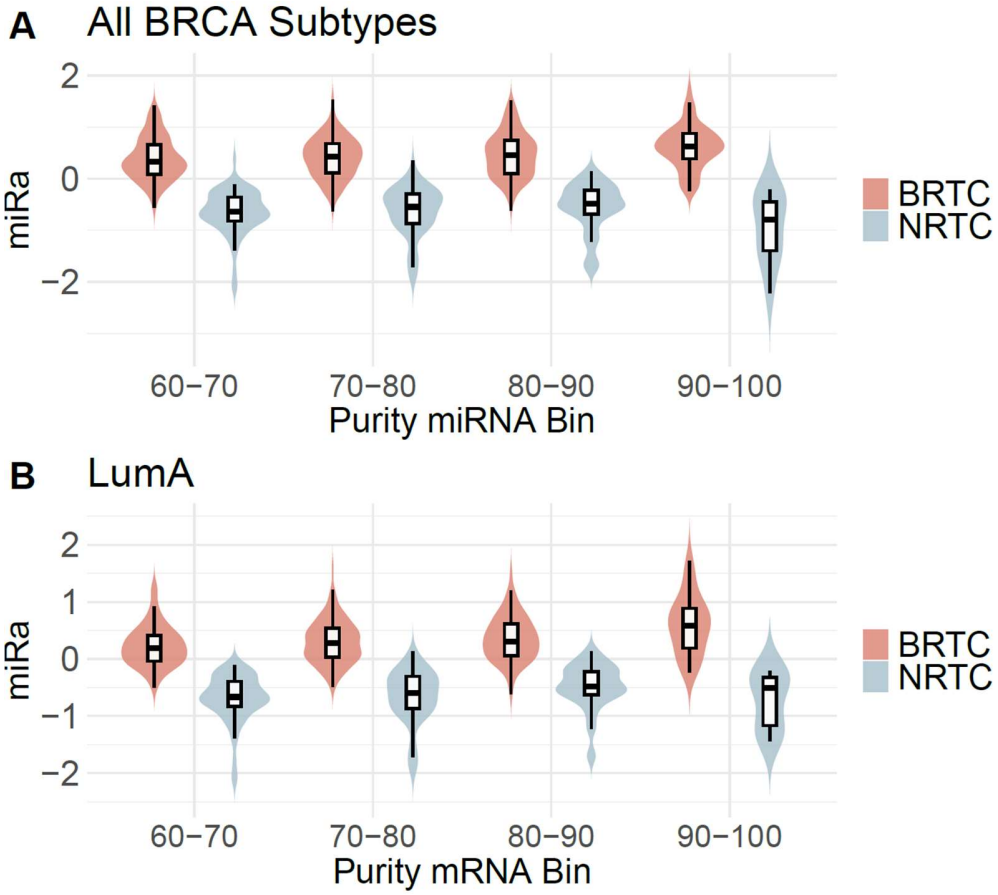
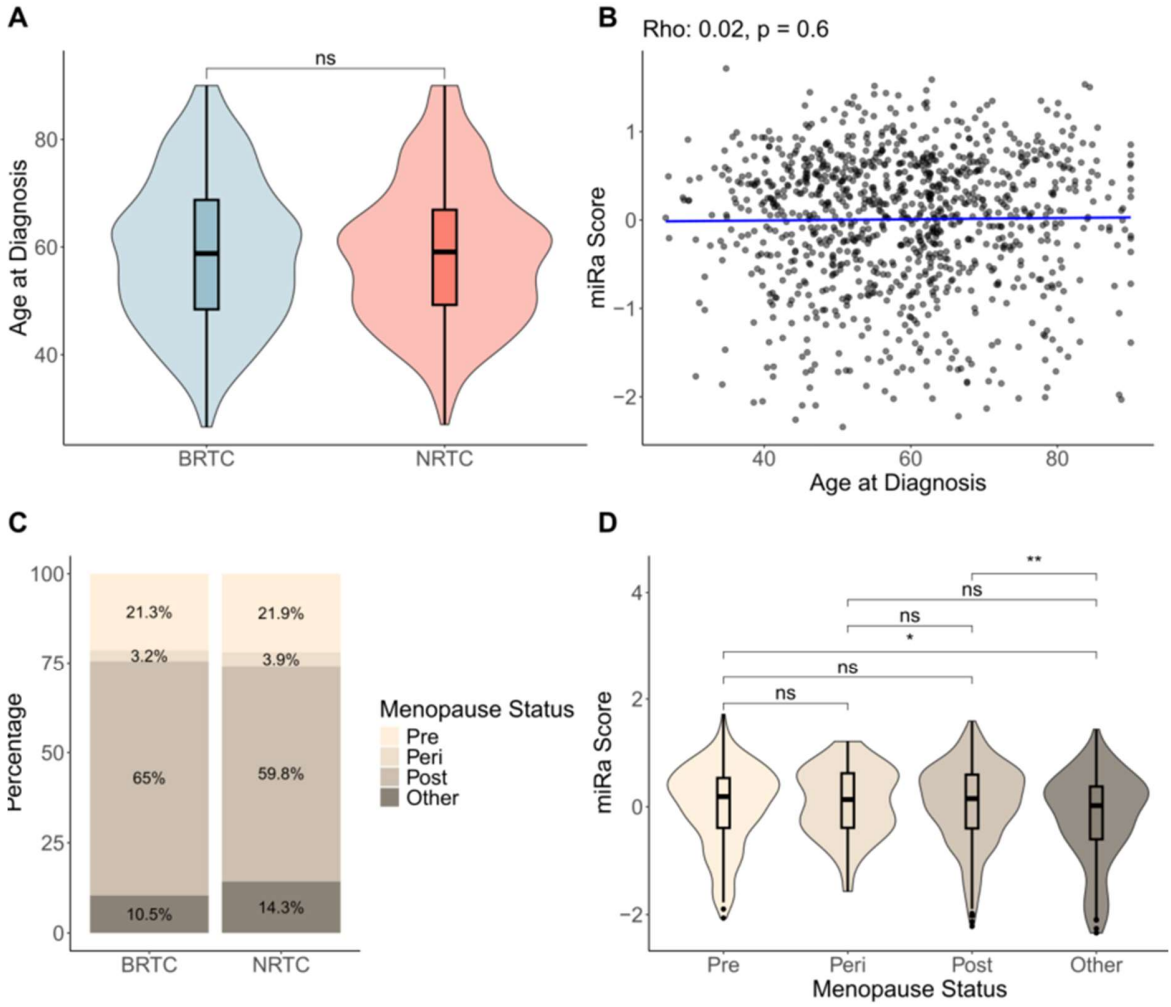
Supplementary Figures

Supplementary Fig. 1.



miRNA arm usage patterns are independent of tumor purity. miRa score distributions across tumor purity bins (60–70%, 70–80%, 80–90%, and 90–100%), stratified by miRNA arm usage cluster (NRTC and BRTC). Analyses are shown for all TCGA-BRCA tumor subtypes (A) and for LumA tumors only (B). Tumor purity estimates were obtained from TCGA annotations. Normal-like subtype samples were excluded the analyses.

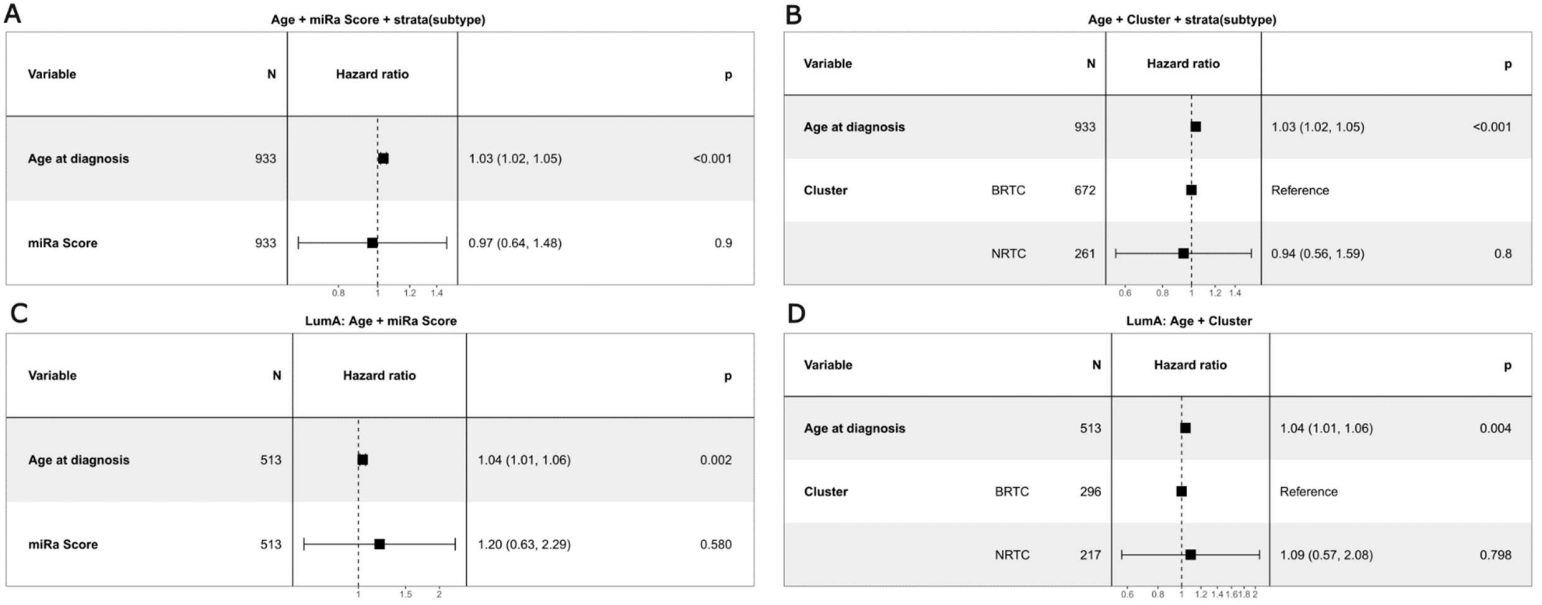
Supplementary Fig. 2.



miRNA arm usage patterns are not associated with patient age or menopausal status. (A) Distribution of patient age at diagnosis for LumA tumors assigned to the NRTC and BRTC clusters. Statistical significance was assessed using a Wilcoxon test. (B) Spearman correlation between miRa score and patient age at diagnosis across LumA tumor samples. Each dot represents an individual sample. The blue line represents the linear regression fit. (C-D) Menopausal status in relation to miRNA arm usage showing the distribution of menopausal status categories (pre-, peri-, post-menopausal and other/unknown) in LumA tumors assigned to the NRTC and BRTC

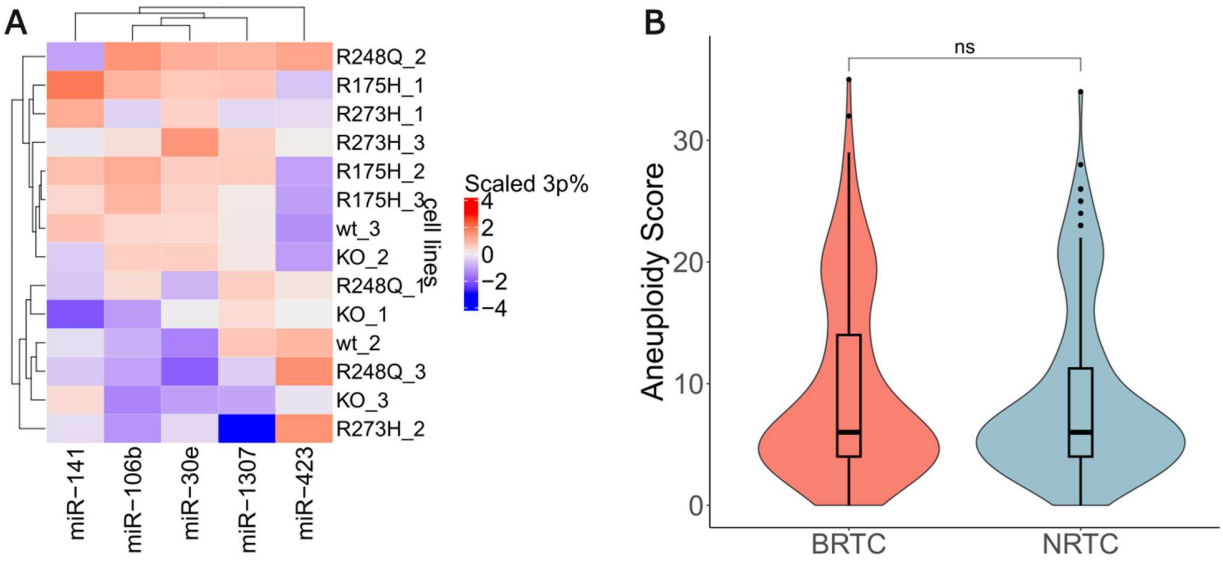
clusters (C), and the corresponding miRa score distributions stratifies by menopausal status (D). Statistical significance was assessed using a Wilcoxon test. Asterisks indicate p-value significance levels: *p < 0.05, **p < 0.01; ns = not significant.

Supplementary Fig. 3.



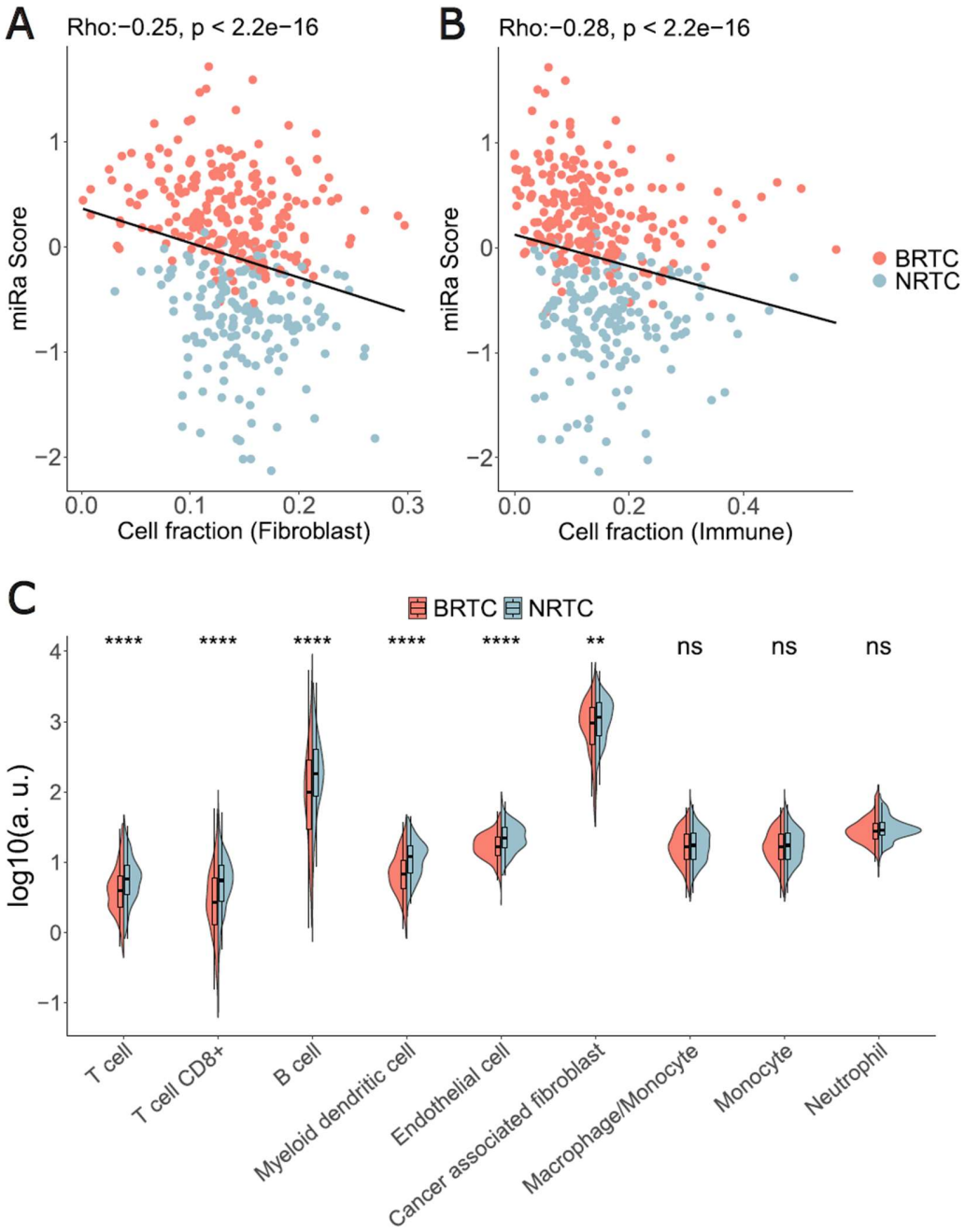
Association of miRa score and arm-usage clusters with overall survival in TCGA-BRCA patients. Forest plots showing hazard rations from Cox models evaluating the association of the miRa score or NRTC and BRTC clusters with overall survival in the TCGA-BRCA cohort. Age at diagnostics was included as covariate in all models. (A-B) Multivariable Cox moles stratified by PAM50 subtype assessing the prognostic value of (A) the miRa score and (B) clusters. (C-D) Cox models restricted to Lumial A breast cancer subtype evaluating the association of the (C) miRa score and (D) clutters with overall survival. Hazard ratios are sown with 95% CI and corresponding p-values.

Supplementary Fig. 4:



miRNA arm usage patterns are not associated with chromosomal aneuploidy. (A) Heatmap showing miRNA arm usage patterns across MCF10A breast epithelial cell lines isogenically expressing different TP53 states (wild-type, knockout and hotspot mutations). Rows represent individual samples and columns represent miRNAs, with values corresponding to the relative 3p arm usage. (B) Aneuploidy scores of LumA tumors assigned o NRTC and NRTC clusters. Aneuploidy scores were derived from TCGA copy number alteration data and reflect the degree of chromosomal instability in each tumor sample. Statistical significance was assessed using a Wilcoxon test. ns = not significant.

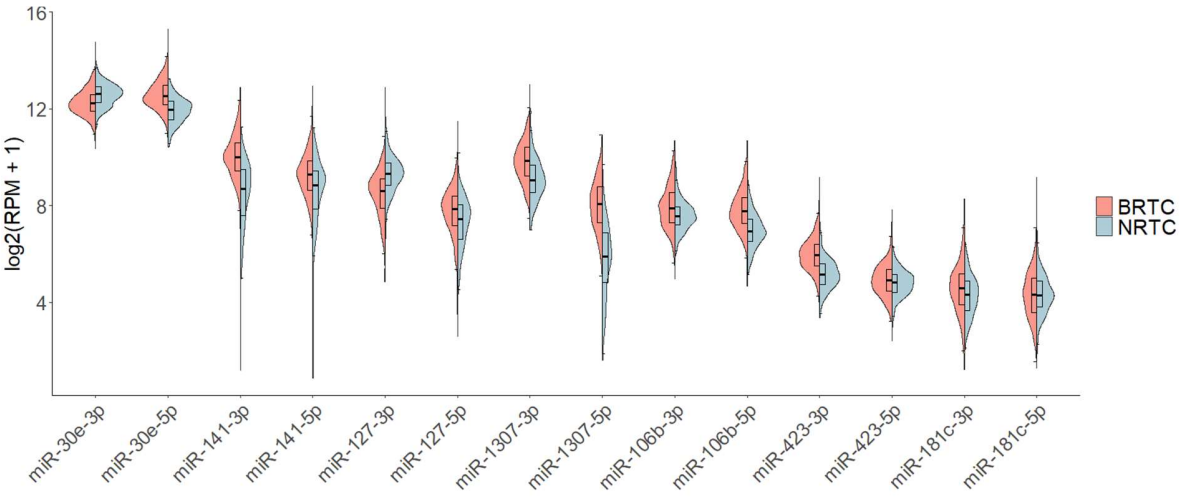
Supplementary Fig. 5



Association of miRNA arm usage with tumor microenvironment composition in Luminal A breast cancer. (A-B) Spearman correlation between the miRa score and the estimated fibroblast (A) and immune cell fractions (B) in LumA tumors inferred using gene-expression- and DNA-

methylation-based cell type deconvolution methods (1, 2). Each dot represents one tumor sample and is colored according to miRNA arm-usage cluster (NRTC or BRTC). The black line represents the linear regression fit. (C) Cell-type fractions were estimated based on DNA methylation profiles using the Houseman method (2) for LumA tumors assigned to NRTC or BRTC clusters. Cell-type fractions are shown on a log10-transformed scale (a.u.). Statistical significance was assessed using a Wilcoxon test. Asterisks indicate p-value significance levels: **p < 0.01, ****p < 0.0001; ns = not significant.

Supplementary Fig 6.



miRNA arm expression patterns across NRTC and BRTC clusters in Luminal A tumors. Distribution of normalized expression levels (log2(RPM + 1)) of the mature 3p and 5p arms of the miRNAs contributing to the miRa score across Luminal A breast cancer patients from the TCGA-BRCA cohort assigned to the NRTC and BRTC clusters.

References

1. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17(1):218.
2. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.