



**Teaching Guide**

Identifying Data					2022/23
<b>Subject (*)</b>	Advanced statistical methods in bioinformatics	<b>Code</b>	614522009		
<b>Study programme</b>	Mestrado Universitario en Bioinformática para Ciencias da Saúde				
Descriptors					
<b>Cycle</b>	<b>Period</b>	<b>Year</b>	<b>Type</b>	<b>Credits</b>	
Official Master's Degree	2nd four-month period	First	Obligatory	6	
<b>Language</b>	SpanishEnglish				
<b>Teaching method</b>	Face-to-face				
<b>Prerequisites</b>					
<b>Department</b>	Matemáticas				
<b>Coordinador</b>	Lopez de Ullibarri Galparsoro, Ignacio	<b>E-mail</b>	ignacio.lopezdeullibbarri@udc.es		
<b>Lecturers</b>	Lopez de Ullibarri Galparsoro, Ignacio	<b>E-mail</b>	ignacio.lopezdeullibbarri@udc.es		
<b>Web</b>	www.master.bioinformatica.fic.udc.es				
<b>General description</b>	Preténdese proporcionar ó alumnado os coñecementos necesarios para abordar unha selección de problemas importantes en Bioinformática dende unha perspectiva eminentemente estatística/probabilística.				

**Study programme competences / results**

Code	Study programme competences / results
A5	CE5 - Development of skills in the management of statistical techniques and their application to data sets from the bioinformatics field.
A6	CE6 - Ability to identify software tools and most relevant bioinformatics data sources, and acquire skill in their use
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
B6	CG1 -Search for and select the useful information needed to solve complex problems, driving fluently bibliographical sources for the field
B7	CG2 - Maintain and extend well-founded theoretical approaches to enable the introduction and exploitation of new and advanced technologies
C3	CT3 - Use the basic tools of the information technology and communications (ICT) necessary for the exercise of their profession and lifelong learning
C6	CT6 - To assess critically the knowledge, technology and information available to solve the problems they face to.

**Learning outcomes**

Learning outcomes	Study programme competences / results		
Coñecer os principios estatísticos da análise de sistemas de altas prestacións para o estudo da expresión xénica	AJ5 AJ6	BJ1 BJ2 BJ6 BJ7	CJ3 CJ6
Comprender os aspectos estatísticos da construción de árbores filoxenéticas	AJ5 AJ6	BJ1 BJ2 BJ6 BJ7	CJ3 CJ6
Coñecer os fundamentos dos principais modelos estocásticos empregados en xenética de poboacións	AJ5 AJ6	BJ1 BJ2 BJ6 BJ7	CJ3 CJ6



Entender as bases probabilísticas e estatísticas dos métodos de análise de secuencias biolóxicas	AJ5	BJ1	CJ3
	AJ6	BJ2	CJ6
		BJ6	
		BJ7	

Contents	
Topic	Sub-topic
1. Análise estatística de sistemas de altas prestacións para o estudo da expresión xénica	Xeneralidades sobre o estudo da expresión xénica con microarrays e RNA-seq. Preprocesamento dos datos de microarrays e RNA-seq. Análise de conglomerados de mostras e xenos. Multiplicidade dos contrastes de hipóteses nos estudos de expresión xénica diferencial. Métodos estatísticos para a análise da expresión xénica diferencial con RNA-seq.
2. Árbores filoxenéticas	Modelos de evolución das secuencias de ADN. Construción e estimación de árbores filoxenéticas: métodos baseados nos conceptos de parsimonia, distancia e máxima verosimilitude. Contrastes sobre filoxenias e métodos relacionados.
3. Introducción ós modelos estocásticos en xenética de poboacións	O modelo de Wright-Fisher. O coalescente: o coalescente básico, contrastes da hipótese de neutralidade, extensións do coalescente.
4. Aspectos estatísticos da análise de secuencias biolóxicas	Sistemas de puntuación de alineamentos. Análise da significación estatística de alineamentos de secuencias.

Planning				
Methodologies / tests	Competencies / Results	Teaching hours (in-person & virtual)	Student?s personal work hours	Total hours
Guest lecture / keynote speech	A5 B1 B7	18	40	58
Problem solving	B2 B6 C6	12	28	40
ICT practicals	A5 A6 B2 C3	12	25	37
Mixed objective/subjective test	A5	3	0	3
Practical test:	B6	0	10	10
Personalized attention		2	0	2

(\*)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	Exposición oral complementada co uso de medios audiovisuais, coa finalidade de transmitir ó estudante os coñecementos teóricos
Problem solving	Seminarios en grupos de tamaño intermedio destinados á resolución de exercicios e problemas
ICT practicals	Resolución de supostos prácticos e teóricos mediante a utilización de software estatístico
Mixed objective/subjective test	Proba con preguntas de tipo test de opcións múltiples e/ou preguntas de resposta breve realizada co fin de avaliar globalmente a adquisición de coñecementos polo estudante
Practical test:	Resolución polo estudante dun caso práctico



## Personalized attention

Methodologies	Description
Guest lecture / keynote speech Problem solving ICT practicals Practical test:	A atención personalizada farase mediante titorías presenciais no despacho do profesor

## Assessment

Methodologies	Competencies / Results	Description	Qualification
Mixed objective/subjective test	A5	Proba que constará de preguntas de tipo test con opcións múltiples e/ou con preguntas de breve resposta	65
Practical test:	B6	Caso práctico resolto polo estudante	35

## Assessment comments

<p>Avaliación na primeira oportunidade</p> <p>Poderá ser tido en conta o seguimento obxectivo do curso e a participación activa pola parte do estudante, reflectíndose na calificación con ata un 20% da nota final. A porcentaxe restante da calificación (entre 80% e 100%) dependerá da avaliación da Proba mixta e da Proba práctica, e calcularase reescalando se é preciso as porcentaxes que figuran na táboa superior.</p> <p>Para superar a materia é necesario acadar unha calificación total mínima de 50 puntos, sendo en todo caso obligatoria a presentación á Proba mixta. Quen non se presente á Proba mixta será calificado como 'non presentado'.</p> <p>Avaliación na segunda oportunidade</p> <p>Farase cos mesmos criterios da primeira oportunidade. O alumnado que houbera presentado a Proba práctica na primeira oportunidade conservará a nota obtida na súa avaliación na primeira oportunidade.</p> <p>O alumnado con recoñecemento de dedicación a tempo parcial e dispensa académica de exención de asistencia que decida non asistir regularmente ás clases, será evaluado nas dúas oportunidades como o resto do alumnado que se atopa nunha situación similar.</p> <p>A realización fraudulenta das probas ou actividades de avaliación, unha vez comprobada, implicará directamente a cualificación de suspenso (0) na convocatoria correspondente, invalidando así calquera cualificación obtida en todas as actividades de avaliación.</p>
--

## Sources of information



<b>Basic</b>	<ul style="list-style-type: none"><li>- Deonier R C, Tavaré S, Waterman M S (2005). Computational Genome Analysis. Springer</li><li>- Ewens W J, Grant G R (2005). Statistical Methods in Bioinformatics. Springer</li><li>- Felsenstein J (2004). Inferring Phylogenies. Sinauer</li><li>- Datta S, Nettleton DS (Editores) (2014). Statistical Analysis of Next Generation Sequence Data. Springer</li><li>- Holmes S, Huber W (2019). Modern Statistics for Modern Biology. Cambridge University Press</li><li>- Korpelainen E, Tuimala J, Somervuo P, Huss M, Wong G (2014). RNA-seq Data Analysis: A Practical Approach. Chapman&amp;Hall/CRC</li><li>- Wakeley J (2008). Coalescent Theory. Freeman</li><li>- R (). Sitio web programa R. <a href="https://cran.r-project.org/">https://cran.r-project.org/</a></li><li>- Bioconductor Project (). Sitio web Bioconductor. <a href="https://www.bioconductor.org/">https://www.bioconductor.org/</a></li></ul>
<b>Complementary</b>	

## Recommendations

Subjects that it is recommended to have taken before

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.