

Bootstrap Searching LC-MS/MS Spectra for Peptides Derived from t(11;22) Reciprocal Translocation in Brain Cancer

Lohith Madireddy

Graduate Student, Computational Biosciences Program

Arizona State University

Abstract

Cancer is not one disease, but many different diseases with a common theme of genetic instability. Glioblastoma Multiforme (GBM) arises from glial cells and is the most common form of brain cancer. GBM often recurs in patients post surgery, radiotherapy and/or chemotherapy. In a subset of patients with recurrent glioma who were treated with radiotherapy, a translocation between chromosome 11 and chromosome 22 was observed (Scheck, Hank, & Chavez, 2004). One hypothesis is that the fusion protein is responsible for tumor growth. I developed a novel method involving bioinformatics to be used with mass spectrometry data from brain tumor-derived peptides to identify the exact break points in chromosomes 11 and 22 with the goal of indentifying the new fusion proteins generated by this translocation. A Perl program was written and run on Saguaro supercomputer to combinatorially generate all fusion polypeptides from the confined break point regions of two chromosomes. Bootstrap methodology was implemented in Python language to randomly sample 1/10,000th number of peptides from the combinatorially generated large fusion polypeptide library of about 2.2 billion sequences. These comparatively smaller bootstrap peptides libraries were searched across the mass spectral data using Agilent Spectrum Mill software to find peptides containing amino acids from chromosomes 11 and

22. The bootstrapping and searching is repeated 100 times to have a statistically reliable estimate of hit rate from these peptide libraries. Once a fusion peptide is identified, DNA sequencing can be performed to identify the start and stop sites of the fusion protein produced as a result of the translocation. Then experiments will be performed to determine if the fusion protein is driving the growth of the tumor.