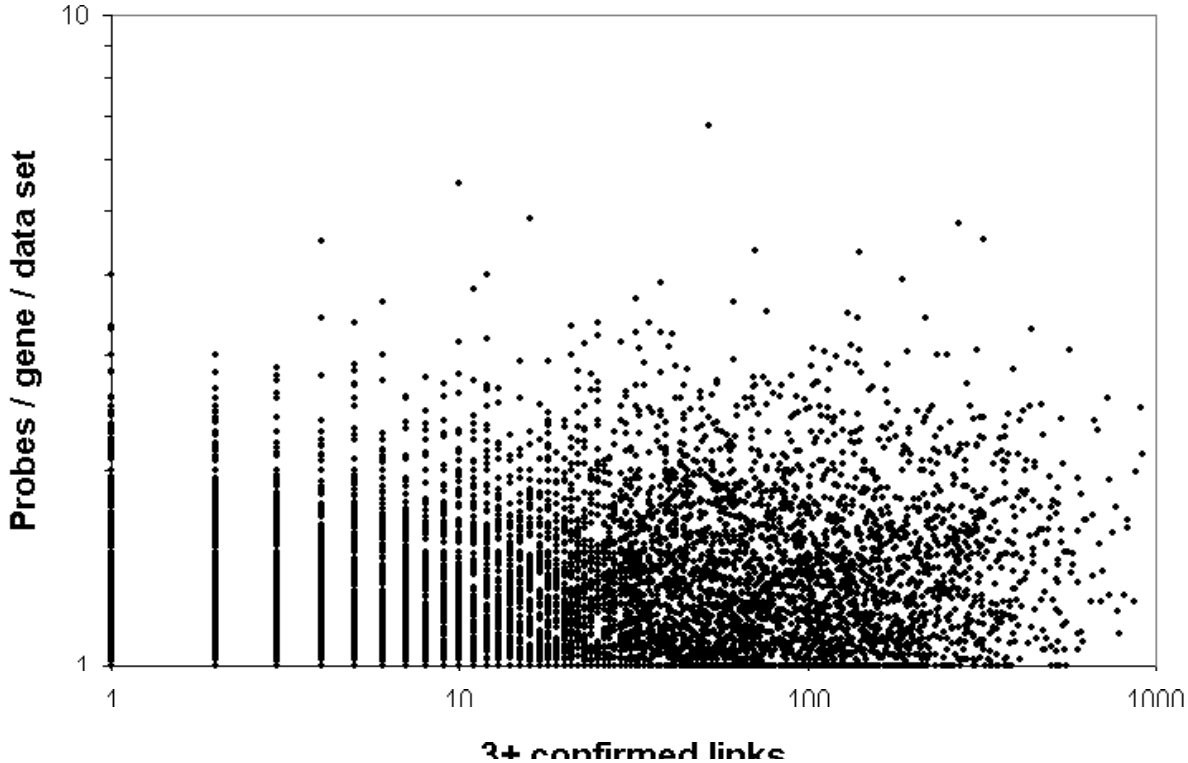
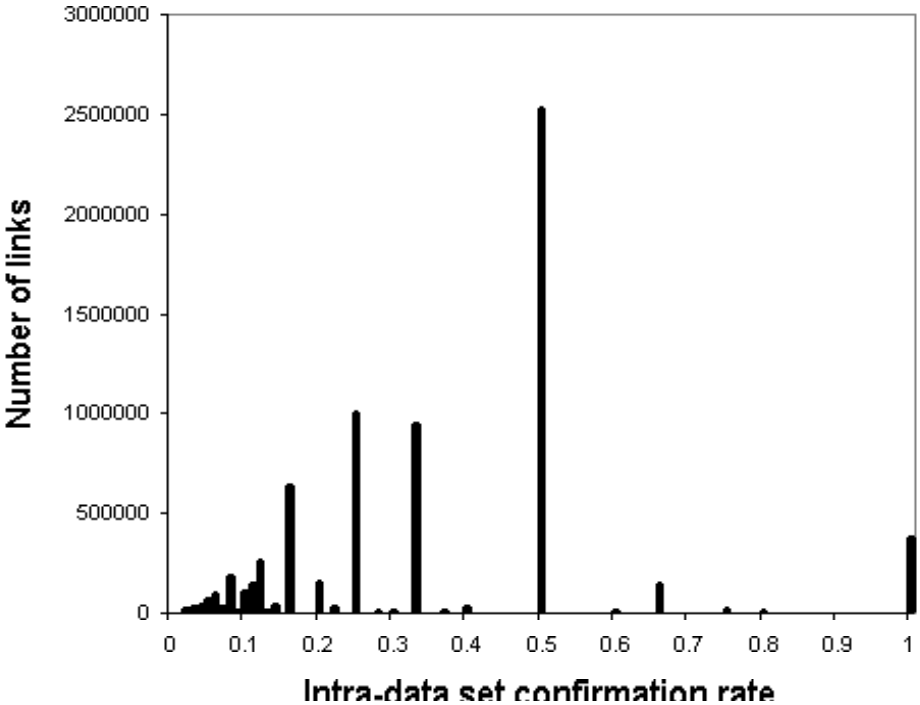
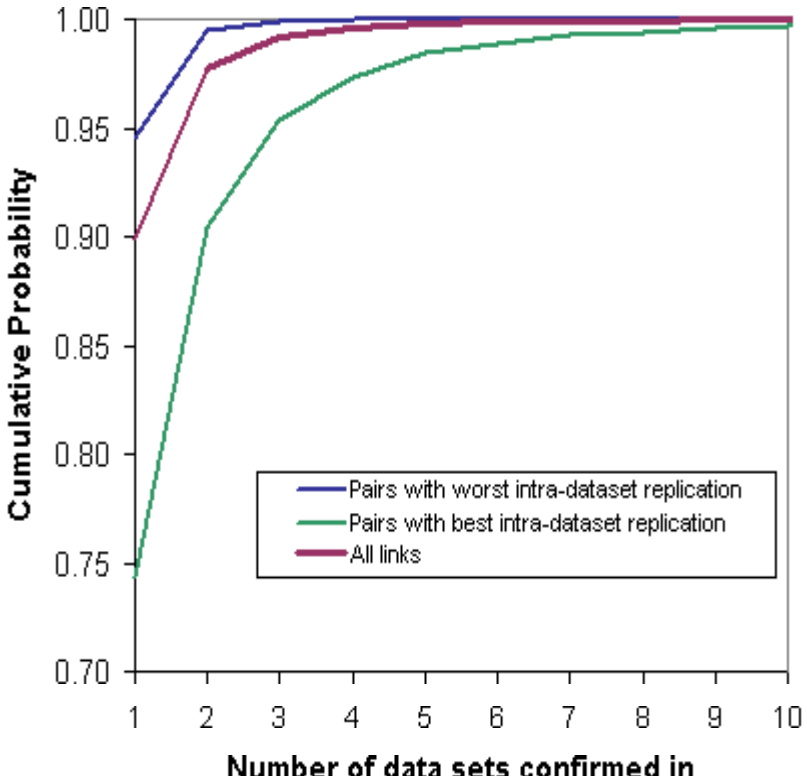
Supplementary figures



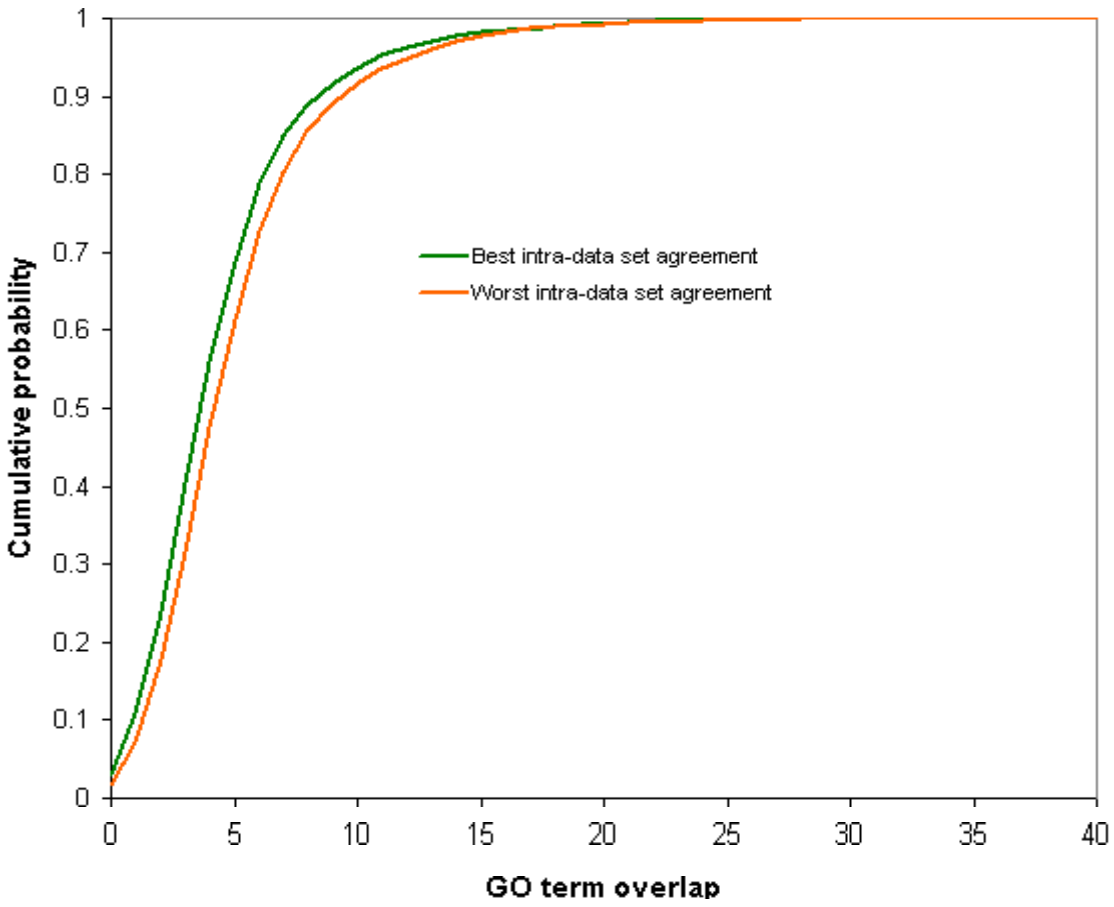
Supplementary Figure A. This figure shows the relationship between the number of links for a gene and the number of probes per data set for the gene. Note that both axes are logarithmic. A value of 2 on the y axis means that the gene has 2 probes per data set on average. Though not obvious from this figure, there is a low positive correlation between the number of links and the number of probes per gene per data set.



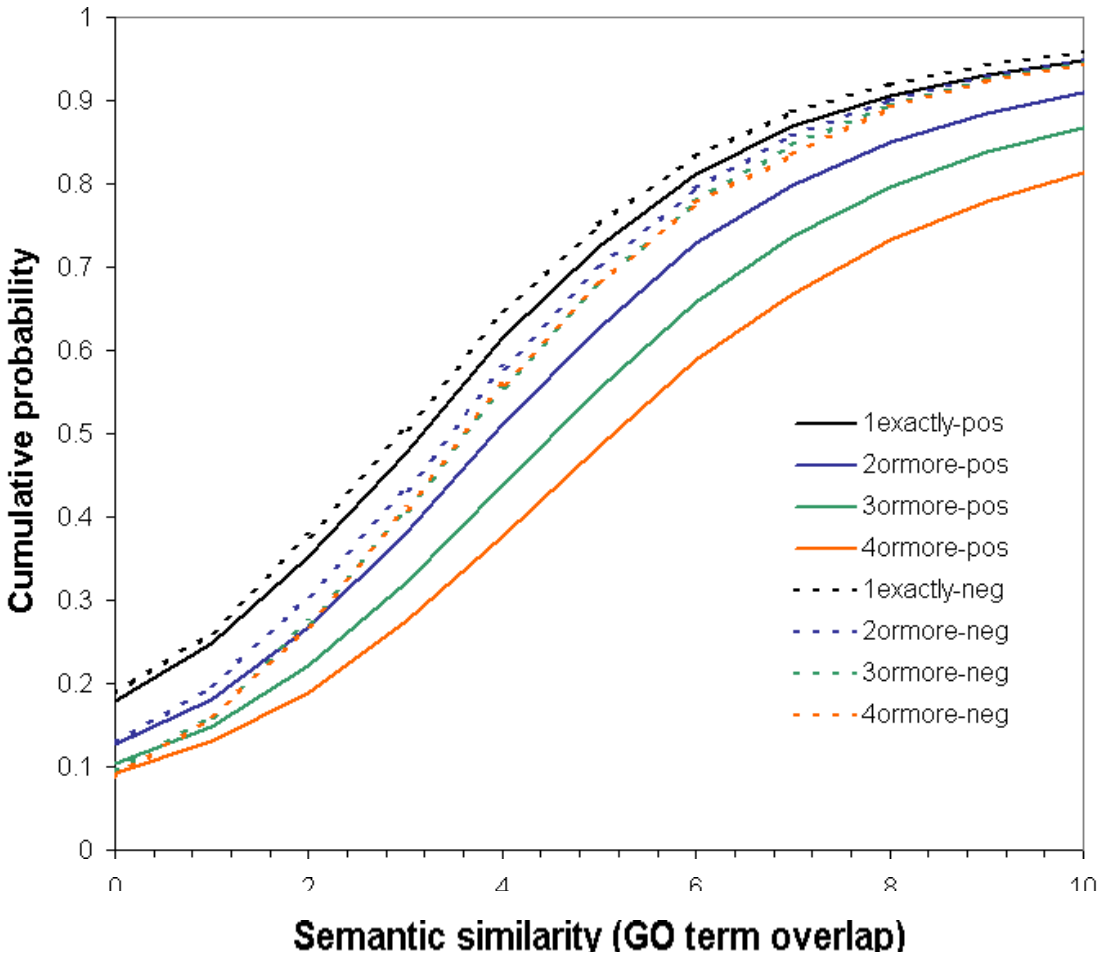
Supplementary Figure B: This figure summarizes the results of analyzing multiple occurrences of the same link within data sets. The intra-dataset confirmation rate is the number of tests for a link that were "successful", divided by the total number of tests for that link. For example, a link might be tested between two genes that have two probes representing them, each. This yields a total of four tests of correlation between the two genes. If two of these tests yielded links that passed our selection criteria, the intra dataset confirmation rate is 0.5



Supplementary Figure C: This graph shows that links that show high intra-dataset confirmation (see Supplementary-figureB) also tend to show high inter-dataset confirmation. The curve for "all links" here is essentially the same data shown in Figure 2A in the paper, plotted as a cumulative probability. As described in the main text of the supplement, "Worst" were those that were tested at least 10 times in a data set, but selected only once (e.g., a confirmation rate less than 0.1). The "best" links were those that were tested at least 4 times and yielded confirmation rates of 1.0. This determination is made on a per-data set basis, so a link might have a high intradataset confirmation rate in one data set but a low rate in another. In this situation, we use the best confirmation rate to represent the link. The graph thus shows that links that tend to have low intra-dataset confirmations also tend not to be confirmed by other data sets. One practial effect this has is that even though the "worst" data sets are numerous, they have a relatively small effect on the links we detect as being confirmed between data sets.



Supplementary Figure D: This figure is similar to Figure 4 in the paper. Here we are examining links that show good and bad intra-dataset confirmation (see legend to Supplementary figure C for definitions of "good" and "bad"). This graph shows that there is no evidence that links showing poor intra-dataset reproducibility are less "functionally informative" than ones showing high reproducibility within data sets. If anything the poorly reproducible links have slightly higher semantic simliarities.



Supplementary Figure E: This figure is similar to Figure 4 in the paper, but shows the analysis for positive correlation links separated from negative correlation links. Though both types of links show increases in the median semantic similarity with increasing confirmation, the effect is much less pronounced for the negative correlations, and appears to show little increase beyond 3 confirmations.