Supplementary Data

We downloaded and analyzed data publicly available from the Cancer Genome Atlas Project (TCGA; http://tcga-data.nci.nih.gov/) for 11 tumor types: Bladder Urothelial Carcinoma, Breast Invasive Carcinoma (BRCA), Brain Lower-Grade Glioma, Colon Adenocarcinoma, Head and Neck Squamous Cell Carcinoma, Kidney Clear Cell Carcinoma, Lung Adenocarcinona, Lung Squamous Cell Carcinoma, Ovarian Serous Cystadenocarcinoma, Rectum Adenocarcinoma, and Uterine Corpus Endometrial Carcinoma. Level 3 Illumina RNASeq, respectively microRNA (miRNA)Seq, were used for mRNA, miRNA expression. For miRNASeq data, we derived the "reads_per_million_miRNA_mapped" values for mature forms for each miRNA from the "isoform_quantification" files.

The clinical information for patients with breast cancer was downloaded from http://tcga-data.nci.nih.gov/docs/publications/brca_2012/ (clinical data associated to the paper: Comprehensive molecular portraits of human breast tumors, Nature, September 27, 2012).

Statistical analyses were performed in R (version 2.14.2) (http://r-project.org/), and the statistical significance was defined as a p-value less than 0.05. The Log-rank test was employed to determine the relationship between expression and overall survival, and the Kaplan–Meyer method was used to generate survival curves. We randomly split the entire population into training/validation cohorts (2/3, 1/3) and for each miRNA, we checked for a relation with the survival as follows. In both cohorts, patients were divided into percentiles according to miRNA expression. Using the training set, we considered any cut-off value between 25th and 75th to significantly split the samples into two groups and checked for statistical significance in the validation set. We then chose the cut-off value to optimally split the samples in both cohorts.

The Spearman's rank-order correlation test was applied to measure the strength of the association between miR-210 and glycerol-3-phosphate dehydrogenase 1-like.