Causal Modeling of TCGA NSCLC and HNSCC Data Identifies Network Drivers of Tumor Immune Subtypes

GNS HEALTHCARE

Rahul K Das[1*], Raymond T Yan[1], Aishwarya Krishnakumar[1], Bruce Church[1], Iya Khalil[1], Xinmeng J Mu[2], Mateusz Maciejewski[3], Craig Davis[2]

[1]GNS Healthcare, Cambridge, MA 02139; [2]Pfizer Inc, San Diego, CA 92121, [3]Pfizer Inc, Cambridge MA, 02139

BACKGROUND

- Immune checkpoint inhibitors (ICI) have achieved unprecedented success in several cancer types, yet only a subset of patients derives clinical benefit. A major focus today is to discover why ICI work or fail, and how they can be improved to reach their hoped-for potential.
- There is an unmet need for biomarkers that will identify patients who are more likely to respond to ICI or develop resistance afterwards.
- The fundamental basis for ICI response is the immunogenicity of a tumor, primarily determined by tumor-immune interactions. Better understanding of tumor immunogenicity and the underlying molecular drivers is imperative to improve clinical outcomes in ICI.
- Bayesian causal machine learning was applied to TCGA Non-Small Cell Lung Carcinoma (NSCLC) and Head and Neck Squamous Cell Carcinoma (HNSCC) primary tumor data to elucidate the molecular drivers/ biomarkers of tumor immunogenicity.

METHODS

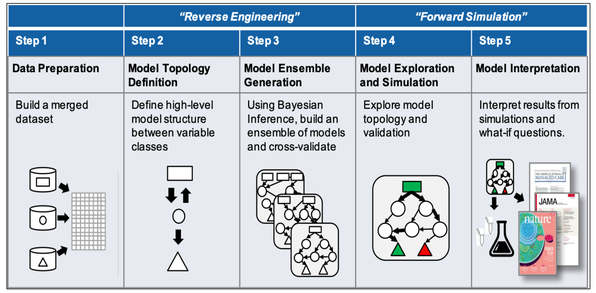

Figure 1: Schematic of REFS™ Reverse Engineering & Forward Simulation Workflow

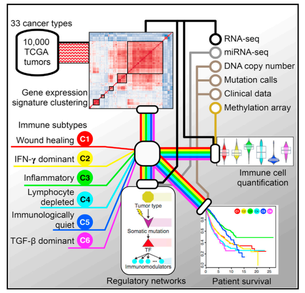

Figure 2: Definition of tumor immunogenicity (immune subtypes)[1]

- Using a Reverse Engineering Forward Simulation (REFS™) platform, ensembles of 256 independent causal models were built on molecular and clinical data from primary tumor samples of 681 NSCLC, 328 Lung Adeno (LUAD) and 353 Lung Squamous Cell carcinoma (LUSC), and 413 HNSCC patients in TCGA to identify causal drivers of tumor immunogenicity. Due to scarcity of molecular drivers in LUSC, results are shown only for LUAD.
- A total of ~ 2500 variables from genomics, transcriptomics, and clinical data were used for the modeling.
- Tumor immunogenicity was defined as six tumor immune subtypes[1]: wound-healing, IFNγ dominant, inflammatory, lymphocyte depleted, immunologically quiet, and TGFβ dominant. TCGA NSCLC patients had only the first three and HNSCC patients had only the first two immune subtypes.
- Causal drivers of tumor immune subtypes were identified from average causal effect (ACE) of the variables, as computed from the REFS counterfactual simulations. ACE was defined as median of posterior distribution of odds ratio (1 vs 0 for discrete; 95th vs. 5th %ile for continuous variables).
- Model performance was evaluated for 5-fold cross-validation measured by precision/recall, accuracy, and AUC.

Bayesian Causal Inference On TCGA HNSCC And NSCLC Tumor Data

- suggests macrophage activation and polarization, which is driven in part by metabolic reprogramming, is the primary driver of tumor immunogenicity.
- generates literature-supported hypotheses regarding molecular biomarkers of patient response and resistance to ICI.
- classifies tumors accurately based on their immunogenicity indicating the power of causal AI in identification of biomarkers for predicting response to ICI and for turning "cold" tumors "hot".

Immune Subtype	Tumor	Causal Driver	Average Causal Effect
Wound healing	LUAD	STK11 mutation	1.3
		Tumor mutation burden (TMB)	4.2
	HNSCC	AKT1/mTOR upregulation	3.4
		STAT6 upregulation	1.3
IFNγ dominant	LUAD	DUSP6 upregulation	16.5
		TMB	2.6
	HNSCC	CXCL13 upregulation	2.2x10^6
Inflammatory	LUAD	DCTD upregulation	16.5

Table 1: Causal drivers of tumor immune subtypes for HNSCC and LUAD

References:

1. Immunity. 2018 Apr 17;48(4):812-830.e14.
2. J Clin Oncol. 2019 Feb 1;37(4):318-327
3. Science. 2018 Oct 12;362(6411)
4. Cancer Discov. 2018 Jul;8(7):822-835.
5. Semin Cancer Biol. 2018 Feb;48:91-103.

***Correspondence:**
rdas@gnshealthcare.com

RESULTS

CAUSAL SUBNETWORKS OF TUMOR IMMUNE SUBTYPES:

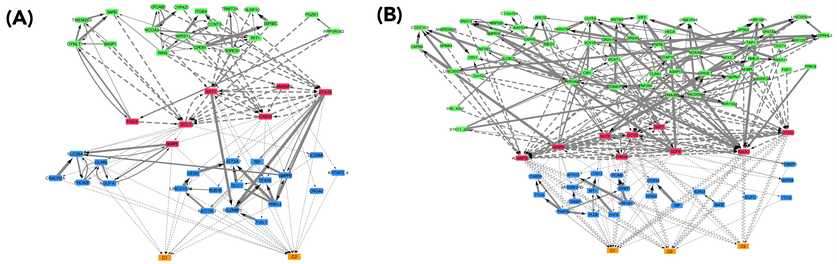

Figure 3: The consensus of reverse engineered causal subnetworks for (A) HNSCC (B) LUAD tumor immune subtypes (shown in orange). Key causal drivers, their upstream and downstream nodes are shown in red, green and blue, respectively.

MODEL PERFORMANCE:

Immune Subtypes	HNSCC				LUAD			
	N	AUC	Precision	Recall	N	AUC	Precision	Recall
Wound healing	106	0.89 (0.86,0.95)	0.65 (0.54,0.73)	0.88 (0.73,1)	68	0.86 (0.85,0.89)	0.5 (0.47,0.58)	0.82 (0.73,1)
IFN γ dominant	303	0.91 (0.88,0.94)	0.95 (0.92,1)	0.81 (0.8,0.92)	112	0.83 (0.75,0.87)	0.72 (0.57,0.83)	0.75 (0.74,0.89)
Inflammatory	-	-	-	-	121	0.94 (0.89,0.95)	0.84 (0.79,0.9)	0.9 (0.73,0.96)

Table 2: Summary of model performance in classifying immune subtypes in 5-fold cross-validation. The numbers for AUC, precision, and recall denote median (minimum, maximum) across the 5 folds.

DISCUSSION:

The identified causal drivers of tumor immunogenicity are in agreement with literature supported hypotheses on determinants of tumor microenvironment, tumor antigenicity and ICI response.

- Upregulation of CXCL13 gene (identified as causal driver of HNSCC IFN γ dominant immune subtype) is part of a T-cell-inflamed Gene Expression Profile (GEP) that predicts response to pembrolizumab across multiple cancers[2].
- STK11 mutation (identified as causal driver of LUAD wound healing immune subtype) is negatively correlated with T-cell-inflamed GEP in LUAD[3] and associated with shorter OS and PFS with PD-1 blockade among KRAS-mutant LUAD[4].
- Recent data derived from clinical trials and pre-clinical mouse models suggests that therapeutic inhibition of the PI3K-AKT-mTOR signaling (upregulation identified as causal driver of HNSCC wound healing immune subtype) augments tumor immunosurveillance by preventing activation of immunosuppressive pathways[5].
- TMB is identified as causal driver of both T-cell inflamed IFN γ dominant and T-cell devoid wound healing immune subtypes in LUAD, consistent with the emerging hypothesis that TMB and T-cell GEP need to be considered together as predictive biomarkers for ICI response.

FUTURE DIRECTIONS

The ability to classify tumors by immunogenicity along with concordance of the identified molecular drivers with literature-supported biomarkers of response/resistance to ICI therapies lends credence to the causal models and indicates potential utility of causal AI for predicting response to ICI therapies. **Validation of these models in independent study cohorts is ongoing.**