FaceBase: Interaction, Visualization, Collaboration, and Computation

Harry Hochheiser

for

the FaceBase Hub Technical Team

U. Pittsburgh Department of Biomedical Informatics

harryh@pitt.edu

www.dbmi.pitt.edu
www.facebase.org

My perspective...

- Computer Science

  Human-Computer Interaction

  Information Visualization

- Building Collaborative Systems

Information Visualization

- " The use of computer-supported, interactive, visual representations of abstract data to amplify cognition" Card, Mackinlay, Shneiderman 1999

- 'abstract' concepts

- No "given" representation

- Support annotation and search tasks

Visualization:
A Picture is worth a thousand words

- Visual bandwidth is huge – more so than for other senses

- Fast, parallel

- Good pattern recognition

- Pre-attentive processing

Appropriate visual representations of data can reveal structure, aid cognition, and facilitate development of understanding.

Harry Hochheiser, harryh@pitt.edu

Information Visualization
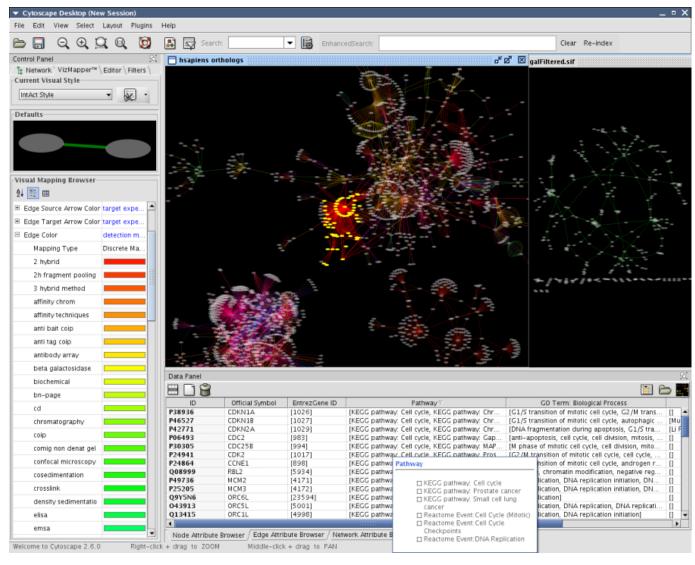
- Visualization +Interactivity

- Rapid, incremental, reversible queries

- Overview → Detail

- Multiple coordinated views

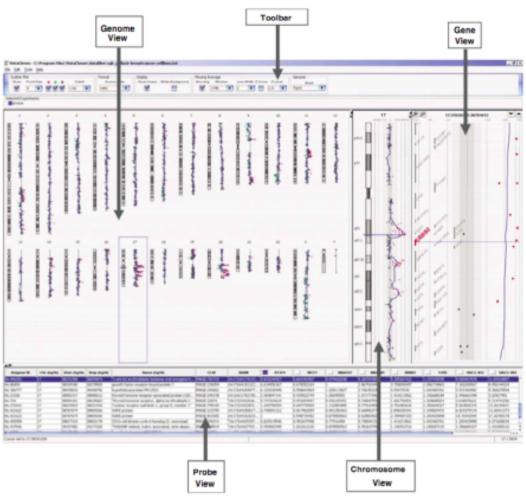 - Alternative representations of data

- Searching, Browsing, Exploration

Networks, Interactions, pathways – Cytoscape www.cytoscape.org

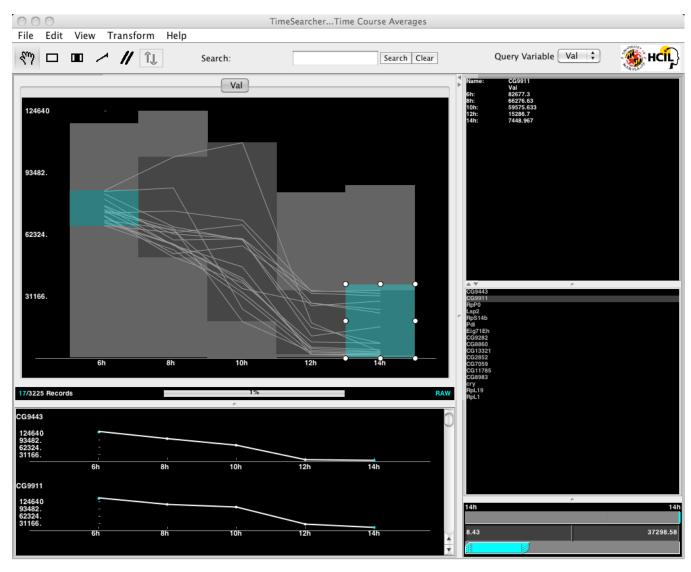# ArrayCGH –
copy number variation – VistaChrom

(Kincaid, et al. 2005)

TimeSearcher & Microarray Data

InfoVis & HCI for FaceBase

- Can we improve

    - Searching and browsing?

    - Understanding of connections between datasets?

    - Integration of diverse data types?

What's Different about FaceBase?

- Range of Data Types

    - Integration of diverse data types: sequences, images, genotypes, phenotypes....

    - Aggressive data sharing plan

- Collaboration Support?

    - Explicit interactions in several projects

    - Similar overlaps in interests → new collaborations?

Beaty

Gene X Environment
Interactions

GWA Studies/
Genotypes

Donahue/Murray develop mouse strains

Marazita/
Weinberg

Morphometry

Genetic Maps

*Imaging: developmental atlases
for mouse & chick.*

Normative image
Database

Cranial Neural Crest
Expression + Images

LCM/FACS
microRNA –
gene atlas

Fraser

Spritz

Images

Shapiro/Satya
provide algorithms/
tools

Chai

Potter

Clouthier/Artinger/
Postlethwait

RNAi Zebrafish/Mouse

Expression
Data

Beaty

Visel

GWAS results for
CNV/Linkage
Analysis

Enhancers,
ChIP-seq, RNA-seq

Sequencing

Human Data

Mouse

Zebrafish

Harry Hochheiser, harryh@pitt.edu

Organizational/Socio-Technical Challenges

- Gary & Judith Olson U.C. Irvine (formerly U. Michigan)

- Science of Collaboratories

    - 5- year, NSF-funded

    - scienceofcollaboratories.org

- **Scientific Collaboration on the Internet** Olson, Zimmerman, & Bos, eds., MIT Press 2008

"From Shared Databases to Communities of Practice"

Bos, et al., in Olson, et al.

- 7 types of collaboratories
  - Shared Instrument
  - <u>Community Data System</u>
  - <u>Open Community Contribution System</u>
    - contributions come in the form of work rather than data
  - Virtual Community of Practice
  - Virtual Learning Community
  - <u>Distributed Research Center</u>
    - "like a university research center, but at a distance. Distributed Research Centers are unified by a topic of interest and joint projects in that area".

  ## Which is FaceBase?

"From Shared Databases to Communities of Practice"

Bos, et al., in Olson, et al.

	Community Data System	**Open Community Contribution System**	**Distributed Research Center**
Technology Issues	Data formats Modeling and visualization	Interoperability, ease of use Data formats	All issues faced by others, including data formats Workplace awareness
Organizational Issues	Motivating contributors Large-scale decision-making	Motivating contributors Quality control	All issues faced by others Cross-institutional IP Career issues for younger investigators?

A Theory of Remote Scientific Collaboration (TORSC)

J. Olson, et al., in G. Olson, et al.

- *How do we demonstrate success?*

- Measures of success

  - *Impact on science* – discoveries, papers, artifacts, research quality, etc.

  - *Science Careers* – diversity, breadth of participation, tenure, quality of life & satisfaction of researchers

  - *Inspiration to others* – new collaboratories, new software

  - *Funding and public perception* – funding renewal, public interest, more funding for collaboratories

  - *Tool use* – Development and demonstrated use/reuse of production-quality software

TORSC: Factors that lead to success

- ## Nature of the work – tightly coupled or not

- ## Common ground
  - *Mutual knowledge* – past experience and terminology
    - Common vocabulary is key
    - Mappings? Zebrafish anatomy ontology has no entry for palate
  - *Beliefs and Assumptions in Management* – hierarchical vs. informal...

Mouse BIRN as a "mutual knowledge" success story

- "Atlas" shows relationship between terms

- Spatial layout – point to areas of interest without using terms

- Use of consensus terminology may not be necessary

  - If there are ways to map between related term

- Ontologies possible, but resource-intensive

TORSC: Factors that lead to success

J. Olson, et al., in G. Olson, et al.

- ## Common ground, continued…

  - *Beliefs and Assumptions in Management* – hierarchical vs. informal…
    - Common style -> aligned interactions and expectations

- ## Collaboration Readiness

  - *Work-related Dimensions:* techie vs. bench scientist tensions

  - *Social dimensions:* do people like working together?
    - Trust?
    - Collective efficacy? - feel empowered to make changes?

TORSC: Factors that lead to success

J. Olson, et al., in G. Olson, et al.

- Management, Planning , and Decision Making

  - *Time and Attention:* Are participants dedicating enough time?

  - *Management*: Is there a good management structure that includes all players?

  - *Communication and Possibilities for Redirection* – is there appropriate dialog that can be used to support changes in focus if needed?

  - *Knowledge Management, Decision Making, Institutional Issues*

- Technological Readiness: "the key is to understand the real needs of the end users, not to push 'cool' technologies on people."

TORSC: Underlying Themes

- Collaboratory success depends upon commitment from involved parties

- It's not just a matter of writing code

- Without meaningful commitment, it won't work

- Use TORSC to see if the necessary support and commitment is there...

TORSC & FaceBase

- Judith & Gary Olson have developed an online assessment of collaboration readiness

- FaceBase to pilot this tool – provide a baseline measurement of readiness

- Please cooperate

  - Give us a list of folks in your labs/institutions who are participating

  - Get these folks to complete the survey (~ 20 minutes)

- Use this to demonstrate value to funders