Using the genome browser (GBrowse) Exercise 2

2.1 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.



From the tools section on the home page or the grey tool bar 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?
- What chromosome are you looking at?
- Where are you on the chromosome?

monu

- 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. 1100000..1200000).
- 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_v4:1100000..1200000).
- Zoom in to a 20Kb region. Select 20Kb from the Scroll/zoom drop down

Browser Select Tracks Snapshots Custom Tracks Preferer	nces
Search	
Landmark or Region:	Annotate Restriction Sites
Tb927_09_v4:1,000,0001,200,000 Search	Save Spapshot Load S Show 1 Mbp
Examples: Tb927_02_v4:25000-55000, Tc00.1047053507641.290,	Show 500 kbp
LbrM.13.0190, LinJ.17:6500175000, LmjF.05.1210, contig136.	Show 200 kbp
Data Source	
TriTrypDB GBrowse v2.48	Show 20 kbp
Overview	Show 10 kbp
Tb927_09_v4	Show 5 kbp
om in in in in in	Show 2 kbp
Region	Show 1 kbp
1050k 1060k 1070k 1080k 1090k 1100k 1110k 1120k 11	130k 1140k 1150k 1160k 1170k Show 200 bp 1210k 1220k 1230k 1240k 1250k
Details	Show 100 bp
Tb927_89_v4: 58 kbp 20 kbp	
TI30k 1140k	1150K 1160K 1170K
Tb09.160.5310 Tb09.160.5340 Tb09.160.5370	Tb09.160.5400 Tb09.160.5420 Tb09.v1.0300 Tb09.v1.0330 Tb09.160.545
Th09.160.5320 Th09.160.5350 Th09.160.5380	Tb09.160.5410 Tb09.160.5430 Tb09.v1.0310 Tb09.160.5440
Then 450 5220 Then 450 5260 Then 450 526	
Select T	racks 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:



- Are there 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:			Annotate Re	estriction Sites 🗘 Configure	e Go								
Tb927_09_v	4:11720001217999	9 Search		Save Snapsh	ot Load Snapshot									
Examples: Tb927_02_v4:25000-55000, Tc00.1047053507641.290, LbrM.13.0190, LinJ.17:6500175000, LmjF.05.1210, contig136.														
Data Source TriTrypDB (e BBrowse v2.48			n: < < 🗕 Show 20 kbp 🕴	🖯 🕂 🔀 😕 🗆 Flip									
Overview	Tb927_09)_v4	450 kbp			-+ + + + + + +	+ + + + + + + + + + + + + + + + + + +							
Region	VII			- H	1	20	311							
Region	<	1070k 1080k	1090k 1100k 1110k	112 Overview	1150k 1	160k 1170k 1180k 1190k 1200k	1210k 1220k 1230k 1240k 1250k							
Details	Tb927_09)_v4: 20 kbp	5 kbp	Zoom Cancel		Tb927 09 v4:11720001217999	(X)							
	1140k 1141k	1142k 1143k	1144k 1145k 1146k	1147k 1148k :	1149k 1150k 1		7k 1158k 1159k 1160k							
	Annotated Genes	(with UTRs in g	ray when available)			Zoom in								
	Tb09.160.	5380	Tb09.16	0.5400	Tb09.160.5410	Recenter on this region	1.0290							
	Tb09.16	0.5390				Dump selection as FASTA								
		_		Coloct Tracks		Submit selection to UCSC BLAT								
				Select Tracks	j clear nignlight	Submit selection to NCBI BLAST								

- **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
 - Hint: type the ID in the "landmark or region" box.
 - What is this gene?

2.2 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. zoom out to 20Kb. What does this region look like? What genes are upstream and downstream of the SBPase?
- Modify the subtracts to remove *Leishmania* species from the view. (click on the link 'showing 26 of 26 subtracks' and uncheck all *Leishmania* species then click on change.

⊨×⊪⊭ : bra contig	1037k 1038k 1039k 1040k 1041k 1042k : Syntenic Sequences and Genes (Shaded by Ortho	1043k 1043k 1043k blo (y) (Showing 26	10468 104, 1048k 1049k 1050k 1051k 1052k 1053k 1054k 1055k 1056k 1 of 26 subtracks)
Select the	584k 585k 586k 587k 588k 589k e tracks you wish to display. Sort the tracks by clicking or by clicking and dragging rows into position.	590k 591k 592 on the column	x 593k 594k 595k 596k 597k 598k 599k 600k 601k 602k 603k
Syntenic	Sequences and Genes (Shaded by Orthology) Sul	otracks	855k 856k 8 <mark>.</mark> 7k 858k 859k 860k 861k 862k 863k 864k
Select	Organism	[©] Туре	
All off All on	All ÷	All ‡	7k 928k 929k 30k 931k 932k 933k 934k 935k 936k 937k
	Leishmania braziliensis	contig	
	Leishmania braziliensis	genes	7k 1578k 1579k 1 80k 1581k 1582k 1583k 1584k 1585k 1586k 1587k
	Leishmania infantum	contig	
	Leishmania infantum	genes	3k 1218k 1219k 219k 1219k 1219k 1220k 1220k 1220k 1220k 1220k 1221k
	Leishmania major strain Friedlin	contig	
	Leishmania major strain Friedlin	genes	758k 759k 60k 761k 762k 763k 764k 765k 766k 767k 768k
	Leishmania mexicana	contig	
	Leishmania mexicana	genes	1.02M
	Leishmania tarentolae Parrot-Tarll	contig	
	Leishmania tarentolae Parrot-Tarll	genes	k 1017k 1049k 1049k 1050k 1051k 1052k 1053k 1053k 1055k 1055k
	Trypanosoma brucei gambiense	contig	K 1047K 1040K 1040K 1030K 1031K 1032K 1033K 1033K 1030K
	Trypanosoma brucei gambiense	genes	
ø	Trypanosoma brucei Lister strain 427	contig	791k 792k 793k 794k 795k 796k 797k 798k 799k 800k 801k
	Trypanosoma brucei Lister strain 427	genes	
ø	Trypanosoma brucei TREU927	contig	50k 49k 48k 47k 46k 45k 44k 43k 42k 41k
Ø	Trypanosoma brucei TREU927	genes	
Ø	Trypanosoma congolense	contig	46k 45k 44k 43k 42k 41k 40k 39k 38k 37k 36k
	Trypanosoma congolense	genes	

- Examine the gene corresponding to the *T. vivax* SBPase. What is it called? 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 date tracks in Gbrowse.

Hint: in this case you are looking at *T. brucei*, so the data tracks you turn on will display data only if they are *T. brucei* datasets. For example, turn on the ChIP-seq coverage plots and turn off the syntenic gene and region tracks.

- 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 gene direction (blue vs. red)?

Now zoom back to 50Kb. Turn off the ChIP-seq tracks and turn on the "Cultured Procyclic and Bloodstream Tb Lister 427 RNA-seq coverage plot (log2)" track (two subtracks). Move to the region around 0.6Mbs of the chromosome (you should be on chromosome 2). Take note of the blue 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 into a gene that looks like it is differentially regulated. What are your conclusions?



- 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 mass spec evidence – which gene in this view has the highest number of peptide hits?

Outer view The set of the set o	Overview																							
off 0.11 0.21 0.31 0.41 0.51 0.61 0.31 1.11 Region 01 0.21 0.31 0.41 0.51 0.61 0.71 0.81 0.51 1.11 Details 01 0.21 0.31 0.41 0.51 0.61 0.71 0.81 0.51 11 1.11 Details 00 500k 500k 600k 700k 800k 800k 800k 1 100 kbp	Overview	Tb927_0	2_v4																					
Region CHAIL CHAIL <t< th=""><th></th><th>0M</th><th>0.11</th><th>0.2M</th><th>0.3M</th><th>0.4M</th><th>·····</th><th>-5M</th><th>0.6</th><th>н н н н н</th><th>++++</th><th>IIIII M</th><th></th><th>1111 8M</th><th></th><th></th><th>+++++ M</th><th>••••</th><th>++++ 1M</th><th>••••</th><th>1</th><th>++++ 1M</th><th>+++</th><th>↔</th></t<>		0M	0.11	0.2M	0.3M	0.4M	·····	-5M	0.6	н н н н н	++++	IIIII M		1111 8M			+++++ M	••••	++++ 1M	••••	1	++++ 1M	+++	↔
Image: Details 0.1M 0.2M 0.3M 0.4M 0.5M 0.7M 0.3M 0.9M 1M 1.1M Details Tb927_82_v4: 508 kbp 100 kbp	Pagion	VII	V.111	V.20	V.01	V.40	~	1.011	V.0		V./		~.	.011		v.5			711		1.			
0H 0.1H 0.2H 0.3H 0.4H 0.5H 0.7H 0.8H 0.9H 1H 1.1H Details Tb927_82_v4: 598 kbp 00k 500k 600k 700k 800k 100 kbp 400k 500k 600k 700k 800k 100 kbp 400k 500k 600k 100 kbp 400k 500k 100 kbp 400k 500k 100 kbp 400k 500k 100 kbp 400k 500k 100 kbp 100 kbp <	Region	(++++++									+++++			<u>.</u>								+++++		\leftrightarrow
Details Tb927_82_v4: 508 kbp 100 kbp 40k 500k 500k 700k 800k 800k 800k 800k 800k 800k 8		OM	0.1M	0.2M	0.3M	0.4M	0	.5M	0.61	1	0.7	M	0.	.8M		0.9	M		1M		1.	,1M		
tbg2/2,62_44: 500 kbp 100 kbp 400k 500k	Details					100 kbp .																		
400k 500k 600k 700k 800k 1 1 1 1 1 1 1 1		Tb927_0	2_v4: 50	0 kbp		TOO KOP F					-													
Constrained Genes (with UTRs in gray when available) Constrained Genes (Cuervo et. al.) Constrained Genes (Cuervo et. al.) Constrained Genes (Cuervo et. al.) Constrained Genes (Panigrahi, et. al.) Constrained Genes (4	ook		500k				500k				70	0k					8	jook	\neg		-	+>
Image: Spec - Peptides (Cuervo et. al.) Image: Spec - Peptides (Panigrahi, et. al.)	🛦 🖬 🕅 🕅 🖬 Anno	tated Genes	s (with UT	Rs in gray v	when availa	ible)																		
Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Panigrahi, et. al.) Image: Spect - Peptides (Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides (Spect - Panigrahi, et. al.) Image: Spect - Peptides	- ENER PART h Ma		Pontidos ((4466689		**						-					490 9 8		
t I I I I I I I I I I I I I I I I I I I			Charletter		itooolio Dha	anhonrataa	ma (Nott o																
Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et. al.) Image: Spec - Peptides (Panigrahi, et.) Image: Spec - Peptides (Panigrahi, et.) Image: Spec - Peptides (Panigrahi, et.) Image: Spec - Peptides (Panigrahi, et.) Image: Spec - Peptides (Panigrahi, et.)		ass spec - L	looustiea		NUSUIC FII		me (i	I I	L. al.)											5.1		1		
1 1	🛦 🖬 🕅 🕅 🖬 Th Ma	ass Spec - F	Peptides (F	Panigrahi, e	et. al.)															1.				
t IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII			ייין ייי	· ·	4441		"	11 F			۱.	"	ï	• •	' D 1	ľ		101	1	Υ.	''h	١ï	η	11
t ⊒ ⊠⊡ 72 12 Unified MS/MS Peptides (see <10k region for details)			'		111			11				1			- Y	- [-1	11			1			11
t ■ ⊠⊡ 73 III Unified MS/MS Peptides (see <10k region for details)								11													1			
t ■ 2 1 Unified MS/MS Peptides (see <10k region for details)																- 1								
T III III IIII IIII IIIIIIIIIIIIIIIIII								1																
t ■ ⊠ □ 7 ■ Unified MS/MS Peptides (see <10k region for details)																								
	🔹 🗖 🕅 🕅 🕅 🛛 🖛 Il Inific	d MS/MS P	antidas (s	00 <10k ror	tion for det	aile)										· '								
		Ju monino i	cpildes (s	CC STORICE		1137																		
																								⁵
																								2.5
								J. J.			Ι.	- 11												
											Щ		Ш	Ц			ЦЩ	ЦЦЦ	Ц		Ш	_11		μL

2.3 Downloading data from GBrowse and uploading your own tracks to GBrowse

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



1. 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 looking at.

2. 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 click on "new" in the section called "upload your own data".
- 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 it location.

Paste the following into the editor (the next window after you click on new), and then click on upload (hint: sometimes you have to zoom in or out a little to see your new glyph):

Brow	ser	S	elec	t Tra	acks		Snap	osho	ts	Custo	om Tra	icks	Pr	eferer	nces									
	Custom Tracks																							
[Helj	o with	up	loadii	ing c	usto	m tra	acks]																	
A	dd cu	sto	m track	ks ye acks	it. : [Fi	om	text]	[Fro	ma	URL] [I	From a f	file]										_		
Cre	ate a	a ne	w tra	ICK																				
G	or	סו				"	¢,	ind	۶r	Go	no"		-	ThO	97	02	v4	.600	າຂດ	0 6	300	aaa		
							50	ιþ¢	71	aci				103		_02		.000		00	000	555		
																							Uploa	d

You should see a new track with your gene displayed.



Now let us load a more complex graphic. For this we are going to use a URL that contains the information we need. In the field "Fetch track file from this URL", enter the following and click on upload (hint: sometimes you have to zoom in or out a little to see your new glyph): http://goo.gl/igfrO

