Bioinformatic and Public Health Research Databases, Collaborative Environments, and Virtual Organizations


Department of Biological Sciences
Department of Computer Science & Engineering

Workshop on Interdisciplinary Biomedical Research
University of Notre Dame
April 10 - 11, 2008
Interdisciplinary Collaboration

• Five year collaboration between the Departments of Biological Sciences and Computer Science & Engineering
  – VectorBase: NIH/NIAID Bioinformatic Resource Center
  – Malaria Transmission Consortium (MTC): Bill & Melinda Gates Foundation project

• International collaborators

• International users of resources
Cyberinfrastructure

• Enabling Technologies for Collaboration
  1. Development of Shared Resources
  2. Use of Shared Resources

• Convergence of
  – Collaboratories
  – Virtual Organizations (VO) / Virtual Teams
  – Web 2.0
  – Computer Supported Collaborative Work (CSCW)
Cyberinfrastructure / VO Functionality

VO Functionality
- Communication
  - Email lists, Listserv
  - Audio Conferences
  - Video Conferences
  - Chat/IM
  - Web Casts
  - Wiki
- Information Sharing
  - Web Pages
  - Bulletin Boards
  - Documents Upload / Download
  - Blast, Clustal, Hammr
  - HPC Pipelines
- Computation / Analysis
  - Data files
  - Data base/SQL
  - Experimental telepresence
  - Sensor nets
- Data Storage / Retrieval
- Experiments
- Data Collection
- Workflow
- Security
- Application Pipeline
- Job Scheduling and distribution
- Authentication
- Encryption
- Privacy
CAUSES OF DEATH

Developed World
- Circ. Syst. 46%
- Cancer 21%
- Resp. Syst. 8%
- Other 23%
- Inf. Dis. 1%

Developing World
- Circ. Syst. 24%
- Cancer 9%
- Resp. Syst. 5%
- Other 9%
- Perinatal 10%
- Inf. Dis. 43%
### Pathogen-Specific Causes of Death by Infectious Diseases (1997)

<table>
<thead>
<tr>
<th>Affection</th>
<th>Annual Mortality</th>
<th>Prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TB</strong></td>
<td>2,900,000</td>
<td>16,300,000</td>
</tr>
<tr>
<td><strong>AIDS</strong></td>
<td>2,300,000</td>
<td>&gt;30,000,000</td>
</tr>
<tr>
<td><strong>Malaria</strong></td>
<td>~2,000,000</td>
<td>&gt;500,000,000</td>
</tr>
<tr>
<td><strong>Arboviruses</strong></td>
<td>200,000</td>
<td>------</td>
</tr>
<tr>
<td><strong>African Trypanosomiasis</strong></td>
<td>100,000</td>
<td>400,000</td>
</tr>
<tr>
<td><strong>Leishmaniasis</strong></td>
<td>80,000</td>
<td>12,000,000</td>
</tr>
<tr>
<td><strong>American Trypanosomiasis</strong></td>
<td>45,000</td>
<td>18,000,000</td>
</tr>
<tr>
<td><strong>Filariasis + Onchocerciasis</strong></td>
<td>------</td>
<td>140,000,000</td>
</tr>
</tbody>
</table>
Malaria Transmission Consortium (MTC)

“The University of Notre Dame has received a $20 million grant from the Bill & Melinda Gates Foundation to support research aimed at evaluating existing malaria control programs and designing improved methods for malaria control. This study represents multiple institutions in multiple countries, carrying out multiple field studies, gathering huge amounts of data, developing appropriate data bases, data analysis and management methods, geographical information systems, maps, data quality assessment, modeling and simulation — all activities that will involve, to a very large extent, contributions from the College of Engineering.” (News Release - September 2007)
Malaria and Global Health

- Malaria is responsible for more than 1 million deaths per year among infants and small children.
- Malaria, TB and HIV/AIDS are top three pathogen-specific causes of human death.
- Malaria is a disease of poverty in the tropics, especially Africa.
Malaria Transmission Consortium

• Bill & Melinda Gates Foundation-funded Project to Implement and Evaluate Malaria Control
• Project Leadership by Notre Dame Colleges of Engineering & Science
• $20 Million, 5-year Program
• Consortium Project
MTC Participants

- Notre Dame - Lead Organization
  - Biological Science & CSE, Center for Global Health and Infectious Diseases, Frank Collins
  - CSE, College of Engineering, Greg Madey
- U.S. Centers for Disease Control
- London School of Tropical Medicine and Hygiene
- Durham University, UK
- Swiss Tropical Institute
- Partners in Indonesia, Tanzania, Kenya, Uganda, and Zambia
Major Objectives of Malaria Transmission Consortium

• Develop Improved and Measures of Malaria Transmission
• Use Impact on Malaria Transmission to Assess Malaria Control Methods
  – Insecticide-impregnated Bed Nets
  – Indoor Residual Insecticide Sprays
  – Combination Treatments
• Assess Mosquito Characteristics that Impact Control Effectiveness
  – Insecticide Resistance
  – Blood Feeding Behavior
• Develop Data Management and Analysis Systems
  – Web-based Relational Database Management System
  – Simulation Modeling
Malaria Life Cycle
Emergence of new mosquitoes everyday

Host-seeking

May encounter any number of different types of hosts

Death while host-seeking

Encounters human with ITN

Encounters unprotected human

Death while feeding

Susceptible mosquitoes may get infected.
Infectious mosquitoes may infect humans.

Figure from Paul Libiszowski
Mosquito Feeding Cycle

- Emergence of new mosquitoes everyday
- Host-seeking
- Alive but no host encountered
- May encounter any number of different types of hosts
- Fed
- ITNs
- Death while host-seeking
- ITNs, IRS, House-screening, Repellants
- IRS
- Ovipositing
- Death while ovipositing
- Death while resting
- Resting
- ITNs
- Death while feeding
- Death while escaping host

Figure from Paul Libiszowski
Activities

• Database design, web interface, deployment

• Geographical Information Systems (GIS)

• Handheld PDA/GPS data collection and transfer to Notre Dame
Activities

- Agent Based Simulation & Modeling
- Data mining, statistical analysis, assessment
- Predictive tools, decision support systems, operations research
VectorBase

“The Center for Global Health & Infectious Diseases at the University of Notre Dame has been awarded a $10 million contract from the National Institute of Allergy and Infectious Diseases (NIAID), an arm of the National Institutes of Health (NIH), to create and maintain a public Bioinformatics Resource Center (BRC) that will manage genomic information on insects and other arthropods that transmit human pathogens.” (News Release - August 2004)
VectorBase

• PI - Frank Collins, Biological Sciences & CSE, ND
• Co-PI’s/Subcontractors
  – EMBL/Imperial College (Fotis Kafatos)
  – EBI Hixton, Cambridge (Ewan Birney)
  – Institute of Molecular Biology, Crete, (Christos Louis)
  – Harvard (William Gelbart)
  – CSE Notre Dame (Greg Madey)
Activities

- Coordinate with authors of “white papers” and sequencing/assembly centers
- Gene annotation pipeline: auto, manual, and community
- Development and deployment of VO functionality
  - Genome browsers, CV & ontologies, search, bioinformatic tools
Species: *Anopheles gambiae* PEST
Genome size: 260 Mb
Status: 4th assembly and annotation
NIAID funded
Species: *Aedes aegypti Liverpool*

Genome size: 1.3 Gb

Status: 1st assembly and annotation

NIAID funded
Species: *Pediculus humanus USDA*
Genome size: 105 Mb
Status: 1st assembly and annotation completed
NHGRI funded
Additional Species in Progress

- *Culex pipiens*: 540 Mb
- *Ixodes scapularus*: 2.1 Gb
- *Rhodnius prolixus*: 670 Mb
- *Glossina morsitans*: 600 Mb
- Sand Flies: ~450 Mb
- 13 *Anopheles* species: 3.4 Gb
- 100 *Anopheles gambiae* genomes: 26 Gb
<table>
<thead>
<tr>
<th>Month</th>
<th>Unique visitors</th>
<th>Number of visits</th>
<th>Pages</th>
<th>Hits</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 2007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Feb 2007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mar 2007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Apr 2007</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>May 2007</td>
<td>6</td>
<td>25</td>
<td>2311</td>
<td>5142</td>
<td>39.75 MB</td>
</tr>
<tr>
<td>Jun 2007</td>
<td>1434</td>
<td>3081</td>
<td>177721</td>
<td>618770</td>
<td>26.00 GB</td>
</tr>
<tr>
<td>Jul 2007</td>
<td>1486</td>
<td>3530</td>
<td>158258</td>
<td>560213</td>
<td>29.67 GB</td>
</tr>
<tr>
<td>Aug 2007</td>
<td>1600</td>
<td>3978</td>
<td>156540</td>
<td>586051</td>
<td>34.06 GB</td>
</tr>
<tr>
<td>Sep 2007</td>
<td>2067</td>
<td>4549</td>
<td>184734</td>
<td>693155</td>
<td>3785.00 GB</td>
</tr>
<tr>
<td>Oct 2007</td>
<td>2332</td>
<td>5806</td>
<td>227348</td>
<td>1153835</td>
<td>62.10 GB</td>
</tr>
<tr>
<td>Nov 2007</td>
<td>2432</td>
<td>5603</td>
<td>261769</td>
<td>959337</td>
<td>106.57 GB</td>
</tr>
<tr>
<td>Dec 2007</td>
<td>1611</td>
<td>3597</td>
<td>136144</td>
<td>536012</td>
<td>56.94 GB</td>
</tr>
<tr>
<td>Total</td>
<td>12968</td>
<td>30169</td>
<td>1304825</td>
<td>5112515</td>
<td>4100.37 GB</td>
</tr>
</tbody>
</table>
The Grand Challenge!
Eliminate Malaria from the Planet

Bill and Melinda Gates Call for New Global Commitment to Chart a Course for Malaria Eradication

New resources and scientific progress help pave the way toward malaria eradication

U.S. presidential candidates urged to sustain and expand President’s Malaria Initiative

SEATTLE -- Bill and Melinda Gates today called on global leaders to embrace "an audacious goal—to reach a day when no human being has malaria, and no mosquito on earth is carrying it." They delivered the call to action at a forum of 300 leading malaria scientists and policymakers from around the world.

"Advances in science and medicine, promising research, and the rising concern of people around the world represent an opportunity to make a difference," said Gates. "Our goal is to make access to diagnostic and treatment a reality for all those who need it, and to ensure that the future is a time without malaria."