

# Goal for this tutorial

- Identify top small noncoding RNAs that are expressed in a particular tissue
- Analyze publicly available brain dataset from ENCODE in hg38, and identify highly expressed sncRNA in this data.
- Compare and validate the analyze results in vivo using a brain dataset from DASHR in hg38
- Note: refer to “SPAR\_output\_and\_report\_page.pdf” if you want to under the outputs more

## Analyze public datasets

Click here to start

### Genome

Human (Dec. 2013 hg38, GRCh38) ▼

### Select DASHR dataset (small RNA-seq) (Human hg38):

Select DASHR dataset ▼

Analyze DASHR

### Select ENCODE data portal dataset (short total RNA-seq) (Human hg38):

frontal-cortex-rep1\_ENCSR000AFS ▼

Analyze ENCODE

## Analyze public datasets

### Genome

Human (Dec. 2013 hg38, GRCh38) ▼

**Step 1: Select hg38 as reference genome**

### Select DASHR dataset (small RNA-seq) (Human hg38):

Select DASHR dataset ▼

Analyze DASHR

### Select ENCODE data portal dataset (short total RNA-seq) (Human hg38):

frontal-cortex-rep1\_ENCSR000AFS ▼

**Step 2: Select from dropdown, a frontal cortex data from ENCODE**

Analyze ENCODE

**Step 3: Click "Analyze ENCODE" to proceed**

# Progress of SPAR



analysis, annotation, visualization of small RNA sequencing experiments  
*discover and characterize small RNAs from sequencing*

Running job **ca8ac37**

**Job id is shown**

Input files:  
[http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE\\_dataportal\\_hg38/frontal-cortex-rep1\\_ENCSR000AFS.pos.bigWig](http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_ENCSR000AFS.pos.bigWig)  
[http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE\\_dataportal\\_hg38/frontal-cortex-rep1\\_ENCSR000AFS.neg.bigWig](http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_ENCSR000AFS.neg.bigWig)

**Input files are listed**

```
Genome: hg38
Fri, 26 Jan 2018 15:37:49 -0500 ..... Loading PLUS strand: http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_E
Fri, 26 Jan 2018 15:37:49 -0500 ..... Loading MINUS strand: http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_
Fri, 26 Jan 2018 15:37:49 -0500 ..... Starting SPAR run
Fri, 26 Jan 2018 15:37:49 -0500 ..... Analysis parameters: min. fold change=2; min. coverage=10; min. size=14; max. size=44
bash run_SPAR.sh SPAR_out/ca8ac37/tracks/raw.pos.bigWig SPAR_out/ca8ac37 pos SPAR_out/ca8ac37/config.hg38.sh 2>&1
Fri, 26 Jan 2018 15:37:49 -0500 ..... Processing job ca8ac37.....
Fri, 26 Jan 2018 15:37:55 -0500 ..... DONE.

Fri, 26 Jan 2018 15:37:55 -0500 ..... Sending results to the web server.....
```

**Progress of SPAR**

**Parameters used in SPAR analyses**

# Landing results page (full view)

## SPAR

analysis, annotation, visualization of small RNA sequencing experiments  
discover and characterize small RNAs from sequencing

### Analysis results for job **ca8ac37**

Input files:  
[http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE\\_dataportal\\_hg38/frontal-cortex-rep1\\_ENCSR000AFS.pos.bigWig](http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_ENCSR000AFS.pos.bigWig)  
[http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE\\_dataportal\\_hg38/frontal-cortex-rep1\\_ENCSR000AFS.neg.bigWig](http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_ENCSR000AFS.neg.bigWig)

Link to results: [http://tesla.pcbi.upenn.edu/~plukcsa/SPAR/SPAR\\_out/ca8ac37](http://tesla.pcbi.upenn.edu/~plukcsa/SPAR/SPAR_out/ca8ac37) (will be kept for 2 weeks)  
Link to results: [Download results \(ZIP\)](#) (will be kept for 2 weeks)

#### RUN SUMMARY

##### Reads vs RNA class

RNA Class	Percentage
Unannotated	54.50%
miRNAprimary	7.19%
mir-3p	7.19%
mir-5p	17.13%
mi-5p3pno	3.22%
piRNA	
rRNA	
snoRNA	
5S RNA	
5S F3	
5S F5	

##### Peaks vs RNA class

RNA Class	Percentage
Unannotated	78.62%
miRNAprimary	
mir-3p	
mir-5p	
mi-5p3pno	
piRNA	
rRNA	
snoRNA	
5S RNA	
5S F3	
5S F5	

[Show more](#)

INTERACTIVE PEAK BROWSER +

VIEW IN GENOME BROWSER +

PLOTS +

DOWNLOAD RESULTS +

RUN LOG +

Questions?  
SPAR@lisanwanglab.org  
Wang Lab | Penn Neurodegeneration Genomics Center | University of Pennsylvania



# Landing results / outputs page (top)



analysis, annotation, visualization of small RNA sequencing experiments  
*discover and characterize small RNAs from sequencing*

## Analysis results for job **ca8ac37**

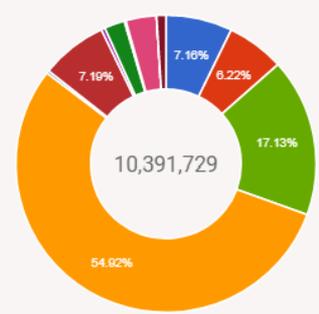
Input files:  
[http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE\\_dataportal\\_hg38/frontal-cortex-rep1\\_ENCSR000AFS.pos.bigWig](http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_ENCSR000AFS.pos.bigWig)  
[http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE\\_dataportal\\_hg38/frontal-cortex-rep1\\_ENCSR000AFS.pos.bigWig](http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/frontal-cortex-rep1_ENCSR000AFS.pos.bigWig)

Link to results: [http://tesla.pcbi.upenn.edu/~pkuksa/SPAR/SPAR\\_out/ca8ac37](http://tesla.pcbi.upenn.edu/~pkuksa/SPAR/SPAR_out/ca8ac37) (will be kept for 2 weeks)  
Link to results: [Download results \(ZIP\)](#) (will be kept for 2 weeks)

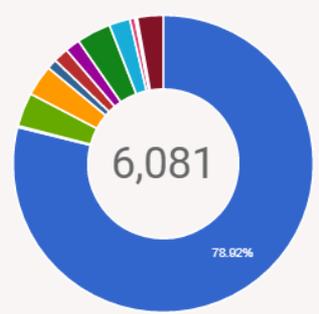
To view a report (in html / pdf): Click  
“Link to  
results: [http://tesla.pcbi.upenn.edu/~pkuksa/SPAR/SPAR\\_out/ca8ac37](http://tesla.pcbi.upenn.edu/~pkuksa/SPAR/SPAR_out/ca8ac37)”

### RUN SUMMARY

Reads vs RNA class



Peaks vs RNA class



- Unannotated
- miRNAprimary
- mir-3p
- mir-5p
- mir-5p3pno
- piRNA
- rRNA
- snoRNA
- tRF3
- tRF5
- tRNA
- scRNA
- snRNA

Show more

# Landing results → SPAR report page

## Index of /~pkuksa/SPAR/SPAR\_out/ca8ac37

<u>Name</u>	<u>Last modified</u>	<u>Size</u>	<u>Description</u>
 <a href="#">Parent Directory</a>		-	
 <a href="#">DASHR_comparison/</a>	26-Jan-2018 15:38	-	
 <a href="#">ENCODE_comparison/</a>	26-Jan-2018 15:39	-	
 <a href="#">figures/</a>	26-Jan-2018 15:39	-	
 <a href="#">inputs/</a>	26-Jan-2018 15:37	-	
 <a href="#">logs/</a>	26-Jan-2018 15:39	-	
 <a href="#">results/</a>	26-Jan-2018 15:39	-	
 <a href="#">tracks/</a>	26-Jan-2018 15:39	-	
 <a href="#">report.html</a>	26-Jan-2018 15:39	40K	
 <a href="#">report.pdf</a>	26-Jan-2018 15:39	1.8M	
 <a href="#">results.html</a>	26-Jan-2018 15:39	2.4K	
 <a href="#">run_summary.html</a>	26-Jan-2018 15:39	584	
 <a href="#">SPAR.hg38.config</a>	26-Jan-2018 15:37	2.0K	

**Report is available in html / pdf**

# SPAR report.pdf

A 15 page report containing tables and figures for download and sharing



Link to results: [http://tesla.pcbi.upenn.edu/~pkuksa/SPAR/SPAR\\_out/ca8ac37](http://tesla.pcbi.upenn.edu/~pkuksa/SPAR/SPAR_out/ca8ac37) (will be kept for 2 weeks)

Link to results: [Download results \(ZIP\)](#) (will be kept for 2 weeks)

## Run Summary

Reads	10,391,735
Expressed small RNA loci	6,081
Reads (annotated)	9,647,188
Reads (unannotated)	744,547
Genes (annotated)	1,214
Called peaks (annotated)	1,282
Called peaks (unannotated)	4,799
Expressed loci length (average)	18.90
Genome coverage (nucleotides)	114,957
RPM (average)	164.45



# Landing results / outputs page (bottom)

After clicking show more at the bottom of the "Run Summary"

Reads	10,291,735
Expressed small RNA loci	6,081
Reads (annotated)	9,647,188
Reads (unannotated)	744,547
Genes (annotated)	1,214
Called peaks (annotated)	1,282
Called peaks (unannotated)	4,799
Expressed loci length (average)	18.90
Genome coverage (nucleotides)	114,957
RPM (average)	164.45

Show less

Summary information on reads

Summary information on peaks

Step 1: Expand this to view the interactive peak browser

- INTERACTIVE PEAK BROWSER +
- VIEW IN GENOME BROWSER +
- PLOTS +
- DOWNLOAD RESULTS +
- RUN LOG +

# Interactive Peak Browser (Output)

The image shows a screenshot of a web application interface. At the top, there is a dark red header bar with the text "INTERACTIVE PEAK BROWSER" on the left and a minus sign on the right. Below this header, the main content area is white. On the left side of this area, there is a link that says "Browse peaks (expressed small RNAs)". A yellow callout box with a black border points from this link to the right. Inside the callout box, the text reads "Click this to go to 'Interactive Peak Browser'". At the bottom of the interface, there are two more dark red header bars. The first one contains the text "VIEW IN GENOME BROWSER" on the left and a plus sign on the right. The second one contains the text "PLOTS" on the left and a plus sign on the right.

INTERACTIVE PEAK BROWSER -

Browse peaks (expressed small RNAs)

Click this to go to "Interactive Peak Browser"

VIEW IN GENOME BROWSER +

PLOTS +

# Interactive Peak Browser (Output)

**Total number of called peaks  
(annotated or unannotated)**

**Click "pointing down error" to select other features to display**

Browse peaks (expressed small RNA)

Download Table

Loci: 1-10 / 6081

Page 1 of 609

Results per page 10

chr	peakID	Expression (raw)	strand	Same 5' end read %	Expression (RPM)	Expression (percentile)	Annotation	Annotated class	Conservatio	mRNA?	lncRNA?	repeat?
Clear			Clear					Clear		Clear	Clear	Clear
chr1	P116	305.5	+	0.973813	29.3985	87.7	chr1:109598939-109598961:+:hsa-miR-197-3p	mir-3p	0.2112	Intergenic	Non-lncRNA	Non-repeat
chr1	P253	121.426	-	0.986206	11.6849	76.6	chr1:10999855-10999945:-:HY1	scRNA	0.0197	Intergenic	Non-lncRNA	Non-repeat
chr1	P119	13.2556	+	0.933779	1.2756	12.33	chr1:111446797-111446916:+:HY1	scRNA	8e-04	Intergenic	Non-lncRNA	Non-repeat
chr1	P121	333.222	+	0.926571	32.0662	88.88	chr1:116671786-116671808:+:hsa-miR-320b	mir-5p3pno	5e-04	Intergenic	Non-lncRNA	Non-repeat
chr1	P7	38	+	0.947368	3.6568	52.05	chr1:1167159-1167181:+:hsa-miR-200b-3p	mir-3p	0.9188	Intergenic	lncRNA Exon	Non-repeat
chr1	P8	15.0556	+	1	1.4488	18.01	chr1:1169054-1169076:+:hsa-miR-429	mir-5p3pno	0.9834	Intergenic	lncRNA Exon	Non-repeat

# Interactive Peak Browser (Output)

Filter snRNA loci with the top 5% of expression

INTERACTIVE PEAK BROWSER

Step 1: Type "95." in this text box and click enter

Number of loci that passed this criteria: 60

Browse peaks (expressed small RNAs)  
Download Table

Loci: 1-10 / 60

chr	peakID	Expression (raw)	strand	Same 5' end read %	Expression (RPM)	Expression (percentile)	Annotation	Annotated class	Conservatio	mRNA?	lncRNA?	repeat?
Clear			Clear			95.		Clear		Clear	Clear	Clear
chr1	P131	1358.61	+	0.587891	130.7399	95.13	chr1:145157156-145157229:+:tRNA-Val-CAC-chr1-16	tRNA	0.3593	Intergenic	Non-lncRNA	Non-repeat
chr1	P150	1400.25	+	0.656677	134.747	95.21	chr1:153671250-153671282:+:piR-35176	piRNA	0.8856	Intergenic	Non-lncRNA	Non-repeat
chr1	P109	1818.27	-	0.590711	174.9733	95.94	chr1:156420353-156420375:-:hsa-miR-9-3p	mir-3p	1	3' UTR Intron	lncRNA Intron	Non-repeat
chr1	P104	1709.2	-	0.809757	164.4774	95.71	chr1:159141610-159141684:-:tRNA-Arg-TCT-4-1	tRNA	1	Intergenic	Non-lncRNA	Non-repeat
chr1	P103	1303	-	0.590763	125.3885	95.02	chr1:159141610-159141684:-:tRNA-Arg-TCT-4-1	tRNA	1	Intergenic	Non-lncRNA	Non-repeat
chr1	P218	1779.94	+	0.574177	171.2848	95.84	chr1:204506526-204506599:+:tRNA-Lys-TTT-2-1	tRNA	0.8393	Intergenic	lncRNA Exon	Non-repeat

# Interactive Peak Browser (Output)

Further filter sncRNA loci that only present in Promoter, 3' or 5' UTR exonic regions



**Step 2: Select "Promoter", "3' UTR Exon", "5' UTR Exon"**

**Number of loci that fulfilled these criteria: 4**

Browse peaks (expressed small RNAs)  
Download Table

Loc: 1-4 / 4 Page 1 of 1 Results per page 10

chr	peakID	Expression (raw)	strand	Same 5' end read %	Expression (RPM)	Expression (percentile)	Annotation	Annotated class	Conservation	mRNA?	lncRNA?	repeat?
chr16	P993	1512.66	+	0.965511	145.5642	95.44	chr16:15643300-15643322:+:hsa-miR-484	mir-5p3pno	0.9999	5' UTR Exon	Non-lncRNA	Non-repeat
chr16	P997	1355.09	+	0.466894	130.4012	95.12	chr16:19498609-19498750:+:U4	snRNA	0.0546	3' UTR Exon	Non-lncRNA	Non-repeat
chr16	P1032	173.59	-	0.358646	167.1132	95.76	chr16:636735-636806:-:tRNA-Gly-CCC-2-2	tRNA	0.9205	Promoter	Non-lncRNA	Non-repeat
chr2	P479	1745.09	-	0.358646	167.9311	95.77	chr2:70248990-70249061:-:tRNA-Gly-CCC-2-1	tRNA	0.8976	Promoter	Non-lncRNA	Non-repeat

**Step 3: Export these 4 filtered sncRNAs**

**Optional: click this link to visualize this peak P993 in UCSC genome browser**



# Finish exploring Interactive Peak Browser, next view peaks in genome browser

**INTERACTIVE PEAK BROWSER** -

**Step 1: click this to minimize this section**

Browse peaks (expressed small RNAs)

Download Table

Loci: 1-4 / 4 Page 1 of 1 Results per page 10

chr	peakID	Expression (raw)	strand	Same 5' end read %	Expression (RPM)	Expression (percentile)	Annotation	Annotated class	Conservation	mRNA?	lncRNA?	repeat?
chr16	P993	1512.66	+	0.965511	145.5642	95.44	chr16:15643300-15643322:+:hsa-miR-484	mir-5p3pno	0.9999	5' UTR Exon	Non-lncRNA	Non-repeat
chr16	P997	1355.09	+	0.466894	130.4012	95.12	chr16:19498609-19498750:+:U4	snRNA	0.0546	3' UTR Exon	Non-lncRNA	Non-repeat
chr16	P1032	1736.59	-	0.358646	167.1132	95.76	chr16:636735-636806:-:tRNA-Gly-CCC-2-2	tRNA	0.9205	Promoter	Non-lncRNA	Non-repeat
chr2	P1479	1745.09	-	0.356899	167.9311	95.77	chr2:70248990-70249061:-:tRNA-Gly-CCC-2-1	tRNA	0.8976	Promoter	Non-lncRNA	Non-repeat

**VIEW IN GENOME BROWSER** +

**Step 2: click this to maximize this section**

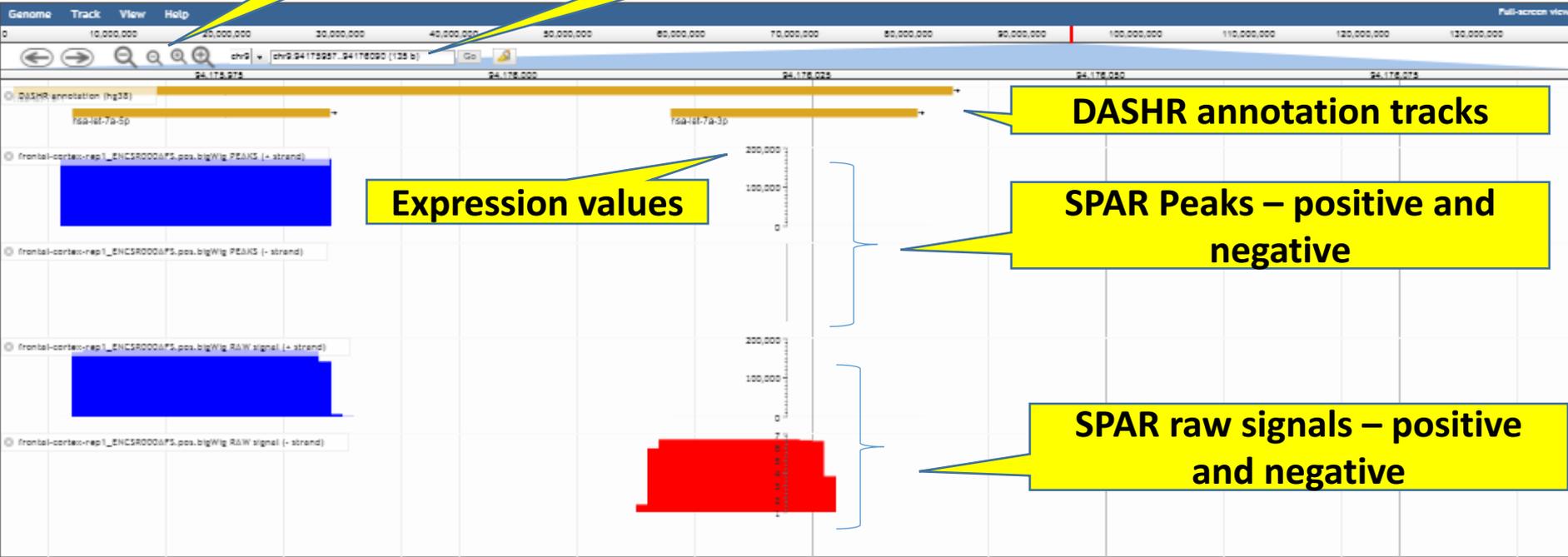
# View in Genome Browser (Output)

Optional: "View SPAR result tracks in UCSC Genome Browser (opens in the new window)" if preferred

Type in genomic coordinates / RNA / gene symbols of interest

VIEW IN GENOME BROWSER **Zoom in or out**

View SPAR result tracks in UCSC Genome Browser (opens in the new window)



DASHR annotation tracks

SPAR Peaks – positive and negative

SPAR raw signals – positive and negative

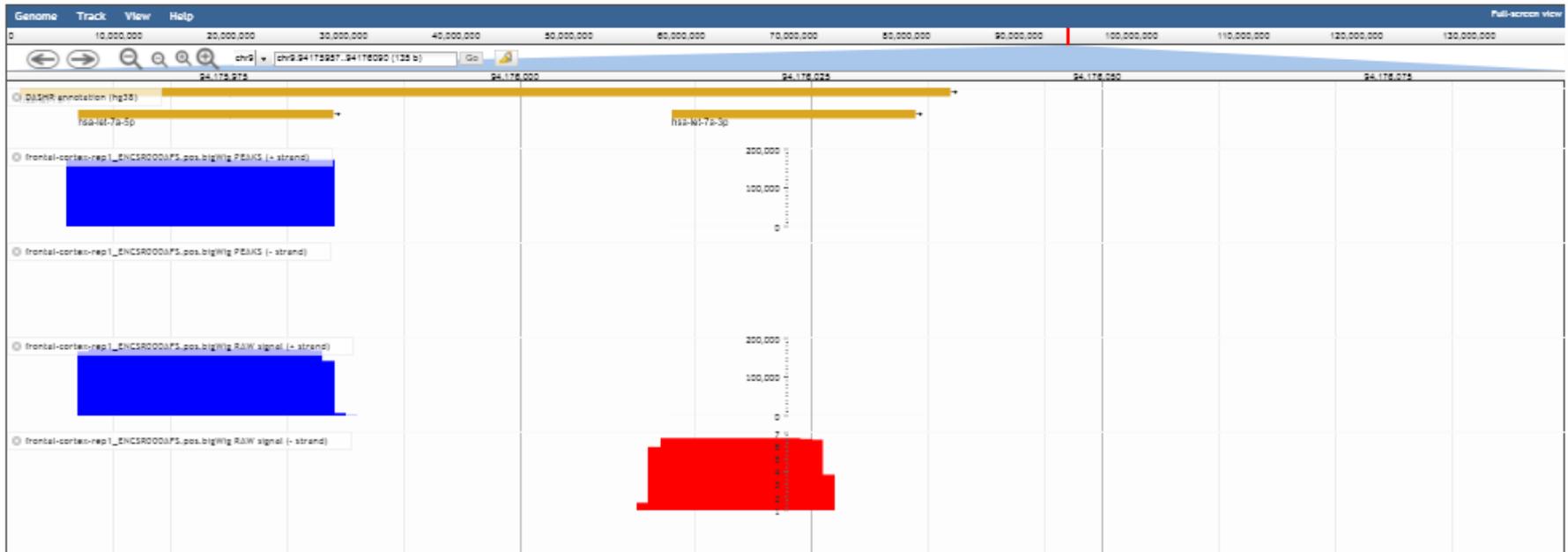
Expression values

# Finish exploring Genome browser, quickly understand the data in a genomewide view

Step 1: click this to minimize this section

VIEW IN GENOME BROWSER

View SPAR result tracks in UCSC Genome Browser (opens in the new window)



PLOTS

Step 2: click this to maximize this section



# Plots (Output)

## PLOTS

4 different plot sessions

### SEGMENTATION

Length distribution of rpm of segmented loci  
Genomic length distribution of identified small rna loci

### LENGTHEXPRESSION

Rpm distribution of identified small rna loci  
Percentile distribution of identified small rna loci across different lengths

### PROCESSING

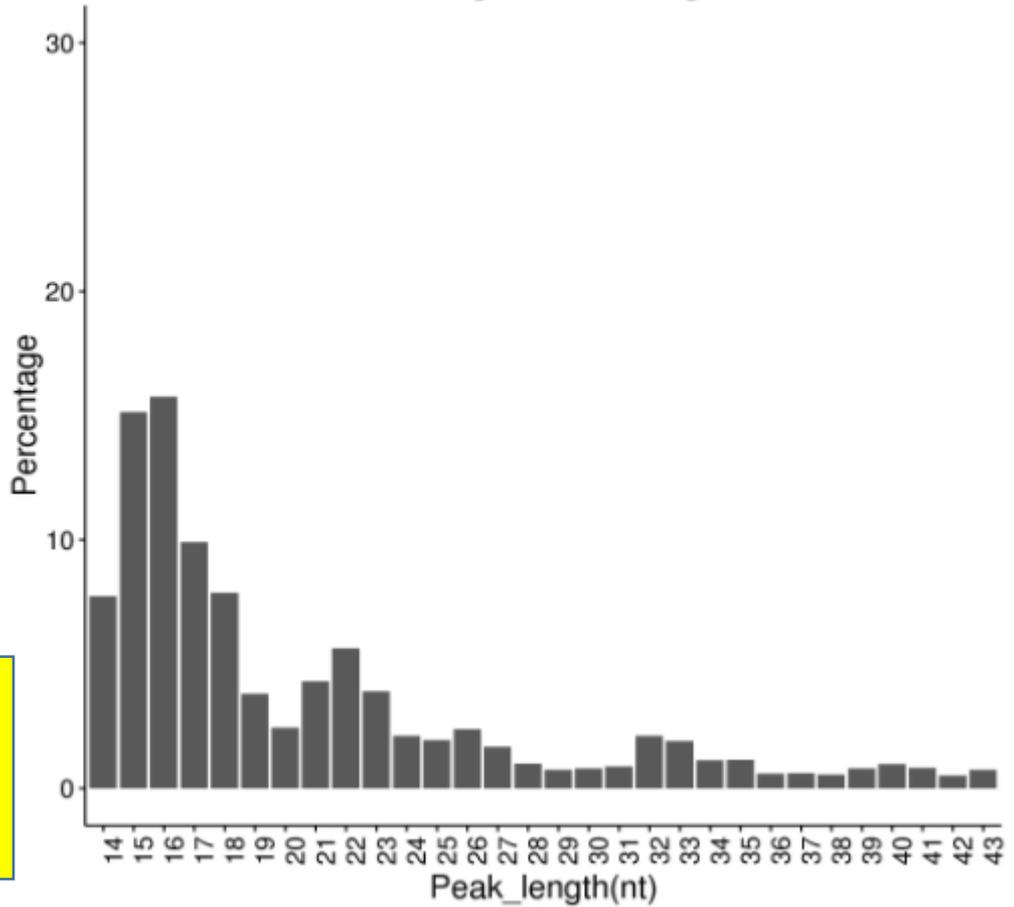
5p end positional offset between identified loci and small rna gene  
Processing specificity at 5p end of identified small rna loci  
Read pipeup at 5p end of identified small rna loci

### GENOMEWIDE

Proportion of expressed annotated small rna genes  
Genomewide distribution patterns of small rna loci  
Circular genome data visualization  
Genomewide distribution of expressed small rna loci  
Proportion of mapped reads across all loci

Click to display,  
corresponding plot will  
be shown in the right

Distribution of loci RPM  
at different lengths after segmentation

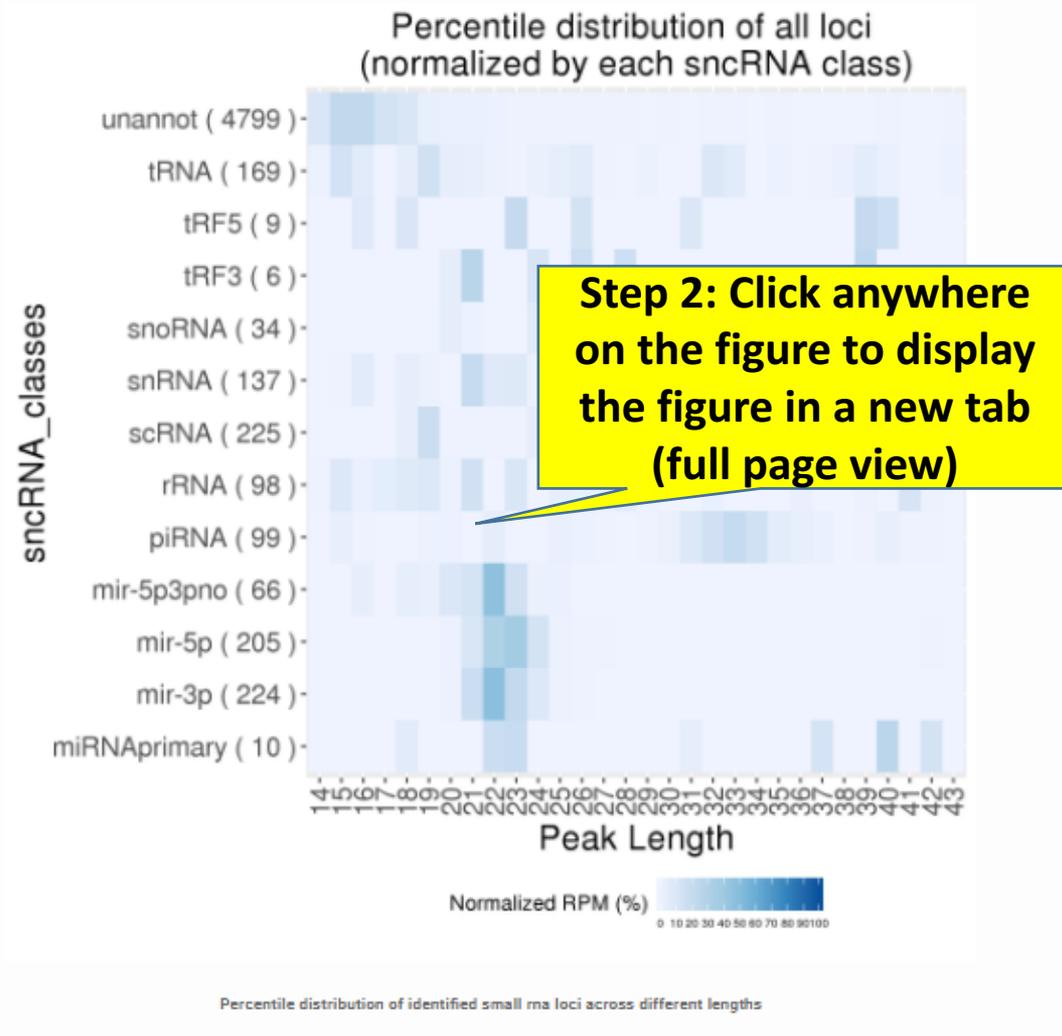


# Plots (Output)

4 different plot sessions

- SEGMENTATION**
  - Length distribution of rpm of segmented loci
  - Genomic length distribution of identified small rna loci
- LENGTHEXPRESSION**
  - Rpm distribution of identified small rna loci
  - Percentile distribution of identified small rna loci across different lengths**
- PROCESSING**
  - 5p end positional offset between identified loci and small rna gene
  - Processing specificity at 5p end of identified small rna loci
  - Read pipeup at 5p end of identified small rna loci
- GENOMEWIDE**
  - Proportion of expressed annotated small rna genes
  - Genomewide distribution patterns of small rna loci
  - Circular genome data visualization
  - Genomewide distribution of expressed small rna loci
  - Proportion of mapped reads across all loci

Step 1: Click to display, corresponding plot will be shown in the right



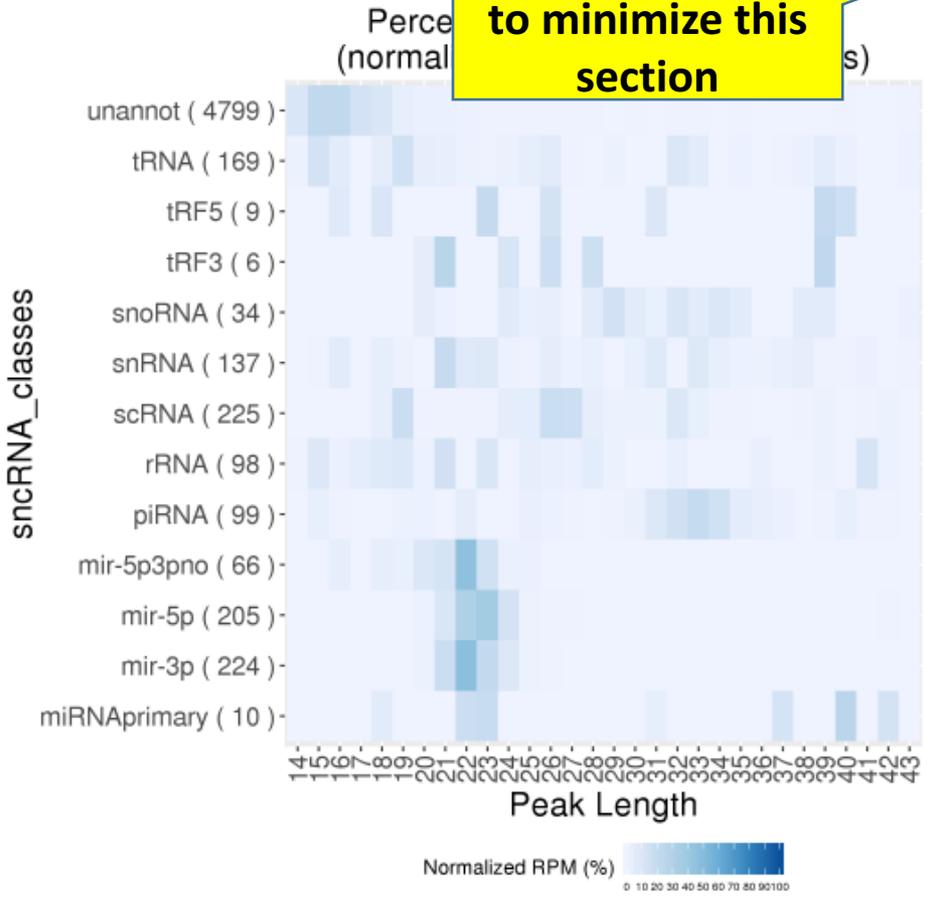
# Finish viewing the plots, want to perform own analyses with SPAR results – visit download

**PLOTS**

- SEGMENTATION**
  - Length distribution of rpm of segmented loci
  - Genomic length distribution of identified small rna loci
- LENGTHEXPRESSION**
  - Rpm distribution of identified small rna loci
  - Percentile distribution of identified small rna loci across different lengths**
- PROCESSING**
  - 5p end positional offset between identified loci and small rna genes
  - Processing specificity at 5p end of identified small rna loci
  - Read group at 5p end of identified small rna loci
- GENOMEWIDE**
  - Proportion of expressed annotated small rna genes
  - Genomewide distribution patterns of small rna loci
  - Circular genome data visualization
  - Genomewide distribution of expressed small rna loci
  - Proportion of mapped reads across all loci

**Download Results**

**Step 1: click this to minimize this section**



**Step 2: click this to maximize this section**

# Download results (Top) – bulk download

DOWNLOAD RESULTS		
Content	Description	Download
Peak tracks	Genome-wide tracks	<ul style="list-style-type: none"> <li>• plus</li> <li>• minus</li> </ul>
Raw signal tracks	Genome-wide tracks	<ul style="list-style-type: none"> <li>• plus</li> <li>• minus</li> </ul>
Peak tables	Tables with detailed information 1. Annotated peaks: peaks overlapping with annotated genes or mature products 2. Unannotated peaks: peaks that do not overlap with annotated snRNA genes or mature products.	<ul style="list-style-type: none"> <li>• Annotated peaks [xls]</li> <li>• Unannotated peaks [xls]</li> <li>• All called peaks [xls] [bed]</li> </ul>
Gene expression	Small RNA gene expression	<ul style="list-style-type: none"> <li>• Gene expression table (xls) [ 3.31 M]</li> </ul>
Comparison with DASHR	Integrated expression tables	<ul style="list-style-type: none"> <li>• Annotated peaks [xls]</li> <li>• Unannotated peaks [xls]</li> <li>• All (Annotated + Unannotated) peaks [xls] [bed]</li> <li>• Peaks not in DASHR [xls]</li> </ul>
Comparison with ENCODE	Integrated expression tables	<ul style="list-style-type: none"> <li>• Annotated peaks [xls]</li> <li>• Unannotated peaks [xls]</li> <li>• All (Annotated + Unannotated) peaks [xls] [bed]</li> <li>• Peaks not in ENCODE [xls]</li> </ul>

**Tracks for browsing**

**Comparison expression tables**

**Step 1: Click to download peaks present in both the analysed data and also expressed in DASHR**

**Step 2: Click to download peaks only in analysed data but not in DASHR**

# Download results (Bottom) – download by sncRNA class

**Step 2: Click to download, e.g. all unannotated loci from SPAR**

**Total number of reads falling onto this RNA class**

DOWNLOAD PEAK TABLES BY RNA CLASS

RNA class	Peaks	Genes	Reads	Percentage of reads
Annotated	1,282	1,214	9,647,188	92.84
Unannotated	4,799	4,799	744,547	7.16
miRNAprimary	10	10	646,698	6.22
mir-3p	224	224	1,779,736	17.13
mir-5p	205	205	5,706,729	54.92
mir-5p3pno	66	66	30,553	0.29
piRNA	99	99	747,600	7.19
rRNA	98	81	42,675	0.41
scRNA	225	215	227,677	2.19
snRNA	137	132	17,896	0.17
snoRNA	34	29	341,572	3.29
tRF3	6	6	1,323	0.01
tRF5	9	9	998	0.01
tRNA	169	138	103,725	1.00

**Step 1: Click to download, e.g. tRNA only loci results from SPAR**

# Download results (Output)

We found earlier that P1032, P993, P997, P1479 are of interest, not sure if these peaks are found in any of the DASHR brain tissues.

First excel file downloaded:

- 1) DASHR\_comparison\_all.xls – if any of the above peaks are found in this file, meaning they can be present in any DASHR tissue; not necessarily expressed in brain.

The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The data table below is as follows:

	A	B	C	D	E	F	G	H	I	J	K	L
1	#peakChr	peakChrS	peakChrE	peakID	peakExpri	peakExpri	annotID	annotRNAclass	brainog1_SRA012516	brainpfc1_GSE43335	brainpfc2_GSE48552	braintgm1_GSE46131
2	chr16	15643300	15643322	P993	1512.66	95.44	chr16:15643300:15643322:+:hsa-miR-484	mir-5p3pno	NA	1083.58	2661.43	594.515
3	chr16	19498716	19498750	P997	1355.09	95.12	chr16:19498609:19498750:+:U4	snRNA	NA	3997.57	NA	NA

P993 and P997 were found in this table. We extracting the four brain related DASHR tissues (now in column I to L). Only column J (a prefrontal cortex brain data) have both peaks expressed.

# Download results (Output)

- Second excel files downloaded:
- 2) not\_in\_DASHR.xls (for sanity check only)

The screenshot shows an Excel spreadsheet titled 'not\_in\_DASHR - Excel'. The ribbon is set to 'Data'. The active cell is A1, containing the text '#peakChr'. The table below has the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	#peakChr	peakChrS	peakChrE	peakID	peakExpress	peakExprPe	annotID	annotRNAclass	brainog1	brainpfc1	brainpfc2	braintgm	overlap	peakOver	DASHRpeakOverlapPct	
2	chr16	636771	636809	P1032	1736.59	95.76	chr16:636735:636806:-:tRNA-Gly-CCC-2-2	tRNA	.	.	.	.	0	0	0	
3	chr2	70249026	70249065	P1479	1745.09	95.77	chr2:70248990:70249061:-:tRNA-Gly-CCC-2	tRNA	.	.	.	.	0	0	0	

**P1032 and P1479 were found in this table, i.e. not expressed in any of the DASHR tissues, but present in the ENCODE prefrontal cortex data.**

**Conclusion – out of the 4 sncRNAs we identified, two were in silico confirmed to be highly expressed across different brain datasets in both DASHR and ENCODE.**

# Run log (Output)

**Step 1: click this to minimize "Download results"**

**DOWNLOAD RESULTS** -

Content	Description	Download
Peak tracks	Genome-wide tracks	<ul style="list-style-type: none"> <li>plus strand [bigBED] [ 382.36 KB ]</li> <li>minus strand [bigBED] [ 384.94 KB ]</li> </ul>
Raw signal tracks	Genome-wide tracks	<ul style="list-style-type: none"> <li>plus strand [bigWig] [ 3.66 MB ]</li> <li>minus strand [bigWig] [ 3.66 MB ]</li> </ul>
Peak tables	Tables with detailed information: 1. Annotated peaks: peaks overlapping snRNA genes or mature products with annotations. 2. Unannotated peaks: peaks that do not overlap with annotated snRNA genes or mature products.	<ul style="list-style-type: none"> <li>Annotated peaks [xls]</li> <li>Unannotated peaks [xls]</li> <li>All called peaks [xls] [bed]</li> </ul>
Gene expression	Small RNA gene expression	Gene expression table (xls) [ 3.31 MB ]
Comparison with DASHR	Integrated expression tables	<ul style="list-style-type: none"> <li>Annotated peaks [xls]</li> <li>Unannotated peaks [xls]</li> <li>All (Annotated + Unannotated) peaks [xls] [bed]</li> <li>Peaks not in DASHR [xls]</li> </ul>
Comparison with ENCODE	Integrated expression tables	<ul style="list-style-type: none"> <li>Annotated peaks [xls]</li> <li>Unannotated peaks [xls]</li> <li>All (Annotated + Unannotated) peaks [xls] [bed]</li> <li>Peaks not in ENCODE [xls]</li> </ul>

**DOWNLOAD PEAK TABLES BY RNA CLASS**

RNA class	Peaks	Genes	Reads	Percentage of reads
Annotated	1,282	1,214	9,647,188	92.84
Unannotated	4,799	4,799	744,547	7.16
miRNAprimary	18	18	646,698	6.22
miR-3p	224	224	1,779,736	17.13
miR-5p	265	265	5,786,729	54.92
miR-3p&pno	66	66	38,553	0.29
piRNA	99	99	747,688	7.19
rRNA	98	81	42,675	0.41
scRNA	225	215	227,677	2.19
snRNA	137	132	17,896	0.17
snRNA	34	29	341,572	3.29
trf3	6	6	1,323	0.01
trf3	9	9	998	0.01
tRNA	169	138	183,725	1.88

**RUN LOG** +

**Step 2: click this to maximize "Run log"**

Questions?  
SPAR@tsanwanglab.org

Wang Lab | Stem Neurodevelopmental Genomics Center | L1000b1

# Run log (Output)

Run log displayed again

## RUN LOG

```
Genome: hg38
Fri, 26 Jan 2018 15:37:49 -0500 ..... Loading PLUS strand: http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/fron
Fri, 26 Jan 2018 15:37:49 -0500 ..... Loading MINUS strand: http://dashr2.lisanwanglab.org/DASHRv2/tracks/hg38/ENCODE_dataportal_hg38/fro
Fri, 26 Jan 2018 15:37:49 -0500 ..... Starting SPAR run
Fri, 26 Jan 2018 15:37:49 -0500 ..... Analysis parameters: min. fold change=2; min. coverage=10; min. size=14; max. size=44
bash run_SPAR.sh SPAR_out/ca8ac37/tracks/raw.pos.bigWig SPAR_out/ca8ac37 pos SPAR_out/ca8ac37/config.hg38.sh 2>&1
Fri, 26 Jan 2018 15:37:49 -0500 ..... Processing job ca8ac37.....
Fri, 26 Jan 2018 15:37:55 -0500 ..... DONE.

Fri, 26 Jan 2018 15:37:55 -0500 ..... Sending results to the web server.....
Fri, 26 Jan 2018 15:39:38 -0500 ..... Processed 10,391,735 reads in 109 seconds (95,337 reads / second)

Fri, 26 Jan 2018 15:39:38 -0500 ..... Creating plots
Fri, 26 Jan 2018 15:39:46 -0500 ..... DONE creating plots
```

SPAR run log

Send questions to this email:  
SPAR@lisanwanglab.org

END