**Course: Advance Bio Informatics**

**Module Title: Databank for Biological Data**

**Module No: 81**

**Databank for Biological Data**

Huge Volume of biological data available due to contribution of research community and high throughput experiments. Biological data has evolutionary nature. Type of data is not known. Eco system can vary the behavior of biological data.

Structure of biological data varies based on eco system. Data is explorative and iterative nature. Different organizations have developed dbs.

These databases can be categorized into two types:

**Primary Databases**

Submitter of the original data has control over it. i.e. he can modify it. e.g. GenBank, Databank of Japan (DDBJ), and IEMBL Nucleotide Format (EMBL).

**Secondary Databases**

Contains results obtained from the analysis of primary databases. RefSeq, UniGene and RefSNP.

These databases use different conventions and formats to store biological data.

**Biological Data Formats**

Plain Sequence Format, FASTA Format, EMBL Format, GCG Format, GCG-RSF (rich sequence format), GenBank Format, IG Format are some of these formats.

**Need of Databank:**

Researchers often need to consult more than one database for their experimental work. Databank can solve their problems in a way that all related data is stored in a single physical database instead of disperse places.