Written in blood

Clinical proteomics

The low-molecular-weight region of the blood proteome is a treasure trove of diagnostic information ready to be harvested by nanotechnology.

Lance A. Liotta, Mauro Ferrari, Emanuel Petricoin

The blood contains a treasure trove of previously unstudied biomarkers that could reflect the ongoing physiologic state of all tissues. Every cell in the body leaves a record of its physiological state in the products it sheds to the blood, either as waste or as signals to neighbouring cells. What some may view as a cellular refuse is really a diagnostic gold mine.

Routine laboratory blood tests sample only a minute fraction of this potential repository, and there are few specific markers for life-threatening diseases such as cancer. The current biomarker repertoire cannot detect treatable early-stage cancer and often misclassifies common benign conditions. In the face of the urgent need for better disease markers, it is unfortunate that the number of new markers submitted for regulatory approval has virtually dried up.

It is time to rethink our approach to biomarker discovery. The quest for a single biomarker for a particular disease has the illusion of analytical simplicity, but makes little sense from a biological perspective. For example, cancer is caused by intrinsically deranged cells of the host — not an external infectious agent. Why should we expect the cancer cell to generate a unique new protein? It is not surprising that we have failed to find an accurate single marker for a disease as heterogeneous as cancer, which comes in hundreds of types and stages and is a product of the tumour–host microenvironment. Instead, why not take advantage of the very complexity of the disease? Genomics researchers have moved beyond one-gene-at-a-time analysis, and are profiling thousands of gene transcripts to generate entire patterns of information. We should be doing the same with protein biomarkers.

The relative cellular abundance of tens of thousands of different proteins, along with their cleaved or modified forms, is a reflection of ongoing physiological and pathological events. For example, cells that succumb to programmed cell death will leave behind a different protein signature from cells dying of oxygen starvation or infectious insult. As tissues are perfused by blood and lymph, proteins and protein fragments passively or actively enter the circulation. Thus, the complex chemistry of the tumour–host microenvironment should generate unique signatures in the blood macroenvironment.

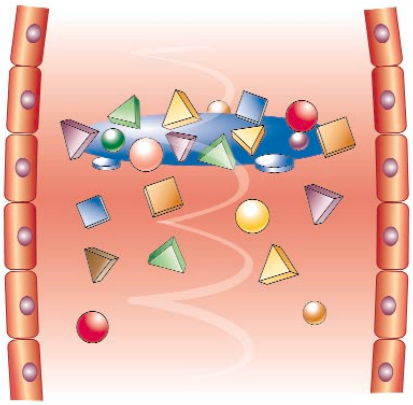
The serum proteome is a complex mixture predominated by high-abundance resident proteins, such as albumin and other carrier proteins, together with proteins that originate from circulating blood cells. Although proteins entering the blood from the surrounding tissue are much less abundant, it is this fraction that is likely to contain most of the undiscovered biomarkers.

Large proteins only enter the bloodstream intact if they are actively secreted, or if the vascular wall becomes permeable owing to disease. But degradation and cleavage can generate fragments small enough to enter the blood passively, producing diagnostic traces. Thus the low-molecular-weight (LMW) region of the blood proteome, which is a mixture of small intact proteins plus fragments of the large proteins, represents all classes of proteins. This treasure trove of diagnostic information has largely been ignored until now.

The size of a protein determines how fast it is cleared from the blood by kidney filtration and uptake by the liver. Small proteins are rapidly cleared, with half-lives of less than a few hours, whereas large proteins have extended half-lives — albumin's, for example, is 19 days. Consequently, the only way a small molecule can stay in circulation is to hitch a ride with a carrier protein. Thus, at any point in time the concentration of a LMW biomarker is a function of the kinetics of its entry and exit, as well as its binding affinity for carrier proteins.

The carrier proteins act as magnets to accumulate, and thereby amplify, low-abundance biomarkers. A trickle of biomarkers entering the blood is mopped up the carrier proteins, which amplify and integrate the history of biomarker production, as a capacitor stores electricity. Existing fractionation methodologies often discard abundant carrier proteins and thus fail to capture this valuable resource — akin to throwing the baby out with the bathwater.

Recognizing the existence of such amplification and enrichment should shift our future discovery efforts toward the constellation of carrier-protein-bound LMW biomarkers. Fortunately, mass spectrometry is a technique ideally suited for detecting and distinguishing thousands of LMW proteins and peptides in seconds. In the near future, we will be able to scan the blood proteome rapidly by mass spectrometry, decipher the buried diagnostic information, and then go directly to a list of the underlying identities in a database. Artificial intelligence-type algorithms can sort through the information contained in thousands of data points, recognizing and exposing disease portraits.



We can envision a future in which the blood-biomarker archive can be monitored with new technology created at the intersection of the fields of artificial intelligence, nanotechnology and proteomics. Advances in microfabrication may provide 'nanoharvesting' agents designed specifically to capture and amplify classes of LMW biomarkers. Imagine a fleet of harvesting agents tailored to monitor specific diseases. Harvesting particles could be administered in the physician's office, and then sampled at a later visit after they have had time to collect their diagnostic cargo. A serum sample containing LMW diagnostic molecules sequestered on the harvesters can now be rapidly analysed with mass spectrometry. The result could be an individual global-health profile rendered at an affordable cost, revolutionizing the field of disease diagnosis and health monitoring. Perhaps there will be a time in which a small sample of blood will reveal an image of the physiological and pathological states of every tissue in the body. ■

Lance A. Liotta is at the National Cancer Institute, 10 Centre Drive, Bethesda, Maryland 20892, USA. Mauro Ferrari is in the Department of Internal Medicine, Dorothy M. Davis Heart and Lung Research Institute, Ohio State University, Columbus, Ohio 43210, USA.
Emanuel Petricoin is at the Food and Drug Administration (FDA) Center for Biologics Evaluation and Research, 8800 Rockville Pike, Bethseda, Maryland 20892, USA.

FURTHER READING

Anderson, N. L. & Anderson, N. G. *Mol. Cell Proteom.* **1,** 845–867 (2002).
Petricoin, E. F. *et al. Lancet* **359,** 572–577 (2002).
Tirumalai, R. *et al. Mol. Cell Proteom.* **2,** 1096–1103 (2003).
Liu, J. & Ferrari, M. *Dis. Markers* **18,** 175–183 (2002).
Mehta, A. *et al. Dis. Markers* **19,** 1–10 (2003).