TITLE: Genomic characterization of two novel *Leuconostoc sp.* strains isolated from Canastra cheese endogenous starter culture

AUTHORS: OLIVEIRA, D. P.; D'OVIDIO, L.; LACORTE, G. A.; CRUVINEL, L. A.; HOFFMANN, C.

INSTITUTION: FOOD RESEARCH CENTER, DEPARTMENT OF FOOD AND EXPERIMENTAL NUTRITION, FACULTY OF PHARMACEUTICAL SCIENCES, UNIVERSITY OF SÃO PAULO, SP, BRAZIL (RUA DO LAGO, 250 – ED. SEMI INDUSTRIAL, BLOCO C, CEP: 05508-080, SÃO PAULO – SP, BRAZIL)

ABSTRACT:
Leuconostoc is one of the main heterofermentative lactic acid bacteria found in endogenous starter cultures used to produce many Brazilian artisanal cheeses, such as Canastra cheese. *Leuconostoc* spp. are known to be able to uptake and metabolize citrate to additional flavor compounds. The co-metabolism of citrate and lactose can lead to production of CO_2 and flavor compounds such as diacetyl (butter flavor), but *Leuconostoc* spp, can only produce diacetyl and acetoin from citrate at acidic pH, and acidification is predominantly driven by lactose fermentation by the lactococci also present in the starter culture. Here isolate two *Leuconostoc* strains and characterize them by whole genome sequencing. The two strains were obtained using decimal dilutions and standard colonize purification techniques, from two distinct Canastra cheese endogenous starter cultures, used by two cheese producers. Genome sequencing was carried out using the Illumina platform, and genomes assemblies SPAdes v3.14.1. Pangenomic analysis was carried out using Anvi'o v.6.2 and 8 published genomes from 7 distinct *Leuconostoc* species. Genome metabolic characterization was performed using GhostKoala and the KEGG database. The genomes have 97.2 of completeness and 1.4 of redundancy. The pangenomic analysis using found 2.970 (73%) gene clusters on the core genome which demonstrated the main preserved metabolisms, such as synthesis of vitamins and citrate. Based on the pangenomic analysis, the two novel strains likely belong to a novel *Leuconostoc* species, as the Average Nucleotide Identities (ANI) for each strain against known species were 75% and 83%, well below the 95% ANI accepted for identification of the same species. Both genomes presented enterocin coding genes, bacteriocins known to act against food borne pathogens, such as *Listeria monocytogenes*. The accessory genome restricted to our novel isolates that was detected in the pangenomic analysis included 1.092 (27%) gene clusters, including those responsible for the synthesis of diacetyl. In conclusion, we have detected a likely novel *Leuconostoc* species, which may be endogenous to the Canastra region. Future studies may determine if the the sensorial profile originating from the presence of diacetyl in Canastra cheese originates from occurrence of *Leuconostoc*.

Keywords: *Leuconostoc* spp, starter culture, genome, Canastra cheese, "pingo".

Funding Agency: CAPES; FAPESP