

Supplementary Material

SigMat: A Classification Scheme for Gene Signature Matching

Jinfeng Xiao¹, Charles Blatti¹ and Saurabh Sinha^{1,2,*}

¹Department of Computer Science, University of Illinois at Urbana-Champaign

²Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign

*To whom correspondence should be addressed.

Supplementary Table 1. Number of drugs shared between pairs of cell lines, so that each cell line has at least 10 experiments on each drug.

	A375	A549	HA1E	HT29	MCF7	PC3	VCAP	HCC515	HEPG2
A375	654	323	442	572	632	619	305	443	131
A549		546	270	413	546	530	384	275	131
HA1E			527	434	453	449	173	402	127
HT29				781	774	748	450	435	133
MCF7					2,278	1,456	969	460	153
PC3						3,440	1,985	447	152
VCAP							3,370	152	50
HCC515								466	129
HEPG2									244

Supplementary Table 2. Number of drugs shared between the training cell lines (columns) and the testing cell lines (rows) in Section 3.6. Each training cell line contains the top 500 drugs with the greatest number of LINCS L1000 experiments on that cell line.

	A375	A549	HA1E	HT29	MCF7	PC3	VCAP
HCC515	400	272	381	288	220	210	88
HEPG2	129	131	125	129	95	97	17

Success@K, K =		1					5					10								
train	test	tune%	SigMat	cor	RF	SVM	KS	random	SigMat	cor	RF	SVM	KS	random	SigMat	cor	RF	SVM	KS	random
A375	HCC515	0.05	0.084	0.084	0.082	0.034	0.003	0.002	0.241	0.186	0.197	0.129	0.014	0.010	0.322	0.247	0.264	0.159	0.031	0.020
	HEPG2	0.5	0.104	0.076	0.078	0.032	0.004	0.002	0.244	0.175	0.191	0.117	0.013	0.010	0.309	0.241	0.258	0.143	0.030	0.020
A549	HCC515	0.05	0.052	0.049	0.039	0.036	0.004	0.002	0.160	0.148	0.105	0.130	0.022	0.010	0.212	0.192	0.153	0.183	0.037	0.020
	HEPG2	0.5	0.053	0.053	0.039	0.036	0.002	0.002	0.170	0.150	0.111	0.123	0.012	0.010	0.239	0.202	0.159	0.170	0.025	0.020
HA1E	HCC515	0.05	0.124	0.120	0.097	0.055	0.003	0.002	0.244	0.234	0.195	0.125	0.020	0.010	0.291	0.291	0.248	0.163	0.038	0.020
	HEPG2	0.5	0.129	0.105	0.079	0.045	0.002	0.002	0.252	0.212	0.166	0.108	0.019	0.010	0.315	0.260	0.207	0.149	0.037	0.020
HT29	HCC515	0.05	0.073	0.058	0.036	0.030	0.002	0.002	0.178	0.142	0.102	0.119	0.020	0.010	0.245	0.202	0.156	0.164	0.040	0.020
	HEPG2	0.5	0.088	0.069	0.038	0.026	0.000	0.002	0.206	0.149	0.103	0.109	0.011	0.010	0.282	0.202	0.154	0.158	0.034	0.020
MCF7	HCC515	0.05	0.149	0.124	0.091	0.044	0.002	0.002	0.295	0.237	0.225	0.116	0.016	0.010	0.365	0.307	0.293	0.156	0.031	0.020
	HEPG2	0.5	0.149	0.127	0.095	0.041	0.003	0.002	0.301	0.244	0.237	0.114	0.017	0.010	0.374	0.313	0.309	0.155	0.035	0.020
PC3	HCC515	0.05	0.053	0.057	0.025	0.025	0.003	0.002	0.135	0.126	0.091	0.094	0.016	0.010	0.185	0.168	0.145	0.142	0.031	0.020
	HEPG2	0.5	0.081	0.074	0.032	0.035	0.005	0.002	0.145	0.145	0.116	0.126	0.016	0.010	0.211	0.176	0.174	0.182	0.037	0.020
VCAP	HCC515	0.05	0.116	0.089	0.071	0.031	0.003	0.002	0.233	0.181	0.162	0.101	0.019	0.010	0.294	0.236	0.224	0.143	0.039	0.020
	HEPG2	0.5	0.088	0.065	0.048	0.029	0.001	0.002	0.192	0.140	0.137	0.091	0.017	0.010	0.254	0.186	0.201	0.130	0.031	0.020
MCF7	HCC515	0.05	0.056	0.057	0.040	0.021	0.007	0.002	0.146	0.161	0.114	0.109	0.020	0.010	0.198	0.220	0.157	0.165	0.031	0.020
	HEPG2	0.5	0.045	0.044	0.039	0.025	0.003	0.002	0.172	0.153	0.114	0.109	0.014	0.010	0.222	0.203	0.152	0.162	0.025	0.020
PC3	HCC515	0.05	0.171	0.135	0.112	0.059	0.001	0.002	0.265	0.241	0.228	0.133	0.018	0.010	0.307	0.289	0.283	0.173	0.040	0.020
	HEPG2	0.5	0.150	0.115	0.087	0.051	0.003	0.002	0.268	0.212	0.204	0.117	0.023	0.010	0.337	0.273	0.264	0.157	0.040	0.020
MCF7	HCC515	0.05	0.043	0.036	0.026	0.021	0.002	0.002	0.111	0.111	0.090	0.060	0.014	0.010	0.169	0.170	0.127	0.083	0.023	0.020
	HEPG2	0.5	0.057	0.043	0.028	0.030	0.002	0.002	0.126	0.115	0.096	0.066	0.015	0.010	0.191	0.194	0.130	0.091	0.021	0.020
PC3	HCC515	0.05	0.094	0.070	0.052	0.025	0.002	0.002	0.202	0.159	0.138	0.086	0.017	0.010	0.248	0.204	0.175	0.117	0.030	0.020
	HEPG2	0.5	0.088	0.063	0.049	0.023	0.003	0.002	0.197	0.137	0.132	0.079	0.016	0.010	0.246	0.184	0.174	0.108	0.030	0.020
VCAP	HCC515	0.05	0.057	0.046	0.039	0.016	0.003	0.002	0.147	0.125	0.110	0.095	0.011	0.010	0.199	0.170	0.161	0.120	0.024	0.020
	HEPG2	0.5	0.050	0.033	0.035	0.015	0.002	0.002	0.129	0.100	0.098	0.090	0.008	0.010	0.183	0.142	0.152	0.112	0.027	0.020
VCAP	HCC515	0.05	0.043	0.060	0.065	0.022	0.004	0.002	0.182	0.182	0.154	0.099	0.016	0.010	0.124	0.228	0.208	0.157	0.030	0.020
	HEPG2	0.5	0.102	0.052	0.059	0.030	0.007	0.002	0.230	0.189	0.168	0.105	0.016	0.010	0.302	0.223	0.225	0.155	0.027	0.020
MCF7	HCC515	0.05	0.062	0.144	0.088	0.006	0.000	0.002	0.238	0.238	0.231	0.119	0.019	0.010	0.288	0.288	0.312	0.206	0.038	0.020
	HEPG2	0.5	0.138	0.150	0.125	0.012	0.000	0.002	0.312	0.238	0.275	0.138	0.000	0.010	0.425	0.288	0.412	0.275	0.025	0.020

Supplementary Table 3. Success@K (K = 1, 5 or 10) for SigMat and other methods under 28 combinations of 7 training cell lines, 2 testing cell lines, and 2 different percentages of testing drugs allowed for SigMat tuning. Best viewed in color.