MICROBIOLOGY ARCHIVES
Original Research Article | ISSN: 3041-590X
Journal homepage: https://microjournal.researchfloor.org/

Molecular identification and fermentation-associated functional gene profiling of bacterial and yeast isolates from traditional and modified tepache: a genotypic basis for starter culture selection

Nnodim, L. C.*, Oriji, E. N., Augustus, B. W., Woko, G. C., and Martins, S. C.

Department of Microbiology, Faculty of Science, Rivers State University, Nkpolu-Oroworukwo, Port Harcourt, Rivers State, Nigeria

ARTICLE INFO

Citation: Nnodim, L. C.; Oriji, E. N.; Augustus, B. W.; Woko, G. C.; and Martins, S. C. (2026). Molecular identification and fermentation-associated functional gene profiling of bacterial and yeast isolates from traditional and modified tepache: a genotypic basis for starter culture selection. *Microbiology Archives, an International Journal.*
DOI: https://doi.org/10.51470/MA.2026.8.2.30

Received 14 April 2026
Revised 17 May 2026
Accepted 19 June 2026
Available Online July 15, 2026

Corresponding Author: **Nnodim, L. C.**
E-Mail: **lasbry.nnodim@rsu.edu.ng**

ABSTRACT

Tepache is a spontaneously fermented, pineapple-based beverage whose microbial consortium remains poorly characterised at the level of individual isolates. Spontaneous fermentation remains the norm for comparable beverages across sub-Saharan Africa, and the resulting batch-to-batch variability obstructs standardisation. This study identified bacterial and yeast isolates from traditional and modified tepache formulations, the latter containing Nigerian botanicals, and screened them for the genes carrying out core fermentative reactions. Genomic DNA was extracted from four isolates, and the 16S rRNA gene and internal transcribed spacer (ITS) region were amplified and sequenced. Polymerase chain reaction screening targeted the yeast genes ADH1, PDC and SUC2 and the bacterial lactate dehydrogenase (D-LDH and L-LDH) and scr sucrose-utilisation loci. Sequence analysis identified both yeast isolates as Saccharomyces cerevisiae, matching strains ZP 541 (98.40%) and NCIM3107 (97.60%). The bacterial isolate from the modified beverage matched Bacillus subtilis CICC10148 at 98.70%, whereas that from the traditional beverage matched Lacticaseibacillus sp. at 95.10%, an identity supporting assignment at the genus level only. Both yeast isolates carried ADH1 (179 bp), PDC (195 bp) and SUC2 (approximately 1320 bp), which define a complete genotypic route from sucrose hydrolysis to ethanol. Both bacterial isolates yielded D-LDH and scr genes as multiple fragments, whereas L-LDH amplified only in the traditional-beverage isolate. The isolates therefore carry a gene complement matched to the sucrose-rich substrate, providing a genotypic rationale for their evaluation as defined starter cultures.

Keywords: *tepache; functional gene screening; 16S rRNA; internal transcribed spacer; starter culture; Saccharomyces cerevisiae; Lacticaseibacillus sp.; Bacillus subtilis.*

1. Introduction

Traditional fermented beverages hold a central place in the food systems of much of the developing world, where they supply calories, micronutrients and organoleptic variety and, in many settings, a measure of microbiological safety derived from acidification and ethanol production. Producers make most of these products by spontaneous fermentation, in which the raw substrate, the processing utensils and the immediate environment recruit the microbial consortium rather than the producer introducing it deliberately. A review of approximately ninety studies covering nine Nigerian traditional and artisanal beverages recorded *Bacillus*, *Escherichia*, *Lactobacillus*, *Staphylococcus* and *Streptococcus* in every product examined, together with the fungal genera *Saccharomyces*, *Aspergillus*, *Candida* and *Penicillium*, and reported a higher prevalence of hygiene indicator organisms in the non-alcoholic beverages, which the authors attributed to incorrect practice during production [14].

That review calls for regulation and for molecular characterisation of the organisms present throughout production and storage. Neither measure, however, addresses the source of the variability itself. A defined starter culture does, and its absence remains the principal technological obstacle to industrialising beverages of this kind.

Producers in Mexico traditionally make tepache, a lightly alcoholic beverage, by fermenting pineapple peel and pulp with unrefined cane sugar. High-throughput sequencing has resolved its microbial ecology in detail. Metabarcoding of the 16S rRNA gene and the ITS region across a 72 h fermentation revealed a two-phase succession in which an initial lactic and alcoholic phase driven by lactic acid bacteria and yeasts gives way to a rise in acetogenic bacteria; ethanol rose from 0.83 to 3.39 g/L and L-lactic acid from 1.42 to 8.77 g/L as total sugars fell from 123.43 to 84.70 g/L [9].

At the close of that fermentation, *Lactobacillus*, *Leuconostoc*, *Acetobacter* and *Lactococcus* dominated the bacterial fraction, and *Saccharomyces*, *Zygosaccharomyces*, *Candida* and *Meyerozyma*, among other genera, dominated the fungal fraction. Community composition is not invariant, however. Metagenomic analysis of homemade tepache instead recovered a community dominated by *Bacillus*, *Meyerozyma* and *Talaromyces*, and detected the tetracycline resistance gene *tetM* in community DNA, which shows that the wild fermentation is neither compositionally stable nor free of risk [2]. The tibicos grains used in tepache production have also yielded lactic acid bacteria with demonstrable probiotic properties, among them a *Lacticaseibacillus paracasei* strain that survived simulated gastrointestinal passage [15].

The beverage has recently attracted interest as a template for adaptation in West Africa, where pineapple is abundant and where processors discard a substantial fraction of the harvest, particularly the peel. A modified formulation that combines pineapple and its peel with *Hibiscus sabdariffa*, *Xylopia aethiopica*, *Piper guineense* and *Monodora myristica* extends the tepache concept into a distinctly Nigerian sensory and phytochemical space. No previous work has examined whether such reformulation alters the fermentative flora, or whether the resulting isolates retain the genetic machinery that efficient fermentation of a sucrose-rich substrate requires.

Culture-independent surveys answer the question of which organisms are present, but not the question of what a given isolate is equipped to do, and strain selection for starter culture development requires the second kind of information. Indigenous *Saccharomyces cerevisiae* strains recovered from spontaneous food fermentations differ markedly in stress tolerance, substrate utilisation and volatile production even when they share a source, so performance is strain-specific rather than species-specific [6]. Controlled fermentations with defined yeast and lactic acid bacteria strains correspondingly improve the safety, physicochemical stability and sensory quality of African beverages relative to spontaneous fermentation [16]. Screening candidate isolates for the genes that carry out the reactions of interest offers a rapid, low-cost first filter in this selection process, particularly in laboratories where genome sequencing is not routinely available.

Three genes define the fermentative capability of yeast on a sucrose-rich substrate. *SUC2* encodes invertase, which hydrolyses sucrose to glucose and fructose and so makes the disaccharide accessible to glycolysis; genes encoding invertase occur at multiple telomeric loci in baker's and distiller's strains, an arrangement that represents an adaptation to sucrose-rich broths [13]. *PDC* encodes pyruvate decarboxylase, which commits pyruvate to acetaldehyde, and *ADH1* encodes the alcohol dehydrogenase that reduces acetaldehyde to ethanol. Chromosomal co-overexpression of *S. cerevisiae PDC1* and *ADH1* alongside the endogenous *GAPDH* gene installed ethanol biosynthesis at 70.8 mg/L in *Yarrowia lipolytica*, a species that does not natively accumulate ethanol, which confirms that these loci constitute the core ethanologenic module [18]. In lactic acid bacteria, the corresponding determinants are the lactate dehydrogenase genes: disruption of *ldh* in *Lacticaseibacillus paracasei* lowers growth rate and acidification capacity and alters the formation of flavour-associated metabolites [7]. Sucrose utilisation in several bacterial lineages proceeds instead through phosphoenolpyruvate-dependent phosphotransferase uptake followed by intracellular hydrolysis, a route that the four-gene *scr* cluster encodes.

Detailed characterisation of that cluster comes from the marine bacterium *Photobacterium damselae* subsp. *damselae*, in which it comprises *scrA*, encoding a phosphotransferase system sucrose-specific IIBC component, *scrK*, encoding a fructokinase, *scrB*, encoding a sucrose-6-phosphate hydrolase, and the repressor *scrR*; deletion of *scrA* alone abolishes sucrose fermentation, and clusters of comparable architecture have entered different lineages independently by horizontal transfer [1].

The present study therefore pursued two objectives: to establish the identity of the dominant bacterial and yeast isolates from traditional and modified tepache by 16S rRNA and ITS sequencing, and to determine whether those isolates carry the functional genes that sucrose utilisation, ethanol production and lactic acid production require. The intention was to move beyond a descriptive inventory of the fermenting flora towards a genotypically justified shortlist of candidate starter organisms.

2. Materials and Methods

2.1. Source and designation of isolates

Isolates were recovered from two spontaneously fermented tepache formulations: a traditional formulation and a modified formulation containing pineapple with peels together with Nigerian botanical adjuncts. Bacterial isolates were obtained on de Man, Rogosa and Sharpe (MRS) agar and yeast isolates on potato dextrose agar (PDA). Four isolates were carried forward for molecular characterisation and were designated according to formulation and cell type: M1 (bacterium, modified), T9 (bacterium, traditional), MY (yeast, modified) and TY (yeast, traditional).

2.2. Genomic DNA extraction

Total genomic DNA was extracted with the ZR Fungal/Bacterial DNA MiniPrep kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. Briefly, 2 mL of culture broth was transferred to a ZR BashingBead Lysis Tube containing 750 µL of lysis solution and processed in a bead beater at maximum speed for at least 5 min. The lysate was centrifuged at 10,000 × g or above for 1 min, and up to 400 µL of the supernatant was passed through a Zymo-Spin IV spin filter at 7,000 × g for 1 min. The filtrate was mixed with 1,200 µL of Fungal/Bacterial DNA Binding Buffer, and 800 µL aliquots of this mixture were loaded onto a Zymo-Spin IIC column and centrifuged at 10,000 × g for 1 min, the step being repeated to process the full volume. The column was washed with 200 µL of DNA Pre-Wash Buffer and then with 500 µL of Fungal/Bacterial DNA Wash Buffer, each wash followed by centrifugation at 10,000 × g for 1 min. DNA was eluted in 100 µL of DNA Elution Buffer by centrifugation at 10,000 × g for 30 s and stored at −20 °C until use.

2.3. Amplification of identification markers

All reactions were carried out in a 25 µL volume containing 12.5 µL of Taq 2X Master Mix (New England Biolabs, M0270), 1 µL each of 10 µM forward and reverse primers, 2 µL of template DNA and 8.5 µL of nuclease-free water.

The bacterial 16S rRNA gene was amplified with primers 27F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1525R (5′-AAGGAGGTGWTCCARCCGC-3′). Cycling comprised initial denaturation at 94 °C for 5 min; 36 cycles of 94 °C for 30 s, 56 °C for 30 s and 72 °C for 45 s; and a final extension at 72 °C for 7 min. The fungal internal transcribed spacer region (ITS1–5.8S–ITS2) was amplified with primers ITS1 (5′-TCCGTAGGTGAACCTGCGG-3′) and ITS4 (5′-TCCTCCGCTTATTGATATGC-3′) under identical conditions, except that the annealing temperature was 54 °C.

2.4. Screening for fermentation-associated functional genes

Yeast isolates were screened for three target genes. *SUC2* was amplified with 5′-GCGATAGACCTTTGGTCCAC-3′ and 5′-GGACCGTGGTAACTCTAAGG-3′; *PDC* with 5′-AGCTAACGCTGCTGTCCCAG-3′ and 5′-GTGGTGAAACCAATGGAACC-3′; and *ADH1* with 5′-CGGTGCTGTTCTAAAGGCC-3′ and 5′-TGGACTTGACGACTTGGTTG-3′.

Bacterial isolates were screened for three targets. The lactate dehydrogenase gene was amplified with 5′-AAGAATTCCCAGACAAGCGTTTAGC-3′ and 5′-AGGAACGTTGGTAATTTCGAAGC-3′, a pair designated D-LDH, and separately with the pair LDH-f2 (5′-AAGACAAGCTTGATGAAATTCATAAGAGTG-3′) and LDH-r2 (5′-ACCCAATTCTTGGGCAATACCTTGG-3′), designated L-LDH. The *scr* sucrose-utilisation cluster was amplified with 5′-GGAGGATCCGTTCCCGAAAATTTCGCAAG-3′ and 5′-ATGAAGCTTGGAATGTATGCTCTCCAACTG-3′. All functional gene reactions used the mix described in Section 2.3, with initial denaturation at 94 °C for 5 min; 36 cycles of 94 °C for 30 s, 53 °C for 30 s and 72 °C for 45 s; and a final extension at 72 °C for 7 min. Throughout this report, the designations *ADH1*, *PDC*, *SUC2*, *D-LDH*, *L-LDH* and *scr* denote the intended primer targets rather than sequence-verified loci.

2.5. Agarose gel electrophoresis

Genomic DNA was resolved on 1% (w/v) agarose and PCR products on 2% (w/v) agarose, each prepared in 1× TAE buffer and stained with 10 µL of EZ-Vision DNA dye. Samples were mixed with loading buffer and loaded alongside a molecular weight ladder: 1 kb for the 16S rRNA amplicons and 50 bp for the ITS and functional gene amplicons. Electrophoresis was carried out at 80 – 150 V for 1 – 1.5 h in 1× TAE, and bands were visualised on a UV transilluminator.

2.6. Sequencing and sequence analysis

Amplicons of the 16S rRNA gene and the ITS region were sequenced on an Applied Biosystems Genetic Analyser 3500 with the BigDye Terminator v3.1 Cycle Sequencing Kit according to the manufacturer's protocol. Chromatograms were inspected, trimmed and assembled in Geneious Prime (Biomatters, Auckland, New Zealand). Consensus sequences were compared with the GenBank non-redundant nucleotide database using BLASTn, and the closest match, pairwise identity, bit score, E-value and aligned coordinates were recorded for each isolate. Identities were interpreted against published rank boundaries: 97.2 – 100% for species and 90.1 – 99.0% for genus in the case of the 16S rRNA gene [10], and 98.41% for species and 96.31% for genus in the case of the yeast ITS region [17]. Functional gene amplicons were scored as present or absent based on a band of the expected size and were not sequenced.

3. Results

3.1. Amplification of identification markers

All four isolates yielded genomic DNA of sufficient quality for amplification. Amplification of the bacterial 16S rRNA gene generated a single, intense product of approximately 1500 bp in both bacterial isolates (Figure 1), and amplification of the fungal ITS region generated a single product of approximately 650 bp in both yeast isolates (Figure 2).

In each case the two lanes migrated to the same position, which indicates that the isolates from the traditional and the modified formulations carried markers of indistinguishable length. All amplicon sizes reported here derive from comparison with the molecular weight ladder and are therefore approximate.

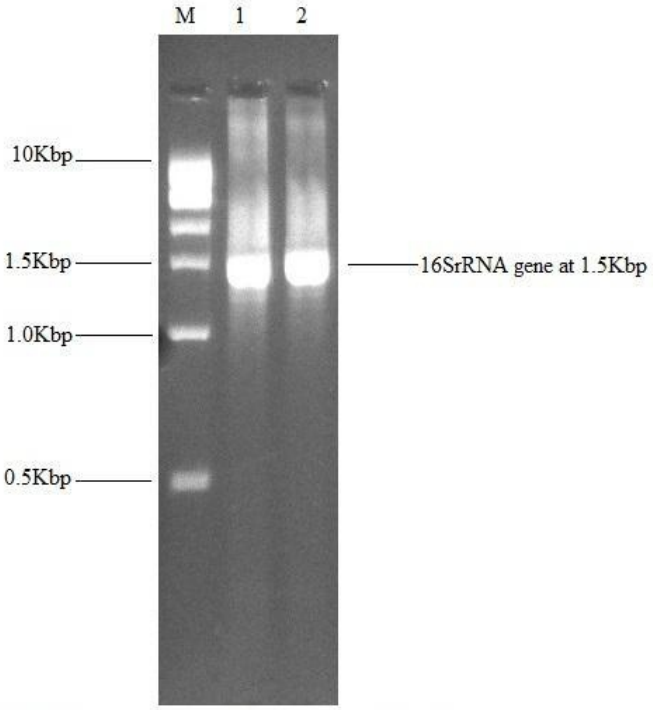

Figure 1. Agarose gel electrophoresis of 16S rRNA gene amplicons (approximately 1500 bp) of bacterial isolates from tepache. Lane M, 1 kb DNA ladder; lane 1, isolate M1 (modified tepache); lane 2, isolate T9 (traditional tepache).

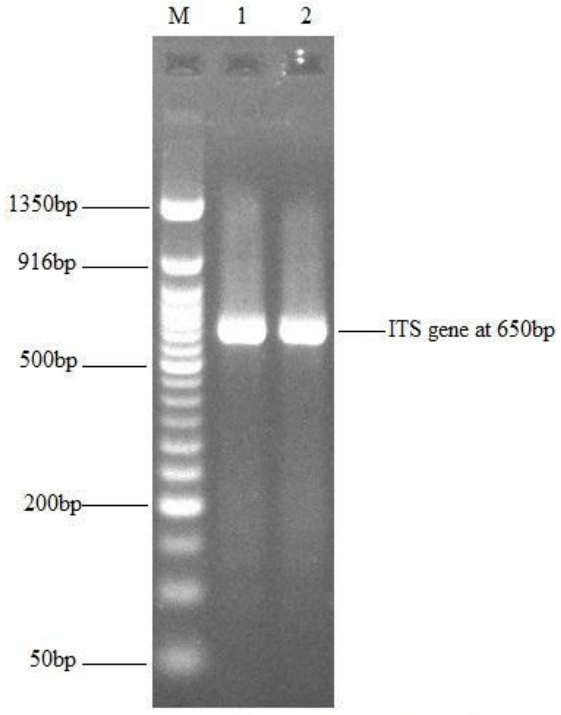

Figure 2. Agarose gel electrophoresis of internal transcribed spacer (ITS) region amplicons (approximately 650 bp) of yeast isolates from tepache. Lane M, 50 bp DNA ladder; lane 1, isolate TY (traditional tepache); lane 2, isolate MY (modified tepache).

3.2. Sequence-based identification of isolates

BLASTn comparison of the assembled consensus sequences with GenBank returned the matches summarised in Table 1. The yeast isolates from both formulations matched *Saccharomyces cerevisiae*, isolate TY matching strain ZP 541 at 98.40% pairwise identity over an 807 bp alignment and isolate MY matching strain NCIM3107 at 97.60% over a 749 bp alignment. The bacterial isolate from the modified beverage, M1, returned *Bacillus subtilis* strain CICC10148 as its closest match at 98.70% identity over a 571 bp alignment, and the bacterial isolate from the traditional beverage, T9, returned *Lacticaseibacillus* sp. at 95.10% identity over an 892 bp alignment. Every match returned an E-value of 0. Section 4.1 evaluates these identities against the rank boundaries set out in Section 2.6.

Table 1: Molecular identification of bacterial and yeast isolates from traditional and modified tepache based on 16S rRNA and ITS sequence homology in the GenBank database

Isolate code	Source	Marker	Identity (%)	Closest GenBank match (accession)	Bit score	E-value	Aligned region
M1	Modified tepache	16S rRNA	98.70	*Bacillus subtilis* strain CICC10148 (AY971358)	813.6	0	102–672
T9	Traditional tepache	16S rRNA	95.10	*Lacticaseibacillus* sp. (CP001084)	1343.6	0	276442–277333
MY	Modified tepache	ITS	97.60	*Saccharomyces cerevisiae* strain NCIM3107 (CP009950)	1260.5	0	452900–453648
TY	Traditional tepache	ITS	98.40	*Saccharomyces cerevisiae* strain ZP 541 (EU145764)	1415.7	0	1–807

Note. *Genus nomenclature for isolate T9 follows the reclassification of Lactobacillus by Zheng et al. [19]; the GenBank record CP001084 is deposited under the earlier name Lactobacillus casei Zhang. The 95.10% identity returned for this isolate supports' assignment at genus level only (see Section 4.1).*

3.3. Fermentation-associated functional genes in the yeast isolates

Both yeast isolates yielded amplicons for all three yeast target genes. Amplification of *ADH1* generated a single sharp band of 179 bp in both isolates (Figure 3), and amplification of *PDC* generated a single band of 195 bp, more intense in isolate MY than in isolate TY (Figure 4). Amplification of *SUC2* generated a high molecular weight band of approximately 1320 bp in both isolates (Figure 5). The three amplicons were of identical size in the two isolates, as Figure 6 summarises.

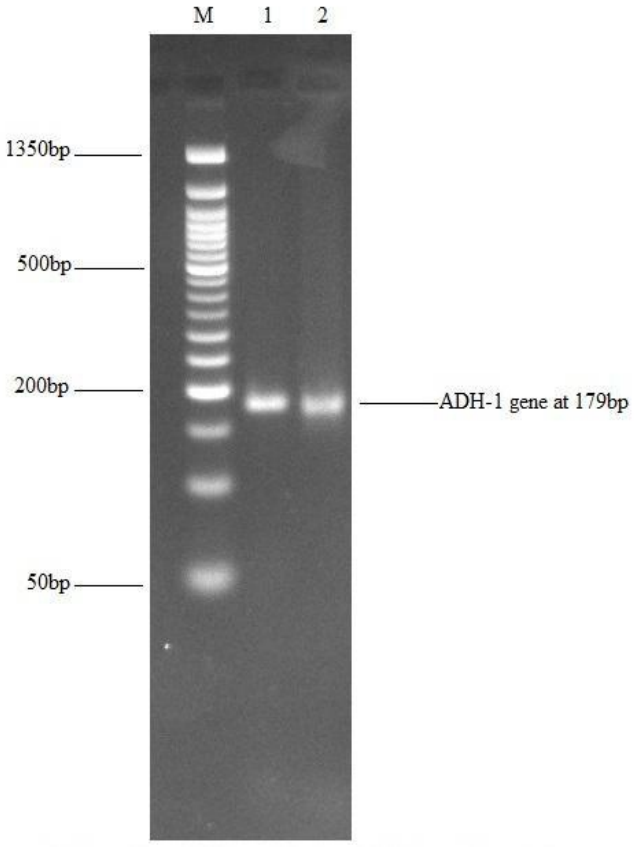

Figure 3. Amplification of the alcohol dehydrogenase (ADH1) gene (179 bp) in yeast isolates from tepache. Lane M, 50 bp DNA ladder; lane 1, isolate TY; lane 2, isolate MY.

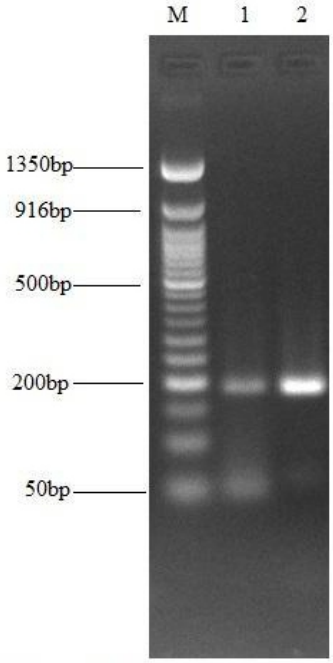

Figure 4. Amplification of the pyruvate decarboxylase (PDC) gene (195 bp) in yeast isolates from tepache. Lane M, 50 bp DNA ladder; lane 1, isolate TY; lane 2, isolate MY.

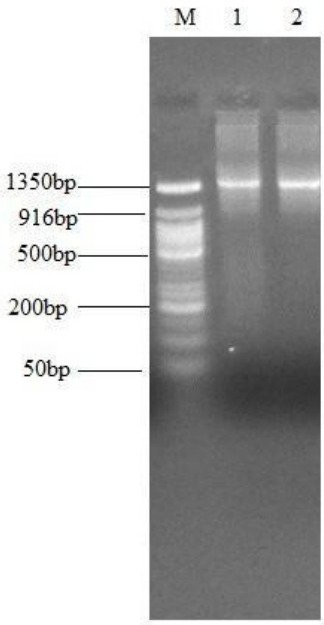

Figure 5. Amplification of the invertase (SUC2) gene (approximately 1320 bp) in yeast isolates from tepache. Lane M, 50 bp DNA ladder; lane 1, isolate TY; lane 2, isolate MY.

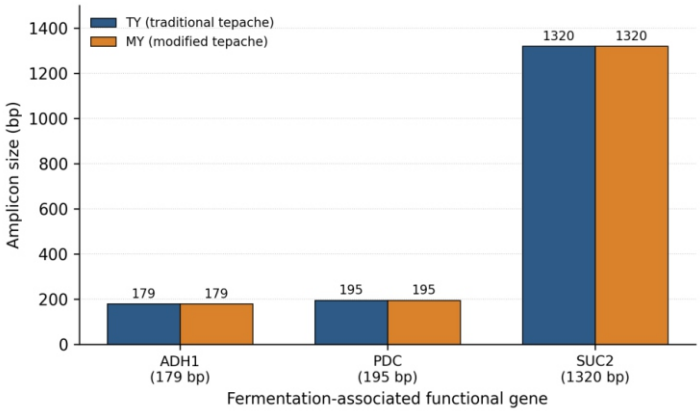

Figure 6. Comparative amplicon sizes of the three fermentation-associated functional genes detected in the two Saccharomyces cerevisiae isolates. Sizes are those estimated from the gels shown in Figures 3 to 5.

3.4. Fermentation-associated functional genes in the bacterial isolates

The bacterial functional gene screens produced multiband rather than single-band profiles. Amplification with the *D-LDH* primer pair generated multiple fragments between approximately 500 and 1200 bp in both bacterial isolates (Figure 7). Amplification with the *L-LDH* primer pair generated a faint multiband profile between approximately 350 and 700 bp in isolate T9 but produced no detectable product in isolate M1 (Figure 8). Amplification of the *scr* cluster generated a reproducible profile of multiple fragments between approximately 300 and 700 bp in both isolates, with two prominent bands at approximately 400 and 550 bp (Figure 9). Table 2 summarises the complete pattern of gene detection across all four isolates.

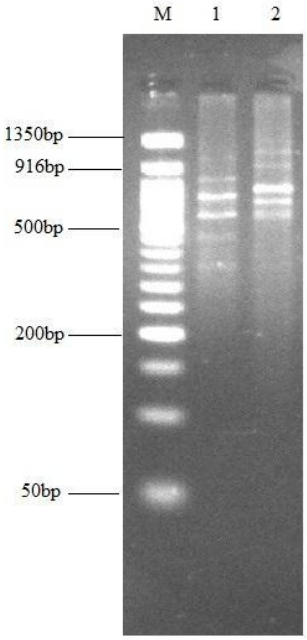

Figure 7. Amplification profile of the lactate dehydrogenase (D-LDH) gene in bacterial isolates from tepache, showing multiple fragments between approximately 500 and 1200 bp. Lane M, 50 bp DNA ladder; lane 1, isolate M1; lane 2, isolate T9.

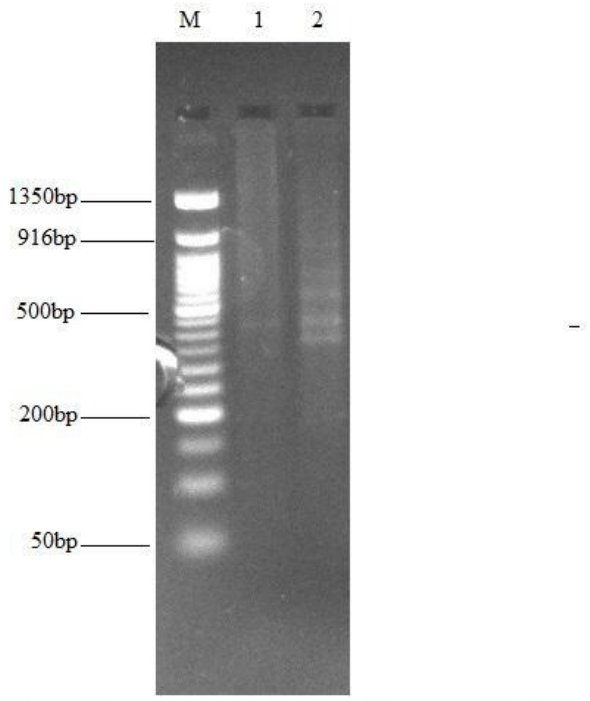

Figure 8. Amplification profile of the lactate dehydrogenase (L-LDH) gene in bacterial isolates from tepache, showing multiple fragments between approximately 350 and 700 bp in isolate T9 only. Lane M, 50 bp DNA ladder; lane 1, isolate M1; lane 2, isolate T9.

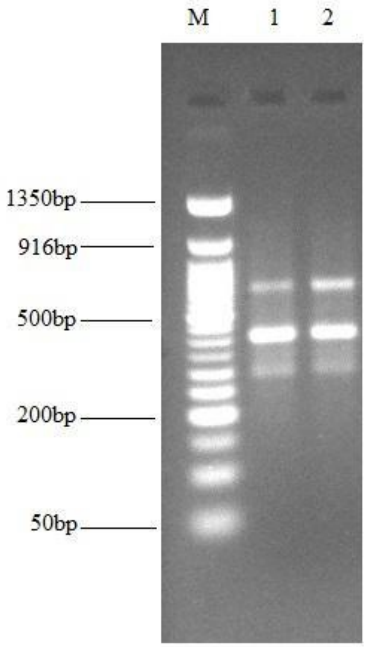

Figure 9. Amplification profile of the scr sucrose-utilisation gene cluster in bacterial isolates from tepache, showing multiple fragments between approximately 300 and 700 bp. Lane M, 50 bp DNA ladder; lane 1, isolate M1; lane 2, isolate T9.

Table 2: Detection of fermentation-associated functional genes in bacterial and yeast isolates from traditional and modified tepache

Isolate	Source	Identification	*ADH1*	*PDC*	*SUC2*	*D-LDH*	*L-LDH*	*scr*
TY	Traditional	*Saccharomyces cerevisiae*	+	+	+	nt	nt	nt
MY	Modified	*Saccharomyces cerevisiae*	+	+	+	nt	nt	nt
T9	Traditional	*Lacticaseibacillus* sp.	nt	nt	nt	+	+	+
M1	Modified	*Bacillus subtilis*	nt	nt	nt	+	–	+

Note. *+, amplicon of the expected size detected; –, no detectable amplicon; nt, not tested, the target being inapplicable to that cell type.*

4. Discussion

4.1. Confidence limits of the sequence-based identifications

The four identifications in Table 1 do not carry equal weight, and stating this explicitly serves the reader better than presenting them as equivalent. An analysis of 19,556 type strains, encompassing 94% of all validly published prokaryotic species and involving more than 191 million pairwise alignments, established that in 90% of cases sequences from the same species share a minimum of 97.2–100% 16S rRNA gene identity, whereas the corresponding range for the same genus is 90.1–99.0% [10]. The most consequential finding of that work is that these boundaries overlap, so 16S identity alone cannot resolve taxonomic rank near the margin.

Applied to the present data, isolate M1 at 98.70% falls within the species range, and its assignment to *Bacillus subtilis* is therefore defensible. The same value nevertheless falls within the upper part of the genus range, and the alignment spanned only 571 bp of an approximately 1500 bp amplicon, so a longer read would strengthen the call. Isolate T9 at 95.10% falls squarely within the genus range and outside the species range. The correct conclusion is that T9 belongs to the genus *Lacticaseibacillus*, and that the species-level attribution to *L. paracasei* remains provisional pending corroboration. Average nucleotide identity or multilocus sequence analysis would settle the point, although inference from average nucleotide identity carries limitations of its own for divergent taxa [8]. The genus name applied here follows the reclassification of *Lactobacillus* into 25 genera on whole-genome phylogenetic evidence, under which *L. paracasei* moved to *Lacticaseibacillus* [19]; the older binomial persists widely in the fermented foods literature and causes avoidable confusion.

The two yeast identifications require the same scrutiny, which the original framing of this study did not apply to them. Barcoding of more than 9,000 yeast isolates places the ITS species boundary at 98.41% and the genus boundary at 96.31% [17]. Isolate TY at 98.40% therefore sits marginally below the species threshold, and isolate MY at 97.60% sits clearly below it, while both sit well above the genus boundary. Assignment of both isolates to the genus *Saccharomyces* is secure. The species attribution to *S. cerevisiae*, although consistent with the closest matches returned and with the ecology of the substrate, would become secure only through sequencing of the D1/D2 domain of the large subunit rRNA gene or through a multilocus approach. We report the species names here as the closest database matches rather than as settled identifications.

4.2. *Bacillus subtilis* in the modified formulation

The recovery of *Bacillus subtilis* as the dominant bacterial isolate from the modified beverage warrants careful interpretation, because metabarcoding surveys of tepache do not rank the genus among the dominant taxa [9]. Metagenomic analysis of homemade tepache, by contrast, recovered a community in which *Bacillus* dominated the bacterial fraction [2]. The organism therefore belongs to the tepache microbial repertoire rather than representing a laboratory artefact, and its prominence here may reflect the botanical adjuncts, the inclusion of peel material, or a spore-forming lifestyle that permits survival of transient thermal and osmotic stress.

Whether its presence is desirable is a separate question. *Bacillus subtilis* drives the fermentation of a range of traditional foods and contributes amylolytic and proteolytic activity, but it also produces biogenic amines, and strains differ substantially in this respect: comparative metabolomic analysis of high- and low-amine-producing strains revealed markedly different volatile and non-volatile profiles, with 2-phenylethylamine detected only in the high-producing strain [11]. The trait is not uniformly negative. A *B. subtilis* strain isolated from gouda-type cheese reduced histamine, tyramine, phenylethylamine, putrescine and cadaverine by 65 – 85% in vitro under defined temperature and pH conditions [3], and *Bacillus* strains screened for low amine production and high degradation capacity have served successfully as starters for fermented soybean products [4]. Any proposal to retain isolate M1 as a starter component must therefore rest on strain-level assessment of amine production, haemolytic activity and enterotoxin gene carriage.

4.3. Lacticaseibacillus sp. in the traditional formulation

The recovery of a *Lacticaseibacillus* isolate from traditional tepache accords with the acidifying phase of the fermentation and with the wider literature on this beverage system. A *Lacticaseibacillus paracasei* strain isolated from the tibicos grains used in tepache production survived simulated gastric and intestinal digestion at rates of approximately 57% and 40%, respectively, exhibited high hydrophobicity, autoaggregation and mucin adhesion, lacked haemolytic activity, and produced antimicrobial and antioxidant activity in cell-free supernatant [15]. The genus therefore represents the most promising probiotic reservoir in this beverage. Safety and technological traits within *Lacticaseibacillus* are nonetheless strain-specific rather than species-specific: a survey of 121 isolates from dairy, sourdough, wine, must and human sources found that, although strains from the same matrix often shared genetic characteristics, phenotypic traits such as antibiotic resistance, biogenic amine production and tolerance of salt, ethanol and low pH varied at the level of the individual strain [5]. Isolate T9 accordingly requires phenotypic characterisation before it can support any probiotic claim.

4.4. The yeast functional gene complement

Both yeast isolates carried *ADH1*, *PDC* and *SUC2*, which is precisely the complement that the substrate in question demands. Tepache is a sucrose-rich medium, and sucrose is not directly fermentable: hydrolysis must come first. *SUC2* encodes the invertase that performs this hydrolysis, and genes encoding invertase occupy multiple telomeric loci in baker's and distiller's strains, an arrangement that represents an adaptation to sucrose-rich industrial broths [13]. Detection of a 1320 bp *SUC2* amplicon in both isolates confirms that the initial, rate-determining step is genetically available to them. The same study offers a useful corrective, however: strains carrying a single locus displayed adequate invertase activity, and the authors concluded that invertase activity does not necessarily limit sucrose fermentation. Presence of the locus is therefore a necessary but not a sufficient condition for efficient performance.

Downstream of hydrolysis, *PDC* and *ADH1* constitute the ethanologenic module. Heterologous expression demonstrates their sufficiency directly: chromosomal co-overexpression of *S. cerevisiae PDC1* and *ADH1* together with the endogenous *GAPDH* gene installed ethanol biosynthesis at 70.8 mg/L in *Yarrowia lipolytica*, a species that does not natively accumulate ethanol [18]. Detection of both loci in isolates TY and MY therefore establishes an unbroken genotypic route from sucrose to ethanol.

Matching transport and hydrolysis capacity to the substrate also pays measurable dividends in practice: introducing a heterologous sucrose transporter together with an invertase into *Clostridium beijerinckii* raised acetone–butanol–ethanol production from sucrose, sugarcane molasses and sugarcane juice by 38.7%, 22.3% and 52.8% respectively [12].

The two yeast isolates proved indistinguishable at every locus examined and produced amplicons of identical size (Figure 6). This equivalence of profile does not establish equivalence of function. Characterisation of 81 indigenous *S. cerevisiae* strains from spontaneous fermentations found that strains from the same source differed in phytase activity, tolerance of high glucose concentrations, salt tolerance and volatile profile, and that genotypic grouping by Rep-PCR and RAPD-PCR did not correlate fully with isolation source; the authors concluded that starter selection for fermented beverages must establish performance strain by strain [6]. Discriminating between TY and MY would require typing methods with strain-level resolution, such as interdelta or microsatellite analysis, rather than presence-absence screening of conserved metabolic loci.

4.5. Bacterial functional genes and the interpretation of multiband profiles

Both bacterial isolates amplified *D-LDH* and the *scr* cluster, and both did so as multiple fragments rather than as single discrete products. Two explanations merit consideration. The first, and in our assessment the more likely, is that the profiles reflect genuine multiplicity of targets. Lactic acid bacteria characteristically carry more than one lactate dehydrogenase gene, encoding L- and D-specific enzymes, and functional analysis in *Lactocaseibacillus* sp. confirms that these loci govern growth rate, acidification capacity and the formation of flavour-associated metabolites [7]. The *scr* locus is likewise not a single gene but a four-gene cluster comprising *scrA*, *scrK*, *scrB* and *scrR*, in which deletion of *scrA* alone abolishes sucrose fermentation, and which exists in more than one version as a consequence of independent horizontal transfer into different lineages [1]. Primers directed at such a cluster may reasonably generate several products.

The second explanation is non-specific amplification, and two features of the reactions support it. The annealing temperature of 53 °C used for the functional gene screens is permissive, and the *scr* primers carry BamHI and HindIII restriction sites at their 5′ ends, which indicates that they were designed for directional cloning rather than for diagnostic screening; the appended non-complementary sequence lowers effective specificity in the early cycles. The detailed characterisation of the *scr* cluster on which these primers rest, moreover, comes from *Photobacterium damselae* subsp. *damselae* [1], a Gram-negative marine bacterium only distantly related to the isolates screened here, so primers derived from that system would prime with limited specificity on a *Lacticaseibacillus* or *Bacillus* template.

We cannot discriminate between these possibilities on the present data. The defensible position is that the multiband profiles establish the presence of sequences homologous to the targets but establish neither the identity nor the number of the loci amplified. Excision and sequencing of individual bands, or gradient PCR to determine an annealing optimum, would resolve the question and should precede any further use of these primer sets.

The one clear differential result is the absence of an *ldhF2* product in isolate M1 against its presence in isolate T9. This outcome accords with the identifications, since the *ldhF2* primers target a lactic acid bacterium and would not prime efficiently on a *Bacillus* template. The result is therefore internally coherent and lends modest independent support to the distinction drawn between the two bacterial isolates.

4.6. Implications for starter culture development

Read together, the results provide a genotypic rationale for a defined two-component starter. The yeast isolates carry an intact route from sucrose hydrolysis through pyruvate decarboxylation to ethanol, and the bacterial isolates carry both lactate dehydrogenase and sucrose-utilisation sequences, a combination that matches the two-phase lactic and alcoholic succession characterising tepache fermentation [9]. Pairing a characterised yeast with a characterised lactic acid bacterium is a strategy of demonstrated benefit in African beverage fermentation, where defined pure and mixed cultures improved safety and physicochemical and sensory quality relative to spontaneous fermentation of palm sap [16]. Given the hygiene concerns documented for spontaneously fermented artisanal beverages in Nigeria [14] and the detection of a tetracycline resistance determinant in homemade tepache [2], the case for replacing spontaneous fermentation with a defined inoculum in any scaled production of this beverage is strong.

4.7. Limitations

Several limitations constrain the inferences that these data support. First, and most importantly, PCR detects the presence of a gene, not its expression or the activity of its product. None of the genotypic findings reported here demonstrates that the corresponding enzymes are produced at useful levels under fermentation conditions; reverse transcription quantitative PCR, enzyme assay, or direct measurement of ethanol and lactate yield would be required for that. Second, the study rests on four isolates recovered by culture-dependent methods on two selective media, and such approaches under-represent the fastidious and non-culturable fraction of a fermentation community relative to metabarcoding [9]; the isolates characterised here therefore represent the culturable dominant flora rather than a census. Third, species-level identification rests on a single locus, which proved insufficient for a species call in isolate T9 and, judged against the ITS thresholds of Vu et al. [17], leaves the species attribution of both yeast isolates short of secure. Fourth, the functional gene amplicons were scored by size and were not sequenced, so target identity is inferred rather than demonstrated. Fifth, the presence-absence calls rest on single reactions scored by amplicon size; replicate amplifications and explicit reporting of no-template and positive amplification controls would strengthen them. Finally, the sequences generated in this study require deposition in GenBank and citation of their accession numbers, and we recommend completing this before publication.

5. Conclusion

This study identified bacterial and yeast isolates from traditional and modified tepache by 16S rRNA and ITS sequencing and screened them for the genes underpinning the core fermentative reactions. Both yeast isolates belong to the genus *Saccharomyces*, with *S. cerevisiae* as the closest species-level match, and each carried *ADH1*, *PDC* and *SUC2*, which together establish a complete genotypic route from sucrose hydrolysis to ethanol on this sucrose-rich substrate.

The bacterial isolate from the modified formulation matched *Bacillus subtilis*, and that from the traditional formulation belongs to the genus *Lacticaseibacillus* sp. at an identity supporting genus- but not species-level assignment. Both bacterial isolates carried lactate dehydrogenase and *scr* cluster sequences, detected as multiple fragments consistent with the known multiplicity of these loci. Reformulation with Nigerian botanicals did not eliminate the functional gene complement that fermentation requires, although it accompanied the recovery of a different bacterial genus. These isolates constitute a genotypically justified shortlist for evaluation as defined starter cultures, subject to the strain-level safety, identification and expression studies identified above.

References

1. Abushattal S, Vences A, Barca AV, Osorio CR. Diverse horizontally-acquired gene clusters confer sucrose utilization to different lineages of the marine pathogen *Photobacterium damselae* subsp. *damselae*. Genes (Basel). 2020;11(11):1244. https://doi.org/10.3390/genes11111244
2. Allshouse T, Amendano M, Caruso B, Del Campo R, Murphy G, Shaffer L, et al. The microbiota of homemade tepache includes antibiotic-resistant microorganisms. MicroPubl Biol. 2025;2025. https://doi.org/10.17912/micropub.biology.001744
3. Butor I, Jančová P, Purevdorj K, Klementová L, Kluz M, Huňová I, et al. Effect of selected factors influencing biogenic amines degradation by *Bacillus subtilis* isolated from food. Microorganisms. 2023;11(4):1091. https://doi.org/10.3390/microorganisms11041091
4. Chen Y, Luo W, Fu M, Yu Y, Wu J, Xu Y, et al. Effects of selected *Bacillus* strains on the biogenic amines, bioactive ingredients and antioxidant capacity of shuidouchi. Int J Food Microbiol. 2023;388:110084. https://doi.org/10.1016/j.ijfoodmicro.2022.110084
5. Colautti A, Ginaldi F, Camprini L, Comi G, Reale A, Iacumin L. Investigating safety and technological traits of a leading probiotic species: *Lacticaseibacillus paracasei*. Nutrients. 2024;16(14):2212. https://doi.org/10.3390/nu16142212
6. de Almeida Silva Vilela J, de Figueiredo Vilela L, Ramos CL, Schwan RF. Physiological and genetic characterization of indigenous *Saccharomyces cerevisiae* for potential use in productions of fermented maize-based beverages. Braz J Microbiol. 2020;51(3):1297-307. https://doi.org/10.1007/s42770-020-00271-8
7. Fang S, Song X, Cui L, Bai J, Lu H, Wang S. The lactate dehydrogenase gene is involved in the growth and metabolism of *Lacticaseibacillus paracasei* and the production of fermented milk flavor substances. Front Microbiol. 2023;14:1195360. https://doi.org/10.3389/fmicb.2023.1195360
8. Gosselin S, Fullmer MS, Feng Y, Gogarten JP. Improving phylogenies based on average nucleotide identity, incorporating saturation correction and nonparametric bootstrap support. Syst Biol. 2022;71(2):396-409. https://doi.org/10.1093/sysbio/syab060
9. Gutiérrez-Sarmiento W, Peña-Ocaña BA, Lam-Gutiérrez A, Guzmán-Albores JM, Jasso-Chávez R, Ruíz-Valdiviezo VM. Microbial community structure, physicochemical characteristics and predictive functionalities of the Mexican tepache fermented beverage. Microbiol Res. 2022;260:127045. https://doi.org/10.1016/j.micres.2022.127045
10. Hackmann TJ. Setting new boundaries of 16S rRNA gene identity for prokaryotic taxonomy. Int J Syst Evol Microbiol. 2025;75(4):006747. https://doi.org/10.1099/ijsem.0.006747
11. Lee K, Kwon SH, Song S, Lee DY, Park MK, Kim YS. Comparative analysis of volatile and non-volatile metabolites derived from *Bacillus subtilis* strains producing different levels of biogenic amines. Metabolites. 2023;13(2):219. https://doi.org/10.3390/metabo13020219
12. Lin L, Zhang Z, Tang H, Guo Y, Zhou B, Liu Y, et al. Enhanced sucrose fermentation by introduction of heterologous sucrose transporter and invertase into *Clostridium beijerinckii* for acetone-butanol-ethanol production. R Soc Open Sci. 2021;8(9):201858. https://doi.org/10.1098/rsos.201858
13. Muller G, de Godoy VR, Dário MG, Duval EH, Alves-Jr SL, Bücker A, et al. Improved sugarcane-based fermentation processes by an industrial fuel-ethanol yeast strain. J Fungi (Basel). 2023;9(8):803. https://doi.org/10.3390/jof9080803
14. Nwaiwu O, Aduba CC, Igbokwe VC, Sam CE, Ukwuru MU. Traditional and artisanal beverages in Nigeria: microbial diversity and safety issues. Beverages. 2020;6(3):53. https://doi.org/10.3390/beverages6030053
15. Romero-Luna HE, Peredo-Lovillo A, Hernández-Mendoza A, Hernández-Sánchez H, Cauich-Sánchez PI, Ribas-Aparicio RM, et al. Probiotic potential of *Lactobacillus paracasei* CT12 isolated from water kefir grains (tibicos). Curr Microbiol. 2020;77(10):2584-92. https://doi.org/10.1007/s00284-020-02016-0
16. Tapsoba F, Kagambèga B, Tarnagda B, Zongo O, Zongo C, Savadogo A. Fermentative capabilities of *Saccharomyces* yeast and lactic acid bacteria strains in pure and mixed cultures during *Borassus akeassii* sap fermentation. Sci Afr. 2023;21:e01793. https://doi.org/10.1016/j.sciaf.2023.e01793
17. Vu D, Groenewald M, Szöke S, Cardinali G, Eberhardt U, Stielow B, et al. DNA barcoding analysis of more than 9,000 yeast isolates contributes to quantitative thresholds for yeast species and genera delimitation. Stud Mycol. 2016;85:91-105. https://doi.org/10.1016/j.simyco.2016.11.007
18. Yu A, Zhao Y, Li J, Li S, Pang Y, Zhao Y, et al. Sustainable production of FAEE biodiesel using the oleaginous yeast *Yarrowia lipolytica*. MicrobiologyOpen. 2020;9(7):e1051. https://doi.org/10.1002/mbo3.1051
19. Zheng J, Wittouck S, Salvetti E, Franz CMAP, Harris HMB, Mattarelli P, et al. A taxonomic note on the genus *Lactobacillus*: description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. Int J Syst Evol Microbiol. 2020;70(4):2782-858. https://doi.org/10.1099/ijsem.0.004107