912P Metabolomic profiling for the early detection of lung cancer

P. Joubert[1], D. Wishart[2], J-F. Haince[3], H. Bach[4], R. Bux[5], P. Tappia[6], B. Ramjiawan[6]

[1]Pathology, Universite Laval, CHU de Quebec, Quebec, QC, Canada; [2]Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada; [3]BioMark-Quebec, BioMark Diagnostics Inc., Richmond, QC, Canada; [4]Division of Infectious Diseases, UBC - The University of British Columbia, Vancouver, BC, Canada; [5]BioMark- Headoffice, BioMark Diagnostics Inc, Richmond, BC, Canada; [6]Asper Clinical Research Institute, St. Boniface Hospital, Winnipeg, MB, Canada

Background: Currently, the five-year survival rate of lung cancer patients is very low, largely attributed to newly diagnosed patients presenting with locally advanced or metastatic disease. The lung cancer five-year survival rate (18.6%) is lower than many other leading cancer sites, such as colorectal (64.5%), breast (89.6%) and prostate (98.2%). The five-year survival rate for lung cancer is 56% for cases detected when the disease is still localized (within the lungs). However, only 16% of lung cancer cases are diagnosed at an early stage. For distant tumors (spread to other organs) the five-year survival rate is only 5%. More than 50% of lung cancer cases die within one year of being diagnosed. Accordingly, early diagnosis is key to the successful treatment, management and care of lung cancer.

Methods: Metabolomic techniques were used to discover and validate plasma biomarkers for the diagnosis of early-stage non-small cell lung cancer (NSCLC). Plasma samples from 599 patients with biopsy-confirmed NSCLC along with age and sex-matched plasma samples from 214 controls were analyzed. A fully quantitative targeted mass spectrometry analysis (targeting 142 metabolites) was performed. The sample set was split into a discovery set and validation set. Metabolite concentrations, clinical data, and smoking history were used to determine optimal sets of biomarkers and optimal regression models for identifying different stages of NSCLC using the discovery sets. The same biomarkers and regression models were used and assessed on the validation models.

Results: Univariate and multivariate statistical analysis identified β-hydroxybutyric acid, LysoPC 20:3, PC ae C40:6, citric acid, and fumaric acid as being significantly different between healthy controls and early stage (I/II) NSCLC. Predictive models with AUC $>$ 0.8 were developed and validated using these metabolites and other clinical data for detecting different stages of NSCLC.

Conclusions: This study has identified and validated a simple, high-performing, metabolite-based test for detecting early stage (I/II) NSCLC patients in plasma. Such an observation will enable a blood-based routine screening test for patients at the highest risk for lung cancer that is cost-effective, accurate and reliable.

Legal entity responsible for the study: BioMark Diagnostics Inc.

Funding: BioMark Diagnostics Inc.; Maunders-McNeill Foundation.

Disclosure: All authors have declared no conflicts of interest.

https://doi.org/10.1016/j.annonc.2022.07.1038

914P Early cancer detection using circulating extracellular vesicles

R. Kurzrock[1], P. Billings[2], H. Balcer[3], J. Hinestrosa[3]

[1]Medicine, Medical College of Wisconsin, Milwaukee, WI, USA; [2]Biological Dynamics, San Diego,CA, USA; [3]Research, Biological Dynamics, Inc., San Diego, CA, USA

Background: For many patients, early detection of cancer strongly correlates with better outcomes. Unfortunately, the cancers with the poorest 5-year survival statistics are those that are most often found at late stages. Correlation between outcomes and detection may justify broad cancer screening if innovative approaches provide compelling results with acceptable performance for general screening or early detection purposes. Several diagnostic companies test blood products to identify markers for disease, e.g. liquid biopsy. Many of these sequence circulating tumor DNA, which may not be detectable in early-stage patients. To overcome this and other limitations of competitive approaches, we have developed a novel platform for disease detection using extracellular vesicle (EV) proteins and machine learning (ML) algorithms.

Methods: Using minimal patient or control volume (240 µL), we isolate EVs from the plasma and probe them for surface and intravesical protein cargoes. The platform used for EV isolation enriches particles (size 50 to 250 nm) with minimal contamination in contrast with other EV isolation technologies. Retrospective samples from patients with stage I or II cancer across sixteen different cancer types, and non-cancer controls (including benign conditions) were evaluated in a pilot early detection study and a ML classifier was developed using EV biomarkers.

Results: The most informative biomarkers were found by analyzing each cancer type individually in comparison to the controls. Using this information, we then employ a suite of machine learning algorithms to create a multi-cancer early-detection (MCED) classifier from the most informative biomarkers and phenotypic/clinical information obtaining an AUC greater than 0.95.

Conclusions: EVs isolated by our method and analyzed by multiomics may be a powerful tool for early cancer detection. Multiple challenges remain for the successful implementation of platforms like ours in screening or early detection of disease. These include partnering with academic clinics to address patient heterogeneity; identify elevated risks for specific diseases (germline mutations, precursor lesions), establish performance metrics and explore if our testing may improve pre- or post-surgical management with curative intent.

Legal entity responsible for the study: The authors.

Funding: Biological Dynamics.

Disclosure: P. Billings: Financial Interests, Institutional, Member of the Board of Directors: Biological Dynamics. H. Balcer, J.P. Hinestrosa: Financial Interests, Institutional, Full or part-time Employment: Biological Dynamics.

https://doi.org/10.1016/j.annonc.2022.07.1039

915P EMLI-ICC: An ensemble machine learning-based proteome and transcriptome integration algorithm for metastasis prediction and risk-stratification in intrahepatic cholangiocarcinoma

C. Ye, R. Chen, Q. Jiang, W. Wu, F. Yan, Q. Li, X. Shuaishuai, Y. Wang, Y. Jia, X. Zhang, P. Shen, J. Ruan

Medical Oncology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China,

Background: Intrahepatic cholangiocarcinoma (ICC) is the second-most-common primary liver cancer with increasing incidence and mortality worldwide. Although tumor metastasis is frequent even in ICC, improved robust strategies to identify patients at high risk for metastasis remains limited. Batch effects and different data types greatly decrease the predictive performances of signatures based on gene/protein expression profiles in inter-laboratory and different data type validation.

Methods: To address this problem and assist in more precise diagnosis, we performed a genome-wide integrative analysis of proteome and transcriptome and developed an ensemble machine learning-based integration algorithm for metastasis prediction and risk-stratification in ICC (EMLI-ICC).

Results: Based on massive proteome and transcriptome data sets, 186 feature (biomarker) genes were selected and used to train the EMLI-ICC algorithm.The metastasis prediction model based on the EMLI-ICC algorithm showed AUC 0.887 for determining the metastasis samples in proteome and transcriptome datasets. To test prediction accuracy of our EMLI-ICC algorithm, we evaluated two RNA-seq datasets (TCGA, and Ahn's data) and found AUC 0.923 and 0.883, respectively. We next analyzed 103 specimens from patients with ICC (metastasis, n = 55; non- metastasis, n = 48), followed by Cox proportional hazard regression analysis, to develop an integrated prognostic gene panel and establish a risk-stratification model for metastasis in ICC.

Conclusions: We report an ensemble machine learning-based integration algorithm for metastasis prediction and risk-stratification that is superior to currently used clinicopathological features in patients with ICC.

Legal entity responsible for the study: The authors.

Funding: 1.National Natural Science Foundation of China; 2.Natural Science Foundation of Zhejiang Province; 3. "Hundred Talents Plan (Clinical Medicine)" Foundation of Zhejiang University.

Disclosure: All authors have declared no conflicts of interest.

https://doi.org/10.1016/j.annonc.2022.07.1040

916P Deep learning-based multimodal ensemble algorithm for multi-cancer detection and classification using cf-WGS

J. Lee[1], T-R. Lee[2], G. Kim[3], J.M. Ahn[2], S.R. Park[4], K-B. Song[5], E. Jun[5], D. Oh[6], J-W. Lee[7], Y.S. Park[8], G-W. Song[9], J-S. Byeon[10], B.H. Kim[11], J. Lee[2], D. Kim[2], C-S. Ki[2], E. Cho[2], J.K. Choi[3]

[1]Department of Bioinformatics, Soongsil University, Seoul, Republic of Korea; [2]Genome Research Center, GC Genome, Yongin, Republic of Korea; [3]Department of Bio and Brain Engineering, KAIST, Daejeon, Republic of Korea; [4]Oncology Dept., Asan Medical Center - University of Ulsan College of Medicine, Seoul, Republic of Korea; [5]Department of Surgery, Asan Medical Center - University of Ulsan, Seoul, Republic of Korea; [6]Department of Radiation Oncology, Samsung Medical Center (SMC) - Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; [7]Department of Obstetrics and Gynecology, Samsung Medical Center (SMC) - Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; [8]Division of Hepatopancreatobiliary Surgery and Liver Transplantation, Seoul National University Hospital, Seoul, Republic of Korea; [9]Division of Hepatopancreatobiliary Surgery and Liver Transplantation, Asan Medical Center, Seoul, Republic of Korea; [10]Gastroenterology, Asan Medical Center - University of Ulsan College of Medicine, Seoul, Republic of Korea; [11]Center for Liver and Pancreatobiliary Cancer, National Cancer Center, Goyang, Republic of Korea

Background: Various cell-free DNA (cfDNA) features have been investigated for their potential use in early cancer detection and cancer type classification. In this study, we developed a multimodal ensemble algorithm which integrates two independent deep learning models in which one uses regional mutational density (RMD) and mutational