Peer Graded Assignment: Bioinformatics Application Challenge

You submitted!
Your work is ready to be reviewed by classmates. Next, you need to review your classmates' work. We'll email you when your grade is ready. Your grade should be ready by **July 27, 11:59 PM PDT**.

Review Classmates' Work

Instructions

My submission

Discussions

Instructions

Every year in the United States, half a million patients contract a ***Staphylococcus* (Staph)** infection after surgery. Many of these patients are infected with drug-resistant strains such as **methicillin-resistant *Staphylococcus aureus* (MRSA)**, which can resist even last-resort antibiotics like Vancomycin and Daptomycin. As a result, MRSA causes over 20,000 deaths a year in the U.S. alone. Since there are over 40 different types of Staph bacteria that could be causing these infections, you want to determine which species is causing a Staph infection in a given patient by isolating this species in the patient and sequencing its genome. After you have sequenced its genome, scientists can start analyzing mutations that have led to antibiotics resistance.

ASSEMBLY

Let's assume that we have isolated bacteria in the patient and generated reads for these bacteria. To assemble the genome from the reads, you will be using the **SPAdes** assembler (Bankevich *et al*, 2012) through Illumina's **BaseSpace** service. Please follow these step-by-step instructions to register on BaseSpace and run SPAdes:

Register: Create an account on BaseSpace here. You will need to fill out all fields.After logging into BaseSpace, you will see the following dashboard and menu at the top of the page. Click on "Projects" and then create a new project. Name it whatever you like. All of the following analysis will be performed under this project.
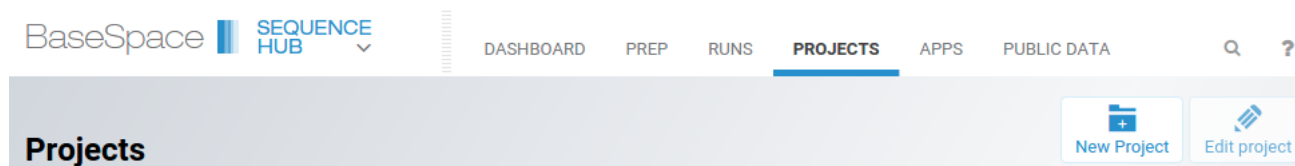
Next, we need to import our data. For this assignment, we will import raw data directly from the SRA database. Click on "Apps" in the top menu and search for the SRA import app.
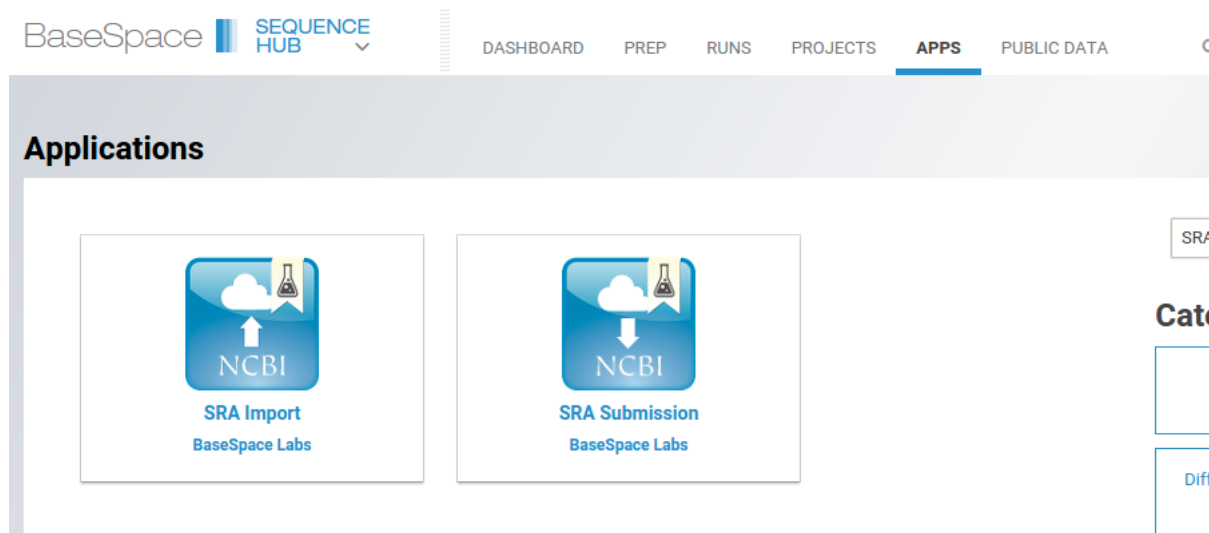
Launch the app and use accession number **SRR643156** to import the data. Click "Save to," select your project, agree to the disclaimer, and click the continue button to import the data. It may take about 30 minutes for the app to begin execution and to load the data. When the app is finished, you will be able to see the imported files and related information in your project folder.

SRA Import v0.0.3
BaseSpace Labs

Analysis Name:

SRA Import 05/19/2016 4:49:36

SRA Accession:

SRR643156　ⓘ

This app is free.

→
Continue

Save Results To:

Select Project(s):

Analysis_data　　　　　　　　×

BaseSpace Labs Disclaimer

BaseSpace Labs:

☑ I acknowledge and agree that (i) this is a BaseSpace Labs App, (ii) I am using it AS-IS without any warranty of any kind, (iii) Illumina has no obligation to provide any technical support for this App, and (iv) Illumina has no liability for my use of this App, including without limitation, any loss of data, incorrect results, or any costs, liabilities, or damages that result from use of this App.

Next, we will use SPAdes to assemble the genome. Go to the App page again and search for SPAdes. Launch SPAdes (**please use version 3.0**) and click "Select Sample(s)." Click on the top sample, labeled **SRR643156**, and click "Confirm." Save results to the project that you made and specify the parameters below (use default settings except for the value of k-mer size). For this homework assignment, you will need to run the program three times (for $k = 25$, $k = 55$, and $k = 85$) to investigate how the choice of parameter k affects the assembly quality. Each run takes about 30 min, but you can queue all three runs at once – you will need to repeat all of the steps in this paragraph for each run. You will see the results appear in your project folder when the app has finished running. While the app runs, please continue reading.

≡
MENU　▕▏　SEQUENCE HUB

Applications

spades

SPAdes Genome Assembler

Algorithmic Biology Lab
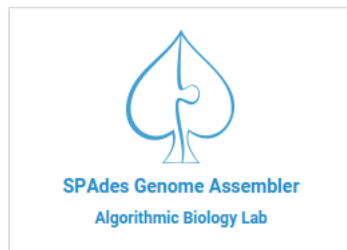
Review criteria　　　　　　　　　　　　　　　　　　　　　　　　**less ⌃**

Everyone enrolled in the course must review at least five other submissions to ensure everyone receives a grade; however, many learners complete more to help their peers who are still waiting.

Copying answers from the rubric is a violation of the Coursera honor code. Please assign 0 points to any question whose text significantly matches the rubric text. If significant matches with the rubric are found for multiple questions, then please assign the entire assignment a 0 and notify the course staff.