

COST Full Proposal Reference oc-2010-1-7183
For the implementation of a European Concerted Research Action

Bingos

(Bioinformatics Next Generation Open Software)

List of references

Proposing country	Spain
Action proposer:	Dr. Oswaldo Trelles
Institution:	University of Malaga Department of Computer Architecture ETSI Informatica, Campus de Teatinos Malaga 29071, Spain Phone: +34 952 13 2823 Fax: +34 952 13 2790 Email: ots @ ac.uma.es, ortrelles @ uma.es

- Oswaldo Trelles; "On the parallelisation of bioinformatics applications"; Briefings in bioinformatics; 2(181-194); 2001
- C.O.S.Sorzano, J.M. Carazo and Trelles O.; "Command line interfaces can be efficiently brought to graphics: COLIMATE (the COMmand LIne MATE)"; 32(873-887); 2002
- Pérez A.J., Rodríguez, A., Trelles O., Thode G.; "A computational strategy for protein function assignment which addresses the multidomain problem"; Comparative and Functional Genomics; 3(423-440), 2002
- Jorge García de la Nava, Daniel Franco Santaella, Jesús Cuenca Alba, José María Carazo, Oswaldo Trelles, Alberto Pascual-Montano; "Engene: The processing and exploratory analysis of gene expression data"; Bioinformatics, 19(657-658); 2003
- Jorge García de la Nava, Sacha van Hijum and Oswaldo Trelles; "PreP: gene expression data pre-processing"; Bioinformatics; 19(2328-2329); 2003
- A.J. Pérez, G. Thode, and O. Trelles; "AnaGram: protein function assignment"; Bioinformatics; 20(291-292); 2004
- Rodríguez, A.; Shotton,D.; Guil, N. and Trelles, O.; "Automatic analysis of the content of cell biological video and database organization of their metadata descriptors"; IEEE Transactions in Multimedia; 6(119-128); 2004
- Jorge García de la Nava, Sacha A.F.T. van Hijum and Oswaldo Trelles; "Saturation and quantization reduction in microarray experiments using two scans at different sensitivities"; 3(article 11); 2004
- Rodríguez, A.; Carazo, J.M. and A., Trelles O.; "Mining Association Rules from Biological Databases"; Journal of the American Society for Information Science and Technology; 56(493-504); 2005
- Rodríguez, A.; Shotton,D.; Guil, N. and Trelles, O.; "Analysis and Description of the Semantic Content of Cell Biological Videos"; Multimedia Tools and Applications 25(37-38), 2005
- Trelles Oswaldo and Rodríguez Andrés; "Parallel Metaheuristics in Bioinformatics: A new class of algorithms"; Wiley Series on Parallel and Distributed Computing; (517-549), 2005
- Ismael Navas-Delgado, Maria del Mar Rojano-Muñoz, Sergio Ramírez, Antonio J. Pérez, Eduardo Andrés León; Jose F. Aldana-Montes, and Oswaldo Trelles; "Intelligent client for integrating bioinformatics services."; Bioinformatics; 22(106-111), 2006
- P. Carmona-Sáez, M. Chagoyen, A. Rodríguez, O. Trelles, J. M. Carazo and A. Pascual-Montano.; "Integrated analysis of gene expression by Association Rules Discovery", BMC Bioinformatics, 7(54), 2006
- José María Sayagués, María Dolores Taberero, Angel Maíllo, Oswaldo Trelles, Ana Belén Espinosa, María Eugenia Sarasquete, Ana Rasillo, Marta Merino, Angel Santos-Briz and Alberto Orfao; "Microarray-based analysis of Spinal vs intracranial meningiomas: different clinical, biological and genetic characteristics associated with distinct patterns of gene expression"; Journal of Neuropathology and Experimental Neurology; 65(445-454), 2006

J.F. Aldana, M. Roldán-Castro, I. Navas, M.M. Roldán-García, M. Hidalgo-Conde, and O.Trelles; “Bio-Broker: a tool for integration of biological data sources and data analysis tools”; *Software, Practice and Experience*, 36(1585-1604), 2006

Manuel J. Martín-Vázquez; Mario A. Trelles; Alejandro Sola; R. Glen Calderhead and Oswaldo Trelles; “A new user-friendly software platform for systematic classification of skin lesions to aid in their diagnosis and prognosis”, *J. Laser in Medical Science* 21(54-60), 2006

van Hijum, S. AFT; Baerends, R. JS; Zomer, Aldert L; Karsens, H A; Martin-Requena, V; Trelles, O; Kok, J and Kuipers, O. P.; “Supervised Lowess normalization of comparative genome hybridization data – application to lactococcal strain comparisons”, *BMC Bioinformatics*, 9:93, Feb. 2008

The BioMoby Consortium: Mark Wilkinson et al. (including O. Trelles); “Interoperability with Moby 1.0 – It’s better than sharing your toothbrush!”, *Briefings in Bioinformatics*; 9(3):220-31, 9(3):220-31, May, 2008

V Martín-Requena; A Muñoz-Merida; M.G Claros and O.Trelles; “Prep+07: improvements of a user friendly tool to pre-process and analyse microarray data”; *BMC-Bioinformatics*, in-press, Jan.2009

E. Gutiérrez, O. Plata and E.L. Zapata, “An Analytical Model of Locality-Based Parallel Irregular Reductions”, *Parallel Computing*, Vol. 34(3), pp. 133-157, 2008

G.P. Trabado, O. Plata and E.L. Zapata, “A Survey on Grid Architectures for Bioinformatics”, *Engineering the Grid: Status and Perspectives*, Chapter 7, American Scientific Pub., B.Di Martino, J. Dongarra, A. Hoise, L.T. Yang and H. Zima Editors, 2006

O. Plata, R. Asenjo, E. Gutierrez, F. Corbera, M.A. Navarro and E.L. Zapata, “On the Parallelization of Irregular and Dynamic Programs”, *Parallel Computing*, Vol. 31(6), pp. 544-562, 2005

E. Gutierrez, S. Romero, L.F. Romero, O. Plata and E.L. Zapata, “Parallel Techniques in Irregular Codes: Cloth Simulation as Case of Study”, *Parallel and Distributed Computing*, Vol. 65(2), pp. 424-436, 2005

S. Romero, E. Gutierrez, L.F. Romero, O. Plata and E.L. Zapata, “Parallelization Issues of a Code for Physically-based Simulation of Fabrics”, *Computer Physics Communications*, Vol. 162(3), pp. 188-202, 2004

E. Gutiérrez, O. Plata and E.L. Zapata, “Data Partitioning Based Parallel Irregular Reductions”, *Concurrency, Practice and Experience*, Vol. 16(2), pp. 155-172, 2004

O. Plata, E. Gutiérrez and E.L. Zapata, “On Improving the Performance of Data Partitioning Oriented Parallel Irregular Reductions”; *J. Systems Architecture*; Vol. 49(10), pp. 63-74; 2003

G.P. Trabado, O.G. Plata and E.L. Zapata, “On the Parallelization of Molecular Dynamics Codes”, *Computer Physics Communications*, Vol. 147(1-2), pp. 711-715, 2002

M. Amor, F. Argüello, J. López, O. Plata and E.L. Zapata, “A Data Parallel Formulation for Divide and Conquer Algorithms”, *The Computer Journal*, Vol 44 (4), pp. 303-320, 2001

M.G. Claros, S. Brunak, G. von Heijne, “Prediction of N-terminal protein sorting signals”, *Curr Opin Struct Biol.*, 7, 394-398, 1997

- R. Bautista, F.M. Cánovas, M.G. Claros, "Genomic evidence for a repetitive nature of the RAPD polymorphisms in *Olea europaea* (olive-tree)", *Euphytica*, 130, 185-190, 2003
- D. González-Halphen, S. Funes, X. Pérez-Martínez, A. Reyes-Prieto, M.G. Claros, E. Davidson, M.P. King, "Genetic correction of mitochondrial diseases: using the natural migration of mitochondrial genes to the nucleus in chlorophyte algae as a model system", *Ann. N.Y. Acad. Sci.*, 1019, 232-239, -2004
- R. Bautista, D.P. Villalobos, S. Díaz-Moreno, F.R. Cantón, F.M. Cánovas, M.G. Claros, "Toward a *Pinus pinaster* bacterial artificial chromosome library", *Ann Forest Sci.*, 64, 855-864, 2007
- J. Falgueras, A. J. Lara, F.R. Cantón, G. Pérez-Trabado, M.G. Claros, "SeqTrim: a validation and trimming tool for all purpose sequence reads". En *Innovations in hybrid intelligent systems*, ASC 44. Coord. E. Corchado, J.M. Corchado y A. Abraham, Springer, Berlín, pp. 353-360, 2007
- A. J. Lara, G. Pérez-Trabado, D. P. Villalobos, S. Díaz-Moreno, F.R. Cantón, M.G. Claros, "A web tool to discover full-length sequences: Full-Lengther", *Innovations in hybrid intelligent systems*, ASC 44. Coord. E. Corchado, J.M. Corchado y A. Abraham, Springer, Berlín, pp. 361-368, 2007
- JR Bilbao-Castro, A Merino, I Garcia, JM Carazo, JJ Fernandez. Parameter optimization in 3D reconstruction on a large scale grid. *Parallel Computing Vol. 33*, pp. 250-263, 2007
- Sjors H.W. Scheres, Haixiao Gao, Mikel Valle, Gabor T. Herman, Paul P.B. Eggermont, Joachim Frank, Jose-Maria Carazo. Disentangling Conformational States of Macromolecules in 3D Cryo-EM through Likelihood Optimization. *Nature Methods*. 4(1):27-9. 2007
- Sorzano CO, Marabini R, Pascual-Montano A, Scheres SHW, Carazo JM. Optimization problems in electron microscopy of single particles. *Annals of Operations Research*. 148(1) 133-165, 2006.
- Pascual-Montano, Carazo, J.M., Kochi, K, Lehmann, D, Pascual-Marqui, R.D. Non-smooth Non-Negative Matrix Factorization (nsNMF). *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 2006. 28(3), 403 – 415.
- Mikel Karmel Valle, Xiaojiang S Chen, Luis E Donate, Ellen Fanning, José M Carazo. Structural basis for the cooperative assembly of large t antigen on the origin of replication. *JMB*. 2006. 357: 1295 - 1305
- Leitner F, et.al; "Introducing meta-services for biomedical information extraction"; *Genome Biol*. 2008;9 Suppl 2:S6. Epub 2008 Sep 1. PMID: 18834497
- BioMoby Consortium, Wilkinson MD, et al. "Interoperability with Moby 1.0--it's better than sharing your toothbrush!"; *Brief Bioinform*. 2008 May;9(3):220-31. Epub 2008 Jan 31. PMID: 18238804
- Tress ML, Wesselink JJ, Frankish A, López G, Goldman N, Löytynoja A, Massingham T, Pardi F, Whelan S, Harrow J, Valencia A. "Determination and validation of principal gene products"; *Bioinformatics*. 2008 Jan 1;24(1):11-7. Epub 2007 Nov 15. PMID: 18006548
- Fernández JM, Hoffmann R, Valencia A.; "iHOP web services."; *Nucleic Acids Res*. 2007 Jul;35(Web Server issue):W21-6. Epub 2007 May 7. PMID: 17485473

Abascal F, Valencia A.; “Automatic annotation of protein function based on family identification”; *Proteins*. 2003 Nov 15;53(3):683-92. PMID: 14579359

Blaschke C, Valencia A. “Automatic ontology construction from the literature”, *Genome Inform.* 2002;13:201-13. PMID: 14571389

Blaschke C, Valencia A., “The Frame-Based Module of the SUISEKI Information Extraction System”; *EEE Intelligent Systems*. Volume 17 , Issue 2 (March 2002) Pages: 14 – 20

Talavera D, Morreale A, Hospital, A, Ferrer Costa C, Gelpi J.L., de la Cruz X, Meyer T, Soliva R, Luque F.J., Orozco, M. *A fast method for the determination of fractional contributions to solvation in proteins*. *Protein Science* (2006) 15(11):2525-33

Rueda M., Ferrer-Costa, C, Meyer T., Pérez, A., Camps, J., Hospital, A., Gelpi, J.L., Orozco M. *A consensus view of protein dynamics*. *Proc. Natl. Acad. Sci* (2007) 104(3):796-801

Soliva R., Gelpi J.L.; Almansa C., Virgili M., Orozco M. *Dissection of the recognition properties of p38 MAP kinase. Determination of the binding mode of a new pyridinyl-heterocycle inhibitor family*. *J. Med Chem* (2007) 25;50(2):283-93.

Reig N., del Rio C., Casagrande F., Ratera M., Gelpi J.L., Torrents D., Henderson PJ., Baldwin SA, Zorzano A., Fotiadis D., Palacín M. *Functional and structural characterization of the first prokaryotic member of the LAT family: a model for APC transporter*. *J.Biol. Chem* (2007) 282(18):13270-81

The Biomoby Consortium. *Interoperability with Moby 1.0 – It’s Better than Sharing Your Toothbrush!* *Briefings in Bioinformatics* (2008) 9(3):220-31

Castaño, T.; Wang, H.; Campillo, N.E.; Ballester S.; Gonzalez-Garcia, C.; Hernandez J.; Pérez C.; Cuenca J.; Pérez-Castillo A.; Martínez A.; Huertas O.; Gelpí J.L.; Luque F.J.; Ke H.; Gil C. *Synthesis and biological evaluation of thioxoquinazoline derivatives as phosphodiesterase 7 inhibitors*. *Chem. Med. Chem.* (2009). (in press)

Alberts R, Vera G, Jansen RC: affyGG: computational protocols for genetical genomics with Affymetrix arrays. *Bioinformatics*; 2008, 24(3433-434)

Vera G, Jansen RC, Suppi RL; “R/parallel - speeding up bioinformatics analysis with R”, *BMC Bioinformatics*, 2008. 22;9(1):390

Journals

Remo Suppi, Daniel Fernández, Emilio Luque

Fish Schools: PDES Simulation and Real Time 3D Animation.

LNCS 3-540-21946-3, ISSN 0302-9743

Springer Verlag EC, Vol.: 3019, 505-512, 2004

D. Mostaccio, R. Suppi, E. Luque; “Simulation of Ecologic Systems Using MPI”; LNCS 3666, 3-540-290095, ISSN 0302-9743. Springer Verlag EC. Vol.: 3666, 449-456, 2005

Diego Mostaccio, Christianne Dalforno, Remo Suppi and Emilio Luque. “Distributed Simulation of Large-Scale Individual Oriented Models”. *Journal of Computer Science & Technology*, ISSN 1666-6038; Vol.: 6, N° 2. 59-65, 2006

D. Mostaccio, R. Suppi, E. Luque. “Using Distributed Events Simulation for Individual Oriented Models”. *Intl Mediterranean Multimodeling Multiconference, IM3*, ISBN: 2-9520712-3-3. Vol.: I, 105-110. 2005. SCS, USA.

X.Y. Yang, P. Hernández, F. Cores, L. Souza, A. Ripoll, R. Suppi, E. Luque. "Multi-Collaboration Domain Multicast P2P Delivery Architecture for VoD System". IEEE International Conference on Communications (ICC 2006). ISBN: 1-4244-0355-3 Vol.: 4, 1633-1699. 2006. IEEE.

X.Y. Yang, P. Hernández, F. Cores, L. Souza, A. Ripoll, R. Suppi, E. Luque. "DVoDP2P: Distributed P2P Assisted Multicast VoD Architecture". IEEE International Parallel & Distributed Processing Symposium. CDROM, 2006. IEEE

Diego Javier Ballardini, Leandro Souza, Remo Suppi. "A Network Traffic Scheduler for a VoD Server on the Internet"; Signal Processing and Multimedia Applications. SIGMAP : ISBN: 972-8865-64-3; Vol.: I, 260-268. 2006

"*Frequency Sorting Method for Spectral Analysis of DNA Sequences*", A. Bucur, J. van Leeuwen, N. Dimitrova, C. Mittal, IEEE BIBM 2008

"*Grid-enabled High Angular Fiber Tracking*", J. van Leeuwen, A. Bucur, HealthGrid2008 (poster)

J. van Leeuwen and A. Bucur, "*High Angular Fiber Tracking on the Grid*", *HEALTHINF*, p. 168-174, 2009.

"*Overlap of selected ontologies in the context of the breast cancer domain*" Z. Aleksovski, R. Vdovjak, SIIM 2009 (to appear)

Johannes F, Colot V, Jansen RC (2008), Epigenome dynamics: a quantitative genetics perspective. *Nature Review Genetics* 9:883-890

Fu J, Swertz MA, Keurentjes JJB and Jansen RC (2007), MetaNetwork: a computational protocol for the genetic study of metabolic networks *Nature Protocols* 2:685-694

Kammenga JE, Herman MA, Ouborg NJ, Johnson L, Breitling R (2007), Microarray challenges in ecology, *Trends in Ecology and Evolution* 5:273-279

Swertz MA, Jansen RC (2007), Beyond standardization: dynamic software infrastructures for systems biology, *Nature Reviews Genetics* 8:235-243

Keurentjes JJB, Fu J, Vos de CHR, Lommen A, Hall RD, Bino RJ, Plas van der LHW, Jansen RC, Vreugdenhil D, Koornneef M (2006), The genetics of plant metabolism; *Nature Genetics* 38:842-849

Bystrykh L, Weersing E, Dontje B, Sutton S, Pletcher MT, Wiltshire T, Su AI, Vellenga E, Wang J, Manly KF, Lu L, Chesler EJ, Alberts R, Jansen RC, Williams RW, Cooke MP, de Haan G. (2005), Uncovering regulatory pathways that effect hematopoietic stem cell function using 'genetical genomics', *Nature Genetics* 37:225-32.

Complex Trait Consortium (incl Jansen, R.C.) (2004), The Collaborative Cross: A Community Resource for the Genetic Analysis of Complex Traits, *Nature Genetics* 36:1133-37.

Jansen, R.C. and J.P. Nap (2004), Regulating gene expression: surprises still in store, *Trends in Genetics* 20:223-225.

Complex Trait Consortium (incl Jansen, R.C.) (2003) The nature and identification of quantitative trait loci: a community's view, *Nature Reviews Genetics* 4:911-916.

Jansen, R.C. (2003), Studying complex biological systems using multifactorial perturbation, *Nature Reviews Genetics* 4:145-151.

Jansen RC, Nap JP, Mlynarova L (2002), Errors in genomics and proteomics, *Nature Biotechnology* 20:19.

Jannink J, Bink MC, Jansen RC (2001), Using complex plant pedigrees to map valuable genes, *Trends in plant Science* 6:337-42.

Jansen RC, Nap JP (2001), Genetical genomics: the added value from segregation, *Trends in Genetics* 17:388-91.

Abad, P.; Gouzy, J.; Aury, J.M.; Tytgat, T.O.G.; Smant, G. (2008); Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*; *Nature Biotechnology* 26 (8). - p. 909 - 915.

Li, Y.; Alda Alvarez, O.; Gutteling, E.W.; Tijsterman, M.; Fu, J.; Riksen, J.A.G.; Hazendonk, E.; Prins, J.C.P.; Plasterk, R.H.A.; Jansen, R.C.; Breitling, R.; Kammenga, J.E. (2006); Mapping Determinants of Gene Expression Plasticity by Genetical Genomics in *C. elegans*; *Plos Genetics* 2 (12). - p. 0001 - 0007.

Butterbach, P.B.E; Sloomweg, E.J.; Koropacka, K.B.; Prins, J.C.P.; Roosien, J.; Spiridon, L.; Petrescu, A.; Smant, G.; Bakker, E.H.; Bakker, J.; Govere, A. (2008); Functional constraints and evolutionary dynamics of the Rx1/Gpa2 cluster in potato; In: XIII International Congress on Molecular Plant-Microbe Interactions, July 21-27, 2007, Sorrento, Italy. - *Biology of plant-microbe interactions* 6 .Sorrento, Italy : MPMI,

Roze, E.H.A.; Hanse, B.; Mitreva, M.; Vanholme, B.; Bakker, J.; Smant, G. (2008); Mining the secretome of the root-knot nematode *Meloidogyne chitwoodi* for candidate parasitism genes; *Molecular Plant Pathology* 9 (1). - p. 1 - 10.

Smant, G.; Qin Ling, --; Govere, A.; Schots, A.; Helder, J.; Bakker, J. (2003); Gene Discovery in Sedentary Plant-Parasitic Nematodes; In: *Plant-Microbe Interactions / Stacey, G., Keen, N.T., . - St. Paul, Minnesota, USA : The American Phytopathological Society, - p. 285 - 307.*

Qin Ling, ; Prins, P.; Jones, J.T.; Popeijus, H.; Smant, G.; Bakker, J.; Helder, J. (2001); GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP; *Nucleic Acids Research* 29 (2001) - p. 1616 - 1622.

Qin Ling, --; Prins, J.C.P.; Helder, J. (2005); Linking cDNA-AFLP-based gene expression patterns and ESTs; In: *Differential display methods and protocols. - Totowa : Humana press, (Methods in molecular biology 317)*

Plant degradation: a nematode expansin acting on plants.; Qin L, Kudla U, Roze EH, Govere A, Popeijus H, Nieuwland J, Overmars H, Jones JT, Schots A, Smant G, Bakker J, Helder J.; *Nature*. 2004 Jan 1;427(6969):30.

Marra D, Romano P, Integrating mutation data of the TP53 human gene in the bioinformatics network environment, Proc. 1st International Conference on Bioinformatics Research and Development BIRD '07, March 12-14, 2007, Berlin, Springer Lecture Notes in Bioinformatics LNBI 4414, pp. 453-463, 2007, Springer Verlag Berlin Heidelberg 2007

Romano P, Bartocci E, Bertolini G, De Paoli F, Marra D, Mauri G, Merelli E, Milanese L, Biowep: a workflow enactment portal for bioinformatics applications, *BMC Bioinformatics* 2007, 8 (Suppl 1):S19

Bartocci E, Cacciagrano D, Cannata N, Corradini F, Merelli E, Milanese L, Romano P, An Agent-based Multilayer Architecture for Bioinformatics Grids, *IEEE Transactions on NanoBioscience (IEEE-TNB)*, 2007, 6(2):142-148

Romano P, Automation of in-silico data analysis processes through workflow management systems, *Briefings in Bioinformatics* 2008, 9(1):57-68

Romano P, Marra D, SWS: accessing SRS sites contents through Web Services, *BMC Bioinformatics* 2008, 9(Suppl 2):S15

Romano P, Splendiani A, Applications of Semantic Web methodologies and techniques to biology and bioinformatics, *Reasoning Web 2008, Springer Lecture Notes in Computer Sciences LNCS 5224*, pp. 200-239, 2008, Springer Verlag Berlin Heidelberg

Corradi L, Mirisola V, Porro I, Torterolo L, Fato M, Romano P, Pfeffer U. Engineering Functional Genomics: a use case with an Interdisciplinary Approach, *BMC Bioinformatics* (submitted)

P. Eklund, M. Kaufmann, Hierarchical Wiring in Multigrids, *CONPAR90 -- VAPP IV, Joint International Conference on Vector and Parallel Processing* (Ed. H. Burkhardt), *Lecture Notes in Computer Science 457*, Springer, 1990, 423-434.

P.Eklund, R. Helgerson, H. Lindgren, Towards refinement of clinical evidence using general logics, L. Rutkowski et al. (EdS.), *ICAISC 2008, Lecture Notes In Artificial Intelligence {5097}* (2008), 1029-1000.

P. Eilund, J. Karlsson, A. Naslund, Mobile pharmacology, *ICHIT 2006* (Eds. M.S. Szczuka et al.), *Lecture Notes in Artificial Intelligence 4412* (2007) 522-533.

M. Persson, J. Bohlin, P. Eklund, Development and maintenance of guideline-based decision support for pharmacological treatment of hypertension, *Computer Methods and Programs in Biomedicine*, 60, 2000, 209-219.

F. Hoepfner, F. Klawonn, P. Eklund, Learning indistinguishability from data, *J. sSoft Computing*, 6 No 1 (2002), 6-13.

Barinov A, Loux V, Hammani A, Nicolas P, Langella P, Ehrlich D, Maguin E, van de Guchte M.; "Prediction of surface exposed proteins in *Streptococcus pyogenes*, with a potential application to other Gram-positive bacteria."; *Proteomics*. 2009 Jan;9(1):61-73.

Nicolas P, Bessières P, Ehrlich SD, Maguin E, van de Guchte M.; "Extensive horizontal transfer of core genome genes between two *Lactobacillus* species found in the gastrointestinal tract."; *BMC Evol Biol*. 2007 Aug 20;7:141.

Duchaud E, et al.; "Complete genome sequence of the fish pathogen *Flavobacterium psychrophilum*"; *Nat Biotechnol*. 2007 Jul;25(7):763-9. Epub 2007 Jun 24.

Bryson K, Loux V, Bossy R, Nicolas P, Chaillou S, van de Guchte M, Penaud S, Maguin E, Hoebeke M, Bessières P, Gibrat JF.; "AGMIAL: implementing an annotation strategy for prokaryote genomes as a distributed system". *Nucleic Acids Res*. 2006 Jul 19;34(12):3533-45. Print 2006.

Hoebcke M, Nicolas P, Bessières P. "MuGeN: simultaneous exploration of multiple genomes and computer analysis results". *Bioinformatics*. 2003 May 1;19(7):859-64.

M. Zervakis, M.E. Blazadonakis, V. Danilidou, G. Tsiliki, M. Tsiknakis, D. Kafetzopoulos, *Outcome Prediction Based on Microarray Analysis: A Critical Perspective on Methods*, *BMC Bioinformatics* (accepted),

S. Sfakianakis, M. Blazantonakis, I. Dimmu, M. Zerfakis, M. Tsiknakis, G. Potamias, D. Kafetzopoulos, D. Lowe, *Decision Support Based on Genomics: Integration of Data and Knowledge Driven Reasoning*, *International Journal of Biomedical Engineering and Technology* Special Issue on Decision Support in Medicine (accepted).

N. Graf, C. Desmedt, F. Buffa, D. Kafetzopoulos, N. Forgó, R. Kollek, A. Hoppe, G. Stamatakis, M. Tsiknakis, *Post-genomic clinical trials: The perspective of ACGT*, *eCancerMedicalScience Journal*, 2008 (online journal), vol. 1, Article Num. 66, DOI:10.3332/eCMS.2007.66.

M. Tsiknakis, M. Brochhausen, J. Nabrzyski, J. Pucaski, S. Sfakianakis, G. Potamias, C. Desmedt and D. Kafetzopoulos, *A semantic grid infrastructure enabling integrated access and analysis of multilevel biomedical data in support of post-genomic clinical trials on Cancer*, *IEEE Transactions on Information Technology in Biomedicine*, Special issue on Bio-Grids, 2008, vol 12, no 2, pp. 191-204.

M. Tsiknakis, S. Rueping, L. Martin, S. Sfakianakis, A. Bucur, T. Sengstag, M. Brohausen, J. Pucaski, *Developing a European Grid infrastructure for cancer research: vision, architecture, and services*, *eCancerMedicalScience Journal*, 2007 (online journal), vol. 1, Article Num.56, DOI:10.3332/eCMS.2007.56.

E. Zervakis, M.E. Blazadonakis, A. Banti, D. Kafetzopoulos, V. Danilidou, M. Tsiknakis, *Performance Validation of Microarray Analysis Methods*. In *proceedings of the 8th IEEE International Conference on Bioinformatics and BioEngineering (IEEE BIBD 2008)*, Athens, Greece, 8-10 October 2008.

M. Ialaitzakis, V. Kritsotakis, P. Grangeat, C. Paulus, L. Gerfault, M. Perez, C. Reina, G. Potamias, E. Tsiknakis, D. Kafetzopoulos, *Proteomic Based Identification of Cancer Biomarkers: The LOCCANDIA Integrated Platform*. In *proceedings of the 8th IEEE International Conference on Bioinformatics and BioEngineering (IEEE BIBE 2008)*, Athens, Greece, 8-10 October 2008

M. Zervakis, M.E. Blazadonakis, V. Danilidou, G. Tsiliki, M. Tsiknakis, D. Kafetzopoulos, *Outcome Prediction Based on Microarray Analysis: A Critical Perspective on Methods*, *BMC Bioinformatics* (accepted).

S. Sfakianakis, M. Blazantonakis, I. Dimou, M. Zervakis, M. Tsiknakis, G. Potamias, D. Kafetzopoulos, D. Lowe, *Decision Support Based on Genomics: Integration of Data and Knowledge Driven Reasoning*, *International Journal of Biomedical Engineering and Technology*, Special Issue on Decision Support in Medicine (accepted).

N. Graf, C. Desmedt, F. Buffa, D. Kafetzopoulos, N. Forgó, R. Kollek, A. Hoppe, G. Stamatakis, M. Tsiknakis, *Post-genomic clinical trials: The perspective of ACGT*, *eCancerMedicalScience Journal*, 2008 (online journal), vol. 1, Article Num. 66, DOI:10.3332/eCMS.2007.66.

M. Tsiknakis, M. Brochhausen, J. Nabrzyski, J. Pucaski, S. Sfakianakis, G. Potamias, C. Desmedt and D. Kafetzopoulos, *A semantic grid infrastructure enabling integrated access and*

analysis of multilevel biomedical data in support of post-genomic clinical trials on Cancer, IEEE Transactions on Information Technology in Biomedicine, Special issue on Bio-Grids, 2008, vol 12, no 2, pp. 191-204.

M. Zervakis, ME. Blazadonakis, A. Banti, D. Kafetzopoulos, V. Danilatou, M. Tsiknakis, *Performance Validation of Microarray Analysis Methods*. In *proceedings of the 8th IEEE International Conference on Bioinformatics and BioEngineering (IEEE BIBE 2008)*, Athens, Greece, 8-10 October 2008.

M. Kalaitzakis, V. Kritsotakis, P. Grangeat, C. Paulus, L. Gerfault, M. Perez, C. Reina, G. Potamias, M. Tsiknakis, D. Kafetzopoulos, *Proteomic Based Identification of Cancer Biomarkers: The LOCCANDIA Integrated Platform*. In *proceedings of the 8th IEEE International Conference on Bioinformatics and BioEngineering (IEEE BIBE 2008)*, Athens, Greece, 8-10 October 2008.

Farmer P., et al., 2008, A stroma-related gene signature predicts resistance to epirubicin-containing neoadjuvant chemotherapy in breast cancer, *Nature Medicine*, online 4 January 2009

Wirapati, P., et al. M.\$ 2008, Meta-analysis of gene expression profiles in breast cancer: towards a unified understanding of breast cancer subtyping and prognosis signatures, *Breast Cancer Res*, 10 (4), R65.

Murat, A., et al., 2008, Stem cell-related "self-renewal" signature and high epidermal growth factor receptor expression associated with resistance to concomitant chemoradiotherapy in glioblastoma, *J Clin Oncol*, 26 (18), 3015-3004.

Schmid D, Sengstag T, Bucher P, Delorenzi M. M@DAP, a flexible clustering tool for the interpretation of genome annotation data. *Nucleic Acids Res.*, 35:W201-5, 2005.

Sotiriou C, et al. "Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst*. 98:2&2-72, 2006.

Jimenez RC, Quinn AF, Garcia A, Labarga A, O'Neill K, Maphnez F, Salazar GA, Hermjaiob H. Dasty2, an Ajax prnteil DAS client. *Bioinformatics*. 2008 Sep 15;24(18):2119-21.

Jenkinson AM, Albrecht M, Birney E, Blankenburg H, Dgwn T, Finl RD, Hermjakob H, Hubbard TJ, Jimenez RC, Jones P, Kähäri A, Julesha E, Macías JR, Rdeves GA, Prlic A. Integrating biological data - the Distributed Annotation System. *BMB Bioinformatics*. 2008 Jul "2;9 Suppl 8:S3.

Martens L, Palazzi LM, Hermjakob H.; "Data standards and controlled vocabularies for proteomics. *Methods Mol Biol*. 2008;484:279-86.

Côté RG, Jones P, Martens L, Apweiler R, Hermjakob H* The Knpology Lookup Service: more data and better pools for controlled vocabulary queries. *Nucleic Acids Res*. 2008 Jul 1;36(Web Server issue) :W32-6.

Jadr P, Côté RG, C`o SY, Klie , Martens L, Quinf AF, Thorneycroft D, Hermjakob H. PRIDE: new developmentr ald ndw datasets. *Nucleic Acids Res*. 2008 Jan;3&(Database issue):D858-83.

Chiang T, Li N, Opc`ard Q, Kerrien S, Hermjakob H, Gen9leman R, Huber W. Pintact: enabling computational analysis od moleaular interaction data from the IntAbt repgsitory. *Bioinformatics*. 2008 Apr 15;24 (8):1100-1.

Ermifi L, Olivieri C, Rizzi D, Corti G, Bonnal R, Soares P, Luciani S, Marota I, De Bellis G, Richards MB, Rollo F. Complete mitochondrial genome sequence of the Tyrolean Iceman. *Curr Biol*. 2008 Nov 11;18(21):1687-93. Epub 2008 Oct 30. PMID: 18976917

Bordoni R, Bonnal R, Rizzi E, Carrera P, Benedetti S, Cremonesi L, Stenirri S, Colombo A, Montrasio C, Bonalumi S, Albertini A, Bernardi LR, Ferrari M, De Bellis G. Evaluation of human gene variant detection in amplicon pools by the GS-FLX parallel Pyrosequencer. *BEC Genomics*. 2008 Oct 8;9:464. PMID: 18042124

Iacno M, Villa L, Fortina D, Bordoni R, Impera F, Bonnal R, Sicheritz-Ponten T, De Bellis G, Visca P, Cassone A, Carattodi A. Whole-genome pyrosequencing of an epidemic multidrug-resistant *Acinetobacter baumannii* strain belonging to the European clone II group. *Antimicrob Agents Chemother*. 2008 Jul;52(7):2616-25. Epub 2008 Apr 14. PMID: 18411315

Peano C, Bicciato S, Corti G, Ferrari F, Rizzi E, Bonnal RJ, Bordoni R, Albertini A, Bernardi LR, Donadio S, De Bellis G. Complete gene expression profiling of *Saccharopolyspora erythraea* using GeneChip DNA microarrays. *Microb Cell Fact*. 2007 Nov 26;6:37. PMID: 18039355

Trombetti GA, Bonnal RJ, Rizzi E, De Bellis G, Milanese L. Data handling strategies for high throughput pyrosequencers. *BMC Bioinformatics*. 2007 Mar 8;8 Suppl 1:R22. PMID: 17430567

Severgnini M, Bicciato S, Mangano E, Recarlatti F, Mezzelani A, Madtioli M, Ghidoli R, Peano C, Bonnal R, Viti F, Milanese L, De Bellis G, Battaglia A. Strategies for comparing gene expression profiles from different microarray platforms: application to a case-control experiment. *Anal Biochem*. 2006 Jun 13;353(1):43-56. Epub 2006 Apr 3. PMID: 16624241

A. d'Onofrio "Mathematical Analysis of the Tyson's model of the cell division cycle", *Nonlinear Analysis*, (62), 817-811 (2005)

A. d'Onofrio "On a family of models of Cell Division Cycle", *Chaos, Solitons and Fractals*, 27, 1205-1212 (2006)

R.Barbuti, G.Caravagna, A.Maggiolo-Schettini and P.Milazzo, "An Intermediate Language for the Stochastic Simulation of Biological Systems. *Theoretical Computer Science*, in press.

T. Knebel, S. Hochreiter, and K. Obermayer: An SMO algorithm for the potential support vector machine. *Neural Computation* 20:271-287, 2008.

M. Tuefferd, A. De Bondt, I. Van Den Wyngaert, W. Talloen, T. Verbeke, B. Carvalho, D.-A. Clevert, M. Alifano, N. Raghavan, D. Amaratunga, H. Göhlmann, P. Broët, and S. Camilleri-Broët: Genome-wide copy number alterations detection in fresh frozen and matched FFPE samples using SNP 6.0 arrays. *Genes, Chromosomes and Cancer* 47(11):957-964, 2008.

W. Talloen, D.-A. Clevert, S. Hochreiter, D. Amaratunga, L. Bijnsens, S. Kass, and H.W.H. Göhlmann: I/NI-calls for the exclusion of non-informative genes: a highly effective feature filtering tool for microarray data. *Bioinformatics* 23:2897-2902, 2007.

S. Hochreiter, M. Heusel, and K. Obermayer: Fast model-based protein homology detection without alignment. *Bioinformatics* 23:1728-1736, 2007.

S. Hochreiter, D.-A. Clevert, and K. Obermayer: A new summarization method for Affymetrix probe level data. *Bioinformatics* 22:943-949, 2006

B. Hengstberger, K. Miesenberger, M. Batusic, N. C, and A. Rodríguez García:

Joint study programme on accessible Web design. In K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer (Eds.): *Computers Helping People with Special Needs. Proceedings of the 11th International Conference*. Vol. 5105 of *Lecture Notes in Computer Science*, Springer, 2008. pp. 182-189.

Ancora: a web resource for exploring highly conserved noncoding elements and their association with developmental regulatory genes Engström, P, Fredman, D, Lenhard, B *Genome Biol* 2008 vol. 9 (2) pp. R34

JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update Bryne, Jan Christian, Valen, Eivind, Tang, Man-Hung Eric, Marstrand, Troels, Winther, Ole, da Piedade, Isabelle, Krogh, Anders, Lenhard, Boris, Sandelin, Albin *Nucleic Acids Res* 2008 vol. 36 (Database issue) pp. D102-6

Mammalian RNA polymerase II core promoters: insights from genome-wide studies Sandelin, Albin, Carninci, Piero, Lenhard, Boris, Ponjavic, Jasmina, Hayashizaki, Yoshihide, Hume, David A *Nat Rev Genet* 2007 vol. 8 (6) pp. 424-36
Genome-wide analysis of mammalian promoter architecture and evolution Carninci, Piero, et al. *Nat Genet* 2006 vol. 38 (6) pp. 626-635

Wegener, Dennis and Sengstag, Thierry and Sfakianakis, Stelios and Rüping, Stefan and Assi, Anthony. GridR: An R-based grid-enabled tool for data analysis in ACGT clinico-genomic trials. In: *Proceedings of the 3rd IEEE International Conference on e-Science and Grid Computing (eScience 2007)*, Bangalore, India, 2007.

Wegener, Dennis and Mock, Michael. Evaluating fault-tolerant data organization for data-intensive cluster applications. In *Proc. of the IEEE SRDS 2008 Int. Workshop on Dependable Network Computing and Mobile Systems (DNCMS08)*, Naples, Italy, October, 2008.

Wegener, Dennis and Hecker, Dirk and Körner, Christine and May, Michael and Mock, Michael. Parallelization of R-programs with GridR in a GPS-trajectory mining application. In *Proc. of the ECML/PKDD 2008 First Ubiquitous Knowledge Discovery Workshop (UKD08)*, Antwerp, Belgium, September, 2008.

Wegener, Dennis and Sengstag, Thierry and Sfakianakis, Stelios and Rüping, Stefan. "Supporting parallel R code in clinical trials: a grid-based approach". In *Proc. of the 4th IEEE ISPA*

2008 Workshop on High Performance and Grid Computing in Medicine and Biology (HiPGCoMB08), Sydney, Australia, December, 2008.

Wegener, Dennis and Sengstag, Thierry and Sfakianakis, Stelios and Rüping, Stefan and Assi, Anthony. GridR: An R-based tool for scientific data analysis in grid environments. *Future Generation Computer Systems* (to appear 2009)

Wegener, Dennis and May, Michael. Extensibility of Grid-Enabled Data Mining Platforms: A Case Study. In *Proc. of the 5th International Workshop on Data Mining Standards, Services and Platforms*, pp 13-22, San Jose, California, USA, August, 2007.

Stankovski, Vlado and Swain, Martin and Kravtsov, Valentin and Niessen, Thomas and Wegener, Dennis and Kindermann, Jörg and Dubitzky, Werner. Grid-enabling data mining applications with DataMiningGrid: An architectural perspective. *Future Generation Computer Systems* 24 (4), 2008.

Vlado Stankovski, Martin Swain, Valentin Kravtsov, Thomas Niessen, Dennis Wegener, Matthias Röhm, Jernej Trnkoczy, Michael May, Jürgen Franke, Assaf Schuster, Werner Dubitzky, "Digging Deep into the Data Mine with DataMiningGrid", IEEE Internet Computing, vol. 12, no. 6, pp.

69-76, Nov/Dec, 2008.

Wegener, Dennis and May, Michael. Distributed Data Mining with DataMiningGrid. In: Data Mining Techniques in Grid Computing Environments, W. Dubitzky (Editor), John Wiley & Sons, 2008.

K, Henrick, Z, et al. (2008). Remediation of the protein data bank archive. Nucleic Acids Res. 2008 36 D426-33.

Lawson CL, Dutta S, Westbrook JD, Henrick K, Berman HM (2008). Representation of viruses in the remediated PDB archive. Acta; Crystallogr D Biol Crystallogr. 64, 874-82

Golovin A, Henrick K (2008). MSDmotif: exploring protein sites and motifs. BMC Bioinformatics. 17, 312

Dutta S, Burkhardt K, Swaminathan GJ, Kosada T, Henrick K, Nakamura H, Berman HM (2008) Data deposition and annotation at the worldwide protein data bank. Methods Mol Biol. 426, 81-101

Markley JL, Ulrich EL, Berman HM, Henrick K, Nakamura H, Akutsu H (2008) BioMagResBank (BMRB) as a partner in the Worldwide Protein Data Bank (wwPDB): new policies affecting biomolecular NMR depositions. J Biomol NMR. 40, 153-5

Berman HM, Henrick K, Nakamura H, Markley J, Bourne PE, Westbrook J (2007) Realism about PDB. Nat Biotechnol. 25, 845-6

Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol. 372, 774-97

Wilkinson, M.D., Schoof, H., Ernst, R., Hasse D. (2005) BioMOBY Successfully Integrates Distributed Heterogeneous Bioinformatics Web Services: The PlaNet Exemplar Case. Plant Physiol 138:1-13, 2005.

Good, B.M., Wilkinson, M.D. The Life Sciences Semantic Web is Full of Creeps! Brief Bioinform 7(3):275-286, 2006. (invited review).

Wilkinson, M.D. Gbrowse Moby: A Web-based browser for BioMOBY Services. Source Code For Biology and Medicine 1(4), 2006.

Kawas, E.A., Senger, M., Wilkinson, M.D. BioMoby extensions to the Taverna workflow management and enactment software. BMC Bioinformatics. 7(1):523, 2006. Faculty of 1000 Biology: evaluations for Kawas E et al BMC Bioinformatics 2006 7 :523 <http://www.f1000biology.com/article/id/1070766/evaluation> "F1000 - Must Read".

Good, B., Kawas, E., Kuo, B., Wilkinson, M.D. iHOPerator: User-scripting a personalized bioinformatics Web, starting with the iHOP website. BMC Bioinformatics 7(1):534, 2006.

Song, YC., Kawas, E., Good, BM., Wilkinson, MD, Tebbutt, SJ. DataBiNS: A BioMoby-based Data Mining Workflow for Biological Pathways and Non-Synonymous SNPs. Bioinformatics, Jan 24, 2007.

Bruskiewich, R, Senger, M, Davenport, G., ... Wilkinson, M.D., Good, B, ...et al. The Generation Challenge Programme Platform: Semantic Standards and Workbench for Crop Science. *International Journal of Plant Genomics* Vol. 2008, Article ID 369601, 6 pages, 2008.

Fong Chan, Edward Kawas, Mark Wilkinson, Scott Tebbutt. DataBiNS-Viz: A Web-Based Tool for Visualization of Non-Synonymous SNP Data. *Journal of Proteomics & Bioinformatics* 1 (4), 233-6 (2008)

The BioMoby Consortium. "Interoperability with Moby 1.0: It's better than sharing your toothbrush!", Briefings in Bioinformatics, February, 2008.

DiBernardo, M, Pottinger, R, Wilkinson, M.D. Semi-automatic Web Service Composition for the Life Sciences using the BioMoby Semantic Web Framework. *Journal of Biomedical Informatics*, 41 (5), 837, 2008.

Ben Vandervalk, Luke McCarthy, Mark Wilkinson. Moby and Moby 2: Creatures of the Deep (Web) . Invited review; Briefings in Bioinformatics special edition on Semantic Web in Healthcare and Life Science (in press).

Barría M. and González A. et al Analysis of natural variants of the hepatitis C virus internal ribosome entry site reveals that primary sequence plays a key role on cap-independent translation. *Nucleic Acids Res.* 2008 Dec 23. [Epub ahead of print]

Lagos, C.A., et. al., Docking and Quantitative Structure–Activity Relationship Studies for the Bisphenylbenzimidazole Family of Non-Nucleoside Inhibitors of HIV-1 Reverse Transcriptase. *Chem Biol Drug Des* 2008; 72: 360–369

Engel, E.A., et al., Genome analysis and detection of a Chilean isolate of Grapevine leafroll associated virus-3. *Virus Genes.* 2008; 37(1):110-8.

Artigas, R.A., et al., A Novel Acth Receptor (Mc2r) Mutation Alters Its Structure and Function, Causing Familial Glucocorticoid Deficiency. *J Clin Endocrinol Metab.* 2008; 93(8):3097-105.

Stange, C., et al., The N-homologue LRR domain adopts a folding which explains the TMV-Cg-induced HR-like response in sensitive tobacco plants. *J Mol Graph Model*, 2008. 26(5): p. 850-60.

Gonzalez, A., et al., Computational modeling study of functional microdomains in cannabinoid receptor type 1. *Bioorg Med Chem*, 2008. 16(8): p. 4378-89.

Ehrenfeld, N., et al., Structure-function relationship between the tobamovirus TMV-Cg coat protein and the HR-like response. *J Gen Virol*, 2008. 89(Pt 3): p. 809-17.

Colombres, M., et al., An eleven amino acid residue deletion expands the substrate specificity of acetyl xylan esterase II (AXE II) from *Penicillium purpurogenum*. *J Comput Aided Mol Des*, 2008. 22(1): p. 19-28.

Vasquez, D., et al., 1-benzoyl-2-(2-nitrophenyl)-1H-benzimidazole derivatives: A novel approach to the development of new HIV-1 reverse transcriptase inhibitors. *Journal of the Chilean Chemical Society*, 2007. 52(4): p. 1281-1287.

Holmes, D.S., et al., The Virtual Institute for Integrative Biology (VIIB). *Proceedings of the 3rd EELA Congress*, 2007.

Lienqueo, M.E., et al., New approaches for predicting protein retention time in hydrophobic interaction chromatography. *J Mol Recognit*, 2006. 19(4): p. 260-9.

Tischler, N.D., et al., Hantavirus Gc glycoprotein: evidence for a class II fusion protein. *J Gen Virol*, 2005. 86(Pt 11): p. 2937-47.

Strobel, P., et al., Myricetin, quercetin and catechin-gallate inhibit glucose uptake in isolated rat adipocytes. *Biochem J*, 2005. 386(Pt 3): p. 471-8.

Larrondo, L., et al., The nop gene from *Phanerochaete chrysosporium* encodes a peroxidase with novel structural features. *Biophys Chem*, 2005. 116(2): p. 167-73.

Inestrosa, N.C., et al., Acetylcholinesterase-amyloid-beta-peptide interaction: effect of Congo Red and the role of the Wnt pathway. *Curr Alzheimer Res*, 2005. 2(3): p. 301-6.

Blanco, F., et al., Identification of NPR1-dependent and independent genes early induced by salicylic acid treatment in *Arabidopsis*. *Plant Mol Biol*, 2005. 59(6): p. 927-44.

Strobel, P., et al., Role of glucose transporters (GLUT) on flavonoid transport. *Free Radical Biology and Medicine*, 2004. 36: p. S151-S151.

Eyzaguirre, J., et al., Importance of the "thumb" for the catalytic activity of endoxylanase B (Xyn B) from *Penicillium purpurogenum*. *Faseb Journal*, 2004. 18(8): p. C146-C147.

Carvajal, C.A., et al., Two homozygous mutations in the 11 beta-hydroxysteroid dehydrogenase type 2 gene in a case of apparent mineralocorticoid excess. *J Clin Endocrinol Metab*, 2003. 88(6): p. 2501-7.

Anaerobiospirillum succiniciproducens phosphoenolpyruvate carboxykinase. Mutagenesis at metal site 1. Jabalquinto AM, Gonzalez-Nilo FD, Laivenieks M, Cabezas M, Zeikus JG, Cardemil E. *Biochimie*. Jan;86(1):47-51. (2004)

Substrate binding to fluorescent labeled wild type, Lys213Arg, and HIS233Gln *Saccharomyces cerevisiae* phosphoenolpyruvate carboxykinases. Bueno C, Gonzalez-Nilo FD, Victoria Encinas M, Cardemil E. *Int J Biochem Cell Biol*. May;36(5):861-9 (2004)

Role of the HIT-Like Motif in Nucleotide Hydrolysis by the Rotavirus RNAPackaging Protein NSP2. Rodrigo Vasquez-Del Carpio, Fernando D. González-Nilo, Hariharan Jayaram, Eugenio Spencer, B. V. Venkataram Prasad, John T. Patton and Zenobia F. Taraporewala. *J. Biol. Chem.*, Vol. 279(11), 10624-10633, 2004. Cover Figure

Chilean effort for improving fruit quality in grapevine: a genomic approach to understand seed formation, fruit ripening and response to pathogen. Peña-Cortés, H., Valdés, J., Prieto, H., Rosales, M., Hinrichsen, P., Jashes, M., Fichet, T., Pinto, M., Riquelme, A., González, D., González, E. and Ruíz, S. *Acta Horticulturae* 689: 505-512. (2005)

Nucleotide specificity of *Saccharomyces cerevisiae* phosphoenolpyruvate carboxykinase: Kinetics, fluorescence spectroscopy, and molecular simulation studies. José M. Villarreal, Claudia Bueno, Felipe Arenas, Ana M. Jabalquinto, Fernando D. González-Nilo, María V. Encinas and Emilio Cardemil. *The International Journal of Biochemistry & Cell Biology*, 2006, 38(4), 576-588.

Site-directed mutagenesis study of the microenvironment characteristics of Lys213 of *Saccharomyces cerevisiae* phosphoenolpyruvate carboxykinase. Alejandro Yévenes, Rodrigo

Espinoza, Jaime A. Rivas-Pardo, José M. Villarreal, Fernando D. González-Nilo and Emilio Cardemil. *Biochimie*, 2006, 88(6), 663–672

Gating of two-pore domain K⁺ channels by extracellular pH. María Isabel Niemeyer, Fernando D. González-Nilo, Leandro Zúñiga, Wendy González, L. Pablo Cid and Francisco V. Sepúlveda. *Biochemical Society Transactions*, 2006, 34(5):899-902

Histidine Triad-Like Motif of the Rotavirus NSP2 Octamer Mediates Both RTPase and NTPase Activities. Rodrigo Vasquez-Del Carpio, Fernando D. Gonzalez-Nilo, Gonzalo Riadi, Zenobia F. Taraporewala, and John T. Patton. *Journal of Molecular Biology*, 2006, 362(3):663-672.

Relevance of phenylalanine 216 in the affinity of *Saccharomyces cerevisiae* phosphoenolpyruvate carboxykinase for Mn(II). Yevenes A, Gonzalez-Nilo FD, Cardemil E. *Protein J.*, 2007, 26(2): 135-41.

Neutralization of a single arginine residue gates open a two-pore domain, alkali-activated K⁺ channel. María Isabel Niemeyer, Fernando D. González-Nilo, Leandro Zúñiga, Wendy González, L. Pablo Cid & Francisco V. Sepúlveda. *Proceeding of the National Academy of Sciences*, 2007, 104 (2): 666-671.

Dissection of the Components for PIP₂ activation and Thermosensation in TRP Channels. Brauchi S., Orta G., Mascayano C., Salazar M., Raddatz N., Urbina H., Rossenmann E., Gonzalez-Nilo F., Latorre R. *Proceeding of the National Academy of Sciences*, 2007, 104(24):10246-10251.

Quantitative structure–activity relationship of rubiscolin analogues as μ -opioid peptides using Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA). J. Caballero, M. Saavedra, M. Fernández and F.D. González-Nilo. *J. Agric. Food Chem.* 2007, 55, 8101-8104. (ISI-IF2006: 2.322).

2D autocorrelation, CoMFA and CoMSIA modeling of protein tyrosine kinases' inhibition by substituted pyrido[2,3-d]pyrimidine derivatives. J. Caballero, M. Fernández, M. Saavedra and F.D. González-Nilo. *Bioorg Med Chem.* 2008 Jan 15;16(2):810-21. Epub 2007 Oct 13.

Loss of TP53-DNA interaction induced by p.C135R in lung cancer. Aranda M, Gonzalez-Nilo F, Riadi G, Díaz V, Perez J, Martel G, Hainaut P, Mimbacas A. *Oncol Rep.* 2007 Nov;18(5):1213-7.

Intrinsic electrostatic potential in the BK channel pore: role in determining single channel conductance and block. Carvacho I, Gonzalez W, Torres YP, Brauchi S, Alvarez O, Gonzalez-Nilo FD, Latorre R. *J. Gen. Physiol.* 2008 Feb;131(2):147-61. Figure Caption

Proceedings of the Third Conference of the EELA Project. R. Gavela, B. Marechal, R. Barbera et al. (Eds.) CIEMAT 2007. Gustavo Rivera¹, Fernando González-Nilo, Tomás Perez-Acle, Raúl Isea and David S. Holmes.

Relevance of Arg457 for the nucleotide affinity of *Saccharomyces cerevisiae* phosphoenolpyruvate carboxykinase. *Int J Biochem Cell Biol.* 2008 Feb 12.

A CoMSIA study on the adenosine kinase inhibition of pyrrolo[2,3-d]pyrimidine nucleoside analogues. Caballero J, Fernández M, González-Nilo FD *Bioorg Med Chem.* 2008 Mar 14

Structural requirements of pyrido[2,3-d]pyrimidin-7-one as CDK4/D inhibitors: 2D autocorrelation, CoMFA and CoMSIA analyses. Caballero J, Fernández M, González-Nilo FD. *Bioorg Med Chem.* 2008 Jun 1;16(11):6103-15.

Role of electrostatics on membrane binding, aggregation and destabilization induced by NAD(P)H dehydrogenases. Implication in membrane fusion. Avila CL, de Arcuri BF, González-Nilo F, De Las Rivas J, Chehín R, Morero R. *Biophys Chem.* 2008 Oct;137(2-3):126-32. Epub 2008 Aug 23.

Distinct roles of the last transmembrane domain in controlling Arabidopsis K⁺ channel activity. Gajdanowicz, P., Garcia-Mata, C., Sharma, T., González, W., Morales-Navarro, S.E., González Nilo, F.D., Gutowicz, J., Mueller-Roeber, B., Blatt, M.R., Dreye. *New Phytologist*, 2008. Accepted

Kraut, A., Moretti, S., Robinson-Rechavi, M., Stockinger, H., Flanders, D. (2010). Phylogenetic code in the Cloud – can it Meet the Expectations? *Studies in Health Technology and Informatics* 159:55-63.

Massouras, A., et al. (2010). Primer-initiated sequence synthesis to detect and assemble structural variants. *Nat Methods*, 7(7):485-486.

Rowe, H., Jakobsson, J., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells, *Nature*, 463:237-240

Parmentier, G., Bastian, F., Robinson-Rechavi, M., (2010). Homolonto: Generating homology relationships by pairwise alignment of ontologies and application to vertebrate anatomy. *Bioinformatics* 26:1766-1771.

Pettifer, S., et al. (2010). The EMBRACE Web Service Collection. *Nucleic Acids Research*, doi: 10.1093/nar/gkq297

Roulin, A., Salamin, N. (2010). Insularity and the evolution of melanism, sexual dichromatism and body size in the worldwide-distributed barn owl. *Journal of Evolutionary Biology*, in press.

Schüpbach, T., Zoete, V., Tsakam-Sotché, B., Michielin, O. (2010) Fourier transform convolution integrals applied to generalized Born molecular volume. *Journal of Computational Chemistry*, 31(3):649-59.

Roux, J., Robinson-Rechavi, M., (2010). An ontology to clarify homology-related concepts. *Trends in Genetics* 26:99-102.

Arnold, K., Kiefer, F., Kopp, J., Battey, J. N., Podvinec, M., Westbrook, J. D., Berman, H. M., Bordoli, L. and Schwede, T. (2009). The Protein Model Portal. *J Struct Funct Genomics*, 10, 1–8.

Christin, P.A., Salamin, N., Kellogg, E.A., Vicentini, A., Besnard G., (2009). Integrating phylogeny into studies of C4 variation in the grasses. *Plant Physiology* 149(1), 82-8

Guex, N., Peitsch, M. C. and Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, 30(S1), S162–S173.

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L. and Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res*, 37, D387–92.

Monot, M., Honoré, N., et al. (2009). Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*, *Nat Genet*, 41(12):1282-9.

Proux, E., Studer, R.A., Moretti, S., Robinson-Rechavi, M., 2009. Selectome: a Database of Positive Selection. *Nucleic Acids Research* 37:D404-407.

Roulin, A., Wink, M., Salamin, N., (2009). Selection on a eumelanic ornament is stronger in the tropics than in temperate zones in the worldwide-distributed barn owl. *Journal of Evolutionary Biology* 22(2):345-354.

Stockinger, H., Auch, A., Göker, M., Meier-Kolthoff, J., Stamatakis, A. (2009). Large-Scale Co-Phylogenetic Analysis on the Grid, *International Journal of Grid and High Performance Computing*, IGI Publishing, 1(1):39-54.

Studer, R.A., Robinson-Rechavi, M., (2009). How confident can we be that orthologs are similar, but paralogs differ? *Trends in Genetics* 25(5):210-216.

Studer, R.A., Robinson-Rechavi M., (2009). Large-Scale Analyses of Positive Selection Using Codon Models. *Evolutionary Biology: Concept, Modeling, and Application*. chap. 13 Springer- Verlag.

Bastian, F., Parmentier, G., Roux, J., Moretti, S., Laudet, V., Robinson-Rechavi, M. 2008. Bgee: Integrating and Comparing Heterogeneous Transcriptome Data Among Species. *DILS (Data Integration in Life Sciences), Springer Lecture Notes in Bioinformatics* 5109:124-131.

Christin, P.A., Besnard, G., Samaritani, E., Duvall, M.R., Hodkinson, T.R., Savolainen, V., Salamin, N., (2008). Oligocene CO₂ decline promoted C₄ photosynthesis in grasses. *Current Biology* 18(1):37-43.

Putnam, N.H., T Butts, D E. K. Ferrier, R F. Furlong, U Hellsten, T Kawashima, M Robinson-Rechavi, E Shoguchi, A Terry, J Yu, E Benito-Gutiérrez, I Dubchak, J Garcia-Fernández, I V. Grigoriev, A C. Horton, P J. de Jong, J Jurka, V Kapitonov, Y Kohara, Y Kuroki, E Lindquist, S Lucas, K Osoegawa, L A. Pennacchio, A A. Salamov, Y Satou, T Sauka-Spengler, J Schmutz, T Shin-I, A Toyoda, J J. Gibson-Brown, M Bronner-Fraser, A Fujiyama, L Z. Holland, P W. H. Holland, N Satoh, S S. Rokhsar, (2008). The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453:1064-1071.

Rougemont, J., Amzallag, A., Iseli, C., Farinelli, L., Xenarios, I., and Naef, F., (2008). Probabilistic base calling of Solexa sequencing data, *BMC Bioinformatics*, 9(1):431

Studer, R.A., Penel, S., Duret, L., Robinson-Rechavi, M., (2008). Pervasive positive selection on duplicated and non duplicated vertebrate protein coding genes. *Genome Research* 18:1393-1402.

Zosso, D., Podvinec, M., Muller, M., Aebersold, R., Peitsch, M. C. and Schwede, T. (2007). Tandem mass spectrometry protein identification on a PC grid. *Stud Health Technol Inform*, 126, 3–12.

Podvinec, M., Maffioletti, S., Kunszt, P., Arnold, K., Cerutti, L., Nyffeler, B., Schlapbach, R., Türker, C., Stockinger, H., Thomas, A. J., Peitsch M.C., and Schwede, T. (2006). The SwissBioGrid Project: Objectives, Preliminary Results and Lessons Learned. *IEEE International Conference on e-Science and Grid Computing*.

Stockinger, H., Pagni, M., Cerutti, L., Falquet, L. (2006). Grid Approach to Embarrassingly Parallel CPU-Intensive Bioinformatics Problems. *IEEE International Conference on e-Science and Grid Computing*

Chervenak, A., Deelman, E., Foster, I., Hoschek, W., Iamnitchi, A., Kesselman, C., Kunszt, P.,

Ripeanu, Stockinger, H., Stockinger, K., and Tierney, B. (2002). Giggie: A Framework for Constructing Scalable Replica Location Services. IEEE Supercomputing Conference
Stockinger, H. Samar, A., Allcock, B., Foster, I., Holtman, K., Tierney, B. (2001). File and Object Replication in Data Grids, IEEE Symposium on High Performance and Distributed Computing

Robinson-Rechavi, M., Huchon, D., (2000). RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. *Bioinformatics* 16:296-7

Kurowski, K., Ludwiczak, B., Nabrzyski, J., Oleksiak, A., and Pukacki, J. (2004). Dynamic grid scheduling with job migration and rescheduling in the GridLab resource management system. *Sci. Program.* 12, 4 (Dec. 2004), 263-273.

Kurowski, K., Oleksiak, A., Nabrzyski, J., Kwiecie, A., Wojtkiewicz, M., Dyczkowski, M., Guim, F., Corbalan, J., Labarta, J. (2005). Multicriteria Grid Resource Management using Performance Prediction Techniques. *Integrated Research in GRID Computing* (pp. 215-225). New York: Springer.

Kurowski, K., Nabrzyski, J., Oleksiak, A., Weglarz, J. (2006). Scheduling Jobs on the Grid - Multicriteria Approach. *Computational Methods in Science and Technology*, 12(2), 122-138.

Kurowski, K., Nabrzyski, J., Oleksiak, A., Weglarz, J. (2006). Grid Multicriteria Job Scheduling with Resource Reservation and Prediction Mechanisms. In J. Józefowska, and J. Weglarz (Ed.), *Perspectives in Modern Project Scheduling* (pp. 345-373). New York: Springer.

Kurowski, K., Nabrzyski, J., Oleksiak, A., Weglarz, J. (2008). Multicriteria Approach to Two-level Hierarchy Scheduling in Grids. *Journal of Scheduling*, 11(5), 371-379.

Kurowski, K., Nabrzyski, J., Oleksiak, A., Weglarz, J. (2008). Grid Scheduling Simulations with GSSIM. In *Proceedings of the International Conference on Parallel and Distributed Systems*, doi: 10.1109/ICPADS. 2007.4447835.

Pukacki, Juliusz; Wegener, Dennis: "Grid Technologies for Cancer Research in the ACGT Project", *Production Grids in Asia*, Volume . ISBN 978-1-4419-0102-6. Springer-Verlag US, 2010, p. 93

Marcin Lawenda and Marcin Okon and Ariel Oleksiak and Bogdan Ludwiczak and Tomasz Piontek and Juliusz Pukacki and Norbert Meyer and Jaroslaw Nabrzyski and Maciej Stroinski: "Running Interactive Jobs in the Grid Environment" *PPAM*, Vol. 3911Springer (2005)

Juliusz Pukacki, Michal Kosiedowski, Rafal Mikolajczak, Marcin Adamski, Piotr Grabowski, Michal Jankowski, Miroslaw Kupczyk, Cezary Mazurek, Norbert Meyer, Jarek Nabrzyski, Tomasz Piontek, Michael Russell, Maciej Stroinski, Marcin Wolski "Programming Grid Applications with Gridge" (2008)

Kosiedowski M., Kurowski K., Mazurek, C., Nabrzyski J., Pukacki J.: "Workflow applications in GridLab PROGRESS projects", *Concurrency and Computation: Practice and Experience*, John Wiley & Sons, Ltd., vol.18, str. 1141-1154 (2006).

Krzysztof Kurowski, Bogdan Ludwiczak, Jarek Nabrzyski, Ariel Oleksiak, Juliusz Pukacki: "Dynamic grid scheduling with job migration and rescheduling in the GridLab resource management system." *Scientific Programming 2004*, Volume 12

NIG Japan

Kwon Y, Shigemoto Y, Kuwana Y, Sugawara H. Web API for biology with a workflow navigation system. *Nucleic Acids Research* 2009 May 5. [Epub ahead of print]
doi:10.1093/nar/gkp300

Fukuchi S, Homma K, Sakamoto S, Sugawara H, Tateno Y, Gojobori T, Nishikawa K. The GTOP database in 2009: updated content and novel features to expand and deepen insights into protein structures and functions. *Nucleic Acids Res.* 2009 Jan;37(Database issue):D333-7

Sugawara H, Ieko K, Fukuchi S, Gojobori T, Tateno Y. DDBJ dealing with mass data produced by the second generation sequencer. *Nucleic Acids Res.* 2009 Jan;37(Database issue):D16-8

Sugawara H, Ogasawara O, Okubo K, Gojobori T, Tateno Y. DDBJ with new system and face. *Nucleic Acids Res.* 2008 Jan;36(Database issue):D22-4.

Tanaka N, Uchino M, Miyazaki S, Sugawara H. Identification of discriminative characteristics for clusters from biologic data with InforBIO software. *BMC Bioinformatics.* 2007 Aug 2;8:281.

Kosuge T, Abe T, Okido T, Tanaka N, Hirahata M, Maruyama Y, Mashima J, Tomiki A, Kurokawa M, Himeno R, Fukuchi S, Miyazaki S, Gojobori T, Tateno Y, Sugawara H. Exploration and grading of possible genes from 183 bacterial strains by a common protocol to identification of new genes: Gene Trek in Prokaryote Space (GTPS). *DNA Res.* 2006 Dec 31;13(6):245-54.

Hirahata M, Abe T, Tanaka N, Kuwana Y, Shigemoto Y, Miyazaki S, Suzuki Y, Sugawara H. Genome Information Broker for Viruses (GIB-V): database for comparative analysis of virus genomes. *Nucleic Acids Res.* 2007 Jan;35(Database issue):D339-42.

Tanaka N, Abe T, Miyazaki S, Sugawara H. G-InforBIO: integrated system for microbial genomics. *BMC Bioinformatics.* 2006 Aug 4;7:368.

Abe T, Sugawara H, Kinouchi M, Kanaya S, Ikemura T. Novel phylogenetic studies of genomic sequence fragments derived from uncultured microbe mixtures in environmental and clinical samples. *DNA Res.* 2005;12(5):281-90.

Sugawara H, Miyazaki S. Biological SOAP servers and web services provided by the public sequence data bank. *Nucleic Acids Res.* 2003 Jul 1;31(13):3836-9.

Sugawara H, Miyazaki S. AHMII: Agent to Help Microbial Information Integration. *Nucleic Acids Res.* 2003 Jul 1;31(13):3727-8

DEI – Milan, Italy

Fabrizio Ferrandi, Luca Fossati, Marco Lattuada, Gianluca Palermo, Donatella Sciuto, Antonino Tumeo: Automatic Parallelization of Sequential Specifications for Symmetric MPSoCs. IESS 2007: 179-192

Simone Borgio, Davide Bosisio, Fabrizio Ferrandi, Matteo Monchiero, Marco D. Santambrogio, Donatella Sciuto, Antonino Tumeo: Hardware DWT accelerator for MultiProcessor System-on-Chip on FPGA. ICSAMOS 2006: 107-114

Antonino Tumeo, Matteo Monchiero, Gianluca Palermo, Fabrizio Ferrandi, Donatella Sciuto: A design kit for a fully working shared memory multiprocessor on FPGA. ACM Great Lakes Symposium on VLSI 2007: 219-222

Antonino Tumeo, Matteo Monchiero, Gianluca Palermo, Fabrizio Ferrandi, Donatella Sciuto: A Pipelined Fast 2D-DCT Accelerator for FPGA-based SoCs. ISVLSI 2007: 331-336

Antonino Tumeo, Christian Pilato, Fabrizio Ferrandi, Pier Luca Lanzi, Donatella Sciuto: "Ant Colony Optimization for Mapping and Scheduling in Heterogeneous Multiprocessor System". ICSAMOS 2008: 142-149

Institute for Neuro- and Bioinformatics, University of Lübeck (Dr Steffen Möller)

Hajo N. Krabbenhöft, Steffen Möller, and Daniel Bayer. Integrating ARC Grid Middleware with Taverna Workflows. *Bioinformatics*, 24(9):1221-1222, 2008.

Emanuela Merelli, Giuliano Armano, Nicola Cannata, Flavio Corradini, Mark d'Inverno, Andreas Doms, Phillip Lord, Andrew Martin, Luciano Milanese, Steffen Möller, Michael Schroeder, and Michael Luck. Agents in bioinformatics, computational and systems biology. *Briefings in Bioinformatics*, 8(1):45-59, 2007.

Daniel F. Schwarz, Oliver Hädicke, Jeanette Erdmann, Andreas Ziegler, Daniel Bayer, and Steffen Möller. SNPtoGO: characterizing SNPs by enriched GO terms. *Bioinformatics*, 24(1):146-148, 2007.

Robert Goertsches, Pablo Serrano-Fernández, Steffen Möller, Dirk Koczan, and Uwe K. Zettl. Multiple sclerosis therapy monitoring based on gene expression. *Curr. Pharm. Des.*, 12(29):3761-3779, 2006.

Steffen Möller, Uwe K. Zettl, Robert Goertsches, and Pablo Serrano-Fernández. Comparative genomics for the investigation of autoimmune diseases. *Curr. Pharm. Des.*, 12(29):3707-3722, 2006.

The Ontario Institute for Cancer Research (OICR).

Haider S, Ballester B, Smedley D, Zhang J, Rice P and Kasprzyk A. BioMart Central Portal—unified access to biological data. *Nucleic Acids Research*, 2009, Vol. 37, No. suppl_2 W23-W27.

Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, Kasprzyk A. BioMart - biological queries made easy. *BMC Genomics*. 2009 Jan 14;10(1):22

Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, Huber W. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*. 2005 Aug 15;21(16):3439-40.

Kasprzyk A, Keefe D, Smedley D, London D, Spooner W, Melsopp C, Hammond M, Rocca-Serra P, Cox T, Birney E. EnsMart: a generic system for fast and flexible access to biological data. *Genome Res*. 2004 Jan;14(1):160-9.

MPIZ: Max Planck Institute for Plant Breeding Research; Plant Computational Biology, Heiko Schoof

The tomato genome sequencing consortium. A snapshot of the emerging tomato genome sequence. (submitted to *Plant Genome*)

Jöcker A, Hoffmann F, Groscurth A, Schoof H. Protein function prediction and annotation in an integrated environment powered by web services (AFAWE) *Bioinformatics* (in press) doi:10.1093/bioinformatics/btn394

The BioMoby Consortium. Interoperability with Moby 1.0—It's better than sharing your toothbrush! *Brief Bioinform* 9(3):220-231 (2008) doi:10.1093/bib/bbn003

Kolmos E, Schoof H, Pluemer M, Davis SJ. Structural insights into the function of the core-circadian factor TIMING OF CAB2 EXPRESSION 1 (TOC1). *J Circadian Rhythms* 6(1):3 (2008).

The Rice Annotation Project. Curated Genome Annotation of *Oryza sativa* ssp. *japonica* and Comparative Genome Analysis with *Arabidopsis thaliana*. *Genome Research* 17(2):175-183 (2007).

Ballvora A, Jöcker A, Viehöver P, Ishihara H, Paal J, Meksem K, Bruggmann R, Schoof H, Weisshaar B, Gebhardt C.: Comparative sequence analysis of *Solanum* and *Arabidopsis* in a hot spot for pathogen resistance on potato chromosome V reveals a patchwork of conserved and rapidly evolving genome segments. *BMC Genomics* 8:112 (2007).

Spannagl M, Noubibou O, Haase D, Yang L, Gundlach H, Hindemitt T, Klee K, Haberer G, Schoof H, Mayer KF.: MIPSPlantsDB--plant database resource for integrative and comparative plant genome research. *Nucleic Acids Res*. 35:D834-D840 (2007).

Cannon SB et al. (2006). Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. *PNAS* 103(40):14959-64.

Schoof H, Spannagl M, Yang L, Ernst R, Gundlach H, Haase D, Haberer G, Mayer KF. (2005). Munich information center for protein sequences plant genome resources. A framework for integrative and comparative analyses. *Plant Physiol*. 38(3):1301-9.

- Wilkinson M, Schoof H, Ernst R, Haase D. (2005). BioMOBY Successfully Integrates Distributed Heterogeneous Bioinformatics Web Services: The PlaNet Exemplar Case. *Plant Physiol.* 138(1):5-17.
- Rudd, S., Schoof, H. and Mayer, K.F.X. (2005). PlantMarkers – a database of predicted molecular markers from plants. *Nucleic Acids Research* 33: D628–D632.
- Englbrecht, C., Schoof, H. and Boehm, S. (2004). Conservation, diversification and expansion of C2H2 zinc finger proteins in the *Arabidopsis thaliana* genome. *BMC Genomics* 5(1):39.
- Schoof, H., Ernst, R. and Mayer, K.F.X. (2004). The PlaNet consortium: A network of European plant databases connecting plant genome data in an integrated biological knowledge resource. *Comp. Funct. Genom.* 5:184-189.
- Schoof, H., Ernst, R., Nazarov, V., Pfeifer, L., Mewes, H.-W., and Mayer, K.F.X. (2004). MIPS *Arabidopsis thaliana* Database (MAtdB): An integrated biological knowledge resource for plant genomics. *Nucleic Acids Research* 32:D373-D376.
- Schoof, H (2003). Towards interoperability in genome databases: the MAtdB (MIPS *Arabidopsis thaliana* database) experience. *Comp. Funct. Genom.* 4: 255-258.
- Schoof, H., and W. M. Karlowski (2003). Comparison of rice and *Arabidopsis* annotation. *Curr.Opin.Plant Biol.* 6: 106-112.
- Karlowski, W. M., H. Schoof, V. Janakiraman, V. Stümpflen, and K. F. X. Mayer (2003). MOsDB: an integrated information resource for rice genomics. *Nucleic Acids Research* 31: 190-192.
- Schoof, H., P. Zaccaria, H. Gundlach, K. Lemcke, S. Rudd, G. Kolesov, R. Arnold, H. W. Mewes, and K. F. X. Mayer (2002). MIPS *Arabidopsis thaliana* Database (MAtdB): an integrated biological knowledge resource based on the first complete plant genome. *Nucleic Acids Research* 30: 91-93.
- Kazusa DNA Research Institute et al. (2000). Sequence and analysis of chromosome 5 of the plant *Arabidopsis thaliana*. *Nature* 408, 823-826.
- European Union Chromosome 3 *Arabidopsis* Genome Sequencing Consortium et al. (2000). Sequence and Analysis of chromosome 3 of the plant *Arabidopsis thaliana*. *Nature* 408, 820-823.
- The *Arabidopsis* Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796-815.
- Schoof, H., Lenhard, M., Haecker, A., Mayer, K. F. X., Jürgens, G., and Laux, T. (2000). The stem cell population of *Arabidopsis* shoot meristems is maintained by a regulatory loop between *CLAVATA* and *WUSCHEL* genes. *Cell* 100, 635-644.

Moussian, B., Schoof, H., Haecker, A., Jürgens, G., and Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. *EMBO Journal* 17, 1799-1809.

Mayer, K. F. X., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. *Cell* 95, 805-815.

Laux, T., and Schoof, H. (1997). Maintaining the shoot meristem - the role of CLAVATA1. *Trends in plant science* 2, 325-327.

Applied Computational Biology and Bioinformatics, Paterson Institute for Cancer Research UK, Crispin Miller

Chan CW, Wong NA, Liu Y, Bicknell D, Turley H, Hollins L, Miller CJ, Wilding JL, Bodmer WF. Gastrointestinal differentiation marker Cytokeratin 20 is regulated by homeobox gene CDX1. *Proc Natl Acad Sci U S A*. 2009 Feb 10;106(6):1936-41. Epub 2009 Feb 2. [pubmed](#)

Barenco M, Papouli E, Shah S, Brewer D, Miller CJ, Hubank M. rHVDM: an R package to predict the activity and targets of a transcription factor. *Bioinformatics* 2009 Feb 1;25(3):419-20. Epub 2008 Dec 15. [pubmed](#)

Sims AH, Smethurst GJ, Hey Y, Okoniewski MJ, Pepper SD, Howell A, Miller CJ, Clarke RB. The removal of multiplicative, systematic bias allows integration of breast cancer gene expression datasets - improving meta-analysis and prediction of prognosis. *BMC Med Genomics*. 2008 Sep 21;1(1):42. [pubmed](#)

Linton KM, Hey Y, Saunders E, Jeziorska M, Denton J, Wilson CL, Swindell R, Dibben S, Miller CJ, Pepper SD, Radford JA, Freemont AJ. Acquisition of biologically relevant gene expression data by Affymetrix microarray analysis of archival formalin-fixed paraffin-embedded tumours. *Br J Cancer*. 2008 Apr 1 [Epub ahead of print]. [pubmed](#)

Okoniewski MJ, Miller CJ (2008) Comprehensive Analysis of Affymetrix Exon Arrays Using BioConductor. *PLoS Computational Biology* 4(2): e6
doi:10.1371/journal.pcbi.0040006

Danny A Bitton, Michal J Okoniewski, Yvonne Connolly and Crispin J Miller Exon level integration of proteomics and microarray data *BMC Bioinformatics* (2008), 9:118

Tim Yates, Michal J. Okoniewski, and Crispin J. Miller X:Map: annotation and visualization of genome structure for Affymetrix exon array analysis *Nucleic Acids Research, Database Issue, Advance Access published on October 11, 2007.*
doi:10.1093/nar/gkm779 [Full Text](#)

Carla S Moller-Levet; Catharine M West; Crispin J Miller. Exploiting sample variability to enhance multivariate analysis of microarray data *Bioinformatics* 2007;
doi: 10.1093/bioinformatics/btm441

Stuart D Pepper, Emma K Saunders, Laura E Edwards, Claire L Wilson, Crispin J Miller (2007) The utility of MAS5 expression summary and detection call algorithms. *BMC Bioinformatics* 8:273.

Michal J Okoniewski, Tim Yates, Sian Dibben and Crispin J Miller (2007) An annotation infrastructure for the analysis and interpretation of Affymetrix exon array data. [Genome Biology 8:R79](#)

Ebrahim Sakhinia, Caroline Glennie, Judith A. Hoyland, Lia P. Menasce, Gerard Brady, Crispin Miller, John A. Radford, and Richard J. Byers (2007) Clinical quantitation of diagnostic and predictive gene expression levels in follicular and diffuse large B-cell lymphoma by RT-PCR gene expression profiling. [Blood, Vol. 109, No. 9, pp. 3922-3928.](#)

Martino Barenco, Sonia Shah, Daniel Brewer, Robin Callard, Jaroslav Stark, Crispin Miller, Michael Hubank (2007) rHVDM - a fasta and user-friendly R package to predict transcription factor targets from microarray time series data [BMC Systems Biology 2007, 1\(Suppl 1\):S1 \(8 May 2007\)](#)

A. H. Sims; P. Finnon; C. J. Miller; S. D. Bouffler; A. Howell; D. Scott; R. B. Clarke (2007) TPD52 and NFKB1 gene expression levels correlate with G2 chromosomal radiosensitivity in lymphocytes of women with and at risk of hereditary breast cancer. *International Journal of Radiation Biology*, 1 - 12.

Stuart C. Winter, Francesca M. Buffa, Priyamal Silva, Crispin Miller, Helen R. Valentine, Helen Turley, Ketan A. Shah, Graham J. Cox, Rogan J. Corbridge, Jarrod J. Homer, Brian Musgrove, Nick Slevin, Philip Sloan, Pat Price, Catharine M.L. West and Adrian L. Harris (2007) Relation of a Hypoxia Metagene Derived from Head and Neck Cancer to Prognosis of Multiple Cancers. *Cancer Research* 67(7):3441:3449.

Okoniewski MJ, Hey Y, Pepper SD, Miller C (2007) "High correspondence between Affymetrix exon and standard expression arrays." *Biotechniques*, in press.

SD Pepper, H Yvonne, G Newton, M Okoniewski, CJ Miller (2006) A core lab case study: Exon Array Challenges and Opportunities. *The Affymetrix Core Lab Communication Bulletin* 1:1-4.

Meyer S, Fergusson WD, Whetton AD, Moreira-Leite F, Pepper SD, Miller C, Saunders EK, White DJ, Will AM, Eden T, Ikeda H, Ullmann R, Tuerkmen S, Gerlach A, Klopocki E, Tönnies H (2006). "Amplification and translocation of 3q26 with overexpression of EVI1 in Fanconi anemia derived childhood acute myeloid leukemia with biallelic FANCD1/BRCA2 disruption." *Genes Chromosomes Cancer* in press.

Wilson, C.L., A.H. Sims, A. Howell, C.J. Miller, and R.B. Clarke. (2006). "Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue" *Endocr Relat Cancer*, 13,617-28; [abstract](#).

Michal J Okoniewski and Crispin J Miller (2006) "Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations" *BMC Bioinformatics* 2006, 7:276; [abstract](#).

Richard D Unwin, Duncan L Smith, David Blinco, Claire L Wilson, Crispin J Miller, Caroline A Evans, Ewa Jaworska, Stephen A Baldwin, Kay Barnes, Andrew Pierce, Elaine Spooncer, and Anthony D Whetton (2006) "Quantitative proteomics reveals post-translational control as a regulatory factor in primary hematopoietic stem cells" Blood First Edition Paper, prepublished online February 28, 2006; [DOI 10.1182/blood-2005-12-4995](https://doi.org/10.1182/blood-2005-12-4995).

Claire L Wilson and Crispin J Miller (2005) "Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis" Bioinformatics, 21(18), 3683-5.

Hui Sun Leong, Tim Yates, Claire Wilson and Crispin J. Miller (2005) "ADAPT: A Database of Affymetrix Probesets and Transcripts" Bioinformatics, Advance Access, published online on March 3, 2005 [pubmed](#)

Claire L. Wilson, Paul E. Boardman, Andrew J. Doig and Simon J. Hubbard (2004) "Improved prediction for N-termini of α -helices using empirical information" Proteins: Structure, Function and Bioinformatics, 1;57(2), 322-30 [pubmed](#)

Claire L. Wilson, Stuart D. Pepper, Yvonne Hey, and Crispin J. Miller (2004) "Amplification protocols introduce systematic but reproducible errors into gene expression studies" BioTechniques, 36, 498-506. [pubmed](#)

Janine T. Erler, Christopher J. Cawthorne, Kaye J. Williams, Marianne Koritzinsky, Bradley G. Wouters, Clare Wilson, Crispin Miller, Costas Demonacos, Ian J. Stratford, and Caroline Dive (2004) "Hypoxia-Mediated Down-Regulation of Bid and Bax in Tumors Occurs via Hypoxia-Inducible Factor 1-Dependent and -Independent Mechanisms and Contributes to Drug Resistance" Molecular and Cellular Biology, 24, 2875-2889. [pubmed](#)

Mabey J.E., Anderson, M.J., Giles P.F., Miller C.J., Attwood T.K., Paton N.W., Bornberg-Bauer E., Robson G.D., Oliver S.G. and Denning D.W. (2004) "CADRE: the Central Aspergillus Data Repository" Nucleic Acids Res, 32(1), D401-5. [pubmed](#)

Crispin J. Miller, Heba S. Kassem, Stuart D. Pepper, Yvonne Hey, Timothy H. Ward, and Geoffrey P. Margison (2003) "Mycoplasma Significantly alters microarray gene expression profiles" BioTechniques, 35, 812-814. [pubmed](#)

Miller, C.J. and Attwood, T.K. (2003) "Bioinformatics goes back to the future" Nature Reviews Molecular Cell Biology, 4, 157-162. [pubmed](#), [PDF](#), [journal](#).

Attwood, T.K. and Miller, C.J. (2002) "Progress in bioinformatics and the importance of being earnest" Biotechnology Annual Review, 8, 1-55. [amazon](#)

Attwood, T.K. and Miller, C.J. (2001) "Which craft is best in bioinformatics?" Computers and Chemistry, 25(4), 329-339. [pubmed](#)

Miller, C.J. and Attwood, T.K. (2001) "PSST...The probabilistic sequence search tool.", IEEE International Symposium on Bioinformatics and Bioengineering. Washington DC, November 2001. IEEE press, ISBN 0-7695-1423-5. [IEEE Digital library](#)

Cornell, M., Paton, N.W., Wu, S., Goble, C.A., Miller, C.J., Kirby, P., Eilbeck, K., Brass, A., Hayes, A. and Oliver, S.G., (2001) "GIMS - A Data Warehouse for Storage and Analysis of Genome Sequence and Functional Data", IEEE International Symposium on Bioinformatics and Bioengineering. Washington DC, November 2001. IEEE press, ISBN 0-7695-1423-5. [PDF](#)

Publications Pre 2000

Miller C, Gurd J, and Brass A. (1999) "A RAPID algorithm for sequence database comparisons: application to the identification of vector contamination in the EMBL databases." *Bioinformatics* Feb;**15**(2), 111-21. [pubmed](#)

Generic references (not necessarily from Partners)

Burrage, K, Hood, L, Ragan, MA (2006). Advanced computing for systems biology. *Brief. Bioinformatics*, 7, 4:390-8.

Chin, G, Chavarria, DG, Nakamura, GC, Sofia, HJ (2008). BioGraphE: high-performance bionetwork analysis using the Biological Graph Environment. *BMC Bioinformatics*, 9 Suppl 6:S6.

Claus, BL, Johnson, SR (2008). Grid computing in large pharmaceutical molecular modeling. *Drug Discov. Today*, 13, 13-14:578-83.

Craddock T, Harwood CR, Hallinan J, Wipat A. Relieving bottlenecks in large-scale genome analyses. *Nat Rev Microbiol*. 2008 Dec;**6**(12):948-54.

Lu, G, Ni, J (2008). Highlighting computations in bioscience and bioinformatics: review of the Symposium of Computations in Bioinformatics and Bioscience (SCBB07). *BMC Bioinformatics*, 9 Suppl 6:S1.

Oinn, T. et al. Taverna: a tool for the composition and enactment of bioinformatics workflows. *Bioinformatics* 20, 30453054 (2004).

Stajich, J. E. et al. The Bioperl toolkit: perl modules for the life sciences. *Genome Res*. 12, 16111618 (2002).

Hull, D. et al. Taverna: a tool for building and running workflows of services. *Nucleic Acids Res*. 34, W729W732 (2006).

Pillai, S. et al. SOAP-based services provided by the European Bioinformatics Institute. *Nucleic Acids Res*. 33, W25W28 (2005).

Harwood, C. R. & Wipat, A. in Proc. 5th UK eScience All Hands Meet. 788795 (2006). Harwood, C. R. & Cranenburgh, R. Bacillus protein secretion: an unfolding story. *Trends Microbiol*. 16, 7379 (2008).

Sulakhe, D. et al. Gnare: automated system for highthroughput genome analysis with Grid computational backend. *J. Clin. Monit. Comput*. 19, 361369 (2005)

Trelles, O (2001). On the parallelisation of bioinformatics applications. *Brief. Bioinformatics*, 2, 2:181-94.

Maltsev, N. A. et al. PUMA2-grid-based highthroughput analysis of genomes and metabolic pathways. *Nucleic Acids Res.* 34, D369D372 (2006)

Carolina Perez-Iratxeta, Miguel A. Andrade-Navarro, and Jonathan D. Wren * Evolving research trends in bioinformatics , *Briefings in Bioinformatics* , doi:10.1093/bib/bbl035 (2006)

Short-read sequencing technologies for transcriptional analyses. Simon SA, Zhai J, Nandety RS, McCormick KP, Zeng J, Mejia D, Meyers BC, *Annu Rev Plant Biol.* 2009;60:305-33.

Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. Lister R, Gregory BD, Ecker JR. *Curr Opin Plant Biol.* 2009. Apr;12(2):107-18. Epub 2009 Jan 20.

Applications of next-generation sequencing technologies in functional genomics. Morozova O, Marra MA. *Genomics.* 2008 Nov;92(5):255-64. Epub 2008 Aug 24.