

# Bacteria from the *Amycolatopsis* genus associated with a toxic bird secrete protective secondary metabolites

---

Received: 26 September 2023

---

Accepted: 3 September 2024

---

Published online: 02 October 2024

---

 Check for updates

---

Elena Seibel<sup>1,2,12</sup>, Soohyun Um<sup>2,3,12</sup>, Kasun H. Bodawatta<sup>4</sup>, Anna J. Komor<sup>5</sup>, Tanya Decker<sup>1</sup>, Janis Fricke<sup>1</sup>, Robert Murphy<sup>6</sup>, Gibson Maiah<sup>7</sup>, Bulisa Iova<sup>8</sup>, Hannah Maus<sup>9</sup>, Tanja Schirmeister<sup>9</sup>, Knud Andreas Jønsson<sup>4,10</sup>, Michael Poulsen<sup>6</sup> & Christine Beemelmans<sup>1,2,11</sup> ✉

Uropygial gland secretions of birds consist of host and bacteria derived compounds and play a major sanitary and feather-protective role. Here we report on our microbiome studies of the New Guinean toxic bird *Pachycephala schlegelii* and the isolation of a member of the *Amycolatopsis* genus from the uropygial gland secretions. Bioactivity studies in combination with co-cultures, MALDI imaging and HR-MS/MS-based network analyses unveil the basis of its activity against keratinolytic bacteria and fungal skin pathogens. We trace the protective antimicrobial activity of *Amycolatopsis* sp. PS\_44\_ISF1 to the production of rifamycin congeners, ciromicin A and of two yet unreported compound families. We perform NMR and HR-MS/MS studies to determine the relative structures of six members belonging to a yet unreported lipopeptide family of pachycephalamides and of one representative of the demiguisins, a new hexapeptide family. We then use a combination of phylogenomic, transcriptomic and knock-out studies to identify the underlying biosynthetic gene clusters responsible for the production of pachycephalamides and demiguisins. Our metabolomics data allow us to map molecular ion features of the identified metabolites in extracts of *P. schlegelii* feathers, verifying their presence in the ecological setting where they exert their presumed active role for hosts. Our study shows that members of the Actinomycetota may play a role in avian feather protection.

The chemical ecology of bird–bacteria interactions has received considerable attention over past decades as wild birds act as reservoirs for emerging zoonotic bacterial pathogens<sup>1</sup>. Birds have evolved a variety of protective measures against predators and microbial pathogens, including sanitary behaviors, toxins, gland secretions, and presumed protective microbial symbionts<sup>2–5</sup>. In particular, the uropygial gland (UG), a holocrine secretory gland situated at the dorsal base of a bird's tail, contributes to waterproofing and safeguarding the plumage and feathers against microbial pathogens and ectoparasites; mainly through the secretion of an odoriferous, antimicrobial preen oil that is

a mixture of waxes and volatiles<sup>6–12</sup>. Studies analyzing the bacterial communities and their possible contributions within UGs and feathers across bird species have documented pronounced prevalence of  $\gamma$ - and  $\alpha$ -Pseudomonadota, occasionally accompanied by Actinomycetota, as well as keratinolytic feather-degrading bacteria (FDB), such as *Bacillus licheniformis* or *Kocuria rhizophila*, that can negatively affect feather integrity and thus likely avian health<sup>3,4,11–14</sup>. Based on co-cultures, it was anecdotally hypothesized that the avian microbiome could encompass protective bacterial species to counteract activities of FDB. In one example, *Enterococcus faecalis* isolates from the

---

A full list of affiliations appears at the end of the paper. ✉ e-mail: [Christine.beemelmans@helmholtz-hips.de](mailto:Christine.beemelmans@helmholtz-hips.de)

Eurasian hoopoe preen oil was indeed inhibitory against FDB due to the secretion of bacteriocins (enterocin MR10)<sup>8,15</sup>.

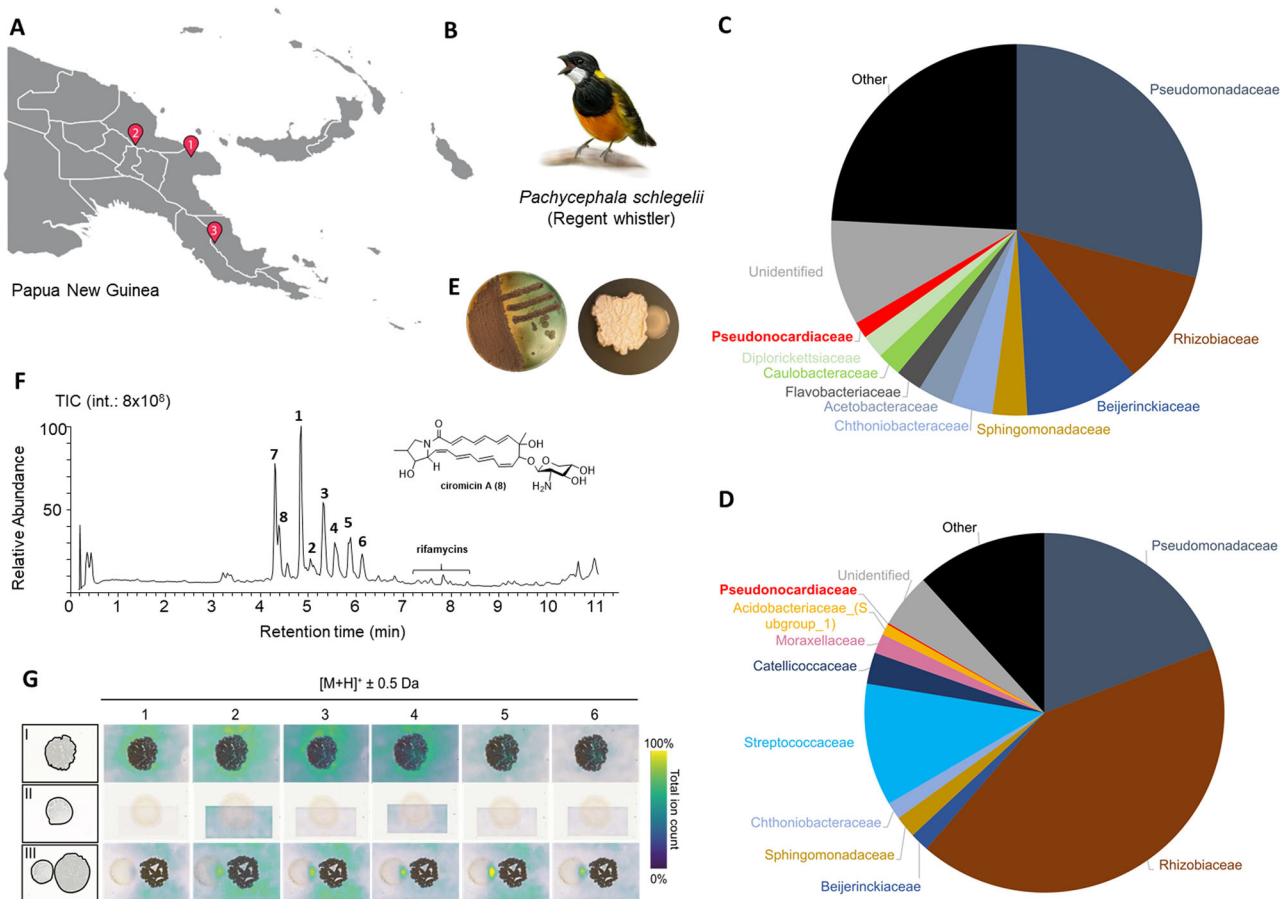
However, despite the growing body of evidence underscoring the importance of bacterial symbionts in avian health, chemical evidence remains scarce. To address this knowledge gap, and to explore whether bacterial symbionts could contribute to the protection against FDB by chemical means, we conducted a comprehensive chemical-ecological study using our model organism, the New Guinean toxic bird *Pachycephala schlegelii* (Regent Whistler) known to integrate the neurotoxic steroid alkaloid batrachotoxin (BTX) into feathers to likely deter predators<sup>2</sup>. Amplicon sequencing allowed us to map the microbial community of UG secretions and feathers of *P. schlegelii* individuals, and guided isolation allowed cultures of microbes with putative protective functions. One isolate belonging to the genus *Amycolatopsis* was singled out for analysis due to its antibacterial activity against keratinolytic bacteria and antifungal activity against skin pathogens. By tracing the bioactivity, we uncovered not only the production of known antibiotics, but also the production of two new compound families of nonribosomal peptide (NRP) origin that had yet unreported structural modifications. By combining phylogenomic, transcriptomic and knock-out studies we identified and verified the biosynthetic gene clusters responsible for the production of the antimicrobial natural products. We also detected molecular ion features of the identified metabolites in extracts of *P. schlegelii* feathers, verifying their presence in the ecological setting where they exert their

presumed active role. To our knowledge, this study represents the first example of the isolation of bioactive NRPs from a bird UG isolate, with two compound classes exhibiting unique structural features and biochemistry, and also provides the first evidence that Actinomycetota may play an important role in maintaining healthy plumage, setting the path for further investigations on natural product producers at the bird-microbe interface.

## Results and discussion

### MiSeq microbiome analysis and isolation of bacterial symbionts

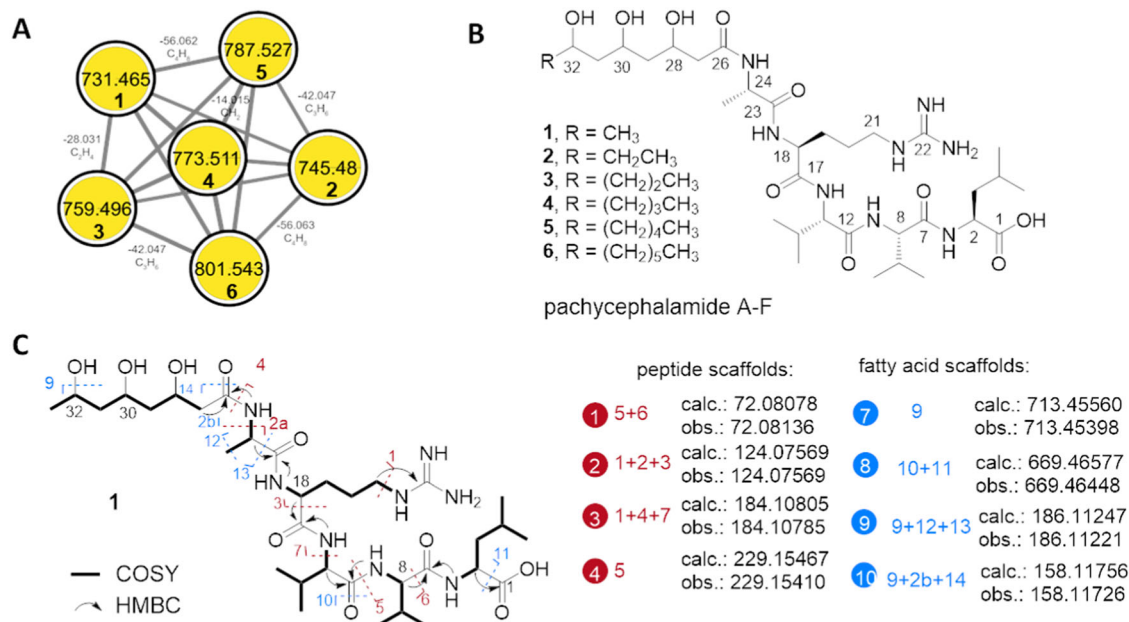
Between 2018 and 2019, we collected twelve *P. schlegelii* individuals from three different sampling sites (Fig. 1A, B). All twelve individuals were subjected to UG microbiome studies, and nine out of twelve individuals were also included in analyses of the feather microbiome (Supplementary Fig. 1, Supplementary Table 1). Illumina MiSeq 16S rRNA amplicon sequencing yielded reads derived from feather samples that were dominated by Pseudomonadota (70.5%), followed by Actinomycetota (10.1%), and Bacteroidetes (5.8%)<sup>3</sup>, while the microbiomes of UG were composed primarily of members of Pseudomonadota (74.1%), followed by Firmicutes (15.5%), and Actinomycetota (3.7%) (Fig. 1C, D; Supplementary Data 1), and thus differed notably from temperate bird species<sup>4</sup>. Subsequently, UG secretions of five *P. schlegelii* specimens were investigated for culturable bacterial strains, which yielded 23 morphologically different bacterial colony forming units, of which isolate PS\_44\_ISF1 was of interest due to its inhibitory



**Fig. 1 | Microbiome study and analysis of the UG isolate *Amycolatopsis* sp. PS\_44\_ISF1.** **A** Sampling locations of *P. schlegelii* specimen in Papua New Guinea; **B** Sketch of *P. schlegelii*; and relative averaged abundances of the ten most abundant bacterial families in **C** feather (nine biological samples) and **D** uropygial glands (UG) (twelve biological samples), including in both cases the Pseudonocardaceae to which *Amycolatopsis* belongs. **E** Representative plate culture of *Amycolatopsis* sp.

PS\_44\_ISF1 alone (top) and in co-cultivation with *B. licheniformis* (bottom).

**F** Representative total ion chromatogram (TIC) of methanolic extracts of PS\_44\_ISF1 showing compounds 1–8 as dominating metabolites with chemical structure of the known metabolite ciromicin A (8), and **G** MALDI-IMS images of *Amycolatopsis* sp. PS\_44\_ISF1 co-cultures with *B. licheniformis* showing the abundance of molecular ion features of compounds 1–6.



**Fig. 2 | Structure elucidation of pachycephalamides A-F.** **A** GNPS-based network of molecular ion features of pachycephalamides A-F (**1–6**); **B** chemical structure of **1–6**; **C** 2D-NMR correlations of pachycephalamide A (**1**) and annotated MS/MS

fragments with matching calculated  $m/z$  values predicted by MassFrontier 8.0 (<sup>1</sup>H-<sup>1</sup>H COSY: Homonuclear correlation spectroscopy; <sup>1</sup>H-<sup>13</sup>C HMBC: Heteronuclear multiple-bond correlation spectroscopy).

activity against other bacterial strains (Fig. 1G). Based on its 16S rRNA sequence, we identified PS\_44\_ISF1 to be closely related to the type strain *Amycolatopsis saalfeldensis* DSM 44993, which belongs to the family Pseudonocardiaceae (Supplementary Figs. 2 and 3, Supplementary Tables 2 and 3)<sup>16,17</sup>.

To enable a more accurate phylogenomic placement of *Amycolatopsis* sp. PS\_44\_ISF1, a hybrid genome assembly was created by combining paired-end shotgun sequencing with long-read sequencing from PacBio and Nanopore technologies (Supplementary Table 4). Subsequently, whole genome-based taxonomic analysis was performed using the Type (Strain) Genome Server (TYGS)<sup>18</sup> as well as digital DNA-DNA-hybridization (dDDH d4 36.8% to closest strain)<sup>19</sup>, which revealed that strain PS\_44\_ISF1 is likely a new species of the *Amycolatopsis* clade C<sup>20</sup> to which its closest relatives *Amycolatopsis* sp. M39<sup>21</sup> and *A. saalfeldensis* DSM 44993 belong (Supplementary Data 2). This finding was supported by the similar but distinct fatty acid profile of the strain compared to closely related type strains (Supplementary Table 5)<sup>22–24</sup>. To assess whether the isolate could thrive within the UG gland or on feathers, we performed keratinolytic and lipolytic activity tests. In both assessments, *Amycolatopsis* sp. PS\_44\_ISF1 showed the capability to utilize feather or oil as a carbon source for growth (Supplementary Figs. 4 and 5). We then re-examined if sequencing reads assigned to *Amycolatopsis* were also detectable in our feather and UG microbiome data, and indeed Pseudonocardiaceae accounted for up to 1.5% of the feathers, and 0.14% of the UG microbiome, which included detectable amplicon sequence variants assigned to *Amycolatopsis* in almost half of the samples (Supplementary Fig. 1E). Altogether, our results were intriguing as members of the Pseudonocardiaceae are well recognized for their defensive symbiotic interactions<sup>25</sup>, most prominently in the fungus-growing ant symbiosis, where *Pseudonocardia* has aided in protection against specialized parasites of the ants' fungal gardens for millions of years<sup>26,27</sup>.

### Antimicrobial properties and secondary metabolome of *Amycolatopsis* sp. PS\_44\_ISF1

We then assessed the antimicrobial properties of *Amycolatopsis* sp. PS\_44\_ISF1 in plate-based co-cultivation assays against ecologically

relevant microbial pathogens, including FDB (*B. licheniformis*, *Pseudomonas monteilii*, *K. rhizophila*), non-feather degrading bacteria (*B. thuringiensis*, *Staphylococcus epidermidis*), and fungal pathogens (*Aspergillus fumigatus*, *Aspergillus nidulans*, *Candida albicans*) (Supplementary Figs. 6–20). Indeed, strain PS\_44\_ISF1 exhibited moderate antifungal activity against the human-pathogenic yeast *C. albicans* and *Aspergillus* spp. as well as growth inhibitory activity against feather-degrading *B. licheniformis* and *K. rhizophila* (Fig. 1G, bottom). We then tested if antimicrobial metabolites caused the observed growth inhibition by analyzing methanolic extracts of the microbial interaction zone using tandem high-resolution mass spectrometry (HR-MS/MS) and global network analyses<sup>28</sup> with interconnected natural-product databases. For dereplication purposes, we also compared the metabolome of *Amycolatopsis* sp. PS\_44\_ISF1 with two of its close relatives, *Amycolatopsis saalfeldensis*<sup>16,17</sup> and *Amycolatopsis* sp. M39 (Supplementary Fig. 21)<sup>21</sup>.

Indeed, a plethora of putative secondary metabolite features were detectable from extracts of *Amycolatopsis* sp. PS\_44\_ISF1 and the interaction zones of co-cultures, including polyene macrolactam ciromicin A (**8**)<sup>29,30</sup>. Furthermore, database dereplication also uncovered known and PS\_44\_ISF1-specific molecular ion features belonging to the antibiotic rifamycin family<sup>31,32</sup>, but which were in low abundance. However, the number of annotated features was low compared to the abundance of detectable strain-specific unknown features, with two clusters dominating the metabolome. One cluster was comprised of  $m/z$  features ranging from  $[M + H]^+$  731.465 to 801.540 (**1–6**, Fig. 2A) in which nodes differed by  $\Delta m/z$  14 (CH<sub>2</sub> units), while the second cluster (**7** with  $[M + H]^+$  694.351) showed a distinct peptide-related MS/MS pattern (Fig. 3A).

To visualize the molecular ion features in microbial co-culture settings, we also performed Matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-IMS) analysis of single and co-cultures<sup>33</sup>. Here, we observed in most cases comparable higher abundances of molecular ion features  $[M + H]^+$  731.5–801.4 (**1–6**) within the interaction zone of microbial co-cultures in comparison to axenic growth, while the abundance of the molecular ion feature  $[M + Na]^+$  716.5 (**7**) and others appeared to remain mainly unchanged (Fig. 1G, Supplementary Figs. 17–20). At this stage, we concluded that



*Amycolatopsis* sp. PS\_44\_ISF1 produces a series of antimicrobial secondary metabolites that likely account for its observed antimicrobial activity.

### Characterization of peptide-based metabolites

To enable the characterization of the two yet unknown compound families, we evaluated the effect of different growth conditions, including eco-mimetic settings, on their production (Supplementary Figs. 22 and 23). Overall, relative production of both compound families appeared highest on PDA agar plate cultivation between 7 and 10 days. Following up on this finding, large scale cultivation (120 PDA plates, 90 × 16 mm) and methanolic extraction was followed by an MS- and NMR-guided isolation strategy using semi-preparative reversed-phase column (C<sub>18</sub>) chromatography (HPLC). During the course of purification, we isolated seven yet unreported metabolites that we named pachycephalamides A-F (**1–6**, 2.1–7.6 mg) and demiguisin (**7**, 2.7 mg) (Fig. 2).

The molecular formula of pachycephalamide A (**1**) was determined as C<sub>33</sub>H<sub>62</sub>O<sub>10</sub>N<sub>8</sub> based on HR-ESI-MS (calcd. [M+H]<sup>+</sup> *m/z* 731.4662, obsd. [M + H]<sup>+</sup> *m/z* 731.4653), which indicated seven degrees of unsaturation. The <sup>1</sup>H and <sup>13</sup>C NMR spectra of **1** (CD<sub>3</sub>OH-*d*<sub>3</sub>) showed well dispersed proton and carbon resonances indicating a pattern of chemical shifts for a typical peptidic compound as bearing five amide/acid carbonyl signals (δ<sub>C</sub> 175.4, 173.9, 173.8, 173.7, and 173.3), five α-carbon signals (δ<sub>C</sub> 61.2, 61.1, 53.2, 51.7, and 51.3), and five α-proton signals (δ<sub>H</sub> 4.26, 4.16, 4.12, 3.96, and 3.84). In addition to these typical peptidic resonances, pachycephalamide A showed an aliphatic signal pattern including three hydroxyl carbon signals (δ<sub>C</sub> 71.8, 69.8, and 65.9). Further analysis of 1D (<sup>1</sup>H and <sup>13</sup>C) and 2D (COSY, HSQC, and HMBC) NMR data together with MS/MS fragmentation patterns enabled the identification of five amino acid residues (Leu, Ala, Arg, and two Val) and a 3,5,7-trihydroxyoctanoic acid moiety (Supplementary Figs. 24–60, Supplementary Tables 6–8).

As the molecular ion features of pachycephalamide B-F (**2–6**) differed by Δ14 Da (CH<sub>2</sub>) with otherwise similar fragmentation pattern, it was deduced that the trihydroxyalkyl acid chain varied only by the number of methylene units (Supplementary Methods). The absolute configuration of each amino acid residue in pachycephalamides was analyzed by LC-MS guided Marfey's method using *L*-FDAA (1-fluoro-2,4-dinitrophenyl-5-alanine amide) as the derivatization agent (Supplementary Tables 9 and 10), and determined as *L*-configuration, which is in line with the *in silico* substrate specificity analysis. Sequence alignment of characterized ketoreductase (KR) domains with the putatively responsible KR domain within the PKS cluster region ctg2\_1673 (*vide infra*) inferred an all B-type KR stereospecificity. Thus, we propose that all three hydroxyl moieties exhibit *R*-configuration. Overall, pachycephalamides are lipopeptides carrying an unusual odd or even numbered trihydroxylated fatty acid tail, which to the best of our knowledge has not yet been reported.

The molecular formula of demiguisin (**7**) was deduced to be C<sub>30</sub>H<sub>47</sub>N<sub>9</sub>O<sub>10</sub> (*m/z* 694.3520 [M + H]<sup>+</sup>, caclcd. for C<sub>30</sub>H<sub>48</sub>N<sub>9</sub>O<sub>10</sub>, 694.3519), indicating 12 degrees of unsaturation (Supplementary Figs. 61–71, Supplementary Table 8). The HSQC NMR spectrum displayed resonances that are typical of a peptide possessing six amino methines, which was consistent with <sup>13</sup>C NMR signals assigned to six amide/acid carbonyl signals. 1D and 2D NMR data analyses enabled us to identify the presence of three unmodified amino acid residues with a glycine (Gly) moiety at the C-terminus and a leucine (Leu), and piperazic acid (Piz) moiety incorporated into the peptide chain. In addition, atypical amino acid signals such as oxygenated methine correlations and olefin signals were detectable, suggesting the presence of three (oxidatively) modified alkenyl amino acid units. Next to the C-terminal glycine, we deduced the presence of the oxygenated amino acid 4,5-dihydroxyhexahydropyridazine-3-carboxylic acid (β,γ-OH-Piz), for which a characteristic proton spin system consisting of a

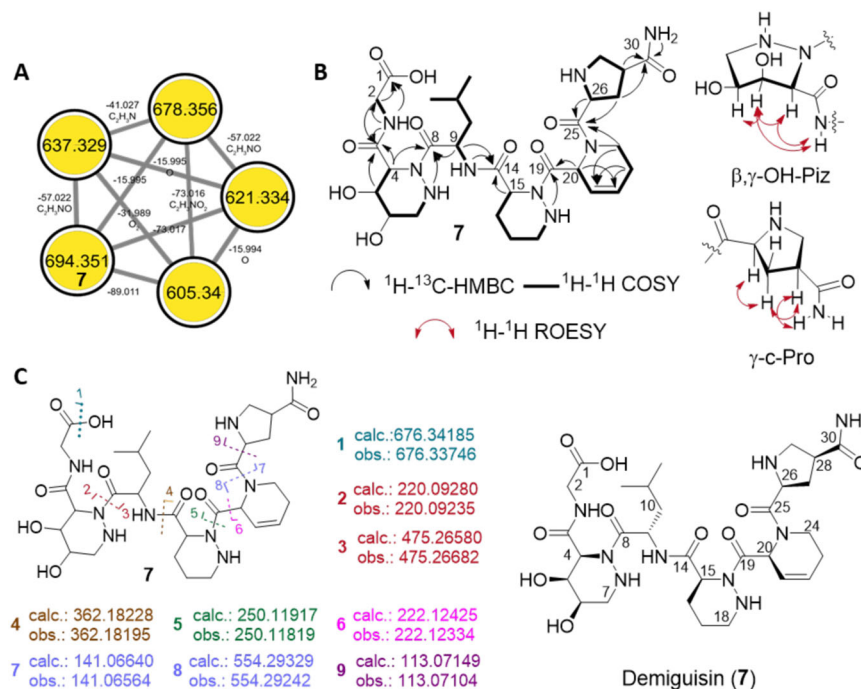
methine proton H-4 (δ<sub>H</sub> 4.76), oxygenated methine protons H-5 (δ<sub>H</sub> 4.17) and H-6 (δ<sub>H</sub> 3.39), an methylene proton H<sub>2</sub>-7 (δ<sub>H</sub> 3.15 and 2.59), and a secondary amine 7-NH (δ<sub>H</sub> 3.17) (H-4/H-5/H-6/H<sub>2</sub>-7/7-NH) was assigned, and which was supported by HMBC correlations from H-4 to C-3 (δ<sub>C</sub> 168.4). The planar structure of the 3,4-dehydropipecolic acid (DHPA, 3,4-dehydro-homoproline) was deduced from COSY and TOCSY correlations between H-20 (δ<sub>H</sub> 5.75)/H-21 (δ<sub>H</sub> 5.88)/H-22 (δ<sub>H</sub> 5.93)/H<sub>2</sub>-23 (δ<sub>H</sub> 2.20 and 2.05)/H<sub>2</sub>-24 (δ<sub>H</sub> 3.96, 3.38) together with the formation of a spin system ranging from the methine carbon C-20 (δ<sub>C</sub> 51.9) to the epsilon methylene carbon C-24 (δ<sub>C</sub> 40.3). HMBC correlations of the α-proton (H-20, δ<sub>H</sub> 5.75) and δ-proton (H-23b, δ<sub>H</sub> 2.05) to γ-olefin carbon (C-21, δ<sub>C</sub> 122.3), and from ε-proton (H-24a, δ<sub>H</sub> 3.96) to β-olefin carbon (C-22, δ<sub>C</sub> 126.9) aligned with the deductions. A partial structure of the third modified amino acid, γ-carbamoyl-proline (γ-c-Pro), was deduced from COSY and TOCSY correlations as proton–proton couplings between the methine proton H-26 (δ<sub>H</sub> 3.98), the β-protons H<sub>2</sub>-27 (δ<sub>H</sub> 2.33 and 1.67), the γ-methine proton H-28 (δ<sub>H</sub> 2.84), and the methylene protons H<sub>2</sub>-29 (δ<sub>H</sub> 3.04 and 2.90) (H-26/H<sub>2</sub>-27/H-28/H<sub>2</sub>-29) were observed. The HMBC correlations from primary amine proton signals (30-NH<sub>2</sub>, δ<sub>H</sub> 7.34 and 6.82), β-proton (H-27a, δ<sub>H</sub> 2.33), and δ-proton (H-29a, δ<sub>H</sub> 3.04) to C-30 (δ<sub>C</sub> 175.7) allowed to deduce the planar structure of γ-c-Pro. The amino acid sequence of Gly–β,γ-OH-Piz–Leu–Piz–DHPA–γ-c-Pro (Fig. 3B), which satisfies the 12 degrees of unsaturation, was inferred from the molecular formula, sequence-typical MS/MS fragments and supported by HMBC and ROESY correlations from α-proton signals to amide carbonyl signals of neighboring amino acids. Marfey's analysis allowed again to determine the absolute configuration of the unmodified amino acids Leu and Piz<sup>34</sup> (Supplementary Fig. 61, Supplementary Table 10).

Overall demiguisin (**7**) represents a linear hexapeptide containing three modified amino acid motifs (β,γ-OH-Piz, 3,4-DHPA and γ-carbamoyl-Pro). Peptides containing Pip or Piz, are regularly found to be part of NRPS-derived peptides and known to possess various biological properties, including antibacterial, antifungal and cytotoxic activity, and can act as biofilm inhibitors<sup>34,35</sup>.

In contrast, α,β-dehydroamino acids are only occasionally found with α,β-(2,3)-dehydropipecolic acid in ulleungamides as one of the few examples<sup>36–38</sup>. Even less common are examples for peptides containing β,γ-dehydroamino acids or γ-carbamoyl proline. Similarly, γ-hydroxy proline<sup>39</sup> – a major constituent of animal collagen – was found only occasionally as a structural feature in natural products, such as antifungal heinamides<sup>40</sup> and pneumocandins<sup>41</sup>, while the structurally close γ-carboxy proline has not yet been reported from NRPS-derived natural products but itself acts as a glutamate uptake inhibitor<sup>42</sup>.

### Biosynthetic considerations and validation

To identify the biosynthetic gene clusters (BGCs) responsible for the production of pachycephalamide A-F (**1–6**) and demiguisin (**7**), we employed a combination of phylogenomic, transcriptomic and mutagenesis studies (Fig. 4). The genome of isolate PS\_44\_ISF1 was first mined using the web-based antiSMASH7.0 platform<sup>43</sup> and build-in NRPSpredictor2<sup>44</sup>, followed by manual curation of the resulting query hits with respect to the number of modules, adenylation (A) and condensation (C) domains and specificity-conferring codes that matched the amino acid sequence incorporated in the peptide structure (Supplementary Tables 11–13, Supplementary Data 3 and 4). We also performed phylogenetic analyses (maximum likelihood)<sup>45</sup> of encoded and reference domains retrieved from the NaPDos database<sup>46</sup> (Supplementary Figs. 72–75). In parallel, we pursued a transcriptomic analysis of RNAseq data from cultures that were grown on either a lipopeptide production medium (PDA) or low production medium (CSA) and then mapped the transcript levels of the encoded genes within the boundaries of the gene clusters region (Supplementary Figs. 76–78).



**Fig. 3 | Structure elucidation of peptide demiguisin (7).** **A** GNPS network showing molecular ion features including the one of demiguisin (7). **B** Proposed absolute structure of 7. **C** 2D-NMR correlation of demiguisin and detected MS/MS fragments with atom numbering for NMR assignments. A set of fragments matching

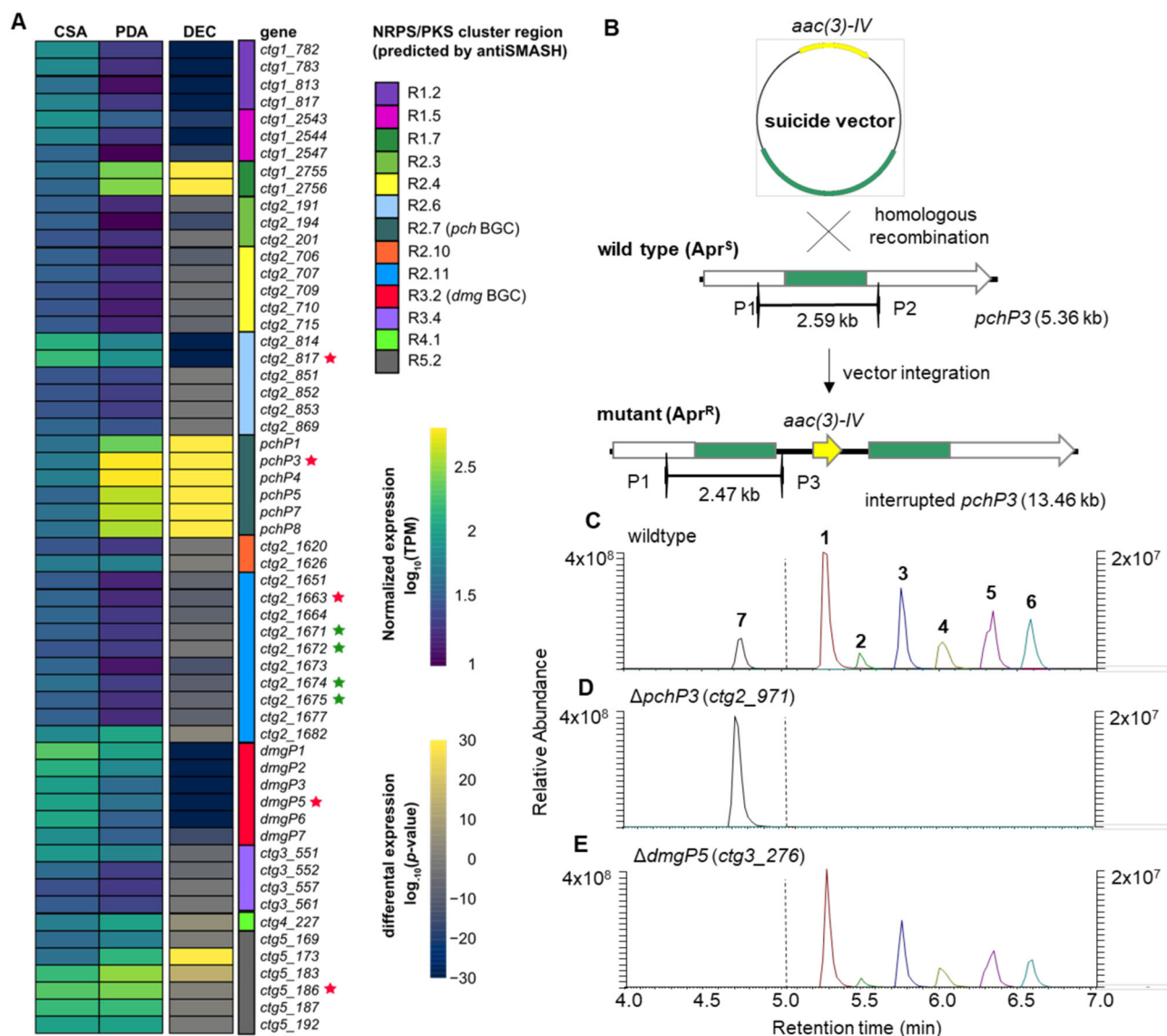
calculated  $m/z$  values predicted by MassFrontier 8.0 ( $^1\text{H}-^1\text{H}$  COSY: homonuclear correlation spectroscopy;  $^1\text{H}-^{13}\text{C}$  HMBC: Heteronuclear multiple-bond correlation spectroscopy;  $^1\text{H}-^1\text{H}$  ROESY: Rotating-frame nuclear Overhauser effect spectroscopy).

First, we focused on the biosynthesis of the lipopeptides pachycephalamides, for which we hypothesized a hybrid gene cluster arrangement composed of a dedicated non-ribosomal peptide synthase (NRPS)-machinery interconnected with either a fatty acid synthase (FAS) or a highly reducing type II PKS enzyme machinery, which was also reported for the structurally related ishigamides and other unsaturated polyene-type natural products<sup>47–51</sup>. Typically, the elongation of the acyl chain through a peptide scaffold necessitates the presence of a starter condensation domain (Cs) located at the initial module of the NRPS assembly line<sup>52,53</sup>. Genome and transcriptome mining uncovered 13 putative NRPS, PKS or NRPS/PKS hybrid candidate cluster regions, but only two of which carried a fitting number of A-domains within an NRPS encoding region (R2.6, R2.11, Supplementary Fig. 77), as well as a neighboring PKS or FAS encoding region. To obtain more insights on the putative candidate clusters, we then analyzed and compared transcript levels of genes within the boundaries of the 13 PKS or NRPS cluster regions (Fig. 4). Two additional NRPS-encoding regions (R2.7 and R5.2) showed differential expression levels and encoded NRPS-like domain architectures that fitted the deduced compound structure. However, none of the regions encoded PKS-related genes up and downstream of the cluster, which led us to hypothesize that these BGC could also be fragmented and/or proceeding via a non-canonical pathway.

We then pursued a knock-out strategy targeting four core NRPS genes and two KS-CLF gene pairs within the four most likely NRPS or PKS-encoding regions, which included the two highly expressed regions R2.7 (*ctg2\_971*) and R5.2 (*ctg5\_186*), as well as the NRPS/PKS hybrid regions R2.6 (*ctg2\_817*) and R2.11 (NRPS: *ctg2\_1663*, KS/CLF: *ctg2\_1671/2*, *ctg2\_1674/5*). For this, a suicide vector was constructed by amplifying an internal fragment (~2 kb) of the target gene, which was then cloned into the SphI/XbaI-linearized *E. coli-Streptomyces* shuttle vector pKJ55 (Fig. 4B). The vector was transferred to *Amycolatopsis* sp. PS\_44\_ISF1 by conjugation using the methylation-deficient helper strain *E. coli* ET12567/pUZ8002<sup>54,55</sup>. As the vector is not able to

integrate (removed integrase) or replicate, it integrates via heterologous recombination, thereby interrupting the target gene in the *Amycolatopsis* genome (Supplementary Figs. 79–81, Supplementary Tables 14 and 15, Supplementary Data 3). Overall, we were able to obtain knock-outs for all four NRPS-encoding genes and verified apramycin-resistant ex-conjugants were analyzed for the production of pachycephalamides by HR-MS/MS (Supplementary Figs. 82–86). Overall, only one NRPS-knock-out of *ctg2\_971* (region 2.7) abolished pachycephalamide production (Fig. 4C), while the relative production levels of other secondary metabolites remained unchanged.

Region 2.7 (*ctg2\_971*) carries the BGC cluster now termed *pch*, which spans approx. ~29 kb (20 open reading frames) and encodes eight NRPS-related genes (*pchPI-PS*), four genes putatively annotated as oxidoreductases (*pchORI-OR4*), four regulator genes (*pchRI-R4*), and several genes with other functions (Fig. 5A). A closer in silico analysis of the encoded NRPS machinery indicated that the four encoded adenylation (A) domains, including one putative iterative domain, showed substrate-conferring codes recognizing Ser (PchP4-A1), Arg (PchP1-A1), Thr (PchP7) and Leu (PchP3-A1). Although PchP4-A1 (Ser) and PchP7 (Thr) showed deviating amino acid conferring codes, A-domain promiscuity<sup>56</sup> might allow the activation of the structurally related Ala and Val, suggesting the formation of the deduced chemical substructure of pachycephalamides. The phylogenetic analyses allowed us to identify a C domain (PchP4-C1) that was classified as an LCL domain (Supplementary Fig. 72), but also showed some relation to the C-domains SpoT10 (sporolide biosynthesis)<sup>57</sup> and SgcC5 (C-1027 biosynthesis)<sup>58</sup>, both catalyzing the condensation of two biosynthetic intermediates (Fig. 5C). Thus, we hypothesized that the PchP4 C-domain could catalyze the condensation (lipoinitiation)<sup>53</sup> reaction of the trihydroxylated PKS/ fatty acid precursor and the first amino acid *L*-Ala, while the second minimal NRPS module encoded by PchP1 could catalyze the downstream condensation with *L*-Arg (Fig. 5D). The NRPS module PchP8 is assumed to catalyze the subsequent biosynthetic step, despite its unusual architecture (C-PCP-C)



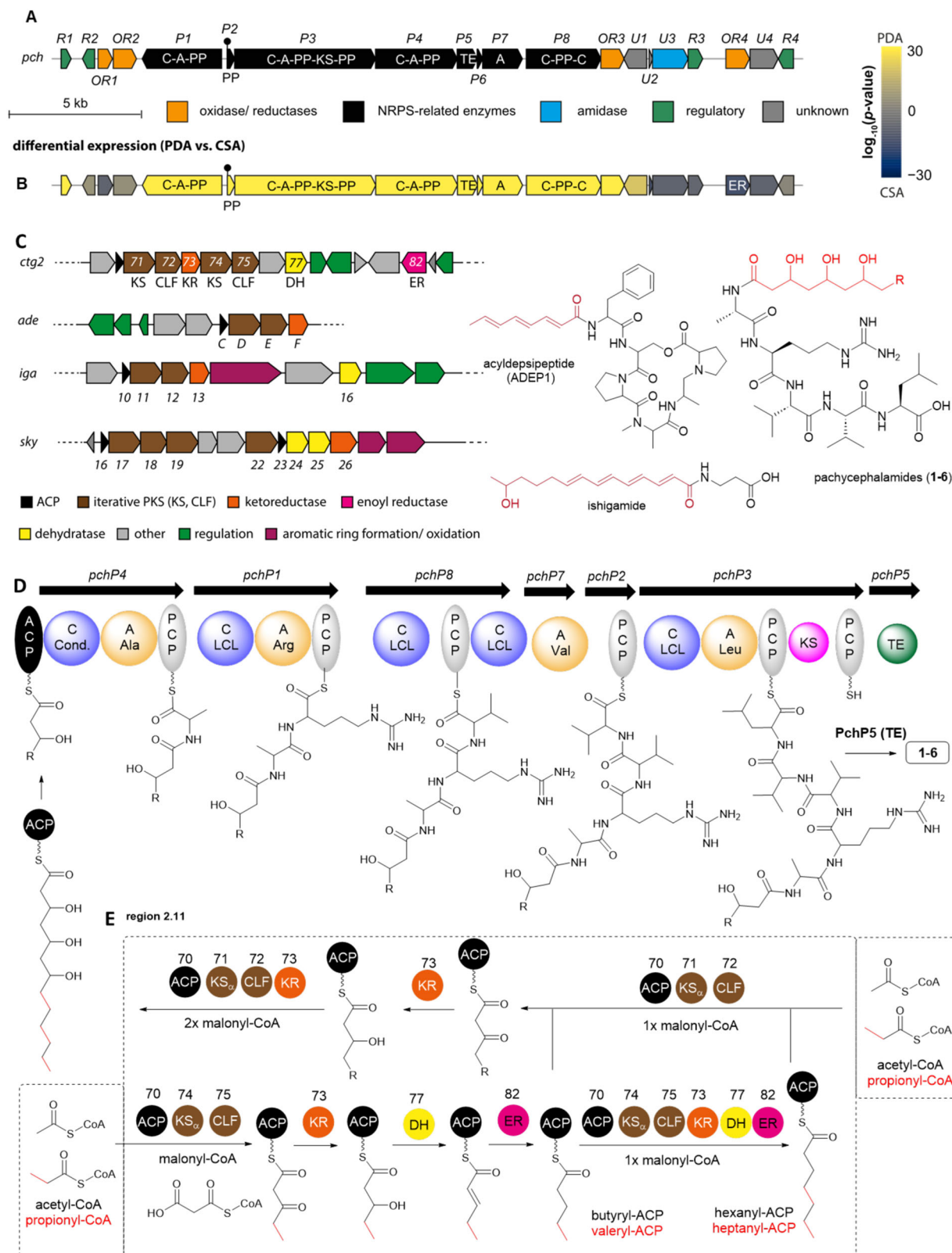
**Fig. 4 | Transcriptomic studies and knock-out experiments for the verification of pachycephalamide and demiguisin biosynthetic gene clusters.** **A** Analysis of the differential expression of NRPS-encoding cluster regions encoded in the PS\_44\_ISF1 genome when comparing growth on a lipopeptide high-production (PDA) and low-production (CSA) medium (DEC differential expression confidence,  $n=1$ , Supplementary Figs. 76–78, source data are provided as a Source Data file). Five NRPS-encoding genes (red star) and two PKS-CLF gene pairs (green star) were

chosen for further knock-out studies to validate the biosynthetic proposals. **B** Schematic illustration of the generation, selection, and screening of PS\_44\_ISF1 knock-out mutants (Apr<sup>S</sup>: apramycin-sensitive, Apr<sup>R</sup>: apramycin-resistant). **C** Extracted ion chromatograms for the detection of compounds 1–7 in the generated knock-out mutants. Compared to the wild-type (**C**), pachycephalamide production was abolished in **D**  $\Delta pchP3$  and production of demiguisin was abolished in mutant **E**  $\Delta dmgP5$ .

lacking adenylation domains. Here, we identified an upstream located stand-alone A domain PchP7 with proposed Thr/Val specificity, which could act iteratively in *trans* and interact with both C domains of PchP7, introducing two Val moieties to the growing peptide scaffold. While iteratively and in *trans* acting domains are unusual, they are not unprecedented. In WS9326A biosynthesis<sup>59</sup>, one C-domain was shown to interact successively with two stand-alone A-domains. Lastly, PchP3 is supposed to incorporate Leu, and the mature lipopeptide is likely released by the stand-alone thioesterase PchP5. Although the *pch* cluster encoded a ketosynthase (KS) domain (PchP3) in its proximity, its function in the biosynthesis of the unusual odd and even-numbered 3,5,7-trihydroxylated fatty acids remains hypothetical as it lacked ketoreduction (KR), dehydration (DH) or enoylreduction (ER) domains. Based on homology searches we propose that the type II PKS of region 2.11 complements the biosynthesis instead, as it harbors two

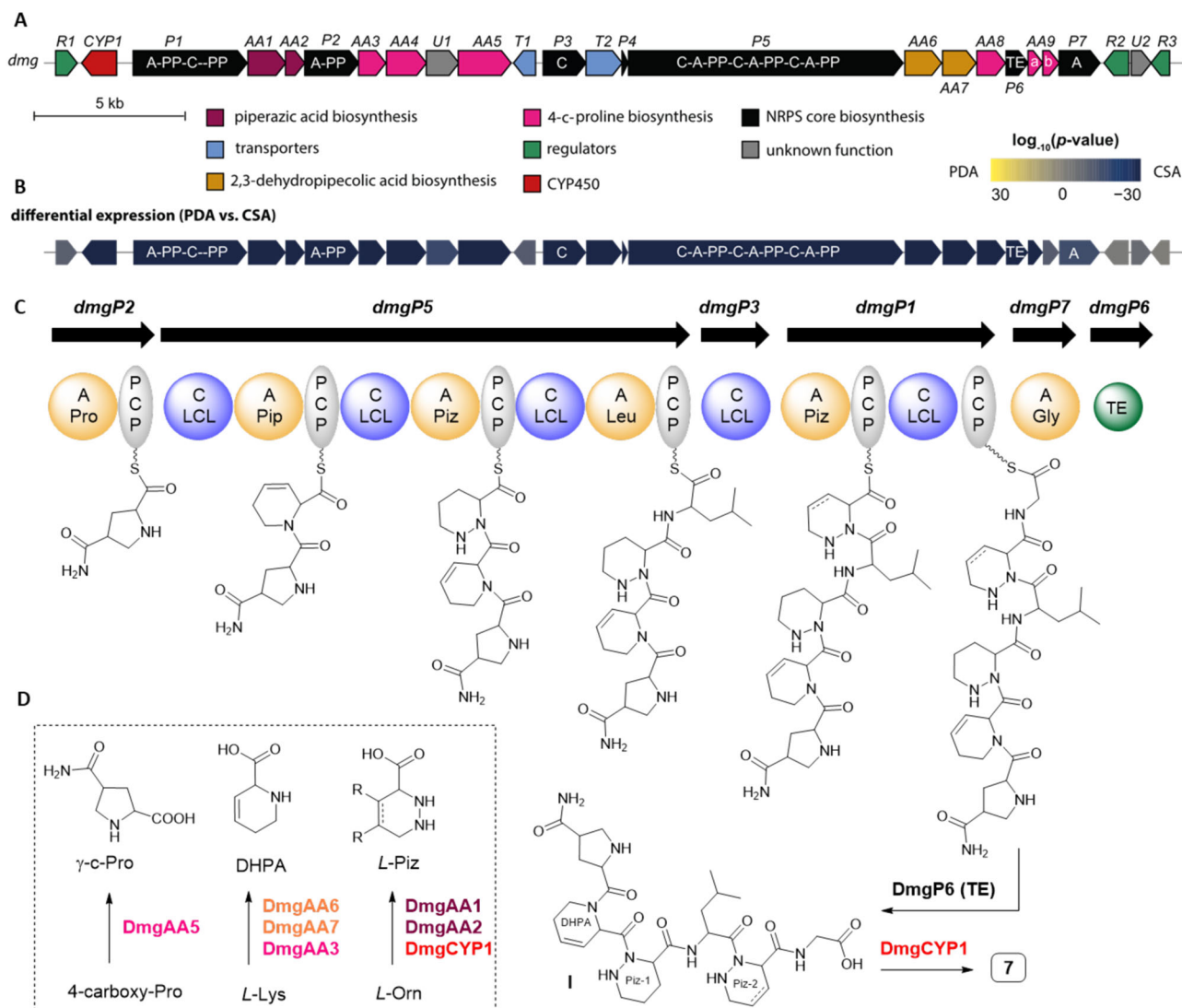
pairs of type II KS domains accompanied with a chain length factor (CLF) with high similarity to the non-aromatic type II PKS, such as skyllamycin (*sky*), ishigamide (*iga*) and ADEP1 (*ade*) (Fig. 6E)<sup>50,52</sup>. The first CFL/KS pair (*ctg2\_1671/2*) co-locates with a discrete KR domain (*ctg2\_1673*), while the second pair (*ctg2\_1674/5*) co-locates with a DH (*ctg2\_1677*) and enoylreductase (ER, *ctg2\_1682*) domain. Based on these findings, we hypothesize that the reducing KS-CLF machinery (*ctg2\_1674/5*) is able to use acetyl-CoA as well as propionyl-CoA. Two additional elongation steps using malonyl-CoA result ultimately in hexanoyl-ACP or heptenyl-ACP. Assuming that these ACP-bound intermediates are recognized by the partially reducing KS-CLF pair *ctg2\_1671/2*, subsequent Claisen condensations followed by reduction of the carbonyl group would allow introduction of the 3,5,7-trihydroxylation pattern into a variety of different fatty acid intermediates. Based on the alignment of ketoreductases, the KR domain in *ctg1\_1673*





**Fig. 5 | Maps of biosynthetic pathways involved in pachycephalamides biosynthesis. A** Organization of the *pch* BGC; **B** differential expression of *pch* genes when comparing growth on a lipopeptide high-production (PDA) and low-production (CSA) medium ( $n=1$ , source data are provided as a Source Data file); **C** comparison of reducing type II PKS clusters and module numbering with type II PKS clusters of skyllamycin (*sky*), ishigamide (*iga*) and ADEP1 (*ade*) and chemical

structures of skyllamycin, ishigamide and ADEP1. **D** NRPS assembly line for the production of pachycephalamides, and **E** interlinked PKS-based pathway responsible for the production of 3,5,7-trihydroxy fatty acid with varying  $\text{CH}_2$ -units (ACP acyl carrier protein, KS ketosynthase, CLF chain length factor, DH dehydratase, ER enoyl reductase, TE thioesterase, C condensation domain, A adenylation domain; presentation of stereochemistry was omitted for clarity).



**Fig. 6 | Biosynthesis of demiguisin (7).** **A** Organization of the *dmG* BGC and **B** differential expression of *dmG* genes when comparing growth on a lipopeptide high-production (PDA) and low-production (CSA) medium (Source data are provided as a Source Data file,  $n = 1$ ); **C** NRPS assembly line for the production of a proposed demiguisin precursor I, including the origin of the modified amino acid

precursors (Supplementary Fig. 87), and **D** proposed transformation/enzymes for precursor biosynthesis (PCP peptidyl carrier protein, C condensation domain, A adenylation domain, TE thioesterase, Pip piperazine, Piz piperazic acid, Gly glycine, Pro proline, Lys lysine, Leu leucine, Orn ornithine, presentation of stereochemistry was omitted for clarity).

was determined as B-type, thus the stereochemistry of three hydroxyl moieties is proposed to be *R* configured (Supplementary Fig. 75)<sup>60</sup>. In a final step, the respective acyl chains could be linked to the peptide backbone by the first C domain in the pachycephalamide NRPS assembly line.

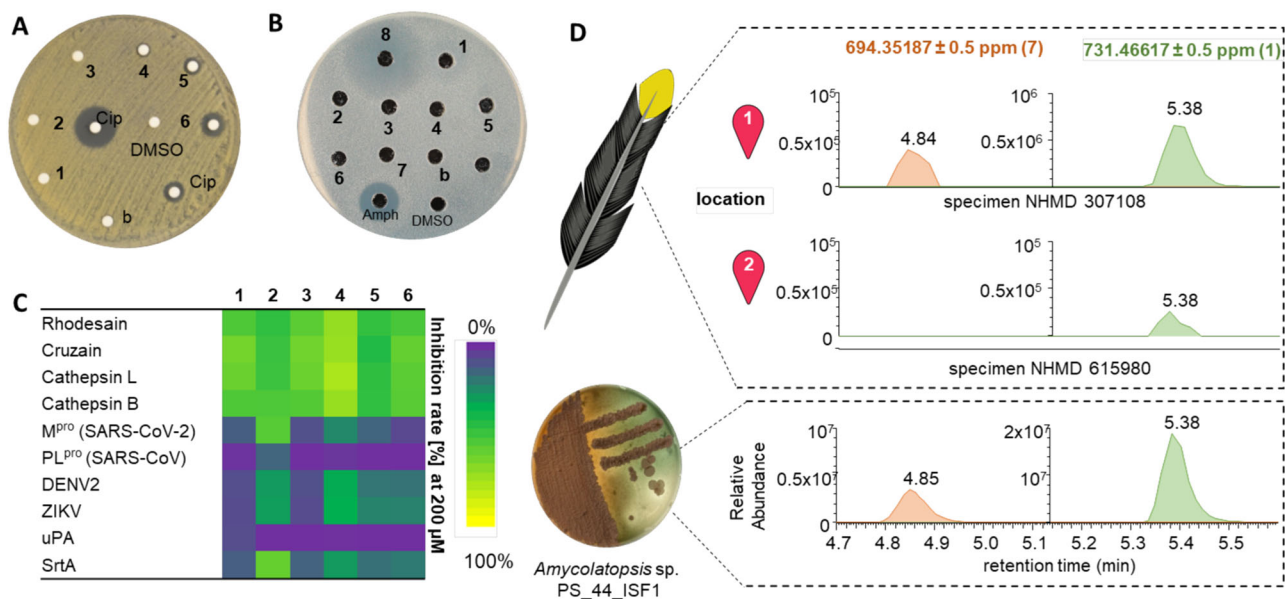
For demiguisin (7) we identified one gene cluster region (*dmG*) encoding a multimodular NRPS carrying six A domains, which we deemed responsible for its biosynthesis (Supplementary Data 4). The gene cluster region *dmG* spanned about 38.2 kb and encoded at least 25 open reading frames (*dmGR1-R3*), including seven NRPS-related genes (*dmGP1-P7*), nine genes encoding enzymes responsible for building block formation and modification (*dmGAA1-AA9*, *dmGCYP1*), five genes coding for transport (*dmGT1-T2*) and regulatory (*dmGR1-R3*) functions, among others (Fig. 6). Transcripts of all genes located within the boundaries of the BGC region were found in RNAseq data from two growth conditions, which correlated with production of demiguisin (7). We then analyzed the A domains encoded in the cluster region for their substrate specificities and found that four of the stand-alone domains showed specificity for Pro/Pip (DmgP1-A1, DmgP2-A1, and

DmgP5-A1/A2), one for Leu (DmgP5-A3) and one for Gly (DmgP7-A1). Of the four A domains proposed to recognize Pro or Pip, two shared a 90% identical specificity-conferring code and were also related to the consensus code of Piz-activating A domains (DmgP1-A1 and DmgP5-A2, Supplementary Table 13)<sup>34,35</sup>.

To verify that this particular NRPS is responsible for the assembly of the peptide backbone of demiguisin (7), we again prepared knock-out mutants by interrupting the key NRPS-encoding gene within the *dmG* cluster (*dmGP5*). As expected, demiguisin production in apramycin-resistant ex-conjugants carrying the interrupted NRPS-encoding gene *dmGP5* was completely abolished (Supplementary Figs. 82–86), while the relative production levels of other secondary metabolites remained unchanged, thus supporting the theory that the *dmG* cluster should be responsible for the biosynthesis of the peptide scaffold of 7.

Based on this dataset, the biosynthetic assembly is proposed as follows (Fig. 6C). The core structure of 7 is likely build first, by activation and loading of the  $\gamma$ -carbamoyl-Pro or  $\gamma$ -carboxyl-Pro ( $\gamma$ -c-Pro) on to NRPS DmgP2, followed by transfer of the PCP tethered  $\gamma$ -c-Pro to





**Fig. 7 | Bioactivity and ecological relevance of compounds.** Disc-diffusion assays of isolated compounds **1–8** against **A** *K. rhizophila* and **B** *C. albicans* (DMSO, blanks (b), and antibiotic controls (Ciprofloxacin: AmpH; Amphotericin B). **C** Inhibition of selected proteases [%] by compounds **1–6** (200  $\mu$ M,  $n = 3$ ). **D** Detection of

compounds **1** and **7** in methanolic extracts of feathers (sketch) of two *P. schlegelii* specimens from different locations with a culture extract of strain PS\_44\_ISF1 as reference.

DmgP5 and the step-wise incorporation of the amino acids DHPA, Piz and Leu (Fig. 6). The second (modified) Piz moiety is presumably activated by the A domain of DmgP1 and coupled to the growing peptide backbone via the catalytic action of the stand-alone C domain DmgP3. Glycine is likely activated by the discrete A domain DmgP7 and incorporated by the C domain of DmgP1. The discrete thioesterase DmgP6 should release the mature peptide backbone. All of the necessary C domains of the NRPS modules were found to cluster with known LCL domains, which supported our deduction from prior analytical datasets (Supplementary Fig. 72). To which degree the respective (discrete) A domains are capable of incorporating natural and modified amino acids or whether the peptide chain is modified after the release from the NRPS remains a topic of investigations.

We also identified genes encoding for the necessary precursor biosynthesis within the boundaries of the cluster (Fig. 6D, Supplementary Fig. 87, Supplementary Data 4). Piperazine acid (Piz) biosynthesis is likely catalyzed by DmgAA1-AA2, which show similarities to *L*-ornithine-N-hydroxylase KtzI (51%) and the heme-dependent piperazate synthase KtzT (55%), commonly present in BGCs of Piz-containing natural products<sup>34,35</sup>. Here, we propose that  $\beta,\gamma$ -hydroxylation of the Piz-moiety might occur via the action of cytochrome P450 monooxygenase DmgCYP1 either before incorporation or after the release from the NRPS. 3,4-Dehydropipecolic acid (DHPA) is likely formed from *L*-lysine via  $\alpha$ -transamination of  $\alpha$ -ketoglutarate catalyzed by the putative *L*-lysine-2-aminotransferase DmgAA6, which shows similarities to the aminotransferases NikC or CrmC reported in the nikkomycin<sup>51</sup> or caerulomycin A<sup>62</sup> biosynthesis (Supplementary Data 4). As proposed in the nikkomycin biosynthesis, the  $\alpha$ -keto intermediate spontaneously cyclizes and dehydrates to  $\Delta^1$ - or  $\Delta^2$ -piperidine-2-carboxylate (P2C), which can be further oxidized by FAD-dependent oxidoreductase DmgAA7 homologous to NikD/CrmD resulting in a  $\Delta^1$ - $\Delta^3$ -P2C intermediate. The encoded reductase DmgAA3 is proposed to reduce the  $\Delta^1$  double bond, yielding DHPA.

The biosynthesis of the  $\gamma$ -carbamoyl proline ( $\gamma$ -c-Pro) moiety remains unclear at this stage as two possible biosynthetic scenarios were deduced from our in silico analysis. In the first possible scenario, condensation reaction of *L*-Ser with oxalacetate by DmgAA4, an enzyme homologous to tryptophan synthase-like enzyme CourR1,

should yield 4-amino-1-oxobutane-1,2,4-tricarboxylic acid, which is then decarboxylated via the action of the decarboxylase DmgAA9a/b to yield  $\Delta^5$ -4-carboxyproline (Supplementary Fig. 74). Reduction of the  $\Delta^5$ -double bond could be induced by DmgAA3, an oxidoreductase sharing homology with GriH in griselimycin or HrmD in hormaomycin biosynthesis<sup>63</sup>. In an alternative biosynthetic route Leu is first hydroxylated by hydroxylase GriE and then subsequently oxidized to an aldehyde by dehydrogenase GriF. The resulting intermediate cyclizes by forming  $\Delta^5$ -4-methyl-Pro which is reduced to 4-methyl-Pro by GriH and is then oxidized to 4-carboxy-Pro via 4-hydroxymethylproline. In both scenarios, amidation of 4-carboxy-Pro could occur either before or after release from the NRPS via the putative amidotransferase DmgAA5, a homolog of amidotransferase SsFD reported in the SF2575 biosynthesis<sup>64</sup>.

### Antimicrobial activity and occurrence in feathers

Peptides containing Pip, Piz or  $\alpha$ -3,4-DHPA, similar to lipopeptides, are well known to have antimicrobial activities and often confer functions in cell motility and biofilm formation<sup>65</sup>. We thus subjected the isolated compounds (**1–7**) to various different bioassays to explore their activity range and pharmacological properties.

For lipopeptides (**1–6**), we found moderate antibacterial activity against several indicator strains as well as the feather-degrading species *K. rhizophila* (Fig. 7A, Supplementary Table 16), which are likely caused by the amphiphilic character and the chain-length dependent tensioactive properties of this substance family (Fig. 7A, Supplementary Fig. 88)<sup>65,66</sup>. In contrast, demiguisin (**7**) appeared to have no antimicrobial effects, while the co-eluting macrolactam ciromycin A (**8**) and rifamycin congeners exhibited antifungal and antibacterial effects as previously reported (Fig. 7B, Supplementary Fig. 89)<sup>30</sup>.

Due to the antibacterial properties of pachycephalamide, we went on to test **1–6** against a panel of different proteases of pharmacological and ecological relevance (Fig. 7C, Supplementary Tables 17–19), including the bacterial transpeptidase sortase A of *Staphylococcus aureus*<sup>67</sup>, which is a well-known virulence factor during infections with these bacteria. In addition, we included cathepsin L, cathepsin B<sup>68</sup> and urokinase (uPA) into the screen<sup>69</sup>, as they are related to many cancer types and tumor-associated fibroblasts and macrophages. We also

included cathepsin-like proteases rhodesain<sup>70</sup> and cruzain<sup>71</sup>, which are central in the life cycle of the parasitic hemoflagellate protists *Trypanosoma brucei* and *Trypanosoma cruzi*, the major cause of African trypanosomiasis (*T. brucei*), also known as “sleeping sickness”, and Chagas disease (American trypanosomiasis, *T. cruzi*). Validated targets for new antivirals, such as the proteases (M<sup>pro</sup>, PL<sup>pro</sup>) of SARS-coronavirus (SARS-CoV) and SARS-CoV-2<sup>72</sup> and the NS2B-NS3 proteases of the flaviviruses ZIKV and DENV<sup>73</sup> (Zika and Dengue) were also tested. A single-concentration assay (200  $\mu$ M) uncovered that compounds **1–6** moderately inhibited the structurally related cysteine proteases rhodesain, cruzain, cathepsin L and cathepsin B in a structure-related fashion. Finally, we tested if metabolic features of *Amycolatopsis* sp. PS\_44\_ISF1 were also detectable in methanolic extracts of feather samples of *P. schlegelii* specimens<sup>2</sup>. Indeed, targeted analysis revealed molecular ion features of the most abundant secondary metabolites **1** and **7** in two specimens collected from two different regions with the same retention time and fragmentation pattern as the isolated and characterized compounds (Fig. 7D, Supplementary Fig. 71).

Wild bird species host intricate microbial communities throughout various body regions, yet our comprehension of the roles these microbiomes play trails behind that of many other vertebrates. Following up on our findings that the relative microbial composition between the UG system and feathers of the toxic bird species *P. schlegelii* show notable differences, we were able to culture a fraction of the predicted UG and feather microbiome including a new member and likely a new species of the genus *Amycolatopsis*, which includes species well recognized for their defensive symbiotic interactions and protection against specialized parasites. With *Amycolatopsis* sp. PS\_44\_ISF1 as a representative of the bird protective microbiome, we analyzed its keratinolytic and lipolytic activities and found evidences for its ability to thrive in the bird gland and feather environment. Subsequent co-cultivation experiments coupled with MALDI-IMS analysis revealed the strains ability to suppress common bacterial and fungal pathogens found in birds, including FDB. The observed antimicrobial activity was attributed to the co-secretion of several classes of natural products, including the antimicrobial polyene macrolactam ciromicin A, the low-abundant ansamycin-derived rifamycin congeners, as well as the two newly discovered peptide classes of NRPS origin, both exhibiting unique structural features. While pachycephalamides A-F (**1–6**) represent pentapeptides with a 3,5,7-trihydroxylated fatty acid (C<sub>8</sub>-C<sub>13</sub>) moiety linked to the N-terminus, demiguisin (**7**) is a hexapeptide containing various modified amino acids, such as  $\beta$ ,  $\gamma$ -hydroxylated piperazic acid, dehydropipecolic acid  $\gamma$ -carbamoylproline. To the best of our knowledge, these represent the first examples of NRPS-derived natural products isolated from a bacterial strain originating from an UG. We further were able to detect members of both peptide classes in feathers of *P. schlegelii*, supporting that the strain and its metabolome might play a beneficial protective role for the host bird. Altogether, our results indicate that along with gland secretions, bacteria derived from uropygial glands and secreting antimicrobial natural products are likely applied onto the feathers to inhibit the growth of harmful microbes.

By integrating phylogenetic, transcriptomic, and mutagenetic studies, we propose that two predominantly non-canonical biosynthetic gene clusters are accountable for the biosynthesis of the two new compound classes. Both clusters include *trans*-acting stand-alone as well as iteratively acting NRPS domains, and a fascinating complementary pair of (iterative) PKS modules. Our molecular biological studies serve now as the foundation for further elucidation of the enzymatic reactions responsible for producing these unusual features. A deeper understanding of these mechanisms will facilitate targeted efforts to identify and exploit these unique features in other natural product producers. Overall, our study also reaffirms the potential inherently present when exploring microbial communities in

underexplored ecological niches, offering opportunities for the identification of new bacterial species, innovative natural product producers, and previously undisclosed biochemical processes.

## Methods

### Ethics

The fieldwork protocol, including ethical standards follows Guidelines to the Use of Wild Birds in Research formulated by The Ornithological Council. Samples were collected in 2018 and 2019 at three locations in Papua New Guinea (PNG) under the research permits no. 99902341112 (to Kasun Bodawatta) and no. 99902260244 (to Knud Jønsson) and exported under export permits no. 019067 in 2018 and no. 019362 and no. 019423 in 2019 by the PNG Immigration & Citizen Service Authority and the Department of Environment and Conservation, Boroko, Papua New Guinea.

### MiSeq sequencing of UG and feather samples

UGs were collected from thirteen *P. schlegelii* specimens obtained from three locations in Papua New Guinea under the research permit 99902341112 in the years 2018 and 2019. Birds were humanely euthanized through rapid cardiac compression following the Natural History Museum of Denmark collection guidelines, and directly dissected on a sterile field tray. While UG were stored in RNALater™ until DNA extractions, ~20  $\mu$ g of breast feather samples from each of the individual were stored in individual Ziploc bags, initially, at room temperature (approx. 20 °C), and once reached the field station, samples were kept in -20 °C until further analyses. DNA from UG and a subset of the feather samples (4–5 breast feathers) were extracted following the manufacture protocol. Samples were sent to the Microbiome Core at the University of Michigan for MiSeq amplicon sequencing, using primer pair SA701 and SB511 targeting the V4 region of the bacterial 16S rRNA gene under an Illumina sequencing platform. For more details, see Supplementary Methods.

### Analysis of MiSeq sequencing data

MiSeq amplicon sequences were analyzed using the QIIME<sup>74</sup> and DADA2<sup>75</sup> pipelines with default settings. Sequences were assigned to amplicon sequence variants (ASVs) at the 100% similarity and subsequently assigned to taxonomy using the SILVA 138.1 bacterial reference database<sup>76</sup>. We further removed archaeal, mitochondrial and chloroplast sequences from the dataset. Using the phyloseq package<sup>77</sup> we analyzed the bacterial alpha diversities (chao 1 richness estimate and Shannon's diversity index) of the microbiomes. We visualized the microbial community compositions (based on Bray-Curtis dissimilarity) using NMDS plots and investigated the statistical difference between microbiomes using adonis 2 test (Permutational multivariate analyses of variance: PERMANOVA) in vegan package (<https://CRAN.R-project.org/package=vegan>).

### Isolation of UG-associated bacteria

UG secretions were applied directly onto a Potato Dextrose Agar medium (PDA, 32 g of PDA mixed with 800 mL of water) containing 50 mg cycloheximide per liter of media. Plates were kept in room temperature under aerobic conditions. Once the colony forming units (CFUs) appeared, morphologically different colonies were isolated and reinoculated in new PDA media without cycloheximide.

### Strain identification by amplicon sequencing

Strain PS\_44\_ISF1 was sub-cultured on MS agar, and single colonies stored in 25% glycerol at -80 °C. Bacterial DNA from the strain was acquired using Qiagen DNeasy® Blood and Tissue kit (Hilden, Germany), with a 24-h incubation period. Initial Sanger sequencing was conducted at the Eurofins genomics (Copenhagen, Denmark), with the bacterial 16S rRNA gene primer pair 27 F and 1492 R to identify the bacterial isolate<sup>78</sup>.

## Whole genome sequencing and analysis

*Amycolatopsis* sp. PS\_44\_ISF1 cultures (50 mL) were prepared in ISP2 broth and cultivated for seven days at 30 °C and shaking at 150 rpm. Bacterial pellets were harvested, frozen in liquid nitrogen and ground to a fine powder. DNA was extracted using DNeasy Plant Mini Kit (Quiagen) according to manufacturer's instructions and subsequently precipitated using isopropanol. The precipitated DNA was washed with 95% ethanol (at least 5 times), dried and dissolved in water before sequencing. Whole-genome sequencing was performed using a 150 bp paired-end shotgun (BGISEQ) and long-read (PacBio sequel) sequencing at BGI. Additionally, Oxford Nanopore technology (Oxford Nanopore Technologies, Oxford, UK) was employed for long-read sequencing. The MinION sequencing library was prepared using the Rapid DNA sequencing kit (SQK-RAD4) according to the manufacturer. DNA sequencing was performed on a MinION Mk1B sequencing device equipped with a R9.4.1 flow cell, which was prepared and run according to the manufacturer.

Sequencing results were checked for quality using FastQC version 0.11.8<sup>79</sup> and MultiQC version 1.7<sup>80</sup>. K-mer depth was calculated using Jellyfish version 2.2.10<sup>81</sup> and Kmer-based estimates of genome size, heterozygosity, and repeat content generated using GenomeScope<sup>82</sup>. Nanopore sequencing raw data was generated using MinkNOW software version 4.0.20 (Oxford Nanopore Technologies), and was base-called and trimmed using Guppy version 4.2.2 (Oxford Nanopore Technologies). The resulting fastq files were filtered using Nanofilt<sup>83</sup>. A hybrid de novo genome assembly, combining PacBio and Oxford Nanopore data, was performed using MaSuRCA version 3.4.1<sup>84</sup>. The resulting draft assembly was then polished with the accurate Illumina reads using the POLCA genome polisher<sup>85</sup>.

## HR-UPLC-ESI-MS/MS analysis

High resolution ultra-high performance liquid chromatography electrospray ionization mass spectrometry (HR-UPLC-ESI-MS) measurements were carried out on a Dionex Ultimate3000 system (Thermo Scientific) combined with a Q-Exactive Plus mass spectrometer (Thermo Scientific) equipped with an electrospray ion (ESI) source and a Luna Omega C18 column (100 × 2.1 mm, particle size 1.6 μm, pore diameter 100 Å). For structure elucidation MS/MS spectra were predicted using CFM-ID 4.0 and compared to the experimental MS/MS spectra. Each sample was analyzed using HR-UPLC-MS<sup>2</sup> measurements using a Dionex Ultimate3000 system (Thermo Scientific) (gradient: 0–1 min, 5% B, 1–7 min, 5–97% B, 7–9 min, 97% B, 9–10 min; 97–5% B, 10–13 min, 5% B) with an injection volume of 5 μl and a flow rate of 0.3 mL/min. All eluents of the UHPLC system were acidified with 0.1% formic acid. H<sub>2</sub>O was used as eluent A while acetonitrile was used as eluent B. For each analytical runs, washing steps and blank runs were included (at least every 5 biological samples) and served as control samples to check for cross-contamination. For biological samples one technical replicate was measured. For structure elucidation MS/MS spectra were predicted using CFM-ID 4.0 or MassFrontier 8.0 (Thermo Fisher Scientific) and compared to the experimental MS/MS spectra.

## Network-based analysis

Tandem mass spectrometry molecular networks were created using the GNPS platform (<http://gnps.ucsd.edu>)<sup>28</sup>. Data (.RAW) were first converted to the .mzML format with MS-Convert<sup>86</sup>. The converted files were used to generate an MS/MS molecular network using the GNPS Data Analysis workflow version release 30. The precursor and fragment ion mass tolerance were set to 0.02 Da and to a product ion tolerance of 0.02 Da. Networks were generated using six minimum matched ion fragments, a minimum cluster size of 2 and a cosine score of 0.8. The remaining parameters were kept at default. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7 and at least 6 matched peaks.

After analysis, data were opened and visualized using Cytoscape 3.8.0 software.

## Co-cultivation experiments

*Amycolatopsis* sp. PS\_44\_ISF1 was cultivated on MS agar for 7 days at 30 °C. Indium-Tin-Oxide (ITO)-coated slides (IntelliSlides, Bruker) were coated by aluminum foil and sterilized by autoclaving. Slides were placed into a petri dish (90 × 16 mm) and 10 ml of PDA were distributed evenly on top of the slide. A spore solution (5 μL, OD<sub>600</sub> of 0.1) was placed in the center of the PDA-coated slides. *Amycolatopsis* cultures were incubated at 30 °C for a total period of six days. After one (=day 2) and three days (=day 4) of incubation 5 μl of a competitor strain culture or spore solution (OD<sub>600</sub> = 0.1) was added in a proximity of 1 cm to the *Amycolatopsis* culture. Competitor strains used included three feather-degrading bacteria (*B. licheniformis* DSM13, *K. rhizophila* DSM11926, *P. monteilii* DSM1388), filamentous fungi (*A. fumigatus*, *A. niger*) and the yeast *C. albicans* SC5314. Incubation of co-cultures were pursued until the end of the experiment. Axenic cultures of *Amycolatopsis* sp. PS\_44\_ISF1 and each competitor strain inoculated at the same time as the co-cultures served as controls. For analysis of metabolite abundance, a defined agar piece was extracted in a pre-weighed beaker after a given time of incubation using MeOH overnight at room temperature. The resulting methanolic extract was dried *in vacuo* and the resulting residue was dissolved in 100% MeOH to a final concentration of 50 μg/ml and subjected to HR-UPLC-ESI-MS analysis. The production of each metabolite of interest was evaluated by the comparison of the peak area using the extracted ion mode (EIC) and normalized to the wet weight of agar extracted (peak area/mg wet weight).

## Matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-IMS)

Cultures on Indium-Tin-Oxide (ITO)-coated slides were dried for 5 h at 37 °C. The dried samples were then sprayed with a saturated solution (20 mg/mL) of universal MALDI matrix (1:1 mixture of 2,5-dihydroxybenzoic acid and α-cyano-4-hydroxy-cinnamic acid; Bruker Daltonics, Bremen Germany) prepared in acetonitrile/methanol/water (70:25:5, v/v/v), using the automatic system ImagePrep device 2.0 (Bruker Daltonics, Bremen Germany) in 60 consecutive cycles (the sample was rotated 180° after 30 cycles) of 41 s (1 s spraying, 10 s incubation time, and 30 s of active drying)<sup>87</sup>. The sample was analyzed in an UltrafleXtreme MALDI TOF/TOF (Bruker Daltonics, Bremen Germany). For further details, see Supplementary Methods.

## Monitoring of product formation under different growth conditions

For analysis of production levels, *Amycolatopsis* sp. PS\_44\_ISF1 pre-cultures (50 ml) were prepared in ISP2 broth and cultivated for seven days at 30 °C and shaking at 150 rpm. Agar plates of different media were inoculated with 100 μl of the pre-culture and further incubated at 30 °C for a given time period. Three replicates of each medium were extracted after a given time of incubation by cutting agar into small pieces (0.5 cm × 0.5 cm). After transfer into a pre-weighed 100 mL beaker, the weight of the transferred agar was determined. Agar pieces were extracted with 40 mL MeOH overnight at room temperature. The solvent was then filtered off and evaporated *in vacuo*. The resulting residue was dissolved in 100% MeOH to a final concentration of 50 μg/ml and subjected to HR-UPLC-ESI-MS analysis. The production of each metabolite of interest was evaluated by the comparison of the peak area using the extracted ion mode (EIC) and normalized to the wet weight of agar extracted (peak area/mg wet weight).

## Large scale cultivation for natural product isolation

For isolation from agar plates, 500 μl 7-day-old pre-cultures of *Amycolatopsis* spp. were streaked on 120 PDA (peptides) agar plates (90 × 16 mm). The cultures were sealed with parafilm and incubated at



30 °C for another 7 days. Plates were cut in small cubes (1 × 1 cm) and extracted with 6 L of methanol overnight. Afterwards, the solvent was filtered off and evaporated *in vacuo*. The dried extract was then redissolved in 20% methanol (MeOH) and subjected to solid phase extraction (SPE). A 10 g C<sub>18</sub> column was activated with 2 column volumes (CV) of 100% MeOH and equilibrated with 20% MeOH in dH<sub>2</sub>O. Afterwards the extract was loaded, washed with 20% MeOH, and eluted with 50% and 100% MeOH (2 CV). All three fractions were concentrated *in vacuo* yielding all together 98.5 g of extract. For peptide isolation, the combined 50% and 100% fractions were used for further processing and first resuspended with Celite in MeOH and then dried *in vacuo* to generate Celite-adsorbed extract in a round flask. The Celite-adsorbed extract was loaded onto 20 g of the pre-packed C18 SPE Sepak resin. The extract was fractionated by elution with a step gradient composed of dH<sub>2</sub>O and methanol (10%, 20%, 40%, 60%, 80%, and 100% MeOH). Macrolactam ciromicin A and demiguisin (7) eluted in the 40% MeOH SPE fractions, while pachycephalamides A-F (1–6) eluted in the 60% and 80% MeOH SPE fractions, and rifamycin congeners in the 100% MeOH SPE fractions. Fractions were purified by semipreparative reversed-phase HPLC (Phenomenex Luna C18(2) 250 × 10 mm column, particle size 5 μm, pore diameter 100 Å, flow rate: 2 mL/min, detection: UV 210 nm, gradient solvent system: 10–40% aqueous acetonitrile over 30 min). For further analytical experimental details, see Supplementary Methods.

**Pachycephalamide A (1):** white solid;  $[\alpha]_D^{25}$  –5.67 (c 0.1 w/v%, MeOH); UV (acetonitrile),  $\lambda_{\max}$  204 nm; IR (ATR)  $\nu_{\max}$  3778, 3213, 2783, 2363, 1694, 1550, 1021 cm<sup>-1</sup>; NMR spectral data, see Supplementary Tables 6 and 7; HRMS (ESI)  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>33</sub>H<sub>63</sub>O<sub>10</sub>N<sub>8</sub> 731.4662, found 731.4653,  $\Delta_{\text{ppm}}$  –0.84.

**Pachycephalamide B (2):** white solid;  $[\alpha]_D^{25}$  –2.28 (c 0.1 w/v%, MeOH); UV (acetonitrile),  $\lambda_{\max}$  204 nm; IR (ATR)  $\nu_{\max}$  3743, 3621, 3311, 2702, 2322, 1616, 1562 cm<sup>-1</sup>; NMR spectral data, see Supplementary Tables 6 and 7; HRMS (ESI)  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>34</sub>H<sub>65</sub>O<sub>10</sub>N<sub>8</sub> 745.4829, found 745.4802,  $\Delta_{\text{ppm}}$  –2.81.

**Pachycephalamide C (3):** white solid;  $[\alpha]_D^{25}$  –135.89 (c 0.1 w/v%, MeOH); UV (acetonitrile),  $\lambda_{\max}$  204 nm; IR (ATR)  $\nu_{\max}$  3742, 2940, 2540, 2362, 2162, 1725, 1549, 1022 cm<sup>-1</sup>; NMR spectral data see Supplementary Tables 6 and 7; HRMS (ESI)  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>35</sub>H<sub>67</sub>O<sub>10</sub>N<sub>8</sub> 759.4985, found 759.4968,  $\Delta_{\text{ppm}}$  –2.23.

**Pachycephalamide D (4):** white solid;  $[\alpha]_D^{25}$  –145.68 (c 0.1 w/v%, MeOH); UV (acetonitrile),  $\lambda_{\max}$  204 nm; IR (ATR)  $\nu_{\max}$  3742, 2942, 2540, 2362, 2162, 2024, 1693, 1549, 1021 cm<sup>-1</sup>; NMR spectral data, see Supplementary Tables 6 and 7; HRMS (ESI)  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>36</sub>H<sub>69</sub>O<sub>10</sub>N<sub>8</sub> 773.5131, found 773.5112,  $\Delta_{\text{ppm}}$  –1.88.

**Pachycephalamide E (5):** white solid;  $[\alpha]_D^{25}$  –129.60 (c 0.1 w/v%, MeOH); UV (acetonitrile),  $\lambda_{\max}$  204 nm; IR (ATR)  $\nu_{\max}$  3743, 3621, 3311, 2702, 2322, 1616, 1562 cm<sup>-1</sup>; NMR spectral data, see Supplementary Tables 6 and 7; HRMS (ESI)  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>37</sub>H<sub>71</sub>O<sub>10</sub>N<sub>8</sub> 787.5298, found 787.5278,  $\Delta_{\text{ppm}}$  –2.53.

**Pachycephalamide F (6):** white solid;  $[\alpha]_D^{25}$  –196.30 (c 0.1 w/v%, MeOH); UV (acetonitrile),  $\lambda_{\max}$  204 nm; IR (ATR)  $\nu_{\max}$  3741, 2939, 2540, 2274, 1693, 1549, 1448, 1022 cm<sup>-1</sup>; NMR spectral data, see Supplementary Tables 6 and 7; HRMS (ESI)  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>38</sub>H<sub>73</sub>O<sub>10</sub>N<sub>8</sub> 801.5444, found 801.5424,  $\Delta_{\text{ppm}}$  –2.07.

**Demiguisin (7):** white solid;  $[\alpha]_D^{25}$  –18.7 (c 0.1 w/v%, MeOH); UV (acetonitrile),  $\lambda_{\max}$  204 nm; IR (ATR)  $\nu_{\max}$  3731, 2938, 2539, 2364, 2024, 1690, 1550, 1020 cm<sup>-1</sup>; NMR spectral data, see Supplementary Table 8; HRMS (ESI)  $m/z$  [M + H]<sup>+</sup> calcd. for C<sub>30</sub>H<sub>47</sub>O<sub>10</sub>N<sub>9</sub> 694.3519, found 694.3520,  $\Delta_{\text{ppm}}$  0.14.

### Biosynthetic pathway analysis

For phylogenetic reconstruction of condensation and ketosynthase domains reference sequences were retrieved from natural product domain seeker database (NaPDos2\_v13b)<sup>46</sup>. For further details, see Supplementary Methods.

### RNA sequencing and analysis

*Amycolatopsis* sp. PS\_44\_ISF1 pre-cultures were prepared in ISP2 broth and cultivated for seven days at 30 °C and shaking at 150 rpm. Then, agar plates (each  $n = 3$ ) of a lipopeptide high-production (PDA) and low-production (CSA) medium were inoculated with 500 μl of the pre-culture. After cultivation for five days at 30 °C, mycelium of each plate was harvested, frozen in liquid nitrogen and stored at –80 °C until shipment. For RNA extraction, samples ( $n = 3$  for CSA, and  $n = 3$  for PDA) were pooled to obtain sufficient RNA amount and quality for sequencing. RNA extraction, sample quality control, library construction and RNA sequencing was performed by Novogene Co., Ltd (Cambridge, UK). Each set of reads (CSA, PDA) were mapped to the annotated genome of PS\_44\_ISF1 using Geneious Prime v2020.1.2 (Biomatters Ltd.). Expression levels depicted by normalized transcript counts (TPM = CDS read count \* mean read length \* 10<sup>6</sup>)/(CDS length \* total transcript count)<sup>88</sup> and differential expression confidence (DEC). DEC is the negative base 10 log of the  $p$ -value, adjusted to be negative for genes that are under-expressed in sample 2 (PDA) compared to sample 1 (CSA), or positive for over-expressed genes. The  $p$ -value corresponds to the probability that, for a given gene, a randomly selected transcript would come from that gene and is calculated as number of transcripts mapped to that gene/total number of transcripts from each sample. The probabilities for each sample are then multiplied to form the  $p$ -value. Afterwards, the expression data for selected core biosynthetic genes from BGC regions encoding for NRPS, PKS or NRPS/PKS hybrids was extracted and imported in RStudio v1.4.1106 (R Foundation). The selected values BGCs and selected core biosynthetic genes with their expression values were summarized and the normalized TPM (log10 transformed) and DEC values were visualized in a heatmap created using pheatmap package v1.0.12 with color schemes generated by viridis package v0.6.2. For details, see Supplementary Figs. 76–78.

### Creation of *Amycolatopsis* sp. PS\_44\_ISF1 knockout mutants

For the construction of single-crossover knockout constructs the integrative *E. coli* – *Streptomyces* shuttle vector pKJ55 was digested with *Xba*I and *Sph*I (NEB, USA). A). Internal fragments (–2 kb) of NRPS or PKS candidate genes were amplified using a standard Phusion® High-Fidelity Polymerase protocol (NEB, USA). Standard PCR conditions were as follows: 98 °C/3 min, 35 cycles (98 °C/30 s, T<sub>a</sub> °C/45 s, 72 °C/60 s per kb), 72 °C/10 min. Due to the high GC-content of the genome, PCR reactions were performed in GC buffer supplemented with 5% DMSO (v/v) with primer annealing temperatures (T<sub>a</sub>) 3–5 °C below primer melting temperature T<sub>m</sub>. Primers contain a 20 bp overhang for subsequent cloning of the PCR product into the linearized pKJ55 vector backbone. Both, linearized pKJ55 and each PCR product were purified from 1% agarose gel using Zymoclean Gel DNA Recovery Kit (Zymo Research, USA). For cloning, Gibson Assembly (NEBBuilder® HiFi DNA Assembly Cloning Kit, NEB) was performed according to the manufacturer's protocol with a molar vector-insert ratio of 1:2 and incubated for 1 h at 50 °C. The reaction mixture was transformed in *E. coli* DH5α chemically competent cells and positive transformants were selected on LB agar with apramycin (50 μg/mL). All constructs were purified using innuPREP Plasmid Mini Kit (JenaAnalytik, Germany) and verified by sequencing using Seq\_pKJ55\_IF\_fwd/ rev (Eurofins Genomics, Germany) (Supplementary Tables 14 und 15).

Knockout mutants of PS\_44\_ISF1 were generated via intergeneric conjugational transfer<sup>54,55,89</sup>. Methylation-deficient *E. coli* ET12567/pUZ8002 was transformed with pKJ55-derived suicide vector via electroporation. For each vector construct three conjugations were prepared as follows: 10 mL ET12567 (pUZ8002/pKJ55-ko) in LB medium (supplemented with 50 μg/mL Apr, 25 μg/mL Kan, 25 μg/mL Cam) were grown at 37 °C to an OD<sub>600</sub> of 0.4–0.5, washed twice with ice-cold LB to remove residual antibiotics, and resuspended in fresh 1 mL LB without antibiotics. PS\_44\_ISF1 spore suspensions (50 μL, 3.5 × 10<sup>9</sup>

spores/mL,  $2 \times 10^8$  spores in 25% glycerol) were added to 500  $\mu$ L of 2x YT medium and heat shocked at 50 °C for 10 min. Equal amounts (v/v) of donor and recipient cells were mixed and the bacteria were pelleted by centrifugation. The medium was discarded, and the pellets were resuspended in residual 100  $\mu$ L of medium. The suspension was streaked on MS agar plates, supplemented with 10 mM MgCl<sub>2</sub> and varying concentrations of CaCl<sub>2</sub> (10, 25, 50 mM)<sup>55,89</sup>. After incubation for 16–20 h, the plates were overlaid with 1 mL of water containing 0.5 mg of nalidixic acid (NA) and 1.25 mg of apramycin, and incubation at 30 °C was continued for 5–7 days. Ex-conjugants were streaked on MS (with 50  $\mu$ g/mL Apr and 25  $\mu$ g/mL NA). For verification of knock out mutants, a PCR using primer pairs PIP2 and PIP3 was conducted. PCR products were then confirmed by Sanger sequencing (Eurofins Genomics, Germany), and production of metabolites verified by LC-HRMS/MS analysis (Supplementary Figs. 82–86). Confirmed mutants were stored as 50% glycerol stocks prior to further use.

### Tensioactivity tests

Isolated compounds were dissolved in water (final concentration 10 mM), of which 20  $\mu$ L were placed on a hydrophobic surface (parafilm 'M'). Tween-20 (10 mM) was used as tensioactive positive control. For better visualization, 0.0025% crystal violet was added to the droplet (Supplementary Fig. 88).

### Disc-diffusion based antimicrobial activity assays

SPE fractions and pure compounds (10 mg/mL in DMSO) were tested for their antimicrobial activity against ecologically relevant indicator strains, including feather-degrading bacteria *B. licheniformis* DSM13, *P. monteilii* DSM1388 and *K. rhizophila* DSM11926, and non-degrading feather isolates *B. thuringiensis* DSM104061 and *S. epidermidis* DSM103867, as well as the skin pathogenic yeast *C. albicans* SC5314. For more experimental details, see Supplementary Methods and Supplementary Tables 16–18.

### Protease inhibition assays

Pure compounds were dissolved in DMSO (final concentration 20 mM). The inhibitory effect of the compounds was determined for the following enzymes: rhodesain, cruzain, cathepsin L, cathepsin B, SARS-CoV2 M<sup>pro</sup>, SARS-CoV PL<sup>pro</sup>, DENV2 NS2B-NS3, ZIKV NS2B-NS3, urokinase (uPA), sortase A (SrtA). Fluorimetric assays were performed at a concentration of 200  $\mu$ M and changes in fluorescence recorded over a period of 10 min in intervals of 30 seconds (exception: SrtA 30 min). For more experimental details, see Supplementary Methods and Supplementary Table 19<sup>70–73,90–92</sup>.

### Reporting summary

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

### Data availability

Supplementary Information contains all details to experimental and analytical data. Analytical data (1D, 2D NMR) is available on Zenodo [<https://doi.org/10.5281/zenodo.13125628>]. The WGS data used in this study are available in the NCBI database under accession code [JANUXO000000000.1](https://doi.org/10.5281/zenodo.13125628). The SRA data used in this study are available in the NCBI database under accession codes [SRR21206835](https://doi.org/10.5281/zenodo.13125628), [SRR21206836](https://doi.org/10.5281/zenodo.13125628) and [SRR21206837](https://doi.org/10.5281/zenodo.13125628). RNA sequencing data used in this study are available in the NCBI database under accession codes [SRR21206833](https://doi.org/10.5281/zenodo.13125628) and [SRR21206834](https://doi.org/10.5281/zenodo.13125628). The amplicon sequencing data generated in this study have been deposited in the NCBI database under accession codes [SAMN30849808](https://doi.org/10.5281/zenodo.13125628), [SAMN30849809](https://doi.org/10.5281/zenodo.13125628), [SAMN30849810](https://doi.org/10.5281/zenodo.13125628), [SAMN30849811](https://doi.org/10.5281/zenodo.13125628), [SAMN30849812](https://doi.org/10.5281/zenodo.13125628), [SAMN30849813](https://doi.org/10.5281/zenodo.13125628), [SAMN30849814](https://doi.org/10.5281/zenodo.13125628), [SAMN30849815](https://doi.org/10.5281/zenodo.13125628), [SAMN30849816](https://doi.org/10.5281/zenodo.13125628), [SAMN30849817](https://doi.org/10.5281/zenodo.13125628), [SAMN30849818](https://doi.org/10.5281/zenodo.13125628), [SAMN30849819](https://doi.org/10.5281/zenodo.13125628), [SAMN30849820](https://doi.org/10.5281/zenodo.13125628), [SAMN30849821](https://doi.org/10.5281/zenodo.13125628), [SAMN30849822](https://doi.org/10.5281/zenodo.13125628), [SAMN30849823](https://doi.org/10.5281/zenodo.13125628), [SAMN30849824](https://doi.org/10.5281/zenodo.13125628), [SAMN30849825](https://doi.org/10.5281/zenodo.13125628), [SAMN30849826](https://doi.org/10.5281/zenodo.13125628),

[SAMN30849827](https://doi.org/10.5281/zenodo.13125628) and [SAMN30849828](https://doi.org/10.5281/zenodo.13125628). We have uploaded the MS-Data to the MAStIVE Server under the accession code MSV000093302 [<https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=7c1ba2e849ed42f88592c69f6a599d83>]. Source data are available in the Source Data file. Source data are provided with this paper.

### References

1. Contreras, A. et al. Epidemiological role of birds in the transmission and maintenance of zoonoses. *Rev. Sci. Tech.* **35**, 845–862 (2016).
2. Bodawatta, K. H. et al. Multiple mutations in the Nav1.4 sodium channel of New Guinean toxic birds provide autoresistance to deadly batrachotoxin. *Mol. Ecol.* **33**, e16878 (2024).
3. Javůrková, V. G. et al. Unveiled feather microcosm: feather microbiota of passerine birds is closely associated with host species identity and bacteriocin-producing bacteria. *ISME J.* **13**, 2363–2376 (2019).
4. Bodawatta, K. H. et al. Great Tit (*Parus major*) Uropygial gland microbiomes and their potential defensive roles. *Front. Microbiol.* **11**, 1735 (2020).
5. Bodawatta, K. H. et al. Specific gut bacterial responses to natural diets of tropical birds. *Sci. Rep.* **12**, 713 (2022).
6. Shawkey, M. D., Pillai, S. R. & Hill, G. E. Chemical warfare? Effects of uropygial oil on feather-degrading bacteria. *J. Avian Biol.* **34**, 345–349 (2003).
7. Reneerkens, J., Versteegh, M. A., Schneider, A. M., Piersma, T. & Burt, E. H. Seasonally changing preen-wax composition: Red Knots' (*Calidris canutus*) flexible defense against feather-degrading bacteria? *Auk* **125**, 285–290 (2008).
8. Ruiz-Rodriguez, M. et al. Symbiotic bacteria living in the hoopoe's uropygial gland prevent feather degradation. *J. Exp. Biol.* **212**, 3621–3626 (2009).
9. Jacob, S. et al. Uropygial gland size and composition varies according to experimentally modified microbiome in Great tits. *BMC Evol. Biol.* **14**, 134 (2014).
10. Moreno-Rueda, G. Preen oil and bird fitness: a critical review of the evidence. *Biol. Rev. Camb. Philos. Soc.* **92**, 2131–2143 (2017).
11. Rodríguez-Ruano, S. M. et al. Seasonal and sexual differences in the microbiota of the hoopoe uropygial secretion. *Genes* **9**, 134 (2018).
12. Grieves, L. A. et al. Olfactory camouflage and communication in birds. *Biol. Rev. Camb. Philos. Soc.* **97**, 1193–1209 (2022).
13. Grieves, L. A., Gloor, G. B., Kelly, T. R., Bernards, M. A. & MacDougall-Shackleton, E. A. Preen gland microbiota of songbirds differ across populations but not sexes. *J. Anim. Ecol.* **90**, 2202–2212 (2021).
14. Braun, M. S., Sporer, F., Zimmermann, S. & Wink, M. Birds, feather-degrading bacteria and preen glands: the antimicrobial activity of preen gland secretions from turkeys (*Meleagris gallopavo*) is amplified by keratinase. *FEMS Microbiol. Ecol.* **94**, fiy117 (2018).
15. Martín-Platero, A. M. et al. Characterization of antimicrobial substances produced by *Enterococcus faecalis* MRR10-3, isolated from the uropygial gland of the hoopoe (*Upupa epops*). *Appl Environ. Microbiol.* **72**, 4245–4249 (2006).
16. Um, S., Seibel, E., Schalk, F., Balluff, S. & Beemelmanns, C. Targeted Isolation of Saalfelduracin B-D from *Amycolatopsis saalfeldensis* Using LC-MS/MS-Based Molecular Networking. *J. Nat. Prod.* **84**, 1002–1011 (2021).
17. Carlsohn, M. R. et al. *Amycolatopsis saalfeldensis* sp. nov., a novel actinomycete isolated from a medieval alum slate mine. *Int. J. Syst. Evol. Microbiol.* **57**, 1640–1646 (2007).
18. Meier-Kolthoff, J. P. & Goker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **10**, 2182 (2019).
19. Wink, J. et al. *Amycolatopsis decaplanina* sp. nov., a novel member of the genus with unusual morphological diversity. *Int. J. Syst. Evol. Microbiol.* **54**, 235–239 (2004).

20. Adamek, M. et al. Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC Genomics* **19**, 426 (2018).
21. Beemelmans, C. et al. Macrotermycins A-D, glycosylated macro-lactams from a termite-associated *Amycolatopsis* sp. M39. *Org. Lett.* **19**, 1000–1003 (2017).
22. Lee, S. D. *Amycolatopsis jejuensis* sp. nov. and *Amycolatopsis halotolerans* sp. nov., novel actinomycetes isolated from a natural cave. *Int. J. Syst. Evol. Microbiol.* **56**, 549–553 (2006).
23. Groth, I. et al. *Amycolatopsis nigrescens* sp. nov., an actinomycete isolated from a Roman catacomb. *Int. J. Syst. Evol. Microbiol.* **57**, 513–519 (2007).
24. Yassin, A. F., Haggel, B., Budzikiewicz, H. & Schaal, K. P. Fatty Acid and Polar Lipid Composition of the Genus *Amycolatopsis*: Application of Fast Atom Bombardment-Mass Spectrometry to Structure Analysis of Underivatized Phospholipids. *Int. J. Syst. Bacteriol.* **43**, 414–420 (1993).
25. Riahi, H. S., Heidarieh, P. & Fatahi-Bafghi, M. Genus *Pseudonocardia*: What we know about its biological properties, abilities and current application in biotechnology. *J. Appl. Microbiol.* **132**, 890–906 (2022).
26. Li, H., Young, S. E., Poulsen, M. & Currie, C. R. Symbiont-mediated digestion of plant biomass in fungus-farming insects. *Annu. Rev. Entomol.* **66**, 297–316 (2022).
27. Heine, D. et al. Chemical warfare between leafcutter ant symbionts and a co-evolved pathogen. *Nat. Commun.* **9**, 2208 (2018).
28. Wang, M. et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **34**, 828–837 (2016).
29. Seibel, E. et al. Genome mining for macrolactam-encoding gene clusters allowed for the network-guided isolation of  $\beta$ -amino acid-containing cyclic derivatives and heterologous production of cirocin A. *Commun. Chem.* **6**, 257 (2023).
30. Derewacz, D. K., Covington, B. C., McLean, J. A. & Bachmann, B. O. Mapping microbial response metabolomes for induced natural product discovery. *ACS Chem. Biol.* **10**, 1998–2006 (2015).
31. August, P. R. et al. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* **5**, 69–79 (1998).
32. Xiao, Y. S. et al. Rifamorpholines A-E, potential antibiotics from locust-associated actinobacteria *Amycolatopsis* sp. Hca4. *Org. Biomol. Chem.* **15**, 3909–3916 (2017).
33. Watrous, J. D. et al. Microbial metabolic exchange in 3D. *ISME J.* **7**, 770–780 (2013).
34. Morgan, K. D., Andersen, R. J. & Ryan, K. S. Piperazic acid-containing natural products: structures and biosynthesis. *Nat. Prod. Rep.* **36**, 1628–1653 (2019).
35. Stephan, P. et al. Directed Evolution of Piperazic Acid Incorporation by a Nonribosomal Peptide Synthetase. *Angew. Chem. Int. Ed. Engl.* **62**, e202304843 (2023).
36. Son, S. et al. Ulleungamides A and B, modified  $\alpha$ , $\beta$ -dehydropepelic acid containing cyclic depsipeptides from *Streptomyces* sp. KCB13F003. *Org. Lett.* **17**, 4046–4049 (2015).
37. Siodlak, D.  $\alpha$ , $\beta$ -Dehydroamino acids in naturally occurring peptides. *Amino Acids* **47**, 1–17 (2015).
38. Berkowitz, D. B., Jahng, W. J. & Pedersen, M. L.  $\alpha$ -Vinyllysine and  $\alpha$ -vinylarginine are time-dependent inhibitors of their cognate decarboxylases. *Bioorg. Med. Chem. Lett.* **6**, 2151–2156 (1996).
39. Li, P. & Wu, G. Roles of dietary glycine, proline, and hydroxyproline in collagen synthesis and animal growth. *Amino Acids* **50**, 29–38 (2018).
40. Heinilä, L. M. P. et al. The structure and biosynthesis of heinamides A1-A3 and B1-B5, antifungal members of the laxaphycin lipopeptide family. *Org. Biomol. Chem.* **19**, 5577–5588 (2021).
41. Houwaart, S., Youssar, L. & Huttel, W. Pneumocandin biosynthesis: involvement of a trans-selective proline hydroxylase. *ChemBioChem* **15**, 2365–2369 (2014).
42. Lewerenz, J., Klein, M. & Methner, A. Cooperative action of glutamate transporters and cystine/glutamate antiporter system Xc- protects from oxidative glutamate toxicity. *J. Neurochem.* **98**, 916–925 (2006).
43. 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).
44. Rottig, M. et al. NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Res.* **39**, W362–W367 (2011).
45. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).
46. Klau, L. J. et al. The Natural Product Domain Seeker version 2 (NaPDoS2) webtool relates ketosynthase phylogeny to biosynthetic function. *J. Biol. Chem.* **298**, 102480 (2022).
47. Du, D., Katsuyama, Y., Shin-Ya, K. & Ohnishi, Y. Reconstitution of a type II polyketide synthase that catalyzes polyene formation. *Angew. Chem. Int. Ed. Engl.* **57**, 1954–1957 (2018).
48. Pohle, S., Appelt, C., Roux, M., Fiedler, H. P. & Sussmuth, R. D. Biosynthetic gene cluster of the non-ribosomally synthesized cyclodepsipeptide skyllamycin: deciphering unprecedented ways of unusual hydroxylation reactions. *J. Am. Chem. Soc.* **133**, 6194–6205 (2011).
49. Grammbitter, G. L. C. et al. An uncommon type II PKS catalyzes biosynthesis of aryl polyene pigments. *J. Am. Chem. Soc.* **141**, 16615–16623 (2019).
50. Masschelein, J. et al. A combination of polyunsaturated fatty acid, nonribosomal peptide and polyketide biosynthetic machinery is used to assemble the zeamine antibiotics. *Chem. Sci.* **6**, 923–929 (2015).
51. Thomy, D. et al. The ADEP Biosynthetic Gene Cluster in *Streptomyces hawaiiensis* NRRL 15010 Reveals an Accessory clpP Gene as a Novel Antibiotic Resistance Factor. *Appl. Environ. Microbiol.* **85**, e01292–19 (2019).
52. Dose, B. et al. Unexpected bacterial origin of the antibiotic icosalide: Two-tailed depsipeptide assembly in multifarious *Burkholderia* symbionts. *ACS Chem. Biol.* **13**, 2414–2420 (2018).
53. Zhong, L. et al. Engineering and elucidation of the lipoinitiation process in nonribosomal peptide biosynthesis. *Nat. Commun.* **12**, 296 (2021).
54. Ishida, K., Fritzsche, K. & Hertweck, C. Geminal tandem C-methylation in the discoid resistomycin pathway. *J. Am. Chem. Soc.* **129**, 12648–12649 (2007).
55. Kieser, T., Foundation, J. I. *Practical Streptomyces Genetics* (John Innes Foundation, 2000).
56. Kaljunen, H. et al. Structural elucidation of the bispecificity of A domains as a basis for activating non-natural amino acids. *Angew. Chem. Int. Ed. Engl.* **54**, 8833–8836 (2015).
57. McGlinchey, R. P., Nett, M. & Moore, B. S. Unraveling the biosynthesis of the sporolide cyclohexenone building block. *J. Am. Chem. Soc.* **130**, 2406–2407 (2008).
58. Van Lanen, S. G., Lin, S., Dorrestein, P. C., Kelleher, N. L. & Shen, B. Substrate specificity of the adenylation enzyme SgcC1 involved in the biosynthesis of the enediyne antitumor antibiotic C-1027. *J. Biol. Chem.* **281**, 29633–29640 (2006).
59. Bernhardt, M., Berman, S., Zechel, D. & Bechthold, A. Role of two exceptional trans adenylation domains and Mbth-like proteins in the biosynthesis of the nonribosomal peptide WS9324A from *Streptomyces calvus* ATCC 13382. *ChemBioChem* **21**, 2659–2666 (2020).



60. Caffrey, P. Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. *ChemBioChem* **4**, 654–657 (2003).
61. Bruckner, R. C., Zhao, G., Venci, D. & Jorns, M. S. Nikkomycin biosynthesis: formation of a 4-electron oxidation product during turnover of NikD with its physiological substrate. *Biochemistry* **43**, 9160–9167 (2004).
62. Zhu, Y. et al. Insights into caerulomycin A biosynthesis: a two-component monooxygenase CrmH-catalyzed oxime formation. *J. Am. Chem. Soc.* **135**, 18750–18753 (2013).
63. Lukat, P. et al. Biosynthesis of methyl-proline containing griselimycins, natural products with anti-tuberculosis activity. *Chem. Sci.* **8**, 7521–7527 (2017).
64. Pickens, L. B. et al. Biochemical analysis of the biosynthetic pathway of an anticancer tetracycline SF2575. *J. Am. Chem. Soc.* **131**, 17677–17689 (2009).
65. Zhang, S. et al. Structural diversity, biosynthesis, and biological functions of lipopeptides from *Streptomyces*. *Nat. Prod. Rep.* **40**, 557–594 (2023).
66. Chen, W.-C., Juang, R.-S. & Wei, Y.-H. Applications of a lipopeptide biosurfactant, surfactin, produced by microorganisms. *Biochemical Eng. J.* **103**, 158–169 (2015).
67. Barthels, F. et al. 2-Sulfonylpyrimidines as privileged warheads for the development of *S. aureus* sortase A inhibitors. *Front. Mol. Biosci.* **8**, 804970 (2021).
68. Rudzińska, M. et al. The role of cysteine cathepsins in cancer progression and drug resistance. *Int. J. Mol. Sci.* **20**, 3602 (2019).
69. Masucci, M. T., Minopoli, M., Di Carluccio, G., Motti, M. L. & Carriero, M. V. Therapeutic strategies targeting urokinase and its receptor in cancer. *Cancers* **14**, 498 (2022).
70. Jung, S. et al. Fluorovinylsulfones and -sulfonates as potent covalent reversible inhibitors of the trypanosomal cysteine protease rhodesain: Structure-activity Relationship, inhibition mechanism, metabolism, and *in vivo* studies. *J. Med. Chem.* **64**, 12322–12358 (2021).
71. Arafet, K. et al. Impact of the recognition part of dipeptidyl nitroalkene compounds on the inhibition mechanism of cysteine proteases cruzain and cathepsin L. *ACS Catal.* **13**, 6289–6300 (2023).
72. Welker, A. et al. Structure-activity relationships of enzamides and isoindolines designed as SARS-CoV protease inhibitors effective against SARS-CoV-2. *ChemMedChem* **16**, 340–354 (2021).
73. Maus, H. et al. SAR of novel benzothiazoles targeting an allosteric pocket of DENV and ZIKV NS2B/NS3 proteases. *Bioorg. Med. Chem.* **47**, 116392 (2021).
74. Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).
75. Callahan, B. J. et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
76. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
77. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013).
78. Miller, C. S. et al. Short-read assembly of full-length 16S amplicons reveals bacterial diversity in subsurface sediments. *PLOS ONE* **8**, e56018 (2013).
79. Wingett, S. W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* **7**, 1338 (2018).
80. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
81. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011).
82. Vurture, G. W. et al. GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).
83. De Coster, W., D’Hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669 (2018).
84. Zimin, A. V. et al. The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).
85. Zimin, A. V. & Salzberg, S. L. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput. Biol.* **16**, e1007981 (2020).
86. Chambers, M. C. et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
87. Hoffmann, T. & Dorrestein, P. C. Homogeneous matrix deposition on dried agar for MALDI imaging mass spectrometry of microbial cultures. *J. Am. Soc. Mass Spectrom.* **26**, 1959–1962 (2015).
88. Wagner, G. P., Kin, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285 (2012).
89. Li, C., Zhou, L., Wang, Y., Zhao, G. & Ding, X. Conjugation of  $\phi$ BT1-derived integrative plasmid pDZL802 in *Amycolatopsis mediterranei* U32. *Bioengineered* **8**, 549–554 (2017).
90. Millies, B. et al. Proline-based allosteric inhibitors of Zika and Dengue virus NS2B/NS3 proteases. *J. Med. Chem.* **62**, 11359–11382 (2019).
91. Barthels, F. et al. Asymmetric disulfanylbenzamides as irreversible and selective inhibitors of *Staphylococcus aureus* sortase A. *ChemMedChem* **15**, 839–850 (2020).
92. Klein, P. et al. Naphthoquinones as covalent reversible inhibitors of cysteine proteases—Studies on inhibition mechanism and kinetics. *Molecules* **25**, 2064 (2020).

## Acknowledgements

This study was funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) under Project-ID 239748522 – CRC 1127 (project A6). We are grateful for the financial support from the Agence Nationale de la Recherche (ANR, ANR-17-CE07-0051-01), the German Research Foundation (DFG, BE 4799/3-1), the Carlsberg Foundation (Distinguished Associate Professor Fellowship, CF17-0248) and the Korean National Research Foundation (2021R111A1A0104960613 and 2018R1A6A1A03023718). We also would like to acknowledge support from the New Guinea Binatang Research Centre in Madang, the PNG National Museum and Art Gallery in Port Moresby, the PNG Immigration & Citizen Service Authority and the Department of Environment and Conservation, Boroko, Papua New Guinea. We also would like to thank Daria Ripin (HKI) for technical assistance, Sven Balluff (HIPS) for conducting initial co-cultivation experiments, Christiane Weigel (HKI) for conducting antimicrobial assays, Heike Heinecke (HKI) for NMR measurements, Philipp Stephan (HKI) for helpful discussions and providing a Piz standard for Marfey’s analysis, and Marie Dayras (HIPS) for the critical review of NMR and MS/MS data.

## Author contributions

E.S., S.U. and K.H.B. conceived the study. E.S., S.U., K.H.B., T.D., A.K., J.F., R.M., G.M., B.I. and H.M. performed the research and analyzed the data. E.S., S.U., K.H.B., T.D. and T.S. analyzed analytical data. K.A.J., G.M., M.P., B.I., T.S. and C.B. supervised the research and acquired funding. E.S., K.H.B., S.U. and C.B. wrote the paper with input from all co-authors. All authors contributed to the review and editing of the paper.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-52316-3>.

**Correspondence** and requests for materials should be addressed to Christine Beemelmans.

**Peer review information** *Nature Communications* thanks Pieter Dorrestein, Michael Wink, and the other, anonymous, reviewer for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

---

<sup>1</sup>Anti-infectives from Microbiota, Helmholtz-Institut für Pharmazeutische Forschung Saarland (HIPS), Campus E8.1, 66123 Saarbrücken, Germany. <sup>2</sup>Chemical Biology of Microbe-Host Interactions, Leibniz institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute (HKI), Beutenbergstraße 11a, 07745 Jena, Germany. <sup>3</sup>Yonsei Institute of Pharmaceutical Sciences, College of Pharmacy Yonsei University, Songdogwahak-ro 85, Incheon 21983, Republic of Korea. <sup>4</sup>Natural History Museum of Denmark, Research and Collections University of Copenhagen, 2100 Copenhagen East, Denmark. <sup>5</sup>Department of Biomolecular Chemistry, Leibniz institute for Natural Product Research and Infection Biology – Hans-Knöll-Institute (HKI), Beutenbergstraße 11a, 07745 Jena, Germany. <sup>6</sup>Section for Ecology and Evolution, Department of Biology, University of Copenhagen, 2100 Copenhagen East, Denmark. <sup>7</sup>The New Guinea Binatang Research Centre, Madang, Papua New Guinea. <sup>8</sup>Papua New Guinea National Museum and Art Gallery, Port Moresby, Papua New Guinea. <sup>9</sup>Institute for Pharmaceutical and Biomedical Sciences (IPBW), Johannes Gutenberg University Mainz, Staudinger Weg 5, 55128 Mainz, Germany. <sup>10</sup>Swedish Museum of Natural History, Department of Bioinformatics and Genetics, P.O. Box 50007, SE-10405 Stockholm, Sweden. <sup>11</sup>Saarland University, Campus, 66123 Saarbrücken, Germany. <sup>12</sup>These authors contributed equally: Elena Seibel, Soohyun Um. ✉ e-mail: [Christine.beemelmans@helmholtz-hips.de](mailto:Christine.beemelmans@helmholtz-hips.de)