**Project 3 of BIO316 – Proteomics Data Analysis**

**Essay 1:** The students will be credited if all the steps and corresponding explanations (as illustrated in the following chart) are described properly in their essays.

Key points for significance and potential application (the students should at least mention the following points):

1. To provide knowledge of roles in biological processes, molecular functions and cellular components for each target protein.
2. To provide knowledge of significantly altered biological processes, molecular functions and cellular components and associated target proteins.

**Target Proteins**

**BLAST**

Proteins with similar amino acid sequences will probably exhibit similar functions (similarity should be at least over 30%). Identification of homologues sequences with known functions is of usage for making educated guess of unknown protein function.

**GO Mapping**

**GO Annotation**

For each candidate GO an annotation score is computed by applying certain annotation rule. GO terms above pre-set threshold will be assigned to the target protein.

**Annotation Augmentation**

**Results Returned**

Including 1) retrieving domain / motif information by InterProScan and transferring corresponding GO terms and 2) augment annotation by ANNEX

**Essay 2:** The students will be credited if all the steps and corresponding explanations (as illustrated in the following chart) are described properly in their essays.

Key points for significance and potential application (the students should at least mention the following points):

1. To provide knowledge of roles in biological pathways for each target protein.
2. To provide knowledge of significantly altered biological pathways and associated target proteins.

**Grouping by KO**

**Scoring by Probability and Heuristics**

**Cut off by Bi-Directional Hit Rate**

**Blast to KEGG GENES**

**Query protein**

**Homologs**

**Ranking of KO**

**Pathway Mapping**

**Ortholog Candidates**

**KO groups**

**Essay 3:** The students should at least mention the following points:

1. In biological systems, proteins do not exist individually but exhibit their functions by interacting with and modulating other molecules.
2. Clustered / connected proteins might share similarities at functional level.
3. The central nodes with high connectivity might be key modulator in the network and biological system.