NANOPORE SEQUENCING ENABLES DETAILED METAGENOMICS OF CLOSELY RELATED BACTERIAL STRAINS USED IN DAIRY INDUSTRY

Denitsa V. Stefanova[1], Daniella Lucena[1], Josué L. Castro-Mejía[1], Yan Hui[1], Göksen Arik[1], Susann Romy Gröbner[1], Paulina Deptula[1], Finn K. Vogensen[1], Dennis S. Nielsen[1], Lukasz Krych[1]

[1]Food Microbiology & Fermentation, Department of Food Science, University of Copenhagen, 1958 Frederiksberg C, Denmark;
[2] Arla Foods amba; Arla Innovation Centre, Agro Food Park 19, DK-8200 Aarhus N

INTRODUCTION

Diacetyl, acetoin, acetate, acetaldehyde and ethanol are central compounds for the quality and distinct flavor profile of many dairy products. These compounds are produced during complex fermentation processes. Understanding such processes are far from trivial, as important technological properties differ at the bacterial strain level interactions playing a strong role for final quality. Many dairy products (e.g., buttermilk and many cheese types) are produced using undefined starter cultures containing several different species further divided into dozens of different strains. Recent developments in long read DNA sequencing offered by Oxford Nanopore Technologies give new opportunities to study complex microbial cultures with precision never seen before. Detailed genomics combined with high-resolution data on aroma formation during fermentations opens new possibilities for developing machine learning-based tools.

We have recovered full, complete genomes (including plasmids) of 60 bacterial strains from a mesophilic starter culture used for fermentation of dairy products. Detailed metagenomics disclosed pivotal, unique, strain specific features of closely related bacteria. This includes unique phage resistance mechanisms, mobile elements and gene clusters playing role in flavour formations of dairy products.

MATERIALS & METHODS

ON-REP SEQ

DAIRY

COMPLETE GENOMES

GENES SYNTENY

WET LAB

1.Propagation and Preparation for DNA extraction of mesophilic starter culture

200 bacterial strains were isolated from mesophilic starter culture and screened with ON-rep-seq (Figure 1). 60 isolates were selected for the Whole Genome sequencing using Oxford Nanopore Technology (ONT).

2. DNA extraction has been performed using Bead-Beat Micro AX Gravity (A&A BIOTECHNOLOGY), according to the manufacturer's procedure. DNA was concentrated and size selected (reads > 5kB) using custom made beads.

3. High molecular weight DNA (~50 ng/µl) was used for library preparation with the Rapid Barcoding Sequencing Protocol (SQK-RBK004). The sequencing was performed using the R9.4.1 flow cells on GridIONx5(ONT).

4. The data were analyzed with newly developed pipeline utilizing open access

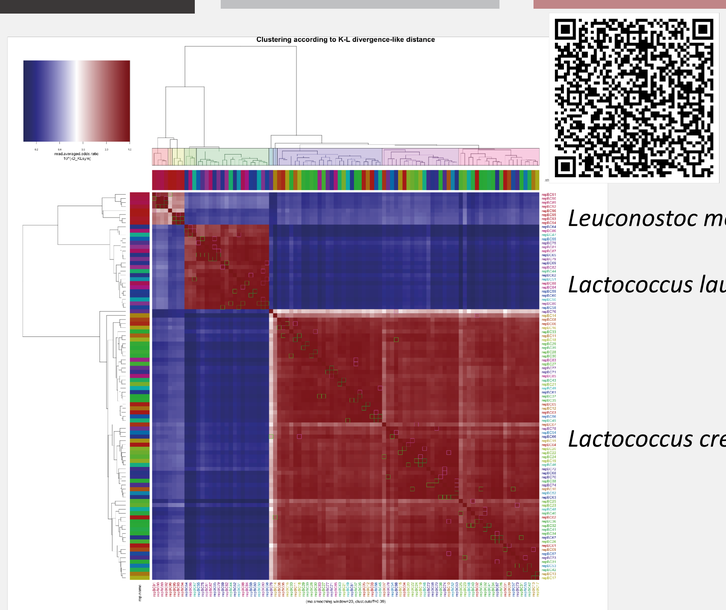
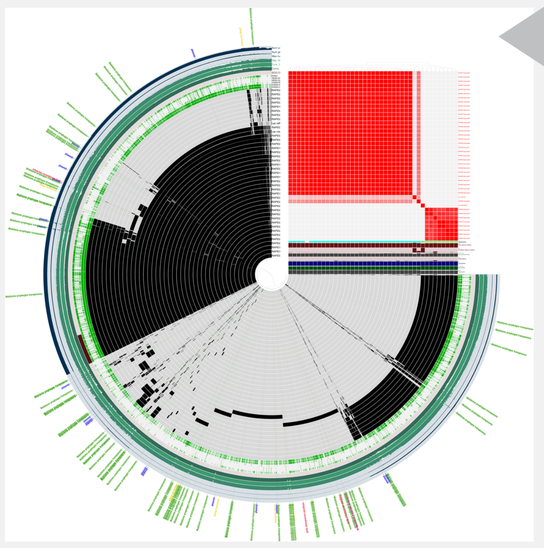


Figure 1. On-rep-seq results. Heatmap showing similarity (10^(-D_KLsym)), and clusters according to cut-off = 0.09 of three major species. Additional subclusters indicate presence of various strains within given species cluster

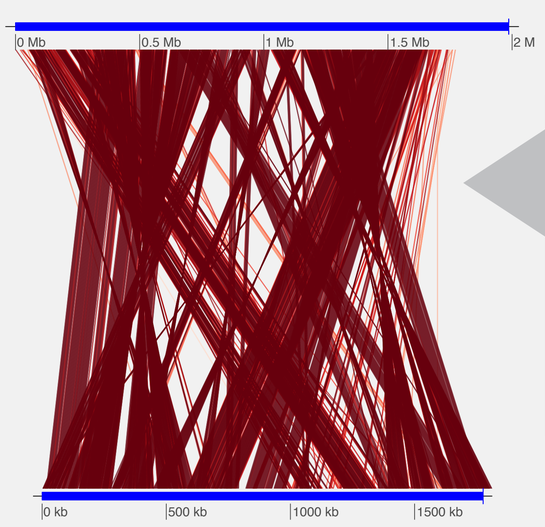
DATA ANALYSIS

RESULTS



Comparative genomics with Cladogram depicting gene clusters of 50 complete bacterial genomes

Overview of the full genome assembly pipeline



Comparison of recovered full genome with the complete genome related bacterium (*Leuconostoc mesentroides*) (ANI > 97%) from the database showing high level of genome rearmaments

SUMMARY

We have recovered 50 full genomes of Lactic acid bacteria including *Leuconostoc mesenteroides*, *Lactococcus laudencis*, *Lactococcus cremoris* *from mesophilic starter culture.* The complete genome recovery of bacterial genomes gives highly detailed picture of the community allowing not only to study the genetical content but also genes synteny.

We observed that all genomes carried Restriction-Modification (RM) systems, and most of the genomes carried Abortive-infection (Abi) systems as well. However, several novel systems such as CBASS, BREX and Nhi were also detected, in a more variable fashion.

Full genomes analysis revealed existence of multiple minor differences (additional gene clusters, inversions, translocations) between closely related strains, yet their function is yet to be investigated.

With this data we seek to elucidate the biodiversity, composition and evolution of starter cultures. Futhermore, by reconstructing metabolic pathways we will look into aroma and flavor formation compounds of high relevance for dairy industry.

UNIVERSITY OF COPENHAGEN

Arla

LC 2020 LONDON CALLING