

My Trip to ISB 2011

The 2nd Indo-US Symposium on Biocomputing

Held at

Taj Gateway

Calicut, India

12 and 13 September, 2011

Trip account by

Sriganesh SRIHARI

Department of Computer Science, National University of Singapore, Singapore

ISB 2011

- Jointly organized by

- National Institute of Technology Calicut (NITC)

(My alma mater – part of the IIT-NIT system under the Central Government of India)



- Indiana University Purdue University Indianapolis (IUPUI)

A collaborative university set up between Indiana University Bloomington and Purdue University West Lafayette

- Funded by NSF, DST/DBT

ISB 2011

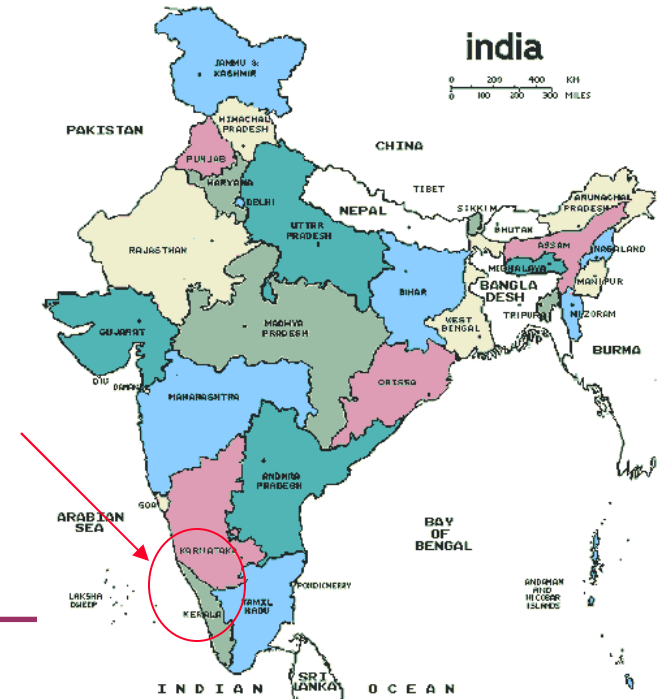
- ISB 2011 was the second symposium in the ISB series
- This year organized as a workshop with lectures from eminent scientists from US and India
- Last year (2010) organized as a conference with proceedings published by Springer/ACM

About IIT-NIT

- Part of the IIT-NIT system in India
 - Set up under the “Institutes of National Importance” Act
 - IITs – completely Central Government
There are 8 “old” and 7 “new” IITs
 - NIT – part Central, part State Government
Each state in India has an NIT
 - Common all India entrance test
Either choose an IIT or choose an NIT close to your home!
In my time (1999), 0.2 million students competed for ~8000 seats!!
~ 2000 in the IIT system and ~ 6000 in the NIT system

NIT Calicut

- Set up in **1961** in the state of “Kerala”
- Southernmost state
- Close to Bangalore and Chennai
- This year is the 50th anniversary



NIT Calicut

- An engineering institute
 - CSE, ECE, EEE, ME, PE, CE, Archi, BT, ChE
 - In 1999, 400 students per batch (very less for a country like India!)
- Predominantly teaching. Very less research
 - Bachelor of Technology (B.Tech.),
 - Master of Technology (M.Tech.),
 - Doctor of Philosophy (PhD)
Very few students (about 5-10 per department)

ISB - Aims and Objectives

- Increased collaborative projects between the US and India
- Setting up of a “Center for Excellence” (COE) at NIT Calicut focusing on Bioinformatics/Computational Biology research
- Producing higher number of research scientists in India

ISB Venue

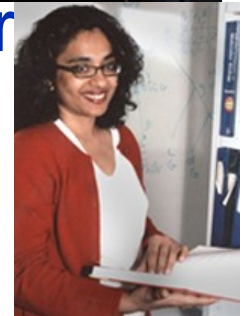
- Taj Hotel in Calicut



ISB 2011

Invited Speakers

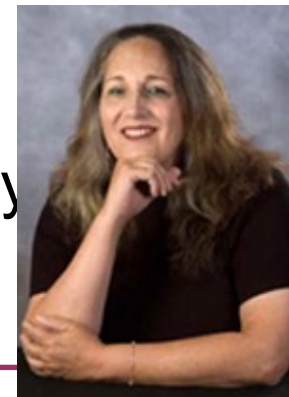
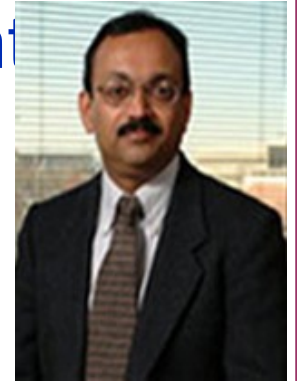
- Keith Dunker (Biologist)
 - Indiana University
- Thomas Ferrin (Pharmaceutical)
 - UCSF
- Mona Singh (Computer Scientist/Software development)
 - Princeton



ISB 2011

Invited Speakers

- Mathew Palakal (Software Development)
 - Indiana University
- Lang Li (Biologist – Clinical)
 - Indiana University Med School
- Tarynn Witten (Biostatistician)
 - Virginia Commonwealth University



ISB 2011

Invited Speakers

- Teresa Przytycka (Computer Scientist)

- NIH/NCBI



- Rajanikanth G. K. (Biologist)

- NIT Calicut



- Pramod Wangikar (Chemical Engineer)

- IIT Bombay



ISB 2011

Invited Speakers

- **G P S Raghava (Biologist)**
 - Institute of Microbial Technology, India
- **Pawan K Dhar (Biologist)**
 - Symbiosis International University, India
- **Ashish Tendulkar (Computer Scientist)**
 - IIT Madras
- **Ramesh Hariharan (Computer Scientist)**
 - IISc Bangalore, CEO Strand Life Sciences Bangalore



ISB 2011

Invited Speakers and Organizers

- Group photo



Keith Dunker's Talk



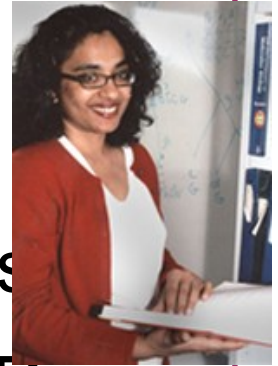
- **Disordered proteins**
 - Not all proteins attain a stable 3D conformation
 - Such proteins are “disordered”
 - Some regions of “structured” proteins can also be disordered
 - Cannot be crystallized using X-ray crystallography
 - Usually represented as an “ensemble” or a probability distribution of different structures
 - Cyclin-CDK complexes

Thomas Ferrin's Talk



- Software development platforms
 - Chimera
Visualization software for sequences to structures
 - Cytoscape
Molecular network visualization
- Did a demo of Chimera
- Explained how important it is to develop such software platforms

Mona Singh's Talk



- Finding ligand binding residues in proteins
- Uses the amino acid sequence and the 3D structure
- Ligand dock into “concavities” in the protein
- Algorithm to find “concavities” from the 3D structure of protein
- Incorporating amino acid information boosts the prediction accuracy

Mathew Palakal's Talk



- Literature mining to build molecular networks
 - Eg. Building protein associations by mining text
- “Bibliomics”
- Developed a framework and NLP based text mining algorithms for mining literature
- Extended to “medical” networks
 - Clinical networks
 - Drug association networks

Lang Li's Talk



- Genetic variation in populations and responses to drugs
- Understanding the effects of drugs on different kind of populations
 - Case studies on breast cancer patients
 - How drugs react with different enzymes in different populations
 - Statistical study

Rajanikanth's Talk



- Therapeutic studies for neurological disorders
- How different drugs impact on different disorders
- Association among the drugs
- Statistical techniques to study diverse populations

Pramod Wangikar's Talk



- Energy from Cynobacteria strains (algae)
- Identify the genes of interest
- Build the gene regulatory networks (GRN)
- “Twig” these networks to improve energy production from these bacteria

- GRN
 - Networks are very large
 - Use operon finding tools to compress the gene space

Raghava's Talk



- Developing therapeutic peptides
- Aim is to develop vaccines
- Usually vaccines are denatured or weakened pathogens
 - These stimulate the immunosystem of the body
- No need to use pathogens
- Use directly only the proteins that can stimulate the immunosystem
- No need of full proteins also. Just use peptides
- Developing such therapeutic peptide database

Pawan Dhar's Talk



- About 98% of the DNA comprises of non-coding regions (junk DNA)
- Why did nature do so?
- Do these regions have any potential?
- Capitalize on the potential of these regions
 - Identified some non-coding DNA regions
 - Expressed these non-coding genes!
 - Developed synthetic proteins!
 - Therapeutic value of these synthetic proteins

Ashish Tendulkar's Talk



- Database search for protein structures
- Developed a NLP-based method
 - Break the interest protein into fragments
 - “Bag of fragments” approach
 - Keep also the positional information
 - Represent these fragments as text
 - Use NLP based text mining approaches to match these fragments with database
 - Retrieve the proteins structures similar to interest protein

Ramesh Hariharan's Talk



- Re-sequencing of whole genomes
 - Different from de novo sequencing
 - Know the reference genome, which is usually of a closely-related species
 - No concepts of contigs here.
 - Take the reads and match it to positions in the reference genome
 - Use suffix arrays compressed using BW Transform to store and search in the reference genome
 - Avadis platform

Tarynn Witten's Talk



- Impact of experimental methods on experimental results
- While doing the experiments, parameters and environment can change!
- Designing theoretical models to capture the results under theoretically “exact” situations and compare them with experimental results
- A very “general” talk

Teresa Przytycka's Talk



- Genotype to Phenotype relationship
- Understanding SNP and mutational changes in populations
- Connecting them to differences in gene regulation and expression
- Connecting this to differences in protein sequence and structure
- Connecting this to phenotypic changes in species!

Break-out Session

- Aim: To start collaborative projects between India and the US.
 - Start a Bioinformatics COE in NITC
- Formed groups of 7 – 10 people
 - 2 Scientists
 - PhD students
 - Discuss your ideas, project collaborations, any help needed, etc.



Lots to learn...

- Amazing talks
- Learnt lots of new things
 - New areas of research
 - New approaches
 - New projects

Lots to Learn...

- **Raghava and Lang Li:**
 - Bioinformatics/Comp Bio is a very interdisciplinary field
 - Easy to “get lost on the way”
 - **Very important to be strong in at least one field – either CS or Biology**
 - Otherwise end up becoming just a “jack of all trades”
 - **If you are a CS student, using only softwares is of NO use!**
 - Biologists can use the softwares themselves – very user-friendly these days

Personal learning...

(My opinions only)

- In India, I saw that the acceptance of bioinformatics/comp bio as a field has lot of hesitancy currently
 - CS depts say that this is not a part of CS
 - Bio depts say that there is no experimental work involved and so not a part of Bio
 - Getting accepted is quite difficult in “traditional” institutes like IITs and NITs

Overall...

- Excellent experience
- Got to meet renowned scientists
 - Particularly Keith Dunker
- Met my professors at NITC
- Visited the campus
 - Lots of new buildings have come up
- Thanks to Prof Leong for partly funding my trip