# Browser Exercises - I Alignments and Comparative genomics

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

#### a. There are two ways to 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





- **b. Go to GBrowse from the TriTrypDB home page.** Explore this page take note of the different sections: Instructions, Search, Overview, Region, Details, Tracks, etc...
- c. 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?
- 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.
- Zoom in to a 20Kb region. Select 20Kb from the Scroll/zoom drop down menu.

Browser Select Tracks Snapshots	Custom Tracks Preferen	ces		
Search	5			
Landmark or Region:		Annotate Restriction Sites	- Co	nfigure Go
Tb927_09_v5:1,000,0001,100,000 Search		Sava Snanshot	Spanshot	
Examples: Tb927_02_v4:25000-55000, Tc00.10470 LbrM.13.0190, LinJ.17:6500175000, LmjF.05.1210,	53507641.290, , contig136.		onapanot	
Data Source TriTrypDB GBrowse v2.48		Scroll/Zoom: < < 💻	Show 100 kbp 💌	i 🕂 ڬ 📂 🗖 Flip
<u>Overview</u>			Show 500 kbp	
Tb927_09_v5			Show 200 kbp	
or i i i i i i i i i i i i i i i i i i i	11	24 2	Show 100 kbp	3M
Equin			Show 50 kbp	
950k 960k 970k 980k 990k	1000k 1010k 1020k 1030k 10	040k 1050k 1060k 1070k 1	Show 20 kbp	1110k 1120k 1130k 1140k 1150k
Details			Show 5 kbp	
Tb927_09_v5: 100 kbp <sup>kbp</sup> ⊢			Show 2 kbp	
1000k 1010k 1020k	1030k 1040k	1050k 1060k	Show 1 kbp	1080k 1090k 1100k
★ 🖬 🖾 🖾 🖾 Annotated Genes (with UTRs in gray whe	n available)		Show 200 bp	
Tb927.9.5940 Tb927.9.6090 Tb9	27.9.6150Tb927.9.6230 Tb927.9	.6320 Tb927.9.6390 Tb927.	Show 100 bp	Tb927.9.6570 Tb927.9.6670
Tb927.9.5950 Tb927.9.6100	Tb927.9.6160 Tb927.9.6290 Tb	927.9.6360 Tb927.9.6420	Tb927.9.6470 T	16927.9.6560 Tb927.9.6620
Tb927.9.5960 Tb927.9.611	.0 Tb927.9.6170 Tb927	.9.6350 Tb927.9.6410 Tb927	7.9.6460 Tb927.9.	6540 Tb927.9.6590
		Tb927.9.6370 Tb927.9.6430	→ ↔ Tb927.9.6490	■→ Tb927.9.6580
↔ DE27 9 5090 Tb927 9	9 6130 Th027 9 6310		€1 5440 Tb927 9 6500.	Tb927 9 6600
	.9.6140	₩		
Tb927.9.6020 ↔				Tb927.9.6660
Tb927.9.6040				
Tb927.9.6030				
Tb927.9.6050				
Tb927.9.6060				
→ Tb927.9.6070				
	Select Tracks	Clear highlighting		

- 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.
- 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								
Landmark or	Region :			An	notate Restriction Site	es 👻 Config	gure Go	
Tb927_09_v5.1	1:1,040,0001,059,9	999 S	earch	Sa	ve Snapshot Load	Snapshot		
Examples : The LbrM.13.0190,	927_02_v5.1:25000 LinJ.17:65001750	-55000, TcChr40 00, LmjF.05.121	-S:14861301536120 ), contig136.	),				
Data Source				Scr	oll/Zoom: < < 💻	Show 20 kbp 👻	🕂 ڬ 😕 🗆 Fiij	р
TriTrypDB GBr	owse v2.48					,		
Overview	Tb927_09_v	v5.1				- · · · · · ·		·
Region	<del>&lt;</del>	970k 980k 99	0k 1000k 1010k 102	0k 1030k 1040k	1050k 1050k 1070k		1110k 1120k 113	) 0k 1140k 1150k
Ditails	T1927_09_0	/5.1: 20 kbp	5 kbp	10493	87			
	1040k 1041k 1	L042k 1043k 104	4k 1045k 1046k 104	7k 1048k 1049k	1050k 1051k 1052k	1053k 1054k 1055k	1056k 1057k 105	8k 1059k 1060k
	Annotated Genes (wi	th UTRs in gray v	vhen available)	Th027 0 6200 T		Th007 0 6400		Th027 0 6440
	Tb927.9	.6350(IMPase)	.9.6360 10927.9.6370	10927.9.6380	Tb927.9.6410	Tb927.9.8420	27.9.6430	Tb927.9.6440
			Se	lect Tracks Clear	highlighting			

Browser	Select Tracks	Snapshots	Custom Track	s Preferences		
Search Landmark or Tb927_09_v5.1 Examples : Tb LbrM.13.0190,	<b>Region</b> : I:10790001111999 v927_02_v5.1:25000 LinJ.17:65001750	-55000, TcChr40 00, LmjF.05.121	<mark>earch</mark> -S:14861301536 ), contig136.	An Si 120,	notate Restriction Sites	Configure Go
Data Source TriTrypDB GBr	owse v2.48			Sci	oll/Zoom: < < 🔤 Show 20 k	ibp 💽 🕂 🔀 💴 🗆 Flip
Overview	Tb927_09_v	5.1	130 kbp		+ + +   + + + + + + + + + + + + + + + +	-+-+-+-++-++++++++++++++++++++++++++++
Region	<del>()</del> 950k 960k	970k 980k 99	0k 1000k 1010	verview	1050k 1050k 1070k 1080k 1090k	1100k 140k 1120k 1130k 1140k 1150k
<u>Details</u>	<b>Tb927_09_v</b> <del>(111111)</del> 1040k 1041k 1	5.1: 20 kbp	5 kbj 6 4k 1045k 1046k 1	ancel 1047k 1048k 1049k	Tb927_09_v5.1:1079000111	1999 X 1057k 1058k 1059k 1060k
	Annotated Genes (wi Tb927.9.632 Tb927.9	th UTRs in gray v ) Tb927 .6350(IMPase)	vhen available) .9.6360 Tb927.9.6	370 Tb927.9.6380	Becenter on this region Dump selection as FASTA	Tb927.9.6440
				Select Tracks Clea	r high Submit selection to UCSC BL Submit selection to NCBI BLA	IST

- **d.** 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?

Hint: 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.



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

- What direction is the gene relative to the chromosome?
- What genes are upstream and downstream of the SBPase?
- Modify the subtracks to remove *Leishmania* species from the view. Click on the link 'showing 40 of 40 subtracks', wait for the popup and uncheck all *Leishmania* and *Crithidia* species. 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 regions 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).

- What does this data show you?

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.

http://www.ncbi.nlm.nih.gov/pubmed/19369410

- 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: T.b.
   mRNAseq Blood Form and Procyclic Form Coverage (Cross) (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, 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?

<u>Overview</u>																																												
		Tb9	27	.02	_v5.	<b>.1</b>																																						~
		omiii	••••	Ċ	).1M	• • •		0.2	2M		0.	.зм		0	.4M			0.	5M			0.6	5M			o.7M			0.	8M		i i i	0.9	1		1	1M .			1.	1M			· /
Region														1																														
		_ <del>≦++</del>	••••	÷÷	+   + \ 1 M	• • •	•••	벐	H H H		÷	1++ 3M		1	-  // M		++++	11	HH EM	++++		H C C C	HH I M			+   + 0. 7M	++++	++++	+	em l	++++	1 + + ;		HH M	+++	++ <u>+</u>	+++ 1 M	-++	+++	1	1 M	+++		↔
Dotoilo		VII		``	·			0.2			ν.	JII		1 4	. 411			v.	JI			0.0				2.711			v.	on			0.51			-				1.	T11			
Details	l.l.l.l.l.	Tb9	27.	.02	_v5	.1:	50	0 1	сbр					10	) kb	P ⊢	+	_	+	-	-			+																				
		←	- 1	400	)k	+	+ +	-	-	+	-	50	)k	+	-	+	+	+	+	60	bok	-		-	+	+	-	-	70	0k	+	+	+	-	$\vdash$	-	-	1005	k t	+	+	+	-	↔
* = × 🔊 🖉	Annota	ted G	ene	es (	with	UT	'Rs i	n a	rav v	vhe	n a	vail	able	5															• •										`					
		-	0-¢#	-	-				-	1	-		-	-		) and a					-00	00		-	-	- 10	-	- 46				-	-	-	•	-	-	2 <b>1</b> 1 1	-	****	atio:	-		
* 🗆 🗙 🔊 🖉	T bruce	ei TRE	U9	27	MS/I	MS	Рер	tide	es Co	pmp	Dara	ativ	e pro	teo	mic	S O	fbl	000	dstr	rear	m a	nd	рго	сус	clic	for	ns (	T.b	. 42	27)	(Urt	bani	iak	et a	il.)							Ι.		
	T bruce	i TRF	-119	27	MS/	MS.	Pen	tide	s Pr	00	clie	CS.	prof	ein	pal	mite	ovla	atio	n C	29-1	13)	(Em	nme	er, e	et a	D		1.			Т	·   .	E.		[ ]				T		11	1.		
					I I				1					1	Pull		1		. 1		,				1	.,		1					1	11	ų į				1	10				Ļ.
					- i						1							'	1	Ϊ.							"1	Ľ					i.	1.					1	Υ.	11			i -
											ι.								i														ł	ł.							1			
																																				ł.								
												Ū.																					į.											
																																	÷.											
																																	ł.											
																								_									1											
	I bruce	HIRE	:09	271	MS/I	MS	Рер	tide	s Bl	000	Istr	ear	n Foi	m (	.yto	SOI		'no	spi	юр	rote	eon	ne (	RU	MP	528	5) (N	ett	et	al.)														
* 🗆 🛛 🖾 🖉	T bruce	ei TRE	U9	27	MS/I	MS	Рер	tide	es Co	pmp	Dara	ativ	e pho	ospi	hop	rote	eon	nes	s of	blo	ods	stre	am	an	d p	FOC	/clic	: fo	rm	s (T	.b. 4	427)	) (U	rba	nia	k ef	t al.	1						
				j.	1	11			31E					•••	Ţ.	- 11		1	'					1		1	1				η.		ι	п	H.		1		н.		1			1
						1																	H		i								<u>۱</u>		H.			1	Π.		11	11	1 i	1
									·   -			1																							ų.								11	
	Thruce			271		II C	Don	tide		0.04	letr	0.01	n for	mo	nrie	ho	d ki	ino	mo	142	7 1	шта	+4		lich	ani	alk of	tal													1			
			.09	211	13/1	13	Pep	uue	5 DI	1 1	50	eai								(42		1110	1	+) (1		aille	in e	1 01	.,					τ.			н.				10		-	1
					Ш÷.	11				1		ij.		Π.	- I		13			٩.					1		17.			ï.	10					i.	4	-11	ŝ, i	11				
					3i -				- 12		ü (			1.					<u>8</u> –	1			i				131	i		i						ï I		11	i [] []	ij.			11	1
					11				1			į.							§								1			ł.		i	i.			i I					- i			1
					11													1	3								ł.						1			i I				1				
					i I														2								ł.						÷.											
					i							1																												1				
																			ι.																	i .								
																			1																						- É			
																																				i l					ļ			
												Ş.																													1			
					1							15									1																							

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

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 if you have data that you would like to display 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): **Gene** "Super Gene" Tb927\_02\_v5.1:600800..639999



You should see a new track with your gene displayed.

Details	Tb927	2_02_v5.1: 100	<sup>5</sup> նե <sup>թը</sup> ——								
	550k	560k	570k	580k	590k	600k	610k	620k	630k	640k	650k
						Super	Gene				
★ 🗖 🛛 📾 🗖 🖬 Anno	tated Gen	ies (with UTRs ir	gray when a	/ailable)							
	Tb927.	.2.2880 Tb927.2.	2940 Tb927.2.	3000 Tb92	7.2.3090 Tb927	.2.3190 Tb927	.2.3280 Tb92	7.2.3310 Tb92	27.2.3370 Tb9	27.2.3460	Tb927.2.3610
		Tb927.2.2900 Tk	927.2.2970(M	(P13) T	b927.2.3160	Tb927.2.32	70 Tb927.2.3	300 Tb927.2.33	30 Tb927.2.3	400	Tb927.2.36
		Tb927.2.2910	Tb927.2.29	90 Tb927.2.	3080 Tb927.2.3:	180(PPR1) T	b927.2.3290	ть927.2.3320	Tb927.2.3390	Tb927.2.3	580(TFIIS2-1)
		Tb927.2.2920	Tb927	.2.3020	Tb927_02_v4.s	noRNA.0100		Tb927.2.	.3340 Tb927.2	.3420	Tb927.
		Tb927.	2.2950	Tb927.2.3030	Þ			<b>+</b>	←∎ Тb927.	.2.3440	
		<b></b> →		>					T ++	b927.2.3480	(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://loquat.rcc.uga.edu/swfeltz/bigwig/TREU927\_Cross\_RNASeq.bw

Browser	Select Tracks	Snapshots	Custom Tracks	Preferences	
					Custom Tracks
					[EuPathDB specific help] [GMOD-GBrowse help]
EuPathD releases.	B Note: our websites	are now retaining	Custom Tracks across user	sessions. When you	log out, you should not need to uplo
There are	no tracks yet.				
Add cu	stom tracks : [From	text] [From a URI	.] [From a file]		
Fetch tra	ck file from this UR	Ĺ			
http://loqu	uat.rcc.uga.edu/swfel	tz/bigwig/TREU927	Cross_RNASeq.bw		Import Remove

Details				10 kbp								
	Tb927_02	_v5.1:	25 kbp	то кор	H							
	<del>( 11111 11111</del> 625k 626k 6	n <b>in minin</b> 27k 628k	. 62 <mark>9</mark> k 6	<mark></mark>	632k 63	<del>******* ******</del> 3k 634k 6		<mark>duuuuduuuud</mark> 37k 638k 639	9k 640k 641k	642k 643k 64	4k 645k 646k	<del>444/44444/4444/4444/4</del> 647k 648k 649k 650k
★ 🗖 🖾 🗟 🖉 🛍 http	p_loquat.rcc.ug	a.edu_s	wfeltz_b	gwig_TR	EU927_CI	ross_RNA	Seq.bw					
	_							E				-
		مادر الأفاذة.	de al lubr		il i siste			. فالتقالف	الأفاقية أروية	in a statistical sector	ing all the state	
★ 🗖 🖾 🗟 🖉 🖾 Ani	notated Genes	(with UT	Rs in gra	ly when a	vailable)							
	Tb927.2.33	370	Tb923	7.2.3400	Tb927.2	.3440	Tb927.2.34	BO(TFIIS2-2)			Tb927.2.30	510 Tb927.2.36
	Tb927.2.3	390		Tb927 ←■→	.2.3420	Tb927.	2.3460	Tb927.2.358	O(TFIIS2-1)		Tb92	7.2.3620

# 4. Designing PCR Primers with GBrowse

Open GBrowse at the genomic location where you want to find 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 Preferences
Search Landmark or Tb927_02_v5 Examples : Tl LbrM.13.0190, Data Source TriTorpDB GB	Region :       .1:1,037,0131,057,012 Search         .9927_02_v5.1:25000-55000, TcChr40-S:14861301536120,       LinJ.17:6500175000, LmjF.05.1210, contig136.         Scroll/Zoom:       Scroll/Zoom:
Overview	Tb927_02_v5.1
Region	
Details	Tb927_02_v5.1: 26 kbp         5 kbp           1038k         1040k         1041k         1043k         1046k         1049k         1050k         1051k         1052k         1053k         1053k         1055k         1055k
	Annotated Genes (with UTRs in gray when available) Tb927.2.5760 Tb927.2.5800(SBPase) Tb927.2.5810 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.
- 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.

< 🤇 — Show 20 kbp 🔽 🕂 🎽 🎾
Tb927_02_v5.1: 20 kbp         1038k 1039k 1043k 1042k 1042k 1043k 1044k 1045k 1046k 1047k 1046k 1047k 1046k 1047k 1046k 1047k 1050k 1051k 1053k 1054k 1055k 105
<ul> <li>Targetting information</li> <li>PCR primers will flank the shaded region.</li> <li>Click on a different sequence feature to change the selection</li> <li>The boundaries of the shaded target region can be adjusted by clicking on the lower scalebar</li> <li>The size of potential PCR products can be adjusted via the 'Product size range' option below</li> </ul>

Primer Size	Min. 18 Op	ot. 20 Max.	. 27
Primer GC%	Min. 20.0 Op	ot. Max.	80.0
D. S. Hu	WIII. 57.5	60.0 Max.	63.0
Product size range:	1050-1350		
Pillierses.			
Design Primers Res	et Parameters Retu	rn to Browser	>

24.00
9.0
8.00
3.00
5