

Editorial

Cancer Biomarker Discovery: Speed-bumps and Tire Shredders

Michael A. Tainsky

Program in Molecular Biology and Genetics, Karmanos Cancer Institute and Department of Pathology, Wayne State University School of Medicine, 110 E. Warren, Detroit, MI, USA

Tel.: +1 313 578 4340; Fax: +1 313 832 7294; E-mail: tainskym@karmanos.org

In this remarkable era of modern molecular biology we have been fortunate to be able to perform high throughput investigations of cancer-specific DNA sequence variations, RNA expression profiling, proteomics, and metabolomics with tremendous expectations for their application to the early detection and treatment of human cancer. The ease of obtaining such high dimensional data and the throughput of these approaches to discover cancer-specific changes in various molecular moieties belie the difficulties in the full development and implementation of these molecules for clinic-ready biomarker tests. The problems can be rooted in the high-dimension data sets that may be over-fitted in the discovery phase of this type of study (too many analytes and too few specimens) or they may be related to the nature of the samples used in discovery and validation that may not properly represent the true clinical implementation. Few recently developed molecular diagnostic tests have passed the regulatory hurdles to FDA-approval for the pre-symptomatic diagnosis of cancer [2] and many tests in use today might not pass in the current regulatory environment especially given the recognition of the over-diagnosis and over-treatment of otherwise indolent cancers [4].

The field of cancer biomarkers is grounded in the principle that if we can identify individuals with cancer at its earliest known stage we can reduce the morbidity and mortality from the disease. Through this credo of early detection as a means to reduce cancer mortality we can look back with confidence at cervical cancer and the reduction of mortality from this disease resulting from the PAP test [1] as evidence that the advances

in molecular diagnostics have enormous potential [3]. The reduction in mortality from cervical cancer was not without its cost as there is certainly over-treatment of what are likely indolent lesions, cervical dysplasia, in order to assure the removal of malignant cells. Therefore sensitive diagnostic tests with low specificity result in costs we currently bear at the expense of morbidity from overtreatment. Because particular cancers with similar presentations can have very different courses of malignant development due to subtle differences in molecular etiology, it is unlikely that a single biomarker could provide an accurate diagnostic test for that disease. Therefore panels of biomarkers will be required to achieve sufficient diagnostic sensitivity to comprehensively identify even a single type of human cancer.

Assuming the appropriate technology is in place, three prominent elements are required for outstanding biomarker discovery, technical validation and prospective clinical validation: specimens, specimens, specimens! One only needs to look at the standard approaches to biomarker discovery to identify specimens as the major barrier to arriving at sensitive and specific biomarker panels for each clinical application. Investigators who discover and validate biomarkers do not generally have access to well-curated prospective case-control samples representative of the future screening population and so biomarkers often fail at the validation stage or in pre-symptomatic clinical trials. Commonly the study design begins with samples obtained at the time of disease diagnosis for the discovery phase of the work. Similar but independent samples are

used for validation. Control samples are frequently obtained in different settings not involving the physiological and psychological stresses of the disease or preparation for clinical interventions. Even more importantly, these cases and controls are not representative of the eventual screening populations so that even the best biomarker discovery project that can discriminate cases from controls may not perform predictively in a pre-symptomatic screening setting. Clearly the field will struggle until biological samples from well-curated prospective case-control studies are available for biomarker discovery with sufficient clinical follow-up to avoid confusing indolent tumors from malignant cancers. These samples from well-curated prospective case-control studies are precious so their use in biomarker discovery requires careful peer-review combined with a common sense attitude that rationally designed studies, using feasible technologies, and requiring only small amounts should be supported as the lifetime of some analytes in these samples in the freezer are not infinite.

Antibodies as analytes have the advantage that they are able to survive longer in these serum/plasma biorepositories than most other proteins. All of these factors need to be taken into account by the gatekeepers of these precious repositories. The recent initiative by funding agencies to support research into the proper development and maintenance of biorepositories is an indication of the recognition of this critical aspect of biomarker research.

Another kind of study design issue in biomarker research, more mechanical in nature, I call the "technical abyss". Highly discriminant biomarker panels may be identified using genomic or proteomic platforms that are suitable only for the discovery stages of a program. Once samples are evaluated using a more clinically-appropriate technical platform, the panel of biomarkers could lose accuracy if one or more essential components of the panel is not amenable to the new device [5] resulting in a loss of sensitivity and/or specificity. For example, a panel of mRNA biomarkers could be discovered in tissue analyses. However about 40% of the mRNA expression differences result in protein expression changes. If these proteins are coordinately regulated with the mRNAs but they are not detectable in body fluids suitable for pre-symptomatic diagnostics, proteins from the biomarker panel will be lost along with the accuracy of the potential composite panel-based test. Proteomics-based biomarker discovery in serum or plasma avoids many of these problems as long as the ultimate clinical bioassay is implemented on a

mass spectroscopy platform. However, a panel of protein biomarkers is more likely to be implemented in an immunoassay-based device. In this situation the barrier to successful translation of that panel of proteomic biomarkers to a clinical laboratory is the development of antibody pairs that function together in an ELISA test. This is non-trivial as antibody recognition of different epitopes by different antibody pairs can result in large discrepancies in the ability to detect each analyte when compared to the original assay. Another technical abyss that may arise is the lack of sensitivity or specificity when the assay is transferred to the clinical laboratory platform which leads to an overall loss of accuracy for the panel of analytes. Therefore it is advisable to use a discovery platform that is as technically similar to the anticipated clinical diagnostics platform as possible.

The proper implementation of sophisticated molecular diagnostics in the near future depends on whether biomarker discovery and validation studies on early stage disease can detect early cancers in pre-symptomatic prospectively accrued samples. Another major issue is whether discovery performed on archival samples will be relevant to the current standard of care and current test subjects. An additional challenge will be the development of customized informatics tools to integrate disparate analyses, such as imaging tests, serum/plasma biomarker tests, and personalized whole genome-based genetics into a single diagnostic recommendation.

In this issue we address several high-throughput approaches to the development, validation and implementation of antibodies as protein biomarkers for the detection of cancer as well as their theranostic value for personalized cancer vaccines. Cancer antibodies are utilized in various realms such as 1) analytical tools to measure candidate protein biomarkers in classical ELISAs and multi-analyte ELISAs on novel platforms such as surfaces of planar microarrays and microbead arrays, 2) as the bait to discover diagnostically useful tumor-associated antigens expressed by the tumor cells or the tumor microenvironment, and 3) as the analytes themselves because autoantibodies are generated by the host immune system in response to tumor associated antigens. This monograph will address critical issues in cancer antibodies and barriers to their usefulness as tools in the oncologist's arsenal.

References

- [1] S.H. Landis, T. Murray, S. Bolden and P.A. Wingo, *Cancer statistics, 1999*, *CA Cancer J Clin* **49** (1999), 8-31.

- [2] J.A. Ludwig and J.N. Weinstein, Biomarkers in Cancer Staging, Prognosis and Treatment Selection, *Nature Reviews Cancer* **5** (2005), 845–856.
- [3] G.N. Papanicolaou, The Cell Smear Method of Diagnosing Cancer, *Am J Public Health Nations Health* **38** (1948), 202–205.
- [4] H.G. Welch, and W.C. Black, Overdiagnosis in cancer, *J Natl Cancer Inst* **102** (2010), 605–613.
- [5] J.W. Zolg and H. Langen, How industry is approaching the search for new diagnostic markers and biomarkers, *Mol Cell Proteomics* **3** (2004), 345–354.