A comprehensive characterization of hyper-morph, hypo-morph, and neo-morph mutations in cancer

PHaNToM: Protein-activity based identification of Hyper-, Hypo-, and Neo-morphic effecTors of Mutations

Somnath Tagore[1], Samuel Tsang[2], Gordon B. Mills[2], and Andrea Califano[1]

[1]Department of Systems Biology, Columbia University Irving Medical Center, New York, NY, USA.2Department of Cell Development and Cancer Biology, Oregon Health and Science University, Knight Cancer Institute, Portland, Oregon, USA

E-Mail: [1]st3179@cumc.columbia.edu, [2]tsangsa@ohsu.edu, [2]millsg@ohsu.edu, [1]ac2248@cumc.columbia.edu

INTRODUCTION

BACKGROUND: Although annotation of genomic mutations is a highly relevant and complex segment of the analysis of sequence-based genomic analyses, currently more than ten million variants lack functional annotation. While computational predictions of variant function are usually integrated into gene-based analyses of rare-variants, there is limited information for assessing variant function in the context of a particular disease.

HYPOTHESIS: Use activity of transcription regulators as a gene reporter assay for asessing the effect of upstream mutations.

GOAL: Computational characterization of somatic mutations as neutral, hyper- (gain-of-function), hypo- (loss-of-function) or neo-morphs based on their effect on the actiivty of their downstream transcriptional regulators, using the VIPER[1].

APPROACH: Our pipeline integrates structural and functional information encompassing six topics: 1) structural domains affected by the mutation, 2) the overlap between mutation-specific TF/co-TFs, 3) differential activity signatures and signatures induced by established hyper-morph, hypo-morph and neutral or neo-morph mutations, 4) in vitro data generated by reporter assays, 5) the VIPER[1]-inferred activity of each protein relative to a validated control, and 6) the fraction of proteins in a sample that are not affected by established hyper-morphs and hypo-morphs (candidate neo-morphs)[2].

RESULTS: PHaNToM considers 25 TCGA cohorts and 17 CCLE tissues (cell-lines) for 3830 Proteins, to identify hyper-, neutral, hypo-, and neo-morph mutations. The pipeline also predicts mutations that phenocopy the effect of other mutations.

METHODS

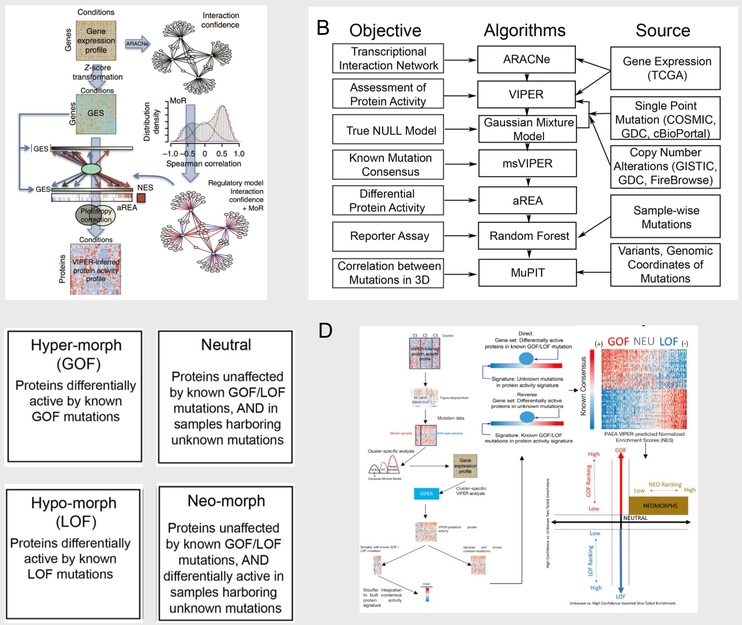
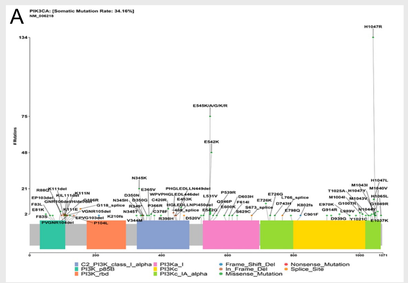


Figure 1: A) VIPER method[1], B) Overall objectives, C) Definitions of GOF, Neutral, LOF, Neo-morph, D) Step-wise method description

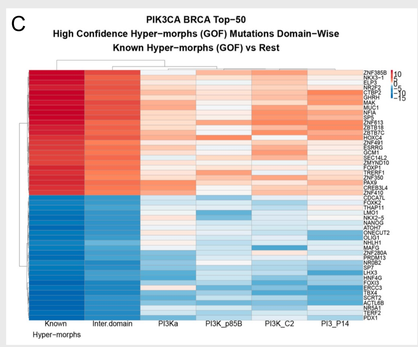
RESULTS

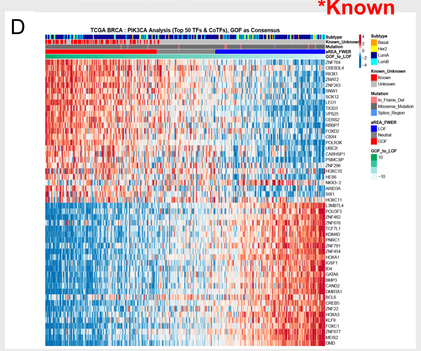


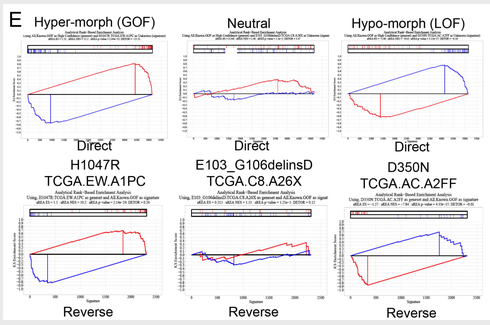
B

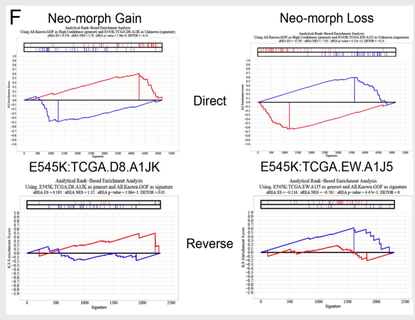
	Mutation Variants
PI3K_p85B	E81K, R88Q, P104L, R108H, E103_G106delinsD, E103_P104del, F70F, F83S, G106_R108del, V105_R108del
PI3K_C2	P366R, R398H, C420R, E453K, L456Afs.13, E469delinsDK, N145N
PI3Ka	P539R, E542G, Q546K, D603H, F614I, S673T, E545K, E542K, Q546R, D520V, E545A, D603H, E600K
PI3_P14	C901F, G914R, D939G, E970K, L989V, M1004I, G1007R, C901F, D939G
Inter Domain Region	E110del, K111E, K111N, K111del, G118D, V344M, N345K, D350G, E726K, L764, Y1021C, T1025A, E1037K, M1040V, M1043I, M1043V, H1047L, H1065L, H1047R, E109_I112del, E109_I112delinsD, L156L, D350N, N345H, N345T, P447_L455del, R349, E726G, E726K, L755L, A1046A, G1049R, H1047Y, H1065L, N1044Y

*Known









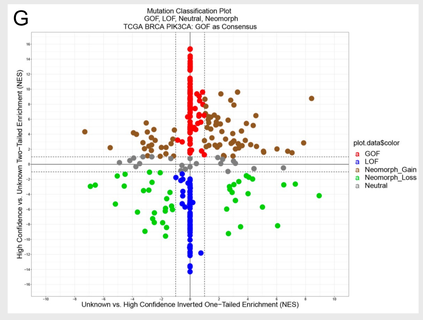


Figure 2: A-B) PIK3CA (BRCA) mutations domain-wise, C) Known GOF vs Unknown, D) Heatmap representing sample with PIK3CA (BRCA) mutations as compared to known PIK3CA GOF, E-F) aREA plots, G) Mutation classification plot

CONCLUSIONS

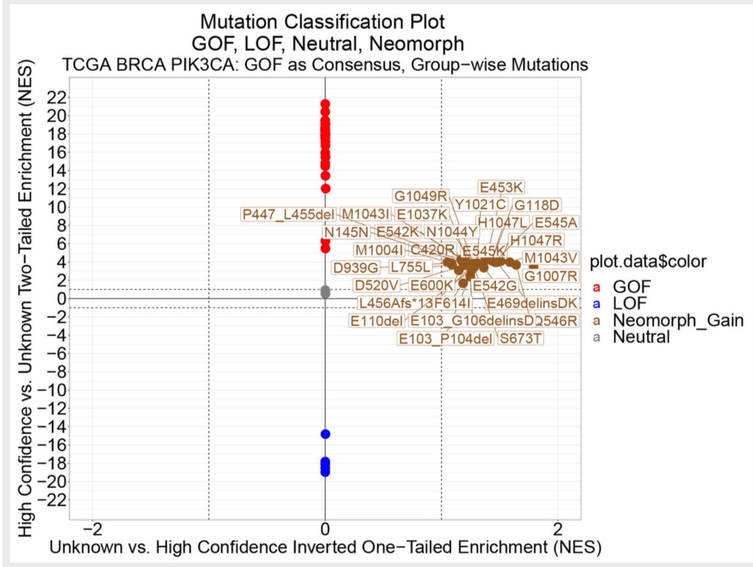


Figure 3: Mutation classification plot for PIK3CA (BRCA) prediction neo-morphs

Here, we present our analysis results on PIK3CA in TCGA Breast Cancer dataset (TCGA-BRCA) by predicting with very high confidence several neo-morphic phenotypes, including the previously-described $PIK3CA^{E545}$, $PIK3CA^{E542}$, $PIK3CA^{H1047}$, $PIK3CA^{Q546K}$ and $PIK3CA^{G1049R}$. Interestingly, $PIK3CA^{E545K}$, classified previously as a gain-of-function mutation (in one TCGA-BRCA sample) or a loss-of-function mutation (in two other TCGA-BRCA samples), is predicted to be a neo-morph based on our approach. Further validation of these mutations using PIK3CA reporter assays led to the identification of several significant hypo-morphic signals in TP53 mutant samples. We defined this phenomenon as mutational mimicry (i.e. mutations in proteins mimicking those in established oncogenes) and we propose it as a tool for predicting tumor sensitivity/resistance to drugs. Currently, experimental validation is ongoing for the predicted hyper-, hypo-, and neo-morphs.

Hyper-morphs (GOF)	Neutral	Hypo-morphs (LOF)	Neo-morph Gain
N345K	F70F	V105_R108del	C420R
E726K	D350G	K111N	E453K
C901F	R108H	N345T	E545K
R349*		E109_I112delinsD	H1047L
T1025A		D350N	E542K
P366R			G118D
P104L			M1043I
P539R			P447_L455del
K111del			H1047R
A1046A			E545A
N345H			E469delinsDK
L764=			M1004I
R88Q			M1043V
E970K			D939G
G106_R108del			Y1021C
D603H			N145N
G914R			G1049R
L989V			E1037K
H1065L			N1044Y
F83S			L755L
V344M			Q546R
M1040V			D520V
L156L			L456Afs*13
R398H			G1007R
K111E			S673T
E726G			E542G
H1047Y			E110del
E109_I112del			E600K
LOF			E103_P104del
LOF			E103_G106delinsD
LOF			F614I
LOF			
LOF			

Table 1: List of potential GOF, Neutral, LOF, Neo-morphs in PIK3CA (BRCA)

ACKNOWLEDGEMENT

We thank Dr. Gordon B. Mills group for their collaboration.

U01 CTD2 for funcding to Dr. Andrea Califano.

BIBLIOGRAPHY

[1]Alvarez, M.J., Shen, Y., Giorgi, F.M., Lachmann, A., Ding, B.B., Ye, B.H., and Califano, A. (2016). Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat Genet 48, 838-847.

[2]Ng, P.K., Li, J., Jeong, K.J., Shao, S., Chen, H., Tsang, Y.H., Sengupta, S., Wang, Z., Bhavana, V.H., Tran, R., et al. (2018). Systematic Functional Annotation of Somatic Mutations in Cancer. Cancer Cell 33, 450-462.e410.