Supplementary Data

Supplementary Methods

List Length Correction 2
Supplementary Figure 1. Determination of the bounds for the ratio of log P-values for equally proportioned enrichments involving different gene list length 3

Rank-Rank Hypergeometric Overlap User's Guide

Rank-Rank Hypergeometric Overlap (RRHO) maps generated by our web application 4
Interpretation of pixel values in RRHO maps 5
Examples of overlap signals using synthetic data 6
Under- versus over-enrichment, converting negative signal to positive signal 8
Statistical effects on the shape and peak of the hypergeometric probability distribution in comparison to the rank density distribution 9
Effect of increasing noise on the strength of the overlap signal in RRHO analysis 10
Comparing RRHO to a standard metric of correlation 11
User-Guide Figure 1: Range of correlated data points analyzed by RRHO and by Spearman rank correlation coefficient 12
User-Guide Table S1: Comparing RRHO to the Spearman rank correlation coefficient 13
Recommendations for multiple hypothesis correction 14
Sample permutations recommendations 15
Sample permutation P-value determination 15
Comparisons of raw heatmaps, multiple hypothesis corrected heatmaps and permutation P-values 16
User-Guide Table S2: Comparison of RRHO analysis summary statistics before and after the application of different multiple hypothesis correction approaches 16
False discovery rate for the overlapping gene list 17

Supplementary Methods

List Length Correction:

Since the hypergeometric distribution is based on counting statistics, similar percentages of overlap between ranked lists will be more statistically significant when the lengths of the lists are longer (due to the statistical effect of larger N). Thus, the length of the list of common genes between two experiments needs to be comparable to compare two maps directly. Consider the following example: if 50 genes were found to be overlapping in two groups of 100 genes chosen from a population of 1000 genes (P-value = 7 x 10^-29), this is far more significant than if 25 genes were found overlapping in two groups of 50 in a population of 500 genes (P-value = 4 x 10^-15), even though the proportion of overlapping genes is the same in both cases. The statistical difference here is real, but it is due to difference in gene number not percent overlap. Since percent overlap is of particular interest when comparing gene signatures, we have developed a technique to scale hypergeometric overlap maps to correct for differences in list length. Two common situations in which maps should be scaled are (i) when comparing maps made with different species, such as when a human-versus-human map is compared to a human-versus-mouse map which will be reduced in size by incomplete homolog mapping between the two species, and (ii) when two experiments are done on different microarray platforms that measure considerably different numbers of genes.

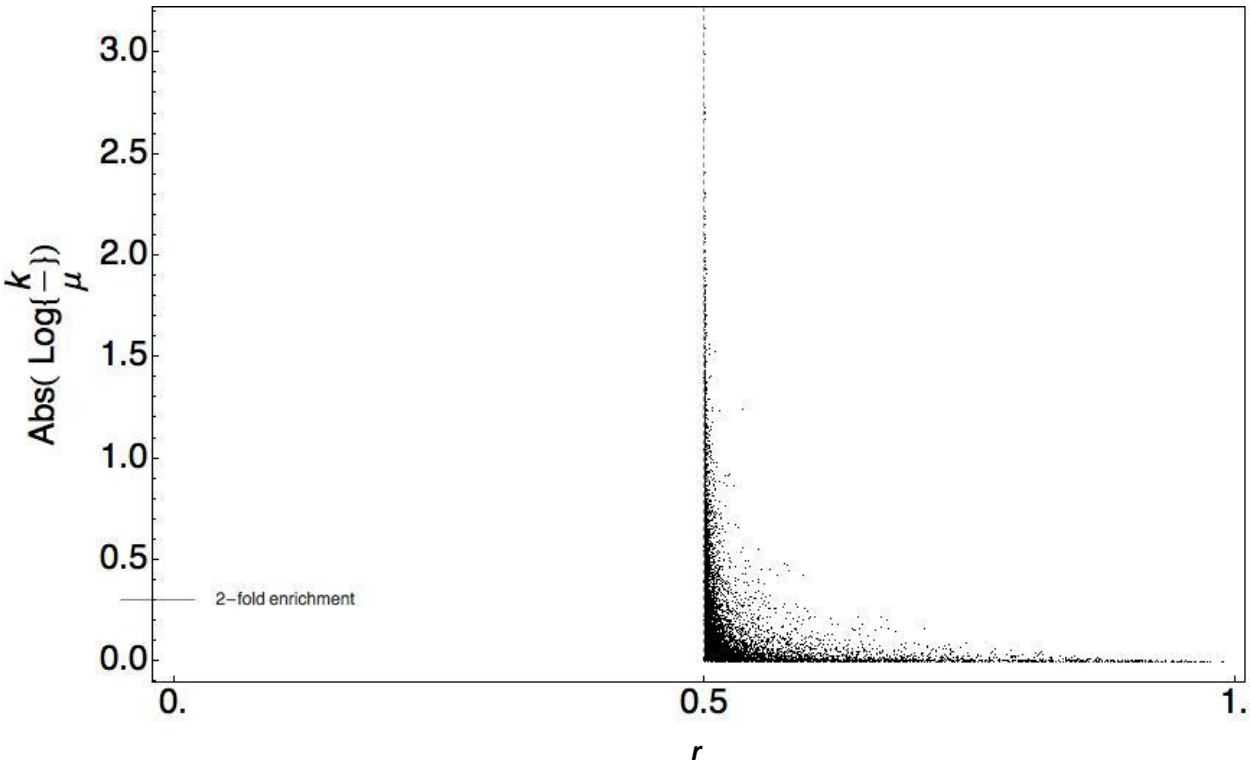
In order to address a P-value correction for differences in list size, we performed a numerical analysis to establish the bounds of the ratio of two log-transformed hypergeometric P-values where one of the P-values is calculated with all factors scaled by some factor, a. That is, if $H(k; s, M, N)$ is the hypergeometric P-value of k overlapping genes observed in two groups of size s and M in a total population of size N, what is the value of:

$$r = \frac{-\log(H(k;s,M,N))}{-\log(H(k\cdot a;s\cdot a,M\cdot a,N\cdot a))}$$

We calculated this ratio using 100,000 random numbers with these constraints: $a >= 1$, $0 <= M <= N$, $0 <= s <= N$, and $0 <= k <= \min(s, M)$. From this analysis, we observed that the ratio r has an upper bound of 1 and approaches $1/a$ when the observed overlap, k, is far from the mean ($\mu = sM / N$) or expectation value of the hypergeometric distribution (Supplementary Fig. 1). As little as a 2-fold enrichment ($\log(k / \mu)$ = log(2) = 0.3), brings the ratio close to $1/a$. Although it appears, by looking at Supp. Fig. 1, that the ratio has an exact lower bound of $1/a$, strictly speaking it does not since a few values are actually slightly less than $1/a$. However, in practice, it does no harm to consider $1/a$ as the lower bound. Therefore when comparing maps from RRHO comparisons performed with gene lists of unequal size, we recommend a 'list length correction' that scales the log P-values by the ratio of list lengths. Typically we reduce the significance of the longer list-based result using:

$$\log P_{longer,\ corrected} = \log P_{longer,\ uncorrected} \cdot (N_{shorter} / N_{longer}),$$

where N_i is the length of the respective list. After the P-values of the longer-list hypergeometric map have been scaled, the maximal log P-value on the resulting maps can be compared to determine which overlap is stronger overall. A second option is to only use the genes common to all experiments when comparing multiple RRHO overlap comparisons, but in some cases this will remove a substantial number of informative gene measurements.



Supplementary Figure 1. **Determination of the bounds for the ratio of log P-values for equally proportioned enrichments involving different gene list length.**

We first defined the ratio (*r*) of log P-values as:

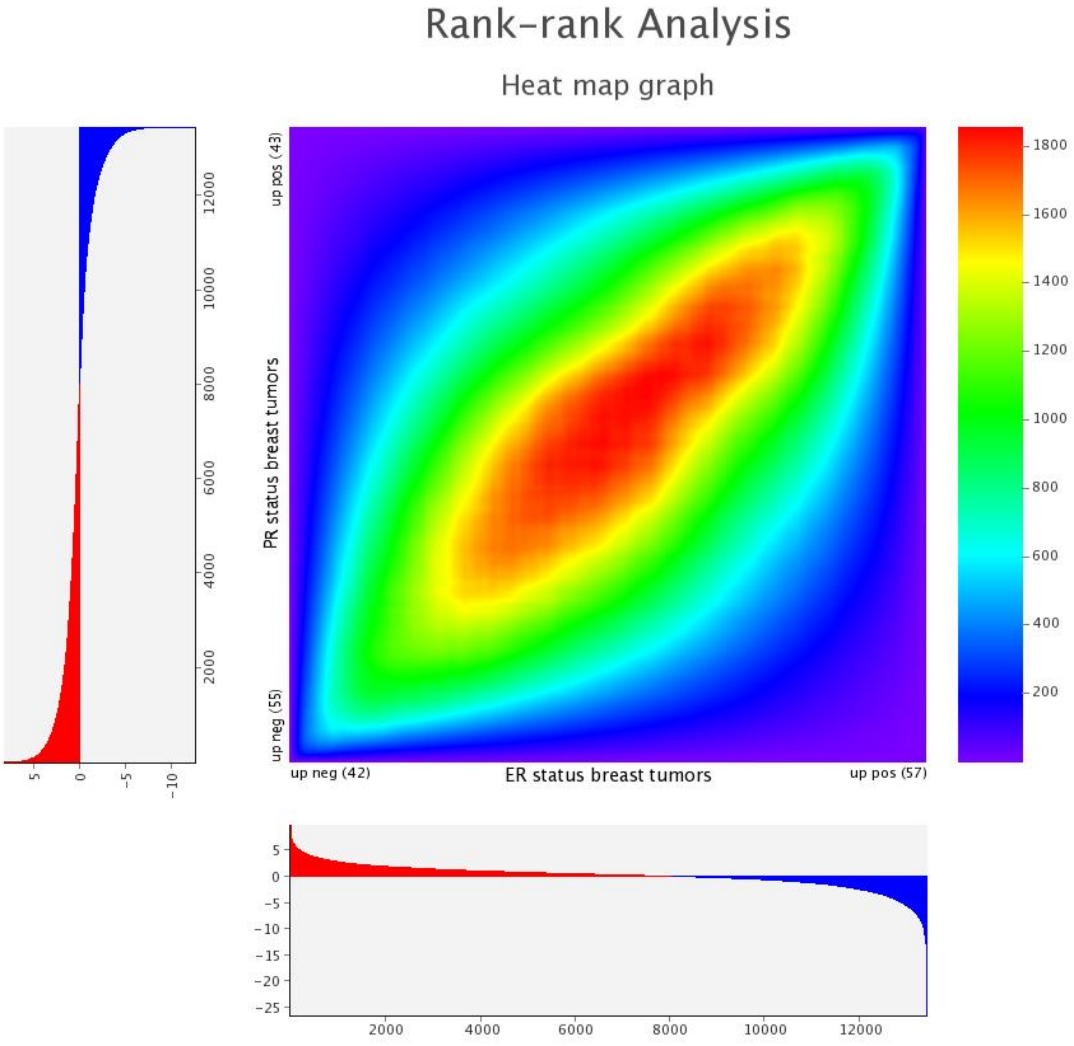
$$r = \frac{-\log(H(k; s, M, N))}{-\log(H(a \cdot k; a \cdot s, a \cdot M, a \cdot N))},$$

where the two enrichment calculations differ in that all factors are scaled by some factor, *a*. We next calculated this ratio using 100,000 random numbers with these constraints: $a >= 1$, $0 <= M <= N$, $0 <= s <= N$, and $0 <= k <= \min(s, M)$. The scatter plot shows the $a = 2$ ($1/a = 0.5$) case, plotting *r* versus the absolute value of $\log(k/\mu)$, where *k* is the overlap observed and $\mu = sM/N$ is the mean or the expectation value of the hypergeometric distribution. This numerical evaluation shows that the ratio is *almost* bounded by the interval [1/a, 1] and that even for cases of small enrichment the ratio tends towards 1/*a*. For example, a 2-fold enrichment ($\log(k/\mu) = \log(2) = 0.3$) results in ratios generally near 1/*a*. Thus our scaling correction of the hypergeometric P-values from RRHO comparisons using different gene list lengths is multiplication of the log P-values of the longer-list case by 1/*a*.

Rank-Rank Hypergeometric Overlap User's Guide

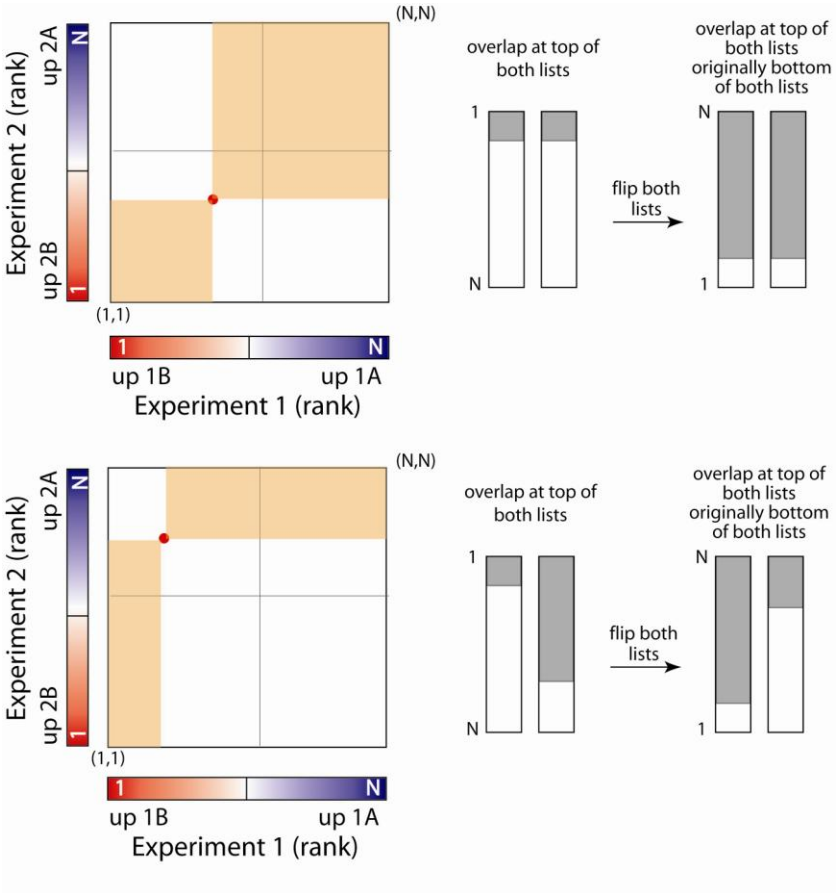
http://systems.crump.ucla.edu/rankrank/

Rank-Rank Hypergeometric Overlap (RRHO) maps generated by our web application:
The current version of RRHO online generates three different 2-dimensional graphical representations of the overlap or correlation between the two input gene-expression signatures (ranked lists of differentially expressed genes). As in Figure 2 of the manuscript, these are i) metric scatter plots, ii) rank-rank scatter plots, and iii) rank-rank hypergeometric overlap (RRHO) heatmaps. Benjamini-Yekutieli corrected hypergeometric maps can also be generated. In our website output, the metric values for differential gene-expression are plotted in a bar graph along the x- and y-axes. These metric plots were removed in the main manuscript figures due to space limitations. An example of the RRHO heatmap format from the website output corresponding to Figure 2C is shown below. The metric used is a signed, $\log_{10}$-transformed t-test P-value. The website has a demo link to see examples of the other plot types. The numbers in parenthesis indicate the number of samples in each experimental class.



Interpretation of pixel values in RRHO maps:

Each pixel in a RRHO map represents the $\log_{10}$-transformed hypergeometric overlap of subsections of two ranked lists. Pixels with positive values (red) indicate a higher than expected number of overlapping genes in the subsections and pixels with negative values (blue) indicate a lower than expected number of overlapping genes. Since the RRHO algorithm slides the rank threshold from rank 1 (top of list) to rank N (bottom of the list), the value of any pixel (coordinates x, y) represents the overlap between genes with ranks 1 to x in the ranked gene list of experiment 1, and 1 to y in the gene list of experiment 2. The symmetry of the enrichment question and of the hypergeometric distribution ensures that enrichment at the top of the two lists is equivalent to enrichment at the bottom of the two lists using the same rank threshold points, $H(k;s,M,N) = H(N-k;N-s,N-M,N)$. In other words, the above threshold regions cannot be over-enriched unless the below threshold regions are also enriched. Therefore each pixel value can be interpreted with regards to the bottom left corner (rank 1, 1) or to the top right corner (rank N, N). Since there are a fixed number of genes in the ranked lists, over-enrichment in the shaded areas will correspond to an under-enrichment in the complementary white areas. This means that the probability for over-enrichment between the top of list 1 and the top of list 2 is equal to the probability for the corresponding under-enrichment between the top of list 1 and the bottom of list 2, or $R(k;s,M,N) = -R(s-k;s,N-M,N)$ due to our sign convention for over- or under-enrichment. All genes from the top of list 1 that do not overlap with the genes above the threshold of the 2nd list must overlap with the genes below the threshold. See Methods for the full definitions of *H*, the hypergeometric cumulative distribution function (CDF) P-value, and *R*, the signed and $\log_{10}$-transformed value of *H*.

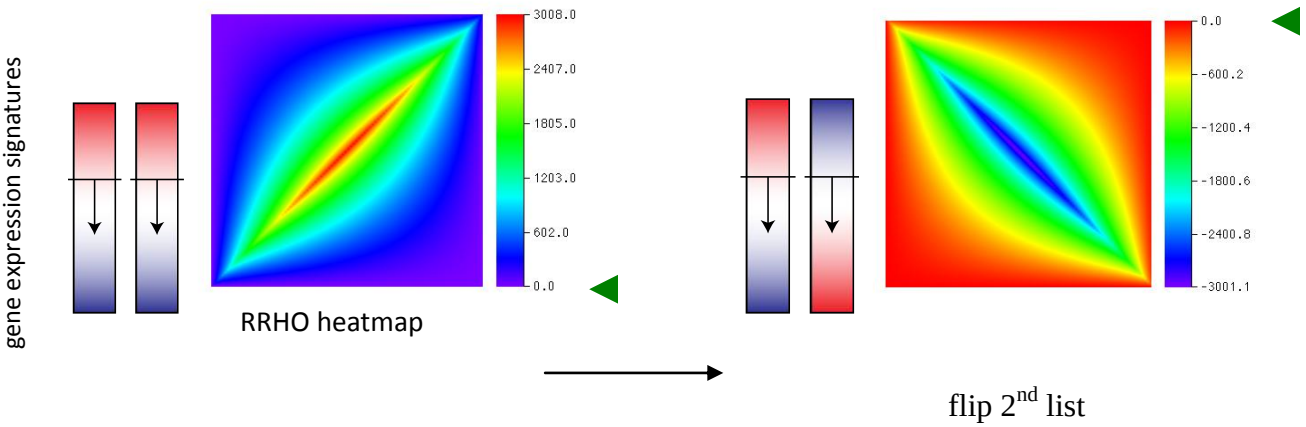


Examples of overlap signals using synthetic data:

To aide interpretation of the RRHO maps, we have created a series of synthetic comparisons to show how and where overlap signal appears. The synthetic RRHO maps and the associated pattern of overlap illustrated in the gene-expression signature (ranked gene list) schematics can be used to better understand the nature of the overlap in a true RRHO comparison result.

Perfect overlap

The first examples below were created by starting with two synthetic gene lists of length 10000 with exactly the same rankings. Perfect correlation is the overlap between these identical ranked lists and anti-correlation is the overlap when the second list is put in reverse order as illustrated by the red to blue gradient in the gene list schematics. In the perfect correlation RRHO heatmap, the most significant (red) area is across the bottom left to top right diagonal of the map (the exact value of the peak significance will in general vary based on the length of the lists). The results of comparing two lists that are the exact reverse of one another shows a strong negative value down the perpendicular diagonal. Green triangles indicate the position in the color scale corresponding to zero. A negative signal is interpreted as seeing less overlapping genes than expected above the corresponding rank thresholds. See further discussion on under- and over-enrichment and the effects of flipping gene lists in the "under- versus over-enrichment" section below.

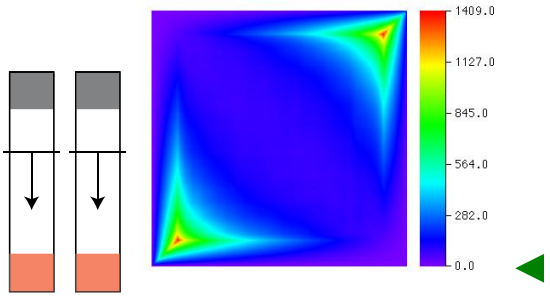


Overlap at the extremes

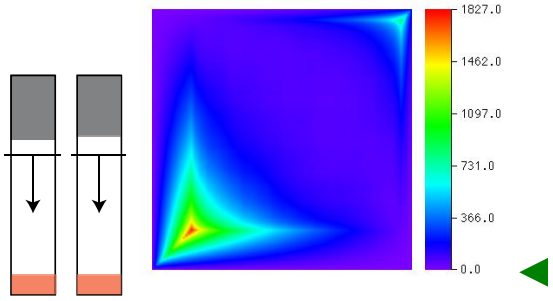
In practice, biological expression signatures tend to show overlap only at the extremes of the ranked lists, as the relatively non-changing genes found in the middle of the list have more rank variation. This is in part because the genome-wide scale of gene-expression platforms means that often many tissue-specific genes are not expressed in the samples tested and thus have noise-driven values. Mis-predicted genes and poorly hybridizing probes also contribute.

For the next examples, the gene rankings in the two lists are first randomly shuffled relative to each other, and then 1000-gene sections of the first list are pulled to the corresponding position of the second list to create the trends indicated in the gene list schematics. In the schematics the shaded regions match perfectly and the un-shaded areas have random order between lists. Green triangles indicate the position in the color scale corresponding to zero. These maps are meant to indicate the main trends from overlapping patterns. In general maps made from experimental data show more variation due to biological and measurement noise.

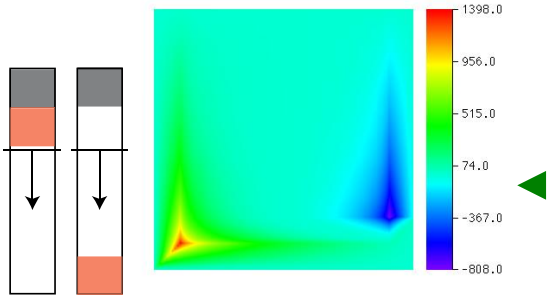
Perfect overlap at the ends of both lists (1000 genes at each end):



More overlap at the tops than bottoms (1500 genes at the top, 500 at the bottom):



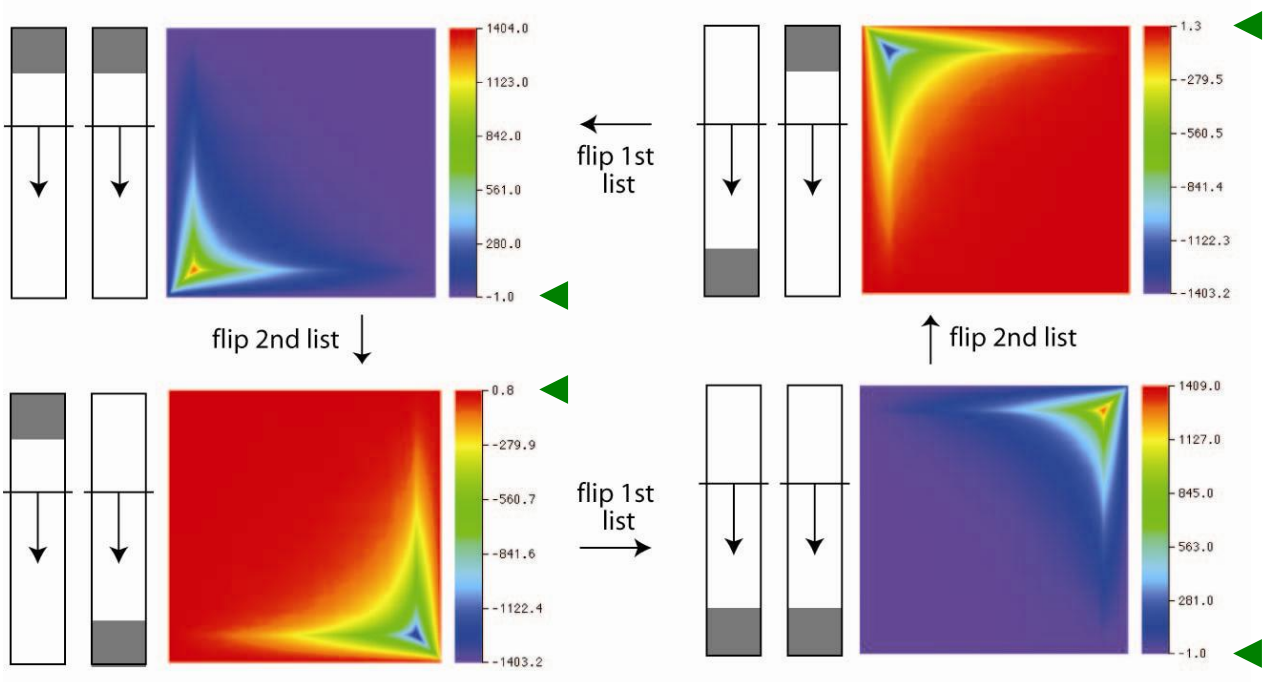
Split signal (1000 genes for each shaded region):



Under- versus over-enrichment, converting negative signal to positive signal:

Since the RRHO approach analyzes the genes above the threshold limits in the ranked lists, flipping a single list will always turn a higher than expected overlap into a lower than expected overlap at the equivalent threshold points (or vice versa). In a mathematical sense, the symmetry properties of the hypergeometric distribution and our sign convention for under- or over-enrichment are such that flipping a list only changes the sign of the direction-signed and $\log_{10}$-transformed hypergeometric CDF: $R(k;s,M,N) = -R(s-k;s,N-M,N)$ (Methods). Thus, a strong negative signal can be interpreted as a strong positive trend in the opposite or flipped sense. In these cases, flipping one list creates a map that illustrates this alternative way of looking at the results. This is analogous to switching classes A and B in one of the experiments, for example switching from an 'up in sensitive' to an 'up in resistant' perspective. This type of switching can make directional enrichment trends easier to describe in relationship to the biological phenotype.

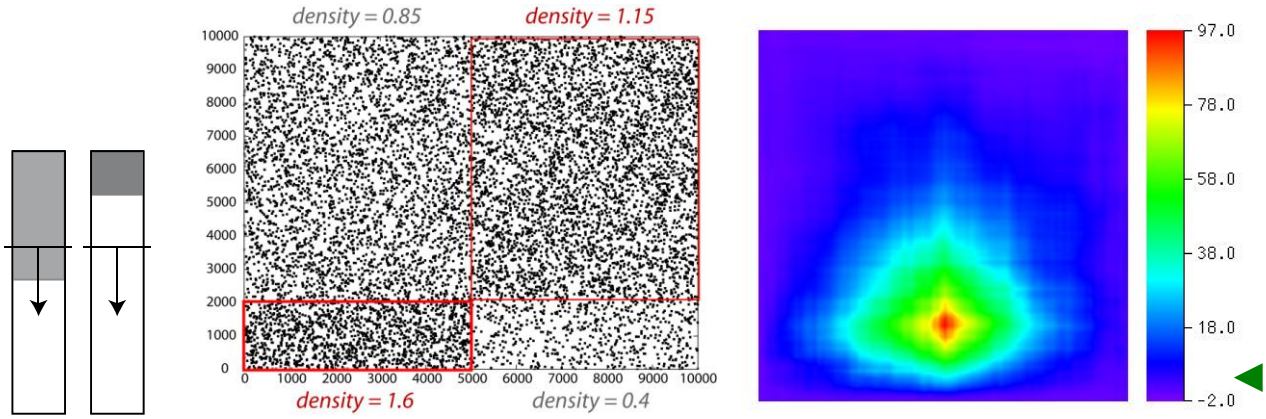
The synthetic heatmaps below illustrate this property by alternatively flipping the two lists. Green triangles indicate the position in the color scale corresponding to zero.



Statistical effects on the shape and peak of the hypergeometric probability distribution in comparison to the rank density distribution.

In this example, we again began with two randomly ranked lists. Then the overlap between the first 5000 genes in ranked list 1 with the top 2000 genes in ranked list 2 was evenly increased using 600 random swaps from the other parts of the ranked lists. By increasing the density in this area (thick red line in rank-rank scatter below), we correspondingly cause points to accumulate in the area below both of these thresholds (thin red line) due to the restriction that for each axis (gene-expression experiment) each rank must be uniquely associated with a gene. Likewise, the density in the other two regions decreases. The density of all four areas normalized to the scatter density of two randomly ordered lists is indicated.

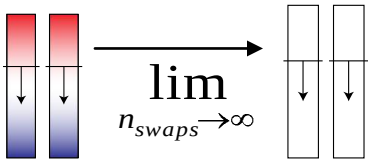
The nature of the hypergeometric distribution is that the P-value of overlap gets more and more significant with increasing sample size and otherwise similar overlap ratios. Therefore, in cases like this where the density of overlapping genes at the top of the list remains high as you slide the rank threshold in rank list 1, the statistical overlap signal will increase along this axis until the region of increased density is fully incorporated. Thus a strong peak in the RRHO map can be shifted away from the bottom left corner (rank 1, 1) in comparison to the larger region of enriched overlap density. This is in part a statistical phenomenon related to the effect of larger numbers improving the significance calculation. For this reason, the rank-rank scatter plots aid in interpretation of the RRHO maps and are accordingly included in our website output.

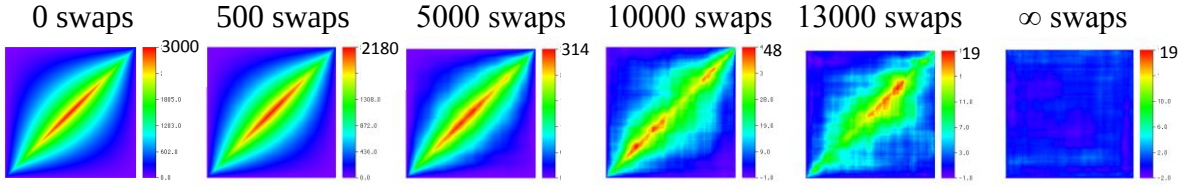


Effect of increasing noise on the strength of the overlap signal in RRHO analysis:

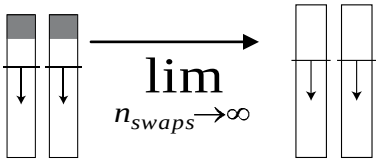
The examples above show results of perfect overlap in part or all of the ranked lists; signals this strong will never be observed in true expression profiling experiments. To model differing levels of noise, we randomly swapped the ranks of an increasing number of genes in the second gene list to show how the overlap map changes. Note that the scale of the maximal $-\log_{10}$(P-value) also decreases as noise is added.

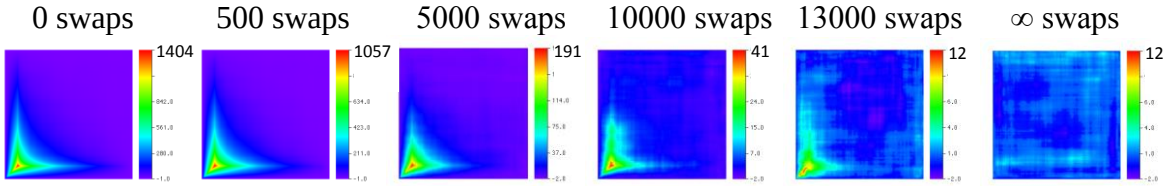
Random swaps in two identical overlapping lists:



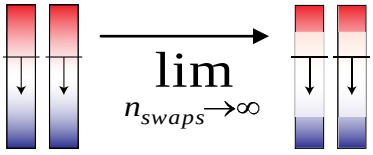


Random swaps in lists overlapping at the top:

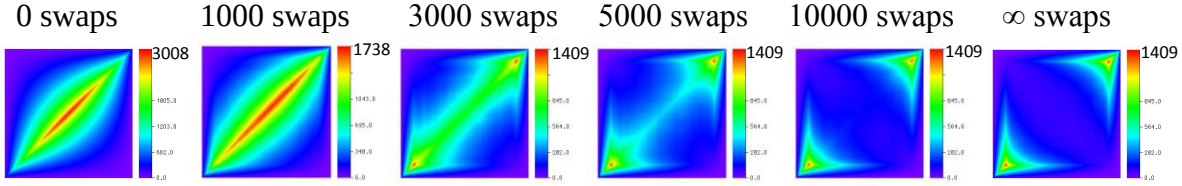




Random swaps in the middle of two identical lists:



Here the random swaps are limited to the genes between rank thresholds 1000 and 9000 in a 10000 gene list.



Comparing RRHO to a standard metric of correlation

To put RRHO analysis in context of existing measures of correlation, we generated synthetic ranked gene lists representing a range of correlation and compared the RHHO maps and maximal hypergeometric P-values to the Spearman rank correlation coefficient.

Synthetic data sets: We first created synthetic data sets containing two ranked lists of *n* elements in each list. The lists begin with identical ranks 1 to *n* (*n*=10000). Then, for each list individually, Gaussian noise was added to every element of the sequence of *n* positive integers to create a set *S*:

$$S = \{s_k\}, \text{ with } s_k = k + N(0,\sigma(k)), \; k=\{1, 2, \ldots, n\}, \text{ and } n=10000,$$

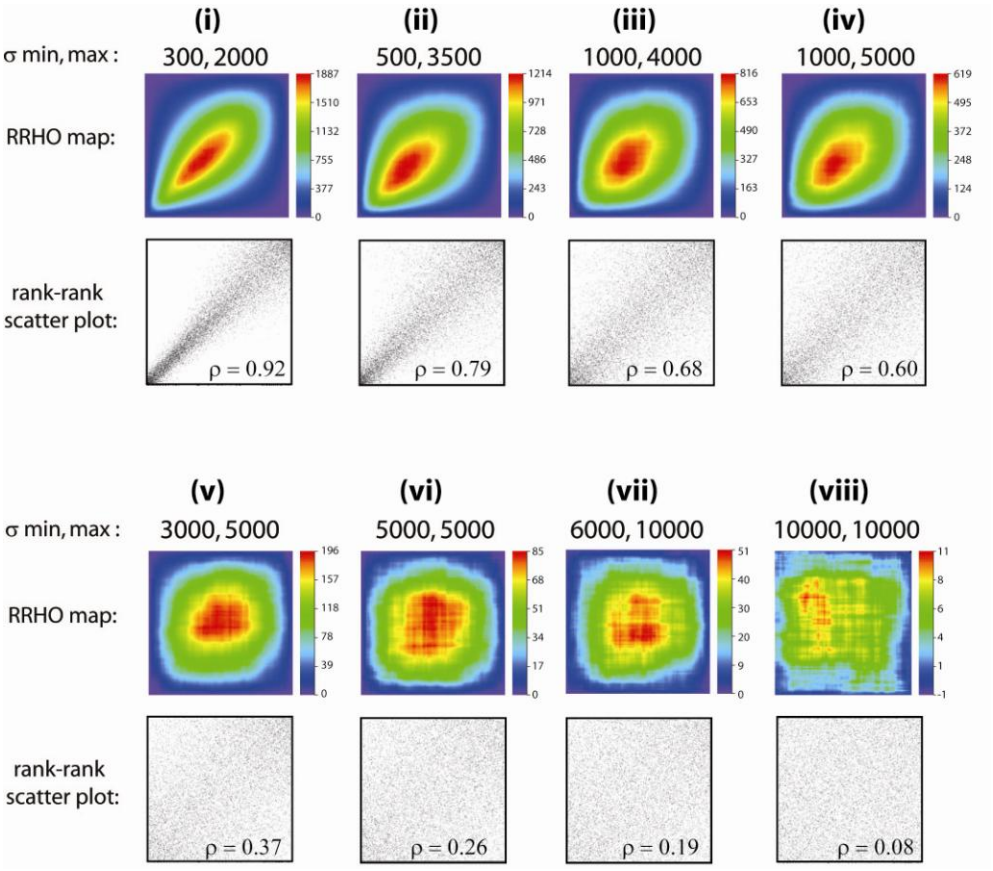
where $N(\mu,\sigma)$ was a normally distributed random variable. The standard deviation $\sigma(k)$ was a function of rank position (k) and two parameters, σ_{min} and σ_{max}, controlling the degree of randomness:

$$\sigma(k) = \sigma_{min} + \left(\frac{k-1}{n-1}\right)^2 (\sigma_{max} - \sigma_{min}).$$

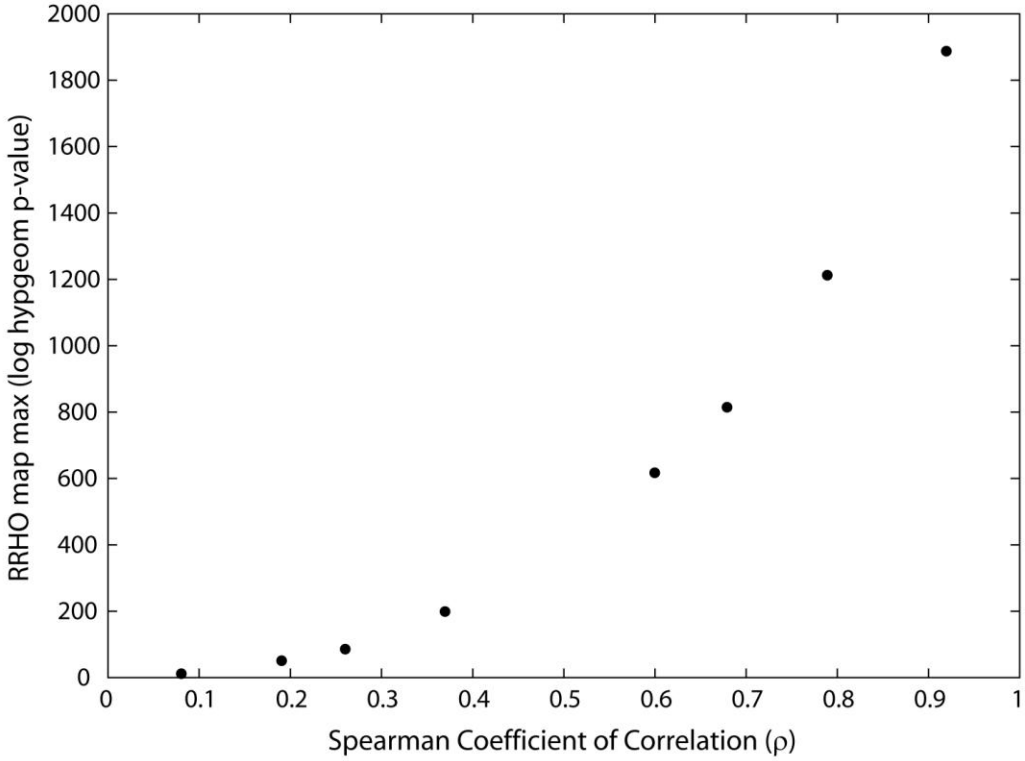
The genes in the set were then ranked based on s_k. When σ_{max} is greater than σ_{min}, more Gaussian noise is added to the bottom (large k) of the initial list than to the top (small k). After separately generating two randomly re-ranked lists using the same parameters, the ranked lists were compared for overlap using RRHO analysis. This synthetic data generation procedure resulted in ranked list comparisons that are reminiscent of true gene-expression data: with higher overlap in the most differentially expressed genes (top of a differential expression-based ranked list), and more random relationships in the non-differentially expressed genes (middle and/or bottom of the list) (User-Guide Figure 1A).

We generated a set of eight synthetic data ranked list comparisons, with list order randomness increasing from cases (i) to (viii) as the sigma parameters σ_{min} and σ_{max} are increased (User-Guide Figure 1A). We next analyzed these data sets using both the RRHO procedure and the Spearman rank correlation coefficient (ρ). User-Guide Figure 1A shows the RRHO map and rank-rank scatter plots of the synthetic data sets. The RRHO summary statistic (maximum log transformed hypergeometric P-value of the overlap map) and Spearman correlation coefficient show a correlated relationship (User-Guide Figure 1B, User-Guide Table S1). The RRHO overlap map provides more detailed information as to the pattern of overlap, especially in cases of weak but statistically significant correlated data. Comparisons between the overlap metrics from RRHO analysis and from the Spearman correlation coefficient for the actual gene-expression data analyzed in the manuscript can be found in User-Guide Table S2.

User-Guide Figure 1A:



User-Guide Figure 1B:



User-Guide Table S1:

Comparing RRHO to the Spearman rank correlation coefficient

Synthetic data set	*σ(k) min*	*σ(k) max*	*Spearman coefficient of correlation (ρ) **	*RRHO map max (log hypgeom P-value)*
i	300	2000	0.92	1887
ii	500	3500	0.79	1214
iii	1000	4000	0.68	816
iv	1000	5000	0.60	619
v	3000	5000	0.37	196
vi	5000	5000	0.26	85
vii	6000	10000	0.19	51
viii	10000	10000	0.08	11

* the P-value significance for the Spearman coefficient of correlation 3.3×10^{-15} for case viii and $< 2.2 \times 10^{-16}$ for all other cases based on the cor.test function of the R statistical package.

Recommendations for multiple hypothesis correction:

In general, multiple hypothesis correction is most relevant when trying to compare the significance of multiple maps from different experiments to one another directly and when dealing with weak overlap cases. The following discussion is based on our experience in using sample permutation-based correction and Benjamini-Yekutieli correction (from the Methods section), and 'list length correction' (Supplementary Methods) in combination with each other on gene-expression data. Combinations of different correction techniques on weak signals should be performed and interpreted with care since each data set may have different characteristics that could add artifacts in weak cases.

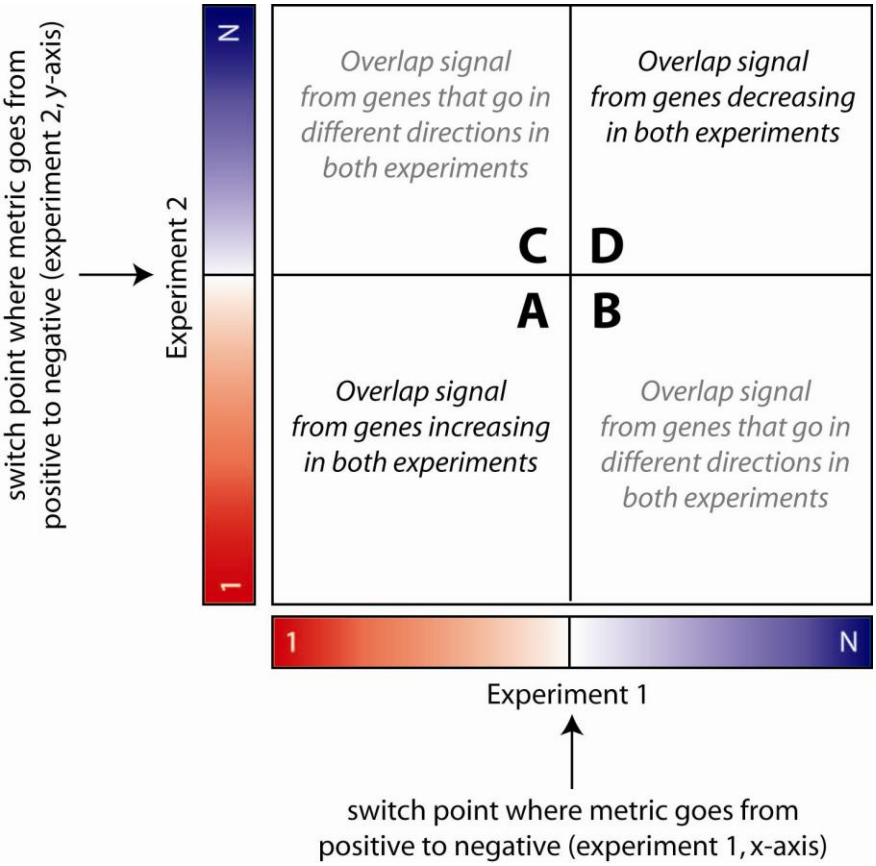
From our experience, the best order of operations for multiple hypothesis correction and map significance comparison is as follows. For all maps, perform the Benjamini-Yekutieli analytical correction, which will give an initial estimated correction. If the Benjamini-Yekutieli corrected hypergeometric map has most significant areas of absolute $\log_{10}$(P-value) intensity 15 or greater, then sample permutations are likely to be significant below a permutation frequency P-value of 0.01 (see User Guide Table S2 below). When the corrected map maxima are less than 15, perform sample permutations when enough samples in each class are available. If the two maps for comparison are made from gene lists that are of considerably different lengths, one can either (a) scale the hypergeometric maps using the list length correction method to account for length differences before applying a Benjamini-Yekutieli correction or (b) remake the hypergeometric maps using only genes common to all signatures in the set of comparisons making sure this restriction is not too limiting on the total gene number.

Sample permutation recommendations:

The following are practical guidelines for performing sample permutations. Sample permutations require high sample numbers in order to ensure that there are adequate sample label arrangements to recreate a random distribution of permutation results to compare to the true overlap case. Sample permutations require computation time as the 2-dimensional hypergeometric calculations need to be done for each shuffle of the sample labels. We recommend having enough samples to allow for 1000 permutations. This requires approximately 6 samples in each class for at least one of the experiments since $N_{permutations} = (n_A+n_B)! / (n_A! \cdot n_B!)$, where n_i is the number of samples in class i. We typically get the same results shuffling the samples in one or both of the experiments, thus we shuffle the labels in the experiment with the most samples and with a more even division of samples between the two classes. For a comparison of gene lists of length 5000, it takes approximately 15 seconds to calculate the 10,000 $((5000/50)^2)$ hypergeometric CDF results required to create one RRHO map with rank threshold step size of 50 using an Intel Xeon 3.2 GHz processor, thus requiring about 4 hours to create and analyze 1000 sample permutation maps on a single computational node.

Sample permutation P-value determination:
After RRHO maps have been created for all sample permutation comparisons, we employ a heuristic to ensure that the permutation cases that show a more significant overlap than the true RRHO map have overlap signal in the same area of the map. User-guide Figure 2 below shows a schematic that describes how we use the rank at which the differential expression metric in each experiment goes from positive to negative to define four separate quadrant-like regions of the map. Areas A and D correspond to overlapping genes that go in the same direction in both experiments. We compare the absolute maximum in these areas as a summary statistic to screen through permutation RRHO maps compared to the true RRHO map. The frequency at which permutation maps have a higher summary statistic than the true map is defined as the permutation P-value. If the observed pattern of overlap in the true RRHO map were much different than the typical pattern seen in the results shown in the manuscript, with the majority of signal in regions A and D, then another summary statistic might be more appropriate. When applying RRHO with genes ranked on unsigned (non-directional) differential expression statistics, we use the absolute maximum of the whole RRHO map as a summary statistic.

User-Guide Figure 2:



Schematic for areas of the RRHO map used to screen sample permutation maps for higher overlap in genes going in the same direction in both experiments. The region borders are defined by the switch points at which differential expression flips from upregulated to downregulated in each of the experiments. For permutation P-values we typically use either the maximum $-\log_{10}$(P-value) for region A (increasing in both experiments), for region D (decreasing), or the sum of the maximum $-\log_{10}$(P-value) in regions A and D (increasing and decreasing) as the map summary statistic (see for example the results in Fig. 4D).

Comparisons of raw heatmaps, multiple hypothesis corrected heatmaps and permutation P-values:

To demonstrate the relative effects of the different multiple hypothesis correction approaches, User-Guide Table S2 compares each correction approach applied to all of the RRHO maps from our manuscript that had enough samples to permit sample permutation analysis. The maximum $-\log_{10}$(P-value) of the RRHO map is listed to denote the strength of overlap between these pairs of experiments. For reference, we also included Spearman Rho rank correlation coefficient and its significance as an example of a rank-based correlation metric that has been previously characterized. In order to compare all of these RRHO maps to one another, it is important to note that the list length of genes measured in both platforms used to make the RRHO maps vary from 2882 to 13433 genes, so we need to apply a list length scaling correction. Here we chose to scale all comparisons to N = 5000 genes for reference. The maximum $-\log_{10}$(P-value) of the list length corrected maps are listed in the table. We next applied Benjamini-Yekutieli corrections to the RRHO maps and compared the maximum $-\log_{10}$(P-value) of these maps to permutation P-values derived from 1000 sample label shuffles of experiment 2. We find that permutation P-values for maps that have list length and Benjamini-Yekutieli corrected map maximum above 10 had a permutation P-value below 0.1 (most are below 0.05). Note that this estimated significance threshold would be different if a different list length reference for the $L_{shorter}/L_{longer}$ list length correction factor. We also note that the trend between corrected maximum map values and permutation P-values do have exceptions, and any results close to these significance thresholds should be interpreted with care. In summary, the Benjamini-Yekutieli correction is a good first estimate, but typically it is less conservative than permutation-based analysis.

User-guide Table S2:

Comparison of RRHO analysis summary statistics before and after the application of different multiple hypothesis correction approaches.

Manu-script figure	*Expt 1-sample number class A,B*	*Expt 2-sample number class A,B*	*Spearman Rho*	*Spearman Rho test P-value*	*list length*	*list length scaling factor (Nref = 5000)*	*RRHO map max*	*RRHO map max with list-length correction*	*RRHO map max with BY correction*	*RRHO map max with list length and BY correction*	**Permutation P-value**
2C	57,42	43,55	0.90	$< 10^{-16}$	13432	0.4	1856	742	1847	739	<0.001
3B (1)	4,4	5,5	0.28	$< 10^{-16}$	2882	1.7	26	45	21	36	0.044
3A	4,4	10,8	0.20	$< 10^{-16}$	11891	0.4	77	31	71	29	0.021
3F	7,16	10,5	0.03	0.58	5371	0.9	33	30	28	25	0.001
3B (2)	4,4	5,5	0.21	$< 10^{-16}$	2882	1.7	17	32	14	24	0.025
2F	57,42	23,62	0.17	$< 10^{-16}$	13432	0.4	63	25	57	23	0.05
3B (3)	4,4	5,5	0.15	5×10^{-15}	2882	1.7	12	21	7	11	0.095
3D	62,11	10,10	0.02	0.12	13432	0.4	24	5.2	18	7.1	0.102
2I	57,42	7,4	0.04	0.004	4691	1.1	9.3	10	4.3	4.7	0.14
3E	5,8	7,8	0.03	0.04	9093	0.5	10	5.1	5.2	2.6	0.033

The RRHO map max values are the maximum $-\log_{10}$(P-value) of the map from the indicated manuscript figure. The table is ordered on list-length and Benjamini-Yekutieli (BY) corrected max map values. The permutation P-value is calculated using the sum of areas A and D in User-Guide Figure 2 as detailed above. The Spearman Rho and Rho test P-value were calculated using the cor.test function in the statistical package R with input being the signed and log-transformed t-test P-values in two experiments used to rank genes for RRHO analysis (R Development Core Team, 2009).

R Development Core Team (2009) R: A Language and Environment for Statistical Computing (Vienna, Austria) Available at: http://www.R-project.org

False discovery rate for the overlapping gene list:

Once the ranks corresponding to the most statistically significant overlap between the two lists are determined, the false discovery rate (FDR) of the observed set of overlapping genes (*k*) can be determined by calculating

$$FDR_{\text{overlapping gene list}} = \frac{k_{expected}}{k_{observed}},$$

where $k_{expected} = \frac{s \times M}{N}$, *s* and *M* are the optimal rank thresholds, and *N* is the full gene list length.

In cases where the overall signature overlap is statistically significant, the FDR of the overlapping gene list is not always small. Thus in applications where the genes on the overlapping list will be studied individually, the FDR should be calculated to help guide interpretation.