



LEARN PYTHON & R FOR BIOINFORMATICS

Prerequisite Terminologies:

In order to have a better understanding of the main topic, you should have the basic concept of the following terms:

- **Genome Browsers**
- **Comparative Genomics**
- **Sequence Variation**
- **Transcriptional Regulation**

Introduction:

Ensembl is a genome browser for vertebrate genomes that support research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl provides annotated genes information, computes multiple alignments, predicts regulatory function and collects disease data. Ensembl aims to provide a centralized resource for geneticists, molecular biologists and other researchers studying the genomes of our own species, vertebrates and model organisms. Ensembl

is one of the several well known genome browsers for the retrieval of genomic information.

Steps:

- Open the homepage of Ensembl database, which you can access from [here](#).
- Ensembl provides a simple interface for the users. On the top of the homepage of Ensembl, you'll find the following tools provided by Ensembl:

BioMart	→ Data-mining tool, exports custom datasets from Ensembl.
BLAST/BLAT	→ Search genomes for your DNA or protein sequence.
VEP	→ Variant Effect Predictor, analyse variants and predict the functional consequences of known and unknown variants.

- For more tools information provided by Ensembl, click on the '**Tools**' option in the menu bar.
- On the right side of the homepage of Ensemble, you'll find **Ensembl Release 100**, which provides following information:
 - GENCODE, which identifies and maps all protein-coding genes within the ENCODE regions.
 - gnomAD, which determines the genomic allele frequencies.
 - New genomes information.
 - Updates genomes.
 - New interface for configuration of multidimensional track hubs.
- In the '**Search**' box, from '**All Species**' bar, you can select any species of vertebrate and get the whole genomic information for it.

- You can also search for specific information instead of genomes such as any particular chromosomal region, disease keyword, gene name, protein name or some sort of SNP information from the search bar below the 'All species'.
- In the 'Favourite Genomes' option, you'll find those genomes that are widely used for the analysis and research work.
- There are other advanced options offers by Ensembl database such as:
 - Compare genes across species
 - Find SNPs and other variants from my gene
 - Gene expression in different tissues
 - Retrieve gene sequence
 - Find a Data Display
 - Use my own data in Ensembl
- Search for the 'BRCA2' which is a breast cancer gene just to demonstrate how Ensembl provides information about a particular search.
 - Once you click on 'BRCA2', you'll get the result summary.
 - Just like other browsers, it'll provide various searches for that particular gene in 'Restriction Category' such as:
 - Gene
 - Transcripts
 - Variants
 - Phenotype
 - Gene Markers and other sorts of information.
 - You can also select information according to 'Restrict species to' options.
 - In the '**Layout**' option, select '**Table**' to visualize the information in a modern way.
 - It provides results in a tabular form, where Ensembl ID, name, species, category and description are given.
 - Select a particular gene from the 'Result Summary' page.

- You'll see various information about that particular gene is given on the result page.

Summary:

In this introductory video of Ensembl database, we came to know about what sort of information does Ensembl database provide. We have learned about its various features and tools that can be utilized for the particular search. We also got to know about the results summary information about a particular gene provided by the Ensembl database.