Computer Vision and Machine Learning Allow for the Prediction of Genomic Instability Using Circulating Tumor Cell Morphology in Triple Negative Breast Cancer Patients

A. Jendrisak, J. Lee, A. Rodriguez, R. Sutton, A. Gill, J. Schonhoft, T. Pramparo, Y. Wang, R. Wenstrup

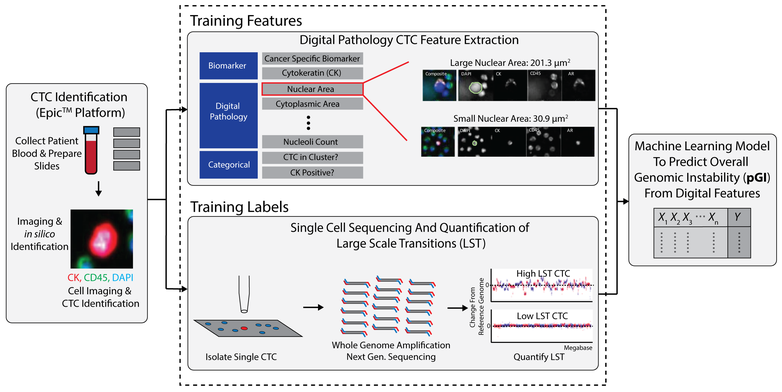
Epic Sciences, Inc., San Diego, CA. epicsciences.com

BACKGROUND

- Genomic scars and HRD gene mutations are biomarkers for PARP inhibitor (PARPi) and platinum agent therapy response in breast cancer.
- Tests for these markers are usually performed by next generation sequencing (NGS) of tumor tissue or circulating tumor DNA (ctDNA) and have limitations such as sample accessibility/availability and under-sampling due to inter-/intra-tumor heterogeneity.
- Previously we predicted GI using CTC phenotypic features without the use of NGS in metastatic castration resistant prostate cancer (mCRPC) pts with high accuracy (>76%) (ASCO 2016). In addition, patients with phenotypically predicted GI (pGI) CTCs had significantly better PSA responses on Abiraterone + Veliparib (92%) vs. Abiraterone alone (23%) (ESMO 2016), and also responded better to platinum agents vs. taxane agents (ESMO 2017).
- Here, we sought to develop and analytically validate an algorithm for predicting CTC pGI in TNBC

METHODS

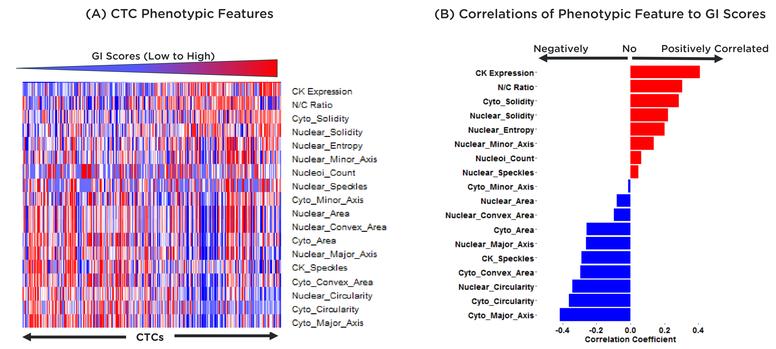
- Training set: 521 CTCs from 26 mCRPC pts were detected with the Epic Sciences CTC platform and analyzed for 19 phenotypic digital pathology features, including protein expression and cell morphology. The same CTCs were single-cell sequenced for the number of large scale transitions (LSTs) as an indicator of GI. A linear regression algorithm to predict GI (aka pLST) by CTC phenotype was developed, cross validated, and utilized to generate a CTC pGI score.
- Test set: 192 CTCs from 26 TNBC blood samples, median of 5 CTCs/pt, were sequenced for GI and phenotypically predicted for pGI.



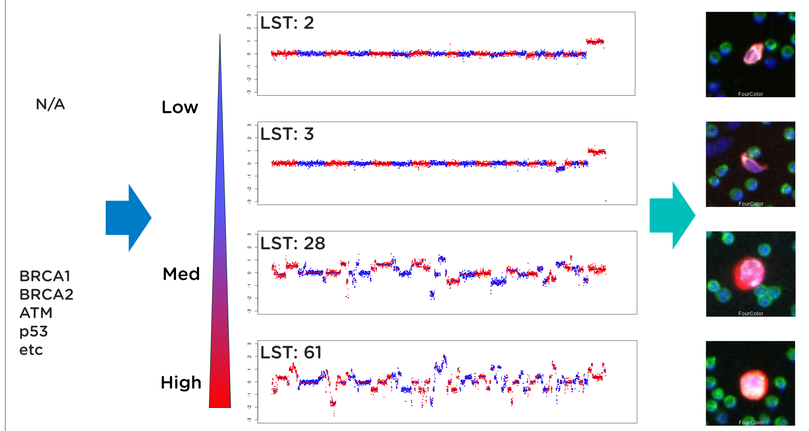
Schematic of Epic CTC platform for CTC identification, single cell sequencing, GI analyses workflow:

- Nucleated cells from patient's blood samples were deposited onto 10-12 glass slides containing ~3 million cells. Slides were IF-stained and scanned automatically at high speed to visualize cytokeratin (an epithelium marker), and CD45 (leukocyte exclusion marker), while DAPI was used as nuclear counterstain.
- CTC identification based on (DAPI+; CK+; CD45-) phenotype was achieved using a multi-parametric digital pathology algorithm. Digital pathology features were extracted.
- Relocated CTCs were individually isolated. Low-pass single cell NGS sequencing was used to determine the number of large scale transitions (LSTs) for each individual CTC genome within the training cohort and used as a ground truth or label for algorithm development.
- Using this set of features and labels, a linear regression model was trained to predict the continuous value per CTC. A cutoff was applied to obtain a binary read-out. Prediction performance was based on binary classification using the same cutoff for the continuous actual and predicted values.
- CTCs that were predicted to have greater than approximately 9 LSTs in total were designated pGI positive (i.e. predicted to be genomically unstable).

PHENOTYPE and GENOTYPE CORRELATION (mCRPC)

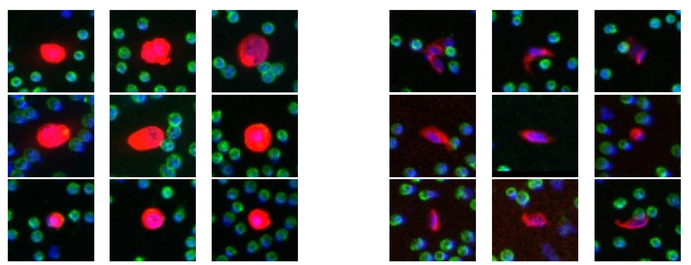


GENOTYPE DRIVING PHENOTYPE (TNBC)



EXAMPLES of TNBC CTCs with HIGH and LOW GI

- Representative images of CTCs with high genomic instability (high LSTs/pGI) and low genomic instability (low LSTs/pGI) and key features associated with the phenotypic classifier



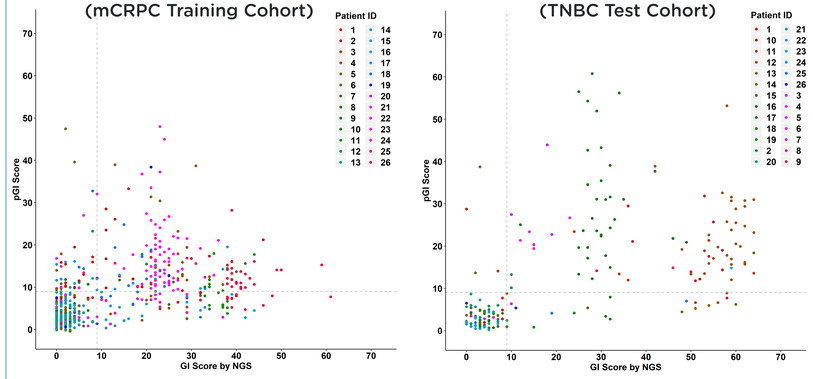
High Genomic Instability (high LSTs/pGI)

- High CK Expression
- High nuclear/cytoplasm ratio
- Low cytoplasmic major axis

Low Genomic Instability (low LSTs/pGI)

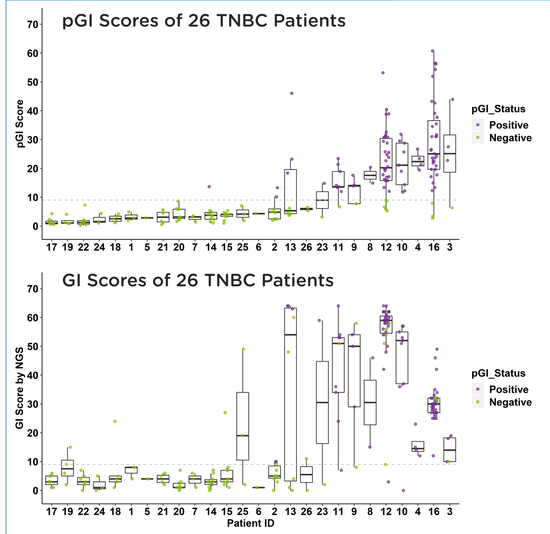
- Low CK Expression
- Low nuclear/cytoplasm ratio
- High cytoplasmic major axis

ANALYTICAL VALIDATION at THE CELL LEVEL



	Sample Size	Sensitivity	Specificity	Accuracy	PPV	NPV
Training - mCRPC cohort	521 CTCs (26 pts)	68%	83%	76%	77%	75%
Test - TNBC cohort	192 CTCs (26 pts)	84%	95%	88%	96%	81%

ANALYTICAL VALIDATION at SAMPLE LEVEL (TNBC)



- CTCs from 26 TNBC patients were sequenced for GI and analyzed for phenotypic GI.
- Patients were ranked from left to right by sample level median pGI, while GI scores followed the same trend.
- Heterogeneous pGI and GI scores were observed in multiple patients

CONCLUSIONS

- Previous studies showed that pGI was an analytically validated biomarker with clinical utility to predict PARPi or platinum therapy response in mCRPC pts.
- Here we show the same test concept can be applied to TNBC.
- Further analytical validation in a larger cohort is ongoing.
- The ability to identify PARPi/platinum sensitivity using an IF CTC staining method for CTC phenotype without the use of NGS will help to stratify patients more rapidly, at reduced cost and aid in the acceleration of drug development.