ISSN 0974-4371

AGRIBIOINFORMATICS

http;//www.iari.res.in/divisions/usi/agribio.php

Vol. 2, No 1 & 2 **A** Quarterly Newsletter for disseminating ICT in Agriculture and Rural development **Jan-June 2009**

From Director's Desk…….

This is an era of "omics". Ever since the successful sequencing of human genome, great advances in biotechnology have taken place. Genomics, Proteomics and Metabolomics contribute in understanding and unraveling the secrets and mystery of life. Computer science, statistics, and biology have given birth to bioinformatics to provide support related to storage, retrieval and analysis of these experimental data in terms of structure, sequence and function. Bioinformatics centers on understanding the molecular world and therefore requires knowledge in the fields of biochemistry, molecular biology, molecular evolution, thermodynamics, biophysics, molecular engineering, and statistical mechanics. It is about converting biological observations to a model that computer will understand. It is therefore a fundamental discipline which makes predictive models of biological systems. In recent years bioinformatics increased the diversity of its data such as DNA and protein sequences, protein structures, microarrays, pathways, bio-images and time series. Grid and web services based approaches have been developed to face new challenges. There are some top challenges which will determine for a long time the future path of bioinformatics such as molecular mechanisms of: transcription, alternative RNA splicing, non coding RNA regulation of gene expression, prediction of protein structure from, signal transduction pathways, organisms development, disease initiation and progression, complete computer representation of cells and organisms and speciation and genome evolution. The newsletter *Agribioinformatics* is aimed at dissemination of bioinformatics related research news in all disciplines of agriculture.

(H. S. Gupta)

CONTENTS

Editorial.....
Bioinformatics.....
Crop Informatics.....
Chemoinformatics.....
Fish Informatics.....
Pest Bioinformatics.....
News Roundup.....
Online Bioinformatics Education.....

EDITORIAL BOARD

EDITORIAL

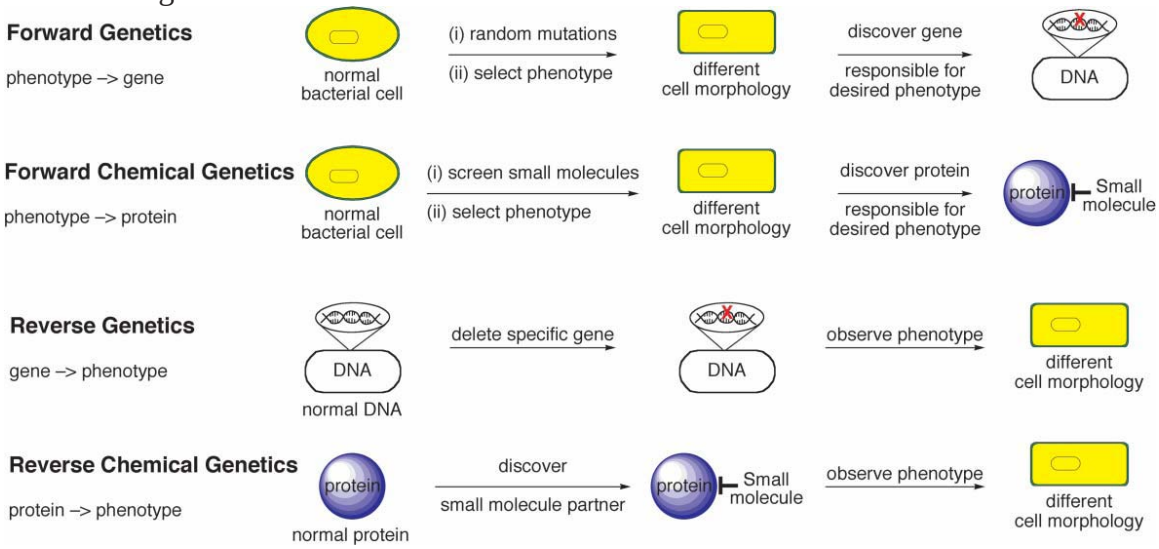
CHEMICAL GENETICS: AN INTERDISCIPLINE OF ENORMOUS POTENTIAL

(C. Devakumar, ICAR)

One is familiar with Chemical Kinetics and Chemical Physics. What on earth is Chemical Genetics? Schreiber (1998) described chemical genetics as a fruit of passion for synthetic organic chemistry. The term "chemical genetics," was coined in the 1990's, when combinatorial chemistry was developed as a fast method to synthesize large compound libraries. There have been many examples of neat chemical genetic solutions to biological problems in the past, long before this term was coined. One of the first examples for successful target identification in biological research was the discovery of tubulin as the target protein for colchicine in the 1960's.

Chemical genetic research consists of three major steps: design and synthesis of the library of compounds to be screened, screening of the compound library in the system of choice and target identification/validation including specificity tests. Biologically active small molecules better known as drugs bind with one or other protein which is a receptor or target protein for this "drug". Organic chemists in pursuit of new or more potent and safer drugs have been using target protein as a test substrate. The drug discovery depends on either *in vitro* enzyme assay or cell line or *in vivo* whole organism.

Just as genetics is divided into forward genetics (involving random mutations followed by phenotypic screening and gene identification) and reverse genetics (involving mutation of a specific gene and phenotype characterisation), chemical genetics is also classified accordingly. Using the same analogy, forward chemical genetics involves the use of small molecules (in place of the 'mutations') to screen for the desired phenotypic effect *on the biological system* under investigation. On the other hand, reverse chemical genetics involves the use of small molecules *against a protein* (gene product) of interest. Genetics in the 'forward' direction is from *phenotype to gene*; in the 'reverse' direction it is from *gene to phenotype*. Similarly, chemical genetics in the 'forward' direction is from *phenotype to protein*; in 'reverse' it is from *protein to phenotype*. In genetics, genomics is involved whereas in chemical genetics proteomics is. The study of small molecular interaction with genome is in the domain of chemical genomics.



A comparison of genetics with chemical genetics

In both forward and reverse chemical genetics, the identification of a selective small molecule followed by detailed biological investigation is required. Biochemical and molecular genetic tools necessary for target (receptor) identification and characterization are involved in chemical genetics. With the advent of molecular biology and advances in protein expression and purification techniques, the role of this new interdiscipline is becoming widespread in the elucidation of gene expression, biochemical pathways etc.

The most visible advantage of using small molecules is that they act on a fast time scale and that their effect can be easily controlled through titration. Other methods, like mutagenesis or RNAi interference, cause either permanent effects or have a temporal resolution at the scale of days.

Moreover, the effect of small molecules is usually equally strong in all cells treated, while methods like antibody microinjection and RNAi are highly variable. These properties of small molecules become relevant in the study of dynamic processes, because high time resolution and synchrony of the effect in all cells is essential here. The most common application of chemical genetics in this context is synchronization and release of cells at a certain cell cycle stage with molecules like aphidicolin, nocodazole, taxol, or thymidine. This cannot be achieved with any of the other mentioned approaches. Another gap filled by chemical genetics is the possibility to inhibit protein activity without physically removing the protein from the system studied.

Ensuring drug specificity is one of the most difficult tasks in chemical genetics beside target identification. The ultimate goal of chemical genetics is to provide biologists with a toolbox of specific small molecules for every protein in a cell. This requires the discovery of sophisticated methods to synthesize complex molecules, to perform elaborate small molecule screens and finally to demonstrate specificity of the identified bioactive compounds. Chemical genetics is an interdisciplinary approach involving cell biology, structural biology, biochemistry, and chemistry to identify or develop new small molecules for the study of biological questions. It is most effective when used in combination with the more traditional tools of the field. It is a very versatile tool that effectively complements the toolbox of biologists but success of its application heavily depends on the right choice of the biological question to be answered. Our institute has both laboratories and scientists to contribute in this field but what is needed is a strong programme with efficient collaborations for a project to be successful. We need to set up core facilities for combinatorial chemistry and high throughput screening platform. Readers may refer to the following articles for a comprehensive knowledge.

1. Smukste I and Stockwell B R (2005). Advances in Chemical Genetics *Annu. Rev. Genomics Hum. Genet*. **6**:86-261.
2. Blackwell H E and Zhao Y (2003). Chemical Genetic Approaches to Plant Biology. *Developmental Cell* **5**:11-19.
3. Florian S, Hümmer H, Catarinella M,and Mayer T U (2007). Chemical genetics: reshaping biology through chemistry. *HFSP Journal* **1**:104-114
4. Specht M K and Shokat M K (2002). The emerging power of chemical genetics. *Current Opinion in Cell Biology* **14**:155-159
5. Yeh J R and Crews C M (2003) Chemical Genetics: Adding to the Developmental Biology Toolbox. *Developmental Cell*, 5: 11–19.

BIOINFORMATICS

(C. Devakumar and H. Chandrasekaran, IARI)

Bioinformatics for Next Generation Sequencing

Next-generation sequencing technologies the Human Genome Project have begun to revolutionize genomics and their effects are becoming increasingly widespread. The1000genomesproject (http://www.1000genomes.org/) will create a new map of genetic variation for human genome going far beyond the detail captured in the HapMap (Haplotype Map of Human Genome). The growing number of robust applications and the steadily falling cost for generating sequence-based data suggest that these next-generation technologies will continue to rapidly open new applications in the biological sciences and generate new opportunities for software and algorithm development. Developing a sound data storage and management solution and creating informatics tools to effectively analyze the copious volume of data are critical to successful application of the technology. Recently, a large number of new software applications and algorithms have been developed to deal with this new data. Several papers, which described methods, to take the short sequences produced by the Illumina Genome Analyzer and Applied Biosystems SOLiD machines and align them to a reference genome have appeared in recent times. This is a crucial and basic requirement for many applications and a variety of techniques have been applied to make the tools sufficiently fast to deal with millions of sequences. Now there are many of these tools[1] available like PatMaN, SeqMap, Slider, SOAP, and SSAKE etc.

Note1 PatMaN: rapid alignment of short sequences to large databases; SeqMap: mapping massive amount of oligonucleotides to the genome; Slider - Maximum use of probability information for alignment of short sequence reads and SNP detection; SOAP: short oligonucleotide alignment program; SSAKE Assembling millions of short DNA sequences.

Bateman A and Quackenbush J (2009) Bioinformatics for next generation sequencing. *Bioinformatics* **25**: 429.

Genome Sequencing Becoming Inexpensive

California-based Biotechnology Company is stated to

to slash the cost of human genome sequencing. It sequenced three human genomes for about $4400 each, at least in the cost of reagents. Such cheap sequencing could hasten studies designed to pinpoint genes underlying complex diseases. The cost of sequencing a human genome at an estimated $300 million in 2003 came down to $60,000 last year. This success is due to both speed and low use of reagents. The sequencing technology works by first chopping the genome into single-stranded DNA fragments and then combining them with snippets of known synthetic DNA so that they form small circles of about 400 bases. Each DNA circle is copied over and over so that the DNA winds itself into a 200-nanometer ball. These "DNA nanoballs" are then washed over a slide patterned with sticky spots designed to hold a single ball at each spot. The company uses a modified version of a conventional sequencing technology that attaches a series of 10-base-long fluorescently labeled DNA snippets to the DNA in the balls, which enables them to work out the sequence of the original DNA. These provide them with billions of DNA sequences, each of which are about 35 bases long. A computer program then reads overlaps between these sequences to reconstruct the overall genome. The technique's advantage is that the nanoballs assemble themselves into an ultra-dense array, thereby holding down reagent costs and speeding the analysis of the fluorescent signals. Using the new technology, it is estimated that a complete human genome can be sequenced in about a day. The Company Officials in collaboration with the Institute for the Systems Biology in Seattle, Washington, to sequence genomes of 100 patients with Huntington's disease, the largest human genome disease association study conducted to date. However, rapid sequencing is prone to errors at the rate of about 10 in every million bases.

http://sciencenow.sciencemag.org/cgi/content/full/2009/1105/3

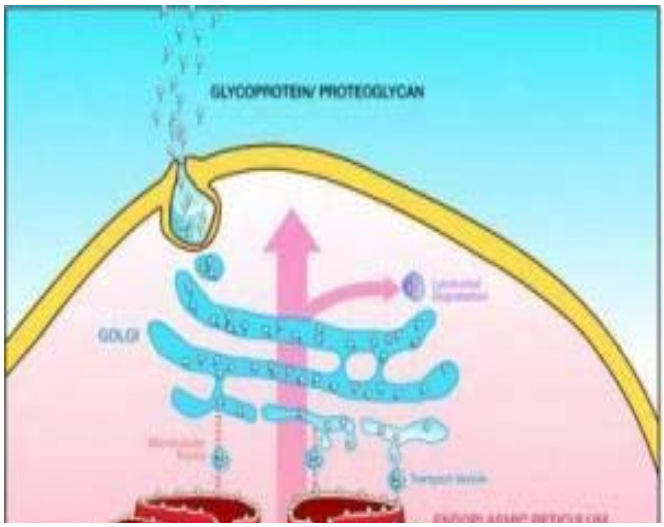
Antibiotic Marker Genes Unlikely to Harm Human Health and the Environment

The European Food Safety Authority provided an overview of the use of selectable antibiotic resistance marker genes in genetically modified plants. EFSA's GMO and Biohazards Panels concluded that based on currently available information, the commonly used antibiotic marker genes nptII and aadA are unlikely to have adverse effects on human health and the environment. In their joint opinion, the panels noted that the transfer of antibiotic resistant genes from GM plants to bacteria have not been shown to occur either in natural conditions or in the laboratory.

According to the report, the key barrier to stable uptake of antibiotic resistance marker genes from GM plants to bacteria is the lack of DNA sequence identity between plants and bacteria. The Panels, however, underlined the limitations in estimating exposure levels and the inability to assign gene transfer to a defined source. According to them, it is not possible to find out precisely from which organism a marker gene present in another organism may have originated. The GMO and Biohazards Panels also considered the clinical importance of the antibiotics to which the marker genes confer resistance. The gene nptII confers resistance to kanamycin, which is used by doctors as a second-line antibiotic for the treatment of infections with multiple drug-resistant tuberculosis (MTB). According to the GMO and Biohazards Panels nptII has not been implicated in resistance to kanamycin in the treatment of MTB.

http://www.newstin.co.in/tag/in/128197515

Artificial Golgi' Apparatus Developed



(Credit: The American Chemical Society)

Structure inside cells helps process and package hormones, enzymes and other substances that allow the body to function normally. The lab-on-a-chip device may lead to a faster and safer method for producing heparin, the popular anticoagulant.

The Golgi organelle is named after Camillo Golgi, the Italian scientist and Nobel Prize winner who discovered the structure in 1898. It is composed of a network of sacs, stacked like a deck of playing cards, located inside cells. Golgi bodies are one of the most poorly understood organelles in the human body.

The researchers describe development of a prototype lab-on-a-chip device that closely mimics the natural Golgi apparatus. They showed in lab tests that the device could quickly and efficiently produce heparin. It did so in an assembly-line fashion using a combination of enzymes, sugars and other raw materials. In future, an "artificial Golgi" could lead to a faster and safer method of heparin production.

Martin G, Gupta M, Xu Y, Akella S, Liu J, Dordickand S and Linhardt R J *(*2009). Toward an Artificial Golgi: Redesigning the Biological Activities of Heparan Sulfate on a Digital Microfluidic Chip. *Journal of the American Chemical Society* **131**:11041–11048.

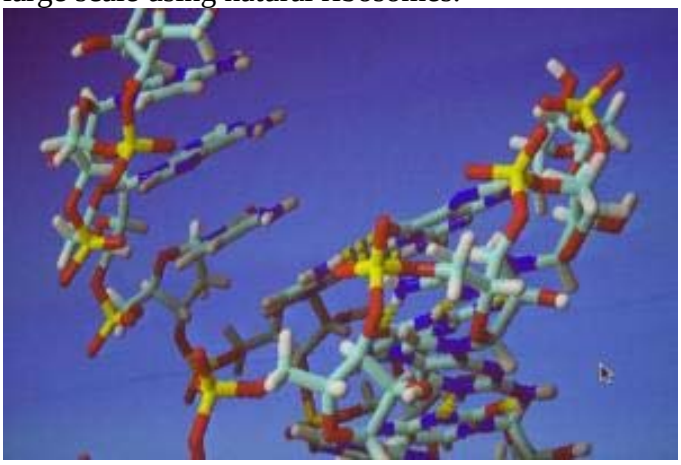
Combinatorial Libraries toward Synthetic Biology

The construction of synthetic gene networks has led to develop a technique that couples libraries of diversified components with computer modeling to guide predictable gene network construction without the back and forth tweaking. The research team, led by James J. Collins, Professor of Biomedical Engineering Boston University, USA focused on ways to speed up the construction process by assembling a library of 20 versions of two gene promotors and a simple synthesis technique to create component libraries for synthetic biology. Each version covered a wide range gene expression. With the activity levels calculated from the component libraries, a computer model was designed to build a basic gene circuit to predict how fluorescent protein expression varied with levels of promoter-inhibiting chemicals. They then built a representative genetic timer using a promoter from each of their libraries and, over time, tracked its behavior. Based on information from one network, they were able to calibrate their model and achieve accurate predictions from all the other possible network combinations. These timers, were effectively genetic toggle switches. The researchers concluded that their method using combinatorial libraries to engineer genetic circuits moves the "tweaking" from the back-end of gene network engineering to the front-end. This approach will help accelerate synthetic biology by yielding many more components for the community.

Ellis T, Wang X and Collins J J (2009). Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature Biotechnology* **27:** 465–471.

Synthetic biology: Ribosomes Cell Protein Machinery Created

Ribosomes are bodies inside of each cell that take the instructions from DNA and use them to create the proteins encoded by specific genes. Proteins are critical to forming the body's structure, including muscles, bones and tendons, and are also critical in its daily functioning, through enzymes, for example, which control metabolism. George Church, a genetics professor at Harvard Medical School and member of Harvard's Origins of Life Initiative, reported the creation of billions of synthetic ribosomes that readily create a long, complex protein called firefly luciferase. Using the bacteria *Escherichia.coli*, Church extracted the bacteria's natural ribosomes, broke them down into their constituent parts, removed the key ribosomal RNA and then synthesized the ribosomal RNA anew from molecules. Industry today manufactures proteins on a large scale using natural ribosomes.



(Credit: Kris Snibbe/Harvard News Office)

One possible use would be to create mirror-image proteins that would be less susceptible to breakdown by enzymes, making them longer-lived. The advance breaks a 40-year period with little progress in artificial ribosome creation. The last significant work in this area was done in 1968, when researchers assembled an artificial ribosome, but in an unusual chemical environment rather than an environment in which protein synthesis normally occurs. The ultimate goal is to create an artificial genome of 151 genes that they believe are the minimum to create a functioning, self-replicating cell.

Alvin Powell, Harvard University (2009, March 9). Toward Synthetic Life: Scientists Create Ribosomes -Cell Protein Machinery. *ScienceDaily*. Retrieved

Synthetic Biology: The Next Biotech Revolution is Brewing

Synthetic biology promises major advances in areas such as biofuels, specialty chemicals, and agriculture and drug products. Rodemeyer examines the benefits and drawbacks of using the existing U.S. regulatory framework for biotechnology to cover the new products and processes enabled by synthetic biology.

According to Rodemeyer, initial synthetic biology products will be relatively simple modifications of current technology and can be addressed by existing biotechnology regulations with only modest revisions. However, as the technology develops, regulatory agencies such as the Environmental Protection Agency and Food and Drug Administration will face challenges in assessing potential risks and the adequacy of controls, especially if complex synthetic microorganisms are released into the environment. Today's risk assessment practices and laws like the Toxic Substances Control Act and Federal Food, Drug, and Cosmetic Act, simply are not designed to handle 21st century technological advances.

DNA of Uncultured Organisms Sequenced Using Single-cell Approach

Scientists from the U.S.A. Department of Energy (DOE) Joint Genome Institute (JGI) and the Bigelow Laboratory for Ocean Sciences have assembled high quality, contamination-free draft genomes of uncultured biodegrading microorganisms using a novel single cell genome sequencing approach. Their approach published offers researchers a new method to access and decipher the information embedded in genomes of interest with only minute quantities of DNA. A technique called fluorescence activated cell sorting was used to pick out individual bacterial cells directly from the environmental sample. The single cells were then lysed and a process called multiple displacement amplification was applied to make millions of copies of the bacterial genomes for sequencing. The resulting flavobacterial genome sequences are approximately 80 to 90 percent complete, a level sufficient, to prove the utility of the technique. The single cell sequencing approach can be applied to organisms from a number of environments, including those microbial communities inhabiting extreme environments, such as hot pools, contaminated soil and those constituting the human microbiome. The technique bypasses the need for culturing before sequencing, he said, because only one cell is needed to decode a genome.

Gary Xie, Cliff Han, Hajnalka Kiss, Jimmy Saw, Pavel Senin, Chi Yang, Alex Copeland,Jan-Fang Cheng, DOE/Joint Genome Institute (2009, April 28). DNA Of Uncultured Organisms Sequenced Using Novel Single-cell Approach. *ScienceDaily*.

Single-Molecule Nano-Vehicles Synthesized: *'Fantastic Voyage'* Not away

Ayusman Sen, Heads Penn State's Department of Chemistry, and Thomas E. Mallouk, Director of the Center for Nanoscale Science at Penn State, have investigated technologies that could realize these remarkable machines whose uses might include delivering medicine to specific tissue, accomplishing surgeries or communicating with the outside world from inside the human body. Though researchers consistently have improved ways to build nano-machines, the stumbling block has been finding a way to power them. Shrinking energy producers--internal combustion engines, electric motors or jet engines--below millimeter dimensions is not an easy task, but researchers may be closer to a fantastic solution. In the 1966 movie *Fantastic Voyage,* scientists shrink a submarine to microscopic size and inject it into the blood stream of a brilliant scientist, who has a blood clot forming in his brain. The nano-sized surgeons then set out to remove the blood clot. All of this, they say, makes the world seen in *Fantastic Voyage* not so far-fetched.

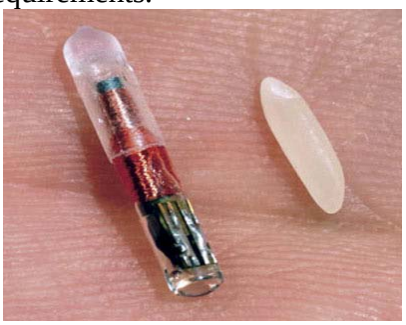
http://www.nsf.gov/cgi-bin/good-bye?http://www.sciam.com/article.cfm?id=how-to-build-nanotech-motors

CROP BIOINFORMATICS

(H. Chandrasekaran , C. Devakumar and Minakshi)

Microchip to Measure Real-Time Water Stress

Cornell University's nanofabrication laboratory is credited with a measuring real-time water stress in living plants. The device will be useful in monitoring irrigation requirements.



(Courtesy photobucket.com)

The device is composed of a slab of hydrogel with nanometer-scale pores, which acts as a synthetic tree that mimics the flow of water inside plants. They have also started developing a multi-use sensor that redirects water flow inside the plant through a shunt. In this case, the sensor could measure the flow of water and mineral nutrients through the plant, in addition to water stress. This multi-use sensor could be implanted throughout all trees in a forest ecosystem to measure water use and nutrient flow on a large scale with unprecedented accuracy.

http://www.news.cornell.edu/stories/July09/plantWaterStress.html

Jatropha Genome Deciphered

California-based Synthetic Genomics Inc. (SGI) and

Asiatic Centre for Genome Technology (ACGT) based in Kuala Lumpur, Malaysia announced that the completion of the first draft of the genome of jatropha (*Jatropha curcas*), an important biofuel\ crop. Researchers used both the traditional Sanger sequencing and next generation sequencing to crack the crop's genome. The genome is around 400 million basepairs long, similar to the size of the rice genome. The teams are now working on annotating the genome to pinpoint particular genes of interest and to discover genetic variations for use in marker assisted breeding. The teams are combining traditional breeding tools with modern plant molecular biology tools, to improve plant yield, oil quality, fertilizer requirements and to enhance stress and disease tolerance.
http://www.syntheticgenomics.com/press/2009-05-20.htm

Transcriptome Analysis of Wheat

Wheat is generally a long-day plant that will go through phase transition from vegetative to floral growth, as days are lengthening in spring and early summer. Wheat cultivars can be classified as either winter or spring varieties depending on whether they require to be exposed to an extended period of cold in order to become competent to flower. Using a growth regime to mimic the conditions that occur during a typical winter in Britain, and a microarray approach to determine changes in gene expression over time the genes of the major pathways involved in floral transition, wheat orthologues and functional equivalents of genes involved in the phase transition in *Arabidopsis* and MADS-box genes that could be identified as such on the Affymetrix genechip wheat genome array were surveyed. It was observed that novel responses of several genes are of major importance in vernalisation-induced phase transition. Several MADS-box genes that might play an important role in the onset of flowering were identified. In addition, responses in genes of the Gibberellin pathway having some role in phase transition, which is more complex, were reported. There is evidence that day-length has an influence on genes that were once thought to respond exclusively to an extended period of cold.

Winfield O M, Lu C, Wilson I D, Coghill J A, and Edwards K J (2009) Cold- and light-induced changes in the transcriptome of wheat leading to phase transition from vegetative to reproductive growth. *BMC Plant Biology* **147**:9-55.

CHEMOINFORMATICS

(C. Devakumar and Minakshi)

ChemSpider-database of Chemical Structures and Property Predictions

Chemspider a chemical database. First launched in March 2007 has been acquired by the Royal Society of Chemistry in May, 2009.The Database is a free access service providing a structure centric community for chemists.

(Courtesy chemspider.com)

Providing access to millions of chemical structures and integration to a multitude of other online services, ChemSpider is the richest single source of structure-based chemistry information. It contains more than 20 million unique molecules from different sources. The database can be updated with user contributions including chemical structure and spectra deposition and user curation. The ChemSpider database has been used in combination with text mining as the basis of chemistry document markup. ChemMantis' (the Chemistry Markup And Nomenclature Transformation Integrated System) uses algorithms to identify and extract chemical names from documents and web pages and converts the chemical names to chemical structures using name-to-structure conversion algorithms and dictionary look-ups in the database. The result is an integrated system between chemistry documents and information look-up via ChemSpider into over 150 data sources. The ChemSpider services are offered free of charge. Search hits include both free information and pointers into commercial databases that may require a subscription for access. A number of services are made available online. These include the conversion of chemical names to chemical structures, the generation of SMILES and InChI strings as well as the prediction of many physicochemical parameters and integration to a web service allowing NMR prediction.
http://www.rsc.org/AboutUs/News/PressReleases/2009/ChemSpider.asp

ChemMine: A Compound Mining Database for Chemical Genomics

To facilitate the incorporation of chemical genomics-based approaches in the discovery process of novel protein functions and gene networks, the ChemMine database (http://bioinfo.ucr.edu/projects/PlantChemBase/search.php) is developed. The first release of this public service provides access to an integrated suite of analysis and information retrieval tools for compound searching, structure-based clustering, descriptor generation (chemical properties), and retrieval of published bioactivity and target protein information. ChemMine centralizes compound structure and activity information from a growing number of public providers and vendors of chemical screening libraries. The incorporation of commercially available compounds provides access to their purchase information. This knowledge can be critical for follow-up studies and assembly of focused libraries in secondary screens when the resources for resynthesis of novel chemicals in larger quantities are limited or do not exist at all. It is expected that the current set of commercial compound collections (over 1 million) will quickly grow. In addition to commercial compounds, most collections from public initiatives are included in ChemMine. These highly annotated compound sets maximize access to bioactivity information, known target proteins, literature, and other useful annotation information, enabling the user to correlate screening results with available biological knowledge. Additional information will be included as it becomes available. Searches for analogs of metabolic compounds are available through the incorporation of the KEGG ligand database. Information about bioactive chemicals and their functional characterization is provided through the data sets from ChEBI, ChemBank, NCI, PubChem, and other providers. The annotations from ChEBI illustrate the growing utility of these services. This initiative was started to provide systematic target associations of small compounds that interfere with processes of living organisms. ChemMine users can retrieve the target protein sequences, 3D structures, and literature for annotated drugs or metabolic molecules that are available or hyperlinked in the UniProt database. Similar drug-to-target associations are available in the data sets from ChemBank and PubChem. The ChemMine project has a strong focus on online services which allow users to utilize most of ChemMine's analysis tools for external compound sets without being restricted to the compound coverage in the database. Since downstream analyses of compounds and their target proteins require the usage of various molecular modeling and computational chemistry programs, ChemMine supports interconversions of the most common structure formats (SDF, SMILES, PDB, etc.) for file exchange with other tools. The libraries from Open Babel (http://openbabel.sourceforge.net) are used for these reformatting steps.

Girke T, Chang Cheng L, and Raikhe N (2005) .ChemMine. A Compound Mining Database for Chemical Genomics1, *Plant Physiology* **138**: 573–577.

Self-assembled Nanowires Could Make Chips Smaller and Faster

Nanowires are attractive building blocks for both electronics and photonics applications. Compound semiconductor nanowires, such as gallium arsenide, are especially desirable because of their better transport properties and versatile heterojunctions. However, a number of challenges – including integration with existing microelectronics – must first be overcome. Researchers at the University of Illinois have found a new way to make transistors smaller and faster. The technique uses self-assembled, self-aligned, and defect-free nanowire channels made of gallium arsenide. In a paper to appear in the IEEE (Institute of Electrical and Electronics Engineers) journal Electron Device Letters, U. of I. electrical and computer engineering professor Xiuling Li and graduate research assistant Seth Fortuna describe the first metal-semiconductor field-effect transistor fabricated with a self-assembled, planar gallium-arsenide nanowire channel. The gallium-arsenide nanowire channel used in the researchers' demonstration transistor was grown by metal organic chemical vapor deposition using gold as a catalyst. The rest of the transistor was made with conventional microfabrication techniques. While the diameter of the transistor's nanowire channel was approximately 200 nanometers, nanowires with diameters as small as 5 nanometers can be made with the gold-catalyzed growth technique. The self-aligned orientation of the nanowires is determined by the crystal structure of the substrate and certain growth parameters. Earlier the nanowires and then transfer-print led on other substrates, including silicon, for heterogeneous integration. In the work presented in the current paper, the researchers grew the gallium-arsenide nanowire channel in place, instead of transferring it. In contrast to the common types of non-planar gallium arsenide nanowires, the researchers' planar nanowire was free from twin defects, which are rotational defects in the crystal structure that decrease the mobility of the charge carriers. Considering their planar, self-aligned and transferable nature, the nanowire channels could help create higher performance transistors for next-

for next-generation integrated circuit applications. The high quality planar nanowires can also be used in nano-injection lasers for use in optical communications.
http://www.sciencedaily.com/releases/2009/04/090420141206.htm

Adam, the first Robot Scientist

Researchers funded by the Biotechnology and Biological Sciences Research Council (BBSRC) have created a 'robot scientist' which they believe is the first machine to have independently discovered new scientific knowledge. The robot, called Adam, is a computer system that fully automates the scientific process Prof Ross King, who led the research at Aberystwyth University, hopes to have teams of human and robot scientists working together in laboratories. The scientists at Aberystwyth University and the University of Cambridge designed Adam to carry out each stage of the scientific process automatically without the need for further human intervention. The robot has discovered simple but new scientific knowledge about the genomics of the baker's yeast *Saccharomyces cerevisiae*. The researchers have used separate manual experiments to confirm that Adam's hypotheses were both novel and correct. Using artificial intelligence, Adam hypothesised that certain gene in baker's yeast code for specific enzymes, which catalyse biochemical reactions in yeast. The robot then devised experiments to test these predictions, ran the experiments using laboratory robotics, interpreted the results and repeated the cycle.

Ross D. King, Jem Rowland, Stephen G. Oliver, Michael Young, Wayne Aubrey, Emma Byrne, Maria Liakata, Magdalena Markham, Pinar Pir, Larisa N. Soldatova, Andrew Sparkes, Kenneth E. Whelan, and Amanda Clare. **The Automation of Science**. *Science*, 2009; 324 (5923): 85-89 DOI: 10.1126/science.1165620

New Protein Designed and Created

Using design and engineering principles learned from nature, a team of biochemists from the University of Pennsylvania School of Medicine have built – from scratch – a completely new type of protein. This protein can transport oxygen, akin to human neuroglobin, a molecule that carries oxygen in brain and peripheral nervous system. Some day this approach could be used to make artificial blood for use on the battlefield or by emergency-care professionals. To build their protein, the Penn team started with just three amino acids, which code for a helix-shaped column. From this, they assembled a four-column bundle with loop that resembles a simple candelabra. They added a heme, a chemical group that contains an iron atom, to bind oxygen molecules. They also added another amino acid called glutamate to add strain to the candelabra to help the columns open up to capture the oxygen. Since heme and oxygen degrade in water, the researchers also designed the exteriors of the columns to repel water to protect the oxygen payload inside. Initially, the team used a columns. When they are satisfied with the sequence, they use the bacterium E. coli as a biological host to make the full protein. The team used chemical tests to confirm that their protein did indeed capture oxygen. When the oxygen did bind to the iron heme molecule in the artificial protein, the solution in which the reaction took place changed color from dark red to scarlet, a color signature almost identical to natural neuroglobin.

Ronald L. Koder, J. L. Ross Anderson, Lee A. Solomon, Konda S. Reddy, Christopher C. Moser, P. Leslie Dutton. **Design and engineering of an O_2 transport protein**. *Nature*, 2009; 458 (7236): 305 DOI: 10.1038/nature07841

FISH BIOINFORMATICS

(C. Devakumar and H. Chandrasekharan)

Robot Fish Could Monitor Water Quality

Robotic fish could give researchers far more precise data on aquatic conditions and deepen our knowledge of critical water supplies and habitats. Tan and Elena Litchman based at MSU's Kellogg Biological Station on Gull Lake in Kalamazoo County, recently won funding from the National Science Foundation to integrate their research. One can be able to obtain at an unprecedented high spatial and temporal resolution. Such data are essential for researchers to have a more complete picture of what is happening under the surface as climate change and other outside forces disrupt the freshwater ecosystems. It will bring environmental monitoring to a whole new level. The robotic fish will carry sensors recording such things as temperature, dissolved oxygen, pollutants and harmful algae. Tan also is developing electronics so the devices can navigate and communicate in their watery environment. The project is very practical and we are designing the fish to be inexpensive so they can be used in various applications like sampling lakes, monitoring aquafarms and safeguarding water reservoirs. This project will greatly advance bio-robotic technology. The robotic fish might detect toxic algal blooms. Some of these algal blooms create poor conditions for fish and exude toxins that also endanger people. The team is building "fins" for robotic fish with electro-active polymers that use electricity to change shape. Similar to real muscle tissue, ion movements twist and bend the polymer when voltage is applied. The effect works in reverse, too -- slender "feelers" could signal maneuvering circuits in a sort of electro-active central nervous system. Infrared sensors also could be used for "eyes" to avoid obstacles. The robots will communicate wirelessly with a docking station after surfacing at programmed intervals and could similarly be linked to other robotic fish for coordinated maneuvers or signal relay. Global

positioning system technology and inertial measurement units will allow precise navigation.

Michigan State University (2009, November 2). Robot Fish Could Monitor Water Quality. *ScienceDaily*. Retrieved December 1, 2009, from http://www.sciencedaily.com /releases/2009/11/091102085825.htm

PEST BIOINFORMATICS

(*Usha Dev)*

Vat: A New Aphid Resistance Gene Sourced from India

Scientists at the French National Institute for Agricultural Research (INRA) have identified a novel gene that shows resistance against the dreaded melon or cotton aphid *Aphis gossypii*. The aphid has emerged as a major problem of farmers growing cucurbits, tomato and citrus trees. Aphids are also the most common vectors of plant viruses. The researchers identified the resistance gene, which they called Vat (virus aphid transmission resistance), in melon lines originating from India. The gene confers a double resistance phenotype: resistance to aphid infestation and resistance to viral transmission. The Vat locus has been successfully introduced to high-yielding commercial melon cultivars. The INRA researchers now plan to introduce the gene to cotton, cucumber and other plant species susceptible to the aphid. They are also looking for orthologues of the Vat gene in species other than melon.

http://www.international.inra.fr/

Insect Gene Expression Responds to Diet

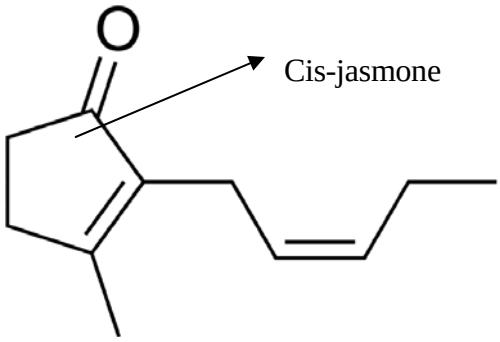
Trichoplusia ni

Dalial Freitak *et* ck Institute for Chemical Ecology studied the effects of dietary bacteria on general gene expression in the cabbage looper caterpillar *(Trichoplusia ni).* Larval feeding on a bacteria-rich diet led to substantial gene expression changes, potentially resulting in a reorganization of the insects' metabolism to maintain organismal homeostasis, not only in the larval but also in the adult stage. The authors explain that having a suite of genes capable of responding to dietary composition allows the cabbage looper to fine-tune its natural defenses. The insects could upregulate immunity related genes like Gloverin, HDD1 and hemolin in response to the presence of bacteria, but leave them switched off in the absence of pathogens. It was further shown that as well as tailoring , the insects were able to pass this information along to their offspring.

Freitak D, Heckel D G and Vogel H (2009). Bacterial feeding induces changes in immune-related gene expression and has trans-generational impacts in the cabbage looper *(Trichoplusia ni*). *Frontiers in Zoology* **6:**7.

A Single Olfactory Receptor Controls Host-specificity of Silkmoth to Mulberry

A jasmine-scented chemical emitted in small quantities by the leaves triggers a single, highly tuned olfactory receptor in the silkworms' antennae. *Bombyx mori*, the domesticated silkmoth, has reduced mouthparts, do not feed, cannot fly, and respond only to a sex pheromone for reproduction. The threshold amount of potent attractant cis-jasmone needed to attract silkworms appears to be significantly lower than the amount of any attractant to food reported for other insect larvae, such as fruit flies and mosquitoes.



Structure of cis-iasmone

Earlier studies had been hindered by a limited understanding of olfaction at the molecular and genomic level. They ultimately found 20 olfactory receptors that are active in the antennae of silkworm larvae. Of those, only one responds strongly to cis-jasmone. In another study, using the olfactory system of the lettuce aphid to investigate volatiles from plants avoided by this insect, (*Z*)-jasmone was found to be electrophysiologically active and also to be repellent in laboratory choice tests. In field studies, repellency from traps was demonstrated for the damson-hop aphid, and with cereal aphids numbers were reduced in plots of winter wheat treated with (*Z*)-jasmone. In contrast, attractant activity was

found in laboratory and wind tunnel tests for insects acting antagonistically to aphids, namely the seven-spot ladybird and an aphid parasitoid. When applied in the vapor phase to intact bean plants, (*Z*)-jasmone induced the production of volatile compounds, including the monoterpene (*E*)-b-ocimene, which affect plant defense, for example by stimulating the activity of parasitic insects. These plants were more attractive to the aphid parasitoid in the wind tunnel when tested 48 h after exposure to (*Z*)-jasmone had ceased. This possible signaling role of (*Z*)-jasmone is qualitatively different from that of the biosynthetically related methyl jasmonate and gives a long-lasting effect after removal of the stimulus. Differential display was used to compare mRNA populations in bean leaves exposed to the vapor of (*Z*)-jasmone and methyl jasmonate. One differentially displayed fragment was cloned and shown by Northern blotting to be up regulated in leaf tissue by (*Z*)-jasmone. This sequence was identified by homology as being derived from a gene encoding an a-tubulin isoform.

K. Tanaka, Y. Uda, Y. Ono, T. Nakagawa, M. Suwa, R. Yamaoka, K. Touhara (2009). Highly selective tuning of a silkworm olfactory receptor to a key mulberry leaf volatile.*Current Biology*, 19: 881-890.

Michael A Birkett, Colin A M Campbell, Keith Chamberlain, Emilio Guerrieri, Alastair J Hick, Janet L Martin, Michaela Matthes, Johnathan A.Napier, Jan Pettersson, John A Pickett, Guy M Poppy, Eleanor M. Pow, Barry J Pye, Lesley E. Smart, George H Wadhams, Lester J Wadhams, and Christine M Woodcock (2009). New roles for cis-jasmone as an insect semiochemical and in plant defense. PNAS 97: 9329–9334.

NEWS ROUNDUP

(H. Chandrasekharan and C. Devakumar)

NAAS Releases State of Indian Agriculture

India's National Academy of Agricultural Sciences (NAAS) has come out with its first annual series of publications on The State of Indian Agriculture. During the opening ceremony of the Foundation Day of NAAS, the Union Minister of Agriculture Shri Sharad Pawar presented the document on the state of different sectors within agriculture and key/critical areas that need urgent attention. It is divided into six chapters: Agriculture Sector, Natural Resources, Farm Inputs and Management, Agricultural Bio security, Policies and Institutions, and Agricultural Research Preparedness. The State of Indian Agriculture is a timely publication that contains analysis of the issues and recommendations which can guide policy makers, thus ensuring national and household nutritional

(Courtesy icar.org.in)

security," he said. The publication contains up-to-date information on the overall status and performance of Indian agriculture during the last 60 years.

Promising Biofuel Made From Commercial Yeasts Upgraded with a New Enzyme

Eckhard Boles, co-founder of the Swiss biofuel company Butalco GmbH and a professor at Goethe-University in Frankfurt, Germany, has discovered a new enzyme, which teaches yeast cells to ferment xylose into ethanol. Xylose is an unused waste sugar in the cellulosic ethanol production process. The researchers have recently filed a patent application for their process. In industrial fermentation processes, the yeast *Saccharomyces cerevisiae* is commonly used for ethanol production. Current bioethanol production technologies can use only the, like glucose, sucrose or starch. However, this technology is in competition with food and feed production. Eckhard Boles, co-founder of the Swiss biofuel company Butalco GmbH and a professor at Goethe-University in Frankfurt, Germany, has searched for ways of teaching the microorganisms to convert waste sugars, xylose and arabinose, into ethanol. Now, Boles and his colleagues have succeeded in genetically modifying industrial yeast strains for converting waste sugars xylose and arabinose into ethanol. Having already succeeded in transforming arabinose into ethanol by genetically modified yeast strains, Boles and his team have now found an efficient way to convert most of the plants energy into biofuel. Step by step they took 12 enzymes from different bacterial organisms and inserted the enzymes into yeast cells. Finally they discovered a new enzyme that even worked in yeast cells from a commercial ethanol plant. In contrast to current cellulosic ethanol technologies the new enzyme can convert xylose in a single step and is not inhibited by other chemical compounds normally present within the yeast cells. The researchers have recently filed a patent application for their process. "This is a break-through in the commercialisation of cellulosic ethanol", comments

Boles. Boles says: "We have successfully demonstrated the conversion of waste sugars into ethanol. However, ethanol is not the best renewable biofuel. There are other alcohols with many more promising properties is now constructing yeast strains to convert plant waste materials into biobutanol, which is being seen as a more superior alternative fuel than ethanol due to its more favourable chemical and physical properties.

Goethe University Frankfurt (2009, March 9). Promising Biofuel Made From Commercial Yeasts Upgraded With A New Enzyme.

ONLINE BIOINFORMATICS EDUCATION

KNOW YOUR BIOINFORMATICS TOOLS: PART I

SWISS MODEL-A Fully Automated Protein Structure Homology- Modelling Server

Prediction of a protein's tertiary structure from its primary structure is of high importance in agriculture (structure prediction of proteins of important crop plants), medicine (in drug design) and biotechnology (in the design of novel enzymes).

SWISS-MODEL is a server for prediction of 3D structures of protein. Started in 1993, it is the most popular free web-based automated modeling facility today. The purpose of this server is to make Protein Modelling accessible to all biochemists and molecular biologists Worldwide. The advantages of using SWISS-MODEL are

- Automated modeling pipeline with improved hierarchical approach for template selection.
- Increased sensitivity of template detection (sequence to profile search using an adapted HHSearch protocol)
- New tools for model and structure quality assessment: Dfire and Qmean global scores; ProQres residue based assessment scores

The server is under constant development to improve the successful implementation of expert knowledge into an easy-browse server. It is designed to work with a minimum of user input, i.e. in the simplest case, only the amino acid sequence of a target protein. As comparative modeling projects can be of different complexity, additional user input may be necessary for some modeling projects, e.g. to select a different template or adjust the target-template alignment. Therefore, the SWISS-MODEL server gives the user the choice between three main interaction modes.1.First approach mode-The 'first approach mode' provides a simple interface and requires only an amino acid sequence as input data. The server will automatically select suitable templates. Optionally, the user can specify up to five template structures, either from the ExPDB library or uploaded coordinate files. The automated modeling procedure will start if at least one modeling template is available that has a sequence identity of more than 25% with the submitted target sequence. However, users need to be aware that the model reliability decreases as the sequence identity decreases and that target-template pairs sharing less than 50% sequence identity may often require manual adjustment of the alignment

2.Alignment mode-In the 'alignment mode' the modeling procedure is initiated by submitting asequence alignment. The user specifies which sequence in the given alignment is the target sequence and which one corresponds to a structurally known protein chain from the ExPDB template library. The server will build the model based on the given alignment.

3.Project mode-The 'project mode' allows the user to submit a manually optimized modeling request to the SWISS-MODEL server. The starting point for this mode is a DeepView project file. It contains the superposed template structures, and the alignment between the target and the templates. This mode gives the user control over a wide range of parameters, e.g. template selection or gap placement in the alignment. Furthermore, the project mode can also be used to iteratively improve the output of the 'first approach mode'.

All homology-modeling methods consist of the following four steps: (i) template selection; (ii) target template alignment; (iii) model building; and (iv) evaluation. These steps can be iteratively repeated, until a satisfying model structure is achieved.

SWISS-MODEL is accessible via a web interface at http://swissmodel.expasy.org or directly as a link from SWISSPROT entries on the ExPASy server (34). The program DeepView (Swiss-PdbViewer) can be downloaded for free at http://www.expasy.org/spdbv/. Depending on the complexity of the modeling task and server workload, it may take a few minutes to several hours for the server to build a model, including energy minimization. The model coordinates and log-files are returned to the user by email. The computational resources for the SWISS-MODEL server are provided by collaboration between the Swiss Institute of Bioinformatics at the Biozentrum Basel (University of Basel, Switzerland) and the Advanced Biomedical Computing Center (NCIFCRF Frederick, MD, USA).

Minakshi
Senior Research Fellow,
Unit of Simulation and Informatics, IARI, New Delhi-110012