

Supplementary results of THD-Triclustering Technique

Table 1: Comparison of THD-Tricluster and TRICLUSTER associated with Process in terms of biological significance for Dataset 1 ($\delta=0.65$).

Cluster no.	For THD-Tricluster		For TRICLUSTER [1]	
	# of genes	GO terms associated with process	# of genes	GO terms associated with process
1	6103	Single organism process (n=3317 and p=1.53e-106), Single organism cellular process (n=2978 and p=8.91e-78), biological regulation (n=1495 and p=1.89e-71), regulation of cellular process (n=1225 and p=8.11e-57), cellular component organization (n=1882 and p=1.00e-56), response to stimulus (n=1008 and p=2.24e-56), single-organism metabolic process (n=1775 and p=1.40e-49),cellular response to stimulus (n=908 and p=2.93e-48)	66	trans- <i>port</i> with n=16, p=0.00332, localization with n=16,p=0.00453, pantothenate metabolism with n=2, p=0.00245, pantothenate biosynthesis with n=2 and p=0.00246
2	6101	Single organism process (n=3369 and p=1.04e-105), Single organism cellular process (n=2976 and p=5.29e-77), biological regulation (n=1494 and p=9.19e-71), regulation of biological process (n=1278 and p=8.51e-59), regulation of cellular process (n=1224 and p=4.20e-56), cellular component organization (n=1880 and p=1.01e-55), response to stimulus (n=1007 and p=1.97e-55), single-organism metabolic process (n=1775 and p=8.52e-50),localization (n=1275 and p=1.19e-47)	52	mRNA polyadenylation with n=2 and p=0.00826, G1/S transition of mitotic cell cycle with n=3 and p=0.00468
3	6102	Single organism process (n=3371 and p=8.53e-107), Single organism cellular process (n=2978 and p=5.62e-78), biological regulation (n=1495 and p=1.47e-71), regulation of biological process (n=1279 and p=1.36e-59), regulation of cellular process (n=1225 and p=6.66e-57), cellular component organization (n=1882 and p=7.63e-57), response to stimulus (n=1008 and p=1.90e-56), single-organism metabolic process (n=1775 and p=1.09e-49),cellular response to stimulus (n=908 and p=2.53e-48)	57	lipid transport with n=2 and p=0.0089
4	488	Single organism process (n=3370 and p=2.23e-106), Single organism cellular process (n=2977 and p=1.37e-77), biological regulation (n=1494 and p=9.19e-71), regulation of biological process (n=1278 and p=8.51e-59), regulation of cellular process (n=1224 and p=4.20e-56), cellular component organization (n=1881 and p=2.44e-56), response to stimulus (n=1007 and p=1.97e-55), single-organism metabolic process (n=1774 and p=3.33e-49),localization (n=1276 and p=2.17e-48)	97	physiological process with n=76 and p=0.0017, organelle organization and biogenesis with n=15 and p=0.00173, localization with n=21 and p=0.00537
5	6098	Single organism process (n=3369 and p=1.81e-106), Single organism cellular process (n=2976 and p=1.33e-77), biological regulation (n=1494 and p=4.39e-71), regulation of biological process (n=1279 and p=5.96e-60), regulation of cellular process (n=1225 and p=3.03e-57), cellular component organization (n=1880 and p=4.49e-56), response to stimulus (n=1007 and p=1.19e-55), single-organism metabolic process (n=1773 and p=6.13e-49),localization (n=1275 and p=6.75e-48)	51	carbohydrate biosynthesis with n=3 and p=0.00946, ubiquitin cycle with n=3 and p=0.00346, protein polyubiquitination with n=2 and p= 0.00796

Table 2: Comparison of THD-Tricluster and TRICLUSTER associated with Function in terms of biological significance for Dataset 1 ($\delta=0.65$).

Cluster no.	For THD-Tricluster		For TRICLUSTER [1]	
	# of genes	GO terms associated with function	# of genes	GO terms associated with function
1	6103	Protein binding (n=606 and p=6.74e-32), transferase activity (n=792 and p=1.01e-20), DNA binding (n=373 and p=2.01e-13), pyrophosphatase activity (n=373 and p=4.30e-12), hydrolase activity, acting on acid anhydrides (n=373 and p=4.30e-12), hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides(n=373 and p= 4.30e-12), sequence-specific DNA binding(n=239 and p=2.49e-11), nucleoside-triphosphatase activity (n=352 and p=9.26e-11), phosphotransferase activity, alcohol group as acceptor(n=176 and p=9.88e-09), oxidoreductase activity (n=272 and p=1.37e-07), macromolecular complex binding (n=198 and p=1.54e-07)	66	lipid transporter activity with n=2 and p=0.00833, ubiquitin conjugating enzyme activity with n=2 and p=0.00833
2	6101	Protein binding (n=605 and p=6.07e-31), transferase activity (n=791 and p=3.03e-20), DNA binding (n=373 and p=1.80e-13), pyrophosphatase activity (n=373 and p=3.87e-12), hydrolase activity, acting on acid anhydrides (n=373 and p=3.87e-12), hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides(n=373 and p= 4.30e-12), sequence-specific DNA binding(n=239 and p=2.49e-11), nucleoside-triphosphatase activity (n=352 and p=8.42e-11), kinase activity (n=192 and p=6.60e-09), hydrolase activity (n=809 and p=1.68e-08), phosphotransferase activity, alcohol group as acceptor (n=175 and p=1.13e-07)	52	phosphatase regulator activity with n=2 and p=0.00397, protein phosphatase regulator activity with n=2,p=0.00397
3	6102	Protein binding (n=606 and p=6.11e-32), transferase activity (n=792 and p=9.16e-21), DNA binding (n=373 and p=1.90e-13), pyrophosphatase activity (n=373 and p=4.08e-12), hydrolase activity, acting on acid anhydrides (n=373 and p=4.08e-12), hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides(n=373 and p= 4.08e-12), sequence-specific DNA binding(n=239 and p=2.40e-11), nucleoside-triphosphatase activity (n=352 and p=8.83e-11), kinase activity (n=193 and p=5.12e-10), hydrolase activity (n=809 and p=1.81e-08), phosphotransferase activity, alcohol group as acceptor (n=176 and p=9.60e-09)	57	lipid transporter activity with n=2 and p=0.00627, antioxidant activity with n=2 and p=0.00797, oxidoreductase activity with n=7 and p=0.00239
4	488	Protein binding (n=606 and p=5.54e-32), transferase activity (n=791 and p=3.03e-20), DNA binding (n=373 and p=1.90e-13), pyrophosphatase activity (n=373 and p=3.87e-12), hydrolase activity, acting on acid anhydrides (n=373 and p=3.87e-12), hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides(n=373 and p= 3.87e-12), sequence-specific DNA binding(n=239 and p=2.31e-11), nucleoside-triphosphatase activity (n=352 and p=8.42e-11), kinase activity (n=193 and p=4.97e-10), hydrolase activity (n=809 and p=1.68e-08), phosphotransferase activity, alcohol group as acceptor (n=176 and p=9.34e-09)	97	deaminase activity with n=2 and p=0.00804, hydrolase activity, acting on carbon-nitrogen but not peptide, bonds with n=4 and p=0.00918, MAP kinase activity with n=2 and p=0.00209, receptor signaling protein serine/threonine kinase activity with n=2, p=0.00964
5	6098	Protein binding (n=606 and p=4.14e-32), transferase activity (n=791 and p=2.24e-20), DNA binding (n=373 and p=1.54e-13), pyrophosphatase activity (n=373 and p=3.32e-12), hydrolase activity, acting on acid anhydrides (n=373 and p=3.32e-12), hydrolase activity, acting on acid anhydrides, in phosphorus-containing anhydrides(n=373 and p= 3.87e-12), sequence-specific DNA binding(n=239 and p=2.06e-11), nucleoside-triphosphatase activity (n=352 and p=7.29e-11), kinase activity (n=193 and p=4.52e-10), hydrolase activity (n=809 and p=1.68e-08), phosphotransferase activity, alcohol group as acceptor (n=176 and p=8.58e-09)	51	not known

Table 3: Comparison of THD-Tricluster and TRICLUSTER associated with Cellular component in terms of biological significance for Dataset 1 ($\delta=0.65$).

For THD-Tricluster			For TRICLUSTER [1]	
Cluster no.	# of genes	GO terms associated with component	# of genes	GO terms associated with component
1	6103	organelle (n=3962 and p=1.63e-97), intracellular organelle (n=3957 and p=1.66e-97), membrane-bounded organelle (n=3620 and p=2.13e-80), intracellular membrane-bounded organelle (n=3618 and p=3.20e-80), cell part (n=4888 and p=9.60e-67), cell (n=4888 and p=9.60e-67), protein complex (n=1433 and p= 5.48e-58)	66	Golgi vesicle with n=2 and p=0.00729
2	6101	organelle (n=3960 and p=6.75e-97), intracellular organelle (n=3955 and p=6.89e-97), membrane-bounded organelle (n=3619 and p=2.43e-80), intracellular membrane-bounded organelle (n=3617 and p=3.67e-80), cell part (n=4886 and p=2.00e-66), cell (n=4886 and p=2.00e-66), protein complex (n=1432 and p=2.08e-57)	52	not known
3	6102	organelle (n=3962 and p=8.22e-98), intracellular organelle (n=3957 and p=8.37e-98), membrane-bounded organelle (n=3620 and p=1.20e-80), intracellular membrane-bounded organelle (n=3618 and p=1.81e-80), cell part (n=4888 and p=4.07e-67), cell (n=4888 and p=4.07e-67), protein complex (n=1433 and p=4.39e-58)	57	microsome with n=2, p=0.00627, vesicular fraction with n=2, p=0.00627, cytoplasm with n=41, p=0.00052, peroxisome with n=3, p=0.00929, microbody with n=3, p=0.00929
4	488	organelle (n=3961 and p=1.67e-97), intracellular organelle (n=3956 and p=1.70e-97), membrane-bounded organelle (n=3619 and p=2.43e-80), intracellular membrane-bounded organelle (n=3617 and p=3.67e-80), cell part (n=4887 and p=5.88e-67), cell (n=4887 and p=5.88e-67), protein complex (n=1432 and p=2.08e-57)	97	membrane with n=29, p=9.36e-06, cell with n=86, p=0.0003, vacuolar membrane with n=6, p=0.0015, endoplasmic reticulum with n=13, p=0.00112, cytoplasm with n=63, p=0.00169 intracellular with n=79, p=0.00209, integral to endoplasmic reticulum membrane with n=3, p=0.00328, nuclear envelope-endoplasmic reticulum network with n=6, p=0.00488, endoplasmic reticulum membrane with n=6, p=0.00289
5	6098	organelle (n=3959 and p=3.46e-97), intracellular organelle (n=3954 and p=3.54e-97), membrane-bounded organelle (n=3617 and p=5.64e-80), intracellular membrane-bounded organelle (n=3615 and p=8.50e-80), cell part (n=4885 and p=5.20e-67), cell (n=4885 and p=5.20e-67), protein complex (n=1431 and p=6.28e-57)	51	not known

Table 4: Comparison of THD-Tricluster and TRICLUSTER associated with Process in terms of biological significance for Dataset 1 ($\delta=0.85$).

Cluster no.	For THD-Tricluster		For TRICLUSTER [1]	
	# of genes	GO terms associated with process	# of genes	GO terms associated with process
1	457	Cytoplasmic translation (n=45 and p=1.66e-13), organonitrogen compound metabolic process (n=127 and p=3.65e-06), oxoacid metabolic process (n=52 and p=6.38e-06), cellular protein metabolic process (n=149 and p=7.26e-05), single organism metabolic process (n=159 and p=0.00828), organonitrogen compound biosynthetic process (n=103 and p=0.00120)	66	trans- <i>port</i> with n=16, p=0.00332, localization with n=16,p=0.00453, pantothenate metabolism with n=2, p=0.00245, pantothenate biosynthesis with n=2 and p=0.00246
2	238	Cytoplasmic translation (n=29 and p=4.04e-10), organonitrogen compound metabolic process (n=79 and p=1.04e-06), single organism process (n=164 and p=1.30e-06), cellular respiration (n=17 and p=1.32e-06), generation of precursor metabolites and energy (n=22, p= 4.66e-06), purine ribonucleoside triphosphate metabolic process (n=15, p=1.25e-05), ATP metabolic process (n=14 and p=6.78e-05), nucleoside triphosphate metabolic process (n=15 and p=7.40e-05)	52	mRNA polyadenylation with n=2 and p=0.00826, G1/S transition of mitotic cell cycle with n=3 and p=0.00468
3	196	Small molecule metabolic process (n=57 and p=1.62e-12), cytoplasmic translation (n=29 and p=2.28e-12), organonitrogen compound metabolic process (n=79 and p=9.40e-12), metabolic process (n=155 and p=2.07e-11),cellular metabolic process (n=151 and p=5.41e-11), oxoacid metabolic process (n=37 and p=9.72e-10), generation of precursor metabolites and energy (n=22, p= 1.12e-07), primary metabolic process (n=138 and p=1.86e-07) organic substance metabolic process (n=143 and p=2.45e-07), single organism process (n=140 and p=3.76e-07), cellular respiration (n=16, p=6.15e-07)	57	lipid transport with n=2 and p=0.0089
4	488	Small molecule metabolic process (n=91 and p=7.59e-08), carboxylic acid metabolic process (n=52 and p=3.91e-05), cytoplasmic translation (n=31 and p=0.00104), oxoacid metabolic process (n=37 and p=9.72e-10), single organism process (n=312 and p=1.53e-07), single organism cellular process (n=281 and p=1.54e-06), generation of precursor metabolites and energy (n=29 and p=0.00050), pyridine nucleotide metabolic process (n=17 and p=0.00532), nucleobase containing small molecule metabolic process (n=36 and p=0.00015)	97	physiological process with n=76 and p=0.0017, organelle organization and biogenesis with n=15 and p=0.00173, localization with n=21 and p=0.00537
5	85	Cytoplasmic translation (n=14 and p=6.80e-06), cellular amide metabolic process (n=26 and p=0.00041), peptide metabolic process (n=23 and p=0.00495), organonitrogen compound metabolic process (n= 31 and p=0.00773)	51	carbohydrate biosynthesis with n=3 and p=0.00946, ubiquitin cycle with n=3 and p=0.00346, protein polyubiquitination with n=2 and p=0.00796

Table 5: Comparison of THD-Tricluster and TRICLUSTER associated with Function in terms of biological significance for Dataset 1 ($\delta=0.85$).

For THD-Tricluster			For TRICLUSTER [1]	
Cluster no.	# of genes	GO terms associated with function	# of genes	GO terms associated with function
1	457	Structural constituent of ribosome (n=52 and p=3.02e-14), structural molecule activity (n=62 and p=5050e-12)	66	lipid transporter activity with n=2 and p=0.00833, ubiquitin conjugating enzyme activity with n=2 and p=0.00833
2	238	oxidoreductase activity (n=30 and p=2.29e-06), structural constituent of ribosome (n=26 and p=6.45e-06), structural molecule activity (n=29, p=0.00073), hydrogen ion transmembrane transporter activity (n=11, p=0.00067), isocitrate dehydrogenase activity (n=4 and p=0.00116)	52	phosphatase regulator activity with n=2 and p=0.00397, protein phosphatase regulator activity with n=2,p=0.00397
3	196	oxidoreductase activity (n=30 and p=2.29e-06), structural constituent of ribosome (n=26 and p=6.45e-06), structural molecule activity (n=29, p=0.00073), electron carrier activity (n=7 and p=0.00024), isocitrate dehydrogenase activity (n=4 and p=0.00116), hydrogen ion transmembrane transporter activity (n=10, p=0.00080), succinate dehydrogenase (ubiquinone) activity (n= 4 and p=0.00154), unfolded protein binding (n=10 and p=0.00333), oxidoreductase activity acting on CH-OH group of donors (n=10 and p=0.00978)	57	lipid transporter activity with n=2 and p=0.00627, antioxidant activity with n=2 and p=0.00797, oxidoreductase activity with n=7 and p=0.00239
4	488	oxidoreductase activity (n=44 and p=4.23e-05), structural molecule activity (n=46 p=0.00253), structural constituent of ribosome (n=33 and p=0.00958)	97	deaminase activity with n=2 and p=0.00804, hydrolase activity, acting on carbon-nitrogen but not peptide, bonds with n=4 and p=0.00918, MAP kinase activity with n=2 and p=0.00209, receptor signaling protein serine/threonine kinase activity with n=2, p=0.00964
5	85	structural constituent of ribosome (n=17 and p=8.69e-08), structural molecule activity (n=19 and p=1.28e-06), repressing transcription factor (n=4 and p=0.00354)	51	not known

Comparison of the results of THD-Tricluster with that of TRICLUSTER [1] in terms of biological processes, molecular functions and cellular component for Dataset 1 for $\delta=0.65$ and $\delta=0.85$ are given in Tables 1-6.

References

- [1] Lizhuang Zhao and Mohammed J Zaki. Tricluster: an effective algorithm for mining coherent clusters in 3d microarray data. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 694–705. ACM, 2005.

Table 6: Comparison of THD-Tricluster and TRICLUSTER associated with Cellular component in terms of biological significance for Dataset 1 ($\delta=0.85$).

For THD-Tricluster			For TRICLUSTER [1]	
Cluster no.	# of genes	GO terms associated with component	# of genes	GO terms associated with component
1	457	cytosolic ribosome (n=149 and p=5.07e-17), cytosolic part (n=56 and p=1.61e-16), ribosomal subunit (n=54 and p=5.13e-14), intracellular (n=376 and p=1.86e-05), ribonucleoprotein complex(n=85 and p=2.78e-05), macromolecular complex(n=179 and p=0.00338)	66	Golgi vesicle with n=2 and p=0.00729
2	238	cytoplasm (n=190, p=5.052e-13), cytosolic ribosome (n=29 and p=1.24e-10), intracellular part (n=215 and p=9.58e-10), intracellular (n=215 and p=1.47e-09), oxidoreductase (n=10 and p=8.36e-06), mitochondrial protein complex (n= 16 and p=0.00027), mitochondrial respiratory chain (n=8 and p=0.00031), intracellular organelle (n=178 and p=0.00013)	52	not known
3	196	cytoplasm (n=169, p=2.39e-18), intracellular part (n=185, p=1.17e-13), intracellular (n=185 and p=1.81e-13), cytosolic ribosome (n=29 and p=7.04e-13),intracellular organelle (n=158 and p=3.64e-08), oxidoreductase complex (n=11 and p=6.90e-08), macromolecular complex (n=98 and p=9.17e-07)	57	microsome with n=2, p=0.00627, vesicular fraction with n=2, p=0.00627, cytoplasm with n=41, p=0.00052, peroxisome with n=3, p=0.00929, microbody with n=3, p=0.00929
4	488	cytoplasm (n=356, p=2.72e-13), cytoplasmic part (n=288, p=2.01e-12), mitochondrion (n=125 and p=6.43e-06), cytosolic part (n=40 and p=1.38e-05), external encapsulating structure (n=20 and p=0.00227), fungal type cell-wall (n=20 and p=0.00227)	97	membrane with n=29, p=9.36e-06, cell with n=86, p=0.0003, vacuolar membrane with n=6,p=0.0015, endoplasmic reticulum with n=13, p=0.00112, cytoplasm with n=63, p=0.00169 intracellular with n=79,p=0.00209, integral to endoplasmic reticulum membrane with n=3, p=0.00328, nuclear envelope-endoplasmic reticulum network with n=6,p=0.00488, endoplasmic reticulum membrane with n=6, p=0.00289
5	85	cystolic ribosome (n=15 and p=2.28e-07), cystolic part (n=18 and 2.01e-08), ribosome (n= 18 and p=1.60e-05),cystolic small ribosomal subunit (n=10 and p=3.76e-07), intracellular part (n=78 and p=0.00072), organelle (n=72 and p=4.97e-05), cell (n=81 and p=0.00014)	51	not known