communications biology

Article

https://doi.org/10.1038/s42003-024-07217-y

Uncovering the arsenal of class II bacteriocins in salivarius streptococci

Check for updates

Julien Damoczi[1], Adrien Knoops[1], Marie-Sophie Martou[1], Félix Jaumaux[2], Philippe Gabant[2], Jacques Mahillon[3], Jan-Willem Veening[4], Johann Mignolet[1,4,5] & Pascal Hols[1,5]

Facing the antibiotic resistance crisis, bacteriocins are considered as a promising alternative to treat bacterial infections. In the human commensal *Streptococcus salivarius*, the production of unmodified bacteriocins (or salivaricins) is directly controlled at the transcriptional level by quorum-sensing. To discover hidden bacteriocins, we harnessed here the unique molecular signatures of salivaricins not yet used in available computational pipelines and performed genome mining followed by orthogonal reconstitution and expression. From 100 genomes of *S. salivarius*, we identified more than 50 bacteriocin candidates clustered into 21 groups. Strain-based analysis of bacteriocin combinations revealed significant diversity, reflecting the plasticity of seven independent loci. Activity tests showed both narrow and broad-spectrum bacteriocins with overlapping activities against a wide panel of Gram-positive bacteria, including notorious multidrug-resistant pathogens. Overall, this work provides a search-to-test generic pipeline for bacteriocin discovery with high impact for bacterial ecology and broad applications in the food and biomedical fields.

Antimicrobial resistance is one of the biggest current challenges facing humankind. In 2019, an estimated 1,270,000 people died from drug-resistance infections[1]. According to the World Health Organization, antibiotic resistance is one of the largest threats to global health. By 2050, more than 10 million deaths are forecasted due to antimicrobial resistance[2]. With the surge of antibiotic resistance, there is a critical need for alternatives to treat bacterial infections. Diverse promising treatments are in development to curtail that phenomenon, such as bacteriophage therapy[3], probiotics[4], and RNA therapeutics[5]. Another alternative is the use of bacteriocins[6–9]. These small antimicrobial peptides, ribosomally synthesized by bacteria, are active against Gram-positive or Gram-negative bacteria but do not affect the producer, which typically encodes a specific immunity mechanism[10]. So far, around a thousand bacteriocins are inventoried in the main bacteriocin database of BAGEL4[11]. Yet, 99% of bacteria are predicted to produce at least one type of bacteriocin, implying that the bacteriocin reservoir available for bacterial infection treatment is massive[12].

Bacteriocins from Gram-positive bacteria are classified into two main groups: class I gathers peptides that undergo a range of post-translational modifications, while class II includes peptides that are unmodified or contain only minor modifications (*e.g.*, disulfide bridges)[13]. Class II bacteriocins are subdivided in four subclasses: classes IIa, IIb, IIc, and IId that correspond to pediocin-like, two-peptide, leaderless, and non-pediocin-like single peptide bacteriocins, respectively[13]. Except class IIc bacteriocins, the majority of class II bacteriocins are characterized by the presence of a leader sequence with a conserved addressing GG motif ($[M|L|V]X_4GG$)[14]. This sequence is essential for the secretion of the bacteriocin outside the cell and maintains the peptide in an inactive form to avoid intracellular toxicity issues[15]. This work will focus on bacteriocins produced by *Streptococcus salivarius*, a Gram-positive lactic acid bacterium and a dominant commensal of the oral cavity[16]. Several studies identified bacteriocinogenic strains of this species as highly potent against pathogenic bacteria. For that reason, several strains are already used as probiotics, such as M18[17], K12[18], and 24SMB[19]. In *S. salivarius*, six class I bacteriocins (SlvA, SlvB, SlvD, SlvE, SlvG32, and Slv9)[20,21] and six class II bacteriocins (BlpK, SlvV, SlvW, SlvX, SlvY, and SlvZ)[21,22] have been identified so far.

In the genus *Streptococcus*, bacteriocin production is controlled by a cell-to-cell communication strategy known as quorum sensing[22–24]. The regulation of class I and class II bacteriocins differs and is species-dependent. Class I bacteriocins as such act as stimulating pheromones and induce their own production through a phosphorylation cascade involving a dedicated two-component system[25,26]. In contrast, the production of class II bacteriocins is regulated by unmodified peptide pheromones that normally do not exhibit direct toxic effects[27]. In streptococci, two regulation systems (*i.e.*,

[1]Biochemistry and Genetics of Microorganisms, Louvain Institute of Biomolecular Science and Technology, Université catholique de Louvain, Louvain-la-Neuve, Belgium. [2]Syngulon, Seraing, Belgium. [3]Laboratory of Food and Environmental Microbiology, Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium. [4]Department of Fundamental Microbiology, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland. [5]These authors contributed equally: Johann Mignolet, Pascal Hols. e-mail: Pascal.Hols@uclouvain.be

ComRS and BlpRH) control the production of class II bacteriocins[27]. In *S. salivarius*, the ComRS system synchronizes competence and predation through upregulation of *comX* and bacteriocin genes, respectively[22]. The precursor ComS is produced, secreted, and matured in the pheromone XIP (*comX*-inducing peptide). At a threshold concentration, XIP is internalized and interacts with the cytoplasmic sensor ComR[28–30]. Then, the complex ComR-XIP activates transcription of target genes through its binding to a well-conserved DNA sequence called ComR-box[22,30]. Through the combined production of bacteriocins and bacteriolytic enzymes, this predation-competence coupling allows the bacteria to ensure the presence of free DNA in the environment before entering the competence state[22,31–34]. In other streptococcal species, the production of bacteriocins is generally under the control of the BlpRH system[23,24,35], which is absent or inactive in *S. salivarius*[22]. The pheromone precursor BlpC is produced, processed at a double-glycine motif, and exported by the dedicated transporter BlpAB or through the competence transporter ComAB[23,36]. When the mature pheromone BlpC* reaches a threshold concentration, it activates the two-component system BlpRH[23]. Upon activation, the trans-membrane histidine kinase BlpH phosphorylates the cytoplasmic response regulator BlpR, which in turn upregulates genes of the *blp* locus, including bacteriocin and self-immunity genes[23,35,37,38].

Generally, the discovery of novel bacteriocins is easier for class I compared to class II. Since class I bacteriocins and modification enzymes are encoded in the same locus, in silico analyses are based on the identification of putative bacteriocin and processing enzyme genes in close genomic vicinity[39,40]. In contrast, class II bacteriocins are generally encoded by small isolated (and usually unannotated) genes, either scattered throughout the genome (e.g., *S. salivarius*)[22] or clustered in a unique locus (e.g., *Streptococcus pneumoniae*)[37]. Consequently, dedicated bioinformatic pipelines need to be developed for their discovery.

In this work, we set up a combined approach based on bioinformatics and synthetic biology to identify undisclosed class II bacteriocins in streptococci. We coupled the in silico recognition of regulatory DNA motifs with specific properties of known bacteriocins to unveil hidden candidates. By testing our pipeline on *S. salivarius* genomes, we discovered 21 bacteriocin groups with multiple variants, including 13 groups never reported before. Using in vivo and in vitro expression systems, we showed that unearthed bacteriocins display various profiles of inhibition against a range of Gram-positive bacteria, including a series of clinically-relevant pathogenic species (e.g., *S. pneumoniae*, *Staphylococcus aureus* and *Enterococcus faecium*).

Results

In silico identification of 21 groups of class II bacteriocins in *S. salivarius*

To disclose the class II bacteriocin content (defined hereafter as the bacteriocinome) from *S. salivarius*, we first selected 100 genomes that were either fully sequenced or displaying a low contig number. The corresponding strains were isolated from 15 different sources (mainly human oral cavity, feces, and blood), across 13 different countries (Supplementary Table 1). We next designed an appropriate multiparametric bioinformatic pipeline (Fig. 1 and Supplementary Software 1). Considering that class II bacteriocins are hard to identify due to their small size and high sequence diversity, we combined three features: regulatory elements, sequence size, and signal sequence property. In several *S. salivarius* strains, the bacteriocin loci are under the direct control of the ComRS system[22]. We therefore scouted the genomes for all open reading frames (ORFs) that include a genuine ComR-box in their upstream genomic region (step 1) (Supplementary Table 2). The ComR-box is a palindromic sequence oriented at the 3' end by a T-rich stretch (T-track)[22,28,29]. Therefore, we restricted the search to all ORFs downstream of the ComR-box T-track at a maximal distance of 4 kb (step 2). Then, we narrowed down the analysis by selecting sequences coding for precursor peptides ranging from 40 to 100 amino acids. According to the BAGEL4 database, this length includes the majority (> 98%) of class II bacteriocin precursors encompassing a GG leader sequence[11] (step 3). We next discarded all the peptide sequences that did not include the conserved sequence motif ([M | L | V]X_4GG) of class II bacteriocins (step 4). We finally retained the peptide sequences with specific lengths for leader sequences (from 14 to 28 amino acids) and associated mature sequences (from 25 to 70 amino acids) (step 5). Then, we manually removed all sequences devoid of a ribosome binding site upstream of the start codon (step 6) (Supplementary Data 1).

We ultimately identified 21 putative bacteriocins, including six from *S. salivarius* (BlpK, SlvV, SlvW, SlvX, SlvY, and SlvZ; positive controls) and two from *Streptococcus thermophilus* (BlpE, and BlpF) that were previously reported[22,41]. Since *S. thermophilus* and *S. salivarius* are two closely related species, the identification of known bacteriocins from *S. thermophilus*

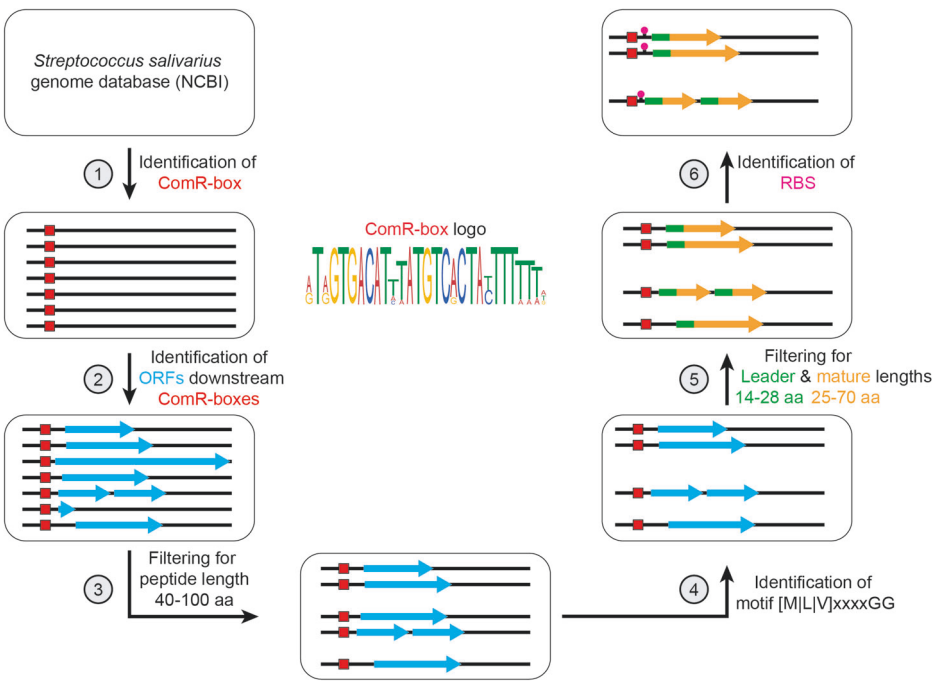

Fig. 1 | Schematic representation of the six steps of the multiparametric bioinformatics pipeline. Red boxes represent the ComR-box and blue arrows indicate the presence of downstream open reading frame(s). Leader sequences, mature sequences, and ribosome binding sites are shown as green squares, orange arrows, and purple dots, respectively.

suggests that they were exchanged through horizontal gene transfer. From this in silico analysis, we unearthed 13 unknown bacteriocin candidates that were named PsnA to PsnM (Pan Salivaricin A to M) (Table 1). We also noticed that several bacteriocins are always encoded in tandem (PsnA-PsnB, PsnC-PsnD, PsnE-PsnF, PsnG-PsnH, and SlvY-SlvZ), suggesting that they could belong to the two-peptide class IIb bacteriocins[42]. Together, these results show that our bioinformatic pipeline can identify a range of hidden class II bacteriocin candidates in *S. salivarius*.

Class II salivaricins are variably distributed among streptococcal species

We next investigated the diversity of identified salivaricins across the genomes of *S. salivarius* and other streptococci that could be found in the oral cavity, such as *S. thermophilus*, *Streptococcus pyogenes*, and *S. pneumoniae* (Fig. 2a). We clustered bacteriocin candidates using BLAST searches and defined variants as homologous peptides with an identity ≥ 50% and at least one different amino acid. In total, we identified more than 50 salivaricin variants (Supplementary Table 3). We observed that ~50% of salivaricin peptides (or their variants) (11/21) are restricted to the species *S. salivarius*, while others are moderately to highly spread (*i.e.*, BlpK) among the three other streptococcal species (Fig. 2a). Strikingly, some bacteriocins are more represented in other species than in *S. salivarius*. For example, BlpK is present in ~25% of *S. salivarius* strains, while it is found in ~75% of strains from *S. thermophilus* and *S. pneumoniae* (Fig. 2a). Although the evolutionary pressure that shapes the specific bacteriocin content for each species remains unclear, we speculate that it is likely related to inter-species competition inside the ecological niche[43]. We also observed a high sequence conservation between variants of each salivaricin within the same species (Supplementary Table 3). For example, BlpK variants display a negative correlation between the phylogenetic distance and the sequence conservation (Fig. 2b and c). This suggests that bacteriocin gene exchanges through horizontal gene transfer between distant streptococcal species are uncommon events.

Together, these analyses show that approximately half of class II salivaricins are restricted to *S. salivarius*, while the other half are variably shared through rare exchange events between streptococci that could be present in the oral cavity.

Bacteriocinome reshuffling results in a high diversity of salivaricin cocktails in *S. salivarius*

We also investigated the salivaricin gene content in all the *S. salivarius* strains that were used during the bioinformatic analysis (Fig. 3 and Supplementary Table 4). First, we observed that 85% of all strains contain between five and seven bacteriocin genes (few strains can harbor up to 10 bacteriocin genes) arrayed in 37 different combinations (Fig. 3a, b). Second, we noticed that several specific bacteriocin genes are more widely distributed than others among *S. salivarius* strains. We identified that SlvY with SlvZ and SlvW are present in 100% and 95% of analyzed strains, respectively (Fig. 3c). In contrast, the bacteriocins PsnK and PsnM are only present in 2% of *S. salivarius* strains (Fig. 3c). We also observed a high co-occurrence for some bacteriocins (e.g., SlvY with SlvZ, PsnA with PsnB), which supports their belonging to the two-peptide class IIb bacteriocins (Fig. 3a). Unexpectedly, we also observed bacteriocin gene exclusion (e.g., PsnL vs. BlpK, PsnI vs. SlvV) (Fig. 3a). This high intra-species variability in salivaricins prompts us to analyze in detail the genomic context of all bacteriocin loci (Fig. 4). We identified seven loci encoding various bacteriocins and immunity proteins, most of them being highly divergent among *S. salivarius* strains (Fig. 4a–g). Analyzing the most variable locus (Fig. 4a), we noticed that the genes *blpK*, *psnL*, *psnJ*, and *psnC-D* are located at the same position, explaining their mutual exclusion at a single genome level (Fig. 4a). Interestingly, we also spotted gene fragments encoding bacteriocin leader sequences without their mature part (Fig. 4a, orange arrows), probably resulting from intra-chromosomal deletion events between directly repeated bacteriocin genes. Bacteriocin loci are highly variable islands framed by two strictly conserved regions. Therefore, we hypothesize that surrounding conserved regions in combination with conserved leader-encoding sequences are involved in double recombination events, potentially mediated by natural transformation. This specific mechanism would account for salivaricin genes intra- (and possibly inter-) species exchanges.

Together, these analyses demonstrate that the cocktail of class II salivaricins is highly variable at the species level and that this diversity partially relies on a specific locus, which is a potential hot spot of recombination.

A majority of class II salivaricin candidates display antibacterial activity

Since our bioinformatics analysis identified 13 salivaricin candidates that were not reported before, we decided to evaluate their activity and prey spectrum. In a first step, we produce them in their mature form using synthetic DNA blocks coupled to an in vitro transcription-translation system, as reported before[44]. Each cell-free sample was directly tested against three reporter strains: *Lactococcus lactis* IL1403, *S. thermophilus* LMD-9, and *S. salivarius* HSISS4. *L. lactis* was chosen for its high sensitivity to a wide range of class II bacteriocins from salivarius group streptococci in spot-on-lawn assays[22,41]. In addition, we used *S. salivarius* and the closely-related species *S. thermophilus* since we expect a narrow spectrum of activity for several salivaricins. Antibacterial activity was identified for 5 out of 13 bacteriocin candidates (Supplementary Fig. 1). In a second step, we decided to assess bacteriocin activity in vivo for all the 21 putative salivaricins. We chose to produce them directly in *S. salivarius* since class II bacteriocins may undergo minor post-translational modifications that require enzymes present in the original host. For instance, 60% of the salivaricin candidates display two cysteine residues, which might be involved in the formation of a disulfide bridge required for their functionality[41] (Table 1). To evaluate the activity of each individual bacteriocin in a single genetic background, we designed fusions between expression-secretion signals of the *blpK* gene and the mature part of each bacteriocin candidate (Fig. 5a). This strategy ensures tight control of expression through the XIP-inducible *blpK* promoter and secretion compatibility between the BlpK leader sequence and the bacteriocin exporter ComA (also XIP-inducible) in *S. salivarius* HSISS4[22] (Fig. 5a). Those fusions were inserted at a permissive ectopic locus in a mutant strain devoid of any residual antibacterial activity (HSISS4 Δ*slv5*)[22].

The collection containing 21 bacteriocin-producing strains was tested for antimicrobial activity against a broad range of Gram-positive bacterial species by spot-on-lawn assays (Fig. 5b and c). Since several bacteriocins belong to the two-peptide class IIb bacteriocins (PsnAB, PsnCD, PsnEF, PsnGH, and SlvYZ), we evaluated their activities by spotting a pool of the two clones encoding each individual peptide (Fig. 5b, c; Supplementary Data 2). In line with the prediction of being two-peptide bacteriocins, no activity was observed when tested individually (Supplementary Fig. 2a). Specifically, the collection was tested against *S. thermophilus*, *Streptococcus vestibularis*, and *S. salivarius* (species from the salivarius group). Other tested streptococci included *S. pneumoniae*, *Streptococcus oralis*, *Streptococcus mitis*, *Streptococcus pyogenes, Streptococcus anginosus*, *Streptococcus sobrinus*, and *Streptococcus mutans*. They are oral streptococci found in different sub-niches of the oral cavity (epithelium surface, oropharynx, dental plaque) and upper airways. Pathobionts from the upper airways (i.e. *Staphylococcus aureus* and *Staphylococcus epidermidis*) and the intestine tract (i.e. *Listeria monocytogenes*, *Enterococcus faecium*, and *Enterococcus faecalis*) were also tested.

We observed that PsnI, PsnL, PsnM, and BlpK have the broadest spectrum of prey species. Except the cariogenic species *S. mutans* and *S. sobrinus* that were insensitive to all tested salivaricins (illustrated for *S. sobrinus* in Supplementary Data 2), all tested oral streptococci were inhibited at a different extent by these four broad-range salivaricins. Remarkably, PsnI, PsnL, and BlpK are active against the pathogenic species *S. pneumoniae* (Fig. 5b, c). As bacteriocin activity could be strain-specific, we evaluated a range of *S. pneumoniae* strains with different serotypes and from different global pneumococcal sequence clusters (Supplementary Table 5 and Data 3), including antibiotic-resistant strains (Fig. 6a). Interestingly, all seven tested strains could be inhibited by PsnI, PsnL, and BlpK. Moreover,

https://doi.org/10.1038/s42003-024-07217-y

Table 1 | Class II bacteriocins of *S. salivarius* identified by in silico analysis

Name	Precursor sequence (aa)	MW (kDa)	GRAVY	pI	Charge	Lys-Arg (%)
PsnA	MLEINMIYSDLSLTE**L**AQIN**GG** ↓ GILSSVAGLVKDTWGTLYSTGRDFGRSVVNAAMP	3.5	0.2	9.3	1	8.8
PsnB	MYREKRNDEELNIND**L**AMIN**GG** ↓ VAPVVIAVGVVATPFVLGAIAGGADEYARKRGHH	3.4	0.7	9.3	2	8.8
PsnC	MTTQTMNNFETLDLEA**L**ANVE**GG** ↓ NRCKDLIFGGALTGAGAGFTGGMAAFGVTAFPGAFVGAHVGAVAGGLACIGASF	5.0	1.0	8.2	1.5	3.7
PsnD	MNVTKYNKMCDAE**L**EIIS**GG** ↓ GKVGEVIGGCLGGMLMAWAAGPVSAGGYTMICATSGLANAYF	4.0	0.9	6.1	0	2.4
PsnE	MKNTFQKIDIKD**L**AIVV**GG** ↓ SIYSSVGYGVGVVTHTVYDFGRGFVDGFRG	3.2	0.3	7.5	0.5	6.7
PsnF	MDTMMMDKFDSLTFDK**L**SEIV**GG** ↓ NVAYDIGYGVGQVSYIAVEILKLRKIR	3.0	0.3	9.7	2	14.8
PsnG	MVLNLENVNCVTLQDEE**L**MMVN**GG** ↓ LAPIVIGGVVVGWKVIGGVAGVIAASGAAGIAAGYYANRP	3.7	1.2	10.0	2	5.0
PsnH	MSKKYELSKKE**L**QEIQ**GG** ↓ VAPIMVAIGLGSLAVASFSSGYKFGTDLARRGR	3.4	0.6	11.4	3	12.1
PsnI	MTTQIINNFNSLNSED**L**SIIE**GG** ↓ GVIGCVAGTAGSAGLGFLTGTSVGTVTFPIVGTVSGGAFGAWSGAGLGMATFCGV	5.0	1.2	5.8	0	0.0
PsnJ	MTTQTINTFETLDFEA**L**ANVE**GG** ↓ KVNWDRLSSSASGAAEGIYFCAASGAPLFTLAPYAIVGCGVVGAGLGYAFPH	5.2	0.6	7.3	0.5	3.8
PsnK	MATKTIENFNTLDFET**L**ASVE**GG** ↓ KVIY*YGNGLYCKSNGGCWV*DWSQTINSILTNSAMNWATRGNAGWHSGGIAP	5.5	−0.3	8.8	2.5	5.9
PsnL	MTTQTMNTFETLDLET**L**ANVE**GG** ↓ GWVKCYAGTIGSALVGSAGGPVGYWGGALVGYATFC	3.5	0.8	8.2	1	2.8
PsnM	MATQTIENFNTLDLET**L**ASVE**GG** ↓ GCSWRGAGGATVQGAIGGAFGGNVVLPVVGSVPGYLAGGVLGGAGGAVAYGATCWWS	5.5	0.7	8.1	1	1.6
SlvV	MTTQIINNFNSLNSED**L**STVK**GG** ↓ ACSFWGATAAVGVGAVGGAIKGGIQTGTWQGAALKGIGYGIKDGITYGLICRY	5.2	0.5	9.3	3	7.5
SlvW	MRTKAYGEELTAET**L**ENVT**GG** ↓ DFVSRPQTLPERLGWNKWWLKRRPPYGD	3.5	−1.3	10.9	3	21.4
SlvX	MKTSEMTQQVLTHQE**L**SVVV**GG** ↓ GCSWNNAAKAALGTSITGLITSGPGGALLGLAGGAVSYGALCWV	4.1	0.8	8.2	1	2.3
SlvY	MINKEMKAAD**L**ASVT**GG** ↓ GWKTNLAIGGLCLASGPIGTMVCLGAYNGYMDSAR	3.6	0.4	8.2	1	5.7
SlvZ	MTKTINNRKNMTTQE**L**EAVS**GG** ↓ VVPWAAISVGIAAAKLTYDLSYAAGKSFYNLTH	3.5	0.6	9.0	1.5	6.1
BlpK	MTTQTMNNFETLDLEA**L**ANVE**GG** ↓ GCSWGGFAKQGVATGVGNGLRLGIKTRTWQGAVAGAAGGAIVGGVGYGATCWW	5.1	0.3	10.3	4	7.5
BlpE	MGTQTIENFNTLDLET**L**ASVE**GG** ↓ RVNWERWGMCGASVAVGATIGGGVALLGC	2.9	0.7	8.1	1	6.9
BlpF	MLTELTYND**L**EVTQ**GG** ↓ AVCRMPNVNSKEGDPLKDGWVSPPKYRSGEAYPMVYIPICAIM	4.8	−0.1	8.2	0	9.3

In the precursor amino-acid (aa) sequences, the predicted cleavage motif ([M | L | V]X_4GG ↓) of the leader sequence and the cysteine residues potentially involved in disulfide bridges are indicated in bold and bod underlined, respectively. The pediocin-like motif YGNG[V | L]$XCX_{4\text{-}5}$CXV of PsnK is indicated in italics. MW, molecular weight (kDa); GRAVY, GRand AVerage of hydropathY, positive and negative values indicate hydrophobic and hydrophilic peptides, respectively; pI, isoelectric point; charge, D and E (-1), H (+ 0.5), K and R (+ 1); and % Lys-Arg, percentage in lysine and arginine. All those parameters were calculated by the bioinformatic script on the mature part of bacteriocins.

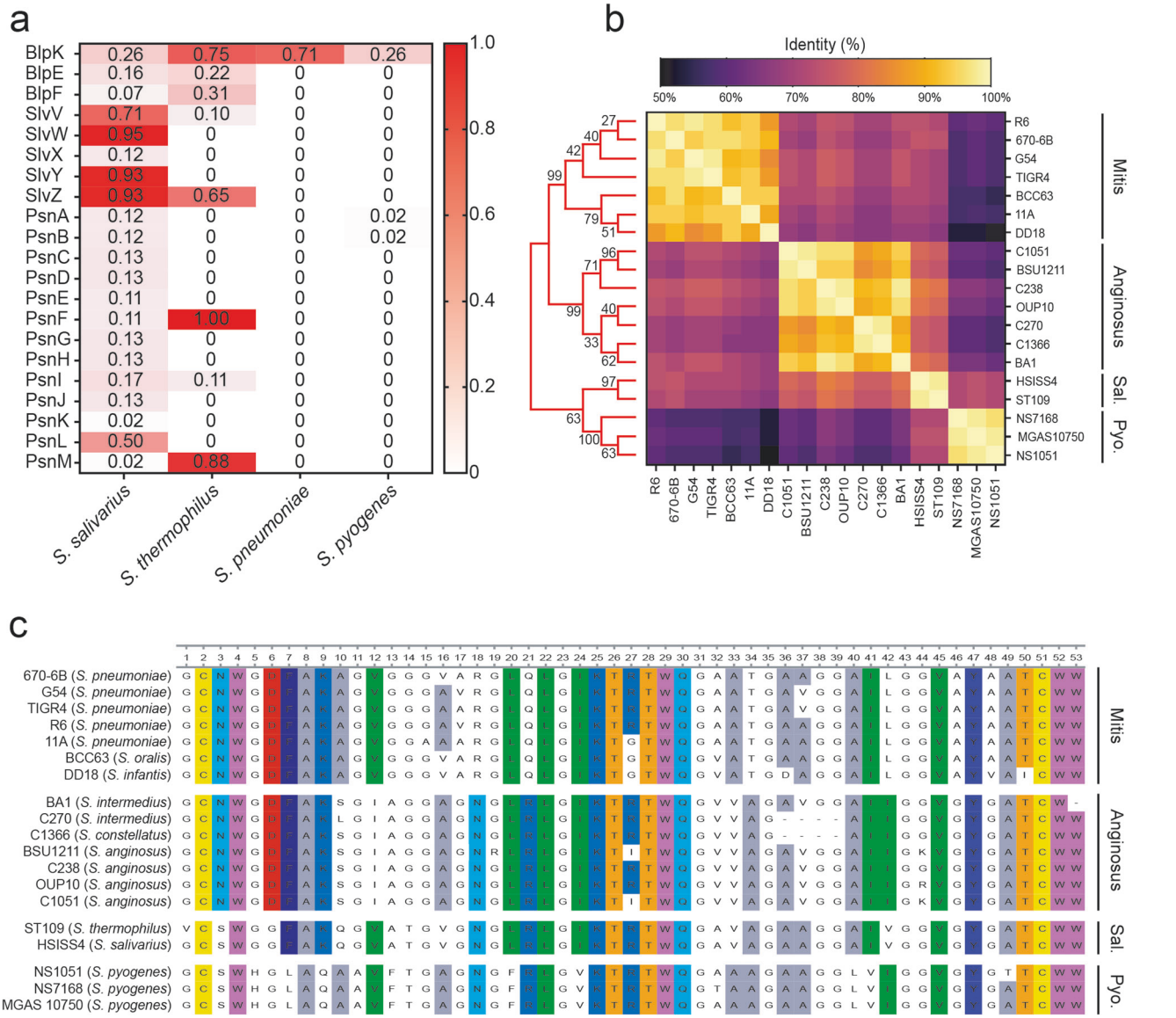

Fig. 2 | Distribution of salivaricins among oral streptococci. a Occurrence of the 21 prototypical salivaricins in *S. salivarius* compared to three streptococcal species of in the oral cavity. The number in each box indicates the frequency of observation of the bacteriocin among all analyzed strains of the species. The analysis was performed on *S. salivarius*, *S. thermophilus*, *S. pneumoniae*, and *S. pyogenes* for 99, 78, 137, and 261 genomes, respectively. **b** Phylogenetic tree and heatmap of sequence identity (%) of natural BlpK variants from various streptococcal strains. The streptococcal groups are indicated on the right (Sal., salivarius; Pyo., pyogenic). BlpK sequences were aligned using MUSCLE and the tree was then generated with MEGA11 using the neighbor-joining method (10,000 bootstrap replicates). Bootstrap values are given at each node. **c** Alignment of natural BlpK variants. The alignment was generated with Clustal Omega and residues were colored (RasMol code) according to physicochemical properties.

these four salivaricins are active against a group of clinically-relevant pathogenic bacteria more phylogenetically distant, such as *Staphylococcus aureus*, *Enterococcus faecium*, and *Listeria monocytogenes*. For instance, PsnL is the most active against *S. aureus*. As reported above for *S. pneumoniae*, we evaluated PsnL activity against a range of antibiotic-resistant *S. aureus* strains (Supplementary Table 5 and Data 3). Out of 12 tested strains, five mono- or multi-antibiotic resistant strains were sensitive to PsnL, including the well-known strain NRS 384 (USA300) (Fig. 6b).

In contrast, several bacteriocins showcased a narrow spectrum of activity. The antimicrobial activity of PsnCD, PsnEF, SlvYZ, PsnJ, and BlpE targeted only streptococcal species, while PsnK was mainly active against *L. monocytogenes*. The anti-listeria activity of PsnK correlates with the presence of a pediocin-like motif (YGNG[V | L]XCX$_{4\text{-}5}$CXV) in its mature sequence, which is a typical feature of class IIa bacteriocins active against the genus *Listeria*[45] (Table 1). Still, PsnAB, BlpF, and SlvW did not show any activity. We can hypothesize that they are inactive, non-bacteriocin peptides, not properly processed, or simply active against bacterial species outside of the tested spectrum.

These results validate a bacterial chassis designed for testing class II bacteriocins in vivo and demonstrate antibacterial activity from a narrow to large spectrum for 17 out of 21 in silico-identified bacteriocin candidates against a large panel of pathobionts.

Enriched salivaricin cocktails potentiate their antibacterial activity

To determine how the composition of bacteriocin cocktails might affect the antibacterial activity of *S. salivarius*, we added PsnJ, PsnK, or PsnL to the native cocktail (BlpK, SlvV, SlvW, SlvX, and SlvYZ) of the strain HSISS4. While PsnL was chosen for its broadest spectrum of prey species, PsnJ and PsnK were selected for their more restricted activity against salivarius group streptococci and *L. monocytogenes*, respectively (Fig. 5b, c). To generate the expanded cocktails, we transferred each of the three cognate fusions

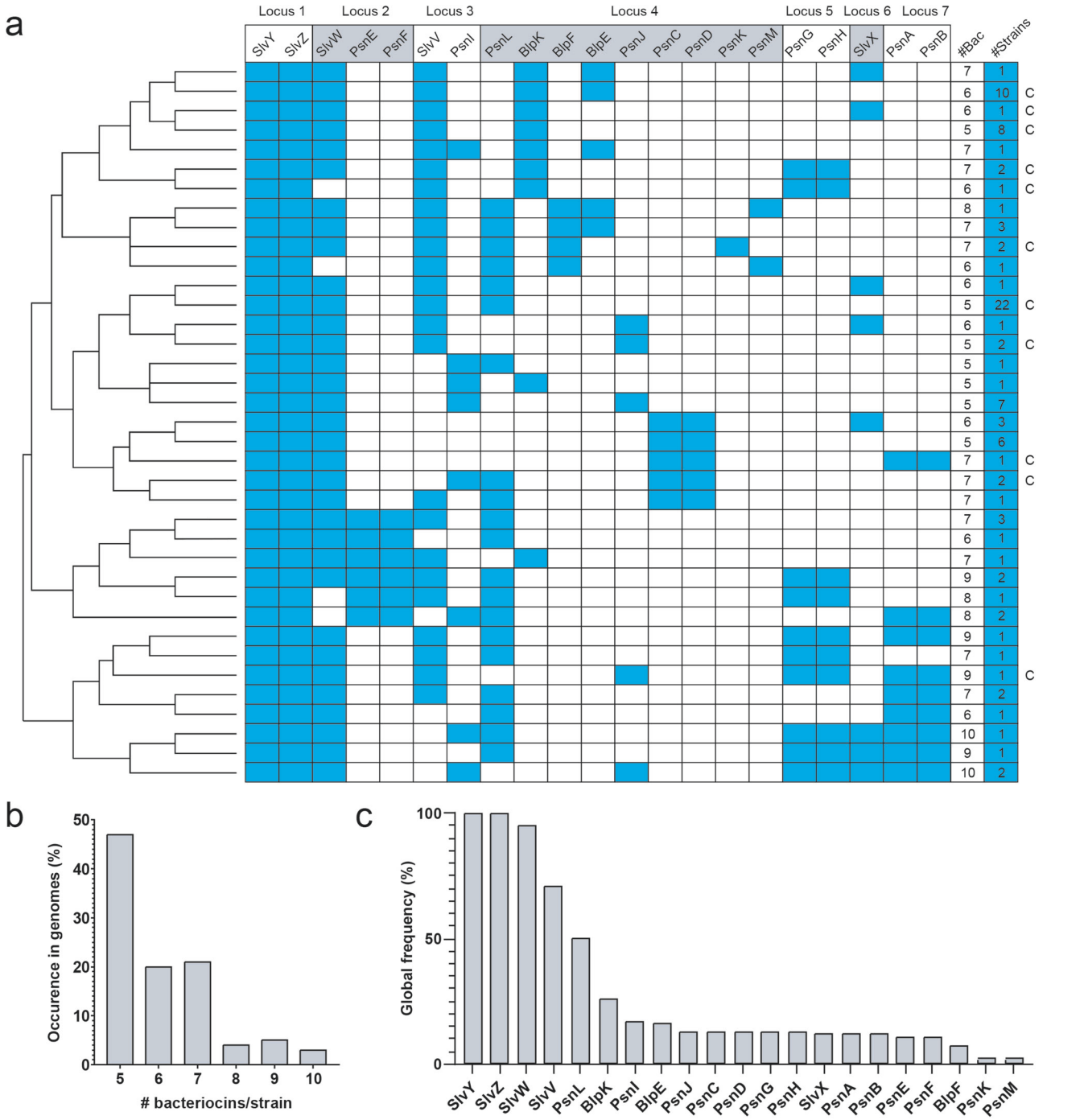

Fig. 3 | Composition of salivaricin cocktails identified in *S. salivarius*. **a** Schematic representation of the 37 salivaricin cocktails identified in seven loci from 100 *S. salivarius* genomes. A blue box indicates the presence of the bacteriocin. The hierarchical clustering in bacteriocin content of the different strains is shown on the left. The C letter on the right indicates that the cocktail was identified in at least one strain for which a closed genome is available. **b** Occurrence (%) of the number of salivaricins among *S. salivarius* strains. **c** Occurrence (%) of each individual salivaricin in *S. salivarius*.

described in the previous section in the wild-type HSISS4 background. The three resulting strains were tested for their antimicrobial activity against the same panel of 12 prey species by spot-on-lawn assays (Fig. 7, Supplementary Fig. 2b, and Supplementary Data 4 and 5). Although the inhibitory activity of the expanded cocktails remains unchanged against most species compared to the native cocktail produced by HSISS4 (Supplementary Fig. 2b), we observed that the addition of a single bacteriocin could increase its activity against specific species (Fig. 7). First, the addition of the narrow-spectrum bacteriocin PsnJ or PsnK significantly increased activity against *S. thermophilus* or *L. monocytogenes*, respectively (Fig. 7). Notably, the anti-listeria effect of PsnK is clearly synergistic with the native cocktail, while the impact of PsnJ seems to be more limited (Fig. 7). As PsnK is the only member of class IIa among all identified salivaricins, its specific mode of action could explain this synergy with at least one native bacteriocin, which have no or weak anti-listeria activity when produced alone (Fig. 5b). Second, the addition of the broad-spectrum bacteriocin PsnL increased the cocktail efficiency against the human pathogens *S. aureus* and *L. monocytogenes*, and the human opportunist *Staphylococcus epidermidis*. While a synergy is observed with the native cocktail against *S. epidermidis*, PsnL alone accounts for most of the activity against *S. aureus* and *L. monocytogenes* (Fig. 7b).

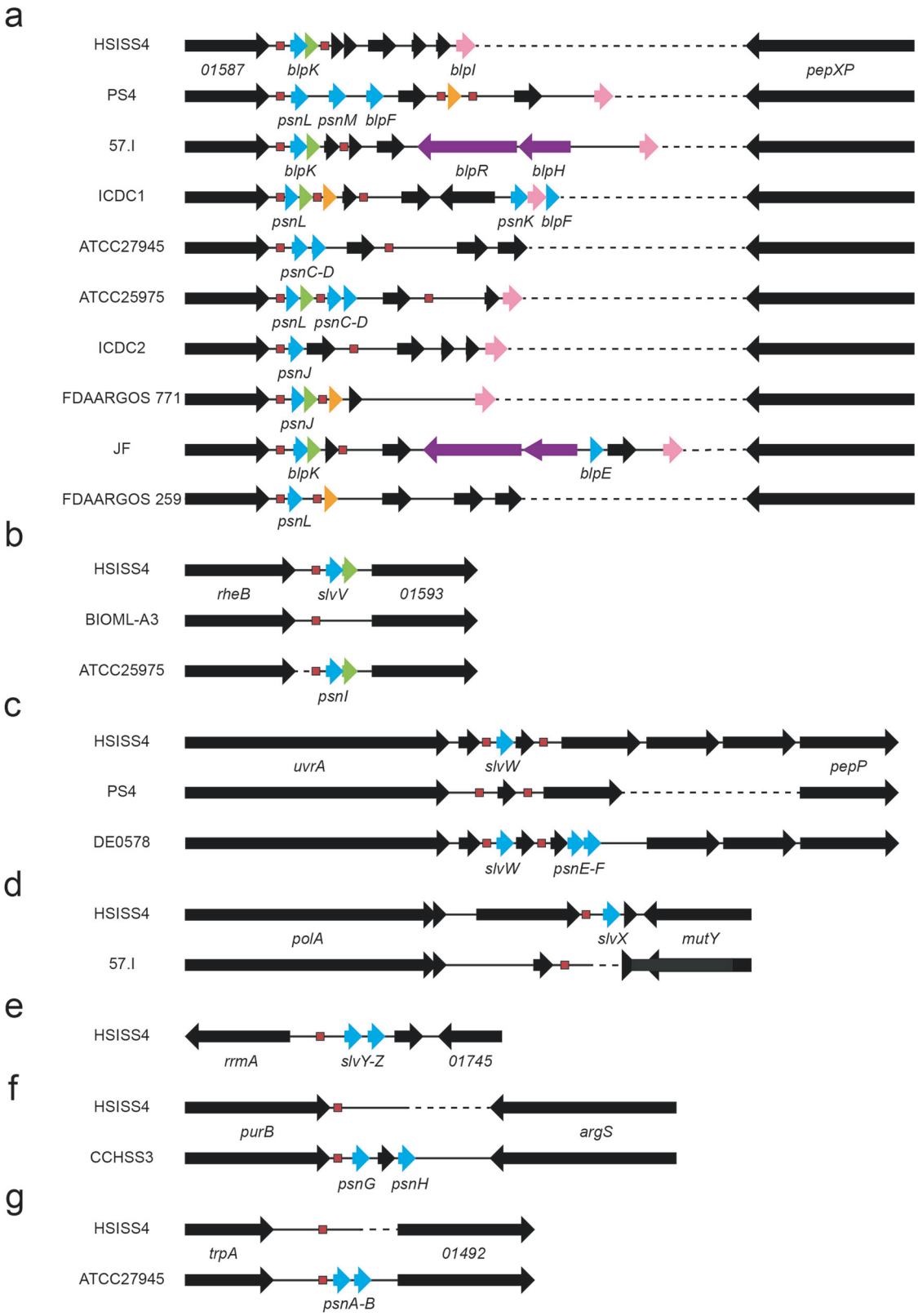

Fig. 4 | Comparison of the seven salivaricin loci in *S. salivarius*. Panels (**a**–**g**) display a representative variability in the seven different bacteriocin loci. A red box indicates the presence of a ComR-box; arrows represent genes encoding bacteriocins (blue), immunity peptides (green), bacteriocin leader sequences without the mature part (orange), fragments of the BlpRH system (purple), homologues of BlpI (pink), and other unrelated products (black). Dotted lines are spaces that were artificially added to align conserved genes in order to better visualize the genetic reorganization of each locus.

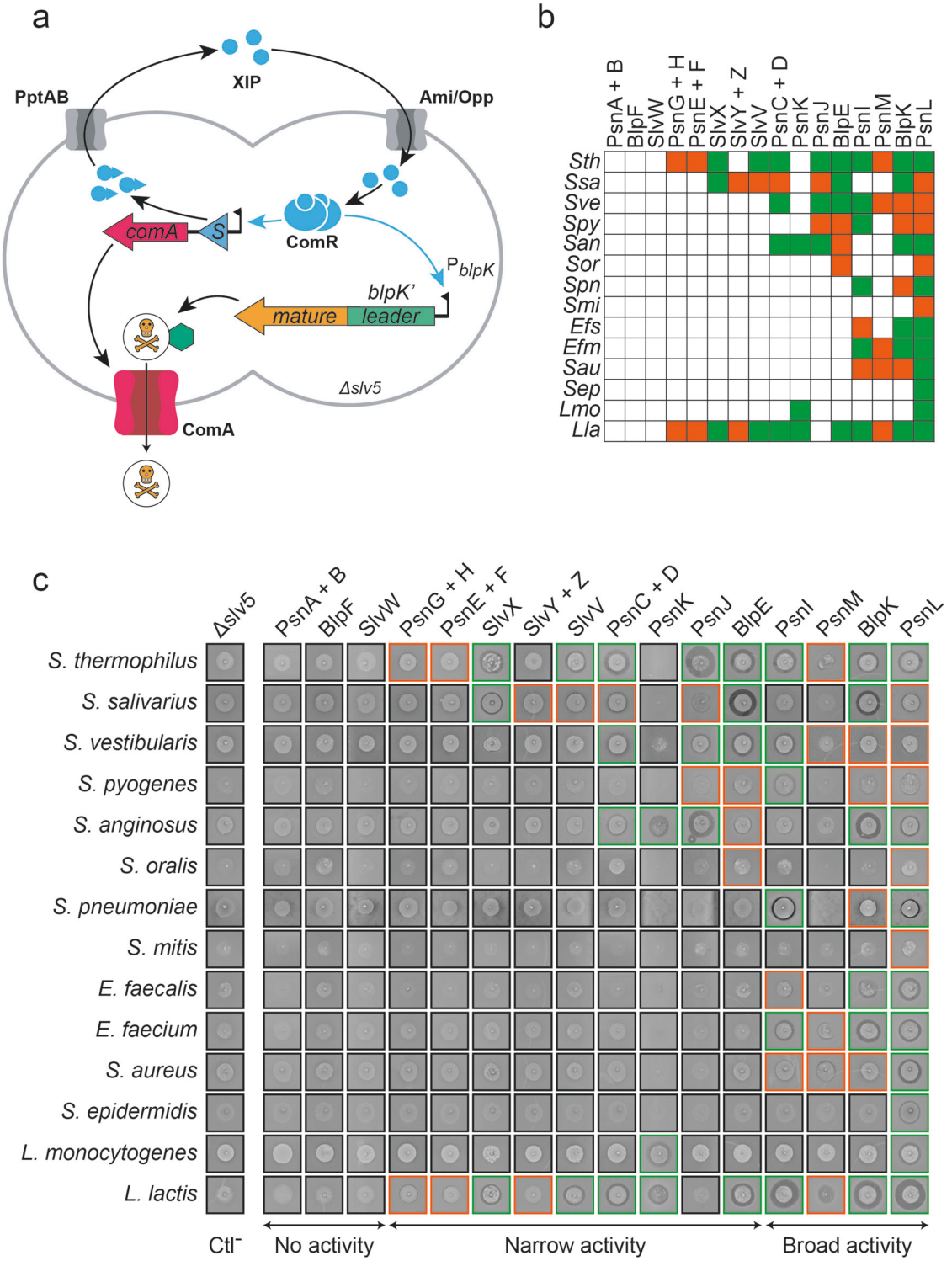

Fig. 5 | Activity assays performed with individual bacteriocins produced by *S. salivarius*. **a** Test strain (Δ*slv5*) for in vivo production of salvaricins. The gene fragment encoding the mature part of a salivaricin (orange arrow box) was fused to expression-secretion signals of *blpK* (P_{blpK}-*blpK*'; green rectangle). The addition of the XIP inducer (blue circles) activates ComR, which will in turn stimulate the production of ComS (positive feedback loop via PptAB-mediated ComS export and Opp/Ami-mediated XIP import) and ComA (bacteriocin exporter, red arrow box). Concomitantly, the ComR-XIP complex activates the production of the hybrid bacteriocin by binding to P_{blpK}. **b** Spectra of salivaricins against 12 Gram-positive bacteria belonging to the *Streptococcus*, *Lactococcus*, *Enterococcus*, *Staphylococcus*, and *Listeria* genera. The list of indicator strains is reported in Supplementary Table 8. In case of *S. pneumoniae* and *S. aureus*, strains PMEN31 and ATCC 6538 were used as indicator strains, respectively. Green, orange, and white boxes refer to high, weak, and absence of activity. **c** Spot-on-lawn assays with the 21 producer strains used to generate the table shown in panel (**b**). Single-peptide producers of class IIa bacteriocins were mixed. High (inhibition zone > 0.5 mm) and weak (inhibition zone ≤ 0.5 mm) activities are highlighted by green and orange lines, respectively. Producers are grouped into three categories: no activity, narrow spectrum, and large spectrum. Strain Δ*slv5* without bacteriocin activity is used as a negative control (Ctl^-). Each inhibition spot is representative of three biologically independent experiments performed in the same conditions (Supplementary Data 2).

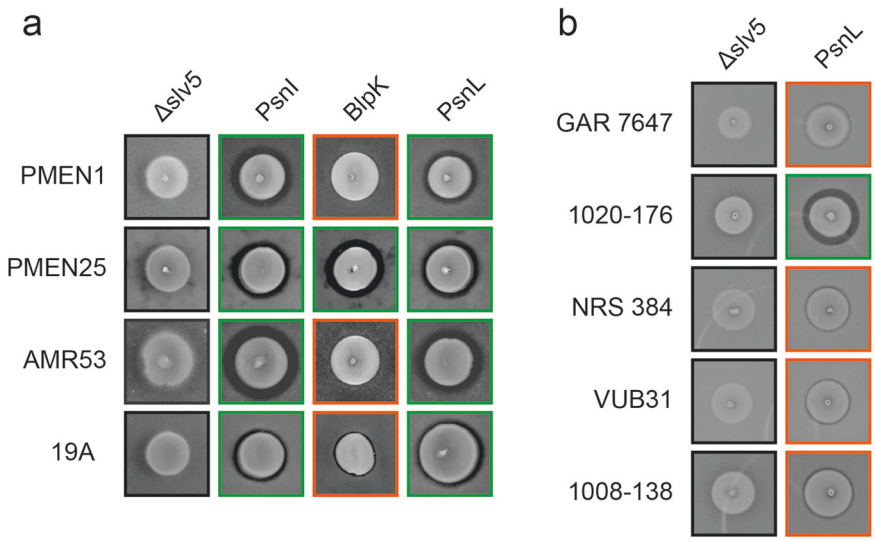

Fig. 6 | Activity assays performed against antibiotic-resistant strains of *S. aureus* and *S. pneumoniae*. Spot-on-lawn assays against antibiotic-resistant strains of *S. pneumoniae* (**a**) and *S. aureus* (**b**). The antibiotic resistance profile of strains is provided in Supplementary Table 8. High (inhibition zone > 0.5 mm) and weak (inhibition zone ≤ 0.5 mm) activities are surrounded by green and orange lines, respectively. Δ*slv5* mutant deprived of bacteriocin activity is used as a negative control. Each inhibition spot is representative of three biologically independent experiments performed in the same conditions (Supplementary Data 3).

Together, these results underline the importance of salivaricin combinations for enhanced efficiency and extended anti-bacterial spectra, as illustrated here for growth inhibition of clinically-relevant pathogens.

Discussion

In the past, novel bacteriocins were discovered by large-scale screening of producer candidates in combination with prey species of interest using the cell growth defect as a readout, followed by bacteriocin purification and identification of their sequences by mass spectrometry. This approach is not only time-consuming but also very restrictive for bacteriocins that are produced in laboratory conditions. This is particularly true for Gram-positive bacteria, and specifically streptococci, where the production of most bacteriocins is strictly induced by a range of transcriptional regulatory systems (RNPP cytoplasmic sensors and/or TCSs)[21,24,27]. Hence, the whole antimicrobial potential of a defined species remains mostly underestimated. To unveil this potential, various bioinformatics approaches have been undertaken, but none of them include the specific feature of genetic regulation as a key parameter to identify functional bacteriocin genes. The addition of this parameter to specific bacteriocin features was expected to strongly increase the power and sensitivity of the bioinformatic analysis (lower number of false negatives and positives). In this work, we fully validated this approach by unveiling the *S. salivarius* bacteriocinome under the control of the ComRS system. We also benchmarked our approach with classical bioinformatic tools used to identify bacteriocins in bacterial genomes, such as BAGEL4[11] and AntiSMASH[46]. For this comparison, we selected representative *S. salivarius* genomes in order to cover the seven identified loci at least once (Supplementary Table 6). Because BAGEL4 and AntiSMASH are searching for the synteny between bacteriocin genes and genes encoding post-translational modification enzymes, immunity factors, or export systems, they can only detect bacteriocins included in the main *blp* locus and globally failed to identify about 50% of the class II bacteriocin genes in *S. salivarius* genomes (Supplementary Table 6). Since the identification of bacteriocin candidates in *S. salivarius* was successful with our bioinformatic pipeline, we quickly evaluated if it could be applied to other streptococcal species whose bacteriocin genes could be controlled by the ComRS system (Supplementary Table 2). We selected one representative genome of three species (*i.e.*, *Streptococcus vestibularis*, *Streptococcus downei*, and *Streptococcus sobrinus*) and identified multiple bacteriocin candidates in each genome (Supplementary Table 7). While bacteriocin candidates of *S. vestibularis* were all previously identified in *S. salivarius*, 12 uncharacterized bacteriocins were discovered in the genomes of the two other species (Supplementary Table 7). Additionally, we also implement our bioinformatic approach to identify streptococcal bacteriocins that could be regulated by the alternative signaling system BlpRH (Supplementary Table 2). As reported above, we used one representative genome of three species (*i.e.*, *Streptococcus pneumoniae*, *Streptococcus oralis*, and *S. thermophilus*) whose bacteriocin genes are predicted to be controlled by the BlpRH system. From the 19 bacteriocin candidates identified in the three genomes, three bacteriocins from *S. oralis* were never reported before (Supplementary Table 7). These two additional in silico analyses show that our multiparametric bioinformatic tool could be broadly applied for the discovery of hidden class II bacteriocins in a range of streptococcal species. Finally, we reckon that this regulation-based screening could be adapted to other quorum sensing systems (or sensory modules) in many bacteria to drastically increase the identified reservoir of bacteriocins and associated functions (*e.g.*, maturation enzymes, immunity proteins, signaling peptides).

The in silico analysis of the *S. salivarius* bacteriocinome revealed interesting features in terms of inter- and intra-species distributions. At the inter-species level, the class II salivaricins are partially shared between the three streptococcal species of the salivarius group. For instance, approximately 40% of salvaricins are common to *S. salivarius* and *S. thermophilus* with a high level of protein identity (Fig. 2a and Supplementary Table 3). Since these two species cross-talk by using very similar XIP pheromones and display a high nucleotide identity at the genome level (90–95%)[16,28], the bi-directional spreading of bacteriocin genes could be strongly favored by natural transformation. By contrast, their low distribution in other streptococcal groups (Fig. 2a) is certainly constrained by the absence of cross-talk and lower genomic sequence conservation[47]. At the intra-species level, we observed a high variability of the salivaricin content (Fig. 3). While class II bacteriocins were reported to be mainly grouped in a single cluster (*blp* locus) in *S. pneumoniae* and *S. thermophilus*[23,37], they are scattered throughout seven unlinked loci in *S. salivarius* (Fig. 4). This large spread at the genome level favors a high variability of bacteriocin cocktails (~ 40 over 100 strains). However, the *blp*-like locus (locus 4), which contains pseudogenes of the BlpRH system[22], remains the more dynamic locus for bacteriocin gene insertion, deletion, or swapping. Indeed, it could be considered as a plasticity platform for bacteriocin gene exchanges as it encodes nine different bacteriocin peptides and up to four in the same strain (Figs. 3a and 4). Since bacteriocin loci could exist as remnants, one question that will need further investigation is the functionality of the different loci in each analyzed strain. The use of ComR boxes as a hint for their discovery is a strong indication for their inducibility by the ComRS system. We also verified the presence of intact genes encoding ComR as well as ComA, the class II bacteriocin transporter, in the 100 selected genomes of *S. salivarius* (Supplementary Table 4). Their presence in all genomes is in favor of bacteriocin functionality since these two proteins are absolutely required for regulation and secretion of class II salivaricins,

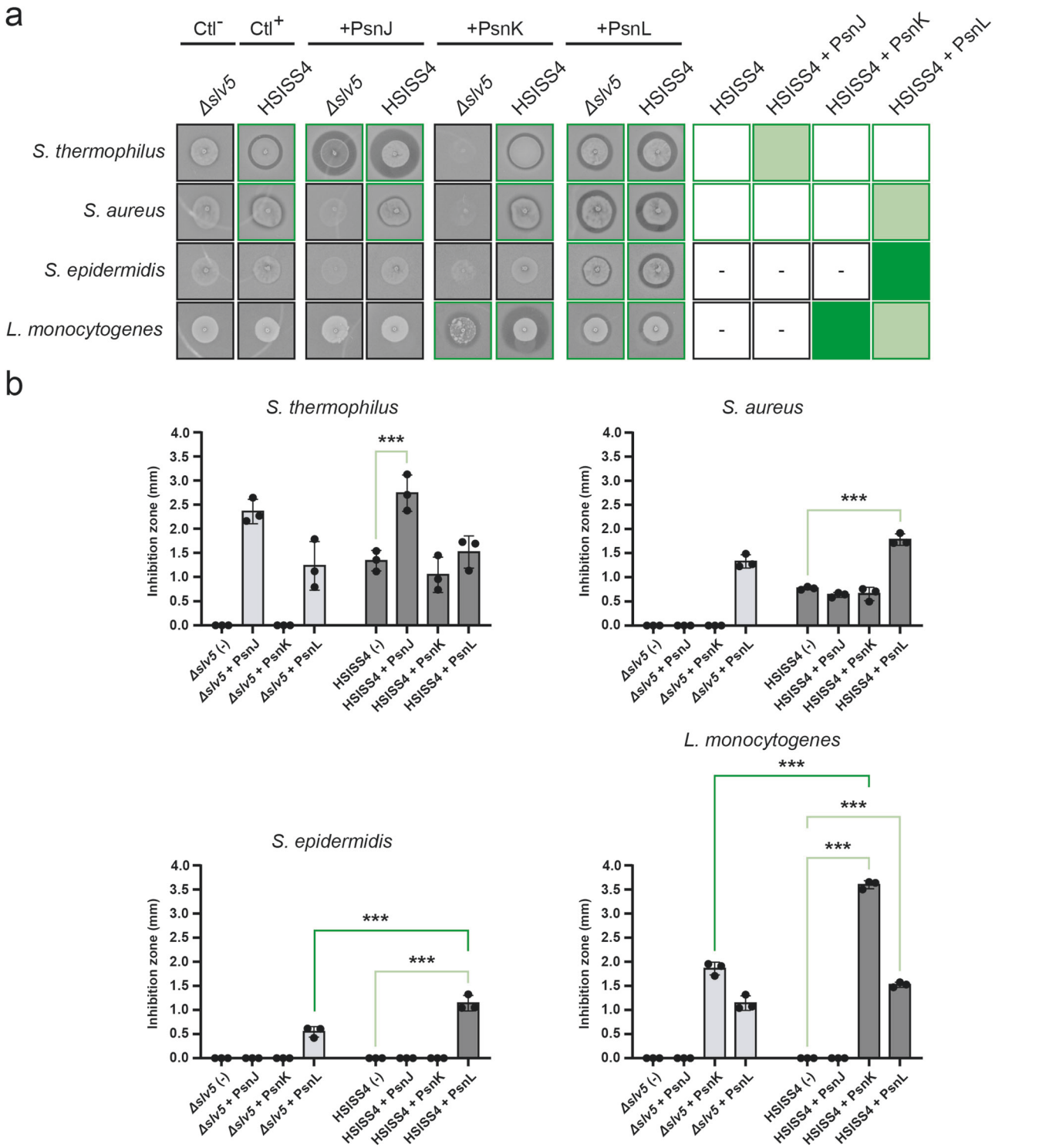

Fig. 7 | Activity assays of modified salivaricin cocktails produced by *S. salivarius* HSISS4. a Spot-on-lawn assays (left panels) of HSISS4 derivatives (wild-type and Δ*slv5*) producing PsnJ, PsnK, or PsnL against four Gram-positive bacteria. The list of indicator strains is reported in Supplementary Table 8. In case of *S. aureus*, strain ATCC 6538 was used as indicator strain. Strains Δ*slv5* and HSISS4 (wild-type) are used as negative (Ctl^-) and positive (Ctl^+) controls, respectively. Bacteriocin producers are surrounded by green lines. A schematic representation of inhibitory spectra from HSISS4 cocktails is shown in the right panel. Boxes with a minus sign indicate that no activity was detected. Boxes surrounded in green indicate high activity (inhibition zone > 0.5 mm). Boxes filled in green indicate a modified activity due to the presence of the additional salivaricin in the cocktail of wild-type HSISS4; light and dark green colors refer to additive and synergic effects, respectively. Each inhibition spot is representative of three independent experiments performed in the same conditions (Supplementary Data 4). **b** Inhibition zone (mm) of HSISS4 wild-type (dark gray bars) or Δ*slv5* (light gray bars) derivatives producing PsnJ, PsnK, or PsnL against four Gram-positive bacteria (list of strains in Supplementary Table 8). Inhibition zones were measured with the ImageJ software. Dots show the values of biologically independent experiments ($n = 3$); mean values ± standard deviations. Statistical analysis was performed using One-Way ANOVA with Tukey's correction. ***, *P*-Value < 0.001 (Supplementary Data 5).

respectively[22]. In addition, the high conservation of the 21 salivaricin genes and the activity of 13 out of 16 bacteriocin modules (one- or two-peptide(s)) argues against a massive decay resulting in relic bacteriocin genes, even if we cannot rule out the presence of some inactive mutated variants. Activation by the XIP peptide combined with transcriptomic studies and systematic inactivation will be required to definitively establish their functionality in a defined strain, as previously performed for the strain HSISS4[22]. In this strain, that contains five out of the seven loci identified in this work, all bacteriocin genes were shown to be expressed (XIP-induced) and required for the global antimicrobial activity of the strain[22].

Our systematic work of testing the activity of each individual bacteriocin peptide sheds light on its respective contribution to bacteriocin cocktails in different strains. In *S. salivarius*, competence and predation are intimately interlinked since both processes are directly controlled by the ComRS system[22]. Competence-induced bacteriocins that kill bacteria of the same species or closely related species could be considered as fratricins or sobrinicins, respectively[33,34,48,49]. Both types of killing peptides (in association or not with bacteriolysins) are able to release exogenous DNA by lysis from neighboring bacteria. Next, free DNA could be captured by the transformation machinery with an opportunity of genome integration, depending on the DNA homology of the prey genetic material. The repertoire of class II salivaricins active against streptococci includes both fratricins and sobrinicins (Fig. 8a).

On one hand, the two-peptide SlvYZ exclusively inhibits *S. salivarius* among streptococci, suggesting that it might be considered as a strict fratricin. Even if its activity is weak, these two peptides are present in all *S. salivarius* genomes analyzed during this study, indicating that they play a key role for sibling competition in *S. salivarius* (Figs. 3a and 5b). On the other hand, SlvX, SlvV, PsnCD, and PsnEF can be seen as narrow-range sobrinicins targeting species of the salivarius group, while PsnJ and BlpE are large-range sobrinicins with an extended repertoire against other streptococcal groups (Figs. 5b and 8a). Interestingly, *S. salivarius* is also equipped with four broad-spectrum bacteriocins of gradual increased activity (i.e., PsnI < PsnM < BlpK < PsnL) against many genera among *Firmicutes* (low-GC Gram-positive bacteria) (Figs. 5b and 8a). Such a large spectrum of prey species may suggest that their roles overcome the competence-predation coupling for genome plasticity. The high occurrence (around 85%) of at least one of these large-spectrum bacteriocins in *S. salivarius* strains suggests an additional role in general predation for niche occupancy, which is in correlation with the high incidence of this species as a commensal in many subniches of the human body (*e.g.*, oral cavity, upper airways, and upper part of the small intestine)[50–52]. As *S. salivarius* is present as a consortium of multiple strains[52], its shared class II bacteriocinome revealed a combinatory strategy of overlapping antimicrobial activities from a (very) narrow to large spectrum against *Firmicutes* (Fig. 8a). Besides a wider inhibition of competitors, the selective pressure that has shaped overlapping activities against a defined species is most probably dictated by an increased inhibitory efficacy (i.e., additive and/or synergistic effects), as shown here by the mimetic transfer of additional bacteriocins in a wild-type background (Fig. 7). Such a dual narrow-&-large spectrum antimicrobial strategy might be the optimal way to maintain an ecological dominance over the microbiota community considering it does not require to assess the status of complex mixed populations (abundance and diversity). Indeed, *S. salivarius* coordinates bacteriocin production through ComRS control[22]. Therefore, it is unlikely that *S. salivarius* alternates production of narrow and broad spectrum bacteriocins in response to particular microbiota cues. Conversely, the advantage of simultaneous production will be two-fold. First, broad-range bacteriocins will impose a burden on the global community growth/survival. This could result in a total dominance in extreme cases when the producer cells initially outnumber the preys[53]. Second, narrow-range bacteriocins will eliminate, in a targeted manner, direct competitors that grapple for equivalent ressources (nutrient or space) or synthesize inhibitory molecules. This concomitant salvo is supposed to minimize disruption of the overall community and save energy consumption from the producer cells (lower amount of competitor cells to kill)[43,53].

To better understand their evolution and mode of action, we also predicted the 3D structure of salivaricins using AlphaFold 2[54–56] (Fig. 8b). Among class IId bacteriocins (single peptide), which include the largest set of active salivaricins, most are predicted to be α-hairpin peptides formed of two anti-parallel α-helices stabilized by a disulfide bond[57]. Although initially considered class IId bacteriocins in our study, BlpF and SlvW are not or weakly structured in correlation with an absence of detected antimicrobial activity (Figs. 5 and 8b). We hypothesized that they could be either helper antimicrobial peptides or uncharacterized signaling peptides. The predicted structures of most class IIb heterodimeric bacteriocins mimics α-hairpin peptides (i.e., PsnAB, PsnEF, and PsnGH) (Fig. 8b) or are the association of two α-hairpin peptides (i.e., PsnCD) as recently reported for the atypical class IIb Gallocin A[58]. Finally, the predicted structure of PsnK is in full accordance with the well-established structures of pediocin-like bacteriocins of class IIa, shown to use the Man-PTS system as a receptor[59,60]. With some exceptions (i.e., BlpE, SlvYZ, and PsnK), the structures of most class II salivaricins are variations or mimetics of α−hairpins, highlighting a possible convergent evolution for their activity. Another feature shared by the majority of salivaricins is the presence of positively-charged residues (Table 1, see global charge and % Lys-Arg), which could be clustered in the α-hairpin linker region (i.e., BlpK) or at C-terminal extremities (i.e., PsnAB, PsnEF, and PsnGH) (Fig. 8b). As previously reported[61], we hypothesized that those residues are key to initiate interactions with negatively charged lipids in the plasma membrane. Except PsnK that is presumed to interact with the Man-PTS, the mode of action of class II salivaricins such as the requirement of a specific receptor or the ability to oligomerize for pore formation will require further investigations.

To conclude, we set up an original combined approach for bacteriocin discovery that integrates their mode of regulation, exploited here for both in silico analysis and in vivo validation of their activities. This approach allowed us to extensively reveal the class II bacteriocinome of *S. salivarius* and offers a high potential to unveil the comprehensive bacteriocinomes of many Gram-positive species that are so far extremely underestimated. This generic approach is key to enrich bacteriocin databases and physical collections of synthetic genes such as PARAGEN[44] and maximally exploit the whole potential of these antimicrobial peptides for many future biocontrol applications. The characterization of the *S. salivarius* bacteriocinome as such also offers many opportunities of development in the probiotics field, such as the prediction of the antimicrobial potential of existing bacteriocin cocktails in various strains or the design of second-generation probiotics with adapted cocktails targeting specific pathogens. Class II bacteriocins as unmodified peptides are easy to synthesize chemically and to engineer by genetic modification for optimizing their efficiency and/or activity spectrum[21]. Hence, the antimicrobial spectra of BlpK, PsnL, PsnI, PsnM, and PsnK against clinically-relevant pathogens and the identification of their numerous natural variants should have a high potential for future exploitation in the food industry and biomedicine.

Methods

Bacterial strains and DNA material

Bacterial strains, plasmids, oligonucleotides, and PCR fragments used in this study are listed in Supplementary Tables 8, 9, 10, and 11, respectively.

Growth conditions

Streptococcus salivarius HSISS4 and derivates were grown without shaking in M17G (glucose 1% [w/v]) medium (Oxoid) or CDMG (chemically-defined medium with glucose 1% [w/v])[51] at 37 °C. All indicator strains were grown without shaking in M17G at 37 °C, except *L. lactis* IL1403 that was grown at 30 °C. Solid plates inoculated with *S. salivarius* were incubated anaerobically (AnaeroGen 2.5 L, Thermo Scientific) at 37 °C. Spectinomycin was added when needed at 200 μg ml^{-1}. Synthetic peptide sXIP (LPYFAGCL; purity of 95%) was supplied by Peptide2.0 Inc. (Chantilly, VA, USA), resuspended in water, and used at a concentration of 250 nM, except if otherwise stated.

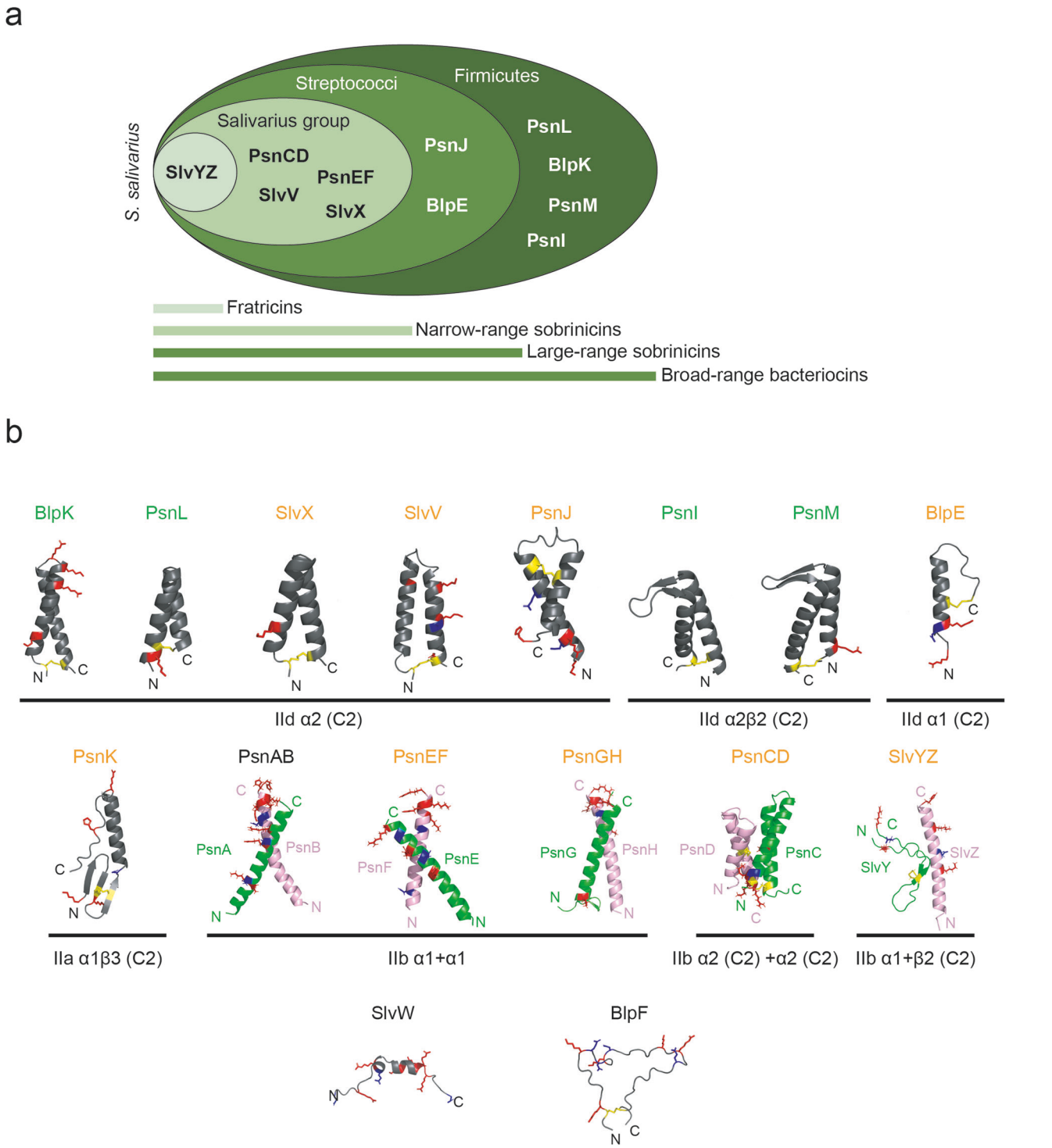

Fig. 8 | Global prey spectra and structure prediction of class II salivaricins. **a** Global overview of the overlapping activity spectra of active salivaricins against streptococci and *Firmicutes*. The activity against *L. lactis* (dairy strain) is not included here since considered as not relevant for the ecology of the digestive tract. **b** Structure prediction by AlphaFold 2 of monopeptides and two-peptide bacteriocins. Predicted 3D structures are organized according to their subclasses (IId, IIb, and IIa) in class II bacteriocins and shared structural elements (α helices and β strands). The presence of a predicted disulfide bond (C2) is also indicated. The color of the bacteriocin name in green, orange, and black refers to a large spectrum of prey species, a narrow spectrum of prey species, and the absence of identified activity, respectively. Monopeptides are in dark gray, while dipeptides are in green and pink. Cysteine, positively charged, and negatively charged residues are colored in yellow, red, and blue, respectively.

Strain construction by natural transformation

S. salivarius HSISS4 was transformed with Gibson assembly products composed of three PCR fragments: (i) the upstream region of $tRNA^{Ser}$ associated with a *blpK* gene fragment containing expression signals and encoding the leader sequence of BlpK; (ii) the downstream region of $tRNA^{Ser}$ associated with the *spc* cassette (spectinomycin resistance); and (iii) the mature sequence of a bacteriocin. PCR fragments were amplified with the Q5 polymerase (NEB, Ipswich, MA, USA), following the manufacturer's recommended protocol. To induce natural transformation, an overnight CDMG preculture was diluted in 500 µl of fresh CDMG at a final OD_{600} of

0.05 and incubated for 135 min at 37 °C. Then, 1 µM of sXIP and linear DNA (Gibson assembly product) were added to the culture, and cells were further incubated for 3 h at 37 °C before plating on M17G agar supplemented with spectinomycin. After transformation, all constructions were verified by DNA sequencing.

Spot-on-lawn assays for bacteriocin activity

To detect bacteriocin activity, 50 µl of overnight cultures of producer strains were diluted in 1 ml of M17G and grown for 4 h at 37 °C. In parallel, a first feeding layer (30 ml) of M17G 1.5% agar supplemented with XIP (250 nM) was cast on a plate. Then, a second layer (15 ml) of M17G 0.3% agar containing the indicator strain was poured on top of it. This layer contains 800 µl of an overnight culture of the indicator strain. On the top of this last layer, 3 µl of bacteriocin producers were spotted. Finally, the plate was incubated overnight in anaerobic conditions at 37 °C.

In vitro production of bacteriocins

Synthetic plasmids (pUC57 derivative) expressing the mature parts of bacteriocins under the control of the T7 promoter were provided by GenScript (Piscataway, NJ). These plasmids (Supplementary Table 9) in combination with the cell-free system PURExpress (NEB) were used to produce bacteriocins in vitro. To synthesize bacteriocins, the reaction mixture was assembled on ice in the following order: 4 µl of solution A, 3 µl of solution B, 2 µl of nuclease-free water, and 1 µl of template DNA (100 ng μl^{-1}). Then, the reaction mixture was incubated for 3 h at 37 °C, and the activity was measured by a spot-on-lawn assay.

Multiparametric bioinformatics script

The script used to perform the bioinformatic analysis was written in R. The code of this script and the data set obtained from the screening of 100 genomes of *S. salivarius* are available in Supplementary Software 1 and Supplementary Data 1, respectively. For *S. salivarius*, a detection limit of ComR-boxes in the upper 75% of identity to the consensus of the identified ComR-box (Supplementary Table 2) was set to avoid identifying pseudo-ComR-boxes. The function predORF available on R was used to predict ORFs and coding sequences (CDSs) in the provided DNA sequences. A maximal distance of 4 kb was set between the ComR-box and the start codon to cover potential operons. For each selected CDS, the script automatically provides DNA and amino acid sequences, the peptide length, sequence and length of the leader and mature parts, the identified regulatory box, the distance between the regulatory box and the start codon. For the predicted mature peptides, the script automatically provides the GRAVY (GRand AVerage of hydropathY) index, the isoelectric point, the charge, the percentage of lysine and arginine, and the molecular weight. We manually searched for the presence of a Shine-Dalgarno sequence (i.e, AAAGGAGGT) in the upstream region of each CDS. Candidate ribosome binding sites were retained if the Shine-Dalgarno sequence contains four consecutive conserved positions with at least three G surrounding the central A nucleotide (hereabove underlined), and a downstream start codon is distant from 5 to 12 nucleotides[62]. The script was calibrated using the genome of *S. salivarius* HSISS4 in which class II salivaricins encoding loci were previously investigated[22]. The script was configured to identify the six known bacteriocins of HSISS4 without detecting any false positive.

In addition to the script, tBLASTn was used on each strain of *S. salivarius* to confirm bacteriocin presence and identify variants. The tBLASTn searches were performed with the default setting, except for the expect threshold parameter, which was increased to 5 to avoid false negative hits.

Bioinformatic tools

Structure predictions of bacteriocins with AlphaFold 2 were obtained from the AlphaFold CoLab notebook[55,56]. Alignment of BlpK sequences were generated using MUSCLE[63] or Clustal OMEGA[64]. Phylogenetic trees were generated with MEGA11[65].

Statistics and reproducibility

Three biologically independent experiments were performed in the same conditions for all bacteriocin assays. Numerical data are the mean values of biological triplicates ± standard deviations. For multiple comparisons, *P* values were calculated using One-Way ANOVA with Tukey's correction.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data generated or analyzed during this study are included in this published article and supplementary files. The data set generated from the analysis of 100 genomes of *S. salivarius* is available in Supplementary Data 1. The source data underlying the figures and Supplementary Figs. can be found in Supplementary Data 2-5.

Code availability

The code of the script used in this work is available as Supplementary Software 1.

Received: 29 May 2024; Accepted: 5 November 2024;
Published online: 14 November 2024

References

1. Murray, C. J. et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).
2. O'Neill, J. Tackling drug-resistant infections globally: final report and recommendations. The government of the United Kingdom. The review on antimicrobial resistance. (2016).
3. Chan, B. K., Abedon, S. T. & Loc-Carrillo, C. Phage cocktails and the future of phage therapy. *Future Microbiol.* **8**, 769–783 (2013).
4. Czaplewski, L. et al. Alternatives to antibiotics-a pipeline portfolio review. *Lancet Infect. Dis.* **16**, 239–251 (2016).
5. Kole, R., Krainer, A. R. & Altman, S. RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.* **11**, 125–140 (2012).
6. Cotter, P. D., Ross, R. P. & Hill, C. Bacteriocins - a viable alternative to antibiotics? *Nat. Rev. Microbiol.* **11**, 95–105 (2013).
7. Joerger, R. D. Alternatives to antibiotics: bacteriocins, antimicrobial peptides and bacteriophages. *Poult. Sci.* **82**, 640–647 (2003).
8. Sang, Y. & Blecha, F. Antimicrobial peptides and bacteriocins: alternatives to traditional antibiotics. *Anim. Health Res. Rev.* **9**, 227–235 (2008).
9. Sugrue, I., Ross, R. P. & Hill, C. Bacteriocin diversity, function, discovery and application as antimicrobials. *Nat. Rev. Microbiol.* **22**, 556–571 (2024).
10. Diep, D. B., Skaugen, M., Salehian, Z., Holo, H. & Nes, I. F. Common mechanisms of target cell recognition and immunity for class II bacteriocins. *Proc. Natl Acad. Sci. USA* **104**, 2384–2389 (2007).
11. van Heel, A. J. et al. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res* **46**, W278–W281 (2018).
12. Klaenhammer, T. R. Genetics of bacteriocins produced by lactic acid bacteria. *FEMS Microbiol. Rev.* **12**, 39–85 (1993).
13. Alvarez-Sieiro, P., Montalban-Lopez, M., Mu, D. & Kuipers, O. P. Bacteriocins of lactic acid bacteria: extending the family. *Appl. Microbiol. Biotechnol.* **100**, 2939–2951 (2016).
14. Yount, N. Y. et al. Discovery of Novel Type II Bacteriocins Using a New High-Dimensional Bioinformatic Algorithm. *Front Immunol.* **11**, 1873 (2020).
15. Patton, G. C., Paul, M., Cooper, L. E., Chatterjee, C. & van der Donk, W. A. The importance of the leader sequence for directing lanthionine formation in lacticin 481. *Biochemistry* **47**, 7342–7351 (2008).

16. Delorme, C., Abraham, A. L., Renault, P. & Guedon, E. Genomics of *Streptococcus salivarius*, a major human commensal. *Infect. Genet. Evol.* **33**, 381–392 (2015).
17. Burton, J. P. et al. Influence of the probiotic *Streptococcus salivarius* strain M18 on indices of dental health in children: a randomized double-blind, placebo-controlled trial. *J. Med. Microbiol* **62**, 875–884 (2013).
18. Di Pierro, F., Adami, T., Rapacioli, G., Giardini, N. & Streitberger, C. Clinical evaluation of the oral probiotic *Streptococcus salivarius* K12 in the prevention of recurrent pharyngitis and/or tonsillitis caused by Streptococcus pyogenes in adults. *Expert. Opin. Biol. Ther.* **13**, 339–343 (2013).
19. Bidossi, A. et al. Probiotics *Streptococcus salivarius* 24SMB and *Streptococcus oralis* 89a interfere with biofilm formation of pathogens of the upper respiratory tract. *Bmc. Infect. Dis.* **18**, 653 (2018).
20. Barbour, A., Wescombe, P. & Smith, L. Evolution of Lantibiotic Salivaricins: New Weapons to Fight Infectious Diseases. *Trends Microbiol* **28**, 578–593 (2020).
21. Hols, P., Ledesma-Garcia, L., Gabant, P. & Mignolet, J. Mobilization of Microbiota Commensals and Their Bacteriocins for Therapeutics. *Trends Microbiol* **27**, 690–702 (2019).
22. Mignolet, J. et al. Circuitry Rewiring Directly Couples Competence to Predation in the Gut Dweller *Streptococcus salivarius*. *Cell Rep.* **22**, 1627–1638 (2018).
23. Fontaine, L. et al. Quorum-sensing regulation of the production of Blp bacteriocins in *Streptococcus thermophilus*. *J. Bacteriol.* **189**, 7195–7205 (2007).
24. Shanker, E. & Federle, M. J. Quorum Sensing Regulation of Competence and Bacteriocins in *Streptococcus pneumoniae* and *mutans*. *Genes (Basel)* **8**, 15 (2017).
25. Jimenez, J. C. & Federle, M. J. Quorum sensing in group A *Streptococcus*. *Front Cell Infect. Microbiol* **4**, 127 (2014).
26. Kuipers, O. P., Beerthuyzen, M. M., de Ruyter, P. G., Luesink, E. J. & De Vos, W. M. Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal transduction. *J. Biol. Chem.* **270**, 27299–27304 (1995).
27. Garcia-Curiel, L., Del, R. L.-C., Rodriguez-Hernandez, A. I. & Chavarria-Hernandez, N. Toward understanding the signals of bacteriocin production by *Streptococcus* spp. and their importance in current applications. *World J. Microbiol. Biotechnol.* **37**, 15 (2021).
28. Fontaine, L. et al. A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* **192**, 1444–1454 (2010).
29. Fontaine, L. et al. Mechanism of competence activation by the ComRS signalling system in streptococci. *Mol. Microbiol.* **87**, 1113–1132 (2013).
30. Fontaine, L., Wahl, A., Flechard, M., Mignolet, J. & Hols, P. Regulation of competence for natural transformation in streptococci. *Infect. Genet. Evol.* **33**, 343–360 (2015).
31. Berg, K. H., Biornstad, T. J., Johnsborg, O. & Havarstein, L. S. Properties and biological role of streptococcal fratricins. *Appl. Environ. Microbiol.* **78**, 3515–3522 (2012).
32. Biornstad, T. J., Ohnstad, H. S. & Havarstein, L. S. Deletion of the murein hydrolase CbpD reduces transformation efficiency in *Streptococcus thermophilus*. *Microbiol. (Read.)* **158**, 877–885 (2012).
33. Claverys, J. P. & Havarstein, L. S. Cannibalism and fratricide: mechanisms and raisons d'etre. *Nat. Rev. Microbiol.* **5**, 219–229 (2007).
34. Veening, J. W. & Blokesch, M. Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nat. Rev. Microbiol.* **15**, 621–629 (2017).
35. de Saizieu, A. et al. Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. *J. Bacteriol.* **182**, 4696–4703 (2000).
36. Kjos, M. et al. Expression of *Streptococcus pneumoniae* Bacteriocins Is Induced by Antibiotics via Regulatory Interplay with the Competence System. *PLoS. Pathog.* **12**, e1005422 (2016).
37. Lux, T., Nuhn, M., Hakenbeck, R. & Reichmann, P. Diversity of bacteriocins and activity spectrum in *Streptococcus pneumoniae*. *J. Bacteriol.* **189**, 7741–7751 (2007).
38. Pinchas, M. D., LaCross, N. C. & Dawid, S. An electrostatic interaction between BlpC and BlpH dictates pheromone specificity in the control of bacteriocin production and immunity in *Streptococcus pneumoniae*. *J. Bacteriol.* **197**, 1236–1248 (2015).
39. Bushin, L. B., Clark, K. A., Pelczer, I. & Seyedsayamdost, M. R. Charting an Unexplored Streptococcal Biosynthetic Landscape Reveals a Unique Peptide Cyclization Motif. *J. Am. Chem. Soc.* **140**, 17674–17684 (2018).
40. Montalban-Lopez, M. et al. New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* **38**, 130–239 (2021).
41. Fontaine, L. & Hols, P. The inhibitory spectrum of thermophilin 9 from *Streptococcus thermophilus* LMD-9 depends on the production of multiple peptides and the activity of BlpG(St), a thiol-disulfide oxidase. *Appl. Environ. Microbiol.* **74**, 1102–1110 (2008).
42. Nissen-Meyer, J., Oppegard, C., Rogne, P., Haugen, H. S. & Kristiansen, P. E. Structure and Mode-of-Action of the Two-Peptide (Class-IIb) Bacteriocins. *Probiotics. Antimicrob. Proteins* **2**, 52–60 (2010).
43. Arbulu, S. & Kjos, M. Revisiting the Multifaceted Roles of Bacteriocins. *Microb. Ecol.* **87**, 41 (2024).
44. Gabant, P. & Borrero, J. PARAGEN 1.0: A Standardized Synthetic Gene Library for Fast Cell-Free Bacteriocin Synthesis. *Front Bioeng. Biotechnol.* **7**, 213 (2019).
45. Eijsink, V. G., Skeie, M., Middelhoven, P. H., Brurberg, M. B. & Nes, I. F. Comparative studies of class IIa bacteriocins of lactic acid bacteria. *Appl. Environ. Microbiol.* **64**, 3275–3281 (1998).
46. Blin, K. et al. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* **51**, W46–W50 (2023).
47. Shanker, E. et al. Pheromone Recognition and Selectivity by ComR Proteins among Streptococcus Species. *PLoS. Pathog.* **12**, e1005979 (2016).
48. Ikryannikova, L. N., Kurbatov, L. K., Soond, S. M. & Zamyatnin, A. A., Jr Harnessing the Potential of Killers and Altruists within the Microbial Community: A Possible Alternative to Antibiotic Therapy? *Antibiotics. (Basel)* **8**, (2019).
49. Kreth, J., Merritt, J., Shi, W. & Qi, F. Co-ordinated bacteriocin production and competence development: a possible mechanism for taking up DNA from neighbouring species. *Mol. Microbiol.* **57**, 392–404 (2005).
50. Jorissen, J. et al. Case-Control Microbiome Study of Chronic Otitis Media with Effusion in Children Points at Streptococcus salivarius as a Pathobiont-Inhibiting Species. *mSystems*. **6**, (2021).
51. Letort, C. & Juillard, V. Development of a minimal chemically-defined medium for the exponential growth of *Streptococcus thermophilus*. *J. Appl. Microbiol.* **91**, 1023–1029 (2001).
52. van den Bogert, B. et al. Comparative genomics analysis of *Streptococcus* isolates from the human small intestine reveals their adaptation to a highly dynamic ecosystem. *PLoS. One* **8**, e83418 (2013).
53. Palmer, J. D. & Foster, K. R. The evolution of spectrum in antibiotics and bacteriocins. *Proc. Natl Acad. Sci. USA* **119**, e2205407119 (2022).
54. McDonald, E. F., Jones, T., Plate, L., Meiler, J. & Gulsevin, A. Benchmarking AlphaFold2 on peptide structure prediction. *Structure* **31**, 111–119 (2023).
55. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

56. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439–D444 (2022).
57. Barthe, P., Rochette, S., Vita, C. & Roumestand, C. Synthesis and NMR solution structure of an alpha-helical hairpin stapled with two disulfide bridges. *Protein Sci.* **9**, 942–955 (2000).
58. Proutiere, A. et al. Gallocin A, an Atypical Two-Peptide Bacteriocin with Intramolecular Disulfide Bonds Required for Activity. *Microbiol. Spectr.* **11**, e0508522 (2023).
59. Fregeau Gallagher, N. L. et al. Three-dimensional structure of leucocin A in trifluoroethanol and dodecylphosphocholine micelles: spatial location of residues critical for biological activity in type IIa bacteriocins from lactic acid bacteria. *Biochemistry* **36**, 15062–15072 (1997).
60. Zhu, L., Zeng, J., Wang, C. & Wang, J. Structural Basis of Pore Formation in the Mannose Phosphotransferase System by Pediocin PA-1. *Appl. Environ. Microbiol.* **88**, e0199221 (2022).
61. Bechinger, B. & Salnikov, E. S. The membrane interactions of antimicrobial peptides revealed by solid-state NMR spectroscopy. *Chem. Phys. Lipids* **165**, 282–301 (2012).
62. Shultzaberger, R. K., Bucheimer, R. E., Rudd, K. E. & Schneider, T. D. Anatomy of *Escherichia coli* ribosome binding sites. *J. Mol. Biol.* **313**, 215–228 (2001).
63. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797 (2004).
64. Madeira, F. et al. Using EMBL-EBI Services via Web Interface and Programmatically via Web Services. *Curr. Protoc.* **4**, e1065 (2024).
65. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).

Acknowledgements

We warmly thank F. Van Bambeke and F. Goormaghtigh for providing *S. aureus* strains, and A. Gillis for the use of the BSL-2+ facility. The work of P.H. was supported by the Belgian National Fund for Scientific Research (FNRS, grants PDR T.0110.18/T.0111.22 and CDR J.0090.21), the Concerted Research Actions (ARC, grants 17/22-084 and 22/27-120) from Federation Wallonia-Brussels, and the Syngulon company. The work of J.W.V. was supported by the Swiss National Science Foundation (SNSF grants 310030_192517, 310030_200792 and NCCR 51NF40_180541). J.D. and A.K. held doctoral fellowships from FNRS (FRIA fellowship). F.J. was supported by SWP Research from Walloon Region (grant N° 8065 Staphcontrol). J.Mi. received funding from the European Union's Horizon 2020 research and innovation program (Marie Skłodowska-Curie grant N° 101018461). P.H. is Research Director at the FNRS.

Author contributions

J.D., P.G., J.Mi., and P.H.; conception and design. J.D., J.Mi., M.S., and F.J.; acquisition of data. J.D. and A.K.; creation of the bioinformatic script. J.D., J.Mi., J.-W.V., and P.H.; analysis and interpretation of data. J.D., P.G., J.Mi., J.-W.V., J.Ma., and P.H., draft or revising the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare the following competing interests: P.H., J.Mi., J.-W.V., P.G., F.J., and J.D. declare that they are listed as inventors on patent(s) or patent application(s) related to bacteriocin production and uses. P.H. is a member of the Scientific Advisory Board of the Syngulon company. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-07217-y.

Correspondence and requests for materials should be addressed to Pascal Hols.

Peer review information *Communications Biology* thanks Shaynoor Dramsi, Suzanne Dawid and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Christopher LaRock and Tobias Goris.