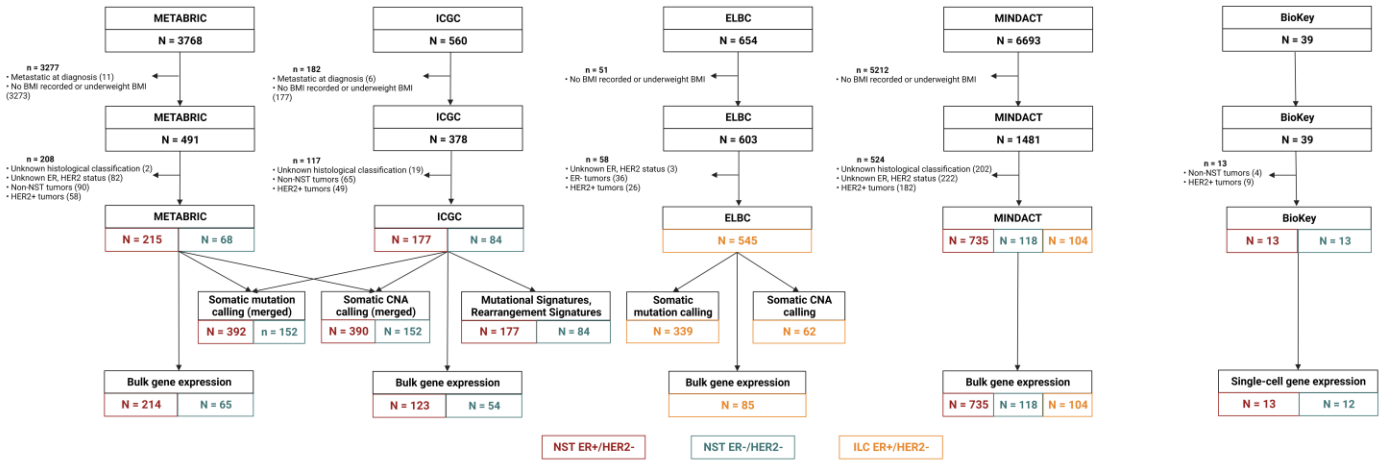
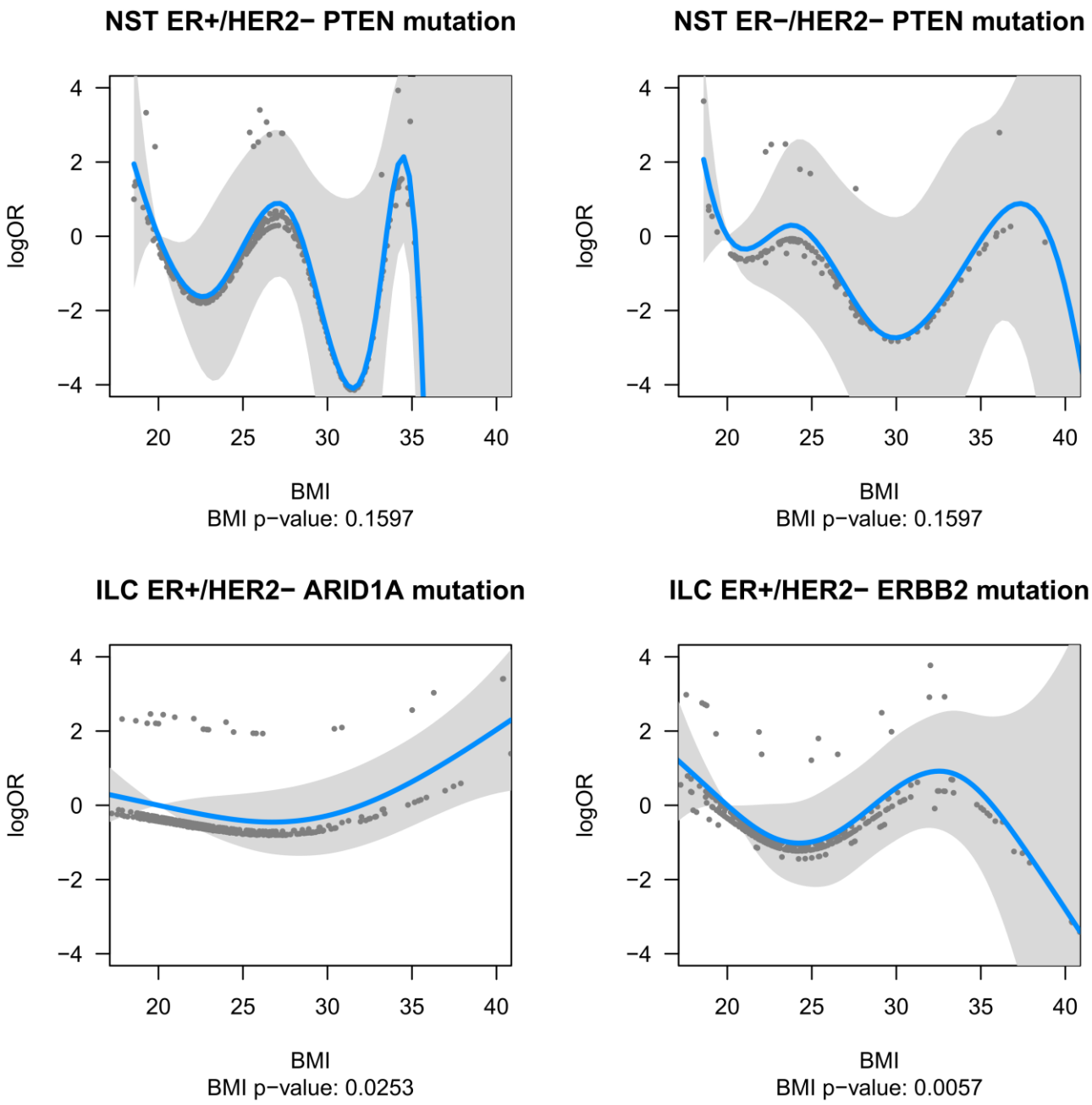
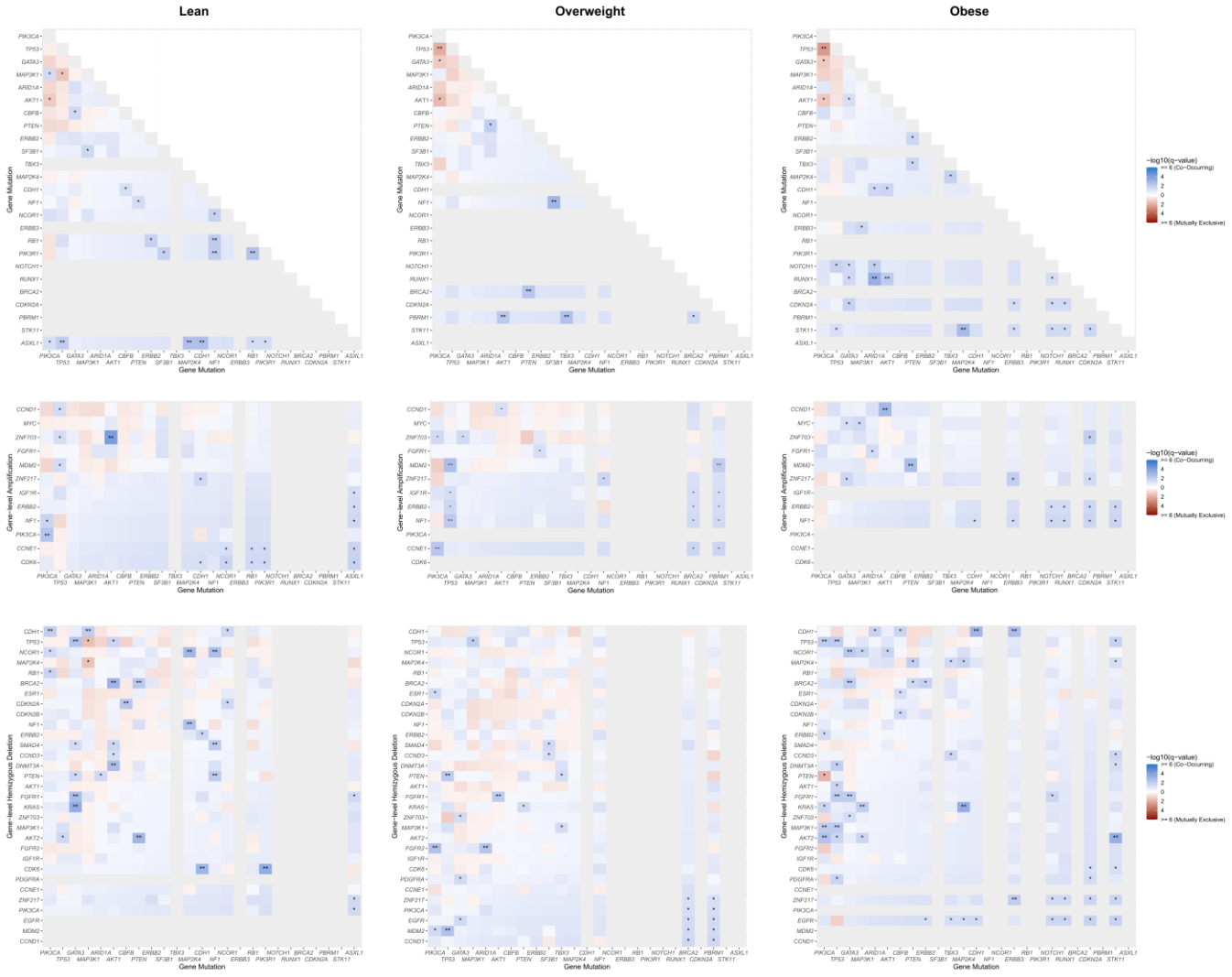


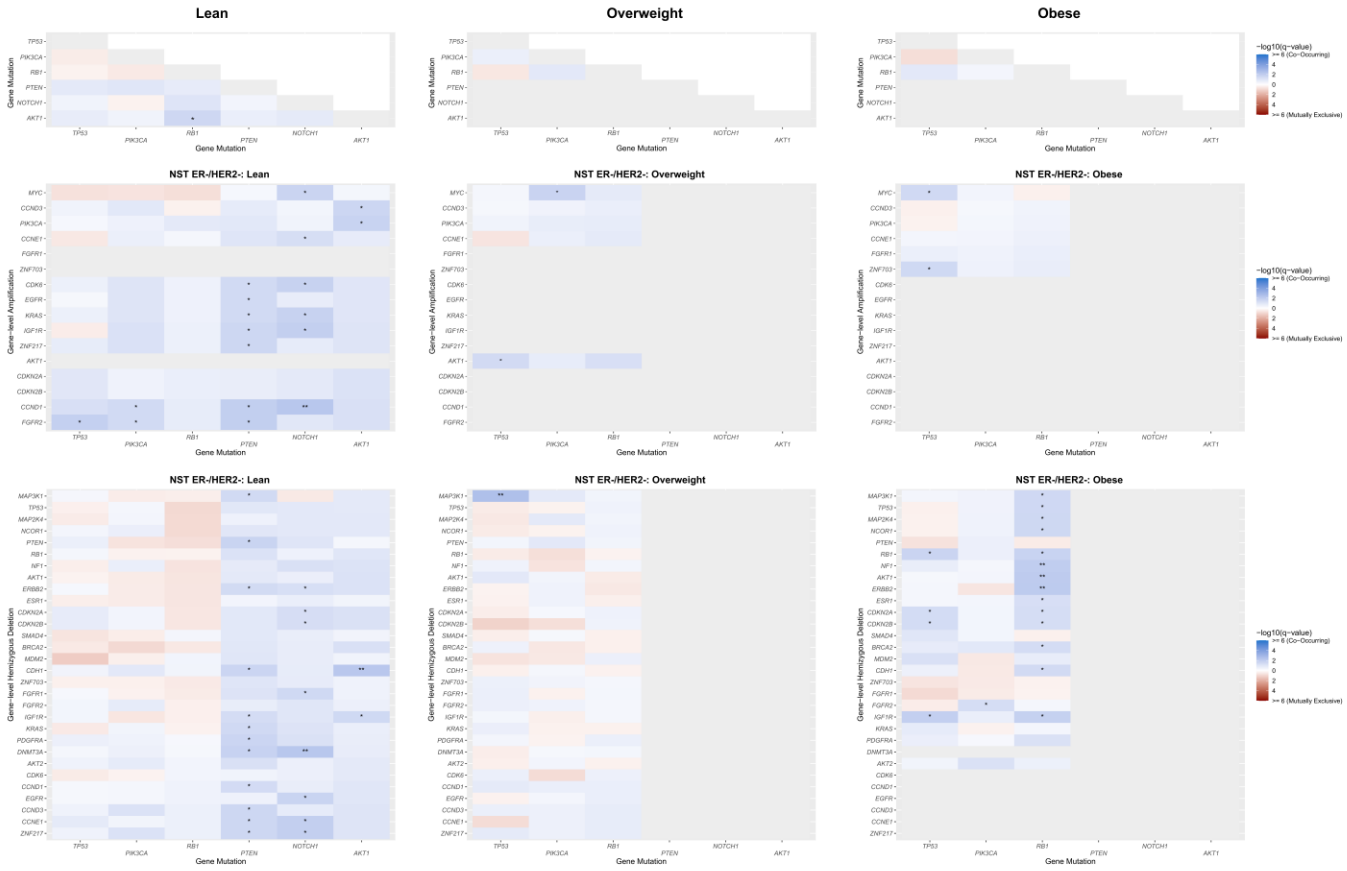
Supplementary Figure 1. Overview of the data flow in the study. Flowchart displaying the data selection and stratification process and showing the final numbers of cases analyzed in the study.



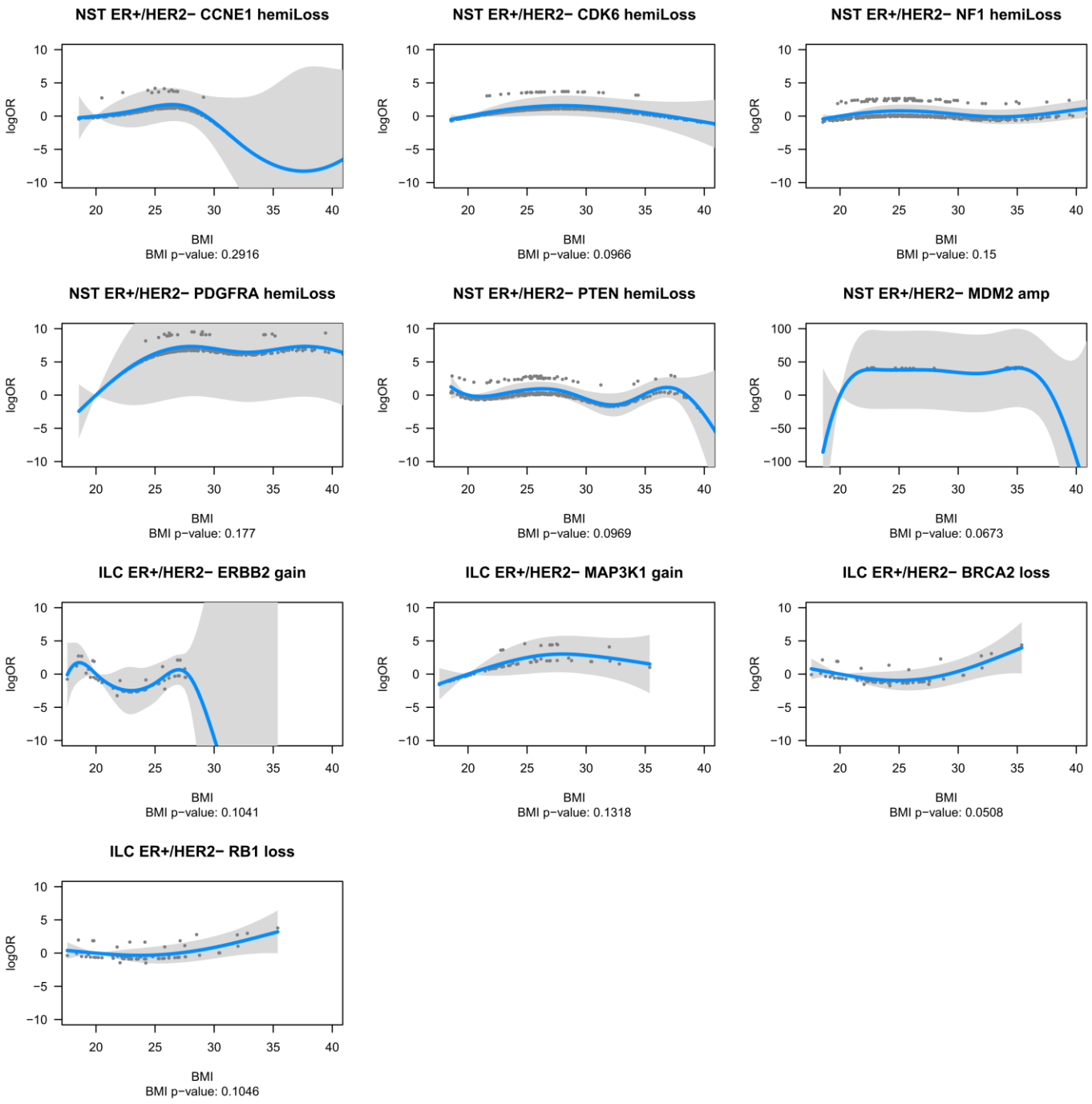
Supplementary Figure 2. Non-linear association of gene mutations with BMI. Plots showing fitted lines of the logOR of gene mutations against continuous BMI (baseline BMI of 20) for those indicated to be better associated with BMI in a non-linear model.



Supplementary Figure 3. 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.

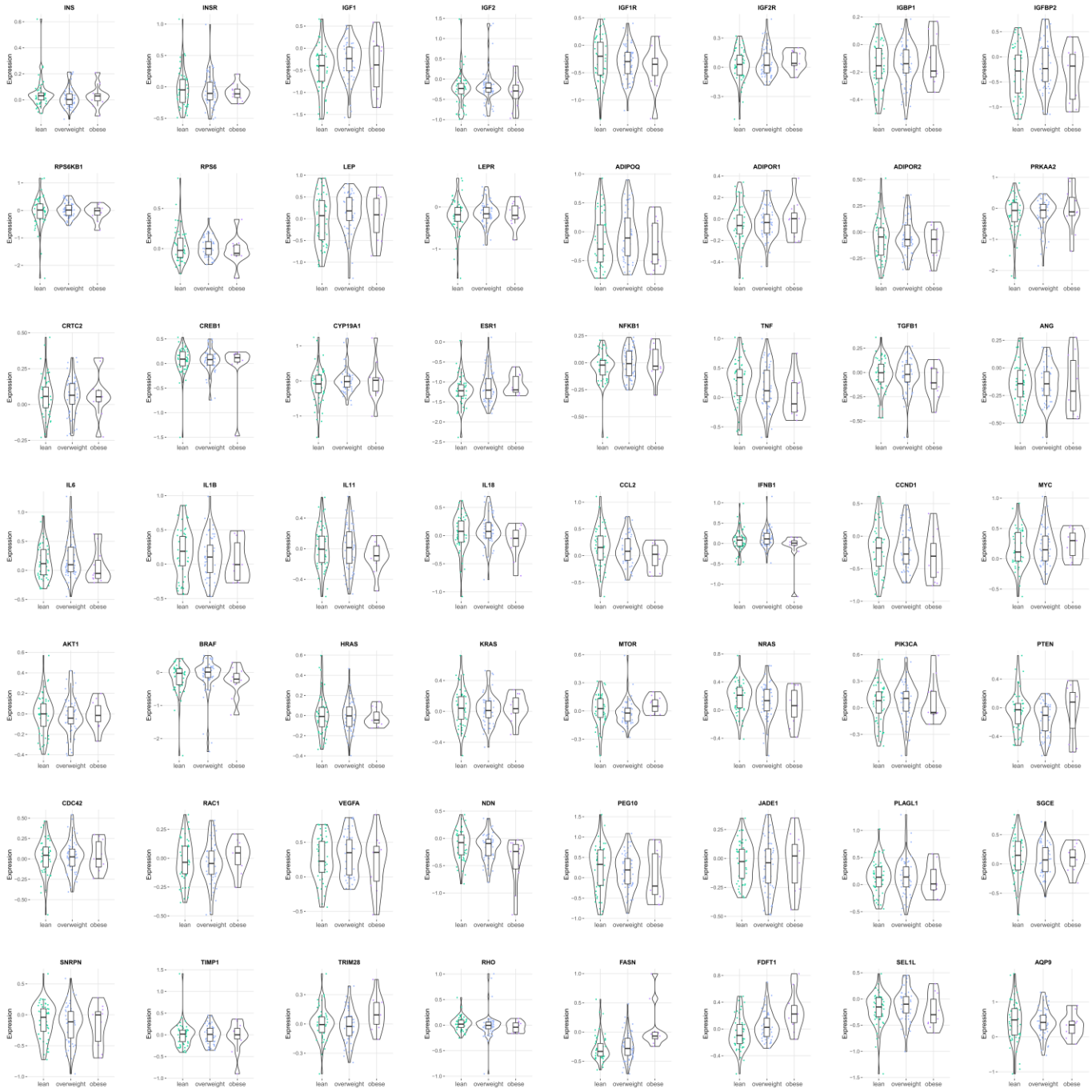


Supplementary Figure 4. Co-occurrence and mutual exclusivity of genomic alterations in NST ER-/HER2- tumors from patients of different BMI categories. Pairwise analyses of recurrent gene mutations and mutations, mutations and CNAs (either amp or hemiLoss) were performed. The color scale represents -log(q-value), 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.

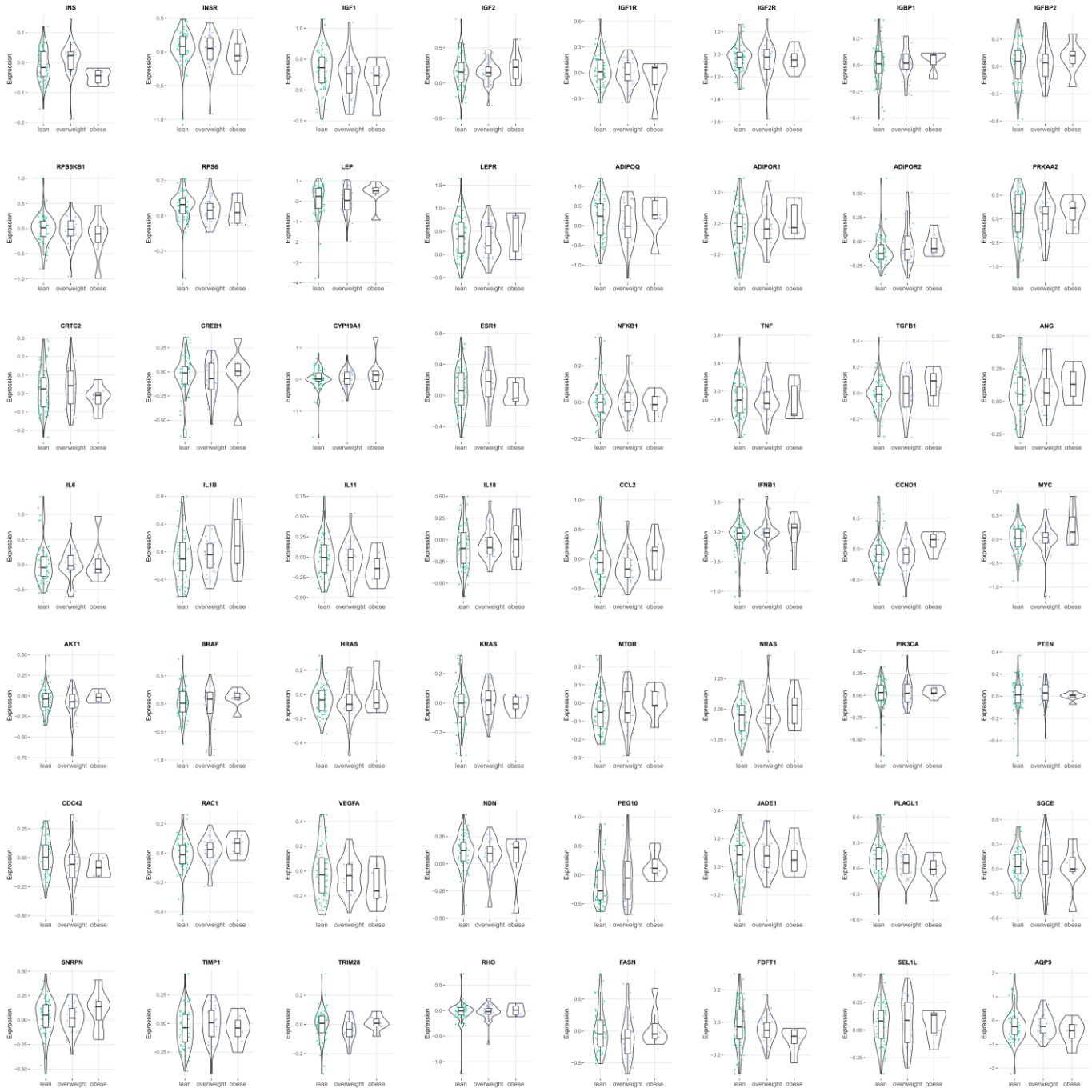


Supplementary Figure 5. Non-linear association of gene-level CNAs with BMI. Plots showing fitted lines of the logOR of gene-level CNAs against continuous BMI (baseline BMI of 20) for those indicated to be better associated with BMI in a non-linear model.

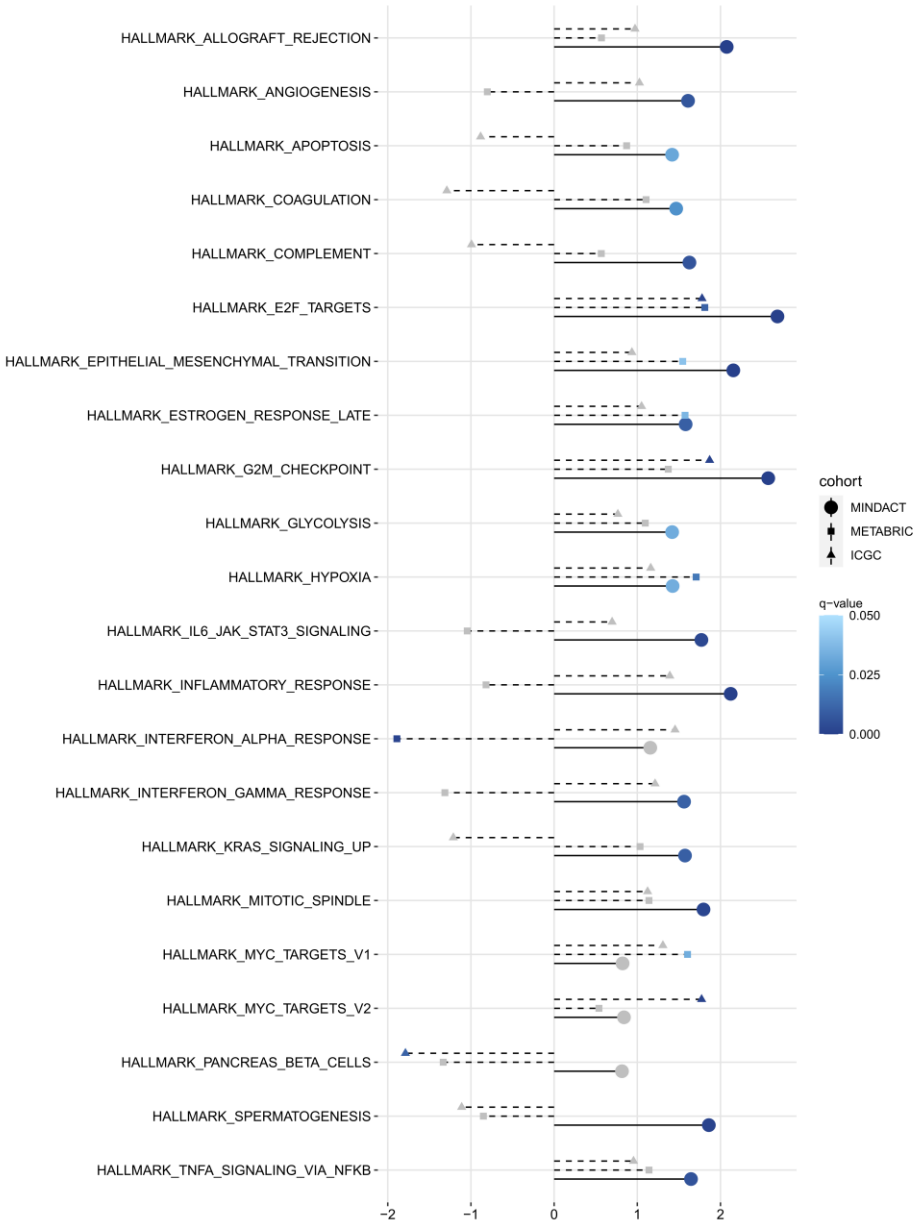
Supplementary Figure 6. Gene expression of relevant genes in the BC-obesity axis in NST ER+/HER2- tumors from patients of different BMI categories in the MINDACT cohort. Violin plots showing the expression levels of selected genes. The expression values (log10-transformed) presented here had been normalized prior to data retrieval.



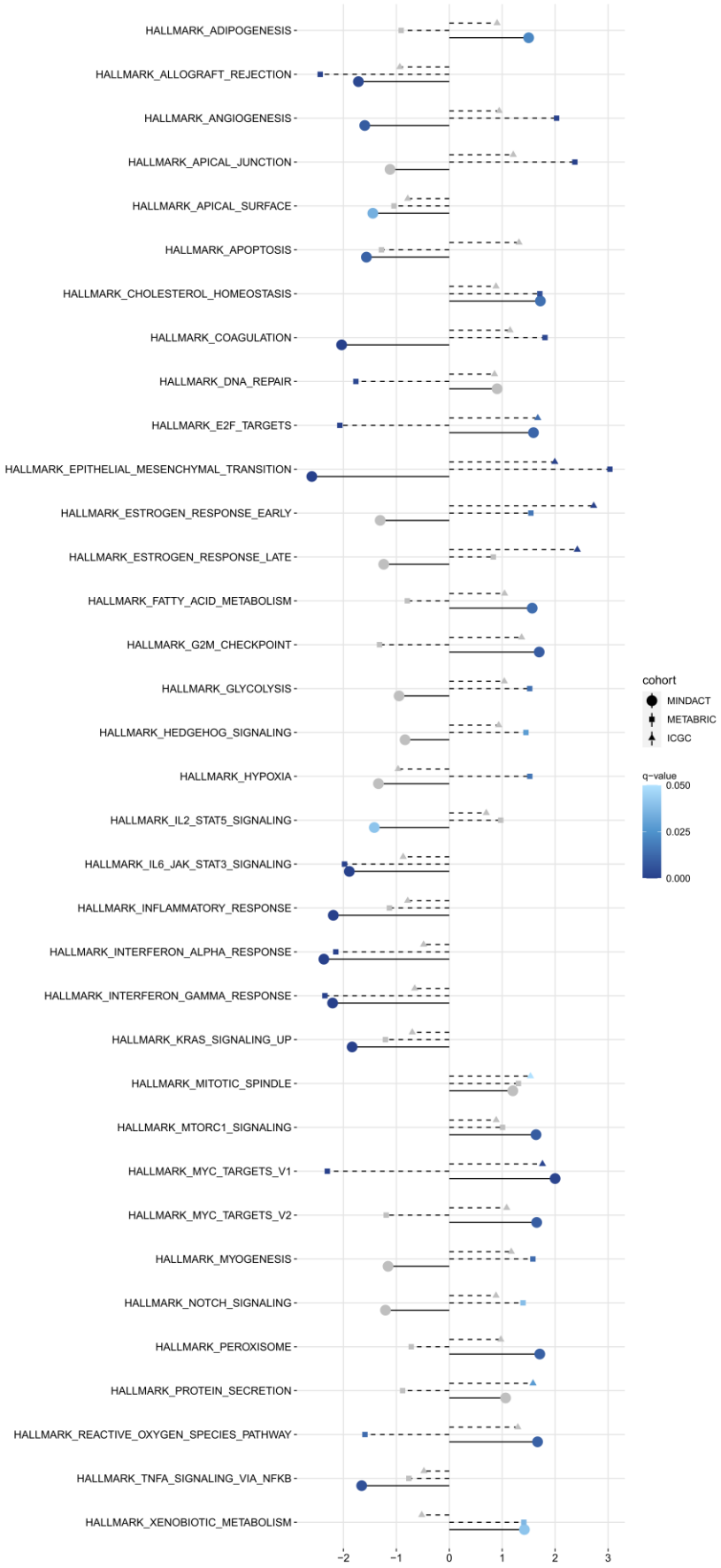
Supplementary Figure 7. Gene expression of relevant genes in the BC-obesity axis in NST ER-/HER2- tumors from patients of different BMI categories in the MINDACT cohort. Violin plots showing the expression levels of selected genes. The expression values (log10-transformed) presented here had been normalized prior to data retrieval.



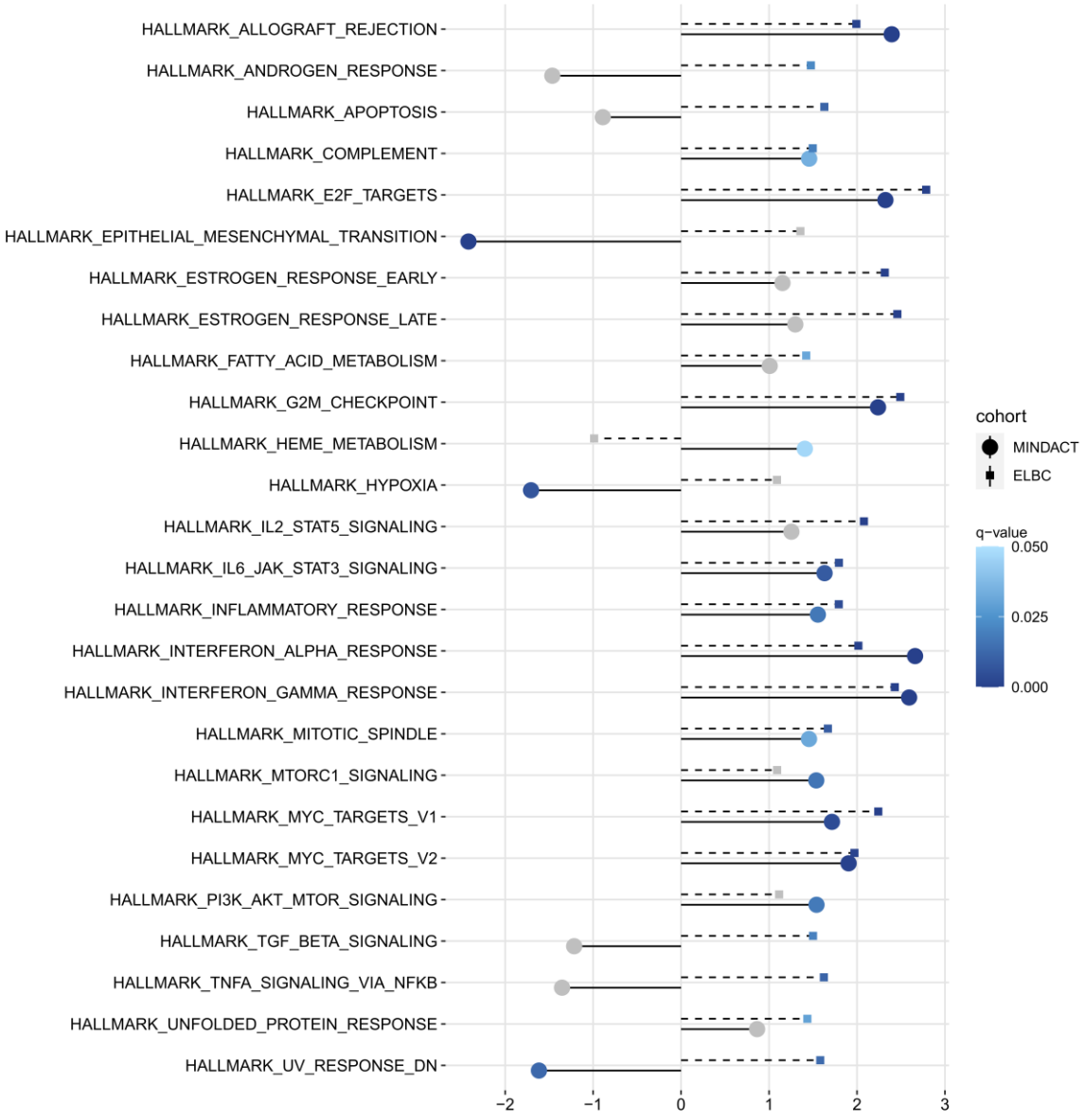
Supplementary Figure 8. Gene expression of relevant genes in the BC-obesity axis in ILC ER+/HER2- tumors from patients of different BMI categories in the MINDACT cohort. Violin plots showing the expression levels of selected genes. The expression values (log10-transformed) presented here had been normalized prior to data retrieval.



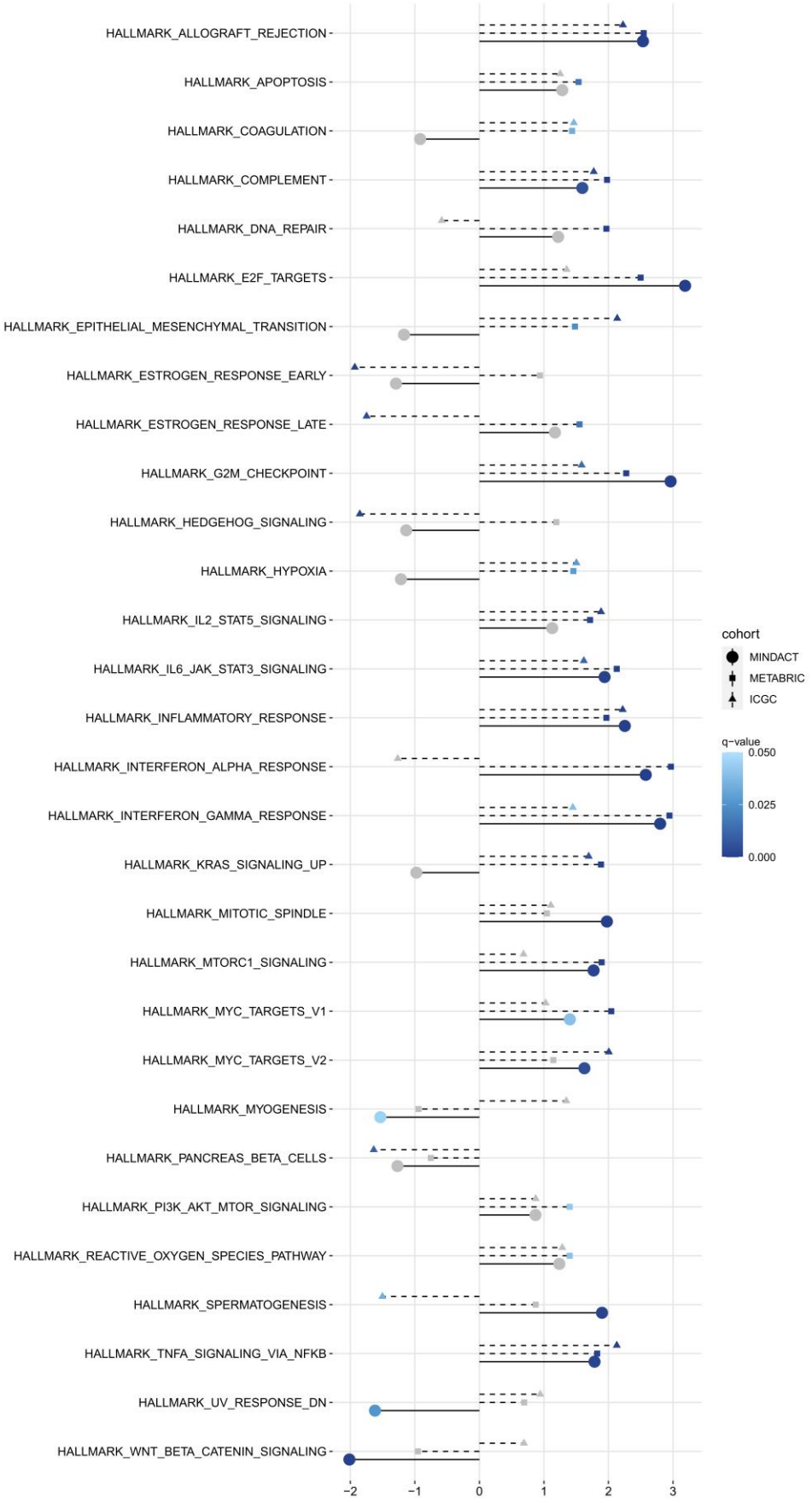
Supplementary Figure 9. Differentially enriched hallmarks in NST ER+/HER2- tumors from obese patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized 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
cohort
MINDACT
METABRIC
ICGC
q−value
0.050
0.025
0.000

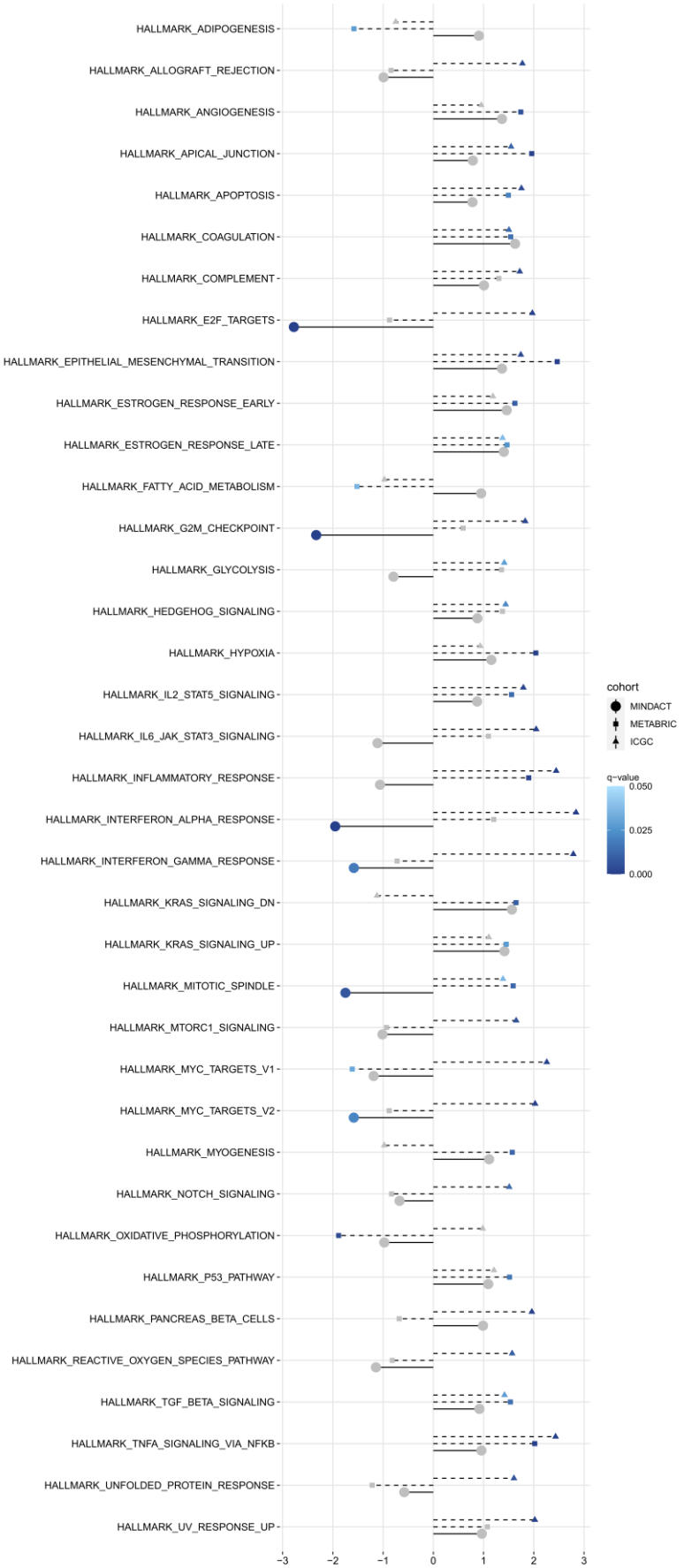
Supplementary Figure 10. Differentially enriched hallmarks in NST ER-/HER2- tumors from obese patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized 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.



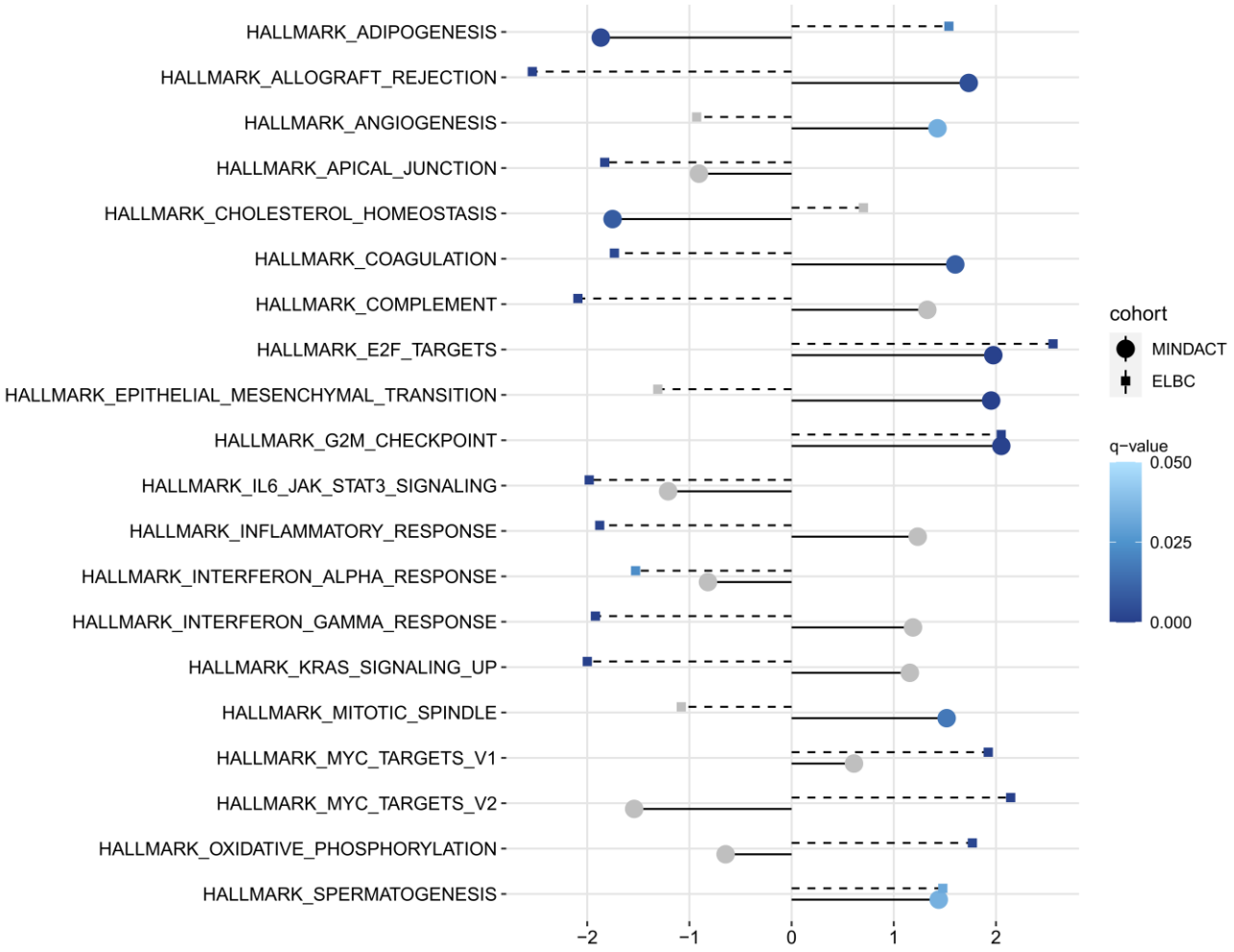
Supplementary Figure 11. Differentially enriched hallmarks in ILC ER+/HER2- tumors from obese patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized 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
cohort
MINDACT
METABRIC
ICGC
q−value
0.050
0.025
0.000

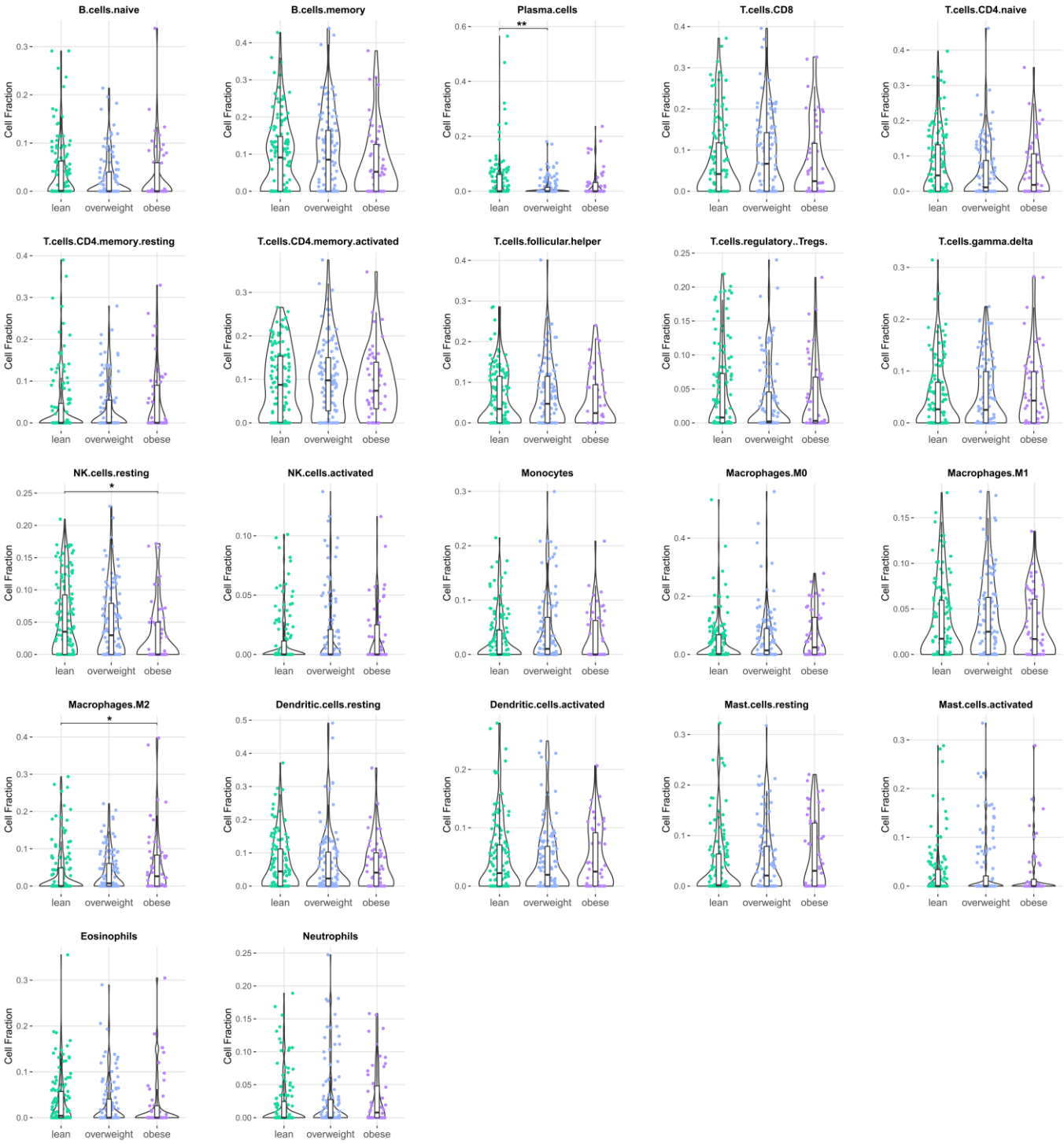
Supplementary Figure 12. Differentially enriched hallmarks in NST ER+/HER2- tumors from overweight patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized 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
cohort
MINDACT
METABRIC
ICGC
q−value
0.050
0.025
0.000

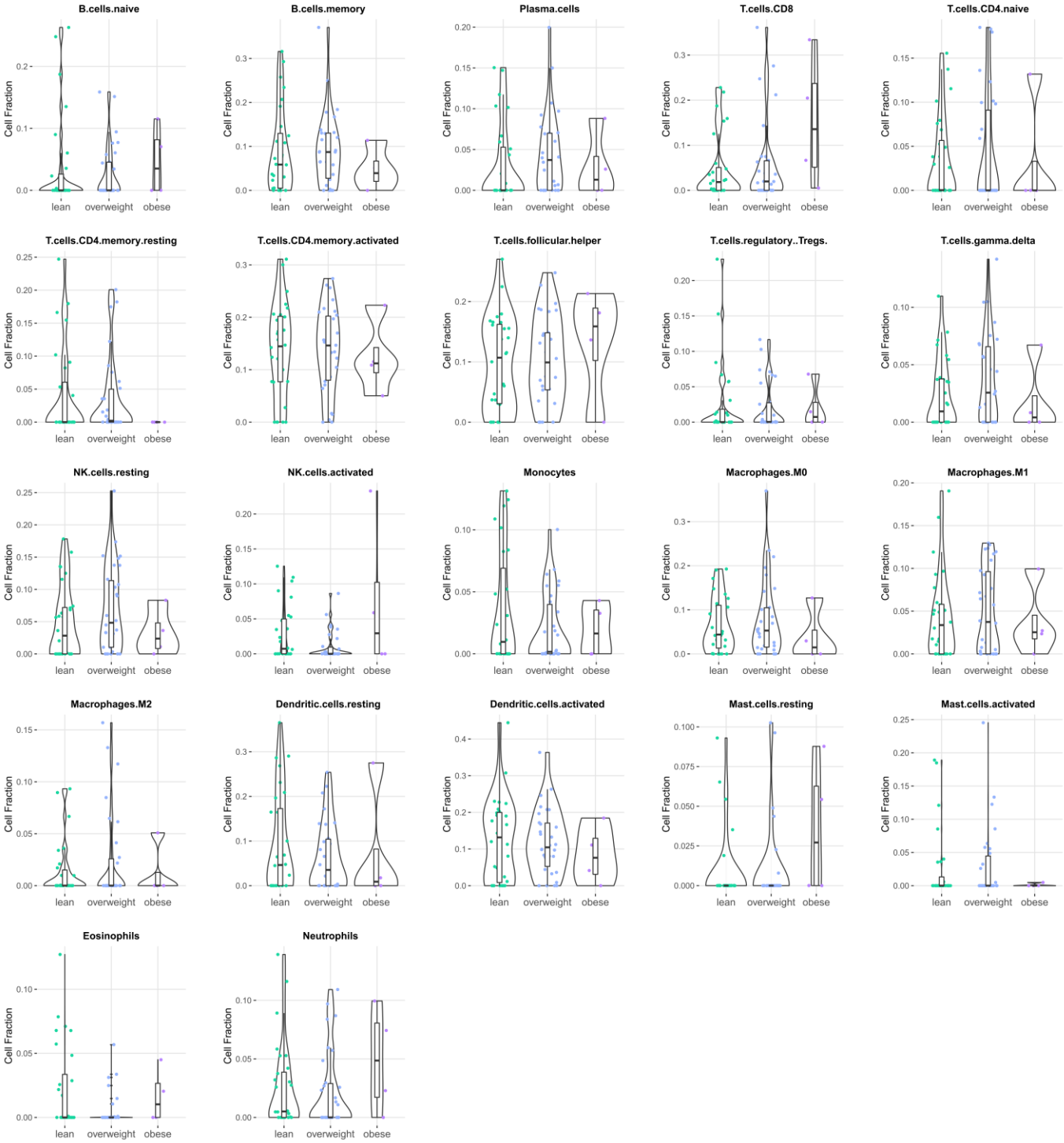
Supplementary Figure 13. Differentially enriched hallmarks in NST ER-/HER2- tumors from overweight patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized 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.



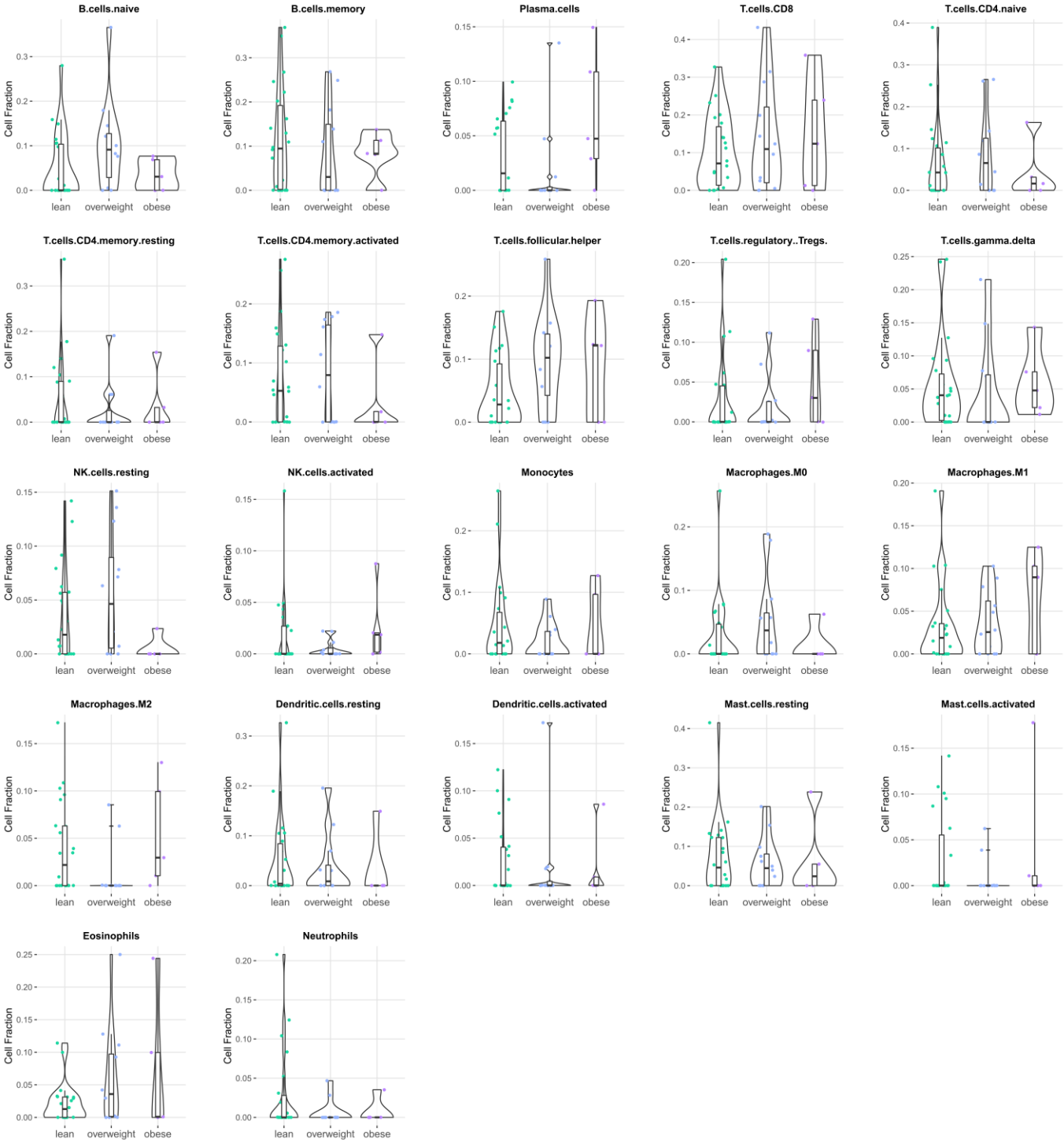
Supplementary Figure 14. Differentially enriched hallmarks in ILC ER+/HER2- tumors from overweight patients compared to lean patients. Lollipop plots displaying differentially enriched molecular hallmarks detected by GSEA (q-value < 0.05) in at least one of the analyzed cohorts. The signs of the normalized 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.



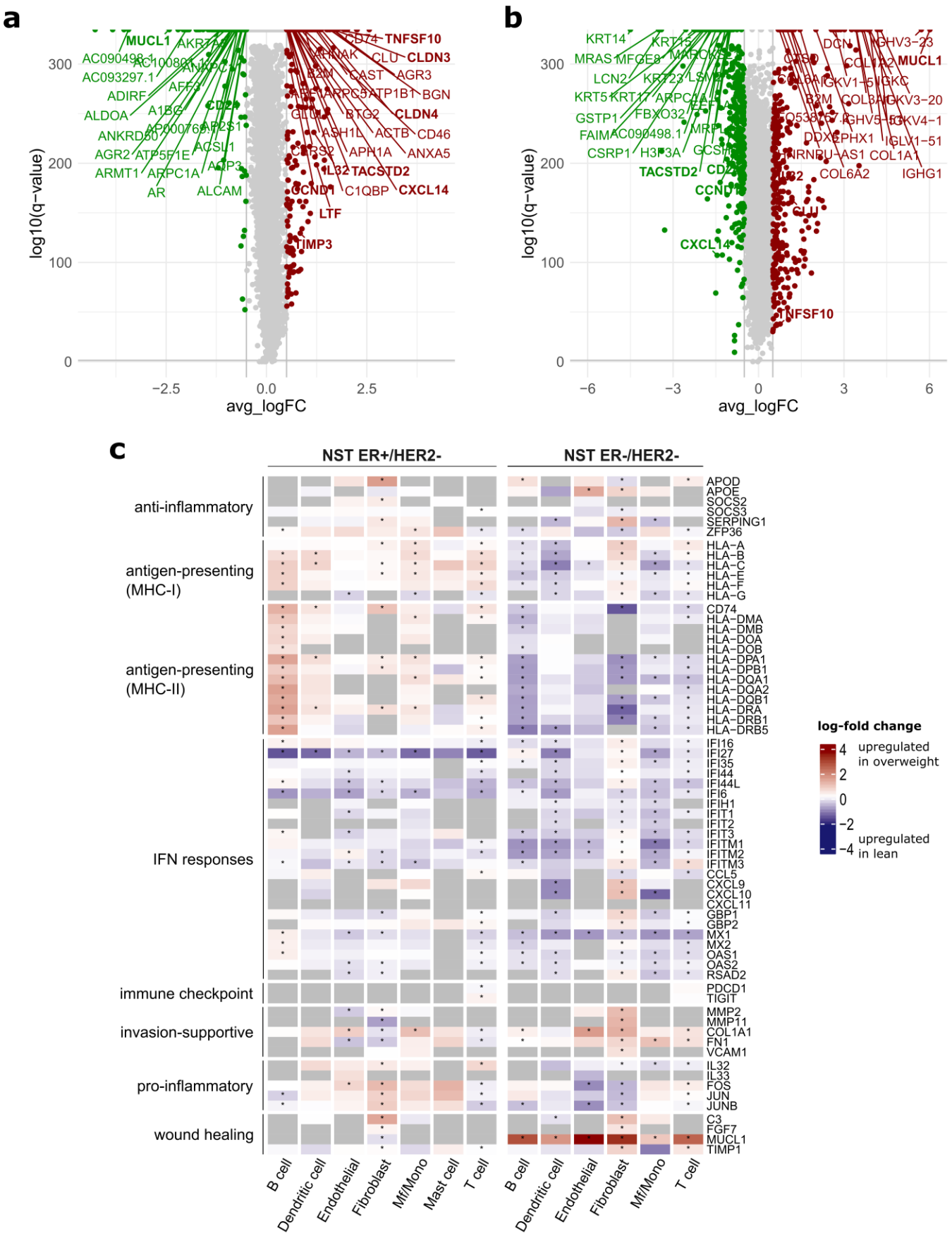
Supplementary Figure 15. Cell fractions of 22 immune cell types in the tumor bulk of NST ER+/HER2- tumors from the MINDACT cohort. Violin plots showing relative frequency of 22 immune cells inferred by deconvolution of the bulk profiling data using CIBERSORTx. p-values were derived from a Wilcoxon's ranked sum test (*, $p < 0.05$; **, $p < 0.01$).



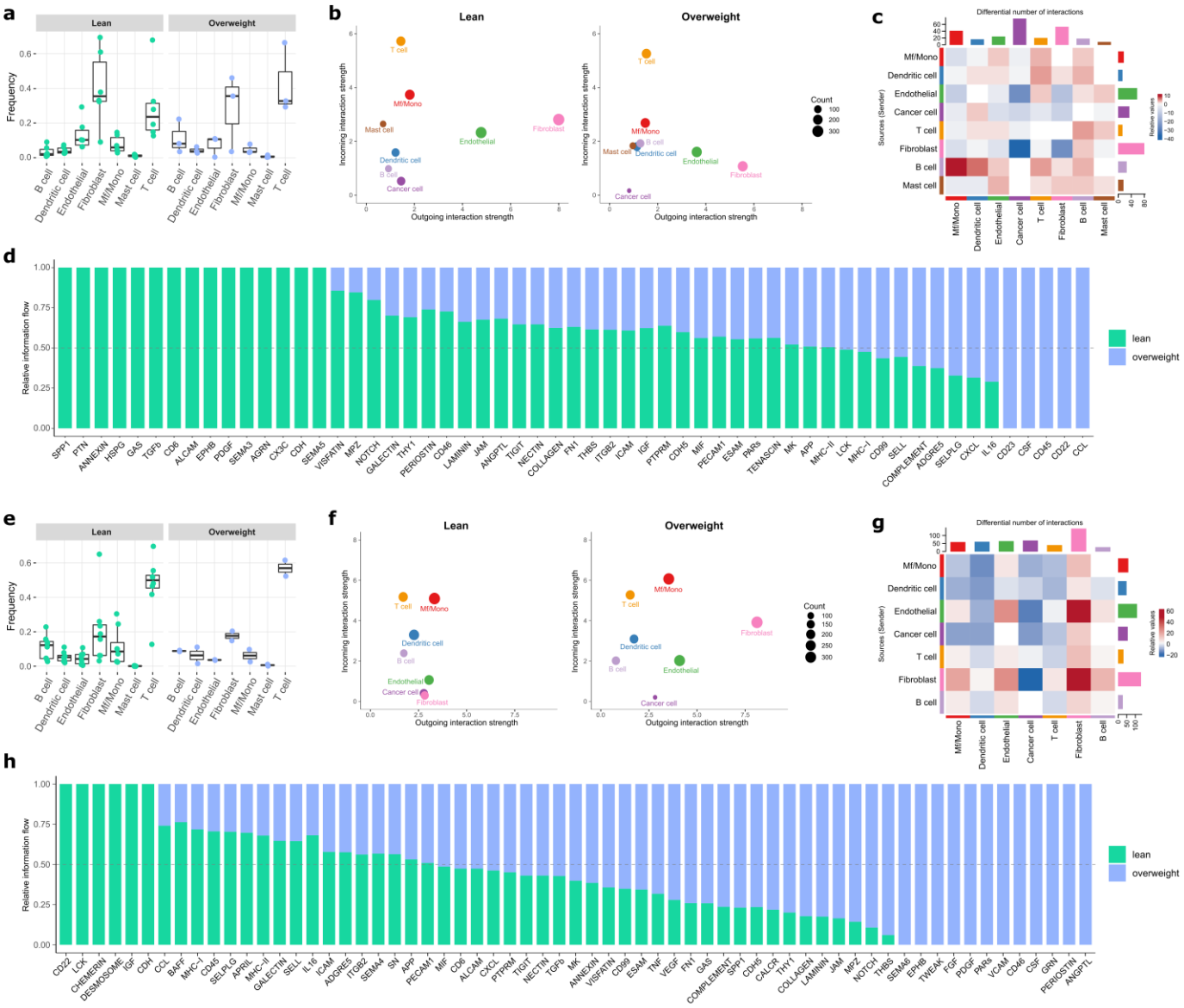
Supplementary Figure 16. Cell fractions of 22 immune cell types in the tumor bulk of NST ER-/HER2- tumors from the MINDACT cohort. Violin plots showing relative frequency of 22 immune cells inferred by deconvolution of the bulk profiling data using CIBERSORTx. p-values were derived from a Wilcoxon's ranked sum test (*, $p < 0.05$; **, $p < 0.01$).



Supplementary Figure 17. 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. p-values were derived from a Wilcoxon's ranked sum test (*, $p < 0.05$; **, $p < 0.01$).


a
b
c
log10(q−value)
avg_logFC
NST ER+/HER2-
NST ER-/HER2-
anti-inflammatory
antigen-presenting (MHC-I)
antigen-presenting (MHC-II)
IFN responses
immune checkpoint
invasion-supportive
pro-inflammatory
wound healing
B cell
Dendritic cell
Endothelial
Fibroblast
Mf/Mono
Mast cell
T cell
log-fold change
upregulated in overweight
upregulated in lean

Supplementary Figure 18. 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-FC 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



Supplementary Figure 19. 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. **(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.

Supplementary Tables - Legends

Supplementary Table 1. Clinical and pathological characteristics of patients with available non-underweight BMI (investigated subset) and all patients in cohorts with incomplete BMI data. Association between clinicopathological variables and BMI availability was assessed using Fisher's exact test for each cohort. p-values were not adjusted for multiple testing.

Supplementary Table 2. Clinical and pathological characteristics of patients with available non-underweight BMI in all cohorts. Association between clinicopathological variables and cohort was assessed using Fisher's exact test for cohorts that were not histology specific (i.e. METABRIC, ICGC, and MINDACT), and two cohorts that were merged in subsequent analyses of genomic alteration data (i.e. METABRIC and ICGC). p-values were not adjusted for multiple testing. The Biokey cohort was excluded from the statistical test due to a small size.

Supplementary Table 3. Clinical and pathological characteristics according to BMI category of patients in all cohorts. Association between clinicopathological variables and categorical BMI was assessed using Fisher's exact test. p-values were not adjusted for multiple testing. The Biokey cohort was excluded from the statistical test due to a small size.

Supplementary Table 4. List of driver genes used to investigate the association of genomic alterations with BMI. Genes harboring base substitutions and indels, or CNAs reported as BC driver events by Nik-Zainal *et al*.

Supplementary Table 5. Oncogenic base substitutions and indels in tumors from METABRIC and ICGC.

Supplementary Table 6. Gene-level copy number alterations in tumors from METABRIC and ICGC.

Supplementary Table 7. Association of recurrent gene mutations with continuous BMI. Merged data from two cohorts, METABRIC and ICGC, was used for analyses of the NST ER+/HER2- and NST ER-/HER2- subtypes, and from the ELBC cohort for the ILC ER+/HER2- subtype. Gene mutations with at least 5 events detected in the respective sub-cohorts were evaluated. Linear association was assessed using Firth's logistic regression model. Potential non-linear association was explored using generalized additive models (GAM) whereby models with (SPLINE) and without (noSPLINE) a spline term were compared using two metrics, AIC and a LR test.

Supplementary Table 8. Association of prevalence of hotspot substitutions and indels with categorical BMI. Merged data from two cohorts, METABRIC and ICGC, was used for analyses of the NST ER+/HER2- and NST ER-/HER2- subtypes. Substitutions and indels with at least one event detected in tumors from lean patients, and at least one event detected in tumors from either overweight or obese patients, in the respective sub-cohorts, were tested.

Supplementary Table 9. Association of recurrent gene-level CNAs with continuous BMI. Merged data from two cohorts, METABRIC and ICGC, was used for analyses of the NST ER+/HER2- and NST ER-/HER2- subtypes, and from the ELBC cohort for the ILC ER+/HER2- subtype. Gene-level CNAs with at least 10 events detected in the respective cohorts were evaluated. Linear association was assessed using Firth's logistic regression model. Potential non-linear association was explored using generalized additive models (GAM) whereby models with (SPLINE) and without (noSPLINE) a spline term were compared using two metrics, AIC and a LR test.

Supplementary Table 10. Association of recurrent gene-level CNAs with categorical BMI. Merged data from two cohorts, METABRIC and ICGC, was used for analyses of the NST ER+/HER2- and NST ER-/HER2- subtypes, and from the

ELBC cohort for the ILC ER+/HER2- subtype. Gene-level CNAs with at least 10 events detected in the respective cohorts were evaluated.

Supplementary Table 11. Association of genome instability and mutational and rearrangement signatures with continuous and categorical BMI. Corresponding data from the ICGC cohort were used for analyses of the NST ER+/HER2- and NST ER-/HER2- subtypes.

Supplementary Table 12. Cell type-specific differentially expressed genes in NST ER+/HER2- and NST ER-/HER2- tumors from obese patients versus lean patients. Data from the BioKey cohort was used for analyses. Mast cells were not evaluated for the NST ER-/HER2- subtype due to low absolute cell counts. Genes with absolute logFC > 0.5 and q-value < 0.05 are shown. A logFC greater than 0 indicates an overexpression of the respective gene in the respective cell type in tumors from obese patients, and vice versa.

Supplementary Table 13. Cell type-specific differentially expressed genes in NST ER+/HER2- and NST ER-/HER2- tumors from overweight patients versus lean patients. Data from the BioKey cohort was used for analyses. Mast cells were not evaluated for the NST ER-/HER2- subtype due to low absolute cell counts. Genes with absolute logFC > 0.5 and q-value < 0.05 are shown. A logFC greater than 0 indicates an overexpression of the respective gene in the respective cell type in tumors from overweight patients, and vice versa.

Supplementary Table 14. Cell type-specific differentially enriched GOBP gene sets in NST ER+/HER2- and NST ER-/HER2- tumors from obese patients versus lean patients. Data from the BioKey cohort was used for analyses. Mast cells were not evaluated for the NST ER-/HER2- subtype due to low absolute cell counts. Gene sets with q-value < 0.1 are shown. A NES greater than 0 indicates an enrichment of the respective gene set in the respective cell type in tumors from obese patients, and vice versa.

Supplementary Table 15. Cell type-specific differentially enriched REACTOME gene sets in NST ER+/HER2- and NST ER-/HER2- tumors from obese patients versus lean patients. Data from the BioKey cohort was used for analyses. Mast cells were not evaluated for the NST ER-/HER2- subtype due to low absolute cell counts. Gene sets with q-value < 0.1 are shown. A NES greater than 0 indicates an enrichment of the respective gene set in the respective cell type in tumors from obese patients, and vice versa.

Supplementary Table 16. Cell type-specific differentially enriched GOBP gene sets in NST ER+/HER2- and NST ER-/HER2- tumors from overweight patients versus lean patients. Data from the BioKey cohort was used for analyses. Mast cells were not evaluated for the NST ER-/HER2- subtype due to low absolute cell counts. Gene sets with q-value < 0.1 are shown. A NES greater than 0 indicates an enrichment of the respective gene set in the respective cell type in tumors from overweight patients, and vice versa.

Supplementary Table 17. Cell type-specific differentially enriched REACTOME gene sets in NST ER+/HER2- and NST ER-/HER2- tumors from overweight patients versus lean patients. Data from the BioKey cohort was used for analyses. Mast cells were not evaluated for the NST ER-/HER2- subtype due to low absolute cell counts. Gene sets with q-value < 0.1 are shown. A NES greater than 0 indicates an enrichment of the respective gene set in the respective cell type in tumors from overweight patients, and vice versa.

Supplementary Table 18. Cell counts in investigated cellular compartments in NST ER+/HER2- and NST ER-/HER2- tumors according to BMI category in the Biokey cohort