APPLIED MEDICAL INFORMATICS Volume 11, No. 3-4/2002

CONTENTS

Tudor Drugan Cristina Drugan Gheorghe Jebeleanu Andrei Achimaş Cadariu Ştefan Ţigan	BIOINFORMATICS AND GENOME BASED MEDICINE	3
Tudor Călinici Daniel Leucuţa Andrei Achimaş Cadariu	THE BIOINFORMATICS WORKSTATION – THE WINDOWS APROACH	10
Dan Istrate Daniel Leucuţa Andrei Achimaş Cadariu	THE BIOINFORMATICS WORKSTATION – THE LINUX APROACH	14
Daniel Leucuţa Andrei Achimaş Cadariu	COMPUTER SIMULATIONS AND VISUALIZATIONS OF MOLECULAR STRUCTURES	19
Delia Carmen Cacuci	PHARMACEUTICAL BIOINFORMATICS AND DRUG DISCOVERY	25
Sorana Bolboacă	THE COMPUTER REVOLUTION IN NEUROBIOLOGY	32
Cristina Drugan Tudor Drugan Ştefan Ţigan	REAL-TIME PCR	39
Cosmina Bondor Tudor Drugan	STATISTICAL METHODS IN BIOINFORMATICS: MARKOV CHAINS	47
Mihaela Iancu	HIDDEN MARKOV MODELS	54
Ramona Gălăţuş	DATABASES CONCEPTS AND BIOINFORMATICS	59
Mădălina Văleanu	BUILDING BIOLOGICAL DATABASES	67
Mădălina Neamţu	ETHICAL ASPECTS OF GENOME RESEARCH AND BANKING	73
Horaţiu Colosi Tudor Drugan Andrei Achimaş Cadariu	MODELING AND SIMULATION IN MOLECULAR BIOLOGY	80
Tudor Drugan Ştefan Ţigan	BIOPERL: A PERL EXTENSION FOR BIOINFORMATICS	84
Keywords		91
Author Index		92

Applied Medical Informatics, vol. 11, No. 3-4/2002, pp: 32-38

Review

THE COMPUTER REVOLUTION IN NEUROBIOLOGY

Sorana Bolboacă

Dept. of Medical Informatics and Biostatistics, University of Medicine and Pharmacy
„Iuliu Haţieganu" Cluj-Napoca

Abstract - *The actual challenge in medicine is the understanding of the most complex organ created during evolution, the human brain. For this, we need to bridge many different levels of description, from molecules to cells and, from systems to organisms, which can be addressed in many disciplines, from anthropology to molecular biology. Actual evolution of information technologies gives as the possibility to look insights into the organization of the brain and to understand brain function. We have now a lot of non-invasive technique used in order to create brain maps as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), magnetoencephalography (MEG), single photon emission computed tomography (SPECT), and positron emission tomography (PET). Using these non-invasive techniques, we can create functional brain imaging in order to be used to map de neural systems that underlie human behavior. In the actual development, is necessary to integrate neuroanatomical maps with molecular details stemming from biochemical, physiological, pharmacological, and molecular studies, such as ion channel and receptor distribution (proteomics), mRNA level distribution (genomics), synaptic connectivity and plasticity. The accumulation of facts and data on brain is growing day-by-day and it is impressive, but it is necessary to work harder in order to understand their meaning and how we can use this information in comprehending the complex human brain function, or in prevention or treatment of the human brain pathology.*

Key words *– neurobiology, neuroinformatics, computer revolution*

Introduction

Nowadays, challenge is the understanding of the most complex organ created during evolution, the human brain. In order to understand the brain we need to bridge many different levels of description, from molecules to cells and, from systems to organisms, which can be addressed in many disciplines, from anthropology to molecular biology. The accumulation of facts and data on brain is growing day-by-day and it is impressive, but the depth of our insight regarding their meaning is much more limited.

In the last few years, we have been assisted at a tremendous advance and progress of information technology. Computers play a central role in the design, execution, and analysis of experiments in medicine. They have changed the dynamics of the modern laboratory in almost every respect, not scientific thinking or methods, nor hypotheses and controversies surrounding discoveries and falsification of models, but the core of science: the actual experiment. [i] Now, computers count the number of cells in a Petri dish, measure the size of cell nuclei in the microscope cross section of a kidney tumor for example, record the electric activity of an isolated neuronal cell, and read the sequence information from an electrophoresis gel after having exposed the radioactive labeled DNA fragments overnight on X-ray films. Other indispensable lab support come from electronic cell sorting machines, video cameras with online control from computers, purifying a product in reverse-phase chromatography, digital oscilloscope, and real-time peak current analysis. Sophisticated computers assist in experimental systems where the human hand and eye would be too slow and inaccurate to guide and record important data. [i]

The information technologies are now being brought to bear on providing insights into the

organization of the brain and in particular to understand brain functions. We can speak now, of potential synergies between neuroscience and information technology, for example:

- Applying advanced information technology methods to deal with the flood of neuroscientific data;
- Developing and applying data analysis methods for the study of the brain;
- Understanding brain function in order to develop information technology.

However, in order to allow the development of this, it is important to realize that we must accept some important challenges at the level of practical science as well as science administration and policy making. Thus, if we want to understand the brain and appreciate the intricate inter-relationship of its multiple levels of functional organization, it will be necessary a truly global collaborative effort, to communicate our ideas and observation beyond the boundaries of particular disciplines in which individual researchers gather their data and a real communication between researchers. [ii]

Even if health informaticians are to become centrally involved in bioinformatics, they need to establish robust collaborations with bioscientists. In the forging of such collaborations, it is important to understand that informaticians can play two general types of informatics roles - specifically, providing informatics support for bioscience research and performing informatics research that uses the bioscience domain as a context for addressing basic informatics research issues. [iii]

Medicine is a beneficiary of computer technology. Non-invasive techniques eliminate the need for surgical procedures, the use of chemicals, or the use of X-rays. Two successful non-invasive techniques are functional magnetic resonance imaging (fMRI) and ultrasound, which make use of the physical properties of specific atoms or molecules.

Human Brain Project

Precision and high-quality measurements in neuroscience were achieved before the advent of the computer. Ramon Y. Cajal (1852-1934), a Spanish neuroanatomist and the 1906 Nobel laureate in medicine was the leader in brain research. He showed that the vertebrate brain was made up of billions of individual cells and neurons and was not a continuous network of fine arteries. Nowadays, new straining technique combine with photography and computer-aided spectroscopy are used to construct anatomical maps of the human brain.

On April 2, 1993, The Human Brain was officially announced by the National Institute of Mental Health. The Human Brain Project [iv] is a broad-based initiative which supports research and development of advanced technologies, and infrastructure support, through cooperative efforts among neuroscientists and information scientists (computer scientists, engineers, physicists, and mathematicians).

The Human Brain Project consists conceptually of three-projects:

- Connectivity Project or Multi-Modal Imaging and Analysis of Neuronal Connectivity;
- Atlas Project or In Vivo Atlases of Brain Development;
- Algorithms Project or 3D Analysis and Visualization.

A number of distinct technologies need to be integrated into these projects, such as chemistry, animal models, computation, Web design, networking, and functional magnetic resonance. These technologies include [i]:

- the design, synthesis, purification, and characterization of novel contrast agents (chemistry core);
- animal acquisition, care, breeding, surgery, and disposal (animal core);
- design and maintenance of a stable and effective computational environment;
- implementation of algorithms associated with goal directed imaging experiments;
- dissemination of the results to provide routine hardware and software support for computers, network, Web page maintenance, and video environments;
- the development and maintenance of an imaging facility that users nuclear magnetic resonance.

In the long term, the Human Brain Project will provide more than just a sophisticated array of information technologies to help scientists understand how various aspects of brain function fit together. It will also make available to researchers powerful models of neural functions, and facilitate hypothesis formulation and electronic collaboration. The technologies and

standards which are developed as part of the Human Brain Project will serve as models for other scientific information tools. The Human Brain Project will, therefore, have impact far beyond the community of brain and behavioral researchers, and this impact will be felt long after the end of the Decade of the Brain.
In order to understand better the complexity of human brain, one of the most important goals of Human Brain Project is the "weaving of the informatics and neuroscience components" of brain research into a single unit.

Neuroinformatic
The brain receives attention from scientists for the next reasons: behavior, consciousness, memory, sleep disorders, perception, or pain.

Neuroinformatics, a new interdisciplinary field, intend to integrate the maps or sequences with functional data. The goal of the field of neuroinformatics is to provide enabling informatics technology that supports neuroscience research at many different levels (table 1) [v]. These include the genetic level, the cellular level, the physiologic and pharmacologic level, and, eventually, the level of behavioral research. The research includes experimental microanatomic studies (e.g. imaging cells) and macroanatomic studies (e.g. imaging brains). Each type of experiment uses different experimental techniques and generates different types of experimental data. [iii]

Levels of Brain Function	**Type of Data**
Behavior	▪ performance quantification ▪ video monitoring ▪ drug testing
Distributed systems	▪ 2D and 3D axon tracing between regions ▪ electrophysiologic recordings (spike timing) ▪ brain imaging ▪ 3D brain maps
Specific regions	▪ 2D and 3D cytoarchitectonics of layers and functional columns ▪ transmitter-receptor localization ▪ anatomic, physiologic, and metabolic maps
Nerve cells	▪ electrophysiologic recordings of action-potential firing patterns and membrane currents
Neuronal components	▪ 3D imaging of axon terminals, growth cones, dendrites, dendritic spines ▪ 3D localization of organelles and synaptic microcircuits
Microcircuits	▪ 3D fine structure and imaging of synaptic patterns ▪ synaptic pharmacology ▪ action-potential firing patterns and synaptic currents, and potentials
Organelles	▪ 2D and 3D fine structure and molecular composition of synapses, mitochondria, microtubules, etc. ▪ recordings of synaptic currents and potentials
Molecules	▪ 3D molecular models of receptors, channels, enzymes and structural proteins ▪ molecular physiology, and pharmacology of transmitters, modulators, hormones, guidance molecules, growth factors and gene-transcription factors
Genes	▪ DNA and protein sequences

Table 1. Experimental neuroscience data at different levels of brain function

.

Neuroinformatics aims are to provide tools for data management and data exchange. The ability to disseminate and exchange data sets and computational tools and models to any interested investigator will contribute to advancement in neuroscience. In this moment, the researchers have some concerns about deep of control in data entry and data quality.

Quality neuroanatomy will profit from the neuroinformatics tools that are developed and progress. Advances in this field will [vi]:

- result in more powerful data acquisition and processing facilities
- improve shape resolution in 3D morphological reconstruction
- enhance the power of quantitative shape analysis

- improve algorithmic tools to synthesize neuronal morphological complexity in greater detail
- understanding of the functional role of dendritic geometrical details in neuronal information processing
- increase our ability to integrate neuronal morphological data with other neuroscience data
- provide tools for morphological data management and exchange with colleague researchers.

Mapping the Human Brain

The success of the Human Genome project and the recent computer-aided brain scanning technique has revitalized attempts to map the human brain both anatomically and functionally. With these maps, it should one day be feasible to co-localizate neural activity at the level of single sensory neurons as a function of motoneural tasks, learning words, or reading symbols.

In order to create a map of human brain it is necessary to superimpose an anatomical and functional map both in space and time. In this filed is very important the level of resolution, because the type of questions that can be asked depend on these level. Functional brain mapping makes use of non-invasive techniques such as:

- SPECT - single photon emission computed tomography / PET - positron emission tomography;
- fMRI – functional magnetic resonance imaging;
- EEG – electroencephalography;
- MEG – magnetoencephalography;
- Optical imaging;
- Neuroanatomical tools.

These tools are used to produce maps composed of cross-sectional images of the brain. After that, cross sections and put it together into a virtual 3D image reconstruction of the brain.

The integration of the spatial and temporal domains represents a unique challenge because of the existence of anatomically distinct processing between regions that communicate across several time scales. In the last decade, have been described a number of different techniques for non-invasive measuring human brain activity. [vii] These can be broadly classified into either homodynamic or metabolic or electromagnetic measurements. Current hemodynamic measurements, particularly functional MRI, provide excellent spatial resolution (millimeters) but are temporally limited by the latency of the hemodynamic response (seconds). [viii, ix] Conversely, electroencephalography (EEG) and magnetoencephalography (MEG) provide excellent temporal resolution (milliseconds) but uncertain spatial localization. [vii] In order to obtained movie of brain activity is need to correlate spatial resolution in the millimeter range (fMRI) and temporal resolution in the millisecond range (EEG, MEG). These brain maps results have a resolution at the neuroanatomical level. The researcher from Harvard Medical School developed The Whole Brain Atlas [x] which put at the disposition of the people interested in a real-time map that compares normal brains with those affected by cerebrovascular diseases (e.g., strokes), neoplastic disease (e.g., tumors), degenerative diseases (e.g., Altzheimer's, Huntington's), and inflammatory or infectious diseases (e.g., multiple sclerosis, AIDS-related dementia, Creutzfeld-Jako, herpes). The maps reflect a high spatial resolution obtained from hemodynamic or metabolic measurements, such as glucose levels and oxygen consumption with the high temporal resolution of their electromagnetic signals. This correlation is not a trivial one because the brain is a complex organ with anatomically distinct processing centers that communicates across large distance and several time scales [vii].

Spatial resolution in millimeters is not good enough for molecular studies of brain functions. Neuroanatomical maps like The Whole Brain Atlas can be complemented with molecular details stemming from biochemical, physiological, pharmacological, and molecular studies, such as ion channel and receptor distribution (proteomics), mRNA level distribution (genomics), synapstic connectivity and plasticity responsible for enormous complexity of neuronal networks. [i]

Three dimensional visualization of brain structures allow to view structure from a perspective that we are otherwise unable to have and give as some useful information.

From Molecules to Neurodegenerative Disorders

Neurodegeneration is a major element in neurodegenerative disorders and is often causing

of disability in many diseases not usually classified as degenerative (e.g., multiple sclerosis, epilepsy, some inborn errors of metabolism, schizophrenia, and even tumors). Core members of neurodegenerative diseases are dementias, Parkinson's disease, motor neurons disease, cerebral degenerations, Hungtington's disease, and prior diseases. [xi]

Neurobiology has made great progress in characterizing the basic function of individual neurons and their components.

Genes and proteins involved with neurodegenerative conditions are being rapidly elucidated, and naming the condition after the protein is an option. [xii, xiii]

Alzheimer disease, the fourth leading cause of death in adults, manifest as a progressive inability to remember facts and events and later to recognize friends and family rise because there are some mutation in four genes, situated on chromosomes 1, 14, 19 an 21. The formation of lesion made of fragmented brain cells surrounded by amyloid-family proteins is characteristic of the disease. [xiv]

Parkinson disease, first described by James Parkinson in 1817, is a neurodegenerative disease that manifests as a tremor, muscular stiffness and difficulty with balance and walking. A classic pathological feature of the disease is the presence of an inclusion body, called the Lewy body, in many regions of the brain.

Mutation in a gene placed on chromosome 4 is responsible of Parkinson. The product of this gene is a protein called alpha-synuclein, a familiar culprit: a fragment of it is a known constituent of Alzheimer disease plaques. [xiv]

Huntington disease is an inherited, degenerative neurological disease that leads to dementia. The mutation in gene responsible of Huntington disease is a characteristic expansion of a nucleotide triplet repeat in the DNA that codes for the protein huntingtin. [xiv]

There are other genes and proteins involved with neurodegenerative conditions. The propagation of basic signaling mechanics of the brain, potentials in neurons, is a well understood nowadays. Action potentials are localized, small voltage changes - measured in milivolts - across the cell membranes and their lifetime is measured in milliseconds. These kind of electric activities are record using analog instruments, remain the tasks of analysis them. Nowadays, using computer software, data analysis can e performed in minutes and we will have the results immediately. The study of ion channel, as functionally and structurally, and the use of molecular biology to unravel genes and genomic organization of channel proteins during the 1990s has greatly improve the understanding of Altzeimer's, Parkinson's and Huntington's diseases.

Molecular neurobiology addresses the structure-function relationship of so-called ion channels, the protein complexes responsible for action potentials. Many of these channel proteins can be linked to neurodegenerative diseases and susceptibility to drugs and toxins.

The cell membrane is an electrical insulator and blocks the movement of ion in the absence of channels. With extremely sensitive electrical amplifiers, the movement of a few thousand ions in every thousandth of a second can be measured and could reveal information about the electrical activity of a cell membrane. Action potentials, the signaling mode of neurons, are the combination of at least three different classes of ion channels; that is means that for each different ion it is present a different protein in order to across the membrane. The ion channels are ion selective.

Nowadays, computers have become indispensable tools in the control and analysis of electrophysiological experiments (because of the nature of electrical activity and the extremely low number of charged molecules moving across the membrane). Electrophysiology was born in 1940s and has grown into an industry. Now, scientists interested in particular proteins or diseases can study specific pharmacology and physiology though commercially available computers that help us to push the limits of experimental systems, especially in terms of very rapid events. Current signals that can be converted into voltage signals or coupled to fast optical events, allowing ever-improving time and amplitude resolution, are reaching the theoretical limit of measuring the activity of a single molecule.

Long-QT-syndrome is a defect in the rhythm control of the heart muscle as the results of structural abnormalities in the potassium channels of the heart, which predispose affected persons to an accelerated heart rhythm. [xv] This can lead to sudden loss of consciousness and may cause sudden cardiac death in teenagers and young adults who are faced with stress ranging from exercise to load sound. [xiv]

The responsible protein is a potassium-selective ion channel that excludes all other biologically important ions like sodium, chloride, calcium, and magnesium. Potassium is found mostly inside cells and outside only at a lower concentration. When a potassium channel is open, potassium ions move from the inside of the cell to the outside. , moving positive charges to the extra-cellular space. This movement will happen as long as the channels are open and until the potassium, concentration is balanced (that means, equal concentration of potassium inside and outside). But, this equilibrium never occurs in normal cells.

The function of ion channels is to pump the potassium ions back into the cell simultaneously shuffling sodium ions out of the cell through the same transport proteins, creating an asymmetric distribution of sodium ions that generates an electrogenic force opposing the flux direction of potassium. After being open a very short time, the potassium channels are inactivate. The active and inactive potassium channels time varies from cell to cell, because there are different potassium channel genes that coexist in a cell's genome. The net effect of potassium channel activity in the repolarization of the resting potential in neurons. In the Long-QT-syndrome, this process is markedly slowed because of mutations in one of the potassium channels genes causing induction of arrhythmia in the heart muscle.

We know today two categories of diseases linked to ion channel defects, that is present it in Table 2. ion channels are also susceptible to toxins and drugs, such as calcium channels for snail toxins and caffeine, acetylcholine receptors for nicotine, and potassium channels for the snake toxin charybdotoxine.

Disease	**Susceptibility to toxins and drugs**
Gap junction protein Cx43; X-linked Charcot-Marie-Tooth disease	Calcium channels; omega-conotoxin (snail toxins), caffeine
Chloride channel CIC-2; Myotonia	Achetylcholine receptor; nicotine, alpha-bungarotoxin
Long-QT-syndrome	Potassium channel; charybdotoxin

Table 2. Diseases linked to the channel defects.

Sequencing and gene mapping progress results in the identification of the genes responsible for the diseases listed in table 2. The National Center for Biotechnology Information [xiv] offers a summary link to genetically determined diseases. From those diseases, eight are linked to ion channels, pumps and transporters, from which the next neurodegenerative diseases:

- Merker's syndrome is an inborn error of metabolism that markedly decreases the cells ability to absorb cooper. Sufferers cannot transport copper that is needed by enzymes involved in making bone, nerve and other structures. The disorder causes severe cerebral degeneration and arterial ahanges, resulting death in infancy.
- Wilson's disease is a rare autosomal recessive disorder of copper transport, resulting in cooper accumulation and toxicity to the liver and the brain. Liver disease in the most common symptom in children and neurological disease is most common in young adults.
- Zellweger syndrome affects infants and usually results in death. These children have unusual problems in prenatal development, an enlarged liver, high level of iron and cooper in the blood. The gene responsible for this disease is PXR1 on chromosome 12. There are also involved peroxisome proteon import receptor and peroxisome biogenesis disorder.

We leave it back the era of analog machines and nowadays, a scientist is not necessary to be a pioneer in electronics and code writing to use a computer. Scientists have a tool to study the function of protein responsible for brain activity at the molecular level.

What is next?

The computer, as well as hardware and as well as software, and the progress of information technology in the next years will grow in parallel with the progress in neurobiology. The development and deployment of expert systems will increase the ease and success of performing routine medical tasks.

In the same time, the neurobiology databases will have data as much as will be possible to integrate informatics and neuroscience components of brain research into a single unit. We will have more and more date about the anatomy and

function of the human brain, about neurodegenerative diseases. Analyzing these data, we will be able to make real progresses in understanding normal brain functionality and in comprehending neurodegenerative diseases, in order to find out new treatments.

Bibliography

[i] Rashidi H. H., Buehler K. L.: **Bioinformatics Basics. Applications in Biological Science and Medicine. The Computer Revolution in Neurobiology.** CRC Press LLC New Work 2002: 143-154

[ii] Amari S.-I., Beltrame F., Bjaalie J. G., Dalkara T., Schutter E. D., Egan G. F., Goddard N. H., Gonzalez C., Grillner S., Herz A., Hoffmann K.-P., Jaaskelainen I., Koslow S. H., Lee S.-Y., Matthiessen L., Miller P. L., Silva F. M. D., Novak M., Ravindranath V., Ritz R., Ruotsalainen U., Sebestra V., Subramaniam S., Tang Y., Toga A. W., Usui S., Pelt J. V., Verschure P., Willshaw D., Wrobel A.: **Neuroinformatics: The integration of Shared Databases and Tools Towards Integrative Neuroscience**. Journal of Integrative Neuroscience December 2002; 1(2): 117-128

[iii] Perry L. Miller: **Opportunities at the Intersection of Bioinformatics and Health Informatics.** J Am Med Inform Assoc. 2000 September; 7 (5): 431–438

[iv] National Institute of Mental Health. Human Brain Project home page: http://www.nimh.nih.gov/neuroinformatics/index.cfm

[v] Shepherd G.M., Mirsky J.S., Healy M.D., et al.: **The Human Brain Project: neuroinformatics tools for integrating, searching, and modeling multidisciplinary neuroscience data.** Trends Neurosci. 1998; 21: 460-8.

[vi] Japp van Pelt: **Quantitative Neuroanatomy and Neuroinformatics.** Network: Comput. Neural Syst. 13 (August 2002) 243-245

[vii] Liu K. A., Belliveau W.J., Dale M.A.: **Spatiotemoral imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations.** Proc. Natl. Acad. Sci. 95 (July 1998): 8945 – 8950.

[viii] Belliveau J.W., Kwong K.K., Kennedy D.N., Baker J.R., Stern C.E., Benson R., Chesler D.A., Weisskoff R.M., Cohen M.S., Tootell R.B., et al.: **Magnetic resonance imaging mapping of brain function. Human visual cortex.** Invest. Radiol.27 Suppl 2 (1992 Dec): S59-65.

[ix] Frahm J., Merboldt K.D., Hanicke W.: **Functional MRI of human brain activation at high spatial resolution.** Magn Reson Med. 1993 Jan; 29(1):139-44.

[x] Havard medical School. The Whole Brain Atlas home page: http://www.med.harvard.edu/AANLIB/home.html

[xi] Williams A.: **Defining neurodegenerative diseases**. BMJ 2002; 324: 1465-1466.

[xii] Polymeropoulous M.H., Lavedan C., Leroy E. Ide S.E., Dejejia A., Dutra A., et al.: **Mutation in the alpha-synuclein gene identified in Families with Parkinson's disease.** Science 1997; 276: 2045-7.

[xiii] Goate A., Chartier-Harlin M.C., Mullan M., Brown J., Crawford F., Fidani L., et al.: **Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease.** Nature 1991; 349: 704-6.

[xiv] National Center of Biotechnology Information. Genes and Diseases http://www.ncbi.nlm.nih.gov/diseases

[xv] Yang W.P., et al.: KvLQT1**, a voltage-gated potassium channel responsible for human cardiac arrhythmias.** Proc. Antl. Acad. Sci. U.S.A., 1997; 94(8): p.4017-21.