



Plenary Speakers

Pages 32 - 42

Abstracts PL-1 – PL-9

THE MOUSE GENOME – DNA TO SYSTEMS BIOLOGY

Ewan Birney

European Bioinformatics Institute, Hinxton, UK

The genome is both the source of the information of an organism and the natural index for much of biology. Ensembl provides an infrastructure for handling the genomes of large organisms, making both accurate predictions of functional regions of the genome and also using the genome as an index to correlate genes, microarray experiments and genetic modifications. Because of the ease of genetic manipulation in the Mouse, the Mouse genome has become the richest source of functional information across all mammals.

In my talk I will expand on the prediction of new functional elements in Ensembl, in particular RNA genes (including miRNAs) and regulatory regions. I will show how integrating large scale functional experiments (eg, Chip on chip) can be coordinated with these other features. I will also expand on the growing comparative genomics resources in vertebrates and how this informs our understanding of mouse biology

A DECADE WITH THE MOUSE GENOME ENCYCLOPEDIA

Yoshihide Hayashizaki

Laboratory for Genome Exploration Research, RIKEN Genomic Sciences Center, Yokohama, Japan

I dedicate this lecture to the research history of our field, the beginning of the path and what we will have to do to follow it further, all in the memory of he who has been forced to step off the road, Verne Chapman, and the collaboration that created the international FANTOM consortium.

Before whole genome sequencing was a realistic concept, we knew that genes were connected to phenotypes and we used positional cloning to realize these relationships. But as our knowledge of the genome grew, we also noted how little we actually knew. We realized that there are two complementary information sources in the genome, the actual sequence and the transcriptome.

Together with Verne, we embarked on a giant quest in genome research. We both shared the belief that technical innovations are necessary for any scientific discovery and embraced the new technology involved in genome scanning and genome mapping, we identified the U2rfp, Reeler, Syrian Hamster and the Grf1 genes. But we couldn't stop there, we discussed a Japanese cDNA project, development of new technologies for full-length cDNA cloning and high-throughput sequencing technologies (RISA), and we discussed the establishment of a platform of knowledge for the science of the 21st century, consisting of clones and databases.

Unfortunately our discussions ended too early, Verne passed away during his last visit to RIKEN, his last commandment was "Mouse is the best sample for cDNA collection, use the C57BL/6 strain".

In keeping with this, FANTOM 1,2 and 3 was organized to annotate genes in mouse and RIKEN compiled the Mouse genome Encyclopedia. From then on, we have continued our quest into the mysteries of life. We have discovered unexpectedly large number of non-coding RNAs and sense/antisense pairs to occur virtually all over our genome and we have found our selves tottering on the brink of the vast RNA Continent, stumbling with the insight that RNA is no longer to be considered only as a passing stage in the protein synthesis, but also as a biologically active molecule on its own.

I give Verne Chapman the credit to have found the beginning of our path, now it is up to us to light the torch of knowledge and follow it all the way to the end.

GENETIC DISSECTION OF RETINOID SIGNALING THROUGH CELL-SPECIFIC TEMPORALLY-CONTROLLED TARGETED MUTAGENESIS

Pierre Chambon
(Institut Clinique de la Souris, Illkirch Cedex, France)

The completion of the human and mouse genome projects has revealed the presence of approximately 30,000 genes in vertebrates and many more encoded mRNAs and proteins. The next formidable and most challenging task is to determine the physiological and pathophysiological functions of these genes within living organisms. This will be achieved by interfering with the expression of individual genes using genetic and pharmacological approaches. In this respect germ line targeted mutagenesis through gene disruption (knock-out) in murine embryonic stem (ES) cells has been widely used during the last decade. However, there are a number of limitations (embryonic lethality, functional redundancies between related genes, compensatory mechanisms during development, impossibility to discriminate between cell-autonomous and non-cell autonomous events) that preclude a simple genetic analysis of gene function in vivo through conventional gene knock-out approaches.

In fact, only targeted spatio-temporally-controlled somatic mutations that are generated in animal models in given cell-type/tissues and at chosen times during pre- and post-natal life, can lead to the functional dissection of complex regulatory networks, thus revealing the physiological and pathophysiological functions of the genes involved. To generate such somatic mutations, we have developed a new genetic tool based on the Cre/lox system. A conditional ligand-inducible Cre recombinase, Cre-ERT2, has been generated by fusing Cre with a mutated ligand binding domain of the human oestrogen receptor (ER) that binds the synthetic ligand tamoxifen, but not estradiol. Upon tamoxifen administration the Cre activity is induced, and “floxed” DNA segments flanked by loxP sites are efficiently excised.

I will illustrate this somatic conditional mutagenesis approach by describing our recent mouse studies that are aimed at dissecting the retinoid signaling pathway, a complex regulatory network involved in vertebrate morphogenesis, organogenesis, growth, cellular differentiation and homeostasis.

PROGRESS ON ASSEMBLING AND ASSESSING THE C57BL/6J MOUSE GENOME SEQUENCE

Deanna M. Church, Richa Agarwala, Jack Wang, Yuri Kapustin, Wratko Hlavina, Donna Maglott, Paul Kitts
(DHHS/NIH/NLM/NCBI; Bethesda, MD)
Stacie Loftus, William J. Pavan
(National Human Genome Research Institute, Bethesda, MD)
Evan Eichler, Ginger Cheng, Eray Tuzun, Xinwei She
(HHMI, University of Washington, Seattle, WA)
The Mouse Genome Sequencing Consortium

The initial draft sequence of the C57BL/6J mouse genome was based on a Whole Genome Assembly (WGA) strategy, with the actual assembly referred to as the MGSCv3. Over the past two years, substantial progress has been made in obtaining clone based sequence and moving toward a finished genome assembly. Several intermediate genome assemblies combining the clone based sequence with the WGA have substantially improved our understanding of the mouse genome. A description of mouse Build 35, which consists of >80% finished sequence, will be presented. The data in this build is based up sequence data available as of Aug 5, 2005.

The process of producing the combined assembly provides us with some information with respect to the quality of the WGA when compared to finished sequence. To more formally assess this, we have aligned the MGSCv3 to mouse builds 34 and 35. Preliminary assessment of these alignments suggests that in general, the WGA is in good agreement with the current finished build. However, there are regions of the current build that do not align well to the initial WGA. Roughly 2% of the MGSCv3 did not align to the Build 34, while approximately 5.4% of the Build 34 sequence was unique with respect to the WGA. In addition, ~37% of contigs that were unmapped in the MGSCv3 can be assigned a unique chromosome location in Build 34. Genome wide analysis of the assembly comparisons using Build 35 will be presented, as well as some specific examples of regions where the clone based sequence is not well represented in the MGSCv3. We will also present additional assessment of Build 35 based on the analysis of Reference Sequence transcripts that do not align to the current build.

It is clear that many of the regions in the finished portions of the current assembly that do not align well to the WGA assembly contain regions of segmental duplications. Assessment of the MGSCv3 suggests that only about 1-2% of the mouse genome is involved in segmental duplication. Assessment of more recent builds has driven this number higher, with estimates of the amount of segmental duplication currently ranging from 3-4%. This number could increase as the genome nears completion.

Lastly, in addition to providing the reference C57BL/6J genome, we have been producing alternate assemblies of finished sequence from other strains, when available. While there is only a small amount of non-C57BL/6J sequence comparison of these to the reference can provide information concerning large-scale genomic polymorphism. Information concerning the comparison of the reference assembly to data from other mouse strains will be presented.

DNA MICROARRAY FOR STEM CELL RESEARCH

Lars Martin Jakt, PhD Takumi Era MD PhD and [Shin-Ichi Nisikawa](#)
Stem Cell Research Group, Center for Developmental Biology, Chuo-ku Kobe, Japan

Our group has been interested in two types of stem cells: resting stem cell(SC) in the adult body such as melanocyte SC and hematopoietic cell, and constitutively self-renewing embryonic stem cells such as ES cells or mesenchymal stem cells. As an approach to investigate molecular mechanisms supporting stem cell systems, we have spent some years to establish a DNA microarray data-base of stem cells and subpopulations derived from various stem cells. The most important issue of this project is how to collect cell populations for analysis. Through efforts over years, we succeeded to purify more than 100 populations and subjected to DNA microarray analysis, that have developed to a continuously expanding database of transcription based on Affymetrix oligonucleotide arrays. In order to handle the data generated in this process we have developed novel analysis systems that allow both analysis and dissemination of the data. Importantly this system allows the data to be accessed concurrently by a number of researchers specializing in various aspects of stem cell research, thus allowing the biologically relevant information to be extracted.

In this workshop, I will describe an ability of our data-base and analysis software by addressing a couple of issues of stem cell research such as stemness genes and fate determination. Secondly, we will describe some genes that were discovered by using our data-base. Indeed, these examples strongly suggest a potential of this data-base as a platform of functional genomics.

THE GENETICS OF COLINEARITY IN MAMMALS

Denis Duboule, Marie Kmita, François Spitz and Basile Tarchini
Faculty of Sciences, University of Geneva and National Research Centre 'Frontiers in Genetics',
Geneva, Switzerland

Mammalian *Hox* genes are required for the proper organization of axial structures, in particular the trunk and the limbs. In these various contexts, combinations of HOX proteins instruct cohorts of cell regarding their developmental fates, and mutations within *Hox* genes lead to alterations in these important patterning processes. Such combinations derive from precisely orchestrated transcripts distributions, which themselves require an appropriate and coherent regulation. This is in part achieved via an enigmatic property, referred to as 'colinearity' whereby the space and time of *Hox* gene activation is reflected in their topographic organization along the DNA clusters.

In order to understand the underlying mechanisms, we have developed two strategies (*TAMERE* and *STRING*) to produce a range of modifications involving the *HoxD* cluster, such as deletions, duplications and inversions, relying on mere genetic approaches. The effects of these rearrangements upon gene transcription were systematically analyzed and allowed to propose a mechanistic model for colinearities during limb development, based upon the existence of different global enhancers located on either sides of the cluster. This model will be discussed in both ontogenetic and phylogenetic contexts.

CHROMATIN AND NUCLEAR RE-ORGANISATION DURING DIFFERENTIATION AND DEVELOPMENT IN THE MOUSE

Céline Morey¹, Anne Wiblin¹, Clémence Kress¹, Marie Kmita², Denis Duboule², and Wendy Bickmore¹

¹MRC Human Genetics Unit, Edinburgh, UK; ²Department of Zoology and Animal Biology, University of Geneva, Switzerland

Chromatin structure and its modification are central to epigenetic mechanisms that direct patterns of gene expression during development. Whereas there have been many studies of the links between transcriptional regulation and nucleosome modifications, chromatin organisation at and beyond the level of the nucleosome is poorly understood. We have revealed Programd changes in large-scale chromatin condensation and nuclear organisation that occur during the differentiation of murine ES cells. This occurs both at the level of global nuclear organisation of heterochromatin, and at specific loci. In particular, there is a decondensation of chromatin, and a re-organisation of chromosome territories, at *Hox loci* (*Chambeyron and Bickmore, 2004, Genes and Dev*). This supports the idea that there is a progressive 'opening' of chromatin structure propagated through *Hox* clusters from 3' to 5', which contributes to the colinear sequential activation of gene expression. These events also occur during embryogenesis. The first changes in chromatin structure and nuclear organisation are detected during gastrulation in the *Hoxb1* expressing posterior primitive streak region. (*Chambeyron et al., 2005, Development*). Later, at E9.5, the differential nuclear organisation of *HoxB* along the anterior-posterior axis of the developing neural tube is also coherent with colinear *Hox* gene expression.

We will discuss the cis- and trans-acting factors that contribute to re-modelling of chromatin structure and nuclear organisation. Using transgenic mice in which *Hox* clusters have been re-arranged, we will discuss whether nuclear re-organisation at *Hox* loci can be uncoupled from transcription per se.

FORWARD GENETICS IN *DROSOPHILA* REVEAL A PROTOTYPICAL INNATE DEFENSE SYSTEM

Jules A Hoffmann, J M Reichhart, D Ferrandon, J L Imler, C Hetru
Insitut de Biologie Moleculaire et Cellulaire, Strasbourg, France

Insects have long been known to be particularly resistant to microbial infection. In the 1980ies it was understood that a hallmark of the defense reactions is the challenge-induced synthesis of a battery of potent small-sized cationic peptides directed against bacteria and/or fungi. *Drosophila*, in particular, transcribes genes encoding seven distinct peptide families active against Gram-positive or Gram-negative bacteria and against fungi. The bulk of the antimicrobial peptide synthesis occurs in the fat body, a functional equivalent of the mammalian liver.

The genetic analysis of the control of gene expression during the immune response of *Drosophila* has unexpectedly shown that the gene cascade directing dorsoventral patterning in the early embryo also serves for immune gene expression. In particular, the Toll receptor which plays a pivotal role in dorsoventral patterning appeared to also control response to fungal and Gram-positive bacterial infection. The defense against Gram-negative bacteria however was found to rely on an independent signaling pathway evocative of the intracellular cascades activated by the TNF-receptor in mammals. A series of mutagenesis screens have now led to a relative comprehensive picture of the response to these infections, from the receptor molecules to the subsequent signaling cascades controlling the expression of effector molecules. Some of these studies paved the way for the discovery of equivalent players in innate immune defenses in mammals. Overwhelmingly, they show striking similarities between immune reactions in insects and innate immunity in mammals pointing to a common evolutionary ancestry of these systems. Studies in the field of insect immunity are now being extended to antiviral and antiparasitic defenses and forward genetics are again a method of choice in this endeavour.

CO-ADAPTATED ALLELES AND THE EVOLUTION OF COMPLEX TRAITS

Gary A Churchill
The Jackson Laboratory

Why are most common disease related traits genetically complex? In this talk, I will consider complex trait genetics from an evolutionary perspective. I will argue that the genetic factors determining complex disease inheritance may be fundamentally different from those involved in Mendelian disorders. Adherence to the concept of disease susceptibility loci may limit our ability to discern the genetic loci and mechanisms that are responsible for variation in common disease-related traits.

Allelic co-adaptation is a mechanism that may be implicated in many instances of complex disease inheritance. I will provide examples from QTL mapping in the mouse to illustrate this point. Our studies of a complex type II diabetes model provide an example in which a mechanism of co-adaptation can be described at the molecular level. I will present evidence that loci with co-adapted alleles may be widespread throughout the laboratory mouse genome. The mixing of diverse subspecies and intense selection during inbreeding have produced a genome wide signature of selection for co-adapted alleles.

These observations suggest that our efforts to identify the “bad” alleles responsible for complex disease traits should be augmented with approaches that consider allelic combinations at multiple loci as the genetic cause of many complex disease traits.

INVOLVEMENT IN THE COMPLEX TRAITS CONSORTIUM

David Threadgill
Department of Genetics, University of North Carolina

The complex trait consortium (CTC) was originally founded in 2000 as a satellite of the International Mammalian Genome Society to plan and develop resources that would enhance the prospects of cloning genes associated with quantitative trait loci (QTL). Although called a consortium, the CTC has no formal membership or organization structure but instead provides an open forum for exchange of ideas and for planning community projects targeting complex traits. More recently, the CTC has begun to publicize the need for mammalian systems genetics, a variant of systems biology that is anchored on the underlying genetic variation present among individuals. As part of this effort, a new resource called the Collaborative Cross (CC), has been designed to complement ongoing single-gene mutagenesis programs. The CC, currently being developed at three sites, Oak Ridge, TN (USA), Perth (Australia), and Nairobi (Kenya), is an 8-way recombinant inbred (RI) panel that encompasses a large proportion of the genetic variation present in *Mus musculus*. This session will highlight the background of the CTC and its annual meeting each spring, review the status of the CC and provide a series of short discussions detailing how the cross was designed, how it is being generated, and how others in the IMGS can participate. Additionally, the utility of the CC for mammalian systems genetics and other CTC initiatives like the Wellcome Trust SNP project will be described. Details on the CTC can be found at www.complextrait.org

