Teaching Guide				
Identifying Data				2022/23
Subject (*)	Advanced processing of biological sequences		Code	614522020
Study programme	Mestrado Universitario en Bioinformática para Ciencias da Saúde			
Descriptors				
Cycle	Period	Year	Type	Credits
Official Master's Degree	1st four-month period	Second	Optional	3
Language	SpanishGalicianEnglish			
Teaching method	Hybrid			
Prerequisites				
Department	Ciencias da Computación e Tecnoloxías da InformaciónComputación			
Coordinador	Bernardo Roca, Guillermo de	E-mail	guillermo.debernardo@udc.es	
Lecturers	Bernardo Roca, Guillermo de Santos Reyes, Jose	E-mail	guillermo.debernardo@udc.es jose.santos@udc.es	
Web	moodle.udc.es			
General description	The course introduces advanced data structures, algorithms and tools for the management of biological sequences. Particularly, it sintroduces compression techniques for biological sequences, graphs and networks, and techniques for protein structure prediction.			

Study programme competences	
Code	Study programme competences
A1	CE1 - Ability to know the scope of Bioinformatics and its most important aspects
A2	CE2 ? To define, evaluate and select the architecture and the most suitable software for solving a problem in the field of bioinformatics
A3	CE3 ? To analyze, design, develop, implement, verify and document efficient software solutions based on an adequate knowledge of the theories, models and techniques in the field of Bioinformatics
A6	CE6 - Ability to identify software tools and most relevant bioinformatics data sources, and acquire skill in their use
A8	CE8 - Understanding the basis of the information of the hereditary material, its transmission, analysis and evolution
A9	CE9 ? To understand the benefits and the problems associated with the sequencing and the use of biological sequences, as well as knowing the structures and techniques for their processing
B1	CB6 - Own and understand knowledge that can provide a base or opportunity to be original in the development and/or application of ideas, often in a context of research
B2	CB7 - Students should know how to apply the acquired knowledge and ability to problem solving in new environments or little known within broad (or multidisciplinary) contexts related to their field of study
B8	CG3 - Be able to work in a team, especially of interdisciplinary nature
C6	CT6 - To assess critically the knowledge, technology and information available to solve the problems they face to.
C7	CT7 ? To maintain and establish strategies for scientific updating as a criterion for professional improvement.

Learning outcomes			
Learning outcomes	Study programme competences		
To know the main state-of-the-art data structures for the compact and self-indexed representation of sequences, and algorithms to manage them.	AJ1 AJ2 AJ9		
To create compressed data structures to develop analysis and alignment tasks on sequences efficiently in time and space.	AJ2 AJ3 AJ6 AJ8	BJ1 BJ2 BJ8	CJ6 CJ7

To know the main issues associated secondary and tertiary protein structure prediction and their importance, as well as the main prediction techniques in the state of the art.	AJ1 AJ2 AJ3 AJ6 AJ9	BJ1	CJ6 CJ7

Contents	
Topic	**Sub-topic**
Compresión de secuencias biolóxicas	Lempel-Ziv Grammar-based compression
Biological sequence indexing	Burrows-Wheeler Transform FM-index Search and assembly applications
Succinct representation of graphs and biological networks	Data structures for compact graph representation Representation of biological networks Applications to biological sequences
Protein structure prediction	Basic concepts on proteins Secondary structure prediction with machine learning techniques Tertiary structure prediction Protein folding models

Planning				
Methodologies / tests	**Competencies**	**Ordinary class hours**	**Student?s personal work hours**	**Total hours**
Guest lecture / keynote speech	A1 A2 A3 A6 A8 A9	11	11	22
Mixed objective/subjective test	A1 A2 A3 A6 A8 A9 B2	4	0	4
ICT practicals	A1 A2 B1 B2 B8 C6 C7	10	38	48
Personalized attention		1	0	1
(*)The information in the planning table is for guidance only and does not take into account the heterogeneity of the students.				

Methodologies	
Methodologies	**Description**
Guest lecture / keynote speech	Lectures where the course contents are exposed
Mixed objective/subjective test	Test to show that the student has acquired the knowledge and skills required during lectures and practice sessions
ICT practicals	Students will complete, individually or in groups, different practical exercises to develop the concepts acquired in the lectures

Personalized attention	
Methodologies	**Description**
ICT practicals	There may be differences among students regarding the knowledge of algotrithms and tehcniques used in the course. Personalized attencion will be provided for all practical work deceloped by the students.

Assessment			
Methodologies	**Competencies**	**Description**	**Qualification**

Mixed objective/subjective test	A1 A2 A3 A6 A8 A9 B2	Constará dunha proba na que deben ser demostrados os coñecementos e competencias adquiridos. Para aprobar a materia globalmente hai que obter unha NOTA MÍNIMA de 1 (sobre 2) nesta proba. Non sendo así, a nota máxima global da materia non será en ningún caso superior a 4,0 e a materia considerarase suspensa.	20
ICT practicals	A1 A2 B1 B2 B8 C6 C7	Os estudantes deberán entregar boletíns cos resultados das prácticas realizadas ou solución aos problemas propostos.	80

Assessment comments
FIRST OPPORTUNITYSudents that do not take the test will obtain a grade of "Non presentado" (absent) SECOND OPPORTUNITYOnly students that have not passed the course in the first opportunity can be evaluated in the second opportunity. In the second opportunity, students that do not retake any part will obtain a grade of "Non presentado" (absent) ADVANCED OPPORTUNITY: The assessment for the advanced opportunity will consist of a written exam that will compute for the 100% of the grade, and will include the knowledge and skills acquired during lectures and practice sessions.ACADEMIC DISPENSATION: Students enrolled part-time with official dispensation from attending classes must contact the teachers within the first two weeks of the course to establish the condition for submitting and defending the practical exercises. PRIMEIRA OPORTUNIDADEOportunidade ganar

Sources of information	
Basic	- N. C. Jones, P. A. Pevzner (2004). An introduction to bioinformatics algorithms. MIT Press - A. Tramontano (2006). Protein structure prediction: Concepts and Applications. Wiley-VCH - V. Mäkinen, D. Belazzougui, F. Cunial, A.I.Tomescu (2015). Genome-scale algorithm design. Cambridge University Press
Complementary	- T.K. Attwood, D.J. Parry-Smith (2002). Introducción a la bioinformática. Pearson educacion

Recommendations
Subjects that it is recommended to have taken before
Data structures and algorithmics for biological sequences/614522013
Subjects that are recommended to be taken simultaneously
Subjects that continue the syllabus
Other comments

(*)The teaching guide is the document in which the URV publishes the information about all its courses. It is a public document and cannot be modified. Only in exceptional cases can it be revised by the competent agent or duly revised so that it is in line with current legislation.