Characterizing Phenotypic Heterogeneity in Small Cell Lung Cancer

Sarah F. Maddox, David J. Wooten, Jonathan Lehman, Pierre Massion, Vito Quaranta
Vanderbilt University

Introduction

Small cell lung cancer (SCLC) is an aggressive tumor type with a strong ability to become resistant to all known treatments and to survive in diverse microenvironments. Proposals to stratify patients based on tumor phenotype have been met with resistance due to unclear clinical relevance, as the "small blue round" SCLC cells are extremely uniform by histopathology. More recently, however, it has become increasingly understood that SCLC tumors exhibit phenotypic heterogeneity implicated in the aggressiveness of the disease.

My central hypothesis is that interactions between these plastic phenotypes form a functional ecosystem that drives growth and controls the overall response to therapy.

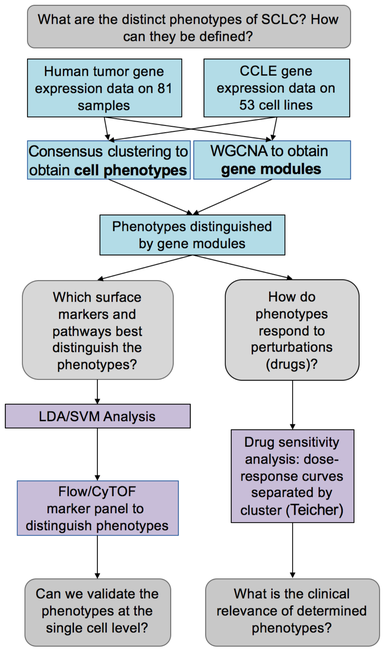


Figure 1. Pipeline for phenotypic heterogeneity analysis of SCLC. The main questions this work addresses are highlighted in gray boxes.

Clustering by Samples and Genes Reveals Four Phenotypes defined by Gene Modules

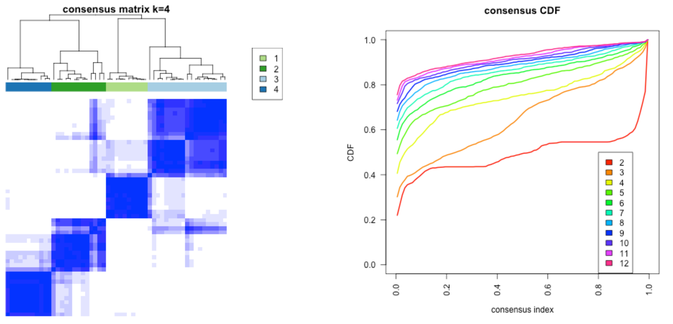


Figure 2. Consensus clustering with k-means algorithm and 1 — Pearson correlation distance metric on cell lines reveals four biologically relevant clusters.

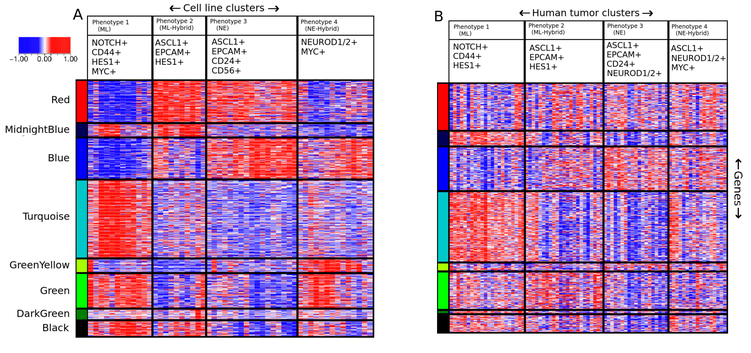


Figure 3. Phenotypic variation across A) 53 SCLC cell lines and B) 81 human tumor samples. Four distinct phenotypes generated by consensus clustering methods above (columns) are described by modules of co-expressed genes generated by Weighted Gene Co-Expression Network Analysis (WGCNA, rows). Similar patterns of co-expression can be seen in the human tumors.

Distinguishing Surface Markers

A)

Module	Ph 2 v Ph 4	Ph 2 v Ph 3	Ph 3 v Ph 4
Grey60	CEACAM6	RET	CEACAM6
Brown	CEACAM5	CLDN4	CEACAM5
Turquoise	THY1	ITGAV	ROR1
Pink	TPBG	CD46	TPBG
Magenta	SPARC	HPCA	SPARC
Blue	JAM2	FZD3	PLXNC1
Black	F3	ICAM3	AHR
Dark magenta			
Green	SLC16A1		SLC16A1
Dark orange	TYRO3		TYRO3
Midnight blue			
White		ADA	
Salmon	CD109	CD109	
Purple	ICAM2	SERPINF2	DPP4
Dark green			
Green yellow		CXCR7	CXCR7
Dark red	CD274	SIGLEC7	
Dark olive green			
Orange			
Yellow		CD86	PIGR
Dark grey			

B)

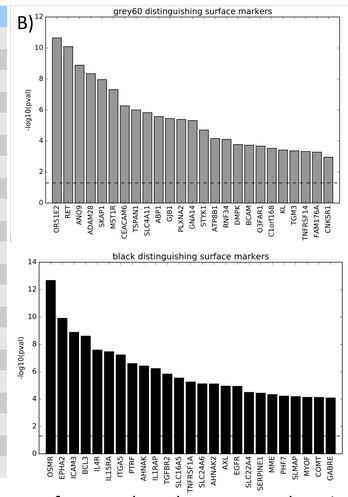


Figure 4. A) Table of the most distinguishing surface markers between each pair of phenotypes listed. Combinations of markers from this table were ranked via analysis below for their ability to maximally distinguish the phenotypes. B) Two example histograms of distinguishing ability (metric: Benjamini-Hochberg corrected p-value for a Mann-Whitney U test between phenotypes) of surface markers from a single gene module.

Drug Sensitivity Analysis by Phenotype

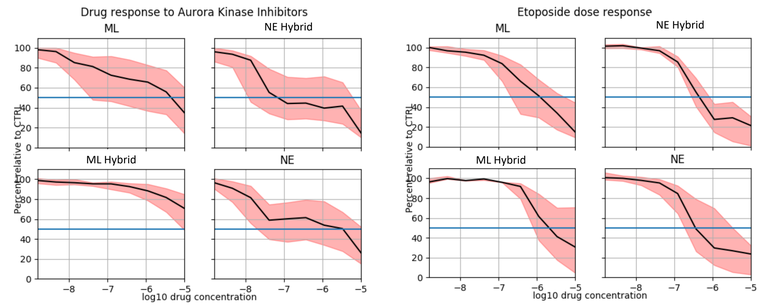


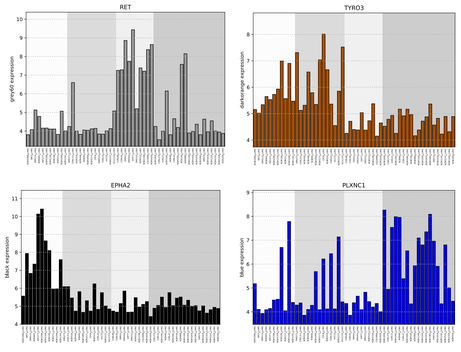
Figure 5. Using previously acquired data from Polley et al. on sensitivity of SCLC to various drugs, we analyzed dose response curves by phenotype. In a significant portion of drugs tested, the NE hybrid phenotype was most sensitive to drug, and the ML hybrid phenotype was most resistant, with NE and ML phenotypes having intermediate sensitivity. Notably, the NE hybrid cell lines showed increased sensitivity to aurora kinase inhibitors, a finding which may corroborate previous research reported in the literature. Etoposide, which is part of the standard treatment for patients with SCLC, showed similar responses when segregated by phenotype.

Support Vector Machine and Linear Discriminant Analysis

Figure 6. A support vector machine (SVM) is a supervised classifier that optimizes a separating hyperplane between groups of data. Both an SVM and a Linear Discriminant Analysis (LDA) were used for this project. SVM was used to determine the best combination of surface markers for distinguishing phenotypes, and LDA was used for dimensionality reduction and projection of experiment results onto the model.

Figure 7. SVM models scored by accuracy using train and test subsets of gene expression data. Average score was calculated from 100 models generated by combinations of three surface markers from a curated list, and the score describes the model's ability to distinguish between phenotypes 1, 2, and 3. Epha2 was added to the marker list to distinguish phenotype 4, which has been used in previous work. Our marker set has a score much higher than average (blue dotted line).

Figure 8. Four chosen surface markers for determining phenotypes. Gene expression data from the CCLE is shown for each of the 53 cell lines for each gene.



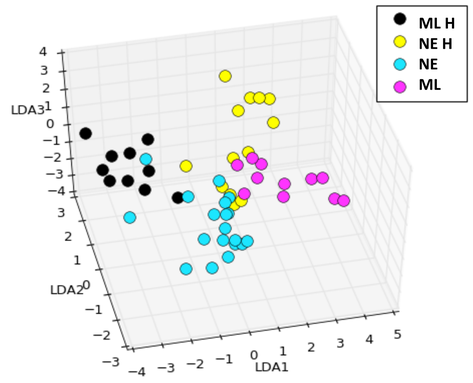


Figure 9. Four phenotypes projected into a reduced-dimensional space by LDA. The model is able to separate the phenotypes.

Future Directions: Flow Cytometry

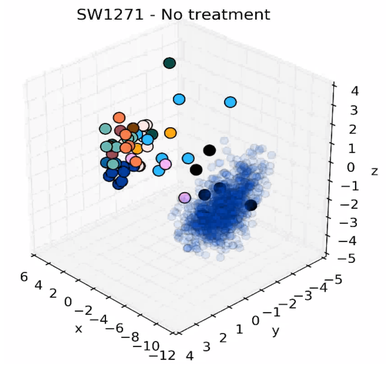


Figure 10. The panel of four cell surface markers optimized to distinguish between and ensure broad coverage of these phenotypes will be used in flow and mass cytometry. This data can then be projected onto the LDA/SVM space to validate the phenotypes at the single cell level. Shown is an example of this type of projection from a preliminary experiment and prior clustering in which the data (dark blue, faint circles) correlates with the "black" phenotype.

Summary and References

- Consensus clustering identified four phenotypic clusters defined by WGCNA gene modules.
- By building SVMs and LDA models based on combinations of a subset of cell surface markers, we are able to distinguish between the four phenotypes using only four markers.
- Flow cytometry allows us to project experimental data at the single cell level into the LDA space defining the separable phenotypes.
- There may be a relationship between phenotype and drug sensitivity; initial analysis suggests ML-hybrid is more resistant than average and NE-hybrid is more sensitive than average. Further analysis is needed.
- Mass cytometry and tissue microarrays will be used to validate phenotypic heterogeneity within human tumors.

(1) Barrentina C., et al. Nature (2012).
(2) George J., et al. Nature (2015).
(3) *Introduction to Support Vector Machines.* Received 31 July 2017 from OpenCV website.
(4) Iorio, F., et al. Cell (2016).
(5) Mollaoglu G., et al. Cancer Cell (2017).
(6) Polley E, et al. J Natl Cancer Inst. (2016).
(7) Udyavar A.R., et al. Cancer Res. (2017).