

A Pilot Study for the Use of Random Peptide Arrays for Biomarker Discovery in Breast Cancer

Johannes Zerweck¹, Antonia Masch¹, Petra Meyer¹, Tobias Knaute¹, Paul von Hoegen¹, Holger Wenschuh¹, Mike Schutkowski² & Ulf Reimer¹

¹JPT Peptide Technologies, Berlin, Germany. ²Martin-Luther-Universität Halle-Wittenberg, Halle, Germany.

Introduction

Autoantibodies against tumor associated antigens (TAA) are considered to be potent biomarkers for cancer.

The identification of patterns of autoantibody reactivity towards multiple TAAs can increase sensitivity of diagnostic assays.

Due to the number of >1700 proteins suspected to be potential biomarkers¹ and the frequency of sequence variations in such proteins a comprehensive screening approach seems difficult.

Diverse random peptide libraries are an alternative to knowledge based peptide libraries.

Library Design

- Random libraries are suitable for identification of distinct antibody epitopes and mimotopes²
- Knowledge based libraries (peptide scans) are less diverse than random libraries
- Library of 40734 peptides biased towards lower hydrophobicity and AA distribution of human proteome

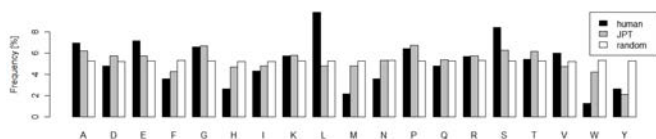


Fig. 1. Amino acid distribution in the human proteome and the random library used.

Final library of 40734 peptides is printed onto 6 peptide microarrays and possesses the following coverage:

- 87 % of all possible tetramers
- 85 % of all tetramers with one gap
- 82 % of all tetramers with two gaps
- 77 % of all tetramers with three gaps

Array Production

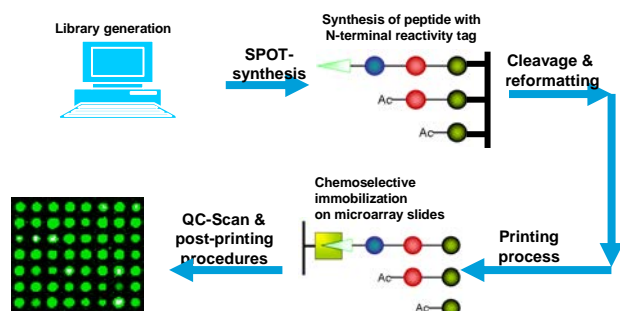


Fig. 2. Schematic representation of the array production process.

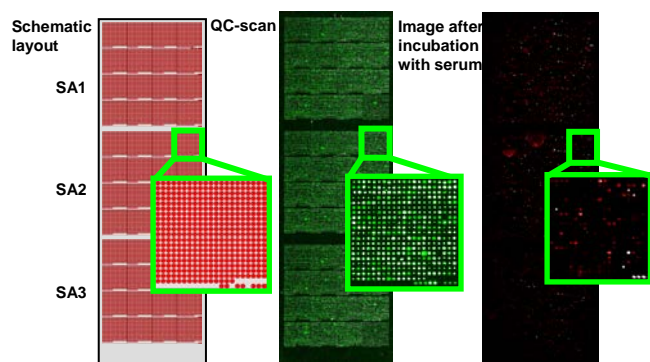


Fig. 3. Layout of peptide microarray and images after printing (QC-scan) and after Serum incubation. Three subarrays are used for improved data quality.

Seroscreening

- 20 plasma samples from patients with breast cancer
- 12 plasma samples from healthy volunteers
- Incubation on 6 library slides (> 200 incubations)
- Evaluation of images using GenePix
- Processing of data and calculation for QC with R

> 1.3 Mio data points were generated

Data Quality

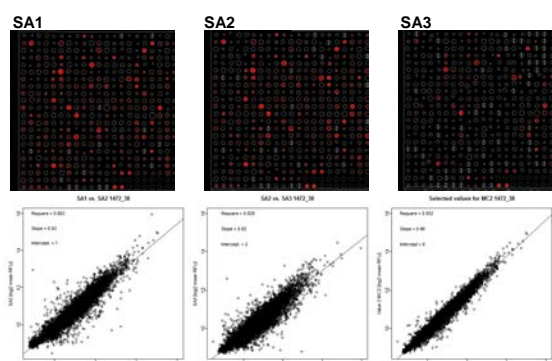


Fig. 4. Typical intra array reproducibility between the three subarrays.

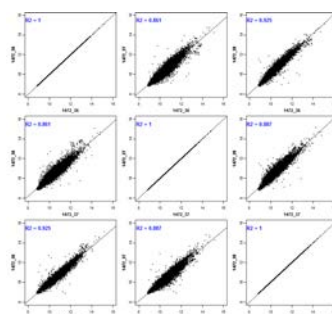


Fig. 5. Reproducibility between different assays for one sample. The scatterplots show the mean signal from three identical subarrays in one experiment vs. another.

Data Evaluation

- Use of log2 for signal intensities
- Normalization of intensities within array series
- Calculation of p-value (Wilcoxon rank sum test)

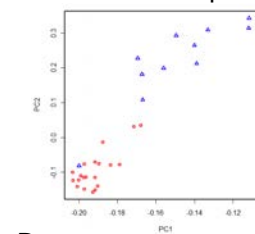


Fig. 6. Principal component analysis of 2183 peptides with a p-value < 0.01 for the average signal in the cancer and control group.

Detection of classes and clusters using self-organizing maps (SOM)

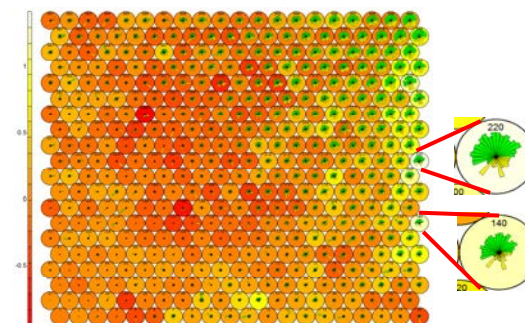


Fig. 7. SOM with 400 nodes. Colors reflect difference between mean signal of both groups. Stars show activity of sera (cancer/normal).

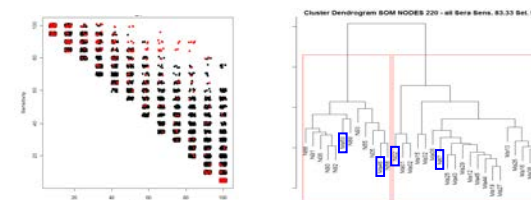


Fig. 8. Left panel: sensitivity and specificity for nodes of original dataset (red) and randomly assigned datasets (black). Cluster dendrogram of a node from Fig. 7. False positives/negatives are highlighted blue.

- Development of a high content and high throughput screening platform for biomarker discovery

platform for biomarker discovery

- Screening of serum samples results in specific signal patterns

References peptides were selected for validation

¹Polanski, M. & Anderson, N.L. (2007) A list of candidate cancer biomarkers for targeted proteomics. *Biomarkers* 12: 449-454

²Reischole, U., Ivancu, C., Schliep, M., Landerhof, C., Goetzke, S., Zahn, C., Hanzel, H., Volkmer-Engert, R., Schneider-Mergener, J. (2002) Identification of distinct antibody epitopes and mimotopes from a peptide array of 5520 randomly generated sequences. *J. Immunol. Methods* 267, 37-51.

Acknowledgements

We are indebted to Martin Löwer of TRON Mainz and Bernhard Renard of RKI Berlin for their intellectual input on data evaluation.

* Correspondence should be addressed to Ulf Reimer: reimer@jpt.com