EURATRANS aims to create fundamental knowledge about the genetic basis of human disease mechanisms. EURATRANS will achieve this goal by developing and apply novel functional genomic tools and generate large-scale genomic data, methodologies and resources (including animal models) tailored to understand the regulatory pathways that underlie epidemiologically important common diseases. Knowledge generated in the programme will be available to the scientific community and to the stakeholders through freely accessible databases and repositories. EURATRANS will pursue specific objectives in three areas that are most relevant to functional genomic research in the rat:

- **Large scale data generation using functional genomics technologies (Pillar 1)**
- WP1.1 Genome sequencing
- WP1.2 Transcriptome inventory - RNA sequencing
- WP1.3 Defining transcriptional initiation complexes and epigenetic modifications
- WP1.4 Quantitative proteome analysis
- WP1.5 Mapping the genetic determinats of the metabonome
- WP1.6 Computational infrastructure for multiple data modalities

- **Building and validating models (Pillar 2)**
- WP 2.1 Building molecular gene regulatory networks
- WP 2.2 Functional validation of key network discoveries

- **Comparative informatics & translation to human (Pillar 3)**
- WP3.1 Comparative analysis of rodent and human gene regulatory networks
- WP3.2 Integration of experimental results with human GWAS data
- WP3.3 Validation of of the discovered pathways in human disease tissues

The generation of large scale data using novel tools, methodologies and resources for functional genome research (Pillar 1) are technological objectives that apply to and overarch all scientific objectives with the aim of generating data appropriate for building and validating models for cardiovascular/metabolic, inflammatory, and psychiatric disorders (Pillar 2) to carry out comparative genomics and bioinformatics for translation to humans (Pillar 3).

The specific objectives of EURATRANS are summarized as follows:

Large scale data generation using functional genomics technologies (Pillar 1)

- Generation of the complete genome sequence for 9 key model rat strains (parental progenitors for crosses used). Based on the existing Brown Norway (BN) reference genome and building on our experience in sequencing the Spontaneously Hypertensive Rat (SHR) genome at >10x coverage using the Illumina/Solexa platform (Partner 1, 2, 4, 5, 10), we will sequence the parental progenitor rat strains for the experimental crosses used in this project (RI strains and HS) using next-generation sequencing technology to identify genetic and structural variation at the genome-scale level.
- Generation of a genome-wide transcriptional catalogue of coding, microRNAs and other non-coding RNAs using next-generation sequencing (RNA-seq) in selected cells and tissues of the key model strains, in RI panels and in rat ES cells. We will also determine the transcriptional initiation complexes in a subset of cells and tissues by ChIP-Seq.
- Identification of the differential methylation state of genomic DNA by methyl-sequencing in the BXH/HXB and LEXF/FXLE RI strains. We will characterize genome-wide chromatin state maps in defined cells and developmental stages in the progenitors of the BXH/HXB, LEXF/FXLE and HS panel.
- Using in vivo stable isotope labeling with amino acids (SILAC), we will quantify the proteome in key cells and tissues of the progenitor strains of the RI and HS lines and in RI strains and map differential protein expression across the RI strains.
- Using key cells and tissues we will generate metabonomic profiles in the key parental strains and two sets of rat recombinant inbred strain panels and map the genetic determinants of the metabonome.
- Consistent data processing such as read mapping, feature identification, and visualisation. Establishment of a bioinformatic research framework integrating multilevel datasets from molecular to physiological. We have established a critical mass of world-class bioinformaticians and computational biologists that will build a bioinformatic and analytical research framework annotating and integrating the generated multilevel DNA variation (genetic & structural), transcription, micro-RNA, metabonomic, proteomic, and epigenomic datasets with extensive sets of physiological data that has been generated in the two RI panels and the HS by Euratools and the Japanese Phenome Project.

Building and validating models (Pillar 2)

- Identify the genetic determinants of new molecular phenotypes using existing high resolution SNP maps in RI and HS animals that were generated through the Euratools and STAR consortia.
- Building molecular gene regulatory networks and identifying the most prominent hubs and nodes. We will apply techniques for QTL and quantitative trait transcript (QTT) network building using correlational and Bayesian models.
- Projection of regulatory networks on (patho-)physiological data to identify molecular pathways underlying specific physiological and disease phenotypes.

- Validation of the key components in disease-related networks through perturbation and manipulation of gene function by loss and gain of function experiments in vivo and in vitro. We will use successfully established techniques including ENU mutagenesis, transposon mediated mutagenesis in rat spermatogonial stem cells, transgenesis, antisense approaches, and emerging ES cell technologies that will allow manipulating the network components in pertubation experiments to validate findings.

Comparative informatics and translation to human (Pillar 3)

- Identify the regulatory expression networks, and epigenomic, metabonomic and proteomic signatures that are conserved between rats, mice and humans.
- Integration of human GWAS in our three focus disease areas with all models generated in order to identify biological pathways and processes that are important for these common human diseases (cardiovascular/metabolic, inflammation, psychiatric).
- Validation of biological insights from the discovery of pathways for the three focus disease areas in human disease tissues. We will use our extensive and preassembled tissue banks and patient cohorts to validate our findings in the human disease process.