BIOINFORMATICS INSTITUTE

The Bioinformatics Institute (BII) was set up by the Agency for Science and Technology Research (A*STAR) in July 2001. Located in the Biopolis, it was conceived as the computational biology research and postgraduate training institute as well as a national resource centre in bioinformatics within the Biomedical Research Council (BMRC) of A*STAR.

The BII focuses on theoretical approaches to understand biomolecular mechanisms that underlie biological phenomena, the development of computational methods to support this discovery process, and experimental verification of predicted molecular and cellular functions of genes and proteins with biochemical methods. Together with the BMRC, A*STAR research institutes and multinational R&D organizations in the Biopolis, the BII is situated in a conducive environment for exchange of scientific knowledge and friendly interaction that will prompt greater collaborations, and position the Biopolis as a notable biomedical R&D hub in Asia and in the world.

DIRECTOR’S WELCOME MESSAGE

“The Bioinformatics Institute has developed and deployed analytical tools and computational techniques for biology research in house and through close collaborations with experimental and clinical groups within and outside the Biopolis and Singapore.”

It is still premature to speak about life sciences as a theoretical scientific discipline. The extrapolation depth is small due to the fragmentary knowledge in a vast space of the unknown. Incremental accumulation of data as a result of hypothesis-driven experiments and observations is still the major source of new insight. Nevertheless, there are a few increasingly important research areas where the application of quantum, statistical, and mathematical concepts has become instrumental for the discovery of new biomolecular mechanisms and for progress in biological theory. This development has been fueled by the emergence of high-throughput experimental techniques (such as DNA sequencing, microscopy techniques, etc.). As a result, researchers can, for the first time, generate so much data that, essentially, the aim of describing living organisms in their totality has become realistic. Yet, the deluge of data is without understanding in terms of biomolecular mechanisms that link genome information and phenotypes. Computational biology has entered a new era characterized by the availability of fully sequenced genomes, as well as increasingly complete gene expression and proteomics datasets that wait for functional interpretation.

The Bioinformatics Institute, which was founded by Dr. Shankaran Raghavan in 2001 and led by myself since August 2007, is on its way to becoming a notable contributor of biologically relevant results and new, efficient computational biology methods to the world-wide scientific effort in the search for yet unknown biomolecular mechanisms, an effort with the goal of applications in medicine and biotechnology. The Institute carries out research in the areas of:

- biomolecular sequence analysis for the prediction of molecular and cellular functions (including the biochemical verification of hypotheses on function);
- biomolecular structure modelling and ligand design;
- gene expression profile analysis at the transcript and proteome levels;
- automated analysis of microscopic pictures from cellular systems (imaging informatics).

The Bioinformatics Institute has developed and deployed analytical tools and computational techniques for biology research in house and through close collaborations with experimental and clinical groups within and outside the Biopolis and Singapore. The members of our Institute are united in making this effort a success and we invite you to join us in this endeavor that will open new horizons in biology for the benefit of mankind.

Dr. Frank Eisenhaber
Director
Bioinformatics Institute, A*STAR

SCIENTIFIC ADVISORY BOARD

The Bioinformatics Institute is headed by a Director who is responsible for providing the scientific and administrative leadership to ensure that its goals are met. The Director is advised by a Scientific Advisory Board (SAB) consisting of eminent scientists in the field appointed for a period of time. The role of the SAB is to provide advice with respect to the institute’s research direction, recruitment of scientific staff and international research interactions. The current members of the SAB are:

Professor John Woolley
Associate Vice Chancellor of Research, University of California, San Diego

Professor Warren Ewens
Professor of Biology, University of Pennsylvania, Philadelphia

Professor Michael Levitt
Professor and Chair, Department of Structural Biology, Stanford University School of Medicine, Stanford

Professor Michael S Waterman
USC Associates Chair in Natural Sciences and Professor of Biological Sciences, Computer Science and Mathematics, University of Southern California, Los Angeles

Professor Eytan Domany
The Henry L Loeb Professorial Chair, Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel
RESEARCH DIVISIONS

Bioinformatics is a multi-disciplinary approach combining computational and biological expertise to analyse biological data (both generic and clinical), to advance biomedical research and development. Bioinformatics is both a science and an engineering art, concerned with the application of mathematics, physical/chemical principles and information technology to solve biological problems.

In the Bioinformatics Institute, there are four methodology-oriented research divisions comprising of research groups lead by independent principal investigators that focus on specific areas of computational biology. The common denominator is the goal of understanding biomolecular mechanisms underlying cellular phenomena, which is the basis for a rational understanding of pathogenesis or for planning biotechnological applications.

**RESEARCH DIVISIONS**

**BIOMOLECULAR FUNCTION DISCOVERY DIVISION**

**Principal Investigators**
- Frank Eisenhaber
- Sebastian Maurer-Stroh
- Georg Schneider
- Manfred Koranda

**GENOME AND GENE EXPRESSION DATA ANALYSIS DIVISION**

**Principal Investigators**
- Vladimir Kuznetsov
- Henry Yang
- Vivek Taravdik

**BIOMOLECULAR MODELLING AND DESIGN DIVISION**

**Principal Investigators**
- Chandra Verma
- Ivana Mihalek

**IMAGING INFORMATICS DIVISION**

**Principal Investigators**
- Lee Hywe Kuan
- Martin Wasser
PROTEIN FUNCTION PREDICTION

Although the complete sequencing of the human and other genomes was almost a decade ago, we find ourselves in the exciting situation that still only a subset of the identified genes is functionally well-characterized. The group embarks on the journey to map the functions in the many remaining uncharacterized genes, and the proteins they code for. This is not sufficient to look at the individual components of complex biological systems but rather consider the context of their interaction with each other to better understand how cells, organs and bodies work.

Based primarily on protein sequence analysis and the analysis of other sequence-associated data, the various aspects of molecular and cellular function (enzymatic activities, posttranslational modifications, disease interactions, pathways, etc.) are predicted. The work by the Protein Sequence Analysis group is strongly facilitated by the ANNOTATOR group which provides a software service for efficient sequence-annotation workflows. Finally, the hypothesis about protein functions are either followed up by experimental collaborations or are validated in the Division's own protein biochemical laboratory.

Protein sequence analysis and studies of genome-associated data is one area of research which the group focuses on. Embedded in the life science cluster Biopolis, the group is open to requests from local and international collaborators. The interaction with a Bioinformatics group like this can have dramatic impact on the research and performance of experimental labs. Often the group's analysis and predictions complement the experimental work, trigger new experiments or speed up findings of biological and medical importance which can be critical in certain competitive fields.

Another research area the group focuses on is the prediction of functional motifs in non-globular regions. Given the huge number of sequences of otherwise uncharacterised protein sequences, computer-aided prediction of posttranslational modifications (PTMs) and translocation signals from amino acid sequence becomes necessary. The group has contributed to this multi-faceted, worldwide effort with the development of prediction for GPI lipid anchor sites, for N-terminal N-myristoylation sites, for fatty acid and geranylgeranyl anchor attachment as well as for the PT5 peroxisomal signal.

Although the substrate protein sequence signals for various PTMs or translocation systems vary dramatically, the group has found that their principal architecture is similar for all the cases studied (refer to Figure 1).

SELECTED PUBLICATIONS


Figures:

- Figure 1: A small stretch of the amino and carboxy termini is buried in the catalytic core of the protein-modifying enzyme (on the binding site of the substrate). This face is usually involved in the interaction with other proteins, which is relatively conserved. This stretch is surrounded by linker segments that connect the parts bound by the enzyme with the rest of the substrate protein. These stretches are, in the end, small with a residue's backbone and polar. Due to the mechanical restrictions of binding to the enzyme, we suggest that most PTMs sites are necessarily embedded into intrinsically disordered regions (except for cases of a catalytically relevant PTM sites). This minimizes the need for tight interaction with other proteins with complex architectures.

- Figure 2: Sequences of the human PT5 peroxisomal signal. The sequence is highly conserved, with a high degree of similarity to other peroxisomal targeting signals. The figure also shows the alignments of this sequence with other peroxisomal targeting signals from different species, highlighting the conserved and variable regions. These alignments help in the prediction of peroxisomal targeting signals in other proteins.

ANNOTATION SOFTWARE DEVELOPMENT

The ever-increasing amount of data flowing into biological databases makes no signs of leveling off. Sequencing technology is improving at an unprecedented rate, bringing down the time it takes to decipher entire genomes to a matter of days. Meanwhile, the analysis of this data by predicting molecular functions is a time-consuming and tedious task. The number of new sequence analysis methods constantly being added to the toolbox of the computational biologist requires knowledge about the vast array of different interfaces, execution parameters and input formats.

The ANNOTATOR group is developing an advanced tool for functional characterization of sequences and strives to establish the ANNOTATOR software environment as the de facto standard in this field. The scope of work includes the integration of established algorithms as well as research into novel heuristics for teasing distant evolutionary relationships. Due to the complex nature of such heuristics, it is necessary to additionally consider aspects of high performance and distributed computing.

The ANNOTATOR, which was initially conceived by the Eisenhaber group at the Institute of Molecular Pathology in Vienna and is now developed and enhanced by the group at BI, provides an integrated environment for the analysis of sequences as well as other biologically relevant entities. Biological objects are represented in a unified data model and long-term persistence in a relational database is supplied by an object-relational mapping layer. Data to be analyzed can be provided in different formats ranging from web-based forms, FASTA formatted flat files to remote import over a SOAP interface.

SELECTED PUBLICATIONS

LIPID POST-TRANSLATIONAL MODIFICATIONS IN BIOMOLECULAR MECHANISMS

The lipid post-translational modification of proteins is a critical and rapidly evolving area of research. Recent advances in high-throughput proteomics and metabolomics have expanded our understanding of the complex regulation of lipid modifications and their functional consequences. In this context, a multidisciplinary approach involving biochemistry, molecular biology, and computational biology is necessary to fully comprehend the implications of these modifications in cellular processes.

SELECTED PUBLICATIONS


EVOLUTION OF PROTEIN STRUCTURE AND FUNCTION

The aim of this group is to reverse engineer the function of a protein through studying its evolution. Bioinformatics is used to get the first intuition of the layout and mechanism of these protein nanomachines, and computer simulation to test, to the extent it currently allows, the reasonability of the interpretation of the data. Ultimately, the goal is to build a straightforward hypothesis which can then be tested experimentally. Therefore, serious effort is invested into developing ways to present the group’s findings in an easily accessible and comprehensive way to experimentalists and colleagues.

In evolution, as in any statistical process, anything that can happen will happen. Compared, however, to the options open to a simple physical system, “can happen” is a somewhat more elaborate condition. While the physics of DNA stability may allow for a mutation, this mutation might severely degrade the stability of the protein it encodes, which in turn may kill the organism carrying the mutation. Another mutation might be irrelevant to the protein stability, but it may adversely impact its interaction with another protein, thus disrupting a pathway in the hapless organism. Keeping that scenario in mind, a combination of mutations performing the same function in living and thriving organisms can be added, in addition to identifying the regions in the protein in which mutations, or certain types of mutation, are conspicuously absent. Since it can be reasonably assumed that mutations do happen sporadically in those places, as they do in all underlying positions in the DNA, it is possible that the carriers were eliminated from the gene pool because the mutation resulted in some disadvantage for the organism, be it on the translational, folding, or protein-protein interaction level.

When trying to understand the protein function, a top-down (reverse engineering) approach may be adopted to aid the search for functional regions among the most conserved ones. While detecting such “conserved” regions in a protein may not be a very challenging bioinformatics task, interpreting their meaning might have to be quite complex. Thus, for example, a study of a protein called K (from the large group of silhouette-related proteins) could not be understood solely from the sequence comparison. The results were turned straight away to experimental colleagues who were able to establish, through site-directed mutagenesis, that several pathways were critically affected, distinct pathways intriguingly assorting with distinct protein regions.

This group’s goal is to push forward, through analogical and explicit computer simulation, the point at which the experiment needs to be invoked, thus bolstering the understanding of proteins and shortening the benchwork time.

SELECTED PUBLICATIONS

1) Mihalek I, Kend I, Lichtenauer O. Background frequencies for residue variability estimates. BLOSUM revisited. BMC Bioinformatics. 2007 Dec 27;8:488 (Epub ahead of print)


COMPUTATIONAL ANALYSIS OF GENOME COMPLEXITY, TRANSCRIPTION REGULATION AND CELLULAR PHENOTYPES

The group attempts to understand the complexities of gene and genome architecture, transcription regulation and their relevance for phenotypic properties of cells. This group develops and applies integrative bioinformatics, probabilistic and computational systems biology methods, which are used in the analysis of sequences and gene expression data, ranging from short regulatory sequences, motifs, individual genes and gene pairs, to the entire genome, gene regulatory networks and pathways.

In collaborations with other groups within the bioinformatics Institute, other ASTAR institutes and notabe overseas groups, this group focuses on the challenges and successful problems of gene expression events in normal and cancerous cells; protein-DNA interactions in transcription regulation; computational genome cartography of transcription factors binding sites; local and distant regulations of transcriptional machinery; direct gene targets for transcription factors and gene expression and transcription and proteome complexes.

The group translates their study of gene expression patterns and analyses of macromolecular structures and interactions to support the classification of diseases, biological validation of in silico findings, individual disease prognosis and relief events, and the identification of novel biologically essential and clinically significant risk factors and its reliable combinations.

Currently, the group is developing computational algorithms and statistical methods for (1) better understanding regulation and co-regulation of transcriptional machinery associated with expression of c-erit, STAT1 and several other key transcription factors, (2) identification and prediction novel evolutionarily conserved regulatory sequences in the human genome and (3) analysis of complex human genome architectures - including cis-sense anti-sense gene pairs. This group is also developing and applying data mining, statistical pattern recognition algorithms, probabilistic models and network biology tools to improve diagnostics, individual prediction and optimization of treatment assignment for breast and lung cancers.

SELECTED PUBLICATIONS


INTEGRATED LARGE SCALE DATA ANALYSIS FOR FUNCTIONAL DISCOVERY

The group’s main research activities have been focused on analysis and integration of genomic data to perform data-driven discovery using an integrated and interdisciplinary approach (Bioinformatics, Systems Biology, & Stem Cell Biology).

Bioinformatics provides computational tools for integrated data mining and annotation, while Systems Biology approaches are employed for the integration of different ‘omics’ data (MPS/SAGE, CDNA microarray, ChIP-chip array, etc) and statistical pathway analysis. Both bioinformatics and Systems Biology approaches are linked with biology for novel gene/pathway identification and experimental validation. Thus, most of the group’s research projects are collaborations with Stem Cell biologists e.g. from the Genome Institute of Singapore (GIS). The group also works with wet-lab biologists in other research areas such as cancer, drug discovery, and neurology.

With the generation of large-scale data by various high-throughput technologies (MPS to SAGE to microarrays for global gene expression profiling, ChIP-chip to ChIP-seq for DNA-protein binding, and two-hybrid assays for protein-protein interactions), there is a need for integrating all these different kinds of data for biological discovery. The group approaches this problem from three aspects: (1) data analysis, (2) data integration, and (3) development of integration software platforms.

In order to obtain more accurate analysis results for large-scale datasets, the group employs several spike-in microarray datasets with a large amount of both spike-in and replicates. Using the benchmark data, the group is able to statistically evaluate various data analysis methods as well as develop novel and effective data analysis methods, in particular for data with a small number of replicates.

SELECTED PUBLICATIONS


EXPRESSION AND SIGNALING IN MESENCHYMAL AND HEMATOPOIETIC STEM CELLS

Adult stem cells have the potential to differentiate into a wide variety of tissue specific cells. These cells can therefore be used to treat a variety of disorders ranging from myocardial infarction to osteoporosis. Although hematopoietic stem cells found in the marrow have been well studied for many decades and are routinely used for treating hematological disorders, there is increased interest in mesenchymal stem cells which are the non hematopoietic stem cells found in the marrow.

Although these cells are already being used clinically, we know very little about the mechanisms these cells use to differentiate to different lineages. This group aims to understand these mechanisms using gene expression analysis and use this information for designing better media suitable for targeted differentiation of mesenchymal stem cells.

The goal of this group is to understand mechanisms used by mesenchymal stem cells to differentiate into fat, cartilage and bone. Using detailed analysis of gene expression by microarray, three signaling pathways were predicted to be critical for differentiation and survival of mesenchymal stem cells. The significance of these pathways was then verified by inhibition of receptor kinases using specific small molecule inhibitors. The figure (left) shows the effect of TGF-β signaling on mesenchymal stem cells. The microarray data predicted that TGF-β Pathway was downregulated in adipogenesis and upregulated in chondrogenesis, in keeping with this prediction, inhibition of TGF-β signaling by 58431542 enhanced the adipogenic differentiation of mesenchymal stem cells and blocked their chondrogenic differentiation.

This project was carried out in collaboration with Immunology and Identification of these pathways has led to the development of the first commercially available serum-free medium for culturing mesenchymal stem cells.

SELECTED PUBLICATIONS


Lee Hwee Kuan  
Postdoctoral Fellow: LAW Yi, WHAM EEXT, YU Weimiao  
Research Associate: YIP Cheon Kang, U ILBIIH

Martin Wasser  
Postdoctoral Fellow: CHIN RYB, KJSTONANUIZ, DLI Tienan  
Research Associate: PJAM Wey Choo

**COMPUTER VISION AND PATTERN DISCOVERY**

The group of Computer Vision and Pattern Discovery for Bioimages focuses on applying advanced computer vision, machine learning and mathematical models to elucidate the complex behaviour of biological systems. The group analyses images from wide-field and confocal microscopes, including image data sets from high-throughput screens. The trend towards quantitative biology has spawned new areas of research, especially in the area of digital imaging where thousands of images are acquired automatically through robotic systems of chemical and cell assay handling. These images are then analyzed and used to create new biological hypotheses that are further validated using other experimental means. The group's contributions to high-throughput, high-content imaging are to provide, accurate and fast computational methods for the data mining of large image data sets.

**Stochastic Level Sets method for cell segmentation:**  
The Mumford-Shah model is one of the most successful image segmentation models for biomedical images. But existing algorithms for the model require a good initial guess to obtain good results and are therefore impractical. The group has developed a well-designed basin-hopping scheme that uses global optimisation to escape from local traps in a way that is much more effective than standard stochastic methods. This algorithm is independent of initial conditions and can compute the global or near global optimal solution.

**High-throughput Neural Cell analysis using a novel level set method:**

The group has developed a novel seeds controlled level set method to process neural cells images, where the segments of the nuclei are used as the seeds for the level set functions. Merging and splitting of the zero level set curves are prevented using dynamic watershed lines. Validation of the proposed method showed that it outperforms CellProfiler and MetaMorph. CellProfiler is one of the best free image analysis software freely developed by the Broad Institute of Harvard and MIT. MetaMorph is a commercial software specially developed for cellular image analysis by MDS Inc. The computational results on a data set of 6000 images show that it is fast and accurate. Other information, such as the number, length and size of the cells can also be automatically extracted. Lastly, the formula can be easily extended to three dimensions.

**Analysis of high-throughput siRNA screens on Human Keratinocytes:**

The group presents a computational framework for detection of keratin 14 aggregates in images of in vitro cultured Keratinocytes when phosphate genes are silenced using siRNA. Phenotypic changes are observed in only 1-2% of the cells and in 10 of the genes silenced causes the keratin intermediate filaments to dissociate and form aggregates. We sort the images according to the amount of keratin aggregates detected. 85% of the top 100 images, out of 1802 images in the training set, are correctly detected.

A total of 20,000 images are processed within a few days on a desktop machine.

**SELECTED PUBLICATIONS**


**LIVE CELL IMAGING AND AUTOMATION OF IMAGE ANALYSIS**

The group is interested in studying animal development using 3D time-lapse microscopy and computer vision. Their principal goals are to develop protocols for live-cell microscopy and software tools for the automated analysis of multi-dimensional image data. The major research activities are directed at constructing the components of a computational pipeline that will be integrated into an intelligent image analysis system. The computational pipeline covers preprocessing, segmentation, feature extraction, classification and cell tracking (figure 1). Currently, the system is directed at the phenotypic characterization of two biological processes in the model system Drosophila melanogaster: (1) Cell cycle progression of embryonic cells and (2) apoptosis and remodeling of muscle cells during metamorphosis.

The study of cells in their natural tissue environment promises to uncover novel insights into the mechanics and regulation of cell proliferation that cannot be observed in homogenous cell culture. The tracking of cell divisions in Drosophila embryos is accomplished by observing live cells labeled with fluorescent proteins using wide-field macroscopy. Desirable features, e.g. DNA content, and identifiable markers such as the phase of the cell cycle will be used to characterize the function of known genes and identifying novel players by genetic screens.

The second biological theme is the deconstruction and re-assembly of tissues during metamorphosis. The group focuses on the muscular system and uses fluorescence live cell imaging to study apoptosis of obesolite and remodeling of persistent larval into adult muscles. The structural organization of muscles is accompanied by initial degrading and later increasing thickness of the muscle fiber. Therefore, studying the dynamics of muscle remodeling in flies might evolve into an animal model for muscle atrophy and hypertrophy.

Observations of real biological processes in living tissues are the main driving force behind the development of novel software tools for quantitative microscopy. Motion in live tissues that cause misalignments of frames in image stacks prompted the development of an algorithm for image registration. The faithful segmentation of object regions is a prerequisite for accurate feature detection and measurement (figure 2). However, inhomogeneous signal intensities and contrast levels in 3D image stacks demand novel strategies for segmentation and object recognition that take into account prior biological knowledge. In order to derive biological meaning from movies, analysis systems need to identify the different action in the scene and characterise their behavior. Machine-learning techniques are an essential tool in classifying the image objects according to cell type or phase of the cell cycle.

**SELECTED PUBLICATIONS**

**BIO-COMPUTING CENTRE (BCC)**

*Overview*

The Bio-Computing Centre's (BCC) mission is to apply, innovate and provide information technology to enable advances and new discovery in the biomedical sciences for the user community within and beyond the Biopolis. Focusing on data-oriented biomedical computing applications linking a variety of data sources from the bench to the bedside, BCC serves as a central resource for data integration, modelling and analysis.

It achieves this by deploying cutting-edge cyber-infrastructure through the provision of software, hardware and human resources within a multi-disciplinary ethos to address the needs of its user community. The BCC serves as a leadership role in a national cyber-infrastructure centre with its staff involved actively in a number of national initiatives such as the A*STAR Computational Resource Centres (A*CRC) ICT Infrastructure at Fusionopolis and SingaREN.

---

**BCC ORGANISATION CHART**

- **BCC**
  - **NETWORK SERVICES**
  - **SYSTEMS ADMINISTRATION**
  - **IT SECURITY**
  - **WEB SERVICES**

---

**AREAS OF EXPERTISE**
NETWORK SERVICES

Team Leader: William LAI
Team Members: Derrick LAU, Darren SEE, OIUA See Hian

The Network team is responsible for the infrastructure planning, design, administration and daily operation of the BIIT Network Operation Centre (NOC) and Data Centre in the Matrix building in the Biopolis. The Data Centre hosts application servers, network and IT equipment for the following Singapore government funded entities and affiliated institutes:

- Bioinformatics Institute, A*STAR
- A*STAR Headquarters
- A*STAR Computational Resource Centre
- A*STAR’s Biomedical Sciences Institutes
- A*STAR’s Biomedical Research Council’s (BMRC) Data Protection Office
- Singapore Centre for Cohort Studies
- Singapore Tissue Network

Roles and Responsibilities

- Oversees the daily operation of the network services of the IT Infrastructure in BIIT
- Involves in network design, planning and monitoring, and helpdesk support
- Provides IT advice and guidance to BMIS institutes or collaborators on how to develop and deploy scalable network designs, solutions, policies, and recommendations.

Project Highlights

- Administration and operation of Biopolis ICT Cyber-Infrastructures (SELECT) project for A*STAR
- Conceptualisation, planning, design, writing, reviewing and implementation of the new Biopolis ICT Infrastructure for A*STAR
- Design and build the Cyber-Infrastructures for the Fusionopolis’ High Performance Computing (HPC) Facility and services (including IT infrastructure, network backbone, network security, network connectivity, telecommunication infrastructure and Internet services)
- Design and setup a DWDM optical network to link up the 2 Data Centres between Biopolis and Fusionopolis

SYSTEMS ADMINISTRATION

Team Leader: Caleb KHOR
Team Members: CHAN Ang Loon, Charlie TAN

The role of the Systems Administration team is to ensure that BIIT’s corporate system, which reside on the campus LAN, are robust and secure. Apart from providing the high quality corporate IT services expected of a technology organization (including both running a fail-safe server infrastructure as well as desktop support), the Systems Administration team also undertakes development efforts as part of its charter to constantly enhance the functionalities, reliability of all the systems under its charge.

Roles and Responsibilities

- Provide technical and systems support to all scientific, IT and corporate users in BIIT

IT SECURITY

Team Leader: James TAY
Team Members: TAY Teck Woe, LIM Swee Teck, ROHNEE TANG

The key role of the IT Security Team is to ensure a robust and secure Biopolis infrastructure, protected from malicious intent from internal as well as external bodies. The team aims to address a variety of complex threats and to provide security solutions to the Bio-Computing Centre on appropriate preventive measures for systems and servers, databases, infrastructure and other IT environments.

Roles and Responsibilities

- Development of the overall systems and network security policies
- Implementation of appropriate systems security measures and procedures
- Investigates abuse, security violations, conducts regular penetration/exploit testing, and inspects code and scripts for loopholes.

WEB SERVICES

Team Leader: ZHOU Shaoxin
Team Members: JOHNSON CHEN, LIM Zhi Yen, MAKELA ANASTASIJA, Zenn QIU

The Web Services team is mainly responsible for the design and maintenance of the BIIT home pages on the internet and the Intranet websites. As the BIIT home page and its associated web pages are the first point of contact between the Institute and public, the team ensures that the electronic information published by BIIT is of high accuracy and quality. The content published is also fundamental in relaying the strong reputation and image of the Institute. The team will also ensure that the information published electronically is accurate, visually appealing, clearly presented and follows closely to the Singapore Government Web Interface standards set by IDA.

Besides servicing the internal research groups within BIIT, the Web Services team is also assisting external parties such as the research institutes, research unit and consortiums within A*STAR in the areas of web and application development, website hosting and graphics design and consultancy.

Roles and Responsibilities

- Design and maintenance of the BIIT Internet and Intranet websites
- Involved in the design, maintenance and development of all web portals associated with the BIIT research groups, including different knowledgebases for Defensins, p53, Stem Cell and Cancer
- Develops, designs and maintains corporate portals for other research institutes and consortiums within A*STAR, as well as collaborators with BIIT
- Other services include graphic and multimedia design, web site design and layout, scanning and photography services, web publishing support and web database hosting.

Projects Highlights

- Corporate portals and hosting of portals for research institutes and consortiums within A*STAR
- Experimental Therapeutics Centre
- Institute of Medical Biology
- Singapore Institute of Clinical Sciences
- Singapore Biomimetic Consortium
- Singapore Immunology Network
- Singapore Stem Cell Consortium
- Biological Resource Centre
- Biopolis Shared Facilities
- Data Protection Office
- Web portals
  - Defensins Knowledgebase (Dii & SERI)
  - Human Intermediate Filament Mutation Database (MBM)
  - p53 Knowledgebase (MBMC & NCC)
  - Gastric Cancer (Gastriclines) Knowledgebase (GIS, NUS and NUH)
  - Singapore Life Science Virtual Grid Community
  - Singapore Human Mutation & Polyomavirus Database (Oc Tan Ee Choo, KK Women’s and Children’s Hospital)
SOFTWARE ENGINEERING

Team Leader: CHEOK Leong Poh
Team Members: NG Wee Thong; Kelvin GOH; KWO Chia Yee; MOHAMED Hanifa; VOOON Kian Loon; AW Siew Cheng; LUA Seow Chin

The Software Engineering team serves as a bridge between the scientists and IT professionals to develop software tools/applications that leverage on the computing, storage and network infrastructure in BII. The Software Engineers in the team work in close collaboration with scientists from A*STAR’s research institutes to address their needs in scientific software solutions. Besides earlier software projects focus on Genomics, Proteomics and Bioimaging, the team will embark on Laboratory Information Management System (LIMS) development in the future.

Roles and Responsibilities

- Translate user requirements into design specification, develop and deliver software solution to meets scientific objective
- Evaluates new technologies and methodologies to enhance discovery platforms developed in collaboration with the scientists
- Provide expertise in software engineering to BII’s research groups
- Actively bridge BII with other A*STAR research institutes with scientific software solutions that leverage on BII’s computing infrastructure

Project Highlights

- Screensaver, collaboration with Institute of Molecular and Cell Biology (IMCB)
- IT collaboration work with Singapore Tissue Network (STN)
- Database for Integrated Stem Cell Research Online (internal collaboration)