Making .html and .pdf from .Rmd (pandoc setup)

Sept. 6, 2023

Testing this out again .

MEME requires:

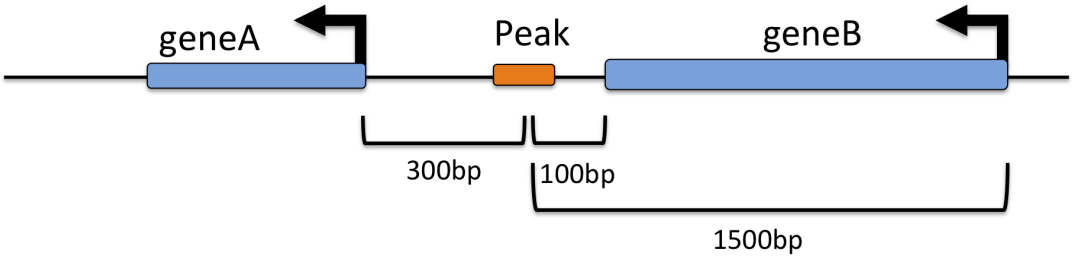

Figure 1: peak assignment problem

section sssddss

```
this is code
```

- '.html' file output

Getting started

A useful skill to have under your belt is to identify RNA-Seq datasets in published papers that you can re-analyze on your own to ask new questions. RNA-seq datasets are abundant, and publicaly accessible. When a group generates RNA-Seq data and publishes a paper with this data, there are always many unexplored analysis that can be done. For example, in the Miura lab, we are interested in Alternative Polyadenylation and Alternative Splicing. Existing RNA-Seq datasets can be re-analyzed using tools such as `Qapa` and `rMATS`. We will perform these analysis as part of this tutorial on a published RNA-Seq dataset.

For more on the usefuless of re-analyzing RNA-Seq data as a tool for learning, see this article by our department Brent Graveley.

Finding Existing RNA-Seq datasets

We will be using a short read RNA-Seq dataset from Kiltschewskij et al., NAR, 2023.

The dataset is a neural differentiation of SHSY5Y cells performed on a NextSeq500 Illumina sequencer. The dataset, deposited in GEO, can be found here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE155432.

The dataset

SRR	SRX	Condition
SRR12352385	SRX8851832	Undiff_1
SRR12352386	SRX8851833	Undiff_2
SRR12352387	SRX8851834	Undiff_3
SRR12352388	SRX8851835	Diff_1
SRR12352389	SRX8851836	Diff_2
SRR12352390	SRX8851837	Diff_3

Obtaining the .sra files and converting to .fastq

These are very large files, and thus downloading from a web browser is not an option. You'll have to use a tool called `prefetch`

This will be performed on the Xanadu HPC and executed with a slurm script `prefetch.sl`

- For a primer on using Xanadu see here

Create a directory called something like `/labs/miura/your_name/SHSY5Y`

Enter into that directory `cd /labs/miura/your_name/SHSY5Y`

create a new file called `prefetch.sl` and edit it `nano prefetch.sl`

Here is the code for `prefetch.sl` that you can then paste in

```
#SBATCH --job-name=prefetch
#SBATCH -N 1
#SBATCH -n 1
#SBATCH -c 1
#SBATCH --partition=general
#SBATCH --qos=general
#SBATCH --mail-type=END
#SBATCH --mem=16G
#SBATCH --mail-user=miura@uchc.edu
#SBATCH --output=/projects/Karlsruhe/SHSY/eofiles/%x.%j.out         #standard output
#SBATCH --error=/projects/Karlsruhe/SHSY/eofiles/%x.%j.err          #standard error log

module load sratoolkit
prefetch SRX8851832 SRX8851833 SRX8851834 SRX8851835 SRX8851836 SRX8851837
```

You will change `miura@uchc.edu` to your email. Change the `--output` and `-error` flags to whatever you like. I suggest `/labs/miura/your_name/SHSY5Y/eofiles`