### modMine... modMine http://intermine.modencode.org - is a data warehouse that integrates modENCODE results with data from FlyBase, WormBase and other sources. Accessing modENCODE data - provides easy navigation through modENCODE experiments to view and export data that have been submitted and vetted in the project. - includes a library of pre-defined template queries to mine modENCODE - allows you to create and analyse lists of data and upload your own lists - is based on the InterMine data warehousing system. Links to summary of all Summary Genomic Distributions of Histone Modifications experiment experiments and data generatedall m about experiment view Access data generated by each submission Export features from all submissions in GFF3, Keyword Search: csv or other formats ordered faceted results Genome Region Search find features in user assigned genomic regions View all submissions for this experiment - each GBrowse Ideogram submission varies Interactive exploratory experimental conditions ew an expression score heatmap interface of the regulatory All the available data regions files can be reached with Expression heatmap for a View and sort table by list of selected genes Explore an hierarchical view of the summary table antibodies, cell lines, development stages, etc Analyze Lists of Data Use Template Queries Enter a list of identifiers submission Description of the submission Submission : seq-AB1791\_H3\_N2\_L3 Experiment Type: ChIP-sea and its location in the context of the modENCODE project. view Detailed information about the submission, with list of experimental factors and line Powered by to more specific data Features can be exported in feature view Single click queries find overlapping and nearby genes for features generated Report page with detailed by the submission. information about every GBrowse image of the feature. View GBrowse tracks and relevant feature and link to Browse all get raw data files from the the complete track protocols used in Web Service API modMine includes a RESTful web service API allowing you to access data directly from Links to relevant data, including the list of features overlapping the programs you write in Perl, Java, Python or Ruby. given feature. See http://intermine.org/wiki/WebServiceQuickStart

# modENCODE

http://www.modencode.org help@modencode.org

Jelena Aleksic<sup>1</sup>, Daniela Butano<sup>1</sup>, Sergio Contrino<sup>1</sup>, Fengyuan Hu<sup>1</sup>, Ellen Kephart<sup>2</sup>, Paul Lloyd<sup>2</sup>, Rachel Lyne<sup>1</sup>, Marc Perry<sup>3</sup>, Peter Ruzanov<sup>3</sup>, Richard Smith<sup>1</sup>, Radek Stepan<sup>1</sup>, Nicole Washington<sup>2</sup>, Quang Thrin<sup>3</sup>, Suzanna Lewis<sup>2</sup>, Gos Micklem<sup>1</sup>, Lincoln Stein<sup>3</sup>.

University of Cambridge, UK, <sup>2</sup> Lawrence Berkeley National Laboratory, <sup>3</sup> Ontario Institute for Cancer Research

- a comprehensive **encyclopedia of genomic functional elements** in the model organisms C. elegans and D. melanogaster
- a consortium of 11 primary projects, divided between worm and fly, spanning the various domains (see data categories).
- extensive collection of metadata, enabling fine-grained control over presentation

- a **portal** (www.modencode.org) to access the data.

## - Faceted dataset search

- Sophisticated data mining
- GBrowse tracks
- Cloud instance
- Traditional ftp site
- Extensive web service API
- Gene networks interactive displays
- Expression heatmaps
- Project wiki
- Blog
- Growing set of explanatory documents
- Web seminars
- Help desk

### **Data Categories**

Chromatin structure 4 experiments 50 data submissions
Copy Number Variation 3 experiments 9 data submissions
Copy Structure 19

Gene Structure 18 experiments 128 data submissions
Histone modification and replacement 5 experiments 254 data

Origins of Replication 1 experiment 3 data submissions

Other chromatin binding sites 6 experiments 284 data submissions

Replication Factors
Replication Timing
2 experiments 16 data submissions
2 experiments 4 data submissions

RNA expression profiling 26 experiments 682 data submissions

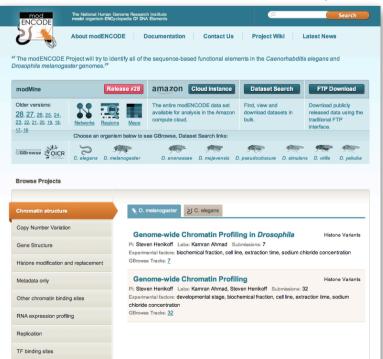
TF binding sites 3 experiments 390 data submissions

## GBrowse

modENCODE portal

Data are shown as a series of **graphical tracks** spanning the worm and fly genomes, alongside core datasets from WormBase and FlyBase. GBrowse is able to show **correlations** among diverse modENCODE data sets. For example, the stage-specific binding of a transcription factor can be visually correlated with stage-specific changes in that factor's putative target genes.

GBrowse can visualize both **continuous quantitative data**, such as ChIP signal plots, and **discrete features** such as gene models and inferred binding sites. When appropriate, the two types of data can be shown simultaneously, for example called transcription factor binding sites superimposed on top of the underlying signal data.



## **Dataset Search**

http://data.modencode.org

The tool make possible to quickly reach the dataset of interest using various **filtering criteria**: organism, category, target element, technique, etc.

It then allows to **download** the dataset, **browse** the relevant GBrowse tracks, **view** it in modMine.

