



Asignatura: Bioinformática
Código: 30122
Centro: Facultad de Ciencias
Titulación: Máster en Biofísica
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3 ECTS

ASIGNATURA / COURSE TITLE

Bioinformática / [Bioinformatics](#)

1.1. Código / [Course number](#)

30122

1.2. Materia / [Content area](#)

Física, Biología, Matemáticas / [Physics](#), [Biology](#), [Mathematics](#)

1.3. Tipo / [Course type](#)

Obligatoria / [Compulsory](#)

1.4. Nivel / [Course level](#)

Posgrado / [Graduate \(postgraduate program\)](#)

1.5. Curso / [Year](#)

1º / [1st](#)

1.6. Semestre / [Semester](#)

Anual / [1st,2nd](#)

1.7. Número de créditos/ [Credit allotment](#)

3 ECTS / 3 [ECTS](#)

1.8. Requisitos previos/ [Prerequisites](#)

The minimum requirements are those necessary to access to the Master.

1.9. Requisitos mínimos de asistencia a las sesiones presenciales / [Minimum attendance requirement](#)

Minimum attendance for theory lectures: 80%. Practical lectures: 100%.



Asignatura: Bioinformática
Código: 30122
Centro: Facultad de Ciencias
Titulación: Máster en Biofísica
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3 ECTS

1.10. Datos del equipo docente/ Faculty data

Ugo Bastolla:
(subject coordinator)

Unidad de Bioinformática.
Centro de Biología Molecular Severo Ochoa
Teléfono:
E-mail: ubastolla@cbm.uam.es
Página web:

1.11. Objetivos del curso/ Course objectives

The aim of the theoretical part of this program is to provide to the students an elemental knowledge on bioinformatics, with an special emphasis on those branches more directly related with biophysics. The program is divided in three different blocks.

Into the first block we present some basic foundations on bioinformatics. Here we show some basic biostatistics procedures, the more relevant databases, sequence analysis techniques and an introduction to protein structure.

Into the second block, we present advanced bioinformatics methods focusing on protein structure dynamics and interactions, that have special importance due to its connection with biophysical techniques.

Finally, all these concepts are merged into an evolutive context, showing the formal parallelism that it is possible to establish with statistical physics. It is shown how molecular evolution can be modelled using this mathematical framework, how it contributes to our understanding and prediction of biological processes and which is its further impact on the development and improvement of bioinformatics tools.

On the other hand, the aim of the practical lectures is to provide to the students the minimum programming skills in order to analyse and solve simple scientific problems with an special emphasis on bioinformatics algorithms.

- Contenidos del programa/ Course contents

1st Block.

- Biostatistics.
- Data Mining and Functional Genomics.
- Sequence Analysis.
- Introduction to Protein Structure. Protein Folding.



2nd Block.

- Molecular Dynamics.
- Normal Mode Analysis.
- Docking.
- Virtual Screening.

3rd Block.

- Molecular Evolution.
- Evolution Rates and Phylogenetics.
- Protein Structure Evolution and back: Classification and Modelling.

1.12. Referencias de consulta / Course bibliography

*** = Almost mandatory
 ** = Recommended to everybody
 * = Recommended to those particularly interested in the field

Papers

- Biostatistics and Data Mining.
 #####

Richard M. Simon, Edward L. Korn, Lisa M. McShane, Michael D. Radmacher, George W. Wright, Yingdong Zhao. 2004. Design and Analysis of DNA Microarray Investigations. Springer*

-Cichocki A, Zdunek R, Phan AH, Amari S. Non-negative matrix and tensor factorizations: Wiley; 2009.**

Draghici, S. Data Analysis Tools for DNA Microarrays. Chapman & Hall /CRC Press. 2003*

- Evolutionary rates: general
 #####

Ho SY, Phillips MJ, Cooper A, Drummond AJ.. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Mol Biol Evol. 2005 Jul;22(7):1561-8. **



Asignatura: Bioinformática
Código: 30122
Centro: Facultad de Ciencias
Titulación: Máster en Biofísica
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3 ECTS

Rocha EP, Smith JM, Hurst LD, Holden MT, Cooper JE, Smith NH, Feil EJ. Comparisons of dN/dS are time dependent for closely related bacterial genomes. *J Theor Biol.* 2006 Mar 21;239(2):226-35. Epub 2005 Oct 18 *

Peterson GI, Masel J. Quantitative prediction of molecular clock and ka/ks at short timescales. *Mol Biol Evol.* 2009 Nov;26(11):2595-603. *

— Evolutionary rates: examples

#####

Bininda-Emonds OR, Cardillo M, Jones KE, MacPhee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A. The delayed rise of present-day mammals. *Nature.* 2007 446:507-12 *

S Dorus, EJ Vallender, PD. Evans, JR. Anderson, SL. Gilbert, M. Mahowald, GJ. Wyckoff, CM. Malcom, and BT. Lahn. (2004) Accelerated Evolution of Nervous System Genes in the Origin of Homo sapiens. *Cell*, Vol 119, 1027-1040, 2004 **

Xia Y, Franzosa EA, Gerstein MB. Integrated assessment of genomic correlates of protein evolutionary rate. *PLoS Comput Biol.* 2009 Jun;5(6):e1000413. *

— Molecular clock

#####

Bromham, L. and Penny, D. 2003. The modern molecular clock. *Nature Reviews Genetics* 4:216-224. *

— Population genetics

#####

Sella G, Hirsh AE. The application of statistical physics to evolutionary biology. *Proc Natl Acad Sci U S A.* 2005 Jul 5;102(27):9541-6. ***

Kussell, E. & Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* 309, 2075-2078 (2005) *

Sergey Kryazhimskiy and Joshua B. Plotkin The dynamics of adaptation on correlated fitness landscapes. *Proc Natl Acad Sci U S A* 2009 *

— Experimental evolution

#####

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature.* 2009 Oct 29;461(7268):1243-7. ***

— Effective population size

#####

Berg, O.G. (1996) Selection Intensity for Codon Bias and the Effective Population Size of *Escherichia coli*,



Genetics 142:1379-1382. *

— Mutation bias

#####

Chen, S.L., Lee, W., Hottes, A.K., Shapiro, L. and McAdams, H. (2004), Codon usage between genomes is constrained by genome-wide mutational processes, Proc. Natl. Acad. Sci. USA 101, 3480-5. *

Duret L, Eyre-Walker A, Galtier N. A new perspective on isochores evolution. Gene. 2006 385:71-4. *

Bastolla, U., Porto, M., Roman, H.E., Vendruscolo, M. (2006). A protein evolution model with independent sites that reproduces site-specific amino acid distributions from the Protein Data Bank. BMC Evol. Biol. 6:43. *

— Protein folding

#####

Hartl FU, Hayer-Hartl M. Converging concepts of protein folding in vitro and in vivo. Nat Struct Mol Biol. 2009 Jun;16(6):574-81. **

Shakhnovich; Farztdinov; Gutin & Karplus Protein folding bottlenecks: A lattice Monte Carlo simulation. Phys Rev Lett, 1991, 67, 1665-1668

Shakhnovich; Proteins with selected sequences fold into unique native conformation. Phys Rev Lett, 1994, 72, 3907-3910 **

— Protein structure evolution

#####

Tokuriki N, Tawfik DS. Stability effects of mutations and protein evolvability. Curr Opin Struct Biol. 2009 Oct;19(5):596-604. **

W.R. Taylor, Evolutionary transitions in protein fold space, Curr Opin Struct Biol 17 (2007), 354-361. *

S.S. Krishna and N.V. Grishin, Structural drift: a possible path to protein fold change, Bioinformatics 21 (2005), pp. 1308-1310. *

S.G. Peisajovich, L. Rockah and D.S. Tawfik,. Evolution of new protein topologies through multistep gene rearrangements, Nat Genet 38 (2006), pp. 168-174. *

England, J. L. & Shakhnovich, E. I. Structural determinant of protein designability. Phys Rev Lett, , 2003, 90, 218101 *

Shakhnovich, B. E.; Deeds, E.; Delisi, C. & Shakhnovich, E. Protein structure and evolutionary history determine sequence space topology. Genome Res, 2005, 15, 385-392 **



Zeldovich, K. B.; Chen, P.; Shakhnovich, B. E. & Shakhnovich, E. I. A first-principles model of early evolution: emergence of gene families, species, and preferred protein folds. PLoS Comput Biol, 2007, 3, e139 *

— Transcription Regulation

#####

X Gu, Zhongqi Zhang and Wei Huang (2005). Rapid evolution of expression and regulatory divergences after yeast gene duplication PNAS 2005 vol. 102 no. 3 707-712 *

— Alternative splicing

#####

Tress ML et al..The implications of alternative splicing in the ENCODE protein complement. Proc. Natl Acad. Sci. USA (2007) 104:5495-5500. *

— Translation

#####

D. Allan Drummond & Claus O. Wilke. The evolutionary consequences of erroneous protein synthesis. Nature Reviews Genetics 10, 715-724 (2009) **

— Aggregation

#####

F. Chiti and C.M. Dobson, Protein misfolding, functional amyloid, and human disease,. Annu. Rev. Biochem. 75 (2006), pp. 333-366. **

— Disordered proteins

#####

Uversky, VN and Dunker, AK (2010) Understanding protein non-folding. Biochim. Biophys. Acta 1804: 1231-64. **

Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, and Obradovic Z. (2002) Intrinsic disorder and protein function, *Biochemistry*, 41, 6573-6582. *

— Chaperones

#####

Sangster TA, Lindquist S, Queitsch C. Under cover: causes, effects and implications of Hsp90-mediated genetic capacitance. *Bioessays*. 2004 Apr;26(4):348-62. Review. **

Tokuriki N, Tawfik DS. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature*. 2009 Jun 4;459(7247):668-73. **

F.U. Hartl and M. Hayer-Hartl,. Molecular chaperones in the cytosol: from nascent chain to folded protein, *Science* 295 (2002), pp. 1852-1858.**

Fares, M.A., Ruiz-Gonzalez, M.X., Moya, A., Elena, S.F., Barrio, E. (2002) Endosymbiotic bacteria: GroEL buffers against deleterious mutations, *Nature* 417:398.*



K.A. Morano (2007) New Tricks for an Old Dog The Evolving World of Hsp70 Ann. N. Y. Acad. Sci. 1113: 1-14. *

— Protein function evolution

#####

Devos, D. and Valencia, A. (2000) Practical limits of function prediction. Proteins 41:98-107. *

Tokuriki N, Stricher F, Serrano L, Tawfik DS. How protein stability and new functions trade off. PLoS Comput Biol. 2008 Feb 29;4(2):e1000002. *

A. Pascual-García, D Abia, R Mendez, GS. Nido, U Bastolla (2009) Quantifying the evolutionary divergence of protein structures: The role of function change and function conservation. Proteins, doi 10.1002/prot.22616 *

— Evolution of protein folding thermodynamics

#####

Bastolla, U., Moya, A., Viguera, E., van Ham, R.C.H.J. (2004) Genomic determinants of protein folding thermodynamics, J. Mol. Biol. 343:1451-1466. *

— Normal Modes Analysis

#####

Molecular Modeling of proteins. Andreas Kukol Editor. Normal Modes and Essential Dynamics. Hayward S, Groot BL. Chp 5: 89-106. From: Methods in Molecular Biology, vol 443. Humana Press 2008.

Computational Biochemistry and Biophysics. Becker O, Mackerell AD, Roux B, Wanatabe M Editors. Normal Mode Analysis of Biological Molecules. Hayward S. Chp 8: 153 – 168. CRC Press 2001.

Go, N., Noguti, T. and Nishikawa, T. (1983). Dynamics of a small globular protein in terms of low-frequency vibrational modes. Proc. Natl. Acad. Sci. USA 80, 3696–3700.

Garcia, A. E. (1992). Large-amplitude nonlinear motions in proteins. Phys. Rev. Lett. 68,2696–2699.

Amadei, A, Linssen, A. B. M. and Berendsen, H. J. C. (1993). Essential dynamics of proteins. Proteins: Struct. Funct. Genet. 17, 412–425.

A. and Go, N. (1995). Harmonicity and anharmonicity in protein dynamics: a normal modes and principal component analysis. Proteins: Struct. Funct. Genet. 23, 177–186

Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. Physical Review Letters 77, 1905–1908.

Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuations dynamics of proteins with an elastic network model. *Biophys J.* 2001; 80: 505-515.

Mendez R, Bastolla U. Torsional Network Model: Normal Modes in Torsion Angle Space Better Correlate with Conformation Changes in proteins. Phys Rev Lett. 2010; 104:228103.



Asignatura: Bioinformática
Código: 30122
Centro: Facultad de Ciencias
Titulación: Máster en Biofísica
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3 ECTS

— Evolution of protein folding thermodynamics

#####

Bernard Munos: Lessons from 60 years of pharmaceutical innovation. Nature Reviews in Drug Discovery, 2009, 8, 959-968. (<http://www.nature.com/nrd/journal/v8/n12/full/nrd2961.html>)

Hugo Kubinyi: The changing landscape in drug discovery. Computational Approaches to Structure Based Drug Design, R. M. Stroud, Ed., Royal Society of Chemistry, London, 2007, 24-45. (<http://www.kubinyi.de/rsc.pdf>)

Caterina Bissantz, Bernd Kuhn, and Martin Stahl: A medicinal chemistry guide to molecular interactions. Journal of Medicinal Chemistry 2010. (<http://pubs.acs.org/doi/abs/10.1021/jm100112j>)

Tjelvar S. G. Olson, Mark A. Willians, William R. Pitt and John E. Ladbury: The thermodynamics of protein-ligand interaction and salvation: insights for ligand design. Journal of Molecular Biology, 2008, 384, 1002-1017.

Wikipedia articles

Mutation
Natural selection
Genetic drift
Fixation (population genetics)
Molecular clock
Phylogenetic tree
Substitution model
Ka/Ks ratio

Markov chain

Protein folding
Protein aggregation
Prion
Chaperone (protein)



Asignatura: Bioinformática
Código: 30122
Centro: Facultad de Ciencias
Titulación: Máster en Biofísica
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3 ECTS

Suggested books

Graur, D., Li, W.H. (2000) Fundamentals of molecular evolution, Sinauer, Sunderland. *

W.J. Ewens, G.R. Grant (2001) Statistical methods in Bioinformatics: An introduction. Springer Verlag, New York **

NH Barton, DEG Briggs, JA Eisen, DB Goldstein and NH Patel (2007) Evolution Cold Spring Harbor Laboratory Press *

2. Métodos docentes/ Teaching methodology

- 1) Theory lectures: Oral presentations by the teachers of the fundamental contents of the subject. Audiovisual material will be available for the students in the Master web page, or in the personal web pages of lecturers.
- 2) Practical lectures: Problems solution, practical exercises to be programmed by the students in the computer's room assisted by teachers.
- 3) Individual analysis of proposed papers and group discussions.

3. Tiempo de trabajo del estudiante / Student workload

		Nº de horas	Porcentaje
Personal attendance	Theory classes	26	60%
	Practical classes	14	
	Papers discussions.	5	
No attendance	Practical exercises, theory study and papers discussions.	30	40%
Total		75 h	100%



Asignatura: Bioinformática
Código: 30122
Centro: Facultad de Ciencias
Titulación: Máster en Biofísica
Nivel: Máster
Tipo: Obligatoria
Nº de créditos: 3 ECTS

4. Métodos de evaluación y porcentaje en la calificación final / Evaluation procedures and weight of components in the final grade

For the regular evaluation session the evaluation consists on a:

- Practical exercise: programming exercises to be solved. (50% final mark)
- Theoretical exercise: Analysis of a proposed reference following some peer-review like guidelines, that will be presented in an individual report. (35 final mark%) This analysis is further completed with a group discussion and the conclusions will be summarized in an oral presentation. (15% final mark).

The extraordinary evaluation session follows the same guidelines than the regular, but in the theoretical exercise the group discussion and oral presentation will be substituted by an oral session where the student must answer some questions about the proposed reference. (15% final mark).

5. Cronograma* / Course calendar

Semana aprox. Week	Contenido Contents	Horas presenciales Contact hours	Horas no presenciales Independent study time
1-2	Practical lectures	14	10
3-10	Theory lectures	26	10
6	Papers discussions	5	10

*Este cronograma tiene carácter orientativo.