Meeting report

Genomics meets nanoscience: probing genes and the cell nucleus at 10^{-9} meters

Joan Politz* and Ana Pombo†

Addresses: *Department of Biochemistry and Molecular Biology, University of Massachusetts Medical Center, 377 Plantation Street, Worcester, MA 01605, USA. †MRC Clinical Sciences Centre, Imperial College School of Medicine, Hammersmith Hospital Campus, Du Cane Road, London W12 0NN, UK.

Correspondence: Ana Pombo. E-mail: ana.pombo@csc.mrc.ac.uk

Published: 25 February 2002

Genome ***Biology*** 2002, **3(3)**:reports4007.1-4007.3

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2002/3/3/reports/4007

A report on the Jackson Laboratory 'Genomics meets nanoscience' meeting, Bar Harbor, USA, 9-12 October 2001.

This meeting brought together scientists from areas of research ranging from genome-sequence analysis, gene expression and organization in the cell nucleus to nanolithography and high-resolution imaging methods. It was recognized that advances from ongoing genome-sequencing projects and three-dimensional analysis of cell nuclei may extend our understanding of gene expression beyond just a few well-studied examples to give a much more comprehensive view of how the genome is organized within the cell nucleus. Research in these areas was predicted to lead to breakthroughs in the study of genome plasticity (for example in stem cells), cancer (and other conditions in which proper gene expression becomes impaired), and gene therapy.

Genome organization and gene expression

Carol Bult (Jackson Laboratories, Bar Harbor, USA) began the meeting by giving an overview of the status of genomic sequencing, management of databases, and structural analysis. The vast amount of sequence data to be analyzed requires versatile tools that can be used to recognize similarities between single sequences as well as to identify interactions among linkage groups, both long-range and in three dimensions. She described current efforts using proteomics to help understand gene regulatory networks and the different sets of expressed proteins in various cellular and tissue structures and reminded us that only about 1-3% of the human genome is thought to code for proteins. Charles Lawrence (Wadsworth Center, Albany, USA) explained the use of computational algorithms to identify promoters in large sequence databases using cross-species comparisons and reported that these methods can correctly identify promoter sequences in *Escherichia coli*. Ross Hardison (Pennsylvania State University, University Park, USA) explained the use of sequence-alignment programs to identify important non-coding genomic regions. As much as 38% of the mouse and human genome is conserved non-coding sequence and is found in relatively evolutionarily stable clusters within genomes, suggesting that these sequences are important 'non-junk' DNA.

Tom Maniatis (Harvard University, Cambridge, USA) then brought the discussion to a more molecular biological level by discussing the pattern of protocadherin gene expression in the human brain. Protocadherins are members of the immunoglobulin superfamily of proteins, and their genes are organized and rearranged like immunoglobulin genes, with 'constant' and 'variable' regions. His work used reverse-transcriptase PCR on individual neurons to achieve a remarkably high level of resolution and suggests that trans-splicing generates cell-specific protocadherin mRNAs. This result initiated discussions about whether the relative positions of the protocadherin loci within the nucleus might be important for the trans-splicing mechanism. Douglas Engel (Northwestern University, Evanston, USA) next discussed how interactions between distant genomic sequences might control gene expression. The best known example is the β-globin gene cluster, which contains a locus-control region separated by about 50 kilobases (kb) from the globin genes that are expressed at different stages of development. Various models have been proposed to account for the regulation of globin transcription by the distant locus-control region: DNA looping, DNA linking and DNA tracking. Engel gave other examples of long-range regulation, noting that this type of regulation is often associated with a distinct class of

genes that are expressed in regulatory hierarchies during development. He pointed out that few co-regulated clusters have been studied in as much detail as the β-globin cluster, and more needs to be done to understand how genes in different chromosomal environments may be regulated.

Genome organization in the three-dimensional nucleus

Thomas Reid (National Cancer Institute, Bethesda, USA) discussed the results of karyotypic analysis of various mouse cancer cell types, showing that certain chromatin segments are translocated and amplified in a pattern characteristic of the cell type. It was clear that knowledge of the three-dimensional arrangement of the genome within the nucleus before and after translocation would be very useful, given that the location of the amplified segments might influence their activity. The imminent establishment of a repository at the National Cancer Institute [http://cgap.nci.nih.gov/Chromosomes/CCAP_BAC_Clones] of 3,000 bacterial artificial chromosome (BAC) clones of average length about 1 megabase (Mb), covering the whole human genome, will be a valuable resource for achieving this aim. Siegfried Janz (National Cancer Institute, Bethesda, USA) discussed how the organization of gene loci within the nucleus might directly affect the generation of translocations that lead to disease. He showed that the t(12;15) translocation in mouse (of which the equivalent translocation in humans is directly correlated with Burkitt's lymphoma) seems to be favored by physical proximity within the nucleus of the two loci at the break points - immunoglobin H (*IgH*) and *c-Myc*. The two homologous *IgH* loci seem to differ in their compaction in the same cell and between different cell types (pre-B cells show more open domains, as seen by spectral precision distance microscopy; see below), and *IgH* loci appear close to *c-Myc* loci in normal B cells. A consensus view was expressed throughout the meeting that knowledge of the three-dimensional disposition of individual genes and the mechanism(s) that control their localization will be essential for a complete understanding of gene expression, as well as to an understanding of changes taking place in disease.

The current techniques for mapping chromatin disposition in the cell nucleus, including modeling algorithms, were discussed by Gregor Kreth (Kirchoff Institute for Physics, University of Heidelberg, Germany) and Jorg Langowski (German Cancer Research Center, Heidelberg, Germany). Although it has been known for some time that chromosomes occupy specific territories within the nucleus, it has been difficult to characterize these in the living nucleus, so information about their potential movements or reorientation during changes in gene expression is very limited. There was substantial discussion of correct ways to perturb living cells in order to study changes in spatial organization. Physicists were especially interested in brainstorming ways to monitor gene expression using inert nanoprobes of various types. The need for both better imaging systems and better data-handling and analysis tools was discussed, along with the limits of resolution inherent to light microscopy (only structures separated by more than half the wavelength of incident light can be resolved).

Two examples of the new advances in imaging were highlighted by Christoph Cremer (Kirchoff Institute for Physics, University of Heidelberg, Germany). Spatially-modulated illumination microscopy, which involves modulated interference of the light from two confluent lasers at the focal point, is allowing determination of the sizes of objects with diameters between 15 and 150 nm labeled with a single fluorochrome. Other parameters, including volume, surface area and roundness, may become measurable *in situ* at the scale of molecular machines (for example, ribosomes are about 30 nm in diameter). Spectral precision distance microscopy allows high-precision monitoring of the positions and distances between objects with distinct spectral signatures, currently with resolution down to 30-50 nm. Furthermore, promising results from virtual microscopy suggest that the distance between objects may be decreased to only a few nanometers. The physicists at the conference also contributed interesting information regarding another hurdle in the use of light microscopy to study the live cell nucleus. Currently, it is difficult to define the absolute dimensions of very small regions within a nucleus during imaging because there is no good reference that can be used to visually calibrate these regions. Michael Grunze (University of Heidelberg, Germany) described the potential of nanolithography to help with this problem, pointing out that 20 nm lines, which could act as such a reference, can be accurately etched using chemical lithography techniques.

Other techniques that provide higher levels of resolution were discussed by two speakers. Aaron Lewis (Hebrew University of Jerusalem, Israel) described progress in the development of near-field microscopy for biological applications. With this technique, resolutions are highest (objects as small as 10 nm can be resolved) in the *z* axis, the dimension for which one usually obtains lowest resolution using more standard techniques. Rasmus Schroeder (Max Planck Institute for Medical Research, Heidelberg, Germany) discussed the use of 'energy-filtering' transmission electron microscopy to study macromolecular structures within the nucleus. This technique allows one to construct molecular models from a small number of images of purified structures, but advances in sample preparation will be needed in order to study the structure of intranuclear bodies.

Nuclear metabolism and biophysical properties

Molecular machines and other nanoscale structures and their study were the subject of a number of talks. Chris Woodcock (University of Massachusetts, Amherst, USA) presented the concepts of 'high-precision' and 'low-precision' structures:

polymerases, ribosomes, and nucleosome core particles are examples of well-defined (high-precision) structures, whereas nucleosome arrays and chromosomes may be better thought of as low-precision nanoscale structures, because their structure cannot be determined to nanometer precision by averaging images of the individual units. He went on to show how high-fidelity transcriptional repression can be attained by a low-precision mechanism, that is, a mechanism involving low-precision structures, using the *STE6* mating-type gene in *Saccharomyces cerevisiae* as a model system. Cryo-electron microscopy was used to determine the structure of the *STE-6* gene with ten bound nucleosomes and to visualize its transcriptional repression by the Tup1 and Ssn6 proteins. These form a molecular barrier that locally prevents the passage of RNA polymerase II, without the formation of a complex with an unique conformation.

Roel van Driel (University of Amsterdam, The Netherlands) gave an overview of the current state of knowledge about the cell nucleus and presented recent data on the distribution of Polycomb-like proteins in euchromatic but transcriptionally inactive regions in mammalian nuclei. He stressed that the disposition of active and inactive sub-chromosomal domains throughout the nucleus affects gene expression directly, and pointed out that new approaches might be needed to study the way the nucleus is organized at high resolution. The chromatin changes taking place during the transition from the inactive to the active state were highlighted by James McNally (National Cancer Institute, Bethesda, USA) with elegant experiments using a glucocorticoid receptor tagged with green fluorescent protein (GFP) and a glucocorticoid-inducible tandem array of over 200 transcription units containing the mouse tumor virus (MMTV) promoter and a *Ras* reporter. Live-cell imaging showed that the tagged loci decondense and become transcriptionally active upon addition of glucocorticoid, and fluorescence *in situ* hybridization experiments showed overlap between the positions of nascent transcripts and open chromatin loci. We (A.P.) described recent work to identify RNA polymerase II and III transcription sites in the nucleus using immunolabeling with anti-polymerase and anti-Bromo-RNA antibodies. An antibody-inhibition method was developed that surpasses the resolution limits of light microscopy; the ability of one antibody to inhibit the access of another to the same compartment was used to quantify the extent of colocalization between RNA polymerases II and III and their associated nascent transcripts. The effective resolution is then of the order of magnitude of the probes used, or the size of antibodies (10-20 nm). Using this technique, we (A.P.) showed that polymerase II and polymerase III transcription sites are distinct.

Centromeres, specialized stretches of chromatin required for chromosome separation, contain repetitive DNA and specific proteins, with ill-understood maintenance. Astrid Visser (Scripps Research Institute, La Jolla, USA) gave details of the way the cell cycle controls the synthesis and stability of centromere proteins and guarantees correct inheritance of centromeres. Constructs of the centromeric CENP-A protein fused to GFP were injected at precise times through the cell cycle; incorporation of CENP-A into centromeres was seen in live cells at all stages of the cell cycle, although endogenous synthesis of CENP-A is restricted to the late S and G2 phases, before its mRNA is degraded in mitosis.

Studies of the mechanism and rate at which RNA and proteins move within the nucleus were discussed by Thoru Pederson (University of Massachusetts Medical School, Worcester, USA), Langowski and Goedele Maertens (Katholieke Universitiet Leuven, Belgium). New applications of fluorescence correlation spectroscopy now allow the measurement of molecular movement within a picoliter-sized confocal volume within the living nucleus. This approach has revealed that macromolecules appear to diffuse within the intrachromosomal space and can move much more rapidly than was previously thought, with diffusion coefficients close to those observed in solution. Pederson showed this to be true for the movement of both intranuclear polyadenylated RNA and ribosomal RNA, Maertens showed it for HIV integrase, and Langowski showed it for GFP when used as a marker to study intranuclear diffusion. Langowski also discussed the application of an anomalous diffusion model to help predict how the apparent diffusion coefficient is affected by the degree of obstruction (transient confinement within physical barriers ('corralling') and/or transient binding interactions) encountered by the diffusing molecules as they move through the chromosomal 'obstacle course'. The need for a truly inert marker molecule to be used in live cell studies was brought up again at this point. For example, dextrans are currently often used as 'inert' marker molecules to estimate diffusion coefficients and the viscosity of various cell compartments, but dextrans do interact with other molecules in the cell and are not truly inert. The nature of a truly inert marker molecule was discussed in some detail, and the physicists in the audience took this need as quite an interesting challenge.

It is our view that the goals of the meeting were achieved in full: by the end of three days, physicists, geneticists, microscopists and cell biologists were engaging freely in discussions that continued well after the end of the meeting. As a result of those discussions, a consortium report is being prepared to help focus attention on potentially highly fruitful areas of research at the interface of the seemingly diverse disciplines of physics, cell biology and genomics.