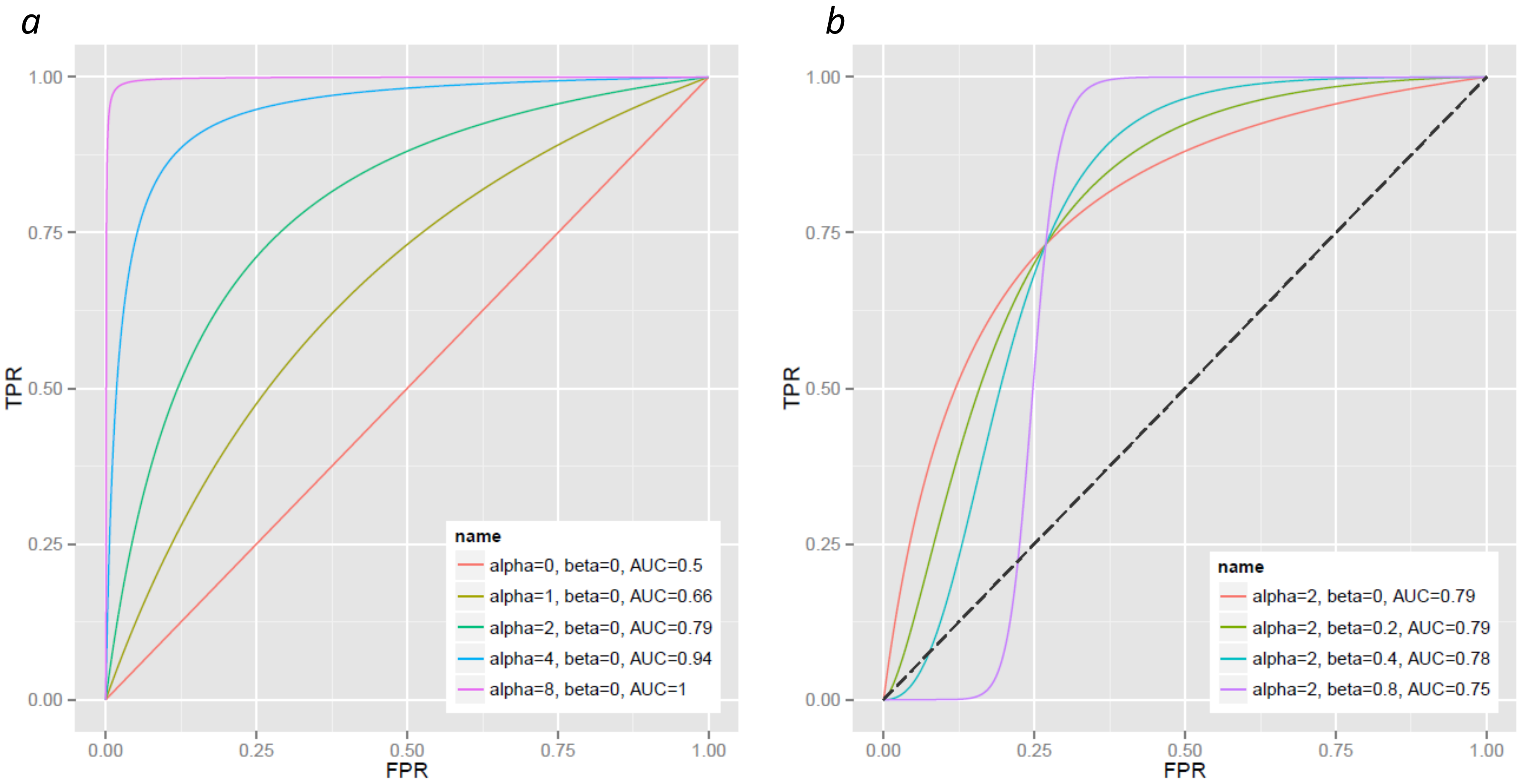
THE LANCET
Respiratory Medicine

Supplementary appendix

Supplement to: Sweeney TE, Braviak L, Tato CM, Khatri P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med* 2016; published online Feb 19. http://dx.doi.org/10.1016/S2213-2600(16)00048-5.



Appendix page 1. Example ROC curves constructed using the method of Kester and Buntinx for a range of (a) alpha and (b) beta, showing the effect of varying the different parameters on both ROC curve shape and AUC. For summary ROC curves, alpha and beta are calculated from a random-effects model from the contributing datasets.

Appendix P2. List of all genes found to be significant (q<0.01, ES>1.5 fold) in multi-cohort analysis, sorted according to absolute summary effect size.

	Number of studies	Summary effect size	Summary std. error	p.value	q.value (FDR)	tau squared	Q	df	resid.hetero p.value
GBP5	3	1.574	0.368	1.90E-05	0.000277495	0.382	35.899	2	1.60E-08
ANKRD22	3	1.443	0.233	6.27E-10	2.65E-08	0.152	30.041	2	3.00E-07
GBP2	3	1.367	0.349	8.93E-05	0.001059851	0.341	33.813	2	4.55E-08
BATF2	3	1.355	0.223	1.24E-09	4.97E-08	0.126	13.94	2	0.000939474
WDFY1	3	1.284	0.248	2.11E-07	5.24E-06	0.161	17.384	2	0.000167895
GBP6	3	1.284	0.302	2.14E-05	0.000308173	0.262	51.137	2	7.87E-12
CD274	3	1.271	0.325	9.14E-05	0.00108001	0.293	29.955	2	3.13E-07
DUSP3	3	1.27	0.393	0.0012232	0.009736512	0.44	43.46	2	3.65E-10
VAMP5	3	1.228	0.231	1.03E-07	2.76E-06	0.138	15.345	2	0.000465486
C1QB	3	1.184	0.141	4.83E-17	7.15E-15	0.039	5.805	2	0.054891213
CASP5	3	1.169	0.252	3.41E-06	6.09E-05	0.167	18.393	2	0.000101397
FLVCR2	3	1.161	0.282	3.75E-05	0.000499422	0.215	23.056	2	9.85E-06
GBP1	3	1.1	0.1	5.58E-28	3.73E-25	0.02	5.984	2	0.050189373
FAM26F	3	1.098	0.201	4.70E-08	1.38E-06	0.099	11.905	2	0.002598863
ETV7	3	1.084	0.207	1.76E-07	4.42E-06	0.107	12.709	2	0.001739083
BRSK1	3	1.076	0.322	0.0008389	0.007183124	0.289	30.567	2	2.30E-07
LAP3	3	1.06	0.097	7.36E-28	4.53E-25	0.01	2.976	2	0.225873562
PSME2	3	1.06	0.121	1.67E-18	3.34E-16	0.024	4.395	2	0.111098018
TAP1	3	1.054	0.122	7.20E-18	1.27E-15	0.025	4.515	2	0.104585959
PSMB9	3	1.046	0.204	2.81E-07	6.76E-06	0.113	24.752	2	4.22E-06
C1QC	3	1.044	0.091	1.16E-30	1.05E-27	0.007	2.676	2	0.262414434
GBP4	3	1.027	0.101	3.89E-24	1.50E-21	0.012	3.231	2	0.198778423
SCO2	3	1.026	0.13	2.34E-15	2.59E-13	0.03	5.083	2	0.078764017
SOCS1	3	1.017	0.19	8.49E-08	2.31E-06	0.087	10.795	2	0.004526958
PSTPIP2	3	1.017	0.216	2.49E-06	4.62E-05	0.119	13.952	2	0.000934147
CACNA1E	3	1.006	0.174	7.67E-09	2.59E-07	0.07	9.096	2	0.010590192
GK	3	0.994	0.125	1.52E-15	1.75E-13	0.036	9.349	2	0.009328877

LHFPL2	3	0.993	0.216	4.16E-06	7.27E-05	0.118	14.068	2	0.000881508
C2	3	0.992	0.157	2.57E-10	1.18E-08	0.053	7.456	2	0.024036629
GADD45B	3	0.984	0.116	2.38E-17	3.77E-15	0.021	4.133	2	0.126640068
PDCD1LG2	3	0.981	0.073	9.46E-41	2.42E-37	0	1.909	2	0.385055613
STAT1	3	0.974	0.122	1.27E-15	1.48E-13	0.037	13.425	2	0.001215437
KCNJ2	3	0.974	0.221	1.09E-05	0.000171987	0.126	14.752	2	0.000626147
AIM2	3	0.97	0.052	1.17E-77	1.80E-73	0	1.544	2	0.462092605
SLC6A12	3	0.969	0.196	7.67E-07	1.62E-05	0.094	11.591	2	0.003041707
P2RY14	3	0.931	0.121	1.77E-14	1.69E-12	0.037	13.394	2	0.001234303
FRMD3	3	0.924	0.073	1.11E-36	2.43E-33	0	0.737	2	0.691754679
C5	3	0.924	0.17	5.78E-08	1.65E-06	0.066	8.835	2	0.012062456
TIFA	3	0.921	0.134	6.91E-12	4.15E-10	0.034	5.498	2	0.064001875
TAP2	3	0.918	0.166	3.40E-08	1.03E-06	0.072	16.864	2	0.00021784
TRAFD1	3	0.917	0.127	4.86E-13	3.61E-11	0.028	4.925	2	0.085235356
SESTD1	3	0.913	0.093	1.66E-22	5.32E-20	0.008	2.871	2	0.238032687
IFI30	3	0.91	0.073	1.14E-35	2.19E-32	0	0.547	2	0.760597893
PARP14	3	0.9	0.073	5.60E-35	9.56E-32	0	0.681	2	0.711289799
MOV10	3	0.897	0.106	2.05E-17	3.36E-15	0.015	3.516	2	0.172397986
TRIM21	3	0.892	0.195	4.75E-06	8.18E-05	0.093	11.592	2	0.00303987
TIMM10	3	0.889	0.094	2.89E-21	7.40E-19	0.009	2.899	2	0.234701745
HIST2H2AC	3	0.888	0.131	1.16E-11	6.70E-10	0.032	5.281	2	0.071319803
CFB	3	0.888	0.166	8.72E-08	2.37E-06	0.062	8.425	2	0.014807989
FBXO6	3	0.873	0.143	8.94E-10	3.69E-08	0.041	6.269	2	0.043511007
EPSTI1	3	0.868	0.117	1.17E-13	9.78E-12	0.022	4.256	2	0.119075413
TAPBPL	3	0.86	0.244	0.0004308	0.004086438	0.158	18.254	2	0.000108691
KREMEN1	3	0.859	0.167	2.89E-07	6.89E-06	0.077	25.914	2	2.36E-06
MICB	3	0.857	0.083	5.55E-25	2.37E-22	0.004	2.389	2	0.302916642
RTP4	3	0.853	0.073	7.57E-32	8.95E-29	0	0.903	2	0.636725031
SLAMF8	3	0.84	0.073	5.04E-31	4.85E-28	0	0.012	2	0.99388584
KCNJ15	3	0.839	0.089	6.65E-21	1.62E-18	0.017	7.388	2	0.024868953

LMNB1	3	0.828	0.208	6.71E-05	0.000829611	0.108	13.273	2	0.001311632
ZNF438	3	0.824	0.19	1.47E-05	0.000223401	0.088	11.153	2	0.00378622
APOL6	3	0.823	0.073	1.70E-29	1.38E-26	0	2.016	2	0.364926861
SP140	3	0.815	0.105	7.51E-15	7.65E-13	0.026	10.16	2	0.006221166
EIF4G3	3	0.814	0.077	4.79E-26	2.37E-23	0.002	2.169	2	0.338071992
IFITM3	3	0.806	0.072	7.72E-29	5.65E-26	0	1.039	2	0.594727199
TRIM56	3	0.803	0.229	0.0004456	0.004206879	0.136	16.167	2	0.000308598
DYNLT1	3	0.801	0.126	2.00E-10	9.40E-09	0.028	4.962	2	0.083652027
ADM	3	0.786	0.102	1.43E-14	1.38E-12	0.013	3.381	2	0.184428485
IFI35	3	0.784	0.072	1.82E-27	1.04E-24	0	1.264	2	0.531438766
STAT2	3	0.776	0.072	5.59E-27	3.07E-24	0	0.275	2	0.871615806
RARRES3	3	0.775	0.101	1.95E-14	1.84E-12	0.013	3.33	2	0.189186466
FGL2	3	0.775	0.089	2.79E-18	5.36E-16	0.006	2.688	2	0.260859625
TNFSF10	3	0.775	0.072	6.90E-27	3.66E-24	0	0.489	2	0.783210975
CEACAM1	3	0.773	0.047	1.12E-60	8.63E-57	0.001	2.348	2	0.309098332
PARP9	3	0.761	0.068	2.63E-29	2.02E-26	0.005	3.029	2	0.219965993
IGF2BP3	3	0.76	0.127	2.08E-09	7.99E-08	0.029	5.066	2	0.079410063
HIST1H3D	3	0.759	0.106	8.91E-13	6.40E-11	0.015	3.624	2	0.163337865
FAM111A	3	0.755	0.225	0.0008015	0.006928204	0.142	31.654	2	1.34E-07
C1QA	3	0.755	0.152	6.64E-07	1.43E-05	0.049	7.245	2	0.026718449
APOL1	3	0.753	0.131	8.64E-09	2.88E-07	0.041	10.731	2	0.004676191
BST2	3	0.752	0.121	5.31E-10	2.31E-08	0.025	4.618	2	0.099336644
EPB41L3	3	0.751	0.231	0.0011138	0.009062602	0.139	16.529	2	0.000257502
ATF5	3	0.751	0.072	1.92E-25	8.94E-23	0	1.899	2	0.387007848
PLAUR	3	0.75	0.218	0.0005877	0.005319643	0.136	44.678	2	1.99E-10
SAMD9L	3	0.736	0.072	1.78E-24	7.21E-22	0	1.527	2	0.466101427
IFIT3	3	0.734	0.119	6.45E-10	2.70E-08	0.035	13.251	2	0.001326337
MR1	3	0.731	0.12	1.19E-09	4.80E-08	0.024	4.604	2	0.100069318
XAF1	3	0.731	0.151	1.27E-06	2.56E-05	0.058	14.216	2	0.000818734
USF1	3	0.73	0.126	6.36E-09	2.19E-07	0.037	9.876	2	0.007168402

DHRS12	3	0.727	0.18	5.62E-05	0.000711263	0.091	30.464	2	2.43E-07
IGSF6	3	0.722	0.193	0.000176	0.001898478	0.091	11.586	2	0.003048454
SLC26A8	3	0.721	0.087	1.73E-16	2.37E-14	0.018	9.623	2	0.008135549
CYBB	3	0.719	0.072	1.60E-23	6.01E-21	0	0.598	2	0.741434758
TRIM22	3	0.711	0.178	6.32E-05	0.000787765	0.074	9.846	2	0.007276785
IL27	3	0.71	0.18	8.26E-05	0.000989568	0.077	10.214	2	0.00605271
RAB24	3	0.71	0.204	0.0004978	0.004628401	0.12	65.274	2	6.66E-15
PSME1	3	0.709	0.133	1.06E-07	2.83E-06	0.043	11.163	2	0.003767782
TMEM140	3	0.706	0.087	6.91E-16	8.51E-14	0.006	2.64	2	0.267161716
SECTM1	3	0.704	0.122	8.03E-09	2.69E-07	0.026	4.756	2	0.092729384
ATG3	3	0.699	0.08	1.89E-18	3.73E-16	0.01	4.029	2	0.133396695
KARS	3	0.695	0.072	3.22E-22	9.90E-20	0	0.887	2	0.641683701
TNFAIP2	3	0.695	0.107	7.09E-11	3.58E-09	0.016	3.684	2	0.158466306
KIF1B	3	0.692	0.208	0.0009061	0.007647731	0.123	41.082	2	1.20E-09
JAK2	3	0.688	0.072	7.94E-22	2.22E-19	0	1.952	2	0.376857509
CNDP2	3	0.687	0.072	8.91E-22	2.45E-19	0	1.008	2	0.604132344
CNIH4	3	0.687	0.21	0.0010986	0.008971636	0.112	13.877	2	0.000969801
IFITM1	3	0.686	0.144	1.86E-06	3.57E-05	0.043	6.542	2	0.037977813
LACTB	3	0.68	0.105	9.33E-11	4.63E-09	0.026	10.365	2	0.005613098
TCN2	3	0.677	0.091	7.97E-14	6.92E-12	0.007	2.797	2	0.246936191
ADCY3	3	0.677	0.175	0.0001066	0.001227128	0.071	9.617	2	0.008161259
ACOT9	3	0.675	0.177	0.0001316	0.001475102	0.083	19.629	2	5.47E-05
UBE2L6	3	0.675	0.087	8.32E-15	8.36E-13	0.016	7.077	2	0.029064067
HPSE	3	0.675	0.062	1.52E-27	8.98E-25	0.003	2.658	2	0.264721861
ALDH1A1	3	0.675	0.075	2.54E-19	5.43E-17	0.008	3.669	2	0.159682224
SQRDL	3	0.67	0.082	3.39E-16	4.41E-14	0.004	2.404	2	0.30064679
RHBDF2	3	0.669	0.19	0.0004199	0.004013253	0.103	44.707	2	1.96E-10
LIMK2	3	0.666	0.149	7.78E-06	0.000126938	0.061	28.033	2	8.18E-07
PRPS2	3	0.663	0.112	3.52E-09	1.30E-07	0.028	7.97	2	0.018591341
IL15	3	0.66	0.128	2.26E-07	5.57E-06	0.044	20.476	2	3.58E-05

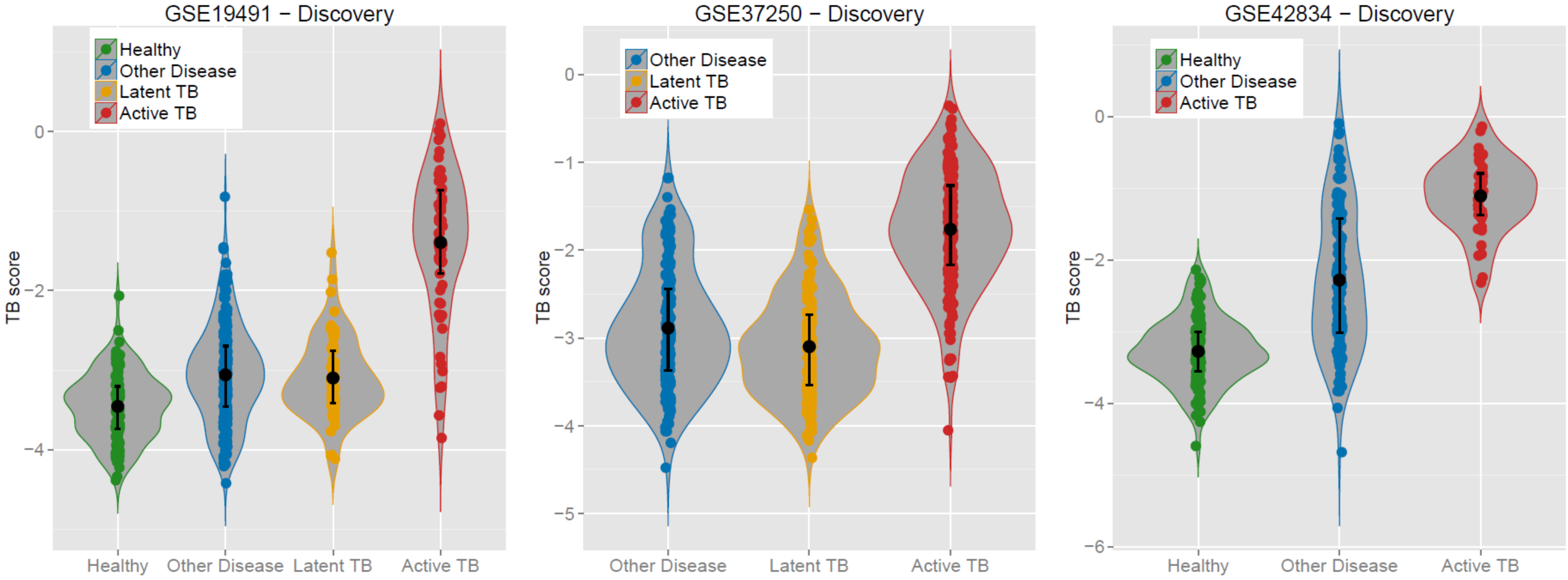
HIST1H2BG	3	0.659	0.072	3.56E-20	8.43E-18	0	1.389	2	0.499284234
GPR65	3	0.659	0.177	0.0001922	0.002051609	0.083	19.71	2	5.25E-05
DTX3L	3	0.658	0.132	6.43E-07	1.39E-05	0.033	5.578	2	0.061489104
CASP1	3	0.657	0.08	3.12E-16	4.11E-14	0.013	6.025	2	0.049158857
SORT1	3	0.651	0.197	0.0009706	0.008094196	0.097	12.296	2	0.00213792
SAT1	3	0.648	0.072	1.29E-19	2.84E-17	0	0.475	2	0.78844407
GPBAR1	3	0.646	0.173	0.0001876	0.002009268	0.079	18.919	2	7.79E-05
KLHDC8B	3	0.645	0.121	1.03E-07	2.76E-06	0.025	4.729	2	0.093986648
TNFSF13B	3	0.642	0.128	5.41E-07	1.20E-05	0.039	10.39	2	0.00554416
TLR7	3	0.637	0.102	3.83E-10	1.70E-08	0.013	3.421	2	0.180731706
OBFC2A	3	0.634	0.123	2.80E-07	6.75E-06	0.027	4.854	2	0.0883023
ZCCHC6	3	0.632	0.071	9.68E-19	1.99E-16	0	0.435	2	0.804551786
ZBP1	3	0.631	0.158	6.26E-05	0.000782546	0.055	7.864	2	0.019606348
XRN1	3	0.63	0.091	4.43E-12	2.77E-10	0.015	5.237	2	0.072900748
IFI6	3	0.628	0.149	2.57E-05	0.000361205	0.057	14.106	2	0.000864661
MFSD7	3	0.627	0.136	3.78E-06	6.66E-05	0.036	5.868	2	0.053173459
KYNU	3	0.624	0.166	0.0001711	0.001846936	0.072	17.507	2	0.000157918
CTSL1	3	0.624	0.127	8.53E-07	1.78E-05	0.041	15.148	2	0.000513591
FAS	3	0.623	0.051	6.81E-35	1.05E-31	0	1.118	2	0.571680966
SRBD1	3	0.62	0.097	1.47E-10	7.08E-09	0.011	3.135	2	0.208523247
BTN3A1	3	0.617	0.079	5.70E-15	5.96E-13	0.009	4.028	2	0.133439215
PLSCR1	3	0.616	0.144	1.81E-05	0.000266265	0.042	6.532	2	0.038164925
SCARF1	3	0.613	0.051	1.50E-33	1.92E-30	0	1.262	2	0.532075935
HLA-DMA	3	0.612	0.171	0.0003354	0.003307439	0.067	9.235	2	0.009879833
FAM20A	3	0.611	0.12	3.48E-07	8.12E-06	0.024	4.626	2	0.098954338
SLITRK4	3	0.611	0.141	1.53E-05	0.000231871	0.04	6.35	2	0.041800713
C5orf15	3	0.61	0.109	1.94E-08	6.09E-07	0.017	3.85	2	0.145871498
ASGR1	3	0.609	0.175	0.0005167	0.004763803	0.072	9.74	2	0.00767213
LMO2	3	0.607	0.141	1.66E-05	0.00024622	0.04	6.302	2	0.042813603
CDS2	3	0.607	0.072	3.41E-17	5.19E-15	0	2.022	2	0.363779764

SIPA1L1	3	0.606	0.071	1.94E-17	3.21E-15	0	1.113	2	0.573209414
CXCL10	3	0.605	0.071	2.32E-17	3.71E-15	0	0.847	2	0.654779493
TMEM180	3	0.601	0.177	0.0006808	0.006030432	0.074	9.911	2	0.007044812
LMTK2	3	0.596	0.097	6.55E-10	2.74E-08	0.01	3.132	2	0.208829662
BAZ1A	3	0.595	0.095	4.49E-10	1.98E-08	0.018	5.788	2	0.055351089
HIST2H2AB	3	0.593	0.071	8.74E-17	1.24E-14	0	0.131	2	0.936483175
MTHFD2	3	0.593	0.042	3.68E-46	1.42E-42	0	1.716	2	0.423924627
FCER1G	3	0.593	0.13	5.28E-06	8.96E-05	0.032	5.428	2	0.066282799
IFNAR1	3	0.587	0.138	2.06E-05	0.000298249	0.038	6.072	2	0.048015736
TMEM51	3	0.587	0.13	5.93E-06	9.88E-05	0.031	5.388	2	0.067607336
CUL1	3	0.586	0.071	2.16E-16	2.94E-14	0	0.727	2	0.695100417
ZNF671	3	-0.586	0.16	0.0002559	0.002630618	0.057	8.177	2	0.016763463
CARD11	3	-0.587	0.179	0.0010118	0.008387598	0.076	10.127	2	0.006322618
WDR6	3	-0.59	0.071	1.23E-16	1.71E-14	0	0.121	2	0.941134192
TLE1	3	-0.592	0.183	0.0011988	0.009586658	0.08	10.587	2	0.00502502
HPCAL4	3	-0.592	0.128	3.77E-06	6.66E-05	0.03	5.262	2	0.071998691
ORAI1	3	-0.592	0.13	5.35E-06	9.07E-05	0.032	5.425	2	0.066378384
OXSR1	3	-0.595	0.139	1.90E-05	0.000277899	0.039	6.173	2	0.045658819
CYBASC3	3	-0.596	0.1	2.20E-09	8.42E-08	0.012	3.305	2	0.191615784
PPM1H	3	-0.596	0.071	6.20E-17	8.99E-15	0	1.749	2	0.417042934
CD28	3	-0.597	0.177	0.000726	0.006372689	0.074	9.895	2	0.007100976
EHBP1	3	-0.599	0.18	0.000856	0.0073093	0.077	10.242	2	0.005971204
TRRAP	3	-0.599	0.071	4.46E-17	6.65E-15	0	1.715	2	0.42432118
GOT2	3	-0.6	0.147	4.78E-05	0.000617274	0.046	6.92	2	0.031431178
PAFAH1B1	3	-0.6	0.137	1.17E-05	0.00018215	0.037	5.975	2	0.050417714
RPS4X	3	-0.6	0.156	0.0001172	0.001335431	0.063	15.471	2	0.000437046
SWAP70	3	-0.601	0.121	7.51E-07	1.59E-05	0.025	4.738	2	0.093564529
ABHD14A	3	-0.602	0.163	0.000213	0.002243088	0.059	8.385	2	0.015105623
CD5	3	-0.603	0.186	0.0011945	0.009567027	0.084	10.979	2	0.004129196
ERP27	3	-0.603	0.071	3.05E-17	4.69E-15	0	1.882	2	0.390277371

HLA-DOB	3	-0.604	0.155	0.0001022	0.001180533	0.053	7.688	2	0.021411388
FAM84B	3	-0.604	0.137	9.91E-06	0.000158002	0.037	5.963	2	0.050710696
AGMAT	3	-0.606	0.14	1.55E-05	0.000233521	0.04	6.258	2	0.043752888
ALDH9A1	3	-0.607	0.174	0.000495	0.004604862	0.071	9.623	2	0.008136561
CD19	3	-0.609	0.093	4.97E-11	2.54E-09	0.009	2.915	2	0.232808084
SIN3A	3	-0.61	0.071	1.24E-17	2.15E-15	0	0.011	2	0.994328331
CD27	3	-0.611	0.132	3.56E-06	6.32E-05	0.033	5.549	2	0.062389228
EP400	3	-0.612	0.071	1.06E-17	1.84E-15	0	0.772	2	0.679885211
FNBP1	3	-0.613	0.072	2.22E-17	3.60E-15	0	2.029	2	0.362558261
TPK1	3	-0.618	0.072	5.76E-18	1.04E-15	0.006	3.388	2	0.183738235
ASF1B	3	-0.621	0.071	3.50E-18	6.48E-16	0	1.831	2	0.400262413
IMPDH2	3	-0.622	0.071	3.26E-18	6.19E-16	0	1.562	2	0.457890027
CD79A	3	-0.623	0.155	5.82E-05	0.000734953	0.065	22.797	2	1.12E-05
SMYD3	3	-0.624	0.085	2.44E-13	1.93E-11	0.005	2.551	2	0.279345088
PLCG1	3	-0.629	0.193	0.0011049	0.009013592	0.101	23.614	2	7.45E-06
TXK	3	-0.631	0.17	0.000211	0.002226883	0.067	9.179	2	0.010158861
SUSD3	3	-0.632	0.138	4.91E-06	8.40E-05	0.038	6.087	2	0.047656284
GZMK	3	-0.633	0.165	0.0001268	0.00142902	0.062	8.624	2	0.013406411
TOMM20	3	-0.633	0.194	0.0011092	0.009044108	0.099	18.054	2	0.000120108
GTF3A	3	-0.637	0.162	8.07E-05	0.000968633	0.058	8.257	2	0.016109102
FAM129C	3	-0.639	0.132	1.34E-06	2.69E-05	0.033	5.584	2	0.061313259
SH2D3A	3	-0.639	0.161	7.01E-05	0.000859076	0.058	8.168	2	0.016841528
KIAA1737	3	-0.64	0.186	0.0005924	0.005352549	0.094	22.007	2	1.66E-05
PEX5	3	-0.641	0.088	3.47E-13	2.65E-11	0.006	2.682	2	0.261557345
AP1M1	3	-0.642	0.076	1.94E-17	3.21E-15	0.001	2.147	2	0.341770914
OLIG1	3	-0.645	0.132	1.06E-06	2.16E-05	0.033	5.521	2	0.063261974
BIN1	3	-0.647	0.145	8.58E-06	0.00013823	0.053	13.388	2	0.00123817
VPREB3	3	-0.649	0.165	8.37E-05	0.00100106	0.062	8.634	2	0.013341139
CALM1	3	-0.653	0.075	4.79E-18	8.77E-16	0.001	2.138	2	0.343393795
NOV	3	-0.66	0.099	2.46E-11	1.34E-09	0.012	3.238	2	0.19806942

SPTAN1	3	-0.662	0.072	3.96E-20	9.09E-18	0	2.012	2	0.365715713
USP11	3	-0.662	0.148	7.50E-06	0.000122738	0.046	6.901	2	0.031730154
MCM5	3	-0.663	0.145	5.13E-06	8.74E-05	0.044	6.69	2	0.035264604
RBBP7	3	-0.665	0.111	2.28E-09	8.67E-08	0.019	4	2	0.13536372
HRK	3	-0.665	0.078	1.66E-17	2.81E-15	0.002	2.236	2	0.326998067
IL27RA	3	-0.668	0.139	1.57E-06	3.09E-05	0.039	6.133	2	0.04658428
SMARCC1	3	-0.67	0.184	0.0002773	0.002811148	0.082	10.682	2	0.004791547
DKC1	3	-0.671	0.167	5.84E-05	0.000736002	0.063	8.764	2	0.012501505
PPIA	3	-0.672	0.195	0.0005621	0.005114589	0.093	11.924	2	0.002574681
SLC9A3R1	3	-0.674	0.186	0.0002964	0.002977053	0.084	10.895	2	0.004307032
CXCR5	3	-0.675	0.096	1.62E-12	1.12E-10	0.018	5.797	2	0.055103
EBF1	3	-0.675	0.178	0.0001489	0.001639758	0.075	9.977	2	0.006815125
SLAMF1	3	-0.676	0.174	9.98E-05	0.001157273	0.07	9.483	2	0.008723683
ACTR1B	3	-0.691	0.131	1.30E-07	3.37E-06	0.032	5.418	2	0.066598642
ZNF329	3	-0.692	0.072	5.42E-22	1.57E-19	0	0.328	2	0.848938988
MOAP1	3	-0.692	0.072	5.23E-22	1.55E-19	0	0.643	2	0.725024483
KLF13	3	-0.697	0.198	0.0004342	0.004108695	0.097	12.236	2	0.00220288
STK38	3	-0.697	0.119	4.16E-09	1.51E-07	0.023	4.51	2	0.104878303
RBL2	3	-0.7	0.135	2.17E-07	5.36E-06	0.035	5.752	2	0.056345857
FCRLA	3	-0.704	0.204	0.0005429	0.004972416	0.104	12.992	2	0.001509736
TRIM28	3	-0.705	0.176	5.97E-05	0.000749577	0.072	9.67	2	0.007944978
MFGE8	3	-0.712	0.078	9.27E-20	2.07E-17	0.002	2.233	2	0.327362543
CD79B	3	-0.713	0.12	3.10E-09	1.16E-07	0.037	13.645	2	0.001088759
MARCKSL1	3	-0.713	0.072	3.75E-23	1.28E-20	0	1.878	2	0.391031164
COL9A2	3	-0.716	0.072	2.15E-23	7.70E-21	0	0.307	2	0.857814779
PRPF8	3	-0.718	0.108	2.57E-11	1.38E-09	0.016	3.746	2	0.153695446
PNOC	3	-0.72	0.207	0.000516	0.004761899	0.108	13.429	2	0.001213308
RNF44	3	-0.727	0.09	9.40E-16	1.13E-13	0.007	2.771	2	0.250251738
SERTAD2	3	-0.731	0.072	2.93E-24	1.16E-21	0	1.48	2	0.477176062
CABIN1	3	-0.734	0.142	2.45E-07	5.98E-06	0.041	6.344	2	0.041930144

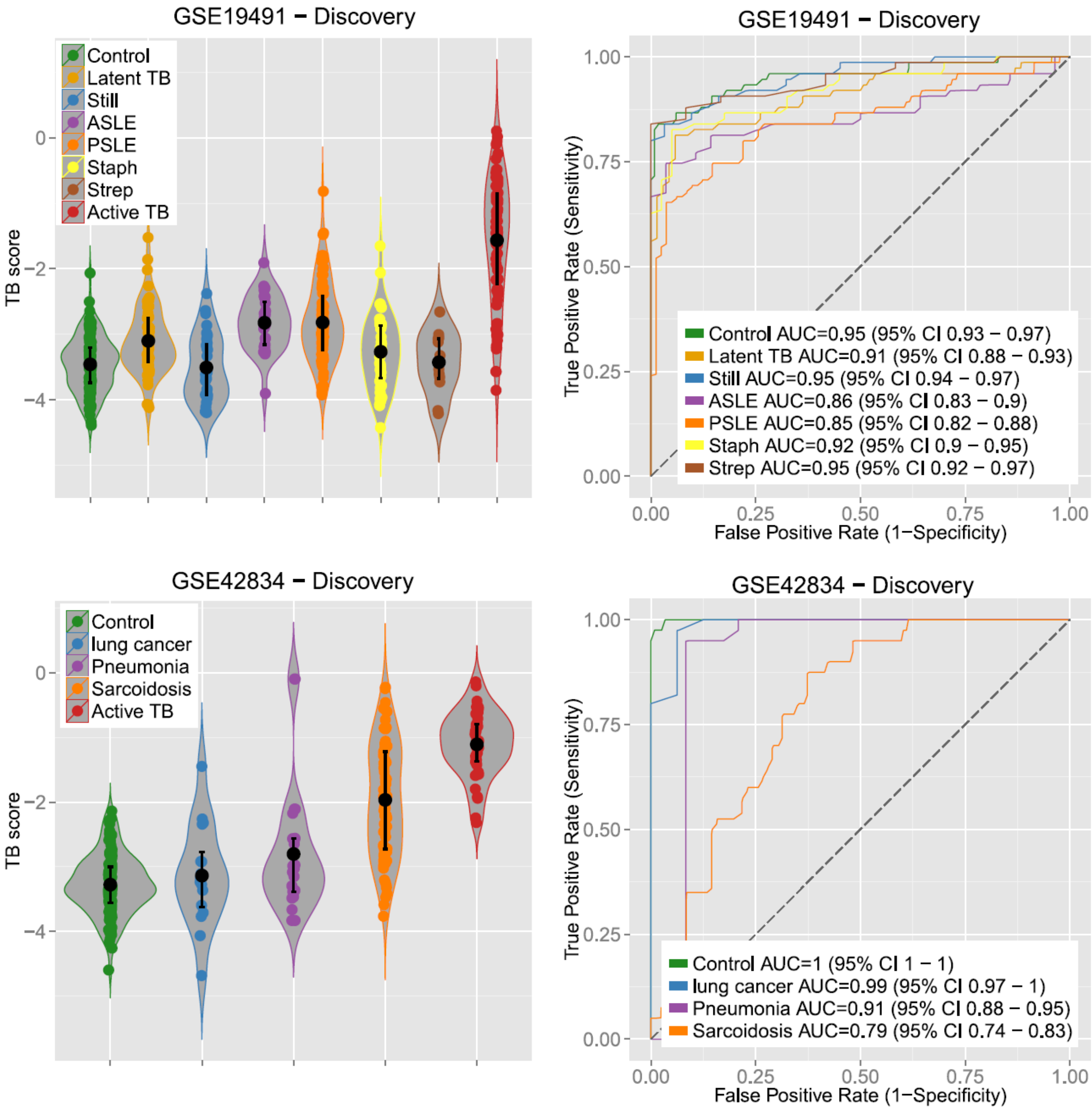
P2RY10	3	-0.734	0.148	7.38E-07	1.57E-05	0.056	13.759	2	0.001028624
NELL2	3	-0.736	0.2	0.0002254	0.00234877	0.099	12.42	2	0.002009736
EML4	3	-0.736	0.228	0.0012195	0.00971643	0.135	16.167	2	0.000308558
SYTL1	3	-0.739	0.111	2.37E-11	1.30E-09	0.018	3.934	2	0.139847946
PFAS	3	-0.739	0.147	4.80E-07	1.08E-05	0.045	6.759	2	0.034065856
KIAA0355	3	-0.74	0.072	9.44E-25	3.92E-22	0	0.644	2	0.724665277
BANK1	3	-0.742	0.213	0.0005011	0.004653322	0.115	14.155	2	0.000843776
TBC1D10C	3	-0.742	0.074	1.83E-23	6.69E-21	0.001	2.082	2	0.35317039
CACNA2D3	3	-0.748	0.123	1.18E-09	4.78E-08	0.026	4.794	2	0.090969012
PITPNC1	3	-0.752	0.198	0.0001516	0.001663865	0.108	24.529	2	4.72E-06
CA5B	3	-0.753	0.126	2.48E-09	9.35E-08	0.029	5.042	2	0.080392388
FLNB	3	-0.757	0.163	3.51E-06	6.25E-05	0.06	8.295	2	0.01580117
ID3	3	-0.768	0.134	1.14E-08	3.71E-07	0.035	5.679	2	0.058453652
EPHA4	3	-0.772	0.072	1.04E-26	5.36E-24	0	1.335	2	0.513032857
TATDN2	3	-0.781	0.226	0.0005524	0.005044102	0.133	15.848	2	0.000361953
ZNF518B	3	-0.789	0.222	0.0003791	0.003687358	0.127	15.276	2	0.000481785
GPX7	3	-0.797	0.221	0.0003113	0.003102844	0.126	15.104	2	0.000525157
MAP7	3	-0.799	0.127	3.57E-10	1.62E-08	0.038	10.078	2	0.006481724
BLK	3	-0.804	0.159	4.27E-07	9.73E-06	0.056	7.829	2	0.019953393
DBP	3	-0.815	0.197	3.37E-05	0.000453602	0.095	11.928	2	0.002569768
OSBPL10	3	-0.817	0.165	7.45E-07	1.58E-05	0.061	8.418	2	0.014864833
FAIM3	3	-0.818	0.19	1.58E-05	0.000237263	0.087	11.088	2	0.003911241
SESN1	3	-0.851	0.236	0.0003109	0.003100158	0.146	17.1	2	0.000193508
MEF2D	3	-0.861	0.162	1.15E-07	3.01E-06	0.058	8.051	2	0.017852313
KLF2	3	-0.881	0.073	1.05E-33	1.47E-30	0	1.369	2	0.504444863
ITPKB	3	-0.884	0.112	2.93E-15	3.20E-13	0.019	3.945	2	0.139101418
GNG7	3	-0.907	0.108	5.38E-17	7.88E-15	0.016	3.697	2	0.157440503
FOXO1	3	-0.951	0.214	9.19E-06	0.000147144	0.117	13.894	2	0.000961316
ARHGEF18	3	-1.027	0.235	1.23E-05	0.000190028	0.144	16.444	2	0.000268628



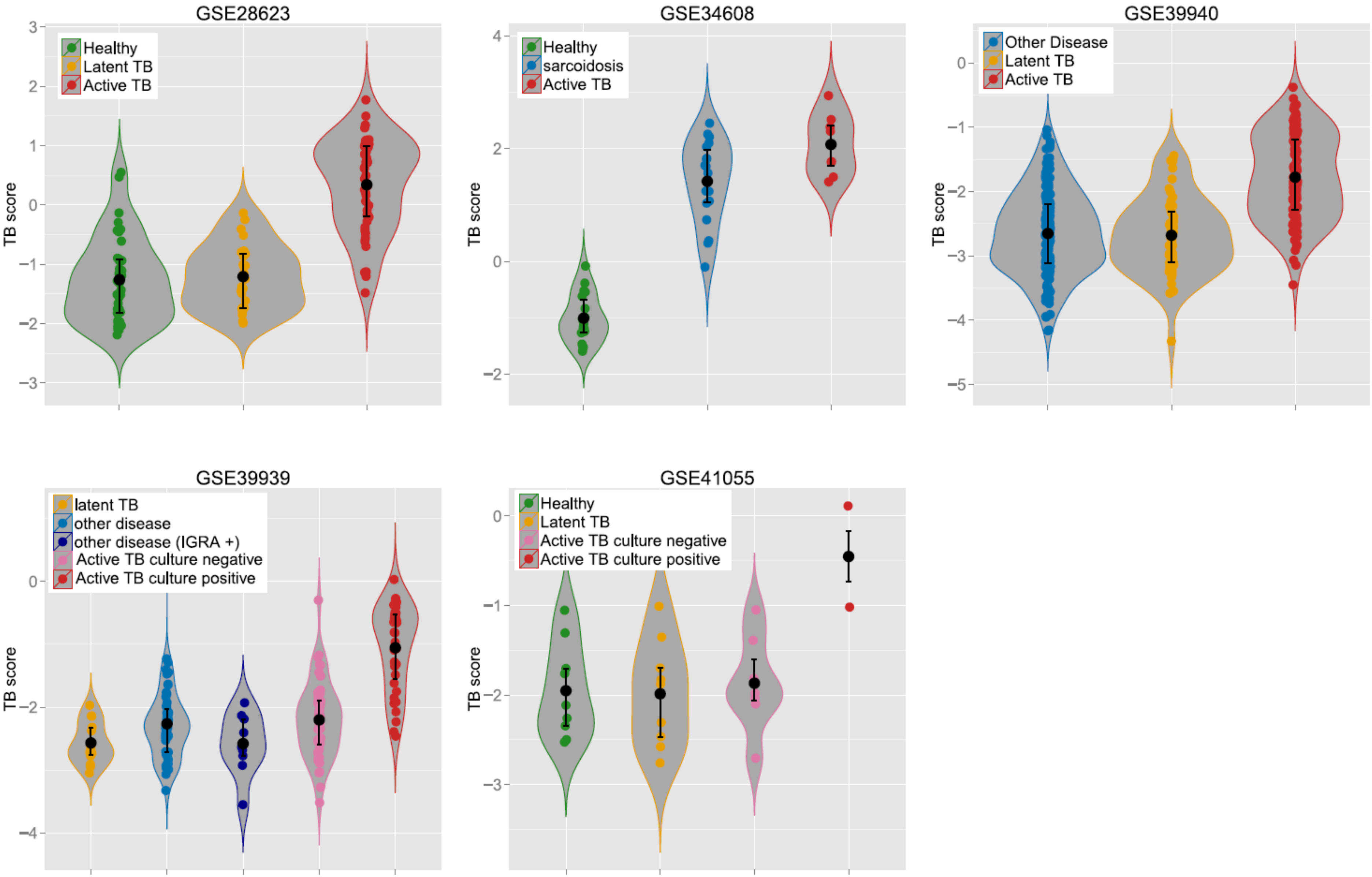
Appendix page 3. Performance of the three-gene set in the discovery datasets. (a-c) Violin plots of GSE19491, GSE32750, and GSE42834, respectively; all comparisons to ATB significant (Wilcoxon p<1e-10). Healthy patients were not included in multi-cohort analysis but are shown here.

Discovery - HC vs. ATB					
Dataset	sensitivity	specificity	PPV	NPV	accuracy
GSE19491	0.885	0.966	0.931	0.942	0.938
GSE42834	0.975	0.983	0.951	0.991	0.981
Discovery - LTB vs. ATB					
Dataset	sensitivity	specificity	PPV	NPV	accuracy
GSE19491	0.885	0.87	0.857	0.896	0.877
GSE37250	0.872	0.832	0.859	0.848	0.854
Discovery - OD vs. ATB					
Dataset	sensitivity	specificity	PPV	NPV	accuracy
GSE19491	0.869	0.855	0.654	0.954	0.858
GSE37250	0.805	0.789	0.809	0.784	0.797
GSE42834	0.8	0.732	0.492	0.918	0.748

Appendix page 4. Test parameters at an automated threshold (maximum sensitivity + specificity) in the discovery datasets. The threshold for each dataset was calculated separately, and then test statistics were generated from the resulting patient classifications. PPV, positive predictive value, NPV, negative predictive value.



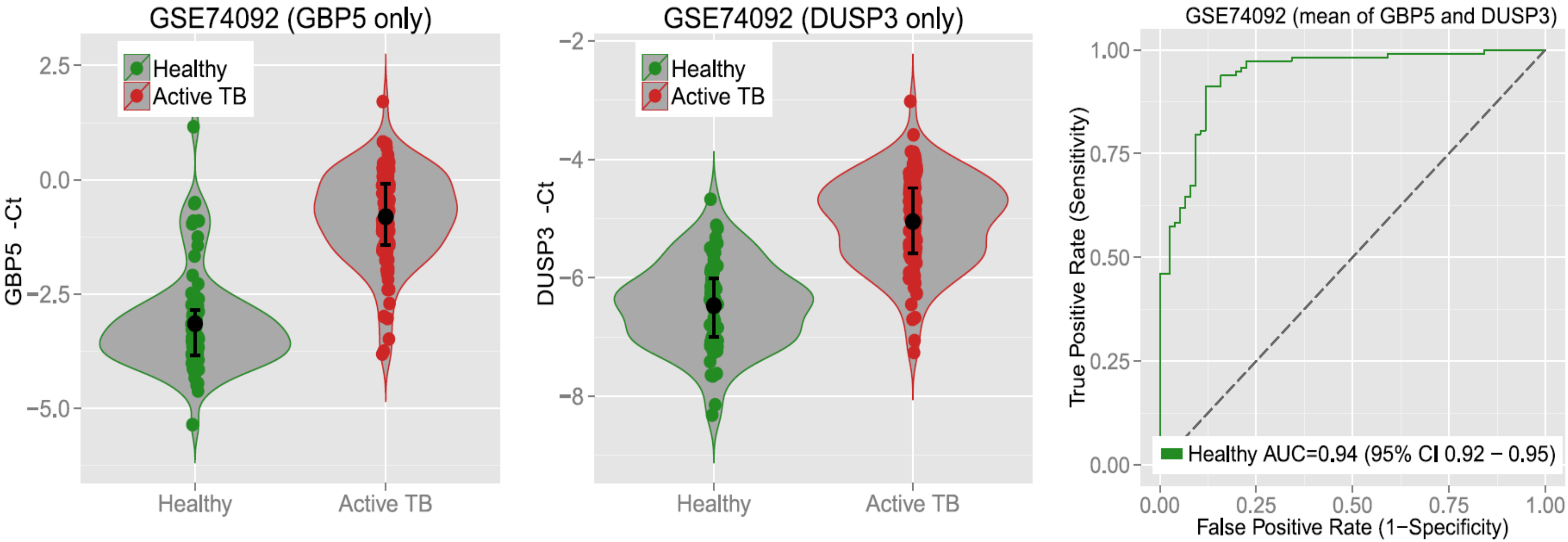
Appendix page 5. Breakdown of ‘Other Disease’ category by disease type in the discovery datasets for which this information is supplied.



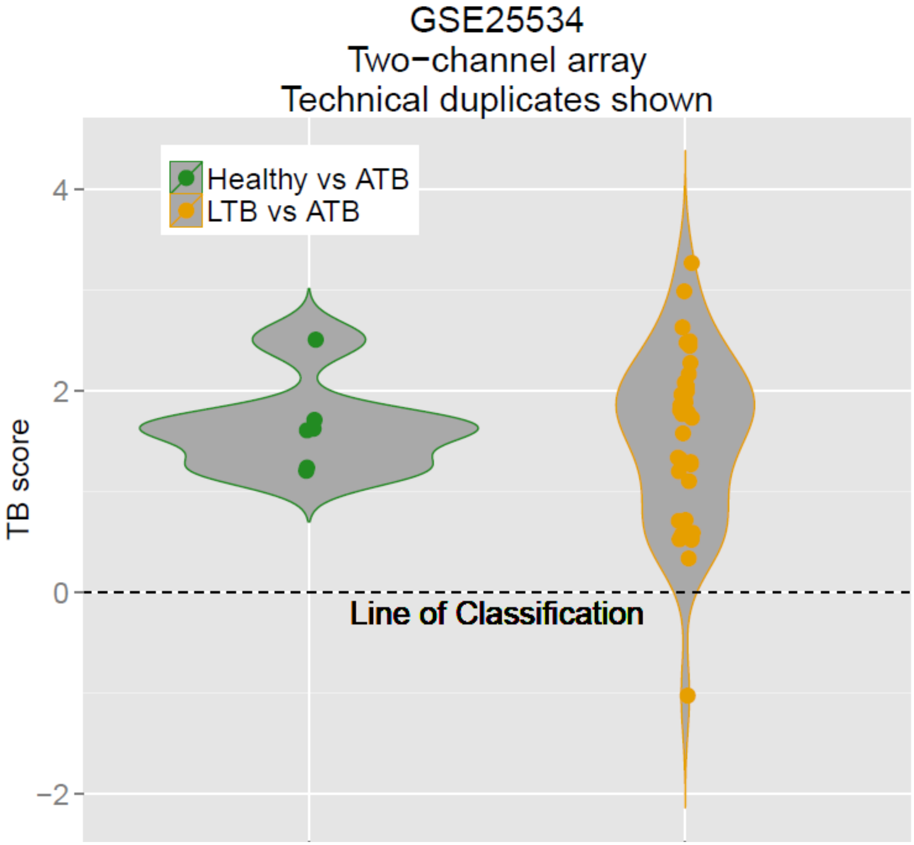
Appendix page 6. Violin plots of validation datasets not shown in Figure 6.

Validation - HC vs. ATB					
Dataset	**sensitivity**	**specificity**	**PPV**	**NPV**	**accuracy**
GSE28623	0.848	0.811	0.848	0.811	0.831
GSE34608	1	1	1	1	1
GSE41055 (children)	1	0.778	0.5	1	0.818
GSE56153	0.611	0.667	0.647	0.632	0.639
Validation - LTB vs. ATB					
Dataset	**sensitivity**	**specificity**	**PPV**	**NPV**	**accuracy**
GSE28623	0.87	0.84	0.909	0.778	0.859
GSE39939 (children)	0.886	0.929	0.969	0.765	0.898
GSE39940 (children)	0.712	0.759	0.859	0.562	0.727
GSE41055 (children)	1	0.889	0.667	1	0.909
Validation - OD vs. ATB					
Dataset	**sensitivity**	**specificity**	**PPV**	**NPV**	**accuracy**
GSE34608	0.5	0.611	0.364	0.733	0.577
GSE39939 (children)	0.771	0.875	0.771	0.875	0.838
GSE39940 (children)	0.685	0.74	0.633	0.781	0.718

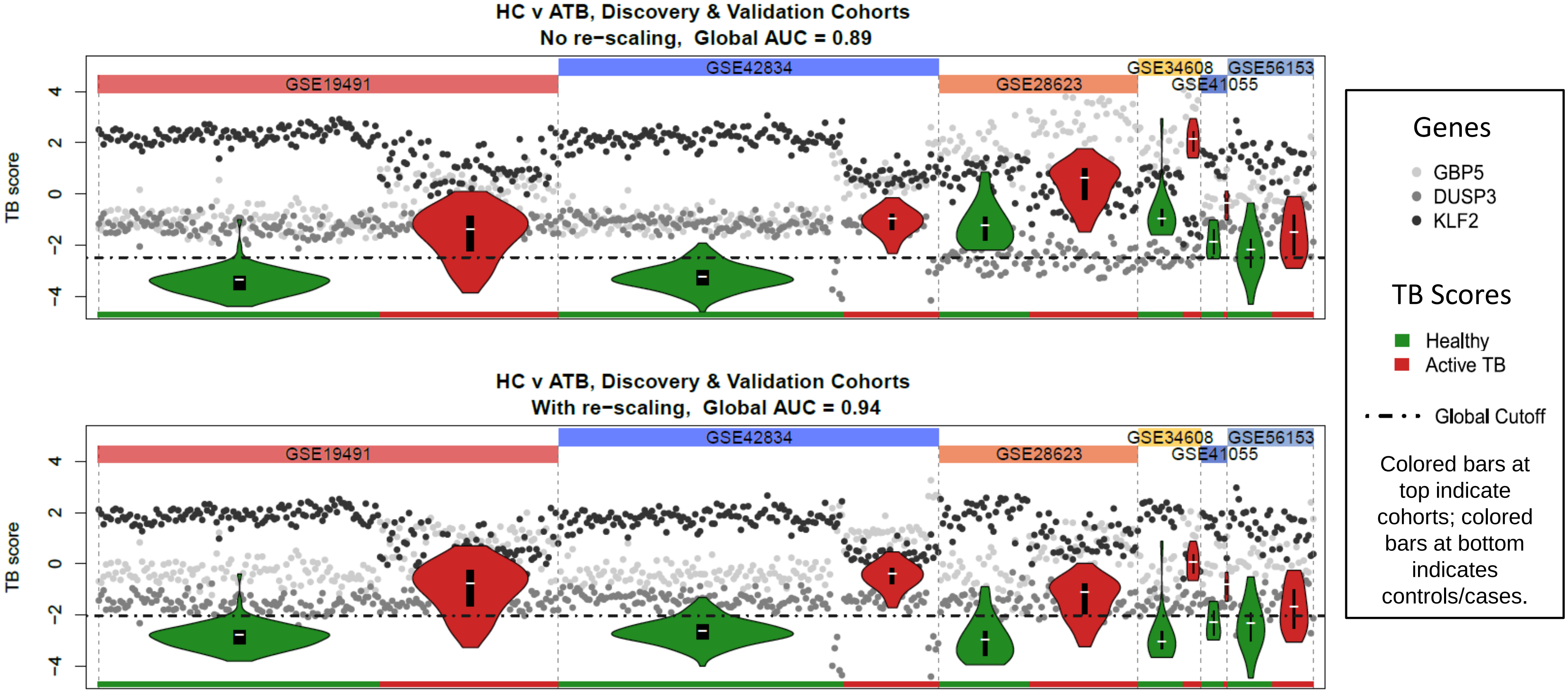
Appendix page 7. Test parameters at an automated threshold (maximum sensitivity + specificity) in the validation datasets. The threshold for each dataset was calculated separately, and then test statistics were generated from the resulting patient classifications. PPV, positive predictive value, NPV, negative predictive value.



Appendix page 8. Performance of *GBP5* and *DUSP3* in the RT-PCR data from GSE74092 (*KLF2* is not present in these data). The TB score from which the ROC is calculated is simply the mean of *GBP5* and *DUSP3*, as *KLF2* is not present.



Appendix page 9. TB score in GSE25534, which utilized a two-channel array, wherein gene expression values represent relative values between the two samples on the array. Here, a positive TB score means that the TB score was greater in the ATB sample on the array than the control (healthy or LTB) sample on the array. A positive TB score for a given array would thus correctly classify that ATB sample vs. that control sample. The violin plots thus indicate that all but one sample are correctly classified by the three-gene set. As with other two-channel array studies, GSE25534 contains technical duplicates, which are shown here.

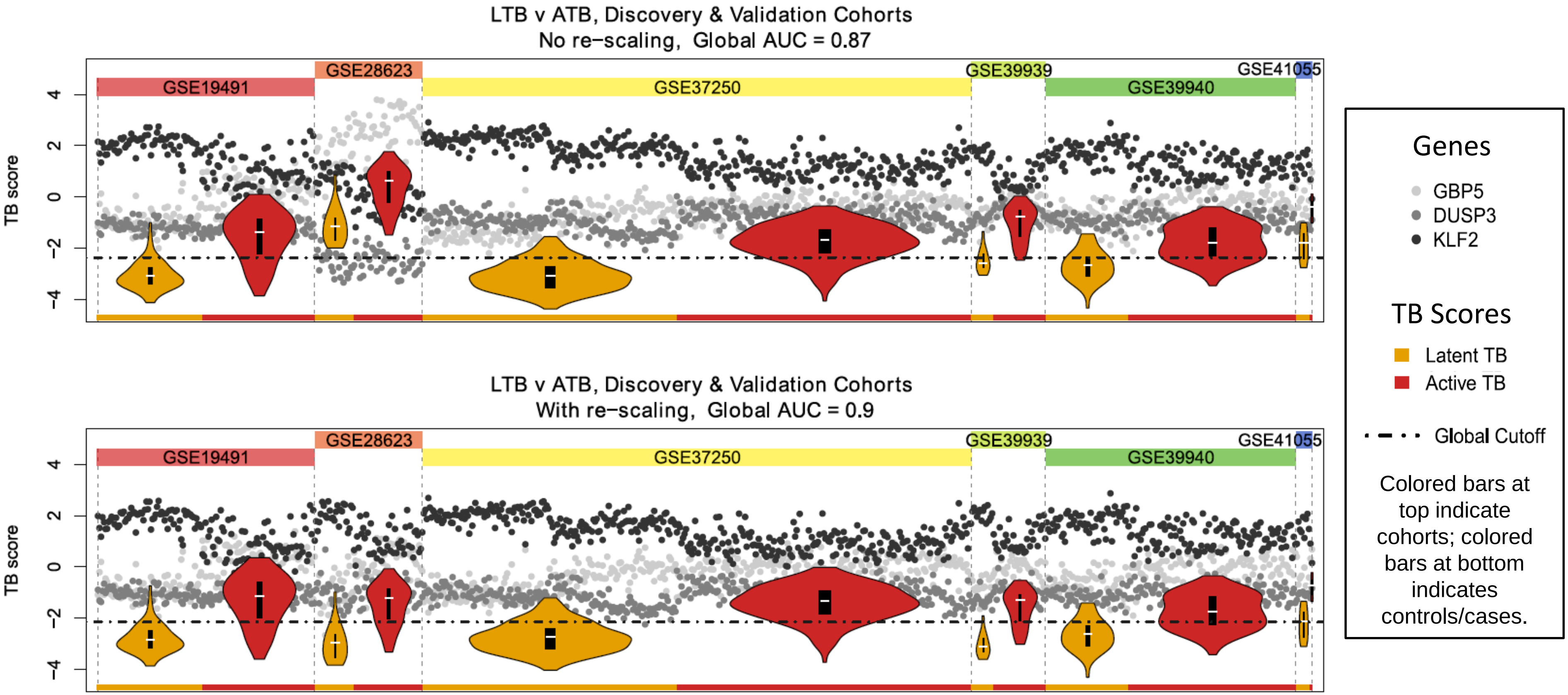


HC v ATB, Discovery & Validation Cohorts
No re−scaling, Global AUC = 0.89

HC v ATB, Discovery & Validation Cohorts
With re−scaling, Global AUC = 0.94

Colored bars at top indicate cohorts; colored bars at bottom indicate controls/cases.

Appendix page 10. Establishing a single global test cutoff in the joint discovery and validation datasets for HC vs ATB. Shown are sample-level normalized gene scores, along with group TB score distributions. (A) genes have not been re-centered to their global mean. (B) Genes have been re-centering to their global mean by subtracting the difference between the dataset mean and the global mean for each gene. Note that between (A) and (B), each gene maintains its distribution within a dataset.

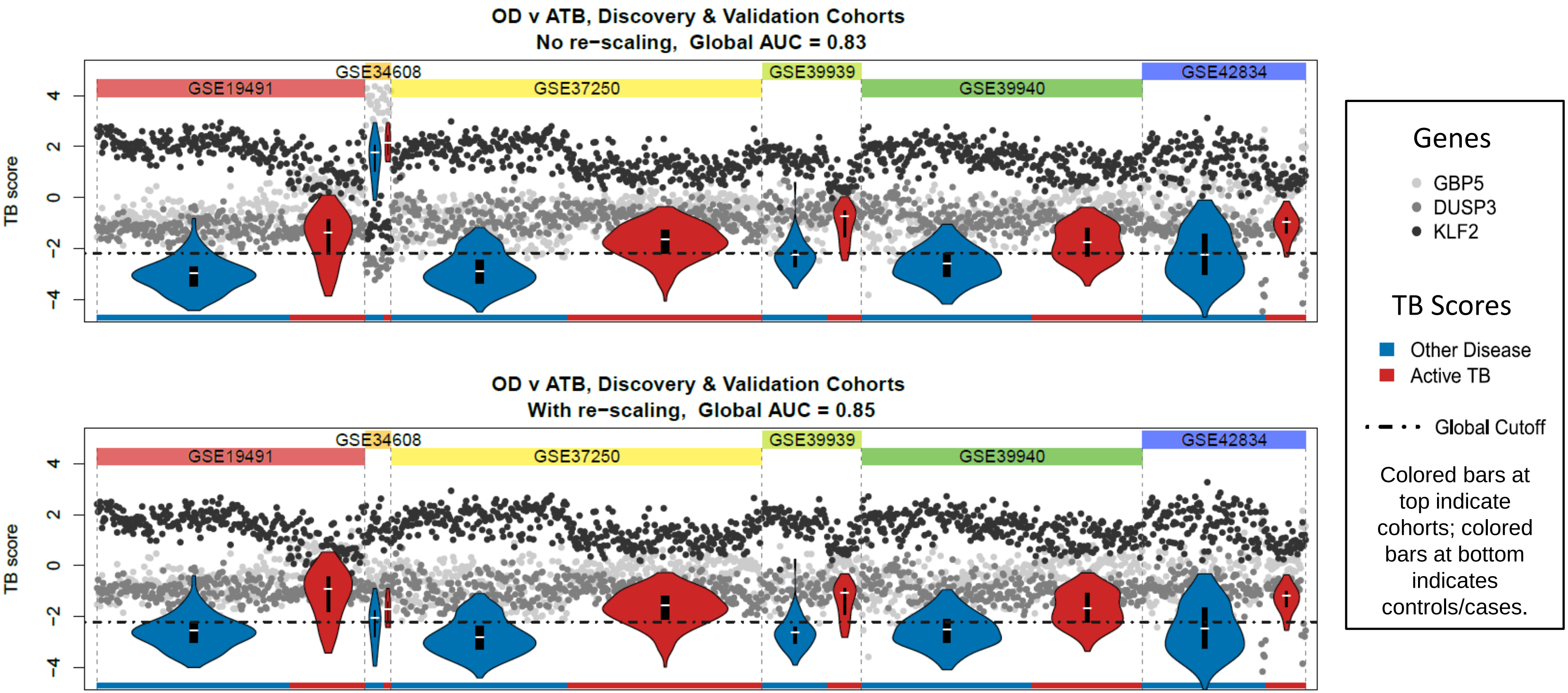


LTB v ATB, Discovery & Validation Cohorts
No re−scaling, Global AUC = 0.87

LTB v ATB, Discovery & Validation Cohorts
With re−scaling, Global AUC = 0.9

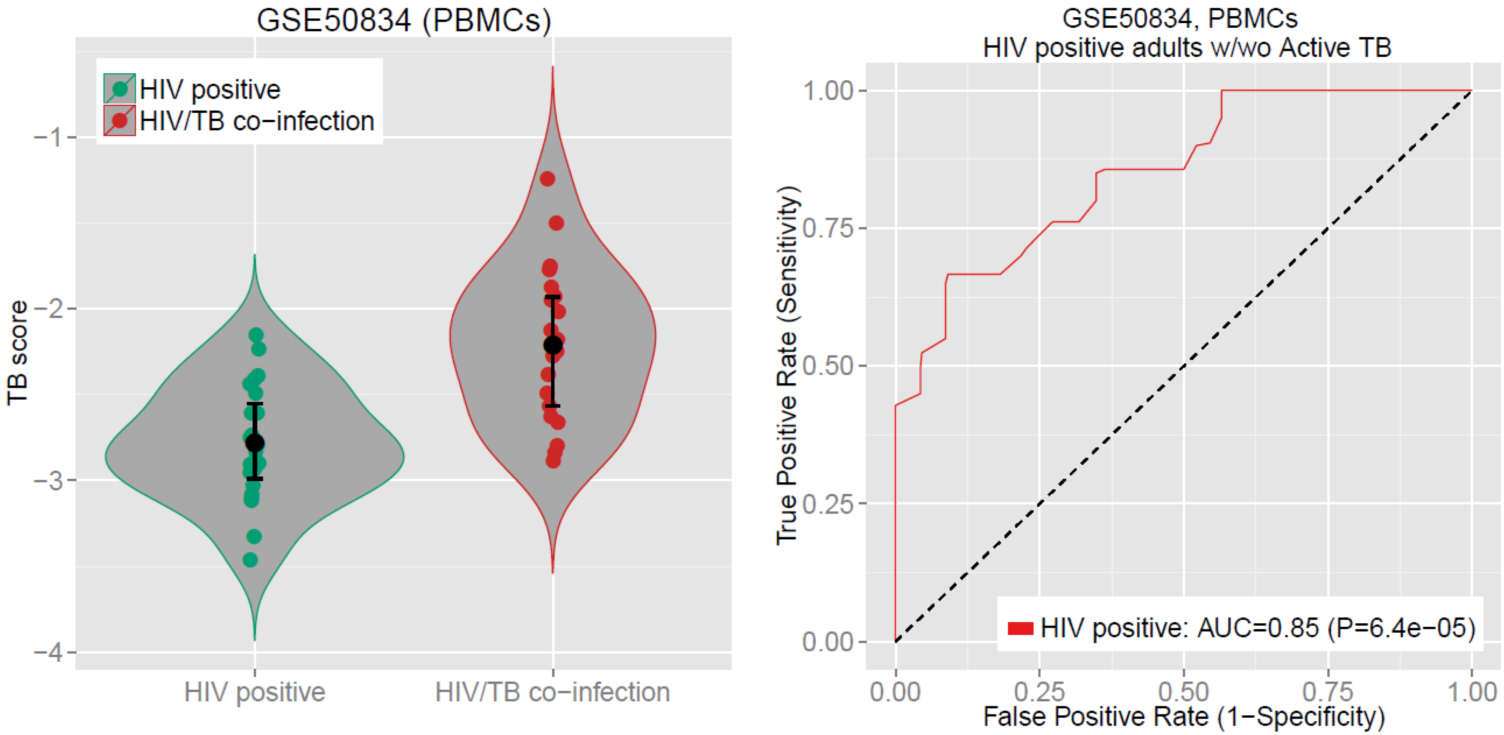
Colored bars at top indicate cohorts; colored bars at bottom indicates controls/cases.

Appendix page 11. Establishing a single global test cutoff in the joint discovery and validation datasets for LTB vs ATB. Shown are sample-level normalized gene scores, along with group TB score distributions. (A) genes have not been re-centered to their global mean. (B) Genes have been re-centering to their global mean by subtracting the difference between the dataset mean and the global mean for each gene. Note that between (A) and (B), each gene maintains its distribution within a dataset.

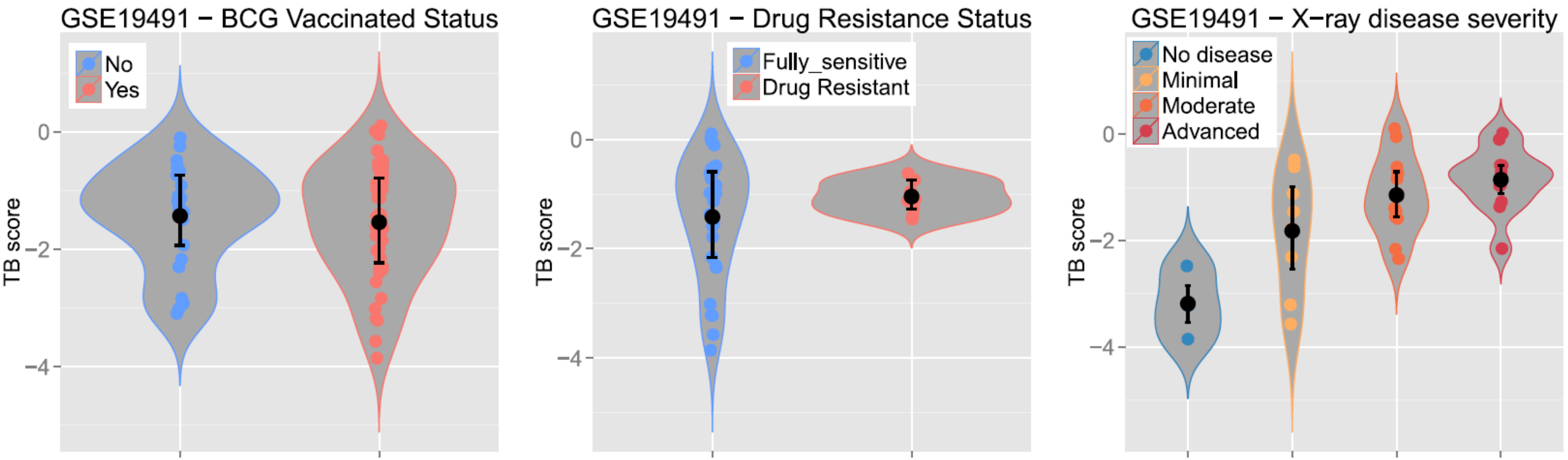


Colored bars at top indicate cohorts; colored bars at bottom indicates controls/cases.

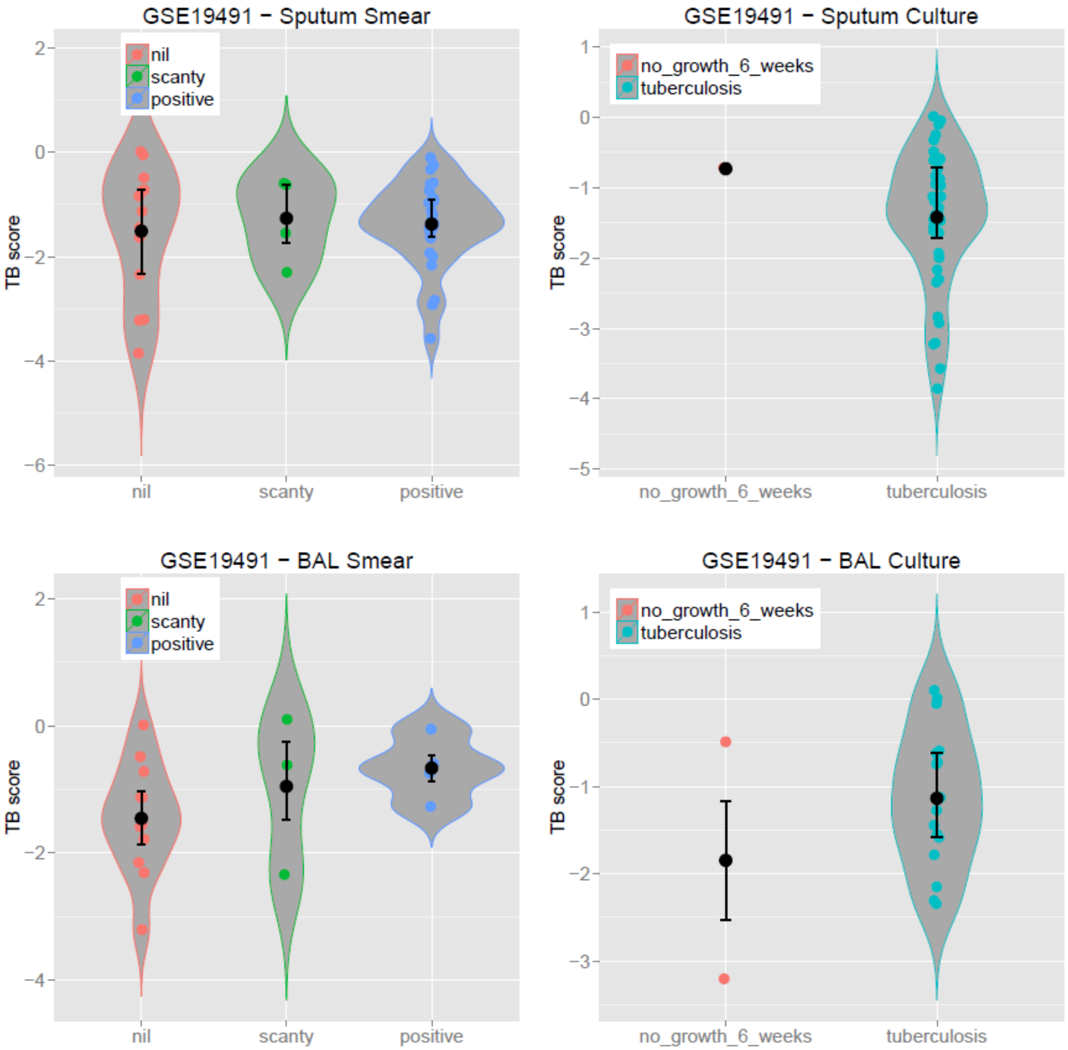
Appendix page 12. Establishing a single global test cutoff in the joint discovery and validation datasets for OD vs ATB. Shown are sample-level normalized gene scores, along with group TB score distributions. (A) genes have not been re-centered to their global mean. (B) Genes have been re-centering to their global mean by subtracting the difference between the dataset mean and the global mean for each gene. Note that between (A) and (B), each gene maintains its distribution within a dataset.



Appendix page 13. GSE50834 compared PBMCs in HIV-positive patients to those with HIV/TB co-infection. (A) The three gene set showed a significant difference between the two groups, with (B) an ROC AUC of 0.85.



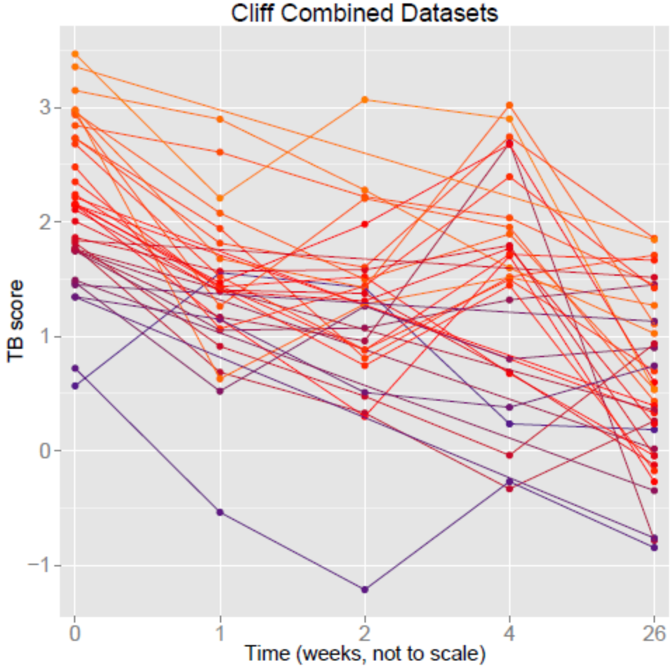
Appendix page 14. In dataset GSE19491, TB score was not affected by either (a) BCG vaccination status or (b) TB drug resistance status (both Wilcoxon p=NS), but (c) increased with X-ray disease severity (JT-test p<0.01).

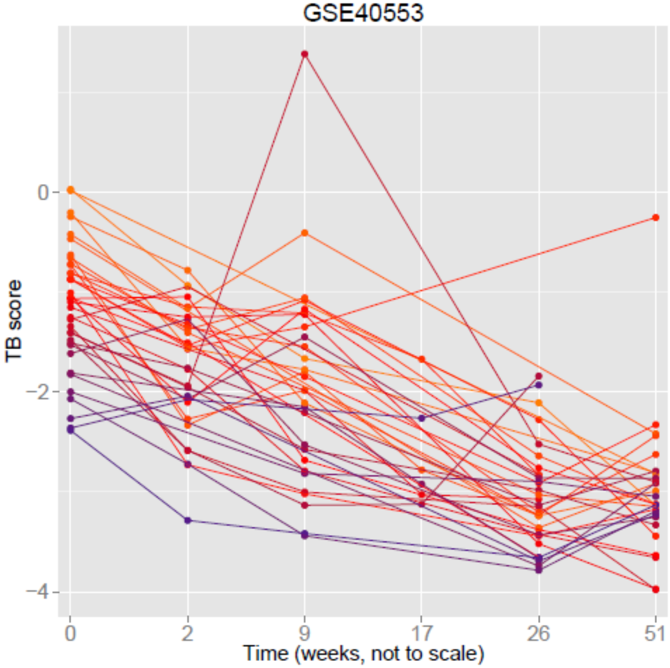


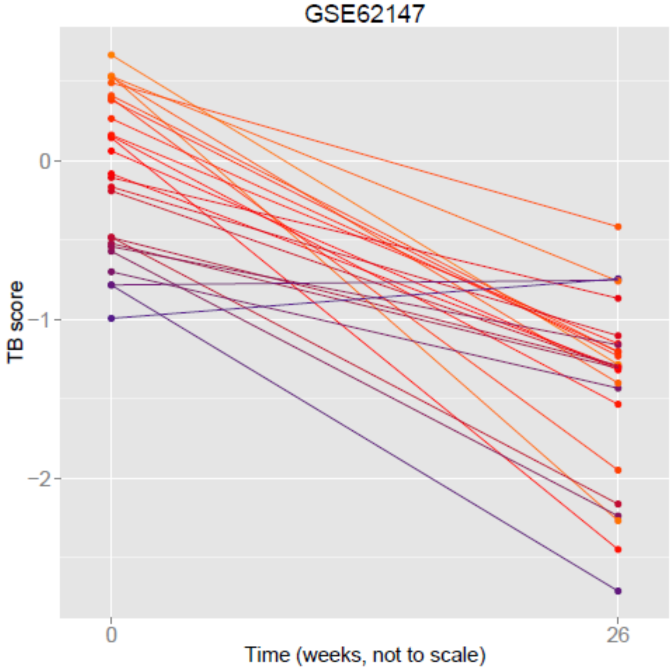
Appendix page 15. TB score in ATB patients in GSE19491 according to (A,B) sputum and (C,D) BAL smear and culture results. There are many patients overlapping between the different figures; no ATB patients had both negative sputum culture and negative BAL culture. There is no significant effect of smear or culture positivity between in any group (Wilcoxon p=NS).

Cliff Combined				
	Estimate	Std. Error	t value	p value
(Intercept)	1.691	0.082	20.698	< 2e-16
time.weeks	-0.044	0.006	-6.812	2.14E-10
Residual standard error: 0.826 on 151 degrees of freedom			Multiple R-squared: 0.2351	
GSE40553				
	Estimate	Std. Error	t value	p value
(Intercept)	-1.569	0.082	-19.11	<2e-16
time.weeks	-0.035	0.003	-10.68	<2e-16
Residual standard error: 0.7834 on 164 degrees of freedom			Multiple R-squared: 0.4101	
GSE56153				
	Estimate	Std. Error	t value	p value
(Intercept)	-1.668	0.158	-10.569	1.89E-14
time.weeks	-0.025	0.009	-2.785	7.49E-03
Residual standard error: 0.7923 on 51 degrees of freedom			Multiple R-squared: 0.132	
GSE62147				
	Estimate	Std. Error	t value	p value
(Intercept)	-0.065	0.105	-0.621	5.37E-01
time.weeks	-0.052	0.006	-9.058	4.03E-12
Residual standard error: 0.5333 on 50 degrees of freedom			Multiple R-squared: 0.6213	

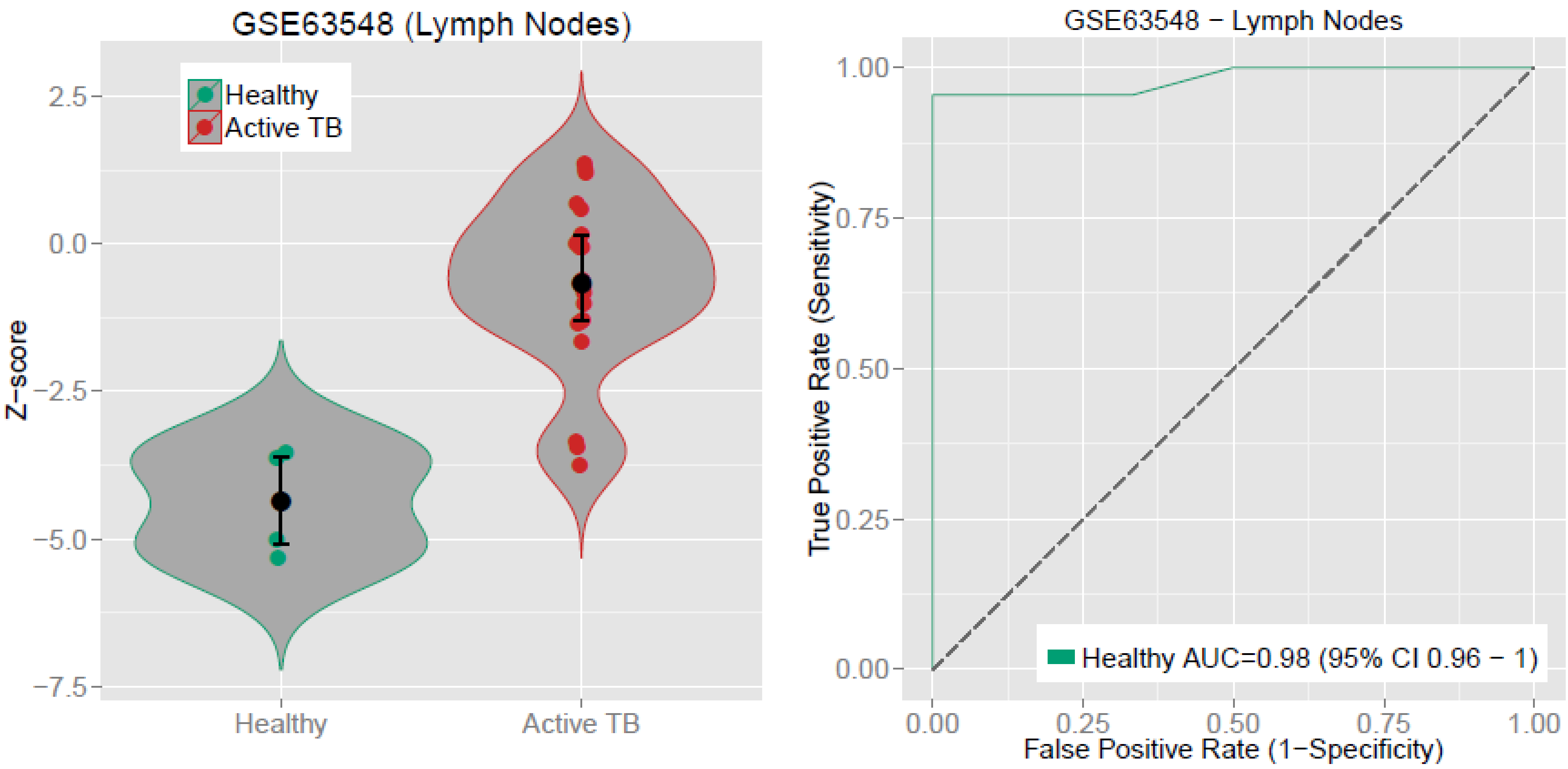
Appendix page 16. Linear regressions of TB score on treatment time in weeks. All four datasets show significant decreases over time.


Cliff Combined Datasets


GSE40553


GSE62147

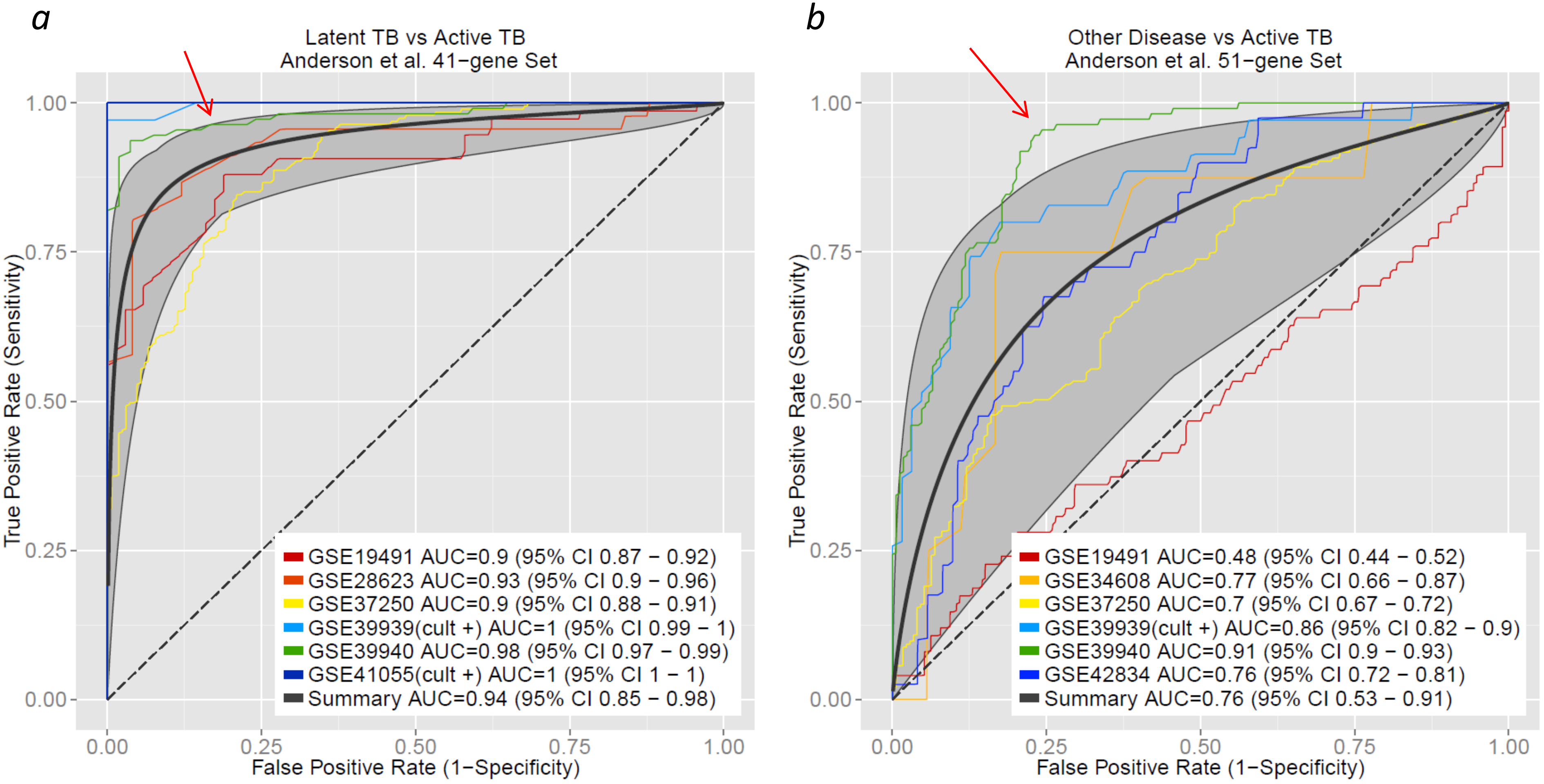
Appendix page 17. Individual ATB patient trajectories of TB score for datasets that track patients longitudinally during treatment. Each patient is colored according to TB score rank at time of diagnosis (0 weeks). Time is not drawn to scale.



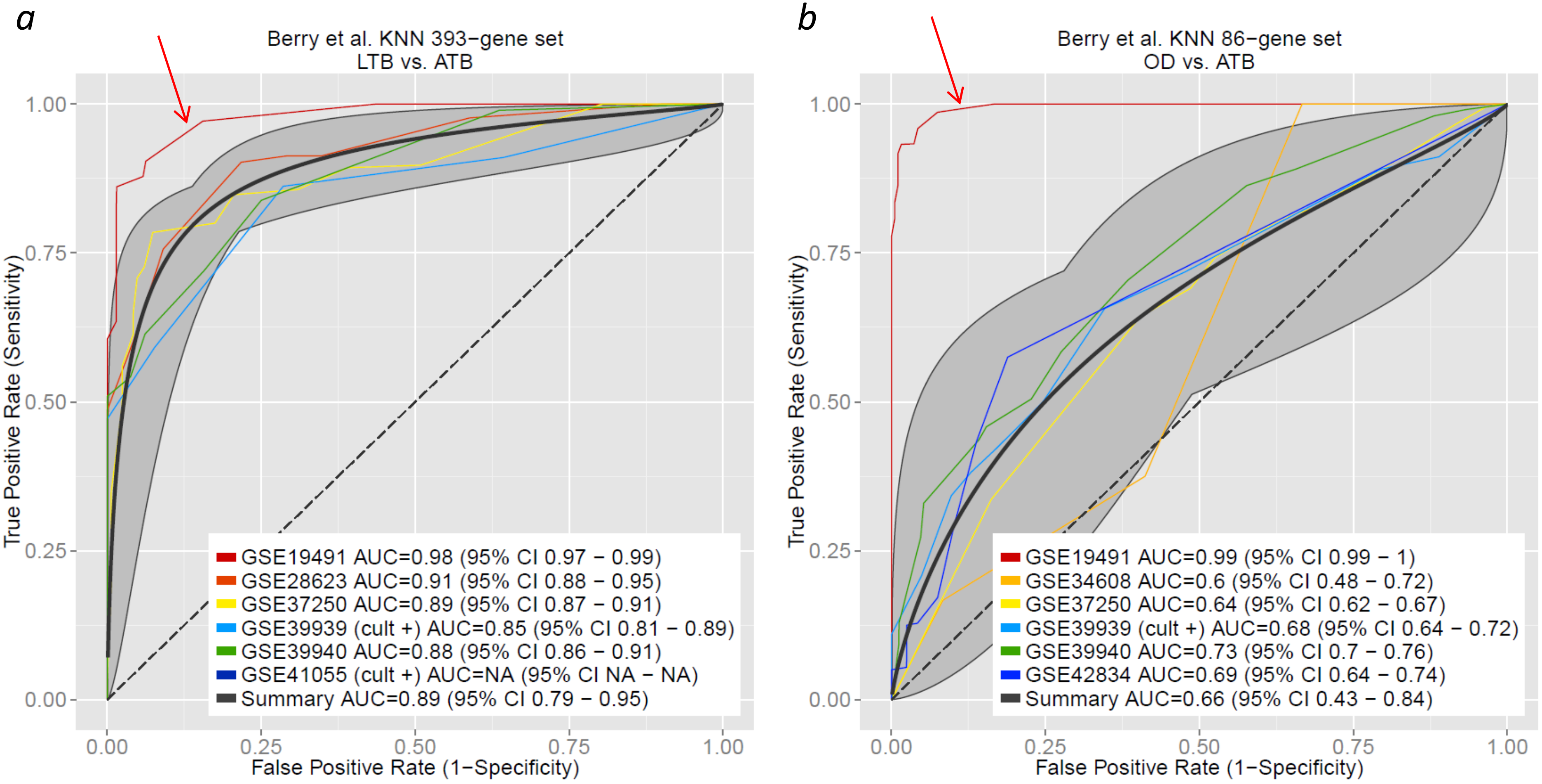
Appendix page 18. GSE63548 compared lymph node tissue between healthy controls and patients with extrapulmonary lymph node TB infections. The three gene set showed (A) a significant difference between the two groups, with (B) an ROC AUC of 0.98.

Manuscript	Discovery Dataset	Discovery Comparison & Gene Sets	Model Type
Anderson et al.	GSE39940	LTB vs ATB - 41 genes OD vs ATB - 51 genes	Difference of means
Berry et al.	GSE19491	LTB vs ATB – 393 genes OD vs ATB – 86 genes	K-nearest neighbors
Bloom et al.	GES42834	OD vs ATB – 144 transcripts	Support vector machines
Kaforou et al.	GSE37250	LTB vs ATB – 27 genes OD vs ATB – 44 genes	Difference of means
Verhagen et al.	GSE41055	HC & LTB vs ATB - 10 genes	Random Forests
Sweeney et al. (present study)	GSE19491, GSE37250, GSE42834	OD & LTB vs ATB – 3 genes	Difference of means

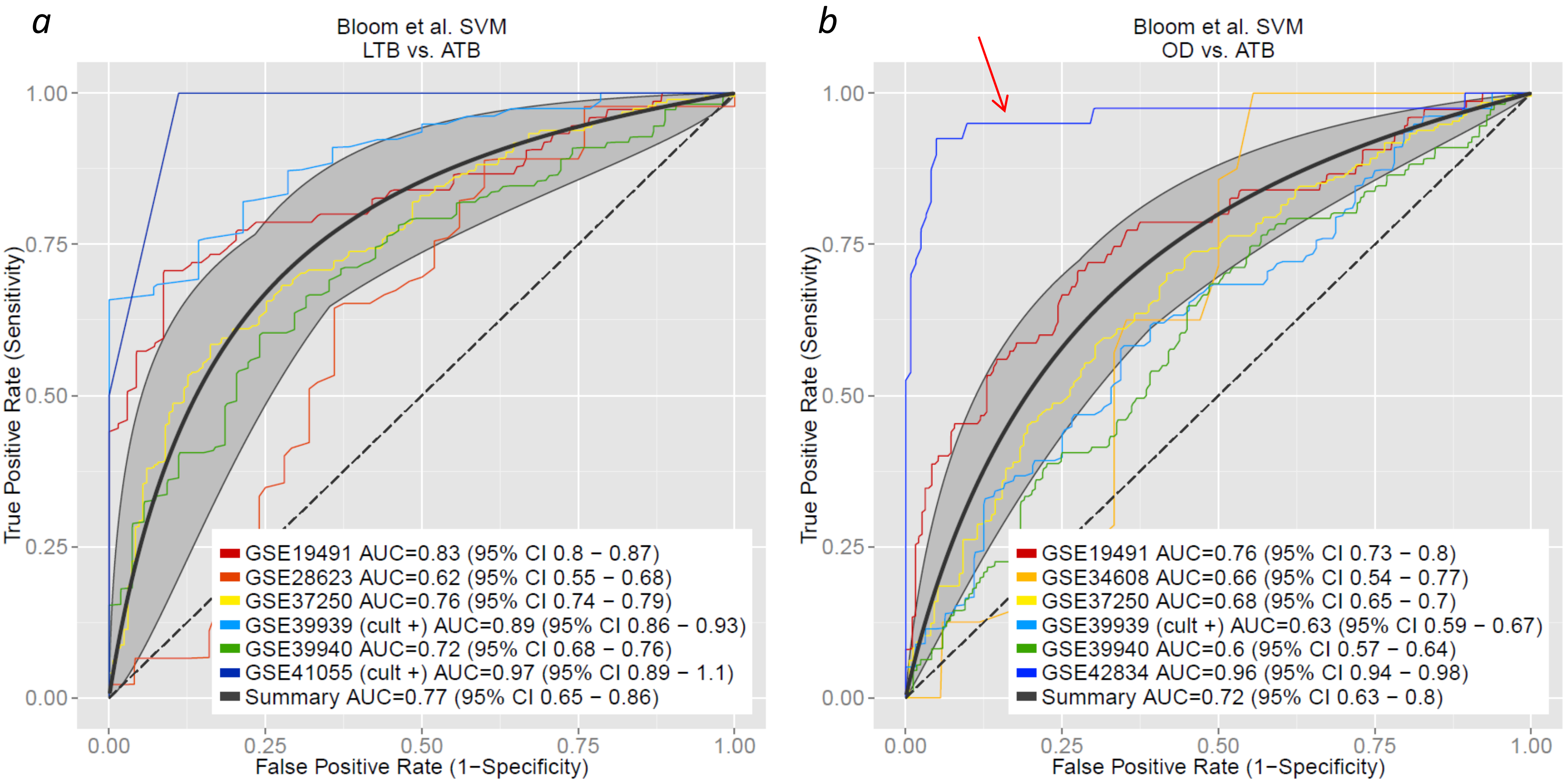
Appendix page 19. An overview of previous proposed host-response diagnostic gene sets discovered in microarray studies.



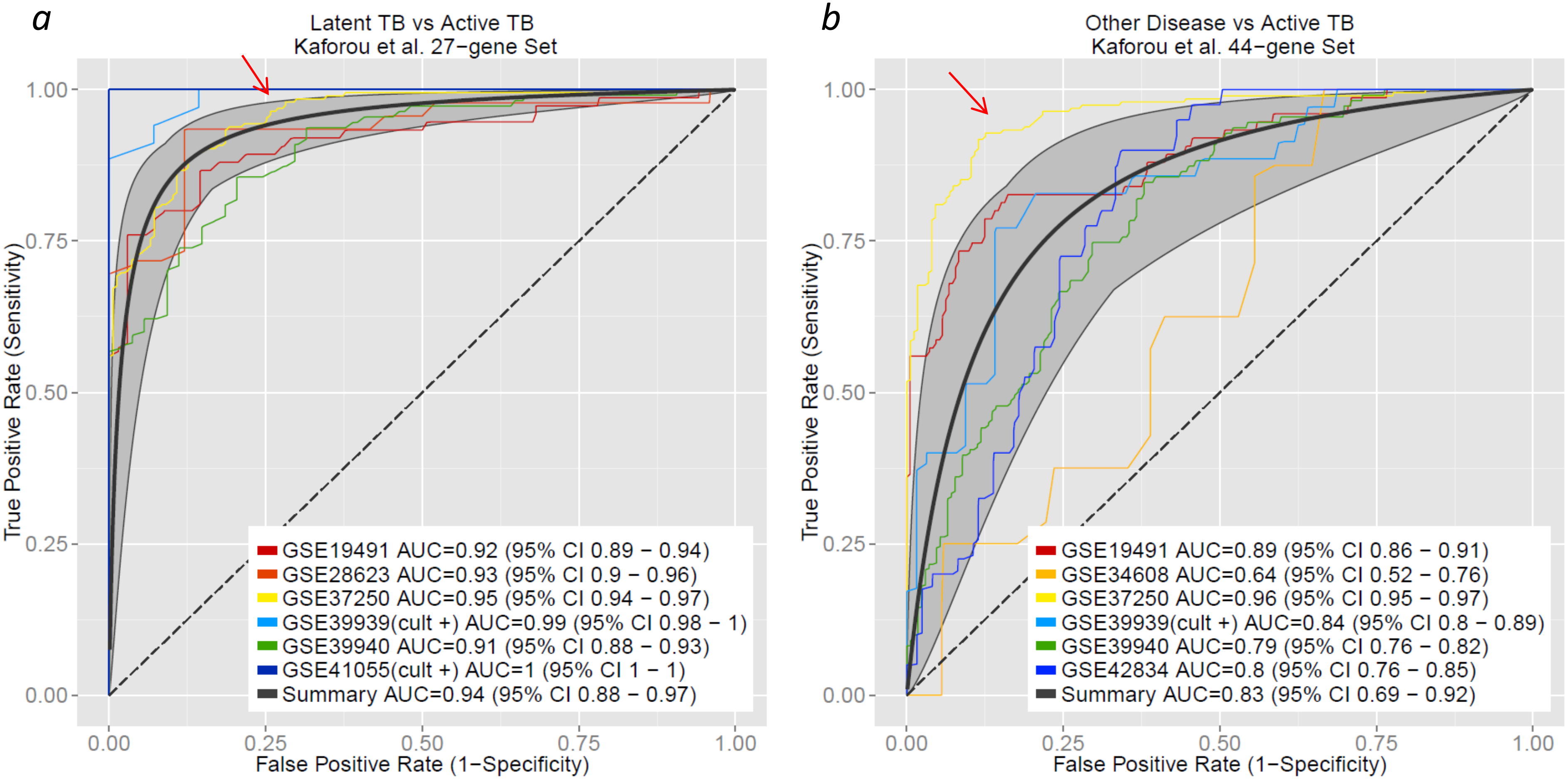
Appendix page 20. Summary ROC plots for the Anderson et al. diagnostic gene sets in all publically available TB gene expression datasets. The red arrows mark the discovery dataset (GSE39940). (a) Latent TB versus active TB; (b) other disease versus active TB. The gene sets were tested with the difference of arithmetic means as in the original paper.



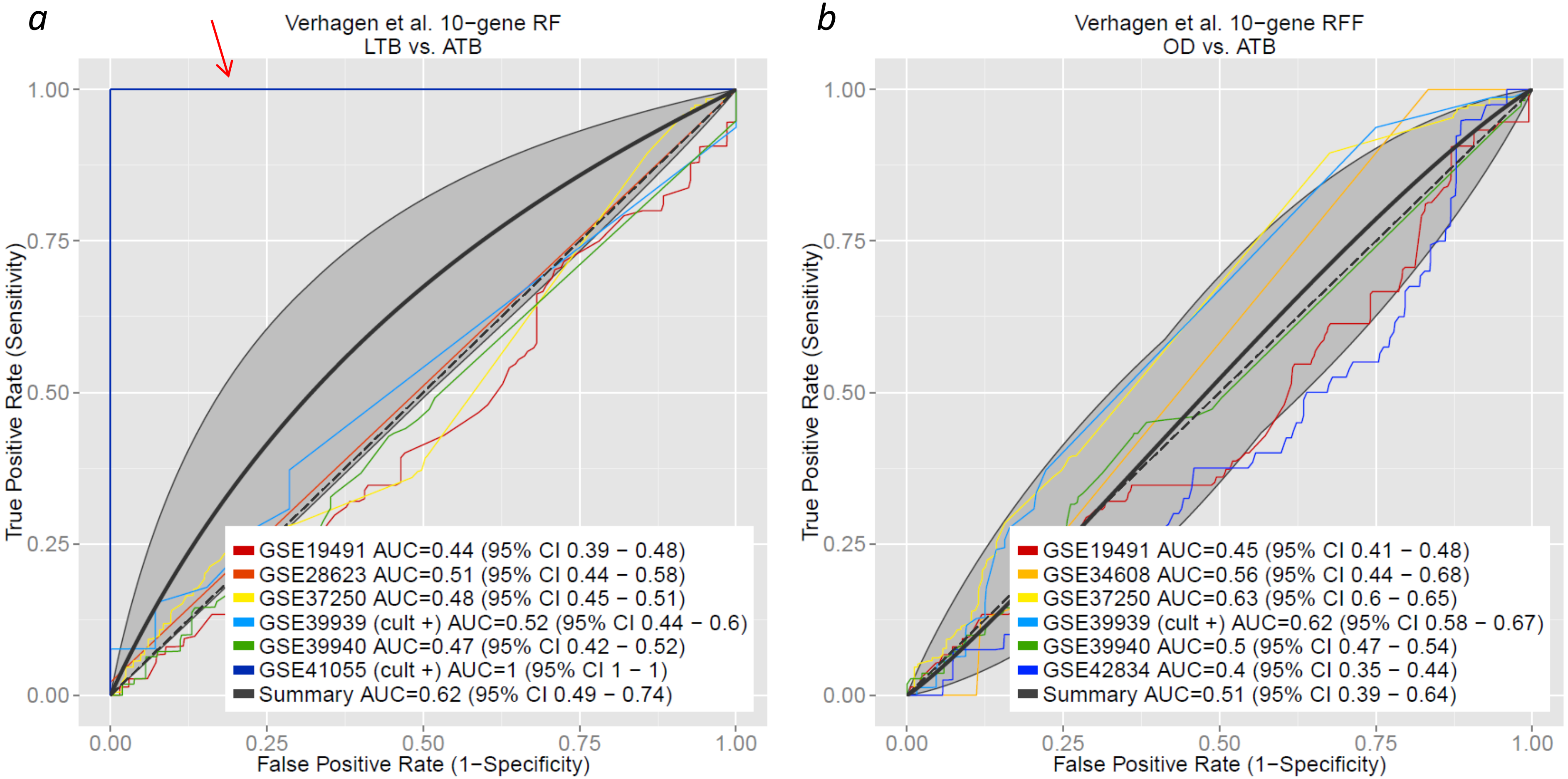
Appendix page 21. Summary ROC plots for the Berry et al. diagnostic gene set in all publically available TB gene expression datasets. The red arrow marks the discovery dataset (GSE19491 (a) Latent TB versus active TB; (b) other disease versus active TB. Each dataset was tested using a K-nearest neighbors classifier built in GSE19491, as in the original paper. ROC curves were built from vote-count thresholds. GSE41055 is listed as ‘NA’ because all votes assigned both classes as LTB, so no thresholding could be done.



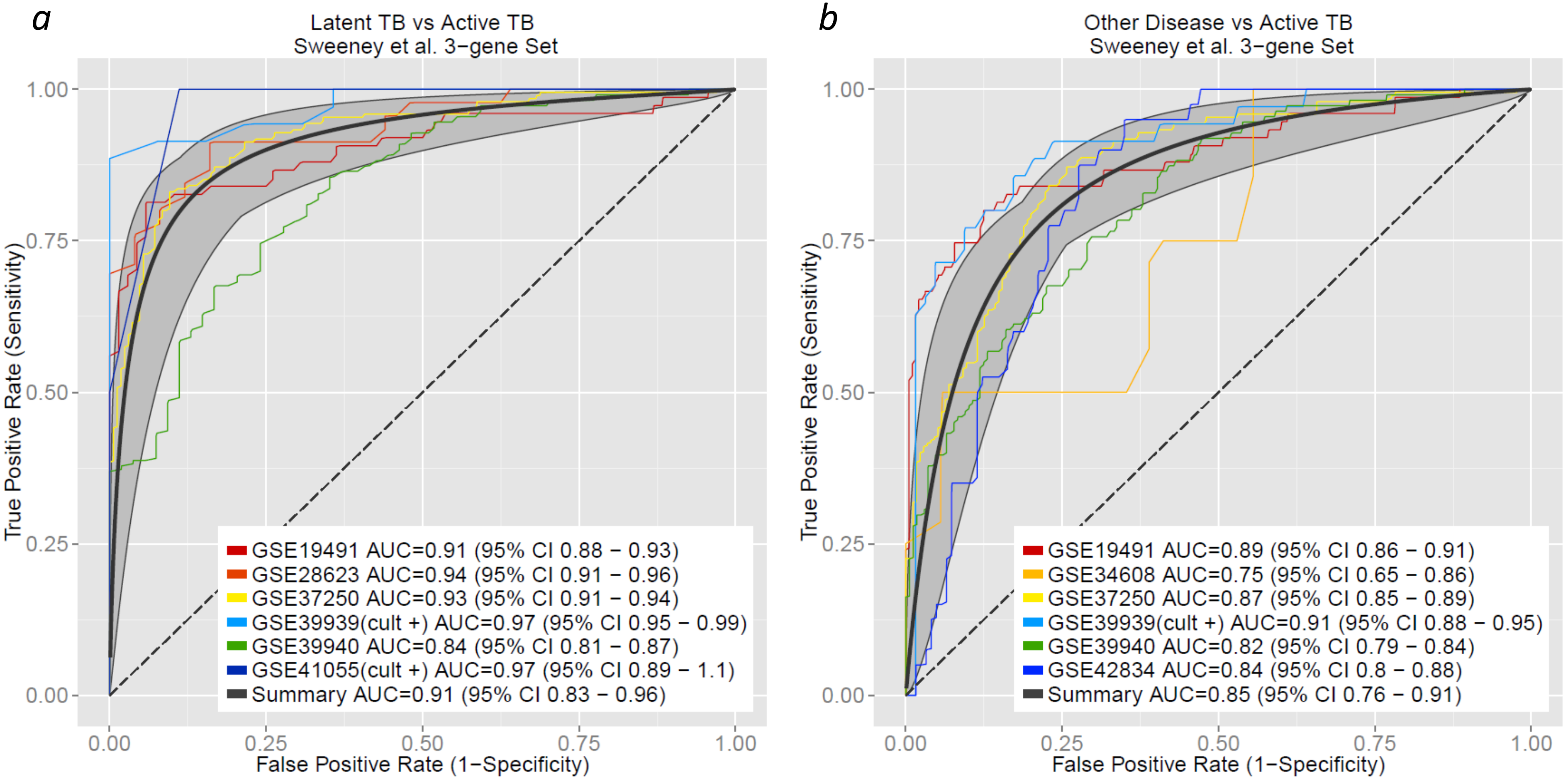
Appendix page 22. Summary ROC plots for the Bloom et al. diagnostic gene set in all publically available TB gene expression datasets. The red arrow marks the discovery dataset (GSE42834). (a) Latent TB versus active TB; (b) other disease versus active TB. Each dataset was tested using a support vector machine model built in GSE42834 using genes in the 144-transcript set, as in the original paper.



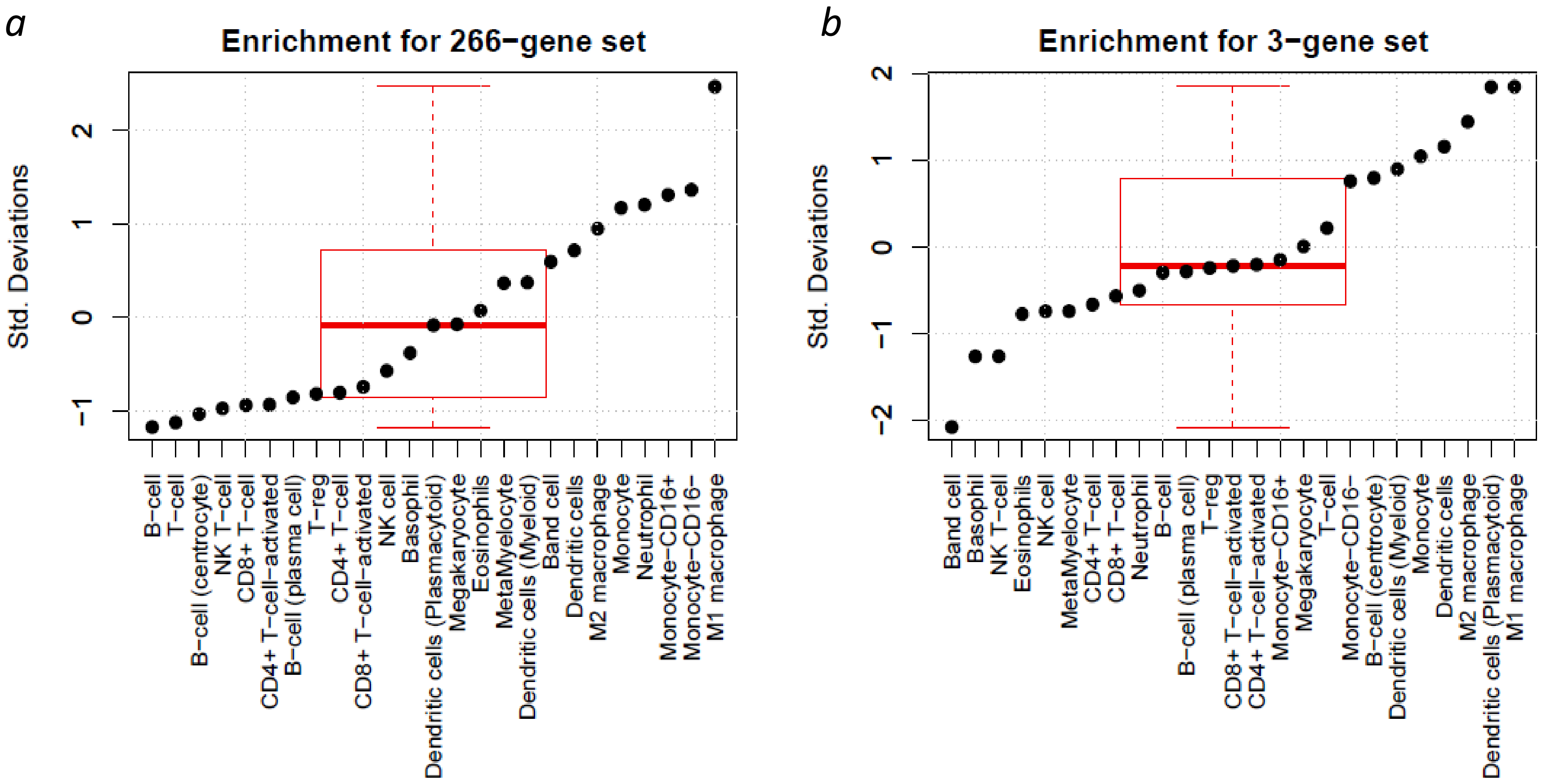
Appendix page 23. Summary ROC plots for the Kaforou et al. diagnostic gene set in all publically available TB gene expression datasets. The red arrow marks the discovery dataset (GSE37250). (a) Latent TB versus active TB; (b) other disease versus active TB. The gene sets were tested with the difference of arithmetic means in each dataset, as in the original paper.



Appendix page 24. Summary ROC plots for the Verhagen et al. diagnostic gene set in all publically available TB gene expression datasets. The red arrow marks the discovery dataset (GSE41055). (a) Latent TB versus active TB; (b) other disease versus active TB. Each dataset was tested against a random forest model built in GSE41055 using the 10-gene set, as in the original paper.



Appendix page 25. The TB score from the three-gene set (as described in the text) is shown. (a) Latent TB versus active TB; (b) other disease versus active TB. This plot is supplied to allow comparison of the generalizability of the three-gene set and method to the other gene sets and methods that have been reported previously.



Appendix page 26. Enrichment profiles of (a) all 266 differentially expressed genes and (b) the 4 diagnostic genes in publically available sorted-cell gene expression profiles. Y-axis shows standard deviations from the mean. Both gene sets are significantly enriched in M1 macrophages compared to other cell types ($p<0.05$).