Use of simple PSSM mutation scoring in the CAGI2012 P16 Challenge

A simple sequence-only scoring function was used to make predictions for the P16 challenge. Firstly a two iteration PSI-BLAST search was carried out against the UNIREF90 sequence data bank with the target sequence (p16INK4a) and a stringent E-value cutoff of 10-16. A full precision PSSM file was then generated using the Impala makemtx program. For each given mutation, the difference between the native amino acid PSSM score and the mutant PSSM score was calculated and then scaled in the range 0.5-1.0 by least-squares regression with a logistic function to the provided training data (Pearson r=0.77).