Microbiome

RESEARCH Open Access

Culturomics of the pig tonsil microbiome identifies new species and an untapped source of novel antimicrobials

Isabela Maria Fernandes de Oliveira[1], Simen Fredriksen[1], Marcela Fernández Gutiérrez[1], Hermie J. M. Harmsen[2], Jos Boekhorst[1], Peter van Baarlen[1] and Jerry M. Wells[1,3*]

Abstract

Background In humans and pigs, altered composition of the microbiota associated with the epithelium of the palatine tonsils has been associated with bacterial or viral infection and lymphoid tissue inflammation. Tonsil lymphoid tissue is important for immunity and considered an important portal of entry for pathogens such as *Streptococcus suis*. Little is known about correlations between tonsil-associated microbiota, tonsillar infections, and the species that might confer colonization resistance against pathogens. Here, we describe a large collection of representative bacterial species from the tonsil surface biofilm and used genome mining and in vitro assays to assess their potential as probiotics to reduce infections by *S. suis* and other pathogens.

Results Data on tonsil microbiota composition from over 100 piglets from 11 farms and 3 countries revealed a core microbiota comprising *Actinobacillus*, *Streptococcus*, and *Moraxella* and 11 other less abundant but prevalent genera. To establish a collection of culturable core species, we plated 5 tonsil swabs taken from healthy piglets on different farms and countries on 8 different media and isolated 518 pure cultures belonging to 23 genera. To identify candidate probiotic strains, we tested for antagonistic activity against a panel of pathogens and in silico genome mining to find biosynthetic gene clusters (BGCs) in isolates that might produce antimicrobial compounds. We identified two novel species with potential probiotic activities: a *Brevibacterium* species and *Corynebacterium* species producing a heat and proteolytically stable lanthipeptide variant of flavucin, inhibiting in vitro growth of the opportunistic pathogens *S. suis* and *Staphylococcus aureus*.

Conclusions We defined the core tonsil microbiota of piglets and cultured representative single bacterial isolates for research on microbiota–host interactions in the oral cavity. Several isolates inhibiting the growth of bacterial pathogens that might be exploited as probiotics to promote colonization resistance were deposited in publicly available strain repositories. Our mining of genomes from cultured isolates suggests that the tonsil microbiota is an untapped source of novel antimicrobials

Keywords Culturomics, Porcine palatine tonsil, Microbiota, Biosynthetic gene cluster, Upper respiratory tract

*Correspondence:
Jerry M. Wells
jerry.wells@wur.nl
Full list of author information is available at the end of the article

Introduction

An important step towards unravelling causal host–microbiome interactions and identifying microbial functions contributing to human and animal health is to isolate and catalogue representative microbes colonizing specific body sites. To obtain a sufficiently diverse collection of microbes, they must be cultured in an unbiased fashion and identified according to standardized taxonomical procedures. Culturomics is the generic term given to such a diverse culturing approach and commonly utilizes matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and 16S rRNA gene sequencing to identify the taxonomy of isolates obtained by repeated selection of single colonies on agar plates. Genome sequencing and functional screening of isolate collections from environmental and host-associated microbial communities can lead to the discovery of novel species and biosynthetic gene clusters (BGCs) producing bioactive molecules that mediate microbe–microbe and microbe-host interactions [1].

Once characterized and archived as a pure culture, strains from specific body niches with symbiotic or beneficial activities can be used for translational research on host-microbe interactions or for biotechnological applications. Identification and (co)culture of niche-specific microorganisms are an important strategy to gain knowledge on community structure and antagonism between symbionts and pathobionts. One successful example of this strategy is the discovery of *Staphylococcus lugdunensis*, which produces lugdunin, a thiazolidine-containing cyclic peptide antibiotic that inhibits colonization by *Staphylococcus aureus* in the nasal cavity. Human nasal colonization by *S. lugdunensis* was associated with a significantly reduced carriage of *S. aureus* [2, 3]. The production of natural antimicrobials such as lugdunin is an adaptative advantage for bacteria and drives ecological processes such as colonization and niche persistence [4, 5]. A previous study of human microbiomes in several body sites revealed specific bacteria harboring BGCs predicted to produce novel bioactive compounds, which led to the discovery of lactocillin, a novel thiopeptide that was active against multiple gram-positive pathogens [1]. Other small molecules produced by the microbiota influence bacterial growth, motility, nutrient acquisition, stress response, and biofilm formation [6].

The palatine tonsils are a less well-studied niche for host microbiota interactions but are important because microorganisms entering the tonsil epithelial crypts are sampled by macrophages and dendritic cells to elicit mucosal immune responses that protect the respiratory tract [7, 8].

The tonsil surface is a major host niche for *Streptococcus suis*, a bacterial species that contains commensal strains but also disease-associated (or virulent) strains that can cause meningitis and sepsis in pigs and humans [9]. Reduced host immunity, coinfections, and stress appear to contribute to the risk of *S. suis* disease, and asymptomatic carriage of pathogenic strains of *S. suis* in piglets is common on farms with a recent history of *S. suis* disease [10]. Recent studies in piglets showed that the abundance of specific species of tonsil-colonizing bacteria is associated with *S. suis* disease, and that reduced microbiota diversity preweaning may predispose to *S. suis* disease postweaning [10, 11]. These findings suggest that oral administration of isolates that antagonize adherence or growth of *S. suis* may help to reduce *S. suis* carriage and risk for invasive disease, as shown for colonization by *S. lugdunensis* and nasal carriage of *S. aureus* [2].

The aim of this study was (i) to determine the core tonsil microbiome of healthy piglets on European farms without a recent history of *S. suis* cases, (ii) perform culturomics using the same tonsil swabs to establish a strain collection representing a substantial part of porcine health-associated core microbiota, and (iii) find novel tonsillar species that antagonize *S. suis* and other pathogens. We isolated 518 pure cultures of bacteria from the tonsillar microbiota belonging to 23 genera and identified different classes of small-molecule-encoding BGCs produced by less well-studied tonsillar microbiota taxa. We also describe two novel species with antagonistic activity against a panel of pathogenic bacteria, one producing a novel antimicrobial flavucin-like lanthipeptide.

Results

16S rRNA gene diversity and core microbiota of healthy piglets

We analyzed a total of 101 microbiota samples obtained as tonsil swabs of piglets from 11 farms located in Germany, the Netherlands, and Spain (Table 1) using 16S rRNA gene V3–V4 amplicon sequencing at a mean depth of 81,028 reads (range: 51,577–180,000 after quality filtering and chimera removal; see the " Material and methods").

DADA2 [12] identified 39,312 unique 16S rRNA gene V3–V4 region amplicon sequence variants (ASVs), of which 33.06% did not have a matching entry in the SILVA v132 database [13] at the species level across the 101 tonsil swab samples. The 3 most abundant genera present in the tonsil microbiota were *Actinobacillus*, *Moraxella*, and *Streptococcus*, adding up to 51.60% of the mean relative abundance with 206 unique ASVs (Fig. 1). At the genus level, after comparing the relative abundance and presence/absence of taxa across samples, we propose that 14 genera with relative abundance greater than 0.1% and presence in ≥ 90% of the samples (before and after

Table 1 Overview of the sampling distribution for the culturomics (ES1, ES2, and NL1) and microbiota analyses

Country	Farm	Age	No. of samples	Farm type
Germany	DE1	Postweaning	6	Commercial farming
Germany	DE2	Postweaning	6	Commercial farming
Germany	DE3	Postweaning	5	Commercial farming
Germany	DE4	Postweaning	4	Commercial farming
Germany	DE5	Postweaning	4	Commercial farming
Germany	DE6	Postweaning	4	Commercial farming
Spain	ES1	Postweaning	15	Commercial farming
Spain	ES2	Preweaning	16	Commercial farming
Spain	ES3	Preweaning	22	Commercial farming
Spain	ES4	Preweaning	10	Commercial farming
Netherlands	NL1	Postweaning	9	Research farm

weaning) can be designated core members of the tonsil microbiota. These core members comprised 80.36% of the total taxonomic diversity of the piglet tonsillar microbiota (Fig. 1).

At the ASV level, the core piglet tonsillar microbiota included 17 ASVs, 2 in the genus *Moraxella*, 7 in the genus *Actinobacillus*, 3 in the genus *Streptococcus*, and 1 each in the genera *Alloprevotella*, *Porphyromonas*, *Rothia*, *Prevotella*, and *Fusobacterium* (Table 2).

To assess bacterial community composition across all samples and farms, principal component analysis (PCA) on Bray–Curtis dissimilarity was conducted (Fig. 2). The first two principal components explained 23.1% (14.5% and 8.6% for PC1 and PC2, respectively) of the total variance in community composition. We found significant differences between microbiota composition and observed that the samples clustered together primarily according to country of origin (PERMANOVA on Bray–Curtis dissimilarity, $p < 0.001$) as well as per farm. Samples from farms in Germany and the Netherlands clustered more closely together compared to samples from Spanish farms in the PCA (Fig. 2). PERMANOVA on Bray–Curtis dissimilarity also detected significantly different microbiota compositions ($p < 0.001$) within each country. The Shannon alpha diversity of the core genera microbiota varied slightly between countries (Additional File 1).

Culturomics approach

To obtain cultivable bacterial isolates of the piglet tonsil-associated microbiota, swab samples from five piglets from three different farms (ES1, ES2, and NL1) were diluted and plated under different culture conditions and on eight different media (see the " Material and methods" section for details). Approximately, 10,000 colonies were screened, and random isolates with morphological variation were picked per plate, yielding 518 distinct bacterial isolates representing 60 species of gram-positive and gram-negative species from 23 genera (Additional Table 2) that were further characterized in vitro (Fig. 3A). Taxonomic classification was initially based on MALDI-TOF profiles and subsequently validated by full 16S rRNA gene sequencing (Fig. 3B/Additional Table 2).

To gain insight into the relative abundance of these 518 cultured isolates as a fraction of taxonomic diversity present in the total (cultured and uncultured) tonsillar microbiota, the full 16S rRNA gene sequences from the 518 isolates cultured were compared against the 16S rRNA V3–V4 region sequences that we obtained across the tonsil microbiota samples (Fig. 4). The taxonomic diversity of the 518 cultured isolates spanned 23 genera and corresponded to 38.89% of the total taxonomic diversity captured by 16S profiling of our tonsillar microbiota dataset. From these cultured isolates, 46.5% ($n = 241$) were prevalent in > 90% of all tonsillar samples that had been obtained at 11 farms from 3 countries. The 16S rRNA sequences corresponding to the 255 other cultured isolates had a prevalence between 60 and 20% across tonsillar microbiota samples and were found at relative abundances between 0.4 and 0.05%. The remaining 22 strains were rarely found in the tonsil microbiota at a mean relative abundance of ≤ 0.09%. The 16S sequences from cultured isolates that mapped to ASVs of core microbiota members were classified as *S. suis*, *Moraxella pluranimalium, Moraxella porci*, and *Rothia nasimurium* (Table 2).

Identification of novel species

Based on 16S rRNA gene analysis, genome-based taxonomic assignment, and cellular fatty acid profiles, we propose two strains as novel species: *Brevibacterium moorei* and *Corynebacterium kozikiae* (Fig. 5). *B. moorei* (Brevibacteriaceae) (strain 68QC2CO) has a genome length of 3.42 Mbp spanning 69 contigs with an average GC content of 66.76%. Two other strains from our culture collection, 50QC2O2 and 91QC2O2, were also classified as *B. moorei* (Fig. 5A). *Brevibacterium rongguiense* 5221 is the closest phylogenetic neighbor, with 97.00% sequence identity to the 16S rRNA gene of *B. moorei*. The average nucleotide identity (ANI) between the 3 *B. moorei* strains and 21 available genomes of other *Brevibacterium* species varied between 78.45% and 77.15% identity, supporting the notion that *B. moorei* is a separate species (Additional File 3). The digital DNA–DNA hybridization (dDDH) sequence identity between *B. rongguiense* 5221 and *B. moorei* was 14.8%. The *Brevibacterium* genus was present in low abundance among the 101 sequenced porcine tonsil microbiota samples from

DE1	DE2	DE3	DE4	DE5	DE6	ES1	ES2	ES3	ES4	NL1	Mean	ASVs	Prevalence	
22.1	29.1	36.5	25.9	31.0	18.4	31.9	38.0	25.5	24.9	24.7	28.0	130	100	***Actinobacillus***
5.6	5.9	6.1	2.8	4.8	7.6	22.8	21.6	28.6	20.3	8.2	12.2	37	100	***Moraxella***
5.5	22.5	17.9	15.4	17.3	12.3	6.6	7.4	11.1	3.7	8.5	11.7	43	100	***Streptococcus***
3.1	2.8	0.7	3.2	3.2	5.9	1.8	2.7	2.7	17.6	8.3	4.7	42	100	***Porphyromonas***
2.2	2.9	1.1	7.4	3.2	7.8	4.2	3.9	3.3	6.5	3.0	4.1	17	100	***Alloprevotella***
9.0	3.9	5.0	2.8	2.5	2.2	4.6	2.4	1.8	1.6	2.2	3.4	45	100	***Prevotella***
2.7	1.2	0.4	6.4	2.4	2.4	1.7	2.4	2.3	4.5	2.7	2.7	13	100	***Fusobacterium***
2.2	1.2	0.7	4.2	3.0	1.5	1.6	1.5	2.6	1.6	4.7	2.2	11	94	***Acinetobacter***
3.8	3.8	4.1	3.8	3.7	1.7	0.6	0.9	0.6	0.1	0.9	2.2	13	92	***Lactobacillus***
1.4	0.8	0.6	1.5	1.4	5.8	0.5	1.2	1.0	2.1	3.2	1.8	22	100	***Bacteroides***
1.2	3.5	0.3	1.5	1.6	2.8	0.8	2.3	1.5	2.5	1.0	1.7	16	99	***Leptotrichia***
0.6	0.8	0.8	0.5	1.5	6.3	1.6	1.0	0.9	2.0	2.1	1.6	3	100	***Rothia***
0.4	3.9	1.1	3.4	1.7	3.7	0.5	0.5	0.4	0.2	0.5	1.5	11	85	*Veillonella*
1.9	1.4	0.3	0.9	1.1	2.0	0.6	0.6	0.9	1.7	0.9	1.1	9	100	***Neisseria***
1.3	0.7	1.6	0.8	0.6	0.5	2.3	0.8	0.6	0.0	0.9	0.9	6	83	*Clostridium*
0.8	0.3	0.1	1.0	0.6	0.7	0.1	0.1	0.3	0.3	4.1	0.8	4	67	*Alysiella*
0.1	0.1	0.2	0.3	0.3	0.4	0.6	1.2	3.5	1.0	0.1	0.7	8	70	*Bergeyella*
0.6	0.2	0.2	0.9	1.0	0.8	0.3	0.4	0.4	0.8	1.3	0.6	4	86	*Conchiformibius*
1.5	0.5	1.2	0.2	0.3	0.1	0.7	0.3	0.1	0.0	0.4	0.5	8	66	*Agathobacter*
1.7	0.6	1.6	0.3	0.4	0.1	0.3	0.1	0.0	0.0	0.3	0.5	5	54	*Subdoligranulum*

Fig. 1 Heatmap displaying the most abundant and prevalent genera across all samples ($n = 101$) collected in three countries in Europe (DE, Germany; NL, the Netherlands; and ES, Spain). Fourteen core genera that occurred in the dataset after applying a 90% prevalence and 0.1% abundance threshold are displayed in bold font. The numbers in columns DE1 to NL1 represent the average relative abundance. ASVs, unique amplicon sequence variants per genus. Mean, overall average of the per-country averages

different farms and countries (prevalence, 7% among all piglets; mean relative abundance, 8.24E-04%). The major membrane fatty acid isolates of *B. moorei* 68QC2CO are anteiso $C_{17:0}$ (43.92%) and anteiso $C_{15:0}$ (21.74%), iso $C_{15:0}$ (14.60%), and iso $C_{16:0}$ (10.60%), with $C_{16:0}$ (3.67%), $C_{17:0}$ (3.82%), $C_{14:0}$ (0.63%), iso $C_{14:0}$ (0.35%), and $C_{18:0}$ (0.37%) in lower amounts.

The second novel species, *C. kozikiae* (isolate 76QC2CO), has a genome length of 2.45 Mbp spanning 57 contigs with an average GC content of 60.36%. Seven other isolates from our culture collection were classified as *C. kozikiae* (strains: 35RC1, 70RC1, 122RC1, 142RC1, 153RC1, 209RC1, 732RC1) (Fig. 5B). *C. kozikiae* isolate 76QC2CO has as closest phylogenetic neighbor "unclassified strains *Corynebacterium* sp." (99% rRNA gene identity) followed by *Corynebacterium vitaeruminis* DMS20294 at 97.00% 16S rRNA gene sequence identity. The dDDH identity between *C. vitaeruminis* DMS20294 and *C. kozikiae* was 14.3%. The ANI value between the *C. kozikiae* strains and 80 other reference species from the genus *Corynebacterium* was between 73.79–79.82%, supporting the notion that *C. kozikiae* is a novel species (Additional File 3). The mean relative abundance of the genus in the pig tonsil microbiota data from 101 piglets was 0.084%. The major cellular fatty acids of *C. kozikiae* isolate 76QC2CO were $C_{16:0}$ (57.98%), $C_{16:1\ w7c}$ (16.15%), $C_{17:0}$ (5.06%), $C_{18:1\ w9c}$ (6.61%), and $C_{14:0}$ (6.43%). The following fatty acids were detected in lower amounts: $C_{18:0}$

Table 2 ASV-level core microbiota from 101 animals on 11 European farms from 3 countries collected in this study. An ASV was scored as part of the core microbiota in this study when it had been scored in ≥ 90% of the samples, and ASV abundance in the respective microbiota sample had been higher than 0.1%

ASV	Taxon	Family	Prevalence (%)	Mean abundance (%)
ASV00006	*Streptococcus suis*	Streptococcaceae	100.00	2.06
ASV00018	*Streptococcus suis*	Streptococcaceae	99.28	1.44
ASV00002	*Moraxella pluranimalium*	Moraxellaceae	98.56	3.85
ASV00013	*Moraxella porci*	Moraxellaceae	98.56	1.72
ASV00020	*Porphyromonas* Unclassified	Porphyromonadaceae	98.56	1.30
ASV00003	*Alloprevotella* Unclassified	Prevotellaceae	97.12	2.73
ASV00027	*Actinobacillus indolicus/H. parasuis*	Pasteurellaceae	97.12	0.78
ASV00021	*Fusobacterium gastrosuis*	Fusobacteriaceae	96.40	1.06
ASV00029	*Prevotella* Unclassified	Prevotellaceae	94.96	0.60
ASV00007	*Actinobacillus porcinus*	Pasteurellaceae	93.53	2.43
ASV00015	*Actinobacillus minor/porcitonsillarum*	Pasteurellaceae	93.53	1.81
ASV00016	*Actinobacillus indolicus/minor*	Pasteurellaceae	92.81	1.38
ASV00008	*Actinobacillus indolicus/H. parasuis*	Pasteurellaceae	92.09	1.42
ASV00023	*Actinobacillus minor/indolicus*	Pasteurellaceae	92.09	0.69
ASV00009	*Actinobacillus porcinus*	Pasteurellaceae	91.37	2.17
ASV00036	*Rothia nasimurium*	Micrococcaceae	91.37	0.53
ASV00019	*Streptococcus suis*	Streptococcaceae	90.65	1.32

(3.22%), $C_{17:1\ w8c}$ (1.08%), iso $C_{10:0}$ (0.95%), $C_{12:0}$ (0.62%), $C_{14:1\ w5c}$ (0.79%), $C_{9:0}$ (0.56%), 3OH $C_{16:0}$ (0.26%), and $C_{18:2\ w6,9c}$ (0.29%).

Discovery of antagonistic bacteria in overlay inhibition assays and genome mining for antimicrobial peptide-encoding genes

We tested all 518 cultured strains for their growth inhibitory effect on a panel of disease-associated and commensal bacteria using overlay inhibition (antagonism) assays (see the " Material and methods"). Of these 518 strains, [14] showed reproducible inhibitory activity in vitro against at least 1 of the target species tested (Table 3, Fig. 3C).

To identify genomic gene clusters predicted to produce natural antibiotics, whole-genome sequencing was carried out for 45 isolates based on their strong and reproducible antagonistic activities and 16S rRNA diversity (Table 3). These 45 isolates belonged to the genera *Brevibacterium* ($n=3$), *Corynebacterium* ($n=16$), *Neisseria* ($n=1$), *Pediococcus* ($n=2$), *Pelistega* ($n=1$), *Staphylococcus* ($n=13$), and *Streptococcus* ($n=9$) (Fig. 6). To identify biosynthetic gene clusters (BGCs) that might produce natural products, we used the in silico analysis tool antiSMASH (the " Material and methods" section).

In the genomes of the 45 isolates, we identified 164 BGCs (68.6% of these had less than 10% homology or no homology with known BGCs) predicted to produce a broad variety of secondary metabolites, such as nonribosomal peptides (NRPs), terpenes, siderophores, polyketides (PKs), bacteriocins, lanthipeptides, and ribosomally synthesized and posttranslationally modified peptides (RiPPs) (Additional Table 4) (Fig. 6).

The genomes of seven *Staphylococcus* strains that inhibited the growth of *S. suis* (Table 3) contained between 4 and 8 BGCs, some of which were predicted to produce antimicrobial compounds. The BGCs identified in the staphylococcal tonsil isolates included NRPSs and type III polyketide synthases (T3PKSs) and clusters producing RiPPs, siderophores, and terpenes. AntiSMASH predicted that specific BGCs identified in *S. aureus* produce siderophores with high identity to the siderophores staphyloferrin (87 to 100% identity) and staphylobactin (87% identity) produced by *S. aureus*. The NRPs producing aureusimine, which is involved in biofilm production [4], and the bacteriocin hyicin 3682 were identified in some porcine isolates of *S. aureus* (Table 3). AntiSMASH predicted 5 BGCs in the porcine *Staphylococcus hyicus* isolates, including clusters with 81% identity to the BGC producing the lanthibiotic suicin 90–1330 and clusters producing the siderophores staphyloferrin (100% identity) and a staphylobactin analogue (51% identity). The genomes of strains from the staphylococcal species *Staphylococcus borealis*, *Staphylococcus cohnii*,

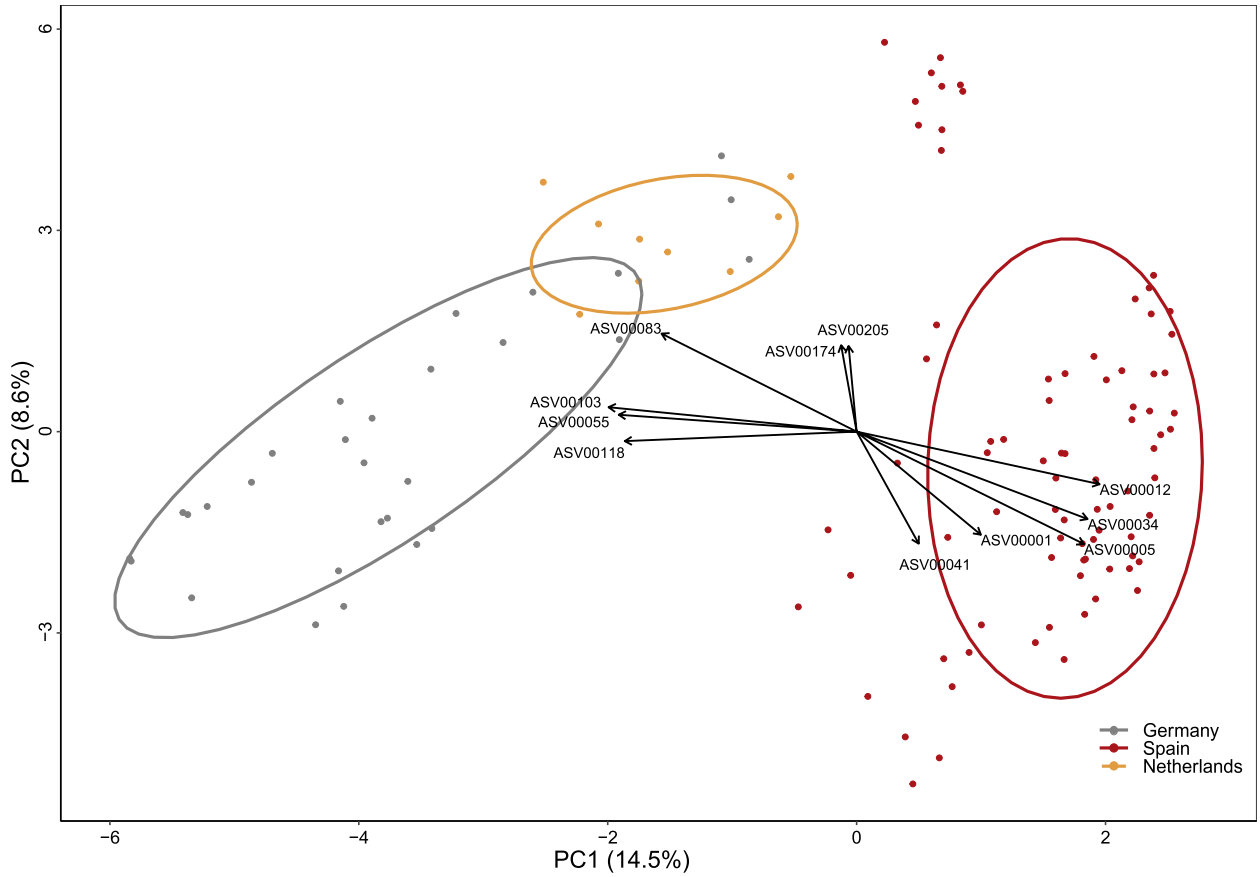

Fig. 2 Principal component analysis (PCA) of bacterial community composition across samples and farms. PCA was performed on Bray–Curtis dissimilarity indices to visualize the variation in bacterial communities. The first two principal components accounted for 23.1% of the total variance (14.5% by PC1 and 8.6% by PC2), demonstrating the diversity in microbial profiles. The proximity of points on the plot indicates the similarity in bacterial composition between the samples, with closer points representing more similar communities. The samples clustered per country, red dots represent all samples from Spanish farms on the right side, gray dots on the left side represent samples from German farms, and yellow dots in the center/upper area represent samples from Dutch farms. Each ellipse represents a 75% confidence interval border for each group. The arrows show how ASVs contribute to the patterns observed in the PCA and how they relate to each other

and *Staphylococcus saprophyticus* were predicted to contain four, seven and four BGCs, respectively. *S. borealis* strains showed antimicrobial activity against *M. pluranimalium* (Table 3). One of the 4 BGCs predicted in the genome of *S. borealis* had 100% identity to a BGC producing the siderophore staphyloferrin A; the other BGCs appear to be novel.

Out of the 16 streptococcal species identified in piglet tonsillar microbiota, 8 *S. suis* strains and 1 *Streptococcus gallolyticus* strain were included for sequencing. The strains from these two species were predicted to produce 10 antimicrobials belonging to different classes, including RiPP-like compounds, compounds produced by T3PKSs, lassopeptides, and lanthipeptides (Fig. 6). T3PKSs were present in all streptococci sequenced in our study. The sequenced *S. suis* genomes carried a gene encoding a radical S-adenosylmethionine (rSAM) enzyme that was predicted by antiSMASH to be involved in biosynthesis of the macrocyclic peptide streptide (75% identity) and the lanthibiotic suicin (36–45% identity). BGCs producing secondary metabolites with 58% identity to linocin-M18 BGC and with 97.81% similarity to the bacteriocin lactococcin 972 BGC were also identified. Although these molecules have been associated with antibiotic activity [5, 15], none of the sequenced *S. suis* strains showed in vitro activity in inhibition assays against the tested targets (Table 3).

The *Neisseria dentiae* genomes contained BGCs predicted to produce arylpropylene, terpene, homoserine (hser)lactone, and RiPP-like compounds. *Pelistega suis* genomes encode BGCs predicted to produce betalactones and terpenes. *Pediococcus pentosaceus* genomes contained BGCs with high identity to T3PKs and RiPPs (Fig. 6). All *Neisseria*, *Pelistega*, and *Pediococcus* strains tested had antimicrobial activity in the antagonism assays, consistent with the prediction that these BGCs

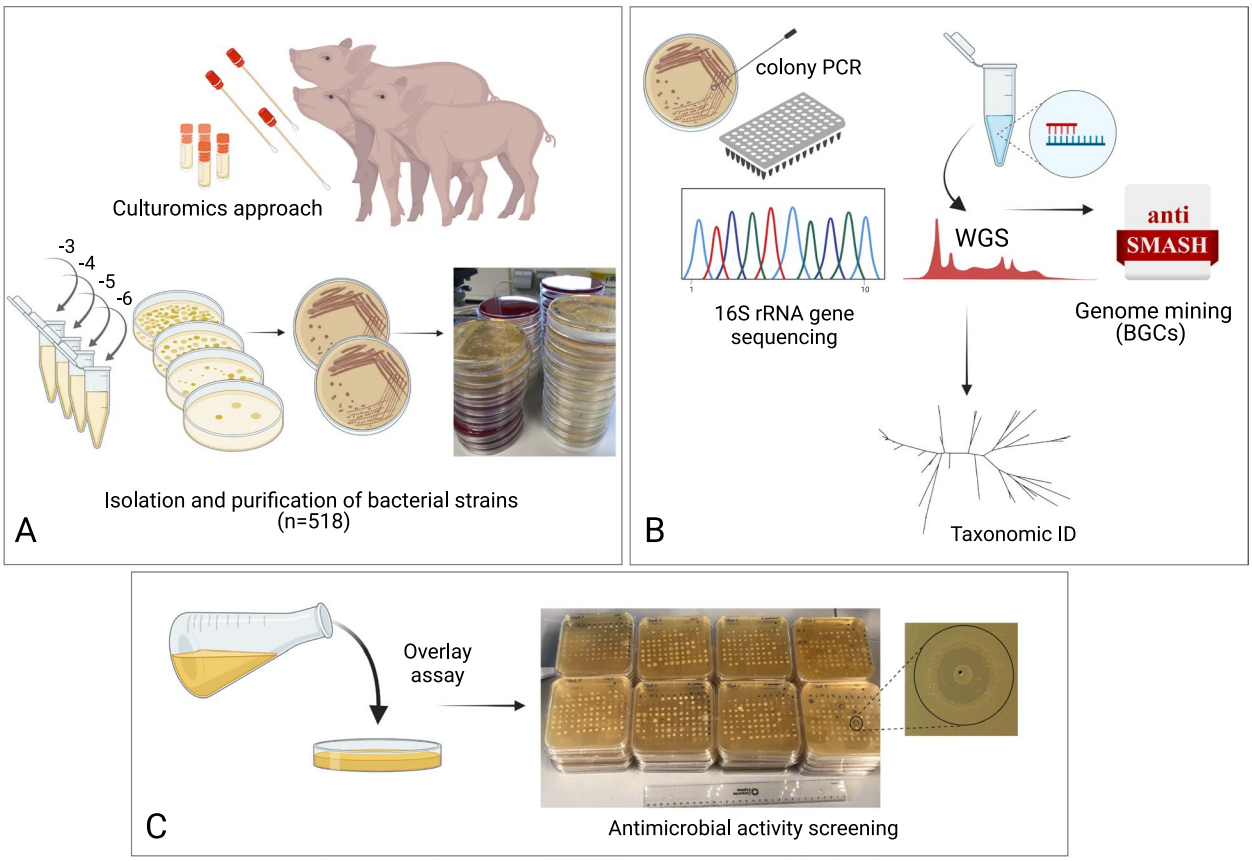

Fig. 3 Schematic overview of steps to obtain cultivable species and identify candidate producers of antimicrobial products from the tonsil microbiome. **A** Processing of pig tonsil swab samples, including isolation and purification steps of bacterial isolates using nine different culture media. **B** Taxonomic identification of pure tonsil isolates using full 16S rRNA gene sequences obtained by colony PCR, whole-genome sequencing (WGS), genome mining, and in silico prediction of production of bioactive compounds by identification and analysis of biosynthetic gene clusters (BGCs). **C** In vitro screening for natural antimicrobial compounds produced by tonsil isolates. Illustration generated using BioRender (https://app.biorender.com)

produced natural antibiotic compounds (Table 3). The three strains from the proposed novel species *B. moorei* were predicted to contain highly conserved NRPSs, but no inhibition zones were observed in the inhibition assays using these strains, possibly due to lack of expression in vitro.

The genomes of the *Corynebacterium* strains reported in Table 3, including the proposed novel species *C. kozikiae*, were predicted to include NRPSs, T1PKSs, T3PKSs, and NAPAA (non-alpha poly-amino acid with NRP domains) and one cluster producing lanthipeptide class I compounds. The isolate 76QC2CO (*C. kozikiae*) was predicted to produce a class I lanthipeptide and displayed strong and consistent inhibitory activity in vitro (Table 3). In agar overlay and agar well-diffusion assays, isolate 76QC2CO inhibited the growth of multiple strains and serotypes of the porcine pathogen *S. suis* and two other closely related streptococci, as well as *S. aureus*. AntiSMASH detected a lanthipeptide class I BGC predicted to synthesize a compound 94% similar to the lantibiotic flavucin, previously identified in the genome of *Corynebacterium lipophiloflavum* DSM 44291 [16] (Fig. 7A). Compared to the original flavucin produced by *C. lipophiloflavum* DSM 44291, the predicted lantibiotic of isolate 76QC2CO contains two amino acid modifications in the leader peptide and five in the core peptide (Fig. 7B); these posttranslational modifications may lead to a mature peptide carrying five amino acid differences compared to the original flavucin. We named this novel flavucin variant flavucin IF76. The predicted modifications were mainly due to substitutions of amino acids with similar biochemical properties, apart from a mutation yielding a different amino acid with a side chain of different sizes. The impact of these alterations displayed in the structure of flavucin IF76 may result in differences in 3D folding and/or in the antimicrobial activity compared to the known flavucin (Fig. 7C).

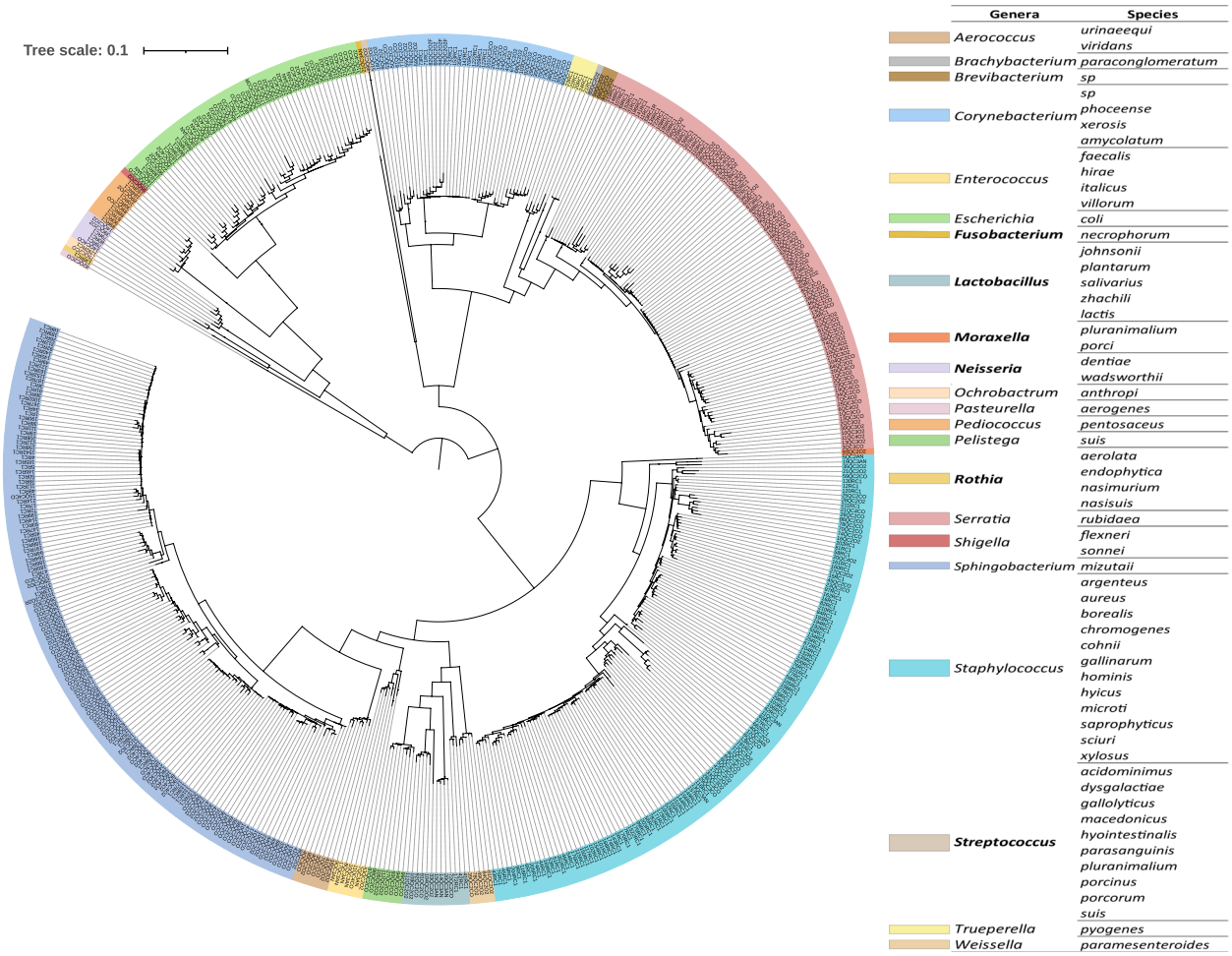

Fig. 4 Phylogenetic tree of 518 bacterial single-colony isolates from the porcine tonsillar microbiota based on full 16S rRNA gene sequences. The tree was visualized using the iTOL platform (https://itol.embl.de/). Genera considered part of the core microbiota (90% prevalence and 0.1% abundance threshold) are displayed in bold font

We obtained and sequenced nine isolates of the proposed novel species *C. kozikiae* (Fig. 5) and found that only *C. kozikiae* isolate 76QC2CO contained the lanthipeptide class I BGC predicted to produce flavucin IF76. Flavucin IF76 is stable at high temperatures (> 95 °C), and proteinase K and trypsin treatments had small effects (< 10%) on peptide antimicrobial activity. Moreover, flavucin IF76 displayed no cytotoxicity to HEK-293 and HEPG2 mammalian cell lines (data not shown).

Finally, strains from the streptococcal species *Streptococcus macedonicus*, *Streptococcus porcinus*, and *Streptococcus porcorum* and strains belonging to the genera *Aerococcus*, *Enterococcus*, *Lactobacillus*, and *Escherichia/Shigella* also showed inhibitory activity against the target panel of pathogens (Table 3); we did not further investigate these isolates because their spectrum of antagonism did not go beyond the activity of other isolates that we had selected to characterize in greater detail.

Discussion

To assess taxonomic diversity in the tonsil microbiota from healthy piglets at 11 farms in Germany, Spain, and the Netherlands, we used two complementary methods to characterize tonsillar bacterial diversity. Using 16S rRNA gene amplicon sequencing, we discovered a core tonsil-associated microbiota that consisted of 14 genera and 17 individual amplicon sequence variants (ASVs) at more than 0.1% abundance in 90% or more of the sampled healthy piglets. By culturomics, we obtained 518 strains belonging to 6 of the 14 core genera. The bacterial taxa that we cultured and identified by 16S rRNA gene sequence profiling were largely consistent with earlier DNA-based taxonogenomic studies on healthy piglets without antibiotic treatment [17, 18], underpinning the relevance of these taxa in the establishment of a health-associated early life microbiome of piglets. The core genera that we found in the 101 piglets have been reported

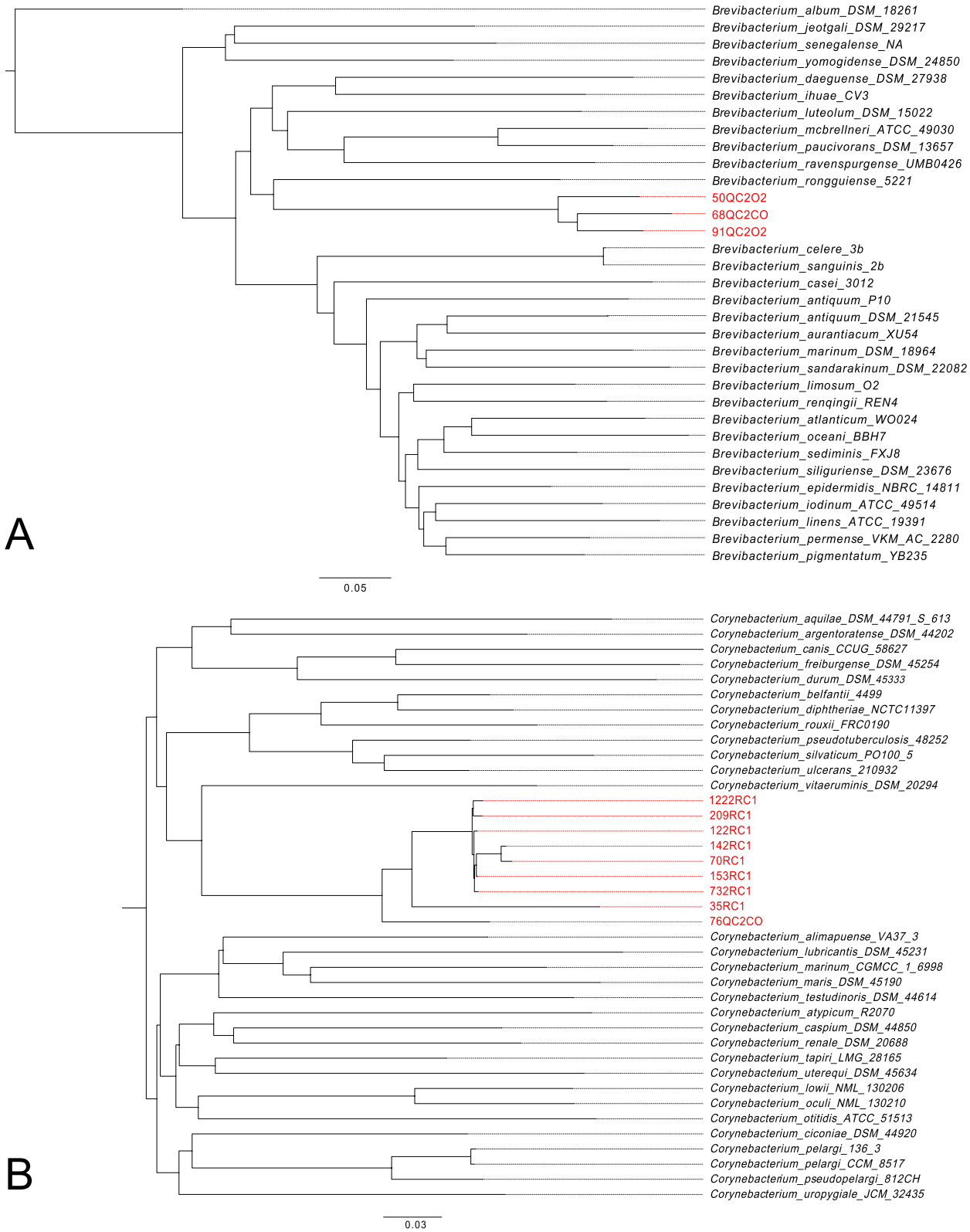

Fig. 5 Phylogenetic tree (core genes; see the "Material and methods") showing the taxonomic position of the two new species: **A** *B. moorei* and **B** *C. kozikiae*. Isolates in red font represent the proposed new species

Table 3 Results of inhibition assays with porcine isolates obtained from tonsil microbiota. Shown are strains (ID code) with inhibitory activity against at least one of the target bacteria tested

		Inhibition assays									
ID	**16S ID**	**SS J28**	**SS S10**	**SS P1/7**	**SP_1 DSM 23759**	**SP_2 DSM 20725**	**SP_3 DSM 29126**	**SA 6538P**	**EC L4242**	**PA DSM 10153**	**MP DSM 22804**
3QC2CO	*Aerococcus urinaeequi*	-	-	-	-	-	-	+	-	-	-
4QC2CO	*Aerococcus urinaeequi*	-	-	-	-	-	-	+	-	-	-
7QC2CO	*Aerococcus urinaeequi*	-	-	-	-	-	-	+	-	-	-
15QC2O2	*Aerococcus urinaeequi*	-	-	-	-	-	-	+	-	-	-
4QC3CO	*Aerococcus viridans*	-	-	-	-	-	-	+	-	-	-
90QC2O2	*Aerococcus viridans*	-	-	-	-	-	-	++	+	-	-
76QC2CO	*Corynebacterium* sp. (WGS)	++	-	++	++	-	++	+	-	-	-
20QC2O2	*Corynebacterium* sp. (WGS)	-	-	-	-	-	-	-	-	±	±
7QC3AN	*Enterococcus faecalis*	-	-	-	±	-	+	-	-	-	-
79QC2CO	*Escherichia coli*	-	-	-	-	-	-	-	++	-	-
85QC2CO	*Escherichia coli*	-	-	-	-	-	-	-	++	-	-
97QC2O2	*Escherichia coli*	-	-	-	-	-	-	-	++	-	-
21QC4O2	*Escherichia coli*	-	-	-	-	-	-	-	+	-	-
22QC4O2	*Escherichia coli*	-	-	-	-	-	-	-	+	-	-
1QC2AN	*Escherichia coli*	-	-	-	-	-	-	-	+	-	-
2QC2AN	*Escherichia coli*	-	-	-	-	-	-	-	+	-	-
4QC2AN	*Escherichia coli*	-	-	-	-	-	-	-	+	-	-
16QC3AN	*Escherichia coli*	-	-	-	-	-	-	-	+	-	-
41RC1	*Lactobacillus salivarius*	±	-	-	-	-	-	-	-	-	-
172RC1	*Lactobacillus salivarius*	-	-	-	-	+	-	-	-	-	-
1292RC1	*Neisseria dentiae* (WGS)	-	-	-	±	-	-	-	-	-	-
105QC2O2	*Pediococcus pentosaceus* (WGS)	-	-	-	-	-	-	-	-	-	+
111QC2O2	*Pediococcus pentosaceus* (WGS)	-	-	-	-	-	-	-	-	-	+
138RC1	*Pelistega suis* (WGS)	-	±	-	-	-	-	-	-	-	-
95QC2O2	*Serratia rubidaea*	-	-	-	-	-	-	-	+	-	-
99QC2O2	*Shigella flexneri*	-	-	-	-	-	-	-	+	-	-
16QC3CO	*Shigella sonnei*	-	-	-	-	-	-	-	+	-	-
21QC3CO	*Shigella sonnei*	-	-	-	-	-	-	-	+	-	-
1252RC1	*Staphylococcus aureus*	±	-	+	-	-	+	-	-	-	-
128RC1	*Staphylococcus aureus*	+	-	+	±	-	-	-	-	-	-
129RC1	*Staphylococcus aureus*	+	-	+	±	-	-	-	-	-	-
131RC1	*Staphylococcus aureus* (WGS)	+	-	+	±	-	-	-	-	-	-

Table 3 (continued)

ID	16S ID	Inhibition assays									
		SS J28	SS S10	SS P1/7	SP_1 DSM 23759	SP_2 DSM 20725	SP_3 DSM 29126	SA 6538P	EC L4242	PA DSM 10153	MP DSM 22804
133RC1	*Staphylococcus aureus* (WGS)	+	-	+	±	-	-	-	-	-	-
137RC1	*Staphylococcus aureus* (WGS)	+	-	+	±	-	+	-	-	-	-
63QC2O2	*Staphylococcus borealis* (WGS)	-	-	-	-	-	-	-	-	-	+
32QC2O2	*Staphylococcus gallinarum*	-	-	-	-	-	-	-	-	-	+
140RC1	*Staphylococcus hyicus*	-	±	-	-	-	-	-	-	-	-
26QC3O2	*Staphylococcus microti*	-	-	-	±	-	-	-	-	-	-
41QC2CO	*Staphylococcus sciuri*	±	-	-	-	-	-	-	-	-	-
8QC2O2	*Staphylococcus sciuri* (WGS)	-	-	-	-	-	-	-	-	+	+
83QC2O2	*Staphylococcus xylosus*	±	-	-	-	-	+	-	-	-	-
12QC3O2	*Streptococcus acidominimus*	±	-	-	±	-	-	-	-	-	-
8RC1	*Streptococcus gallolyticus*	-	-	-	-	+	-	-	-	-	-
9RC1	*Streptococcus gallolyticus*	-	-	-	-	+	±	-	-	-	-
74RC1	*Streptococcus gallolyticus*	-	-	-	-	+	+	-	-	-	-
99RC1	*Streptococcus gallolyticus*	-	-	-	-	+	-	-	-	-	-
150RC1	*Streptococcus gallolyticus*	-	-	-	-	+	-	-	-	-	-
171RC1	*Streptococcus gallolyticus*	-	-	-	-	+	-	-	-	-	-
13QC3AN	*Streptococcus gallolyticus*	+	+	+	-	+	±	-	-	-	-
75RC1	*Streptococcus gallolyticus*	-	-	-	-	+	+	-	-	-	-
782RC1	*Streptococcus gallolyticus*	-	-	-	-	+	+	-	-	-	-
143RC1	*Streptococcus gallolyticus*	-	-	-	-	+	-	-	-	-	-
149RC1	*Streptococcus gallolyticus*	-	-	-	-	+	-	-	-	-	-
205RC1	*Streptococcus gallolyticus*	-	-	-	-	+	-	-	-	-	-
151RC1	*Streptococcus gallolyticus* (WGS)	-	-	-	-	+	-	-	-	-	-
12QC3AN	*Streptococcus hyointestinalis*	+	+	+	-	+	-	-	-	-	-
56RC1	*Streptococcus macedonicus*	-	-	-	-	+	-	-	-	-	-
40QC2CO	*Streptococcus parasanguinis*	±	-	-	-	-	-	-	-	-	-
5QC2AN	*Streptococcus porcinus*	+	-	+	±	-	±	-	-	-	-
33QC3O2	*Streptococcus porcorum*	-	-	-	-	-	-	-	±	-	-

Inhibition zones were determined after overnight incubation. *WGS* whole genome sequence available. $^{-}$Absence of inhibitory activity, $^{\pm}$inhibition zone of 2-mm diameter (weak), $^{+}$inhibition zone between 3- and 5-mm diameter (weak), $^{++}$inhibition zone between 5- and 8-mm diameter. *Streptococcus suis* (*SS*), *Streptococcus porci* (SP_1), *Streptococcus porcinus* (SP_2), *Streptococcus parasuis* (SP_3), *Escherichia coli* (*EC*), *Pasteurella aerogenes* (*PA*), *Moraxella pluranimalium* (*MP*)

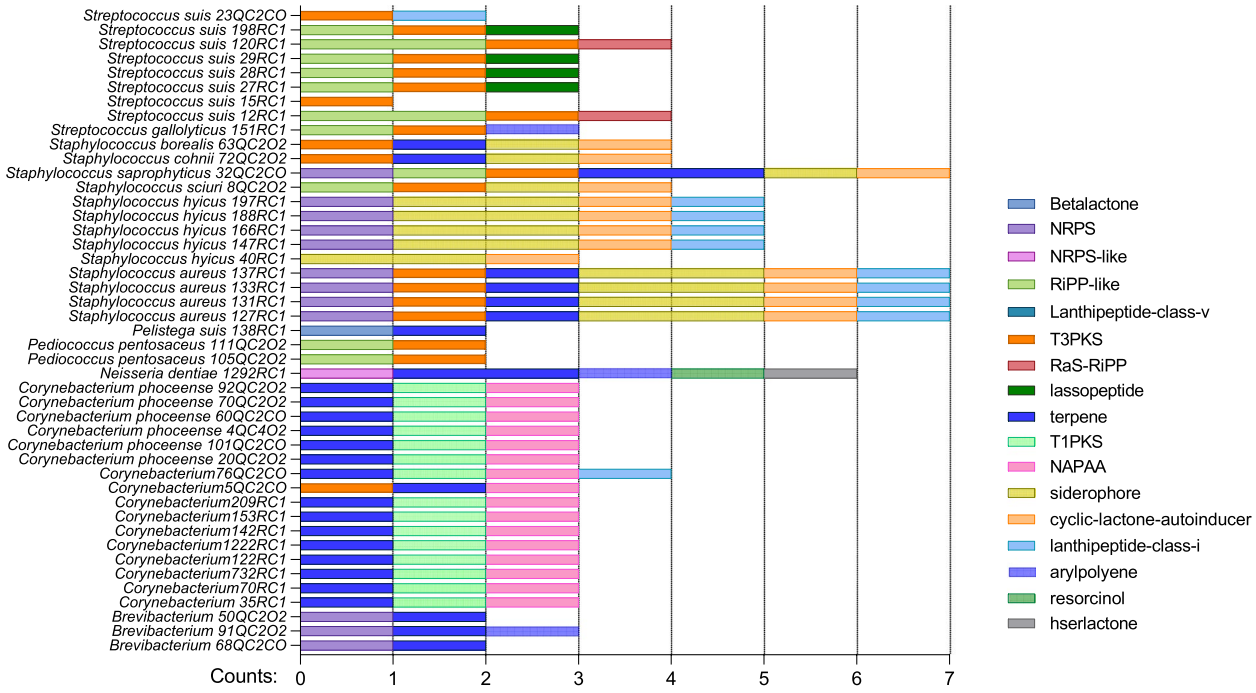

Fig. 6 Biosynthetic gene clusters predicted using antiSMASH from genome sequences of 45 pig tonsil isolates that were selected by different colony morphologies and antagonistic activity in bioassays. Different classes of secondary metabolites are represented with different colors. The *x*-axis shows how many of the different biosynthetic gene clusters (BCGs) are present in the genomes of each strain listed on the *y*-axis (extended table provided in Additional Table 4). Gene clusters of the same type are represented by the same color and sometimes appear two or more times in the same genome when the sequences of the genes within the BGC are not identical (e.g., there may be two different BGCs producing a siderophore ion chelating molecule)

by others to be members of the tonsil microbiota of healthy piglets with no recent history of respiratory disease [19], reinforcing the idea they might be symbionts associated with a health-associated microbiota. Other genera we uncovered at lower prevalence (≤16%) and abundance (≤0.1%) included *Pasteurella*, *Haemophilus*, and *Peptostreptococcus*, genera also previously reported to be core members of the tonsil microbiota of piglets. However, some species within these genera contain disease-associated strains causing infection in piglets of different ages and with hosts in compromised health status [20]. Minor differences in the abundance of certain genera (Fig. 1) might relate to the transport of pigs, country- and/or farm-specific differences in antimicrobial usage, animal feed compositions, and various environmental conditions and farming practices.

Our collection of 518 bacterial pure cultures covers a substantial proportion of the high bacterial biodiversity associated with the porcine tonsil surface, representing 60 species from 23 genera identified by 16S rRNA gene sequences. This includes two novel species belonging to the genera *Corynebacterium* and *Brevibacterium* harboring BGCs producing bioactive metabolites likely involved in microbiota ecology.

Our cultured bacterial tonsil collection contained aerobic and facultative anaerobic bacteria, of which 86% were gram-positive bacteria from the phyla Firmicutes (new name: Bacillota) and Actinobacteria (new name: Actinomycetota), and 14% were gram-negative species from the phyla Proteobacteria (new name: Pseudomonadota) and Bacteroidetes (new name: Bacteroidota). Our collection did not include members of the genus *Actinobacillus*, even though it is an abundant genus in the tonsil microbiota. The failure to culture species of *Actinobacillus* was probably due to their fastidious growth requirements and long incubation times needed for colony formation [21].

Strains belonging to the genera *Moraxella*, *Streptococcus*, *Fusobacterium*, *Lactobacillus*, *Rothia*, and *Neisseria* were present in the microbiota in ≥90% of all 101 tonsil samples, and species from these genera cultured comprised 71% of our collection. These microbiota members might be associated with relevant ecological interactions within the niche. Therefore, our collection of sequenced and biologically characterized strains will be a valuable resource for future research on the health-associated tonsillar microbiome. The remainder of the cultured strains (29%) featured low-abundance ASVs (<0.1% relative abundance). Although we focused on characterizing

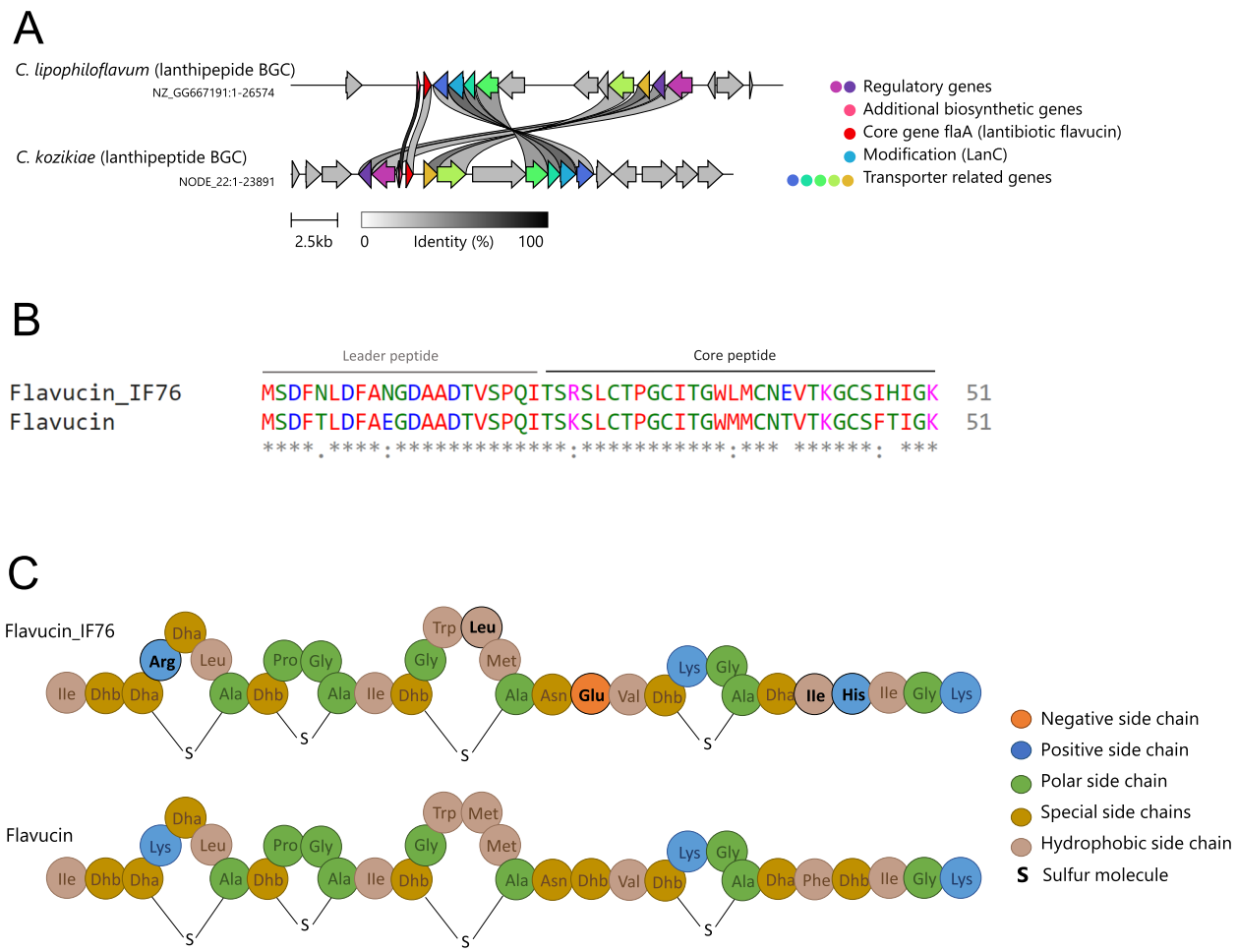

Fig. 7 Summary of the major features of flavucin IF76, a novel variant of the class I lanthipeptide flavucin. **A** Conserved genes present in the biosynthetic gene clusters (BGCs) producing the class I lanthipeptide of the species *C. kozikiae* and *C. lipophiloflavum* predicted to produce flavucin; homologous genes from the flavucin IF76 and original flavucin BGCs are labelled with the same color. Cluster organization was visualized using clinker (https://github.com/gamcil/clinker). **B** Amino acid sequences of the two variant flavucin peptides produced by the genomes of *C. kozikiae* and *C. lipophiloflavum*, displaying the two modifications present in the leader peptide and the five modifications present in the core peptide. **C** Schematic representation of the structure of the lanthipeptide flavucin IF76 and the original flavucin; the chains of amino acid residues containing the five amino acid differences are highlighted in black

the most abundant taxa for practical and applied reasons, transient or less abundant species may also contribute to the establishment of health-associated microbial communities through several mechanisms, such as blocking invasion of niches by disease-associated species (thus avoiding dysbiosis) and contributing to microbiota metabolic potential by increasing genetic and enzymatic diversity within the population [22, 23].

It is important to note that the relative abundances of bacterial taxa in the 101 reported samples were determined through sequencing of the 16S rRNA V3–V4 amplicon, a method that, while effective for taxonomic profiling, provides no insight into functional genetic repertoires of uncultured or unclassified microorganisms. More comprehensive methods like shotgun metagenomic sequencing offer a distinct advantage, by delivering a broader and more detailed view of the genetic diversity that is present within a microbial community, thus enabling in-depth analysis of microbial genera and species, their genes, gene families, predicted enzymatic activities and biochemical pathways, and functional correlations within the microbial ecosystem [24]. However, DNA extracted from tonsil swabs can contain high percentage of host DNA greatly increasing the amount of sequence data needed to get coverage of microbial genomes.

The microbiota of human and animal body sites has been previously recognized as a source of potentially novel and bioactive secondary metabolites [6, 25]. The

ability to produce natural products can provide adaptive and competitive advantages within different environments, and strains producing interesting bioactive metabolites may form a relevant resource for pharmaceutical drug discovery [26]. Whole-genome sequencing of 45 strains belonging to 13 species from the genera *Brevibacterium*, *Corynebacterium*, *Streptococcus*, *Staphylococcus*, *Escherichia*, *Neisseria*, *Pelistega*, and *Pediococcus* revealed the presence of 164 BGCs, predicted to produce compounds with antimicrobial activities. Twelve of these 45 strains displayed inhibitory activity in vitro (Table 3; Fig. 6). Eight of the species carrying BGCs predicted to produce antibacterials did not show any inhibitory activity. This could be due to (i) lack of expression under our culture conditions, (ii) lack of activity against the target bacteria tested in our study, or (iii) because the produced secondary metabolites had functions other than antimicrobial activity [27]. The proposed novel species *C. kozikiae* strain 76QC2CO is a promising example of the potential of the tonsil microbiota in the discovery of commensal isolates able to antagonize pathogens such as *S. suis*. *C. kozikiae* strain 76QC2CO was predicted to produce the lanthipeptide flavucin variant IF76 and showed consistent in vitro antimicrobial activity against five of the gram-positive target bacteria (Table 3) and other features such as thermostability and no cytotoxicity to mammalian cells.

We could not demonstrate a negative correlation between abundance of the flavucin IF76-producing *C. kozikiae* and *S. suis* because the former species is not present in the 16S rRNA gene V3–V4 amplicon reference database but might be part of the "unclassified" group. As discussed above, it might be possible to demonstrate a negative correlation by shotgun metagenomics sequencing. However, a high amount of sequencing data is likely to be required to achieve sufficient coverage of microbial genomes. Even though flavucin 1F76 expressing *C. kozikiae* inhibits *S. suis* in vitro, a negative correlation may not be observed because in oral biofilms some species interact more closely than others due to the different growth conditions created in biofilms and the spatial structure of the microbial communities in the tonsil biofilm.

Sixty-two BGCs, including NRPSs, T3PKSs, and clusters producing siderophores, were identified in all five species of *Staphylococcus* that we cultured. *S. aureus* and *S. hyicus* include opportunistic pathogenic strains that reside in the upper respiratory tract and skin of pigs [28, 29]. Both species have been associated with the production of antimicrobials with activity against other (pathogenic) bacteria, and we observed that our cultured isolates from these species showed inhibitory activity in our antagonism screens (Table 3). In the genome sequences from our cultured *S. aureus* strains, antiSMASH predicted clusters producing the siderophores staphyloferrin and staphylobactin. The production of siderophores in *S. aureus* increased the survival rates of *S. aureus* strains in iron-deficient host environments compared to mutants lacking the genes necessary for siderophore production [30] and provided adaptive advantages in the colonization of niches low in iron [30]. Some *S. aureus* strains also produce class II bacteriocins, such as aureocin and lanthibiotics (Bac1829, C55a/b, and MP1102), which are active against a wide range of gram-positive and gram-negative bacteria including staphylococci, *S. suis*, *Corynebacterium pseudotuberculosis*, *Haemophilus parasuis*, and *Pasteurella multocida* [31–33]. Twenty-six other BGCs were predicted in 8 *S. suis* strains and 1 *S. gallolyticus* strain from our culture collection (Fig. 7). S-Adenosyl-L-methionine (rSAM), a RiPP-associated radical molecule that has been identified by antiSMASH in most genomes from these nine streptococcal strains, may be involved in the biosynthesis of natural products that function as chemical messengers in involved in the biosynthesis of naturalproducts that function as chemical messengers in quorum sensing and intraspecies communication [34, 35].

Most of the genomes from our cultured tonsillar microbiota isolates with in vitro antimicrobial activity contained BGCs that had previously been described in microbiota from animal and human body sites (Additional Table 4). Nonetheless, 68.6% of BGCs identified in the genomes from tonsillar bacteria that we cultured had less than 10% homology or no homology with known BGCs. The natural products produced by these BGCs might provide lead molecules that could be developed into novel therapeutic molecules for the treatment of infectious diseases.

Material and methods

Animals and swab samples

A total of 101 samples were obtained from 101 piglets collected on 15 farms in 3 countries: Germany, the Netherlands, and Spain (Table 1). The 101 healthy piglets were randomly selected 1 week before weaning (3 weeks old; timepoint − 1) and 3 weeks after weaning (timepoint + 3), and tonsillar biofilms were collected from the surface of the palatine tonsil using Puritan HydraFlock Swabs (Daklapack, Europe). Samples were collected from piglets between 3 and 8 weeks old. After sampling, swabs were kept on ice, and upon reaching the laboratory, they were stored at − 80 °C until further processing. Samples from five animals from three farms (ES1, ES2, and NL1) were used for culturomics. Swabs used for culturomics were transferred to a tube containing 2 mL of sterile buffered peptone water (BPW), a nonselective preenrichment

medium, containing 15% glycerol used to preserve bacterial viability. Samples were stored at − 80 °C within 4 h of collection. Swabs used for DNA extraction and microbiota analyses were placed directly after sampling into 2-mL vials containing garnet beads with 1 mL of PowerBead Solution (Qiagen).

16S rRNA gene amplicon sequencing of microbiota samples

The V3–V4 region of the 16S rRNA gene was amplified with primers 341F (CCTAYGGGRBGCASCAG) and 806R (GGACTACNNGGGTATCTAAT) and 250-bp paired-end sequencing with Illumina HiSeq 2500 or NovaSeq 6000 sequencing platforms. Reads were trimmed with cutadapt 2.3 [36] using default settings before being processed in the DADA2 [12] v1.4 workflow for paired-end data. Taxonomy was assigned with the SILVA database v132 [13]. Amplicon sequence variants (ASVs) with taxonomic assignment as eukaryotes or chloroplasts were discarded, and read counts were rarefied to the minimum library size (51,577 reads). Alpha and beta diversity were calculated using R packages phyloseq [37] and vegan [38], and the adonis function in vegan was used to perform permutational analysis of variance (PERMANOVA) test; this multivariate ANOVA calculates the strength and statistical significance of sample groupings based on a dissimilarity matrix. The criteria used to define the core microbiota were a minimum abundance of 0.1% per sample and at least 90% prevalence over the 101 samples.

Isolation and culture conditions

For culturomics, swabs in transport medium and 15% glycerol were taken from storage at − 80 °C, and for each sample, dilutions between 10^{-3} and 10^{-6} were plated onto eight different culture media: sheep blood agar (SBA) (Becton Dickinson, Heidelberg/Germany), brain heart infusion (BHI) (Becton Dickinson, Heidelberg/Germany), M17 (Sigma–Aldrich/Germany), de Man, Rogosa, and Sharpe (MRS) (VWR International, Leuven/Belgium), Mueller Hinton (MH) (Oxoid Ltd., Basingstoke), Todd-Hewitt Broth (THB) (Oxoid, Basingstoke/UK), Luria–Bertani (LB) medium (Merck, Darmstadt/Germany), and MacConkey (MC) medium (Merck, Darmstadt/Germany). Single colonies were picked and transferred to the corresponding liquid medium. Pure cultures were grown on solid 1.5% w/v agar medium and incubated at 37 °C under aerobic conditions with and without 5% CO_2 or at 37 °C in an anaerobic chamber (BACTRON, Sheldon Manufacturing) between 18 and 48 h.

Bacterial identification

MALDI-TOF MS was used first for taxonomic identification of pure cultures. To this end, samples were spotted in target grids and covered with 1 µL of matrix solution (saturated α-cyano acid-4-hydroxycinnamic acid in 50% acetonitrile and 2.5% trifluoroacetic acid). An isolate was considered correctly identified at the species level when the colonies' spectra had a score ≥ 1.9 and identified at the genus level when the spectra had a score ≥ 1.7 [39]. Isolates for which whole-genome sequences were available had 16S RNA genes identified using Barrnap (version 0.9; https://github.com/tseemann/barrnap). When no genome sequence was available, taxonomic assignment was performed by sequencing near-full-length (> 1100 bp) amplicons of the 16S rRNA gene obtained by colony PCR with the universal primers 27F (5′-AGAGTTTGATCMTGG-3′) and 1492R (5′-ACGGGCGGTGTGTRC-3′). The following conditions were used for amplification using a Bio-Rad thermocycler: 1 cycle of 98 °C for 30 min, followed by 35 cycles of 98 °C for 10 min, 60 °C for 15 min, 72 °C for 20 min, and 1 cycle of 72 °C for 2 min. Prior to colony PCR, colonies were resuspended in 40 µL of lysis buffer, and 2 µL was used as a template to which 5 µL of Q5 reaction buffer (5×), 0.5 µL of dNTPs (10 mM), 1.25 µL of each primer (10 µM), and 0.25 µL Q5 Hot-Start High-Fidelity DNA Polymerase (New England Biolabs Inc.) were added and nuclease-free water to an end volume of 25-µL reaction mix. Expected sizes of PCR products were confirmed on a 1.5% agarose gel using a 1-kb DNA Ladder (Invitrogen by Thermo Fisher Scientific). For each isolate, the corresponding 16S rDNA sequence was searched against 16S rDNA sequences provided by the NCBI database by BLAST [40], and taxonomy was assigned using the SILVA database v132 using the criteria below. For phylogenetic analysis, multiple alignments of 16S rRNA gene sequences were generated using MUSCLE [41], and the corresponding phylogenetic trees were generated using maximum likelihood in FastTree v2.1 [42] (1000 replicates). Consensus trees were visualized and annotated using the interactive Tree Of Life (iTOL) v5 [43].

Novel taxa

For the initial delineation of novel species, we used BLAST similarity searches of full-length 16S rRNA genes at a threshold of 98.7% sequence identity. The Genome-to-Genome Distance Calculator 3.0 (GGDC) (http://ggdc.dsmz.de) was used to estimate genome-based delineation of species by digital DNA–DNA hybridization (dDDH), and FastANI [44] was used to calculate the average nucleotide identity (ANI) with a threshold of 95% [45, 46]. Species-level delineation was further confirmed by membrane fatty acid profiles [47]. For two candidate

novel species, under their respective optimal growth conditions, cellular fatty acid profiles were determined at the Leibniz Institute DSMZ. Approximately, 300 mg (wet weight) of cell biomass was extracted according to the standard protocol of the Microbial Identification System (MIDI Inc., version 6.21). The composition of cellular fatty acids was identified by comparison with the TSBA6 6.21 library.

- *Brevibacterium moorei* sp. nov.: *B. moorei* was named in honor of the African-American Dr. Ruth Ella Moore for her contribution to the field of bacteriology.
- *Corynebacterium kozikiae* sp. nov.: *C. kozikiae* was named in honor of the African-American microbiologist Dr. Ariangela J. Kozik for her contribution to human respiratory microbiome research.

Inhibition assays and crude extraction of antimicrobial peptides

Pathogen inhibition screening (or antagonism) assays were performed using overnight (o/n) liquid cultures generated from pure bacterial colonies of the 518 cultured tonsil isolates. Five microliters of each o/n culture was spotted onto the appropriate agar medium to allow growth of bacterial colonies. Agar media with inoculated colonies were then overlaid with soft agar (0.75% w/v agar) that contained approximately 10E+05 CFU/mL of one of the following: disease-associated *S. suis* P1/7, *S. suis* S10 and its isogenic unencapsulated mutant J28, *S. aureus* ATCC 6538P, *Pasteurella aerogenes* DMS 21448, *M. pluranimalium* DSM 22804, *Escherichia coli* L4242, or a *ΔtolC* mutant derivative from *E. coli* MG1061 that does not produce the outer membrane channel protein TolC, a key component of multidrug efflux and type I secretion transporters [48]. We also checked for inhibition of three streptococci that commonly occur in the microbiota of healthy piglets: *Streptococcus porci* strain DSM23759, *Streptococcus porcinus* strain DSM20725, and *Streptococcus parasuis* strain DSM29126. The antimicrobial activity of cultured strains was determined by the presence of clearly visible zones of growth inhibition around the colonies after overnight incubation.

To further characterize the activity identified in *C. kozikiae* strain 76QC2CO, glycerol stocks of isolate 76QC2CO were subcultured in BHI media overnight at 37 °C. From this overnight culture, BHI agar plates were inoculated and incubated for 48 h at 37 °C to obtain a lawn of growth. A volume of 5-mL sterile NaCl 0.9% solution was added, and bacteria were detached from the agar surface by 5-min incubation at room temperature on a horizontal shaker followed by gentle scraping with a sterile plastic T-shaped spreader (VWR International, Amsterdam, the Netherlands). The cell suspension was collected in sterile 15-mL Falcon tubes and centrifuged for 15 min at 3200×g at 4 °C. Bacterial cell pellets were resuspended in 70% 2-propanol (IPA) and 0.1% trifluoracetic acid (TFA) and stirred for 6 h at 4 °C. The IPA-TFA cell extract (CE) was obtained by centrifugation for 15 min at 10,000×g at 4 °C and pressed through a 0.2-µm sterile syringe filter. To partially purify the antimicrobial peptide from the CE, the IPA-TFA fraction was removed by rotary evaporation, and the remaining sample was used for agar diffusion assays and thermal sensitivity and cytotoxicity assays. BHI agar (0.75% w/v agar) was cooled to 50 °C and seeded with a selected target bacterium (approximately 1×10^6 CFU mL^{-1}) (Table 3). A volume of 25 mL was transferred to sterile Petri dishes, and after solidification, wells (4.6-mm diameter) were made using a disposable sterile Pasteur pipette (Corning®). A volume of 50 µL of the CE was dispensed into the wells, and the plates were incubated at 37 °C overnight. Antimicrobial activity was determined by the presence and size of clearly visible zones of growth inhibition around the wells as described previously [49].

Genome sequencing and genome mining for antimicrobial peptide-encoding genes

Whole-genome sequencing (WGS) was carried out for 45 tonsillar isolates that were selected based on the presence of inhibition zones and differences in colony color, size, and shape. Isolates that had shown antimicrobial activity against target bacteria were selected for WGS and genome mining. Prior to DNA extractions, strains were grown overnight in BHI broth at 37 °C. Genomic DNA was extracted using the PowerSoil Genomic Purification Kit (Qiagen) according to the manufacturer's recommended protocol. Recovery of high-molecular-weight DNA was assessed on a 0.8% agarose gel (Sigma–Aldrich) in 1×TAE buffer [Tris–HCl 40 mM, 20-mM acetic acid, and 1-mM EDTA (pH 8)]; DNA was stained with 25 µg/mL SYBR Safe and quantified using the Qubit dsDNA Broad-Range (Invitrogen) assay and Invitrogen Qubit Fluorometer (Thermo Fisher Scientific).

We obtained full genome sequences for 45 tonsil isolates from 5 piglets randomly selected from different litters on 2 Spanish farms with a high health status. Genome sequencing was performed on an Illumina HiSeq 2000 platform (Illumina, Inc.) at MicrobesNG, Birmingham, UK. Reads were trimmed using Trimmomatic 0.30 software with a sliding window set at Q15 [50]. Genome assembly was performed using SPAdes version 3.7 [51]. All genomes were annotated using Prokka 1.14.6 [52]. For identification and annotation of biosynthetic

gene clusters (BGCs), draft genomes of the 45 bacterial strains were analyzed using antiSMASH 6.0.1 [53] using default analysis settings. BAGEL4 was used to mine the genomes for RiPPs and their precursor peptides [54].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40168-025-02064-3.

Additional file 1. Shannon alpha diversity of the core genera microbiota between countries.

Additional file 2. Core genera.

Additional file 3. The average nucleotide identity (ANI) between the three *B. moorei* strains and 21 available genomes of other *Brevibacterium* species and between the *C. kozikiae* and 85 other genomes of *Corynebacterium* species.

Additional file 4. Includes: Additional Table 2 -Taxonomy ID and 16S rRNA gene sequence at Genbank and Additional Table 4: antiSMASH results for 164 BGCs in 45 isolates

Acknowledgements
We thank IRTA, Programa de Sanitat Animal, Centre de Recerca en Sanitat Animal (CReSA) in Spain, Schothorst Feed Research in the Netherlands and anonymized farms that participated in field sampling and contributed to the success of this study, and Marian Figge, from the Westerdijk Fungal Biodiversity Institute (Utrecht, the Netherlands), for her support in the nomeclature and getting the isolates archived.

Authors' contributions
I.M.F.O., P.v.B., and J.M.W. conceptualized the study. I.M.F.O. and S.F., performed the data analysis and interpretation of results under the supervision of M.F.G., H. H, J.B., P.v.B. and J.M.W. The experiments and culturomic approach were performed by I.M.F.O. The microbiota analyses performed by S.F. The manuscript was written by I.M.F.O., P.v.B., J.M. W with input from all coauthors. All authors contributed to the article and approved the final manuscript.

Funding
The authors gratefully acknowledge funding from the Program for Innovative Global Prevention of ***Streptococcus suis*** (PIGSs) Project, supported by European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement no. 727966, and CARTNET project MSCA-ITN-ETN European Training Network under Grant Agreement ID 765147.

Data availability
No datasets were generated or analysed during the current study.

Ethics approval and consent to participate
This study uses samples obtained for diagnostic procedures performed according to the ethical principles and guidelines covered by EU Directive 2010/63/EU.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Host-Microbe Interactomics Group, Animal Sciences Department, Wageningen University and Research Centre, Wageningen, The Netherlands. [2]Department of Medical Microbiology and Infection Prevention, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [3]Department of Veterinary Medicine, University of Cambridge, Cambridge, UK.

Received: 3 December 2023 Accepted: 10 February 2025
Published online: 29 March 2025

References

1. Donia MS, Cimermancic P, Schulze CJ, Brown LCW, Martin J, Mitreva M, et al. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. Cell. 2014;158(6):1402–14. https://doi.org/10.1016/j.cell.2014.08.032.
2. Nowrotek M, Jałowiecki Ł, Harnisz M, Płaza GA. Culturomics and metagenomics: in understanding of environmental resistome. Front Environ Sci Eng. 2019;13(3):40.
3. Zipperer A, Konnerth MC, Laux C, Berscheid A, Janek D, Weidenmaier C, et al. Human commensals producing a novel antibiotic impair pathogen colonization. Nature. 2016;535(7613):511–6.
4. Wyatt MA, Wang W, Roux CM, Beasley FC, Heinrichs DE, Dunman PM, Magarvey NA. Staphylococcus aureus nonribosomal peptide secondary metabolites regulate virulence. Science. 2010;329(5989):294–6.
5. LeBel G, Vaillancourt K, Frenette M, Gottschalk M, Grenier D. Suicin 90–1330 from a nonvirulent strain of Streptococcus suis: a nisin-related lantibiotic active on gram-positive swine pathogens. Appl Environ Microbiol. 2014;80(17):5484–92.
6. Barber CC, Zhang W. Small molecule natural products in human nasal/oral microbiota. J Industr Microbiol Biotechnol. 2021;48(3–4):kuab010.
7. Kernaghan S, Bujold AR, MacInnes JI. The microbiome of the soft palate of swine. Anim Health Res Rev. 2012;13(1):110–20.
8. Horter DC, Yoon K-J, Zimmerman JJ. A review of porcine tonsils in immunity and disease. Anim Health Res Rev. 2003;4(02):143–55.
9. Obradovic MR, Segura M, Segalés J, Gottschalk M. Review of the speculative role of co-infections in Streptococcus suis-associated diseases in pigs. Vet Res. 2021;52(1):49.
10. Fredriksen S, Neila-Ibáñez C, Hennig-Pauka I, Guan X, Dunkelberger J, de Oliveira IF, et al. Streptococcus suis infection on European farms is associated with an altered tonsil microbiome and resistome. Microb Genom. 2024;10(12):001334.
11. Niazy M, Hill S, Nadeem K, Ricker N, Farzan A. Compositional analysis of the tonsil microbiota in relationship to Streptococcus suis disease in nursery pigs in Ontario. Animal Microbiome. 2022;4(1):10.
12. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13(7):581–3.
13. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2012;41(D1):D590–6.
14. Oliveira IMFd, Ng DYK, van Baarlen P, Stegger M, Andersen PS, Wells JM. Comparative genomics of Rothia species reveals diversity in novel biosynthetic gene clusters and ecological adaptation to different eukaryotic hosts and host niches. Microb Genom. 2022;8(9):mgen000854.
15. Martínez B, Rodríguez A, Suárez JE. Lactococcin 972, a bacteriocin that inhibits septum formation in lactococci. Microbiology. 2000;146(4):949–55.
16. van Heel AJ, Kloosterman TG, Montalban-Lopez M, Deng J, Plat A, Baudu B, et al. Discovery, production and modification of five novel lantibiotics using the promiscuous nisin modification machinery. ACS Synth Biol. 2016;5(10):1146–54.
17. Cortes LCP, LeVeque RM, Funk J, Marsh TL, Mulks MH. Development of the tonsillar microbiome in pigs from newborn through weaning. BMC Microbiol. 2018;18(1):35.
18. Lowe BA, Marsh TL, Isaacs-Cosgrove N, Kirkwood RN, Kiupel M, Mulks MH. Microbial communities in the tonsils of healthy pigs. Vet Microbiol. 2011;147(3–4):346–57.
19. Lowe BA, Marsh TL, Isaacs-Cosgrove N, Kirkwood RN, Kiupel M, Mulks MH. Defining the "core microbiome" of the microbial communities in the tonsils of healthy pigs. BMC Microbiol. 2012;12(1):20.
20. MacInnes JI, Gottschalk M, Lone AG, Metcalf DS, Ojha S, Rosendal T, et al. Prevalence of Actinobacillus pleuropneumoniae, Actinobacillus suis, Haemophilus parasuis, Pasteurella multocida, and Streptococcus suis in representative Ontario swine herds. Can J Vet Res. 2008;72(3):242.
21. Zbinden R, Graevenitz AV. Actinobacillus, Capnocytophaga, Eikenella, Kingella, Pasteurella, and other fastidious or rarely encountered gram-negative rods. Manual Clin Microbiol. 2011;11:574–87.
22. Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, Kurm V, et al. Where less may be more: how the rare biosphere pulls ecosystems strings. ISME J. 2017;11(4):853–62.

23. Lynch MDJ, Neufeld JD. Ecology and exploration of the rare biosphere. Nat Rev Microbiol. 2015;13(4):217–29.
24. Durazzi F, Sala C, Castellani G, Manfreda G, Remondini D, De Cesare A. Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. Sci Rep. 2021;11(1):3030.
25. Wylensek D, Hitch TCA, Riedel T, Afrizal A, Kumar N, Wortmann E, et al. A collection of bacterial isolates from the pig intestine reveals functional and taxonomic diversity. Nat Commun. 2020;11(1):6389.
26. Kenshole E, Herisse M, Michael M, Pidot SJ. Natural product discovery through microbial genome mining. Curr Opin Chem Biol. 2021;60:47–54.
27. Hoskisson PA, Fernández-Martínez LT. Regulation of specialised metabolites in Actinobacteria–expanding the paradigms. Environmental microbiology reports. 2018;10(3):231–8.
28. Wang L, Wu Z-w, Li Y, Dong J-g, Zhang L-y, Liang P-s, et al. Profiling and identification of novel immunogenic proteins of Staphylococcus hyicus ZC-4 by immunoproteomic assay. PLoS ONE. 2016;11(12):e0167686.
29. Asadi S, Jamali M. Assessment the frequency of Staphylococcus aureus golden methicillin- resistant (MRSA) and vancomycin-resistant VRSA in determining the MIC using E-test. Immunol Disorder Immunother. 2017;02(01):1–3.
30. Dale SE, Doherty-Kirby A, Lajoie G, Heinrichs DE. Role of siderophore biosynthesis in virulence of Staphylococcus aureus: identification and characterization of genes involved in production of a siderophore. Infect Immun. 2004;72(1):29–37.
31. Netz DJA, Bastos MdCdF, Sahl H-G. Mode of action of the antimicrobial peptide aureocin A53 from Staphylococcus aureus. Appl Environ Microbiol. 2002:68(11):5274–80.
32. Crupper SS, Iandolo JJ. Purification and partial characterization of a novel antibacterial agent (Bac1829) produced by Staphylococcus aureus KSI1829. Appl Environ Microbiol. 1996;62(9):3171–5.
33. Zhao F, Yang N, Wang X, Mao R, Hao Y, Li Z, et al. In vitro/vivo mechanism of action of MP1102 with low/nonresistance against Streptococcus suis type 2 strain CVCC 3928. Front Cell Infect Microbiol. 2019;9:48.
34. Benjdia A, Balty C, Berteau O. Radical SAM enzymes in the biosynthesis of ribosomally synthesized and post-translationally modified peptides (RiPPs). Front Chem. 2017;5:87.
35. Schramma KR, Bushin LB, Seyedsayamdost MR. Structure and biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. Nat Chem. 2015;7(5):431–7.
36. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal. 2011;17(1):10–2.
37. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE. 2013;8(4): e61217.
38. Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J, Borman T. vegan: Community Ecology Package. R package version 2.6-10. 2025. https://github.com/vegandevs/vegan .
39. Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nat Microbiol. 2016;1(12):1–8.
40. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):421.
41. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5(1):113.
42. Price MN, Dehal PS, Arkin AP. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS ONE. 2010;5(3): e9490.
43. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256–9.
44. 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):5114.
45. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol. 2014;12(9):635–45.
46. Hugenholtz P, Chuvochina M, Oren A, Parks DH, Soo RM. Prokaryotic taxonomy and nomenclature in the age of big sequence data. ISME J. 2021;15(7):1879–92.
47. Kaneda T. Iso-and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. Microbiol Rev. 1991;55(2):288–302.
48. Hockett KL, Baltrus DA. Use of the soft-agar overlay technique to screen for bacterially produced inhibitory compounds. JoVE (Journal of Visualized Experiments). 2017;119: e55064.
49. Usha Nandhini S, Sudha S, Anusha Jeslin V, Manisha S. Isolation, identification and extraction of antimicrobial compounds produced by Streptomyces sps from terrestrial soil. Biocatal Agric Biotechnol. 2018;15:317–21.
50. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
51. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.
52. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068–9.
53. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, et al. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. 2019;47(W1):W81-W7.
54. van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. Nucleic Acids Res. 2018;46(W1):W278–81.

Publisher's Note