

## NAHEP-CAAST

ICAR-Indian Agricultural Research Institute, New Delhi 110 012  
Application form for NAHEP sponsored Training programme on  
"High Dimensional Genome data Analysis by R and Open Source  
Tools" from November 01-11, 2019

(Website: <http://nahep-caast.iari.res.in>)

| 1. Name   | :       |      |                               |                        |
|---|---------|------|-------------------------------|------------------------|
| 2. Gender (Male or Female or others)  | :       |      |                               |                        |
| 3. Division and Degree programme  | :       |      |                               |                        |
| 4. Age and date of birth  | :       |      |                               |                        |
| 5. Category (Gen/OBC/SC/ST/ PC)   | :       |      |                               |                        |
| 6. Communication address  | :       |      |                               |                        |
| 7. Phone & Email  | :       |      |                               |                        |
| 8. Permanent address (For use in case of emergency)   | :       |      |                               |                        |
| 9. Educational qualifications (From Graduation onwards)   | :       |      |                               |                        |
| Degree  | Subject | Year | Percentage of marks/ Division | Name of the University |
| i.  |         |      |                               |                        |
| ii.   |         |      |                               |                        |
| iii.  |         |      |                               |                        |
| iv.   |         |      |                               |                        |
| v.  |         |      |                               |                        |
| <b>10. Research activities</b>  |         |      |                               |                        |
| a. Area of PhD or MSc research and title of approved thesis title (Attach ORW)  | :       |      |                               |                        |
| b. Indicate the future plans on utilizing the technical expertise gained from the winter school programme in your research (Attach Separate Sheet if necessary) | :       |      |                               |                        |
| c. Indicate whether you have attended any similar training programme earlier  | :       |      |                               |                        |
| 11. Write in brief (not exceeding 100 words) about the expected benefits of this training.  | :       |      |                               |                        |

Signature of the Applicant

Forwarding note by Chairman/Guide

Endorsement & Seal of the Professor

## Organizers

### Course Director

**Dr. A. R. Rao**

Core Team Member,  
NAHEP-Centre for Advanced *Agricultural Science and Technology* (CAAST) and  
Principal Scientist & Professor (Bioinformatics),  
Centre for Agricultural Bioinformatics (CABin),  
ICAR-Indian Agricultural Statistics Research Institute,  
Library Avenue, Pusa, New Delhi-110012,  
Email: [ar.rao@icar.gov.in](mailto:ar.rao@icar.gov.in),  
Phone: 011-25847121-24/Extn.4330, Mob: 09999422935

### Course Coordinators

**Mr. Sanjeev Kumar**

Associate Team Member, NAHEP-CAAST and  
Scientist, ICAR-IASRI, New Delhi-110012  
Email: [sanjeevbunty@gmail.com](mailto:sanjeevbunty@gmail.com)

**Dr. Soumen Paul**

Associate Team Member, NAHEP-CAAST and  
Scientist, ICAR-IASRI, New Delhi-110012  
Email: [soumen.pal@icar.gov.in](mailto:soumen.pal@icar.gov.in)

**Dr. Prabina Kumar Meher**

Scientist, ICAR-IASRI, New Delhi-110012  
Email: [meherprabin@yahoo.com](mailto:meherprabin@yahoo.com)

### Research Staff

**Dr. Tanmaya Kumar Sahu,**

Research Associate, NAHEP-CAAST, ICAR-IASRI,  
New Delhi-110012, Email: [tanmayabioinfo@gmail.com](mailto:tanmayabioinfo@gmail.com)

**Mrs. Manjeet Kaur**

Chief Technical Officer, CABin, ICAR-IASRI,  
New Delhi-110012, Email: [mkjeet.31@gmail.com](mailto:mkjeet.31@gmail.com)



National Agricultural Higher  
Education Project (NAHEP)

Sponsored  
Training programme

High Dimensional Genome  
data Analysis by R and  
Open Source Tools

November 01 - 11, 2019



Organized by

Center for Advanced Agricultural Science and  
Technology (CAAST)

ICAR-Indian Agricultural Statistics Research Institute  
Indian Agricultural Research Institute  
Pusa Campus, New Delhi- 110012

## About NAHEP-CAAST

Centre for Advanced Agricultural Science and Technology (CAAST) is a new initiative and student centric sub-component of World Bank sponsored National Agricultural Higher Education Project (NAHEP) granted to IARI to provide a platform for strengthening educational and research activities of post graduate and doctoral students. CAAST theme for IARI is **Genomic assisted crop improvement and resource management** that specifically aims at inculcating genomics literacy and skills among the students of National Agricultural Research Education and Extension (NAREE) System.

### Objectives

- To develop online teaching facility and online courses for enhancing the teaching and learning efficiency, and scientific communications skills
- To develop and/or strengthen state-of-the art next-generation genomics and phenomics facilities for producing quality PG and Ph.D. students
- To develop collaborative research programmes with institutes of international repute and industries in the area of genomics and phenomics
- To enhance the skills of faculty and PG students of IARI and NARES
- To generate and analyze big data in genomics and phenomics of crops, microbes and pests for genomics augmentation of crop improvement and management

### Rationale

Advances in information technology and computational methods are driving the mathematical sciences forward. In particular, Artificial Intelligence and Big Data analytics have revolutionized the bioinformatics based genome data analysis. High-throughput genomic technologies have made it possible to generate massive -omics data for understanding the complex phenomena involved in biological systems. With the availability of next generation sequencing (NGS) technologies, voluminous structured and unstructured data is now available. High-dimensional genomic data analysis is also challenging due to presence of noise and biases. Thus, computational analysis of such high-dimensional data often includes identification and correction of hidden biases, dimensionality reduction and application of AI / machine learning techniques for mining the hidden information to answer unsolved problems.

Moreover, combined analysis of high dimensional genome data from various sources to interpret the biological insights is highly difficult. However, to some extent, these difficulties can be overcome by use of open source software like "R" and tools based on R programming language. Thus, the aims of this training are: (i) To train post graduate students in handling high dimensional -omics data (ii) To familiarize the students with R and other open source software for analysis of -omics data (iii) to upgrade analytical skills of the participants.

### Venue

Discipline of Bioinformatics, PG School, IARI & Centre for Agricultural Bioinformatics (CABin), ICAR-Indian Agricultural Statistics Research Institute, Library Avenue, Pusa, New Delhi.

### Duration

01 – 11, November 2019

### Who Can Participate?

The Indian Agricultural Research Institute, New Delhi, invites applications from M.Sc. and Ph.D. students of ICAR-DUs/SAUs/CAUs for 10 days **Training programme on "High Dimensional Genome Data Analysis by R and Open Source Tools"** sponsored by **NAHEP-Centre for Advanced Agricultural Science and Technology (CAAST)**, Indian Council of Agricultural Research, New Delhi.

M.Sc. and Ph.D. degree students from any branch of Agriculture, Veterinary Sciences, Fisheries, Basic Sciences, *etc.* under National Agricultural Research Education and Extension (NAREE) System are eligible to apply. The number of participants will be limited to **TWENTY FIVE only**.

### Registration Fees

No registration fee is to be paid; the programme is fully sponsored by NAHEP-CAAST

### Food and Accommodation

Students will be paid travel fare to and from by rail restricted to the maximum of AC III Tier. Free lodging will be provided to the participants during the training program and boarding will be provided as per World Bank guidelines

### How to Apply

Complete application form in the prescribed format forwarded by the Professor/Dean or equivalent should reach the Course Director, NAHEP-Centre for Advanced Agricultural Science and Technology (CAAST), CABin, ICAR-IASRI, Pusa, New Delhi on or before **18<sup>th</sup> October 2019**.

### Theory Lectures

Room No. 109, Committee Room,  
Discipline of Bioinformatics, PG School, IARI,  
Centre for Agricultural Bioinformatics (CABin),  
ICAR-Indian Agricultural Statistics Research Institute

### Practical Lectures

Room No. 110, Training Lab,  
Discipline of Bioinformatics, PG School, IARI,  
Centre for Agricultural Bioinformatics (CABin),  
ICAR-Indian Agricultural Statistics Research Institute

### Course Contents

The course has been structured in a series of modules with the class room lectures and practicals on computers, including the demonstration of R software and tools based on R programming language. The course manual will be provided to all the participants. Participants would be given opportunity to share their research experiences in group discussions.

#### **Module – 1: Basics of R**

Data management, I/O format, File management.

#### **Module – 2: Univariate data analysis**

Descriptive statistics, testing hypotheses: t, F,  $\chi^2$ , design of experiments: CRD, RBD, LSD, BIBD, Lattice design, GXE & Biplot, Analysis of breeding data, ANOVA, Linear Models.

#### **Module – 3: Multivariate data analysis (unsupervised)**

Cluster Analysis, Principal Component Analysis, Discriminant Analysis, MANOVA.

#### **Module – 4: Supervised learning**

ANN, SVM, Random Forest and their application in the analysis of genomic data.

#### **Module – 5: Whole genome data analysis**

QTL, Genome Wide Association Study (GWAS), Genomic Selection (GS), transcriptomics, Genome assembly, Annotation, Variant (SNPs) Analysis, Genome editing.