



PRIME EuroMouse meeting and SAG Priorities for mouse functional genomics research in Europe 6th July 2007

PRIME sponsored the EuroMouse meeting at Heathrow on 5th and 6th July to bring together researchers from the EU-funded mouse projects in Europe. The aim of the meeting was to foster collaboration and links between the projects and their resources and to debate the future priorities for mouse functional genomics research in Europe. The PRIME Scientific Advisory Group then met to synthesise the discussion and to determine the key priorities. Their summary recommendations are given below. In some cases the proposed strategic priorities represent a continuation and further development of major themes already initiated. For the others, they represent new opportunities often arising from resource generation and research already underway.

RECOMMENDATIONS

Mutant libraries

1. The current ES cell knock-out and conditional libraries underway as part of the EUCOMM/KOMP/NorCOMM programmes are being produced in C57BL/6N cells. There is a need to develop mutant ES libraries on different backgrounds further enhancing the power of the mouse mutant resource and creating rich opportunities to further explore gene function. A first step to achieving this goal is the production of robust ES cell lines from other background strains of mice. 10 – 12 strains should be selected from the 40 strains being used in the phenome project.
2. EUCOMM and the other international mouse knock-out projects (KOMP and NorCOMM) are aiming to produce ES cells with a knock-out for each the 25,000 or so mouse genes. It is anticipated that these projects will produce ES cell lines for about 80% of the total. The remainder will present a greater challenge that will require extra time and resources. It will be important to ensure continuity of the mutant production pipelines and to engage them in the finishing effort and to apply the developed infrastructure and skills to subsequent efforts in generating mutant libraries on additional genetic backgrounds (see above).

Cre zoos

The development of extensive cre zoos will be pivotal if we are to take full advantage of the conditional mutant libraries being produced by the International Mouse Mutagenesis Consortium. Current cre resources are fragmented and poorly characterised. There is a need to coordinate the production of cre zoos, ensuring that they are produced on uniform genetic backgrounds, are well phenotyped and expression patterns are carefully determined. A two-stage process is required:

1. A Coordination Action project to coordinate the dissemination of extant resources and capture existing knowledge on for example expression patterns.
2. This should be complemented by an Integrated Project to develop new lines using standardized approaches and employing in depth characterization. Expression data from EurExpress could be used to target tissues and identify appropriate loci.

Research is also needed on inducible cre to further develop the technologies in a systematic way and investigate reversibility.

Proteome

It is timely to undertake a major project to map the mouse core proteome. The acquisition of a core proteome map from a few well-chosen cell types will be an enormously valuable tool in exploring and understanding the functional landscape of mammalian cells. The development of genome-wide knock-out resources in the EUCOMM, KOMP and NorCOMM projects along with technical advances that will allow genome-wide tagging opens up new approaches to cataloguing the mammalian proteome. Research on appropriate tags is underway and, building on developments in high throughput mass spectrometry technology, rapid progress can be made in large-scale studies of the mammalian proteome. The project would include the determination of proteome maps in several tissues and where possible develop comparisons between the mouse and human proteome. Any large-scale programme will benefit from international cooperation akin to the arrangements developed in the International Mouse Mutagenesis Consortium.

ENCODE

The ENCODE project aims to provide a complete catalog of functional elements in the human genome. This project, with funding lead by NHGRI in the US, has passed its pilot phase and is moving into the scale up phase in 2007. There is now an opportunity to initiate a mouse ENCODE project. It would save considerable duplication of effort worldwide as many researchers would not have to perform expensive gene specific tests e.g. examining DNaseI sites around a gene of interest. It would enhance the mouse as model for human and mammalian biology. Moreover, when coupled with the excellent genetic tool kit in mouse, many observations can be followed up with specific targeted experiments. The use of novel technologies such as Archea-express to investigate the functionality of the extensive transcriptome of the mammalian genome should also be explored as part of the wider mouse ENCODE. Overall, the proposed project would

- Provide a genome-wide comparative functional dataset, allowing a deeper insight into the comparative biology in mouse and human, and underpinning and delivering a better understanding of disease mechanisms.
- Benefit enormously from experience in human ENCODE, and in particular that the bulk of the analysis of the human ENCODE pilot was undertaken in Europe.
- Take advantage of the critical mass of European mouse resources from mutagenesis pipelines to informatics analysis.

It will be important to ensure that we bring together and synergise the activities of the communities involved with generating and analyzing mouse ENCODE and the communities involved in mutagenesis and functional annotation of the mouse genome. For this reason, we would propose a project in two parts. A first phase would envisage a comprehensive genome scan of the most significant genome sequence elements. Given advances in technology in genome element annotation this could cover a significant portion of the mouse genome. This phase would also include pilot projects in mutagenesis and functional annotation. This would be followed in a second phase by use of the mouse toolkit for more comprehensive functional studies of particular elements with a focus in all of the genetic projects on disease related genes.

Phenotyping, archives and InfraFrontier

The future goal of functional genomics is to determine the function of all of the genes in the human genome and their involvement with disease. The mouse is a key model system to address this undertaking and the availability of comprehensive mutant ES cell libraries is a major step towards this goal. However, phenotyping of the mutant lines remains a formidable challenge to which Europe, through the EUMORPHIA and EUMODIC programmes, is making a

leading contribution. Additionally, through the InfraFrontier programme, Europe aims to enhance the infrastructure required for large-scale production and archiving of mutants and their phenotyping. EUMODIC is an ongoing pilot programme funded through FP6 to apply standardized primary phenotyping methods to determine the phenotype of around 650 mouse mutants. An International Mouse Phenotyping Consortium is being formed and, building on the EUMORPHIA developments and the experience of EUMODIC, will develop global plans for scale-up and approaches to systematically phenotype thousands of mutants. EUMODIC and its constituent members will play a leading role in this endeavour and it will be important for the European consortia to continue their efforts in large-scale systematic phenotyping efforts and take advantage of the progress and lead already made in mouse phenotyping approaches in the European arena. In addition, it is important to continue to improve high-throughput phenotyping approaches, including for example the introduction of automation, in order to provide significant enhancements in scale. Another key part of this effort will be the formation and maintenance of extensive archives of mutant mouse lines stored as frozen sperm and embryos. These archives have been formed under EMMA, but will need to be greatly expanded to store all of the mutant lines produced in InfraFrontier and to be able to meet demand for these from scientists worldwide.

Sequencing diverse rodent genomes

The study of the function of genetic variation in the human population has benefited enormously from the study of other diverse genomes, including amongst others apes and primate species. The identification of constrained and non-constrained sequence allows a number of insights into sequence evolution and function. Re-sequencing costs are now plummeting to a point where it is timely to consider projects that would tackle the sequencing of numerous additional species particularly species focused on understanding better the genetic basis of sequence variation in response to environment. Given the pivotal role of the mouse as a disease model, there would be enormous synergies and opportunities to sequencing a large number of representative species from the order Rodentia. Aside from the relationship to the experimental organism, the house mouse, there are several important reasons for considering this order as a focus for sequencing. The order Rodentia represents 42% of mammalian species. Rodents are represented on all continents, bar Antarctica, and occupy a very diverse habitat range from rain forest to desert. They are critical to many ecosystems where they are often an important food source. Importantly, the order is known to undergo very rapid karyotypic evolution, a phenomenon that is not understood but may reflect partly the organization of rodent populations in demes. The lab mouse belongs to the family Muridae, which includes the Rat and the Apodemus (field mouse) species. It would be appropriate to sequence a number of genomes from this family, along with a selection of representative species from other families representing both evolutionary and habitat diversity. The analysis of the sequenced genomes would allow a fascinating insight into the evolution of this order and into mammalian genome evolution more generally and its relationship to environmental and other pressures.

Informatics – European Networks of Mouse Databases

A number of databases have been developed under EC funding. These range from large-scale depositories of data from projects such as EMMA to databases developed in individual disease research projects such as EUSynapse and EUCLOCK. The CASIMIR project is cataloguing these databases and determining information about their structure that will allow them to be coordinated. Prototype links are being established under ENSEMBL. These developments form the basis for a new Integrated Project, the European Mouse Information Network (EMIN), which will provide a federated 'one-stop shop' for information on European mouse resources and research. One of the key aims of EMIN will be to ensure that data and data exchange standards for mouse phenotyping, including development of common SOP, ontology and phenotype data

exchange standards are developed and established across the European network and further afield.

Informatics - Linking mouse phenotypes to human disease

The current large-scale mouse research projects are focused at producing a knock-out for each mouse gene, determining the phenotype of resulting mice and relating this to human disease. The development and analysis of pre-clinical models is a key driver in the translation engine, providing important resources for clinical genetics, experimental medicine and therapeutic efficacy testing. The utility of these models will significantly depend upon our ability to relate mouse phenotypes to human disease phenotypes. Establishing linkages between these phenotype sets is non-trivial and suffers from the lack of appropriate informatics methodologies, particularly in the area of determining ontological relationships. It will be important to initiate a project to study and apply methods to assist in mapping mouse and human disease phenotypes and widening our pool of pre-clinical models with consequent benefits for the understanding of disease mechanisms.

Training

We envisage two critical areas where training is a priority if we are to ensure the skills base and expertise to underpin a vibrant scientific sector encompassing functional genomics, genetics and disease mechanism studies.

1. Skills in mouse genetics, biology and physiology are key to a vigorous community at the forefront of international studies in functional genomics and translational research. The development of the mouse clinics and associated centres offering specific expertise in mouse phenotyping provides numerous opportunities to train people in mouse genetics, biology and physiology. These training courses may best be provided under a Marie-Curie project to provide a hub and spoke training within and between the European mouse clinics.
2. Pathological investigation of the many disease models generated through the International Mouse Mutagenesis consortium will be vital if we are to fully exploit the resources created, particularly in the translational dimension. The lack of skilled pathologists with experience of mouse pathology is of acute concern in both Europe and elsewhere. Training opportunities are required for research scientists and pathologists, importantly harnessed to the phenotyping centres and clinics already underway or planned. The opportunities to co-fund mouse pathology training by industry should be explored and in addition any new training developments should emerge in consultation with official bodies that can provide accreditation for these courses.

PRIME

The PRIME project has performed a central role bringing together the diverse EuroMouse projects as well as providing a platform to determine priorities and the direction in large-scale mouse mutagenesis, phenotyping and archiving projects. In addition, PRIME has played a significant role in allowing the community to scope out future priorities and for new projects to emerge, particularly in the characterization and analysis of disease models. The project has also had a wider role in promoting the value of the mouse as a model of human disease and its role in pre-clinical studies and translational research. The effectiveness of such a Co-ordination Action underlines the importance of continuing the project and ensuring the community continues to work towards cohesive goals at the international forefront.