*Article*

# High-throughput Sequencing Analysis of the Actinobacterial Spatial Diversity in Moonmilk Deposits

Marta Maciejewska[1], Magdalena Całusińska[2], Luc Cornet[3], Delphine Adam[1], Igor S. Pessi[1], Sandrine Malchair[4], Philippe Delfosse[2], Denis Baurain[3], Hazel H. Barton[5], Monique Carnol[4], and Sébastien Rigali[1*]

[1] InBioS - Centre for Protein Engineering, Institut de Chimie B6a, University of Liège, Liège, Belgium

[2]Environmental Research and Innovation Department, Luxembourg Institute of Science and Technology, Belvaux, Luxembourg

[3] InBioS - PhytoSYSTEMS, Eukaryotic Phylogenomics, University of Liège, Liège, Belgium

[4] InBioS - Plant and Microbial Ecology, Botany B22, University of Liège, Liège, Belgium

[5] Department of Biology, University of Akron, Akron, Ohio, United States of America

[*] Correspondence: srigali@ulg.ac.be; Tel.: + 32 4 366 98 30

**Abstract:** Moonmilk are cave carbonate deposits that host a rich microbiome including antibiotic-producing Actinobacteria making these speleothems appealing for bioprospecting. Here we investigated the taxonomic profile of the actinobacterial community of three moonmilk deposits of the cave "Grotte des Collemboles" via high-throughput sequencing of 16S rRNA amplicons. Actinobacteria was the most common phylum after Proteobacteria, ranging from 9 to 23% of the total bacterial population. Next to actinobacterial OTUs attributed to uncultured organisms at the genus level (~44%), we identified 47 actinobacterial genera with *Rhodoccocus* (4 OTUs, 17%) and *Pseudonocardia* (9 OTUs, ~16%) as the most abundant in terms of absolute number of sequences. Streptomycetes presented the highest diversity (19 OTUs, 3%), with most of OTUs unlinked to the culturable *Streptomyces* strains previously isolated from the same deposits. 43% of OTUs were shared between the three studied collection points while 34% were exclusive to one deposit indicating that distinct speleothems host their own population despite their nearby localization. This important spatial diversity suggests that prospecting within different moonmilk deposits should result in the isolation of unique and novel Actinobacteria. These speleothems also host a wide range of non-streptomycetes antibiotic-producing genera, and should therefore be subjected to methodologies for isolating rare Actinobacteria.

**Keywords:** antibiotics; geomicrobiology; Illumina sequencing; microbiome diversity; *Streptomyces;* Actinobacteria

## 1. Introduction

Molecular approaches evaluating microbial communities in caves have revealed a level of diversity greater than initially expected [1]. Microorganisms have been found to inhabit virtually all subterranean niches, including cave walls, ceilings, speleothems, soils, sediments, pools and aquifers

Peer-reviewed version available at Antibiotics 2018, 7, 27; doi:10.3390/antibiotics7020027



[2]. Cave bacteria often represent novel taxonomic groups [3–7], which are frequently more closely related to other cave-derived bacterial lineages than to the microbiota of other environments [8–10].

Among cave speleothems, moonmilk draws a particular scientific attention due to its distinctive crystal morphology. The origins of various moonmilk crystalline habits, including monocrystalline rods, polycrystalline chains and nano-fibers, are tentatively attributed to the moonmilk indigenous microbial population [11]. Among a moonmilk microbiome comprising Archaea, Fungi, and Bacteria [9,10,12–19], the indigenous filamentous fungi [20] and Actinobacteria [11,21] are believed to mediate moonmilk genesis with cell surfaces promoting $CaCO_3$ deposition [11,20,21]. Actinobacteria were additionally reported to be metabolically capable of inducing favorable conditions for $CaCO_3$ precipitation or even directly precipitating carbonate minerals [12,21]. Members of the phylum Actinobacteria are routinely found in this speleothem [9,10,12–14,18,19], as well as in the other subterranean deposits within limestone caves [3,8,22,23], volcanic caves [24–26] and ice caves [27]. The broad distribution of Actinobacteria in the subsurface systems stimulates investigation to understand the factors driving their existence in mainly inorganic and highly oligotrophic environments and the processes which enable them to mediate speleogenesis. The successful adaptation of Actinobacteria to a wide range of environments could probably be a consequence of their broad-spectrum metabolism that includes prolific secreted hydrolytic systems capable of generating nutrient sources from various substrates, along with their extraordinary faculty to produce specialized metabolites (metal-chelators, antimicrobials, hormones, etc.) [28].

As recently reported, moonmilk Actinobacteria represent novel microorganisms, a discovery that opens great avenues for bioprospecting of novel drugs [6,10,18]. Rooney et al. (2010) showed that spatially separated moonmilk speleothems in Ballynamintra Cave are inhabited by taxonomically distinct fungal and bacterial communities. Instead, in our attempt to isolate moonmilk-dwelling Actinobacteria for assessing their potential at participating in the genesis of these speleothems [21] and at producing antimicrobial compounds [10], we only recovered members of the genus *Streptomyces*. Such a dominance of streptomycetes was rather unexpected according to other moonmilk microbial diversity studies performed through culture-dependent [10,12,13,18] and culture-independent approaches using clone libraries [9], Denaturing Gradient Gel Electrophoresis (DGGE) fingerprinting [14,16,17], Automated Ribosomal Intergenic Spacer Analysis (ARISA) [13], and, more recently, high-throughput sequencing (HTS) [19]. The actinobacterial genera identified in those studies included *Rhodococcus*, *Pseudonocardia*, *Propionibacterium*, *Nocardia*, *Amycolatopsis*, *Saccharothrix*, *Geodermatophilus*, *Mycobacterium*, *Aeromicrobium*, *Kribella*, *Nocardioides*, *Actinomycetospora*, *Nonomuraea*, *Euzebya*, *Rubrobacter* and *Arthrobacter*, in addition to *Streptomyces*. Nonetheless, the diversity of the moonmilk actinobacterial microbiome still remains largely unknown and, beyond evaluating '*what and how much have we missed in our culture-dependent bioprospecting approach*' [10], a major important question that arises is '*to which extent moonmilk-dwelling Actinobacteria are different between the moonmilk deposits within a single or in different caves?*'.

In this work we carried out a comparative (HTS) of 16S (SSU) rRNA gene from DNA extracted from spatially separated moonmilk deposits within the same cave, "Grotte des Collemboles" (Springtails' Cave) in Comblain-au-Pont, Belgium (Figure S1), in order to draw a detailed taxonomic picture of the intra-phylum diversity. Identifying the presence of rare Actinobacteria and unveiling to which degree they exhibit a spatial variability would help

Peer-reviewed version available at Antibiotics 2018, 7, 27; doi:10.3390/antibiotics7020027



82 determining whether it is worth prospecting from different moonmilk deposits to isolate unique and
83 novel natural compound producers.
84
85

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2018 doi:10.20944/preprints201802.0090.v1

Peer-reviewed version available at Antibiotics 2018, 7, 27; doi:10.3390/antibiotics7020027



## 2. Results

### 2.1. Actinobacterial abundance within the whole moonmilk bacterial microbiome

Libraries spanning the V4-V6 variable regions of the 16S rRNA gene using universal bacterial primers were used to assess the proportion of Actinobacteria in comparison to the whole bacterial community of three moonmilk deposits of the cave "Grotte des Collemboles" (Table S1a). The observed bacterial communities differed in species richness, evenness, and diversity between the three sampling points (Table 1). Phylotype richness (total number of OTUs per site) was the highest in COL4 (1863 OTUs), followed by COL1 and COL3, with 1332 and 1161 OTUs, respectively (Table 1, Figure 1a). Across the three sampling points, we found a total of 2301 different OTUs amongst which 710 (31%) were common to all the deposits (Figure 1a). Interestingly, pair-wise comparison revealed highly similar percentages (~31.7 ± 0.53%) of shared bacterial OTUs between moonmilk deposits (Table 2, Figure 1a). A total of 956 OTUs (42%) were found to be exclusive to one sampling site, with COL4 having the highest number of unique bacterial phylotypes (584 OTUs), along with the most diverse bacterial population, as reflected by the highest diversity indices (Table 1, Figure 1a).

**Table 1.** Richness, specificity, diversity, and evenness of the bacterial and actinobacterial communities in the three moonmilk deposits of the "Grotte des Collemboles".

| | Site | Total OTUs (richness) | Unique OTUs (specificity) | Inverse Simpson index (diversity) | Simpson index (evenness) |
|---|---|---|---|---|---|
| **Bacteria** | **COL1** | 1332 | 238 (17.9%) | 13.23 | 0.01 |
| | **COL3** | 1161 | 134 (11.6%) | 58.94 | 0.05 |
| | **COL4** | 1863 | 584 (31.3%) | 155.31 | 0.08 |
| **Actinobacteria** | **COL1** | 150 | 14 (9.3%) | 6.21 | 0.04 |
| | **COL3** | 147 | 15 (10.2%) | 7.74 | 0.05 |
| | **COL4** | 211 | 54 (25.6%) | 24.13 | 0.11 |

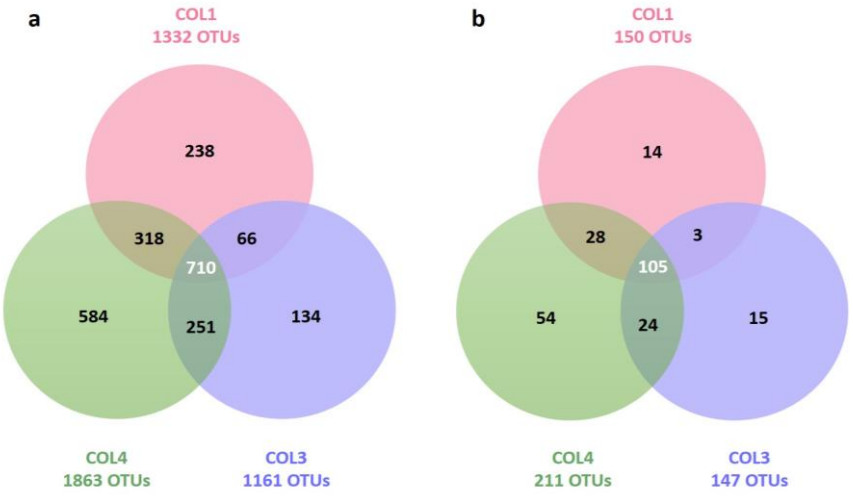


**Figure 1.** Venn diagrams showing the numbers of shared and unique bacterial (a) and actinobacterial (b) OTUs between the three moonmilk sampling points (COL1, COL3, COL4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2018 doi:10.20944/preprints201802.0090.v1

Peer-reviewed version available at Antibiotics 2018, 7, 27; doi:10.3390/antibiotics7020027



**Table 2.** Pair-wise comparisons of shared OTUs between the moonmilk deposits.

| | COL1 and COL3 | COL1 and COL4 | COL4 and COL3 |
|---|---|---|---|
| **Bacteria** | 776/2493 (31.1%) | 1028/3195 (32.2%) | 961/3024 (31.8%) |
| **Actinobacteria** | 108/297 (36.4%) | 133/361 (36.9%) | 129/358 (36.0%) |

Bacterial OTUs were grouped into 21 phyla and 18 candidate phyla (Table S2, Figure 2). Actinobacteria represented 9, 23, and 10% of the total bacterial population in COL1, COL3, and COL4, respectively (Table S2, Figure 2a). In terms of abundance, they were the most common phylum after Proteobacteria, which accounted for 52, 34 and 30% of the total community in COL1, COL3, and COL4, respectively (Table S2, Figure 2a). The other major phyla of the moonmilk microbiome included Acidobacteria, Nitrospirae, Chloroflexi, Gemmatimonadetes, Planctomycetes, Latescibacteria, Verrucomicrobia, Zixibacteria, Armatimonadetes, Bacteroidetes, and Parcubacteria (Table S2, Figure 2a). Together these phyla constituted 93.4%, 94.7%, and 91.5% of the total community in COL1, COL3, and COL4, respectively (Table S2, Figure 2a). The remaining phyla (with a relative abundance < 1%) were pooled as 'other' (Figure 2a), and included most of the candidate divisions identified in this study (Figure 2b). Sequences which could not be affiliated to any bacterial phylum accounted for 4, 3 and 5% of the sequences in COL1, COL3, and COL4, respectively (Table S2). Some fraction of the moonmilk microbial diversity still remains to be discovered for all the three sampling sites, as rarefaction curves did not reach a plateau (Figure S2a).

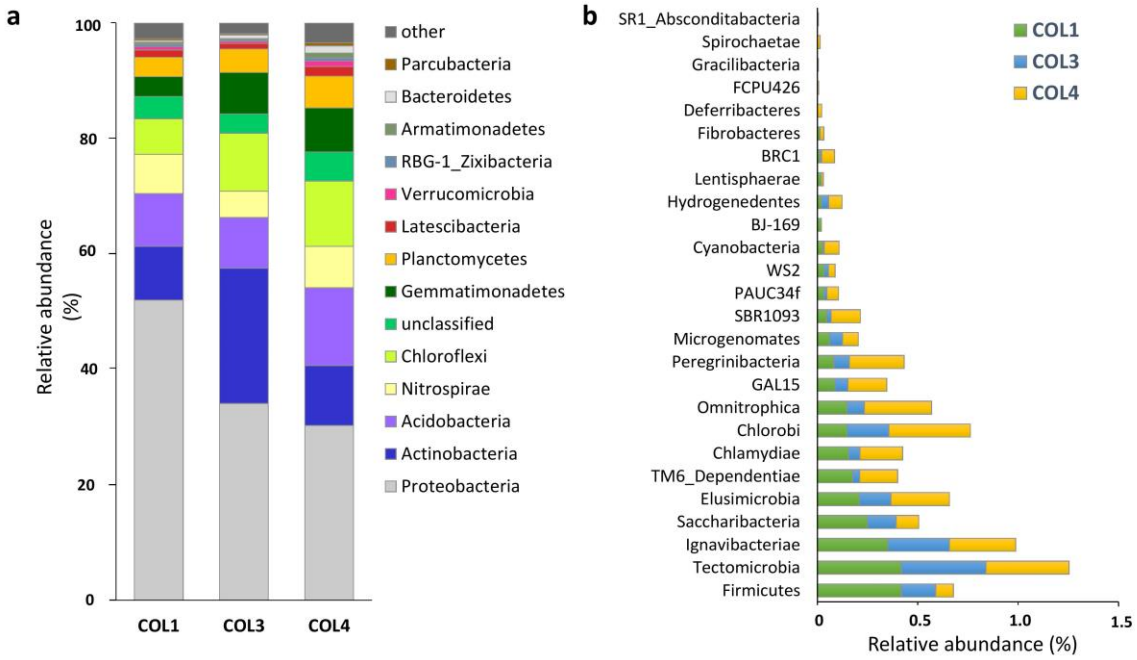


**Figure 2.** Taxonomic profiles of the moonmilk-associated microbiome at the phylum level across the three moonmilk sampling points (COL1, COL3, COL4). The main phyla of the microbiome are presented on the left (a), while the pattern of low-abundance taxa, named as 'other' (with relative abundance <1%) is displayed on the right (b).

### 2.2. *Actinobacterial diversity in moonmilk deposits*

Evaluation of the actinobacterial profile was performed with libraries spanning the V6-V7 variable regions of 16S rRNA gene and using modified Actinobacteria-specific primers (Table S1a). The specificity of the primers was confirmed by the detection of only 1, 0.2, and 2% of non-actinobacterial sequences in COL1, COL3, and COL4, respectively (Figure S3). In contrast to the bacterial dataset, the diversity of Actinobacteria appeared to be exhaustively sampled with the phylum-specific primers (Figure S2b).

The diversity indices for Actinobacteria showed the same trends as the ones observed for the whole Bacteria domain, *i.e.*, evenness and diversity were the highest in COL4, followed by COL3 and COL1 (Table 1). Phylotype richness was the highest in COL4 with 211 OTUs, followed by COL1 and COL3 with 150 OTUs and 147 OTUs, respectively (Figure 1b and Table 1). Among the 243 different OTUs, 105 OTUs (43%) were found in all the three studied moonmilk deposits (Figure 1b). Hence, the moonmilk-associated actinobacterial community appeared to be more conserved than the moonmilk-associated bacterial population (31%, Figure 1a). If we also include OTUs shared at least between two sampling points, the level of conservation rises to 66% of OTUs for Actinobacteria, and 58% for Bacteria. Still, 34% of the 243 OTUs (14, 15, and 54 OTUs in COL1, COL3, and COL4, respectively) remained specific to a moonmilk deposit, despite the close localization of collection points within the studied cave (Figure 1b). COL4 was characterized not only with the highest number of unique phylotypes (54 OTUs) (Figure 1b), but also with the most diverse population, as revealed by diversity indices (Table 1). As observed for the bacterial dataset, pair-wise comparisons showed highly similar percentages (~36.4 ± 0.41%) of shared actinobacterial OTUs between moonmilk deposits (Table 2).

Taxonomic assignment of actinobacterial OTUs revealed the presence of two major classes - Acidimicrobiia and Actinobacteria, next to low-abundant Thermoleophilia class (Table S3, Figure 3a). Acidimicrobiia was represented by one single order, the Acidimicrobiales, which dominated sample COL4 constituting 55.3% of the population (Table S3, Figure 3a). The Acidimicrobiales order consisted of two families, *i.e.*, Acidimicrobiaceae and Iamiaceae (Table S3). The Actinobacteria class was represented by 15 orders, with Corynebacteriales dominating in COL1, and Pseudonocardiales in COL3 and COL4 (Table S3, Figure 3b). The most abundant families among the Actinobacteria class were Pseudonocardiaceae, Nocardiaceae, and Streptomycetaceae (Table S3). The proportion of unclassified and uncultured sequences at the family level ranged from 9% in COL3, to 25% in COL1, and 53% in COL4 (Table S3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2018 doi:10.20944/preprints201802.0090.v1

Peer-reviewed version available at Antibiotics 2018, 7, 27; doi:10.3390/antibiotics7020027



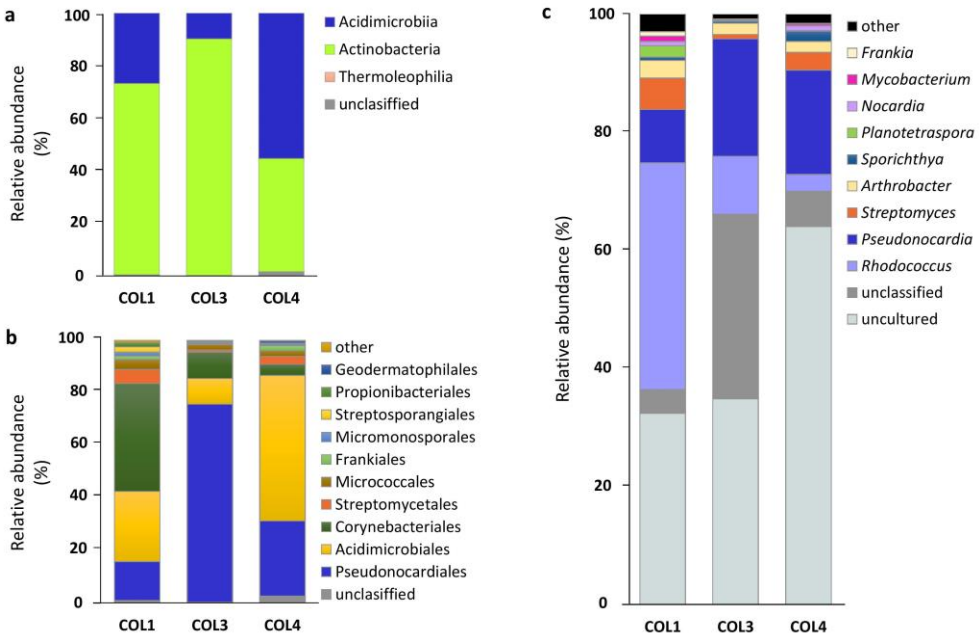


**Figure 3.** Taxonomic profiles of moonmilk-associated Actinobacteria at different taxonomic levels - (a) class; (b) order; (c) family - observed across the three moonmilk-sampling points (COL1, COL3, COL4). 'Other' includes orders and families with relative abundance <1%.

Among 28 families, a total of 47 genera were identified across the investigated samples (Table S3 and Table 3), with 35 genera identified for the first time in moonmilk (Table 3). COL1 was dominated by *Rhodococcus* (38.37%), while uncultured and unclassified Actinobacteria were the most abundant in COL3 and COL4 (Table 3). When only known genera were taken into account, *Pseudonocardia* prevailed in those samples, accounting for 20% and 18% of the population in COL3 and COL4, respectively (Table 3). Other genera, which constituted more than 1% of the population in at least one moonmilk deposit, included *Streptomyces, Arthrobacter, Sporichthya, Planotetraspora, Nocardia, Mycobacterium,* and *Frankia* (Table 3). While accounting in average for only 3% of the actinobacterial community, streptomycetes displayed the highest diversity with 19 OTUs identified across the three moonmilk deposits (Table 3).

Some taxa showed important differences in their relative abundance between investigated samples, particularly *Rhodococcus*, which was approximately four and fourteen times more abundant in COL1 than in COL3 and COL4, respectively (Table 3). The *Streptomyces* genus represented only 0.8% of the population in COL3, while it was detected at the level of 5.3% and 3% in the COL1 and COL4, respectively (Table 3). An important discrepancy in the relative abundance between speleothems was also observed for the genera *Planotetraspora*, *Mycobacterium,* and *Frankia,* whereas some taxa (e.g. *Pseudoclavibacter*, *Lentzea*, *Propionibacterium*) were exclusively found in a single sampling site (Table 3).

189 **Table 3**. Actinobacterial genera pattern in moonmilk deposits of the "Grotte des Collemboles"
190 based on 16S rRNA amplicon libraries.
191

| | COL1 | | | COL3 | | | COL4 | | | TOTAL | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Seq. | % | OTUs | Seq. | % | OTUs | Seq. | % | OTUs | Seq. | Av. % | Diff. OTUs |
| **Uncultured** | 46346 | 32.38 | 65 | 58903 | 34.80 | 63 | 92444 | 63.98 | 90 | 197693 | 43.7 | 96 |
| ***Rhodococcus*** | 54920 | 38.37 | 3 | 16636 | 9.83 | 3 | 4102 | 2.84 | 3 | 75658 | 17.0 | 4 |
| *Pseudonocardia* | 12913 | 9.02 | 7 | 33640 | 19.87 | 8 | 25545 | 17.68 | 9 | 72098 | 15.5 | 9 |
| **Unclassified** | 5762 | 4.03 | 18 | 52908 | 31.26 | 20 | 8668 | 6.00 | 26 | 67338 | 13.8 | 32 |
| *Streptomyces* | 7628 | 5.33 | 10 | 1292 | 0.76 | 11 | 4347 | 3.01 | 18 | 13267 | 3.0 | 19 |
| ***Arthrobacter*** | 4323 | 3.02 | 4 | 3214 | 1.90 | 4 | 2649 | 1.83 | 5 | 10186 | 2.3 | 5 |
| *Sporichthya* | 771 | 0.54 | 1 | 556 | 0.33 | 2 | 2515 | 1.74 | 2 | 3842 | 0.9 | 2 |
| *Planotetraspora* | 2769 | 1.93 | 1 | 294 | 0.17 | 1 | 113 | 0.08 | 1 | 3176 | 0.7 | 1 |
| ***Nocardia*** | 1023 | 0.71 | 2 | 157 | 0.09 | 2 | 1212 | 0.84 | 2 | 2392 | 0.5 | 2 |
| *Mycobacterium* | 1299 | 0.91 | 2 | 127 | 0.08 | 2 | 440 | 0.30 | 3 | 1866 | 0.4 | 3 |
| *Frankia* | 1067 | 0.75 | 2 | 77 | 0.05 | 1 | 126 | 0.09 | 3 | 1270 | 0.3 | 3 |
| *Luedemannella* | 555 | 0.39 | 2 | 212 | 0.13 | 2 | 276 | 0.19 | 2 | 1043 | 0.2 | 2 |
| *Longispora* | 371 | 0.26 | 1 | 374 | 0.22 | 1 | 110 | 0.08 | 1 | 855 | 0.2 | 1 |
| *Agromyces* | 310 | 0.22 | 2 | 136 | 0.08 | 2 | 355 | 0.25 | 2 | 801 | 0.2 | 2 |
| *Actinoplanes* | 591 | 0.41 | 2 | 85 | 0.05 | 2 | 111 | 0.08 | 2 | 787 | 0.2 | 2 |
| *Nakamurella* | 416 | 0.29 | 1 | 114 | 0.07 | 1 | 101 | 0.07 | 1 | 631 | 0.1 | 1 |
| *Nocardioides* | 360 | 0.25 | 4 | 52 | 0.03 | 3 | 158 | 0.11 | 8 | 570 | 0.1 | 9 |
| *Geodermatophilus* | 78 | 0.05 | 1 | 70 | 0.04 | 2 | 374 | 0.26 | 2 | 522 | 0.1 | 2 |
| *Catellatospora* | 95 | 0.07 | 3 | 70 | 0.04 | 2 | 141 | 0.10 | 4 | 306 | 0.07 | 4 |
| *Kribbella* | 120 | 0.08 | 1 | 36 | 0.02 | 1 | 135 | 0.09 | 2 | 291 | 0.07 | 2 |
| *Kocuria* | 261 | 0.18 | 1 | 23 | 0.01 | 1 | - | - | - | 284 | 0.1 | 2 |
| *Actinomyces* | 247 | 0.17 | 5 | 1 | 0.001 | 1 | 6 | 0.004 | 1 | 254 | 0.06 | 5 |
| *Corynebacterium* | 151 | 0.11 | 1 | 32 | 0.02 | 3 | 10 | 0.01 | 2 | 193 | 0.04 | 5 |
| *Rhizocola* | - | - | - | 27 | 0.02 | 1 | 145 | 0.10 | 1 | 172 | 0.06 | 1 |
| *Microbacterium* | 108 | 0.08 | 1 | 48 | 0.03 | 1 | 5 | 0.003 | 1 | 161 | 0.04 | 1 |
| *Iamia* | 107 | 0.07 | 1 | - | - | - | 31 | 0.02 | 1 | 138 | 0.05 | 1 |
| *Pseudoclavibacter* | 138 | 0.10 | 1 | - | - | - | - | - | - | 138 | 0.1 | 1 |
| *Lentzea* | - | - | - | - | - | - | 111 | 0.08 | 1 | 111 | 0.08 | 1 |
| *Aeromicrobium* | 73 | 0.05 | 1 | - | - | - | 28 | 0.02 | 2 | 101 | 0.04 | 2 |
| ***Amycolatopsis*** | 86 | 0.06 | 1 | - | - | - | 5 | 0.003 | 1 | 91 | 0.03 | 2 |
| *Cryptosporangium* | - | - | - | 71 | 0.04 | 1 | 10 | 0.01 | 1 | 81 | 0.02 | 1 |
| *Glycomyces* | 35 | 0.02 | 1 | - | - | - | 45 | 0.03 | 1 | 80 | 0.03 | 1 |
| *Streptosporangium* | 61 | 0.04 | 1 | - | - | - | 19 | 0.01 | 1 | 80 | 0.03 | 1 |
| *Smaragdicoccus* | 43 | 0.03 | 1 | - | - | - | 34 | 0.02 | 1 | 77 | 0.03 | 1 |
| ***Propionibacterium*** | 57 | 0.04 | 1 | - | - | - | - | - | 0 | 57 | 0.04 | 1 |
| *Kineosporia* | 44 | 0.03 | 1 | - | - | - | 8 | 0.01 | 1 | 52 | 0.02 | 1 |
| *Jatrophihabitans* | - | - | - | 21 | 0.01 | 1 | 24 | 0.02 | 1 | 45 | 0.01 | 1 |
| *Promicromonospora* | - | - | - | 22 | 0.01 | 1 | 21 | 0.01 | 1 | 43 | 0.01 | 2 |
| *Millisia* | - | - | - | 22 | 0.01 | 1 | 6 | 0.004 | 1 | 28 | 0.01 | 1 |
| *Rothia* | 2 | 0.001 | 1 | 13 | 0.01 | 1 | 7 | 0.005 | 1 | 22 | 0.005 | 1 |
| *Tessaracoccus* | - | - | - | 17 | 0.01 | 1 | - | - | - | 17 | 0.01 | 1 |
| *Acidothermus* | - | - | - | - | - | - | 16 | 0.01 | 1 | 16 | 0.01 | 1 |
| *Marmoricola* | - | - | - | - | - | - | 14 | 0.01 | 2 | 14 | 0.01 | 2 |
| *Dermacoccus* | - | - | - | - | - | - | 11 | 0.008 | 1 | 11 | 0.01 | 1 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***Ponticoccus*** | - | - | - | 8 | 0.005 | 1 | - | - | - | 8 | 0.005 | 1 |
| ***Stackebrandtia*** | - | - | - | - | - | - | 8 | 0.006 | 1 | 8 | 0.01 | 1 |
| ***Umezawaea*** | - | - | - | - | - | - | 2 | 0.001 | 1 | 2 | 0.001 | 1 |
| ***Actinospica*** | - | - | - | - | - | - | 1 | 0.001 | 1 | 1 | 0.001 | 1 |
| ***Propionimicrobium*** | - | - | - | - | - | - | 1 | 0.001 | 1 | 1 | 0.001 | 1 |

For each taxon, the number of obtained sequences (Seq.) and their relative abundance (%), together with the number of OTUs, are given. The total number of sequences, average relative abundance and total number of different OTUs obtained per genus are shown in the last three columns. Taxa in red were reported for the first time in moonmilk deposits in this study. Taxa in blue were detected in moonmilk deposits in this work and in the HTS-based study of Dhami et al [19]. Taxa in black and underlined represent the ones which were also detected in other moonmilk microbial diversity studies [12–15,18]. Cases filled in grey highlight the most abundant genera in each studied sampling point. Abbreviations: Seq. - number of sequences identified.

### *2.3. Analysis of the most abundant actinobacterial OTUs*

In order to obtain more information about the most dominant moonmilk-dwelling Actinobacteria a detailed analysis was conducted for the 41 most abundant OTUs (~17% of all the OTUs) accounting together for 90% (413,739 out of 456,878) of the sequences obtained via our HTS approach (Table 4). Out of the subset of 41 OTUs, 16 phylotypes belonged to the class Acidimicrobiia, most of them being uncultured at the family level, and the remaining 25 OTUs belonged to the class Actinobacteria (Table 4). In the latter case, all the OTUs were associated with major families previously identified in moonmilk deposits, including Pseudonocardiaceae, Propionibacteriaceae, Micrococcaceae, Nocardiaceae, Streptomycetaceae, and Streptosporangiaceae (Table 4). Only 16 OTUs could be classified at the genus level and were affiliated to genera *Rhodococcus, Pseudonocardia, Arthrobacter, Sporichthya, Streptomyces, Planotetraspora* and *Nocardia* (Table 4).

Taking into account the spatial differences in terms of the most abundant taxa across the cave, COL1 was highly dominated by OTU1, affiliated to the genus *Rhodococcus* and accounting for 38% of the total population in this speleothem (Table 4). This phylotype highly outnumbered other two *Rhodococcus* OTUs detected in COL1 (Table 3), which together constituted only 0.1% (data not shown). The predominant phylotypes identified in speleothems COL3 and COL4 were OTU2 representing an unclassified Pseudonocardiaceae in COL3 (29%), and OTU4 representing uncultured bacterium from Acidimicrobiia class in COL4 (11%) (Table 4). Among known genera, *Rhodococcus* (OTU1, 10%) was also prevailing in COL3, while *Pseudonocardia* (OTU262, 6%) was found to be the most abundant in COL4 (Table 4).

40 out of 41 OTUs were present in all the three studied moonmilk deposits, often with an extreme variation in terms of their relative abundance across the different collection points. This is well demonstrated by OTU2 (Pseudonocardiaceae, unclassified at the genus level), which largely dominated the actinobacterial community in COL3 (29%) while only representing 0.3% of the actinobacterial microbiome in COL1 (Table 4).

**Table 4**. The relative abundance (%) and taxonomy assignment of the most abundant actinobacterial OTUs found across moonmilk samples within the "Grotte des Collemboles".

| OTU | COL1 | COL3 | COL4 | Av. % | Class | Family | Genus |
|---|---|---|---|---|---|---|---|
| **OTU1** | 38.28 | 9.68 | 2.74 | 16.90 | Actinobacteria | Nocardiaceae | *Rhodococcus* |
| **OTU2** | 0.31 | 28.89 | 2.13 | 10.44 | Actinobacteria | Pseudonocardiaceae | unclassified |
| **OTU8** | 0.75 | 13.96 | 0.58 | 5.10 | Actinobacteria | Pseudonocardiaceae | uncultured |
| **OTU4** | 3.35 | 1.13 | 11.47 | 5.32 | Acidimicrobiia | uncultured | uncultured |
| **OTU3** | 0.52 | 7.80 | 4.87 | 4.40 | Actinobacteria | Pseudonocardiaceae | *Pseudonocardia* |
| **OTU262** | 3.45 | 4.12 | 5.89 | 4.49 | Actinobacteria | Pseudonocardiaceae | *Pseudonocardia* |
| **OTU6** | 7.53 | 1.00 | 3.98 | 4.17 | Acidimicrobiia | uncultured | uncultured |
| **OTU12** | 2.97 | 6.61 | 1.82 | 3.80 | Actinobacteria | Pseudonocardiaceae | uncultured |
| **OTU5** | 3.53 | 1.70 | 5.87 | 3.70 | Acidimicrobiia | uncultured | uncultured |
| **OTU13** | 0.46 | 4.91 | 3.65 | 3.01 | Actinobacteria | Pseudonocardiaceae | *Pseudonocardia* |
| **OTU98** | 0.99 | 1.07 | 6.94 | 3.00 | Acidimicrobiia | uncultured | uncultured |
| **OTU203** | 1.73 | 0.31 | 5.92 | 2.66 | Acidimicrobiia | uncultured | uncultured |
| **OTU432** | 3.68 | 1.40 | 1.66 | 2.25 | Actinobacteria | Pseudonocardiaceae | *Pseudonocardia* |
| **OTU142** | 0.72 | 1.91 | 2.35 | 1.66 | Actinobacteria | Pseudonocardiaceae | unclassified |
| **OTU7** | 0 | 1.70 | 3.31 | 1.67 | Actinobacteria | Pseudonocardiaceae | uncultured |
| **OTU19** | 2.06 | 0.52 | 1.65 | 1.41 | Actinobacteria | Micrococcaceae | *Arthrobacter* |
| **OTU190** | 0.46 | 0.31 | 2.84 | 1.20 | Acidimicrobiia | uncultured | uncultured |
| **OTU10** | 0.81 | 0.29 | 1.68 | 0.93 | Acidimicrobiia | Acidimicrobiaceae | uncultured |
| **OTU9** | 0.24 | 0.15 | 2.39 | 0.93 | Acidimicrobiia | uncultured | uncultured |
| **OTU251** | 0.77 | 0.65 | 1.12 | 0.85 | Actinobacteria | Pseudonocardiaceae | *Pseudonocardia* |
| **OTU14** | 0.54 | 0.31 | 1.73 | 0.86 | Actinobacteria | Sporichthyaceae | *Sporichthya* |
| **OTU30** | 2.11 | 0.13 | 0.31 | 0.85 | Actinobacteria | Streptomycetaceae | *Streptomyces* |
| **OTU360** | 0.40 | 0.17 | 1.68 | 0.75 | Acidimicrobiia | uncultured | uncultured |
| **OTU24** | 1.93 | 0.17 | 0.08 | 0.73 | Actinobacteria | Streptosporangiaceae | *Planotetraspora* |
| **OTU11** | 0.38 | 0.03 | 1.72 | 0.71 | Acidimicrobiia | uncultured | uncultured |
| **OTU20** | 0.71 | 0.47 | 0.81 | 0.67 | Acidimicrobiia | Iamiaceae | uncultured |
| **OTU47** | 0.75 | 0.26 | 0.98 | 0.66 | Acidimicrobiia | uncultured | uncultured |
| **OTU15** | 0.30 | 0.17 | 1.48 | 0.65 | Actinobacteria | Streptomycetaceae | *Streptomyces* |
| **OTU192** | 0.01 | 1.65 | 0.01 | 0.56 | Actinobacteria | Pseudonocardiaceae | uncultured |
| **OTU16** | 0.24 | 0.51 | 1.11 | 0.62 | Acidimicrobiia | uncultured | uncultured |
| **OTU50** | 0.88 | 0.54 | 0.37 | 0.60 | Actinobacteria | Pseudonocardiaceae | uncultured |
| **OTU22** | 1.61 | 0.16 | 0.06 | 0.61 | Actinobacteria | Propionibacteriaceae | unclassified |
| **OTU23** | 0.27 | 1.19 | 0.10 | 0.52 | Actinobacteria | Micrococcaceae | *Arthrobacter* |
| **OTU54** | 0.13 | 0.92 | 0.46 | 0.50 | Actinobacteria | Pseudonocardiaceae | *Pseudonocardia* |
| **OTU18** | 0.48 | 0.12 | 1.04 | 0.54 | Acidimicrobiia | uncultured | uncultured |
| **OTU25** | 0.56 | 0.09 | 0.83 | 0.49 | Actinobacteria | Nocardiaceae | *Nocardia* |
| **OTU21** | 0.80 | 0.19 | 0.34 | 0.45 | Actinobacteria | Streptomycetaceae | *Streptomyces* |
| **OTU99** | 1.05 | 0.09 | 0.17 | 0.44 | Actinobacteria | Streptomycetaceae | *Streptomyces* |
| **OTU36** | 0.22 | 0.04 | 0.96 | 0.41 | Acidimicrobiia | uncultured | uncultured |
| **OTU44** | 0.80 | 0.14 | 0.26 | 0.40 | Acidimicrobiia | uncultured | uncultured |
| **OTU32** | 0.22 | 0.03 | 0.92 | 0.39 | Actinobacteria | unclasiffied | unclassified |

### *2.4. Comparison of moonmilk Streptomyces OTUs and Streptomyces strains isolated via the culture-dependent approach*

The true diversity of microbial communities is known to be strongly biased by cultivation-based methods in comparison to molecular techniques, therefore we wanted to assess how much of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 February 2018 doi:10.20944/preprints201802.0090.v1

Peer-reviewed version available at Antibiotics 2018, 7, 27; doi:10.3390/antibiotics7020027



the *Streptomyces* moonmilk-dwelling community we managed to isolate in our previous bioprospection work [10]. For this purpose, we compared the 16S rRNA sequences of the 19 *Streptomyces* OTUs retrieved from the HTS approach with the sequences of the 31 previously isolated *Streptomyces* phylotypes (MM strains), which were trimmed to the corresponding V6-V7 variable regions of HTS amplicons. Figure 4 presents the phylogenetic tree generated by maximum likelihood with all the 252 nt 16S rRNA sequences from the *Streptomyces* phylotypes (MM strains) and OTUs. The identity threshold for clustering sequences in the same branch of the tree was fixed to 97%, *i.e.*, the same threshold as the one used to define OTUs in our HTS approach (see methods for details). As deduced from the generated phylogenetic tree, the 31 isolated *Streptomyces* strains matched with only 5 of the 19 *Streptomyces* OTUs, suggesting that the isolated strains represent a minor fraction of the *Streptomyces* species dwelling the moonmilk deposits of the studied cave. Expectedly, Figure 4 further shows that we isolated *Streptomyces* species that are associated with the most abundant *Streptomyces* OTUs, e.g., OTU15, OTU21, OTU30, and OTU99 (Table 4), which together represent 79% of *Streptomyces* sequences retrieved by our HTS approach. Moreover, 21 out of the 31 phylotype strains (68%) clustered together with OTU21 (Figure 4). Finally, two *Streptomyces* isolates, *i.e.*, MM24 and MM106, did not cluster with any of the identified *Streptomyces* OTUs (Figure 4).

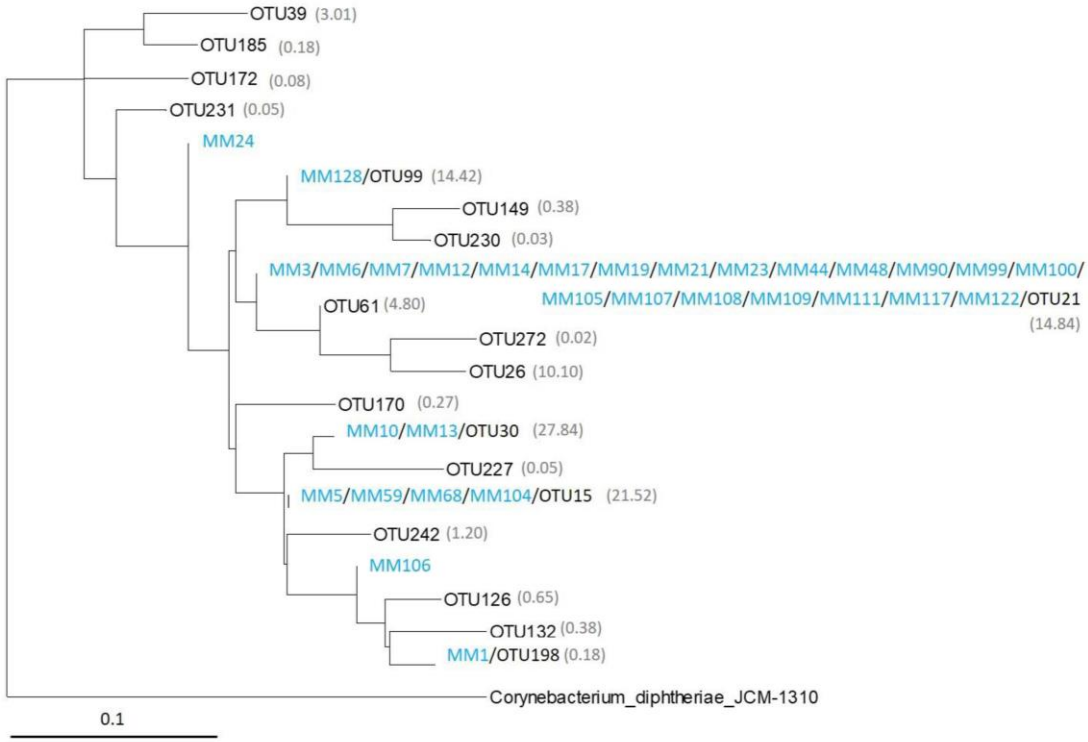


**Figure 4.** Phylogenetic relationships between culturable and non-culturable *Streptomyces* originating from moonmilk of "Grotte des Collemboles". The tree was inferred by maximum likelihood. Scale bar is in substitution per site. Numbers between brackets reflect the predicted mean abundance of *Streptomyces* OTUs in the studied deposits based on the percentage of sequences retrieved from the HTS analysis. *Streptomyces* phylotypes isolated in our previous bioprospection study (MM strains) are marked in blue.

## 3. Discussion

### New insights into moonmilk bacterial diversity revealed by high-throughput sequencing

Previous investigations on the moonmilk microbiome revealed a very diverse microbial community in these deposits [9,13–17,19]. The high-throughput sequencing approach used in this work complemented previous findings by providing an in-depth picture of the bacterial population, together with a detailed taxonomic fingerprint of the phylum Actinobacteria.

Comparison of the bacterial diversity in moonmilk between earlier investigations and the present work is limited to some extent by the differences in experimental procedures, such as DNA isolation and PCR-based approaches, and the sensitivity of the sequencing techniques. Nonetheless, the profile of the major taxonomic groups found in this work is consistent with that observed for the moonmilk communities in the caves "Grotta della Foos" and "Bus della Genziana" in Italy, which were obtained from 16S rRNA clone libraries [9]. All the phyla detected in the above-mentioned caves, including Bacteroidetes, Acidobacteria, Chloroflexi, Planctomycetes, Verrucomicrobia, Actinobacteria, Firmicutes, Nitrospirae, Chlorobi, Proteobacteria, WS3 (now Latescibacteria), were also identified in the cave "Grotte des Collemboles", although their relative abundance varied between the studies. While Proteobacteria were found to be the most abundant phylum in both cases, the second most abundant population identified in Italian caves was the phylum Bacteroidetes, which constituted a minor part of the bacterial community in the present study. The Actinobacteria population was found to be an important part of the moonmilk microbiome in the "Grotte des Collemboles" (from 9 to 23%), but instead represented only a minor fraction (<2%) of the bacterial population in the two Italian caves investigated by Engel et al. (2013). Very recently, a study by Dhami et al. has reported the moonmilk microbiome profile in the Australian "Lake Cave" using a HTS approach [19], as in this work. The presence of Proteobacteria, Actinobacteria, Acidobacteria, Chloroflexi, Nitrospirae, Gemmatimonadetes, Firmicutes and Bacteroidetes was detected in the moonmilk deposit of the "Lake Cave", similarly to the "Grotte des Collemboles". However, many of the low-abundance taxa identified in the Belgian cave were not reported, possibly because phylogenetic profiles were based on different regions of 16S rRNA gene – V3/V4 for the "Lake Cave" and V6-V7 for the "Grotte des Collemboles". Interestingly, unlike in Italian and Belgian caves, the "Lake Cave" moonmilk deposit was strongly dominated by Actinobacteria, which were over twice more abundant than Proteobacteria [19].

The highly sensitive HTS amplicon sequencing approach employed in this work revealed the presence of twenty-six phyla within the moonmilk microbiome that had not been previously described in this speleothem. These included Zixibacteria (formerly RBG-1), Armatimonadetes (formerly OP10), and Parcubacteria (formerly OD1) among the main phyla of moonmilk microbiome (Figure 2a), which have been previously reported from other subterranean environments [24,40–44], and twenty-three low-abundant taxa that were found below the level of 1%, and included many candidate divisions (Figure 2b).

This new study uncovered a surprisingly diverse Actinobacteria taxonomic profile that demonstrates the limitations of our previous cultivation-based screening, in which only *Streptomyces* species could be isolated from the three moonmilk deposits [10]. Here, a total of 47 actinobacterial genera from 28 families were identified across the investigated samples. Beyond the previously reported members of the Actinomycetales family – including *Nocardia* and *Rhodococcus*

(Nocardiaceae) [15,18], *Pseudonocardia, Amycolatopsis* and *Saccharothrix* (not identified in our study) (Pseudonocardiaceae) [12,14,19], *Propionibacterium* (Propionibacteriaceae) [14], *Streptomyces* (Streptomycetaceae) [10,12,18,19], *Arthrobacter* (Micrococcaceae) [13], *Mycobacterium* (Mycobacteriaceae) [19], *Nocardioides, Aeromicrobium* and *Kribbella* (Nocardioidaceae) [19], and *Geodermatophilus* (Geodermatophilaceae) [19] – 35 other genera were identified in the moonmilk deposits of the "Grotte des Collemboles". The population of each investigated sample was also found to include representatives of the Acidimicrobiia class, which were not previously reported in moonmilk. Their presence in all the three sampling sites, with an abundance up to 55% in COL4, and the dominance of the unclassified Acidimicrobiia phylotype (OTU 4) within the community of COL4, suggest that the chemical composition of the investigated moonmilk would be particularly suitable for the development of the representatives of this class of Actinobacteria, of which the ecology and metabolism are still largely unknown.

**Moonmilk deposits as appealing source of novel producers of bioactive compounds**

Extreme environmental niches have recently become the main targets for intense bioprospecting, as they are expected to host diverse yet-unknown microorganisms, which could offer unexplored chemical diversity. While *Streptomyces* are reported as the most prolific 'antibiotic makers', advances in cultivation and characterization of rare Actinobacteria revealed similarly promising capabilities for the production of bioactive natural compounds [45–47]. The results obtained in this work suggest a significant biodiversity of the moonmilk-dwelling actinobacterial population, with a wide spectrum of rare genera. Next to *Streptomyces*, other members of Actinobacteria with valuable secondary metabolism were detected at a high proportion, such as *Pseudonocardia*, *Amycolatopsis*, *Streptosporangium*, *Nocardia*, *Nocardioides*, and *Rhodococcus*. Such findings clearly prompt to apply appropriate selective cultivation methods to isolate rare Actinobacteria from moonmilk deposits.

Moreover, particular importance should be also focused on Acidimicrobiia, which constituted an important part of the community in the studied deposits. Members of this class are a recently identified taxonomic unit [48] considered to represent an early-branching lineage within the phylum [49]. Because of their phylogenetic isolation and novelty, they are likely to hide a yet-uncovered valuable bioactive arsenal.

A great potential of moonmilk as a source of diverse and metabolically beneficial Actinobacteria is illustrated by the comparison of *Streptomyces* isolated in our previous study and the *Streptomyces* phylotypes identified in this work (Figure 4). Most *Streptomyces* OTUs are phylogenetically distinct from culturable representatives (Figure 4), indicating that still a great number of species remain to be isolated. On the other hand, our culture-dependent study identified phylotypes (MM24 and MM106, Figure 4) that were not found via the HTS approach, confirming that both strategies are complementary and should be used in parallel for microbial diversity assessment [50,51]. In addition, next to the identification methods themselves, our data suggests that the diversity level can be also biased by the identity threshold used for OTU definition. The tree revealed that a single OTU (OTU21, Figure 4) clustered together with most of the phylotypes deduced from MLSA (multilocus sequence analysis), each most likely representing a distinct species [10]. This indicates that the 97% sequence homology threshold applied for the comparative analysis of V6-V7 regions of

16S RNA gene largely underestimated the number of *Streptomyces* species dwelling a studied environmental niche.

Additionally, our results showed that different moonmilk deposits host their own indigenous microbial population and thus that each individual speleothem can be a source of great biodiversity. Indeed, only 43% of actinobacterial phylotypes were common to all the samples collected at close distances. Consequently, our findings imply that bioprospecting within different moonmilk deposits - from different caves or within the same cave - could result in the isolation of unique and novel natural compounds producers.

## 4. Materials and Methods

### Site description and sampling

The cave "Grotte des Collemboles" (Springtails' Cave), located in Comblain-au-Pont (GPS coordinates 50° 28' 41" N, 5° 36' 35" E), Belgium (Figure S1, [21] for full description), was formed in Visean limestone and has a shape of a 70-m long meander. White to brown-orange moonmilk deposits are found on the walls in the first narrow chamber located at the entrance of the cave, as well as in the narrow passages leading deeper into the cave (Figure S1). Moonmilk samples used for total DNA extractions were aseptically collected in January 2012 from three spatially separated locations along about a 20 m transect in the cave. Soft moonmilk speleothems were scratched with sterile scalpels into sterile Falcon tubes from the wall in the first chamber, adjacent to the cave entrance (COL4), and from the walls in a narrow passage after the first chamber (COL1, COL3) (Figure S1). COL4 was located approximately 6 m from COL1 and 20 m from COL3 (Figure S1). Samples were immediately transferred to the laboratory, freeze-dried on a VirTis Benchtop SLC Lyophilizer (SP Scientific, Warminster, PA, USA) and stored at -20°C.

### Total DNA extraction and 16S rRNA gene amplicon high-throughput sequencing

The metagenetic approach applied in this work was performed on DNA extracted from three moonmilk deposits (COL1, COL3, and COL4) originating from the "Grotte des Collemboles". Environmental genomic DNA isolation was carried out from 200 mg of the freeze-dried moonmilk samples COL1, COL3, and COL4 (Figure S1), using the PowerClean Soil DNA kit (MoBio, CA, USA), according to manufacturer's instructions. The integrity of purified DNA was assessed by agarose gel electrophoresis (1% w/v) and the dsDNA concentration was evaluated by Qubit fluorometer (Invitrogen, MA, USA).

The 16S rRNA gene amplicon libraries were generated using bacterial (S-D-Bact-0517-a-S-17/S-D-Bact-1061-a-A-17 spanning V4-V6 region; [29]) and actinobacterial (Com2xf/Ac1186r, spanning V6-V7 region; [30]) specific primer pairs. The Illumina platform-compatible dual index paired-end approach was designed as previously described [31] (detailed description provided in the Table S1a). Each forward and reverse primer consisted of an Illumina-compatible forward/reverse primer overhang attached to the 5′ end. Additionally, a heterogeneity spacer of four degenerate nucleotides (Ns) was added to the forward primer, between the primer overhang and the locus-specific sequence. The Illumina barcodes and sequencing adapters were added during the subsequent cycle-limited amplification step using Nextera XT Index kit (Illumina, CA, USA). Triplicated PCR reactions were performed for each sample in 25 µl volume containing 2.5 µl of total DNA, 5 µl of each primer (1 µM), 12.5 µl of 2X Q5 High-Fidelity Master Mix (New England Biolabs,

Peer-reviewed version available at Antibiotics 2018, 7, 27; doi:10.3390/antibiotics7020027



UK). Amplification conditions for each set of primers are listed in Table S1b. The triplicated amplicons were visualized on 3% agarose gel, pulled, purified with Agencourt AMPure XP beads (Beckman Coulter, CA, USA) and quantified with the Qubit HS dsDNA assay kit (Invitrogen, MA, USA) before being processed for index ligation, using the Nextera XT Index kit (Illumina, CA, USA). The PCR amplifications were performed with the same enzyme and cycling conditions as described above [31], with the total number of cycles reduced to eight and the annealing temperature of 55°C. The resulting amplicons were purified with the Agencourt AMPure XP magnetic beads (Beckman Coulter, CA, USA), quantified and pooled in equimolar concentrations. The library concentration was quantified by qPCR using Kappa SYBR FAST kit (Kappa Biosystems, MA, USA) and subsequently the library was normalized to 4 nM, denaturated and diluted to the final concentration of 8 pM. The resulting pool was mixed with the PhiX control and subjected to 2x300 bp paired-end sequencing on Illumina MiSeq platform (Illumina, CA, USA). Raw sequences were deposited in NCBI Sequence Read Archive (SRA) database under the Bioproject PRJNA428798 with accession numbers SRX3540524 - SRX3540529.

**16S rRNA amplicon analysis**

16S rRNA amplicon analysis was based for both Bacteria and Actinobacteria on forward reads only, owing to the poor quality of reverse reads. Quality trimming (prohibiting mismatches and ambiguities, ensuring a minimum quality score of 20 and removing the four degenerate nucleotides from the 5′ end) was carried out using CLC Genomic Workbench (Qiagen, Germany). USEARCH [32] was applied for length trimming (minimum length = 240 nt) and dereplication. Operational taxonomic units (OTUs) for both bacterial and actinobacterial datasets were defined using a 97% identity threshold on 16S rRNA sequences. OTUs were clustered using the UPARSE algorithm [33] and their taxonomic position was assigned by MOTHUR [34] with SILVA v128 database [35]. OTUs were further classified using BLASTN [36] analyses against a local mirror of NCBI nt database (downloaded on 09-Aug-2017), through manual and automatic analyses. For the automatic approach, a Last Common Ancestor (LCA) classification was performed with a custom parser mimicking the MEGAN algorithm [37], which we developed for analyses of genome contamination (Cornet et al., 2017, under review). A maximum number of 100 hits per OTU were taken into account. To consider a BLASTN hit, the E-value threshold was set at 1e-15, the minimum identity threshold a 95.5%, the minimum bit score at 200 and the bit score percentage threshold at 99% of the best hit. These thresholds were defined through preliminary analyses (data not shown). When BLASTN hits are too numerous, the MEGAN-like algorithm frequently yields high-ranking LCAs (e.g. Bacteria) that are not informative in practice. To minimize this effect, we decided to skip uncultured/unclassified hits whenever other, more informative, hits also passed the thresholds. Moreover, when computing LCAs, we only considered the most frequent taxa provided that they represented ≥95% of the (up to 100) accumulated BLASTN hits, so as to avoid uninformative classifications due to a few (possibly aberrant) outliers.

Normalized OTU abundance data was used to calculate $\alpha$- and $\beta$-diversity estimators using MOTHUR [34]. Community richness, evenness, diversity and differential OTU abundance between samples were calculated using sobs, Simpson index, inverse Simpson index and Venn diagrams, respectively.

The 19 OTUs identified as *Streptomyces* were combined to 31 sequences (16S rRNA region V6-V7) from previously isolated *Streptomyces* phylotypes (MM strains) and dereplicated with the UCLUST algorithm [32] using an identity threshold of 97%. This yielded 21 clusters, to which we added the homologous region of *Corynebacterium diphteriae* JCM-1310 as outgroup. A multiple sequence alignment was built with MUSCLE [38] (default parameters) and then analyzed with PhyML [39] under a K80+$\Gamma_4$ model. Due to the limited amount of phylogenetic signal (short sequences from very related organisms), the resolution of the tree was low (bootstrap proportions <50 for nearly all nodes; data not shown).

**Supplementary Materials**

**Figure S1:** Localization of the "Grotte des Collemboles" (Springtails' Cave) together with the cave map and visualization of the moonmilk deposit sampling points.

**Figure S2:** Rarefaction curves of OTUs clustered at 97% sequence identity across the three moonmilk-sampling points for Bacteria (a) and Actinobacteria (b).

**Figure S3.** Taxonomic profile of bacterial phyla generated with Actinobacteria-specific primers. Note the high specificity of Actinobacteria primers.

**Table S1:** Details of the PCR primers used for community profiling of moonmilk samples (a) and -PCR conditions used for 16S rRNA amplification from moonmilk samples (b).

**Table S2:** Relative abundance (%) of bacterial phyla identified in the three moonmilk deposits in "Grotte des Collemboles". Low-abundant taxa with relative abundance <1% are marked in red.

**Table S3:** Relative abundance (%) of the phylum Actinobacteria at different taxonomic levels identified in the three moonmilk deposits in the "Grotte des Collemboles".

**Acknowledgments:** The authors are grateful to Luc Willems for the introduction to the subject and help with sampling. MM, DA, and LC work was supported by a Research Foundation for Industry and Agriculture (FRIA) grant. MC and PD were supported by the Luxembourg National Research Fund (FNR CORE 2011 project GASPOP; C11/SR/1280949: Influence of the Reactor Design and the Operational Parameters on the Dynamics of the Microbial Consortia Involved in the Biomethanation Process). Computational resources ("durandal" grid computer) were funded by three grants from the University of Liège, "Fonds spéciaux pour la recherche," "Crédit de démarrage 2012" (SFRD-12/03 and SFRD-12/04) and "Crédit classique 2014" (C-14/73) and by a grant from the F.R.S.-FNRS "Crédit de recherche 2014" (CDR J.0080.15). This work is supported in part by the Belgian program of Interuniversity Attraction Poles initiated by the Federal Office for Scientific Technical and Cultural Affairs (PAI no. P7/44). SR is a Research Associate at Belgian Fund for Scientific Research (F.R.S-FNRS). The authors declare no conflict of interest. We dedicate the work to the memory of Leonard Maculewicz (1936-2017) who always supported our work with great enthusiasm.

**Author Contributions:** MM, MaC, SM, MC, and SR designed and performed experiments. Bioinformatic analyses were performed by MM, MC, MaC, LC, DB, and SR. Data were analyzed by all authors. The manuscript was written and/or corrected by all authors.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References


1. Engel, A. S. Microbial diversity of cave ecosystems. In *Geomicrobiology: Molecular and Environmental Perspective*; 2010; pp. 219–238, ISBN 9789048192038.
2. Barton, H. A.; Northup, D. E. Geomicrobiology in cave environments: Past, current and future perspectives. *J. Cave Karst Stud.* **2007**, *69*, 163–178.
3. Groth, I.; Vettermann, R.; Schuetze, B.; Schumann, P.; Saiz-Jimenez, C. Actinomycetes in Karstic caves of northern Spain (Altamira and Tito Bustillo). *Journal of Microbiological Methods*; **1999**; Vol. 36, pp. 115–122, doi:10.1016/S0167-7012(99)00016-0.
4. Jurado, V.; Kroppenstedt, R. M.; Saiz-Jimenez, C.; Klenk, H. P.; Mouniée, D.; Laiz, L.; Couble, A.; Pötter, G.; Boiron, P.; Rodríguez-Nava, V. *Hoyosella altamirensis* gen. nov., sp. nov., a new member of the order Actinomycetales isolated from a cave biofilm. *Int. J. Syst. Evol. Microbiol.* **2009**, *59*, 3105–3110, doi:10.1099/ijs.0.008664-0.
5. Jurado, V.; Groth, I.; Gonzalez, J. M.; Laiz, L.; Saiz-Jimenez, C. *Agromyces salentinus* sp. nov. and *Agromyces neolithicus* sp. nov. *Int. J. Syst. Evol. Microbiol.* **2005**, *55*, 153–157, doi:10.1099/ijs.0.63199-0.
6. Maciejewska, M.; Pessi, I. S.; Arguelles-Arias, A.; Noirfalise, P.; Luis, G.; Ongena, M.; Barton, H.; Carnol, M.; Rigali, S. *Streptomyces lunaelactis* sp. nov., a novel ferroverdin A-producing *Streptomyces* species isolated from a moonmilk speleothem. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **2015**, *107*, 519–531, doi:10.1007/s10482-014-0348-4.
7. Lee, S. D.; Kang, S. O.; Hah, Y. C. *Hongia* gen, nov., a new genus of the order Actinomycetales. *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 191–199, doi:10.1099/00207713-50-1-191.
8. Ortiz, M.; Neilson, J. W.; Nelson, W. M.; Legatzki, A.; Byrne, A.; Yu, Y.; Wing, R. A.; Soderlund, C. A.; Pryor, B. M.; Pierson, L. S.; Maier, R. M. Profiling Bacterial Diversity and Taxonomic Composition on Speleothem Surfaces in Kartchner Caverns, AZ. *Microb. Ecol.* **2013**, *65*, 371–383, doi:10.1007/s00248-012-0143-6.
9. Engel, A. S.; Paoletti, M. G.; Beggio, M.; Dorigo, L.; Pamio, A.; Gomiero, T.; Furlan, C.; Brilli, M.; Dreon, A. L.; Bertoni, R.; Squartini, A. Comparative microbial community composition from secondary carbonate (moonmilk) deposits: implications for the *Cansiliella servadeii* cave hygropetric food web. *Int. J. Speleol.* **2013**, *42*, 181–192. doi:10.5038/1827-806X.42.3.2
10. Maciejewska, M.; Adam, D.; Martinet, L.; Naômé, A.; Calusinska, M.; Smargiasso, N.; De Pauw, E.; Barton, H.; Carnol, M.; Hanikenne, M.; Hayette, M.-P.; Delfosse, P.; Baurain, D.; Rigali, S. A Phenotypic and Genotypic Analysis of the Antimicrobial Potential of Cultivable *Streptomyces* isolated from Cave Moonmilk Deposits. *Front. Microbiol.* **2016**, doi:10.3389/fmicb.2016.01455.
11. Cañaveras, J. C.; Cuezva, S.; Sanchez-Moral, S.; Lario, J.; Laiz, L.; Gonzalez, J. M.; Saiz-Jimenez, C. On the origin of fiber calcite crystals in moonmilk deposits. *Naturwissenschaften* **2006**, *93*, 27–32, doi:10.1007/s00114-005-0052-3.
12. Cañaveras, J. C.; Hoyos Gómez, M.; Sánchez-Moral, S.; Sanz Rubio, E.; Bedoya, J.; Hoyos, V.; Groth, I.; Schumann, P.; Laiz Trobajo, L. Microbial Communities Associated With Hydromagnesite and Needle-Fiber Aragonite Deposits in a Karstic Cave (Altamira, Northern Spain). *Geomicrobiol. J.* **1999**, *16*, 9–25, doi:10.1080/014904599270712.
13. Rooney, D. C.; Hutchens, E.; Clipson, N.; Baldini, J.; McDermott, F. Microbial Community

Diversity of Moonmilk Deposits at Ballynamintra Cave, Co. Waterford, Ireland. *Microb. Ecol.* **2010**, *60*, 753–761, doi:10.1007/s00248-010-9693-7.

14. Portillo, M. C.; Gonzalez, J. M. Moonmilk Deposits Originate from Specific Bacterial Communities in Altamira Cave (Spain). *Microb. Ecol.* **2011**, *61*, 182–189, doi:10.1007/s00248-010-9731-5.
15. Reitschuler, C.; Lins, P.; Schwarzenauer, T.; Spotl, C.; Wagner, A. O.; Illmer, P. New Undescribed Lineages of Non-extremophilic Archaea Form a Homogeneous and Dominant Element Within Alpine Moonmilk Microbiomes. *Geomicrobiol. J.* **2015**, *32*, doi: 10.1080/01490451.2015.1025317.
16. Reitschuler, C.; Spötl, C.; Hofmann, K.; Wagner, A. O.; Illmer, P. Archaeal Distribution in Moonmilk Deposits from Alpine Caves and Their Ecophysiological Potential. *Microb. Ecol.* **2016**, *71*, 686–699, doi:10.1007/s00248-015-0727-z.
17. Reitschuler, C.; Lins, P.; Wagner, A. O.; Illmer, P. Cultivation of moonmilk-born non-extremophilic Thaum and Euryarchaeota in mixed culture. *Anaerobe* **2014**, *29*, 73–79, doi:10.1016/j.anaerobe.2013.10.002.
18. Axenov-Gibanov, D. V.; Voytsekhovskaya, I. V.; Tokovenko, B. T.; Protasov, E. S.; Gamaiunov, S. V.; Rebets, Y. V.; Luzhetskyy, A. N.; Timofeyev, M. A. Actinobacteria isolated from an underground lake and moonmilk speleothem from the biggest conglomeratic karstic cave in Siberia as sources of novel biologically active compounds. *PLoS One* **2016**, *11*, doi:10.1371/journal.pone.0149216.
19. Dhami, N. K.; Mukherjee, A.; Watkin, E. Characterisation of Mineralogical-Mechanical-Microbial properties of calcitic speleothems and the *in vitro* biomineralization potential of associated microbial communities. *Front. Microbiol.* **2018**, *9*, 40, doi:10.3389/FMICB.2018.00040.
20. Bindschedler, S.; Milliere, L.; Cailleau, G.; Job, D.; Verrecchia, E. P. Calcitic nanofibres in soils and caves: a putative fungal contribution to carbonatogenesis. *Geol. Soc. London, Spec. Publ.* **2010**, *336*, 225–238, doi:10.1144/SP336.11.
21. Maciejewska, M.; Adam, D.; Naômé, A.; Martinet, L.; Tenconi, E.; Calusinska, M.; Delfosse, P.; Hanikenne, M.; Baurain, D.; Compère, P.; Carnol, M.; Barton, H. A.; Rigali, S. Assessment of the potential role of *Streptomyces* in cave moonmilk formation. *Front. Microbiol.* **2017**, *8*, doi:10.3389/fmicb.2017.01181.
22. Nimaichand, S.; Devi, A. M.; Tamreihao, K.; Ningthoujam, D. S.; Li, W. J. Actinobacterial diversity in limestone deposit sites in Hundung, Manipur (India) and their antimicrobial activities. *Front. Microbiol.* **2015**, *6*, doi:10.3389/fmicb.2015.00413.
23. Wu, Y.; Tan, L.; Liu, W.; Wang, B.; Wang, J.; Cai, Y.; Lin, X. Profiling bacterial diversity in a limestone cave of the western Loess Plateau of China. *Front. Microbiol.* **2015**, *6*, doi:10.3389/fmicb.2015.00244.
24. Northup, D. E.; Melim, L. A.; Spilde, M. N.; Hathaway, J. J. M.; Garcia, M. G.; Moya, M.; Stone, F. D.; Boston, P. J.; Dapkevicius, M. L. N. E.; Riquelme, C. Lava Cave Microbial Communities Within Mats and Secondary Mineral Deposits: Implications for Life Detection on Other Planets. *Astrobiology* **2011**, *11*, 601–618, doi:10.1089/ast.2010.0562.
25. Riquelme, C.; Hathaway, J. J. M.; Dapkevicius, M. de L. N. E.; Miller, A. Z.; Kooser, A.; Northup, D. E.; Jurado, V.; Fernandez, O.; Saiz-Jimenez, C.; Cheeptham, N. Actinobacterial diversity in volcanic caves and associated geomicrobiological interactions. *Front. Microbiol.*

**2015**, *6*, doi:10.3389/fmicb.2015.01342.

26. Cheeptham, N.; Sadoway, T.; Rule, D.; Watson, K.; Moote, P.; Soliman, L. C.; Azad, N.; Donkor, K. K.; Horne, D. Cure from the cave: Volcanic cave actinomycetes and their potential in drug discovery. *Int. J. Speleol.* **2013**, *42*, 35–47, doi:10.5038/1827-806X.42.1.5.
27. Tebo, B. M.; Davis, R. E.; Anitori, R. P.; Connell, L. B.; Schiffman, P.; Staudigel, H. Microbial communities in dark oligotrophic volcanic ice cave ecosystems of Mt. Erebus, Antarctica. *Front. Microbiol.* **2015**, *6*, doi:10.3389/fmicb.2015.00179.
28. Chater, K. F.; Biró, S.; Lee, K. J.; Palmer, T.; Schrempf, H. The complex extracellular biology of *Streptomyces*. *FEMS Microbiol. Rev.* **2010**, *34*, 171–198, doi:10.1111/j.1574-6976.2009.00206.x.
29. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F. O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, doi:10.1093/nar/gks808.
30. Schäfer, J.; Jäckel, U.; Kämpfer, P. Development of a new PCR primer system for selective amplification of Actinobacteria. *FEMS Microbiol. Lett.* **2010**, *311*, 103–112, doi:10.1111/j.1574-6968.2010.02069.x.
31. Goux, X.; Calusinska, M.; Fossépré, M.; Benizri, E.; Delfosse, P. Start-up phase of an anaerobic full-scale farm reactor - Appearance of mesophilic anaerobic conditions and establishment of the methanogenic microbial community. *Bioresour. Technol.* **2016**, *212*, 217–226, doi:10.1016/j.biortech.2016.04.040.
32. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461, doi:10.1093/bioinformatics/btq461.
33. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **2013**, *10*, 996–8, doi:10.1038/nmeth.2604.
34. Schloss, P. D.; Westcott, S. L.; Ryabin, T.; Hall, J. R.; Hartmann, M.; Hollister, E. B.; Lesniewski, R. A.; Oakley, B. B.; Parks, D. H.; Robinson, C. J.; Sahl, J. W.; Stres, B.; Thallinger, G. G.; Van Horn, D. J.; Weber, C. F. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541, doi:10.1128/AEM.01541-09.
35. Pruesse, E.; Peplies, J.; Glöckner, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **2012**, *28*, 1823–1829, doi:10.1093/bioinformatics/bts252.
36. Boratyn, G. M.; Schäffer, A. A.; Agarwala, R.; Altschul, S. F.; Lipman, D. J.; Madden, T. L. Domain enhanced lookup time accelerated BLAST. *Biol. Direct* **2012**, *7*, doi:10.1186/1745-6150-7-12.
37. Huson, D. H.; Auch, A. F.; Qi, J.; Schuster, S. C. MEGAN analysis of metagenomic data. *Genome Res.* **2007**, *17*, 377–386, doi:10.1101/gr.5969107.
38. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797, doi:10.1093/nar/gkh340.
39. Guindon, S.; Dufayard, J. F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321, doi:10.1093/sysbio/syq010.
40. Castelle, C. J.; Hug, L. A.; Wrighton, K. C.; Thomas, B. C.; Williams, K. H.; Wu, D.; Tringe, S. G.; Singer, S. W.; Eisen, J. A.; Banfield, J. F. Extraordinary phylogenetic diversity and metabolic

versatility in aquifer sediment. *Nat. Commun.* **2013**, *4*, doi:10.1038/ncomms3120.

41. Schabereiter-Gurtner, C.; Saiz-Jimenez, C.; Pinar, G.; Lubitz, W.; Rolleke, S. Phylogenetic diversity of bacteria associated with Paleolithic paintings and surrounding rock walls in two Spanish caves (Llonin and La Garma). *FEMS Microbiol. Ecol.* **2004**, *47*, 235–247, doi:10.1016/S0168-6496(03)00280-0.
42. Pedersen, K.; Arlinger, J.; Ekendahl, S.; Hallbeck, L. 16S rRNA gene diversity of attached and unattached bacteria in boreholes along the access tunnel to the Äspö hard rock laboratory, Sweden. *FEMS Microbiol. Ecol.* **1996**, *19*, 249–262, doi:10.1016/0168-6496(96)00017-7.
43. Zhou, J.; Gu, Y.; Zou, C.; Mo, M. Phylogenetic diversity of bacteria in an earth-cave in Guizhou Province, Southwest of China. *J. Microbiol.* **2007**, *45*, 105–112.
44. De Mandal, S.; Chatterjee, R.; Kumar, N. S. Dominant bacterial phyla in caves and their predicted functional roles in C and N cycle. *BMC Microbiol.* **2017**, *17*, 90, doi:10.1186/s12866-017-1002-x.
45. Tiwari, K.; Gupta, R. K. Rare actinomycetes: A potential storehouse for novel antibiotics. *Crit. Rev. Biotechnol.* **2012**, *32*, 108–132, doi: 10.3109/07388551.2011.562482.
46. Choi, S. S.; Kim, H. J.; Lee, H. S.; Kim, P.; Kim, E. S. Genome mining of rare actinomycetes and cryptic pathway awakening. *Process Biochem.* **2015**, *50*, 1184–1193, doi: 10.1016/j.procbio.2015.04.008.
47. Rigali, S.; Anderssen, S.; Naômé, A.; van Wezel, G. P. Cracking the regulatory code of biosynthetic gene clusters as a strategy for natural product discovery. *Biochem. Pharmacol.* **2018**, doi: 10.1016/j.bcp.2018.01.007.
48. Norris, P. Class II. Acidimicrobiia class. nov. In *Bergey's Manual of Systematic Bacteriology v.5*; M. Goodfellow, P. Kampfer, H.-J. Busse, M.E. Trujillo, K.-I. Suzuki, W. L. and W. B. W., Ed.; Springer New York, 2012; p. 1968.
49. Bull, A. T. Actinobacteria of the extremobiosphere. In *Extremophiles Handbook*; 2011; pp. 1203–1240, ISBN 9784431538981.
50. Stefani, F. O. P.; Bell, T. H.; Marchand, C.; Providencia, I. E. de la; Yassimi, A. El; St-Arnaud, M.; Hijri, M. Culture-Dependent and -Independent Methods Capture Different Microbial Community Fractions in Hydrocarbon-Contaminated Soils. *PLoS One* **2015**, doi:10.1371/journal.pone.0128272.
51. Vaz-Moreira, I.; Egas, C.; Nunes, O. C.; Manaia, C. M. Culture-dependent and culture-independent diversity surveys target different bacteria: A case study in a freshwater sample. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **2011**, *100*, 245–257, doi:10.1007/s10482-011-9583-0.