1 Postprint of Food Microbiology, Vol. 94, April 2021, 103652

2 DOI: <https://doi.org/10.1016/j.fm.2020.103652>

3
4 **Genotypic and phenotypic diversity among *Lactobacillus plantarum* and *Lactobacillus***
5 ***pentosus* isolated from industrial scale cucumber fermentations**
6

7 Running title: lactobacilli biodiversity in pickles

8
9 Ilenys M. Pérez-Díaz^{1,*}, Suzanne D. Johanningsmeier¹, Kartheek Anekella², Christian G. Pagán-
10 Medina¹, Lesley Méndez-Sandoval², Consuelo Arellano³, Robert Price², Katheryne V.
11 Daughtry², Michelle Borges², Chloe Bream², Lauren Connelly², Susan E. Dieck², Meredith T.
12 Levi², Erin K. McMurtrie^{2,4}, Rickey E. Smith, Jr.², Jeannette C. Theora^{2,4}, Paige Wendland²,
13 Francisco Gómez-Rodríguez⁵, Francisco Noé Arroyo-López⁵
14

15 ¹USDA-Agricultural Research Service, Food Science Research Unit, Raleigh, North Carolina,
16 USA; ²North Carolina State University, Department of Food, Bioprocessing, and Nutrition
17 Sciences, Raleigh, North Carolina, USA; ³North Carolina State University, Department of
18 Statistics, Raleigh, North Carolina, USA; ⁴North Carolina State University, Department of
19 Chemistry, Raleigh, North Carolina, USA; ⁵Department of Food Biotechnology, Instituto de la
20 Grasa, (CSIC), Campo Universitario Pablo de Olavide, Seville, Spain
21

22 Mention of a trademark or proprietary product does not constitute a guarantee or warranty of the
23 product by the U. S. Department of Agriculture or North Carolina Agricultural Research Service,

24 nor does it imply approval to the exclusion of other products that may be suitable. USDA is an
25 equal opportunity provider and employer.

26 * Corresponding author Pérez-Díaz, Ilenys.Perez-Diaz@usda.gov.

ABSTRACT

The *Lactobacillus plantarum* and *Lactobacillus pentosus* genotypes existing in industrial-scale cucumber fermentations were defined using rep-PCR-(GTG)₅. The ability of each genotype to ferment cucumbers under various conditions was evaluated. Rep-PCR-(GTG)₅ was the technique capable of illustrating the most intraspecies discrimination compared to the sequencing of housekeeping genes (*recA*, *dnaK*, *pheS* and *rpoA*), MLST and RAPD with primers LP1, OPL5, M14 and COC. Ten genotypic clusters were defined for the 199 *L. pentosus* tested and three for the 17 *L. plantarum* clones. The ability of the 216 clones genotyped and 37 additional cucumber fermentation isolates, of the same species, to rapidly decrease the pH of cucumber juice medium under various combinations of sodium chloride (0 or 6%), initial pH (4.0 or 5.2) and temperatures (15 or 30 °C) was determined using a fractional factorial screening design. A reduced fermentation ability was observed for the *L. plantarum* strains as compared to *L. pentosus*, except for clone 3.2.8, which had a ropy phenotype and aligned to genotypic cluster A. *L. pentosus* strains belonging to three genotypic clusters (B, D and J) were more efficient in cucumber juice fermentation as compared to most *L. plantarum* strains. This research identified three genetically diverse *L. pentosus* strains and one *L. plantarum* as candidates for starter cultures for commercial cucumber fermentations.

Keywords: Cucumber fermentation; Lactic acid bacteria; Starter culture; Screening design; RAPD; Rep-PCR

1. Introduction

The development of starter cultures for commercial cucumber fermentation becomes imperative with the inception of technologies for low-salt pickle processing in the U.S (Fleming et al., 2002; McFeeters and Pérez-Díaz, 2010). While starter cultures are widely used by the dairy industry worldwide, this practice is less common in the commercial production of fermented vegetables. The majority of pickle processors currently rely on the indigenous microbiota and the application of cover brines with at least 0.96 M (5.6%) sodium chloride (NaCl) to achieve high quality, stable and safe fermentations. Hypertonic cover brines aid in the inhibition of unwanted pathogenic and spoilage associated microbes naturally present in the fruits at the initiation of a cucumber fermentation, thus creating a competitive advantage for the indigenous and desired lactic acid bacteria (LAB) (Davidson, 2001; Shelef and Seiter, 2005). The use of starter cultures in low-salt cucumber fermentations is expected to induce a rapid lactic acid production that results in a quick significant reduction of pH, which in turn propels the early exclusion of undesired microbes (McMurtrie et al., 2019; Pérez-Díaz et al., 2020). Utilization of selected starter culture has resulted in consistently accelerated food fermentations (Di Cagno et al., 2008, 2009; Ivey et al., 2013; Leroy and De Vuyst, 2004; Panagou et al., 2008; Tufariello et al., 2015). Moreover, it has been demonstrated that food fermentations with autochthonous starter cultures are more efficient as compared to allochthonous starter cultures or spontaneous fermentation (Di Cagno et al., 2008, 2009; Panagou et al., 2008; Tufariello et al., 2015). When autochthonous starter cultures are used, the organoleptic properties and nutritional attributes of fermented vegetables are favorably impacted and retained (Di Cagno et al., 2008, 2009, 2013; Panagou et al., 2008; Tufariello et al., 2015). Autochthonous bacteria have been observed to possess an intrinsic ability to produce metabolites such as bacteriocins and exopolysaccharides,

that are tailored specifically for their establishment during fermentation (Di Cagno et al., 2013). Some of the metabolites specifically produced by autochthonous starter cultures in their native food matrix are also often considered to add value to the final product (Di Cagno et al., 2009).

Utilization of autochthonous starter cultures of *Pediococcus cerevisiae* (currently reclassified as *P. pentosaceus*), *Lactobacillus plantarum* and *Lactobacillus brevis* has been shown to result in high acidity, cell densities and lower pH in pure cucumber fermentations with 0.92 to 0.96 M (5.4 -5.6% wt/vol) NaCl, conducted with fruits subjected to gamma radiation and blanching (Etchells et al., 1964) and used for controlled fermentations of vegetables in 0 to 1.1 M (0 to 6.5% wt/vol) NaCl (Etchells et al., 1976; Fleming et al., 1978, 1988). At the time, *L. plantarum* was considered the most prevalent bacterial species in commercial cucumber fermentations (Etchells and Jones, 1946; Pederson and Albury, 1956). Conversely, a recent study applying high throughput sequencing technology and culture based microbial classification methods found that *L. pentosus*, *L. brevis* and *P. pentosaceus* prevail in cucumber fermentations brined with 1.06 M NaCl (Pérez-Díaz et al., 2016). Discrepancies among the LAB species found in the two studies mentioned above are mostly due to the progress made in the identification and classification of this group of microbes between 1964 and the present time (Zanoni et al., 1987). An ideal autochthonous starter culture for cucumber fermentation would grow rapidly and vigorously by converting glucose and fructose, sugars naturally present in the fruit juice, to lactic acid exclusively (Etchells et al., 1973). Homofermentative metabolism excludes the formation of carbon dioxide from sugars, and consequently reduces the incidence of the formation of hollow cavities inside the fermented fruits, a processing defect known as bloated cucumber (Daeschel et al., 1984). The ability of the aforementioned species to efficiently ferment cucumbers under the

various conditions of temperature, initial fermentation pH and salt content representative of the commercial process is unknown.

Cucumber fermentations worldwide are often performed in open-top tanks using either fresh or recycled cover brine solutions. Such tanks are located in the field and exposed to outdoor temperatures. The active fermentation period proceeds during the late spring to early fall when outdoor temperatures range from 10 to 32 °C (50 to 90 °F) in the United States. The need to reduce the chloride content in waste waters generated in selected tank yards has forced the continuous recycling of cover brine solutions (McFeeters, 1977), which mildly acidify the fermentation at its start. Commercially, cucumbers in fresh cover brines would typically equilibrate to an initial pH between 6.5 and 4.5 (personal communication with pickle processors) depending on the buffering capacity of the fruits and the vinegar content of the full strength cover brine. The use of recycled cover brines may reduce the initial cover brine pH given its acidic content (personal communication with pickle processors). Therefore, the ideal starter culture for low salt cucumber fermentations should conduct a rapid fermentation under a variety of temperatures and varied initial pH. It is the objective of this study to screen a collection of *L. plantarum* and *L. pentosus* isolated from industrial-scale-cucumber fermentations (Pérez-Díaz et al., 2016) for genotypic diversity and the ability to rapidly initiate fermentation in cucumber juice medium (CJM). The *L. pentosus* and *L. plantarum* strains used in this study were obtained from two commercial fermentations that were geographically distant and performed with either fresh or recycled cover brine solutions (Pérez-Díaz et al., 2016). Four methods were considered to scrutinize genotypic diversity including the sequencing of housekeeping genes (*recA*, *dnaK*, *pheS* and *rpoA*), Multi Locus Sequence Typing (MLST), Random Amplified Polymorphic DNA (RAPD) with primers LP1, OPL5, M14 and COC and rep-PCR-(GTG)₅. A fractional factorial

screening design was used to evaluate the ability of the strains to ferment in a CJM with variable initial pH and salt concentration incubated at either 30 or 15 °C.

2. Materials and methods

2.1 Culture collection

A group of 243 *L. plantarum* and *L. pentosus* cultures isolated from the cover brine solutions of two commercial cucumber fermentation tanks were selected for this study. All of the cultures used in this study were streaked for purification on Lactobacilli deMan Rogosa and Sharpe (MRS; Cat No. 288130, Difco™ Laboratories, Detroit, MI, USA) agar from the MRS agar plates inoculated with the fermentation cover brine. Pure colonies were used for the preparation of frozen stocks in MRS broth containing 15% glycerol (Cat No. G5516, Sigma, St. Louis, MO, USA). The 243 isolates were identified by the sequencing of the 16S rDNA and *recA* amplification as described by Pérez-Díaz et al. (2016). This group of bacteria represents the most abundant species isolated during the peak of commercial fermentations (days 1, 3 and 7) (Pérez-Díaz et al. 2016). Fresh cultures were prepared in MRS broth from frozen stocks to perform all the analysis described below.

2.2 L. plantarum and L. pentosus genotyping and analysis

DNA extractions were successfully completed for a group of 216 isolates of which 199 had been identified as *L. pentosus* and 17 as *L. plantarum* (Pérez-Díaz et al. (2016). DNA extractions were attempted for an additional group of 37 isolates that had been identified as *L. pentosus*. Out of the 37 additional DNA extractions, 17 generated low quality DNA, 9 generated

DNA preparations of low concentration and 11 were initiated with insufficient cells. All DNA extractions were initiated from pure cultures in MRS broth using the MasterPure™ DNA purification kit (Cat No.: MCD85201, Epicentre, Madison, WI, USA) following the manufacturer's instructions. The DNA extracted from the 216 isolates was stored at -20 °C until further use.

Seven *L. pentosus* and *L. plantarum* strains out of the 216 in the collection were used to evaluate the application of housekeeping gene sequencing (Ghotbi et al., 2011; Huang et al., 2010; Nasser et al., 2007; Yu et al., 2012) for the discriminative identification to the strain level. The housekeeping genes used to discriminate members of the *L. plantarum* cluster (*L. plantarum*, *L. pentosus* and *L. paraplantarum*) were *recA*, *dnaK*, *pheS* and *rpoA* which were amplified as described in Table 1 (Huang et al., 2010; Naser et al., 2007; Torriani et al., 2001). Amplicon sequences were obtained using Sanger sequencing (Eton Biosciences, Research Triangle Park, NC, USA). The nucleotide sequences obtained were deposited in Genbank. Sequence data was formatted and analyzed using the BioEdit 7 software (www.mbio.ncsu.edu/bioedit). The sequences obtained were subjected to the basic local alignment search tool (BLAST) (Altschul et al., 1990; Benson et al., 2002; Pruitt et al., 2002) using the non-redundant nucleotide database (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) to confirm the identity of the strains (data not shown). Phylograms and similarity coefficients for the amplicon sequences obtained for *recA*, *pheS*, *rpoA* and *dnaK* were constructed using the Maximum Likelihood method based on the Tamura-Nei model with a bootstrap value of 1000 (Tamura and Nei, 1993). All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6 (Tamura et al., 2013).

MLST, RAPD with primers LP1, OPL-05, M14 and COC and rep-PCR-(GTG)₅ were also tested, with the subset of the seven strains used for the testing of the housekeeping genes, for their efficiency to discriminate intra-species diversity. *L. pentosus* ATCC 8041, *L. plantarum* WCSF1 and *L. plantarum* ATCC 14917 were used as reference strains (Table 2). MLST based on six *L. plantarum* housekeeping genes was performed as described by de las Rivas et al. (2006). MLST data was processed as described above for the housekeeping genes. Nucleotide sequences obtained were deposited in Genbank (Table 4). Results of the cultures identification based on the nucleotide sequence matches using the NCBI-BLAST and corresponding percentage similarity to the housekeeping genes in the database are presented in Table 5. RAPD was conducted using primers LP1 (Johansson et al., 1995; Plengvidhya et al., 2004), OPL-05 (Torriani et al., 1996), M14 (Zapparoli et al., 2000) and COC (Cocconcelli et al., 1995; Spano et al., 2002) applying the thermocycler conditions previously described (Cocconcelli et al., 1995; Johansson et al., 1995; Plengvidhya et al., 2004; Spano et al., 2002; Torriani et al., 1996; Zapparoli et al., 2000). Magnesium chloride was added to the RAPD-PCR reaction mixtures from a 10 mM stock solution (Genesee Scientific, San Diego, CA, USA) to 2.15 μ M for COC, 2.2 μ M for LP1, 2.05 μ M for OPL-05 and 2.2 μ M for M14. RAPD-PCR reactions had a total volume of 25 μ L. A BioRad PCR Master Mix (BioRad, Hercules, CA, USA) was used together with 10 mM rehydrated primer stocks synthesized by IDT Inc. (San Jose, CA, USA) and 1 μ L of whole genomic DNA suspended in nuclease free water (Ambion®, Life Technologies, Grand Island, NY, USA). Rep-PCR-(GTG)₅ was conducted as described by Versalovic et al. (1994) with some modifications on the PCR cycles and time. The conditions used for the Rep-PCR-(GTG)₅ were: initial denaturation at 95 °C for 3 min, followed by 29 cycles that consist of denaturation (95 °C for 60 s), annealing (40 °C for 60 s), extension (65 °C for 4 min), and final

extension at 65 °C for 16 min. Rep-PCR reactions were performed in 20 µl total volumes with the same components used for RAPD as described. Rep-PCR-(GTG)₅ was additionally conducted for the remaining 209 clones out of the 216 destined for phenotyping that had been subjected to DNA extraction.

Amplicons were run in a 1% agarose gel electrophoresis supplemented with 0.1% SYBR® safe (Invitrogen™, Life Technologies, Carlsbad, CA, USA) for 75 min at 85V using a 0.5X Tris, boric acid and EDTA buffer. Amplifications were performed in duplicate. The DNA markers used was a 1kb ladder from Promega (Madison, WI, USA). The gel images obtained from the gel electrophoresis runs were digitized, normalized, and analyzed using BioNumerics 7.6.3 software (Applied Maths, Belgium). Similarity matrices of the densitometric curves from each sample were calculated using the Pearson or DICE correlation coefficients, as indicated in the text, and clustered by unweighted pair group method with arithmetic averages (UPGMA).

2.3 Preparation of the inocula and CJM for the fermentation ability assay

The pure cultures were transferred from frozen stocks to CJM containing 2% NaCl (Morton Salt Inc., Chicago, IL, USA) and incubated under aerobic, static conditions for 4 d at 30 °C prior to the inoculation of the experimental media. At the time of transfer to CJM containing 2% NaCl, each isolate was concomitantly streaked onto MRS agar to verify purity. The CJM was composed of 60% (v/v) cucumber juice (CJ) and 40% (v/v) distilled water. CJ was prepared using fresh pickling cucumbers procured locally and a commercial juice maker (Warring Commercial, Madison, CT, USA). The CJ was strained through cheesecloth to remove large particulates prior to adding NaCl (Morton Pickling Salt) and diluting with distilled water. The

pH of the CJM with 2% salt was not adjusted (pH 5.6-5.8). CJM was spun at 4000 rpm for 60 min at ambient temperature using a bucket rotor on an Eppendorf Centrifuge, Model 5810 (Hamburg, Germany), and filter-sterilized using 0.22 micron filtration units (Nalgene®-Rapid Flow™, Thermo Scientific, Santa Clara, CA, USA). Aliquots of 1.2 mL of the filter-sterilized CJM containing 2% pickling salt were aseptically transferred to sterile microcentrifuge tubes to serve as the medium for the proliferation of the individual cultures. Cultures were calibrated for inoculation into the experimental CJM by measuring optical density at 600 nm using a spectrophotometer (OD₆₀₀, Thermo Spectronic Genesys20 Model 4001/4, Chelmsford, MA, USA). An initial OD₆₀₀ of 0.05 was targeted during inoculation, which was estimated to be equivalent to 10⁴ CFU/mL. Cultures were diluted 10X prior to measuring OD₆₀₀, unless turbidity corresponded to an OD₆₀₀ of less than 0.60.

2.4 Screening design for the fermentation ability assay and CJM modifications

A group of 243 pure cultures, composed of the 216 cultures used for genotyping and 37 additional isolates from commercial cucumber fermentations, was tested for their ability to ferment cucumbers under various pH, NaCl, and temperature conditions using CJM as a model system. A fractional factorial screening design (JMP version 10.0, SAS, Cary, NC, USA) was used to define various treatment combinations to be evaluated as follows: Treatment 1 (T1)- (pH 5.4; 0% NaCl; 15 °C incubation), Treatment 2 (T2)- (pH 4.0; 6% NaCl; 15 °C incubation), Treatment 3 (T3)- (pH 5.4; 6% NaCl; 30 °C incubation) and Treatment 4 (T4)- (pH 4.0; 0% NaCl; 30 °C incubation).

CJ and CJM were prepared as described above. The pH of the experimental CJM was adjusted using an Accumet® Research 25 pH meter (Fisher Scientific, Carlsbad, CA, USA)

equipped with a Gel-Filled Pencil-Thin pH Combination Electrode (Accumet, Fisher Scientific) and minimal volumes of 3N HCl (Spectrum Chemicals, New Brunswick, NJ, USA). The unadjusted pH of the experimental CJM without salt was 5.4 ± 0.1 . Lowering the pH to 4.0 in treatments 2 and 4 formed precipitates which were eliminated subsequently in the pellet after the centrifugation step. The CJM were spun at 4,000 rpm for 30 min at ambient temperature (Eppendorf Centrifuge Model 5810). The CJM were filtered-sterilized using 0.22 micron filtration units (Nalgene®-Rapid Flow™, Thermo Scientific). Aliquots of 1 mL of the experimental CJM were aseptically dispensed into sterile microcentrifuge tubes and incubated at the corresponding treatment temperatures 48 h prior to inoculation. The experimental CJ media was inoculated with the pre-calibrated cultures prepared as described above and incubated for 48 h at either 15 or 30 °C under static aerobic conditions.

The ability of each isolate to rapidly initiate fermentation under each of the four treatment conditions was assessed by measuring the fermentation pH (Denver Instruments, model UB-5, Bohemia, NY, USA) after 48 h incubation time. pH meters were combined with Gel-Filled Pencil-Thin pH Combination Electrodes (Accumet Fisher Scientific, Pittsburgh, PA, USA) so that pH measurements could be made directly from the CJM cultures in microcentrifuge tubes. This experiment was performed in duplicate using different lots of cucumbers for the growth and experimental media and independent pre-calibrated inocula. The pH upper threshold values for selection of rapid fermenters for T1, T2, T3 and T4 were 4.8, 3.9, 3.4 and 3.5 respectively. A manual selection of the isolates that were able to reduce the pH below experimentally determined threshold values for any combination of three treatments was made. Hierarchical two-way Cluster Analysis (HCA) of the *L. plantarum* and the *L. pentosus* clones was performed based on their average 48 h fermentation pH in each of the four treatments (JMP version 10.0 SAS

Institute). A screening analysis was conducted to determine the independent variables that most significantly impacted the ability of each culture to rapidly initiate fermentation (JMP version 10.0, SAS), and the number of isolates significantly impacted by each variable ($P < 0.05$) was presented in a Venn diagram and used to calculate the percentage of cultures impacted by each variable.

3. Results and Discussion

The genotypic characterization of a subgroup of clones using housekeeping genes, MLST, RAPD or rep-PCR-(GTG)₅ illustrated, to the extent possible, the magnitude of the genetic differences existing within the species of interest. Rep-PCR-(GTG)₅ generated the finer discrimination power for intra-species biodiversity. The genotypic clustering diverged from the phenotypic clustering using fermentation ability profile as an indicator. Nevertheless, three distinct *L. pentosus* strains and one *L. plantarum* were identified as candidates for starter cultures in commercial cucumber fermentations based on the results obtained from the rep-PCR-(GTG)₅ and the fermentation ability assay.

3.1 Definition of phenotypic biodiversity among *L. plantarum* and *L. pentosus* clones isolated from commercial cucumber fermentations

Members of the *L. plantarum* phylogenetic group, composed of *L. plantarum*, *L. pentosus* and *L. paraplantarum*, are indistinguishable from their *16S rRNA* gene sequences given that they share 99% similarity (Bringel et al., 2001; Collins et al., 1991; Ghotbi et al., 2011; Singh et al., 2009; Torriani et al., 2001; Yu et al., 2012). Amplification and sequencing of ubiquitous and conserved protein coding genes such as *recA*, *pheS*, *rpoA* and *dnaK* have been proposed and used

for the discrimination of *L. plantarum*, *L. pentosus* and *L. paraplantarum* (Anukam et al., 2013; Ghotbi et al., 2011; Huang et al., 2010; Kotik-Kogan et al., 2005; Park et al., 2010; Torriani et al., 2001). The amplification of *recA*, coding for the enzyme recombinase A (Bianco and Kowalczykowski, 1998), generates amplicons of 318, 218 and 107 bp from *L. plantarum*, *L. pentosus* and *L. paraplantarum* genomic DNA, respectively (Anukam et al. 2013; Ghotbi et al. 2011; Torriani et al. 2001). Six clones and the *L. pentosus* reference strains generated amplicon sizes representative of *L. pentosus* (218 bp) among the subgroup of seven clones and three reference strains tested. The sequencing of the amplicons confirmed that those six strains belong to *L. pentosus* and that the *L. plantarum* reference strain LA0045 had to be reclassified as *L. pentosus*. One clone out of the seven in the subgroup, 3.2.8, and the reference strain *L. plantarum* WCFS1 generated amplicons of 318 bp representative of the *L. plantarum* species. The sequences from the amplicons derived from clone 3.2.8 DNA and the reference strain *L. plantarum* WCFS1 had 100% similarity to the *L. plantarum recA* upon alignment. A phylogram constructed based on the *recA* sequence homology showed that 6 cultures and LA0445 are 100% similar to the *recA* sequence derived from the reference strain *L. pentosus* ATCC8041 (Figure 1). Similarly, clone 3.2.8 clustered with the *L. plantarum* reference cultures WCFS1 and ATCC14917 (Figure 1). As expected, the amplification of *recA* and sequencing allowed the discrimination of the clones to the species level and enabled the intraspecies segregation of strains (Figure 1).

As expected, *pheS* produced a finer differentiation with respect to percentage similarity among various clones as compared to *recA*, *dnaK*, coding for a heat shock protein 70 (Huang et al., 2010), and *rpoA*, coding for a DNA-directed RNA polymerase α -subunit (Park et al., 2010) (Figure 1). *pheS* codes for the phenylalanine t-RNA synthetases involved in translation of

phenylalanine (Kotik-Kogan et al., 2005). *pheS* and *rpoA* are exceptionally conserved and widely used for evolutionary based identification purposes of various organisms (Kotik-Kogan et al., 2005; Naser et al., 2007; Park et al., 2010). *pheS* and *rpoA* sequences have an interspecies gap of only 10% and 5%, respectively, and thus considered an effective tool for identification and discrimination of closely related lactobacilli (Naser et al., 2007). Alignment of amplicon sequences for the housekeeping genes *pheS*, *rpoA* and *dnaK* derived from the subgroup of seven clones confirmed the species identity, *rpoA* and *dnaK* discriminated among strains and *pheS* presented a finer strain level discrimination (Table 3 and Figure 1). Sequences of the amplicons obtained for *pheS*, *rpoA*, and *dnaK* from six clones, excluding 3.2.8, presented at least 99% similarity to the *L. pentosus* homolog (alignments not shown; Table 3). Once again, the characteristics of clone 3.2.8 derived amplicons were in line with those expected for *L. plantarum*. Three phylogenetic clusters were defined based on the maximum likelihood method from the more polymorphic *pheS* amplicon-sequences distinguishing among clones isolated from a single habitat (Figure 1).

MLST is an extension of the multilocus enzyme electrophoresis method but targets alleles of different housekeeping gene loci (Farfán et al., 2002). Instead of using electrophoretic mobility to distinguish strain level differences, MLST targets differences in housekeeping gene sequences. MLST has been used to determine strains of multiple bacteria of public health significance (Dingle et al., 2001; Farfán et al., 2002; Helgason et al., 2004) as well as some LAB such as *L. plantarum* (de las Rivas et al., 2006) and *Oenococcus oeni* (de las Rivas et al., 2004). In this study we followed the MLST method described by de las Rivas et al. (2006) for *L. plantarum* given that the design of universal primers for the homologous genes in *L. pentosus* would have been limited by the number of genome sequences available for such species to date.

The MLST analysis conducted included the amplification of *mutS* (mismatch repair protein), *purK1* (ATPase subunit of phosphoribosylaminoimidazole carboxylase), *ddl* (D-alanine-D-alanine Ligase), *gyrB* (B subunit of DNA gyrase), *gdh* (glutamate dehydrogenase) and *pgm* (phosphoglucomutase). The majority of the strains (6) subjected to the analysis shared 93-99% similarity to homologous genes in previously identified *L. pentosus* (Tables 4 and 5). Only the 3.2.8 clone presented gene similarities to those reported for previously characterized *L. plantarum* (Tables 4 and 5). Interestingly, primers designed for the amplification of *gyrB* and *pgm* from *L. plantarum* by de las Rivas et al. (2006) did not successfully function for the amplification of the *L. pentosus* counterparts. Thus, utilization of the MLST approach designed for *L. plantarum* could not be tested for the strain level differentiation of *L. pentosus* given the lack of amplification of certain genes. Primers specifically targeting the *L. pentosus gyrB* and *pgm* would have to be designed and tested to determine if this technique can effectively discriminate *L. pentosus* strains. A new reference database for *L. pentosus* MLST patterns would be needed.

The LA0445 strain isolated from cucumber fermentations in 1988 and originally classified as *L. plantarum* based on the partial sequence of the 16S *rRNA* gene (Fleming et al., 1988; McDonald et al., 1993) showed similarity to the *L. pentosus* reference cultures more so than to the *L. plantarum* reference strain. LA0445 clustered with the *L. pentosus* strains isolated from cucumber fermentations based on the sequences for *recA*, *pheS*, *rpoA* and *dnaK* (Figure 1). Although the LA0445 16S *rRNA* gene sequence shares 100% similarity with the *L. plantarum* in the NCBI BLAST database (Table 5), its MLST profile obtained using primers designed for *L. plantarum*, represented that of *L. pentosus* strains (Table 5). The LA0445 *mutS*, *purK1*, *ddl* and *gdh* genes are 99, 99, 99, and 86% similar to the homologous genes in the *L. pentosus* reference

strain ATCC8041 isolated from sauerkraut by Fred et al. in 1921. Thus, we proposed that this strain be re-classified as *L. pentosus* LA0445.

RAPD was originally proposed by Williams et al. (1990) as a DNA polymorphism assay for eukaryotic organisms based on the amplification of random DNA segments with single primers of arbitrary nucleotide sequences. The method was adapted to LAB using a 10-mer primer (Cocconcelli et al., 1995). RAPD with primers P1, P2 and P3 was proposed by Bringel et al. (2001) for the discrimination of *L. plantarum* strains into two subgroups and the segregation of such species from *L. pentosus* and *L. paraplantarum*. Thus, genomic variability among the selected cultures was characterized using RAPD (Spano et al., 2002). Utilization of primers M14 and OPL5 (Torriani et al., 1996; Zapparoli et al., 2000) for RAPD did not yield band patterns for the seven clones tested under the conditions of the assay (data not shown). PCR amplification using primer LP1 (Plengvidhya et al., 2004) was useful for the generation of band patterns that differentiated a few *L. pentosus* clones (data not shown). RAPD using the COC primer (Cocconcelli et al., 1995) yielded defined band patterns illustrating the heterogeneity among the *L. pentosus* isolated from distant habitats and commercial cucumber fermentations (Figure 2).

Rep-PCR-(GTG)₅ amplifies regions between interspersed repetitive elements and was introduced to differentiate microbes using PCR and gel electrophoresis fractionation (Versalovic et al., 1994). Application of rep-PCR-(GTG)₅ for the discrimination of the clones of interest yielded defined band patterns showing the heterogeneity among the *L. pentosus* and *L. plantarum* clones (Figure 3). Similarity coefficients from 22 to 99 and 80 to 99 were calculated for the band patterns generated using the (GTG)₅ primer among the *L. pentosus* and *L. plantarum* clones, respectively (Figure 3).

The intraspecies genetic diversity among the remaining 209 lactobacilli clones was evaluated using rep-PCR-(GTG)₅, given that such technique has been more commonly used for the discrimination of *L. plantarum* and *L. pentosus* strains with reproducible results (Gevers et al., 2001; Kostinek et al., 2005; Panagou et al., 2008; Pino et al., 2018; Tamang et al., 2005). Three and ten genotypic clusters were found among the *L. plantarum* and *L. pentosus* strains isolated from commercial cucumber fermentations, respectively (Figures 4 and 5). The clustering obtained is in agreement with those typically generated to document the biodiversity of *L. plantarum* and *L. pentosus* strains isolated from fermented foods typically illustrating 8 to 11 clusters based on rep-PCR-(GTG)₅ derived band patterns (Gevers et al., 2001; Kostinek et al., 2005; Panagou et al., 2008; Pino et al., 2018; Tamang et al., 2005).

3.2 Phenotypic biodiversity among L. plantarum and L. pentosus clones isolated from commercial cucumber fermentations

The mean pH of the CJM at the end point of incubation resulting from the fermentation by the isolates inoculated in treatments T1, T2, T3 and T4 were 4.80 ± 0.32 , 3.87 ± 0.12 , 3.59 ± 0.46 and 3.54 ± 0.32 , respectively. The initial pH of the CJM significantly affected the ability of 66% of the *L. plantarum* and *L. pentosus* cultures to rapidly initiate fermentation ($P < 0.05$). Temperature alone affected approximately 60% of the *L. plantarum* and *L. pentosus* cultures, while no cultures were solely influenced by NaCl concentration (Figure 6). All the 101 *L. plantarum* and *L. pentosus* cultures that appeared to be impacted by NaCl concentration were also influenced by initial pH and/or temperature, suggesting that the other two variables and potential synergies between the three variables were more useful in screening for robust starter culture candidates.

Sixty-six percent of the initial pool of 243 bacterial clones collected from two independent commercial cucumber fermentations lacked the ability to rapidly decrease the experimental CJM pH in varied salt concentration, initial pH, and incubation temperature. The remaining 34% of the cultures were able to ferment in all four treatment combinations, with a few performing well in only three treatments (Figure 7). A significant number of the *L. plantarum* and *L. pentosus* cultures that performed well in all treatments, fermented most rapidly in the experimental CJM that had an initial pH of 5.4, 6% NaCl, and was incubated at 30 °C (T3), as illustrated by the primarily green 4th column from the left in the HCA dendogram (Figure 7). Screening using only the T3 treatment, which represents a common cucumber fermentation model, would have eliminated only 9 to 12% of the pure cultures from the group of candidates (Figure 7, cluster 7). Furthermore, selecting only the top fermenters under ideal conditions may have resulted in the exclusion of isolates that have better potential for performance in commercial environments that are suboptimal. For example, several *L. plantarum* and *L. pentosus* isolates in clusters 4, 5 and 6 of Figure 7, are shown in dark green for T3, indicating that a low fermentation pH was achieved in a short time period. However, these isolates did not perform well under the other conditions presented, as indicated by the red shading for other treatments for the cultures in these clusters (Figure 7). The most robust starter culture candidates grouped together in clusters 1, 2, and 3. The cluster analysis also highlighted one particular *L. pentosus* strain in cluster 3 (14.2.1 RB) that exhibited robust fermentation ability under all conditions tested.

Supplementary figure 1 merges the genotypic, rep-PCR-(GTG)₅ dendograms shown in Figures 4 and 5 with the phenotype of each individual clone as determined by the fermentation ability experiment (Figure 7). The fermentation ability phenotypes for each clone was populated

into the genotypic Rep-PCR-(GTG)5 dendrogram to illustrate the distribution of phenotypes among the genotypic clusters. The dendrogram in Figure S1 combines the dendrograms shown in Figure 4 and 5 and the color-coded phenotype corresponding to each individual clone to the right of the bands. The dendrogram clusters in Figure S1 were defined with a 40% cut-off value using Bionumerics as described above. *L. plantarum* clones grouped in genotypic clusters A, I and K with 4, 2 and 11 members, respectively. Genotypic cluster A is composed of 4 *L. plantarum* clones, including clone 3.2.8, and several *L. pentosus*, suggesting that some *L. plantarum* are genetically closer to *L. pentosus* than most others, which grouped in cluster K with 11 members. Best *L. pentosus* fermenters (Figure 7, cluster 1, 2, and 3) are scattered across the genotypic dendrogram with more clones aligning to clusters B, D, and J (Figure S1). The remaining clones are scattered across clusters A, C, E, F, G, H, and I. Interestingly, 77% (43 out of 56) of the robust fermenters, including *L. pentosus* and *L. plantarum*, in the fermentation ability cluster 1 (Figure 7) show two high molecular weight bands of around 1,600 bp and 1,000 bp (Figure S1). The other 25% show at least three distinctive bands. The inspection of bands for all the robust fermenters show no discernable pattern that might indicate a relationship between the amount of bands or molecular weight and fermentation ability. In addition, each cluster in the genotypic dendrogram contains various phenotypic clusters.

Furthermore, the Rep-PCR-(GTG)5 phylogenetic dendrograms published by Tamang et al. (2005) and Pino et al. (2018) for lactobacilli isolated from traditional vegetable fermentations were compared to those generated in this study (Figure S1). Tamang et al. (2005) and Pino et al. (2018) presented dendrograms with 15 and 60 samples, respectively, and used the Pearson correlation coefficient and UPGMA for clustering. The electrophoresis parameters applied by these two groups differed from the ones used in this study. However, the clusters from the

dendrogram by Tamang et al. (2005) are comparable with clusters A, H, and I in Figure S1, based on band patterns with at least 8 to 11 bands per isolate. The various clusters from Pino et al. (2018) are comparable to all of the clusters in this study, ranging from 3 to 11 bands per isolate. This demonstrates the ability and reproducibility of Rep-PCR-(GTG)₅ across different and varied parameters and clustering methods. However, based on the results a complete view of biodiversity is dependent on the amount of samples used in the study. A large sample size (>60 isolates) is needed to detect all biotypes.

Genotypic discrimination of the clones isolated from commercial cucumber fermentations resulted in the identification of 10 and 3 clusters or strains of *L. pentosus* and *L. plantarum*, respectively. These clusters do not align with the strains ability to ferment CJM. A combination of the effective tools used here for genotyping may prove useful as a strain discriminatory tool for *L. pentosus* and *L. plantarum* that more closely resembles fermentation robustness. A combination of the sequencing of the housekeeping gene *pheS*, *pgm* and *gyrB*, RAPD-COC and rep-PCR-(GTG)₅ may serve as an improved genotypic discriminatory tool for strains of *L. pentosus*.

The phenotypic clustering of the strains of interest based on their ability to ferment cucumbers demonstrated that the *L. pentosus* species is capable of completing the bioconversion more effectively under various conditions of pH, salt and temperature than *L. plantarum*. The effect of temperature and initial pH is more significant than the salt effect on the growth of *L. pentosus* in cucumber fermentations. This study identifies three *L. pentosus* strains and one *L. plantarum* strain for further scrutiny in the development of autochthonous starter cultures for commercial cucumber fermentations.

4. Conclusion

This study identified the *L. plantarum* and *L. pentosus* genotypes present in industrial-scale cucumber fermentations and determine their ability to ferment cucumbers under various conditions. The genotypic clustering of 216 clones was facilitated by rep-PCR-(GTG)₅, a technique that illustrated the most intraspecies biodiversity among RAPD, housekeeping gene sequencing and MLST. The genotypic clustering including 10 and 3 clades for the 216 *L. pentosus* (199) and *L. plantarum* (17) clones, respectively, significantly differ from the phenotypic clustering based on fermentation ability. A reduced fermentation ability was observed for the *L. plantarum* strains as compared to *L. pentosus*, except for clone 3.2.8, which had a ropy phenotype. *L. pentosus* strains belonging to three genotypic clusters (B, D and J) were more efficient in CJM fermentation as compared to most *L. plantarum* strains. *L. pentosus* strains isolated from cucumber cover brine samples collected on days 1 and 3 fermented CJM more efficiently than those isolated on day 7. This research identified three genetically diverse *L. pentosus* strains and one *L. plantarum* as candidates for starter cultures for commercial cucumber fermentations able to acidify under various conditions.

Acknowledgements

The authors thank Ms. Janet Hayes and Ms. Jane Caldwell at the USDA-ARS, SEA Food Science Research Unit located in Raleigh, NC for technical assistance, Mr. Clyde Manuel for his assistance with the MEGA6 software and Ms. Sandra Parker for administrative assistance. We also thank Dr. Brian Farkas and Dr. Gabriel Harris with the Department of Food, Nutrition and Bioprocessing Sciences at North Carolina State University for facilitating the participation of several undergraduate students in this research.

435

436 **Funding sources**

437 This research did not receive any specific grant from funding agencies in the public, commercial,
438 or not-for-profit sectors.

439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J Mol Biol* **215(3)**, 403-10.
- Anukam, K.C., Macklaim, J.M., Gloor, G.B., Reid, G., Boekhorst, J., Renckens, B., van Hijum, S.A., Siezen, R.J., 2013. Genome sequence of *Lactobacillus pentosus* KCA1: Vaginal isolate from a healthy premenopausal woman. *PloS one* **8(3)**, e59239.
- Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Rapp, B.A., Wheeler, D.L., 2002. GenBank. *Nucleic Acids Res* **30(1)**, 17-20.
- Bianco, P.R., Kowalczykowski, S.C., 1998. RecA protein. In Encyclopedia of Life Science. pp. 1-11. London: Nature Publishing Group (<http://www.els.net>).
- Bringel, F., Qune, P., Tailliez, P., 2001. Polyphasic investigation of the diversity within *Lactobacillus plantarum* related strains revealed two *L. plantarum* subgroups. *Syst. Appl. Microbiol.* **24(4)**, 561-571.
- Cocconcelli, P., Porro, D., Galandini, S., Senini, L., 1995. Development of RAPD protocol for typing of strains of lactic acid bacteria and enterococci. *Lett. Appl. Microbiol.* **21(6)**, 376-379.
- Collins, M.D., Rodrigues, U., Ash, C., Aguirre, M., Farrow, J.A.E., Martinez-Murcia, A., Phillips, B.A., Williams, A.M., Wallbanks, S., 1991. Phylogenetic analysis of the genus *Lactobacillus* and related lactic acid bacteria as determined by reverse transcriptase sequencing of 16S rRNA. *FEMS Microbiol. Lett.* **77(1)**, 5-12.
- Daeschel, M.A., McFeeters, R., Fleming, H., Klaenhammer, T., Sanozky, R., 1984. Mutation and selection of *Lactobacillus plantarum* strains that do not produce carbon dioxide from malate. *Appl Environ Microbiol* **47(2)**, 419-20.

Davidson, P.M., 2001. Chemical preservatives and natural antimicrobial compounds, Food microbiology: Fundamentals and frontiers. Doyle MP, Beauchat LR, Montville TJ, editors. Washington, DC: ASM Press.

de las Rivas, B., Marcobal, Á., Muñoz, R., 2004. Allelic diversity and population structure in *Oenococcus oeni* as determined from sequence analysis of housekeeping genes. *Appl Environ Microbiol* DOI: 10.1128/AEM.70.12.7210-7219.2004

de las Rivas, B., Marcobal, Á., Muñoz, R., 2006. Development of a multilocus sequence typing method for analysis of *Lactobacillus plantarum* strains. *Microbiol.* **152**(1), 85-93.

Di Cagno, R., Coda, R., De Angelis, M., Gobbetti, M., 2013. Exploitation of vegetables and fruits through lactic acid fermentation. *Food Microbiol* **33**, 1-10.

Di Cagno, R., Surico, R. F., Minervini, G., De Angelis, M., Rissello, C. G., Gobbetti, M., 2009. Use of autochthonous starters to ferment red and yellow peppers (*Capsium annum L.*) to be stored at room temperature. *Int. J. Food Microbiol.* **130**, 108-116.

Di Cagno, R., Surico, R.F., Siragusa, S., De Angelis, M., Paradiso, A., Minervini, F., De Gara, L., Gobbetti, M., 2008. Selection and use of autochthonous mixed starter for lactic acid fermentation of carrots, French beans or marrows. *Int J Food Microbiol* **127**, 220-228.

Dingle, K.E., Colles, F.M., Wareing, D.R.A., Ure, R., Fox, A.J., Bolton, F.E., Bootsma, H.R., Willems, J.L., Urwin, R., Maiden, M.C.J., 2001. Multilocus sequence typing system for *Campylobacter jejuni*. *J. Clin. Microbiol.* **39**, 14-23.

Etchells, J.L., Bell, T.A., Fleming, H.P., Kelling, R.E., Thompson, R.L., 1973. Suggested procedure for the controlled fermentation of commercially brined pickling cucumbers-The use of starter cultures and reduction of carbon dioxide accumulation. *Pickle Pak Sci* **3**(1), 4-14.

485 Etchells, J.L., Bell, T.A., Fleming, H.P., Thompson, R.L., 1976. Controlled bulk vegetable
 486 fermentation. U. S. No. 3,932,674.

487 Etchells, J.L., Costilow, R.N., Anderson, T.E., Bell, T.A., 1964. Pure culture fermentation of
 488 brined cucumbers. *Appl. Microbiol.* **12(6)**:523-535.

489 Etchells, J.L., Jones, I.D., 1946. Characteristics of lactic acid bacteria from commercial
 490 cucumber fermentations. *J. Bacteriol.* **52(5)**, 593-599.

491 Farfán, M., Miñana-Galbis, D., Fusté, M.C., Lorén, J.G., 2002. Allelic diversity and population
 492 structure in *Vibrio cholerae* O139 bengal based on nucleotide sequence analysis. *J. Bacteriol.*
 493 DOI: 10.1128/JB.184.5.1304-1313.2002

494 Fleming, H.P., Humphries, E.G., Thompson, R.L., McFeeters, R.F., 2002. BAG-IN-BOX
 495 TECHNOLOGY: Storage stability of process-ready, fermented cucumbers. *Pickle Pak Sci.*
 496 **8(1)**, 14-18.

497 Fleming, H.P., McFeeters, R.M., Daeschel, M.A., Humphries, E.G., Thompson, R.L., 1988.
 498 Fermentation of cucumbers in anaerobic tanks. *J. Food Sci.* **53(1)**, 127-133.

499 Fleming, H.P., Thompson, R.L., Bell, T.A., Hontz, L.H., 1978. Controlled fermentation of sliced
 500 cucumbers. *J. Food Sci.* **43(3)**, 888-891.

501 Fred, E.B., Peterson, W.H., Anderson, J.A., 1921. The characteristics of certain pentose destroying
 502 bacteria, especially as concerns their action on arabinose and xylose. *J. Bio. Chem.*, 48(2),
 503 385-412.

504 Gevers, D., Huys, G., Swings, J., 2001. Applicability of rep-PCR fingerprinting for identification
 505 of *Lactobacillus* species. *FEMS Microbiol. Lett.* **205**, 31-36.

506 Ghotbi, M., Soleimani-Zad, S., Sheikh-Zeinoddin, M., 2011. Identification of *Lactobacillus*
 507 *pentosus*, *Lactobacillus paraplantarum* and *Lactobacillus plantarum* in Lighvan cheese with

508 4 month ripening period by means of *recA* gene sequence analysis. *Afr J Biotechnol* **10**,
509 1902-1906.

510 Helgason, E., Tourasse, N. J., Meisal, R., Caugant, D. A., Kolsto, A. B., 2004. Multilocus
511 sequence typing scheme for bacteria of the *Bacillus cereus* group. *Appl Environ*
512 *Microbiol* **70**, 191–201.

513 Hols, P., Slos, P., Dutot, P., Reymund, J., Chabot, P., Delplace, B., Delcour, J., Mercenier, A.,
514 1997. Efficient secretion of the model antigen M6-gp41E in *Lactobacillus plantarum* NCIMB
515 8826. *Microbiol.*, **143(8)**, 2733-2741.

516 Huang, C.-H., Lee, F.-L., Liou, J.-S., 2010. Rapid discrimination and classification of the
517 *Lactobacillus plantarum* group based on a partial *dnaK* sequence and DNA fingerprinting
518 techniques. *Antonie Van Leeuwenhoek* **97(3)**, 289-296.

519 Ivey, M., Massel, M., Phister, T.G., 2013. Microbial interactions in food fermentations. *Ann Rev*
520 *Sci Technol* **4**, 141-162.

521 Johansson, M.L., Quednau, M., Molin, G., Ahrné, S., 1995. Randomly amplified polymorphic
522 DNA (RAPD) for rapid typing of *Lactobacillus plantarum* strains. *Lett. Appl. Microbiol.* **21**,
523 155-159.

524 Kostinek, M., Pukall, R., Rooney, A.P., Schillinger, U., Hertel, C., Holzapfel, W.H., Franz,
525 C.M.A., 2005. *Lactobacillus arizonensis* is a later heterotypic synonym of *Lactobacillus*
526 *plantarum*. *Int. J. Sys. Evol. Microbiol.* **55**, 2485-2489. doi: 10.1099/ij.s.0.63880-0

527 Kotik-Kogan, O., Moor, N., Tworowski, D., Safro, M., 2005. Structural basis for discrimination
528 of L-phenylalanine from L-tyrosine by phenylalanyl-tRNA synthetase. *Struct* **13(12)**, 1799-
529 807.

530 Leroy, F., De Vuyst, L., 2004. Lactic acid bacteria as functional starter cultures for the food
531 fermentation industry. *Trends Food Sci Technol* **15**, 67-78.

532 McDonald, L.C., Shieh, D.H., Fleming, H.P., McFeeters, R.F., Thompson, R.L., 1993.
 533 Evaluation of malolactic-deficient strains of *Lactobacillus plantarum* for use in cucumber
 534 fermentations. *Food Microbiol* **10(6)**, 489-499.
 535 McFeeters, R.F., 1977. Reuse of fermentation brines in the cucumber pickling industry executive
 536 summary. United States Environmental Protection Agency, Doc. No. 5771.
 537 McFeeters, R.F., Pérez-Díaz, I.M., 2010. Fermentation of cucumbers brined with calcium
 538 chloride instead of sodium chloride. *J Food Sci* **75(3)**, C291-C296.
 539 McMurtrie, E.K., Johanningsmeier, S.D., Breidt, F., Price, R.E., 2019. Effect of brine
 540 acidification on fermentation microbiota, chemistry and texture quality of cucumbers
 541 fermented in calcium or sodium chloride brines. *J Food Sci* **84(5)**: 1129-1137.
 542 Naser, S.M., Dawyndt, P., Hoste, B., Gevers, D., Vandemeulebroecke, K., Cleenwerck, I.,
 543 Vancanneyt, M., Swings, J., 2007. Identification of lactobacilli by *pheS* and *rpoA* gene
 544 sequence analyses. *Int J Syst Evol Microbiol* **57(12)**, 2777-2789.
 545 Panagou, E.Z., Schillinger, U., Franz, C.M.A.P., Nychas, G.J.E., 2008. Microbiological and
 546 biochemical profile of cv. Conservolea naturally black olives during controlled fermentation
 547 with selected strains of lactic acid bacteria. *Food Microbiol.* **2**, 348–358. doi:
 548 10.1016/j.fm.2007.10.005.
 549 Park, H.K., Yoon, J.W., Shin, J.W., Kim, J.Y., Kim, W., 2010. *rpoA* is a useful gene for
 550 identification and classification of *Streptococcus pneumoniae* from the closely related
 551 viridans group streptococci. *FEMS Microbiol Lett* **305(1)**, 58-64.
 552 Pederson, C.S., Albury, M.N., 1956. Variations in bacterial flora of low salt cucumber brines.
 553 *Appl. Microbiol.* **4**, 259-263.

554 Pérez-Díaz, I.M., Dickey, A. N., Fitria, R., Ravishankar, N., Hayes, J., Campbell, K., Arritt, F.,
 555 2020. Modulation of the bacterial population in commercial cucumber fermentations by
 556 brining salt type. *J. Appl. Microbiol.* *submitted*
 557 Pérez-Díaz, I.M., Hayes, J.S., Medina-Pradas, E., Anekella, K., Daughtry, K.V., Dieck, S., Levi,
 558 M., Price, R., Butz, N., Lu, Z., Azcarte-Peril, M., 2016. Reassessment of the succession of
 559 lactic acid bacteria in commercial cucumber fermentations and physiological and genomic
 560 features associated with their dominance. *Food Microbiol.* **63**, 217-227 DOI:
 561 10.1016/j.fm.2016.11.025.
 562 Pino, A., De Angelis, M., Todaro, A., Van Hoorde, K., Randazzo, C. L., Caggia, C., 2018.
 563 Fermentation of *Nocellara etnea* table olives by functional starter cultures at different low
 564 salt concentrations. *Front. Microbiol.* **9**, 1125.
 565 Plengvidhya, V., Breidt, Jr., F., Fleming, H., 2004. Use of RAPD-PCR as a method to follow the
 566 progress of starter cultures in sauerkraut fermentation. *Int J Food Microbiol* **93(3)**, 287-296.
 567 Pruitt, K., Brown, G., Tatusova, T., Maglott, D., 2002. Chapter 18. The Reference Sequence
 568 (RefSeq) Database. In The NCBI handbook [Internet] ed. McEntyre, J. and Ostell, J.
 569 Bethesda, MD: National Center for Biotechnology Information (US). [Updated 2012 Apr 6].
 570 Singh, S., Goswami, P., Singh, R., Heller, K.J., 2009. Application of molecular identification
 571 tools for *Lactobacillus*, with a focus on discrimination between closely related species: A
 572 review. *LWT - Food Sci Technol* **42(2)**, 448-457.
 573 Shelef LA, Seiter J., 2005. Indirect and miscellaneous antimicrobials, Antimicrobials in
 574 food. 3rd ed. Davidson PM, Sofos JN, Larry BA, editors. Boca Raton, FL: Taylor and
 575 Francis; pp. 573–598.

Spano, G., Beneduce, L., Tarantino, D., Zapparoli, G., Massa, S., 2002. Characterization of *Lactobacillus plantarum* from wine must by PCR species-specific and RAPD-PCR. *Lett Appl Microbiol* **35**(5), 370-374.

Tamang, J., Tamang, B., Schillinger, U., Franz, C.M.A.P., Gores, M., Holzapfel, W.H., 2005. Identification of predominant lactic acid bacteria isolated from traditionally fermented vegetable products of the eastern Himalayas. *Int. J. Food Microbiol.* **105**, 347–356.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10**(3), 512-526.

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**(12), 2725-2729.

Torriani, S., Felis, G.E. and Dellaglio, F., 2001. Differentiation of *Lactobacillus plantarum*, *L. pentosus*, and *L. paraplantarum* by *recA* gene sequence analysis and multiplex PCR assay with *recA* gene-derived primers. *Appl Environ Microbiol* **67**(8), 3450-4.

Torriani, S., Van Reenen, C., Klein, G., Reuter, G., Dellaglio, F., Dicks, L., 1996. *Lactobacillus curvatus* subsp. *curvatus* subsp. nov. and *Lactobacillus curvatus* subsp. *melibiosus* subsp. nov. and *Lactobacillus sake* subsp. *sake* subsp. nov. and *Lactobacillus sake* subsp. *carnosus* subsp. nov., new subspecies of *Lactobacillus curvatus* Abo-Elnaga and Kandler 1965 and *Lactobacillus sake* Katagiri, Kitahara, and Fukami 1934 (Klein *et al.* 1996, emended descriptions), respectively. *Int J Syst Bacteriol* **46**(4), 1158-1163.

Tufariello, M., Durante, M., Ramires, F.A., Grieco, F., Tommasi, L., Perbellini, E., Falco, V., Tasioula-Margari, M., Logrieco, A.F., Mita, G., Bleve, G., 2015. New process for production of fermented black table olives using selected autochthonous microbial resources. *Front. Microbiol.* **6**:1007–1022. doi: 10.3389/fmicb.2015.01007.

599 Versalovic, J., Schneider, M., De Bruijn, F.J., Lupski, J.R., 1994. Genomic fingerprinting of
 600 bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol Cell Biol*
 601 **5(1)**, 25-40.

602 Williams, J.G., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S.V., 1990. DNA
 603 polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids*
 604 *Res.* **18(22)**, 6531-6535.

605 Yu, J., Sun, Z., Liu, W., Bao, Q., Zhang, J., Zhang, H., 2012. Phylogenetic study of
 606 *Lactobacillus acidophilus* group, *L. casei* group and *L. plantarum* group based on partial
 607 *hsp60*, *pheS* and *tuf* gene sequences. *Eur Food Res Technol* **234(6)**, 927-34.

608 Zapparoli, G., Reguant, C., Bordons, A., Torriani, S., Dellaglio, F., 2000. Genomic DNA
 609 fingerprinting of *Oenococcus oeni* strains by pulsed-field gel electrophoresis and randomly
 610 amplified polymorphic DNA-PCR. *Curr Microbiol* **40(6)**, 351-355.

611 Zanoni, P., Farrow, J.A.E., Phillips, B.A., Collins, M.D., 1987. *Lactobacillus pentosus* (Fred,
 612 Peterson, and Anderson) sp. nov., nom. Rev. *Int. J. Syst. Bacteriol.* **37(4)**, 339-34.

Table 1: Primer concentrations and PCR conditions used for genotyping *Lactobacillus plantarum* and *Lactobacillus pentosus* strains using housekeeping genes.

Target Genes	Primers Concentrations (μ M)	PCR Conditions
<i>16S rRNA</i> (Amann and others, 1995)	0.4	94 °C for 3 min, 25 cycles of 94 °C for 1 min, 57 °C for 2 min and 72 °C for 2 min and a final extension at 72 °C for 5 min
Housekeeping Genes of Interest		
<i>recA</i> (Torriani et al., 2001)	0.25	94 °C for 3 min, 30 cycles of 94 °C for 30 s, 56 °C for 10 s and 72 °C for 30 s and a final extension at 72 °C for 5 min
<i>dnaK</i> (Huang et al., 2010)	0.25	94 °C for 5 min, 35 cycles of 94 °C for 1 min, 58 °C for 1 min and 72 °C for 1.5 min and a final extension at 72 °C for 7 min
<i>pheS</i> (Naser et al., 2007)	0.5	95 °C for 5 min, 30 cycles of 95 °C for 35 s, 60 °C for 1.15 min and 72 °C for 1.15 min and a final extension at 72 °C for 7 min
<i>rpoA</i> (Naser et al., 2007)	0.5	95 °C for 5 min, 30 cycles of 95 °C for 35 s, 56 °C for 1.15 min and 72 °C for 1.15 min and a final extension at 72 °C for 7 min

Table 2. Description of the LAB used as reference in this study.

Genus	Species	ID Number(s)	Sources	References
<i>Lactobacillus</i>	<i>plantarum</i>	LA1196; ATCC BAA-793; NCIMB 8826; WCFS1	Saliva	Hols et al., 1997
<i>Lactobacillus</i>	<i>plantarum</i>	LA0070; ATCC 14917T	Pickle Cabbage	ATCC < PA Hansen < Roy Techn. Coll., Copenhagen< S. Orla-Jensen (<i>Streptobacterium plantarum</i>) https://www.atcc.org/products/all/14917.aspx#history
<i>Lactobacillus</i>	<i>pentosus</i>	LA0233; ATCC 8041T	Sauerkraut	Fred et al., 1921
<i>Lactobacillus</i>	<i>plantarum</i>	LA0445; BI0007, MOP3	Commercial cucumber fermentation	Fleming et al., 1988

Table 3: Accession numbers for the 16S *rRNA*, *recA*, *dnaK*, *pheS*, and *rpoA* sequences obtained from seven lactobacilli clones and the reference cultures used in this study.

Isolate ID	Accession Numbers for the Partial Gene Sequences Obtained				
	<i>16S rRNA</i>	<i>recA</i>	<i>dnaK</i>	<i>pheS</i>	<i>rpoA</i>
1.2.11	KT025927	KT027550	KT124106	KT124117	KT124128
1.2.13	KT025928	KT027551	KT124107	KT124118	KT124129
1.8.6	KT025929	KT027552	KT124108	KT124119	KT124130
1.8.9	KT025930	KT027553	KT124109	KT124120	KT124131
3.8.24	KT025931	KT027554	KT124110	KT124121	KT124132
3.2.37	KT025932	KT027555	KT124111	KT124122	KT124133
LA0445 (Reference Strain)	KT025933	KT027556	KT124112	KT124123	KT124134
<i>L. pentosus</i> ATCC 8041	KT025934	KT027557	KT124113	KT124125	KT124135
3.2.8	KT025935	KT027558	KT124114	KT124126	KT124136
<i>L. plantarum</i> WCFS1	KT025936	KT027560	KT124115	KT124124	KT124137
<i>L. plantarum</i> ATCC 14917	KT025937	KT027559	KT124116	KT124127	KT124138

Table 4: Accession numbers for the MLST amplicon-sequences obtained from seven lactobacilli clones and reference strains.

Isolates	Accession Numbers for the Partial Gene Sequences Obtained					
	<i>mutS</i>	<i>purk1</i>	<i>ddl</i>	<i>gdh</i>	<i>gyrB</i>	<i>pgm</i>
1.2.11	KT185461	KT185452	KT185479	KT185471		
1.2.13	KT185462	KT185453	KT185480	KT185472		
1.8.6	KT185463	KT185454	KT185481	KT185473		
1.8.9	KT185464	KT185455	KT185482	KT185474	NA*	NA
3.8.24	KT185465	KT185456	KT185483	KT185475		
3.2.37	KT185466	KT185457	KT185484	KT185476		
LA0445 (Reference Strain)	KT185467	KT185458	KT185485	KT185477		
3.2.8	KT185468	KT185459	KT185486	KT185478	KT185488	KT185490
<i>L. plantarum</i> WCFS1	KT185469	KT185460	KT185487	KT185479	KT185489	KT185491

* No Amplicon

Table 5: Identification of lactobacilli clones based on the partial sequence for the 16S *rRNA* and the MLST amplicon-sequences.

Isolated Culture ID	Best Match / Sequence Similarity (%)						
	16S <i>rRNA</i>	<i>mutS</i>	<i>purk1</i>	<i>ddl</i>	<i>gdh</i>	<i>gyrB</i>	<i>pgm</i>
1.2.11	<i>L. plantarum</i> / 100	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 98	<i>L. pentosus</i> / 98	NA*	NA
1.2.13	<i>L. plantarum</i> / 100	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 98	<i>L. pentosus</i> / 94	NA	NA
1.8.6	<i>L. plantarum</i> / 100	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 98	<i>L. pentosus</i> / 98	NA	NA
1.8.9	<i>L. plantarum</i> / 100	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 93	<i>L. pentosus</i> / 98	NA	NA
3.8.24	<i>L. plantarum</i> / 100	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 98	NA	NA
3.2.37	<i>L. plantarum</i> / 100	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 98	<i>L. pentosus</i> / 98	NA	NA
LA0445	<i>L. plantarum</i> / 100	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 99	<i>L. pentosus</i> / 86	NA	NA
3.2.8	<i>L. plantarum</i> / 100	<i>L. plantarum</i> / 99	<i>L. plantarum</i> / 99	<i>L. plantarum</i> / 98	<i>L. plantarum</i> / 99	<i>L. plantarum</i> / 99	<i>L. plantarum</i> / 100
<i>L. plantarum</i> WCFS1	<i>L. plantarum</i> / 100	<i>L. plantarum</i> / 100	<i>L. plantarum</i> / 99	<i>L. plantarum</i> / 100	<i>L. plantarum</i> / 100	<i>L. plantarum</i> / 100	<i>L. plantarum</i> / 100

*No Amplicon

Figure 1: Dendrograms and similarity coefficients for the amplicons obtained for *recA*, *pheS*, *rpoA* and *dnaK* from the subgroup of seven isolates were constructed using the Maximum Likelihood method based on the Tamura-Nei model with a bootstrap value of 1000. All positions containing gaps and missing data were eliminated. *L. pentosus* clusters are identified by letters.

Figure 2: Phylogenetic dendrogram and RAPD-COC band pattern for the subgroup of seven isolates. Percent similarity were calculated using the Pearson correlation coefficient. The clustering was performed using the Unweighted Pair Group Method with the Arithmetic Mean (UPGMA; Bionumerics 6.6, Applied Maths, Sint-Martens-Latem, Belgium).

Figure 3: Phylogenetic dendrogram and rep-PCR-(GTG)₅ band pattern for the subgroup of seven isolates. Percent similarity were calculated using the Pearson correlation coefficient. The clustering was performed using the Unweighted Pair Group Method with the Arithmetic Mean (UPGMA; Bionumerics 6.6, Applied Maths, Sint-Martens-Latem, Belgium).

Figure 4: Rep-PCR-GTG₅ phylogram for *L. plantarum*. UPGMA dendrogram obtained for the *L. plantarum* clones included in this study using the DICE coefficient. *L. plantarum* clusters are identified by the greek letters: π , τ , ψ . The color legend to the right describes the phenotypes for each clone in the genotypic clusters and follows the same color assignment shown in Figure 7.

Figure 5: Rep-PCR-GTG₅ dendrogram for *L. pentosus*. UPGMA dendrogram using DICE coefficient. *L. pentosus* clusters are showed by the greek letters: α , β , γ , δ , ϵ , η , θ , λ , μ , ρ . The color legend to the right describes the phenotypes for each clone in the genotypic clusters and follows the same color assignment shown in Figure 7.

Figure 6: Venn diagram of the number of clones significantly affected by one or more fermentation variable (n=243).

Figure 7: Hierarchial Cluster Analysis (HCA) of *L. plantarum* and *L. pentosus* clones isolated from commercial cucumber fermentations. Two-way clustering of clones based on 48 h fermentation pH under varying conditions of salt, initial pH, and temperature.

T1 (0% NaCl, pH 5.4, 15°C)	T2 (6% NaCl, pH 4.0, 15°C)
4.02	3.4
4.1752	3.4939
4.3305	3.5878
4.4857	3.6817
4.6409	3.7756
4.7962	3.8695
4.9199	3.9206
5.0437	3.9717
5.1675	4.0228
5.2912	4.0739
5.415	4.125
T4 (0% NaCl, pH 4.0, 30°C)	T3 (6% NaCl, pH 5.4, 30°C)
3.03	3.21
3.1318	3.2851
3.2336	3.3603
3.3354	3.4354
3.4372	3.5106
3.539	3.5857
3.6392	3.6606
3.7394	3.7354
3.8396	3.8103
3.9398	3.8851
4.04	3.96