

<b>Title</b>	Impacts of seasonal housing and teat preparation on raw milk microbiota: a high-throughput sequencing study
<b>Author(s)</b>	Doyle, Conor J.; Gleeson, David; O'Toole, Paul W.; Cotter, Paul D.
<b>Publication date</b>	2016-11-04
<b>Original citation</b>	Doyle, C. J., Gleeson, D., O'Toole, P. W. and Cotter, P. D. (2017) 'Impacts of Seasonal Housing and Teat Preparation on Raw Milk Microbiota: a High-Throughput Sequencing Study', Applied and Environmental Microbiology, 83(2). doi:10.1128/aem.02694-16
<b>Type of publication</b>	Article (peer-reviewed)
<b>Link to publisher's version</b>	<a href="http://dx.doi.org/10.1128/aem.02694-16">http://dx.doi.org/10.1128/aem.02694-16</a> Access to the full text of the published version may require a subscription.
<b>Rights</b>	© 2016 American Society for Microbiology. All Rights Reserved. This is the Accepted manuscript version of the article published in its final form in Appl. Environ. Microbiol. doi: 10.1128/AEM.02694-16
<b>Item downloaded from</b>	<a href="http://hdl.handle.net/10468/3995">http://hdl.handle.net/10468/3995</a>

Downloaded on 2019-01-23T22:42:00Z

1 **High-throughput sequencing highlights the significant influence of seasonal**  
2 **housing and teat preparation on the raw milk microbiota**

3 Conor J Doyle<sup>ab</sup>, David Gleeson<sup>c</sup>, Paul W O'Toole<sup>bd</sup>, Paul D Cotter<sup>ad\*</sup>

4

5 a Teagasc Food Research Centre, Moorepark, Fermoy, Cork, Ireland

6 b School of Microbiology, University College Cork, Cork, Ireland

7 c Teagasc Animal and Grassland Research and Innovation Centre, Moorepark, Fermoy, Cork,  
8 Ireland

9 d APC Microbiome Institute, Cork, Ireland

10

11 **Running title:**

12 Factors influencing the raw milk microbiota

13

14 **For correspondence.**

15 \*[paul.cotter@teagasc.ie](mailto:paul.cotter@teagasc.ie); Tel. (+353) 25 42694.

16

17 **Abstract**

18 In pasture-based systems, changes in dairy herd habitat due to seasonality results in the  
19 exposure of animals to different environmental niches. These niches contain distinct  
20 microbial communities that may be transferred to raw milk, with potentially important food  
21 quality and safety implications for milk producers. It is postulated that the extent to which  
22 these microorganisms are transferred could be limited by the inclusion of a teat preparation  
23 step prior to milking. Here high-throughput sequencing, of a variety of microbial niches on  
24 the farm, is employed to study the patterns of microbial movement through the dairy  
25 production chain and, in the process, investigate the impact of seasonal housing and  
26 inclusion/exclusion of teat preparation regime on the raw milk microbiota from the same  
27 herd over two sampling periods, i.e., indoor and outdoor. Beta diversity and network  
28 analyses showed that environmental and milk microbiotas separated depending on whether  
29 they were sourced from an indoor or outdoor environment. Within these respective  
30 habitats, similarities between the milk microbiota and that of teat swab samples and, to a  
31 lesser extent, faecal samples were apparent. Indeed, SourceTracker identified the teat  
32 surface as the most significant source of contamination, with herd faeces being the next  
33 most prevalent source of contamination. In milk from cows grazing outdoors, teat prep  
34 significantly increased the numbers of total bacteria present. In summary, sequence-based  
35 microbiota analysis identified possible sources of raw milk contamination, and highlighted  
36 the influence of environment and farm management practices on the raw milk microbiota.

37

38

39 **Importance**

40 The composition of the raw milk microbiota is an important consideration from both a  
41 spoilage and food safety perspective and has implications for milk targeted for direct  
42 consumption and for downstream processing. Factors which influence contamination have  
43 been examined previously, primarily through the use of culture-based techniques. This  
44 manuscript describes the extensive application of high throughput DNA sequencing  
45 technologies to study the relationship between the milk production environment and the  
46 raw milk microbiota. Results highlight that the environment in which the herd was kept was  
47 the primary driver of the composition of the milk microbiota composition.

48

49 **Introduction**

50 The impact of the dairy farm environment on the microbial composition of raw milk and raw  
51 milk products has been appreciated for some time (1). There are numerous niches that  
52 collectively constitute the dairy farm environment and these harbour a vast array of  
53 microbes. The transfer of microbes from the farm environment to raw milk can be  
54 influenced by a number of factors including farmer hygiene, husbandry practices, herd  
55 health, and herd housing (2). In turn, the microbial composition of raw milk is critically  
56 important to its quality, processability and safety.

57 The microbiota composition of dairy farm niches and of raw milk has typically been  
58 examined using traditional plate cultivation-based techniques. These culture-based assays  
59 are still widely used by industry and target specific phenotypes, e.g. ability to grow at or  
60 survive exposure to particular temperatures (psychrotrophs (3), mesophiles (4),  
61 thermotolerants (5), or capacity to produce proteases, lipases or other enzymes (6)) or species  
62 known to be human pathogens (2). Using these culture-based techniques, Vacheyrou  
63 previously examined possible routes of microbial transfer in farms supplying raw milk for  
64 Comte style cheese, revealing that the extent to which milk was contaminated varied  
65 depending on the type of barns used to house animals (2). However, recent advances in  
66 molecular microbiology, and in high throughput DNA sequencing (HTS) in particular, have  
67 allowed for a more in-depth analysis of the flow of microbes through environments (7-12).

68 Indeed, a study of two artisan cheese-making plants observed that spatial diversification  
69 within both plants was indicative of “functional adaptations” by microbial communities  
70 colonising different fomites within each plant. Spatial diversification between plants  
71 confirms the phenomenon of a unique production plant (“house”)-associated microbiota,  
72 which was postulated to influence the distinct organoleptic properties of products from  
73 each facility (11). The facility-specific microbiota developed as a result of the selection  
74 pressure introduced by the individual cheese-making processing methods (11). The  
75 observation of a niche-specific functional adaptation has also been observed in the  
76 microbiota of a winery, with the additional observation that the community was influenced  
77 by seasonality (12).

78 The present proof of concept study focuses on the Irish dairy farm system, which is primarily  
79 a pasture based system, in which herds are grazed on pasture for the majority of their  
80 lactation curve. However, during the winter months, herds are housed indoors. The  
81 transition between environments is an important consideration for dairy producers as it is  
82 accompanied by changes in exposure to microbes from different niches in the environment  
83 as well as dietary changes. Previous, culture-based, efforts to address this question have  
84 noted elevated spore counts in bulk tank milk collected from a number of mid-West  
85 American farms during summer months on American farms (13), although elevated  
86 numbers of sporeformers can also be an issue when cows are housed indoors if poor quality  
87 silage is used (14). Our study also investigates the impact that teat preparation has on the  
88 microbiology of raw milk. This farm management practice has been shown to reduce  
89 bacterial counts in milk previously (15) but its impact on the raw milk microbiota has not  
90 been reported.

91 Based on the results of the studies highlighted above, and in the context of the seasonal  
92 milk production system applied in Ireland (all cows calved within a 12 week period), it is  
93 reasonable to assume that cattle are exposed to niche-specific microbes when housed  
94 indoors during winter months, and that these environmental microbes differ significantly  
95 from that present when the herd is grazing on pasture during the summer. Such differences  
96 would be expected, in turn, to impact on the raw milk microbiota. Specifically, we examined  
97 the influence that seasonal housing and grazing conditions have on the microbiota of raw  
98 cows' milk. We also examined the influence that the farm management practice of teat  
99 preparation (prep) has on the raw milk microbiota in both environments. To address these  
100 questions, we applied HTS and a Bayesian inference algorithm to examine environmental  
101 sources of bacteria, as well as seasonal changes to the raw milk microbiota driven by  
102 changes in habitat.

103

#### 104 **Materials and methods**

#### 105 **Treatment and Sample collection**

106 Samples were collected from the same herd of Holstein-Friesian dairy cows (n=60) from the  
107 Moorepark Research Farm (Fermoy, Co Cork, Ireland) during February (Average days in milk;  
108 ADIM= 140) and May (ADIM=200)) of 2015. The milking parlour and equipment were  
109 cleaned after each milking as outlined previously(16). Sampling phases corresponded to  
110 when cows were housed indoors (February) and outdoors on pasture (May). During the  
111 indoor sampling period (February) cows were fed grass silage within a cubicle house with  
112 automatic scraper cleaning of the central passageway. Cubicle beds were fitted with rubber  
113 mats with a daily allowance of ground limestone added to the backend of the cubicle. Cows  
114 managed in the outdoor sampling period (May) grazed on perennial ryegrass pasture on a  
115 24h rotational grazing regime. The herd was milked in a 30-unit, 80-degree side-by-side  
116 milking parlour (Dairymaster, Causeway, Co Kerry, Ireland). Although most studies  
117 incorporating molecular methods focus only on the bulk tank milk (BTM), in this instance,  
118 milk from three individual cows was also tested. Three cows with a somatic cell count lower  
119 than 100,000 cells/mL were chosen for specific individual sampling before commencement  
120 of the study and were used throughout the study. Milk and teat swab samples were  
121 collected twice weekly from these three cows throughout the study during the morning  
122 milking.

123 Two pre-milking teat preparation treatments were applied within each sampling phase. One  
124 treatment comprised of washing teats with running water, drawing of foremilk, and an  
125 application of a pre-milking teat disinfectant (Deosan Teat-foam) (Deosan, Johnson Diversey  
126 (Ireland) Ltd, Jamestown RD, Finglas 11, Dublin) followed at least 30 seconds later by drying  
127 using individual paper towels, prior to attaching the milking cluster (prep). The second  
128 treatment involved no teat preparation prior to cluster attachment for milking (non prep).  
129 For both indoor and outdoor sampling periods, the teat treatments applied were as follows:  
130 week one, all animals had teats prepped prior to milking; week two, animals were not  
131 prepped; week three, teats were prepped prior to milking and week four no teat  
132 preparation was carried out. All cows in the herd were subjected to each teat preparation  
133 treatment at each day of sampling. Environmental samples (faeces, bedding, silage grass  
134 and surface soil) were collected twice a week on day 1 and day 3, apart from the teat swab  
135 samples, which were collected after the teat preparation treatment was applied and prior to  
136 cluster attachment for milking on days 2 and 4. Microbial DNA was extracted from all

137 samples using the Powersoil kit (Mobio, Carlsbad CA). Due to the different sample types, the  
138 pre-processing protocol for samples varied. At morning's milking on day 2 and 4 of each  
139 sampling week, all four teats from the cows were swabbed using one sterile cotton swab  
140 per teat (Sarstedt, Ireland). Swabs were dipped in a solution of 3ml of NaCl (0.09%) prior to  
141 swabbing to improve recovery (17). Swabs were drawn across the teat orifice and up the  
142 side of each teat avoiding contact with the udder hair. The four swabs from each cow were  
143 then pooled in a NaCl solution (12 ml) in a sterile 15 ml falcon tube (Sarstedt, Ireland) and  
144 vortexed for 2 minutes. This resulted in one teat pool for each cow sampled at each time  
145 point. The pool, including liquid and swab heads, was then centrifuged for 5 minutes at 900  
146 x g to separate the swab heads from the liquid. The supernatant was then removed and  
147 transferred to another sterile 15 ml falcon tube. Each pool was then centrifuged at 5444 x g  
148 for 30 minutes at 4 °C. The supernatant was then carefully removed and the resulting pellet  
149 was dissolved in the lysis solution from the Powersoil microbead tubes.

150 Milk samples from the selected three cows were collected within sterilized sampling bottles  
151 using the Weighall milk meter on days 2 and 4 of each sampling week (DairyMaster,  
152 Causeway, Co Kerry, Ireland). 60 ml of individual milk was used for each extraction. BTM  
153 samples representing the complete herd were collected after the morning milking on days 2  
154 and 4. These were collected using 30 ml sterile blue dippa sample tubes (Ocon chemicals).  
155 60 ml of the BTM was used for each extraction. For both individual milk and BTM, milk was  
156 aseptically transferred to 15ml Falcon tubes (Sarstedt, Ireland), and centrifuged at 5444 x g  
157 for 30 minutes at 4 °C. The fat layer was carefully removed and the supernatant was  
158 decanted. The resulting pellets were then washed using sterile PBS and centrifuged at  
159 14,000 x g for 1 minute. The four pellets for each individual milk and BTM sample were then  
160 pooled, to give four samples (three individual milk samples and one BTM sample). Cell  
161 pellets were then dissolved in the lysis solution from the microbead tubes from the  
162 Powersoil kit.

163 For faecal pool samples, a pool of the herd's faecal samples was created at each day of  
164 sampling. Two faecal pools were collected on each week of sampling on day 1 and 3. To  
165 make this pool, equivalent amounts of faecal material were collected from 5 random cow  
166 pats and the pool was then homogenised for 2 minutes by vortexing at full speed. DNA was  
167 extracted from 250mg of this faecal pool.



168 Surface soil samples were collected on days 1 and 3 from the paddock from which the herd  
169 were grazing. These samples were collected, taking care to avoid collecting faeces or grass  
170 using a disposable spatula (VWR, Ireland), 250mg of surface soil was used for the soil sample  
171 extractions. For bedding, silage and grass samples, 20 g of material was aseptically collected  
172 using sterile forceps (VWR, Ireland) and scissors (for grass samples) (Medguard, Co. Meath  
173 Ireland) and stored in stomacher bags. For bedding samples 4g of bedding material was  
174 collected from 5 cubicles from which the herd had been occupying to create a 20g bedding  
175 sample, two bedding samples were collected on each week of the indoor sampling period.  
176 For silage samples 20g of silage was collected from where the herd was feeding, two silage  
177 samples were collected on each week of the indoor sampling period. For grass samples, 20g  
178 of grass was aseptically collected from the paddock in which the herd had been grazing  
179 when outdoors; two grass samples were collected on each week of the outdoor sampling  
180 period. Then 180 ml of sterile PBS was added to each stomacher bag and the samples were  
181 homogenised in a stomacher. The resultant mixture was then aliquoted into 50 ml falcon  
182 tubes and centrifuged at 900 x G for 5 minutes to remove solids. Following this, the  
183 supernatant was filtered through 0.45 µm nitro cellulose filter membrane (Merck Millipore).  
184 After filtration, the membrane was aseptically cut into microbead tubes (Powersoil kit) using  
185 a sterile scissors and forceps.

186 The sample numbers collected included surface soil (n=8), faeces (n=16, 8 indoor pools and  
187 8 outdoor pools), silage (n=8) and bedding (n=8), as well as teat swabs (n=48, of which 40  
188 subsequently yielded amplicons - 10 indoor no prep [INP], 11 indoor prep [IP], 11 outdoor  
189 prep [OP] and 8 outdoor no prep [ONP]), individual milk samples (n=48, of which 47  
190 subsequently yielded amplicons - 12 INP, 12 IP, 11 OP and 12 ONP), bulk tank milk (BTM;  
191 n=14, 4 INP, 3 IP, 3 ONP, and 4 OP) and grass (n=8).

192 After pre-processing of the samples had been pre-processed and lysis solution added, C1  
193 solution lysis solution (preheated to 60°C) was added to all samples, and followed  
194 incubation for 10 minutes at 60°C with vortexing every two minutes for 30 seconds. After  
195 this incubation, samples were mechanically lysed at full speed for 10 minutes using a  
196 TissueLyser (Qiagen) and then processed as per Powersoil kit protocol. DNA was quantified  
197 and quality checked by gel electrophoresis and spectrophotometry on a nanodrop 1000  
198 instrument (Thermo Fisher Scientific Inc).

199 **16S rRNA amplicon sequencing**

200 The V3-V4 variable region of the 16S rRNA gene was amplified from the 149 DNA extracts  
201 using the 16S metagenomic sequencing library protocol (Illumina). PCR reactions were  
202 completed on the template DNA. Initially, the DNA was amplified with primers specific to  
203 the V3-V4 region of the 16S rRNA gene which also incorporates the Illumina overhang  
204 adaptor (Forward primer 5'  
205 TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG; reverse primer 5'  
206 GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC) (18). Each PCR  
207 reaction contained DNA template (~10–12ng), 5 µl forward primer (1 µM), 5 µl reverse  
208 primer (1 µM), 12.5 µl 2X Kapa HiFi Hotstart ready mix (Anachem, Dublin, Ireland), PCR  
209 grade water to a final volume of 25µl. For environmental samples (surface soil, faecal, silage,  
210 swabs, bedding, and grass) PCR amplification was carried out as follows: heated lid 110°C,  
211 95°C x 3mins, 25 cycles of 95°C x 30s, 55°C x 30s, 72°C x 30s, then 72°C x 5mins and held at  
212 4°C was used. For milk samples the same cycling parameters were used, except 32 cycles  
213 were used instead of 25 cycles. PCR products were visualised using gel electrophoresis (1X  
214 TAE buffer, 1.5% agarose, 100V) and cleaned using AMPure XP magnetic beads (Labplan,  
215 Dublin, Ireland). Following this, a subsequent PCR reaction was completed on the purified  
216 DNA (5µl) to index each of the samples, allowing samples to be pooled for sequencing on  
217 three flow cell and subsequently demultiplexed for analysis. Samples were indexed  
218 randomly to prevent any run bias in analysis. Two indexing primers (Illumina Nextera XT  
219 indexing primers, Illumina, Sweden) were used per sample. Each PCR reaction contained 5µl  
220 index 1 primer (N7xx), 5µl index 2 primer (S5xx), 25µl 2x Kapa HiFi Hot Start Ready mix, 10µl  
221 PCR grade water. PCRs were completed as described above, with 8 amplification cycles. PCR  
222 products were visualised using gel electrophoresis and subsequently cleaned (as described  
223 above). Samples were quantified using the Qubit (Bio-Sciences, Dublin, Ireland); along with  
224 the broad range DNA quantification assay kit (BioSciences) and samples were then pooled in  
225 an equimolar fashion. The pooled sample was run on the Agilent Bioanalyser for quality  
226 analysis prior to sequencing. The sample pool (4nM) was denatured with 0.2N NaOH, then  
227 diluted to 4pM and combined with 10% (v/v) denatured 4pM PhiX, prepared following  
228 Illumina guidelines. Samples were sequenced on the MiSeq sequencing platform in the  
229 Teagasc sequencing facility, using a 2 x 250 cycle V3 kit, following standard Illumina

230 sequencing protocols. Reads were deposited in the SRA database under the accession number  
231 PRJEB16770.

### 232 **Bioinformatic and statistical analysis**

233 250 base pair paired-end reads were assembled using FLASH (FLASH: fast length adjustment  
234 of short reads to improve genome assemblies) (19). Further processing of paired-end reads  
235 including quality filtering based on a quality score of > 25 and removal of mismatched  
236 barcodes and sequences below length thresholds was completed using QIIME(20). A total of  
237 32,766,563 reads were generated post filtering, with an average of 219,909 per sample.  
238 Denoising, chimera detection and clustering into operational taxonomic units (OTUs) (97%  
239 identity) were performed using USEARCH v7 (64-bit)(21). OTUs were aligned using PyNAST  
240 (python nearest alignment space termination; a flexible tool for aligning sequences to a  
241 template alignment) and taxonomy was assigned using BLAST against the SILVA SSURef  
242 database release 111. Samples were then rarefied to an even depth of sequences per  
243 sample. Alpha diversity was generated in QIIME and the compareGroups function (22) was  
244 then used to determine any statistically significant differences ( $P < 0.05$ ) and  
245 generate standard deviations between samples based on conditions using the ANOVA test..  
246 Beta diversity was calculated in R, using Phyloseq (23) and Bray Curtis distances. Principal  
247 coordinate analysis (PCoA) plots were visualised using ggplot2 (24). Confidence ellipses were  
248 generated using stat\_ellipse in the ggplot2 package (24). Network analysis was also carried  
249 out using phyloseq and ggplot2. The SourceTracker algorithm (9) was also used to  
250 investigate possible sources of environmental contamination in milk from both sampling  
251 periods. SourceTracker analysis was carried out at a depth of 13500, with 100 burn-ins and  
252 10 re-starts. The compareGroups function was used in R to compare differences in microbial  
253 composition between individual milk, teat swab and faecal pool samples; the Kruskal Wallis  
254 test was applied in this instance with Benjamini-Hochberg corrections (25), to highlight any  
255 statistically significant differences ( $P < 0.05$  after correction).

### 256 **Quantitative PCR**

257 Quantitative PCR (qPCR) was carried out on individual milk samples to determine total  
258 bacteria levels in each sample using 16S rRNA gene. qPCR was carried out as described  
259 previously (26) except for the use of the equivalent volume of Kappa SYBR fast (Roche

260 Diagnostics) was used instead of SYBR green for the present study. Samples, negative  
261 controls (where template DNA was replaced with PCR-grade water) and standards were run  
262 in triplicate (technical replicates).

263

## 264 **Results**

### 265 **Microbiota alpha and beta diversity of raw milk, teat surface swabs and environmental** 266 **samples cluster according to habitat**

267 Samples were collected from the same herd over two sampling periods. Sampling phases  
268 corresponded to when the herd was housed indoors and outdoors on pasture, respectively.  
269 Across both sampling phases, milk samples were collected from teat prepared (prepped)  
270 and non-teat prepped samples. Samples were also classified as either a potential 'source' of  
271 microorganisms or a 'sink' (a sample that is liable to contain bacteria originating from a  
272 source). Milk samples both from individual cows and BTM were classified as sinks and all  
273 environmental samples were classified as sources. After sequencing, the alpha and beta  
274 diversity of the bacterial populations present was investigated.

275 Alpha diversity is the diversity in each sample, using species richness and evenness to  
276 calculate the diversity in each environment. There was no significant difference in alpha  
277 diversity between the microbiotas of individual indoor and outdoor milk samples from non-  
278 prepped animals. Similarly, there was no significant difference in the alpha diversity of the  
279 microbiota of indoor milk sourced from animals who underwent teat prep and those that  
280 did not. However, the alpha diversity of the outdoor milk microbiota was significantly higher  
281 in OP samples relative to IP samples (P=0.016 Simpsons diversity index, P=0.008 Shannon diversity  
282 index; Table 1). A corresponding analysis of the alpha diversity of the microbiota of the teat  
283 surface revealed significantly greater diversity (chao1, Shannon, PD whole tree and  
284 observed species) among OP samples relative to IP samples (P=<0.01, 0.026, <0.01 and  
285 <0.01, respectively; Table 1). No other significant differences in the alpha diversity of teat  
286 microbiota samples were observed.

287 Beta diversity is the diversity between different samples; it provides a measure of  
288 dissimilarity between samples. The Bray Curtis Principle Coordinate plot of beta diversity

289 (Fig.1A) depicts all samples from this study with data points coloured by sample origin and  
290 shaped according to their designation as source or sink. In this plot it can be observed that  
291 samples (soil, grass, bedding, silage, teat surface indoor, teat surface outdoor, faecal indoor  
292 pool, faecal outdoor pool, indoor milk, outdoor milk [individual and BTM]) form clusters,  
293 which in turn are further separated from one another based on habitat (outdoor/indoor).  
294 More specifically, there is a clear separation between samples depending on whether they  
295 were collected from an indoor or an outdoor environment. Faeces, teat, individual milk  
296 samples and BTM samples also separate based on which environment they were sampled  
297 from (indoor/outdoor) (Fig.1A). There are more similarities between samples taken from the  
298 same habitat. This includes environmental samples (grass and soil [outdoor] and bedding  
299 and silage [indoor]), as seen by the overlaps in the ellipses. Within both habitats, it is  
300 apparent that there is an overlap between data points representing the milk sample  
301 microbiota and that of teat swab samples, reflecting similarities in their beta diversity  
302 (Fig.1A). Teat prep did not result in further sub-clusters within the milk or teat samples  
303 (Fig.S1). Faecal pool samples from both habitats separate from one another and are located  
304 in relatively close proximity to the corresponding milk and teat samples from the same  
305 environment (Fig.1A).

#### 306 **Network analysis shows relationships between raw milk and environmental samples**

307 Network plots are a useful graphical tool to illustrate relationships between microbiota  
308 datasets. The nodes in this network plot represent samples, and the edges that connect  
309 nodes indicate correlations between samples. The network analysis shows relationships that  
310 exist between the environmental samples and milk samples (Fig.1B). Consistent with beta  
311 diversity data, it is particularly notable that, of the environmental microbiota samples, the  
312 faecal pools and teat microbiota are most closely related to the microbiota of the milk  
313 samples, thereby identifying faeces and the teat surface as important sources of  
314 contamination. These relationships reflect the habitat (indoor or outdoor) from which the  
315 samples were collected. There are more edges linking indoor faecal pool samples with  
316 indoor BTM samples, than outdoor faecal pool samples with outdoor BTM. Some of the  
317 outdoor milk samples are not linked to any of the outdoor sources by edges. This suggests  
318 that these niches are not substantial sources of microbial contaminants in these milk  
319 samples.

320 **SourceTracker analysis further highlights the contribution of faecal and teat sources to the**  
321 **raw milk microbiota**

322 The SourceTracker model assumes that each individual community (milk, soil, grass, faeces,  
323 teat, bedding and silage) is a mixture of communities deposited from other known or  
324 unknown source environments and, using a Bayesian approach, the model provides an  
325 estimate of the proportion of the community originating from each of the different sources.  
326 When a community contains a mixture of taxa that do not match any of the potential source  
327 environments studied, that portion of the community is assigned to an “unknown” source.  
328 The analysis revealed that the teat surface was the most significant contributor of microbes  
329 in milk samples regardless of habitat or teat preparation. Teat surface contaminants  
330 constitute a higher proportion of total contaminants in indoor milk compared to outdoor  
331 milk, both for individual and for BTM samples. Faeces was the next most important source  
332 of contaminants, and had a greater influence on indoor, than outdoor, milk samples,  
333 particularly in BTM samples (Fig.2).

334 **Taxonomic analysis of raw milk, teat surface and herd faecal microbiota**

335 Graphs representing the microbiota at Family level in the various sample sets are provided  
336 in the supplementary data (Supplementary Fig. 2-3). The compareGroups function was used  
337 in R to compare differences in microbial composition between samples. OTUs that differ  
338 significantly can be found in the supplementary material (Tables S1-S3). In milk samples  
339 from individual animals that did not undergo a teat prep treatment, it was noted that indoor  
340 samples contained higher relative proportions of, for example, *Eremococcus*, *Ruminococcus*,  
341 *Prevotella*, uncultured *Corynebacteriales* bacterium, and *Ruminococcaceae* Incertae Sedis  
342 ( $P=0.012$ ,  $0.012$ ,  $0.02$ ,  $0.022$ ,  $0.028$ , respectively) and lower proportions of *Pseudomonas*,  
343 *Acinetobacter*, *Lactococcus* and *Tumebacillus* ( $P=0.003$ ,  $0.008$ ,  $0.002$  and  $0.014$  respectively),  
344 relative to outdoor milk samples. qPCR analysis to determine total bacterial numbers  
345 showed that there was significantly more bacteria in indoor milk samples than the  
346 equivalent outdoor milk samples ( $P=0.003$ ) (Table 2). When the corresponding milk samples  
347 from individual teat prepped animals were compared, it was noted that 25 genera were  
348 present in significantly different proportions in indoor milk samples relative to outdoor-milk  
349 samples. Sixteen of these OTUs were higher in indoor samples, these include *Eremococcus*,

350 *Alloiococcus*, *Trichococcus*, *Prevotella*, and *Psychrobacter*, which were all more abundant in  
351 indoor samples ( $P=0.001$ ,  $0.001$ ,  $0.001$ ,  $0.02$ , and  $0.019$ , respectively). Nine OTUs were  
352 higher in PO samples, including *Flavobacterium*, *Sphingomonas* and *Tumebacillus* ( $P= 0.009$ ,  
353  $0.014$ , and  $0.021$  respectively). There was no significant difference in total bacterial numbers  
354 between the indoor and outdoor milk samples from teat prepped cows ( $P=0.598$ ) (Table 2).

355 The taxonomic data also facilitated an analysis of the specific effects of teat prep on the  
356 bacterial composition of the milk produced. In indoor milk samples from individual animals,  
357 it was noted that proportions of *Pseudomonas* were higher in samples from cows which had  
358 undergone teat prep ( $P=0.035$ ) suggesting that, among the indoor teat microbiota,  
359 *Pseudomonas* was relatively less sensitive to the antimicrobial effects of the teat prep in  
360 indoor samples. qPCR analysis demonstrated that there was no significant difference in total  
361 bacterial numbers because of the teat prep ( $P=0.758$ ) (Table 2). *Pseudomonas*, *Lactococcus*  
362 and *Lactobacillus* were among nine genera present in outdoor milk samples that were  
363 influenced by teat prep. In the case of the aforementioned genera, proportions were higher  
364 in samples when no teat prep was carried out ( $P=0.011$ ,  $0.025$ , and  $0.03$ , respectively).  
365 There were significantly fewer total bacteria in milk samples from non-prepped animals  
366 samples compared to samples from prepped animals in the outdoor environment ( $P=0.004$ )  
367 (Table 2).

368 The microbiota composition of the teat swabs was also assessed and it was established that,  
369 in samples where teat prep did not occur, 18 genera differed significantly in their relative  
370 abundance between indoor and outdoor samples. *Trichococcus*, *Proteiniphilum*, and  
371 *Eremococcus*, as well as *Corynebacterium*, were more abundant in indoor samples ( $P= 0.012$ ,  
372  $0.021$ ,  $0.044$ , and  $0.039$ , respectively) while a further 11 OTU's were present in significantly  
373 higher proportions in outdoor samples. In samples where teat preparation was carried out,  
374 60 genera differed significantly between indoor and outdoor samples. Twenty-one of these,  
375 including *Eremococcus*, *Proteiniphilum*, *Corynebacterium*, *Psychrobacter* *Bifidobacterium*,  
376 *Trichococcus* and *Prevotella*, were significantly higher in indoor samples ( $P= 0.001$ ,  $0.001$   
377  $0.002$ ,  $0.002$   $0.003$ ,  $0.004$ , and  $0.005$ , respectively) and thirty-nine genera, including  
378 *Stenotrophomonas*, *Xanthomonas* and *Rhizobium*, ( $P= 0.001$ ,  $0.001$ , and  $0.003$ , respectively)  
379 were significantly higher in outdoor samples. Among the outdoor teat samples, there were  
380 no significant differences between prepped and non-prepped samples. Among the

381 corresponding indoor teat samples, proportions of *Variovorax* and *Devosia* were higher in  
382 teat samples which were not treated ( $P=0.033$  and  $0.043$ ) (Supplementary table 2).

383 Additionally, it is noteworthy from the stacked bar charts (Fig S1 (B) and (D)) that the  
384 composition of individual milk samples differs considerably from that of BTM. More  
385 specifically, higher proportions of *Micrococcaceae* and *Flavobacteriaceae* are observed in all  
386 individual milk sample types and *Prevotella* and *Rikenellaceae* were higher in BTM samples.

387 Finally, the availability of faecal pool samples from both the indoor and outdoor  
388 environment facilitated a comparison of their composition. At the genus level 15 genera,  
389 including *Prevotella*, *Bacteroides* and *Treponema*, were higher in indoor faecal pool samples  
390 ( $P=0.001$ ,  $0.002$ , and  $0.021$ ) and a further eight genera, including *Phocaeicola* and  
391 *Paludibacter*, were higher in outdoor faecal pool samples ( $P=0.027$  and  $0.036$ )  
392 (Supplementary table 3).

393

#### 394 Discussion

395 The objective of this proof of concept study was to harness the power of next-generation  
396 DNA sequencing technologies to investigate the influence that seasonal housing and teat  
397 preparation have on the raw milk microbiota from individual cows and in BTM.  
398 Furthermore, information potentially revealing the extent to which different microbial  
399 niches in the milk production environment influence the microbiota of raw milk was also  
400 generated. While, in the past, culture-based investigations to study the source of  
401 microorganisms in raw milk have primarily focused on BTM, in this instance samples from a  
402 small subset of individual animals was also included. While analysis did not reveal  
403 differences between the microbiota alpha diversity of indoor and outdoor milk samples,  
404 beta diversity analysis highlighted a clear separation between samples that are sourced  
405 from an indoor versus an outdoor environment. No distinct separation pattern was  
406 observed when samples were coloured by teat preparation treatment (Fig S1). Thus, this  
407 analysis demonstrates that habitat had a greater impact on the raw milk microbiota than  
408 teat preparation.



409 The SourceTracker algorithm was used as a complementary means of identifying the likely  
410 source within the dairy farm environment (soil, silage, bedding, grass, teat, and faeces) of  
411 bacteria ultimately found in raw milk and, in the process, also reveals the influence of  
412 seasonal housing and farm management practices. Regardless of habitat or treatment, teat  
413 surface was again identified as the greatest contributor to the raw milk microbiota, followed  
414 by faeces. This is consistent with a previous (culture-based) study, which proposed that the  
415 teat skin was a source of microbial populations in raw milk and that farm management and  
416 animal grazing practices influenced the diversity and microbiota of raw milk(27).

417 The taxonomic results also show that habitat had a much greater influence on the raw milk  
418 and teat microbiota than teat prep. For instance, in milk samples from cows that were not  
419 subjected to teat prep, Gram positive and gut-associated genera were higher in indoor,  
420 relative to outdoor milk, such as *Ruminococcus*, *Eremococcus*, *Ruminococcaceae* Incertae  
421 Sedis and uncultured *Corynebacteriales* were higher in indoor, relative to outdoor, samples.  
422 *Ruminococcus* and *Ruminococcaceae* Incertae Sedis are both gut-associated genera  
423 although, from a dairy perspective, *Ruminococcaceae* Incertae Sedis has previously been  
424 found in continental type cheese (28) and *Ruminococcus* has been detected in raw milk (29)  
425 , and in this study these were in higher proportions in INP milk compared to ONP. While,  
426 relatively little is known about the uncultured *Corynebacteriales*, the cultured equivalent  
427 contains species known to cause mastitis (30) as well as others that are found on the surface  
428 of surface-ripened cheese (31). Similarly, the other genus noted, *Eremococcus*, has not been  
429 well characterised, although a typed strain does exist, having been isolated from the vaginal  
430 discharge of a thoroughbred horse (32). Proportions of the Gram negative genus *Prevotella*,  
431 which is typically gut-associated was also higher in indoor samples while, for the outdoor  
432 samples, the Gram negative genera *Pseudomonas* and *Acinetobacter*, as well as the Gram  
433 positive genus *Lactococcus*, were among those that were more dominant. *Pseudomonas*  
434 and *Acinetobacter* are both dairy spoilage-associated genera (6) that can have a negative  
435 impact on dairy product quality. Lactococci are best known for their positive contribution to  
436 the production of fermented dairy products, but can also be isolated from outdoor  
437 environments such as grass (33). These results indicate that indoor milk is more likely to  
438 have higher proportions of host/gut associated microbes than outdoor milk while,

439 unsurprisingly, outdoor milk is more likely to contain higher proportions of environmental  
440 bacteria.

441 For milk samples from cows that were teat prepped prior to milking, LAB, such as  
442 *Eremococcus*, *Alloiococcus*, and *Trichococcus*, as well as *Psychrobacter*, are also in a  
443 significantly higher proportion in IP samples. Interestingly, *Alloiococcus* has not been  
444 described in raw milk previously, having instead being associated with human ear infections  
445 (34). *Trichococcus* has been found in raw milk and dairy waste (35) and *Psychrobacter* have  
446 previously been found in teat apexes (36) and in cheese (37). Again, in the corresponding OP  
447 milk samples soil bacteria such as *Flavobacterium*, *Sphingomonas* and *Tumebacillus* where  
448 in higher proportions. This indicates that outdoor milk is more likely to contain increased  
449 proportions of soil associated microbes, while indoor milk is more likely to have higher  
450 proportions of host/gut bacteria. The proportions of LAB found in the milk appear to be low  
451 in comparison to other studies (29), this is perhaps due to the protocol used which did not  
452 incorporate enzymatic lysis.

453 In teat swab samples, Gram positive genera such as *Corynebacterium*, *Trichococcus* and  
454 *Eremococcus* and Gram negative genera such as *Proteiniphilum* were significantly higher in  
455 NPI samples compared to NPO samples. *Proteiniphilum* has previously been associated with  
456 the faeces of dairy cattle (38). A number of soil type OTU's were observed to be significantly  
457 elevated in NPO, relative to NPI teat swab samples. This indicates that the transmission of  
458 soil type bacteria to the teat is greater in periods where cows are grazing outdoors,  
459 potentially leading to subsequent transmission from the teat to milk. In teat samples that  
460 were prepped, *Corynebacterium*, *Eremococcus* and *Trichococcus* were again more abundant  
461 in IP teat samples. *Bifidobacterium* was also present in greater proportions in these samples.  
462 Although *Bifidobacterium* is typically associated with the gastrointestinal tract (GIT) of warm  
463 blooded mammals (38), it may be significant that prep has previously been shown to cause  
464 an increase in Actinobacteria proportions on the teat surface (15). With regard to Gram  
465 negative bacteria, *Proteiniphilum*, *Psychrobacter* and *Prevotella*, were all significantly more  
466 abundant in IP teat swab samples compared to OP samples. In outdoor samples that were  
467 teat prepped, many soil type bacteria, including *Rhizobium*, *Xanthomonas*, and  
468 *Stenotrophomonas*, were significantly more prevalent compared to OP samples. Thus, soil-

469 type bacteria, also noted on the surface of ONP teat surface, persist even when teat prep  
470 occurs.

471 Using the data generated, it possible to assess the impact of teat preparation on the milk  
472 and teat microbiota composition by comparing data from animals that were/were not  
473 subjected to a treatment (during the same season). In milk samples, lactic acid bacteria,  
474 such as *Lactococcus* and *Lactobacillus*, and *Pseudomonas* were higher in NPO samples,  
475 suggesting that the application teat prep significantly reduced the numbers of these  
476 microbes in raw milk. There were no significant differences between PO and NPO teat swab  
477 samples. Among indoor teat samples, soil type *Proteobacteria*, such as *Variovorax* and  
478 *Devosia*, were more abundant in NPI, relative to PI teats. *Variovorax* has previously been  
479 found in hay (2), and *Devosia* has previously been found in raw milk (39). It was surprising to  
480 note that teat prep increased the numbers of total bacteria in both indoor and outdoor milk.  
481 Alpha diversity was also found to have increased in milk from cows where teats were  
482 prepped prior to milking compared to milk from cows where teat preparation was omitted.  
483 It may be that the teat preparation process, including forestripping and drying, weakens the  
484 attachment of commensal and contaminating teat canal bacteria and results in their being  
485 shed into the milk in greater numbers. This result contrasts findings from culture based  
486 analysis on the impact of teat prep on raw milk, which found that it reduced bacterial  
487 diversity or counts respectively (15, 40). Further studies will be required to re-examine the  
488 influence that teat preparation has on the raw milk microbiota. Another important  
489 consideration is that the farm used in this study is a research farm where stringent hygiene  
490 practices are upheld. This could perhaps limit the impact that teat preparation has on the  
491 raw microbiota

492 There were considerable differences observed between the individual milk and BTM  
493 microbiotas (Fig S1). This may be due to microorganisms in the BTM being acquired from  
494 the milking machine and pipes. Indeed, this possibility has been highlighted previously (41)  
495 but not in the context of DNA-based analysis. Further explorations to definitively establish  
496 the basis for these differences is merited.

497 The availability of faecal microbiota data from multiple samples also facilitated comparative  
498 analysis of these samples. It was apparent that the beta diversity of the herd faecal pool

499 microbiota differed significantly from the two sampling periods. From a taxonomic  
500 perspective, eight genera were found to be significantly higher in outdoor herd faecal  
501 samples and fifteen genera were found to be significantly higher in indoor herd faecal pool  
502 samples. *Treponema*, *Prevotella* and *Bacteroides* were among the gut-associated genera  
503 that were more prevalent in indoor samples. *Treponema* has previously been associated  
504 with digital dermatitis in cattle (42) and in the bovine rumen (43). *Phocaeicola* and  
505 *Paludibacter* have also been positively associated with valerate in the rumen previously (44),  
506 and were higher in outdoor samples. This difference in faecal microbiota may be influenced  
507 by habitat, host physiological changes or by dietary changes associated with the differing  
508 habitats. It is also possible that transmission of bacteria from faecal origin may differ based  
509 on habitat due to the differences in the microbiota seen here.

510 Here, high-throughput DNA sequencing has facilitated the analysis of the microbiota of raw  
511 milk samples in parallel with samples from the dairy farm environment. The results provide  
512 a more detailed insight into the composition of these microbial populations while also  
513 allowing an examination of the relationship between the microbiota of these environments  
514 and of raw milk. This analysis highlights that herd habitat is a significant driver for milk  
515 microbiota composition, and that teat prep has a much more limited impact on the raw milk  
516 microbiota. In the process it is made apparent that high-throughput sequencing can be an  
517 extremely insightful tool to help better understand the movement of microbes from the  
518 environment into the food chain.

519

#### 520 **Acknowledgments**

521 Sincere gratitude is due to John Paul Murphy and all the staff at the Animal and Grassland  
522 Research and Innovation Centre Dairy Unit, Teagasc, Moorepark, Fermoy. Thanks also to  
523 Fiona Fouhy and Fiona Crispie of the Food Bioscience department in Teagasc Moorepark.  
524 The project was funded through a Teagasc Walsh Fellowship ([2013030](#)) to CD and  
525 internal Teagasc funding ([RMIS6364](#)) to PC.

526

527

528 **References**

- 529 1. **Sevi A, Albenzio M, Muscio A, Casamassima D, Centoducati P.** 2003. Effects of litter  
530 management on airborne particulates in sheep houses and on the yield and quality of ewe  
531 milk. *Livestock production science* **81**:1-9.
- 532 2. **Vacheyrou M, Normand A-C, Guyot P, Cassagne C, Piarroux R, Bouton Y.** 2011. Cultivable  
533 microbial communities in raw cow milk and potential transfers from stables of sixteen  
534 French farms. *International journal of food microbiology* **146**:253-262.
- 535 3. **Vithanage NR, Dissanayake M, Bolge G, Palombo EA, Yeager TR, Datta N.** 2016. Biodiversity  
536 of culturable psychrotrophic microbiota in raw milk attributable to refrigeration conditions,  
537 seasonality and their spoilage potential. *International Dairy Journal* **57**:80–90.
- 538 4. **Mhone TA, Matope G, Saidi PT.** 2011. Aerobic bacterial, coliform, *Escherichia coli* and  
539 *Staphylococcus aureus* counts of raw and processed milk from selected smallholder dairy  
540 farms of Zimbabwe. *International Journal of Food Microbiology* **151**:223-228.
- 541 5. **Doyle CJ, Gleeson D, Jordan K, Beresford TP, Ross RP, Fitzgerald GF, Cotter PD.** 2015.  
542 Anaerobic sporeformers and their significance with respect to milk and dairy products.  
543 *International journal of food microbiology* **197**:77-87.
- 544 6. **Hantsis-Zacharov E, Halpern M.** 2007. Culturable psychrotrophic bacterial communities in  
545 raw milk and their proteolytic and lipolytic traits. *Applied and environmental microbiology*  
546 **73**:7162-7168.
- 547 7. **Kembel SW, Meadow JF, O'Connor TK, Mhuireach G, Northcutt D, Kline J, Moriyama M,  
548 Brown G, Bohannon BJ, Green JL.** 2014. Architectural design drives the biogeography of  
549 indoor bacterial communities. *PloS one* **9**:e87093.
- 550 8. **Flores GE, Bates ST, Knights D, Lauber CL, Stombaugh J, Knight R, Fierer N.** 2011. Microbial  
551 biogeography of public restroom surfaces. *PLoS One* **6**:e28132.
- 552 9. **Knights D, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, Bushman FD,  
553 Knight R, Kelley ST.** 2011. Bayesian community-wide culture-independent microbial source  
554 tracking. *Nature methods* **8**:761-763.
- 555 10. **Bokulich NA, Bergsveinson J, Ziola B, Mills DA.** 2015. Mapping microbial ecosystems and  
556 spoilage-gene flow in breweries highlights patterns of contamination and resistance. *eLife*  
557 **4**:e04634.
- 558 11. **Bokulich NA, Mills DA.** 2013. Facility-specific “house” microbiome drives microbial  
559 landscapes of artisan cheesemaking plants. *Applied and environmental microbiology*  
560 **79**:5214-5223.
- 561 12. **Bokulich NA, Ohta M, Richardson PM, Mills DA.** 2013. Monitoring seasonal changes in  
562 winery-resident microbiota. *PloS one* **8**:e66437.
- 563 13. **Buehner KP, Anand S, Garcia A.** 2014. Prevalence of thermotolerant bacteria and spores on 10  
564 Midwest dairy farms. *Journal of dairy science* **97**:6777-6784.
- 565 14. **Gleeson D, O'Connell A, Jordan K.** 2013. Review of potential sources and control of  
566 thermotolerant bacteria in bulk-tank milk. *Irish Journal of Agricultural and Food Research*  
567 **52**:217-227.
- 568 15. **Verdier-Metz I, Michel V, Delbes C, Montel M-C.** 2009. Do milking practices influence the  
569 bacterial diversity of raw milk? *Food Microbiology* **26**:305-310.
- 570 16. **O'Connell A, Ruegg P, Jordan K, O'Brien B, Gleeson D.** 2016. The effect of storage  
571 temperature and duration on the microbial quality of bulk tank milk. *Journal of dairy science*  
572 **99**:3367-3374.
- 573 17. **Landers TF, Hoet A, Wittum TE.** 2010. Swab type, moistening, and preenrichment for  
574 *Staphylococcus aureus* on environmental surfaces. *Journal of clinical microbiology* **48**:2235-  
575 2236.

- 576 18. **Fouhy F, Deane J, Rea MC, O'Sullivan Ó, Ross RP, O'Callaghan G, Plant BJ, Stanton C.** 2015.  
577 The Effects of Freezing on Faecal Microbiota as Determined Using MiSeq Sequencing and  
578 Culture-Based Investigations. *PLoS one* **10**:e0119355.
- 579 19. **Magoč T, Salzberg SL.** 2011. FLASH: fast length adjustment of short reads to improve  
580 genome assemblies. *Bioinformatics* **27**:2957-2963.
- 581 20. **Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N,  
582 Pena AG, Goodrich JK, Gordon JI.** 2010. QIIME allows analysis of high-throughput  
583 community sequencing data. *Nature methods* **7**:335-336.
- 584 21. **Edgar RC.** 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*  
585 **26**:2460-2461.
- 586 22. **Subirana IV, Sanz J, Lucas H, Peñafiel G, Giménez J. D.**(2013). compareGroups: Descriptive  
587 analysis by groups: R package version 2.0. Retrieved September 2013.
- 588 23. **McMurdie PJ, Holmes S.** 2013. phyloseq: an R package for reproducible interactive analysis  
589 and graphics of microbiome census data.
- 590 24. **Wickham H, Chang W, Wickham MH.** 2013. Package 'ggplot2'.
- 591 25. **Benjamini Y, Hochberg Y.** 1995. Controlling the false discovery rate: a practical and powerful  
592 approach to multiple testing. *Journal of the royal statistical society Series B*  
593 (Methodological):289-300.
- 594 26. **Fouhy F, Guinane CM, Hussey S, Wall R, Ryan CA, Dempsey EM, Murphy B, Ross RP,  
595 Fitzgerald GF, Stanton C, Cotter PD.** 2012. High-throughput sequencing reveals the  
596 incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic  
597 treatment with ampicillin and gentamicin. *Antimicrobial agents and chemotherapy* **56**:5811-  
598 5820.
- 599 27. **Verdier-Metz I, Gagne G, Bornes S, Monsallier F, Veisseire P, Delbès-Paus C, Montel M-C.**  
600 2012. Cow teat skin, a potential source of diverse microbial populations for cheese  
601 production. *Applied and environmental microbiology* **78**:326-333.
- 602 28. **O'Sullivan DJ, Cotter PD, O'Sullivan O, Giblin L, McSweeney PL, Sheehan JJ.** 2015. Temporal  
603 and spatial differences in microbial composition during the manufacture of a Continental-  
604 type cheese. *Applied and environmental microbiology* **81**:2525-2533.
- 605 29. **Quigley L, McCarthy R, O'Sullivan O, Beresford TP, Fitzgerald GF, Ross RP, Stanton C, Cotter  
606 PD.** 2013. The microbial content of raw and pasteurized cow's milk as determined by  
607 molecular approaches. *Journal of Dairy Science* **96**:4928-4937.
- 608 30. **Hogan J, Smith K, Todhunter D, Schoenberger P.** 1988. Rate of environmental mastitis in  
609 quarters infected with *Corynebacterium bovis* and *Staphylococcus* species. *Journal of dairy  
610 science* **71**:2520-2525.
- 611 31. **Beresford TP, Fitzsimons NA, Brennan NL, Cogan TM.** 2001. Recent advances in cheese  
612 microbiology. *International Dairy Journal* **11**:259-274.
- 613 32. **Collins MD, Jovita MR, Lawson PA, Falsen E, Foster G.** 1999. Characterization of a novel  
614 Gram-positive, catalase-negative coccus from horses: description of *Eremococcus coleocola*  
615 gen. nov. sp. nov. *International Journal of Systematic and Evolutionary Microbiology*  
616 **49**:1381-1385.
- 617 33. **Alemayehu D, Hannon JA, McAuliffe O, Ross RP.** 2014. Characterization of plant-derived  
618 lactococci on the basis of their volatile compounds profile when grown in milk. *International  
619 journal of food microbiology* **172**:57-61.
- 620 34. **Aguirre M, Collins M.** 1992. Phylogenetic analysis of *Alloiooccus otitis* gen. nov., sp. nov.,  
621 an organism from human middle ear fluid. *International Journal of Systematic and  
622 Evolutionary Microbiology* **42**:79-83.
- 623 35. **Rasolofoa EA, St-Gelais D, LaPointe G, Roy D.** 2010. Molecular analysis of bacterial  
624 population structure and dynamics during cold storage of untreated and treated milk.  
625 *International journal of food microbiology* **138**:108-118.

- 626 36. **Braem G, De Vliegher S, Verbist B, Heyndrickx M, Leroy F, De Vuyst L.** 2012. Culture-  
627 independent exploration of the teat apex microbiota of dairy cows reveals a wide bacterial  
628 species diversity. *Veterinary microbiology* **157**:383-390.
- 629 37. **Quigley L, O'Sullivan O, Beresford TP, Ross RP, Fitzgerald GF, Cotter PD.** 2012. High-  
630 throughput sequencing for detection of subpopulations of bacteria not previously associated  
631 with artisanal cheeses. *Applied and environmental microbiology* **78**:5717-5723.
- 632 38. **Kim M, Wells JE.** 2016. A Meta-analysis of Bacterial Diversity in the Feces of Cattle. *Current*  
633 *microbiology* **72**:145-151.
- 634 39. **Baur C, Krewinkel M, Kranz B, von Neubeck M, Wenning M, Scherer S, Stoeckel M, Hinrichs**  
635 **J, Stressler T, Fischer L.** 2015. Quantification of the proteolytic and lipolytic activity of  
636 microorganisms isolated from raw milk. *International Dairy Journal* **49**:23-29.
- 637 40. **McKinnon CH, Rowlands GJ, Bramley AJ.** 1990. The effect of udder preparation before  
638 milking and contamination from the milking plant on bacterial numbers in bulk milk of eight  
639 dairy herds. *Journal of Dairy Research* **57**:307-318.
- 640 41. **Quigley L, O'Sullivan O, Stanton C, Beresford TP, Ross RP, Fitzgerald GF, Cotter PD.** 2013.  
641 The complex microbiota of raw milk. *FEMS microbiology reviews* **37**:664-698.
- 642 42. **Trott DJ, Moeller MR, Zuerner RL, Goff JP, Waters WR, Alt DP, Walker RL, Wannemuehler**  
643 **MJ.** 2003. Characterization of *Treponema phagedenis*-like spirochetes isolated from  
644 papillomatous digital dermatitis lesions in dairy cattle. *Journal of clinical microbiology*  
645 **41**:2522-2529.
- 646 43. **Bekele AZ, Koike S, Kobayashi Y.** 2011. Phylogenetic diversity and dietary association of  
647 rumen *Treponema* revealed using group-specific 16S rRNA gene-based analysis. *FEMS*  
648 *microbiology letters* **316**:51-60.
- 649 44. **Mao S, Zhang R, Wang D, Zhu W.** 2012. The diversity of the fecal bacterial community and  
650 its relationship with the concentration of volatile fatty acids in the feces during subacute  
651 rumen acidosis in dairy cows. *BMC Veterinary Research* **8**:1-13.
- 652
- 653

654 **Table 1: Alpha diversity differences between individual milk and teat swab samples**

Milk								
	INP	ONP	IP	OP	P value INP vs ONP	P value IP vs OP	P value INP vs IP	P value ONP vs OP
<b>chao1</b>	3139 (1271)	2733 (833)	3017 (703)	3328 (784)	0.721	0.867	0.99	0.445
<b>Simpson</b>	0.98 (0.02)	0.95 (0.05)	0.98 (0.02)	0.98 (0.02)	0.036	0.885	0.983	0.016
<b>Shannon</b>	8.25 (1.07)	7.49 (1.17)	8.26 (1.07)	9.02 (0.80)	0.309	0.361	1	0.008
<b>PD whole tree</b>	90.3 (29.4)	70.5 (27.2)	93.8 (26.1)	86.3 (23.8)	0.304	0.918	0.99	0.521
<b>observed species</b>	2914 (1232)	2525 (784)	2791 (706)	3036 (752)	0.726	0.922	0.988	0.547
Teat								
	INP	ONP	IP	OP	P value INP vs ONP	P value IP vs OP	P value INP vs IP	P value ONP vs OP
<b>chao1</b>	3373 (792)	4307 (1172)	2949 (536)	4791 (1219)	0.187	<0.001	0.742	0.699
<b>Simpson</b>	0.99 (0.01)	0.99 (0.00)	0.99 (0.01)	0.99 (0.00)	0.99	0.716	0.962	0.997
<b>Shannon</b>	8.54 (0.67)	8.84 (0.41)	8.44 (0.48)	9.17 (0.67)	0.695	0.026	0.977	0.612
<b>PD whole tree</b>	125 (27.0)	157 (37.7)	107 (17.8)	174 (39.9)	0.156	<0.001	0.589	0.665
<b>observed species</b>	3194 (767)	4090 (1119)	2725 (500)	4526 (1188)	0.19	<0.001	0.655	0.741

655

656 Numbers in the brackets represent standard deviations. NPI= No prep indoor; NPO=No prep  
 657 outdoor; PI= Prep indoor; PO= Prep outdoor

658



659

660 Table 2: (A )qPCR determination of total bacteria numbers for individual milk samples, (B) results of  
661 comparison total bacterial numbers present in individual milk samples from different conditions

<b>A</b>		
Sample Type	Total bacteria (copies of 16S rRNA gene)	
INP	335500	
IP	424333	
ONP	49600	
OP	416000	

  

<b>B</b>		
Comparison	P values	
INP vs IP	0.758	
INP vs ONP	0.003	
IP vs OP	0.598	
ONP vs OP	0.004	

662

663 NPI= No prep indoor; NPO=No prep outdoor, PI= Prep indoor, PO= Prep outdoor

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679 Figure1 : (A)Bray-Curtis PCoA plot of milk and environmental samples, (B) Bray-Curtis Network plot  
680 of milk and environmental samples. SourceSink indicates if a sample is classified as a potential  
681 source of contamination or a sink for contaminating communities. ENV\_dif indicated the sample  
682 origin

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707 Figure2: SourceTracker results highlight the percentages of inferred sources of contamination in  
708 BTM and individual milk samples  
709  
710



