**Course: Advanced Bioinformatics**

**Module title: Sequence Input Output**

**Module no. : 195**

This document represents the Sequences in BioJava and Input Output Operations.

No need to write a new parser if we can avoid it

* BioJava supports
  + Genbank, GenPept, RefSeq, EMBL, SwissProt, PDB, Fasta, ABI, LocusLink, Unigene (requires Java 1.4)
  + GAME, AGAVE
  + Blast, Fasta, HMMER (models and results), BlastXML, MEME, Phred
  + OBDA, BioIndex, BioSQL, DAS, GFF, XFF
  + Ensembl (with biojava-ensembl package)
* StAX/ Tag value
* RMI and Serialization

Simple I/O

* Most of BioJava’s simpler I/O operations are conveniently wrapped up behind static methods from the SeqIOTools class.
* SeqIOTools can read and write:
  + Fasta (protein or DNA)
  + EMBL
  + GenBank (flat file and XML)
  + SwissProt
  + GenPept
  + MSF (protein or DNA)
  + Fasta Alignments

Sequence I/O

SeqIOTools Reader Methods

SequenceIterator i = SeqIOTools.readGenbank(br);

SequenceIterator i = SeqIOTools.readGenpept(br);

SequenceIterator i = SeqIOTools.readSwissprot(br);

SequenceIterator i = SeqIOTools.readEmbl(br);

etc...

SequenceIterator i = (SequenceIterator)

SeqIOTools.fileToBiojava("fasta", "dna", br);

Alignment a =

(Alignment) SeqIOTools.fileToBiojava("MSF", "rna", br);

SeqIOTools Reader Methods

SequenceIterator i = SeqIOTools.readGenbank(br);

SequenceIterator i = SeqIOTools.readGenpept(br);

SequenceIterator i = SeqIOTools.readSwissprot(br);

SequenceIterator i = SeqIOTools.readEmbl(br);

etc…

SequenceIterator i = (SequenceIterator) SeqIOTools.fileToBiojava("fasta", "dna“, br);

Alignment a =

(Alignment) SeqIOTools.fileToBiojava(“MSF", “rna“, br);