TCSS 592: Group Project Spring 2017

Project by:

Ashwini Shinde
Iswarya Sridharan
Saranya
Yu Zhuang

Description:

What is Orange?

- **Orange** is an open‒source data visualization, machine learning and data mining tool

- Add‒ons for bioinformatics and text mining

What is widget?

- Orange components are called widgets

- Orange widget provide GUI to orange data mining

- It includes widgets for data entry and processing

Bowtie2

- Bowtie 2 is an ultrafast and memory-efficient tool for aligning sequencing reads to long reference sequences.

- Instruction: **bowtie2** [**options**]* −x ⟨**bt2−idx**⟩ {−**1** ⟨**m1**⟩ −**2** ⟨**m2**⟩ | −**U** ⟨**r**⟩} −**S** [⟨**hit**⟩]

  ❖ *i.e. bowtie2 -x chr1 -U SRR494099.fq -S result_test.sam*

- Input:

  ❖ The basename of the bowtie2 index for the reference genome. *i.e. chr1*

  ❖ Comma-separated list of files containing unpaired reads to be aligned *i.e. SRR494099.fq*

- Output: File to write SAM alignments to. *i.e. result_test.sam*

Implementation:

Data preparation: Builds a Bowtie index from a set of DNA sequences.

- Instruction: bowtie2-build [options]* <reference_in> <bt2_base>
- ❖ *i.e. bowtie2-build chr.fa chr1*

- Input: A comma-separated list of FASTA files containing the reference sequences to be aligned to. *i.e. chr.fa*

- Output: A set of 6 files with suffixes .1.bt2, .2.bt2, .3.bt2, .4.bt2, .rev.1.bt2, and .rev.2.bt2. *i.e. chr1.1.bt2, chr1.2.bt2, ..., chr1.rev.1.bt2, chr1.rev.2.bt2*

Implementation Details:

- Automatically filter several Bowtie Index files as a whole:
- ❖ *i.e. chr1.1.bt2, chr1.2.bt2, …, chr1.rev.1.bt2, chr1.rev.2.bt2* ➔ *chr1*

- Command of Bowtie2 using on Docker needs "bash -c" prefix.
- ❖ *i.e. Commands = "bash -c \"" + ' '.join((str(w) for w in parameters)) + "\""*

- The names of volumes can't be the same, because the DS of volumes is a dictionary.
- ❖ *i.e. volumes= {self.host_fastq_dir: self.container_fastq_dir,*
- ❖ *self.host_index_dir: self.container_index_dir}*

Demo

Conclusion:

Learning Part:

How to add a widget in Orange using biocontainers on Docker in details.

- Pull the useful image from the Docker

- Pass the host file path to Docker

- GUI designe using PyQt5

Challenging Part:

Setting up the environment was the challenging part. We got it on MAC but was pretty tedious on windows

Work Distribution

Yu Zhuang - setting up the environment, Implementation of bowtie2 widget and run in orange

Ashwini - setting up the environment, Implementation of bowtie2 widget and run in orange

Iswarya - setting up the environment, Implementation of bowtie2 widget and run in orange

Credit:

Thanks to Jiaming Hu' s help and valuable suggestions

References:

https://sourceforge.net/projects/bowtie-bio/files/bowtie/

http://computing.bio.cam.ac.uk/local/doc/bowtie2.html

https://github.com/BioDepot/BioDepot-workflow-builder

Thank you.