The Lal Research Group of the Cologne Center for Genomics, Cologne Germany is hosting a three-day coding workshop (December 19th-21st 2018) on:

Basic Coding, Computational Genetics, and Applications in Medical Research

Molecular techniques have evolved and allow researchers today to gather a massive amount of data. Frequently, the data is too complex or big to be analyzed without a computer and only shallow analyses can be performed with available software. Basic coding skills are beneficial/essential for data analysis in today's biomedical research landscape. This observation holds true particularly for medical genetic/genomic research. With the introduction and evolution of next-generation sequencing platforms, it is now feasible to analyze the exome and genome of thousands and soon millions of people. The biological and clinical interpretation of variation in disease-associated genes lags far behind data generation. Many efficient computational approaches for the prediction of protein function and the assessment of variant pathogenicity have been developed but to use them efficiently and integrate them into workflows basic programming skills are mandatory.

What is the work about?

In this workshop, we will provide a brief introduction into basic coding and will give an overview of current computational approaches in translational genomics and computational biology. The workshop is organized in two modules. **Module 1** is "stand alone" and can be registered without the **Module 2**. **Module 2** requires prior developed programming skills or attendance at **Module 1**. The content will be taught in short lectures followed by hands-on coding sections supported by an individual mentor for small groups of 2-4 participants. One week in advance of the workshop, teaching material will be shared with participants and preparation for the workshop is expected.

Module 1: By the end of the module participants will be able to:

- Use basic bash shell commands in a Unix/Linux command line environment;
- Download and install R/Python;
- Understand basic operations in R/Python environments;
- Read and transform data from multiple sources and formats;
- Understand, manipulate and explore different types of R/Python data objects such as vectors, matrices, and data frames;
- Find help about a given command and explore similar commands;
- Use R/Python script files to organise R/Python commands;
- Display summary statistics and basic plots; and
- Download, install and find documentation for additional R/Python libraries.

Module	Date	Time	Module topic	Content
1	WED 19th	9a-5p	Basic of programming using Unix/Linux, R, and Python	• Bash/Linux introduction • Basic programming in R • Basic programing in Python

Module 2: By the end of the module participants will be able to:

- Understand the major sources of genetic variation in the genome;

- Understand different types of genetic data and discovery technologies;
- Quality control different types of genetic and genomic data;
- Carry-out genetic association analyses while considering population structure;
- Utilise online bioinformatic resources to explore the functional properties of genetic variation;
- Interpret variant pathogenicity in the context of commonly-used guidelines and changes in protein structure features;
- Utilise results from genetic studies to investigate clinical phenotype correlation.

Module	Date	Time	Module topic	Content
2	THU 20th	9a-1p	Analysis of genomic data	• Variant calling • Quality control of genotyping and sequencing data • Ethnicity stratification • Association analysis
2	THU 20th	2-5p	Functional prediction of patient variants	• Annotation of genetic and biological database information • Annotation and interpretation of prediction scores
2	FRI 21st	9a-1p	From variants to proteins	• Review of current variant interpretation guidelines • Mapping and modeling of genetic variants in protein structures
2	FRI 21st	2-5p	Genotype-phenotype correlation analysis	• Correlation of genetic variants, prediction scores, protein features, and clinical data

Costs: Attendance to the workshop is free.
Venue: The workshop will take place at CMMC large seminar room (map).
Mentors: Computational biologists from: i) University of Cologne, Germany; ii) Cleveland Clinic/Case Western, USA; iii) Broad Institute of MIT and Harvard, USA; iv) University of Luxembourg, Luxembourg.
Target audience: Graduate students, postdoctoral scholars, clinical scientists, and principal investigators currently working with genetic data, or about to embark on projects that require analysis of such data.
Prerequisites: All participants are expected to bring a laptop to the workshop (more details later) For **Module 1** no prior programming experience is required. **For Module 2 attendance at Module 1 or basic familiarity with UNIX/Linux and a scripting language (Python or R) are required**.
How to apply: Due to space constraints, we can only accept a limited number of participants. Please send a short quarter page personal statement/essay about why you want to participate, which module(s) you want to join and to which degree you fulfill the workshop prerequisites to **dlal@broadinstitute.org** .