Keynote speakers

Martial Marbouty *(Pasteur Institute, G5 group Spatial Regulation of Genomes)**:***
Metagenome assembly and characterization using high-throughput chromosome conformation capture data

Abstract:
Metagenome sequence analysis relies principally on compositional approaches, which hypothesise that sequences sharing similar characteristics (GC%, codon bias, etc.) Although these approaches have generated important results, they remain somehow limited and do not allow the full characterization and understanding of the genetic composition of a complex microbial population. Contact genomics, which aimed at exploiting the 3D physical signature of genomes to solve their sequence, has the potential to alleviate or improve some of these caveats. To explore the genomic content of bacterial population at a new level of resolution we have recently developed meta3C (Marbouty et al. 2014), a derivative protocol of the chromosome conformation capture (3C; Dekker et al. 2002) assay that aims at deciphering the average 3D organization of a genome. Using controlled mixes of bacterial or yeast species, we showed that the frequent collisions experienced by DNA molecules sharing similar cellular compartment can be measured through meta3C and conveniently used to assemble larger scaffolds of the genomes present in a metapopulation. Here i will present new data obtained from a complex mix of bacterial species of the gut microbiota of mouse. Meta3C was performed directly onto a sample of mouse feces, unveiling hundreds of genomic compartments, hence species, some sharing similar DNA elements. We will discuss the different way to explore this network and the results in light of the promising potential of the approach for future applications.
Joint work with Baudry L., & Koszul R.

Rayan Chikhi:
de Bruijn graphs of sequencing data

Asbtract:
The keynote will be a gentle introduction to de Bruijn graphs. Although these are purely mathematical objects, we will focus on their usage in bioinformatics. De Bruijn graphs are indeed widely used in algorithms to analyze DNA and RNA sequencing data. Formally, a de Bruijn graph is a directed graph, where nodes are all substrings of length k present in the data (k-mers). Edges represent exact overlaps of length k-1 between pairs of nodes. These overlaps between k-mers precisely reflect the structure of the sequenced genome(s) or transcriptome(s). Note that de Bruijn graphs can be constructed for both model and non-model organisms: generally, the nodes of the graph are all k-mers present in a set of sequencing reads. Analysis algorithms often rely on the property that repetitions, variants and sequencing artifacts appear as special structures in the de Bruijn graph.

We will give an overview of how de Bruijn graphs look in practice, and their main applications: de novo genome, metagenome and transcriptome assembly.We will also talk about emerging applications: reference-free variants detection and RNA-seq quantification. Finally, we will see some statistical aspects of k-mer abundance histograms.

Axel Munk *(Georg August University Goettingen and Max Planck Institut for Biophysical Chemistry, Goettingen)**:***
False positives in microRNA target prediction

Abstract:
Statistical MUltiscale Change point Estimation (SMUCE) is an inference tool for estimation and inference about a change-point regression function and its main characteristics location, size, and number of jumps. SMUCE detects these features on all scales in an optimal fashion and provides confidence statements for all quantities.
The methodology can be extended to a variety of other situations, including time series change

point regression, phase regression, heterogeneous models, and blind source separation problems. Fast computation of SMUCE via accelerated dynamic programming is addressed and the R package StepR is introduced. Several ion channels which regulate the protein transport at the cell membrane will be analyzed and we report on some new findigs for gramicidin channels.
This is joint work with M. Behr, H. T. Hotz, Li, K. Frick, F. Pein and H. Sieling and the Steinem lab at the Institute for Biomolecular Chemistry, University of Goettingen.

Hervé Seitz:
False positives in microRNA target prediction

Abstract:
According to the current view, each microRNA regulates hundreds of genes. Computational tools aim at identifying microRNA targets, usually selecting evolutionarily conserved microRNA binding sites. Here we provide evidence that such predictions are often biologically irrelevant. Focusing on miR-223-guided repression, we observed that it is often smaller than inter-individual variability in gene expression among wild-type mice, suggesting that most predicted targets are functionally insensitive to the microRNA. Accordingly, we found a significant correlation between predictors of gene dose-sensitivity and the conservation of microRNA target sites. Secondly, many conserved microRNA binding sites are conserved in a microRNA-independent fashion: sequence elements may be conserved for other reasons, while being fortuitously complementary to microRNAs. Thirdly, some mRNAs can efficiently titrate microRNAs, providing yet another reason for microRNA binding site conservation despite their inefficiency in target repression. Collectively, our data suggest that the role of microRNAs has been over-estimated.