**Course: Advanced Bioinformatics**

**Module title: IR in biological Databases**

**Module no. : 171**

Biological data can have following type of information.

* Evolutionary info.
* Genomic info.
* Structural info.
* Expression info.
* Functional info.

Due to such variety, retrieval of biological data using current IR approaches in not straight forward. There are specialized IR systems built for retrieval of biological data. Several types of databases exist for biological data such as

* Sequence databases
* Sequence analysis
* Functional genomics
* Literature databases
* Structural databases
* Metabolic pathway databases
* Specialized databases

Searching for information about genes and their prod

Gene and gene product databases are often organized by sequence. Genomic sequence encodes all traits of an organism. Gene products are uniquely described by their sequences. Similar sequences among biomolecules indicates both similar function and an evolutionary relationship

Macromolecular sequences provide biologically meaningful keys for searching databases.

Exact matches are rare (even uninteresting in many cases), so often goal is to retrieve a set of similar sequences. Both small (mutations) and large (required for function) differences within “similar” can be interesting

Information about Gene:

* Genomic (chromosomal location, allelic information, regulatory regions, etc.)
* Structural (known structure? structural domains? etc.)
* Functional (molecular, cellular & disease)