### Browser Exercises - I

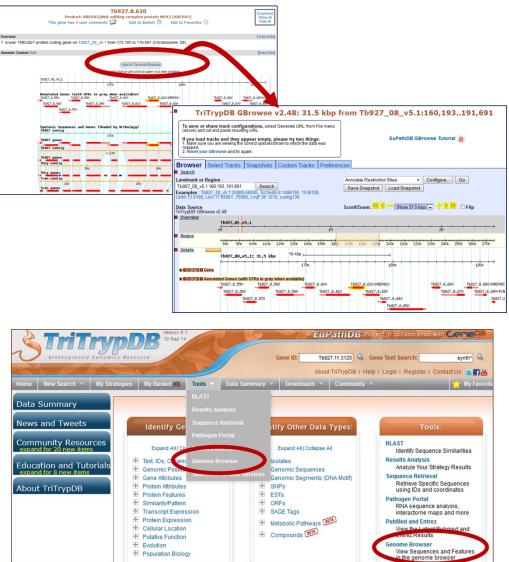
### Alignments and Comparative genomics

### Navigating to the Genome Browser (GBrowse) Note: For this exercise use http://www.tritrypdb.org

### a. Navigate to GBrowse from TriTrypDB.

From record pages, like a gene page, genomic sequence or EST page, click on the "View in Genome Browser" link. You can also use the Tools section on the homepage or the grey toolbar in the header section

- Go to GBrowse from the TriTrypDB home page. Explore this page – take note of the different sections: Instructions, Search, Overview, Region, Details, Tracks,



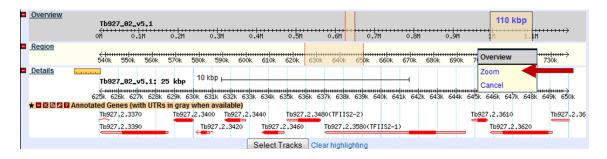
etc...

### b. Explore a genomic sequence and the annotation track in GBrowse.

- Look at the "Landmark or Region" box. What information does the "Landmark or Region" box contain? The Landmark or Region box should read – Tb927\_02\_v5.1:625,000..650,000.
- What chromosome is displayed?
- What location of the chromosome is displayed?

Browser Select Tracks Snapshots C	Custom Tracks Preferences					
search						
Landmark or Region : Annotate Restriction Sites						
Tb927_02_v5.1_625.000.650.000 Search Save Snapshot Load Snapshot						
Examples : Tb927_02_v5.1:25000-55000, TcChr40-S	-S:148613015 6120.					
Lt M.13.0190, LinJ.17:6500175000, LmjF.05.1210, (						
Data Source	Scroll/Zoom: < < 🔤 Show 25 kbp 🖃 🕂 🔀 🗔 Flip					
Tri ypDB GBrowse v2.48						
Tb927_02_v5.1	олан о.4и о. <del></del>	>				
	0.3M 0.4M 0.5M 0.6M 0.7M 0.8M 0.9M 1M 1.1M					
540k 550k 560k 570k 580	i80k 590k 600k 610k 620k <mark>630k 640k 65</mark> 0k 660k 670k 680k 690k 700k 710k 720k 73	30k				
Details	10 kbp					
10327_02_VJ+1; 25 KUP						
625k 626k 627k 628k 629k 630k	<del></del>	<del>, </del>				
★ 🖬 🖾 🖾 🖾 Annotated Genes (with UTRs in gray whe	vhen available)					
Tb927.2.3370 Tb927.2.3	2.3400 Tb927.2.3440 Tb927.2.3480(TFIIS2-2) Tb927.2.3610	Tb927.2.36				
Tb927,2.3390	Tb927.2.3420 Tb927.2.3460 Tb927.2.3580(TFIIS2-1) Tb927.2.3620					
	Select Tracks Clear highlighting					

- Move to a different genomic region on this chromosome. For example, visit the right arm of this chromosome.
  - Hint: change the coordinate numbers in the "landmark or region" box to correspond to an area in that region. Look at the overview to give you an indication of the total size of this chromosome, ie. 1000000..1100000).
  - **OR** highlight the area representing approx. 1000000-1100000 on the scale in the Overview section and then choose zoom from the popup.



- Move to chromosome 9. How did you do this?
  - Hint: Change the chromosome number in the "landmark or region" box. It should look like this: Tb927\_09\_v5.1:1,000,000..1,100,000. The Tb927 09 v5.1 portion is the genomic sequence ID for chromosome 9.
- Zoom in to a 20Kb region. Select 20Kb from the Scroll/zoom drop down menu.

Browser Select Tracks Snapshots Custon	n Tracks Preferences		
Search			
Landmark or Region:	Annotate	Restriction Sites -	Configure Go
Tb927_09_v5:1,000,0001,100,000 Search	Save S	napshot Load Snapshot	
Examples: Tb927_02_v4:25000-55000, Tc00.10470535076- LbrM.13.0190, LinJ.17:6500175000, LmjF.05.1210, contig	41.290,	,,	
Data Source TriTrypDB GBrowse v2.48	Scroll/Zoo	om: K K Show 100 kb Show 1 Mbp	p 🚽 🕂 😕 💴 🗆 Flip
<u>Overview</u>		Show 500 kb	n
Tb927_09_v5		Show 200 kb	
OM i i i i i i i i i i i i i i i i i i i	M	2 <sup>11</sup> Show 100 kb	
Region		Show 50 kbp	
950k 960k 970k 980k 990k 100 <mark>0k</mark>	1010k 1020k 1030k 1040k 1050k	1060k 1070k 1 Show 20 kbp	1110k 1120k 1130k 1140k 1150k
Details		Show 5 kbp	
Tb927_09_v5: 100 kbp		Show 2 kbp	
1000k 1010k 1020k	1030k 1040k 1050k	1060k Show 1 kbp	1080k 1090k 1100k
★ 🖬 🕅 🖾 🕼 Annotated Genes (with UTRs in gray when availa		Show 200 bp	
Tb927.9.5940 Tb927.9.6090 Tb927.9.615			Tb927.9.6570 Tb927.9.6670
Tb927.9.5950 Tb927.9.6100 Tb927.9	9.6160 Tb927.9.6290 Tb927.9.6360	Tb927.9.6420 Tb927.9.64	70 Tb927.9.6560 Tb927.9.6620
Tb927.9.5960 Tb927.9.6110 Th	927.9.6170 Tb927.9.6350 Tb	927.9.6410 Tb927.9.6460 Tb9	
Tb927.9.5970 Tb927.9.6120		370 Tb927.9.6430 Tb927.9.6	
Tb927.9.5990 Tb927.9.6130	Tb927.9.6310 Tb927.9	9.6380 Tb927.9.6440 Tb927.9	.6500 Tb927.9.6600
Tb927.9.6010 Tb927.9.6140			b927.9.6530 Tb927.9.6650
Tb927.9.6020 ➡→	<b>~</b>		Tb927.9.6660
H→ Tb927.9.6040			<b>·→</b>
Tb927.9.6030			
↔ Tb927.9.6050			
←■ Tb927.9.6060			
Tb927.9.6070			
	Select Tracks Clear high	lighting	

- What genes are in this region? Mouse over the gene graphics and look at the popups.
- Explore the ruler tool. Click on the ruler to engage then drag it across the window. The ruler tool displays the nucleotide coordinates of the ruler's solid center line. This is very useful for comparing between the annotation data track and others that we will add later.

Browser	Select Tracks	Snapshots	Custom Tracks	Preferences				
Search								
Landmark or Region :					notate Restriction Si	ites 👻 Co	onfigure Go	
Tb927_09_v5.1	1:1,040,0001,059,9	99 5	earch	S	we Snapshot Loa	d Snapshot		
Examples : Tb927_02_v5.1:25000-55000, TcChr40-S:14861301536120, LbrM.13.0190, LinJ.17:6500175000, LmjF.05.1210, contig136.								
Data Source TriTrypDB GBro	owse v2.48			Scr	oll/Zoom: <mark>&lt;&lt; &lt;</mark> 🗕	Show 20 kbp	- <mark>- &gt; &gt;&gt;</mark> -	lip
Overview	Tb927_09_v	5.1			+ + + + + + + + + + + + + + + + + + +		-+ -+ -+ -+ -+	
Region	<del>()</del> ek 960k	970k 980k 99	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	0k 1030k 1040k	1050k 1050k 1070k	1080k 1090k 110	 00k 1110k 1120k 1	
Ditails	7 327_09_v	5.1: 20 kbp	5 kbp 🛌	1049				
	1040k 1041k 1	.042k 1043k 104	4k 1045k 1046k 104	7k 1048k 1049k	1050k 1051k 1052k	1053k 1054k 105	55k 1056k 1057k 1	1058k 1059k 1060k
	Annotated Genes (wi			E	-			
	Tb927.9.632	0 Tb927	.9.6360 Tb927.9.6370	Tb927.9.6380	b927.9.6390	Tb927.9.6	420	Tb927.9.6440
		.6350(IMPase)		E	Tb927.9.641	0	16927.9.6430	Tb927.9.645
			Se	lect Tracks Clea	rhighlighting			

- There are other ways to move and zoom. Try highlighting an area along the scale in the overview, region or details sections of GBrowse.

Browser	Select Tracks	Snapshots	Custom Tracks	Preferences		
Search     Search						
Data Source TriTrypDB GBr	owse v2.48			Scroll/Zoom: ≤ 🗲 Show 20 kbp 💽 🕂 Ӯ 🖉 Flip		
Overview	Tb927_09_v	/5.1	430 kbp	· · · · · · · · · · · · · · · · · · ·		
Region	<del>(</del>	970k 980k 99	ok 1000k 1010	view		
■ <u>Details</u> .		<b>75.1: 28 kbp</b>	5 kb Canc 4k 1045k 1046k 104			
* = × 6 2 2 /	Annotated Genes (wi Tb927.9.632	th UTRs in grav	when available)	Tb927.9.6380 Tb927 P.6440 Dump selection as FASTA 00 Tb927.9.6440 Units selection to UCSC BLAT Submit selection to VCSC BLAT Submit selection to NCBI BLAST		

- What if you want to go to a specific gene in Gbrowse? Try to figure out how to go to this gene: Tb927.2.5800
- Type the ID in the "landmark or region" box. The landmark box has a search function that supports gene IDs. What else does it support?
- What is this gene?

### 2. Exploring data tracks in GBrowse

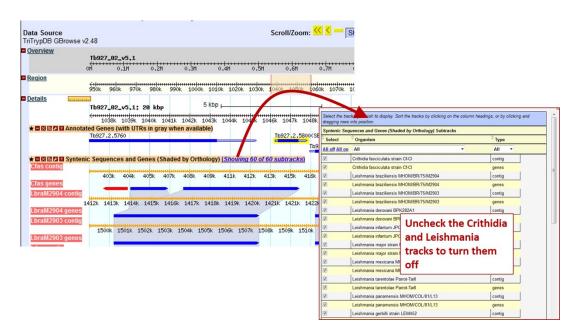
## a. Is the region containing the sedoheptulose-1,7-bisphosphatase (SBPase) gene syntenic in all kinetoplastids?

- Go to the "Select Tracks" section and turn on the track called "Syntenic Sequences and Genes". The browser is automatically updated with tracks you select. Note that this track contains multiple subtracks.

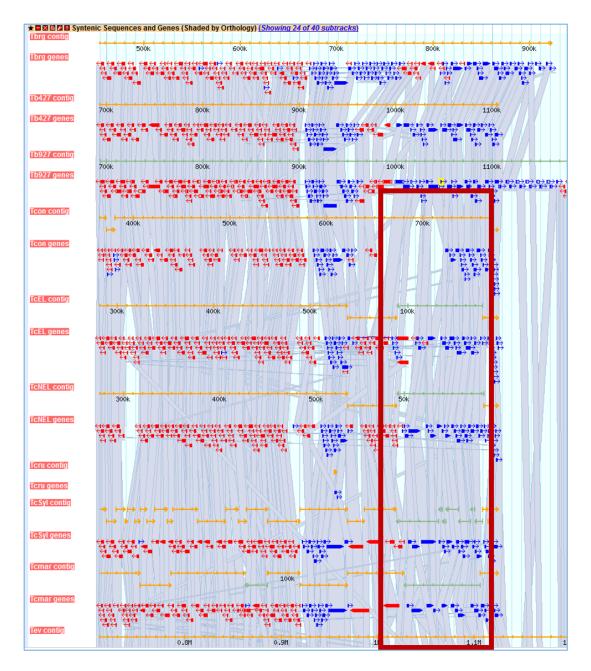
File	e → Help →					
•	TriTrypDB GBrowse v2.48: 20 kbp from Tb					
	To save or share track configurations, select 'Generate URL' from File menu (above) and cut and paste resulting URL If you load tracks and they appear empty, please try two things: 1. Make sure you are viewing the correct species/strain to which the data was mapped. 2. Reset your GBrowse and try again.					
Bro	owser Select Tracks Snapshots Custom Tracks Preferences					
	ack to Browser Show Active Tracks Only Show Favorites Only <u>Gene Models</u> ■ <u>A. Annotated</u> ■ <u>All on</u> All off					
	Annotated Genes (with UTRs in gray when around 00 [7]					
	B. Synteny All on All off     Syntenic Sequences and Genes (Stated by					
	<u>C. Amino Acid Analysis</u> All on All off					

- Return to the browser by clicking the "Browser" tab and **zoom out to 20Kb**.
- What does this region look like?
- What direction is the SBPase gene relative to the chromosome?

- What genes are upstream and downstream of the SBPase?
- Modify the subtracks to remove all species apart from *Trypanosoma* from the view. Click on the link 'showing 70 of 70 subtracks', wait for the popup and uncheck all the species you do not want to see. **Then click Change**.



- Examine the gene corresponding to the *T. vivax* SBPase in the synteny track. Hover the image to find the gene name in the popup. Why is it a fragment? What could be some possible reasons for this?
- Zoom out to 50KB. Look at the genomic sequence for *T. congolense* why does the synteny look like it does?
- Zoom out to 500KB what could you conclude about this region in *T. congolense*? (See image on next page if needed).
- You will also notice that some of the genomes have contigs that are not contiguous. Why is that?
- Mouse over the two contigs and look at the information in the popups do these pieces belong to the same chromosome? What does this mean?

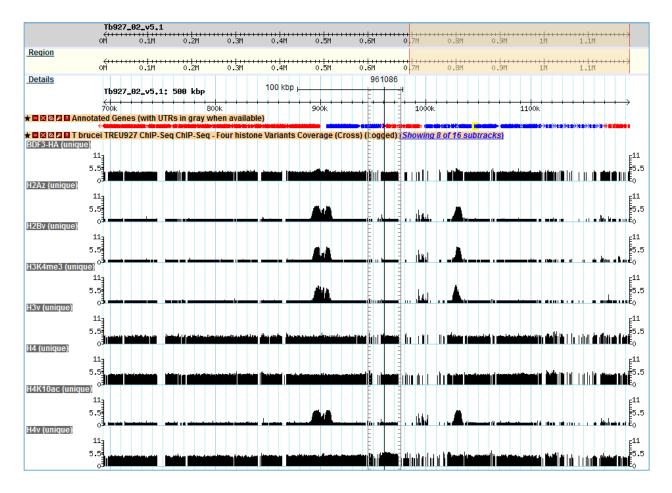


### b. Exploring other data tracks in Gbrowse.

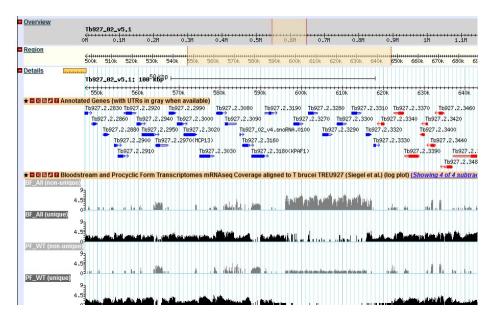
In this example we are viewing *T. brucei* SBPase, so the data tracks you turn on will display data only if the data is aligned to the *T. brucei* genome.

Turn on the ChIP-seq coverage plots and turn off the syntenic gene and region tracks. The data track is called: ChIP-Seq - Four histone Variants ChIP-Seq Coverage aligned to T brucei TREU927 (Cross) (log plot). For this experiment, chromatin was immunoprecipitated using several different histone antibodies. The DNA that precipitated with the histone was sequenced and aligned to the T. brucei TREU927 genome. Peaks in the sequence coverage plots represent areas of histone binding and transcription start sites. <u>http://www.ncbi.nlm.nih.gov/pubmed/19369410</u>

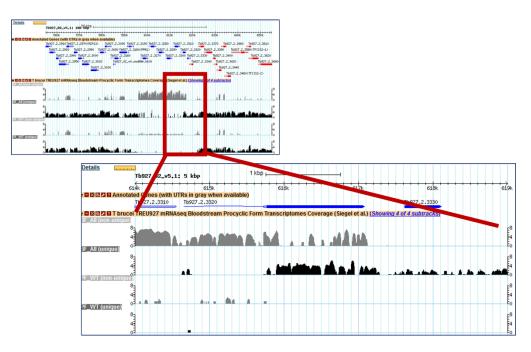
- What does this data show you?
- Roughly how many polycistronic units does this chromosome have? Zoom out to the entire chromosome.
- Do the ChIP-seq peaks correlate with the direction of gene transcription (blue vs. red)?



- Now zoom back to 50Kb. Turn off the ChIP-Seq tracks and turn on the track called: Bloodstream and Procyclic Form Transcriptomes mRNAseq Coverage aligned to T brucei TREU927 (Siegel et al.) (log plot)(4 subtracks).
- Move to the **region around 0.6Mbs of the chromosome** (you should be on chromosome 2) and turn on all four subtracks. Take note of the black and grey bars in the coverage plots. What do you think the grey bars indicate?

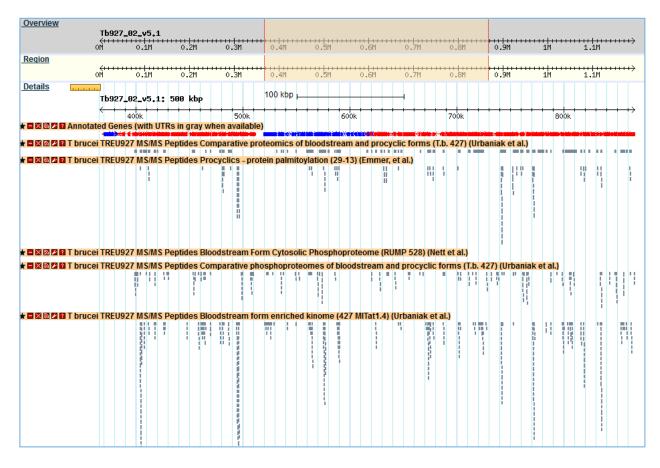


- Now zoom out to 100Kb do you see a difference between the blood forms and procyclics?
- Zoom in to a gene that looks like it is differentially expressed. What are your conclusions? Are the reads supported by unique or non-unique reads?



- Can you turn on additional tracks that may give some more support to your conclusions?
   Hint: turn on the EST and *T. brucei* protein expression evidence tracks.
  - Is there any proteomics evidence for this region?
  - How about EST evidence? Click on an EST graphic (glyph) to get additional information.

- Turn off the RNA-seq graphs and make sure the *T. brucei* protein expression evidence tracks are on. **Zoom out to 500Kb**. Explore the evidence for gene expression based on mapped peptides from proteomics experiments – which gene in this view has the highest number of peptide hits?



# 3. Downloading data from GBrowse and uploading your own tracks to GBrowse

•	TriTryp	DB GBrowse v2.39: 20 kbp from T	b927_02_v4:602,523622,522
NOTE: If you load 1. Make sure you 2. Reset gbrowse	tracks and they appear emp are viewing the correct spec by clicking on the red Reset	by, you can try two things to resolve this issue: issistrain to which the data was mapped. link, then try again.	
Browser Se	elect Tracks Custom T	racks Preferences	
Search			
Landmark or R	egion:	Annotate	Restriction Sites Configure Go
Tb927_02_v4:61			Restriction Sites
Examples: Tb9	27_02_v4:1-55000, Tc00.1 75000, LmjF.05.1210, cont		Descended EASTA File
Linj. 17.05001	75000, Linjr.05.1210, con	Download Download	nanaric and a second se
Data Source		Caralian Tanana kanal TDE 1007	x   Show 20 kbp - Show 20 kbp
TriTrypDB GBro EuPathDB GBro		Species: Trypanosoma brucei TREU927	-
Overview	wse rutorial	ID: Tb927.2.3280	
	Tb9:		
	OM	Description: 65 kDa invariant surface glycoprotein (10C8.4	40) m 0.8M 0.9M 1M 1.1M
Region		Location: join(	
	<del>(****</del> 5	600482601593, 601594602904)	x 640k 650k 660k 670k 680k 690k 700k 710k
Details	and the second s	UTR: 3 600482601593	
	Tb9:		
	<l< th=""><th>Boelister Logiste sove genes.</th><th>615k 616k 617k 618k 619k 621k 622k</th></l<>	Boelister Logiste sove genes.	615k 616k 617k 618k 619k 621k 622k
	* B X B Z B Frame U an	Download: CDS I protein	by the second se
		C Brown Latene Page	Tb927_02_v4:619400620299 (X)
	* E 🛛 🖬 🖾 🖻 Annotated G	with UTRs in gray when available)	Zoom in
	Tb92	Z.3280 Tb927.2.3300 Tb927.2.33	10 Descent of the spice
	-	Tb927.2.3290	Dump selection as FASTA .3340
		Select Tracks Clear highlig	abting Oddina deletion to boot BLAT
L		Select Tracks Clear highlig	Submit selection to NCBI BLAST

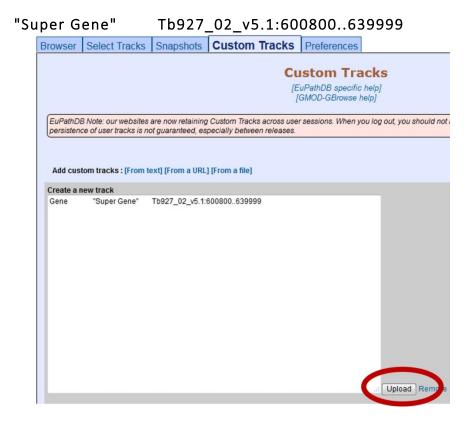
You can download data from GBrowse in multiple ways and formats.

 The Report and Analysis drop down menu allows you to select a format for the download file that will contain the all the features that you have displayed in the region you are viewing.
 Highlighting a section of the Details scale allows you to retrieve a FASTA dump of the nucleotide sequence from this region. You can also use this same tool to submit a sequence to NCBI Blast.

3. Mousing over a gene will reveal a popup window with the option to get the coding (CDS) or amino acid sequence of that gene.

- Uploading your own tracks is also possible. One reason to upload your own tracks is to display your own data on a chromosome or genomic segment and view it in the context of gene models and other data. To do this you have to follow some rules to ensure that the file you are uploading can be understood by GBrowse. In this exercise we will only go through a couple of simple examples to give you an idea of the possibilities. There are many online resources if you wish.
- Imagine that you have cloned a new gene and you would like to display it in GBrowse. Click on the "Custom Tracks" GBrowse tab and add a custom track "From text".

There are many types of formats that can be used. For this example we are going to tell GBrowse that we have a gene and a few things about the gene, like its location. Paste the following into the editor (the next window after you click on Add Custom Tracks From Text), and then click on upload (hint: sometimes you have to zoom in or out a little to see your new glyph):

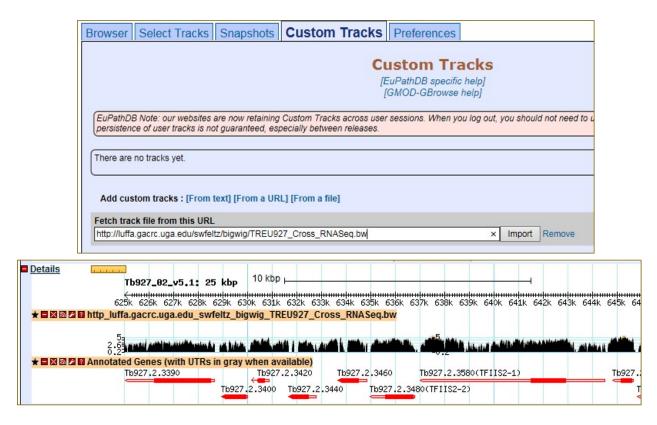


You should see a new track with your gene displayed.

	Tb927_02_v5.1:			····						····· <u>·</u> 2
	50k 560k	570k	580k	590k	600k	610k	620k	630k	640k	650k
🗕 🛛 🖾 ピ Gene					Super	Gene				
🗙 🗟 🖉 🛛 Annotat	ed Genes (with UTF	ts in grav when a	vailable)							
	Tb927.2.2880 Tb927	.2.2940 Tb927.2	.3000 Tb927	.2.3090 Tb927.	2.3190 Tb927.	.2.3280 Tb92	7.2.3310 Tb92	27.2.3370 Tb	927.2.3460	Tb927.2.3
	Tb927.2.2900	Tb927.2.2970(M	CP13)	927.2.3160	Tb927.2.32	70 Tb927.2.3	300 Tb927.2.33	30 Tb927.2.:	3400	Tb927.2.
	Tb927.2.291	0 Tb927.2.29	90 Tb927.2.30	080 Tb927.2.318	30(PPR1) T	b927.2.3290	Tb927.2.3320	Tb927.2.3390	Tb927.2.358	30CTFIIS2
	Tb927.2.2	920 Tb927	.2.3020 1	[b927_02_v4.sn	oRNA.0100		Tb927.2	.3340 Tb927.:	2.3420	Tb9
	Tb9	27.2.2950	Tb927.2.3030				11117		.2.3440	
		▶	-					<b>←</b>	Tb927.2.3480(	TFIIS2-2

 Now let us load a more complex graphic, a bigwig file of some RNA Sequencing data. For this we posted the file to a public site and are using the URL to direct GBrowse to the file location. In the field "Fetch track file from this URL", enter the following and click Import: http://luffa.gacrc.uga.edu/swfeltz/bigwig/TREU927\_Cross\_RNASeq.bw

Gene



### 4. Designing PCR Primers with GBrowse

Open GBrowse at the genomic location for your new primers.

- Go to gene page of a gene you want to design primers for and use the 'View in GBrowse' button. This example is using SBPase, Tb927.2.5800 and we have zoomed out to 20K.
- OR open GBrowse from the home page and then enter genomic coordinates of your region in the landmark region.

Choose "Design PCR Primers" from the drop down menu and then **click GO**.

- This opens the Design Primers application.

Browser   Select Tracks   Snapshots   Custom Tracks   Prefere	ences			
Search				
Landmark or Region :           Tb927_02_v5.1:1,037,0131,057,012         Search           Examples : Tb927_02_v5.1:25000-55000, TcChr40-S:14861301536120,           LbrM.13.0190, LinJ.17:6500175000, LmjF.05.1210, contig136.	Annotate Restriction Sites Design PCR primers Download Decorated FASTA File Download Track Data			
Data Source TriTrypDB GBrowse v2.48	Scroll/Zoom: K 🤇 — Show 20 kbp 🔽 🕂 🔀 💴 🗆 Flip			
■ <u>Overview</u> Tb927_68_v5.1 (++++++++++++++++++++++++++++++++++++	n 0.6N 0.7N 0.8N 0.9N 1n 1.1N			
Region 550k 960k 970k 980k 990k 1000k 1010k 1020k 1030	4 1040K 1050K 1050K 1070K 1080K 1090K 1100K 1110K 1120K 1130K 1140K			
Details Tb927_82_v5.1: 28 kbp 5 kbp				
1038k 1039k 1040k 1041k 1042k 1043k 1044k 1045k	1046k 1047k 1048k 1049k 1050k 1051k 1052k 1053k 1054k 1055k 1056k 1057k			
★ 🛛 🖾 🗗 🖓 Annotated Genes (with UTRs in gray when available)				
Tb927.2.5760	Tb927.2.5800(SBPase) Tb927.2.5820			
	Tb927.2.5810			

Choose a target:

- The graphic is interactive. To choose a target, highlight an area on the scale. You can zoom in with the controls in the upper left corner. The PCR primers that you design with this application will flank the shaded region.
- Once you choose a target, the Product size range is automatically updated in the parameter table at the bottom of the page.
- You can choose to customize the primer design using other parameters.

Click DESIGN PRIMERS to run the application.



#### Targetting information

- PCR primers will flank the shaded region.
  Click on a different sequence feature to change the selection
- The boundaries of the shaded target region can be adjusted by clicking on the lower scalebar
   The size of potential PCR products can be adjusted via the 'Product size range' option below

