Missing proteins in Chromosome 16 - Spanish HPP

ProteoRed Carlos III Networked Proteomics Platform

SpHPP Spanish Human Proteome Project

Manuel Fuentes[1], Noelia Dasilva[1], Felipe Clemente[2], María Luisa Hernáez[2], Paula Díez[1], Maria Gonzalez-Gonzalez[1], Alberto Orfao[1], Felix Elortza[5], Fernando Corrales[4], Juan Pablo Albar[3], Concha Gil[2].

[1]Cancer Research Center. University of Salamanca-CSIC, IBSAL. Campus Miguel de Unamuno s/n. 37007 Salamanca. Spain. [2]ProteoRed-ISCIII. Dpt. Microbiology & Proteomics Unit. University Complutense. Madrid. Spain. [3] ProteoRed-ISCIII. National Center of Biotechnology. CSIC. Madrid. Spain. [4]ProteoRed-ISCIII. Center for Applied Medical Research (CIMA), Spain. [5]ProteoRed-ISCIII. CIC bioGUNE. Bilbao. Spain.

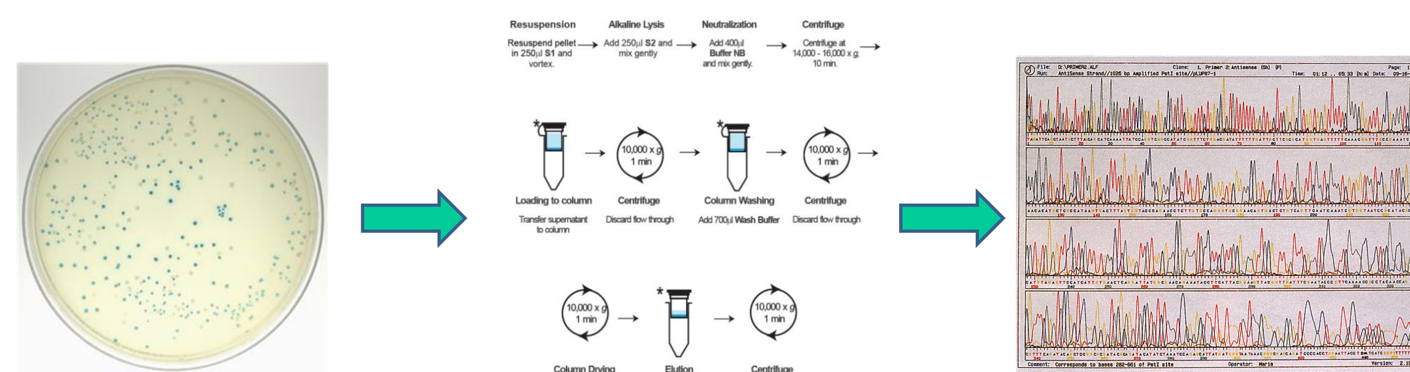
AIM

In the scope of the HPP project, there is a special situation for the proteins that had not spectral and/or expression evidence, those are called "Unknown proteins".These proteins must be studied specifically to get enough information to build MRM quantitation methods. The Spanish HPP consortium has developed a specific protocol for unknown proteins in order to get MSMS information for the MRM method.

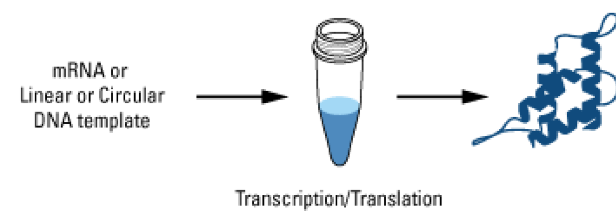
MATERIALS AND METHODS

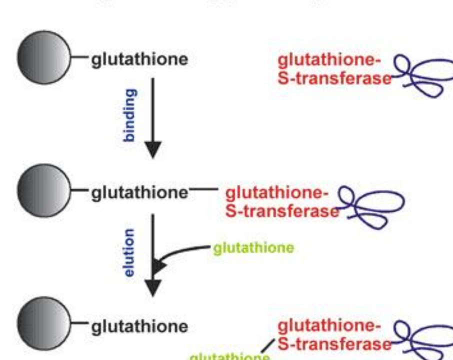
Clones and plasmids

All the clones used are from pANT7_cGST clone collection distributed by Plasmid repository at Arizona State University Biodesign Institute. Each clone contains an in-frame fused C-terminal GST tag. Each bacterial clone was grown overnight in 5 mL of Luria broth with 100 µg/mL ampicillin. Plasmid DNA was extracted using Mini-Prep Kit from Promega, following manufacture instructions. All plasmids were sequence to confirm the identity of the insert.



DNA isolation and sequence verification





Protein translation and enrichment

Protein IVVT production

Proteins were synthesized from plasmid DNA using Human In vitro Protein Expression kit (Thermo) following manufacture s protocol with a few minimal modifications in order to adapt for 1.5 mL eppendorf tubes.

Protein Capture Conditions

To enrich the GST-fusion proteins, 125 µL of glutathione-sepharose 4B (GE Healthcare) was added to each reaction tube and was rocked end-over-end for 16 h at 4ºC. After that, binded protein were washed 3 times with DPBS and two times with 50 mM ammonium bicarbonate.

Protein Digestion

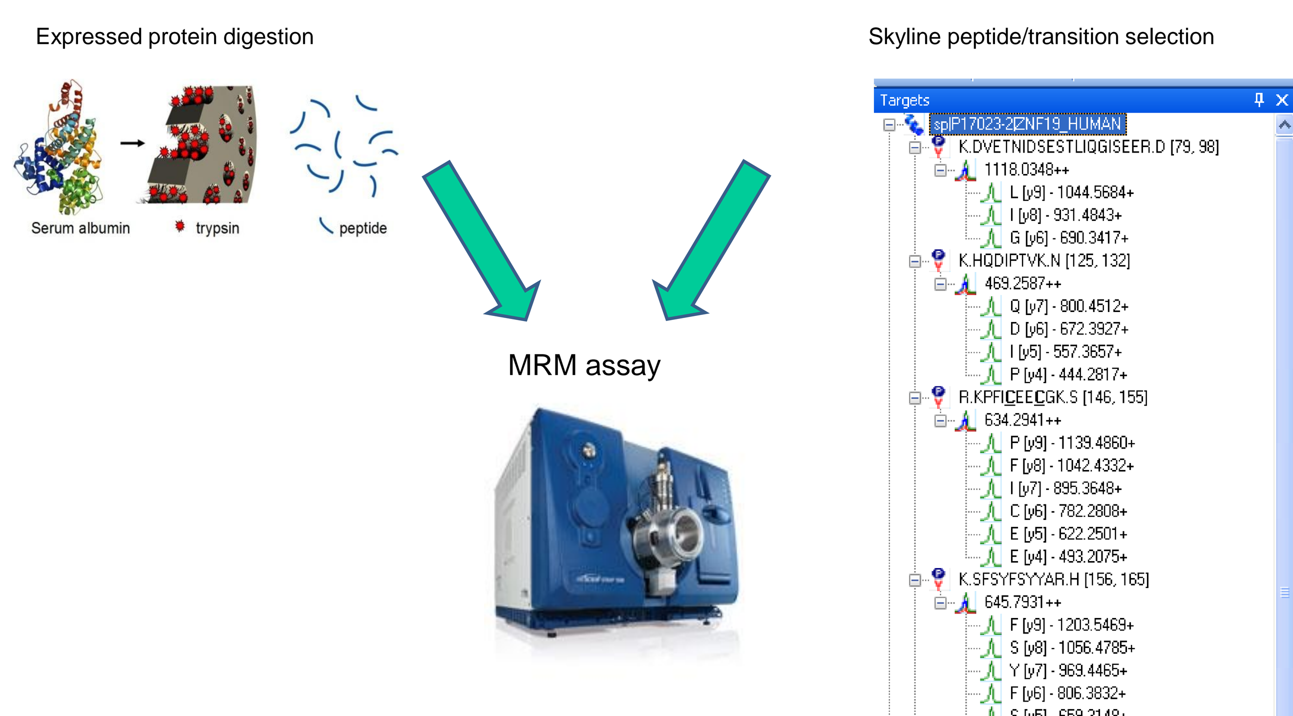
Proteins were suspended in 50 ul 50 mM ammonium bicarbonate. The digestion in solution was performed as reported before (Monteoliva et al., 2011). Proteins were reduced with 5 mM DTT and alkylated with 55 mM iodoacetamide. 0,4 micrograms of trypsin (proteomics grade; Roche Applied Science) were added to the samples, and incubated overnight at 37 C. After digestion, the samples were desalted with POROS R2 resin (ABSciex, Framingham, MA).

MRM methods

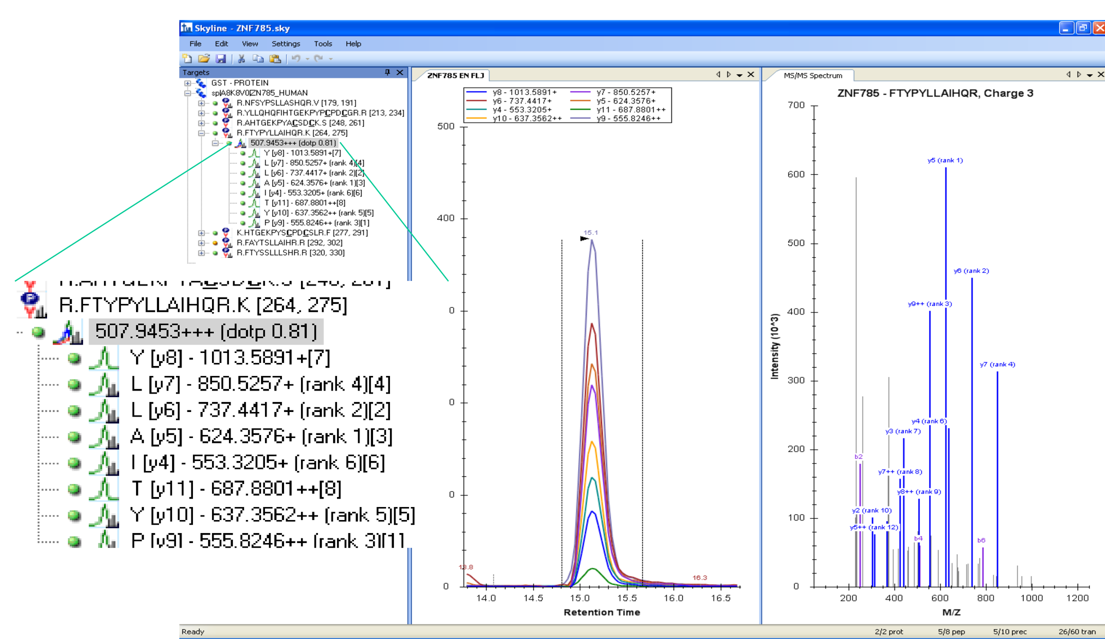
The MRM transitions for each protein were selected using the Skyline software (MacCoss Lab, University of Washington). When possible, peptides were choosen among that that matched the settings: no trypsin misscleavages, 7-25 aminoacids in length, excluding that containing Met, Trp and Cys. Transitions between 300m/z and 1250 m/z for each peptide were selected among the *y* fragments of 2+ and 3+ precursors. MRM transition list for 5500 QTRAP were generated by default and used without any modification.

LCMS methods

The digests were analyzed using an Eksigent Nanoflow LC system connected to the nanoSpray source of a 5500 QTRAP. 5 microliters of the sample was directly loaded onto a Eksigent trap column (NanoLC Trap Set ChromXP C18-3um, 120 A, 350 um x 0.5 mm) at 3 ul/min. Following 5 min of loading, the samples were fractionated onto a 75 µm 15 cm ChromXP nano LC column (Eksigent) at 300 nL/min. Solvent A was 100% water 0.1% formic acid. Solvent B was 100% acetonitrile 0.1% formic acid. The gradient used was 5-65% B in 30 min. The 5500 QTRAP was operated in MRM mode. The MRM transition intensities were used to trigger dependent enhanced product ion (EPI) scans.



Protein digestion, MRM method generation and MRM assay



Transitions matching and verification by MS/MS spectrum

Results Analysis

Wiff files generated by the 5500 QTRAP were analyzed in Skyline software to confirm the peptides and transitions. The same files were analyzed in Protein Pilot software (AB Sciex) to get the group archives used by skyline to generate the spectral libraries. The peptides identified by MS/MS spectra were annotated as "identified peptide". MRM methods were polished with the spectral and transition data.

RESULTS

All the peptides selected for MRM were tested to be proteotypic against the in house database used. Four proteins assayed for IVTT protein expression did not yield results in the MRM assay, we can explain it because of the insufficient DNA used for the IVTT (data not shown). The most relevant information about the MRM experiments are collected in the Table 1. With the MS/MS spectra obtained in the MIDAS assay we confirm the peptide sequence ("Identified peptides" column).

Swissprot ID	Protein	Skyline selected peptides	Found peptides	number of 2+ precursors	number of 3+ precursors	Total found precursors	Total Transitions	Identified peptides
Q8WV35	LRRC29 Leucine-rich repeat-containing protein 29	9	9	7	4	11	53	9
Q8TEW6	DOK4 Docking protein 4	11	9	6	4	10	44	8
O60359	CACNG3 Voltage-dependent calcium channel gamma-3 subunit	12	9	7	4	11	43	8
Q96H86	ZNF764 Zinc finger protein 764	12	6	4	3	7	22	7
Q6UXU4	GSG1L Germ cell-specific gene 1-like protein	7	5	4	4	8	33	7
A8K8V0	ZNF785 Zinc finger protein 785	10	7	3	6	9	58	7
Q96A59	MARVELD3 MARVEL domain-containing protein 3	6	6	4	4	8	34	6
Q8IUW3	SPATA2L Spermatogenesis-associated protein 2-like protein	9	9	7	5	12	64	6
Q6PII5	HAGHL Hydroxyacylglutathione hydrolase-like protein	10	7	5	3	8	43	6
Q9HBE5	IL21R Interleukin-21 receptor	12	8	7	1	8	36	5
Q9NXF8	ZDHC7 Palmitoyltransferase ZDHHC7	7	2	0	2	2	10	2
Q9NWW0	HCFC1R1 Host cell factor C1 regulator 1	7	4	4	0	4	17	2
Q96B96	TMEM159 Promethin	3	3	3	1	4	28	2
Q8WTQ4	C16ORF78 Uncharacterized protein C16orf78	10	8	5	3	8	40	2
Q8TDN1	KCNG4 Potassium voltage-gated channel subfamily G member 4	11	3	3	0	3	12	2
Q8TB05	FAM100A UBA-like domain-containing protein 1	1	1	1	1	2	8	2
O75324	SNN Stannin	3	2	2	1	3	28	2
Q8TAZ6	CKLFSF2 CKLF-like MARVEL transmembrane domain-containing protein 2	6	2	1	2	3	12	1
Q8N635	C16ORF73 Meiosis-specific with OB domain-containing protein	14	4	4	1	5	26	1
Q8IZF4	GPR114 Probable G-protein coupled receptor 114	11	3	3	1	4	18	1
Q9BTX3	TMEM208 Transmembrane protein 208	3	1	1	0	1	5	0
Q8WVE7	TMEM170A Transmembrane protein 170A	6	1	1	0	1	6	0
Q6PL45	C16ORF79 BRICHOS domain-containing protein 5	7	5	3	2	5	22	0
P17023	ZNF19 Zinc finger protein 19	15	3	2	1	3	11	0

Table 1. MRM results for 24 unknown proteins assayed. Identified peptides is the sume of 2+ and 3+ precursors with Paragon identified spectra.

CONCLUSIONS

Twenty four out of 28 proteins assayed yield results in the MRM experiment. Twenty out of 24 were detected with at least, 3 ion precursors, and 20 out of 24 were confirmed with at least one MS/MS spectrum.

The approach used to study the unknown proteins allows us to obtain the MRM data (peptides and transitions) necessary for the detection, identification and quantification of these proteins in complex mixtures, and can be applied to any set of proteins.