frontiers in Microbiology

Genome analysis of two Pseudonocardia phylotypes associated with Acromyrmex leafcutter ants reveals their biosynthetic potential

Neil A. Holmes[2], Tabitha M. Innocent[1], Daniel Heine[3], Mahmoud Al Bassam[2], Sarah F. Worsley[2], Felix Trottmann[3], Elaine H. Patrick[2], Douglas W. Yu[2, 4], J. Colin Murrell[5], Morten Schiøtt[1], Barrie Wilkinson[3*], Jacobus J. Boomsma[1*], Matthew I. Hutchings[2*]

[1]Centre for Social Evolution, University of Copenhagen, Denmark, [2]School of Biological Sciences, University of East Anglia, United Kingdom, [3]Department of Molecular Microbiology, John Innes Centre, United Kingdom, [4]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, China, [5]School of Environmental Sciences, University of East Anglia, United Kingdom

Submitted to Journal:
Frontiers in Microbiology

Specialty Section:
Antimicrobials, Resistance and Chemotherapy

ISSN:
1664-302X

Article type:
Original Research Article

Received on:
15 Oct 2016

Accepted on:
08 Dec 2016

Provisional PDF published on:
08 Dec 2016

Frontiers website link:
www.frontiersin.org

Citation:
Holmes NA, Innocent TM, Heine D, Al_bassam M, Worsley SF, Trottmann F, Patrick EH, Yu DW, Murrell J, Schiøtt M, Wilkinson B, Boomsma JJ and Hutchings MI(2016) Genome analysis of two Pseudonocardia phylotypes associated with Acromyrmex leafcutter ants reveals their biosynthetic potential. *Front. Microbiol.* 7:2073. doi:10.3389/fmicb.2016.02073

Genome analysis of two *Pseudonocardia* phylotypes associated with *Acromyrmex* leafcutter ants reveals their biosynthetic potential

Neil A. Holmes[1], Tabitha M. Innocent[2], Daniel Heine[3], Mahmoud Al Bassam[1], Sarah F. Worsley[1], Felix Trottmann[3], Elaine H. Patrick[1], Douglas W. Yu[1,4], J. Colin Murrell[5], Morten Schiøtt[2], Barrie Wilkinson[3], Jacobus J. Boomsma[2], Matthew I. Hutchings[1]

[1]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk, United Kingdom.

[2]Centre for Social Evolution, University of Copenhagen, Copenhagen, Denmark.

[3]Department of Molecular Microbiology, John Innes Centre, Norwich Research Park, Norwich, Norfolk, United Kingdom.

[4]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Kunming, Yunnan,China.

[5]School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk, United Kingdom.

Correspondence:
Professor Jacobus Boomsma, Email: jjboomsma@bio.ku.dk
Professor Matt Hutchings, Email: m.hutchings@uea.ac.uk
Professor Barrie Wilkinson, Email: barrie.wilkinson@jic.ac.uk

Manuscript Length: 6672
Number of Figures: 8

Keywords: Leafcutter ants, antibiotics, actinomycetes, *Pseudonocardia*, nystatin, polyene, genome mining, *Acromyrmex*

Abstract.

The attine ants of South and Central America are ancient farmers, having evolved a symbiosis with a fungal food crop >50 million years ago. The most evolutionarily derived attines are the *Atta* and *Acromyrmex* leafcutter ants, which harvest fresh leaves to feed their fungus. *Acromyrmex* and many other attines vertically transmit a mutualistic strain of *Pseudonocardia* and use antifungal compounds made by these bacteria to protect their fungal partner against co-evolved fungal pathogens of the genus *Escovopsis*. *Pseudonocardia* mutualists associated with the attines *Apterostigma dentigerum* and *Trachymyrmex cornetzi* make novel cyclic depsipeptide compounds called gerumycins, while a mutualist strain isolated from derived *Acromyrmex octospinosus* makes an unusual polyene antifungal called nystatin P1. The novelty of these antimicrobials suggests there is merit in exploring secondary metabolites of *Pseudonocardia* on a genome-wide scale. Here we report a genomic analysis of the *Pseudonocardia* phylotypes Ps1 and Ps2 that are consistently associated with *Acromyrmex* ants collected in Gamboa, Panama. These were previously distinguished solely on the basis of 16S rRNA gene sequencing but genome sequencing of five Ps1 and five Ps2 strains revealed that the phylotypes are distinct species and each encodes between 11-15 secondary metabolite biosynthetic gene clusters (BGCs). There are signature BGCs for Ps1 and Ps2 strains and some that are conserved in both. Ps1 strains all contain BGCs encoding nystatin P1-like antifungals, while the Ps2 strains encode novel nystatin-like molecules. Strains show variations in the arrangement of these BGCs that resemble those seen in gerumycin gene clusters. Genome analyses and invasion assays support our hypothesis that vertically transmitted Ps1 and Ps2 strains have antibacterial activity that could help shape the cuticular microbiome. Thus, our work defines the *Pseudonocardia* species associated with *Acromyrmex* ants and supports the hypothesis that *Pseudonocardia* species could provide a valuable source of new antimicrobials.

1. Introduction.

Almost all antibiotics currently in clinical use are derived from the secondary metabolites of a group of soil bacteria called actinomycetes, but the discovery of these strains and their natural products (NPs) peaked in the 1950s. Since then, problems of rediscovery of extant strains and compounds has led to a decline in the discovery of new classes of metabolites. As a result, few new antibiotics have made it to market in the last 50 years. More recently, however, the field of NP discovery has been revitalised by large scale genome sequencing, which has revealed that actinomycetes express less than 25% of their secondary metabolite biosynthetic gene clusters (BGCs) *in vitro* (Doroghazi *et al.,* 2014). Genome mining for novel BGCs in soil actinomycetes isolated over the last 80 years, plus new actinomycete strains isolated from under-explored environments, promises to yield thousands of new NPs, including new anti-infective drugs (Katz and Baltz, 2016). One promising new approach is to genome mine strains that have co-evolved with their eukaryotic hosts. Such symbiotic relationships are known as protective mutualisms, because the plant or animal host houses, feeds, and sometimes vertically transmits the bacteria in exchange for antibiotics that protect them against infection (Clardy *et al.,* 2009; Kaltenpoth, 2009; Seipke *et al.,* 2011b). Arguably, the best characterized examples are the protective mutualisms between the attine ants of South and Central America and their vertically transmitted strains of *Pseudonocardia* (Cafaro *et al.,* 2011; Caldera and Currie, 2012).

The common ancestor of attine ants developed fungiculture 50-60 million years ago, leading to the tribe Attini, which consists of ca. 15 genera and more than 230 species (Nygaard *et al.,* 2016; Schultz and Brady, 2008). The most evolutionary derived attines are

the genera *Acromyrmex* and *Atta* which are known as leafcutter ants because they actively cut fresh leaves and feed them to their co-evolved symbiotic fungus *Leucoagaricus gongylophorus*. In return for the maintenance and manuring services provided by the ant farmers, the fungal cultivar produces gongylidia that are rich in fats and sugars and which the ants harvest as the sole food source for their larvae and queen (De Fine Licht *et al.,* 2014; Schiøtt *et al.,* 2010). *Acromyrmex* ants grow *Pseudonocardia* on their cuticles (Figure 1), housed in specialized crypts that are connected to subcuticular glands through which the hosts likely provide nutrients (Currie *et al.,* 2006)(Andersen *et al.,* 2013). The *Pseudonocardia* strain provides the ants with at least one antifungal compound and the ants use it to protect their fungus against disease, in particular from the specialized fungal pathogen *Escovopsis* (Barke *et al.,* 2010; Currie *et al.,* 2003; Oh *et al.,* 2009; Sit *et al.,* 2015).

Pseudonocardia are considered rare actinomycetes because they are hard to isolate from soil, and there are relatively few (<20) available genome sequences for this genus (Barke *et al.,* 2010; 2011; Sit *et al.,* 2015). Like other filamentous actinomycetes, they grow as multicellular hyphae and reproduce under nutrient stress conditions by erecting aerial hyphae that undergo cell division to form spores. The spores provide an effective dispersal mechanism for these non-motile bacteria. In the well-studied actinomycete genus *Streptomyces*, antibiotic production occurs at the onset of sporulation, but little is known about the developmental life-cycle of *Pseudonocardia* or the secondary metabolites they encode. It was recently discovered that *Pseudonocardia* mutualists of the lower attine *Apterostigma dentigerum* and the basal higher non-leafcutter attine *Trachymyrmex cornetzi* make cyclized depsipeptide antifungals called dentigerumycin and gerumycins, respectively, which are closely related in structure (Oh *et al.,* 2009; Sit *et al.,* 2015). In addition, we have previously reported that a *Pseudonocardia* mutualist of *Acromyrmex octospinosus* ants collected in Trinidad also makes a polyene antifungal named nystatin P1 (Barke *et al.,* 2010). Nystatin P1 is closely related to another polyene called NPP, which was subsequently identified in a strain called *Pseudonocardia autotrophica* of unknown origin (Lee *et al.,* 2012). These polyenes are assembled by a polyketide synthase (PKS) and are very different to gerumycins, but they are closely related to the widely used antifungal agent nystatin A1 made by *Streptomyces noursei* with nystatin P1 having an extra hexose sugar attached to its mycosamine moiety when compared to nystatin A1 (Barke *et al.,* 2010; Brautaset *et al.,* 2000). The BGC for nystatin P1 and NPP encodes an additional glycosyl transferase which attaches the second deoxysugar to the mycosamine already attached to the nystatin backbone (Barke *et al.,* 2010; Kim *et al.,* 2015). These additional tailoring enzymes have potential for bioengineering nystatin A1 producers, because the addition of an additional deoxysugar makes NPP 300 times more soluble in water (Lee *et al.,* 2012). Polyenes bind to ergosterol and form channels in the fungal cell membrane, which ultimately kills the fungus. They have broad spectrum activity, and resistance to polyenes is rare, including for clinically important molecules like amphotericins, which likely makes them valuable defence molecules against pathogens like *Escovopsis*, which are specifically adapted to exploit attine cultivars.

In this work we used whole-genome sequencing and analysis to investigate the secondary metabolites encoded by *Pseudonocardia* mutualists of *Acromyrmex echinatior* ants. All *Acromyrmex* colonies were collected from a population in Gamboa, Panama and are associated with one of two *Pseudonocardia* phylotypes distinguished by 16S rRNA gene sequencing and named Ps1 and Ps2. All ants within a single colony share the same phylotype, making the two strains almost completely mutually exclusive at the colony level (Andersen *et al.,* 2013; Poulsen *et al.,* 2005). Little is known about these strains, but previous studies have found clear differences between Ps1 and Ps2 at the 16S rRNA gene level (Poulsen *et al.,* 2005). Also, older worker ants carrying a Ps1 strain are associated with a more diverse host cuticular microbiome than those carrying Ps2 (Andersen *et al.,* 2013), and

there are some colony-level behavioural differences in the ant hosts that are related to disease control performance (Andersen *et al.,* 2015). We reasoned this could be due to differences in the *Pseudonocardia* symbiont growth rates, the antibiotics they encode, or both, and this prompted us to investigate the differences between the genomes of the two phylotypes.

In this study we report the first detailed analysis of the Ps1 and Ps2 *Pseudonocardia* phylotypes that are associated with *Acromyrmex* leafcutter ants and show that they are sufficiently different as to represent two distinct species with distinct BGC profiles. Ps1 strains encode 14-15 BGCs, of which seven are conserved in Ps1 but not found in Ps2 strains. Ps2 strains encode 11-15 BGCs, and five of these are unique to Ps2. Sic BGCs are shared between Ps1 and Ps2 strains and they all encode nystatin-like gene clusters. However, the Ps1 strains encode nystatin P1, while the Ps2 strains are predicted to make novel variants of nystatin. We also show that both phylotypes have antibacterial activity *in vitro*, which may help them to monopolize the entire cuticular microbiomes of large workers (Andersen *et al.,* 2013) as predicted by Scheuring & Yu (2012). The antibacterial NPs cannot be predicted from the BGC sequences alone, which suggests they have the potential to yield new classes of antibiotics.

2. Materials and methods

2.1 Ant collection and bacterial isolations.

This study used a total of 23 *Acromyrmex echinatior* colonies collected during fieldwork in the Gamboa area of Soberania National Park, Panama, between 2001 and 2014. Colonies were subsequently maintained at the University of Copenhagen, in climate-controlled rooms (ca. 70% humidity, 25°C), with the exception of the Ae707 sub-colony, which was set up in the laboratory for ~12 months at the University of East Anglia, and fed bramble leaves, apple, and dry rice, similar to the colonies maintained in Copenhagen. *Pseudonocardia* strains were isolated from five *A. echinatior* worker ants from each of the 23 captive colonies by direct contact of a needle with the laterocervical plates. Bacteria were cultured from five single worker ants taken from each colony (Figure 1). They were plated onto Lennox agar (20g agar, 10g Tryptone, 5g Yeast Extract, 5g NaCl made up to 1L with distilled water and then autoclaved) and incubated at 30°C until small white *Pseudonocardia* colonies were visible. Isolates were re-streaked until pure and then spread for confluent lawns on MS agar (20g agar, 20g mannitol, 20g soya flour made up to 1L with tap water and autoclaved twice) to induce sporulation. Spores were collected by gentle washing in 3mL sterile 20% (v/v) glycerol using sterile cotton buds and then transferred to 2mL screw cap tubes and stored at -20°C. Phylotype identification of *Pseudonocardia* was performed by PCR amplification of the 16S rRNA gene using primers PRK341F (5'-CCT ACG GGR BGC ASC AG-3') and MPRK806R (5'-GGA CTA CNN GGG TAT CTA AT-3').

2.2 Antibacterial bioassays.

To simulate challenges to *Pseudonocardia* dominated leafcutter ant cuticles and thereby assess invasion of *Pseudonocardia* microbiomes by other bacteria, we devised the following bioassays. *Pseudonocardia* spores were spread onto sterile cellophane disks placed on top of MS agar plates and grown at 30°C for seven days until they covered the plates in a confluent lawn. The cellophane disks were then peeled off to remove the *Pseudonocardia* and the agar plates were replenished with nutrients by adding 200µl LB nutrient broth and then air dried to get them ready for the bioassays. Unicellular bacterial test strains were grown from frozen glycerol stocks and grown overnight in 10mL Lennox broth at 37°C.

They were then sub-cultured 1:100 in 10mL fresh Lennox broth. Subcultures were grown for ~4.5 hours, after which an OD_{600} was measured, where $OD_{600} = 1$ was assumed to be = 8 x 10^9 cells. Working stocks were then made of the test strains to 10^8 cells. Equivalently, titred spore stocks of *Streptomyces* (Kieser *et al.,* 2000) were diluted to a working concentration of 10^8 spores. Subsequently, 10µl of the working stocks of all 10 bacterial test strains (five unicellular and five *Streptomyces*) were added to the plates by pipetting and the plates were incubated for 2 days at 30°C.

2.3 Genome sequencing and assembly.

High quality draft genomes of strains Ae150A_Ps1, Ae168_Ps1, Ae263_Ps1, Ae356_Ps1, Ae331_Ps2, Ae406_Ps2, Ae505_Ps2 and Ae706_Ps2 were obtained using Illumina technology. The two strains Ae707_Ps1 and Ae717_Ps2 were sequenced using PacBio to obtain high quality draft genomes to use as a reference for comparison. *Pseudonocardia* spores were inoculated onto Lennox agar plates covered with sterilized cellophanes. After 2-3 weeks of incubation at 30°C the cellophanes were removed and the mycelium was scraped off into sterile tubes. The salting out method was used to prepare genomic DNA from these mycelium samples (Kieser *et al.,* 2000). Minor modifications included the addition of achromopeptidase and DNase-free RNase (Qiagen) to the lysozyme step, as well as longer incubation steps for lysozyme and proteinase K treatment. DNA was quantified and quality checked using a Nanodrop 2000c spectrophotometer. Illumina sequencing of DNA was carried out at the DNA Sequencing Facility, Department of Biochemistry, University of Cambridge, UK, using TruSeq PCR-free and Nextera Mate Pair libraries and a MiSeq 600 sequencer. Genome assembly was performed using Roche Newbler v3.0, scaffolds were polished using PILON version 1.13, and reads were mapped using Burrows-Wheeler transformation version 0.7.12-r1039 (Walker *et al.,* 2014). PacBio sequencing of DNA was carried out at the Earlham Institute, Norwich Research Park, Norwich, UK, using SMRT bell adaptor libraries and PacBio standard large insert conditions, with 2-3 SMRT cells per sample. PacBio data were assembled using SMRT analysis software (Pacific Biosciences of California, Inc.) incorporating HGAP3 (Chin *et al.,* 2013). Assembled genomes were submitted to antiSMASH for BGC analysis and individual BGCs were compared across all strains using MultiGeneBlast (Weber *et al.,* 2015). Each BGC was downloaded as a Genbank file from the antiSMASH output and amino acid sequences from those clusters were used to generate a database. The database was then searched using individual BGCs as query sequences using the MultiGeneBlast algorithm. 16S rRNA, *rpsL* and *rpoB* sequences were aligned using ClustalW 2.1 (Larkin *et al.,* 2007). Blast analysis was carried out using ncbi-blast-2.2.31+ (Camacho *et al.,* 2009).

3. Results

3.1 Isolation of mutualist *Pseudonocardia* strains.

The large callow workers of *Acromyrmex echinatior* are covered in filamentous *Pseudonocardia* that are visible as a whitish covering on the cuticle (Figure 1) and can be cultured on nutrient agar. For a few weeks after hatching the filamentous bacterial growth can also be seen all over the worker ant cuticles and the rest of the body using scanning electron microscopy (Figure 1B)(Poulsen *et al.,* 2003). The *A. echinatior* workers typically have a concentration of *Pseudonocardia* on their laterocervical plates, so we scraped the plates of 1-2 workers taken from 23 separate *A. echinatior* colonies to isolate their mutualist strains. We consistently isolated either a single Ps1 or Ps2 *Pseudonocardia* strain from 22 out of the 23

ant colonies but for colony Ae_707, the only colony reared in Norwich rather than Copenhagen, we isolated both Ps1 and Ps2 strains. This result matched an earlier survey (Andersen *et al.,* 2015), which had a single double infection in a sample of 19 colonies of Panamanian *A. echinatior* and *A. octospinosus*. In total, we isolated 24 strains of *Pseudonocardia*. We then selected five Ps1 strains and five Ps2 strains for genome sequencing (Table 1); these included 6 strains originating from colonies also used in the Andersen *et al.,* (2015) study. Under laboratory culture conditions, Ps2 strains generally grew faster and generated more biomass than Ps1 strains, but all the *Pseudonocardia* isolates grew weakly compared to *Streptomyces* species and to *Pseudonocardia* strains isolated previously from *A. octospinosus* worker ants (Barke *et al.,* 2010). Only one of the *Pseudonocardia* strains sequenced grew in liquid culture, and we could not conjugate any vectors into the strains.

3.2 Ps1_Ae707 and Ps2_A706 strains have antibacterial activity.

Culture-dependent and independent studies have shown that *Pseudonocardia* dominate the cuticle of *Acromyrmex echinatior* ants, but closely related *Streptomyces albus* strains have also been isolated from *Acromyrmex* ants collected from different locations in South America (Barke *et al.,* 2010; Haeder *et al.,* 2009; Sen *et al.,* 2009). It is still not clear if these are mutualist strains, but they have been shown to produce antibiotics on the surface of worker ants using MALDI-TOF, which means they are growing and metabolically active on the ant cuticle (Schoenian *et al.,* 2011). Domination of the ant cuticle by *Pseudonocardia* is facilitated by vertical transmission and inoculation of worker ants within 24 hours of eclosing. Subsequently, the *Pseudonocardia* blooms over the entire surface of the ant before eventually shrinking back to the laterocervical plate when large workers are c.a. 6 weeks old (Poulsen *et al.,* 2003) when worker ants leave the nest to begin foraging (Marsh *et al.,* 2014). One way in which the *Pseudonocardia* mutualist might inhibit other bacteria from gaining a foothold on the cuticle is by making antibacterial compounds in addition to the antifungals used to suppress *Escovopsis* (Barke *et al.,* 2011; Scheuring and Yu, 2012). To test this hypothesis, we devised resistance bioassays in which *Pseudonocardia* strains were grown on sterile, porous cellophane disks placed on top of agar plates, allowing secondary metabolites produced by *Pseudonocardia* to permeate the agar.

After seven days, the cellophane and *Pseudonocardia* lawns were removed, the agar was replenished with nutrients by applying Lennox broth, and the surfaces of the agar plates were dried before inoculation with test strains of unicellular bacteria or filamentous *Streptomyces* species. Some of these were isolated previously from fungus-growing ant nests whereas some were soil-isolated *Streptomyces* strains (Table 1). The cellophane disks allowed exchange of nutrients and secondary metabolites, including antibiotics that might inhibit the growth of other bacteria. Results showed that both Ps1 and Ps2 strains could partially inhibit the growth of Gram-negative and Gram-positive bacteria, but the Ps1 strains appeared to be more potent, at least under the conditions used here. This suggests that both strains are producing antibacterial compounds, and since we know that typically ≤25% of BGCs are expressed *in vitro* it seemed likely that in nature they would exhibit more potent activity against other strains, particularly since the genomes of all 10 sequenced *Pseudonocardia* strains encode multiple potential bacteriocins. Consistent with our predictions (Barke *et al.,* 2011), only the unicellular bacterial strains were inhibited, whereas growth of the *Streptomyces* strains was unaffected or even enhanced compared with the control plate (Figure 2). This is consistent with the fact that *Streptomyces* strains carry multiple antibiotic resistance genes and can invade the ant cuticular microbiome (Barke *et al.,*

2011; Wright, 2007).

3.3 Ps1 and Ps2 strains are phylogenetically distinct.

For all Ps1 and Ps2 strains, their draft genomes consist of one large contig representing the majority of the chromosome (Supplementary Table 1). The additional, smaller contigs per strain could represent either unassembled parts of the chromosome or extrachromosomal plasmids. Alignment of full length 16S rRNA and *rpoB* and *rpsL* genes (Supplementary Material) showed that the Ps1 and Ps2 strains are more similar within (99.93-100%) than between strains (97.61-97.74%). For *rpoB*, Ps1 conservation was 99.74-100.00%, Ps2 conservation was 99.91-100.00%, and between Ps1 and Ps2 there was 94.96-95.10% identity. For the relatively short *rpsL* gene, both Ps1 and Ps2 conservation were 100% similar within each lineage, whereas between Ps1 and Ps2, mean similarity was 92.8%. We therefore conclude that the Ps1 and Ps2 phylotypes represent two different species of *Pseudonocardia.* Since the nystatin P1 producer (Ps1 phylotype) was first characterized in detail from the *Acromyrmex octospinosus* system, we suggest the name *Pseudonocardia octospinosus* for Ps1 and *Pseudonocardia echinatior* for Ps2.

3.4 Ps1 and Ps2 mutualists can be grouped according to their secondary metabolite BGCs.

To determine which secondary metabolites are encoded by the *Pseudonocardia* mutualists of *A. echinatior*, we used antiSMASH 3.0 to identify the BGCs in all 10 sequenced genomes (Figure 3, Supplementary Figures 1-10)(Weber *et al.,* 2015). Although in some cases we know that antiSMASH has called the BGCs bigger than they really are, we have adopted a consistent approach of presenting all the data from the antiSMASH analysis because for most BGCs we have no idea where they start and end. AntiSMASH outputs indicated that the Ps1 strains encoded for the production of 14-15 secondary metabolites, while the Ps2 strains encoded 11-15 (Figure 3). In order to verify the similarity of the clusters between strains, we used MultiGeneBlast (Figure 4-6)(Weber *et al.,* 2015). Through shared homology and gene architecture, we deduced that six of the BGCs were shared between all 10 strains, seven are unique to Ps1, and five are unique to Ps2 (Table 2). The BGCs shared by all 10 *Pseudonocardia* strains (clusters A-F) include the *ectABCD* operon (Figure 4F, cluster F) which encodes biosynthesis of the osmoprotectants ectoine and 5'-hydroxyectoine which are common in both Gram-positive and Gram-negative bacteria and often found in actinomycetes (Galinski *et al.,* 1985). There was a distinct bacteriocin BGC present in all Ps1 and Ps2 strains (Figure 4E, cluster E), and these proteinaceous molecules usually have antimicrobial activity, which can range from narrow to broad spectrum (Cotter *et al.,* 2013; Gross and Morell, 1971). Encoded within the cluster is a peptide with a TIGR04222 domain predicted to be processed as a ribosomally encoded peptide, several nucleases and a sulfur transferase plus accessory protein that could modify the peptide backbone. Ps1 and Ps2 strains share two terpene encoding BGCs (Figures 4B & D, clusters B & D), hydrocarbons comprised of isoprene-derived units that have diverse bioactivities including antimicrobial activity (Gallucci *et al.,* 2009; Gershenzon and Dudareva, 2007). One of these terpene BGCs (cluster B) has similarity to carotenoid BGCs, and encodes a polyprenyl synthetase and a phytoene synthase (Richter *et al.,* 2015), while the other shared terpene BGC (cluster D) contains a terpene cyclase closely related to a lycopene cyclase found in carotenoid biosynthesis. This cluster also shares similarity to a conserved cluster in *Rhodococcus* bacteria. The Ps1 and Ps2 strains also share an oligosaccharide cluster which encodes for

enzymes associated with the biosynthesis of deoxysugars and glycosyltransferases, but for which we are unable to predict a product (Figure 4A, cluster A).

All five Ps1 strains share three non-ribosomal peptide synthetase (NRPS) BGCs that are not present in the Ps2 strains (clusters I, K & M). We propose that these BGCs synthesize metal binding siderophore molecules, likely to sequester ferrous iron or other metal ions, particularly since these BGCs include have transporters and other siderophore associated genes. The cluster for Ps1 Siderophore 1 (Figure 5C, cluster I) is predicted to have the initial peptide arrangement of 2,3-DHB-ser-ala-(ala)-orn (second ala missing from Ae150A, Ae168, Ae263 and Ae356). Due to the predicted incorporation of a catechol unit (2,3-dihydroxybenzoic acid; 2,3-DHBA), a feature commonly found in siderophores and amino acids likely to interact with metal ions such as serine and ornithine, we propose a role in iron-acquisition for the secondary metabolite related to this cluster. The cluster for Ps1 Siderophore 2 (Figure 5E, cluster K) has similarity with the erythrochelin BGC which was first identified in *Saccharopolyspora erythraea* (Lazos *et al.,* 2010; Robbel *et al.,* 2010). The cluster for Ps1 Siderophore 2 has an NRPS that we predict generates a peptide arrangement of orn-ser-orn-orn, similar to the sequence for erythrochelin. However, there are a further 2-3 additional adenylation domains encoded by additional genes. Finally, the cluster for Ps1 siderophore 3 (Figure 5G, cluster M) we predict makes a tripeptide, with arrangement 2,3-DHBA then a polar residue either ser or thr, and lastly orn or arg. The last two amino acids we predict are epimerized to the D-configuration. We presume that iron is limited on ant cuticles, so siderophores would potentially play an important role for resource acquisition in these mutualist strains; efficient scavenging of iron and other metals by these species may also function to inhibit the growth of other microbes and therefore contribute to an anti-microbial phenotype. All Ps1 strains shared a novel cluster that was not present in any of the sequenced Ps2 strains (Figure 5A, cluster G), which contains several glycosyl transferases and enzymes capable of activating carboxylic acids (adenylation-like domains), and makes an unknown product. Ps1 strains also had an additional BGC (Figure 5B, cluster H) that we are uncertain makes a secondary metabolite and may be an artifact of the antiSMASH search algorithm. All the Ps1 strains encoded a further two BGC clusters that encode possible NRPSs. One cluster (Figure 5D, cluster J) encodes two adenylating enzymes, one of these shows similarity to an NRPS. This cluster is classified by antiSMASH as a bacteriocin cluster on account of containing a peptide that has similarity to the bacteriocin lincocin M18 produced by *Brevibacterium linens* (Valdés-Stauber and Scherer, 1994). The other BGC (Figure 5F, cluster L) that was maintained across all Ps1 strains has NRPS-like genes and possibly encodes an adenylation domain, a carrier protein and a cyclase. We predict it catalyzes synthesis of a dipeptide.

All five Ps2 strains share five BGCs that are not present in Ps1 strains (Table 2, Figure 6), including one bacteriocin BGC (Figure 6A, cluster N) which also has similarity to the bacteriocin lincocin M18 but only 74-78% similarity to the gene in the Ps1 strains (cluster L). This cluster does not have the same surrounding genes, including the adenylating enzymes. All Ps2 strains have a BGC (Figure 6B, cluster O) that has two adenylating proteins that could act as an NRPS as well as several oxidative proteins. All Ps2 strains also contained a genomic island containing BGCs with Type 1 PKS (T1 PKS) and NRPS genes (Figure 6C, cluster P). In Ae717_Ps2 and Ae505_Ps2, this genomic island was called as two separate and adjacent BGCs by antiSMASH, but they were called as one TIPKS-NRPS hybrid cluster in Ae331_Ps2, Ae406_Ps2 and Ae706_Ps2. It is likely that these are in fact three separate BGCs. In Ae717_Ps2 cluster 5 (cluster P part 1) makes a pentapeptide (predicted sequence; ala-thr-orn-ser/thr-orn). Being rich in threonine and ornithine residues suggests the resulting

peptide is likely to bind metal. Consistent with this, proteins immediately adjacent include a siderophore interacting protein member involved in metal acquisition/utilisation. Ae717_Ps2 cluster 5 (cluster P left) is separated with cluster 6 (cluster P centre/right) by apparent primary metabolism genes. Ae717_Ps2 cluster 6 can be split into two subclusters (cluster P centre/right); subcluster 6A (cluster P centre) is an NRPS-PKS hybrid consisting of modules for assembly of a glycine starter unit followed by five PKS extension steps. The PKS appears to have an unprecedented hybrid in *cis* and in *trans* acyltransferase architecture. Ae717_Ps2 subcluster 6A also has a type 1 glycosyl transferase and other deoxysugar genes associated with it. Subcluster 6B (cluster P right) encodes three co-linear NRPS genes that likely make a hepta/hexa-peptide including a 2,3-DHBA unit followed by cys-ala-ser/thr-orn-(D-orn) residues (the second orn is absent from Ae331_Ps2, A406_Ps2 and Ae706_Ps2). This arrangement strongly suggests a metal binding peptide likely to act as a siderophore. Ps2 strains also had a unique terpene BGC (Figure 6D, cluster Q) with some similarity to the brasilicardin BGC. Brasilicardin is a diterpenoid molecule possessing immunosuppressive activity produced by *Nocardia brasiliensis* (Hayashi *et al.,* 2008; Shigemori *et al.,* 1998). The Ps2 cluster could make a molecule with a cyclic core and additional elaboration. The last Ps2-specific BGC is predicted to encode biosynthesis of a lassopeptide (Figure 6E, cluster R), a ribosomally encoded peptide that is post-translationally modified. Lassopeptides form a distinct topology where the *N*-terminus is bound covalently to an aspartate or glutamate side chain further back in the peptide. The *C*-terminal end of the molecule is then threaded through the ring forming the "lasso" structure The Ps2 BGC encodes an asparagine synthase that could perform the peptide cyclization, and there is a GCN5 *N*-acetyltransferase that could perform the proteolytic action of peptide processing. There is a small open reading frame upstream of the asparagine synthase that could provide the starting peptide molecule. Lassopeptides belong to a class of natural product known as ribosomally-synthesized and post-translationally-modified peptides (RiPPs) and have been recognized as an important class of molecules because their intrinsic protease resistance gives them great potential as scaffolds for drug design (Hegemann *et al.,* 2015; Piscotta *et al.,* 2015). This BGC is encoded on a fragment of the genome sequence that we predict is a plasmid; this lassopeptide BGC could therefore be passed around a population of *Pseudonocardia* strains in a functionally dependent manner in response to environmental circumstances.

In addition to the BGCs that were found in either Ps1, Ps2 or both there were several BGCs specific to individual strains or subsets of strains (Table 2, Supplementary Figure 11). BGC 10 in strain Ae707_Ps1 is unique (Supplementary Figure 11A, cluster S), and antiSMASH did not find similarity with any known BGCs. Inspection of this cluster finds an AMP-dependent synthetase/ligase that may function as an NRPS. Strain Ae717_Ps2 encoded a unique BGC (Supplementary Figure 11B cluster T) with similarity to the BGC cluster for production of the galbonolides from *Streptomyces galbus* (Karki *et al.,* 2010). Galbonolides are anti-fungal macrolactones made by an iterative T1PKS. The Ae717_Ps2 BGC encodes an appropriate iterative T1PKS and other associated genes to make a related anti-fungal macrolactone. Strain Ae717_Ps2 also contained an additional BGC encoding for a T1PKS (Supplementary Figure 11C, cluster U), and only Ps2 strain Ae505_Ps2 shared this cluster. We predict the product of this BGC will be a highly unsaturated pentaketide derived from a 3-amino-5-hydroxybenzoic acid (AHBA) starter unit and four extension sites using malonyl-CoA as the substrate. This is highly likely to be modified by additional post-PKS genes encoded in the cluster. Ae717_Ps2 contained a unique T2PKS BGC that is highly likely to encode for a glycosylated T2PKS (Supplementary Figure 11D, cluster V). The DNA fragment encoding this T2PKS contains similarity to plasmid sequences so could be a recent addition to the genome of this strain. A BGC predicted to make a lantipeptide

(Supplementary Figure 11E, cluster W) is found in a subset of the Ps1 strains (Ae150A_Ps1, Ae168_Ps1, Ae263_Ps1 and Ae356_Ps1). Lantipeptides are RiPPs with posttranslational modifications in which amino acid residues are cross-linked via thioether bridges. The Ps1 lantibiotic clusters contain lantibiotic biosynthesis enzymes including the dehydratase and cyclase activities that are required to form lanthionine bridges (Karakas Sen *et al.,* 1999; Li *et al.,* 2006).

3.5 Ps1 and Ps2 phylotypes encode different polyene antifungals.

Perhaps the most intriguing feature of all 10 sequenced *Pseudonocardia* strains is the presence of nystatin-like biosynthesis gene clusters in their genome (Figure 4C). All Ps1 strains likely encode molecules similar to Nystatin P1 and NPP, previously reported as secondary metabolites from an *A. octospinosus* mutualist strain (Barke *et al.,* 2010) and from *P. autotrophica* (Kim *et al.,* 2009), respectively. As we obtained high quality PacBio sequenced genomes of Ae707_Ps1 and Ae717_Ps2, we used the nystatin-like clusters from each as templates for comparison in order to compare these BGCs across all 10 strains. The nystatin cluster from Ae707_Ps1, proves to be mostly identical to that encoded by the *A. octospinosus* mutualist strain we identified previously (Barke et al 2010) (data not shown). However, for comparison we performed bioinformatics analysis (Supplementary Table 5) against the *npp* cluster from *Pseudonocardia autotrophica* (Kim *et al.,* 2009) and demonstrate a high level of similarity between the amino acid sequences. In contrast, the nystatin-like cluster from Ae717_Ps2 shows less overall conservation with the *Pseudonocardia autotropica* NPP cluster (Supplementary Table 5). This prompted us to perform a detailed analysis and comparison of the polyene BGC for both phylotypes.

We have analysed the PKS architecture and predicted the assembled polyketide product of the nystatin cluster from Ae707_Ps1 (Figure 7). The polyene cluster has a total of 19 PKS modules and contains one ketoreductase domain (module 13) and two dehydratases (module 17 and 18) which are predicted to be inactive due to a lack of the catalytic tyrosine (ketoreductase) and histidine (dehydrogenase) motifs respectively. This is in conjunction with the biosynthetic cluster for NPP found in *Pseudonocardia autotrophica* (Kim *et al.,* 2009). Analysis of the KR sequence motifs (Keatinge-Clay, 2007) revealed all but one of the stereocentres being consistent with the configuration in nystatin (Brautaset *et al.,* 2000). The KR responsible for reduction leading to the hydroxyl group at position 19 shows an LDD motif (compared to the catalytically inactive LDA motif at this position in the original nystatin producer *Streptomyces noursei*). Due to the presence of potentially active LDD and W motifs in module 10 we cannot predict the stereochemistry with absolute certainty. We previously documented the presence of a second glycosyl transferase (NypY) for addition of a second hexose unit to nystatin P1 (Barke *et al.,* 2010). The NPP cluster encodes an additional glycosyl transferase (NppY) predicted to add *N*-acetyl-glucosamine (Kim *et al.,* 2015). Accordingly, Ae707_Ps1 and all the Ps1 strains encode an additional glycosyl transferase.

In contrast to Ps1, Ps2 strains encode a nystatin-like biosynthetic pathway that differs considerably from nystatin A1, NPP and nystatin P1 (Figure 8). First, the cluster seems to lack several modules compared to the BGC in Ae707_Ps1, including the loading module, the first two elongation modules, and a module responsible for the incorporation of a fully reduced acetate unit. This presumably results in a considerable structural re-arrangement of the polyketide scaffold and in a conjugated tetraene moiety in the southern part of the molecule. Instead of the loading KS^S, there is a CoA ligase (CAL) domain encoded in the

cluster, potentially responsible for providing the starter unit. Involvement of CoA ligases in PKS assembly lines is rare and usually associated with the recruitment of aromatic starter units. The CAL domain is predicted to accept AHBA as its substrate, and consistent with this a gene belonging to the 3-amino-5-hydroxybenzoic acid synthase (AHBAS) family can be found downstream of the PKS coding sequence. However, we cannot completely rule out the incorporation of a distinct but structurally related aromatic moiety into the polyene. Likewise, there appears to be a considerable re-arrangement in the later module architecture with the incorporation of an additional, fully reduced acetate unit (module 14) and the substitution of a fully reduced acetate unit with an ethylene moiety (module 12). These changes presumably lead to an altered cyclization pattern for the final product. Moreover, careful analysis of all KR sequence motifs suggests a significantly altered configuration at most of the predicted stereocenters compared to nystatin. Even though the KR in module 9 contains a VDD motif instead of LDD we suggest that catalytic activity remains after the replacement of leucine by valine, and propose the assembly of an *S*-configured hydroxyl group at position 17. We suggest a potential macrolactonization with the phenol group of AHBA for cleaving off the final polyketide chain but cannot rule out the formation of the alternative amide bond and macrolactamization instead. Another intriguing feature is that all polyene clusters found in the PS2 strains contain an additional methyltransferase and an *O*-methyltransferase, a feature rather uncommon for polyene antifungals. In contrast to the cluster in all Ps1 strains, the Ps2 polyene clusters do not encode a second glycosyltransferase and therefore potentially produce a mono-glycosylated product. Based on this analysis it is highly likely that Ps2 encodes a structurally distinct polyene macrolide with altered biological properties. Unfortunately, we have not been able to induce the expression of this silent BGC in any of the Ps2 strains so far.

To compare BGCs of the *Pseudonocardia* strains sequenced by Illumina we mapped the PKS amino acid sequences of either Ae707_Ps1 or Ae717_Ps2 to genomes sequences of the other Ps strains using blastx (Supplementary Table 6) and closely inspected antismash outputs and multigene blast (Figure 4C). Despite all the *Pseudonocardia* strains containing genes for the biosynthesis of nystatin-like molecules, there is a disparity between the genomic arrangements of the nystatin-like clusters in the *A. echinatior* associated strains that we analyzed. We observe something of a paralogous evolutionary example in the variance of the genomic arrangement of the hybrid NRPS/PKS BGCs for the gerumycins produced by *Pseudonocardia* mutualists of *Apterostigma* and *Trachymyrmex* ants, where the clusters are spatially separated, in some cases on plasmids (Sit *et al.,* 2015). Ae707_Ps1 is the only Ps1 strain to contain all the PKS genes in a single BGC. The other Ps1 strains have clusters split into two different locations on the chromosome and are also are missing two PKS modules (Supplementary Table 6) suggesting they make molecules with a smaller macrolide ring size. In contrast the Ps2 phylotype strains Ae331_Ps2, Ae406_Ps2, Ae706_Ps2 and Ae717_Ps2 all contain full intact clusters at a single locus on the chromosome. However, the Ae505_Ps2 BGC contains only half of the genes expected for the nystatin-like cluster and may encode for a shorter polyketide molecule or may be inactive. The polyene PKS architectures are essentially identical except for the aforementioned Ae505_Ps2 example.

4. Discussion.

The Ps1/Ps2 symbiosis with Panamanian *Acromyrmex* species is one of the best-studied mutualisms involving actinomycete bacteria. The two *Pseudonocardia* phylotypes were first recognized as distinct more than 10 years ago and were subsequently found to be shared by sympatric populations of *A. echinatior* and *A. octospinosus* in Panama (Andersen *et al.,* 2015; Poulsen *et al.,* 2005) and *A. volcanus* (Sapountzis *et al.,* 2015). Not only are the two

phylotypes shared between congener ant species, they are also consistently found in ca. 50/50 frequencies across colonies, and this is reflected in our genome sequencing of strains for the present study (Andersen *et al.,* 2015). This is a remarkable natural distribution in light of the findings we present here on how different the two strains are at the genome level. There is some quantitative evidence (Currie *et al.,* 2003) that growth patterns on the cuticles of large workers respond to *Escovopsis* infections, but whether there are differences between the Ps1 and Ps2 strain in the efficiency of their defenses has not been studied, although recent cross fostering experiments suggests such differences might exist (Andersen *et al.,* 2015).

Here we have defined these phylotypes by genome sequencing five representatives from each and demonstrated that these strains are sufficiently distantly related as to be classified as separate species (see Supporting Information). We have thus named them *P. octospinosus* (Ps1) and *P. echinatior* (Ps2), and shown that they are remarkably different in the type of secondary metabolites that they can produce. Seven BGCs are found only in *P. octospinosus* strains whereas five BGCs are only found in *P. echinatior* strains suggesting that ~50% of the secondary metabolites are unique to each species. Only six BGCs are shared between the two *Pseudonocardia* species and some of these, such as ectoine, are likely part of the core genome. Our results suggest that there are two major ways in which to maintain a defensive cuticular microbiome and that both – as the field data also suggests – appear to create colonies of comparable health and reproductive fitness that coexist in the same populations. It is interesting that Ps1 and Ps2 make different siderophore/iron-binding molecules rather than conserving the same set. As cross-inoculation is feasible and produces only subtle changes in ant behavior (Andersen *et al.,* 2015), it might be that both strains sequester iron with about equal efficiency and that the glands whose ducts end in the cuticular crypts where *Pseudonocardia* grows, secrete the iron to maintain the different types of *Pseudonocardia* biofilms on large workers of *Acromyrmex* leafcutter ants. Clarifying these enigmatic issues remains a challenge to be resolved in future research.

In terms of antifungal metabolites, we obtained a detailed assessment of all the BGCs encoded by *P. octospinosus* and *P. echinatior*. Strains of *P. octospinosus* all encode forms of nystatin P1 that we first identified in a *Pseudonocardia* strain associated with *Acromyrmex octospinosus* workers collected in Trinidad (Barke *et al.,* 2010). We also compared our 10 genome sequences with those for *Pseudonocardia* strains isolated from *Apterostigma* and *Trachymyrmex* ants, which make cyclic depsipeptide antifungals called gerumycins (Oh *et al.,* 2009; Sit *et al.,* 2015). We were unable to identify the hybrid NRPS/PKS BGC responsible for the biosynthesis of gerumycin compounds in either *P. echinatior* or *P. octospinosus* strains. Symbionts of *Apterostigma* ants have recently been shown to encode nystatin-like antifungals called selvamicins (Van Arnam *et al.,* 2016) and strain AL041005-10 isolated from *Trachymyrmex cornetzi* also encodes a BGC resembling the Nystatin P1 BGC. *Apterostigma* are lower attines and operate a more primitive fungiculture than *Trachymyrmex* and *Acromyrmex,* both of which have sophisticated methods to farm fungus material so it is intriguing that they all use polyenes. Resistance to polyene antifungals is rare and this might make them useful molecules in the fight against co-evolving *Escovopsis* parasites.

Evolution of secondary metabolite production in the system of fungus farming ants appears to be an interesting area, with *Pseudonocardia* genome analysis providing abundant insights. As we have found here structural arrangements of BGCs may be under constant selective pressure and rearrangements of nystatin polyene BGCs either to spatially separated positions or adaption of module arrangement and makeup is of significant note. We also

identified several BGCs appearing on DNA with similarity to plasmids, suggesting that *Pseudonocardia* associated with fungus farming ants may be inclined to pass around BGCs via horizontal gene transfer. In regard to evolution in fungus farming ant systems, genome sequencing recently revealed that *Escovopsis weberi* has a reduced genome, most likely due to its role as a parasite on the fungal garden of attine ants (de Man *et al.,* 2016). Despite loss of many genes, however, *E. weberi* has maintained a number of genes involved in secondary metabolite biosynthesis, suggesting that a microbial war is active in the nests of attine ants. It cannot be ruled out that compounds produced directly by attine ants or the fungal garden strain *Leucoagaricus gongylophorus* are included in this chemical warfare as the fungal symbiont and the cuticular *Pseudonoardia* are vertically co-transmitted by default. It will be interesting in the future to examine the compounds produced by *E. weberi* and *Pseudonocardia* mutualists grown in competition on agar plates and on fungus gardens or ant cuticles using advanced imaging mass spectrometry techniques.

Acknowledgements.

The work was supported by NERC Grants NE/M015033/1 and NE/M014657/1 awarded to MIH and BW. JJB was supported by the European Research Council (ERC Advanced grant 323085), and TMI by a Marie Curie Individual European Fellowship (IEF grant 627949). We thank Kim Findlay at the John Innes Centre Bioimaging Facility for SEM, and David Nash at the Centre for Social Evolution for photographs, of *Acromyrmex echinatior* worker ants. The Smithsonian Tropical Research Institute provided logistic help and facilities to work in Gamboa, and the Autoridad Nacional del Ambiente yel Mar gave permission to sample and export ants from Panama. We thank the DNA Sequencing Facility, Department of Biochemistry, University of Cambridge, UK, especially Dr Markiyan Samborskyy and the Earlham Institute, Norwich Research Park, Norwich, UK for DNA sequencing and assembly services.

References

Andersen, S. B., Hansen, L. H., Sapountzis, P., Sørensen, S. J., and Boomsma, J. J. (2013). Specificity and stability of the *Acromyrmex*- *Pseudonocardia* symbiosis. *Mol Ecol* 22(16), 4307–4321. doi:10.1111/mec.12380.

Andersen, S. B., Yek, S. H., Nash, D. R., and Boomsma, J. J. (2015). Interaction specificity between leaf-cutting ants and vertically transmitted *Pseudonocardia* bacteria. *BMC Evol Biol* 15, 27. doi:10.1186/s12862-015-0308-2.

Barke, J., Seipke, R. F., Gruschow, S., Heavens, D., Drou, N., Bibb, M. J., et al. (2010). A mixed community of actinomycetes produce multiple antibiotics for the fungus farming ant *Acromyrmex octospinosus*. *BMC biology* 8, 109. doi:10.1186/1741-7007-8-109.

Barke, J., Seipke, R. F., Yu, D. W., and Hutchings, M. I. (2011). A mutualistic microbiome: How do fungus-growing ants select their antibiotic-producing bacteria? *Commun Integr Biol* 4, 41–43. doi:10.4161/cib.4.1.13552.

Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A.-M., Challis, G. L., Thomson, N. R., James, K. D., et al. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141–147. doi:10.1038/417141a.

626 Brautaset, T., Sekurova, O. N., Sletta, H., Ellingsen, T. E., StrŁm, A. R., Valla, S., et al.
627 (2000). Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei*
628 ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway.
629 *Chem Biol* 7, 395–403.

630 Cafaro, M. J., Poulsen, M., Little, A. E. F., Price, S. L., Gerardo, N. M., Wong, B., et al.
631 (2011). Specificity in the symbiotic association between fungus-growing ants and
632 protective *Pseudonocardia* bacteria. *Proc. R. Soc. B* 278, 1814-1822.
633 doi:10.1098/rspb.2010.2118.

634 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009).
635 BLAST+: architecture and applications. *BMC Bioinformatics* 10, 421. doi:10.1186/1471-
636 2105-10-421.

637 Caldera, E. J., and Currie, C. R. (2012). The Population Structure of Antibiotic-Producing
638 Bacterial Symbionts of *Apterostigma dentigerum* Ants: Impacts of Coevolution and
639 Multipartite Symbiosis. *Am Nat* 180, 604–617. doi:10.1086/667886.

640 Chin, C.-S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013).
641 Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing
642 data. *Nat Methods* 10, 563–569. doi:10.1038/nmeth.2474.

643 Clardy, J., Fischbach, M. A., and Currie, C. R. (2009). The natural history of antibiotics. *Curr*
644 *Biol* 19, R437–41. doi:10.1016/j.cub.2009.04.001.

645 Cotter, P. D., Ross, R. P., and Hill, C. (2013). Bacteriocins - a viable alternative to
646 antibiotics? *Nat Rev Micro* 11, 95–105. doi:10.1038/nrmicro2937.

647 Currie, C. R., Bot, A., and Boomsma, J. J. (2003). Experimental evidence of a tripartite
648 mutualism: bacteria protect ant fungus gardens from specialized parasites. *Oikos* 101, 91-
649 102. doi:10.1034/j.1600-0706.2003.12036.x.

650 Currie, C. R., Poulsen, M., Mendenhall, J., Boomsma, J. J., and Billen, J. (2006). Coevolved
651 crypts and exocrine glands support mutualistic bacteria in fungus-growing ants. *Science*
652 311, 81–83. doi:10.1126/science.1119744.

653 De Fine Licht, H. H., Boomsma, J. J., and Tunlid, A. (2014). Symbiotic adaptations in the
654 fungal cultivar of leaf-cutting ants. *Nat Commun* 5, 5675. doi:10.1038/ncomms6675.

655 de Man, T. J. B., Stajich, J. E., Kubicek, C. P., Teiling, C., Chenthamara, K., Atanasova, L.,
656 et al. (2016). Small genome of the fungus *Escovopsis weberi*, a specialized disease agent
657 of ant agriculture. *Proc Natl Acad Sci USA* 113(13), 3567-3572.
658 doi:10.1073/pnas.1518501113.

659 Doroghazi, J. R., Albright, J. C., Goering, A. W., Ju, K.-S., Haines, R. R., Tchalukov, K. A.,
660 et al. (2014). A roadmap for natural product discovery based on large-scale genomics and
661 metabolomics. *Nat Chem Biol* 10, 963-968. doi:10.1038/nchembio.1659.

662 Galinski, E. A., Pfeiffer, H. P., and Trüper, H. G. (1985). 1,4,5,6-Tetrahydro-2-methyl-4-
663 pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic
664 bacteria of the genus *Ectothiorhodospira*. *Eur. J. Biochem.* 149, 135–139.

Gallucci, M. N., Oliva, M., and Casero, C. (2009). Antimicrobial combined action of terpenes against the food-borne microorganisms *Escherichia coli*, *Staphylococcus aureus* and *Bacillus cereus*. *Flavour Fragr. J.* 24, 348-354. doi:10.1002/ffj.1948.

Gershenzon, J., and Dudareva, N. (2007). The function of terpene natural products in the natural world. *Nat Chem Biol* 3, 408–414. doi:10.1038/nchembio.2007.5.

Gross, E., and Morell, J. L. (1971). The structure of nisin. *J Am Chem Soc* 93, 4634–4635.

Haeder, S., Wirth, R., Herz, H., and Spiteller, D. (2009). Candicidin-producing *Streptomyces* support leaf-cutting ants to protect their fungus garden against the pathogenic fungus *Escovopsis*. *Proc Natl Acad Sci USA* 106, 4742–4746. doi:10.1073/pnas.0812082106.

Hayashi, Y., Matsuura, N., Toshima, H., Itoh, N., Ishikawa, J., Mikami, Y., et al. (2008). Cloning of the gene cluster responsible for the biosynthesis of brasilicardin A, a unique diterpenoid. *J Antibiot* 61, 164–174. doi:10.1038/ja.2008.126.

Hegemann, J. D., Zimmermann, M., Xie, X., and Marahiel, M. A. (2015). Lasso Peptides: An Intriguing Class of Bacterial Natural Products. *Acc. Chem. Res.* 48, 1909-1919. doi:10.1021/acs.accounts.5b00156.

Kaltenpoth, M. (2009). Actinobacteria as mutualists: general healthcare for insects? *Trends Microbiol* 17, 529–535. doi:10.1016/j.tim.2009.09.006.

Karakas Sen, A., Narbad, A., Horn, N., Dodd, H. M., Parr, A. J., Colquhoun, I., et al. (1999). Post-translational modification of nisin. The involvement of NisB in the dehydration process. *Eur. J. Biochem.* 261, 524–532.

Karki, S., Kwon, S.-Y., Yoo, H.-G., Suh, J.-W., Park, S.-H., and Kwon, H.-J. (2010). The methoxymalonyl-acyl carrier protein biosynthesis locus and the nearby gene with the beta-ketoacyl synthase domain are involved in the biosynthesis of galbonolides in *Streptomyces galbus*, but these loci are separate from the modular polyketide synthase gene cluster. *FEMS Microbiol. Lett.* 310, 69–75. doi:10.1111/j.1574-6968.2010.02048.x.

Katz, L., and Baltz, R. H. (2016). Natural product discovery: past, present, and future. *J Ind Microbiol Biotechnol* 43, 155-176. doi:10.1007/s10295-015-1723-5.

Keatinge-Clay, A. T. (2007). A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem Biol* 14, 898–908. doi:10.1016/j.chembiol.2007.07.009.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000). *Practical Streptomyces genetics.* Norwich: John Innes Centre.

Kim, B.-G., Lee, M.-J., Seo, J., Hwang, Y.-B., Lee, M.-Y., Han, K., et al. (2009). Identification of functionally clustered nystatin-like biosynthetic genes in a rare actinomycetes, *Pseudonocardia autotrophica*. *J Ind Microbiol Biotechnol* 36, 1425–1434. doi:10.1007/s10295-009-0629-5.

Kim, H.-J., Kim, M.-K., Lee, M.-J., Won, H.-J., Choi, S.-S., and Kim, E.-S. (2015). Post-PKS tailoring steps of a disaccharide-containing polyene NPP in *Pseudonocardia autotrophica*. *PLoS ONE* 10(4), e0123270. doi:10.1371/journal.pone.0123270.

703 Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam,
704 H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948.
705 doi:10.1093/bioinformatics/btm404.

706 Lazos, O., Tosin, M., Slusarczyk, A. L., Boakes, S., Cortés, J., Sidebottom, P. J., et al.
707 (2010). Biosynthesis of the putative siderophore erythrochelin requires unprecedented
708 crosstalk between separate nonribosomal peptide gene clusters. *Chem Biol* 17, 160–173.
709 doi:10.1016/j.chembiol.2010.01.011.

710 Lee, M.-J., Kong, D., Han, K., Sherman, D. H., Bai, L., Deng, Z., et al. (2012). Structural
711 analysis and biosynthetic engineering of a solubility-improved and less-hemolytic
712 nystatin-like polyene in *Pseudonocardia autotrophica*. *Appl Microbiol Biotechnol* 95,
713 157–168. doi:10.1007/s00253-012-3955-x.

714 Li, B., Yu, J. P. J., Brunzelle, J. S., Moll, G. N., van der Donk, W. A., and Nair, S. K. (2006).
715 Structure and mechanism of the lantibiotic cyclase involved in nisin biosynthesis.
716 *Science* 311, 1464–1467. doi:10.1126/science.1121422.

717 Marsh, S. E., Poulsen, M., Pinto-Tomás, A., and Currie, C. R. (2014). Interaction between
718 Workers during a Short Time Window Is Required for Bacterial Symbiont Transmission
719 in *Acromyrmex* Leaf-Cutting Ants. *PLoS ONE* 9(7), e103269.
720 doi:10.1371/journal.pone.0103269.s002.

721 Nygaard, S., Hu, H., Li, C., Schiøtt, M., Chen, Z., Yang, Z., et al. (2016). Reciprocal
722 genomic evolution in the ant–fungus agricultural symbiosis. *Nature Communications* 7,
723 1–9. doi:10.1038/ncomms12233.

724 Oh, D.-C., Poulsen, M., Currie, C. R., and Clardy, J. (2009). Dentigerumycin: a bacterial
725 mediator of an ant-fungus symbiosis. *Nat Chem Biol* 5, 391–393.
726 doi:10.1038/nchembio.159.

727 Piscotta, F. J., Tharp, J. M., Liu, W. R., and Link, A. J. (2015). Expanding the chemical
728 diversity of lasso peptide MccJ25 with genetically encoded noncanonical amino acids.
729 *Chem. Commun.* 51, 409–412. doi:10.1039/C4CC07778D.

730 Poulsen, M., Bot, A. N. M., and Boomsma, J. J. (2003). The effect of metapleural gland
731 secretion on the growth of a mutualistic bacterium on the cuticle of leaf-cutting ants.
732 *Naturwissenschaften* 90, 406–409. doi:10.1007/s00114-003-0450-3.

733 Poulsen, M., Cafaro, M., Boomsma, J. J., and Currie, C. R. (2005). Specificity of the
734 mutualistic association between actinomycete bacteria and two sympatric species of
735 *Acromyrmex* leaf-cutting ants. *Mol Ecol* 14, 3597–3604. doi:10.1111/j.1365-
736 294X.2005.02695.x.

737 Richter, T. K. S., Hughes, C. C., and Moore, B. S. (2015). Sioxanthin, a novel glycosylated
738 carotenoid, reveals an unusual subclustered biosynthetic pathway. *Environ Microbiol* 17,
739 2158–2171. doi:10.1111/1462-2920.12669.

740 Robbel, L., Knappe, T. A., Linne, U., Xie, X., and Marahiel, M. A. (2010). Erythrochelin--a
741 hydroxamate-type siderophore predicted from the genome of *Saccharopolyspora*
742 *erythraea*. *FEBS Journal* 277, 663–676. doi:10.1111/j.1742-4658.2009.07512.x.

743 Sapountzis, P., Zhukova, M., Hansen, L. H., Sørensen, S. J., Schiøtt, M., and Boomsma, J. J.
744 (2015). Acromyrmex Leaf-Cutting Ants Have Simple Gut Microbiota with Nitrogen-
745 Fixing Potential. *Appl Environ Microbiol* 81, 5527–5537. doi:10.1128/AEM.00961-15.

746 Scheuring, I., and Yu, D. W. (2012). How to assemble a beneficial microbiome in three easy
747 steps. *Ecology Letters* 15, 1300–1307. doi:10.1111/j.1461-0248.2012.01853.x.

748 Schiøtt, M., Rogowska-Wresinska, A., Roepstorff, P., and Boomsma, J. J. (2010). Leaf-
749 cutting ant fungi produce cell wall degrading pectinase complexes reminiscent of
750 phytopathogenic fungi. *BMC Biology* 8, 156. doi:10.1186/1741-7007-8-156.

751 Schoenian, I., Spiteller, M., Ghaste, M., Wirth, R., Herz, H., and Spiteller, D. (2011).
752 Chemical basis of the synergism and antagonism in microbial communities in the nests of
753 leaf-cutting ants. *Proc Natl Acad Sci USA* 108, 1955–1960.
754 doi:10.1073/pnas.1008441108.

755 Schultz, T. R., and Brady, S. G. (2008). Major evolutionary transitions in ant agriculture.
756 *Proc Natl Acad Sci USA* 105, 5435–5440. doi:10.1073/pnas.0711024105.

757 Seipke, R. F., Barke, J., Heavens, D., Yu, D. W., and Hutchings, M. I. (2013). Analysis of the
758 bacterial communities associated with two ant-plant symbioses. *MicrobiologyOpen* 2,
759 276–283. doi:10.1002/mbo3.73.

760 Seipke, R. F., Crossman, L., Drou, N., Heavens, D., Bibb, M. J., Caccamo, M., et al. (2011a).
761 Draft Genome Sequence of Streptomyces Strain S4, a Symbiont of the Leaf-Cutting Ant
762 *Acromyrmex octospinosus*. *J Bacteriol* 193, 4270–4271. doi:10.1128/JB.05275-11.

763 Seipke, R. F., Kaltenpoth, M., and Hutchings, M. I. (2011b). *Streptomyces* as symbionts: an
764 emerging and widespread theme? *FEMS Microbiol Rev* 36, 862-876. doi:10.1111/j.1574-
765 6976.2011.00313.x.

766 Sen, R., Ishak, H. D., Estrada, D., Dowd, S. E., Hong, E., and Mueller, U. G. (2009).
767 Generalized antifungal activity and 454-screening of *Pseudonocardia* and *Amycolatopsis*
768 bacteria in nests of fungus-growing ants. *Proc Natl Acad Sci USA* 106, 17805–17810.
769 doi:10.1073/pnas.0904827106.

770 Shigemori, H., Komaki, H., and Yazawa, K. (1998). Brasilicardin A. A Novel Tricyclic
771 Metabolite with Potent Immunosuppressive Activity from Actinomycete *Nocardia b*
772 *rasiliensis*. *J Org Chem* 63(20), 6900-6904. doi:10.1021/jo9807114.

773 Sit, C. S., Ruzzini, A. C., Van Arnam, E. B., Ramadhar, T. R., Currie, C. R., and Clardy, J.
774 (2015). Variable genetic architectures produce virtually identical molecules in bacterial
775 symbionts of fungus-growing ants. *Proc Natl Acad Sci USA* 112, 13150–13154.
776 doi:10.1073/pnas.1515348112.

777 Valdés-Stauber, N., and Scherer, S. (1994). Isolation and characterization of Linocin M18, a
778 bacteriocin produced by *Brevibacterium linens*. *Appl Environ Microbiol* 60, 3809–3814.

779 Van Arnam, E.B., Ruzzini, A.C., Sit, C.S., Horn, H. *et al.* (2016). Selvamicin, an atypical
780 antifungal polyene from two alternative genomic contexts. *Proc Natl Acad Sci USA*
781 www.pnas.org/cgi/doi/10.1073/pnas.1613285113.
782

783 Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014).
784 Pilon: an integrated tool for comprehensive microbial variant detection and genome
785 assembly improvement. *PLoS ONE* 9(11), e112963. doi:10.1371/journal.pone.0112963.

786 Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., et al. (2015).
787 antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene
788 clusters. *Nucleic Acids Res* 43, W237–43. doi:10.1093/nar/gkv437.

789 Wright, G. D. (2007). The antibiotic resistome: the nexus of chemical and genetic diversity.
790 *Nat Rev Micro* 5, 175–186. doi:10.1038/nrmicro1614.

791 **Table 1. Bacterial strains used in this work. The names of the ten Panamanian strains**
792 **are combinations of ant colony identifiers and *Pseudonocardia* phylotypes.**

Strain	Description	GenBank Accession No.	Source
Ae150A_Ps1	*Pseudonocardia* Ps1 mutualist strain	MCIJ00000000	This study
Ae168_Ps1	*Pseudonocardia* Ps1 mutualist strain	MCIK00000000	This study
Ae263_Ps1	*Pseudonocardia* Ps1 mutualist strain	MCIL00000000	This study
Ae356_Ps1	*Pseudonocardia* Ps1 mutualist strain	MCIN00000000	This study
Ae707_Ps1	*Pseudonocardia* Ps1 mutualist strain	MCIR00000000	This study
Ae331_Ps2	*Pseudonocardia* Ps2 mutualist strain	MCIM00000000	This study
Ae406_Ps2	*Pseudonocardia* Ps2 mutualist strain	MCIO00000000	This study
Ae505_Ps2	*Pseudonocardia* Ps2 mutualist strain	MCIP00000000	This study
Ae706_Ps2	*Pseudonocardia* Ps2 mutualist strain	MCIQ00000000	This study
Ae717_Ps2	*Pseudonocardia* Ps2 mutualist strain	MCIS00000000	This study
S4	*Streptomyces* S4	CADY00000000	(Seipke *et al.,* 2011a)
KY5	*Streptomyces* KY5	N/S	(Seipke *et al.,* 2013)
Streptomyces lividans	Soil-derived *Streptomyces*	N/S	John Innes Centre, Norwich, NR4 7UH, UK
Streptomyces coelicolor M145	Soil-derived *Streptomyces*	AL645882	John Innes Centre, Norwich, NR4 7UH, UK (Bentley *et al.,* 2002)
Streptomyces venezuelae ATCC	Soil-derived *Streptomyces*	NC_018750	John Innes Centre,

10712			Norwich, NR4 7UH, UK.
KY8	KY8 *Kocuria sp.*	N/S	(Seipke *et al.,* 2013)
KY12	KY12 *Pseudomonas*	N/S	(Seipke *et al.,* 2013)
KY15	KY15 *Bacillus*	N/S	(Seipke *et al.,* 2013)
KY17	KY17 *Serratia*	N/S	(Seipke *et al.,* 2013)
KY20	KY20 *Staphylococcus*	N/S	(Seipke *et al.,* 2013)

793

794 **Table 2. BGCs encoded by the five Ps1 and five Ps2 strains sequenced in this study.**

BGCs shared by Ps1 and Ps2 strains				
Cluster ID	Ae707Ps1 BGC no.	Ae717Ps2 BGC no.	Type	Predicted product and function
A	1	13	Oligosaccharide	Unknown
B	5	9	Terpene	Carotenoid
C	7	8	T1 PKS	Nystatin-like polyene
D	8	7	Terpene	Carotenoid
E	11	11	Bacteriocin	Probably antibacterial
F	14	12	Ectoine	Osmoprotectant
BGCs shared by Ps1 strains				
Cluster ID	Ae707Ps1 BGC no.	Ae717Ps2 BGC no.	Type	Predicted product and function
G	2	-	Oligosaccharide	Unknown
H	3	-	Other	Unknown
I	4	-	NRPS	Siderophore
J	6		Bacteriocin	Possible NRPS
K	9	-	NRPS	Siderophore
L	12	-	Other	Possible NRPS
M	13	-	NRPS	Siderophore
BGCs shared by Ps2 strains				
Cluster ID	Ae707Ps1 BGC no.	Ae717Ps2 BGC no.	Type	Predicted product and function
N	-	2	Bacteriocin	Probably antibacterial
O	-	3	Other	Possible NRPS
P	-	5+6	T1 PKS-NRPS	Genomic Island possibly encoding siderophores.
Q	-	10	Terpene	Weak similarity to the terpene brasilicardin A
R	-	14	Lassopeptide	Unknown
Other BGCs				
Cluster	Strain	Cluster no.	Type	Predicted product and function

ID				
S	Ae707_Ps1	10	Other	Unknown
T	Ae717_Ps2	1	Iterative T1 PKS	Similarity to galbonolides; antifungal macrolactones
U	Ae717_Ps2	4	T1 PKS	Unknown
V	Ae717_Ps2	15	T2 PKS	Unknown
W	Ae150A_Ps1	12	Lantipeptide	Unknown

795

796 **Figure legends**

797

798 **Figure 1. Imaging of filamentous bacteria associated with *A. echinatior* ants. (A)**
799 Photographs of an *A. echinatior* callow large worker ant reveals a dusty white covering of
800 bacteria. (Image: David Nash). **(B)** Scanning electron microscopy of an *A. echinatior* worker
801 reveals filamentous bacteria dominate the microbiome (Images: Kim Findlay). Top left panel:
802 wide view of a worker, scale bar equals 200 µm. Top right panel: the laterocervical plates,
803 scale bar equals 100 µm. Bottom left panel: zoomed in view of the filamentous bacteria on
804 the laterocervical plates, scale bar equals 10 µm. Bottom right panel: high magnification
805 image of filamentous *Pseudonocardia* on the surface of an *A.echinatior* worker, scale bar
806 equals 2 µm.

807

808 **Figure 2. Resistance assays using *Pseudonocardia* strains show that both phylotypes**
809 **have antibacterial activity..** 1-5 are *Streptomyces* strains isolated originally from either soil
810 or ant associated sources. 6-10 are unicellular bacteria isolated from ant associated sources.
811 *Streptomyces* are undisturbed in growth on lawns of *Pseudonocardia* Ae707_Ps1 and
812 Ae717_Ps2 relative to the blank control, but unicellular bacteria show suppression.

813

814 **Figure 3. Secondary metabolite biosynthetic gene clusters in the genomes of**
815 ***Pseudonocardia* strains as predicted by antismash.** Clusters appear as ordered by
816 antismash of the DNA assemblies of *Pseudonocardia* genomes. Abbreviations; Non-
817 ribosomal peptide synthetase (NRPS), Polyketide synthases (PKS), ribosomally synthesized
818 and post-translationally modified peptides (RiPPs).

819

820 **Figure 4. Multigene blast outputs showing conservation of biosynthetic gene**
821 **clusters shared between Ps1 and Ps2 phylotypes.** Shared clusters include; cluster A:
822 oligosaccharide cluster **(A)**, cluster B: a terpene cluster with similarity to carotenoid BGCs
823 **(B)**, cluster C: a nystatin-like T1PKS cluster **(C)**, cluster D: a second terpene cluster with
824 similarity to carotenoid BGCs **(D)**, cluster E: a bacteriocin cluster **(E)** and cluster F: a cluster
825 predicted to make the osmoprotectant ectoine **(F)**.

826

827 **Figure 5. Multigene blast outputs showing conservation of biosynthetic gene**
828 **clusters shared between all Ps1 phylotypes.** Ps1 clusters include; cluster G & H: two
829 clusters encoding unknown products **(A & B)**, cluster I: an NRPS cluster for Ps1 siderophore
830 1 **(C)**, cluster J: a bacteriocin cluster encoding a possible NRPS **(D)**, cluster K: an NRPS
831 cluster for Ps1 siderophore 2 **(E)**, cluster L: a possible NRPS cluster encoding an unknown
832 product **(F)** and cluster M: an NRPS cluster for Ps1 siderophore 3 **(G)**.

Figure 6. Multigene blast outputs showing conservation of biosynthetic gene clusters shared between all Ps2 phylotypes. Ps2 clusters include; cluster N: a bacteriocin cluster **(A)**, cluster O: a possible NRPS cluster encoding an unknown product **(B)**, cluster P: a genomic island with T1PKS and NRPS genes **(C),** cluster Q: a terpene cluster with similarity to the brasilicardin cluster **(D)** and cluster R: a cluster for production of a possible lassopeptide **(E)**.

Figure 7. Organisation of the nystatin NPP biosynthesis gene cluster in Ae707_Ps1 and model for its biosynthesis deduced from the PKS assembly line. Abbreviations; acyl carrier protein (ACP), acyl transferase (AT), dehydratase (DH), enoylreductase (ER), ketoreductase (KR), ketosynthase (KS), *nys*-like loading KS (KS_S), thioesterase (TE). *Domains labelled with an asterisk indicate inactive domains.

Figure 8. Organisation of the nystatin-like biosynthesis gene cluster in Ae717_Ps12 and model for its biosynthesis deduced from the PKS assembly line. Abbreviations; CoA ligase (CAL), acyl carrier protein (ACP), acyl transferase (AT), dehydratase (DH), enoylreductase (ER), ketoreductase (KR), ketosynthase (KS), thioesterase (TE). *Domains labelled with an asterisk indicate inactive domains.

Figure 01.TIF

Figure 1.

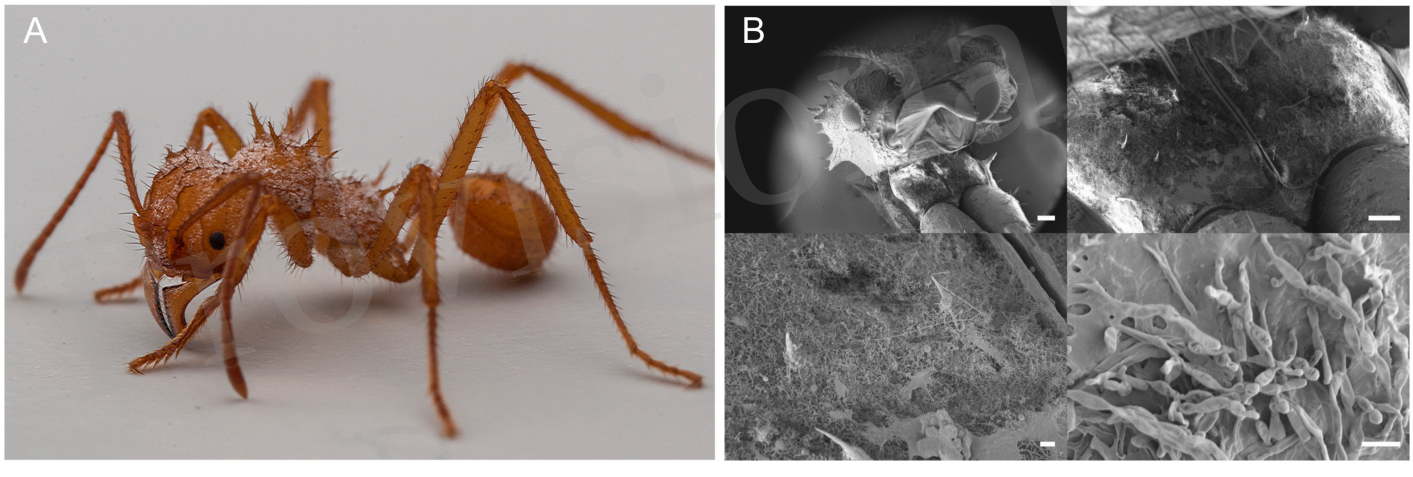
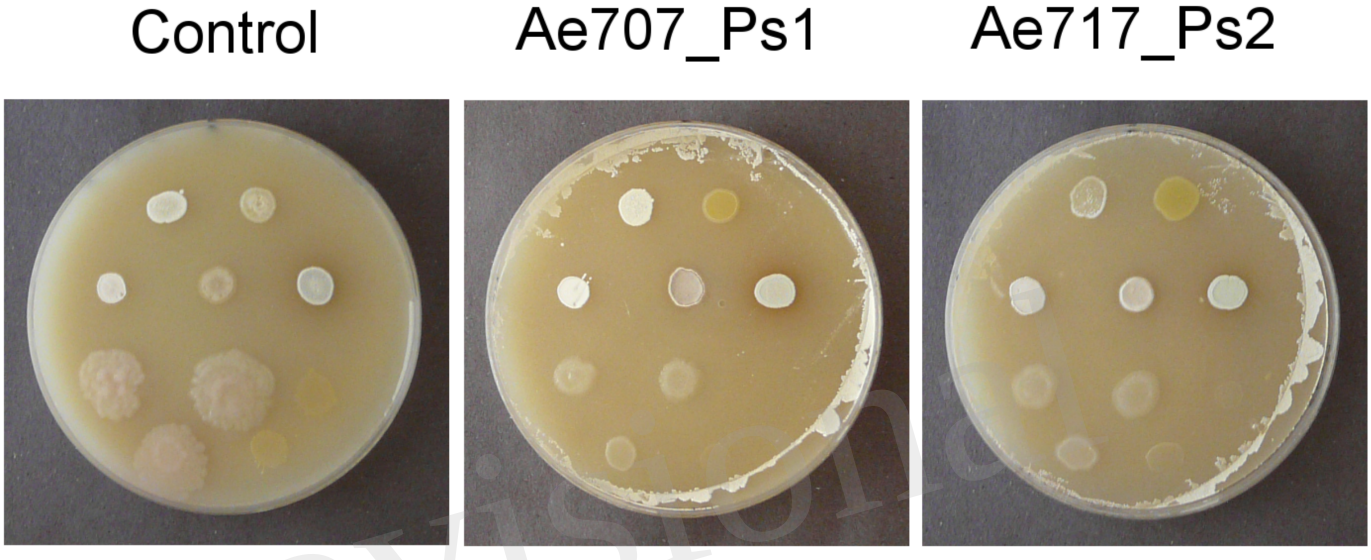
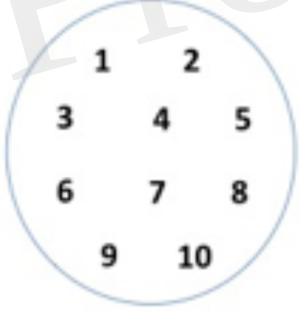


Figure 02.TIF

Figure 2.





1 = S4
2 = KY5
3 = *S. lividans*
4 = *S. coelicolor*
5 = *S. venezuelae*
6 = KY8 *Kocuria*
7 = K12 *Pseudomonas*
8 = KY15 *Bacillus*
9 = KY17 *Serratia*
10 = KY20 *Staphylococcus*

Figure 03.TIF

Figure 3.

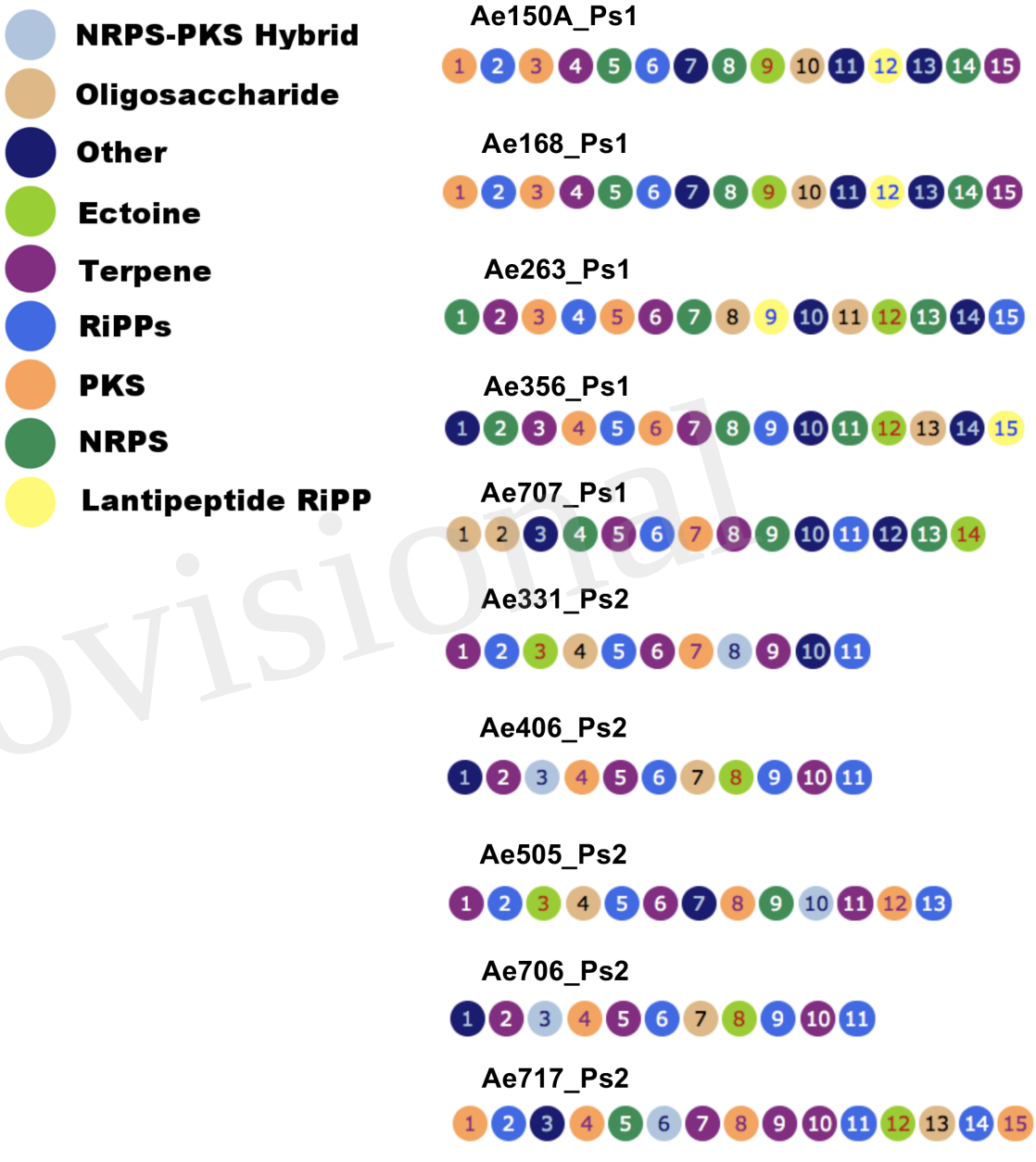


Figure 4

Figure 04.TIF

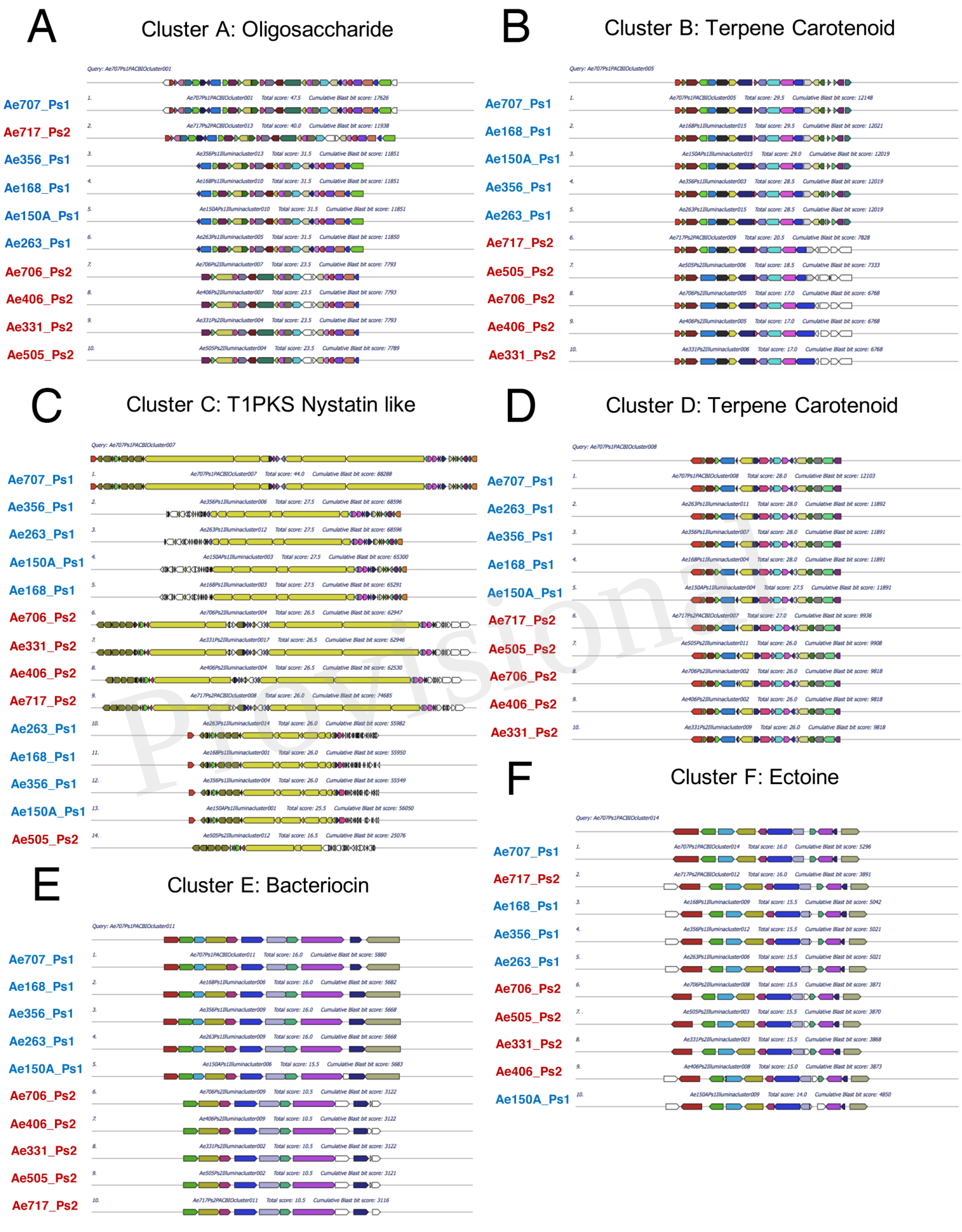


Figure 05.TIF

Figure 5

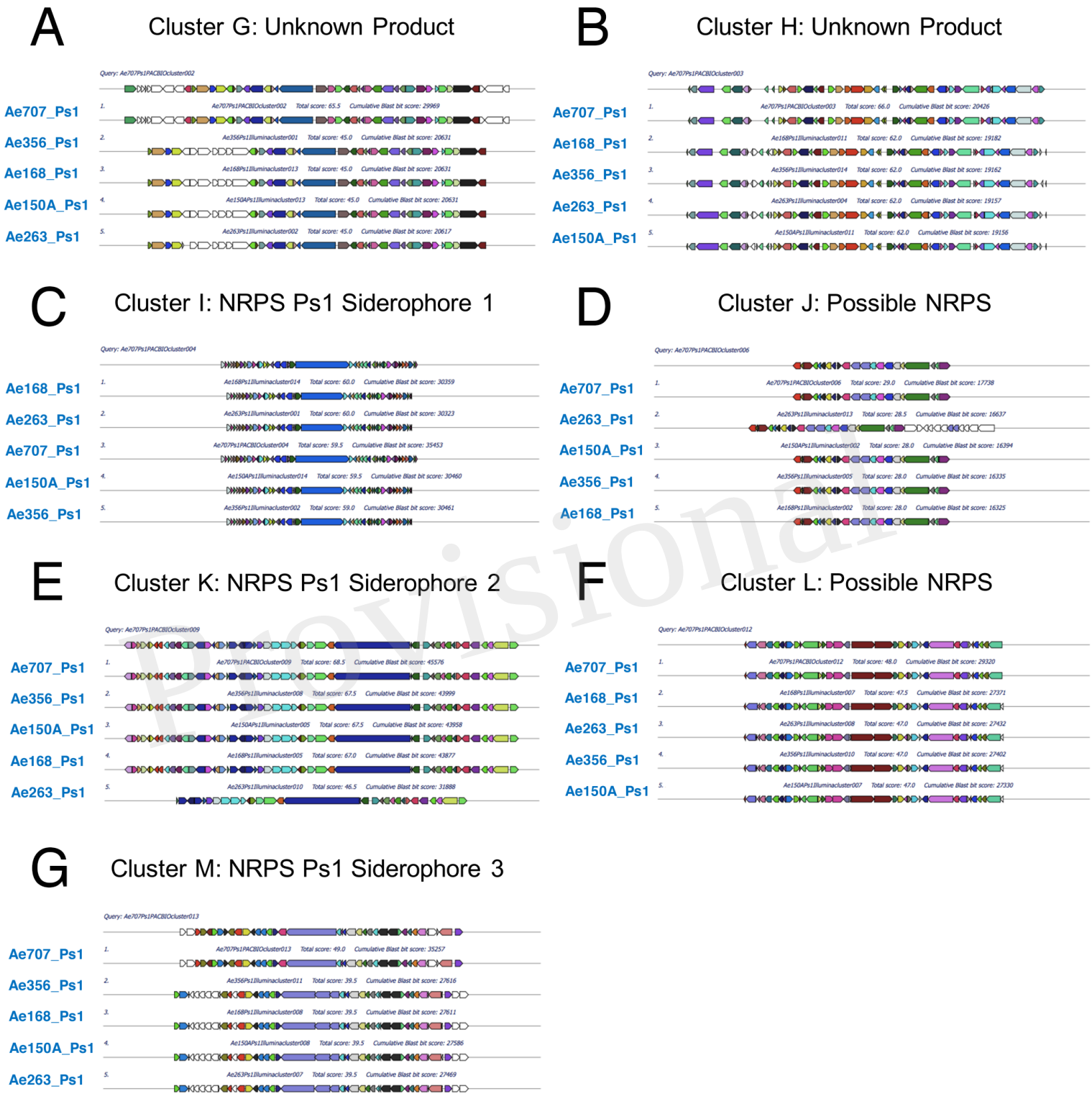


Figure 06.TIF

Figure 6

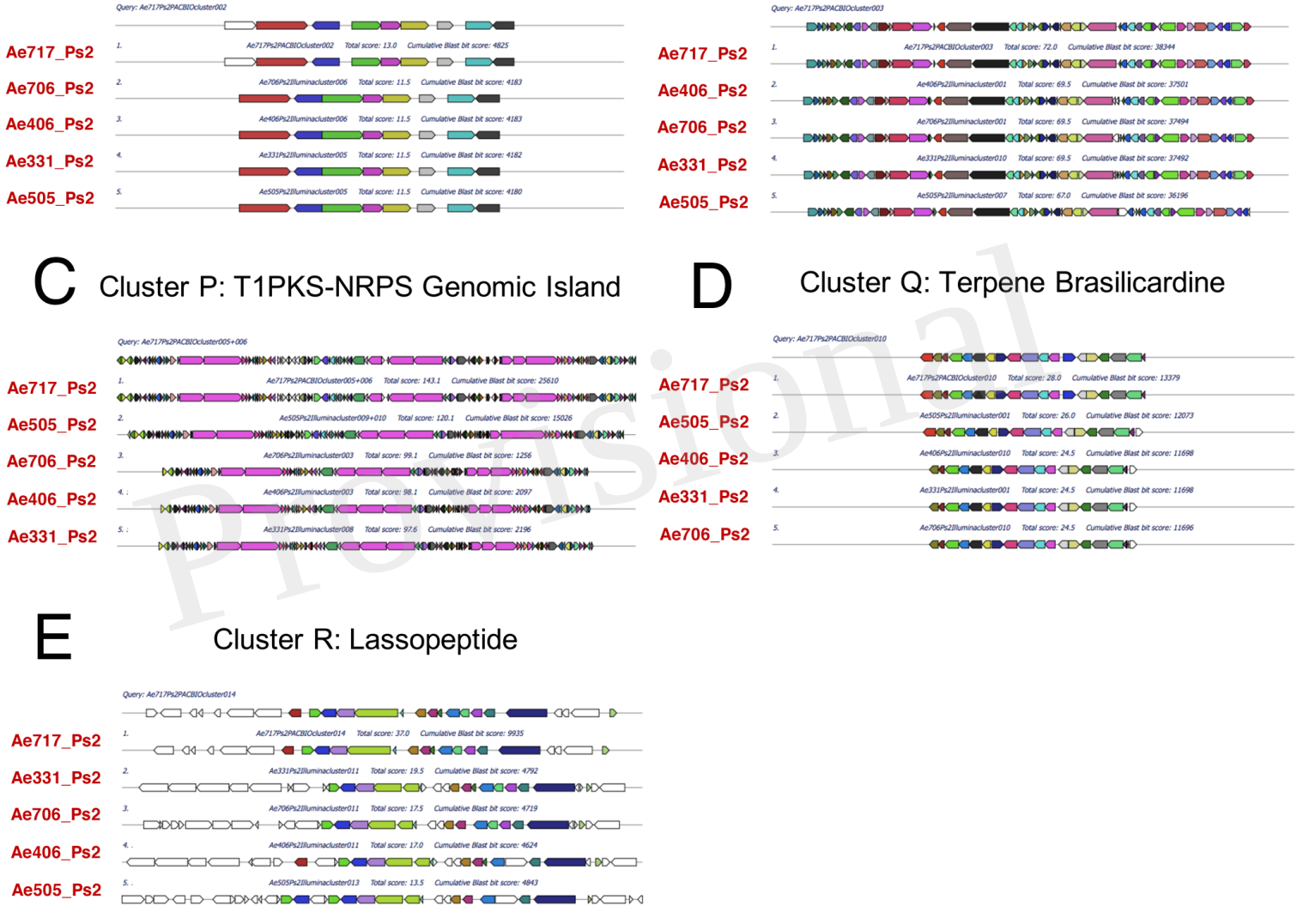


Figure 07.TIF

Figure 7

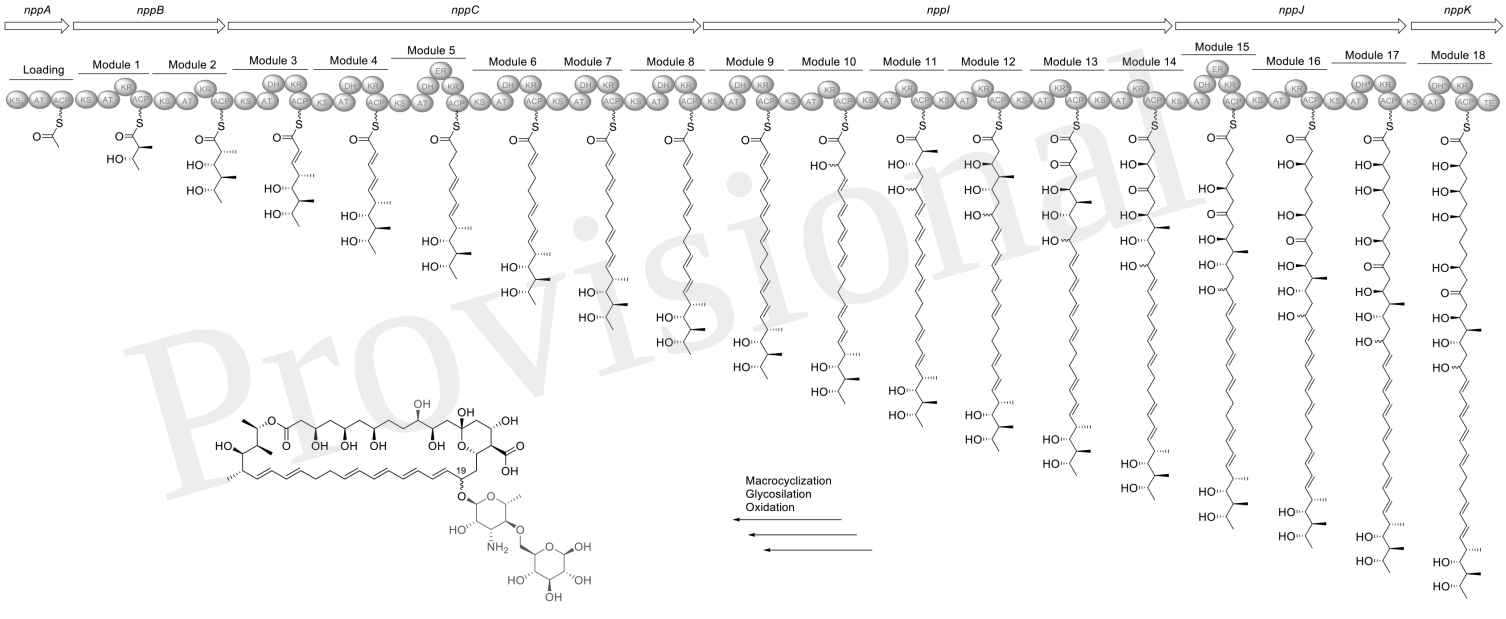


Figure 08.TIF

Figure 8