

Latest Prostate Recurrent Mutation Results

Lucas Lochovsky

2014-01-27

Overview

Results

- Combine multiple prostate datasets
 - Berger, Grasso, Korbel, Barbieri, Baca
- Find significantly mutated KEGG pathways and HPRD interaction pairs
 - Filter results list compared to last time

Methods

- Convert data from LARVA exon runs into gene numbers
 - Apply to prostate and glioma datasets

LARVA(all prostate, genes) With Pathway Membership Data

Gene list sorted by number of pathways (top 20)

| gene name | num_samp | num_var | # pathways |
|-----------|----------|---------|------------|
| PIK3CA | 6 | 1 | 34 |
| HRAS | 2 | 0 | 31 |
| BRAF | 3 | 0 | 26 |
| TP53 | 32 | 5 | 21 |
| GSK3B | 3 | 0 | 18 |
| MAPK10 | 2 | 0 | 18 |
| EGF | 2 | 0 | 15 |
| CTNNB1 | 3 | 0 | 14 |
| MAPK14 | 2 | 0 | 14 |
| PLCB3 | 3 | 1 | 13 |
| ADCY3 | 3 | 0 | 13 |
| PLCB1 | 2 | 0 | 13 |
| ADCY8 | 4 | 0 | 12 |
| ADCY9 | 3 | 0 | 12 |
| PLA2G4A | 2 | 0 | 12 |
| PTEN | 8 | 0 | 11 |
| CACNA1C | 4 | 0 | 11 |
| ACTB | 2 | 0 | 11 |
| ACTG1 | 2 | 0 | 11 |
| ADCY1 | 2 | 0 | 11 |

Gene list sorted by number of samples mutated (top 20)

| gene name | num_samp | num_var | # pathways |
|-----------|----------|---------|------------|
| TP53 | 32 | 5 | 21 |
| TTN | 23 | 0 | 2 |
| RYR3 | 10 | 0 | 2 |
| ATM | 9 | 0 | 3 |
| DMD | 8 | 1 | 4 |
| PTEN | 8 | 0 | 11 |
| CTNNA2 | 8 | 0 | 6 |
| RYR1 | 7 | 0 | 2 |
| PIK3CA | 6 | 1 | 34 |
| RYR2 | 6 | 0 | 5 |
| CACNA1E | 6 | 0 | 3 |
| COL6A3 | 6 | 0 | 2 |
| TUBA3C | 6 | 0 | 2 |
| AR | 5 | 2 | 3 |
| GRIN2B | 5 | 0 | 6 |
| LAMB1 | 5 | 0 | 4 |
| C3 | 5 | 0 | 3 |
| DGKB | 4 | 1 | 3 |
| IDH1 | 4 | 1 | 3 |
| CARD11 | 4 | 1 | 2 |

LARVA(all prostate, KEGG)

List of pathways with significantly enriched/depleted number of recurrently mutated exons.

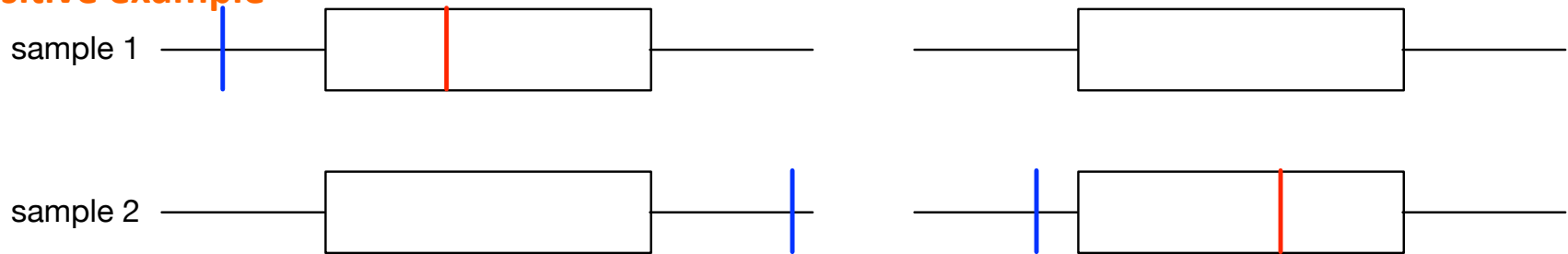
| KEGG pathway | nsamp | nannot | ngene | nvar | # annotations mutated | rand avg | # annotations mutated | p-value | Enrichment/Depletion |
|---|-------|--------|-------|------|-----------------------|----------|-----------------------|-------------|----------------------|
| kegg huntingtons disease.txt | 71 | 12 | 23 | 8 | | 3.90 | | 7.39E-161 | Enr |
| kegg endometrial cancer.txt | 45 | 18 | 14 | 6 | | 2.80 | | 5.05E-92 | Enr |
| kegg hypertrophic cardiomyopathy hcm.txt | 70 | 14 | 26 | 1 | | 3.00 | | 4.62E-35 | Enr |
| kegg mapk signaling pathway.txt | 95 | 26 | 42 | 5 | | 8.30 | | 2.05E-30 | Enr |
| kegg neurotrophin signaling pathway.txt | 48 | 14 | 13 | 6 | | 5.50 | | 2.74E-26 | Enr |
| kegg adherens junction.txt | 56 | 10 | 20 | 0 | | 3.60 | | 2.50E-22 | Enr |
| kegg small cell lung cancer.txt | 64 | 11 | 18 | 6 | | 3.60 | | 3.40E-16 | Enr |
| kegg arrhythmogenic right ventricular cardiomyopathy arvc.txt | 69 | 10 | 28 | 1 | | 4.40 | | 4.98E-10 | Enr |
| kegg axon guidance.txt | 78 | 7 | 27 | 1 | | 12.90 | | 7.97E-09 | Dep |
| kegg dilated cardiomyopathy.txt | 77 | 15 | 30 | 1 | | 6.30 | | 1.05E-08 | Enr |
| kegg melanoma.txt | 49 | 15 | 13 | 6 | | 3.70 | | 1.82E-08 | Enr |
| kegg tight junction.txt | 74 | 16 | 31 | 1 | | 9.20 | | 2.60E-08 | Enr |
| kegg pancreatic cancer.txt | 42 | 12 | 12 | 6 | | 2.90 | | 4.33E-08 | Enr |
| kegg viral myocarditis.txt | 55 | 6 | 21 | 1 | | 2.70 | | 1.28E-07 | Enr |
| kegg olfactory transduction.txt | 96 | 38 | 41 | 3 | | 78.60 | | 1.23E-06 | Dep |
| kegg prostate cancer.txt | 60 | 16 | 16 | 8 | | 5.50 | | 3.43E-06 | Enr |
| kegg ubiquitin mediated proteolysis.txt | 52 | 2 | 12 | 0 | | 6.3 | | 0.0003509 | Dep |
| kegg p53 signaling pathway.txt | 42 | 8 | 7 | 6 | | 3.3 | | 0.000458012 | Enr |
| kegg pathways in cancer.txt | 134 | 30 | 60 | 8 | | 22.4 | | 0.001173363 | Enr |
| kegg hematopoietic cell lineage.txt | 45 | 7 | 17 | 1 | | 3.2 | | 0.002089279 | Enr |
| kegg abc transporters.txt | 50 | 5 | 16 | 0 | | 3.7 | | 0.002278175 | Enr |
| kegg cell adhesion molecules cams.txt | 67 | 15 | 26 | 2 | | 11.2 | | 0.003320943 | Enr |
| kegg apoptosis.txt | 44 | 8 | 7 | 6 | | 4.2 | | 0.003320943 | Enr |
| kegg oocyte meiosis.txt | 47 | 3 | 12 | 2 | | 9.5 | | 0.003600326 | Dep |
| kegg leukocyte transendothelial migration.txt | 53 | 12 | 16 | 1 | | 6.8 | | 0.005393725 | Enr |
| kegg cell cycle.txt | 57 | 9 | 15 | 5 | | 4.8 | | 0.01300724 | Enr |
| kegg purine metabolism.txt | 61 | 4 | 17 | 0 | | 8.7 | | 0.02760817 | Dep |
| kegg vascular smooth muscle contraction.txt | 65 | 4 | 27 | 2 | | 11.4 | | 0.03483315 | Dep |
| kegg glioma.txt | 43 | 10 | 8 | 6 | | 5.2 | | 0.04243526 | Enr |

LARVA Network Analysis Refinement

- Use pairs of interacting genes as annotation sets
 - Use recurrently mutated annotation set numbers for prioritization
- Strengthen prioritization criteria so that, in addition to the multi-sample criterion, the mutations must also span both partners

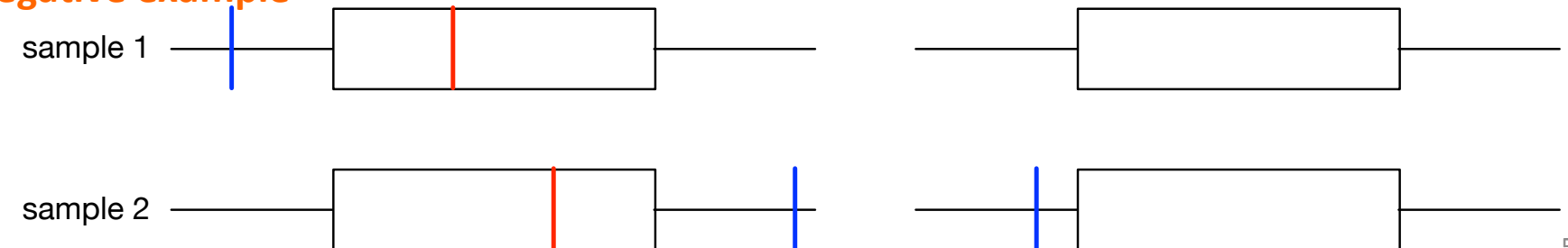
Pair of interaction partners: mutations span multiple samples, and both partners are mutated

Positive example



Pair of interaction partners: mutations span multiple samples, but only one partner is mutated.

Negative example



LARVA(all prostate, HPRD)

- List of genes with number of partners, where both the listed gene and its partner were mutated
 - Idea is that the set of two genes is recurrently mutated
 - This list is a truncation to only those with 15 or more mutated partners

| Gene | # Partners Also Mutated |
|--------|-------------------------|
| TP53 | 52 |
| SRC | 49 |
| PRKCA | 42 |
| CASP3 | 35 |
| RB1 | 34 |
| SP1 | 31 |
| PTK2 | 31 |
| SHC1 | 30 |
| CREBBP | 29 |
| AR | 27 |
| RHOA | 26 |
| PRKCD | 25 |
| RXRA | 24 |
| PTPN6 | 23 |
| LRP1 | 23 |
| VIM | 22 |
| SYK | 20 |
| MYC | 20 |
| VAV1 | 19 |
| HRAS | 19 |
| CBL | 19 |
| PRKACA | 18 |
| LRP2 | 17 |
| ZAP70 | 16 |
| SMAD4 | 16 |
| MAPK14 | 16 |
| ACTN2 | 16 |
| DLG4 | 15 |

LARVA(glioma, genes) with KEGG pathway numbers

Grade 2

| Gene name | nsamp | nvar | # pathways |
|-----------|-------|------|------------|
| IDH1 | 5 | 1 | 3 |
| TP53 | 4 | 0 | 21 |
| HLA-DRB1 | 1 | 0 | 12 |
| ADCY5 | 1 | 0 | 9 |
| HLA-C | 1 | 0 | 9 |
| CTNNA2 | 1 | 0 | 6 |
| CACNA2D3 | 1 | 0 | 5 |
| MAP2K3 | 1 | 0 | 5 |
| DMD | 1 | 0 | 4 |
| ANPEP | 1 | 0 | 3 |
| CR2 | 1 | 0 | 3 |
| CSNK2A1 | 1 | 0 | 3 |
| MAPKAPK2 | 1 | 0 | 3 |
| PIK3C3 | 1 | 0 | 3 |
| PML | 1 | 0 | 3 |
| PTCH2 | 1 | 0 | 3 |
| HPD | 1 | 0 | 2 |
| MYH7B | 1 | 0 | 2 |

Grade 4

| Gene name | nsamp | nvar | # pathways |
|-----------|-------|------|------------|
| TP53 | 6 | 0 | 21 |
| TTN | 6 | 0 | 2 |
| ARHGAP5 | 5 | 75 | 2 |
| IDH1 | 4 | 1 | 3 |
| PTEN | 4 | 0 | 11 |
| MAP2K3 | 3 | 1 | 5 |
| CACNG3 | 2 | 1 | 5 |
| CAMK2G | 2 | 1 | 10 |
| GIT1 | 2 | 1 | 3 |
| HSPA1L | 2 | 1 | 4 |
| MTHFD1 | 2 | 1 | 2 |
| C7 | 2 | 0 | 3 |
| EGFR | 2 | 0 | 20 |
| HLA-DQB1 | 2 | 0 | 11 |
| PTPN11 | 2 | 0 | 8 |
| RYR3 | 2 | 0 | 2 |

Which genes are recurrently mutated, and participate in multiple pathways? 7

LARVA(glioma, KEGG) with recurrently mutated gene numbers

- All grade 2 pathways mutated in multiple samples had only one recurrently mutated gene each

Grade 4 pathways ordered by # recurrently mutated genes (top 20)

| Pathway | nsamp | nannot | nvar | num_genes |
|---|-------|--------|------|-----------|
| kegg_mapk_signaling_pathway.txt | 18 | 6 | 4 | 7 |
| kegg_olfactory_transduction.txt | 15 | 5 | 4 | 5 |
| kegg_glioma.txt | 14 | 3 | 1 | 4 |
| kegg_cell_adhesion_molecules_cams.txt | 15 | 2 | 0 | 4 |
| kegg_epithelial_cell_signaling_in_helicobacter_pylori_infection.txt | 7 | 1 | 1 | 3 |
| kegg_erbB_signaling_pathway.txt | 9 | 2 | 1 | 3 |
| kegg_gnrh_signaling_pathway.txt | 10 | 2 | 2 | 3 |
| kegg_calcium_signaling_pathway.txt | 11 | 1 | 1 | 3 |
| kegg_endometrial_cancer.txt | 12 | 2 | 0 | 3 |
| kegg_neurotrophin_signaling_pathway.txt | 13 | 3 | 1 | 3 |
| kegg_melanoma.txt | 14 | 2 | 0 | 3 |
| kegg_prostate_cancer.txt | 15 | 2 | 0 | 3 |
| kegg_neuroactive_ligand_receptor_interaction.txt | 16 | 1 | 1 | 3 |
| kegg_endocytosis.txt | 16 | 2 | 2 | 3 |
| kegg_pathways_in_cancer.txt | 17 | 2 | 0 | 3 |
| kegg_focal_adhesion.txt | 18 | 2 | 75 | 3 |
| kegg_antigen_processing_and_presentation.txt | 9 | 1 | 1 | 2 |
| kegg_non_small_cell_lung_cancer.txt | 9 | 2 | 0 | 2 |
| kegg_wnt_signaling_pathway.txt | 9 | 3 | 1 | 2 |
| kegg_alzheimers_disease.txt | 10 | 0 | 0 | 2 |

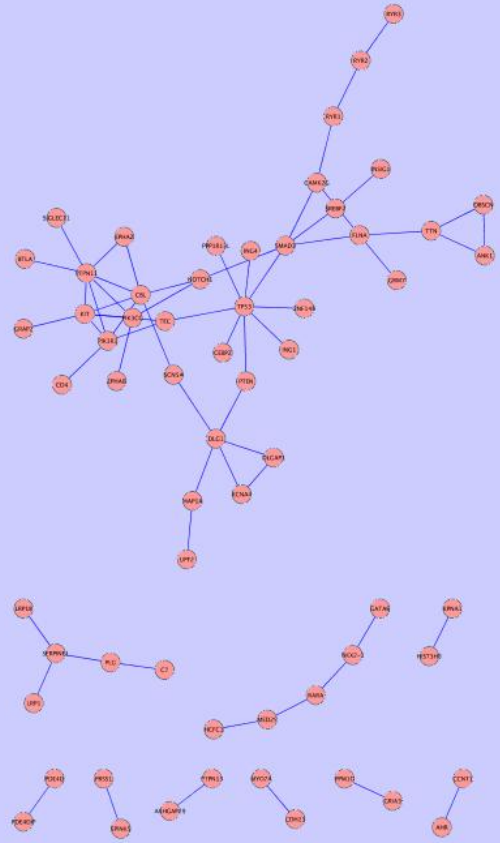
LARVA(glioma, KEGG) with recurrently mutated gene numbers

Grade 4 pathways ordered by # samples mutated (top 20)

| Pathway | nsamp | nannot | nvar | num_genes |
|--|-------|--------|------|-----------|
| kegg_mapk_signaling_pathway.txt | 18 | 6 | 4 | 7 |
| kegg_focal_adhesion.txt | 18 | 2 | 75 | 3 |
| kegg_pathways_in_cancer.txt | 17 | 2 | 0 | 3 |
| kegg_neuroactive_ligand_receptor_interaction.txt | 16 | 1 | 1 | 3 |
| kegg_endocytosis.txt | 16 | 2 | 2 | 3 |
| kegg_olfactory_transduction.txt | 15 | 5 | 4 | 5 |
| kegg_cell_adhesion_molecules_cams.txt | 15 | 2 | 0 | 4 |
| kegg_prostate_cancer.txt | 15 | 2 | 0 | 3 |
| kegg_glioma.txt | 14 | 3 | 1 | 4 |
| kegg_melanoma.txt | 14 | 2 | 0 | 3 |
| kegg_dilated_cardiomyopathy.txt | 14 | 1 | 1 | 2 |
| kegg_regulation_of_actin_cytoskeleton.txt | 14 | 1 | 1 | 2 |
| kegg_small_cell_lung_cancer.txt | 14 | 2 | 0 | 2 |
| kegg_neurotrophin_signaling_pathway.txt | 13 | 3 | 1 | 3 |
| kegg_hypertrophic_cardiomyopathy_hcm.txt | 13 | 1 | 1 | 2 |
| kegg_endometrial_cancer.txt | 12 | 2 | 0 | 3 |
| kegg_p53_signaling_pathway.txt | 12 | 2 | 0 | 2 |
| kegg_calcium_signaling_pathway.txt | 11 | 1 | 1 | 3 |
| kegg_systemic_lupus_erythematosus.txt | 11 | 0 | 0 | 2 |
| kegg_chronic_myeloid_leukemia.txt | 11 | 2 | 0 | 2 |

13 pathways here are common to last slide's top 20, so gene order and sample order produce similar results here.

LARVA(grade 4 glioma, HPRD)



LARVA(grade 4 glioma, HPRD)

- DAVID results for 38-node component:
 - These nodes are enriched for phosphoproteins and membrane proteins (SP_PIR_KEYWORDS)
 - KEGG pathways:
 - T cell receptor signalling pathway: 7 genes
 - Neurotrophin signalling pathway: 7 genes
 - 5 genes already implicated in glioma
 - Overlapping cancers: chronic myeloid leukemia, endometrial cancer, melanoma, pancreatic cancer, small cell lung cancer, prostate cancer, non small cell lung cancer, renal cell carcinoma
- DAVID results for 5-node component, consisting of:
 - RARA
 - GATA6
 - NKX2-1
 - HCFC1
 - MED25
 - SP-PIR indicates that these are DNA-binding transcription activators. (localized to nucleus)
 - So they're regulators
 - Backed up by GO
 - NCBI indicates no prior publication connecting these to glioma.
 - NKX2-1 was last year published as a novel biomarker for lung cancer.

Future Work

- Modify LARVA-SAM to compute stats on gene level
 - Currently only works on exon level