**295123 - 2951333 - Bioinformatics**

**Teaching staff**

**Coordinator:** Samir Kanaan Izquierdo

**Others:** María de los Ángeles Maqueda González

**Prior skills**

Basic Programming Skills

**Requirements**

Data analysis and pattern recognition

**Degree competences to which the subject contributes**

**Specific:**

CEMUEII-18. Design innovative solutions in biomedicine through the use of computer tools for design, modeling and computational simulation (Specific competence of the Healthcare and Biomedical Applications specialty).

**Generical:**

CGMUEII-01. Participate in technological innovation projects in multidisciplinary problems, applying mathematical, analytical, scientific, instrumental, technological and management knowledge.

CGMUEII-05. To communicate hypotheses, procedures and results to specialized and non-specialized audiences in a clear and unambiguous way, both orally and through reports and diagrams, in the context of the development of technical solutions for problems of an interdisciplinary nature.

**Transversal:**

05 TEQ. TEAMWORK. Being able to work as a team player, either as a member or as a leader. Contributing to projects pragmatically and responsibly, by reaching commitments in accordance to the resources that are available.

06 URI. EFFECTIVE USE OF INFORMATION RESOURCES. Managing the acquisition, structure, analysis and display of information from the own field of specialization. Taking a critical stance with regard to the results obtained.

03 TLG. THIRD LANGUAGE. Learning a third language, preferably English, to a degree of oral and written fluency that fits in with the future needs of the graduates of each course.

**Teaching methodology**

The methodology of the course combines theory lessons, laboratory sessions and autonomous learning through the development of projects and the analysis of real applications.

**Learning objectives of the subject**

- Properly design a bioinformatics analysis workflow, paying attention to the specific issues of the input data type to be analyzed
### Study load

<table>
<thead>
<tr>
<th></th>
<th>Total learning time: 150h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours large group:</td>
<td>22h</td>
</tr>
<tr>
<td>Hours medium group:</td>
<td>0h</td>
</tr>
<tr>
<td>Hours small group:</td>
<td>22h</td>
</tr>
<tr>
<td>Guided activities:</td>
<td>4h</td>
</tr>
<tr>
<td>Self study:</td>
<td>102h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>14.67%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>14.67%</td>
</tr>
<tr>
<td></td>
<td>2.67%</td>
</tr>
<tr>
<td></td>
<td>68.00%</td>
</tr>
</tbody>
</table>
### Content

<table>
<thead>
<tr>
<th><strong>Biology background</strong></th>
<th><strong>Learning time:</strong> 18h</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Theory classes: 4h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 2h</td>
</tr>
<tr>
<td></td>
<td>Self study : 12h</td>
</tr>
</tbody>
</table>

**Related activities:**
- Home activity 1: Analysis of mRNA sequences using command-line
- Laboratory session 1: Genome databases and browsers

**Specific objectives:**
- Learn key aspects of genetics: structure and function of nucleic acids and chromosomes.
- Learn the influence of genetics on diseases
- Know current types of functional genomic data: from 1D to 3D genome perspectives.
- Understand basic annotations in a genome. GENCODE project.
- Understand and interpret the information retrieved from a genome browser.

<table>
<thead>
<tr>
<th><strong>Sequence alignment</strong></th>
<th><strong>Learning time:</strong> 26h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theory classes: 4h</td>
</tr>
<tr>
<td></td>
<td>Laboratory classes: 6h</td>
</tr>
<tr>
<td></td>
<td>Self study : 16h</td>
</tr>
</tbody>
</table>

**Description:**
- Local and global alignment algorithms
- Database similarity searches
- Protein motifs

**Related activities:**
- Understand the implications of carrying out a local or global alignment.
- Know the family of BLAST algorithms, local alignment tools.
- Get familiar with substitution matrices: PAM and BLOSUM families
- Understand the steps of progressive alignment.
- Identification of conserved motifs given several sequences.
## Analysis of Transcriptomics Data

**Learning time:** 26h  
Theory classes: 4h  
Laboratory classes: 6h  
Self study : 16h

**Description:**  
- Gene expression and regulation. Techniques and tools  
- Processing: normalization, filtering and clustering  
- Gene expression databases

**Related activities:**  
- Laboratory 4: microarrays  
- Laboratory 5: RNA-sequencing

## Knowledge databases and enrichment analysis

**Learning time:** 22h  
Theory classes: 4h  
Laboratory classes: 6h  
Self study : 12h

**Description:**  
- Biological pathways: KEGG, Reactome  
- Gene ontologies: GO  
- Protein interactions: STRINGdb  
- Enrichment analysis

**Related activities:**  
- Laboratory 6: Biological pathways and protein interactions  
- Laboratory 7: Gene Ontology. Enrichment analysis

## Advanced techniques and applications

**Learning time:** 20h  
Theory classes: 2h  
Laboratory classes: 2h  
Self study : 16h

**Description:**  
- Proteomics and metabolomics  
- Epigenomics  
- Machine learning methods

**Related activities:**  
- Laboratory 8: metabolomics analysis pipeline  
- Home assignment 2: epigenomics
Qualification system

Laboratory reports: 40%
Homework assignments: 20%
Final application project: 40%

Bibliography

Basic:


Others resources:

Hyperlink

KEGG
KEGG: Kyoto encyclopedia of genes and genomes

Reactome
Reactome

Bioconductor
Bioconductor

Gene Ontology
Gene Ontology