

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

Gregory R. Madey, Jenica Abrudan, Robert Bruggner, Ryan  
E. Butler, Scott Christley, Ryan C. Kennedy, Nathan A.  
Konopiniski, Neil F. Lobo, Eric O. Stinson, Matthew J. Van  
Antwerp, Maria F. Unger, Ying Zhou & Frank H. Collins

Department of Biological Sciences

Department of Computer Science & Engineering

***Workshop on Interdisciplinary Biomedical Research***

***University of Notre Dame***

***April 10 - 11, 2008***

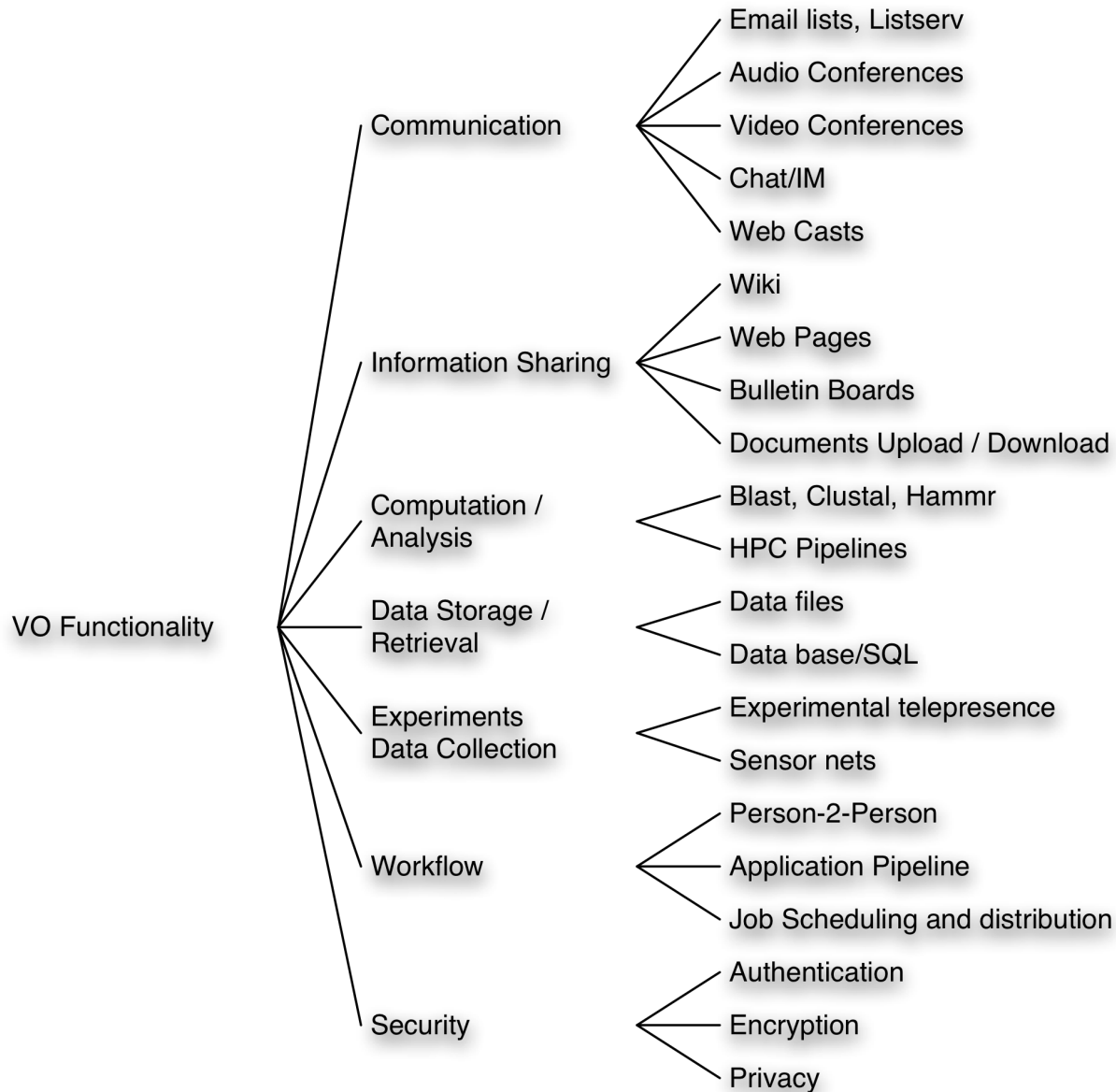
# Interdisciplinary Collaboration

- 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)

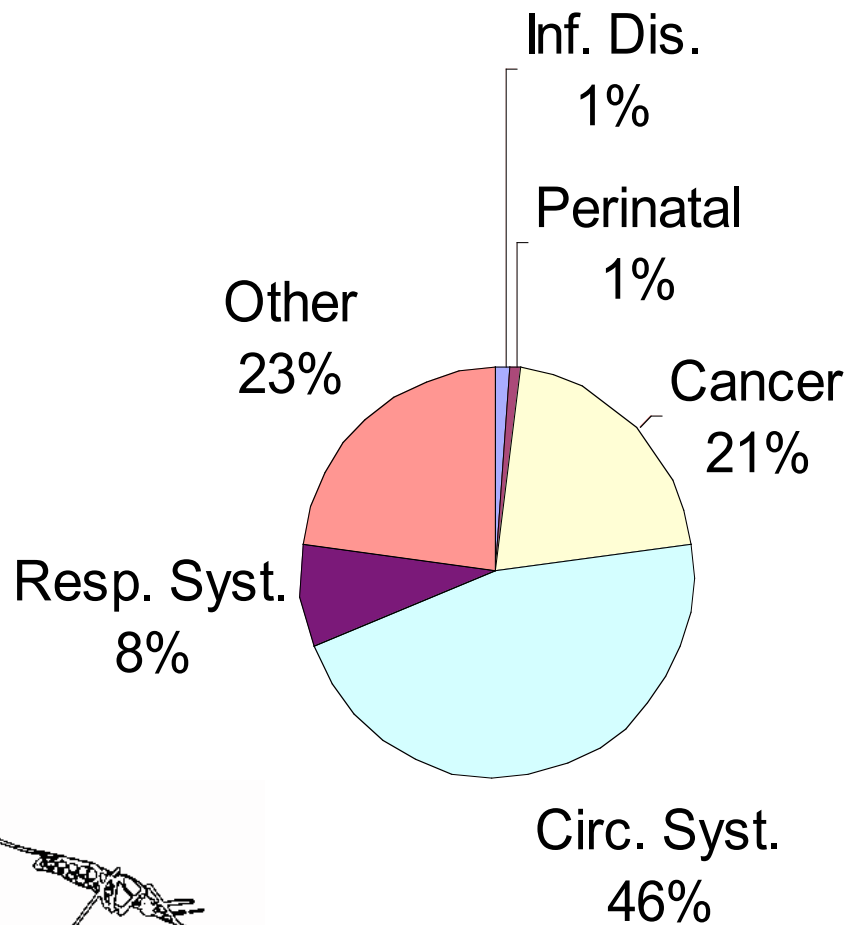
# Cyberinfrastructure / VO Functionality



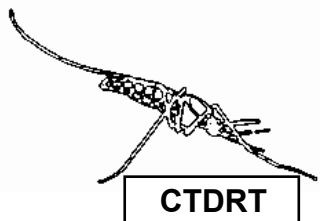
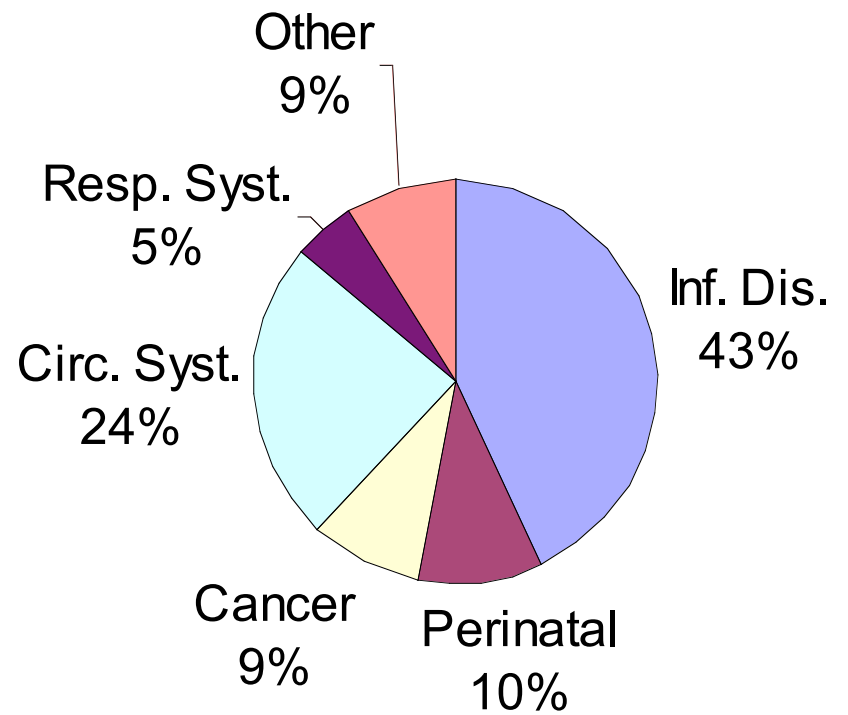


# CAUSES OF DEATH

## Developed World

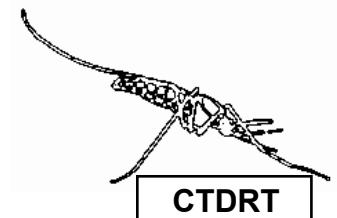


## Developing World



## Pathogen-Specific Causes of Death by Infectious Diseases (1997)

	Annual Mortality	Prevalence
<b>TB</b>	<b>2,900,000</b>	<b>16,300,000</b>
<b>AIDS</b>	<b>2,300,000</b>	<b>&gt;30,000,000</b>
<b>Malaria</b>	<b>~2,000,000</b>	<b>&gt;500,000,000</b>
<b>Arboviruses</b>	<b>200,000</b>	<b>-----</b>
<b>African Trypanosomiasis</b>	<b>100,000</b>	<b>400,000</b>
<b>Leishmaniasis</b>	<b>80,000</b>	<b>12,000,000</b>
<b>American Trypanosomiasis</b>	<b>45,000</b>	<b>18,000,000</b>
<b>Filariasis + Onchocerciasis</b>	<b>-----</b>	<b>140,000,000</b>



# Malaria Transmission Consortium (MTC)

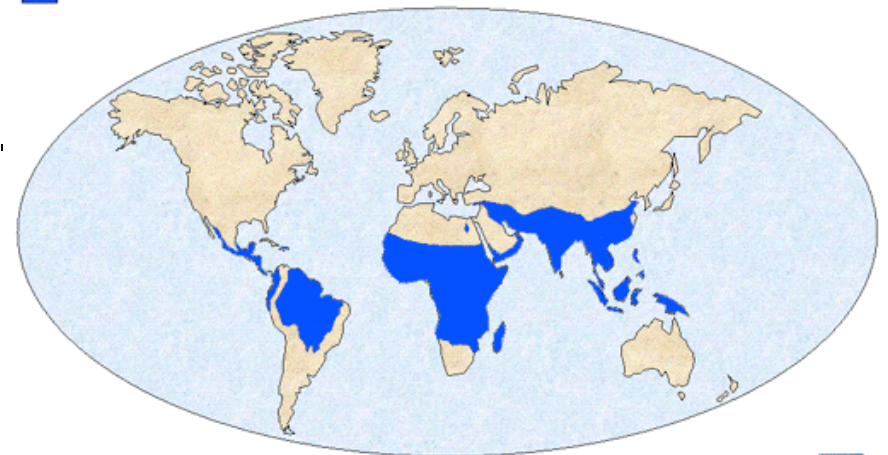
“The University of Notre Dame has received a \$20 million grant from the Bill & Melinda Gates Foundation to support research aimed at evaluating existing malaria control programs and designing improved methods for malaria control.... This study represents multiple institutions in multiple countries, 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.



■ Distribution of Malaria



# 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

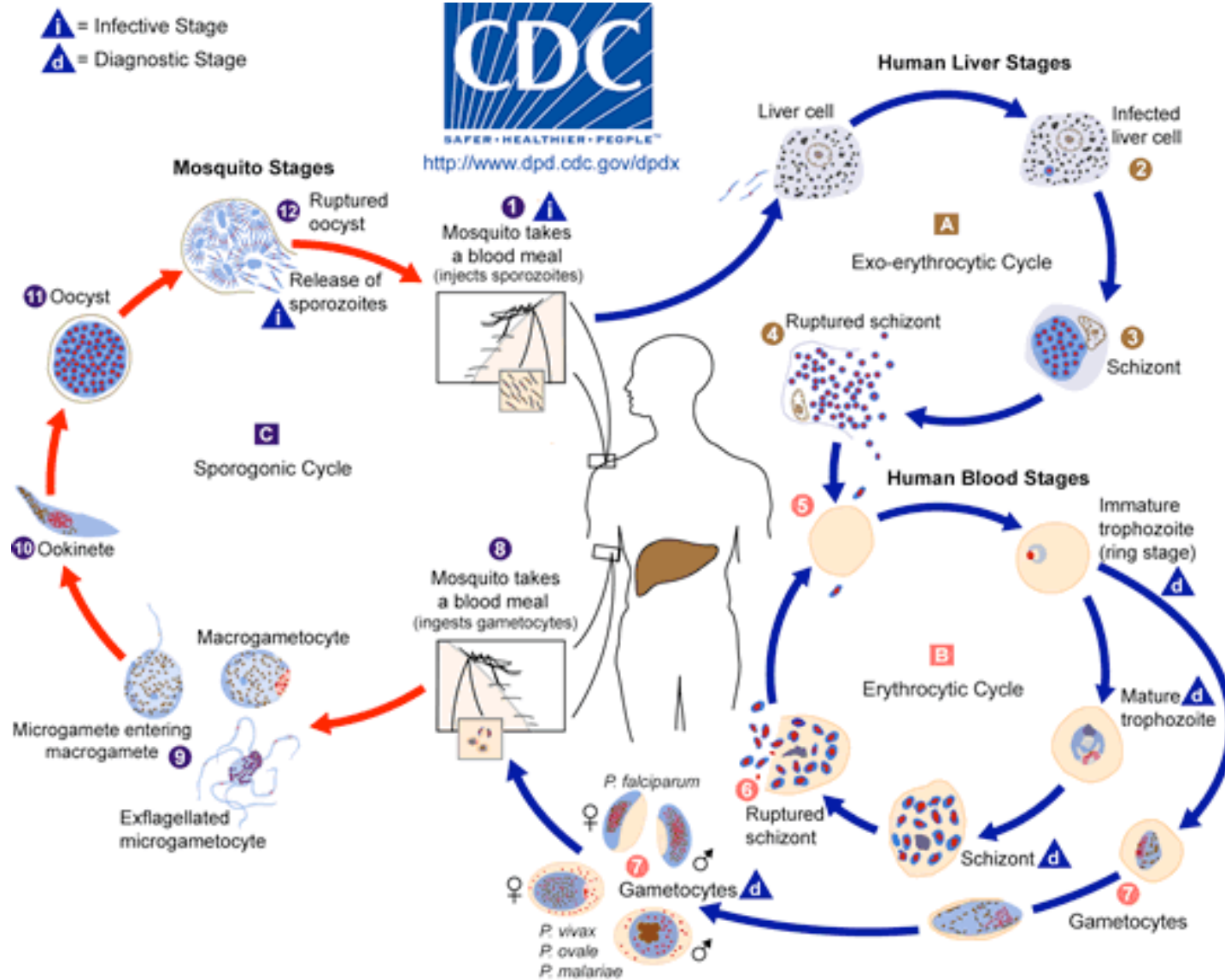
# MTC Participants

- 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





# Mosquito Feeding Cycle

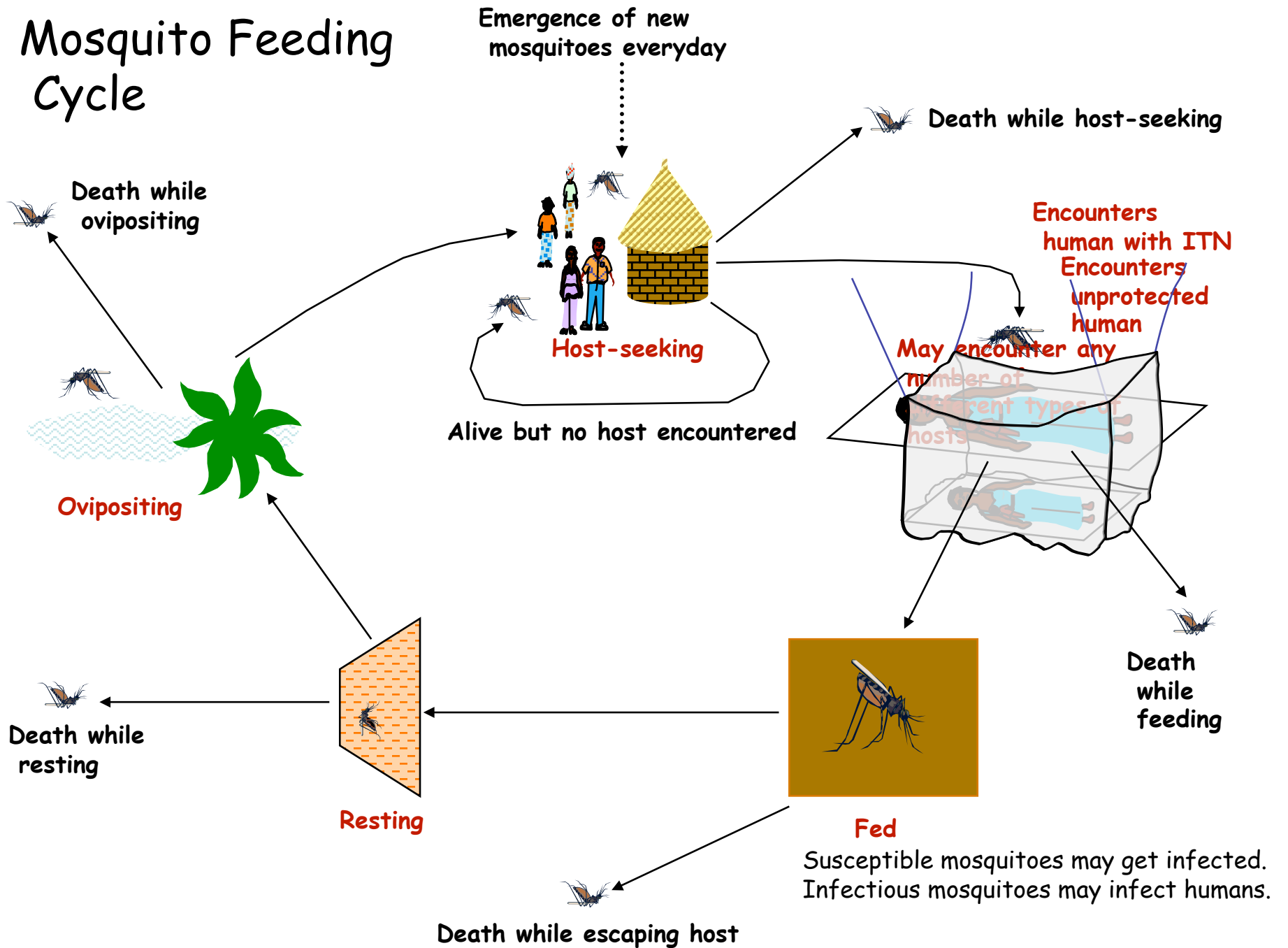


Figure from Paul Libiszowski

# Mosquito Feeding Cycle

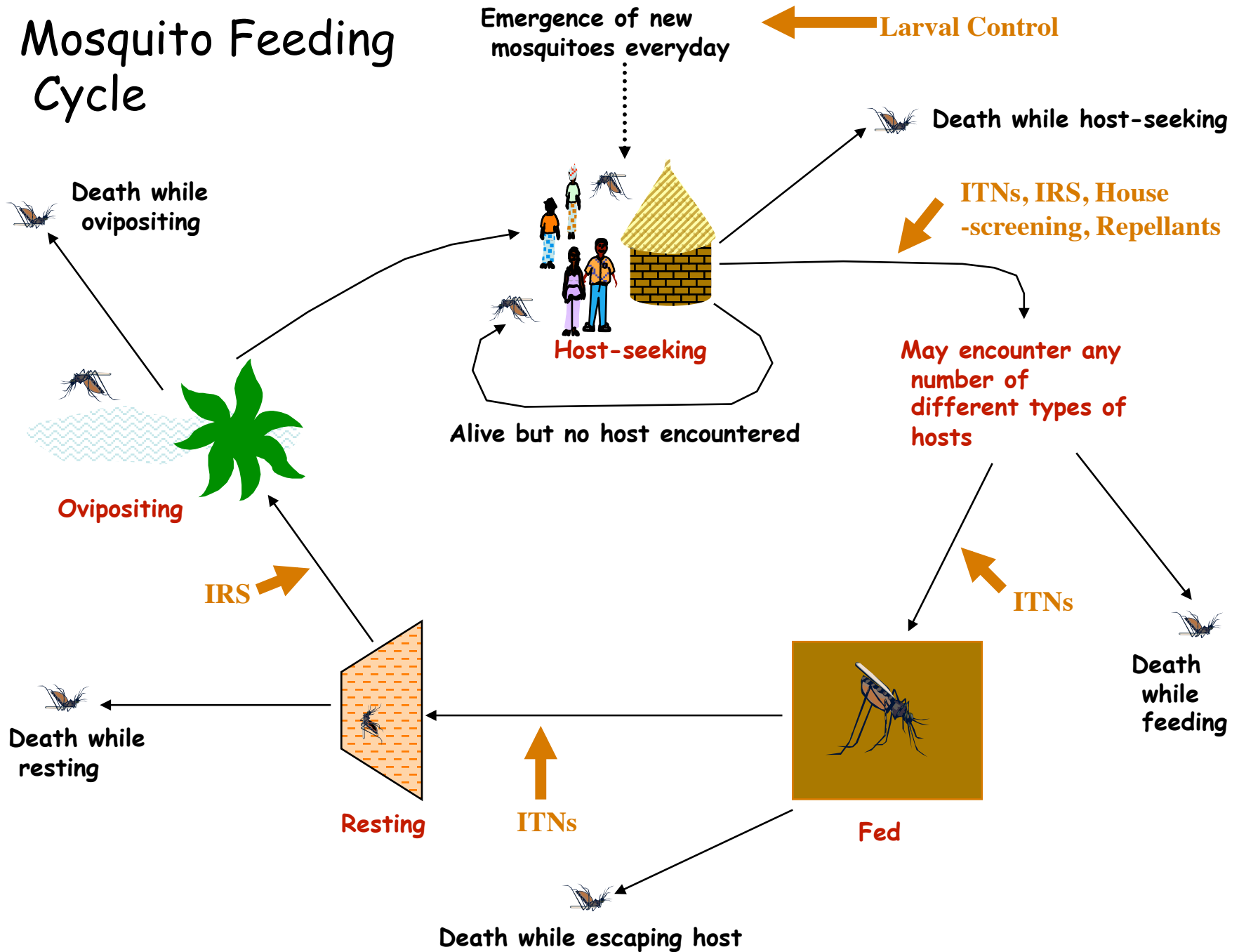
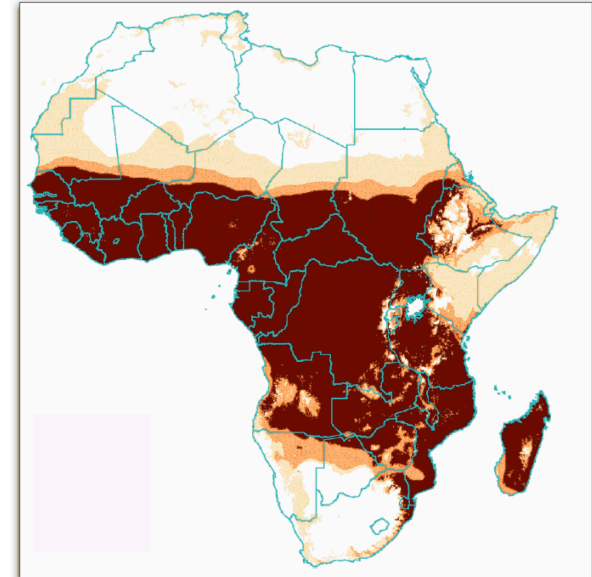


Figure from Paul Libiszowski

# Activities

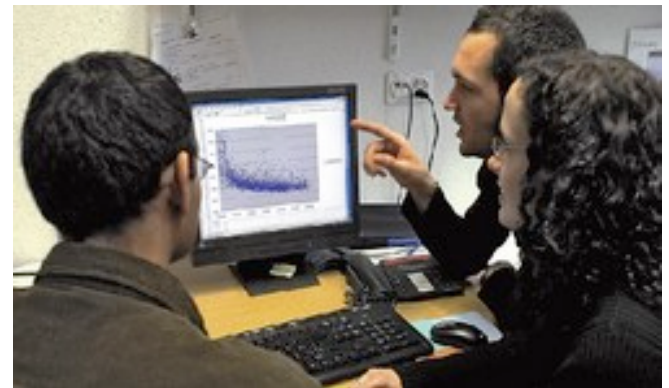
- Database design, web interface, deployment
- Geographical Information Systems (GIS)



- 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

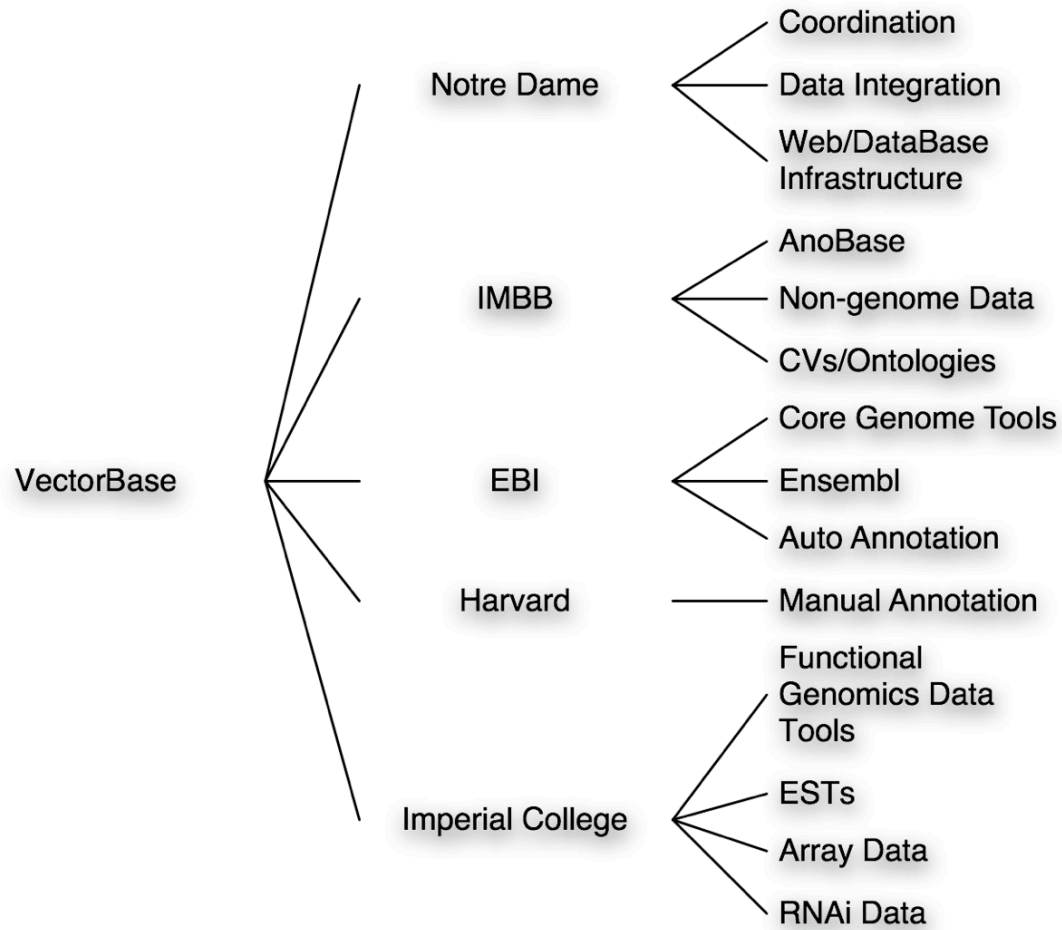
- PI - Frank Collins, Biological Sciences & CSE, ND
- Co-PI's/Subcontractors
  - EMBL/Imperial College (Fotis Kafatos)
  - EBI Hixton, Cambridge (Ewan Birney)
  - Institute of Molecular Biology, Crete, (Christos Louis)
  - Harvard (William Gelbart)
  - CSE Notre Dame (Greg Madey)



# 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

# Activities





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 scapularis*: 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

## Welcome To VectorBase!

An NIAID Bioinformatics Resource Center for Invertebrate Vectors of Human Pathogens.



### VectorBase Services:

#### Organisms »

- [A. gambiae](#), [A. aegypti](#), [I. scapularis](#), [C. pipiens](#), [P. humanus](#), [Additional Organisms](#)

#### Data Sections »

[Sequence Data](#), [Gene Expression Data](#), [Images](#)

#### Tools »

- [BLAST](#), [ClustaW](#), [HMMER](#), [Controlled Vocabulary Search](#), [Manual Annotation Submission Tool](#)

#### Retrieve Data »

[Download Data](#), [BioMart](#)

#### VectorBase Documentation »

[Help Wiki](#), [FAQ](#), [BRC Information](#), [Mailing Lists](#)

#### Additional Resources »

[Links](#), [VectorBase Publications](#), [Genome White Papers](#)



[C. pipiens CpipJ1.1 gene build & genome browser released](#)

### Recent News:

**Mar 14th, 2008:**

[Ixodes scapularis genome sequence available »](#)

**Feb 19th, 2008:**

[Ixodes electronic mailing list available »](#)

**Feb 19th, 2008:**

[Culex pipiens gene set now in GenBank »](#)

**Feb 19th, 2008:**

[BioMart back online »](#)

### VectorBase Quick Search

### Contact VectorBase



Need help or have feedback? Please contact us at [info@vectorbase.org](mailto:info@vectorbase.org)

### VectorBase Newsletter



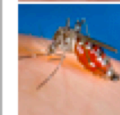
January 2008 - #2

[More Information](#)

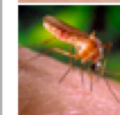
### Current Genome Releases



[A. gambiae PEST](#)  
AgamP3.4, June 2007



[A. aegypti Liverpool](#)  
AaegL1.1, August 2007



[C. pipiens JHB](#)  
CpipJ1.1, October 2007

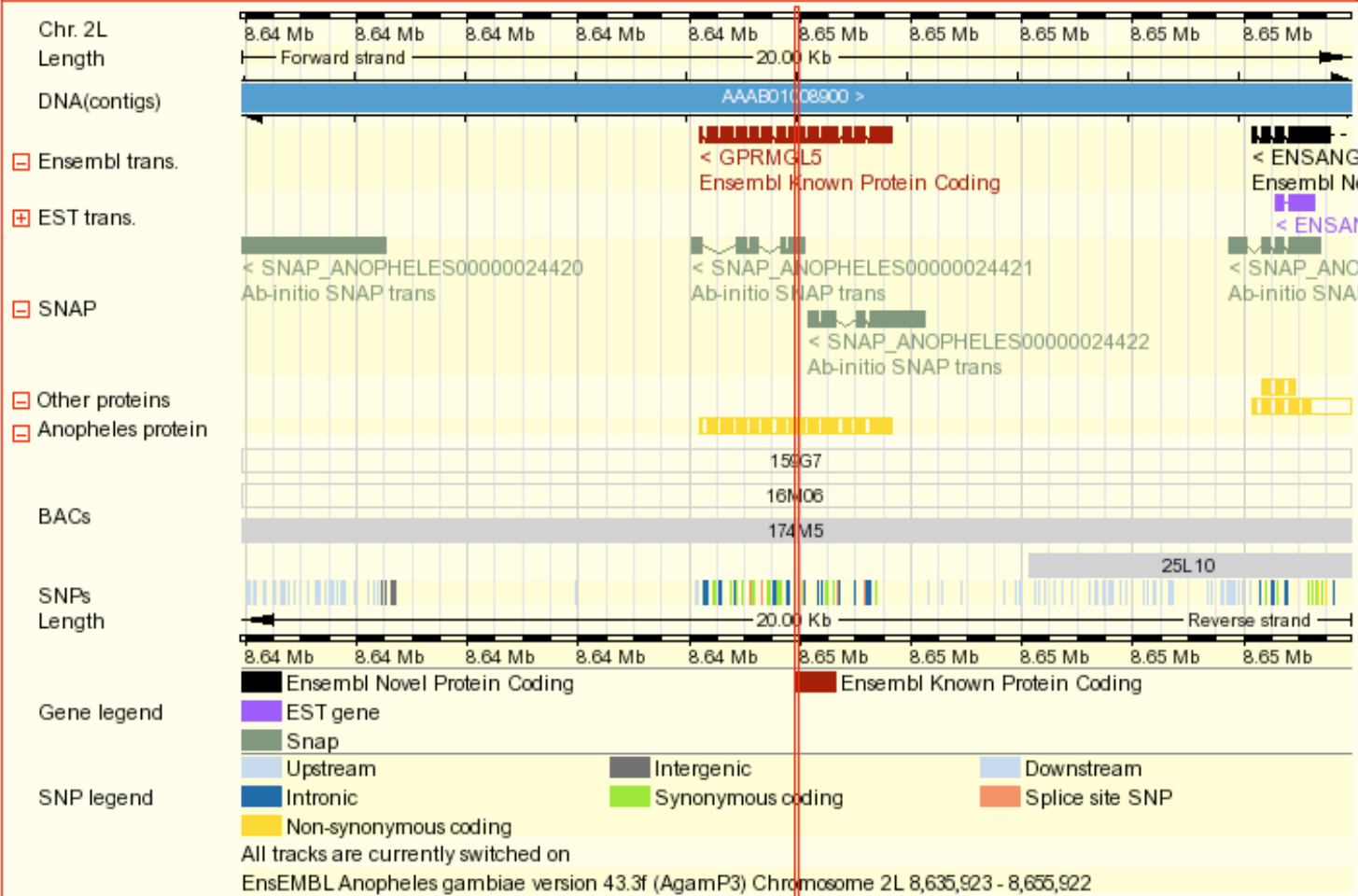
Detailed view

Features ▾ Comparative ▾ DAS Sources ▾ Repeats ▾ Decorations ▾ Export ▾ Image size ▾ Help ▾

Jump to region 2L : 8635923 - 8655922 Refresh Band: Refresh

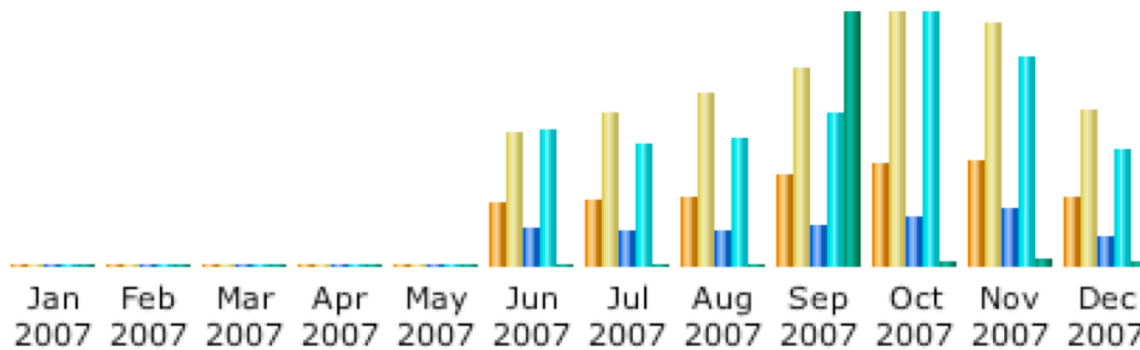
Zoom

<< 5MB < 2MB < 1MB < Window Window > 1MB > 2MB > 5MB >>





## Monthly history



Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2007	0	0	0	0	0
Feb 2007	0	0	0	0	0
Mar 2007	0	0	0	0	0
Apr 2007	0	0	0	0	0
May 2007	6	25	2311	5142	39.75 MB
Jun 2007	1434	3081	177721	618770	26.00 GB
Jul 2007	1486	3530	158258	560213	29.67 GB
Aug 2007	1600	3978	156540	586051	34.06 GB
Sep 2007	2067	4549	184734	693155	3785.00 GB
Oct 2007	2332	5806	227348	1153835	62.10 GB
Nov 2007	2432	5603	261769	959337	106.57 GB
Dec 2007	1611	3597	136144	536012	56.94 GB
<b>Total</b>	<b>12968</b>	<b>30169</b>	<b>1304825</b>	<b>5112515</b>	<b>4100.37 GB</b>

# The Grand Challenge!

## Eliminate Malaria from the Planet

October 17, 2007

### 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