QE Prep presentation

Edmund Miller

2023-02-27 Mon

Overview

- nf-core bytesize talk on nf-test
- nf-core maintenance

Aims

- Aim 1: Create a best practice secondary analysis pipeline for nascent transcripts
- Aim 2: Take advantage of New Developments to improve eRNA annotation
- Aim 3: Dive into the "junk" DNA

Aim 1: Create a best practice secondary analysis pipeline for nascent transcripts

- Creation of the pipeline is most of the preliminary results
- Quantifying is the main focus of this aim

Aim 1 Projects

- nascent pipeline
- Improvements to nascent transcript identification

Aim 2: Quantifying of New Developments to improve eRNA annotation

- How much better are PINTS, groHMM, and homer specifically for eRNAs
  - eRNAs identified
- How much improvement are we getting with better aligners?
  - RNA-seq counts
- How much improvement are we getting with improved Reference genomes
  - Any more reads that were picked up due to improved bases?
  - eRNAs identified in previously unmapped regions(aim 3)

Aim 2: Success criteria

- Success criteria, we expect as we progress through time these will have improved
    - Is this a matrix like

	hg19	hg38	CHM13
bowtie2			
bwa			
bwa2			
hisat			
STAR			

Aim 3: Dive into the "junk" DNA

- CHM13 analysis
  - Are there any active enhancers in these regions?
- Viralintegration and CHM13
  - How does this differ from Alyssa's work? While she's focused on the viruses present and biological inferences/consequences, I'm quantifying how many reads we're throwing away that could be recycled.

Aim 3 Projects

- eRNA splicing
- Simple eRNA kinetics in new regions