Elucidation of Pathways in Cancer Associated with Survival

Andrew M. Gross, Janusz Dutkowski, Trey Ideker

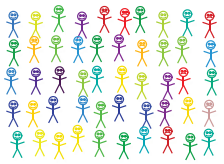
Bioinformatics and Systems Biology Program, Department of Medicine, University of California- San Diego, CA

When seeking to isolate the oncogenic drivers from the spectrum of genomic alterations in cancer, current methods achieve power by citing the presence of a selective advantage in tumor cells and show that these driver events are over-represented verses a background model. The inherent limitation of such methods is their inability to find rare genomic alterations. The use of network and pathway based methods has been proposed to solve this problem, but such methods are underpowered for annotating significant gene-sets based on selectivity as a criteria.

We have developed a series of methods for the characterization of such sub-network and pathway level biomarkers not based on prevalence of mutations, as has been done previously, but by associations of altered pathways with clinical and molecular phenotypes.

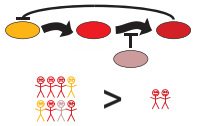
We have systematically applied this method to all available TCGA datasets using mSigDB pathways, uncovering associations recapitulating known hallmarks of cancer alongside a number of novel pathways.

Problem	What events differentiate people with cancer?
Intuition	Different types of patients should have different clinical and molecular phenotypes.
Solution	Stratify cohort, look for changes in phenotype.

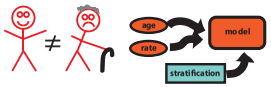


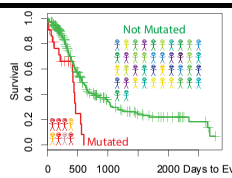
Question What events differentiate people with cancer?

Problem Events are sparse across cohort.
Solution Collapsing onto pathways improves sample size and power.



Problem Global variables such as age and mutation rate confound analysis of genes and pathways.
Solution Modeling covariates along with genomic events acounts for local and global contributions to phenotype.

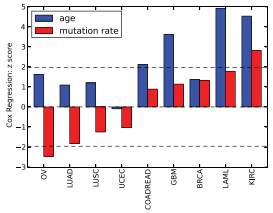




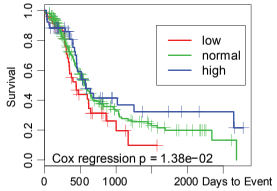
Result Pathway level events can be associated with phenotypes such as patient survival.

Many Pathways have Survival Associations

Global Variables are Associated with Survival

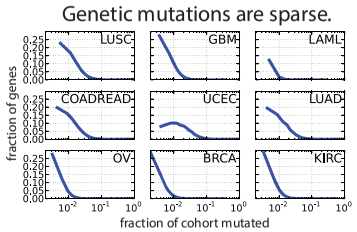


Z-scores for age and mutation-rate association with patient survival assessed by Cox regression.

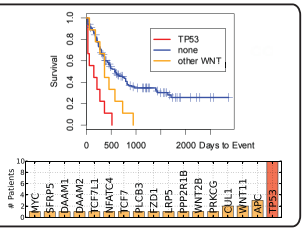


Patients stratified by mutation rate in the **Ovarian** cohort. Here low and high represent +/- 1 standard deviation from the mean.

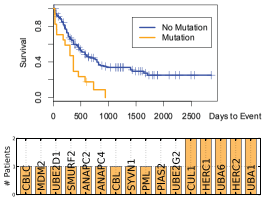
Pathway Level Analysis Can Characterize Sparse Mutations



In the **LAML** dataset TP53 is mutated in 10 patients who are associated with poor survival. Expanding to the **WNT pathway** adds an additional 10 patients with a similar phenotype.



Often mutations are very sparse across the pathway but have a common effect. The **ubiquitin mediated proteolysis pathway** contains 138 genes, in **LAML** 19 of those contain mutations covering 20 patients. While this pathway is definately not under selection, taken together these patients have poor prognosis.

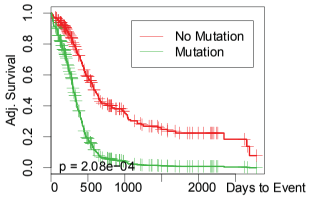


Cox PH Model Adjusts Global Covariates

- Non-parametric survival function adjusted with parametric covariates
- Assumption of time-independence of covariates
- Assumes a proportional increase with increasing values of covariates (linear in log- space)

```
model = Surv(time, event) ~ covariates + interactions
```

	hazard ratio	p-value
pathway	3.10	0.0002
age	1.01	0.0740
rate	0.67	0.0033
pathway:age	0.99	0.7071
pathway:rate	0.71	0.6077



Conclusion

Selectivity is useful for finding genes that cause cancer, but we are also interested in markers for cancer progression and treatment success.
Integration of sparse genetic events over networks and pathways boosts our power to uncover novel targets.
Finding **associations** with clinical phenotypes allows for annotation of new genes and pathways.
While rare genomic alterations elude single gene centered analysis, their characterization through pathway and network based approaches is an important milestone on the path towards patient-tailored treatment of cancer.