Supplementary Methods

Discordant

Initial Parameters

The parameters used to separate the class components are denoted by b in Figure 1. Group v=1,2 has a unique parameter b_v to allow for different mixture model variances for each group. The parameter b_v is the standard deviation of the Fisher transformed z scores. Observations between $-b_v$ and b_v are set to component 0, observations to the left of $-b_v$ are set to component – and observations to the right of b_v are set to component +. Based on these assignments, the mean and variance of the observations in each component were used to determine the initial parameters $\mu_0, \mu_1, \mu_2, \sigma_0, \sigma_1, \sigma_2$ for group 1, $\eta_1, \eta_2, \tau_0, \tau_1, \tau_2$ for group 2. The EM algorithm (Dempster et al., 1977) is then used to iteratively update parameters in the M-step and the posterior probability for each class and group in the E-step (Figure 1).

Adjustments were made from the original Lai, et. al algorithm. In Lai et al., z scores derived from differential expression in microarray experiments have a $N(0,1)$ distribution, whereas the Fisher-transformed z scores of correlation coefficients tend to have a smaller standard deviation. It was determined using a grid approach that having b set to the standard deviation of the z scores improved convergence of the EM algorithm and resulted in a model with a better fit than when b was set to 1.

Estimation of Parameters in M-step

Each parameter contained in θ and weights π are estimated in a similar way. The symbols i and j are classes -, 0 and + for groups 1 and 2 respectively.

For parameters in group 1:

$$\pi_i = \frac{\sum_{k=1}^{K}\sum_{j=0}^{2} q_{ij}^{r}(k)}{K}$$

$$\mu_i^r = \frac{\sum_{k=1}^{K}\sum_{j=0}^{2} q_{ij}^{r}(k)\cdot z^{(1)}(k)}{\sum_{k=1}^{K}\sum_{j=0}^{2} q_{ij}^{r}(k)}$$

$$\sigma_i^r = \frac{\sum_{k=1}^{K}\sum_{j=0}^{2} q_{ij}^{r}(k)\cdot \left(z^{(1)}-\mu_i^r\right)^2}{\sum_{k=1}^{K}\sum_{j=0}^{2} q_{ij}^{r}(k)}$$

For parameters in group 2:

$$\pi_j = \frac{\sum_{k=1}^{K}\sum_{i=0}^{2} q_{ij}^{r}(k)}{K}$$

$$\eta_j^r = \frac{\sum_{k=1}^{K}\sum_{j=0}^{2} q_{ij}^{r}(k)\cdot z^{(2)}(k)}{\sum_{k=1}^{K}\sum_{i=0}^{2} q_{ij}^{r}(k)}$$

$$\tau_j^r = \frac{\sum_{k=1}^{K}\sum_{j=0}^{2} q_{ij}^{r}(k)\cdot \left(z^{(2)}-\mu_j^r\right)^2}{\sum_{k=1}^{K}\sum_{i=0}^{2} q_{ij}^{r}(k)}$$

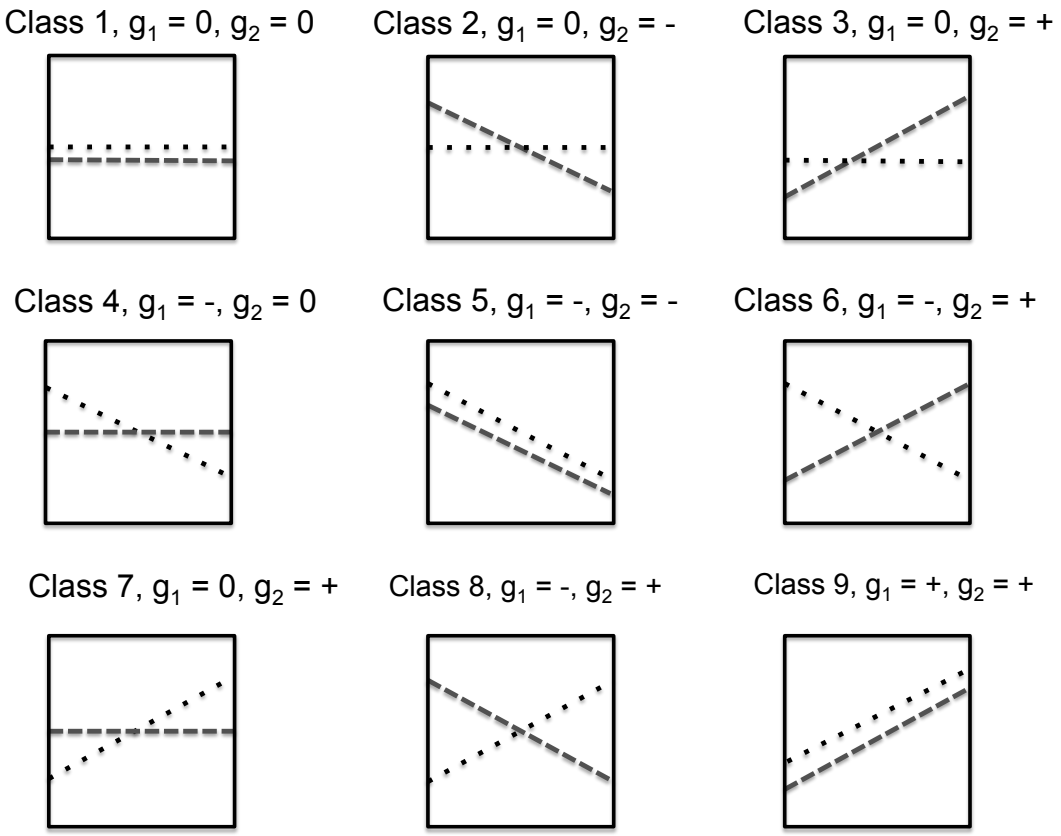
EBcoexpress

Initial Parameters

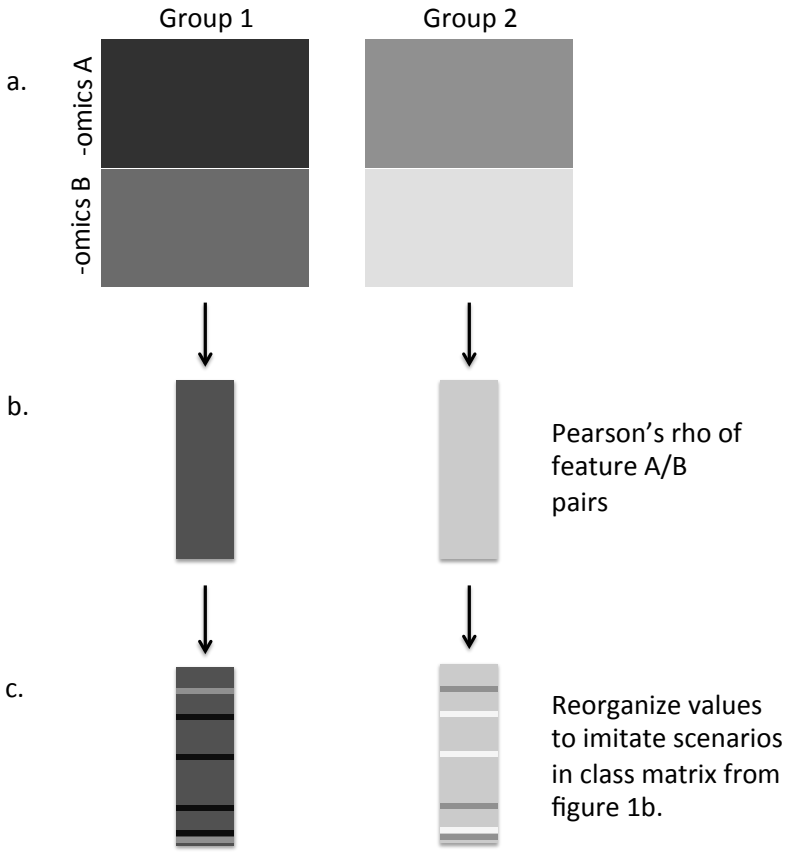
Initial hyperparameters in EBcoexpress can either be determined by using the normal mixture modeling function mclust() by default in the EBcoexpress R package (Dawson et al., 2012), or by a grid approach. We did the latter for the simulations since the hyperparameters determined by mclust produced posterior probabilities that had small range and were noninformative. The hyperparameters we determined by the grid approach were based on three components and produced a more even distribution of posterior probabilities. Since EBcoexpress takes long periods to run with large datasets, only ten percent of the data was randomly selected for ten iterations when determining appropriate hyperparameters. The final output of EBcoexpress is either the posterior probability of differential coexpression or equivalent coexpression

Hardware

Biological data was run on a server with 11 Intel(R) Xeon(R) CPU E5-2640 0 2.50GHz processors, 90 GB of memory and hard disk drive able to hold 41 GB. Simulations were run on a MacBook Pro with a 2.7 GHz Intel Core i7 and 8 GB of memory. Run-time reported for biological data is measured on the larger server.



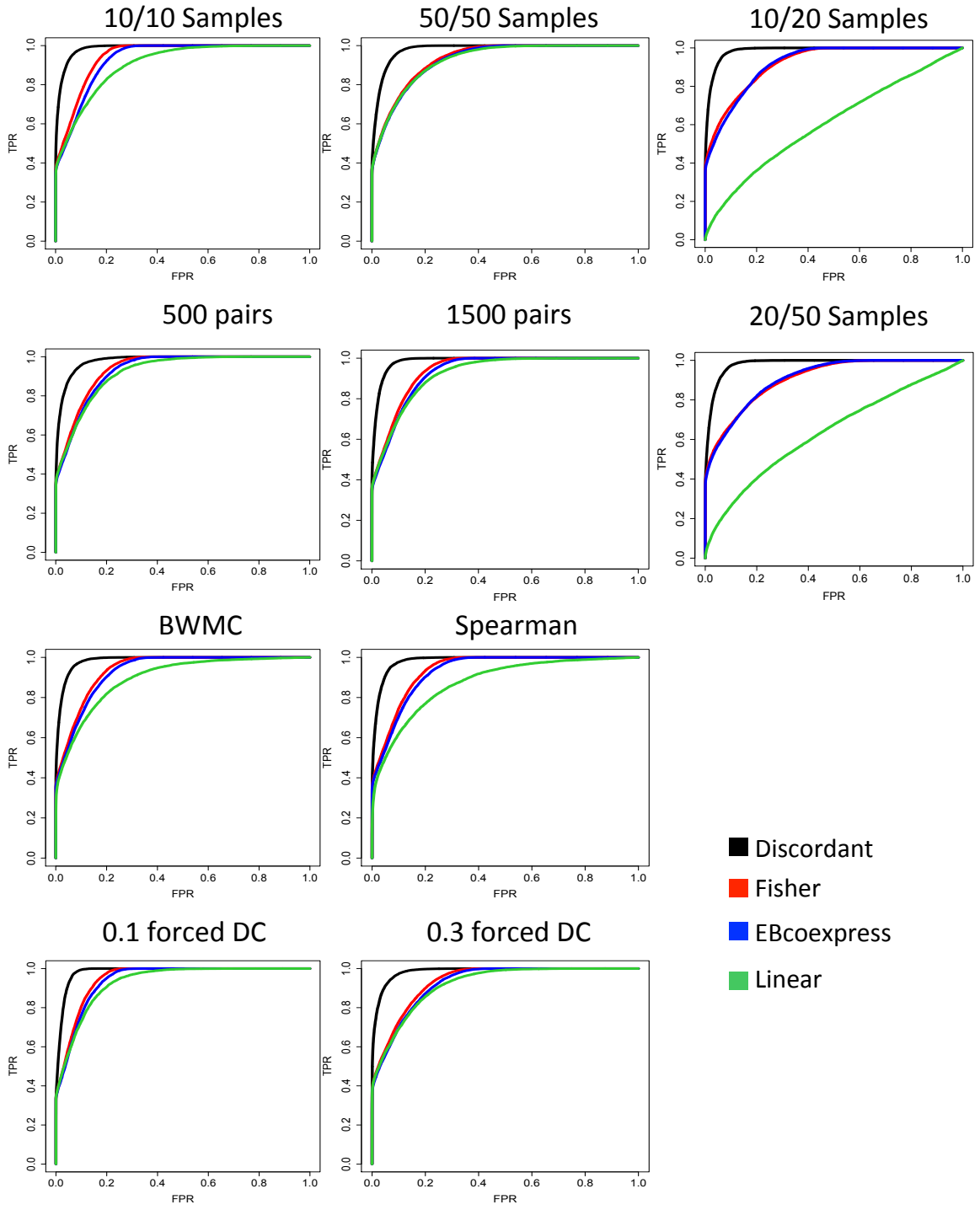
Supplementary Figure 1. Visualization of classes from class matrix in figure 1b. Group 1 dashes, group 2 dots.



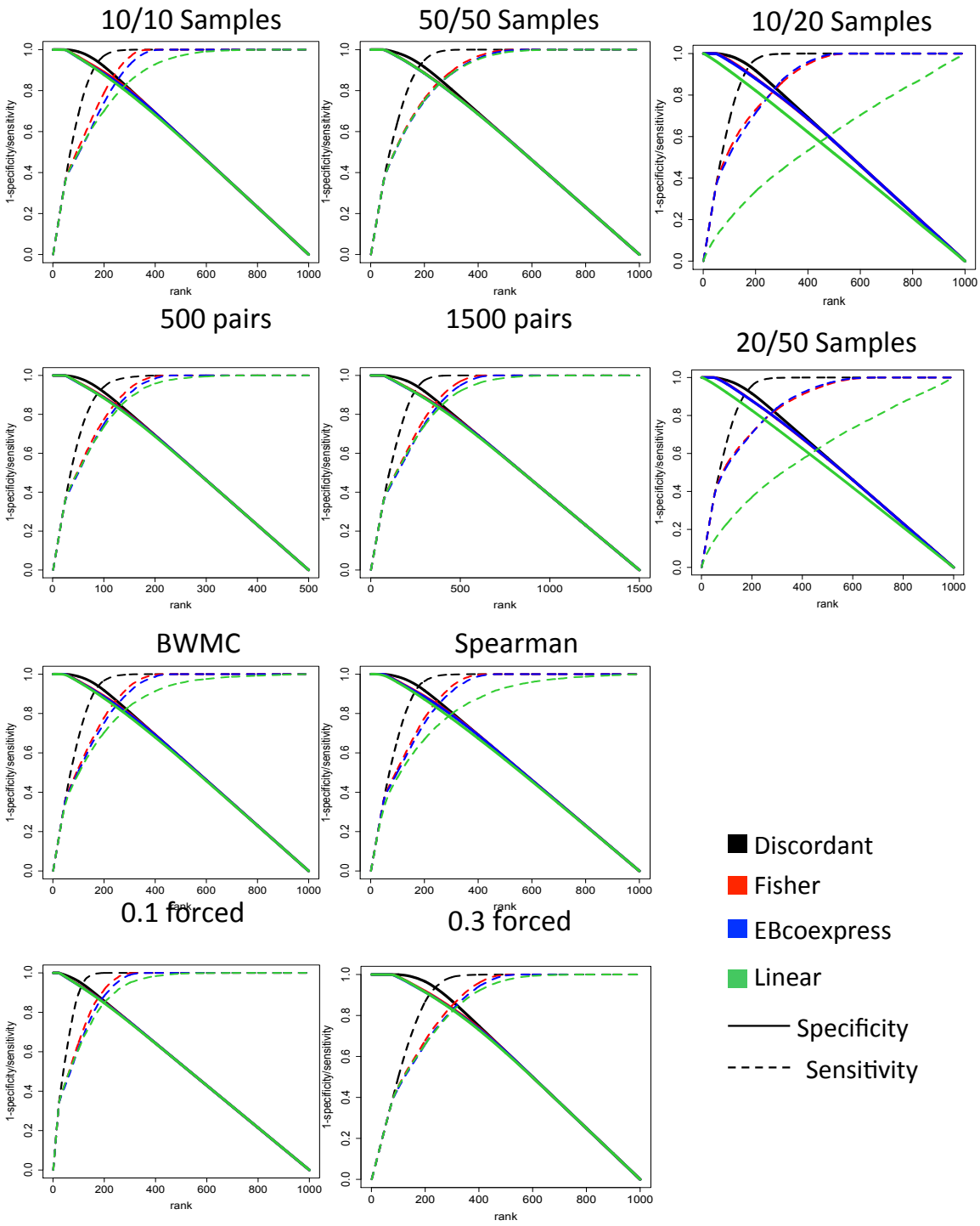
Supplementary Figure 2. Pipeline of simulations. Simulations were set up in order to easily control the number of feature pairs that were DC, the total number of feature pairs and sample size. Also, we wished to examine the effect the data had when using two different type of –omics.
a. Two matrices are formed for each group, where the two matrices associate with A or B omics, for a total of 4 matrices shaded in the figure. Each of the four matrices are simulated using mvrnorm in the R package MASS where the mean was 0 and covariance matrix is a diagonal matrix. This results in independent samples/subjects and independent omics features.
b. Correlation coefficients are determined between the mth feature in omics A and the nth feature in omics B, so there are M X N feature pairs, where M is the number of omics A features and N is the number of omics B features. As in the real data, this creates dependencies among some of the feature pairs since a feature occurs in more than one pair.
c. The number of DC pairs is determined by sectioning out the most significant correlation coefficients. These are swapped to imitate the scenarios in the class matrix from figure 1b, e.g. a positive correlation coefficient is paired with 0 in order to imitate class 3 or 6.

Sample Size	Forced DC Pairs	Feature Size	Correlation Method
10 10	0.1	500	Spearman
20 20	0.2	1000	Pearson
50 50	0.3	1500	Biweight MidCorrelation

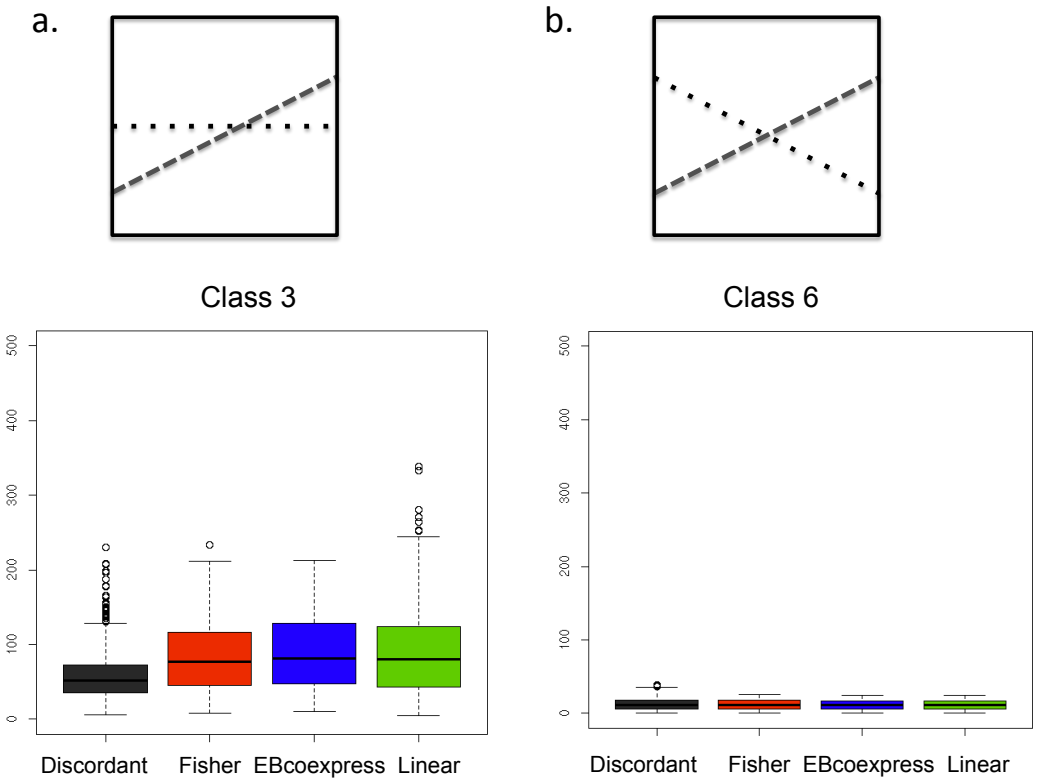
Supplementary Table 1. Summary of Simulation Alteration. Highlight red is the standard all alterations were based on and reported in Figure 4 and 5. ROC, sensitivity and specificity curves for other parameters found in Supplementary Figures 4 and 5.



Supplementary Figure 3. ROC curves with changes in simulation parameters. Title tells parameter changed from shaded red row in Supplementary Table 1.



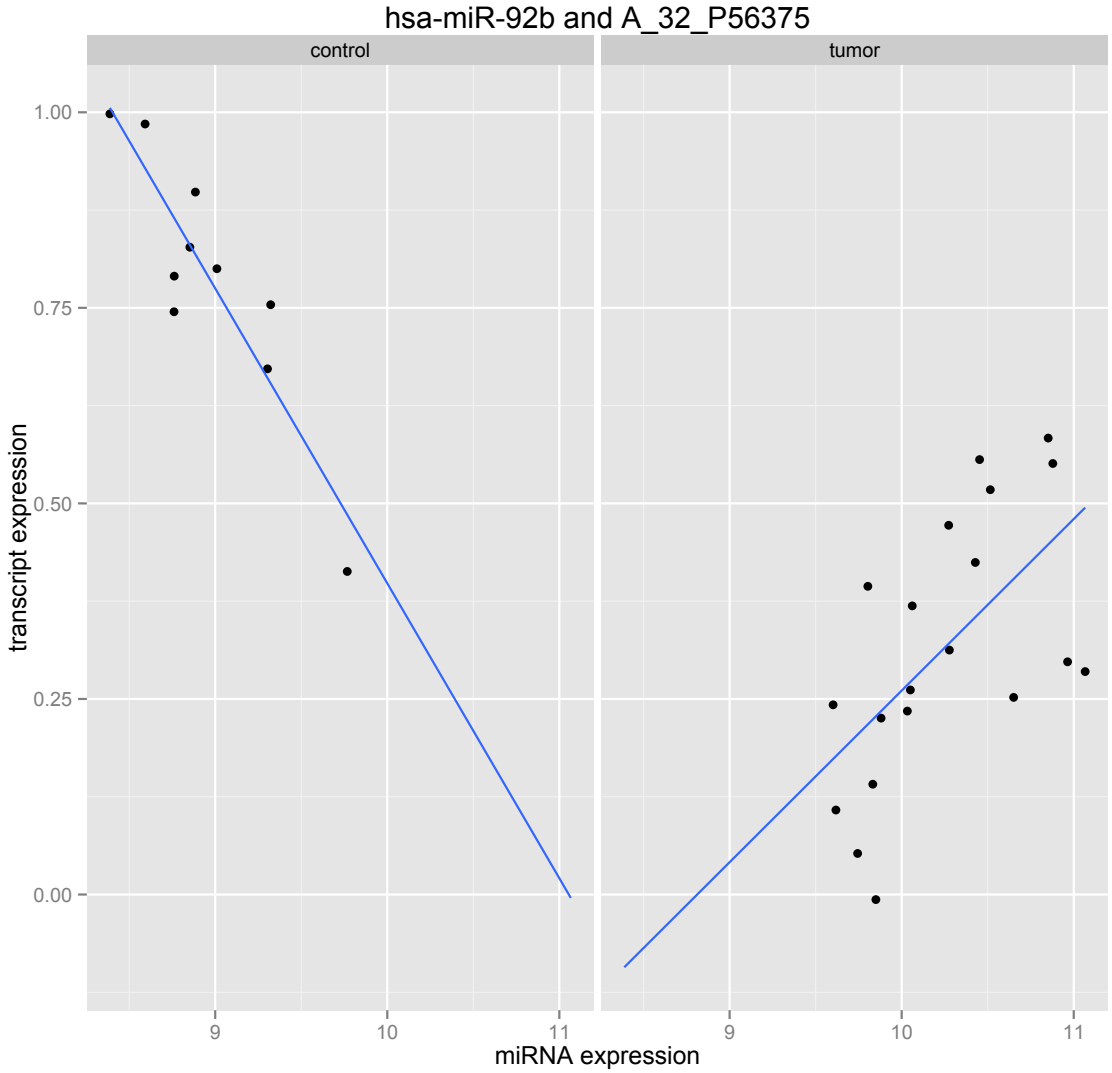
Supplementary Figure 4. Sensitivity/1-Specificity curves with changes in simulation parameters. Title tells parameter changed from shaded red row in Supplementary Table 1.



Supplementary Figure 5. Distribution of class rank for all methods. group 1 and 2 are dashed and dotted respectively. a. Class 3 is an example of disrupted DC where group 1 has positive correlation, and group 2 has no correlation. Discordant has a smaller distribution of rank for class 3 compared to other method. b. Class 6 is an example of cross DC where group 1 has positive correlation and group 2 has negative correlation. All four methods have similar distribution of rank for class 6.

Method	statistic	hsa-miR-124a	hsa-miR-137	hsa-miR-326	hsa-miR-92b
Discordant	Rank	223	1081	472	83
	1 - pp	2.51e-7	1.16e-6	5.04e-7	1.05e-7
	1 - q-value	4.98e-7	2.39e-6	1.02e-6	2.02e-7
EBcoexpress	Rank	585	1862	628	185
	1 - pp	3.54e-3	9.86e-3	3.84e-3	1.25e-3
	1 - q-value	6.66e-3	0.184	7.23e-3	2.44e-3
Fisher	Rank	799	1282	803	238
	p-value	3.61e-6	7.43e-6	3.63e-6	6.24e-7
	FDR	0.109	0.139	0.109	0.629
miRNA –Dep. Linear	Rank	3246	297	768	69
	p-value	5.65e-5	1.29e-6	6.34e-6	8.95e-8
	FDR	0.420	0.103	0.198	0.031
Transcript – Dep. Linear	Rank	467	8807	4	1108
	p-value	1.85e-5	4.72e-4	1.97e-7	4.78e-5
	FDR	0.950	1	0.578	1

Supplementary Table 2. Summary of unique GBM-related miRNAs top ranked pair with a transcript. Shown is rank, p-value/1 - posterior probability (1 - pp) and FDR/q-value for Discordant, EBcoexpress and Fisher.



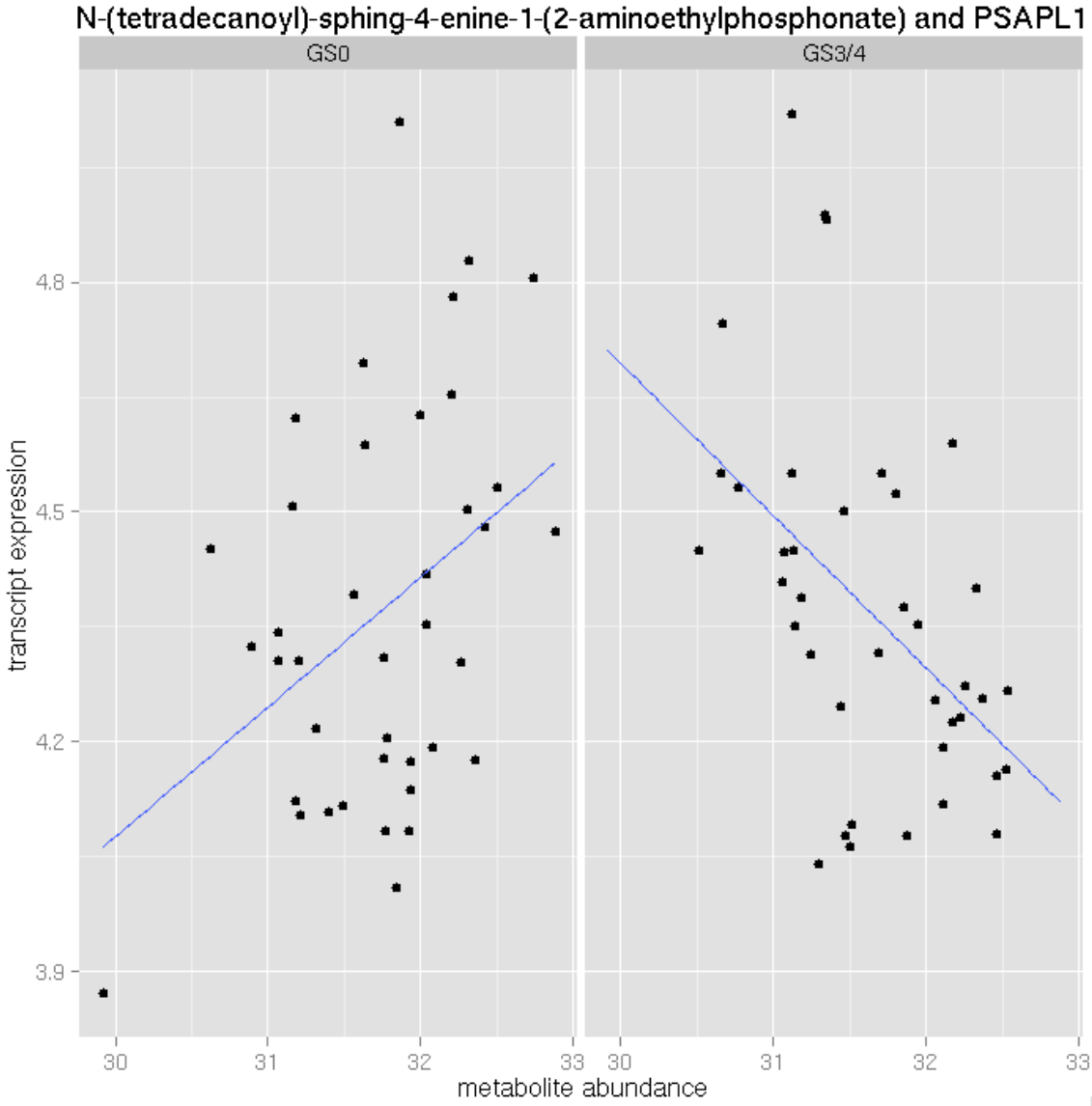
Supplementary Figure 7. Top example of unique GBM-related miRNA differential correlation in GBM data in Discordant (probe was not annotated, probe name in title).

Gene	Connections	Annotations
AGAP2	60	Anti-apoptotic effects of nerve growth factor
CRY2	39	Circadian protein
GRIN1	34	Glutamate receptor, form ligand-gated ion channel
UPF3A	34	Part of post-splicing multiprotein complex involved in mRNA decay and nuclear export.

Suppl. Table 3. Summary of gene hubs with top connections in pairs with q-value 1.0e-4 in GBM analysis.

	Discordant			EBcoexpress			Fisher		
metabolite	Rank	pp	1 - q-value	Rank	pp	1 - q-value	Rank	p-value	FDR
*1-3Galalpha1-3Galalpha1-3Galalpha1-4Galbeta1-4Glcbeta-Cer(d18:1/24:1(15Z))	272483	6.87e-3	1.31e-2	312901	0.385	0.496	998686	0.216	1
3-O-Sulfogalactosylceramide (d18:1/18:1(9Z))	105638	2.91e-3	5.54e-3	89993	0.253	0.345	99173	0.096	1
C18-OH Sulfatide	854304	1.96e-2	3.79e-2	554288	0.452	0.566	523347	0.173	1
Cer(d18:1/22:1(13Z))	158937	4.21e-3	8.02e-3	165520	0.314	0.417	91285	0.093	1
Cer(d18:1/23:0)	626399	1.47e-2	2.83e-2	372631	0.406	0.517	384322	0.155	1
Cer(d18:1/24:1(15Z))	70023	2.02e-3	3.82e-3	66774	0.226	0.311	53879	0.077	1
Cer(d18:2/23:0)	198078	5.14e-3	9.80e-3	372969	0.406	0.517	102053	0.097	1
Ceramide (d18:1/18:0)	161785	4.27e-3	8.14e-3	321358	0.388	0.499	70169	0.085	1
Ceramide (d18:1/22:0)	15721	5.49e-4	1.02e-3	15698	0.123	0.178	16451	0.050	1
Ceramide (d18:1/25:0)	134584	3.62e-3	6.90e-3	137213	0.295	0.394	78423	0.088	1
CerP(d18:1/24:1(15Z))	169281	4.45e-3	8.49e-3	255398	0.362	0.470	796240	0.200	1
Ganglioside GA1 (d18:1/16:0)	409392	9.96e-3	1.91e-2	441601	0.426	0.538	944945	0.212	1
Ganglioside GM1 (18:1/9Z-18:1)	168094	4.42e-3	8.44e-3	290467	0.377	0.486	457741	0.165	1
Ganglioside GM3 (d18:1/14:0)	663578	1.55e-2	2.99e-2	733865	0.486	0.599	423699	0.161	1
Ganglioside GM3 (d18:1/16:0)	1160391	2.60e-2	5.07e-2	633445	0.468	0.581	733809	0.195	1
GlcAbeta-Cer(d18:1/18:0)	927875	2.11e-2	4.10e-2	566984	0.455	0.568	1098109	0.224	1
Glucosylceramide (d18:1/25:0)	731165	1.69e-2	3.27e-2	527497	0.447	0.560	676768	0.189	1
Lactosyceramide (d18:1/18:1(9Z))	147435	3.93e-3	7.49e-3	217794	0.344	0.451	262888	0.136	1
N-(tetradecanoyl)-sphing-4-enine-1-(2-aminoethylphosphonate)	376192	9.22e-3	1.77e-2	432036	0.423	0.535	226664	0.129	1
NeuGcalpha2-3Galbeta1-4GlcNAcbeta1-3Galbeta1-4Glcbeta-Cer(d18:1/24:1(15Z))	472751	1.13e-2	2.18e-2	545833	0.451	0.564	1043493	0.220	1
N-Lignoceroylsphingosine	1871161	4.09e-2	8.05e-2	1334581	0.556	0.665	1877890	0.269	1
N-Palmitoylsphingosine	520229	1.24e-2	2.38e-2	319315	0.388	0.498	441743	0.163	1
SM(d16:1/17:0)	130542	3.52e-3	6.71e-3	107532	0.270	0.365	298258	0.142	1
SM(d18:0/24:1(15Z))	1207572	2.70e-2	5.27e-2	1803152	0.590	0.695	1106829	0.224	1
SM(d18:1/12:0)	12906	4.64e-4	8.58e-4	14048	0.117	0.169	95758	0.095	1
SM(d18:1/14:0)	3146	1.45e-4	2.61e-4	3449	0.060	0.090	23730	0.057	1
SM(d18:1/16:1)	1855826	4.06e-2	7.99e-2	2791220	0.637	0.735	2001704	0.276	1
SM(d18:1/20:1)	2624635	5.68e-2	1.12e-1	2090254	0.606	0.709	1590396	0.254	1
SM(d18:1/23:0)	140351	3.76e-3	7.17e-3	222316	0.453	0.453	188406	0.121	1
SM(d18:1/24:1(15Z))	198733	5.15e-3	9.83e-3	146886	0.402	0.402	334705	0.148	1
SM(d18:2/14:0)	493438	1.18e-2	2.27e-2	473547	0.547	0.547	523556	0.173	1
SM(d18:2/15:0)	214413	5.52e-3	1.05e-2	403837	0.527	0.527	74789	0.087	1
SM(d18:2/23:0)	144149	3.85e-3	7.34e-3	206117	0.338	0.444	75113	0.087	1
Sphingosine	187977	4.90e-3	9.34e-3	156783	0.308	0.410	451307	0.164	1
Sphingosine 1-phosphate	319191	7.94e-3	1.52e-2	394444	0.712	0.524	554183	0.176	1
Sphingosine-1-phosphocholine	99476	2.76e-3	5.24e-3	109241	0.272	0.367	461599	0.166	1
Trihexosylceramide (d18:1/16:0)	954799	2.17e-2	4.21e-2	536542	0.449	0.562	869371	0.206	1

Supplementary Table 4. Summary of sphingolipid metabolite top ranked pair with a sphingolipid-related gene. Shown is rank, p-value/1 - posterior probability (1 - pp) and FDR/q-value for Discordant, EBcoexpress and Fisher.



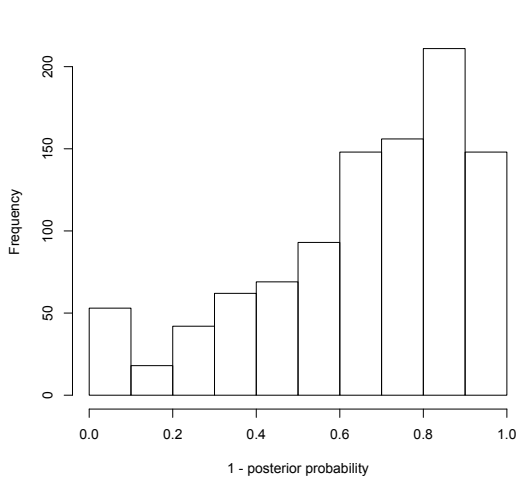
Supplementary Figure 8. Top example of sphingolipid-related differentially coexpressed pair in Discordant.

Type	Name	Connections	Annotation
Gene	LOC284561	247	unknown
Gene	SARDH	265	Sarcosine dehydrogenase
Gene	IGHG1	294	Immunoglobulin
Metabolite	C20 H33 N9 P2 S	2222	Unknown
Metabolite	L-valine	1667	Amino acid

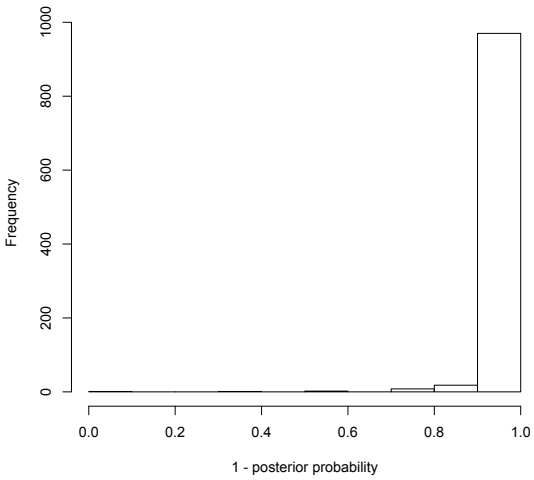
Supplementary Table 5. Summary of metabolite gene hubs with pairs with q-value 2.0e-3 in COPD analysis.

Method	GBM	COPD
Discord	19.26 hours	2.16 days
EBcoexpress	10.42 days	39.53 days
Fisher	6.35 seconds	18.85 seconds
Linear	16.15 hours	N/A

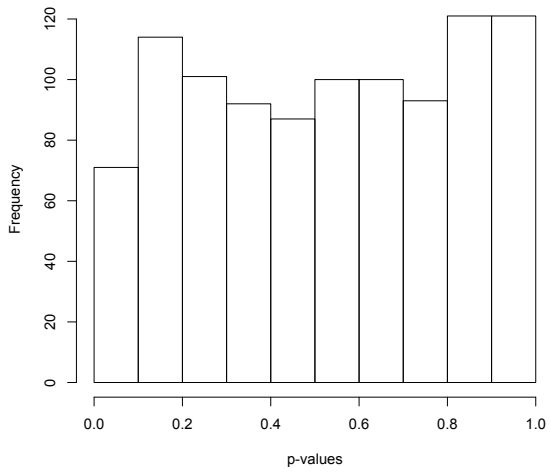
Supplementary Table 5. Run-time on 11 Intel(R) Xeon(R) CPU E5-2640 0 2.50GHz processors, 90 GB of memory.

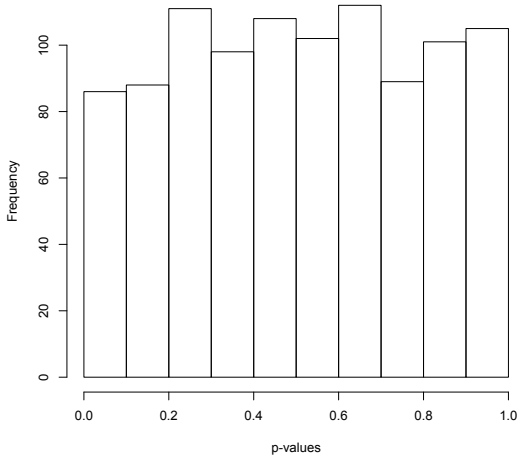


a.



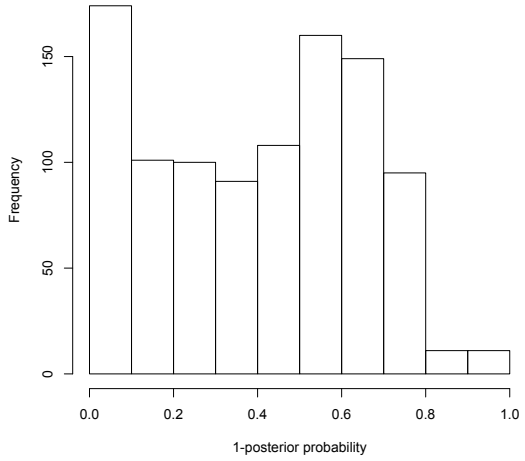
b.



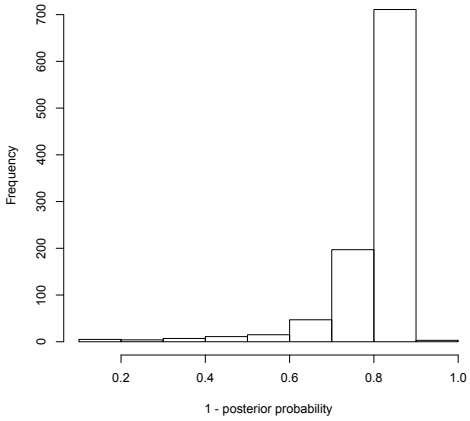


c.

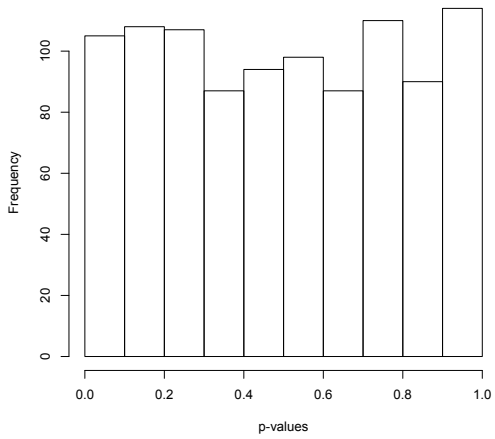
Supplementary Figure 9. Distributions of simulated p-values and posterior probabilities (1000 pairs, 0.2 DC, sample size 20 for both groups). a. Discordant. b. EBcoexpress c. Fisher d. Linear models



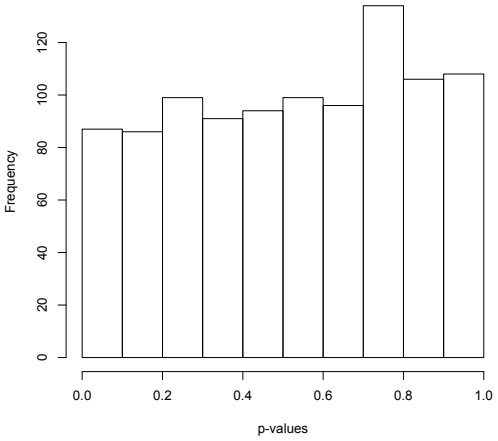
a.



b.

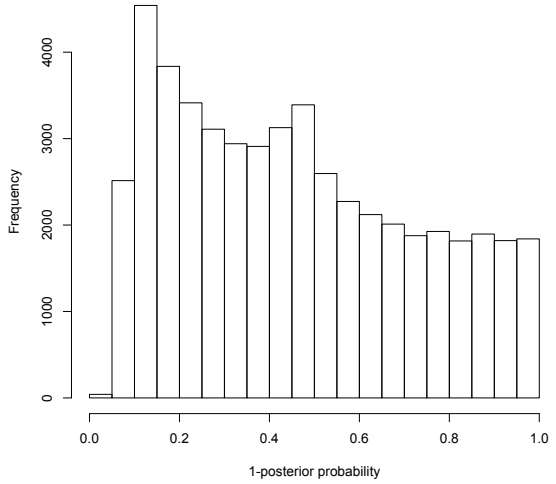


c.

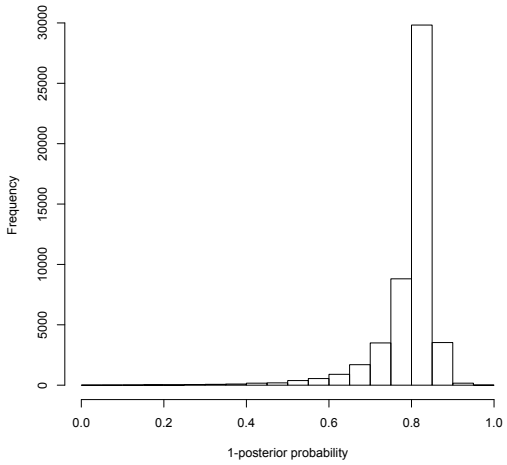


d.

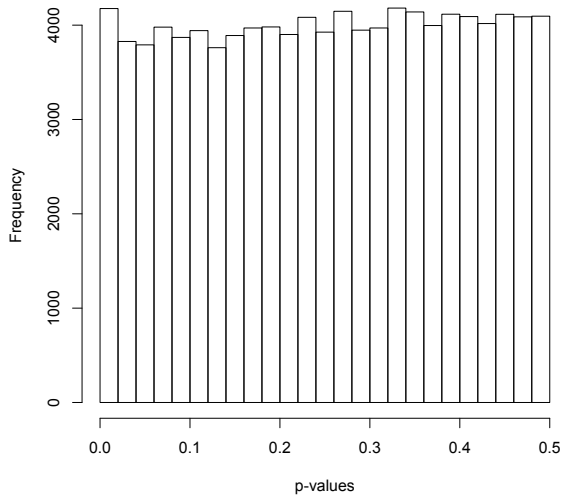
Supplementary Figure 10. Sample distributions of simulated p-values and posterior probabilities from TCGA GBM data. a. Discordant. b. EBcoexpress c. Fisher d. Linear model



a.



b.



c.

Supplementary Figure 11. Sample distributions of simulated p-values and posterior probabilities from COPD data. a. Discordant. b. EBcoexpress c. Fisher