Uncovering Stochastic and Drug-Induced Phenotypic Transitions in Small Cell Lung Cancer

Sarah Maddox, David Wooten, Carlos Lopez, Vito Quaranta, Vanderbilt University

ABSTRACT

Small cell lung cancer (SCLC) is an aggressive tumor type that relapses at high rates and metastasizes early. Despite these poor outcomes, the standard of care for SCLC has not changed in decades, because the mechanisms of acquired resistance are not well understood. Within the past decade, the role of intratumoral heterogeneity in tumor aggressiveness has become more apparent, as cooperation between subtypes of SCLC can potentiate metastasis [Calbo, 2011] and treatment evasion [Jahchan, 2016]. The discovery of multiple SCLC phenotypes necessitates further analysis to map the phenotypic space of SCLC. Consensus clustering and weighted gene co-expression network analysis (WGCNA) reveal 4 clusters, including a novel neuroendocrine variant (NEv2) distinguished by gene modules and drug response, and a transcriptional factor network is able to explain these four steady states. Analysis of human and mouse tumors suggests that SCLC phenotype is plastic, both stochastically and in response to drug. We use RNA velocity of single cell RNA-seq data to understand and quantify movement through the phenotypic landscape, and well as make predictions of perturbations that may induce those transitions. Characterization of the transitions between core SCLC phenotypes and predictions of perturbation strategies may suggest new methods of personalized treatment.

CLUSTERING CELL LINES BY GENE EXPRESSION REVEALS 4 BROAD SCLC SUBTYPES

Figure 1 Consensus clustering of 50 CCLE SCLC cell lines reveals 4 transcriptional subtypes of SCLC (left), including 3 neuroendocrine subtypes (NE, NEv1, NEv2) and one non-neuroendocrine subtype (non-NE). The non-NE subtype is most similar to non-SCLC lung cancer cell lines by gene expression (right). Both NEv1 and NEv2 are similar to the canonical NE subtype, but differ greatly from each other.

CURRENT BIOMARKERS FROM LITERATURE ARE UNABLE TO RESOLVE FOUR SUBTYPES

Three of the four subtypes seem to correlate with previously-identified phenotypes from the literature; however, gene signatures, not single biomarkers, are necessary to fully under the intratumoral heterogeneity of SCLC.

Figure 2 Expression of common genes from literature on SCLC phenotypic heterogeneity. Markers are unable to completely distinguish subtypes, especially NE and NEv2.

GENE ONTOLOGY ANALYSIS REVEALS NEv2 IS ENRICHED IN DRUG METABOLISM AND IMMUNO-MODULATORY PROPERTIES

To understand the complex distinguishing features of each subtype, we applied GSEA to the WGCNA module eigengenes that significantly differentiated the subtypes (ANOVA, fdr-corrected p-val, <0.05). NEv2 (high in green module, 3C) showed enrichment of "drug catabolism" and "immune response" terms. NEv2 is the least sensitive to an array of drugs.

Figure 3 A. Gene expression across subtypes can be distinguished by WGCNA modules. B. A Gene Ontology Map (GO map) describing enriched terms in modules that significantly differentiate the four subtypes. C. Green and Yellow module-specific GO maps, respectively.

SUBTYPES ARE REPRESENTED ACROSS HUMAN AND MOUSE TUMORS

Figure 4 Using a nearest neighbor analysis, scRNA-seq data from two independent mouse tumors shows correlations with human cell line subtypes. CIBERSORT (Newman, 2015) was used to detect proportions of subtypes in human tumors before and after treatment (Drapkin, 2018).

IN SILICO SIMULATION OF GENE REGULATORY NETWORK RECAPITULATES FOUR SUBTYPES

Genetic mutations are unable to distinguish between phenotypic subpopulations of SCLC completely; thus, we investigate a transcription factor (TF) network to explain the stability of multiple subtypes. Using TFs from each significant module and the literature, we simulated a network with quasi-Bayesian Boolean rules calibrated to cell line data. This led to 10 attractors that match our four expected phenotypes.

Figure 5 (Left) Gene regulatory network including TFs relevant to SCLC phenotypic heterogeneity. (Right) Simulations of the network are able to recapitulate the expected subtypes. Attractors are defined probabilistically as those states that are more probable than any neighboring state.

IN SILICO PERTURBATIONS SUGGEST STRATEGIES FOR REPROGRAMMING SCLC CELLS TOWARD SENSITIVITY

Boolean perturbations (TF held at 0 or 1) allow us to test the outcome of silencing (0) or activating (1) a specific gene. These perturbations suggest experimental strategies for destabilizing a subtype, or "pushing" a cell away from that attractor state. When applied to resistant cell types, such as NEv2 cells, reprogramming may provide a successful strategy for SCLC treatment.

Figure 6 (Left) Examples of destabilization strategies for each subtype. Perturbation allows the simulation to leave a specified state's 4-TF neighborhood faster than a random walk, so a smaller number of steps is required, shifting the distribution of walk lengths to the left. (Right) Destabilization strategies for multiple subtypes. Four-circle Venn diagram shows which perturbations will destabilize (>20% average change in walk length) at least one subtype.

HOW PLASTIC ARE SCLC SUBTYPES?

Figure 7 (Top) RNA velocity [La Manno, 2018] gradient field for scRNA-seq of a triple knockout (TP53, RB1, P130) mouse SCLC tumor. Within the main "island," cells seem to move from a canonical NE phenotype (purple) towards an NEv2 phenotype (orange). Green = NEv1; Magenta = non-NE. (Bottom) Interestingly, CALCA phase portraits show repression in NEv2 and activation in NE cells, and cells are moving from a low-CALCA, activation state to a high-CALCA, repression state.

CONCLUSIONS AND FUTURE DIRECTIONS

- Consensus clustering identified four phenotypic clusters defined by WGCNA gene modules, with differential drug sensitivity and gene expression patterns.
- NE Variant 2 may be able to modulate the immune system and better resist perturbations by drug through enriched drug metabolism.
- Stable states in a quasi-Boolean GRN recapitulate empirical results, and perturbations of these states suggest strategies for reprogramming.
- RNA velocity analysis suggests relatively plastic NE and NEv2 states, and CALCA expression dynamics show movement toward a high CALCA state (NEv2).
- Pseudo-temporal single cell data along an RNA velocity trajectory will be used to constrain the GRN model and define movement through the SCLC phenotypic landscape.

REFERENCES

Calbo et al. (2011) *Cancer Cell*. 19:244-256.
Jahchan et al. (2016) *Cell Reports*. 16:644-656.
Langfelder, et al. (2008) *BMC Bioinformatics*. 9:559.
Lim et al. (2017) *Nature Letters*. 545:360-364.
Mollaoglu et al. (2017) *Cancer Cell*. 31:270-285.
Williamson et al. (2016) *Nature Communications*.
Huang et al. (2018) *Genes and Development*. 32:915-928.
Drapkin et al. (2018) *Cancer Discovery*.
Wooten & Maddox et al. (2018) *Preprint on BioRxiv*.
Udyavar et al. (2017) *Cancer Research*.
Newman et al. (2015) *Nature Methods*. 12:453-457.
La Manno et al. (2018) *Nature Letters*.
Polley, et al. (2016) *J Natl Cancer Inst*.

This work is supported by the NCI Integrative Cancer Biology Program Grants U54CA113007 and U01CA215845, and NSF Fellowship DGE - 1445197.

NATIONAL CANCER INSTITUTE