Novel *Siphoviridae* phage PMBT4 belonging to the group b *Lactobacillus delbrueckii* subsp. *bulgaricus* phages

Sabrina Sprotte[1], Olakunle Fagbemigun[2], Erik Brinks[1], Gyu-Sung Cho[1], Eoghan Casey[3], Folarin A. Oguntoyinbo[2,4], Horst Neve[1#], Jennifer Mahony[3], Douwe van Sinderen[3], Charles M.A.P. Franz[1*]

[1]Department of Microbiology and Biotechnology, Max Rubner-Institut, Federal Research Institute of Nutrition and Food, Hermann-Weigmann-Str. 1, 24103 Kiel, Germany; [2]Department of Microbiology, Faculty of Science, University of Lagos, Akoka, Lagos Nigeria; [3]School of Microbiology & APC Microbiome Ireland, University College Cork, Western Road, Cork T12 YT20, Ireland; [4]A.R. Smith Department of Chemistry and Fermentation Sciences, Appalachian State University, 730 River Street, Boone, NC, 28608, USA; oguntoyinbofa@appstate.edu

Authors S. Sprotte, O. Fagbemigun and E. Brinks share first authorship.

[#] retired from Max Rubner-Institut

Corresponding Author: Charles.Franz@mri.bund.de; Tel.: +49 431 609 2340.

Highlights:

- A novel *Cequinquevirus* PMBT4 was identified from African fermented milk product by electron microscopy and high-throughput sequencing (HTS).
- Complete genome sequence of the virus was obtained by HTS.
- The *Cequinquevirus* PMBT4 shares the highest NT identity with *Lb. delbrueckii* group b phage c5.

- No KIS gene element was found in the genome of phage PMBT4 despite the presence of observed collar structures on some phages previously suggested to be encoded by these genes.

Abstract. A novel *Lactobacillus delbrueckii* bacteriophage PMBT4 was isolated from the Nigerian fermented milk product *nono*. The phage possesses a long and thin, non-contractile tail and an isometric head, indicating that it belongs to the *Siphoviridae* family. A neck passage structure (`collar`), previously hypothesized to be encoded by two genes located in the *Lactobacillus delbrueckii* phage LL-K insertion sequence (KIS) element, as well as in two additional *Lb. delbrueckii* phages Ld17 and Ld25A, could also be observed on an estimated 1-5% of phage particles by transmission electron microscopy. However, neither mapping of high throughput sequencing data to KIS element genes from *Lb. delbrueckii* phages LL-K, Ld17 and Ld25A nor PCR amplification of the KIS element genes could corroborate the presence of these genes in the PMBT4 genome. The PMBT4 genome consists of 31,399 bp with a mol% GC content of 41.6 and exhibits high (95-96%) sequence homologies to *Lb. delbrueckii* phages c5, Ld3, Ld25A and Ld17, which assigned PMBT4 as a new member of this genus, i.e. the *Cequinquevirus* genus.

Keywords: *Lactobacillus delbrueckii* phage, genome, group b, KIS element

1. Introduction

Nono is a fermented cow milk product produced in the northern parts of Nigeria. It is predominantly prepared and sold on local markets by the Hausa/Fulani cattle herdsmen (Ogbonna, 2011). *Nono* is a spontaneously (lactic acid) fermented beverage (Okagbue and Bankole, 1992). So far, there have not been many studies performed on the diversity of LAB associated with *nono* fermentation. Banwo et al. (2012) showed that quite diverse lactic acid bacteria (LAB) are associated with *nono* production.

These LAB were identified as *Lactiplantibacillus plantarum, Enterococcus* and *Pediococcus* spp., while Okagbue and Bankole (1992) identified *Lactococcus lactis* subsp. *lactis* biovar *diacetylactis, Levilactobacillus brevis* and the yeast *Saccharomyces cerevisiae* to be associated with the fermentation. Knowledge of the bacteria involved is, however, a pre-requisite for successful starter culture development. In our previous study on key LAB in Nigerian *nono* samples, we identified *Lactobacillus* (*Lb*.) *helveticus*, *Limosilactobacillus fermentum*, *Streptococcus thermophilus* and *Lb. delbrueckii* subsp. *bulgaricus* to be predominant species associated with the fermentation (Fagbemigun et al., 2021). To achieve starter culture development for *nono* fermentation, therefore, suitable strains belonging to these species with appropriate technological properties should be considered. An important criterion for starter strain selection is that the strains should preferentially be resistant to bacteriophage infection. In the present study, therefore, we investigated Nigerian *nono* samples for the presence of phages with lytic activity against *Lb. delbrueckii* subsp. *bulgaricus* strains.

Currently, the databases include the genomes of nine phages infecting *Lb. delbrueckii* that are arranged in five genetically diverse groups (i.e. a to e): i) phage LL-H and temperate phage phiJB (group a) (Guo et al., 2016; Mikkonen et al., 1996), ii) phage LL-Ku (Riipinen et al., 2011), phage c5 (Riipinen et al., 2011), phages phiLdb, Ld3, Ld17 and Ld25A (group b) (Wang et al. 2010; Casey et al. 2014), iii) temperate phage JCL1032 (group c) (Riipinen et al., 2011), iv) phage 0235 (group d) (Munsch-Alatossava and Alatossava, 2013) and v) phage Ldl1 (group e) (Casey et al., 2015). Among these phages, c5, Ld3, Ld17, Ld25A, LL-Ku and phiLdb have recently been classified as *Cequinquevirus* phages within the family *Siphoviridae* (Walker et al., 2019). Their baseplates show similarity in organization and morphology to those of lactococcal P335 phages, where it has been associated with strong multiple binding to the carbohydrate receptor in the host cell wall (Casey et al., 2014). In addition, two group b phages (i.e. Ld17 and Ld25A) possess a genomic region comprised of two putative ORFs, which resembles the KIS element (LL-K insertion sequence) of *Lb. delbrueckii* subsp. *lactis* phage LL-K (Forsman and Alatossava, 1994). It is suggested that at least one of the genes in the

78 KIS element is a structural gene specifying a collar structure that may also play a role in host range
79 specificity (Casey et al., 2014).

80 In the present study, we report on the characterization of the *Lb. delbrueckii* subsp.
81 *bulgaricus* phage PMBT4 that was isolated from Nigerian fermented *nono* with respect to its
82 morphology, host range and genome sequence.

83

84 **2. Materials and Methods**

85 **2.1 *Nono* sampling and phage isolation**

86 Twenty-six *nono* samples (fermented for 24 – 72 h) were collected from various local markets in four
87 local Nigerian states. From all samples (with high acidity, i.e. pH < 4.6), 3-5 ml were centrifuged at
88 17,500 x*g* (20-35 min, 4°C) and the supernatants were filtered through a 0.45 µm Filtropur S filter
89 (Sarstedt, Germany). The double-layer plaque assay was used for phage screening, plaque
90 purification, phage titer determination, and host range analysis. Ten microliters of each filtrate were
91 spotted on MRS soft (top) agar (7.5 g l^{-1}) (de Man, Rogosa and Sharpe (De Man et al., 1960) VWR,
92 Darmstadt, Germany), inoculated with ca. 1 x 10^7 colony-forming units (cfu) ml^{-1} of a *Lb. delbrueckii*
93 subsp. *bulgaricus* strain isolated from *nono*. MRS agar was used as bottom agar. To both - MRS soft
94 agar and MRS agar - 20 mM $CaCl_2$ and 1% glycine were added. For determination of phage titers, the
95 filtered phage sample was diluted in a ten-fold dilution series and 100 µl of appropriate dilutions
96 were added to 0.3 ml of a *Lb. delbrueckii* subsp. *bulgaricus* 2-h culture (10% inoculation; final $OD_{620\,nm}$
97 ca. 0.5) together with 100 µl of 100 mM $CaCl_2$. After 10-min adsorption time at room temperature,
98 the suspension was then added to 3 ml of 50°C warm MRS soft agar (20 mM $CaCl_2$, 1% glycine),
99 mixed and then poured on an MRS agar plate with 20 mM CaCl2, 1% glycine. Plates were incubated
100 anaerobically at 37°C for 18 h (Anaerocult jars [Merck, Darmstadt, Germany]) and AnaeroGen 3.5 L
101 sachets [Thermo Scientific, Schwerte, Germany]). Bacteriophage titers were determined as plaque-

forming units (pfu) ml^{-1} in duplicate. Finally, phage PMBT4 was purified and concentrated by CsCl gradient centrifugation as described in detail elsewhere (Sambrook and Russell, 2001).

The bacteriophage lytic activity spectrum was tested using a panel of nine *Lb. delbrueckii* subsp. *bulgaricus* strains isolated from *nono* (results not shown) and various strains from culture collections (Table 1). These strains were routinely propagated in MRS broth at 37°C. Bacteriophages could only be isolated from a 48-h fermented *nono* sample ("*sallah*") from the Kano State in Nigeria.

2.2 Transmission electron microscopy (TEM) analysis

For TEM analysis, PMBT4 purified by CsCl gradient ultracentrifugation was dialysed against SM buffer (20 mM Tris-HCl pH 7.2, 10 mM NaCl, 20 mM $MgSO_4.7H_2O$). Negative staining was performed with 2% (w/v) uranyl acetate on ultra-thin carbon films as described previously in detail (Casey et al., 2014). Transmission electron microscopy was performed at an acceleration voltage of 80 kV (Tecnai 10, FEI Thermo Fisher Scientific, The Netherlands), and micrographs were acquired with a MegaView G2 CCD camera (Emsis, Muenster, Germany).

2.3 Phage genome sequencing and analysis

For phage PMBT4 DNA isolation, a modified protocol based on the peqGOLD Bacterial DNA Mini Kit (VWR) was used. Briefly, 1.5 ml of phage lysate (10^{11} pfu ml^{-1}) were mixed with each of 1 µl of RNAse-free DNAse (10 mg ml^{-1}) and RNAse (10 mg ml^{-1}) and incubated for 2 h at 37°C. The samples were filtered through a 0.45 µm pore-size filter and centrifuged for 2 h (15,500 x*g*, 4°C). After discarding the supernatant, the phage pellet was resuspended in 400 µl DNA Lysis Buffer T from the DNA isolation kit. Subsequently, 2 µl Proteinase K (20 mg ml^{-1}) were added and the mixture was incubated for 3 h at 37°C. The sample was then mixed with 200 µl DNA Binding Buffer by inverting gently, and then loaded onto a kit column. After centrifugation, the column was washed twice and dried according to the manufacturer's protocol. For the elution step, 50 µl Elution Buffer heated to 70°C were added and incubated for 10 min at room temperature on the column.

For genome sequencing, the Nextera XT DNA Library Preparation Kit and the MiSeq Reagent Nano Kit v2 were used according to the manufacturer's instructions (Illumina, Munich, Germany) and sequencing was performed on a MiSeq (Illumina) sequencer. Assembly of generated reads to contigs was performed using Geneious 9.1.2 (Kearse et al., 2012) and SPAdes 3.11.0 (Bankevich et al., 2012) . The subsequent annotation was performed with RAST (Aziz et al., 2008).

The proteomes of PMBT4 and related *Lb. delbrueckii* subsp. *bulgaricus* phages were compared using all-against-all, bi-directional BLAST alignments (Altschul et al., 1990). An alignment cut-off E-value of 0.0001, and a similarity cut-off level of at least 30% amino acid identity across 80% of the sequence length were applied. Based on this analysis, the closest relatives of PMBT4 were identified. The proposed functional annotations were further investigated by performing structural homology searches via HHpred, TMHMM v2.0 (http://www.cbs.dtu.dk/services/TMHMM/) and Pfam.

2.4 PCR-based analysis for the presence of a KIS element in PMBT4

To analyze whether the PMBT4 genome harbors a KIS element similar to those located in the genomes of *Lb. delbrueckii* subsp. *lactis* phages LL-K (acc. no. AY739900.2 (Forsman and Alatossava, 1994; Riipinen et al., 2011)), Ld17 (acc. no. NC_025420) and Ld25A (acc. no. NC_025415), two PCR reactions with primers that bind within the putative genes for a glycerophosphoryl diester phosphodiesterase (ORF16) and an antireceptor (ORF17) in the phage PMBT4 genome, which would flank a putative KIS element if present, were used (Table 2). First, phage DNA was isolated from 1.5 ml of a high titer lysate (ca. 10^{11} pfu ml^{-1}) as described above. For amplification of genes, the DreamTaq Green PCR Mastermix (Thermo Scientific) was used according to the manufacturer's protocol. One µl (~30 ng) of the kit-purified phage DNA was mixed with primers, mastermix and ddH_2O and PCR-amplified using the following steps: 1) 95°C for 3 min, 2) 95°C for 30 s, 3) 54°C for 30 s, 4) 72°C for 1 min and 5) 72°C for 10 min. Steps 2) to 4) were repeated 35 times.

3. Results and Discussion

3.1. Phage PMBT4 morphology and host range

Phage PMBT4 was assigned by transmission electron microscopy (Fig. 1) to the *Siphoviridae* family with a long but thin, non-contractile 167 nm tail and an isometric head (diameter: 58 nm) (Table 3). This phage shows a unique morphology as it has an unusually large neck passage (collar) structure (width: 24 nm, Table 3 & Fig. 1), which appears larger than those of other *Lb. delbrueckii* phages (Casey et al., 2014). However, this collar was only present on a minority of the phage particles observed under the electron microscope (estimated at 1-5%). The majority of the phages had clearly lost this structural decoration. In addition, phage PMBT4 possessed a short tail fiber (length: 16 nm), protruding under the large baseplate (height: 15 nm; Fig. 1 & Table 3). Six globular appendages were visible on micrographs revealing a bottom-view on the baseplate complex (Fig. 1). With this morphology, phage PMBT4 particles lacking a collar resembled phage *Lb. delbrueckii* Ld3, while those with a collar were similar to the *Lb. delbrueckii* phages Ld17 and Ld25A (Casey et al., 2014). In contrast to phages Ld3 and Ld17 (Casey et al., 2014), which could only infect one or three of the *Lb. delbrueckii* subsp. *bulgaricus* strains tested in this study (Table 1), phage PMBT4 exhibited a relatively wide host range, as it infected nine out of the 21 (i.e. 43%) *Lb. delbrueckii* strains.

3.2. Phage PMBT4 genome analysis

A total of 81,162 paired-end reads (2x251 bp) were generated by sequencing with MiSeq, from which 80,932 reads were *de novo* assembled into a single contig with a total length of 31,399 bp. On average, the assembled genome showed more than 500-fold coverage. Annotation with RAST resulted in the identification of 50 coding sequences (CDS). Phages that infect strains of the *two Lb. delbrueckii* subspecies *bulgaricus* and *lactis*, respectively, are currently classified into five distinct groups (i.e., groups a, b, c, d and e) based on DNA homology (Casey et al., 2014). Phages LL-Ku, c5 and Ld3, Ld17 and Ld25A are group b phages based on their sequence homology, and these were isolated from dairy plants in Finland and a yoghurt production facility in France (Accolas and

176 Spillmann, 1979; Alatossava and Pyhtila, 1980) and more recently from whey samples from yoghurt
177 production facilities in Jordan and Turkey and from Gorgonzola cheese production in Italy (Casey et
178 al., 2014), respectively (Table 4).

179 Phage PMBT4 (isolated from Nigerian *nono* in this study) showed high genome sequence
180 homology, as well as similar numbers and organization of genes with group b *Lb. delbrueckii* phages
181 (Fig. 2). High genome homologies to phage c5 (96% identity/ 93% coverage), Ld3 (95% identity, 84%
182 coverage), Ld25A (96% identity, 82% coverage) and Ld17 (95% identity and 83% coverage) could be
183 detected, while only slightly lower homologies were observed when phage PMBT4 was compared to
184 phages phiLdb (92% identity, 88% coverage) and LL-Ku (94% identity, 90% coverage).

185 These *Siphoviridae* phages (classified in 2018 as members of the genus *Cequinquevirus* by the
186 International Committee on the Taxonomy of Viruses, ICTV; (Walker et al., 2019)) possess genomes
187 with cohesive ends that vary in size from 29 to ca. 34 kbp. The largest genome among the isolates is
188 that of the virulent phage phiLdb (Wang et al., 2010) with a size of 33,996 bp. The mol% GC values
189 were very similar and ranged from 41.5 to 42.2 % (Table 4). Phage PMBT4 genome displayed a typical
190 gene organization, with genes associated with morphogenesis, replication and lysis being organized
191 within modules. The morphogenesis module starting from the portal protein-encoding gene (ORF3)
192 to the tail component protein-encoding genes (ORF14) is well conserved among the type b *Lb.*
193 *delbrueckii* phages (Casey et al., 2014) with many of the predicted proteins sharing >95% amino acid
194 identity (Fig. 2).

195 ORF3 is predicted to encode the portal protein; ORF4 a capsid maturation function and ORF5
196 is predicted to encode the major capsid protein (Fig. 2, Suppl. Table 1) with 100 % structural
197 relatedness to that of the coliphage HK97 (PDB No. 3QPR_D). ORFs 6-9 encode small proteins and
198 based on their genomic location and structural relatedness (in the case of ORF7 and 8), these
199 proteins are predicted to encode head-tail joining functions. ORF10 is predicted to encode the tail
200 terminator protein based on structural homology searches (99.3 % probability; PDB NO. 6TE9_F).

ORF11 encodes the major tail protein (96.6 % probability; PDB No. 6XGRJ). ORF12 is of unknown function; however, based on its genomic location, it is likely a chaperone for the tail tape measure protein. ORF13 of PMBT4 possesses two predicted transmembrane domains based on transmembrane modelling predictions and bears structural similarity to the tail tape measure protein of the *Staphylococcus aureus* phage 80 (99.5 % probability; PDB No. 6V8I_BF) (Fig. 3). We propose that ORF14 of PMBT4 encodes the distal tail (Dit) protein based on structural homology searches. This protein is predicted to have 100 % structural similarity to the Dit protein of the *Bacillus subtilis* phage SPP1. It is a small protein comprising 234 aa (Fig. 3) and does not possess any identifiable carbohydrate binding domains and is, therefore, considered a "classical" Dit. Downstream of *dit* ORF14, is a gene encoding a protein with unknown function (ORF15). This protein bears structural similarity to a baseplate protein of the lactococcal P335 phage, TP901-1 (98.1% probability) and is, therefore likely to form part of the distal tail structure of PMBT4. ORF16 is predicted to encode a glycerophosphoryl diester phosphodiesterase that is believed to function as the tail associated lysin (Tal), while ORF17 is predicted to encode the putative antireceptor. The glycerophosphoryl diester phosphodiesterase (GDPD) was reported to be a structural component of the baseplate from phage Ld17 and possesses a domain with structural similarity to GDPDs encoded by multiple bacteria (100% probability) (Cornelissen et al., 2016). The putative antireceptor genes share 70-89% sequence identity between the group b phages PMBT4, Ld3, Ld25A and Ld17 (Fig 2). This protein exhibits two identifiable domains with similarity to tail tip proteins: at the amino terminal end, 324 aa bear structural relatedness to a protein within the *Staphylococcus* phage 80 tail tip complex (99.9% probability; PMD No. 6V8I_AE) while the C-terminus contains a region with structural relatedness to the coliphage T4 baseplate protein Gp10 (99.1% probability; PMD No. 2FKK_A) (Fig. 3).The carboxy-termini of these antireceptors exhibit most variability (consistent with previous studies) (Fig. 4). The repeat region starting at position 441 of the antireceptor protein was previously noted to be present in phage Ld17 while absent in phage Ld25A (and PMBT4, see Figure 4), and was suggested to possibly play a role in host recognition (Casey et al. 2014). This, or other differences in specific amino acid residues within the binding domain (which has so far not been elucidated) may explain the unique

host range of the phages (Table 1) and the apparent broader host ranges of phages PMBT4 and Ld25 (Casey et al. 2014). Alternatively, host-factors such as CRISPR-Cas spacers that may correspond to the genomes of phages Ld3 and Ld17 with the narrower host range, or restriction modification systems which may recognize specific sequences that are less abundant in phages PMBT4 or phage Ld25A and thus allow a broader host range, may be responsible for the differences in host range observed.

3.3. Phage PMBT4 does not encode a KIS-element.

The genomes of phages Ld25A and Ld17 have two additional genes located between the putative glycerophosphoryl diester phosphodiesterase and antireceptor genes: a gene encoding a collagen repeat protein and a putative adsorption protein gene. These genes are absent in the other group b *Lb. delbrueckii* phages, including phage PMBT4, and were previously described in *Lb. delbrueckii* subsp. *lactis* phage LL-K to encode a KIS element (LL-K insertion sequence) (Forsman and Alatossava, 1994), which is believed to encode a neck passage structure (collar) and a putative adsorption protein that might also be involved in host range determination (Casey et al., 2014). Casey et al. (2014) speculated that the presence of the two KIS element genes in some bacteriophages may be the result of acquisition events from other *Lb. delbrueckii* phages, or alternatively that they were deleted from those phages who are missing these genes. Interestingly, phage PMBT4 seems to have a significantly broader host range when tested for lytic activity against a variety of *Lb. delbrueckii* strains compared to the phages Ld17 and Ld3 (Table 1), even though it lacks the KIS element genes. However, our electron microscopic study of PMBT4 phage particles revealed a low amount (ca. 1-5%) of phages which contained a collar structure below the head. This observation suggests that phage PMBT4 possesses a KIS like element. Several attempts to isolate phage types equipped with such a collar were not successful, when 50 PMBT4 phage derivatives were isolated from single plaques and analysed by electron microscopy, i.e. none of the phages thus assessed showed the presence of a collar (data not shown). Consequently, we searched for unassembled reads which contain the KIS element genes in the raw high-throughput sequence data. Therefore, the total raw reads of phage PMBT4 were mapped directly to the KIS element regions of phages Ld25A and Ld17, but no sequence

match was obtained (data not shown). To confirm the absence of the KIS element in the phage PMBT4 genome, two PCR assays with phage PMBT4 specific primers (Table 2), that should theoretically flank the KIS element genomic region between ORF16 and ORF17, were performed. In detail, the following combinations were used i) primer 99fw and 100rev, which would yield a 1,753-bp PCR product in the presence of a KIS element but ii) only a 69-bp product in case of its absence, iii) primer 99fw and 101rev, which would result in a 1,902-bp PCR product should a KIS element be absent, but iv) if present, the PCR product would be 3,586 bp in size. The result of the PCR (Fig. 4) showed that no KIS element could be found in the phage PMBT4 genome, as only the combination of primer 99fw and 101rev resulted in a corresponding PCR product, representing the 1,902-bp DNA region. Furthermore, as expected, the combination of primers 99fw and 100rev resulted in a small PCR product representing the 69-bp product. Hypothetically, the presence of a defective prophage on the genome of the host strain might be able to supply the collar protein to PMBT4 *in trans*. However, when checking the chromosomal DNA of the sequenced host strain *Lb. delbrueckii* subsp. *bulgaricus* Nono-21:328M (Cho et al., 2020), no collagen repeat-containing protein (CRP) gene or adsorption protein (AdP) gene, which make up the mobile genetic element termed the KIS element, could be detected on the two incomplete prophages of 10.2 kbp and 32.2 kbp identified on the chromosome, respectively (results not shown). To conclude, our results suggest that the phage PMBT4 collar structure observed for some PMBT4 phage particles under the electron microscope did not derive from a KIS like element. The alignment of genome sequences in Fig. 2 clearly shows that there is higher diversity in genes on the right half of the genome (genes located downstream of the antireceptor protein gene) when compared to genes on the left half of the genome. Also, there were numerous genes on the right arm of *L. delbrueckii* phage genomes which encode hypothetical proteins with unknown function. Thus, it may be conceivable that there are genes in this region which may encode the observed collar. However, the Blast search results of the genes (Suppl. Table 1) and those reported in the ORF table for phage Ld25A by Casey et al. (2014) do not allow the determination of one or more candidate genes responsible for encoding such proteins.

4. Conclusions

A novel *Siphoviridae* bacteriophage infecting a relatively wide range of *Lb. delbrueckii* subsp. *bulgaricus* strains was isolated from a Nigerian fermented milk product called *nono*. Based on genome sequencing, the phage could be assigned to the group b *Lb. delbrueckii* phages and the genome size of 31,399 bp and the 41.6 mol% GC content compared well to the characteristics of these group b bacteriophages. The close relationship to phages c5, Ld3, Ld17, Ld25A and LL-Ku based on genome and morphology similarity clearly revealed that phage PMBT4 belongs to the genus *Cequinquevirus*. The genomic analysis further revealed the presence of two genes (ORF16 and 17), which encode a glycerophosphoryl diester phosphodiesterase and a putative antireceptor, respectively. These genes were also found to be present in genomes of two other members of the group b *Lb. delbrueckii* phages, i.e. Ld17 and Ld25A, where they encompass the so-called KIS element. Importantly, electron microscopic studies of the phages that possess KIS elements in their genomes, showed that the phages produce a neck passage structure that was hypothesized to play a role in a relatively broader host range (Casey et al., 2014). To test this hypothesis that the KIS element genes encode the neck passage structure also in phage PMBT4, the complete set of raw sequence data from the genomic study of PMBT4 were investigated in order to determine unassembled contigs in which the KIS element genes were present. Furthermore, PCR primers flanking the KIS element were used to identify this DNA element in the genomic DNA pool isolated from the phage lysate, which should contain phages with neck passage structure as well as phages without these structures. In this study, the absence or presence of a neck passage structure could not be correlated to either absence or presence of genomic sequence DNA contigs, or by a differential PCR amplification targeting the KIS element genes. Therefore, the question of which open reading frame(s) encode(s) a neck passage structure, and whether the loss of this structure is associated with gene loss, still remains unanswered. Our results suggest that the KIS element genes do not appear to be the genes associated with the neck passage structure observed for phage PMBT4

306 particles under TEM. Protein analyses studies may in future provide a definite answer to the
307 protein(s) and the gene(s) that form the biological basis for this observed structure.

308

309 **Credit Authorship Contribution Statement:** Conceptualization, OF, HN, FO, CF; methodology, FO, EB,
310 EC, SS, JM; software, OF, EB, G-SC, EC, JM; validation, EC, EB, JM, DvS; formal analysis, OF, SS, EB, G-
311 SC,EC, JM; investigation, OF, EB, SS, JM; resources, CF, FO, HN, DvS; data curation, EB, G-SC, EC, SS,
312 JM, DvS, HN, FO, CF; writing: original draft preparation, OF, SS, HN, JM, FO, DVS, CF; writing—review
313 and editing, OF, EB, EC, SS, FO, HN, JM, DvS, CF; visualization, OF, EB, SS, JM, HN, DvS, CF supervision,
314 FO, HN, JM, DvS, CF project administration, FO, CF; funding acquisition, FO, CF. All authors have read
315 and agreed to the published version of the manuscript.

316 **Funding.** This research was funded in part by Alexander von Humboldt Stiftung, grant number 3.4-
317 IPDEU/1154993.

318 **Acknowledgments.** The authors wish to thank Angela Back and Inka Lammertz for excellent technical
319 assistance.

320 **Conflicts of Interest.** The authors declare no conflict of interest.

321 **Data availability:** The genome of phage PMBT4 has been deposited in GenBank under the accession
322 no. MG913376. The version described in this paper is the version MG913376.1.

323 **Figure Legends**

324 **Fig. 1.** Transmission electron micrographs of the *Lb. delbrueckii* subsp. *bulgaricus* phage PMBT4. The
325 arrows indicate the central tail fiber protruding beneath the baseplate complex (1a-b, d-e). The
326 majority of phage particles did not possess collar (neck passage) structures (1a-c). Unusually large
327 collar structures were detected on a low number (ca. 1-5%) of PMBT4 phage particles (see triangles
328 in 1d-e). In 1c, the six-fold symmetry of the globular subunits of a baseplate complex is visible on a
329 PMBT4 phage particle with a folded baseplate complex.

330 **Fig. 2**. Genomic comparison of *Lb. delbrueckii* phage PMBT4 with other group b members. Predicted
331 functions associated with the gene products are coded according to the colored boxes on the right.
332 The leftward region of the genome is associated with the structural components of the phage tail
333 and capsid (purple, blue and yellow arrows). The rightward end of the genome is associated with
334 replication (green) and lysis (red) functions. The sequence similarity between gene products of the
335 phages is indicated by shaded grey/black boxes with the percentage of identity (% aa) indicated by
336 distinct grey-scale colors as indicated in the figure.

337 **Fig. 3.** HHPred outputs for the predicted tail structural proteins. Based on structural homology, it is
338 possible to identify the (A) tail tape measure protein encoded by ORF13; (B) a “classical” or non-
339 evolved distal tail protein encoded by ORF14; (C) a tail-associated lysin (encoded by ORF 16) with an
340 associated glycerophosphate phosphodiester phosphodiesterase domain and; (D) the putative
341 receptor binding/antireceptor protein (encoded by ORF17) with domains associated with tail tip and
342 baseplate functions at the amino- and carboxy-termini, respectively.

343 **Fig. 4** Exclusion of a KIS element in the PMBT4 genome by PCR. PCR assays were performed with
344 primers 99for and 100rev and 99for and 101rev, respectively, and 5 µl of each sample were
345 separated with an 0.8% agarose gel for 55 min at 80 V. 1: Gene Ruler 1 kb plus (Thermo Scientific); 2:
346 Phage PMBT4 DNA with primer 99fw and 100rev; 3: Phage PMBT4 DNA (duplicate) with primer 99fw
347 and 100rev; 4: negative control primer 99fw and 100rev; 5: Phage PMBT4 DNA with primer 99fw and

101rev; 6: Phage PMBT4 DNA (duplicate) with primer 99fw and 101rev; 7: negative control with primer 99fw and 101rev; 8: Gene Ruler Mix (Thermo Scientific).

Figures

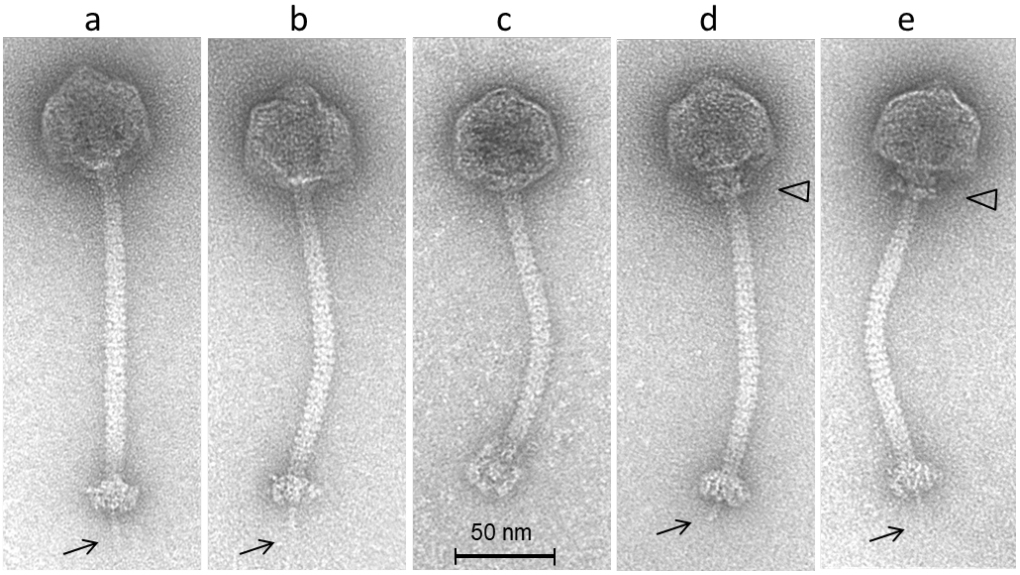


Fig. 1

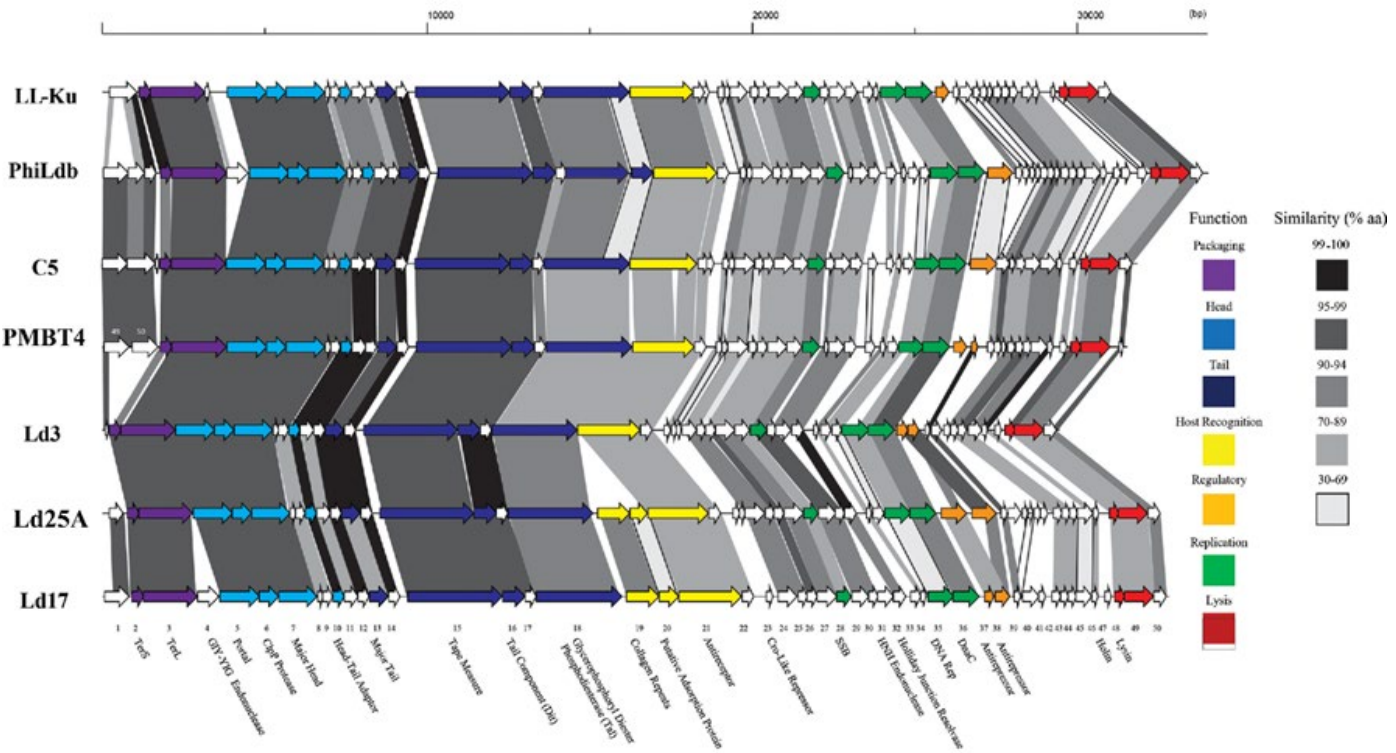
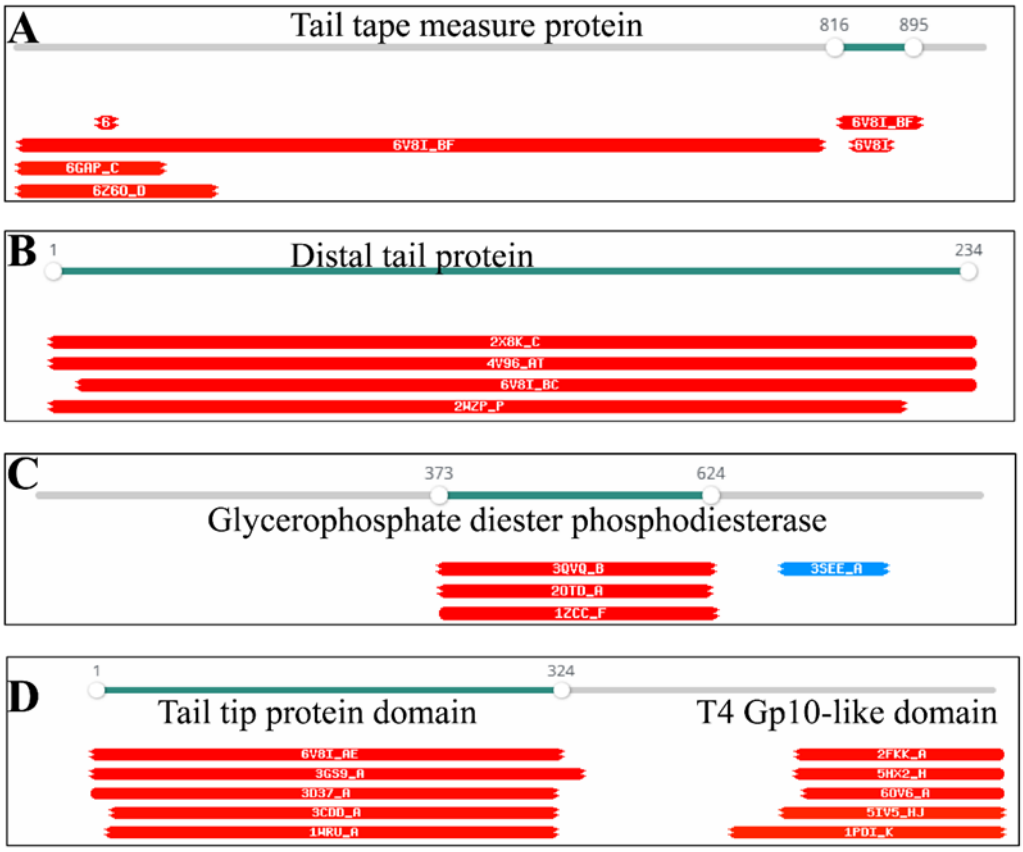
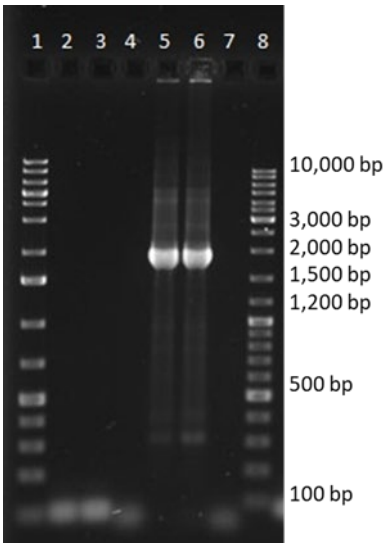


Fig. 2



357

358 Fig. 3



359

360 Fig. 4

361

362 **Tables**

363
364
365 **Table 1.** Host ranges of *Lb. delbrueckii* subsp. *bulgaricus* phages PMBT4 (this study), Ld3 and Ld17
366 (Casey et al., 2014).
367
368

Lb. delbrueckii subsp. *bulgaricus* strain	Phage PMBT4	Phage Ld3 [a]	Phage Ld17 [a]	Strain source / reference
Nono-21:328M	● [b]	-	-	*Nono* isolate (this study)
MBT 92063 3038a	-	-	-	Institute's strain collection [c]
MBT 92063-PM11	●	-	-	Institute's strain collection
MBT 92067	-	-	-	Institute's strain collection
MBT 92068	-	-	-	Institute's strain collection
MBT 92197 Vitus	●	-	●	Institute's strain collection
MBT 92235	●	-	-	Institute's strain collection
DSMZ 20081	●	-	-	DSMZ Braunschweig (DE)
MBT 92375	-	-	-	Institute's strain collection
MBT 92376	-	-	-	Institute's strain collection
MBT 92378	-	-	-	Institute's strain collection
MBT 92059	-	-	-	Institute's strain collection
Y532-2Lb	-	-	-	Institute's strain collection
Y532-HLB-1M	-	-	-	Institute's strain collection
Y532-HLA-2M	-	-	-	Institute's strain collection
CHCC3984	●	-	●	Chr. Hansen strain collection
CHCC3606	●	-	-	Chr. Hansen strain collection
Jo1-1	●	-	-	Institute's strain collection
Jo231-1	-	-	-	Institute's strain collection
Ldb3	-	● [b]	-	(Casey et al., 2014)
Ldb17	●	-	● [b]	(Casey et al., 2014)

369
370 ● Lysis [a]Casey et al., (2014).
371 - no lysis
372 [b]Strain used as host strain for phage propagation
373 [c]from Max Rubner-Institut
374
375

376 **Table 2.** Primers used in this study.
377

Primer	binding site	Sequence 5`→3`	T_M
99fw	3`end glycerophosphoryl diester phosphodiesterase gene (ORF16)	GCAATCTTCCTCTAGCGG	58.8
100rev	5`end antireceptor gene (ORF17)	CGGTAATCCCGAAAACTCGT	57.3
101rev	3`end antireceptor gene (ORF17)	CCGCTAAATAAGGTGGCATG	57.3

378

379

380

381 **Table 3.** Dimensions of phage PMBT4.

Structure measured	Phage dimensions (nm)	Phage particles measured
Head diameter	57.6 ± 1.7	14
Collar[a] height	8.5 ± 0.8	16
Collar[a] width	24.1 ± 1.7	16
Tail length	166.7 ± 2.8	14
Tail width	11.6 ± 0.4	14
Baseplate height	14.5 ± 0.8	22
Baseplate width	24.3 ± 1.6	22
Baseplate globular structures diameter	9.4 ± 0.7	18
Baseplate distal fiber length	15.7 ± 1.2	14

382 [a] When present

383

384

385 **Table 4.** Genomic features of group b *Lb. delbrueckii* phages Ld3, Ld17, Ld25A, PMBT4, c5, phiLdb and
386 LL-Ku.

Characteristic	Ld3	Ld17	Ld25A	PMBT4	c5	phiLdb	LL-Ku
Length (bp)	29,616	32,975	32,799	31,399	31,841	33,996	31,080
No. of ORFs	49	50	51	50	50	59	51
GC content (mol%)	42.2	41.97	42.2	41.6	41.9	42.0	41.5
Origin product	Jordan yoghurt	Italy gorgonzola	Turkey yoghurt	Nigeria *nono*	France yoghurt	China yoghurt	Finland cheese whey

387

388

References

Accolas, J., Spillmann, H., 1979. Morphology of bacteriophages of *Lactobacillus bulgaricus*, *L. lactis* and *L. helveticus* J. Appl. Microbiol. 47, 309-319.

Alatossava, T., Pyhtila, M.J., 1980. Characterization of a new *Lactobacillus lactis* bacteriophage. IRCS Medical Science, Library Compendium 8(5), 297-298.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215(3), 403-410.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9, 75.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19(5), 455-477.

Banwo, K., Sanni, A., Tan, H., Tian, Y., 2012. Phenotypic and genotypic characterization of lactic acid bacteria isolated from some Nigerian traditional fermented foods. Food Biotechnol. 26, 124-142.

Casey, E., Mahony, J., Neve, H., Noben, J.P., Dal Bello, F., van Sinderen, D., 2015. Novel phage group infecting *Lactobacillus delbrueckii* subsp. *lactis*, as revealed by genomic and proteomic analysis of bacteriophage Ldl1. Appl. Environ. Microbiol. 81(4), 1319-1326.

Casey, E., Mahony, J., O'Connell-Motherway, M., Bottacini, F., Cornelissen, A., Neve, H., Heller, K.J., Noben, J.P., Dal Bello, F., van Sinderen, D., 2014. Molecular characterization of three *Lactobacillus delbrueckii* subsp. *bulgaricus* phages. Appl. Environ. Microbiol. 80(18), 5623-5635.

Cho, G.S., Fagbemigun, O., Brinks, E., Adewumi, G.A., Oguntoyinbo, F.A., Franz, C.M.A.P., 2020. Draft genome sequences of *Lactobacillus helveticus*, *Lactobacillus fermentum*, and *Lactobacillus delbrueckii* strains from African fermented nono. Microbiol. Resour. Announc. 9(1) e01342-19.

Cornelissen, A., Sadovskaya, I., Vinogradov, E., Blangy, S., Spinelli, S., Casey, E., Mahony, J., Noben, J.P., Dal Bello, F., Cambillau, C., van Sinderen, D., 2016. The baseplate of *Lactobacillus delbrueckii* bacteriophage Ld17 harbors a glycerophosphodiesterase. J. Biol. Chem. 291(32), 16816-16827.

De Man, J.C., Rogosa, M., Sharpe, M.E., 1960. A medium for the cultivation of lactobacilli. J. Appl. Bacteriol. 23(1), 130-135.

Fagbemigun, O., Cho, G.S., Rösch, N., Brinks, E., Schrader, K., Bockelmann, W., Oguntoyinbo, F.A., Franz, C.M.A.P., 2021. Isolation and characterization of potential starter cultures from the Nigerian fermented milk product *nono*. Microorganisms 9 (640).

Forsman, P., Alatossava, T., 1994. Repeated sequences and the sites of genome rearrangements in bacteriophages of *Lactobacillus delbrueckii* subsp. *lactis*. Arch. Virol. 137(1-2), 43-54.

Guo, T., Zhang, C., Xin, Y., Xin, M., Kong, J., 2016. A novel chimeric prophage vB_LdeS-phiJB from commercial *Lactobacillus delbrueckii* subsp. *bulgaricus*. J. Ind. Microbiol. Biotechnol. 43(5), 681-689.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12), 1647-1649.

Mikkonen, M., Raisanen, L., Alatossava, T., 1996. The early gene region completes the nucleotide sequence of *Lactobacillus delbrueckii* subsp. *lactis* phage LL-H. Gene 175(1-2), 49-57.

Munsch-Alatossava, P., Alatossava, T., 2013. The extracellular phage-host interactions involved in the bacteriophage LL-H infection of *Lactobacillus delbrueckii* ssp. *lactis* ATCC 15808. Front. Microbiol. 4, 408.

Ogbonna, I., 2011. Microbiological analyses and safety evaluation of *nono*: A fermented milk product consumed in most parts of Northern Nigeria. Int. J. Dairy Sci. 6, 181-189.

Okagbue, R., Bankole, M., 1992. Use of starter cultures containing *Streptococcus diacetilactis, Lactobacillus brevis* and *Saccharomyces cerevisiae* for fermenting milk for production of Nigerian *nono*. World J. Microbiol. Biotechnol. 8, 251-253.

Riipinen, K.A., Forsman, P., Alatossava, T., 2011. The genomes and comparative genomics of *Lactobacillus delbrueckii* phages. Arch. Virol. 156(7), 1217-1233.

Sambrook, J., Russell., D.W., 2001. Molecular cloning: A laboratory manual. 3rd ed. Cold Springer Harbor Laboratory Press, Cold Springer Harbor, N.Y.

Walker, P.J., Siddell, S.G., Lefkowitz, E.J., Mushegian, A.R., Dempsey, D.M., Dutilh, B.E., Harrach, B., Harrison, R.L., Hendrickson, R.C., Junglen, S., Knowles, N.J., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Nibert, M., Rubino, L., Sabanadzovic, S., Simmonds, P., Varsani, A., Zerbini, F.M., Davison, A.J., 2019. Changes to virus taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2019). Arch. Virol. 164(9), 2417-2429.

Wang, S., Kong, J., Gao, C., Guo, T., Liu, X., 2010. Isolation and characterization of a novel virulent phage (phiLdb) of *Lactobacillus delbrueckii*. Int. J. Food. Microbiol. 137(1), 22-27.

Literaturverzeichnis

Casey, Eoghan; Mahony, Jennifer; O'Connell-Motherway, Mary; Bottacini, Francesca; Cornelissen, Anneleen; Neve, Horst et al. (2014): Molecular characterization of three *Lactobacillus delbrueckii* subsp. *bulgaricus* phages. In: *Applied and environmental microbiology* 80 (18), S. 5623–5635. DOI: 10.1128/AEM.01268-14.

Wang, Shaohua; Kong, Jian; Gao, Chen; Guo, Tingting; Liu, Xiaoyong (2010): Isolation and characterization of a novel virulent phage (phiLdb) of *Lactobacillus delbrueckii*. In: *International journal of food microbiology* 137 (1), S. 22–27. DOI: 10.1016/j.ijfoodmicro.2009.10.024.