CHEMISTRY & BIODIVERSITY

Supporting Information

NetCellMatch: Multiscale Network-Based Matching of Cancer Cell Lines to Patients Using Graphical Wavelets

Neel Desai,* Jeffrey S. Morris, and Veerabhadran BaladandayuthapaniPart of a Special Collection on

Supplementary File: NetCellMatch

S.1 Data Pre-Processing

For each cancer type, cell line and patient data were pre-processed using the ComBat method. For proteomic markers with missing values, the median for that marker was imputed for patients with missing values. If the entire column for that proteomic marker was missing, the marker was omitted from analysis for that cancer type.

S.2 NetCellMatch Software

R-code for NetCellMatch is available at: https://bayesrx.github.io/. Software includes a worked example (for patients and cell lines with Breast Cancer) along with all functions and their corresponding descriptions via R-markdown files

S.3 Scatterplots for Proteomic Drivers

We display below scatterplots for proteomic drivers referenced in NetCellMatch but not displayed in the body of the paper
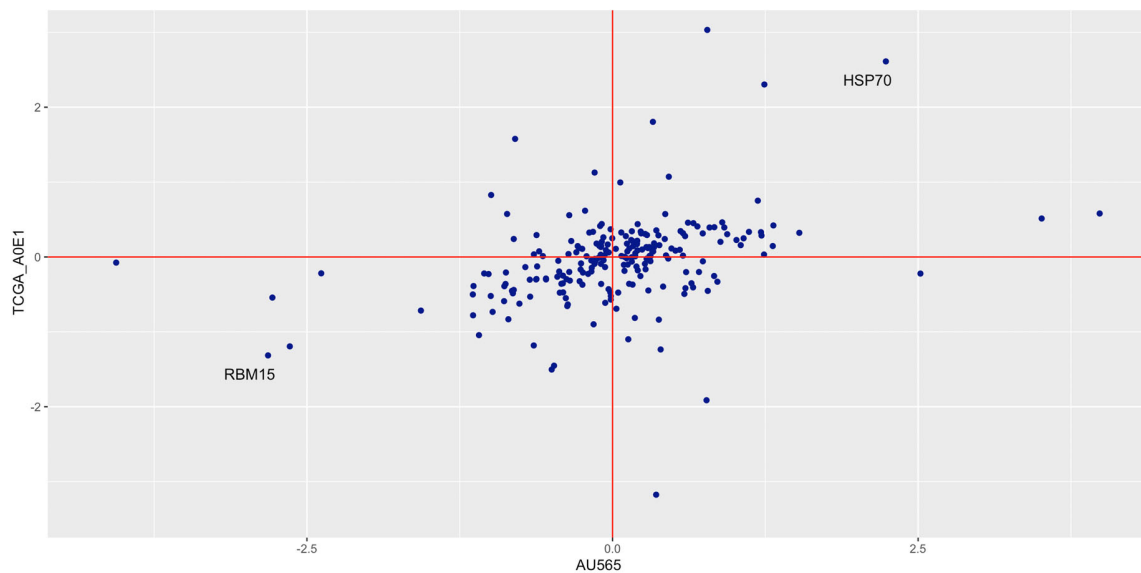
Breast Cancer

Figure S.3.1 - Scatterplot comparing proteomic expression levels of potential avatar cell line AU565 on the x-axis and top connecting patient TCGA_A0E1. 2 biomarkers are highlighted as being mutually aberrantly expressed: RBM15 and HSP70.
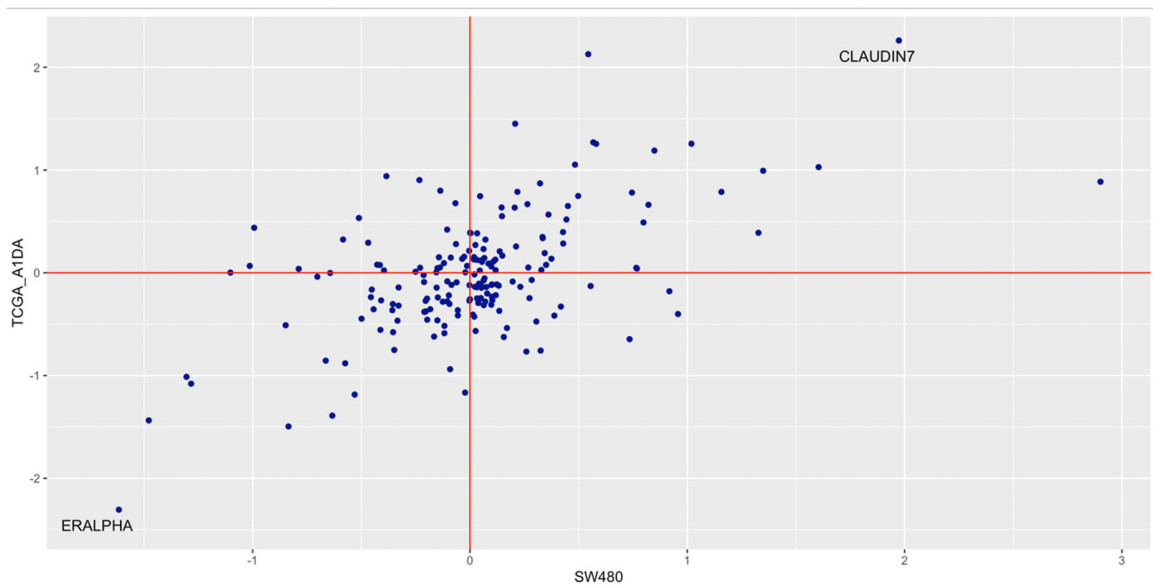
Colorectal Cancer

Figure S.3.2 - Scatterplot comparing proteomic expression levels of potential avatar cell line SW480 on the x-axis and top connecting patient TCGA_A1DA. 2 biomarkers are highlighted as being mutually aberrantly expressed: CLAUDIN7 and ER-α.
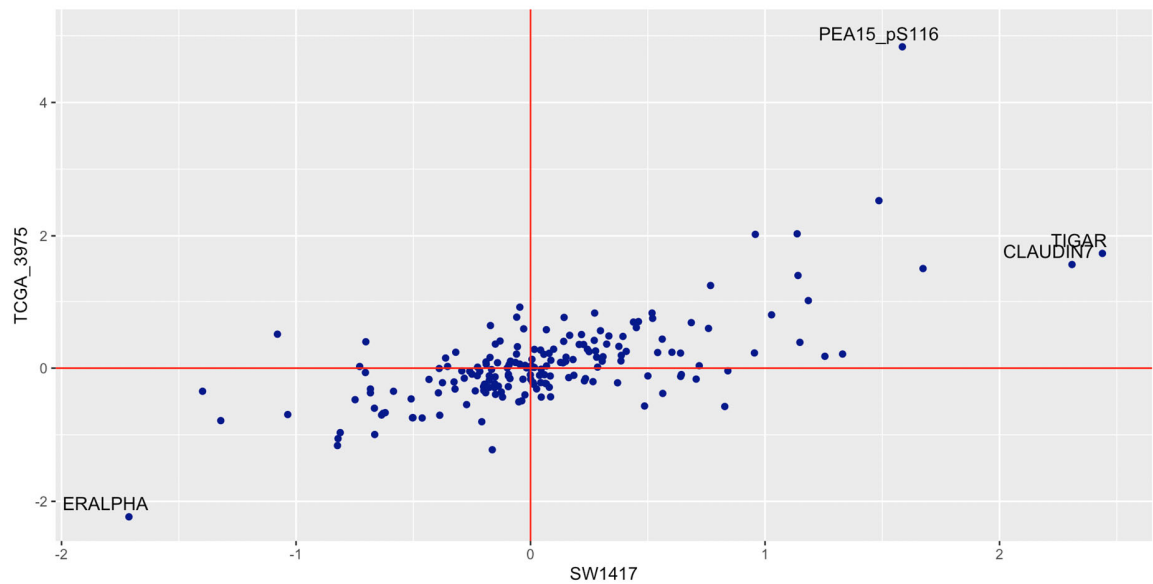

Figure S.3.3 - Scatterplot comparing proteomic expression levels of potential avatar cell line SW1417 on the x-axis and top connecting patient TCGA_3975. 4 biomarkers are highlighted as being mutually aberrantly expressed: ER-α, PEA15_pS116, CLAUDIN7, and TIGAR
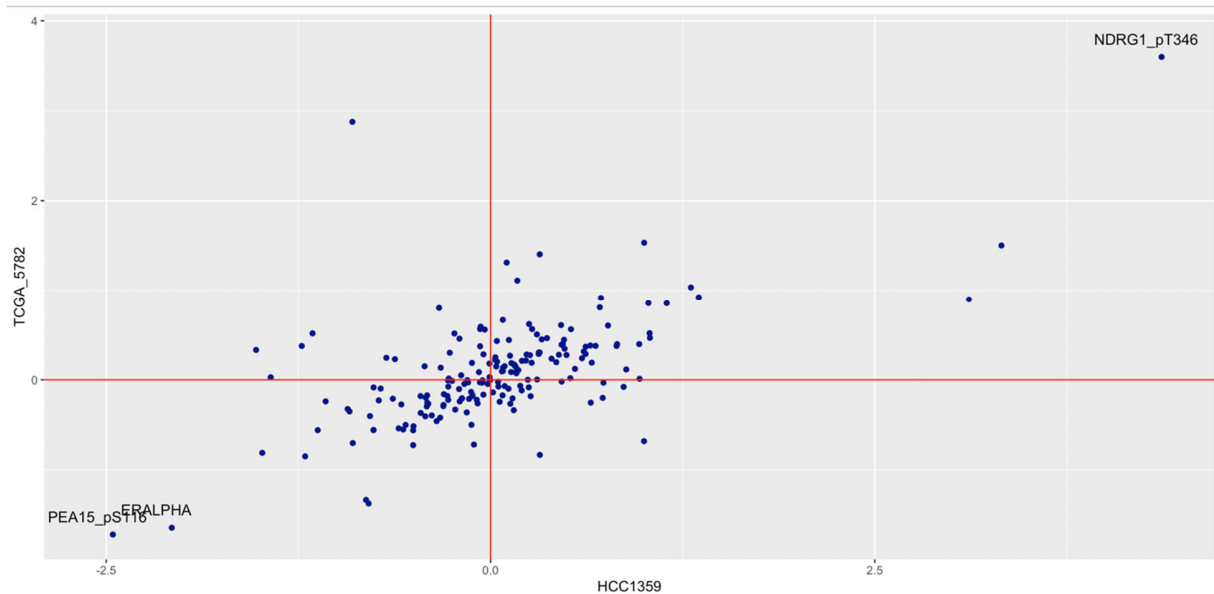
Lung Cancer

Figure S.3.4 - Scatterplot comparing proteomic expression levels of potential avatar cell line HCC1359 on the x-axis and top connecting patient TCGA_5782. 3 biomarkers are highlighted as being mutually aberrantly expressed: PEA15_pS116, ER-α, and NDRG1_pT346

S4. Cell Line and Patient Information

Information regarding cell lines is obtained from https://www.atcc.org/products and https://web.expasy.org/cellosaurus

Detailed information regarding patient tumor samples can be found at https://portal.gdc.cancer.gov/

S.5 Comprehensive Matching Scores for NetCellMatch and for Aggregated Hierarchical Clustering Approaches

For each of Breast, Colon, and Lung Cancer we list below all non-zero NBMS scores for analyses run in the main body of this paper. To compare NBMS scores derived from NetCellMatch, we also present all non-zero normalized matching scores derived by aggregating hierarchical clustering results across dendrogram cut points using two different distance metrics: correlation distance and Euclidean distance. The number of cut points aggregated across was

determined by the combined number of patients and cell lines available for that particular cancer type and average linkage was used wherever hierarchical clustering was performed. We conclude this section by comparing the 3 sets of results. It should be noted that all information for reading patient tumor labels can be found at:
https://docs.gdc.cancer.gov/Encyclopedia/pages/TCGA_Barcode/.

Breast Cancer

NetCellMatch

Patient Tumor	Cell Line	NBMS
TCGA-A2-A4S2_BRCA	ZR75T	1
TCGA-D8-A27H_BRCA	HS578T	0.80952381
TCGA-BH-A0E1_BRCA	AU565	0.71428571
TCGA-A8-A091_BRCA	SUM190	0.57142857
TCGA-D8-A1JM_BRCA	HCC1395	0.52380952
TCGA-B6-A0IO_BRCA	BT474	0.42857143
TCGA-BH-A0BQ_BRCA	MDAMB134VI	0.42857143
TCGA-A8-A08Z_BRCA	HCC70	0.38095238
TCGA-C8-A26Y_BRCA	HCC1395	0.33333333
TCGA-EW-A1OZ_BRCA	MDAMB157	0.33333333
TCGA-B6-A0IO_BRCA	MDAMB361	0.33333333
TCGA-A2-A4RX_BRCA	BT549	0.33333333
TCGA-C8-A26Y_BRCA	MDAMB468	0.28571429
TCGA-D8-A1JM_BRCA	BT549	0.28571429
TCGA-AR-A2LJ_BRCA	MCF10A	0.28571429
TCGA-C8-A26Y_BRCA	HCC70	0.23809524
TCGA-AC-A8OQ_BRCA	CAL51	0.23809524
TCGA-C8-A26Y_BRCA	CAL51	0.23809524
TCGA-A2-A0CV_BRCA	MDAMB175VII	0.23809524
TCGA-D8-A1JF_BRCA	EVSAT	0.23809524
TCGA-BH-A42T_BRCA	UACC893	0.23809524
TCGA-A2-A0T4BRCA	184B5	0.19047619
TCGA-AR-A1AM_BRCA	MDAMB468	0.19047619
TCGA-D8-A142_BRCA	JIMT1	0.19047619
TCGA-BH-A1EN_BRCA	MFM223	0.19047619
TCGA-BH-A0HK_BRCA	UACC732	0.19047619
TCGA-E9-A1R3_BRCA	BT474	0.14285714

TCGA-AR-A2LJ_BRCA	184B5	0.14285714
TCGA-D8-A27R_BRCA	CAL148	0.14285714
TCGA-BH-A0C0_BRCA	CAL148	0.14285714
TCGA-BH-A208_BRCA	CAL51	0.14285714
TCGA-D8-A1XT_BRCA	CAL51	0.14285714
TCGA-D8-A1XT_BRCA	JIMT1	0.14285714
TCGA-EW-A2FR_BRCA	JIMT1	0.14285714
TCGA-D8-A27R_BRCA	MFM223	0.14285714
TCGA-OL-A5DA_BRCA	AU565	0.14285714
TCGA-D8-A1JM_BRCA	HCC1187	0.14285714
TCGA-OL-A66O_BRCA	MCF7	0.14285714
TCGA-BH-A1FN_BRCA	HCC1954	0.14285714
TCGA-D8-A142_BRCA	HCC1954	0.14285714
TCGA-OL-A66O_BRCA	BT474	0.0952381
TCGA-AC-A6NO_BRCA	BT474	0.0952381
TCGA-BH-A42T_BRCA	HCC70	0.0952381
TCGA-AO-A0JD_BRCA	HDQP1	0.0952381
TCGA-AN-0XO_BRCA	MDAMB175VII	0.0952381
TCGA-BH-A0W4_BRCA	MDAMB361	0.0952381
TCGA-GM-A5PV_BRCA	MDAMB415	0.0952381
TCGA-S3-AA11_BRCA	MDAMB453	0.0952381
TCGA-LL-A9Q3_BRCA	ZR7530	0.0952381
TCGA-S3-AA14_BRCA	ZR7530	0.0952381
TCGA-E2-A15G_BRCA	ZR7530	0.0952381
TCGA-BH-A42T_BRCA	SUM225	0.0952381
TCGA-A8-A08J_BRCA	EVSAT	0.0952381
TCGA-E9-A1NA_BRCA	EVSAT	0.0952381
TCGA-C8-A1HK_BRCA	EVSAT	0.0952381
TCGA-D8-A1XD_BRCA	MCF7	0.0952381
TCGA-EW-A2FR_BRCA	HCC38	0.0952381
TCGA-A8-A08L_BRCA	HCC1954	0.0952381
TCGA-A8-A08X_BRCA	HS578T	0.0952381
TCGA-AO-A0JD_BRCA	CAL851	0.0952381
TCGA-BH-A0B4_BRCA	HCC1428	0.0952381
TCGA-BH-A42T_BRCA	HCC1428	0.0952381
TCGA-E2-A15R_BRCA	HCC1500	0.0952381
TCGA-GM-A5PX_BRCA	BT474	0.04761905
TCGA-AO-A03M_BRCA	HCC70	0.04761905
TCGA-A2-A4RX_BRCA	HCC70	0.04761905

TCGA-A7-A0DA_BRCA	HCC70	0.04761905
TCGA-AN-A0XN_BRCA	184B5	0.04761905
TCGA-GM-A5PX_BRCA	184B5	0.04761905
TCGA-AO-A12D_BRCA	HDQP1	0.04761905
TCGA-A2-A4RX_BRCA	HDQP1	0.04761905
TCGA-D8-A1JM_BRCA	HDQP1	0.04761905
TCGA-EW-A6S9_BRCA	CAL120	0.04761905
TCGA-D8-A1XW_BRCA	CAL148	0.04761905
TCGA-AR-A251_BRCA	HCC1395	0.04761905
TCGA-E2-A1B0_BRCA	HCC1395	0.04761905
TCGA-D8-A27R_BRCA	CAL51	0.04761905
TCGA-BH-A0C0_BRCA	CAL51	0.04761905
TCGA-A1-A0SJ_BRCA	MDAMB134VI	0.04761905
TCGA-D8-A27V_BRCA	MDAMB134VI	0.04761905
TCGA-AC-A6IX_BRCA	MDAMB157	0.04761905
TCGA-E9-A244_BRCA	MDAMB157	0.04761905
TCGA-E2-A15P_BRCA	MDAMB157	0.04761905
TCGA-C8-A135_BRCA	MDAMB231	0.04761905
TCGA-A1-A0SF_BRCA	MDAMB361	0.04761905
TCGA-S3-AA11_BRCA	MDAMB361	0.04761905
TCGA-A7-A0DB_BRCA	MDAMB361	0.04761905
TCGA-E2-A15G_BRCA	MDAMB361	0.04761905
TCGA-BH-A0AU_BRCA	MDAMB415	0.04761905
TCGA-EW-A424_BRCA	MDAMB415	0.04761905
TCGA-BH-A42V_BRCA	MDAMB415	0.04761905
TCGA-E2-A15G_BRCA	MDAMB415	0.04761905
TCGA-D8-A27V_BRCA	MDAMB415	0.04761905
TCGA-BH-A0C3_BRCA	JIMT1	0.04761905
TCGA-B6-A0IG_BRCA	ZR7530	0.04761905
TCGA-A7-A0DB_BRCA	ZR7530	0.04761905
TCGA-OL-A5DA_BRCA	ZR75T	0.04761905
TCGA-A8-A091_BRCA	SUM225	0.04761905
TCGA-E9-A24A_BRCA	SUM225	0.04761905
TCGA-E9-A226_BRCA	SUM225	0.04761905
TCGA-E9-A227_BRCA	SUM225	0.04761905
TCGA-A7-A0DB_BRCA	SUM225	0.04761905
TCGA-A2-A0CK_BRCA	HCC2218	0.04761905
TCGA-A8-A09R_BRCA	MFM223	0.04761905
TCGA-BH-A0C0_BRCA	MT3	0.04761905

TCGA-AC-A6NO_BRCA	MT3	0.04761905
TCGA-A8-A06N_BRCA	AU565	0.04761905
TCGA-A7-A0DA_BRCA	HCC1187	0.04761905
TCGA-GM-A3NY_BRCA	MCF10A	0.04761905
TCGA-4H-AAAK_BRCA	MCF7	0.04761905
TCGA-S3-A6ZH_BRCA	MCF7	0.04761905
TCGA-A2-A0T1_BRCA	HCC1954	0.04761905
TCGA-D8-A1JH_BRCA	HCC1954	0.04761905
TCGA-A2-A3KD_BRCA	BT483	0.04761905
TCGA-AR-A250_BRCA	BT483	0.04761905
TCGA-C8-A275_BRCA	DU4475	0.04761905
TCGA-UU-A93S_BRCA	DU4475	0.04761905
TCGA-A2-A1G0_BRCA	EFM19	0.04761905
TCGA-E9-A1RE_BRCA	EFM19	0.04761905
TCGA-4H-AAAK_BRCA	EFM19	0.04761905
TCGA-AC-A6NO_BRCA	EFM192A	0.04761905
TCGA-E2-A15O_BRCA	UACC893	0.04761905
TCGA-E2-A14Q_BRCA	CAMA1	0.04761905
TCGA-AR-A5QN_BRCA	CAMA1	0.04761905
TCGA-A2-A0T2_BRCA	HCC1569	0.04761905
TCGA-AO-A03L_BRCA	HCC1428	0.04761905
TCGA-D8-A1XC_BRCA	T47D	0.04761905
TCGA-GM-A5PX_BRCA	184A1	0.04761905
TCGA-A8-A091_BRCA	UACC812	0.04761905
TCGA-A7-A0DB_BRCA	UACC812	0.04761905
TCGA-E9-A1R3_BRCA	UACC812	0.04761905
TCGA-A7-A0CJ_BRCA	SUM190	0.04761905
TCGA-E2-A15T_BRCA	HCC1500	0.04761905

Table S.5.1: Table showing all non-zero Network-Based Matching Scores (NBMS) for patient-cell line pairs for breast cancer patient tumors and cell line samples. Scores were formed using 150 logarithmically spaced scales, the same as in the main body of the text.

Hierarchical Clustering – Correlation Distance

Patient Tumor	Cell Line	Matching Score
TCGA-AC-A62X_BRCA	HCC1187	1
TCGA-E2-A107_BRCA	MDAMB453	0.694267516

TCGA-AO-A0JG_BRCA	HCC1500	0.614649682
TCGA-D8-A140_BRCA	EFM19	0.598726115
TCGA-E2-A14S_BRCA	HCC1500	0.522292994
TCGA-C8-A1HK_BRCA	BT474	0.49044586
TCGA-AC-A6NO_BRCA	MDAMB361	0.452229299
TCGA-A2-A0CQ_BRCA	HCC1500	0.407643312
TCGA-BH-A18N_BRCA	EFM19	0.356687898
TCGA-A2-A0CU_BRCA	EFM19	0.343949045
TCGA-BH-A0DS_BRCA	T47D	0.299363057
TCGA-C8-A12Y_BRCA	EFM19	0.286624204
TCGA-E9-A5UP_BRCA	T47D	0.27388535
TCGA-BH-A42T_BRCA	MDAMB134VI	0.270700637
TCGA-AR-A0TV_BRCA	MDAMB175VII	0.267515924
TCGA-D8-A27V_BRCA	MDAMB134VI	0.248407643
TCGA-AC-A6NO_BRCA	MDAMB134VI	0.24522293
TCGA-B6-A0IG_BRCA	BT474	0.242038217
TCGA-AR-A2LJ_BRCA	184B5	0.242038217
TCGA-B6-A0IG_BRCA	SUM225	0.213375796
TCGA-EW-A1J3_BRCA	BT474	0.171974522
TCGA-E2-A15O_BRCA	EFM19	0.136942675
TCGA-E2-A15M_BRCA	MDAMB134VI	0.108280255
TCGA-LL-A442_BRCA	MDAMB361	0.089171975
TCGA-E9-A1RE_BRCA	MDAMB361	0.085987261
TCGA-E2-A15T_BRCA	HCC1500	0.082802548
TCGA-AR-A2LH_BRCA	MDAMB468	0.076433121
TCGA-BH-A0EI_BRCA	184B5	0.066878981
TCGA-AC-A6NO_BRCA	BT474	0.054140127
TCGA-E2-A1LS_BRCA	HCC38	0.054140127
TCGA-C8-A134_BRCA	MDAMB468	0.050955414
TCGA-B6-A0IO_BRCA	DU4775	0.050955414
TCGA-E9-A243_BRCA	HCC1187	0.044585987
TCGA-E2-A15O_BRCA	MDAMB361	0.041401274
TCGA-C8-A26Y_BRCA	BT474	0.035031847
TCGA-GM-A5PX_BRCA	MDAMB134VI	0.02866242
TCGA-C8-A130_BRCA	MDAMB175VII	0.01910828
TCGA-C8-A26Y_BRCA	CAL148	0.012738854
TCGA-AR-A24Q_BRCA	HCC1395	0.012738854
TCGA-D8-A1JM_BRCA	HCC1187	0.012738854
TCGA-E2-A158_BRCA	HCC1187	0.00955414

Patient Tumor	Cell Line	Matching Score
TCGA-E2-A1L8_BRCA	T47D	0.006369427

Table S.5.2: Table showing all non-zero normalized matching scores for breast cancer patient-cell line pairs formed from aggregating across dendrogram cut points using hierarchical clustering with distance metric correlation distance. The dendrogram was cut at as many points as the total number of patient tumors and cell lines combined.

Hierarchical Clustering – Euclidean Distance

Patient Tumor	Cell Line	Matching Score
TCGA-AC-A62X_BRCA	HCC1187	1
TCGA-AN-A0XO_BRCA	MDAMB415	0.99749373
TCGA-B6-A0IG_BRCA	BT474	0.63157895
TCGA-E2-A107_BRCA	MDAMB453	0.61904762
TCGA-S3-AA14_BRCA	CAMA1	0.58145363
TCGA-A2-A0CU_BRCA	EFM19	0.49122807
TCGA-E2-A14X_BRCA	COLO824	0.43609023
TCGA-BH-A0C0_BRCA	UACC812	0.37844612
TCGA-D8-A1JM_BRCA	HCC70	0.30075188
TCGA-BH-A0BQ_BRCA	CAMA1	0.28571429
TCGA-AN-A0XO_BRCA	MDAMB175VII	0.27568922
TCGA-AO-A03L_BRCA	MDAMB415	0.27568922
TCGA-AR-A1AY_BRCA	CAL51	0.26817043
TCGA-AR-A0TQ_BRCA	SUM225	0.26315789
TCGA-AR-A0TQ_BRCA	SKBR3	0.24561404
TCGA-A8-A07W_BRCA	EFM19	0.2406015
TCGA-D8-A1XQ_BRCA	CAL51	0.22556391
TCGA-AO-A03L_BRCA	DU4775	0.21553885
TCGA-BH-A18N_BRCA	EFM19	0.21052632
TCGA-C8-A273_BRCA	UACC812	0.21052632
TCGA-A8-A09I_BRCA	BT474	0.19047619
TCGA-AC-A6NO_BRCA	MDAMB175VII	0.18045113
TCGA-D8-A140_BRCA	EFM19	0.15789474
TCGA-AC-A6NO_BRCA	UACC812	0.15789474
TCGA-LL-A5YO_BRCA	HCC70	0.15538847
TCGA-AQ-A54N_BRCA	MT3	0.14536341
TCGA-C8-A1HK_BRCA	BT474	0.13784461

TCGA-A7-A0DA_BRCA	CAL51	0.13784461
TCGA-C8-A134_BRCA	HCC1395	0.13283208
TCGA-E2-A14X_BRCA	184B5	0.13032581
TCGA-D8-A27V_BRCA	MDAMB134VI	0.12280702
TCGA-UU-A93S_BRCA	BT474	0.10526316
TCGA-A8-A08L_BRCA	HCC70	0.10526316
TCGA-AC-A6NO_BRCA	BT474	0.09774436
TCGA-E9-A1RB_BRCA	MFM223	0.09022556
TCGA-A8-A08L_BRCA	CAL148	0.08521303
TCGA-AC-A6NO_BRCA	MDAMB134VI	0.0802005
TCGA-AR-A2LJ_BRCA	184B5	0.07769424
TCGA-EW-A1OZ_BRCA	SUM225	0.0726817
TCGA-BH-A0EI_BRCA	MCF10A	0.06516291
TCGA-AC-A62X_BRCA	HCC70	0.05513784
TCGA-BH-A42T_BRCA	184B5	0.05513784
TCGA-A7-A0DA_BRCA	HCC70	0.05012531
TCGA-AC-A6NO_BRCA	184B5	0.05012531
TCGA-AC-A6NO_BRCA	CAMA1	0.03258145
TCGA-AO-A0JD_BRCA	MFM223	0.02506266
TCGA-BH-A0EI_BRCA	184B5	0.02255639
TCGA-AC-A6NO_BRCA	MDAMB361	0.01002506

Table S.5.3: Table showing all non-zero normalized matching scores for breast cancer patient-cell line pairs formed from aggregating across dendrogram cut points using hierarchical clustering with distance metric Euclidean distance. The dendrogram was cut at as many points as the total number of patient tumors and cell lines combined.

Lung Cancer

NetCellMatch

Patient Tumor	Cell Line	NBMS
TCGA-21-5782_LUSC	HCC1359	1
TCGA-35-4123_LUAD	H1869	1
TCGA-78-7540_LUAD	DFCI024	0.875
TCGA-97-7554_LUAD	H2073	0.875
TCGA-51-6867_LUSC	HCC95	0.875
TCGA-55-8620_LUAD	H1385	0.75
TCGA-MP-A4SY_LUAD	H838	0.75
TCGA-44-3917_LUAD	H322M	0.75
TCGA-56-A4BX_LUSC	H661	0.75
TCGA-78-7153_LUAD	HOP62	0.625
TCGA-O2-A52Q_LUSC	HCC366	0.5
TCGA-86-6851_LUAD	HCC366	0.5
TCGA-4B-A93V_LUAD	A549	0.5
TCGA-86-7701_LUAD	H727	0.5
TCGA-NJ-A55A_LUAD	H920	0.5
TCGA-34-8455_LUSC	HCC364	0.5
TCGA-64-1680_LUAD	HCC4019	0.5
TCGA-50-5072_LUAD	H1623	0.5
TCGA-77-7335_LUSC	H2087	0.5
TCGA-MP-A4T7_LUAD	H1573	0.5
TCGA-78-7149_LUAD	H2342	0.5
TCGA-44-A47G_LUAD	H1355	0.375
TCGA-55-6971_LUAD	H1385	0.375
TCGA-78-7155_LUAD	A549	0.375
TCGA-78-7149_LUAD	H2122	0.375
TCGA-78-7155_LUAD	H378	0.375
TCGA-78-7162_LUAD	HCC1359	0.375
TCGA-78-7155_LUAD	DMS79	0.375
TCGA-44-7659_LUAD	HCC15	0.375
TCGA-50-5941_LUAD	HCC364	0.375
TCGA-49-4494_LUAD	HCC4011	0.375
TCGA-69-7760_LUAD	H1373	0.375
TCGA-50-5930_LUAD	H1792	0.375

TCGA-L9-A50W_LUAD	H2110	0.375
TCGA-55-6971_LUAD	H358	0.375
TCGA-MP-A4T4_LUAD	HCC1195	0.375
TCGA-98-A53J_LUSC	H324	0.375
TCGA-66-2770_LUSC	HCC3051	0.375
TCGA-56-7221_LUSC	HCC1833	0.375
TCGA-NC-A5HF_LUSC	H1648	0.375
TCGA-73-4670_LUAD	H2342	0.375
TCGA-43-8115_LUSC	H661	0.375
TCGA-49-6767_LUAD	H1703	0.375
TCGA-NJ-A55R_LUAD	HOP92	0.25
TCGA-35-4123_LUAD	HCC366	0.25
TCGA-55-6971_LUAD	HCC366	0.25
TCGA-98-7454_LUSC	H1355	0.25
TCGA-56-5898_LUSC	H1385	0.25
TCGA-78-7154_LUAD	A549	0.25
TCGA-39-5028_LUSC	H2122	0.25
TCGA-49-4507_LUAD	H2122	0.25
TCGA-94-8035_LUSC	H378	0.25
TCGA-44-7660_LUAD	H378	0.25
TCGA-97-A4M5_LUAD	H727	0.25
TCGA-44-6147_LUAD	PC9	0.25
TCGA-97-7554_LUAD	DFCI024	0.25
TCGA-49-4501_LUAD	DFCI032	0.25
TCGA-44-6147_LUAD	DFCI032	0.25
TCGA-39-5016_LUSC	HCC1438	0.25
TCGA-85-A5B5_LUSC	HCC4017	0.25
TCGA-37-4129_LUSC	HCC4018	0.25
TCGA-NC-A5HR_LUSC	SHP77	0.25
TCGA-56-7730_LUSC	H1838	0.25
TCGA-63-A5MN_LUSC	H1869	0.25
TCGA-94-7943_LUSC	H1993	0.25
TCGA-77-6842_LUSC	H2110	0.25
TCGA-66-2769_LUSC	H125	0.25
TCGA-80-5608_LUAD	H290	0.25
TCGA-44-7660_LUAD	H720	0.25
TCGA-85-A53L_LUSC	H810	0.25
TCGA-95-7039_LUAD	HCC1313	0.25
TCGA-60-2696_LUSC	HCC1313	0.25

TCGA-78-7154_LUAD	HCC1313	0.25
TCGA-55-8506_LUAD	HCC2108	0.25
TCGA-77-6844_LUSC	HCC2814	0.25
TCGA-50-5072_LUAD	HCC78	0.25
TCGA-78-7161_LUAD	H1435	0.25
TCGA-NC-A5HI_LUSC	H1437	0.25
TCGA-97-7553_LUAD	SKLU1	0.25
TCGA-44-A4SS_LUAD	H1666	0.25
TCGA-44-6147_LUAD	H1666	0.25
TCGA-44-6147_LUAD	H2286	0.25
TCGA-NJ-A55R_LUAD	H820	0.25
TCGA-78-7154_LUAD	H1650	0.25
TCGA-95-7039_LUAD	H441	0.25
TCGA-58-A46J_LUSC	H1581	0.25
TCGA-63-5128_LUSC	H1648	0.25
TCGA-55-8507_LUAD	H1648	0.25
TCGA-43-6143_LUSC	H2085	0.25
TCGA-55-1594_LUAD	H2085	0.25
TCGA-90-7769_LUSC	H2882	0.25
TCGA-95-7567_LUAD	H1781	0.25
TCGA-97-7938_LUAD	H2172	0.25
TCGA-56-5897_LUSC	H2172	0.25
TCGA-55-8506_LUAD	HOP62	0.125
TCGA-35-4123_LUAD	HOP62	0.125
TCGA-NJ-A55R_LUAD	HOP62	0.125
TCGA-55-6971_LUAD	HOP62	0.125
TCGA-44-6147_LUAD	HOP92	0.125
TCGA-49-6742_LUAD	H1355	0.125
TCGA-49-4494_LUAD	H1355	0.125
TCGA-44-6147_LUAD	H1355	0.125
TCGA-56-A5DS_LUSC	H1385	0.125
TCGA-22-4591_LUSC	H1385	0.125
TCGA-50-5930_LUAD	A549	0.125
TCGA-44-2657_LUAD	H1944	0.125
TCGA-97-A4M5_LUAD	H1944	0.125
TCGA-77-6845_LUSC	SKMES1	0.125
TCGA-69-7764_LUAD	H727	0.125
TCGA-80-5608_LUAD	H920	0.125
TCGA-22-5482_LUSC	H920	0.125

TCGA-MP-A4T6_LUAD	H969	0.125
TCGA-86-7954_LUAD	H969	0.125
TCGA-78-7154_LUAD	H969	0.125
TCGA-73-4658_LUAD	DFCI024	0.125
TCGA-44-A47F_LUAD	DFCI032	0.125
TCGA-44-6777_LUAD	DFCI032	0.125
TCGA-98-A539_LUSC	HCC1359	0.125
TCGA-55-6971_LUAD	DMS79	0.125
TCGA-85-6175_LUSC	HCC1438	0.125
TCGA-86-8055_LUAD	HCC2279	0.125
TCGA-64-1676_LUAD	HCC2279	0.125
TCGA-44-6777_LUAD	HCC2279	0.125
TCGA-56-7730_LUSC	HCC364	0.125
TCGA-49-AARR_LUAD	HCC4011	0.125
TCGA-78-7153_LUAD	HCC4017	0.125
TCGA-90-7766_LUSC	HCC4018	0.125
TCGA-39-5031_LUSC	HCC4018	0.125
TCGA-58-A46L_LUSC	HCC4018	0.125
TCGA-66-2794_LUSC	HCC4018	0.125
TCGA-MP-A4TK_LUAD	H1373	0.125
TCGA-63-A5MY_LUSC	SHP77	0.125
TCGA-44-3398_LUAD	H1693	0.125
TCGA-NJ-A55R_LUAD	H1693	0.125
TCGA-60-2696_LUSC	H1792	0.125
TCGA-67-3771_LUAD	H1792	0.125
TCGA-44-7670_LUAD	H1792	0.125
TCGA-85-8048_LUSC	H1793	0.125
TCGA-56-8309_LUSC	H1838	0.125
TCGA-55-6971_LUAD	H1838	0.125
TCGA-86-6851_LUAD	H1869	0.125
TCGA-50-5066_LUAD	H596	0.125
TCGA-05-5423_LUAD	H1975	0.125
TCGA-93-7348_LUAD	H2073	0.125
TCGA-MP-A4SV_LUAD	H2110	0.125
TCGA-21-1076_LUSC	H125	0.125
TCGA-MP-A4TC_LUAD	H2228	0.125
TCGA-63-A5MU_LUSC	CALU3	0.125
TCGA-44-6775_LUAD	H290	0.125
TCGA-60-2725_LUSC	H460	0.125

TCGA-4B-A93V_LUAD	H460	0.125
TCGA-86-8585_LUAD	SW1573	0.125
TCGA-60-2696_LUSC	SW1573	0.125
TCGA-22-5471_LUSC	SW1573	0.125
TCGA-22-0940_LUSC	HCC2374	0.125
TCGA-NJ-A4YF_LUAD	HCC2374	0.125
TCGA-39-5027_LUSC	H810	0.125
TCGA-55-6971_LUAD	H810	0.125
TCGA-97-A4M5_LUAD	H838	0.125
TCGA-05-4418_LUAD	HCC1313	0.125
TCGA-50-5930_LUAD	HCC1313	0.125
TCGA-62-A46O_LUAD	HCC1195	0.125
TCGA-33-6738_LUSC	HCC1195	0.125
TCGA-05-4427_LUAD	HCC1195	0.125
TCGA-55-6969_LUAD	HCC2108	0.125
TCGA-35-4122_LUAD	HCC2108	0.125
TCGA-MP-A4T4_LUAD	HCC2108	0.125
TCGA-46-6026_LUSC	H522	0.125
TCGA-95-7567_LUAD	HCC2429	0.125
TCGA-44-7660_LUAD	HCC2814	0.125
TCGA-77-6844_LUSC	H324	0.125
TCGA-63-7022_LUSC	HCC78	0.125
TCGA-86-6851_LUAD	HCC78	0.125
TCGA-50-6597_LUAD	H1395	0.125
TCGA-NJ-A55O_LUAD	H1435	0.125
TCGA-62-A46O_LUAD	H1437	0.125
TCGA-05-4427_LUAD	H1437	0.125
TCGA-66-2794_LUSC	H1437	0.125
TCGA-49-AARE_LUAD	SKLU1	0.125
TCGA-91-6848_LUAD	SKLU1	0.125
TCGA-05-4410_LUAD	HCC3051	0.125
TCGA-64-5779_LUAD	A427	0.125
TCGA-85-8664_LUSC	H1770	0.125
TCGA-75-6207_LUAD	H1819	0.125
TCGA-58-A46J_LUSC	H1819	0.125
TCGA-73-4675_LUAD	H2126	0.125
TCGA-55-6712_LUAD	H2023	0.125
TCGA-MP-A4TK_LUAD	H2023	0.125
TCGA-50-5072_LUAD	H2250	0.125

TCGA-05-4418_LUAD	H2250	0.125
TCGA-NJ-A7XG_LUAD	H2258	0.125
TCGA-75-7031_LUAD	H2087	0.125
TCGA-55-8620_LUAD	H211	0.125
TCGA-55-1594_LUAD	H2170	0.125
TCGA-60-2709_LUSC	H292	0.125
TCGA-58-A46L_LUSC	H292	0.125
TCGA-MP-A4T4_LUAD	H292	0.125
TCGA-62-8398_LUAD	H2195	0.125
TCGA-58-A46L_LUSC	H2195	0.125
TCGA-44-6777_LUAD	H820	0.125
TCGA-38-A44F_LUAD	H2347	0.125
TCGA-22-5480_LUSC	H1650	0.125
TCGA-39-5029_LUSC	H1650	0.125
TCGA-34-8455_LUSC	H441	0.125
TCGA-77-8153_LUSC	H2085	0.125
TCGA-73-7498_LUAD	H1573	0.125
TCGA-L4-A4E6_LUAD	H1573	0.125
TCGA-MN-A4N4_LUAD	H1573	0.125
TCGA-91-8497_LUAD	HCC827	0.125
TCGA-MN-A4N4_LUAD	HCC827	0.125
TCGA-62-A46P_LUAD	H1781	0.125
TCGA-90-7766_LUSC	H1781	0.125
TCGA-56-A5DS_LUSC	H1781	0.125
TCGA-58-A46K_LUSC	HCC95	0.125
TCGA-21-1078_LUSC	H1734	0.125
TCGA-86-6851_LUAD	H1734	0.125
TCGA-50-5072_LUAD	H1568	0.125
TCGA-78-7154_LUAD	H1568	0.125
TCGA-85-7844_LUSC	H2172	0.125
TCGA-44-6777_LUAD	H1651	0.125
TCGA-05-4430_LUAD	H1703	0.125
TCGA-J2-8192_LUAD	H1563	0.125
TCGA-44-A47G_LUAD	H2135	0.125

Table S.5.4: Table showing all non-zero Network-Based Matching Scores (NBMS) for patient-cell line pairs for lung cancer patient tumors and cell line samples. Scores were formed using 150 logarithmically spaced scales, the same as in the main body of the text.

Hierarchical Clustering – Correlation Distance

Patient Tumor	Cell Line	Matching Scores
TCGA-85-6175_LUSC	H596	1
TCGA-55-1594_LUAD	H2073	0.952380952
TCGA-95-7567_LUAD	H324	0.952380952
TCGA-75-6207_LUAD	HCC364	0.704761905
TCGA-85-A4QQ_LUSC	H125	0.666666667
TCGA-MN-A4N4_LUAD	H920	0.533333333
TCGA-NJ-A4YF_LUAD	H2122	0.507936508
TCGA-44-6147_LUAD	H920	0.507936508
TCGA-62-8394_LUAD	H1819	0.488888889
TCGA-05-5715_LUAD	HCC1438	0.473015873
TCGA-44-7669_LUAD	SHP77	0.466666667
TCGA-68-7757_LUSC	HOP92	0.450793651
TCGA-60-2716_LUSC	H322M	0.441269841
TCGA-85-6560_LUSC	HCC15	0.425396825
TCGA-55-5899_LUAD	H460	0.412698413
TCGA-77-6844_LUSC	HCC2814	0.40952381
TCGA-95-A4VN_LUAD	HCC78	0.406349206
TCGA-75-7027_LUAD	H1437	0.396825397
TCGA-39-5028_LUSC	HCC2814	0.39047619
TCGA-MN-A4N4_LUAD	HCC193	0.371428571
TCGA-50-5044_LUAD	HCC1438	0.368253968
TCGA-O2-A52N_LUSC	HCC827	0.361904762
TCGA-49-4494_LUAD	H920	0.349206349
TCGA-21-1082_LUSC	EKVX	0.346031746
TCGA-MP-A4T4_LUAD	H1734	0.320634921
TCGA-38-4627_LUAD	H1563	0.311111111
TCGA-69-A59K_LUAD	H2170	0.285714286
TCGA-95-7567_LUAD	H1693	0.279365079
TCGA-38-4625_LUAD	H1793	0.263492063
TCGA-69-7760_LUAD	HCC364	0.244444444
TCGA-NJ-A55R_LUAD	HCC2814	0.244444444
TCGA-55-1594_LUAD	HCC827	0.244444444

TCGA-49-4494_LUAD	HCC2279	0.238095238
TCGA-75-7027_LUAD	H460	0.171428571
TCGA-37-A5EL_LUSC	H23	0.146031746
TCGA-44-6145_LUAD	H1944	0.136507937
TCGA-78-7540_LUAD	PC9	0.123809524
TCGA-4B-A93V_LUAD	SHP77	0.120634921
TCGA-56-5897_LUSC	HCC2450	0.117460317
TCGA-38-4627_LUAD	HOP62	0.111111111
TCGA-MP-A4T4_LUAD	HCC364	0.111111111
TCGA-60-2696_LUSC	HOP92	0.107936508
TCGA-94-7033_LUSC	HCC2814	0.104761905
TCGA-55-6971_LUAD	H1385	0.098412698
TCGA-44-6147_LUAD	H1944	0.095238095
TCGA-NJ-A55R_LUAD	H23	0.092063492
TCGA-94-7557_LUSC	HCC78	0.082539683
TCGA-85-7698_LUSC	EKVX	0.079365079
TCGA-55-6971_LUAD	HOP62	0.076190476
TCGA-37-4129_LUSC	H1385	0.073015873
TCGA-4B-A93V_LUAD	H1385	0.06984127
TCGA-49-AAQV_LUAD	H1734	0.066666667
TCGA-44-6147_LUAD	HOP62	0.053968254
TCGA-85-7698_LUSC	H23	0.053968254
TCGA-78-7155_LUAD	H322M	0.038095238
TCGA-78-7155_LUAD	H2110	0.034920635
TCGA-NJ-A55R_LUAD	H2073	0.031746032
TCGA-38-4627_LUAD	HOP92	0.025396825
TCGA-21-1078_LUSC	H727	0.025396825
TCGA-78-7153_LUAD	H1993	0.015873016
TCGA-35-4123_LUAD	H322M	0.012698413
TCGA-62-8398_LUAD	H1395	0.006349206

Table S.5.5: Table showing all non-zero normalized matching scores for lung cancer patient-cell line pairs formed from aggregating across dendrogram cut points using hierarchical clustering with distance metric correlation distance. The dendrogram was cut at as many points as the total number of patient tumors and cell lines combined.

Hierarchical Clustering – Euclidean Distance

Patient Tumor	Cell Line	Matching Scores
TCGA-95-7567_LUAD	H324	1
TCGA-75-6207_LUAD	HCC364	0.689655172
TCGA-85-A4QQ_LUSC	H125	0.655172414
TCGA-86-8358_LUAD	SHP77	0.468390805
TCGA-44-7660_LUAD	H460	0.462643678
TCGA-94-8491_LUSC	H322M	0.442528736
TCGA-55-1594_LUAD	H1573	0.425287356
TCGA-44-6147_LUAD	HCC366	0.393678161
TCGA-85-8049_LUSC	H23	0.370689655
TCGA-75-7031_LUAD	H1781	0.359195402
TCGA-44-7660_LUAD	H1355	0.344827586
TCGA-85-6560_LUSC	HCC15	0.33045977
TCGA-21-1082_LUSC	H596	0.33045977
TCGA-18-5592_LUSC	H2073	0.275862069
TCGA-49-4507_LUAD	H1650	0.267241379
TCGA-55-6971_LUAD	HOP62	0.224137931
TCGA-85-8049_LUSC	H125	0.224137931
TCGA-44-8119_LUAD	H727	0.221264368
TCGA-95-7948_LUAD	H324	0.215517241
TCGA-85-6561_LUSC	H460	0.212643678
TCGA-56-8309_LUSC	HCC2814	0.189655172
TCGA-05-4397_LUAD	HCC2429	0.183908046
TCGA-50-5044_LUAD	HOP62	0.181034483
TCGA-56-5897_LUSC	H1355	0.172413793
TCGA-98-7454_LUSC	HCC366	0.163793103
TCGA-O2-A52N_LUSC	H2073	0.16091954
TCGA-97-A4M5_LUAD	HCC366	0.158045977
TCGA-21-1082_LUSC	HCC366	0.155172414
TCGA-62-8398_LUAD	HCC1195	0.155172414

TCGA-78-7143_LUAD	H1734	0.152298851
TCGA-4B-A93V_LUAD	SHP77	0.143678161
TCGA-44-7660_LUAD	H125	0.140804598
TCGA-95-7039_LUAD	HCC364	0.137931034
TCGA-69-A59K_LUAD	H1819	0.126436782
TCGA-60-2714_LUSC	H1869	0.117816092
TCGA-77-8154_LUSC	H1650	0.112068966
TCGA-55-1594_LUAD	H1373	0.097701149
TCGA-L9-A5IP_LUAD	H322M	0.091954023
TCGA-55-1594_LUAD	H2073	0.086206897
TCGA-56-5897_LUSC	H838	0.086206897
TCGA-55-6971_LUAD	H727	0.083333333
TCGA-34-5231_LUSC	H2073	0.083333333
TCGA-95-7948_LUAD	H920	0.08045977
TCGA-78-7154_LUAD	H1568	0.077586207
TCGA-49-4507_LUAD	H2122	0.07183908
TCGA-L9-A5IP_LUAD	PC9	0.07183908
TCGA-56-A5DS_LUSC	HCC2814	0.07183908
TCGA-55-6971_LUAD	SHP77	0.066091954
TCGA-05-4418_LUAD	H1437	0.063218391
TCGA-44-6147_LUAD	H1734	0.060344828
TCGA-44-6147_LUAD	HOP62	0.051724138
TCGA-50-5931_LUAD	H378	0.051724138
TCGA-69-A59K_LUAD	CALU3	0.051724138
TCGA-49-4507_LUAD	H1437	0.043103448
TCGA-55-1596_LUAD	CALU3	0.040229885
TCGA-44-7660_LUAD	HOP62	0.031609195
TCGA-55-6971_LUAD	H1385	0.028735632
TCGA-50-5044_LUAD	HCC1438	0.025862069
TCGA-69-7760_LUAD	HCC1195	0.025862069
TCGA-35-4123_LUAD	H1568	0.020114943
TCGA-66-2794_LUSC	H1568	0.020114943
TCGA-39-5019_LUSC	H460	0.014367816

Table S.5.6: Table showing all non-zero normalized matching scores for lung cancer patient-cell line pairs formed from aggregating across dendrogram cut points using hierarchical clustering with distance metric Euclidean distance. The dendrogram was cut at as many points as the total number of patient tumors and cell lines combined.

Colorectal Cancer

NetCellMatch

Patient Tumor	Cell Line	NBMS
TCGA-DM-A1DA_COAD	SW480	1
TCGA-AZ-4308_COAD	SW1417	0.94444444
TCGA-QL-A97D_COAD	COLO205	0.77777778
TCGA-G4-6303_COAD	SW480	0.77777778
TCGA-AZ-4684_COAD	SW837	0.66666667
TCGA-G4-6321_COAD	SW48	0.66666667
TCGA-AA-3986_COAD	LS174T	0.55555556
TCGA-AA-A029_COAD	COLO320	0.5
TCGA-AA-A01Q_COAD	H716	0.5
TCGA-AA-A02J_COAD	HT29	0.44444444
TCGA-AA-A029_COAD	HT29	0.44444444
TCGA-D5-6931_COAD	SW480	0.44444444
TCGA-CM-6166_COAD	H747	0.38888889
TCGA-AA-A01V_COAD	CX1	0.38888889
TCGA-AD-6889_COAD	HCT8	0.38888889
TCGA-G4-6303_COAD	SW837	0.38888889
TCGA-AA-3688_COAD	LOVO	0.33333333
TCGA-DM-A1D0_COAD	SW1417	0.27777778
TCGA-DM-A1D0_COAD	H716	0.27777778
TCGA-G4-6297_COAD	LOVO	0.27777778
TCGA-G4-6297_COAD	SNUC2B	0.22222222
TCGA-AA-3819_COAD	CL14	0.22222222
TCGA-AA-A022_COAD	COLO201	0.22222222
TCGA-DM-A1D0_COAD	KM12	0.22222222
TCGA-AA-3553_COAD	MDST8	0.22222222
TCGA-AZ-6599_COAD	SNUC1	0.22222222
TCGA-AA-A02H_COAD	HCC2998	0.22222222
TCGA-AA-3519_COAD	LOVO	0.22222222

TCGA-DM-A1D8_COAD	SW620	0.22222222
TCGA-AA-3553_COAD	SNUC2B	0.16666667
TCGA-A6-2686_COAD	SW1417	0.16666667
TCGA-AA-3548_COAD	CL40	0.16666667
TCGA-AZ-6599_COAD	CL40	0.16666667
TCGA-SS-A7HO_COAD	COLO205	0.16666667
TCGA-F4-6854_COAD	KM12	0.16666667
TCGA-AM-5820_COAD	KM12	0.16666667
TCGA-CK-6746_COAD	LS411N	0.16666667
TCGA-AA-3552_COAD	MDST8	0.16666667
TCGA-CM-6679_COAD	C2BBE1	0.16666667
TCGA-AA-3814_COAD	C2BBE1	0.16666667
TCGA-AA-3814_COAD	HT29	0.11111111
TCGA-D5-6537_COAD	HT55	0.11111111
TCGA-CA-5797_COAD	SW1417	0.11111111
TCGA-AA-A00N_COAD	H716	0.11111111
TCGA-F4-6856_COAD	LS123	0.11111111
TCGA-DM-A1D8_COAD	SW837	0.11111111
TCGA-A6-5657_COAD	LS174T	0.11111111
TCGA-A6-2684_COAD	LS174T	0.11111111
TCGA-AF-6784_COAD	LS174T	0.11111111
TCGA-AZ-4323_COAD	SW48	0.11111111
TCGA-A6-6653_COAD	SW48	0.11111111
TCGA-AA-3548_COAD	HT29	0.05555556
TCGA-AZ-6599_COAD	HT29	0.05555556
TCGA-CM-4746_COAD	DLD1	0.05555556
TCGA-DM-A1D8_COAD	DLD1	0.05555556
TCGA-DM-A1DA_COAD	DLD1	0.05555556
TCGA-AA-3975_COAD	SW1417	0.05555556
TCGA-CA-6717_COAD	SW1417	0.05555556
TCGA-G4-6309_COAD	CL34	0.05555556
TCGA-AA-3973_COAD	CL40	0.05555556
TCGA-AA-3814_COAD	COLO201	0.05555556
TCGA-DM-A1D8_COAD	H747	0.05555556
TCGA-AD-6889_COAD	H747	0.05555556
TCGA-AA-A010_COAD	CX1	0.05555556
TCGA-AD-6889_COAD	CX1	0.05555556
TCGA-CM-6680_COAD	LS411N	0.05555556
TCGA-DM-A280_COAD	MDST8	0.05555556

Patient Tumor	Cell Line	Matching Score
TCGA-DM-A1D6_COAD	SW480	0.05555556
TCGA-CM-6674_COAD	H716	0.05555556
TCGA-AA-3517_COAD	GP5D	0.05555556
TCGA-AA-3511_COAD	GP5D	0.05555556
TCGA-AA-A022_COAD	C2BBE1	0.05555556
TCGA-CK-6747_COAD	C2BBE1	0.05555556
TCGA-DM-A28A_COAD	SW837	0.05555556
TCGA-CM-6679_COAD	HT115	0.05555556

Table S.5.7: Table showing all non-zero Network-Based Matching Scores (NBMS) for patient-cell line pairs for colorectal cancer patient tumors and cell line samples. Scores were formed using 150 logarithmically spaced scales, the same as in the main body of the text.

Hierarchical Clustering – Correlation Distance

Patient Tumor	Cell Line	Matching Score
TCGA-AD-6889_COAD	LS411N	1
TCGA-AA-A029_COAD	HT29	0.550561798
TCGA-AD-6889_COAD	DLD1	0.348314607
TCGA-G4-6297_COAD	HT29	0.213483146
TCGA-AM-5820_COAD	HCT116	0.202247191
TCGA-AA-A01V_COAD	HT29	0.191011236
TCGA-AA-A029_COAD	COLO205	0.157303371
TCGA-A6-6140_COAD	COLO205	0.134831461
TCGA-DM-A1D8_COAD	SW837	0.134831461
TCGA-G4-6299_COAD	H747	0.101123596
TCGA-DM-A1D0_COAD	MDST8	0.101123596
TCGA-AA-A029_COAD	CACO2	0.08988764
TCGA-CA-6717_COAD	HCT116	0.078651685
TCGA-AA-3982_COAD	LOVO	0.056179775
TCGA-A6-6653_COAD	LS174T	0.04494382
TCGA-A6-6653_COAD	H716	0.033707865
TCGA-DM-A1D0_COAD	CACO2	0.02247191

Table S.5.8: Table showing all non-zero normalized matching scores for colorectal cancer patient-cell line pairs formed from aggregating across dendrogram cut points using hierarchical clustering with distance metric correlation distance. The

Hierarchical Clustering – Euclidean Distance

Patient Tumor	Cell Line	Matching Score
TCGA-AA-A029_COAD	HT29	1
TCGA-AD-6889_COAD	LS411N	0.966101695
TCGA-AF-6784_COAD	SW837	0.745762712
TCGA-CA-6715_COAD	SW837	0.457627119
TCGA-AA-A029_COAD	CACO2	0.415254237
TCGA-AA-A01V_COAD	HT29	0.313559322
TCGA-G4-6298_COAD	DLD1	0.194915254
TCGA-DM-A1DA_COAD	DLD1	0.186440678
TCGA-DM-A1D0_COAD	CACO2	0.152542373
TCGA-AZ-6607_COAD	H747	0.152542373
TCGA-AF-6784_COAD	CACO2	0.118644068
TCGA-G4-6298_COAD	SW1417	0.050847458
TCGA-DM-A1D0_COAD	SW48	0.016949153

Table S.5.9: Table showing all non-zero normalized matching scores for colorectal cancer patient-cell line pairs formed from aggregating across dendrogram cut points using hierarchical clustering with distance metric Euclidean distance. The dendrogram was cut at as many points as the total number of patient tumors and cell lines combined.

We note that NBMS scores from NetCellMatch produce cell line rankings that are noticeably different than those produced from both aggregated hierarchical clustering approaches, which tend to show more similarities. This highlights the potential ability of NetCellMatch to produce unique insights by aggregating clustering results along a topologically meaningful path. It should be noted,

however, that some cell lines are prominently scored across all three clustering approaches (e.g. → Breast: BT474).