LSHG-CT-2005-019015

EURATools

European Rat Tools for Functional Genomics

http://www.euratools.eu

PUBLISHABLE FINAL ACTIVITY REPORT (FAR)

Date of preparation:	**18.11.2010** **Final Version**
Period: covered from:	**Reporting Periods 1 to 4** **01.03.2006 (M1) to** **31.08.2010 (M54)**
Start date of the project: Duration of the project:	01.03.2006 54 months
Project coordinator: Project coordinator organisation:	Prof. Timothy J. Aitman Medical Research Council Clinical Sciences Centre

Table of Contents

Involved Institutions and Principal Investigators

Nr.	*Institution*	*Acronym*	*Principal Investigators*
1	Medical Research Council, Clinical Sciences Centre (CSC), London	MRC	Timothy Aitman Laurence Game Stuart Cook
2	Max-Delbrück-Center for Molecular Medicine (MDC), Berlin	MDC	Norbert Hübner Michael Bader
3	Colleges of Medicine, and Science and Engineering, University of Edinburgh	UEDIN	John Mullins Ian Wilmut
4	Institute of Physiology, Czech Academy of Sciences, Prague	CAS	Michal Pravenec Vladimír Landa Vladimír Křen
5	European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton	EMBL-EBI	Ewan Birney Xosé Fernández
6	Division of Cardiovascular and Medical Sciences, University of Glasgow	UGL	Anna Dominiczak
7*	The department of Medical Inflammation Research, Lund University	ULUND	*Rikard Holmdahl
8	Institute of Molecular Biology and Medicine, Université Libre de Bruxelles	ULB	Claude Szpirer
9	Commissariat à l'Energie Atomique, Evry	CEA	Mark Lathrop Ivo Gut Jean Weissenbach
10	The Wellcome Trust Centre for Human Genetics, University of Oxford	UOXF	Dominique Gauguier Jonathan Flint
11	Medical Psychology Unit, School of Medicine, Autonomous University of Barcelona	UAB	Alberto Fernandez-Teruel
12	Karolinska Institutet, Center for Molecular Medicine, Stockholm	KI	Tomas Olsson **Rikard Holmdahl
13	Institute of Zoology, Chinese Academy of Sciences, Beijing	IOZ CAS	Qi Zhou
14	Max-Planck-Institute for Molecular Genetics, Berlin	MPIMG	Richard Reinhardt
15	Unité de Biologie du Développement, National Institute for Agricultural Research, Paris/Jouy-en-Josas	INRA	Jean-Paul Renard, Nathalie Beaujean
16	genOway SA, Lyon	genOway	Alexandre Fraichard, Jean Cozzi
17	CXR Biosciences Limited, Dundee	CXR	Roland Wolf, Simon Plummer

** until February 2008, ** from March 2008*

(1) PROJECT EXECUTION

Introduction

The EURATools project started 2006 to develop reagents and tools for high-throughput analysis of the rat genome that will accelerate progress in rat genetics; to create a research environment conducive to rapid identification of genes and pathways underlying complex rat disease phenotypes; to integrate new and existing data, and to develop new tools that will increase the utility and accessibility of rat genome databases.

The consortium was continuously supported by an external Scientific Advisory Board (Alan Archibald (Roslin Institute, Edinburgh), Allen Cowley (Medical College of Wisconsin), Philip Iannaccone (University of Chicago) and Klaus Lindpaintner (Roche; Chair)) and the overall success of the project is well judged from their final review:

> "There was unanimous acclaim that the program had been run in an exemplary fashion, and that it represents a true showcase-piece for similar consortia. In particular, the SAB commented favorably on:
>
> - The skill with which the PI's were able to draw in young researchers to take ownership of (parts of) the program, extending confidence all the way towards allowing them to organize their own young investigator (YEIS) meetings, strongly promoting scientist exchange through travelling fellowship programs, and successfully instilling a sense of European identity as a scientific community in them;
> - The flexibility, pragmatism, and sense of accountability with which the group has embraced technological developments and taken advantage of them by applying them as coursecorrections to the original research plans, leveraging emerging convergences of technologies and thus ensuring and enhancing the program's success;
> - The degree to which SMEs were involved in several subprojects in truly synergistic and mutually contributing and beneficial fashion;
> - The overall tangible sense of enthusiasm and level of truly cooperative work the consortium was able to cultivate, and which resulted in a truly European collaboration in the best of the FP6- funders intentions;
> - An impressive publication record; and last, but certainly not least
> - The accomplishment of significantly raising the "stature" of the rat as a fully viable model organism in the genomic age."
>
> *(from the Final Annual Review, SAB Members' Comments and Report).*

Overall, the EURATools project can claim the successful positional cloning of more than five rat complex disease genes, three of which were reported in a special focus issue of *Nature Genetics* in 2008, with the imminent delivery of several additional disease genes in the closing phases of the Project. The funding has led to a substantial growth in genome resources, including new cDNA sequences, millions of new sequence variants and platforms for high throughput genotyping, the genome sequence of the spontaneously hypertensive rat and, in the last year, several other rat strains, and an outstanding set of genome resources co-ordinated by the European Bioinformatics Institute. The training given in informatics has meant that rat genetics is now carried out hand-in-hand with informatics and expert use of genome resources. This is a major change from the start of the Project, when rat genome resources were scarce and use of informatics was the exception rather than the rule.

A further major achievement has been the discovery of a gene network for inflammation in multiple rat tissues that is conserved in humans and that has been found to contribute to the pathogenesis of the human autoimmune disease, type 1 diabetes. This highly integrated project was achieved by combined work of three partner institutes (London, Berlin, Prague) through integrated use of gene expression, biological resources, genome tools, informatics, and Bayesian statistics.

Further details are outlined for each project Activity below.

Activity 1: Genome Tools – Annotation and Functional Anaysis for Genes and Complex Disorders

Activity 1 Workplan

WP 1.1 Transcriptome of full length cDNAs

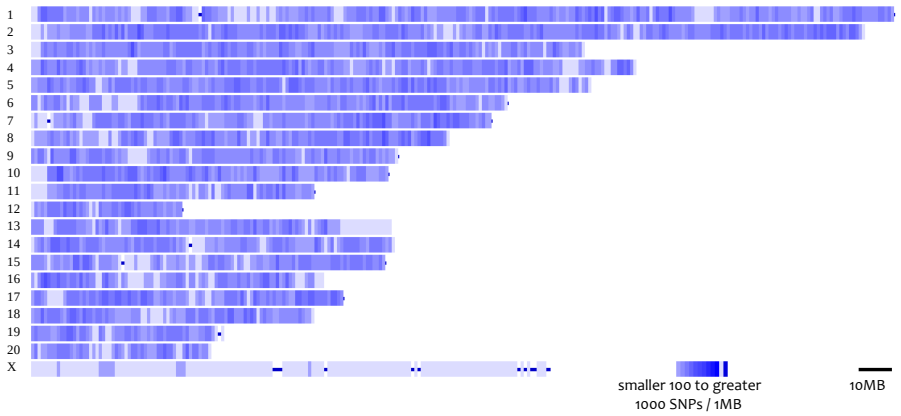
Task 1.1-1	Construction of full length cDNA libraries
Task 1.1-2	Sequencing of libraries
Task 1.1-3	Initial sequence analysis full length cDNA
Task 1.1-4	Finishing of non redundant FL cDNAs
Task 1.1-5	Annotation, display and integration with ENSEMBL

WP 1.2 SNPs and haplotypes for assisting complex trait analysis

Task 1.2.1	Identification of common and divergent haplotypes across 50 strains
Task 1.2.2	High density genotyping of SNPs in heterogeneous stock
Task 1.2-3	Development and evaluation of dense SNP genotyping tool

It was the first goal of this Activity to **improve the diversity and annotation of the rat genome (Workpackage 1)**. To achieve this, full length cDNA libraries from 10 tissues (brain, foetus, liver, skeletal muscle, spleen, aorta, heart, kidneys embryo; planned were 8) were constructed and sequenced through more than 300.000 full-length cDNA clones with a variety of modern, state-of-the-art sequencing techniques.
At a time when the project was planned, the number of rat cDNAs in various databases including Ensembl was quite low (<8000), especially when compared to the number of human or even mouse cDNAs listed in the databanks. By now, the overall number of rat cDNAs has increased greatly (>15000) and yet EURATools investigators made a major contribution to overall cDNAs sequenced and even exceeded the number of **"new" cDNAs** depicted on top of that (> 9500, instead of 2500-4000 expected in the beginning).

The second major goal of this Activity was to **facilitate the identification of complex trait disease genes in the rat (Workpackage 1.2)**. We have developed a first **high throughput genotyping tool** for the rat. Using this tool in the first phase of the project, 20K SNPs were genotyped in a large number of commonly used rat strains and crosses to construct **ancestral haplotype maps**. Since there is little insight into genetic variation among inbred rats coming from the reference genome sequence, these resources enabled us (and others) to reveal deeper insight into genetic variation among inbred rats, providing the tools to study functional involvement of genes in patho-physiological processes in this widely used model organism. **Genome wide association studies in heterogeneous stock or outbred populations** require higher density genotyping platforms. We thus set up to develop a new SNP array allowing to genotype at least 500K rat SNPs. As we wanted to cover the spectrum of variation over many different inbred rat strains, SNP selection was performed using *in silico* predicted SNPs from strains where sequence information was available and which belong to different branches of phylogeny. For the final array design we have selected 550K high quality SNPs and 1 mio CNV probes. This **new array** will significantly enhance the toolbox for genetic and genomic investigations in the rat and will facilitate studies interrogating the characterization of genetic variation, copy number variation, allele specific gene expression and DNA methylation throughout the rat genome.



Workpackage 1.2: Distribution of SNPs on the new high density SNP array.

All data has been or is being deposited in the ENSEMBL databases and also contributes to the official genome assembly at Baylor College of Medicine.

Most important publications resulting from Activity 1

- STAR consortium (2008): **SNP and haplotype mapping for genetic analysis in the Rat**. *Nat Genet*, 40(5), 560-566.

Activity 2: Nuclear Transfer in the Rat

Activity 2 Workplan

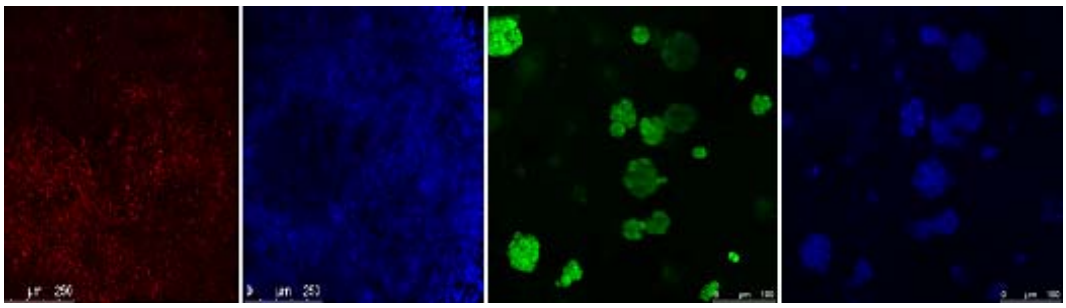
WP 2.1	***Characterisation of early rat embryos***
Task 2.1-1	*Biochemical characterisation of the rat embryo before/after cloning*
Task 2.1-2	*Epigenetic characterisation of embryos before/after cloning*
WP 2.2	***Preparation/optimisation of nuclear donor cells***
Task 2.2-1	*Selection and provision of ES-like cells for testing as nuclear donors*
Task 2.2-2	*Growth of rat fibroblasts and optimisation of long-term survival*
WP 2.3	***Screening potential nucleus donor cells***
Task 2.3-1	*Screening of potential nucleus donor cells - method I*
Task 2.3-2	*Screening of potential nucleus donor cells – method II*
Task 2.3-3	*Interspecies studies of rat nuclear transfer*

The overall aim of this Activity was to optimise and facilitate **practical and efficient germline modification procedures in the rat** by refinement and adaptation of existing but highly inefficient nuclear transfer protocols.

To **learn more about the early events associated with remodelling** of somatic or embryonic nuclei when introduced into enucleated oocytes **and about reprogramming** processes under different donor conditions, rat embryos have been characterised biochemically and epigenetically (**Workpackage 2.1**). No data was available for rat embryos concerning a major cell cycle event: the S-phase or replication phase. We therefore tested different protocols required to detect this event by immunostaining in rat embryos and an appropriate protocol to **follow-up replication in rat embryos** has been set up. We successfully established the time-window of each replication phase at these stages and detected four major phases during the S-phase: (A) the first pattern shows numerous and well-dispersed replication foci all over the nucleoplasm, (B) replication foci are smaller and more dense, as well as more concentrated in the central region of the pronuclei,

(C) this pattern is characterized by peri-nucleolar fluorescent rings and (D) when only small peripheral replication sites remain.

In order to **optimise nuclear donor cells for nuclear transfer in the rat** it was the objective of this Activity to generate ES cell lines with desirable features as nuclear transfer donors and to assess ES-like cell lines (ExS cells) in long term culture (**Workpackage 2.2**). We characterized two cell lines (C5 and B10) which were obtained from rat blastocysts in medium with serum and LIF. In the B10 cell line we found the expression of genes known to be expressed in trophoblast, Cdx-2, cytokeratin-7, and Hand-1. Moreover, B10 cells invaded the trophectodermal layer upon injection into rat blastocysts. In contrast to mouse Trophoblast Stem (TS) cells proliferation of B10 cells occurred independently of FGF4. Cells of the C5 line expressed traditional markers of extraembryonic endoderm (XEN) cells, in particular, GATA-4, but also the pluripotency markers SSEA-1 and Oct-4. C5 cell proliferation exhibited dependence on LIF, which is not known to be required by mouse XEN cells. Our results confirm and extend findings about **differences between blastocyst-derived cell lines of rat and mice**. Our data show the important roles of feeder cells, LIF and inhibitors for the maintainance of cell pluripotency *in vitro* which led to new cell lines with typical ES cell morphology. These cells show a positive immunostaining for the ES-cell specific markers Oct4 and Nanog. Furthermore, we have established a simplified ES culture protocol, using chemical inhibition of only the MEK pathway, which produces rat ES cells that can be transmitted through the germ line. We have also generated *hprt* knock-out rat ES cells using lines derived from inbred (Fischer 344) and outbred (Sprague Dawley) strains of rats.



Oct **Hoechst** **Nanog** **Hoechst**

Workpackage 2.2: Characterization of rat ESC by staining with antibodies against Oct4 and Nanog, both commonly used ES-cell markers. The nuclei were stained with the DNA-specific dye Hoechst.

Finally, to actually **establish nuclear transfer in the rat**, ES like-cells with improved developmental ability were selected with the goal to produce viable fetuses. For this, also embryo handling and efficient rat embryo culture system *in vitro* were established and optimised (Workpackage 2.3). As of February 2009, there have been no reports of successful and routine nuclear transfer in the rat and this technology remains highly specialist and non-generic in nature. However, there have been **several major advances in the community** in parallel, including the reports of rat stem cells that have been demonstrated to be competent for passage through the germline (Buehr, M. et al. Cell 135, 1287-1298 and Li, P. et al. Cell 135, 1299-1310 (2008)). In the light of those breakthrough achieved, we reconsidered the original strategy and focussed efforts towards alternative sources of stem cells since.

A further goal was the production of cloned rat blastocyst *in vitro* by interspecies or serial nuclear transfer which may constitute alternative nuclear transfer procedures to **improve the rat cloning efficiency**. Interspecies nuclear transfer and serial nuclear transfer procedures were designed and setup. New new rat cell lines were isolated using the 2i methodology, germ-line competent rat embryonic stem cell lines were generated, G418-resistant clones were transfected and selected and homologous recombination in rat ES cells was successfully targeted.

Most important publications resulting from Activity 2

- Popova E, Bader M, Krivokharchenko A (2008): **Efficient production of nuclear transferred rat embryos by modified methods of reconstruction.** Mol Reprod Dev. 76, 208-216.
- Chuykin I., Lapidus I., Popova E., Vilianovich L., Mosienko V., Alenina N., Binas B., Chai G., Bader M., Krivokharchenko A. **Characterization of trophoblast and extraembryonic endoderm cell lineages derived from rat preimplantation embryos**. *PLoS ONE,* 5(3):e9794.

Activity 3: Rat Biological Resources and Toxigenetics

Activity 3 Workplan

WP 3.1	***Provision of a repository of live and frozen rat strains***
Task 3.1-1	*Breeding, phenotyping and distribution of rat tissues and animals*
Task 3.1-2	*Cryopreservation of rat embryos*
Task 3.1-3	*Revitalisation of rat embryos*
WP 3.2	***Toxicogenetics and pharmacogenetics***
Task 3.2-1	*Drug toxicity in SHR and BN rats*
Task 3.2-2	*Mapping the genetic determinants of drug toxicity*
Task 3.2-3	*Mapping pharmacogenetic responses for cardio-metabolic traits*
WP 3.3	***Rat heterogeneous stock (HS)***
Task 3.3-1	*Obtain phenotype data for HS animals*
Task 3.3-2	*Genotype and analyze HS animals*

As a central resource, the **EURATools Rat Repository** was established in Prague and provided the rat animals and tissues for many experiemtns throughout the consortium (**Workpackage 3.1**). More than 500 embryos were **cryopreserved** from multiple inbred, congenic, transgenic or mutant lines including all strains of the BXH/HXB recombinant inbred (RI) panel. The effectiveness of revitalisation of strains from frozen embryos, though not required/requested during the project runtime, was demonstrated to be about 25% in embryos that were frozen over 10 years and about 47% in embryos that were frozen less than 10 years. To increase the value of existing and new biological resources, new phenotypes were generated and multiple physiological phenotypes were mapped in heterogeneous outbred rat stocks, permitting fine mapping and inter-phenotype correlations across disease areas. This includes lipid and carbohydrate metabolism, blood pressure and heart mass. Combining linkage and expression profile analyses several pathophysiological QTL were identified at the the molecular level as quantitative trait genes (QTG): mutated *Srebf1*, *Cd36*, and *Ogn* as genetic determinants predisposing to hepatic steatosis, hypertension and cardiac hypertrophy, respectively. Also, new congenic and conplastic rat lines were produced and characterised, including the SHR-mt^{BN} conplastic strain that has been used to provide evidence for the role of mitochondrial genome variation in the pathogenesis of several risk factors for type 2 diabetes (ref. 17).

In order to **map loci which regulate drug-induced hepato- and nephrotoxicity** and to identify candidate genes, also the BXH/HXB RI strains were used (**Workpackage 3.2**). An extensive litereature study for 21 compounds from 5 drug classes (glitazones, statins, COX2 inhibitors/NSAIDS, anti-cancer drugs and PPAR gamma agonists) and from drug compounds that cause rat toxicity with a human counterpart revealed 10 candidates for toxicity studies in the parental BN and SHR strains and 8 of them were actually tested. 7 of those 8 showed significant differences between the progenitor rat strains (BN and SHR) in response at the level of plasma toxicity biomarkers and the full mapping of underlying susceptibility genes in BXH/HXB RI strains was ultimately done for two of them, paracetamol/acetaminophen and valproic acid. Mapping results revealed a suggestive **QTL associated with paracetamol toxicity** on rat chromosome 2

in the vicinity of *Pld1* candidate gene. Another significant quantitative trait locus (QTL) on chromosome 2 (LRS=17.6) associated with **valproic acid (VPA) toxicity** (designated Vpa1) was also found. The Vpa1 QTL on chromosome 2 is colocalized with *Cth* (cystathionase) candidate gene that plays an important role in glutathione (GSH) metabolism. Testing VPA associated hepatotoxicity in the BN.SHR-chr.2 congenic strain confirmed the presence of Vpa1 QTL. In addition, the BN.SHR-chr.2 congenic strain versus BN showed reduced liver concentrations of GSH, increased levels of lipoperoxidation products, conjugated dienes and thiobarbituric acid-reactive substances (TBARS) as well as increased activity of anti-oxidant enzymes.
Also identified within this workpackage was a highly significant QTL associated with with **blood pressure response to captropril** treatment on chromosome 15.

A panel of heterogeneous stock rats was established in Barcelona to achieve the **high resolution genetic mapping of the determinants of behavioural, haematological, cardiovascular, diabetes and immunological phenotypes (Workpackage 3.3)**. It had been shown in the mouse that mapping of QTLs in a heterogeneous stock (HS) gives sub-centimorgan QTL localisation and comparable resolution was demonstrated in the rat. In a formidable collaborative effort more than 2000 HS rats were phentotyped accordingly and the data is being analysed together with high density SNP data of the rats contributed by WP 1.2.

Most important publications resulting from Activity 3

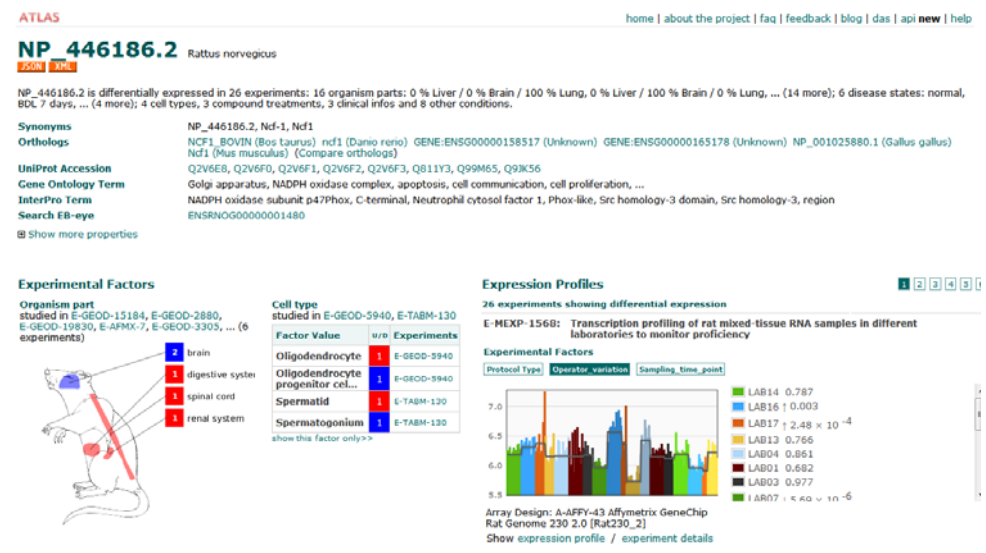
- Pravenec M, Hyakukoku M, Houstek J, Zidek V, Landa V, Mlejnek P, Miksik I, Dudová-Mothejzikova K, Pecina P, Vrbacký M, Drahota Z, Vojtiskova A, Mracek T, Kazdova L, Oliyarnyk O, Wang J, Ho C, Qi N, Sugimoto K, Kurtz TW (2007): **Direct linkage of mitochondrial genome variation to risk factors for type 2 diabetes in conplastic strains.** *Genome Res*, 17(9), 1319-26.
- Pravenec M, P.C.Churchill PC, M.C.Churchill MC, O.Viklicky O, L.Kazdova L, T.J.Aitman TA, E.Petretto E, Hubner N, Wallace CA, Zimdahl H, Zidek V, Landa V, Dunbar J, Bidani A, K.Griffin K, Qi N, Maxova M, Kren V, Mlejnek P, Wang J, Kurtz TW (2008): **Identification of renal /Cd36 /as a determinant of blood pressure and risk for hypertension.** *Nat Genet*, 40(8):952-4.
- Pravenec M, Kazdova L, Landa V, Zidek V, Mlejnek P, Simakova M, Jansa P, Forejt J, Kren V, Krenova D, Qi N, Wang J, Chan D, Aitman TJ, Kurtz TW (2008): **Identification of mutated Srebf1 as a QTL influencing risk for hepatic steatosis in the spontaneously hypertensive rat.** *Hypertension* 51,148-153.
- Johnson MD, He L, Herman D, Wakimoto H, Wallace CA, Zidek V, Mlejnek P, Musilova A, Simakova M, Vorlicek J, Kren V, Viklicky O, Qi NR, Wang J, Seidman CE, Seidman J, Kurtz TW, Aitman TJ, Pravenec M (2009): **Dissection of chromosome 18 blood pressure and salt-sensitivity quantitative trait loci in the spontaneously hypertensive rat.** *Hypertension*, 54(3):639-45
- Liska F, Snajdr P, Sedová L, Seda O, Chylíková B, Slámová P, Krejcí E, Sedmera D, Grim M, Krenová D, Kren V. (2009): **Deletion of a conserved noncoding sequence in Plzf intron leads to Plzf down-regulation in limb bud and polydactyly in the rat.** *Dev Dyn.* 238(3):673-84.
- Johannesson M, Lopez-Aumatell R, Stridh P, Diez M, Tuncel J, Blazquez G, Martinez E, Canete A, Vicens-Costa E, Graham D, Copley RP, Hernandez-Pliego P, Beyeen AD, Öckinger J, Fernandez C, Gulko PS, Brenner M, Tobeña A, Guitart-Masip M, Gimenez-Llort L, Dominiczak A, Holmdahl R, Gauguier D, Olsson T, Mott R, Valdar W, Redei E, Fernandez-Teruel A, Flint F (2009): **A resource for the simultaneous high-resolution mapping of multiple quantitative trait loci in rats: the NIH heterogeneous stock.** *Genome Research*, 19(1):150-8.

Activity 4: Rat Genome Informatics

Activity 4 Workplan

WP 4.1	***Development of data mining resources***
Task 4.1-1	*Compendium of rat gene expression data*
WP 4.2	***User survey, training and use-case development***
Task 4.2-1	*Yearly European workshops*
Task 4.2-2	*Structured outreach sessions*

It was the objective of this Activity to **integrate new and existing genome data and to develop data mining resources** by co-ordinating genome sequence, gene models, mapping resources and expression data (**Workpackage 4.1**).
The scope of the compendium has been enhanced with the launch of The **ArrayExpress Atlas of Gene Expression** at the European Bioinformatics Institute (http://www.ebi.ac.uk/gxa/). Now we provide comprehensive gene expression annotation across multiple experiments (133 studies, 4858 assays) and different biological conditions (with a meta-analytical ranking them). Condition-specific queries of gene expression across multiple data sets are possible. Users can query for a gene or a set of genes by name, synonym, Ensembl identifier, GO term or, alternatively, for a biological sample property or condition, e.g. tissue type, disease name, developmental stage, compound name or identifier. Queries for both genes and conditions are also possible. The Atlas takes data directly from the Array Express Archive of Functional Genomics Experiments including publicly available datasets from MiMiR. These datasets are curated, mapped to the latest genome builds and the experimental conditions are systematized and mapped to the Experimental Factor Ontology. Statistical computations are performed, providing P-values linking each gene to each experimental condition in the study.



Workpackage 4.1: Screenshot from the Gene Expression Atlas displaying expression level of Ncf1 gene in rat from 26 different experiments

In order to **advance rat genome informatics and to integrate world wide efforts**, Rat Genome Coordination meetings were organised (Workpackage 4.2). Major contributors to rat bioninformatics include representatives of Baylor College of Medicine (Human Genome Sequencing Center, where the reference genome sequence is located; George Weinstock and Richard Gibbs), NCBI (Kim Pruit), the three main genome browsers (Ensembl, UCSC) the Rat Genome Database (MCW Miwlaukee, Howard Jacob, Mary Shimoyama, Liz Worthey) and EURATools researchers. EURATools is now perceived as the European voice of the rat community! In addition, **use-case interactive courses for bench scientists** to **disseminate**

expertise in genome data mining and use of genome query tools with special emphasis on the rat resources in Ensembl took place on eight occasions and involved both basic (single day) and advanced (multiple day) workshops, always with very positive feedback from participants.

Most important publication resulting from Activity 4

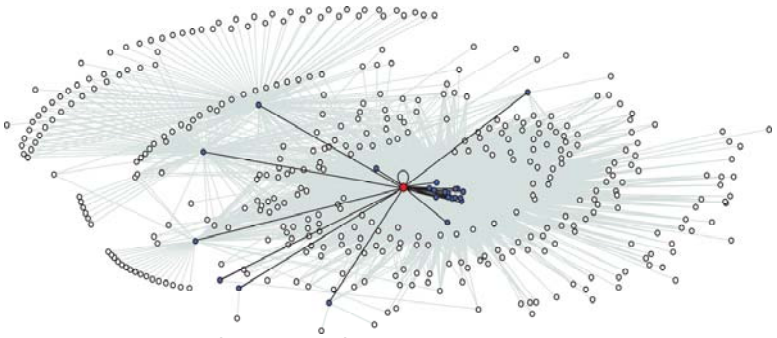
- Twigger SN, Pruitt KD, Fernandez-Suarez XM, Karolchik D, Worley K, Maglott D, Brown G, Weinstock G, Gibbs R, Kent J, Birney E, Jacob HJ (2008): **What everybody should know about the Rat Genome and its online resources.** Nat Genet, 40(5), 523-527

Activity 5: Gene Expression Analysis as a Tool for Gene Identification

Activity 5 Workplan

WP 5.1	***Gene expression in congenic strains***
Task 5.1-1	*Gene identification using minimal congenic strain*
WP 5.2	***Gene expression in recombinant inbred strains***
Task 5.2-1	*Collection of tissues*
Task 5.2-2	*Expression profiling and linkage analysis*
Task 5.2-3	*Expression data extraction, normalisation, and analysis*
Task 5.2-4	*Comparison of genome wide gene expression data from different tissues*
WP 5.3	***Microarray data management***
Task 5.3-1	*Data warehousing*
Task 5.3-2	*Data mining and statistical analysis*
Task 5.3-3	*Bioinformatics*
WP 5.4	***Proteomics***
Task 5.4-1	*Differential in-gel electrophoresis (DiGE) analysis of tissues of congenic rat strains*
Task 5.4-2	*Proteomic analysis of protein complexes formed by QTL gene products*

The first objective of this Activity was to assist in **prioritising candidate QTL genes** within narrow or "minimal" congenic intervals (**Workpackage 5.1**). Six microarray and three supporting sequence capture experiments were selected following consortium internal calls and include the search for genes involved in macrophage activation, glomerulonephritis, right ventricular hypertrophy, malignant hypertension, cardiac remodelling, MOG induced experimental autoimmune encephalomyelitis, pristane-induced arthritis, T cell selection and MHC class I expression. The identified candidate genes are now in the laboratory **validation** phase (see also Activity 6). The generated data has helped to identify angiotensin-converting enzyme (ACE) as the modifier of hypertension-induced tissue microvascular injury in a rat model of hypertension (ref Liu X et al, 2009, Angiotensin-converting enzyme is a modifier of hypertensive end organ damage, J Biol Chem, 284:15564-72). In addition, the entire **SHR genome was sequenced** using Illumina short read technology, the first disease model to have been sequenced by short-read technology, giving an overall10.7 fold coverage revealing major coding sequence mutations in over 600 SHR genes. These new sequence and transcript data are being integrated with results from Activities 1 and 4 and used to accelerate gene discovery (Activity 6).



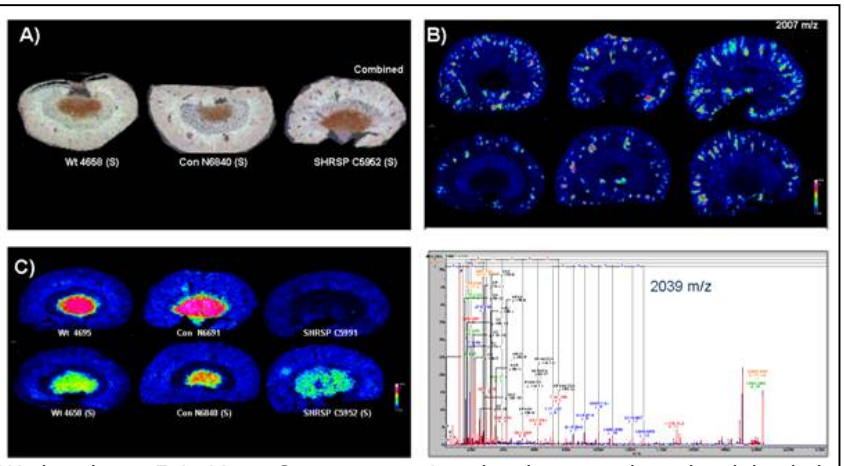
Workpackage 5.2: Irf-7 driven inflammatory network *(IDIN),* the expanded *IDIN:* Nodes represent genes; the node representing Irf7 is coloured red and its predicted targets are coloured blue. Edges connect genes that are either predicted Irf7 targets (black) or show significant Pearson correlation (FDR,0.1%) to one of the predicted targets (grey).

In order to **map the genetic determinants of gene expression** in SHR rats we used expression profiling from several tissues derived

from the BXH/HXB panel of RI strains and integrated this data into a functional genomic approach, in the context of the interplay of the whole organism (**Workpackage 5.2**). In addition, we apply a method for the detection of genotype dependent co-expression genome wide and are developing the construction of a tissue-specific transcriptional network. We have implemented an **analysis pipeline** which can be used to perform every step starting from RMA normalized expression data: filtering, network reconstruction, module discovery by clustering, functional enrichment analysis and genetic mapping of transcriptional modules to the genome. One outstanding result was the identification of a conserved trans-acting regulatory locus that underlies an inflammatory gene expression network in macrophages and also confers susceptibility in humans to autoimmune type 1 diabetes (see figure above).

Both workpackages above were accompanied by the development and use of sophisticated tools for microarray data warehousing, analysis, visualising and mining to **analyse massive data sets** across tissues, time points, sex, cell type, organ, and disease phenotype (**Workpackage 5.3**). Data mining and bioinformatics initiatives have been integrated across partner institutes for biological discovery.

In order to **monitor protein expression, functional modifications and interactions** at the protein level, the proteomic approach, notably 2D-gel electrophoresis and mass spectrometry (MS), complements the microarray-based expression profiling at the transcript level (**Workpackage 5.4**). 19 unique proteins were identified in kidney between 2k and WKY compared to SHRSP, with further 40 proteins identified between WKY and SHRSP. In heart, 81 spots have been selected for identification by mass spectrometry. The peptide pattern of urine collected from SP.WKYGla2k congenic strains revealed 4 targets for MS-MS identification and subsequent confirmation.



Workpackage 5.4: Mass Spectrometry Imaging in non-salt and salt-loaded strains. A) Combined images showing all mass filters giving an overview of the distribution of markers. B) Single mass filter selected (2007 m/z) using heat-map to show regions of highest intensity. Mass seen to be distributed consistently across all kidneys. C) Single mass filter selected (2039 m/z) using heat-map to show regions of highest intensity. Mass seen to vary significantly between SHRSP kidney and that of WKY and 2k congenic strains. Salt treatment also seen to affect mass intensity. Spectra – annotated ms/ms spectra collected directly off tissue (without enzyme digestion) identifies the marker 2039 m/z as histone H1.

Most important publications resulting from Activity 5

- Heinig M, Petretto E, Wallace C, Bottolo L, Rotival M, Lu H, Li Y, Sarwar R, Langley SR, Bauerfeind A, Hummel O, Lee YA, Paskas S, Rintisch C, Saar K, Cooper J, Buchan R, Gray EE, Cyster JG; Cardiogenics Consortium, Erdmann J, Hengstenberg C, Maouche S, Ouwehand WH, Rice CM, Samani NJ, Schunkert H, Goodall AH, Schulz H, Roider HG, Vingron M, Blankenberg S, Münzel T, Zeller T, Szymczak S, Ziegler A, Tiret L, Smyth DJ, Pravenec M, Aitman TJ, Cambien F, Clayton D, Todd JA, Hubner N, Cook SA (2010): **A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk.** *Nature*. 467(7314):460-4.

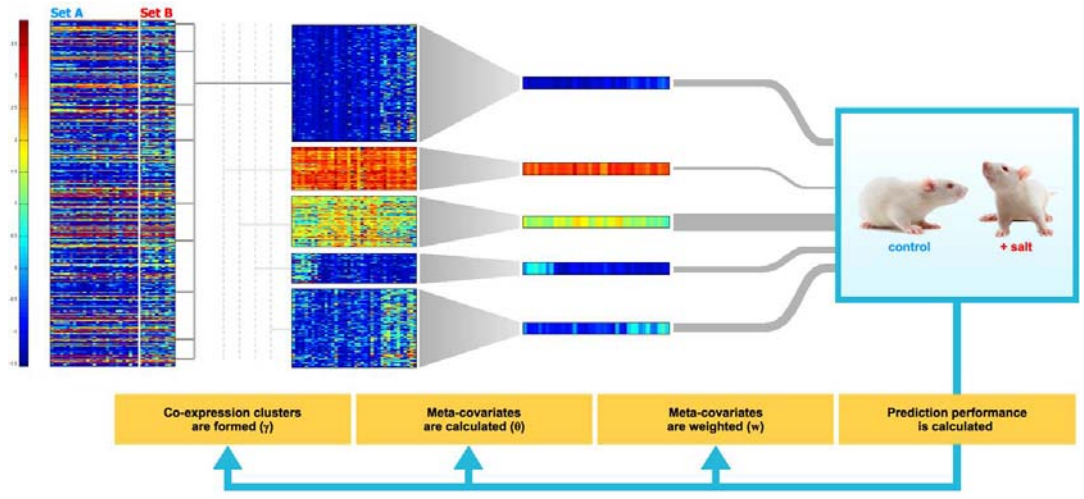
- Monti J, Fischer J, Paskas S, Heinig M, Schulz H, Gösele C, Heuser A, Fischer R, Schmidt C, Schirdewan A, Gross V, Hummel O, Maatz H, Patone G, Saar K, Vingron M, Weldon SM, Hammock BD, Rohde K, Dietz R, Cook SA, Schunck W-H, Luft FC, Hubner N (2008): **Soluble epoxide hydrolase (Ephx2) is a susceptibility gene for heart failure in a rat model of human disease,** *Nat Genet*, 40(5), 529-537

Activity 6: Rat Gene Discovery and Application of the Rat Model to Human Disease

Activity 6 Workplan

WP 6.1	***Cardiovascular diseases***
Task 6.1-1	*High fidelity cardiovascular phenotyping*
Task 6.1-2	*Further development and selection of minimal congenic strains*
Task 6.1-3	*Transcriptomics analysis of minimal congenic strains*
Task 6.1-4	*Cloning by position and functional analysis*
WP 6.2	***Inflammatory diseases***
Task 6.2-1	*Establish a phenotyping platform*
Task 6.2-2	*Positional cloning*
Task 6.2-3	*Genetic interactions*
Task 6.2-4	*Animal models for drug testing*

The overall objective of this Activity is the **in depth functional and interaction analysis** and **transfer of discoveries from rat to human** via the **identification and positional cloning of new genes** and their associated pathways **involved in cardiovascular diseases (Workpackage 6.1**, especially hypertension) **and inflammatory diseases (Workpackage 6.2**, especially rheumatoid arthritis, multiple sclerosis, autoimmune glomerulonephritis and bronchial asthma). Available consortium expertise, methodology and data were used and further developed and include phenotyping, refinement of minimal congenic strain models, comparative genome analysis, recombination, high-throughput sequencing and disease testing, expression analysis, fine mapping in hetergoeanous stock rats and proteomics. Main achievements from **Workpage 6.1** (cardiovascular diseases) include the generation of SHRSP derived congenic substrains through the dissection of congenic intervals on rat chromosome 2, 3 and 14 with additional phenotypic characterisation of these strains including



Workpackage 6.1: The meta-covariate method. Expression data are used to form clusters of probes (clustering is represented by the DxK matrix of responsibilities γ). N-dimensional meta-covariates (θk) are calculated from the clusters are used to make predictions in a probit regression model (with regression coefficients wk). The novelty of our method is highlighted in turquoise: the prediction performance is used to update γ, θk, and wk, thereby iteratively improving the clustering structure and the prediction performance.

kidney and cardiac function. Two significantly differentially expressed genes mapping to chromosome 14 ***(Cxcl 13*** (chemokine ligand 13) and ***Abcg3*** (ATP binding cassette G3)) were identified by analysis of the SP.WKYGla14a heart time series experiment via Affymetrix exon chip profiling. A novel statistical method (meta-covariate analysis) has also been used to analyse Affymetrix microarray gene expression profiling from the SP.WKYGla2a renal, salt/no salt experiment to identify genes and pathophysiological pathways involved in salt-sensitive hypertension. Functional and canonical pathway analysis of the most significant cluster implicated transcriptional activation and circadian rhythm signalling. In addition, SHRSP ***Gstm1*** transgenic rats have been generated, which demonstrate significantly reduced systolic blood pressure. These transgenic rats will allow investigation of the functional relevance of ***Gstm1*** in the development of hypertension. Also, pathophysiological and transcriptomic landscapes of diabetes and obesity spontaneous susceptibility in congenic strains of the GK rat. Diabetes and obesity positional and functional candidate genes were identified through pathophysiological and transcriptomic screening of congenic strains of the GK rat.

In **Workpackage 6.2** (inflammatory diseases) genome sequencing of the strains with importance for inflammatory disorders (DA/Harlan, DA/Zfv, PVG and E3/Rhd) has been carried out resolving several QTLs. The Pia7 congenic has been minimized down to the APLEC gene cluster and by negative exclusioon a new gene has bee positonal identified, DCAR1. It has been shown that this gene not only controls arthritis but also T cell activation. Interaction between several QTLs and the Ncf1 gene has been demonstrated and the role of Ncf1 has been identified in models of Guiliian Barre. The cusative SNP in the Ncf1 gene has been positoned by both new congneics and in vitro transfections. There has been further minimization of a congenic fragment with a spontaneous mutation in chromosome 9. The region has been mapped with BAC clones and genome sequenced. Further dissection of the MHC region has been carried out and three minimized congenics identified. We have succeeded to idnetifed MHC class II B as a cuasative gene and the fucntional role of this is currently under study. Thi is certianly a wellknown gene in autoimmunity but the psotioned of the gene in a disease model that is not depdent ona specifc antigen is new and is of high similarity to the human disaese. In addition, positional cloning ofVav1 has been carried out in the rat and translationed studies have identified a functional role of the human homologue. Exon and transcriptome profiling of EAE has also been carried out in the rat. A chemokine gene cluster was identified and associated to human disease. Gene regions regulation the microglial response to injury have been described in a BN x L cross. Also, a narrow QTL regulating TNF production in the rat has been identified. Finally, in a herpes encephalitis rat model we are down to a three gene segment with a strong positional candidate.

Most important publications resulting from Activity 6

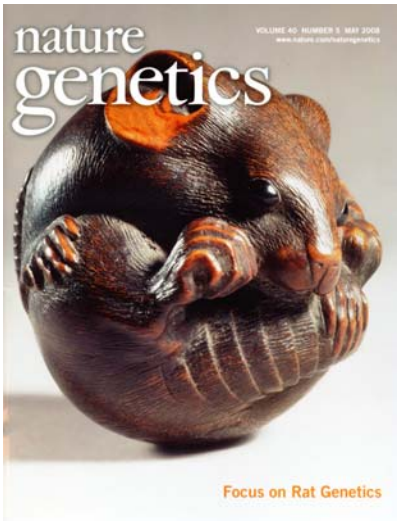
- Yogi A; Callera GE; Aranha AB; Graham D, McBride M; Dominiczak A.F, Touyz **Sphingosine-1-phosphate-induced inflammation involves receptor tyrosine kinase transactivation in vascular cells: upregulation in hypertension.** Hypertension *in press*.
- Hopcroft LE, McBride MW, Harris KJ, Sampson AK, McClure JD, Graham D, Young G, Holyoake TL, Girolami MA, Dominiczak AF. **Predictive response-relevant clustering of expression data provides insights into disease processes.** Nucleic Acids Res. 2010 Jun 22.
- Fairley S, McClure JD, Hanlon N, Irving R, McBride MW, Dominiczak AF, Hunt E. Mapping Affymetrix Microarray Probes to the Rat Genome via a Persistent Index. International Journal of Knowledge Discovery in Bioinformatics 2010; 1:1 pp 48-65.
- Koh-Tan HH, Graham D, Hamilton CA, Nicoll G, Fields L, McBride MW, Young B, Dominiczak AF. **Renal and vascular glutathione S-transferase mu is not affected by pharmacological intervention to reduce systolic blood pressure.** J Hypertens. 2009 Aug;27(8):1575-84.
- Graham D, McBride MW, Gaasenbeek M, Gilday K, Beattie E, Miller WH, McClure JD, Polke JM, Montezano A, Touyz RM, Dominiczak AF (2007). **Candidate genes that**

determine response to salt in the stroke-prone spontaneously hypertensive rat: congenic analysis. Hypertension. 50(6):1134-41.

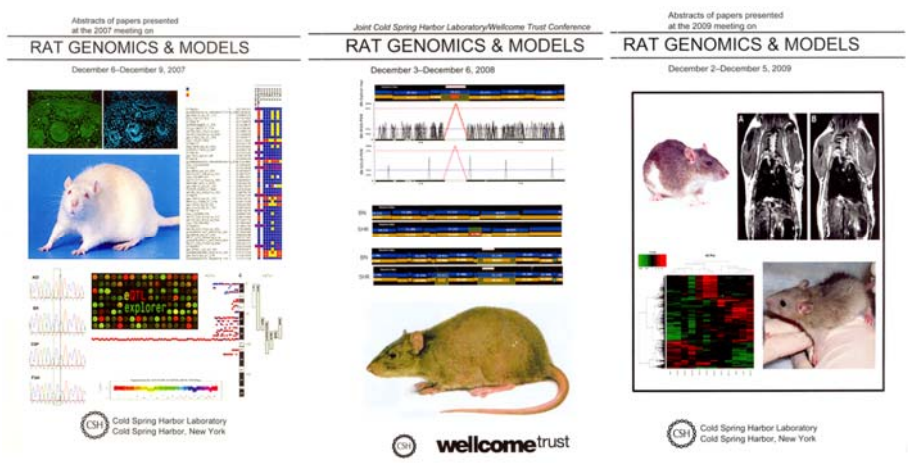
- Graham D, Huynh NN, Hamilton CA, Koh-Tan HHC, Smith RAJ, Cocheme HM, Murphy MP, Dominiczak AF. **The mitochondria targeted antioxidant MitoQ10 improves cardiovascular function and attenuates left ventricular hypertrophy**. Hypertension. 54:322-328. 2009
- Rintisch C, Kelkka T, Norin U, Lorentzen JC, Olofsson P, Holmdahl R. **Finemapping of the arthritis QTL Pia7 reveals co-localization with Oia2 and the APLEC locus.** Genes Immun. 2010 Mar 4; in press.
- Rintisch C, Forster M, Holmdahl R. **Detection of arthritis-susceptibility loci, including Ncf1, and variable effects of the major histocompatibility complex region depending on genetic background in rats.** Arthritis Rheum. 2009 Feb; 60(2):419-427.
- Huberle A, Beyeen AD, Ockinger J, Ayturan M, Jagodic M, de Graaf KL, et al. **Advanced intercross line mapping suggests that ncf1 (ean6) regulates severity in an animal model of guillain-barre syndrome.** J Immunol. 2010; 182(7):4432-4438.
- Rintisch C, Holmdahl R. **DA rats from two colonies differ genetically and in their arthritis susceptibility.** Mamm Genome. 2008 Jun; 19(6):420-428.
- Rintisch C, Ameri J, Olofsson P, Luthman H, Holmdahl R. **Positional cloning of the Igl genes controlling rheumatoid factor production and allergic bronchitis in rats.** Proc Natl Acad Sci U S A. 2008 Sep 16; 105(37):14005-14010.
- Holmdahl R. **IL-21 and autoimmune disease - hypothesis and reality?** Eur J Immunol. 2008 Jul; 38(7):1800-1802.
- Gelderman KA, Hultqvist M, Olsson LM, Bauer K, Pizzolla A, Olofsson P, et al. **Rheumatoid arthritis: the role of reactive oxygen species in disease development and therapeutic strategies.** Antioxid Redox Signal. 2007 Oct; 9(10):1541-1567.

(2) DISSEMINATION AND USE

The scientific achievemens of the EURATools project have been summarised above and published in major scientific journals and disseminated in more than 200 conference contributions (oral and poster presentations). More than 90 scientific publications acknowledge the FP6 funding, many of which cited above. A complete publication list is available on the website, www.euratools.eu. The outstanding highlight certainly is represented by the ***Natue Genetic*** **Focus issue on Rat Genetics** (published May 2008). EURATools investigators co-authored seven of the articles in this focus issue, including reports on positional cloning of genes for heart failure, cardiac mass and autoimmune glomerulonephritis, and a **community view on "Progress and prospects in rat genetics"**, supported by 257 named members of the rat genetics and genomics community. The editorial accompanying this issue commented "Although the mouse is still the mammalian genetic model of choice, the gap may be closing". EURATools researchers also have organised and contributed strongly to the annual conference "**Rat Genomics & Models**" (2007 and 2009 in Cold Spring Harbor, USA; 2008 in Hinxton, UK; 2010 in Kyoto, Japan), where approximately 30 to 40 % of the meeting presentations have been from EURATools researchers.



Nature Genetics **Focus on Rat Genetics, May 2008**



Rat Genomics & Models Meetings – Abstract Book Covers

One of the most notable highlights of EURATools Project has been the support for and co-ordination of the activities of early stage investigators. Two **Young EURATools Investigator Symposia** were initiated by, organised by and held exclusively for young investigators. All young investigators, from students and technicians to junior postdocs (approx. 30 on each occasion) presented their own data, showing a level of enthusiasm for rat genetics and genomics that could not have been anticipated and has given momentum to the activities of the Consortium. The enthusiasm has been furthered by the award of (mainly early stage) travel **fellowships**, which have enabled 14 exchange visits between EURATools laboratories (and in one case a collaborating laboratory).

In summary it can be said that the consortium is now perceived as the **European voice of the rat research community** which is also underlined by the trust the reviewers and the European Commission has to the core members of the EURATools consortium by funding the **FP7 project EURATRANS** (European large-scale functional genomics in the rat for translational research; April 2010 to September 2014; http://www.euratrans.eu). EURATRANS uses key resources, data and expertise from EURATools and brings it to the next level by using state-of-the-art and emerging large-scale technologies and advanced computation in an expanded multi-disciplinary approach to identify gene networks and genomic mechanisms underlying common diseases. EURATRANS will continue to use the rat as a model system to identify the major functional pathways underlying human inflammatory, cardiovascular and metabolic, and behavioral disorders.