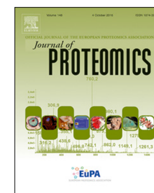
ARTICLE IN PRESS

Journal of Proteomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Proteomics

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



Discrimination and quantification of homologous keratins from goat and sheep with dual protease digestion and PRM assays

Chen Miao[a,b], Yunfei Yang[a], Shanshan Li[c], Yufeng Guo[a], Wenqing Shui[c,d,**], Qichen Cao[a,*]

[a] Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] iHuman Institute, Shanghai Tech University, Shanghai 201210, China
[d] School of Life Science and Technology, Shanghai Tech University, Shanghai 201210, China

ARTICLE INFO

Keywords:
Mass spectrometry
Keratin
Keratin-associated proteins (KAPs)
Species identification
Homology
Hair fiber

ABSTRACT

Mass spectrometry (MS) technology has a special advantage in species determination for protein-rich samples which requires identification of species-specific peptides. However, for species discrimination of highly homologous proteins, it remains challenging to select the species unique peptides with routine proteomics approaches. In this work, we chose keratins and keratin-associated proteins (KAPs) present in cashmere fibers from goat and wool fibers from sheep as targets, to develop a dual-protease digestion workflow based on in-silico and experimental analysis. Combined usage of Glu-C and trypsin proteases showed the best digestion performance for MS identification of keratins and KAPs from different species. The parallel reaction monitoring (PRM) technique was implemented to validate and quantify the selected species discriminable peptides. The fiber composition of both blended animal hair fibers and industrial textile fabrics were successfully determined with the PRM assay. Furthermore, we identified over 360 peptides from the cashmere fiber beyond the current Uniprot goat proteome database. We expect our new workflow would improve the identification and quantification of keratin and KAPs, and provide inspiration for distinguishing other highly homologous proteins. We also anticipate the set of species-specific peptides from keratin or KAPs validated in this work would benefit the quality assessment for industrial fiber materials and textile products.

Significance: Discriminating species from highly homologous proteins is challenging for MS-based shotgun proteomics. The large percentage of overlapped protein sequence hinders the identification of the species unique peptides. In this work, we aimed to discriminate sample species between goat and sheep from keratins and keratin-associated proteins (KAPs). A dedicated workflow was developed to boost the exposure and quantification of species discriminable peptides. The dual-proteases digestion approach was optimized based on amino acid sequence analysis and protein in-silico digestion analysis. The PRM assays were established to validate and quantify the selected species unique peptides. Additionally, we have identified about 360 novel candidate peptides complementary to the current goat protein sequence database. We expect our workflow would improve the species discrimination for highly homologous proteins and benefit the proteomics study of keratin and KAPs in the human proteome.

1. Introduction

Nowadays, mass spectrometry (MS) is becoming a leading technology for identification and quantification of proteins and is increasingly utilized for species discrimination of protein-rich samples [1–7]. MS based methods distinguish one species from another by detecting the species-specific proteins or identifying the difference of protein sequence between different species. To-date, two basic MS approaches are commonly applied in protein identification, based on whether intact proteins or the peptides of protein digestion are introduced into MS instruments. The former is termed as “Top-down” proteomics and the latter is “bottom-up” or “shotgun” proteomics. The bottom-up MS approach is widely applied in a high-throughput manner, while, the top-down MS approach was not so frequently adopted [8, 9]. Due to the fact that, the bottom-up approach detects only peptides, it's important to identify the unique peptides to confirm the presence of specific proteins

* Corresponding author.
** Corresponding author at: iHuman Institute, Shanghai Tech University, Shanghai 201210, China.
E-mail addresses: shuiwq@shanghaitech.edu.cn (W. Shui), cao_qc@tib.cas.cn (Q. Cao).

https://doi.org/10.1016/j.jprot.2018.07.010
Received 23 March 2018; Received in revised form 3 July 2018; Accepted 13 July 2018
1874-3919/ © 2018 Elsevier B.V. All rights reserved.

and species [10]. However, for proteins in highly homologous or evolutionary conservative species, that often share high percentages of amino acid sequence homology, the amount of the unique peptides is limited. Therefore, it is very challenging to distinguish between species that contain large numbers of homologous proteins.

Keratin, a superfamily of structural and ubiquitous proteins, has been found from hair to the nuclei of cells [5, 11, 12]. Due to its lightweight and sturdy properties, keratin-materials have been made into various luxurious textiles or artware. Species identification of keratins is therefore important for the quality assessment of the keratinous products. For keratins, one of the significant features is the high degree of homology found within each family, which can amount to 92% between some members of the Type I keratins and 85% among the Type II keratins [13]. In hair, the keratins are embedded in a matrix of keratin-associated proteins (KAPs) [14], a large group of proteins that contains more than 20 families, with most proteins having multiple isoforms. Similar degrees of homology has also been observed for the KAPs [15]. Because of the difficulties in MS identification owing to the substantial sequence homology, coverages of keratins and KAPs are below average in the reported human proteome draft datasets [16], and it is even more challenging to quantify the expression of different isoforms of keratins and KAPs [17, 18].

To address the challenges of identification and quantification of highly homologous keratin and KAPs, in this work, we took goat cashmere and sheep wool fiber as benchmark samples, and a dedicated shotgun proteomics workflow was developed to distinguish the fiber originated species. We first selected the proper protease for protein digestion by carrying out the amino acid sequence analysis and protein in-silico digestion analysis. The best protease combination was confirmed experimentally and was further used to identify the unique peptides of keratin and KAPs from specific species. To maximize the sensitivity and reproducibility for quantifying these unique peptides in the background of high abundance shared peptides, a Parallel Reaction Monitoring (PRM) [19, 20] based MS assay has been developed. The robustness of the PRM assay was evaluated by the blend samples composed of animal fibers originated from distinct species. In total, two species discriminating peptides were successfully quantified in industrial textile samples. We anticipate that the workflow presented in this article could be further implicated in the study of highly homologous proteins in the human proteome. And the PRM assays for the keratin discriminating peptides developed in this work will benefit the quality control (QC) of industrial textiles, which are traditionally examined manually and depend on the expertise of the operator heavily [6].

2. Materials and methods

2.1. Protein extraction from fiber sample

Inner Mongolia white cashmere (C) and fine wool (F) fiber were used for untargeted marker screening and three commercial textile fiber was used for targeted marker verification. Fiber samples were pretreated before protein extraction. Detailed procedure has been introduced in the previous article [21]. Briefly, about 5 mg of pretreated fiber sample was cut up with scissor and added to the extraction buffer (25 mM Tris-HCl, 5 M urea, 2.4 M thiourea, 5% DTT, pH 8.5), slightly shaking for 16 h at 55 °C. Residual fiber sample was then removed by centrifugation at 12000*g* for 5 min. The protein of the supernatant was precipitated with acetone in a protein-to-acetone ratio of 1:6 (v/v) for at least 2 h under −20 °C and centrifuged at 3000 rpm for 5 min at 4 °C. The protein pellets were further washed with cooled acetone for three times. The supernatant was removed, and the residual acetone was volatilized under nitrogen until about 10 μL left. The protein pellets were re-suspended with re-dissolve buffer (20 mM Tris-HCl, 8 M urea). The protein concentration was determined using the Bradford assay and then the samples were kept at −80 °C.

2.2. Protein digestion

About 100 μg protein solution was reduced with 10 mM dithiothreitol (DTT) at 37 °C for 4 h in a thermo-shaker at 550 rpm. The alkylation was performed with 40 mM iodoacetamide (IAA) at room temperature in darkness for 40 min. Additional 30 mM DTT was added to consume the excess IAA, followed by vortexing and incubation at 37 °C for 40 min. The protein digestion was conducted in diverse conditions: i. For trypsin digestion, urea was diluted with 50 mM ammonium bicarbonate to a final concentration of less than 1 M, proteins were digested using sequencing grade modified trypsin (Promega, Madison, USA) at an enzyme-to-protein ratio of 1:100 (w/w) at 37 °C for 4 h, followed by adding fresh trypsin at 1:50 (w/w) before incubating at 37 °C overnight. ii. For trypsin-chymotrypsin digestion, urea concentration was adjusted to less than 1 M with 50 mM ammonium bicarbonate, chymotrypsin (Promega, Madison, USA) was added firstly at an enzyme-to-protein ratio of 1:150 (w/w) reacting at 25 °C for 2 h and fresh chymotrypsin was added at a ratio of 1:150 (w/w) at 25 °C for 12 h, then trypsin was added at the ratio of 1:150 (w/w) at 37 °C for 3 h and additional trypsin was added at the same amount incubating overnight. iii. For trypsin-Glu-C digestion, Glu-C and trypsin were added simultaneously or sequentially. When digestion reacted simultaneously, 50 mM ammonium bicarbonate was used to dilute urea to a concentration of less than 1 M, trypsin and Glu-C were then added at an enzyme-to-protein ratio of 1:100 (w/w) and 1:40 (w/w) respectively at 37 °C for 4 h, followed by adding fresh trypsin at the ratio of 1:50 (w/w) and Glu-C at the ratio of 1:40 (w/w) before incubation at 37 °C overnight. When digestion reacted sequentially, urea was diluted with phosphate buffered solution (pH 7.5) to a final concentration of less than 1 M, Glu-C was firstly added at an enzyme-to-protein ratio of 1:60 (w/w) at 25 °C for 4 h and then fresh Glu-C was added at the same amount incubating for 13 h, hereafter, trypsin was added at the ratio of 1:100 (w/w) at 37 °C for 4 h, followed by additional trypsin added at the ratio of 1:50 (w/w) before incubation at 37 °C overnight. All the digestion described above were terminated by adding formic acid to pH 3. The protein lysate was desalted with C18 tip columns (Nest, MA, USA) and lyophilized under vacuum.

2.3. Nano LC-MS/MS analysis

The peptides were dissolved with LC mobile phase A (2% acetonitrile, 0.1% formic acid) and about 1 μg peptide was subjected to nano-LC-MS/MS analysis with an Eksigent Nano LC coupled to Triple-TOF™ 5600 mass spectrometer (SCIEX, USA) with a nano-electrospray ionization source. The peptides were firstly loaded onto a C18 (5 μm) trap column and then switched to an in-house packed 150 mm × 75 μm analytical column with Reprosil-Pur Basic C18 (3 μm) sorbent. LC mobile phase buffer B was composed of 2% water and 0.1% formic acid in acetonitrile. Peptides were separated with a 120 min discontinuous gradient of 5% - 95% buffer B with a flow rate of 300 nL/min. For untargeted proteomics analysis, the instrument was operated in the data-dependent acquisition (DDA) mode. Precursor MS scan range was set to 350–1500 m/z with ion charge states of 2–5. The top 40 strongest precursor ions were fragmented with 22 s of dynamic exclusion time. The targeted proteomic analysis was conducted in parallel reaction monitoring (PRM) mode [22]. The acquisition list was generated from candidate marker peptides selected based on the DDA results. The collision energy was optimized for each peptide precursor to obtain high-quality MS/MS spectra. The top 3 to 6 product ions by intensity were used to construct the transitions for quantitating individual peptide precursors. All the transitions were validated using the mProphet algorithm [23].

2.4. DDA data processing for peptide markers screening

Obtained DDA data were searched against target-decoy database

Download English Version:

https://daneshyari.com/en/article/7633217

Download Persian Version:

https://daneshyari.com/article/7633217

Daneshyari.com