

Supplementary Information (SI)

SI Materials and Methods

ROC Analysis of Predictive Power Between Experiments

ROC Analysis was performed using one of the top-performing methods from the correlation analysis (Subtractive/Centered/SampleMerge=mean). For each pair of experiments (A, B) the values of their respective disease state vectors were converted to ranks. The ability of experiment B to predict experiment A was then quantified using various rank-percentile cutoffs, starting at 5% and moving up by increments of 5%, and counting the number of up-and-down-regulated genes in experiment A predicted by experiment B at the percentile cutoff threshold. We then use trapezoidal integration to compute the raw AUC value for the pair.

Evaluating the Effect of Semantic Annotation Distance and Biological Concordance

For each pair of experiments in the data set (A, B) the UMLS concept hierarchy table (MRHIER) was queried to determine the paths from the tissue terms annotating A and B to the root of the hierarchy. The semantic distance was then calculated as the number of traversals required to reach a common parent concept between the tissue terms annotating A and B. In cases where there was more than one path to the root of the concept hierarchy, the path offering the nearest common parent concept was used.

Statistical Analyses

All statistical analyses were conducted using the R language interpreter version 2.7.2 (<http://www.R-project.org>). Spearman correlation values were converted z-scores using the Fisher transformation prior to hypothesis testing. The difference in the means of z-transformed correlations across the four disease categories was evaluated using one-way ANOVA. The pair-wise differences in z-transformed correlation distributions between the four disease/tissue categories were evaluated using Tukey's HSD to maintain the same family-wise error rate across multiple comparisons.

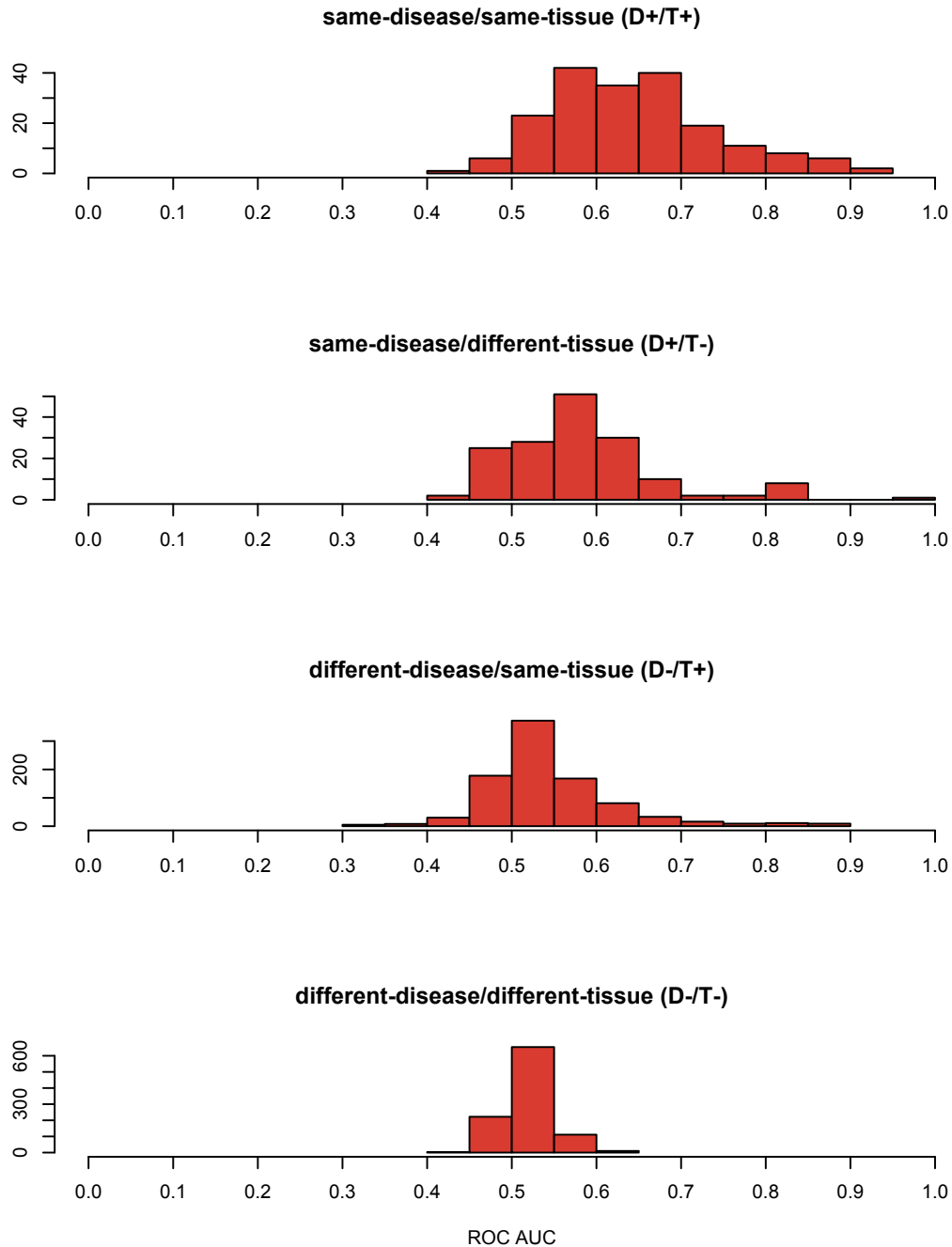
Supplementary Table S1

Normalize	Collapse Probes	Aggregate Arrays	Disease vs. Control	D+/T+ $\bar{x}(rs)$	D+/T+ $\bar{x}(p)$	D+/T- $\bar{x}(rs)$	D+/T- $\bar{x}(p)$	D-/T+ $\bar{x}(rs)$	D-/T+ $\bar{x}(p)$	D-/T- $\bar{x}(rs)$	D-/T- $\bar{x}(p)$	ANOVA p-value
CenterNorm	NoCollapse	NoAggregate	SubtractiveDiff	2.19E-01	8.03E-22	5.63E-02	1.62E-09	4.29E-02	1.01E-07	6.60E-03	5.84E-03	0.00E+00
CenterNorm	MeanCollapse	MeanAggregate	SubtractiveDiff	2.19E-01	8.03E-22	5.58E-02	1.62E-09	4.26E-02	1.12E-07	6.52E-03	5.79E-03	0.00E+00
CenterNorm	NoCollapse	NoAggregate	SubtractiveDiff	2.19E-01	8.03E-22	5.63E-02	1.62E-09	4.29E-02	1.01E-07	6.60E-03	5.84E-03	0.00E+00
CenterNorm	MeanCollapse	NoAggregate	SubtractiveDiff	2.19E-01	8.03E-22	5.61E-02	1.62E-09	4.29E-02	1.20E-07	6.60E-03	5.83E-03	0.00E+00
CenterNorm	MeanCollapse	NoAggregate	SubtractiveDiff	2.19E-01	8.03E-22	5.61E-02	1.62E-09	4.29E-02	1.20E-07	6.60E-03	5.83E-03	0.00E+00
CenterNorm	NoCollapse	MeanAggregate	SubtractiveDiff	2.19E-01	8.03E-22	5.58E-02	1.62E-09	4.26E-02	1.12E-07	6.51E-03	5.79E-03	0.00E+00
NoNorm	MeanCollapse	NoAggregate	SubtractiveDiff	2.17E-01	8.45E-24	5.98E-02	1.05E-07	4.24E-02	2.86E-07	5.96E-03	9.15E-03	0.00E+00
NoNorm	NoCollapse	NoAggregate	SubtractiveDiff	2.17E-01	8.45E-24	5.98E-02	1.05E-07	4.27E-02	2.86E-07	6.00E-03	9.22E-03	0.00E+00
NoNorm	NoCollapse	MeanAggregate	SubtractiveDiff	2.17E-01	6.65E-24	5.87E-02	1.05E-07	4.23E-02	3.44E-07	5.96E-03	9.09E-03	0.00E+00
NoNorm	MeanCollapse	NoAggregate	SubtractiveDiff	2.17E-01	8.45E-24	5.98E-02	1.05E-07	4.24E-02	2.86E-07	5.96E-03	9.15E-03	0.00E+00
NoNorm	MeanCollapse	MeanAggregate	SubtractiveDiff	2.17E-01	7.20E-24	5.87E-02	1.05E-07	4.23E-02	3.89E-07	5.98E-03	9.15E-03	0.00E+00
NoNorm	NoCollapse	NoAggregate	SubtractiveDiff	2.17E-01	8.45E-24	5.98E-02	1.05E-07	4.27E-02	2.86E-07	6.00E-03	9.22E-03	0.00E+00
RankNorm	NoCollapse	MeanAggregate	SubtractiveDiff	1.98E-01	0.00E+00	4.72E-02	4.36E-05	3.50E-02	1.90E-04	6.01E-03	5.14E-02	0.00E+00
RankNorm	MeanCollapse	MeanAggregate	SubtractiveDiff	1.95E-01	0.00E+00	4.72E-02	4.68E-05	3.50E-02	1.90E-04	6.03E-03	5.15E-02	0.00E+00
RankNorm	MeanCollapse	NoAggregate	SubtractiveDiff	1.95E-01	0.00E+00	4.56E-02	5.68E-05	3.63E-02	1.90E-04	5.97E-03	5.12E-02	0.00E+00
RankNorm	MeanCollapse	NoAggregate	SubtractiveDiff	1.95E-01	0.00E+00	4.56E-02	5.68E-05	3.63E-02	1.90E-04	5.97E-03	5.12E-02	0.00E+00
RankNorm	NoCollapse	NoAggregate	SubtractiveDiff	1.95E-01	0.00E+00	4.60E-02	5.68E-05	3.63E-02	1.90E-04	5.97E-03	5.10E-02	0.00E+00
RankNorm	NoCollapse	NoAggregate	SubtractiveDiff	1.95E-01	0.00E+00	4.60E-02	5.68E-05	3.63E-02	1.90E-04	5.97E-03	5.10E-02	0.00E+00
CenterNorm	MeanCollapse	MedianAggregate	SubtractiveDiff	1.90E-01	1.45E-15	5.58E-02	6.06E-08	3.38E-02	5.70E-06	5.87E-03	1.15E-02	0.00E+00
NoNorm	MeanCollapse	MedianAggregate	SubtractiveDiff	1.85E-01	1.59E-16	4.70E-02	1.72E-06	3.66E-02	9.52E-06	5.33E-03	1.74E-02	0.00E+00
CenterNorm	NoCollapse	MedianAggregate	SubtractiveDiff	1.79E-01	5.20E-13	4.60E-02	3.22E-06	3.16E-02	9.52E-06	5.42E-03	1.87E-02	0.00E+00
NoNorm	NoCollapse	MedianAggregate	SubtractiveDiff	1.66E-01	1.09E-14	4.76E-02	5.59E-07	3.54E-02	6.30E-05	5.07E-03	2.60E-02	0.00E+00
RankNorm	MeanCollapse	MedianAggregate	SubtractiveDiff	1.65E-01	5.33E-15	4.51E-02	6.39E-05	2.55E-02	1.95E-03	5.25E-03	8.13E-02	0.00E+00
CenterNorm	MeanCollapse	MeanAggregate	FoldDiff	1.60E-01	4.09E-14	2.65E-02	5.39E-03	3.18E-02	1.76E-04	4.70E-03	5.58E-02	0.00E+00
CenterNorm	NoCollapse	MeanAggregate	FoldDiff	1.60E-01	4.09E-14	2.65E-02	5.39E-03	3.16E-02	1.89E-04	4.69E-03	5.52E-02	0.00E+00
NoNorm	NoCollapse	MeanAggregate	FoldDiff	1.59E-01	1.64E-14	2.54E-02	5.44E-03	3.05E-02	2.00E-04	4.72E-03	5.31E-02	0.00E+00
NoNorm	MeanCollapse	MeanAggregate	FoldDiff	1.59E-01	2.38E-14	2.54E-02	5.44E-03	2.96E-02	1.89E-04	4.60E-03	5.34E-02	0.00E+00
CenterNorm	MeanCollapse	NoAggregate	FoldDiff	1.58E-01	2.27E-13	2.43E-02	4.13E-03	3.23E-02	1.71E-04	4.30E-03	5.97E-02	0.00E+00
CenterNorm	MeanCollapse	NoAggregate	FoldDiff	1.58E-01	2.27E-13	2.43E-02	4.13E-03	3.23E-02	1.71E-04	4.30E-03	5.97E-02	0.00E+00
RankNorm	NoCollapse	MedianAggregate	SubtractiveDiff	1.56E-01	4.67E-12	4.10E-02	9.58E-04	2.46E-02	4.62E-03	5.17E-03	9.58E-02	0.00E+00
NoNorm	MeanCollapse	NoAggregate	FoldDiff	1.54E-01	1.98E-13	1.95E-02	1.03E-02	3.00E-02	1.30E-04	4.38E-03	5.33E-02	0.00E+00
NoNorm	MeanCollapse	NoAggregate	FoldDiff	1.54E-01	1.98E-13	1.95E-02	1.03E-02	3.00E-02	1.30E-04	4.38E-03	5.33E-02	0.00E+00
NoNorm	NoCollapse	NoAggregate	FoldDiff	1.52E-01	2.99E-12	1.79E-02	1.26E-02	2.91E-02	5.37E-04	4.13E-03	7.07E-02	0.00E+00
NoNorm	NoCollapse	NoAggregate	FoldDiff	1.52E-01	2.99E-12	1.79E-02	1.26E-02	2.91E-02	5.37E-04	4.13E-03	7.07E-02	0.00E+00
CenterNorm	NoCollapse	NoAggregate	FoldDiff	1.52E-01	7.08E-12	2.12E-02	1.15E-02	3.15E-02	4.49E-04	4.01E-03	7.41E-02	0.00E+00
CenterNorm	NoCollapse	NoAggregate	FoldDiff	1.52E-01	7.08E-12	2.12E-02	1.15E-02	3.15E-02	4.49E-04	4.01E-03	7.41E-02	0.00E+00
NoNorm	MeanCollapse	MedianAggregate	FoldDiff	1.52E-01	1.37E-11	2.27E-02	3.91E-03	2.38E-02	2.24E-03	4.29E-03	7.27E-02	0.00E+00
CenterNorm	MeanCollapse	MedianAggregate	FoldDiff	1.50E-01	4.98E-11	2.47E-02	2.65E-03	2.35E-02	2.48E-03	4.08E-03	7.62E-02	0.00E+00
NoNorm	MeanCollapse	MaxAggregate	SubtractiveDiff	1.41E-01	1.99E-10	3.71E-02	1.17E-03	2.84E-02	4.72E-04	6.81E-03	3.43E-02	5.35E-319

CenterNorm	NoCollapse	MedianAggregate	FoldDiff	1.36E-01	8.12E-10	1.94E-02	1.77E-02	2.29E-02	5.37E-03	3.99E-03	9.76E-02	0.00E+00
NoNorm	NoCollapse	MaxAggregate	SubtractiveDiff	1.36E-01	5.10E-09	3.85E-02	2.72E-03	2.66E-02	7.91E-04	6.38E-03	4.29E-02	2.53E-307
NoNorm	NoCollapse	MedianAggregate	FoldDiff	1.34E-01	2.63E-10	2.11E-02	8.57E-03	2.26E-02	3.77E-03	4.03E-03	9.16E-02	0.00E+00
CenterNorm	MeanCollapse	MaxAggregate	SubtractiveDiff	1.32E-01	7.50E-10	3.49E-02	1.14E-03	2.57E-02	1.15E-04	6.25E-03	2.62E-02	1.14E-271
CenterNorm	NoCollapse	MaxAggregate	SubtractiveDiff	1.27E-01	2.56E-09	3.31E-02	1.88E-03	2.28E-02	2.84E-04	5.98E-03	3.40E-02	2.90E-262
CenterNorm	MeanCollapse	MaxAggregate	FoldDiff	1.21E-01	3.83E-07	2.36E-02	1.11E-02	2.24E-02	3.09E-03	4.99E-03	9.22E-02	0.00E+00
NoNorm	MeanCollapse	MaxAggregate	FoldDiff	1.17E-01	6.47E-08	2.23E-02	1.52E-02	2.24E-02	2.71E-03	4.59E-03	8.71E-02	0.00E+00
NoNorm	NoCollapse	MaxAggregate	FoldDiff	1.15E-01	1.51E-07	2.48E-02	1.24E-02	2.08E-02	3.85E-03	4.36E-03	1.01E-01	0.00E+00
CenterNorm	MeanCollapse	MinAggregate	SubtractiveDiff	1.15E-01	4.79E-07	3.23E-02	1.02E-03	3.02E-02	5.55E-05	3.64E-03	2.97E-02	8.10E-233
CenterNorm	NoCollapse	MaxAggregate	FoldDiff	1.13E-01	7.75E-07	2.70E-02	1.03E-02	2.27E-02	5.69E-03	4.69E-03	1.07E-01	0.00E+00
RankNorm	MeanCollapse	MaxAggregate	SubtractiveDiff	1.05E-01	4.64E-06	3.45E-02	9.79E-03	2.25E-02	1.58E-02	4.85E-03	1.28E-01	0.00E+00
RankNorm	MeanCollapse	MeanAggregate	FoldDiff	1.03E-01	2.21E-06	2.93E-02	8.33E-03	1.92E-02	1.81E-02	4.14E-03	1.33E-01	0.00E+00
RankNorm	NoCollapse	MeanAggregate	FoldDiff	1.03E-01	2.21E-06	2.92E-02	8.33E-03	1.92E-02	1.78E-02	4.15E-03	1.33E-01	0.00E+00
RankNorm	NoCollapse	MaxAggregate	SubtractiveDiff	1.02E-01	2.42E-05	3.13E-02	2.04E-02	2.18E-02	1.70E-02	4.24E-03	1.31E-01	0.00E+00
NoNorm	MeanCollapse	MinAggregate	SubtractiveDiff	9.91E-02	4.38E-05	3.14E-02	2.19E-03	2.11E-02	5.24E-04	2.89E-03	3.69E-02	3.61E-204
RankNorm	MeanCollapse	NoAggregate	FoldDiff	9.45E-02	6.53E-06	2.80E-02	6.52E-03	1.75E-02	2.42E-02	3.54E-03	1.47E-01	0.00E+00
RankNorm	MeanCollapse	NoAggregate	FoldDiff	9.45E-02	6.53E-06	2.80E-02	6.52E-03	1.75E-02	2.42E-02	3.54E-03	1.47E-01	0.00E+00
CenterNorm	NoCollapse	MinAggregate	SubtractiveDiff	9.41E-02	5.14E-05	2.46E-02	1.41E-02	2.44E-02	1.67E-03	3.19E-03	6.56E-02	8.79E-210
RankNorm	NoCollapse	NoAggregate	FoldDiff	9.32E-02	1.38E-05	2.45E-02	9.53E-03	1.80E-02	2.75E-02	3.43E-03	1.58E-01	0.00E+00
RankNorm	NoCollapse	NoAggregate	FoldDiff	9.32E-02	1.38E-05	2.45E-02	9.53E-03	1.80E-02	2.75E-02	3.43E-03	1.58E-01	0.00E+00
RankNorm	MeanCollapse	MinAggregate	SubtractiveDiff	8.37E-02	3.99E-05	2.36E-02	2.36E-02	1.88E-02	3.50E-02	2.90E-03	1.60E-01	0.00E+00
RankNorm	MeanCollapse	MedianAggregate	FoldDiff	7.88E-02	6.69E-04	2.26E-02	2.68E-02	1.56E-02	4.75E-02	3.45E-03	1.75E-01	0.00E+00
NoNorm	NoCollapse	MinAggregate	SubtractiveDiff	7.80E-02	3.50E-04	2.27E-02	5.95E-03	1.78E-02	5.45E-03	2.55E-03	7.38E-02	1.08E-189
RankNorm	NoCollapse	MinAggregate	SubtractiveDiff	7.39E-02	4.69E-04	1.86E-02	3.49E-02	1.56E-02	6.19E-02	2.55E-03	2.01E-01	0.00E+00
CenterNorm	MeanCollapse	MinAggregate	FoldDiff	7.31E-02	7.05E-04	1.73E-02	2.16E-02	1.93E-02	1.56E-02	2.21E-03	1.28E-01	1.88E-246
NoNorm	MeanCollapse	MinAggregate	FoldDiff	7.21E-02	1.47E-03	1.70E-02	2.61E-02	1.37E-02	2.08E-02	2.19E-03	1.18E-01	2.72E-187
RankNorm	NoCollapse	MedianAggregate	FoldDiff	7.19E-02	1.90E-03	2.44E-02	2.95E-02	1.57E-02	6.68E-02	3.49E-03	1.89E-01	0.00E+00
RankNorm	MeanCollapse	MaxAggregate	FoldDiff	6.17E-02	7.92E-03	1.68E-02	1.19E-01	1.23E-02	4.83E-02	3.11E-03	1.57E-01	2.34E-224
RankNorm	NoCollapse	MaxAggregate	FoldDiff	5.72E-02	1.63E-02	1.51E-02	1.24E-01	1.12E-02	6.48E-02	2.81E-03	1.64E-01	1.61E-195
NoNorm	NoCollapse	MinAggregate	FoldDiff	5.18E-02	8.94E-03	1.03E-02	6.50E-02	1.07E-02	5.54E-02	1.51E-03	1.85E-01	1.33E-164
CenterNorm	NoCollapse	MinAggregate	FoldDiff	5.01E-02	9.46E-03	1.03E-02	9.47E-02	1.35E-02	6.26E-02	1.46E-03	1.95E-01	2.47E-210
RankNorm	MeanCollapse	MinAggregate	FoldDiff	4.30E-02	5.88E-02	1.19E-02	1.13E-01	7.11E-03	1.47E-01	2.43E-03	2.18E-01	6.83E-165
CenterNorm	NoCollapse	NoAggregate	TtestDiff	3.83E-02	6.69E-02	1.07E-02	1.30E-01	8.13E-03	1.48E-01	1.14E-03	2.76E-01	2.74E-178
CenterNorm	NoCollapse	NoAggregate	TtestDiff	3.83E-02	6.69E-02	1.07E-02	1.30E-01	8.13E-03	1.48E-01	1.14E-03	2.76E-01	2.74E-178
RankNorm	NoCollapse	MinAggregate	FoldDiff	3.82E-02	7.13E-02	1.19E-02	1.90E-01	6.66E-03	1.94E-01	1.85E-03	2.83E-01	4.69E-159
NoNorm	NoCollapse	NoAggregate	TtestDiff	3.58E-02	7.05E-02	1.22E-02	1.68E-01	6.62E-03	1.74E-01	6.08E-04	2.85E-01	1.33E-154
NoNorm	NoCollapse	NoAggregate	TtestDiff	3.58E-02	7.05E-02	1.22E-02	1.68E-01	6.62E-03	1.74E-01	6.08E-04	2.85E-01	1.33E-154
CenterNorm	MeanCollapse	NoAggregate	TtestDiff	3.21E-02	9.79E-02	1.06E-02	3.74E-01	7.05E-03	1.66E-01	1.33E-03	2.94E-01	3.11E-103
CenterNorm	MeanCollapse	NoAggregate	TtestDiff	3.21E-02	9.79E-02	1.06E-02	3.74E-01	7.05E-03	1.66E-01	1.33E-03	2.94E-01	3.11E-103
NoNorm	MeanCollapse	NoAggregate	TtestDiff	3.07E-02	8.45E-02	1.32E-02	1.33E-01	5.31E-03	1.56E-01	5.89E-04	2.96E-01	9.67E-80
NoNorm	MeanCollapse	NoAggregate	TtestDiff	3.07E-02	8.45E-02	1.32E-02	1.33E-01	5.31E-03	1.56E-01	5.89E-04	2.96E-01	9.67E-80
RankNorm	NoCollapse	NoAggregate	TtestDiff	2.70E-02	1.90E-01	5.25E-03	3.07E-01	4.86E-03	3.21E-01	5.11E-04	3.81E-01	8.39E-105
RankNorm	NoCollapse	NoAggregate	TtestDiff	2.70E-02	1.90E-01	5.25E-03	3.07E-01	4.86E-03	3.21E-01	5.11E-04	3.81E-01	8.39E-105

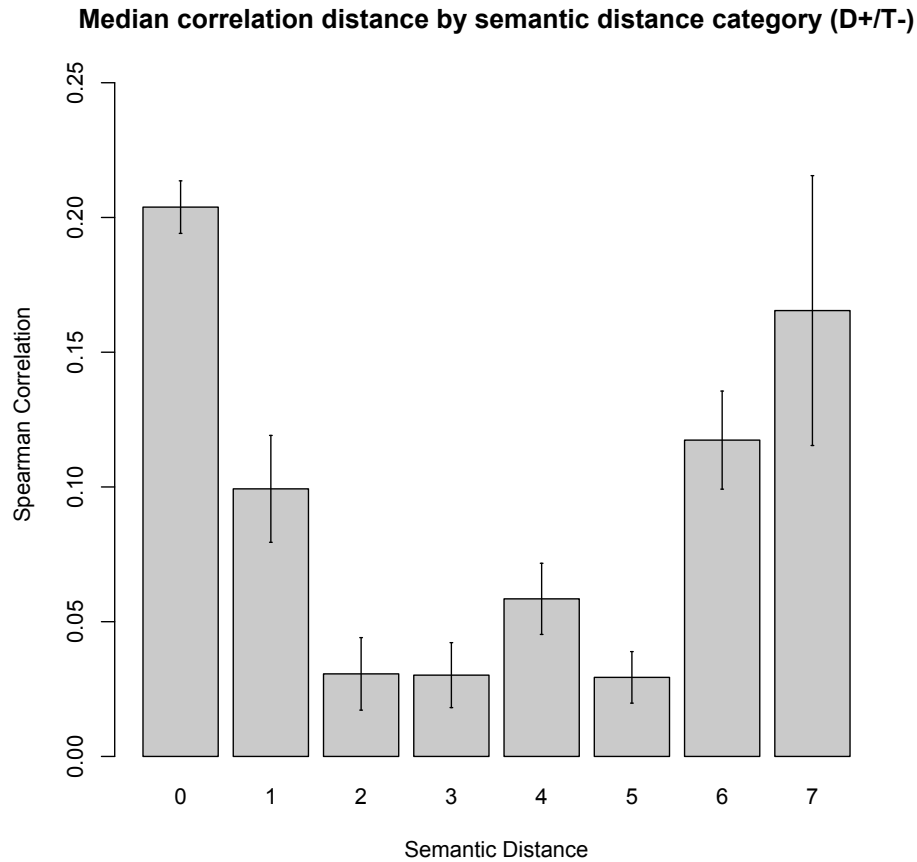
RankNorm	MeanCollapse	NoAggregate	TtestDiff	2.21E-02	2.15E-01	8.65E-04	3.97E-01	3.81E-03	3.73E-01	3.02E-04	4.35E-01	1.30E-35
RankNorm	MeanCollapse	NoAggregate	TtestDiff	2.21E-02	2.15E-01	8.65E-04	3.97E-01	3.81E-03	3.73E-01	3.02E-04	4.35E-01	1.30E-35

Figure S1



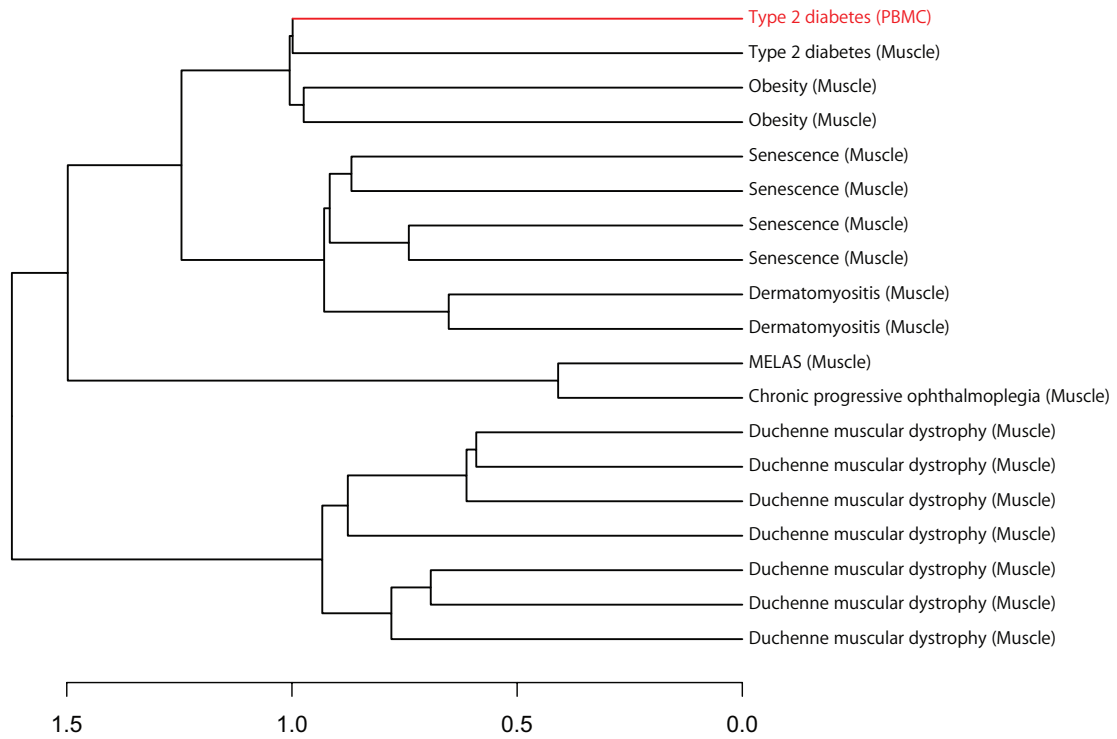
Distributions of raw Receiver Operating Characteristic (ROC) Area Under Curve (AUC) values within each disease/tissue category. For each pair of experiments A and B, the AUC was computed based on the ability of experiment B to predict the percentile of genes in experiment A. Biologically mismatched experiments (D-/T-) performed similar to a random classifier, however D+/T+, D+/T- and D-/T+ performed significantly better than random (t-test p-values $< 2.2 \times 10^{-16}$). We also observe that the distribution of the AUC values for the D+/T- category is significantly differentiated above that of the D-/T+ category (t-test p-value 9.35×10^{-07}).

Figure S2



Mean Spearman correlation for each semantic distance category between tissue terms. The semantic distance between tissue terms is computed as the minimum number of traversals up the ontology hierarchy required to reach a common parent term. As expected, experiments measuring the same tissue (Semantic Distance = 0) exhibit the highest mean correlation. Importantly, the semantic distance does not appear to have a bearing on the correlation value between dissimilar tissues, and therefore we can assert that the annotation scheme does not confound interpretation of the results.

Figure S3



Hierarchical clustering of multiple microarray experiments measuring various disease states in skeletal muscle integrated with an experiment from peripheral blood. The experiments are clustered using the correlation distance of the full disease-state transcriptomes. The diseases cluster consistently within-tissue according to the disease state. Integration of the Type 2 diabetes experiment (GSE9006) measured in peripheral blood (red) results in the clustering of both Type 2 diabetes experiments across-tissue.