Infrafrontier

Mouse models and phenotyping data for the European biomedical research community

Michael Raess and Martin Hrabé de Angelis
Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Germany

www.infrafrontier.eu

infrafrontier@helmholtz-muenchen.de

Mice as models for human diseases

Mouse and man share 95% of their genetic make-up. Mice are easy to keep and breed in the laboratory and researchers have developed a comprehensive toolbox for altering the mouse genome. This is why mice are ideal models for human diseases such as Diabetes, Osteoporosis, Asthma, Alzheimer's disease or Depression, as documented by an exponentially increasing number of scientific publications on mouse models and the Nobel Prize in Medicine in 2007. To provide a source for new mouse models and support the research activities in the field of functional genomics, the systematic mutagenesis of the approximately 25.000 genes in the mouse genome is currently underway, coordinated by the International Mouse Knockout Consortium (IKMC). These developments create an enormous demand for access to a systematic functional and molecular characterisation of the mouse mutants. Furthermore, new mouse models for human diseases must be made available to entire European mouse genetics, biomedical and translational research community [1].

Infrafrontier

It is clear that this tremendous task cannot be fulfilled by individual research facilities or on the national level alone. This is the rationale for the European project *Infrafrontier* (The European infrastructure for phenotyping and archiving of model mammalian genomes, www.infrafrontier.eu), which is coordinated at the Helmholtz Zentrum München by Prof. Hrabé de Angelis. *Infrafrontier* is on the European roadmap for research infrastructures of ESFRI (European Strategy Forum for Research Infrastructures, http://www.cordis.europa.eu/esfri) and receives funding from the EC's Seventh Framework Program. It will organise a pan-European research infrastructure to increase the capacities for systemic phenotyping and archiving of mouse models. The *Infrafrontier* consortium currently contains 22 partners (representing 14 phenotyping and archiving centres, 1 bioinformatics institute and 12 European ministries and funding agencies) from 10 different European countries. Six new partners will join *Infrafrontier* in the near future, extending the project to Austria, Czech Republic and Canada (Fig.1).

Phenomefrontier and Archivefrontier

In *Infrafrontier*, the four existing European primary phenotyping centres (or mouse clinics, see below) in Germany, France and the U.K. will associate with three emerging phenotyping centres in Spain, Italy and the Czech Republic, and with the Toronto-based Centre for Modeling Human Disease (CMHD) to form *Phenomefrontier*, a sustainable pan-European research infrastructure, providing capacities and access to mouse model phenotyping in Europe and around the globe.

The second major part of *Infrafrontier*, called *Archivefrontier*, aims to increase the capacities for the archiving and distribution of mouse models on a sustainable basis. To achieve this, EMMA, the European Mouse Mutant Archive (www.emmanet.org, [2]) will be extended and upgraded. Several new archiving nodes in Europe and one in Canada will be added to the EMMA network (Fig. 2).

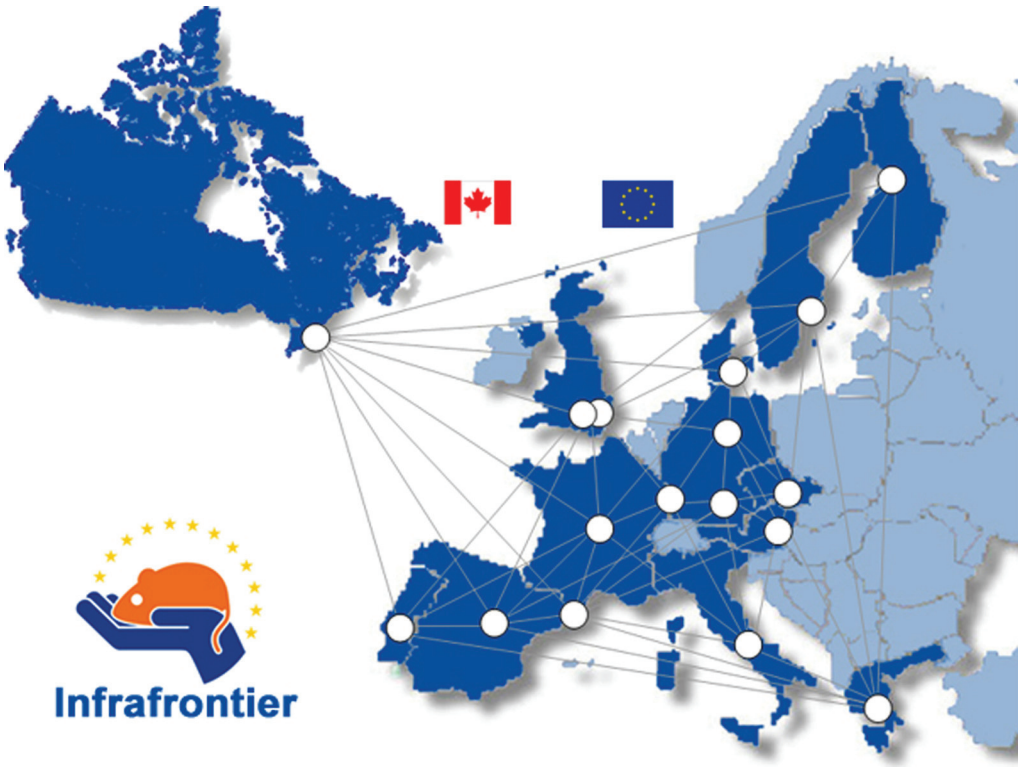

Figure 1. The members of the *Infrafrontier* consortium are phenotyping and archiving facilities as well as ministries, research councils and funding agencies in 12 European countries and Canada.

Systemic phenotyping and archiving / distribution of mouse models

Mouse Clinics use a whole-system approach to obtain a comprehensive picture of the systemic effects of mutations in the mouse genome: mutant mice are screened for alterations in bone and cartilage development, neurology and behaviour, clinical chemistry, immunology, energy metabolism and many more, comprising all essential organ systems [3]. At least 320 parameters are measured per mutant mouse line. This *systemic phenotyping* approach broadens our view of the functions of individual genes. Indeed, as recently reported by the German Mouse Clinic [1, 4], in the mouse lines characterised so far most mutations affected more than one organ system. Novel phenotypes were discovered in more than 95% of the cases, even in mouse lines that had been used in research for many years. This clearly shows the value of the approach. The existing European mouse clinics collaborate in the European pilot project EUMODIC (The European Mouse Disease Clinic, www.eumodic.org) to systemically phenotype 650 mutant mouse lines, following standardised phenotyping protocols that are detailed in the EMPReSS database (European Mouse Phenotyping Resource of Standardised Screens, www.empress.har.mrc.ac.uk, [5]).

Mouse models are archived and distributed to the biomedical research community by EMMA. The currently ten European partners of EMMA run an archive of together more than 1600 cryo-preserved mouse lines. Frozen sperm or embryos of about 300 mouse lines are added each year. In 2008 EMMA distributed upon request more than 360 mouse lines (as frozen material or life mice) to research institutions around the world.

Managing phenotyping and archiving data acquisition

High-throughput phenotyping and archiving would not be possible without the development of sophisticated computational tools for the management of complex logistics and the handling of large amounts of data. The phenotyping and archiving centres organised in *Infrafrontier* use several different software solutions, most of them developed in-house by dedicated bioinformatics groups.

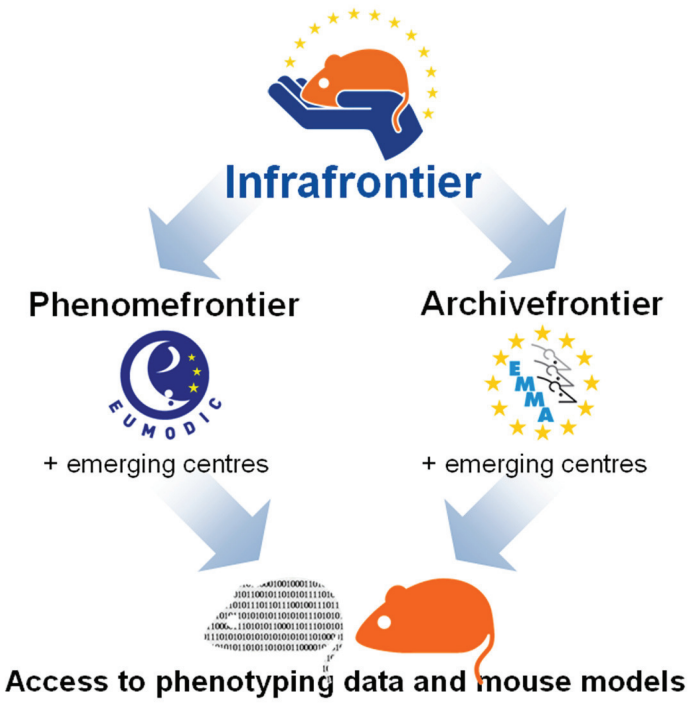

Figure 2. *Infrafrontier* organises two complementary pan-European research infrastructure networks: Phenomefrontier for large-scale systemic phenotyping, Archivefrontier for archiving and dissemination of mouse models. Both networks build on existing European initiatives and will include new facilities.

- **Mouse and facility management systems:** Management of mouse keeping, mouse breeding programs including pedigree information, cage capacities, specific housing or dietary requirements, as well as the documentation of all individual mouse fates have to be accomplished.
- **Workflow management systems:** High-throughput systemic mouse phenotyping requires the logistic management of multi-parallel test pipelines following a standardised workflow. Similarly, cryo-preservation and re-derivation of frozen material by the mouse archives involves complex parallel workflows. An additional task is the management of intellectual property or material transfer agreements associated with the different mouse lines.
- **Laboratory information management systems, storage device and sample management:** In the systemic phenotyping pipelines at least 320 different parameters are measured for each mouse line and their results recorded in a relational database. These results have to be matched with SOPs and metadata describing e.g. experimental and environmental conditions. In the archiving centres the storage capacities of the cryo-tanks or freezers, as well as the storage locations of each individual cryo-sample and its aliquots (including the back-up systems) have to managed.

An example of such a software solution is MausDB. It is used in the German Mouse Clinic and the German EMMA node. MausDB is realised as a LAMP system (**L**inux as operating system, **A**pache web server, **M**ySQL database, **P**erl as programming language) and accomplishes most of the tasks listed above [6]. MausDB is an open-source software that is freely available from the download section of the Institute of Experimental Genetics of the Helmholtz Zentrum München (http://www.helmholtz-muenchen.de/en/ieg/downloads/index.html).

Accessing phenotyping and archiving data

Phenotyping data from mutant mouse lines is available at EuroPhenome (www.europhenome.org, [7]), an open-access platform which contains the data collected by the EUMODIC project. This MySQL relational database offers a data browser

based on PHP, JSP and AJAX and several tools for data visualisation and data mining. An important future step will be the mapping of the phenotyping parameters onto phenotype ontologies that can be included into genome databases like Ensembl, Gene Ontology (GO) and Mouse Genome Informatics (MGI) [8, 9]. Information on mouse lines available from EMMA can be found on the archive's website (www.emmanet.org). The mouse line descriptions contain information on the genetic background, the phenotype and genotype descriptions that are cross-linked to the MGI database.

Bioinformatics activities in Infrafrontier

Infrafrontier's bioinformatics activities have, besides the maintenance of the project's web resources, two major objectives. The first is a comprehensive survey on IT systems for managing facilities, mice and data that are used by phenotyping and archiving facilities around the world. This will lead to a report with recommended IT solutions for new facilities. Moreover, suggestions for a common nomenclature will be made that will facilitate exchange of mouse line data between the facilities. The second objective is the definition of minimum information standards required for an integration of the existing databases for phenotyping, archiving and mouse production data (EuroPhenome, EMMA, EUCOMM [the European part of the IKMC]).

Since data is a major asset of *Infrafrontier*, it is important that the project keeps strong links with other information-oriented European initiatives, such as the coordination action CASIMIR (Coordination and sustainability of International Mouse Informatics Resources, www.casimir.org.uk), the EMBRACE Network of Excellence (A European Model for Bioinformatics Research and Community Education, www.embracegrid.info), and the ESFRI project ELIXIR (European Life Sciences Infrastructure for Biological Information, www.elixir-europe.org), as well as the other ESFRI research infrastructure projects in the biological and medical sciences.

Conclusions

In order to understand the fundamental processes governing living organisms in health and disease, modern life sciences employ increasingly sophisticated technologies and produce increasingly complex and large datasets. High-throughput phenotyping and archiving of mouse models of human diseases are good examples for this development. Infrafrontier will provide access to a sustainably funded research infrastructure to meet the growing demand for these services in the biomedical research community. Infrafrontier in concert with the other ESFRI biological and medical research infrastructure initiatives will promote cutting-edge research ranging from basic science to the translation of results into novel drugs and treatments. Their success is likely to change the face of biomedical research in the European Research Area.

References

1. Beckers J, Wurst W, Hrabé de Angelis M (2009) Towards better mouse models: enhanced genotypes, systemic phenotyping and envirotype modelling. Nat Rev Genet 10: 371-380.
2. Hagn M, Marschall S, Hrabé de Angelis M (2007) EMMA - the European mouse mutant archive. Brief Funct Genomic Proteomic 6: 186-92.
3. Brown SDM, Hancock JM, Gates H (2006) Understanding mammalian genetic systems: The challenge of phenotyping in the mouse. Plos Genetics 2: 1131-1137.
4. Fuchs H, Gailus-Durner V, Adler T, et al. (2009) The German Mouse Clinic: A Platform for Systemic Phenotype Analysis of Mouse Models. Current Pharmaceutical Biotechnology 10: 236-243.
5. Brown SDM, Chambon P, Hrabé de Angelis M, et al. (2005) EMPReSS: standardized phenotype screens for functional annotation of the mouse genome. Nature Genetics 37: 1155-1155.
6. Maier H, Lengger C, Simic B, et al. (2008) MausDB: An open source application for phenotype data and mouse colony management in large-scale mouse phenotyping projects. BMC Bioinformatics 9.
7. Mallon AM, Blake A, Hancock JM (2008) EuroPhenome and EMPReSS: online mouse phenotyping resource. Nucleic Acids Research 36: D715-D718.
8. Hancock JM, Adams NC, Aidinis V, et al. (2007) Integration of mouse phenome data resources. Mammalian Genome 18: 157-163.
9. Beck T, Morgan H, Blake A, et al. (2009) Practical application of ontologies to annotate and analyse large scale raw mouse phenotype data. BMC Bioinformatics 10: S2.