

Regulatory landscapes of specific miRNAs are conserved between cell lines and primary tumors

Hanwen Zhu¹, Boting Ning^{2*}

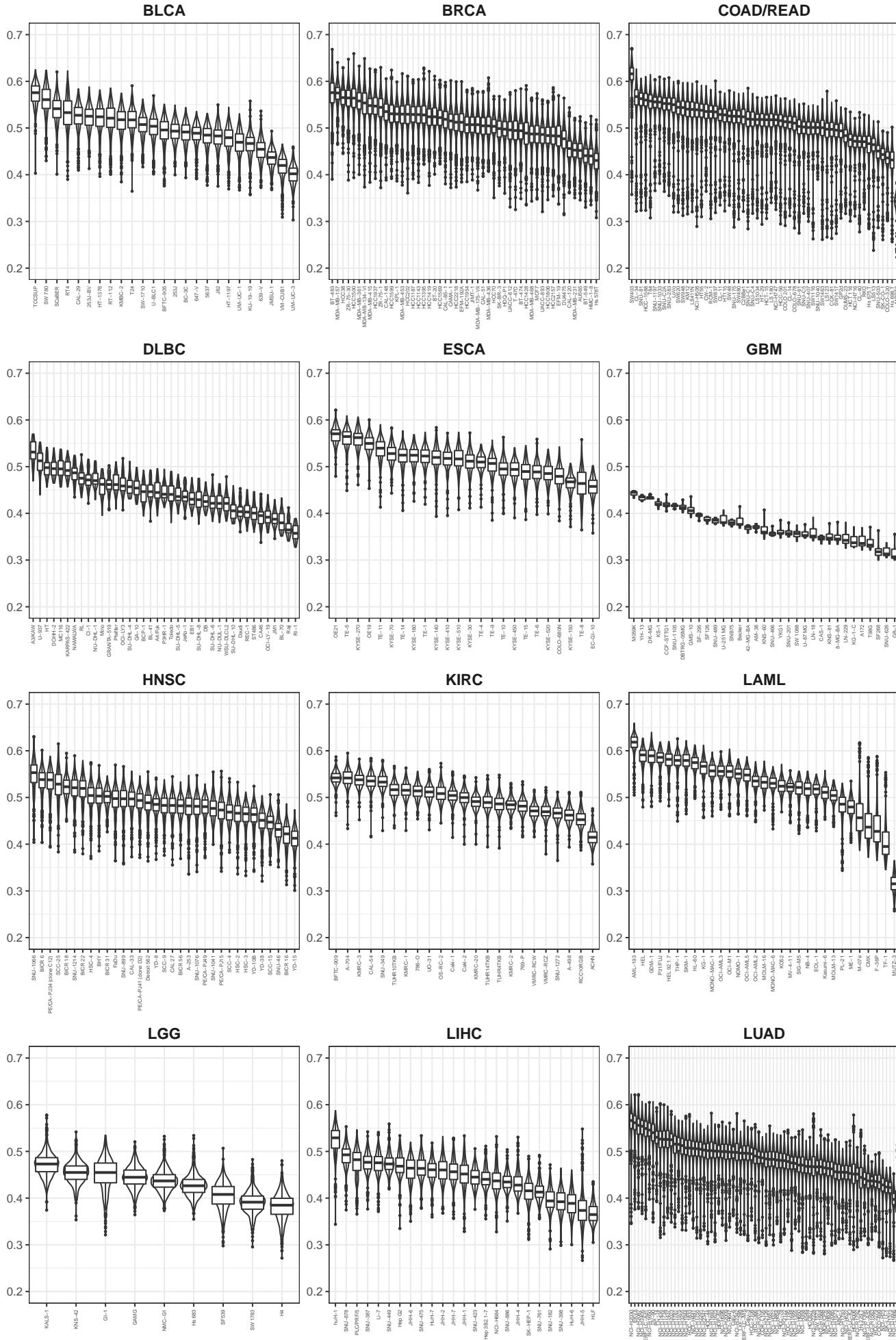
¹YK Pao School, Shanghai, China.

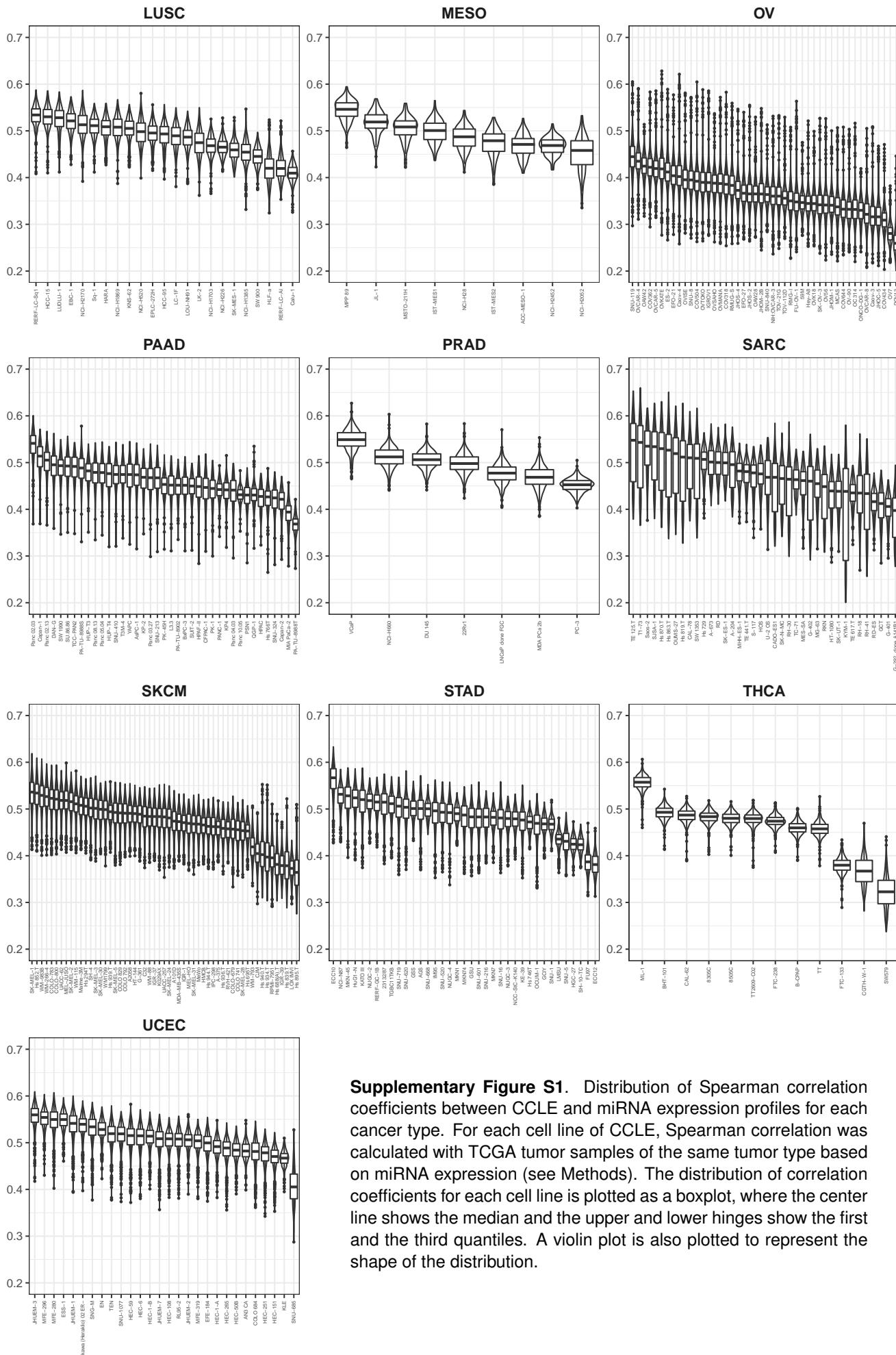
²Boston University School of Medicine, Boston, MA, USA.

*email: ningb@bu.edu

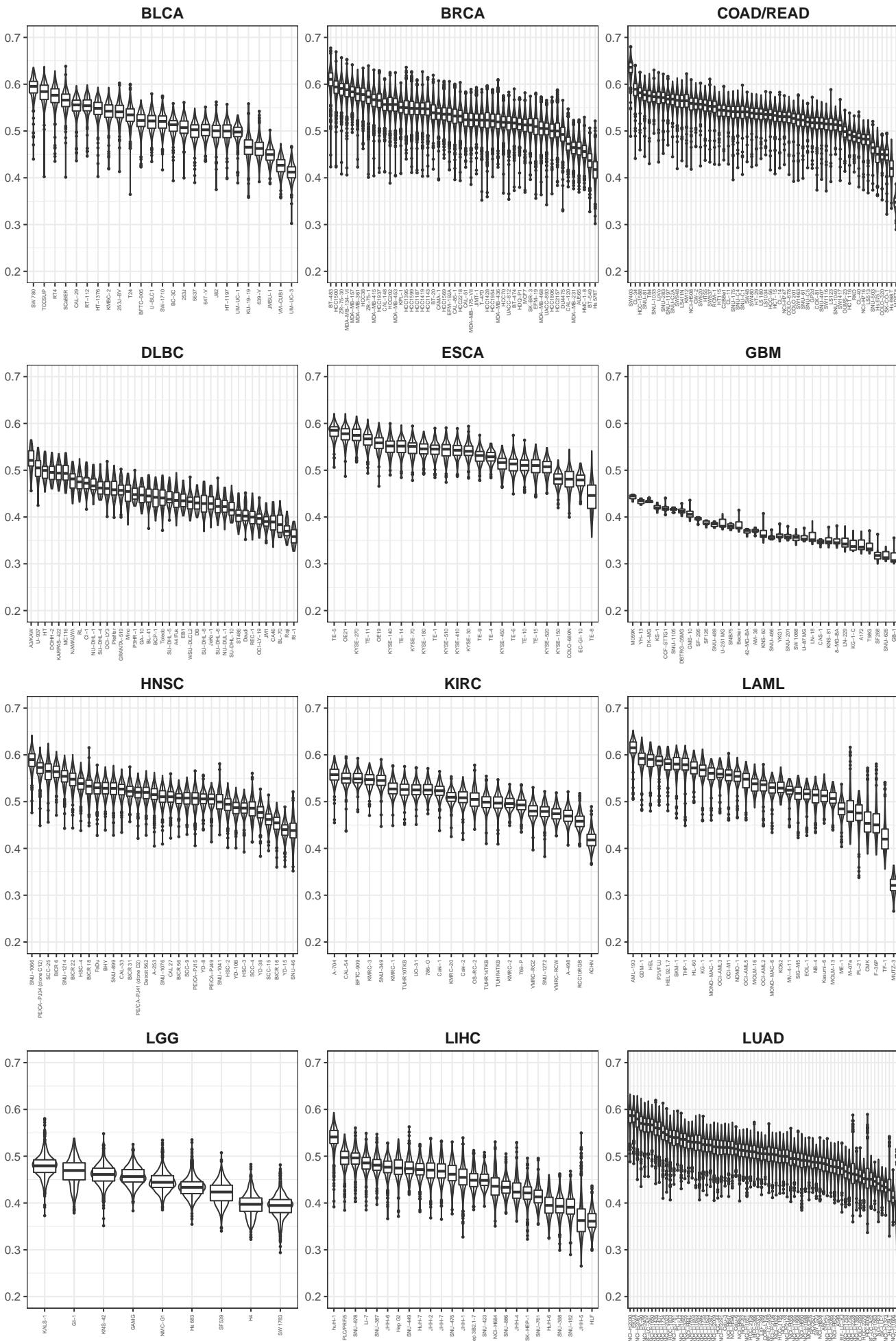
Study Abbreviation	Study Name
BRCA	Breast invasive carcinoma
COADREAD	Colorectal adenocarcinoma
KIRC	Kidney renal clear cell carcinoma
THCA	Thyroid carcinoma
HNSC	Head and neck squamous cell carcinoma
UCEC	Uterine corpus endometrial carcinoma
LUAD	Lung adenocarcinoma
PRAD	Prostate adenocarcinoma
LGG	Brain lower grade glioma
LUSC	Lung squamous cell carcinoma
OV	Ovarian serous cystadenocarcinoma
STAD	Stomach adenocarcinoma
SKCM	Skin cutaneous melanoma
BLCA	Bladder urothelial carcinoma
LIHC	Liver hepatocellular carcinoma
KIRP	Kidney renal papillary cell carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
SARC	Sarcoma
ESCA	Esophageal carcinoma
LAML	Acute myeloid leukemia
PCPG	Pheochromocytoma and paraganglioma
PAAD	Pancreatic adenocarcinoma
TGCT	Testicular germ cell tumors
THYM	Thymoma
KICH	Kidney chromophobe
MESO	Mesothelioma
UVM	Uveal melanoma
ACC	Adrenocortical carcinoma
UCS	Uterine carcinosarcoma
DLBC	Lymphoid neoplasm diffuse large b-cell lymphoma
CHOL	Cholangiocarcinoma
GBM	Glioblastoma multiforme
NB	Neuroblastoma
MM	Multiple myeloma
LCML	Chronic myelogenous leukemia
MB	Medulloblastoma
CLL	Chronic lymphocytic leukemia

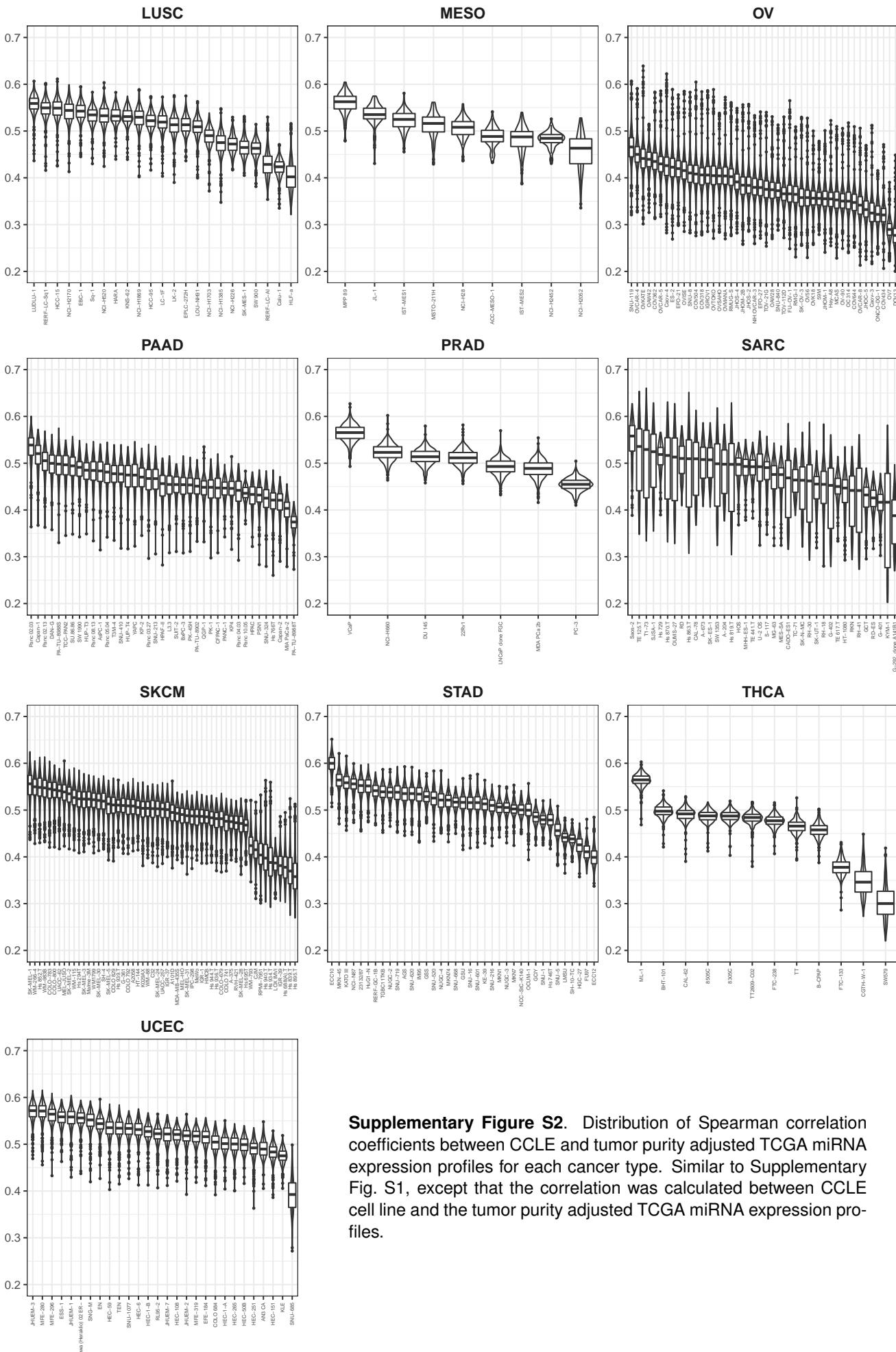
Supplementary Table S1. TCGA and CCLE tumor type abbreviations.

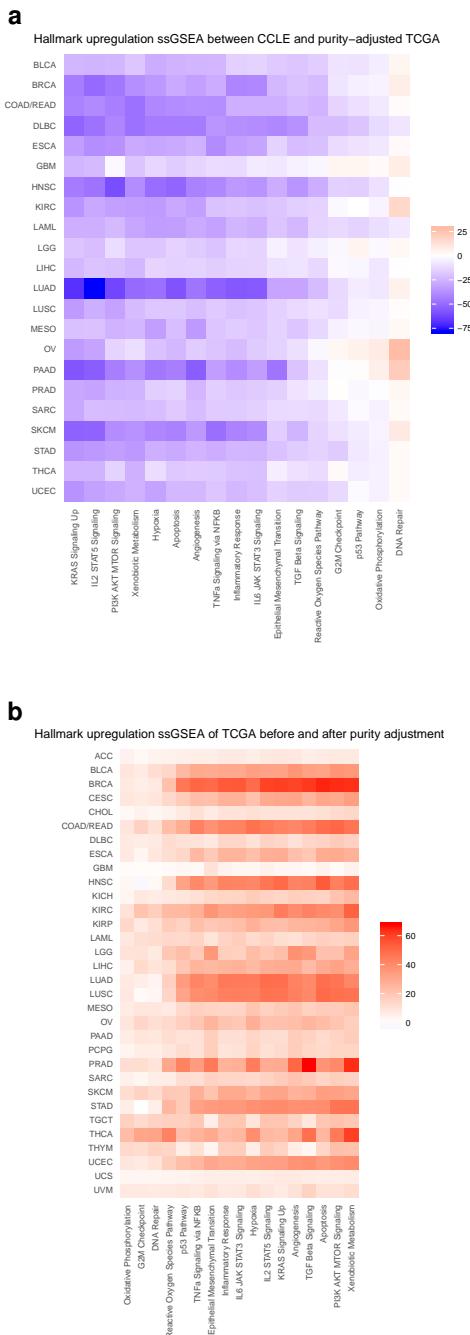




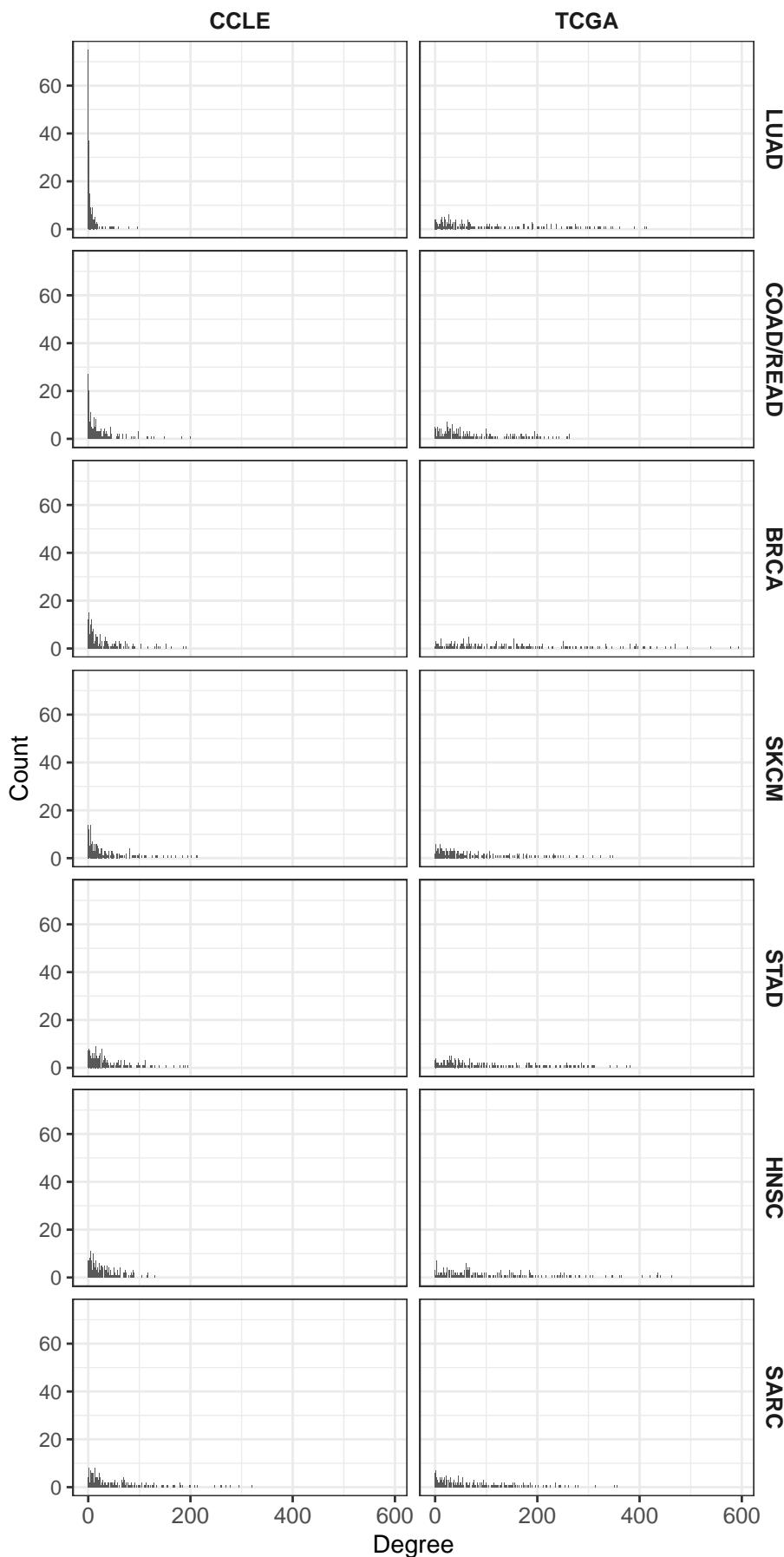
Supplementary Figure S1. Distribution of Spearman correlation coefficients between CCLE and miRNA expression profiles for each cancer type. For each cell line of CCLE, Spearman correlation was calculated with TCGA tumor samples of the same tumor type based on miRNA expression (see Methods). The distribution of correlation coefficients for each cell line is plotted as a boxplot, where the center line shows the median and the upper and lower hinges show the first and the third quartiles. A violin plot is also plotted to represent the shape of the distribution.



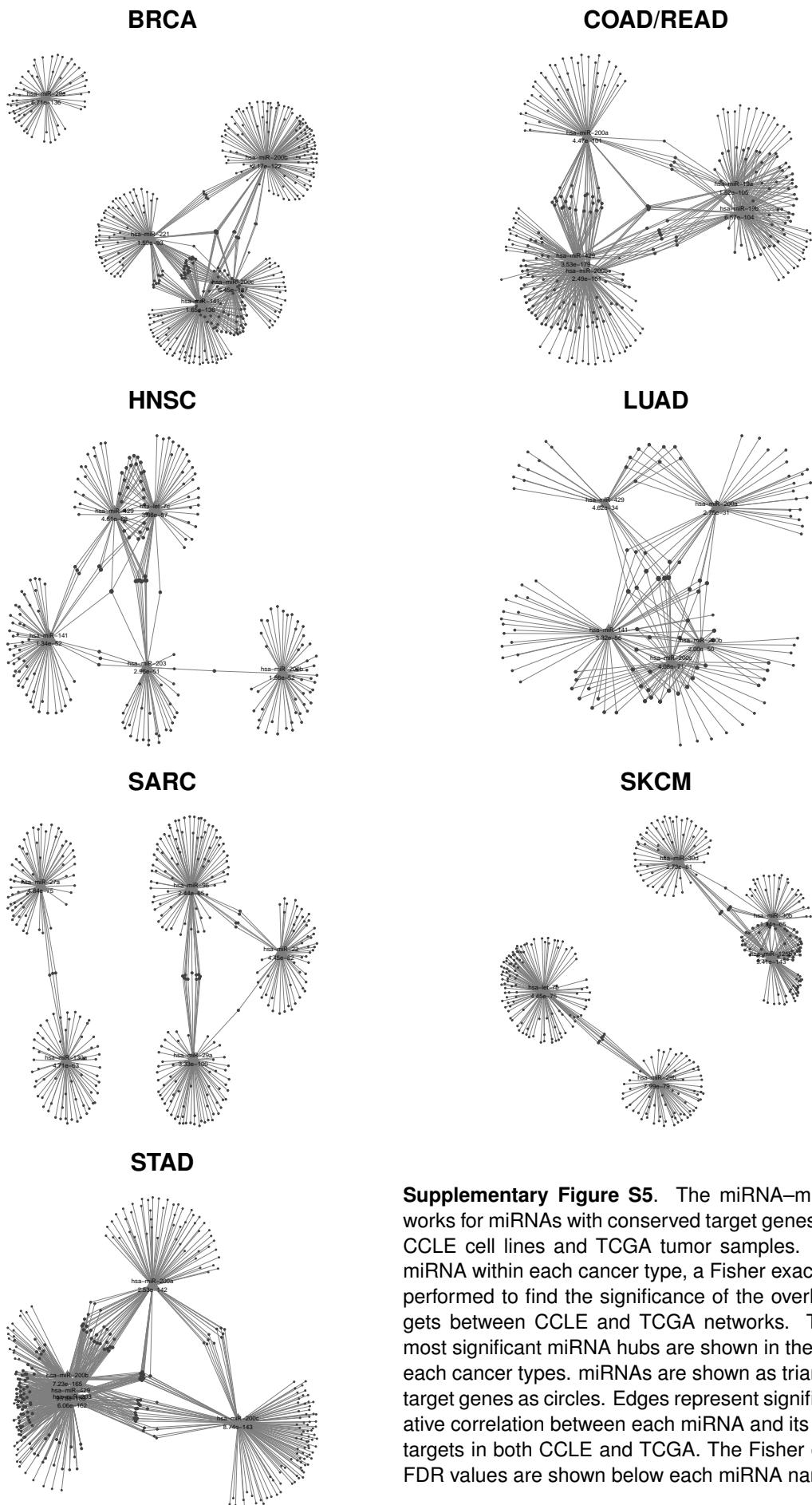




Supplementary Figure S3. Comparison of GSVA scores for cancer hallmark upregulation based on the miRNA expression of TCGA and CCLE. The t -tests of the GSVA scores of upregulated hallmarks of CCLE minus those of tumor-purity adjusted TCGA (**a**) and the scores of TCGA before purity adjustment minus those of TCGA after adjustment were calculated (**b**) and visualized as heatmaps.



Supplementary Figure S4. The miRNA out-degree distribution in target networks of selected cancer types. The degree distribution is plotted for CCLE and TCGA separately. The selection criteria for cancer types and methods for constructing miRNA–mRNA networks are described in the Methods section.



Supplementary Figure S5. The miRNA–mRNA networks for miRNAs with conserved target genes between CCLE cell lines and TCGA tumor samples. For each miRNA within each cancer type, a Fisher exact test was performed to find the significance of the overlap of targets between CCLE and TCGA networks. The top 5 most significant miRNA hubs are shown in the figure for each cancer types. miRNAs are shown as triangles and target genes as circles. Edges represent significant negative correlation between each miRNA and its predicted targets in both CCLE and TCGA. The Fisher exact test FDR values are shown below each miRNA name.