

Description of Additional Supplementary Files

Title: Supplementary Data 1: Study list

Description: Overview of studies identified during our search.

Title: Supplementary Data 2: List of the ADC targets

Description: Overview of identified targets with target name used in article, more commonly used protein name if available and if the target has been used in a breast cancer specific trial.

Title: Supplementary Data 3: Repeated sampling

Description: Data of the repeated sampling experiment. The best model is reported in column C. Of note, genes having linear and non-linear as best model were among the lowest expressed genes or changes were too subtle to be further considered.

Abbreviations: const: constant, lin: linear, nnlin: nonlinear, RT: Room Temperature

Title: Supplementary Data 4.1-4.4: Data figure 3 and supplementary figure 5B

Description: Data subgroup analysis observing differences in target gene expression.

- S6.1: subgroup analysis of metastases vs primary untreated samples
- S6.2: subgroup analysis of ILC vs IBC-NST metastases
- S6.3: subgroup analysis of HR+/HER2- vs TNBC metastases
- S6.4: subgroup analysis of ER+ vs ER- metastases

p-values were derived from a two-sided Wald's test. Source data and exact p-values/FDR are provided as a Source Data file.

Abbreviations: CI: confidence Interval, ER: estrogen receptor, FDR: false discovery rate, HR: hormone receptor, IBC-NST: invasive breast carcinoma of no special type, ILC: invasive lobular carcinoma, M : metastatic, NonSig: non-significant, P : primary untreated, Sig: significant, TNBC: triple negative breast carcinoma

Title: Supplementary Data 5: Data figure 4 and supplementary figure 6

Description: Data subgroup analysis observing differences in organ specific expression. p-values were derived from a two-sided Wald's test. Source data and exact p-values/FDR are provided as a Source Data file.

Abbreviations: CI: confidence Interval, FDR: false discovery rate, NonSig: non-significant, Sig: significant

Title: Supplementary Data 6.1-6.3: Data figure 5 and supplementary figure 7

Description: Data of the weighted correlation.

- S8.1: Weighted correlation across all metastases
- S8.2: Weighted correlation across all metastases in patients with HR+/HER2- BC
- S8.3: Weighted correlation across all metastases in patients with TNBC

Source data are provided as a Source Data file.

Abbreviations: HR: hormone receptor, NA: not applicable, TNBC: triple negative breast carcinoma

Title: Supplementary Data 7: Evidence of statistical significance in overlap of gene expression between normal tissue and metastases after 10,000 permutations.

Description: Data describes the original expression overlap per target between normal tissue and metastases as a continuous variable. This information was also reported as a category in figure 6B. P-values represent the evidence of statistical significance of the reported overlap when 10,000 permutations were applied on a pooled dataset containing expression values of both metastases and normal tissue where sampling was performed between two groups that reflected the sample size of metastases (n = 749 samples from 30 patients) and normal tissue (n = 96 samples from 24 patients). p-values were derived from a two-sided Wald's test. Source data and exact p-values/FDR are provided as a Source Data file.

Title: Supplementary Data 8: immunohistochemical assessment

Description: Data on immunohistochemical staining of targets of interest with corresponding mRNA expression values. Information on tissue type is available as organ of origin (column C) and ICD-O-3-code (column D). mRNA expression of the target is available in column E-L. Percentage of positive cells per target of interest is available in column M and O-U with details on HER2 status in column N. If a sample was used in figure 8 (metastatic samples) or supplementary figure 9 (metastatic and normal samples) this is mentioned in column V.

Abbreviations: ICD: international classification of diseases for oncology

(<https://www.who.int/standards/classifications/other-classifications/international-classification-of-diseases-for-oncology>), TROP2: trophoblast cell surface antigen 2, VTCN1: V-set domain-containing T-cell activation inhibitor 1, HER: human epidermal growth factor receptor, FN1: fibronectin 1, MUC1: mucin 1, FGFR2: fibroblast growth factor receptor 2, pc: percentage