Use of Non-Natural Sequence Peptide Arrays to Discover Cancer NeoAntigens For Vaccines and Diagnostics

Phillip Stafford, Josh Richer, Stephen Albert Johnston, Luhui Shen

Center for Innovations in Medicine, The Biodesign Institute, Arizona State University, Tempe, Arizona, 85287

Abstract

The humoral response to tumors could reveal tumor antigens recognized by immunosurveillance. We have developed the immunosignatures diagnostic technology to detect this tumor specific response. The technology profiles reactivity of sera antibodies on the HT330K chip with 330K addressable non-natural sequence peptides. We have shown that subjects with specific cancers have specific immunosignatures that are distinguished from healthy people and other cancers. Here we demonstrate that the signatures can be "deciphered" to indicate specific neo-antigens for any cancer.

Neo-antigens caused by mistranscription in tumors can be the sources of tumor antigens. We have set up a database of predicted chimeric tumor antigens encoded by trans-splicing transcripts analyzed from EST databases. These transcripts encode the neo-epitopes by producing frame shift (FS) peptides from shifted reading frames of the downstream genes. We hypothesize that these predicted neo-antigens are frequently expressed in cancers and recognized by the immune system. We predict that these neo-antigens can be deciphered by their immunosignatures.

To test the hypothesis, we assayed the sera from glioblastoma patients (GBM, n=16) and normal controls (NC, n=16) on the HT330K chip. The 1,000 most significantly different peptides between GBM and NC were selected. Recurrent motifs in these peptides were across searched in the neo-antigen database. The seven most frequently matched antigens were selected for further study: 2 of them were encoded by the in-frame fusion transcripts and others were encoded by the FS fusion transcripts.

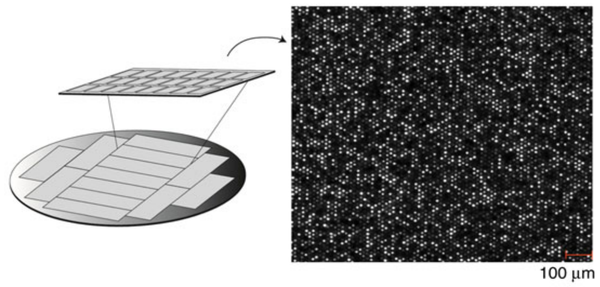
To validate the antibody reactivity of 7 deciphered neo-antigens, three overlapping peptides encompassing each of the antigens were synthesized and printed with unrelated peptides on the CIM10K array, which contains total 10,000 peptides. We analyzed NC (n=69) and GBM (n=19) with the CIM10K array. The positive cut off value of each peptide is the average plus two fold standard deviations of the NC. A positive sample was indicated as positive to at least one peptide of the antigen. The highest positive rates of 7 antigens is 11.6% in NC and 57.9% in GBM. 34.8% NC were positive to at least one antigen, 15.9% to more than one and only 4.3% positive to more than two antigens. No NC was positive to more than three antigens. This indicated the reactivity to these antigens is random in NC. The positive percentage in GBM were 89.5%, 73.7% and 47.4% respectively. 26.3% GBM was positive to more than three antigens. The implication is that cancer patients had much higher reactivity to these putative neo-antigens.

To test if other cancers elicited reactions to these neo-antigens, we analyzed breast cancer sera (BC, n=95) with the CIM10K. Similar to GBM, 83.15% BC were positive to at least one antigen and 34.7% were positive to more than 3 antigens. Five of seven antigens had positive reactivity in more than 25% of patients in both GBM and BC.

To further characterize these neo-antigens, we designed primers for RT-PCR analysis. We detected and sequence confirmed five fusion transcripts in GBM cDNAs. Two of the FS transcripts were also detected in multiple BC cDNAs and several normal tissues with lower level. Our interpretation is that the processing of the FS variants is primarily post-transcriptional in normal tissue but this processing is defective in tumors.

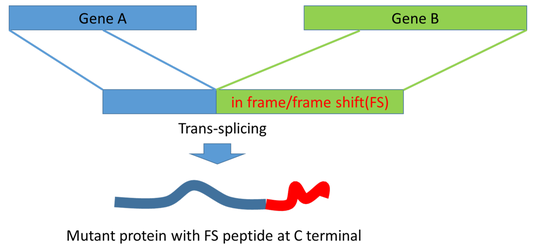
Here we showed that the neo-antigens from predicted fusion transcripts are the targets of the cancer humoral immune response. The specific antibody response was only in cancer patients indicating that the expression of these antigens in normal tissues was below the sensitivity of the immune system. Most importantly, this work indicates that we can efficiently decipher neo-antigens from the immunosignatures of cancer subjects. This may be a new source of antigens for diagnosis and cancer vaccines.

330K random peptide array

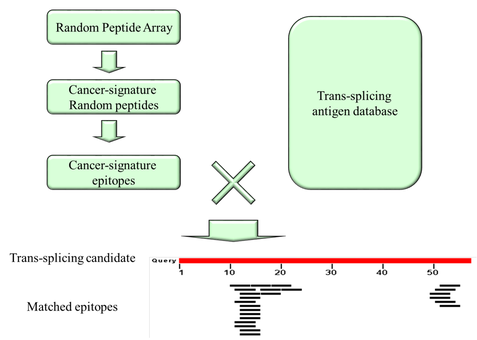


***In situ* synthesis of 330,000 random peptides for each array in 0.49 cm² on silicon wafers.** Each array arrays contain >27% of possible pentamers and 83% of possible tetramers[1].

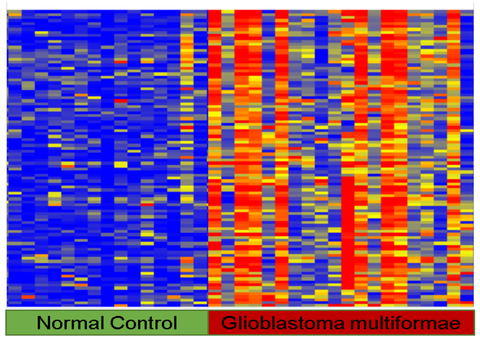
Trans-splicing encodes neo-antigen



Strategy for neo-antigen discovery



Analysis of 330K array and neo-antigen search



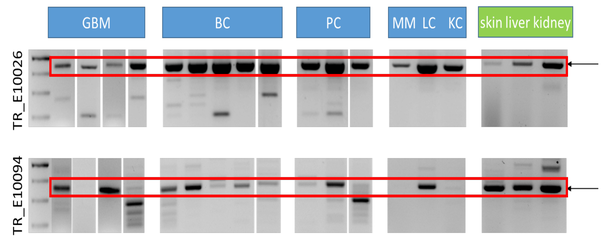
B

TR_E10501

ILYKRLKVPP**LSHPPGPGGVPGAL**MECQ**LHPG**GHPFPAPGAEAQGWRPV**WDVRPW**RA

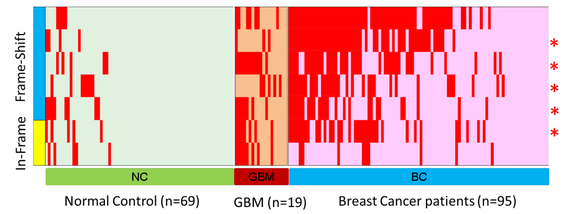
A. Heatmap of portion GBM specific immunosignature. The entire top 1,000 significant peptides were used to across search the trans-splicing antigen database. **B**. TR_E10501 is an example of the most frequently matched FS pepetide from the trans-splicing database. The red amino acids are the matched epitopes (4mer and 5mer). Some epitopes matched multiple times.

Sequence analysis of the predicted neo-antigens

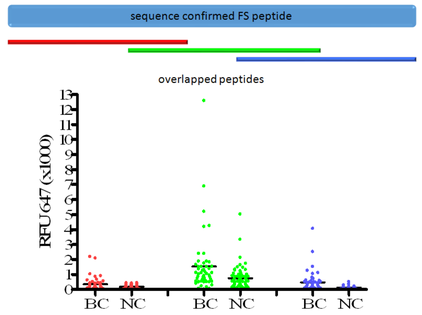


Example of RT-PCR analysis of FS candidates in cancer cell lines, primary cancers and normal tissues.

Verify the response with the original antigen peptides



Peptides of seven most frequently matched trans-splicing antigens were synthesized and printed on the array. All of the seven antigens presented higher IgG reactive in GBM and BC than in normal controls. The antigens labeled with red star were sequence confirmed in GBM cDNAs.



Example of sequence confirmed FS antigen that was further confirmed on the NSB9 surface array for high affinity and specific IgG analysis in both normal control and breast cancer patients[2].

Future Plan

Here we showed an efficient strategy for common cancer neo-antigen deciphering with a non-natural sequence peptide array. We will improve the method for epitope analysis and antigen database search for more accurate and efficient prediction. Further investigation of vaccine and cancer diagnosis with theses candidates is in the process.

Reference

1. Legutki, J. B., et al. (2014). "Scalable high-density peptide arrays for comprehensive health monitoring." Nat Commun **5**: 4785.

2. Stafford, P., et al. (2012). "Physical characterization of the "immunosignaturing effect"." Mol Cell Proteomics **11**(4): M111 011593.