Ecology and genomics of Actinobacteria and their specialised metabolism

Bergeijk, D.A. van

Bioactive *Pseudonocardia* within the microbiome of zebrafish

Doris A. van Bergeijk, Robert Whelan, Hannah Augustijn, Victor Carrión, Bjørn Koch, Gilles P. van Wezel, Herman P. Spaink

Abstract

Filamentous Actinobacteria are well known for their production of bioactive specialised metabolites. Many gifted producers have been isolated from the microbiomes of eukaryotic organisms, such as plants, marine organisms, and insects. Actinobacteria play an important role in the protection of these organisms against pathogens. Yet, the bioactive and functional potential of filamentous Actinobacteria within the microbiome of their eukaryotic hosts remains poorly characterised. Zebrafish larvae are excellent systems to characterise the impact of host-microbe interactions in health and disease. In this study, we investigated antibiotic-producing Actinobacteria associated with this model system. We analysed the gut microbiome of adult zebrafish and isolated Actinobacteria from zebrafish larvae, belonging to the less common actinobacterial genera *Microbacterium, Dermacoccus, Kocuria,* and *Pseudonocardia.* Genome sequencing of *Pseudonocardia* sp. ZF1 revealed much uncharacterised biosynthetic potential and high similarity to ant-associated *Pseudonocardia*. *Pseudonocardia* sp. ZF1 had antimicrobial activity against both Gram-positive and Gram-negative bacteria. Further research is required to identify the bioactive metabolites produced by the *Pseudonocardia* isolate and characterise is functional role *in vivo*.

Introduction

The bacterial world is home to a great number of specialised chemists, many of which belong to the phylum Actinobacteria [1,27]. Their specialised metabolites display an extraordinary chemical diversity, allowing Actinobacteria to thrive in a wide range of ecosystems [2]. Within these environments, siderophores facilitate the uptake of iron, compatible solutes protect against osmotic stress, and antibacterial metabolites function as signalling molecules or as chemical weaponry to outcompete other organisms [42]. Importantly, the beneficial effects provided by these metabolites are not limited to the producer, but can also provide advantage to the many multicellular organisms that host Actinobacteria [45-47]. For example, antibiotic-producing Actinobacteria are consistently associated with disease-suppressive soils, in which plants suffer less from soil-borne pathogens [47]. Another well characterised example is the fungus-farming ant system where *Pseudonocardia* reside in the cuticle of ants and protect the fungal cultivar against different pathogenic fungi through the production of bioactive metabolites [288-290].

The interactions that take place between resident bacteria, host and invading pathogens can shape microbial enzymatic potential and select for efficacious and relevant bioactivity [291]. It is therefore not surprising that microbiomes are a promising source of bioactive molecules. Sequencing of the human microbiome revealed the presence of many biosynthetic gene clusters (BGCs) across different bacterial phyla, including Actinobacteria [181,193]. Additionally, several novel bioactive molecules have been isolated from host-associated bacteria such as the novel antifungal agent cyphomycin produced by insect-associated *Streptomyces* [43], the anthracycline antibiotic keyicin from *Micromonospora* isolated from the sea squirt microbiome [292], and the antibiotic lugdunin from the human nasal commensal *Staphylococcus lugdunensis* [99]. Notably, human nasal carriage of *S. lugdunensis* was associated with reduced nasal carriage of *Staphylococcus aureus*, illustrating a potential role for antibiotic-producing commensal bacteria in the protection against (opportunistic) pathogens in humans [99].

Within the phylum Actinobacteria, most interest has been directed towards the health-promoting effects of the probiotic *Bifidobacteria* within the animal and human microbiome [293-295]. Yet, much remains to be discovered about the functional and bioactive potential of antibiotic-producing Actinobacteria within the animal microbiome. Moreover, a growing body of evidence shows that host-associated signalling molecules play a role in the activation of specialised metabolite production [296-300]. For example, the human opioid dynorphin stimulates the production of the virulence factor pyocyanin in *Pseudomonas aeruginosa* [242], and plant- and animal-associated stress hormones stimulate antibiotic and siderophore production in *Streptomyces* [44,301] (Chapter 4). We hypothesise that upon host stress, such as infection,

host stress molecules might serve as a 'cry for help' and activate production of specialised metabolites by Actinobacteria that can aid in the protection against invading pathogens [2]. To test this hypothesis, models are needed to study whether such interactions occur *in vivo* and whether Actinobacteria can provide protection against invading pathogens.

Over the past years, zebrafish larvae have proven to be a relevant model for microbiome research *in vivo* [302-305]. Potentially, they can serve as a model to explore the functional role of Actinobacteria within the microbiome of animals and to study the involvement of host-microbe interactions in the activation of specialised metabolite production. For this purpose, we analysed the presence of Actinobacteria within the gut microbiome of adult zebrafish and isolated Actinobacteria from zebrafish larvae, including a bioactive *Pseudonocardia* sp. with much uncharacterised biosynthetic potential. This study provides a first step towards exploring the functional role of antibiotic-producing Actinobacteria within the animal microbiome. Further research is required to identify the bioactive metabolites produced by the *Pseudonocardia* isolate and test its protective effect *in vivo*.

Results

Actinobacteria in the zebrafish microbiome

We aimed to isolate zebrafish-associated Actinobacteria to provide a first step towards using zebrafish as an *in vivo* model to explore the bioactive and functional potential of Actinobacteria within the animal microbiome. To get an overview of the genera that are present in the gut microbiome of zebrafish, 16S rRNA amplicon sequencing was performed. For this, the intestine of three adult zebrafish were removed and homogenised, and total DNA extracted. The DNA samples were analysed by 16S rRNA amplicon sequencing of the V3-V4 hypervariable region using the MiSeq Illumina sequencing platform. Similar to previous reports, the community was numerically dominated by Proteobacteria (Fig. 1A). Actinobacteria comprised on average 0.97% of the filtered reads (Fig. 1B). Sequence variants were assigned to the genera *Collinsella, Corynebacterium, Dermacoccus, Slackia, Bifidobacterium, Rothia*, *Nesterenkonia, Kocuria, Brevibacterium, Actinomyces*, *Marmicola, Cutibacterium, Lawsonella, Micrococcus, Turicella*, *Pseudonocardia,* and *Rhodococcus*.

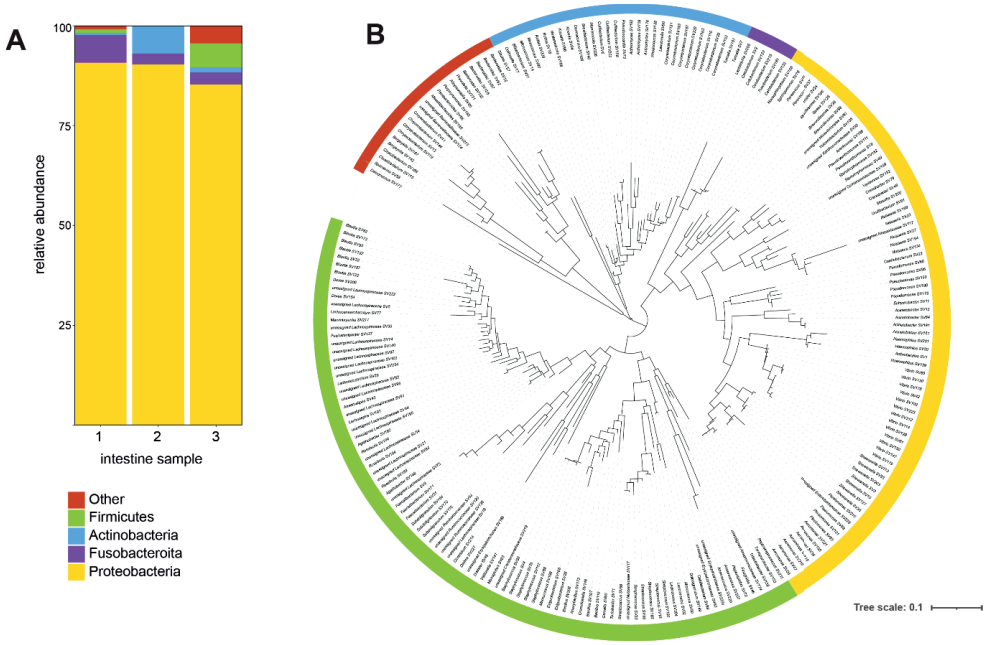

Figure 1. 16s rRNA taxonomic profiling of the adult zebrafish intestinal microbiome. (A) Taxonomic composition bar plot illustrating the phylum-level composition of the adult intestinal microbiota. (B) Midpoint-rooted maximum likelihood phylogenetic tree of adult zebrafish gut microbiota amplicon sequence variants.

For the isolation of Actinobacteria, zebrafish larvae at 5 days post fertilization (dpf) were washed in sterile PBS and tissue was homogenised. The homogenate was then serially diluted and plated onto media selective for Actinobacteria. Bacteria were isolated based on several morphological characteristics (non-motile, filamentous growth, and/or dry colony morphology) and grown on different media for phenotypic discrimination. Based on this, 12 strains were selected. Sequencing of the 16S rRNA gene revealed that six isolates belonged to the phylum Actinobacteria spanning five genera, namely *Dermacoccus, Kocuria, Microbacterium*, *Micrococcus*, and *Pseudonocardia* (Table 1). The other isolates belonged to the genera *Sphingomonas, Mesorhizobium*, and *Staphylococcus*. All of these isolated genera were detected in the amplicon sequencing data of the adult zebrafish gut, except *Microbacterium.*

Table 1. Isolated bacteria and their closest neighbour using (partial) 16S rRNA gene sequencing

Isolate	Top-hit EzBioCloud	similarity	Completeness
ZF0	*Staphylococcus warneri* ATCC 27836	100%	47.5
ZF1	*Pseudonocardia carboxydivorans* Y8	99.93%	99.9
ZF4	*Sphingomonas paucimobilis* NBRC 13935	100%	49.6
ZF9	*Micrococcus luteus* NCTC2665	100%	48.4
ZF22	*Sphingomonas sanguinis* NBRC 13937	99.86%	49.7
ZF23	*Sphingomonas yabuuchiae* GTC 868	100%	49.6
ZF27	*Microbacterium hominis* NBRC 15708	99.14%	48.4
ZF30	*Mesorhizobium terrae* NIBRBAC000500504	100%	49.8
ZF31	*Kocuria arsenatis* CM1E1	100%	48.2
ZF33	*Microbacterium hominis* NBRC 15708	99.14%	48.4
ZF37	*Sphingomonas sanguinis* NBRC 13937	99.86%	49.7
ZF50	*Dermacoccus nishinomiyaensis* DSM 20448	99.57%	48.5

Genomic analysis of *Pseudonocardia* sp. ZF1

The richness of bacterial genomes in terms of BGC diversity, has been correlated with key organismal features, including multicellularity [1]. Of the isolates, only *Pseudonocardia* sp. ZF1 exhibited filamentous growth and this strain was therefore selected for further characterisation. *Pseudonocardia* sp. ZF1 produced a cream-coloured aerial spore mass and orange substrate mycelia (Fig. 2A). No diffusible pigments were produced. Also in liquid, no diffusible pigments were produced and growth was characterised by the formation of a branched mycelium and pellets (Fig. 2B).

To gain insights into the relatedness to known *Pseudonocardia* isolates, we obtained the complete genome sequence using a combination of Nanopore and Illumina sequencing. Combined assembly resulted in a draft genome with a G+C content of 73.9% consisting of 3 contigs of 5.66 Mb, 494 kb, and 25.6 kb in length, all of which were predicted to be circular (Table S1, Fig. 3). The 5.66 Mb contig represents the genome sequence, while the smaller contigs (ZFp1 and ZFp2) likely represent extrachromosomal plasmids as they contained many genes for mobile genetic elements, such as transposases and integrases (Fig. 3).

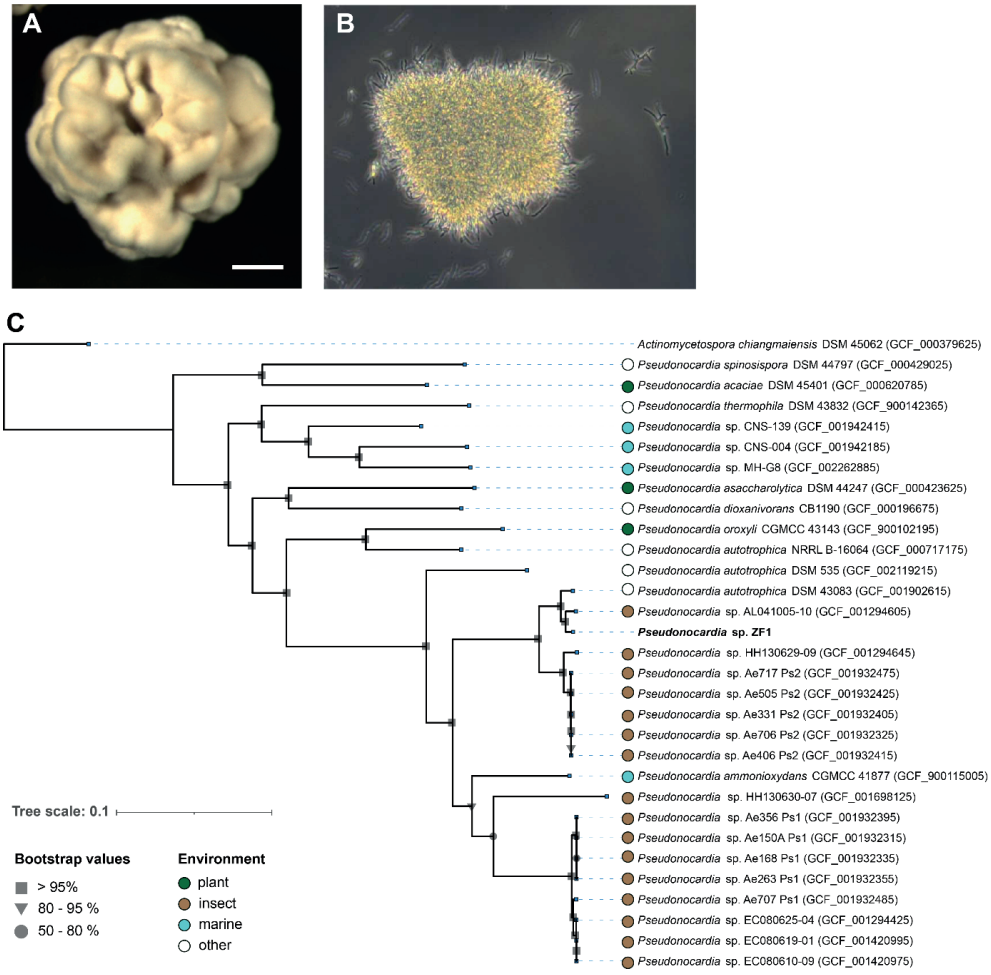

Figure 2. Phenotypic and taxonomic characterization of *Pseudonocardia* sp. ZF1. (A) *Pseudonocardia* sp. ZF1 produces cream-coloured colonies with an orange substrate mycelium when grown on ISP2 agar. Scale bar: 1 mm **(B)** In liquid ISP2, growth is characterised by the formation of a branched mycelium and pellets. **(C)** Genome-scale phylogenetic tree based on Multi-Locus Sequence Analysis of *Pseudonocardia* sp. ZF1 and different members of the genus *Pseudonocardia*. The tree was constructed using autoMLST [306]. *Actinomycetospora chiangmaiensis* DSM 45062 was used as an outgroup. Symbols indicate the bootstrap values in percentage of a total of 1000 bootstrap. Colours indicate from which environment the strains have been isolated.

A phylogenetic tree was constructed based on whole-genome sequences using autoMLST [306] (Fig. 2C). *Pseudonocardia* sp. ZF1 formed a monophyletic group with the fungus-growing ant-associated *Pseudonocardia* sp. AL041005-10 [48] and *Pseudonocardia autothrophica* DSM 43083 [307]. To assess the genomic distance between *Pseudonocardia* sp. ZF1 and its closest neighbour *Pseudonocardia* sp. AL041005-10, the average nucleotide identity (ANI) was calculated in an all-to-all genome comparison, revealing a similarity score of 97.21%. The species delineation threshold typically lies around 95% gene identity [232], suggesting that *Pseudonocardia* sp. ZF1 and *Pseudonocardia* sp. AL041005-10 belong to the same species. Comparative genome analysis of *Pseudonocardia* sp. ZF1 and *Pseudonocardia* sp. AL041005-10 indeed showed high similarity between the strains. No plasmid sequences similar to plasmids ZFp1 and ZFp2 were found in *Pseudonocardia* sp. AL041005-10 (Fig. S1).

Biosynthetic potential of zebrafish isolate *Pseudonocardia* sp. ZF1

Actinobacteria isolated from the microbiomes of different eukaryotic hosts have promising bioactive potential. To study the biosynthetic potential of *Pseudonocardia* sp. ZF1, the genome sequence was analysed using antiSMASH [217]. 15 BGCs were identified, the majority of which was located on the chromosome. Four BGCs were located on plasmid ZFp1, and no BGCs were identified on ZFp2 (Table S2). Only one BGC matches a known BGC with 100% similarity, specifying the compatible solute ectoine, which is commonly found in bacteria [223]. Notably, two-third of the BGCs had less than 25% homology to BGCs in the MIBiG database (Table S2).

We then wondered whether these BGCs were unique to *Pseudonocardia* sp. ZF1 or whether they were shared among *Pseudonocardia* spp. For this, the BGCs were analysed using cblaster [309], a tool for finding clusters of co-located sequences using BLAST searches. The majority of the BGCs were shared with other *Pseudonocardia* species, such as *Pseudonocardia alni* DSM44104, *Pseudonocardia* sp. ICBG1034, and *Pseudonocardia autotrophica* DSM 43083 (Table S2). However, for the BGCs located on plasmid ZFp1, almost no homology was found, also when the search was limited to only a small subset of genes of the predicted BGCs. In the comparative genome analysis of *Pseudonocardia* sp. ZF1 and its closest neighbour *Pseudonocardia* sp. AL041005-10, the BGCs on ZFp1 were not shared (Fig. S1). Notably, the best hit for BGC2.2 was found on plasmid pPA12743CP of *Pseudonocardia autotrophica* NBRC 12742, further validating ZFp1 most likely to be a plasmid.

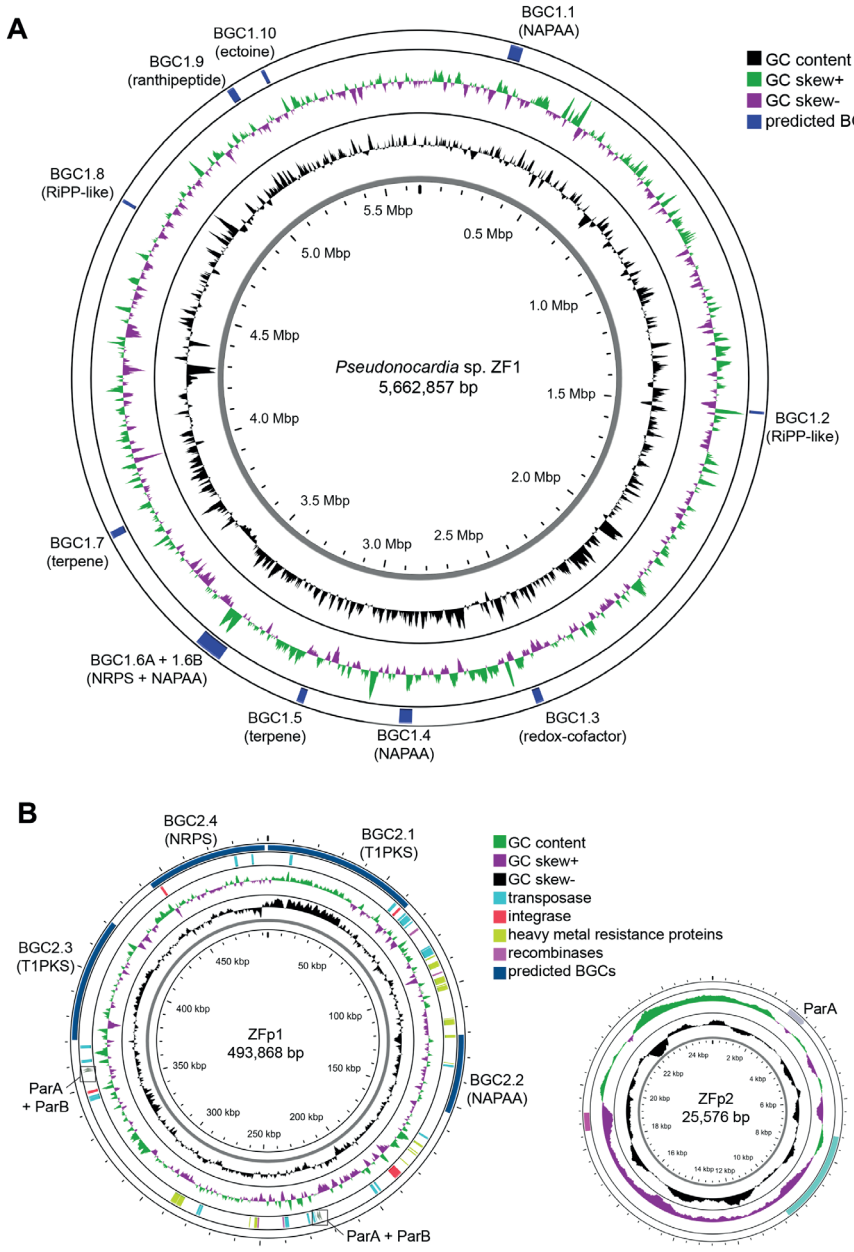

Figure 3. Schematic visualization of the circular chromosome of *Pseudonocardia* sp. ZF1 and the extrachromosomal elements. A-D) Representations of the chromosome (**A**) and plasmids ZFp1 and ZFp2 (**B**) were obtained using the CGview server [308]. In addition to GC content and CG skew, the location of the BGCs predicted by antiSMASH [217], and several plasmid elements are shown. Additionally, the ParA and ParB partition proteins are indicated. NAPAA: non-alpha poly-amino acids like e-Polylysin, NRPS: non-ribosomal peptide synthetase cluster, RiPP: ribosomally synthesised and post-translationally modified peptide product cluster, T1PKS: type 1 polyketide synthase.

Next, the antibiotic-producing potential of *Pseudonocardia* sp. ZF1 and the other isolated Actinobacteria was assessed. The isolates were grown on the nutrient-rich growth medium Nutrient Agar and the nutrient-poor Minimal Medium. Plates were inoculated with spots from glycerol stocks of the different isolates and after seven days of growth, plates were overlaid with soft agar containing the Gram-positive *Bacillus subtilis* 168 or the Gram-negative *Escherichia coli* ASD19. Of the isolates, only *Pseudonocardia* sp. ZF1 inhibited the growth of the tested indicator strains (Fig. 4). To identify the nature of the metabolites produced by *Pseudonocardia* ZF1, the strain was streaked on MM and NA agar plates, followed by ethyl acetate (EtOAc) extraction of the metabolites from the spent agar after seven days of growth. However, the crude extracts did not exhibit bioactivity. Also dereplication based on MS/MS spectra against the GNPS spectral library did not result in annotation of any known bioactive compounds [220]. Unpublished proteomics data indicates that three predicted BGCs are expressed when *Pseudonocardia* sp. ZF1 is grown on MM medium: BGC1.2 with no similarity to any known BGCs in the MIBiG repository, BGC1.6A with 60% similarity to the scabichelin BGC of *Streptomyces scabiei* 87.22, and the ectoine BGC1.10. Future research is needed to identify the bioactive metabolites and the cognate BGC, and analyse whether *Pseudonocardia* sp. ZF1 produces these molecules *in vivo.*

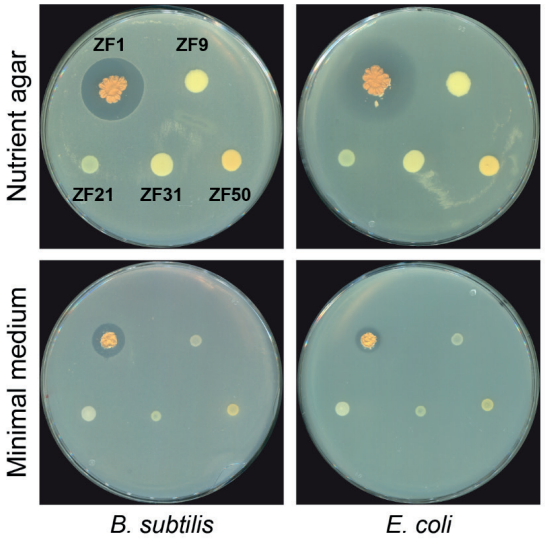

Figure 4. Antimicrobial activity of zebrafish isolates. After 4 days of growth, plates were overlaid with an indicator strain to test for antimicrobial activity (n = 3). Of the isolates, only *Pseudonocardia* sp. ZF1 inhibited the growth of *B. subtilis* and *E. coli.*

Discussion

As nature's medicine makers, Actinobacteria are welcome inhabitants of the microbiomes of eukaryotes, among others offering protection against invading pathogens. They likely do so in response to host-associated signaling molecules [2,44,301]. While a lot of attention is directed towards plant- and insect-Actinobacteria symbioses, the role of bioactive Actinobacteria within the microbiome of vertebrate organisms remains poorly characterised. Zebrafish larvae have been extensively used to study the role of the microbiome in health and disease. In this study, we explored the zebrafish microbiome for antibiotic-producing Actinobacteria to provide a first step towards using zebrafish as an *in vivo* model to study the bioactive and functional potential of Actinobacteria within animal microbiome.

16S rRNA amplicon sequencing revealed the presence of several actinobacterial genera commonly found in the human microbiome, such as *Bifidobacterium*, *Collinsella*, *Actinomyces*, *Rothia,* and *Corynebacterium* [310]. While several of these genera have received much attention with regard to their role in digestion and their abundance in correlation to health and disease, little is known about their specialised metabolites. Yet, a systematic analysis of the biosynthetic gene clusters in the human microbiome highlighted several BGCs that are harboured by members of *Corynebacterium. Actinomyces,* and *Rothia* [181]. Also, *Rothia* sp. isolated from the upper respiratory tract of weaned healthy piglets produce valinomycin *in vitro* and *in vivo* [311]. These findings highlight that much remains to be discovered about commensal Actinobacteria and their specialised metabolites in the animal microbiome. We failed to isolate members of these genera from zebrafish larvae. This is likely because of the isolation method used (aerobic and focused on filamentous Actinobacteria) and the young age of the zebrafish larvae. Future isolation studies will therefore focus on adult zebrafish and the use of different isolation methods and media suitable for both aerobic and anaerobic organisms.

Of the isolates, only *Pseudonocardia* sp. ZF1 exhibited filamentous growth, which has been linked to the propensity to produce bioactive metabolites and a high richness in BGC diversity. Therefore, the genome of *Pseudonocardia* sp. ZF1 was sequenced, resulting in three contigs, predicted to be circular. Contig 1 likely represents the chromosome, while the two smaller contigs (designated ZFp1 and ZFp2) may represent extrachromosomal plasmids. Indeed, different *Pseudonocardia* spp. carry one or multiple plasmids and may horizontally acquire BGCs first on their plasmids, which are later integrated in the chromosome [48,312,313]. Indeed, almost all predicted BGCs on contig 1 were shared with other *Pseudonocardia* spp., while the predicted BGCs on contig 2 were mostly unique to *Pseudonocardia* sp. ZF1. Due to the high levels of BGC diversity, plasmids of *Pseudonocardia* might be fruitful targets for drug discovery.

Pseudonocardia sp. ZF1 displays high similarity to *Pseudonocardia* sp. AL041005-10, isolated from the fungus-growing ant *Trachymyrmex cornetzi* [48]. The association of *Pseudonocardia* spp. with ants is one of the most well studied insect-microbe symbioses. Different *Pseudonocardia* sp. are present in the ant cuticle and aid in the protection of the ant's fungal garden against different fungal parasites [288-290]. Potentially, *Pseudonocardia* (or other bioactive Actinobacteria) can play a role in the protection of fish against fungal infections. Fungal infections by oomycetes, such as *Saprolegnia*, can cause severe losses of freshwater fish and their eggs in both nature and commercial fish farms [45,314]. Indeed, high abundance of specific Actinobacteria on salmon eggs has been correlated to a low incidence of *Saprolegnia* [45]. Of these Actinobacteria, *Streptomyces* isolates displayed antifungal activity against *Saprolegnia* while isolates of the genus *Frondihabitans* effectively inhibited its attachment to the salmon eggs, suggesting multiple protective mechanisms. It remains to be tested whether *Pseudonocardia* may also play a role in the protection of zebrafish against oomycetes.

Pseudonocardia spp. have mainly been isolated from soil, plants, marine sediments, and insects [315]. In this study, we have now isolated a fish-associated *Pseudonocardia* from zebrafish larvae. Using 16S rRNA amplicon sequencing we also confirmed a low abundance of *Pseudonocardia* within the intestine of adult zebrafish. This corresponds to several other studies analysing the microbiome composition of larvae and adult zebrafish [316-318]. Isolation of *Pseudonocardia* sp. ZF1 strengthens the finding that *Pseudonocardia* spp. are associated with zebrafish. The bioactivity of the isolate against Gram-positive and Gram-negative bacteria may suggest a protective role for *Pseudonocardia* sp. ZF1 against bacterial infections. Future studies are needed to identify the bioactive metabolites produced and to analyse whether the molecules are produced *in vivo*. Additionally, it would be of interest to test whether inoculation of zebrafish eggs and larvae with *Pseudonocardia* sp. ZF1 can protect against fungal and bacterial pathogen invasion.

Taken together, we show that antibiotic-producing *Pseudonocardia* with much uncharacterised potential are associated with zebrafish, providing a first step towards the use of zebrafish as an *in vivo* model to explore the functional potential of bioactive Actinobacteria within the animal microbiome.

Methods

16S rRNA gene amplicon sequencing of microbial content zebrafish intestine

Zebrafish were handled in compliance with animal welfare regulations and maintained according to standard protocols (http://zfn.org**).** The use of adult zebrafish was approved by the local animal welfare committee (DEC) of the University of Leiden (license number:

AVD1060020171767) and adhered to the international guidelines specified by the EU Animal Protection Directive 2010/63/EU. Adult zebrafish (n = 3) were euthanized using tricaine methanesulfonate (MS-222), the intestines were extracted surgically, transferred in one piece to a sterile Eppendorf tube containing 500 μL PBS, and cut lengthwise with sterilised scissors. Samples were vortexed thoroughly to free intestinal contents, after which the host tissue was removed. Next, samples were centrifuged at maximum speed, and supernatant was removed. DNA was extracted from the remaining pellet using phenol:chloroform extraction.

16S rRNA gene amplicon sequencing was performed on the V3-V4 hypervariable region of the microbial content of the adult zebrafish intestine using primers 341F (5'-CCTACGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGTATCTAATCC-3'). 2 x 300 bp paired-end demultiplexed sequenced reads were generated using the Miseq sequencing platform by BaseClear, The Netherlands. Reads were quality filtered based on the sequence centres Illumina Chastity filtering and FASTQC quality control tool (v 0.11.5) The adult intestinal zebrafish data comprised of three samples, with the number of reads retrieved for each sample ranging from 9,173 to 92,725**.**

The Qiime2 (v2020.8) bioinformatics platform [319] was used to analyse the 16S rRNA sequencing data. DADA2 (v1.10.0) [320] was used to model and correct amplicon sequence errors, providing amplicon sequence variants (ASV) and associated abundances. The primers were trimmed, and the forward and reverse sequences were truncated to remove low-quality regions (positions 288 and 222 respectively). MAFFT (v7.471) [321] was used to perform multiple sequence alignment and FastTree (2.1.10) [322] was used to generate a phylogenetic tree. Taxonomy was assigned using a naïve Bayesian classifier and using the Silva v138 ribosomal RNA reference database [323], which was made Qiime2 compatible using the RESCRIPt package. ASVs assigned to the taxonomic labels chloroplast and mitochondria were removed from the analysis.

Isolation of Actinobacteria from zebrafish larvae

At five days post fertilization, 100 larvae of ABxTL wild-type zebrafish were washed in sterile PBS and transferred to a 1.5 mL SafeLock tube containing 100 μL sterile PBS and 5 zirconium oxide beads (1.0 mm diameter). The larvae were anesthetised for 2 min on ice, homogenised for 2 x 1 min in a tissue homogenizer (Bullet Blender model Blue-CE; Next Advance) and cooled on ice in between. For the isolation of Actinobacteria, 20 µL homogenate was serial diluted (10^{-1} – 10^{-6}) in sterile PBS, and 100 µL of the dilutions were plated on selective media containing nystatin (50 μg/mL) and nalidixic acid (10 μg/mL) for the inhibition of fungi and Gram-negative bacteria, respectively. The media were Actinomycete Isolation Agar (Difco), Czapek Dox (Difco), ISP2 medium (DSMZ #987) . ISP5 medium (DSMZ #993), Kuster's agar

[324], Trypton Yeast Extract agar (DSMZ #680), Minimal Medium [227] and Humic Acid agar [226]. Plates were incubated at 30 °C and isolates were obtained by loop-streaking on fresh agar media until pure. For stock preparation, isolates were grown in ISP2 liquid medium and stocks were prepared according to Kieser et al. [227] and stored at −80 °C.

16S rRNA-based bacterial identification

Genomic DNA was isolated and purified from overnight liquid-cultures as described previously [227]. The 16S rRNA genes of the isolates were amplified in polymerase chain reactions (PCR) using the universal primers 8F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3'). Each PCR reaction (50 µL) contained 5 µL 5x PFU buffer, 5 µL DNTP mix (20 mM), 1.5 µL of each primer (10 µM) 5 µL DMSO, 1 µL PFU polymerase, and 1 µg gDNA in dH_2O. Amplification was achieved under the following conditions: an initial denaturation step of 3 min at 95 °C, followed by 35 cycles of denaturation (30 sec at 95 °C), annealing (30 sec at 56 °C), and extension (2 min at 72 °C), and a final extension step of 7 min at 72 °C. Sequencing was done by BaseClear, Leiden using Sanger sequencing with primer 8F. A 700 bp high quality region of the obtained sequence chromatograms was selected using Chromas Lite v2.01 and compared against the EZBioCloud 16S rRNA Database.

Genome sequencing

Strains were cultured in TSBS:YEME (50:50) at 30 °C with 200 rpm shaking speed [227]. Genomic DNA was isolated by phenol-chloroform extraction as described previously [227] and sent to be commercially sequenced at Future Genomics Technologies, The Netherlands. Genomes were sequenced using Illumina and Nanopore sequencing. Hybrid assembly was performed for each isolate using Unicycler (v0.4.0.7) [229]. Briefly, Unicycler performs a SPAdes assembly using the Illumina reads and then scaffolds the assembly graph using long reads. Unicycler polishes its final assembly with Illumina reads and uses Pilon to reduce the rate of small base-level errors [230]. The genome was annotated with the NCBI Prokaryotic Genome Annotation Pipeline and the genome sequence has been deposited at GenBank under accession number JAJNBY000000000.

Phylogenetic, bioactive potential and comparative genomics analysis

A phylogenetic tree was constructed based on whole-genome sequences using autoMLST [306]. AntiSMASH v6.0 [217] was used under default settings to predict BGCs from *Pseudonocardia* sp. ZF1 and its closest neighbour *Pseudonocardia* sp. AL041005-10. For the comparative genome analysis between these strains, BLASTP was used to identify similarity, where after MCScanX [235] was utilised to predict gene collinearity and synteny. Finally, the coordinates of the collinear regions and predicted BGCs were used as input for the visualizing tool Circos

v0.69-8 [236]. Cblaster [309] with default settings was used to compare the predicted BGCs of *Pseudonocardia* sp. ZF1 against NCBI sequence databases.

Antimicrobial activity assays

Strains were manually spotted (2 μL) on Nutrient Agar (Difco) and Minimal Medium agar plates (MM) supplemented with 0.5% mannitol and 1% glycerol (w/v) as non-repressing carbon sources. Both media were supplemented with 25 mM TES buffer. After four days of incubation, plates were overlaid with soft LB agar (1.8% w/v agar) containing one of the indicator strains pre-grown in liquid LB to exponential phase (OD600 = 0.4–0.6) and incubated overnight at 37 °C (± 18 hours). *Bacillus subtilis* 168 and *Escherichia coli* ASD19 were used as indicator strains for antimicrobial activity and were cultured in LB media at 37 °C.

Supplementary information for Chapter 6

Table S1. Genomic features of *Pseudonocardia* sp. ZF1

	Length (bp)	GC (%)	tRNAs	rRNAs
Contig 1	5,662,857	74.1	52	12
Contig 2	493,868	71.9		
Contig 3	25,576	70.4		

Table S2. Overview of BGCs predicted by antiSMASH 217 and analysis of homology using cblaster.

Proto-cluster	Type	Most similar known cluster	Top hit cblaster	Average identity
Contig 1				
1.1	NAPAA	CC-1065 (6%)	*Pseudonocardia pini* ICBG1122	87%
1.2	RiPP-like	-	*Pseudonocardia* sp. ICBG1034	99%
1.3	Redox-cofactor	Lankacidin C (13%)	*Pseudonocardia* sp. ICBG601	94%
1.4	NAPAA	Streptobacin (11%)	*Pseudonocardia antarctica* DSM 44749	99%
1.5	Terpene	SF2575 (6%)	*Pseudonocardia alni* DSM 44104	99%
1.6A*	NRPS	Scabichelin (60%)	*Pseudonocardia autotrophica* DSM 43083	100%
1.6B	NAPAA	Stenothricin (18%)	*Pseudonocardia alni* DSM 44104	99%
1.7	Terpene	Isorenieratene (28%)	*Pseudonocardia alni* DSM 44104	98%
1.8	RiPP-like	-	*Pseudonocardia* sp. ICBG1034	98%
1.9	Ranthipeptide	-	*Pseudonocardia autotrophica* DSM 43083	99%
1.10	Ectoine	Ectoine 100%	*Pseudonocardia* sp. AL041005-10	99%
Contig 2				
2.1*	T1PKS	Aculeximycin (41%)	No hits	-
2.2*	NAPAA	-	*Pseudonocardia autotrophica* NBRC 12743 plasmid pPA12743CP DNA	97%
2.3*	T1PKS	Coelimycin P1 (29%)	*Frankia* sp. BMG5.36	27%
2.4*	NRPS-like, NRPS	Paulomycin (5%)	*Pseudonocardia spinosispora* DSM 44797	49%

*Search was limited to only a small subset of genes of the predicted BGC as a search with the whole predicted cluster did not provide any hits

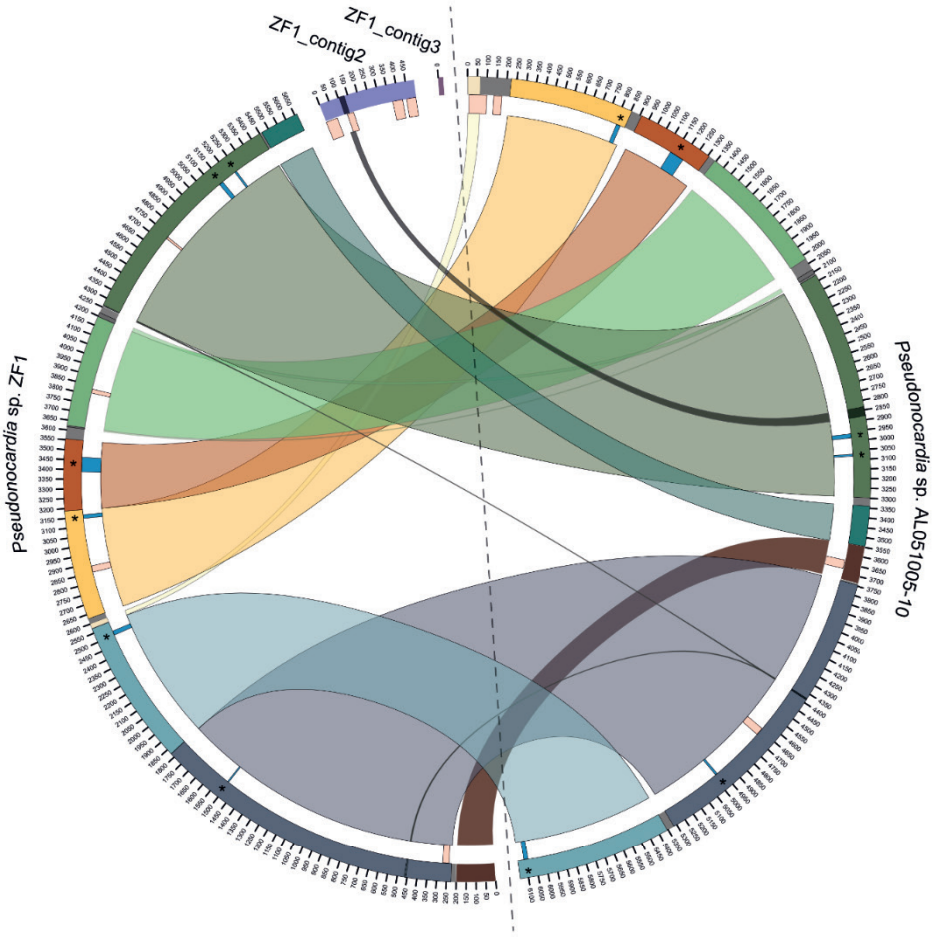

Figure S1. Synteny plot comparing *Pseudonocardia* sp. ZF1 to *Pseudonocardia* sp. AL41005-10. Genome sequences were compared by pairwise BLASTP and collinearity calculation method. Similarly coloured areas represent high similarity regions. The location of predicted BGCs by antiSMASH v6.0 are indicated in pink. Shared BGCs are indicated in blue and with an asterisk