

# INFOBIOMED

## NEWSLETTER

N.5/6 MARCH/JUNE 2006



```
TCCTCCTAGCAACGCGGCCCTCTCGGCTCTTAAGGCGTGACTTGCCATGGATGTTCTGCC  
TTTGGAAACCAAGTGGGTGGTCACAGCACAGAGGCTGTTGGTTTCATGGCTGTCACATGCAC  
AGAGACAGCCCTGTCCACTGAGCCTGTTCTGTGGCAGAAGGAACCCCTAGCCTGGACC  
TCCTTGGGTGGAAACACCAGCCCTCAGCTCTGACTGTCATTTTCAGATGTGAAGAGAA  
ACACAGTGTGCTGATGTTGTATTGCCGTGAACAACGAGAGTTTCTTTGACTTC  
TAACTGCCTTAAAGCTAGAGCCCGGCATTAGGATGGCAGTGGTGGTAATGAGGTTATG
```

**BIOMEDICAL INFORMATICS  
TO SUPPORT INDIVIDUALISED HEALTHCARE**

**SUPPORT INDIVIDUALISED HEALTHCARE**



# SUMMARY

## NEWS

PG. 2

In search of Microarray Standards	pg. 2
A cyber infrastructure network for Europe: ESF Forward Look launched in order to keep Europe at the forefront of computational science	pg. 2
caBIG participation: Management, Exchange and Integration of Data for Cancer Research on the Grid	pg. 3
Genomic Researchers head to New England schools for National DNA Day	pg. 3
Nanotechnology may find disease before it starts	pg. 4
Molswitch: a Molecular Magnetic Switch that links the Biological and Silicon Worlds	pg. 5
Researchers simulate complete structure of Virus on a computer	pg. 5
IMIA Biosurveillance Task Force Initiated	pg. 6
Web model of Influenza-Host Lifecycles will aid scientists in creating anti-viral drugs	pg. 6
Faster computation of Haplotypes provides insight into genetic basis of Human Disease	pg. 7

## INFOBIOMED PROFILES

PG. 8

ACTA and VU University Medical Center	pg. 8
---------------------------------------	-------

## RESOURCES

PG. 9

The National Genetics education and Development Centre website	pg. 9
GenePING: secure, scalable management of personal genomic data	pg. 9
Stratagene launches Pathway Architect (TM) Software	pg. 10
Office of genomics and disease prevention (OGDP) at CDC	pg. 10

## EVENTS

PG. 11

## BMI INITIATIVES IN EUROPE

PG. 11

Network of excellence Semantic Mining	pg. 11
---------------------------------------	--------

## INTERVIEW

PG. 13

## INFOBIOMED PILOT

PG. 14

## NEWS

### LAB TOOLS

#### In search of Microarray Standards

**Source:** The scientist, April 2006 issue 4. In Search of Microarray Standards, an industry/academia/government coalition puts microarray reproducibility to the test. Jeffrey M. Perkel.

<http://www.the-scientist.com/toc/2006/4/>

DNA microarrays are useful tools widely used nowadays in many laboratories, however due to the lack of standards; the results obtained in one lab may significantly vary from those obtained in another. To tackle this problem a group of researchers of the US Food and Drug Administration's National Center for Toxicological Research, led by Leming Shi, is working on the Microarray Quality Control (MAQC) project. They are preparing ten papers to be submitted jointly to Nature Biotechnology for their publication next September.

The aim of the MAQC project is clearly stated in its website (<http://www.fda.gov/nctr/science/centers/toxicoinformatics/maqc/>): *"to establish QC metrics and thresholds for objectively assessing the performance achievable by various microarray platforms and evaluating the advantages and disadvantages of various data analysis methods"*.

The article by Jeffrey M. Perkel published in the Scientist mentions three basic goals of this project: to select and mass-produce a pair of publicly available standard RNA samples; to collect replicate datasets using those samples on eight array and three alternative quantitative gene expression platforms; and finally to analyze that data to produce a benchmark standard. It also includes comments by some of the researchers involved in the project such as Jensen from the University of Massachusetts, Boston who mentions the importance of the support given by the big companies like Affymetrix, Agilent, and Illumina, who could work together *"for the good of the industry"*.

It is thought that once the results of the project are published the labs will be able to run the standard

RNAs that are available from Ambion and Stratagene and check their own results against the published ones. This does not mean that there will be soon an official MAQC verification mark. However, this project might encourage others to develop a test useful for verifying results.

Although this might not solve all data issues, it might be another important step towards understanding the variability in results between and within labs and it will provide a needed push forward to the field.

#### A cyber infrastructure network for Europe - ESF Forward Look launched in order to keep Europe at the forefront of computational science

**Source:** European Science Foundation News release - issued 13/04/2006

<http://www.bioinfo-online.net/modules/news/article.php?storyid=803>

Posted by bioinfo on 2006/4/14

The European Science Foundation issued a press release announcing the launch of a new Forward Look during this year (2006) that will boost computational science in Europe.

Europe still has an important role in this field *"particularly for the simulation of materials at the atomic scale"* but it has lacked up to now of a coordinated European action. It is thought that this initiative will tackle this gap by helping keep Europe in its leading role in the development of computational techniques, programs and hardware. The program will be a milestone towards the creation of a coordinated cyber infrastructure at a European level by uniting efforts to provide policy makers with clear guidelines of the needs of researchers. It will set the basis for developing the mentioned coordinated European cyber infrastructure which will provide researchers with an 'e-environment' of supporting services, more programs, codes and hardware to carry out their work. *"This infrastructure also has the potential to create a platform for code developers in Europe to work in a collaborative and integrated way"*.

The press release included a comment by Professor Berend Smit of CECAM (Centre Européen de Calcul Atomique Moléculaire) in Lyon, France, the principal scientist who proposed the forward look. He explained that there are three main incentives for this Forward Look. He said that the first one is to raise awareness that this field is changing and that there is a need for this infrastructure to be put in place. The second emphasizes the importance of making sure that the development of this infrastructure is done at the European level, so that all of the support systems will be effective throughout Europe to enable all scientists to work more effectively and to make collaborative work possible. And the third incentive would be that if this infrastructure is successfully developed, we should see future scientific advances, which would not be possible in the current climate.

Another advantage of this Forward Look mentioned is that it aims to encourage the scientific community to connect directly with the policy makers.

### caBIG participation: Management, Exchange, and Integration of Data for Cancer Research on the Grid



**Sources:** [http://bmi.osu.edu/areas\\_and\\_projects/caBIG.cfm](http://bmi.osu.edu/areas_and_projects/caBIG.cfm),  
caBIG web site: <https://cabig.nci.nih.gov/> and  
Mobius Project website: <http://msc.osu.edu/projects/mobius.cfm>

The National Cancer Institute (NCI) announced in July 2003 the creation of the Cancer Biomedical Informatics Grid (caBIG), a biomedical informatics network for cancer research. The **cancer Biomedical Informatics Grid**, or **caBIG™**, is described in its web page as “a voluntary network or grid connecting individuals and institutions to enable the sharing of data and tools, creating a World Wide Web of cancer research. The goal is to speed the delivery of innovative approaches for the prevention and treatment of cancer”. For this

purpose the network will connect many institutions, people, tools and data. This will pose a big challenge; it has to develop a software system capable of handling biological data from different and heterogeneous sources, considering that this variation will require to set clear standards. The approach is to leverage existing technologies for the standardization of data representation and semantics and couple them to a common data integration architecture related with cancer and such a way to create a World Wide Web of cancer research that also be available to non scientific people.

Due to the great amount of disparate data sources the project must face the challenges extend beyond the structure and semantics and also count with the evolution of data types. The caBIG project agreed on the use of XML for metadata and data exchange format.

Some of the other members of this initiative are Ohio State University Medical Center’s BMI and Multiscale Computing Laboratory. They contributed their Mobius framework to the caBIG. Mobius is a project dedicated to develop “an array of tools and middleware components to coherently share and manage data and metadata in a Grid and/or distributed computing environment”. The caBIG infrastructure was developed taking the Mobius architecture as a framework, and its first public version (0.5) was released September 9, 2005.

### Genomic researchers head to new England schools for National DNA Day.

### Celebration Encourages Exploration of Genomic Careers



**Source:** National Human Genome Research Institute (NHGRI);  
NIH News <http://www.genome.gov/18516592>

The National Human Genome Research Institute (NHGRI) from the National Institutes of Health (NIH) participated in the fourth annual National DNA Day (April 25). This celebration began in April 2003 with the commemoration of the completion of the Human Genome Project and the anniversary of the discovery of DNA's double helix in 1953. The main purpose is to bring genetics closer to high school students by giving them the opportunity to access the information directly from professionals.

Many schools, mostly in the New England area, participated and were visited by the NHGRI's DNA Day Ambassadors. A list of all the participant schools in the different states is available at <http://www.genome.gov/18516588>. The content of the meetings were quite diverse just like the background of the ambassadors, ranging from basic to clinical research.

The NHGRI, the American Society of Human Genetics, the Genetics Society of America, the Genetic Alliance and the National Society of Genetic Counselors collaborate in the events taking place this day with the participation of more groups each year, for example this year researchers from the University of Vermont got involved in this initiative.

This day is the perfect opportunity for "NHGRI's DNA Day Ambassadors" to transmit their personal knowledge and experiences in genomic research to students of dozens of American high schools, exciting their interest in genetics and informing them about possible career options. Some local genetic experts from the American Society of Human Genetics participated in the DNA Day together with the NHGRI Ambassadors. In their case their collaboration will extend in time, through their Genetic Mentorship Network, that gives the teachers the opportunity of using them as a resource throughout the school year.

Barbara Biesecker, a genetic counsellor at NHGRI participated in the development of a multimedia webcast presentation called "Genomics: Towards a Healthier You" that shows the relevance of genetics in health and information about genetic

counselling accessible through a video postcard in NHGRI website: <https://genome.gov/DNAday>. Other features that can be found in the DNA day Web site are free educational tools on genetics and genomics available for teachers and students and a transcript of the online chat that took place on April 25th.

## Nanotechnology May Find Disease before it Starts

### Source:

[http://www.bio.com/newsfeatures/newsfeatures\\_research.jhtml;jsessionid=EEYWADVUW1AJBR3FQLMCFEWHUWBNSIV0?cid=18900006](http://www.bio.com/newsfeatures/newsfeatures_research.jhtml;jsessionid=EEYWADVUW1AJBR3FQLMCFEWHUWBNSIV0?cid=18900006) (04/25/06) and Liu J, Levine AL, Mattoon JS, Yamaguchi M, Lee RJ, Pan X, Rosol TJ. Nanoparticles as image enhancing agents for ultrasonography. *Phys Med Biol*. 2006 May 7;51(9):2179-2189. Epub 2006 Apr 10.

A recent paper by Jun Liu and his collaborators published in the journal *Physics in Medicine and Biology* shows for the first time that nanoparticles in the body enhance and improve the quality of the images obtained in ultrasound. This study (supported by the Susan G. Komen Breast Cancer Foundation, the National Cancer Institute, the National Center for Research Resources and the National Science Foundation) has as its main long term objective to use this method for very early diagnosis of serious diseases such as cancer.

Knowing that nanoparticles are small enough to go beyond capillary vasculature, researchers wanted to find out if ultrasound would be able to sense them. Liu and his group injected a polystyrene nanoparticles suspension in a saline solution into the tail vein of each mouse. The mice were then anesthetized and ultrasound images of their livers, where the particles accumulated, were taken every five minutes for 90 minutes. These images were compared with those taken from the control group (injected only with saline solution). Ultrasound images of the test group were brighter as the result of the grey scale enhancement effect caused by the sound waves bouncing off a hard surface.

Future works will focus in other places of the body trying to make the particles accumulate in any tissue researchers want to study, eventually allowing to see changes in the tissues. This will permit an early diagnosis of diseases, even at the cellular level said Jun Liu in the news article in bio.com. This technique might in the future provide a means of detecting tumours thus diminishing the need to use invasive procedures such as biopsy to do so. Another potential use of nanoparticles is as drug agents delivering the product directly where it is needed. This might help decrease the side effects of chemotherapy.

### MOL SWITCH - A Molecular Magnetic Switch that links the Biological and Silicon Worlds



**Source:**

[http://www.bio.com/newsfeatures/newsfeatures\\_research.jhtml?cid=18800036](http://www.bio.com/newsfeatures/newsfeatures_research.jhtml?cid=18800036) and <http://www.nanonet.org.uk/molswitch/> (Information Society Technologies, European Commission)

The MOL SWITCH project is a three year project, EU funded under the European Commission's FET (Future and Emerging Technologies) initiative of the IST programme. The project ended at the end of 2005. In its webpage the consortium says that the main aim of the project was *"to produce a nano-actuator, based on a biological molecular motor that moves DNA, fuelled by ATP, which is able to detect the movement of biocompatible magnetic nano-particles attached to the DNA. This will allow the development on an integrated system that could be used for single-molecule DNA sequencing"*. The project had two phases, the demonstration of a Mol switch device and the demonstration of single molecule DNA sequencing.

Although the use of Nanotechnology is very complex, the project team developed a Mol Switch Device. They used a microfluidics chip that works with a number of DNA-based motors. They

translocate DNA, they initially bind DNA at a specific site and then they translocate the rest of the DNA through the bound motor. The main element is that *"the free end of the DNA is attached to a paramagnetic bead, which can be used to stretch the DNA, but can also produce a molecular dynamo effect allowing an electronic output from the moving magnet"*.

The potential number of applications for the device is very large, flow-control valves, pumps, positioning drives, motors, switches, relays and biosensors. One important application for the Mol Switch is DNA sequencing. In their work the research team used the Mol-Switch with time-resolved fluorescence for DNA sequencing but *"more work needs to be done"* said Dr Keith Firman, Reader in Molecular Biotechnology at Portsmouth University and coordinator of the project. He added that they are applying for a new project under the [European Union's] New and Emerging Science and Technology (NEST) scheme.

### Researchers Simulate Complete Structure of Virus. On A Computer

**Source:**

<http://www.bio.com/realms/research.jhtml?realmld=3&cid=17900003>

Computational biologist researches and crystallographers from the University of Illinois at Urbana-Champaign and the University of California respectively have developed a computer program applying engineering knowledge to study if an entire life form works in silico, in this case viruses and their infection, proliferation and re-infection into other host cells mechanisms. More specifically in this study, the satellite tobacco mosaic virus was selected for its simplicity and small size. They used reverse engineering to study the environment surrounding the virus and its dynamics of providing an unprecedented view of these processes. This work was published in the Journal Structure in March.

They created artificial nanomachines made of capsids, each a *"small protein shell that contains*

a viral building plan, a genome, in the form of DNA or RNA". According to co-author Peter Freddolino, a doctoral student in biophysics and computational biology at Illinois in the comments published in bio.com news "the simulations followed the life of the satellite tobacco mosaic virus, but only for a very brief time".

Another co-author Klaus Schulten, Swanlund Professor of Physics at Illinois said that "a big first step has been taken to 'test fly' living organisms" and that "this step will assist modern medicine as we continue to learn more about how viruses live." All simulations were carried out by a computer program operated by the National Center for Supercomputing Applications in Schulten's Theoretical and Biophysics Group's lab at the Beckman Institute for Advanced Science and Technology.

## IMIA Biosurveillance Task Force Initiated



**Source:** Methods of Information in Medicine 3/2006

The real possibility of a pandemic of any emerging infectious disease such as the influenza H5N1 is making people realise the need to standardize and coordinate biosurveillance for public health reporting infrastructure.

It is thought that these efforts should not only include the evaluation of terminology, information interchange and security standards but it should consider "open source options that have potential to serve as a world-wide infrastructure to support bio-surveillance".

In this line of thought and in accordance with the proposed IMIA Strategic Planning "Nancy Lorenzi, IMIA President has authorized a task force comprised of Working Group chairs to evaluate such collaborative efforts for biosurveillance". This task force will have two specific initial activities: to develop a paper regarding the creation of a world wide open-source informatics infrastructure and "to evaluate requirements and encourage

an open source sharing mechanism for genomic identification of bacterial and viral pathogens".

The Task Force will be chaired by Floyd Eisenberg, MD, MPH, currently serving on the IMIA General Assembly as representative of Siemens Medical Solutions, who also serves on the IMIA Strategic Planning Committee. Other members include leadership from eight IMIA Working Groups as follows:

- :: Open Source Health Informatics - Dr. **Peter J. Murray**, PhD, RN, CertEd, FBCS CITP  
Lincoln, Lincolnshire, UK
- :: Informatics in Genomic Medicine (IGM) - Dr. **Fernando J. Martin-Sanchez**, PhD - Madrid, Spain
- :: Intelligent Data Analysis and Data Mining (WG 03)  
Dr. **Riccardo Bellazzi**, PhD, Prof. - Pavia, Italy
- :: Security in Health Information Systems (WG 04)  
Dr. **Jochen R. Moehr**, MD, PhD - Victoria, BC, Canada
- :: Primary Health Care Informatics (WG 05) - **Sheila Teasdale** - University of Nottingham, Nottingham, UK
- :: Health Informatics for Development (WG 09) - Prof. Dr. **George I. Mihalas**, PhD - Timisoara, Romania
- :: Health Information Systems (WG 10) - Prof. Dr. **Klaus Kuhn** - Munich University of Technology, Munich, Germany
- :: Standards in Health Care Informatics (WG 16) - Prof. **Michio Kimura** - MD, PhD - Hamamatsu University, Hamamatsu, Japan

## Web Model of Influenza-Host Lifecycles Will Aid Scientists in Creating Anti-viral Drugs

**Source:** bio.com News 01/31/06  
(<http://www.bio.com/realm/research.jhtml?realmid=3&cid=16500026>) and the Reactome webpage ([www.reactome.com](http://www.reactome.com))

The online map of the influenza lifecycle is a part of the Reactome project. The availability of this map will give researchers the opportunity to look at specific reactions. Individual reactions are linked with associated reactions in diagrams resembling constellations giving the diagrams the nickname "starry sky".

The influenza lifecycle map has been developed by the University of Texas Southwestern Medical Center at Dallas as part of the BioHealthBase Bioinformatics Resource Center project, in a collaboration between Dr. Scheuermann's lab and a bioinformatics software engineering team at Northrop Grumman Information Technology funded by the National Institute of Allergy and Infectious Diseases.

*"Understanding how influenza reacts with its host is really critical for the rational design of anti-viral drugs and new vaccines,"* said Dr. Richard Scheuermann to bio.com news. He also mentioned the importance that *"the influenza pathway component and a simultaneous HIV database going online mark the first time Reactome has displayed interactions between an infectious pathogen and its host"* and he says that *"in the future, the goal is to build out all of the host-pathogen reactions"*. The influenza database will be updated as more information is gathered by the researchers.

The Reactome project was created in 2004. This project is a collaboration between Dr. Lincoln Stein's laboratory at Cold Spring Harbor Laboratory in New York, the laboratory of Dr. Ewan Birney of the European Bioinformatics Institute, part of the European Molecular Biology Laboratory, in Hinxton, UK and the Gene Ontology Consortium. The main aim for the project is to develop a curated resource of core pathways and reactions in human biology that are available in the project's webpage [www.reactome.com](http://www.reactome.com). The purpose is to help researchers and other interested stakeholders understand the concepts and biochemical and cell death pathways primarily in humans. The information in the databases is constantly updated and cross-referenced with PubMed, GO and with the sequences in the databases at NCBI, Ensembl and UniProt.

## Faster Computation of Haplotypes Provides Insight into Genetic Basis of Human Disease

### Source:

<http://www.bio.com/realm/research.jhtml?realmId=3&cid=14400003>

To better understand human variation and the genetic basis of complex diseases it is important to deepen in the understanding of the individual variation at the DNA level and in particular haplotypes and the possible connections with diseases. Up to now a great number of single nucleotide polymorphisms SNPs have been discovered and are currently available in different databases one of them being the NCBI's dbSNP public database. To obtain information about the different haplotypes the data should be processed and up till recently all available methods were expensive and costly.

A recent study by experts of two California research Institutes has solved the problem at a relatively low cost. The team from the University of California San Diego (UCSD) and the International Computer Science Institute (ICSI) used a computational tool to run a HAP algorithm to compute the entire dbSNP database of the National Center for Biotechnology Information (NCBI), part of National Institute of Health's National Library of Medicine. It took only 24 hours to process all 286 million human genotypes in this database.

The research team has seen that the HAP algorithm works well with diverse data sets by taking into account their different characteristics, and that it is able to deal with different types of data (related and unrelated individuals). They also used the algorithm to work with three genotypes when they are from the mother, the father and their child.

The information obtained will provide, according to lead author Eleazar Eskin, a professor of computer science and engineering at the University of California, San Diego who is affiliated with the California Institute for Telecommunications and Information Technology (Calit2), a deeper understanding of the data which *"will improve the design of studies that look for associations between certain genes and disease or inherited conditions"*. The results of this work were published in the Journal Genome Research, in a special issue on "Human Genetic Variation".

## INFOBIOMED PROFILES

### ACTA and VU University Medical Center



Periodontitis is a destructive inflammatory disease of the supporting tissues of the teeth. It is caused by bacteria and affects a substantial part of the human population. Over the age of 45 years it is a major cause for tooth loss. This research program is devoted to clarify the individual differences in susceptibility to periodontitis and to investigate the mechanisms involved in periodontal breakdown.

The Departments of Periodontology and Oral Microbiology at the Academic Centre for Dentistry in Amsterdam (ACTA), and the Laboratory of Immunogenetics at the VU Universiteit Medical Center collaborate with the aim to continue investigation into the etiology of periodontitis. The collaboration will focus on the relative importance of modifying disease genes, bacteriological and environmental factors (in particular smoking and stress). Epidemiological studies are well underway at these centres. Existing data banks need integration, further genotyping and a modern bioinformatics approach for data analysis. The developed Periodontitis Datawarehouse will be submitted to the Network of Excellence (NoE) INFOBIOMED, with the purpose that biomedical informatics can contribute to studies into the multifactorial etiology of periodontitis, and also very importantly, can determine new classification systems. Vice versa, periodontitis is an excellent model for studies in biomedical informatics due to its complexity and clear interactions between host genetics, bacterial infection and environmental factors. Knowledge obtained in periodontitis will enhance studies into other chronic inflammatory diseases.

The ACTA-group within INFOBIOMED consists of scientists from ACTA and VU Universiteit Medical Center. Both institutes are located in the West part of Amsterdam, The Netherlands. ACTA was founded in 1984 through a merger of the two dentistry faculties of the Universiteit van Amsterdam and the Vrije Universiteit, Amsterdam.

ACTA conducts scientific research, educates graduate and post-graduate students and provides patient care in the field of dentistry. It has a

staff of 540 and a student body of 850. With a student intake of 158 in 2001, ACTA is one of the biggest dentistry education & training programmes in the world.

Dental research at ACTA focuses on the study of health and diseases of the tissues in and around the oral cavity. Besides infectious diseases like dental caries and periodontal inflammatory processes, attention is paid to the development, function and regenerative capacities of the hard tissues in general, dysfunction of the masticatory system and diseases of salivary glands and oral mucosa. It is the general aim to improve strategies for diagnosis and prevention of diseases and functional repair of the affected tissues in and around the oral cavity.

The Vrije Universiteit hospital at its official opening in 1966 adopted the motto 'Medicina Misericordiae Ministra' – Medicine is the servant of mercy. The motto characterizes the Christian tradition from which the Vrije Universiteit hospital sprang. The hospital was finally able to open its doors to patients thanks to the Association for Christian Higher Education and the Christian women's organization 'Vrouwen VU Hulp' (Women's Aid for the VU). These organizations were also involved in the foundation of the Faculty of Medicine in 1950. The VU University Medical Center has been known by this name since 1 January 2001. That date marks the integration of the VU hospital and the Faculty of Medicine of the Vrije Universiteit.

As an academic centre, the VU Universiteit Medical Center is specialized in helping patients with uncommon clinical pictures and those who require highly intensive and complex care. These types of care are referred to as 'tertiary clinical care' and 'tertiary referral-centre care'.

Its tertiary clinical care and tertiary referral-centre care functions emphasize both the VU Universiteit Medical Center academic nature and its leading role in the Dutch health service. This often involves unique diagnosis and treatment, and specialized research. Naturally, the VU Universiteit Medical Center also functions as a hospital, with an outpatients department, where patients can obtain normal specialized care. The other principal tasks of the VU Universiteit Medical Center are scientific research, academic programmes and training.

## RESOURCES

### British NHS



**Source:** The National Genetics Education and Development Centre website (<http://www.geneticseducation.nhs.uk/>)

The National Genetics Education and Development Center website does not only give technical support to those that are learning or teaching genetics within the healthcare system but it also provides them with educational material. In addition, it serves as an information resource for those centres providing genetic resources and those clinicians that apply genetics in their clinical practice. It carries out several tasks in conjunction with other groups in the UK. These tasks are, as described in their website:

- :: "To identify genetics knowledge, skills and attitudes useful for clinical roles
- :: To facilitate the integration of genetics into curricula and courses
- :: To develop a framework for competencies in genetics
- :: To identify and develop resources appropriate to the needs of health professionals"
- :: To provide support and aid in the learning process sprung from the initiatives to develop services.

This center works with specialists in different disciplines such as dieticians, physicians, healthcare workers such as nurses, pharmacists and service development project managers to provide the necessary information about the attitudes around genetics and to develop educational programs. The center organizes seminars and conferences. One of the objectives of the center is to keep the needs of the patients in mind. The center encourages patient participation to better learn what patients expect healthcare professionals to know

about genetics, to learn from their experiences and to get opinions and suggestions that might help improve the website.

*"The Centre was funded by the Department of Health as one of the major initiatives of the 2003 Genetics White Paper, **Our Inheritance, Our Future - Realising the potential of genetics in the NHS** which set out the Government's strategy for ensuring that the potential benefits of genetics are realised by the NHS. A major aspect of this plan is to improve the understanding of genetics among all health professionals and its role in modern healthcare".*

### GenePING: secure, scalable management of personal genomic data

**Source:** <http://www.biomedcentral.com/1471-2164/7/93> article by Ben Adida and Isaac S Kohane. BMC Genomics 2006, 7:93 doi:10.1186/1471-2164-7-93

The PING (Personal Internetworked Notary and Guardian) allows patients and health-care providers to share health record data with access control rules defined by the patient. All data is exchanged in XML with publicly-defined schemas, and the protocol is implemented using XML over HTTP (preferably HTTPS).

A new extension of PING called GenePING that fits into the standard PING architecture and is available at <http://ping.chip.org/geneping.html> has been developed. This is a new personal health record system that securely supports personal health information and efficiently shares it among authorized health care providers, also allowing the retrieval of individual data points such as SNPs or rare mutations when requested. One of the most important objectives of GenePING is to protect patient's privacy.

GenePING is also storage system for genomic data and it is able to create large genome datasets with clinical and genomic personal characteristics. This is an important feature since nowadays genomic data is becoming a significant part of the decision support system for clinicians, hen-

ce the need of developing systems that allow the storage of individual patient's data derived from full genome expression profile and genotyping.

### Stratagene Launches PathwayArchitect(TM) Software



**Source:**

<http://news.biohealthmatics.com/PressReleases/2006/01/18/000000004026.aspx>

Stratagene Corporation has developed a new software package called PathwayArchitect(TM) Software. It is supported by MySQL(R) and in the future also by Oracle(R) databases. The software is compatible with all platforms (Mac(R), Windows(R), and Linux(R)). It is also integrated with Stratagene's ArrayAssist(R) suite of software for microarray data analysis. These technologies are co-developed with Strand Life Sciences. PathwayArchitect(TM) Software is focused in the study of biologic pathways and it aims to become the new standard in pathway analysis. It has a biology-focused workflow and a database that includes more than one million biological interactions. It can also analyze biological interactions through its Relevance Interaction Network, an algorithm that identifies the most relevant protein molecule related in complex biology process. Stratagene has included in all its software products a new scripting engine that allows "scripting of complex functions and user-defined workflows."

### Office of Genomics and Disease Prevention (OGDP) at CDC



**Source:** Informatica y Salud and <http://www.cdc.gov/genomics/default.htm>

The Office of Genomics and Disease Prevention (OGDP) at Centers of Disease Controls and Prevention (CDC) was established in 1997 and its website (<http://www.cdc.gov/genomics/default.htm>) provides information about human genomic discoveries. Its vision is "to improve population health and prevent disease through the application of genomic information" and its goals are:

- :: "Integrate genomics into public health research,
- :: Assess the value of family history and genomic tests for population health, and
- :: Incorporate genomics into public health practice".

To go beyond gene discovery and to be able to apply new genetics knowledge in support of public health, Clinical and epidemiological studies need to be able to evaluate the interaction between genetics and the environment as a cause of diseases and to assess the validity and usefulness of genetic tests. Also, there is a need for public health policies to be able to answer social, ethical and legal challenges associated to genetics as well as to guarantee the equity in the access to genetic services. Another aspect to the integration of genetic information in public health is the education of people working in the public health field as well as all aspects related to data and information communication to the general public.

With a relevant role of public health authorities, discoveries in human genetics could be translated into disease prevention programs and in actions focused on avoiding diseases based on the individuals' genetic make-up. The activities carried out by the OGDP show the scale of the challenge faced by public health nowadays. As an example, the Human Genome Epidemiology Network or HuGENet is a global collaboration of individuals and organizations dedicated to the evaluation of the impact of human genetic variation in the health of the population and in how genetic information can be used for improving health and to prevent disease.

The commitment of the OGDP is to answer to the challenge that is the use of genetic information to improve health and prevent diseases in this century.

## EVENTS



**4th International DNA Sampling Conference**  
June 4 - 7, 2006 – Montreal - Canada  
<http://www.humgen.umontreal.ca/events/dnasampling/>

**SFTeHIN: Society for Telemedicine and eHealth in Nigeria**  
June 19 - 21, 2006 - Abuja, Nigeria  
<http://www.sftehin.org/>

**CBMS 2006: 19th IEEE International Symposium on Computer-Based Medical Systems**  
June 22 - 23, 2006 - Salt Lake City, Utah, United States  
<http://cbms2006.ece.byu.edu/>

**IPMU 2006: Information Processing and Management of Uncertainty in Knowledge-Based Systems**  
July 2 - 7, 2006 - Paris, France  
<http://ipmu2006.lip6.fr/index.php>

**DILS 2006 - 3rd International Workshop on Data Integration for the Life Sciences**  
July 20 - 27, 2006 - Hinxton, United Kingdom  
<http://www.informatik.hu-berlin.de/dils2006/>

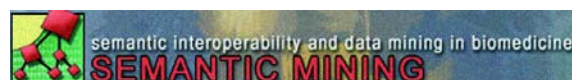
**11th International Congress on Human Genetics**  
August 6 - 10, 2006 - Brisbane, Australia  
<http://www.ichg2006.com/>

**ECCB 06: 5th European Conference on Computational Biology**  
September 10 - 13, 2006 - Eilat, Israel  
<http://www.eccb06.org/>

**BioDigital 2006**  
October 11 - 13, 2006 - Freiburg, Germany  
<http://www.messe-freiburg.de/?src=kq1025poe>

## BMI INITIATIVES IN EUROPE

### Network of Excellence Semantic Mining



<http://www.semanticmining.org/>

The aim of the Network of Excellence "Semantic Interoperability and Data Mining in Biomedicine - SEMANTIC MINING" is to establish Europe as the international scientific leader in medical and biomedical informatics. The long-term goal of the network is the development of generic methods and tools supporting the critical tasks of the field; data mining, knowledge discovery, knowledge representation, abstraction and indexing of information, semantic-based information retrieval in a complex and high-dimensional information space, and knowledge-based adaptive systems for provision of decision support for dissemination of evidence based medicine. The NoE is a response to the strategic objectives addressed in the IST call 1, areas "eHealth" and "Semantic-based Knowledge Systems".

The general objective of the network is to bridge gaps in the European research infrastructure and to facilitate cross-fertilisation between scientific disciplines. Traditionally academic departments in the domain have their roots either in computer science, system engineering (including a variety of engineering disciplines) or in a medical or clinical context. The network is composed of partners from these scientific areas, all bringing their experience and in-depths knowledge together into a common framework. An important aspect of this is the merging of medical or clinical informatics and bioinformatics including the new fields of genomics and proteomics. Another bridging activity addressed is knowledge-transfer and co-operation between academia and organisations in the health and welfare sector, such as standardisation bodies and public and user-driven health care organisations.

The NoE is based on the partnership of 25 participants from 11 European countries. SemanticMining is coordinated by Linköping University, Sweden.

## Joint Research

The research activities in SemanticMining have for the two first years been focused around seven areas (work packages):

- :: principles in ontology engineering (WP21),
- :: evaluation of SNOMED CT (WP22),
- :: impact of ontologies on health statistics (WP23),
- :: terminology systems in laboratory medicine (WP25),
- :: the construction of a multi-lingual medical dictionary (WP20),
- :: text mining and information retrieval in bioinformatics (WP24),
- :: the concept-based electronic health record (WP26).
- :: As a result of interaction between work packages, two new cross-work package activities have been added for the last phase of the project:
- :: medical terminology for laymen (WP27 ~ WP20, WP26).
- :: terminology services (WP28 ~ WP21, WP22, WP26).

Researchers in the network play an influential role in the process of harmonisation and further development of terminology systems. Examples of areas of interaction are the Gene Ontology, the Foundational Model of Anatomy, and SNOMED CT. Part of the network objectives is also an active interaction with standardisation bodies such as CEN TC251, ISO and IMIA. The research carried out under the auspices of this NoE will also address the need for approaches in Europe which will bridge language barriers and facilitate access for non-English native persons to the large scientific corpus of texts written in English.

Examples of external relations during the first two years of SemanticMining are:

- :: standardisation activities performed in e.g. CEN TC251 and HL7,
- :: developers of the Foundational Model of Anatomy (FMA),
- :: developers of the Gene Ontology (GO),
- :: developers of SNOMED CT,
- :: developers of CNPU and LOINC in the area of laboratory medicine.

## Integration and dissemination

During 2005, the mobility program has taken off. 13 PhD students have been awarded a mobility grants, and have spent from one week up to three months at another partner site. Continuation of the mobility program, also in cooperation with the NoEs in the eHealth area is planned for 2006.

A significant role of this NoE is to provide educational material based on both the workshops and the research. It is hoped that this educational material available through the web site ([www.semanticmining.org](http://www.semanticmining.org)) will be useful, not only to students associated with SemanticMining, but also for the general public, and other interested parties. A series of links to more specific information and websites will be provided, where possible, for various topics of relevance for SemanticMining. The traditional way of sharing research results is through scientific publication. During the last year, there has been a significant increase in co-authored research papers within the network. As can be seen in the next section (Joint research) the network contributes to the harmonisation and further development of terminology systems via close interaction with standardisation bodies.

A Summer School was successfully held once again at lake Balaton in Hungary. The program included workshops on the semantic web, text mining, terminologies in laboratory medicine, and the so-called boundary problem between information models and terminology systems. In the first summer school 2004, the activities included a PhD student poster session. In the second summer school, the PhD students were even more in focus during a successful two-day doctoral consortium. The two day event included 21 student presentations with feedback from senior researchers, group discussions on methods, and advice of students for each other.

Based on the very positive experiences of these summer events, a joint summer conference is now being planned together with the two other NoEs in the eHealth area, INFOBIOMED and BIOPAT-TERN.

## INTERVIEW

### Bio-sketch Martin Dugas



**1967;** Born in Straubing (Germany).

**1987;** Medical school at Technical University of Munich, Germany school of computer science (simultaneously).

**1992;** Medical studies in Zurich (Switzerland), London (UK) and Montpellier (France).

**1994;** M.D. Technical University of Munich.

**1996;** M.S. Computer science; Assistant professor, Department of Medical informatics, University of Munich.

**2000;** Board certificate "Medical Informatics" (GMDS); IT consultant for Munich-Harvard-Alliance.

**2001;** Expert evaluator eHealth appointed by European Commission; Member of EU expert group "Prospective analysis of the relationships and synergy between medical informatics and bioinformatics"; Course director "Clinical data analysis" at UMIT (University for Health Informatics and Technology Tyrol); Board certificate "Biometrics"; Textbook Medical Informatics & Bioinformatics.

**2004;** Siemens Medical Solutions Health Services, Malvern, USA: Product manager research & development.

**2005;** Professor of Medical Informatics, University of Münster, Germany.

### Questions:

#### *Brief description of the place where you work and the main research topics in which you are involved.*

I'm working at the Department of Medical Informatics and Biomathematics which belongs to the University of Münster, Germany. In my team we are managing the hospital information system of the University Hospital of Münster, a tertiary care center with approximately 1500 beds.

My main research topics are analysis of complex medical databases, in particular integration of clinical and genetic data with microarrays. I have been working in several projects in oncology and cardiology, particularly dedicated to microarray-based diagnostics.

#### *What do you understand by Biomedical Informatics?*

A new interdisciplinary field emerging from the intersection of Medical Informatics and Bioinformatics.

[Maojo V, Martin-Sanchez F, Billhardt H, Iakovidis I, Kulikowski C: Establishing an Agenda for Biomedical Informatics. *Methods Inf Med* 2003; 42: 121-125]

Biomedical Informatics applies methods primarily from computer science and biostatistics to medicine and biology, especially with respect to molecular medicine. A typical application field is integration of clinical and genetic data.

#### *Which new relevant developments have you seen lately in the field of Biomedical Informatics, particularly in Europe?*

I'm aware of some biomedical informatics projects in Europe which are moving from academic research towards industrial application, for instance regarding microarray-based diagnostics in medicine. This field is very competitive at the international level and I hope Europe will play an important role.

#### *In your opinion, which are the new trends and future challenges for Biomedical Informatics?*

People with heterogeneous backgrounds like computer science, engineering, medicine, biology or biochemistry are entering this field. This is beneficial to reach critical mass, but on the other hand it is challenging because there is not yet a common academic basis for biomedical informatics in Europe.

#### *Could you briefly compare the current situation Biomedical Informatics in Europe vs USA; which are the strengths and weaknesses of each?*

In USA there are several training programs in biomedical informatics (e.g. Columbia, Stanford) which are driven by experts in the field of medicine. In Europe there is a stronger focus on basic research involving biology and computer science. More involvement of clinical experts would be beneficial in Europe to develop projects with real medical relevance.

## INFOBIOMED PILOT

### Genomics and chronic inflammation: Periodontitis

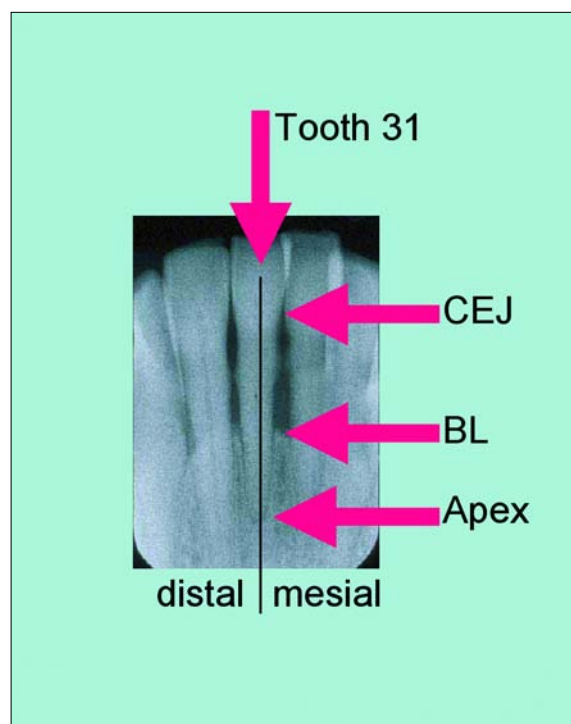
Common diseases, such as cardiovascular disease, cancer, metabolic imbalances, and chronic inflammatory illnesses impose a major drain on society, both in terms of financial and human costs. They are thus priority targets for current biomedical research. Chronic inflammation is of particular interest, since beyond obvious inflammatory states, it may well be a contributory factor in a range of common disorders.

Chronic inflammatory diseases have a complex pathogenesis and have a multifactorial etiology, involving complex interactions between multiple genetic loci, infectious agents and environmental (behavioral) factors such as diet, smoking habits and physical exercise. The general paradigm is that certain individuals are genetically more susceptible than others to the infectious and/or environmental risk factors. This innate variability then also explains why patients vary in age of onset, severity and response to medical treatment.

Periodontitis is a chronic inflammatory disease of the supporting tissues of the teeth. If left untreated, teeth may show exposed root surfaces, in conjunction with red, swollen gums that easily bleed. Dental radiographs reveal periodontal (alveolar) bone loss around the teeth that is due to the inflammation process (Figure 1); teeth will become mobile and migrate, and will eventually exfoliate. Patients with periodontitis experience problems with chewing due to tooth mobility and loss of teeth; they have bad breath and suffer from important subjective and objective esthetic problems. Dental practitioners provide labor-intensive diagnostic and treatment sessions to periodontitis patients, including periodontal surgery.

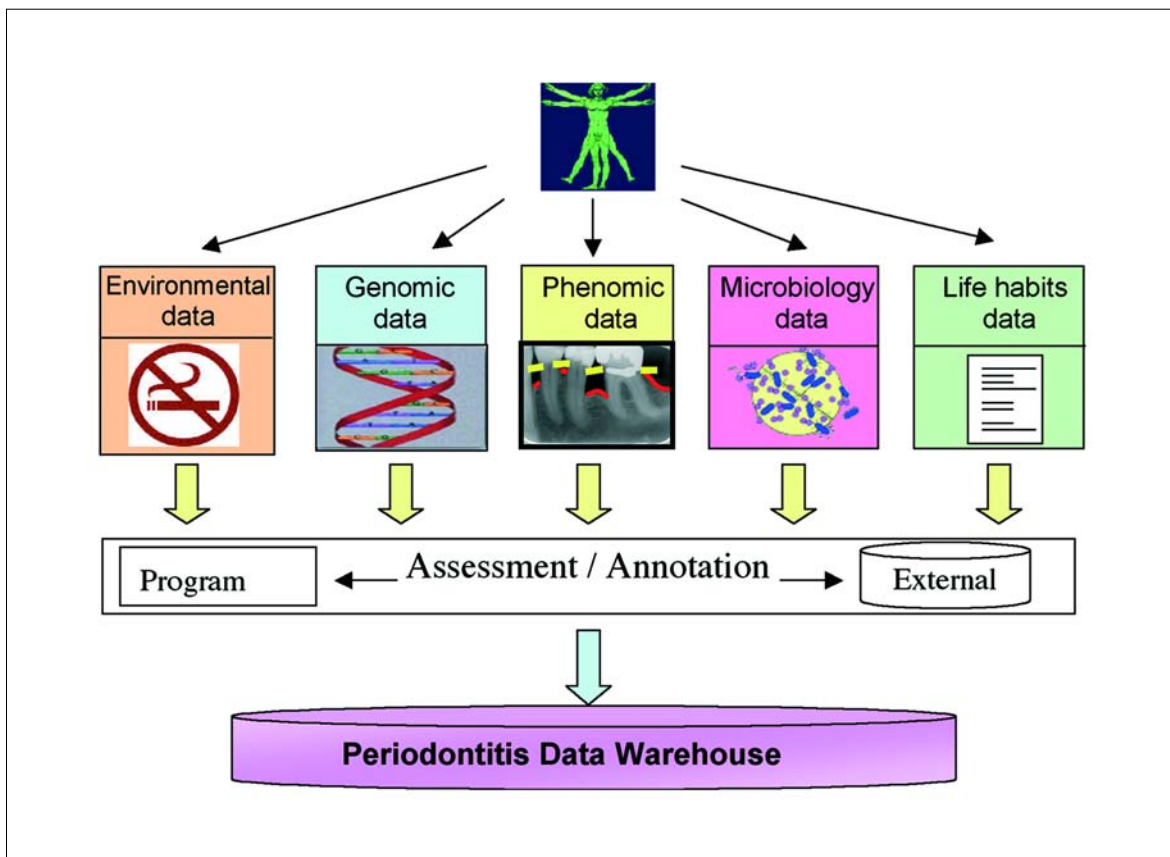
Periodontitis has a relative high prevalence in the population. It has been estimated that about 10% of the total adult population and about 30% of individuals over the age of 50 years, suffer from severe periodontitis. Recent data suggest that periodontitis is associated with increased risk for

cardiovascular diseases, possibly through the elevation of the acute phase reactant C-reactive protein (CRP) or other systemic markers of inflammation. The systemic reactions to periodontitis may come about through systemic dissemination of oral bacteria. There are strong indications that the inflamed and ulcerated pocket epithelium forms an easy port of entry for oral microorganisms.



**Figure 1** Example of a dental radiograph of tooth 31 (lower left central incisor) and surrounding hard tissue structures showing important landmarks: the cemento-enamel junction (CEJ), the apex and the periodontal bone level (BL) on the mesial site of the tooth. Periodontal bone levels in the healthy situation are located 1-2 mm apical from the CEJ; in this example there is about 55% periodontal bone lost on the mesial surface.

Research objectives of ACTA within INFOBIOMED network

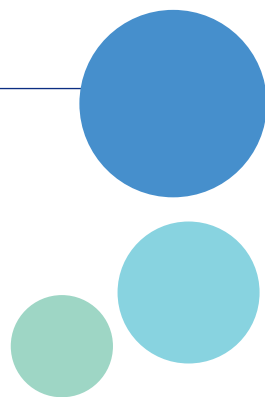


Periodontitis is an excellent model for biological and medical informatics (BMI), because of its multifactorial etiology, *i.e.* infectious component, genetics, environmental factors. Data can be easily obtained due to the relative high disease prevalence, plus the fact that periodontitis is not a life threatening disease and that no invasive procedures are necessary to obtain biological samples. Furthermore patients are often very willing to collaborate. Diseased and healthy tissues, genomic DNA and access to the history of infections and other relevant data through the patient records, are all accessible. Integration of various data sources will give new insights in the pathophysiology, will enable us to make new disease classifica-

tion schemes and ultimately will result in risk profiling and screening. To meet these objectives, the disciplines periodontology, MI and BI will to be integrated.

The goals of BMI within the INFOBIOMED consortium for the improvement of human health are, among others, the integration of genetic and clinical information with preventive, diagnostic and therapeutic purposes. Within the consortium of INFOBIOMED, the complex disease periodontitis will serve as a case study and model of integration of BI and MI. Figure 2 summarizes the factors involved in this disease and their interrelations providing a structure for the complexity of its study.

BIOMEDICAL INFORMATICS  
TO SUPPORT INDIVIDUALISED HEALTHCARE





BIOMEDICAL INFORMATICS TO SUPP

<http://www.infobiomed.org>

