

## Some statistical models for genomic analysis

Stéphane. Robin

**1st course.** We will first consider the analysis of CGH (comparative genomic hybridization) arrays which allow us to detect chromosomal aberrations. We will introduce segmentation models and consider their inference from both algorithmic and statistical point of views. We will also consider model selection issues associated with their analysis.

**1st practical.** A first practical will be dedicated to the analysis of some real CGH profiles.

**2nd course.** We will then consider « high density » array analysis in which we often face probe classification issues. Mixture models are a natural way to deal with such unsupervised classification problem. These models also enable us to account for some spatial dependency of the signal.

**3rd course.** Multiple testing issues arise in the analysis of most « high throughput » technology. We typically want to control the overall type I error rate to avoid numerous false positive. We will recall classical multiple testing procedure such as Bonferonni or FDR. We will also consider the estimation of the local FDR that can be stated in a mixture model context.

**2<sup>nd</sup> practical.** A second practical will be dedicated to the implementation of the E-M algorithm, in view of mixture model inference.

**4th course.** Networks constitute a natural way to describe how the elements present in the cell interact with each other. The understanding of the organisation or the topology of this network requires the definition of proper and realistic random graph models. We will present some of these models and consider the problems raised by their inference.