ASA Assignment 2

Describe briefly but clearly mode of action:

Performed BLAST search using NCBI on the Swissprot/Uniprot database using the query sequence ID 7160686.

How you did what you did:

From the main BLAST NCBI page, clicked on protein blast under the Basic BLAST header

Entered the query sequence accession number: 7160686

Changed the database to UniProtKB/Swiss-Prot(swissprot)

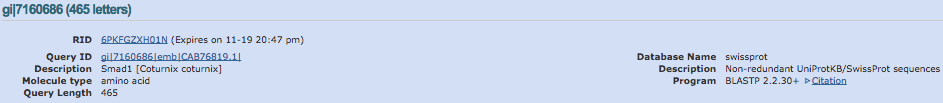
Selected blastp algorithm

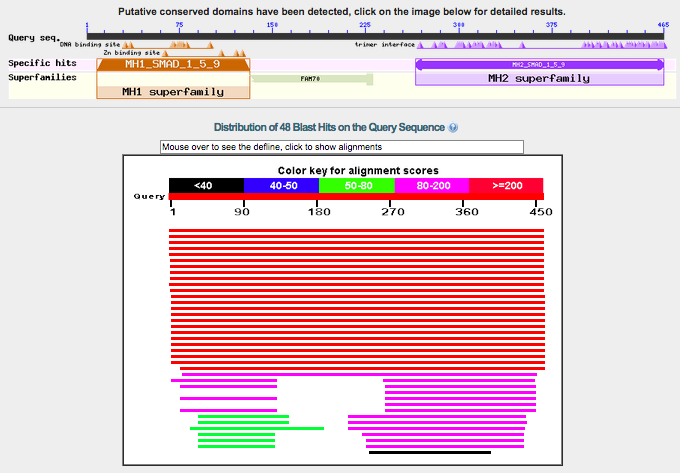
Clicked BLAST

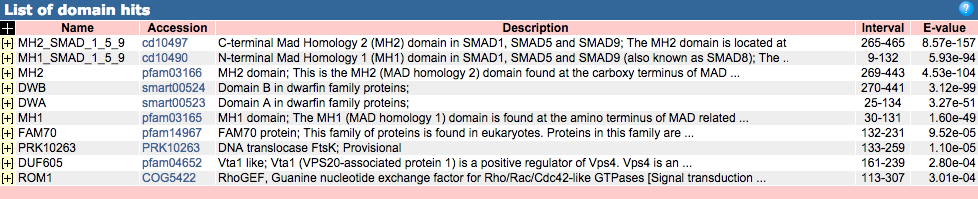
What you based your decisions and conclusions:

BLAST

38 sequences with significant alignments

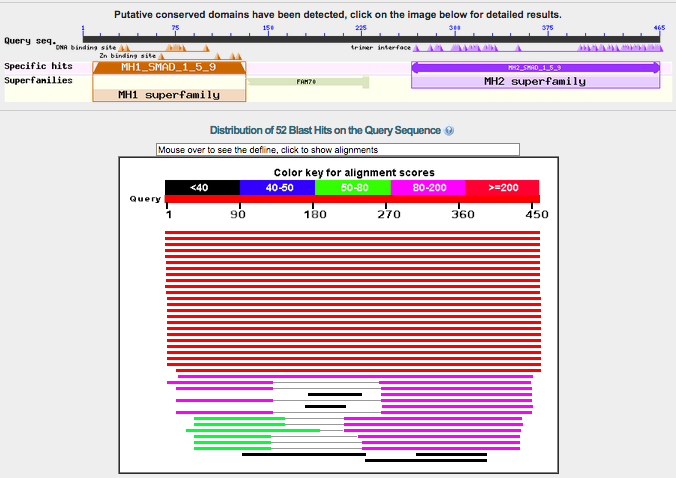






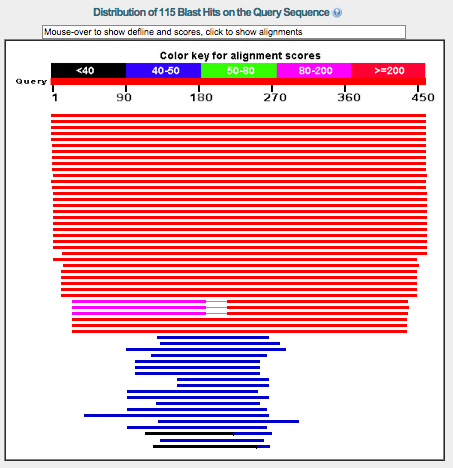
PSI-BLAST Iteration 1 with max 500





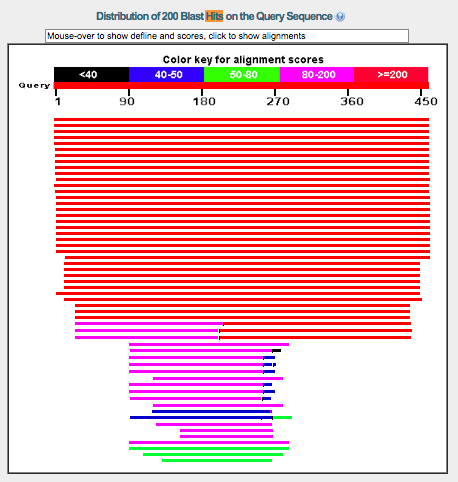
PSI-BLAST iteration 2 with max 500

37 hits



49 hits

PSI-BLAST iteration 3



173 significant alignments

On consecutive PSI-Blast iterations you will see other hits appearing; are they all from the SMAD family?

No they are not all from the SMAD family (take picture)

If not, what do you think is happening here? How many more useful sequences did PSI-Blast give you compared to BLAST?

PSI-BLAST detects distant homologs. PSI-BLAST is finding hits that are in different families. The more hits is important because we are getting more homologs because profile constructing properties of PSI-BLAST (elaborate more later)

The SMAD proteins contain a DNA-binding domain (MH1). Also many of the PSI-Blast hits are DNA binding (such as, transcription elongation factor, Zinc Finger, transcription factor, helicase). Based on that, could you consider these hits as putative (very) distant homologs of the SMAD proteins?

Potentially.

From the Blast output, select all SMAD proteins and download them as FASTA; you need them unaligned but it doesn’t hurt to also download the BLAST (master-slave) alignment.

Done

Before you start making and analyzing your own alignments, you can have a sneak preview of the specificity detection from the Multi-Harmony webserver. Run the Example Dataset (this happens to be the SMAD family). For easy comparison of the alignment column (residue) numbering, the example includes ‘1KHX’ as a reference PDB structure to the webserver advanced input – this adds a column with the corresponding residue numbering to the output table. The resulting output table should look similar to the one in the Pirovano *et al.* (2006) paper.

Part 2: Alignment and Analysis and specificity scoring

Praline-PSI-BLAST

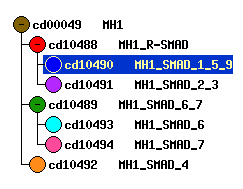
E-value 0.01, iterations: 3,

Secondary structure prediction: PSIRED

Transmembrane structure prediction: PHOBIUS

Report briefly on where your (four) alignments differ most, and which parts are consistently aligned in a similar way irrespective of alignment method. You may want to review the domain structure as reported by PSI-BLAST. When comparing using the alignment viewer (JalView), please be aware that alignment tools typically change the order of sequences in the alignment produced.

Subfamily hierarchy from PSI-BLAST output:



MH1:

**N-terminal Mad Homology 1 (MH1) domain in SMAD1, SMAD5 and SMAD9 (also known as SMAD8)**

The MH1 is a small DNA-binding domain present in SMAD (small mothers against decapentaplegic) family of proteins, which are signal transducers and transcriptional modulators that mediate multiple signaling pathways. MH1 binds to the DNA major groove in an unusual manner via a beta hairpin structure. It negatively regulates the functions of the MH2 domain, the C-terminal domain of SMAD. This MH1 domain is found in SMAD1, SMAD5 and SMAD9, all closely related receptor regulated SMADs (R-SMADs). SMAD1 plays an essential role in bone development and postnatal bone formation through activation by bone morphogenetic protein (BMP) type 1 receptor kinase. SMAD5 is involved in bone morphogenetic proteins (BMP) signal modulation and may also play a role in the pathway involving inhibition of hematopoietic progenitor cells by TGF-beta. SMAD9 mediates the differentiation of mesenchymal stem cells (MSCs) into tendon-like cells by inhibiting the osteogenic pathway.

MH2:

**C-terminal Mad Homology 2 (MH2) domain in SMAD1, SMAD5 and SMAD9**

The MH2 domain is located at the C-terminus of the SMAD (small mothers against decapentaplegic) family of proteins, which are signal transducers and transcriptional modulators that mediate multiple signaling pathways. The MH2 domain is responsible for type I receptor interaction, phosphorylation-triggered homo- and hetero-oligomerization, and transactivation. It is negatively regulated by the N-terminal MH1 domain, which prevents it from forming a complex with SMAD4. SMAD1, SMAD5 and SMAD9 (also known as SMAD8), are receptor regulated SMADs (R-SMADs). SMAD1 plays an essential role in bone development and postnatal bone formation through activation by bone morphogenetic protein (BMP) type 1 receptor kinase. SMAD5 is involved in BMP signal modulation and may also play a role in the pathway involving inhibition of hematopoietic progenitor cells by TGF-beta. SMAD9 mediates the differentiation of mesenchymal stem cells (MSCs) into tendon-like cells by inhibiting the osteogenic pathway.

Praline quality sum : 53692.75491318107

ClustalW quality sum: 58362.0244615078

MUSCLE quality sum: 61809.965611457825

PSI-Praline quality sum: 63356.90463542938