Journal of Proteomics 142 (2016) 114–121

Contents lists available at ScienceDirect

Journal of Proteomics

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

An integrated cell line-based discovery strategy identified follistatin and kallikrein 6 as serum biomarker candidates of breast carcinoma

Alain Mangé [a,b,c,d], Lampros Dimitrakopoulos [e], Antoninus Soosaipillai [e], Peter Coopman [a,b,c,d], Eleftherios P. Diamandis [e], Jérôme Solassol [a,b,c,d,*]

[a] IRCM, Institut de Recherche en Cancérologie de Montpellier, Montpellier, F-34298, France
[b] INSERM, U1194, Montpellier, F-34298, France
[c] Université de Montpellier, Montpellier, F-34090, France
[d] Institut régional du Cancer de Montpellier, Montpellier, F-34298, France
[e] Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 11 January 2016
Received in revised form 7 April 2016
Accepted 27 April 2016
Available online 7 May 2016

Keywords:
Breast cancer
Biomarkers
Proteomics

ABSTRACT

Secreted proteins constitute a relevant source of putative cancer biomarkers. Here, we compared the secretome of a series of four genetically-related breast cancer cell lines as a model of aggressiveness using quantitative mass spectrometry. 537 proteins (59.5% of the total identified proteins) predicted to be released or shed from cells were identified. Using a scoring system based on i) iTRAQ value, ii) breast cancer tissue mRNA expression levels, and iii) immunohistochemical staining (public database), a short list of 10 candidate proteins was selected. Using specific ELISA assays, the expression level of the top five proteins was measured in a verification set of 56 patients. The four significantly differentially expressed proteins were then validated in a second independent set of 353 patients. Finally, follistatin (FST) and kallikrein 6 (KLK6) in serum were significantly higher (p-value < 0.0001) in invasive breast cancer patients compared with non-cancerous controls. Using specific cut-off values, FST distinguished breast cancer samples from healthy controls with a sensitivity of 65% and an accuracy of 68%, whereas KLK6 achieved a sensitivity of 55% and an accuracy of 61%. Therefore, we concluded that FST and KLK6 may have significance in breast cancer detection.

Biological significance: Discovery of new serum biomarkers that exhibit increased sensitivity and specificity compared to current biomarkers appears to be an essential field of research in cancer. Most biological markers show insufficient diagnostic sensitivity for early breast cancer detection and, for the majority of them, their concentrations are elevated only in metastatic forms of the disease. It is therefore essential to identify clinically reliable biomarkers and develop effective approaches for cancer diagnosis. One promising approach in this field is the study of secreted proteins through proteomic analysis of in vitro progression breast cancer models. Here we have shown that FST and KLK6 are elevated in breast cancer patient serum compared to healthy controls. We expect that our discovery strategy will help to identify cancer-specific and body-fluid-accessible biomarkers.

1. Introduction

Early detection of breast cancer, so as to diagnose and treat cancer in its state prior to clinical symptoms and/or metastasis, may greatly impact the treatment and prognosis of patients with this common, but deadly, malignancy. Clinical breast exam and breast self-exam did not show a clear benefit to increase early cancer detection. Diagnostic mammography can often help find breast cancer at an early stage. However, it can also miss some cancers. False-negative mammograms can be attributed to technical or interpretive reasons including the absence of calcifications and, more importantly, high breast density [1,2]. In addition, sometimes more tests are needed to find out if something found on a mammogram is or is not cancer. Therefore, there is an urgent need to develop complementary approaches to improve the sensitivity and specificity of early breast cancer detection.

As whole blood is considered to provide a dynamic representation of an individual's physiological and pathological status, human serum/plasma represents the most extensively studied biological matrix in the quest for cancer biomarkers [3]. Known serum-based tumor markers, such as CA15.3 or BR27.29, cannot be used for breast cancer detection. These markers, having insufficient predictive value as an early detection blood diagnostic assay, are only recommended for monitoring treatment response and disease recurrence of patients with metastatic disease [4]. Therefore, the search for specific disease-

Abbreviations: FDR, false discovery rate; GO, gene ontology; MMTS, methyl methanethiosulfonate; SCX, strong cation exchange; TCEP, tris-(2-carboxyethyl) phosphine.

* Corresponding author at: Department of Biopathology, CHU Arnaud de Villeneuve, 191 Avenue du Doyen Giraud, 34295 Montpellier Cedex 5, France.
E-mail address: j-solassol@chu-montpellier.fr (J. Solassol).

http://dx.doi.org/10.1016/j.jprot.2016.04.050
1874-3919/

associated biomarker signatures is of particular interest since they could be applied in a standard clinical setting. Biomarker discovery for this disease is still very much in its discovery phase.

Multiple approaches have been developed that hold promise for the identification of serum biomarkers. Among them, quantitative proteomics yields information that specifically recognizes the differences between samples. Numerous studies have already shown that this methodology can be used to uncover proteomic expression patterns linked with cancer, and some expression patterns have shown high promise to discover new biomarkers of early-stage cancers [5]. The challenges of blood proteomics, stemming from the complexity of the fluid, have led researchers to seek alternate sources for the discovery of circulating cancer biomarkers [6]. In tumor progression, the "secretome" proteins released or shed by cells, tissues, or organisms through various pathways, act as mediators of cancer cell-host communication in the cancer microenvironment [7]. These proteins may be detected by analyzing the conditioned media of cultured cell lines derived from specific cancer types. Interestingly, these proteins are often present at relative high concentration in proximity to tumor site but are also more likely to end up in body fluids such as serum/plasma in sufficient concentration to be detectable. However, search for cancer biomarkers using cell lines established from different individuals is complicated by differences among the cancer cell donors, their origin, the passage number and culture conditions, creating huge variations that may be unrelated to normal and malignant behaviors. A comparison of malignant and non-malignant cell line variants from the same lineage will avoid this problem.

Here, we report a progression-related differential secretome analysis between MCF10 genetically-related breast cancer cell lines [8–11] using 2D nanoLC-MS/MS and isobaric tags for relative and absolute quantitation (iTRAQ) labeling technology. We identified a total of 903 proteins. Among them, 109 were found to be present at significantly elevated levels in breast cancer cell lines compared to normal or premalignmant cell lines. The differential expression of selected proteins was further validated by specific ELISA assays in a large independent cohort of invasive breast cancer patients. We demonstrated that the serum levels of two proteins (follistatin and kallikrein 6) were significantly higher in breast cancer patients versus healthy controls.

2. Methods

2.1. Patient selection

Sera from female breast cancer patients were prospectively collected between 2005 and 2009 at the "Institut du Cancer de Montpellier", France, after obtaining written informed consent (Inserm RBM03–63 cohort). The study protocol was approved by the institutional review board. The healthy controls were collected at in the same institute during the same period, and were sex- and age-matched. All patients and healthy controls are Caucasian. The verification population set consisted of 56 subjects: 28 women whith histopathologic diagnosis of breast cancer, and 28 healthy controls with negative mammograms, negative physical breast exams for at least 4 years and no history of prior malignancy, immunodeficiency, autoimmune disorder, hepatitis or HIV infection. The validation population set included 353 patients: 241 patients with a histopathologic diagnosis of breast cancer, and 112 healthy controls exempt of any breast, infectious and/or inflammatory disease. Sera from patients with breast cancer were collected at the time of cancer diagnosis and just prior surgery. All samples were collected, processed, and stored in a similar fashion. Blood samples were centrifuged at 1250 g for 5 min, and sera were then stored at −80 °C. Detailed clinical and pathological information of breast cancer patients are listed in Table 1.

Table 1
Clinicopathologic characteristic of patients with invasive breast cancer.

Characteristics	Verification Set	Validation set
	N = 28 (%)	*N* = 241 (%)
Age (years) median, [min–max]	68 [38–86]	72, [37–97]
Histotype		
Ductal	28 (100)	222 (92.1)
Lobular	–	19 (7.9)
Tumor size		
T1	15 (53.6)	115 (47.7)
T2	10 (35.7)	112 (46.5)
T3	1 (3.6)	4 (1.7)
T4	2 (7.1)	10 (4.1)
Histological grade		
I	4 (14.3)	39 (16.2)
II	18 (64.3)	126 (52.3)
III	5 (17.8)	72 (29.9)
Missing	1 (3.6)	4 (1.7)
Lymph node status		
Negative	14 (50.0)	121 (50.2)
Positive	14 (50.0)	120 (49.8)
Estrogen receptor		
Negative	1 (3.6)	46 (19.1)
Positive	27 (96.4)	195 (80.9)
Progesterone receptor		
Negative	6 (21.4)	60 (24.9)
Positive	19 (67.9)	181 (75.1)
Missing	3 (10.7)	0
Her-2 overexpression		
Negative	19 (67.9)	138 (57.3)
Positive	2 (7.1)	25 (10.4)
Missing	7 (25.0)	78 (32.4)
Classification		
Luminal A	19 (67.8)	127 (52.7)
Luminal B	1 (3.6)	14 (5.8)
Enriched Her-2	1 (3.6)	11 (4.6)
Others	–	11 (4.6)
Missing	7 (25)	78 (32.3)

2.2. Cell cultures and conditioned medium

The human MCF10A (non-tumorigenic), MCF10·NeoT (premalignant; tumorigenic), MCF10.DCIS (tumorigenic and locally invasive) and MCF10·CA1d (tumorigenic and metastatic) breast cancer cell lines were purchased from Asterand (MCF10.DCIS), the Barbaba Ann Karmanos Cancer Institute (MCF10·NeoT and MCF10·CA1d) and ATCC (MCF10A). Cells were maintained in DMEM/F12 (1:1) supplemented with 5% horse serum, 10 μg/mL insulin, 25 ng/mL epidermal growth factor, 0.5 μg/mL hydrocortisone, and 100 ng/mL cholera toxin. Culture media were supplemented with penicillin/streptomycin (100 U/mL). Cells were cultured at 37 °C in 5% CO_2. DMEM/F12, horse serum and penicillin/streptomycin were obtained from Gibco and insulin, EGF, hydrocortisone and cholera toxin from Sigma-Aldrich. Cells were grown to 60% confluence in 100 mm culture dishes and were rinsed three times for 15 min with serum free medium. The cells were then incubated in the serum free medium at 37 °C for 18 h. The cells were 95–100% viable after the serum free growth, as determined by trypan blue exclusion counting (Supplemental Fig. 1A). The conditioned medium was centrifuged for 10 min at 800 g to remove suspended cells. Samples were then concentrated using Amicon Ultra Centrifugal Filter Unit with a 10 kDa cut-off (Millipore). The protein concentration was measured in triplicate using the Micro BCA Kit (Pierce). Equal loading of proteins onto protein gels showed the same pattern of bands for all the samples (Supplemental Fig.1B).

2.3. iTRAQ reagent labeling and mass spectrometry analysis

The experimental design used for this study is illustrated in Fig. 1. iTRAQ labeling and mass spectrometry analysis were performed as previously described [12]. Briefly, 80 μg protein of each sample were digested using trypsin before iTRAQ labeling. Labeled peptides were separated on an IPG Drystrip 24 cm, pH 3–10 using the Agilent 3100 OFFGEL Fractionator (Agilent), and all fractions were analyzed by nanoLC-MS/MS using a MALDI TOF/TOF 4800 mass spectrometer