



AlgoB

the seminar you can't miss

<http://algeb.univ-mlv.fr/seminar>

# Spaced seeds improve metagenomic classification

Gregory Kucherov

LIGM Paris-Est

Tuesday, April 28, 2015 at 11:00 in 4B084

Laboratoire d'Informatique Gaspard-Monge  
Université Paris-Est Marne-la-Vallée  
5 bd Descartes, Champs-sur-Marne

**Abstract:** Metagenomics is a powerful approach to study genetic content of environmental samples that has been strongly promoted by NGS technologies. To cope with massive data involved in modern metagenomic projects, recent tools (Kraken, LMAT) rely on the analysis of k-mers shared between the read to be classified and sampled reference genomes. Within this general framework, we show that spaced seeds provide a significant improvement of classification capacity as opposed to traditional contiguous k-mers. We support this thesis through a series of different computational experiments, including simulations of large-scale metagenomic projects.