# Practical Bioinformatics
## (BIPG511/711) Summer I – 1 credit
### Instructor: Robert J. Trumbly, Ph.D.
### Tues-Thurs: 2:00 PM-4:00 PM Room: HEB119

Session 1. NCBI databases
1. Using NCBI resources: Entrez
2. Knowing location of important databases on the internet. Are there important resources in addition to NCBI?
3. Understanding the different types of genomic and sequence databases: Genbank, Entrez, Gene, Refseq, Unigene

Session 2. Resources at the European Bioinformatics Institute (EBI) and Ensembl.
1. Databases at EBI: EMBL-Bank, UniProt, ArrayExpress, InterPro
2. Ensembl- genome browser.
3. Biomart. An interface that allows access to related tables of information

Session 3. UCSC Genome Browser
1. Use BLAT to locate sequences within their genomic context.
2. Download sequences corresponding to specific gene regions: exons, intron, promoters
3. Examine sequence conservation.
4. Use CSC Table Browser to retrieve genome-associated information.
5. Create custom tracks in the UCSC Table Browser.

Session 4. Sequence alignment
1. Dot matrix sequence comparisons.
2. Amino acid substitution matrices: PAM and BLOSUM.
3. Global and local alignments by dynamic programming.

Session 5 BLAST & database searching
1. Understanding the BLAST algorithm
2. Evaluating significance of BLAST searches.
3. Position-Specific Iterated (PSI)-BLAST: creating sequence profiles for database searching.

Session 6 Multiple sequence alignment (MSA)
1. Algorithms for MSA.
2. Recent developments in MSA.
3. Performing MSA with CLUSTAL X.

Session 7. Microarray analysis of gene expression I
1. Comparison of microarray platforms, recent developments
2. Identifying differential gene expression
3. Accessing microarray databases, primarily GEO (Gene Expression Omnibus)

Session 8. Microarray analysis of gene expression II
1. Comparison and annotation of gene lists.
2. Overrepresentation  analysis: enrichment of functional categories in co-expressed genes.
3. Gene ontology and pathways.

Session 9. Transcription factor binding sites n(TFBS)

1. Defining TFBS with increasing information: from consensus to PSSMs
2, Identifying TFBS in sequences: TRANSFAC and MATCH
3. Finding over-represented TFBS in coexpressed genes and chip-chip peaks

Session 10. Galaxy tools: a website for retrieval and comparison of genomic sequences and annotations.
1. Retrieve sequences from genome corresponding to given sequence coordinates.
2. Compare sequence regions of chip-chip peaks and regions upstream of genes.

Session 11. Taverna: access to web services
1. Introduction to web services
2. Integrating web services
3. Work flows