

modENCODE data and tools for the community

www.modencode.org

Gos Micklem
University of Cambridge



BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY



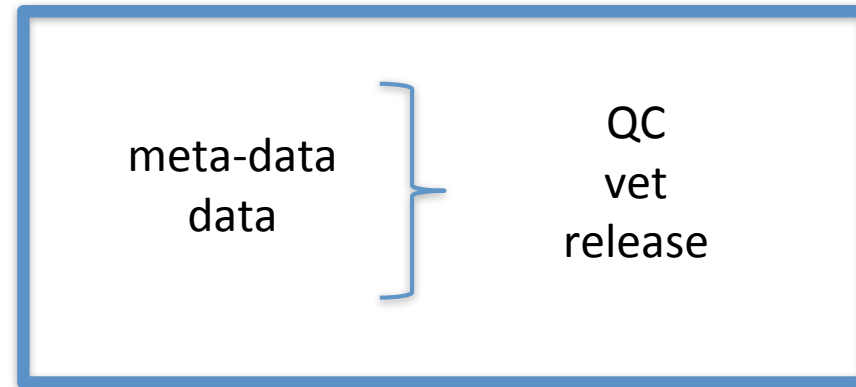
**UNIVERSITY OF
CAMBRIDGE**

modENCODE DCC Data Flow

modENCODE DCC
data wranglers



modENCODE DCC pipeline



submit data &
meta-data



Faceted Browser

Dataset	Organism	Technique	Target Element	Assay Factor	Conditions	PI	ID
3-prime-UTR>All stages,integrated-gene-model	C. elegans	integrated-gene-model	mRNA	3-prime-UTR	All stages	Piano, F.	2745
3-prime-UTR.dat	C. elegans	RACE	mRNA	3-prime-UTR	Dauer Larvae	Piano, F.	2334
11(m47)Dauer Larvae;20 degree Celsius;RACE	C. elegans	RACE	mRNA	3-prime-UTR	20 degree Celsius	Piano, F.	2464
3-prime-UTR.dat	C. elegans	RACE	mRNA	3-prime-UTR	Dauer Larvae	Piano, F.	2335
11(m170)Dauer Larvae;20 degree Celsius;RACE	C. elegans	RACE	mRNA	3-prime-UTR	20 degree Celsius	Piano, F.	2461
3-prime-UTR.dat	C. elegans	RACE	mRNA	3-prime-UTR	Dauer Larvae	Piano, F.	2339
71(m172)Dauer Larvae;20 degree Celsius;RACE	C. elegans	RACE	mRNA	3-prime-UTR	20 degree Celsius	Piano, F.	2462
3-prime-UTR.dat	C. elegans	RACE	mRNA	3-prime-UTR	Dauer Larvae	Piano, F.	2337
8(m540)Dauer Larvae;20 degree Celsius;RACE	C. elegans	RACE	mRNA	3-prime-UTR	20 degree Celsius	Piano, F.	2463

data.modencode.org
ftp.modencode.org

modMine

intermine.modencode.org

Amazon/Bionimbus

www.bionimbus.org

modENCODE Data Volume

2317 of 3763 datasets released: ~6 TB

Final freeze: expect ~20-25 TB altogether

modENCODE Data Volume

2317 of 3763 datasets released: ~6 TB

Final freeze: expect ~20-25 TB altogether

Post-laptop era

Nuisance to download

GEO/SRA (crude), WormBase/ FlyBase (refined)

Amazon/ BioNimbus (all)



“The National Human Genome Research Institute (NHGRI) model organism ENCyclopedia Of DNA Elements (modENCODE) Project will try to identify all of the sequence-based functional elements in the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes.”

modMine release #25	amazon Cloud instance	Dataset Search Datasets	FTP Download Get FTP data
Older versions: 25 , 24 , 23 , 22 , 21 , 20 , 19 , 18 , 17 , 16	The entire modENCODE data set available for analysis in the Amazon compute cloud.	Find, view and download datasets in bulk.	Download publicly released data using the traditional FTP interface.
Choose an organism below to see GBrowse, Dataset Search links:			
	<i>C. elegans</i>	<i>D. melanogaster</i>	<i>D. ananassae</i>
		<i>D. mojavensis</i>	<i>D. pseudoobscura</i>
		<i>D. simulans</i>	<i>D. virilis</i>
		<i>D. yakuba</i>	

Browse Projects

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

D. melanogaster

C. elegans

Genome-wide Chromatin Profiling

Histone Variants

Pi: **Steven Henikoff** Labs: **Kamran Ahmad, Steven Henikoff** Submissions: **32**

Experimental factors: **developmental stage, biochemical fraction, cell line, extraction time, sodium chloride concentration**

GBrowse Tracks: [32](#)

Internet Explorer users: If this page is too slow for you, consider Firefox, Safari or Chrome. Please do not abort script execution.

Faceted Browser: data.modencode.org

No datasets selected

Search:

Search Filters

Organism

- 792 D. melanogaster
- 312 C. elegans
- 4 D. mojavensis

Project Category

- 203 Gene Structure
- 203 RNA expression profiling
- 164 TF binding sites
- 171 Other chromatin binding sites
 - 43 Chromatin structure
- 311 Histone modification and replacement
 - 9 Copy Number Variation
- 19 Replication

Genomic Target Element

Technique

1119 data sets

Dataset ▲

- 3-prime-UTR;daf-11(m47);Dauer Larva;20 degree celsius;RACE
- 3-prime-UTR;daf-2(e1370);Dauer Larva;20 degree celsius;RACE
- 3-prime-UTR;daf-7(e1372);Dauer Larva;20 degree celsius;RACE
- 3-prime-UTR;daf-9(m540);Dauer Larva;20 degree celsius;RACE
- 3-prime-UTR;him-8(e1489);Adult Male;post-L1 stage larvae;20 degree celsius;RACE
- 3-prime-UTR;N2;Adult 70 hr post-L1 larvae;20 degree celsius;RACE
- 3-prime-UTR;N2;L1 8 hr post-L1 stage larvae;20 degree celsius;RACE

Internet Explorer users: If this page is too slow for you, consider Firefox, Safari or Chrome. Please do not abort script execution.

No datasets selected

Search:

Search Filters

Organism 1

- 792 D. melanogaster
- 312 C. elegans
- 4 D. mojavensis

Project Category

- 131 Gene Structure
- 128 RNA expression profiling
- 87 TF binding sites
- 143 Other chromatin binding sites
- 29 Chromatin structure
- 248 Histone modification and replacement
 - 9 Copy Number Variation
- 19 Replication

Genomic Target Element

Technique

792 data sets filtered from 1119

Data

5-prime-UTR;Y cn bw sp;Mixed

AGO2;Oregon-R;Larvae 3rd in

AGO2;S2-DRSC;Late Embryo cell-line;ChIP-chip

ASH1;ML-DmBG3-c2;Larvae 3 cell-line;ChIP-chip

ASH1;S2-DRSC;Late Embryo cell-line;ChIP-chip

bab-1;Embryos 0-12 hr;embryo

BEAF-32;Embryos 0-12 hr;Ch

Internet Explorer users: If this page is too slow for you, consider Firefox, Safari or Chrome. Please do not abort script execution.

No datasets selected

Search:

Search Filters

Organism 1

- 87 **D. melanogaster**
- 77 C. elegans

Project Category 1

- 131 Gene Structure
- 128 RNA expression profiling
- 87 **TF binding sites**
- 143 Other chromatin binding sites
- 29 Chromatin structure
- 248 Histone modification and replacement
- 9 Copy Number Variation
- 19 Replication

Genomic Target Element

Technique

87 data sets filtered from 1119 originally (Re...

Dataset ▲

- bab-1;Embryos 0-12 hr;embryo;ChIP-chip
- bonus;Embryos 16-24 hr;ChIP-seq
- brahma;Y cn bw sp;Pupae;ChIP-chip
- Brakeless;Embryos 0-12 hr;embryo;ChIP-
- Brakeless;Embryos 0-4 hr;ChIP-chip
- cad;7T-CAD;Adult Female 3days;ChIP-se
- cad;7T-CAD;Adult Female;ChIP-seq
- cad;7T-CAD;Embryos 0-4 hr;ChIP-chip
- cad;7T-CAD;Embryos 0-4 hr;ChIP-seq
- cad;7T-CAD;Embryos 4-8 hr;ChIP-seq
- cad;7T-CAD;Larvae L3 stage;ChIP-seq

Internet Explorer users: If this page is too slow for you, consider Firefox, Safari or Chrome. Please do not abort script execution.

No datasets selected

Search:

Search Filters

Organism

87 **D. melanogaster**

77 C. elegans

Project Category

131 Gene Structure

128 RNA expression profiling

87 **TF binding sites**

143 Other chromatin binding sites

29 Chromatin structure

248 Histone modification and replacement

9 Copy Number Variation

19 Replication

Genomic Target Element

Technique

Search Filters

Organism

Project Category

Genomic Target Element

Technique

Principal Investigator

Assay Factor

Developmental Stage

Strain

Cell Line

Tissue

Compound

Temperature

ally (Re

IP-chip

hip

o;ChIP-

ip

ChIP-se

eq

chip

-seq

-seq

cad:ZT-CAD:Larvae L3 stage:ChIP-seq

Internet Explorer users: If this page is too slow for you, consider Firefox, Safari or Chrome. Please do not abort script execution.

- 2569: Brakeless;Embryos 0-4 hr;ChIP-chip (*D. melanogaster*)
- 609: Brakeless;Embryos 0-12 hr;embryo;ChIP-chip (*D. melanogaster*)
- 628: bab-1;Embryos 0-12 hr;embryo;ChIP-chip (*D. melanogaster*)

Search:

[Browse Worm Tracks](#)
[Browse Fly Tracks](#)
[Download](#)
[View in ModMine](#)
[List Cloud Files](#)
[Clear All](#)

Search Filters

 Organism 1

- 87 *D. melanogaster*
- 77 *C. elegans*

 Project Category 1

- 131 Gene Structure
- 128 RNA expression profiling
- 87 TF binding sites
- 143 Other chromatin binding sites
- 29 Chromatin structure
- 248 Histone modification and replacement
- 9 Copy Number Variation
- 19 Replication

 Genomic Target Element

 Technique

87 data sets filtered from 1119 originally ([Reset All Filters](#))

 Showing items 1-20, Max items:
[Next 20>>](#)

<input type="checkbox"/> Dataset	Organism	Technique	Target Element	Assay Factor	Conditions	PI	ID
<input checked="" type="checkbox"/> bab-1;Embryos 0-12 hr;embryo;ChIP-chip	<i>D. melanogaster</i>	ChIP-chip	Transcriptional Factor	bab-1	Embryos 0-12 hr	White, K.	628
<input type="checkbox"/> bonus;Embryos 16-24 hr;ChIP-seq	<i>D. melanogaster</i>	ChIP-seq	Transcriptional Factor	bonus	Embryos 16-24 hr	White, K.	3393
<input type="checkbox"/> brahma;Y cn bw sp;Pupae;ChIP-chip	<i>D. melanogaster</i>	ChIP-chip	Transcriptional Factor	brahma	Pupae	White, K.	602
<input checked="" type="checkbox"/> Brakeless;Embryos 0-12 hr;embryo;ChIP-chip	<i>D. melanogaster</i>	ChIP-chip	Transcriptional Factor	Brakeless	Embryos 0-12 hr	White, K.	609
<input checked="" type="checkbox"/> Brakeless;Embryos 0-4 hr;ChIP-chip	<i>D. melanogaster</i>	ChIP-chip	Transcriptional Factor	Brakeless	Embryos 0-4 hr	White, K.	2569
<input type="checkbox"/> cad;7T-CAD;Adult Female 3days;ChIP-seq	<i>D. melanogaster</i>	ChIP-seq	Transcriptional Factor	cad	Adult Female 3days	White, K.	3231
<input type="checkbox"/> cad;7T-CAD;Adult Female;ChIP-seq	<i>D. melanogaster</i>	ChIP-seq	Transcriptional Factor	cad	Adult Female	White, K.	2625
<input type="checkbox"/> cad;7T-CAD;Embryos 0-4 hr;ChIP-chip	<i>D. melanogaster</i>	ChIP-chip	Transcriptional Factor	cad	Embryos 0-4 hr	White, K.	902
<input type="checkbox"/> cad;7T-CAD;Embryos 0-4 hr;ChIP-seq	<i>D. melanogaster</i>	ChIP-seq	Transcriptional Factor	cad	Embryos 0-4 hr	White, K.	2626
<input type="checkbox"/> cad;7T-CAD;Embryos 4-8 hr;ChIP-seq	<i>D. melanogaster</i>	ChIP-seq	Transcriptional Factor	cad	Embryos 4-8 hr	White, K.	2637
<input type="checkbox"/> cad;7T-CAD;Larvae L3 stage;ChIP-seq	<i>D. melanogaster</i>	ChIP-seq	Transcriptional Factor	cad	Larvae L3 stage	White, K.	3403



Search modENCODE Data Sets

Internet Explorer users: If this page is too slow for you, consider Firefox, Safari or Chrome. Please do not abort script execution.

20 Selected Datasets:

- 2437: AMA-1;N2;Fed L1 stage larvae;20 degree celsius;ChIP-seq (*C. elegans*)
- 2435: AMA-1;N2;Early Embryos;20 degree celsius;ChIP-seq (*C. elegans*)
- 343: AMA-1;Mixed Embryos;20 degree celsius;ChIP-chip (*C. elegans*)
- 33: LY-2;OP217;Larvae L3 stage;20 degree celsius;ChIP-seq (*C. elegans*)

Search:

Search Filters

Organism

1

921 *D. melanogaster*

440 *C. elegans*

8 *D. pseudoobscura*

Project Category

155 Gene Structure

65 RNA expression profiling

103 TF binding sites

29 Other chromatin binding sites

14 Chromatin structure

76 Histone modification and replacement

Genomic Target Element

440 data sets filtered from 1384 originally ([Reset All Filters](#))

Showing items 1-20, Max items:

[Next 20>>](#)

<input checked="" type="checkbox"/> Dataset	Organism	Technique	Target Element	Assay Factor	Conditions	PI	ID
<input checked="" type="checkbox"/> 3-prime-UTR;All stages;integrated-gene-model	<i>C. elegans</i>	integrated-gene-model	mRNA	3-prime-UTR	All stages	Piano, F.	2745
<input checked="" type="checkbox"/> 3-prime-UTR;daf-11(m47);Dauer Larvae;20 degree celsius;RACE	<i>C. elegans</i>	RACE	mRNA	3-prime-UTR	Dauer Larvae 20 degree celsius	Piano, F.	2334 and 2464
<input checked="" type="checkbox"/> 3-prime-UTR;daf-2(e1370);Dauer Larvae;20 degree celsius;RACE	<i>C. elegans</i>	RACE	mRNA	3-prime-UTR	Dauer Larvae 20 degree celsius	Piano, F.	2335 and 2461
<input checked="" type="checkbox"/> 3-prime-UTR;daf-7(e1372);Dauer Larvae;20 degree celsius;RACE	<i>C. elegans</i>	RACE	mRNA	3-prime-UTR	Dauer Larvae 20 degree celsius	Piano, F.	2336 and 2462
<input checked="" type="checkbox"/> 3-prime-UTR;daf-9(m540);Dauer Larvae;20 degree celsius;RACE	<i>C. elegans</i>	RACE	mRNA	3-prime-UTR	Dauer Larvae 20 degree celsius	Piano, F.	2337 and 2463



“The National Human Genome Research Institute (NHGRI) model organism ENCyclopedia Of DNA Elements (modENCODE) Project will try to identify all of the sequence-based functional elements in the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes.”

modMine release #25	amazon Cloud instance	Dataset Search	Datasets	FTP Download	Get FTP data			
Older versions: 25 , 24 , 23 , 22 , 21 , 20 , 19 , 18 , 17 , 16	The entire modENCODE data set available for analysis in the Amazon compute cloud.	Find, view and download datasets in bulk.		Download publicly released data using the traditional FTP interface.				
Choose an organism below to see GBrowse, Dataset Search links:								
	<i>C. elegans</i>	<i>D. melanogaster</i>	<i>D. ananassae</i>	<i>D. mojavensis</i>	<i>D. pseudoobscura</i>	<i>D. simulans</i>	<i>D. virilis</i>	<i>D. yakuba</i>

Browse Projects

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

D. melanogaster

C. elegans

Genome-wide Chromatin Profiling

Histone Variants






















Pi: Steven Henikoff Labs: Kamran Ahmad, Steven Henikoff Submissions: 32

Experimental factors: developmental stage, biochemical fraction, cell line, extraction time, sodium chloride concentration









GBrowse Tracks: [32](#)

ftp.modencode.org

Index of /

Name	Size	Date Modified
 C.elegans/		12/20/11 8:26:00 PM
 D.ananassae/		12/20/11 8:26:00 PM
 D.melanogaster/		12/20/11 8:26:00 PM
 D.mojavensis/		12/20/11 8:26:00 PM
 D.pseudoobscura/		12/20/11 8:26:00 PM
 D.simulans/		12/20/11 8:26:00 PM
 D.virilis/		12/20/11 8:26:00 PM
 D.yakuba/		12/20/11 8:26:00 PM
 DATA_SNAPSHOTS.txt	1.9 kB	2/1/12 5:50:00 PM
 DATA_SNAPSHOTS.txt~	2.1 kB	2/1/12 5:49:00 PM
 MANIFEST.txt 	3.3 MB	12/20/11 8:26:00 PM
 README	8.7 kB	11/28/11 7:24:00 PM
 README.submitters	206 B	9/6/11 12:00:00 AM
 WITHDRAWN_STUDIES.txt	820 B	12/8/11 7:21:00 PM
 all_files/		12/9/11 4:31:00 PM
 metadata.csv 	5.6 MB	12/20/11 8:26:00 PM
 modENCODE_publication_policy.pdf	50.5 kB	8/9/10 12:00:00 AM
 publications/		12/11/11 1:27:00 PM
 robots.txt	26 B	10/18/11 12:00:00 AM

Index of /C.elegans/

Name	Size	Date Modified
 [parent directory]		
 Chromatin-Structure/		12/20/11 8:26:00 PM
 Histone-Modification/		12/20/11 8:26:00 PM
 Non-TF-Chromatin-binding-factor/		12/20/11 8:26:00 PM
 Transcriptional-Factor/ 		12/20/11 8:26:00 PM
 mRNA/		12/20/11 8:26:00 PM
 small-RNA/		12/20/11 8:26:00 PM

Index of /C.elegans/Transcriptional-Factor/

Name	Size	Date Modified
 [parent directory]		
 ChIP-seq/ 		12/20/11 8:26:00 PM

Index of /C.elegans/Transcriptional-Factor/ChIP-seq/

Name	Size	Date Modified
 [parent directory]		
 computed-peaks_gff3/		12/20/11 8:26:00 PM
 coverage-graph_wiggle/		12/20/11 8:26:00 PM
 gene-model_gff3/		12/20/11 8:26:00 PM
 raw-seqfile_fastq/		12/20/11 8:26:00 PM



“The National Human Genome Research Institute (NHGRI) model organism ENCyclopedia Of DNA Elements (modENCODE) Project will try to identify all of the sequence-based functional elements in the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes.”

modMine release #25	amazon Cloud instance	Dataset Search Datasets	FTP Download Get FTP data
Older versions: 25 , 24 , 23 , 22 , 21 , 20 , 19 , 18 , 17 , 16	The entire modENCODE data set available for analysis in the Amazon compute cloud.	Find, view and download datasets in bulk.	Download publicly released data using the traditional FTP interface.
Choose an organism below to see GBrowse, Dataset Search links:			
	C. elegans	D. melanogaster	D. ananassae
	D. mojavensis	D. pseudoobscura	D. simulans
		D. virilis	D. yakuba

Browse Projects

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

D. melanogaster

C. elegans

Genome-wide Chromatin Profiling

Histone Variants

Pi: Steven Henikoff Labs: Kamran Ahmad, Steven Henikoff Submissions: 32

Experimental factors: developmental stage, biochemical fraction, cell line, extraction time, sodium chloride concentration

GBrowse Tracks: [32](#)

<< [Back to Browser](#)

- Search for Specific Tracks

Search: Stick to top when scrolled

Clear

Search

- Tracks

- Chromatin Structure** ●
- DNA replication**
- Gene Structure**
- RNA Expression Profiling**
- Transcription Factors**
- non-modENCODE Reference Data**
- Analysis** *All on* *All off*

<< [Back to Browser](#)


- Search for Specific Tracks

Search: Stick to top when scrolled

- Tracks

- Chromatin Structure

Chromatin Proteins

Histone Modifications 

Histone Modifying Enzymes (Grouped by Cell Line) All on All off

Nucleosome Structure All on All off

DNA replication

Gene Structure

RNA Expression Profiling

Transcription Factors

non-modENCODE Reference Data

Analysis All on All off

<< [Back to Browser](#)

- Search for Specific Tracks

Search: Stick to top when scrolled

Clear

Search

- Tracks

- Chromatin Structure

+ Chromatin Proteins

- Histone Modifications

+ ChIP-chip (Grouped by Cell Line) All on All off

+ ChIP-Seq All on All off

+ ChIP-chip All on All off

+ Histone Modifying Enzymes (Grouped by Cell Line) All on All off

+ Nucleosome Structure All on All off

+ DNA replication

+ Gene Structure

+ RNA Expression Profiling

+ Transcription Factors

+ non-modENCODE Reference Data

+ Analysis All on All off

<< [Back to Browser](#)

Search for Specific Tracks

Search: Stick to top when scrolled

Tracks

Chromatin Structure

Chromatin Proteins

Histone Modifications

ChIP-chip (Grouped by Cell Line) All on All off

ChIP-Seq All on All off

ChIP-Seq: Histone Modifications in Adult Female [4 of 4 subtracks selected]

ChIP-Seq: Histo
subtracks selected

ChIP-Seq: Histone Modifications in Adult Male [5 of 5 subtracks selected]

ChIP-Seq: Histo
subtracks selected

ChIP-Seq: Histone Modifications in embryo 0-4hr [6 of 6 subtracks selected]

ChIP-Seq: Histo
subtracks selected

ChIP-Seq: Histone Modifications in embryo 12-16hr [6 of 6 subtracks selected]

ChIP-Seq: Histo
subtracks selected

ChIP-chip All on All off

Histone Modifying Enzymes (Grouped by Cell Line) All on All off

Nucleosome Structure All on All off

DNA replication

Gene Structure

RNA Expression Profiling

Transcription Factors

non-modENCODE Reference Data

Analysis All on All off

<< [Back to Browser](#)

Search for Specific Tracks

Search: Stick to top when scrolled

Clear

Search

Tracks

Chromatin Structure

Chromatin Proteins

Histone Modifications

ChIP-chip (Grouped by Cell Line) All on All off

ChIP-Seq All on All off

ChIP-Seq: Histone Modifications in Adult Female [4 of 4 subtracks selected]

ChIP-Seq: Histone Modifications in Adult Male [5 of 5 subtracks selected]

ChIP-Seq: Adult Female H3K4Me1 ChIP-seq (White project) Details
Staged Drosophila (y; bw cn sp) adult... [Click for more](#)

ChIP-Seq: Histone Modifications in Embryo [12 of 12 subtracks selected]

ChIP-chip All on All off

Histone Modifying Enzymes (Grouped by Cell Line) All on All off

Nucleosome Structure All on All off

DNA replication

Gene Structure

RNA Expression Profiling

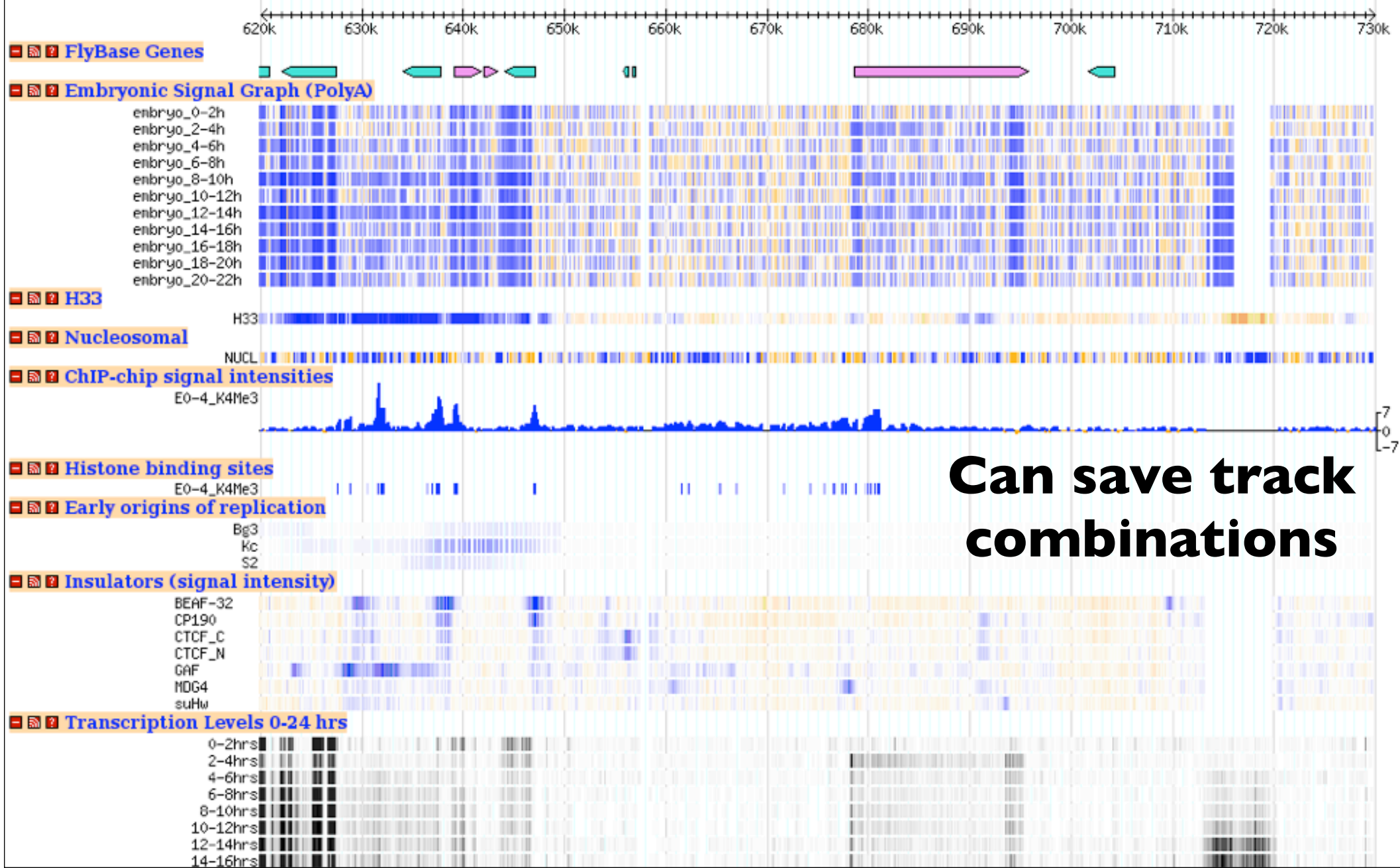
Transcription Factors

non-modENCODE Reference Data

Analysis All on All off



GBrowse



Can save track combinations



“The National Human Genome Research Institute (NHGRI) model organism ENCYclopedia Of DNA Elements (modENCODE) Project will try to identify all of the sequence-based functional elements in the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes.”

modMine

release
#25

amazon

Cloud instance

Dataset Search

Datasets

FTP Download

Get FTP data

Older versions:

[25](#), [24](#), [23](#), [22](#), [21](#),
[20](#), [19](#), [18](#), [17](#), [16](#)

The entire modENCODE data set available for analysis in the Amazon compute cloud.

Find, view and download datasets in bulk.

Download publicly released data using the traditional FTP interface.

Choose an organism below to see GBrowse, Dataset Search links:



C. elegans



D. melanogaster



D. ananassae



D. mojavensis



D. pseudoobscura



D. simulans



D. virilis



D. yakuba

Browse Projects

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

D. melanogaster

C. elegans

Genome-wide Chromatin Profiling

Histone Variants

Pi: Steven Henikoff Labs: Kamran Ahmad, Steven Henikoff Submissions: 32

Experimental factors: developmental stage, biochemical fraction, cell line, extraction time, sodium chloride concentration

GBrowse Tracks: [32](#)

Browse all modENCODE data

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

RNA expression profiling

Replication

www.modmine.org

Chromatin structure



Nucleosome mapping

Project Chromatin Function - Tison Lab | Lab: Tison Lab

2 GBrowse tracks
1 entry in GEO



Search

Enter names, identifiers or keywords for genes, proteins, pathways, ontology terms, etc. (e.g. **zen**, **pha-4**, **DNA binding**).

Or search **modENCODE** experiments by type, lab name, antibody, etc. (e.g. **RNA-seq**, **Snyder**, **CP190**).



Genomic Region Search

Explore a genomic region for features found by the **modENCODE** project.



[Genomic Region Search](#)

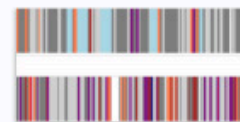
Fly Gene Expression



View an expression score heatmap for any list of fly genes. See an [example](#).

To upload your own list of genes, use the form above or go to the 'Lists' tab and click on '**Upload**' to create and name the new list. You can also use any of modMine's queries to create a list.

Fly Chromatin states



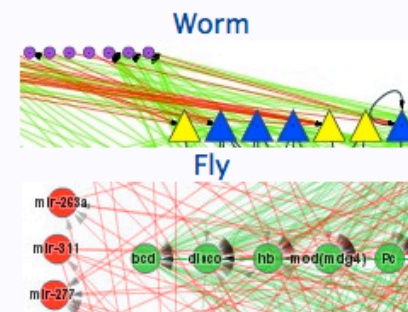
GBrowse Ideograms



Park Lab Viewer

Includes folded view and also data about DHS, TSS, replication, etc.

Regulatory Network



Explore an hierarchical view of the physical regulatory networks.



Upload Lists

Enter a list of identifiers.

CG9151, FBgn0000099, CG3629, TfiIB, Mad, CG1775, CG2262, TWIST_DROME, tinman, runt, E2f,

[advanced](#)



Use Template Queries

Get started with **powerful queries** using our predefined searches. These customizable templates have been designed around common tasks performed by our biologist community. To see how they work, why not try a template from our [examples page](#)?



Analyze Lists of Data

Explore and **Analyze**. Upload lists of identifiers to use in queries and discover relationships in our analysis widgets. See an [example](#).



Browse all modENCODE data

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

RNA expression profiling

Replication

www.modmine.org

Chromatin structure



Nucleosome mapping

Project Chromatin Experiment - 1 track 1 field - 1 label - 1 view 1 table

2 GBrowse tracks
1 entry in GEO

Search

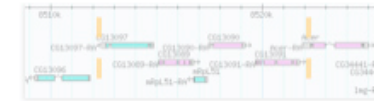
Enter names, identifiers or keywords for genes, proteins, pathways, ontology terms, etc. (e.g. **zen**, **pha-4**, **DNA binding**).

Or search **modENCODE** experiments by type, lab name, antibody, etc. (e.g. **RNA-seq**, **Snyder**, **CP190**).



Genomic Region Search

Explore a genomic region for features found by the **modENCODE** project.



[Genomic Region Search](#)

Fly Gene Expression



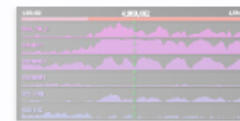
View an expression score heatmap for any list of fly genes. See an [example](#).

To upload your own list of genes, use the form above or go to the 'Lists' tab and click on 'Upload' to create and name the new list. You can also use any of modMine's queries to create a list.

Fly Chromatin states



GBrowse Ideograms



Park Lab Viewer

Includes folded view and also data about DHS, TSS, replication, etc.

Regulatory Network



Explore an hierarchical view of the physical regulatory networks.



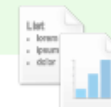
Upload Lists

Enter a list of identifiers.



Use Template Queries


Get started with **powerful queries** using our predefined searches. These customizable templates have been designed around common tasks performed by our biologist community. To see how they work, why not try a template from our [examples page](#)?



Analyze Lists of Data

Explore and Analyze. Upload lists of identifiers to use in queries and discover relationships in our analysis widgets. See an [example](#).



 **Search**

Enter names, identifiers or keywords for genes, proteins, pathways, ontology terms, etc. (e.g. **zen**, **pha-4**, **DNA binding**).

Or search **modENCODE** experiments by type, lab name, antibody, etc. (e.g. **RNA-seq**, **Snyder**, **CP190**).

- Antibody names: **PolII, H3K4me1, CP190**
- Lab names: **Reinke, Snyder**
- Combine terms with AND/AND NOT: **fly AND embryo**

Hits by Category

Submission: 72

Hits by Experiment

Chromatin Binding Site Mapping: 47

ChIP-Seq Identification of *C. elegans*

TF Binding Sites: 22

RNA-seq support of the ChIP data: 3

Type	Details
Submission	Snyder LIN-11_GFP_L2 modE <i>design: binding_site_identificati</i> <i>experimentDate: March 2, 2009</i> <i>publicReleaseDate: November 27, 2</i> <i>embargoDate: August 16, 2010</i>
Submission	Snyder UNC-130_GFP_L1 mo <i>design: binding_site_identificati</i> <i>experimentDate: March 2, 2009</i> <i>publicReleaseDate: November 27, 2</i> <i>embargoDate: August 16, 2010</i>
Submission	Snyder LIN-39_GFP_L3 modE <i>design: binding_site_identificati</i> <i>experimentDate: March 2, 2009</i> <i>publicReleaseDate: November 27, 2</i> <i>embargoDate: August 16, 2010</i>



Summary for selected Submission[↔ Link](#)

title H3K18ac.S2
design binding_site_identification_design
DCCid modENCODE_292
experimentDate May 15, 2008
publicReleaseDate October 11, 2009
embargoDate November 12, 2009

description We aim to determine the locations of 125 chromosomal proteins and histone modifications across the Drosophila melanogaster genome. The... [\[View all\]](#)

experimentType ChIP-chip
qualityControl biological_replicate
replicate biological_replicate

antibodies 1 Antibody [\[details...\]](#)

Class	name	hostOrganism	
Antibody	H3K18ac	Rabbit	[details...]

[\[show in table...\]](#)

arrays 1 Array [\[details...\]](#)

cellLines 1 CellLine [\[details...\]](#)

Class	name	tissue	species	
CellLine	S2-DRSC	embryo-derived cell-line	D. melanogaster	[details...]

[\[show in table...\]](#)

databaseRecords 4 DatabaseRecord

developmentalStages 1 DevelopmentalStage [\[details...\]](#)

Class	name	
DevelopmentalStage	late embryonic stage	[details...]

[\[show in table...\]](#)

experiment 1 Experiment [\[details...\]](#)

experimentalFactors 2 ExperimentalFactor

Class	type	name	
ExperimentalFactor	cell line	S2-DRSC	[details...]
ExperimentalFactor	antibody	H3K18ac	[details...]

[\[show in table...\]](#)**Further information for this Submission****Lab:** Sarah Elgin – Washington University

Project: Chromosomal Proteins – Karpen

Organism: D. melanogaster

Experiment: [Genomic Distributions of Histone Modifications](#)**Submission description**

We aim to determine the locations of 125 chromosomal proteins and histone modifications across the Drosophila melanogaster genome. The proteins and modifications under study are involved in basic chromosomal functions such as DNA replication, gene expression, gene silencing, and inheritance. We will perform Chromatin Immunoprecipitation (ChIP) using genomic tiling arrays. We will initially assay localizations using chromatin from three cell lines and two embryonic stages, and will then extend the analysis of a subset of proteins to four additional animal tissues/stages.

Features generated by this submission:

Feature type	View data	Export
BindingSite <small>def</small>	11924	TAB DELIMITED GFF3 SEQUENCE

Find overlapping features:

Find which overlap the features generated by this submission.

Find nearby genes:

Find Genes that have features generated by this submission located within

GBrowse tracks: [Karpen_HISMODS_S2/H3K18Ac.S2](#)

Data files: [repset.3506196.smoothedM.v](#)
[repset.3506196.smoothedM.c](#)
[36.Mvalues.wig](#)
[80.Mvalues.wig](#)
[80.pirrota_230_S12_H3K18Ac](#)
[85.pirrota_242_S12_Input_](#)
[36.pirrota_188_A3H3K18Ac.C](#)
[38.pirrota_190_A3_Input.CB](#)

[Download ALL data files](#)**Lists**

Summary for selected Submission

title H3K18ac.S2
design binding_site_identification_design
DCCid 292
experimentDate 15 May 2008
publicReleaseDate 11 October 2009
description We aim to determe the locations of 125 chromosomal proteins and histone modifications across the Drosophila melanogaster genome. The... [View all]
embargoDate Thu Nov 12 00:00:00 GMT 2009
experimentType ChIP-chip
qualityControl biological_
replicate biological_

antibodies

arrays

Class	name
Array	Affymetrix Drosophila Tiling Arrays v2.0R
Array	Affymetrix Drosophila Tiling Arrays v2.0R

[show in table...]

cellLines

Class	name	ti
CellLine	S2-DRSC	embryo-deri

[show in table...]

databaseRecords

developmentalStages

Class	
DevelopmentalStage	late emb:

[show in table...]

experiment

1 Experiment [details...]

experimentalFactors

2 ExperimentalFactor

Class	type	name	
ExperimentalFactor	cell line	S2-DRSC	[details...]
ExperimentalFactor	antibody	H3K18ac	[details...]

[show in table...]

strains

0 Strain

tissues

1 Tissue [details...]

Further information for this Submission

Lab: Sarah Elgin - Washington University

Project: Chromosomal Proteins - Karpen

Organism: D. melanogaster

Experiment: Genomic Distributions of Histone Modifications

Submission description

We aim to determe the locations of 125 chromosomal proteins and histone modifications across the Drosophila melanogaster genome. The proteins and modifications under study are involved in basic chromosomal functions such as DNA replication, gene expression, gene silencing, and inheritance. We will perform Chromatin ImmunoPrecipitation (ChIP) using genomic tiling arrays. We will initially assay localizations using chromatin

Features generated by this submission:

Feature type	Count	View data	Export
BindingSite	11924	RESULTS TABLE	TAB DELIMITED GFF3 SEQUENCE

Find overlapping features:

Find which overlap the features generated by this submission.

Find nearby genes:

Find Genes that have features generated by this submission located within

Lists

Lists in which this Submission can be found:

Browse metadata for this submission (click to toggle) ▶

Step	Inputs	Applied Protocol	Outputs
1	developmental_stage	Embryo 2-4hr OR	--> next Step
	organism	expt.687	Cage population
	strain_or_line	Oregon-R-modENCODE	
1	developmental_stage	Embryo 2-4hr OR	--> next Step
	strain_or_line	Oregon-R-modENCODE	Cage population
	organism	expt.690	
1	developmental_stage	Embryo 2-4hr OR	--> next Step
	strain_or_line	Oregon-R-modENCODE	Cage population
	organism	expt.686	
1	developmental_stage	Embryo 2-4hr OR	--> next Step
	strain_or_line	Oregon-R-modENCODE	Cage population
	organism	expt.650	
2	output from step 1 -->	Chromatin prep from fixed D.m. embryos	--> next Step
2	output from step 1 -->	Chromatin prep from fixed D.m. embryos	--> next Step
2	output from step 1 -->	Chromatin prep from fixed D.m. embryos	--> next Step
2	output from step 1 -->	Chromatin prep from fixed D.m. embryos	--> next Step
3	output from step 2 -->		--> next Step
	antibody	H3K18Ac (new lot)	ChIP
	chromatin	prep.9273392	
3	output from step 2 -->		--> next Step
	chromatin	prep.9273392	ChIP
	antibody	[no value]	
3	output from step 2 -->		--> next Step
	antibody	H3K18Ac (new lot)	ChIP
	chromatin	prep.8061212	
3	output from step 2 -->		--> next Step
	antibody	[no value]	ChIP
	chromatin	prep.8061212	
4	output from step 3 -->	WGA	--> next Step
4	output from step 3 -->	WGA	--> next Step
4	output from step 3 -->	WGA	--> next Step
4	output from step 3 -->	WGA	--> next Step
5	output from step 4 -->		--> next Step
	ADF	Affymetrix Drosophila Tiling Arrays v2.0R	Hybridization to Affy arrays
5	output from step 4 -->		--> next Step
	ADF	Affymetrix Drosophila Tiling Arrays v2.0R	Hybridization to Affy arrays
5	output from step 4 -->		--> next Step
	ADF	Affymetrix Drosophila Tiling Arrays v2.0R	Hybridization to Affy arrays

Growth

**Chromatin
preps**

ChIP

Hybridisation

5	output from step 4 -->		Hybridization to Affy arrays	--> next Step	
	ADF	Affymetrix Drosophila Tiling Arrays v2.0R			
6	output from step 5 -->		Array Scanning Protocol	CEL	687.pirrota_932_Early2_H3K18ac.CEL.zip
6	output from step 5 -->		Array Scanning Protocol	CEL	
6	output from step 5 -->		Array Scanning Protocol	CEL	
6	output from step 5 -->		Array Scanning Protocol	CEL	
7	CEL	687.pirrota_932_Early2_H3K18ac.CEL.zip	M-value normalization	GEO_record	[no value]
				WIG	687.Mvalues.wig
7	CEL	690.pirrota_935_Early2_Input_20cyc_.CEL.zip	M-value normalization	GEO_record	[no value]
				WIG	
7	CEL	686.pirrota_931_Early1_H3K18ac.CEL.zip	M-value normalization	GEO_re	
				WIG	
7	CEL	650.pirrota_879_E_early_1_Input.CEL.zip	M-value normalization	GEO_record	[no value]
				WIG	686.Mvalues.wig
8	WIG	687.Mvalues.wig	Smoothed M-value enrichment profiles	WIG	repset.9273880.smoothedM.wig
	smoothing_factor	500			
8	WIG	686.Mvalues.wig	Smoothed M-value enrichment profiles	WIG	
	smoothing_factor	500			
9	WIG	repset.9273880.smoothedM.wig	Regions of significant enrichment	GFF3	
	false_discovery_rate	1.00E-003			
	minRun	1000			
	maxGap	100			

Scanning

Normalisation

Enriched regions

Browse all modENCODE data

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

RNA expression profiling

Replication

www.modmine.org

Chromatin structure



Nucleosome mapping

Project Chromatin Function - Tison Lab | Lab: Tison Lab

2 GBrowse tracks
1 entry in GEO



Search

Enter names, identifiers or keywords for genes, proteins, pathways, ontology terms, etc. (e.g. **zen**, **pha-4**, **DNA binding**).

e.g. zen, pha-4

Search

Or search **modENCODE** experiments by type, lab name, antibody, etc. (e.g. **RNA-seq**, **Snyder**, **CP190**).

e.g. RNA-seq, CP190

Experiment Search



Genomic Region Search

Explore a genomic region for features found by the **modENCODE** project.



[Genomic Region Search](#)

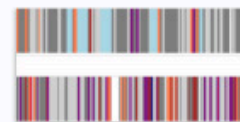
Fly Gene Expression



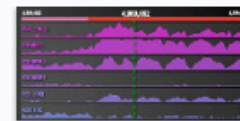
View an expression score heatmap for any list of fly genes. See an [example](#).

To upload your own list of genes, use the form above or go to the 'Lists' tab and click on 'Upload' to create and name the new list. You can also use any of modMine's queries to create a list.

Fly Chromatin states



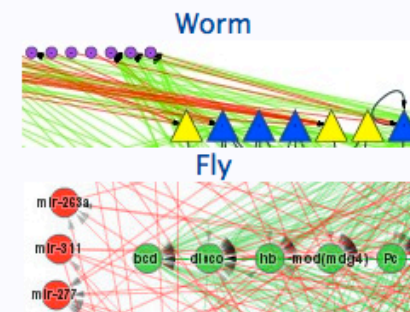
GBrowse Ideograms



Park Lab Viewer

Includes folded view and also data about DHS, TSS, replication, etc.

Regulatory Network



Explore an hierarchical view of the physical regulatory networks.



Upload Lists

Enter a list of identifiers.

Gene

CG9151, FBgn0000099, CG3629, TfiIB, Mad, CG1775, CG2262, TWIST_DROME, tinman, runt, E2f,

advanced

upload



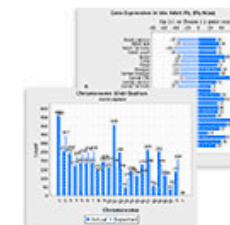
Use Template Queries

Get started with **powerful queries** using our predefined searches. These customizable templates have been designed around common tasks performed by our biologist community. To see how they work, why not try a template from our [examples page](#)?



Analyze Lists of Data

Explore and Analyze. Upload lists of identifiers to use in queries and discover relationships in our analysis widgets. See an [example](#).



lists in modMine

Create a new list

Select the type of list to create and either enter in a list of identifiers or upload identifiers from a file. A search will be performed for all the identifiers in your list.

- Separate identifiers by a **comma**, **space**, **tab** or **new line**.
- Qualify any identifiers that contain whitespace with double quotes like so: "even skipped".

Select Type:

Gene

for Organism:

D. melanogaster

Type/Paste in
identifiers

[\(click to see an example\)](#)

```
zfh1  
opa  
mef2  
cad  
ush  
ftz-f1  
apt  
prd  
tin  
exd  
lola
```

or Upload
identifiers from a
.txt file...

Browse...

fly gene expression from list

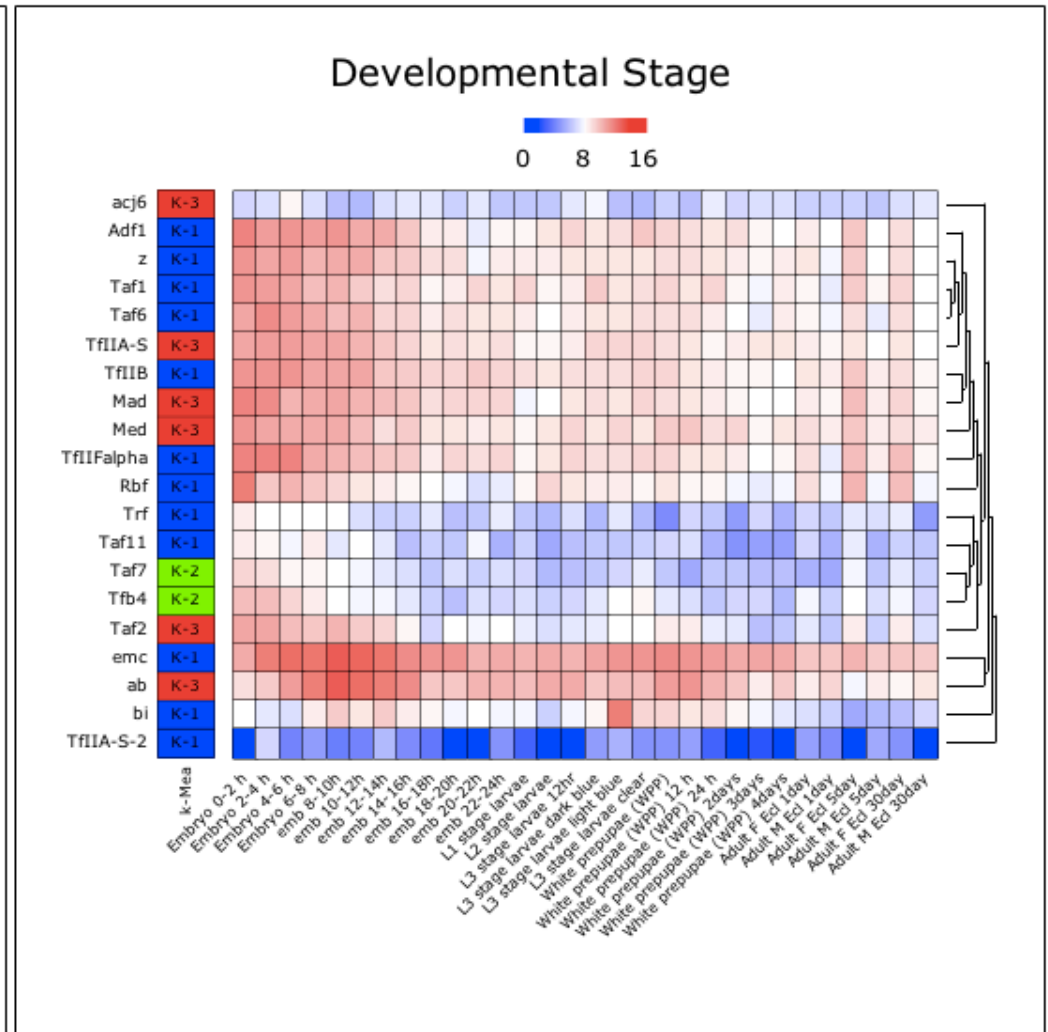
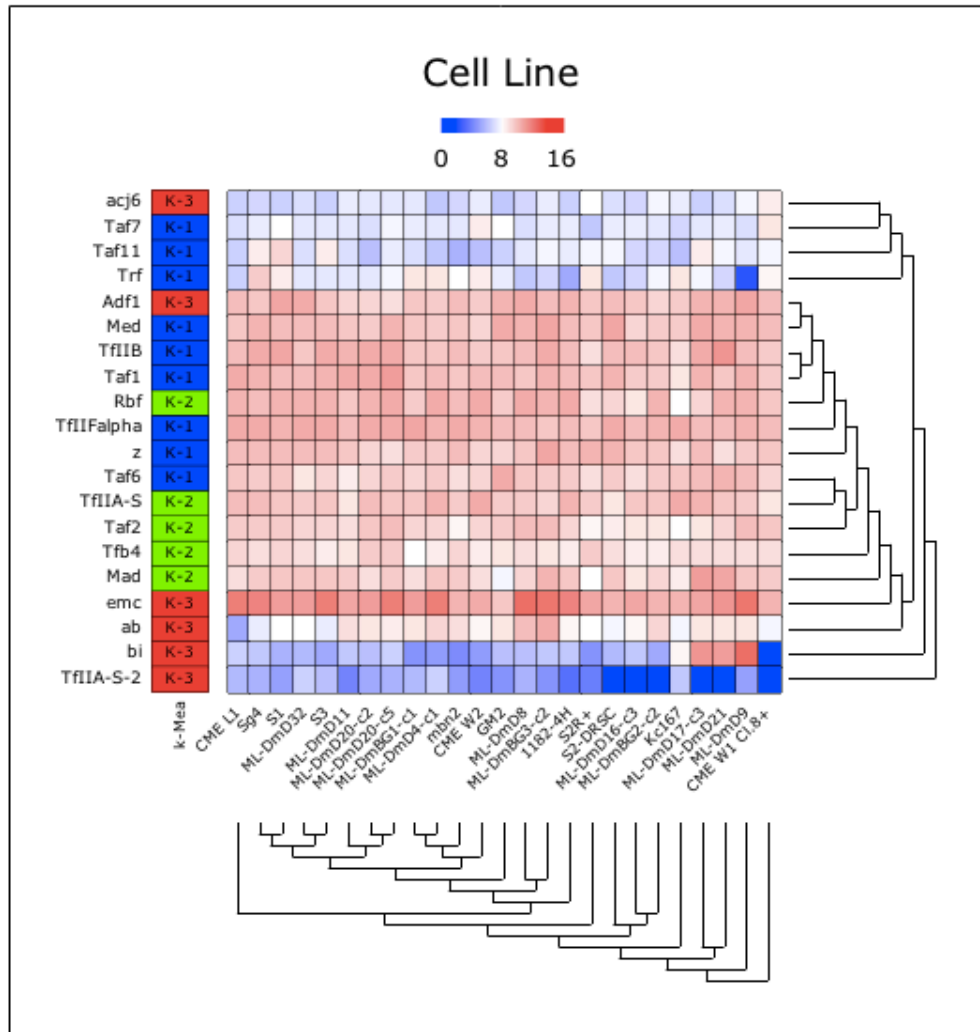
Drosophila melanogaster Gene Expression Scores

These expression levels are derived from RNA-seq data from the [Celniker group](#) and are \log_2 of the actual value. Heatmap visualization powered by [canvasXpress](#), learn more about the [display options](#).

Click to see/hide the expression maps ▶

Cell Line Clustering - Hierarchical: and K-means:

Developmental Stage Clustering - Hierarchical: and K-means:



Statistical enrichment

GO terms

Gene Ontology Enrichment close

GO terms enriched for items in this list. For more information about the math used in these calculations, see [here](#).

Number of Genes in this list not analysed in this widget: 0

Options

Multiple Hypothesis Test Correction Holm-Bonferroni

Ontology: biological_process

Maximum value to display 0.05

[View](#) [Download](#)

<input type="checkbox"/>	GO Term	p-Value	
<input type="checkbox"/>	transcription from RNA polymerase II promoter [GO:0006366]	9.5473E-20	19
<input type="checkbox"/>	regulation of transcription, DNA-dependent [GO:0006355]	2.9302E-18	21
<input type="checkbox"/>	regulation of RNA biosynthetic process [GO:2001141]	2.9302E-18	21
<input type="checkbox"/>	transcription, DNA-dependent [GO:0006351]	1.7978E-17	21
<input type="checkbox"/>	RNA biosynthetic process [GO:0032774]	1.896E-17	21
<input type="checkbox"/>	regulation of RNA metabolic process [GO:0051252]	2.1084E-17	21
<input type="checkbox"/>	regulation of macromolecule biosynthetic process [GO:0010556]	2.5361E-17	21
<input type="checkbox"/>	regulation of cellular macromolecule biosynthetic process [GO:2000112]	2.5361E-17	21
<input type="checkbox"/>	regulation of cellular biosynthetic process	7.6227E-17	21

Publications

Publication Enrichment close

Publications enriched for genes in this list. For more information about the math used in these calculations, see [here](#).

Number of Genes in this list not analysed in this widget: 0

Options

Multiple Hypothesis Test Correction Holm-Bonferroni

Maximum value to display 0.05

[View](#) [Download](#)

<input type="checkbox"/>	Publication	p-Value	
<input type="checkbox"/>	Genes encoding Drosophila melanogaster RNA polymerase II general transcription factors: diversity in TFIIA and TFIID components contributes to gene-specific transcriptional regulation. [PubMed:10908585]	2.3803E-19	12
<input type="checkbox"/>	ATP-dependent nucleosome remodeling and histone hyperacetylation synergistically facilitate transcription of chromatin. [PubMed:11279013]	7.5664E-15	9
<input type="checkbox"/>	Transcription properties of a cell type-specific TATA-binding protein, TRF. [PubMed:9335336]	3.1763E-13	8
<input type="checkbox"/>	Novel Mediator proteins of the small Mediator complex in Drosophila SL2 cells. [PubMed:12021283]	3.9784E-13	9
<input type="checkbox"/>	Reconstitution of chromatin transcription with purified components reveals a chromatin-specific repressive activity of p300. [PubMed:16415879]	6.7802E-11	7
<input type="checkbox"/>	Design and function of transcriptional switches in Drosophila. [PubMed:12225917]	6.9285E-11	9
	Drosophila head segmentation factor buttonhead		

Browse all modENCODE data

Chromatin structure Copy Number Variation Gene Structure Histone modification and replacement Metadata only Other chromatin binding sites RNA expression profiling Replication

www.modmine.org

Chromatin structure



Nucleosome mapping

Project: Chromatin Function - Tison 1 lab | Lab: Tison 1 lab

2 GBrowse tracks
1 entry in GEO



Search

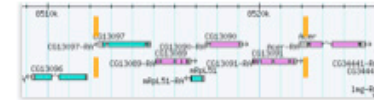
Enter names, identifiers or keywords for genes, proteins, pathways, ontology terms, etc. (e.g. **zen**, **pha-4**, **DNA binding**).

Or search **modENCODE** experiments by type, lab name, antibody, etc. (e.g. **RNA-seq**, **Snyder**, **CP190**).



Genomic Region Search

Explore a genomic region for features found by the **modENCODE** project.



[Genomic Region Search](#)

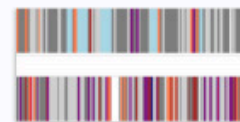
Fly Gene Expression



View an expression score heatmap for any list of fly genes. See an [example](#).

To upload your own list of genes, use the form above or go to the 'Lists' tab and click on **'Upload'** to create and name the new list. You can also use any of modMine's queries to create a list.

Fly Chromatin states



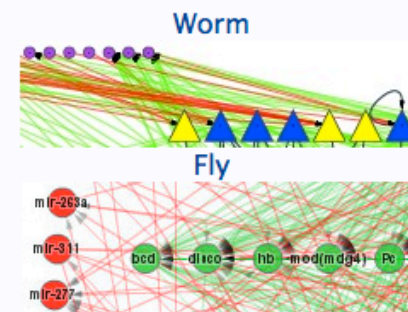
GBrowse Ideograms



Park Lab Viewer

Includes folded view and also data about DHS, TSS, replication, etc.

Regulatory Network



Explore an hierarchical view of the physical regulatory networks.



Upload Lists

Enter a list of identifiers.



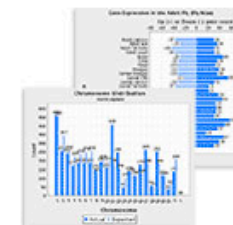
Use Template Queries

Get started with **powerful queries** using our predefined searches. These customizable templates have been designed around common tasks performed by our biologist community. To see how they work, why not try a template from our [examples page](#)?



Analyze Lists of Data

Explore and **Analyze**. Upload lists of identifiers to use in queries and discover relationships in our analysis widgets. See an [example](#).



[Contact Us](#)

Overlap features search from a new list of Genome Regions

Select the organism, experiments and feature types to create and either enter in a list of genome regions or upload genome regions from a file. A search will be performed for all the genome regions in your list.

- Genome regions in the following formats are accepted:
 - **chromosome:start..end**, e.g. *2L:11334..12296*
 - **chromosome:start-end**, e.g. *2R:5866746-5868284* or *chrII:14646344-14667746*
 - **tab delimited**
- Both **base coordinate** (e.g. **BLAST**, **GFF/GFF3**) and **interbase coordinate** (e.g. **UCSC BED**, **Chado**) systems are supported, e.g. for a DNA piece **GCCATGTA**, the position of the **ATG** in interbase is [3, 6], and in base coordinates is [4, 6]. Users need to explicitly select one. By default, the base coordinate is selected.
- Each genome region needs to take a **new line**.
- Only experiments with features are listed below.
- Right click an **experiment** in the tree to go to experiment report page.

1. Select Organism: genome version: 5.43

2. Select Experiments:

- Chromatin structure**
 - Copy Number Variation**
 - Differential Replication of Polytene Chromosomes
 - Gene Structure**
 - RT-PCR
 - 5' RACE
 - Annotation of the developmental transcriptome of *Drosophila melanogaster*
 - Validation of *Drosophila* transcripts by full-length cDNA screening and sequencing
 - Confirmation of *Drosophila* transcripts by full-length cDNA screening
 - Annotation of *Drosophila* splice junctions by RNA-seq
 - Gene Model Prediction
 - Histone modification and replacement**
 - Genomic Distributions of Histone Modifications
 - Chromatin Binding Site Mapping of Modified Histone Proteins in *D. melanogaster*



Browse all modENCODE data

Chromatin structure

Copy Number Variation

Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

RNA expression profiling

Replication

www.modmine.org

Chromatin structure



Nucleosome mapping

Project Chromatin Function - Tison Lab | Lab: Tison Lab

2 GBrowse tracks
1 entry in GEO



Search

Enter names, identifiers or keywords for genes, proteins, pathways, ontology terms, etc. (e.g. **zen**, **pha-4**, **DNA binding**).

e.g. zen, pha-4

Search

Or search **modENCODE** experiments by type, lab name, antibody, etc. (e.g. **RNA-seq**, **Snyder**, **CP190**).

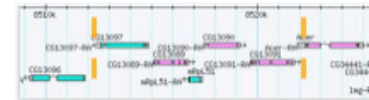
e.g. RNA-seq, CP190

Experiment Search



Genomic Region Search

Explore a genomic region for features found by the **modENCODE** project.



[Genomic Region Search](#)

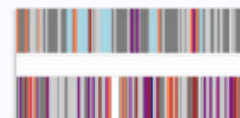
Fly Gene Expression



View an expression score heatmap for any list of fly genes. See an [example](#).

To upload your own list of genes, use the form above or go to the 'Lists' tab and click on 'Upload' to create and name the new list. You can also use any of modMine's queries to create a list.

Fly Chromatin states



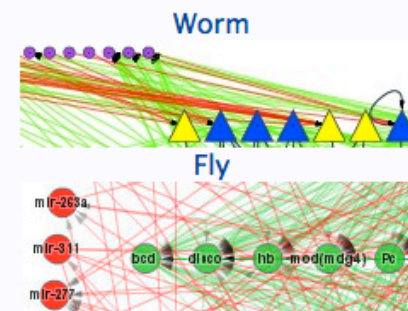
GBrowse Ideograms



Park Lab Viewer

Includes folded view and also data about DHS, TSS, replication, etc.

Regulatory Network



Explore an hierarchical view of the physical regulatory networks.



Upload Lists

Enter a list of identifiers.

Gene

CG9151, FBgn0000099, CG3629, TfiIB, Mad, CG1775, CG2262, TWIST_DROME, tinman, runt, E2f,

advanced

upload



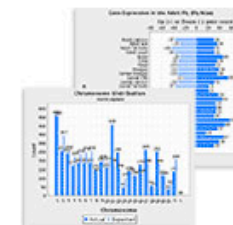
Use Template Queries

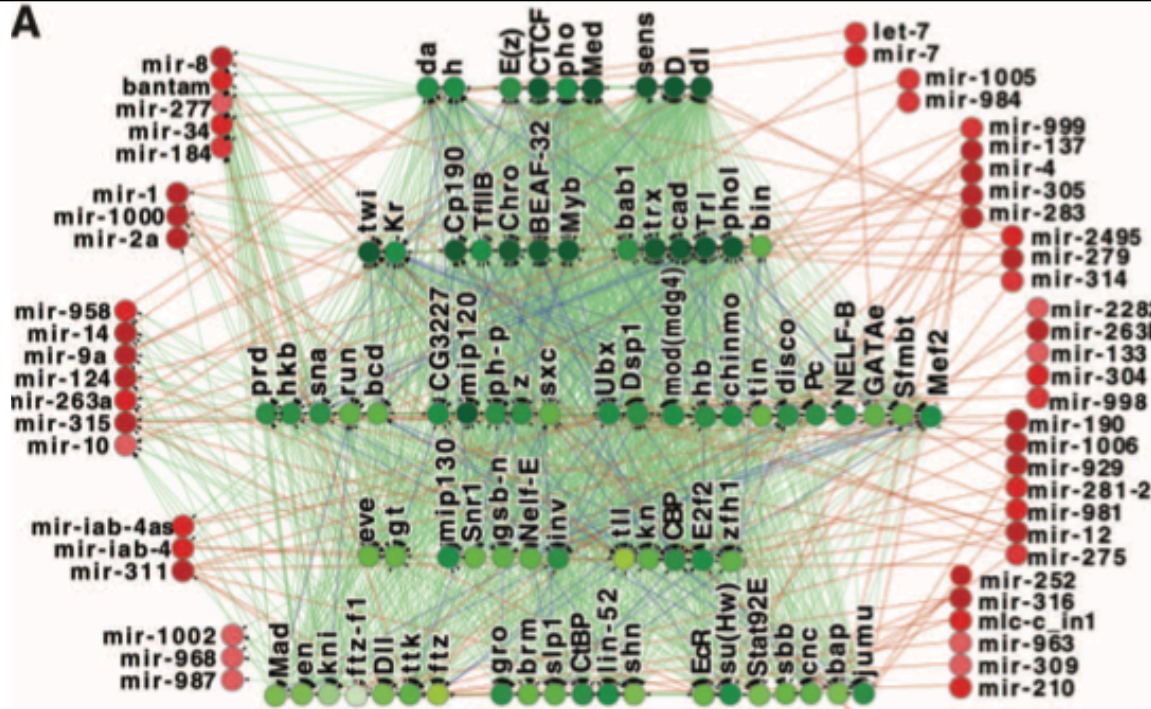
Get started with **powerful queries** using our predefined searches. These customizable templates have been designed around common tasks performed by our biologist community. To see how they work, why not try a template from our [examples page](#)?



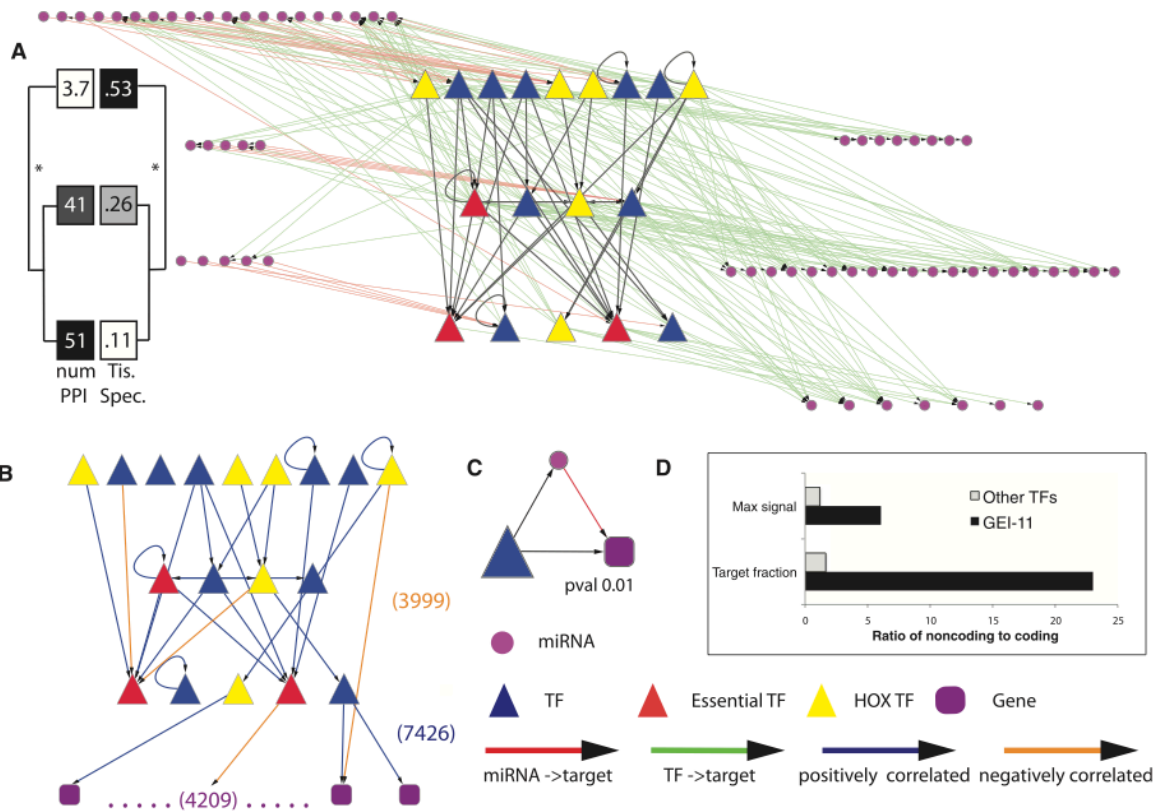
Analyze Lists of Data

Explore and Analyze. Upload lists of identifiers to use in queries and discover relationships in our analysis widgets. See an [example](#).



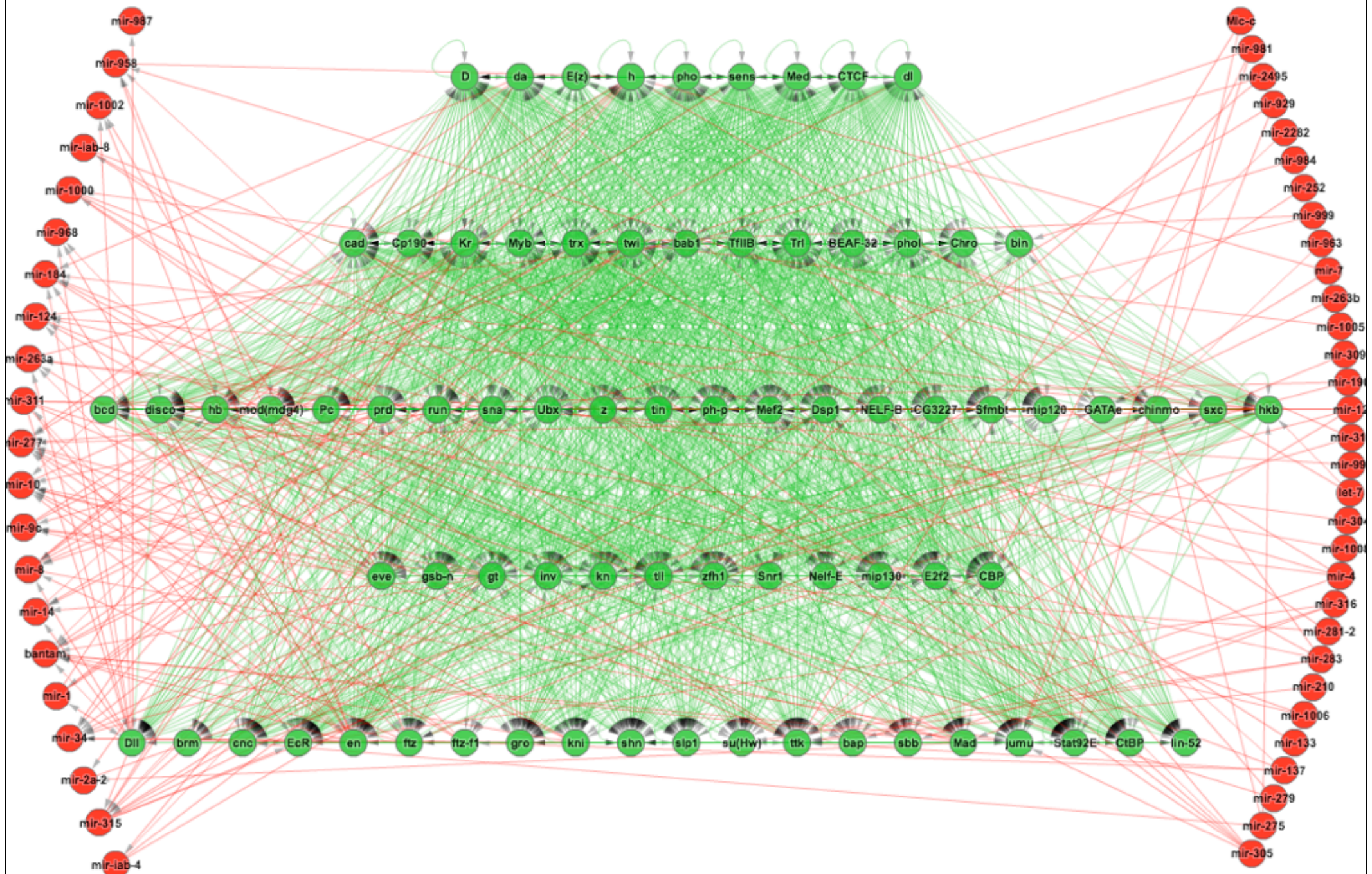


Science paper figures



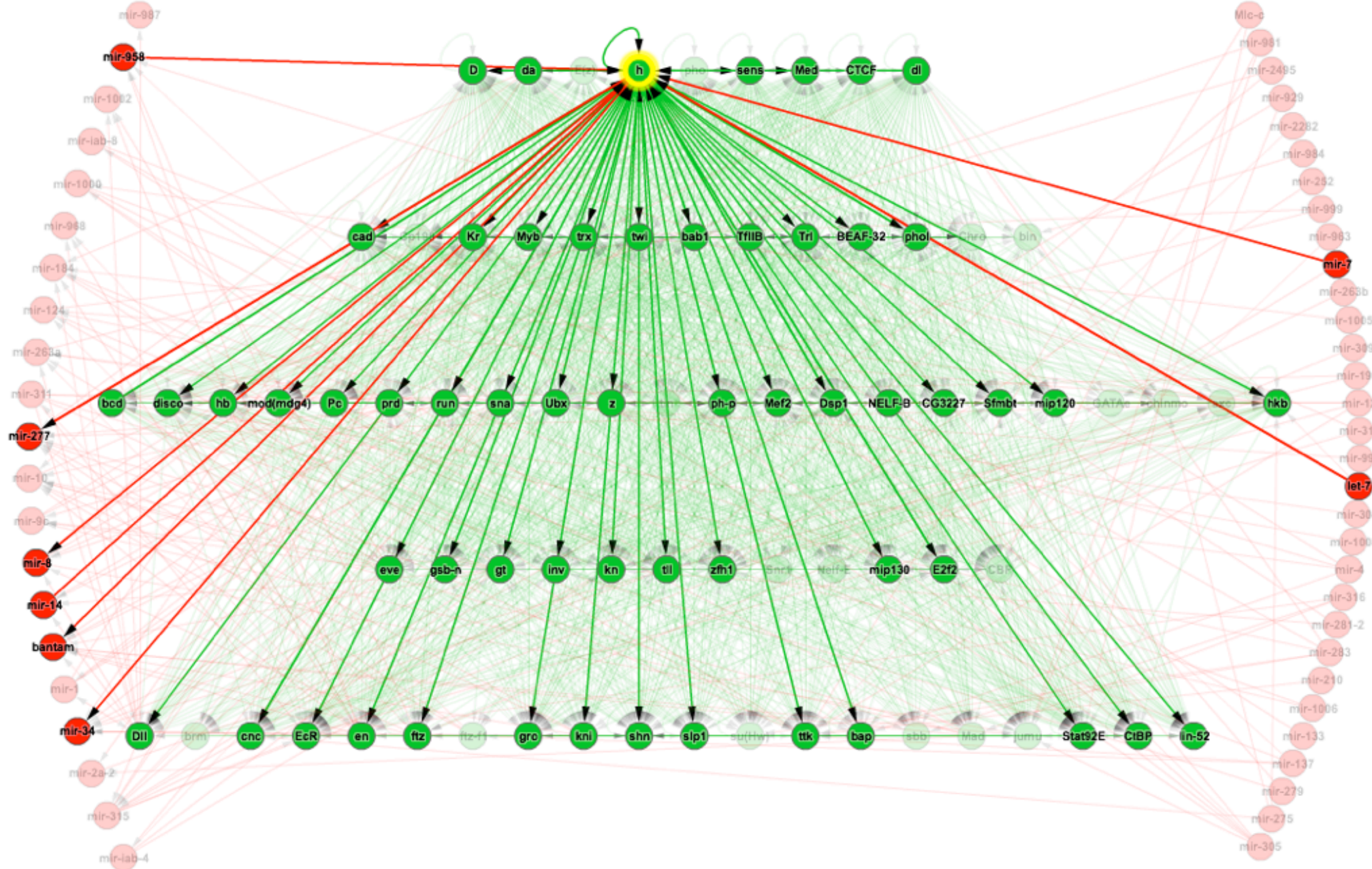
Hierarchical View Of Drosophila Physical Regulatory Network

About the network..













Hierarchical View Of Drosophila Physical Regulatory Network

About the network...





“The National Human Genome Research Institute (NHGRI) model organism ENCYclopedia Of DNA Elements (modENCODE) Project will try to identify all of the sequence-based functional elements in the *Caenorhabditis elegans* and *Drosophila melanogaster* genomes.”

modMine release #25	amazon Cloud Instance	Dataset Search Datasets	FTP Download Get FTP data					
Older versions: 25 , 24 , 23 , 22 , 21 , 20 , 19 , 18 , 17 , 16	The entire modENCODE data set available for analysis in the Amazon compute cloud.	Find, view and download datasets in bulk.	Download publicly released data using the traditional FTP interface.					
Choose an organism below to see GBrowse, Dataset Search links:								
 	 C. elegans	 D. melanogaster	 D. ananassae	 D. mojavensis	 D. pseudoobscura	 D. simulans	 D. virilis	 D. yakuba

Browse Projects

Chromatin structure


Copy Number Variation


Gene Structure

Histone modification and replacement

Metadata only

Other chromatin binding sites

 [D. melanogaster](#)

 [C. elegans](#)

Genome-wide Chromatin Profiling

Histone Variants

Pi: [Steven Henikoff](#) Labs: [Kamran Ahmad](#), [Steven Henikoff](#) Submissions: 32

Experimental factors: [developmental stage](#), [biochemical fraction](#), [cell line](#), [extraction time](#), [sodium chloride concentration](#)

GBrowse Tracks: [32](#)

http://aws.amazon.com/datasets/8042906995278110



“amazon modENCODE data”

Sign Up

My Account / Console

English

AWS Products & Solutions

Public Data Sets



Developers

Support

Browse By Category

- Astronomy
- Biology
- Chemistry
- Climate
- Economics
- Encyclopedic
- Geographic
- Mathematics



Model Organism Encyclopedia of DNA Elements (modENCODE)

Public Data Sets > Model Organism Encyclopedia of DNA Elements (modENCODE)

A collection of data from the modENCODE project (<http://www.modencode.org>)

Details

Submitted By: modENCODE DCC (help@modencode.org)

Size: 5 TB

Created On: April 19, 2012 9:25 PM GMT

Last Updated: April 24, 2012 9:18 PM GMT

NOTE: these snapshots only contained released data up to December 2011

AMI = Amazon Machine Image

Mount everything, GBrowse, just data

Pay as you go

www.bionimbus.org

Bionimbus Cloud

Bionimbus is a cloud-based system for managing, analyzing and sharing genomic data.

[News](#) [About Bionimbus](#) [Public Data](#) [Using Bionimbus](#) [Registered Users](#) [Support](#) [Sponsors](#)

About Bionimbus

Overview. Bionimbus is an open source cloud-based system for managing, analyzing and sharing genomic data that has been developed by the Institute for Genomics and Systems Biology (IGSB) at the University of Chicago. Bionimbus is designed to support next-generation sequencing instruments and integrates technology for the analyzing and transporting large datasets. There is an open source version of Bionimbus available to those who wish to set up their own clouds. There is also a Bionimbus Community Cloud operated by the 501(c)(3) [Open Cloud Consortium's](#) Open Science Data Cloud that includes a variety of public genomics and related data.

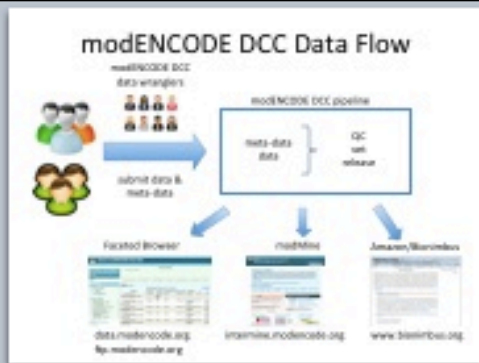
• **Bionimbus Community Cloud.** There is a [Bionimbus Community Cloud](#), which research collaborators can log into and use.

Bionimbus uses the [Open Cloud Consortium's](#) Open Science Data Cloud (OSDC) for its infrastructure. The first generation of Bionimbus used 7 racks of equipment containing approximately 3000 cores and 1 PB of disk that Yahoo! [donated](#) to the OSDC. Currently, we are using approximately six racks of equipment that the [Gordon and Betty Moore Foundation](#) has funded. [Cisco](#) has provided access to the Cisco C-Wave so that we can connect the four OSDC data centers together with 10 Gbps wide area networks.

• **Virtual machine images.** We develop and maintain [Bionimbus machine images](#) that can be run on:

- i) The Bionimbus Community Cloud.
- ii) Public clouds such as Amazon's.

Search Bionimbus Cloud



1

The screenshot shows the homepage of www.modencode.org. It features a search bar, navigation tabs for 'Genomes', 'Genes', 'Proteins', 'Transcripts', 'Chromatin', and 'Other'. Below the search bar, there are sections for 'Browse Projects' and 'Genome-wide Chromatin Profiling'.

2

The screenshot shows the 'Search modENCODE Data Sets' page. It displays a table of search results with columns for 'Name', 'Size', and 'Date Modified'. The table lists various data sets such as 'C. elegans', 'Drosophila', and 'Mus musculus'.

3

The screenshot shows the 'Index of /' page for ftp.modencode.org. It lists various files and directories with columns for 'Name', 'Size', and 'Date Modified'. Files include 'C. elegans', 'Drosophila', and 'Mus musculus'.

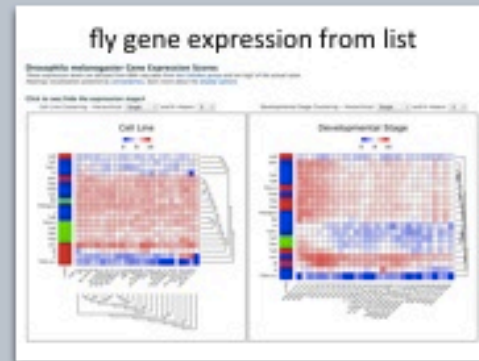
4

The screenshot shows the GBrowse genome browser interface. It displays multiple tracks for genomic data, including gene models, repeats, and other features. A text box in the bottom right corner states 'Can save track combinations'.

5

The screenshot shows the homepage of www.modmine.org. It features a search bar and several analysis tools, including 'Fly Data Exploration', 'Fly Chromatin tracks', and 'Regulatory Network'.

6



7



8

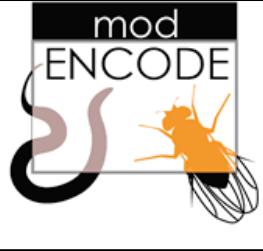
The screenshot shows the Amazon S3 console interface. It displays the 'amazon modENCODE data' bucket. Below the bucket name, there is a note: 'NOTE: these snapshots only contained released data up to December 2011'. At the bottom, it says 'AMI = Amazon Machine Image Mount everything, GBrowse, just data Pay as you go'.

9

The screenshot shows the homepage of www.bionimbus.org. It features a search bar and information about the 'Bionimbus Cloud' service, which provides access to genomic data.

10

help@modencode.org



Acknowledgments

modENCODE DCC:

Nicole Washington, Seth Carbon, Ellen Kephart, Paul Lloyd, Chris Mungall, E.O. Stinson, Suzanna Lewis (LBNL)

Daniela Butano, Sergio Contrino, Fengyuan Hu, Rachel Lyne, Kim Rutherford, Richard Smith, Gos Micklem (Cambridge)

Angie Hinrichs, Jim Kent (UCSC)

Marc Perry, Peter Ruzanov, Quang Trinh, Zheng Zha, Lincoln Stein (OICR)

All the modENCODE data producers