East River Metatranscriptomics

Run this R script to analyze the environmental RNA (eRNA) samples from a set of water samples in order to explore the microbes detected as well as what we can learn about eRNA stability and the impact of filtration in these water samples.

Introduction

In this explore module, we'll compare the microbial composition of water samples from the East River, collected just by the Rockefeller University (RU) campus in New York City. These water samples were collected on July 5, 2021 in anticipation of receiving student samples for Rockefeller's Summer Science Research Program (SSRPv; v for virtual, since this occurred during the COVID-19 pandemic).

Since students from across the country would be sending samples to RU, we wanted to control for the amount of time water samples would spend in transit. We reasoned that because RNA is so short lived (**add link to RNA metatranscriptomics "Know" section**), it was likely that the RNA composition could change once a sample was taken from its environment and shipped to us. To mimic this transit time, we took two samples from the East River and kept one (called "roomtemp") for 2 days at room temperature to mimic mailing, before freezing it. The other sample, called "frozen", was placed in a freezer immediately after the sample was taken from the river.

Next, both samples underwent the same isolation procedure at the same time to get RNA. Each sample was passed through a filter to trap bacteria and single-cell eukaryotes, and from the flowthrough we isolated RNA which was expected to contain viruses (Figure?).

RNA from both samples was used to make sequencing libraries and then processed using [CZID](<http://CZID.org>) to get taxon level microbial composition. So in total, we have 4 samples we'll be comparing:

Sample	**Condition**
Frozen	filter
Frozen	flowthrough
Roomtemp	filter
Roomtemp	flowthrough

Let's now explore some data. For this explore module, we'll start by firing up R in RStudio, and retrieving the data. This module assumes some basic familiarity with R.

This project and analysis is provided by Dr. Joseph Luna as a RockEDU Fellow while working as a Postdoctoral Fellow in the Rice Lab at The Rockefeller University. Follow along with Joe as he takes you through these data and their analysis...

Materials

RStudio

- Open RStudio. (Learn more here to start downloading the components for the free RStudio package if you don't have it yet)

R Script

- Download the provided R script "EastRiver.Rmd" from the Save & Share menu on this page and load it into RStudio. Run each section and read the comments to walk through the analysis with Joe.

Protocol

1. Run each step of the script, reading the comments and viewing the displayed outputs as you go.
2. Once you get to line 160, now it's your turn to step in and do your own analysis of the data.
3. After going through these analyses, feel free to modify the analysis to ask related questions, graph a particular subset of these data, etc. Where does your curiosity take you?

Exploration

As you look at the table of species most lost from the Roomtemp samples (as generated from running the script; a snapshot of which is shown below), consider some of the following questions:

What are the viruses that we seem to be losing at roomtemp and what do you think they normally infect?

Which eukaryotes exhibit the greatest changes comparing the frozen to roomtemp samples?

Show 10 entries Search:

	tax_id	genus_tax_id	name	category	is_phage	sample	condition	
1	2559284	-200010744	Roseobacter phage CRP-5	viruses	true	roomtemp	flowthrough	7
2	186617	-200	uncultured marine virus	viruses	false	roomtemp	flowthrough	7
3	1283076	-200010662	Pelagibacter phage HTVC008M	viruses	true	roomtemp	flowthrough	1
4	1922559	-200	Beihai picorna-like virus 17	viruses	false	roomtemp	flowthrough	8
5	2559285	-200010744	Roseobacter phage CRP-6	viruses	true	roomtemp	flowthrough	9
6	1407671	-200010699	uncultured Mediterranean phage uvMED	viruses	true	roomtemp	flowthrough	5

Showing 1 to 10 of 235 entries Previous 1 2 3 4 5 … 24 Next