Identification and characterization of lactic acid bacteria and yeasts of PDO Tuscan bread sourdough by culture dependent and independent methods

Michela Palla, Caterina Cristani, Manuela Giovannetti, Monica Agnolucci*

Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy

RUNNING HEAD: Lactic acid bacteria and yeasts of PDO Tuscan bread sourdough

*Corresponding author: M. Agnolucci, Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy

Phone: +39.0502216647, Fax: +39.0502220606, e-mail address: monica.agnolucci@unipi.it

Highlights

- The microbiota of PDO Tuscan bread sourdough was isolated and molecularly identified.
- A distinctive tripartite microbial association characterized the PDO Tuscan bread sourdough.
- A large number of *Lactobacillus sanfranciscensis* (98) and *Candida milleri* (65) strains were isolated.
- Among yeasts, *Candida milleri* dominated over *Saccharomyces cerevisiae.*

Abstract

Sourdough fermentation has been increasingly used worldwide, in accordance with the demand of consumers for tasty, natural and healthy food. The high diversity of lactic acid bacteria (LAB) and yeast species, detected in sourdoughs all over the world, may affect nutritional, organoleptic and technological traits of leavened baked goods. A wide regional variety of traditional sourdough breads, over 200 types, has been recorded in Italy, including special types selected as worthy of

either Protected Geographical Indication (PGI) or Protected Designation of Origin (PDO), whose sourdough microbiota has been functionally and molecularly characterized. As, due to the very recent designation, the microbiota of Tuscan bread sourdough has not been investigated so far, the aim of the present work was to isolate and characterize the species composition of LAB and yeasts of PDO Tuscan bread sourdough by culture-independent and dependent methods. A total of 130 yeasts from WLN medium and 193 LAB from both mMRS and SDB media were isolated and maintained to constitute the germplasm bank of PDO Tuscan bread. Ninety six LAB from mMRS medium and 68 yeasts from WLN medium were randomly selected and molecularly identified by ARDRA (Amplified Ribosomal DNA Restriction Analysis) and PCR-RFLP analysis of the ITS region, respectively, and sequencing. The yeast identity was confirmed by 26S D1/D2 sequencing. All bacterial isolates showed 99% identity with *Lactobacillus sanfranciscensis*, 65 yeast isolates were identified as *Candida milleri*, and 3 as *Saccharomyces cerevisiae*. Molecular characterization of PDO Tuscan bread sourdough by PCR-DGGE confirmed such data. The distinctive tripartite species association, detected as the microbiota characterizing the sourdough used to produce PDO Tuscan bread, encompassed a large number of *L. sanfranciscensis* and *C. milleri* strains, along with a few of *S. cerevisiae*. The relative composition and specific physiological characteristics of such microbiota could potentially affect the nutritional features of PDO Tuscan bread, as suggested by the qualitative functional characterization of the isolates. Investigations on the differential functional traits of such LAB and yeast isolates could lead to the selection of the most effective single strains and of the best performing strain combinations to be used as starters for the production of baked goods.

Keywords: ARDRA, ITS-RFLP, PCR-DGGE, *Lactobacillus sanfranciscensis*, *Candida milleri*, *Saccharomyces cerevisiae.*

1. Introduction

Sourdough fermentation represents one of the oldest biotechnologies used by humans to ferment cereals for bread production. In recent times, it has been increasingly used worldwide, in accordance with the demand of consumers for tasty, natural and healthy foods (Gobbetti and Gänzle, 2013). Indeed, sourdough fermentation improves bread sensory, structural and nutritional properties, and prolongs its shelf-life (Arendt et al., 2007; Katina et al., 2005; Minervini et al., 2014). In addition, microbial metabolism during sourdough fermentation positively affects several bread functional features, producing bioactive compounds, such as peptides, beta-glucans and other exopolisaccharides (Gobbetti et al., 2014).

A large number of lactic acid bacteria (LAB) and yeast species, establishing positive interactions and often stable associations, has been isolated from sourdoughs all over the world, *i.e.* about 80 bacterial and 20 yeast species (De Vuyst and Neisens, 2005; Gänzle and Ripari, 2016). Notwithstanding, only a few species characterize single batches of sourdough, where no more than 6 different species have been usually identified. The typical most commonly detected bacterial species are *Lactobacillus brevis, Lactobacillus pontis, Lactobacillus reuteri, Lactobacillus plantarum* and *Lactobacillus sanfranciscensis*, the latter representing the dominant member of the microbiota, as it has been isolated in more than 75% of sourdoughs globally (Ganzle and Ripari, 2016). The most common sourdough yeast is *Saccharomyces cerevisiae*, although, as reviewed by De Vuyst et al. (2016) other species may occur in spontaneously developed stable sourdoughs (in decreasing order of abundance): *Candida humilis/Candida milleri, Wickerhamomyces anomalus, Torulaspora delbrueckii, Kazachstania exigua, Pichia kudriavzevii and Candida glabrata.* Some of these species, such as *K. exigua, C. humilis* and *C. milleri*, are maltose-negative yeasts reported to form stable mutualistic associations with *L. sanfranciscensis,* which is able to hydrolyze maltose (De Vuyst et al., 2014).

Several studies investigating the microbiota of sourdough from different countries, *e.g.* Belgium (Scheirlinck et al., 2007), France (Ferchichi et al., 2008; Vera et al., 2012), Turkey (Dertli et al., 2016), China (Liu et al., 2016; Zhang et al., 2011), showed that the diversity of microbial

communities depends on process technologies, types of flour and other ingredients traditionally associated with local culture and origin. Such diversity is at the basis of differential metabolic products, affecting nutritional, organoleptic and technological traits of leavened baked goods. In particular, traditional or type I sourdough is characterized by a spontaneous fermentative process, based on continuous backslopping, carried out by LAB and yeasts originating from the flour, other dough ingredients and the environment. Type I sourdough is utilized to produce various leavened baked products, such as San Francisco bread, French bread, rye bread, Altamura bread (Corsetti, 2013).

A wide regional variety of traditional sourdough breads, over 200 types, has been recorded in Italy (INSOR 2000), including special types selected as worthy of either Protected Geographical Indication (PGI) (*Coppia Ferrarese*, *Pane Casareccio di Genzano*, *Cornetto di Matera*) or Protected Designation of Origin (PDO) (*Pagnotta del Dittaino* and *Pane di Altamura*). The sourdoughs of such traditional regional breads have been functionally and molecularly characterized, *i.e. Cornetto di Matera* (Zotta et al., 2008), *Pane di Altamura* (Ricciardi et al., 2005), along with breads from Abruzzo (Valmorri et al., 2006; 2010), Marche (Osimani et al., 2009), Molise (Reale et al., 2005), Sicily (Pulvirenti et al., 2001) and Sardinia (Catzeddu et al., 2006).

Recently, *Pane Toscano* has obtained the PDO status from the European Community (Commission implementing regulation (EU) 2016/303 of 1 March 2016, Official Journal of the European Union L 58 of 04 March 2016). PDO Tuscan bread has to be manufactured by a typical method generally adopted in Tuscan bakeries. It requires the exclusive use of sourdough starters, water, the absence of added salt, and type '0' soft-wheat flour from wheat varieties grown in Tuscany, according to the production guideline described in the EU Regulation for PDO Tuscan bread (Official Journal of the European Union C 235 of 14 August 2013). Due to the very recent designation, the microbiota of Tuscan bread sourdough has not been investigated so far. The aim of the present work was to isolate and characterize the species composition of lactic acid bacteria and yeasts of PDO Tuscan bread sourdough. To this aim, we utilized *i*) a culture-independent approach,

105 Polymerase Chain Reaction (PCR) Denaturating Gradient Gel Electrophoresis (DGGE), a
106 molecular technique able to avoid underestimates deriving from the constraints of culture conditions
107 and from the presence of microorganisms in Viable But Non-Culturable (VBNC) state; *ii*) a culture-
108 dependent approach to isolate and molecularly identify LAB and yeast species; *iii*) a preliminary
109 qualitative screening to characterize LAB and yeasts with functionally important traits.
110
111 **2. Materials and methods**
112 *2.1. Sourdough sampling*
113 The sourdough analyzed in this study was collected from the Consortium of Promotion and
114 Protection of Tuscan Sourdough Bread - *Consorzio Pane Toscano a Lievitazione Naturale* (CPT).
115
116 *2.2. Microbiological analysis and isolation of LAB and yeasts*
117 Three samples of about 10 g of PDO Tuscan bread sourdough were homogenized in a sterile
118 stomacher bag containing 90 mL of saline-peptone water (9 g/L NaCl, 1 g/L bacteriological
119 peptone, Oxoid, Milan, Italy) for 2 min at 260 rpm, using a Stomacher (Stomacher 400, Laboratory
120 Blender). Further, a tenfold serial dilution (10^{-1} to 10^{-6}) was carried out in the same solution and
121 aliquots (100 µL) were added in triplicate into a Petri dish containing the agar media listed below.
122 LAB were counted on Sourdough Bacteria (SDB) (Kline and Sugihara, 1971) medium and on
123 mMRS agar (de Man, Rogosa, Sharpe, 1960) modified by adding 20 g/L maltose and 50 mL/L fresh
124 yeast extract and adjusted to pH 5.6. To inhibit yeast growth, media were supplemented with 100
125 mg/L cycloheximide. Inoculated plates were incubated for 7 days at 28 °C under anaerobic
126 conditions (AnaeroGen, Oxoid). Yeasts were counted on Wallerstein Laboratory Nutrient (WLN)
127 agar (Oxoid, Basingstoke, UK) and on Yeast Extract Peptone Dextrose (YEPD) agar. Both media
128 were added with 100 mg/L chloramphenicol and incubated at 28 °C for 48 h.
129 LAB were randomly selected picking up at least 20 colonies from each plate of both media
130 and purified by streaking four times onto the same medium used for isolation. About 15 yeast

colonies were randomly selected from each plate of WLN medium on the basis of phenotypic colony characteristics and then purified as described above.

Each strain was named with the acronym of the Collection of the Department of Agriculture, Food and Environment of the University of Pisa (IMA, International Microbial Archives), followed by a progressive number plus "Y" or "LAB" for yeasts or bacteria, respectively. Purified strains were stored at –80 °C in the appropriate broth medium, supplemented with 20% (w/v) glycerol.

2.3. Molecular identification of LAB and yeast isolates

DNA of isolates and reference strains, listed in Table 1, was extracted from microbial liquid cultures grown at 28 °C using "MasterPure™ Yeast DNA Purification Kit" (Epicentre®) according to the manufacturer's protocols. LAB strains were identified by Amplified Ribosomal DNA Restriction Analysis (ARDRA) and yeast strains by ITS region amplification and its Restriction Fragment Length Polymorphism (RFLP) analysis. Amplification reactions were carried out in a final volume of 50 µL, containing 5 µL of 10X Ex Taq Buffer (Takara Biotechnology), 0.2 mM of each dNTP (Takara Biotechnology), 0.5 µM of each primer (Eurofins), 1.25U of Takara Ex Taq polymerase (Takara Biotechnology) and 10-20 ng of DNA. The 16S rRNA gene was amplified using 27f (5'-GAG AGT TTG ATC CTG GCT CAG-3') and 1495r (5'-CTA CGG CTA CCT TGT TAC GA-3') primers (Lane 1991; Weisburg et al., 1991) and ITS region was amplified using ITS1 (5'-TCC GTA GGT GAA CCT GCG G-3') and ITS4 (5'-TCC TCC GCT TAT TGA TAT GC-3') primers (White et al., 1990). PCR amplifications were carried out with an iCycler-iQ Multicolor Real-Time PCR Detection System (Bio-Rad) using the following conditions: 94 °C initial denaturation for 1 min; 35 amplification cycles of 30 s at 94 °C, 30 s at annealing temperature, 45 s at 72 °C; final extension at 72 °C for 5 min. The annealing temperatures, for bacteria and yeasts, were 60 and 55 °C, respectively. The presence of amplicons was confirmed by electrophoresis in 1.5% (w/v) agarose gel stained with ethidium bromide (0.5 µg/mL).

156 LAB 16S rRNA gene amplicons were digested at 37 °C overnight with the restriction
157 endonucleases *Alu*I, *Hinf*I and *Hae*III (BioLabs), while yeast ITS amplicons were digested with
158 *Hae*III and *Hinf*I (BioLabs) enzymes. The restriction fragments were separated (at 50 V for 2 h) on
159 1.8% (w/v) agarose gels stained with ethidium bromide (0.5 µg/mL) in Tris-borate-EDTA buffer
160 (Sigma-Aldrich). A 100 bp DNA ladder (BioLabs) was used as a molecular weight marker. All gels
161 were visualized and captured as TIFF format files by the Liscap program for Image Master VDS
162 system (Pharmacia Biotech). Isolates and reference strains profiles were digitally processed and
163 analysed with BioNumerics software version 7.5 (Applied Maths, St-Martens-Latem, Belgium).
164 ARDRA and ITS-RFLP profiles from all digestions were respectively combined, for LAB and
165 yeasts, and the similarity calculated on the basis of the Dice's coefficient. For cluster analysis,
166 unpaired group method with arithmetic average (UPGMA) trees with highest resampling support, in
167 a permutation sample of size 200, were constructed.
168 The identification of isolates was confirmed by sequencing LAB 16S and yeast ITS1-5.8S-
169 ITS2 and 26S D1/D2 rRNA gene amplicons. The 26S D1/D2 was amplified using the primers NL1
170 and NL4, as reported by Kurtzman and Robnett (1998).
171 In particular, PCR products were purified with the UltraClean PCR CleanUp kit (CABRU)
172 according to the manufacturer's protocol, quantified and 5' sequenced by BMR Genomics (Padova,
173 Italy). Sequences were analyzed using BLAST on the NCBI web
174 (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The related sequences were collected and aligned using
175 MUSCLE (Edgar, 2004a, b), and phylogenetic trees were constructed using the Neighbor-Joining
176 method based on the kimura 2-parameter model (Kimura, 1980) in Mega 6 software (http://
177 www.megasoftware.net/) with 1000 bootstrap replicates. The sequences were submitted to the
178 European Nucleotide Archive under the accession numbers from LT605080 to LT605156 and from
179 LT718652 to LT718656.
180
181 *2.4. Molecular characterization of PDO Tuscan bread sourdough by PCR-DGGE*

DNA was extracted from the three PDO Tuscan bread sourdough samples by using "Power Soil DNA Isolation kit" (MO-BIO Laboratories) according to the manufacturer's protocol and stored at -20 °C until further analyses.

For the analysis of LAB communities, the V3-V4 region of 16S rRNA gene was amplified with the Lac1 (5'-AGC AGT AGG GAA TCT TCC A-3') and Lac2 (5'-ATT YCA CCG CTA CAC ATG-3') primers (Walter et al., 2001). The primer Lac2 had at its 5' end an additional 40-nucleotide GC-rich tail (5'-CGC CCG GGG CGC GCC CCG GGC GGC CCG GGG GCA CCG GGG G-3'). For the analysis of yeasts, an approximately 250 bp long fragment of D1/D2 region of the 26S rRNA gene was amplified using NL1 (5'-GCC ATA TCA ATA AGC GGA GGA AAA G-3') and LS2 (5'-ATT CCC AAA CAA CTC GAC TC-3') primers (Cocolin et al., 2000). An additionally GC clamp (5'-GCG GGC CGC GCG ACC GCC GGG ACG CGC GAG CCG GCG GCG G-3') was added to the forward primer NL1. PCR amplifications were performed as previously described, using 10-20 ng of DNA. Amplification conditions were: 94 °C initial denaturation for 1 min; 35 amplification cycles of 30 s at 94 °C, 30 s at annealing temperature, 30 s at 72 °C; final extension at 72 °C for 5 min. The annealing temperatures for bacteria and yeasts, were 58 and 55 °C, respectively. The presence of amplicons was confirmed by electrophoresis in 1.5% (w/v) agarose gel stained with ethidium bromide (0.5 µg mL^{-1}).

For the DGGE analysis, amplicons were separated in 8% (w/v) polyacrylamide gels with a 35–50% and 30–56% urea-formamide gradient, for bacteria and yeasts, respectively, using the DCode™ Universal Mutation Detection System (BioRad). A composite mix of bacterial 16S rRNA gene fragments and fungal 26S rRNA gene fragments from reference strains (Table 1), were added as bacterial and fungal reference DGGE markers (M). Gels were run and visualized as described in Agnolucci et al. (2013). DGGE fragments were cut out from the gels for sequencing. DNA was extracted by eluting for 3 days in 50 µL 10 mM TE at 4 °C. One µL of the supernatant diluted 1:100 was used to re-amplify the LAB V3-V4 and the yeast D1/D2 regions of the DNA according to the PCR protocol described above, using Lac2 and NL1 primers without GC clamp. PCR products were

than purified by UltraClean PCR CleanUp Kit (MO-BIO Laboratories) according to the manufacturer's protocol, quantified and 5' sequenced at the BMR Genomics (Padova, Italy). Sequences were analyzed using BLAST on the NCBI web and submitted to the European Nucleotide Archive under the accession numbers from LT605157 to LT605165.

2.5. Qualitative functional characterization of LAB and yeast isolates

To test phytase activity, yeasts were grown on Phytate Screening Medium (PSM) (Jorquera et al., 2008). The same medium added with 50 mL/L of fresh yeast extract at pH 5.6 was used for LAB. To eliminate false positive results, plates were counterstained as described by Bae et al. (1999). Protease activity of yeasts and LAB was assessed on YEPD without Peptone and on mMRS, containing 2% skim milk, respectively. Isolates were than incubated at 28 °C for 8 days. After incubation the formation of halo zones around microbial colonies indicated phytase and protease capacity. Yeast amylase activity was tested on YEPD without Dextrose containing 10 g/L soluble starch (Fluka), pH 6.46. LAB amylase activity was screened on mMRS agar without maltose, containing 10 g/L soluble starch (Fluka), pH 5.6. Yeasts and LAB were than incubated for 7 days at 30 °C and 37 °C, respectively, and starch hydrolysis was revealed by the disappearance of the blue colour of the medium around the microbial colonies after addition of Lugol iodine solution (Sigma-Aldrich).

3. Results

3.1. Microbiological analysis of PDO Tuscan bread sourdough

Lactic acid bacteria counts on mMRS and SDB media were $2.33\pm0.54\cdot10^{9}$ CFU/g and $1.03\pm0.22\cdot10^{9}$ CFU/g, respectively. Yeast numbers were approximately 100 fold lower than LAB, showing colony counts of $2.14\pm0.17\cdot10^{7}$ on WLN and $2.21\pm0.37\cdot10^{7}$ CFU$\cdot$g^{-1} on YEPD. A total of 130 yeasts were isolated from WLN medium as such medium allowed the differentiation among colony morphologies, while 193 LAB, showing only one colony morphology, were isolated from

234 both mMRS and SDB media. All the isolates were maintained to constitute the germplasm bank of
235 PDO Tuscan bread. Ninety six LAB from mMRS medium and 68 yeasts from WLN medium were
236 randomly selected and molecularly characterized at species level.
237
238 *3.2. Molecular identification of LAB and yeast isolates*
239 The 96 LAB and 68 yeasts selected were molecularly identified by ARDRA and PCR-RFLP
240 analysis of the ITS region, respectively, and sequencing. The expected 16S rRNA gene amplicon of
241 1468 pb was obtained for all the 96 LAB isolates. The amplified product was then subjected to
242 restriction analysis using the enzyme *Alu*I, *Hae*III and *Hinf*I. *Alu*I produced two different ARDRA
243 profiles consisting of three bands (210, 240 and 620 bp) (52% of isolates) and four bands (210, 240,
244 620 and 870 bp) (48% of isolates). *Hinf*I produced three different ARDRA profiles consisting of six
245 bands (70, 90, 120, 320, 400 and 1000 bp) in 51% of isolates, five bands (70, 90, 120, 320 and 1000
246 bp) in 48% of isolates and seven bands (70, 90, 120, 320, 400, 1000 and 1430 bp) in one isolate.
247 *Hae*III produced only one ARDRA profile of three bands (125, 240, 1330 bp) in all isolates
248 (Supplementary Fig. S1). The dendrogram created by combining the three different ARDRA
249 profiles grouped the isolates into two main clusters with a similarity level of 89.9%. The first one
250 consisted of 49 isolates, which clustered at similarity level of 100% with the reference strain *L.*
251 *sanfranciscensis* DSMZ 20451, and of the isolate IMA 23LAB which separately branched at
252 similarity of 95.2%. The remaining 46 isolates, showing 100% similarity, grouped in the second
253 cluster (Fig. 1). Five isolates from each of the two clusters, along with the isolate IMA 23LAB,
254 were subjected to 16S rRNA gene sequence analysis. All isolates showed 99% identity with *L.*
255 *sanfranciscensis* strain ATCC 27651 (DSMZ 20451) despite the presence of a polymorphism
256 within 16S rRNA gene detected by ARDRA analysis (Table 2).
257 The ITS region of the 68 yeast isolates was successfully amplified. An amplicon of about
258 650 bp, was obtained for 65 isolates and one of about 840 bp for the remaining three isolates (IMA
259 19Y, IMA 36Y and IMA 105Y). Digestion of 650 pb amplicons with *Hinf*I produced two fragments

of 350 and 300 bp in all the 65 isolates and in the reference strains *C. milleri* DBVPG 6753^T, *C. humilis* DBVPG 7219^T and *C. humilis* DBVPG 6754 (Supplementary Fig. S2). Using *Hae*III all the 65 isolates produced the same restriction pattern (425 and 230 bp) as the reference strains *C. milleri* DBVPG 6753^T, while *C. humilis* DBVPG 7219^T type strain showed three fragments of 310, 210 and 110 bp. Conversely, the reference strain *C. humilis* DBVPG 6754 showed two fragments of 425 and 230 bp as obtained for the *C. milleri* DBVPG 6753^T. The other three isolates, producing 840 bp amplicons, showed a profile of 310, 230, 165 and 130 bp with *Hae*III and of 350 and 120 bp with *Hinf*I. The dendrogram, created by combining the size of the ITS amplicons and the different restriction profiles, grouped all the 65 isolates with the reference strains *C. milleri* DBVPG 6753^T and *C. humilis* DBVPG 6754 in a main cluster with a 100% of similarity (Fig. 2). *C. humilis* DBVPG 7219^T branched separately at 73% similarity. In a separate cluster (33.5% similarity) the isolates IMA 19Y, IMA 36Y and IMA 105Y grouped with *S. cerevisiae* ATCC 32167 reference strain at a similarity level of 100%. The 840 bp amplicons were then subjected to restriction analysis by the enzyme *Hpa*II and a profile of 700 and 130 bp, corresponding to that of *S. cerevisiae* (Fernández-Espinar et al., 2000), was obtained. The BLAST analysis of ITS1-5.8S-ITS2 and 26S D1/D2 region sequences confirmed the identity of the three isolates IMA 19Y, IMA 36Y and IMA 105Y as *S. cerevisiae* (Table 2).

To determine the species affiliation of our 65 *Candida* isolates, all 650 bp ITS amplicons were sequenced and examined using BLAST and phylogenetic trees analyses. The dendrogram obtained by the analysis of the ITS sequences of our isolates, and those of *C. humilis* and *C. milleri* reference strains available in GeneBank, is reported in Fig. 3. Results showed that all our ITS sequences grouped with those of *C. milleri* reference strains with 100% identity. The BLAST analysis of 26S D1/D2 region sequences of two representative isolates (IMA 11Y and IMA 33Y) confirmed the identity of our isolates as *C. milleri* (Table 2).

3.3. Molecular characterization of PDO Tuscan bread sourdough by PCR-DGGE

Microbial community diversity of PDO Tuscan bread sourdough was further investigated by PCR-DGGE. The amplification of the V3-V4 region of 16S rRNA gene of LAB and of the partial D1/D2 domain of 26S rRNA gene of yeasts from the three sourdough samples produced a DNA fragment of approximately 326 and 250 bp, respectively. DGGE profiles of bacterial PCR products (Supplementary Fig. S3a) showed a main fragment corresponding to that of the reference strain *L. sanfranciscensis* DSMZ 20451. DGGE profiles of the yeast community (Supplementary Fig. S3b) showed two main fragments migrating one at the same position of that of *S. cerevisiae* ATCC 32167 and one at the same position of both *C. humilis* DBVPG 7219^{T} and *C. milleri* DBVPG 6753^{T}. The bands from each of the two DGGE gels were excised, sequenced and affiliated to bacterial and yeast species by using BLAST analysis. Results revealed the presence of *L. sanfranciscensis* associated with *C. milleri/C. humilis* and *S. cerevisiae* (Supplementary Fig. S3 and Table 2).

3.4. Qualitative functional characterization of LAB and yeast isolates

Phytase, amylase and protease proprieties of our isolates were assessed by plate assays. The three strains identified as *S. cerevisiae* were able to solubilize phytate, hydrolyze starch and digest casein. Among *C. milleri* strains, 50% showed protease activity, while only 6% and 2% were able to solubilize phytate and to hydrolyze starch, respectively. Specifically, *C. milleri* IMA 33Y showed all the three activities tested (Table 3). Concerning lactic acid bacteria, the ability to solubilize phytate was detected in 19% of *L. sanfranciscensis* strains, while none was able to digest casein and to hydrolyze starch (Table 3).

4. Discussion

Here, a peculiar tripartite species association of *Candida milleri*, *Saccharomyces cerevisiae* and *Lactobacillus sanfranciscensis* was detected as the microbiota characterizing the sourdough of PDO Tuscan bread, by using a multimodal approach.

Microbiological analysis showed the occurrence of LAB and yeasts at a level of about 10^9 and 10^7 CFU/g respectively, revealing a yeasts/LAB ratio of 1:100, consistently with previous data reported for sourdoughs (Gobbetti, 1998). LAB and yeast counts were not affected by the medium used. Among all the isolates, 96 LAB and 68 yeasts were selected and preliminarly functionally characterized for phytase, amylase and protease activity. In particular, the ability to solubilize phytate was found in 18 out of the 96 *L. sanfranciscensis* isolates, while among yeasts this trait was detected in all the *S. cerevisiae* isolates and in 3 out of the 65 *C. milleri* isolates. Such results are consistent with a study of Nuobariene et al. (2012), where phytase activity was found predominant among *S. cerevisiae* isolates and for the first time detected in one *C. humilis* isolate. Moreover, the three *S. cerevisiae* isolates were found positive to the amylase activity assay, a trait which has been scarcely investigated in *S. cerevisiae* sourdough isolates. Consistently with our results, Osimani et al. (2009), analysing 36 yeast strains isolated from sourdoughs collected in the Marche region, found 21 *S. cerevisiae* isolates able to hydrolyze starch, with different levels of activity. However, further investigations are needed in order to exploit the most important functional properties of yeast sourdough isolates for the production of baked goods.

The molecular identification of the selected 96 LAB was carried out by ARDRA and sequence analysis. *Hinf*I and *Alu*I restriction enzymes revealed a polymorphism within the 16S rRNA gene, allowing us to separate the isolates into three different groups, assigned by sequence analysis to the same species, *L. sanfranciscensis*. Such an intraspecific diversity is consistent with data reported by Foschino et al. (2001) showing the same 16S rRNA gene polymorphism in *L. sanfranciscensis* strains isolated from some Italian sourdough samples, when *Hinf*I was used. Interestingly, in the sourdough of PDO Tuscan bread the only lactic acid bacterial species identified was *L. sanfranciscensis*, dissimilarly from previous findings on bacterial communities characterizing other Italian sourdoughs, that detected *L. sanfranciscensis* as the predominant LAB species, though associated with several facultatively heterofermentative species, such as *Lactobacillus plantarum* and *Lactobacillus alimentarius* (Minervini et al., 2012; Yazar and

Tavman, 2012). Moreover, also in sourdoughs used for other typical Tuscan breads, *i.e.* Bozza Pratese and Pane di Altopascio Tradizionale, *L. sanfranciscensis* was found associated with *Lactobacillus paralimentarius* and *Lactobacillus gallinarum*, respectively (Minervini et al., 2012).

The molecular identification of the selected 68 yeasts showed the occurrence of *S. cerevisiae* and *C. milleri* as dominant yeast species in PDO Tuscan bread sourdough, in agreement with previous works reporting that these species are the most frequently found in spontaneously developed stable sourdoughs (De Vuyst et al., 2016). *C. milleri* was the prevalent species in our samples, representing 96% of the isolates identified, whereas *S. cerevisiae* represented the remaining 4%. Such data supplement those obtained from Pagnotta del Dittaino PDO sourdough, where a yeast other than *S. cerevisiae*, *C. humilis*, was the only dominant species recovered (Gullo et al., 2003), which occurred in association with the LAB species *L. sanfranciscensis* and *Enterococcus durans* (Minervini et al., 2012). Similar results were reported also for Pane di Cappelli sourdough, where *C. humilis* and *L. sanfranciscensis* were associated with *L. plantarum* (Minervini et al., 2012). By contrast, other studies on sourdoughs used to produce typical Italian baked goods, detected *S. cerevisiae* as the prevalent species. In particular, Corsetti et al. (2001), analysing 25 different sourdoughs from Apulia region, reported the widespread presence of *S. cerevisiae*, which was also the only yeast species isolated from the Altamura bread sourdoughs (Ricciardi et al., 2005). Accordingly, Valmorri et al. (2010), in 20 sourdoughs collected from artisan bakeries throughout Abruzzo region, identified 85% of the isolates as *S. cerevisiae*, associated with *C. milleri* (11%), *Candida krusei* (2.5%), and *Torulaspora delbrueckii* (1%). Moreover, *S. cerevisiae* was the only dominant species in three out of four sourdoughs in Marche region (Osimani et al., 2009) and in Northern Italy (Iacumin et al., 2009).

It is important to note that the same species association of *C. milleri*, *S. cerevisiae* and *L. sanfranciscensis* characterizing the sourdough used to produce PDO Tuscan bread, was previously described only in traditional Italian sweet baked goods sourdoughs, such as Colomba, Legaccio, Panettone and Veneziana (Lattanzi et al., 2013; Venturi et al., 2012; Vernocchi et al., 2004).

Although in the sourdough context *C. humilis* is often distinguished from *C. milleri* by ITS-RFLP analysis using the *Hae*III enzyme (Pulvirenti et al., 2001), in our work such analysis was not able to discriminate between *C. milleri* and *C. humilis* since the reference strain *C. humilis* DBVPG 6754 showed only one *Hae*III restriction site, corresponding to that expected for *C. milleri*. Similar results were obtained by Vigentini et al. (2014) who found some isolates with only one *Hae*III restriction site, as shown by *C. milleri,* which were positioned closer to *C. humilis* after ITS sequencing. However, ITS region sequence analysis of our 65 *Candida* isolates showed that all our isolates belonged to the species *C. milleri.* The reference strain *C. humilis* DBVPG 6754 grouped in a homogeneous sub-cluster of *C. humilis*, which encompassed other isolates possessing sequences with intermediate traits between the two species. Our data confirm that the ITS sequence analysis is the most reliable method for a correct identification of these two sibling species, as suggested by Vigentini et al. (2014). Interestingly, a phylogenetic analysis of the ITS sequences of *C. milleri* and *C. humilis* collected from GeneBank allowed us to note that all strains affiliated to *C. milleri* were recovered from sourdoughs, whereas those identified as *C. humilis* were isolated also from different fermented foods, such as cacao, bantu beer, tequila etc. (Supplementary Fig. S4). It is tempting to speculate that the species *C. milleri* may represent a key species characteristic of the sourdough environment. Finally the identification of our isolates as *C. milleri* was further confirmed by 26S D1/D2 sequencing. As reported in the literature, *C. humilis* and *C. milleri* differ only by one substitution in the D1/D2 domains of 26SrDNA. Accordingly, our sequences show a C base as *C. milleri* CBS 6897^T (U94923.1) instead of a T base found in *C. humilis* CBS 5658^T (U69878.1) (Kurtzman and Robnett, 2003; Lachance et al., 2011). Recently Jacques et al. (2016), using a multigenic analysis based on the comparison of coding genes D1/D2 LSU rRNA, RPB1, RPB2 and EF-1α, have reconsidered *C. humilis* and *C. milleri* as conspecific and reassigned them to the genus *Kazachastania* proposing the new combination *Kazachastania humilis* (E.E. Nel & Van der Walt) Jacques, Sarilar &Casaregola comb.nov. On the other hand, as the level of intraspecific diversity among the tested *C. milleri* strains was higher than the divergence between *C. humilis* CBS 5658^T

and *C. milleri* CBS 6897^T (Vigentini et al., 2014), the same authors suggested that such isolates may be part of a species complex.

The microbial community composition of PDO Tuscan bread sourdough was also investigated by PCR-DGGE, a method widely utilized to investigate the microbial diversity of many different fermented foods (Cocolin et al., 2013). The same species composition obtained by microbiological analysis was detected in our sourdough samples, where *L. sanfranciscensis, C. milleri* and *S. cerevisiae* were the dominant species. However, by PCR-DGGE we could not discriminate between *C. milleri* and *C. humilis* as the only one base, differing between these two species, is located out of the D1/D2 fragment amplified by the primers used.

The consistency of PCR-DGGE data with those obtained by culture-dependent methods, was reported also by other authors, who studied the microbiota characterizing the sourdoughs used to produce traditional baked goods from different Italian regions *i.e.* Molise (Gatto and Torriani, 2004), Sicily (Randazzo et al., 2005), Abruzzo (Settanni et al., 2006), Northern Italy (Iacumin et al., 2009), Campania (Palomba et al., 2011), confirming that PCR-DGGE is a rapid, economic and efficient tool to investigate yeast and LAB species diversity in the sourdough ecosystem. Moreover, this technique can be conveniently applied to investigate the stability of the microbial communities of specific sourdoughs, particularly those used to produce baked goods protected by the PGI or PDO marks (Palla et al., 2015).

In conclusion, here, for the first time, a distinctive tripartite microbial association, represented by yeast and LAB species characterizing the sourdough used to produce PDO Tuscan bread was detected. The association encompassed a large number of *L. sanfranciscensis* and *C. milleri* strains, along with a few of *S. cerevisiae*, whose relative composition and specific physiological characteristics could potentially affect the organoleptic, rheological, nutritional and nutraceutical features of Tuscan bread, as suggested by the qualitative functional characterization of the isolates. Further investigations on the differential functional traits of the LAB and yeast strains isolated from PDO Tuscan bread sourdough are the next essential steps, in order to exploit the

biotechnological potential of the most effective single strains after assessing their complementary and/or synergistic activities and to select the best performing strain combinations to be used as starters for the production of functional baked goods.

Acknowledgments

This work was supported by the Tuscany Region PIF Project INNOVAPANE "Processi innovativi per la produzione del Pane Toscano a lievitazione naturale (DOP)" (D.D. n. 2260 - 12.06.2013). The authors wish to thank Dr. Benedetta Turchetti, curator of the DBVPG Industrial Yeasts Collection (University of Perugia, Italy), for sharing information on DBVPG strains.

References

Agnolucci, M., Cristani, C., Battini, F., Palla, M., Cardelli, R., Saviozzi, A., Nuti, M., 2013. Microbially-enhanced composting of olive mill solid waste (wet husk): bacterial and fungal community dynamics at industrial pilot and farm level. Bioresour. Technol. 134, 10-16.

Arendt, E.K., Ryan, L.A., Dal Bello, F., 2007. Impact of sourdough on the texture of bread. Food Microbiol. 24, 165-174.

Bae, H.D., Yanke, L.J., Cheng, K.J., Selinger, L.B., 1999. A novel staining method for detecting phytase activity. J. Microbiol. Methods 39, 17-22.

Catzeddu, P., Mura, E., Parente, E., Sanna, M., Farris, G.A., 2006. Molecular characterization of lactic acid bacteria from sourdough breads produced in Sardinia (Italy) and multivariate statistical analyses of results. Syst. Appl. Microbiol. 29, 138-144.

Cocolin, L., Bisson, L.F., Mills, D.A., 2000. Direct profiling of the yeast dynamics in wine fermentations. FEMS Microbiol. Lett. 189, 81-87.

Cocolin, L., Alessandria, V., Dolci, P., Gorra, R., Rantsiou, K., 2013. Culture independent methods to assess the diversity and dynamics of microbiota during food fermentations. Int. J Food Microbiol. 167, 29–43.

Corsetti, A., 2013. Technology of sourdough fermentation and sourdough applications. In: Gobbetti, M., Gänzle, M. (Eds.), Handbook on Sourdough Biotechnology. Springer, New York, pp. 85-104.

Corsetti, A., Lavermicocca, P., Morea, M., Baruzzi, F., Tosti, N., Gobbetti, M., 2001. Phenotypic and molecular identification and clustering of lactic acid bacteria and yeasts from wheat (species *Triticum durum* and *Triticum aestivum*) sourdoughs of Southern Italy. Int. J. Food Microbiol. 64, 95-104.

De Man, J.C., Rogosa, D., Sharpe, M.E., 1960. A medium for the cultivation of lactobacilli. J. Appl. Microbiol. 23, 130-135.

Dertli, E., Mercan, E., Arıcı, M., Yılmaz, M.T., Sağdıç, O., 2016. Characterisation of lactic acid bacteria from Turkish sourdough and determination of their exopolysaccharide (EPS) production characteristics. LWT-Food Sci. Technol. 71, 116-124.

De Vuyst, L., Harth, H., Van Kerrebroeck, S., Leroy, F., 2016. Yeast diversity of sourdoughs and associated metabolic properties and functionalities. Int. J. Food Microbiol. 239, 26-34.

De Vuyst, L., Neysens, P., 2005. The sourdough microflora: biodiversity and metabolic interactions. Trends Food Sci. Technol. 16, 43-56.

De Vuyst, L., Van Kerrebroeck, S., Harth, H., Huys, G., Daniel, H.M., Weckx, S., 2014. Microbial ecology of sourdough fermentations: diverse or uniform? Food Microbiol. 37, 11-29.

Edgar, R.C., 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics 5, 1.

Edgar, R.C., 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32, 1792-1797.

Felsenstein J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39, 783-791.

Ferchichi, M., Valcheva, R., Oheix, N., Kabadjova, P., Prévost, H., Onno, B., Dousset, X., 2008. Rapid investigation of French sourdough microbiota by restriction fragment length polymorphism

of the 16S-23S rRNA gene intergenic spacer region. World J. Microbiol. Biotechnol. 24, 2425-2434.

Fernández-Espinar, M.T., Esteve-Zarzoso, B., Querol, A., Barrio, E., 2000. RFLP analysis of the ribosomal internal transcribed spacers and the 5.8S rRNA gene region of the genus *Saccharomyces*: a fast method for species identification and the differentiation of flor yeasts. Antonie van Leeuwenhoek, 78, 87-97.

Foschino, R., Arrigoni, C., Mora, D., Galli, A., 2001. Phenotypic and genotypic aspects of *Lactobacillus sanfranciscensis* strains isolated from sourdoughs in Italy. Food Microbiol. 18, 277-285.

Gänzle, M., Ripari, V., 2016. Composition and function of sourdough microbiota: From ecological theory to bread quality. Int. J. Food Microbiol. 239, 19-25.

Gatto, V., Torriani, S., 2004. Microbial population changes during sourdough fermentation monitored by DGGE analysis of 16S and 26S rRNA gene fragments. Ann. Microbiol. 54, 31-42.

Gobbetti, M., 1998. The sourdough microflora: interactions of lactic acid bacteria and yeasts. Trends Food Sci. Technol. 9, 267-274.

Gobbetti, M., Gänzle, M., 2013. Handbook on sourdough biotechnology, first ed. Springer, New York.

Gobbetti, M., Rizzello, C.G., Di Cagno, R., De Angelis, M., 2014. How the sourdough may affect the functional features of leavened baked goods. Food Microbiol. 37, 30-40.

Gullo, M., Romano, A.D., Pulvirenti, A., Giudici, P., 2003. *Candida humilis* - dominant species in sourdoughs for the production of durum wheat bran flour bread. Int. J. Food Microbiol. 80, 55-59.

Iacumin, L., Cecchini, F., Manzano, M., Osualdini, M., Boscolo, D., Orlic, S., Comi, G., 2009. Description of the microflora of sourdoughs by culture-dependent and culture-independent methods. Food Microbiol. 26, 128-135.

493 Istituto Nazionale di Sociologia Rurale, 2000. Atlante dei prodotti tipici: il pane (p. 13). Agra-
494 Rai Eri, Rome, Italy.
495 Jacques, N., Sarilar, V., Urien, C., Lopes, M.R., Morais, C.G., Uetanabaro, A.P.T., Tinsley C.R.,
496 Rosa C.A., Sicard D., Casaregola, S., 2016. Three novel ascomycetous yeast species of the
497 *Kazachstania* clade, *Kazachstania saulgeensis* sp. nov., *Kazachstania serrabonitensis* sp. nov. and
498 *Kazachstania australis* sp. nov. Reassignement of *Candida humilis* to *Kazachastania humilis* fa
499 comb nov and *Candida pseudohumilis* to *Kazachstania pseudohumilis* fa comb. nov. *International*
500 Int. J. Syst. Evol. Microbiol. 66, 5192-5200.
501 Jorquera, M.A., Hernández, M.T., Rengel, Z., Marschner, P., de la Luz Mora, M., 2008. Isolation
502 of culturable phosphobacteria with both phytate-mineralization and phosphate-solubilization
503 activity from the rhizosphere of plants grown in a volcanic soil. Biol. Fertil. Soils 44, 1025-1034.
504 Katina, K., Arendt, E., Liukkonen, K.H., Autio, K., Flander, L., Poutanen, K., 2005. Potential of
505 sourdough for healthier cereal products. Trends Food Sci. Technol. 16, 104-112.
506 Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions
507 through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 111-120.
508 Kline, L., Sugihara, T.F., 1971. Microorganisms of the San Francisco sour dough bread process
509 II. Isolation and characterization of undescribed bacterial species responsible for the souring
510 activity. Applied Microbiol. 21, 459-465.
511 Kurtzman, C.P., Robnett, C.J., 1998. Identification and phylogeny of ascomycetous yeasts from
512 analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. Antonie Van
513 Leeuwenhoek 73, 331-371.
514 Kurtzman, C.P., Robnett, C.J., 2003. Phylogenetic relationships among yeasts of the
515 'Saccharomyces complex' determined from multigene sequence analyses. FEMS Yeast Res. 3, 417-
516 432.

517 Lachance, M.A., Boekhout, T., Scorzetti, G., Fell, J.W., Kurtzman, C.P., 2011. Candida
518 Berkhout. In: Kurtzman, C., Fell, J.W., Boekhout, T. (Eds.), The Yeasts: a Taxonomic Study, fifth
519 ed., vol. 2. Elsevier, Amsterdam (NL), pp. 987-1278.
520 Lane, DJ., 1991. 16S/23S rRNA sequencing. In: Stackebrandt, E., Goodfellow, M. (Eds),
521 Nucleic acid techniques in bacterial systematics. Wiley, New York, pp. 115-175.
522 Lattanzi, A., Minervini, F., Di Cagno, R., Diviccaro, A., Antonielli, L., Cardinali, G., Cappelle,
523 S., De Angelis, M., Gobbetti, M., 2013. The lactic acid bacteria and yeast microbiota of eighteen
524 sourdoughs used for the manufacture of traditional Italian sweet leavened baked goods. Int. J. Food
525 Microbiol. 163, 71-79.
526 Liu, T., Li, Y., Chen, J., Sadiq, F.A., Zhang, G., Li, Y., He, G., 2016. Prevalence and diversity of
527 lactic acid bacteria in Chinese traditional sourdough revealed by culture dependent and
528 pyrosequencing approaches. LWT-Food Sci. Technol. 68, 91-97.
529 Minervini, F., De Angelis, M., Di Cagno, R., Gobbetti, M., 2014. Ecological parameters
530 influencing microbial diversity and stability of traditional sourdough. Int. J. Food Microbiol. 171,
531 136-146.
532 Minervini, F., Di Cagno, R., Lattanzi, A., De Angelis, M., Antonielli, L., Cardinali, G., Cappelle,
533 S., Gobbetti, M., 2012. Lactic acid bacterium and yeast microbiotas of 19 sourdoughs used for
534 traditional/typical Italian breads: interactions between ingredients and microbial species diversity.
535 Appl. Environ. Microbiol. 78, 1251-1264.
536 Nuobariene, L., Hansen, Å. S., Arneborg, N., 2012. Isolation and identification of phytase-active
537 yeasts from sourdoughs. LWT-Food Sci. Technol. 48, 190-196.
538 Osimani, A., Zannini, E., Aquilanti, L., Mannazzu, I.M., Comitini, F., Clementi, F., 2009. Lactic
539 acid bacteria and yeasts from wheat sourdoughs of the Marche region. Ital. J Food Sci. 21, 269-286.
540 Palla, M., Cristani, C., Giovannetti, M., Agnolucci, M., 2015. Identificazione molecolare di
541 batteri lattici e lieviti caratterizzanti l'impasto acido per la produzione di Pane Toscano a
542 lievitazione naturale. Industrie Alimentari 54, 5-11.

543 Palomba, S., Blaiotta, G., Ventorino, V., Saccone, A., Pepe, O., 2011. Microbial characterization
544 of sourdough for sweet baked products in the Campania region (southern Italy) by a polyphasic
545 approach. Ann. Microbiol. 61, 307-314.
546 Pulvirenti, A., Caggia, C., Restuccia, C., Gullo, M., Giudici, P., 2001. DNA fingerprinting
547 methods used for identification of yeasts isolated from Sicilian sourdoughs. Ann. Microbiol. 51,
548 107-120.
549 Randazzo, C.L., Heilig, H., Restuccia, C., Giudici, P., Caggia, C., 2005. Bacterial population in
550 traditional sourdough evaluated by molecular methods. J. Appl. Microbiol. 99, 251-258.
551 Reale, A., Tremonte, P., Succi, M., Sorrentino, E., Coppola, R., 2005. Exploration of lactic acid
552 bacteria ecosystem of sourdoughs from the Molise region. Ann. Microbiol. 55, 17-22.
553 Ricciardi, A., Parente, E., Piraino, P., Paraggio, M., Romano, P., 2005. Phenotypic
554 characterization of lactic acid bacteria from sourdoughs for Altamura bread produced in Apulia
555 (Southern Italy). Int. J. Food Microbiol. 98, 63-72.
556 Saitou, N., Nei, M., 1987. The neighbor-joining method: A new method for reconstructing
557 phylogenetic trees. Mol. Biol. Evol. 4, 406-425.
558 Scheirlinck, I., Van der Meulen, R., Van Schoor, A., Vancanneyt, M., De Vuyst, L., Vandamme,
559 P., Huys, G., 2007. Influence of geographical origin and flour type on diversity of lactic acid
560 bacteria in traditional Belgian sourdoughs. Appl. Environ. Microbiol.73, 6262-6269.
561 Settanni, L., Valmorri, S., van Sinderen, D., Suzzi, G., Paparella, A., Corsetti, A., 2006.
562 Combination of multiplex PCR and PCR-denaturing gradient gel electrophoresis for monitoring
563 common sourdough-associated *Lactobacillus* species. Appl. Environ. Microbiol. 72, 3793-3796.
564 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular
565 Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.
566 Valmorri, S., Settanni, L., Suzzi, G., Gardini, F., Vernocchi, P., Corsetti, A., 2006. Application
567 of a novel polyphasic approach to study the lactobacilli composition of sourdoughs from the
568 Abruzzo region (central Italy). Lett. Appl. Microbiol. 43, 343-349.

569 Valmorri, S., Tofalo, R., Settanni, L., Corsetti, A., Suzzi, G., 2010. Yeast microbiota associated
570 with spontaneous sourdough fermentations in the production of traditional wheat sourdough breads
571 of the Abruzzo region (Italy). Antonie Van Leeuwenhoek 97, 119-129.
572 Venturi, M., Guerrini, S., Vincenzini, M., 2012. Stable and non-competitive association of
573 *Saccharomyces cerevisiae*, *Candida milleri* and *Lactobacillus sanfranciscensis* during manufacture
574 of two traditional sourdough baked goods. Food Microbiol. 31, 107-115.
575 Vera, A., Ly-Chatain, M.H., Rigobello, V., Demarigny, Y., 2012. Description of a French natural
576 wheat sourdough over 10 consecutive days focussing on the lactobacilli present in the microbiota.
577 Antonie van Leeuwenhoek 101, 369-377.
578 Vernocchi, P., Valmorri, S., Gatto, V., Torriani, S., Gianotti, A., Suzzi, G., Guerzoni, M.E.,
579 Gardini, F., 2004. A survey on yeast microbiota associated with an Italian traditional sweet-
580 leavened baked good fermentation. Food Res. Int. 37, 469-476.
581 Vigentini, I., Antoniani, D., Roscini, L., Comasio, A., Galafassi, S., Picozzi, C., Corte, L.,
582 Compagno, C., Dal Bello, F., Cardinali, G., Foschino, R., 2014. *Candida milleri* species reveals
583 intraspecific genetic and metabolic polymorphisms. Food Microbiol. 42, 72-81.
584 Walter, J., Hertel, C., Tannock, G.W., Lis, C. M., Munro, K., Hammes, W.P., 2001. Detection of
585 *Lactobacillus, Pediococcus, Leuconostoc*, and *Weissella* species in human feces by using group-
586 specific PCR primers and denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 67,
587 2578-2585.
588 Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J., 1991. 16S ribosomal DNA
589 amplification for phylogenetic study. J. Bacteriol. 173, 697-703.
590 White, T.J., Bruns, T., Lee, S., Taylor, J.W., 1990. Amplification and direct sequencing of fungal
591 ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White T.J.
592 (Eds.), PCR protocols: a guide to methods and applications. Academic Press Inc., New York, pp.
593 315-322.

Yazar, G., Tavman, Ş., 2012. Functional and technological aspects of sourdough fermentation with *Lactobacillus sanfranciscensis*. Food Eng. Rev. 4, 171-190.

Zhang, J., Liu, W., Sun, Z., Bao, Q., Wang, F., Yu, J., Chen, W., Zhang, H., 2011. Diversity of lactic acid bacteria and yeasts in traditional sourdoughs collected from western region in Inner Mongolia of China. Food Control 22, 767-774.

Zotta, T., Piraino, P., Parente, E., Salzano, G., Ricciardi, A., 2008. Characterization of lactic acid bacteria isolated from sourdoughs for *Cornetto*, a traditional bread produced in Basilicata (Southern Italy). World J. Microbiol. Biotechnol. 24, 1785-1795.

620 **Table 1**

621 Lactic acid bacteria and yeast reference strains used in this study.

Strains[a]	*Source of isolation*
Lactobacillus panis DSMZ 6035^T	Sourdough
Lactobacillus sanfranciscensis DSMZ 20451^T	San Francisco sourdough
Lactobacillus fermentum DSMZ 20052^T	Fermented beets
Lactobacillus brevis DSMZ 20054^T	Faeces
Lactobacillus plantarum IMA B23	Boza
Lactobacillus curvatus IMA LB51	Sourdough
Saccharomyces cerevisiae ATCC 32167	Unknow
Dekkera bruxellensis IMA 1L	San Giovese Tuscan wine
Candida milleri DBVPG 6753^T	San Francisco sourdough
Candida humilis DBVPG 7219^T	Bantu beer
Candida humilis DBVPG 6754	Sourdough, Finland
Kazachstania exigua DBVPG 6956	Wheat sourdough, Italy

622

623 ^{T}Type Strain.

624 [a]DSMZ=Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany;
625 IMA=International Microbial Archives, Department of Agriculture, Food and Environment,
626 University of Pisa, Pisa, Italy; ATCC=American Type culture Collection, Manassas, Virginia,
627 USA; DBVPG=International Collection of Department of Agricultural, Food and Environmental
628 Science, University of Perugia, Perugia, Italy.

629

630 **Table 2**

631 Lactic acid bacteria and yeasts characterizing the PDO Tuscan bread sourdough: identification by

632 sequencing of different representative isolate genes and of DNA fragments in DGGE profiles.

Isolates	*Genes*	*Accession number*	*Species in NCBI database*	*Sequence identity*
IMA 2LAB, 3, 13, 19, 23, 29, 44, 64, 67, 93, 97	16S	From LT605146 to LT605156	*L. sanfranciscensis* DSMZ 20663 (X76331) *L. sanfranciscensis* ATCC 27651^{T} (X76327)	99%
IMA 105Y	ITS1-5.8S-ITS2	LT605145	*S. cerevisiae* ATCC 834 (KU729072)	98%
IMA 19Y, 36, 105	26S D1/D2	From LT718652 to LT718654	*S. cerevisiae* CBS 2962 (KY109317)	100%
IMA 11Y, 33	26S D1/D2	From LT718655 to LT718656	*C. milleri* CBS 6897 (KY106585)	100%
DGGE fragments				
5LAB, 14LAB, 29LAB	16S V3-V5	From LT605157 to LT605159	*L. sanfranciscensis* ATCC 27651^{T} (X76327) *L. sanfranciscensis* DSMZ 20663 (X76331)	99%
1Y, 9Y, 19Y	partial 26S D1/D2	From LT605160 to LT605162	*C. milleri* NRRL Y-7245 ^{T} (U94923) *C. humilis* NRRL Y-17074 ^{T} (U69878)	99% 99%
3Y, 15Y, 20Y	partial 26S D1/D2	From LT605163 to LT605165	*S. cerevisiae* CTBRL121 (JX423567)	100%

633
634

635 **Table 3**

636 Phytase, protease and amylase activities of lactic acid bacteria and yeasts isolated from the PDO
637 Tuscan bread sourdough.

Isolates	*Phytase activity*	*Protease activity*	*Amylase activity*
Lactic acid bacteria			
IMA 1LAB; 3; 5-13; 17-21; 23-26; 28-38; 40-49; 51; 52; 56-58; 60-62; 64; 66-75; 77-82; 84-86; 88-92; 94-96; 99	-	-	-
IMA 2LAB; 22; 39; 53; 63; 76; 87; 93	+	-	-
IMA 14-16LAB; 27; 55; 59; 65; 97	-/+	-	-
IMA 83LAB; 98	++	-	-
Yeasts			
IMA 1Y; 122	+	-/+	-
IMA 2-3Y; 9-10; 12; 17-18; 20; 34; 40; 48-49; 103-104; 106-108	-	-/+	-
IMA 4-8Y; 13-16; 21-26; 28; 30-31; 35; 38; 41; 43; 44-47; 50-51; 109-120	-	-	-
IMA 11Y; 32; 37	-	+	-
IMA 19Y	+	+	+
IMA 27Y; 29	-	++	-
IMA 33Y	+	-/+	++
IMA 36Y	++	-/+	++
IMA 105Y	+	-/+	+

638

639 Lactic acid bacteria. Phytase activity: - = no activity (halo = 0mm), +/- = low activity (halo ≤ 1mm),
640 + = moderate activity (1mm < halo ≤ 4mm), ++ = high activity (halo > 4mm).

641 Yeasts. Phytase activity: - = no activity (halo = 0mm), +/- = low activity (halo ≤ 1mm), + =
642 moderate activity (1mm < halo ≤ 15mm), ++ = high activity (halo > 15mm); protease activity: - =
643 no activity (halo = 0mm), +/- = low activity (halo ≤ 1mm), + = moderate activity (1mm < halo ≤
644 5mm), ++ = high activity (halo > 5mm); amylase activity: - = no activity (halo = 0mm), +/- = low
645 activity (halo ≤ 1mm), + = moderate activity (1mm < halo ≤ 4mm), ++ = high activity (halo >
646 4mm).

647

648

649

650 **FIGURE CAPTIONS**

651 **Fig. 1.** Dendrogram obtained from UPGMA analysis, using Dice's coefficient, of combined
652 ARDRA profiles of 96 LAB isolates and LAB reference strains. The scale indicates the similarity
653 level.

654

655 **Fig. 2.** Dendrogram obtained from UPGMA analysis, using Dice's coefficient, of combined ITS-
656 RFLP profiles and ITS amplicons size of 68 yeast isolates and yeast reference strains. The scale
657 indicates the similarity level.

658

659 **Fig. 3.** Dendrogram showing multiple sequence alignment of ITS1-5.8S-ITS2 region of *Candida*
660 *milleri* strains isolated from the PDO Tuscan bread sourdough. The evolutionary history was
661 inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum
662 of branch length = 0.54023792 is shown. The percentage of replicate trees in which the associated
663 taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches
664 (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the
665 evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were
666 computed using the Kimura 2-parameter method (Kimura, 1980) and are in the units of the number
667 of base substitutions per site. The rate variation among sites was modeled with a gamma
668 distribution (shape parameter = 1). The analysis involved 83 nucleotide sequences. All positions
669 with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing
670 data, and ambiguous bases were allowed at any position. There were a total of 524 positions in the
671 final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013). Numbers in
672 parentheses are accession numbers of published sequences.

673

674

Figure

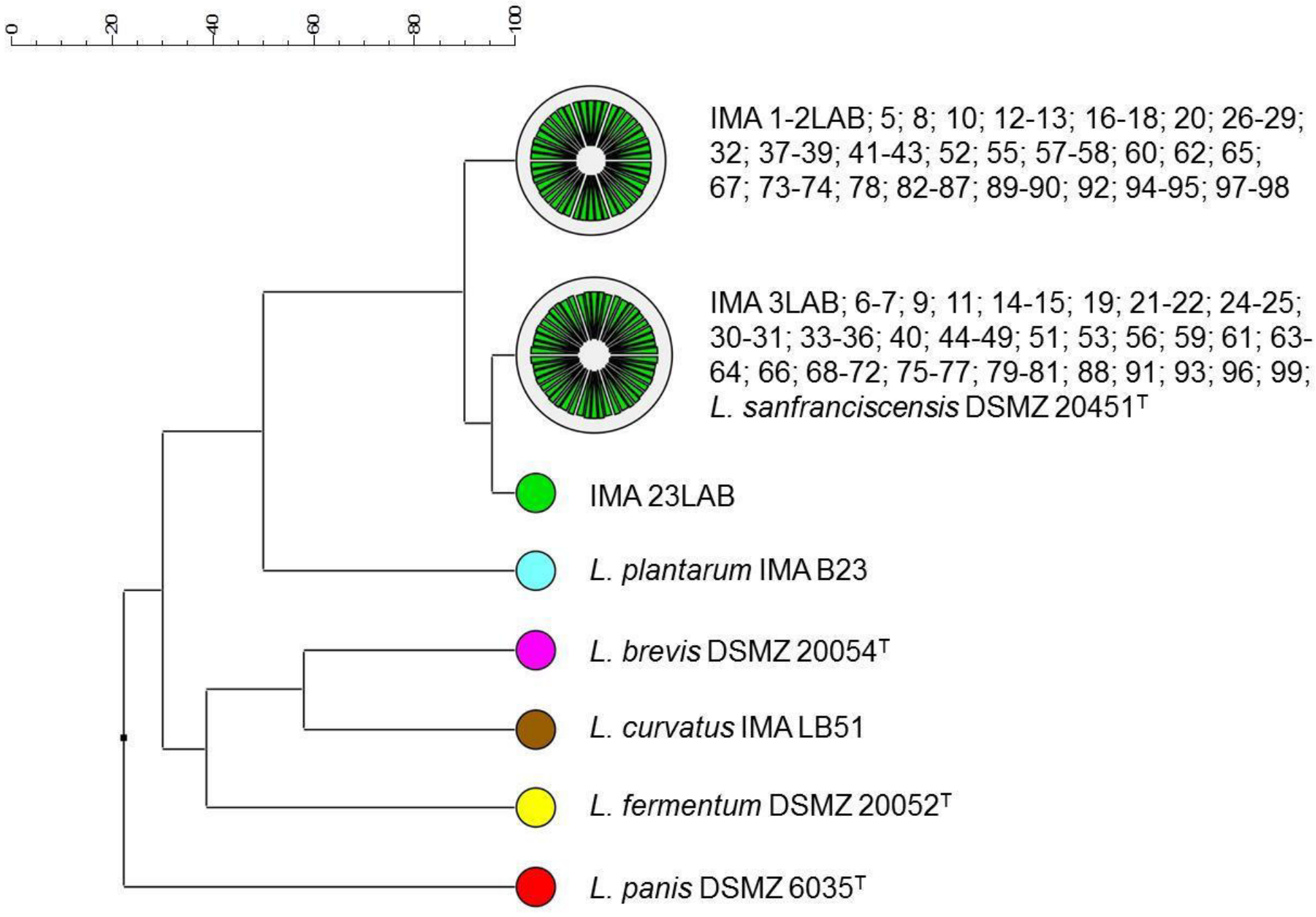

0
20
40
60
80
100
IMA 1-2LAB; 5; 8; 10; 12-13; 16-18; 20; 26-29; 32; 37-39; 41-43; 52; 55; 57-58; 60; 62; 65; 67; 73-74; 78; 82-87; 89-90; 92; 94-95; 97-98
IMA 3LAB; 6-7; 9; 11; 14-15; 19; 21-22; 24-25; 30-31; 33-36; 40; 44-49; 51; 53; 56; 59; 61; 63-64; 66; 68-72; 75-77; 79-81; 88; 91; 93; 96; 99; L. sanfranciscensis DSMZ 20451T
IMA 23LAB
L. plantarum IMA B23
L. brevis DSMZ 20054T
L. curvatus IMA LB51
L. fermentum DSMZ 20052T
L. panis DSMZ 6035T

Fig. 1

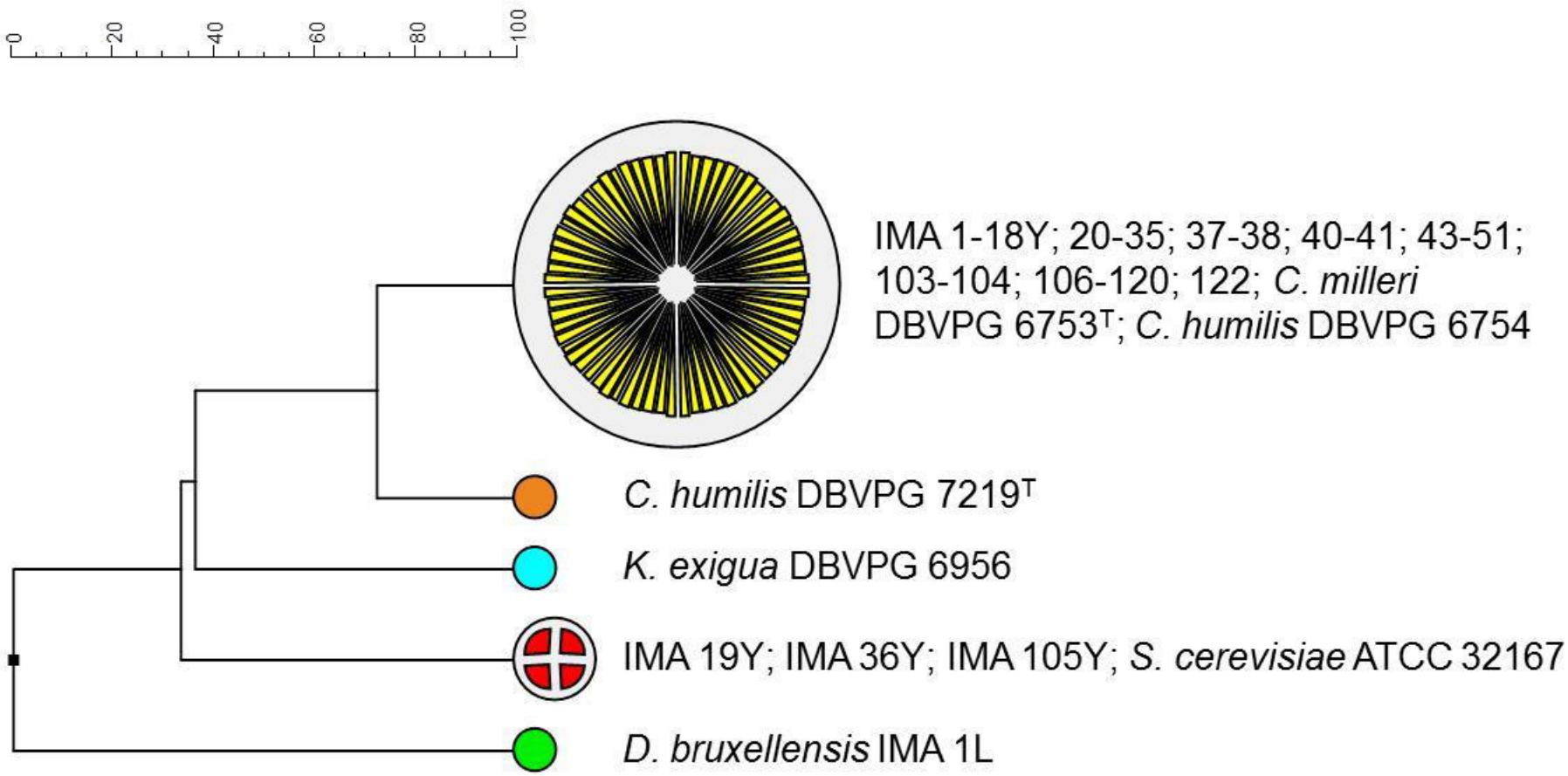

0
20
40
60
80
100
IMA 1-18Y; 20-35; 37-38; 40-41; 43-51; 103-104; 106-120; 122; C. milleri DBVPG 6753T; C. humilis DBVPG 6754
C. humilis DBVPG 7219T
K. exigua DBVPG 6956
IMA 19Y; IMA 36Y; IMA 105Y; S. cerevisiae ATCC 32167
D. bruxellensis IMA 1L

Fig. 2

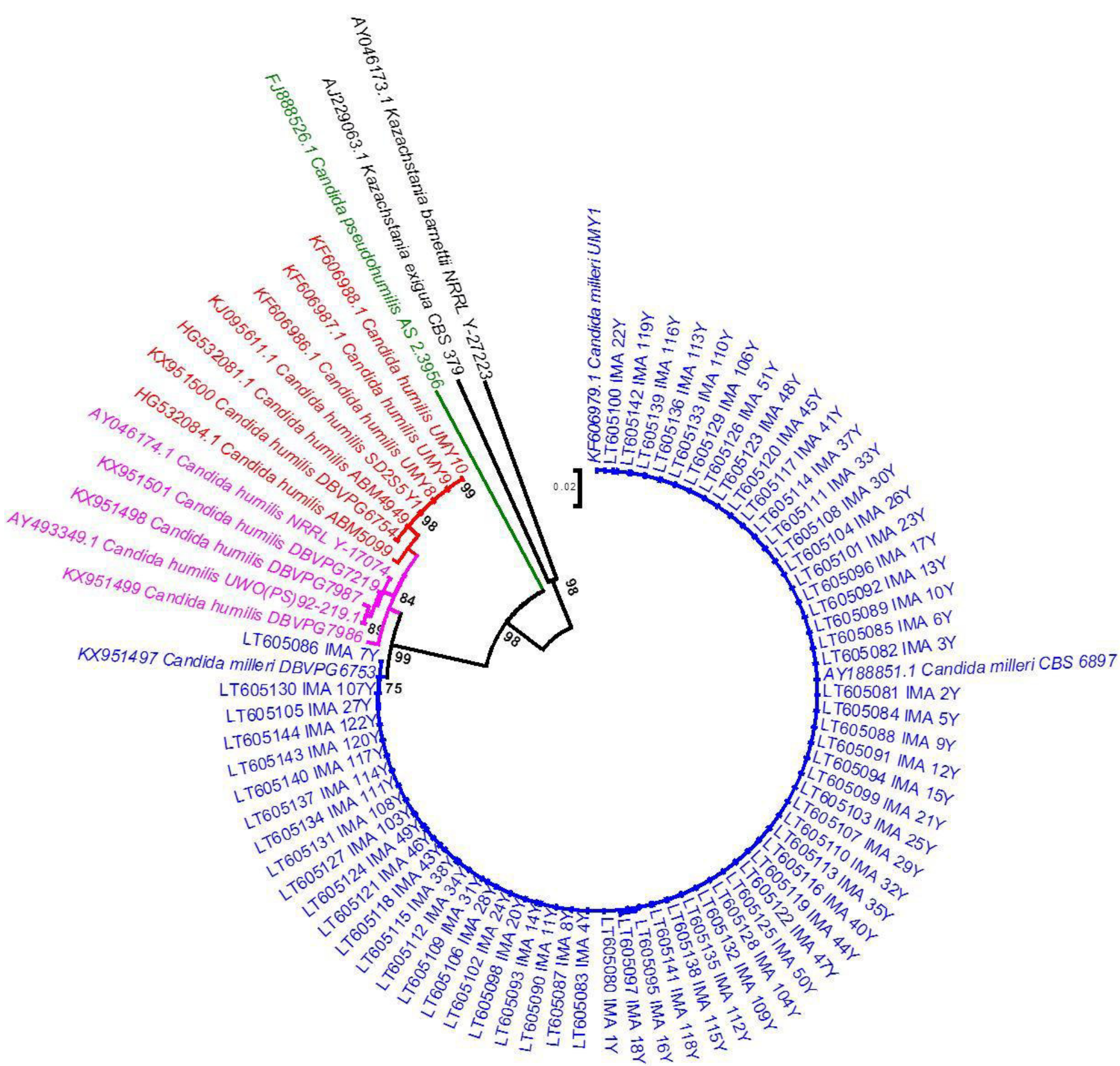
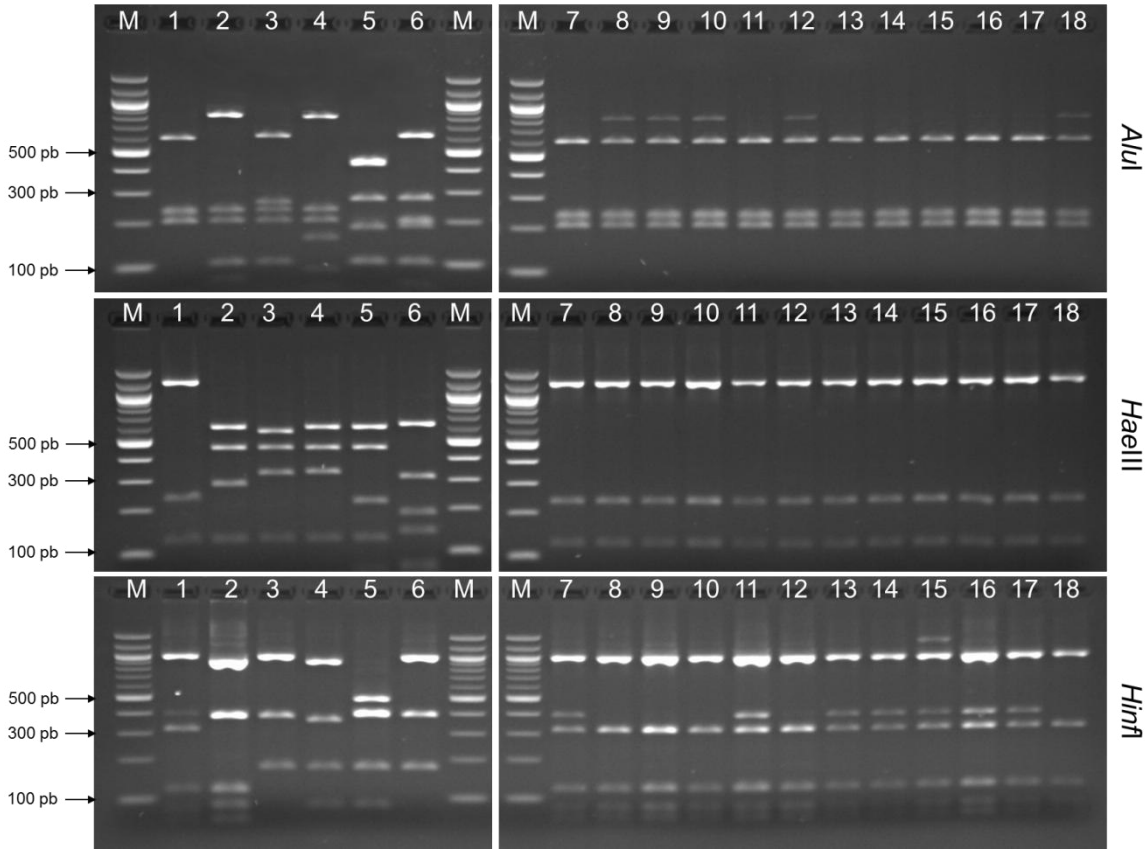

AY046173.1 Kazachstania barnettii NRRL Y-27223
AJ229063.1 Kazachstania exigua CBS 379
FJ888526.1 Candida pseudohumilis AS 2.3956
KF606988.1 Candida humilis UMY10
KF606987.1 Candida humilis UMY9
KF606986.1 Candida humilis SD2S5Y1
KJ095611.1 Candida humilis ABM4949
HG532081.1 Candida humilis DBVPG6754
KX951500 Candida humilis ABM5099
HG532084.1 Candida humilis NRRL Y-17074
AY046174.1 Candida humilis DBVPG7219
KX951501 Candida humilis DBVPG7987
KX951498 Candida humilis UWO(PS)92-219.1
AY493349.1 Candida humilis DBVPG7986
KX951499 Candida humilis DBVPG6753
LT605086 IMA 7Y
LT605130 IMA 107Y
LT605105 IMA 27Y
LT605144 IMA 122Y
LT605143 IMA 120Y
LT605140 IMA 117Y
LT605137 IMA 114Y
LT605134 IMA 111Y
LT605131 IMA 108Y
LT605127 IMA 103Y
LT605124 IMA 49Y
LT605121 IMA 46Y
LT605118 IMA 43Y
LT605115 IMA 38Y
LT605112 IMA 34Y
LT605109 IMA 31Y
LT605106 IMA 28Y
LT605102 IMA 24Y
LT605098 IMA 20Y
LT605093 IMA 14Y
LT605090 IMA 11Y
LT605087 IMA 8Y
LT605083 IMA 4Y
LT605080 IMA 1Y
LT605097 IMA 18Y
LT605095 IMA 16Y
LT605141 IMA 118Y
LT605138 IMA 115Y
LT605135 IMA 112Y
LT605132 IMA 109Y
LT605128 IMA 104Y
LT605125 IMA 50Y
LT605122 IMA 47Y
LT605119 IMA 44Y
LT605116 IMA 40Y
LT605113 IMA 35Y
LT605110 IMA 32Y
LT605107 IMA 29Y
LT605103 IMA 25Y
LT605099 IMA 21Y
LT605094 IMA 15Y
LT605091 IMA 12Y
LT605088 IMA 9Y
LT605084 IMA 5Y
LT605081 IMA 2Y
AY188851.1 Candida milleri CBS 6897
LT605082 IMA 3Y
LT605085 IMA 6Y
LT605089 IMA 10Y
LT605092 IMA 13Y
LT605096 IMA 17Y
LT605101 IMA 23Y
LT605104 IMA 26Y
LT605108 IMA 30Y
LT605111 IMA 33Y
LT605114 IMA 37Y
LT605117 IMA 41Y
LT605120 IMA 45Y
LT605123 IMA 48Y
LT605126 IMA 51Y
LT605129 IMA 106Y
LT605133 IMA 113Y
LT605136 IMA 116Y
LT605139 IMA 119Y
LT605142 IMA 22Y
KF606979.1 Candida milleri UMY1
0.02
99
98
84
89
99
75
98
98

Fig. 3

Identification and characterization of lactic acid bacteria and yeasts of PDO Tuscan bread sourdough by culture dependent and independent methods

Michela Palla, Caterina Cristani, Manuela Giovannetti, Monica Agnolucci*
Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy
*Corresponding author: M. Agnolucci, Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy Phone: +39.0502216647, Fax: +39.0502210606, e-mail address: monica.agnolucci@unipi.it

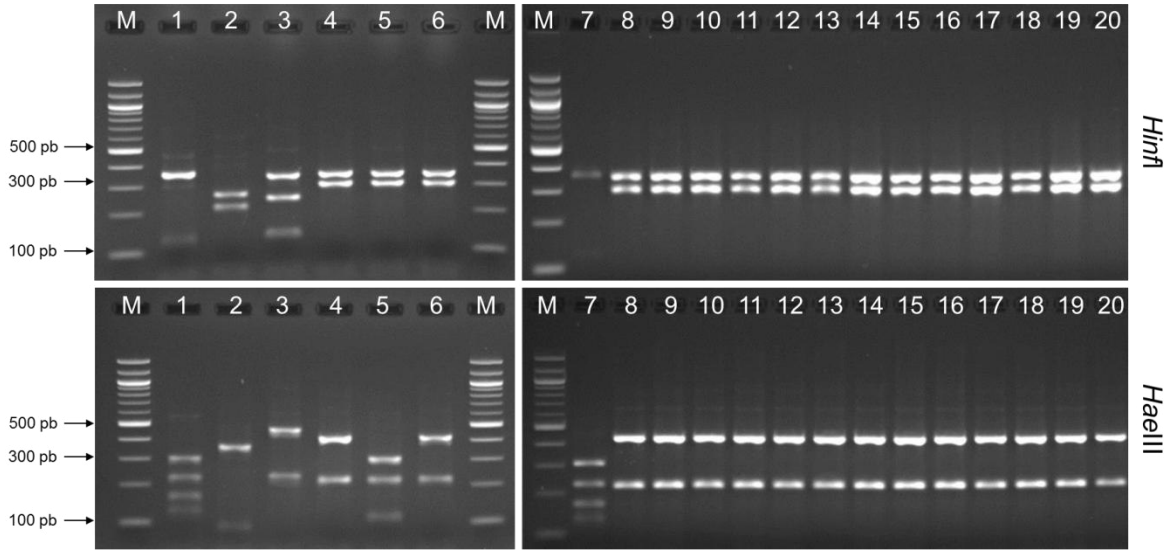
Fig. S1. Electrophoresis of *Alu*I, *Hae*III and *Hinf*I ARDRA patterns of LAB reference strains and isolates. 1: *L. sanfranciscensis* DSMZ 20451^T; 2: *L. brevis* DSMZ 20054^T; 3: *L. plantarum* IMA B23; 4: *L. curvatus* IMA LB51; 5: *L. fermentum* DSMZ 20052^T; 6: *L. panis* DSMZ 6035^T; 7-18: IMA 15LAB-IMA 26LAB; M: Marker 100bp (BioLabs).



Identification and characterization of lactic acid bacteria and yeasts of PDO Tuscan bread sourdough by culture dependent and independent methods

Michela Palla, Caterina Cristani, Manuela Giovannetti, Monica Agnolucci*
Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy
*Corresponding author: M. Agnolucci, Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy Phone: +39.0502216647, Fax: +39.0502210606, e-mail address: monica.agnolucci@unipi.it

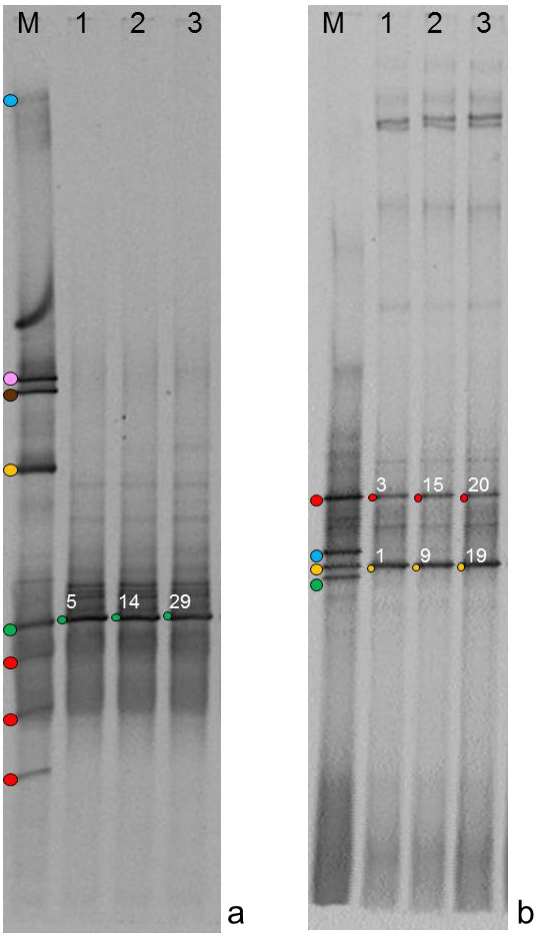
Fig. S2. Electrophoresis of *Hae*III and *Hinf*I RFLP patterns of the ITS regions of yeast reference strains and isolates. 1: *S. cerevisiae* ATCC 32167; 2: *D. bruxellensis* IMA 1L; 3: *K. exigua* DBVPG 6956; 4: *C. milleri* DBVPG 6753^{T}; 5: *C. humilis* DBVPG 7219^{T}; 6: *C. humilis* DBVPG 6754; 7-20: IMA 19Y-IMA 35Y; M: Marker 100bp (BioLabs).



Identification and characterization of lactic acid bacteria and yeasts of PDO Tuscan bread sourdough by culture dependent and independent methods

Michela Palla, Caterina Cristani, Manuela Giovannetti, Monica Agnolucci*
Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy
*Corresponding author: M. Agnolucci, Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy Phone: +39.0502216647, Fax: +39.0502210606, e-mail address: monica.agnolucci@unipi.it

Fig. S3. DGGE profiles of the microbial community of PDO Tuscan bread sourdough. (a) LAB 16S rRNA gene V3-V4 region; 1, 2, 3: sourdough replicates; marker (M): ● *Lb. plantarum* IMA B23, ● *Lb. brevis* DSMZ 20054, ● *Lb. curvatus* IMA LB51, ● *Lb. fermentum* DSMZ 20052, ● *Lb. sanfranciscensis* DSMZ 20451, ● *Lb. panis* DSMZ 6035. (b) Yeast 26S rRNA gene D1/D2 region; 1, 2, 3: sourdough replicates; marker (M): ● *S. cerevisiae* ATCC 32167, ● *K. exigua* DBVPG 6956, ● *C. milleri* DBVPG 6753/*C. humilis* DBVPG 7219, ● *D. bruxellensis* IMA 1L. The numbers indicate sequenced DNA fragments and the colored circles their species affiliation.



Identification and characterization of lactic acid bacteria and yeasts of PDO Tuscan bread sourdough by culture dependent and independent methods

Michela Palla, Caterina Cristani, Manuela Giovannetti, Monica Agnolucci*
Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy
*Corresponding author: M. Agnolucci, Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy Phone: +39.0502216647, Fax: +39.0502210606, e-mail address: monica.agnolucci@unipi.it

Fig. S4. Dendrogram showing multiple sequence alignment of ITS1-5.8S-ITS2 region of *C. milleri* and *C. humilis* collected from GeneBank. The evolutionary history was inferred using the Neighbor-Joining method. Evolutionary analyses were conducted in MEGA6. For each sequence is reported its source of isolation and the relative number of *Hae*III restriction sites.