

Internship projects offered by BMR Genomics

BMR Genomics started its activity in 2004, offering its expertise in DNA sequencing and bioinformatics to universities, businesses and customers. To keep the pace with the continuous advancements of genomics, BMR spends a considerable effort in research and development, in particular for data analysis. With this premise, BMR is proposing two projects for Data Science Internships.

1. Development and application of AI methods for the analysis of gut microbiome communities.

It has been calculated that the number of bacteria in our body is higher than the number of human cells. Indeed, in our gut there are many species of bacteria that have implications with different processes of intestine, immunity system and brain. The species of bacteria and their genomes can be easily established by DNA sequencing, but the correlations between the different bacterial communities and the consequent conditions of our body is a challenging problem. The aim of this project is to apply approaches of machine learning and AI to improve the interpretation of the metagenomic data.

2. Analysis of data from Genome Wide Association Studies (GWAS). Many GWAS data are becoming available in the public domain. In general, GWAS data include a large number of individuals (several tens of thousands) and a very large number of genetic loci (several hundreds of thousands). The aim of GWAS is to find correlations between single nucleotide polymorphisms (SNPs) present in these loci and quantitative traits such as risks of diseases, food intolerance, disposition to physical activity and much more. In most cases the variants that are identified by GWAS are not causative variants, but they are close to them. Since the rate of crossing over is relatively low, they tend to remain associated. In other words, several hundreds of thousands of variants may seem a lot, but our genome contains more than 3 billion bases. We have a very partial picture. The aim of this project is to apply imputation methods to fill the gaps and identify some causative variants.