

- 1) Go to (<http://epic.gs.washington.edu/modERN/>) in your web browser.
- 2) The project and website description will be displayed (highlighted in orange) by clicking on that tab.

Fly Reference Genome

R5
R6

Worm Reference Genome

ce10
ce11

Project and Website Description
Worm By Gene
Worm By LifeStage
Fly By Gene
Fly By LifeStage
File Descriptions

modERN (model organism Encyclopedia of Regulatory Networks) is an offshoot of the former modENCODE project. The goal of this new consortium is to streamline our efforts to perform ChIP-seq for transcription factors (TFs) in both worm and fly, and to begin to determine the functionality of the identified binding sites in gene regulation. This site organizes and provides all the ChIP-seq data files generated for transcription factors in worm and fly for both modENCODE and modERN projects.

Users can search for datasets in worm and fly by TF or by life-stage in their chosen reference genome. A tab containing information on the file descriptions is also provided. Users can access the data files using the drop down arrow from their chosen transcription factor. A brief description of the file appears as you scroll over each file type. Multiple files can be downloaded at one time using the "Download Selected Files" button at the upper left of the screen and all peak files from each organism can be downloaded using the button "Download All Peak Files." Datasets for orthologous TFs have been linked. Links to Wormbase/Flybase, the BDGP insitu homepage, the UCSC Genome Browser, Caenorhabditis Genetics Center (CGC), Bloomington Stock Center, and the DCC ENCODE site for each TF/dataset are also provided.

- 3) Choose your organism and reference genome (highlighted in orange) by clicking on the circle (indicated by red arrow) and then choose whether you want to view the datasets by gene name or by lifestage by clicking on the tab. Shown is "Worm by gene" (highlighted in orange). The steps to follow to access all the data and information are the same for fly. The datasets are displayed in alphabetical order. Scroll down to find your favorite gene/TF. Scroll to the right to see all available features.

Fly Reference Genome

R5
R6

Worm Reference Genome

ce10
ce11

Project and Website Description
Worm By Gene
Worm By LifeStage
Fly By Gene
Fly By LifeStage
File Descriptions

Download Selected Files
Download All Peak Files

Worm ChipSeq

Identifier	Type	Strain	Orthologs	Micrograph	ExpMetaData	InputMetaData	Browser Tracks	Add'l Info	Download
* aha-1	TF							WormBase	All Stages for this Gene
* ahr-1	TF							WormBase	All Stages for this Gene
* alr-1	TF		toe					WormBase	All Stages for this Gene
* aly-1	NonTF							WormBase	All Stages for this Gene
* aly-2	NonTF							WormBase	All Stages for this Gene
* ama-1	NonTF		Rpl1215					WormBase	All Stages for this Gene
* B0035.1	TF							WormBase	All Stages for this Gene
* B0261.1	TF		CG9305					WormBase	All Stages for this Gene
* B0310.2	TF		CG8319 : sen					WormBase	All Stages for this Gene
* blmp-1	TF							WormBase	All Stages for this Gene
* C04F5.9	TF		CG8089 : CG					WormBase	All Stages for this Gene
* C08G9.2	TF							WormBase	All Stages for this Gene
* C16A3.4	TF							WormBase	All Stages for this Gene
* C34F6.9	TF							WormBase	All Stages for this Gene
* ceh-14	TF							WormBase	All Stages for this Gene
* ceh-16	TF		en					WormBase	All Stages for this Gene
* ceh-18	TF							WormBase	All Stages for this Gene
* ceh-2	TF		ems					WormBase	All Stages for this Gene
* ceh-24	TF		tin					WormBase	All Stages for this Gene

- 4) To view a specific dataset, in this case TBX-2 , click on the blue arrow to the left of the gene (indicated by red arrow) and then all lifestages available for TBX-2 will be displayed. For each gene, we provide the type of gene (under Type, TF or nonTF), whether it has an ortholog (under Orthologs) and a link to Wormbase (or Flybase for fly) for additional information (under Add'l info).

Fly Reference Genome	<span style="color: blue;">●</span> R5 <span style="color: orange;">●</span> R6
Worm Reference Genome	<span style="color: blue;">●</span> ce10 <span style="color: orange;">●</span> ce11

Project and Website Description	Worm By Gene	Worm By LifeStage	Fly By Gene	Fly By LifeStage	File Descriptions
---------------------------------	--------------	-------------------	-------------	------------------	-------------------

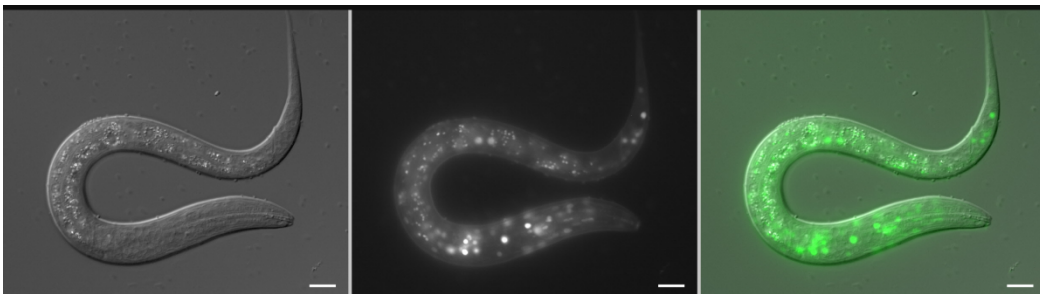
Download Selected Files	Download All Peak Files
-------------------------	-------------------------

Worm ChipSeq									
Identifier	Type	Strain	Orthologs	Micrograph	ExpMetaData	InputMetaData	Browser Tracks	Add'l Info	Download
<div> <div></div> <div>tbx-2</div> </div>	TF		<a href="#">prg-1</a>					<a href="#">WormBase</a>	<a href="#">All Stages for this Gene</a>
<div> <div></div> <div>L1 larva</div> </div>	Stage	OP159		<a href="#">View</a>	<a href="#">ENCSR238YNR</a>	<a href="#">ENCSR484UJP</a>	<a href="#">UCSC</a>		<a href="#">All DataTypes for this Stage</a>
<div> <div></div> <div>L3 larva</div> </div>	Stage	QP159		<a href="#">View</a>	<a href="#">ENCSR344UEX</a>	<a href="#">ENCSR335VIF</a>	<a href="#">UCSC</a>		<a href="#">All DataTypes for this Stage</a>
<div> <div></div> <div>late embryonic</div> </div>	Stage	QP159		<a href="#">View</a>	<a href="#">ENCSR162MNX</a>	<a href="#">ENCSR896VBS</a>	<a href="#">UCSC</a>		<a href="#">All DataTypes for this Stage</a>
<div> <div></div> <div>tbx-7</div> </div>	TF							<a href="#">WormBase</a>	<a href="#">All Stages for this Gene</a>
<div> <div></div> <div>tbx-9</div> </div>	TF							<a href="#">WormBase</a>	<a href="#">All Stages for this Gene</a>

- 5) For each stage, we provide the strain name (under Strain), images of the worm taken before collection for worm (under Micrograph), and links to the Experiment summary at the ENCODE DCC site for the ChIP replicates (under ExpMetaData) and the input control (InputMetaData). There are also links to the UCSC browser to view the signal track.

- 6) Worm image of TBX-2 taken at the L1 stage before collection after clicking “view” under Micrograph.



7) Experiment summary page at the ENCODE DCC site for the TBX-2 L1 dataset after clicking the link under “ExpMetaData.” Each dataset has a unique identifier, in this case “ENCSR238YNR.” This page provides detailed information on the dataset such as the sequencing platform used, collection information, all raw and processed files, worm images, and a ChIP protocol. The page also

link to the instrument control page, and information on the instrument, and

ENCODE Data Encyclopedia Materials & Methods Help

EXPERIMENTS / CHIP-SEQ / CAENORHABDITIS ELEGANS / WHOLE ORGANISM

## Experiment summary for ENCSR238YNR

Status: released

Summary		Attribution	
<b>Assay:</b>	ChIP-seq	<b>Lab:</b>	Valerie Reinke, Yale
<b>Target:</b>	<a href="#">eGFP-tbx-2</a>	<b>Award:</b>	<a href="#">U41HG007355</a> (Robert Waterston, UW)
<b>Biosample summary:</b>	Caenorhabditis elegans strain OP159 (unc-119(ed3) III; wgl159) whole organism hermaphrodite L1 larva 6 hours post synchronization at L1 larva starved after bleaching stage expressing fusion protein	<b>Project:</b>	modERN
<b>Biosample Type:</b>	whole organisms	<b>Aliases:</b>	kevin-white:TBX-2-OP159-L1
<b>Replication type:</b>	isogenic	<b>Date released:</b>	2016-05-10
<b>Description:</b>	ChIP-seq of transgenic worms expressing TBX-2-eGFP fusion proteins. The IP was performed using an anti-GFP antibody.	<b>Submitter comment:</b>	ENCOFF158CYV, ENCOFF986IYR, ENCOFF578RIA, ENCOFF118QGO were revoked because a software bug accidentally truncated the file content. The submitters will provide replacement files.
<b>Nucleic acid type:</b>	DNA		
<b>Fragmentation method:</b>	sonication (generic microtip)		
<b>Size selection method:</b>	SPRI		
<b>Strand specificity:</b>	Non-strand-specific		
<b>Platform:</b>	<a href="#">HiSeq 2500</a>		
<b>Controls:</b>	<a href="#">ENCSR484UJP</a>		

8) Experiment summary page at the ENCODE DCC site for the TBX-2 L1 control dataset after clicking the

ENCODE Data Encyclopedia Materials & Methods Help

EXPERIMENTS / CHIP-SEQ / CAENORHABDITIS ELEGANS / WHOLE ORGANISM

## Experiment summary for ENCSR484UJP

Status: released

Summary		Attribution	
<b>Assay:</b>	ChIP-seq	<b>Lab:</b>	Valerie Reinke, Yale
<b>Target:</b>	<a href="#">Control</a>	<b>Award:</b>	<a href="#">U41HG007355</a> (Robert Waterston, UW)
<b>Biosample summary:</b>	Caenorhabditis elegans strain OP159 (unc-119(ed3) III; wgl159) whole organism hermaphrodite L1 larva 6 hours post synchronization at L1 larva starved after bleaching stage expressing fusion protein	<b>Project:</b>	modERN
<b>Biosample Type:</b>	whole organisms	<b>Aliases:</b>	kevin-white:TBX-2-OP159-L1-INPUT
<b>Replication type:</b>	unreplicated	<b>Date released:</b>	2016-05-10
<b>Description:</b>	This is the ChIP-seq control of transgenic worms expressing TBX-2-eGFP fusion proteins.		
<b>Nucleic acid type:</b>	DNA		
<b>Fragmentation method:</b>	sonication (generic microtip)		
<b>Size selection method:</b>	SPRI		
<b>Strand specificity:</b>	Non-strand-specific		
<b>Platform:</b>	<a href="#">HiSeq 2500</a>		

9) You can also view the datasets by lifestage by choosing “Worm by Lifestage” (highlighted in orange) or “Fly by Lifestage”. Shown are all the stages that have been collected for worm. You can download all the data associated with each stage by choosing “All TFs for this Stage” in the download column indicated by the red arrow (this will take some time!).

Fly Reference Genome	<input checked="" type="radio"/> R5 <input checked="" type="radio"/> R6
Worm Reference Genome	<input checked="" type="radio"/> ce10 <input checked="" type="radio"/> ce11

Project and Website Description	Worm By Gene	Worm By LifeStage	Fly By Gene	Fly By LifeStage	File Descriptions
---------------------------------	--------------	-------------------	-------------	------------------	-------------------

Download Selected Files	Download All Peak Files
-------------------------	-------------------------

Worm ChIPSeq										
Identifier	Type	Strain	Orthologs	Micrograph	ExpMetaData	InputMetaData	Browser Tracks	Add'l Info	Download	
* L1 larva	Stage								All TFs for this Stage	
* L2 larva	Stage								All TFs for this Stage	
* L3 larva	Stage								All TFs for this Stage	
* L4 larva	Stage								All TFs for this Stage	
* L4/young adult	Stage								All TFs for this Stage	
* early embryonic	Stage								All TFs for this Stage	
* late embryonic	Stage								All TFs for this Stage	
* midembryonic	Stage								All TFs for this Stage	
* mixed stage (embryonic)	Stage								All TFs for this Stage	
* young adult	Stage								All TFs for this Stage	

10) Shown below are all the genes/TFs collected at the L1 stage in worm. Scroll down for more.

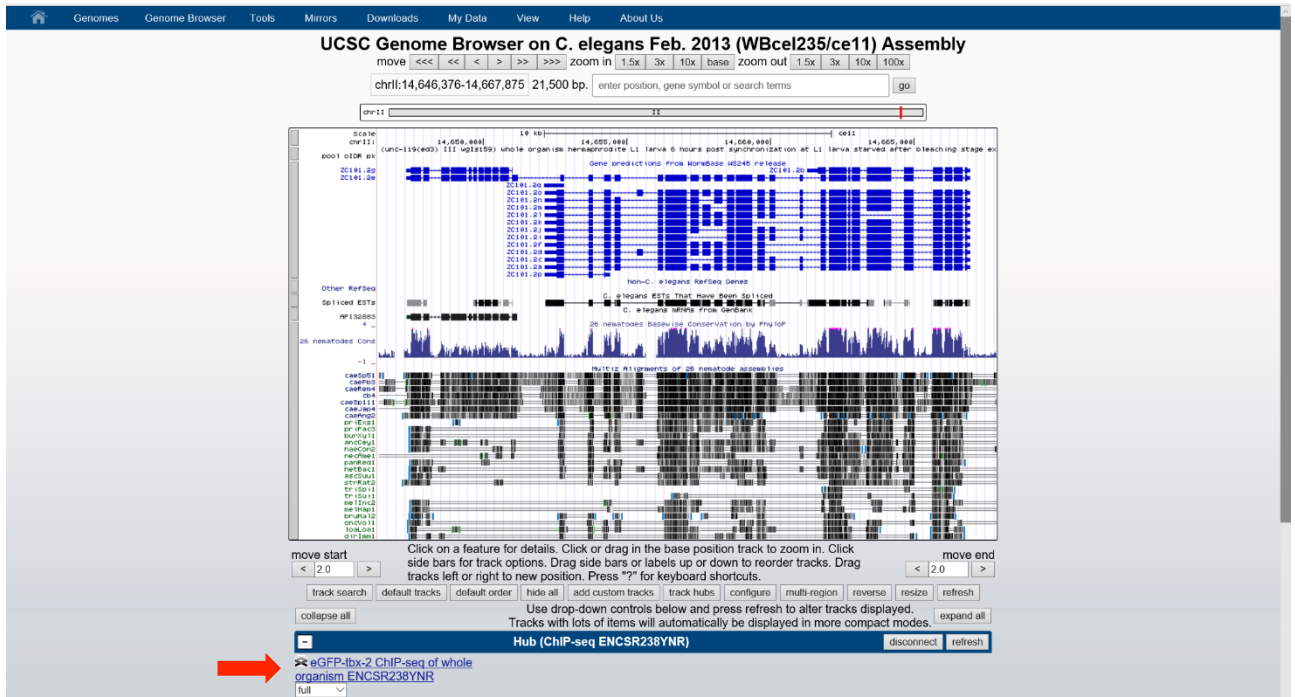
Fly Reference Genome	<input checked="" type="radio"/> R5 <input checked="" type="radio"/> R6
Worm Reference Genome	<input checked="" type="radio"/> ce10 <input checked="" type="radio"/> ce11

Project and Website Description	Worm By Gene	Worm By LifeStage	Fly By Gene	Fly By LifeStage	File Descriptions
---------------------------------	--------------	-------------------	-------------	------------------	-------------------

Download Selected Files	Download All Peak Files
-------------------------	-------------------------

Worm ChIPSeq										
Identifier	Type	Strain	Orthologs	Micrograph	ExpMetaData	InputMetaData	Browser Tracks	Add'l Info	Download	
* L1 larva	Stage								All TFs for this Stage	
* C04F5.9	TF	QP525	CG8089 : CG3	View	ENCSR917KLO	ENCSR748EUE	UCSC	WormBase	All DataTypes for this Gene	
* C16A3.4	TF	QP345			ENCSR025VUG	ENCSR181OGQ	UCSC	WormBase	All DataTypes for this Gene	
* F16B12.6	TF	QP114			ENCSR592DAH	ENCSR622TPD	UCSC	WormBase	All DataTypes for this Gene	
* Y116A8C.19	TF	QP532		View	ENCSR334YYP	ENCSR411ODP	UCSC	WormBase	All DataTypes for this Gene	
* ZK337.2	TF	QP355			ENCSR453VRK	ENCSR724QCI	UCSC	WormBase	All DataTypes for this Gene	
* aha-1	TF	QP124			ENCSR977DAI	ENCSR489KVE	UCSC	WormBase	All DataTypes for this Gene	
* aly-2	NonTF	QP217			ENCSR666MXI	ENCSR903ZHS	UCSC	WormBase	All DataTypes for this Gene	
* ama-1	NonTF	N2	Rpl1215		ENCSR881KXY	ENCSR338IOQ		WormBase	All DataTypes for this Gene	
* blmp-1	TF	QP109			ENCSR232QHU	ENCSR065BRX	UCSC	WormBase	All DataTypes for this Gene	
* ceh-82	TF	QP212			ENCSR725SQB	ENCSR032JLQ	UCSC	WormBase	All DataTypes for this Gene	
* ceh-90	TF	QP210		View	ENCSR891WMQ	ENCSR934KWG	UCSC	WormBase	All DataTypes for this Gene	
* ces-1	TF	OP174			ENCSR767XEC	ENCSR831TIJ		WormBase	All DataTypes for this Gene	
* dpl-1	TF	YL425			ENCSR219QTR	ENCSR675QOS	UCSC	WormBase	All DataTypes for this Gene	
* dpl-1	TF	OP105			ENCSR944YBH	ENCSR165EZK	UCSC	WormBase	All DataTypes for this Gene	
* dsc-1	TF	QP522	QdsH : toe : C	View	ENCSR489SGT	ENCSR190ONG	UCSC	WormBase	All DataTypes for this Gene	
* efl-1	TF	N2			ENCSR336KUP	ENCSR006GXP	UCSC	WormBase	All DataTypes for this Gene	
* efl-1	TF	MT10430			ENCSR878HTM	ENCSR239PCG	UCSC	WormBase	All DataTypes for this Gene	
* efl-1	TF	YL418			ENCSR898YRV	ENCSR294SZY	UCSC	WormBase	All DataTypes for this Gene	

- 11) To view the signal track for the TBX-2 L1 dataset click the UCSC browser track link (under Browser tracks). Only datasets that have a UCSC track available will have a link. To visualize the track, click the link indicated by the red arrow.



- 12) Make sure that the signal is set to full (indicated by the arrow). The default is that the pooled replicates with input subtracted signal track is displayed. To view the called peaks that passed IDR (scoring), click on the "Optimal IDR thresholded peaks" track indicated by the black arrow. Make sure that the peak track is set to "dense." Click submit below on the left to view.

The screenshot shows the UCSC Genome Browser 'Track Settings' page for the eGFP-tbx-2 ChIP-seq dataset. The page shows various track settings including 'Optimal IDR thresholded peaks' (indicated by a black arrow) and 'Signal' (indicated by a red arrow). The 'Signal' dropdown is set to 'full'.

**Track Settings**

**eGFP-tbx-2 ChIP-seq of whole organism ENCSR238YNR**

Maximum display mode: **full** Submit Cancel Reset to defaults

Select views (help): **Optimal IDR thresholded peaks** dense Peaks (hide) Signal **full**

Select subtracks by multiple variables:

Replicates: ☒ Pooled ☐ Replicate 1 ☐ Replicate 2

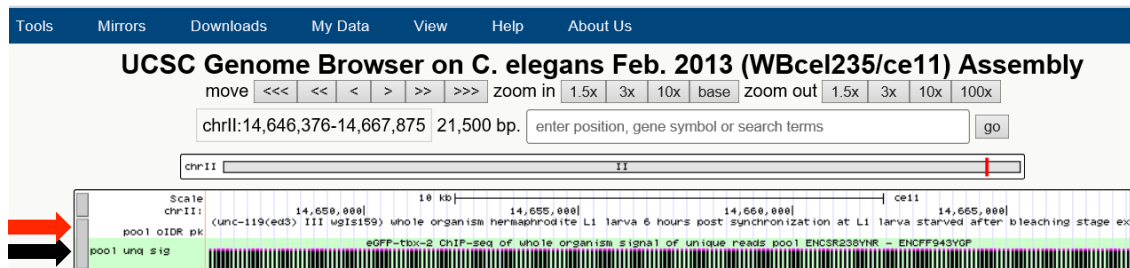
List subtracks: ☐ Only selected/visible ☒ all (2 of 7 selected)

Biosample	Targets	Replicates	Views	Track Name
<input checked="" type="checkbox"/> dense	whole organism eGFP-tbx-2 Pooled	Optimal IDR thresholded peaks	eGFP-tbx-2 ChIP-seq of whole organism optimal idr thresholded peaks pool ENCSR238YNR - ENCF852JBM	
<input checked="" type="checkbox"/> hide	whole organism eGFP-tbx-2 Pooled	Peaks	eGFP-tbx-2 ChIP-seq of whole organism peaks pool ENCSR238YNR - ENCF244FMV	
<input checked="" type="checkbox"/> full	whole organism eGFP-tbx-2 Pooled	Signal	eGFP-tbx-2 ChIP-seq of whole organism signal of unique reads pool ENCSR238YNR - ENCF943YGP	
<input type="checkbox"/> hide	whole organism eGFP-tbx-2 Replicate 1	Peaks	eGFP-tbx-2 ChIP-seq of whole organism peaks rep1 ENCSR238YNR - ENCF186GHH	
<input type="checkbox"/> hide	whole organism eGFP-tbx-2 Replicate 1	Signal	eGFP-tbx-2 ChIP-seq of whole organism signal of unique reads rep1 ENCSR238YNR - ENCF061TSO	
<input type="checkbox"/> hide	whole organism eGFP-tbx-2 Replicate 2	Peaks	eGFP-tbx-2 ChIP-seq of whole organism peaks rep2 ENCSR238YNR - ENCF039LLO	
<input type="checkbox"/> full	whole organism eGFP-tbx-2 Replicate 2	Signal	eGFP-tbx-2 ChIP-seq of whole organism signal of unique reads rep2 ENCSR238YNR - ENCF825WJF	

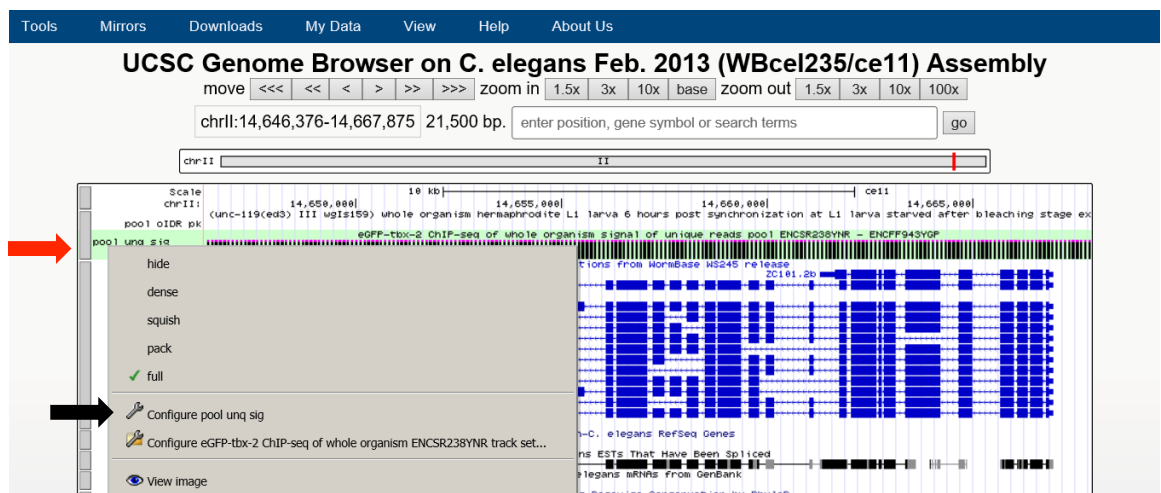
2 of 7 selected

Submit

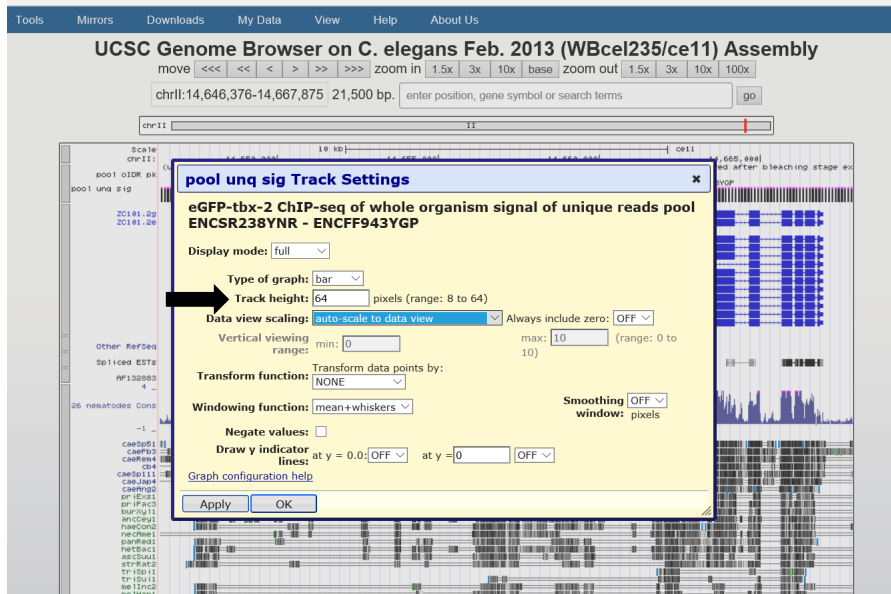
- 13) Shown is the signal track for the pooled replicates with background subtracted (black arrow) and the called peaks track shown as black bars (red arrow).



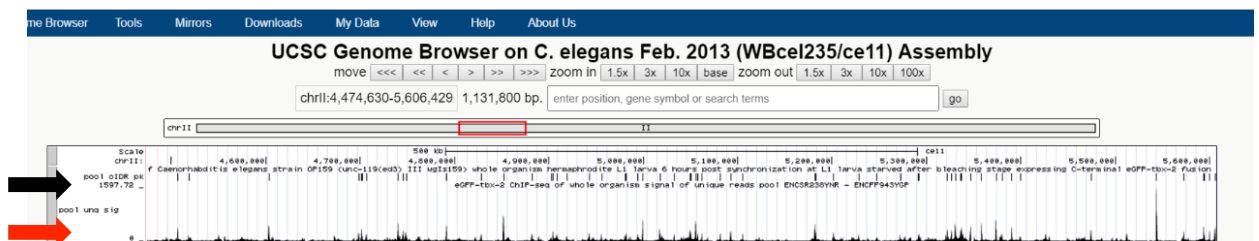
- 14) To change the track height, right click the grey bar on the left of the track (red arrow). Make sure that just the signal track is highlighted in green. Click on "Configure pool unq sig" (black arrow).



- 15) Change the track height to your desired height (black arrow). Make sure the Data view scaling below is set to “auto-scale to data view.” Click “apply” on the lower left and then “OK” to view the changes.



- 16) Shown is the pooled replicates with input subtracted signal track (red arrow) and the called peaks that passed scoring (Optimal IDR thresholded peaks) track as black bars above (black arrow).



17) To view the individual replicates, click on replicates 1 and 2 (indicated by the red arrow). You can turn off the individual replicate peak tracks by unclicking the boxes on the left or by setting the Peaks to “hide.” To view the called peaks that passed IDR (scoring), click on the “Optimal IDR thresholded peaks” track indicated by the black arrow. Make sure that it is set to “dense.” Click submit below on the left to view. Follow the same directions given for the pooled replicates on the previous page to adjust the track height.

UCSC Genome Browser Track Settings

eGFP-tbx-2 ChIP-seq of whole organism ENC238YNR Track Settings

Maximum display mode: full Submit Cancel Reset to defaults

Select views (help):  
Optimal IDR thresholded peaks dense Peaks hide Signal full

Select subtracks by multiple variables:  
Replicates: Pooled ☒ Replicate 1 ☒ Replicate 2

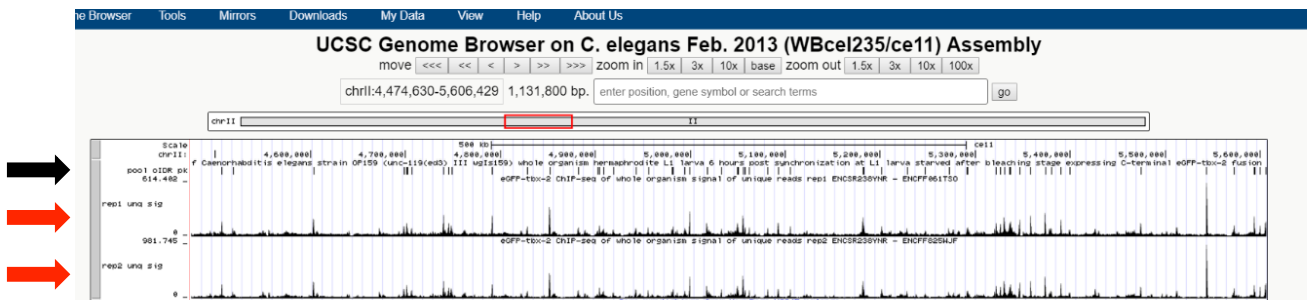
List subtracks: only selected/visible all (3 of 7 selected)

Biosample <sup>1</sup>	Targets <sup>2</sup>	Replicates <sup>3</sup>	Views <sup>4</sup>	Track Name <sup>5</sup>
<input checked="" type="checkbox"/> dense	whole organism eGFP-tbx-2 Pooled	Optimal IDR thresholded peaks	<input checked="" type="checkbox"/>	eGFP-tbx-2 ChIP-seq of whole organism optimal idr thresholded peaks pool ENC238YNR - ENCFF652JBM
<input checked="" type="checkbox"/> hide	whole organism eGFP-tbx-2 Pooled	Peaks	<input type="checkbox"/>	eGFP-tbx-2 ChIP-seq of whole organism peaks pool ENC238YNR - ENCFF244FMV
<input checked="" type="checkbox"/> full	whole organism eGFP-tbx-2 Pooled	Signal	<input checked="" type="checkbox"/>	eGFP-tbx-2 ChIP-seq of whole organism signal of unique reads pool ENC238YNR - ENCFF943YGP
<input checked="" type="checkbox"/> hide	whole organism eGFP-tbx-2 Replicate 1	Peaks	<input type="checkbox"/>	eGFP-tbx-2 ChIP-seq of whole organism peaks rep1 ENC238YNR - ENCFF186GHH
<input checked="" type="checkbox"/> full	whole organism eGFP-tbx-2 Replicate 1	Signal	<input checked="" type="checkbox"/>	eGFP-tbx-2 ChIP-seq of whole organism signal of unique reads rep1 ENC238YNR - ENCFF061TSO
<input checked="" type="checkbox"/> hide	whole organism eGFP-tbx-2 Replicate 2	Peaks	<input type="checkbox"/>	eGFP-tbx-2 ChIP-seq of whole organism peaks rep2 ENC238YNR - ENCFF039LLO
<input checked="" type="checkbox"/> full	whole organism eGFP-tbx-2 Replicate 2	Signal	<input checked="" type="checkbox"/>	eGFP-tbx-2 ChIP-seq of whole organism signal of unique reads rep2 ENC238YNR - ENCFF825WJF

3 of 7 selected

Submit

18) Shown are the signal tracks for each replicate (indicated by the red arrows) and the “Optimal IDR thresholded peaks” track as black bars above (indicated by the black arrow).





- 19) To view all the available raw and processed files for the dataset, click on the blue arrow next to the L1 stage under TBX-2. As you scroll over each file type, a brief description appears. The description for alignments is shown as an example. To download individual file types, click on “All files for this DataType” in the download column or you can download “All Datatypes for this Stage” or “All Stages for this Gene.”

tbx-2	TF		org-1					WormBase	All Stages for this Gene
L1 larva	Stage	OP159		View	ENCSR238YNR	ENCSR484UJP	UCSC		All DataTypes for this Stage
alignments	These files contain the sequence reads aligned to the genome with an alignment score greater than 30.								All Files for this DataType
control normalized signal	DataType								All Files for this DataType
optimal idr thresholded peaks	DataType								All Files for this DataType
peaks	DataType								All Files for this DataType
read-depth normalized signal	DataType								All Files for this DataType
reads	DataType								All Files for this DataType
signal of unique reads	DataType								All Files for this DataType

- 20) A list of all the files and their descriptions can be found under File Descriptions (highlighted in orange below).

Fly Reference Genome
R5
R6

Worm Reference Genome
ce10
ce11

Project and Website Description
Worm By Gene
Worm By LifeStage
Fly By Gene
Fly By LifeStage
File Descriptions

### File name descriptions for modERN database

**Alignments:**

These files contain the sequence reads aligned to the genome with an alignment score greater than 30.

**Alignment file for rep1:** CEH-34\_OP524\_WA\_LE\_IP\_GoatV-aGFP\_Rep1.bam

**Alignment file for rep2:** CEH-34\_OP524\_WA\_LE\_IP\_GoatV-aGFP\_Rep2.bam

**Alignment file for input:** CEH-34\_OP524\_WA\_LE\_Input\_NA\_Rep1.bam

**Control normalized signal:**

Replicates and pooled replicates with input subtracted to view signal tracks in UCSC genome browser or IGB. Allows you to view signal above input.

**Rep1 input subtracted:**  
CEH-34\_OP524\_WA\_LE\_IP\_GoatV-aGFP\_Rep1.tagAlign\_VS\_CEH-34\_OP524\_WA\_LE\_Input\_Rep0.tagAlign.bw

**Rep2 input subtracted:**  
CEH-34\_OP524\_WA\_LE\_IP\_GoatV-aGFP\_Rep2.tagAlign\_VS\_CEH-34\_OP524\_WA\_LE\_Input\_Rep0.tagAlign.bw

**Pooled replicates input subtracted:**  
CEH-34\_OP524\_WA\_LE\_IP\_Rep0.tagAlign\_VS\_CEH-34\_OP524\_WA\_LE\_Input\_Rep0.tagAlign.bw