

Indo-US Workshop on Genomics and Bioinformatics to Explore Human Microbial Ecology in Health and Diseases

September 6-8, 2017

Translational Health Science and Technology Institute (THSTI), DBT, India

Workshop Organizers

- Translational Health Science and Technology Institute (THSTI), DBT, India
- NIAID Division of Microbiology and Infectious Diseases (DMID), NIH
- NIAID PATRIC Bioinformatics Resource Center/University of Chicago/ANL
- NIAID Genomic Center for Infectious Diseases at J. Craig Venter Institute(JCVI)

Background:

The idea of a human microbiome workshop emerged as a follow up of the collaboration between NIAID and THSTI (more specifically the DBT-Inter-Institutional Program on Preterm Birth). The same was considered for support during the Joint Working Group Meeting of Indo-US Vaccine Action Programme, held on April 11-12 at NIAID USA. The central aim of this workshop would be to build analytical capacity amongst young scientists in THSTI and other research institutes across India.

Importance of the workshop:

The human body is populated with trillions of microbial species from all the three domains of life (*Archaea*, *Bacteria* and *Eukarya*). Bacteria, in particular, play crucial role in several aspects of our physiology like dietary nutritional uptake, immune homeostasis, tissue and organ function, response to therapeutic interventions as well as in seemingly unrelated aspects like neurological behavior and antimicrobial resistance. Concurrently, dysbioses in the composition of these microbial communities residing in the different body sites have been implicated in health and the onset of several diseases and metabolic disorders.

Obtaining a holistic picture of body site-specific microbiomes from healthy individuals and identifying 'key' agents/factors of dysbiosis is therefore crucial for diagnosing susceptibilities to various diseases including infectious diseases and clinical complications and microbiome-based therapies.

The advent of the Next Generation Sequencing (NGS) technologies has revolutionized the way microbiomes are being investigated. Rather relying on a culture based approach, researchers can sequence the DNA content of entire microbial communities in a cost-effective

manner. However, the primary challenge pertains to the analysis of this microbial community sequence data. Metagenomic sequence data is not only huge in terms of volume, but also complex given the presence of genome sequences from various microbes residing in the sample isolation site. Making meaningful inferences on the taxonomic make-up, functional attributes and the interaction dynamics of microbial communities from high throughput sequencing data requires several downstream analysis steps involving specialized algorithms and multivariate statistical analysis techniques.

Objectives:

The present workshop aims to advance microbiome research in India by providing training on the analysis of high throughput sequencing data and obtaining importance inferences on the

- (a) Diversity and richness of microbial ecology.
- (b) Dynamics of microbial population in specific niche.
- (c) Interactions among microbial members and infectious diseases.
- (d) Functional attributes in the microbial genomes, especially **antimicrobial resistance** traits.

The emphasis will be on “hands on sessions” for the participating scientists who will be encouraged to work on their own data sets during these sessions.

Expected Outcomes:

- (1) Hands on training in analyzing high throughput sequencing data.
- (2) Better understanding of microbial ecology and its impact in health and diseases.
- (3) Understand different aspects of human microbiome research and its implication in public well-being.
- (4) Develop national and international collaborations for future research.

AGENDA

Wednesday, September 6, 2017

- 9:00 am** **Welcome Remarks**
Gagandeep Kang, THSTI
- 9:15am** **Objectives and expected outcome of the workshop**
Shinjini Bhatnagar, THSTI
- 9:30 am** **Overview of NIH-NIAID Genomics and Bioinformatics Programs**
Maria Y. Giovanni, US NIH-NIAID
- 10.00am** **Overview of NIAID PATRIC Bioinformatics Resources**
Rick Stevens, PATRIC
- 10:30 am** **Overview of NIAID Genomic Center for Infectious Diseases at JCVI**
Karen Nelson, Lauren Brinkac, Indresh Singh, JCVI
- 11:00am** **Human Microbiome Research in THSTI: Understanding and challenges**
Bhabatosh Das, THSTI, Faridabad
- 11.30am** **Insights into human gut microbiome**
Yogesh S. Shouche, NCCS, Pune
- 12.00 am** **Insights into human skin microbiome**
Souvik Mukherjee, NIBMG, Kalyani
(Open session ends)

12:30 pm Lunch

Hands-On Sessions

- 1:30 pm** **Getting started in PATRIC**
Tom Brettin, PATRIC
- Signing-up for PATRIC user account
 - Navigating the site
 - Using the PATRIC workspace
 - Uploading / editing data
- 2:00 pm** **Assembling Genomes in PATRIC**
Maulik Shukla, PATRIC
- Uploading reads and read libraries
 - Selecting assembly methods for comparison
 - Queuing jobs and retrieving results

- Analyzing assembly output and comparisons

2:45 pm

Sequencing and Data QC

Lauren Brinkac and Indresh Singh, JCVI

- Evaluating quality of data
- Trimming low quality or contaminant data
- Genome QC (identify and remove contaminants)
- Genome Typing using MASH

3:30 pm

Afternoon Break

3:45 pm

Annotating Genomes in PATRIC

Marcus Nguyen, PATRIC

- Uploading or selecting contigs in workspace
- Selecting genome metadata and annotation options
- Queuing jobs and retrieving results
- Previewing an annotated genome

4:30 pm

Question and Answer

PATRIC and JCVI

5:00 pm

Day 1 Adjourns

Thursday, September 7, 2017

Hands-On Sessions

9:00 am

Phenotype Predicting Using Machine Learning

Marcus Nguyen, PATRIC

- Introducing AMR Classifiers

9:30 am

Comparative Genomics

Rebecca Wattam, PATRIC

- Selecting genomes for a study
- Visualizing compared genomes
- Downloading and analyzing results

10:30 am

Morning Break

11:00 am

Genome Typing and PanGenome Analysis

Lauren Brinkac and Indresh Singh, JCVI

- Genome Typing using LOCUST/MASH
- Setting-up PanGenome
- Running PanGenome

- PanGenome Visualization and Interpretation
- AMR Analysis

12:00 pm Lunch

1:00 pm Microbiome (16S and ITS) Data Analysis

Lauren Brinkac and Indresh Singh, JCVI

- 16S and ITS analysis
- Interpreting Data and Visualization

2:15 pm Afternoon Break

2:30 pm SNP and MNP Variation service

Rebecca Wattam, PATRIC

- Uploading reads data
- Selecting genomes and mapping
- Downloading and using data in PATRIC

3:30 pm Question and Answer

PATRIC and JCVI

4:30 pm Day 2 Adjourns

Friday, September 8, 2017

Hands-On Sessions

9:00 am RNA-seq and Differential Gene Expression Analysis in PATRIC

Maulik Shukla, PATRIC

- Uploading reads and specifying experimental conditions
- Queuing jobs and monitoring progress
- Filtering differentially expressed genes based on significance
- Viewing results in heat-map viewer
- Exploring public gene expression in datasets available in PATRIC

9:45 am Analysis of shotgun metagenomic data

Lauren Brinkac, Harinder Singh and Indresh Singh, JCVI

- Shotgun Metagenomic Assembly & Annotation
- Interpreting Data and Visualization

12:00 pm Lunch and Distribution of Certificates

1:00 pm	Optional Breakout: Using the PATRIC Unix/Mac Command Line Interface <i>Tom Bretin, PATRIC</i>
2:00 pm	Analyzing THSTI Datasets: Instruction One-on-One <i>JCVI/PATRIC/THSTI</i>
3:00 pm	Conclusions and future directions Prof. MK Bhan, Prof. Maria Giovanni, Prof. Partha P. Majumder