AI Powered-platform to Predict Gene Modifications from Prostate and Breast Cancer Whole Slide Images

Wei Huang[1,2], Parag Jain[2], Chensu Xie[2], Hassan Muhammad[2], Hirak Basu[2,3], George Wilding[2] and Rajat Roy[2]

[1] University of Wisconsin School of Medicine and Public Health, Madison, WI; [2] PathomIQ, Cupertino, CA; [3] MD Anderson Cancer Center, Houston, TX

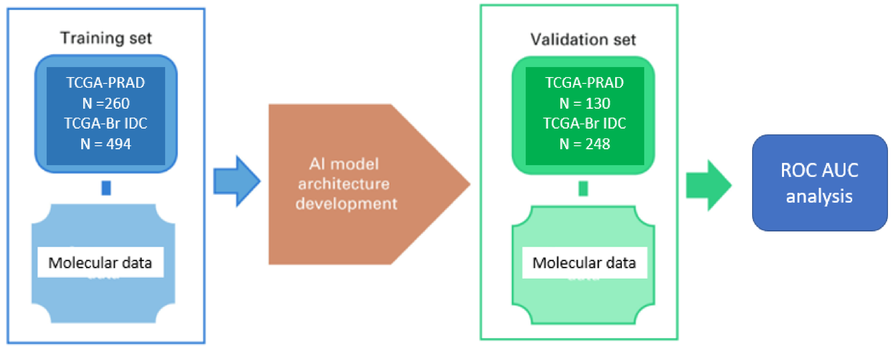
INTRODUCTION

- Data have shown that targeted therapies early in the Breast (BrCa) and prostate cancer (PCa) disease course improve outcomes in many patients.
- It has paramount importance to be able to identify patients who would benefit from targeted therapy in a timely fashion.
- Features at the cellular and molecular levels in BrCa and PCa pathogenesis and progression in each patient are divergent and complex, which complicates the timing and sequence determination for drug selection strategies to dynamically combat resistance mechanisms in the progressing BrCa and PCa.
- Molecular testing or sequencing are commonly used to determine the tumor's underlying molecular composition for patient stratification, which is time-consuming and costly process.
- It is well known that the complex molecular alterations in cancer progression are often reflected in subtle morphologic variations.
- There are a few early successes predicting microsatellite instability (MSI) status in colorectal (CRC) and gastric cancer (GC), FGFR3 mutation in bladder cancer, MYOD1 and PAX3/7-FOXO1 in rhabdomyosarcoma, etc. directly from hematoxylin and eosin (H&E) stained whole slide images (WSI)
- There is little published evidence of predicting cancer genotypes directly from H&E stained WSI, especially for BrCa and PCa.
- We have built a morphology-based and AI-powered platform able to extract and discriminate morphological features at visual and subvisual levels for cancer genotyping, risk stratification and outcome prediction with high accuracy that addresses the needs for optimal treatment decision-making in a cost effective and timely fashion.

EXPERIMENT DESIGN

A cohort of 390 prostate cancer and 742 invasive breast cancer patients with known molecular status of key genes, such as TP53, PIK3CA, MYC, ERBB2 (her2), TMPRSS2-ERG fusion and PTEN from The Cancer Genome Atlas (TCGA) were included in this study. H&E stained WSI of the patients' cancer tissue sections were available at 20x and 40x magnifications. The WSI from the PCa and BIDC cohorts were split 2:1 into a training and test cohort, respectively.

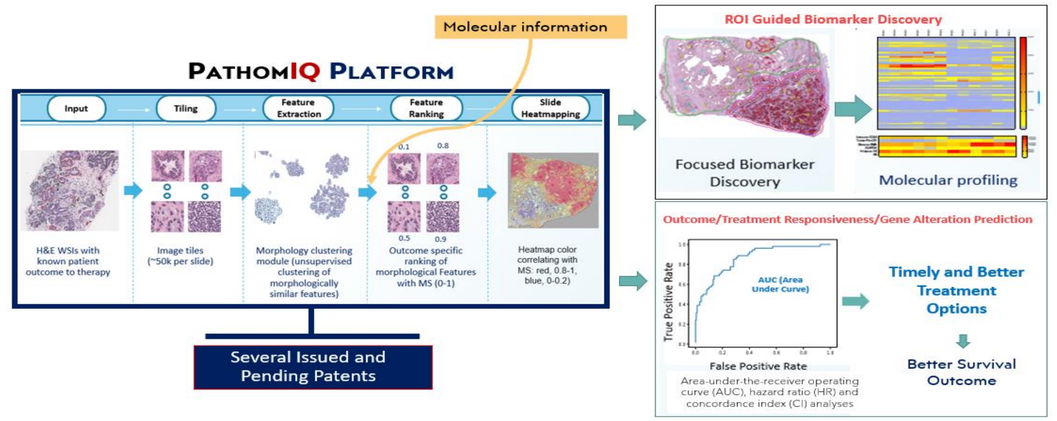
Figure 1. Study Workflow: Prediction of PCa and Breast IDC Genetic Alterations



EXPERIMENT DESIGN

Our platform technology involved two different deep Convolutional Neural Network (DCNN) architectures. The platform first divided each WSI into multiple tiles. Each tile was then analyzed using a DCNN that graded the tile and generated a high dimensional vector to provide a mathematical representation of the morphology based on the associated molecular phenotype. The combination of high dimensional vectors across the WSI was then fed into a second DCNN that generated a morphological score, which predicted whether the gene under consideration was wild type or underwent any modification (Figure 2).

Figure 2. Platform Workflow for Gene Alteration Prediction



RESULTS

Our platform has achieved 70 - 80% accuracy as defined by the Area under the Curve for the receiver operating characteristics curves for the genetic markers on the test cohorts (Table 1). Our platform can predict genotypes/molecular alterations directly from H&E stained WSI with high accuracy. (Table 1).

Table 1. Genotype Prediction for Breast Cancer (BrCa) and Prostate Cancer (PCa) from WSI

Cohort	Gene	Training Set (n)		Test Set (n)		ROC AUC Score
		Modification/Loss	Intact/Wild Type	Modification/Loss	Intact/Wild Type	
BrCa	TP53	196	302	96	149	80
	PIK3Ca	153	344	75	170	74
	MYC	68	414	33	204	77
	ERBB2	66	416	32	205	78
PCa	TMPRSS2-ERG	104	157	51	78	70
	PTEN	57	201	28	99	73

CONCLUSION

- Our platform can predict genotypes/molecular alterations directly from H&E stained WSI with high accuracy.
- This technology presents a novel, practical and cost-effective approach for cancer molecular classification and risk stratification, enabling timely and optimal treatment decision-making for positive clinical outcomes.

REFERENCES

1. Milewski D, Jung H, Brown GT, Liu Y, Somerville B, Lisle C, et al. Predicting Molecular Subtype and Survival of Rhabdomyosarcoma Patients Using Deep Learning of H&E Images: A Report from the Children's Oncology Group. Clin Cancer Res. 2023;29(2):364-78.
2. Loeffler CML, Ortiz Bruechle N, Jung M, Seillier L, Rose M, Laleh NG, et al. Artificial Intelligence-based Detection of FGFR3 Mutational Status Directly from Routine Histology in Bladder Cancer: A Possible Preselection for Molecular Testing? Eur Urol Focus. 2022;8(2):472-9.
3. Huang W, Randhawa R, Jain P, Hubbard S, Eickhoff J, Kummar S, et al. A Novel Artificial Intelligence-Powered Method for Prediction of Early Recurrence of Prostate Cancer After Prostatectomy and Cancer Drivers. JCO Clin Cancer Inform. 2022;6:e2100131.