Mass spectrometry-based profiling of lysine acetylation and arginine methylation for biomarker discovery in triple negative breast cancer

Alissa J. Schunter, Hongbo Gu, Jian Min Ren, Vicky Yang, Matthew P. Stokes
Cell Signaling Technology, Inc., Danvers MA 01923

INTRODUCTION

Triple negative breast cancer (TNBC) is the most aggressive type of breast tumor and there are currently no approved targeted therapies. Better biomarkers are needed for early detection and for therapeutically informative subtyping of this genetically heterogeneous disease. We recently reported a methodology to enrich post-translationally modified peptides from serum (1) and found acetyl lysine (AcK) and mono-methyl arginine (Rme) to be some of the most abundant post-translational modifications (PTMs) in the sera of cancer patients. This method has the advantage of profiling serum samples without prior depletion of abundant serum proteins, a major limitation of current proteomic methods. To develop a biomarker signature of TNBC, these two PTMs were profiled in the sera of 10 patients with stage I-IIA TNBC and 10 healthy female controls.

METHODS

Antibody enrichment and liquid chromatography mass spectrometry (LC-MS) were used to profile lysine acetylation and arginine mono-methylation in TNBC patient sera. Serum proteins were trypsin digested prior to immunoaffinity enrichment of the modified peptides with PTM-specific antibodies. The enriched peptides were analyzed by LC-MS and relative abundance of peptides across samples was measured using label-free quantification.

CONCLUSIONS

The PTMScan anti-acetyl lysine and monomethyl lysine antibodies successfully enriched hundreds of modified peptides directly from serum without prior depletion of highly abundant proteins. Rme enrichment allowed for even deeper coverage of lower-abundance proteins than AcK. Broad down-regulation of acetyl lysine was observed across TNBC patient samples, which suggests a pattern of either acetyltransferases becoming inactivated or deacetylases becoming over-active. AcK profiles were found to correctly differentiate TNBC patients from healthy controls, whereas Rme profiles do not correlate to disease status. Several AcK peptides are under further investigation to confirm their utility as potential biomarkers. This work shows that clinically relevant biomarker profiles can be distinguished in in a non-invasive serum-based assay using a PTM-specific antibody.

REFERENCES

1. Gu H, Ren JM, Jia X, Levy T, Rikova K, Yang V, et al. Quantitative Profiling of Post-translational Modifications by Immunoaffinity Enrichment and LC-MS/MS in Cancer Serum without Immunodepletion. Mol Cell Proteomics. 2016;15:692–702. Available from: http://www.mcponline.org/lookup/doi/10.1074/mcp.O115.052266

Experimental Design

Acetyl lysine (AcK) and monomethyl arginine (Rme)- modified peptides were enriched from the sera of ten patients with TNBC and ten healthy volunteers, following the PTMScan workflow as shown in **Figure 1**. The patients' clinical characteristics are described in **Table 1**.

Figure 1

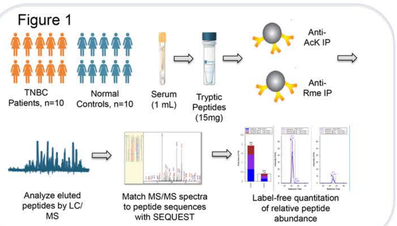

Figure 1. PTMScan workflow. Serum was collected from ten patients diagnosed with TNBC plus ten healthy volunteers. The sera were purchased from ProteoGenex (Culver City, CA). Equal aliquots of serum containing 15mg of total protein were diluted 1:3 in 9M urea buffer digested with trypsin, then desalted and lyophilized. Acetyl lysine (AcK)- and monomethyl arginine (Rme)- containing peptides were enriched by immunoprecipitation with modification-specific antibodies. The enriched peptides were analyzed by LC-MS in duplicate 90 min injections on a Fusion Lumos Orbitrap mass spectrometer. MS/MS spectra were assigned to peptide sequences using the SEQUEST algorithm, and label-free relative quantification was performed by extracting precursor ion peak areas using Skyline. High-scoring peptide integrations were manually verified in Skyline for additional confidence.

Table 1. Patient characteristics

Code	Age	Grade	TNM	Stage	ER	PR	Her2 neu
P1	49	G3	T2N0M0	IIA	negative	negative	0
P2	57	G2	T1cN0M0	I	negative	negative	1+
P3	41	G3	T1cN0M0	I	negative	negative	0
P4	42	G3	T1bN0M0	I	negative	negative	0
P5	38	G3	T2N0M0	IIA	negative	negative	1+
P6	49	N/A	T1cN0M0	I	negative	negative	1+
P7	36	G2	T2N0M0	IIA	negative	negative	1+
P8	58	G2	T1cN0M0	I	negative	negative	0
P9	34	G3	T1cN0M0	I	negative	negative	1+
P10	35	G3	T2N0M0	IIA	negative	negative	1+

All Patients:
- Diagnosed with infiltrating ductal carcinoma
- Female
- Caucasian
- Non-smokers
- BMI<30

Table 1. Patient Characteristics. All 10 patients were diagnosed with infiltrating ductal carcinoma and tested negative for estrogen receptor (ER) and progesterone receptor (PR). All ten were Caucasian females, nonsmokers, and had BMIs below 30.

AcK and Rme Peptide Identifications

The Rme enrichment identified the most modification sites, 728 unique peptides on 530 proteins. AcK enrichment yielded 579 modified peptides originating from 217 proteins. Albumin accounted for just 14% of the AcK peptides and 3% of the Rme peptides, which is low considering that albumin typically accounts for approximately 50% of total serum protein content. The gene ontology annotations of the modified proteins are shown in **Figure 2**. Highly abundant proteins such as albumin, various immunoglobulins, and fibrinogen contribute more than half the peptides in the "transporter", "unknown function", and "secreted protein" categories. The Rme immunoprecipitation was better able to enrich lower-abundance proteins than the AcK immunoprecipitation, as seen by the greater fraction of proteins with functions in transcription, RNA binding, translation, and so on.

Figure 2A

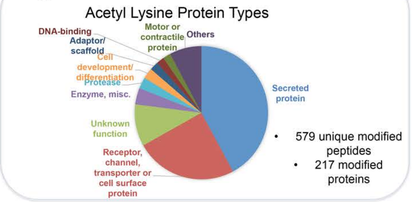

Figure 2B

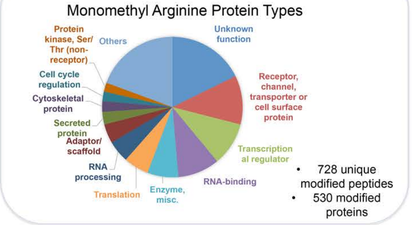

Figure 2. Types of proteins identified with AcK (**A.**) or Rme (**B**) modifications. The charts show the number of modified peptides from each protein type as a percentage of all modified peptides identified. The "unknown function" category includes several immunoglobulins, albumin is categorized as a transporter, and the "secreted protein" category includes several high-abundance proteins such as fibrinogen, alpha-1-antitrypsin, haptoglobin, and complement factors.

AcK Peptides Significantly Down-Regulated

Average peak areas were calculated for patients and controls for each modified peptide. Two-sided t-tests were performed for each peptide with a Benjamini-Hochberg correction for multiple comparisons. Peptides were excluded from this analysis when peak areas were missing for five or more patients or normal individuals. In general, AcK abundances were lower in patients than in controls, as seen by the skew towards negative fold changes in **Figure 3A**. Four AcK peptides were found significantly downregulated with an adjusted p-value <0.01 and an additional 54 peptides were significant at the p<0.05 level. **Figure 3B** shows the log2-fold changes and t-test significance values for Rme peptides. Although some peptides showed large magnitude fold changes (greater than +/- 4-fold), none were significant at the p<0.05 level.

Figure 3A

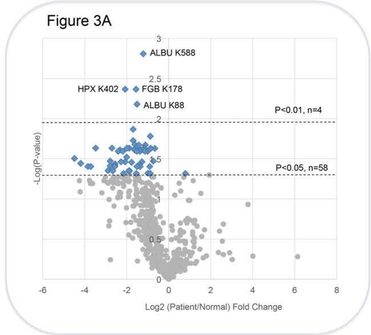

Figure 3B

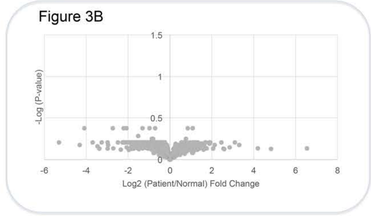

Figure 3. Log_2 -fold changes of average patient modified peptide abundance over normal abundance vs t-test significance. Modified peptides with significant p-values are shown in blue diamonds and all other peptides are shown in gray circles. **A.** AcK modified peptides. The gene names and acetylated lysine sites of selected peptides are also indicated. **B.** Rme modified peptides.

Patients and Normals Differentiated by AcK Profiles

Unsupervised hierarchical clustering was performed on modified peptides for which there were no missing values (n=518 AcK peptides, 680 Rme peptides). **Figure 4** shows heatmaps with this clustering with the Z-scores of the modified peptide abundances. AcK profiles were sufficient to separate the TNBC patients from the normal controls with just two exceptions out of 20 samples. Patient 10 (P10) clustered separately from both the normals and the other patients; this was due to globally higher LCMS intensities across both modified and unmodified peptides. Normal 10 (N10) was clustered closest to P4 and P1.

The AcK peptides clustered into two distinct groups. In the top portion of the heatmap were peptides that were broadly lower in abundance in the patients than in the normal controls. The lower portion has a few peptides that are more abundant in the patients than in the controls, but mostly consists of peptides that do not show any clear pattern of expression.

The Rme clustering was not able to distinguish patients from normal controls and the heatmap shows a random distribution of peptide abundance. This contrasts with the clear patterns in the AcK dataset.

Figure 4A AcK

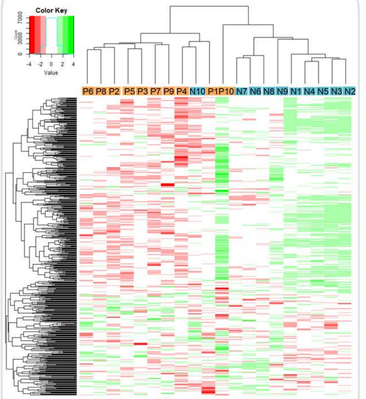

Figure 4B Rme

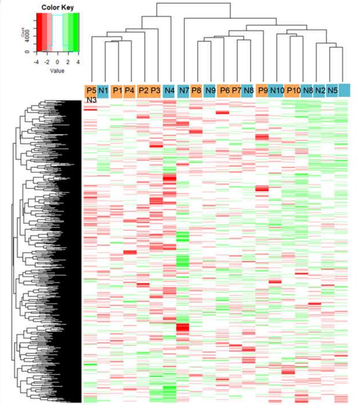

Figure 4. Heatmap of (**A.**) acetyl lysine and (**B.**) monomethyl arginine peptides. The normalized peak areas were log10- transformed and then scaled across each row (peptide) so that the mean individual for each peptide is shown in white and those with higher or lower abundance are shown in green or red, respectively. Hierarchical clustering was performed using Euclidean distance and the dendrogram was built with the complete linkage method. The columns are the individual patients or normal controls, while each row is a modified peptide.

Cell Signaling TECHNOLOGY®

For Research Use Only. Not For Use in Diagnostic Procedures

Alissa J. Schunter
email: alissa.schunter@cellsignal.com
www.cellsignal.com/posters