Serum proteomics profiling—a young technology begins to mature

Kevin R Coombes, Jeffrey S Morris, Jianhua Hu, Sarah R Edmonson & Keith A Baggerly

During the three years since the US National Cancer Institute–Food and Drug Administration (NCI-FDA) proteomics group published their seminal[1] (but flawed[2–4]) study using mass spectrometry to profile the serum proteome of ovarian cancer patients, more than 60 published studies have applied similar technology to a wide range of cancers and other diseases. Although some studies have looked directly at tumor tissue, most have targeted easily accessible fluids such as serum. A few studies used high-resolution instruments with a matrix-assisted laser desorption and ionization (MALDI) ion source and a time-of-flight (TOF) ion detector. The majority, however, have relied on lower resolution instruments and surface-enhanced laser desorption and ionization (SELDI), a MALDI variant that uses commercially prepared chromatographic surfaces (ProteinChips; Ciphergen, Fremont, CA, USA) to separate proteins.

These studies have a common goal: to identify biomarker patterns in the proteome that can be used for diagnosis, prognosis or monitoring of disease. The preponderance of evidence from these studies seems to suggest that proteomic profiling is able to detect anonymous protein peaks that are expressed differently in cancer patients than in healthy individuals[5–7]. The studies also provide evidence that we may be able to discover patterns that are useful for prediction of the presence or absence of disease[8–11]. For serum proteomics to realize its full potential, however, several problems in sensitivity and reproducibility remain unresolved. This article outlines some possible improvements in experimental design and data analysis that might help in this regard.

Kevin R. Coombes, Jeffrey S. Morris, Jianhua Hu and Keith A. Baggerly are at the Section of Bioinformatics, Department of Biostatistics and Applied Mathematics, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Box 447, Houston, Texas 77030, USA and Sarah R. Edmonson is in the Department of Family and Community Medicine, Baylor College of Medicine, Houston, Texas 77098, USA. e-mail: kcoombes@mail.mdanderson.org

Potential problems

It is important to recognize the provisional nature of the conclusions of most serum proteomic studies carried out to date. In most cases where the protein peaks have been identified, they have turned out to be well-known, acute-phase proteins. Critics of this application of mass spectrometry have claimed that it is inherently limited in its depth of coverage, with a dynamic range that prevents detection of low-abundance proteins[12,13]. As for predictive patterns, no two independent studies have found the same patterns to date. Moreover, none of the patterns has been validated in an independent study by another laboratory. In one study, both the sensitivity and the specificity of the pattern declined significantly when samples were processed in the same laboratory after a delay of several months[14].

Our experience at the M.D. Anderson Cancer Center suggests that it can be difficult to obtain stable, reproducible mass spectrometry results over time and across laboratories. In one unsuccessful study, SELDI experiments were performed to analyze the serum proteome from 247 patients with five different subtypes of cancer, with 40, 60, 65, 62 and 20 patients with each subtype. Simple hierarchical clustering (average clustering with distance based on Pearson correlation) of the data produced six clusters. Unfortunately, the clusters matched the run dates of the samples, not the biological subtypes of cancer[15]. The same clustering was observed using repeated runs of a quality control sample, which was run concurrently with the experimental samples. Because some of the biological subtypes were confounded with the run order, it was impossible in this data set to separate technological effects from biological contrasts of interest. Similar problems appear to have affected the NCI-FDA studies[4,16].

The seminal NCI-FDA study used mass spectrometry to analyze the serum proteome of patients with ovarian cancer, cells of which are shown above in a scanning electron micrograph (magnification: 414×).

In another MALDI study of the serum proteome of 50 cancer patients carried out by our groups at M.D. Anderson[15], hierarchical clustering produced two clusters. In this case, the sample collection protocol had been changed after the first 20 samples had been collected; the clusters matched the sample collection protocols. The change in protocols resulted in reproducible changes in the serum proteome, and these changes were significantly larger than those resulting from different histological subtypes of cancer[15].

The source of these problems appears to be inextricably linked to the promise of the technology. Mass spectrometry can be exquisitely sensitive to changes in the part of the proteome that it can measure accurately. Differences in sample collection or sample handling affect the proteome, to a degree that can dominate biological changes. In addition, as with many sensitive instruments, mass spectrometers can be temperamental. They provide the most accurate and highest resolution results over a very small

subset of their mass range, where they have been carefully focused. Relatively small changes to the operating conditions can be amplified to produce fairly large differences in mass spectra, making it difficult to maintain consistent, reproducible results.

Potential solutions

Reproducibility can be enhanced in several ways: by improving instrumentation, experimental protocols, study design or analysis tools. Although improved instruments might help, the studies conducted to date using higher resolution MALDI instruments have not convincingly demonstrated that they produce more robust proteomics patterns than the SELDI approach. This may change as the technology evolves. Performing peptide profiling instead of protein profiling is promising because it concentrates all measurements in a small region of extremely high mass resolution[17,18]. In addition, high-end Fourier transform mass spectrometry is being evaluated for its reproducibility in serum proteomic profiling[19].

The Early Detection Research Network (EDRN) is currently conducting a multi-institutional study to develop and validate protocols to produce reproducible mass spectra[20]. Their initial report computed the coefficient of variation in mass location, signal-to-noise ratio and normalized intensity for three protein peaks, concluding that between-laboratory reproducibility was as good as within-laboratory reproducibility[21]. It is still unclear, however, if reproducibility of the location and intensity of a few peaks is enough to ensure the reproducibility of the entire spectrum. It is possible that more global measures of reproducibility (that take into account all or most peaks) will be required.

The need for better study design and experimental design for proteomic profiling has been discussed elsewhere[15,22]. Epidemiologists and statisticians have known the necessary design principles for a long time. It often seems, however, that these principles are forgotten in the excitement that accompanies the first applications of a new technology. The spectrum of subjects in the study, both with and without disease, must reflect the actual population. Validation data sets must be independent and large. Care must be taken to avoid confounding biological variables of interest (e.g., presence or absence of disease) with technological factors (e.g., sample collection or run order) that might bias the results. Appropriate measures to minimize confounding and bias include running blinded tests, randomization and blocking.

The current state of the art in the analysis of serum proteome profiling experiments allows considerable room for improvement. At one extreme, the NCI-FDA proteomics group performs minimal preprocessing, and they allow the intensity at every measured mass-to-charge (m/z) value to be a potential feature for distinguishing cancer patients from healthy controls[1,23,24]. Because their experimental design confounded biology with technology[2–4,16], their approach is not yet firmly established. At the other extreme, a majority of published studies perform preprocessing and peak detection using software from Ciphergen, the manufacturer of the SELDI instrument[25]. In our opinion, the Ciphergen software is overly conservative when peak finding, and its baseline correction algorithm produces biased estimates of peak heights. These algorithmic weaknesses can reduce the effective sensitivity of the instrument below its true capabilities and can hamper its reproducibility.

Researchers have been actively developing better methods for processing mass spectra. Yasui *et al.*[26,27] have used a super smoother to find peaks, but had to settle for binary indicators of presence or absence in lieu of quantification. Our group[28] introduced an iterative method to perform simultaneous background correction and peak finding, but this method required manual adjustment of numerous parameters. Qu *et al.*[29] have applied wavelets for data reduction. Because the wavelet coefficients do not directly correspond to physical quantities, however, they are difficult to interpret. We also used wavelets for noise reduction[30], especially in concert with the use of the mean spectrum to borrow strength across spectra[31]. Our studies benefited from a computer model based on the physics of a mass spectrometer, which used a simple model of the detector[32]. In another approach, Sauve and Speed[33] have used dynamic programming to improve calibration across multiple spectra and morphological filters for baseline correction.

Elsewhere, Malyarenko *et al.*[34] have approached mass spectra from the vantage of time series. They obtained better estimates of baseline by realizing that its primary source resided in the physics of the ion detector. They also used deconvolution filters to smooth the signal and obtain better resolution of mass peaks. We expect that incorporating knowledge of the physical and chemical properties of mass spectrometry instruments will lead to further improvements in analytical methods.

Conclusions

Mass spectrometry profiling of the serum proteome remains an exciting tool with the potential to transform medicine. As often happens with a young technology, the spectacular claims of early studies have not held up under closer scrutiny. Nevertheless, numerous studies have provided evidence that this technology can be used to uncover proteomic expression patterns linked to a disease state. Careful studies are now underway to develop reproducible protocols. An increased awareness of the importance of sound experimental design is enhancing the ability of studies to yield reproducible results. Rapid improvements are being made in developing methods to analyze the data produced by mass spectrometry protein profiling experiments. The next step will be to show that these improved protocols, designs and analytical tools produce results that are sufficiently reliable and reproducible to change the way patients are treated.

1. Petricoin, E.F. *et al. Lancet* **359**, 572–577 (2002).
2. Sorace, J.M. & Zhan, M. *BMC Bioinformatics* **4**, 24 (2004).
3. Baggerly, K.A., *et al. Bioinformatics* **20**, 777–785 (2004).
4. Baggerly, K.A., *et al.* Signal in noise: evaluating reported reproducibility of serum proteomics tests for ovarian cancer. *J. Natl. Cancer Inst.*, **97**, in press (2005).
5. Vlahou, A. *et al. Clin. Chem.* **50**, 1438–1441 (2004).
6. Pusztai, L. *et al. Cancer* **100**, 1814–1822 (2004).
7. Lin, Z., Jenson, S.D., Lim, M.S. & Elenitoba-Johnson, K.S. *Mod. Pathol.* **17**, 670–678 (2004).
8. Zhang, Z. *et al. Cancer Res.* **64**, 5882–5890 (2004).
9. Wadsworth, J.T. *et al. Clin. Cancer Res.* **10**, 1625–1632 (2004).
10. Soltys, S.G. *et al. Clin. Cancer Res.* **10**, 4806–4812 (2004).
11. Xiao, Z. *et al. Cancer Res.* **64**, 2904–2909 (2004).
12. Diamandis, E.P. *Expert Rev. Mol. Diagn.* **4**, 575–577 (2004).
13. Diamandis, E.P. *Mol. Cell Proteomics* **3**, 367–378 (2004).
14. Rogers, M.A. *et al. Cancer Res.* **63**, 6971–6983 (2003).
15. Hu, J., *et al. Brief. Funct. Genomic. Proteomic.* **3**, 322–331 (2005).
16. Baggerly, K.A., *et al. Endocr. Relat. Cancer* **11**, 583–584 (2004).
17. Villanueva, J. *et al. Anal. Chem.* **76**, 1560–1570 (2004).
18. Koomen, J.M. *et al. Rapid Commun. Mass Spectrom.* **18**, 2537–1848 (2004).
19. Johnson, K.L., *et al. Anal. Chem.* **76**, 5097–5103 (2004).
20. Grizzle, W.E. *et al. Dis. Markers* **19**, 185–195 (2004).
21. Semmes, O.J. *et al. Clin. Chem.* **51**, 102–112 (2005).
22. Ransohoff, D.F. *Nat. Rev. Cancer* **4**, 309–314 (2004).
23. Petricoin, E.F. 3rd *et al. J. Natl. Cancer Inst.* **94**, 1576–1578 (2002).
24. Conrads, T.P. *et al. Endocr. Relat. Cancer* **11**, 163–178 (2004).
25. Fung, E.T. & Enderwick, C. *Biotechniques* **Suppl**, 34–41 (2002).
26. Yasui, Y. *et al. J. Biomed. Biotechnol.* **2003**, 242–248 (2003).
27. Yasui, Y. *et al. Biostatistics* **4**, 449–463 (2003).
28. Coombes, K.R. *et al. Clin. Chem.* **49**, 1615–1623 (2003).
29. Qu, Y. *et al. Biometrics* **59**, 143–151 (2003).
30. Coombes, K.R. *et al.* Improved peak detection and quantification of mass spectrometry data acquired from surface-enhanced laser desorption and ionization by denoising spectra with the undecimated discrete wavelet transform. *Proteomics*, in press (2005).
31. Morris, J.S., Coombes, K.R., Koomen, J., Baggerly, K.A. & Kobayashi, R. *Bioinformatics*, published online 26 January 2005.
32. Coombes, K.R., *et al.* Understanding the characteristics of mass spectrometry data through the use of simulation. *Cancer Informatics*, in press.
33. Sauve, A.C. & Speed, T.P. Normalization, baseline correction and alignment of high-throughput mass spectrometry data *Proceedings Gensips 2004*, in press (2005).
34. Malyarenko, D.I. *et al. Clin. Chem.* **51**, 65–74 (2005).