Mathematical methods in systems biology

Genes in 3D

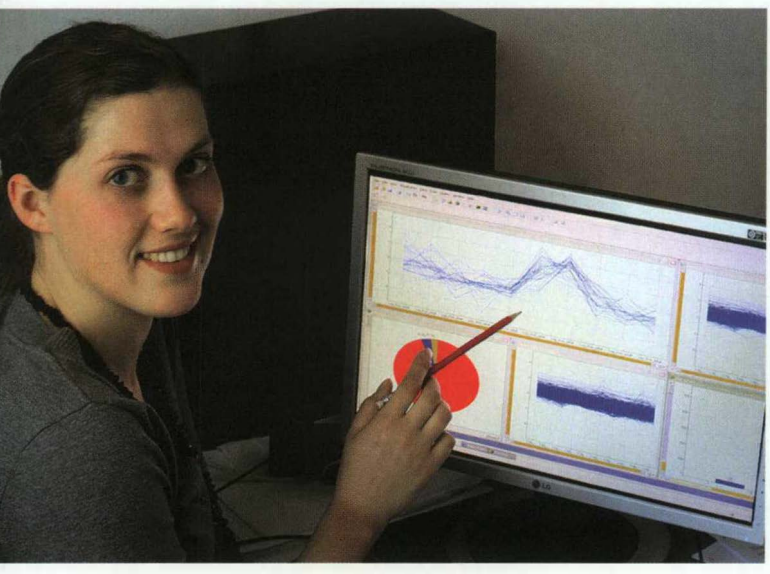
TERESA SZCZEPIŃSKA
Laboratory of Bioinformatics and Systems Biology
Nencki Institute of Experimental Biology, Warsaw
Polish Academy of Sciences
t.szczepinska@nencki.gov.pl

Mathematical and computational tools can seek out hidden correlations within a huge volume of existing biological data, teaching us more about the architecture of the cellular nucleus

The cellular nucleus is an extraordinarily well-ordered system. Inside are the chromosomes, complexes of DNA and proteins that get precisely distributed during cell division. The chromosomes become condensed during cell division but unfold inside the nucleus during the interphase (the time between cell divisions), enabling the information they carry – the genes – to be read off and expressed. The DNA molecules together with the proteins and other molecules attached to them contain all the biological information necessary for cell function.

Kissing genes

Despite their great diversity of structure and function, all the cells in an organism contain identical DNA molecules.



Zofia Pilat

***In silico* analysis (via computer simulation) enables us to posit new hypotheses about the spatial organization of chromosomes. Here: the author in her laboratory**

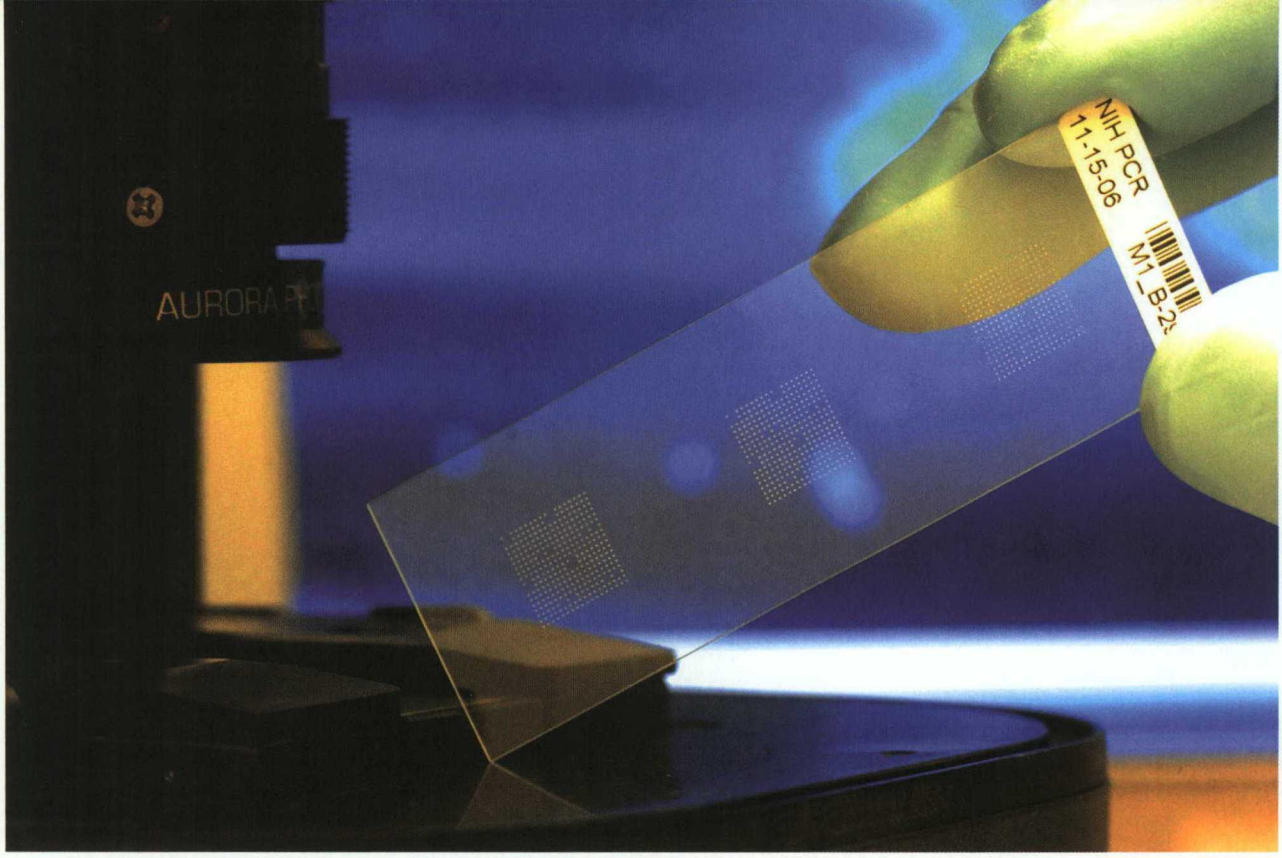
The only differences found in the genetic material of cells from different tissues involve modifications of the nitrogenous bases which make up the DNA, the attached proteins, and other molecules, and the specific details of how the DNA is packed. A chromosome never becomes completely unfolded and its various parts may be condensed to varying degrees. More and more research is now showing that the location which genes occupy within the nucleus constitutes an important factor in the organization of genetic material.

An unfolding chromosome does not expand to fill the entire nuclear space. Instead, it remains limited to a certain specific area. The "territories" of the different chromosomes overlap only to a small degree. A given chromosome's position is not identical in every cell of a given tissue type, but certain preferences have been observed. This is a matter of probability, not an absolute rule: certain chromosomes are simply seen to more frequently occupy certain positions relative to the nuclear center and to one other. Such distributions seem to be characteristic of tissue types, and their non-randomness encourages us to study the role they play in cell function.

It has been observed that despite the restricted territories that the chromosomes occupy, certain fragments of them may sometimes become more unraveled, altering the position of the genes they contain. Certain genes located on different chromosomes are known to share certain functions. The alteration of their position causes a change in their activity. Such "gene kissing," when specific genes on different chromosomes draw spatially close to one another, may either stimulate or silence their action. Why is that? There are regions within the nucleus where genetic material gets decoded more intensely, and others that are largely closed off for decoding. A gene being brought closer to such an inactive region has been observed to cause its silencing (being temporarily switched off).

Chromosome territories

The existence of chromosome territories significantly reduces the possible physical interaction between genes. If two chromosomes are located next to one another, it is much simpler for certain genes on them to become focused together. But if the chromosomes are situated in opposite parts of the nucleus, such proximity may be nearly impossible. Gene activation and decoding are performed by complicated machinery comprised of molecules that work together, forming complexes. If the genes that are meant to be active at a given moment were to be grouped



Photoresearchers/BE&W

The technique of studying gene activity using microarrays enables us to measure the expression of many genes simultaneously

together spatially, their expression would be easier: at such locations the necessary molecules would occur in higher concentrations, thereby improving the efficiency of the process.

Nuclear architecture

Evidence that the architecture of the cellular nucleus affects the expression of genes encourages us to visualize genetic material in three dimensions. The gene regulation system now being explored by scientists may reveal new elements if it is viewed in spatial terms. The non-random distribution of chromosomes within the nucleus and the impact of their positioning on gene activity must be related in a special way to the distribution of genes, both within a given chromosome and between chromosomes. Changes in a gene's position within the genome or its location on the chromosome are known to be able to affect the cell functions related to that gene. However, it is not clear when such a "displacement mutation" will trigger significant functional changes.

Studying the position of a specific gene within the nucleus is difficult and costly. It is performed only on a small scale: for a limited number of genes and tissues. That will presumably change once molecular biology techniques become more advanced, but already today computational analysis methods can be used to glean a better understanding of the phenomenon.

Following the sequencing of the human genome, we are now aware of the position of all known genes within the genome - both the chromosome number and their location on that chromosome. The technique of studying gene expression using microarrays enables us to measure the intensity of gene expression for many genes simultaneously, in practice for all known genes. A sizable amount of such experimental data is already publicly available, for many healthy as well as diseased tissues - a rich resource that can be drawn upon without having to carry out additional experiments.

Advanced mathematical and computational tools allow a huge volume of biological data to be analyzed, seeking to uncover hidden correlations. There are some 30,000 known human genes, microaray experiments are numbered in the hundreds, and each of them may yield data from dozens of microarrays. Bioinformatics techniques can harness huge datasets of this sort (collected but not yet fully analyzed). By asking new questions about experiments already carried out by other researchers and by analyzing correlations discovered within gene expression data, we can develop new hypotheses *in silico* (via computer simulation) concerning the spatial organization of chromosomes.

Computer-enhanced biology

A better grasp of the system of dependencies between gene chromosomal position and expression would be extraordinarily useful, helping us to better understand the significance of the position of chromosomes within the nucleus and interactions between chromosomes, and to predict such interactions in the future. It would also clarify the significance of gene displacement mutations, which play a role in the development of cancer cells. ■

Further reading:

Meaburn K.J., Misteli T. (2007). Cell biology: Chromosome territories. *Nature, 445*, 379-381.

Cremer T., Cremer M., Dietzel S., Müller S., Solovei I., Fakan S. (2006). Chromosome territories - a functional nuclear landscape. *Current Opinion in Cell Biology, 18*, 307-316.

Carter D.R.F., Eskiw Ch., Cook P.R. (2008). Transcription factories. *Biochemical Society Transactions, 36*, 585-589.

Cavalli G. (2007). Chromosome kissing. *Current Opinion in Genetics & Development, 17*, 443-450.