

September, the 5th 2018

Title: Mixture regression to identify latent pathways for genetic association studies

Context & objectives

Applications are invited for an internship at the Institut Pasteur within the Statistical Genetics group in the Center of Bioinformatics, Biostatistics and Integrative Biology (C3BI).

Large genome-wide association studies (GWAS) have been successful in identifying thousands of significant genetic associations for multiple traits and diseases. While the application of standard univariate approaches for detecting new genetic variants associated with traits and diseases will continue as sample size increases, a range of multi-traits analyses have been proposed in parallel to address questions about the joint genetic architecture of human phenotypes. In particular, several recent studies have discussed the application of Mendelian Randomization (MR) principle to GWAS summary statistics to assess potential causal relationship between human phenotypes. Our group recently explored similar approaches based on mixture regression with a primary objective of determining latent classes across genetic variants. The goal of this internship is to continue the development and application of the approach.

The internship involves statistical method development and application, and the optimization of their implementation. The objectives of the candidate will be multiple folds:

- 1) Understanding the basis of genetic-phenotypes associations and familiarizing with genetics data.
- 2) Understanding the basis of Mendelian Randomization applied to genetic association studies summary statistics.
- 3) Implementing and optimizing a mixture regression model for the inference of latent classes for genetic variants.
- 4) Assessing the approach in simple simulation studies.
- 5) Performing real data application of the approach in GWAS summary statistics from over 50 phenotypes.

The selected candidate will be mentored by Dr. Hugues Aschard, but will also work with members of our research group and international collaborators involved in the project. She/he will have access to all resources at Pasteur, including in particular the High Performance Computing Cluster which includes over 2,000 cores.

QUALIFICATIONS

The position requires advance knowledge in statistics and computer sciences. The applicants should therefore have substantial educational background in Statistics/Biostatistics, Bioinformatics, Computer Science or other relevant disciplines. The call specifically addresses master 2 students and 3rd year engineer school students.

ADDITIONAL INFORMATION

Interested applicants should send their curriculum vitae, a brief cover letter, and contact information from at least one referee (e.g. teacher or previous internship mentor) to Dr. Hugues Aschard (hugues.aschard@pasteur.fr). More information about our group and the C3BI can be found here <https://research.pasteur.fr/en/team/statistical-genetics/> and here <https://research.pasteur.fr/en/center/c3bi/>.