

INFOBIOMED

NEWSLETTER

N.2 JUNE 2005



```
TCCTCCTAGCAACGCGGCCCTCTCGGCTCTTAAGGCGTGACTTGCCATGGATGTTCTGCC  
TTTGGAAACCAAGTGGGTGGTCACAGCACAGAGGCTGTTGGTTTCATGGCTGTCACATGCAC  
AGAGACAGCCCTGTCCACTGAGCCTGTTCTGTGGCAGAAGGAACCCCTAGCCTGGACC  
TCCTTGGGTGGAACACCAGCCCTCAGCTCTGACTGTCATTTTCAGATGTGAAGAGAA  
ACACAGTGTGCTGATGTTGTATTGCCGTGAACAACGAGAGTTTCTTTGACTTC  
TAACTGCCTTAAAGCTAGAGCCCGGCATTAGGATGGCAGTGGTGGTAAATGAGGTATG
```

**BIOMEDICAL INFORMATICS
TO SUPPORT INDIVIDUALISED HEALTHCARE**

PORT INDIVIDUALISED HEALTHCARE

SUMMARY

NEWS

PG. 2

Researchers develop promising new gene network analysis method	pg. 2
Scientists Use Antispam Technology in AIDS Vaccine Work	pg. 2
The University of Cambridge, Cancer Research Technology, Cancer Research UK, and Perlegen Sciences Collaborate to Analyze Thousands of DNA Samples from Breast Cancer Patients	pg. 3
UK Computer Scientists Identify Future IT Challenges	pg. 3
Expression Project for Oncology (expO) completes first phase of standardized gene expression analyses	pg. 3
Standardized Microarrays May Bring Us One Step Closer to Personalized Medical Treatment	pg. 4
Medicine Gets Personal	pg. 4
First Version of Healthchip® for Cancer Prediction	pg. 5
Multiple Sclerosis - Serono Announces Major Milestone in Identifying The Genes Involved	pg. 5
Genomic analysis to become tool for studying trauma patients	pg. 6
Microarrays as phenotype	pg. 6
Computational Tool Predicts How Drugs Work In Cells	pg. 7
A new method sequence thousands of genomes simultaneously	pg. 7
The Viking's route of the Multiple Sclerosis	pg. 8
New computer program uses brain scans to assess risk of Alzheimer's	pg. 9
NIH Awards Four Microarray Centers \$25M for Neuroscience Research	pg. 9

INFOBIOMED PROFILES

PG. 10

Custodix	pg. 10
UPM - Polytechnical University of Madrid	pg. 10

RESOURCES

PG. 11

Genetic Science Learning Center	pg. 11
NCI Creates Gene Expression Database of Normal Human Organ Tissue	pg. 11

EVENTS

PG. 14

COMMENT

PG. 15

NEWS

Researchers develop promising new gene network analysis method

<http://medicalnewstoday.com/medicalnews.php?newsid=23965>

A multidisciplinary research team led by Leon Cooper and John Sedivy from Brown University together with colleagues at Università di Bologna and Tel Aviv University has set out to develop a sensitive and reliable tool for analysing connections between genes, a map of the links, based on an experiment published in "Proceedings of the National Academy of Science" in May.

In this article the authors studied a cancer protein implicated in switching on genes that ultimately induce cells to multiply.

It is believed that mapping the interactions between thousands of genes is critical to understanding human development and disease, and also aids in the design of new drugs.

The statistical analysis of gene activity experiments was tested by two methods to try to model this network: a linear Markov model (commonly used) and a correlation method based on network theory, which has been used to explain complex systems such as power grids and neural networks. They demonstrated the second model is a more effective analytical tool.

Sedivy said that the network theory has been applied to the genomes of simple organisms (yeast) yielding large amounts of useful information. The plan is to use "the time series experiments and the combination of statistic and network theory" to a much more complex organism. This is "a new approach to studying gene expression."

Scientists Use Antispam Technology in AIDS Vaccine Work

<http://www.bio-itworld.com/news/022305-report7589.html>

Computer Scientists from Microsoft Corp, David Heckerman and Nebojsa Jojic, in collaboration with bioengineers from Australia's Royal Perth Hospital and the University of Washington in Seattle are using computer algorithms for fighting spam to study HIV's mutation patterns. These researchers believe is a key factor that will lead to design broadly effective vaccines against AIDS (Acquired Immune Deficiency Syndrome).

Just as spam merchants make little changes in the words that are blocked by filters, indeed HIV mutates rapidly and in tiny ways that keeps it one step ahead of the body's immune system.

For that reason, these scientists believe that software used to train anti-spam filters to recognize subtly altered words - "VIAGRA" and "V1AGXA" for instance - might be used in DNA sequence mutations and in the future might provide clues as to how a vaccine could recognize and destroy HIV in the body.

A vaccine against HIV designed at Perth and the University of Washington is actually in the laboratory testing phase, using immune cell samples taken from HIV-infected patients. Initial results will probably be ready later this year. Scientists working in this project think that this technique might be applicable to other mutating viruses such as hepatitis C.

The University of Cambridge, Cancer Research Technology, Cancer Research UK, and Perlegen Sciences Collaborate to Analyze Thousands of DNA Samples from Breast Cancer Patients

<http://www.perlegen.com/>

The University of Cambridge, Cancer Research Technology, Cancer Research UK, and Perlegen Sciences Inc. will collaborate to design a study that will be a genome-wide scan for common predisposing genetic variants that are associated with susceptibility to breast cancer. This way they could explain a significant portion of the variation in breast cancer risk in the population.

This new study that will include over 200 million individual genotypes of DNA samples hopes to provide a more comprehensive understanding of the genetic basis of breast cancer that may help to identify more precisely women at high risk and improve the prevention, detection and treatment of the disease. Initial results will be validated by a parallel analysis coordinated by University of Cambridge and collaborators.

UK Computer Scientists Identify Future IT Challenges

http://www.bio-itworld.com/news/012505-_report7262.html

The British Computer Society (BCS) is dedicated for the information technology (IT) in the field of computers and information systems and is responsible for setting standards for the IT profession.

A Grand Challenge for research in science or engineering pursues a goal that is recognized one or two decades in advance; its achievement is a major milestone in the advance of knowledge or technology, celebrated not only by the researchers themselves but also by the wider scientific community and the general public.

A model that could explain the behaviour of living organisms is one of the academic challenges that

are of commercial interest. For this purpose, the scientists suggest to develop a verifying compiler that that proves automatically that a program is correct before allowing it to run.

The UK Computing Research Committee began the Grand Challenges Exercise in 2002 by appointing a Programme Committee to organize and conduct it, beginning with a Grand Challenges Workshop that culminated in an academic conference in March 2004 with the publication of the report "Grand Challenges in Computing Research". There are seven chapters to the document, one for each challenge defined.

They will need further discussion and refinement in the future. Their titles are:

- In Vivo-in Silico (iViS): to examine organisms in virtual reality;
- Science for global ubiquitous computing;
- Memories for life: managing information over a human lifetime;
- Scalable ubiquitous computing systems;
- The architecture of brain and mind;
- Dependable systems evolution;
- Journeys in non-classical computation.

Expression Project for Oncology (expO) completes first phase of standardized gene expression analyses

http://www.intgen.org/expo_phase1_release.html

In cancer research, the International Genomics Consortium's (IGC's) Expression Project for Oncology (expO) seeks to integrate longitudinal clinical annotation with gene expression data of 60 tumor specimens for a unique and powerful portrait of human malignancies, providing critical perspective on diagnostic markers, prognostic indicators, and therapeutic targets.

The clinically annotated dataset is available in the public domain through the National Center for Biotechnology Information web site at <http://www.ncbi.nlm.nih.gov/geo/>.

The main purpose of ExpO is the one to standardize the conditions of the analysis of expression of genes and their clinical annotation. In the future the deidentified tumor collection will reach around the 3,000 samples of a broad spectrum of malignancies and 500-1,000 normal tissues. Also, to improve the quality of the gathered samples they are unified through the different medical centers that participate in the study. It is thought that the results obtained will be of benefit to researchers worldwide.

The funding for this project comes from several pharmaceutical firms (Major sponsors are Bristol-Myers Squibb, GlaxoSmithKline, Wyeth and IBM and other sponsors are Pfizer, Aventis, Johnson & Johnson, Flinn Foundation, ILEX Oncology, and Genomics Collaborative Inc. Research support from Affymetrix, AmeriPath, the Georgia Cancer Coalition, Agilent Technologies, Amersham Biosciences, and Telik Inc) and from Arizona State University, the City of Phoenix, and from Maricopa County in Arizona.

Standardized Microarrays May Bring Us One Step Closer to Personalized Medical Treatment

<http://www.nih.gov/news/pr/apr2005/niehs-21.htm>.

A recent study suggests that doctors will probably soon be able to analyze patient's genes expression linked to a specific diseases to select the best treatments by using microarrays or gene chip technology.

At present there are a lot of different methods and protocols in microarray experiments. Researchers believe that when methods and protocols are unified and standardized, variation between lab results will be minimized and the advantages around this technology will become more apparent and will improve clinical practice. The TRC (Toxico Genomics Research) Consortium studies the causes of variation in the gene expression experiments within and between labs, as well as within and between microarray platforms, reaching a standardized process that will lead to more consistent results. In particular the aim of this study is to analyze the variability of

12 microarrays, which were either produced commercially or produced in-house by each lab, by using mouse RNA.

Medicine Gets Personal

<http://www.the-scientist.com/2005/4/25/34/1>

The National Institutes of Health (NIH) through the Pharmacogenetics Research Network (PGRN) are promoting the discovery of new genes involved in drug metabolism and its effects. In March, the Food and Drug Administration (FDA) released a guideline that indicates how drug companies must submit genomic data and it also shows the evaluation procedure. But there are limitations in this regulatory process that can affect directly the companies involved in pharmacogenetics (study of how inherited DNA variations, typically in just a few genes, affect drug metabolism or toxicity) and pharmacogenomics (technologies that can be used in high-throughput screening) research.

It will be necessary many years before applying genetic tests in the clinic routine and it will be also necessary to carry out studies that convince doctors to use them, like Howard McLeod, a pharmacogeneticist at Washington University in St. Louis said.

The pharmacogenetics could save drug companies millions per drug in the development of new drugs.

There are new guidelines coordinated by Webster chairs that involve the genomics group at Pharmaceutical Research and Manufactures of America (PhRMA) and trade group's that show procedures by which companies submit genomics data to a newly formed interdisciplinary pharmacogenomic reviewed group via a "safe harbor" in which data would not be used for regulatory decision-making. The agency and the drug-development would be both benefited by this process.

Since 2003, according to Marisa Papaluca-Amati, deputy head of sector for safety and efficacy of medicines, The European Agency for the Evaluation of Medicinal Products (EMA) has an expert group through the Committee for

Proprietary Medicinal Products (CPMP) to clarify terminology and make recommendations for pharmacogenetic sample collection in clinical trials.

The FDA is actually working in the development of another guidelines that would be discussed by several companies like the Biotechnology Industry Organization, The Drug Information Association, and PhRMA, at the meeting in Bethesda, Md. in May.

The FDA has published the pharmacogenomics data submission guidance document at [<http://www.fda.gov/cder/genomics>]

The first genetic lab test made by Roche Molecular Systems Inc. was approved by the FDA in December 2004, and was used to genotype cytochrome P450 involved in drug metabolism. But this microarray doesn't still give enough information about the drug dosage.

The public research tool database PharmGKB.org, a nationwide collaborative research consortium that comprising 10 research groups, was developed by Stanford University with funding from the NIH and it is part of the NIH PGRN. Its aim is to aid researchers in understanding how genetic screen for variants among collected 4500 individual patient cohort contributes to differences in reactions to drugs and like Rochelle Long, PGRN's program director says "The whole purpose of the initiative is to link genotype to phenotype: Where do genetic variations matter for drug responses, and where can we robust research correlations and store information in a knowledge base?"

Richard Hockett, from the department of diagnostic and experimental medicine at Eli Lilly thinks that it could be at least 2,000 variants of these genes that can affect drug metabolism. In March Eli Lilly and ParAllele SioScience of South San Francisco announced the development of just such a chip (MegAllele D-MET chip) that will be used to screen in patients in Lilly's Phase I trials after FDA approval.

Roche is working on other automated AmpliChip test, which involves the p53 gene, a tumor suppressor gene that is often mutated in patients with cancer.

"We think every publicly funded clinical trial should contain pharmacogenetics," says Mary

Relling, a PGRN member who chairs the department of pharmaceutical sciences at St. Jude Children's Research Hospital, Memphis, Tenn. "We should be getting DNA and appropriate consent from patients on every trial that's supported by tax dollars. Otherwise, 20 years from now we will have made very little progress."

First Version of Healthchip® for Cancer Prediction

<http://medicalnewstoday.com/medicalnews.php?newsid=23224>

GenoMed is a high-growth biotechnology company that is dedicated to basic research to improve patient outcomes by identifying disease-genes. This company has recently introduced its Healthchip® made up of thousands of cancer-associated genotypes by single nucleotide polymorphisms (SNPs) that has the potential to serve as an early warning system for the top six common cancer in Caucasians: breast, colon, lung, ovarian, pancreatic, and prostate.

Multiple Sclerosis - Serono Announces Major Milestone in Identifying The Genes Involved

<http://www.medicalnewstoday.com/medical-news.php?newsid=21871>

Researchers at the Serono Genetics Institute (SGI) have studied Multiple Sclerosis (MS), a chronic, inflammatory condition of the nervous system and they have identified 80 genes involved in the pathology that could allow the new drugs developing and therapies in the future.

The technique based on a 40% genome scan comparing the genetic profile of a total of 1,800 people with MS and healthy individuals in different populations. SGI used the Affymetrix GeneChip technology to scan over 100,000 SNPs (single nucleotide polymorphisms) to identify the genes involved in MS, comparing the genetic profile of cases and controls. The next step is to continue this endeavour, applying next generation GeneChip technology to scan over 500,000

SNPs and thus complete the MS Whole Genome Scan during 2006.

"We are excited about this significant step forward in building a complete inventory of genes involved in MS," said Professor Daniel Cohen, Vice-President and Worldwide Head of Genetics at Serono. "The completion of the MS Whole Genome Scan in 2006 will lead to a comprehensive catalogue of potential MS drug targets providing a basis for the future development of innovative MS therapies."

Genomic analysis to become tool for studying trauma patients

<http://medicalnewstoday.com/medicalnews.php?newsid=21499>

Genome analysis has an ample spectrum of applications even it is set to provide a means for identifying pathways that are important to critically ill patients in the intensive care unit (ICU), and discovering target effects of compounds. By combining data that interpret changes in genes, proteins, metabolites, it can provide a means of diagnosis and evaluating drugs that alter disease outcome.

Researchers demonstrated that these state-of-the-art techniques produced useful insights into the health of critically ill patients. Results of this study were published in March in Proceedings of the National Academy of the Sciences journal 102 (13).

"It's an exciting time because our field has experienced frustration with some of these questions, many of which have important ramifications for how we treat patients," says J. Perren Cobb, the paper's lead author.

The new study, conducted by Cobb and his colleagues at the Washington University School of Medicine in the Inflammation and Host Response to Injury Large Scale Collaboration network, asked whether the technology could detect differences in the activity levels of genes in critically ill patients versus healthy patients.

They applied DNA microarray to study 34 illness patients and 23 healthy controls. They showed with this technology that there are changes in

gene activity levels in patients that affect can reprogram white blood cells, immune system cells that circulate in the bloodstream. The new approach will allow the investigators for the first time to monitor neutrophil gene activity genome-wide in injured patients.

As scientists understanding of the interaction between multiple genes involved in inflammatory responses becomes more complete, they may be able to develop more effective ways to treat critically ill people even save lives.

This research is supported by funds from the National Institute of General Medical Sciences.

The procedures and protocols established for critical care research are available at the group's web site, <http://www.gluegrant.org>

Microarrays as phenotype

<http://medicalnewstoday.com/medicalnews.php?newsid=22630>

"Microarrays provide a method of quantifying the expression and order of genes in a particular genome - acting as a surrogate measure of cell physiology", said researchers at Baylor College of Medicine (BCM) in a report that appears online in April in the journal Nature Genetics.

"Microarray data are good phenotypes to determine the order of genes and are a good surrogate measure of cell status," said Dr. Gad Shaulsky, associate professor of molecular and human genetics at BCM.

In the microarray technique, the probe sequences (DNA) are immobilized on the surface, at a separation of a few micrometers so that is possible to place many different probes on a small single surface of one square centimeter. The sample (DNA or RNA) from the cells under the study is usually labeled with a fluorescent dye that can be detected by a light scanner that scans the surface of the chip.

The Shaulsky's study performed in the amoeba *Dictyostelium discoideum* provides comprehensive cellular profiles (a particular phenotype) that can be used to identify gene function. For example, they showed that the activation of the protein kinase A (PKA), an essential regulator of gene expression and cell differentiation during multicel-

lular development of *Dictyostelium discoideum* occurs when the organism encounters starvation. "We pretended we did not know the order of genes in the pathway," said Shaulsky in the article. "We were able to reconstruct the pathway from the microarray data. This means the microarray provides a good phenotype that is quantitative. We can prove that gene A comes before gene B and give mathematical support for these findings."

"This is a proof of principle that we set out to do - assessing the function of unknown genes is feasible," said Shaulsky. "It can be done with a microarray phenotype."

Other participants in the study are Drs. Nancy Van Driessche, Ezgi O. Booth, Paul Hill and Adam Kuspa, all of Baylor College of Medicine; and Janez Demsar, Peter Juvan and Blaz Zupan of the Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia.

Computational Tool Predicts How Drugs Work In Cells

<http://www.bio.com/realms/research.jhtml?realmid=3&cid=9000002>

A team of biomedical engineers and chemist at Boston University, Tim Gardner and James Collins from College of Engineering's Department of Biomedical Engineering (BME) and its Center for BioDynamics in collaboration with Center for Chemical Methodology and Library Development (CMLD) have developed a mathematical algorithm that predicts the precise effects a given compound will have on a cell's molecular components or chemical processes enables to facilitate drug design that acts specifically on their targets avoiding side effects. This research was published in March in *Nature Biotechnology*.

Until now the development of new drugs have limitations since they would have multiple effects on multiple genes.

This mathematical process used first a reverse-engineering approach to elucidate the regulatory networks that involves genes in the organism. After that they test the network models results to

predict the drugs effects. Finally they verify the method studying the possible targets of an anti-cancer compound, PTBS that inhibit growth in baker's yeast and also in human small lung carcinoma cells acting on thioredoxin and thioredoxin reductases.

A new method sequence thousands of genomes simultaneously

Diario Médico (Sección Tecnología) 29 de Abril 2005

To fully enhance the understanding of biological process implicated in diseases, there is a need for the study of the simultaneous expression of a great amount of genes and therefore elucidating genetic changes in expressed genes in many samples of cancer, or understanding the different responses that people have to drug treatment, so as to better adapt medications to the needs of individual patients.

One of the world's leading scientists, Nobel Prize winner, Dr Sydney Brenner, has devised a new method for obtaining sequence information from thousands of genomes simultaneously. Current technologies can only analyze one genome at a time.

Dr. Brenner said that in order for disease research to provide more immediate benefits for society, we do not have to know everything about every gene, that what is most important is to discover the variants in genes that contribute to human disease.

This method, to be developed by a new English biotechnology company called Population Genetics Technologies, is expected to reduce significantly the cost of studying large populations of genomes. The new technology will enable users to discover extremely quickly much information about such gene variants from studies of whole populations and can be used also for a broad range of complex biological problems requiring many parallel analyses. This is because this method will allow the mixing of thousands of samples in one test tube and the simultaneous interrogation of all of them in one experiment.

This technique will permit the development of

new medicine since “it might enable the discovery of mutations, rare in a clinical trial population, but responsible for serious deleterious side effects that are discovered only when the drug is very broadly prescribed. Patients that are potentially subject to such side effects could be screened if these mutations are determined” said Dr Brenner.

The Wellcome Trust, an independent research-funding charity, will promote the research of Population Genetics Technologies.

The Viking 's route of the Multiple Sclerosis

http://www.elpais.es/articulo.html?d_date=&xref=20050329elpepisa_1&type=Tes&anchor=elpepisa

For the first time in Multiple Sclerosis (MS) researchers at the Serono Genetics Institute (SGI) localized in France have identified 80 genes involved in this chronic, inflammatory condition of the nervous system and is the most common, non-traumatic, neurological disease in young adults. The completion of the MS Whole Genome Scan in 2006 will lead to a comprehensive catalogue of potential MS drug targets. In addition, the knowledge of genetics in MS provides a basis for better designing safer and more effective drugs and enabling physicians to address unmet needs and potentially better match treatments to the individual patient

The MS is a multifactorial disease that may affect approximately two million people worldwide (30.000 in Spain) and genetics plays a significant role in the disease.

This study, directed by Professor Daniel Cohen, Vice-President and Worldwide Head of Genetics at Serono that participated in the Human Genome Project, was based in bioinformatics and biostatistical techniques that analyzed single nucleotide polymorphism (SNPs) between patients and controls.

MS is prevalent disease of temperate latitudes and of the western hemisphere. Principally, it is a disease in Europe, North America, Australia and New Zealand. Spain has a lower prevalence (60 per 100.000). “Regions between north and south

of 40-60 degrees latitude have a markedly higher incidence” says Oscar Fernandez (head of Neurology laboratory of Carlos Haya Hospital, Malaga). He bases his hypothesis in Charles Poser’s publication in Acta Neurologica Scandinavica that indicates that the highest prevalence rates for multiple sclerosis are found in Iceland, Scandinavia, the British Isles, and the countries settled by their inhabitants and their descendants, that is, the United States, Canada, Australia, and New Zealand. This suggests that the Vikings may have been instrumental in disseminating the genetic susceptibility to the disease in those areas as well as in other parts of the world. The Vikings raided in most European countries and also migrated to Asia and EEUU.

The human genome is estimated to contain 30.000 genes and 3.000 million base pairs or nucleotides and almost all (99.9%) nucleotide bases are exactly the same in all people. This small variation or single-base DNA differences (SNPs) can change the protein sequence and this way have a role in the development of any disease or affect drug metabolism.

This large-scale association study was performed in a French, Swedish and American population, including a total of 900 people with MS and an equivalent number of healthy individuals. Researchers at SGI used the Affymetrix GeneChip technology to scan over 100,000 SNPs (single nucleotide polymorphisms) to identify the genes involved in MS, comparing the genetic profile of cases and controls. The next step is to continue this endeavor, applying next generation GeneChip technology to scan over 500,000 SNPs and thus complete the MS Whole Genome Scan during 2006.

Among the genes confirmed in the inflammatory response of MS were ones that coded for the Human Leukocyte Antigen (HLA). HLA genes are central to the immune response and are involved in recognition of infectious agents (e.g. bacteria, viruses). The HLA-DR2 gene is more prevalent in control group than in patients with MS, therefore it is a protector factor against this disease, but like Hiller from the head of Neurology division of Karolinska Institute, Sweden, says its mechanisms are still unknown. This scientific study that will be published in Genes and Immunity journal

also mentions that LAG3 and IL7R genes can influence in MS.

Two genes were described for the first time involved in the neuro-degenerative pathway included neuregulin (NRG1) and Presenillin 2 was also shown to have been pivotal in Alzheimer's.

Oscar Fernandez coordinator of the International congress celebrated in Madrid in March says that the Canadian consensus (guidance in the treatment of MS) facilitates the doctors a uniform treatment based on the scientific knowledge.

New computer program uses brain scans to assess risk of Alzheimer's

http://www.azprensa.com/noticias_ext.php?idreg=15432

New York University School of Medicine has developed a brain scan-based computer program that allows researchers to standardize and computer automate the sampling of PET (technique which produces a three dimensional image or map of functional processes) brain scans.

The new program measures metabolic activity in a region of the brain affected in the early stages of Alzheimer's disease. Researchers demonstrated that reductions in brain metabolism in healthy individuals were associated with the later development of the memory robbing disease.

The study followed 53 healthy subjects between 54 and 80 years old, for at least 9 years and in some cases for as long as 24 years. Altogether there were 136 PET scans.

The researchers found that the baseline hippocampus glucose metabolism was significantly reduced in patients who develop Alzheimer. They think that this brain measure is the only that actually exist that can predicted the future cognitive decline.

"Right now, we can show with great accuracy who will develop Alzheimer's nine years in advance of symptoms, and our projections suggest we might be able to take that out as far as 15 years," says Dr. de Leon, main author of the study.

NIH Awards Four Microarray Centers \$25M for Neuroscience Research

<http://www.genomeweb.com/>

The National Institutes of Health (NIH) Neuroscience Microarray Consortium, supported by the NIH Neuroscience Blueprint, a framework to enhance cooperative activities among these Institutes, combines technology resources from the Translational Genomics Research Institute (Tgen) in Phoenix, AZ, Duke University in Durham, NC, the University of California in Los Angeles and Yale University in New Haven, Connecticut.

The Microarray Consortium was initially funded in 2002 by the National Institute of Neurological Disorders and Stroke and the National Institute of Mental Health with \$9 million from the NIH. This award is part of a greater \$25 million grant that TGen will share with three other microarray centers supported by the NIH Neuroscience Blueprint.

The consortium provides the infrastructure to enable investigators to perform microarray experiments on several different platforms at reagent cost, and to offer assistance with experimental design and bioinformatic analysis of the data. TGEN, DUKE, & UCLA delivered a new 5AM Solutions that provides software products and technology services for the biomedical research industry to facilitate more rapid progress in medical research by connecting researchers all over the world with, and ensure they can easily combine and share genomic data. It is available at:

<http://www.5amsolutions.com/>

INFOBIOMED PROFILES

Custodix



Custodix is a private limited company established in 2000 specialized in privacy protection solutions for data collection.

Today Custodix is recognized as one of the most advanced and reliable Trusted Third Party (TTP) service providers in the Healthcare sector providing both consultancy and technical privacy protection solutions. Its privacy services are supported by a combination of high-level technical and legal expertise in privacy protection and e-security and a sound experience with a number of advanced e-security solutions such as PKI, secure VPN, time-stamping and more.

Custodix has an international customer-base, deploying services in Europe, Asia and Australia. Customers are both commercial companies and governmental research organisations.

In its short existence, Custodix has already been involved in several e-privacy and e-security IST projects (FP5, FP6) and other funded research project (e.g. support from the Flemish government), including:

- PRIDEH (2001 - 2003)
Privacy Enhancement in Data-Management in e-Health
- PRIDEH-GEN (2001 - 2003)
Privacy Enhancement in Data Management in E-Health for GENomics
- PPeP (2002)
Privacy Protection in e-Pharma
- OPUS (2001 - 2004)
Optimal Public Procurement Services in the Healthcare Marketplace
- INFOBIOMED (2004 - 2007)
Network of Excellence aiming to integrate Medical Informatics (MI) and Bioinformatics (BI)

UPM - Polytechnical University of Madrid

The Artificial Intelligence (AI) Lab at the School of Computer Science, UPM (Polytechnical University of Madrid, in English) is a non-profit, public, research laboratory dedicated to 3 main areas of R&D:

- Biomedical Informatics research (distributed databases, data mining, genetic workstation, guidelines and protocols, emergency management, virtual surgery, image processing)
- Artificial Intelligence (ontologies, machine learning, games, search models, knowledge representation and acquisition)
- Neural networks and DNA-based computing

The laboratory was created in 1984 and during the last 15 years has obtained 3 international grants, from multinational companies such as XEROX and Hewlett-Packard (2), for an amount of over a million Euro. The laboratory has also obtained more than 15 grants from Spanish agencies and has participated directly in 2 ESPRIT projects (ACKNOWLEDGE and VALID). Members of the laboratory have also participated in various EU-funded additional research projects.

The AI Lab has two full professors, 4 associate professors, 2 assistant professors, 10 research associates, and around 30 graduate and undergraduate students. Collaborations include exchanges with foreign institutions, such as Harvard University (where six members of the lab have worked during the last years during mid and long-term stays), Stanford University, Georgia Tech, in the USA, and research agreements with hospitals and health institutions in Spain. The last contact has been signed with the Ministry of Defence, Spain, for the development of software to represent and disseminate over Internet emergency protocols

Collaboration with the industry includes funding and support from computer companies (E.g., Xerox and HP), and from Spanish software companies (e.g., Bilbomatica, CETTICO, and others). The AI lab has obtained various awards for excellence in research from Non-profit scientific organizations in Spain, such as the Spanish Society for Health Informatics that has given its main research award of the years 1998 and 1999 to the lab.

RESOURCES

Genetic Science Learning Center

<http://gslc.genetics.utah.edu/units/pharma/>

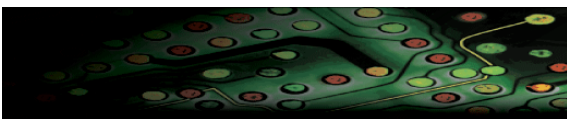


In an attempt to bring the achievements in the study of the Human Genome closer to people, the University of Utah has developed a web site, The Genetic Science Learning Center, that aids in the teaching and learning process in the field of Genetic Science. This resource uses a clear and concise language to describe and transmit scientific information. Their motto "Helping people to understand how genetics affect their lives and society" clearly expresses their main goal. They achieve this by explaining how a DNA molecule is constructed, how a gene is transcribed and translated or doing a practice of how to put an enzyme to work, just to mention a few examples of what can be done in this site.

The visit to this website is encouraged for people interested in leaning Genetic Science. It manages to combine a user-friendly aesthetic style with good contents, all in a fluid, simple and direct language.

NCI Creates Gene Expression Database of Normal Human Organ Tissue

<http://home.ccr.cancer.gov/oncology/oncogenomics/>



Researchers at the National Cancer Institute (NCI), part of the National Institutes of Health (NIH), have built the largest open-source database for normal tissue from human organs. Scientists searching for genes that go awry and cause disease can use the NCI database as a crucial point of refer-

ence because it pinpoints which genes are expressed in many of the body's major organs under normal conditions (without known disease). Scientists can compare the genes from their own biological samples to this dictionary of normal expression. "Genes identified by the database as abnormally active in a particular disease could become potential targets, guiding researchers to better candidates for new drug therapies, immune-based vaccine treatments, and potential biomarkers to help with diagnosis," explained Javed Khan, M.D., chief of the Oncogenomics Section of NCI's Pediatric Oncology Branch. A study validating the database appears in the March 2005 issue of *Genome Research*.

"The NCI database is an important addition to the growing body of knowledge about gene expression in normal human tissues," added James Jacobson, Ph.D., acting branch chief of the Diagnostics Research Branch in NCI's Division of Cancer Treatment and Diagnosis. "These data give investigators a baseline against which to compare gene expression data obtained from tumor or other disease specimens, and should be a valuable resource for the research community."

The normal organ database uses a technology known as gene expression microarrays, more commonly known as gene chips, to provide a kind of fingerprint that researchers and clinicians can use to compare cells and tissue they suspect may have cancerous or other malfunctioning genes. To create these fingerprints, Khan and his team assembled a complementary DNA (cDNA) microarray, using a pair of glass slides on which thousands of known genes have been printed in tiny spots. Cells can be tested by manipulating them so that genes activated in the cell will match up with the known gene samples, like two pieces of Velcro attaching to each other. The cellular genes are treated with fluorescence and literally light up the gene dots on the chip. The light pattern is then measured with a special type of microscope and the results are fed into a computer for analysis.

Gene expression microarrays have been used in numerous applications, including identifying novel genes associated with certain cancers, classifying tumors, and predicting patient outcome. Another NCI-funded study recently demonstrated that

microarray analysis of identical tissue samples at geographically separate laboratories can produce the same quality of results as those done within a single lab. The normal organ database takes that one step further, enabling scientists and clinicians to compare the gene expression results for their own tissue or genes of interest to a baseline standard that represents a generic picture of normal gene activity, organ by organ, in the human body. Users of the array on the new NCI web site (<http://home.ccr.cancer.gov/oncology/oncogenomics/>) will find expression profiles for 18,927 genes, which include most of the genes that are known to help direct basic activities of the human body.

Recently the Human Genome Project revealed a surprisingly low number of human genes (20,000-25,000), and Khan said it had been previously reported that "only a fraction of that, perhaps 10,000 genes, are actively transcribed in normal cell processes." Thus it becomes strategically useful to characterize this essential backdrop. "The normal organ database provides a platform that may help scientists find new targets in the cells of previously incurable cancers. The driving force of research in our section is to translate genomic information to the clinic. The goal is to save lives and improve the quality of life for children with high-risk cancer."

Until now, no publicly available, normal human organ database has used so many tissue samples (158), or included samples of tissue from different parts of the same organs from multiple donors. Tissue samples were harvested an average of 11 hours after death, from males and females of different ethnic groups, ranging from ages 3 months to 39 years old.

The very large cDNA microarray they constructed has more than 42,000 detectors built into two chips using verified cDNA libraries upon which any other researchers currently rely. Analyzing the organ tissue with this tool allowed Khan and his team to identify 18,927 genes that constitute their database. "We found that each organ had a unique expression level profile," said Khan, "and, remarkably, any truly random subset of 1,000 genes could distinguish one organ from another." Each organ revealed a very distinct profile of active genes, different from all others. However,

the gene profiles from different organs that share similar biological functions also showed patterns of expression. For example, though the cerebrum and the cerebellum are two distinct parts of the brain, located apart from each other and doing very different jobs, their gene expression profiles reflected their commonality as part of the nervous system. Similarly, "muscle contraction" genes were found in skeletal muscle, smooth muscle tissue, and the heart - all organs that share a common way of functioning.

To illustrate the kind of useful data that can emerge from using this tool, Khan's team analyzed 100 samples of the most common pediatric solid tumor cancer, neuroblastoma (NB), which accounts for 7 percent to 10 percent of all childhood cancers. Even though the tumor samples were taken from a variety of patients with different stages of cancer, the database kicked out a list of 19 genes that were consistently overexpressed compared to normal brain tissue.

GenMAPP Gene Map Annotator and Pathway Profiler

A team of scientists at the J. David Gladstone Institutes has unveiled a new version of GenMAPP, a widely used software program designed to help biomedical scientists view and analyze genome-scale data sets in the context of biological pathways.

GenMAPP 2.0, short for Gene Map Annotator and Pathway Profiler, marks the first major revision of the program, which was developed and launched by Gladstone scientists in 2002. With the program having been freely available at www.GenMAPP.org to all researchers since its debut, it has now become a standard means of depicting and sharing biological data and pathway information.

As GenMAPP developer Bruce Conklin, MD, points out, a single genomics experiment can yield enough data to fill a large telephone book, and methods for organizing and analyzing the data are desperately needed.

"Genomic experiments can easily overwhelm a scientist with data," explained Conklin, an investigator at the Gladstone Institute of Cardiovascular

Disease and UCSF associate professor of medicine, molecular and cellular pharmacology. "GenMAPP organizes the data by biological process, a scheme that most biologists understand, and allows us to find new connections that we would not have seen otherwise."

From its beginnings, GenMAPP has been designed for viewing and analyzing gene expression data on biological pathways and other groupings of genes. GenMAPP 2.0 incorporates a variety of new features, many of them suggested by users, including:

- A flexible format that accepts many different gene ID systems from resources for many species, including human, *Drosophila* (fruit fly), mouse, rat, zebrafish, *C. elegans* (a microscopic roundworm), and *S. cerevisiae* (yeast).
- Species-specific gene databases that show relationships between the various gene ID systems in the database. For example, genes on the MAPP (GenMAPP files that represent biological pathways or groupings of genes) may use a single common ID type and the expression data sets may be annotated with a completely different ID type, but GenMAPP provides an internal database that can connect the two gene IDs.
- Assistance in creating unique Gene Databases for any species, as well as customization of existing Gene Databases.
- The ability to export (as HTML) entire sets of MAPPs, including information from the researcher's Expression Dataset, enabling convenient, interactive display of data on web sites.

By viewing genes in the context of a known biological process, GenMAPP makes it possible to make sense of data that might otherwise be difficult to interpret. The most widespread alternative analytical method, hierarchical clustering, groups genes without knowledge of the gene's function, but it can miss small changes in expression. In fact, the two methods complement each other in interpreting biological data.

GenMAPP was developed with grant support from the NIH. Any scientist can use it to modify MAPPs to fit other hypotheses, to design new pathways, or to share the data with others in the research community.

The GenMAPP site has logged over 10,000 registrations to download the program, and GenMAPP has been cited as a resource in upwards of 50 publications to date.

"We have been very pleased with the widespread acceptance and use of GenMAPP," said Conklin. "This new version was created in response to comments from those many users, and I am excited about what it will do for biomedical research here and around the world."

EVENTS

GSAC 2005: Genomes, Medicine and the Environment.

Hilton Head Marriott Beach & Golf Resort
(October 17 – 19, 2005)

<http://www.venterininstitute.org/gsac/>

"Symposium Pharmacogenomics and Pharmacogenetics: Current Challenges and Bioinformatics Support".

11th and 12th of July 2005. Barcelona, Spain.

Anybody interested in participating in this Symposium please contact grib@imim.es.

IEEE Computational Systems Bioinformatics Conference.

8-11 August 2005. Stanford, California, USA.

<http://conferences.computer.org/bioinformatics/>

Beyond Genome 2005: The Future of Medicine

Monday, June 13, 2005 to Thursday, June 16, 2005.

The Fairmont Hotel, San Francisco, CA, United States

<http://www.worldpharmacongress.com/>

Personalized Medicine Europe: Health, Genes & Society

Tel-Aviv University, Tel-Aviv, Israel, Tel-Aviv University, Tel-Aviv, Israel

June 19-21, 2005

<http://www.functionalgenomics.org.uk/sections/activities/2005/Livshits/info.htm>

18th IEEE International Symposium on Computer-Based Medical Systems (IEEE CBMS 2005)

June 23-24, 2005 at Trinity College Dublin, Ireland

<http://conferences.computer.org/CBMS2005/>

ISMB 2005: 13th Annual International conference on Intelligent Systems for Molecular Biology

June 25-29, University of Michigan Bioinformatics

<http://www.iscb.org/ismb2005/>

MIE 2005: The XIX International Congress of the European Federation for Medical Informatics

28 August -1 September, Geneva Switzerland

<http://www.mie2005.net/>

7th International Meeting on Single Nucleotide Polymorphism and Complex Genome Analysis

Hinckley Island Hotel, Leicestershire, 22-24 September 2005

<http://snp2005.nci.nih.gov/home.cfm>

AMIA 2005: American Medical Informatics Association - Annual Symposium

October 22-26, Washington DC

<http://www.amia.org/meetings/annual/current/main.html>

International Symposium on Health Informatics and Bioinformatics, Turkey 05

Belek, Antalya, TURKEY, 10 - 12 November 2005

<http://hibit05.ii.metu.edu.tr/>

COMMENT

by Fernando Martín Sanchez, Isabel Hermosilla and Francisco Javier Martín

INFOBIOMED Network aims to set the foundations for Biomedical Informatics (BMI) in Europe.

It is funded by the Directorate-General Information Society of the European Commission within the 6th Framework Programme for Research and Technological Development.

Bioinformatics (BI) and Medical Informatics (MI) have come closer bridging the gap in genomic medicine through the use of Information and Communication Technologies. INFOBIOMED is situated in the imaginary limits that border BI and MI to create a single area of interoperability and compatibility called BMI.

The main purpose is to shape the common European structure of the discipline that integrates the huge amount of genetic information that came out of the human genome into daily clinical practice thanks to the use of information and communication technologies (ICT) to improve Genomic Medicine.

The idea of the Network originally rises from the agenda established by the BIOINFOMED project (1) that identified 18 research lines relevant for the development of this new discipline.

Figure 1, below, accurately summarizes the architecture proposed in BIOINFOMED where all possible synergies between BI and MI were analysed. In the bottom plane and supporting all the structure are the ICT enabling the interaction by sharing both common elements of both disciplines.

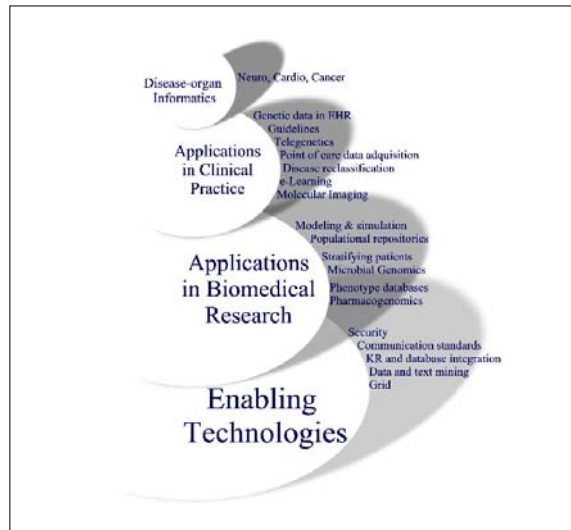


Fig. 1 Technologies

In the level immediately above are those technologies that are applied in Biomedical Research. They are focused towards gathering information that could be subject to potential application in Clinical Practice, that represents the next level.

The final stage of this figurative pyramid is Disease Organ Informatics, specific models for the pathology to which they are destined and it crystallizes in clear and solid examples such as Neuroinformatics, Cardioinformatics or Cancer Informatics.

The INFOBIOMED Network of Excellence (NoE)(2,3) titled "Structuring European Biomedical Informatics to Support Individualised Healthcare" has as its main goal to carry out the vertical integrative work of four pilot applications focused on different areas: pharmainformatics, infectious diseases, chronic inflammation and colon cancer with genomics.

The network will use existing methods, tools and technologies and will also aim to develop new ones needed to stimulate the convergence of the two areas, BI and MI, that complement each other in solving real problems in research and healthcare practice.

COMMENT

More specifically, the challenge of INFOBIOMED is to infer, from the experience of the four pilots, the basis that will establish collaboration lines for the discovery and creation of novel diagnostic and therapeutic methods based on techniques and common tools in BI and MI.

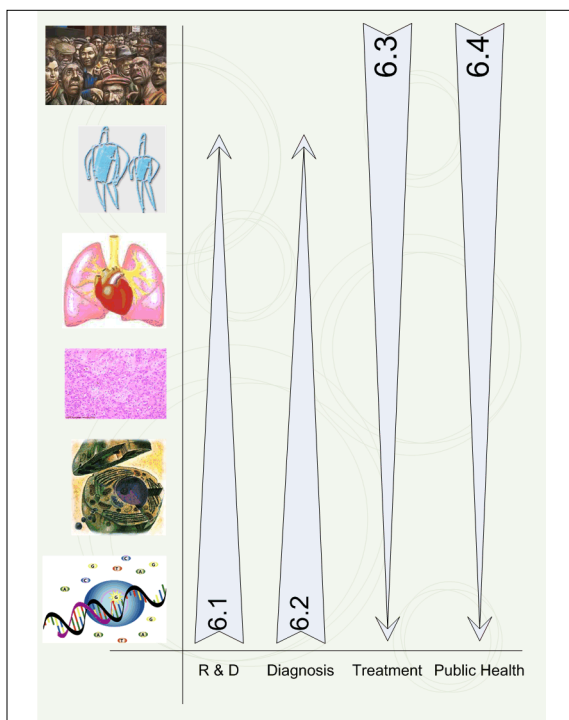


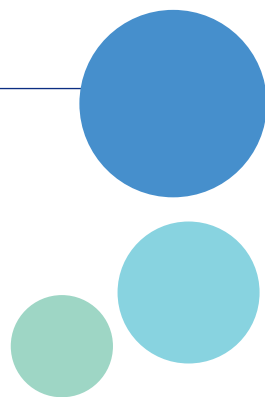
Fig. 2 INFOBIOMED Pilots and their relations to health sciences from the population to the molecular level

(1) Fernando Martín-Sánchez et al. "Synergy between Medical Informatics and Bioinformatics: facilitating genomic medicine for future health care". *J Biomed Inform.* 2004 Feb; 37(1):30-42.

(2) Sanz F, Diaz C, Martin-Sanchez F, Maojo V. "Structuring European biomedical informatics to support individualized healthcare: current issues and future trends". *Medinfo.* 2004;2004:803-7

(3) F. Martin-Sanchez, I. Hermosilla, F.J. Vicente and the INFOBIOMED Consortium The European NoE INFOBIOMED: Mapping Biomedical Informatics Technologies to Research and Clinical Practice Applications. IMACS July 2005, Paris, France.

BIOMEDICAL INFORMATICS
TO SUPPORT INDIVIDUALISED HEALTHCARE



BIOMEDICAL INFORMATICS TO SUPP

<http://www.infobiomed.org>

