NRC Research Press

ARTICLE

Endophytic microbial diversity in coffee cherries of *Coffea arabica* from southeastern Brazil

Marcelo N.V. Oliveira, Thiago M.A. Santos, Helson M.M. Vale, Júlio C. Delvaux, Alexander P. Cordero, Alessandra B. Ferreira, Paulo S.B. Miguel, Marcos R. Tótola, Maurício D. Costa, Célia A. Moraes, and Arnaldo C. Borges

Abstract: The microbiota associated with coffee plants may play a critical role in the final expression of coffee quality. However, the microbial diversity in coffee cherries is still poorly characterized. Here, we investigated the endophytic diversity in cherries of *Coffea arabica* by using culture-independent approaches to identify the associated microbes, ultimately to better understand their ecology and potential role in determining coffee quality. Group-specific 16S rRNA and 26S rRNA genes polymerase chain reaction – denaturing gradient gel electrophoresis and clone library sequencing showed that the endophytic community is composed of members of the 3 domains of life. Bacterial sequences showing high similarity with cultured and uncultured bacteria belonged to the *Betaproteobacteria*, *Gammaproteobacteria*, and *Firmicutes* phyla. Phylogenetic analyses of cloned sequences from *Firmicutes* revealed that most sequences fell into 3 major genera: *Bacillus*, *Staphylococcus*, and *Paenibacillus*. Archaeal sequences revealed the presence of operational taxonomic units belonging to *Euryarchaeota* and *Crenarchaeota* phyla. Sequences from endophytic yeast were not recovered, but various distinct sequences showing high identity with filamentous fungi were found. There was no obvious correlation between the microbial composition and cultivar or geographic location of the coffee plant. To the best of our knowledge, this is the first report demonstrating internal tissue colonization of plant fruits by members of the Archaea domain. The finding of archaeal small-subunit rRNA in coffee cherries, although not sufficient to indicate their role as active endophytes, certainly expands our perspectives toward considering members of this domain as potential endophytic microbes.

Key words: coffee cherries, endophytes, PCR–DGGE, rRNA libraries.

Résumé : Le microbiote associé aux plants de café peut jouer un rôle important dans l'expression finale de la qualité du café. Toutefois, la diversité du microbiote des cerises de café est encore mal caractérisée. Nous avons examiné ici la diversité des endophytes des cerises de *Coffea arabica* en utilisant des approches indépendantes de la culture afin d'identifier les microbes associés, ultimement pour comprendre leur écologie et leur rôle potentiel dans la détermination de la qualité du café. Une PCR–DGGE spécifique au groupe des gènes de l'ARNr 16S et 26S et le séquençage d'une banque de clones ont montré que la communauté des endophytes était composée de membres des trois domaines du vivant. Des séquences bactériennes présentant une forte similarité avec des bactéries cultivables et non cultivables appartenaient aux phylums des *Betaproteobacteria*, *Gammaproteobacteria* et *Firmicutes*. Les analyses phylogéniques des séquences clonées des *Firmicutes* révélaient que la plupart des séquences correspondaient à trois genres principaux : *Bacillus*, *Staphylococcus* et *Paenibacillus*. Les séquences d'archées révélaient la présence d'UTO appartenant aux phylums des *Euryarchaeota* et *Crenarchaeota*. Des séquences de levures endophytes n'étaient pas récupérées, mais plusieurs séquences distinctes montrant une identité élevée avec les champignons filamenteux ont été trouvées. Il n'y avait pas de corrélation évidente entre la composition du microbiote et le cultivar ou la localisation géographique du plant de café. Au meilleur de notre connaissance, il s'agit du premier rapport qui démontre la colonisation du tissu interne de plantes à fruits par des membres du domaine des archées. La découverte de petites sous-unités d'ARNr dans les cerises de café, même si elle ne suffit pas à indiquer leur rôle comme endophytes actifs, élargit certainement nos perspectives pour considérer les membres de ce domaine comme microbes endophytes potentiels. [Traduit par la Rédaction]

Mots-clés : cerises de cafe, endophytes, PCR–DGGE, banques d'ARNr.

Introduction

Coffee is one of the most traded agricultural commodities in the world and an important source of revenue for Brazil. The 2 main species of coffee grown for commercial use are *Coffea arabica* (arabica-type coffee) and *Coffea canephora* (robusta-type coffee). Brazil is the world's largest producer of arabica-type coffee, accounting for approximately 36% of total production and more than 27% of world exports (USDA 2011; International Coffee Organization 2012).

Coffee quality, which determines its marketability and price, is a complex trait resulting from multifactorial determinants such as plant genetics, pedoclimatic conditions, and pre- and post-harvest treatments like roasting and brewing (Martin et al. 1999; Montavon et al. 2003; De Los Santos-Briones and Hernández-Sotomayor 2006; Leroy et al. 2006). Although most of the studies directed at improving coffee quality have focused on physiological and agronomic aspects of the coffee plant (Montavon et al. 2003; Leroy et al. 2006), the microbiota associated with coffee plants may play a critical role in the final expression of coffee quality (Pasin et al. 2011).

More recently, the endophytic microbiota present in coffee plants, and especially in coffee cherries, has received considerable

Received 13 November 2012. Revision received 16 January 2013. Accepted 17 January 2013.

M.N.V. Oliveira,* T.M.A. Santos,* H.M.M. Vale, J.C. Delvaux, A.P. Cordero, A.B. Ferreira, P.S.B. Miguel, M.R. Tótola, M.D. Costa, C.A. Moraes, and A.C. Borges. Departamento de Microbiologia/Instituto de Biotecnologia Aplicada à Agropecuária, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil – 36570-000.

Corresponding author: Arnaldo Chaer Borges (e-mail: chaer@ufv.br).

*These authors contributed equally to this article.

Table 1. Description of samples and sampling sites of coffee cherries from *Coffea arabica*.

Sample	Cultivar	Location	Coordinates	Elevation[b] (m)
625	Catuaí Vermelho	Viçosa, UFV[a] Experimental Field Station	20°44′3.48″S, 42°50′57.2″W	676
628	Bourbon Amarelo	Viçosa, UFV Experimental Field Station	20°44′29.5″S, 42°50′47.4″W	687
40B	Catucaí	Araponga, Serra do Boné farm	20°84′78.7″S, 42°65′65.3″W	936
629	Catuaí Vermelho	Araponga, Itatiaia farm	20°40′32.4″S, 42°34′13.6″W	956
630	Catuaí Amarelo	Araponga, Itatiaia farm	20°39′55.2″S, 42°34′13.6″W	956
631	Bourbon Vermelho	Araponga, Santo Antônio farm	20°39′55.2″S, 42°30′44.2″W	1067
632	Bourbon Amarelo	Araponga, Santo Antônio farm	20°39′55.2″S, 42°30′44.2″W	1067
643	Catuaí Amarelo	Araponga, Serra do Boné farm	20°39′16.3″S, 42°27′32.0″W	1013
644	Catuaí Vermelho	Araponga, Serra do Boné farm	20°39′16.3″S, 42°27′32.0″W	1013
646	Catuaí Vermelho	Araponga, Serra do Boné farm	20°39′57.9″S, 42°30′45.4″W	1189

[a]UFV, Universidade Federal de Viçosa.
[b]Elevation measured as metres above sea level.

attention, when its diversity and potential contribution to positive attributes of brewed coffee began to be appreciated. Classically, the endophytic microbiota has been profiled by cultural-dependent methods (Sakiyama et al. 2001; Vega et al. 2005; Silva et al. 2008; Miguel et al. 2013). However, because the majority of microorganisms are not amenable to cultivation under standard laboratory conditions (Amann et al. 1995; Handelsman et al. 1998; Handelsman 2005), we anticipate that the real complexity of the endophytic microbiota in coffee cherries is largely underestimated and a phylogenetic description of the endophytes is far from complete. Molecular analysis of microbial communities (assemblages of microorganisms that share the same environment) has proven to be a powerful strategy for studying endophytic microbial diversity in a variety of agronomic plants (Sun et al. 2008; Ulrich et al. 2008; West et al. 2010).

The purpose of this study was to provide a broad description of the diversity of endophytes in coffee cherries of *C. arabica* from southeastern Brazil by using a combination of molecular-based approaches. Characterizing the microbial communities associated with this important agricultural plant will enable a future understanding of the microbial ecology and its potential role in determining coffee quality. This approach contrasts most of published data on coffee microbiology that deals either with coffee diseases or coffee grain spoilage.

Material and methods

Coffee cherries sampling and superficial decontamination

Mature coffee cherries from 5 cultivars of *C. arabica* L. ('Catuaí Vermelho','Catuaí Amarelo', 'Bourbon Vermelho', 'Bourbon Amarelo', and 'Catucaí'), grown under field conditions, were sampled from the Universidade Federal de Viçosa experimental field stations and from 3 commercial farms near Viçosa, Minas Gerais State, Brazil (Table 1) during the harvest season. Minas Gerais is the main coffee producer state in Brazil, and farms located at higher altitudes, such as those used in this study, are known to produce a better coffee beverage than farms located at lower altitudes. Within each coffee plantation, a representative stand with good productivity was chosen. In each plant, sampling involved cutting the intact branch to preserve the integrity of the cherries. Samples were maintained on ice during transportation, and subsequently, the cherries were vacuum-packed and frozen at –20 °C until processing. This process was shown to be effective at preserving coffee cherries for long-term storage at lower temperature. Eight healthy representative whole cherries were prewashed in running tap water, soaked in neutral detergent, and rinsed in running distilled water. The cherries' surface was decontaminated aseptically (Sakiyama et al. 2001). Briefly, the cherries were immersed twice in distilled water and once in 50 mmol·L^{-1} potassium phosphate buffer, pH 7.0. The cherries were then immersed in 70% (*v*/*v*) ethanol for 1 min, soaked for 5 min under vigorous agitation in 5% (*v*/*v*) sodium hypochlorite containing 0.05% (*v*/*v*) Tween 80, and quickly flamed after immersion in 70% (*v*/*v*) ethanol. The cherries were immersed in 50 mmol·L^{-1} potassium phosphate buffer, pH 7.0, for 15 min, and the process of surface decontamination described above was repeated. Finally, the cherries were washed twice in distilled water and individually incubated for 72 h in 10 mL of R2A broth (containing, per litre of distilled water, 0.5 g of Bacto-Peptone, 0.5 g of casamino acids, 0.5 g of yeast extract, 0.5 g of dextrose, 0.5 g of soluble starch, 0.3 g of dipotassium phosphate, 0.05 g of $MgSO_4 \cdot 7H_2O$, 0.3 g of sodium pyrophosphate). Duplicates of 100 μL aliquots of the final washing water were spread on R2A agar. Absence of microbial growth in both R2A broth and agar was used to confirm the success of the superficial decontamination.

DNA extraction

Eight representative and superficially decontaminated cherries of each plantation were individually macerated in liquid nitrogen, and the fine powder obtained from each cherry was pooled. Next, 830 μL of extraction buffer (4% (*m*/*v*) cetyltrimethylammonium bromide (CTAB), 1.4 mol·L^{-1} NaCl, 20 mmol·L^{-1} EDTA, 100 mmol·L^{-1} Tris–HCl, pH 8.0, 1% (*m*/*v*) polyvinylpyrrolidone), 70 mmol·L^{-1} β-mercaptoethanol, and 125 μg of lysozyme (Sigma-Aldrich, St. Louis, Missouri) were added to the resulting mixed samples. Additionally, to enhance archaeal DNA extraction, 30 μg of proteinase K (Sigma) and 30 μg of pronase E (Sigma) were added to the mixture and incubated at 37 °C for 1 h with occasional agitation. Subsequently, the mixture was incubated at 65 °C for 20 min and submitted to deproteinization with phenol – chloroform – isoamyl alcohol (25:24:1). The aqueous phase was collected after centrifugation for 5 min at 10 000*g*, and DNA was precipitated with 0.6 volume (approximately 420 μL) of isopropanol at room temperature for 30 min. The mixture was centrifuged for 5 min at 10 000*g*; the DNA pellet was washed twice with cold 70% ethanol and resuspended in 30 μL of TE buffer (10 mmol·L^{-1} Tris–HCl, pH 8.0, 1 mmol·L^{-1} EDTA). DNA concentration and purity were evaluated by optical density at the wavelengths 260 and 280 nm. Integrity was assessed through a 0.8% (*m*/*v*) agarose gel electrophoresis, staining with 0.5 μg·mL^{-1} ethidium bromide, and visualization with a Stratagene Eagle Eye II Trans Iluminator and Imaging System (Stratagene, Cedar Creek, Texas, USA).

Amplification and denaturing gradient gel electrophoresis (DGGE) analysis of rRNA gene fragments

To examine the diversity of the natural microbial community in coffee cherries by DGGE, a nested PCR strategy was used. In a first round of PCR, primers targeting the bacterial (phylum-specific primers) or archaeal (Arch344f/927r) 16S rRNA genes or the yeast (NL1/NL4) 26S rRNA genes (Table 2) were used. The specific fragments generated were separately used as templates for a second PCR (nested PCR) primed by conserved group-specific DGGE primers containing a long GC clamp incorporated into the forward primer (Table 2), targeting the bacterial (F984GC/R1378), archaeal (Arch344fGC/517r), or yeast (NL1GC/LS2) ribosomal RNA. All reac-

Table 2. Sequences of the 16S and 26S rRNA gene-targeted primers used in this study.

Proposed target	Primer name, primer sequence (5′→3′)	Primer sequence reference	Size of amplified product (bp)	Initial denaturation	Amplification	Final extension
Archaea	Arch344f (344–363[a]), ACGGGGCGCAGCAGGCGCGA	Raskin et al. 1994	605	94 °C, 5 min	10 cycles at 94 °C, 30 s; 61–56 °C 30 s; 72 °C, 1 min; and 30 cycles at 94 °C, 30 s; 56 °C, 30 s; 72 °C, 1 min	72 °C, 7 min
	927r (927–948), CCCGCCAATTCCTTTAAGTTTC	Jurgens et al. 1997				
Archaeal nested PCR[c]	Arch344fGC, gc0.1-ACGGGGCGCAGCAGGCGCGA		190			
	517r (517–534), ATTACCGCGGCTGCTGG	Lane et al. 1985				
Bacteria						
Alphaproteobacteria	F203α (174–203[a]), CCGCATACGCCCTACGGGGGAAAGATTTAT	Gomes et al. 2001	1340	94 °C, 5 min	25 cycles at 94 °C, 1 min; 56 °C, 1 min; 72 °C, 2 min	72 °C, 10 min
	R1492 (1492–1513), TACGGYTACCTTGTTACGACTT	Heuer et al. 1997				
Betaproteobacteria	F948β (931–948), CGCACAAGCGGTGGATGA	Gomes et al. 2001	583	94 °C, 5 min	25 cycles at 94 °C, 1 min; 61 °C, 1 min; 72 °C, 2 min	72 °C, 10 min
	R1492 (1492–1513), TACGGYTACCTTGTTACGACTT					
Gammaproteobacteria Enteric bacteria	fD2 (8–27), AGAGTTTGATCATGGCTCAG	Weisburg et al. 1991	1505	94 °C, 2 min	30 cycles at 94 °C, 15 s; 52 °C, 30 s; 72 °C, 1 min	72 °C, 7 min
	rP1 (1492–1512), ACGGTTACCTTGTTACGACTT					
Low G+C Gram-positive bacteria	BLS342F (325–342), CAGCAGTAGGGAATCTTC	Blackwood et al. 2005	1082	95 °C, 3 min	35 cycles at 94 °C, 30 s; 57 °C, 30 s; 72 °C, 1.5 min	72 °C, 7 min
	1392R (1392–1406), ACGGGCGGTGTGTACA	Amann et al. 1995				
Actinomycetales	F243HGC (226–243), GGATGAGCCCGCGGCCTA	Heuer et al. 1997	1288	95 °C, 5 min	35 cycles at 94 °C, 1 min; 63 °C, 1 min; 72 °C, 2 min	72 °C, 10 min
	R1492 (1492–1513), TACGGYTACCTTGTTACGACTT					
Bacterial nested-PCR[c]	984GC (968–984), gc0.1-AACGCGAAGAACCTTAC	Heuer et al. 1997	434	94 °C, 5 min	20 cycles at 94 °C, 1 min; 53 °C, 1 min; 72 °C, 2 min	72 °C, 10 min
	R1378 (1378–1401), CGGTGTGTACAAGGCCCGGGAACG					
Yeast	NL1 (63–86[b]), GCATATCAATAAGCGGAGGAAAAG	Cocolin et al. 2000 Prakitchaiwattana et al. 2004	580	95 °C, 5 min	35 cycles at 94 °C, 1 min; 52 °C, 1 min; 72 °C, 2 min	72 °C, 10 min
	NL4 (624–642), GGTCCGTGTTTCAAGACGG					
Yeast nested-PCR[c]	NL1GC, gc0.2-GCATATCAATAAGCGGAGGAAAAG		223	95 °C, 5 min	30 cycles at 95 °C, 1 min; 52 °C, 1 min; 72 °C, 2 min	72 °C, 7 min
	LS2 (266–285), ATTCCCAAACAACTCGACTC	Cocolin et al. 2000				
GC clamps	gc0.1, CGCCCGGGGCGCGCCCCGGGCGGGGCGGGGGCACGGGGGG	Nubel et al. 1996				
	gc0.2, CGCCCGCCGCGCGCGGCGGGCGGGGCGGGGGC	Prakitchaiwattana et al. 2004				

[a]*Escherichia coli* numbering.
[b]Yeast primers positions are based on the *Saccharomyces cerevisiae* 26S rRNA gene.
[c]These primers were used in the nested-PCR using as template the product of amplification generated by bacterial phylum-specific primers.

tions were carried out as described above, with the thermal cycling conditions given in Table 2. Appropriate PCR controls and blanks were included and the products were analyzed by electrophoresis in a 1.2% (*m*/*v*) agarose gel stained with 0.5 μg·mL^{-1} ethidium bromide before DGGE analysis.

Profiling of the amplified sequences was conducted as described previously (Muyzer et al. 1993) using a DCode Universal Mutation Detection System apparatus (Bio-Rad, Hercules, California, USA). The PCR products were loaded onto a 6% (for Bacteria) or 8% (for Archaea and yeast) polyacrylamide (37.5:1, acrylamide – *N*,*N*′-methylenebisacrylamide) (Sigma) parallel denaturing gradient gel in 1× TAE buffer (40 mmol·L^{-1} Tris–HCl, pH 8.0, 20 mmol·L^{-1} acetic acid, 1 mmol·L^{-1} EDTA, pH 8.0) composed of 0.09% (*v*/*v*) TEMED (*N*,*N*,*N*′,*N*′-tetramethylenediamine) and 0.07% (*m*/*v*) ammonium persulfate. The denaturing gradient was optimized empirically to 40%–58% (for Bacteria) and 40%–60% (for Archaea and yeast) urea/formamide (100% denaturant contained 7 mol·L^{-1} urea and 40% (*v*/*v*) deionized formamide). The electrophoresis was carried out in 1× TAE buffer at 50 V for 16 h at a constant temperature of 60 °C. The DNA fragments were stained for 20 min in 1× TAE buffer with 1× SYBR Gold (Invitrogen, Carlsbad, California, USA). Images of the gels were obtained under UV light using an L-PIX CHEMI Molecular Imaging system (Loccus Biotechnology, São Paulo, SP, Brazil). Sample lanes were analyzed using BioNumerics version 6.0 (Applied Maths, Inc., Austin, Texas, USA).

For analysis of bacterial endophytes, individual DNA bands from the polyacrylamide gels were excised and incubated overnight at 4 °C in 50 μL of sterile deionized water. A 5 μL aliquot of DNA extracted from each band was used as the template in a PCR amplification with the F984 (no GC clamp) and R1378 primers as described above. This procedure was performed only for Bacteria because of the smaller amplicon size for Archaea (190 bp) and Fungi (223 bp), which would result in sequences smaller than 200 bp after sequencing, impairing their deposition in the NCBI database. The purity and relative position of the reamplified DNA were confirmed by DGGE as previously described in this work. PCR products were treated with ExoSapIT (USB, Cleveland, Ohio, USA) according to the manufacturer's instructions and directly sequenced.

Cloning and sequencing

Amplification of rRNA gene fragments from metagenomic DNA extracted from coffee cherries was performed using primers specific for *Firmicutes* (BLS342F/1392R), Archaea (Arch344f/927r), and yeast (NL1/NL4) (Table 2) as described above. The sample 644 was used to construct the *Firmicutes* and *Fungi* library because that sample showed the PCR–DGGE profiles with highest richness (number of operational taxonomic units (OTUs) considered as the number of DGGE bands). The sample 40B was used for the *Archaea* library because that was the only sample with visible amplification product using archaeal-specific primers described above, whereas for all other samples, the amplicon could be detected only after a nested PCR reaction (Table 2), resulting in shorter fragments, about 190 bp, that could not be deposited in the NCBI database. Amplification products were electrophoresed through a 1.2% (*m*/*v*) agarose gel, stained with 0.5 μg·mL^{-1} ethidium bromide, and visualized under UV light. Positive results were considered to be a pool of amplicons that had the expected molecular sizes. PCR products were cloned into a pGEM-T Easy vector system (Promega, Madison, Wisconsin, USA). Competent *Escherichia coli* cells were prepared using the calcium chloride method (Sambrook and Russell 2001) and grown aerobically overnight on solid Luria–Bertani (LB) medium containing ampicillin (50 μg·mL^{-1}) and X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside, 50 μg·mL^{-1}) (Sigma) at 37 °C. After overnight incubation, individual white colonies were randomly picked from the plates, placed into 5 mL of liquid LB containing ampicillin (50 μg·mL^{-1}), and grown aerobically for 16–18 h at 37 °C. Plasmids were extracted using the Wizard SV Minipreps DNA Purification kit (Promega), and the cloned fragments were sequenced using M13 vector primers using the Sanger sequencing method. All clones containing inserts of the correct size were stored in LB medium containing 20% (*v*/*v*) glycerol at –80 °C. In most cases, only one strand of the DNA fragments was sequenced. This criterion proved to be sufficient for taxonomic identification of the cloned fragments obtained using the BLAST (Basic Local Alignment Search Tool) search function (Altschul et al. 1990). Both DNA strands were sequenced in those cases in which the sequence could not be easily assigned to a particular taxonomic group or the nucleotide sequence could not be clearly determined. The resulting electropherograms and sequences were analyzed using SEQUENCHER software version 4.1.4 (Gene Codes Corporation, Ann Arbor, Michigan, USA). Chimeric sequences were identified by Bellerophon (Huber et al. 2004). Both chimeric sequences and sequences yielding weak signals were excluded from further analysis.

Estimation of the size of the clone libraries and phylogenetic analysis

For phylogenetic purposes, sequences in the clone libraries, constructed for *Firmicutes*, Archaea, and Fungi, sharing more than 97% sequence identity were considered as a single OTU. The clone coverage in each library was calculated according to the following equation: $C = 1 - (n/N)$, where n is the number of singleton OTU, and N is the total number of clones screened in the clone library (Good 1953). Rarefaction curves for clone libraries (Colwell and Coddington 1994) were calculated using ESTIMATES version 7.5 software (http://viceroy.eeb.uconn.edu/EstimateS).

To study the evolutionary relationships among the sequences obtained in the 16S rDNA clone libraries, sequences were compared with sequences from the nucleotide collection database stored in GenBank using the BLAST algorithm (Altschul et al. 1990). All sequences with high identity, as identified by the search, were imported into MEGA 4 software and aligned with other 16S rRNA gene sequence fragments using ClustalW (Larkin et al. 2007). The alignments were manually trimmed, and calculation of the phylogenetic trees was based on these sequence alignments using the Maximum Parsimony method (Tamura et al. 2011). Evolutionary distances were computed using the p-distance method. To check the robustness of the resulting tree and the statistical significance levels of the interior nodes, bootstrap analysis with 1000 replicates was carried out and values above 50% were reported.

Nucleotide sequence accession numbers

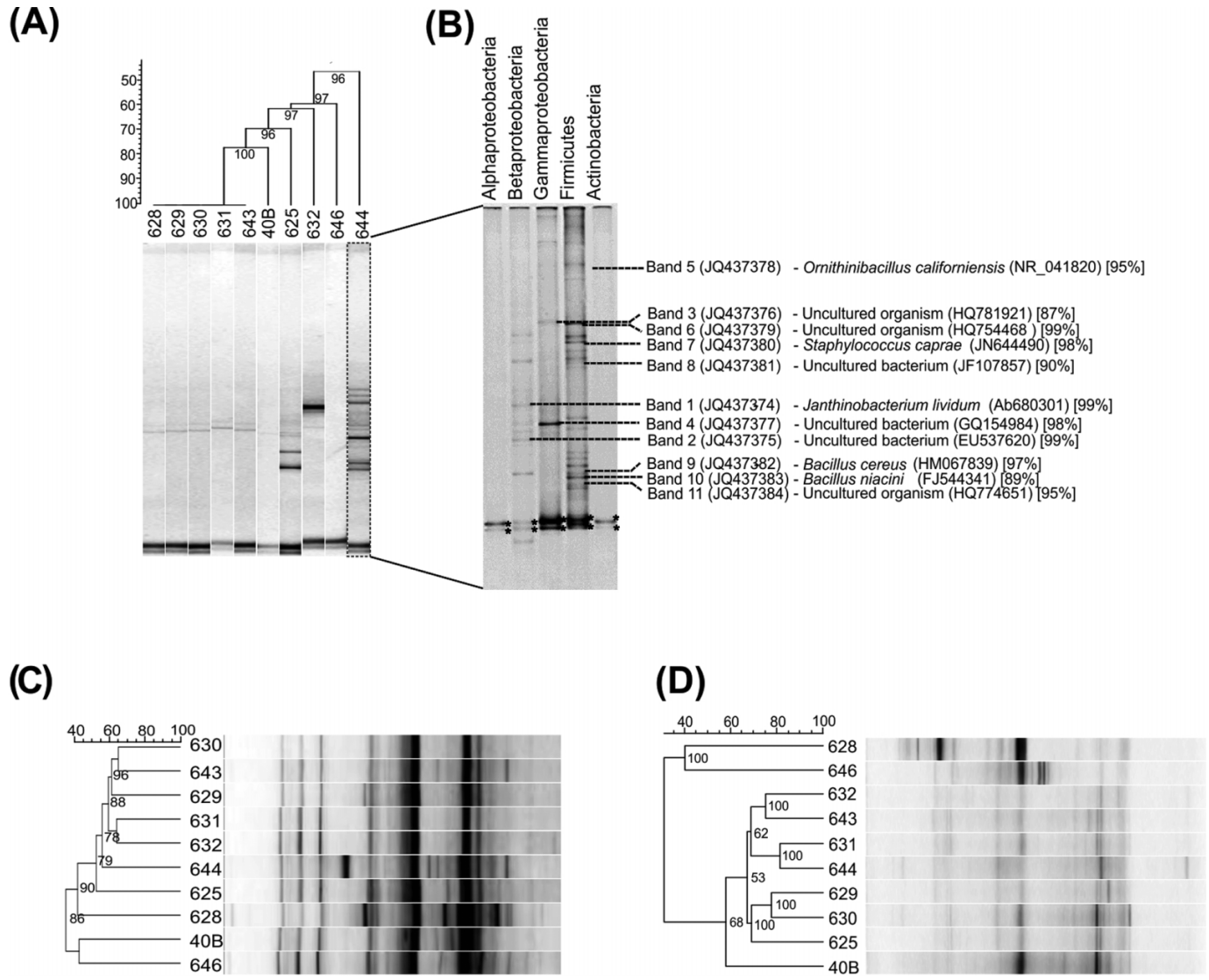
The sequences reported in this study have been submitted to GenBank under accession Nos. JQ417214–JQ417239, JN004109–JN004134, JQ437374–JQ437388, JQ736752–JQ736776, and JQ417214–JQ417239.

Results

Diversity of endophytes of coffee cherries by PCR–DGGE

The DGGE banding profile showed homogeneity of the endophytic bacterial community, and the number of bands detected indicated a relatively small diversity (Fig. 1A). The direct approach, in which bacterial universal primers (984GC/R1378) were used in a single PCR round, indicated that the coffee cherries from most cultivars (samples 40B, 628, 629, 630, 631, 643, and 646) appeared to contain 2 main bands and at least 2 additional minor bands. In contrast, samples 625, 632, and 644 showed additional OTUs, with sample 644 exhibiting the most complex profile, as demonstrated by the number of bands on the DGGE gel (Fig. 1A). The evident richness of sample 644 (Catuaí Vermelho) encouraged a detailed analysis of the various bacterial groups potentially present, and

Fig. 1. Clustering analysis of PCR–DGGE profiles generated from coffee cherries metagenomic DNA from different farms using (A) universal primers targeting the bacterial 16S rRNA gene. (B) The highlighted sample was used for further characterization of the microbial diversity by using phyla-specific primers and by sequencing of selected bands. For each sequenced band the closest match at NCBI database is shown, followed by its accession number and the percent identity. The asterisks (*) indicate bands showing identity with DNA of chloroplast. (C) Primers targeting the fungal D1/D2 domain of the 26S rRNA gene. (D) Universal primers targeting the archaeal 16S rRNA gene. The sample ID No. is indicated beside each lane. The comparison of the PCR–DGGE profiles was generated with BioNumerics software (version 6.0). The scale indicates the level of similarity based on Dice's coefficient.



we used an indirect approach (2-step nested PCR–DGGE) for evaluation of the bacterial diversity profile in this sample.

The nested PCR approach provided a high-resolution overview of the complexity of the microbial communities present and, in addition, allowed analysis of minority populations (Fig. 1B). Fifty-three bands were identified with BioNumerics (version 6.0), with each one likely corresponding to a different OTU. The *Firmicutes* phylum exhibited the most complex profile, followed by the *Betaproteobacteria* and *Gammaproteobacteria* classes. Only 22 of 53 bands were successfully sequenced after purification from the polyacrylamide gel. Apparently, the approach did not completely exclude contamination by chloroplast DNA, since the sequences of 10 bands (Fig. 1B) showed high identity (98%) with chloroplast DNA of *C. arabica*. In addition, the presence of OTUs belonging to *Alphaproteobacteria* and *Actinobacteria* remained uncertain because the sequence bands showed identity with chloroplast DNA as well as with bacterial sequences. Sequencing of the remaining bands showed that the OTUs from *Betaproteobacteria* class (bands 1 and 2), *Gammaproteobacteria* class (bands 3 and 4), and *Firmicutes* (bands 5 to 12) had high similarity to sequences of cultured or uncultured bacteria belonging to the respective classes and phylum (Figs. 1A and 1B). Band 9 presented in the profile corresponding to *Firmicutes* showed no similarity with known sequences previously submitted to the GenBank database.

Profiling of the samples with yeast-specific primers revealed a more complex pattern and apparent homogeneity of the community in the different samples when compared with the bacterial profile (Fig. 1C). Inspection of the DGGE patterns showed that most bands are equally present at the same relative position, and there is no obvious variation in the intensity of the bands. Sample 644 (Catuaí Vermelho), which showed some additional major bands, grouped externally with samples 646, 632, 625, 40B, 643, 631, 630, 629, and 628.

Table 3. Clones from each library constructed using bacterial and archaeal 16S rRNA and fungal 26S rRNA isolated from coffee cherries, and its closest match at NCBI database, percent identity, and clone abundance.

Sample	Library	Clone ID	GenBank acc. No.	Sequence affiliation (GenBank ID)	Sequence identity (%)	% of clones
644	Firmicutes	F1.23	JQ437385	*Bacillus firmus* (AJ491843)	96	44
		F3.14	JQ437387	*Staphylococcus epidermidis* (AM157439)	98	8
		F3.47	JQ437388	*Paenebacillus* sp. (DQ444989)	100	4
		F6.4	JQ437386	*Bacillus cereus* (EU871042)	98	16
				Coffea arabica chloroplast	98	28
644	Fungi	LEM 1-43	JQ417223	*Microdiplodia* sp. (DQ377913)	98	2
		LEM 1-1	JQ417215	*Mycosphaerella keniensis* (DQ246239)	99	10
		LEM 1-10	JQ417216	Endophytic fungi (EU687141)	99	2
		LEM 1-14	JQ417214	Endophytic fungi (EF420023)	99	4
		LEM 1-21	JQ417220	*Mycosphaerella pini* (EF114697)	98	2
		LEM 2-14	JQ417232	*Glomerella acutata* (DQ286133)	99	4
		LEM 2-20	JQ417236	*Phaeophaeria* sp. (EF590319)	99	2
				Coffea arabica	98	74
40B	Archaea	40B0.2	JN004110	*Halococcus* sp. BIGigoW09 (AM902587)	99	35
		40B0.8	JN004117	*Haloferax* sp. FIB210_9 (EU308271)	96	15
		40B0.10	JN004119	*Halobacterium volcanii* (K00421)	99	15
		40B0.14	JN004110	Uncultured crenarchaeote (AY454553)	99	11
		40B0.13	JN004123	*Methanobrevibacter woesei* (DQ445717)	97	8
		40B0.16	JN004124	Uncultured archaeon clone (FJ584387)	98	8
		40B0.24	JN004125	Uncultured *Thermoplasma* sp. (AM292020)	97	8

The archaeal profiles showed a relatively small diversity, with most cultivars containing 2 or 3 main bands and few additional minor bands (Fig. 1D). Cluster analysis grouped samples from 3 Catucaí cultivars (625, 629, and 630), grown below 1013 m above sea level, with samples at the same altitude forming a subgroup with higher similarity. Sample 40B (Catucaí) formed a separate cluster, sharing less than 60% similarity with the other samples. Different from what was observed for the yeast profile, sample 646 grouped together with sample 40B in an independent cluster.

For all 3 domains studied, the different DGGE profiles showed no obvious correlation with altitude of the plant or coffee cultivar, as supported by clustering analysis.

Sample size and OTU diversity in the 16S rRNA gene libraries

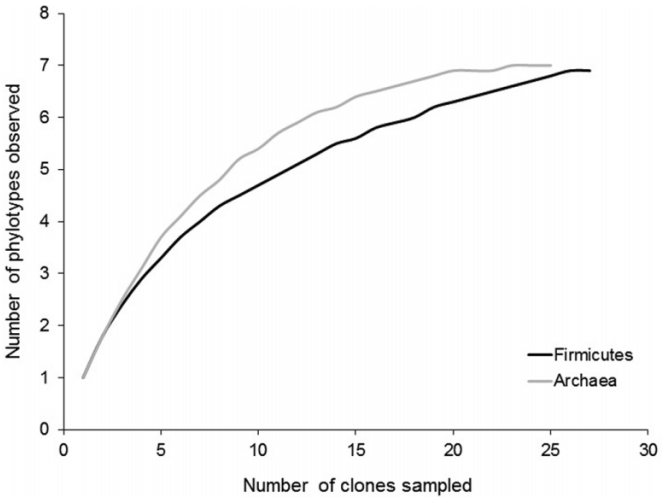
In total, 25 and 26 clones from the *Firmicutes* and archaeal 16S rRNA gene libraries were screened, respectively. The calculated rarefaction curves for both libraries tended to plateau, showing that the number of clones screened in the bacterial clone library was sufficient to reveal most of the sequence types within the library and to reasonably describe the diversity of both groups (Fig. 2). Clone coverage was 96% and 100% for *Firmicutes* and *Archaea*, respectively (Table 3), and these values suggest than the number of unique sequence types sampled from this library approached the total number of unique sequences within the library. Sequencing of 50 clones of the constructed 26S rRNA library revealed sequences related only to filamentous fungi (Table 3), and these were not further analyzed for phylogenetic purposes.

Phylogenetic composition of the *Firmicutes* and archaeal endophytic communities

Two rRNA libraries were constructed from 2 samples using the amplification product of group-specific primers for *Firmicutes* (sample 644) and Archaea (sample 40B).

Phylogenetic analysis of the clone sequences revealed that the majority of the sequences fell into 3 major genera of the *Firmicutes* phylum (≥97% identity): *Bacillus* (60% of sequences), *Staphylococcus* (8%), and *Paenibacillus* (4%). About 28% of the sequences had high identity with unidentified uncultured bacteria or chloroplasts (Table 3). About 44% of the sequences clustered in the *Bacillus* group showed high identity (96%) with *Bacillus firmus*, whereas 16% showed high identity (98%) with *Bacillus cereus*. The remaining clones presented sequences showing high identity with *Staphylococcus epidermidis* and *Paenibacillus* sp. (Table 3 and Fig. 3A).

Fig. 2. Rarefaction curves indicating 16S rRNA gene diversity within the *Firmicutes* and archaeal clone libraries. The frequency of different sequences sharing ≥97% identity is plotted as a function of the number of clones screened.

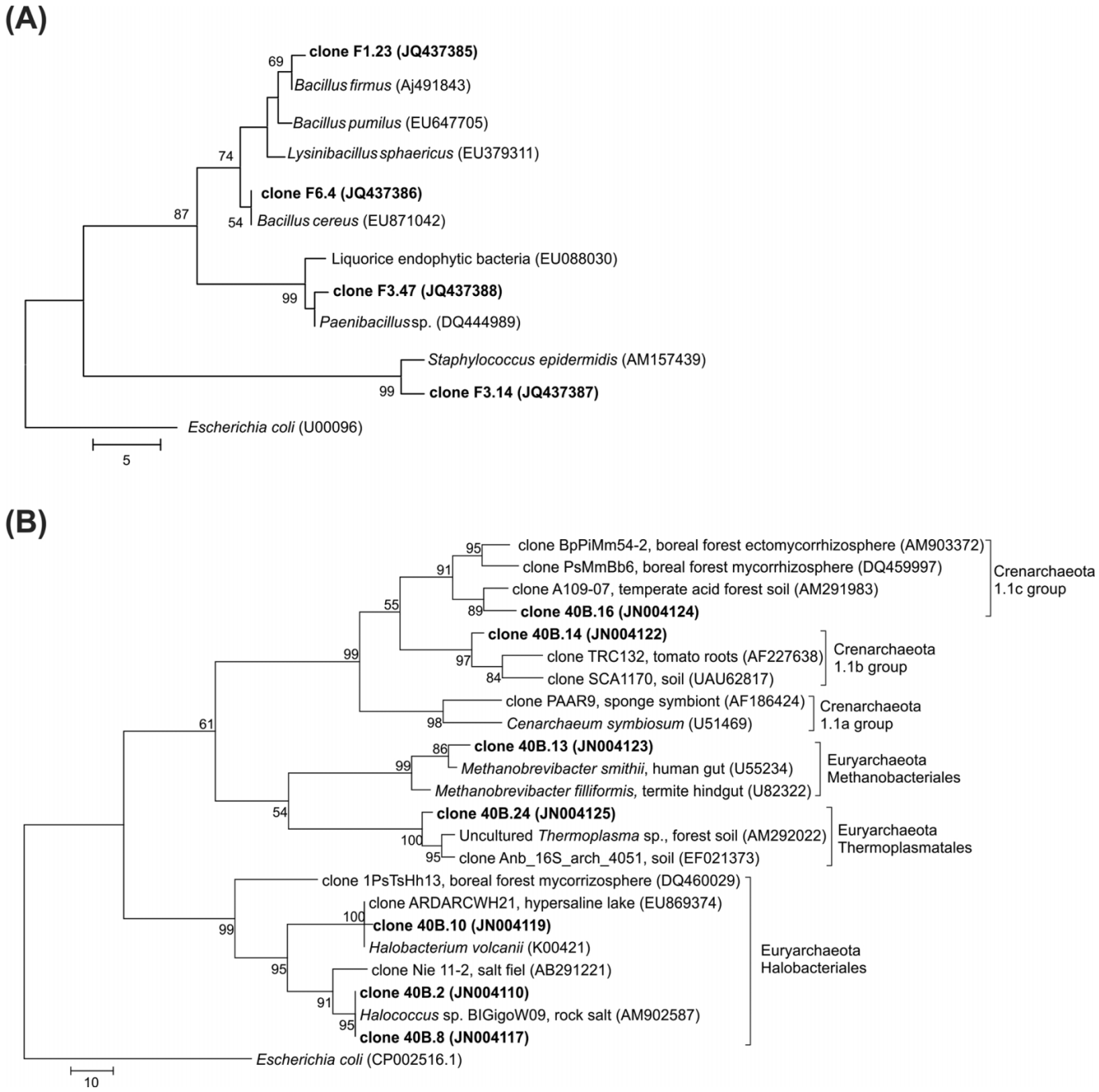


Phylogenetic analysis of the archaeal sequences revealed the presence of 7 OTUs, most of which (80.8% of sequences) belong to the phylum *Euryarchaeota* and are similar to sequences of halophilic archaea (65.4%) (Fig. 3B). In addition, 2 other sequences belonging to *Euryarchaeota* were identified and clustered with sequences of the orders *Methanobacteriales* (7.7%) and *Thermoplasmatales* (7.7%). Only 2 OTUs identified belong to the phylum *Crenarchaeota*, specifically to the groups 1.1b (11.5%) and 1.1c (7.7%), which are abundant in soils (Fig. 3B).

Discussion

Endophytic bacteria are ubiquitous in most plant species (Hallmann et al. 1997), and their frequent occurrence in important agricultural plants and relevance to crop production systems is just beginning to be appreciated more widely. In addition to the

Fig. 3. Maximum parsimony phylogenetic tree of *Firmicutes* (A) and *Archaea* (B) 16S rRNA gene sequences rooted by the 16S rRNA gene sequence of *Escherichia coli*. Bootstrap values are shown for nodes with over 50% support. The sequences found in this study are presented in bold script. The accession numbers of each sequence extracted from the GenBank database are shown after each sequence name.



economic aspects, the study of endophytic microorganisms has academic value, concerning the discovery of new microbial species. The study of coffee microbiology has been classically restricted to the investigation of undesirable microbiota, whereas the potential positive roles of some microorganisms intrinsically associated with the plant (such as endophytes) have been underappreciated.

Indeed, the endophytic microbiota present in coffee has not been extensively investigated and only a few culture-dependent studies have been published (Vega et al. 2005; Silva et al. 2008; Vilela et al. 2010), including a recent study prospecting for plant growth promoters and biocontrol agents (Silva et al. 2012). Hence, it is likely that most endophytic microorganisms present in coffee plants, and especially in coffee cherries, are still largely unknown. Despite previous efforts to isolate and characterize the endophytes present in coffee cherries to better understand their ecology and physiology, a phylogenetic description of the microbiota present in the plant is far from complete. Our study used a combination of culture-independent methods to reveal previously uncharacterized features of the microbiota present in coffee cherries.

The nested PCR approach, used to facilitate the analysis of the 16S and 26S rRNA gene fragments of different microbial groups by DGGE, allowed the analysis of minority populations present in

coffee cherries. A similar approach has been used to monitor changes in the actinomycete community of a potato rhizosphere (Heuer et al. 1997) and bacterial communities in different wastewater treatment plants (Boon et al. 2002). In the present study, sequence analysis of bacterial DNA obtained after DGGE revealed for the first time, to the best of our knowledge, the possible presence of uncultured bacteria in the endophytic community of coffee cherries.

A variety of members of the genus *Bacillus* (Vega et al. 2005; Silva et al. 2008) and several species of *Paenibacillus* (Sakiyama et al. 2001) were previously isolated from coffee cherries. The results of our BLAST search and phylogenetic analysis of clones from our *Firmicutes* library showed that some of the identified OTUs have not been observed previously as endophytic in coffee cherries. In a similar study that investigated the diversity of endophytic bacteria in sugarcane, sequences affiliated with *Staphylococcus epidermidis* were found in the clone library (Velazquez et al. 2008), but information about the occurrence of *Staphylococcus* members as endophytes in coffee cherries is unavailable so far. Additionally, the DNA sequences of some clones showed low identity with cultured bacteria but high identity with sequences of uncultured bacteria or chloroplasts (98% of identity). We do not rule out the possibility that these sequences are simply from contaminating plant tissue, but it is reasonable to speculate that these sequences could be from members of the *Cyanobacteria*, which shares high identity values with chloroplast 16S rRNA. However, that hypothesis should be tested by isolating those microorganisms from coffee cherries and sequencing their 16S rRNA.

We have previously isolated endophytic yeasts from coffee cherries, and gas chromatography of fatty acid methyl ester (MIDI–FAME) identified the isolates as being members of the genera *Candida* and *Lipomyces* (H.M.M. Vale, unpublished data). However, our efforts to amplify 26S rRNA gene fragments from yeasts to construct rRNA library were unsuccessful, since only filamentous fungi related sequences were recovered from sequencing (Table 3). One possible explanation is that both the diversity and the abundance of endophytic yeasts are relatively low when compared with bacterial groups, or the primers used to amplify yeast 26S rRNA gene fragments are biased toward filamentous fungi. Members of the genera *Debaryomyces*, *Candida*, and *Pichia* have previously been isolated from coffee cherries during different stages of maturation (Silva et al. 2000) and during natural coffee fermentation (Silva et al. 2000, 2008). The association of yeasts with coffee cherries is very important during fermentation (Avallone et al. 2001), but their role as endophytes during coffee fruit maturation remains unknown.

Although archaeal species have been identified endophytically inhabiting plant roots (Sun et al. 2008), to the best of our knowledge, this is the first report demonstrating internal tissue colonization of plant fruits by members of this domain. Most archaeal sequences identified in our clone library were phylogenetically affiliated (identity ranging from 96% to 99%) with halophilic archaea (phylum *Euryarchaeota*, order *Halobacteriales*) (Fig. 3B), which require a high-salt environment for growth (Martin et al. 1999). Salt concentration in coffee cherries is insignificant, but the concentration of oligosaccharides ranges from 50% to 55% of the fruit dry weight (De Castro and Marraccini 2006). It is likely that such "halophile-like" archaea may be adapted to control osmotic stress in the fruit "environment" by using mechanisms similar to those employed by their halophilic relatives in high-salt environments.

Nonthermophilic *Crenarchaeota* belonging to groups 1.1b and 1.1c represent up to 3% of the total microbiota in different soils (Ochsenreiter et al. 2003; Auguet and Casamayor 2008). The distribution of archaea of group 1.1c is more restricted and mainly found in acid soils (Yrjala et al. 2004; Bomberg and Timonen 2007). Most Brazilian soils are acidic, including the areas of coffee plantations. So, it is likely that members from both groups are present in the soil of the sites where our samples were collected. In principle, the ubiquitous distribution of Archaea in nature and the relatively recent revelation of Archaea occurrence in nonextreme environments (Ochsenreiter et al. 2003; Bomberg and Timonen 2009), including among members of the human microbiota (Miller et al 1982; Belay et al. 1990; Kulik et al. 2001), might allow us to suggest archaeal prevalence in the fruit environment and even the possibility of archaeal species having an active role in the endophytic community. That archaeal small-subunit rDNA was detected in our library from coffee cherries of Catucaí is not sufficient to indicate their role as active endophytes, but it certainly expands our perspectives toward considering members of this domain as potential endophytes of plants.

Coffee flavor and aroma are the result of complex chemical transformations in the coffee fruit. The green bean during the maturation process has only a faint smell but contains basically all the precursors necessary to generate coffee aroma during postharvesting processing (Montavon et al. 2003; Lindinger et al. 2009). It is likely that the incidence of certain microbial groups in coffee cherries and the biochemical consequences of their presence have significant positive contributions to the improvement of coffee quality. The origin of endophytic microorganisms in fruits of coffee is unknown, but there are several possible routes of entry in the plant, such as during seed germination, flowering, or through stomata or wounds in the plant tissue, especially in the roots, where microorganisms live in close association with the tissue (Hardoim et al. 2008). Once inside plant tissue, how endophytes reach coffee cherries and if there are driving forces selecting microbial colonization in specific parts of plant is a very interesting topic that must be elucidated to further our knowledge on endophytic-coffee plant interactions and their effects on coffee quality.

While we understand the limitations of our study, we reinforce that this is an effort to better characterize the microbiota in coffee fruits and understand the possible ecology in this system. Thus, generalizations of the results observed here to other similar settings should be carefully considered. In fact, it is likely that the farm environment, as well as the preharvesting techniques, exerts a large influence on the structure and richness of the endophytic community inside the fruit. Therefore, it is not surprising that that we did not detect members that are commonly found as endophytes in coffee fruits (Vega et al. 2005; Silva et al. 2008, 2012; Vilela et al. 2010). Even though we cannot completely rule out intrinsic bias of metagenomic studies, particularly during DNA extraction and PCR, which might account for this observation, culture-independent diversity estimates are considerably less biased than culture-based methods, as the latter tend to target fast-growing organisms and exclude fastidious microorganisms. Therefore, further research is needed to compare the microbiota present in coffee fruits from farms located in different regions.

The comprehensive description of the diversity of endophytes in coffee cherries and the study of their ecology is essential to fully understand the production and conversion of beneficial metabolic precursors that will yield high-quality brewed coffees and their unique flavors. This study indicates preliminary data on microbial diversity in coffee cherries, opening the avenues for in-depth investigation on coffee plants distinguished by numerous properties and for further analyses carried out through the synergic application of cultivation-dependent and -independent methods. Here, we characterized the diversity of endophytic microbes in coffee cherries of *C. arabica* and revealed a microbial composition more diverse than previously thought. In addition, it showed that there is no obvious correlation between the microbial composition and cultivar or geographic localization of the coffee plant. Finally, it suggested, for the first time, the presence of members of the Archaea inhabiting the internal fruit tissue.

Acknowledgements

This work was partially funded by Illycaffè S.p.A. and supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

References

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. 1990. Basic local alignment search tool. J. Mol. Biol. **215**(3): 403–410. doi:10.1006/jmbi.1990.9999. PMID:2231712.

Amann, R., Ludwig, W., and Schleifer, K. 1995. Phylogenetic identification and in-situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59**(1): 143–169. PMID:7535888.

Auguet, J.C., and Casamayor, E.O. 2008. A hotspot for cold crenarchaeota in the neuston of high mountain lakes. Environ. Microbiol. **10**(4): 1080–1086. doi:10.1111/j.1462-2920.2007.01498.x. PMID:18215160.

Avallone, S., Guyot, B., Brillouet, J.M., Olguin, E., and Guiraud, J.P. 2001. Microbiological and biochemical study of coffee fermentation. Curr. Microbiol. **42**(4): 252–256. doi:10.1007/s002840110213. PMID:11178725.

Belay, N., Mukhopadhyay, B., Conway de Macario, E., Galask, R., and Daniels, L. 1990. Methanogenic bacteria in human vaginal samples. J. Clin. Microbiol. **28**(7): 1666–1668. PMID:2199527.

Blackwood, C.B., Oaks, A., and Buyer, J.S. 2005. Phylum- and class-specific PCR primers for general microbial community analysis. Appl. Environ. Microbiol. **71**(10): 6193–6198. PMID:10.1128/AEM.71.10.6193-6198.2005.

Bomberg, M., and Timonen, S. 2007. Distribution of cren- and euryarchaeota in scots pine mycorrhizospheres and boreal forest humus. Microb. Ecol. **54**(3): 406–416. doi:10.1007/s00248-007-9232-3. PMID:17334967.

Bomberg, M., and Timonen, S. 2009. Effect of tree species and mycorrhizal colonization on the archaeal population of boreal forest rhizospheres. Appl. Environ. Microbiol. **75**(2): 308–315. doi:10.1128/AEM.01739-08. PMID:18978075.

Boon, N., De Windt, W., Verstraete, W., and Top, E. 2002. Evaluation of nested PCR–DGGE (denaturing gradient gel electrophoresis) with group-specific 16S rRNA primers for the analysis of bacterial communities from different wastewater treatment plants. FEMS Microbiol. Ecol. **39**(2): 101–112. doi:10.1111/j.1574-6941.2002.tb00911.x. PMID:19709189.

Cocolin, L., Bisson, L., and Mills, D. 2000. Direct profiling of the yeast dynamics in wine fermentations. FEMS Microbiol. Let. **189**(1): 81–87. doi:10.1016/S0378-1097(00)00257-3. PMID:10913870.

Colwell, R., and Coddington, J. 1994. Estimating terrestrial biodiversity through extrapolation. Philos. Trans. R. Soc. Lond. B Biol. Sci. **345**(1311): 101–118. doi:10.1098/rstb.1994.0091. PMID:7972351.

De Castro, R.D., and Marraccini, P. 2006. Cytology, biochemistry and molecular changes during coffee fruit development. Braz. J. Plant Physiol. **18**(1): 175–199. doi:10.1590/S1677-04202006000100013.

De Los Santos-Briones, C., and Hernández-Sotomayor, S.M.T. 2006. Coffee biotechnology. Braz. J. Plant Physiol. **18**(1): 217–227. doi:10.1590/S1677-04202006000100015.

Gomes, N., Heuer, H., Schonfeld, J., Costa, R., Mendonca-Hagler, L., and Smalla, K. 2001. Bacterial diversity of the rhizosphere of maize (*Zea mays*) grown in tropical soil studied by temperature gradient gel electrophoresis. Plant Soil, **232**(1–2): 167–180. doi:10.1023/A:1010350406708.

Good, I.J. 1953. The population frequencies of species and the estimation of population parameters. Biometrika, **40**(3–4): 237–264. doi:10.1093/biomet/40.3-4.237.

Hallmann, J., Quadt-Hallmann, A., Mahaffee, W.F., and Kloepper, J.W. 1997. Bacterial endophytes in agricultural crops. Can. J. Microbiol. **43**(10): 895–914. doi:10.1139/m97-131.

Handelsman, J. 2005. Sorting out metagenomes. Nat. Biotechnol. **23**(1): 38–39. doi:10.1038/nbt0105-38. PMID:15637617.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., and Goodman, R.M. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem. Biol. **5**(10): R245–R249. doi:S1074-5521(98)90108-9. PMID:9818143.

Hardoim, P.R., van Overbeek, L.S., and van Elsas, J.D. 2008. Properties of bacterial endophytes and their proposed role in plant growth. Trends Microbiol. **16** (10): 463–471. doi:10.1016/j.tim.2008.07.008. PMID:18789693.

Heuer, H., Krsek, M., Baker, P., Smalla, K., and Wellington, E. 1997. Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl. Environ. Microbiol. **63**(8): 3233–3241. PMID:9251210.

Huber, T., Faulkner, G., and Hugenholtz, P. 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics, **20**(14): 2317–2319. doi:10.1093/bioinformatics/bth226. PMID:15073015.

International Coffee Organization. 2012. Available from http://www.ico.org [accessed 10 November 2012].

Jurgens, G., Lindstrom, K., and Saano, A. 1997. Novel group within the kingdom Crenarchaeota from boreal forest soil. Appl. Environ. Microbiol. **63**(2): 803–805. PMID:9023962.

Kulik, E.M., Sandmeier, H., Hinni, K., and Meyer, J. 2001. Identification of archaeal rDNA from subgingival dental plaque by PCR amplification and sequence analysis. FEMS Microbiol. Let. **196**(2): 129–133. doi:10.1111/j.1574-6968.2001.tb10553.x. PMID:11267768.

Lane, D., Pace, B., Olsen, G., Stahl, D., Sogin, M., and Pace, N. 1985. Rapid-determination of 16S ribosomal-RNA sequences for phylogenetic analyses. Proc. Natl. Acad. Sci. U.S.A. **82**(20): 6955–6959. doi:10.1073/pnas.82.20.6955. PMID:2413450.

Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, H., et al. 2007. Clustal W and clustal X version 2.0. Bioinformatics, **23**(21): 2947–2948. doi:10.1093/bioinformatics/btm404. PMID:17846036.

Leroy, T., Alpizar, E., Dufour, M., and Etienne, H. 2006. Coffee (*Coffea* sp.). Methods Mol. Biol. **344**: 191–208. doi:10.1385/1-59745-131-2:191. PMID:17033063.

Lindinger, C., Pollien, P., de, Vos, R., Tikunov, Y., Hageman, J., Lambot, C., et al. 2009. Identification of ethyl formate as a quality marker of the fermented off-note in coffee by a nontargeted chemometric approach. J. Agric. Food. Chem. **57**(21): 9972–9978. doi:10.1021/jf901673d. PMID:19817421.

Martin, D., Ciulla, R., and Roberts, M. 1999. Osmoadaptation in archaea. Appl. Environ. Microbiol. **65**(5): 1815–1825. PMID:1981742.

Miguel, P.S.B., Delvaux, J.C., Oliveira, M.N.V., Monteiro, L.C.P., Freitas, F.S., Costa, M.D., et al. 2013. Diversity of endophytic bacteria in the fruits of *Coffee canephora*. Afr. J. Microbiol. Res. **7**(7): 586–594. doi:10.5897/AJMR12.2036.

Miller, T.L., Wolin, M.J., Conway de Macario, E., and Macario, A.J. 1982. Isolation of *Methanobrevibacter smithii* from human feces. Appl. Environ. Microbiol. **43**(1): 227–232. PMID:6798932.

Montavon, P., Duruz, E., Rumo, G., and Pratz, G. 2003. Evolution of green coffee protein profiles with maturation and relationship to coffee cup quality. J. Agric. Food Chem. **51**(8): 2328–2334. doi:10.1021/jf020831j. PMID:12670177.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. **59**(3): 695–700. PMID:7683183.

Nubel, U., Engelen, B., Felske, A., Snaidr, J., Wieshuber, A., Amann, R., et al. 1996. Sequence heterogeneities of genes encoding 16S rRNAs in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. J. Bacteriol. **178**(19): 5636–5643. PMID:8824607.

Ochsenreiter, T., Selezi, D., Quaiser, A., Bonch-Osmolovskaya, L., and Schleper, C. 2003. Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. Environ. Microbiol. **5**(9): 787–797. doi:10.1046/j.1462-2920.2003.00476.x. PMID:12919414.

Pasin, L.A.A.P., Abreu, M.S.d., and Souza, I.P. 2011. Influence of the fungi population on the physicochemical and chemical composition of coffee (*Coffea arabica* L.). Ciência e Tecnologia de Alimentos, **31**: 681–687.

Prakitchaiwattana, C.J., Fleet, G.H., and Heard, G.M. 2004. Application and evaluation of denaturing gradient gel electrophoresis to analyse the yeast ecology of wine grapes. FEMS Yeast Res. **4**(8): 865–877. doi:10.1016/j.femsyr.2004.05.004. PMID:15450194.

Raskin, L., Stromley, J., Rittmann, B., and Stahl, D. 1994. Group-specific 16S ribosomal-RNA hybridization probes to describe natural communities of methanogens. Appl. Environ. Microbiol. **60**(4): 1232–1240. PMID:7517128.

Sakiyama, C., Paula, E., Pereira, P., Borges, A., and Silva, D. 2001. Characterization of pectin lyase produced by an endophytic strain isolated from coffee cherries. Lett. Appl. Microbiol. **33**(2): 117–121. doi:10.1046/j.1472-765x.2001.00961.x. PMID:11472518.

Sambrook, J., and Russell, D.W. 2001. Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA.

Silva, C., Schwan, R., Dias, E., and Wheals, A. 2000. Microbial diversity during maturation and natural processing of coffee cherries of *Coffea arabica* in Brazil. Int. J. Food Microbiol. **60**(2–3): 251–260. doi:10.1016/S0168-1605(00)00315-9. PMID:11016614.

Silva, C., Batista, L., Abreu, L., Dias, E., and Schwan, R. 2008. Succession of bacterial and fungal communities during natural coffee (*Coffea arabica*) fermentation. Food Microbiol. **25**(8): 951–957. doi:10.1016/j.fm.2008.07.003. PMID:18954729.

Silva, H.S.A., Tozzi, J.P.L., Terrasan, C.R.F., and Bettiol, W. 2012. Endophytic microorganisms from coffee tissues as plant growth promoters and biocontrol agents of coffee leaf rust. Biol. Control, **63**: 62–67. doi:10.1016/j.biocontrol.2012.06.005.

Sun, L., Qiu, F., Zhang, X., Dai, X., Dong, X., and Song, W. 2008. Endophytic bacterial diversity in rice (*Oryza sativa* L.) roots estimated by 16S rDNA sequence analysis. Microb. Ecol. **55**(3): 415–424. doi:10.1007/s00248-007-9287-1. PMID:17690836.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**(10): 2731–2739. doi:10.1093/molbev/msr121. PMID:21546353.

Ulrich, K., Ulrich, A., and Ewald, D. 2008. Diversity of endophytic bacterial communities in poplar grown under field conditions. FEMS Microbiol. Ecol. **63**(2): 169–180. doi:10.1111/j.1574-6941.2007.00419.x. PMID:18199082.

USDA. 2011. Coffee: world markets and trade. Circular Series 2011. Available from http://www.usda.gov [accessed 10 November 2012].

Vega, F., Pava-Ripoll, M., Posada, F., and Buyer, J. 2005. Endophytic bacteria in *Coffea arabica* L. J. Basic Microbiol. **45**(5): 371–380. doi:10.1002/jobm.200410551. PMID:16187260.

Velazquez, E., Rojas, M., Lorite, M., Rivas, R., Zurdo-Pineiro, J., Heydrich, M., and Bedmar, E. 2008. Genetic diversity of endophytic bacteria which could be find in the apoplastic sap of the medullary parenchym of the stem of healthy sugarcane plants. J. Basic Microbiol. **48**(2): 118–124. doi:10.1002/jobm.200700161. PMID:18383223.

Vilela, D.M., Pereira, G.V., Silva, C.F., Batista, L.R., and Schwan, R.F. 2010. Molecular ecology and polyphasic characterization of the microbiota associated with semi-dry processed coffee (*Coffea arabica* L.). Food Microbiol. **27**(8): 1128–1135. doi:10.1016/j.fm.2010.07.024. PMID:20832694.

Weisburg, W., Barns, S., Pelletier, D., and Lane, D. 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. **173**(2): 697–703. PMID:1987160.

West, E.R., Cother, E.J., Steel, C.C., and Ash, G.J. 2010. The characterization and diversity of bacterial endophytes of grapevine. Can. J. Microbiol. **56**(3): 209–216. doi:10.1139/w10-004. PMID:20453907.

Yrjala, K., Katainen, R., Jurgens, G., Saarela, U., Saano, A., Romantschuk, M., and Fritze, H. 2004. Wood ash fertilization alters the forest humus Archaea community. Soil Biol. Biochem. **36**(1): 199–201. doi:10.1016/j.soilbio.2003.09.003.