

Application of a biomarker discovery platform to identify novel panels of biomarkers for the diagnosis of prostate cancer



Jens Koopmann, Colin Wheeler, Nicholas Workman, John Anson and Michael B. McAndrew

Introduction

The issues surrounding the use of prostate-specific antigen (PSA) in the diagnosis of prostate cancer (PCa) are well documented (sensitivity of 86% and specificity of 33% at 4 ng/ml)¹. There is therefore a need for a diagnostic test with greater discriminatory power to reduce overtreatment and under-diagnosis. The development of autoantibodies associated with PCa has been described². In general, the appearance of such antibodies can precede disease symptoms by many years, making them attractive as potential biomarkers for early diagnosis.

We have developed a unique biomarker discovery platform based on a "functional protein" array which utilises correctly folded proteins³. This platform has been used to detect autoantibodies⁴ in prostate cancer serum samples and to identify their cognate antigens. We have taken advantage of the multiplex nature of array assays and applied powerful data analysis strategies to identify panels of biomarkers which may have clinical utility in the diagnosis of prostate cancer.

Materials & Methods

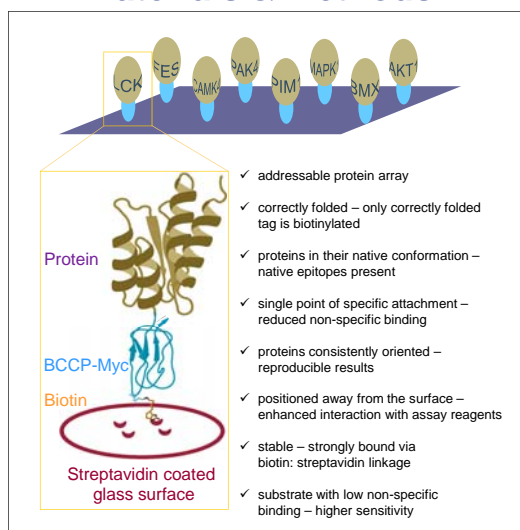


Figure 1: Functional protein array platform.

Protein array

The discovery platform is based around a unique "functional protein" array technology⁴ which has the ability to display native, discontinuous epitopes (Figure 1). The proteins are full-length, expressed with a folding tag and screened for folding before being arrayed in a specific, oriented manner designed to conserve native epitopes. The array content represents multiple functional and disease pathways and many have been previously reported to elicit autoimmune responses in cancer patients. In this study, 925 proteins were screened whereas the current version of the array includes 1340 proteins. Arrays were screened using serum samples as described previously⁴.

Study design

- Case cohort (n=73): patients diagnosed with prostate cancer, stages I-IV.
- Control cohort (n=60): 23 individuals diagnosed with benign prostate hyperplasia and 37 healthy donors

Results

Assay reproducibility

More than 140 QC steps were used to monitor performance of the arrays, assay and data analysis. The reproducibility of the platform is exceptionally good, making it possible to screen statistically meaningful numbers of clinical samples. An identical pooled serum sample was assayed on 82 separate occasions over 2 months which showed an overall CV of 1.3% following normalization (Figure 2A). The Pearson correlation for the same array data showed a similarity of 0.96 (Figure 2B) demonstrating remarkable consistency from assay to assay. Figure 3 demonstrates that identical serum samples show high reproducibility in the assay and can be clearly identified even when assayed using different batches of arrays.

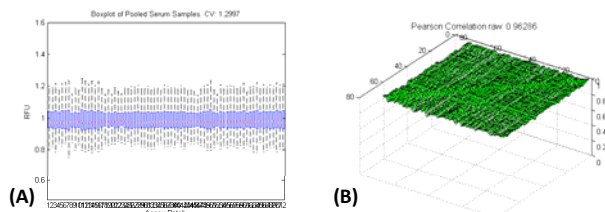


Figure 2: Performance of array in serum binding assay. (A) Boxplot of overall CV for a pooled serum sample, assayed 82 times, and (B) Pearson correlation of 0.96 for a pooled serum sample, assayed 82 times.

Results (cont'd)

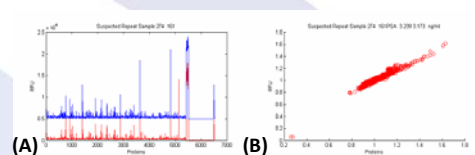


Figure 3: Identical serum samples show high reproducibility. (A) Reactivity of identical sample assayed by two operators in independent assays (red and blue), separated to show reactivity (reactivity profile overlays almost perfectly). (B) Plotting one assay against the other shows high reproducibility.

Selection and performance of biomarker sets

We have investigated several strategies for data normalisation, marker selection and data classification. Importantly, rather than simply generating a list of individual markers which are generally more reactive in case vs control, our approach is focused on building panels of biomarkers where the data from each biomarker is correlated with each other, to produce a robust panel giving high sensitivity and specificity for disease.

All experimental raw data were normalised. Biomarker selection was performed using a genetic algorithm and panels were cross-validated until the sensitivity and specificity converged. Two independent permutation assays were used to confirm that the chosen biomarker panels are related to the disease status of the sera and not to an inherent bias in the data. The performance of classifiers was assessed by referring to a combined sensitivity and specificity (S+S score).

Biomarker panels consist of n members (where n = 1 to 15). As the number of biomarkers in the panels increases, the performance increases as indicated by S+S score and shown by the ROC curve (Figure 4A).

Selection and performance of biomarker sets

The best performing panels have combined S+S scores of greater than 1.8 and classify the samples with both a sensitivity and specificity exceeding 90% (Figure 4B). In comparison, PSA has an S+S score of -1.2 (i.e. sensitivity of 86% and specificity of 33%¹ based on multiple studies⁵). Note that since PSA is used at the point on the ROC curve where sensitivity is close to its maximum, the specificity of the test is very low. In comparison, for a panel of 15 autoantibody biomarkers the curve is steep and the sensitivity approaches its plateau when specificity is still very high (~90%).

Ten of these 15 biomarkers have previously been linked to cancer in general or prostate cancer in particular. The remaining five biomarkers have not previously been reported to have an association with cancer. Further development and validation of these panels is ongoing in a larger study involving 1800 samples.

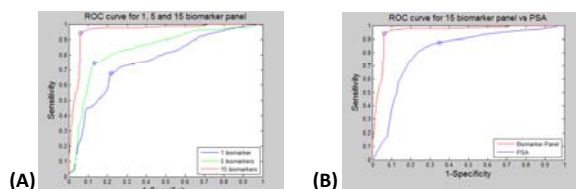


Figure 4: Biomarker panel performance. (A) ROC curves comparing the performance of autoantibody biomarker panels incorporating one, five or fifteen biomarkers. (B) ROC curve of the most highly performing panel, compared with PSA.

Conclusions

- The combination of functional protein microarray platform with multiplexed panels of assays, rigorous QC processes and powerful data analysis methodology enables the rapid identification of highly sensitive and specific biomarker panels and is applicable to a range of cancers.
- A set of autoantibody biomarkers which distinguish prostate cancer from control samples with both sensitivity and specificity above 90%.
- The performance of these panels is considerably better than existing diagnostic tests (i.e. PSA) in distinguishing case from control serum samples.
- A number of proteins previously associated with cancer have been identified along with novel biomarkers.
- Further development and validation of these panels is ongoing in a larger study involving 1800 samples.

Visit www.ogt.co.uk for more information.

References:

1. Hoffman *et al* (2002) BMC Family Practice 3:19
2. Wang *et al* (2005) N. Engl. J. Med. 353:1224
3. Koopmann *et al* (2004) in "Protein Microarrays", pp. 401-420, ed. Mark Schena
4. Grnjatic *et al* (2009) J. Immunol. Methods 341:50
5. Harvey *et al* (2009) BMC Urology 9:14