#2783P

Comprehensive Epigenomic Profiling of Plasma ctDNA Enables Assessment of Drug Target Expression in SCLC

Khoi Nguyen*, Daniel Karl*, Jonathan Beagan*, Justin Finkle, Anthony D'Ippolito, **Aparna Gorthi**, Travis Clark, Amy Donahue, Vy Tran, Tyrone Tamakloe, Charlene O'Brien, Michael Coyne, Jenna I. Wurster, Kristian Cibulskis, Corrie Painter, Mike Zhong, Humphrey Gardner, J. Carl Barrett, Matthew L. Eaton

Precede Biosciences, Boston, MA

BACKGROUND

Small cell lung cancer (SCLC) is an aggressive malignancy with limited therapeutic options. Emerging therapies, such as targeted small molecules, radio targeted ligands, and immune-conjugates, highlight a need for target expression focused diagnostics. Tissue biopsies are challenged by accessibility, tumor necrosis, and comorbidities, while PET scanning with target specific tracers can be challenging to scale and access. Additionally, an unmet need in SCLC care is the identification of patient-specific targets to guide therapy, along with approaches to capture evolving tumor biology as it responds to treatment. To address these limitations, we leveraged epigenomic profiling of circulating chromatin to infer RNA-seq-level expression from plasma ctDNA, enabling prediction of over >1,600 expression-based targets, among them emerging immune-conjugate candidates and SCLC subtypes, at clinically relevant tumor fractions, tunable to lower ctDNA.

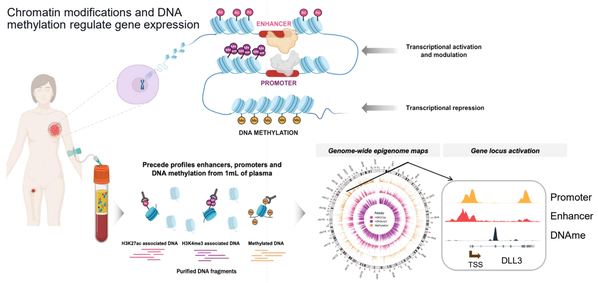
METHODS

We developed transcript prediction models from epigenomic signals and validated in ctDNA from SCLC patients (pts). Models were trained to predict RNA-seq-level expression and evaluated for their ability to classify ASCL1 (A), NEUROD1 (N), POU2F3 (P), and Triple Negative (TN) subtypes, and to quantify >1,600 genes, including clinically relevant Antibody-Drug Conjugate (ADC) targets, from 1mL of plasma from 50 pts. Predicted expression levels were validated by recapitulating known SCLC correlation networks in patient plasma.

RESULTS

Figure 1. Epigenomic Platform

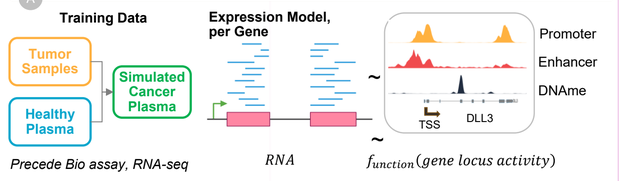
Comprehensive Epigenomic Platform Offers Dynamic Resolution into Target & Pathway Biology from 1ml of Plasma



Cell free DNA derived from tumors exist in circulation as chromatin fragments that faithfully maintain tumor-associated epigenetic modifications on the histones and DNA. Antibodies against active enhancers, active promoters and DNA methylation are used to enrich for associated DNA fragments from 1mL plasma and sequenced to capture the underlying transcriptional state of the tumor cells (Baca et al.).

Figure 2. Gene Locus Expression Models in Simulated Cancer Plasma

Training & Testing

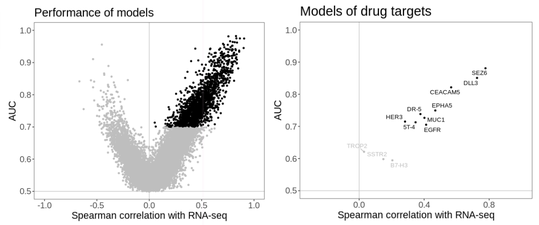


Precede Bio assay, RNA-seq

(Left) Schematic representation of the approach used to generate data to train gene locus expression models. Simulated cancer plasma is created by mixing tumor-derived chromatin with healthy plasma *in silico*.

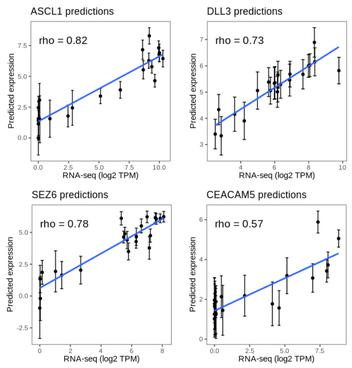
(Right) For each gene, an expression model is trained to predict tumor RNA transcript levels as a function of gene activity in plasma, integrating enhancer, promoter and DNA methylation signals.

Cross-validated Performance Train Across a Panel of Breast Cancer Samples



Precede Bio has quantitatively trained and evaluated >1600 gene expression models applicable to ~60-80% of Stage IV SCLC patients.[1] (Left) Performance of gene models at simulated 10% ctDNA. (Right) Model performance of potential ADC targetable genes. This set contains potential ADC targetable genes in SCLC including **SEZ6, DLL3, CEACAM5, HER3**, and **EGFR**, as well as emerging targets across other indications.

Performance of Models for Clinically Relevant Genes in SCLC



Correlation between actual and predicted values of expression in simulated 10% ctDNA plasma samples for 4 key SCLC genes: **ASCL1, DLL3, SEZ6**, and **CEACAM5** (Spearman's rho, $p<1\times10^{-8}$). The x-axis represents actual expression measured by RNA-seq of the tumor sample before plasma simulation, while the y-axis shows the model's predicted gene expression from cfDNA epigenomic features in simulated 10% ctDNA plasma samples generated from that sample. Error bars represent the standard deviation of the model predictions across simulated biological replicates.

Refining Additional Features Stabilize Predictions to Lower Tumor Fractions

Example of model improvement through feature refinement using the DLL3 gene model. Pearson correlation of DLL3 expression predictions across simulated ctDNA fractions to per-sample mean estimates at 20% ctDNA demonstrate that incorporating additional features can maintain gene expression predictions to lower ctDNA%. This strategy can be applied to refine performance for many individual gene models, with a stable performance down to ~1% ctDNA for DLL3, for example.

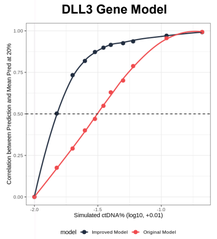
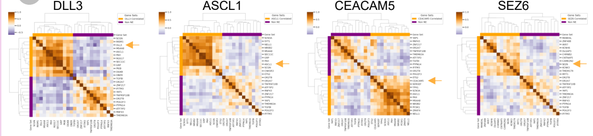


Figure 3. Predicted Gene Expression from SCLC Patient Plasma

Recapitulates Known Co-expression Patterns in SCLC



Tumor gene expression was predicted from 50 plasma samples of extensive-stage SCLC. Genes co-expressed with **DLL3, ASCL1, CEACAM5**, or **SEZ6**, as well as non-neuroendocrine (non-NE) genes, clustered as expected. Target associated gene sets were defined using the transcripts with the highest Spearman correlations from 120 previously published SCLC tumor tissue RNA-seq profiles.[2] Non-NE genes were identified as transcripts negatively correlated with **INSM1, CHGA**, and/or **SYP**. Arrows indicate the location of the relevant gene in the clustered heatmap.

SCLC Subtype-defining Transcription Factors (TFs) & ADC Targets Is Consistent with Expected Patterns

Predicted gene expression was derived from plasma of 37 SCLC patients with >10% ctDNA. *(Left)* Correlation plots show **INSM1** and **CHGA** (indicators of neuroendocrine differentiation) in relation to the subtype-defining transcription factors **ASCL1, NEUROD1**, and **POU2F3**. INSM1, ASCL1 and NEUROD1 are commonly used to classify neuroendocrine (NE) tumors, while POU2F3 expression marks non-NE tumors.[3] *(Right)* While **SEZ6** has been shown to be elevated in several subtypes, it has been demonstrated that **DLL3** is most strongly elevated in the ASCL1 subtype.[4] These observations are recapitulated in SCLC gene prediction correlations from plasma. Significance is determined using a Pearson correlation.

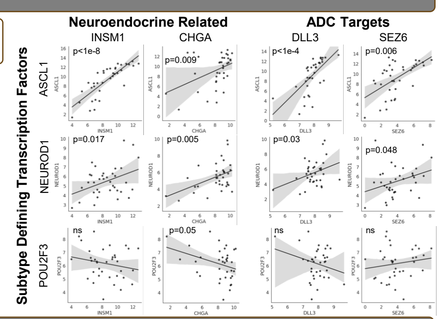
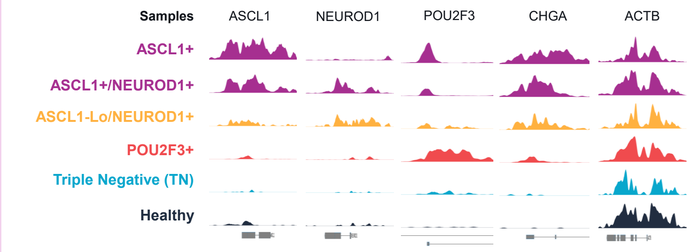


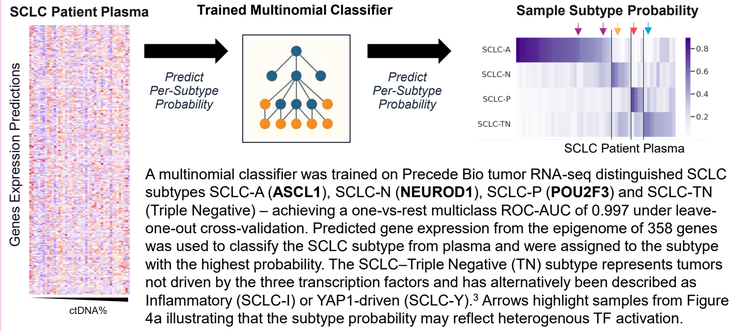
Figure 4. Epigenomic Activation at Defining TFs

Reveals Expected SCLC Subtypes and Evidence of Tumor Heterogeneity



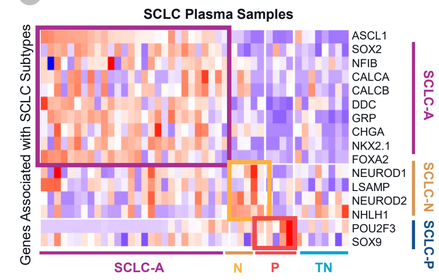
Epigenomic activation signals at genomic loci of subtype-defining transcription factors in SCLC plasma, normalized to **ACTB**. The top sample shows strong **ASCL1** activation, consistent with NE identity by **CHGA** activity, but lacks **NEUROD1** signal. The next two samples exhibit co-activation of **ASCL1** and **NEUROD1**, with ASCL1 predominating in the second, and **NEUROD1** in the third. The fourth sample shows specific **POU2F3** activation, while the fifth is quiescent across all three TFs. Healthy plasma is shown as a reference.

Models Estimate SCLC Subtype Probability and Capture Mixed TF Activity



A multinomial classifier was trained on Precede Bio tumor RNA-seq distinguished SCLC subtypes SCLC-A (**ASCL1**), SCLC-N (**NEUROD1**), SCLC-P (**POU2F3**) and SCLC-TN (Triple Negative) – achieving a one-vs-rest multiclass ROC-AUC of 0.997 under leave-one-out cross-validation. Predicted gene expression from the epigenome of 358 genes was used to classify the SCLC subtype from plasma and were assigned to the subtype with the highest probability. The SCLC–Triple Negative (TN) subtype represents tumors not driven by the three transcription factors and has alternatively been described as Inflammatory (SCLC-I) or YAP1-driven (SCLC-Y).[3] Arrows highlight samples from Figure 4a illustrating that the subtype probability may reflect heterogenous TF activation.

Predicted Subtyping Gene Expression Networks Captured in SCLC Plasma

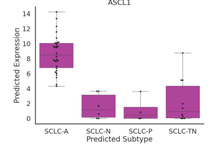


Expression predictions for genes with known associations to SCLC subtypes across SCLC patient plasma samples. SCLC patient samples (columns) were clustered by top predicted SCLC subtype using multinomial classifier. Expression predictions of each gene were z-scored across samples in heatmap for visualization.

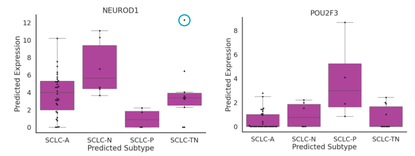
Top Subtype Prediction per Sample, Multinomial Classifier

Normalized Gene Expression Prediction

Predicted Expression Recapitulates Expected TF Patterns Across SCLC Subtypes

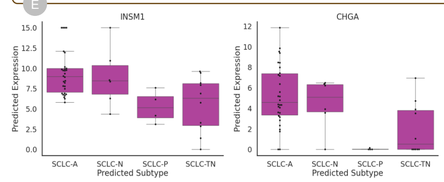


Predicted **ASCL1** expression differed significantly across subtypes (Kruskal–Wallis H=31.2, $p<1\times10^{-6}$), with higher expression in SCLC-A compared to the other groups (post-hoc pairwise Mann–Whitney U, FDR-adjusted p<0.01). **NEUROD1** expression also varied across subtypes (KW H=10.5, p=0.04), with significant elevation in SCLC-N relative to SCLC-P or SCLC-TN (FDR-adjusted p<0.05), but not compared to SCLC-A. Similarly, **POU2F3** expression differed across subtypes (KW H=8.3, p=0.04), although post-hoc tests did not reach significance after FDR correction when comparing SCLC-P to other subtypes.



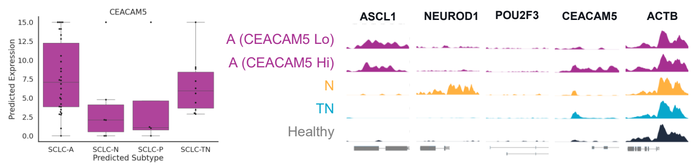
Highlighted SCLC-TN sample with high predicted NEUROD1 expression has a near equal SCLC-N, SCLC-P and SCLC-TN probability (24.3%, 29.3%, and 29.5%, respectively), suggesting that the gene regulatory network makes this sample particularly challenging to subtype. Box plots represent the median and interquartile range (IQR). Whiskers extend to the most extreme data within 1.5x the IQR.

Elevated Predicted Expression of Neuroendocrine (NE) Marker Genes in NE-associated Subtypes



INSM1 expression differed significantly across subtypes (Kruskal–Wallis H=12.8, p=0.005), with higher levels in predicted SCLC-A compared to SCLC-P or SCLC-TN (post-hoc Mann–Whitney U, FDR-adjusted p=0.012). **CHGA** also varied across subtypes (KW H=14.7, p=0.002), with elevated expression in SCLC-A and SCLC-N compared to SCLC-P (MWU FDR-adjusted p<0.04) and in SCLC-A compared to SCLC-TN (MWU FDR-adjusted p=0.008).

Figure 5. CEACAM5 Expression and Epigenomic Activity Extends Beyond SCLC-A, Enabling Stratification Beyond Subtype Boundaries



CEACAM5 is an emerging ADC target in SCLC and has been linked to **ASCL1** gene expression.[5] *(Left)* Predicted expression of **CEACAM5** appeared enriched in the SCLC-A subtype but did not reach statistical significance across groups. *(Right)* Epigenomic enhancer profiling confirmed dynamic **CEACAM5** activity in SCLC-A and revealed activation in additional subtypes, including SCLC-TN. Together, these findings indicate that **CEACAM5** expression extends beyond subtypes and may provide a basis for patient stratification in **CEACAM5**-targeted therapies.

CONCLUSIONS

SCLC is an aggressive malignancy with limited therapeutic options and few tools to guide treatment selection. A major unmet need in this setting is the ability to identify and stratify patients based on targetable biology. We demonstrate that plasma-based epigenomic profiling enables robust prediction of gene expression and molecular subtyping directly from ctDNA, maintaining performance at low ctDNA fractions, including clinically relevant ADC targets such as SEZ6, DLL3 and CEACAM5. This approach not only recapitulates known subtype biology but also resolves intra-tumoral heterogeneity and can capture dynamic changes through the course of treatment. By providing a versatile, non-invasive platform for tumor characterization, longitudinal monitoring, and therapy selection, this strategy has the potential to support patient identification and treatment guidance in SCLC's evolving therapeutic landscape.

Disclosures The first and presenting authors (Khoi Nguyen and Aparna Gorthi) are employed by Precede Biosciences and have no other interests to be declared.

References

1. Husain, H. et al. *JCO Precis. Oncol.* **6**, e2200261 (2022), and internally acquired cohort.
2. George, Julie et al. "Comprehensive genomic profiles of small cell lung cancer." *Nature* vol. 524,7563 (2015): 47-53. doi:10.1038/nature14664
3. Rudin, C. M. et al. *Nat. Rev. Cancer* **19**, 289–297 (2019).
4. Ajay, A. et al. *npj Precis. Oncol.* **9**, 1 (2025).
5. Tlemsani, C. et al. *Cell Rep.* **33**, 108296 (2020).

Contact J. Carl Barrett, PhD
Precede Biosciences
Email: carl.barrett@precede.bio