



**HAL**  
open science

## Lebanese fermented goat milk products: From tradition to meta-omics

I. Ammoun, C.I. Kothe, N. Mohellibi, Catherine Béal, R. Yaacoub, Pierre Renault

► **To cite this version:**

I. Ammoun, C.I. Kothe, N. Mohellibi, Catherine Béal, R. Yaacoub, et al.. Lebanese fermented goat milk products: From tradition to meta-omics. Food Research International, 2023, 168, pp. 112762. 10.1016/j.foodres.2023.112762 . hal-04057222

**HAL Id: hal-04057222**

**<https://hal.science/hal-04057222>**

Submitted on 11 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## 26 1 Abstract

27 Ambriss, Serdaleh and Labneh El Darff are traditional Lebanese products made from  
28 fermented goat's milk from the Bekaa Valley, the Chouf Mountains and Southern Lebanon,  
29 respectively. A questionnaire completed by 50 producers of these products showed that they are  
30 prepared by periodic percolation either by milk or by Laban in amphorae or goatskins during the  
31 lactation season. Production is carried out on a small scale and in a limited number of production  
32 units, often by elderly people, resulting in a real risk of disappearance of these products and loss  
33 of the corresponding microbial resources will be lost. In this study, 34 samples from 18  
34 producers were characterized by culture-dependent and -independent analyses. The results  
35 obtained by these two methods were radically different, especially for Ambriss and Serdaleh, the  
36 latter revealing the co-dominance of *Lactobacillus kefiranofaciens*, a fastidious-growing species,  
37 and *Lactococcus lactis* in a viable but not culturable state. Overall, their composition is  
38 reminiscent of kefir grains. However, phylogenomic and functional analyses of the genomes of  
39 the key species *Lb. kefiranofaciens* show that they differed from those of kefir, for example in  
40 their polysaccharide genes, explaining the absence of grains. Labneh El Darff displays a  
41 dominance of *Lactobacillus delbrueckii*, probably due to the addition of Laban. In addition,  
42 several zoonotic pathogens were found, including *Streptococcus parasuis*, which was dominant  
43 in one sample and which MAG analysis revealed it acquired lactose utilization genes through  
44 horizontal gene transfer. The contamination of the herd with *Mycoplasma agalactiae* in the  
45 Chouf region was also revealed by MAG analysis of the Serdaleh samples. Antibiotic resistance  
46 genes were detected in most of the samples, particularly in the Serdaleh samples, where the  
47 dominant *L. lactis* strains possessed a plasmid with a multi-resistance island. Finally, this study  
48 paves the way for further analyses to shed light on the resilience of these ecosystems established  
49 in amphora or in goatskins and to improve hygiene practices for milk production.

50

51 **Keywords:** Microbial diversity, kefir, metagenomics, genomics, ecology, VBNC, metagenome-  
52 assembled genomes, LAB, *Lactobacillus kefiranofaciens*, *Streptococcus parasuis*,  
53 *Mycoplasma agalactiae*

54

## 55 2 Introduction

56 The Mediterranean region is well known for its goat farming, providing about 21% of the  
57 world's goat milk supply (Miller & Lu, 2019; Dubeuf et al., 2016). Due to the nature of its  
58 composition, organoleptic characteristics and healthy attributes, goat milk has been the focus of  
59 interest in recent years (Prosser, 2021; Clark and Mora García, 2017). This sector has received  
60 increasing attention, particularly in Lebanon, and continues to grow (Kayouli, 2015). Goat milk  
61 products are highly appreciated by local consumers (Balaa & Marie, 2008), especially for their  
62 specific “goaty taste” (Semaan et al., 2011). Compared to cow milk, its casein proteins are more  
63 easily digested and it is better tolerated by people with mild lactose intolerance. It has a higher  
64 content of short-chain fatty acids in milk fat (Shu et al., 2016; Zenebe et al., 2014) and contains  
65 minerals such as magnesium, calcium and phosphorus (Abbas et al., 2014).

66 These goat milk products are made either from standardized milk at the industrial scale,  
67 such as Laban (Chammas et al., 2006a), Labneh (Serhan et al., 2016), Shankleesh (Serhan &  
68 Mattar, 2013; Toufeili et al., 1995), Kishk (Salameh et al., 2016), or artisanally from raw milk,  
69 such as Darfiyeh (Hosri & El Khoury, 2004), Aricheh (Serhan & Mattar, 2013), Serdaleh (Tabet  
70 et al., 2019), Ambriss (Dimassi et al., 2020) and Labneh El Darff (Semaan et al., 2011). The  
71 diversity of these dairy products not only enriches the Lebanese culinary culture and pantry  
72 (Massaad, 2017), but also represents an important source of empowerment for women in poor  
73 rural areas. Female-headed households effectively contribute to the development of the goat  
74 rearing sector and the manufacturing of goat milk products (Kayouli, 2015). The market share of  
75 this sector is not well characterized; however, 70% of all small ruminant industries are located in  
76 rural areas with limited market access (Saadeh, 2016). Among these fermented goat milk  
77 products, Ambriss, Serdaleh and Labneh El Darff, made with milk from the Black Baladi race  
78 that grazes in the Bekaa Valley, the Chouf Mountains and Southern Lebanon, respectively  
79 (**Figure 1A**), are real candidates to be awarded the PGI (Protected Geographical Origin) label  
80 according to the report of the SRVA (Service Romand de Vulgarisation Agricole; IDEAS  
81 Centre, 2005; Ministère de l'Economie et du Commerce de la République Libanaise, 2006).  
82 However, it appears that these products are in danger of disappearing, mainly due to commercial,  
83 regulatory and food safety concerns (Serhan & Mattar, 2018; Semaan et al., 2011). In fact, these  
84 products are made on a seasonal basis, at small or medium production scales and according to  
85 traditional methods (**Figure 1B**), by successive fermentation of raw goat milk in amphorae for  
86 Ambriss (Dimassi et al. 2020) and Serdaleh (Serhan and Mattar, 2013) and Laban in goatskins

87 for Labneh El Darff (Serhan et al. 2016; Nasbimana et al. 2005).

88 In the Mediterranean region, traditional fermented milk products are closely related to the  
89 area in which they are produced and the cultural history and heritage (Boyazoglu & Morand-  
90 Fehr, 2001). This fact has drawn the attention of many studies due to the unique microbial  
91 diversity of these products (Rhaiem et al., 2016; Abdalla & Hussain, 2010; Ayad et al., 2004; El  
92 Soda et al., 2003). Numerous genera of lactic acid bacteria (LAB) such as *Lactobacillus*,  
93 *Lactococcus*, *Enterococcus*, *Leuconostoc* and *Streptococcus* are generally involved in these  
94 transformations. However, only a few studies have described the native microbiota in the  
95 Lebanese fermented milk products, and the characterization of the isolates has been restricted to  
96 phenotypic and biochemical analyses. The first attempts were recorded on Laban, where 96  
97 strains of *S. thermophilus* and *L. bulgaricus* were characterized (Chammas et al., 2006b). Serhan  
98 et al. (2009) then revealed the presence of the LAB *Streptococcus thermophilus*, *Enterococcus*  
99 spp., *Lactococcus lactis* subsp. *lactis* and *cremoris*, *Lactiplantibacillus plantarum*, *Lactobacillus*  
100 *curvatus* in Darfiyeh cheese using culture-dependent methods. Moreover, Dib et al. (2012)  
101 isolated 25 different strains in Serdaleh, Laban and Darfiyeh cheese belonging to the species *Lp.*  
102 *plantarum*, *Lacticaseibacillus paracasei*, *Lacticaseibacillus casei* and *Leuconostoc*  
103 *mesenteroides*

104 Until now, these traditional Lebanese products have only been characterized using  
105 cultural methods, which provide a patchy view of the microbiota, whereas new methods for  
106 characterizing food ecosystems are now available, based on techniques such as 16S and shotgun  
107 metagenomic analyses (Walsh et al., 2022). Approaches based on these techniques have already  
108 been used in a study on traditional Brazilian cheeses and provided a global picture of their  
109 microbiota, with taxonomic resolution close to the strain level for dominant populations and  
110 functional data at the gene level (Kothe et al., 2021). We therefore decided to develop a better  
111 understanding of the microbiota of Ambriss, Serdaleh and Labneh El Darff using these new  
112 methods to improve our knowledge of their processes, and to thus develop strategies to preserve  
113 them and boost their development. To do this, 34 samples of fermented goat milk products from  
114 18 producers were characterized using a combination of culture-dependent and culture-  
115 independent analyses, including genomic analyses of isolated strains and metagenomic  
116 assemblies, as well as functional analyses based on deep metagenomic profiling. The results  
117 were interpreted with regard to the manufacturing processes and physico-chemical characteristics  
118 of the products.

### 119 3 Materials and Methods

#### 120 3.1 Development of a producer's survey

121 A producers' survey was conducted over 50 breeders or producers from the three selected  
122 production regions – the Bekaa Valley, the Chouf Mountains and Southern Lebanon – for the  
123 artisanal products of Ambriss, Serdaleh and Labneh El Darff, respectively. The questionnaire  
124 consisted of 34 questions (**Supplementary datafile 1**), covering three main areas: (i) socio-  
125 demographic aspects; (ii) breeding activities; and (iii) production activities, including  
126 commercialization. The questions included general and socio-economic information about  
127 producers, information regarding livestock size, goat breed, breeding type and practices, milk  
128 production volumes, as well as the type of dairy product produced, the production season, the  
129 raw materials used, production volumes, prices and marketing practices. In addition, a detailed  
130 observation of the processing was carried out in order to build a common production flow chart  
131 for each product type, showing the existing differences between the producers as well.

#### 132 3.2 Sample collection, physic-chemical and microbiological analyses

133 Producers were first selected on the basis of the survey at the different production sites  
134 according to two main criteria: their production was small-scale and they had been producing  
135 their product for more than 15 years. In total, 34 artisanal samples from 18 producers were  
136 collected during two consecutive production seasons in 2018 and 2019 (**Table S1**).

137 Samples of fermented goat milk products (10 g) were taken aseptically and homogenized  
138 for 60 s in a laboratory blender (Blender LB20E, Waring, Torrington, CT, USA) with 90 mL and  
139 20 mL of distilled water, for pH and salinity tests, respectively. The pH was measured using a  
140 digital pH meter (GLP 21, Crison instrument, Barcelona, Spain). The salinity content was  
141 determined using chloride test strips (30-600 mg/L; QuanTabTitrator, Hach, Loveland, CO,  
142 USA). All measurements were performed in triplicate. Samples were preserved in phosphate-  
143 buffered saline (PBS 15X, pH 8.7), and stored at -20°C for further analysis.

144 For culture-dependent analyses, samples were diluted and homogenized in buffer peptone  
145 water (2%). They were spread in triplicate on MRS (pH 6.2, Sigma-Aldrich, St. Louis, MO,  
146 USA) and M17 (pH 6.5, Sigma-Aldrich, St. Louis, MO, USA) agar plates. Plates were incubated  
147 for 48 h, anaerobically for MRS at 30°C and 42°C to allow mesophilic and thermophilic  
148 *Lactobacillus* counts, respectively, and aerobically for M17 at 30°C for lactococci.

149 Specific procedures were applied to isolate *L. lactis* and *Lactobacillus kefiranofaciens*  
150 strains. Culture media were supplemented with 2% salt and 10% goat milk for *L. lactis*, or with  
151 25% goat milk and adjusted at pH 5.3 for *Lb. kefiranofaciens*, and incubated anaerobically at  
152 30°C for 10 days.

153 Based on their distinct morphology (shape, color and size), colonies were selected and  
154 purified by repetitive streaking on the same medium. Isolates were assessed for their catalase  
155 activity and Gram coloration according to the criteria described by Kandler and Weiss (1986), as  
156 well as for their ability to produce gas from glucose by inverted tubes (Zúñiga et al., 1993).  
157 Biochemical identification and carbohydrate fermentation patterns were determined by API 50  
158 CH and API 20 STREP test kits (BioMérieux Inc., Marcy l'Etoile, France) following the  
159 manufacturer's recommendations.

160 In order to identify the different morphotypes, isolates were suspended in 400 µL Tris-  
161 EDTA (TE 1X, pH 8.0) buffer in a 2-mL tube. DNAs was lysed using a bead-beater (24  
162 Microtubes BeadBlaster, Thomas Scientific, NJ., USA), with 100 mg zirconium beads of 0.1 and  
163 0.5-mm diameters and homogenized for 20 s at 4.5 m/s. Once extracted, amplification of the  
164 coding gene for 16S rRNA was carried out for 30 cycles on a final mix of 25 µL containing the  
165 following universal primers: Bact-27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-  
166 GGTTACCTTGTTACGACTT-3)', using Taqara DNA Taq polymerase (Almeida et al., 2014).  
167 Products were purified using a purification kit (Gen Elute PCR clean up, Sigma, St. Louis, MO,  
168 USA). Genotypic identification was carried out by blasting the different sequences of the isolated  
169 strains using the NCBI database.

### 170 3.3 Genomic DNA extraction and sequencing

171 DNA from five isolates, chosen as representative of the main species detected in our  
172 products, was extracted from the bacterial cells according to the protocol described by Kothe et  
173 al. (2021). DNA sequencing was carried out on an Illumina HiSeq at GATC-Biotech (Konstanz,  
174 Germany) to generate about 5-6 million paired-end reads (150 bases in length). For each strain,  
175 the paired-end reads were merged and *de novo* assembly was performed using SPAdes 3.9  
176 (Bankevich et al., 2012). Only contigs with length > 300 bp and coverage > 100 were considered  
177 for further study. Annotations were performed with Rapid Annotation using the Subsystem  
178 Technology server (Aziz et al., 2012).

### 179 3.4 DNA extraction from the products

180 Samples (0.5 g) were diluted 1:1 (w/v) in guanidinium thiocyanate 4M solution (Sigma-  
181 Aldrich, MO, USA) with TrisHCl 0.1M and processed in the same way as for the isolate DNA  
182 extraction. The mixture was then supplemented with 10% N-Lauryl sarcosine (Sigma-Aldrich,  
183 MO, USA), vortexed and centrifuged at 4°C and 14,000 g for 30 min. Supernatant were  
184 discarded and extraction was continued on the emerged pellet, as described by Almeida et al.  
185 (2014). DNA was visually evaluated on 0.8% agarose gel and quantified by a fluorometer (Qubit  
186 2.0, ThermoFisher Scientific) supplied with the Qubit dsDNA HS Assay Kit (ThermoFisher  
187 Scientific).

### 188 3.5 Metagenetic analysis

189 Bacterial diversity was analyzed by sequencing the amplified region V3-V4 of the 16S  
190 rRNA gene using primers V3F and V4R (V3F: 5'-CCTACGGGAGGCAGCAG-3'/ V4R: 5'-  
191 GGACTACHVGGGTWTCTAAT-3'). The PCR was performed with MTP Taq DNA  
192 Polymerase (Sigma-Aldrich, St. Louis, MO, USA), and cycling conditions were: 94°C for 1 min,  
193 followed by 30 cycles of amplification at 94°C for 1 min, 65°C for 1 min, and 72°C for 1 min,  
194 with a final extension step of 10 min at 72°C. Sequencing was performed with the V3 Illumina  
195 MiSeq kit, as described in Poirier et al. (2018), and fasta files were generated at the end of the  
196 run (MiSeq Reporter software, Illumina, USA).

197 The quality of the raw data was evaluated with FastQC (Wingett & Andrews, 2018) and  
198 the sequences were imported into the FROGS (Find Rapidly OTUs with Galaxy Solution)  
199 pipeline (Escudié et al., 2018) to obtain the Operational Taxonomic Units (OTUs). The  
200 sequences were filtered by length (480-580 bp) and then pooled into OTUs with SWARM (Mahé  
201 et al., 2014) with the distance parameter of 1. Chimeras were removed with VSEARCH (Rognes  
202 et al., 2016) and OTUs with at least 0.005% in the whole dataset were retained (Bokulich et al.,  
203 2013). OTUs were affiliated with the 16S SILVA 138.1 database (Quast et al., 2013). Alpha-  
204 diversity and beta-diversity analyses were performed with R Studio v.3.6.1 using the phyloseq  
205 and ggplot2 packages (v1.30.0). Statistics were performed on alpha-diversity using Kruskal-  
206 Wallis and linear model regression. DESeq2 was applied to characterize statistically significant  
207 differentially abundant OTUs in the different products (Love et al., 2014). The PCoA was plotted  
208 with the ward.D2 method and Permanova tests were performed between samples.

### 209 3.6 Taxonomic composition by shotgun metagenomics study

210 Sixteen fermented products (six Ambriss, six Serdaleh and four Labneh) were selected on  
211 the basis of the metagenetic results for more in-depth analysis. The metagenomic DNA was  
212 sequenced using Illumina HiSeq2500 technology at GATC-Biotech (Konstanz, Germany).  
213 Metagenomic reads corresponding to the *Capra hircus* genome were filtered out by reads  
214 mapping against the *Capra hircus* reference genome (LWLT000000000) using Bowtie2  
215 (Langmead & Salzberg, 2012). The proportion of goat reads in the samples was then assessed  
216 using samtools flagstat (Li et al., 2009). On the basis of the fastq file generated after host read  
217 filtration, we first estimated microbial composition by mapping the samples. Reads were mapped  
218 against the representative clade-specific marker catalogue contained in the MetaPhlAn tool v.  
219 3.0.4 (Truong et al., 2015).

220 In addition, we performed taxonomic profiling using an assembly-based marker gene  
221 analysis, which allows non-supervised binning of metagenomes using metaSPAdes v.3.9  
222 (Bankevich et al., 2012). Genes were then predicted using Prodigal (v.2.6.3) and marker genes  
223 were extracted using fetchMG, v.1.0 (Sunagawa et al., 2013). Taxonomic assignment was carried  
224 out with the COG0012, *ychF* gene. Summary species composition plots were created in R v.3.6.1  
225 using the package ggplot2 v.3.3.2.

### 226 3.7 Metagenome-assembled genome (MAG) analyses

227 Genome binning was performed using MetaBAT2-2.12.1 (Kang et al., 2015), with a  
228 minimum contig size of 1,500 nucleotides and the default settings. The quality of the resulting  
229 prokaryotic bins was assessed with CheckM (Parks et al., 2015), and MAGs < 80% completeness  
230 and/or > 10% contamination were excluded. Several MAG assemblies were refined by additional  
231 binning, such as blastn analysis with reference genomes of related species. Contigs were then  
232 filtered by percentage of the length covered, identity and coverage levels. Finally, manual curing  
233 of questionable contigs (new, low coverage of homology on references, etc.) were performed by  
234 blastn and blastx analysis against the nr/nt and nr NCBI databases, respectively. Annotations  
235 were performed using the Rapid Annotation with Subsystem Technology server.

### 236 3.8 Phylogenomic and functional analyses

237 The Automatic Multi-Locus Species Tree web server (Alanjary et al., 2019) was used to  
238 determine closely related genomes based on core gene alignments of the sequenced genomes and  
239 recovered MAGs. The closest species were inferred based on the percentage of average

240 nucleotide identity (ANI) calculated using FastANI, v.1.31 (Jain et al., 2018). The kSNP (v.3.0)  
241 was used to perform the phylogenomic analysis using distances based on single nucleotide  
242 polymorphism variations for the most representative species present in MAGs (Gardner et al.,  
243 2015) with the maximum likelihood option and kmer size of 21. Genomes of the corresponding  
244 species were downloaded from Genbank and combined with our MAG to compute this analysis.  
245 Gene content analysis and metabolic reconstruction was performed on SEED using sequence and  
246 function-based comparison tools (Aziz et al., 2012).

247 Functional analysis on metagenomes was performed by read mapping profiling with the  
248 sequences of the 16 metagenomes from which the *Cellulosimicrobium cellulans* and goat reads  
249 were removed to avoid bias in the gene quantification. The analysis of microbial pathways was  
250 performed by HUMAnN 3.0 (Beghini *et al.*, 2021)(REF) in order to evidence bacteriocin  
251 production genes, potentially health-associated gene clusters or aroma pathways. To screen the  
252 potentially health-associated gene clusters, we followed essentially a methodology described  
253 previously (Leech *et al.*, 2020). except the functional profiles were established with HumanN3.  
254 These genes were binned into one of the following categories: survival, modulation, and  
255 colonization in order to create a Heatmap using Pheatmap. Clusters counts displayed in the  
256 heatmap were not absolute counts but scaled so that the values could be compared between the  
257 samples. The genes potentially implied in bacteriocin production were searched with the term  
258 “bacteriocin” and those of the diacetyl pathways with “Acetoin, Diacetyl, Acetolactate, Citrate  
259 synthase, Citrate lyase, Citrate \[pro-3S\]-lyase and Oxaloacetate decarboxylase”.

260 The search for antibiotic resistance (ABR) genes was performed by read mapping against  
261 the CARD database (Jia et al., 2017; McArthur et al., 2013) using the PATRIC web server  
262 (Antonopoulos et al., 2019). The presence of particular genes in reference genomes, ABR and  
263 genes of technological interest was determined using the PATRIC web server, the  
264 FoodMicrobiome tool (<https://migale.jouy.inra.fr/foodMicrobiome/>), and blast analysis.  
265 Additionally, FoodMicrobiome was used to detect the subdominant populations of given species  
266 with high sensitivity and reliability and to discriminate between subspecies.

267

### 268 3.9 Data availability

269 Raw sequences were deposited on the European Nucleotide Archive (ENA) under the  
270 BioProject ID PRJNA860091 for shotgun metagenomic, amplicon and genomic reads. The  
271 MAGs are available in doi: 10.17632/7w4rd76ndj.2.

272

## 273 4 Results

### 274 4.1 Goat farmer and dairy producer's survey

275 According to the results of the questionnaire, 25% of the respondents were identified as  
276 both producers and farmers, and 75% as producers only. Production depended on female  
277 housekeepers (60%) aged between 40 and 70 years old, and had limited access to the local  
278 market place (38%). Most of producers (88%) used only goat milk compared to 4% who used a  
279 mixture of goat and cow milk, while the rest used goat Laban in addition to goat milk as the  
280 primary dairy material. Producers who reared their own goat herd were able to process larger  
281 milk quantities for the three fermented products (135 to 7,500 L/season) than producers who  
282 purchased their goat milk (140 to 1,890 L/season). Farming and production parameters for each  
283 product are summarized in **Table S1**.

284 Ambriss, Serdaleh and Labneh El Darff were made with milk from the Black Baladi race  
285 that grazes in the Bekaa Valley, the Chouf Mountains and Southern Lebanon, respectively  
286 (**Figure 1A**). All production phases were completely manual, the production scale was small to  
287 medium (Third Quartile = 216.5 kg/season) and both tools and materials were artisanal.  
288 Production takes place during the period when the goats produce milk, i.e., from June to  
289 November. The manufacturing process of Serdaleh and Ambriss lasts up to 6 months and  
290 consists of fermenting raw goat milk on a daily basis in terracotta jars, except for 20% of  
291 Serdaleh producers who used plastic barrels. The jars were first soaked with salt, periodically  
292 refilled with fresh raw milk and salt, with periodic whey evacuation from a hole at their bottom  
293 (**Figure 1B and S1**). The manufacturing of Labneh El Darff is conducted in a three-stage process  
294 consisting of: (i) Laban production, which is prepared from raw goat milk inoculated with a  
295 previously homemade inoculum produced by backslopping; (ii) filling the “Darff” (bag formed  
296 by the whole goatskin, but sometimes cloth fabric) with the produced Laban; and (iii) straining  
297 of the whey from the suspended Darff during a period of 6 weeks (**Figure 1B and S2**). The  
298 manufacturing processes of these products are carried out at ambient temperature.

## 299 4.2 Physico-chemical and culture-dependent analyses

### 300 4.2.1 Temperature, pH and salinity

301 Temperature, pH and salinity levels were measured at the time the samples were collected  
302 (**Table S2**). The ambient temperature at which the products were developed varied between 28°C  
303 and 33°C for Ambriss and Labneh El Darff, but was lower ( $p < 0.05$ ) for Serdaleh (between  
304 20°C to 25°C). The pH values showed little variability within the same product type, unlike  
305 salinity, which displayed greater intra-group variability (**Figure 2A**). The pH values were  
306 comprised between 3.01 and 4.80. They were higher ( $p < 0.05$ ) in Ambriss (4.08 to 4.80) when  
307 compared to Serdaleh (3.01 to 4.12) and Labneh El Darff (3.62-4.62). The Serdaleh product had  
308 the highest level ( $p < 0.05$ ) of salinity (2.69 to 5.04%) compared to Ambriss (0.87 to 2.24%) and  
309 Labneh El Darff (1.10 to 1.75%).

### 310 4.2.2 Culture-dependent analyses

311 The examination of culture results shows a significant variability in the different microbial  
312 enumerations (**Figure 2B**). We observed a dominance of mesophilic lactobacilli (MRS 30°C,  
313 5.4-6.8 Log CFU/g) over lactococci (M17 30°C, 0.5-5.20 Log CFU/g) in the three products.  
314 Thermophilic lactobacilli counts were highly variable in Labneh el Darff (MRS 42°C, 1-6.2 Log  
315 CFU/g) compared to its other counterparts.

316 Overall, 279 strains belonging to 13 species groups were genetically identified by 16S  
317 rRNA from 380 isolates. The most commonly cultivated species in the three fermented milks  
318 was *Lactiplantibacillus plantarum*. This species appeared to be dominant in Ambriss (38% of  
319 isolates in this product) and Serdaleh (77%), and co-dominant with other species in Labneh El  
320 Darff (16%). In Ambriss, we also identified a strong prevalence of the species *Enterococcus*  
321 *durans* and *Lacticaseibacillus rhamnosus* (18% of each), in Labneh El Darff, *Levilactobacillus*  
322 *brevis* was the most frequently obtained cultured species (32% of isolates), followed by *Lb.*  
323 *delbrueckii* and *L. lactis* (16% of each) (**Figure 2C**).

## 324 4.3 Taxonomic composition using amplicon sequencing

325 The 16S rRNA amplicon-sequencing analyses were clustered in 209 OTUs, whose  
326 abundance was over the threshold of 0.005% in the whole dataset. The rarefaction curves of most  
327 of the analyzed samples reached the saturation plateau, indicating that the sequencing depth was  
328 sufficient (**Figure S3**).

329 The alpha-diversity analysis revealed lower richness of species in Labneh El Darff ( $p <$   
330 0.05) compared to the other two products (**Figure 3A**). Moreover, the Chao1, Shannon and  
331 InvSimpson indices, which measure evenness, confirmed this low diversity. The Principal  
332 Coordinates Analysis (PCoA) showed that while Labneh El Darff samples were clearly  
333 differentiated, samples of Serdaleh and Ambriss could not be clearly separated and thus likely  
334 shared a similar bacterial microbiota (**Figure 3B**). Indeed, the analysis of their taxonomical  
335 composition (**Figure 3C** and **Table S3**) showed that *Lactobacillus delbrueckii* was the main  
336 species present in Labneh El Darff samples, representing 55 to 82% of the counts, whereas it was  
337 undetected or at levels below 0.4% in the samples of Ambriss and Serdaleh. Moreover, the  
338 samples of the latter products contained *L. lactis*, *Lb. kefiranofaciens* and *Lentilactobacillus*  
339 *kefiri* as the major species and, to a lesser extent, *Lentilactobacillus diolivorans*-group.  
340 Differential analysis of OTUs showed that Serdaleh samples had greater *L. lactis* abundance than  
341 Ambriss ( $p_{\text{adj}} < 0.01$ ). Remarkably, in the PCoA (**Figure 3B**), these two products appeared to be  
342 spread along Axis 1 as a function of the proportion between Lactococcaceae and  
343 Lactobacillaceae species. In particular, a group of 12 Serdaleh samples contained *L. lactis* at  
344 levels of 58-83%, while the other samples were intermixed on the other part of this axis.  
345 Moreover, the 13 samples of Serdaleh collected in 2019 contained low levels (0.67-2.2%) of  
346 *Mycoplasmopsis agalactiae* or *M. bovis* (**Table S3**). Finally, three Ambriss samples showed  
347 divergent compositions compared to the others: A3-b was dominated by *Lentilactobacillus*  
348 *buchneri*, A4 by the *S. parasuis-sanguinis* group ( $> 53\%$  of the reads), and A6 by high amounts  
349 of species poorly present in other samples such as *L. gasseri* (~36%), *Lc rhamnosus* (~23%) and  
350 Enterobacteriaceae ( $> 18\%$  of reads). The other species, usually detected at low levels in the  
351 various samples, mainly belonged to the Gammaproteobacteria class, such as *Aeromonas sp.*,  
352 *Citrobacter sp.*, *Klebsiella sp.* and *Escherichia-Shigella*.

#### 353 4.4 Refined microbiota by metagenomics

354 For a refined taxonomical analysis, 16 samples were selected to perform shotgun  
355 metagenomic approaches. Such analyses allow a taxonomic identification up to the species-strain  
356 level and the simultaneous assessment of the level of reads corresponding to bacteria and fungi.  
357 Moreover, these analyses make it possible to detect eventual virulence factors in pathogenic  
358 species, antibiotic resistance genes and genes of technological interest.

359 Food metagenomic samples may contain animal reads, which may significantly bias the  
360 analysis when they are in high abundance. Since the present products are manufactured from

361 goat milk, we tested for the presence of *Capra hircus* reads that were detected at levels ranging  
362 from 2.3 to 72% (**Table S4**). They were notably abundant in Labneh El Darff samples where  
363 they accounted for more than 50%, as a consequence of the process mode where the product is  
364 drained in a goatskin.

#### 365 4.4.1 Species-level assignments by shotgun metagenomics

366 In a first approach, the MetaPhlan tool, which identifies species by mapping the  
367 metagenomic reads against a genomic database, was used to determine the taxonomic  
368 composition of the samples. It yielded 98 bacterial and one eukaryotic species (**Table S5**).  
369 Species composition obtained by this analysis was relatively in accordance with those obtained  
370 with the 16S rRNA amplicon. Interestingly, it resolved most taxonomical assignment  
371 ambiguities for lactic acid bacteria such as *Lt. kefir*, *Lt. diolivorans*, *Lactiplantibacillus*  
372 *plantarum*, *Lactobacillus helveticus* and *S. thermophilus*, and also for contaminants such as *S.*  
373 *parauberis* and *M. agalactiae*. It also detected the yeast *Saccharomyces cerevisiae*, but at low  
374 levels and in only three samples.

375 In order to characterize microbiota independently of fixed references, we applied a marker  
376 gene analysis from the assembled metagenomes, which yielded 104 bacterial and 15 fungal  
377 species (**Table S6**). The overall composition is shown in **Figure 4**. Although the main bacterial  
378 species identified here were consistent with the previous analysis made by read profiling, some  
379 differences merit highlighting. Whereas MetaPhlan attributed 35% of the reads of A4 to *S.*  
380 *parauberis*, the marker gene analysis detected only 6.9% of this species and further detected  
381 another streptococcal species, *S. parasuis*, which retained 55% of the reads. This species was  
382 also detected at low levels in four other samples, while *S. parauberis* was found in eight samples,  
383 generally at low levels. This analysis also confirmed the frequent presence at low levels of  
384 several species of Enterobacteriaceae such as *E. coli*, *K. pneumoniae* and *C. freundii*, and of  
385 Enterococcaceae such as *E. faecalis* and *E. faecium*. Lastly, it also established that *L. lactis*  
386 subsp. *lactis* was detected in all samples, whereas the *cremoris* subspecies was detected at  
387 concentrations ten times lower in S8 and S10-b. A specific analysis by read mapping confirmed  
388 the presence of *L. lactis* subsp. *cremoris* at 5-15% lower abundance than that of the *lactis*  
389 subspecies in most samples.

390 Finally, 14 species of yeasts whose genomes are absent from the MetaPhlan database were  
391 identified by marker gene analysis. The most frequent and abundant were *Pichia kudriavzevii*

392 and, to a lesser extent, *S. cerevisiae* and *Kluyveromyces marxianus*. Nevertheless, yeast relative  
393 abundance was generally below 1%, except in Labneh El Darff where it reached 8%. A more  
394 sensitive analysis performed by read mapping (**Table S7**) confirmed the presence of these yeasts  
395 at low levels in more samples, as well as that of *Torulaspota delbrueckii*.

#### 396 4.4.2 MAGs and strain-level analyses of LAB

397 A total of 62 prokaryotic Metagenome Assembled Genomes (MAGs) of high quality were  
398 reconstituted from the metagenome assemblies (**Table S8**). The majority of the MAGs  
399 correspond to Lactobacillaceae such as *Lt. diolivorans* (12), *Lt. kefir* (10), *Lb. kefiranofaciens*  
400 (9), *Lp. plantarum* (5) and *L. gasseri* (1). The two yoghurt species *Lb. delbrueckii* subsp.  
401 *bulgaricus* (1) and *S. thermophilus* (1) were also recovered from Labneh El Darff and MAGs of  
402 *L. lactis* subsp. *lactis* (5) were reconstituted from Ambriss and Serdaleh products. In addition, the  
403 results included MAGs of the goat pathogen *M. agalactiae* (6) from Serdaleh samples, and of  
404 milk contaminants such as *S. parauberis* (3), *S. parasuis* (1), *Escherichia coli* (1), *Klebsiella*  
405 *pneumoniae* (1), *Aeromonas caviae* (2) and *Acetobacter spp.* (2).

406 Further detailed analysis was performed with the lactococci and lactobacilli MAGs and  
407 genome sequences of the most frequent and sometimes dominant species in the three products  
408 (**Table S9**). Phylogenomic analysis of the *L. lactis* FME150 genome and the five *L. lactis* MAGs  
409 was performed with 215 reference genomes of this species, representing strains from its two  
410 main subspecies (*lactis* and *cremoris*) and isolated from different environments (dairy, vegetable,  
411 animal, etc.) (**Figure 5A**). In this analysis, five groups could be highlighted, one of which  
412 included all the representatives of the *cremoris* subspecies (brown branches), and three were  
413 formed by (i) industrial dairy strains (in red), (ii) diacetyllactis biovariant strains (in blue), and  
414 (iii) dairy strains isolated from traditional products (in violet). The last group (green branches),  
415 which was the most diverse, contained environmental strains (plants, animals, etc.). The five *L.*  
416 *lactis* MAGs from Ambriss and Serdaleh were included in this last group (highlighted by  
417 external red arrows).

418 A similar study was done with *Lb. kefiranofaciens*, *Lt. kefir*, *Lt. diolivorans* and *Lp.*  
419 *plantarum*. The *Lb. kefiranofaciens* FME151 genome and the eight *Lb. kefiranofaciens* MAGs  
420 assembled from Serdaleh (6) and Ambriss (3) metagenomes were compared to seven genomes  
421 present in the NCBI database, including those of the two type strains for the subspecies  
422 *kefiranofaciens* and *kefirgranum*. These strains were isolated from kefir of different origins.

423 ANI analysis showed that all strains shared at least 98.1% identity and could be divided into two  
424 main groups, one made by kefir strains and the second by Ambriss and Serdaleh strains, where  
425 ANI were ~99.2-99.4% within each group and ~98.2-98.4% between groups. Further  
426 phylogenomic analyses (**Figure 5B**) showed that two groups could be differentiated within kefir  
427 strains, a first one with three strains sharing over 99.9% identity and including the *Lb.*  
428 *kefiranofaciens* subsp. *kefiranofaciens* type strain, and a second one with four strains that were  
429 more divergent between themselves, sharing 99.0 to 99.8% identity and including the *Lb.*  
430 *kefiranofaciens* subsp. *kefirgranum* type strain. Lastly, Ambriss and Serdaleh strains also formed  
431 two groups sharing 99.4 to 99.8% ANI within and 99.1 to 99.3% outside of their group. Analysis  
432 of gene composition confirmed and refined these results. First, it showed a very high relatedness  
433 of three *Lb. kefiranofaciens* subsp. *kefiranofaciens* strains from kefir that shared more than 95%  
434 of their genes (> 2,325 out of ~2,440 genes), while within each of the other groups, the strains  
435 were more divergent between themselves (87% of shared genes in the *kefirgranum* group, 88%  
436 in the Serdaleh group and 92% in the Ambriss group). Seventeen EPS-related genes involved in  
437 kefiran production present in a 14.4-kb region were present only in the three former strains,  
438 whereas only the first part of the operon containing the genes of enzymes involved in EPS  
439 regulation and priming was conserved in the other strains. Furthermore, in the latter strains, the  
440 genes involved in polymerization, chain length determination and export that could be  
441 characterized appeared to be different in each strain, indicating that they produce  
442 polysaccharides other than kefiran.

443 The *Lt. kefir* FME71 genome and the ten *Lt. kefir* MAGs assembled from Ambriss (3),  
444 Serdaleh (6) and Labneh El Darff (1) were compared to the ten reference genomes present in the  
445 NCBI database, including seven strains isolated from kefir, two from human gut and one from  
446 yoghurt. This analysis showed that all strains shared very high ANI values (> 99.55%),  
447 indicating that they were closely related (**Table S10**). Analysis of their phylogenomic tree did  
448 not indicate clustering as a function of the origin of the strains (**Figure S4A**). Furthermore, the  
449 analysis of the core/pan genome showed that all strains shared ~2,148 genes out of ~2400. The  
450 analysis of the specific genes in the different strains indicated that they code mostly for  
451 “hypothetical proteins” (50%) and mobile element proteins (25%).

452 ANI analysis of *Lb. diolivorans* FME74 and of our 12 MAG strains with that of the type  
453 strain (isolated from maize silage) and one MAG from an ultra-filtered-milk-permeate  
454 fermenting bioreactor showed that these strains could be clustered into four groups (**Figure**

455 **S4B**), displaying at least ~97.7 to 98.4% ANI . Serdaleh MAGs formed a group of very closely  
456 related strains (ANI > 99.9%), Ambriss and Labneh MAGs formed a second homogenous group  
457 (ANI > 99.1%), whereas the bioreactor MAG and the type strain were isolated.

458 The sequences of *Lp. plantarum* FME91 strain and of the five MAGs obtained from  
459 Serdaleh were analyzed with a set of 145 *Lp. plantarum* strains isolated from different  
460 ecosystems, including several types of food, animals, gut, etc. The five MAGs from Serdaleh  
461 appeared to be almost identical (ANI > 99.97%), but unrelated to that of FME91 from Ambriss  
462 (ANI ~92.7%), and they ranged between 95.7 and 99.7-99.9, with an average of 98.7 with the  
463 strains of the set (**Table S10**). An overview of the phylogenetic relationship between these strains  
464 is given by the tree obtained by kSNP analysis (**Figure S4C**). Strains were dispersed irrespective  
465 of their environmental origin. The Ambriss strain FME91 was closely related to those isolated  
466 from milk or artisanal dairy products, whereas the Serdaleh strains were closely related to those  
467 isolated from meat, fruits and vegetables.

#### 468 4.4.3 Detection of metabolic features by functional profiling

469 Microbial pathway analysis was carried out by HUMAnN 3.0 in order to highlight  
470 interesting features in the microbial communities such as bacteriocin production, clusters of  
471 genes potentially associated with health or aromatic pathways. Bacteriocin genes were searched  
472 in the 16 metagenomes, and 29 potential bacteriocin-encoding genes were identified (**Table**  
473 **S11**). The most frequent and abundant were Lactococcin 972 family bacteriocin and Helveticin,  
474 which are mainly from *L. lactis* and *Lb. kefiranofaciens*, respectively. Other frequent potential  
475 bacteriocin-encoding genes are Plantaricin and Thiazolpeptide-type bacteriocin, likely present in  
476 *Lp plantarum* and *Lt diolivorans*, respectively.

477 Potentially health-associated gene were divided into three broad categories, “survival”,  
478 “colonization” and “modulation”, as presented previously (Leech *et al.*, 2020). Briefly, these  
479 gene clusters were shown to be important for surviving in the gut conditions, colonizing the gut  
480 microbiome and stimulating immune system or microbiota products. While most samples  
481 displayed relatively similar profiles, several samples appeared to have different profiles (**Figure**  
482 **S5**). For example, Ambriss A6 presented the highest ratio of survival clusters, while Ambriss A4  
483 presented the lowest ratio for all clusters. A search for overrepresented genes in Ambriss A6  
484 highlighted the presence of genes from *Lb. gasseri*, a species that was found only in this sample.

485 The low scores in Ambriss A4 could be due to the dominance of *S. parasuis* and the low  
486 abundance of LAB in this sample.

487 Finally, genes involved in diacetyl metabolism were searched through citrate utilization,  
488 acetolactate and acetoin-butanol production. Since genomic references were known for dominant  
489 species of these microbiota, the corresponding genes could be assigned at the species level. It  
490 confirmed that the citrate-lyase cluster was absent from *L. lactis* strains but present in *Lb.*  
491 *kefiranofaciens* in our samples. Finally, the diacetyl degrading enzymes acetolactate  
492 decarboxylase (EC 4.1.1.5) and acetoin (diacetyl) reductase (EC 1.1.1.304) were found in *L.*  
493 *lactis*, *Lt. kefiri* and *Lt. diolivorans*, respectively.

#### 494 4.4.4 Detection of pathogens by metagenomic analysis

495 Marker gene analysis showed that undesirable species were present in several Ambriss  
496 and Serdaleh samples. All Serdaleh samples collected in 2019 contained low levels (1-4%) of the  
497 goat pathogen *M. agalactiae*. Moreover, zoonotic streptococcal species, namely *S. parauberis*  
498 and *S. parasuis*, were detected at 0.5-2% in several samples of Ambriss and Serdaleh and,  
499 remarkably, at 7 and 55%, respectively, in Ambriss A4. *S. aureus* was detected at low levels in  
500 Serdaleh S10-b. Finally, Enterobacteria such as *E. coli* and *K. pneumoniae* were detected in  
501 about half of the samples at levels < 1% (except for *K. pneumoniae* present at 2.7% in Ambriss  
502 A1). Deep profiling analysis in the metagenome samples was performed to detect the presence at  
503 low abundance of pathogen species potentially found in dairy products (**Table S7**). No *Brucella*  
504 *abortus*, *Listeria monocytogenes*, *Mycobacterium vaccae* or *Staphylococcus agalactiae* were  
505 detected. However, *S. aureus* was detected at low levels in five Serdaleh samples.

506 Further analyses were performed on these undesirable species, which were sufficiently  
507 covered to yield MAGs. An ANI analysis of these MAGs with representative genomes of these  
508 species confirmed the taxonomic classification previously obtained by the *ychF* marker. For  
509 example, *M. agalactiae*, *S. parauberis* and *K. pneumoniae* MAGs share 99.8, 98.5 and 98.4%  
510 ANI with the type strains of their species, *E. coli* MAG; 99.1% with *E. coli* K12 laboratory  
511 strain, *S. parasuis* MAG; and 97-98% with four genome references of this species present in the  
512 NCBI database.

513 Gene content analysis was performed with *S. parasuis* MAG\_A4, which appears to be  
514 dominant in Ambriss A4. In particular, genes involved in lactose metabolism were searched to  
515 explain the fitness of this strain in milk fermentation. Among the two systems used for lactose

516 utilization, MAG-A4 contains the Lac-PTS system, including ten genes encoding the Lac-PTS  
517 transporter, the phospho-beta-galactosidase, the tagatose pathway, glucokinase, a repressor and a  
518 gene of unknown function (LacX). A blast search with these genes provided a hit at 97-99%  
519 nucleotide identity on its entire length with an ICESsuJH1308-1 mobile element present in *S.*  
520 *suis* JH1308 and a lower identity with several other *S. suis* strains. However, none of the nine *S.*  
521 *parasuis* genomes available contained these genes. These data show that *S. parasuis* MAG\_A4  
522 acquired lactose utilization genes by horizontal transfer, possibly from *S. suis*.

523 Interestingly, MAGs from *M. agalactiae* were of very good quality, probably due to  
524 particular features of their sequences, such as genomic guanine-cytosine content (~29.7%) and  
525 di-tri-nucleotide composition, allowing a better binning. In particular, MAGs from Serdaleh S8  
526 and S9 consisted of only four and ten contigs, respectively. ANI analysis indicated that the  
527 strains from the Serdaleh samples were very closely related to each other (ANI > 99.9%), as well  
528 as to the type strain (ANI ~99.85%; **Table S10**). Phylogenomic analysis confirmed the close  
529 relationship of *M. agalactiae* Serdaleh MAGs together and with PG2<sup>T</sup> (**Figure 5C**). Furthermore,  
530 gene content analysis indicates that these two MAGs differed from PG2<sup>T</sup> by less than 20 genes  
531 encoding hypothetical proteins.

#### 532 4.4.5 Detection of antibiotic resistance genes

533 A mapping-based approach against a comprehensive collection of antibiotic resistance  
534 genes was performed. A total of 18 ABR genes were identified, belonging to seven different  
535 classes of antibiotics (**Figure 6**). They were detected in almost all samples, and notably those  
536 from Serdaleh had higher levels and more diverse genes than those from Ambriss. Concerning  
537 the Labneh samples, only sample L18 presented detectable amounts of ABR. However, it cannot  
538 be concluded that L15, L16 and L17 samples did not contain ABR because there were not  
539 enough useful reads in these samples due to the presence of goat DNA.

540 Overall, the most abundantly detected antibiotic classes were aminoglycosides, MLS  
541 (Macrolide, Lincosamide, Streptogramin), nucleosides and tetracycline. In particular, Serdaleh  
542 samples contained high levels of aminoglycoside (*aad(6)* and APH(3'')-IIIa), MLS (*ermB*),  
543 nucleosides (*sat-4*) and tetracycline (*tetS*). The coverage of these genes is similar to that of *L.*  
544 *lactis*, which was the highly dominant species in these samples. Since this level of coverage  
545 allows a reliable level of assembly, we searched for contigs carrying these determinants by  
546 BLAST analysis in Serdaleh samples. This search pointed out two contigs of 12,360 and 3,370

547 nucleotides in sample S10-b, with coverages ~2 times that of *L. lactis*, and carrying, respectively,  
548 the four first determinants and *tetS*. The first contig is ~98% identical on 76-85% of its length to  
549 large plasmids from *E. faecium*, *E. faecalis* and *Staphylococcus pseudointermedius*. The second  
550 contig is identical on its full length to regions of plasmids and genome drafts of various food  
551 bacteria such as *L. raffinolactis*, *L. lactis*, *S. thermophilus*, *Carnobacterium divergens* and  
552 *Listeria innocua*. The other Serdaleh metagenomes contained contigs of comparable sizes  
553 indicating similar events.

554 The metagenome of Serdaleh S10-b also contained a ~6.6 kb contig carrying the *tetL*,  
555 *tetM* and *cat* genes that was identical on its entire length to regions of plasmids from *E. faecalis*.  
556 The coverage of this contig is 40 times lower compared to the two former ones, which is a level  
557 corresponding to that of *Lt. kefiri*, *Lb. kefiranofaciens* and *Lt. diolivorans*. None of the other  
558 metagenomes contained these three genes together on one contig. Nevertheless, samples S8, S7-  
559 b, S8-b, S9-b and S11-b contained contigs of ~5 and 0.9 kb carrying *tetLM* and *cat* genes,  
560 respectively. The coverages of these contigs were significantly different in the other assemblies  
561 where *tetLM* coverage was 2 to 10 times higher than that of *cat* genes, indicating that their  
562 organization might be different from that found in S10-b.

563 Lastly, in Ambriss A4, *tetM* was found on a ~32 kb contig of *S. parasuis*, *ermB* on a 7 kb  
564 contig sharing 90% identity on 3 kb with plasmids of various Lactobacillaceae, *cat* on a ~3 kb  
565 contig sharing 93% identity to *L. garviae* plasmids, *aad* and *sat* on a ~23 kb contig sharing 93%  
566 identity on ~12 kb (including *aad* and *sat*) of the *S. agalactiae* B111 chromosome and 99%  
567 identity on 4-6 kb of *E. faecalis* plasmids and the *S. pseudointermedius* chromosome.

## 568 **5 Discussion**

569 Traditional Lebanese fermented goat milk products are perceived in Lebanon as having  
570 an important historical and cultural heritage, but their microbiota are still poorly characterized. A  
571 detailed characterization of 50 producers allowed us to select those who used genuine traditional  
572 processes for the production of Ambriss, Serdaleh and Labneh El Darff. The manufacturing  
573 processes were established and the physico-chemical characteristics of the products were  
574 determined. Our survey also made it possible to define the socio-economic profiles of the  
575 producers, which confirmed the danger of extinction of these products, thus making it possible to  
576 require a Protected Designation of Origin label. This study explored the microbial biodiversity of

577 34 samples of three fermented goat milk products from three regions by combining  
578 microbiological, genomic and metagenomic approaches.

579 The samples analyzed here shared similar technical properties for pH, salinity,  
580 temperature and microbial counts as those of the same products previously studied (Dimassi et  
581 al., 2020; Serhan et al., 2016; Nasbimana et al., 2005; Semaan Hajj et al., 2011), and were  
582 therefore probably (Semaan Hajj et al., 2011)(Semaan Hajj et al., 2011)(Semaan Hajj et al.,  
583 2011)(Semaan Hajj et al., 2011) representative of these products in Lebanon.

#### 584 5.1 Microbial originality of traditional Lebanese fermented goat milk products

585 An overall view of the microbiota is provided by metagenetic analysis, which deciphers  
586 species or lower taxonomic levels. This study clearly separated Labneh samples from those of  
587 Ambriss and Serdaleh. Labneh contained a relatively simple microbiota, strongly dominated by  
588 *Lb. delbrueckii*, which was almost absent from the other products. Furthermore, it contained *Lt.*  
589 *diolivorans*, *Lt. kefiri* and *S. thermophilus* to a lesser extent. The dominance of *Lb. delbrueckii* is  
590 due to the way Labneh is inoculated by Laban, whose composition is that of yoghurt made of *Lb.*  
591 *delbrueckii* and *S. thermophilus* (Chammas et al., 2006). The dominance of *L. bulgaricus* is  
592 probably due to its better resistance to acidity than *S. thermophilus* (Béal & Chammas, 2012).  
593 Similar observations were made with yoghurt and Doogh (also known as Ayran), a traditional  
594 Iranian fermented ewe milk product (Davati & Hesami, 2021). The presence of *Lt. diolivorans*  
595 and *Lt. kefiri*, which are also found in Ambriss and Serdaleh, might be due to the specificity of  
596 the local goat milk. Moreover, the conditioning of Labneh in a goatskin may also influence the  
597 microbial composition by releasing nutrients or other metabolic products, as indicated by the  
598 very high level of goat DNA obtained from the metagenomic data for Labneh. Interestingly, *Lt.*  
599 *diolivorans* is usually found in vegetable and fruit fermentation, whereas it is occasionally  
600 reported in dairy products (Tamang et al. 2005; C. Zhang et al., 2010; Zheng et al., 2020). To the  
601 best of our knowledge, it was isolated from one raw milk in Brazil (Agostini et al., 2018),  
602 traditional yoghurts in Iran (RoushanZadeh et al., 2014), traditional fermented dairy products in  
603 Mongolia (Yu et al., 2011), and detected in metagenomic samples of Koumiss (You et al., 2022;  
604 Kozhakhmetov et al., 2014). The present study is thus the first one reporting the genome of *Lt.*  
605 *diolivorans* isolated from a dairy product.

606 Ambriss and Serdaleh samples contained, as dominant species, *L. lactis* and *Lb.*  
607 *kefiranofaciens*, which were almost absent from Labneh, followed by *Lt. kefiri* and, to a lesser

608 extent, by *Lt. diolivorans* and *Lp. plantarum*. Interestingly, this composition was reminiscent of  
609 that of dairy kefir, a fermented milk with health properties, which is produced by the inoculation  
610 of a stable and complex microbial community made of lactic acid bacteria and yeasts entrapped  
611 in gelatinous grains (Blasche et al., 2021; Prado et al., 2015; Nalbantoglu et al., 2014). In this  
612 product, the dominance of *Lb. kefiranofaciens*, which otherwise grows poorly alone in milk, is  
613 attributed to the metabolic cooperation it develops with the community, and the synthesis of an  
614 exopolysaccharide.

615 In this respect, growth conditions in Ambriss and Serdaleh, where the microbiota  
616 remained essentially in a coagulated milk matrix, displayed strong similarities with the one  
617 encountered in the kefir grain, which may thus explain the dominance of this particular  
618 *Lactobacillus* and of several sub-dominant lactic acid bacteria. MAGs and genome comparisons  
619 of *Lb. kefiranofaciens* and *Lt. kefir* confirm the very close phylogenomic relationship between  
620 kefir, Ambriss and Serdaleh strains. Their gene content mainly differed by unknown functions  
621 and mobile elements. In *Lb. kefiranofaciens* strains isolated from Ambriss and Serdaleh and in  
622 *Lb. kefirgranum* subspecies strains, only the first six genes encoding the general functions of  
623 initiating and controlling EPS synthesis were conserved, whereas glycosyl transferase conferring  
624 the specific structure of EPS units differed. This fact explains why these strains did not produce  
625 kefiran, the specific polysaccharide for kefir grain formation. Lastly, metabolic cooperation in  
626 Ambriss and Serdaleh certainly differed from that of kefir, as suggested by the low level of  
627 *Leuconostoc* and *Acetobacter* species, which play an important role in milk fermentation and  
628 citrate metabolism (Laëtitia et al., 2014). Nevertheless, the low input of milk on a the coagulated  
629 milk matrix gradually offers an environment for this “slow growing species” that is *Lb.*  
630 *kefiranofaciens* (Blasche et al., 2021; Georgalaki et al., 2021). Since the microbiota of Ambriss  
631 and Serdaleh were not inoculated, the environmental origin of *Lb. kefiranofaciens* in these  
632 products still remains unknown. Interestingly, this species was also found to be dominant in  
633 several samples of Airag, a traditional Mongolian fermented mare milk made in horse-hide  
634 containers (Watanabe et al., 2008) and in Koumiss (M. Zhang et al., 2020), neither of which  
635 contain grain.

636 In the present set of samples, the balance between *L. lactis* and *Lb. kefiranofaciens*  
637 appeared to be the main factor differentiating Ambriss and Serdaleh, the former generally  
638 containing lower levels of *L. lactis* and higher levels of *Lb. kefiranofaciens*. Nevertheless, the  
639 samples of these two similarly processed products, originating from two distinct regions, could

640 not be clearly separated since some samples of Serdaleh were also dominated by *Lb.*  
641 *kefiranofaciens*. Although Serdaleh samples were obtained at lower temperatures, displayed  
642 higher salt content, and reached lower pH in comparison to Ambriss samples, these factors do  
643 not fully explain this population shift. Alternatively, there may be subtle differences in  
644 production methods, such as the quantity of milk added at once, or a potential mixing that  
645 modifies the speed of milk infiltration through the product, which may affect the metabolism of  
646 the microbial communities in different ways, as discussed earlier. For example, if a significant  
647 quantity of the milk poured into the jar is not mixed with the curd, it may favor a rapid  
648 planktonic development of *L. lactis*, compared to the slow solid-state development where *Lb.*  
649 *kefiranofaciens* and *Lt. kefiri* compete more effectively.

650         A second important outcome of this study is the detection of uncultivated species as  
651 dominant microbiota in Ambriss and Serdaleh. While the initial failure to isolate *Lb.*  
652 *kefiranofaciens* from the samples can be explained by the fact that this species is tedious to grow  
653 and requires specific media and growth conditions (Georgalaki et al., 2021), the lack of recovery  
654 of *L. lactis* is less understandable. Indeed, this bacterium is routinely cultured from different  
655 environments such as plant or dairy products (Soundharrajan et al., 2021; Li et al., 2020). In this  
656 work, we thus tested the hypothesis that *L. lactis* cells were viable but not-cultivable (VBNC)  
657 (Oliver, 2010) and would thus require specific conditions to recover growth throughout their  
658 transfer from the product to M17 medium. Actually, a transfer of several hours in media  
659 containing goat milk allowed a straightforward isolation of strains of this species. This result  
660 indicates that the production process of Ambriss and Serdaleh induced a particular physiological  
661 state in *L. lactis* cells that hampers its growth recovery, thus corresponding to the definition of  
662 VBNC. Indeed, during the Ambriss and Serdaleh process, various environmental factors (such as  
663 pH, salinity and temperature) and the physiological age of the culture could induce the entry of  
664 bacterial cells into a VBNC state. Interestingly, the present process is still reversible, which is  
665 rarely the case (Dong et al., 2020), and hence offers a reproducible model for the study of the  
666 entry into the VBNC state. Further work would be required to understand why *L. lactis* growth  
667 recovery was deprived when directly plated on M17.

668

## 669 5.2 Potential safety concerns

670         A wide metagenomic search for bacteria potentially involved in public health issues  
671 indicated the presence of DNA of several kinds of undesirable species. Among hygiene

672 indicators, Enterobacteria were detected at low levels in about half of the samples, as well as *S.*  
673 *aureus*, which was detected at low levels, mostly in Serdaleh samples. These results are  
674 comparable to those of a recent study on raw cow milk where, from a total of 195 samples  
675 collected from different Lebanese farms, 65% and 17% of the samples were non-compliant with  
676 national standards for *Escherichia coli* and *S. aureus*, respectively (Joubrane et al., 2022).  
677 Enterobacteriaceae are often reported as a typical bacterial population in raw milk from many  
678 mammalian origins (Ramos & Nascimento, 2021). They can also be transmitted to milk and  
679 dairy products by fecal contamination during the milking process or because of poor hygiene  
680 conditions (Giraffa, 2003). *S. aureus* was highly prevalent in some products and was the most  
681 frequent contaminant of milk and goat milk products (Oliveira et al., 2020; Xing et al., 2016).  
682 The transmission of *S. aureus* to milk is due to its ability to form biofilms and to remain on  
683 surfaces and utensils, its presence being mainly attributed to poor handling during processing  
684 (Kümmel et al., 2016).

685         A striking finding was the presence of several zoonotic streptococcal species such as *S.*  
686 *parauberis* and *S. parasuis*, which were detected at 0.5-2% in several samples of Ambriss and  
687 Serdaleh, respectively, and even at 7 and 55% in Ambriss A4. *S. parasuis* is an important  
688 pathogen for humans and pigs (Nomoto et al., 2015) and was already isolated as a contaminant in  
689 raw camel milk. Intriguingly, in Ambriss A4, *S. parasuis* was strongly dominant over the  
690 traditional lactic acid bacteria, which accounted for only 10% of the population. The fact that the  
691 pH of this sample has a value similar to that of the other samples indicates that the strain of *S.*  
692 *parasuis* present in this sample is an acidifying strain. Indeed, our analysis revealed the presence  
693 of ten genes encoding a Lac-PTS system in its genome, similar to the one that allows efficient  
694 lactose utilization in several lactic acid bacteria such as *L. lactis* (Aleksandrak-Piekarczyk,  
695 2013). These genes were probably acquired by a horizontal gene transfer from *S. suis*.  
696 Adaptation to lactose utilization in milk has already been observed for other non-starter  
697 streptococci, in particular, *S. infantarius* in fermented dairy products from East and West Africa  
698 (Jans et al., 2013) and in artisanal cheese from Brazil (Kothe et al., 2021).

699         Serdaleh samples contained low levels (1-4%) of the goat pathogen *M. agalactiae*. This  
700 bacterium is one of the causal agents of contagious agalactia, a serious enzootic disease of small  
701 ruminants, which notably occurs in the Mediterranean Basin (Kumar et al., 2014). Once it has  
702 infected the animal, it disseminates through the blood to different mucosa, including the  
703 mammary gland, and can consequently be found in milk. *M. agalactiae* has already been

704 detected in artisanal cheeses via non-cultural techniques (Martín-Platero et al., 2009). Its  
705 presence here is, therefore, a sign that the herd in the Chouf region was infected with this  
706 pathogen.

707 Finally, transmissible antibiotic resistance genes, which are considered to be a growing  
708 threat to public health (Ghimpețeanu et al., 2022; Yadav & Kapley, 2021), were detected. In  
709 particular, Serdaleh samples contained a multi-resistance cassette carrying aad(6) and APH(3'')-  
710 IIIa, ermB and sat-4, as well as tetracycline resistance genes. Since these antibiotics are used in  
711 prophylaxis for *M. agalactiae*, it is likely that such treatment contaminated the milk and favored  
712 the emergence of antibiotic resistant lactic acid bacteria strains.

## 713 **6 Conclusion**

714 In this study, we analyzed three artisanal Lebanese fermented products (Ambriss,  
715 Serdaleh and Labneh El Darff) that are processed from the raw milk of the Black Baladi goat, a  
716 local breed. The process consists of the daily introduction of salted milk or Laban into a  
717 terracotta jar or a goatskin over several months, together with the draining of the whey, until the  
718 container is full. This production remains highly artisanal and offers opportunities for the  
719 empowerment of women, but it is in danger of extinction due to the low volumes that are  
720 produced, the age of the producers and the absence of marketing. In industrial products that are  
721 inoculated with starters or by backslopping, the microbiota rapidly develops in large batches of  
722 milk, whereas in the current products, it develops slowly during percolation of the milk that is  
723 added periodically. During batch growth, fast-growing microorganisms are highly favored,  
724 whereas in a dense community, slow-growing bacteria can compete more effectively, as is the  
725 case with the well-described kefir grain where *Lb. kefiranofaciens* clearly predominates despite  
726 its poor growth performance in milk. Interestingly, kefir is promoted as a probiotic beverage and  
727 *Lb. kefiranofaciens* is presented as a key factor in this attribute (Georgalaki et al., 2021).  
728 Although no studies have yet proven this, traditional Lebanese goat milk products are believed to  
729 deliver health benefits, and it would thus be interesting to further test this possibility. These  
730 products are also interesting as sources of biodiversity because they contain strains of *Lb.*  
731 *kefiranofaciens* and *Lt. kefiri*, which are genetically quite different from those present in kefir  
732 that are highly conserved.

733 Furthermore, progress has to be made to ensure the safety of these products from herd  
734 prophylaxis to milk chain management via a global action in order to avoid contamination by

735 potential pathogens and the spread of antibiotic resistance genes. Finally, this study also provides  
736 key elements related to the microbiology and microbiota functioning of these products that will  
737 facilitate the development of strategies to improve these processes and preserve their identity.

738

## 739 **7 Acknowledgements**

740 This work was supported by the PHC CEDRE program 2018 (French-Lebanese campus),  
741 project number 40316UJ), the Lebanese University Grant Research program, and the French  
742 Dairy Interbranch Organization (CNIEL, Centre National Interprofessionnel de l'Economie  
743 Laitière), Paris, France. The authors are grateful to the INRAE MIGALE bioinformatics facility  
744 (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi:  
745 10.15454/1.5572390655343293E12) for providing computing and storage resources.

746

## 747 **8 References**

- 748 Abbas, H., Fatma, H., Abd El-Gawad, M., & Enab, A. K. (2014). Physicochemical  
749 Characteristics of Goat's Milk. *Life Science*, *11*(1), 307–317.  
750 <http://journal.stainkudus.ac.id/index.php/equilibrium/article/view/1268/1127>
- 751 Abdalla, M. O. M., & Hussain, S. I. K. (2010). Enumeration and Identification of Microflora in  
752 Roub, A Sudanese Traditional Fermented Dairy Product. *British Journal of Dairy Sciences*,  
753 *1*(2), 30–33.
- 754 Agostini, C., Eckert, C., Vincenzi, A., Machado, B. L., Jordon, B. C., Kipper, J. P., Dullius, A.,  
755 Dullius, C. H., Lehn, D. N., Sperotto, R. A., Pozzobon, A., Granada, C. E., Maciel, M. J., &  
756 Volken de Souza, C. F. (2018). Characterization of technological and probiotic properties of  
757 indigenous *Lactobacillus* spp. from south Brazil. *3 Biotech*, *8*(11), 0.  
758 <https://doi.org/10.1007/s13205-018-1469-7>
- 759 Alanjary, M., Steinke, K., & Ziemert, N. (2019). AutoMLST: An automated web server for  
760 generating multi-locus species trees highlighting natural product potential. *Nucleic Acids*  
761 *Research*, *47*(W1), W276–W282. <https://doi.org/10.1093/nar/gkz282>
- 762 Aleksandrak-Piekarczyk, T. (2013). Lactose and  $\beta$ -Glucosides Metabolism and Its Regulation  
763 in *Lactococcus lactis*: A Review. *Lactic Acid Bacteria - R & D for Food, Health and*

- 764 *Livestock Purposes, May*. <https://doi.org/10.5772/50889>
- 765 Almeida, M., Hébert, A., Abraham, A. L., Rasmussen, S., Monnet, C., Pons, N., Delbès, C.,  
766 Loux, V., Batto, J. M., Leonard, P., Kennedy, S., Ehrlich, S. D., Pop, M., Montel, M. C.,  
767 Irlinger, F., & Renault, P. (2014). Construction of a dairy microbial genome catalog opens  
768 new perspectives for the metagenomic analysis of dairy fermented products. *BMC*  
769 *Genomics*, *15*(1). <https://doi.org/10.1186/1471-2164-15-1101>
- 770 Ayad, E., Nashat, S., El-Sadek, N., Metwaly, H., & El-Soda, M. (2004). Selection of wild lactic  
771 acid bacteria isolated from traditional Egyptian dairy products according to production and  
772 technological criteria. *Food Microbiology*, *21*(6), 715–725.  
773 <https://doi.org/10.1016/j.fm.2004.02.009>
- 774 Aziz, R. K., Devoid, S., Disz, T., Edwards, R. A., Henry, C. S., Olsen, G. J., Olson, R.,  
775 Overbeek, R., Parrello, B., Pusch, G. D., Stevens, R. L., Vonstein, V., & Xia, F. (2012).  
776 SEED Servers: High-Performance Access to the SEED Genomes, Annotations, and  
777 Metabolic Models. *PLoS ONE*, *7*(10), 1–10. <https://doi.org/10.1371/journal.pone.0048053>
- 778 Balaa, R. El, & Marie, M. (2008). Sustainability of the Lebanese small ruminant dairy products  
779 supply chain. *Security, July*, 6–10.
- 780 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V.  
781 M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi,  
782 N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome  
783 assembly algorithm and its applications to single-cell sequencing. *Journal of Computational*  
784 *Biology*, *19*(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>
- 785 Béal, C., & Chammas, G. (2012). Production of Laban. In editors. Hui YH, Özgül Evranuz E  
786 (Ed.), *Handbook of animal-based fermented food and beverage technology*. (2nd Editio, pp.  
787 181–212). CRC Press.
- 788 Beghini, F., McIver, L. J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan,  
789 A., Manghi, P., Scholz, M., Thomas, A. M., Valles-Colomer, M., Weingart, G., Zhang, Y.,  
790 Zolfo, M., Huttenhower, C., Franzosa, E. A., & Segata, N. (2021). Integrating taxonomic,  
791 functional, and strain-level profiling of diverse microbial communities with biobakery 3.  
792 *ELife*, *10*, 1–42. <https://doi.org/10.7554/eLife.65088>

- 793 Blasche, S., Kim, Y., Mars, R. A. T., Machado, D., Maansson, M., Kafkia, E., Milanese, A.,  
794 Zeller, G., Teusink, B., Nielsen, J., Benes, V., Neves, R., Sauer, U., & Patil, K. R. (2021).  
795 Metabolic cooperation and spatiotemporal niche partitioning in a kefir microbial  
796 community. *Nature Microbiology*, 6(2), 196–208. [https://doi.org/10.1038/s41564-020-](https://doi.org/10.1038/s41564-020-00816-5)  
797 00816-5
- 798 Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., Mills, D. A.,  
799 & Caporaso, J. G. (2013). Quality-filtering vastly improves diversity estimates from  
800 Illumina amplicon sequencing. *Nature Methods*, 10(1), 57–59.  
801 <https://doi.org/10.1038/nmeth.2276>
- 802 Boyazoglu, J., & Morand-Fehr, P. (2001). Mediterranean dairy sheep and goat products and their  
803 quality: A critical review. *Small Ruminant Research*, 40(1), 1–11.  
804 [https://doi.org/10.1016/S0921-4488\(00\)00203-0](https://doi.org/10.1016/S0921-4488(00)00203-0)
- 805 Chammas, G., Saliba, R., Corrieu, G., & Béal, C. (2006). Characterisation of lactic acid bacteria  
806 isolated from fermented milk “laban.” *International Journal of Food Microbiology*, 110(1),  
807 52–61. <https://doi.org/10.1016/j.ijfoodmicro.2006.01.043>
- 808 Chammas, I. G., Saliba, R., & Béal, C. (2006). Characterization of the fermented milk “Laban”  
809 with sensory analysis and instrumental measurements. *Journal of Food Science*, 71(2).  
810 <https://doi.org/10.1111/j.1365-2621.2006.tb08919.x>
- 811 Ciccarelli, F. D., Doerks, T., Von Mering, C., Creevey, C. J., Snel, B., & Bork, P. (2006).  
812 Toward automatic reconstruction of a highly resolved tree of life. *Science*, 311(5765),  
813 1283–1287. <https://doi.org/10.1126/science.1123061>
- 814 Clark, S., & Mora García, M. B. (2017). A 100-Year Review: Advances in goat milk research.  
815 *Journal of Dairy Science*, 100(12), 10026–10044. <https://doi.org/10.3168/jds.2017-13287>
- 816 Davati, N., & Hesami, S. (2021). 16S rRNA metagenomic analysis reveals significant changes of  
817 microbial compositions during fermentation from ewe milk to doogh with antimicrobial  
818 activity. *Food Biotechnology*, 35(3), 179–198.  
819 <https://doi.org/10.1080/08905436.2021.1939045>
- 820 Dib, H., Semaan, H., Mrad, R., Ayoub, J., Choueiry, L., Moussa, H., & Bitar, G. (2012).  
821 Identification et évaluation de l’effet probiotique des bactéries lactiques isolées dans des

- 822 fromages caprins traditionnels. *Lebanese Science Journal*, 13, 43–58.
- 823 Dimassi, O., Iskandarani, Y., Afram, M., Akiki, R., & Rached, M. (2020). Production and  
824 physicochemical properties of labneh anbaris, a traditional fermented cheese like product, in  
825 Lebanon. *International Journal of Environment, Agriculture and Biotechnology*, 5(3), 509–  
826 516. <https://doi.org/10.22161/ijeab.53.3>
- 827 Dong, K., Pan, H., Yang, D., Rao, L., Zhao, L., Wang, Y., & Liao, X. (2020). Induction,  
828 detection, formation, and resuscitation of viable but non-culturable state microorganisms.  
829 *Comprehensive Reviews in Food Science and Food Safety*, 19(1), 149–183.  
830 <https://doi.org/10.1111/1541-4337.12513>
- 831 Dubeuf, J., Aw-Hassan, A., Chentouf, M., Mena, Y., Pacheco, F., & Boutonnet, J. (2016). The  
832 Mediterranean sheep and goat sectors between constants and changes over the last decade  
833 future challenges and prospects. *Options Méditerranéennes. Série a, Séminaires*  
834 *Méditerranéens*, 115, 43–52.
- 835 El Soda, M., Ahmed, N., Omran, N., Osman, G., & Morsi, A. (2003). Isolation, identification  
836 and selection of lactic acid bacteria cultures for cheesemaking. *Emirates Journal of Food*  
837 *and Agriculture*, 15(2), 51–71. <https://doi.org/10.1016/j.numecd.2011.02.008>
- 838 Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., Maman, S.,  
839 Hernandez-Raquet, G., Combes, S., & Pascal, G. (2018). FROGS: Find, Rapidly, OTUs  
840 with Galaxy Solution. *Bioinformatics*, 34(8), 1287–1294.  
841 <https://doi.org/10.1093/bioinformatics/btx791>
- 842 Gardner, S. N., Slezak, T., & Hall, B. G. (2015). kSNP3.0: SNP detection and phylogenetic  
843 analysis of genomes without genome alignment or reference genome. *Bioinformatics*,  
844 31(17), 2877–2878. <https://doi.org/10.1093/bioinformatics/btv271>
- 845 Georgalaki, M., Zoumpopoulou, G., Anastasiou, R., Kazou, M., & Tsakalidou, E. (2021).  
846 *Lactobacillus kefiranofaciens*: From isolation and taxonomy to probiotic properties and  
847 applications. *Microorganisms*, 9(10). <https://doi.org/10.3390/microorganisms9102158>
- 848 Ghimpețeanu, O. M., Pogurschi, E. N., Popa, D. C., Dragomir, N., Drăgotoiu, T., Mihai, O. D.,  
849 & Petcu, C. D. (2022). Antibiotic Use in Livestock and Residues in Food—A Public Health  
850 Threat: A Review. In *Foods* (Vol. 11, Issue 10). <https://doi.org/10.3390/foods11101430>

- 851 Giraffa, G. (2003). Functionality of enterococci in dairy products. *International Journal of Food*  
852 *Microbiology*, 88(2–3), 215–222. [https://doi.org/10.1016/S0168-1605\(03\)00183-1](https://doi.org/10.1016/S0168-1605(03)00183-1)
- 853 Hosri, C., & El Khoury, N. (2004). Valoriser le fromage de chèvre traditionnel “ Darfiyeh ” pour  
854 aider au developpement de la région montagnarde nord libanaise. *Options*  
855 *Méditerranéennes*, 61(A), 201–206.
- 856 IDEAS Centre. (2005). *Proposal for the protection of Geographical indication in Lebanon*  
857 *Report on the intention of IDEAS Centre 1204 Geneva Switzerland Beirut , 7-11 February*  
858 *2005* (Issue February).
- 859 Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High  
860 throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries.  
861 *Nature Communications*, 9(1), 1–8. <https://doi.org/10.1038/s41467-018-07641-9>
- 862 Jans, C., Kaindi, D. W. M., Böck, D., Njage, P. M. K., Kouamé-Sina, S. M., Bonfoh, B.,  
863 Lacroix, C., & Meile, L. (2013). Prevalence and comparison of *Streptococcus infantarius*  
864 subsp. *infantarius* and *Streptococcus gallolyticus* subsp. *macedonicus* in raw and fermented  
865 dairy products from East and West Africa. *International Journal of Food Microbiology*,  
866 *167*(2), 186–195. <https://doi.org/10.1016/j.ijfoodmicro.2013.09.008>
- 867 Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., Lago, B. A., Dave,  
868 B. M., Pereira, S., Sharma, A. N., Doshi, S., Courtot, M., Lo, R., Williams, L. E., Frye, J.  
869 G., Elsayegh, T., Sardar, D., Westman, E. L., Pawlowski, A. C., ... McArthur, A. G. (2017).  
870 CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic  
871 resistance database. *Nucleic Acids Research*, 45(D1), D566–D573.  
872 <https://doi.org/10.1093/nar/gkw1004>
- 873 Joubrane, K., Jammoul, A., Daher, R., Ayoub, S., & El, M. (2022). Microbiological  
874 contamination , antimicrobial residues , and antimicrobial resistance in raw bovine milk in  
875 Lebanon. *International Dairy Journal*, 134(July), 105455.  
876 <https://doi.org/10.1016/j.idairyj.2022.105455>
- 877 Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately  
878 reconstructing single genomes from complex microbial communities. *PeerJ*, 2015(8), 1–15.  
879 <https://doi.org/10.7717/peerj.1165>

- 880 Kayouli, C. (2015). LEBANON RECOVERY FUND Recovery and Rehabilitation of the Dairy  
881 Sector in Bekâa Valley and Hermel-Akkar Uplands1-FINAL PROGRAMME- LRF-21  
882 OSRO/LEB/901/UNJ –. In *United Nation* (Vol. 53, Issue 9).  
883 <https://doi.org/10.1017/CBO9781107415324.004>
- 884 Kothe, C. I., Bolotin, A., Kraïem, B. F., Dridi, B., & Renault, P. (2021). Unraveling the world of  
885 halophilic and halotolerant bacteria in cheese by combining cultural, genomic and  
886 metagenomic approaches. *International Journal of Food Microbiology*, 358(November  
887 2020). <https://doi.org/10.1016/j.ijfoodmicro.2021.109312>
- 888 Kozhakhmetov, S., Tynybayeva, I., Baikhanova, D., Saduakhasova, S., Shakhabayeva, G.,  
889 Kushugulova, A., Nurgozhin, T., & Zhumadilov, Z. (2014). Metagenomic Analysis of  
890 Koumiss in Kazakhstan. *Central Asian Journal of Global Health*, 3(February 2016), 8–10.  
891 <https://doi.org/10.5195/cajgh.2014.163>
- 892 Kumar, A., Rahal, A., Chakraborty, S., Verma, A. K., & Dhama, K. (2014). *Mycoplasma*  
893 *agalactiae*, an etiological agent of contagious agalactia in Small ruminants: A review.  
894 *Veterinary Medicine International*, 2014, 7–9. <https://doi.org/10.1155/2014/286752>
- 895 Kümmel, J., Stessl, B., Gonano, M., Walcher, G., Bereuter, O., Fricker, M., Grunert, T., Wagner,  
896 M., & Ehling-Schulz, M. (2016). *Staphylococcus aureus* entrance into the Dairy Chain:  
897 Tracking *S. aureus* from dairy cow to cheese. *Frontiers in Microbiology*, 7(OCT), 1–11.  
898 <https://doi.org/10.3389/fmicb.2016.01603>
- 899 Laëtitia, G., Pascal, D., & Yann, D. (2014). The Citrate Metabolism in Homo- and  
900 Heterofermentative LAB: A Selective Means of Becoming Dominant over Other  
901 Microorganisms in Complex Ecosystems. *Food and Nutrition Sciences*, 05(10), 953–969.  
902 <https://doi.org/10.4236/fns.2014.510106>
- 903 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature*  
904 *Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- 905 Leech, J., Cabrera-Rubio, R., Walsh, A. M., Macori, G., Walsh, C. J., Barton, W., Finnegan, L.,  
906 Crispie, F., O’Sullivan, O., Claesson, M. J., & Cotter, P. D. (2020). Fermented-Food  
907 Metagenomics Reveals Substrate-Associated Differences in Taxonomy and Health-  
908 Associated and Antibiotic Resistance Determinants. *MSystems*, 5(6), 1–21.  
909 <https://doi.org/10.1128/msystems.00522-20>

910 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., &  
911 Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*,  
912 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>

913 Li, W., Ren, M., Duo, L., Li, J., Wang, S., Sun, Y., Li, M., Ren, W., Hou, Q., Yu, J., Sun, Z., &  
914 Sun, T. (2020). Fermentation Characteristics of *Lactococcus lactis* subsp. *lactis* Isolated  
915 From Naturally Fermented Dairy Products and Screening of Potential Starter Isolates.  
916 *Frontiers in Microbiology*, 11(August), 1–12. <https://doi.org/10.3389/fmicb.2020.01794>

917 Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and  
918 dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 1–21.  
919 <https://doi.org/10.1186/s13059-014-0550-8>

920 Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: Robust and  
921 fast clustering method for amplicon-based studies. *PeerJ*, 2014(1), 1–13.  
922 <https://doi.org/10.7717/peerj.593>

923 Martín-Platero, A. M., Maqueda, M., Valdivia, E., Purswani, J., & Martínez-Bueno, M. (2009).  
924 Polyphasic study of microbial communities of two Spanish farmhouse goats' milk cheeses  
925 from Sierra de Aracena. *Food Microbiology*, 26(3), 294–304.  
926 <https://doi.org/10.1016/j.fm.2008.12.004>

927 Massaad, B. (2017). *Mouneh: Preserving Foods for the Lebanese Pantry* (1rst ed.). Interlink  
928 Books.

929 McArthur, A. G., Wagglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., Bhullar, K.,  
930 Canova, M. J., De Pascale, G., Ejim, L., Kalan, L., King, A. M., Koteva, K., Morar, M.,  
931 Mulvey, M. R., O'Brien, J. S., Pawlowski, A. C., Piddock, L. J. V., Spanogiannopoulos, P.,  
932 ... Wright, G. D. (2013). The comprehensive antibiotic resistance database. *Antimicrobial  
933 Agents and Chemotherapy*, 57(7), 3348–3357. <https://doi.org/10.1128/AAC.00419-13>

934 Ministre de l'Economie et du Commerce Republique Libanaise. (2006). Qu'est-ce qu'une  
935 Indication Géographique (IG)? *Bulletin Des Indications Géographiques*, 1–6.

936 Nalbantoglu, U., Cakar, A., Dogan, H., Abaci, N., Ustek, D., Sayood, K., & Can, H. (2014).  
937 Metagenomic analysis of the microbial community in kefir grains. *Food Microbiology*, 41,  
938 42–51. <https://doi.org/10.1016/j.fm.2014.01.014>

- 939 Nomoto, R., Maruyama, F., Ishida, S., Tohya, M., Sekizaki, T., & Osawa, R. (2015). Reappraisal  
940 of the taxonomy of streptococcus suis serotypes 20, 22 and 26: *Streptococcus parasuis* sp.  
941 nov. *International Journal of Systematic and Evolutionary Microbiology*, *65*(2), 438–443.  
942 <https://doi.org/10.1099/ijs.0.067116-0>
- 943 Oliveira, A. P. D. de, Costa, M. M. da, Nogueira, D. M., & Dias, F. S. (2020). Characterisation  
944 of *Staphylococcus aureus* strains from milk and goat cheese and evaluation of their  
945 inhibition by gallic acid, nisin and velame of the Brazilian caatinga. *International Journal*  
946 *of Dairy Technology*, *73*(2), 345–356. <https://doi.org/10.1111/1471-0307.12673>
- 947 Oliver, J. D. (2010). Recent findings on the viable but nonculturable state in pathogenic bacteria.  
948 *FEMS Microbiology Reviews*, *34*(4), 415–425. [https://doi.org/10.1111/j.1574-](https://doi.org/10.1111/j.1574-6976.2009.00200.x)  
949 [6976.2009.00200.x](https://doi.org/10.1111/j.1574-6976.2009.00200.x)
- 950 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM:  
951 Assessing the quality of microbial genomes recovered from isolates, single cells, and  
952 metagenomes. *Genome Research*, *25*(7), 1043–1055. <https://doi.org/10.1101/gr.186072.114>
- 953 Prado, M. R., Blandón, L. M., Vandenberghe, L. P. S., Rodrigues, C., Castro, G. R., Thomaz-  
954 Soccol, V., & Soccol, C. R. (2015). Milk kefir: Composition, microbial cultures, biological  
955 activities, and related products. *Frontiers in Microbiology*, *6*(OCT), 1–10.  
956 <https://doi.org/10.3389/fmicb.2015.01177>
- 957 Prosser, C. G. (2021). Compositional and functional characteristics of goat milk and relevance as  
958 a base for infant formula. *Journal of Food Science*, *86*(2), 257–265.  
959 <https://doi.org/10.1111/1750-3841.15574>
- 960 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F.  
961 O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing  
962 and web-based tools. *Nucleic Acids Research*, *41*(D1), 590–596.  
963 <https://doi.org/10.1093/nar/gks1219>
- 964 Ramos, G. L. de P. A., & Nascimento, J. D. S. (2021). Antibiotic resistance profile and detection  
965 of degradative enzymes by enterobacteriaceae isolated from raw goat milk. *Germs*, *11*(2),  
966 211–220. <https://doi.org/10.18683/germs.2021.1258>
- 967 Rhaiem, N., Chahboun, N., Inekach, S., & Ouhssine, M. (2016). Identification and

- 968 characterization of lactic acid bacteria isolated from cow milk and olives brine. *Journal of*  
969 *Materials and Environmental Science*, 7(5), 1504–1509.  
970 <https://doi.org/10.3923/pjbs.2006.1242.1249>
- 971 Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open  
972 source tool for metagenomics. *PeerJ*, 2016(10), 1–22. <https://doi.org/10.7717/peerj.2584>
- 973 RoushanZadeh, S., Eskandari, M. H., Shekarforoush, S. S., & Hosseini, A. (2014). Phenotypic  
974 and genotypic diversity of dominant lactic acid bacteria isolated from traditional yoghurts  
975 produced by tribes of Iran. *Iranian Journal of Veterinary Research*, 15(4), 347–352.  
976 <https://doi.org/10.22099/ijvr.2014.2590>
- 977 Saadeh, L. (2016). Composition of Dairy Potential Output per Year Lebanese Dairy Sector. In  
978 *Blominvest Bank*.
- 979 Salameh, C., Banon, S., Hosri, C., & Scher, J. (2016). An overview of recent studies on the main  
980 traditional fermented milks and white cheeses in the Mediterranean region. *Food Reviews*  
981 *International*, 32(3), 256–279. <https://doi.org/10.1080/87559129.2015.1075210>
- 982 Semaan, E. H., Dib, H., Abi Ramia, R., & Chedid, M. (2011). Caractérisation chimique et  
983 Qualité Bactériologique de Produits Laitiers Caprins Traditionnels Libanais. *Lebanese*  
984 *Science Journal*, 12(1), 21–29.
- 985 Semaan Hajj, E., Dib, H., Abi Ramia, R., & Chedid, C. (2011). Characterisation Chimique Et  
986 Quaalité Bactériologique De Produits Laitiers. *Lebanese Science Journal*, 12(1), 21–29.
- 987 Serhan, M., Cailliez-Grimal, C., Borges, F., Revol-Junelles, A. M., Hosri, C., & Fanni, J. (2009).  
988 Bacterial diversity of Darfiyeh, a Lebanese artisanal raw goat's milk cheese. *Food*  
989 *Microbiology*, 26(6), 645–652. <https://doi.org/10.1016/j.fm.2009.04.012>
- 990 Serhan, M., & Mattar, J. (2013). Characterization of four lebanese artisanal goat milk cheeses:  
991 Darfiyeh, Aricheh, shankleesh and Serdale by physico-chemical, microbiological and  
992 sensory analyses. *Journal of Food, Agriculture and Environment*, 11(3–4), 97–101.
- 993 Serhan, M., & Mattar, J. (2018). The Goat Dairy Sector in Lebanon. *Goat Science*.  
994 <https://doi.org/10.5772/intechopen.70077>
- 995 Serhan, M., Mattar, J., & Debs, L. (2016). Concentrated yogurt (Labneh) made of a mixture of  
996 goats' and cows' milk: Physicochemical, microbiological and sensory analysis. *Small*

- 997 *Ruminant Research*, 138, 46–52. <https://doi.org/10.1016/j.smallrumres.2016.04.003>
- 998 Shu, G., Bao, C., Chen, H., Wang, C., & Yang, H. (2016). Fermentation optimization of goat  
999 milk with *Lactobacillus acidophilus* and *Bifidobacterium bifidum* by Box-Behnken design.  
1000 *Acta Scientiarum Polonorum, Technologia Alimentaria*, 15(2), 151–159.  
1001 <https://doi.org/10.17306/J.AFS.2016.2.15>
- 1002 Soundharrajan, I., Yoon, Y. H., Muthusamy, K., Jung, J. S., Lee, H. J., Han, O. K., & Choi, K. C.  
1003 (2021). Isolation of *Lactococcus lactis* from whole crop rice and determining its probiotic  
1004 and antimicrobial properties towards gastrointestinal associated bacteria. *Microorganisms*,  
1005 9(12). <https://doi.org/10.3390/microorganisms9122513>
- 1006 Sunagawa, S., Mende, D. R., Zeller, G., Izquierdo-Carrasco, F., Berger, S. A., Kultima, J. R.,  
1007 Coelho, L. P., Arumugam, M., Tap, J., Nielsen, H. B., Rasmussen, S., Brunak, S., Pedersen,  
1008 O., Guarner, F., De Vos, W. M., Wang, J., Li, J., Doré, J., Dusko Ehrlich, S., ... Bork, P.  
1009 (2013). Metagenomic species profiling using universal phylogenetic marker genes. *Nature*  
1010 *Methods*, 10(12), 1196–1199. <https://doi.org/10.1038/nmeth.2693>
- 1011 Tabet, E., Salameh, C., Nehme, M., Daher, J., & Hosri, C. (2019). Innovation aspects of Serdaleh  
1012 , a traditional Lebanese cheese produced from raw extensive goat ' s milk. *Innovations for*  
1013 *Sustainability in Sheep and Goats.Options Méditerranéennes A*, 123, 299–304.
- 1014 Toufeili, I., Shadarevian, S., Artinian, T., & Tannous, R. (1995). Ripening changes and sensory  
1015 properties of bovine, caprine and ovine shankleesh. *International Dairy Journal*, 5(2), 179–  
1016 189. [https://doi.org/10.1016/0958-6946\(95\)92209-M](https://doi.org/10.1016/0958-6946(95)92209-M)
- 1017 Truong, D., Franzosam, E., & Timothy, L. (2015). MetaPhlan2 for enhanced metagenomic  
1018 taxonomic profiling. *Nature Publishing Group*, 12(10), 902–903.  
1019 <https://doi.org/10.1038/nmeth.3589>
- 1020 Walsh, L. H., Coakley, M., Walsh, A. M., O'Toole, P. W., & Cotter, P. D. (2022). Bioinformatic  
1021 approaches for studying the microbiome of fermented food. *Critical Reviews in*  
1022 *Microbiology*, 0(0), 1–33. <https://doi.org/10.1080/1040841X.2022.2132850>
- 1023 Watanabe, K., Fujimoto, J., Sasamoto, M., Dugersuren, J., Tumursuh, T., & Demberel, S.  
1024 (2008). Diversity of lactic acid bacteria and yeasts in Airag and Tarag, traditional fermented  
1025 milk products of Mongolia. *World Journal of Microbiology and Biotechnology*, 24(8),

- 1026 1313–1325. <https://doi.org/10.1007/s11274-007-9604-3>
- 1027 Xing, X., Zhang, Y., Wu, Q., Wang, X., Ge, W., & Wu, C. (2016). Prevalence and  
1028 characterization of *Staphylococcus aureus* isolated from goat milk powder processing  
1029 plants. *Food Control*, *59*, 644–650. <https://doi.org/10.1016/j.foodcont.2015.06.042>
- 1030 Yadav, S., & Kapley, A. (2021). Antibiotic resistance: Global health crisis and metagenomics.  
1031 *Biotechnology Reports*, *29*, e00604. <https://doi.org/10.1016/j.btre.2021.e00604>
- 1032 You, L., Yang, C., Jin, H., Kwok, L. Y., Sun, Z., & Zhang, H. (2022). Metagenomic features of  
1033 traditional fermented milk products. *Lwt*, *155*(December 2021), 112945.  
1034 <https://doi.org/10.1016/j.lwt.2021.112945>
- 1035 Yu, J., Wang, W. H., Menghe, B. L. G., Jiri, M. T., Wang, H. M., Liu, W. J., Bao, Q. H., Lu, Q.,  
1036 Zhang, J. C., Wang, F., Xu, H. Y., Sun, T. S., & Zhang, H. P. (2011). Diversity of lactic  
1037 acid bacteria associated with traditional fermented dairy products in Mongolia. *Journal of*  
1038 *Dairy Science*, *94*(7), 3229–3241. <https://doi.org/10.3168/jds.2010-3727>
- 1039 Zenebe, T., Ahmed, N., Kabeta, T., & Kebede, G. (2014). Review on Medicinal and Nutritional  
1040 Values of Goat Milk. *Academic Journal of Nutrition*, *3*(3), 30–39.  
1041 <https://doi.org/10.5829/idosi.ajn.2014.3.3.93210>
- 1042 Zhang, C., Brandt, M. J., Schwab, C., & Gänzle, M. G. (2010). Propionic acid production by  
1043 cofermentation of *Lactobacillus buchneri* and *Lactobacillus diolivorans* in sourdough. *Food*  
1044 *Microbiology*, *27*(3), 390–395. <https://doi.org/10.1016/j.fm.2009.11.019>
- 1045 Zhang, M., Dang, N., Ren, D., Zhao, F., Lv, R., Ma, T., Bao, Q., Menghe, B., & Liu, W. (2020).  
1046 Comparison of Bacterial Microbiota in Raw Mare’s Milk and Koumiss Using PacBio Single  
1047 Molecule Real-Time Sequencing Technology. *Frontiers in Microbiology*, *11*(October), 1–  
1048 13. <https://doi.org/10.3389/fmicb.2020.581610>
- 1049 Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., O’toole,  
1050 P. W., Pot, B., Vandamme, P., Walter, J., Watanabe, K., Wuyts, S., Felis, G. E., Gänzle, M.  
1051 G., & Lebeer, S. (2020). A taxonomic note on the genus *Lactobacillus*: Description of 23  
1052 novel genera, emended description of the genus *Lactobacillus* beijerinck 1901, and union of  
1053 *Lactobacillaceae* and *Leuconostocaceae*. *International Journal of Systematic and*  
1054 *Evolutionary Microbiology*, *70*(4), 2782–2858. <https://doi.org/10.1099/ijsem.0.004107>

1055 Zúñiga, M., Pardo, I., & Ferrer, S. (1993). An improved medium for distinguishing between  
1056 homofermentative and heterofermentative lactic acid bacteria. *International Journal of*  
1057 *Food Microbiology*, 18(1), 37–42. [https://doi.org/10.1016/0168-1605\(93\)90005-2](https://doi.org/10.1016/0168-1605(93)90005-2)

## 1058 **9 Figure legends**

1059 **Figure 1.** Geographical origin (A) and schematic processing diagram (B) of Ambriss, Serdaleh  
1060 and Labneh El Darff. Ambriss and Serdaleh are made in terracotta jars periodically filled with  
1061 raw goat's milk, and Labneh El Darf in goatskin filled with Laban, which is either a homemade  
1062 yoghurt of raw goat milk. The detailed flow charts are presented in Figure S1 and S2.

1063 **Figure 2.** Physico-chemical and microbiological characteristics of Ambriss, Serdaleh and  
1064 Labneh El Darff: Temperature, pH and salinity (A); boxplot representation of microbial counts  
1065 on M17 and MRS (B); and histogram of isolated species (C).

1066 **Figure 3.** Metagenetic analysis based on OTU data of Ambriss, Serdaleh and Labneh El Darff:  
1067 Boxplots of alpha-diversity indices (A); Principal coordinate analysis among bacterial  
1068 communities identified by OTUs (B); bacterial composition plot depicting the relative  
1069 abundances of the 20 main OTUs (C). Presence of different letters (a, b, c) above the box-plots  
1070 indicate their statistical significance compared to the others

1071 Colors in (A) and (B) refer to the different fermented milk products, i.e., yellow for Ambriss,  
1072 blue for Serdaleh and red for Labneh El Darff.

1073 **Figure 4.** Relative abundance of microbial composition of Ambriss, Serdaleh and Labneh El  
1074 Darff.

1075 Values were calculated from the coverage of the *ychF* marker gene assembled in the  
1076 metagenomes. The asterisk (\*) indicates a low depth of the sequences for L15, L16 and L17  
1077 samples.

1078 **Figure 5.** Phylogenomic tree for *L. lactis* (A), *Lb. kefiranofaciens* (B) and *M. agalactiae* (C).

1079 The different groups of *L. lactis* are highlighted. The colors of the branches make it possible to  
1080 visualize cremoris subspecies (brown), environmental sp. lactis strains (green), industrial dairy  
1081 sp. lactis (red), *diacetylactis* biovariant (blue) and *lactis* sp from traditional dairy products  
1082 (violet). The internal circle presents the environment from which the strains were isolated (blue:

1083 milk; green: vegetable; purple: animal). Ambriss and Serdaleh MAGs and the isolated strain  
1084 genome are highlighted by external red arrows and names are shaded in gray

1085

1086 **Figure 6.** Heatmap showing distribution of the 18 antibiotic resistance genes detected in reads/kb  
1087 within the 13 metagenomic samples.

1088 MLS: Macrolide, Lincosamide and Streptogramin.

## 1089 **10 Supplementary Material**

1090 **Figure S1:** Flow chart of the Ambriss and Serdaleh manufacturing process

1091 **Figure S2:** Flow chart of the Labneh El Darff manufacturing process

1092 **Figure S3.** Rarefaction curves depicting the depth of 16S sequencing and species richness for the  
1093 data obtained from Ambriss, Serdaleh and Labneh El Darff.

1094 The x-axis represents the sequencing depth (reads) and the y-axis represents the estimated OTU  
1095 richness detected at the species level.

1096 **Figure S4:** Phylogenomic tree for *Lb. kefirano*, *Lt. kefir* (A), *Lt. diolivorans* (B) and *Lp.*  
1097 *plantarum* (C).

1098 **Figure S5:** Heatmap showing the presence of potentially health-associated gene clusters across  
1099 the Lebanese fermented goat milk product samples. Gene clusters are binned according to the  
1100 potential ability of the microbiota to colonize the gastrointestinal tract, survive transit to the gut,  
1101 and modulate the host phenotype. Each line is standardized across all samples, allowing only  
1102 products to be compared with each other.

1103 **Table S1:** Producer survey results of production parameters, rearing practices and price  
1104 indicators related to Ambriss, Serdaleh and Labneh El Darff.

1105 **Table S2:** Physico-chemical parameters and culture counts of Ambriss, Serdaleh and Labneh El  
1106 Darff samples as a function of their producers and production dates.

1107 **Table S3.** Percentage (%) of bacterial operational taxonomic units (OTUs) in the samples.

1108 **Table S4.** Percentage (%) of *Capra hircus* reads aligned in each fermented food metagenome.

1109 **Table S5.** Relative abundance of microbial composition of 16 traditional fermented goat milk  
1110 metagenomes from Lebanon using the MetaPhlAn taxonomic assignment method.

1111 **Table S6.** Relative abundance of microbial composition of 16 traditional fermented goat milk  
1112 metagenomes from Lebanon using *ychF* marker gene assignment.

1113 **Table S7.** Relative abundance of microbial composition of 16 traditional fermented goat milk  
1114 metagenomes from Lebanon using the *FoodMicrobiome* tool.

1115 **Table S8.** Quality of prokaryotic MAGs

1116 **Table S9.** Genomes sequenced and assembly statistics

1117 **Table S10.** FastANI values for MAGs, isolated strain genomes and NCBI reference genomes for  
1118 *L. lactis*, *Lb. kefiranofaciens*, *Lt. kefir*, *Lt. diolivorans*, *L. plantarum*, *M. agalactiae* and *S.*  
1119 *parasuis*.

1120 **Table S11.** Bacteriocin profile (in gene counts) of 16 traditional fermented goat milk  
1121 metagenomes from Lebanon.

1122