Genome-resolved metagenomics reveals a phylogenetically cohesive *Acetilactobacillus*-like species complex dominating stingless bee pot honey

Aurora Xolalpa-Aroche,[1,2,*] Haydeé Contreras-Peruyero,[3] Enrique J. Delgado-Suárez,[1,*] David I. Hernández-Mena,[4] Wilson I. Moguel-Chin,[5] J. Fausto Rivero-Cruz,[2] Rodrigo A. Velarde,[8] Elizabeth Ortiz-Vázquez,[6] Blanca E. Rivero-Cruz,[7] Jose Abel Lovaco Flores,[9] Lorena Rodríguez Orduña,[10] Cuauhtémoc Licona Cassani, Francisco Barona-Gómez[12] and Nelly Sélem-Mojica[3,*]

[1]Departamento de Medicina Preventiva, Universidad Nacional Autónoma de México Facultad de Medicina Veterinaria y Zootecnia, Ciudad Universitaria, 04510, Ciudad de México, México, [2] Centro de Innovación para el Desarrollo Apícola Sustentable en Quintana Roo (CIDASQROO), Universidad Intercultural Maya de Quintana Roo, Carretera Muna Felipe Carrillo Puerto, La Presumida José María Morelos, 77870, Quintana Roo, México, [3]Centro de Ciencias Matemáticas, Universidad Nacional Autónoma de México, Antigua Carretera a Pátzcuaro 8701, Col. Ex Hacienda San José de la Huerta Morelia, 58089, Michoacán, México, [4]Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Ciudad de México, México, [5]Facultad de Medicina Veterinaria y Zootecnia, Campus de Ciencias Biológicas y Agropecuarias, Universidad Autónoma de Yucatán, km 15.5 Carretera Mérida-Xmatkuil,Mérida, 97135, Yucatán, México, [6], Tecnológico Nacional de México/ITMérida, Avenida Tecnológico s/n Km. 4.5, Mérida, 97118, Yucatán, México, [7]Departamento de Farmacia, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510, Ciudad de México, México, [8]Departamento de Investigación y Desarrollo, Latin American Society for Bee Research, , , SOLATINA, Uruguay, [9]Unidad Irapuato, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV), Libramiento Norte Carretera Irapuato León Kilómetro 9.6, Carr Panamericana Irapuato León, 36821, Guanajuato, México, [10]Industrial Genomics Laboratory, Centro de Biotecnología FEMSA, Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey, Av Eugenio Garza Sada 2501, Sur, 64849, Nuevo León, México, Integrative Biology Research Unit, The Institute for Obesity Research, Tecnológico de Monterrey, Av Eugenio Garza Sada 2501, Sur, 64849, Nuevo León, México and [12]Evolution of Microbial Chemodiversity Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE, Leiden, The Netherlands

*Corresponding author. nselem@matmor.unam.mx, aurora.xolalpa@uimqroo.edu.mx, enriquedelgado.suarez@gmail.com

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Pot honey, the honey produced by stingless bees, is valued for its antimicrobial capacity, which may be influenced by its microbial content. While Lactobacillaceae species are commonly associated with honeybees and honey microbiomes, most studies have focused on *Apis mellifera*, leaving pot honey microbial diversity largely unexplored. We present the first pot honey shotgun metagenomic analysis from bee species *Melipona beecheii* and *Scaptotrigona mexicana*. We reconstructed 24 metagenome-assembled genomes (MAGs), 15 of which lacked close matches to any described species, showing ≤81% Average Nucleotide Identity (ANI) to available reference genomes. Phylogenetic analyses resolved these MAGs into four well-defined clades (intraclade ANI > 99%, interclade ANI ≤ 81%), consistent with four novel species within the family Lactobacillaceae. GTDB-Tk classification placed MAG clades 1 and 2 closest to *Nicoliella*, and clades 3 and 4 closest to *Acetilactobacillus*. We validated the presence of these lineages in honey by sequencing three isolates that clustered within MAG clade 2. Aminoacid similarity (AAI/cAAI) indicates the presence of two genus-level lineages: one occupying a transitional genomic space near *Nicoliella*, and a second representing an undescribed genus. The genomic similarity of our MAGs and isolates to those from pot honey or larval food in Malaysia, Brazil, and Australia suggests these taxa are closely associated with stingless bees and may contribute to honey properties. By reducing the genomic underrepresentation of evolutionarily divergent sister clades related to *Nicoliella* and *Acetilactobacillus*, our genome-resolved analyses reveal a globally distributed, phylogenetically cohesive Lactobacillaceae species complex dominating pot honey.

Key words: Stingless bees, honey microbiome, *Acetilactobacillus* , *Nicoliella*, *MAG*

Introduction

For centuries, indigenous groups have managed stingless bees of the tribe Meliponini to produce pot honey [1, 2], which is recognized for their antimicrobial properties [3, 4], therapeutic potential [5, 6] , and traditional use in treating ailments [6, 7]. Growing interest in stingless bee products such as honey [3, 4, 8], and propolis [9], highlights the potential of beekeeping to promote family [10] and community economic development [11, 12]. Beyond their economic value, stingless bees contribute significantly to ecosystem services,

© The Author .

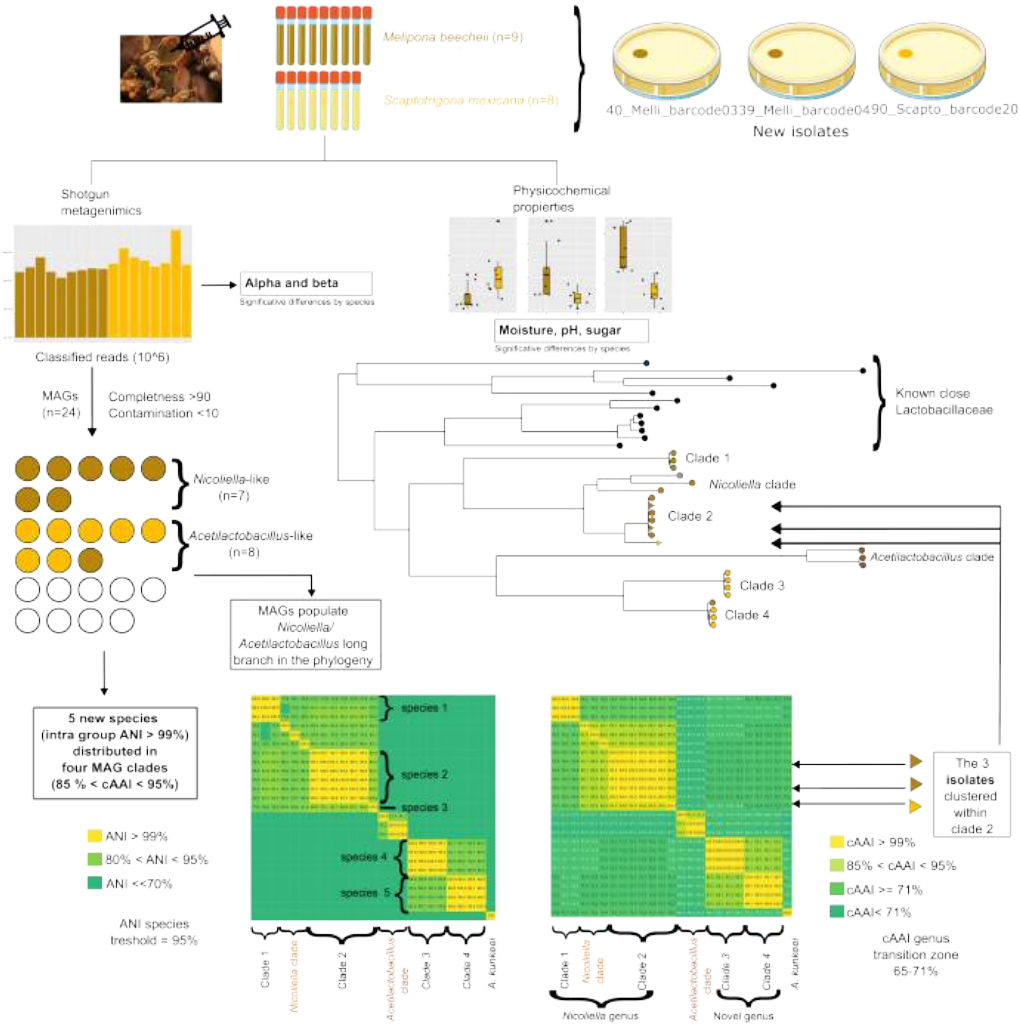

Fig. 0. Genome-resolved analyses reveal that pot honey is dominated by a phylogenetically cohesive Lactobacillaceae species complex. From 17 samples, 15 MAGs and three isolates were clustered into four previously undescribed clades, forming five ANI-defined species and two cAAI-defined genus-level lineages related to *Nicoliella* and *Acetilactobacillus*. Physicochemical and microbiological properties separate seventeen honey samples.

providing pollination of native crops [13, 14] in tropical regions. Meliponiculture, the keeping of stingless bees, supports local livelihood and safeguards biocultural heritage by integrating traditional ecological knowledge, biodiversity conservation, and cultural identity.

The therapeutic effects of honey [15], and particularly of pot honey [5] have been attributed to its diverse bioactive compounds [16]. However, organisms isolated from pot honey also exhibit antimicrobial properties [17, 18, 19], making their identification particularly relevant. Metagenomics and metabarcoding have transformed bacterial ecology by enabling the characterization of microbial communities and the discovery of uncultivated species [20, 21]. In food-associated microbiomes shotgun metagenomics enables the reconstruction of metagenome assembled genomes (MAGs) providing insights into community composition [22, 23] and metabolic potential [24]. Although, the honey microbiome has recently gained attention, shotgun metagenomic studies have so far focused exclusively on *Apis mellifera* [25, 26, 27, 28]. In contrast, the microbiota of stingless bees and their honey (pot honey) has been examined only trough microbial isolation [29, 30] and 16S metabarcoding [31]. Despite the advantages of shotgun metagenomics to asses genomic variability [32], and trace honey origin and authenticity [28], no such studies have been conducted on pot honey, leaving its microbial diversity largely unexplored.

Members of the Lactobacillaceae family are common components of both *A. mellifera* [33] and Meliponini [34] bee microbiome, which has often been associated with bee health [35, 36, 30]. Lactobacillaceae also colonize honey, either produced by Meliponini [37] or by *A. mellifera* [25, 26, 28, 38]. Several Lactobacillaceae bacterial species have been identified in the honey and honey-crop through isolate-based 16S sequencing [30]. In pot honey, isolation [? 19] and a few metabarcoding studies [31] have shown the presence of two sister Lactobacillaceae genera: *Nicoliella* [? ?] and *Acetilactobacillus* [31]. *Nicoliella spurreliana* was isolated from the honey and homogenate of Australian stingless bee *Tetragonula carbonaria* [?], the other species in the genus is *Nicoliella lavandulae* isolated from lavandula [?]. *Acetilactobacillus jinshanensis* is the only known species [39] of its genus and was first isolated and consistently found

in solid vinegar [40, 23, 41] having a recognized role in fermentation [42]. The identification of reported close relatives to *A. jinshanensis* such as *Lactobacillus sp.* Sy-1 isolated from Malaysian pot honey [43]—and the reconstruction of a MAG corresponding to *A. jinshanensis brasiliensis* in stingless bee larval food [44], suggest that the presence of a closely related *Acetilactobacillus*-like lineages ubiquitous to stingless bee-associated environments. However, assessing the microbial composition and its abundance distribution of pot honey remains largely unexplored.

To address this knowledge gap and to characterize the bacterial diversity we applied shotgun metagenomics in honey produced by two stingless Mexican bee species: *Melipona beecheii* and *Scaptotrigona mexicana* (Fig. 1). In seventeen pot honey samples our analyses revealed abundant Lactobacillaceae distributed in four clades with intraclade fastANI >99%. Our reconstructed MAGs share a maximum of 81% ANI with known public genomes, indicating a previously uncharacterized diversity of *Acetilactobacillus*-like species in these honey. Moreover, three strains isolated and sequenced from the same samples showed ANIs ranging from 84% to 99% to our MAGs confirming the presence of these previously uncharacterized microorganisms.

Data summary

Metagenomes and genomes available in SRA BioProject PRJNA1235305. MAGs deposited on Zenodo:`https://zenodo.org/records/18446951`. Phylogenetic analysis available in `https://microreact.org/project/8oNqsaGQm5fhFNgAiNobAe-honey-tree-v2`. Code available at: `https://github.com/HaydeePeruyero/honey_diversity`. Authors confirm that all supporting data, code, and protocols have been provided in the article or through supplementary data files.

Methods

FIELD SAMPLING AND SAMPLE PROCESSING

Honey samples. During the 2022 harvest we selected meliponaries exploiting *M. beecheii* in the Yucatán Peninsula and *S. mexicana* in Veracruz (Table S1, Fig. 1a). Samples comprised several beekeeping methods: *S. mexicana* honey was harvested using modern boxes, blended honey, and traditional pitchers (Fig. 1b); *M. beecheii* using boxes, blends, and log hives (Fig. 1c). Meliponaries relied exclusively on their own resources during the previous year, with no supplementation of sugar or *A. mellifera* honey. Honey was aseptically extracted directly from closed pots using sterile syringes under strict hygiene conditions. For each sample, honey from one to three randomly selected colonies was pooled to a final volume of 320mL, divided into eight 40mL aliquots in sterile 50mL tubes, transported on ice at $10-15°C$, and stored at $-40°C$ at the CIDAS-QROO-UIMQROO. All samples were used in subsequent analyses.

Physicochemical Analyses. Parameters including moisture, pH, total sugars, electrical conductivity, ash content, color (Pfund scale), and hydroxymethylfurfural (HMF) were measured as described by Xolalpa [45], methodological details are provided in the supplementary material (Table S2).

Statistical analyses were performed using a significance level of α=0.05 on 17 samples (9 *S. mexicana*, 8 *M. beecheii*). Normality (Shapiro–Wilk test and residual inspection) and homogeneity of variances (Levene's and Bartlett's tests) were assessed prior to each analysis. When assumptions were met, one-way ANOVA with Tukey's HSD was used; otherwise, Kruskal–Wallis followed by Dunn–Bonferroni and Wilcoxon–Holm pairwise tests were applied, (Table S3). All analyses were conducted in R.

DNA EXTRACTION AND PURIFICATION

Metagenomic DNA extraction. For each sample, 40g were pretreated with TissueLyser, following standard honey extraction protocols [46] [47]; Extensive details are provided in the supplementary material (Table S4). Subsequently, extraction was performed using the DNeasy Blood & Tissue™ Kit (Qiagen), following the manufacturer's instructions. DNA yields exceeded 500ng, with concentrations >20ng/μL and 260/280 ratios >1.7, indicating high purity (QIAxpert, Qiagen).

Shotgun Metagenomic Sequencing. DNA from seventeen honey samples (Table S1) was dried using an Eppendorf® Vacuum Concentrator Plus (1hr) to remove residual water. Final concentrations ranged from 500–1000ng. Shotgun metagenomic sequencing was conducted at Macrogen, Inc. Libraries were prepared with the TruSeq Nano DNA Kit (Illumina) using 350bp inserts and sequenced on the NovaSeq 6000 platform, yielding 150bp paired-end reads.

Microbial isolation and extraction Bacterial were isolated from honey samples following a protocol adapted from *Lactobacillus sp.* Sy-1 isolation [43] (Table S5). Isolates were preserved in MRS broth containing 20% (v/v) glycerol at $-80°C$. For biomass production, frozen stocks were revived on MRS agar and a single colony was inoculated into liquid MRS broth and incubated at 30°C for 24hr prior to cell harvesting and genomic DNA extraction. Genomic DNA was extracted following a phenol–chloroform protocol with minor modifications (Table S5).

Sequencing. For long-read sequencing, libraries were prepared using the Rapid Sequencing Kit V14, SQK-RAD114 Oxford Nanopore Technologies (ONT) according to the manufacturer's instructions. Sequencing was carried out on a MinION device using a 24h sequencing run with default parameters, and data acquisition was performed using the MinKNOW graphical user interface (version 5.3.6).

BIOINFORMATIC ANALYSES

Metagenomes. From quality control to diversity analysis, we followed "The Carpentries Metagenomics Pipeline"[48] modifying the assembly tool. Every metagenome have $>6\times10^7$ raw reads (Table S1). FastQC v0.1 [49] evaluated rawreads quality. Trimmomatic v0.39[50] `SLIDINGWINDOW:25:28`, `MINLEN:35`, `ILLUMINACLIP: TruSeq_LT_CD:2:40:15` removed low-quality nucleotides and adapters. Trimmed reads were assembled with MEGAHIT v1.2.9[51] (30 threads). For taxonomy classification, we used kraken2 v2.1.2 [52], and database PlusPFP (`https://benlangmead.github.io/aws-indexes/k2`) [53]. Results integrated in a biom file with Kraken-biom v.1.2.0[54]. **Diversity**. Beta diversity (Bray-Curtis distance) was visualized via NMDS, revealing between-sample variation. Alpha diversity, reflecting within-sample richness, was evaluated with Observed richness, Shannon, and Chao1 indices. Shannon accounts for abundance distribution, while Chao1 estimates unseen diversity from rare OTUs. Plots were generated in R v4.4.2 using phyloseq v1.50.0[55].

MAGs. We obtained MAGs with MaxBin v2.2 [56] and assessed their quality with CheckM v1.2.1[57]. All MAGs were

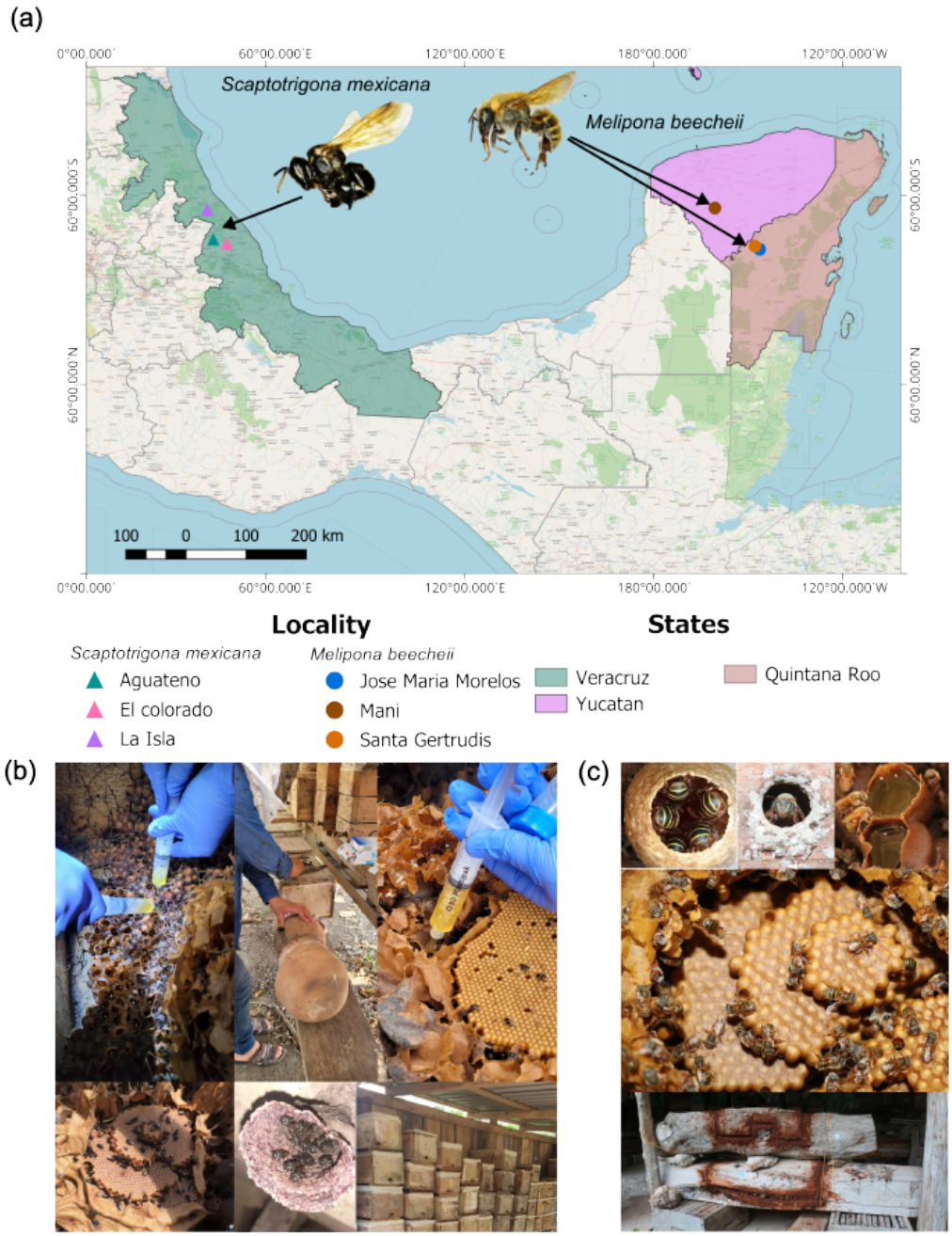

Fig. 1. Honey from *S. mexicana* and *M. beecheii* were collected in Mexico. (a) Samples come from three states and six different localities. (b) Sampling *S. mexicana* honey from different zootechnical managements, including boxes and pitcher hives. (c) *M. beecheii* bees' honey samples come from blended honey, beehive boxes, and traditional "jobones" beehive boxes, and traditional "jobones", nest sections carved or cut from tree trunks, used historically as hives for stingless bees.

refined using GC/coverage profiling (seqkit v2.11.0[?]; and then, clade-specific reference-guided mapping (minimap2 v2.30-r128[?]; bowtie2 v2.5.[?], BAM processing (samtools v1.18[?]) was performed for each clade, and the highest-quality, unrefined MAG served as the reference genome. Finally we applied iterative outlier-based filtering. Only refined MAGs with completness > 90% and contamination < 10% were retained ($n = 24$) (Table 1).

Genome Assembly. Whole-genome sequencing of bacterial isolates was performed to validate MAG reconstruction. For each isolate, ONT adapters were removed from raw reads using Porechop 0.2.4 [?]. NanoStat 1.6.0 [?] assessed read length and quality statistics, while Filtlong 0.2.1 [?] removed the worst 1% of the reads. A de novo assembly approach was applied on genomes with depth >10X. First, Minimap 2.27[?] identified read overlaps (ava-ont preset). Then, Miniasm 0.3 [?] assembled the overlaps and Minipolish 0.2.0 [?] polished

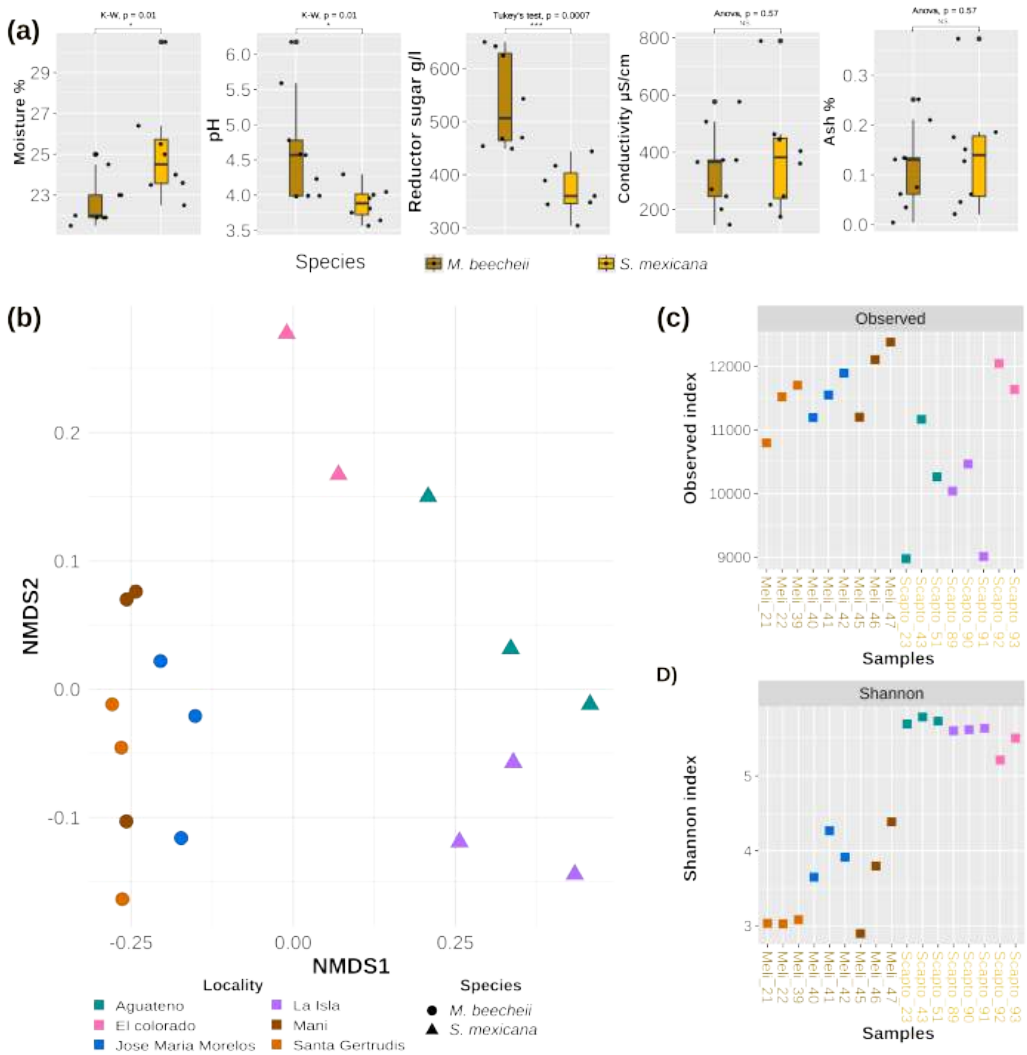

Fig. 2. The diversity of honey's microbial communities is grouped by bee origin. (a) Honey moisture, pH, sugars, conductivity, and ashes are shown in box plots according to species. (b) Species distribution and abundance are linearly separated in beta distance according to bee species. (c) *M. beecheii* has more alpha diversity than *S. mexicana*, measured by the observed index. (d) Contrary to taxonomic units, *S. mexicana* has more alpha diversity when measured by the Shannon Index.

them. Assembly contiguity was assessed with the stats.sh script from BBTools 39.20 [?] while completeness was evaluated with BUSCO 5.5.0 [?] using *Lactobacillales* ODB10 as reference.

Phylogenetic analysis. For each MAG and genome we identified its closest match in the Genome Taxonomy Database (GTDB) with GTDB-Tk v2.5.2[58]. MAGs and genomes closest to *Nicoliella* or *Acetilactobacillus* were considered in posterior analysis. The core genome was calculated using Anvi'o v7[59] for MAGs (n=15), genomes from isolates (n=3), and the following public genomes: *Acetilactobacillus* (n=3), *Nicoliella* (n=2), *Lactobacillus sp.* Sy-1, and the Lactobacillaceae genomes closest to *Acetilactobacillus* according to the 2020 family reorganization (n=11) [39]. *Lentilactobacillus* senioris DSM24302 was included here as an outgroup, in accordance with its position on the 2020 reorganization phylogeny (Table S6). A **phylogenetic tree** was reconstructed using 129 core single-copy protein families obtained by setting (Anvio [59]), `anvi-gen-phylogenomic-tree`, specifically `anvi-get-sequences-for-gene-clusters` with `1=min=max` number of genes per genome). Sequences were aligned using MUSCLE v3.8.15 [?]. Evolutionary model was selected (`-m MFP`) in IQ-TREE v2.1.4-beta [60] with LG+F+R4 identified as the best-fit model according to the BIC criterion. Ultrafast Bootstrap support was applied with 1000 replicates (`-bb 1000`), and branch support was further evaluated using approximate likelihood ratio test (`-alrt 1000`).

Identity indexes. For each pair of MAGs and genomes in the pangenome FastANI v1.34 [61] computed ANI. Average Aminoacid identity (AAI) was calculated pairwise by EzAAI [?], while conserved Average Aminoacid Identity (cAAI) was approximated by calculating the aminoacid identity over the 129 conserved single copy protein families previously described. Results were visualized in a heatmap using Python v3.9.19 packages Seaborn v0.13.2 and Matplotlib v3.9.0.

Metapangenome. A metapangenome uses competitive mapping to align metagenomic reads against a set of genomes [62]; [63]. All *Acetilactobacillus*-like MAGs (n=8,

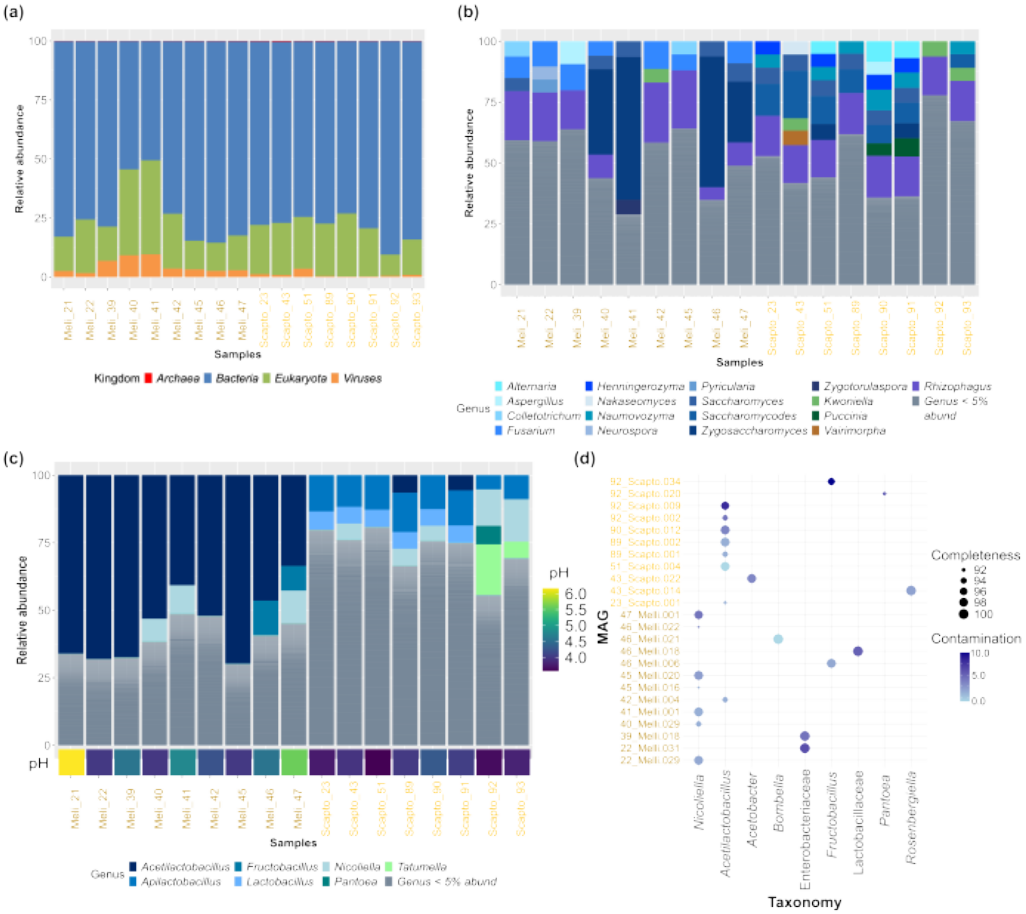

Fig. 3. *Acetillactobacillus*-like lineage is consistently found in honey microbial communities. (a) Most classified reads are from the kingdoms of bacteria, followed by Eukaryota, and viruses. (b) Distribution of fungal genera, whereby those above 5% belonging to phyla Ascomycota are shown in blue, those belonging to *Basidiomycota* in green, those belonging to *Microsporidia* in brown, and those corresponding to *Mucoromycota* in purple. (c) Bacterial reads above 5% comprise the genera *Acetilactobacillus*, *Apilactobacillus*, *Fructobacillus*, *Lactobacillus*, and *Nicoliella*, from the family Lactobacillaceae within the Bacillota phylum (blue), and genera *Tatumella* and *Pantoea* within the *Pseudomonadota phylum* (green). (d) Twenty-four MAGs were recovered with more than 90% of completeness and less than 10% of contamination. Its closest genus, according to the GTDB, is on the x-axis, *Acetilactobacillus*

1 from *M. beecheii* and 7 *S. mexicana*), all *Nicoliella*-like MAGs (n=7 from *M. beecheii*), our three isolates, available *Nicoliella* genomes (n=2), *Lactobacillus* sp. Sy-1, available *A. jinshanensis* genomes (n=3), and one *A. kunkeei*, were included to create an Anvi'o v.8 [59] genome database. Open reading frames (ORFs) were identified via Prodigal v2.6.3 [64]. Gene families were calculated from the genome database using `anvi-pan-genome -ncbi-blast --mcl-inflation=10`. Honey metagenomes were mapped to MAGs and genomes with bowtie2 [65] with default parameters (`--sensitive`). Functional annotation across genomes was performed using ORFs and by annotating ORFs using BLASTp searches against the NCBI COG 2024 database [?]. Subsequently, the program `anvi-display-functions` identified differentially occurring functions across all genomes based on the COG24 PATHWAY annotation. To reduce noise and focus on conserved functional signals, only functions present in at least three genomes were retained (`--min-occurrence 3`).

Results

Honey from *M. beecheii* and *S. mexicana* exhibited distinct physicochemical properties and microbial communities.

Samples were grouped by bee species to compare their physicochemical and microbiological profiles. Significant differences were found in: pH, moisture and sugar content, (Fig. 2a). *M. beecheii* honey had an average pH of 4.65 ± 0.76, while *S. mexicana* averaged 3.88 ± 0.23 (Kruskal-Wallis test, K-W p-value=0.01), both within the reported range of $3.5 - 5.5$ for pot honey [66]. Moisture content was higher in *S. mexicana* $25.12 \pm 2.5\%$ than in *M. beecheii* $22.75 \pm 1.25\%$ (K-W p-value=0.01) while reducing sugar concentration was lower in *S. mexicana* 372.28 ± 47.65g/L compared to *M. beecheii* 537.58 ± 8.86g/L (Tukey's test, p-value=0.0007) (Table S2). No significant differences were found in electrical conductivity, ash content, color, or hydroxymethylfurfural levels.

Beta diversity separated samples by bee species (Fig. 2b), with tighter clustering in *M. beecheii* than in *S. mexicana*, (PERMANOVA pseudo-F = 42.362, p-value=0.0009; 1,000 permutations), indicating a strong species effect on honey

Table 1. Genomic features of MAGs with completeness > 90% contamination < 10% and all isolates.

Genome ID	Group	Completeness	Contamination	Size (Mbp)	ANI to Closest known genome	Closest genus in GTDB	Max AAI to known genomes (*Nicoliella* or *Acetilactobacillus*)	Max core-gene AAI(%, *Nicoliella*)	Core positions compared
40_Melli.029*	1	93.1	1.72	1.4	78.14	*Nicoliella*	68.75	76.49	38027
45_Melli.016*	1	91.38	1.72	1.7	78.20	*Nicoliella*	69.16	76.45	38303
46_Melli.022*	1	91.38	3.45	1.5	78.13	*Nicoliella*	69.19	76.45	38303
22_Melli.029*	2	98.28	2.04	1.9	80.48	*Nicoliella*	80.13	89.29	38289
40_Melli_barcode03	2	97	n.a.	1.9	80.78	*Nicoliella*	80.19	89.25	38300
45_Melli.020*	2	98.28	2.51	2.1	80.35	*Nicoliella*	80.18	89.27	38251
47_Melli.001*	2	97.41	4.31	2.1	80.18	*Nicoliella*	80.18	89.27	38252
39_Melli_barcode04	2	96	n.a.	1.9	80.65	*Nicoliella*	80.18	89.24	38243
41_Melli.001*	2	98.28	1.72	1.8	80.39	*Nicoliella*	80.39	89.30	38269
90_Scapto_barcode20	2	87	n.a.	1.5	81.37	*Nicoliella*	80.64	88.90	37902
51_Scapto.004*	3	98.28	0	1.3	< 77	*Acetilactobacillus*	66.10	71.12	38201
89_Scapto.002*	3	98.28	1.72	1.2	< 77	*Acetilactobacillus*	66.08	71.11	38201
90_Scapto.012*	3	98.28	3.61	1.2	< 77	*Acetilactobacillus*	66.08	71.11	38201
92_Scapto.009	3	96.55	7.68	1.1	< 77	*Acetilactobacillus*	66.18	71.13	38151
42_Melli.004*	4	93.1	1.72	1.2	< 77	*Acetilactobacillus*	65.96	71.57	38254
23_Scapto.001*	4	91.54	1.72	1.2	< 77	*Acetilactobacillus*	65.82	71.60	38244
89_Scapto.001*	4	93.26	1.72	1.2	< 77	*Acetilactobacillus*	65.81	71.58	38247
92_Scapto.002*	4	93.1	4.31	1.2	< 77	*Acetilactobacillus*	65.89	71.59	38244
39_Melli.018*	-	99.69	4.36	3.2	n.a.	*Enterobacteriaceae*	-	-	-
43_Scapto.014*	-	99.84	2.3	3.6	n.a.	*Rosenbergiella epipactidis*	-	-	-
43_Scapto.022*	-	98.28	3.45	1.9	n.a.	*Acetobacter*	-	-	-
46_Melli.006*	-	98.28	2.04	1.6	n.a.	*Fructobacillus*	-	-	-
46_Melli.021*	-	100	0	2.3	n.a.	*Bombella*	-	-	-
92_Scapto.020*	-	91.68	4.89	5.0	n.a.	*Pantoea ananatis*	-	-	-
46_Melli.018	-	100	5.17	2.8	n.a.	Lactobacillaceae	-	-	-
92_Scapto.034	-	94.83	9.25	1.3	n.a.	*Fructobacillus*	-	-	-
22_Melli.031	-	99.69	3.03	2.9	n.a.	*Enterobacteriaceae*	-	-	-

* Indicates a high quality MAGs with < 5% of contamination and > 90% of completeness.

Note. Completness of isolates was estimated with BUSCO.

Note. ANI, AAI, and cAAI have as maximum match *Nicoliella*, and this are the values reported.

Note. Values reported as < 77 in ANI correspond to values below the reporting threshold of FastANI.

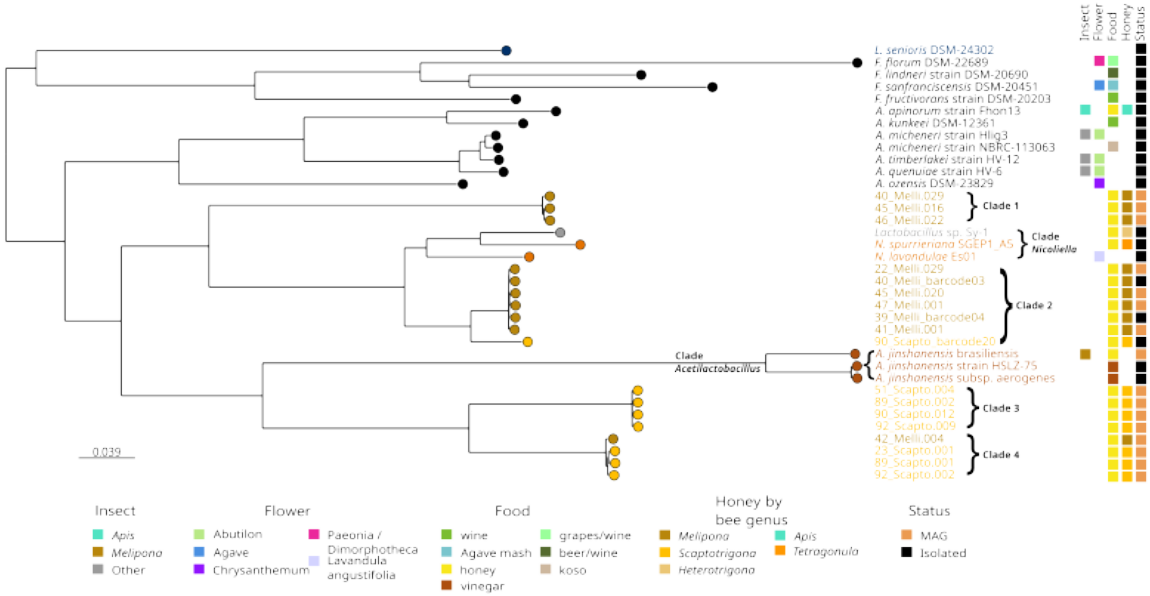

Fig. 4. Phylogenomic tree based on 129 core proteins showing four clades of honey-associated Lactobacillaceae MAGs and isolates related to *Nicoliella* and *Acetilactobacillus*. MAGs cluster with pot honey–derived isolates from multiple geographic regions. Colors denote isolation source, honey-producing bee genera, and genome status (MAG or isolate). The phylogeny was reconstructed from an alignment of 40,955 amino acid positions using the LG+F+R4 substitution model. Branch support was assessed using ultrafast bootstrap with 1,000 replicates and the approximate likelihood ratio test

microbiomes. In contrast, zootechnical practices did not affect clustering (Fig. S1a). Bee species and geography are confounded, as *M. beecheii* samples came from Yucatán and Quintana Roo and *S. mexicana* from Veracruz; thus, microbiological differences may also reflect variation in floral resources or honey physicochemical properties. Alpha diversity differed significantly between bee species with *M. beecheii* honey showing higher richness as measured by observed diversity (Welch t-test, p-value=0.025, Fig. 2c) and Chao1 (ANOVA with post hoc Tukey, p-value=0.0015, Fig. S1b) while *S. mexicana* showed higher evenness based on Shannon index (K–W, p-value=0.005, Fig. 2d). These patterns reflect differences in community evenness, with *S. mexicana* harboring more evenly distributed microbial communities and *M. beecheii* often being dominated by a single lineage.

Previously undescribed *Nicoliella*/*Acetilactobacillus*-like Lactobacillaceae are consistently found in honey microbiome.

Reads were classified using Kraken2, with $\sim 10^6$ reads assigned per sample (Fig. S2). Of these, 76.51% were bacterial, 20.52% eukaryotic, 2.82% viral, and 0.14% archaeal (Fig. 3a). Bacterial reads were dominated by Bacillota and Pseudomonadota (Fig. 3a). Fungi represented a minor fraction of the community, averaging 0.46% of reads (Fig. S4), and were mainly Ascomycota followed by Mucoromycota, Basidiomycota, and Microsporidia (Fig. 3b). Bacterial communities dominated the honey microbiome, with Lactobacillaceae as the most abundant family. Genera exceeding 5% relative abundance included *Acetilactobacillus*, *Apilactobacillus*, *Fructobacillus*, *Lactobacillus*, and *Nicoliella* (Bacillota), and *Pantoea* and *Tatumella* (Pseudomonadota) (Fig. 3c). Kraken-based taxonomic assignment indicated that *M. beecheii* honey was dominated by *Acetilactobacillus* (52.72%), whereas *S. mexicana* honey contained higher proportions of *Lactobacillus* (4.75%) and *Apilactobacillus* (9.77%) which has been found in flowers, bees, and hives [67].

To improve species-level resolution, metagenomic bins were refined into MAGs and filtered to retain genomes with >90% completeness and <10% contamination, yielding 24 MAGs across both stingless bee species, of which 21 met MIMAG high-quality criteria (>90% completeness and <5% contamination) [**?**]. GTDB-based taxonomic assessment affiliated these MAGs with *Nicoliella*, *Acetilactobacillus*, *Rosenbergiella*, *Fructobacillus*, *Bombella*, *Pantoea*, and *Acetobacter* (Fig. 3d). In particular, the sister genera *Nicoliella* (n=7) and *Acetilactobacillus* (n=8) were assigned by GTDB as the closest relatives of fifteen of these MAGs; however, GTDB did not report ANI to any named species (Table1), reflecting the absence of sufficiently close reference genomes. Phylogenomic reconstruction and ANI clustering further resolved these genomes into four distinct, likely undescribed species-level clades (Fig. 5).

Isolation of two *Nicoliella*-like lineages consistent with distinct ANI-based species.

To provide experimental evidence for the presence in the honey of *Nicoliella*/*Acetilactobacillus*-like taxa, bacterial strains were isolated by picking individual colonies from dilution plates. To capture MAG diversity, colonies were isolated from samples yielding MAGs from each of the four clades: 40_Melli (clades 1–2), 90_Scapto (clade 3), and 89_Scapto (clade 4), as well as from sample 39_Melli, which yielded no MAGs. Isolates were sequenced using long-read Nanopore technology, and only datasets with >15× coverage were retained. Strains 40_Melli_barcode03 and 39_Melli_barcode04 clustered within MAG clade 2 (∼99% ANI), while 90_Scapto_barcode20 showed its closest affiliation to

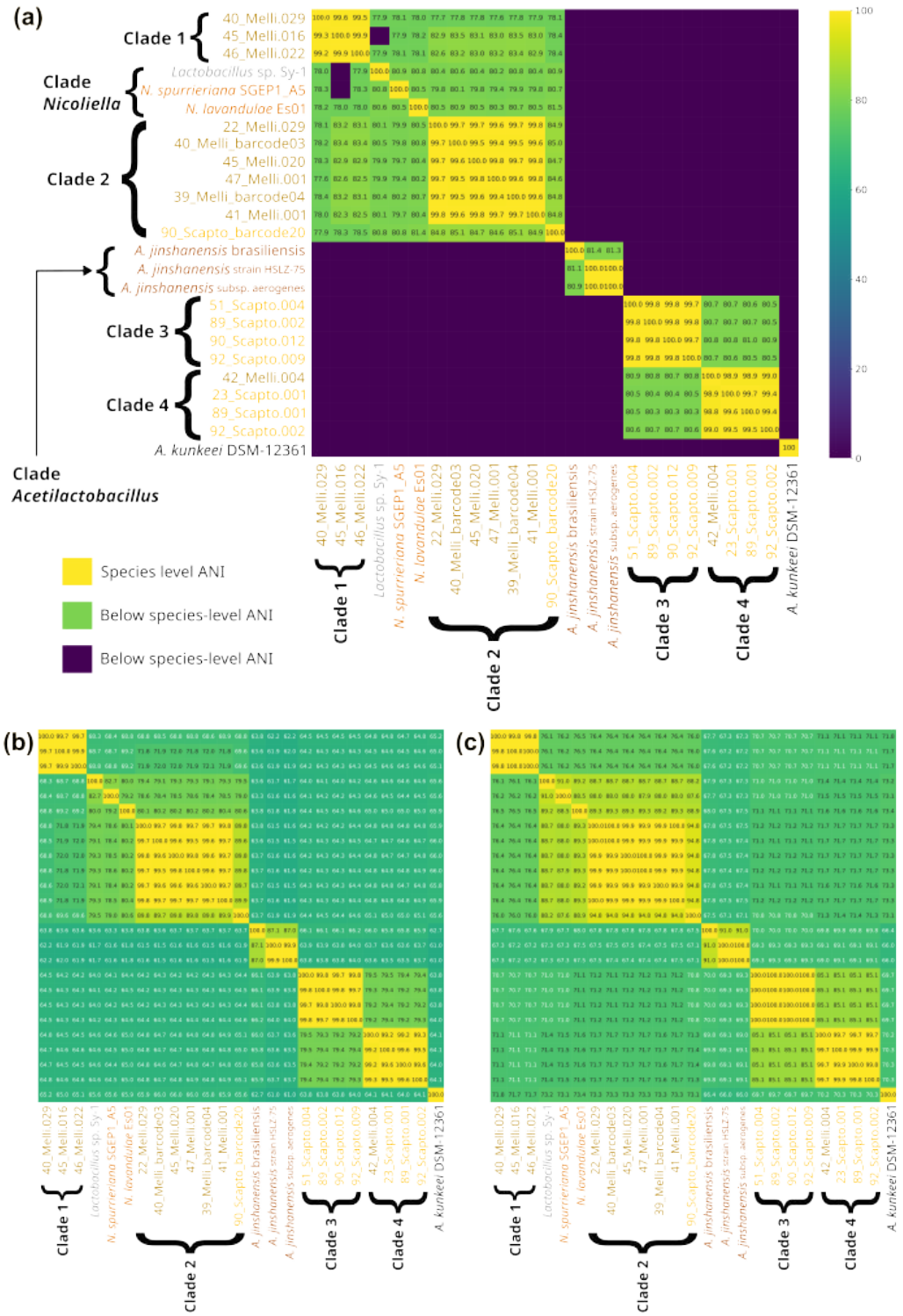

Fig. 5. MAGs and isolates are clustered in four clades, spaning 5 species and 2 genus. **(a)** ANI values above 98% (highlighted in yellow) indicate species-level similarity, while interclade ANI values below ∼ 81% (green) reflect deep genomic divergence. MAGs and isolates cluster into four major clades, each showing high intraclade ANI (> 99%), except for Clade 2, where the inclusion of isolate 90_Scapto_barcode20 results in lower intraclade ANI (∼ 96%). All MAGs exhibit ANI values < 81% relative to described Lactobacillaceae species, supporting their separation into multiple species-level lineages. **(b)** Clades 1, 3, and 4 display AAI values within the genus-level transition zone (≈ 65–71%) relative to *Nicoliella*, indicating substantial evolutionary divergence, whereas Clade 2 shows higher similarity. **(c)** Clades 3, and 4 fall below the genus-level transition range when compared to *Nicoliella* and *Acetilactobacillus* (65–71% cAAI). Together with phylogenomic and ecological evidence, cAAI supports Clades 3 and 4 as candidates for a previously uncharacterized genus.

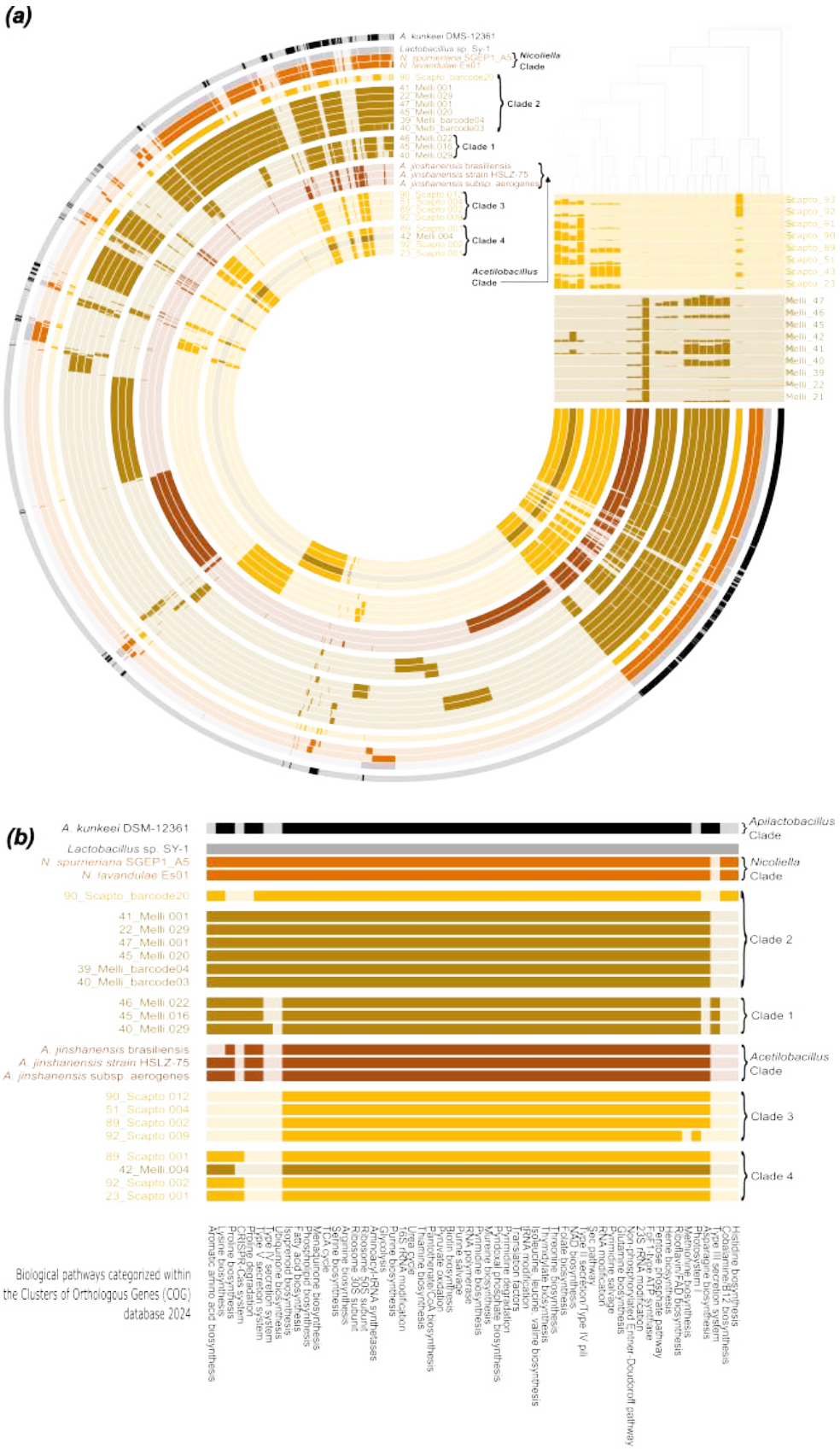

Fig. 6. Metapangenome and functional analyses reveal clade-specific population structure and functions. (a) Presence–absence patterns of gene families cluster MAGs and isolates into four clades consistent with ANI and phylogeny, and distinct from *Nicoliella* and *Acetilactobacillus*. Competitive read recruitment highlights clade-specific mapping patterns across metagenomic samples. (b) Functional profiling based on COG categories shows clade-specific differences in amino acid biosynthesis and secretion systems.

clade 2 at lower similarity (85% ANI), (Table 1). These results demonstrate a close genomic correspondence between MAGs and cultured isolates, supporting the presence of two *Nicoliella*-like lineages consistent with distinct ANI-based species previously unclassified in the honey microbiome.

Phylogeny resolves MAGs into four ANI-defined, previously uncharacterized species-level clades

To assess the phylogenetic placement of MAGs and isolates, we assembled a genome dataset comprising the closest relatives of the genera *Nicoliella* and *Acetilactobacillus*. We retrieved all *Acetilactobacillus* closest relatives (n=11) following the reclassification of Lactobacillaceae[39], we included as outgroup *Lentilactobacillus senioris* (Table S6). Public genomes included vinegar isolates *A. jinshanensis* HSLZ-75 [41] and *A. jinshanensis aerogenes*[68]; as well as *A. jinshanensis* brasiliensis, a MAG from stingless bee larval food [44]. We further added pot honey derived isolates related to *Acetilactobacillus*: *Lactobacillus* sp. *Sy-1* from Malaysian *Heterotrigona itama* honey [43], and *Nicoliella spurreliana* SGEP1_A5 [**?**] from Australian honey. In addition, we included *Nicoliella lavandulae* Es01 [**?**], a flower-derived isolate. Using this dataset, we reconstructed a multilocus phylogeny based on 129 core protein families (Fig. 4).

Phylogenomic analysis grouped our MAGs and isolates into four clades, separating *Nicoliella* and *Acetilactobacillus* related genomes from other Lactobacillaceae members. *Nicoliella* genomes and *Lactobacillus* sp. *Sy-1* were closer to our MAGs Clade 2, while Clades 3 and 4 were closer to *Acetilactobacillus* clade, with all *Acetilactobacillus* forming a distinct lineage. A pair of genomes with ANI>95% are considered the same species;[39, 61], here, ANI values above 98% are shown in yellow (Fig. 5). All MAGs showed ANI<81% to any described Lactobacillaceae species, intraclade ANI exceeded 99% except in Clade 2 with 96% ANI due to isolate 90_Scapto_barcode20 (Table S7), while interclade ANI values were <81% (Table S8). Excluding 90_Scapto_barcode20 isolate from Clade 2 increased intraclade ANI to >99% and reduced variance to $\leq$0.4%, resolving the genomes generated in this study into five species-level lineages, one represented solely by isolate 90_Scapto_barcode20. Mean intra-clade ANI values for *Nicoliella* (87.11 $\pm$ 9.11%) and *Acetilactobacillus* (91.62 $\pm$ 9.34%) clades were below the 95% species threshold, indicating the presence of multiple species in both clades (Fig. 6). Among the three available *A. jinshanensis* genomes, the *brasiliensis* strain is the most phylogenetically divergent, with intra-clade ANI values below 81%, suggesting that its taxonomic placement may warrant reassessment.

Genomic, ecological, and phylogenetic evidence for genus-level divergence.

Criteria for defining novel genera in the Lactobacillaceae include: monophyletic clades, higher intra-than inter-genus AAI and cAAI with limited overlap, and shared ecological or physiological traits [39**?**]. All four MAG groups identified here form monophyletic clades in the phylogenomic tree. Clades 1 and 2 share a common ancestor with *Nicoliella*, whereas Clades 3 and 4 form a distinct lineage with a deeper common ancestor related to *Acetilactobacillus*. Clades 3 and 4 do not cluster within *Nicoliella* unless *Acetilactobacillus* is included, indicating substantial evolutionary divergence. Ecological and physiological differences further distinguish the clades. Isolates within clade 2 were obtained under culture conditions used for *Lactobacillus* sp. *Sy-1*, while no isolates were obtained within Clades 3 and 4. This suggests distinct physiological requirements consistent with clades 3 and 4 affinity to *Acetilactobacillus*. Genome size differences support clades separation, as Clades 3 and 4 have smaller genomes ($\sim$1.2Mbp) than clades 1 (1.4–1.6Mbp), clade 2 (1.5–2.0Mbp), *Nicoliella* ($\sim$2.0Mbp), and *Acetilactobacillus* ($\sim$1.6Mbp) (Table 1).

In the Lactobacillaceae, inter- and intragenus cAAI values are typically separated by a transition zone of approximately 65–71% [**?**]. AAI values for Clades 1, 3, and 4 against their closest described genus, *Nicoliella* fall within this range, suggesting possible genus-level divergence. Inter clade values showed that Clade 1 (AAI= 68.83 $\pm$ 0.27%, cAAI= 76.31 $\pm$ 0.15%), Clade 3 (AAI= 64.35 $\pm$ 0.13%, cAAI= 71.06 $\pm$ 0.05%), and Clade 4 (AAI= 64.74 $\pm$ 0.26%, cAAI= 71.53 $\pm$ 0.05%) lie at the genus transition zone (Table S8, Fig. 5b). In contrast, Clade 2 displayed higher similarity to *Nicoliella* (AAI= 79.42$\pm$ 0.86%, cAAI= 88.57 $\pm$ 0.66%) but remained clearly a distinct clade. Intraclade AAI and cAAI values were consistently high ($\approx$97–99%), indicating strong internal coherence (Table S6, Fig. 5c).

Overall, Clades 1 and 2 are consistent with divergent species within a *Nicoliella*-related lineage, based on shared ancestry, high inter-clade cAAI (76–88%) against Nicoliella, and similar cultivation requirements. In contrast, Clades 3 and 4 show deeper phylogenetic divergence, lower AAI and cAAI values relative to *Nicoliella* ($\approx$64%; $\approx$71% cAAI), smaller genomes, and distinct ecological signatures. Together, these features support Clades 3 and 4 as candidates for a previously uncharacterized genus within the Lactobacillaceae, pending formal taxonomic validation.

Metapangenome reinforces ANI- and phylogeny-defined clades through population structure and functional differentiation.

To evaluate the population structure of *Nicoliella*/ *Acetilactobacillus*-like MAGs and isolates in metagenomic reads, we constructed a metapangenome [62, 63] integrating read recruitment and gene family diversity (Fig. 6). The pangenome resolved core and variable protein families across genomes, and although the topology of the presence–absence-based cladogram differed from the phylogeny, MAGs and isolates largely clustered into the same clades, except for 90_Scapto_barcode20, which grouped with *Nicoliella* genomes and *Lactobacillus sp.* Sy-1 instead of Clade 2. Reads from *M. beecheii* honey mapped mainly to MAGs in clades 1–2, with minimal recruitment to clade 4 and only one MAG in clade 3 (42_Melli), whereas *S. mexicana* honey mapped primarily to genomes in clade 4 and secondarily to the ones in clade 3. Among *A. jinshanensis* genomes, only *A. jinshanensis* brasiliensis recruited reads from *M. beecheii*, while other *A. jinshanensis* genomes showed little or no recruitment.

The main functional differences among clades involve amino acid biosynthesis and secretion systems. *Nicoliella* genomes encode complete pathways for the biosynthesis of histidine, asparagine, proline, lysine, and aromatic amino acids; in contrast, all MAGs and isolates lack histidine patwhay except for barcode20. Clade 1 additionally lacks asparagine patwhay, whereas Clade 3 and *A. jinshanensis brasiliensis* lack lysine and aromatic amino acid patwhay, with Clade 3 also mising proline patwhay. Ubiquinone degradation is absent

in *Acetilactobacillus* and in Clades 1, 3, and 4. Regarding secretion systems, Clade 1 is the only group encoding a Type III, while Type V secretion systems are present in *Nicoliella*, *Acetilactobacillus*, and Clades 1 and 2; Clades 3 and 4 lack detectable secretion systems (Fig. 6).

Discussion

In contrast to the well-studied honey microbiome of *Apis mellifera* [28] pot honey has been examined mainly through 16S ampicillins [31] and isolates [? 43] Using 17 shotgun metagenomes and three isolates, we showed that pot honey microbiomes are dominated by Lactobacillaceae [69] [25, 26, 28, 31]. Although Kraken2 classified most reads as *Acetilactobacillus*, genome-resolved analyses indicated that these reads represent previously uncharacterized *Acetilactobacillus*-like lineages.

Shotgun metagenomics yielded 14 high-quality MAGs (≥90% completeness and ≤5% contamination) closely related to the sister genera *Nicoliella* (n=7) and *Acetilactobacillus* (n=7) but low ANI values to any named species (≤81%). Phylogenomic analysis based on 129 conserved protein families resolved MAGs and isolates into four monophyletic clades, two near *Nicoliella* and two near *Acetilactobacillus*. ANI-based clustering identified five species-level groups, with Clade 2 divided into two species: the first integrated by four MAGs and two isolates (intra clade ANI >99%) and the second represented by a single isolate, with ANI <84% to any genome. Together with Clade 1, these two species from Clade 2 represent three previously undescribed species-level groups closely related to the *Nicoliella* clade.

The 2020 revision of the former *Lactobacillus* lineage split it into 23 genera within the *Lactobacillaceae* [39]; *Nicoliella*, described in 2023 from pot honey, was therefore not included. In our dataset, the 3 species level groups in Clades 1 and 2 fall within the *Nicoliella* genus-level AAI/cAAI transition zone, consistent with possible affiliation with *Nicoliella* genus. In contrast, the 2020 revision included *Acetilactobacillus* as a genus represented by a single species, *A. jinshanensis*, characterized by a long evolutionary branch and sparse genomic representation; currently, only three genomes are available—two vinegar isolates and one MAG from stingless bee larval food. In our study, MAGs clades 3 and 4 clustered adjacent to the *Acetilactobacillus* clade. These *Acetilactobacillus*-like MAGs showed low AAI/cAAI values relative to available *A. jinshanens*is genomes supporting the interpretation that these MAGs represent a previously undescribed genus-level lineage and expanding the currently underrepresented *Acetilactobacillus*-related diversity. In the metapangenome reads from *M. beecheii* honey mapped preferentially to the *brasiliensis* genome and only marginally to other *A. jinshanensis* genomes, supporting the idea that honey harbors *Acetilactobacillus*-like lineages distinct from those associated with vinegar. This pattern suggests that previous reports of *Acetilactobacillus* in honey may reflect limitations of available reference genomes rather than true taxonomic identity.

Together, the recovery of three *Nicoliella*-related isolates and four closely related yet clearly distinct MAG clades distributed around *Nicoliella* and *Acetilactobacillus* clades, suggest that the apparent dominance of *Acetilactobacillus* in honey requires revision to account for previously unrecognized species within this lineage. Genome-resolved analyses further reveal that Lactobacillaceae dominance in pot honey reflects a complex population structure composed of multiple, closely related species-level lineages with broad geographic distribution.

Finally, ecological context may help explain the observed patterns. Honey represents a relatively acidic niche, consistent with the known physiology of *Acetilactobacillus*-related taxa. The inability to recover representatives of clades 3 and 4 as isolates may reflect more stringent growth requirements, such as higher acidity or anaerobic conditions, which is consistent with the significantly lower pH observed in honey from *S. mexicana* when compared to *M. becheii*. *A. jinshanensis* was originally isolated from solid-state vinegar, with an optimal pH range of 3.0–5.0 [41], and subsequent strains have been recovered from acidic fermented products, suggesting that related honey-associated lineages may require similarly conditions for cultivation.

Our *Nicoliella*-like MAGs and isolates cluster with honey- and stingless bee–associated isolates from Malaysia and Australia, as well as with a flower-derived isolate from Spain, whereas the *Acetilactobacillus*-like Clades 3 and 4 group with a MAG recovered from stingless bee larval food in Brazil. Functionally, these lineages differ in amino acid biosynthesis and secretion pathways, suggesting varying degrees of ecological specialization and host association. Future work focused on the formal taxonomic and physiological characterization should explore genomic signatures of specialized metabolism, including genes involved in production of antimicrobial compounds.

Taken together, our results indicate that the pot honey microbiome is dominated by a phylogenetically cohesive complex of closely related species with global distribution. From a human consumption perspective, the dominance and consistent presence of these lineages across geographically distant regions suggest a conserved microbial signature that could inform authenticity assessments, complementing physicochemical markers currently used to evaluate pot honey. The combination of phylogenetic coherence, ecological specialization, and broad geographic occurrence underscores their relevance in the honey environment and motivates further functional and taxonomic investigation.

Conclusion

Pot honey is dominated by a group of phylogenetically cohesive Lactobacillaceae species complex with global distribution, uncovered through genome-resolved analyses.

Competing interests

No competing interest is declared.

Author contributions statement

Conceptualization, conception, and design of study, writing, and submission of manuscript A.X.A. and E.J.D.S.; methodology. A.X.A., E.J.D.S., N.S.M and B.E.R.C.; analysis of DNA extraction and preparation for sequencing A.X.A, D.I.H.M, W.I.M.C; C.L.C. and L.R.O. isolate miscroorganisms; Statistical analysis, A.X.A. and H.C.P.; analysis of metagenomic data H.C.P., N.S.M and A.X.A; metapangenome analysis J.A.L.F; physicochemical analyses. A.X.A. and H.C.P.; writing and preparation of the original draft, A.X.A, N.S.M, H.C.P, and E.J.D.S.; writing and review and editing A.X.A, N.S.M, H.C.P,

and E.J.D.S.; Writing of the article with contributions from D.I.H.M., W.I.M.C., B.E.R.C, J.F.R.C., F.B.G., B.E.O.V., H.C.P.; Management and acquisition of financial funds for the project, A.X.A, and E.J.D.S. All authors have read and accepted the published version of the manuscript.

Funding information

Aurora Xolalpa-Aroche received funding from Finca la Isla, the Consejo Quintanarroense de Ciencia y Tecnología (COQCyT) 2021–2023, the UIMQROO, and the Consejo Nacional de Humanidades, Ciencias y Tecnología (CONAHCyT). Haydeé Contreras Peruyero was supported by the UNAM Postdoctoral Program (POSDOC) at the Centro de Ciencias Matemáticas, UNAM, as well as the UNAM-PAPIIT Research Grant IN114323 and the CONAHCyT Postdoctoral Fellowship 2025.

Acknowledgments

Our gratitude to the beekeepers for providing honey samples from their valuable hives and to Emil-Neme and Edith-Pazos for their sampling support. We thank Carlos-Ucan, Omar-Sánchez, Araceli-López, Rubén-Busambra, Cristian-Suárez, and Areli-Chimal from CIDAS-QROO/UIMQROO for their assistance in fieldwork, physicochemical analyses, and DNA extraction, Jorge-Cime and Grecia-Hernández for the technical support at CISEI/INSP. Thanks to Ángel-Chi for his assistance in analyzing territorial data, to Ricardo-Ayala, Ruben-Busambra, Cristian-Suárez, Areli-Chimal, Juan-Mayo, and Omar-Sánchez for their bee photographs to Anton-Pashkov for his bioinformatic support.

References

1. José Luis Domínguez_castro and Roberto Patiño-Ramírez. A brief introduction of melipona honey cultivation in the Mayan culture. In Elizabeth de la Luz Ortiz Vázquez, Jorge Carlos Ruiz Ruiz, Denis Israel Magaña Ortiz, and Jesús Manuel Ramon Sierra, editors, *Stingless Bee's Honey from Yucatan: Culture, Traditional Uses and Nutraceutical Potential*, pages 1–10. Nova Science publishers, 2016.
2. Elizabeth H. Paris, Veronica Briseño Castrejon, Debra S. Walker, and Carlos Peraza Lope. The Origins of Maya Stingless Beekeeping. *Journal of Ethnobiology*, 40(3):386–405, September 2020. Publisher: SAGE Publications.
3. Fausto Cabezas-Mera, Ariana C. Cedeño-Pinargote, Eduardo Tejera, José M. Álvarez Suarez, and António Machado. Antimicrobial activity of stingless bee honey (Tribe: Meliponini) on clinical and foodborne pathogens: A systematic review and meta-analysis. *Food Frontiers*, 5(3):964–993, 2024. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/fft2.386.
4. Irina Villacrés-Granda, Dayana Coello, Adrián Proaño, Isabel Ballesteros, David W. Roubik, Gabriela Jijón, Genoveva Granda-Albuja, Silvana Granda-Albuja, Reinier Abreu-Naranjo, Favian Maza, Eduardo Tejera, Ana M. González-Paramás, Pedro Bullón, and José M. Alvarez-Suarez. Honey quality parameters, chemical composition and antimicrobial activity in twelve Ecuadorian stingless bees (Apidae: Apinae: Meliponini) tested against multiresistant human pathogens. *LWT*, 140:110737, April 2021.
5. Ankush S. Gadge, Dhananjay V. Shirsat, Parakkattu S. Soumia, Chandrashekhar L. Pote, M. Pushpalatha, Trupti Rajesh Pandit, Ram Dutta, Satish Kumar, S. V. Ramesh, Vijay Mahajan, and Vadivelu Karuppaiah. Physiochemical, biological, and therapeutic uses of stingless bee honey. *Frontiers in Sustainable Food Systems*, 7, January 2024. Publisher: Frontiers.
6. Genoveva R. Ocampo Rosales. Medicinal Uses of Melipona beecheii Honey, by the Ancient Maya. In Patricia Vit, Silvia R. M. Pedro, and David Roubik, editors, *Pot-Honey: A legacy of stingless bees*, pages 229–240. Springer, New York, NY, 2013.
7. Jesús Ramón-Sierra, Denis Magaña-Ortiz, Elizabeth Ortiz, and Jorge Ruiz Ruiz. Medicinal uses of Melipona Bee Honey in Mayan Communities of Yucatán. In Elizabeth de la Luz Ortiz Vázquez, Jorge Carlos Ruiz Ruiz, Denis Israel Magaña Ortiz, and Jesús Manuel Ramon Sierra, editors, *Stingless Bee's Honey from Yucatan: Culture, Traditional Uses and Nutraceutical Potential*, pages 11–26. Nova Science publishers, 2016.
8. Tatiana Colombo Pimentel, Michele Rosset, Janaina Maria Batista de Sousa, Louise Iara Gomes de Oliveira, Isis Meireles Mafaldo, Maria Manuela Estevez Pintado, Evandro Leite de Souza, and Marciane Magnani. Stingless bee honey: An overview of health benefits and main market challenges. *Journal of Food Biochemistry*, 46(3):e13883, 2022.
9. Márcia Aparecida Sanches, Ana Maria S Pereira, and José Eduardo Serrão. Pharmacological actions of extracts of propolis of stingless bees (Meliponini). *Journal of Apicultural Research*, 56(1):50–57, jan 2017.
10. José Ek and Miguel Magaña. Características socioeconómicas de la meliponicultura en localidades rurales del estado de Yucatán. *Textual*, pages 205–232, 2023.
11. Rogério Marcos de Oliveira Alves. Production and Marketing of Pot-Honey. In Patricia Vit, Silvia R. M. Pedro, and David Roubik, editors, *Pot-Honey: A legacy of stingless bees*, pages 541–556. Springer, New York, NY, 2013.
12. Leydiane da Conceição Lazarino, Matheus Galvão Brito, Igor Evangelista do Carmo, Paulo Roberto Antunes de Mello Affonso, Raymundo José de Sá Neto, and Ana Maria Waldschmidt. Meliponicultura: potencialidades e limitações para a conservação de abelhas nativas e redução da pobreza. *Diversitas Journal*, 6(2):2217–2236, May 2021. Number: 2.
13. Ricardo Ayala, Victor H Gonzalez, and Michael S Engel. Mexican stingless bees (Hymenoptera: Apidae): diversity, distribution, and indigenous knowledge. In Patricia Vit, Silvia Pedro, and David Roubik, editors, *Pot-honey: a legacy of stingless bees*, pages 135–152. Springer New York, New York, NY, 2013.
14. Ester Judith Slaa, Luis Alejandro Sánchez Chaves, Katia Sampaio Malagodi-Braga, and Frouke Elisabeth Hofstede. Stingless bees in applied pollination: practice and perspectives. *Apidologie*, 37(2):293–315, March 2006. Publisher: EDP Sciences.
15. Liliana Luca, Daniela Pauliuc, and Mircea Oroian. Honey microbiota, methods for determining the microbiological composition and the antimicrobial effect of honey – A review. *Food Chemistry: X*, 23:101524, October 2024.
16. Patricia Vit, Margarita Medina, and María Eunice Enríquez. Quality standards for medicinal uses of

Meliponinae honey in Guatemala, Mexico and Venezuela. *Bee World*, 85(1):2–5, jan 2004.

17. Pham T Lanh, Bui TT Duong, Ha T Thu, Nguyen T Hoa, and Dong Van Quyen. Comprehensive analysis of the microbiome in apis cerana honey highlights honey as a potential source for the isolation of beneficial bacterial strains. *PeerJ*, 12:e17157, 2024.
18. Tobias C Olofsson, Èile Butler, Pawel Markowicz, Christina Lindholm, Lennart Larsson, and Alejandra Vásquez. Lactic acid bacterial symbionts in honeybees–an unknown key to honey's antimicrobial and therapeutic activities. *International Wound Journal*, 13(5):668–679, 2016.
19. Syariffah Nuratiqah Syed Yaacob, Huyop , Fahrul, Kamarulzaman Raja Ibrahim , Raja, , and Roswanira Abdul Wahab. Identification of Lactobacillus spp. and Fructobacillus spp. isolated from fresh Heterotrigona itama honey and their antagonistic activities against clinical pathogenic bacteria. *Journal of Apicultural Research*, 57(3):395–405, May 2018. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00218839.2018.1428047.
20. Cindy J. Castelle and Jillian F. Banfield. Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell*, 172(6):1181–1197, March 2018.
21. Laura A. Hug, Brett J. Baker, Karthik Anantharaman, Christopher T. Brown, Alexander J. Probst, Cindy J. Castelle, Cristina N. Butterfield, Alex W. Hernsdorf, Yuki Amano, Kotaro Ise, Yohey Suzuki, Natasha Dudek, David A. Relman, Kari M. Finstad, Ronald Amundson, Brian C. Thomas, and Jillian F. Banfield. A new view of the tree of life. *Nature Microbiology*, 1(5):1–6, April 2016. Publisher: Nature Publishing Group.
22. Niccolò Carlino, Aitor Blanco-Míguez, Michal Punčochář, Claudia Mengoni, Federica Pinto, Alessia Tatti, Paolo Manghi, Federica Armanini, Michele Avagliano, Coral Barcenilla, Samuel Breselge, Raul Cabrera-Rubio, Inés Calvete-Torre, Mairéad Coakley, José F. Cobo-Díaz, Francesca De Filippis, Hrituraj Dey, John Leech, Eline S. Klaassens, Stephen Knobloch, Dominic O'Neil, Narciso M. Quijada, Carlos Sabater, Sigurlaug Skírnisdóttir, Vincenzo Valentino, Liam Walsh, Avelino Alvarez-Ordóñez, Pablo Alvarez, Livio Antonielli, Elke Arendt, Federica Armanini, Aurelie Aubry, Jacob Baelum, Coral Barcenilla, Alejandro Belanche, Yaiza Benavent-Gil, Tony Blake, Aitor Blanco-Míguez, Radhika Bongoni, Mickael Boyer, Fiona Brennan, Samuel Breselge, Helgi Briem, Derek Butler, Inés Calvete-Torre, Omar Cristobal Carballo, Mireille Cardinal, Niccolò Carlino, Christian Chervaux, Christine Chopin, Natallia Clotaire, Mairead Coakley, José Francisco Cobo-Díaz, Jim Codd, Stephen Conroy, Karla Fabiola Corral-Jara, Karla-Fabiola Corral-Jara, Paul D. Cotter, Gerard Coyne, Chris Creevey, Patricia D. Cuevas, Brendan Curran, Susana Delgado, Liesbeth Derde, Muriel Derrien, Danilo Ercolini, Ruth Gomez Exposito, María Mercedes López Fernández, Francesca De Filippis, Daniel Fordham, Hubert Galy, Asimenia Gavriilidou, Oddur Gunnarsson, Buck Hanson, Gerben Hermes, Rongcai Huang, Sharon Huws, Israel Ikoyi, Alice Jaeger, Ian Jeffery, Marc Jérôme, Pierre-Alexandre Juan, David Kenny, Annelies Kers, Karim-Franck Khinouche, Stuart Kirwan, Eline S. Klaassens, Stephen Knobloch, Kristinn Kolbeinsson, Laetitia Kolypczuk, Tanja Kostic, Fabio Ledda, John Leech, Doerte Lehmann, Françoise Leroi, Eva Lewis, Johanna Ley, Eva Lucic, Kieran Lynch, Sabrina Mace, Iain MacLaren-Lee, Lisa Mahler de Sanchez, Juergen Marchart, Abelardo Margolles, Viggó Thór Marteinsson, Giulia Masetti, Fiona McGovern, Noirin McHugh, Steven McLoughlin, Dara Meehan, Lars Mølbak, Thomas Monin, Javier Moreno, Diego Morgavi, Steven Morrison, Steffen Müench, Ana Rute Ramos Neves, Emma Neylon, Laura Nyhan, Rhona O'Kelly, Dominic O'Neil, Paul O'Toole, Abimael Ortiz-Chura, Juan Manuel Palma, Edoardo Pasolli, Delphine Passerini, Milica Pastar, Federica Pinto, Walter Pirovano, Olga Plans, Marion Policht, Aurel Pop, Bianca Pop, Milka Popova, Miguel Prieto, Narciso M. Quijada, Antje Reiss, Pedro Romero, Patricia Ruas-Madiedo, Francesco Rubino, Raul Cabrera Rubio, Lorena Ruiz, Angela Ryan, Clodagh Ryan, Carlos Sabater, Aylin Sahin, Cecile Salaun, Fernanda Godoy Santos, Carolin Schneider, Nicola Segata, Evelyne Selberherr, Angela Sessitsch, Sigurlaug Skírnisdóttir, Hauke Smidt, Paul Smith, Markus Sprenger-Haussels, Ilma Tapio, Julien Tap, Vincenzo Valentino, Martin Wagner, Aaron Walsh, Liam Walsh, Sinead M. Waters, Spike Willcocks, David R. Yáñez-Ruiz, Tianhai Yan, Min Yap, Emanuele Zannini, Véronique Zuliani, Avelino Alvarez-Ordóñez, Francesco Asnicar, Gloria Fackelmann, Vitor Heidrich, Abelardo Margolles, Viggó Thór Marteinsson, Omar Rota Stabelli, Martin Wagner, Danilo Ercolini, Paul D. Cotter, Nicola Segata, and Edoardo Pasolli. Unexplored microbial diversity from 2,500 food metagenomes and links with the human microbiome. *Cell*, 187(20):5775–5795.e15, October 2024. Publisher: Elsevier.
23. Jie Wu, Ning Zhao, Qin Li, Kui Zhao, Meiling Tu, Jianlong Li, Kaidi Hu, Shujuan Chen, Shuliang Liu, and Aiping Liu. Metagenomics and Metagenome-Assembled Genomes: Analysis of Cupei from Sichuan Baoning Vinegar, One of the Four Traditional Renowned Vinegars in China. 14(3), 2025.
24. Rubing Du, Wu Xiong, Lei Xu, Yan Xu, and Qun Wu. Metagenomics reveals the habitat specificity of biosynthetic potential of secondary metabolites in global food fermentations. *Microbiome*, 11(1):115, May 2023.
25. Samuele Bovo, Anisa Ribani, Valerio Joe Utzeri, Giuseppina Schiavo, Francesca Bertolini, and Luca Fontanesi. Shotgun metagenomics of honey DNA: Evaluation of a methodological approach to describe a multi-kingdom honey bee derived environmental DNA signature. *PLOS ONE*, 13(10):e0205575, oct 2018.
26. Samuele Bovo, Valerio Joe Utzeri, Anisa Ribani, Riccardo Cabbri, and Luca Fontanesi. Shotgun sequencing of honey DNA can describe honey bee derived environmental signatures and the honey bee hologenome complexity. *Scientific Reports*, 10(1):9279, 2020.
27. Julien Crovadore, François Gérard, Romain Chablais, Bastien Cochard, Karl Kristian Bergman Jensen, and François Lefort. Deeper Insight in Beehives: Metagenomes of Royal Jelly, Pollen, and Honey from Lavender, Chestnut, and Fir Honeydew and Epiphytic and Endophytic Microbiota of Lavender and Rose Flowers. *Genome Announcements*, 5(22):10.1128/genomea.00425–17, June 2017. Publisher: American Society for Microbiology.
28. Carlos Sabater, Inés Calvete, Xenia Vázquez, Lorena Ruiz, and Abelardo Margolles. Tracing the origin and authenticity of spanish pdo honey using metagenomics and machine learning. *International Journal of Food Microbiology*, 421:110789, 2024.

29. Amaury Andrade-Velásquez, Humberto Hernández Sánchez, Lidia Dorantes-Álvarez, Beatriz Palmeros-Sánchez, Raymundo Torres-Moreno, Dolores Hernández-Rodríguez, and Guiomar Melgar-Lalanne. Honey characterization and identification of fructophilic lactic acid bacteria of fresh samples from melipona beecheii, scaptotrigona pectoralis, plebeia llorentei, and plebeia jatiformis hives. *Frontiers in Sustainable Food Systems*, 7:1113920, 2023.
30. Alejandra Vásquez, Eva Forsgren, Ingemar Fries, Robert J Paxton, Emilie Flaberg, Laszlo Szekely, and Tobias C Olofsson. Symbionts as major modulators of insect health: lactic acid bacteria and honeybees. *PloS one*, 7(3):e33188, 2012.
31. Alan Emanuel Silva Cerqueira, Helena Santiago Lima, Lívia Carneiro Fidélis Silva, Tomás Gomes Reis Veloso, Sérgio Oliveira de Paula, Weyder Cristiano Santana, and Cynthia Canêdo da Silva. Melipona stingless bees and honey microbiota reveal the diversity, composition, and modes of symbionts transmission. *FEMS Microbiology Ecology*, 100(7):fiae063, jul 2024.
32. Kirsten M. Ellegaard and Philipp Engel. Genomic diversity landscape of the honey bee gut microbiota. *Nature Communications*, 10(1):446, jan 2019.
33. Quinn S. McFrederick, William T. Wcislo, Douglas R. Taylor, Heather D. Ishak, Scot E. Dowd, and Ulrich G. Mueller. Environment or kin: whence do bees obtain acidophilic bacteria? *Molecular Ecology*, 21(7):1754–1768, April 2012.
34. Hauke Koch, Dharam P. Abrol, Jilian Li, and Paul Schmid-Hempel. Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. *Molecular Ecology*, 22(7):2028–2044, 2013. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.12209.
35. Eva Forsgren, Tobias C Olofsson, Alejandra Vásquez, and Ingemar Fries. Novel lactic acid bacteria inhibiting paenibacillus larvae in honey bee larvae. *Apidologie*, 41(1):99–108, 2010.
36. K Hariprasath, S Mohankumar, M Sudha, N Saranya, and VR Saminathan. The role of honeybee gut and honey microbiome in sustainable bee and human health. *Journal of Pure & Applied Microbiology*, 19(1), 2025.
37. Samra Basharat, Tiantian Meng, Lixin Zhai, Asif Hussain, Sahibzada Muhammad Aqeel, Salman Khan, Obaid Ullah Shah, and Xiangru Liao. Bacterial diversity of stingless bee honey in Yunnan, China: isolation and genome sequencing of a novel acid-resistant Lactobacillus pentosus (SYBC-MI) with probiotic and L. tryptophan producing potential via millet fermentation. *Frontiers in Bioengineering and Biotechnology*, 11, December 2023. Publisher: Frontiers.
38. Vladimíra Kňazovická. Diversity of bacteria in Slovak and foreign honey, with assessment of its physico-chemical quality and counts of cultivable microorganisms. *Journal of Microbiology, Biotechnology and Food Sciences*, 9:414–421, nov 2019.
39. Jinshui Zheng, Stijn Wittouck, Elisa Salvetti, Charles M A P Franz, Hugh M B Harris, Paola Mattarelli, Paul W O'Toole, Bruno Pot, Peter Vandamme, Jens Walter, Koichi Watanabe, Sander Wuyts, Giovanna E Felis, Michael G Gänzle, and Sarah Lebeer. 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. *International Journal of Systematic and Evolutionary Microbiology*, 70(4):2782–2858, 2020.
40. Liangqiang Chen, Huizhen Zheng, Keqi Cheng, Chao Li, Xing Qin, Guozheng Wang, Fan Yang, Hai Du, Li Wang, and Yan Xu. Deciphering the acidophilia and acid resistance in Acetilactobacillus jinshanensis dominating baijiu fermentation through multi-omics analysis. *Food Microbiology*, 125:104655, 2025.
41. Yongjian Yu, Xin Li, Junhong Zhang, Li-Juan Chai, Zhen-Ming Lu, and Zheng-Hong Xu. Lactobacillus jinshani sp. nov., isolated from solid-state vinegar culture of Zhenjiang aromatic vinegar. *Antonie van Leeuwenhoek*, 113(1):43–54, 2020.
42. Aiping Liu, Jie Wu, Jianlong Li, Qin Li, Ning Zhao, Kaidi Hu, Shuliang Liu, Giuseppe Blaiotta, and Jiang Zhou. Uncovering the microbial community dynamics and metabolic pathways of primary organic acids in Sichuan Baoning vinegar through metagenomics. *World Journal of Microbiology and Biotechnology*, 41(3):91, March 2025.
43. Syariffah Nuratiqah Syed Yaacob, Fahrul Huyop, Mailin Misson, Roswanira Abdul Wahab, and Nurul Huda. Exploring the genome of Lactobacillaceae spp. Sy-1 isolated from Heterotrigona itama honey. *PeerJ*, 10:e13053, mar 2022.
44. Ana Carolina Costa Santos, Serena Mares Malta, Raquel Cristina Cavalcanti Dantas, Nina Dias Coelho Rocha, Vasco Ariston de Carvalho Azevedo, and Carlos Ueira-Vieira. Antimicrobial activity of supernatants produced by bacteria isolated from Brazilian stingless bee's larval food. *BMC Microbiology*, 22(1):127, 2022.
45. Aurora Xolalpa-Aroche, David I Hernández-Mena, Wilson I Moguel-Chin, Haydeé Contreras-Peruyero, Blanca E Rivero-Cruz, Elizabeth Ortiz-Vázquez, J Fausto Rivero-Cruz, Rodrigo A Velarde M., and Enrique J Delgado-Suárez. Physicochemical Properties of Two Mexican Stingless Bee Honeys to Strengthen Their Biocultural Value. *Revista Brasileira de Farmacognosia*, 34(5):1129–1139, 2024.
46. Sónia Soares, Joana S. Amaral, Maria Beatriz P.P. Oliveira, and Isabel Mafra. Improving dna isolation from honey for the botanical origin identification. *Food Control*, 48:130–136, 2015. Recent Advances of Food Analysis.
47. Hans-Ulrich Waiblinger, Marc Ohmenhaeuser, Stefanie Meissner, Miriam Schillinger, Klaus Pietsch, Ottmar Goerlich, Joachim Mankertz, Kathrin Lieske, and Hermann Broll. In-house and interlaboratory validation of a method for the extraction of DNA from pollen in honey. *Journal für Verbraucherschutz und Lebensmittelsicherheit*, 7(3):243–254, sep 2012.
48. Claudia Zirión-Martínez, Diego Garfias-Gallegos, Tania Vanessa Arellano-Fernandez, Aarón Espinosa-Jaime, Edder D Bustos-Díaz, José Abel Lovaco-Flores, Luis Gerardo Tejero-Gómez, J Abraham Avelar-Rivas, and Nelly Sélem-Mojica. A data carpentry-style metagenomics workshop. *Journal of Open Source Education*, 7(72):209, feb 2024.
49. S. Andrews. Fastqc. a quality control tool for high throughput sequence data. `http://www.bioinformatics.babraham.ac.uk/projects/fastqc/`, 2010. Accessed: 2025-06-05.
50. Anthony M. Bolger, Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, aug 2014.
51. Dinghua Li, Chi-Man Liu, Ruibang Luo, Kunihiko Sadakane, and Tak-Wah Lam. MEGAHIT: an ultra-fast

single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10):1674–1676, may 2015.

52. Derrick E Wood and Steven L Salzberg. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*, 15(3):R46, 2014.
53. Derrick E Wood, Jennifer Lu, and Ben Langmead. Improved metagenomic analysis with Kraken 2. *Genome Biology*, 20(1):257, 2019.
54. S Dabdoub. kraken-biom: Enabling interoperative format conversion for Kraken results (Version 1.2) [Software], 2016.
55. Paul J McMurdie and Susan Holmes. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE*, 8(4):e61217, apr 2013.
56. Yu-Wei Wu, Blake A Simmons, and Steven W Singer. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*, 32(4):605–607, feb 2016.
57. Donovan H. Parks, Michael Imelfort, Connor T. Skennerton, Philip Hugenholtz, and Gene W. Tyson. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, 25(7):1043–1055, jul 2015.
58. Pierre-Alain Chaumeil, Aaron J Mussig, Philip Hugenholtz, and Donovan H Parks. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*, 36(6):1925–1927, mar 2020.
59. A Murat Eren, Evan Kiefl, Alon Shaiber, Iva Veseli, Samuel E Miller, Matthew S Schechter, Isaac Fink, Jessica N Pan, Mahmoud Yousef, Emily C Fogarty, Florian Trigodet, Andrea R Watson, Özcan C Esen, Ryan M Moore, Quentin Clayssen, Michael D Lee, Veronika Kivenson, Elaina D Graham, Bryan D Merrill, Antti Karkman, Daniel Blankenberg, John M Eppley, Andreas Sjödin, Jarrod J Scott, Xabier Vázquez-Campos, Luke J McKay, Elizabeth A McDaniel, Sarah L R Stevens, Rika E Anderson, Jessika Fuessel, Antonio Fernandez-Guerra, Lois Maignien, Tom O Delmont, and Amy D Willis. Community-led, integrated, reproducible multi-omics with anvi'o. *Nature Microbiology*, 6(1):3–6, 2021.
60. Bui Quang Minh, Heiko A Schmidt, Olga Chernomor, Dominik Schrempf, Michael D Woodhams, Arndt Von Haeseler, and Robert Lanfear. Iq-tree 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution*, 37(5):1530–1534, 2020.
61. Chirag Jain, Luis M. Rodriguez-R, Adam M. Phillippy, Konstantinos T. Konstantinidis, and Srinivas Aluru. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications*, 9(1):5114, nov 2018.
62. Daniel R Utter, Gary G Borisy, A Murat Eren, Colleen M Cavanaugh, and Jessica L Mark Welch. Metapangenomics of the oral microbiome provides insights into habitat adaptation and cultivar diversity. *Genome Biology*, 21(1):293, 2020.
63. Tom O. Delmont and A. Murat Eren. Linking pangenomes and metagenomes: the Prochlorococcus metapangenome. *PeerJ*, 6:e4320, jan 2018.
64. Doug Hyatt, Gwo-Liang Chen, Philip F LoCascio, Miriam L Land, Frank W Larimer, and Loren J Hauser. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11(1):119, 2010.
65. Ben Langmead and Steven L Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4):357–359, 2012.
66. Stefan Bogdanov, Kaspar Ruoff, and Livia Persano Oddo. Physico-chemical methods for the characterisation of unifloral honeys: a review. *Apidologie*, 35(Suppl. 1):S4–S17, 2004.
67. Deon P Neveling, Akihito Endo, and Leon M T Dicks. Fructophilic Lactobacillus kunkeei and Lactobacillus brevis Isolated from Fresh Flowers, Bees and Bee-hives. *Current Microbiology*, 65(5):507–515, 2012.
68. Xingjie Wang, Kaidi Hu, Fang Liu, Juan Mou, Jinghui Lai, Mengmei Zhang, Song Wang, Qin Li, Jianlong Li, Aiping Liu, Xiaolin Ao, Li He, Shujuan Chen, Yong Yang, and Shuliang Liu. Isolation and characterization of a gas-producing and acid-resistant bacterium from spoiled vinegar. *International Journal of Food Microbiology*, 394:110167, 2023.
69. Zirui Ray Xiong, Jonathan H. Sogin, and Randy W. Worobo. Microbiome analysis of raw honey reveals important factors influencing the bacterial and fungal communities. *Frontiers in Microbiology*, 13, January 2023. Publisher: Frontiers.