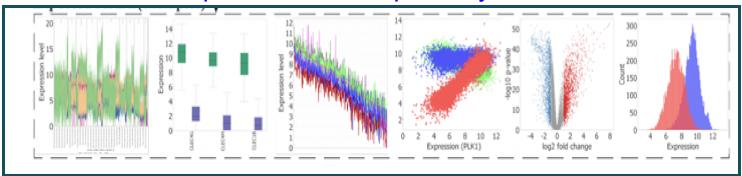
A hands-on workshop in COVID-19 RNA-Seq data analysis in an HPC environment



The goal of this workshop is to train people in *reproducible, high throughput* bioinformatics analysis of RNA-Seq datasets, as crucial to biological discovery.

March 29, 10 am-1 pm CST (4-7 pm UTC)

Session 1 Refresher in python.

Optional pre-workshop session. Refresher in python. Leader, Jeff Haltom.

March 29, 1:30 pm-3 pm CST (7:30-9 pm UTC)

Session 2. pyrpipe. Introduce best-practices for RNA-Seq analysis pipelines and detail popular

RNA-Seq processing tools like SRA Tools, Trimgalore, STAR, and stringtie. Describe anaconda environment manager, a popular package manager for maximizing reproducibility of RNA-Seq results. Hands-on analysis with **pyrpipe** (Singh et al., 2021), an efficient, flexible python package for writing RNA-Seq pipelines. We will describe pyrpipe, and detail how to customize it to process RNA-Seq data from COVID-19 studies. Leader, Urminder Singh.

March 29, 3 pm-4 pm CST (9-10 pm UTC) Session 3.Integrating pyrpipe with the Snakemake workflow manager We will focus on integrating pyrpipe with Snakemake workflow manager to write scalable pipelines. Leader, Urminder Singh.

Python Code ./params #1 create an SRA object sobj=SRA("SRR976159") #2 create trim_galore object sobi tgalore=Trimgalore() #3 create a STAR object star=Star(index="star_ir star.yaml #4 create StringTie object #\$ tringtie="Stringtie(guide="ref.gtf") #5 RNA-seq processing sobj.trim(tgalore).align(star).assemble(stri stringtie stringtie.yar #6 Import orfipy command in python orfipy=Runnable(command='orfipy') orfipy orfipy.yam #7 pass stringtie.gtf to orfipy orfipy.run(srr.assembly) python script.py --threads 10 Shell \$ fasterq-dump -O SRR976159 -o SRR976159.fastq -e 10 -f SRR976159 \$ trim_galore --cores 10 --paired -o SRR976159 SRR976159/SRR976159_1.fastq \ Execute SRR976159/SRR976159_2.fastq \$ STAR --runThreadN 10 --genomeDir star_index --outSAMtype BAM SortedByCoordinate --readFilesIn SRR976159/SRR976159_1_tg.fastq SRR976159/SRR976159_2_tg.fastq \ --outFileNamePrefix SRR976159/ stringtie -p 10 -G ref.gtf -o SRR976159/Aligned.sortedByCoord.out.gtf SRR976159/Aligned.sortedByCoord.out.bam 6 orfipy SRR976159/Aligned.sortedByCoord.out.gtf Analysis reports and summary pyrpipe Extensive debugging information. Reports Benchmark stats and charts. pyrpipe_diagnostic MultiQC reports supported.

March 30, 10-12 CST (4-6 pm UTC)

Session 4. MetaOmGraph (MOG). Downstream user-friendly exploratory analysis and visualization of RNA-Seq data. MOG (Singh et al, 2020) is a Java platform tool to enable analysis and visualization of large datasets. MetaOmGraph seamlessly integrates with R and with pyrpipe.

Leader, Priyanka Bhandari

No coding required for Session 4

For each session, users will run programming exercises on real RNA-Seq data from COVID-19 studies.

Workshop is open to 30 participants.

. . . .

COVID-19 International Research Team



Register at `https://forms.gle/MgP3DjtMj5ddBdBx9`

In conjunction with the 3rd COV-IRT Symposium https://www.cov-irt.org/symposium-3/