Senior Training Fellowship

1. Name: RV Shaji

2. Department: Haematology

3. Year of STF: 2017

4. Objectives stated in the application for STF:

Learn bioinformatics methods to analyze the data generated by next generation sequencing, specifically, exome sequencing and RNA sequencing for service and research.

5. Center visited for STF & Mentor : University of Chicago, Dr. John Cunningham

6. Short description of training:

During this training I learnt the basic methods for analyzing the data obtained after next generation sequencing. The training involved exome sequencing analysis to identify the mutations present in the patients with haematological diseases and RNA sequencing to analyze the gene expression changes in different cell types.

I used Galaxy (usegalaxy.org) to develop pipelines for the analysis. I analyzed 8 exome sequencing datasets and 9 RNA sequencing datasets. By exome sequencing I could find disease causing mutations in 7 samples and RNA sequencing analysis of cells obtained from patients with Fanconi anaemia showed distinct gene expression patterns, which were later confirmed by real time-PCR.

Galaxy provides the interphase to develop the workflows for the analysis of all the types of NGS datasets. In this basic training, I have learnt all the protocols and the parameters required for NGS data analysis.

7. Plans to implement objectives on return to CMC:

The protocols and the workflows for exome sequencing and RNA sequencing have already been implemented in the department, after my return. Now we have established an amplicon based NGS, a new method for the diagnosis of the inherited haematological diseases. The exome sequencing bioinformatics analysis strategy was applied for the amplicon NGS data analysis. The amplicon NGS data results could provide accurate diagnosis of Facnoni anaemia, Diamond Blackfan aanemia and congenital Dyserythropoietic anaemia, thalassaemias, haemophilia. This helped us to become the most comprehensive molecular diagnosis centre for haematological diseases in the country.

Exome sequencing analysis has been carried out in the department for 24 samples for patients with Fanconi anaemia to identify new genes involved in this disease. We could identify mutations in 22 patients, and these results were comparable to that obtained using commercial software.

The training helped me to become confident to do bioinformatics analysis of the NGS data generated from the NGS equipment we have in our department. I have trained a research fellow in the basic bioinformatics analysis. He is going to register for PhD. He will get trained in the various bioinformatics centers in the country to establish comprehensive bioinformatics methods required for NGS. I have also trained the post-doctoral fellows and technicians in the basics bioinformatic analysis required for detection of pathogenic mutations in the patients with haematological disease.

I have initiated conversations with Dr. Jayaseelan of Department of Bioinformatics to establish a bioinformatics programme in the institute. We are in the process of acquiring a computer required for large data set analysis. After that is complete, we will start collaborative projects between Department of Haematology and Biostatistics. This will help in establishing a bioinformatics programme in our institute. We will also submit proposals to DBT for capacity building. This will be completed by December 2019.

25/3/2019

Date:

Signature

Dr. R. V. SHAJI, Ph.D.,
Professor of Molecular Biology
Department of Haematology,
Christian Medical College,
VELLORE - 632 004, T.N., IND