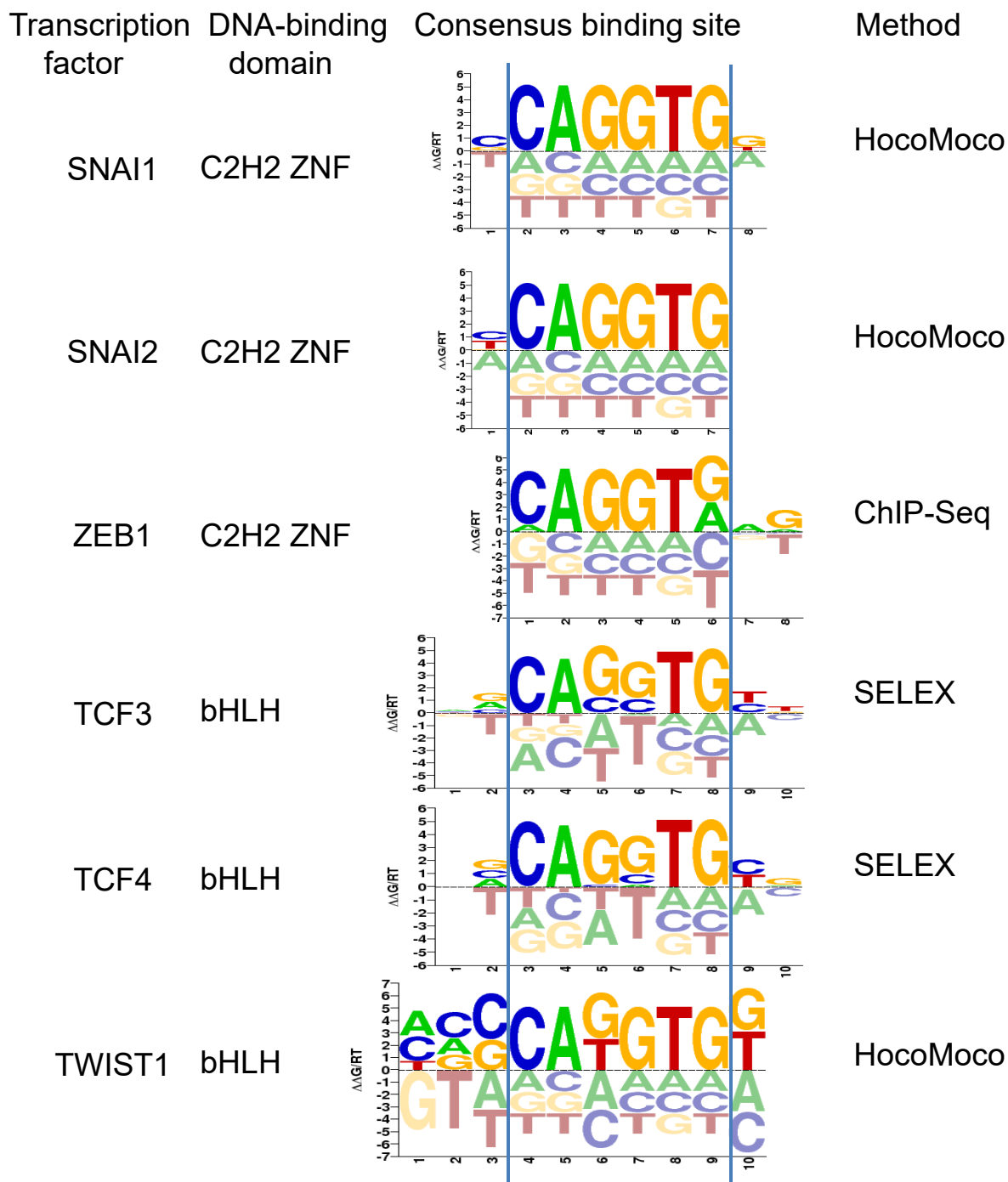


Functional hierarchy and cooperation of EMT master transcription factors in breast cancer metastasis

Joseph B. Addison<sup>1,\$</sup>, Maria A. Voronkova<sup>1,\$</sup>, James H. Fugett<sup>1,\$</sup>,  
Chen-Chung Lin<sup>1</sup>, Nathaniel C. Linville<sup>1</sup>, Brandon Trinh<sup>1</sup>, Ryan H.  
Livengood<sup>2</sup>, Matthew B. Smolkin<sup>2</sup>, Michael D. Schaller<sup>1</sup>, J. Michael  
Ruppert<sup>1</sup>, Elena N. Pugacheva<sup>1</sup>, Chad J. Creighton<sup>3</sup>, and Alexey V.  
Ivanov<sup>1,\*</sup>

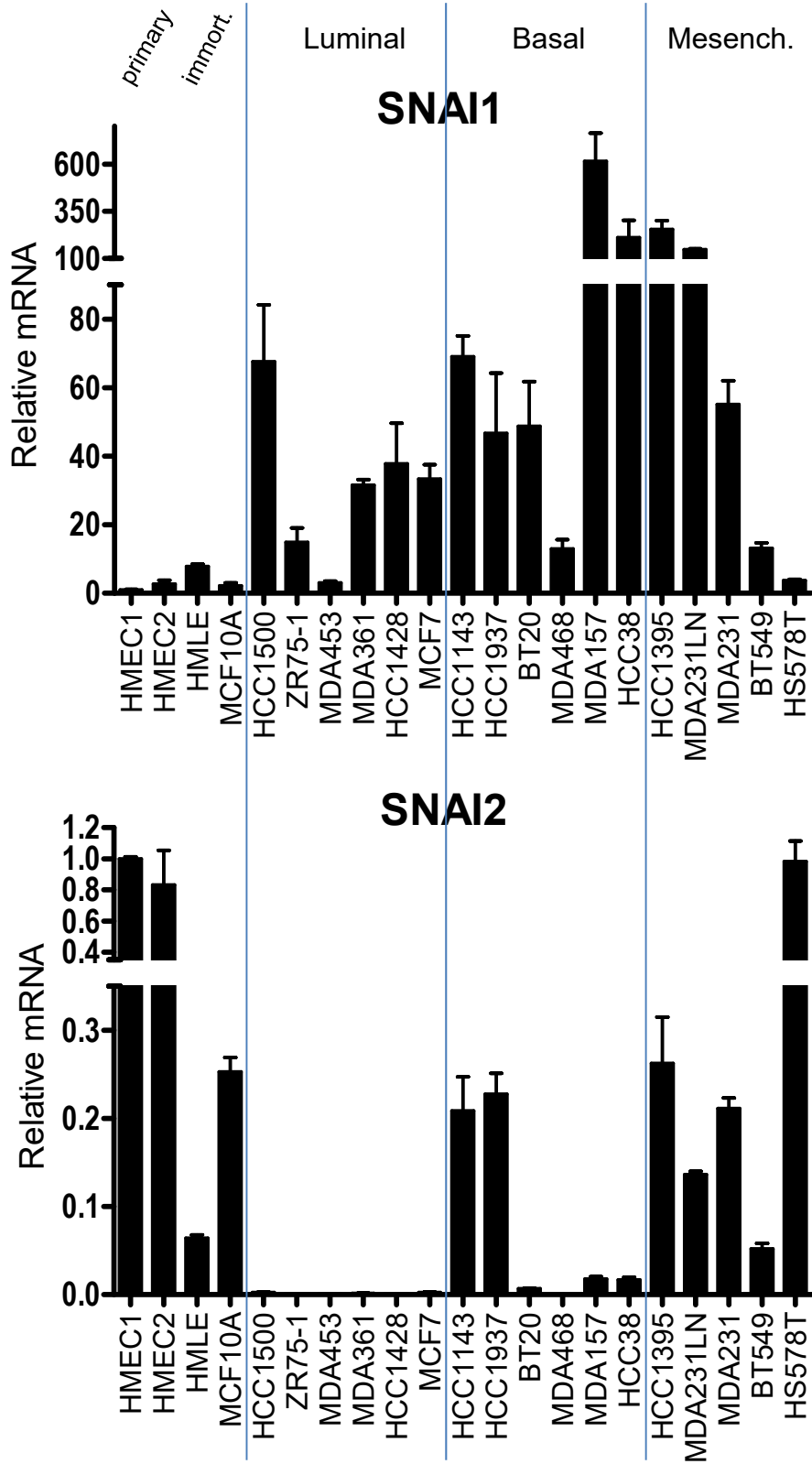
# Supplementary Figure 1

## DNA consensus binding sites for EMT-TFs



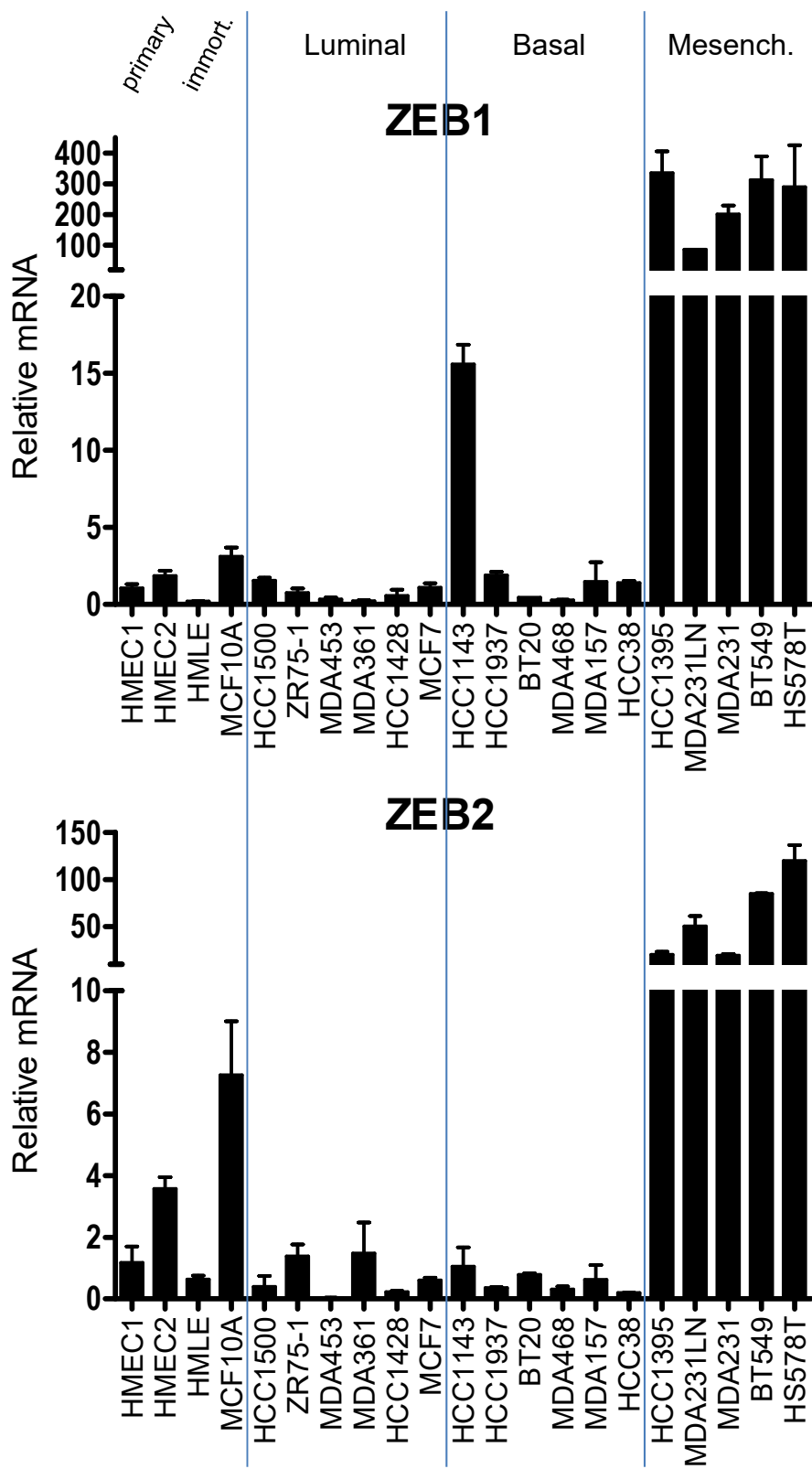
**Supplementary Figure 1. The EMT-TFs bind similar E-box DNA consensus sequences.** DNA consensus binding sites for human EMT-TFs presented as Logo diagrams. Source: CIS-BP, <http://cisbp.cabr.utoronto.ca/>

Supplementary Figure 2A

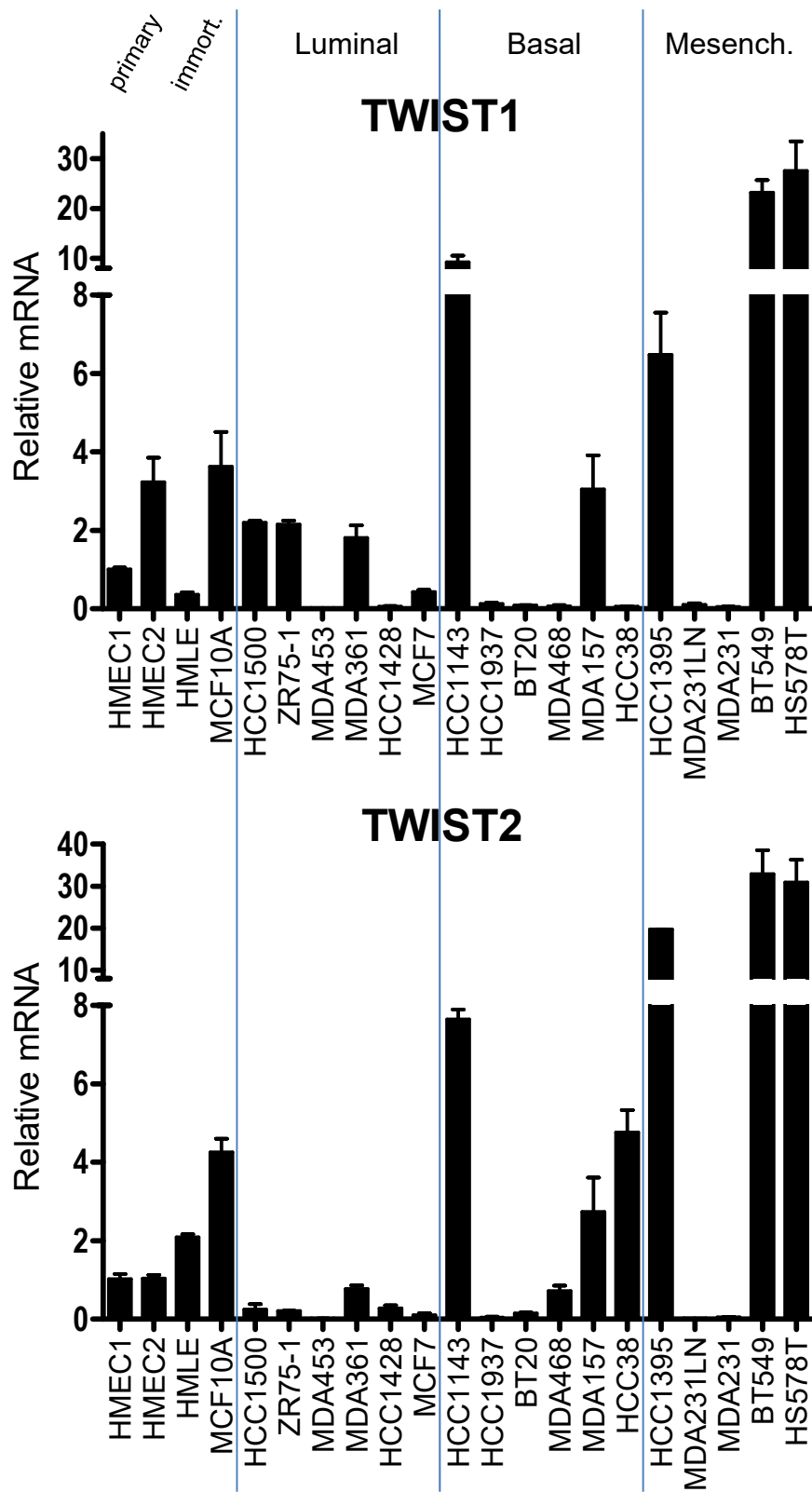


**Supplementary Figure 2. Expression of ZEB1 and ZEB2 strongly correlates with the mesenchymal phenotype in breast cancer cell lines.** Relative mRNA levels of the EMT-TFs in breast cancer cell lines shown at the X-axis determined by RT-qPCR. The mean values of the EMT-TFs expression in HMEC1 cells were set to 1. The data are mean  $\pm$  SEM of three biological replicates. HMEC1 and HMEC2 are primary human mammary epithelial cells; immort. – immortalized cell lines; Luminal, Basal and Mesenchymal cell line groups are indicated on the top.

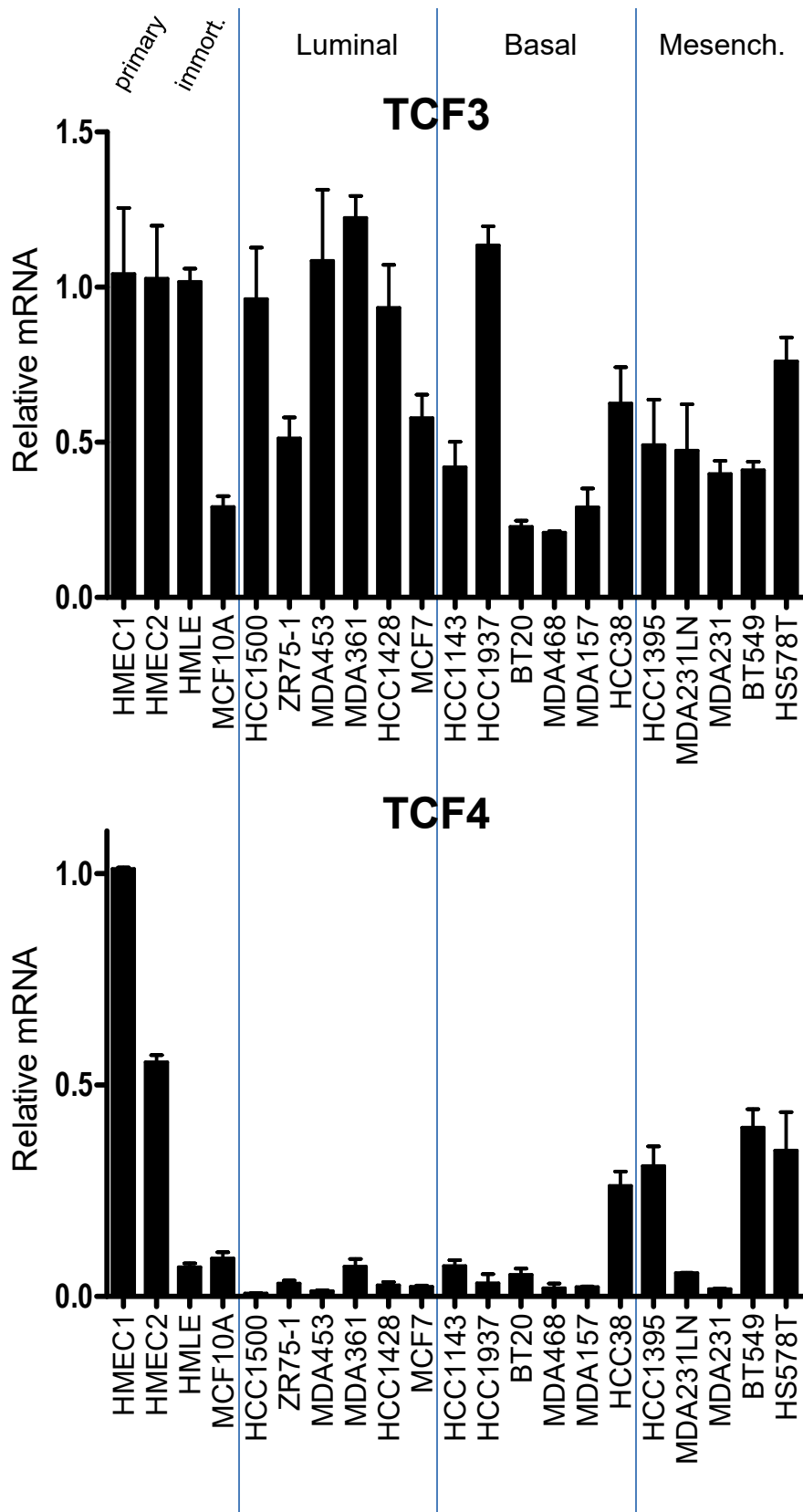
Supplementary Figure 2B



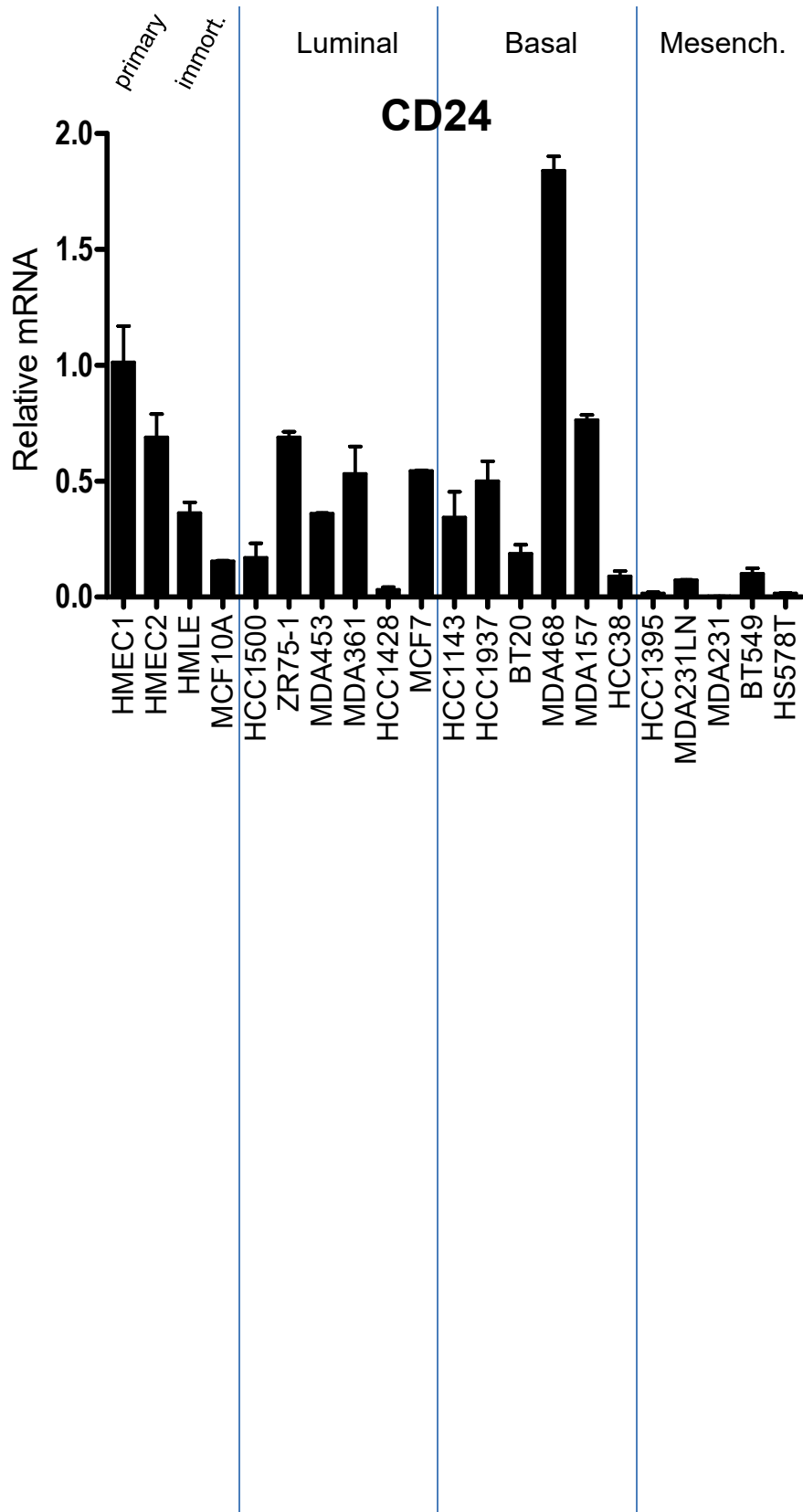
Supplementary Figure 2C



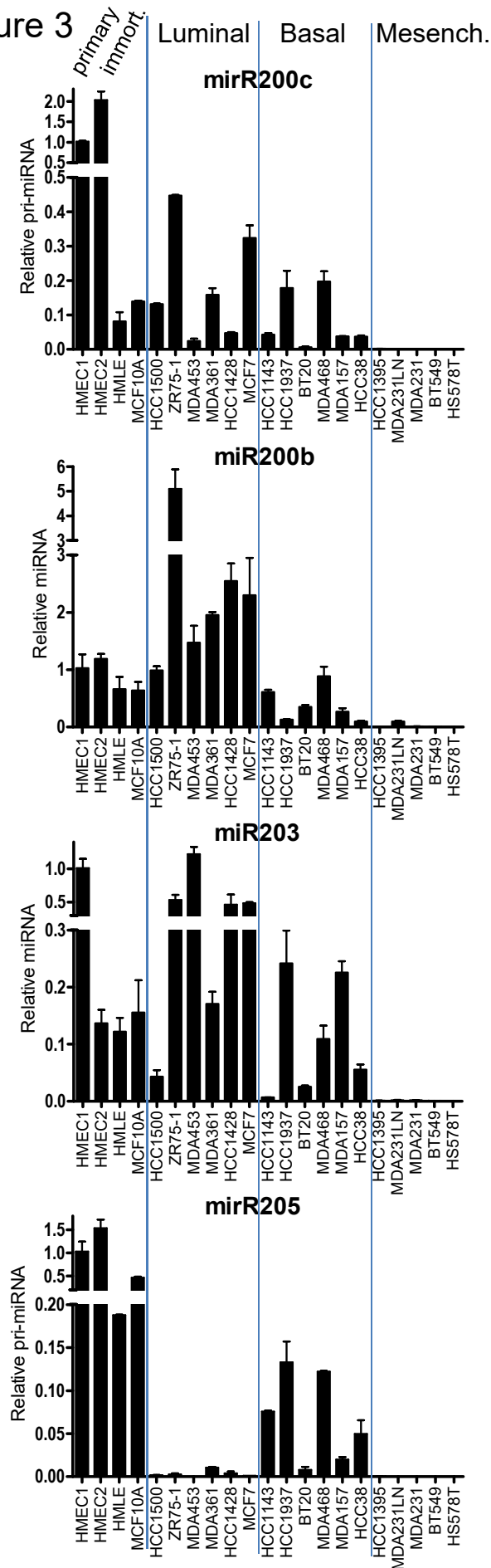
Supplementary Figure 2D



Supplementary Figure 2E



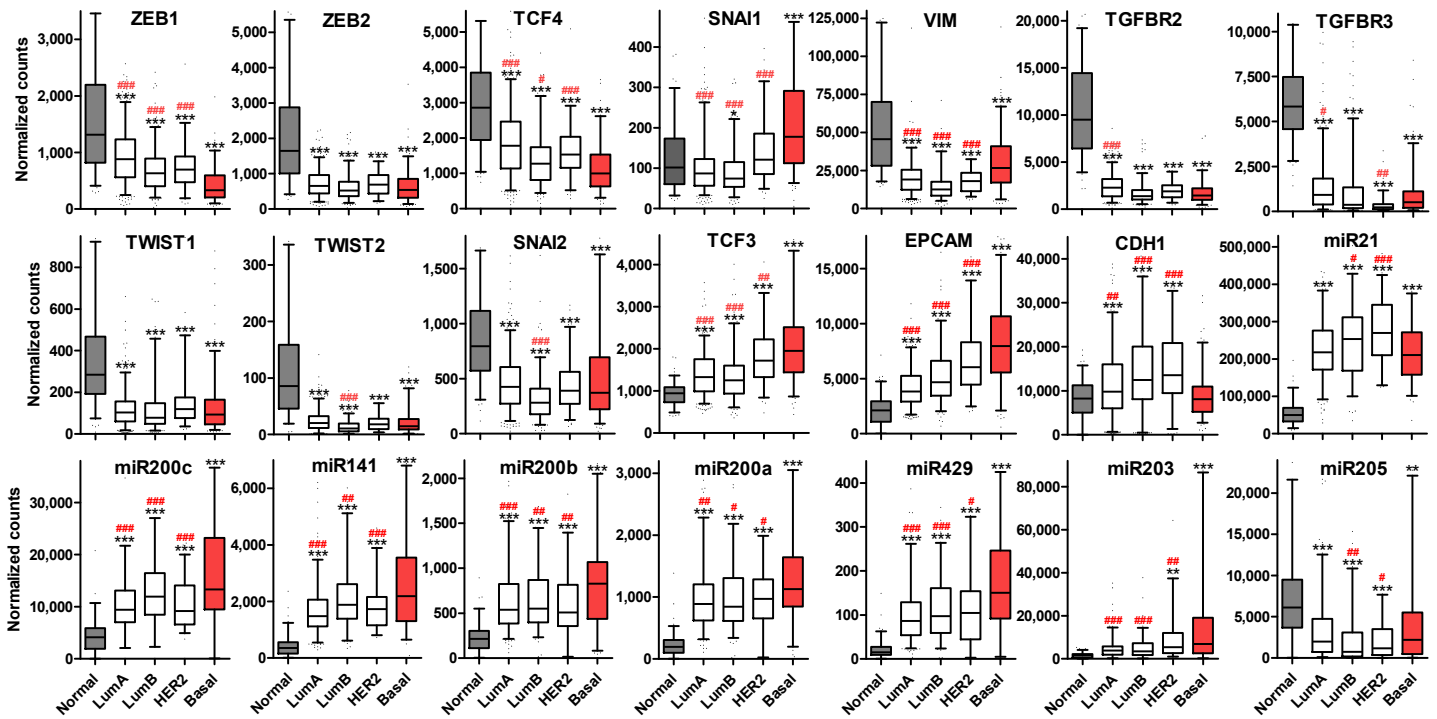
Suppl. Figure 3



**Supplementary Figure 3. Low expression of miR200s/203/205 microRNAs strongly correlates with the mesenchymal phenotype in breast cancer cell lines.** Relative levels of the indicated microRNAs in breast cancer cell lines shown at the X-axis determined by RT-qPCR. The mean values of the microRNAs expression in HMEC1 cells were set to 1. The data are mean + SEM of three biological replicates. HMEC1 and HMEC2 are primary human mammary epithelial cells; immort. – immortalized cell lines; Luminal, Basal and Mesenchymal cell line groups are indicated on the top.



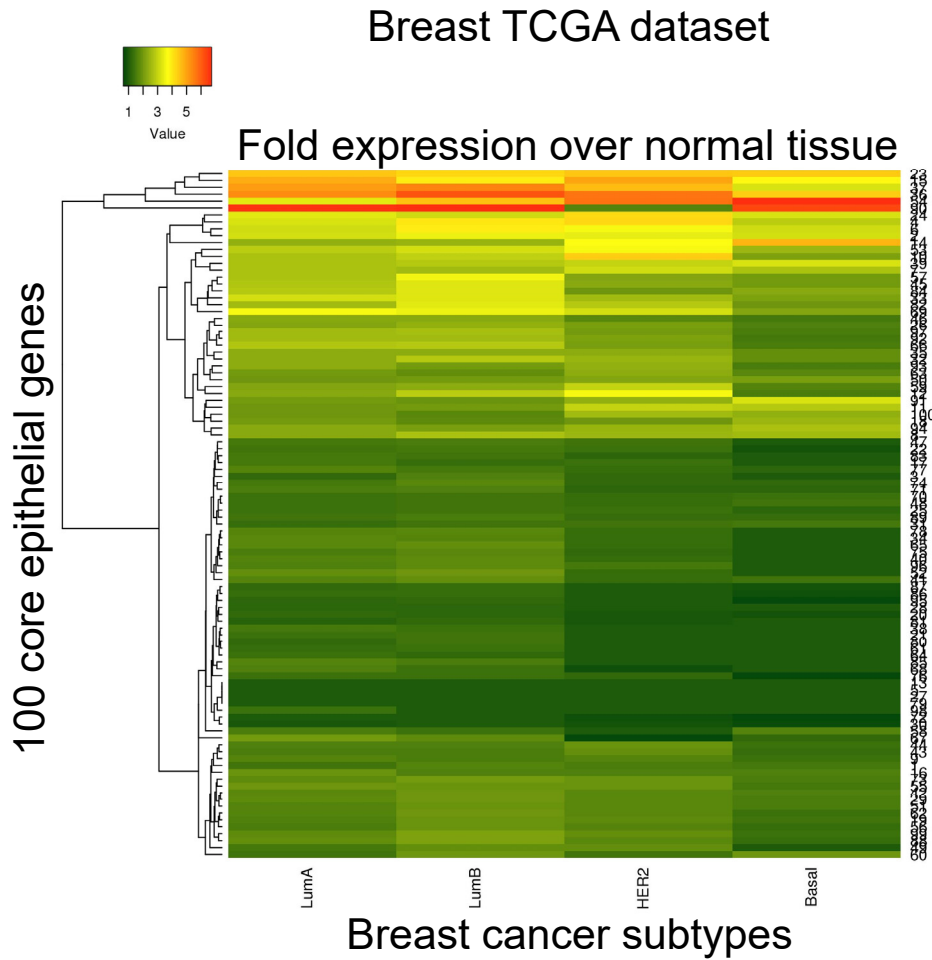
## Supplementary Figure 4



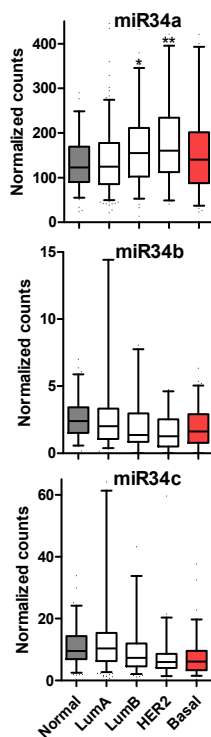
**Supplementary Figure 4. Differential expression of EMT-TFs and epithelial-specific microRNAs in human breast cancer subtypes from the TCGA Breast dataset.** Relative expression levels of the indicated genes in normal breast tissue and four breast cancer subtypes. Data shown as box plots with mean and 5-95% percentile. ANOVA with Dunnett's post-test, black asterisks (\*) indicate statistical significance in comparison to normal tissue (grey box plots), red pound signs (#) indicate statistical significance in comparison to basal subtype (red box plots), \* or # -  $p < 0.05$ , \*\* or ## -  $p < 0.01$ , \*\*\* or ### -  $p < 0.001$ .

# Supplementary Figure 5

## A



## B

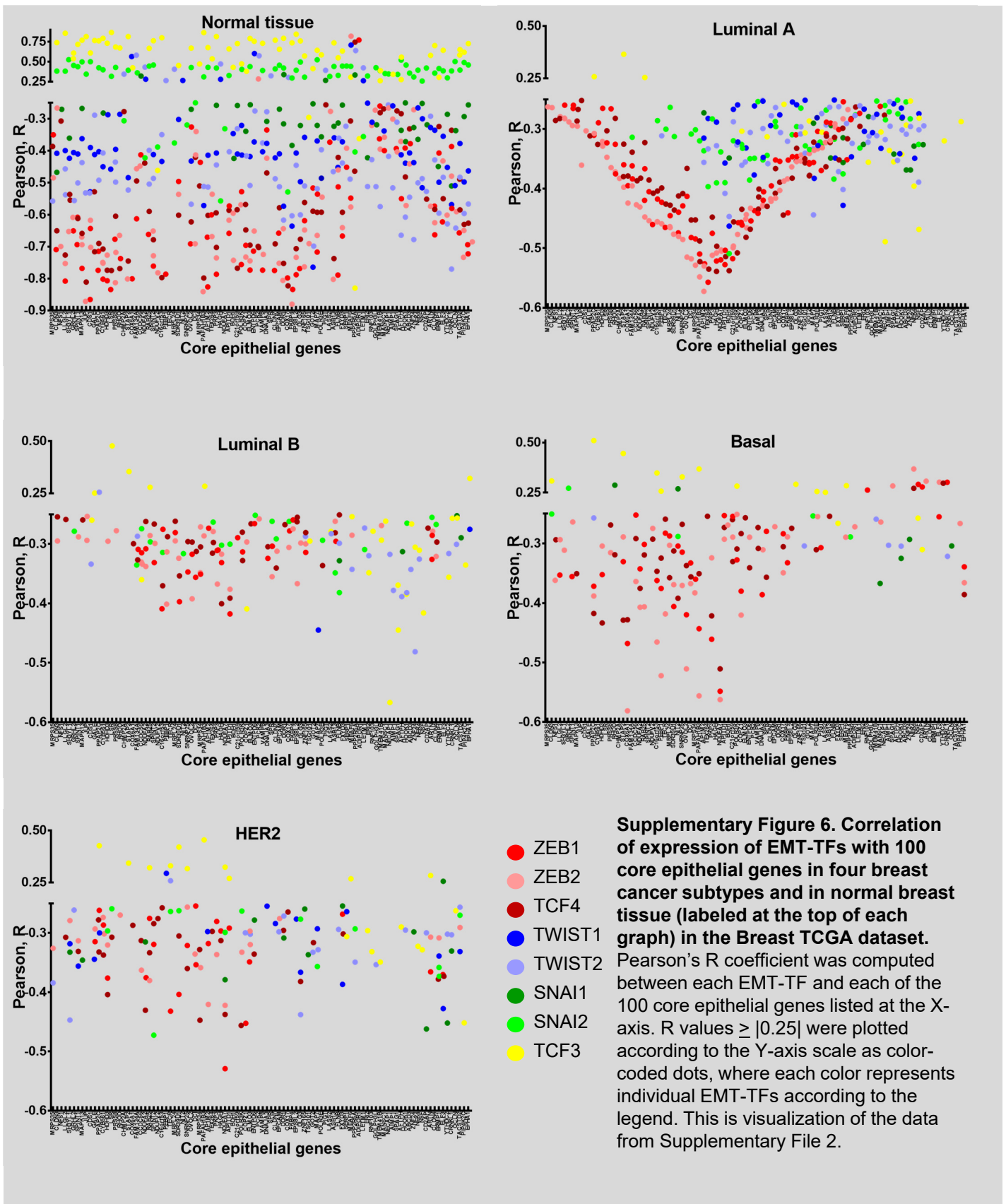


**Supplementary Figure 5. A, Higher expression of the majority of the 100 core epithelial genes in four intrinsic breast cancer subtypes compared to normal breast tissue.** Heat map of expression for 100 core epithelial genes listed in Supplementary File 1 was visualized by HeatMapper (<http://heatmapper.ca/expression/>) using Average Linkage and Manhattan Distance measurement methods.

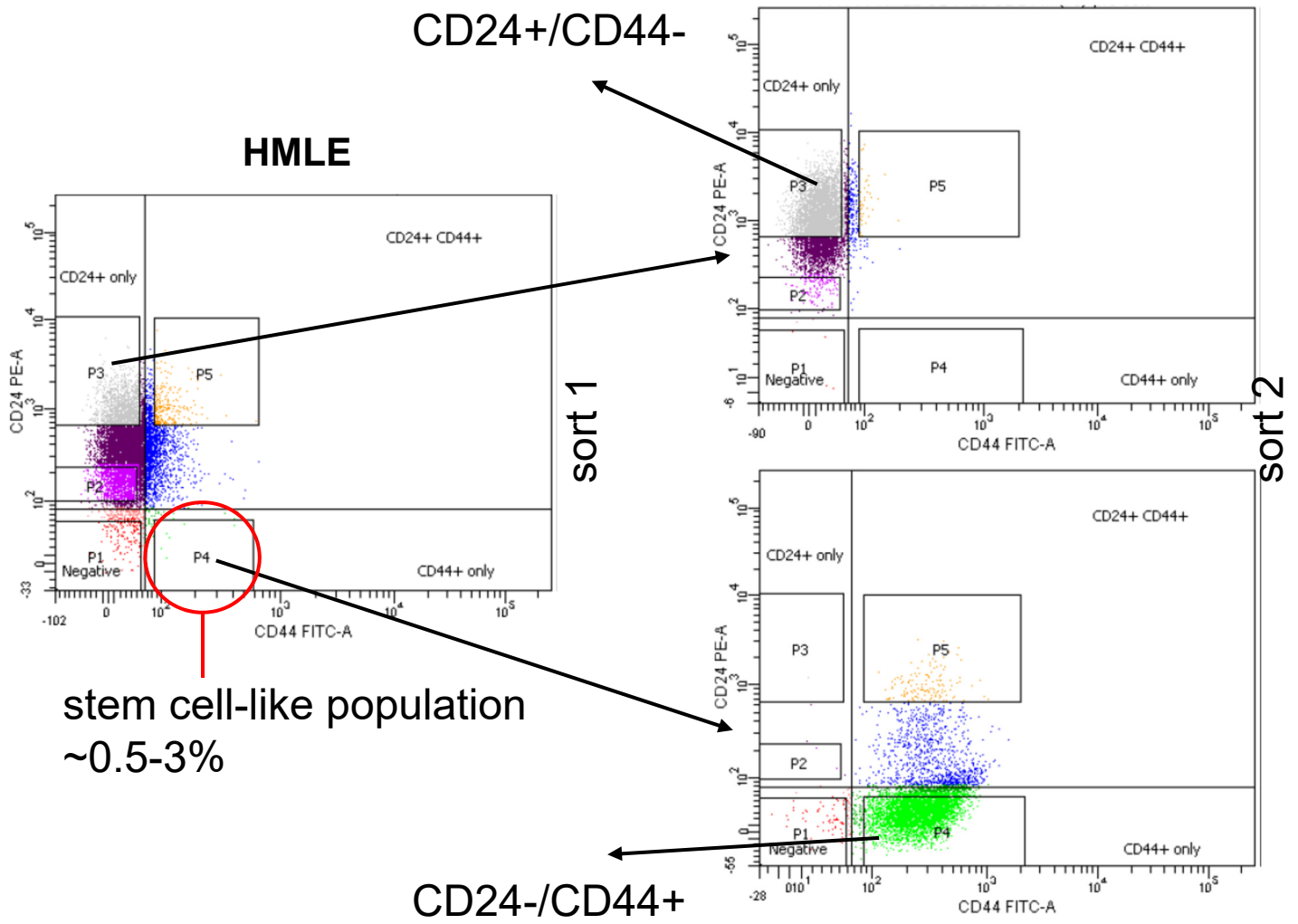
**B, Expression of the mir34 family microRNAs is similar between basal and other breast cancer subtypes.** Relative expression levels of the miR34 family microRNAs in normal breast tissue and four breast cancer subtypes. Data shown as box plots with mean and 5-95% percentile. ANOVA with Dunnett's post-test compared to normal tissue, \* -  $p < 0.05$ , \*\* -  $p < 0.01$ .

# Supplementary Figure 6

## Negative correlation of 100 core epithelial genes with eight EMT-TFs



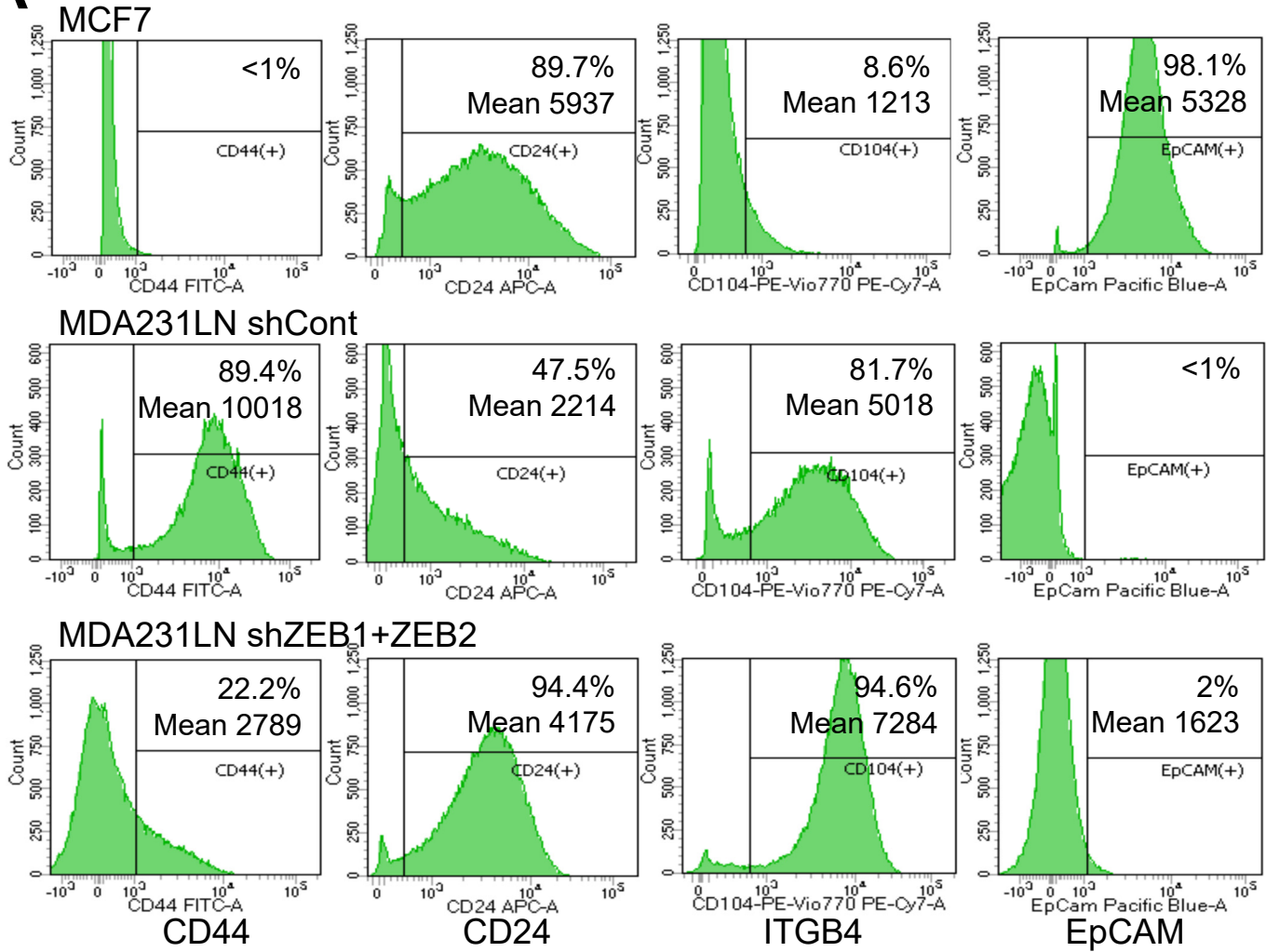
## Supplementary Figure 7



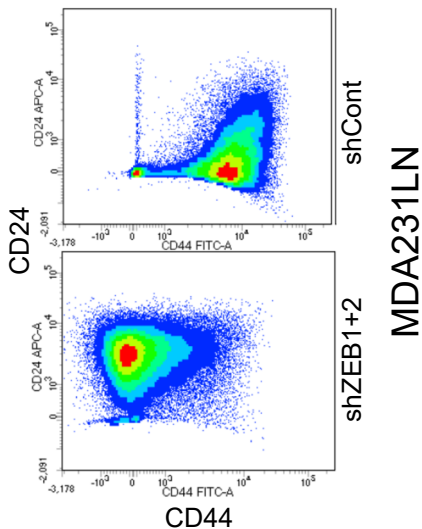
**Supplementary Figure 7. Fluorescence activated cell sorting (FACS) of CD24+/CD44- and CD24-/CD44+ subpopulations of HMLE cells.** Parental HMLE cells were profiled with CD24-PE and CD44-FITC antibodies and subpopulations of CD24-/CD44+ stem-like cells in the P4 quadrant and of CD24+/CD44- cells in the P3 quadrant were sorted out (left panel, sort 1). To insure relative homogeneity after in vitro expansion these cells were profiled again and the corresponding subpopulations were sorted out again (right panel, sort 2). CD24-PE (BD Pharmingen, 555428, CD44-FITC (BD Pharmingen, 555478) were used.

# Supplementary Figure 8

## A

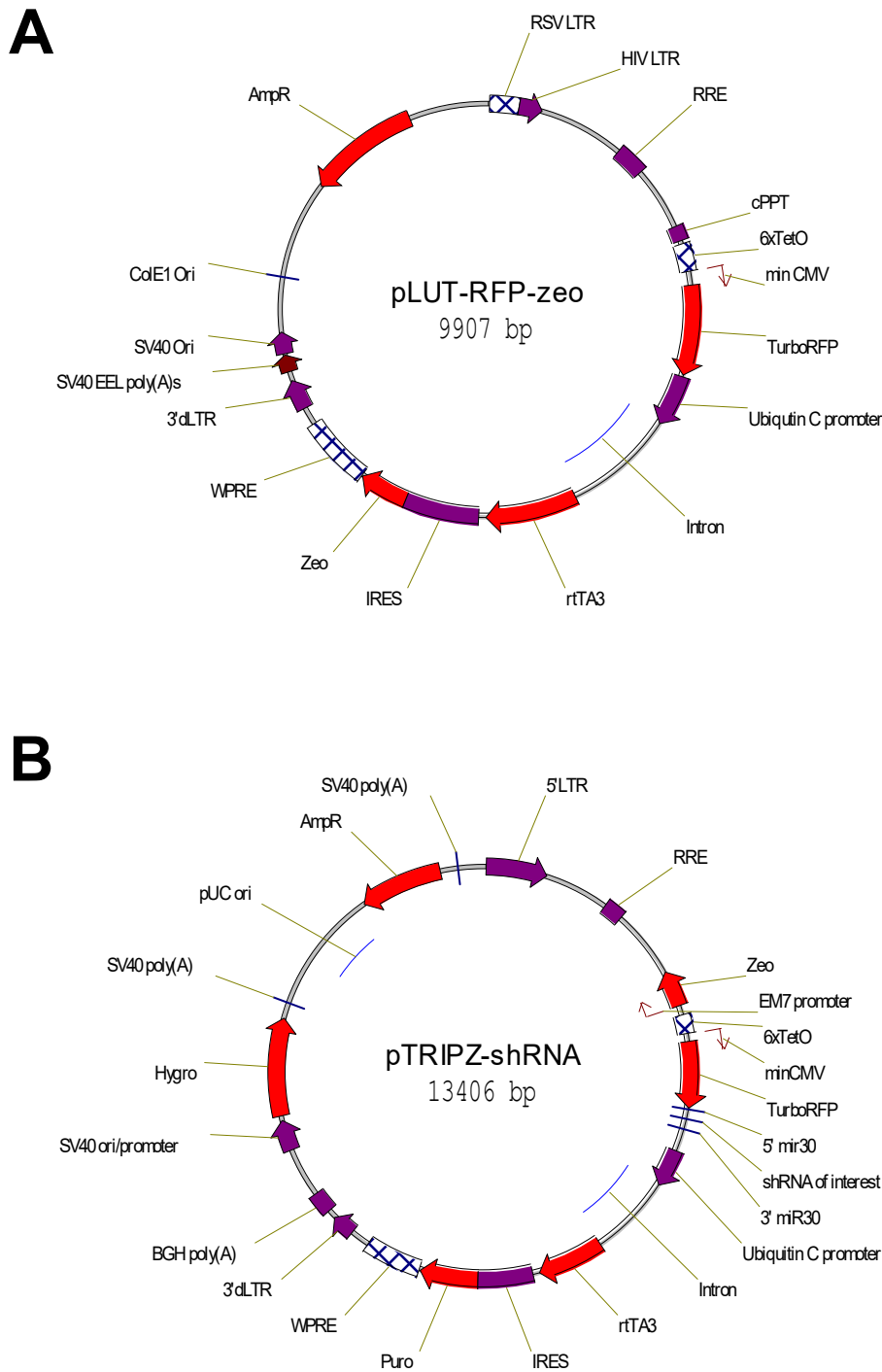


## B



**Supplementary Figure 8. Knockdown of ZEB1+ZEB2 in MDA231LN cells leads to characteristic change in expression of stemness markers. (A)** Flow cytometry analysis of MDA231LN/shCont and shZEB1+ZEB2 cells, and MCF7 epithelial cells for comparison: CD44-FITC, CD24-APC, CD104/ITGB4-PE-Vio770 and CD326/EpCAM-VioBlue cell-surface staining. Numbers indicate the percent of marker positive population, mean fluorescence is also shown. Single live cells were gated using FSC vs SSC, FSC-A vs FSC-W, and by Near-IR Live/Dead stain **(B)** Same cells as in (A) plotted with CD44 and CD24 axes.

## Supplementary Figure 9



**Supplementary Figure 9. Functional maps of Dox-inducible lentiviral expression vectors pLUT (A) and pTRIPZ (B).** Reverse tetracycline-controlled transactivator version 3 (rtTA3)-IRES-zeo cassette is expressed from the constitutive human Ubiquitin C promoter. EMT-TFs ORFs were cloned in pLUT instead of turboRFP. ORF or shRNA of interest are expressed from the tetracycline-dependent TRE promoter (six tetO operator sequences upstream of the minimal CMV promoter).