Wash U Human Epigenome Browser

Xin Zhou
Washington University in St. Louis

GMOD meeting, April 5-6
“data flood”
Challenge with visualizing big public data sets

Encyclopedia of DNA Elements  >3000

Roadmap Epigenomics Project  >1500

... and so many more ...
Challenge with visualizing big public data sets

* Metadata is critical
  - assay type
  - sample
  - processing methods

and so many more ...

Encyclopedia of DNA Elements

>3000

>1500

ROADMAP epigenomics project
Challenge with visualizing big public data sets

* Metadata is critical
  - assay type
  - sample
  - processing methods

* Web display solution is nice

A lab's own data,
Array/sequencing data sets
Files amount to 1TB
Challenge with visualizing big public data sets

* Integrate with public data sets
* Web display solution is nice

* Metadata is critical
  - assay type
  - sample
  - processing methods

A lab’s own data, Array/sequencing data sets
Files amount to 1TB

and so many more ...

Encyclopedia of DNA Elements
>3000

ROADMAP epigenomics project
>1500
Our solution
Brief history

2007-2009 Ting Wang and “UCSC Cancer Genomics Browser”
Brief history

2007-2009 Ting Wang and “UCSC Cancer Genomics Browser”

2010 June Xin joined Ting's new lab @ Wash U
Start of new genome browser development
for human Roadmap Epigenomics Project
Brief history

2007-2009 Ting Wang and “UCSC Cancer Genomics Browser”

2010 June Xin joined Ting's new lab @ Wash U
     Start of new genome browser development
     for human Roadmap Epigenomics Project

2010 Nov Xin's training at Santa Cruz
2007-2009 Ting Wang and “UCSC Cancer Genomics Browser”

2010 June Xin joined Ting's new lab @ Wash U
Start of new genome browser development for human Roadmap Epigenomics Project

2010 Nov Xin's training at Santa Cruz

2011 June “Wash U Human Epigenome Browser” went online
“VizHub” for Roadmap project
Brief history

2007-2009 Ting Wang and “UCSC Cancer Genomics Browser”

2010 June Xin joined Ting’s new lab @ Wash U
Start of new genome browser development
for human Roadmap Epigenomics Project

2010 Nov Xin’s training at Santa Cruz

2011 June “Wash U Human Epigenome Browser” went online
“VizHub” for Roadmap project

2011 Dec Paper accepted by Nature Methods
“big data sets” visualization

- “Need for Speed”
  - BigWig, bigBed
  - Parallel programming
“big data sets” visualization

- “Need for Speed”
  - BigWig, bigBed
  - Parallel programming
- Ajax and client-side rendering
  - Configurable rendering style
“big data sets” visualization

- “Need for Speed”
  - BigWig, bigBed
  - Parallel programming
- Ajax and client-side rendering
  - Configurable rendering style
- Metadata vocabulary
  - Metadata color map
  - Facet browsing
“big data sets” visualization

- “Need for Speed”
  - BigWig, bigBed
  - Parallel programming
- Ajax and client-side rendering
  - Configurable rendering style
- Metadata vocabulary
  - Metadata color map
  - Facet browsing
- Custom track
Visualization and analysis

Bells and whistles

- Genomic juxtaposition
- Gene Set View
- Genome Snapshot
Visualization and analysis

Gene Set View

- start at gene/region: CYP4Z1
- stop at gene/region: CYP2C8
- One pixel spans 350 base pairs
Visualization and analysis

Genome Snapshot
Future work
Acknowledgement

Thank you!

Pamela A.F. Madden, PhD

Current Position
Professor of Psychiatry

Education and Training
B.S., Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania, 1978
M.S., Child Development/Child Care, University of Pittsburgh, Pittsburgh, Pennsylvania, 1983
Ph.D., Psychology in Education, University of Pittsburgh, Pittsburgh, Pennsylvania, 1988

Ting Wang
twang@genetics.wustl.edu
Principle Investigator
Assistant Professor, Department of Genetics

Thank you!