

Pioneers in Proteomics

Dr. Gilbert Omenn

On the challenges of studying proteins

There are some special challenges with proteomics and proteins more generally. The protein assays that we do have have been developed and optimized one protein at a time. So, here we're trying to identify thousands of proteins simultaneously with methods that sort of do the best they can but cannot possibly optimize for each one because first of all the concentrations of the proteins represent an enormous dynamic range. If we're studying plasma, albumin is present, representing 50 percent of the total protein mass at a concentration of nine orders of magnitude higher than cytokines, which can be detected with sensitive immunoassays but not with standard proteomics. Then most of those proteins undergo all kinds of changes as a function of physiology, menstrual cycle, nutrition, exercise, behaviors, other behaviors, and then, of course, pathological and pharmacological influences. This is the power of proteomics as we will practice it in the not too distant future. But for the present, these are obstacles because they introduce such variability that you have to capture and describe it, analyze it, as well as detecting the proteins.

On the study of genomics vs. proteomics

The genes are found in basically the same number of copies in the chromosomes of every nucleated cell. So, you can measure the genes in a buccal smear from the mouth or a skin biopsy or a blood sample or a tissue sample, you'll find the same genes. There's a little variation by copy number and repeats and a few other details I'll neglect. So, it's the same in different tissues, and the copy numbers are very similar -- just the opposite of proteins. But to understand the value of what we've gained from this tremendous human genome project, all the SNPs and HapMap -- we have to find a way of relating genetic variation, and what we think are genetic predispositions to various diseases to all these other influences -- what we call the environment -- the external environment, the internal environment, behaviors, everything that may change what genes do through their genetically coded products, the proteins.

...in plasma we have the potential to detect proteins from cells all through the body. But that means that they are at some concentration low or high in the cells -- and only what gets out -- and what survives degradation and survives filtration through the kidney can be still present and possibly detected after all the dilution in 4 liters of plasma. It's a good challenge.

On the need for standardization

...we, too, have come to realize that you can waste an awful lot of high powered effort on unreliable samples. Standardization of medical practice and laboratory practice is a tall challenge. We know from current medical practice that there's a lot of resistance, and there's a

lot of haphazard sample collection. For some analytes, it doesn't matter. For many analytes, a tremendous effort goes on in the laboratory after the sample is collected to try to calibrate, put in internal standards, try to adjust for all these other variables. It's not the most efficient way. I believe that we will have more and more understanding of the importance of the collection, handling and storage of the sample. We must find ways to have robust analyses that do not get ruined by small variations in the real world because, after all, samples are taken under different circumstances, and people are busy attending to patients. Many things have to happen. So, a combination of a recognition of the importance of the sample and therefore the quality of the sample, the addition of quality assurance to eliminate certain samples that really have become useless and to eliminate from archived specimen bank samples that are no longer useful, but otherwise to include internal standards or other ways of making reliable measurements. All this will advance and the cancers too with these cooperative programs, the HUPO plasma proteome project and other HUPO initiatives where we're stressing standard operating procedures and protocols and working through consensus to develop them -- and the FDA and other regulatory and reimbursement agencies who will surely contribute to what is a felt need.

On eliminating variation

...every kind of source of technical or database variation, which is extraneous to the comparison of the specimens themselves we must eliminate. And to the extent that we cannot eliminate, we must recognize them.

We must know if you're comparing a sample from one person on a sample from another person. If they are different -- if it's because this person has a disease, which is associated with those protein biomarkers, and this person does not have that disease or the precursor to that disease, it is not acceptable if they could have the same difference based upon how the sample was collected, or how long it stood at 25 degrees, or what speed it was spun at in the centrifuge, or whether there were protease inhibitors added or not, or whether it was fractionated on this or that column, or whether you analyze with one mass spectrometer versus another or put through a different search engine and a different database or even a different version of the same database. This is an evolving target. Our gene and protein databases are evolving regularly.

On the validation of biomarker candidates

New diagnostic tests of broad clinical value are infrequent. There's a misunderstanding about how rapidly you can go from the initial reports of promising biomarker candidates to confirmed findings -- first in the same lab with the same and different specimens and then in other labs. And after confirmation, there are multiple steps for validation.

What's usually done is to compare a small number of patients with a small number of controls, possibly collected under very different circumstances and measure the sensitivity detection of the cancers and specificity of not calling "normals" cancers. But the more important value for a validation is what's called the predictive positive value -- the positive predictive value. This means if you're screening 10,000 women for ovarian cancer of whom fewer than 10 would be expected to have the cancer, you must have a result which has such high specificity that you don't have almost all false positives and that you do find the rare or infrequent true positive.

This is standard work in the validation and development of diagnostic tests. It's not just protein biomarkers. It's true for almost any kind of a test.

On moving markers into the clinical lab

We will identify individual markers and probably more likely, more valuably, panels of several markers -- not a huge number -- but several -- that can be analyzed by multiplex methods. This basically means ELISA, which every clinical lab runs, and affinity capture antibody methods of an ELISA. This is a nontrivial task, and the ELISA vendors spend years optimizing each single assay. So, analyzing for five or eight simultaneously is tricky. And the risk of high background due to some degree of cross reactivity and some interferences in the assays is realistic and has to be assessed very carefully. But I think in the near term, the emerging biomarkers will be analyzed in the clinical labs by platforms to which they've been adapted to suit the labs. In the somewhat longer term, there are some quite exciting new mass spec based methods, which I think may prove extremely valuable and even practicable. After all, reference labs if not the regular clinical labs, do run mass spec for a whole lot of small molecules now. And if we could develop these prototypical glycopeptides spiking methods and some other methods that are loosely now called multiple reaction monitoring with particular peptides that can be identified, I think those have the potential to be high throughput specific high resolution mass spec and those instruments in turn could be adapted for clinical laboratories -- at least reference clinical laboratories.

On the need to personalize medicine

The underlying biological variability must not be forgotten. When we say a patient has pancreatic cancer or breast cancer or whatever, we are talking about extremely heterogeneous patients. One of the biggest failures of medicine has been to lump too many patients together who are really different in important ways. Patients who look the same -- whose tumors look the same, for example, under the microscope from the pathologist and in the x-ray for the radiologist and in the surgical field for the surgeon, we now know by molecular profiles can be very different. We need to understand more of what that means in terms of the causation and in terms of the intervention for treatment and even more important for perhaps for prevention. But I have absolute certainty that the heterogeneity of patients and of tumors will become a bigger and bigger issue, which will be addressed productively by proteomics and by genomics and epigenomics -- gene expression.

The field, which works on differences in response to drugs -- we call pharmacogenomics. Pharmacogenetics is the broader term, actually -- genomics means the particular modern version and certain aspects of genetics. Pharmacogenetics is a term that was coined independently -- simultaneously almost -- in the 1950s by four leading medical geneticists of the time, including my mentor, Dr. Arno Motulsky, who is still active. Fifty years later, we do not have too many examples where we actually do test before administering drugs even though, as you just said, we know that patients differ -- sometimes dramatically -- in both the efficacy of the drug and the side effects from the drug. And we have some very successful drugs that have been taken off the market because a small percentage of patients had unacceptable adverse effects. If they could be identified in advance and be spared that

Clinical Proteomic Technologies Initiative for Cancer

Pioneers of Proteomics

Dr. Gilbert Omenn

exposure and be treated with something else or a drug to be developed for them, then all these others could benefit from that drug. So, incentives are high and still the progress has been slow. Well, with the new molecular signatures of pharmaco- and toxicogenomics and toxicoproteomics, there's a lot of hope that this could be advanced more rapidly. But it demonstrates that we must find ways of differentiating within patient groups who will respond because they have a certain carcinogenic pathway activated or a certain pathway of tumor suppression blocked. And we know more about this -- we know a tremendous more from the last 30 years since the war on the cancer about cancer biology. But we are still struggling to turn it to practical application, and hopefully these new methods will make a big difference.