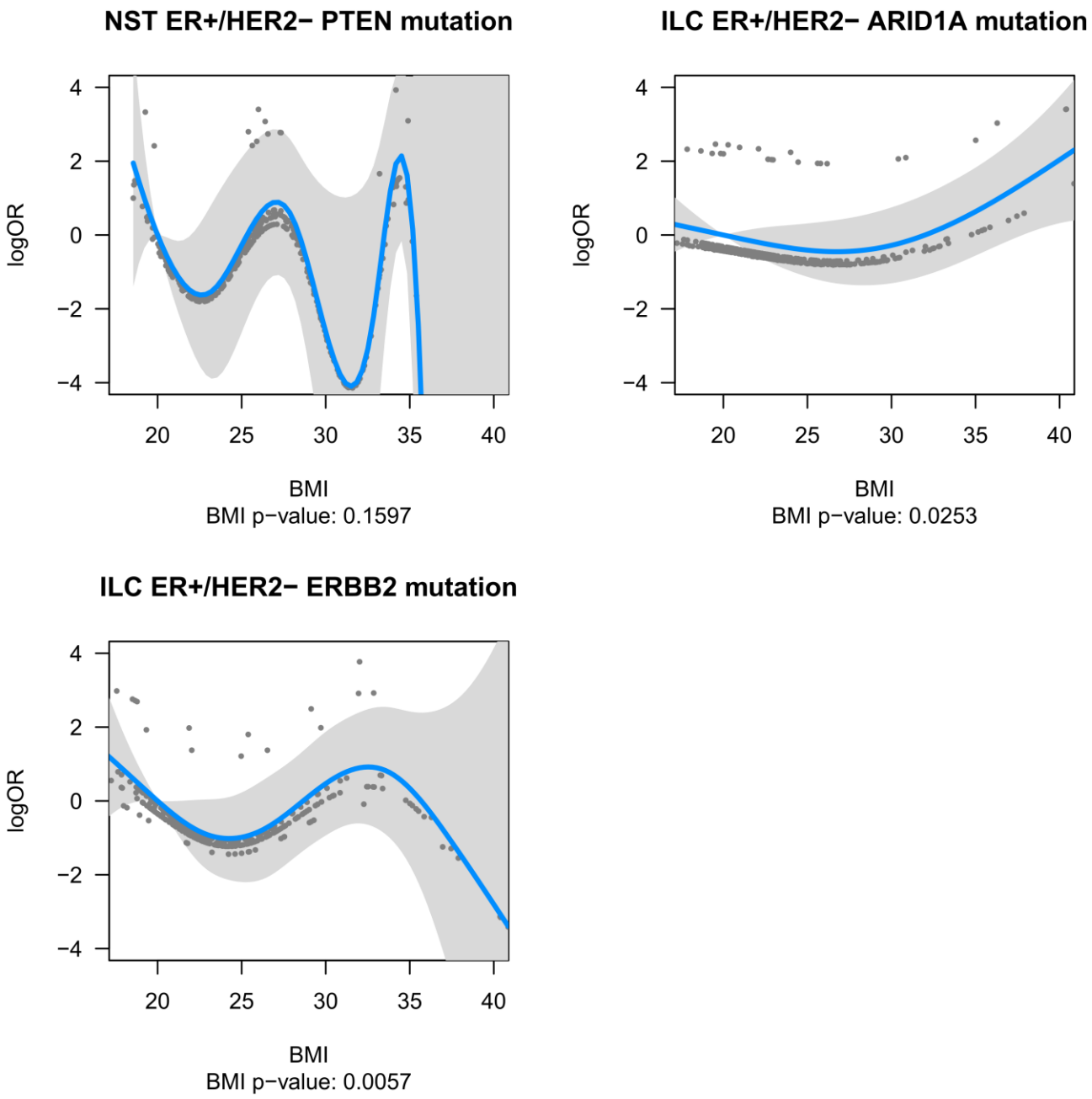
Supplementary Figure 2. Association of BMI as a categorical variable with oncogenic mutations of breast cancer-specific driver genes in NST ER+/HER2- patients. Forest plots showing the associations evaluated using Firth's logistic regression between BMI, as a categorical variable (overweight vs lean, 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. Color-coded boxes indicate point estimates of odds ratios, and whiskers indicate their corresponding 95% confidence intervals. All statistical tests were two-sided. p-values shown were not corrected for multiple testing.



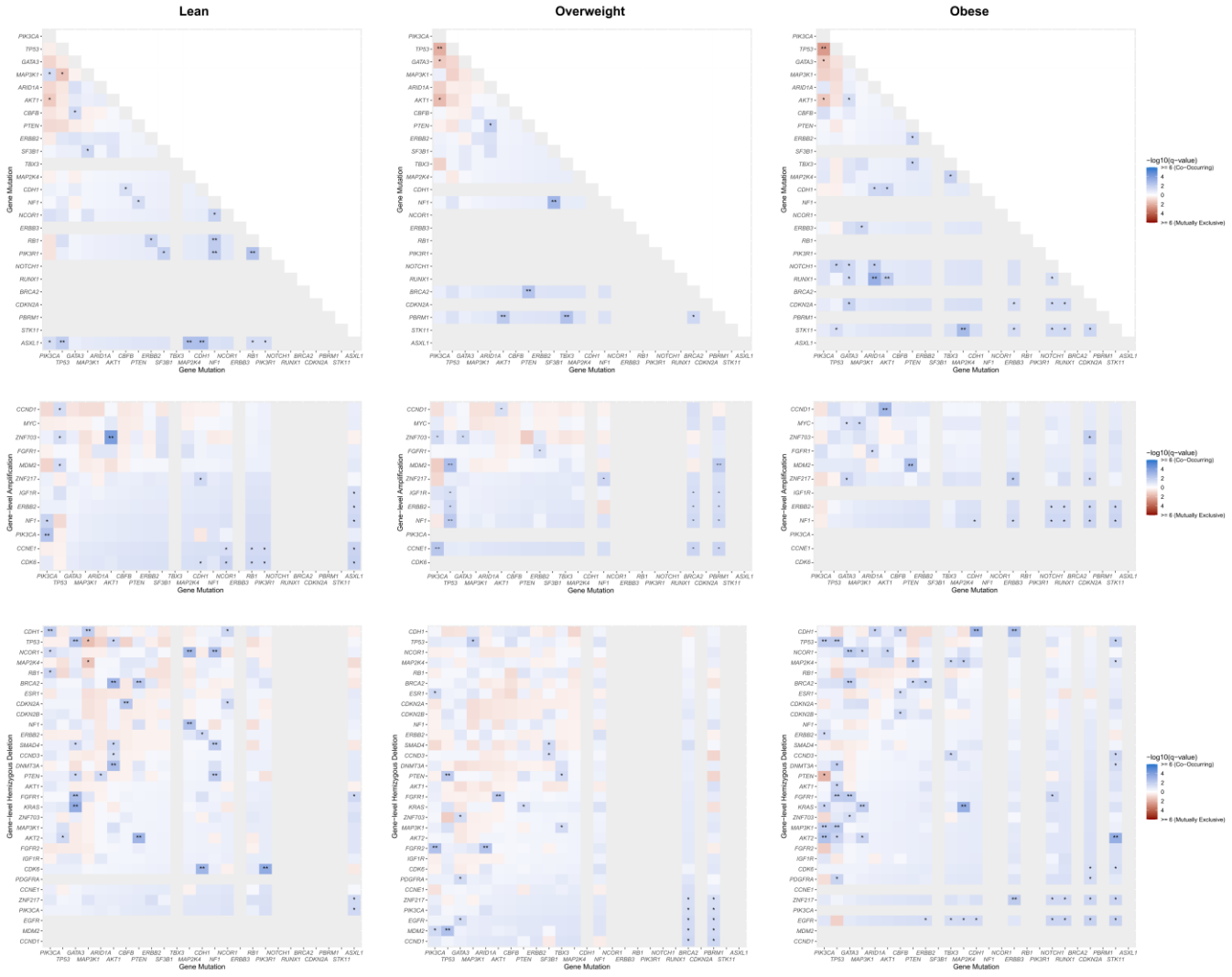
Supplementary Figure 3. Association of BMI with oncogenic mutations of breast cancer-specific driver genes in NST ER-/HER2- patients. Forest plots showing the associations evaluated using Firth's logistic regression between BMI, as a categorical variable (overweight vs lean, and obese vs lean), and driver 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. Color-coded boxes indicate point estimates of odds ratios, and whiskers indicate their corresponding 95% confidence intervals. All statistical tests were two-sided. p-values shown were not corrected for multiple testing.



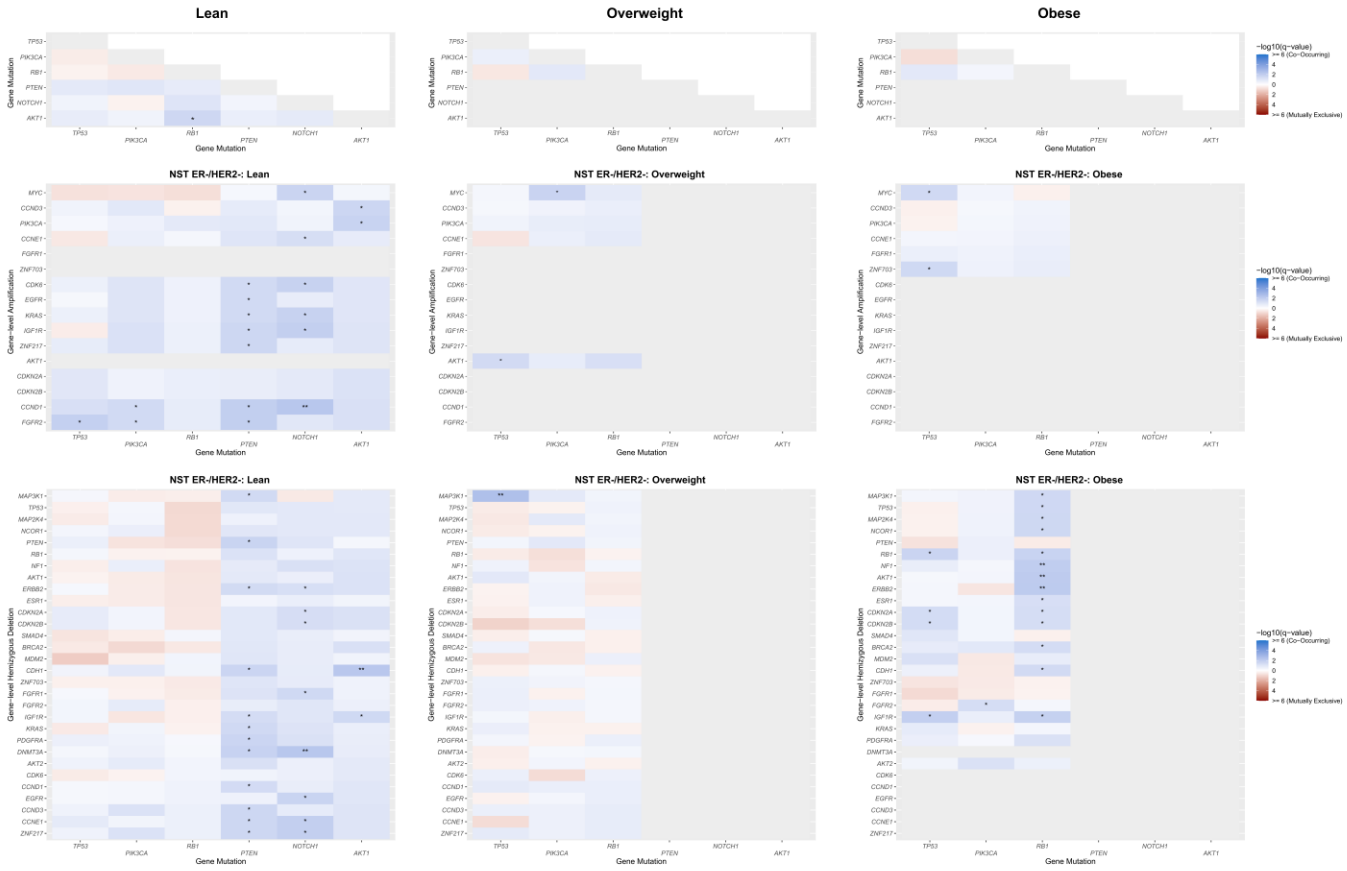
Supplementary Figure 4. Association of BMI with oncogenic mutations of breast cancer-specific driver genes in ILC ER+/HER2- patients. Forest plots showing the associations evaluated using Firth's logistic regression between BMI, as a categorical variable (overweight vs lean, and obese vs lean), and driver gene mutations. Data from the ELBC cohort was used. Gene mutations with at least 5 events detected in the respective cohorts were evaluated. Color-coded boxes indicate point estimates of odds ratios, and whiskers indicate their corresponding 95% confidence intervals. All statistical tests were two-sided. p-values shown were not corrected for multiple testing.



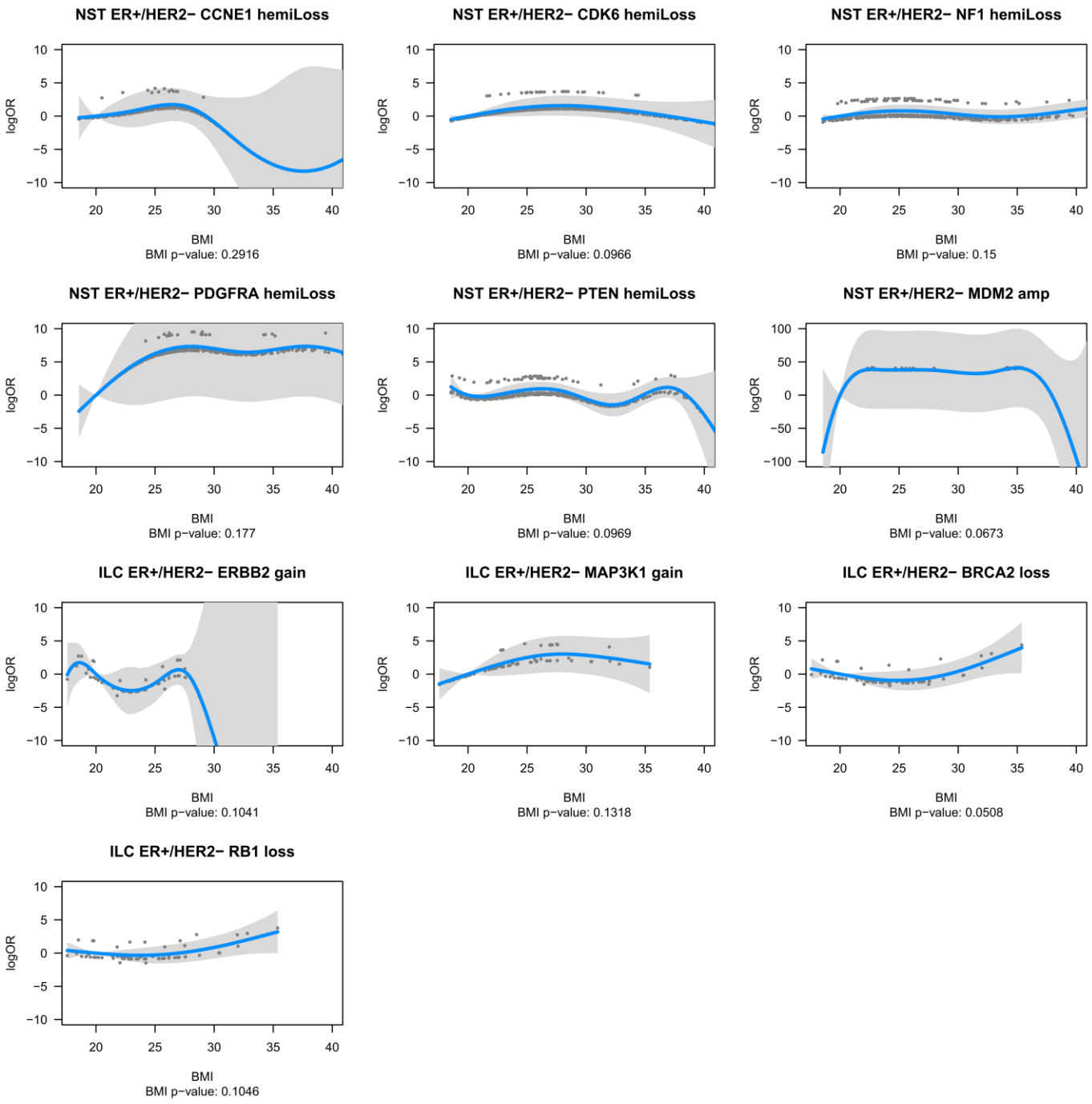
Supplementary Figure 5. Non-linear association of gene mutations with BMI. Plots showing fitted lines of the logOR of gene mutations against continuous BMI (baseline BMI of 20) for those indicated to be better associated with BMI in a non-linear model. Data are presented for point estimates of log-odds ratio (logOR; blue solid line) and their 95% confidence interval (grey shaded band). p-values shown were determined for coefficients estimated by generalized additive models fitted with a spline term of BMI and adjusted for cohort, age, and tumor grade. All statistical tests were two-sided.



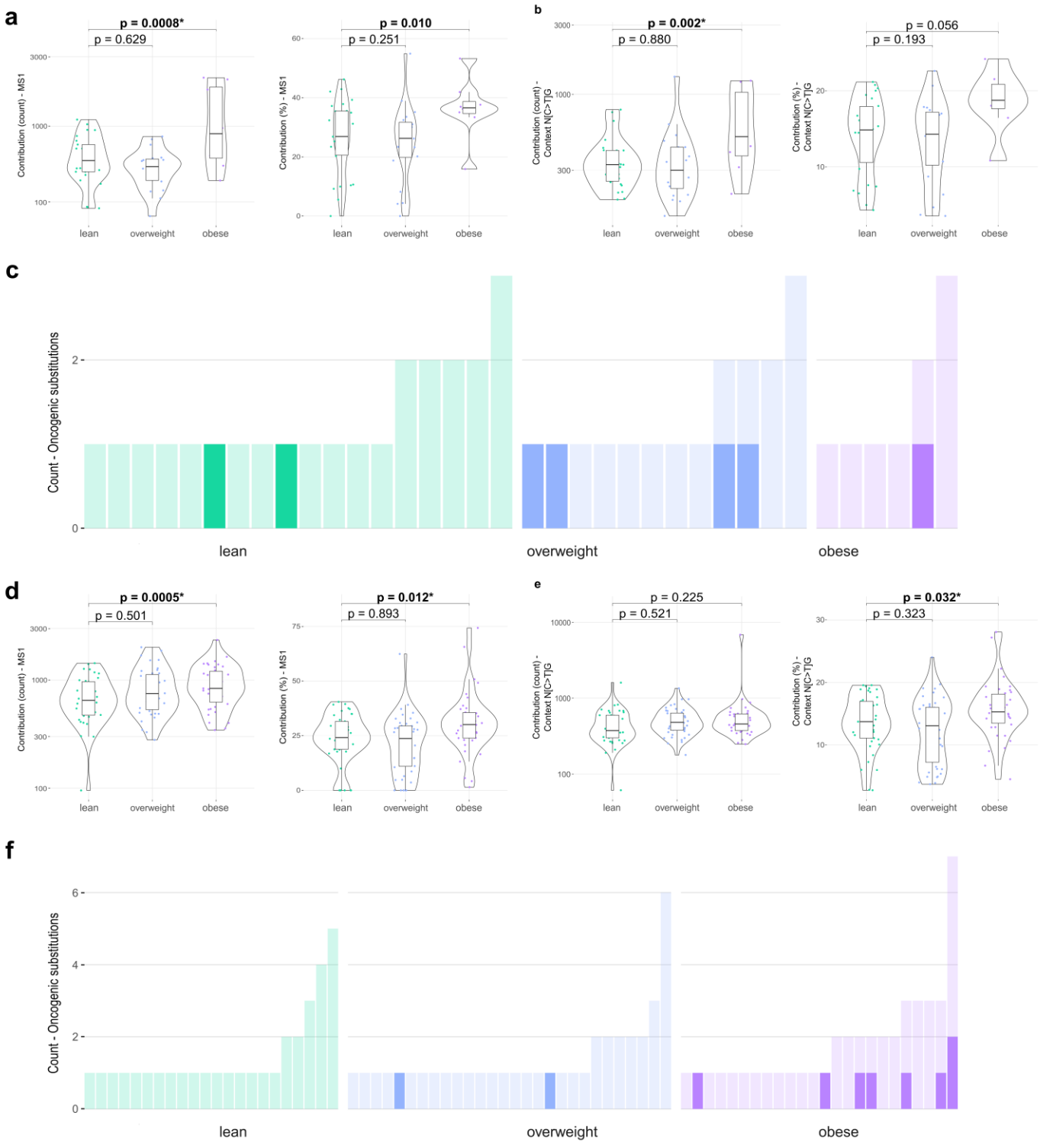
Supplementary Figure 6. Co-occurrence and mutual exclusivity of genomic alterations in NST ER+/HER2- tumors from patients of different BMI categories. Pairwise analyses of recurrent gene mutations and mutations, mutations and CNAs (either amp or hemiLoss) were performed. The color scale represents -log(q-value), which were derived from a Poisson-Binomial distribution-based test and adjusted for multiple testing using the Benjamin-Hochberg method. A grey cell indicates that one of the alterations in the pair did not have enough number of events (3 events) in the respective sub-cohort and thus was not analyzed. Merged data from two cohorts, METABRIC and ICGC, was used.



Supplementary Figure 7. Co-occurrence and mutual exclusivity of genomic alterations in NST ER-/HER2- tumors from patients of different BMI categories. Pairwise analyses of recurrent gene mutations and mutations, mutations and CNAs (either amp or hemiLoss) were performed. The color scale represents -log(q-value), which were derived from a Poisson-Binomial distribution-based test and adjusted for multiple testing using the Benjamin-Hochberg method. A grey cell indicates that one of the alterations in the pair did not have enough number of events (3 events) in the respective sub-cohort and thus was not analyzed. Merged data from two cohorts, METABRIC and ICGC, was used.

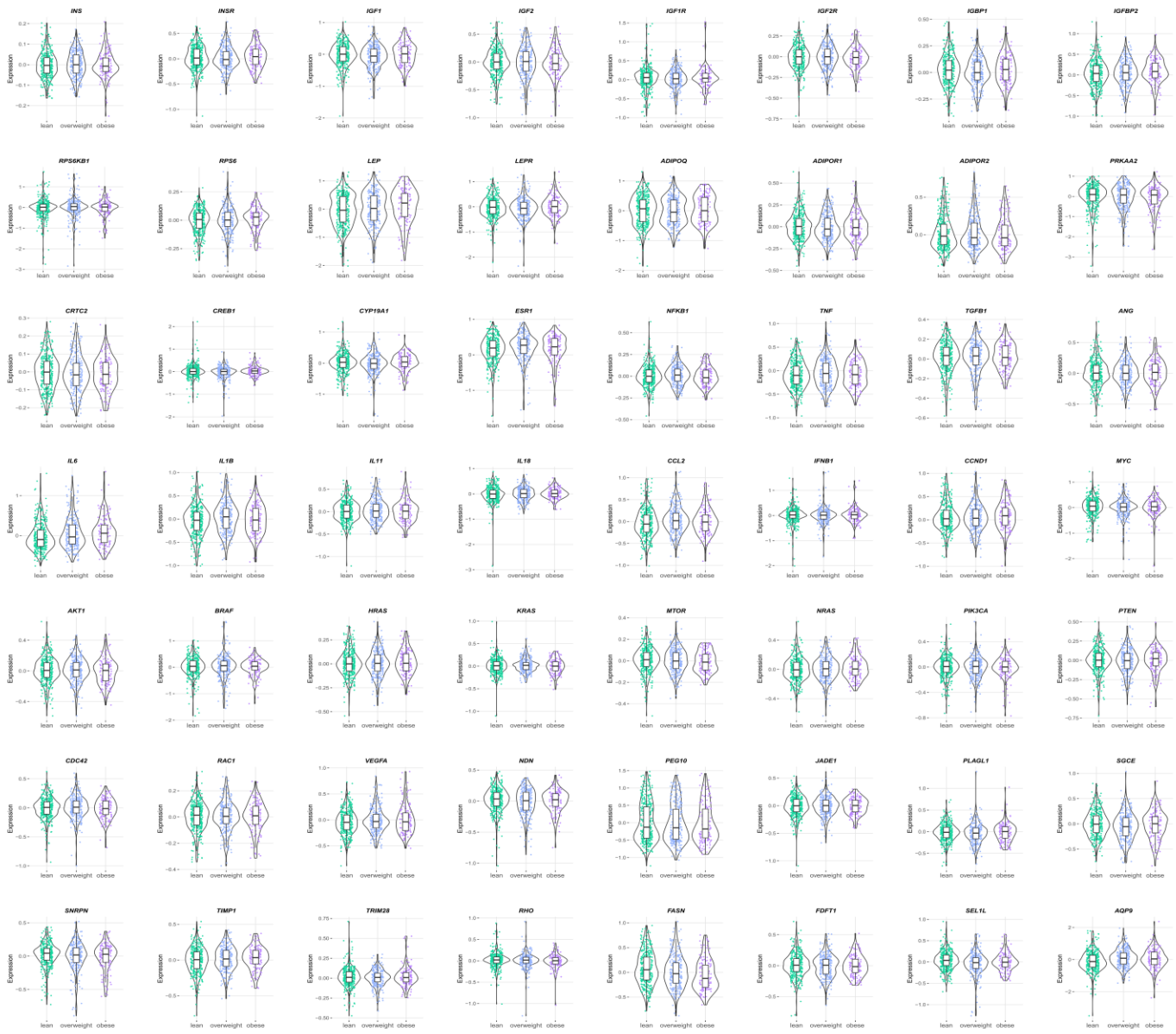


Supplementary Figure 8. Non-linear association of gene-level CNAs with BMI. Plots showing fitted lines of the logOR of gene-level CNAs against continuous BMI (baseline BMI of 20) for those indicated to be better associated with BMI in a non-linear model. Data are presented for point estimates of log-odds ratio (logOR; blue solid line) and their 95% confidence interval (grey shaded band). p-values shown were determined for coefficients estimated by generalized additive models fitted with a spline term of BMI and adjusted for cohort, age, and tumor grade. All statistical tests were two-sided.



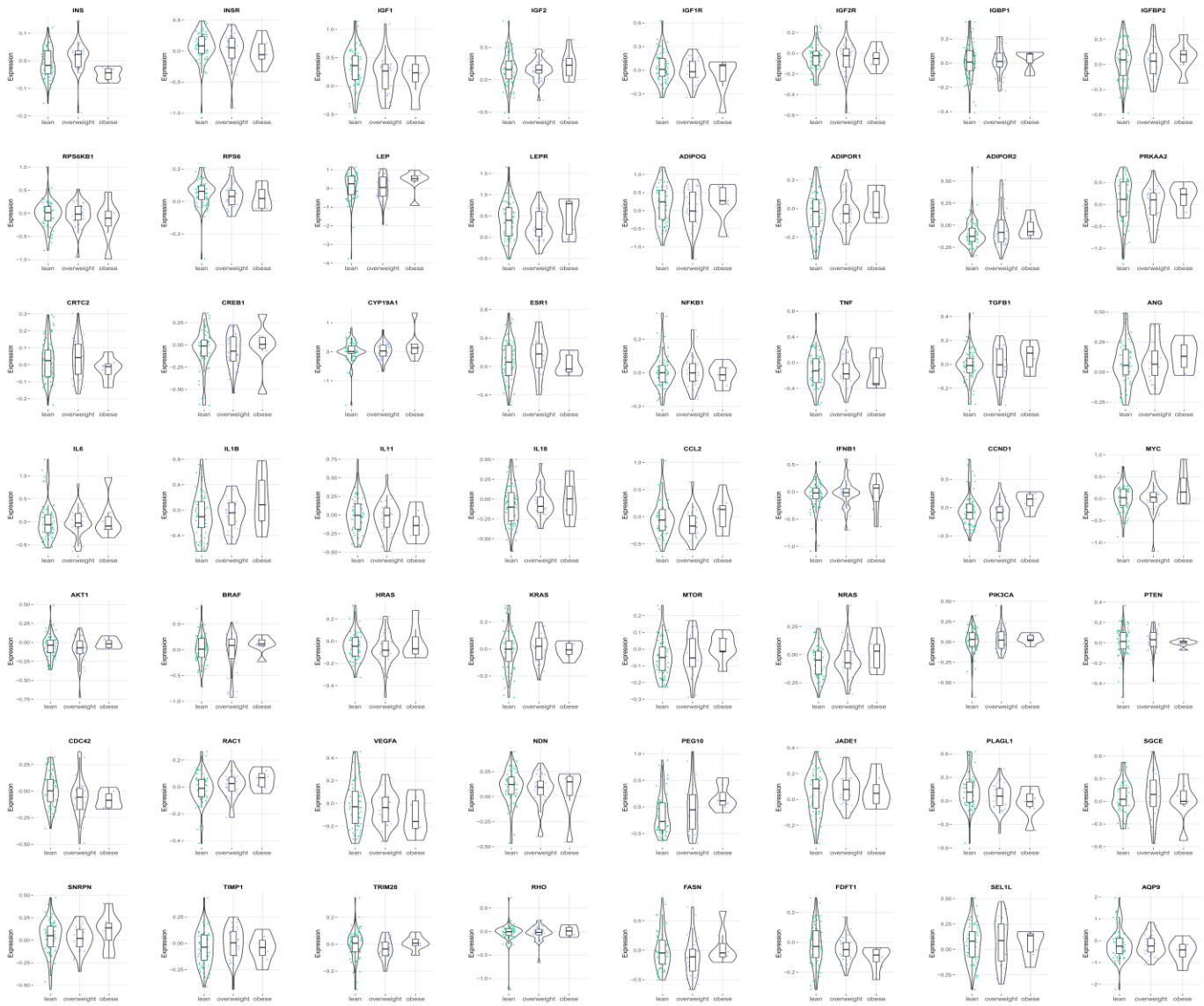
Supplementary Figure 9. Association of BMI with the age-associated mutational signature (Signature.1) in patients with NST ER+/HER2- in different age groups, ≤50 (a-c) and >50 years (d-f), from the ICGC cohort. **a-b, d-e** Violin/box plots showing the contribution in absolute count (left) and relative percentage (right) of the mutational signature 1 (MS1) (a, d) and the sequence context N[C>T]G (b, e) according to BMI categories (age ≤50 lean, n = 26; overweight, n = 20; obese, n = 8; age >50 lean, n = 41; overweight, n = 44; obese, n = 38). In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the

maxima and minima excluding outliers. All statistical tests were two-sided. Wald test p-values determined for coefficients estimated by linear regressions adjusted for age continuous) and tumor grade (G3 vs G1/G2), are reported and were not corrected for multiple testing. **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.

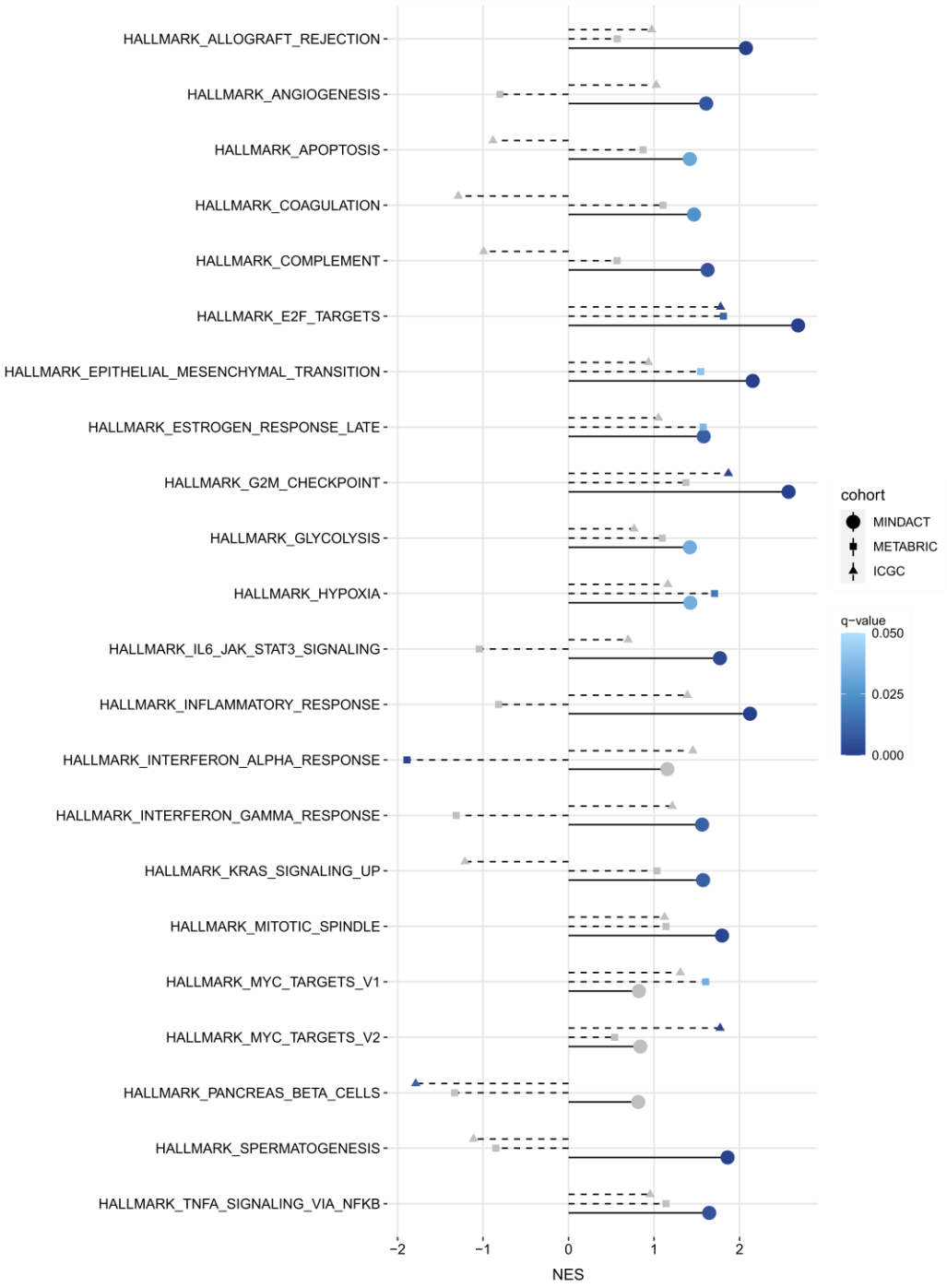


Supplementary Figure 10. Gene expression of relevant genes in the BC-obesity axis in NST ER+/HER2- tumors from patients of different BMI categories in the MINDACT cohort. Violin/box plots showing the expression levels of selected genes according to BMI categories (lean, n = 354; overweight, n = 250; obese, n = 131). The expression values (log10-transformed) presented here had been normalized prior to data retrieval. In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the maxima and minima excluding outliers.

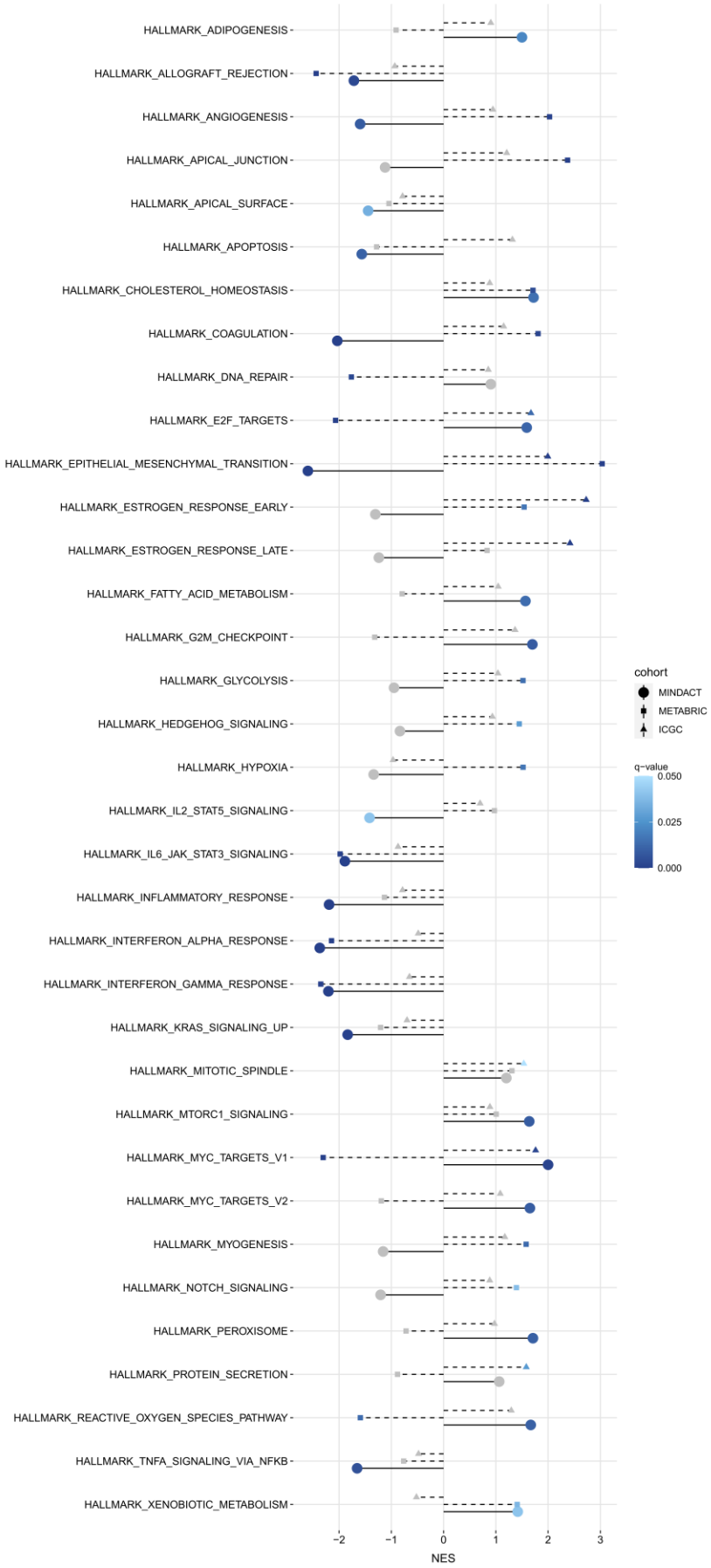
Supplementary Figure 11. Gene expression of relevant genes in the BC-obesity axis in NST ER-/HER2- tumors from patients of different BMI categories in the MINDACT cohort. Violin/box plots showing the expression levels of selected genes according to BMI categories (lean, n = 53; overweight, n = 54; obese, n = 11). The expression values (log10-transformed) presented here had been normalized prior to data retrieval. In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the maxima and minima excluding outliers.



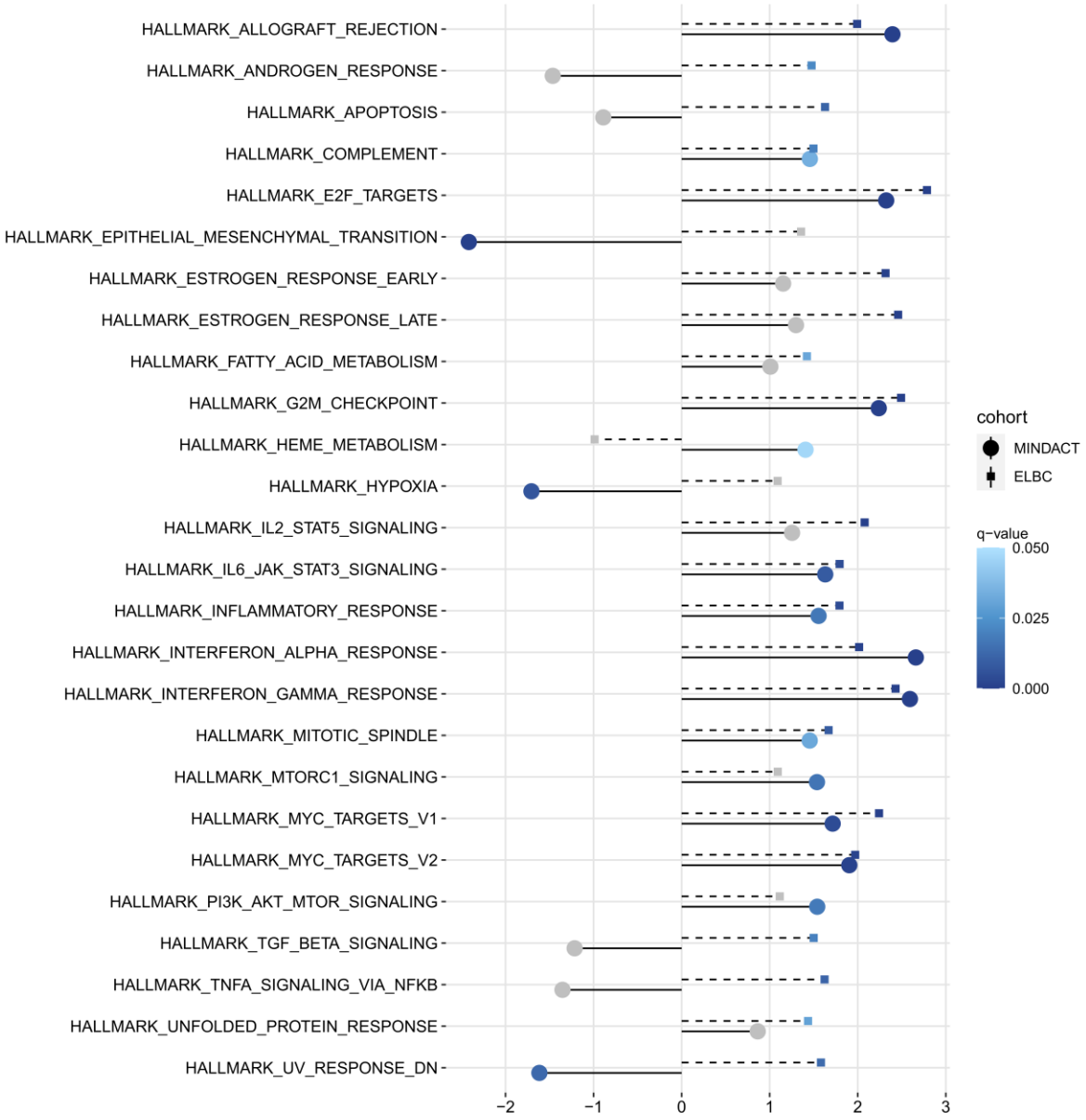
Supplementary Figure 12. Gene expression of relevant genes in the BC-obesity axis in ILC ER+/HER2- tumors from patients of different BMI categories in the MINDACT cohort. Violin/box plots showing the expression levels of selected genes according to BMI categories (lean, n = 65; overweight, n = 32; obese, n = 7). The expression values (log10-transformed) presented here had been normalized prior to data retrieval. In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the maxima and minima excluding outliers.



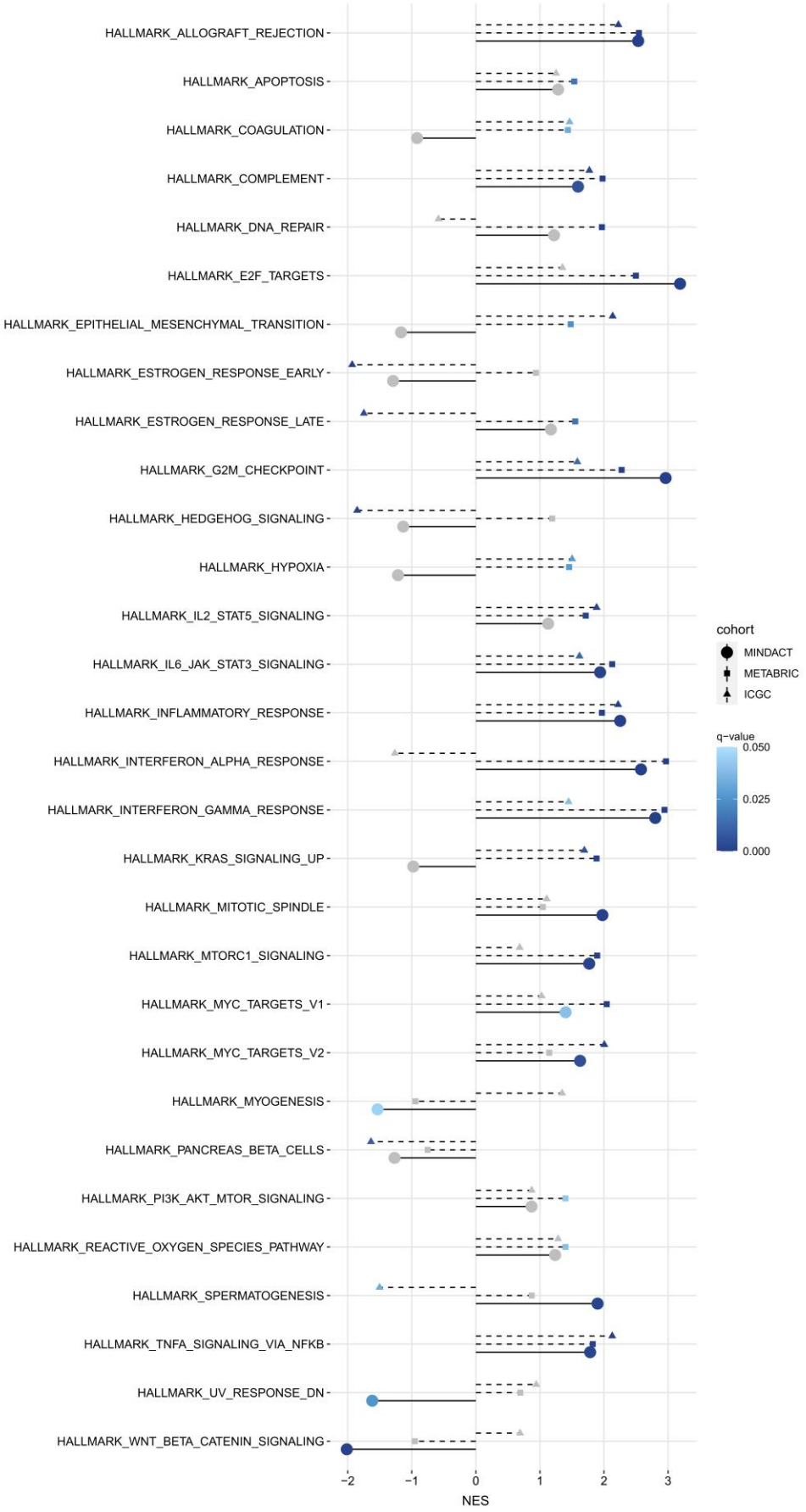
Supplementary Figure 13. Differentially enriched hallmarks in NST ER+/HER2- tumors from obese patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized enrichment score (NES) indicate the orientation of the differential enrichments (positive: enriched in tumors from obese patients, negative: enriched in tumors from lean patients). Lollipops with solid segments represent the main cohort discussed in the main text (i.e. MINDACT) and those with dashed segments represent the other cohorts.


HALLMARK_ADIPOGENESIS
HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_ANGIOGENESIS
HALLMARK_APICAL_JUNCTION
HALLMARK_APICAL_SURFACE
HALLMARK_APOPTOSIS
HALLMARK_CHOLESTEROL_HOMEOSTASIS
HALLMARK_COAGULATION
HALLMARK_DNA_REPAIR
HALLMARK_E2F_TARGETS
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
HALLMARK_ESTROGEN_RESPONSE_EARLY
HALLMARK_ESTROGEN_RESPONSE_LATE
HALLMARK_FATTY_ACID_METABOLISM
HALLMARK_G2M_CHECKPOINT
HALLMARK_GLYCOLYSIS
HALLMARK_HEDGEHOG_SIGNALING
HALLMARK_HYPOXIA
HALLMARK_IL2_STAT5_SIGNALING
HALLMARK_IL6_JAK_STAT3_SIGNALING
HALLMARK_INFLAMMATORY_RESPONSE
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_KRAS_SIGNALING_UP
HALLMARK_MITOTIC_SPINDLE
HALLMARK_MTORC1_SIGNALING
HALLMARK_MYC_TARGETS_V1
HALLMARK_MYC_TARGETS_V2
HALLMARK_MYOGENESIS
HALLMARK_NOTCH_SIGNALING
HALLMARK_PEROXISOME
HALLMARK_PROTEIN_SECRETION
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY
HALLMARK_TNFA_SIGNALING_VIA_NFKB
HALLMARK_XENOBIOTIC_METABOLISM
−2
−1
0
1
2
3
NES
cohort
MINDACT
METABRIC
ICGC
q−value
0.050
0.025
0.000

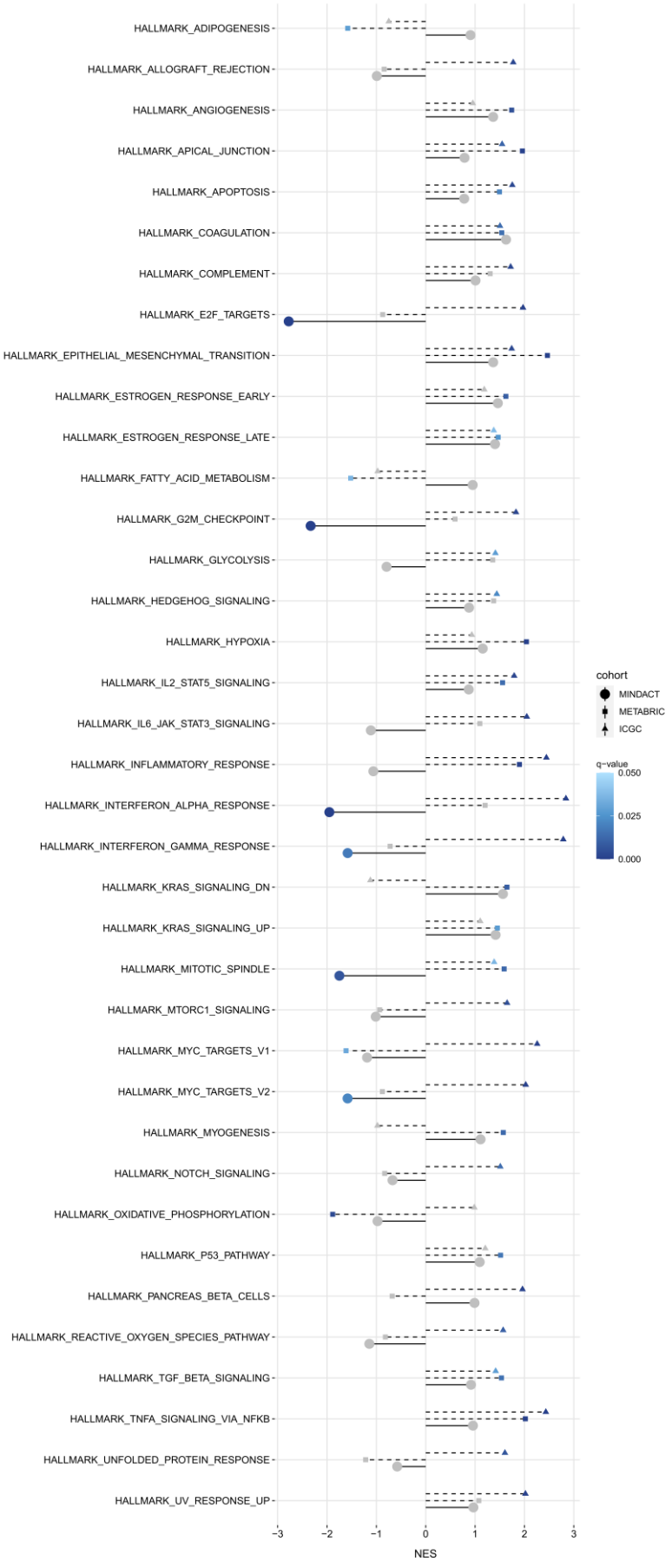
Supplementary Figure 14. Differentially enriched hallmarks in NST ER-/HER2- tumors from obese patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized enrichment score (NES) indicate the orientation of the differential enrichments (positive: enriched in tumors from obese patients, negative: enriched in tumors from lean patients). Lollipops with solid segments represent the main cohort discussed in the main text (i.e. MINDACT) and those with dashed segments represent the other cohorts.



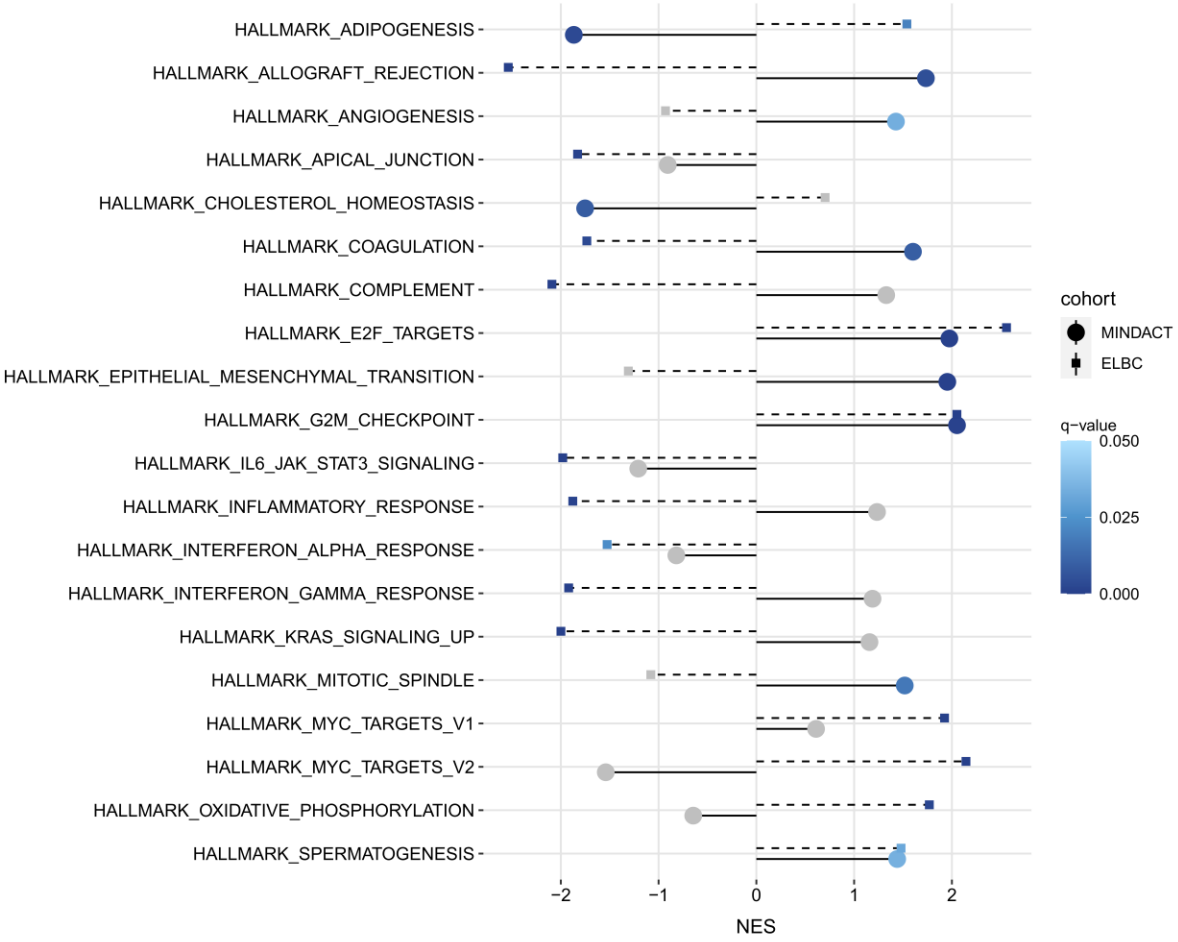
Supplementary Figure 15. Differentially enriched hallmarks in ILC ER+/HER2- tumors from obese patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized enrichment score (NES) indicate the orientation of the differential enrichments (positive: enriched in tumors from obese patients, negative: enriched in tumors from lean patients). Lollipops with solid segments represent the main cohort discussed in the main text (i.e. MINDACT) and those with dashed segments represent the other cohorts.


HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_APOPTOSIS
HALLMARK_COAGULATION
HALLMARK_COMPLEMENT
HALLMARK_DNA_REPAIR
HALLMARK_E2F_TARGETS
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
HALLMARK_ESTROGEN_RESPONSE_EARLY
HALLMARK_ESTROGEN_RESPONSE_LATE
HALLMARK_G2M_CHECKPOINT
HALLMARK_HEDGEHOG_SIGNALING
HALLMARK_HYPOXIA
HALLMARK_IL2_STAT5_SIGNALING
HALLMARK_IL6_JAK_STAT3_SIGNALING
HALLMARK_INFLAMMATORY_RESPONSE
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_KRAS_SIGNALING_UP
HALLMARK_MITOTIC_SPINDLE
HALLMARK_MTORC1_SIGNALING
HALLMARK_MYC_TARGETS_V1
HALLMARK_MYC_TARGETS_V2
HALLMARK_MYOGENESIS
HALLMARK_PANCREAS_BETA_CELLS
HALLMARK_PI3K_AKT_MTOR_SIGNALING
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY
HALLMARK_SPERMATOGENESIS
HALLMARK_TNFA_SIGNALING_VIA_NFKB
HALLMARK_UV_RESPONSE_DN
HALLMARK_WNT_BETA_CATENIN_SIGNALING
−2
−1
0
1
2
3
NES
cohort
MINDACT
METABRIC
ICGC
q−value
0.050
0.025
0.000

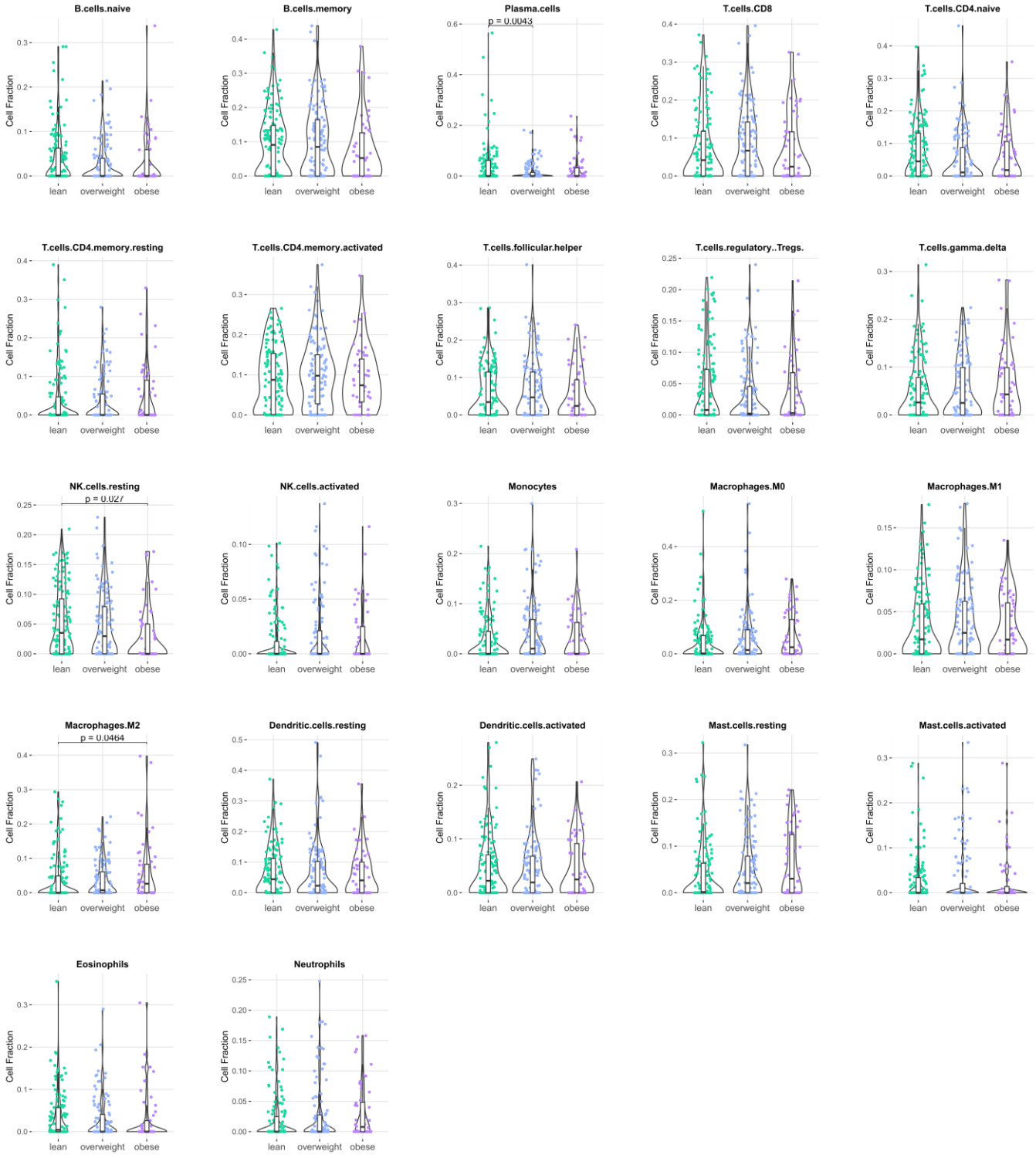
Supplementary Figure 16. Differentially enriched hallmarks in NST ER+/HER2- tumors from overweight patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized enrichment score (NES) indicate the orientation of the differential enrichments (positive: enriched in tumors from overweight patients, negative: enriched in tumors from lean patients). Lollipops with solid segments represent the main cohort discussed in the main text (i.e. MINDACT) and those with dashed segments represent the other cohorts.


HALLMARK_ADIPOGENESIS
HALLMARK_ALLOGRAFT_REJECTION
HALLMARK_ANGIOGENESIS
HALLMARK_APICAL_JUNCTION
HALLMARK_APOPTOSIS
HALLMARK_COAGULATION
HALLMARK_COMPLEMENT
HALLMARK_E2F_TARGETS
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
HALLMARK_ESTROGEN_RESPONSE_EARLY
HALLMARK_ESTROGEN_RESPONSE_LATE
HALLMARK_FATTY_ACID_METABOLISM
HALLMARK_G2M_CHECKPOINT
HALLMARK_GLYCOLYSIS
HALLMARK_HEDGEHOG_SIGNALING
HALLMARK_HYPOXIA
HALLMARK_IL2_STAT5_SIGNALING
HALLMARK_IL6_JAK_STAT3_SIGNALING
HALLMARK_INFLAMMATORY_RESPONSE
HALLMARK_INTERFERON_ALPHA_RESPONSE
HALLMARK_INTERFERON_GAMMA_RESPONSE
HALLMARK_KRAS_SIGNALING_DN
HALLMARK_KRAS_SIGNALING_UP
HALLMARK_MITOTIC_SPINDLE
HALLMARK_MTORC1_SIGNALING
HALLMARK_MYC_TARGETS_V1
HALLMARK_MYC_TARGETS_V2
HALLMARK_MYOGENESIS
HALLMARK_NOTCH_SIGNALING
HALLMARK_OXIDATIVE_PHOSPHORYLATION
HALLMARK_P53_PATHWAY
HALLMARK_PANCREAS_BETA_CELLS
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY
HALLMARK_TGF_BETA_SIGNALING
HALLMARK_TNFA_SIGNALING_VIA_NFKB
HALLMARK_UNFOLDED_PROTEIN_RESPONSE
HALLMARK_UV_RESPONSE_UP
−3
−2
−1
0
1
2
3
NES
cohort
MINDACT
METABRIC
ICGC
q−value
0.050
0.025
0.000

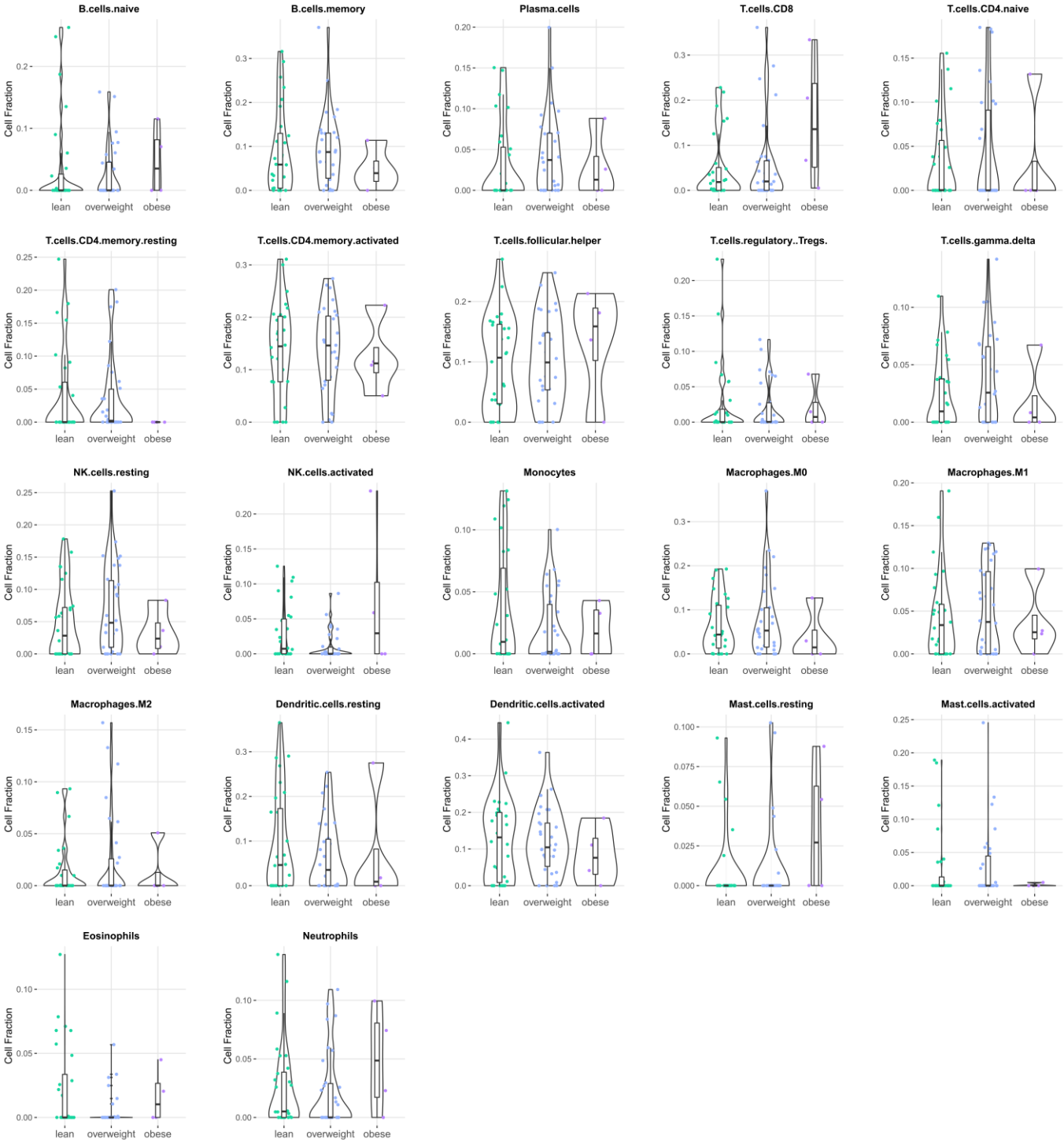
Supplementary Figure 17. Differentially enriched hallmarks in NST ER-/HER2- tumors from overweight patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized enrichment score (NES) indicate the orientation of the differential enrichments (positive: enriched in tumors from overweight patients, negative: enriched in tumors from lean patients). Lollipops with solid segments represent the main cohort discussed in the main text (i.e. MINDACT) and those with dashed segments represent the other cohorts.



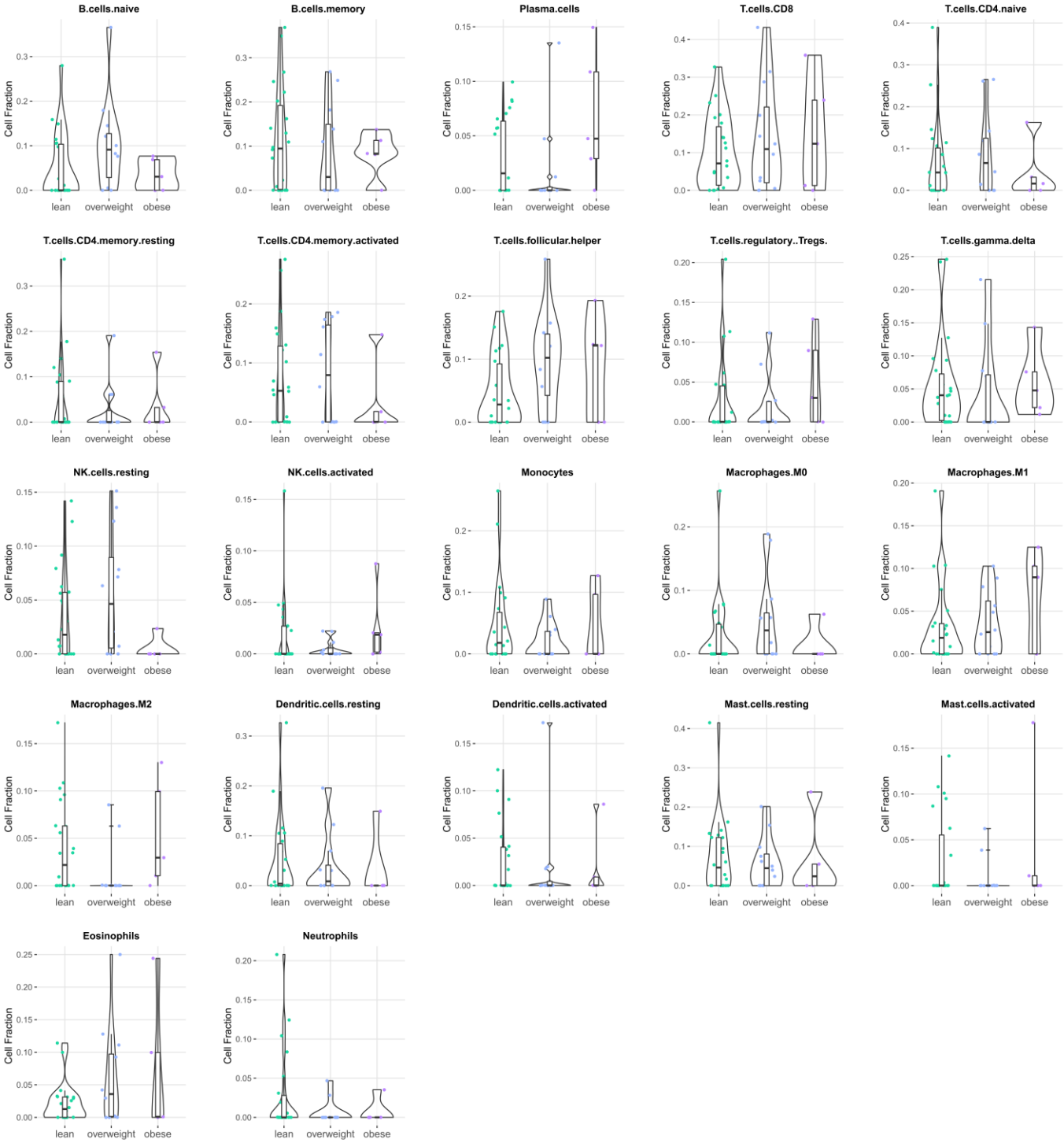
Supplementary Figure 18. Differentially enriched hallmarks in ILC ER+/HER2- tumors from overweight patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized enrichment score (NES) indicate the orientation of the differential enrichments (positive: enriched in tumors from overweight patients, negative: enriched in tumors from lean patients). Lollipops with solid segments represent the main cohort discussed in the main text (i.e. MINDACT) and those with dashed segments represent the other cohorts.



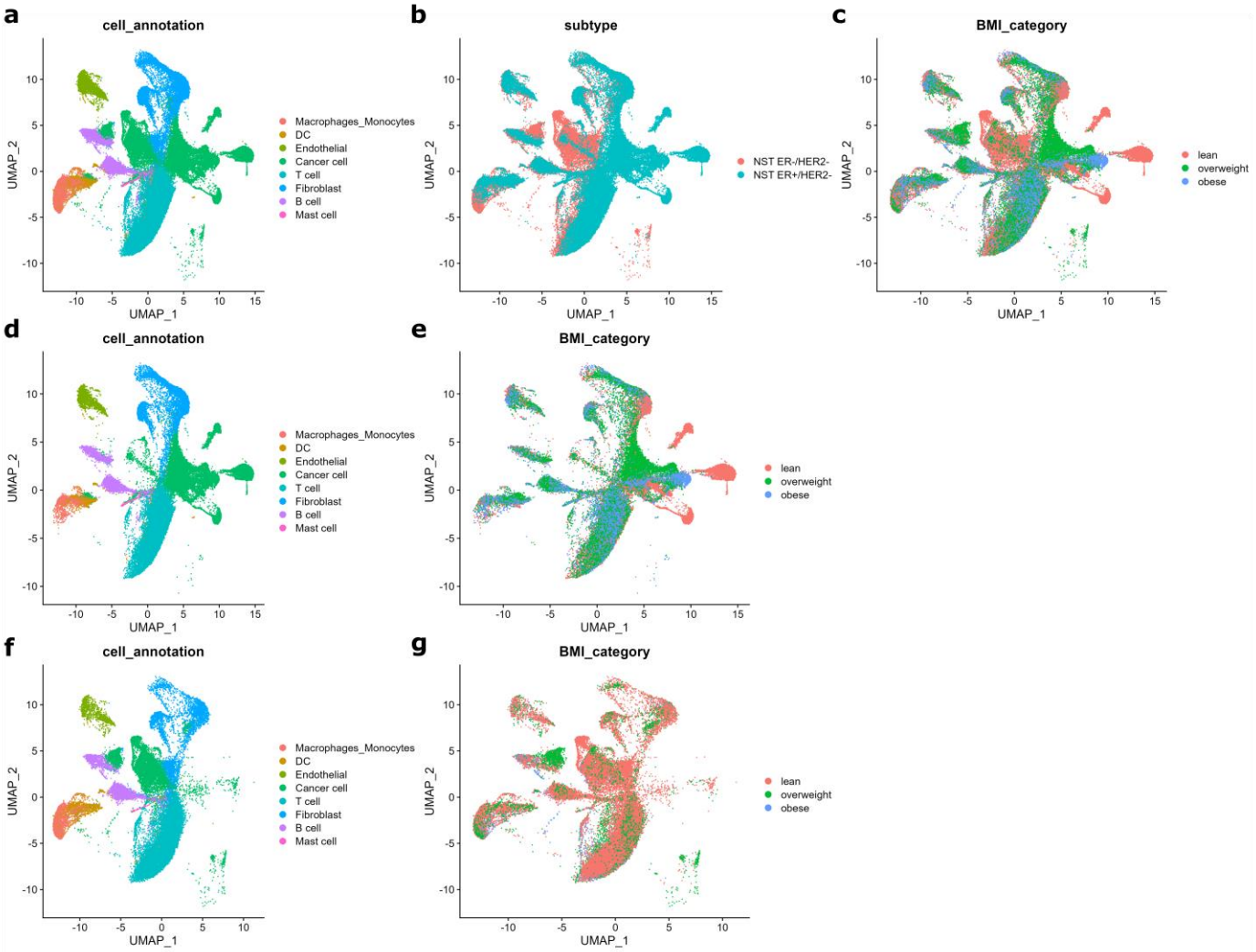
Supplementary Figure 19. Cell fractions of 22 immune cell types in the tumor bulk of NST ER+/HER2- tumors from the MINDACT cohort. Violin/box plots showing relative frequency of 22 immune cells inferred by deconvolution of the bulk profiling data using CIBERSORTx according to BMI categories (lean, n = 354; overweight, n = 250; obese, n = 65). In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the maxima and minima excluding outliers. Comparisons with p-value < 0.05 are indicated. p-values shown were derived from Wilcoxon's ranked sum tests. All statistical tests were two-sided.



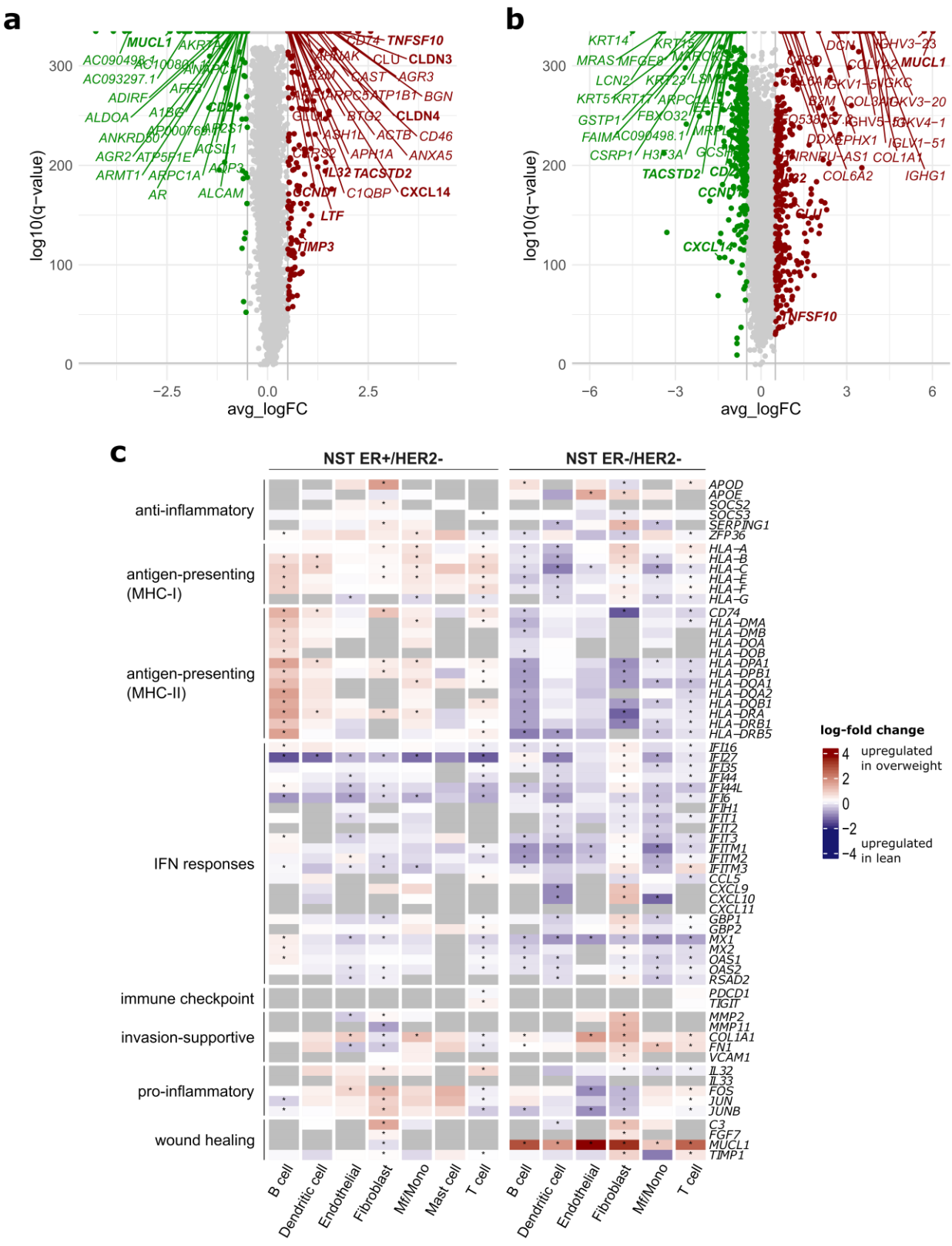
Supplementary Figure 20. Cell fractions of 22 immune cell types in the tumor bulk of NST ER-/HER2- tumors from the MINDACT cohort. Violin plots showing relative frequency of 22 immune cells inferred by deconvolution of the bulk profiling data using CIBERSORTx according to BMI categories (lean, n = 53; overweight, n = 54; obese, n = 11). In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the maxima and minima excluding outliers.



Supplementary Figure 21. Cell fractions of 22 immune cell types in the tumor bulk of ILC ER+/HER2- tumors from the MINDACT cohort. Violin plots showing relative frequency of 22 immune cells inferred by deconvolution of the bulk profiling data using CIBERSORTx according to BMI categories (lean, n = 65; overweight, n = 32; obese, n = 7). In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the maxima and minima excluding outliers.



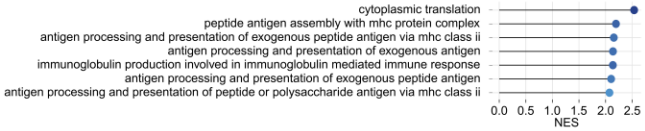
Supplementary Figure 22. Uniform Manifold Approximation and Projection (UMAP) of cells in tumors from NST ER+/HER2- and NST ER-/HER2- patients from the Biokey cohort. a-c UMAP of cells from all patients color-coded according to cell type (a), tumor subtype (b), and BMI category of the patient (c). **d-g** UMAPs of cells from NST ER+/HER2- (d-e) and NST ER-/HER2- (f-g) color-coded according to cell type and BMI category.



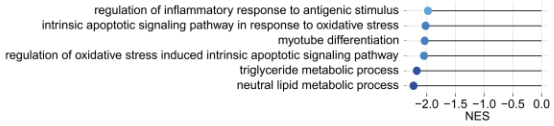
Supplementary Figure 23. Cell type-specific differential gene expression in overweight versus lean patients with NST ER+/HER2- and NST ER-/HER2- from the BioKey cohort. a-b Volcano plots highlighting cancer cell-specific DEGs between overweight and lean in the NST ER+/HER2- (a) and NST ER-/HER2- (b) subgroups. Genes with absolute logFC > 0.5 and q-value < 0.05 are colored (red: upregulated in cancer cells from overweight patients, green: upregulated in

cancer cells from lean patients). The top 20 up-regulated (sorted by q-value), 20 down-regulated genes, and genes discussed in the main text (in bold) are labeled. **c** Heatmaps showing differential expression of a selection of genes involved in several immune and cancer pathways in non-malignant cell populations. The cell color is scaled based on the log-fold change (logFC) values (overweight vs lean) estimated by the MAST test. Gray cells indicate genes not being tested due to expression in less than 10% of the corresponding cell type in both BMI categories. *, q-value < 0.05

B cell

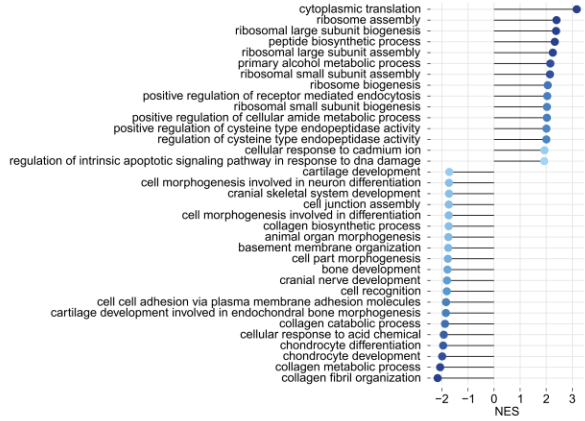


Denditic cell

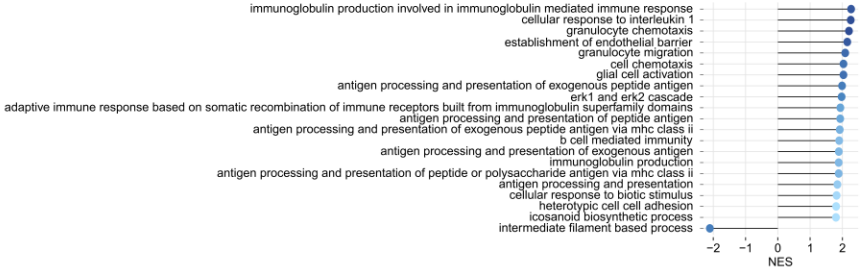


Endothelial cell

Fibroblast

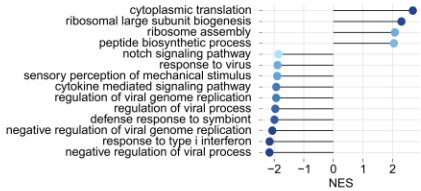


Mf/Mono

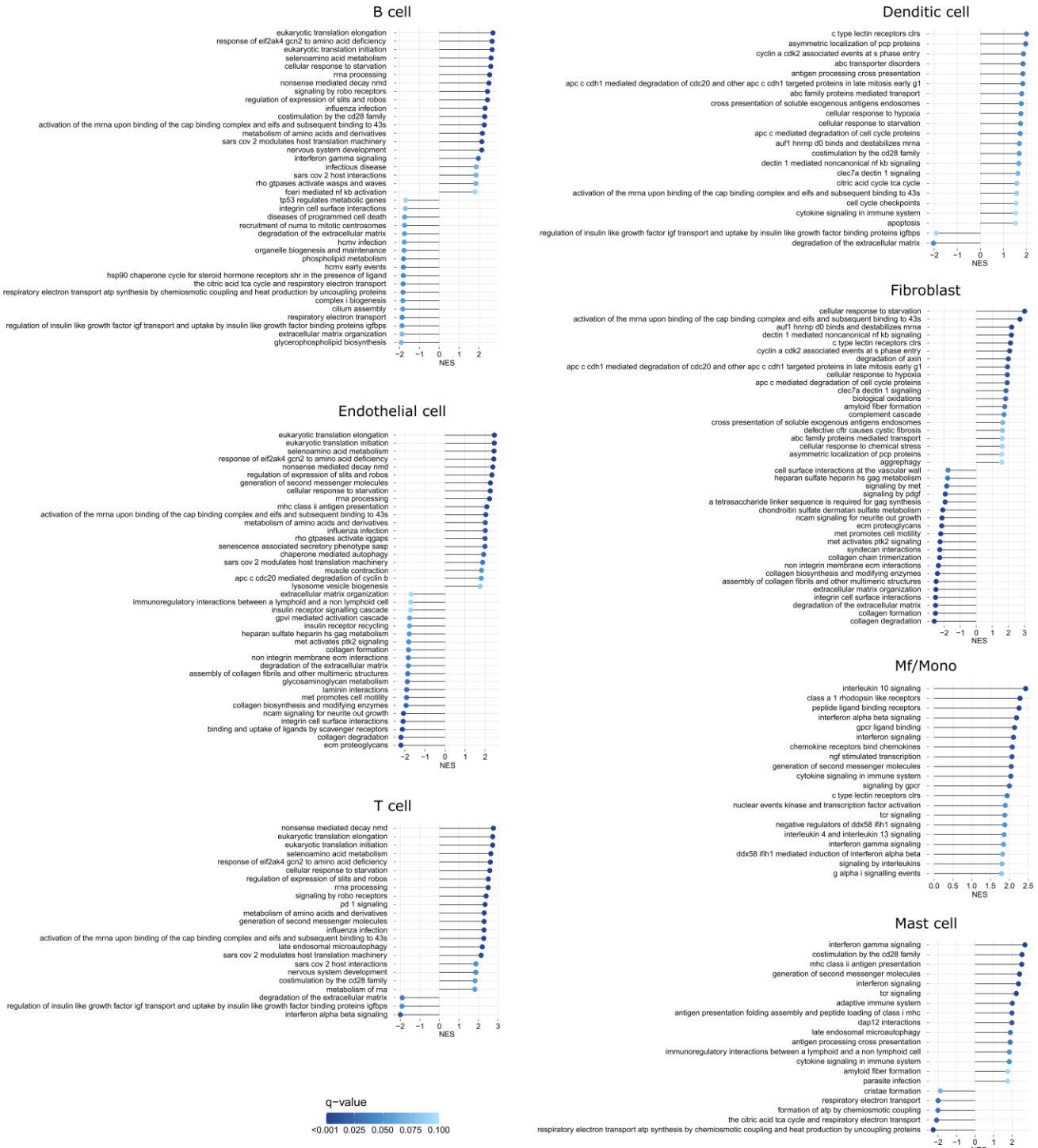


Mast cell

T cell

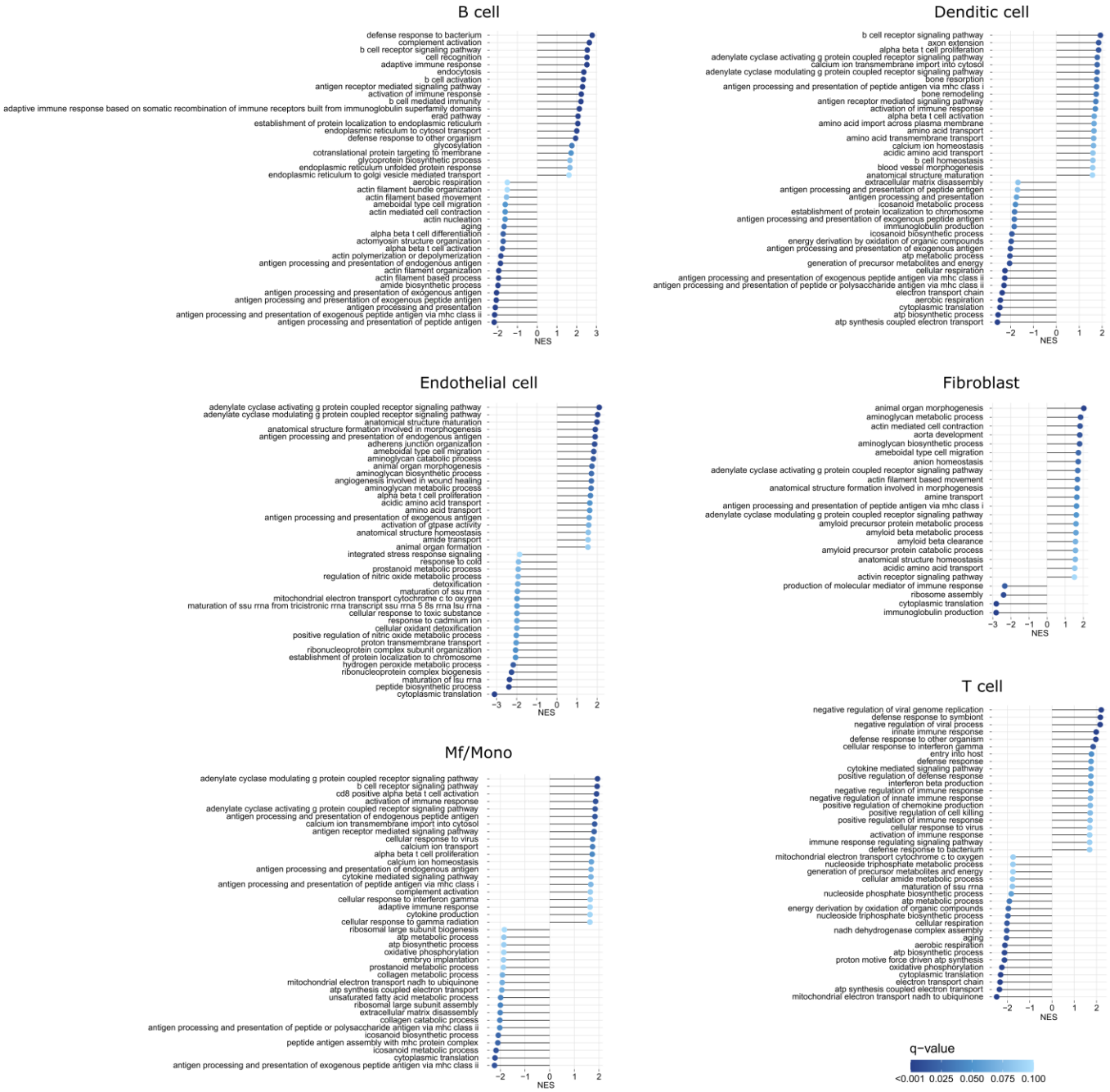


Supplementary Figure 24. Gene set enrichment analysis (GSEA) on Gene Ontology biological process (GOBP) gene sets of non-malignant cells from obese versus lean patients with NST ER+/HER2- from the BioKey cohort. Lollipop plots displaying top 20 upregulated (sorted by q-value, NES > 0) and top 20 downregulated (NES < 0) GOBP gene sets according to BMI category (obese vs lean) detected by GSEA (q-value < 0.1). p-values were computed by permutation and adjusted for multiple testing using the Benjamin-Hochberg method (presented as q-values).

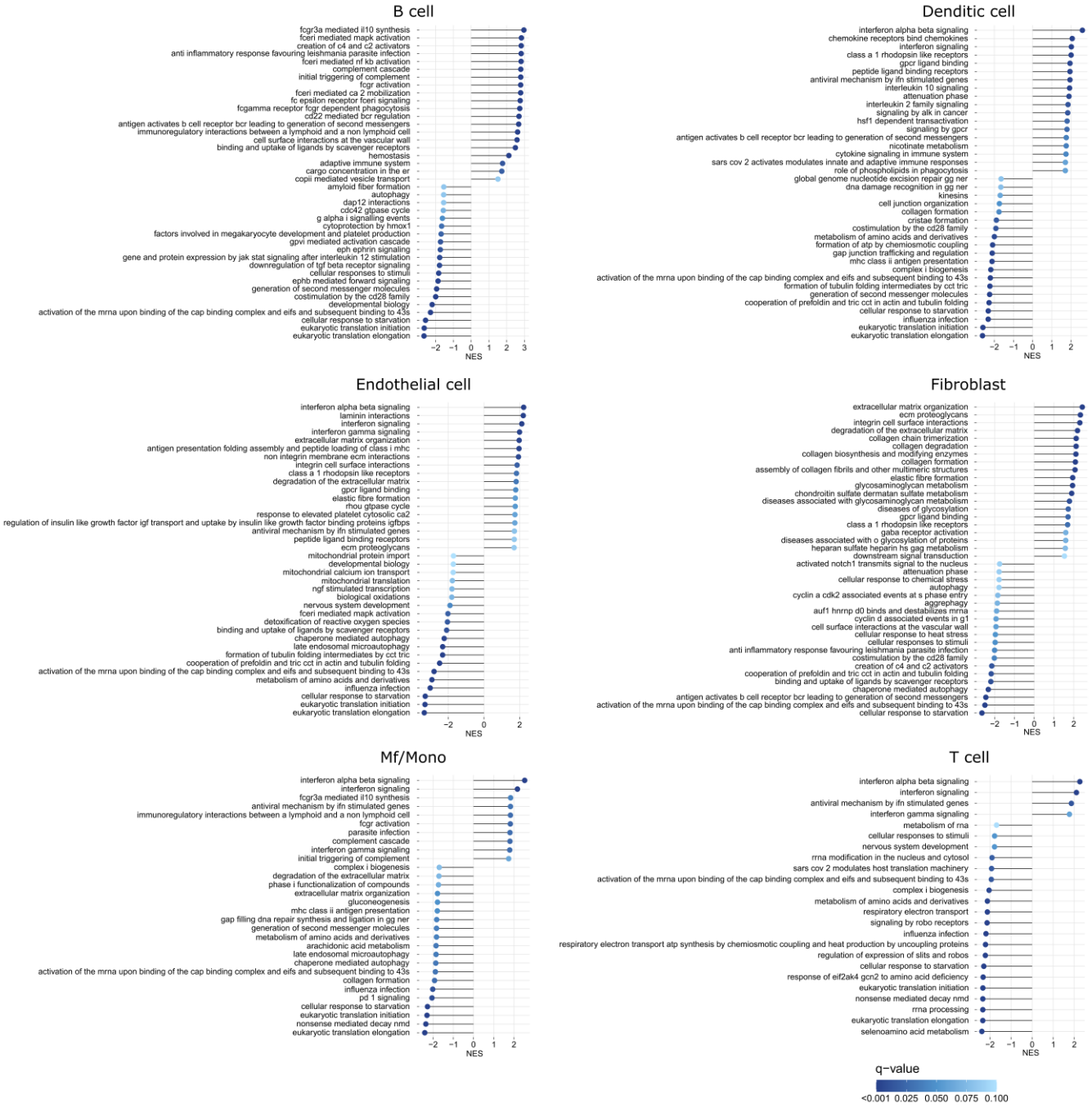


Supplementary Figure 25. Gene set enrichment analysis (GSEA) on REACTOME gene sets of non-malignant cells from obese versus lean patients with NST ER+/HER2- from the BioKey cohort. Lollipop plots displaying top 20 upregulated (sorted by q-value, NES > 0) and top 20 downregulated (NES < 0) REACTOME gene sets according to BMI category (obese vs lean) detected by GSEA (q-value < 0.1). p-values

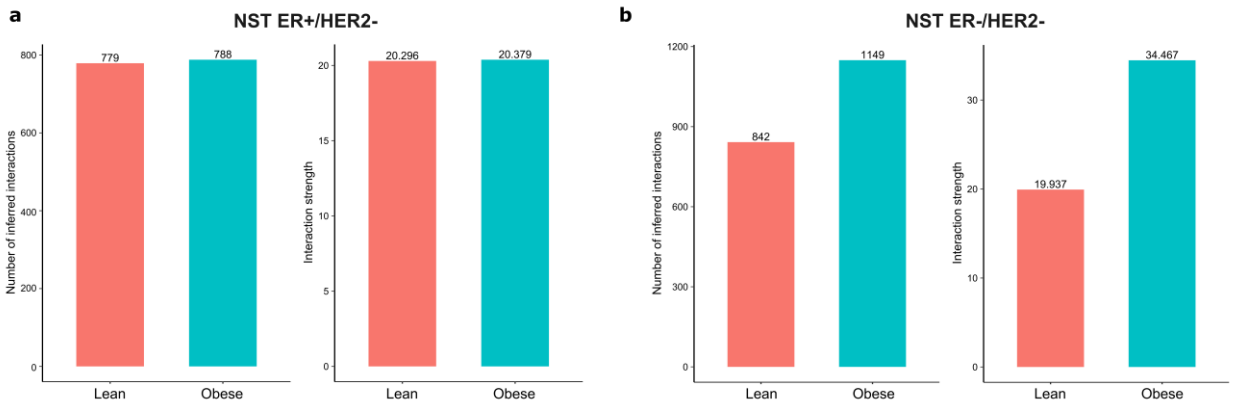
were computed by permutation and adjusted for multiple testing using the Benjamin-Hochberg method (presented as q-values).



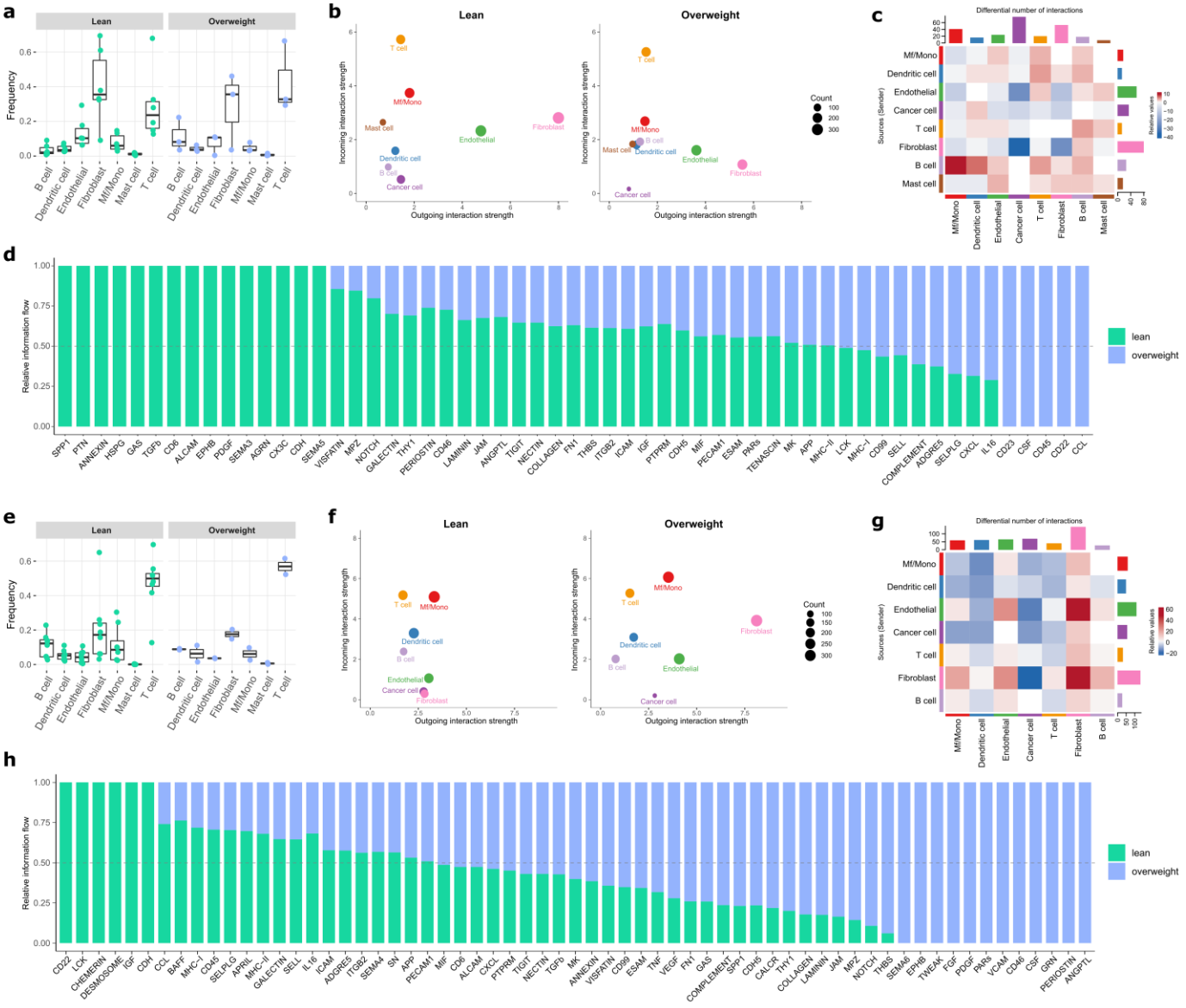
Supplementary Figure 26. Gene set enrichment analysis (GSEA) on GO biological process gene sets of non-malignant cells from obese versus lean patients with NST ER-/HER2- from the BioKey cohort. Lollipop plots displaying top 20 upregulated (sorted by q-value, NES > 0) and top 20 downregulated (NES < 0) GOBP gene sets according to BMI category (obese vs lean) detected by GSEA (q-value < 0.1). p-values were computed by permutation and adjusted for multiple testing using the Benjamin-Hochberg method (presented as q-values).



Supplementary Figure 27. Gene set enrichment analysis (GSEA) on REACTOME gene sets of non-malignant 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). p-values were computed by permutation and adjusted for multiple testing using the Benjamin-Hochberg method (presented as q-values).



Supplementary Figure 28. 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.



Supplementary Figure 29. Differences in the cell composition and intercellular interactions within the tumor microenvironment between overweight and lean patients with NST ER+/HER2- (a-d) and NST ER-/HER2- (e-h) tumors from the BioKey cohort. a, e Frequency of non-malignant cell types relative to the non-malignant cell pool for lean and overweight patients (NST ER+/HER2- lean, n = 6; overweight, n = 3; NST ER-/HER2- lean, n = 8; overweight, n = 2). In each boxplot, the box denotes the range from the 25th to the 75th percentile, the center line indicates the median value, and the whiskers specify the maxima and minima excluding outliers. **b, f** CellChat-derived outgoing and incoming interaction strength from and to cancer cells and non-malignant cells in tumors from lean and overweight patients. **c, g** Differential number of interactions detected by CellChat between lean and overweight. **d, h** Relative Information flow of signaling pathways in the intercellular communication network in tumors from lean and overweight patients. The relative information flow was estimated by the sum of CellChat-derived communication

probability between all pairs of cell compartments in the network. Signaling pathways with non-zero information in at least one of the BMI categories are shown.