Analytica Chimica Acta xxx (2018) 1–9

Contents lists available at ScienceDirect

Analytica Chimica Acta

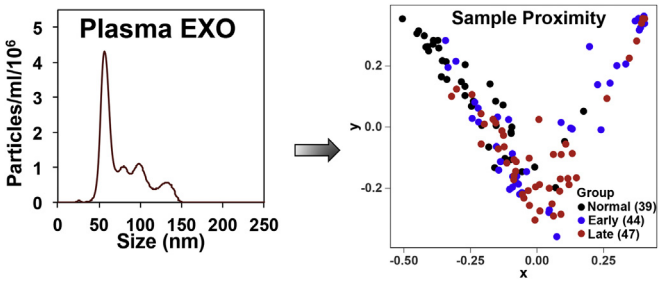
journal homepage: www.elsevier.com/locate/aca

Exosomal lipids for classifying early and late stage non-small cell lung cancer

Teresa W.M. Fan [a, *], Xiaofei Zhang [b], Chi Wang [c], Ye Yang [a, 1], Woo-Young Kang [a], Susanne Arnold [d], Richard M. Higashi [a], Jinze Liu [b], Andrew N. Lane [a, **]

[a] Center for Environmental and Systems Biochemistry (CESB), Markey Cancer Center, Department of Toxicology and Cancer Biology, University of Kentucky, United States
[b] Department of Computer Science and Markey Cancer Center, University of Kentucky, United States
[c] Department of Biostatistics and Markey Cancer Center, University of Kentucky, United States
[d] Markey Cancer Center, University of Kentucky, United States

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 1 November 2017
Received in revised form
16 February 2018
Accepted 19 February 2018
Available online xxx

Keywords:
Non-small cell lung cancer
Ultrahigh resolution Fourier transform mass spectrometry
Exosomal lipid profiling
Random forest
LASSO

ABSTRACT

Lung cancer is the leading cause of cancer deaths in the United States. Patients with early stage lung cancer have the best prognosis with surgical removal of the tumor, but the disease is often asymptomatic until advanced disease develops, and there are no effective blood-based screening methods for early detection of lung cancer in at-risk populations. We have explored the lipid profiles of blood plasma exosomes using ultra high-resolution Fourier transform mass spectrometry (UHR-FTMS) for early detection of the prevalent non-small cell lung cancers (NSCLC). Exosomes are nanovehicles released by various cells and tumor tissues to elicit important biofunctions such as immune modulation and tumor development. Plasma exosomal lipid profiles were acquired from 39 normal and 91 NSCLC subjects (44 early stage and 47 late stage). We have applied two multivariate statistical methods, Random Forest (RF) and Least Absolute Shrinkage and Selection Operator (LASSO) to classify the data. For the RF method, the Gini importance of the assigned lipids was calculated to select 16 lipids with top importance. Using the LASSO method, 7 features were selected based on a grouped LASSO penalty. The Area Under the Receiver Operating Characteristic curve for early and late stage cancer versus normal subjects using the selected lipid features was 0.85 and 0.88 for RF and 0.79 and 0.77 for LASSO, respectively. These results show the

Abbreviations: AUROC, area under the receiver operator characteristic curve; UHR-FTMS, ultra high-resolution Fourier transform mass spectrometry; LASSO, Least Absolute Shrinkage and Selection Operator; NSCLC, non-small cell lung cancer; PCA, Principal Component Analysis; OPLS-DA, Orthogonal partial least square discriminant analysis.

* Corresponding author. Rm 523 Lee T. Todd Jr. Building, 789 S. Limestone St., University of Kentucky, Lexington, KY 40536, United States.
** Corresponding author. Rm 525 Lee T. Todd Jr. Building, 789 S. Limestone St., University of Kentucky, Lexington, KY 40536, United States.
E-mail addresses: twmfan@gmail.com (T.W.M. Fan), andrew.lane@uky.edu (A.N. Lane).
[1] Present Address: Ye Yang, Urologic Oncology Branch, National Cancer Institute, Bethesda MD.

https://doi.org/10.1016/j.aca.2018.02.051

value of RF and LASSO for metabolomics data-based biomarker development, which provide robust an independent classifiers with sparse data sets. Application of LASSO and Random Forests identifies lipid features that successfully distinguish early stage lung cancer patient from healthy individuals.

1. Introduction

1.1. Lung cancer is difficult to diagnose at early stage

Lung cancer is by far the leading cause of cancer deaths in the U.S [1,2]. with an estimated 224,390 new cases and 158,080 deaths in 2016 [3]. Kentucky now leads the nation both in terms of lung cancer incidence and mortality, with the Appalachian population showing even higher incidence and mortality rates [4]. Most lung cancer patients are diagnosed at advanced stages due to the silent nature of the early stage disease. Although the five-year survival rate of localized lung cancer is ~55% with proper surgical intervention [5,6], that of advanced stage disease drops to ~4%. Presently, there is no robust low-cost blood-based screening method for detecting asymptomatic early stage lung cancer. Current imaging or cytology-based methods are impractical for screening at-risk populations for lung cancer, as they are not sufficiently accurate, cost-effective or non-invasive [7,8]. Although low dose helical CT screens have recently been reported to decrease lung cancer mortality by 20% in comparison to chest x-ray screening, there remains a high false positive rate [9]. Thus, techniques to detect and reliably screen lung cancer at its earliest stage in at-risk populations are urgently needed to improve survival and quality of life for lung cancer patients.

Non-small cell lung cancer (NSCLC) is the dominant form (ca. 85%) of lung cancer, and comprises many subtypes with different sets of oncogenic drivers such as mutant *KRAS, EGFR, LKB1, EML4-ALK (*adenocarcinomas*), PIK3CA, NRF2* (squamous cell carcinomas), *cMYC* overexpression and inactivation of *TP53* via mutations (both subtypes) [10–14], and numerous other genetic aberrations yet to be functionally defined. It is becoming clear that one of the key functions of these oncogenic drivers lies in reprogramming specific metabolic events in cancer cells to promote their proliferation, survival and metastasis. Thus, metabolic reprogramming in cancer has been recently recognized as a hallmark of cancer [15]. However, the global metabolic networks, and lipidomics in particular, modulated by these drivers and/or other undefined genetic aberrations are poorly characterized in NSCLC.

1.2. Lung cancer lipid metabolism is a rich ground for biomarker discovery

We have performed lipid profiling of paired CA and NC lung tissues using UHR-FTMS, “UHR” defined here as MS with sufficient resolving power to resolve the many hundreds of lipid species and their ^{13}C isotopologues (same lipid differing only in the number of ^{13}C atoms). A large fraction of our NSCLC tissue collections analyzed are classified as early stage adenocarcinoma (AdC) or squamous cell carcinoma (SqCC) [16]. We noticed consistent differences in the lipid profiles of paired CA versus NC lung tissues e.g. sphingomyelins (SM), ceramides, phosphatidylserines (PS), and cholesterol esters (Fan, Higashi & Lane unpublished data), which could reflect the altered expression of many lipid metabolic genes evident in lung [17] and other tumors [18].

1.3. Exosomes and microvesicles carry tumor cell-derived bioactive materials

Interestingly, both SM and PS have been linked to lipid microparticles (MP) shed from cells [19]. MP such as exosomes (EXO) and microvesicles (MV) can be shed from many different cell types, most notably immune cells and tumor cells, into the circulating blood. EXO are multivesicular bodies originating from the endosomal membrane, and are released upon fusion with the plasma membrane while MV are formed by outward budding and fission of the plasma membrane. Both types of lipidic MP are thought to mediate extracellular communications such as immune activation or suppression [19,20]. MP derived from cancer cells including lung cancer cells can carry a variety of bioactive proteins (e.g. epidermal growth factor receptor, EGFR; vascular endothelial growth factor, VEGF; integrins; Fas ligand; latent membrane protein, LMP-1; angiogenic factor tetraspanin; macrophage migration inhibitory factor or MIF) and microRNAs to promote tumor growth/invasion/metastasis as well as to enact immune evasion [19,21–31] and drug resistance [32–35]. Although largely unexplored, exosomal lipids derived from cancer cells have been shown to elicit apoptosis in sensitive cells via inhibition of the Notch-1 pathway [36] but activate the Akt survival pathway via promoting the NFκB-SDF1-CXCR4 axis in resistant cells [37]. Melanoma cells cultured under acidic conditions released EXO with a higher SM content, and were shown to have a higher capacity for cell fusion and delivery of caveolin-1 (tumor promoting) to less aggressive melanoma cells than neutral EXO [38]. Moreover, blocking CE buildup interferes with exosomal uptake [39] and has anti-cancer effects [18], while ceramide buildup is important for exosomal biogenesis [31] and triggers cancer cell death [40]. Thus, there are vital functions of lipids in exosomal biogenesis and interactions with the tumor microenvironment (TME) to influence tumor development and progression.

Recently, exosomal components such as microRNA and proteins have been shown to be promising diagnostic tools in human cancers including lung cancer [41–45]. However, it is unclear if these components can be generally useful in classifying lung cancer, as the microRNA signatures did not differ qualitatively between lung cancer and normal subjects [46] while the accuracy of protein markers for advanced stage NSCLC detection was only 75%. Such limitations do not meet the specificity and sensitivity requirements for lung cancer screening at early stages.

We have procured blood plasma samples from 39 normal and 91 NSCLC subjects (44 early stage and 47 late stage) for EXO isolation and lipid profiling using UHR-FTMS. We also applied two advanced multivariate statistical methods, Random Forest (RF) and Least Absolute Shrinkage and Selection Operator (LASSO) to perform supervised clustering analysis of the EXO lipid profiles. The Area Under the Receiver Operating Characteristic curve (AUROC) of normal versus early and late stage NSCLC using the top 16 (for RF) or top 6 (for LASSO) lipid features was 0.85 and 0.88 or 0.79 and 0.77, respectively. These results showed that selected lipid species of plasma EXO discriminated normal from early and late stage NSCLC and demonstrate the value of RF and LASSO for metabolomics-based biomarker development.

Please cite this article in press as: T.W.M. Fan, et al., Exosomal lipids for classifying early and late stage non-small cell lung cancer, Analytica Chimica Acta (2018), https://doi.org/10.1016/j.aca.2018.02.051