AI-based quantitation of cancer cell and fibroblast nuclear morphology reflects transcriptomic heterogeneity and correlates with survival in breast cancer

Poster # P4-09-08

John Abel[1], Christian Kirkup[1], Juhyun Kim[1], Filip Kos[1], Ylaine Gerardin[1], Sandhya Srinivasan[1], Jacqueline Brosnan-Cashman[1], Ken Leidal[1], Sanjana Vasudevan[1], Deepta Rajan[1], Suyog Jain[1], Aaditya Prakash[1], Harshith Padigela[1], Jake Conway[1], Neel Patel[1], Ben Trotter[1], Limin Yu[1], Amaro Taylor-Weiner[1], Emma Krause[1], Matthew Bronnimann[1], Laura Chambre[1], Ben Glass[1], Chintan Parmar[1], Stephanie Hennek[1], Archit Khosla[1], Murray Resnick[1], Andy Beck[1], Fedaa Najdawi[1], Michael G. Drage[1], Ilan Wapinski[1]

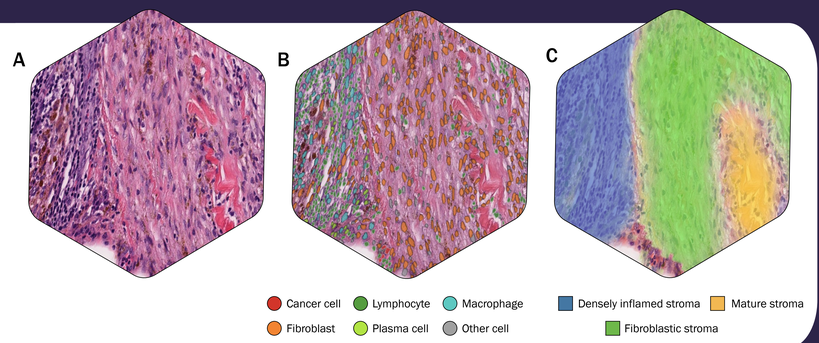
[1]PathAI, Boston, MA.

STUDY BACKGROUND

- Morphological features of cancer cell nuclei are routinely used to assess disease severity and prognosis, and prior work has linked cancer nuclear morphology to genomic alterations[1-3]. In addition, the cellular composition of cancer-associated stroma (CAS) has been linked to prognosis in several cancer types, including breast cancer[4].
- Quantitative analyses of 1) nuclear features of cancer cells and other tumor-resident cell types, such as cancer-associated fibroblasts (CAFs), and 2) composition of CAS may reveal novel biomarkers for prognosis and treatment response.
- Here, we applied a nucleus detection and segmentation algorithm, a cell classification model, and a stromal subdivision model to hematoxylin and eosin (H&E)-stained whole slide images (WSIs) of breast cancer specimens, enabling the assessment of features related to nuclear morphology and stromal composition.

PATHAI MODELS

Figure 3. Examples of model performance in breast cancer. A) H&E, B) nuclear segmentation and cell identification, and C) stromal subtype identification.



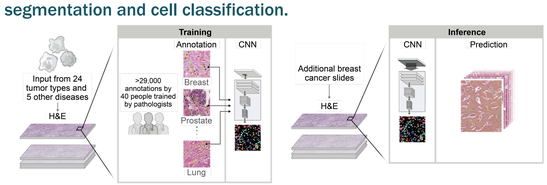
CONCLUSIONS

- The nuclear morphologies of breast cancer cells and CAFs reflect underlying genomic and transcriptomic properties of the tumor and associate with patient outcome.
- The histological appearance of the stroma differs between molecular subtypes of breast cancer, and features associated with macrophages were associated with outcome in patients with HER2+ disease.
- The application of digital pathology analysis of breast cancer histopathology slides enables the integrative study of genomics, transcriptomics, tumor morphology, and overall survival to support research into disease biology research and biomarker discovery.

METHODS

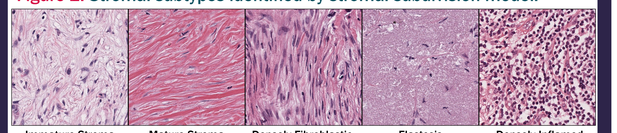
Model Training and Deployment. Convolutional Neural Network models for 1) nucleus detection and segmentation and 2) cell classification (Figure 1) were deployed on H&E-stained WSIs from The Cancer Genome Atlas (TCGA) breast cancer dataset (primary surgical resections; N=890). Separate models were trained to segment regions of stromal subtypes, such as inflamed and fibroblastic stroma (Figure 2). Examples of model output are shown in Figure 3.

Figure 1. Machine learning model training and deployment for nuclear segmentation and cell classification.



Exploratory analyses. Nuclear features (area, axis length, eccentricity, color, and texture) were computed and aggregated across each slide to summarize slide-level nuclear morphology for each cell type[5]. Next-generation sequencing-based metrics of genomic instability (N=774) and gene expression (N=868) were acquired for each case. Gene set enrichment analysis was performed using the Molecular Signatures Database[6]. Spearman correlation was used to relate nuclear features to genomic instability metrics, and linear regression was used to assess the relationship between nuclear features and bulk gene expression. Stromal features showing differences across breast cancer subtypes were identified by Kruskal-Wallis analysis, followed by Mann-Whitney U test between pairs of breast cancer subtypes. Multivariable Cox regression with age and ordinal tumor stage as covariates was used to find association between overall survival (OS) and nuclear features and stromal features. All reported results were adjusted for false discovery rate via the Benjamini-Hochberg procedure.

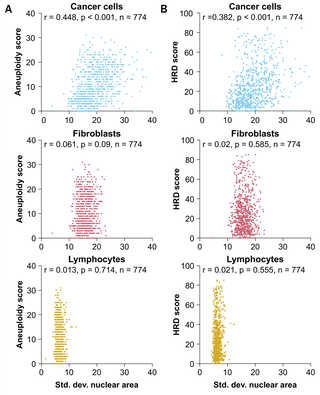
Figure 2. Stromal subtypes identified by stromal subdivision model.



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

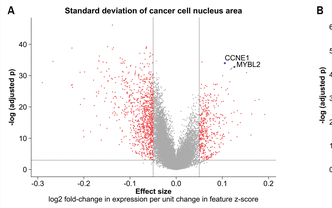
RESULTS

Figure 4. Genomic instability is associated with variability in nuclear area.



- Variation in cancer cell nuclear area, a quantitative metric related to pathologist-assessed nuclear pleomorphism, was calculated by the standard deviation of the nuclear area of cancer cells across a WSI.
- Standard deviation of nuclear area was associated with two metrics of genomic instability, A) aneuploidy score (r=0.448) and B) homologous recombination deficiency (HRD) score (r=0.382).
- In contrast, the variability in CAF and lymphocyte nuclear areas did not correlate with either metric of genomic instability (all r<0.1, p>0.05).

Figure 5. Nuclear morphology is related to gene expression.

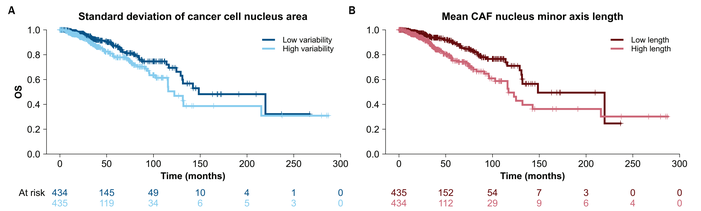


Reactome Gene Set Name	p-value	FDR q-value
Cell Cycle	1.13×10^{-32}	2.15×10^{-29}
Cell Cycle, Mitotic	3.13×10^{-31}	2.98×10^{-28}
FCGR Activation	1.1×10^{-28}	6.99×10^{-26}
Creation of C4 and C2 Activators	2.38×10^{-28}	1.13×10^{-25}
Initial Triggering of Complement	5.62×10^{-27}	2.14×10^{-24}

Reactome Gene Set Name	p-value	FDR q-value
Collagen Formation	6.32×10^{-18}	1.2×10^{-14}
Collagen Biosynthesis and Modifying Enzymes	1.4×10^{-17}	1.34×10^{-14}
Extracellular Matrix Organization	3.03×10^{-14}	1.93×10^{-11}
Assembly of Collagen Fibrils and other Multimeric Structures	1.81×10^{-10}	8.62×10^{-8}
Degradation of the Extracellular Matrix	6.34×10^{-8}	2.42×10^{-5}

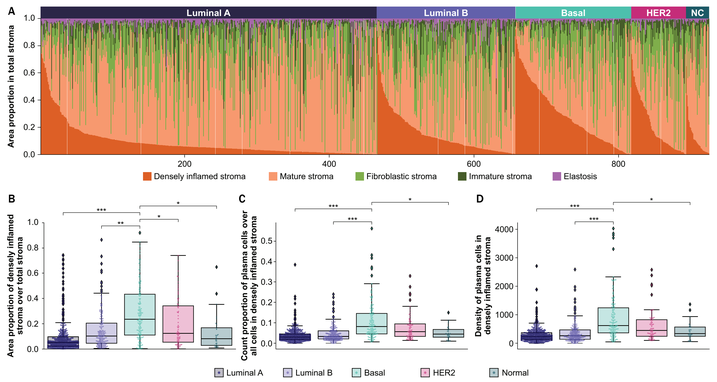
- Nuclear HIFs were compared to RNAseq-derived gene expression signatures for A) cancer cells and B) CAFs.
- Gene set enrichment analysis revealed associations between A) variation in cancer cell nuclear area and the expression of cell cycle and proliferation pathway genes and B) CAF nuclear size and the expression of gene sets involved in extracellular matrix remodeling and degradation. Genes of interest representing cell cycle/proliferation (CCNE1, MYBL2) and collagen regulation (RCN3[7]) are highlighted; the top five significant pathways are shown.

Figure 6. Nuclear morphology associates with overall survival in breast cancer.



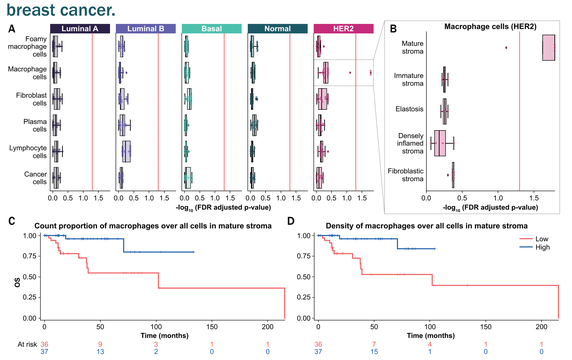
- Overall survival in breast cancer based on nuclear features was assessed by multivariable Cox regression with age and ordinal tumor stage as covariates.
- Stratifying breast cancer by low vs. high cancer cell nuclear area variability reveals that patients with low variability have improved OS compared to patients with high variability (p=0.05; A).
- Stratifying breast cancer by low vs. high CAF nuclear area reveals that patients with smaller CAF nuclei have improved OS compared to patients with larger CAF nuclei (p=0.002; B).

Figure 7. Molecular subtypes of breast cancer show unique stromal compositions.



- The relative area proportions of densely inflamed stroma, mature stroma, fibroblastic stroma, immature stroma, and elastosis were assessed in the four molecular subtypes of breast cancer, as well as non-cancerous tissue (NC) (A).
- Patients with basal-like breast cancer show larger areas of densely inflamed stroma, compared to patients with other subtypes of breast cancer (B).
- Patients with basal-like breast cancer also show higher number (C) and density (D) of plasma cells in densely inflamed stroma, compared to patients with other breast cancer subtypes.
- Significance was calculated by Mann-Whitney U test: ***: p<1e-10, **: p<1e-5, *: p<0.05

Figure 8. Macrophage-related features associate with OS in HER2+ breast cancer.



- Examination of stromal features revealed an association between macrophages and OS, specifically in HER2+ breast cancer (A). Further exploration revealed that this phenomenon was driven by macrophages in mature stroma (B).
- In HER2+ breast cancer, elevated count proportion (C) and density of macrophages in mature stroma (D) corresponded to improved OS (p=0.016 and 0.016, respectively).

REFERENCES

1. Zink, D., et al. *Nat Rev Cancer*. 2004; 4:677–687.
2. Fischer, E.G. *Acta Cytol*. 2020; 64:511–519.
3. Chow, K.-H., et al. *Nat Rev Cancer*. 2012; 12:196–209.
4. Wu, S.Z., et al. *The EMBO J*. 2020; 39:e104063.
5. Abel, J., et al. *Cancer Res*. 2022; 82(12_Supplement):464.
6. Liberzon, A., et al. *Bioinformatics*. 2011; 27:1739-1740.
7. Martinez-Martinez, E., et al. *Sci Rep*. 2017; 7:12192.

ACKNOWLEDGMENTS

We thank Bioscience Communications for assistance with figure design.

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

PathAI

San Antonio Breast Cancer Symposium, December 6-10, 2022, San Antonio, TX, USA