Analyzing the unexplained regions spreadsheet

During a previous section of this tutorial, a spreadsheet named *unexplained_regions* was generated. This spreadsheet contains locations where reads map to the genome but are not annotated by the transcript database, in this case, RefSeqGene. The *unexplained_regions* spreadsheet is potentially very interesting as it may contain novel findings.

- Right click column 6. Average Coverage and select Sort Descending from the menu
- Select Find Overlapping Genes from the Tools option in the command toolbar (Figure 1)



Figure 6. Selecting Find Overlapping Genes from Tools in the command toolbar

- Select Add a new column with the gene nearest to the region in the Find Overlapping Genes dialog (Figure 2)
- Select OK

🥵 Find Overlapping Genes	×
Find Overlapping Genes Select a method for annotating regions with genomic features.	
Add a New Column with the Gene Nearest to the Region	
○ Create a New Spreadsheet with Genes that Overlap with the Regions	
	_
OK Cancel	

Figure 7. Find Overlapping Genes

• Select RefSeqTranscripts – 2017-05-02 from the Output Overlapping Features dialog (Figure 3)

Please note that it is recommended that you annotate with the same database used when you performed mRNA quantification.

Select OK

🥵 Output Overlapping Features	×
Report regions from the specified database well-annotated set of sequences, including genomic DNA, transcripts, and proteins.	
Download required. Click OK to download the file	î
O RefSeq Transcripts 80 - 2017-02-06	
The Reference Sequence (RefSeq) collection aims to provide a comprehensive, integrated, non-redundant, well-annotated set of sequences, including genomic DNA, transcripts, and proteins.	
Download required. Click OK to download the file	
RefSeq Transcripts 81 - 2017-05-02	
The Reference Sequence (RefSeq) collection aims to provide a comprehensive, integrated, non-redundant, well-annotated set of sequences, including genomic DNA, transcripts, and proteins.	
O RefSeq Transcripts 82 - 2017-08-02	
The Reference Sequence (RefSeq) collection aims to provide a comprehensive, integrated, non-redundant, well-annotated set of sequences, including genomic DNA, transcripts, and proteins.	
Download required. Click OK to download the file	
Custom	
	~
	>
Manage available annotations	
OK Can	cel

Figure 8. Select the database to search for overlapping features

The closest overlapping feature and the distance to it is now included as columns 7. Overlapping Features and 8. Nearest Features in the unexplained_regi ons spreadsheet.

Right-clicking on a row header and selecting Browse to Location will show the reads mapped to the chromosome. For this tutorial, a couple of genes are selected to show regions that are located after a known gene or in the intron of a gene.

- Right-click row 39 and select Browse to location from the pop-up menu
 Select the *Chromosome View* tab to view a region within an intron of UNC45B. This may be a novel exon (Figure 4)

Partek Genomics Suite - 1/unexplained_regions (RNA-Seq_regions)	sults.unexplai	ned.regions *)	- 🗆 ×
<u>File</u> <u>V</u> iew			
			Workflows RNA-Seq ~
Analysis X Chromosome View X Gene Ontology Browser (1	I/Diff_Exp_and	$_{\rm I_Alt_Splice/1)} \times$	-RNA-Seq X
Tracks		dr17:33477126-33479361	└─ Import
Tracks			Import and Manage Samples 🛛 🗸
RefSeq Transcripts 81 - 2017-05-02 (hg19) (+)		RefSeg Transcripts 81 - 2017-05-02 (+)	Add Sample Attributes
Isoform proportion (1/transcripts) Rom Profile (Heart)			Choose Sample ID Column 🗸
Bam Profile (Brain)			∨ qa/qc
Bam Profile (Muscle)		Isoform Proportion 🗧 Heart 🗖 Brain 📕 Muscle 📕 Liver	Alignments per Read 🗸
Bam Profile (Liver)		NMM 7316731	\succ Analyze Known Genes
Regions (1/unexplained_regions (RNA-Seq_results.unexp			mRNA Quantification 🗸
		$\sim \sim $	Differential Expression Analysis
· ·		lle est	Alternative Splicing Analysis
	0		Create Gene List 🗸
New Track Remove Track			> Allele-Specific Analysis
Profile Color Labels		Brain	\sim Visualization
Gene	0		Chromosome View 🗸
		Muscle	Cluster Based on Significant Genes
	0		✓ Biological Interpretation
		liver	Gene Set Analysis 🗸
	0		Pathway Analysis
		1/unexplained regions (RNA-Seg results.unexplained.regions)	
	Heart		
	Muscle		
Apply Reset	•	J J C Q €	Related: Analyze a Partek Flow project
			1)

Figure 9. A region within an intron of UNC45B that might be an novel exon

• Right-click row **12576** and select **Browse to location** to go to a region that starts 1 bp after CD82.

• Select () several times to zoom out slightly

This peak may represent an extended exon (Figure 5).

Partek Genomics Suite - 1/unexplained_regions (RNA-Seq_response)	sults.unexplai	ned.regions *)		- 0	×
File View				Maddaus DNA Cas	
Analysis V. Charmone View V. Cons Ontoles - Barrar (1	(Diff Fun and	All C-11		DNA Com	Ň
Analysis X Chromosome view X Gene Ontology Browser (1,	/Dim_exp_and	Ait_spice/ I) X		- KNA-Seq	\sim
Tracks		chr 11:44640536-44642320 V		✓ Import	
Tracks				Import and Manage Samples	~
RefSeq Transcripts 81 - 2017-05-02 (hg19) (+)		RefSeq Transcripts 81 - 2017-05-02 (+)		Add Sample Attributes	~
Isoform proportion (1/transcripts)				Choose Sample ID Column	~
Bam Profile (Heart)		CD82		✓ QA/QC	
Bam Profile (Muscle)		Isoform Proportion 📕 Heart 🗖 Brain 📕 Muscle 📕 I	iver	Alignments per Read	~
Bam Profile (Liver)			1461	Analyze Known Genes	· ·
Regions (1/unexplained_regions (RNA-Seq_results.unexp				mRNA Quantification	
		NM_001024844		Differential Formation Application	
				Differential Expression Analysis	~
×		Heart		Alternative Splicing Analysis	
	0			Create Gene List	~
New Track Remove Track		Brain		> Allele-Specific Analysis	
Profile Color Labels	0			✓ Visualization	
Gene				Chromosome View	~
		Muscle		Cluster Based on Significant Genes	
	0			✓ Biological Interpretation	
		Liver		Gene Set Analysis	~
	0			Pathway Analysis	-
		1/uneveloped regions (PNA Sec. results uneveloped regions)			
		T/unexplained_regions (RNA-Seq_results.unexplained.regions)			
	Heart				
	Muscle				
Apply Reset					
Арру Кесс	<u> </u>		÷.	Related: Analyze a Partek Flow project	

Figure 10. A region that starts 1 bp after CD82 that might represent an extended exon

While RefSeq was used to identify overlapping features, the choice of which database to use will depend on the biological context of your experiment. For example, you may wish to utilize promoter or miRNA databases if you are interested in regulation of expression.

Additional Assistance

If you need additional assistance, please visit our support page to submit a help ticket or find phone numbers for regional support.



Your Rating: ☆☆☆☆☆ Results: ★★★★ 34 rates