

High-throughput Computational and Experimental Biology Strategy in Identifying Tumor Expressing CAMs

Anguraj Sadanandam, Michelle L. Varney and Rakesh K. Singh
 Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE

Email: Asadanandam@unmc.edu

Organ-specific homing of tumor cells includes receptor-ligand interactions of cell adhesion molecules (CAMs), which stimulate or inhibit the cellular growth. In this report we identified unique CAMs using sequence and pattern analysis with organ-specific peptides as query against mouse genome and proteome sequences. The peptides were seven amino acids in length, and were affinity selected against specific organs utilizing *in vivo* phage display peptide library in NOD-SCID mice. Various databases were used for the analyses including Local Mouse Cell Adhesion Molecule (LMCAM) database developed by keyword search. Thirty annotated CAMs corresponding to eleven different organ-specific peptides were identified using the bioinformatics analysis. One such identified protein, SEMA5A is reported for the first time to be expressed in human pancreatic cancer cell lines. This combined strategy of experimental and computation biology is an initial approach in identifying novel tumor-specific molecules, thereby paving the way for complete understanding of their role in various processes of tumor metastasis and making organ-specific targeting possible using these peptides.

Background: Metastasis is a complex disease with devastating effects. However, most patients succumb not to the primary tumor, but as a result of the metastatic lesions. The pattern of metastasis is predictable; and depends on the intrinsic properties of the tumor and its cross-talk with the specific host microenvironment. This cross-talk involves the interaction of Cell Adhesion Molecules (CAMs) specific for tumors and organs. These interactions determine the organ-specific metastasis. Certain residues called Critical residues (4 – 7 amino acids) are essential for these CAM specific interactions, whereas other surrounding residues act as scaffold. These critical residues can be used as antagonist against organ-specific microenvironment to block metastasis. Therefore, our objective is to identify - these residues utilizing phage-display peptide library and the CAMs containing these peptides using bioinformatics strategies,

and to confirm their tumor specific expression.

Materials and Methods: The organ-specific peptides were selected using phage-display library and *in vivo* biopanning in NOD-SCID mice devoid of phagocytic cells. We use different computational strategy to identify the unique CAMs using these organ-specific peptide sequences (Figure 1) and processed through different filtering algorithms. (Figure2).

Results and Discussion: *In vivo* screening of a heptapeptide library of M13 phage resulted in eleven unique and redundant peptides from 100 random clones, and they were found to bind four different organs (Figure 3). Among the screened peptides, two were unique for liver, four for lung, three for bone marrow

Figure 1.

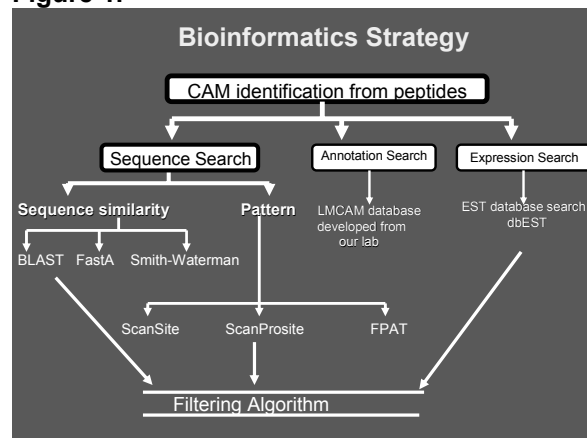
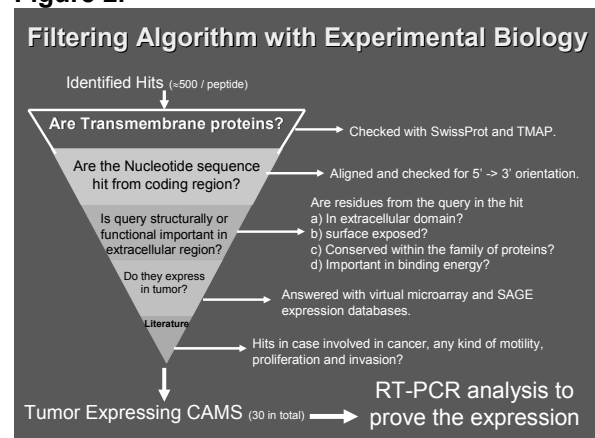


Figure 2.



and one for brain. Phages containing the peptides, NAFTPDY and KXXQVY[V/G/A] (X – any amino acid, [X/B] – any of the given amino acid) was found common for three of the four organs – liver, bone marrow and brain. The reason for these common peptides in liver and bone marrow may be that they are the common sites of metastasis, but that for the brain is unknown.

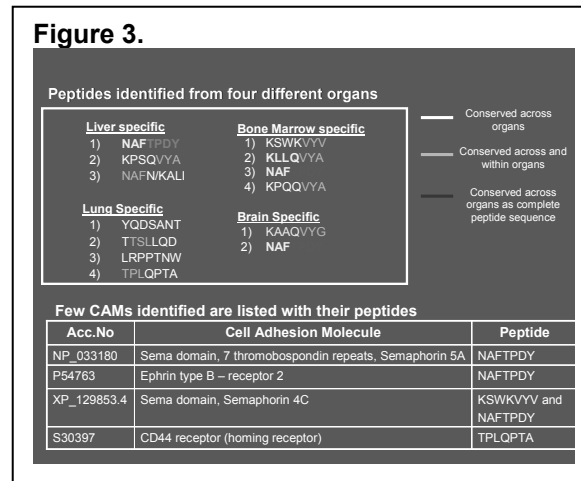
Grouping conserved peptides: The Clustal W alignment of peptides showed patterns among the sequences, and they were grouped not only based on the sequence but also based on within or among organ patterns. One alignment of peptides from all the four organs revealed that the tripeptides NAF and VY[V/G/A] ([X/B] – any of the given amino acids) were conserved among three organs – bone marrow, brain and liver. Also, these peptides were conserved within the organs - liver and bone marrow respectively. Further, the tetrapeptide K[A/L, A/L]Q was conserved among brain-specific and bone marrow-specific peptides. It is interesting to note that lung-specific peptides did not have any continuous peptide motifs conserved. These less conserved patterns may be due to greater diversity of peptides within the phage population in the peptide library. In addition, the discontinuous tri and hepta peptides, TXL and KXXQVY[V/G/A] (X – any amino acid) were conserved within the lung-specific and bone marrow-specific peptides respectively, and these motifs were searched using regular expressions. The conserved patterns of the peptides within and among organs may represent the binding of regular CAMs to the organ-specific proteins whereas non-conserved peptides, such as those in the lung represent the binding of tissue-specific CAMs.

Identifying specific CAMs: The filtering algorithm was used to identify specific hits to proteins involved in organ-specific interaction of cancer cells. First, the proteins were subjected to a process that eliminated those that were not transmembrane proteins. This was done because it was believed that phage interacted with CAMs present on the surface of the endothelial cells of the organs, and CAMs are transmembrane in nature. Second, the transmembrane proteins that showed match in the coding strand were selected. Third, it was important that the residues that matched should present in the extracellular domains and critical for binding. This is based on the idea that if the residues of the peptide ligands are surface exposed and conserved throughout its protein family then the chances that they play a functional or structural role in the identified protein hit is high [1]. This is an indirect evidence for the peptides to mimic its natural binding partners [1] and the direct evidence will be from site-directed mutagenesis. Fourth, these proteins were subjected to a virtual expression database search to make sure that they were of importance in cancer metastasis. Lastly, an extensive

literature search was done to confirm the results from the fourth step.

One example of a protein that remained the peptide, NAFTP in its extracellular domain was Semaphorin 5A (SEMA5A). Homology modelling of SEMA5A with identified similar protein, SEMA4D [1] showed that Phenylalanine (F) is surface exposed. The peptide is found to be conserved throughout the Semaphorin family of proteins. The virtual expression of

Figure 3.



Sema5A using the NCI60 database and one of the EST file showed that it is expressed in cells originated from metastases including PC-3 (prostrate cancer cell line) and Capan-1 (pancreatic cancer cell line). Sema4c, a protein of same Semaphorin family was identified by two different organ-specific peptides – liver and bone marrow. The virtual expression of Sema4C showed high expression in breast cancer. The structural and functional characteristic of SEMA4C is similar to that of SEMA4D [1]. Another hit resulted in the identification of Ephrin family of proteins, which have conserved NAFTPDY peptide in a discontinuous fashion. Among the residues of the peptide, Aspartic acid (D) is surface exposed and important for binding with its receptor [2].

In conclusion, we have develop unique approaches combining phage- display technology and bioinformatics to identify tumor-specific molecules, thereby paving the way for complete understanding of role of the identified proteins in various steps of tumor metastasis and making organ specific targeting possible using these peptides.

References:

1. Loce CA et al., Nat. Struct.Biol. 2003, 10:843-808.
2. Himanen JP et al., Nature, 1998, 396:486-491.