Artificial intelligence (AI)-based classification of stromal subtypes reveals associations between stromal composition and prognosis in NSCLC

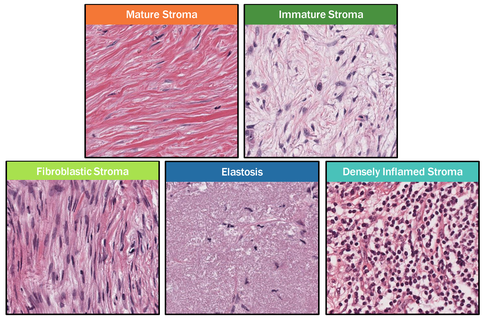
Abstract #5447

STUDY BACKGROUND

- Cancer-associated stroma (CAS) has long been appreciated as an important histological feature in many cancer types. Recently, single-cell molecular analyses have revealed the heterogeneity of CAS. Furthermore, the cellular composition of CAS has been linked to prognosis in several cancer types, including non-small cell lung cancer (NSCLC)[1].
- While pathologists can manually classify CAS based on the architecture of the extracellular matrix and the cells within it, measurement of these regions is difficult and not reproducible.
- To this end, we have developed an artificial intelligence (AI)-based model to sub-classify CAS in hematoxylin and eosin (H&E)-stained whole slide images (WSIs). Tissue and cell human interpretable features (HIFs) extracted from our model were assessed for their association with clinicopathologic features (e.g., stage and overall survival) and their ability to predict known stromal gene expression signatures.

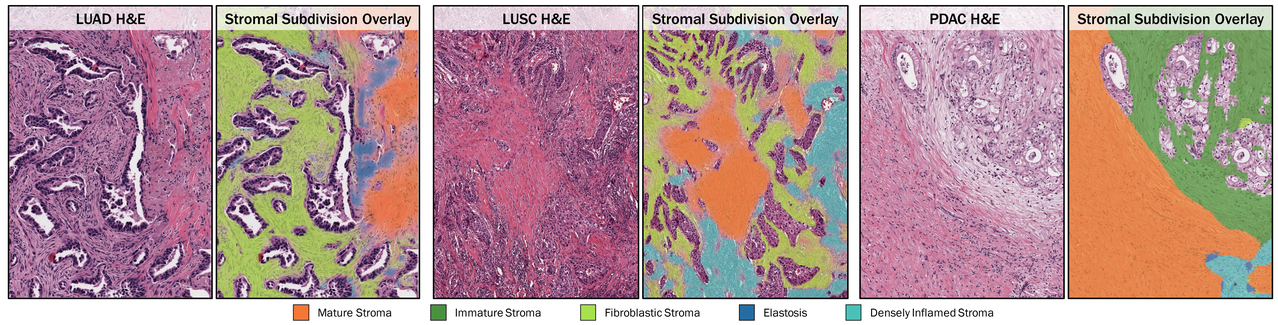
PATHAI STROMAL SUBTYPE MODELS ACROSS CANCERS

Figure 1. Stromal subtypes identified by stromal subdivision model.



All images are of H&E-stained specimens that were captured at 30X magnification with the same pixel dimensions (484 x 484 pixels).

Figure 2. Example of stromal subdivision model performance in LUAD, LUSC, and PDAC.

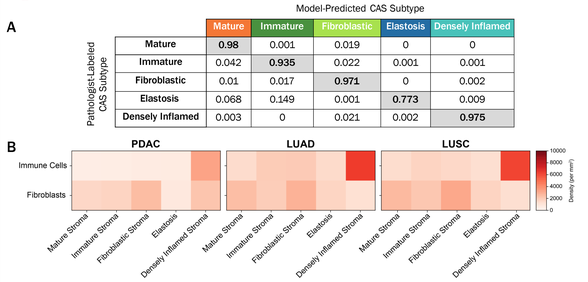


Overlay images show model-identified mature stroma, immature stroma, fibroblastic stroma, elastosis, and densely inflamed stroma superimposed on each H&E image. The absence of overlay indicates the presence of normal tissue, cancer epithelium, or necrosis.

METHODS

Model Training and Development. We developed a convolutional neural network-based model to classify CAS as immature, mature, densely inflamed, fibroblastic, or elastosis (Fig. 1). This model was trained using manual pathologist-derived annotations (N=3019) of H&E-stained whole slide images (WSIs) of pancreatic ductal adenocarcinoma (PDAC) obtained from the TCGA (N=126). This stromal subdivision model was deployed on H&E-stained lung adenocarcinoma (LUAD; N=468) and lung squamous cell carcinoma (LUSC; N=430) WSIs. Model output is depicted in Fig. 2. Model performance was assessed by qualitative review by expert pathologists and quantitative evaluation of model predictions compared to pathologist CAS label (Fig. 3A) and characterized by cell content analysis (Fig. 3B).

Figure 3. Evaluation of stromal subdivision model performance.



A) Agreement between model-predicted and pathologist label of stromal subtypes. B) Quantification of cell types present in model-predicted stromal subtypes. Fibroblasts and immune cells (macrophages, lymphocytes, and plasma cells) were examined. As expected, fibroblastic and densely inflamed stroma show greater densities (per mm^2) of fibroblasts and immune cells, respectively.

Exploratory Analyses. HIFs were extracted from the stromal subdivision model (e.g., proportional area of mature relative to total stroma). We evaluated each HIF independently for its association with tumor ordinal stage, and with overall survival (OS) after adjusting for age, sex, and tumor stage. These stromal HIFs, along with cell and tissue features, were also used to predict molecular biomarkers, such as the immune-related TGEP signature[2] and CAF-related LRCC15-CAF[3] and pCAF signatures[4] using univariate linear regression and multivariable random forest (RF) regression models. HIF impact on the RF model predictions was calculated using Shapley Additive Explanations (SHAP values).

RESULTS

Figure 4. Stromal composition of LUAD, LUSC, and PDAC.

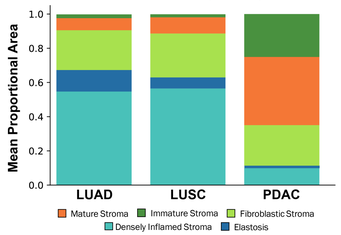
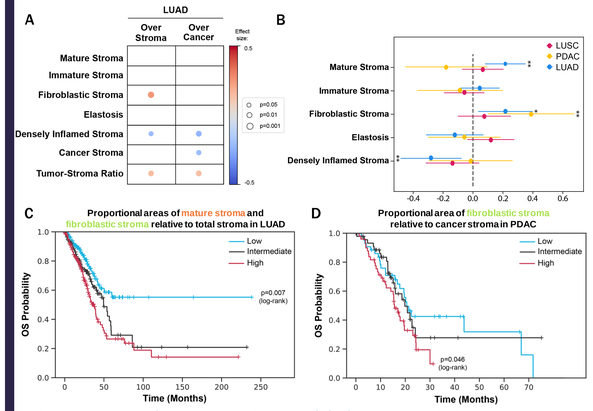
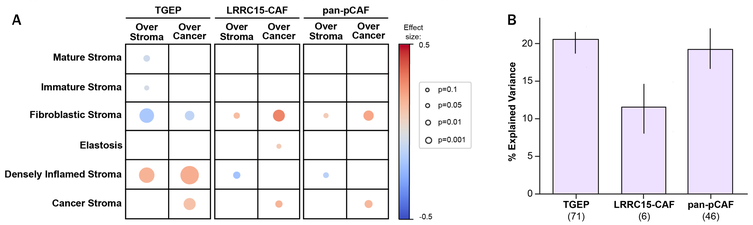


Figure 5. Effect of stromal composition on stage and overall survival.

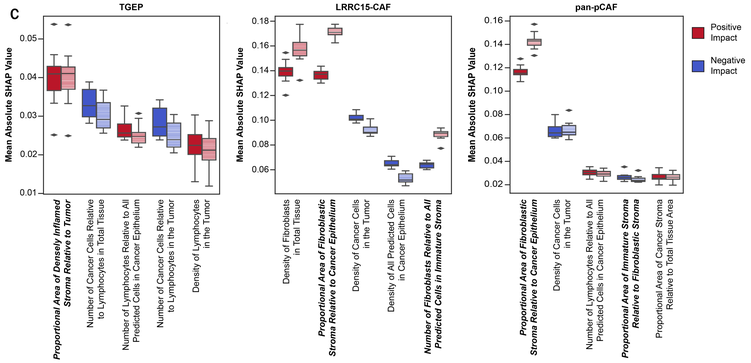


A) Association of proportional area of CAS subtypes (relative to cancer or stroma) with stage in LUAD. No significant relationships between stromal composition and stage were observed in LUSC or PDAC (not shown). B) Association of model-predicted stromal composition and overall survival. C) In LUAD, tertile analyses revealed an association of low combined proportional areas of mature and fibroblastic stroma relative to total cancer stroma with improved OS. D) In PDAC, a similar approach revealed that elevated fibroblastic stroma relative to cancer stroma associated with poor OS. The observed association in LUAD passed multiple hypothesis testing, while the association in PDAC did not.

Figure 6. Association of stromal subtypes with stromal gene expression in LUAD.



A) Published gene expression signatures[2-4] were regressed against the proportional areas of model-predicted stromal subtypes. B) Average explained variance in RF model performance in the test set. Error bars show the range in explained variance for each model, while the numbers in parentheses indicate the number of HIFs used in each model.



C) Mean absolute values of SHAP, representing the additive contribution of each individual feature in presence of other features, were calculated for all HIFs in the prediction of TGEP, LRRC15-CAF, and pan-pCAF scores. The five most impactful HIFs for the prediction of signatures in the test set (striped boxes; N=93) and training set (solid boxes; N=370) are shown. Positive or negative impact of each feature represents the correlation between SHAP and feature values. Stromal HIFs (bold text) were highly impactful for predicting these signatures, especially pan-pCAF (previously associated with poor prognosis in multiple tumor types[4]), which was particularly dependent on the amount of fibroblastic stroma relative to cancer.

CONCLUSIONS

We developed a first of its kind model to predict CAS subtypes in H&E-stained tissue. Model-predicted stromal HIFs were associated with stage in LUAD and OS in LUAD and PDAC. Model extracted HIFs were also observed to associate with stromal gene expression in LUAD. Notably, enrichment of fibroblastic and densely inflamed stromal subtypes led to opposite effects on tumor stage, OS, and gene expression. Current efforts are exploring the utility of stromal HIFs for predicting additional clinically relevant parameters, such as treatment response, in various solid tumor types.

AUTHORS

Fedaa Najdawi[1], Sandhya Srinivasan[1], Neel Patel[1], Nhat Le[1], Michael G. Drage[1], Christian Kirkup[1], Ylaine Gerardin[1], Chintan Parmar[1], Jacqueline Brosnan-Cashman[1], Michael Montalto[1], Andrew H. Beck[1], Archit Khosla[1], Ilan Wapinski[1], Ben Glass[1], Murray Resnick[1], Matthew Bronnimann[1]

[1]PathAI, Boston, MA.

CONTACT

Fedaa Najdawi: fedaa.najdawi@pathai.com
Matthew Bronnimann: matthew.bronnimann@pathai.com

REFERENCES

1) Lambrechts, D., *et al. Nat Med.* 2018; 24(8):1277-1289.
2) Ayers, M., *et al. J Clin Invest.* 2017; 127(8)2930-2940.
3) Dominguez, C.X., *et al. Cancer Discov.* 2020; 10(2):232-253.
4) Galbo, P.M., *et al. Clin Cancer Res.* 2021; 27(9):2636-2647.

ACKNOWLEDGMENTS

We would like to thank the members of PathAI Pathology Network who contributed towards this work.

This poster template was developed by SciStories LLC. https://scistories.com/

PathAI

AACR Annual Meeting, April 14-19, 2023, Orlando, FL, USA