

Equipo Investigador / *Research Team*

Investigadores / *Researchers:*

Marc Martí-Renom
Hernán J. Dopazo
Ana Victoria Conesa Cegarra
Fátima Al-Shahrour
Juan Antonio Gabaldón Esteve
Javier Santoyo López (CIBERER)
Emidio Capriotti
Davide Bau
Francisco García García
David Montaner González (INB)
Joaquín Tárraga Jiménez (INB)

Predoctorales / *Predoctoral Students:*

Eva Alloza Anguiano
Stefan Goetz (CIBERER)
Leonardo Arbiza Brustin

Técnicos / *Technicians:*

Pablo Escobar López
José Carbonell Caballero
Ignacio Medina Castelló (CIBERER)
Rafael Carlos Jimenez Domenech (CIBERER)
François Serra
Marina Marcet Houben
Jaime Huerta Cepas (INB)
Adriana Cucchi (Auxiliar Admin. / *Administrative assistant*)



Colaboradores / *Collaborators:*

Pablo Mínguez Paniagua
Jordi Durbán
Stefania Bossi
José Díaz Riquelme
Inmaculada Rapado
Luba Pardo
Sven Wichert
Chiara Russo
Mateus Patricio
Johanes van Heerden
Patricia Sebastián

Maya Topf
Karin Kjeldahl
María Cíprano
Marta Diestro
Giulia Gentile
Salvador Capella Gutiérrez
Carlos Garrido Ledesma
Giuseppe Marceddu
Víctor Martínez González
Julian Mensch
Vivian Nogueira
Sonia Tarazona

Fig. Red de conexiones entre proteínas correspondientes a varios procesos biológicos relacionados con cáncer.

Fig. Network connecting proteins belonging to several biological processes related to cancer.





Departamento / *Department*

Bioinformática y Genómica *Bioinformatics and Genomics*

Laboratorio / *Laboratory*

Bioinformática *Bioinformatics*

Responsable / *Team Leader*

Joaquín Dopazo Blázquez (jdopazo@cipf.es)

DESCRIPCIÓN DE LA ACTIVIDAD INVESTIGADORA / *SUMMARY OF RESEARCH*

Hoy en día, los principales avances en biomedicina vienen de la genómica y la bioinformática, entendidas bajo la perspectiva de la biología de sistemas. Con esta filosofía, el objetivo general de nuestro departamento es relacionar la estructura genómica y las mutaciones (a través de nuestra unidad de genómica comparativa) con sus efectos a nivel regulatorio, celular y fenotípico (genómica funcional) tratando de entender los mecanismos de acción subyacentes (genómica estructural).

El análisis del enorme volumen de datos producido por las técnicas de alto rendimiento (microarrays, secuenciación de nueva generación, GWAS, etc.) es reconocido hoy en día como uno de los principales cuellos de botella para el uso eficiente de las metodologías genómicas. Nuestro departamento está firmemente involucrado en el desarrollo de algoritmos y herramientas bioinformáticas innovadoras en contacto directo con investigadores de laboratorio. Por ello, hemos desarrollado programas para el análisis de la expresión génica (GEPAS, el servidor de web más usado en su género), filogenómica (Phylemon), interpretación funcional de experimentos a escala genómica (Babelomics), alineamientos estructurales de RNAs (SARA), etc.

Formamos parte de tres importantes iniciativas nacionales: el instituto de bioinformática (INB), el CIBER de enfermedades raras y la red española de cáncer (RTICC). También estamos involucrados en varios consorcios internacionales, como el MAQCII, que intenta establecer el uso correcto de los microarrays con propósito pronóstico y diagnóstico, el SEQC, que intenta establecer controles de calidad en las metodologías de secuenciación de nueva generación o el STAR, que busca la caracterización de la variabilidad en el genoma de la rata (Saar et al., 2008 Nat. Genet).

Today's biomedicine is thriving on the wings of genomics and bioinformatics, understood under the prism of systems biology. Following this philosophy, the general objective of our department is to relate the mutations and genomic structure (through our Comparative Genomics unit) to their effect at cellular, regulatory and phenotypic level (Functional Genomics) trying to understand the mechanism of action (Structural Genomics).

The analysis of the huge amounts of data produced by the new high-throughput technologies (microarrays, next generation sequencing, GWAS, etc.) is nowadays recognised as the main bottleneck hindering the efficient and extensive use of genomics technology. Our department has a strong commitment to the development of innovative analysis algorithms and bioinformatics tools, in close contact with laboratory researchers. Thus we have developed state-of-the-art programmes for gene expression data analysis (GEPAS, the most extensively used web server in its genera), phylogenomics (Phylemon), functional interpretation of genome-scale experiments (Babelomics), structural alignment of RNAs (SARA), etc.

We also form part of three large initiatives in Spain: the Spanish Institute of Bioinformatics (INB), the CIBER of rare diseases and the Spanish Cancer Network, as well as being involved in several international consortia such as the MAQCII, which aims to establish best practices in the use of microarrays for prognostic and diagnostic purposes, the SEQC, which aims to establish quality control criteria in next generation sequencing methodologies, and the STAR consortium, which targets the characterisation of the variability in the rat genome (Saar et al., 2008 Nat. Genet).

Colaboraciones científicas / *Scientific Collaborations*

Centro / *Centre*: Centro Nacional de Investigaciones Oncológicas, Madrid

Investigador / *Researcher*: Miguel Ángel Piris

Centro / *Centre*: Banco Nacional de Líneas celulares - CIPF, Valencia

Investigador / *Researcher*: Carlos Simón Vallés

Centro / *Centre*: CNB, Madrid

Investigador / *Researcher*: José Miguel Martínez Zapater

Centro / *Centre*: IBMCP-UPV, Valencia

Investigador / *Researcher*: Antonio Granell

Centro / *Centre*: IRB, Barcelona.

Investigador / *Researcher*: Manuel Palacín.

Centro / *Centre*: University of Cambridge, Reino Unido

Investigador / *Researcher*: Antonio Vidal-Puig

Centro / *Centre*: Sanger Center, Reino Unido

Investigador / *Researcher*: A. L. Toribio

Centro / *Centre*: Centre de Regulació Genómica, Barcelona.

Investigador / *Researcher*: Cayetano González

Centro / *Centre*: Division of Pathway Medicine, University of Edinburgh, Reino Unido

Investigador / *Researcher*: Peter Ghazal y Tom Freeman.

Centro / *Centre*: University of Cape Town, Sudáfrica

Investigador / *Researcher*: Niccola Illing

Centro / *Centre*: VIB Switch Laboratory, Vrije Universiteit Brussel, Bélgica

Investigador / *Researcher*: Joost Schymkowitz,

Centro / *Centre*: EMBL-Hamburg, Alemania

Investigador / *Researcher*: Mathias Wilmanns

Centro / *Centre*: Pontifica Universidad de Chile, Chile

Investigador / *Researcher*: F. Melo.

Centro / *Centre*: Universitat Pompeu Fabra, Barcelona

Investigador / *Researcher*: B. Oliva.

Centro / *Centre*: St. George Hospital, University of New South Wales, Australia

Investigador / *Researcher*: S. Krillis.

Centro / *Centre*: CUBIC, Columbia University, EEUU y Centro Nacional de Investigaciones Oncológicas, Madrid

Investigador / *Researcher*: B. Rost and A. Valencia.

Centro / *Centre*: Unitat de Biologia Evolutiva, Universitat Pompeu Fabra, Barcelona

Investigador / *Researcher*: Arcadi Navarro.

Centro / *Centre*: Evolución Molecular y Bioinformática, Universidad de Vigo

Investigador / *Researcher*: David Posada

Centro / *Centre*: Grupo de Ecología Reproductiva y Evolución, Museo Nacional de Ciencias Naturales, Madrid

Investigador / *Researcher*: Montserrat Gomendio

Centro / *Centre*: Department of Veterinary Basic Sciences, The Royal Veterinary College London, Reino Unido

Investigador / *Researcher*: Eduardo Roldán

Centro / *Centre*: Grupo de Investigación en Biología Evolutiva, Facultad de Ciencias, Universidad de Buenos Aires, Argentina

Investigador / *Researcher*: Esteban Hasson

LÍNEAS DE INVESTIGACIÓN / *LINES OF RESEARCH*

- Genómica funcional: Relaciones entre la expresión de los genes, sus mutaciones y la función, dentro de la biología de sistemas, en el contexto de las enfermedades.
- Genómica comparativa: Análisis de los patrones y procesos ocurridos durante la evolución del genoma y su aplicación en salud humana y enfermedades.
- Genómica estructural: Uso de las leyes de la física y la evolución para desarrollar y aplicar métodos bioinformáticos para entender y caracterizar la regulación de la célula.
- *Functional genomics: Decipher the interplay between gene expression, mutations and function under the systems biology prism, in the context of diseases.*
- *Comparative genomics: Analysis of patterns and processes which have occurred during the evolution of our genome, and their application to human health and disease.*
- *Structural genomics: Use of the laws of physics and evolution to develop and apply computational methods to understand and characterise cell regulation beyond proteins.*

PUBLICACIONES / *PUBLICATIONS 2008*

1. Bonifaci N, Berenguer A, Diez J, Reina O, Medina I, Dopazo J, Moreno V, Pujana MA. (2008); Biological processes, properties and molecular wiring diagrams of candidate low-penetrance breast cancer susceptibility genes. Ed. BMC Med Genomics. 18;1(1):62.
2. Pieper, U. Eswar, N. Webb, B.M. Eramian, D. Kelly, L. Barkan, D.T. Carter, H. Mankoo, P. Karchin, R. Marti-Renom, M.A. Davis, F.P. Sali, A. (2008); MODBASE, a database of annotated comparative protein structure models, and associated resources. Ed. Nucleic Acid Research. 37:D347-54.
3. Dopazo, J. (2008); Formulating and testing hypotheses in functional genomics. Ed. Artif Intell Med. 2008 Sep 11. [Epub ahead of print] PMID: 18789659.
4. Hoogerwerf, W.A., Sinha, M., Conesa, A., Luxon, B.A., Cassone, V.M. (2008); Transcriptional Profiling of mRNA Expression in the Mouse Distal Colon. Ed. Gastroenterology. 135(6):2019-29.
5. Horcajadas, J.A., Minguéz, P., Dopazo, J., Riesewijk, A., Esteban, F.J., Domínguez, F., Polman, J., Mosselman, S., Giudice, L., Pellicer, A. and Simon, C. (2008); Controlled ovarian stimulation induces a functional genomic delay of the endometrium with potential clinical implications. Ed. J. Clin. Endocr. Metabol. 93(11):4500-10.
6. Conesa, A., Bro, R., García-García, F., Prats, J.M., Goetz, S., Kjeldahl, K., Montaner, D., Dopazo, J. (2008); Direct functional assessment of the composite phenotype through multivariate projection strategies. Ed. Genomics. 92:373:383.
7. Espadaler, J., Eswar, N., Querol, E., Aviles, F.X., Sali, A., Marti-Renom, M.A., Oliva, B. (2008); Prediction of enzyme function by combining sequence similarity and protein interactions. Ed. BMC Bioinformatics. 9(1) 249.
8. Capriotti, E., Marti-Renom, M.A. (2008); RNA structure alignment by a unit-vector approach. Ed. Bioinformatics. 24 i112-i118.
9. Botton, A., Galla, G., Conesa, A., Bachem, C., Ramina, A., Barcaccia, G. (2008); Large-scale Gene Ontology analysis of plant genome- and transcriptome-derived sequences retrieved by AFLP technology. Ed. BMC Genomics 9:347.
10. Al-Shahrour, F., Carbonell, J., Minguéz, P., Goetz, S., Conesa, A., Tárraga, J., Medina, I., Alloza, E., Montaner, D., Dopazo, J. (2008); Babelomics: advanced functional profiling of transcriptomics, proteomics and genomics experiments. Ed. Nucleic Acids Res. 36:341-346.
11. Panjkovich, A., Melo, F., Marti-Renom, M.A. (2008); Evolutionary potentials: structure specific knowledge-based potentials exploiting the evolutionary record of sequence homologs. Ed. Genome Biology. 9:R68.
12. Saar, K., Beck, A., Bihoreau, M.T., Birney, E., Brocklebank, D., Chen, Y., Cuppen, E., Demonchy, S., Dopazo, J., Flicek, P., Foglio, M., Fujiyama, A., Gut, I.G., Gauguier, D., Guigo, R., Guryev, V., Heinig, M., Hummel, O., Jahn, N., Klages, S., Kren, V., Kube, M., Kuh, H., Kuramoto, T., Kuroki, Y., Lechner, D., Lee, Y.A., Lopez-Bigas, N., Lathrop, G.M., Mashimo, T., Medina, I., Mott, R., Patone, G., Perrier-Cornet, J.A., Platzer, M., Pravenec, M., Reinhardt, R., Sakaki, Y., Schilhabel, M., Schulz, H., Serikawa, T., Shikhagaie, M., Tatsumoto, S., Taudien, S., Toyoda, A., Voigt, B., Zelenika, D., Zimdahl, H. & Hubner, N. (2008); SNP and haplotype mapping for genetic analysis in the rat. Ed. Nat. Genet. 40:560-566.
13. Tárraga, J., Medina, I., Carbonell, J., Huerta-Cepas, J., Minguéz, P., Alloza, E., Al-Shahrour, F., Vegas-Azcarate, S., Gotz, S., Escobar, P., García-García, F., Conesa, A., Montaner, D., Dopazo, J. (2008); GE-

- PAS, a web-based tool for microarray data analysis and interpretation. Ed. *Nucleic Acids Res.* 36(Web Server issue):W308-14.
14. Wilkinson MD, Senger M, Kawas E, Bruskiwich R, Gouzy J, Noiro C, Bardou P, Ng A, Haase D, Saiz Ede A, Wang D, Gibbons F, Gordon PM, Sensen CW, Carrasco JM, Fernández JM, Shen L, Links M, Ng M, Opushneva N, Neerinx PB, Leunissen JA, Ernst R, Twigger S, Usadel B, Good B, Wong Y, Stein L, Crosby W, Karlsson J, Royo R, Páraga I, Ramírez S, Gelpi JL, Trelles O, Pisano DG, Jimenez N, Kerhornou A, Rosset R, Zamacola L, Tárraga J, Huerta-Cepas J, Carazo JM, Dopazo J, Guigo R, Navarro A, Orozco M, Valencia A, Claros MG, Pérez AJ, Aldana J, Rojano MM, Fernandez-Santa Cruz R, Navas I, Schiltz G, Farmer A, Gessler D, Schoof H, Groscurth A. (2008); Interoperability with Moby 1.0--it's better than sharing your toothbrush!. Ed. *Brief Bioinform.* 9(3):220-31.
 15. Robert Stierum, Ana Conesa, Wilbert Heijne, Ben van Ommen, Karin Junker, Mary P. Scott, Roger J. Price, Clive Meredith, Brian G. Lake, John Groten. (2008); Transcriptome analysis provides new insights into liver changes induced in the rat upon dietary administration of the food additives butylated hydroxytoluene, curcumin, propyl gallate and thiabendazole. Ed. *Food and Chemical Toxicology*, 46(8):2616-28.
 16. Götz, S., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Terol, J., García-Gómez, J.M., Robles, M., Talón, M., Dopazo, J., Conesa, A. (2008); High-throughput functional annotation and data mining with the Blast2GO suite. Ed. *Nucleic Acids Res.* 36:3420-3435.
 17. Agudo, M., Sobrado, P., Pérez-Marin, M.C., Lönnngren, U., Conesa, A., Cánovas, I., Salinas-Navarro, M., Miralles-Imperial, J., Hallböök, F., Vidal-Sanz, M. (2008); Time course profiling of retinal transcriptome after optic nerve transection and crush. Ed. *Molecular Vision* 14:1050-1060.
 18. Conesa, A., Goetz S. (2008); Blast2GO: A comprehensive suite for functional analysis in plant genomics. Ed. *Journal of Plant Genomics*, 2008;2008:619832.
 19. Montero-Conde, C., Martín-Campos, J.M., Lerma, E., Gimenez, G., Martínez-Guitarte, J.L., Combalá, N., Montaner, D., Matías-Guiu, X., Dopazo, J., de Leiva, A., Robledo, M., Mauricio, D. (2008); Molecular profiling related to poor prognosis in thyroid carcinoma. Combining gene expression data and biological information. Ed. *Oncogene*. 27(11):1554-61.
 20. Reumers J., Conde L, Medina I, Maurer-Stroh S, Van Durme J, Dopazo J, Rousseau F and Schymkowitz J (2008); Joint annotation of coding and non-coding single nucleotide polymorphisms and mutations in the SNPeff and PupaSuite databases. Ed. *Nucleic Acids Res.* 36(Database issue):D825-9.
 21. Huerta-Cepas, J., Bueno, A., Dopazo, J., Gabaldon, T.(2008); PhyloDB: A database for complete collections of gene phylogenies. Ed. *Nucleic Acids Res.* D491-6.
 22. Capriotti, E. Marti-Renom, M.A.(2008); Computational RNA structure prediction. Ed. *Current Bioinformatics* 3(1)32-45.
 23. Valls, J., Grau, M., Sole, X., Hernandez, P., Montaner, D., Dopazo, J., Peinado, M.A., Capella, G., Pujana, M.A., Moreno, V. (2008) ; CLEAR-test: Combining inference for differential expression and variability in microarray data analysis. Ed. *Journal of Biomedical Informatics*. 41(1):33-45.
 24. Capriotti, E., Arbiza, L., Casadio, R., Dopazo, J., Dopazo, H., Marti-Renom, M.A. (2008); The use of estimated evolutionary strength at the codon level improves the prediction of disease related protein mutations in human. Ed. *Human Mutation*. 29(1):198-204.

LIBROS / *BOOKS*

Autores / *Authors*: Dopazo, H.

Título / *Title*: Selective Constraints on Human Disease Mutations and Polymorphisms in Handbook of Human Molecular Evolution

Editorial / *Publisher*: Hildegard Kehrer-Sawatzki & David N. Cooper. John Wiley & Sons, Ltd. UK.

Autores / *Authors*: Dopazo, H.

Título / *Title*: Selective Constraints and Human Disease Genes: Evolutionary and Bioinformatic Approaches in Encyclopedia of Life Science

Editorial / *Publisher*: John Wiley & Sons, Ltd. UK.

Autores / *Authors*: Gabaldón T., Gil R., Peretó J., Latorre A. and Moya A.

Título / *Title*: The core of a minimal gene set: insights from natural reduced genomes in Protocells: bridging nonliving and living matter

Editorial / *Publisher*: S. Rasmussen et. al., The MIT press.

Autores / *Authors*: Gabaldón, T.

Título / *Title*: Comparative genomics-based prediction of protein function. in Methods in Molecular Biology, Vol. 439: Genomics Protocols.

Editorial / *Publisher*: M. Starkey and R. Elaswarapu, Humana press

Autores / *Authors*: Dopazo J and Al-Shahrour, F.

Título / *Title*: Expression and Microarrays Methods in Molecular Biology. Bioinformatics vol 2. Structure, function and applications.

Editorial / *Publisher*: Humana Press. NJ, USA.

Autores / *Authors*: Capriotti, E., Marti-Renom, M.A.

Título / *Title*: Assessment of protein structure predictions in Computational Structural Biology.

Editorial / *Publisher*: World Scientific

Autores / *Authors*: Marti-Renom, M.A. Capriotti, E. Shindyalov, I. Bourne, P.

Título / *Title*: Structural Comparison and Alignment in Structural Bioinformatics

Editorial / *Publisher*: Wiley- Blackwell ISBN: 978-0-470-18105-8