BIOENGINEERING

Multiplexed Protein Analysis

Yesim Gökmen-Polar and Sunil Badve*

Multiplexed protein analysis using an antibody-DNA barcoding approach could accelerate early detection and monitoring of cancer biomarkers in patient samples (Ullal *et al.*, this issue).

Genomic and transcriptomic high-throughput technologies permit detailed analysis of DNA and RNA molecules. These technologies have not only revealed the genetic diversity of cancer at the inter- and intratumor levels but also contributed to the documentation of prognostic and predictive utility of the identified gene signatures. The principal assumption in all of these studies is that alterations in nucleic acids (DNA and RNA) will be accompanied by alterations in the corresponding proteins. "Omics" studies have also documented that noncoding "junk" DNA is translated, and the RNAs generated from these have important roles in governing transcription and protein synthesis. A comprehensive understanding of the complexity of gene-protein relationships in disease states is necessary. In this issue of *Science Translational Medicine*, Ullal and colleagues report on a new technology that profiles the protein expression from liquid biopsies, such as fine-needle aspirates (FNAs), with single-cell sensitivity, to link the genotypic alterations with phenotypic changes and identify pathway-specific protein markers with clinical utility (*1*).

PROTEIN ANALYSIS AND MULTIPLEXING

Comprehensive analysis of proteins is seldom used in clinical practice. Targeted antibody-based methods, such as immunohistochemistry (IHC), are routinely applied to tissue samples for the diagnosis of disease and for prognostic and predictive purposes. Analysis of singleton proteins, one at a time, still remains the standard practice in clinical laboratories. The semiquantitative nature of IHC can be improved by use of fluorescent tags, which are commonly used in interphase fluorescent in situ hybridization (FISH) analyses of gene amplification and fusion events. However, the use of the multiplex fluorescence-based methods has been largely restricted to flow cytometric analysis of cell-surface markers for the diagnosis and classification of leukemias and lymphomas. Fluorescence resonance energy transfer (FRET) methods have been used to study protein-protein interactions, particularly in the analysis of human epidermal growth factor (HER) family proteins in breast cancer (*2*). Recent advances in digital microscopy and the ability to store and integrate the images can permit analysis of a number of markers from a single formalin-fixed, paraffin-embedded (FFPE) tissue section (*3*).

Most multiplex methods for protein analysis sacrifice cellular and subcellular localization in order to analyze multiple proteins at the same time. An overview of the currently available methods for multiplex protein profiling is demonstrated in Fig. 1. Cell lysates and biologic fluids are mostly analyzed with antibody-based method in which the reaction is visualized through colorimetric or fluorescent techniques. Variations that include immobilization of the substrate or target, or having the reaction occurring on the surface of beads, permit simultaneous identification of a large number of proteins (*4*). Arrays generated by spotting antibodies or cell lysates [such as reverse phase protein assay (RPPA)] have also been used for the analysis of large numbers of samples or proteins (*5*). Non–antibody-based high-throughput techniques include liquid chromatography–tandem mass spectrometry (LC-MS/MS) and selective or multiple reaction monitoring (SRM; MRM) (*4*). MRM is becoming more popular and has been applied to archival (FFPE) materials (*6*). Despite these advances, the accurate quantification of cellular protein expression is one of the most challenging areas in biotechnology and medicine. Although LC-MS/MS–based techniques have advanced substantially, they still have considerable technical limitations (*4*). Specifically, they are not ideal for detection of low-abundance proteins. Depletion of highly abundant proteins or sample purification can introduce biases in results owing to nonspecific binding.

SCANNING BARCODES

In this issue of *Science Translational Medicine*, Ullal *et al.* (*1*) describe a new approach to analyze the expression profiles of multiple proteins from limited amounts of FNAs

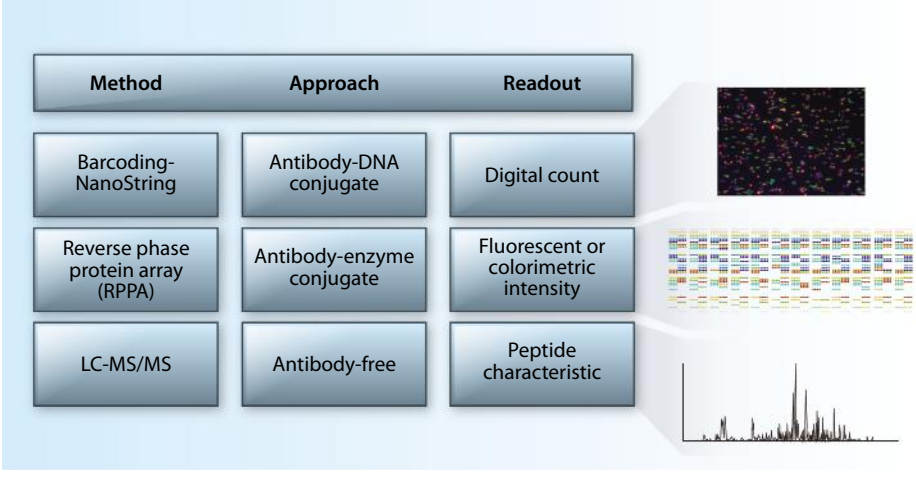

Fig. 1. Methods for multiplex protein profiling. The expression levels of multiple proteins can be measured by using different approaches. Barcoding-NanoString combines digital detection (NanoString's nCounter) with antibody-DNA conjugates. RPPA is a high-throughput antibody-based technique that uses colorimetric or fluorescent assay intensity. High-throughput antibody-free techniques consist of LC-MS/MS, which measures label-free peak peptide intensities, or stable-isotope labeling by tagging the mass of a protein or peptide. Recent modifications include SRM and MRM.

Department of Pathology and Laboratory Medicine, Indiana University School of Medicine, Indianapolis, IN 46202, USA.
*Corresponding author. E-mail: sbadve@iupui.edu

in patients with lung adenocarcinoma. The method couples the DNA barcode–antibody technology reported previously by Agasti *et al.* (*7*) with NanoString's nCounter system (*8*) to capture and count individual proteins in cell lines or clinical samples. Briefly, lysed cells from cancer cell lines or from cancer patients [positive for epithelial cell adhesion molecule (EpCAM) and cytokeratin (CK) and negative for CD45] were stained with a mixture of more than 90 pooled antibodies conjugated to small (~70 oligomer) distinct DNA barcodes via photocleavable linkers. Upon photocleavage at 365 nm, the DNA barcodes released from the cell lysates were processed using NanoString's fluorescent DNA barcoding platform. After hybridization to fluorescent barcodes, DNA was imaged on a cartridge using a digital analyzer for data collection (Fig. 1). Color codes on the surface of the cartridge were counted and converted into a proteomic profile by normalizing to DNA per antibody and housekeeping proteins.

Using their new method, the authors documented the ability to measure intra- and intertumoral heterogeneity of cancer cells at single-cell sensitivity (*1*). For six patients with biopsy-proven lung adenocarcinoma, the authors compared the intertumor protein heterogeneity against the patients' genotypes, including common lung cancer mutations such as *KRAS*, *EGFR*, and *EML4-ALK*, demonstrating a future ability to draw genotypic-phenotypic correlations with this technology. Toward clinical utility, they showed that their technology could have the potential to monitor and predict outcomes to targeted therapies representative of specific cancer pathways, such as the phosphoinositide 3-kinase (PI3K) pathway. In particular, tumors from four patients with metastatic cancers were analyzed before and after PI3K inhibitor treatment so as to correlate response with key proteins of PI3K pathway. Furthermore, they analyzed five drug-naive lung cancer patients—all with various PI3K mutations, who eventually received PI3K inhibitors—to determine the top markers that differentiated responders from nonresponders among the protein panel, suggesting further utility of this technology in the clinical setting.

This digital DNA barcoding–based assay for multiplex protein analysis has several advantages. It successfully provides a multiplexed, quantitative method in a high-throughput manner for limited amounts of clinical samples. Prior approaches have had limited multiplex capacity, were semiquantitative, and involved amplification, which has the potential to introduce biases. Ullal *et al.* (*1*) present an antibody-based approach that has the ability to detect multiple proteins at single-cell sensitivity (correlations greater than 0.95). Being amplification-free, it also improves the detection of low-abundance proteins without increasing background noise. Last, the approach promises to bridge the gap between targeted protein expression and gene profiling by permitting simultaneous analysis.

INTO THE CLINIC

The work described by Ullal *et al.* (*1*) opens new avenues for the application of multiplexed protein analysis in the clinic, which may provide comprehensive understanding of the functional importance of the existing genomic technologies. However, this study also raises several questions with regards to its utility in the clinical setting. Antibody-based methods are subject to variability caused by preanalytic, analytic, and post-analytic standardization, which may affect the interpretation of data. Preanalytical factors (*9*)—such as sample choice, collection, and processing—as well as analytical methods—such as assay conditions and normalization—are major variables that determine the utility of a method of choice in the clinical setting. The current method is restricted to cellular lysates and liquid biopsies, and several questions about how best to integrate this quantitative assay to routine practice need to be considered. First, is this technique superior to mass spectrometry–based techniques, such as SRM and MRM? These are relatively well established for blood-derived specimens and have been applied to blood plasma and FFPE specimens (*6*). The antibody-free approach is a strong plus; however, the technique is not efficient at detecting low-abundance (typically lower than 1 μg/ml) proteins.

Second, it is not clear whether this technology is applicable to other clinical samples. Ullal and coauthors provide evidence of its use in FNA samples, but the utility in FFPE samples needs to be documented because formalin-induced cross-linking leads to alterations of the protein structure and masking of the epitopes. FFPE samples obtained from different institutions may have differing degrees of protein preservation. Additionally, the affinity and avidity of the antibodies vary with species of origin and may lead to disparate results (*10*).

Third, how will the assay contend with tissue heterogeneity? The authors currently overcome this limitation by performing cell sorting and analyzing a single compartment in cancer cell lines and liquid biopsies. Tissues are, however, composed of epithelial and nonepithelial (stromal) elements; differences in the composition of these could result in marked variations in results. Actin, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and histone H3 are not expressed in a compartment-specific manner. Methods that permit normalization of data in a compartment-specific manner (epithelial versus stromal) need to be devised.

Another question to consider is whether the technique will withstand the extensive clinical heterogeneity seen in larger cohorts. For instance, the degree of sample-to-sample and within-sample variation needs to be documented by using specific normalization controls; the use of total protein loading as a control cannot be justified for clinical assays. Although barcoding permits normalization of the DNA reads, variations in cellular characteristics may result in substantial discrepancies in the data obtained from the same number of cells. A more elaborate, consistent, and unbiased approach will be necessary.

FUTURE DIRECTIONS

The ideal technique for protein analysis should enable quantification in an accurate, precise, sensitive, and specific manner for each protein in the panel. The technique described by Ullal *et al.* (*1*) fulfills these essential requirements. In addition, its major strengths include the following: First, this multiplex protein analysis is applicable to a limited amount of cells, involving both single- and bulk-cell populations. In particular, detection of proteins in rare cells and clonal populations, such as stem cells, highlights the particular clinical potential of this technique among the high-throughput protein-profiling methods. Second, its potential clinical utility in companion diagnostics is emphasized by the monitoring of cancer markers and predicting response to therapies. Third, the potential for simultaneous analyses of gene-protein levels from the same patient material will help unravel the complexity of treatment responses among the diverse cases harboring the same mutations.

Further adaptation of the assays described by Ullal *et al.* (*1*) might be necessary for widespread clinical use. The tech-

nology is not adapted to FFPE samples; this is critical for analysis of patients from clinical trials that are the cornerstone of clinical decision-making. Besides technical issues, the assay needs to be validated in larger patient cohorts in order to document the clinical utility with respect to prognosis and prediction of the treatment outcome. Given the complexity of the high-throughput protein multiplex data, this may not be an easy task. New algorithms are needed to determine key driver pathways of functional importance and to identify the dose-response relationship (or relationships). This will enable precision therapeutics, resulting in the prescription of the right therapy to the right patient. Additionally, it should be easy to repeat the assay on tumor biopsy samples so as to monitor the dynamic nature of cancer cells and the adaptation that the tumor cells undergo to acquire resistance to therapies.

The current study by Ullal *et al.* (*1*) is based on liquid biopsies and FNAs and represents a major step forward in achieving the goal of cell-specific multiplexed protein analysis. The ability of this technology to permit simultaneous proteomic and genomic analysis is something that needs to be explored further. One could envision application of these methods to neoadjuvant clinical trials by using "window of opportunity" design to develop a comprehensive understanding of the molecules and pathways associated with the likelihood of response to novel therapies.

REFERENCES

1. A. V. Ullal,, V. Peterson, S. S. Agasti, S. Tuang, D. Juric, C. M. Castro, R. Weissleder, Cancer cell profiling by barcoding allows multiplexed protein analysis in fine-needle aspirates. *Sci. Transl. Med.* **6**, 219ra9 (2014).
2. A. Lipton, L. Goodman, K. Leitzel, J. Cook, J. Sperinde, M. Haddad, W. J. Köstler, W. Huang, J. M. Weidler, S. Ali, A. Newton, E. M. Fuchs, A. Paquet, C. F. Singer, R. Horvat, X. Jin, J. Banerjee, A. Mukherjee, Y. Tan, Y. Shi, A. Chenna, J. Larson, Y. Lie, T. Sherwood, C. J. Petropoulos, S. Williams, J. Winslow, G. Parry, M. Bates, HER3, p95HER2, and HER2 protein expression levels define multiple subtypes of HER2-positive metastatic breast cancer. *Breast Cancer Res. Treat.* **141**, 43–53 (2013).
3. M. J. Gerdes, C. J. Sevinsky, A. Sood, S. Adak, M. O. Bello, A. Bordwell, A. Can, A. Corwin, S. Dinn, R. J. Filkins, D. Hollman, V. Kamath, S. Kaanumalle, K. Kenny, M. Larsen, M. Lazare, Q. Li, C. Lowes, C. C. McCulloch, E. McDonough, M. C. Montalto, Z. Pang, J. Rittscher, A. Santamaria-Pang, B. D. Sarachan, M. L. Seel, A. Seppo, K. Shaikh, Y. Sui, J. Zhang, F. Ginty, Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11982–11987 (2013).
4. Q. Fu, F. S. Schoenhoff, W. J. Savage, P. Zhang, J. E. Van Eyk, Multiplex assays for biomarker research and clinical application: Translational science coming of age. *Proteomics Clin. Appl.* **4**, 271–284 (2010).
5. R. Tibes, Y. Qiu, Y. Lu, B. Hennessy, M. Andreeff, G. B. Mills, S. M. Kornblau, Reverse phase protein array: Validation of a novel proteomic technology and utility for analysis of primary leukemia specimens and hematopoietic stem cells. *Mol. Cancer Ther.* **5**, 2512–2521 (2006).
6. R. W. Sprung, M. A. Martinez, K. L. Carpenter, A. J. Ham, M. K. Washington, C. L. Arteaga, M. E. Sanders, D. C. Liebler, Precision of multiple reaction monitoring mass spectrometry analysis of formalin-fixed, paraffin-embedded tissue. *J. Proteome Res.* **11**, 3498–3505 (2012).
7. S. S. Agasti, M. Liong, V. M. Peterson, H. Lee, R. Weissleder, Photocleavable DNA barcode-antibody conjugates allow sensitive and multiplexed protein analysis in single cells. *J. Am. Chem. Soc.* **134**, 18499–18502 (2012).
8. G. K. Geiss, R. E. Bumgarner, B. Birditt, T. Dahl, N. Dowidar, D. L. Dunaway, H. P. Fell, S. Ferree, R. D. George, T. Grogan, J. J. James, M. Maysuria, J. D. Mitton, P. Oliveri, J. L. Osborn, T. Peng, A. L. Ratcliffe, P. J. Webster, E. H. Davidson, L. Hood, K. Dimitrov, Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26**, 317–325 (2008).
9. A. J. Rai, C. A. Gelfand, B. C. Haywood, D. J. Warunek, J. Yi, M. D. Schuchard, R. J. Mehigh, S. L. Cockrill, G. B. Scott, H. Tammen, P. Schulz-Knappe, D. W. Speicher, F. Vitzthum, B. B. Haab, G. Siest, D. W. Chan, HUPO Plasma Proteome Project specimen collection and handling: Towards the standardization of parameters for plasma proteome samples. *Proteomics* **5**, 3262–3277 (2005).
10. S. Badve, I. T. Vladislav, B. Spaulding, A. Strickland, S. Hernandez, L. Bird-Turner, C. Dodson, B. Elleby, T. Phillips, EP1: A novel rabbit monoclonal antibody for detection of oestrogen receptor α. *J. Clin. Pathol.* **66**, 1051–1057 (2013).

10.1126/scitranslmed.3008407

Citation: Y. Gökmen-Polar, S. Badve, Multiplexed protein analysis. *Sci. Transl. Med.* **6**, 219fs3 (2014).