Environmental sources along natural cave ripening shape the microbiome and metabolome of artisanal blue-veined cheeses

Elena Alexa

Universidad de León

José F Cobo-Díaz

Universidad de León

Erica Renes

Universidad de León

Tom F O'Callaghan

University College Cork

Kieran Kilcawley

Teagasc - The Irish Agriculture and Food Development Authority

David Mannion

Teagasc

Iwona Skibinska

Teagasc

Lorena Ruiz

CSIC

Abelardo Margolles

CSIC

Paula Fernández-Gómez

Universidad de León <https://orcid.org/0000-0001-9647-5251>

Adrián Álvarez-Molina

Universidad de León

Fiona Crispie

Teagasc - The Irish Agriculture and Food Development Authority

Paula Puente-Gómez

Universidad de León

Mercedes López

Universidad de León

Miguel Prieto

Universidad de León

Paul Cotter

Teagasc - The Irish Agriculture and Food Development Authority <https://orcid.org/0000-0002-5465-9068>

Avelino Alvarez-Ordóñez (✉ aalvo@unileon.es)
aalvo@unileon.es

Article

Keywords: cheese microbiome, ripening in caves, environmental microbiome. metabolomics, metagenomics,

Posted Date: August 11th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1150288/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

---

1 **Environmental sources along natural cave ripening shape the microbiome and metabolome of**
2 **artisanal blue-veined cheeses**

3 Elena A. Alexa^{*1}, José F. Cobo-Díaz^{*1}, Erica Renes^{*1}, Tom F. O'Callaghan², Kieran Kilcawley³, David
4 Mannion³, Iwona Skibinska³, Lorena Ruiz^{4,5}, Abelardo Margolles^{4,5}, Paula Fernández-Gómez¹, Adrián
5 Alvarez-Molina¹, Paula Puente-Gómez¹, Fiona Crispie^{3,7}, Mercedes López^{1,6}, Miguel Prieto^{1,6}, Paul D.
6 Cotter^{3,7,8}, Avelino Alvarez-Ordóñez^{1,6}

7
8 ¹ Department of Food Hygiene and Technology, Universidad de León, León, Spain

9 ² School of Food and Nutritional Sciences, University College Cork, Cork, T12 Y337, Ireland

10 ³ Teagasc Food Research Centre, Fermoy, Co. Cork, Ireland

11 ⁴ Dairy Research Institute, Spanish National Research Council, Instituto de Productos Lácteos de Asturias-
12 CSIC, Villaviciosa, Spain

13 ⁵ MicroHealth Group, Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Oviedo, 33011
14 Asturias, Spain

15 ⁶ Institute of Food Science and Technology, Universidad de León, León, Spain

16 ⁷ APC Microbiome Ireland, University College Cork, Cork, Ireland

17 ⁸ VistaMilk, Ireland

18
19 ^{*}The authors equally contributed to the work.

20
21
22
23
24
25 **Keywords:** cheese, microbiome, metabolome, processing environments, source tracking

26
27
28 Correspondence should be addressed to A. Alvarez-Ordóñez. Phone: 0034 987291274; e-mail:
29 aalvo@unileon.es

30 ABSTRACT

31

32 **Background:** Microorganisms colonising processing environments can significantly impact food quality and
33 safety. Cheese ripened in natural caves usually contains a highly diverse microbial community which can
34 contribute to the final product's specific flavor and taste through the production of secondary metabolites.
35 Here we describe a detailed longitudinal study assessing the impact of cave ripening on the microbial
36 succession and cheese metabolome across different producers of *Cabrales* blue-veined cheese.

37 **Results:** Both the producer and cave in which cheeses were ripened significantly influenced the cheese
38 microbiome and metabolome. *Lactococcus* and the former *Lactobacillus* genus, among other taxa, showed
39 high abundance in cheeses at initial stages of ripening, either coming from the raw material, starter culture
40 used and/or the environment of processing plants. Along cheese ripening in caves, these taxa were displaced
41 by other bacteria, such as *Tetragenococcus*, *Corynebacterium*, *Brevibacterium*, *Yaniella* and *Staphylococcus*,
42 predominantly originating from cave environments (mainly food contact surfaces), as demonstrated by
43 source tracking analyses, strain analysis at read level and the characterization of 613 metagenome assembled
44 genomes, which also allowed identifying new species from various bacterial taxa. *Tetragenococcus koreensis*
45 and *T. halophilus* were detected at high abundance in cheese for the first time using whole metagenomic
46 sequencing, associated with the occurrence of various metabolites. Furthermore, *Tetragenococcus* showed a
47 high level of horizontal gene transfer with other members of the cheese microbiome, mainly with
48 *Lactococcus* and *Staphylococcus*, involving genes related to carbohydrate metabolism functions, indicating
49 that these transfer events may have mediated the adaptation of *Tetragenococcus* to the dairy environments.

50 **Conclusion:** Overall, we demonstrated that processing environments can be a source of non-starter
51 microorganisms of relevance to ripening of artisanal fermented foods. Our study highlights that cave
52 environments represented an important source of non-starter microorganisms with a relevant role in the
53 ripening of these artisanal blue-veined cheeses, and identifies among them novel taxa and taxa not previously
54 regarded as being dominant components of the cheese microbiome (*Tetragenococcus* spp.), providing very
55 valuable information to the authentication of this Protected Designation of Origin artisanal cheese in the
56 future.

57

58 **Keywords:** cheese microbiome, ripening in caves, environmental microbiome. metabolomics,
59 metagenomics,

60 BACKGROUND

61 Artisanal cheeses, including those which are recognised in the EU with Protected Designation of Origin
62 (PDO), are highly valuable premium dairy products due to their unique sensory attributes relative to mass-
63 produced cheeses. Their quality, shelf-life and safety are dependent on a number of factors, but are especially
64 influenced by the diverse microbial communities they harbor, which vary from the core to the rind of the
65 cheese and evolve throughout the cheese-making and ripening process [1,2]. Moreover, they can reflect
66 differences among producers or regions in a manner that can be applied to verify the authenticity of PDO
67 products [3].

68 Despite the substantial efforts devoted to study microbial succession patterns during the production and
69 ripening of a wide variety of cheeses, certain aspects relating to microbial interactions and their functional
70 implications remain largely unknown [4]. Historically, the microbiome of artisanal cheeses has been
71 characterized using culture-based methods, which have several limitations [3,4]. High throughput sequencing
72 methodologies (*e.g.*, amplicon sequencing or shotgun whole metagenome sequencing) have transformed the
73 way we study microbial ecology in complex ecosystems, such as those of artisanal cheeses [3,5–8], as they
74 provide access to difficult-to-culture or viable but not culturable microorganisms [1,6], and facilitate the
75 mechanistic understanding of community assembly in such a way that allows studying temporal dynamics of
76 microbial communities at an unprecedented level of detail [6]. Nevertheless, there is also a general consensus
77 that metagenomic studies need to be complemented to, for example, investigate the microbial and chemo-
78 sensory relationships from a multi-omics perspective [9].

79 Some previous studies have highlighted that cheese and other fermented foods can serve as useful models for
80 elucidating the determinants of microbial succession and interactions in complex communities, including
81 understanding the molecular processes and the key metabolically active microbes that contribute to the
82 development of unique sensorial characteristics [6,7,10]. In the present work the microbiome-metabolome
83 interplay in *Cabrales* cheese, a PDO artisanal Spanish blue-veined cheese, was studied. *Cabrales* cheese is
84 produced in the mountainous area of “*Picos de Europa*” (*Principado de Asturias*, Northern Spain) from raw
85 cow, sheep and/or goat milk. This cheese is ripened for up to 5 months in natural caves above 800 meters
86 altitude, with >90% humidity and a temperature ranging from 6°C to 10°C, where it develops its unique
87 intense flavor, creaminess texture, and blue-veins from the rind to the core [11]. It is our hypothesis that the
88 unique microenvironments in these natural caves promote the transfer from the environment to the cheese of
89 specific bacteria, yeasts and molds that impart distinctive quality attributes to the ripened blue-veined
90 cheeses. To test this hypothesis, we undertook a large-scale longitudinal study investigating the structure and
91 functional potential of microbial communities of cheeses (cores and rinds) during the ripening process in
92 natural caves by strategically combining classic microbiological methods, whole-metagenome shotgun
93 sequencing and metabolomic analyses. Through this approach, here we provide a cutting-edge insight into
94 the relationships between traditional cheese-making environments, microorganisms prevailing in cheese and
95 the occurrence of metabolites affecting cheese quality and safety (*e.g.*, volatile compounds and biogenic
96 amines).

97 RESULTS

98 Shared characteristics and patterns of succession

99 The experimental setup included cheeses from three different producers who used the same cave for
100 ripening. In addition, for one producer, cheeses derived from a single production batch were ripened in three
101 different caves, in order to identify microbiome signatures associated with the respective cave micro-
102 environments (Fig. S1).

103 Regardless of producer or cave used for ripening, cheeses maintained certain shared characteristics and
104 microbial succession patterns. Before cheeses entered the ripening caves (Stage1; 30 days post-manufacture)
105 they showed a mild acidic pH (5.41-6.58), and some degree of drying (water activity of 0.828-0.924).
106 *Cabrales* cheeses are routinely rind-washed throughout ripening in the caves, where the relative humidity is
107 >90%. Consequently, significant ($p<0.05$) increases in cheese water activity, relative humidity and pH, and
108 decreases in dry matter and salt content occurred thereafter, marked mostly on cheese rinds at intermediate
109 (Stage2) and final (Stage3) stages of ripening (Fig. S2A). Total fat and protein contents remained fairly
110 stable through ripening (Fig. S2A).

111 Forty-two volatile compounds, belonging to eight chemical classes (Table S1), and seven biogenic amines
112 were detected. While the profile of volatile compounds and biogenic amines was highly variable depending
113 on the producer and ripening cave (as described in detail below), a global view shows that the total content in
114 volatile compounds increased on cheese rinds as ripening proceeded while no significant changes were
115 observed for the cheese cores (Fig. S2B). The biogenic amines content significantly increased with ripening
116 time for both core and rind samples (Fig. S2B).

117 Regarding the microbiome composition, counts of total aerobic bacteria, lactic acid bacteria,
118 *Enterobacteriaceae*, and yeasts and molds reached maximum levels at Stage1 and then gradually declined
119 during ripening, both for cheese cores and rinds (Fig. S3A). At Stage1, cheeses were dominated by lactic
120 acid bacteria, mainly *Lactococcus* and members of the former *Lactobacillus* genus, and by the yeast
121 *Debaryomyces* and the fungi *Penicillium* and *Geotrichum* (Fig. S3B). Whereas fungal populations remained
122 fairly stable along ripening, a significant decrease was evident in the relative abundance of *Lactococcus* and
123 *Lactobacillus*, among others (Fig. S3B), while other taxa, such as *Tetragenococcus*, *Brevibacterium* and
124 *Corynebacterium* emerged with high relative abundances at Stage2 and Stage3, especially at rind level (Fig.
125 S3B).

126 Metabolome and microbiome differences by producer

127 A principal component analysis, performed to reflect all of the volatile compounds and biogenic amines,
128 identified that the producer variable had a significant influence on the metabolomic profile of cheeses, both
129 for the core and rind (adonis, $p=0.001$), with the first two components explaining 80.40 % and 74.80 % of
130 the total variation, respectively (Fig. 1A,B). Focusing on the eight most abundant volatile compounds, some
131 trends were apparent; higher levels of 2-heptanone, 2-nonanone, 2-heptanol and 2-pentanone were found in
132 cheeses from Producer3; higher concentrations of ethyl hexanoate and ethyl butanoate were found in
133 Producer2 cheeses; and the abundance of hexanoic and butanoic acids were lower for Producer3 cheeses

134 (Fig. 1C, Table S2). Regarding amines, tryptamine and tyramine were more abundant on Producer2 and
135 Producer3 cores and cadaverine on Producer1 rinds (Fig. 1D, TableS2).

136 **Figure 1. Metabolome of blue-veined cheeses during ripening in the communal cave.**
137 Biplots of the principal component analysis of metabolites (including biogenic amines and volatiles) found in
138 cheese core **(A)** and cheese rind **(B)** samples ripened in the communal cave. The biplots show the most
139 significant metabolites contributing to the metabolome of cheeses ripened in Cave2 (communal cave) by the
140 three producers. Colored concentration ellipses (size determined by a 0.95-probability level) show the
141 observations grouped by producer. Contrib (Contribution) indicates the average value of percentages of
142 contribution to the 2 first axis for each variable.
143 Barplots presenting volatile compounds **(C)** Absorbance units (AU) (quantitatively minor volatile
144 compounds were grouped as “Other volatiles”); and biogenic amines content **(D)**. Each bar represents the
145 mean value (n=3) for blue-veined cheese samples (cores and rinds separately) belonging to Cave2
146 (communal cave) for each Producer (Producer1, Producer2 and Producer3) during three ripening stages
147 (Stage1, Stage2 and Stage3).

149 The cheese producer also had a significant influence on the bacterial and fungal taxonomic profile of
150 samples (Fig. 2A,B, Fig. S4A,B), both for cheese cores (adonis, $p=0.001$ for both bacteria and fungi) and
151 rinds (adonis, $p=0.001$ for bacteria and $p=0.018$ for fungi). At cheese core level, several bacterial genera
152 within the main ones such as *Lactobacillus*, *Corynebacterium*, and *Enterococcus*, among others, were
153 associated with Producer2, while *Lactobacillus* and *Tetragenococcus* were predominantly linked to
154 Producer3 (Fig. 2A,C, Table S2). At cheese rind level, several bacterial genera were clearly associated with
155 Producer2 (e.g., *Corynebacterium* and *Tetragenococcus*) while *Lactococcus* was associated with Producer1
156 and Producer3 (Fig. 2B,C). Regarding fungi, core samples were highly abundant in *Penicillium*, but, for
157 Producer2, similar relative abundances of *Geotrichum* and *Penicillium* were observed. At rind level,
158 *Penicillium* and *Debaryomyces* showed high abundances in samples from Producer1 and Producer2, and
159 *Debaryomyces* was the dominant genus for Producer3 (Fig. 2D, Table S2).

160 **Figure 2. Microbiome of blue-veined cheeses during ripening in the communal cave.**

161 Biplots of the principal component analysis of the forty most abundant bacterial genera found in the cheese
 162 core (A) and cheese rind (B) samples ripened in the communal cave depending on the Producer. The biplots
 163 show the most significant bacterial genera contributing to the microbiome of cheeses ripened in Cave2
 164 (communal cave) by the three producers. Colored concentration ellipses (size determined by a 0.95-
 165 probability level) show the observations grouped by producer. Contrib (Contribution) indicates the average
 166 value of percentages of contribution to the 2 first axis for each variable.

167 Barplots representing the relative abundance of the 19 most relevant bacterial (C) and yeasts and molds (D)
 168 genera. Other minority bacterial and fungal genera are grouped into “Other”. Each bar represents the average
 169 value (n=3) for blue-veined cheese samples from each Producer.

170

171

172 **Metabolome and microbiome differences by ripening cave**

173 For Producer2, the ripening cave also had a significant influence on the metabolomic profile of cheese cores
174 and rinds (adonis, $p=0.001$), with the first two components explaining 71.3 % and 73.2 % of the total
175 variation, respectively (Fig. 3A,B). Focusing on the eight most abundant volatile compounds, Cave3 cheeses
176 differed most due to lower levels of ethyl hexanoate and ethyl octanoate, and higher levels of 2-heptanone on
177 cores and rinds and hexanoic acid, butanoic acid and octanoic acid on rinds (Fig. 3C, Table S3). Regarding
178 amines, tyramine was more abundant in Cave2 cores, tryptamine in Cave2 cores and Cave3 rinds, and
179 cadaverine in Cave1 cores and rinds (Fig. 3D, Table S3).

180 **Figure 3. Metabolome of blue-veined cheeses from Producer2 during ripening in the three different**
181 **caves.** Biplots of the principal component analysis of metabolites (including biogenic amines and volatiles)
182 found in cheese core **(A)** and cheese rind **(B)** samples ripened by Producer2 in the three different caves. The
183 biplots show the most significant metabolites contributing to the metabolome of cheeses in the three different

184 caves. Colored concentration ellipses (size determined by a 0.95-probability level) show the observations
185 grouped by cave. Contrib (Contribution) indicates the average value of percentages of contribution to the 2
186 first axis for each variable. Barplots presenting volatile compounds **(C)** Absorbance units (AU)
187 (quantitatively minor volatile compounds were grouped as “Other volatiles”) and biogenic amines content
188 **(D)**. Each bar represents the mean value (n=3) for blue-veined cheese samples (cores and rinds separately)
189 before cave ripening (Stage1) and during ripening (Stage2 and Stage3) in the three different natural Caves
190 (Cave1, Cave2, Cave3).

191

192 The cave also had a significant influence, although lower than that found for the variable producer, on the
193 taxonomic profile of cheese cores (adonis, $p=0.007$ for bacteria and $p=0.049$ for fungi) and rinds (adonis,
194 $p=0.027$ for bacteria and 0.653 for fungi) (Fig. 4A,B, Fig. S4C,D). The most relevant differences among
195 caves in the abundance of some particular taxa were those observed for *Lactococcus* and *Lactobacillus*,
196 which were more abundant on core samples from Cave1 and Cave3, respectively, at Stage3;
197 *Tetragenococcus*, *Corynebacterium* and *Staphylococcus*, which were more abundant on rind samples from
198 Cave1, Cave2 and Cave3, respectively (Fig. 4A,C, Table S3); and *Debaryomyces*, which was significantly
199 more abundant on cheese rinds from Cave1 than on those from the other two caves (Fig. 4D, Table S3).

200 **Figure 4. Microbiome of blue-veined cheeses from Producer2 during ripening in the three different**
 201 **caves.** Biplot of the principal component analysis of the forty most abundant bacterial genera found in the
 202 cheese core **(A)** and cheese rind **(B)** samples ripened in the three different caves. The biplots show the most
 203 significant bacterial genera contributing to the cheese microbiome in the three different caves. Colored
 204 concentration ellipses (size determined by a 0.95-probability level) show the observations grouped by cave.
 205 Contrib (Contribution) indicates the average value of percentages of contribution to the 2 first axis for each
 206 variable. Barplots representing the relative abundance of the 19 most relevant bacterial **(C)** and yeasts and
 207 molds **(D)** genera. Other minority bacterial and fungal genera are grouped into “Other”. Each bar represents
 208 the average value (n=3) for blue-veined cheese samples in each Cave.

210 The cave microbiome strongly shapes the rind cheese microbiome

211 The characterisation of the microbial communities prevailing in some primary sources that could determine
 212 the cheese microbiome (e.g., milk, curd, different factory processing environments and cave environments)

213 showed that milk was dominated by *Pseudomonas*, followed by *Acinetobacter*, *Staphylococcus* or
214 *Lactococcus*, at different levels depending on the producer (Fig. S5). Similar profiles but with a higher
215 abundance of some lactic acid bacteria, such as *Lactococcus*, were found in curd samples (Fig. S5).
216 *Brevibacterium*, *Psychrobacter*, *Pseudomonas*, *Acinetobacter*, *Penicillium*, *Debaryomyces* and
217 *Kluyveromyces* were the dominant genera on food processing environments (both food contact and non food
218 contact) of the three cheese producing plants (Fig. S5). *Brevibacterium*, *Corynebacterium* and
219 *Debaryomyces* were the dominant genera in cave food contact surfaces, while *Penicillium* and a wide range
220 of non-dominant bacterial and fungal genera prevailed on the caves' non food contact surfaces (Fig. S5).
221 A source tracking analysis revealed that the bacterial composition of cheese cores, and of cheese rinds at
222 Stage1, was mainly determined, for Producer1 and Producer2, by the curd microbiota, and for Producer3 by
223 the microbiota of food contact environments from the processing plant (Fig. 5). The main bacterial genera
224 traced back to the curd and dairy food contact surfaces were *Lactococcus*, *Lactobacillus* and, in the case of
225 Producer2, *Hafnia* (Fig. 5). Over subsequent stages of ripening, the curd microbiota still represented an
226 important source for the cheese core microbiome, although the cave environment had an important role in
227 shaping the microbiome of cheese cores on Cave2 at Stage2 and Stage3, mainly being a source of
228 *Tetragenococcus* and *Corynebacterium* for Producer1 and Producer3, and *Corynebacterium* for Producer2
229 (Fig. 5). Other less abundant genera (*e.g.*, *Brachybacterium*, *Alkalibacterium*, *Staphylococcus*) were also
230 traced back to the cave environments. Regarding cheese rinds, their bacterial microbiota at Stage2 and
231 Stage3 were traced back mainly to the cave environments, which were revealed as a likely source of
232 *Brevibacterium*, *Corynebacterium* and *Tetragenococcus* (on Stage3) and other minority genera (on Stage2),
233 for Producer1; *Tetragenococcus* in Cave1 and Cave3, together with *Corynebacterium* in Cave2, for
234 Producer2; and *Tetragenococcus*, *Brevibacterium* and other minority genera (on Stage2) and
235 *Tetragenococcus* (on Stage3) for Producer3 (Fig. 5).

Figure 5. Taxonomical (bacteria) source attribution of cheese samples calculated by SourceTracker2 software. Piechart plots represent bacterial sources (on columns, *source* samples) for the cheese bacterial community composition (on rows, *sink* samples). The ratio of the piechart is proportional to the percentage of *source* sample influence on *sink* sample (indicated on the y-axis). The colors within each piechart indicate the percentage of genus influence for each source-sink pair. Only *source* samples and the 16 main genera with significant influence were represented, while other genera were grouped as “Other”.

A source tracking analysis undertaken on the fungal communities showed that food contact surfaces from Cave2 were a source of *Debaryomyces* in rind samples from Producer1 and Producer3, and milk was a source of *Geotrichum* in core samples from Producer2 (Fig. S6).

Strain-level analyses confirmed the relevance of cave environments as a source of bacteria

A total of 110 high-quality metagenome assembled genomes (MAGs), 227 medium quality MAGs with completeness > 90%, and 276 medium-quality MAGs with completeness of 50-90% were obtained. They were classified into 70 genera, with the main genera represented being *Lactococcus* (103 MAGs), *Corynebacterium* (74 MAGs), the former *Lactobacillus* (68 MAGs), *Tetragenococcus* (54 MAGs), *Bifidobacterium* (30 MAGs), *Staphylococcus* (28 MAGs), *Brevibacterium* (28 MAGs) and *Yaniella* (22 MAGs). Cheese samples were the ones with the highest number of MAGs (413 out of 613 MAGs), while raw materials, processing environments at factory level and cave environments yielded 43, 74 and 83 MAGs, respectively (Fig. 6A, Fig. S7).

MAGs from several putative new species were obtained on samples from cave surfaces or cheeses. These included 14 *Brevibacterium* sp. MAGs most closely related to *B. sandarakinum*, 8 *Corynebacterium* sp. MAGs related to *C. nuruki*, 6 *Corynebacterium* sp. MAGs related to *C. glyciniphilum*, 2 *Psychrobacter* sp. MAGs and one *Tetragenococcus* sp. MAG. closely related to *T. koreensis*. Remarkably, despite the low

relative abundance of *Yaniella* found at read level, a total of 22 MAGs from this genus, closely related to the recently discovered Candidatus *Yaniella excrementarium*, were obtained (Fig. 6A, Fig. S7). Furthermore, 3 MAGs assigned to the former genus *Lactobacillus*, obtained from cheese core samples from Producer3 at Stage1, were not assigned at species level, with *L. selangorensis* and *L. camelliae* being the most closely related species (Fig. 6A).

Figure 6. Phylogenetic tree of MAGs obtained by per-sample assembly and binning of metagenomic reads. (A) Phylogenetic tree constructed with the ANI distance matrix generated by using the dRep software. Concentric circles indicate, from inside to outside, the quality of MAGs; producer, cave and surface origin of the sample; genus and species classification of the MAGs; and importance of the MAGs. Only the 24 most abundant taxa at genus level (or above) are indicated, while the rest are grouped in “Other”. Some taxa are classified in higher taxonomic levels (phylum, family or order) according to the deep level of taxonomic assignment obtained by using the CAT/BAT software. Indicated species were taxonomically re-assigned by building additional phylogenetic trees employing reference genomes (RefSeq database) or other representative genomes from the NCBI. The last circle indicates clusters of interest for cave and producer effects on the microbiome. Only new species within the genera of interest (*Brevibacterium*, *Corynebacterium*, *Lactobacillus*, *Tetragenococcus* and *Yaniella*) were marked with red color on the “species” circle. **(B)** Principal Coordinate Analysis of the MAGs functional composition by level 3 of KEGG Orthology classification. Only functional groups belonging to 09100 Metabolism were employed. Only MAGs of interest due to their importance related to the producer and/or cave influence were employed.

It is worth noting that MAGs assigned to *Lactococcus*, a dominant genus at Stage1, presented phylogenetic differences by producer, with *Lactococcus lactis* MAGs from Producer2 clustering separately from MAGs from other producers, which was also observed in the strain-level StrainPhlan population genomics analysis (Fig. S8). Moreover, *L. piscium/carnosum* MAGs were exclusively found for Producer2 and *L. raffinolactis*

283 and *L. garvieae* for Producer3 (Fig. 6A). On the contrary, the phylogenetic tree corresponding to the strain-
284 level StrainPhlan population genomics analysis detected *L. raffinolactis* on both Producer2 and Producer3,
285 with two clear clusters (Fig. S8). On the other hand, different closely related MAGs from *Corynebacterium*
286 *casei*, *Staphylococcus equorum*, *Brevibacterium* sp., *Tetragenococcus halophilus*, *T. koreensis* and *Yaniella*
287 sp. were obtained from both cheese and cave environmental samples. Moreover, in the case of cheese MAGs
288 from *Brevibacterium*, *Tetragenococcus* and *Yaniella*, they were only obtained from cheeses at Stage2 and
289 Stage3 (Fig. 6A, Fig. S7,S8). Similar results were observed in the strain-level StrainPhlan population
290 genomics analyses, with the absence of these genera at Stage1, while they were within the main ones found
291 at Stage2 and Stage3 (Fig. S9). It was also interesting to see that two different clusters of *S. equorum*, with
292 differentiated functional potential, were detected colonizing food processing environments within factories
293 and cheeses at Stage1, and cave environments and cheeses at Stage2 and Stage3, respectively (Fig. 6A, Fig.
294 S10).

295 The comparison of the MAGs from *T. koreensis* and *T. halophilus* with the available genomes from the NCBI
296 showed a clear separation of *T. koreensis* genomes from PDO *Cabrales* (current study) and *Picón Bejes-*
297 *Tresviso* cheese [12] from others from non-dairy food isolates (thick juice, soy sauce, fermented or salted
298 seafood, fermented vegetables), with similar observations for *T. halophilus* (Fig. S11).

299 The analysis of the functional potential of MAGs evidenced that *Lactococcus*, *Lactobacillus*, *Staphylococcus*
300 and *Tetragenococcus* had a relatively similar functional profile, which was characterised by the high
301 abundance of functions related to the metabolism of carbohydrates, while *Corynebacterium*, *Brevibacterium*
302 and *Yaniella* MAGs had a very different functional profile, characterised by the high abundance of different
303 functions mainly related to protein and fatty acids metabolism (Fig. 6B).

304

305 **The cheese microbiome is associated with the cheese metabolome**

306 Correlation analyses undertaken with respect to cheese samples from stages 2 and 3 of ripening revealed that,
307 among the species for which MAGs were generated, at core level, *Lactococcus* (*L. lactis*, *L. piscium*)
308 *Lactobacillus* (mainly *L. paracasei*) and *Bifidobacterium* (*B. crudilactis*, *B. mongoliense*) were those that had
309 the strongest positive correlations with the abundance of various volatile compounds and biogenic amines
310 (Fig. 7A). The most remarkable associations found were for *L. lactis* with pentanoic and hexanoic acids; *L.*
311 *piscium* with 3-methyl-butanal, propanoic acid, ethyl-butanoate, propyl-butanoate, trimethylpyrazine, ethyl-
312 octanoate, putrescine and cadaverine; *L. paracasei* with 3-methyl-butanal, propyl-butanoate,
313 trimethylpyrazine and isobutyl-decanoate; *B. crudilactis* with 3-methyl-butanal and trimethylpyrazine; and
314 *B. mongoliense* with trimethylpyrazine, isoamyl-hexanoate, isopentyl-octanoate, isobutyl-decanoate and
315 tryptamine (Fig. 7A,B). At rind level, these species also showed positive correlations with the abundance of
316 certain volatiles and amines, but other representative taxa, traced back to the cave environment, such as
317 *Tetragenococcus* (mainly *T. halophilus*), *Staphylococcus equorum*, and *Corynebacterium* (*C. flavesens*, *C.*
318 *variabile*) were also strongly associated with some cheese volatile and amine markers (Fig. 7B). The most

319 remarkable associations were found for *T. halophilus* with propyl-butanoate, 2-heptanol, trimethylpyrazine,
 320 2-pentylbutyrate, 2-nonanone, 2-nonanol, ethyl-octanoate, 2-undecanone, ethyl-decanoate, isobutyl-
 321 decanoate, ethyl dodecanoate, putrescine and tyramine; *S. equorum* with 2-heptanone, 2-heptanol, 2-
 322 octanone, 2-nonanol, γ -hexalactone, 2-undecanone, isopenthyloctanoate, isoamyl-hexanoate and putrescine;
 323 and *C. variabile* and *C. flavescentis* with pentanoic and hexanoic acids (Fig.7B). Interestingly, both at cheese
 324 core and cheese rind level, with the exception of *Brevibacterium* and the former *Lactobacillus* genus, species
 325 from the same genera showed highly different microbiome-volatilome correlation patterns while species
 326 belonging to the same genera, with the exception of *Corynebacterium* and *Tetragenococcus*, had similar
 327 correlations with amine concentrations (Fig. 7).

328 **Figure 7. Spearman correlation between abundance of metabolites and bacterial species.** Heatmaps
 329 representing significant ($p < 0.05$) Spearman correlations between the abundance of volatiles and biogenic
 330 amines analyzed and species abundance at read level for cheese core (A) and cheese rind (B) samples. Only
 331 those bacterial species of special interest with metagenomic assembled genomes (MAGs) obtained by using
 332 the metaSPAdes + METABAT2 pipeline were employed for correlation analyses. Cheese samples from
 333 Stage1 (before cave ripening) were discarded for correlation analyses.

335   **The cheese microbiome is rich in Horizontal Gene Transfer events**

336   In order to identify among the main bacterial taxa cues of microbiome acclimatisation to cave and cheese
337   microenvironments, signs of horizontal gene transfer (HGT) events were sought in the available MAGs. A
338   total of 23,001 HGT events, containing 67,411 coding regions, were detected between 56 taxa above species
339   level, with *Lactococcus*, *Tetragenococcus* and *Staphylococcus* being the main genera involved (59.3, 52.0
340   and 23.3% of total HGT events, respectively), and *Lactococcus-Tetragenococcus*, *Staphylococcus-*
341   *Tetragenococcus* and *Lactococcus-Staphylococcus* as the main HGT pairs (29.6, 10.9 and 9.4% of total HGT
342   events, respectively) (Fig. 8A). HGT events were identified from MAGs belonging to all surfaces sampled,
343   but cheese core and rind samples at stages 2 and 3 of ripening contributed the most (56.6% of HGT events)
344   (Fig. 8B).

Figure 8. Horizontal Gene Transfer (HGT) events. HGT events (identity > 99.9%, length > 500 bp) detected between MAGs belonging to different genera (**A**) and classified by surface, producer and cave (**B**). The thickness of the circle internal lines is proportional to the number of HGT events detected for each genera pair. The “Themselves” group corresponds with HGT events between different genera from the same surface, producer or cave group. (**C**) Percentage of coding regions (CDS) detected within HGT sequences belonging to different functions of the level2 and level3 of KEGG Orthology functional classification. Only 45% of CDS detected, with KO code assigned, were represented. (**D**) Relation between the length of HGT events and the number of CDS detected in them. Circle colors indicate the genera pair for each HGT event, while the size of the circles is square-root proportional to the number of HGT events with the same length and CDS number.

Up to 35.3% of the HGT events were associated with plasmid sequences, according to plasflow analysis, while only 0.5% carried relaxase encoding genes, associated with plasmid mobilization. Moreover, 11.7, 8.1 and 0.9% harboured prophages, transposases and integrases, respectively. Although 55.5% of the coding regions could not be assigned to ko codes of the KEGG Orthology (KO) database, 10.6, 7.2 and 5.4% were assigned at level2 to the groups *Protein families: signaling and cellular processes*, *Carbohydrate metabolism* and *Protein families: genetic information processing*, respectively. At level3 of KO classification, *Prokaryotic defense system*, *Galactose metabolism*, *Glycolysis/Gluconeogenesis*, *Purine metabolism* and *Fructose and mannose metabolism* were within the dominant functions (Fig. 8C). Finally, 68 HGT-associated contigs were longer than 20kb, containing from 12 to 45 coding regions (CDS) (Fig. 8D), and clustered into 26 groups. *Lactococcus-Tetragenococcus* and *Alkalibacterium-Tetragenococcus* were the dominant pairs within these longest HGT events, with 14 and 4 clusters, respectively (Table S4). Half of the obtained HGT-clusters contained genes related to mobile genetic elements such as plasmids, transposases, phages and integrons; 3 clusters harboured genes related to β -lactams resistance, and up to 9 clusters contained genes related to protein, carbohydrate or lipid metabolism (Table S4).

369

370 DISCUSSION

Our study found that two different groups of microorganisms prevailed in *Cabrales* PDO cheese. The first group comprises various taxa (e.g., *Lactococcus*, *Lactobacillus*, *Bifidobacterium*, *Penicillium*, *Geotrichum*, *Debaryomyces*) that show a high abundance in the cheeses from the initial stages of production, prior to the cheeses entering the ripening caves. These microorganisms very likely originate from the raw materials, the starter cultures used and/or the dairy plant processing environments. This group of microbes, with the exception of *Debaryomyces*, occurred at high abundances in cheese cores over the whole ripening period and were associated with the occurrence of various volatile compounds in the cheese core, highlighting in some cases intraspecific phylogenetic diversity depending on the cheese producer. However, as ripening progressed within the cave this group of microbes was gradually displaced, especially on cheese rinds, by the second group of microbes, which comprises several taxa (e.g., *Tetragenococcus*, *Corynebacterium*, *Brevibacterium*, *Yaniella*, *S. equorum*) not found (or present at very low abundance) in milk, curd, or environments within the processing plant and cheeses at the initial stages of ripening. These microbes are associated with the occurrence of some volatile compounds in rinds, linked to unique aroma notes, as discussed below in greater detail, and are very likely sourced within the cave environments, as evidenced by SourceTracker predictions and the close genomic relationships among some strains, such as some representatives of *S. equorum*, *T. korrensis*, *T. halophilus* and *C. casei*, as revealed by the analysis of MAGs from cheese and cave samples.

The role of some members of the *Cabrales* cheese microbiome in the ripening process is clear based on previous studies for this sort of cheeses. Indeed, *Lactococcus*, *Lactobacillus*, *Staphylococcus*, *Corynebacterium*, *Brevibacterium*, *Penicillium*, *Geotrichum* or *Debaryomyces* have all previously been

391 detected at high abundance in blue-veined rind washed cheeses, and in some cases also in cheese processing
392 environments [5,6,13], with very clear contributions to cheese fermentation. Moreover, some of these taxa,
393 such as *Corynebacterium* or *Brevibacterium*, play a particularly important role in aroma development during
394 cheese ripening [5,13,14]. In our study, new strong microbiome-volatilome associations were uncovered for
395 these taxa. Indeed, *Lactococcus* showed strong positive correlations with the abundance of carboxylic acids,
396 which impart cheese, rancid, pungent, sharp, sweaty, malty/chocolate, fruity or buttery flavour notes,
397 depending on strain to strain variability [15]; and *L. piscium*, *B. crudilactis* and *L. paracasei* strongly
398 correlated with the occurrence of propyl butanoate and ethyl octanoate, responsible for sweet, fruity, and
399 floral notes [16]; and of 3-methyl butanal, responsible for the generation of malty, cheesy, dark-chocolate
400 notes [16].

401 On the other hand, other taxa of the *Cabrales* cheese microbiome have received much less attention in the
402 past or have been detected at high relative abundance in cheese in our study for the first time. Specifically,
403 this is the case of one low abundant species of *Yaniella* and two species from *Tetragenococcus* (*T. koreensis*
404 and *T. halophilus*), found in very high abundance in the latest stages of *Cabrales* cheese ripening. Although
405 the presence of *Yaniella* in food related environments has been recently described [2,17], its relevance during
406 cheese ripening remains unclear. Here we describe a total of 22 *Yaniella* sp. MAGs (8 of them from cave
407 environments), closely related to the recently discovered *Candidatus Yaniella excrementarium*, for which
408 further research is needed to understand its contribution to the quality and safety of *Cabrales* cheese.
409 *Tetragenococcus* has been isolated from Brie cheese [18], two Mexican cheeses [19] and, very recently, from
410 two Spanish traditional blue-veined cheeses [12]. Only *T. koreensis* strains have been very recently isolated
411 from *Cabrales* cheese, while we are reporting both *T. koreensis* and *T. halophilus*, with different core-rind
412 distribution, in the current study. These two *Tetragenococcus* species here detected have a functional profile
413 similar to that of some traditional lactic acid bacteria (e.g., *Lactococcus*, *Lactobacillus*) and possibly displace
414 them due to their better acclimatization to the microenvironments associated with our study.
415 *Tetragenococcus* are halo-alkaliphilic bacteria and the relatively high salt content and progressive increase in
416 pH along ripening, especially at rind level, can provide ideal conditions to promote their growth [20].
417 Remarkably, the dominance of one of the two *Tetragenococcus* species over the other depended on both the
418 producer and ripening cave, and when both *Tetragenococcus* species co-occurred, *T. halophilus*, with a
419 higher optimum pH for growth of 6.1-7.0 [18], showed tropism for the cheese rind and *T. koreensis*, with an
420 optimum pH for growth of 5.7-6.5 [18], for the cheese core. Interestingly, these taxa were associated with the
421 co-occurrence of various metabolites in the cheese, which suggests that they significantly contribute to some
422 of the quality attributes of this distinctive cheese. Specifically, *T. koreensis* was associated with a high
423 content of carboxylic acids in core cheese samples and *T. halophilus* showed strong positive correlations
424 with some ketones (2-nonanone, 2-undecanone), secondary alcohols (2-nonanol, 2-heptanol), and esters
425 (propyl-butanoate, 2-pentylbutyrate, ethyl-octanoate, ethyl-decanoate, ethyl dodecanoate, isobutyl-
426 decanoate, among others), which originate from β -oxidation of fatty acids and/or transamination and
427 esterification reactions and are responsible for strong musty and blue cheese notes [16]. Moreover, the

428 identification of these taxa in cheeses from a very restricted geographical region suggests that they may be
429 used as microbiome markers for PDO authentication purposes.

430 Finally, the results provide evidence of high levels of horizontal gene transfer among bacteria belonging to
431 different genera, suggesting that these events can mediate strain adaptation to the cheese/cave
432 microenvironments, as has been previously proposed for cheese rind microbiomes [21]. A relatively high
433 proportion of such HGT-associated sequences harboured mobile genetic elements. Likewise, high levels of
434 phage and transposase associated genes have been previously found on MAGs obtained from Swiss Gruyère
435 cheese [22]. Remarkably, extensive HGT was observed for *Tetragenococcus* with *Lactococcus* and
436 *Staphylococcus*, and the fact that many associated genes were related to carbohydrate metabolism functions
437 (e.g., galactose metabolism, glycolysis/gluconeogenesis, and fructose and mannose metabolism) indicates
438 that these transfer events may have mediated the adaptation of *Tetragenococcus* to the cheese dairy
439 environments. Furthermore, the high diversity of lactose metabolism genes and operons previously found in
440 *Tetragenococcus* [12,18] and the presence of mobilization elements and sequences with high nucleotide
441 identity to *Streptococcus* and *Staphylococcus* surrounding those loci suggest the likely acquisition by
442 *Tetragenococcus* through HGT of these clusters from other species from the dairy environment [18],
443 considering also the high frequency of HGT events in cheese microbiomes [21].

444 CONCLUSIONS

445 Overall, our study highlights that cave environments represent an important source of non-starter
446 microorganisms with a relevant role in the ripening of artisanal blue-veined cheeses, and identifies among
447 them various novel taxa and taxa not previously regarded as being dominant components of the cheese
448 microbiome (*Tetragenococcus* spp.), providing very valuable information to the authentication of this PDO
449 artisanal cheese in the future.

450 MATERIALS AND METHODS

451 *Cheese production, sampling strategy and sample collection*

452 *Cabrales* cheese is the most popular traditional blue-veined cheese in Spain and has a Protected Designation
453 of Origin (PDO) status since 1981. This cheese is produced in "Picos de Europa" (a region in Northern Spain
454 with altitudes above 800 meters) from raw cow, sheep or goat milk or a mixture of two or all three types of
455 milk. The traditional cheese-making process of *Cabrales* cheese involves curdling mixtures of evening and
456 morning milk at 28 - 30°C using rennet exclusively of animal origin. Curds are cut to hazelnut grain size and
457 placed in cylindrical moulds at room temperature, being turned upside down several times in order to drain
458 them off without applying pressure. Then, dry salt is added to the cheese surface and cheeses are taken to a
459 ripening chamber where they remain at room temperature for 15-30 days. After this time, cheeses are placed
460 in natural caves, which are characterized for being deep, with a north-facing entry, and for having at least
461 two openings with flowing water in order to create an airstream. These conditions lead to very high humidity
462 (>90%), with temperatures ranging from 6°C to 10°C. Cheeses must remain in the caves for two to five
463 months and during this time they are periodically turned and washed.

464 In each of the three cheese processing facilities the following samples were taken for analysis: 200 mL of
465 milk and curd samples (three replicates) during the cheese-making process, swab samples from the
466 processing environments entering in contact with the cheese (e.g., work tables, cheese vat, cheese moulds),
467 swab samples from non-food contact surfaces (e.g., drains, floors, walls), and three cheeses (2.0 kg each)
468 taken 30 days after cheese making, just before leaving the factory to the ripening cave/s.

469 In the natural caves, the following samples were taken: swab environmental samples (food contact and non-
470 food contact surfaces, including drains if available), and three cheeses (2.0 kg each cheese) for each of the
471 producers at both the intermediate (90 to 120 days of ripening) and final (130 to 190 days of ripening)
472 ripening stages.

473 Environmental samples were collected by swabbing using sterile sponges pre-moistened with 10 mL of
474 neutralizing buffer, which were placed individually in sterile bags (3M, Minnesota, USA).

475 Appropriate personal protective equipment and gloves were used during all samplings to avoid cross-
476 contamination. At each collection point, the samples were placed in a cooling box containing ice packs and
477 transported to the laboratory within two to three hours. A detailed description of the samples taken is
478 provided in Fig. S1.

479 Upon arrival at the lab, the cheese rinds were scraped with a sterile knife until the cheese paste was visible
480 (≈ 0.7 cm). A different sterile knife was used to prepare cheese core samples in order to avoid any possible
481 cross contamination from the rind.

482

483 ***Culture-based microbiological analyses***

484 For microbiological analyses, 10 g of cheese (rind and core, separately), milk or curd were homogenized
485 with 90 mL of Buffered Peptone Water (BPW; Merck, Germany) for 2 min in a Stomacher 400 Lab Blender
486 (Seward Medical, London, UK), while 10 mL of BPW was directly added to the sterile bags containing swab
487 environmental samples. Serial decimal dilutions were then prepared in BPW and spread in triplicate on the
488 following media: (i) Plate Count Agar (PCA; Pronadisa, Spain) for mesophilic bacteria, incubated for 24±2 h
489 at 37°C; (ii) De Man Rogosa Sharpe agar (MRS; Merck) for lactic acid bacteria (LAB), incubated for 24±2 h
490 at 30°C; (iii) Oxytetracycline Glucose Yeast Extract Agar (OGYEA; Pronadisa) for yeasts and molds,
491 incubated for 5 days at 25°C; and (iv) Violet Red Bile Glucose Agar (VRBGA; Pronadisa) for
492 *Enterobacteriaceae*, incubated for 24±2 h at 37°C. The results were expressed as means and standard
493 deviations of log CFU (colony-forming units) per gram from three independent replicates.

494 ***Proximate analysis***

495 Milk samples were analyzed for their content in total solids, fat, protein, lactose and somatic cells by using
496 Milkoscan FT2 (Foss Electric, Hillerod, Denmark).

497 pH and titratable acidity of milk, curd and blue-veined cheese samples were determined according to AOAC
498 International (1990b) methods.

499 In the case of cheese rind and cheese core samples, total solids, fat, protein and salt contents were determined
500 according to the International standards 004 (FIL-IDF 2004), 221 (FIL-IDF 2008), 20-1 (FIL-IDF 2001), and
501 935.43 (AOAC 1990a), respectively. Water activity (*a_w*) was analyzed instrumentally using an Aqua Lab
502 Dew Point Analyzer CX-2 (Decagon Devices, Pullman, WA, USA). All samples were analysed in duplicate.

503 ***Biogenic amines analysis***

504 Biogenic amines in cheese samples were determined following the method described by Redruello et al. [23],
505 as follows: one gram of each cheese sample was homogenized with 10 mL of 0.1 M HCl-0.2% 3,3
506 'thiodipropionic acid (Merck, Madrid, Spain) and then centrifuged at 5000 *g* for 20 min. This mixture was
507 kept in an ultrasonic bath Bransonic 221 (Branson Ultrasonics S.A, Danbury, USA) for 30 min and then
508 centrifuged at 5000 *g* for 20 min. The resultant supernatant was deproteinized by passing it through
509 ultrafiltration inserts (Amicon Biomax 5K; Merck) and centrifuging at 3500 *g* for 60 min. Then, 20 µL were
510 derivatized with diethyl ethoxymethylenemalonate (DEEMM; Merck). L-2-aminoadipic acid (Merck) was
511 used as an internal standard. The chromatography system consisted of an H-Class Acquity UPLCTM system
512 (Waters, Massachusetts, USA) coupled to a photodiode array detector. The separation of biogenic amines
513 was carried out using a Waters Acquity UPLCTM BEH C18 column (1.7 µm particle size, 100 mm × 2.1 mm
514 I.D.) held at 35°C. The mobile phase was made up of 25 mM acetate buffer pH 6.7 plus 0.02 % sodium azide
515 (eluent A), methanol (eluent B) and acetonitrile (eluent C). One µL sample was injected and the flow rate
516 was set up at 0.45 mL/minute according to the linear gradient used by Redruello et al. [23]. The target

517 compounds were identified by their retention times and their spectral characteristics at 280 nm, and were
518 quantified using the internal standard method. Data were acquired and analyzed using the software Empower
519 2 (Waters). Each sample was analysed in duplicate.

520 ***Volatile compounds analysis***

521 Volatile-compounds profiling of the blue-veined cheese samples was carried out using the headspace solid-
522 phase microextraction gas chromatography-mass spectrometry method (HS-SPME-GC-MS) as described by
523 Panthi et al. [24], with some modifications. Briefly, 0.5 g of each cheese sample was placed in a 20-mL
524 amber screw-capped headspace vial with a silicone/PTFE liner (Apex Scientific Ltd., Co. Kildare, Ireland).
525 Extraction of volatile compounds was performed using a single SPME fiber 50/30 μ m
526 Divinylbenzene/Carboxen/Polydimethylsiloxane (DVB/CAR/PDMS; Agilent Technologies Ireland Ltd.,
527 Cork, Ireland), which was exposed to the HS for 20 min at a depth of 21 mm at 40°C under the condition of
528 pulsed agitation, 5 s ON and 2 s OFF at 350 rpm, using a Shimadzu AOC-5000 injection system (Mason
529 Technology, Dublin, Ireland). The fiber was retracted, injected via a merlin microseal, in splitless mode, into
530 the inlet of a Shimadzu 2010 Plus GC (Mason Technology) and desorbed for 2 min at 250°C. Separation of
531 volatile compounds was carried out using a DB-624 UI column (1.80 μ m particle size, 60 m x 0.32 mm I.D.;
532 Agilent Technologies Ltd.). Helium was used as a carrier gas at a constant flow of 1.2mL/min. The
533 temperature of the column oven was set as follows: the initial temperature was 40°C, then increased at 5°C/
534 min to 230°C (15 min hold), and further increased at 15°C/minute to 260°C (5 minutes hold), yielding a total
535 GC run time of 65 min. The detector was a Shimadzu TQ8030 MSD triple quadrupole mass spectrometer
536 (Mason Technology) and it was used in single quadrupole mode. The ion source temperature was 220°C, the
537 interface temperature was set at 260°C, and the MS mode was electronic ionization (70 V), with the mass
538 range scanned between 35 and 250 amu. A set of external standards were run at the start and end of the
539 sample set and abundances were compared to known amounts to ensure that both the SPME extraction and
540 MS detection were performing within specifications. In addition, autotune of the GCMS was performed
541 before the analysis to ensure optimal GC-MS performance. Post-analysis processing of the obtained total ion
542 chromatograms was carried out using GCMSsolution (Shimadzu, Japan) and batch processing of samples
543 was performed using R software (RStudio. Inc., Boston, MA). Volatile compounds were identified by: i)
544 mass spectrum comparison to the National Institute of Standards and Technology (NIST) 2011 mass spectral
545 library, ii) the automated mass spectral deconvolution and identification system (AMDIS), iii) an internal
546 library created in GC-MS Solutions software (Shimadzu, Japan) with target and qualifier ions using linear
547 retention indices as described by van Den Dool and Kratz [25], and vi) linear retention indexes matching
548 against peer-reviewed publications, where possible, to confirm compound identification. Each cheese sample
549 was analysed in triplicate.

550 ***Whole metagenome sequencing analyses***

551 ***DNA extraction***

Prior to DNA extraction, curd, cheese, and milk samples were prepared by homogenising 10 g or mL in 90 ml of BPW, and centrifuging at 6500 *g* for 8 min. Then, pellets were washed with 50 mL of BPW, followed by a new centrifugation at 6500 *g* for 8 min to harvest the associated microbiota. For the environmental samples, pools of five swabs from each sample category were mixed with 10 mL of BPW. After thorough homogenisation, the BPW solutions were centrifuged at 6500 *g* for 8 min. Pellets from all samples were kept at -80 °C until further use. Total metagenomic DNA was extracted from the pellets by using the DNeasy PowerSoil PRO kit (Qiagen GmbH, Germany) following the manufacturer's instructions, but conducting a double elution step with 25 µL of 10 mM Tris-HCl, in order to improve DNA yields.

Library construction and shotgun sequencing

Extracted DNA was employed to prepare 150 bp paired-end sequencing libraries using the Illumina Nextera XT Library Preparation Kit (Illumina Inc., San Diego, CA). Sequencing was performed on the Illumina NextSeq 500 platform using a NextSeq 500/550 High Output Reagent kit v2 (300 cycles), in accordance with the standard Illumina sequencing protocols.

Reads quality filtering and annotation

Quality filtering of raw reads was performed with AfterQC v0.9.6 [26] using default parameters. Reads matching to cow and sheep genomes were removed using a combination of bowtie2 v2.3.4.1 [27], samtools v1.9 [28] and bedtools v2.29.0 (<https://bedtools.readthedocs.io/en/latest/#>) according to a pipeline previously published (<http://www.metagenomics.wiki/tools/short-read/remove-host-sequences>). The taxonomic assignment of filtered reads was done using Kraken2 v2.0.8-beta [29] with the maxikraken2-1903_140GB database (https://lomanlab.github.io/mockcommunity/mc_databases.html) for bacterial taxonomy, while a custom database with kraken2-microbial database plus 20 reference fungal genomes from species found often in cheese environments (Table S5) was employed for fungal taxonomy assignment, due to the absence of major fungal species on the maxikraken2-1903_140GB database.

Reads assembly and MAGs binning

Each sample was independently subjected to *de-novo* metagenomic assembly through metaSPAdes v3.13 [30] using default parameters. Filtered reads were mapped against contigs higher than 1000 bp obtained from the same sample using bowtie2 v2.3.4.1 [27] with the *--very-sensitive-local* parameter. The *jgi_summarize_bam_contig_depths* script, from MetaBAT2 v2.12.1 [31], was used to calculate contigs depth values, mandatory for per-sample contig binning based on tetranucleotide frequency and the contig abundance protocol, which was performed using MetaBAT2 and the option '*-m 1500*' [31].

The quality of MAGs was estimated using the lineage-wf workflow of CheckM (v1.1.3) [32] and cmseq (<https://github.com/SegataLab/cmseq>), to classify MAGs as high quality MAGs (>90% completeness, <5% contamination, <0.5 % non-synonymous mutations) and medium quality MAGs (>50 completeness, <5% contamination) for further analysis.

586 **MAGs analysis**

587 MAGs were assigned to taxonomy groups using the CAT/BAT pipeline v.5.1.2 [33] and
588 CAT_prepare_20200618 database. Those MAGs assigned to the bacterial genera *Lactococcus*, *Lactobacillus*,
589 *Tetragenococcus*, *Staphylococcus*, *Corynebacterium*, *Brevibacterium* and *Yaniella* were taxonomically
590 assigned at species level using as reference the available genomes from these genera on the NCBI
591 (<ftp://ftp.ncbi.nlm.nih.gov/genomes/>), which were selected and downloaded using the download_genomes
592 pipeline (https://github.com/JoseCoboDiaz/download_genomes). The MAGs and NCBI genomes for each
593 selected genus were employed for the calculation of Average Nucleotide Identity (ANI) values by using
594 dRep v2.6.2 [34]. Taxonomic assignment at species level was done by manual inspection of ANI
595 phylogenetic trees obtained by dRep software. Moreover, the *Mdb.txt* output file obtained from all ANI
596 calculations was transformed into a distance matrix, which was employed for the construction of a
597 phylogenetic tree plot by *ggtree* R-package, using the UPGMA clustering algorithm (*hclust*
598 *method="average"*).

599 The functional annotation of MAGs with completeness higher than 90% was performed by using eggNOG-
600 mapper v2.1.5 [35] and the MMseqs algorithm [36]. The “KEGG_ko” column from the eggNOG-mapper
601 output was employed to build the count matrix of KOs per MAG, where the information of the 3 hierarchical
602 levels of the KEGG database were added, and only those KOs belonging to *09100 Metabolism* were kept for
603 further analysis.

604 Additionally, all *Tetragenococcus* spp. genomes available at the NCBI (at 20th June, 2022) were downloaded
605 to build phylogenetic trees, using the previously explained pipeline by dRep and *ggtree*, in order to confirm
606 that new species (or at least not whole genome sequenced species) were detected, and to compare *T.*
607 *koreensis* and *T. halophilus* genomes with those obtained in the current study.

609 **Strain-level analysis through an assembly-free approach**

610 A strain-level population genomics analysis was undertaken for species of special interest, according to the
611 analysis of MAGs, using the StrainPhlan pipeline [37,38]. Prior to StrainPhlan analysis, metaphlan2 [39] was
612 run for each sample, using the paired option as input and the bowtie2out command to obtain the alignment
613 files in .sam format. The sample2makers.py script was employed to obtain the markers files for each sample,
614 and extract_markers.py was used to extract clade markers for each species of interest, for further StrainPhlan
615 analysis, according to the software designers (<https://github.com/biobakery/biobakery/wiki/strainphlan1>),
616 and using as reference the corresponding species reference genome from the NCBI (Table S6). Phylogenetic
617 trees were plotted by *ggtree* R-package, using the UPGMA clustering algorithm (*hclust method="average"*)
618 and the *RaxML_bestTree* output file from StrainPhlan analysis.

620 **Detection of Horizontal Gene Transfer (HGT) events**

621 A custom blast database was built with all MAGs obtained. A Blastn analysis of the MAGs fasta files,
622 containing all MAG sequences correctly labeled on the headers to discriminate to which MAGs each contig
623 belonged, was performed against the MAGs database with a 99.9% identity cut-off. The blastn output was

624 filtered to remove hits where query and subject belonged to the same genus and to remove duplicated hits.
625 Hits with alignment length shorter than 500 bp were also discarded for further analysis. HGT sequences were
626 extracted by bedtools using the query sequences position from the blastn output, and eggNOG-mapper v2.1.5
627 [35] was employed, as previously described for MAGs, to functionally characterise HGT sequences.
628 Detection of CDS related to integrons, transposases or phages was done by searching “integron”,
629 “transposase” and “phage” on the eggNOG-mapper output file, while detection of plasmids was done by
630 searching “relaxase” and by employing PlasFlow (<https://github.com/smaegol/PlasFlow>) as a second
631 approach. The clustering of HGT sequences was done using vsearch [40] and 0.99 as identity cutoff. Chord
632 diagrams for HGT events were plotted using the circlize R-package.

634 ***Diversity and taxonomy plots and statistical analysis***

635 Only those taxa belonging to the kingdoms Bacteria and Fungi were used for further analyses, which were
636 executed in parallel.

637 The β -diversity was estimated by Principal Component Analysis using Bray-Curtis dissimilarity and the
638 *vegdist* function. Within-group dispersion was evaluated through the *betadisper* function. Both functions are
639 located in the R package *vegan* (<https://github.com/vegandevs/vegan>). The effects of cave, producer and
640 sample type (which includes ripening stage and cheese part for cheese samples, and food contact vs non food
641 contact surfaces for environmental samples) on sample dissimilarities were determined by permutational
642 multivariate analysis of variance using distance matrices (PERMANOVA) with the *adonis* function in the R
643 package *vegan* (<https://github.com/vegandevs/vegan>). The *compare_means* function in the R package
644 *ggpubr* was used to include statistically significant differences on boxplot figures, which were plotted by
645 using the R package *ggplot2* (<https://github.com/tidyverse/ggplot2>).

646 Comparisons between multiple groups of samples for taxonomy were performed by using the Kruskal Wallis
647 test and the *post hoc* Wilcoxon signed-rank test. P-values were adjusted through the Benjamini & Hochberg
648 method [41] and significance was established at $p < 0.05$.

649 Correlation analyses were performed using the Spearman correlation coefficient, calculated with the *rcorr*
650 function in the R package *Hmisc* (<https://hbiostat.org/R/Hmisc/>).

651 All analyses were carried out using R version 3.6.2 (<https://cran.r-project.org/>).

653 ***Source tracking analysis***

654 Genera relative abundance matrices were employed to attribute the main sources of the cheese microbial
655 populations by using SourceTracker2 [42] with the *--per_sink_feature_assignments* parameter. The source-
656 to-sink genera contribution dataset was reorganized by using *in-house* scripts
657 (https://github.com/JoseCoboDiaz/piecharts_tax_st2) to plot a matrix of piecharts that indicates the influence
658 of each source to each sink (by the size of the piechart) and the percent influence of each genera in each
659 source-to-sink pair.

663 **List of abbreviations**

664 PDO: Protected designation of origin; MAGs: metagenome assembled genomes; HGT: horizontal gene
 665 transfer; KO: KEGG Orthology; CDS: coding regions; BPW: buffered peptone water; PCA: plate count
 666 agar; MRS: De Man Rogosa Sharpe; LAB: lactic acid bacteria; OGYEA: oxytetracycline glucose yeast
 667 extract agar; VRBGA: violet red bile glucose agar; CFU: colony-forming units; AW: water activity;
 668 DEEMM: diethyl ethoxymethylenemalonate; HS-SPME-GC-MS; headspace solid-phase microextraction gas
 669 chromatography-mass spectrometry method; DVB/CAR/PDMS:
 670 Divinylbenzene/Carboxen/Polydimethylsiloxane; NCBI: National Center for Biotechnology Information;
 671 ANI: average nucleotide identity.

672

673 **Acknowledgements**

674 AAM is grateful to the Ministry of Science, Innovation and Universities of the Spanish Government for
 675 awarding him a pre-doctoral grant (FPI BES-2017-079753). PFG is grateful to Junta de Castilla y León and
 676 the European Social Fund (ESF) for awarding her a pre-doctoral grant (BOCYL-D-15122017-4). B.
 677 Redruello is acknowledged for her assistance in biogenic amines analysis. The authors wish to acknowledge
 678 the cheese producers for their availability and support throughout the study.

679

680 **Authors' contributions**

681 The study was designed by L.R., A.M., M.L., M.P., P.D.C. and A.A.O. Sampling activities were performed
 682 by E.A.A., P.F.G., A.A.M., P.P.G., and A.A.O. Metabolomic procedures were performed by E.R., T.F.O.,
 683 K.K., D.M., I.S., L.R. and A.M. Culture-dependent analyses were performed by E.A.A., P.F.G., A.A.M. and
 684 P.P.G.. DNA isolation activities were performed by E.A.A. Whole metagenome sequencing was done by F.C.
 685 and P.D.C. Reads filtering was performed by J.F.C.D. and further computational and statistical analysis were
 686 performed by E.A.A., J.F.C.D. and E.R. Finally, E.A.A., J.F.C.D., E.R. and A.A.O. participated in the
 687 manuscript writing. All authors contributed to its revision and approved the final version.

688

689 **Funding**

690 The laboratory of A.A.O., M.L. and M.P. have received funding by the Ministry of Science and Innovation of
 691 the Spanish Government, under grant numbers AGL2016-78085-P and PID2020-118813GB-I00, and the
 692 European Commission under the European Union's Horizon 2020 research and innovation program under
 693 grant agreement No 818368.

694

695 **Availability of data and materials**

696 The raw reads fastq files from shotgun metagenome sequencing of the 133 samples and the genome fasta
 697 files from the 613 medium and high-quality MAGs obtained are available under NCBI BioProject ID
 698 PRJNA764004, with physical BioSample accession numbers from SAMN21465692 to SAMN21465824 and

699 SRA accession numbers from SRR15927537 to SRR15927669 for raw reads fastq files, and BioSample
700 accession numbers from SAMN23026855 to SAMN23027467 for MAGs-genome fasta files.

701

702 **DECLARATIONS**

703 **Ethics approval and consent to participate**

704 Not applicable

705 **Consent for publication**

706 Not applicable

707 **Competing interests**

708 The authors declare that they have no competing interests.

709 REFERENCES

710

- 711 1. Yeluri Jonnala BR, McSweeney PLH, Sheehan JJ, Cotter PD. Sequencing of the cheese microbiome and
712 its relevance to industry. *Front Microbiol.* 2018;9:1–12.
- 713 2. Kamilari E, Tomazou M, Antoniadis A, Tsaltas D. High throughput sequencing technologies as a new
714 toolbox for deep analysis, characterization and potentially authentication of protection designation of origin
715 cheeses? *Int J Food Sci. Hindawi*; 2019;2019:1–15.
- 716 3. Penland M, Falentin H, Parayre S, Pawtowski A, Maillard MB, Thierry A, et al. Linking Pélardon artisanal
717 goat cheese microbial communities to aroma compounds during cheese-making and ripening. *Int J Food*
718 *Microbiol.* 2021;345:1–15.
- 719 4. Mayo B, Rodríguez J, Vázquez L, Flórez AB. Microbial interactions within the cheese ecosystem and their
720 application to improve quality and safety. *Foods.* 2021;10:1–28.
- 721 5. Quijada NM, Schmitz-Esser S, Zwirzitz B, Guse C, Strachan CR, Wagner M, et al. Austrian raw-milk
722 hard-cheese ripening involves successional dynamics of non-inoculated bacteria and fungi. *Foods.* 2020;9:1–
723 21.
- 724 6. Wolfe BE, Button JE, Santarelli M, Dutton RJ. Cheese rind communities provide tractable systems for in
725 situ and in vitro studies of microbial diversity. *Cell [Internet]. Elsevier Inc.*; 2014;158:422–33. Available
726 from: <http://dx.doi.org/10.1016/j.cell.2014.05.041>
- 727 7. Bertuzzi AS, Walsh AM, Sheehan JJ, Cotter PD, Crispie F, McSweeney PLH, et al. Omics-Based Insights
728 into Flavor Development and Microbial Succession within Surface-Ripened Cheese. *mSystems.* 2018;3:1–
729 15.
- 730 8. Walsh AM, Macori G, Kilcawley KN, Cotter PD. Meta-analysis of cheese microbiomes highlights
731 contributions to multiple aspects of quality. *Nat Food [Internet]. Springer US*; 2020;1:500–10. Available
732 from: <http://dx.doi.org/10.1038/s43016-020-0129-3>
- 733 9. Afshari R, Pillidge CJ, Dias DA, Osborn AM, Gill H. Cheesomics: the future pathway to understanding
734 cheese flavour and quality. *Crit Rev Food Sci Nutr [Internet]. Taylor & Francis*; 2020;60:33–47. Available
735 from: <https://doi.org/10.1080/10408398.2018.1512471>
- 736 10. Walsh AM, Crispie F, Kilcawley K, O'Sullivan O, O'Sullivan MG, Claesson MJ, et al. Microbial
737 Succession and Flavor Production in the Fermented Dairy Beverage Kefir. *mSystems.* 2016;1:1–16.
- 738 11. Flórez AB, Mayo B. Microbial diversity and succession during the manufacture and ripening of
739 traditional, Spanish, blue-veined Cabrales cheese, as determined by PCR-DGGE. *Int J Food Microbiol.*
740 2006;110:165–71.
- 741 12. Rodríguez J, González-Guerra A, Vázquez L, Fernández-López R, Flórez AB, de la Cruz F, et al.
742 Isolation and phenotypic and genomic characterization of *Tetragenococcus* spp. from two Spanish traditional
743 blue-veined cheeses made of raw milk. *Int J Food Microbiol.* 2022;371:1–11.
- 744 13. Wadhawan K, Steinberger AJ, Rankin SA, Suen G, Czuprynski CJ. Characterizing the microbiota of
745 wooden boards used for cheese ripening. *JDS Commun [Internet]. Elsevier Inc.*; 2021;2:171–6. Available
746 from: <http://dx.doi.org/10.3168/jdsc.2020-0014>

747 14. Frétin M, Martin B, Rifa E, Isabelle VM, Pomiès D, Ferlay A, et al. Bacterial community assembly from
748 cow teat skin to ripened cheeses is influenced by grazing systems. *Sci Rep*. 2018;8:1–11.

749 15. Thierry A, Valence F, Deutsch SM, Even S, Falentin H, Le Loir Y, et al. Strain-to-strain differences
750 within lactic and propionic acid bacteria species strongly impact the properties of cheese—A review. *Dairy*
751 *Sci Technol*. 2015;95:895–918.

752 16. Bertuzzi AS, McSweeney PLH, Rea MC, Kilcawley KN. Detection of Volatile Compounds of Cheese
753 and Their Contribution to the Flavor Profile of Surface-Ripened Cheese. *Compr Rev Food Sci Food Saf*.
754 2018;17:371–90.

755 17. Salazar JK, Carstens CK, Ramachandran P, Shazer AG, Narula SS, Reed E, et al. Metagenomics of
756 pasteurized and unpasteurized gouda cheese using targeted 16S rDNA sequencing. *BMC Microbiol. BMC*
757 *Microbiology*; 2018;18:1–13.

758 18. Unno R, Matsutani M, Suzuki T, Kodama K, Matsushita H, Yamasato K, et al. Lactic acid bacterial
759 diversity in Brie cheese focusing on salt concentration and pH of isolation medium and characterisation of
760 halophilic and alkaliphilic lactic acid bacterial isolates. *Int Dairy J* [Internet]. Elsevier Ltd; 2020;109:1–13.
761 Available from: <https://doi.org/10.1016/j.idairyj.2020.104757>

762 19. Morales F, Morales JI, Hernández CH, Hernández-Sánchez H. Isolation and partial characterization of
763 halotolerant lactic acid bacteria from two mexican cheeses. *Appl Biochem Biotechnol*. 2011;164:889–905.

764 20. Justé A, Lievens B, Frans I, Marsh TL, Klingenberg M, Michiels CW, et al. Genetic and physiological
765 diversity of *Tetragenococcus halophilus* strains isolated from sugar- and salt-rich environments.
766 *Microbiology*. 2008;154:2600–10.

767 21. Bonham KS, Wolfe BE, Dutton RJ. Extensive horizontal gene transfer in cheese-associated bacteria.
768 *Elife*. 2017;6:1–23.

769 22. Somerville V, Lutz S, Schmid M, Frei D, Moser A, Irmeler S, et al. Long-read based de novo assembly of
770 low-complexity metagenome samples results in finished genomes and reveals insights into strain diversity
771 and an active phage system. *BMC Microbiol. BMC Microbiology*; 2019;19:1–18.

772 23. Redruello B, Ladero V, Cuesta I, Álvarez-Buylla JR, Martín MC, Fernández M, et al. A fast, reliable,
773 ultra high performance liquid chromatography method for the simultaneous determination of amino acids,
774 biogenic amines and ammonium ions in cheese, using diethyl ethoxymethylenemalonate as a derivatising
775 agent. *Food Chem* [Internet]. Elsevier Ltd; 2013;139:1029–35. Available from:
776 <http://dx.doi.org/10.1016/j.foodchem.2013.01.071>

777 24. Panthi RR, Sundekilde UK, Kelly AL, Hennessy D, Kilcawley KN, Mannion DT, et al. Influence of herd
778 diet on the metabolome of Maasdam cheeses. *Food Res Int* [Internet]. Elsevier; 2019;123:722–31. Available
779 from: <https://doi.org/10.1016/j.foodres.2019.05.026>

780 25. Dool H Van den, Kratz PD. A generalization of the retention index system including linear temperature
781 programmed gas-liquid partition chromatography. *J Chromatogr*. 1963;II:463–71.

782 26. Chen S, Huang T, Zhou Y, Han Y, Xu M, Gu J. AfterQC: Automatic filtering, trimming, error removing
783 and quality control for fastq data. *BMC Bioinformatics*. 2017;18:91–175.

784 27. Langmead B, Wilks C, Antonescu V, Charles R. Scaling read aligners to hundreds of threads on general-

785 purpose processors. *Bioinformatics*. 2019;35:421–32.

786 28. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format
787 and SAMtools. *Bioinformatics*. 2009;25:2078–9.

788 29. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol. Genome*
789 *Biology*; 2019;20:1–13.

790 30. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. MetaSPAdes: A new versatile metagenomic
791 assembler. *Genome Res*. 2017;27:824–34.

792 31. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: An adaptive binning algorithm for
793 robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;2019:1–13.

794 32. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: Assessing the quality of
795 microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25:1043–55.

796 33. Von Meijenfeldt FAB, Arkhipova K, Cambuy DD, Coutinho FH, Dutilh BE. Robust taxonomic
797 classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol. Genome*
798 *Biology*; 2019;20:1–14.

799 34. Olm MR, Brown CT, Brooks B, Banfield JF. DRep: A tool for fast and accurate genomic comparisons
800 that enables improved genome recovery from metagenomes through de-replication. *ISME J [Internet]*.
801 *Nature Publishing Group*; 2017;11:2864–8. Available from: <http://dx.doi.org/10.1038/ismej.2017.126>

802 35. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2:
803 Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol Biol*
804 *Evol*. 2021;38:5825–9.

805 36. Mirdita M, Steinegger M, Breitwieser F, Söding J, Levy Karin E. Fast and sensitive taxonomic
806 assignment to metagenomic contigs. *Bioinformatics*. 2021;37:3029–31.

807 37. Truong DT, Tett A, Pasolli E, Huttenhower C, Segata N. Microbial strain-level population structure &
808 genetic diversity from metagenomes. *Genome Res*. 2017;27:626–38.

809 38. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic,
810 functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife*. 2021;10:1–
811 42.

812 39. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for enhanced
813 metagenomic taxonomic profiling. *Nat Methods*. 2015;12:902–3.

814 40. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: A versatile open source tool for
815 metagenomics. *PeerJ*. 2016;2016:1–22.

816 41. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to
817 Multiple Testing. *J R Stat Soc Ser B*. 1995;57:289–300.

818 42. Dan K, Kuczynski J, Charlson ES, Zaneveld J, Mozer MC, Collman RG, et al. Bayesian community-
819 wide culture-independent microbial source tracking. *Nat Methods*. 2013;8:761–763.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.pdf](#)
- [FigureS2.pdf](#)
- [FigureS3.pdf](#)
- [FigureS4.pdf](#)
- [FigureS5.pdf](#)
- [FigureS6.pdf](#)
- [FigureS7.pdf](#)
- [FigureS8.pdf](#)
- [FigureS9.pdf](#)
- [FigureS10.pdf](#)
- [FigureS11.pdf](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)
- [TableS4.xlsx](#)
- [TableS5.xlsx](#)
- [TableS6.xlsx](#)