

CRISPRscope: a comparative analysis pipeline of CRISPR diversity in cheese related genomes and metagenomes

Thibault Schowing

Master thesis in Bioinformatics and Computational Biology

Bacteria resist phage infections and other mobile genetic elements by incorporating a fragment of the foreign DNA, called a spacer, in their CRISPR loci, enabling recognition and cleavage of the sequence upon new infections. Phage resistance is vital for industrial processes requiring bacteria as an infection can quickly destroy the necessary balance in bacterial communities or even obliterate an entire population causing irreversible damages to the end products. The CRISPR systems in genome assemblies and in metagenomes have been extensively studied and it remains an active research domain. However, the variations of CRISPR diversity between genome assemblies and metagenomes remain unknown. We profiled CRISPR loci and cas genes among 1831 assemblies of 12 Lactic Acid Bacteria species and 7 metagenomes related to cheese core. We showed that Cas types are widely shared among the different species as well as among the metagenomes but are yet more diversely represented in the metagenomes. Repeats have been shown to be species specific while spacers show a high strain specificity, confirming that each individual adapts to its specific phages encounters. If the spacers are known, their target phages remain mysterious for an important part of them. Ultimately, spacer and target dynamics in the metagenomes showed a Piggyback-The-Winner distribution, predicting that phages integrate into their hosts genomes as prophages when microbial abundances and growth rates are high.

Superviseur : Rémy Bruggmann