Biomarker discovery strategy for trisomy 21 using iTRAQ and 4800 plus MALDI TOF/TOF

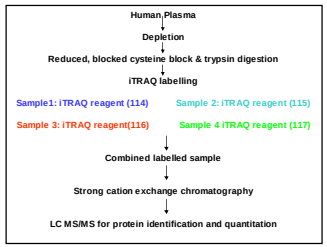
Chinnapapagari Satheesh Kumar Reddy, Varaprasad Kolla, Wolfgang Holzgreve & Sinuhe Hahn
Laboratory for Prenatal Medicine and Gynaecological Oncology,University Women's Hospital Basel/Department of Research, Basel. Switzerland

Introduction

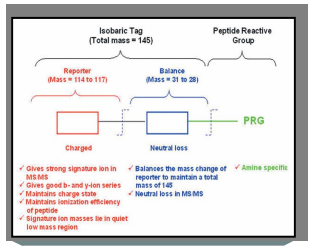
Plasma proteins serve as good indicators of disease as there are representative proteins from several cellular processes and thus a potential source for biomarker discovery. The large dynamic range of plasma proteins makes the analysis very challenging, as a large number of low abundance proteins are masked by a few high abundance proteins. Down's syndrome or Trisomy 21 is the most common genetic cause of human mental retardation, affecting approximately 1 in 700 live births. The presence of the chromosome 21 in triplicate should increase the amount of cell proteins coded by genes located on this chromosome and many other proteins by regulator genes, present in a three fold dosage, and by several indirect effects of trisomy. A series of biochemical findings, including deteriorated glucose, lipid, purine, methionine/homocysteine and folate metabolism have been reported but no systematic study on protein profiling in Normal or Trisomy 21 has been conducted so far. In plasma proteomics non-gel based method has a great advantage over the traditional gel based method. We hereby, report an efficient strategy for biomarker discovery in human plasma using iTRAQ labelling in conjunction with 4800 MALDI TOF/TOF. This approach can assist in the detection of low abundance proteins and the discovery of informative biomarkers

Method

Albumin and other immunoglobulin constitute almost 99% of plasma proteins. These protein were depleted from two trisomy (n=2) and two gestational age matched controls (n=2), using ProteoPrep 20 Plasma Immunodepletion Kit from Sigma. The depleted plasma was then digested with trypsin (ABI) and was label with the single iTRAQ labels (ABI) (114,115,116,117). The labelled sample were pooled and were subjected to a strong cationic exchange (SCX) afterwards the fraction were desalted and loaded on ABI 4800 plus MALDI TOF/TOF to obtain the mass spectra

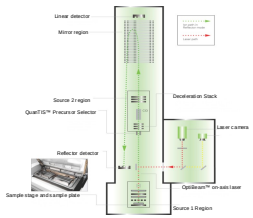


Work flow for non-gel based method



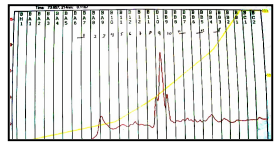
iTRAQ chemistry

4800 MALDI TOF/TOF

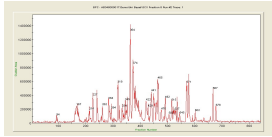


Result & Discussion

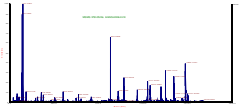
The 'dynamic range' in the plasma is a very promising entity, but with traditional 2D-gel electrophoresis there are many limitations when attempting to visualize low abundant proteins



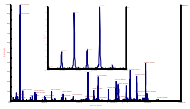
SCX fraction



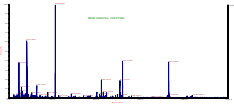
SCX Fraction 9 Base Peak Chromatogram



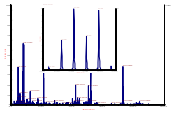
MS/MS of Pregnancy zone protein iTRAQ ratio 115/114: 1.88



MS/MS of Pregnancy zone protein iTRAQ ratio 115/114: 1.88



MS/MS of C4b-binding protein (C4bp) iTRAQ ratio 115/114: 1.54



MS/MS of C4b-binding protein (C4bp) iTRAQ ratio 115/114: 1.54

We got 4070 peaks from SCX fraction number two and three, 2185 peaks from SCX fraction eight and 3682 peaks from SCX fraction nine. Using two different databases (Swissprot and NCBI) 248 and 235 proteins were identified respectively with a minimum confidence of 95%.

N	Acc.No	Name	115:114	116:114	117:114
1	P02751	Fibronectin	1.1971	1.0035	1.2867
2	P08603	Complement factor H	1.1124	0.9757	1.1437
3	P02790	Hemopexin	1.1332	1.1216	1.1219
4	Q14624	Inter-alpha-trypsin	1.1184	0.976	1.1728
5	P01031	Complement C5	1.1635	1.0105	1.259
6	P00751	Complement factor B	1.1672	1.0022	1.1785
7	P00734	Prothrombin	1.1225	0.9627	1.1739
8	P20742	Pregnancy zone protein	1.8957	1.0317	1.9356
9	P04217	Alpha-1B-glycoprotein	1.161	1.1333	1.1332
10	P00738	Haptoglobin	1.1082	0.956	1.1823
11	P07358	Complement C8	1.1493	0.9992	1.2199
12	P04003	C4b-binding protein	1.5373	1.0679	1.4962
13	P02743	Serum amyloid P	1.1788	0.9445	1.3181
14	P35858	Insulin-like growth factor	1.2439	1.0304	1.0777
15	P01834	Ig kappa chain	1.3518	1.0287	1.2934

Up regulated proteins

N	Acc.No	Name	115:114	116:114	117:114
1	P04114	Apolipoprotein B-100	0.8236	0.9488	0.851
2	P02768	Serum albumin	0.6172	0.9787	0.6365
3	P02647	Apolipoprotein A-I	0.7255	0.9751	0.7059
4	|000137	Cationic trypsin	0.8346	1.4225	0.8646
5	P06396	Gelsolin	0.8919	1.0167	0.9032
6	P01023	Alpha-2-macroglobulin	0.7287	1.0217	0.7119
7	P02787	Serotransferrin	0.7467	1	0.7223
8	P06727	Apolipoprotein A-IV	0.6879	0.9315	0.7112
9	P43652	Afamin	0.886	0.956	0.9324
10	P02749	Beta-2-glycoprotein 1	0.8925	0.9815	0.9809
11	P25311	Zinc-alpha-2-glycoprotein	0.8746	0.8962	0.9294
12	P01024	Complement C3	0.8093	0.9994	0.8159
13	P02671	Fibrinogen alpha chain	0.8086	1.007	0.8514
14	P02649	Apolipoprotein E	0.7412	0.8196	0.8843
15	P35908	Keratin, type II	0.8038	1.5298	0.5923

Down regulated proteins

Conclusion

- Simple workflow labels peptides allowing rapid progression to LC/MS/MS analysis and easy data interpretation for relative and absolute quantitation
- Expand proteome coverage by labeling all peptides to extract more detailed information from samples
- Analysis of up to four different biological samples simultaneously in a single experiment
- Increase confidence in identification and quantitation by tagging multiple peptides per protein to gain more statistically significant information
- Absolute quantitation across numerous sample states
- Enhance low-level analysis as a result of the signal amplification from the additive fragmentation of labeled isobaric peptides