

Article

A Simple Method for Assessing Diversity and Dynamics of Microbial Community: Comparison of Dairy Phages from Industrial and Spontaneous Fermentation

Agnieszka Olejnik-Schmidt , Bernadeta Pietrzak , Iwona Kawacka , Klaudia Malak, Weronika Wawrzyniak and Marcin Schmidt * 

Department of Food Biotechnology and Microbiology, Poznan University of Life Sciences, Wojska Polskiego 48, 60-627 Poznań, Poland; agnieszka.olejnik-schmidt@up.poznan.pl (A.O.-S.); bernadeta.pietrzak@up.poznan.pl (B.P.); iwona.kawacka@up.poznan.pl (I.K.); 49702@uppoz.onmicrosoft.com (K.M.); weronika.wawrzyniak@up.poznan.pl (W.W.)

* Correspondence: marcin.schmidt@up.poznan.pl; Tel.: +48-61-8466024

Featured Application: Assessing diversity and dynamics of low-complexity microbial communities, e.g., in fermentation processes.

Abstract: Background: The dairy industry heavily relies on fermentation processes driven in high proportion by *Lactococcus lactis*. The fermentation process can be perturbed or even stopped by bacteriophage activity, leading to complete loss of fermentation batch or decreased quality product. The monitoring of the phage diversity and dynamics in the process allows implementing protective measures (e.g., starter rotation) to maintain unperturbed production. Methods: Universal primers were used to amplify sequences of the 936, c2, and P335 *Lactococcus* phage types. The amplicons were sequenced with the Sanger method and obtained degenerate sequences were analyzed using a simple bioinformatic pipeline in the R environment. Results: The most prevalent phage type is 936, followed by P335, whereas the c2 type is less frequent. Conclusions: Curd cheeses prepared on non-pasteurized milk based on native milk microbiota had a higher diversity of phages distinct from those found in dairy plants. Sanger sequencing of heterogenous amplicons generated on metagenome DNA can be used to assess low-complexity microbiota diversity.

Keywords: phage; dairy; *Lactococcus*; diversity; community



Citation: Olejnik-Schmidt, A.; Pietrzak, B.; Kawacka, I.; Malak, K.; Wawrzyniak, W.; Schmidt, M. A Simple Method for Assessing Diversity and Dynamics of Microbial Community: Comparison of Dairy Phages from Industrial and Spontaneous Fermentation. *Appl. Sci.* **2021**, *11*, 8915. <https://doi.org/10.3390/app11198915>

Academic Editors: Ester Betoret and Noelia Betoret

Received: 17 August 2021

Accepted: 20 September 2021

Published: 24 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The dairy industry heavily relies on fermentation processes driven in high proportion by *Lactococcus lactis*. The fermentation process can be perturbed or even stopped by bacteriophage activity leading to huge economic losses [1–3]. Among bacteriophages infecting *L. lactis* [4], those belonging to species c2, 936, or P335 are most commonly encountered in dairy plants [5,6]. Most in-house dairy plants' laboratories use traditional, indirect methods for phage detection in their routine analyses. These methods are based on activity tests (assessment of a decrease in acid production or a reduction of an indicator compound). Additionally, standard microbiological methods, i.e., plaque assays or spot tests, are used for direct phage detection [1,7]. Molecular methods (mostly PCR-based) give the advantage of short time from sampling to results, high sensitivity, and independence of sensitive strain [8]. Apart from detection methods, analysis of diversity and dynamics of *lactococcal* bacteriophages is crucial for preventing the spread of infections in dairy environments. Knowledge of phage strains presence in milk or dairy environment may allow to plan starter culture rotation.

Analysis of microbial community composition and dynamics especially in the case of unculturable bacteria or viruses always posed a challenge. Especially, lack of knowledge on

growth requirements or undetermined permissive host is a limiting factor. Today in the era of next-generation sequencing (NGS), the problem can be solved both in the determination of community composition [9] and its functional profile [10]. However, NGS is still an expensive method. Its ability to produce a massive output of data in a relatively short time can sometimes be a disadvantage. The relatively low price per sample can be achieved only by filling the whole throughput (lane, chamber, or chip of the sequencer) with samples. In case of low complexity samples or targeted approaches to use the whole throughput, extensive barcoding is necessary. It is often necessary to collect a large number of samples to cover the whole throughput of a chip to run the sequencing. The NGS technique was already successfully used to study microbial communities in dairy technologies [11,12]. In this respect, we looked for a technique that may substitute NGS in the case of low-complexity communities. Here, we describe a novel approach that was used to compare the diversity of *Lactococcus* phages from dairy industrial and spontaneous fermentation processes. The new method is based on well-established and affordable techniques accompanied by simple bioinformatics analysis to study the diversity of the microbial community.

2. Materials and Methods

2.1. Cheese Samples and DNA Isolation

A series of curd cheese samples derived from Polish cheese production facilities ($n = 25$) made from cow's milk and spontaneous non-pasteurized milk fermentations from in-house cheese production ($n = 14$) made from cow's ($n = 8$), goat's ($n = 5$) or sheep's ($n = 1$) milk were collected in 2017–2018. Metagenomic DNA was isolated from the samples of cheeses using Genomic Mini AX Food Kit (A&A Biotechnology, Gdańsk, Poland), according to the kit procedure, and used as the template for PCR reaction.

2.2. PCR Amplification and Sequencing

A previously described multiplex polymerase chain reaction (PCR) system [13] with later modification [5] was employed to screen for the phages present in the cheeses. In cases in which multiple of the three major groups of *Lactococcus* phages, i.e., the 936, c2, or P335 groups were identified in the same sample, separate PCR reactions were performed with single type-specific primers pair for the phage group (936A: 5'-TCAATGGAAGACCAAGCGGA-3' and 936B: 5'-GTAGGAGACCAACCAAGCC-3', or c2A: 5'-CAGGTGTAAGAGTTCGAGAACT-3' and c2C: 5'-TCAGATAACGCGCCTGAATC-3', or P335A: 5'-GAAGCTAGGCGAATCAGTAA-3' and P335B: 5'-GATTGCCATTGCGCTCTGA-3' [5,13]). The amplicons were purified with Clean-Up Kit or Gel-Out Kit (A&A Biotechnology, Gdańsk, Poland), in case of multiple amplicons, obtained and submitted to Genomed S.A. (Warsaw, Poland) for Sanger sequencing.

2.3. Sequence Analysis

The sequencing trace files (.ab1) were inspected for sequencing quality using Sequence Scanner Software 2 (Applied Biosystems, Warsaw, Poland). The sequences were submitted to BLAST service analysis [14] to recover the best-matched sequences set from the NCBI database. The sequence set obtained from BLAST results was used to build a local database and reference matrix composed of sequence distance measures. The measures were calculated in the local alignment of the ambiguous sequence separately to all sequences from the BLAST result sequences set. A principal component analysis (PCA) plot further simplified the interpretation of the community diversity. The bioinformatics analyses were performed in R [15] using DECIPHER [16] and PCAtools [17] libraries.

3. Results

3.1. Phage Survey

The curd cheese samples derived from Polish dairy facilities ($n = 25$) and spontaneous milk fermentations from in-house cheese production ($n = 14$) from different locations were tested for the presence of *Lactococcus* phages with multiplex PCR [13]. All of the dairy-plant-derived cheeses were tested positive for both P335 and 936 phages, and only 16%

of them contained additional c2 phages (Figure 1, lanes 1–12). All cheeses derived from spontaneous non-pasteurized milk fermentations contained the 936 phage type, 93% contained the P335 phage type, and 14% contained the c2 phage type (Figure 1, lanes 13–26). However, the multiplex PCR reactions made with metagenomic DNA from cheeses prepared with spontaneous fermentations resulted in the generation of multiple non-specific amplicons. Sequence analysis of selected non-specific amplicons (Supplementary Materials: File_S1.fas) indicated their significant homology to *Lactococcus* phages. Non-specific size amplicons generated with primer pair specific to the c2 phage type purified and sequenced ($n = 8$) had sequence of highest homology to *Lactococcus* phage CHPC967 (*Siphoviridae*; 88.01% sequence identity), *Lactococcus* phage CHPC959 (*Siphoviridae*; 98.80% sequence identity), *Lactococcus* phage vB_LacS_15 (*Siphoviridae*; 98.28% and 97.62% sequence identity), *Siphoviridae* sp. ctGas1 (*Siphoviridae*; 90.99%, 90.75%, and two of 91.13% sequence identity); with primer pair specific to the P335 phage type ($n = 2$), *Lactococcus* phage Tuc2009 (*Siphoviridae*; 84.53% sequence identity), *Lactococcus* phage bIL286 (*Siphoviridae*; 82.71% sequence identity); nucleotide BLAST search performed on 11 August 2021. This indicates the possible existence of non-dairy plant environment-adopted phages.

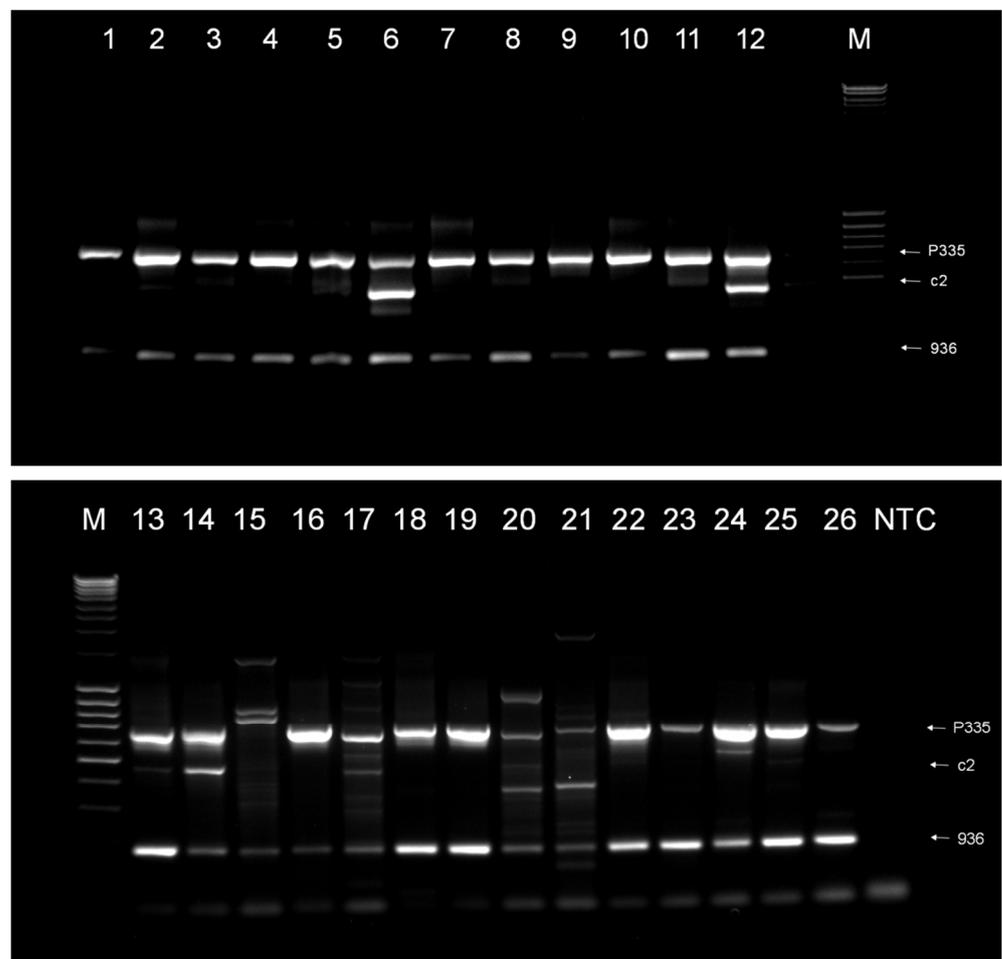


Figure 1. Exemplary multiplex PCR results for detection of lactococcal phages in curd cheeses from dairy plant mass production (lanes 1–12) and spontaneous milk fermentation from in-house cheese production (lanes 13–26). Arrows indicate positions of specific amplicons for P335, c2, and 936 *Lactococcus* phage types (M—molecular weight marker, NTC—non-template control).

3.2. Phage Diversity Analysis between Samples

To compare the diversity of *Lactococcus* phages between samples, we performed PCR amplification in separate PCR reactions with single type-specific primers pair for each of

the phage types. In the case in which several samples of metagenomic DNA were obtained from spontaneous fermentation, multiple amplicons were generated with the use of single pair of type-specific primers. Such amplicons were purified from gel slabs cut out from electrophoretic separation gel before submission to sequencing. Inspection of sequencing trace files indicated ambiguous nucleotide positions (data not shown) that resulted from sequencing of heterogenic amplicons obtained on metagenomic DNA. We performed multiple alignments of the obtained sequences with DECIPHER library [16] and with a substitution matrix accounting for ambiguous nucleotides for all IUPAC nucleic acid codes based on match and mismatch parameters (match = 5, mismatch = -4) and calculated a distance matrix for each sequence set with Jukes–Cantor substitution model used for distance correction [18]. Each element of the distance matrix corresponds to the dissimilarity between two sequences in the sequence set.

Comparison of the phylogenetic trees generated for each of the *Lactococcus* phage types (Figures 2–4) indicated that the P335 type phages from dairy plant cheeses cluster together, whereas the phages from spontaneous fermentations are more divergent.

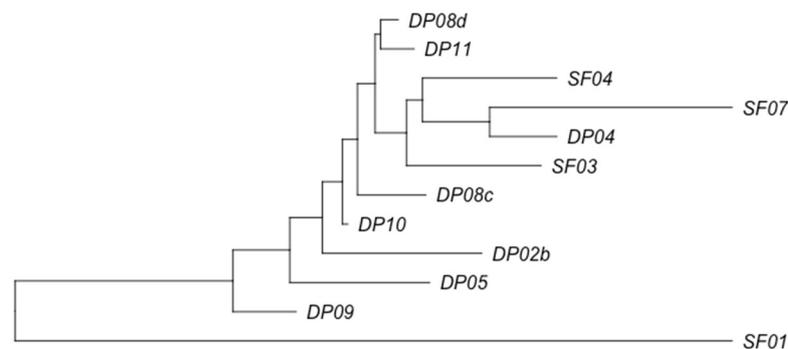


Figure 2. Phylogenetic tree obtained for sequences of amplicons generated with *Lactococcus* 936 phage-type primers with metagenomic DNA from curd cheeses from dairy plants (DP-) and spontaneous in-house fermentations (SF-).

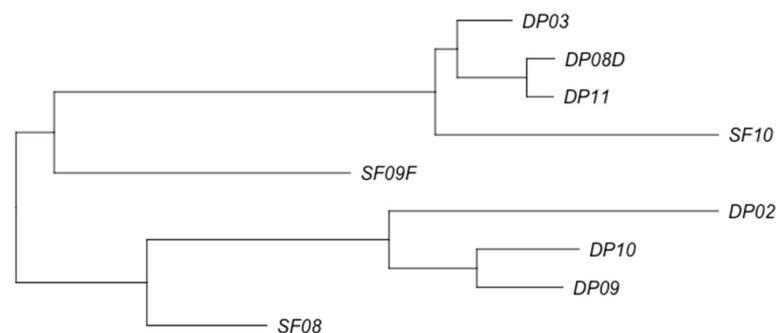


Figure 3. Phylogenetic tree obtained for sequences of amplicons generated with *Lactococcus* c2 phage-type primers with metagenomic DNA from curd cheeses from dairy plants (DP-) and spontaneous in-house fermentations (SF-).

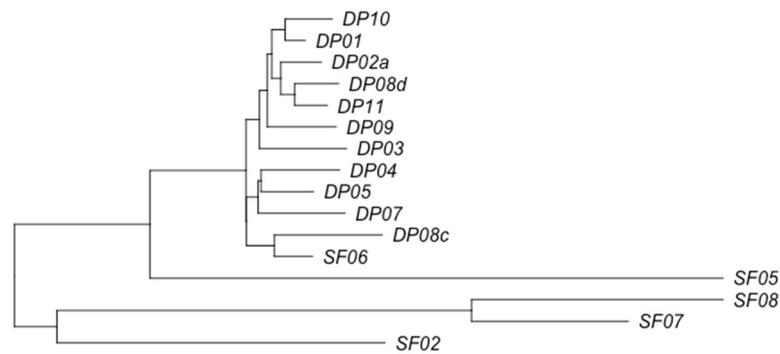


Figure 4. Phylogenetic tree obtained for sequences of amplicons generated with *Lactococcus* P335 phage-type primers with metagenomic DNA from curd cheeses from dairy plants (DP-) and spontaneous in-house fermentations (SF-).

3.3. Global Phage Diversity Analysis

The BLAST search of the obtained phage sequences resulted in a hit list of *Lactococcus* phage sequences of high sequence identity. However, the degenerate nucleotide codes are treated by the BLAST algorithm as mismatches in nucleotide alignment [14]. The hit lists for all of the sequence searches were combined, and duplicate records were removed to build a local database. The local database was used to calculate the reference matrix of sequence distance measures with DECIPHER library [16], with substitution matrix accounting for ambiguous nucleotides for all IUPAC nucleic acid codes, as described previously. This way, we created sample composition-independent measures for global phage diversity analysis. The global distance measures were used for principal component analysis (PCA) to visualize differences and similarities in phage-type composition in analyzed samples. Sequences derived from BLAST search hits used for distance measures cluster tightly together (Figure 5), and metagenome-derived sequences form on the plot distant, more dispersed clusters separate for dairy plant and spontaneous fermentation cheeses with single outliers DP02(c2), SF01(936), and SF06(P335).

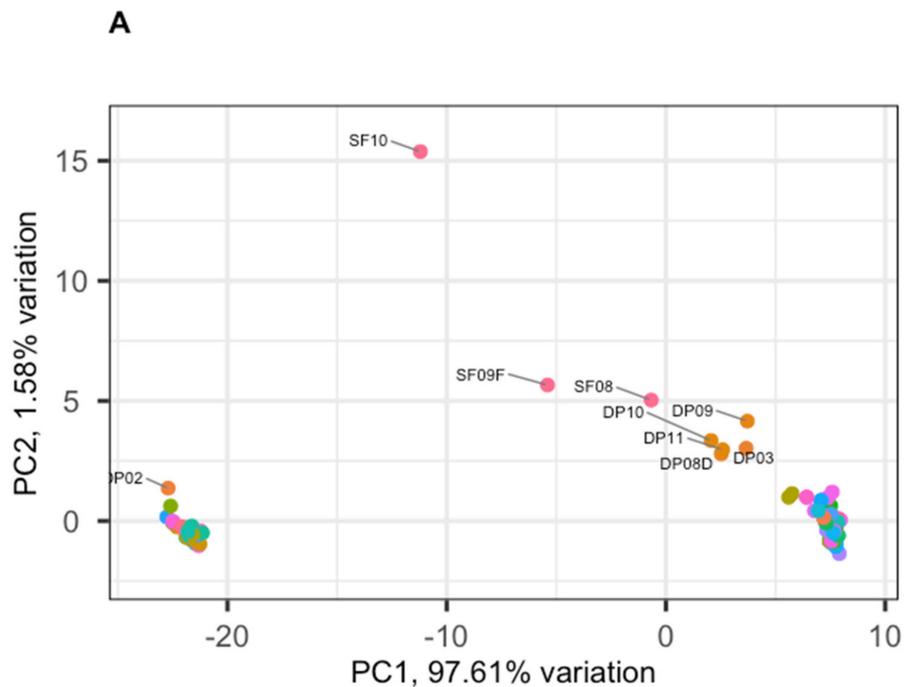


Figure 5. Cont.

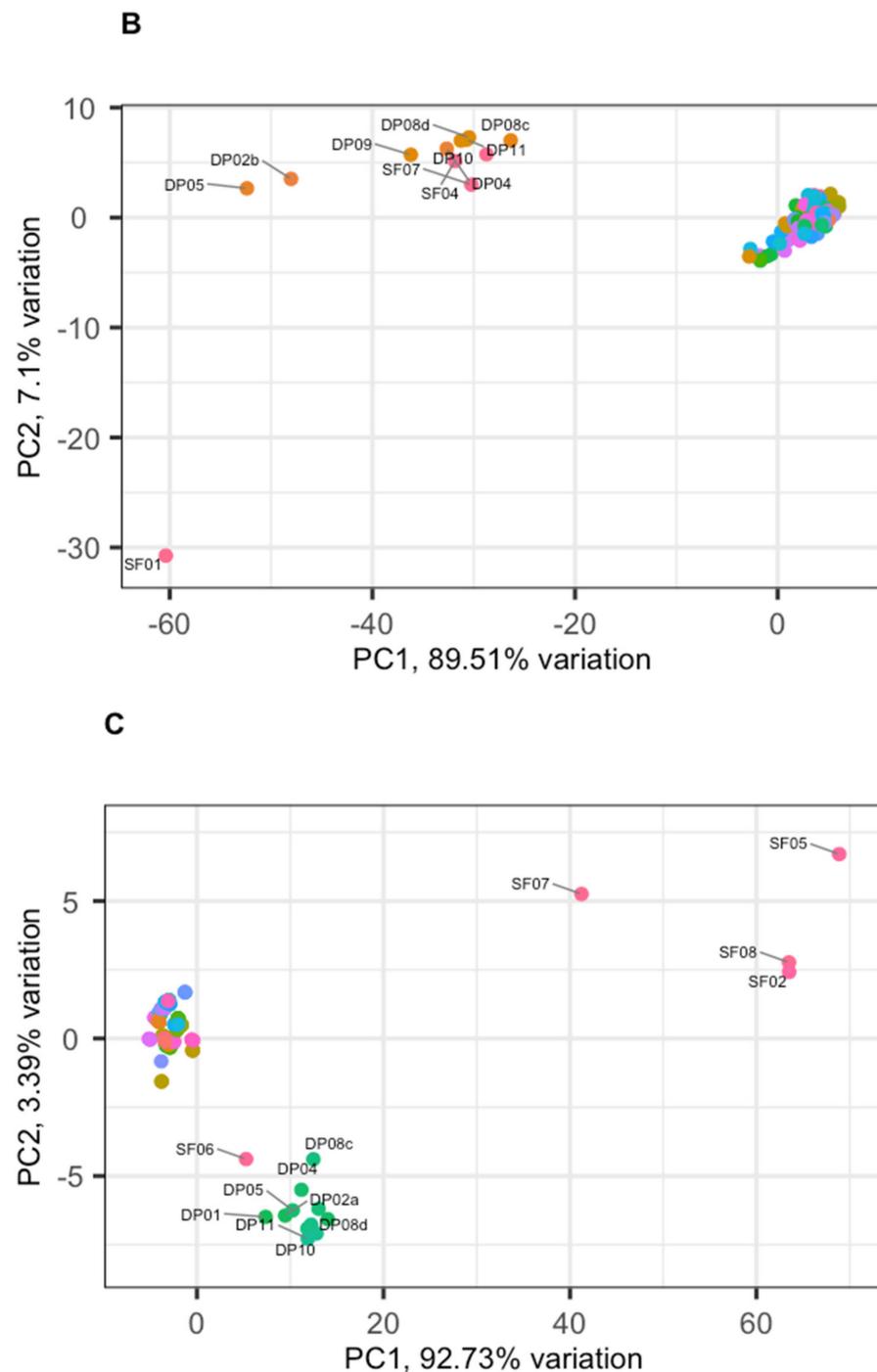


Figure 5. Principal component analysis biplots for sample composition-independent diversity comparison based on sequence distance measures as obtained for *Lactococcus* phage: c2 (A), 936 (B), and P335 (C) types. Labeled points represent phage sequences amplified on metagenomic DNA from dairy plants (DP-) or spontaneous fermentation (SF- produced curd cheeses; unlabeled points represent sequences derived from BLAST search hits.

4. Discussion

Our study indicated a higher prevalence of the *Lactococcus* phage types detected in curd cheeses from Polish dairy plants than previous reports [6,19], which indicates that the problem of phage prevalence in dairy plant environments persists. We also surveyed in-house prepared cheeses from non-pasteurized milk based on spontaneous fermentation resulting from native milk microbiota. Based on our findings, we can assume that the native

milk microbiota contains a much higher diversity of phages and *L. lactis* strains (as sources of prophages). Especially in the case of c2 and P335 phage types, we obtained non-specific size amplicons that were of phage origin. The P335 nonspecific amplicons might be derived from prophages of P335-type or phage DNA remnants in chromosomes of *L. lactis*. The higher diversity of spontaneous fermentation cheeses can be explained by the sensitivity of some phages to thermal treatment [20]. We did not find any correlation between milk type and phage diversity measures.

Here, we described also a simple and easy way to study the diversity and dynamics of a low-complexity microbial community with well-established and affordable techniques and a simple bioinformatics pipeline. Previously reported methods employed laborious and susceptible strains, requiring methods involving phage genomic DNA isolation and sequencing [20], low-resolution sequence analysis with RFLP [6,21], or MLST [22]. Our method involved PCR amplification of metagenomic DNA using universal primers hybridizing to conserved regions spanning variable sequences (e.g., in 16S rDNA, ITS, etc.). The obtained pool of amplicons was sequenced with the Sanger method, resulting in a read-out sequence with ambiguous bases at several positions. Care had to be taken to submit amplicons of the same length for Sanger sequencing to prevent shifts of homologous sequence regions or the possibility of taking into account only selected parts of the homologous regions proximal to the sequencing primer. In case of large differences in amplicon length, electrophoretic separation and gel purification could solve the problem. The resulting sequence corresponds to the consensus sequence, as it would be generated by bioinformatics sequence alignment of several homologous sequences. The sequences from different samples can be aligned, then the distance between them calculated with DECIPHER library [16] using the substitution matrix accounting for ambiguous nucleotides for all the IUPAC nucleic acid codes. Such analysis allows comparing samples composition or dynamics (e.g., time series) in form of phylogenetic tree or PCA plots. To generate the PCA plots on a global scale (sample composition-independent context), a local database of homologous sequences is needed. The local database can be easily prepared by submitting a sequence to BLAST search and downloading the BLAST hits. The BLAST algorithm accepts the IUPAC ambiguity codes; however, it treats them as mismatches in the alignment. Therefore, a significant amount of information is lost. The sequence set built on the BLAST hits is used to build a local database and a reference matrix composed of sequence distance measures. The measures were calculated in the local alignment of the ambiguous sequence separately to all sequences from the BLAST recovered sequences set. A phylogenetic tree was generated showing the most similar samples. A principal component analysis plot further simplified the interpretation of samples diversity or dynamics.

The metagenome DNA from dairy plant and spontaneous fermentation in-house prepared curd cheeses analyzed for diversity within major types of *Lactococcus* phages revealed the distinct composition of phages detected in cheeses from dairy plant and in-house productions and higher diversity of phages from spontaneous fermentation on non-pasteurized milk, as compared to starter culture-based productions.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/app11198915/s1>, a text file with the script for bioinformatic analyses in R environment (applsci.R), a text file with the sequences in FASTA format (File_S1.fas), and data objects for R environment (image_applsci.Rdata).

Author Contributions: Conceptualization, visualization, supervision, and project administration, M.S.; methodology, formal analysis, data curation, resources, funding acquisition, and writing—original draft preparation, M.S. and A.O.-S.; laboratory work, B.P., I.K., K.M., and W.W.; writing—review and editing, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data reported are available in Supplementary Materials.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Marcó, M.B.; Moineau, S.; Quiberoni, A. Bacteriophages and dairy fermentations. *Bacteriophage* **2012**, *2*, 149–158. [CrossRef]
2. Kleppen, H.P.; Bang, T.; Nes, I.F.; Holo, H. Bacteriophages in milk fermentations: Diversity fluctuations of normal and failed fermentations. *Int. Dairy J.* **2011**, *21*, 592–600. [CrossRef]
3. Schmidt, M.T.; Olejnik-Schmidt, A.K.; Zareba, A.; Pezacki, M.; Wojewoda, I.; Grajek, W. Induction of loci mutation during lactococcus lactis spontaneous conversion to bacteriophage-insensitive phenotype. *Food Biotechnol.* **2010**, *24*, 332–348. [CrossRef]
4. Mahony, J.; van Sinderen, D. Current taxonomy of phages infecting lactic acid bacteria. *Front. Microbiol.* **2014**, *5*. [CrossRef]
5. Deveau, H.; Labrie, S.J.; Chopin, M.C.; Moineau, S. Biodiversity and classification of lactococcal phages. *Appl. Environ. Microbiol.* **2006**, *72*, 4338–4346. [CrossRef]
6. Oliveira, J.; Mahony, J.; Hanemaaijer, L.; Kouwen, T.R.H.M.; van Sinderen, D. Biodiversity of bacteriophages infecting *Lactococcus lactis* starter cultures. *J. Dairy Sci.* **2018**, *101*, 96–105. [CrossRef] [PubMed]
7. Fernández, L.; Escobedo, S.; Gutiérrez, D.; Portilla, S.; Martínez, B.; García, P.; Rodríguez, A. Bacteriophages in the dairy environment: From enemies to allies. *Antibiotics* **2017**, *6*, 27. [CrossRef]
8. Del Rio, B.; Martín, M.C.; Ladero, V.; Martínez, N.M.D.; Fernández, M.A.M. Bacteriophages in Dairy Industry: PCR Methods as Valuable Tools. In *Bacteriophages*; InTech: London, UK, 2012.
9. Cattonaro, F.; Spadotto, A.; Radovic, S.; Marroni, F. Do you cov me? Effect of coverage reduction on metagenome shotgun sequencing studies. *F1000Research* **2020**, *7*. [CrossRef]
10. Pacwa-Płociniczak, M.; Biniecka, P.; Bondarczuk, K.; Piotrowska-Seget, Z. Metagenomic Functional Profiling Reveals Differences in Bacterial Composition and Function During Bioaugmentation of Aged Petroleum-Contaminated Soil. *Front. Microbiol.* **2020**, *11*. [CrossRef] [PubMed]
11. Zotta, T.; Ricciardi, A.; Condelli, N.; Parente, E. Metataxonomic and metagenomic approaches for the study of undefined strain starters for cheese manufacture. *Crit. Rev. Food Sci. Nutr.* **2021**. [CrossRef]
12. Frantzen, C.A.; Holo, H. Unprecedented diversity of lactococcal group 936 bacteriophages revealed by amplicon sequencing of the portal protein gene. *Viruses* **2019**, *11*, 443. [CrossRef]
13. Labrie, S.; Moineau, S. Multiplex PCR for detection and identification of lactococcal bacteriophages. *Appl. Environ. Microbiol.* **2000**, *66*, 987–994. [CrossRef] [PubMed]
14. Eric, S.D.; Nicholas, T.K.D.D.; Theophilus, K.A. Bioinformatics with basic local alignment search tool (BLAST) and fast alignment (FASTA). *J. Bioinforma. Seq. Anal.* **2014**, *6*, 1–6. [CrossRef]
15. Ihaka, R.; Gentleman, R. R: A Language for Data Analysis and Graphics. *J. Comput. Graph. Stat.* **1996**, *5*, 299–314. [CrossRef]
16. Wright, E.S. Using DECIPHER v2.0 to analyze big biological sequence data in R. *R J.* **2016**, *8*, 352–359. [CrossRef]
17. PCAtools GitHub—kevinblighe/PCAtools: PCAtools: Everything Principal Components Analysis. Available online: <https://github.com/kevinblighe/PCAtools?s=03> (accessed on 8 June 2021).
18. Som, A. Theoretical foundation to estimate the relative efficiencies of the Jukes-Cantor+gamma model and the Jukes-Cantor model in obtaining the correct phylogenetic tree. *Gene* **2006**, *385*, 103–110. [CrossRef] [PubMed]
19. Szczepańska, A.K.; Hejnowicz, M.S.; Kołakowski, P.; Bardowski, J. Biodiversity of *Lactococcus lactis* bacteriophages in Polish dairy environment. *Acta Biochim. Pol.* **2007**, *54*, 151–158. [CrossRef] [PubMed]
20. Mahony, J.; Moscarelli, A.; Kelleher, P.; Lugli, G.A.; Ventura, M.; Settanni, L.; van Sinderen, D. Phage biodiversity in artisanal cheese wheys reflects the complexity of the fermentation process. *Viruses* **2017**, *9*, 45. [CrossRef] [PubMed]
21. Lavelle, K.; Martínez, I.; Neve, H.; Lugli, G.A.; Franz, C.M.A.P.; Ventura, M.; Dal Bello, F.; van Sinderen, D.; Mahony, J. Biodiversity of streptococcus thermophilus phages in global dairy fermentations. *Viruses* **2018**, *10*, 577. [CrossRef] [PubMed]
22. Moisan, M.; Moineau, S. Multilocus sequence typing scheme for the characterization of 936- like phages infecting *Lactococcus lactis*. *Appl. Environ. Microbiol.* **2012**, *78*, 4646–4653. [CrossRef] [PubMed]