Disrupted PPI Subnetworks in Cancer: Analysis of Prostate Cancer Subtype Expression Data and TCGA Cancer Mutation Data

LL of this.getSubgroupMeeting() of this.getLab() of this.getInstitution()

2012-06-18

Overview: Cancer Expression Samples

- Expression data for four subtypes of prostate cancer
  - ERG+ and SPOP+
  - ERG+ only
  - SPOP+ only
  - Negative for both
- Comparisons
  - (ERG+ and SPOP+) vs. negative
  - ERG+ vs. negative
  - SPOP+ vs. negative

Overview: Methods

- Find which genes are differentially expressed between the two conditions
- Find which of them are connected in the protein-protein interactome
- Results:
  - Subnetwork size distribution
  - % of interactome hubs/bottlenecks disrupted in cancer
- Comparison to datasets with randomized expression values
  - Each gene's expression value is randomly assigned to another gene (uniform probability distribution)

Overview: Cancer Mutation Samples

- Starting with exome point mutations of 8 TCGA cancers (verified by mapping coordinates to consensus CDS regions)
  - AML, BRCA, COAD, EMC, GBM, OV, PRAD, READ
- Take all mutations that occur in at least 2% of samples
  - Lowest frequency for which each cancer had at least one mutation
- Map to PPI interactome
- Results:
  - Subnetwork size distribution
  - % of interactome hubs/bottlenecks disrupted in cancer
- Randomized dataset generation is ongoing

Prostate Cancer Subtypes: Subnetwork Size Distributions

(ERG+ and SPOP+) vs. (ERG- and SPOP-)

13,618 genes mapped to super-network
13,507 subnetworks

Subnetwork distribution:

Subnetwork size	Count
1	13,434
2	59
3	9
4	2
6	1
12	1
13	1

(ERG+) vs. (ERG- and SPOP-)

8153 genes mapped to super-network
6775 subnetworks

Subnetwork distribution:

Subnetwork size	Count
1	6737
2	35
3	2
1340	1

(SPOP+) vs. (ERG- and SPOP-)

9,368 genes mapped to super-network
8,991 subnetworks

Subnetwork distribution:

Subnetwork size	Count
1	8,944
2	39
3	3
4	3
5	1
320	1

Prostate Cancer Subtypes: Hub Summary

ERG+ and SPOP+

Data	Hubs	% of HPRD hubs	% of PPI-mapped nodes	Max degree	Min degree	Avg degree
Actual	268	14.18%	5.44%	172	10	18.1903
Avg random	134.8	7.13%	3.41%	41	10	14.77502

ERG+

Data	Hubs	% of HPRD hubs	% of PPI-mapped nodes	Max degree	Min degree	Avg degree
Actual	42	2.22%	2.12%	35	10	13.5
Avg random	184.4	9.76%	4.05%	56	10	15.42844

SPOP+

Data	Hubs	% of HPRD hubs	% of PPI-mapped nodes	Max degree	Min degree	Avg degree
Actual	218	11.53%	6.57%	269	10	18.1239
Avg random	161.6	8.55%	3.81%	49.9	10	15.13297

Prostate Cancer Subtypes: Bottleneck Summary

ERG+ and SPOP+

Data	Bnecks	% of HPRD bnecks	% of PPI-mapped nodes	Max betweenness	Min betweenness	Avg betweenness
Actual	305	16.14%	6.19%	1083602.00300	12177.26355	34731.77218
Avg random	169	8.94%	4.28%	131278.88199	12194.73229	24380.17979

ERG+

Data	Bnecks	% of HPRD bnecks	% of PPI-mapped nodes	Max betweenness	Min betweenness	Avg betweenness
Actual	85	4.50%	4.29%	90429.08842	12290.44124	22591.96137
Avg random	213.7	11.31%	4.71%	194765.29938	12179.9837	26615.43565

SPOP+

Data	Bnecks	% of HPRD bnecks	% of PPI-mapped nodes	Max betweenness	Min betweenness	Avg betweenness
Actual	242	12.80%	7.30%	2055159.90400	12177.26355	40647.30257
Avg random	192.4	10.18%	4.54%	199359.88900	12179.33809	25641.21837

TCGA Exome Mutations: Subnetwork Size Distribution

- A trend of extremes: Each cancer had a single large subnetwork, a cornucopia of singletons, and a small number of 2, 3, and 4-node subnets

Cancer	Size of Dominant subnetwork	Singletons	Total Nodes	% in dominant subnet	% singletons
AML	N/A	72	72	0.00%	100.00%
BRCA	517	695	1238	41.76%	56.14%
COAD	288	296	597	48.24%	49.58%
EMC	8889	1219	10112	87.91%	12.05%
GBM	N/A	29	32	0.00%	90.63%
OV	N/A	1	1	0.00%	100.00%
PRAD	175	210	387	45.22%	54.26%
READ	1084	1212	2349	46.15%	51.60%

TCGA Exome Mutations: Hub and Bottleneck Summary

Data	Hubs	% of HPRD hubs
AML	11	0.58%
BRCA	136	7.20%
COAD	92	4.87%
EMC	1170	61.90%
GBM	11	0.58%
OV	1	0.05%
PRAD	41	2.17%
READ	326	17.25%

Data	Bnecks	% of HPRD bnecks
AML	9	0.48%
BRCA	131	6.93%
COAD	83	4.39%
EMC	1135	60.05%
GBM	9	0.48%
OV	1	0.05%
PRAD	47	2.49%
READ	305	16.14%

Followup

- Generate random mutation datasets to compare to findings from the TCGA exome mutations
- So far, a preliminary investigation has been made to catalog the pathways that intersect mutated/ differentially expressed nodes
  - Using DAVID Bioinformatics portal
  - To be completed
- Complete computation of hub and bottleneck statistics for the TCGA exome mutation datasets