

Table 1. Pathway Coverage.

	Pathways Covered (total 1320)	Mean Coverage (proportion of Pathway Genes)	Median Coverage (proportion of Pathway Genes)	Mean Multiplicity
1500 data-driven gene list (after Steps 1-3)	659	0.12	0.08	7.73
S1500 gene set (after Step 1-4)	1320	0.26	0.25	10.38
S1500+ (v1)	1320	0.43	0.43	11.44
Random 1500	541 (443, 695)	0.07 (0.05, 0.09)	0.07 (0.05,0.09)	5.88 (5.60, 6.19)
Random 2739	852 (759, 946)	0.13 (0.11, 0.15)	0.12 (0.11,0.14)	5.87 (5.66, 6.18)
L1000 [32]	906	0.17	0.16	12.97

Note: Pathways covered are calculated relative to the 1320 canonical pathways (MSigDB genesets) in MSigDB version 4.0. The pathway level coverage is defined as fraction of genes from pathway that overlap selected gene set. Mean and Median coverage values are derived from pathway level coverage of 1320 canonical pathways from MSigDB (v4.0). The gene level multiplicity metric represents number of pathways a gene is part of. Mean multiplicity is computed using gene level multiplicity metrics across all selected genes. For “Random 1500” and “Random 2739” gene sets parenthesized values represent mean and range (min, max) across 20 alternative randomizations.