Diego Diez

Date of birth: March 19, 1972
Nationality: Spanish
Civil status: Married
Interests: Aikido, music, reading and sports

Address: Kanakusahara 6-1-205, 611-0002 Kohata, Uji, Japan
Phone/Fax: +81 (0) 774 31 4096
E-mail: diego10ruiz@gmail.com
URL: http://www.mendeley.com/profiles/diego-diez/

Current position

Position	Postdoctoral researcher
Organism	Kyoto University
Institute	Institute for Chemical Research, Bioinformatics center
Department	Chemical Life (Prof. Kanehisha and Prof. Goto laboratory)
Address	Gokasho, Uji 611-0011, Japan
Phone	+81 (0) 774 38 3296
Fax	+81 (0) 774 38 3269
E-mail	diez@kuicr.kyoto-u.ac.jp
URL	http://web.kuicr.kyoto-u.ac.jp/∼diez

Academic qualifications

Ph.D.	Biochemistry, *Universidad Autonoma de Madrid*, April 2006.
M.Sc.	Biochemistry, *Universidad Autonoma de Madrid*, February 2004.
B.Sc	Biochemistry, *Universidad Complutense de Madrid*, February 2000.

Languages

Spanish	Native
English	High
Japanese	Medium (JLPT level 3)

Other training

- *Introduction to the use and programing in R(S) language*. CSIC Bioinformatic Network course. May 2003.
- *14th Basic Radiactivity Course ("Curso Basico de Radiactividad")* at Instituto de Investigaciones Biomédicas "Alberto Sols", CSIC/UAM. December 2nd-4th, 2003
- *Animal Research Training Course* at Instituto de Investigaciones Biomédicas, "Alberto Sols", CSIC/UAM, 2003.
- *Programing Analyst Course.* Programming techniques on Visual Basic, C++, Java and Databases. Kernel School Center (500 hours). February - July 2000.

Publications

$\star$ Paper highlighted in "Traslational highlights from Endocrinology" in *The Journal of clinical endocrinology and metabolism*, 2011. 96(3): p. 870.

† Best paper award at *GIW2009: 20th International Conference on Genome Informatics*. 2009, Yokohama, Japan.

$\oint$ Top ten downloaded papers from *Molecular Biosystems* Journal during three consecutive months after publication.

1. **Diez D**, Sanchez-Jimenez F, and Ranea JA, *Evolutionary expansion of the Ras switch regulatory module in eukaryotes*. Nucleic Acids Res, 2011. 39(13): p. 5526-37.

2. $\star$ Grijota-Martinez C, **Diez D**, Morreale de Escobar G, Bernal J, and Morte B, *Lack of action of exogenously administered t3 on the fetal rat brain despite expression of the monocarboxylate transporter 8*. Endocrinology, 2011. 152(4): p. 1713-21.

3. Kirwan G, **Diez D**, Haeggström J, Goto S, and Wheelock C, *Systems Biology Approaches for Investigating the Relationship Between Lipids and Cardiovascular Disease*. Current Cardiovascular Risk Reports, 2011. 5(1): p. 52-61.

4. Morte B, Ceballos A, **Diez D**, Grijota-Martinez C, Dumitrescu AM, Di Cosmo C, Galton VA, Refetoff S, and Bernal J, *Thyroid hormone-regulated mouse cerebral cortex genes are differentially dependent on the source of the hormone: a study in monocarboxylate transporter-8- and deiodinase-2-deficient mice*. Endocrinology, 2010. 151(5): p. 2381-7.

5. Morte B, **Diez D**, Auso E, Belinchon MM, Gil-Ibanez P, Grijota-Martinez C, Navarro D, Morreale de Escobar G, Berbel P, and Bernal J, *Thyroid Hormone Regulation of Gene Expression in the Developing Rat Fetal Cerebral Cortex: Prominent Role of the Ca2+/Calmodulin-Dependent Protein Kinase IV Pathway*. Endocrinology, 2010. 151(2): p. 810-820.

6. $\oint$ **Diez D**, Wheelock AM, Goto S, Haeggstrom JZ, Paulsson-Berne G, Hansson GK, Hedin U, Gabrielsen A, and Wheelock CE, *The use of network analyses for elucidating mechanisms in cardiovascular disease*. Mol Biosyst, 2010. 6(2): p. 289-304.

7. **Diez D**, Hayes N, Joannin N, Normark J, Kanehisa M, Wahlgren M, Wheelock CE, and Goto S, *varDB: A database of antigenic variant sequences-Current status and future prospects*. Acta Trop, 2010. 114(3): p. 144-151.

8. Wheelock CE, Wheelock AM, Kawashima S, **Diez D**, Kanehisa M, van Erk M, Kleemann R, Haeggstrom JZ, and Goto S, *Systems biology approaches and pathway tools for investigating cardiovascular disease*. Mol Biosyst, 2009. 5(6): p. 588-602.

9. † Hayes CN, **Diez D**, Joannin N, Kanehisa M, Wahlgren M, Wheelock CE, and Goto S, *Tools for investigating mechanisms of antigenic variation: new extensions to varDB*. Genome Inform, 2009. 23(1): p. 46-59.

10. Folkersen L, **Diez D**, Wheelock CE, Haeggstrom JZ, Goto S, Eriksson P, and Gabrielsen A, *GeneRegionScan: a Bioconductor package for probe-level analysis of specific, small regions of the genome*. Bioinformatics, 2009. 25(15): p. 1978-9.

11. Hayes CN, **Diez D**, Joannin N, Honda W, Kanehisa M, Wahlgren M, Wheelock CE, and Goto S, *varDB: a pathogen-specific sequence database of protein families involved in antigenic variation*. Bioinformatics, 2008. 24(21): p. 2564-5.

12. **Diez D**, Grijota-Martinez C, Agretti P, De Marco G, Tonacchera M, Pinchera A, de Escobar GM, Bernal J, and Morte B, *Thyroid hormone action in the adult brain: gene expression profiling of the effects of single and multiple doses of triiodo-L-thyronine in the rat striatum*. Endocrinology, 2008. 149(8): p. 3989-4000.

13. **Diez D**, Alvarez R, and Dopazo A, *Codelink: an R package for analysis of GE healthcare gene expression bioarrays*. Bioinformatics, 2007. 23(9): p. 1168-9.

Professional experience

My research interests are focused on the understanding of cellular networks, how they are constructed and regulated, and how they are dynamically shaped during evolution. Both during my Ph.D. and as a postdoctoral researcher I have been working on different aspects of gene regulation related to the actions of thyroid hormones in the brain. I have also performed some work on evolutionary aspects of signaling networks. As a postdoctoral researcher I initially focused my research activities on studying antigenic variation networks in malaria parasites, in particular the transcriptional networks associated with antigenic variant gene families. Later, I got interested on systems biology approaches to the elucidation of cellular networks, in particular, integration of high-throughput datasets (e.g. gene expression and protein-protein interaction networks) and the discovery of disease-associated altered subnetworks. Since then I have focused my research on the application of network approaches to understand inflammatory diseases like atherosclerosis and asthma. Below is a more detailed summary of my main research achievements and ongoing projects.

Application of systems biology approaches to investigate cardiovascular disease

Cardiovascular disease (CVD) is the first cause of premature death in industrialized societies. The main cause of CVD is atherosclerosis, an inflammatory response in the wall of the arteries due to the accumulation of low density lipoproteins. In collaboration with clinicians at Karolinska University Hospital, we used microarray hybridizations from a databank of patients that had preventive carotid endarterectomy. This database contains information about over 60 phenotypic and clinical variables from patients before and after surgery. Our focus was on clinically relevant variables like stability of the plaque (probability to suffer a clinical episode- e.g. heart attack- after surgery), LDL/HDL ratio, glomerular filtration rate (kidney function), treatment with statins (a cholesterol lowering drug), and phenotypical traits like age and gender. Our initial study focused on gender differences in cardiovascular disease. Males have higher susceptibility than females of the same age rage to suffer atherosclerosis. In a network analysis combining transcriptional data with literature mining to obtain associations between selected genes, we identified a module that linked 'small GTPase signaling', 'immune response' and 'lipid homeostasis' Gene Ontology terms via a few hub genes, specifically RAB10 and SOAT1. Among the selected genes, APOC1 was highlighted as an autosomal gene upregulated in males compared to females. Among all the terms associated with APOC1, 'phospholipid efflux' showed the strongest evidence, suggesting a role of this gene and associated processes in gender-related differences in atherosclerosis susceptibility (Diez et al. Molecular Biosystems, 2010). In a follow up study we included information about other variables (stability, statins, age, LDL/HDL, etc) and combined transcriptomics and protein-protein interaction data. In this work we identified regulatory modules that link clinical manifestations to transcriptional pathways, providing insight into the mechanisms of atherosclerosis action. The results for these studies will be published on the near future.

Identification of network modules driving disease mechanisms in asthmatics

Asthma is an inflammatory disease of the airways, with genetic and environmental factors, which is characterized by variable and recurring symptoms (wheezing, coughing, chest tightness, shortness of breath), airflow obstruction and bronchospasm. As of 2009 it affected 300 million people worldwide, being the cause of over 250,000 deaths. One common treatment for asthmatics is the use of glucocorticoids, because these molecules inhibit the inflammatory response and reduce the symptoms. Unfortunately glucocorticoids have also associated a long list of side effects due to its pleiotropic actions. Although the biological mechanisms of glucocorticoids are well understood, a complete understanding of alternative regulatory actions may help design specific drugs that avoid the development of side effects. This project is in collaboration with a group of clinicians at University of California, San Francisco. We use a microarray dataset of asthmatics before and after treatment with either placebo or the glucocorticoid fluticasone propionate (Flovent), and combine the use of multivariate statistics (MVA) with network analysis techniques. We analyzed the microarray data using O-PLS (orthogonal projection to latent structures), and have developed a method to derive a probabilities from the correlations to the O-PLS model. This P-values representing transcriptional changes associated with fluticasone actions were combined with signaling network information provided by protein-protein interactions. A network analysis methods was applied to extract the most significant module (subnetwork) representing the core of glucocorticoid actiones. Our method was able to retrieve a module that correctly links the actions of glucocorticoids with suppression of the immune response, providing a

mechanistic explanation to flovent clinical action (Kirwan G*, Diez D* et al., * equally contributing authors, manuscript in preparation). These results settle the basis for applying this methodology in other scenarios.

In an ongoing study we are analyzing microarray data from asthmatics with high and low levels of circulating IL-13. Low IL-13 asthmatics show a reduced Th2 response and respond poorly to glucocorticoid treatment, a characteristic that has serious implications during asthma exacerbation episodes. We are applying the previously described workflow to elucidate the mechanisms behind IL-13 associated differences. We integrated microarray data from control and asthmatics with low and high IL-13 levels, with information about protein-protein interactions. Network analysis provided optimal subnetworks and additional suboptimal modules, which may provide the clue behind different populations of asthmatics. The results of this project will be publish in the near future (Diez D et al., manuscript in preparation).

Reconstruction of thyroid hormone transcriptional networks in the brain

As a Ph.D. student in Dr. Juan Bernal's laboratory we were interested on the effects of thyroid hormones in the brain. To this aim we worked on the identification of effectors and regulators of thyroid hormone (TH) actions in the fetal and adult brains. As a postdoctoral researcher I have continued my collaboration with Dr. Bernal's laboratory, in particular, to elucidate the role of MTC8 transporter in thyroid hormone regulated gene expression. MCT8 is a thyroid hormone (dual T4/T3) transporter whose mutations are associated to a genetic disorder previously known as Allan-Herndon-Dudley syndrome.

Several sources of evidence suggested a role of maternal thyroid hormone in the developing brain before the onset of the fetal thyroid gland. To clarify this hypothesis, and since most of the thyroid hormone actions are mediated by the regulation of gene expression, we tried to identify genes regulated by maternal THs in the fetal brain. We used a rat animal model and tested different groups of control and hypothyroid animals, untreated and treated with THs. We measured the expression of genes in fetal brain tissue (cortex, hippocampus and cerebellum) using Codelink whole genome microarrays. Several parameters to assess the hypothyroid status of the animals were determined from fetal and maternal blood and liver (e.g. TSH, T3 and T4 levels). We identified many genes with thyroid hormone-dependent altered expression. Our results suggested that TH of fetal origin controls TH-regulated gene expression in the fetal brain, at least during the second half of gestation. A prominent role of the Camk4/Creb1 pathway (in which Camk4 activates the transcription factor Creb1) was discovered, with 56 of the differentially expressed genes (including Camk4) containing Creb1 regulatory elements (CRE) in their promoter regions (Morte B, Diez D et al. Endocrinology, 2009). Following these studies we observed a lack of effect of T3 (but not T4) on the expression of Camk4 on the fetal brain. We hypothesized an impairment of the T4/T3 transporter MCT8 and studied its patterns of expression in fetal cortex. Using cortex from fetal brain we found that T3 had no effect on the expression of several thyroid hormone regulated genes, including Camk4, Sema3c and Slc7a3, in spite the expression of MCT8. This suggests that T4 is the sole source of T3 in the fetal brain and casts doubts about the role of MCT8 as a T3 transporter in the fetus (Grijota-Martinez C, Diez D, et al. Endocrinology, 2011).

In the adult brain, we were motivated by some findings that linked TH to neurological disease states in the adult brain, where alterations in TH levels have been associated with depression and bi-polar disorders. We performed Affymetrix microarray analysis on adult hypothyroid rats after administering T3 (active form of TH). We found, for the first time, a set of genes whose expression was dependent on the levels of thyroid hormones in the adult brain, including many genes already regulated by TH in the postnatal age. Our results suggested that THs regulate several signaling pathways in the adult brain (Diez et al. Endocrinology, 2008). We then studied the role of MCT8 transporter in T3/T4 homeostasis in the brain. Using a MCT8 and DIO2 (Deionidase 2) knock-out and a MCT8/DIO2 double mutant we studied the effect on the expression of several genes previously identified as regulated by TH using microarray analysis. We found that disruption of the MCT8 receptor did not affect gene expression. However, MCT8/DIO2 double mutant did results in an altered expression patter, suggesting that lack of disruption by the MCT8 KO was most likely due to some DIO2-related compensatory mechanism (Morte B, Ceballos A, Diez D et al. Endocrinology, 2010).

Evolution of signaling networks in eukaryotes

Understanding cellular networks is critical to the effective development of therapeutic drugs that can alleviate symptoms and reduce mortality. Current networks evolved from common ancestors and were shaped to accommodate species, environmental and other restrictions. Understanding the evolution of cellular networks is also important to realize their differences in diverse organisms, which can help bring results obtained in model organisms to humans in translational medicine efforts. In collaboration with Dr. Juan Antonio Garcia-Ranea at Malaga University we studied the evolution of the Ras signaling network. Ras proteins regulate many aspects of cellular biology, including proliferation and differentiation pathways. The Ras switch includes the small GTP-binding proteins of the Ras family and the associated regulatory proteins (GEFs and GAPs). Distortion of this pathway is associated with disease, most remarkably cancer. In this project, we analyzed the evolutionary expansion of this module in eukaryotes, with a special emphasis on the metazoan clade. We found that the module's size correlates with organism's complexity, and is negatively correlated to the evolutionary distance to humans. Our subfamily based analysis identified some genes that drive the expansion, and most likely have been implicated in the appearance of cellular functions associated with complex organisms. Some organisms like the protists D. discoideum and M. brevicollis present different domains composition and architectures, suggesting divergent evolution of signaling pathways (Diez et al. NAR, 2011).

In a follow up study we are using the phylogenetic data obtained in our previous work to analyze the conservation between phylogenetic signal and signaling networks. Using protein-protein interaction data and phylogenetic distances we are investigating the correlation between evolutionary distance and distance in the signaling network for pairs of proteins. This study will enable to discover how conserved are signaling networks in different species, from an evolutionary perspective, and will be important to estimate the relevance of results obtained from model organisms.

Evolution of the pir superfamily in Plasmodium

When I started at the Bioinformatics Center in Kyoto, I began to work on aspects related to antigenic variation. Antigenic variation is a mechanism to evade the immune system of the host that is shared among many pathogens. Different variant gene families are involved, with usually no sequence similarity between different species. However, it is believed that the general mechanisms used to generate antigenic variability are restricted, and are common to many pathogens. Malaria causes over 500 million infections and between 1 3 million deaths every year. Plasmodium parasites, the causative agent of malaria, contain different antigenic variable gene families with no sequence similarity (var, rif/stevor, vir, kir, cir, bir and yir). Their mutually exclusive presence in different pathogen groups (human, primate and rodents) led to the hypothesis that these gene families evolved from a superfamily of antigenic variant sequences (referred to as pir) that have diverged. The main source of evidence it the mutually exclusive presence in different pathogens and some weak motif conservation found in a subgroup of sequences. However, the functional role of these proteins in regulating parasite virulence processes is not well understood. We performed a comparative/phylogenetic analyses on these gene families and found evidence that members of the vir/kir gene family, which is exclusive of primate parasites, are also present in the rodent parasites. We also found that these genes do not constitute a multi-gene family in the rodent parasites, suggesting that, unlike the primate homolgs, they are not involved in antigenic variation. Detailed motif and motif-architecture analysis shows that they are not members of the cir/bir/yir gene family (the rodent-specific gene family). Our results challenge the current hypothesis about the evolution of the pir gene family in plasmodium, and poses implications about the use of animal models to study malaria. A manuscript presenting these results will be submitted shortly (Diez et al., manuscript in preparation).

Antigenic variation database (varDB)

The study of antigenic variation presents several challenges, mainly because the inherently variability present in some of these gene families, and because the different mechanisms associated with antigenic variation in different pathogens. To stimulate the development of comparative analyses we created varDB (http://www.vardb.org), a database of antigenic variant gene families (Hayes CN, Diez D et al. Bioinformatics, 2009). The database is updated every three months including new sequences deposited in Genbank, as well as new species and gene families. The latest version of varDB, released on April 1st, 2011, contained information about 49 gene families, in 32 organisms and 23 diseases. Our commitment is to maintain this resource as a free tool for the research community, and to expand

its functionality in the future. varDB is the only existing resource about antigenic variation that covers multiple species and diseases, and its importance was highlighted in a publication by Allred et al in Trends in Parasitology (2009): “varDB: common ground for a shifting landscape“. My contribution to varDB focused on the development of a pipeline that identifies antigenic variant sequences from diverse sources, including Genbank and PlasmoDB. This pipeline is the core of the varDB sequence mining architecture (Diez et al. Acta Tropica, 2010).

Research positions as a postdoctoral researcher:

- Postdoctoral researcher. Bioinformatics center, Institute for Chemical Research, Kyoto University. Project: Integrated Database Project (TOUGOU). January 2009 to present.

  Investigating the use of systems biology approaches to the elucidation of mechanisms in cardiovascular disease and asthma. Integration of whole genome gene expression signatures with protein-protein integration information.

- Postdoctoral researcher. Bioinformatics center, Institute for Chemical Research, Kyoto University. Project: KAKENHI. September 2009 to December 2009.

  Investigating the use of systems biology approaches to the elucidation of mechanisms in cardiovascular disease. Integration of whole genome gene expression signatures with protein-protein integration information.

- Postdoctoral researcher. Bioinformatics center, Institute for Chemical Research, Kyoto University. *Japanese Society for the Promotion of Science (JSPS) fellowship*. September 2007 to August 2009.

  Investigating the use of systems biology approaches to the elucidation of mechanisms in cardiovascular disease. Application to co-transcriptional network estimation and literature mining techniques.

- Postdoctoral researcher. Bioinformatics center, Institute for Chemical Research, Kyoto University. Project: *21st Century COE Program from the Ministry of Education, Culture, Sports, Science and Technology of Japan* . April 2007 to August 2007.

  Evolutionary aspects of antigenic variation in different pathogenic species. Development of varDB and study of evolution of antigenic variant gene families.

- Postdoctoral researcher. Bioinformatics center, Institute for Chemical Research, Kyoto University. Project: *Strategic Japanese-Swedish Cooperative Program on ”Multidisciplinary Bio”. Probing the Plasmodium falciparum genome*. May 2006 to March 2007.

  Analysis of antigenic variation in the malaria parasite.

Experience as an independent researcher

I have initiated and am currently conducting research based on my original ideas. Below is a summary of the past/current projects:

- **Developing of methods to study cell signaling using multivariate methods and network tools.** My recent research interest are centered on the analysis of cellular networks. To this aim my collaborators and I were interested on the application of multivariate statistical methods. This methods, commonly used in fields like chemmometrics, may represent and advantage over univariate methods. This is specially true for complex diseases, like cardiovascular disease, that are the results of a mixture of genetic, environmental and life style factors. The main obstacle to the use of MVA is the integration with existing tools to mine signaling networks. To overcome this limitation I devised and developed a method to derive probabilities from the distribution of p(corr) (correlation of variables to an O-PLS model). This methodology uses a null dataset to determine the distribution of p(corr) under the null hypothesis. Then, test the p(corr) from the real variables to determine the associated probability. The methodology shows consistent results with univariate statistical methods in the univariate case. I am planning to apply this methodology to more complex scenarios (e.g. asthma IL-13 project, see above)

- **Integration of -omics datasets to reconstruct cellular networks.** To understand cellular networks we need to recreate models that represent this networks more faithfully. One limitation of the methods applied in my previous projects is that protein-protein interaction networks do not provide a direct link to the genes regulated by transcription factors, limiting the insight that can be extracted from the subnetwork results. To overcome this limitation I am working on methods that enable to integrate gene regulatory network information (transcription factor-target gene) from experimental (Chip-on-ChIP, ChIP-seq, etc) and computational predictions. This will create a very dense highly connected starting network, that will provide subnetworks more interpretable from a mechanistic perspective.
- **Evolution of the Ras switch in eukaryotes.** Understanding how signaling networks were shaped during evolution is critical to understand their current structure, regulation and when and why they fail in disease. In this project I pursued to get insight into the evolution of an important regulatory module in eukaryote signaling pathways. The approach combines phylogenetic, sequence and gene features analyses. I was responsible of a large percentage of ideas that led to this project, and consequently I hold responsibility for this work as corresponding author in the published paper.
- **Analysis of the pir gene family in Plasmodium.** After identifying an oddity in the distribution of antigenic variant *vir/kir* gene families in the rodent parasites, I hypothesized that this may represent a challenge to the current classification of these gene families in the Plasmodim genus. These results can represent a breakthrough in understanding evolution of antigenic variation in plasmodium species. However, the high level of variability associated with these gene families makes it challenging to derive conclusions from regular sequence analysis tools. To get insight into the origin and evolution of these families, I designed a bioinformatics approach to analyze this large set of sequences by comparing their domain composition and architectures. I will be corresponding author in the future publication.

Grants and fellowships awarded

I wrote the following awarded fellowships:

- *Japanese Society for the Promotion of Science (JSPS) fellowship*. Project: Analysis of antigenic variation in malaria. September 2007 to August 2009. Bioinformatics center, Institute for Chemical Research, Kyoto University.
- *I3P predoctoral fellowships. Ministerio de Ciencia y Tecnologa, I3P*. Project: *Bioinformatics Applied to Expression Arrays Analysis*. February 2002 to March 2006.

I have participated writing the following awarded grant applications:

- STING (The Swedish Foundation for International Cooperation in Research and Higher Education) grant. Project: “Development of bioinformatics methods for the analysis of hyper-variable gene families and probing of their effects upon host-pathogen interactions and parasite virulence“. August 2007 to July 2011.
- VINNOVA/JSPS (The Swedish Governmental Agency for Innovation Systems/Japanese Society for the Promotion of Science) grant. Project: “Systems biology approaches to understanding atherosclerosis“.

Scientific management experience

I have experience mentoring Ph.D. students and postdoctoral researchers. Usually I am involved on giving advice to young pre-doctoral students in an informal way, whenever they need help. Besides this general role, I have been/are more specifically involved in mentoring the following students/researchers:

- From the end of 2010 I am helping Dr. Nicolas Joanning, who is a JSPS postdoctoral researcher at Bioinformatics Center. Dr. Joanning obtained his Ph.D. at Karolinska Institute and is interested on antigenic variation in malaria parasites. My main role as a mentor is to give him advise, and specific help about bioinformatics methods aiming to accomplish his research project.

- From the end of 2010 I am mentoring Dr. Gemma Kirwan, who is a JSPS postdoctoral researcher at Bioinformatics Center, on the project: "Multivariate methods applied to the study of atherosclerosis". Dr. Kirwan got her Ph.D. in Applied Chemistry at the Royal Melbourne Institute of Technology University, Melbourne. Now she is trying to apply multivariate statistical methods to integrate transcriptomics and metabolomics data from an atherosclerosis animal model. Because of her background on chemometrics I am responsible on mentoring her about the particularities and complexities associated with biological cellular and signaling networks, as well on the technical problems associated with microarray analysis.
- In 2010 I supervised Dr. Michihiro Tanaka, who was a Ph.D. student at Bioinformatics Center at that time, for the project: "Transcriptome analysis of the development of insulin resistance". In this project a dataset with transcriptomics and metabolomics data from three different tissues was analyzed using an integrative systems biology approach. The results from this work were presented in domestic conferences and it was used during a presentation at a job interview, that enable him to obtain a postdoctoral research position at CBR (Computational Biology Research Center at Tokyo, Japan).
- In 2009 I supervised Lasse Forlkensen, who is a Ph.D. student at the Center for Molecular Medicine, Karolinska Institute, for the project: "GeneRegionScan: a Bioconductor package for probe level analysis of specific, small regions of the genome". In this project he developed a bioinformatics method to analyze expression microarrays. His approach enables to investigate the effect of gene expression on splicing event as well as SNPs. The results from this interaction where published in the journal Bioinformatics. He is now applying these methods to understand inflammatory mechanisms of cardiovascular disease.

Other merits

Teaching experience

I participated in several coursers aiming to teach programming languages to medical and biological students, both before and after obtaining my doctorate degree.

- *Systems Biology and the Omics Cascade (Course 2143): Introduction to programming Perl*. June 9-13, 2008. Karolinska Institutet, Sweden.
- *Introduction to the use of bioinformatics tools: The Perl Language*. Doctorate course (30 hours), 2005. Universidad Autonoma de Madrid, Spain.
- *Introduction to the use of bioinformatics tools: The Perl Language*. Doctorate course (30 hours), 2004. Universidad Autonoma de Madrid, Spain.
- *Introduction to the use of bioinformatics tools: The Perl Language*. Doctorate course (30 hours), 2003. Universidad Autonoma de Madrid, Spain.
- *Basic Techniques on Biomedical Research: In Situ Hibridization*. Instituto de Investigaciones Biomedicas. November 2000.
- *Introduction to the C Language*. Course of the Biophysics group. Universidad Complutense de Madrid. March 2000.

Summary of computer & bioinformatics skills

OS	UNIX (Linux, Solaris, Irix), MacOSX and Windows XP
Languages	R, Ruby, Perl, MATLAB, C/C++ and Visual Basic
Web	HTML, CSS, SQL, CGI
Databases	MySQL, Postgresql, Access
Bioinformatics	R+Bioconductor, Bioperl, Bioruby, Alignment tools (e.g. ClustalW, MAFFT), Sequence tools (e.g. HMMER, BLAST, MEME), Jalview, TreeDyn, VMD
Other	LaTeX, Office Suite, EndNote, Adobe Illustrator and Photoshop

Summary of laboratory skills

Basic techniques	Design and planification of experiments. RNA extraction and purification from tissues, RNA electrophoresis and transfer techniques, PCR, cloning of fragments into plasmid, Northern blot and laboratory animal care.
Histology	In situ hybridization, Inmunohistochemistry, Nissl.
Microarray	Preparation and hybridization of slides for spotted cDNA microarrays. Scanning and microarray image analysis.

Developed bioinformatics software

- *codelink*: an R package for the analysis of Codelink microarray plaform. Published in Bioconductor in 2006.
- codelink annotation packages, for several array types (mouse, rat, human, whole genome, etc). Updated every six months.
- *GeneRegionScan* a package for the investigation of microarray data and splicing or SNP associated events.
- *seqminer:* a pipeline in Ruby to identify sequences from diverse databases. This method was applied to the varDB project (http://www.vardb.org).
- Several R packages aiming to assist on sequence, microarray and network analysis on the R platform. Not published.

Conferences

† Oral presentation
⋆ Best paper/poster awarded

1. **Diez D**, Wheelock A, Wheelock C, and Goto S. *Identification of pathways mediating glucocorticoid actions in asthmatics*, in ICR Symposium. 2011. Uji, Kyoto, Japan.
2. Hiranuka K, **Diez D**, Joannin N, Kanehisa M, and Goto S. *Clustering analysis of antigenically variable gene families of protozoan parasites*, in ICR Symposium. 2011. Uji, Kyoto, Japan.
3. Kirwan G, **Diez D**, Goto S, Wheelock A, and Wheelock C. *OPLS as a variable selection tool: a comparative study*, in Metabolomics. 2011. Cains, Australia.
4. Hiranuka K, **Diez D**, Goto S, and Kanehisa M. *Detection of Segmental Gene Conversion in Antigenic Variant Multigene Families of Protozoan Parasites*. in IBSB2010: The Tenth International Workshop on Bioinformatics and Systems Biology. 2010. Kyoto, Japan.
5. Grijota-Martinez C, **Diez D**, Escobar GMd, Gil-Ibañez P, Morte B, and Bernal J. *Thyroid hormone transport and action in the fetal rat brain*. in 14th International Thyroid Congress. 2010. Paris, France.
6. Morte B, **Díez D**, Ausó E, Martin-Belinchón M, Grijota-Martínez C, Navarro D, Escobar GMd, Berbel P, and Bernal J. *Papel de la hormona tiroidea materna y fetal en la expresión génica del cerebro fetal*. in SENDIMAD: VIII Congreso de la Sociedad de Endocrinología, Nutrición y Diabetes de la Comunidad de Madrid. 2009. La Granja, Segovia.
7. Morte B, Auso E, **Diez D**, Martín M, Grijota-Martinez C, Navarro D, Morreale G, Berbel P, and Bernal J. *Regulation of gene expresion in the foetal rat brain by thyroid hormone*. in Regulation of gene expresion in the foetal rat brain by thyroid hormone. 2009. Lisboa.
8. ⋆ † Hayes CN, **Diez D**, Joannin N, Kanehisa M, Wahlgren M, Wheelock CE, and Goto S, *Tools for investigating mechanisms of antigenic variation: new extensions to varDB*, in GIW2009: 20th International Conference on Genome Informatics. 2009, Yokohama, Japan.

9. **Diez D**, Hayes N, and Goto S. *Evolution of antigenic variant gene families within Plasmodium species*, in ISMB/ECCB2009: 17th Annual International Conference on Intelligent Systems for Molecular Biology & 8th European Conference on Computational Biology. 2009, Stockholm, Sweden.

10. Hayes CN, **Diez D**, Joannin N, Kanehisa M, Wahlgren M, Wheelock CE, and Goto S. *Using labeling to associate clinical data and sequence variation in varDB*, in JSBi2008: Japanese Society of Bioinformatics Conference. 2008, Osaka, Japan.

11. **Diez D**, Hayes N, Joannin N, Kanehisa M, Wahlgren M, Wheelock CE, and Goto S. *A framework for mining GenBank: Implementation for the varDB pro ject*, in GIW2008: 19th International Conference on Genome Informatics. 2008, Gold Coast, Australia.

12. Hayes CN, **Diez D**, Joannin N, Kanehisa M, Wahlgren M, Wheelock CE, and Goto S. *Improving alignment of hypervariable regions for varDB*, in GIW2008: 19th International Conference on Genome Informatics. 2008. Gold Coast, Australia.

13. **Diez D**, Hayes N, Joannin N, Wheelock C, Wahlgren M, and Goto S. *Alignment of sequences with large hyper-variable regions and different domain architectures*, in KI-Japan seminar. Malaria research: diversity and control. 2008, Stockholm, Sweden.

14. Hayes N, **Diez D**, and Goto S. *An algorithm for identifying distributed sequence motifs* , in APBC2008: The Sixth Asia Pacific Bioinformatics Conference. 2008, Kyoto, Japan.

15. Hayes N, **Diez D**, Kanehisa M, Wahlgren M, Wheelock C, and Goto S. *varDB: a sequence database of protein families involved in antigenic variation in pathogens*, in GIW2007: The Seventeenth International Conference on Genome Informatics 2007, Matrix 30, Biopolis Street, Singapore.

16. † **Diez D**, Hayes N, Kanehisa M, and Goto S. *Alignment of sequences with large hyper-variable regions*, in JSBi2007: Japanese Society of Bioinformatics Conference. 2007, MIRAIKAN, Tokyo Academic Park, Odaiba, Tokyo, Japan.

17. **Diez D**, Ribacke U, Wheelock C, Wahlgren M, and Goto S. *Analysis of the sub-telomeric region of the malaria parasite P. falciparum using aCGH data*, in GIW2006: The Seventeenth International Conference on Genome Informatics 2006, Pacifico Yokohama, Japan.

18. Morte B, **Diez D**, and Bernal J. *The importance of thyroid hormone in fetal brain gene expression*, in 48 Cogress of the Spanish Society of Endocrinology and Nutrition. 2006, Sevilla, Spain.

19. **Diez D,** Pazos F, Valencia A, Bernal J, and Garcia-Ranea JA. *HMM profile to identify new members of the Ras superfamily and its regulators*, in ECCB05: Fourth European Conference in Computational Biology. 2005, Madrid, Spain.

20. **Diez D**, Morte B, Lavado R, Lecea Ld, Rey FEd, Escobar GMd, and Bernal J. *Whole genome analysis of thyroid hormone effect on fetal brain gene expression*, in Functional Genomics and Disease, 2nd ESF Functional Genomics Conference. 2005, Oslo, Norwey.

21. ⋆ Morte B, **Diez D**, Grijota MC, Agretti P, Marco GD, Tonacchera M, Pinchera A, and Bernal J. *Identification of genes sensible to thyroid hormone in the adult brain*, in SEEN: 47 Congreso de la Sociedad Espanola de Endocrinología y Nutrición. 2005, Madrid, Spain.

22. † **Diez D**, Morte B, Lavado R, Lecea Ld, Rey FEd, Escobar GMd, and Bernal J. *Effect of thyroid hormone in fetal brain gene expression*, in Gene Signatures Symposia. 2005, Madrid, Spain.

23. **Diez D**, Morte B, Lavado R, Lecea Ld, Rey FEd, Escobar GMd, and Bernal J. *Influence of thyroid hormone in fetal brain gene expression*, in SENDIMAD: III Congreso de la Sociedad de Endocrinología, Nutrición y Diabetes de la Comunidad de Madrid. 2004, Aranjuez, Spain.

24. **Diez D**, Morte B, Lavado R, Lecea Ld, Rey FEd, Escobar GMd, and Bernal J. *Influence of thyroid hormone in fetal brain gene expression*, in SENC: X Congreso de la Sociedad Española de Neurociencia. 2003, Lleida, Spain.

25. **Diez D**, Morte B, Lavado R, Lecea Ld, Rey FEd, Escobar GMd, and Bernal J. *Effect of thyroid hormone in fetal brain gene expression*, in SEEN: 45 Congreso de la Sociedad Española de Endocrinología y Nutrición. 2003, Caceres, Spain.

26. † **Diez D**, Morte B, Lavado R, Lecea Ld, Rey FEd, Escobar GMd, and Bernal J. *Influence of thyroid hormone in fetal brain gene expression*, in Hormones And Genome: Basic research towards improvement of diagnostic and therapeutic medicine. European Comission, Human Potential Programme, High Level Scientific Confereces, A series of High Level Scientific Conferences. First event: 'Thyroid and Gonadal Hormones in Health and Disease'. 2002, Blaubeuren, Germany.

27. Sagarna JJ, Ayuso L, Morte B, **Diez D**, Peral B, Perez-Aranda A, and Bernal J. *Development of cDNA microarrays to analyze gene expression*, in BioTechnica 2001 International Trade Fair for Biotechnology. 2001, Hannover, Germany.

Other courses and conferences attended

- *Life: Molecular integration & Biological Diversity* 20th IUBMB International Congress of Biochemistry and Molecular Biology and 11th FAOBMB Congress. Kyoto, Japan. June 18-23, 2006.
- Fourth Bioinformatics Conference. Universidad de La Coruna, A Coruna, Spain. September 13-15, 2003.
- *DNA Arrays 2003* CNIO Meeting. Centro Nacional de Investigaciones Oncologicas, Madrid, Spain. May 9, 2003.
- *Introduction to the use and programing in R(S) language* CSIC Bioinformatic Network course. May 2003.
- *Molecular Dynamics* CSIC Bioinformatic Network course. December 9, 2002.
- *Symposium on Bioinformatics And Computational Biology* Third Bioinformatic Conference. Centro de Investigacion del Cancer (CSIC/Universidad de Salamanca), Salamanca, Spain. September 18-21, 2002.
- *Molecular Modelling: Introduction to threading and homology modelling techniques* CSIC Bioinformatic Network course. July 10-11, 2002.
- *Tissue Arrays 2002* CNIO Meeting. Centro Nacional de Investigaciones Oncolgicas, Madrid, Spain. June 12, 2002.
- *Bioinformatics and Computational Biology*, Fundacion BBVA Workshop. April 25-26, 2002.
- *Basic Course in Radioactivity* (10 hours). Service of Radiological Protection, Instituto de Investigaciones Biomdicas. December 2001.
- *IGF/Insulin: From Development to aging* Symposium, Instituto de Investigaciones Biomdicas. July 2001.
- Programing Analyst Course. Kernel School Center (500 hours). February 2000.
- *Salir de la Tierra para conocer nuestros Orgenes*. Fundacion General Unversidad Complutense course (30 hours). July 26-30, 1999.
- Spanish Bioinformatics Network and Glaxo Wellcome Open Meeting. March 24-25, 1999.

Membership

- Member of the Japanese Society of Bioinformatics (JSBi) since 2008.

Contact information for 3 referees

- Dr. Craig Wheelock
  Associate professor at Karolinska Institute
  Email: craig.wheelock@ki.se

  Dr. Craig Wheelock has been a mentor and collaborator for the last five years. He currently holds an assistant professor position at the Division of Physiological Chemistry II Department of Medical Biochemistry and Biophysics, Karolinska Institutet, where he is applying metabolomics-based technologies to probe diseases of inflammation and dyslipidemia. Dr. Wheelock can reference my experience working on cardiovascular disease and inflammatory diseases in general, using systems biology approaches to analyze cellular networks, my experience as independent researcher and scientific management, as well as my experience participating in grant writing. Dr. Wheelock was also involved on the initial varDB project and so he can give references on that project as well. Dr. Wheelock can also reference my experience training Ph.D. candidates at doctoral courses.

- Dr. Susumu Goto
  Associate professor at Kyoto University
  Email: goto@kuicr.kyoto-u.ac.jp Dr. Susumu Goto has been my supervisor and collaborator for the last five years. He is associate professor at Kanehisa Laboratory, Kyoto University, where he leads his own research group. He supervised the development of the varDB project and can give references on that. He can give references as well on my work mentoring students and as independent researcher. Dr. Goto can also reference my work on systems biology approaches to understand inflammatory diseases.

- Dr. Juan Bernal Carrasco
  Research professor at Instituto de Investigaciones Biomedicas “Alberto Sols” CISC/UAM
  Email: jbernal@iib.uam.es

  Dr. Juan Bernal has been my supervisor during my pre-doctoral time, and is now a collaborator on different projects related to the elucidation of thyroid hormone actions in the brain. He can give reference about my experience with experimental approaches, including animal handling and transcriptomics methods. Dr. Bernal can also reference my experience on transcriptional network analysis and my work teaching at doctoral courses for Ph.D. candidates.