Molecular & Biochemical Parasitology 211 (2017) 9–14

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Proteomic investigation of human cystic echinococcosis in the liver

Rémi Longuespée[a,b,*], Rita Casadonte[a], Mark Kriegsmann[b], Petra Wandernoth[c], Katharina Lisenko[d], Gabriel Mazzucchelli[e], Michael Becker[f], Jörg Kriegsmann[a,c,g]

[a] Proteopath GmbH, Trier, Max Planck Strasse 17, 54296 Trier, Germany
[b] Institute of Pathology, University of Heidelberg, Im Neuenheimer Feld 224, 69120 Heidelberg, Germany
[c] Molecular Pathology Trier, Max Planck Strasse 17, 54296 Trier, Germany
[d] Department of Internal Medicine V, University of Heidelberg, Im Neuenheimer Feld 410, 69120 Heidelberg, Germany
[e] Mass Spectrometry Laboratory (LSM), Systems Biology and Chemical Biology, GIGA-Research, University of Liège, Allée du 6 aoÛt 11, 4000 Liège, Belgium
[f] Bruker Daltonik GmbH, Bremen, Germany, Fahrenheitstrasse, 28359 Bremen, Germany
[g] MVZ for Histology, Cytology and Molecular Diagnostics Trier, Max Planck Strasse 5, 54296 Trier, Germany

ARTICLE INFO

Article history:
Received 4 October 2016
Received in revised form
21 November 2016
Accepted 9 December 2016
Available online 13 December 2016

Keywords:
Histoproteomics
Cystic echinococcosis
Laser microdissection
Parasitic disease
MALDI IMS

ABSTRACT

Cystic echinococcosis (CE) is a pandemic infectious disease caused by the tapeworm *Echinococcus granulosus* that forms cysts in different organs such as lungs and liver. Imaging examination and serological tests have some drawbacks such as low sensitivity. In this study, we used an up-to-date workflow of laser microdissection-based microproteomics and matrix-assisted laser desorption/ionization time-of-flight imaging mass spectrometry in order to depict the proteomic pattern of CE in the liver. This investigation revealed specific markers of a parasitic cyst in liver. This proteomic pattern could facilitate diagnosis of CE in the future.

1. Introduction

Cystic echinococcosis (CE) is an endemic parasitic disease which is caused by infection with metacestodes (larval stage) of the tapeworm *Echinococcus granulosus* (Eg) [1]. Humans as well as herbivores such as sheeps, goats or cattle are accidental intermediate hosts that become infected by ingested water, food, or soil contaminated with faeces from infected carnivores such as dogs or foxes which are definite hosts. Upon ingestion of the eggs, oncopheres are released that have the capacity to penetrate the human intestinal wall where they gain access to the portal venous system and subsequently various organs where the oncospheres develop into cysts. The liver (70%) and lungs (20%) are most commonly affected. However, cyst formation can occur anywhere in the body (10%). Since the clinical course is relatively quiescent and patients present with mild symptoms that are dependent on the cyst location and size, CE poses challenges for accurate and timely diagnosis. Additionally it can be expected that increased emigration of populations where prevalence rates are as high as 5–10% such as in certain parts of South Africa, Eastern Europe, Russia, the Middle East, and China raise the possibility to attend patients affected by CE also in western countries [2]. Diagnosis is usually achieved by a combination of clinical history, serum analyses and imaging observation such as ultrasound, computed tomography or magnetic resonance tomography. Whereas in early infections, serological studies are used, in late stages imaging studies are usually performed [3]. Although the host does produce detectable humoral and cellular responses, serological studies suffer from low sensitivities ranging from 30 to 80% depending on the assay employed [4]. In some cases the diagnosis of CE is delayed until histopathological examination of the resection specimen, especially in uncommon locations of echinococcal cysts for example ossal, intramuscular, retroperitoneal or spinal [5–9]. To confirm the diagnosis of CE a Periodic acid Schiff (PAS)

Abbreviations: ACN, Acetronitrile; CA, Citric acid; CE, Cystic echinococcosis; Eg, *Echinococcus granulosus*; ECM, Extra cellular matrix; FFPE, Formalin fixed paraffin embedded; fr/fr, Fresh/frozen; GO, Gene ontology; HIAR, Heat-induced antigen retrieval; HIV, Human immunodeficiency virus; HLA, Human leucocyte antigen; IMS, Imaging mass spectrometry; LMD, Laser microdissection; LC–MS/MS, Liquid chromatography-mass spectrometry/mass spectrometry; MALDI, Matrix-assisted laser desorption-ionization; PAS, Periodic acid Schiff; ID, Protein identification; ROI, Region of interest; TIC, Total ion count; UPLC, Ultra-performance liquid chromatography.

* Corresponding author at: Im Neuenheimer Feld 224, 69120 Heidelberg, Germany.
E-mail address: remi.longuespee@med.uni-heidelberg.de (R. Longuespée).

http://dx.doi.org/10.1016/j.molbiopara.2016.12.002
0166-6851/

reaction is routinely performed since the tapeworm stains PAS-positive. Treatment mostly relies on surgery for the removal of the cysts and chemotherapy [1].

CE has been studied on the genetic level previously [10,11]. Even though proteomic investigations have been performed in the past [12–20], few of new protein sequences of Eg have been reported in databases. The proteome of Eg is indeed still poorly described in universal databases such as Uniprot, with only few protein sequences available. Moreover, very few studies have been executed on human samples.

In the present study, we applied up-to-date methods for the proteomic analysis of formalin-fixed and paraffin-embedded (FFPE) tissues, the standard processing method to store tissues in pathology institutes worldwide.

We used laser microdissection (LMD)-based liquid chromatography tandem mass spectrometry (LC–MS/MS) for the simultaneous extraction and digestion of proteins from specific tissue regions in a patient with CE: the cyst liquid containing the worm, the cystic wall and the surrounding liver tissue. Additionally, we performed matrix-assisted laser desorption/ionization time-of-flight (MALDI) imaging mass spectrometry (IMS) to correlate proteomic data and morphology. This is the first study to perform LMD and IMS on FFPE tissue from a patient with CE and, to the best of our knowledge, the first report combining LMD and IMS for the diagnosis of CE.

2. Material and methods

2.1. Preanalytics

The tissue sample was provided by the University of Heidelberg with institutional ethical review approval. After surgery, the tissue was fixed in buffered formalin, embedded in paraffin, and sectioned [21]. PAS reaction was performed for the visualization of polysaccharides, glycogen, glycoproteins and glycolipids in the tissue section.

Materials for LMD analyses and IMS were described previously [21,22].

2.2. LMD and LC–MS/MS

Different regions of interest (ROIs) of the tissue section, including the cyst fluid containing Eg (“cyst” in the article), the cystic wall and the surrounding liver tissue (“liver” in the article) were collected in three different tubes (Fig. 1**, panel A).**

LMD was used as a sampling method for this investigation as it appeared to be the best one for the accurate selection of the different parts composing the tissue. The tissues were processed following a procedure that avoids tissue loss during biochemical preparation [21]. This method was specifically designed for FFPE tissues. This relies on the biochemical processing including the proteolytic digestion of fixed proteins directly from the laser microdissected tissue piece, in a tube.

Peptides were subsequently separated by reverse-phase LC using ultra performance liquid chromatography (UPLC) 2D nanoAcquity (Waters, Corp., Milford, USA), and analyzed using a Thermo Scientific Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, USA) [21]. A first step of one dimension LC–MS/MS analysis with a 180 min run with 200 ms of accumulation of MS peaks for further MS/MS, and the selection of 10 of the major MS peaks for MS/MS, was performed to estimate the quantity of the peptide content in each sample. 2D LC–MS/MS analyses (RP x RP, high and low pH) included three isocratic elution steps on which 180 min LC gradient were performed, as described before [21]. The time for accumulation of MS peaks for further MS/MS was 200 ms 10 of the major MS peaks were selected for MS/MS.

2.3. Data processing

MaxQuant version 1.5.2.8 was used for protein identifications (IDs) from raw files. MS/MS spectra were analyzed to retrieve peptide sequences and to compare those to databanks, for identification. Andromeda search engine with the Uniprot database from *Homo sapiens* and Eg (released versions 2015-04) was used for interrogation. The following modifications were set for identification: N-ter acetylation, oxidation of methionines, deamidation as variable modifications, carbamidomethylation of the cysteines as fixed modification. The maximum number of missed cleavages was set at seven amino acids. At least two peptides per protein, including one unique peptide were required for identification.

A Venn diagram was made from the list of the proteins identified in each of the samples in order to differentiate the ones that are common between different ROIs and the ones that are only found in a specific ROI [23]. The Uniprot entry of the proteins identified in each compartment were listed in the “GeneVenn” website (http://genevenn.sourceforge.net/). Since three lists can be inserted, a Venn diagram containing the number of proteins commonly found in the different compartments of interest in the liver infected by Eg could be obtained. The Venn diagram was then manually reconstructed in order to adapt the size of the overlapping cells proportionally to the number of proteins.

Gene ontology (GO) annotation of the proteins specifically found in the different tissue parts was performed using Blast2GO 3.2. For each compartment of the infected liver tissue, a FASTA file listing the identified proteins was created. Each of the sequences was blasted against sequences of the Swissprot database, using the National Center for Biotechnology Information (NCBI) Blast service. The blasted sequences were then mapped in order to link the protein IDs to the GO database. The GO database contains millions of functionally annotated gene products for hundreds of species. The molecular function, biological process and cellular component were retrieved from the GO database. Level 3 GO were selected for representation. The number of sequences that were assigned to each GO was reported, and the nine most represented GO were used to create a histogram.

Since two species were checked for the analysis, no filter for taxonomy was applied.

2.4. IMS

FFPE tissue section was subjected to *in situ* trypsin digestion, sprayed with α-Cyano-4-hydroxycinnamic acid matrix for IMS analysis as described before [22,24] and analyzed using a rapifleX MALDI Tissuetyper mass spectrometer (Bruker, Bremen, Germany), with data acquired at 50 μm spatial resolution. IMS dataset was imported into SCiLS Lab (SCiLS GmbH, Bremen, Germany) for visualization of images and statistical analysis. 206 monoisotopic peaks from peptides were detected and used for the establishment of the peak list. Total ion count (TIC) normalization was used as well as baseline removal using Top Hat algorithm. Molecular images in Fig. 2 were represented with 50% transparency in order to show tissue regions from the optical image. Hierarchical clustering was chosen for the segmentation process to group the most similar spectra or sets of spectra into clusters.

The identification of the peptides mapped by IMS was made by comparing the masses of the monoisotopic peptides with the one obtained from the LC–MS data. As mainly monocharged peptides were obtained by IMS, the LC–MS values corresponding to [M + H] were used for correlation with IMS values. “M” corresponds to the mass of the peptides reported in the peptide identification table