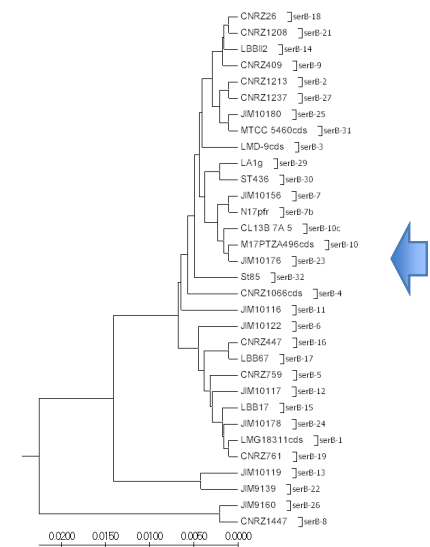


DIVERSITY OF *STREPTOCOCCUS THERMOPHILUS* POPULATIONS BY PHOSPHOSERINE PHOSPHATASE (ser-B) POLYMORPHISMS

Teresa Zotta^{*1}, Annamaria Ricciardi², Francesca De Filippis³, Danilo Ercolini³, Eugenio Parente⁴

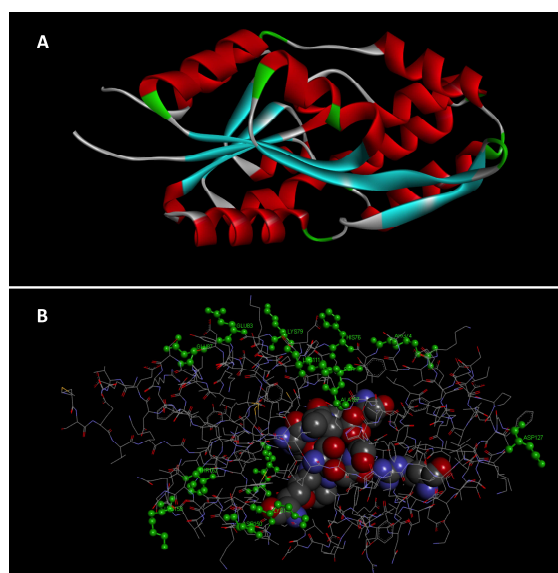
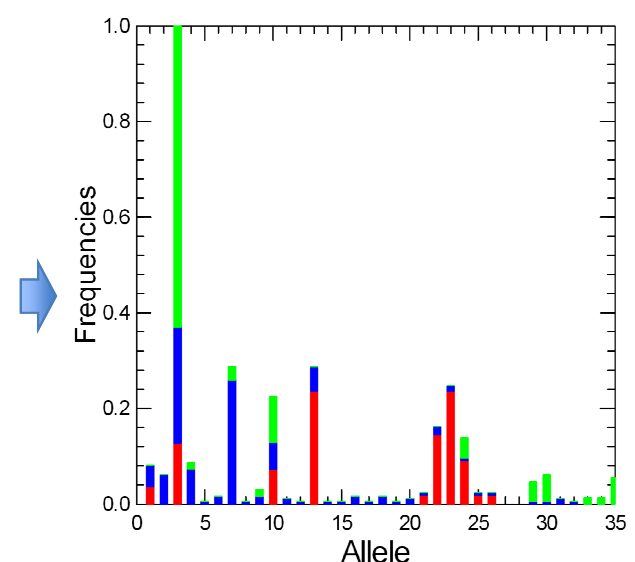
¹ Istituto di Scienze dell'Alimentazione-CNR, Avellino; ² Scuola SAFE, Università degli Studi della Basilicata, Potenza; ³ Dipartimento di Agraria, Università degli Studi di Napoli Federico II, Portici; ⁴ Dipartimento di Scienze, Università degli Studi della Basilicata, Potenza.

Streptococcus thermophilus is one of the most important dairy starters, and it is used in defined starters for the production of cheeses and fermented milks. It also plays a fundamental role in the production of traditional cheeses, and its sources may be the milk, cheese-making environment, and artisanal starter cultures. Monitoring the population structure of this species is of great interest in traditional cheeses produced using artisanal starters, because of the high diversity of these microbial communities. The genotypic diversity of *S. thermophilus* has been assessed by a variety of methods, including Multilocus Sequence Typing (MLST). One of the genes used as a target for MLST (<http://bigsd.bweb.pasteur.fr/streptococcus/streptococcus.html>), the phosphoserine phosphatase gene (*serB*), has high nucleotide diversity. This makes it a promising candidate for *S. thermophilus* biotyping (El Sharoud et al., 2012; Parente et al., 2016).



The partial *serB* sequences of *S. thermophilus* strains from our culture collection (isolated from artisanal starters, traditional cheeses and yoghurt; green) were compared with those available in databases (MLST database, blue; El Sharoud et al., red) and a limited number of alleles detected at a high frequency was found:

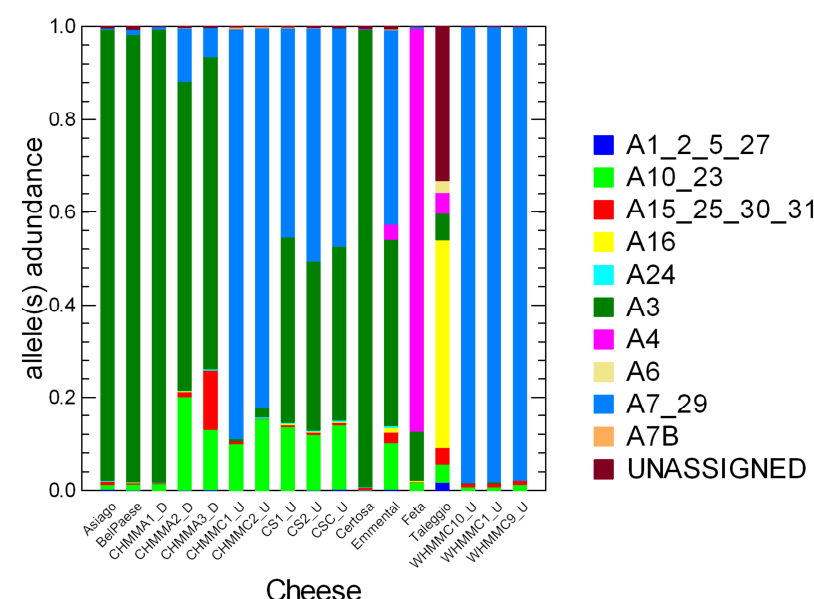
- Three sequence types were the most frequent in public databases and in our collection: both collections are dominated by cheese and starter isolates.
- Overall, one allele (*serB-3*) is the most frequent in cheese, yoghurt and starters. Other alleles (*serB-7*, *10*, *23*) are frequent in strains obtained from natural starters.
- Allele frequencies significantly differ for strains obtained in different geographical areas, although the food source may contribute to confounding this pattern.



3D modelling of the predicted protein structure revealed that amino acid changes (B in green) in different alleles are unlikely to interfere with activity

A *serB*-targeted high-throughput sequencing approach for a quantitative monitoring of *S. thermophilus* biotypes was developed (Parente et al., 2016) and used for a selection of cheeses:

- The average length of the sequences was relatively short (median value 471, range 441-476) and denoising revealed frequent sequencing errors in the form of insertions.
- The method was able to discriminate some sequence types (including *serB-3*) but not closely related alleles.
- In cheeses produced with defined strain starters, *serB-3* was the most frequent allele, but a higher diversity was found when artisanal starters (_U in the figure) were used.



Relative abundance of *serB* allele groups (A) in commercial cheeses, detected by HTS. CHMMA, cow milk high moisture mozzarella cheese. WHMMC, water-buffalo high moisture mozzarella cheese. CS, Caciocavallo Silano.

Conclusions: Even if frayed by some limitations (inability to discriminate all *serB* alleles, inability to detect new alleles due to sequencing errors) the high throughput sequencing method that we developed is still the only cultivation based approach for the detection of the structure of *S. thermophilus* populations in food environments.