



Published in final edited form as:

Geomicrobiol J. 2014 ; 31(3): 221–235. doi:10.1080/01490451.2012.752424.

Diversity of Ammonia Oxidation (*amoA*) and Nitrogen Fixation (*nifH*) Genes in Lava Caves of Terceira, Azores, Portugal

Jennifer J. Marshall Hathaway¹, Robert L. Sinsabaugh¹, Maria De Lurdes N. E. Dapkevicius², and Diana E. Northup^{1,*}

¹Department of Biology, University of New Mexico, Albuquerque, New Mexico, USA

²CITA-A, Departamento de Ciências Agrárias, Universidade dos Açores, Angra do Heroísmo, Portugal

Abstract

Lava caves are an understudied ecosystem in the subterranean world, particularly in regard to nitrogen cycling. The diversity of ammonia oxidation (*amoA*) and nitrogen fixation (*nifH*) genes in bacterial mats collected from lava cave walls on the island of Terceira (Azores, Portugal) was investigated using denaturing gradient gel electrophoresis (DGGE). A total of 55 samples were collected from 11 lava caves that were selected with regard to surface land use. Land use types above the lava caves were categorized into pasture, forested, and sea/urban, and used to determine if land use influenced the ammonia oxidizing and nitrogen fixing bacterial communities within the lava caves. The soil and water samples from each lava cave were analyzed for total organic carbon, inorganic carbon, total nitrogen, ammonium, nitrate, phosphate and sulfate, to determine if land use influences either the nutrient content entering the lava cave or the nitrogen cycling bacteria present within the cave. *Nitrospira*-like sequences dominated the ammonia-oxidizing bacteria (AOB) community, and the majority of the diversity was found in lava caves under forested land. The nitrogen fixation community was dominated by *Klebsiella pneumoniae*-like sequences, and diversity was evenly distributed between pasture and forested land, but very little overlap in diversity was observed. The results suggest that land use is impacting both the AOB and the nitrogen fixing bacterial communities.

Keywords

amoA; bacteria; lava caves; *nifH*; nitrogen cycling

Introduction

The cycling of nitrogen provides energy and nutrients for many types of microorganisms. Organisms responsible for nitrogen transformations have been well studied in soil and aquatic habitats, yet new discoveries are still being made, such as the discovery of *Archaea* that perform ammonia oxidation (Collins et al. 2008; McCarthy et al. 2007; Prosser and Nicol 2008; Ross et al. 2009; Sooksa-nguan et al. 2009; Treusch et al. 2005; Venter et al.

* Address correspondence to Diana E. Northup, Department of Biology, MSC03 2020, University of New Mexico, Albuquerque, New Mexico 87131, USA; dnorthup@unm.edu.

2004). Certain nitrogen cycling processes can provide energy as well as nutrients for microbial communities that grow in oligotrophic environments, such as caves. For closed or semiclosed environments, nitrogen fixation by *Bacteria* and *Archaea* can be the main source of bioavailable nitrogen for other organisms, while ammonia from organic matter mineralization can be a source of energy for chemolithotrophic organisms. Studies of lithotrophic ammonia oxidation in subsurface environments are rare (Chen et al. 2009; Simon and Benfield 2002; Spear et al. 2007). More generally, there has been little study of nitrogen cycling or the diversity of the organisms that carry out nitrogen transformations in subsurface environments.

We investigated the diversity of *amoA* and *nifH* genes that encode the enzymes that mediate nitrogen transformation processes in microbial mats collected in lava caves of Terceira, Azores, Portugal. Nitrogenase, the enzyme responsible for the fixation of nitrogen, is partially encoded by the gene *nifH* (Dean and Jacobson 1992; Ueda et al. 1995). Nitrogen fixation is an energetically expensive process that is preformed anaerobically. It has not been directly measured in oligotrophic cave environments, but putative nitrogen fixing microbial taxa have been found in carbonate caves (Barton et al. 2004; Dichosa 2008; Northup et al. 2003).

The limiting step in nitrification, the transformation of ammonia to nitrate, is the conversion of NH_4^+ to NH_2OH (Kowalchuk and Stephen 2001). This process is controlled by the enzyme ammonia mono-oxygenase, which is partly encoded by the highly conserved gene *amoA*. Ammonia is not stable in oxic environments and is quickly oxidized by ammonia-oxidizing bacteria (AOB) and archaea; however, moderate levels of ammonia have been found in ferromanganese deposits in Spider and Lechuguilla Caves in New Mexico, giving evidence of the potential availability of substrates for ammonia oxidation (Northup et al. 2003). The authors hypothesize that the ammonium is bound to the clay particles in the ferromanganese deposits and hence is biologically unavailable. However, the presence of ammonia may be an indication of nitrogen fixation.

Analyses of nitrogen cycling have not been conducted in subterranean habitats even though nitrogen limitation is thought to be a major constraint on biological productivity (Barton and Jurado 2007). Simon and Benfield (2002) looked at nitrogen levels in a karst cave stream, and found that nitrogen was not as limiting as carbon in that aquatic environment. However, there is little data regarding nitrogen availability in nonaquatic cave environments (Fliermans and Schmidt 1977; Jones et al. 2008; Northup et al. 2003, Snider 2010). Two studies have looked for and found the presence of ammonia-oxidizing genes in the subsurface. Spear et al. (2007) found archaeal *amoA* in a mine adit and Chen et al. (2009) found bacterial *amoA* in a carbonate cave. Neither study investigated if the genes were being expressed in the subsurface. No studies to date have looked specifically for the presence of nitrogen fixation genes in cave environments.

Lava caves are a unique subset of subsurface environments, and are formed during a volcanic eruption. One major mechanism for the formation of lava caves is that as the erupting lava cools on the surface, a river of molten lava continues to flow underneath. When the eruption ceases, the river of lava flows downhill, leaving an empty tube behind.

Lava caves are generally more shallow subterranean environments than limestone caves (Palmer 2007). Their shallowness, combined with the tendency for lava caves to have cracks in the overlying volcanic rocks, results in more connectivity to the surface compared to many limestone caves (Howarth, 1996).

The extensive colorful microbial mats that cover the walls and ceilings of lava caves have long been described, but the composition and diversity of the mats is just being explored with culture-independent techniques (Garcia et al. 2009; Moya et al. 2009; Northup et al. 2008, 2011; Snider 2010; Snider et al. 2009; Staley and Crawford 1975; Stoner and Howarth 1981). A comparison of 16S rRNA gene bacterial diversity from lava caves in Hawai'i and New Mexico in the United States, and the Azores, Portugal, showed Actinobacteria, Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Deltaproteobacteria and Acidobacteria present in all caves studied (Northup et al. 2011).

In an analysis of 16S rRNA bacterial gene clone libraries, Actinobacteria such as *Frankia* and Alphaproteobacteria such as *Mesorhizobium* sequences are found (Hathaway unpublished data). Additionally *Nitrosospira* sp. and *Nitrosomonas* sp. were found, indicating that there is the possibility to do both nitrogen fixation and ammonia oxidation within lava caves (Hathaway unpublished data). Studies of other types of basaltic environments have shown that lava is deficient in both carbon and nitrogen, and bacteria that can fix nitrogen play an important role in the establishment of other bacterial communities in these environments (King 2003; Mason et al. 2009).

We investigated the presence of key genes involved in nitrogen fixation and ammonia oxidation in 11 lava caves on Terceira Island in the Azores, Portugal. Terceira has lava caves across the island, some underneath extensive cow pastures and some underneath forested areas of both native and exotic plants. The soils overlying the lava caves are classified as andisols (Madeira et al. 2007). Because of the high precipitation rate and high hydrologic connectivity between the lava caves and the surface, we hypothesized that nutrient inputs to lava caves would vary with land use and elevation. In response to these differences, we predicted that the diversity of *nifH* and *amoA* genes in the microbial mats of lava caves would show a complementary pattern of variation.

Materials and Methods

Sample Site Description

Terceira is located in the Atlantic Ocean, in the center of the Azorean island chain at 38° 44' N, 27° 17' W, approximately 1,500 km off the coast of Portugal. Eleven lava caves were selected to represent a range of elevation, annual precipitation, and surface land use (Figure 1 and 2; Table 1).

Sample Collection

Samples of microbial mats were collected using aseptic methods in February and June 2008. Samples of yellow, white and tan, the most frequently observed colors of microbial mats, were collected from each cave, as well as microbial mats of less frequently observed mat colors, such as pink, grey and black. Samples were selected for collection based on

uniformity of color (see Figure 2 for an example microbial mat). Samples were covered with sucrose lysis buffer to preserve the DNA (Giovannoni et al. 1990) and transported to the lab where they were stored at -80°C until DNA was extracted. Water for nutrient analyses and soil samples for carbon/nitrogen analyses were collected in June 2008 and July 2009. One water sample was collected per cave with the exception of Gruta de Balcoes, where two water samples were collected. Results from the analysis of the water from Gruta de Balcoes were averaged. The limited water sampling was due to the scarcity of dripping water in the caves. Samples were kept at 4°C until analysis. Temperature and humidity data (wet bulb/dry bulb) were collected throughout the cave with an IMC Digital Thermometer probe (Wittenberg, WI, USA). This instrument does not allow for accurate humidity measurements above 95%, but can be calibrated, which increased accuracy below 95%.

Water and Soil Chemistry

Dissolved organic carbon (DOC) in infiltrating water was collected and passed through a $0.45\ \mu\text{m}$ filter and preserved at pH 2 using HCL on site, as described in Simon et al. (2007). Organic and inorganic carbon water samples were analyzed using the persulfate digestion method as described in Clescerl et al. (1999) on a Shimadzu TOC-5050A instrument (Shimadzu Corporation, Kyoto, Japan). Amounts of sulfate, nitrate, and phosphate were analyzed using a Dionex Ion Chromatograph DX-100 (Dionex, Sunnyvale, CA, USA) as described in Pfaff et al. (1997). The amount of ammonia in the water samples was analyzed using a Technicon (1973) AutoAnalyzer II (Technicon, Tarrytown, NY, USA).

One soil sample was collected from the floor near the entrance of each lava cave and one soil sample was collected from deeper within the interior of the lava cave. Percent nitrogen and percent carbon in soil were determined by high temperature combustion, the resulting gases were eluted on a gas chromatography column and detected by thermal conductivity and integrated to yield carbon and nitrogen content. Analyses were performed on a ThermoQuest CE Instruments NC2100 Elemental Analyzer (ThermoQuest Italia S.p.A., Rodano, Italy) (Pella 1990a; Pella 1990b). The results from soil samples collected from within one lava cave were tested for statistically significant differences then averaged.

DNA Extraction and PCR Amplification

DNA was extracted and purified from 55 samples from the 11 lava caves on Terceira, using the MoBio PowerSoil DNA Isolation Kit using the manufacturer's protocol, with the modification of bead beating instead of vortexing to break open cells (MoBio, Carlsbad, CA). Extractions with no sample added were performed as negative controls. Samples were then screened for the two functional genes, *amoA* and *nifH*, using polymerase chain reaction (PCR).

For *amoA*, the primers *amoA1F* and *amoA2R* were used under conditions described in Table 2. Reactions were carried out in a $25\text{-}\mu\text{L}$ reaction mixture containing $1\times$ PCR buffer with 2 mM Mg^{2+} , 0.2 mM each dNTP, $0.4\ \mu\text{M}$ of each primer, $5\ \mu\text{g}$ BSA and 0.75 U of TaKaRa Ex Taq (TaKaRa, Shiga, Japan).

For *nifH*, anested PCR was required with primers FGPH19 and PolR used in the first round PCR and AQER and PolF used in the second round (Table 2). Reactions were carried out in a 50- μ L reaction mixture containing 1X PCR buffer with 1.5 mM Mg^{2+} , 0.8 μ M of each primer, 0.2 mM each dNTP, 5 μ g BSA and 1U AmpliTaq LD (Applied Biosystems, Foster City, CA, USA). Three microliters of PCR product from round one was used in the second round PCR, with the same conditions as above, except the primer concentration was decreased to 0.4 μ M.

Denaturing Gradient Gel Electrophoresis

For any sample that was positive for the gene of interest, a denaturing gradient gel electrophoresis (DGGE) was performed. The samples were amplified using the same primer sets, with a GC clamp attached to the 5' end of the forward primer using the same PCR conditions as previously described (Table 2). Samples were run on an 8% (w/v) bis-acrylamide gel with a urea-formamide gradient of 40% to 70% (w/v) for 16 hrs at 110 V in 17 L of 60°C 1 \times TAE buffer on a CBS DGGE-1001 (CBS Scientific, Del Mar, CA). Gels were run in duplicate, with one for band excising and one for imaging.

Gels were stained in 1X SYBRGold (Molecular Probes, Eugene, OR), and imaged on a Sygene InGenius Bio Imager (Sygene, Frederick, MD). Individual bands were then excised, reamplified under the same conditions as the initial amplification, with only the second primer set used for *nifH*, and cleaned using the MoBio's UltraClean PCR-Clean-up (MoBio, Carlsbad, CA). Bands that were successfully reamplified were then sequenced with Big Dye Terminator v1.1 using 5 μ L of PCR product (Applied Biosystems, Foster City, CA), and sequenced on an ABI 3130 sequencing machine (Applied Biosystems, Foster City, CA).

Sequences were edited for quality using Sequencher 4.9 (Gene Codes, Ann Arbor, MI), and then aligned using MUSCLE (www.ebi.ac.uk/Tools/muscle; Edgar 2004). Sequences from the same sample with >97% similarity were defined as the same operational taxonomic unit (OTU) using mothur version 1.17 (Schloss et al. 2009). Community analyses were also preformed in mothur. Parsimony analyses were performed using PAUP (version 4.0b10, distributed by Sinauer; <http://paup.csit.fsu.edu/>) with a bootstrap analysis conducted on 1000 re-sampled datasets. Sequences were submitted to the Gen-Bank database and assigned accession numbers HM461331-HM461407 for *amoA* sequences, and HM461408-HM461553 for *nifH* sequences.

Results

Water and Soil Chemistry

The water and soil chemistry results are summarized in Table 3. Samples from under forested land were not significantly different from those from under pasture-land in regards to the amount of organic carbon (water $p = 0.99$, soil $p = 0.59$) and nitrogen (water $p = 0.49$, soil $p = 0.93$) in material entering the lava cave. None of the samples were statistically significantly different in terms of the amount of SO_4^{2-} ($p=0.80$), NO_3^- ($p=0.72$) and NH_4^+ ($p=0.91$) as determined by the water analysis. The only element that was statistically significantly different was PO_4^- ($p=0.017$).

Diversity of *amoA* Genes

A total of 14 (25%) of the 55 samples tested were positive for the *amoA* gene, from six of the 11 lava caves, three under forested land, one under sea/urban land and two under pasture-land. Most samples had approximately 6–10 bands per lane, although some lanes, such as GBO1-1 from Gruta Branca Opala, had considerably more, with over 15 bands. Seventy-eight bands were excised and sequenced from the 14 samples (Figure 3, Table 4). Using a 97% similarity cutoff, 15 unique OTUs were designated. Ten OTUs were found exclusively in lava caves underlying forested land, one exclusively in a lava cave under a pasture, three were shared between lava caves under pastures and lava caves under forests, and one was shared between lava caves under forests and the sea/urban lava cave (Figure 4A). Thirteen of the 15 unique OTUs were similar to known species of *Nitrosospora*, and the other two OTUs were similar to *Nitrosomonas* sequences (Figure 5). Several of the closest relatives to sequences recovered in this study were uncultured soil bacterium clones. *Nitrosospora*-like sequences were seen in all samples except the one from Gruta das Agulhas, the lava cave that opens into the sea. There were no OTUs shared by all samples.

Diversity of *nifH* Genes

A total of 30 of the 55 (55%) samples screened were positive for the *nifH* gene, from all 11 lava caves surveyed. Most samples had approximately 5–7 bands, but GBO54 and GB1927 had considerably more. One hundred forty-six bands were sequenced, and were classified into 22 unique OTUs using a 97% similarity cut-off (Figure 6, Table 4). Nine of the OTUs were unique to the lava caves under forested land, seven were unique to the lava cave under pasture land, and one OTU was unique to the sea/urban lava cave (Figure 4B).

There were two OTUs shared by the lava caves under forested and pasture lands, one shared by the pasture and sea/urban lava cave, one shared by the forested and sea/urban lava cave, and one OTU shared by all three land use categories, a *Klebsiella pneumoniae*-like sequence. Sixty-one of the 146 bands (41.7%) of the sequences recovered were identified as *K. pneumoniae*-like sequences. *K. pneumoniae*-like sequences were found in 13 samples from eight different lava caves. Twenty-two of the 146 sequences (15.1%) were identified as *Zoogloea oryzae*-like. Nitrogen-fixing bacteria from three phyla of bacteria were found: *Actinobacteria* (2 sequences), *Firmicutes* (12 sequences), and four *Proteobacteria* classes, *Alphaproteobacteria* (25 sequences), *Betaproteobacteria* (38 sequences), *Deltaproteobacteria* (2 sequences) and *Gammaproteobacteria* (67 sequences) (Figure 7). There were no OTUs shared by all samples.

Discussion

The water and soil chemistry showed little statistical differences in the nutrient content of material entering the lava caves based on land use (Table 3). This is probably due to small sample sizes, however the high p values for most nutrients indicate that that samples probably are not different. The only exception is the phosphorus levels, with a $p = 0.017$. In general, soils from forest and cleared pasture sites within the same biome may be expected to differ in nitrogen and carbon content (Cenciani et al. 2009; Chaer et al. 2009; Lauber et al. 2008). However, no clear trends have emerged from these studies as to the magnitude of

soil carbon and nitrogen responses to land use regime. To our knowledge, there are no previous studies that examine the differences in carbon/nitrogen in water that percolates through pasture soil and forested soil into caves. Future studies will gather additional data points with which to explore these parameters under different seasonal time frames.

It is important to note that although the water and soil chemistries were not significantly different for the different land use categories, flux rates were not measured in this study. This may limit the extent to which the molecular results can be interpreted in terms of the water and soil chemistries. However, there were some interesting trends that indicated differences in the diversity of nitrogen cycling genes in the underlying lava caves of different land use types. Lava caves under forested land were more likely to have samples positive for the presence of *amoA* genes than lava caves under pastures; three of the four lava caves under forested land were positive, yet only two of the six lava caves occurring under pastures contained samples positive for *amoA*.

The lava caves under forested land also contained a larger number of OTUs, 14 of the 15 OTUs identified (Figure 4). The same sample from a forested lava cave also had the most bands in the DGGE gel for both *amoA* and *nifH*, a sample from Gruta Branca Opala. Although these results are not conclusive, they suggest that lava caves that lie below forested land, in contrast to those below pastures, contain a greater diversity of AOB. We will investigate whether this AOB diversity under forested lands is also observed in our other field sites in Hawai'i and New Mexico in future studies.

Factors such as the nature of the soil microbial community above the lava cave and the effect of fertilizers on these communities may influence the AOB communities in the lava caves. A future study will compare and contrast the soil microbial communities above the caves with those within the caves. Other factors such as the pH of the soil, vegetation, soil nutrient levels and salinity, have been shown to influence the diversity of *Nitrosospira* in different soil communities (Fierer et al. 2009). The difference in diversity may also be due to a greater flux of ammonia into forested caves due to deeper rooting of the plants above. Further studies are needed to conclusively show this difference, and to document the nitrogen flux into these lava caves.

The lava caves of Terceira were dominated by *Nitrosospira*-like sequences. This result may be influenced by primer bias, but the sea cave did show non-*Nitrosospira*-like sequences, rather than *Nitrosospira*-like sequences. *Nitrosospira* are lithoautotrophic bacteria that commonly dominate terrestrial AOB communities (Hayatsu et al. 2008). As many of the closest relatives to bacteria living in lava cave microbial mats also are found in soil environments, an obvious colonization source for the lava caves, it is reasonable to assume that the AOB communities would follow this trend (Garcia et al. 2009; Moya et al. 2009; Northup et al. 2008; Snider 2010; Snider et al. 2009). However, we have also found that lava caves contain many novel organisms (some less than 85% similarity to known sequences), a finding that we are exploring to determine differences among overlying soil communities with those found in the underlying caves (Hathaway unpublished data). *Nitrosospira* has been found to be more diverse in soils that were not treated with urea fertilizer, compared to those that were (Webster et al. 2002). The use of chemical fertilization and the large number

of cattle in the pastures may be limiting the diversity of the AOB communities in the pasture land lava caves.

Four of the 15 OTUs, representing 54 of the 78 sequences, group within *amoA* Cluster 1, which is commonly found in marine environments. However, studies have also found representatives from this cluster in soil environments (Fierer et al. 2009; Schmidt et al. 2007). This may indicate the strong marine influence on the terrestrial island soils in the lava caves, or may be due to the moist nature of lava caves, most of which have 95–100% humidity year-round. The other OTUs grouped with *Nitrosospira* Clusters 2, 3 and 4, all of which are commonly found in soil environments (Hayatsu et al. 2008) and volcanic soils (Cluster 3) in Hawai'i (King and Nanba 2008).

Although multiple studies have found that several environmental factors can influence the AOB composition, similar ecosystem types do not appear to harbor the same AOB communities (Fierer et al. 2009). Furthermore, Fierer et al. (2009) found no significant correlations between amount of nitrogen available and AOB composition across multiple soil ecosystems. This may indicate that although lava caves are similar on the ecosystem scale, each is unique and will harbor its own specific community of AOB.

The dominance of *Nitrosospira*-like sequences in these lava caves is in contrast to the finding of AOB in Movile Cave, a closed-system carbonate cave in Romania near the Black Sea, which was dominated by *Nitrosomonas*-like sequences (Chen et al. 2009). However, this study collected water samples that included floating microbial mats, not microbial mats from the walls of the cave. Although *Nitrosomonas* can be found in soils, it is often associated with water samples, especially contaminated water. Movile Cave has a unique atmosphere, with high amounts of hydrogen sulfide, carbon dioxide, methane, decreased oxygen, has been shown to be isolated from surface influences (Sarbu et al. 1996), and is a very different environment than that found in the lava caves.

One lava cave, Gruta das Agulhas, had only *Nitrosomonas*-like sequences, with no *Nitrosospira*-like sequences. This lava cave is unique in respect to the other lava caves sampled, as it is located under an urban setting, and opens into the sea. At high tide, waves crash a few meters into the entrance of the lava cave. The sample that was positive for *amoA* was collected approximately 15 m from the entrance of the lava cave. The proximity to the ocean may influence the microbial communities in this region. Kowalchuk et al. (1997) studied the AOB of coastal sand dunes. They found *Nitrosomonas*-like sequences only in dunes closest to the ocean, while *Nitrosospira*-like sequences were found in all dunes, and it is thought that salinity affects the type of AOB present in an ecosystem. The influence of the ocean spray affected the AOB community in the Kowalchuk et al. study (1997), and very well may be influencing the AOB community in Gruta das Agulhas.

The most common sequences recovered from the *nifH* DGGE bands were related to *K. pneumoniae*, which is a Gram-negative bacterium belonging to the *Enterobacteriaceae* of *Gammaproteobacteria*. *K. pneumoniae* has a wide range of habitats including soil, vegetation, water, and as pathogens to many mammals. In nonclinical environments, *K. pneumoniae* is known to contribute to both biochemical and geochemical processes and can

be a major component of the microbiota (Brisse et al. 2006). Hunter et al. (2004) showed that coliforms such as *Escherichia coli* can persist in cave environments, despite the low nutrients available. In a more clinical setting, *K. pneumoniae* commonly infects cattle, and was shown to be shed with fecal matter in up to 80% of cows tested in the U.S. (Brisse and van Duinkerken 2005; Munoz et al. 2006). Due to the large number of cattle that inhabit Terceira, it is not hard to imagine them as a source of the *K. pneumonia* in the caves.

In contrast to our *nifH* finding, no *Klebsiella*-like sequences were recovered in our recent 16S rRNA analyses of the same lava caves (Hathaway, unpublished data). This may be due to the limited number of sequences, approximately 60, from each clone library. Moreover, in a more extensive survey of the caves from Lava Beds National Monument, California, USA using 454 sequencing no *Klebsiella*-like sequences were recovered in microbial mats or surface soils above (Northup unpublished results). However, there are no cattle grazing above these caves. We believe that future 454 sequencing of the Terceira microbial mats will help to answer the question of whether *K. pneumonia* is really present in the caves. The presence of *K. pneumoniae* in the lava caves also highlights one of the limitations of this study. We acknowledge that the putative presence of both *amoA* and *nifH* does not confirm activity of these genes within the lava caves. The *K. pneumonia*-like sequences could be remnant DNA from surface contaminants, and not actual living cells in the caves. Future studies will include next generation sequencing to increase the number of sequences analyzed from each cave. This will allow for more in depth analysis of the species in the cave. Future work should also include RNA studies to determine which, if any, of the genes found in this study are actively being expressed in the caves.

Many of the other closest relatives to the *nifH* sequences recovered are found in soil environments, many of which were uncultured. It is common to find closely related sequences from soil environments in 16S rRNA studies of lava cave microbial communities, and *nifH* genes appear to follow this trend as well (Garcia et al. 2009; Moya et al. 2009; Northup et al. 2008, 2011; Snider et al. 2009; Snider 2010). One close cultured relative that comprised 15% of the recovered sequences was *Zoogloea oryzae*, a nitrogen fixing bacterium isolated from rice paddy soil (Xie and Yokota, 2006). Other species in the *Zoogloea* genus are found in marine environments, as were some of the uncultured closest relatives in this study (Figure 7). This genus also has been found in our recent analyses of microbial mats from Lava Beds National Monument, California, USA (Northup, unpublished results). The shallow nature of lava caves, combined with the probably high hydrologic connectivity and the large amounts of surface rainfall (1000–2300 mm) may also be influencing the community structure of *nifH* organisms, allowing a diverse community with close relatives from both soil and marine environments.

The most diverse samples for *nifH* came from Gruta da Branca Opala (GBO) and Gruta dos Buracos (GB). There is nothing in the soil or water chemistry to suggest that these two lava caves are more similar to each other than any of the other lava caves (Table 3). Furthermore, Gruta da Branca Opala is under forested land, while Gruta dos Buracos is under pasture land. There may be other environmental factors besides water or soil chemistry that may be affecting microbial community composition. Studies of terrestrial lava flow volcanic rock have shown that the chemical composition of the lava rock influences the resulting bacterial

communities (Gomez-Alvarez et al. 2007). Future studies will include analysis of underlying basaltic substrate chemistry as a possible determining factor of microbial diversity.

There were twice as many samples positive for *nifH* as there were for *amoA*. This may be due to the low amounts of ammonia coming into the lava caves in rain-water (Table 3). It is probable that any available nitrogen, from the cattle feces or other sources, is being utilized by the soil plant matter or soil microorganisms, and that little bioavailable nitrogen is entering the lava caves. The lack of bioavailable nitrogen would increase the importance of nitrogen fixing bacteria in the establishment and sustainability of microbial mats, as has been suggested for surface lava flows (King 2007).

This study is, to our knowledge, the first report to describe the *amoA* gene diversity in lava caves, and the first to detect the presence of the *nifH* gene in any cave environment. The *amoA* sequences were mostly *Nitrosospora*-like sequences, which are commonly found in soil, although *Nitrosomonas*-like sequences were also found. Despite the lack of significant differences in the water and soil chemistries in regards to land use, there were unique OTUs for both *amoA* and *nifH* in the pasture and forested caves. *nifH* was more widely found, and was dominated by a *K. pneumoniae*-like sequences, which is most likely coming from bovine fecal matter washing into the lava caves. Although this study does not show that these genes are being expressed, it provides a framework upon which future studies can build and is the first step in piecing together the microbial role in cycling nitrogen within caves. Future studies will explore other nitrogen cycle genes within the caves and compare these findings to the diversity found in soils that overlie the caves, and will investigate whether these processes are actually occurring within the caves.

Acknowledgments

We would like to thank A. Dapkevicius, A.R. Varela, F. Pereira, I.R. Amorim and A.F. Rodrigues for help with field work in Terceira. For permits and access to the lava caves we thank the Os Montanheiros, and the landowners.

Funding

This work would not have been possible without the funding provided by the Graduate Research Allocation Committee at UNM Biology, UNM Biology Grove Scholarship, the Student Research Allocation Committee at UNM, the National Speleological Society, the New Mexico Space Grant Consortium, and Kenneth Ingham Consulting. We acknowledge support from the UNM Molecular Biology Facility, which is supported by NIH grant number P20RR018754. This work was supported by Fundação Ciência e Tecnologia (FCT), under the project PTDC/AMB/70801/2006 (M. L. Dapkevicius, P.I.).

References

- Barton HA, Jurado V. What's up down there? Microbial diversity in caves. *Microbe*. 2007; 2(3):132–138.
- Barton HA, Taylor MR, Pace NR. Molecular phylogenetic analysis of a bacterial community in an oligotrophic cave environment. *Geomicrobiol J*. 2004; 21:11–20.
- Brise, S.; Grimont, F.; Grimont, PAD. The genus *Klebsiella*. In: Dworkin, M.; Falkow, S.; Rosenberg, E.; Schleifer, KH.; Stackebrandt, E., editors. *Prokaryotes: A Handbook on the Biology of Bacteria*. Third Edition. Vol. 6. New York: Springer; 2006. p. 159-196.
- Brise S, van Duijkeren E. Identification and antimicrobial susceptibility of 100 *Klebsiella* animal clinical isolates. *Vet Microbiol*. 2005; 105:307–312. [PubMed: 15708829]

- Cenciani K, Lambais MR, Cerri CC, de Azevedo LCB, Feigl BJ. Bacteria diversity and microbial biomass in forest, pasture and fallow soils in the Southwestern Amazon Basin. *Rev Bras Cienc Solo*. 2009; 33:907–916.
- Chaer G, Fernandes M, Myrold D, Bottomley P. Comparative resistance and resilience of soil microbial communities and enzyme activities in adjacent native forest and agricultural soils. *Microb Ecol*. 2009; 58:414–424. [PubMed: 19330551]
- Chen Y, Wu LQ, Boden R, Hillebrand A, Kumaresan D, Moussard H, Baciú M, Lu Y, Murrell JC. Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME J*. 2009; 3:1093–1104. [PubMed: 19474813]
- Clescerl, LS.; Greenberg, AE.; Eaton, AD., editors. *Standard Methods for the Examination of Water and Wastewater*. Denver, CO: American Public Health Association; 1999.
- Collins SL, Sinsabaugh RL, Crenshaw C, Green L, Porras-Alfaro A, Stursova M, Zeglin LH. Pulse dynamics and microbial processes in aridland ecosystems. *J Ecol*. 2008; 96:413–420.
- Dean, DR.; Jacobson, MR. Biochemical genetics of nitrogenase. In: Stacey, G.; Burris, RH.; Evans, HJ., editors. *Biological Nitrogen Fixation*. New York: Chapman and Hall; 1992. p. 763-834.
- Demba-Diallo M, Willems A, Vloemans N, Cousin S, Vandekerckhove TT, de Lajudie P, Neyra M, Vyverman W, Gillis M, Van der Gucht K. Polymerase chain reaction denaturing gradient gel electrophoresis analysis of the N₂-fixing bacterial diversity in soil under *Acacia tortilis ssp. raddiana* and *Balanites aegyptiaca* in the dryland part of Senegal. *Environ Microbiol*. 2004; 6:400–415. [PubMed: 15008817]
- Dichosa, A. Biogenicity and Microbial Community Composition of Desert Varnish and Cave Ferromanganese Deposits. Ph.D. Dissertation. Albuquerque, NM: University of New Mexico; 2008.
- Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucl Acids Res*. 2004; 32:1792–1797. [PubMed: 15034147]
- Fierer N, Carney KM, Horner-Davine MC, Megonigal JP. The biogeography of ammonia oxidizing bacterial communities in soil. *Microb Ecol*. 2009; 58:435–445. [PubMed: 19352770]
- Fliermans CB, Schmidt EL. *Nitrobacter* in Mammoth Cave. *Int J Speleol*. 1977; 9:1–19.
- Garcia MG, Moya M, Spilde MN, Stone FD, Northup DE. Discovering new diversity in Hawaiian lava tube microbial mats. *P Int Cong Speleol*. 2009; 1:364–369.
- Giovannoni SJ, Delong EF, Schmidt TM, Pace NR. Tangential flow filtration and preliminary phylogenetic analysis of marine Picoplankton. *Appl Environ Microb*. 1990; 56:2572–2575.
- Gomez-Alvarez V, King GM, Nusslein K. Comparative bacterial diversity in recent Hawaiian volcanic deposits of different ages. *FEMS Microbiol Ecol*. 2007; 60:60–73. [PubMed: 17381525]
- Hayatsu M, Tago K, Saito M. Various players in the nitrogen cycle: Diversity and functions of the microorganisms involved in nitrification and denitrification. *Soil Plant Nutr*. 2008; 54:33–45.
- Howarth, FG. A comparison of volcanic and karstic cave communities. In: Oromi, P., editor. In: *Proceedings of the 7th International Symposium on Vulcanspeleology*, Canary Islands, November 1994. Barcelona: Forimpres, S.A.; 1996. p. 63-68.
- Hunter AJ, Northup DE, Dahm CN, Boston PJ. Persistent Coliform contamination in Lechuguilla Cave pools. *J Cave Karst Stud*. 2004; 66(3):102–110.
- Jones DS, Lyon EH, Macalady JL. Geomicrobiology of bioer-miculations from the Frasassi Cave system, Italy. *J Cave Karst Stud*. 2008; 70:78–93.
- King GM. Contributions of atmospheric CO and hydrogen uptake to microbial dynamics on recent Hawaiian volcanic deposits. *Appl Environ Microbiol*. 2003; 69:4067–4075. [PubMed: 12839783]
- King GM. Chemolithotrophic bacteria: Distributions, functions and significance in volcanic environments. *Microbes Environ*. 2007; 22:309–319.
- King GM, Nanba K. Distribution of atmospheric methane oxidation and methanotrophic communities on Hawaiian volcanic deposits and soils. *Microbes Environ*. 2008; 23(4):326–330. [PubMed: 21558726]
- Kowalchuk GA, Stephen JR. Ammonia-oxidizing bacteria: A model for molecular microbial ecology. *Ann Rev Microbiol*. 2001; 55:485–529. [PubMed: 11544365]

- Kowalchuk GA, Stephen JR, DeBoer W, Prosser JI, Embley TM, Woldendorp JW. Analysis of ammonia-oxidizing bacteria of the beta subdivision of the class Proteobacteria in coastal sand dunes by denaturing gradient gel electrophoresis and sequencing of PCR-amplified 16S ribosomal DNA fragments. *Appl Environ Microb*. 1997; 63:1489–1497.
- Laubert CL, Strickland MS, Bradford MA, Fierer N. The influence of soil properties on the structure of bacterial and fungal communities across land-use types. *Soil Biol Biochem*. 2008; 40:2407–2415.
- Madeira, M.; Pinheiro, J.; Madruga, J.; Monteiro, F. Soils of volcanic systems in Portugal. In: Arnalds, Ó.; Óskarsson, H.; Bartoli, F.; Buurman, PGS.; García-Rodeja, E., editors. *Soils of Volcanic Regions in Europe*. Berlin: Springer; 2007. p. 69–81.
- Mason OU, Di Meo-Savoie CA, Van Nostrand JD, Zhou JZ, Fisk MR, Giovannoni SJ. Prokaryotic diversity, distribution, and insights into their role in biogeochemical cycling in marine basalts. *ISME J*. 2009; 3:231–242. [PubMed: 18843298]
- McCarthy JJ, Yilmaz A, Coban-Yildiz Y, Nevins JL. Nitrogen cycling in the offshore waters of the Black Sea. *Estuar Coast Shelf S*. 2007; 74:493–514.
- Moya M, Garcia MG, Spilde MN, Northup DE. Composition of bacterial mats in El Malpais, National Monument, New Mexico, USA: Comparison and Contrasts with Bacterial Communities in Hawai'i Lava Tubes. *Proceeding P Int Cong Speleol*. 2009; 2:709–713.
- Munoz MA, Ahlstrom C, Rauch BJ, Zadoks RN. Fecal shedding of *Klebsiella pneumoniae* by dairy cows. *J Dairy Sci*. 2006; 89:3425–3430. [PubMed: 16899675]
- Northup DE, Barns SM, Yu LE, Spilde MN, Schelble RT, Dano KE, Crossey LJ, Connolly CA, Boston PJ, Natvig DO, Dahm CN. Diverse microbial communities inhabiting ferromanganese deposits in Lechuguilla and Spider Caves. *Environ Microbiol*. 2003; 5:1071–1086. [PubMed: 14641587]
- Northup, DE.; Connolly, CA.; Trent, A.; Peck, VM.; Spilde, MN.; Welbourn, WC.; Natvig, DO. *AMCS Bull*. Vol. 19. New Mexico, USA: El Malpais National Monument; 2008. The nature of bacterial communities in Four Windows Cave; p. 119–125.
- Northup DE, Melim L, Spilde MN, Hathaway JJM, Garcia M, Moya M, Stone F, Boston P, Dapkevicius M. Lava cave microbial communities within mats and secondary mineral deposits: Implications for life detection on other planets. *Astrobiol*. 2011; 11(7):601–618.
- Palmer, AN. *Cave Geology*. Dayton, OH: Cave Books; 2007.
- Pella E. Elemental organic analysis. Part 1. Historical developments. *Am Lab*. 1990a; 22:116–125.
- Pella E. Elemental organic analysis. Part 2. State of the art. *Am Lab*. 1990b; 22:28–32.
- Pfaff, JD.; Hautman, DP.; Munch, DJ. Method 300.1 Determination of inorganic anions in drinking water by ion chromatography. Cincinnati, OH: National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency; 1997.
- Poly F, Monrozier LJ, Bally R. Improvement in the RFLP procedure for studying the diversity of *nifH* genes in communities of nitrogen fixers in soil. *Res Microbiol*. 2001; 152:95–103. [PubMed: 11281330]
- Prosser JI, Nicol GW. Relative contributions of archaea and bacteria to aerobic ammonia oxidation in the environment. *Environ Microb*. 2008; 10:2931–2941.
- Ross DJ, Scott NA, Lambie SM, Trotter CM, Rodda NJ, Townsend JA. Nitrogen and carbon cycling in a New Zealand pumice soil under a manuka (*Leptospermum scoparium*) and kanuka (*Kunzea ericoides*) shrubland. *Aust J Soil Res*. 2009; 47:725–736.
- Rothauwe JH, Witzel KP, Liesack W. The ammonia monooxygenase structural gene *amoA* as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. *Appl Environ Microbiol*. 1997; 63:4704–4712. [PubMed: 9406389]
- Sarbu SM, Kane TC, Kinkle BK. A chemoautotrophically based cave ecosystem. *Science*. 1996; 272(5270):1953–1955. [PubMed: 8662497]
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. Introducing mothur: Open source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 2009; 75:7537–7541. [PubMed: 19801464]

- Schmidt CS, Hultman K, Robinson D, Killham K, Prosser JI. PCR profiling of ammonia-oxidizer communities in acidic soils subjected to nitrogen and sulphur deposition. *FEMS Microbiol Ecol.* 2007; 61:305–316. [PubMed: 17573939]
- Shen J, Zhang L, Zhu Y, Zhang J, He J. Abundance and composition of ammonia-oxidizing bacteria and ammonia-oxidizing archaea communities of an alkaline sandy loam. *Environ Microbiol.* 2008; 10:1601–1611. [PubMed: 18336563]
- Simon KS, Benfield EF. Ammonium retention and whole-stream metabolism in cave streams. *Hydrobiologia.* 2002; 482:31–39.
- Simon KS, Pipan T, Culver DC. A conceptual model of the flow and distribution of organic carbon in caves. *J Cave Karst Stud.* 2007; 69:279–284.
- Simonet P, Grojean MC, Misra AK, Nazaret S, Cournoyer B, Normand P. Frankia genus-specific characterization by polymerase chain reaction. *Appl Environ Microbiol.* 1991; 57:3278–3286. [PubMed: 1781685]
- Snider, J. Comparison of microbial communities on roots, ceilings and floors of two lava tube caves in New Mexico. Master's Thesis. Albuquerque, NM: University of New Mexico; 2010.
- Snider JR, Moya M, Garcia MG, Spilde MN, Northup DE. Identification of the microbial communities associated with roots in lava tubes in New Mexico and Hawai'i. *P Int Cong Speleol.* 2009; 2:718–723.
- Sooksa-nguan T, Thies JE, Gypmantasiri P, Boonkerd N, Teaumroong N. Effect of rice cultivation systems on nitrogen cycling and nitrifying bacterial community structure. *Appl Soil Ecol.* 2009; 43:139–149.
- Spear JR, Barton HA, Robertson CE, Francis CA, Pace NR. Microbial community biofabrics in a geothermal mine adit. *Appl Environ Microbiol.* 2007; 73:6172–6180. [PubMed: 17693567]
- Staley JT, Crawford R. The biologist's chamber: Lava tube slime. *Cascade Caver.* 1975; 14:20–21.
- Stoner, MF.; Howarth, FG. Community structure and niche differentiation in Hawaiian lava tubes. In: Dombois, DM.; Bridges, KW.; Carson, HL., editors. *Island Ecosystems: Biological Organization in Selected Hawaiian Communities.* Stroudsburg, PA: Hutchinson Ross Publ. Co.; 1981. p. 318–336.
- Technicon. *Technicon Industrial Methods.* Tarrytown, NY: Technicon Industrial Systems; 1973. Ammonia in water and wastewater.
- Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, Schleper C. Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ Microbiol.* 2005; 7:1985–1995. [PubMed: 16309395]
- Ueda T, Suga Y, Yahiro N, Matsuguchi T. Remarkable N₂-fixing bacterial diversity detected in rice roots by molecular evolutionary analysis of nifH gene sequences. *J. Bacteriol.* 1995; 177:1414–1417. [PubMed: 7868622]
- Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu DY, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO. Environmental genome shotgun sequencing of the Sargasso Sea. *Science.* 2004; 304:66–74. [PubMed: 15001713]
- Webster G, Embley TM, Prosser JI. Grassland management regimens reduce small-scale heterogeneity and species diversity of beta-proteobacterial ammonia oxidizer populations. *Appl Environ Microbiol.* 2002; 68:20–30. [PubMed: 11772604]
- Xie CH, Yokota A. *Zoogloea oryzae* sp. nov., a nitrogen-fixing bacterium isolated from paddy soil, and reclassification of strain ATCC 19263 as *Crabtreeella saccharophila* gen. nov., sp. nov. *ISME J.* 2006; 56:619–624.

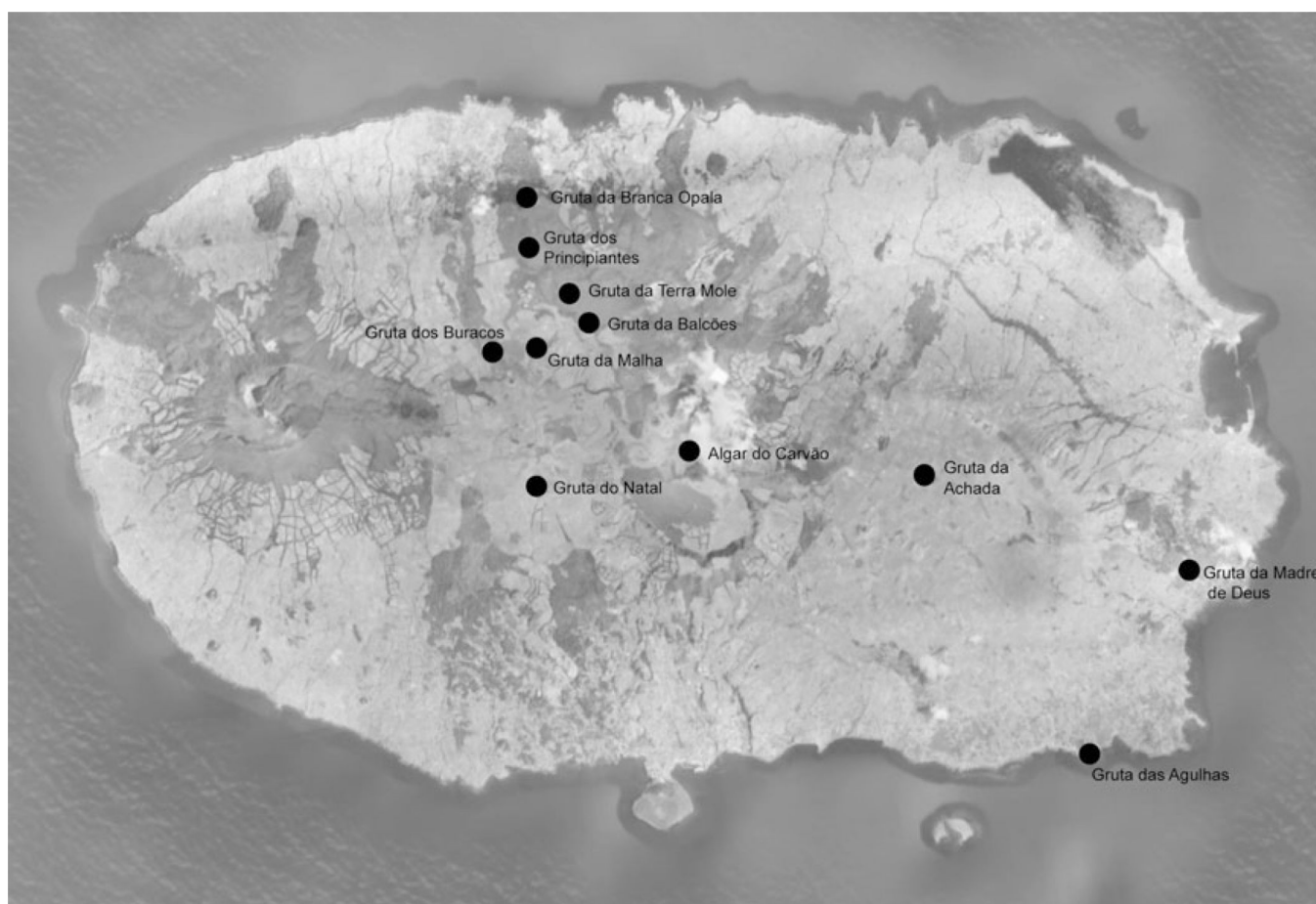


Fig. 1.
Map of Terceira Island, with the location of the lava caves.



Fig. 2.

Images showing different land use practices above the cave and bacterial mats within the lava caves of Terceira. (A) Above Gruta da Madre de Deus; (B) Bacterial mat within Gruta da Madre de Deus; (C) Above Gruta da Balcões; (D) Bacterial mat within Gruta da Balcões. © Kenneth Ingham. Reproduced by permission of Kenneth Ingham. Permission to reuse must be obtained from the rightsholder.

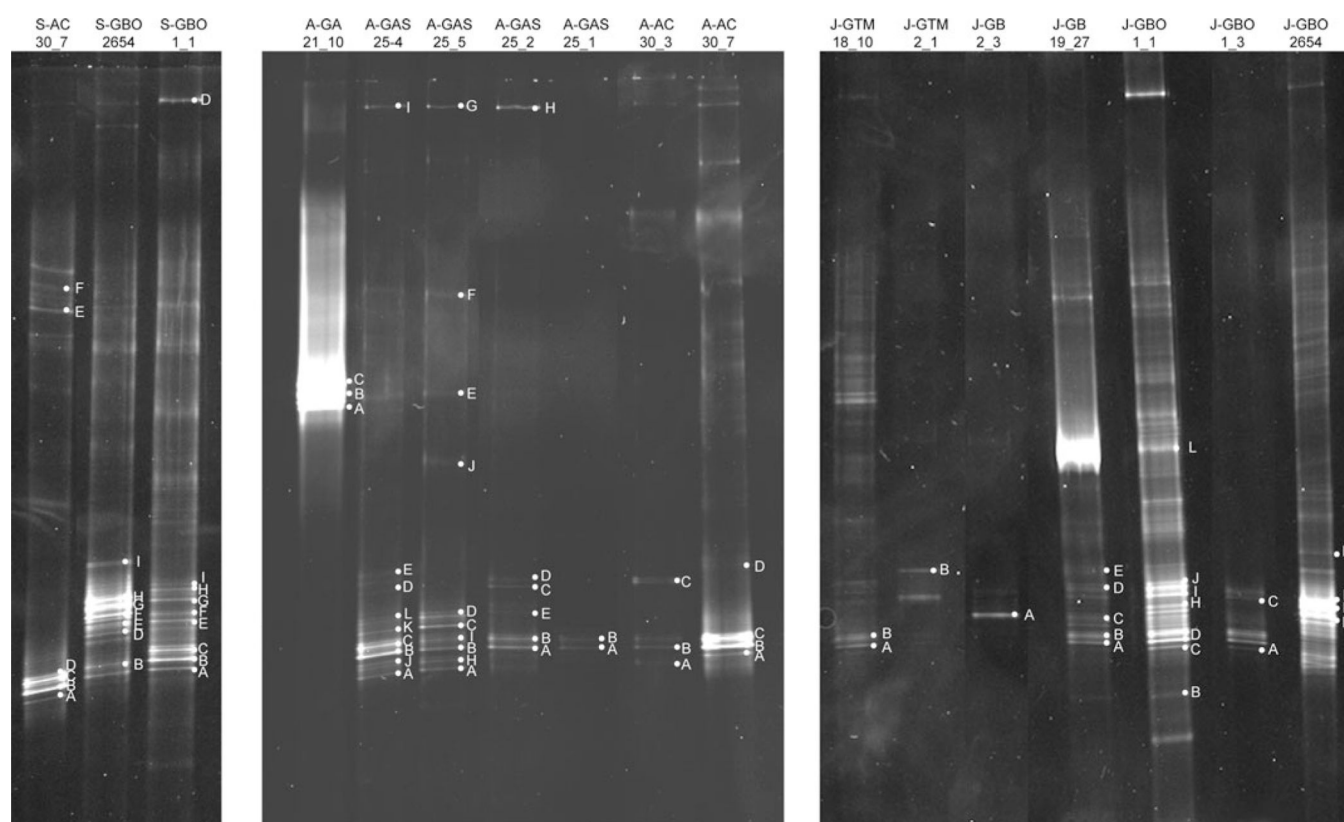


Fig. 3.
 DGGE image of *amoA* gels. Sample names are along the top and individual bands sequenced are labeled within each gel. S, A, and J designate different gels runs.

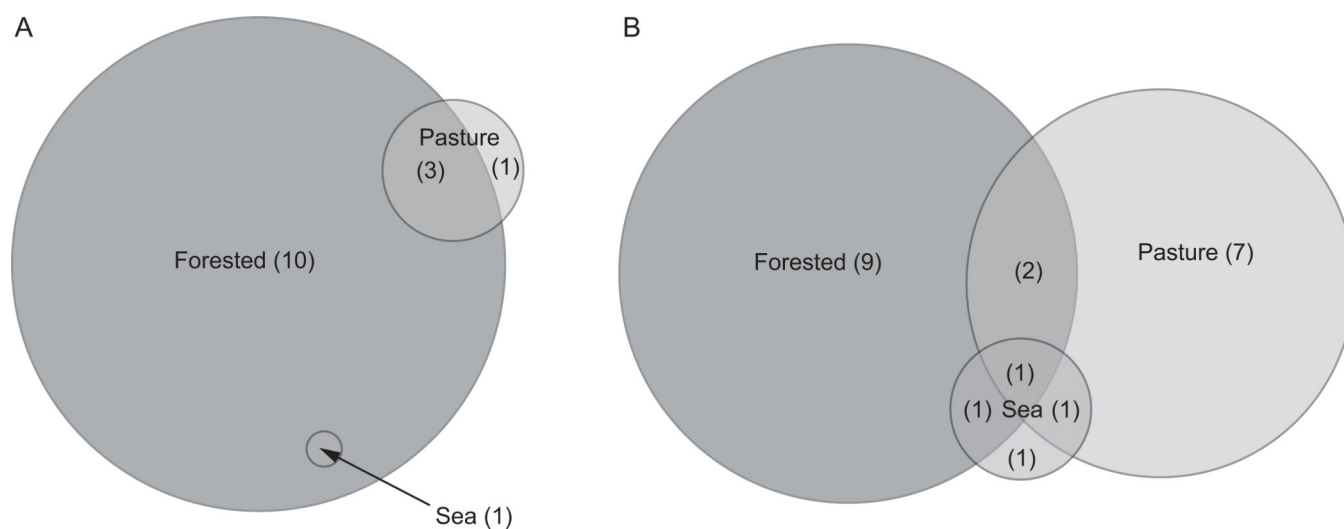
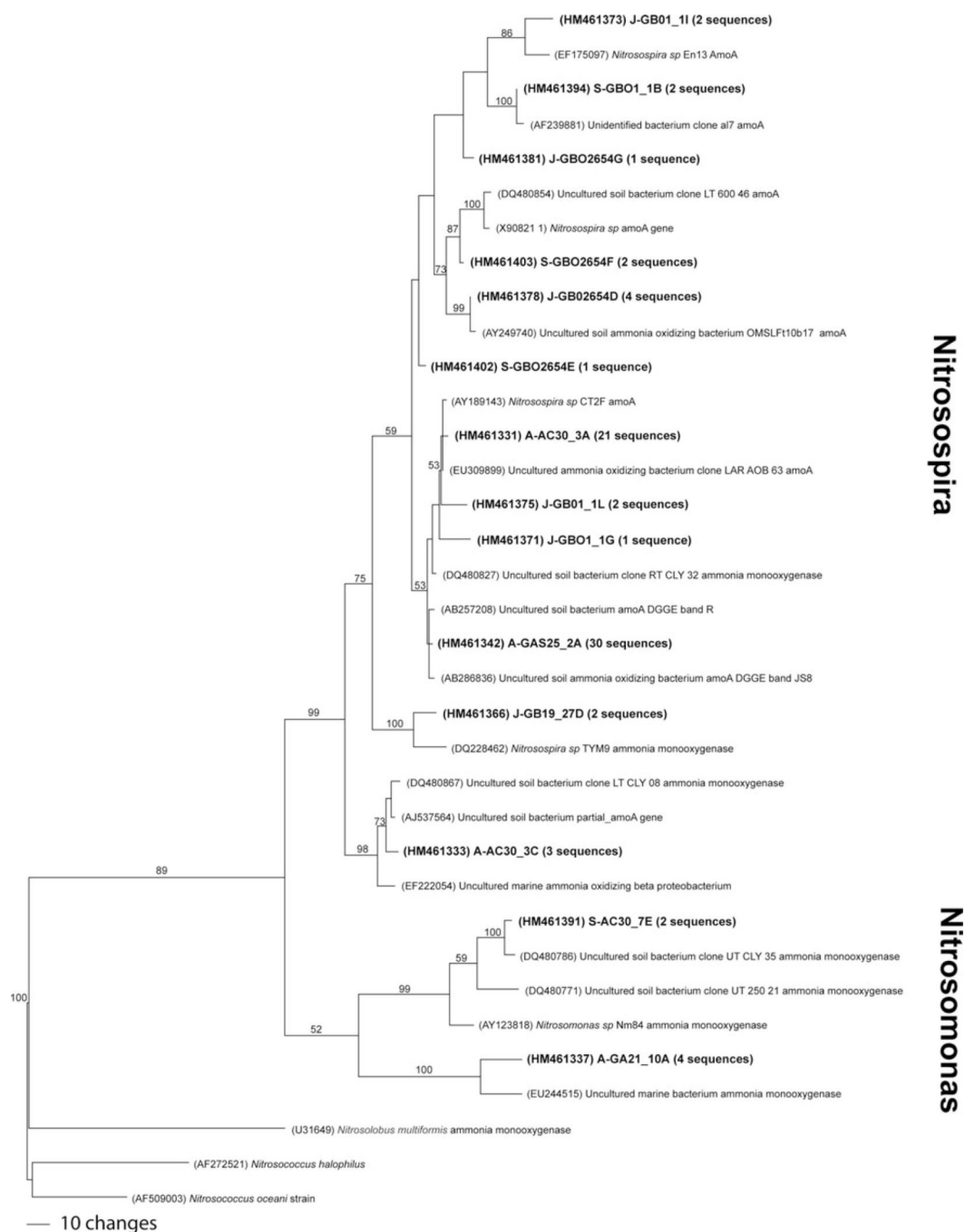


Fig. 4. Venn diagrams showing unique and overlapping OTUs for (A) *amoA* and (B) *nifH* DGGE results, based on the land use above the lava cave. The number in the parentheses indicated the number of unique OTUs for that land use category.

**Fig. 5.**

One of the most parsimonious trees of *amoA* OTUs. Bootstrap values for 1000 replicates are shown. Number in parentheses indicates how many sequences fell within the OTU.

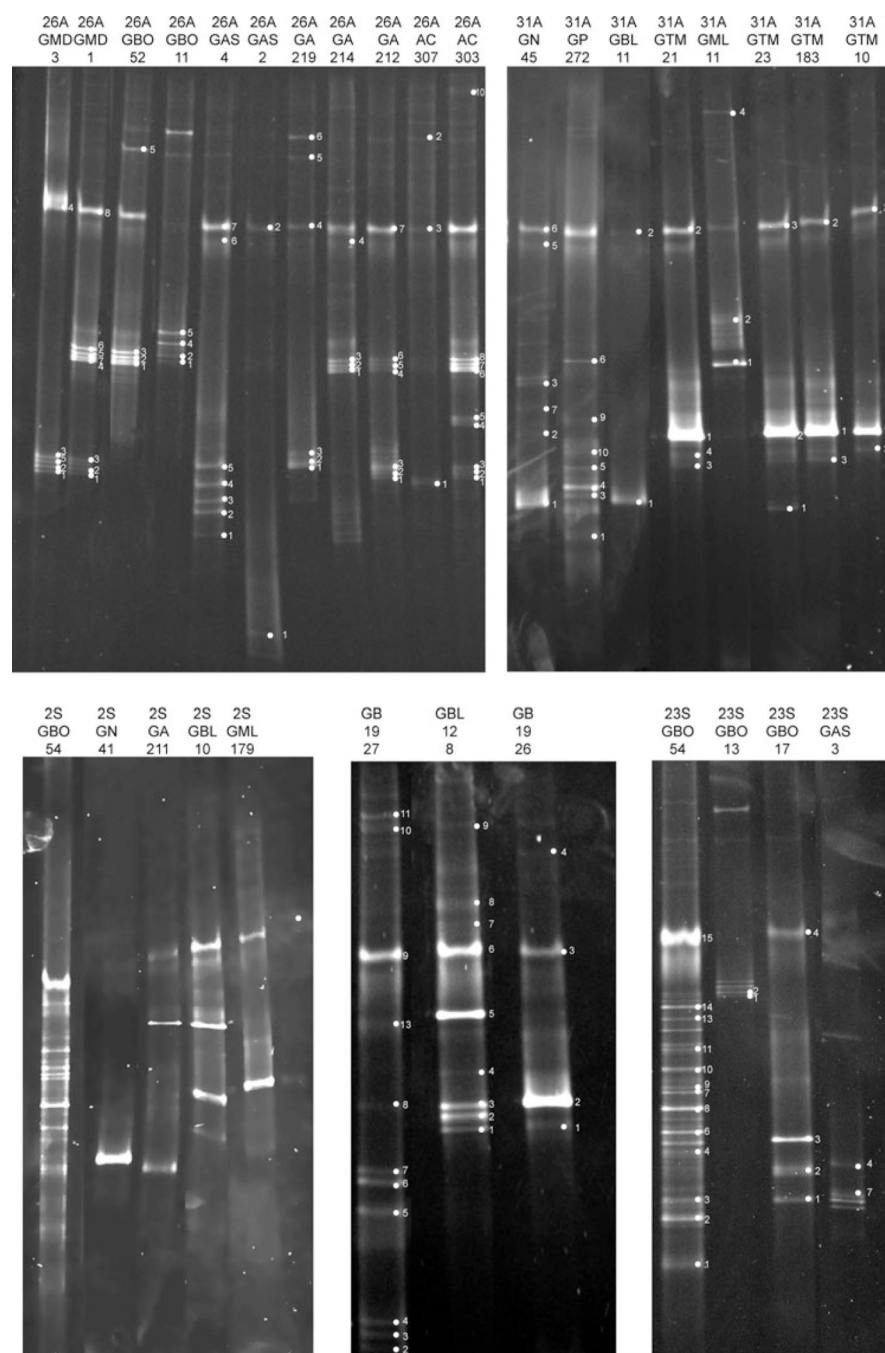
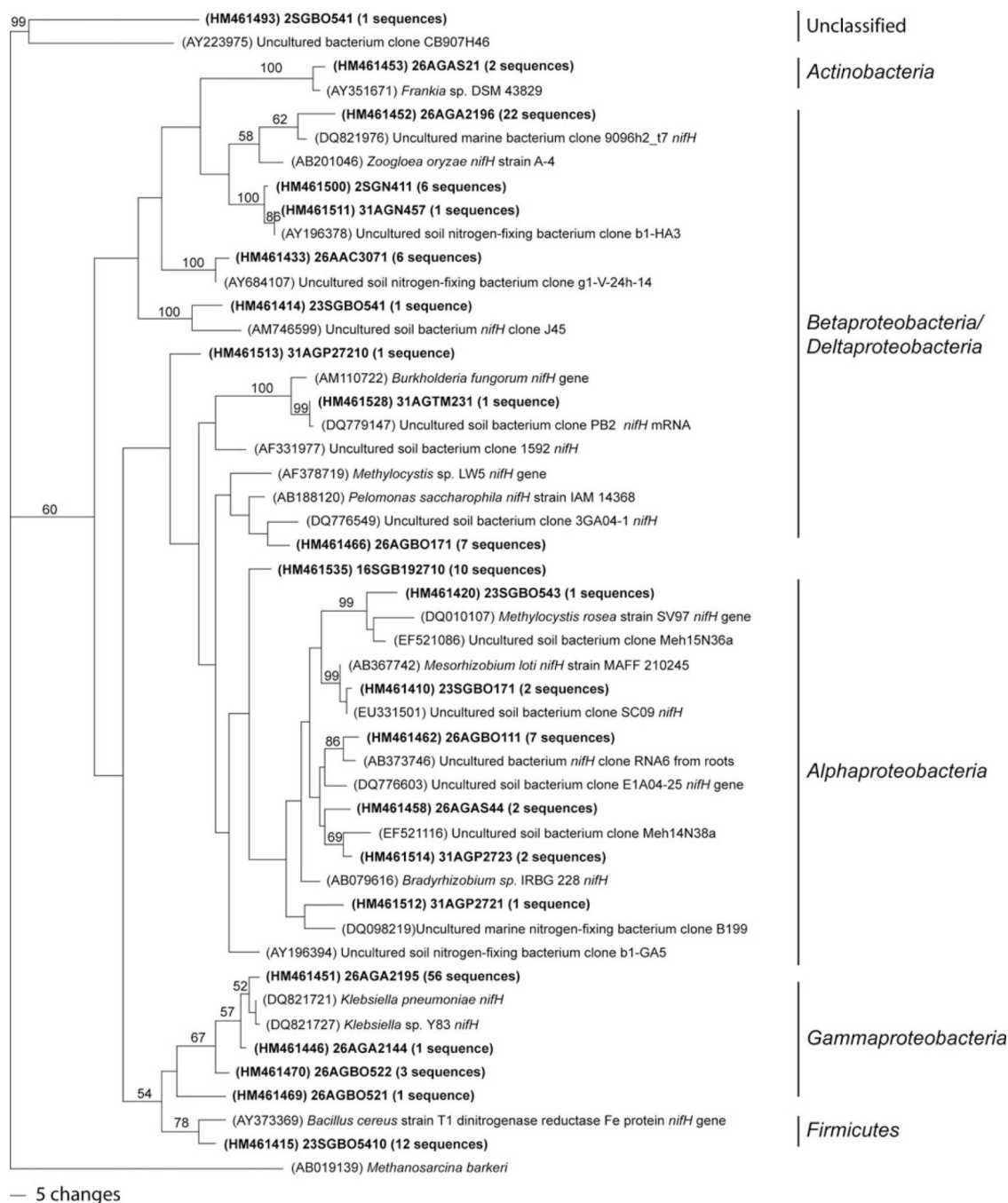


Fig. 6.
 DGGE image of *nifH* gels. Sample names are along the top and individual bands sequenced are labeled within each gel. 26A, 31A, 2S, 16S and 23S designate different gel runs.

**Fig. 7.**

One of the most parsimonious trees of *nifH* OTUs. Bootstrap values for 1000 replicates are shown. Number in parentheses indicates how many sequences fell within the OTU.

Table 1

Abiotic factors associated with each cave

| Cave Name | Elevation (m) | Surface Precipitation (mm) | Temperature (°C) | Relative Humidity (%) | Land Use |
|-------------------------|---------------|----------------------------|------------------|-----------------------|-----------|
| Algar do Carvão | 585 | 2303 | 11.3 | N/A | Forested |
| Gruta da Branca Opala | 255 | 1400 | 15.0 | 95.2 | Forested |
| Gruta da Madre de Deus | 59 | 1050 | 14.6 | 100.0 | Forested |
| Gruta da Achada | 330 | 1635 | 14.9 | 98.5 | Forested |
| Gruta das Agulhas | 1 | 1015 | 22.1 | N/A | Sea/Urban |
| Gruta do Natal | 551 | 2253 | 15.7 | N/A | Pasture |
| Gruta da Terra Mole | 387 | 1809 | 14.9 | 99.8 | Pasture |
| Gruta da Malha | 507 | 2135 | 15.5 | 99.0 | Pasture |
| Gruta dos Buracos | 475 | 2034 | 15.7 | 99.4 | Pasture |
| Gruta dos Principiantes | 346 | 1728 | 15.4 | 98.5 | Pasture |
| Gruta da Balcões | 422 | 1967 | 16.0 | 99.5 | Pasture |

Land use describes how the land above the lava cave is used by humans or the nature of the vegetation above the lava cave where no human use is present.

Table 2

Primers and thermocycling conditions

| Primer | Sequence 5' to 3' | Length of amplicon | Amplification conditions | Reference |
|-------------------------|--|--------------------|--|--|
| amoA1F | GGGGTTTCTACTGGTGGT | 490 | 5 min at 95°C, followed by 35 cycles of 45 s at 94°C, 45 s at 55°C, 1 min at 72°C, with a final extension at 72°C for 10 minutes | Rotthauwe et al. (1997) |
| amoA2R | CCTCKGSAAAGCCTTCTTC | | | Rotthauwe et al. (1997) |
| amoA-1F-GC | CGCCCGCCGCGCCCGC GCCCCGCCCCGCCGCC CCGCCCGGGGTTTC TACTGGTGGT | | | Shen et al.(2008) |
| FGPH19 | TACGGCAARGGTGGNATHG | 432 | 2 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 57°C, 2 min at 72°C, final extension of 5 min at 72°C | Simonet et al. (1991) |
| PolR | ATSGCCATCATYTCRCCGGA | | | Poly et al. (2001) |
| PolF AQER PolF GC | TGCGAYCCSARGCBGACTC GACGATGTAGATYTCCTG CGCCCGCCGCGCGGCGGGC GGGGCGGGGGCACG GGGGGTGCGAYCCSAR GCBGACTC | 321 | 2 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 57°C, 2 min at 72°C, final extension of 5 min at 72°C | Poly et al. (2001) Poly et al. (2001) Demba-Diallo et al. (2004) |

Table 3

Water and soil chemistry

| Cave | Abbreviation | Water | | | | | Soil | | | Land Use | |
|-------------------------|--------------|------------------------------|-------------------------------|------|------|------------------------------|------------------------------|-------|--------|----------|-----------|
| | | PO ₄ ⁻ | SO ₄ ²⁻ | TOC | IC | NO ₃ ⁻ | NH ₄ ⁺ | TN | %N | | %C |
| Algar do Carvão | AC | 0.0374 | 4.10 | 0.86 | 4.14 | 0.08 | 0.060 | 0.31 | 0.1414 | 3.012 | Forested |
| Gruta da Branca Opala | GBO | 0.0312 | 1.42 | 1.24 | 0.16 | 2.12 | 0.040 | 2.46 | 0.8104 | 8.494 | Forested |
| Gruta da Madre de Deus | GMD | 0.0696 | 5.55 | 5.40 | 2.37 | 0.86 | 0.020 | 1.38 | 0.1480 | 2.056 | Forested |
| Gruta da Achada | GAS | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | Sea/Urban |
| Gruta das Agulhas | GA | 0.0035 | 37.29 | 1.10 | 2.35 | 6.92 | 0.090 | 10.20 | 0.0930 | 1.397 | Pasture |
| Gruta do Natal | GN | 0.0350 | 6.36 | 2.32 | 4.82 | 0.75 | 0.080 | 1.86 | 0.1873 | 1.998 | Pasture |
| Gruta da Terra Mole | GTM | 0.0199 | 5.43 | 1.19 | 0.09 | 1.03 | 0.020 | 1.26 | 0.7076 | 10.151 | Pasture |
| Gruta da Malha | GML | 0.0027 | 3.54 | 1.45 | 2.11 | 3.75 | 0.040 | 5.96 | 0.1691 | 3.864 | Pasture |
| Gruta dos Buracos | GB | 0.0029 | 1.87 | 8.10 | 2.80 | 0.28 | 0.030 | 1.08 | 0.2342 | 4.559 | Pasture |
| Gruta dos Principiantes | GP | 0.0013 | 1.94 | 0.78 | 1.85 | 1.78 | 0.030 | 2.20 | 0.3047 | 4.357 | Pasture |
| Gruta da Balcões | GBL | 0.0030 | 5.19 | 1.17 | 2.29 | 0.46 | 0.030 | 1.17 | 0.6622 | 7.871 | Pasture |

Table 4

| A. <i>amoA</i> DGGE band identification | | | | | |
|---|---------------------------|-------------------------------|--|----------------------|--|
| Sample Name | Number of Bands Sequenced | Number of Phylotypes Obtained | Phylogenetic Groups Identified | Phylogroup Number(s) | Phylogroup Accession Number(s) |
| AC30-3 | 3 | 2 | <i>Nitrosospora</i> , <i>Nitrosospora</i> | 1,2 | HM461331, HM461333 |
| AC30-7 | 6 | 2 | <i>Nitrosospora</i> , <i>Nitrosomonas</i> | 1, 11 | HM461331, HM461391 |
| GA21-10 | 3 | 1 | <i>Nitrosomonas</i> | 3 | HM461337 |
| GAS25-1 | 2 | 1 | <i>Nitrosospora</i> | 1 | HM461331 |
| GAS25-2 | 5 | 1 | <i>Nitrosospora</i> | 4 | HM461342 |
| GAS25-4 | 7 | 1 | <i>Nitrosospora</i> | 4 | HM461342 |
| GAS25-5 | 7 | 2 | <i>Nitrosospora</i> , <i>Nitrosomonas</i> | 3,4 | HM461337, HM461342 |
| GB02-3 | 1 | 1 | <i>Nitrosospora</i> | 2 | HM461333 |
| GB19-27 | 5 | 3 | <i>Nitrosospora</i> , <i>Nitrosospora</i> , <i>Nitrosospora</i> | 1,4,8 | HM461331, HM461342, HM461366 |
| GB001-1 | 16 | 6 | <i>Nitrosospora</i> , <i>Nitrosospora</i> , <i>Nitrosospora</i> , <i>Nitrosospora</i> | 1,4,5,6, 9,12 | HM461331, HM461342, HM461375, HM461394 |
| GB001-3 | 1 | 1 | <i>Nitrosospora</i> | 4 | HM461342 |
| GB026-54 | 16 | 5 | <i>Nitrosospora</i> , <i>Nitrosospora</i> , <i>Nitrosospora</i> | 4,7,10, 13, 14 | HM461342, HM461378, HM461381 |
| GTM18-10 | 2 | 1 | <i>Nitrosospora</i> | 1 | HM461331 |
| GTM2-1 | 1 | 1 | <i>Nitrosospora</i> | 2 | HM461333 |
| B. <i>nifH</i> DGGE band identification | | | | | |
| Sample Name | Number of Bands Sequenced | Number of Phylotypes Obtained | Phylogenetic Groups Identified | Phylogroup Number(s) | Phylogroup Accession Number(s) |
| AC30-3 | 9 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7 | HM461451 |
| AC30-7 | 3 | 1 | <i>Gammaproteobacteria</i> | 12 | HM461433 |
| GA21-1 | 3 | 1 | <i>Gammaproteobacteria</i> | 12 | HM461433 |
| GA21-2 | 7 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7 | HM461451 |
| GA21-4 | 4 | 2 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7,8 | HM461451, HM461446 |
| GA21-9 | 6 | 2 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Gammaproteobacteria</i> | 7,22 | HM461451, HM461452 |
| GAS25-2 | 2 | 1 | <i>Actinobacteria</i> (<i>Frankia</i>) | 2 | HM461453 |

| B. <i>nifH</i> DGGE band identification | | | | | |
|---|---------------------------|-------------------------------|--|----------------------|--|
| Sample Name | Number of Bands Sequenced | Number of Phylotypes Obtained | Phylogenetic Groups Identified | Phylogroup Number(s) | Phylogroup Accession Number(s) |
| GAS25-3 | 2 | 1 | <i>Betaproteobacteria</i> (<i>Burkholderiales</i>) | 15 | HM461466 |
| GAS25-4 | 7 | 2 | <i>Alphaproteobacteria</i> , <i>Alphaproteobacteria</i> | 20,21 | HM461462, HM461458 |
| GB19-26 | 4 | 2 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Gammaproteobacteria</i> | 7,22 | HM461451, HM461452 |
| GB19-27 | 10 | 1 | <i>Betaproteobacteria</i> | 16 | HM461535 |
| GBL12-10 | 4 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7 | HM461451 |
| GBL12-11 | 2 | 1 | <i>Betaproteobacteria</i> (<i>Burkholderiales</i>) | 15 | HM461466 |
| GBL12-8 | 9 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7 | HM461451 |
| GBO01-1 | 4 | 2 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Alphaproteobacteria</i> | 7,20 | HM461451, HM461462 |
| GBO01-3 | 2 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 5 | HM461415 |
| GBO01-7 | 7 | 4 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Deltaproteobacteria</i> , <i>Alphaproteobacteria</i> | 7,9,15,18 | HM461451, HM461470, HM461466, HM461410 |
| GBO26-52 | 4 | 3 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 6,7,9 | HM461469, HM461451, HM461470 |
| GBO26-54 | 10 | 4 | Unclassified, <i>Deltaproteobacteria</i> , <i>Firmicutes</i> (<i>Bacillus</i>), <i>Betaproteobacteria</i> | 1,3,5,13 | HM461493, HM461414, HM461415, HM461420 |
| GMD20-1 | 8 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7 | HM461451 |
| GMD20-3 | 5 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7 | HM461451 |
| GML17-11 | 3 | 1 | <i>Gammaproteobacteria</i> | 22 | HM461452 |
| GML17-9 | 2 | 1 | <i>Gammaproteobacteria</i> (<i>Klebsiella</i>) | 7 | HM461451 |
| GN04-1 | 1 | 1 | <i>Gammaproteobacteria</i> | 10 | HM461500 |
| GN4-5 | 5 | 2 | <i>Gammaproteobacteria</i> , <i>Gammaproteobacteria</i> | 10,11 | HM461500, HM461511 |
| GP27-2 | 6 | 5 | <i>Betaproteobacteria</i> , <i>Gammaproteobacteria</i> (<i>Klebsiella</i>), <i>Alphaproteobacteria</i> , <i>Alphaproteobacteria</i> , <i>Alphaproteobacteria</i> | 4,7,17,19,21 | HM461513, HM461451, HM461512, HM461516, HM461458 |
| GTM2-1 | 4 | 1 | <i>Gammaproteobacteria</i> | 22 | HM461452 |

| B. <i>nifH</i> DGGE band identification | | | | | |
|---|---------------------------|-------------------------------|--|---------------------|-------------------------------|
| Sample Name | Number of Bands Sequenced | Number of Phylotypes Obtained | Phylogenetic Groups Identified | Phylotype Number(s) | Phylotype Accession Number(s) |
| GTM2-3 | 3 | 2 | <i>Betaproteobacteria (Burkholderiales), Gammaproteobacteria</i> | 14, 22 | HM461528, HM461452 |
| GTM018-10 | 3 | 1 | <i>Gammaproteobacteria</i> | 22 | HM461452 |
| GTM18-3 | 3 | 1 | <i>Gammaproteobacteria</i> | 22 | HM461452 |