

Whole genome sequences & analysis of a novel group of bacteriophages infecting the dairy bacterium *Streptococcus thermophilus*

B. McDonnell⁴, L. Hanemaaijer⁵, T. Kouwen⁵, J. Mahony⁴, H. Neve², J. Noben¹, D. Van Sinderen⁴⁻³
¹Biomedical Research Institute, Hasselt University, Diepenbeek, Belgium ²Max Rubner Institut, Kiel, Germany ³APC Microbiome Institute ⁴School of Microbiology, University College Cork, Cork, Ireland ⁵DSM Food Specialties, Delft, The Netherlands

Streptococcus thermophilus is a globally employed dairy bacterium used mainly in the production of cheese, yoghurt and related products. The Generally Regarded As Safe (GRAS) status of this species and favourable acidification and texturisation properties have ensured its widespread use in the international dairy industry. Despite a relatively recent general improvement in sanitary standards, increased technical knowledge and genetic information, contamination of dairy production lines by bacteriophages is a persistent and costly problem.

A total of thirteen genomes of phages infecting *S. thermophilus* have been published to date with a large degree of conservation observed within the defined groupings. Whole genome sequencing of these phages has enabled their genome-wide nucleotide-level comparison and elucidation of their putative mechanisms of evolution in dairy fermentations. Traditionally, the classification of phages of *S. thermophilus* has been based on 1) morphology, i.e. as Siphoviridae, 2) a variable genomic region thought to be (at least in part) responsible for host determination (VR2 region) or 3) the mode of DNA packaging and major structural protein content (cos-type or pac-type). More recently, a morphologically distinct and genetically divergent *S. thermophilus* phage (named 5093) containing neither cos- or pac-defining structural elements (nor a confirmed receptor binding protein-encoding gene) was described, necessitating the creation of a third group ('5093-type'). The genomic content of phage 5093 highlights the genetic plasticity in, and potential for diverse lineages of *S. thermophilus* phages.

Here, we present the complete genome sequences of four novel phages capable of infecting *S. thermophilus* BMD2, an industrial dairy starter strain. These phages appear to have undergone considerable genetic exchange with an unknown *L. lactis* phage of the P335 species. Comparative genomic analysis was performed on all four phages with archetypal phages infecting *S. thermophilus* and *L. lactis*. The phages were also characterized by microscopic analysis and adsorption analysis on industrial strains of both *S. thermophilus* and *L. lactis*. The proposed structural proteins of one representative phage of this group were also confirmed by mass spectrometry.