Correlation of immune phenotypes derived from H&E-stained whole slide images with prognosis and response to checkpoint inhibitors in NSCLC

Nhat Le[1]*, Laura Dillon[2]*, Ciyue Shen[1], Matthew Bronnimann[1], Jun Zhang[1], Jennifer Hipp[1], Tan Nguyen[1], Michael Griffin[1], Michael Nercessian[1], Jacqueline Brosnan-Cashman[1], Robert Egger[1], Jean-Philippe Guégan3, Antoine Italiano[4], Xinwei Sher[2], G. Travis Clifton[2]*, Bahar Rahsepar[1]*

[1]PathAI, Inc., Boston, MA
[2]Incendia Therapeutics, Boston, MA
[3]Explicyte, Bordeaux, France
[4]Institut Bergonié, Bordeaux, France
* Contributed equally to study

Background
The classification of tumors as inflamed, excluded or desert based on spatial patterns of tumor infiltrating lymphocytes (TILs) is a potential biomarker of patients likely to respond to checkpoint inhibitors (CPI). However, the subjectivity of manual methods to assess these immune phenotypes (IPs) and poor standardization in the methods and thresholds to define IPs have hampered their clinical adoption. Here, we describe a data-driven approach to inform IP threshold selection on hematoxylin and eosin (H&E)-stained whole slide images (WSI) by maximizing differences in overall survival (OS) between IPs.

Methods
A model to classify the IPs of NSCLC samples from H&E images was developed using PathExplore™ models applied to a TCGA non-small cell lung cancer (NSCLC) cohort of LUAD (N=459) and LUSC (N=424). TIL densities were extracted within cancer and stroma for 0.01 mm^2 patches tiled across each WSI. Cut-offs to define cancer and stroma patches as hot or cold were defined based on the 75^{th} and 50^{th} percentiles, respectively, of cancer TIL (cTIL) densities (386 cTIL/mm^2) and stroma TIL (sTIL) densities (423 sTIL/mm^2) in a TCGA cohort of 4,082 H&E-stained WSI from 10 epithelial tumor types. Hierarchical fitting yielded optimal thresholds minimizing the p-values of OS differences between IPs, leading to classifications of inflamed (iIP, >40% cancer hot patches), excluded (eIP, ≤40% cancer hot patches; >45% stromal hot patches) and desert (dIP, ≤40% cancer hot patches; ≤45% stromal hot patches). The model was then deployed in a clinical cohort of PD-(L)1 inhibitor-treated NSCLC patients (N=95) enrolled in the BIP precision medicine study (NCT02534649; Institut Bergonié, Bordeaux, France). Model-predicted IPs were compared to progression-free survival (PFS) and OS. FDR correction was done with Benjamini-Hochberg.

Results
In the TCGA NSCLC cohort, model-predicted iIP (N=196) and eIP (N=607) patients had significantly better OS compared to dIP (N=80; HR=0.53, p=0.003 and HR=0.59, p=0.003, respectively). In the clinical cohort, cTIL density and fraction of hot epithelial patches were significantly associated with PFS (HR=0.64, q=0.04 and HR=0.69, q=0.04, respectively). PFS was significantly shorter in model-predicted eIP patients (N=46) compared to iIP (N=39; HR=0.54, p=0.045). Notably, in PD-L1(-) patients (N=43, tumor proportion score ≤1%), iIP

patients had longer PFS than eIP and dIP patients (HR=0.35, p=0.02). No difference in PFS was observed for PD-L1(+) patients.

Conclusions

We developed a data-driven approach for predicting IPs using patch-level TIL features. Model-predicted IPs were prognostic in a TCGA NSCLC dataset and predictive of PFS in a CPI-treated clinical NSCLC cohort. Association of IP and PFS was independent of PD-L1 status, potentially allowing the identification of PD-L1(-) patients who may derive greater benefit from CPI.