INTERNATIONAL JOURNAL OF
SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

RESEARCH ARTICLE
Eilers *et al*., *Int. J. Syst. Evol. Microbiol.* 2023;73:006038
DOI 10.1099/ijsem.0.006038

MICROBIOLOGY SOCIETY

OPEN ACCESS

Lactobacillus isalae sp. nov., isolated from the female reproductive tract

Tom Eilers, Jelle Dillen, Sarah Ahannach, Leonore Vander Donck, Nele Van de Vliet, Stijn Wittouck and Sarah Lebeer*

Abstract

A novel strain of the genus *Lactobacillus*, named AMBV1719^T, was isolated from the vagina of a healthy participant in our large-scale citizen science project on the female microbiome, named Isala. Phylogenetic analysis showed that the 16S rRNA gene of AMBV1719^T is most similar to that of *Lactobacillus taiwanensis* with a sequence similarity of 99.873%. However, a genome-wide comparison using average nucleotide identity (ANI) revealed that isolate AMBV1719^T showed the highest ANI with *Lactobacillus paragasseri* JCM 5343^T, with a value of only 88.17%. This low ANI value with the most closely related strains known to date indicated that AMBV1719^T represents a distinct species. This strain has a limited ability to degrade carbon sources compared to *Lactobacillus gasseri*, indicating its adaptation to the host. Its genome has a length of 2.12 Mb with a G+C content of 34.8 mol%. We thus propose the name *Lactobacillus isalae* sp. nov. for this novel species, with AMBV1719^T (=LMG 32886^T=CECT 30756^T) as the type strain.

INTRODUCTION

In humans, the main habitat of lactobacilli is in the female reproductive tract, particularly the vagina. However, these bacteria are also found in other parts of the human body, such as the gut [1], nose [2] and skin [3], as well as in the intestines of other vertebrate and invertebrate hosts, on plants, and in fermented foods and feed [4]. Lactobacilli are facultative anaerobic, Gram-positive, non-spore-forming, rod-shaped bacteria that belong to the family *Lactobacillaceae*, order *Lactobacillales*, class *Bacilli*, phylum *Bacillota*. The term lactobacilli denotes the former genus *Lactobacillus*, which includes genetically distinct groups that were recently reclassified into 25 genera [5]. The current genus *Lactobacillus*, with its type species *Lactobacillus delbrueckii*, includes 44 species that are diverse in their phylogeny, ecology and metabolism [5]. These bacteria are mainly isolated from vertebrate and invertebrate hosts as well as dairy fermentations. Although they all ferment hexoses through the Embden–Meyerhoff pathway, specific strains of this genus also have the ability to ferment niche-specific carbohydrates including fructans and glycogen.

In the vagina, four species of lactobacilli are typically found in high abundances, namely *Lactobacillus crispatus*, *Lactobacillus iners*, *Lactobacillus jensenii* and *Lactobacillus gasseri*. At least one of these species is present in up to 80% of women and dominance of lactobacilli is considered to be an optimal vaginal state [6]. Members of the species *L. gasseri* are commonly found, apart from in the female reproductive tract, in the oral cavity and intestinal tract, also in wounds, urine, blood, carious dentine and pus of patients suffering from septic infections [5]. *L. gasseri* is reported to have various health benefits, such as antimicrobial activity (e.g. lactic acid production) and immunomodulation, while some strains also harbour strain-specific bacteriocins [7]. Advances in sequencing technology have enabled the description of closely related species that are distinct from *L. gasseri*. For example, the novel species *Lactobacillus paragasseri* JCM 5343^T was recently described based on whole-genome sequencing analyses with a 16S rRNA gene sequence similarity of 99.9% to *L. gasseri* [8].

Author affiliations: [1]Department of Bioscience Engineering, Research Group Environmental Ecology and Applied Microbiology, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium.
*Correspondence: Sarah Lebeer, sarah.lebeer@uantwerpen.be

Keywords: citizen science; female health; novel species; vaginal microbiome.

Abbreviations: ANI, average nucleotide identity; GTDB, genome taxonomy database; MAGs, metagenome-assembled genomes; MRS, Man, -Rogosa and -Sharpe.

The EMBL-EBI accession numbers of the 16S rRNA gene and the genome sequences of AMBV1719^T are OX442440 and SAMEA112195920, respectively.

Table 1. Average nucleotide identity (ANI) and 16S rRNA sequence identity of AMBV1719^{T} and the GTDB reference genomes from *Lactobacillus*, including their corresponding accession numbers, species and NCBI strain names

The official type strains are mentioned with a ^{T}. Note that, in some cases, there is an _X suffix which indicates that the GTDB has split a species because one or more genomes are too dissimilar to the type strain (as determined by ANI) to be part of the same species. When the ANI was lower than 70, fastANI was unable to find enough coverage (the default minimum is 50 matched fragments for trusting ANI) [19]. The GTDB reference strains also contain Metagenome-assembled genomes(MAGs), so not all 16S rRNA genes were available. NA, Not available.

Species	Strain	Genome	rRNA (%)	ANI (%)
Lactobacillus paragasseri	JCM 5343^{T}	GCF_003584685.1	99.618	88.2683
Lactobacillus gasseri	ATCC 33323^{T}	GCF_000014425.1	99.618	87.9729
Lactobacillus taiwanensis	DSM 21401^{T}	GCF_001436695.1	99.873	87.4852
Lactobacillus johnsonii	ATCC 33200^{T}	GCF_000159355.1	99.809	86.1453
Lactobacillus sp. 910589175	MGBC162746	GCA_910589175.1	NA	86.0295
Lactobacillus sp. 910589675	MGBC166701	GCA_910589675.1	NA	85.1701
Lactobacillus hominis	CRBIP 24.179^{T}	GCF_000296835.1	98.854	80.634
Lactobacillus intestinalis	DSM 6629^{T}	GCF_024397795.1	92.621	80.2636
Lactobacillus sp. 014803895	Mgla_MAG_30	GCA_014803895.1	NA	79.7756
Lactobacillus crispatus	ATCC 33820^{T}	GCF_018987235.1	92.557	79.5406
Lactobacillus amylovorus	DSM 20531^{T}	GCF_002706375.1	92.498	79.5188
Lactobacillus sp. 014803905	Mgla_MAG_37	GCA_014803905.1	NA	79.5104
Lactobacillus rodentium	DSM 24759^{T}	GCF_003423665.1	97.145	79.4358
Lactobacillus sp. 007570935	LL6	GCF_007570935.1	93.125	79.3817
Lactobacillus colini	DSM 101872^{T}	GCF_017874575.1	97.008	79.3138
Lactobacillus hamsteri	JCM 6256^{T}	GCF_000615445.1	92.635	79.152
Lactobacillus kitasatonis	JCM 1039^{T}	GCF_000615285.1	92.498	79.1516
Lactobacillus panisapium	DSM110155^{T}	GCF_943193025.1	92.366	79.0926
Lactobacillus amylolyticus	DSM 11664 ^{T}	GCF_000178475.1	92.112	79.0526
Lactobacillus ultunensis	DSM 16047^{T}	GCF_001436305.1	92.557	79.0107
Lactobacillus acidophilus	NBRC 13951^{T}	GCF_001591845.1	92.553	78.9408
Lactobacillus sp. 022483245	UW_TS_LAC3_1	GCA_022483245.1	93.325	78.9016
Lactobacillus helveticus	DSM 20075^{T}	GCF_000160855.1	92.812	78.8856
Lactobacillus gallinarum	DSM 10532^{T}	GCF_001434975.1	92.939	78.7609
Lactobacillus xujianguonis	HT111-2^{T}	GCF_004009905.1	92.63	78.7523
Lactobacillus sp. 930989465	ERR3414580	GCA_930989465.1	NA	78.7362
Lactobacillus kalixensis	DSM 16043^{T}	GCF_001434335.1	92.176	78.7132
Lactobacillus mulieris	c10Ua161M^{T}	GCF_007095465.1	93.893	78.5997
Lactobacillus psittaci	DSM 15354^{T}	GCF_000425905.1	94.084	78.584
Candidatus Lactobacillus pullistercoris	F6-686^{T}	GCA_018883635.1	NA	78.5414
Lactobacillus pasteurii	CRBIP 24.76^{T}	GCF_000297025.1	92.994	78.489
Lactobacillus gigeriorum	CRBIP 24.85^{T}	GCF_000296855.1	93.163	78.4763
Lactobacillus corticis	B40^{T}	GCF_018327645.1	93.316	78.4357
Lactobacillus iners	DSM 13335^{T}	GCF_000160875.1	95.217	78.2969

Continued

Table 1. Continued

Species	Strain	Genome	rRNA (%)	ANI (%)
Lactobacillus jensenii	DSM 20557^T	GCF_001436455.1	93.957	78.2511
Lactobacillus kefiranofaciens	ATCC 43761^T	GCF_900103655.1	91.661	78.2302
Lactobacillus apis	LMG 26964^T	GCF_002837055.1	91.974	78.1011
Lactobacillus laiwuensis	F551-2^T	GCF_019972835.1	90.944	77.8403
Lactobacillus sp. 016100975	W8174	GCF_016100975.1	90.875	77.7699
Lactobacillus acetotolerans	DSM 20749^T	GCF_001436775.1	92.571	77.746
Lactobacillus kimbladii_B	Dan47	GCF_014323605.1	91.001	77.6668
Lactobacillus kimbladii	Hma2^T	GCF_000970755.1	93.248	77.6666
Candidatus Lactobacillus pullistercoris	44987^T	GCF_002916935.1	91.456	77.6138
Lactobacillus huangpiensis	F306-1^T	GCF_019972815.1	91.007	77.4882
Lactobacillus melliventris	Hma8^T	GCF_000970775.1	93.612	77.4865
Lactobacillus kullabergensis	Biut2^T	GCF_000967195.1	90.88	77.4406
Lactobacillus helsingborgensis	Bma5^T	GCF_000970855.1	90.99	77.3939
Lactobacillus sp. 003692965	ESL0246	GCF_003692965.1	91.013	77.261
Lactobacillus bombicola	R-53102^T	GCF_900112665.1	91.076	77.0621
Lactobacillus sp. 945290125	SRR11267073	GCA_945290125.1	NA	76.3113
Lactobacillus nasalidis	YZ01^T	GCF_016587775.1	91.128	NA
Lactobacillus delbrueckii	DSM 20074^T	GCF_001433875.1	90.963	NA
Lactobacillus absiana	YH-C36a	GCA_017309015.1	90.878	NA
Lactobacillus equicursoris	DSM 19284^T	GCF_001434815.1	90.758	NA
Lactobacillus delbrueckii_B	W1P13.016	GCA_004556255.1	85.548	NA
Lactobacillus sp. 945281075	SRR11267042	GCA_945281075.1	NA	NA
Lactobacillus sp. 002418055	UBA5804	GCA_002418055.1	NA	NA
Lactobacillus sp. 944327175	E27	GCA_944327175.1	NA	NA
Lactobacillus sp. 002417825	UBA5815	GCA_002417825.1	NA	NA
Lactobacillus sp. 905214545	ERR1430435	GCA_905214545.1	NA	NA
Lactobacillus sp. 022641965	UW_TS_LAC6_1	GCA_022641965.1	NA	NA

This manuscript describes a novel species closely related to *L. gasseri* and *L. paragasseri*. This species was first isolated from the vagina of a healthy participant in our large-scale citizen science project on the female microbiome in Belgium, which we named Isala (https://isala.be/en/ [9], after Isala Van Diest (1842–1916), who was the first female physician in Belgium.

SAMPLING, ISOLATION AND ECOLOGY

Isolate AMBV1719^T was isolated from the human vagina of a healthy participant in the Isala citizen science project [9]. The participant was 26 years old and lived in the Antwerp metropolitan area, Belgium. The study was approved by the Ethical Committee of the Antwerp University Hospital/University of Antwerp (B300201942076) and registered online at https://clinicaltrials.gov with the unique identifier NCT04319536. The participant self-sampled with an ESwab (Copan) and the samples were transported using the national mail services (BPost) at room temperature. After 1 day of transport, the sample, which was stored in Amies medium transport buffer, was stored with 25% glycerol at −80 °C until further processing. The sample was then plated on de Man–Rogosa–Sharpe (MRS) agar and grown for 2 days at 37 °C and 5% CO_2. Single colonies were picked and re-streaked four times to obtain a pure culture. It was then picked for colony PCR (forward primer, 27F 5′-AGAGTTTGATCMTGGCTCAG-3′; reverse primer, 1492R 5′-GGTTACCTTGTTACGACTT-3′) [10] and this fragment

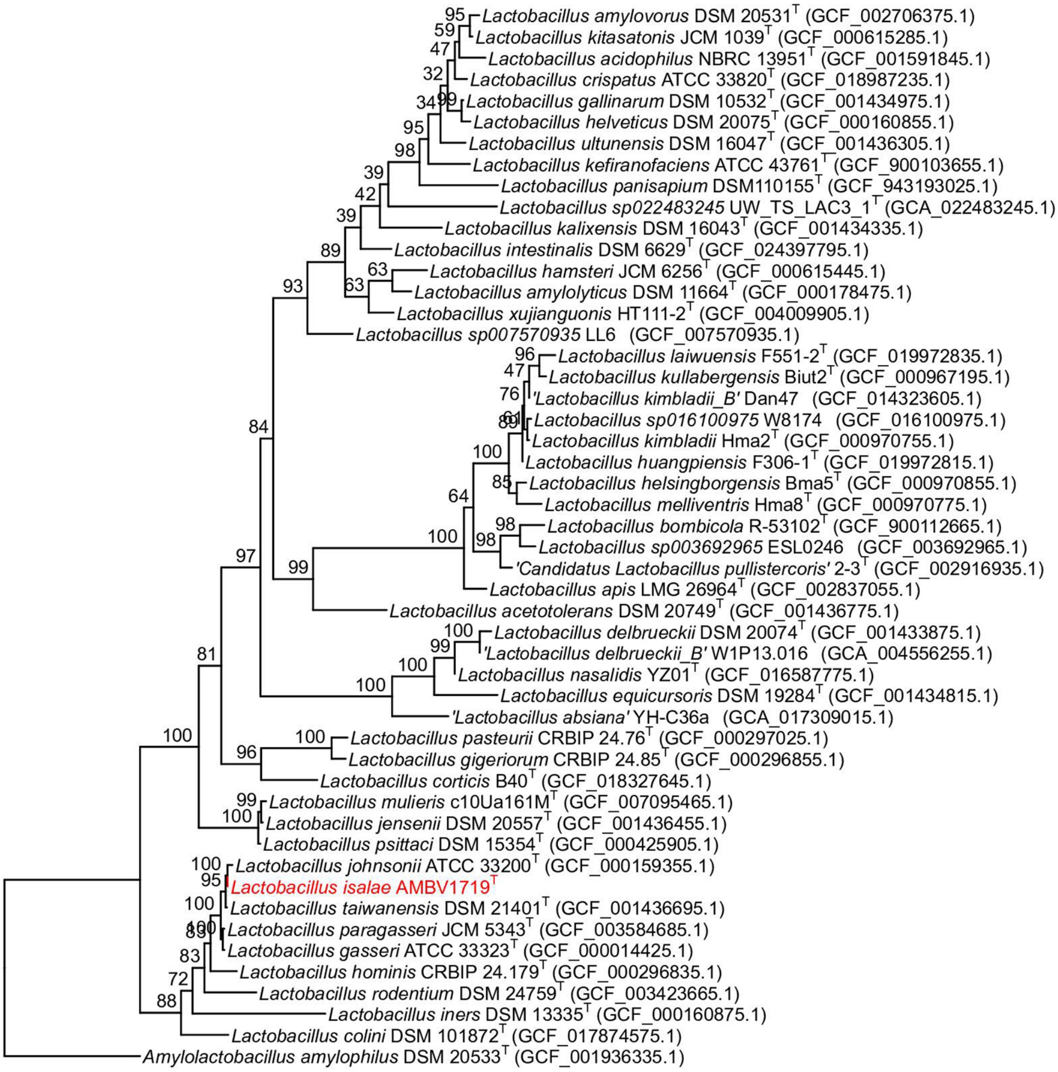

Fig. 1. 16S rRNA tree of all species of the genus *Lactobacillus* and the outgroup *Amylolactobacillus amylophilus*. All representative strains (Table 1) were downloaded from the GTDB [23]. 16S rRNA genes were extracted from the genomes using Barrnap (https://github.com/tseemann/barrnap) and aligned using MAFFT [26]. The maximum-likelihood tree was reconstructed with IQ-TREE with the GTR+F+G4 substitution model [27]. The tree was rooted with tidygenomes [28] and visualized with ggtree [29]. Non-validly published species are indicated between quotes ('')

was subsequently sequenced using Sanger sequencing at the Genetic Service Facility (VIB, Wilrijk, Belgium). The genus of the strain was identified based on close matches using NCBI BLAST [11] and EzBioCloud [12], positioning this isolate in the genus *Lactobacillus*.

GENOME FEATURES

The DNA of AMBV1719^T was extracted using a method based on the P3 protocol by Alimolaei and Golchin with an additional ampicillin treatment (final ampicillin concentration of 0.1 mg ml^{-1} for 1 h) to improve cell lysis [13]. Whole genome sequencing was performed using the Nextera XT DNA Sample Preparation kit (Illumina) and the Illumina MiSeq platform (2×250 cycles) at the Laboratory of Medical Microbiology (University of Antwerp, Antwerp, Belgium) according to Wuyts *et al.* [14]. In addition to short-read sequencing, nanopore sequencing was also performed on the MinION Mk1C apparatus. For long-read sequencing, DNA was extracted using the MagAttract HMW DNA kit and sequenced using the Ligation sequencing kit (LSK-109) with an additional Native Barcoding Expansion 1–12 (EXP-NBD104) on an R9 chemistry flowcell (FLO-MIN106D). Long reads were assembled using Flye [15] and subsequently polished three consecutive times using Illumina reads with Polypolish [16]. This analysis resulted in the complete assembly of the genome. The assembly is available from the European Nucleotide Archive (project PRJEB57255). The genome of isolate AMBV1719^T has a total

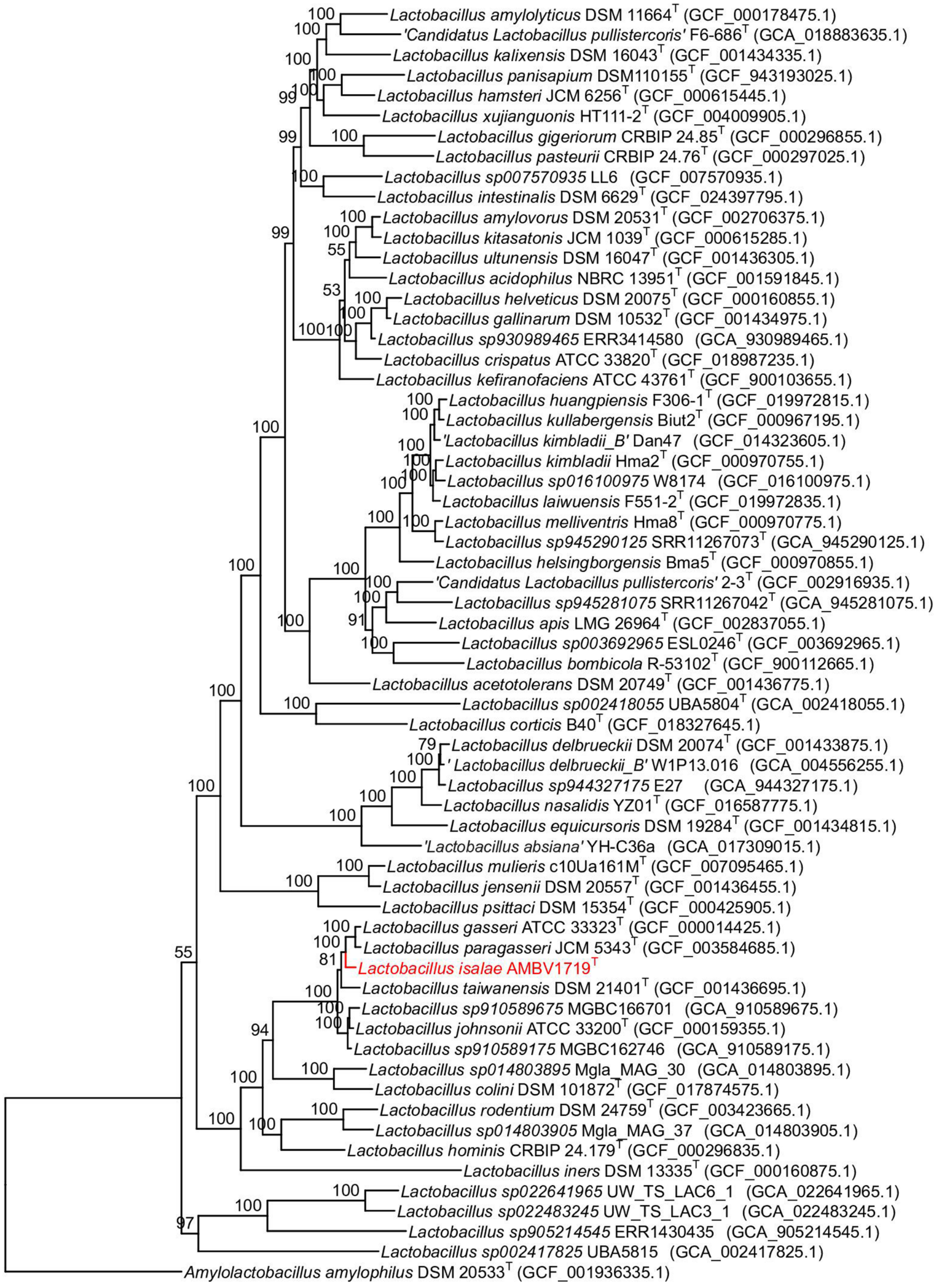

Fig. 2. Core genome phylogenetic tree of the genus *Lactobacillus* with outgroup *Amylolactobacillus amylophilus* based on all single-copy core genes of the genus *Lactobacillus*. All reference strains of this genus were downloaded from GTDB [23] and genes were predicted with Prodigal [24]. Core genes were determined, aligned and concatenated with SCARAP [25]. 407 core genes from all representatives were aligned using MAFFT [26] and a maximum-likelihood tree was reconstructed with IQ-TREE with the LG+F+G4 amino acid substitution model [27]. The tree was rooted with tidygenomes [28] and visualized with ggtree [29]. Non-validly published species are indicated between quotes ('')

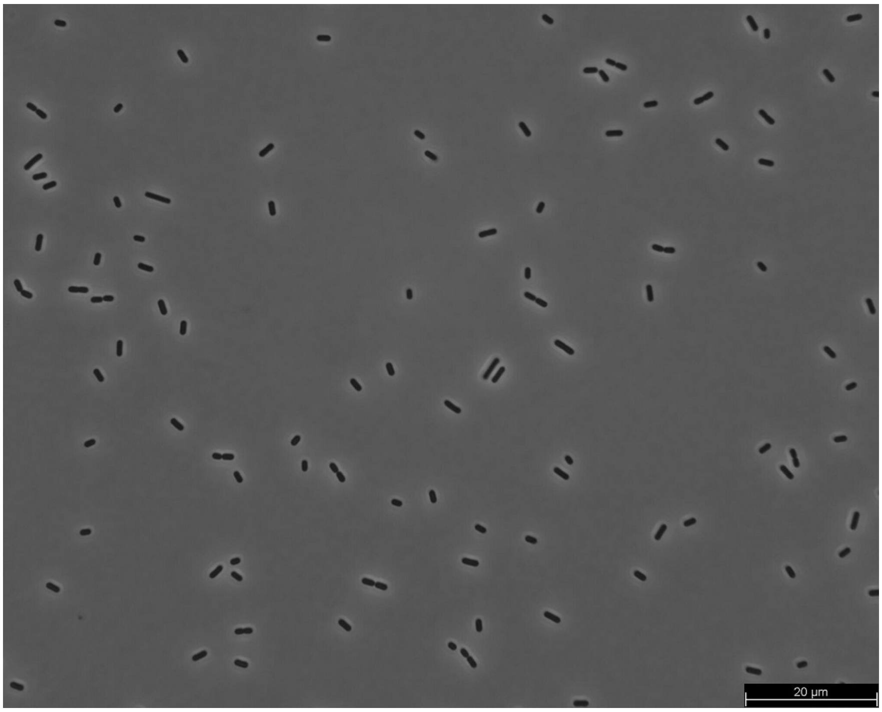

Fig. 3. Light microscope picture (Leica DMi8) at a magnification of ×1000 of AMBV1719^T grown for 24 h on 37 °C +5% CO_2 in MRS medium.

length of 2.12 Mb and a G+C content of 34.8 mol% without any plasmids present. The completeness and contaminations were assessed with CheckM and reported as 99.22 and 0.58%, respectively. However, it is important to note that this evaluation is a prediction and contains a standard error [17]. Annotation with Prokka [18] identified 2058 genes in total with 1959 CDSs and 21 rRNA genes in seven operons, 77 tRNA and one tmRNA genes.

POSITION WITHIN THE GENUS *LACTOBACILLUS*

To classify AMBV1719^T using the 16S rRNA gene, the rRNA genes of AMBV1719^T and representative strains of the genus *Lactobacillus* (Table 1) were downloaded from the Genome Taxonomy Database (GTDB) and were extracted from their genomes with Barrnap and an in-house script. To accurately determine the phylogenic relatedness of the novel species with other members of the genus *Lactobacillus*, 19 non-validly published genomes were included. The 16S rRNA genes were then compared using BLASTN [11]. This method confirmed that the 16S rRNA gene of *L. taiwanensis* was the most similar to that of AMBV1719^T (99.873%), followed by *Lactobacillus johnsonii* (99.809%), *L. gasseri* (99.618%) and *L. paragasseri* (99.618%) (Table 1).

The average nucleotide identity (ANI) between strain AMBV1719^T and representative strains of the genus *Lactobacillus* was calculated using fastANI [19]. This method indicated that *L. paragasseri* JCM 5343^T (GCF_003584685.1) and *L. gasseri* ATCC 33323^T were most similar to strain AMBV1719^T, with ANI values of 88.17 and 87.84%, respectively (Table 1). These ANI values are lower than the 94–95% threshold for a novel species [20, 21]. Due to insufficient coverage, fastANI was unable to calculate ANI values for an ANI under 70%. Generally, fastANI requires a minimum 50 matched fragments to trust ANI calculations. The discrepancy between the percentage identity of the 16S rRNA gene and the ANI confirmed that relying solely on 16S rRNA identification is not sufficient for the discovery of novel species, as described by Fox *et al.* [22].

To gain better understanding of the position of this novel species within the phylogeny of the genus *Lactobacillus*, a 16S rRNA gene tree (Fig. 1) and a core genome tree (Fig. 2) were reconstructed. The core genome tree was inferred using all reference strains of *Lactobacillaceae* downloaded from GTDB [23]. Genes were predicted using Prodigal [24], and single copy core genes were identified, aligned and concatenated with SCARAP [25]. This resulted in 407 core genes. The 16S rRNA genes and 407 single-copy core genes were aligned using MAFFT [26] and a maximum-likelihood tree was reconstructed with IQ-TREE with the GTR+F substitution model and the LG+F+G4 amino acid substitution model, respectively [27]. The 16S rRNA gene and core genome trees were both rooted with tidygenomes [28] with *Amylolactobacillus amylophilus* DSM 20533 (GCF_001936335.1) as the outgroup and visualized using ggtree [29]. The 16S rRNA gene tree (Fig. 1) and the core genome

Table 2. Overview of the ability of carbon source utilization of strain AMBV1719^T and reference strain *Lactobacillus gasseri* ATCC 33323^T, measured using the API (bioMérieux) assay which shows acidification of the medium based on a colour reaction where yellow corresponds with complete acidification (+), green with partial acidification (+/−) and blue with no growth on this carbon source (−)

	Strain AMBV1719^T		***L. gasseri* ATCC 33323^T**			**Strain AMBV1719^T**		***L. gasseri* ATCC 33323^T**	
	24 h	**48 h**	**24 h**	**48 h**		**24 h**	**48 h**	**24 h**	**48 h**
Control	–	–	–	–	Aesculin ferric citrate	–	–	–	–
Glycerol	–	–	–	–	Salicin	–	+/–	+	+
Erythritol	–	–	–	–	Cellobiose	–	–	+	+
D-Arabinose	–	–	–	–	Maltose	–	+	+	+
L-Arabinose	–	–	–	–	Lactose	–	–	+	+
D-Ribose	–	–	–	–	Melibiose	–	–	–	–
D-Xylose	–	–	–	–	Sucrose	+	+	+	+
L-Xylose	–	–	–	–	Trehalose	–	+	+	+
D-Adonitol	–	–	–	–	Inulin	–	–	–	–
Methyl β-D-xylopyranoside	–	–	–	–	Melezitose	–	–	–	+
D-Galactose	–	+/–	+/–	+	Raffinose	–	–	–	–
D-Glucose	+	+	+	+	Starch	–	+/–	–	+
D-Fructose	–	–	+	+	Glycogen	–	–	–	–
D-Mannose	–	+/–	–	+	Xylitol	–	–	–	–
L-Sorbose	–	–	–	–	Gentiobiose	–	+/–	+	+
L-Rhamnose	–	–	–	+/–	Turanose	–	–	–	–
Dulcitol	–	–	–	–	D-Lyxose	–	–	–	–
Inositol	–	–	–	–	D-Tagatose	–	+/–	+/–	+
D-Mannitol	–	–	–	–	D-Fucose	–	–	–	–
D-Sorbitol	–	–	–	–	L-Fucose	–	–	–	–
Methyl α-D-mannopyranoside	–	–	–	–	D-Arabitol	–	–	–	–
Methyl α-D-glucoside	–	–	–	–	L-Arabitol	–	–	–	–
N-Acetylglucosamine	–	–	–	+	Potassium gluconate	–	–	–	–
Amygdalin	–	–	+/–	+	Potassium 2-keto-gluconate	–	–	–	–
Arbutin	–	–	–	+/–	Potassium 5-keto-gluconate	–	–	–	–

tree (Fig. 2) show similarities; specifically, the novel species consistently appears in the same subclade as *L. paragasseri*, *L. gasseri*, *L. johnsonii* and *L. taiwanensis*. However, the trees also show relevant differences, indicating that the 16S rRNA gene tree may not accurately represent closely related species. For instance, in the 16S rRNA tree, *L. johnsonii* and *L. taiwanensis* are the most closely related species to *L. isalae*, while in the core genome tree, *L. paragasseri* and *L. gasseri* are.

CELL AND COLONY MORPHOLOGY AND ARRANGEMENT

In MRS broth, the strain was able to grow at 20–45 °C (optimum, 37 °C), after undergoing testing at temperatures ranging from 4 to 50 °C. When plated out on MRS agar and grown overnight at 37 °C, the bacteria form pointy, convex, undulate, white colonies. Upon Gram staining and examination under a light microscope at a magnification of ×1000 (Olympus CX41),

Table 3. The relative fatty acid contents (%) of strain AMBV1719^T and its closely related type strains

Strains: 1, AMBV1719^T; 2, *Lactobacillus gasseri* ATCC 33323^T; 3, *Lactobacillus paragasseri* JCM 5343^T [8]. –, Not detected; ND, not determined.

Fatty acid	**1**	**2**	***L. paragasseri* Jcm 5343^T**
Saturated:			
$C_{9:0}$	1.4	NA	NA
$C_{12:0}$	–	–	0.5
$C_{14:0}$	1.2	–	0.7
$C_{16:0}$	9.5	19.5	31.9
$C_{18:0}$	4.2	2.0	3.5
$C_{19:0}$ iso	1.4	2.0	1.8
Unsaturated:			
$C_{18:1}\omega 9c$	52.9	65.9	49.5
Ambiguous peaks:			
$C_{16:1}\omega 7c/C_{16:1}\omega 6c$	–	1.2	2.2
$C_{18:1}\omega 7c/C_{18:1}\omega 6c$	1.6	8.0	9.5
$C_{19:0}$ cyclo $\omega 10c/C_{19:1}\omega 6c$	27.7	NA	NA

the bacteria were observed to be Gram-positive, long, narrow rods that were in either single cells or in clusters or chains of two or three cells (Fig. 3). No spores were found after spore staining and examination under the microscope.

GROWTH OPTIMUM AND PHYSIOLOGICAL PROPERTIES

Strain AMBV1719^T was isolated on MRS medium. The strain was able to grow in a microaerobic atmosphere with 5.0% CO_2, but also grew well in both anaerobic and aerobic atmospheres. Growth of the strain was assessed across a range of pH levels from pH 2.5 to 11.5 and it was found that it could grow at pH 4.5–8.5 with an optimal pH of 7. The optimal salt concentration for growth was determined to be 0%, and it was able to grow in salt concentrations up to 4%. Catalase activity was determined under four conditions, as described by Legein *et al.* [30]. The conditions were a combination of aerobic and anaerobic, with and without menaquinone supplementation. To confirm whether AMBV1719^T is catalase-positive, the degradation of H_2O_2 was monitored. An overnight culture (2 µl) was inoculated onto a microscope slide and a few droplets of 3% H_2O_2 solution were added. No reaction was observed for the four conditions, indicating that the strain is catalase-negative.

CARBON SOURCE UTILIZATION

Carbon source utilization was determined using the API 50 CH system test (bioMérieux). Strain AMBV1719^T and *L. gasseri* ATCC 33323^T were grown on MRS agar overnight and were inoculated in the API 50 CHL medium. The strips were incubated at 37 °C and were read after 24 and 48 h, and the results are shown in Table 2. Carbon source utilization differed between strain AMBV1719^T and *L. gasseri*. For instance, AMBV1719^T does not ferment D-fructose, amygdalin, cellobiose, lactose, *N*-acetylglucosamine or melezitose, while *L. gasseri* fermented these carbon sources after 24 h, with the exception of *N*-acetylglucosamine and melezitose that were only fermented after 48 h. Fermentation by AMBV1719^T was slightly positive for D-galactose, D-mannose, salicin, starch, β-gentiobiose and D-tagatose after 48 h, while for *L. gasseri* they were strongly positive. L-Rhamnose and arbutin were not fermented by AMBV1719^T, but slightly fermented by *L. gasseri*. Strain AMBV1719^T is a homofermentative strain, because it does not ferment ribose, arabinose or gluconate [4].

FATTY ACID ANALYSIS

To obtain the cellular fatty acid composition (Table 3), standard protocols were followed as described in the MIDI Microbial Identification System [31] and performed by the Spanish Type Culture Collection (CECT, Valencia, Spain). The cellular fatty acid content was analysed using gas chromatography with an Agilent 6850 chromatographic unit, following the MIDI Microbial Identification System's TSBA6 method. The results were then identified with the Microbial Identification Sherlock software package [32]. Prior to analysis, strain AMBV1719^T was incubated on MRS medium at 37 °C for 24 h in an anaerobic generator system. The most abundant fatty acids found in the membrane of isolate AMBV1719^T were the unsaturated fatty acid $C_{18:1}$ ω9*c* and saturated $C_{16:0}$. Compared to its two most closely related species, strain AMBV1719^T had a slightly different saturated fatty acid composition, with relatively lower levels of $C_{16:0}$ and higher levels of $C_{18:0}$.

PROPOSAL OF *LACTOBACILLUS ISALAE* SP. NOV.

We propose that isolate AMBV1719^T represents a new species of the genus *Lactobacillus* based on the low ANI values between this isolate and the reference strains of the most closely related species *L. gasseri* and *L. paragasseri*. The proposed name for this new species is *Lactobacillus isalae* sp. nov.

DESCRIPTION OF *LACTOBACILLUS ISALAE* SP. NOV.

Lactobacillus isalae (i'sa.lae. N.L. fem. gen. n. isalae, named after our Belgian citizen science project, Isala, which was named after Isala van Diest, the first female Belgian physician).

Cells are homofermentative, Gram-positive, long, narrow rod-shaped and can form chains and clusters. They are able to grow in MRS media with either 5.0% CO_2, micro-aerobically, anaerobically or aerobically between 20 and 45 °C (optimum, 37 °C) and at pH 4.5–8.5 (optimum, pH 7). Tolerates salt concentrations up to 4%, but optimal growth occurs with no added salt. When grown on MRS agar, this novel species forms pointy, convex, undulate, white colonies. The type strain is able to ferment D-glucose, maltose, sucrose and trehalose. Strain AMBV1719^T also slightly ferments D-galactose, D-mannose, salicin, starch, gentiobiose and D-tagatose. The type strain is not able to ferment glycerol, erythritol, D-arabinose, L- arabinose, D-ribose, D-xylose, L-xylose, D-adonitol, methyl β-D-xylopyranoside, D-fructose, L-sorbose, L-rhamnose, dulcitol, inositol, D-mannitol, D-sorbitol, methyl D-mannopyranoside, methyl α-D-glucoside, *N*-acetyl-glucosamine, amygdalin, arbutin, aesculin ferric citrate, cellobiose, lactose, melibiose, inulin, melezitose, raffinose, glycogen, xylitol, turanose, D-lyxose, D-fucose, L-fucose, D-arabitol, L-arabitol, gluconate, 2-keto-gluconate or 5-keto-gluconate. Due to its limited carbon source usage, AMBV1719^T can be considered a very host-adapted strain.

The type strain, AMBV1719^T (=LMG 32886^T=CECT 30756^T), was isolated from the female reproductive tract. The EMBL-EBI accession numbers of the 16S rRNA gene and the genome sequences of AMBV1719^T are OX442440 and SAMEA112195920, respectively.

Funding information
T.E., J.D., S.A. and S.L. are funded by an ERC grant (Lacto-Be, 85600) in which the habitat adaptation of lactobacilli is investigated. N.V.D.V. is partially funded by the University of Antwerp (IOF service platform FFI220247 and IOF FFI220319). S.W. is partially funded by the University of Antwerp (IOF FFB220205) and an ERC grant (Lacto-BE, 85600). S.A. is partially funded by the Special Research Fund of the Universiteit Antwerpen (UAntwerpen BOF; DOCPRO 37054 grant). L.V.D is funded by the research foundation Flanders (FWO-SB;1SD0622N).

Acknowledgements
The authors would like to acknowledge Ines Tuyaerts and Sam Bakelants for practical work. We also would like to thank Prof. Giovanna Felis for her taxonomic advice.

Author contributions
T.E.: conceptualization, investigation, writing – original draft, project administration and visualization. N.V.D.V: investigation, writing –editing and practical work J.D.: DNA sequencing, writing – review and editing S.A.: Isala project manager and writing L.V.D: isolation of AMBV1719 and writing. S.W.: providing pangenome tools, supervision, methodology and writing – review and editing. S.L.: conceptualization, funding acquisition, resources, supervision and writing – review and editing

Conflicts of interest
There are no conflicts of interest related to this work. S.L. is an academic board member of ISAPP (isappscience.org) and chairperson and cofounder of YUN (yun.be), but these organisations were not involved in this work.

References
1. Marco ML, de Vries MC, Wels M, Molenaar D, Mangell P, *et al*. Convergence in probiotic *Lactobacillus* gut-adaptive responses in humans and mice. *ISME J* 2010;4:1481–1484.
2. De Boeck I, van den Broek MFL, Allonsius CN, Spacova I, Wittouck S, *et al*. *Lactobacilli* have a niche in the human nose. *Cell Rep* 2020;31:107674.
3. Lebeer S, Oerlemans EFM, Claes I, Henkens T, Delanghe L, *et al*. Selective targeting of skin pathobionts and inflammation with topically applied *Lactobacilli*. *Cell Rep Med* 2022;3:100521.
4. Duar RM, Lin XB, Zheng J, Martino ME, Grenier T, *et al*. Lifestyles in transition: evolution and natural history of the genus *Lactobacillus*. *FEMS Microbiol Rev* 2017;41:S27–S48.
5. Zheng J, Wittouck S, Salvetti E, Franz CMAP, Harris HMB, *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;70:2782–2858.
6. France M, Alizadeh M, Brown S, Ma B, Ravel J. Towards a deeper understanding of the vaginal microbiota. *Nat Microbiol* 2022;7:367–378.
7. Selle K, Klaenhammer TR. Genomic and phenotypic evidence for probiotic influences of *Lactobacillus* gasseri on human health. *FEMS Microbiol Rev* 2013;37:915–935.
8. Tanizawa Y, Tada I, Kobayashi H, Endo A, Maeno S, *et al*. *Lactobacillus paragasseri* sp. nov., a sister taxon of *Lactobacillus gasseri*,

based on whole-genome sequence analyses. *Int J Syst Evol Microbiol* 2018;68:3512–3517.

9. **Lebeer S, Ahannach S, Wittouck S, Gehrmann T, Eilers T, *et al*.** Citizen-science map of the vaginal microbiome. *In Review*. DOI: 10.21203/rs.3.rs-1350465/v1
10. **Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, *et al*.** Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* 2008;74:2461–2470.
11. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
12. **Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, *et al*.** Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
13. **Alimolaei M, Golchin M.** A comparison of methods for extracting plasmids from a difficult to lyse bacterium: *Lactobacillus casei*. *Biologicals* 2017;45:47–51.
14. **Wuyts S, Allonsius CN, Wittouck S, Thys S, Lievens B, *et al*.** Comparative genome analysis of *Lactobacillus mudanjiangensis*, an understudied member of the *Lactobacillus plantarum* group. *Microb Genom* 2019;5:e000286.
15. **Kolmogorov M, Yuan J, Lin Y, Pevzner PA.** Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 2019;37:540–546.
16. **Wick RR, Holt KE.** Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 2022;18:e1009802.
17. **Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW.** CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
18. **Seemann T.** Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
19. **Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S.** High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:1–8.
20. **Konstantinidis KT, Tiedje JM.** Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci*2005;102:2567–2572.
21. **Richter M, Rosselló-Móra R.** Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci*2009;106:19126–19131.
22. **Fox GE, Wisotzkey JD, Jurtshuk P.** How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J Syst Bacteriol* 1992;42:166–170.
23. **Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P-A, *et al*.** GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 2022;50:D785–D794.
24. **Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, *et al*.** Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;11:1–11.
25. **Wittouck S, Wuyts S, Meehan CJ, van Noort V, Lebeer S.** A genome-based species taxonomy of the *Lactobacillus* genus complex. *mSystems* 2019;4:e00264-19.
26. **Nakamura T, Yamada KD, Tomii K, Katoh K.** Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* 2018;34:2490–2492.
27. **Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ.** IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.
28. **De Boeck I, Wittouck S, Martens K, Spacova I, Cauwenberghs E, *et al*.** The nasal mutualist *Dolosigranulum pigrum* AMBR11 supports homeostasis *via* multiple mechanisms. *iScience* 2021;24:102978.
29. **Yu G, Smith DK, Zhu H, Guan Y, Lam TTY.** Ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 2017;8:28–36.
30. **Legein M, Wittouck S, Lebeer S.** *Latilactobacillus fragifolii* sp. nov., isolated from leaves of a strawberry plant (*Fragaria* x *ananassa*). *Int J Syst Evol Microbiol* 2022;72.
31. **Sasser M.** Bacterial identification by gas chromatographic analysis of fatty acid methyl esters (GC-FAME). In: *MIDI Technical Note 101*. 1990.
32. **MIDI Inc.** Sherlock Microbial Identification System Operating Manual; 2008. www.midi-inc.com [accessed 2 January 2023].