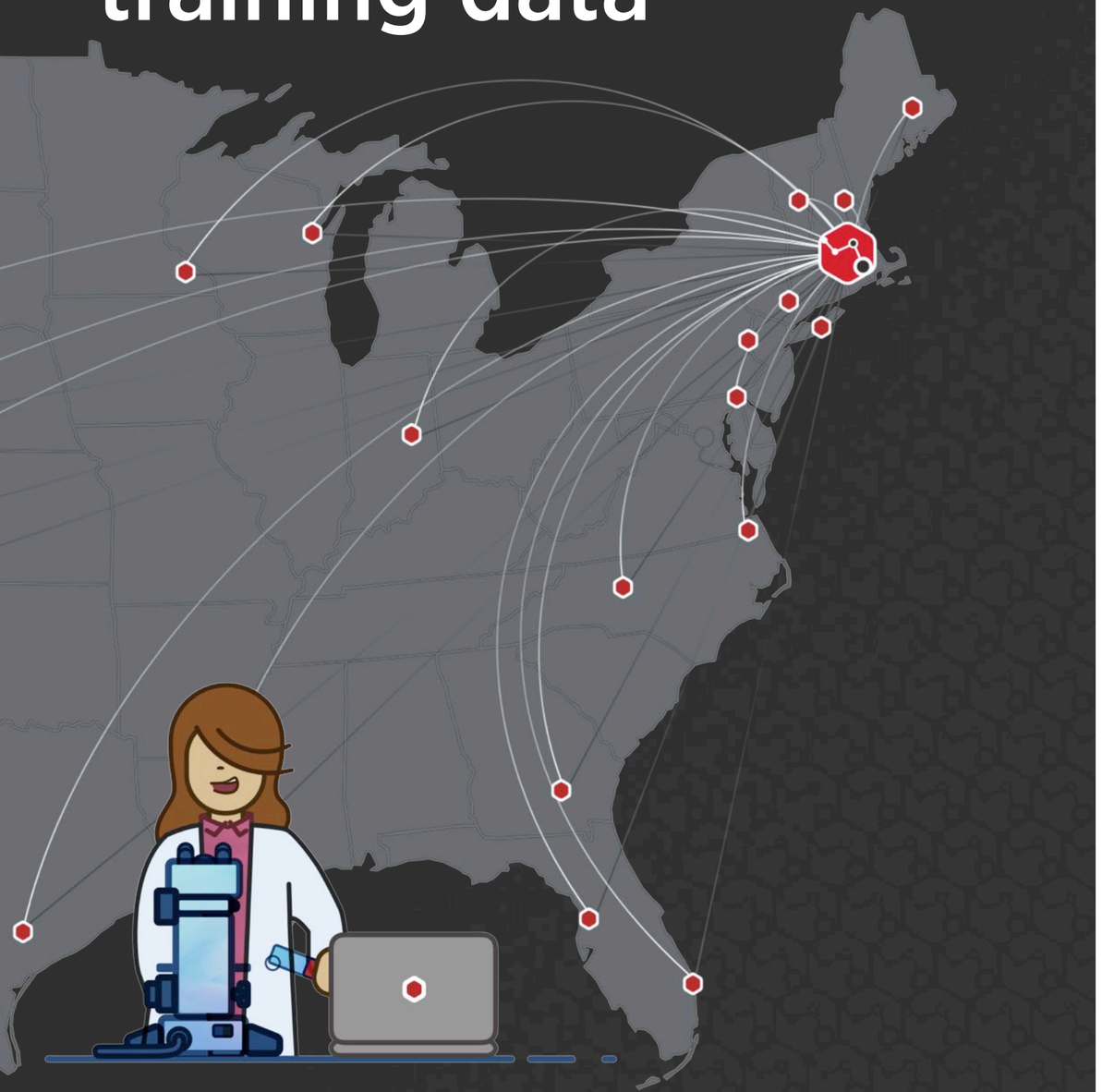
Can we do better?

- Deep learning is data hungry
  - Need 10s of thousands of precise cell annotations

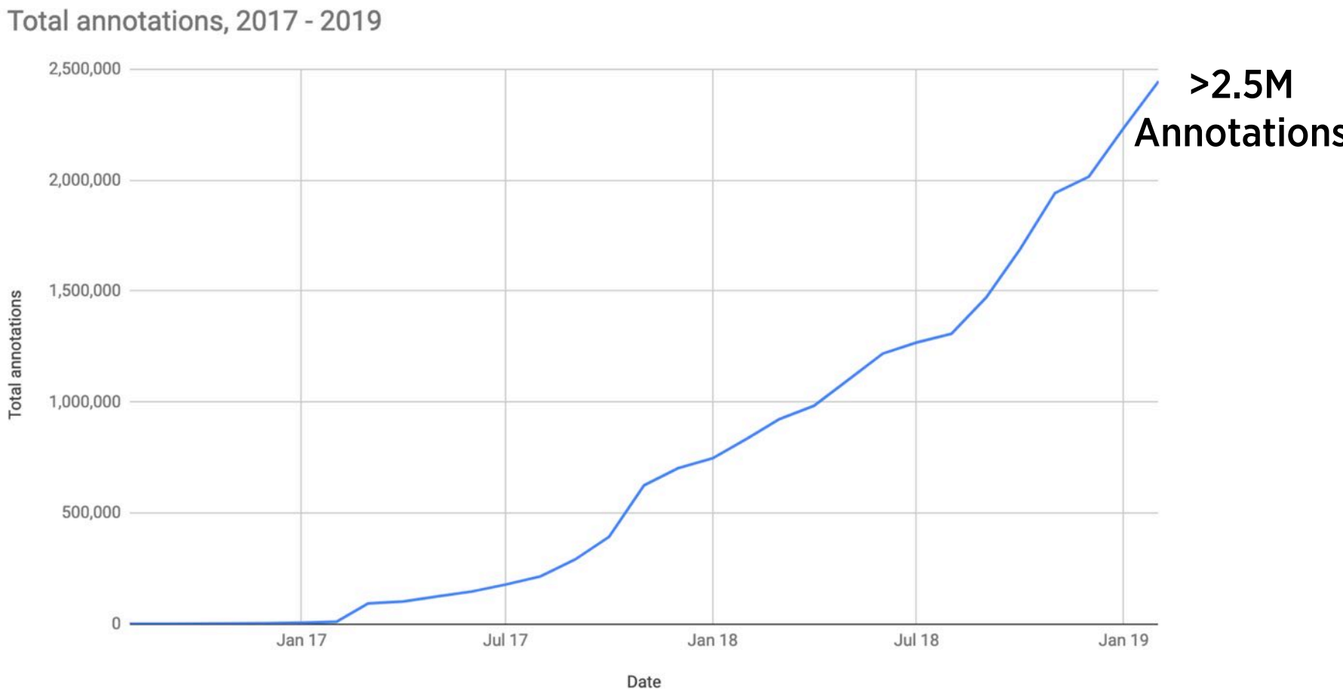
First, we need the data

PathAI

Board-certified training data



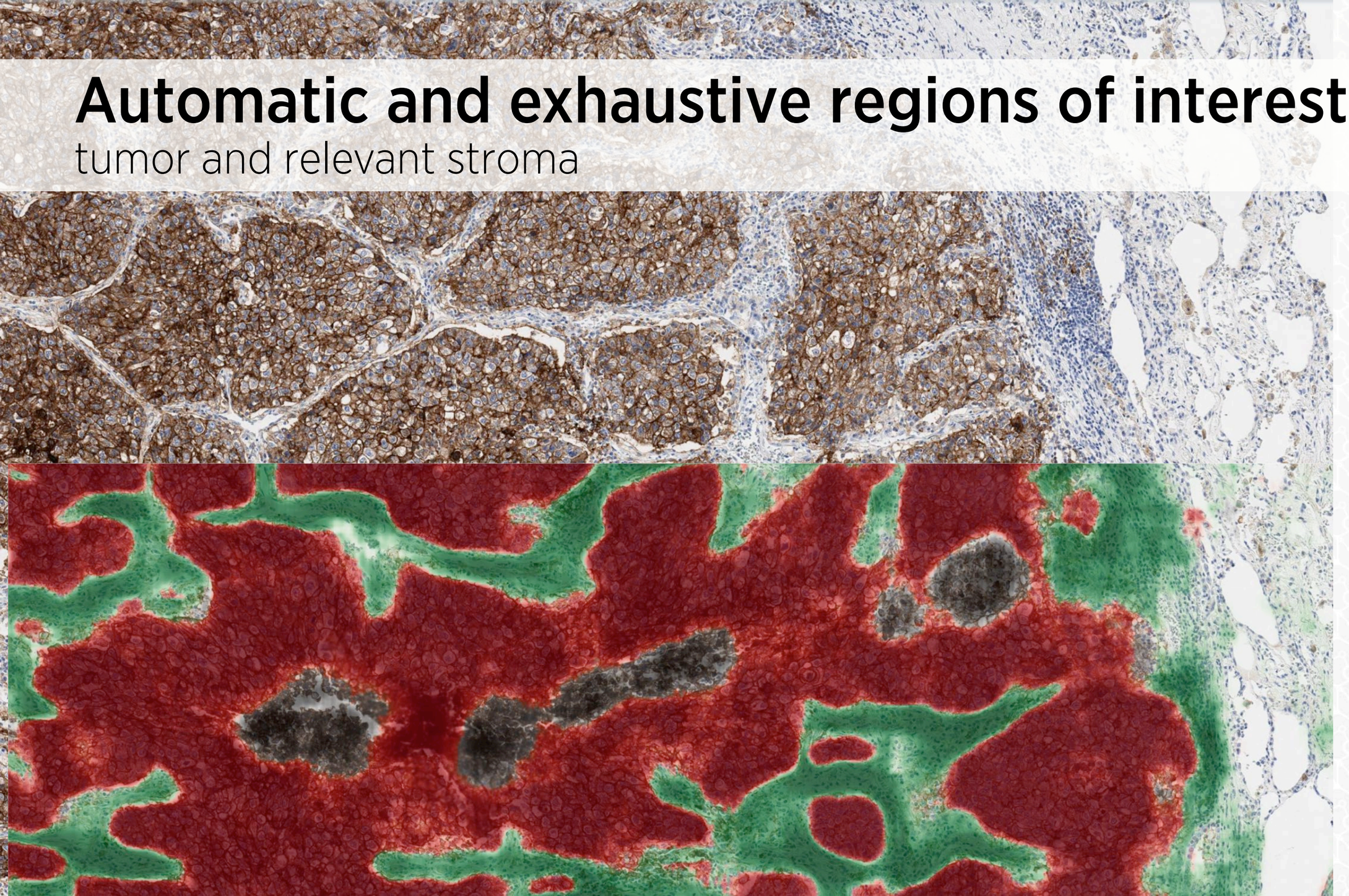
Working with pathologists around the country to generate high-quality annotations



>2.5M Annotations

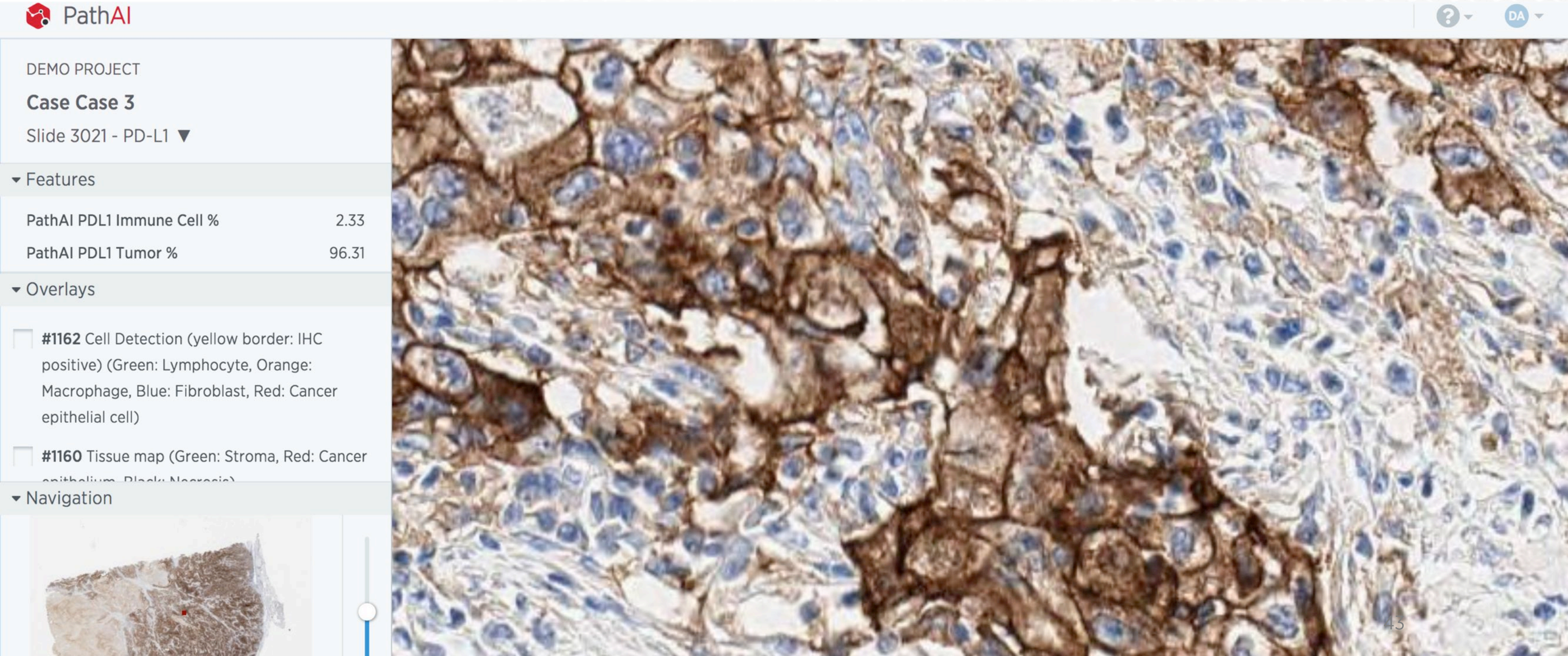
Automatic and exhaustive regions of interest

tumor and relevant stroma



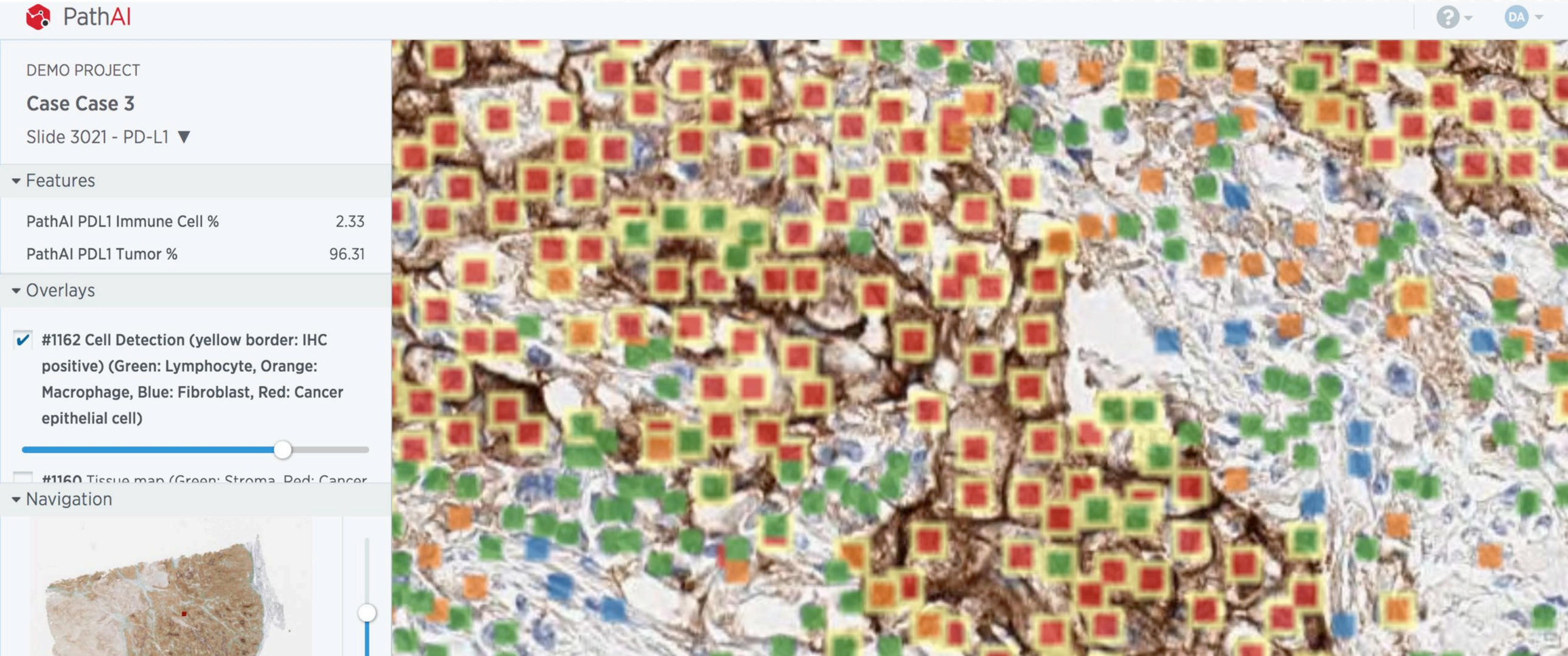
IHC expression difficult to detect

on immune and tumor cells



Exhaustive automated classification

Cell type and cellular IHC positivity classification

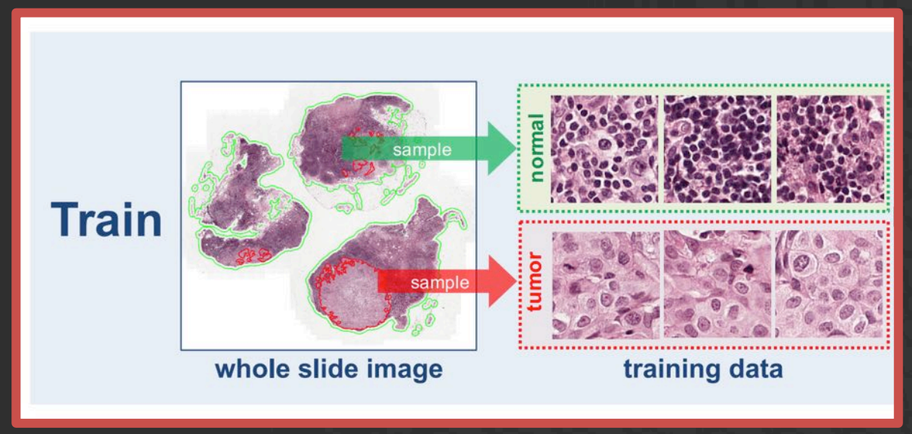


Quantitative and reproducible PD-L1 scoring

- Manual review: few hundred cells over a few arbitrary high-power fields of view
- Automated analysis: exhaustive classification of 10k-1million cells

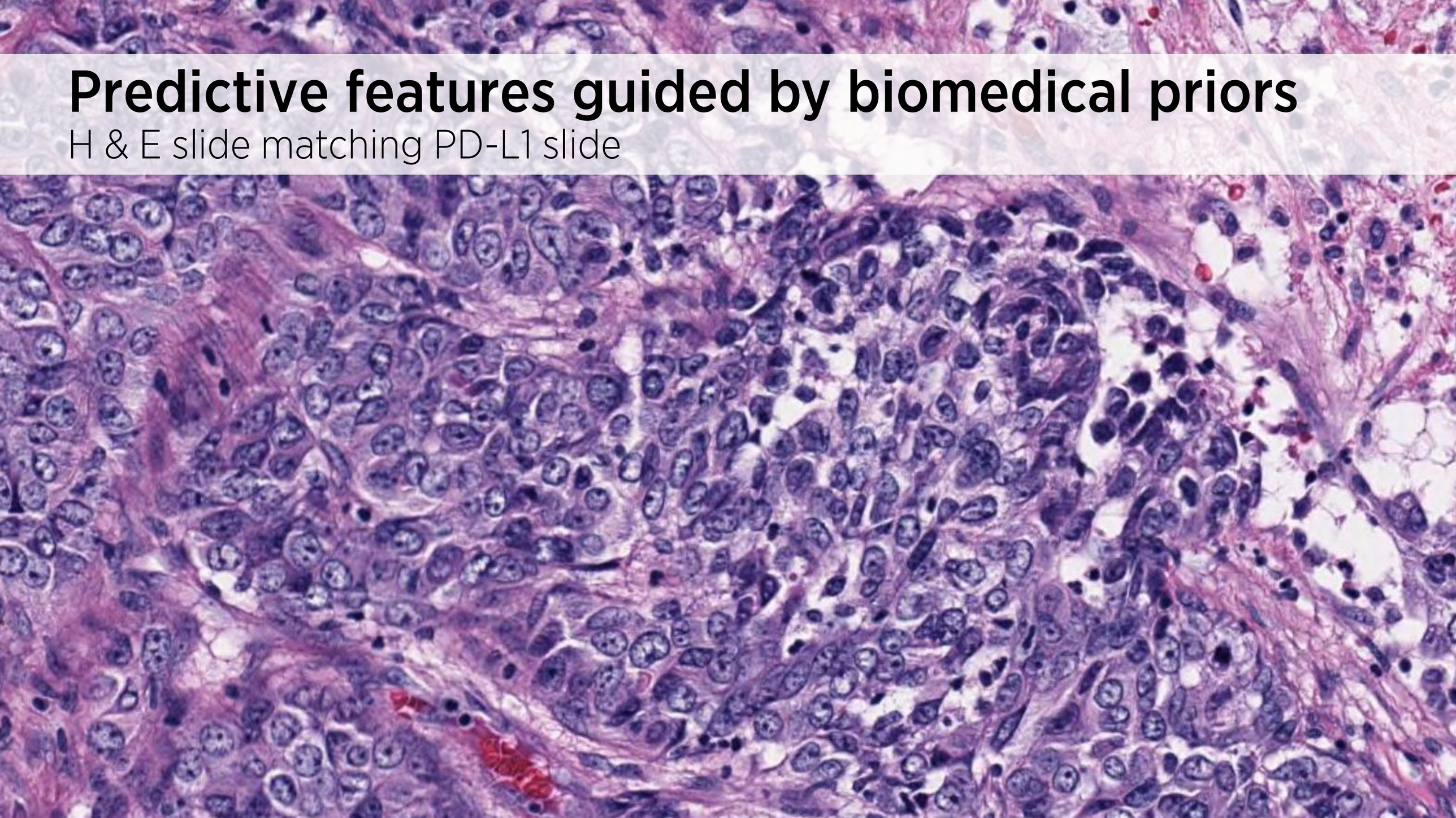
Taking it further
From quantitative assay to patient prediction

- PD-L1 scoring alone reduces billions of pixels to 1-2 numbers.
- Can we identify additional relevant information?
  - Using data from randomized controlled clinical trials
- However: Millions of patches, *hundreds* of patients



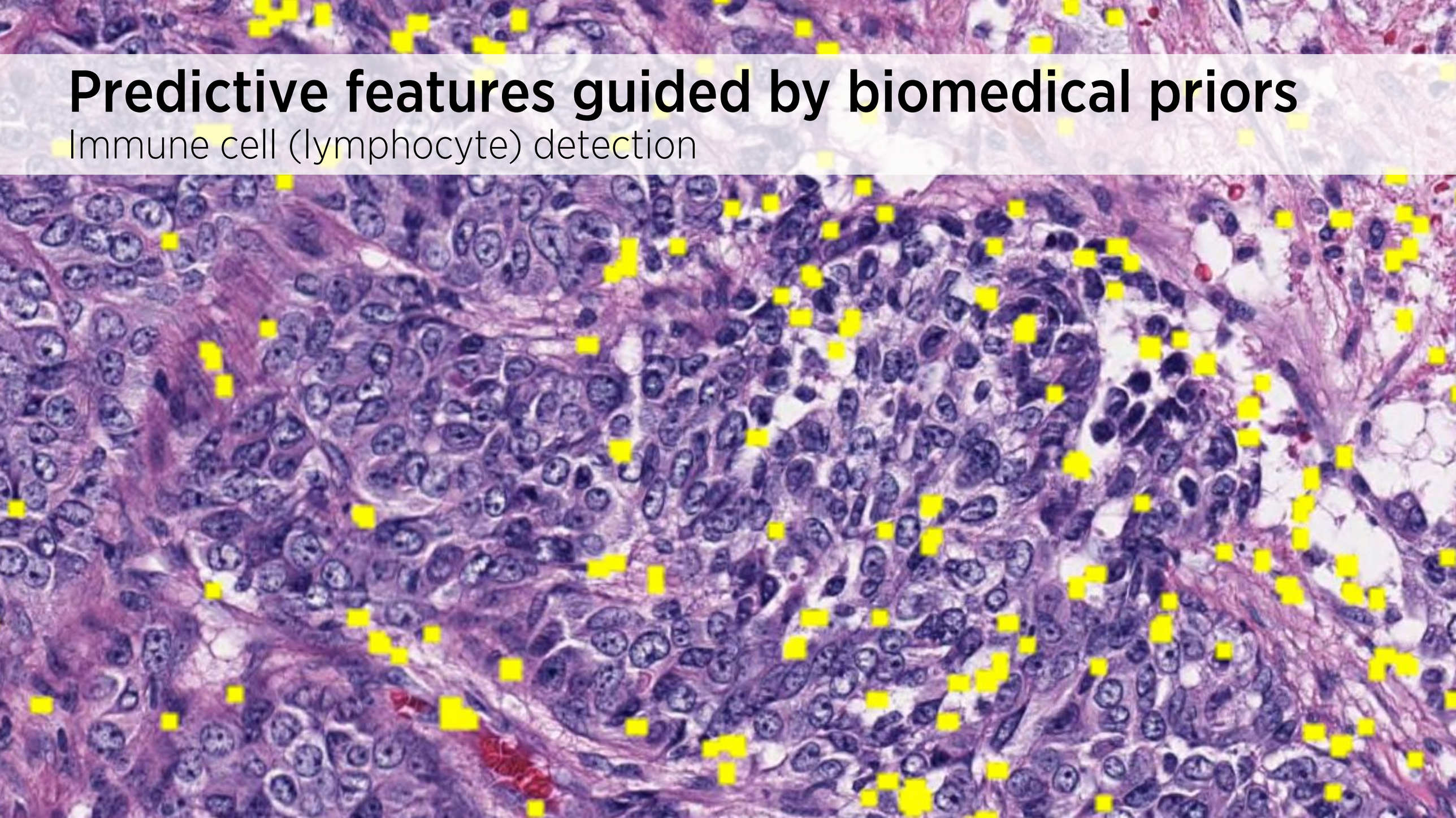
Predictive features guided by biomedical priors

H & E slide matching PD-L1 slide



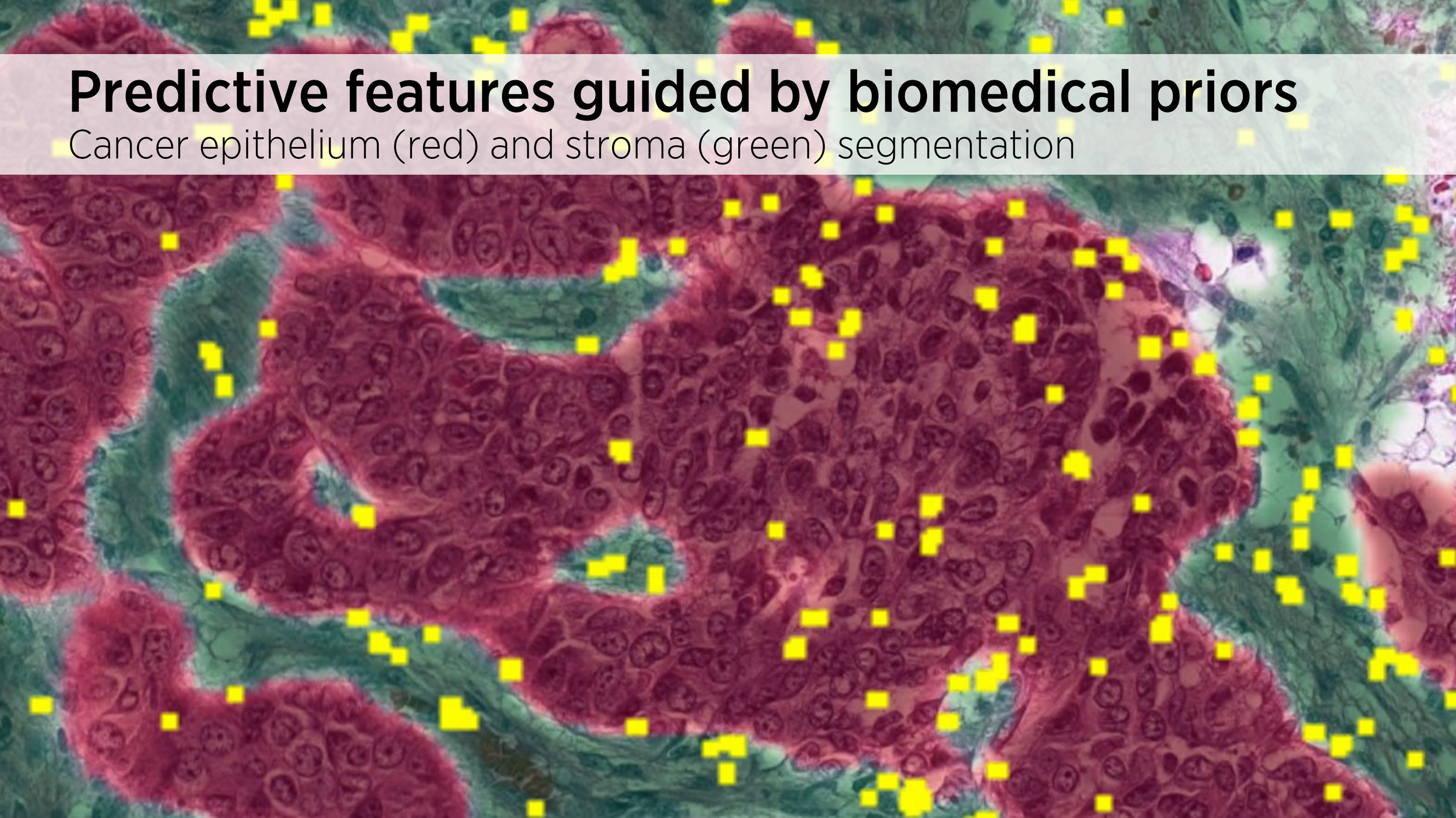
Predictive features guided by biomedical priors

Immune cell (lymphocyte) detection



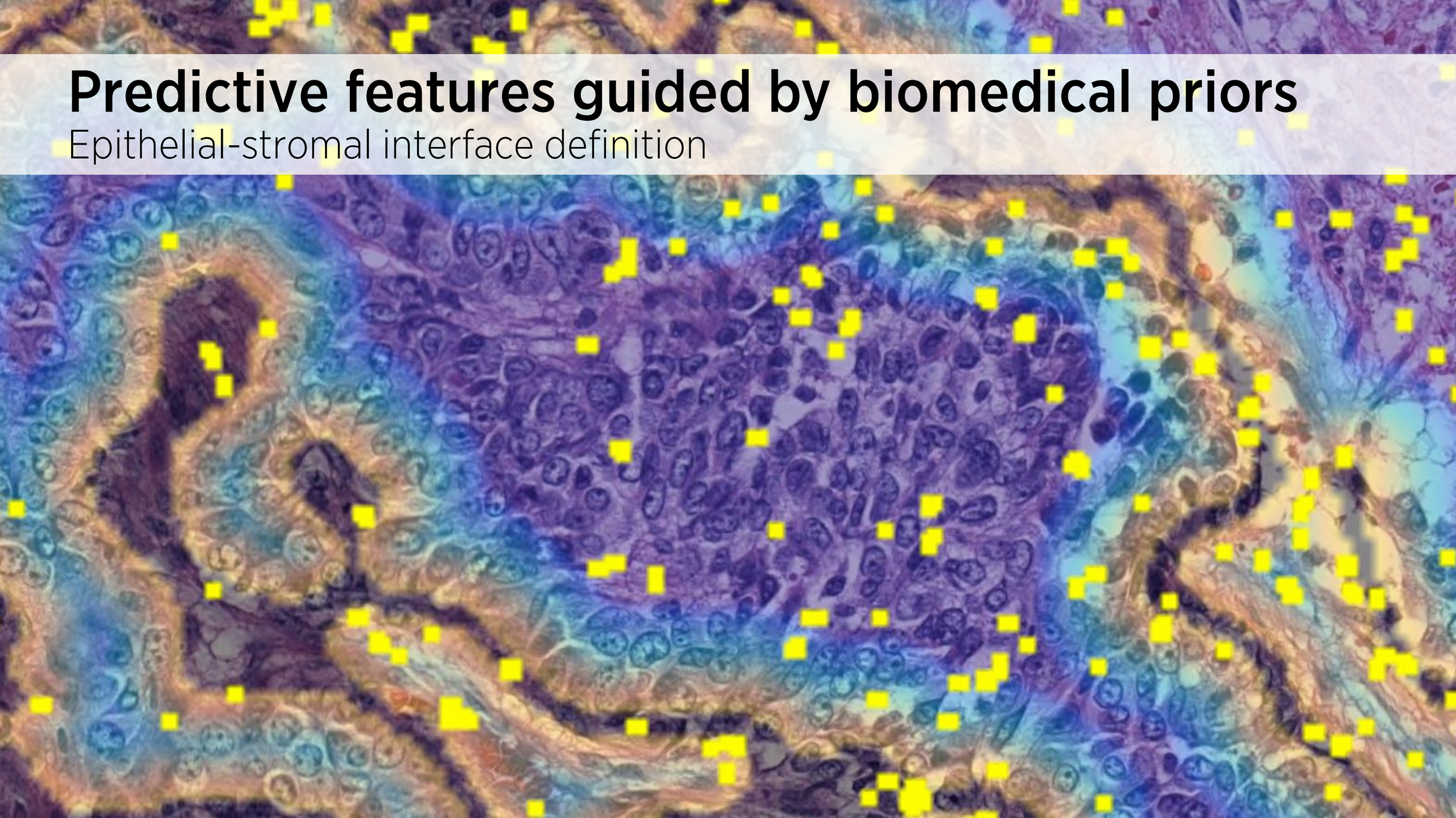
Predictive features guided by biomedical priors

Cancer epithelium (red) and stroma (green) segmentation

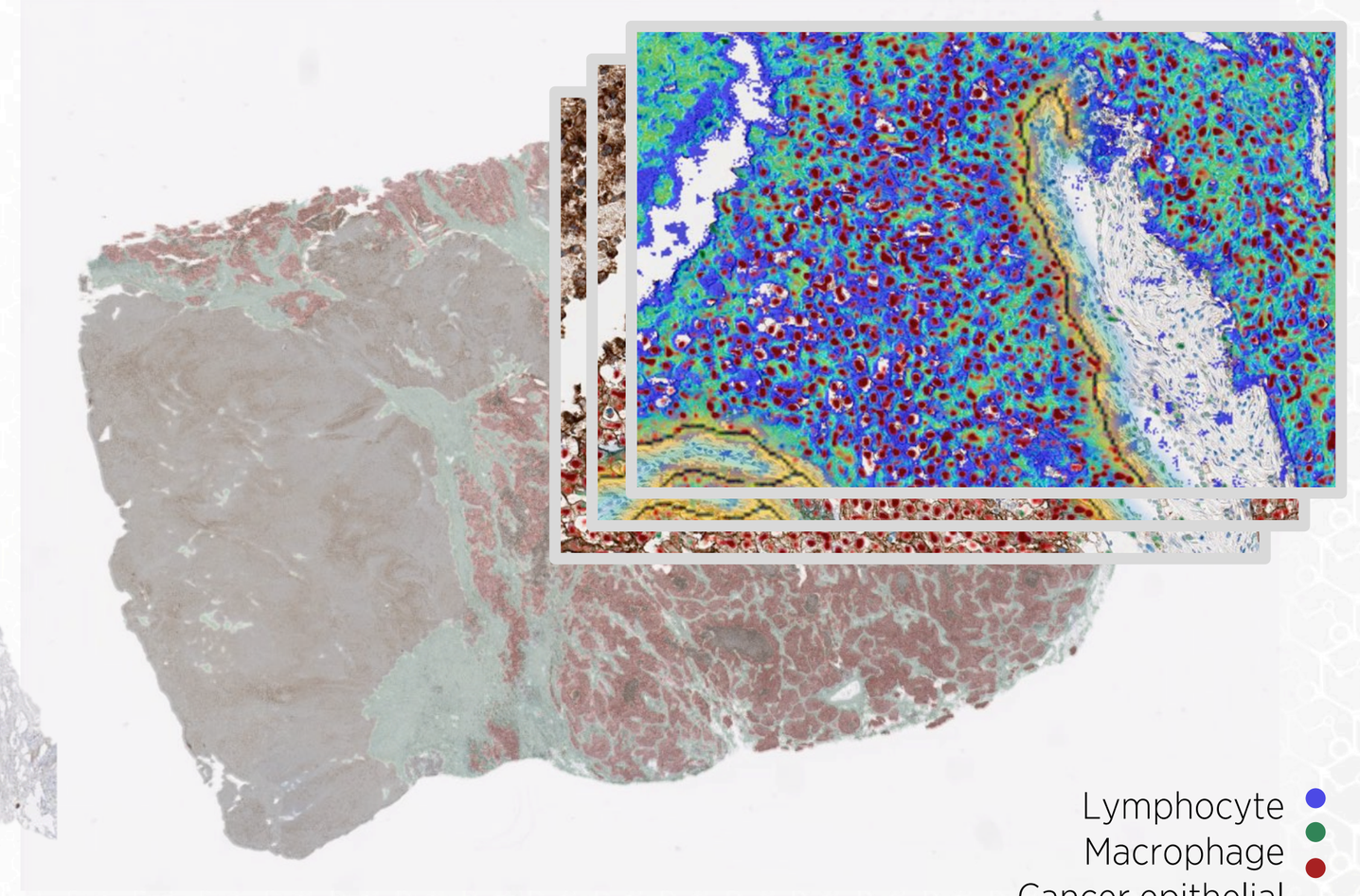


Predictive features guided by biomedical priors

Epithelial-stromal interface definition



Cell-type specific, tissue context-aware IHC-quantification



Lymphocyte
Macrophage
Cancer epithelial cell

Data-driven identification of pathological phenotypes associated with drug response

Total number of macrophages in epithelial/stroma interface (80um)

Total number of macrophages in epithelial/stroma interface (120um)

Total number of macrophages in invasive margin (250um)

Total number of lymphocytes in epithelial/stromal interface on H&E stain

Total number of plasma cells in epithelium on H&E stain

Total number of plasma cells in stroma on H&E stain

Tumor (epithelium + stroma) area on H&E stain

Total number of plasma cells in epithelial/stroma interface (40um)

Total number of plasma cells in epithelial/stroma interface (80um)

Area (mm^2) of epithelial/stroma interface (80um) target positive cancer cells on target stain

Area (mm^2) of epithelial PDL-1 positive macrophages on target stain

Necrosis area on target stain

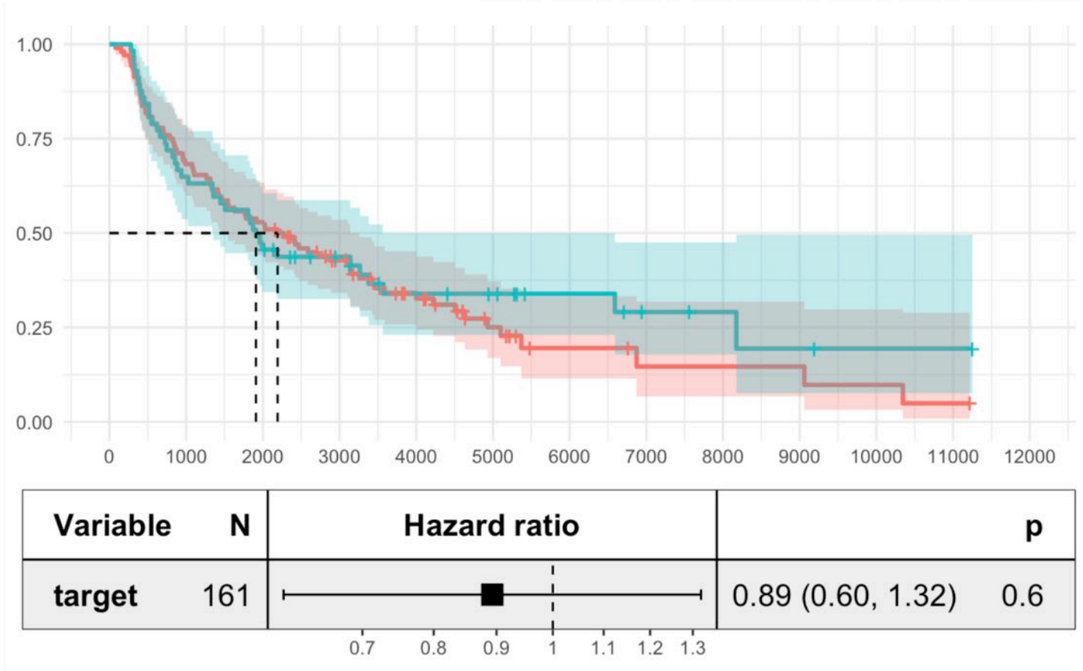
Proportion of tumor infiltrating lymphocytes engaged by target positive macrophages

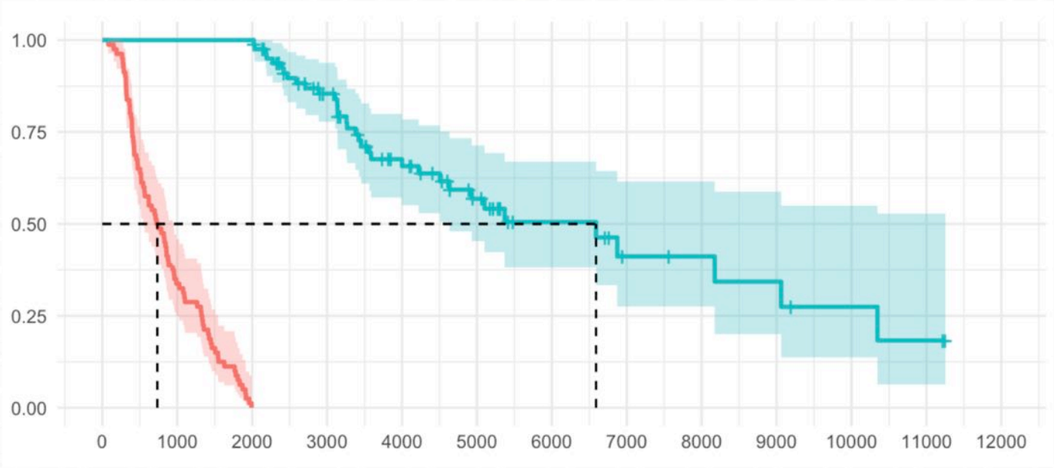
Stroma area on target stain

Tissue area on target stain

Multivariate models predictive of IO response

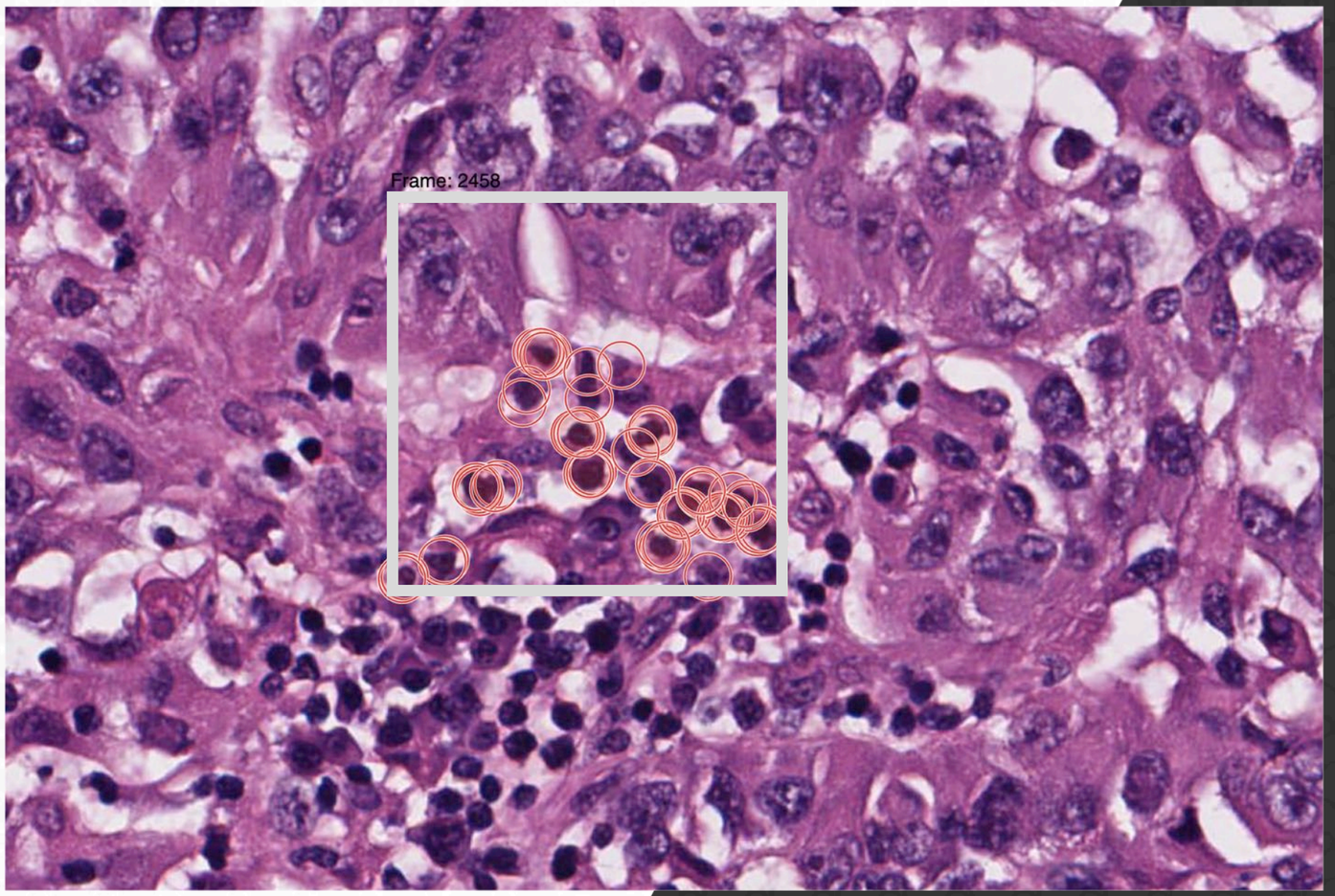
- Low *n*, interpretability and measures of uncertainty valuable:
  - No deep learning (gasp!)
- Feature importance/selection in these models can provide disease insight
  - *Now we're doing things pathologists can't rather than automating / improving what they already can*





Note: KM curves for illustration only

How do we know these features are *correct?*

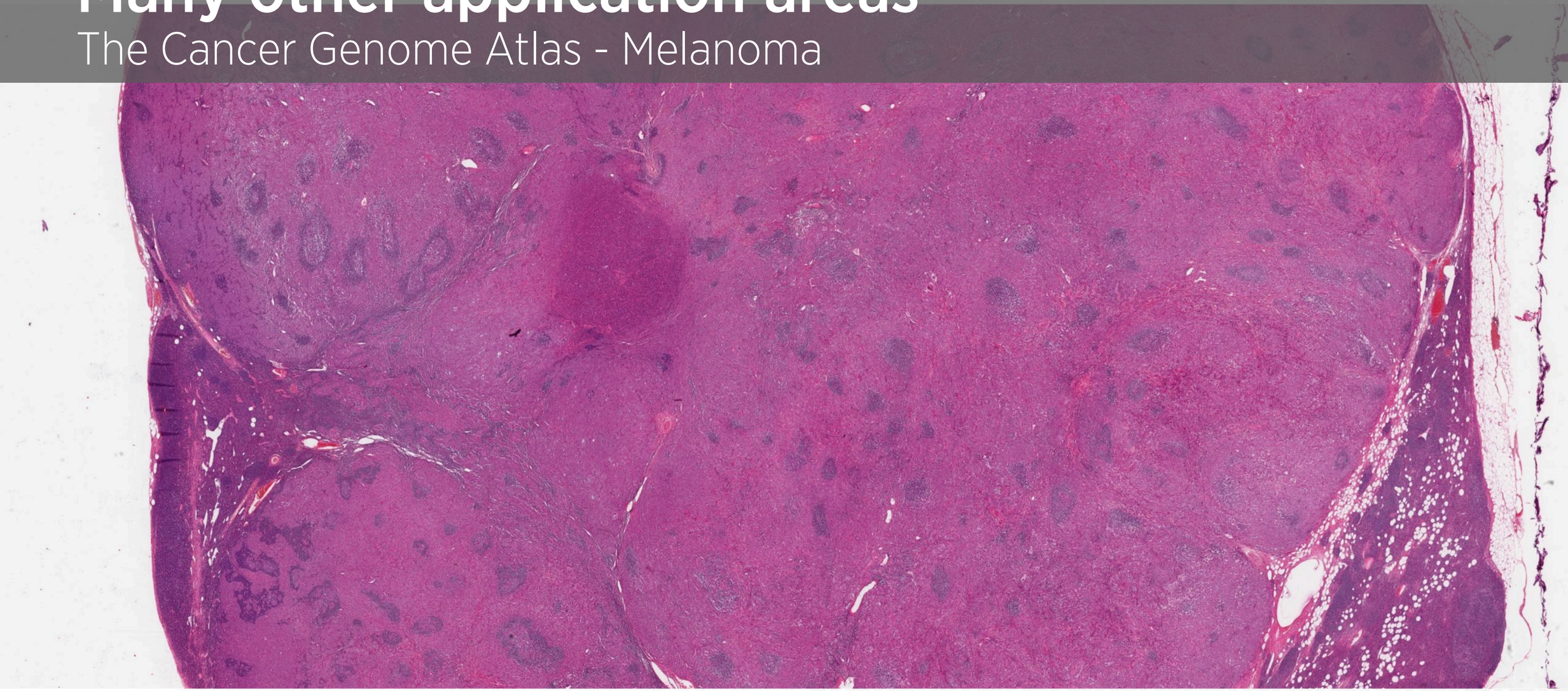


Frames

Validation by exhaustive consensus

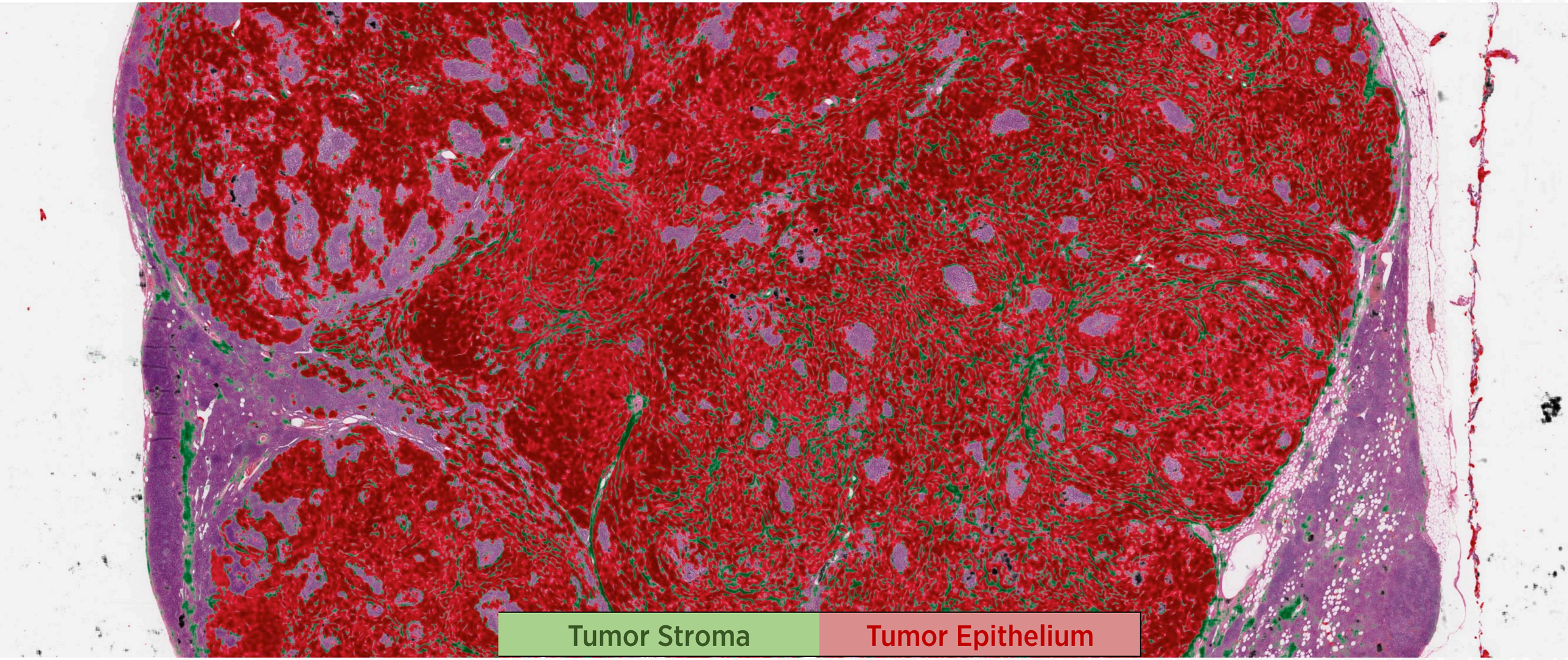
Many other application areas

The Cancer Genome Atlas - Melanoma



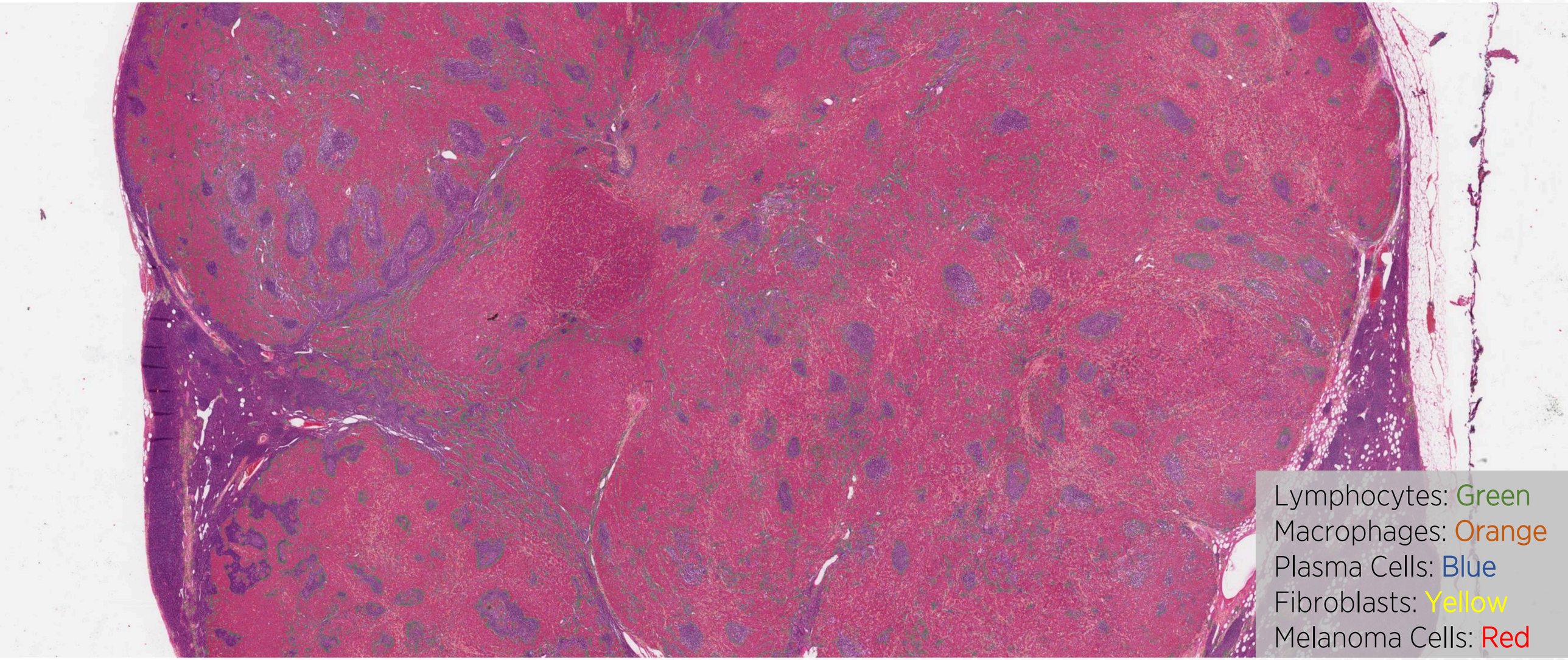
TCGA-EE-A2GL, Malignant Melanoma

Melanoma Tissue Map



TCGA-EE-A2GL, Malignant Melanoma

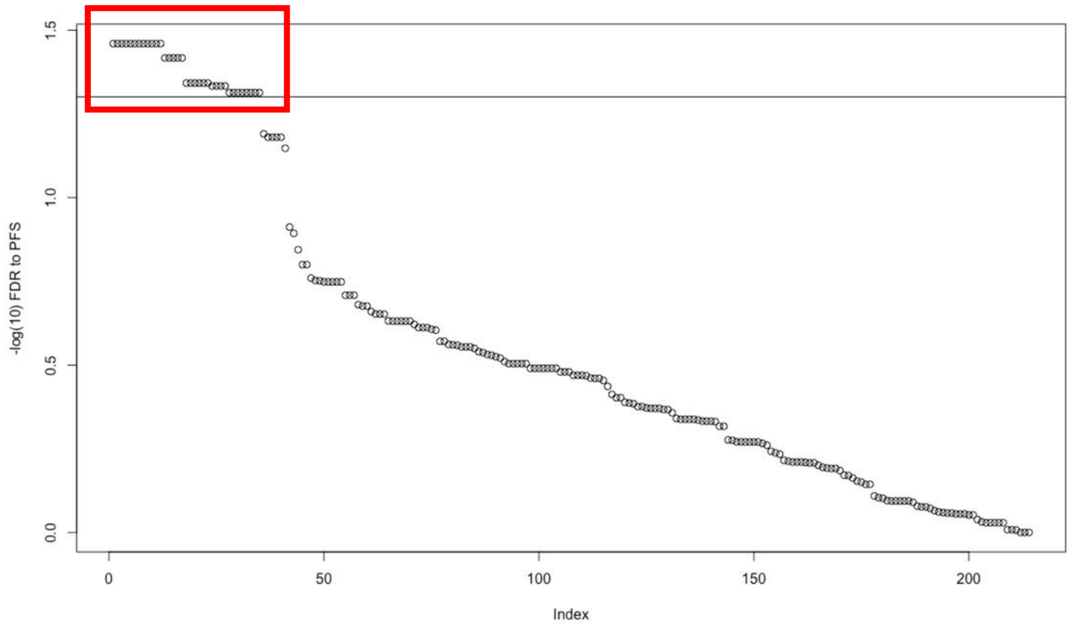
Melanoma Cell Map



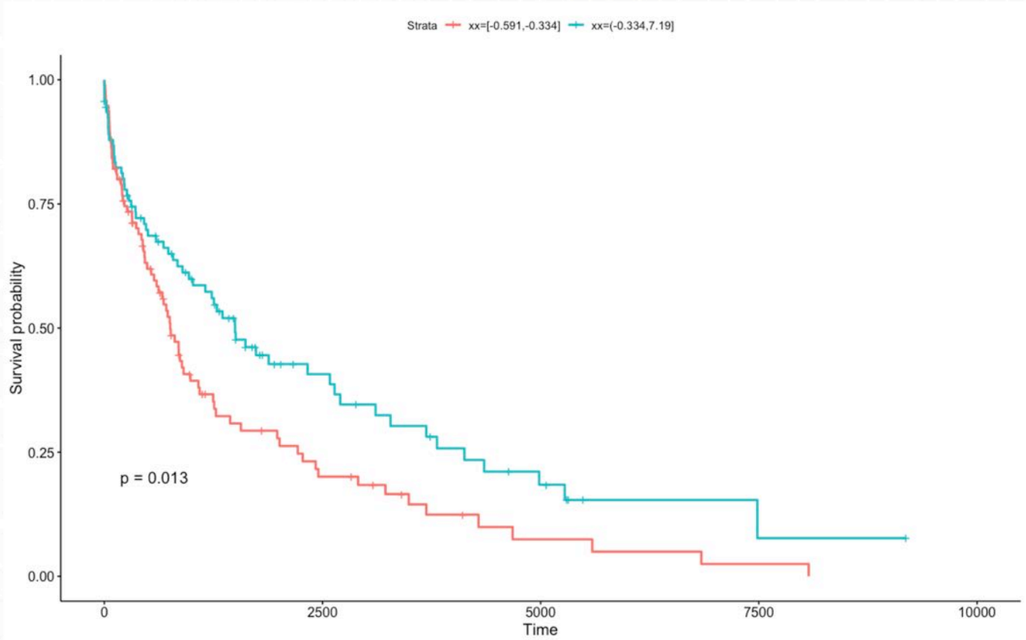
TCGA-EE-A2GL, Malignant Melanoma

Exhaustive analysis of cellular features in TCGA to enable data-driven identification of pathological predictors of survival in malignant melanoma

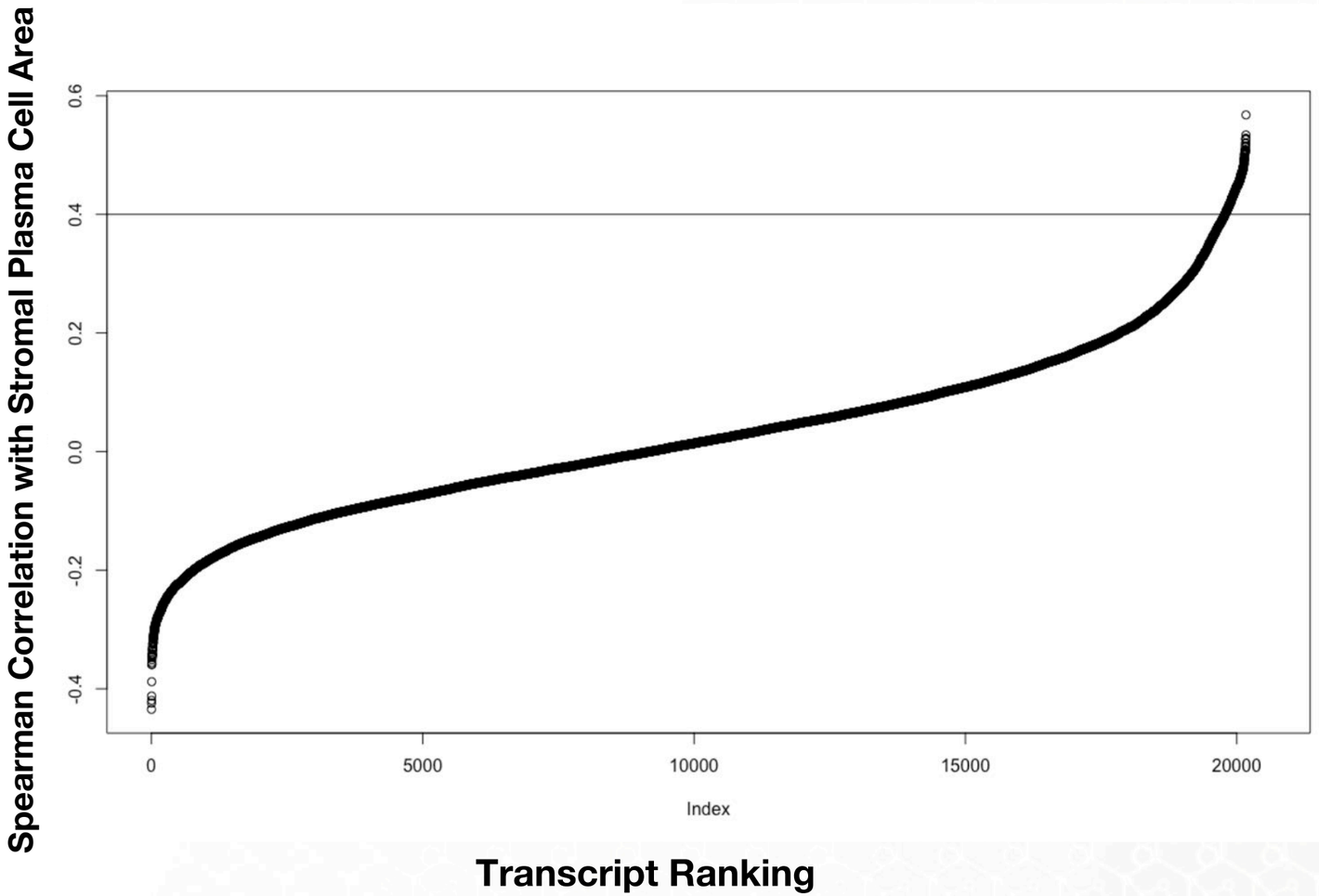
Pathological phenotypes with FDR < 5% for association with Progression Free Survival



Increased area of stromal plasma cells associated with improved survival in melanoma



Data-driven identification of transcriptional signature underlying stromal area of plasma cells in melanoma



Top-ranking transcripts associated with stromal area of plasma cells

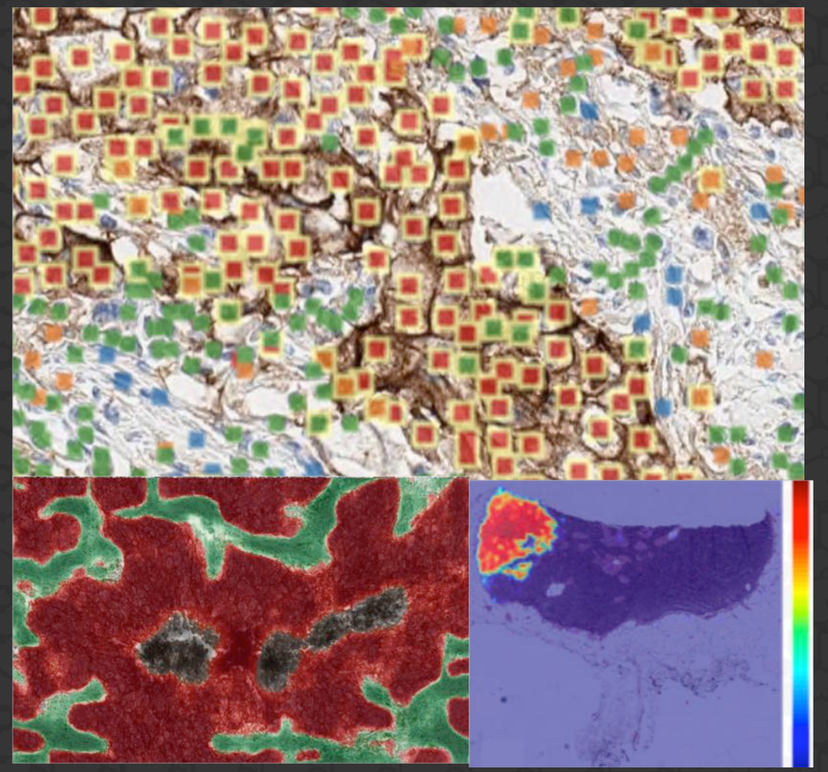
Gene	Correlation
REC8	0.57
GPR174	0.53
CD38	0.53
LAX1	0.53
TOX	0.53
AKAP5	0.53
C8orf80	0.52
JSRP1	0.52
IGJ	0.52
TNFRSF17	0.51
EAF2	0.51

Stromal plasma cell area RNA signature strongly enriched for immune genes

Gene Set Name	Description	FDR q-value
REACTOME_IMMUNE_SYSTEM	Genes involved in Immune System	7.62E-57
REACTOME_ADAPTIVE_IMMUNE_SYSTEM	Genes involved in Adaptive Immune System	6.02E-42
PID_TCR_PATHWAY	TCR signaling in naive CD4+ T cells	4.24E-30
REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON_LYMPHOID_CELL	Genes involved in Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	6.07E-26
KEGG_PRIMARY_IMMUNODEFICIENCY	Primary immunodeficiency	7.98E-24
PID_IL12_2PATHWAY	IL12-mediated signaling events	9.27E-24
PID_CD8_TCR_PATHWAY	TCR signaling in naive CD8+ T cells	9.27E-24
KEGG_CELL_ADHESION_MOLECULES_CAMS	Cell adhesion molecules (CAMs)	3.00E-22
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	Cytokine-cytokine receptor interaction	6.38E-22
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION	Intestinal immune network for IgA production	3.37E-21
REACTOME_TCR_SIGNALING	Genes involved in TCR signaling	3.24E-20
REACTOME_PD1_SIGNALING	Genes involved in PD-1 signaling	3.44E-19
REACTOME_COSTIMULATION_BY_THE_CD28_FAMILY	Genes involved in Costimulation by the CD28 family	5.48E-19

Another AI plus: scalability

- Same pipeline for any solid tumor type
  - Contrast to traditional approach: hand-crafted algorithms.



PathAI for Immuno-oncology

PathAI platform has been applied to:

- Non-small cell lung cancer (Adenocarcinoma)
- Non-small cell lung cancer (Squamous Cell Carcinoma)
- Small cell carcinoma of the lung
- Urothelial carcinoma of the bladder
- Head and neck squamous cell carcinoma
- Melanoma
- Breast cancer
- Prostate cancer
- Colon cancer

>30 IO-IHC biomarkers studied

IHC images processed

10,000+

Number of Annotations

2.5 Million+

PDL1 IHC cells classified

1 Billion+

In 2018, PathAI classified ~15x the number of cells that all US pathologists could perform in a year

PathAI

Extensive Slide Search & Data Standardization

Slides Search

Filter Images

Choose criteria

TCGA

TCGA

Any case

Any stain

Any group

Original file name

Overlays:

Yes No Either

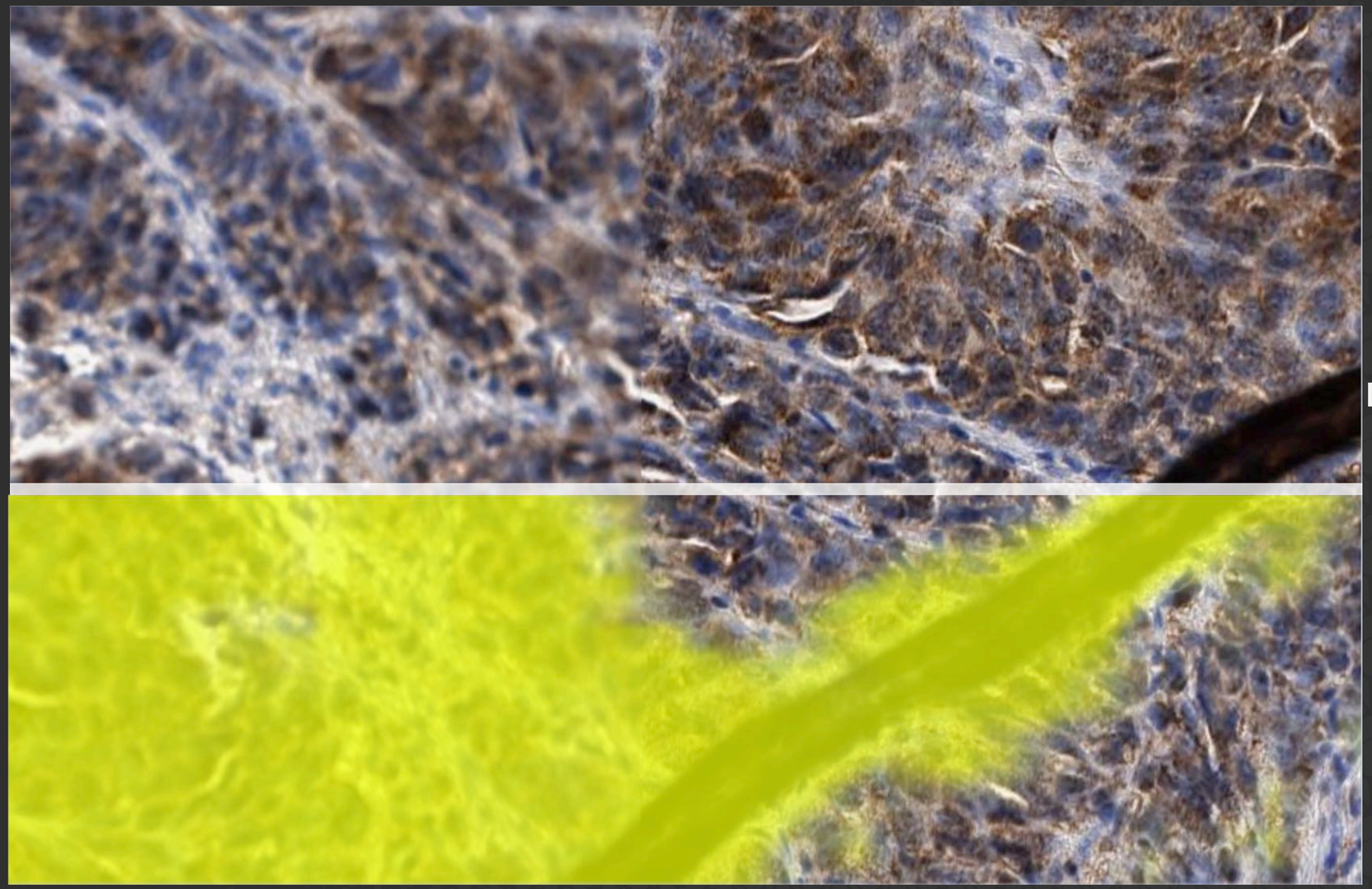
Annotations:

Yes No Either

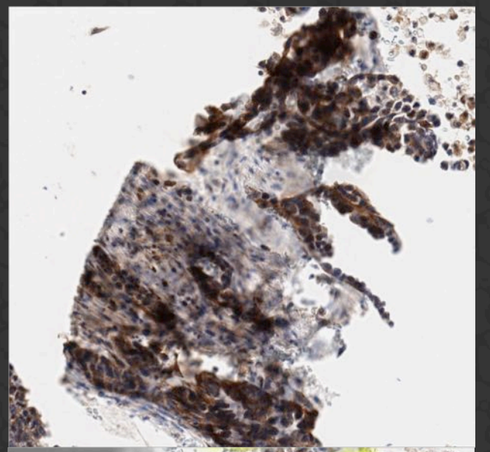
30872 matching images Clear filters

Automated quality control

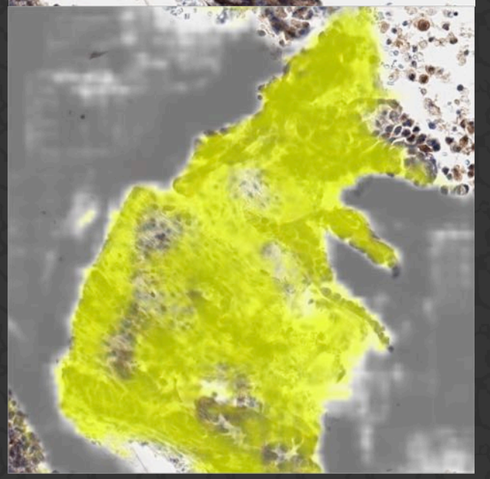
Blurred areas



Folded / damaged tissue

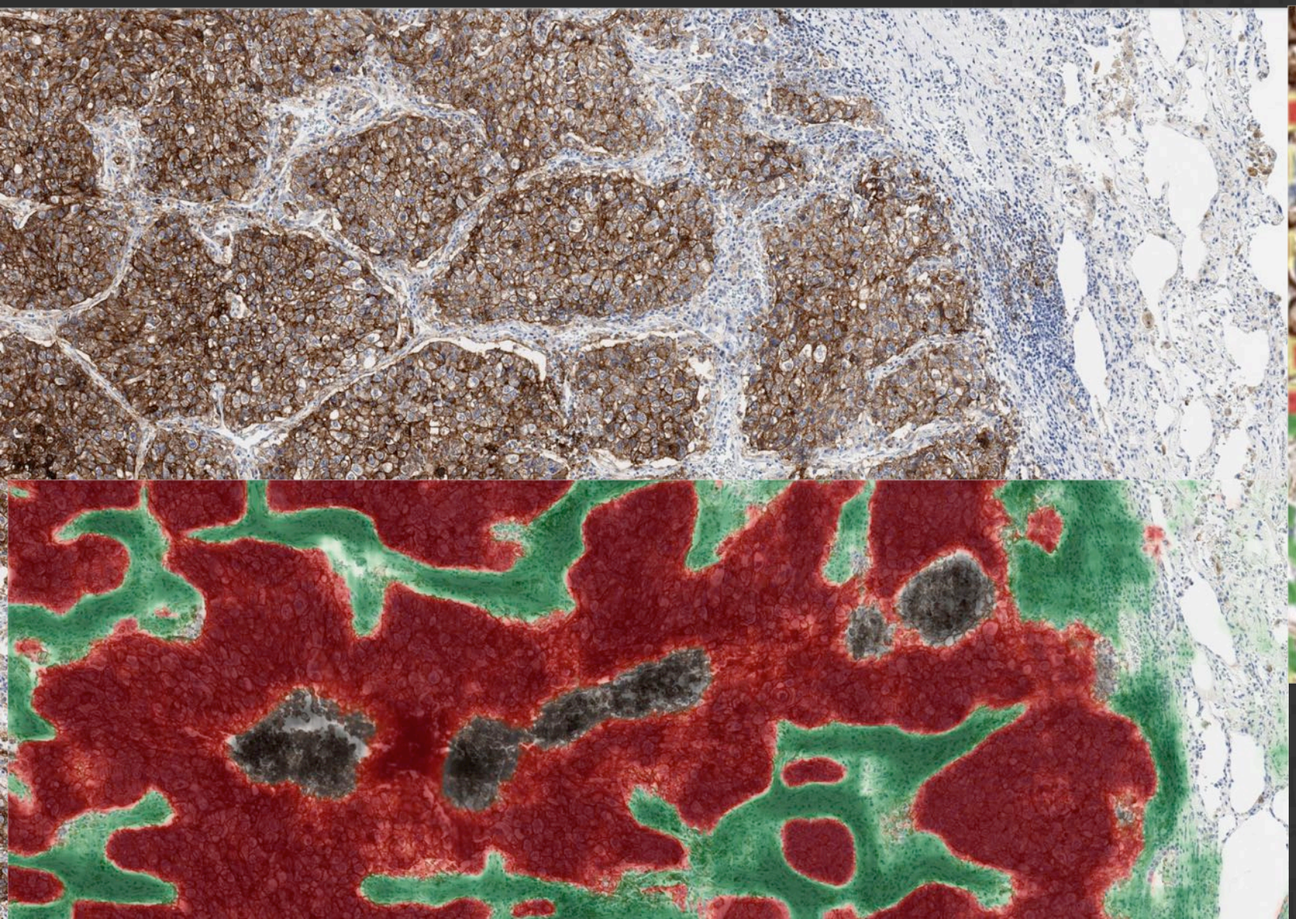


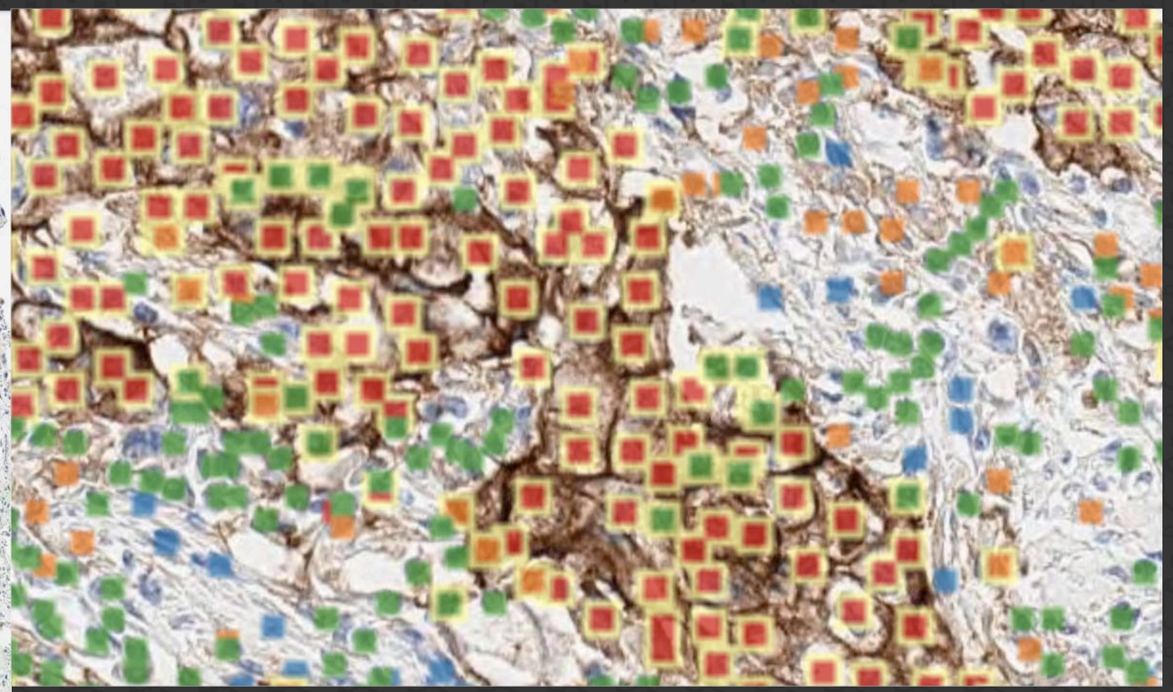
Debris



Annotate, train and deploy task-specific models

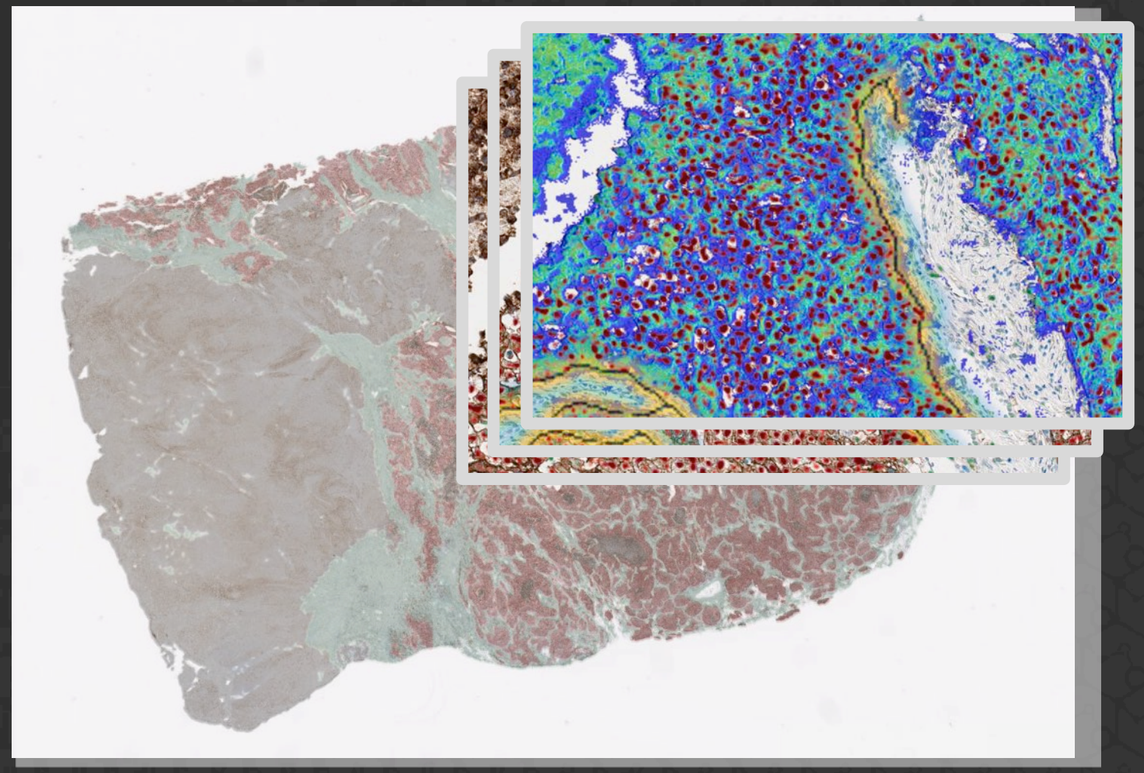
- Determined by partner needs



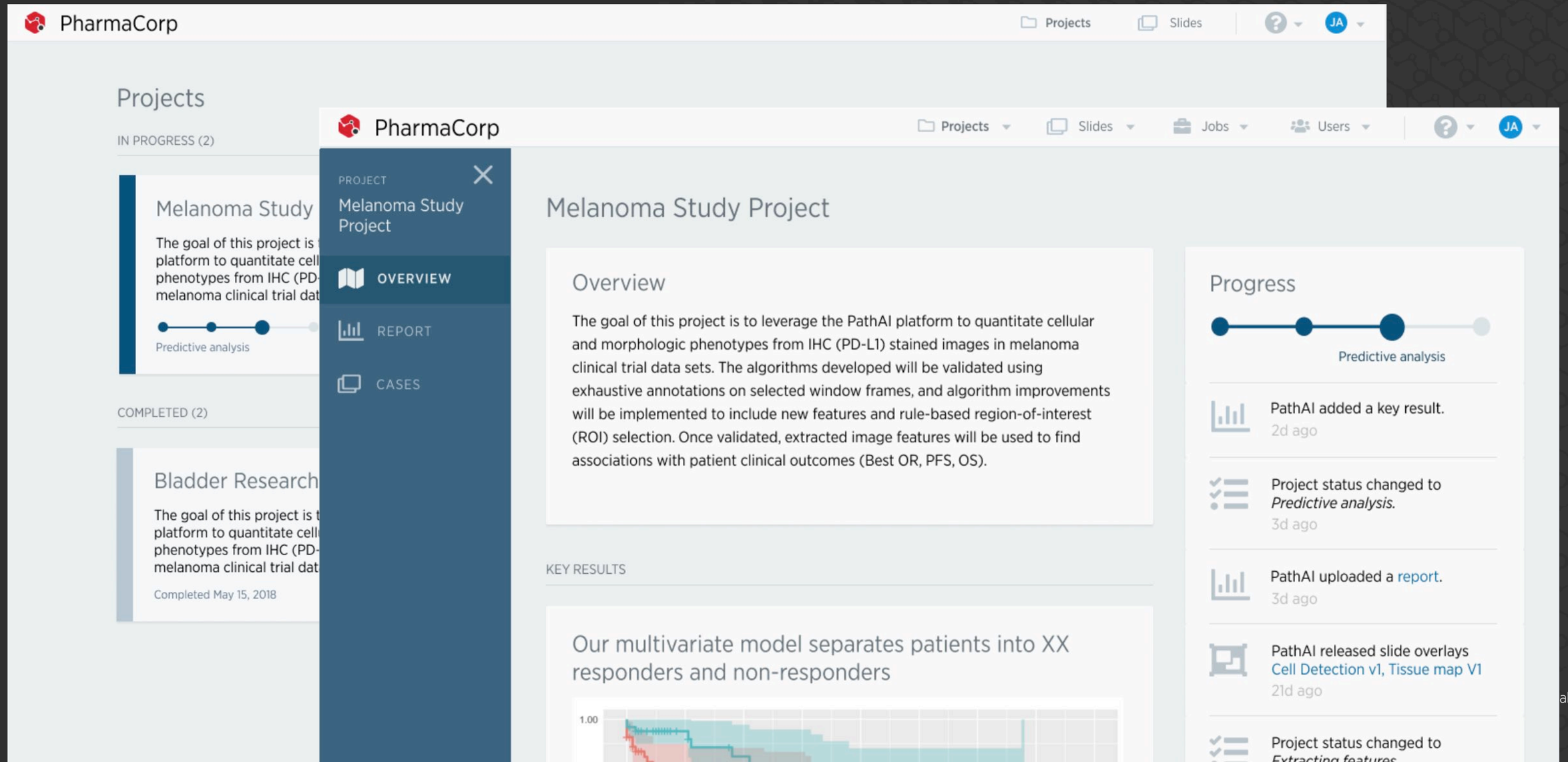


Interpretable feature extraction

- Hypothesis & data driven



Interactive Reports & Live Project Progress



The PathAI Deep Learning Process

Whole-Slide Images + Data

Transmit training data securely to the PathAI cloud

Annotations

Network of board-certified pathologists to provide ground truth consensus

Deep Learning Analysis

Cell detection, tissue & region classification

Deep Learning Feature Analysis

Over 200 relevant features extracted, measured and analyzed

Assay Validated

Identified features of significance reduced to practice

Assay Deployed

Analyze samples, quantified & visual results delivered

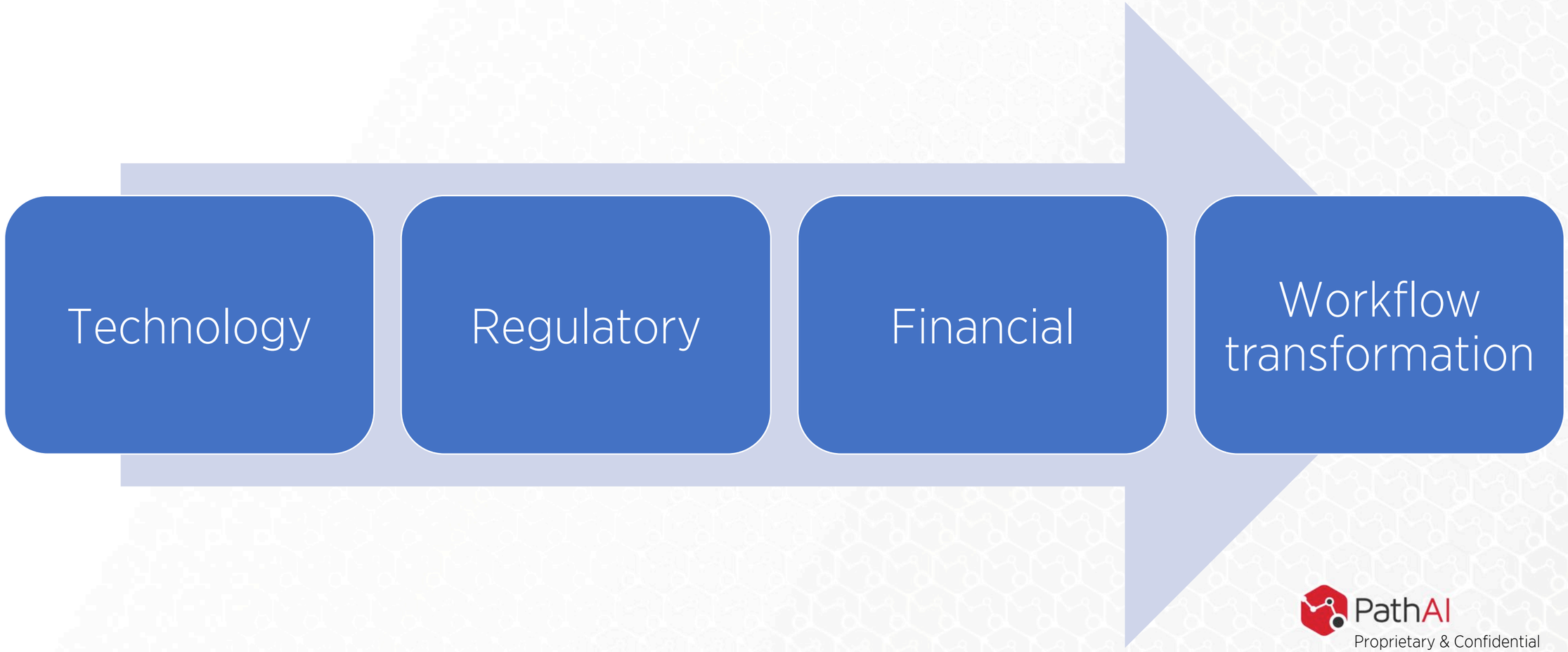
We can execute process in 4 – 8 weeks for new assays

AI in medicine

Some closing thoughts

- ML in the real world:
  - Building the right dataset is 75% of the challenge
- Modern ML: engineering and empirical science
  - Rigorous validation is key
- Ideas and algorithms vs. teams and infrastructure

Core challenges and road ahead



Key Takeaways

- Researchers have been working on AI for pathology for ~30 years
- In the past 5 years, advances in:
  - Availability of digital data
  - Access to large-scale computing resources
  - Major algorithmic advances (e.g., Deep CNNs)
- AI works extremely well when these 3 factors are all available and fails when they are not

Key Takeaways

- AI-powered pathology is broadly applicable across all image-based tasks in pathology and enables integration with other structured data types (e.g., ‘Omics)
- As AI and digital pathology are incorporated into clinical workflow, they will offer significant operational and efficiency advantages
- AI will drive improvements in the accuracy and predictiveness of pathology leading to research advances and improved care for patients

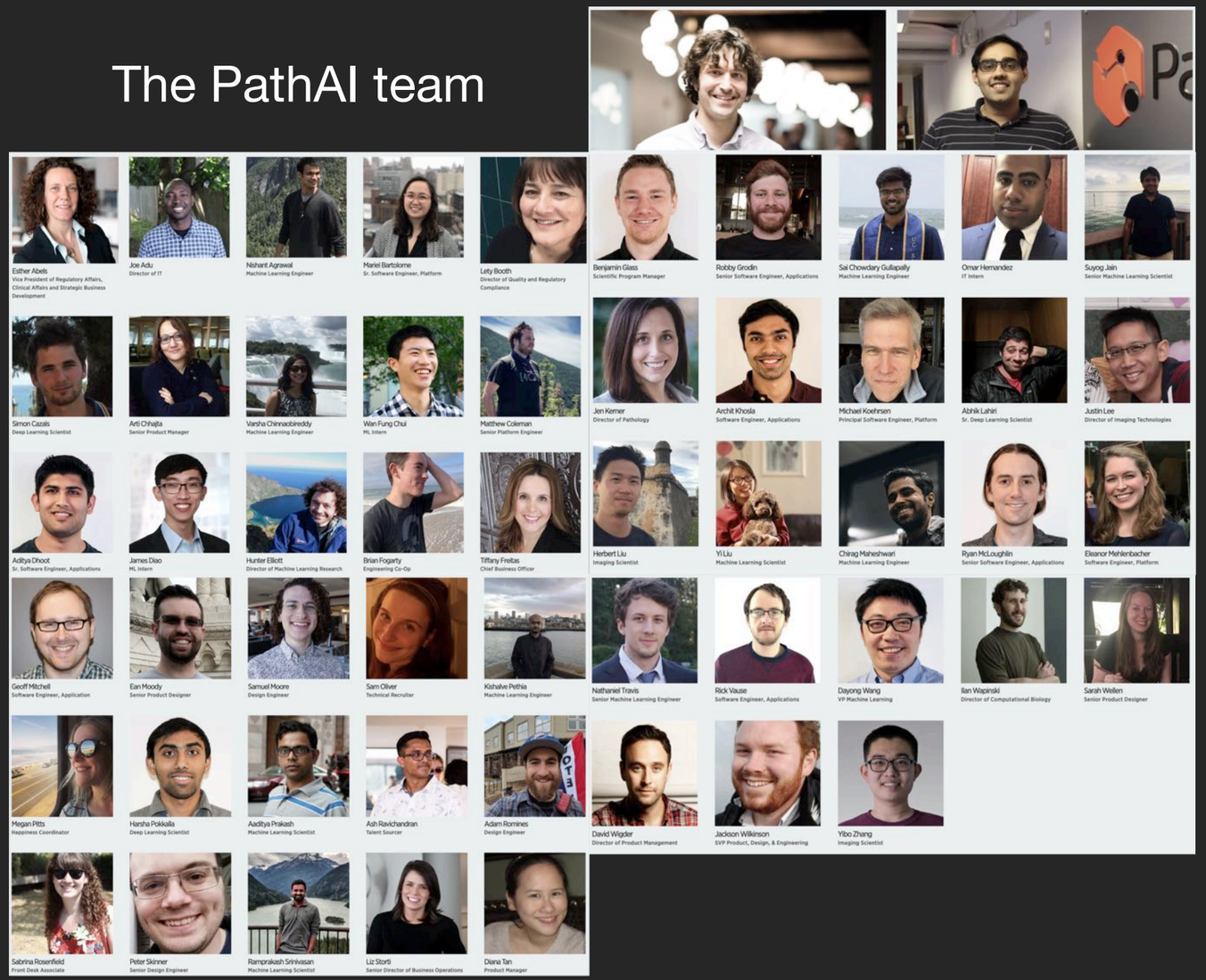
“In the Future…” (1987)

- “Integrated information systems, patient care management by exception, decision support tools, and, in the future, "artificial intelligence" assists can all be expected to become staples of pathology practice, especially impacting those pathologists who choose to be responsive to the new practice milieu of medical information science.”

“Using the computer to optimize human performance in health care delivery. The pathologist as medical information specialist.”
(Arch Pathol Lab Med. 1987)

Thank you!

The PathAI team



Hunter Elliott

Director, Machine Learning Research

pathai.com
Opportunities for ML engineers/scientists, software engineers, pathologists,…