Recognition of Distinct Protein Isoforms and Isoform-Specific Quantitation

Poster ThP 540

Wilfred H. Tang, Sean L. Seymour, Alex Loboda, Christie L. Hunter, Daniel A. Schaeffer, Applied Biosystems|MDS Sciex, 850 Lincoln Centre Dr., Foster City, CA 94404

ABSTRACT

We present here software which intelligently performs protein assembly/grouping and relative protein expression for samples containing related protein isoforms.

INTRODUCTION

In any biological sample, many proteins are actually part of protein families consisting of several variants or isoforms. The presence of multiple closely-related isoforms in a sample complicates the interpretation of results of "bottom-up" experiments because some digest peptides are common to multiple isoforms. The protein assembly/grouping problem is to determine the list of proteins actually present in the sample given a list of identified peptides with confidences. This issue is ignored or crudely handled by most existing software, which forces researchers into laborious manual inspection and analysis of the data when studying closely-related isoforms. Pro Group Software provides a solution to the protein grouping problem. In addition, solving this problem enables Pro Group Software to calculate more accurate isoform-specific protein expression ratios.

MATERIALS AND METHODS

Samples and Mass Spectrometry (Liver Microsome): Proteins from mouse liver microsome (phenobarbital-treated and untreated control) were digested and labeled with iTRAQ™ reagents according to standard protocol. The sample was then fractionated by strong cation exchange chromatography, and each fraction was analyzed by LC/MS on a 3200 Q TRAP® System (hybrid triple quadrupole linear ion trap). To confirm quantitation accuracy, expression changes for selected cytochrome P450 proteins were also measured using highly sensitive targeted Multiple Reaction Monitoring (MRM) absolute quantitation experiments using isotope-labeled custom-synthesized peptides.[1,2]

Samples and Mass Spectrometry (Serum): Serum from four different individuals were partially depleted of high abundance proteins (albumin, immunoglobulin), digested with trypsin, and labeled with iTRAQ™ reagents according to standard protocol. The sample was then fractionated by strong cation exchange chromatography, and each fraction was analyzed by LC/MS on a QSTAR® System.

Software: Peptides were identified from MS/MS spectra by database search using a variety of search engines, commercial (Pro QUANT) as well as research-grade.

Pro Group Software solves the protein grouping problem by determining the minimal number of proteins truly required to explain the identified peptides. Multiple protein isoforms are reported only if there is sufficient distinct peptide evidence. This approach leverages peptide confidences to detect true isoforms while preventing the false detection of isoforms based on weak evidence.

The quantitation portion of Pro Group Software calculates a relative expression ratio for each protein based on measured relative peptide expression ratios. Only the peptides unique to a single isoform are used to calculate that isoform's expression ratio. Peptides common to multiple isoforms almost certainly do not provide an accurate estimate of a single isoform's expression ratio.

RESULTS – MOUSE LIVER MICROSOMES

The mouse liver microsome samples are rich in cytochrome P450 proteins. This poses a special challenge to protein grouping and quantitation software due to the high degree of sequence similarity between different cytochrome P450 isoforms. The cytochrome P450 proteins are of great interest to the pharmaceutical industry due to their major role in the metabolism of drugs and other foreign chemicals.

The iTRAQ reagent LC/MS experiment + Pro Group Software yielded 248 proteins found with ProtScore above 1.3 (corresponding to a 95% confidence threshold). 20 distinct cytochrome P450 isoforms were identified. 9 of those cytochrome P450 isoforms (together with 3 other control proteins) were also quantitated by targeted MRM experiments. The graph to the right shows that there is good agreement between the two quantitation methods.

Figure 1. Graph Comparing Quantitation by iTRAQ Reagent LC/MS + Pro Group Software Vs. Quantitation by Targeted MRM

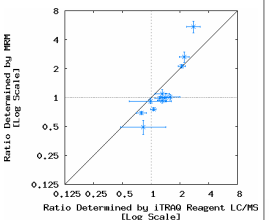
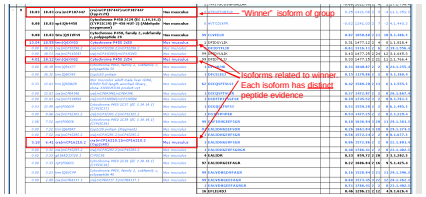


Table 1. Comparison of Quantitation by iTRAQ Reagent LC/MS + Pro Group Software Vs. Quantitation by Targeted MRM

Protein Isoform	Drug-Induced:Control Ratio (Determined by iTRAQ Reagent LC/MS + Pro Group Software)	Drug-Induced:Control Ratio (Determined by Targeted MRM)
Cytochrome P450 2a12	1.30 (95% CI: 1.17 – 1.46)	0.93 (95% CI: 0.88 – 0.99)
Cytochrome P450 2b10	2.78 (95% CI: 2.37 – 3.26)	5.42 (95% CI: 4.71 – 6.23)
Cytochrome P450 2c29	2.23 (95% CI: 1.98 – 2.50)	2.63 (95% CI: 2.32 – 2.97)
Cytochrome P450 2c40	0.98 (95% CI: 0.59 – 1.64)	0.92 (95% CI: 0.86 – 0.98)
Cytochrome P450 2d9	0.83 (95% CI: 0.48 – 1.42)	0.49 (95% CI: 0.41 – 0.58)
Cytochrome P450 2e1	1.06 (95% CI: 0.99 – 1.13)	0.76 (95% CI: 0.73 – 0.79)
Cytochrome P450 2f2	0.79 (95% CI: 0.70 – 0.88)	0.69 (95% CI: 0.66 – 0.72)
Cytochrome P450 3a11	2.09 (95% CI: 1.91 – 2.29)	2.12 (95% CI: 2.02 – 2.22)
Cytochrome P450 4a10	1.58 (95% CI: 1.26 – 1.99)	1.01 (95% CI: 0.94 – 1.08)
Corticosteroid beta-dehydrogenase	1.22 (95% CI: 1.09 – 1.38)	0.98 (95% CI: 0.93 – 1.03)
Triglyceride transfer protein	1.43 (95% CI: 1.28 – 1.60)	1.00 (95% CI: 0.90 – 1.11)
Microsomal GST	1.30 (95% CI: 1.09 – 1.55)	1.10 (95% CI: 1.00 – 1.21)

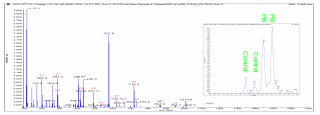
The cytochrome P450 2c subfamily exhibits particularly interesting behavior. Phenobarbital dosing leads to significantly increased expression of cytochrome P450 2c29 but has little or no effect on the expression level of cytochrome P450 2c40 despite the high sequence similarity of the two isoforms (68% sequence identity).

Figure 2. Pro Group Report for Protein Group #9 (Cytochrome P450 2c Proteins)



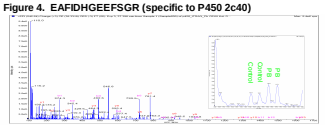
In the Pro Group Report, we see that the iTRAQ reagent LC/MS experiment found distinct peptide evidence for 4 different cytochrome P450 2c isoforms. In calculating expression ratios, only the peptides specific to a single isoform are used.

Figure 3. GSFPMAEK (specific to P450 2c29)



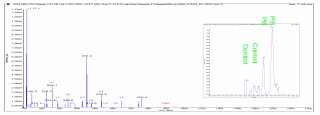
The peptide GSFPMAEK is specific to the cytochrome P450 2c29 isoform. The 114 and 115 peaks correspond to control mice, while the 116 and 117 peaks correspond to phenobarbital-dosed mice. The expression level of this peptide is significantly up-regulated in the phenobarbital-dosed mice compared to the control mice.

Figure 4. EAFIDHGEEFSGR (specific to P450 2c40)



The peptide EAFIDHGEEFSGR is specific to the cytochrome P450 2c40 isoform. There is little difference in the expression level of this peptide in phenobarbital-dosed vs. control mice.

Figure 5. NLGMGK (common to P450 2c29 and P450 2c40)



The peptide NLGMGK is common to both the cytochrome P450 2c29 and the cytochrome P450 2c40 isoforms. Most likely, cytochrome P450 2c29 is more abundant than cytochrome P450 2c40 (more peptides identified, higher ProtScore). Using NLGMGK in calculating expression ratios would greatly skew the results for cytochrome P450 2c40 and would skew to a lesser extent the results for cytochrome P450 2c29.

RESULTS – SERUM

For the serum experiment, the immunoglobulins offer a challenge to protein grouping software due to the high sequence similarity among these proteins. For the three immunoglobulin G (IgG) isoforms discussed below, there is 88% sequence identity among the three isoforms.

Figure 6. Pro Group Report for Protein Group #17 (IgG Proteins)

Table 2. Relative IgG Protein Expression Ratios for Four Individuals

Protein	Individual #2: Individual #1	Individual #3: Individual #1	Individual #4: Individual #1
IgG1	0.64 (95% CI: 0.57 – 0.71)	0.12 (95% CI: 0.07 – 0.21)	0.18 (95% CI: 0.09 – 0.37)
IgG2	0.85 (95% CI: 0.43 – 1.67)	0.18 (95% CI: 0.05 – 0.60)	0.25 (95% CI: 0.08 – 0.76)
IgG4	0.99 (95% CI: 0.62 – 1.58)	0.98 (95% CI: 0.75 – 1.20)	1.00 (95% CI: 0.60 – 1.67)

IgG1 and IgG2 are highly down-regulated in individuals #3 and #4 (116 and 117 iTRAQ reagent labeled) compared to individuals #1 and #2 (114 and 115 iTRAQ reagent labeled). Interestingly, the highly homologous IgG4 protein does not exhibit such high variability in expression pattern among the four individuals. To confirm these results, it may be advisable to repeat this experiment without immunoglobulin depletion during sample handling. Nevertheless, this experiment demonstrates the ability of Pro Group Software to detect differential regulation of very closely-related protein isoforms.

CONCLUSIONS

Homologous proteins pose a special challenge to protein grouping software. The Pro Group Software approach to solving this problem is to require sufficient distinct peptide evidence in order to declare the presence of multiple protein isoforms. We have demonstrated here Pro Group Software's ability to declare related isoforms when appropriate and also to accurately quantitate these isoforms.

REFERENCES

1. Hunter, C. L.; Webb, S.; DuPont, D.; Pillai, S.; Hunt, T.;Kitteringham, N.; Jenkins, R.; Pennington, S. *Proceedings of the 52nd ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, TN, May 23 - 27, 2004.
2. Hunter, C. L.; Webb, S.; Settineri, T. A.; Tang, W. H.; Kitteringham, N.; Jenkins, R.; Pennington, S. *53rd ASMS Conference on Mass Spectrometry and Allied Topics*, San Antonio, TX, June 5 - 9, 2005, Poster MP33.

TRADEMARKS/LICENSING