



Centurion
UNIVERSITY
*Shaping Lives...
Empowering Communities...*

Ensembl Genomes





Centurion
UNIVERSITY

Shaping Lives...
Empowering Communities...

History:

- 1999: Start of Ensembl project for the human genome project.
- 2000: First release of data and web interface.
- 2009: First release of Ensembl Genomes.
- 2011: Ensemble v65 : 63 genomes.
- 2011: Ensembl genomes v12 : 335 genomes.
- Ensembl: EBI and welcome trust Sanger Institute.
- Ensembl Genomes : EBI



Centurion
UNIVERSITY

*Shaping Lives...
Empowering Communities...*

What is Ensembl Genomes?

- New portal offering integrated access to genome-scale data for non-vertebrate species, developed using the Ensembl genome annotation and visualization platform.
- Consist of five sub portals (for bacteria, protists, fungi, plants and invertebrate metazoa) designed to complement the availability of vertebrate genomes in Ensembl.



Centurion
UNIVERSITY

Shaping Lives...
Empowering Communities...

Types:

- Ensembl Genome consists of :
 - a) EnsemblFungi
 - b) EnsemblProtists
 - c) EnsemblePlants
 - d) EnsembleMetazoa
 - e) EnsemblBacteria
- Each site contain data for selected species from their domain, chosen for their scientific interest.
- Information available about each species including the assembly version and annotation methods used, the overall composition of the genome.



Centurion
UNIVERSITY

*Shaping Lives...
Empowering Communities...*

Annotation:

- Inclusion of species depends on various criteria (model Organisms? / community interest / demand? Funding? Completeness / quality of genome assembly ?)
- A broad taxonomic coverage is aimed for.

Ensembl:

- Annotation in-house by the Ensembl team.

Ensembl Genomes:

- Annotation preferably by or in collaboration with the scientific community for the species in question.



Centurion
UNIVERSITY

*Shaping Lives...
Empowering Communities...*

Data:

- Genomic sequences.
- Gene/ transcript / protein models.
- External reference.
- Mapped cDNAs, proteins, microarray probes, BACs, markers etc.
- Comparative Data: orthologs and paralogs, protein families
- Variation Data: sequence variant , structuralvariant.
- Regulatory Data: "best guess" set of regulatory elements.



Centurion
UNIVERSITY

*Shaping Lives...
Empowering Communities...*

Thank You