1st INCF Workshop on

Genetic Animal Models for Brain Diseases

Stockholm, Sweden 13-14 December 2009

Olaf Riess and Holm Graessner
Authors
Olaf Riess and Holm Graessner, University of Tübingen, Germany

Scientific organizer
Olaf Riess, University of Tübingen, Germany

Participants
Christian Desaintes, European Commission
Gerd Kempermann, University of Dresden, Germany
Trygve Leergaard, University of Oslo, Norway
Hitoshi Okamoto, RIKEN Brain Science Institute
Jonathan Pollock, NIH, USA
Pasko Rakic, Yale School of Medicine, USA
Dennis Selkoe, Harvard Medical School, USA
Guus Smit, Vrije University Amsterdam, Netherlands
Monte Westerfield, University of Oregon, USA
Rob Williams, University of Tennessee, USA

Observers
Jan Bjaalie, former Executive Director of INCF
Sten Grillner, Chair of the Governing Board of INCF
Ulf Indal, INCF National Node Norway
Ewelina Knapska, INCF National Node Poland
Matt Nolan, INCF National Node UK
Hui Wang, Deputy Director INCF
1. Executive summary

The INCF Secretariat organized a workshop to focus on the “role of neuroinformatics in the processes of building, evaluating, and using genetic animal models for brain diseases” in Stockholm, December 13–14, 2009. Eight scientists specialized in the fields of neuroinformatics, database, ontologies, and brain disease participated together with two representatives of the National Institutes of Health and the European Union, as well as three observers of the national INCF nodes of Norway, Poland, and the United Kingdom. Further, the INCF Governing Board chair, Prof. S. Grillner, and the former executive director of the INCF, Prof. J. Bjaalie, contributed to defining the format and general objectives of the workshop.

Findings

The participants of the workshop reported on the perceived needs and challenges in the field that can be addressed by neuroinformatics approaches and stimulation or coordination by the INCF. A key finding that emerged from the discussions is the increased need for standardized and interoperable database resources, in which different modality phenotype data from available disease models and wild type animals can be integrated. The scope of the discussions was defined by the different expertise represented by the participants.

Recommendations

The following recommendations were made:

- Encourage, support and fund work on existing databases and information systems, which need to be linked and more easily accessible for researchers to gather information on gene function, expression, variations, and differences among species to allow easier planning and evaluation of animal models. There are numerous databases that cover some aspects of these needs, but gene expression networks in different models, in particular, are not easily accessible, and data are difficult to compare among different models. Most importantly, connecting and integrating software is needed.

- Encourage, support and fund work on database systems for reporting phenotypes of models. Such databases could be comparable to the DNA gene banks. This requires standardization of phenotype data, which needs to be done first, and mandatory reporting of data to the databases (e.g. via agreements with journals for publication). INCF should discuss with funding agencies the need for such a database and for coordinated actions.

- Encourage, support, and fund work on connecting “phenotype” databases of models with “molecular” databases. During development of the phenotype databases one should ensure that they are directly linked and integrated with existing and future molecular databases. In particular, transcriptomic and proteomic databases are still under development and need to be integrated. INCF should inform funding agencies of the need for such coordinated actions.

- Clearly, the generation of “phenotype” databases requires standards for describing behavior in different strains and species. Thus, INCF should encourage, support, and fund work on databases of baseline data from wild-type strains for both behavioral and molecular data. INCF should discuss with funding agencies the need for such databases and coordinated actions.

- Encourage, support, and fund the generation of MANGO – a Mouse adult neurogenesis gene ontology database. As the role of neurogenesis in initiation and progression of neurodegeneration has not yet been defined, and in particular as the rate of neurogenesis in different species and different genetic background needs to be studied in detail, the participants supported the idea of MANGO. Close coordination of this activity is needed among European, North American, and Asian funding organizations. INCF should inform funding agencies of the need for such a database and coordinated actions.

- Continue to coordinate work with animal models on various aspects of neuroscience (such as neurogenesis, linking cell populations and developmental stages across species, etc.) to bring together more experts on one subtopic. For example, the INCF could arrange and coordinate further workshops.

- Arrange annual workshops on large-scale modeling to develop standards for automated home cage behavior. Current software development is poor, read-outs are not standardized and technologies are not comparable. Different technologies are on the market, but no comparison has been made. Minimal and optimal standards need to be defined to support comparisons among publications. This is also absolutely necessary for comparison of preclinical treatment studies.
Because training and teaching are absolutely crucial for all recommendations, INCF should also consider calling for funded programs to enhance the entry of bachelors- and masters-degree level scientists and technicians who can expedite the construction, population, and curation of the databases envisioned herein. This could provide an impetus for more young people to enter bioinformatics science and would not necessarily require immediate expansion of PhD and MD level training programs for these purposes.
2. Introduction

Neurodegenerative diseases represent one of the major health care challenges of this century. Despite a continuously increasing percentage of individuals suffering from these diseases world wide, a therapy is not available for most of the conditions. With the rising incidence of neurodegenerative disease, costs for medical and sociological care will become significant financially for developed countries. Although these change have been well known for many years or even decades, relatively small amounts of money have been invested in brain research in Europe, compared to the US or Japan. Furthermore, the pharmaceutical industry is more interested in drug development than drug delivery into the brain and has developed a huge market, such as drugs against cancer. Historically, limitations of previous models of toxic neurodegeneration, such as MPTP or 6-OHDA rodent and monkey models of Parkinson’s Disease (PD; reviewed in Maries et al. 2003, The role of alpha-synuclein in Parkinson’s disease: Insights from animal models. Nat Rev Neurosci 4:727-738; and Capitanio and Emborg 2008, Contributions of non-human primates to neuroscience research. The Lancet 371:1126-34), and deduced therapeutic neuroprotective strategies have widely failed in the clinic (reviewed in Mandel et al. 2003, Neuroprotective strategies in Parkinson’s disease. An Update on progress. CVS Drugs 17:729-762).

With the recent identification of genetic causes of several neurodegenerative diseases, we have a vastly better chance to develop novel, genetically modified models which may mimic important aspects the human condition much more closely than the earlier toxic models. These higher translational values of such genetic models (as examples see reviews Dawson et al. 2010, Genetic animal models of Parkinson’s disease. Neuron 66:646-661; Ashe and Zahs 2010, Probing the biology of Alzheimer’s disease in mice. Neuron 66:631-645) will make them more appropriate tools for preclinical testing. However, they first need to be characterized behaviorally and neuropathologically in great detail. For neuropathological characterization we recommend the report from the 1st INCF workshop on Mouse and rat brain digital atling systems (www.incf.org), even though numerous aspects of neuropathology, such as sensitivity of different strains and species or even cell types to toxic agents such as mutant or overexpressed protein, need to be defined.

In addition to questions about which strain (genetic background) is best (pharmacological studies typically use out-bred strains, whereas genetic and behavioral studies use inbred strains), there is a serious dearth of behavioral databases and standardized behavioral protocols (SOPs), and there is no general agreement on readouts for human neurological symptoms in rodents. Again, neuroinformatics is required to define standardized read outs and to create appropriate databases.

Several scientists contributed their personal views about the needs in the field of neurodegeneration and the contribution of neuroinformatics. Because this topic is rather complex, we will cover only those aspects that are represented by the scientific expertise of the participants of the INCF workshop held in Stockholm December 2009.

3. Concepts

To define the “Role of Neuroinformatics in the process of generating, evaluating, and using genetic animal models for brain diseases” more precisely, participants of the workshop discussed various readouts and levels of investigation of models. Here, we define the terms used in the context of the models.

Animal model

Modeling plays a major role in neuroscience and in particular in neuroinformatics. A broad definition of models is provided by the report of the 1st INCF workshop on “Large-scale modeling of the nervous system” held in December 2006 in Stockholm (www.incf.org). At the subsequent INCF meeting in 2009, we focused on animal models of brain diseases. For animal models, we usually distinguish between biological models that provide information about the biological function or dysfunction of the studied gene and disease model that mimic a human disease. In principle, disease causing mutations found in human patients are introduced into the germline (so called knock in) of a mouse (or other species), the gene is destroyed (knock out model), or the endogenous protein is overexpressed (transgenic). For each technique, modifications enable the gene or the mutation to be expressed at a specified time or in a specific cell type. In addition to rodents, non-vertebrates, such as Drosophila and C. elegans, are used to study brain diseases. Recently, zebrafish have provided a novel species to study neurodegeneration (see Ringberg Symposium at ringberg.web.med.uni-muenchen.de/schedule/index.html). For cognitive (dys)function, monkeys are used in rare cases.

Genetic background

For most human diseases, there is wide variation in the age of onset or rate of progression even if two patients carry the same disease causing mutation. This is usually explained by environmental and genetic background factors. In laboratory animals, environment can be controlled fairly well. However, different genetic lines have gene variants that lead to amino acid differences or to different levels, time course, or spatial pattern of gene product expression.
In humans, two individuals (except of monoygotic twins) can have approximately 3 million different polymorphisms. To minimize the influence of genetic variation in disease manifestation or progression models, inbred lines are frequently used. Inbred lines are generated by extensive inbreeding to reduce genetic variation. In contrast, outbred lines are used to maintain large genetic variability within a population.

**Biomarker**

Relative to animal models of human diseases, biomarkers are biological compounds that can be easily obtained and measured and that are specific for a disease or a particular stage of a disease (as an example of biomarkers in neurodegeneration see for instance review by Maetzler et al. 2009, Progression of Parkinson’s disease in the clinical phase: potential markers. Lancet Neurol 8:1158-71). Metabolites, proteins, or RNA can provide biomarkers, as well as imaging technologies, such as PET.

**Automated home cage analyses**

Most of behavioral analyses depend on the experience of the investigator. Due to handling differences, read out of behavior may vary dramatically among investigators or even more significantly among different facilities (as an example read: Crabbe et al. 1999, Genetics of mouse behavior: Interactions with laboratory environment. Science 284: 1670-2). These insights led to the development of the so-called automated home cage environment, which allows behavioral analysis without significant handling by the personnel. Different systems are on the market, one uses a transponder implanted under the skin of the animal that allows analysis of several animals in one cage and thus also study models in their social environment. The more widely used cages track one animal per cage using infrared beams incorporated into the frame of the cage (for review see: Spruijt and Dev-Visser 2006, Advanced behavioral screening: automated home cage ethology. Drug Discovery Today: Technologies 3: 231-7).

**Neurogenesis**

The knowledge that new neurons are generated in adult mammalian brain is recent (Reynolds and Weiss 1992, Generation of neurons and astrocytes from isolated cells of the adult mammalian nervous system. Science 255:1707-10). These cells are predominantly generated in the subventricular zone of the lateral ventricles and in the subgranular zone of the hippocampus. They may have a major importance in memory and learning (Kempermann et al 2004, Functional significance of adult neurogenesis. Curr Opin Neurobiol 14:186-91). Impaired neurogenesis may play a role in several neurodegenerative diseases (references).

---

### 4. Workshop Discussions

During the workshop we concentrated on how neuroinformatics can help overcome current limitations of models of neurological diseases. We divided the discussion into several parts:

**4A. Models and Species:**

Choice of species (monkey, mouse, rat, zebrafish, C. elegans, Drosophila; for instance, should monkeys always be considered as models for brain diseases and should they be the last species to be modified). The workshop participants had a lively debate about which species to focus on in creating comprehensive database(s). The discussion concentrated on differences as well as similarities among different species. Some aspects of mice make them less advantageous, including the remarkable variability among currently used mouse strains, their highly inbred nature, and their limitations in regards to a) size (e.g. with respect to CSF and blood sampling for biomarker studies) and b) difficulty sophisticated behavioral testing and in vivo electrophysiology. The advent of genetically manipulated rats means that this species could become increasingly more useful for standardized data collection of normal aging and as a comparator for studies of mouse models of disease. A similar comprehensive effort could be performed with one or two invertebrate animals (worms or flies), and some such studies are already well-underway. Neuroanatomical differences of mice, monkeys, and humans were discussed. There is reason to suggest that one should model particular human diseases only in species that have similar neuroanatomical features. As an example, species lacking the subtypes of interneurons that are affected in schizophrenia in humans are likely to be inappropriate models of this disease. On the other hand, similarities between species exist for other models mimicking human diseases, in particular with respect to protein structure. Thus, many aspects of a human disease can also be studied in species distantly related to humans such as C. elegans, zebrafish, and Drosophila. In some cases, it is difficult to predict which species should be used as a primary choice. On the other hand, for example, marmosets are recommended in Alzheimer’s Disease (AD) research because most of them develop plaques unlike most dogs or vervet monkeys. Of course, it is important to define the read out that the model should provide.
• Choice of strain (inbred, outbred) and unsolved issues such as pharmacogenomics, neurogenesis etcetera. Is there a need to use different species/strains for different diseases or different stages of disease? As an important example, differences in neurogenesis in various mouse strains were discussed. We acknowledge that there is only limited information about strain differences at this point and strain differences are not a major focus of current research. Gerd Kemperman crossed BXD80 strains to discover neurogenesis loci, distinguishing cis genes (with feedback regulation) vs. trans genes (without feedback). In general, the importance of neurogenesis in neurological diseases is not well studied and needs to be investigated widely among different species and in the different genetic backgrounds. For the INCF, the generation of a mouse adult neurogenesis gene ontology database was recommended.

• Standardization of phenotyping (e.g. automated home cage systems, selection of paradigms). Since the readouts of “conventional” behavioral phenotyping differ substantially between laboratories, even when the same strains, equipment, and conditions are used (Crabbe et al. 1999), there is a need for standardized behavioral phenotyping. Automated (high-content) home cage behavioral analyses have emerged as promising new tools to overcome some of these limitations, since human intervention is minimized compared to stand alone testing (comparative analysis). However, behavioral analyses need extensive and rigorous descriptions of operating procedures, and thus metadata recording. The, INCF could make a great contribution to the establishment of formats and standards for metadata depositions in this field.

• Comprehensive collection of factors modifying the phenotype requires in depth analysis of multiple parameters. This can be achieved by gathering multiple data parameters through High Content Screening approaches (HCS) that are information dense, and medium to high throughput. The fact that HCS can be operationalized with well-described standard operating procedures makes it suitable for cross-lab, cross-experiment, and cross-animal model comparison. For animal models, HCS can be used in transcriptomics, proteomics, and behavioral analysis. However, standards need to be defined.

• Rating scales for phenotypes (comparable to human disease readouts). For disease models, and rodents in particular, existing limitations of behavioral screens become even more pronounced because there are no guidelines for standard ways to measure movement disorders (e.g. dystonic, ataxic, choreatic, parkinsonian, et cetera.). Automated home cage analyses may also provide new insights (for example, small differences in rodent behavior may be better recognized during the night).

• Biomarkers. The potential of biomarkers for defining disease, disease onset, or disease progression has not yet been fully explored in human patients. This is in part due to the high variability of the readouts and limited access to tissues, in particular brain tissue. Here, animal models are preferred over humans because tissue availability is not an issue, ethical issues are reduced, and developmental brain diseases and standardization (sex, age, brain region, genetic background etc.) are relatively easily achieved. Imaging analyses, such as PET, tracers, and antibodies can be validated as potential biomarkers before applying to humans. The participants discussed whether monkeys are preferred over rodents systems for biomarker studies because they are evolutionary more closely to humans. Monkeys, however, provided only a limited number of disease models. Standard operating protocols are necessary, as well as databases to store and compare studies.

• Neurometabolism and neurotransmitters. Neurotransmitters (can also be considered biomarkers) can be studied only in models. Although electrophysiology was not a topic of the workshop, it is closely linked to neurometabolism and neurotransmitters and can be used in vivo in models. In vitro, specific neuronal or glial cell populations can be isolated from disease models, grown, and recorded with patch electrodes under various conditions, thus contributing significant information about diseases.

• Generation of databases for phenotypes and biomarkers of disease models. Currently, no database and no standards exist for reporting phenotypes and models, not even for a particular disease with the exception of some lay organizations that have collected data from the literature on existing mouse models (CHDI for Huntington’s disease or Michael J Fox Foundation for Parkinson’s disease). These data are, however, not collected in a standardized way and the raw data are not reported. The participants strongly recommend also reporting negative data on models, which should also be entered into databases to preclude repeated attempts to characterize specific models.

• Databases connecting molecular with behavioral phenotypes are lacking. Together with the lack of databases on phenotyping, there is a strong desire for databases which have the potential to link all large data sets from these models, including transcriptomics, proteomics, neurometabolites, imaging data, etc. Such higher level systems will expectedly be instrumental for the integration of different data modalities into new knowledge.
4B. Optimizing Models:

Given the problem of choosing which genetic background and readout to use in wild-type animals, the lack of databases to store extensive phenotype data, and the absence of a connectivity database with molecular data sets, the problems with genetic models are obvious and difficult to solve: How to define different diseases based on symptoms as specific readouts? All diseases in humans have a range of early onset signs which may differ among patients. Also, not every patient manifests all signs. Furthermore, some diseases manifest differently at different ages (for instance Huntington's disease in early age manifests spastic rather than barely choreatic movement disturbances). For animal models this is even more complex. We do not have choreatic, ataxic, or dystonic symptoms, but rather movement disturbances in general. Subtle signs which can be provoked in humans by asking to perform specific tasks may thus be difficult or not at all visualized in models. Sophisticated behavioral tests can be done only in species like rodents or monkeys.

In addition to these biological limitations, we also have model-dependent problems to be solved. Some of these issues are (i) for transgenic animals, publication of the level of expression of the transgene, (ii) clear description of the promoter used, (iii) DNA integration sites, (iv) inventory of the phenotypes over time and not only until first symptoms appear, (v) description of unchanged or normal or phenotypes, (vi) description of the genetic background, and (vii) cross-breeding into different background strains. There is currently no recommendation for which strains to use for different diseases. Although most genetic association studies need to be confirmed in a different population, phenotype studies in mice are typically conducted only once and in only one laboratory. For studies of cognition, are mice really the appropriate species? Recently, the tendency is to generate complex genetic models to mimic human neuropathology and analyze of genetic pathways better, but as a result, the readouts become more complex.

Anything learned about disease models can be overlaid with what is culled from the healthy animal databases. Thus, for INCF, creation and phenotyping of animal models of specific diseases necessitates creating useful databases of "normal" (healthy) animals over their lifespan.

What is needed is an internationally accepted standards for how to generate models (transgenic cDNA, transgenic BAC, knock in, knock out, level of expression) and analyze models.

What is needed for good models of neurodegenerative diseases?

• Centralized neuropathology facilities; standards optimized for translation to and comparison with human neuropathology
• Typical human phenotype to be “translated” into an animal model
• Cell death / opposed to neuronal dysfunction
• Useful readouts (biomarkers) for evaluation of experimental treatment and preclinical studies

Discussion extracts

• It was pointed out that that the standards maintained in human clinical trials needs to be introduced to the field of animal model research (database of protocols). The emphasis should be on pathway analysis, from neuroinformatics, functional genomics (gene-chip, mass spectrometry, antibody array), and behavioral studies of cognitive decline, to establishment of database systems holding a range of phenotype data (in vivo histobiochemistry, monitoring of disease proteins over time using microdialysis).

• Models also have a lot of variation precluding robust translation from models to humans. Reference populations to deal with this issue (for both humans and models) are necessary.

Preliminary conclusion:

We need a discussion about whether it is favorable to use several species and one readout or one species and several readouts. Most likely, pre-clinical studies should include many read-outs and several models, preferentially even in different species.

There is great need for comparing different existing models of a disease in one laboratory or even more than one laboratory. This is not ground breaking research, but it needs to be done. It is difficult to get funding agencies to finance this work. This work requires coordination of different disease groups in an integrated manner. Also before models are used for pre-clinical studies, they need to be confirmed by an independent, certified behavior laboratory. Finally, a consensus conference of experts is needed to choose a few models to be used and to get funding agencies in agreement. Additionally challenges would be to achieve agreement among researchers and to supply the models. Because data need to be publicly available, strong input from the INCF in database creation would be preferable.

4C. Databases and Ontology:

Several important questions were discussed: How well are current databases prepared to reflect the complexity of the brain in terms of gene regulation and protein networks, including epigenomics? How can they be linked to behavior? How can behavior of models be
linked to human symptoms? Understanding progression of a disease in humans requires many patients, which need to be investigated with sophisticated methods over a long time. Do diseases progress in a similar way in animal models? Based on normal functional networks, how deep is our understanding, and how can the INCF contribute to brain diseases, particularly neurodegenerative brain diseases?

The workshop participants appeared to share the concern that it is very difficult to compare and integrate data from existing databases, because of different formats and rules for populating the databases and inconsistent use of ontologies. Although it seems impossible to change what already exists, new databases should draw upon accumulated experiences, and strive to adhere to systematic standards for their construction.

Discussion extracts

- We are missing ways of defining and comparing molecular phenotypes (between species) as well as systematic digital atlas resources for disease models. We need resources for identifying morphological differences between species. Because most neurodegenerative diseases in humans are age related, one should run longitudinal studies until late ages of the models. Although preferentially done in monkeys, this model is still not widely used and may have been avoided due to ethical concerns. However, digital atlases of morphology, histopathology, and gene expression should be built for rhesus monkey as well.
- Ontologies lack standards to describe phenotypes in models. Comparisons between species are fundamentally impossible. Links among altered cell types, neurophysiology, and phenotypes does not yet exist. The participants questioned whether the data of the Jackson Laboratory would be useful for a first step in this direction.
- It was suggested to introduce a standardized vocabulary as well as methodologies for animal models and the construction of translation tables between models and human disease.
- The question was raised whether it would be possible to visualize activities of particular neurons associated with conditioning experiments, for instance expression of genes in the habenula (in fish) might be related to psychiatric diseases. It was suggested that experiments could go back and forth between fish and human as a means to contribute to understanding psychiatric diseases.
- One should introduce dimensions for evaluation of species: axes, relevance, efficiency, phenotypes, and age. The reuse of model data has not really been foreseen. It was emphasized that differences among humans (many genetic subpopulations) and the many treatments used necessitate a robust translational bridge from model to human (work with types rather than subpopulations). Models have a lot of variation as well, making robust translation from models to humans difficult. Thus, as mentioned under point 1, above, use of reference populations to deal with this issue (for both humans, and models) is strongly encouraged. However, translation will be needed between populations.

- Anatomical atlases should be linked with gene expression, proteomic, etc. atlases and also with behavior and imaging. To this end well-defined, interoperable spatial reference systems (such as the mouse brain Waxholm space, and Thalairach coordinates in primates) are needed.
- “Google earth for the mouse brain” as a neuroscience information framework is needed.
- An important step would be to register all databases at one Internet site to support easier access to existing data (www.komp.org).

5. Concluding Remarks

Compared to recent studies in humans (large-scale genetic studies including whole genome sequencing, epigenomics, epidemiology, and sophisticated phenotyping with biomarkers related to the transcriptome and to imaging), characterization of animal models lags significantly behind. Although it is generally accepted not to use control individuals from a different population in human GWA (genome-wide association) studies, we still do not know precisely which strain of mice or rat is the best model for neurodegenerative diseases. Phenotyping of wild-type strains is lacking, as well as databases and developmental information. Technology is still developing. For imaging, tracer load is mostly experimental. The genomes of various species and strains within a species, and their genetic backgrounds are largely not yet available. Here one would need a combined effort to have all information on phenotyping and genome from all potential species and different genetic backgrounds (mice, rat, C. elegans, Drosophila, fish, and some monkeys) and a consortium which decides the primary focus. The most recent report on sex-specific parent-of-origin allelic expression in the mouse brain (Gregg et al. 2010, Sex-specific parent-of-origin allelic expression in the mouse brain. Science 329:682-5; Gregg et al. 2010, High-resolution analysis of parent-of-origin allelic expression in the mouse brain. Science 329:643-8) highlights the necessity to truly integrate genetic and behavioural data into a single database.
6. Workshop Program

Roles of Neuroinformatics in the process of building, evaluating and using genetic animal models for brain diseases.

Chair: Olaf Riess

PROGRAM

Sunday, December 13, 2009

12:00 – 13:00  Lunch and Welcome
13:00 – 13:30  Introduction of INCF and the Goal of the Workshop
              (Sten Grillner, Chair of INCF Governing Board)
              (Olaf Riess, University of Tuebingen)
13:30 – 14:00  What makes a good genetic model for human brain disorders
              (Olaf Riess, University of Tuebingen)
14:00 – 14:30  Use of genetically modified mouse model to study the biochemical
              mechanisms of neurodegeneration in Alzheimer’s and Parkinson’s
              (Dennis Selkoe, Harvard Medical School)
14:30 – 15:00  Use of non-human primate model to identify specific networks of
              gene expressions in areas of the developing cerebral cortex involved in the higher brain functions
              (Pasko Rakic, Yale School of Medicine)
15:00 – 15:30  Coffee
15:30 – 16:00  Zebrafish and mouse models of human Usher syndrome and zebrafish database to develop
              informatics support for zebrafish models of human diseases
              (Monte Westerfield, University of Oregon)
16:00 – 16:30  Zebrafish as a model animal to study neural circuits against fear
              (Hitoshi Okamoto, RIKEN Brain Science Institute)
16:30 – 17:00  A new paradigm for animal models in personalized medicine
              (Rob Williams, University of Tennessee)
17:00 – 17:30  Systems biology of activity-dependent brain plasticity
              (Gerd Kempermann, University of Dresden)

Monday, December 14, 2009

9:00 – 9:30  High content systems analysis of brain diseases
             (Guus Smit, Vrije University Amsterdam)
9:30 – 10:00 Anatomical phenotyping of transgenic models
             (Trygve Leergaard, University of Oslo)
10:00 – 10:30 Informatics and genetics programs supported by NIH
             (Jonathan Pollock, Division of Genetics, National Institute on Drug Abuse, National Institute
             of Health, USA)
10:30 – 11:00 The EC contribution towards elucidation of human disease using mouse models
             (Christian Desaintes, Directorate General for Research, European Commission and Co-chairman,
             the steering committee of the International Knockout Mouse Consortium)
11:00 – 11:30 Coffee
11:30 – 12:30 Discussion: state-of-the-art for genetic animal models for brain diseases and current role of
              neuroinformatics in the process
12:30 – 13:30 Lunch
13:30 – 17:00 Discussion:
              (a). future developments for genetic animal models and roles of neuroinformatics
              (b). potential contributions of INCF
7. Appendix

Sustainability Issues – summaries of speaker presentations

Numerous recommendations of the participants have been integrated into sections 1-5 of the report. Here, we provide more discussion of individual perspectives and views.

Hitoshi Okamoto

1: Establishment of a brain atlas for precise correlation of individual parts of the brain among different model animals

The triune brain theory has advocated that the brain has expanded its capacity by adding new parts during evolution. This process has culminated in the generation of the neocortex in the mammalian brain. However, the recent discovery of remarkable conservation among all vertebrates, i.e. teleost fish, birds, and mammals, of the expression patterns of the so-called tool kit genes, the essential genes that mastermind patterning of the brain during embryonic development, has challenged this prevailing view. Conserved expressions of these genes in the pallium and subpallium among different species has strongly suggested that the basic structures of the telencephalon derived from the pallium and subpallium, such as those corresponding to the basal ganglia, hippocampus, amygdala, and isocortex are shared more universally by all vertebrates than previously thought. This idea has forced us to think that many complex behaviors of animals may be regulated by neural circuits analogous to the circuit that controls the so-called higher brain functions of the human brain.

- This new view has suggested that we may be able to study the neural mechanisms of at least some of complex behaviours controlled by higher brain functions and the mental illness caused by malfunction of such by using non-mammalian vertebrates, which have smaller and simpler brains that are more easily amenable to genetics than mammalian brains.
- Zebrafish (Danio rerio) is now getting much attention as a newly emerging model animal for studies of neural circuit functions both in normal and impaired states, such as goal directed behaviour, decision making, choice of proper fear responses against aversive stimuli, and psychiatric diseases such as depression, PTSD, and drug addiction.
- The community of the people using zebrafish has already made great efforts to compile information systematically about the whole genome sequence, expression patterns of almost all genes and markers genes in transgenic lines, and phenotypes of the mutants. A new initiative to establish an international resource of lines with mutations in all genes is underway. The results of such efforts should be linked with the international neuroinformatics initiative.
- It is now imminently necessary to provide information that can facilitate comparisons among different model animals. Currently, it is not clear how each part of the zebrafish brain precisely corresponds to the parts of the mammalian brain. For example, it has only recently become clear from the study of gene expression patterns and afferent and efferent neural connections that the ventral habenula of the zebrafish diencephalon corresponds to the lateral habenula of mammals. Knowledge accumulated by such efforts should be annotated in the brain atlas of zebrafish to enable phylogenetic comparisons.

This recommendation should also be extended to other emerging model animals such as song birds.

2: Neural connectomics

Technology such as neural circuit genetics or optogenetics, with which we can conditionally inactivate or activate of any parts of neural circuits using genetic manipulation and viral transfection has opened an entirely new field of neuroscience that enables deeper understanding of brain function on the basis of causality rather than parallelism. These technologies are already feasible in various animals such as mice, rat, primates, and zebrafish. To make this technology more available, it is important to annotate each part (nucleus) of the brain map with information such as available Cre-lines in the case of mouse or enhancer trap lines in the case of zebrafish.

Efforts to establish whole-brain neural connectomics both at light and electron microscopic levels will soon require a new technology to handle the huge amount of information. Such technology is essential not just for professional researchers who integrate the information for reconstruction of 3-D images, but also for general users with only a limited capacity to access the large amounts of neural connectomics data through the conventional Internet.

3: Establishment of centralized site for researchers to find information around the world

Currently, many information sites are run by independent agents. Although we can find such sites with various searching software, it is difficult for non-experts to find which sites are useful for which particular purposes. For examples, modern neuroscience needs access to new technologies that are being developed continu-
ously. It would be convenient to provide centralized site where researchers can obtain advice as to which sites to visit to find information on particular subjects, such as optogenetics, virus vectors, etc. This may be possible by designating individuals to collect and update information regularly.

Dennis Selkoe

Long-term recommendations (difficult to achieve and require new funds):

1. INCF should strongly encourage standardization of methods and terms during construction of new databases for animal models internationally. Consider issuing a set of recommended guidelines and standards for animal database construction through a “white paper” written by 10-20 leading neuroinformaticians convened by INCF, with publication in a peer-reviewed journal.

2. INCF should work with other neuroinformatics efforts to lead the creation of a comprehensive database focused on one or two commonly used strains of healthy rats. Try to achieve a consensus as to which strain(s) to choose. The healthy rat database should include data on the genome (and possibly the epigenome), transcriptome, proteome, and neuronal connectome of the normal rat brain [organized by brain regions and later by discrete neuronal populations (nuclei)]. It could be complemented by a database of results of standardized behavioral tests performed on healthy rats at four ages: 3, 9, 18 and 30 months (if possible).

Short-term recommendations:

INCF should consider convening a consensus group of leading informatics-oriented investigators who use mouse models of human neurodegenerative diseases, especially for Alzheimer’s disease and Parkinson’s disease. The goal would be to evaluate critically which existing mouse models are the most compelling and useful and which should be deemphasized. A white paper could be issued, recognizing that some in the field may choose to disregard it (to their peril!). This workshop could also discuss the utility of rats for neurodegenerative disease modelling at this point in time.

Gerd Kempermann

Standard operating procedures (SOPs) are useful, when they serve the community to reach consensus about results from a particular experiment. Only with clear standards can the results from different groups become comparable and the work cumulative. On the other hand, too much focus on SOPs can prevent originality and steer a field into wrong directions. SOPs should also not become yet another type of artificial hurdle for publications. What journals and grant agencies should enforce are standardized descriptions of materials and methods (preferably based on ontology). The tendency of many major journals to neglect the importance of clear descriptions of experimental procedures needs to be counteracted.

At a very fundamental level the elementary data sets, be it gene arrays or proteomics data require careful and obsessively curated annotations. Otherwise many attempts for systems biology will ultimately fail. Exon arrays and deep sequencing, for example, generate information about great numbers of previously unknown splice variants. Posttranslational modifications further dilute the identity of a particular gene. So far, array companies have paid little to no attention to annotations, leading the existence of large numbers of incorrectly assigned probes and thus false conclusions. Such errors multiply, when based on such data the attempt is made to integrate across experimental domains and platforms.

With the rise of systems biology and the availability of enormous data sets spanning diverse domains of science, one single analysis might draw from results obtained from experiments based on very different, potentially conflicting assumptions. At the same time, with the change in scale, from molecular and below to organismic and cultural, new rules and “laws” govern and any attempt of synthesis is destined to generate conflicts or paradoxes. While it has been widely acknowledged that databases need to learn to “talk to each other” in order to integrate the many different available types of information, less attention has yet been given to the theoretical concepts of how such integration across scales is at all possible. Rules for this type of analysis need to be developed, otherwise the attempted integration will lead to spurious connections and the resulting relationships cannot be interpreted.

Monte Westerfield

One of the major impediments to developing animal models of human disease is the availability of information about human disease phenotypes. It is relatively straightforward to develop an animal model when the disease causative genes are known. However, developing animal models of human diseases with unknown genetic causes is frustratingly difficult. There have been a few recent attempts to make human disease phenotypes “computable”, including the Human Phenotype Ontology (HPO), the human disease ontology (DO), and the Phenotype And Trait Ontology (PATO), but to date, there is no systematic approach and, hence, very few
diseases have been annotated. The development of multiple seemingly parallel ontologies and methods suggests that standards are needed.

Similarly, there are no uniform guidelines for the types of assays used for phenotypic characterization of animal models of human diseases. Here, too, a lack of standards means that information obtained in one study may not be easily compared with data from other studies, nor can it be searched or analyzed computationally.

Recommendations:
1. The INCF should coordinate the establishment of ontologies and standards for describing human brain disease phenotypes.
2. The INCF should coordinate the establishment of recommended standards for phenotypic analyses of animal models of human brain diseases and for the metadata that describe how the assays were conducted, the units of measurement, data formats, etc.

Trygve B. Leergaard
Anatomical phenotyping of transgenic models

Transgenic rodents models of neurodegenerative disease are considered to be powerful tools for experimental investigation of fundamental neuropathological processes and for evaluation of potential therapeutic interventions. The value of such models is largely related to the degree with which a human disease phenotype is reproduced in the animals. As with human patients suffering from neurodegenerative disease, animal model phenotypes are typically characterized or ‘diagnosed’ by behavioral analyses, neuroimaging, and postmortem histopathological examination.

Although invasive experimental manipulations and detailed histological examinations can be routinely performed in the rodent brain, its relatively small size is a limiting factor for neuroimaging approaches. There are different motivations for determining the anatomical phenotype of a transgenic rodent model, and the choice of methods varies according to purpose. General validation of transgenic models is often based on histological detection of disease specific markers, such as e.g. regional neurodegeneration or protein accumulations. Exploration of underlying pathological processes at level of molecules, cells, or brain systems, will usually require more sophisticated experimental approaches. In context of evaluating disease progress and the effect of (therapeutic) experimental interventions, sensitive in vivo biomarkers are needed as benchmark readouts. To this end, positron emission tomography and different magnetic resonance imaging (MRI) techniques have been employed with variable success, with diffusion based MRI possibly appearing as the most promising. Thus, anatomical phenotyping approaches cover a wide range of methods and levels, which can be utilized in context of exploratory mapping studies or more focused hypothesis driven investigations. There are many parallels to human clinical investigations, but the lack of consensus about which criteria a transgenic animal should fulfill to qualify as a specific disease model, and the inevitable variations in phenotype, make neuro-anatomical investigations in transgenic animals more challenging. In contrast to traditional qualitative and descriptive anatomical characterizations in normal animals, phenotyping efforts in transgenic animals require quantitative readouts and higher numbers of experiments to allow statistical group comparisons. The combination of increased efforts and results with restricted validity, give added costs and risks as compared to similar investigations in normal animal.

Novel methods including high-resolution MRI, automated image analyses, histological image databases, and digital brain atlases will likely allow more efficient anatomical screening. It should therefore be possible to industrialize anatomical phenotyping to a certain degree, for example by establishing high-throughput pipelines producing numerical data about e.g. brain region volumes, specific cell numbers, or chemoarchitectonic features. However, given the multiple levels complexity in this field, specifically tailored hypothesis driven (multi-modal) investigations will remain fundamental, and should probably be conducted within consortia of collaborating experts.

Guus Smit
High Content Analysis

Identifying novel molecular targets in animal models of disease needs standardization and in depth analysis of multiple parameters. This can be achieved through High Content Screening approaches. HCS is not the same as High Throughput screening; HCS gathers data on many parameters, is information dense, and is medium to high throughput. The fact that HCS can be operationalized with well-described standard operating procedures makes it suitable for cross-lab, cross-experiment and cross-of animal model comparison. HCS can be used in systems biology approaches for a multilevel analysis, e.g., using Cellular assays, Proteomics analysis and Behavioral assays.

Cellular assays

HC cellular technologies, such as the ArrayScan (Celomics), are valuable functional screening intermediates
for neuronal genes e.g., in involved in neuronal growth and survival. HC cellular technologies can be used for many types of intracellular assays (e.g. protein translocation, receptor activation, etc). Through these technologies new molecular disease targets will come available which can be followed up for drug screening.

Informatics Needs: The data that come available at various sites needs standardized minimal data set descriptions. At present these are not yet routinely provided. In particular combinatorial tests for genetic modifiers would benefit from metadata data, data sharing and across platform operability. Also a central portal, gathering data from cellular screening stored in local databases, is not available.

**Synaptic proteomics**

The synapse is a unique neuronal organelle which function can be accessed through various different technologies. This is of relevance since synaptopathologies, diseases of the synapse, are found in all areas of neurological and neuropsychiatric diseases. The synapse has experimental and computational challenges. For these it is necessary to establish quantities, stoichiometries, interactions of synaptic proteins through new mass spectrometric and proteomics technologies.

Informatics Needs: Synaptic proteins need to be annotated and databased (G2C, SBS-DB, etc). Through European funding of synapse-focused consortia (e.g. Eurospin, and SynSys) it is expected that the basis for adequate synaptic protein ontology will be laid. A European database for the synapse will be generated where future development will be towards integrating different types of information (e.g. molecular, physiology, perturbations).

Synapses have a unique composition, depending on various dimensions, e.g., genotype (animal model), brain area, age, stimulus, etc. As such, data on synaptic proteins should relate to spatial info (e.g. by linking it to a brain atlas), should have proper metadata description for genotype, disease mutation, brain area studied, protocols for studying molecules and physiology.

A specific area of interest will be to map the dynamic properties of synapses, e.g., changes in phospho-proteome signatures or stimulus-dependent dynamics of protein complexes. This requires description of types of stimuli (protocols), which again needs metadata description. Most of the aforementioned requirements are still in the early days of development.

A systems approach of the synapse requires coupling to synaptic physiological data obtained. This type of integrated analysis of data does not exist yet, and needs specific tools to be developed.

**Behavioral analysis**

Behavioral analyses need rigorous description of operating procedures, and thus metadata deposition. Automated home cage behavioral analysis (high content) is coming up as a new promising tool, as it is minimally depending on human intervention compared to stand alone testing (comparative analysis).

Informatics Needs: A specific requirement for high content behavioral analysis is the use of multivariate statistics on long-term multiday behavioral data sets.

There is local storage of data and for instance for genetic resources there is GeneNetwork available for correlative analysis and data sharing. Standardization of methods is still not complete, even given various EU projects in this area. Specific need is the central deposition of behavioral (meta-)data.

**Informatics challenges 2010-2020:**

A specific challenge for informatics will be how to deal with ambiguity of data to facilitate data exchange. In order to allow assessment of data quality the ‘raw’ data needs to be available. In various areas there is a strong need for high quality central storage. In some areas different databases will be developed in parallel. In both cases efficient search tools, consistent data formats, error checking and integration between databases is of utmost importance. A real challenge for the informatics community will be to integrate data from different technologies gathered at different levels of analysis; this is maybe the biggest and most important hurdle to take. Finally a serious threat to data gathering and management is data loss after funding ceases. Therefore, long-term funding strategies for integration portals and databases are essential.

**On anatomical phenotyping of animal models (TBL1)**

Approaches to characterize the anatomical phenotype of animal models span a wide range of levels and methods, the choice which depends on the context, model and questions to be addressed. Anatomical, or neuropathological, descriptions are typically needed for a) validation of reproduction of disease-specific traits, b) exploration of pathological processes underlying the disease, and c) specification of (in vivo) biomarkers that can be used as readout for disease progression or stage. Compared to anatomical investigations in ‘normal’ animal populations, the use of genetically modified models introduces additional complexity and additional unknown parameters. This can relate both to variability of gene expression, disease phenotype, and comorbidity. Further, as traditional journal reports typically describe the results of hypothesis driven research in text supplied with selected representative illustrations, only
fraction of the obtained materials are available. As life events and normal aging can have profound influence on phenotype expression, it is highly important that relevant biographical information are provided and that the effect of aging is mapped in some reference populations. Together, these factors challenge the planning and interpretation of experimental research on transgenic animal models. Wider access to more complete underlying (raw) data, including reference materials and negative data, could contribute to reduce these problems (see also Boline et al., 2007; Nature Precedings: doi:10.1038/npre.2007.1046.1).

Recommendations (pertaining to anatomical phenotyping)

General requirements for all brain-derived data should be that:

- Underlying raw data are shared, including ‘negative’ or ‘normal’ descriptions, if possible full (whole brain) data should be included.
- Minimal metadata are provided (should be specified, cfr. recommendations in the field; see, e.g. Gibson et al., 2008: proceedings.nature.com/documents/1720/ version/1).

There is a need for:

- As neurodegenerative diseases occur with increasing age, a comprehensive mapping of aging animals (rodent) is needed to describe normal aging processes.
- Recommendations for minimum workflows for validation / phenotyping of transgenic models.
- Consensus in the field about certain models can give necessary focus to obtain full descriptions.
- All reported data should be related to a defined 3-D framework (standard atlas space, Waxholm space for mice).

Because assignment of spatial information is essential, it might be possible to delineate some standards or levels of increasing precision:

1. Text descriptions should employ recognized anatomical ontology / terminology (see e.g. Bota and Swanson, 2007; www.incf.org/documents/workshop-reports/ incfworkshop-1st-NeuroanatomicalNomenclature.pdf).
2. For tomographical / histological data, slice / section orientation should be defined relative to atlas space (such as flat brain skull position or Waxholm space).
3. Data should be assign standard reference brain 3-D coordinates.
4. If possible, data should be registered / warped / normalized to a 3-D reference space.

Olaf Riess

We are concentrating on two major issues related to animal models of neurodegenerative diseases: (1) what are the best models and how to generate these, and (2) how do we design and perform preclinical treatment studies to achieve most helpful results for subsequent human clinical trials.

In several ways, animal models have considerable advantages over human patients for studies of brain diseases: (i) one can perform numerous treatment studies in parallel, even if the drug is not optimized for toxicity, (ii) neuropathology can be studied in detail during initiation and progression of the disease, which is basically not possible in humans, (iii) new markers, such as proteins or antibodies, that might be helpful in patients for diagnosis can be tested (e.g. see synuclein labeling of Lewy bodies in PD), (iv) neurotransmitters can be studied more easily, and (v) even for imaging, despite the small brain and resolution of the images, new tracers can be developed.

For issue (1), model generation, currently BAC overexpression is typically used to mimic the expression pattern of the human gene more closely. Also, knockin technology is frequently used in several neurodegenerative diseases such as AD and PD. Currently there is a lack of comparability of existing models. Here, INCF could stimulate initiatives to investigate expression levels, regions of expression, cellular expression, neuropathology such as degeneration of aggregates, stability of phenotype, and progression among the different models. Standardization of behavioural or neuropathological readouts do not exist, also there is no grading system as is usual for human patients. First recommendations for phenotyping strategies of mice (Crawley 2008) and of a scoring system to phenotype transgenic rats (Korenova et al. 2009) have been made without general agreement on which terms to use.

Contribution of the INCF: Stimulating comparative investigation of disease models, discussing the merits of each species, development of standardization, development of databases for animal models, and communicating the importance of standardization and comparability to funding agencies.

The second important issue is how to measure efficacy of preclinical treatment trials in diseases for which basically no effective treatment exists and how to adapt models and read-outs for future human trials. Imaging is widely available for mice, rats, and also for nonhuman primates. Tracer concentrations are sometimes not measured, and resolution and variance of the models are still issues (as in humans). Models can be used
to test new tracers before use in humans. However, no standardization exists and no guideline about whether imaging analyses such as MRI, PET or DTI-MRI should be included in descriptions of the models. Repetitive studies are not required (as they are in human trials) even though reproducibility is always an issue. For phenotyping, the question remains how comparable conventional phenotyping is compared to automated phenotyping. What are specific readouts in animals for symptoms in human patients (e.g. we lack a choreatic mouse or rat). This needs to be defined before initiating comparative preclinical trials. However, for some diseases, such as ALS, guidelines for preclinical drug interventions have been established by a consortium of experts (Ludolph et al. 2007). For each model it needs (1) to be discussed, which biomarker can be used to monitor treatment success, (2) to be defined which specific readouts are wanted, which are reproducible, how to reach standardization, and what scoring system should be used.

Contribution of the INCF: Bring groups of experts together to develop guidelines for preclinical testing of models for specific diseases. Develop databases and encourage editors of scientific journals to accept “negative” results. Work towards standards that provide a baseline for publishing animal models of neurodegeneration.

A third point is which species should be used for neurodegenerative disease studies. Each species has advantages and limitations. Each model needs to be evaluated for use as a mimic of particular aspects of a disease or for use as a broadly accepted model. In addition to mice, transgenic rat can be used as models of neurodegeneration. In the field of neurodegeneration, rats have some advantages compared to mice for neurophysiological studies, for neurotransplantation, for some behavioural tests such as learning, for continuous biomarker sampling such as blood and liquor, for higher resolution in vivo imaging, and for pharmacogenomics because rats respond differently than mice.

Contribution of INCF: Bring experts in rat models of neurodegeneration together to discuss these issues.

Christian Desaintes

The European Commission’s contribution towards the elucidation of human disease using mouse models

During Framework Programme 6 (FP6), covering the period 2002-2006, the European Commission has been supporting a wide-ranging and substantial programme on mouse functional genomics involving more than a dozen large-scale multi-year projects. Some of these projects (such as MUGEN, EVI-GENORET, EuroHear, etc.) are using the mouse as a model to understand fundamental biological processes relevant to human health and diseases, while several others (EUCOMM, EUMODIC, EURExpress, EMMA, etc.) are generating comprehensive resources for the scientific community that should ultimately allow the functional annotation of the mammalian genome.

Many of these projects are centred on the programme of large-scale mutagenesis of coding genes. They encompass molecular phenotyping at the tissue level, the development and improvement of technologies for high-throughput mutagenesis, the production of conditional mutations in ES cells, the derivation of mutant mice and their phenotypic determination, and finally the archiving and distribution of resources (vectors, mutant ES cells and mice).

The European mutagenesis programme is coordinated with similar efforts in North America grouped under the International Knockout Mouse Consortium (IKMC). This initiative includes three funding agencies (European Commission, Genome Canada, and the United States National Institutes of Health) and their funded mutagenesis programmes (EUCOMM, NorCOMM, KOMP) as well as the trap resources that have been generated by the Texas Institute of Genomic Medicine. In FP7 (covering the period 2007-2013), the European Commission continues to support several projects where the mouse plays a key role in elucidating the complexity of basic biological processes at a system level as a way to understand the molecular basis of human development in health and disease (among which several neurological disorders). The European Commission also continues to support the IKMC with the funding of five projects currently (I-DCC, CREATE, Infrafrontier, EMMA service and PhenoScale) that promote the best use of the IKMC resources.

A call for proposals was recently published for possible funding of an additional large project that should valorise the IKMC mutant resources and make them available to the biomedical scientific community as a unique and comprehensive tool to address the role of each gene in development, health, and disease. Despite this considerable effort, further ambitious actions, such as the systematic phenotyping of mutants, could be considered in a worldwide endeavour, to fully capitalise on the investments made in the IKMC, thereby providing the scientific community with an invaluable resource for understanding human health and disease. A conference might also be organised in spring 2010 with the aim of demonstrating the complementary nature of mouse models and medical research on human diseases (e.g. neurodegeneration, cancer, cardiovascular disease, diabetes). A strong focus of this conference should be on the identification and discussion of mutual requirements to unravel the molecular basis of the disease mechanism.
References:


