

CURRICULUM VITAE

PERSONAL INFORMATION

Name: Ramón Díaz Uriarte
Date of birth: 30 October 1968
Nationality: Spanish
Address: Spanish National Cancer Center, Melchor Fernández Almagro 3, E-28029, Madrid (Spain)
Telephone: +34 912 246 900
Email: rdiaz02@gmail.com
Web page: <http://ligarto.org/rdiaz>

EDUCATION

1998 – 2000: MSc Statistics, University of Wisconsin-Madison, USA.
1996 – 2000: PhD Zoology, University of Wisconsin-Madison, USA.
1993 – 1996: MSc Biometry, University of Wisconsin-Madison, USA.
1987 – 1993: Licenciatura con grado ($\approx BSc + MSc$) in Biology, Universidad Autónoma de Madrid, Madrid, Spain.

OTHER PROFESSION QUALIFICATIONS

- Fellow Royal Statistical Society (2002).
- GradStat, Royal Statistical Society (2002).

PROFESSIONAL EXPERIENCE

11/2006 – Present: Part-time Associate Professor of Bioinformatics, Universidad Autónoma de Madrid.

05/2002 – Present: Staff Scientist, Spanish National Cancer Center (CNIO, Centro Nacional de Investigaciones Oncológicas). Responsibilities:

- Analysis and design of cDNA and aCGH array experiments.
- Collaboration on the analysis of data from cancer research experiments.
- Development and implementation of statistical techniques for microarray experiments.
- Teaching of statistics.

08/2001 – 05/2002: Technical Knowledge Manager, Inner Strategic Research (Madrid, Spain). Responsibilities:

- Statistical consulting in marketing research.
- Teaching of statistics.
- Development, implementation, and adaptation, of statistical techniques of use in marketing research.

01/2001 – 07/2001: Data Mining Expert, NewKnow Network (Madrid, Spain). Responsibilities:

- Evaluation, comparison, and benchmarking of data mining algorithms.
- Combination and modification of data mining algorithms and statistical methods.

PUBLICATIONS

A total of 40 articles in peer-reviewed journals and four chapters in books (34 articles and 3 chapters since joining CNIO in May 2002). Except for one, all of the papers in journals with impact factor where R. Diaz-Uriarte is corresponding author have been published in journals within the top 25 % of their corresponding category (most of them in “Mathematical and Computational Biology”).

Articles in peer-reviewed journals with impact factor

- *¹ Rueda OM, Diaz-Uriarte R. 2009 a. Detection of recurrent copy number alterations in the genome: taking among-subject heterogeneity seriously. **BMC Bioinformatics**, 10: 308
- * Rueda OM, Diaz-Uriarte R. 2009 b. Finding Recurrent Copy Number Alteration Regions: A Review of Methods. **Current Bioinformatics**. In press.
- * Rueda OM, Diaz-Uriarte R. 2009 c. RJaCGH: Bayesian analysis of aCGH arrays for detecting copy number changes and recurrent regions. **Bioinformatics**, 25: 1959-1960
- Babel I, Barderas R, Diaz-Uriarte R, Martínez-Torrecuadrada JL, Sánchez-Carbayo M, and Casal JI. 2009. Identification of tumour-associated autoantigens for the diagnosis of colorectal cancer in serum using high-density protein microarrays. **Molecular & Cellular Proteomics**, 8: 2382-2395.
- * Morrissey E, Diaz-Uriarte, R. 2009. Pomelo II: finding differentially expressed genes. **Nucl. Acids.Res.**, 37: W581-W586
- Sequeira-Mendes J, Diaz-Uriarte R, Apedaile A, Huntley D, Brockdorff N, Gomez M. 2009. Transcription Initiation Activity Sets Replication Origin Efficiency in Mammalian Cells. **PLoS Genetics**, 5: e1000446
- Infante A, Laresgoiti U, Fernández-Rueda J, Fullaondo A, Galán J, Diaz-Uriarte R, Malumbres M, Field SJ, Zubiaga AM. 2008. E2F2 represses cell cycle regulators to maintain quiescence. **Cell Cycle**, 7: 3915-3927.
- Alibés, A., Cañada, A., Diaz-Uriarte, R. 2008. PaLS: filtering common literature, biological terms and pathway information. **Nucl. Acids Res.**, 36: W364-W367.
- * Diaz-Uriarte, R. 2008. SignS: a parallelized, open-source, freely available, web-based tool for gene selection and molecular signatures for survival and censored data. **BMC Bioinformatics**, 9: 1-11.
- * Rueda, O. M. and Diaz-Uriarte, R. 2008. A response to Yu et al. 'A forward-backward fragment assembling algorithm for the identification of genomic amplification and deletion breakpoints using high-density single nucleotide polymorphism (SNP) array', BMC Bioinformatics 2007, 8: 145. **BMC Bioinformatics** 2007, 8: 394-400
- * Rueda, O. M., and Diaz-Uriarte, R. 2007 a. Flexible and accurate detection of genomic copy-number changes from aCGH. **PLoS Computational Biology**, 3: 1115-1122
- * Diaz-Uriarte, R., Alibés, A., Morrissey, E. R., Cañada, A., Rueda, O.M., Neves, M. L. 2007. Asterias: integrated analysis of expression and aCGH data using an open-source, web-based, parallelized software suite. **Nucleic Acids Research**, 35: W75-W80
- Alibés, A., Yankilevich, P., Cañada, A., and Diaz-Uriarte, R.. 2007. IDconverter and IDClight: Conversion and annotation of gene and protein IDs. **BMC Bioinformatics**, 8: 1-9.
- Moneo, V., Serelde, B. G., Leal, J.F.M., Blanco-Aparicio, C., Diaz-Uriarte, R., Aracil, M., Tercero, J.C., Jimeno, J., and Carnero, A. 2007. Levels of p27kip1 determine Aplidin sensitivity. **Therapeutics**, 6: 1310-1316.
- Rodríguez, A., Villuendas, R., Yáñez, L., Gómez, M. E., Diaz-Uriarte, R., Pollán, M., Hernández, N., de la Cueva, P., Marín, M. C., Swat, A., Ruiz, E., Cuadrado, M. A., Conde, E., Lombardía, L., Cifuentes, F., Gonzalez, M., García-Marco J. A., and Piris, M. A. 2007. Molecular heterogeneity in chronic lymphocytic leukemia is dependent on BCR signaling: clinical correlation. **Leukemia**, 21: 1984-1991.
- Collado, M., Garcia, V., Garcia, J.M., Alonso, I., Lombardía, L., Diaz-Uriarte, R., López, L. A., Zaballos, A., Bonilla, F., and Serrano, M. 2007. Genomic Profiling of Circulating Plasma RNA for the Analysis of Cancer. **Clinical Chemistry**, 53, 1860-1863.
- * Diaz-Uriarte, R. 2007. GeneSrF and varSelRF: a web-based tool and R package for gene selection and classification using random forest. **BMC Bioinformatics**, 8: 328-343.

¹Papers where where R. Diaz-Uriarte is the corresponding author are marked with asterisks.

- * Diaz-Uriarte, R., Alvarez de Andrés, S. 2006. Gene selection and classification of microarray data using random forest. **BMC Bioinformatics**, 7:3.
- Vaquerizas, J., Conde, L., Yankilevich, P., Cabezon, A., Minguéz, P., Diaz-Uriarte, R., Al-Shahrour, F., Herrero, J., Dopazo, J. 2005. GEPAS, an experiment-oriented pipeline for the analysis of microarray gene expression data **Nucleic Acids Research**, 33: W616-W620.
- Al-Shahrour, F., Diaz-Uriarte, R., Dopazo, J. 2005. Discovering molecular functions significantly related to phenotypes by combining gene expression data and biological information. **Bioinformatics** , 21: 2988-2993.
- Meléndez, B., Martínez-Delgado, B., Cuadros, M., Fernández, V., Diaz-Uriarte, R., Benítez, J.. 2005. Identification of amplified and highly expressed genes in amplicons of the T-cell line huT78 detected by cDNA microarray CGH. **Molecular Cancer**, 4.
- Alvarez de Andrés, S., Diaz-Uriarte, R., Osorio, A., Barrosa, A., Paz, M.F., Honrado, E., Rodríguez, R., Urioste, M., Valle, L., Diez, O., Cigudosa, J. C., Dopazo, J., Esteller, M., Benítez, J. 2005. A predictor based on the somatic changes of the BRCA1/2 breast cancer tumors identifies the non-BRCA1/2 tumors with BRCA1 promoter hypermethylation. **Clinical Cancer Research**, 11: 1146-1153.
- Meléndez, B., Diaz-Uriarte, R., Martínez-Ramírez, A., Fernández-Piqueras, J., Rivas, C., Dopazo, J., Martínez-Delgado, B., Benítez, J. 2004. Gene expression analysis on chromosomal regions of gain or loss in genetic material detected by comparative genomic hybridization. **Genes Chromosomes and Cancer**, 41: 353-365.
- * Vaquerizas, J., Dopazo, J., Diaz-Uriarte, R. 2004. DNMAID: web-based diagnosis and normalization for microarray data. **Bioinformatics**, 20: 3656-3658.
- Herrero, J., Vaquerizas, J., Al-Shahrour, F., Conde, L., Mateos, A., Santoyo, J., Diaz-Uriarte, R., Dopazo, J. 2004. New challenges in gene expression data analysis and the extended GEPAS. **Nucleic Acids Research**, 32: W485-W491.
- Al-Shahrour, F., Diaz-Uriarte, R., Dopazo, J. 2004. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. **Bioinformatics** 20: 578-580.
- Martinez-Delgado, B., Meléndez, B., Cuadros, M., Alvarez, J., Castrillo, J.M., Ruiz de la Parte, A., Mollejo, M., Bellas, C., Diaz-Uriarte, R., Lombardía, L., Al-Shahrour, F., Domínguez, O., Cascon, A., Robledo, M., Rivas, C., Benitez, J. 2004. Expression Profiling of T-Cell Lymphomas Differentiates Peripheral and Lymphoblastic Lymphomas and Defines Survival Related Genes **Clinical Cancer Research** , 10: 4971-4982.
- Moreno-Bueno, G., Sánchez-Estévez, C., Cassia, R., Rodríguez-Perales, S., Diaz-Uriarte, R., Domínguez, O., Hardisson, D., Andujar, M., Prat, J., Matias-Guiu, X., Cigudosa, J.C., Palacios, J. 2003. Differential Gene Expression Profile in Endometrioid and Nonendometrioid Endometrial Carcinoma: STK15 Is Frequently Overexpressed and Amplified in Nonendometrioid Carcinomas. **Cancer Research**, 63: 5697-5702.
- Herrero, J., Al-Shahrour, F., Diaz-Uriarte, R., Mateos, Á., Vaquerizas, J.M., Santoyo, J., Dopazo, J. 2003 a. GEPAS, a web-based resource for microarray gene expression data analysis **Nucleic Acids Research**, 31: 3461-3467.
- Herrero, J., Diaz-Uriarte, R., Dopazo, J. 2003 b. An approach to inferring transcriptional regulation among genes from large-scale expression data. **Comparative and Functional Genomics**, 4: 148-154.
- Herrero, J., Diaz-Uriarte, R., Dopazo, J. 2003 c. Gene Expression Data Preprocessing. **Bioinformatics**, 19: 655-656.
- Díaz-Uriarte, R. 2002. Incorrect analysis of crossover trials in animal behaviour research. **Animal Behaviour**, 63: 815-822.
- Díaz-Uriarte, R. 2001. Territorial intrusion risk and antipredator behaviour: a mathematical model. **Proceedings of the Royal Society of London**, Series B, 268: 1165-1173.
- Díaz-Uriarte, R. 1999. Anti-predator behaviour changes following an aggressive encounter in the lizard *Tropidurus hispidus*. **Proceedings of the Royal Society of London**, Series B, 266: 2457-2464.
- Garland, T. Jr., & Díaz-Uriarte, R. 1999. Polytomies and independent contrasts: an examination of the bounded degrees of freedom approach. **Systematic Biology**, 48: 547-558.
- Díaz-Uriarte, R., & Garland, T. Jr. 1998. Effects of branch length errors on the performance of phylogenetically independent contrasts. **Systematic Biology**, 47: 654-672.
- Bauwens, D., & Díaz-Uriarte, R. 1997. Covariation of reproductive traits of lacertid lizards: a comparative study. **The American Naturalist**, 149: 91-111.

- Díaz-Urriarte, R., & Garland T. Jr. 1996. Testing hypotheses of correlated evolution using phylogenetically independent contrasts: sensitivity to deviations from Brownian motion. **Systematic Biology**, 45: 27–47.
- Díaz, J. A., Díaz-Urriarte, R., Rodríguez, A. 1996. Influence of behavioral thermoregulation on the use of vertical surfaces by Iberian wall lizards, *Podarcis hispanica*. **Journal of Herpetology**, 30: 548–552.

Articles in peer-reviewed journals without impact factor

- Urduinguio RG, Lopez-Serra L, Lopez-Nieva P, Alaminos M, Diaz-Urriarte R, Fernandez AF, Esteller, M. 2008. Mecp2-Null Mice Provide New Neuronal Targets for Rett Syndrome. **PLoS ONE**, 3: e3669.
- * Díaz-Urriarte, R. and Rueda, O. M. 2007 b. ADaCGH: a parallelized web-based application and R packages for the analysis of aCGH data. **PLoS ONE**, 8: 394-400.
- * Alibés A, Morrissey E, Cañada A, Rueda O, Casado D, Yankilevich P, Diaz-Urriarte, R. 2007. Asterias: a parallelized web-based suite for the analysis of expression and aCGH data **Cancer Informatics**, 3: 1-9.

Book chapters

- * Díaz-Urriarte, R. 2005. Supervised methods with genomic data: a review and cautionary view. Pp. 193–214 en F. Azuaje, and J. Dopazo (eds.), **Data analysis and visualisation in genomics and proteomics**, John Wiley & Sons, New York.
- * Díaz-Urriarte, R, Al-Shahrour, F. & Dopazo, J. 2003. Use of GO terms to understand the biological significance of microarray differential gene expression data. Pp. 233-247 en K. F. Johnson and S. M. Lin (eds.), **Methods of microarray data analysis (CAMDA 2002)**, Kluwer Academic Press, New York.
- Al-Shahrour, F., Herrero, J., Mateos, Á., Santoyo, J., Díaz-Urriarte, R., Dopazo, J. 2003. Using Gene Ontology on genome-scale studies to find significant associations of biologically relevant terms to group of genes. pp. 43-52 in **Neural Networks for Signal Processing, XIII**, IEEE Press, New York.
- Garland, T. Jr., Martin, K. L. M., & Díaz-Urriarte, R. 1997. Reconstructing ancestral trait values using squared-change parsimony: plasma osmolarity at the origin of amniotes. Pp. 425–501 in: S. S. Sumida & K. L. M. Martin (eds.), **Amniote origins: completing the transition to land**, Academic Press, San Diego.

BIOINFORMATICS AND STATISTICS SOFTWARE

Developer and programmer of over nine applications and R packages for the analysis of genomic data (ADaCGH, Asterias, DNMad, FatiGO, GeneSrF, Pomelo and Pomelo II, RJaCGH, SignS, Tnasas), including analysis of differential expression, aCGH data, class prediction, and normalization. Most applications use parallel and distributed computing (with MPI) and have a web-based interface. Author also of programs for the analysis of phylogenetic data (PHYLOGR, contributions to PDAP), recording of behavioral data, and evaluation of evolutionarily stable strategies using genetic algorithms.

Main programming languages used for the above applications: R, C, C++, Python, Pascal, Reverse Polish Lisp.

Management of computational systems for bioinformatics and statistics

Systems administrator and manager of two Beowulf clusters of 30 nodes (one of them Intel Xeon based, the other AMD Opteron based) at CNIO (used both for research and application web-serving), including design of configuration, installation and maintenance of operating system, application migration, load balancing, and parallelization of computations (using MPI).

PI OF COMPETITIVELY FUNDED SCIENTIFIC PROJECTS

- “Regiones recurrentes y paisajes genómicos y epigenéticos: integración con árboles oncogenéticos”. Ministerio de Ciencia e Innovación. 2009-2012. PI: Ramón Diaz-Uriarte.
- “Desarrollo de modelos estadísticos para el análisis conjunto de arrays de CGH y de expresión génica para el estudio de neoplasias humanas”. Fundación Mutua Madrileña Automovilista. 07/2005-07/2008. PI: Ramon Diaz-Uriarte.
- “Perfil de expresión genética en carcinoma no microcítico de pulmón: utilidad de las biopsias endoscópicas, correlación con la histología y predicción pronóstica en pacientes operables.” Fondo de Investigación Sanitaria. 01/2005-12/2007. PI: José Palacios and Ramon Diaz-Uriarte.
- “Extracción de conocimiento de los microarrays de ADN usando métodos estadísticos.”. Ministerio de Ciencia y Tecnología. 12/2003-12/2006. PI: Ramón Diaz-Uriarte.

INVITED CONFERENCES AND COURSES

Invited conferences in statistical meetings (e.g., X Conferencia Española de Biometría), and teacher in several official master programs (Bioinformatics and Computational Biology, Molecular Oncology, Molecular Pathology, etc).

OTHER PROFESSIONAL ACTIVITIES AND AWARDS

- Ramón y Cajal researcher (December 2002)
- Two year scholarship from “Fundación La Caixa” (September 1993 – June 1995) for postgraduate work in the USA.
- **Member of** Royal Statistical Society, American Statistical Association, International Biometric Society.
- **Reviewer for** over 19 scientific journals, including Bioinformatics, Nature Biotechnology, IEEE Transactions in Computational Biology and Bioinformatics, Genome Biology, BMC Bioinformatics, Statistical Applications in Genetics and Molecular Biology, Journal of Multivariate Analysis, Electronic Journal of Statistics.

January 2010