

# Biomarker discovery – the need for new generation peptide-protein microarrays

The progression from health to disease is marked by significant biological changes within an individual<sup>1</sup>. Clinically presenting symptoms, however, can be non-specific and variable enough to hinder diagnosis, and may appear only after a disease has already become well-established and consequently more difficult to treat<sup>2</sup>.

**B**iomarkers are extremely valuable tools for clinical scientists looking to overcome such challenges. A biomarker is classified as a substance or molecule indicating a particular biological process, which can be an extremely specific surrogate for a certain disease state. Changes in biomarker levels can predict clinical outcomes which would otherwise take months or years to detect conventionally in the clinic. The most robust biomarkers indicate factors essential to the underlying process of the disease being detected or treated. For example, in slow-progressing fibrotic diseases, biomarker monitoring can be vital in guiding the course of treatment by providing short-term indications of disease progression and effects of on-going therapy. This is equally important in clinical trials, where biomarker state or abundance can replace a traditional clinical endpoint, more qualitatively and often with greater accuracy<sup>3</sup>.

Despite the importance of biomarkers to our understanding of immunology, the field of biomarker discovery has progressed slowly<sup>4</sup>. To date, the FDA has approved only two biomarker panels

for use as a prognostic tool for breast cancer and only one for ovarian cancer<sup>5</sup>. Due to the complexity of biological pathways and the heterogeneity between individuals, a single biomarker does not have a high enough predictive value to be a reliable diagnostic tool as it may only present consistently in a certain proportion of the population in response to an antigen or tumour, or may contain enough variation to elude reliable detection.

With a recent shift in immunology and oncology towards studying networks and interference across whole systems, the concepts of systematic relationships are now being applied to biomarker discovery<sup>6</sup>. In particular, multilevel biomarker panels, detecting markers at different informational levels including microRNAs, molecular lipid species and proteolytic degradation products could provide unparalleled diagnostic, prognostic and theranostic information. However, to successfully target therapies towards the specific patient population in which they will have the most benefit requires a huge leap in the speed of biomarker discovery and validation.

**By Dr Nikolai Schwabe and Katherine Catchpole**

---



**Figure 1**  
Ixodes genus ticks are the  
carriers of Lyme disease

Oncology is a field which stands to particularly benefit from an increasingly routine use of biomarkers, as significant patient benefit is gained from early tumour detection and subsequent tailored treatment<sup>7,8</sup>. The challenge here is to identify biomarkers with the sensitivity to detect (or which are produced as a result of) very small tumour masses, and also with the specificity to allow distinction between related cancerous and benign conditions<sup>5,9</sup>.

The human proteome has been a strong focus for biomarker discovery, which is most accessible in the form of plasma or serum derived from blood samples. While only a handful of proteins are currently used in routine clinical diagnoses, the abundance and structure of the plasma protein pool may change in ways indicative of many, if not all, human diseases<sup>10</sup>. Proteomics therefore has the potential to deliver an immediate and accessible means to assess health and disease states in real-time, including building patterns of activity levels of chronic conditions<sup>4</sup>. Epitope patterns of different patient subsets can be highly informative both in predicting patient prognosis and in gaining insight into the immune system's battle against infection, in turn leading to the development of new therapies for disease treatment or management. For instance, serological analysis of autoimmune patients has revealed dominant epitopes and provided evidence for the presence of antibody specificities associated with specific disease subsets, invaluable for guiding increasingly individualised treatment plans. In general the 'antibodyome', a profile of many different circulating antibodies, can indicate the presence of infectious diseases or cancer. Equally, this can indicate the

body's response to human intervention, such as dosing with biological and other drugs. The challenge which screening technology faces here is the need for flexibility to detect multiple protein or antibody specificities and their relative frequency at different disease states.

Banked and archived sample collections of patient plasma and sera have the potential to be important sources for biomarker discovery in this context, although often only very small volumes of sample are available, for example as 'leftovers' from previous studies. This still represents a valuable resource, however, as sera can retain much reactivity even after several decades of storage<sup>11</sup>.

Immunoassays using, for example, monoclonal antibodies or other detection reagents, are one of the most commonly used methods for the detection and validation of protein biomarkers, and significant advancements in their quality and application range have been made as research groups continue to develop them. Creating traditional immunoassays, such as ELISAs, can be a labour-intensive task. Further, if multiple protein-protein interactions are to be studied on a broader range of samples, such assays require large amounts of reagents and analytes. There is, however, the opportunity to miniaturise such assays following proof of principle<sup>11</sup>. Assays can be transferred to specialist microarray platforms at dedicated core facilities in order to gain very significant economies of scale, both in terms of sample and reagent usage as well as in overall processing cost.

Microarray platforms offer high degrees of multiplexing while consuming very small sample volumes of just a few microlitres over the entire array, compared to the millilitres consumed by the traditional 96-well immunoassays. They are, therefore, ideal for tackling the challenges associated with biomarker discovery, collecting statistically powerful volumes of data with their ability to detect tens of thousands of interactions simultaneously in a multiplex format<sup>1,12</sup>. The reduced sample consumption makes them ideal for analysing the small volumes of sera stored in historic sample banks, as well as in paediatric or animal studies where it is only possible to collect limited sera samples. Once transferred to one of these next-generation platforms, the improved efficiency and speed makes it possible to investigate and compare samples from many more individuals than would have been possible on the original assay set-up.

With the advent of new powerful and scalable assay technologies, projects that were previously prohibitive in terms of cost and sample usage are now possible. Processing up to 40,000 features per

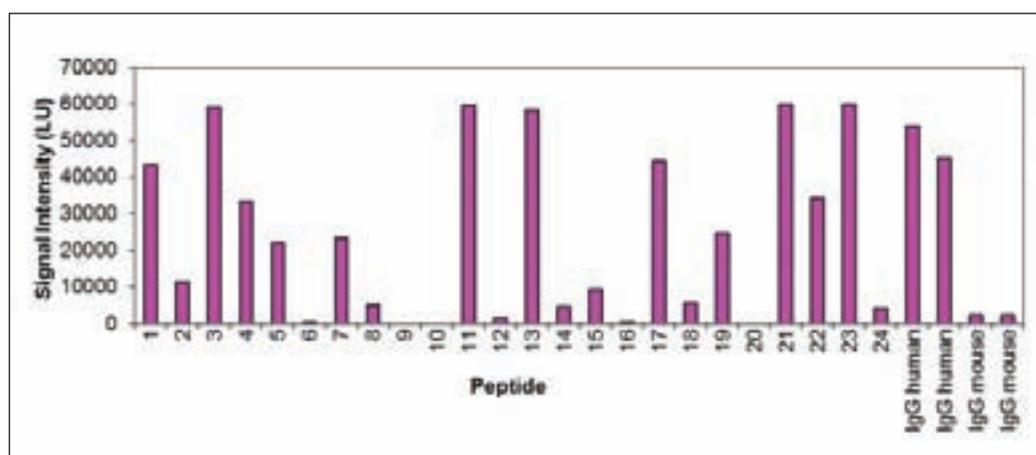
array, projects can be accommodated that demand scale to achieve the required statistical power efficiently and, in fact, can now be delivered at a cost less than equivalent studies on a legacy 96-well format. Microarray analysis equipment has also evolved to such an extent that it now exceeds the detection limits of standard immunoassay reading equipment. This means that more accurate and sensitive measurements can be achieved that reach into the single pictogram per ml range alongside very attractive dynamic ranges.

A recent case study comes from New York Medical College, from which P. Arnaboldi and his team used a commercial high throughput ligand binding assay service as part of their investigation into biomarker-based diagnosis of Lyme disease<sup>13</sup>. Lyme disease is endemic in many regions of North America and Europe, spread by ticks of the genus *Ixodes* which carry *Borrelia* bacteria. If untreated it can lead to permanent neurological and musculoskeletal damage, but early-stage infection can be resolved with a course of antibiotics. Early identification and treatment is therefore critical to patient outcome, but clinical diagnosis is limited by Lyme's varying and non-specific symptoms. Diagnosis relies on detecting antibodies against the bacterium or bacterial proteins in serological assays, as the presence of the *Borrelia* bacterium itself cannot be detected by classic microbiologic methods.

The current two-tier assay used for Lyme diagnosis, an ELISA followed by a western blot, is unreliable at detecting early infection. The target proteins lack in both sensitivity and specificity, with significant sequence variation and a high level of cross-reactivity with antigens from other bacteria. The assay combination fails to identify the presence of

early Lyme in up to 50% of cases. Arnaboldi's team hope to identify at least five peptide epitopes which are both highly specific to *Borrelia* and highly conserved, and thereby develop a sensitive multi-peptide diagnostic assay which will detect the presence of human-generated antibodies against Lyme. The team focused on the *Borrelia* surface protein OspC as a potential source of epitopes. OspC was chosen as it is required for bacterial transmission from ticks to humans, and therefore is always present in Lyme infection. However, with a high degree of variation within the sequence of the protein as a whole, Arnaboldi used the assay service to further investigate whether the OspC protein contained any epitopes which were highly conserved in *Borrelia* and provoked an antibody response from sera of Lyme-infected patients.

Eight Lyme-positive patients were tested against an overlapping library of 37 peptides from the bacterial OspC protein to discover a conserved epitope, OspC1, which is indicative of Lyme infection in humans and could form the basis of a diagnostic biomarker panel. From the data generated, Arnaboldi and colleagues were able to identify three peptides which bound to more than 50% of the Lyme samples, with the entire process taking place over just five weeks. Subsequent ELISAs confirmed that one of these epitopes, OspC1, positively detects Lyme in more than 75% of early Lyme patients and is highly conserved. It is therefore a good candidate to form part of an assay to reliably detect early-stage Lyme and ensure timely treatment of the disease. Dr Arnaboldi and co-workers hope to identify several more of these highly specific, conserved *Borrelia* epitopes, for an assay with high diagnostic specificity which could



**Figure 2:** Results from a test serum sample for binding to 12 peptide pairs (native and citrullinated) from RA-associated proteins. Asterisk indicates result from one citrullinated fibrinogen peptide sequence (peptide 13), shown next to result from the native sequence (peptide 14)

## References

- 1 Mehan, MR et al. Highly multiplexed proteomic platform for biomarker discovery diagnostics, and therapeutics. *Adv Exp Med Biol.* 2013 735:283-300.
- 2 Das, N, Biswas, B, Khera, R. Membrane-bound complement regulatory proteins as biomarkers and potential therapeutic targets for SLE. *Adv Exp Med Biol.* 2013 735:55-81.
- 3 Wallis, RS et al. Tuberculosis biomarkers discovery: developments, needs, and challenges. *Lancet Infect Dis.* 2013 Apr; 13(4):362-72.
- 4 Vives-Pi, M et al. Biomarkers for diagnosis and monitoring of celiac disease. *J Clin Gastroenterol.* 2013 Apr; 47(4):308-13.
- 5 McDermott, JE et al. Challenges in biomarker discovery: combining expert insights with statistical analysis of complex omics data. *Expert Opin Med Diagn.* 2013 Jan; 7(1):37-51.
- 6 Mayr, M et al. MicroRNAs within the continuum of postgenomics biomarker discovery. *Arterioscler Thromb Vasc Biol.* 2013 Feb; 33(2): 206-14.
- 7 Deyati, A et al. Challenges and opportunities for oncology biomarker discovery. *Drug Discov Today.* 2012 Dec (epub).
- 8 Chung, L, Baxter, RC. Breast cancer biomarkers: proteomic discovery and translation to clinically relevant assays. *Expert Rev Proteomics.* 2012 Dec; 9(6):599-614.
- 9 Honda, K et al. Proteomic approaches to the discovery of cancer biomarkers for early detection and personalized medicine. *Jpn J Clin Oncol.* 2013 Feb; 43(2):103-9.
- 10 Anderson, NL, Anderson, NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics.* 2002 Nov; 1(11):845-67.

Continued on page 56

Continued from page 55

**11** Marshall, RP, Simpson, JK, Lukey, PT. Strategies for biomarker discovery in fibrotic disease. *Biochim Biophys Acta*. 2013 Jan(epub).

**12** Makridakis, M, Vlahou, A. Secretome proteomics for discovery of cancer biomarkers. *J Proteomics*. 2010 Nov;73(12):2291-305.

**13** Arnaboldi, PM et al. Outer surface protein C peptide derived from *Borrelia burgdorferi sensu stricto* as a target for serodiagnosis of early Lyme disease. *Clin Vaccine Immunol*. 2013 Apr;20(4):474-81.

**14** Shreffler, WG et al. Microarray immunoassay: association of clinical history, in vitro IgE function, and heterogeneity of allergenic peanut epitopes. *J Allergy Clin Immunol*. 2004 Apr;113(4):776-82.

**15** Klareskog, L et al. Immunity to citrullinated proteins in rheumatoid arthritis. *Annu Rev Immunol*. 2008 26:651-75.

significantly improve the ability to diagnose Lyme disease, and increase positive patient outcomes<sup>13</sup>.

Shreffler et al analysed patient sera using a peptide microarray and found that patients with a greater diversity of epitopes had a history of more severe allergic reactions<sup>14</sup>. And for diseases with well-characterised diversification in antigen reactivity as a result of immune progression, ie Lyme, where the antibody profile changes from presenting as primarily IgM to primarily IgG as the disease progresses, this information is invaluable for identifying the stage of the disease<sup>13</sup>.

Other recent work carried out at Karolinska University Hospital in Stockholm, Sweden in the field of Rheumatoid Arthritis (RA) explored the association of disease specific protein modification called citrullination in RA<sup>15</sup>. The onset of RA is associated with MHC class II-dependent activation of adaptive immunity. A variety of proteins (such as vimentin, alpha-enolase, type II collagen and fibrinogen) undergo post-translational modification of arginine to citrulline over time and a working hypothesis is that this may be aggravated by smoking or bacterial infection. Autoantibodies towards these modified proteins may complex with their target citrullinated proteins as part of a multistep process in RA disease progression.

Increasing the number of B cell epitopes on candidate autoantigens is central when dissecting this complex disease and with that intent researcher Lena Israelsson turned to a B cell epitope mapping service where a series of novel and previously identified peptides from RA-associated proteins were synthesised in both their native and their citrullinated variants, then spotted on to multiple microarray slides for screening. Sera collected from RA patients were assayed for binding to the paired native and citrullinated peptides. Across the set of patient samples, differences were apparent between citrullinated peptides and their unmodified counterparts.

For each of the peptide pairs screened, it was evident that the patient antibodies bound more strongly to the citrullinated variant than the native peptide. The peptide array based mapping is sufficiently sensitive to highlight these differences in binding. While further validation work is required, these results provisionally support the hypothesis that citrullinated epitopes could be important for the development of RA. The data helped shortlist some of the novel peptides for further studies at the Karolinska Institute<sup>15</sup> (Figure 2).

In summary, protein microarrays generate very powerful data sets, providing detailed and quantitative analysis of protein-protein interactions. The

level of assay multiplexing achievable with this kind of technology makes it useful for a broad range of applications, including epitope mapping of monoclonal antibodies and vaccine candidate protein or reverse immunology analysis of finding new vaccine targets based on immune responses observable in patient sera. This provides valuable information for disease monitoring, diagnostics and the development of new therapies.

Most importantly, while developing an immunoassay complete with reagents for detection is a difficult and labour-intensive process, next-generation microarrays allow for easy transfer of almost any existing assay to a microarray set-up where it benefits from the increased sensitivity and power of high-throughput technology. An efficient and cost-effective way to do this is by outsourcing to dedicated facilities, rather than investing the time and labour costs of running these assays in-house and the expenditure on specialised equipment which multiplexed assays require.

This wealth of information beginning to be extracted from new array technology has come with the new challenge of how to extract meaningful molecular signatures of biological processes – biomarkers – from the complex datasets generated which are full of non-useful ‘noise’. Data- and knowledge-driven metrics are being developed by the bioinformatics field to identify and statistically assess biomarker candidates<sup>11</sup>. **DDW**

---

*Dr Nikolai Schwabe is CEO and co-founder of UK-based private biotech company ProImmune, a leader in services for understanding immune responses. Prior to founding ProImmune, Nik was responsible for Business Development and Intellectual Property at Bookham Technology plc, a fibre-optic components maker and before that, a financial analyst at Salomon Brothers International in London, advising clients on financial portfolio strategies. Nik holds a DPhil in Theoretical Physics from Oxford University.*

*Katherine Catchpole works for Dr Nikolai Schwabe's team at ProImmune. Prior to joining ProImmune, Katherine completed her studies in Biological Sciences with a BA from Oxford University.*