

# Diversity of the total bacterial community associated with Ghanaian and Brazilian cocoa bean fermentation samples as revealed by a 16 S rRNA gene clone library

Tamara Garcia-Armisen · Zoi Papalexandratou ·  
Hugo Hendryckx · Nicholas Camu · Gino Vrancken ·  
Luc De Vuyst · Pierre Cornelis

Received: 16 March 2010 / Revised: 25 May 2010 / Accepted: 25 May 2010 / Published online: 18 June 2010  
© Springer-Verlag 2010

**Abstract** Cocoa bean fermentation is a spontaneous process involving a succession of microbial activities, starting with yeasts, followed by lactic acid bacteria and acetic acid bacteria. So far, all microbiological studies about cocoa bean fermentation were based on culture-dependent (isolation, cultivation, and identification), or, more recently, culture-independent (PCR-DGGE, or polymerase chain reaction denaturing gradient gel electrophoresis) methods. Using a metagenomic approach, total DNA was extracted from heap and box fermentations at different time points and from different locations (Ghana and Brazil, respectively) to generate a 16 S rDNA clone library that was sequenced. The sequencing data revealed a low bacterial diversity in the

fermentation samples and were in accordance with the results obtained through culture-dependent and a second, culture-independent analysis (PCR-DGGE), suggesting that almost all bacteria involved in the fermentation process are cultivable. One exception was the identification by 16 S rDNA library sequencing of *Gluconacetobacter* species of acetic acid bacteria that were not detected by the two other approaches. The presence of *Enterobacteriaceae* related to *Erwinia/Pantoea/Tatumella*, as revealed by 16 S rDNA library sequencing, suggests an impact of these bacteria on fermentation.

**Keywords** Cocoa fermentation · Lactic acid bacteria · Acetic acid bacteria · Enterobacteria · 16 S rRNA gene clone library sequencing · PCR-DGGE

Tamara Garcia-Armisen and Zoi Papalexandratou contributed equally to this work.

T. Garcia-Armisen · H. Hendryckx · P. Cornelis (✉)  
Microbial Interactions, Department of Molecular and Cellular  
Interactions, Flanders Institute of Biotechnology (VIB),  
Vrije Universiteit Brussel,  
Pleinlaan 2,  
1050 Brussels, Belgium  
e-mail: pcornel@vub.ac.be

T. Garcia-Armisen · Z. Papalexandratou · N. Camu ·  
G. Vrancken · L. De Vuyst  
Research Group of Industrial Microbiology and Food  
Biotechnology (IMDO), Faculty of Sciences and Bio-engineering  
Sciences, Vrije Universiteit Brussel,  
Brussels, Belgium

*Present Address:*

N. Camu  
Fermentation Innovation, Barry Callebaut,  
Aalstersestraat 122,  
9280 Lebbeke, Belgium

## Introduction

Cocoa beans are the principal raw material for chocolate production (Beckett 2009). They originate as seeds embedded in a mucilaginous pulp in fruit pods of the tree *Theobroma cacao* L. Raw cocoa beans have to be fermented, dried, and roasted to obtain the desired characteristic cocoa flavor and taste (Thompson et al. 2001). After removal of the beans from the pods, the first step in cocoa processing is a spontaneous 3- to 6-day fermentation of the pulp-bean mass, carried out in heaps or boxes, wherein a succession of microbial activities of yeasts, lactic acid bacteria (LAB), and acetic acid bacteria (AAB) takes place (Schwan and Wheals 2004). During fermentation, ethanol and acetic acid diffuse into the beans, and this, in combination with the heat produced during fermentation, causes the death of the seed embryo as well

as the end of the fermentation. This step in turn initiates physical and biochemical changes in the beans, leading to the formation of precursor molecules for the development of a characteristic flavor and color of the beans (Hansen et al. 1998). During the last decade, knowledge about the spontaneous cocoa bean fermentation process has increased (Ardhana and Fleet 2003; Schwan and Wheals 2004; Jespersen et al. 2005; Nielsen et al. 2005; Camu et al. 2007; Nielsen et al. 2007b; De Vuyst et al. 2008; Kostinek et al. 2008; Camu et al. 2008a; Camu et al. 2008b). Some of these and all older studies have typically involved cultivation of microorganisms, colonies being counted to obtain a quantitative picture of the populations, followed by their isolation and identification mainly by reference to phenotypic characteristics (Ardhana and Fleet 2003). Such conventional isolation procedures inevitably require some amount of prior knowledge about the types of microorganisms that may be present in the population to be studied. Selective media and culture conditions used in these studies will enhance recovery of some species, but others may be poorly recovered and some may not be recovered at all on standard laboratory media (Giraffa and Neviani 2001). To obtain a reliable identification of isolates, multiple phenotypic and genotypic techniques are often combined in a polyphasic approach, wherein the shortcomings of one method are compensated by the use of another method (Jespersen et al. 2005; Nielsen et al. 2005; Camu et al. 2007; Nielsen et al. 2007b; Camu et al. 2008a; Camu et al. 2008b). In recent years, researchers have started to use culture-independent techniques to identify possibly uncultivable but potentially important players in the cocoa bean fermentation process (Jespersen et al. 2005; Nielsen et al. 2005, 2007b; Camu et al. 2007, 2008a, b). These methods are based on nucleic acids, such as the 16 S rRNA gene in the case of bacteria, to obtain both a qualitative and a semiquantitative picture of a microbial community without the need to isolate its single components. Access to rRNA gene databases and comparative gene sequence analysis software has simplified the identification of microorganisms (Pace et al. 1986). A widely used method, denaturing gradient gel electrophoresis (DGGE), has shown to be of great potential for the study of microbial population dynamics in food (Ercolini 2004; Ercolini et al. 2006; Fontana et al. 2006; Jany and Barbier 2008). For bacteria, polymerase chain reaction (PCR)-DGGE analysis is based on the separation of PCR-amplified 16 S rDNA gene fragments in a gradient of denaturing agents, providing a 'genetic fingerprint' that is representative of the bacterial population in a sample. This technique has also been used as part of a polyphasic analysis of the bacteria and yeast populations associated with cocoa bean fermentation (Nielsen et al. 2005; Camu

et al. 2007; Nielsen et al. 2008; Camu et al. 2008b; Daniel et al. 2009). Polyphasic, bacterial analyses of Ghanaian cocoa bean fermentation processes have revealed the occurrence of the LAB species *Enterococcus casseliflavus*, *Fructobacillus pseudoficulneus*, *Lactobacillus plantarum*, and *Leuconostoc pseudomesenteroides* and of various *Acetobacter* species in the beginning of the fermentation and the dominance by *Lactobacillus fermentum* and *Acetobacter pasteurianus* towards the end (Camu et al. 2007; Nielsen et al. 2007b; Camu et al. 2008b). Also, this approach has allowed the description of the new LAB species *Lactobacillus cacaonum* (De Bruyne et al. 2009a), *Lactobacillus fabifermentans* (De Bruyne et al. 2009a), *Lactobacillus ghanensis* (Nielsen et al. 2007a), *Weissella fabaria* (De Bruyne et al. 2009b), and *Weissella ghanensis* (De Bruyne et al. 2008), and the new AAB species *Acetobacter fabarum* (Cleenwerck et al. 2008), *Acetobacter ghanensis* (Cleenwerck et al. 2007), and *Acetobacter senegalensis* (Camu et al. 2007; Ndoye et al. 2007).

Another culture-independent approach to explore the bacterial species diversity of natural ecosystems is the cloning and sequencing of ribosomal RNA (rRNA) genes of the whole community. This can be accomplished by their nearly complete amplification, making use of primers considered as universal for bacteria (Lane 1991) or Archaea (Baker et al. 2003). Sequencing of such 16 S rRNA gene clone libraries, hereafter referred to as 16 S rDNA libraries, are an appropriate tool for a more precise picture of the species diversity in a sample taken from a given ecosystem. Such libraries are now extensively used to reveal the diversity of microbial life (Acinas et al. 2004; Tringe et al. 2005; von Mering et al. 2007), but have never been constructed before to analyze food fermentation processes in general and the cocoa bean fermentation process in particular. Therefore, common DNA extraction and PCR amplification methods have to be adopted and optimized to be able to work with good yields of high-quality DNA from a complex fermentation matrix (Marchesi et al. 1998).

The aim of the present study was to use the 16 S rDNA library sequencing approach, as part of a whole-metagenome study of the cocoa bean fermentation process (unpublished results), as an attempt to confirm the cultivable bacterial communities and identify possible unknown players (cultivable or not) in this not completely elucidated fermentation. Therefore, samples from two different locations (Ghana and Brazil) and taken at different times were used. The present study further aimed to reveal the total bacterial community richness and refine the estimation of the relative abundance of each group of bacteria in the cocoa bean fermentation samples, whether or not reflecting their estimated functional role in this process, by comparing the 16 S rDNA library sequencing data with optimized 16 S rRNA-PCR-DGGE analysis data.

## Materials and methods

### Cocoa bean fermentation samples

Field experiments were set up in Ghana (main crop of 2005) and Brazil (main crops of 2006 and 2007) to sample cocoa bean heap and box fermentations, respectively, according to a fixed time schedule for determination of the bacterial diversity of the fermentation process, both culture-dependently and culture-independently, as described previously (Camu et al. 2007; Zoi Papalexandratou, Gino Vrancken, Nicholas Camu, and Luc De Vuyst, unpublished results). Multiple microbiological and chemical analyses of the fermentation samples were carried out as described previously (Camu et al. 2007). Six cocoa bean fermentation samples were selected for construction of a 16 S rDNA library (Table 1) corresponding to time points with a maximum bacterial population density as revealed by plating (see below); although complete cocoa bean fermentation processes last 3–6 days, the greatest changes of the microbial communities occur during the first 3 days with a maximum population of LAB at around 48 h and of AAB at around 84 h of fermentation (Camu et al. 2007, 2008b). One sample was collected at 60 h of a representative Ghanaian cocoa bean heap fermentation (heap 5) carried out in 2005 (Camu et al. 2007); two samples were mixtures of two samplings each carried out at 48 and 84 h of two Brazilian cocoa bean box fermentations performed at two different farms (F1 and F2) in 2006, as each time a similar diversity was expected (Zoi Papalexandratou, Nicholas Camu, and Luc De Vuyst; unpublished results); two samples were mixtures of two samplings each carried out at 48 and 84 h of two Brazilian cocoa bean box fermentations performed at two different farms (B1 and B2) in 2007, as each time a similar diversity was expected (Zoi Papalexandratou, Gino Vrancken, and Luc De Vuyst; unpublished results). As a low-diversity microbial ecosystem was expected, a sixth sample composed of the mixture of samples at 48 h from the Brazilian box fermentations (2006 and 2007) was used. For 16 S rRNA-PCR-DGGE analysis, all samples mentioned above were analyzed individually.

### Direct extraction of DNA from fermented cocoa bean samples

DNA was extracted directly from the fermented cocoa bean samples according to the protocol of Camu et al. (2007) with some modifications. Briefly, 20 g of frozen pulp-bean samples were washed four times with 50 ml of phosphate-buffered saline. The combined fluids ( $\pm 200$  ml) were agitated at 4 °C for 1 h and subsequently centrifuged at  $100 \times g$  at 4 °C for 10 min to remove large particles. The supernatant was further centrifuged at  $8,000 \times g$  at 4 °C for

20 min to pellet the bacterial cells, which were subsequently frozen at  $-20^{\circ}\text{C}$  for at least 1 h. The thawed pellet was washed in 1 ml of TES buffer (6.7%, w/v, sucrose; 50 mM Tris-HCl; 1 mM EDTA, pH 8.0) and resuspended in 300  $\mu\text{l}$  of STET buffer (8%, w/v, sucrose; 5%, w/v, Triton X-100; 50 mM Tris-HCl, pH 8.0; 50 mM EDTA). Seventy-five microliters of lysis buffer [TES containing  $1,330 \text{ U ml}^{-1}$  mutanolysin (Sigma-Aldrich, St. Louis, MO, USA) and  $100 \text{ mg ml}^{-1}$  lysozyme (Sigma-Aldrich)] and 100  $\mu\text{l}$  of a solution of proteinase K [2.5 mg/ml of TE buffer (50 mM Tris-HCl; 1 mM EDTA, pH 8.0); VWR International, Darmstadt, Germany] were added, and the suspension was incubated at 37 °C for 1 h. After the addition of 40  $\mu\text{l}$  of preheated (37 °C) 20% (w/v) SDS in TE buffer, cells were incubated at 37 °C for 10 min, followed by 60-min incubation at 65 °C. One hundred microliters of TE buffer were added and the lysate was extracted with one volume of phenol-chloroform-isoamyl alcohol (49:49:1) (Sigma-Aldrich) for 30 s. Phases were separated by microcentrifugation (13,000 rpm at 4 °C for 5 min) using Phase Lock Gel tubes (Eppendorf AG, Hamburg, Germany). The aqueous phase was carefully mixed with 70  $\mu\text{l}$  of 5 M NaCl and 1 ml of isopropanol, and the DNA was precipitated on ice for at least 15 min. The DNA was collected by microcentrifugation (13,000 rpm at 4 °C for 30 min), and the pellet was washed in ice-cold 70% (v/v) ethanol. The DNA was air-dried and resuspended in 100  $\mu\text{l}$  of TE buffer. Five microliters of DNase-free RNase solution ( $10 \text{ mg ml}^{-1}$ ; Fermentas, St. Leon-Rot, Germany) was added, and the whole suspension was incubated at 37 °C for 10 min. The final samples were stored at  $-20^{\circ}\text{C}$  until further use.

### 16 S rRNA gene clone library construction

#### *Polymerase chain reaction*

Three pairs of primers that have been reported as ‘universal’ for bacteria were tried to amplify bacterial rRNA genes in the DNA extracted from the cocoa bean fermentation samples: 27f (5'-AGAGTTTGATCMTGGC TC-3') and 1492r (5'-GGYTACCTTGTTACGACTT-3') (Frank et al. 2008), 27f and 1392r (5'-CGGAACATGTGM GGCGGG-3') (Lane 1991), and 63f (5'-CAGGCCTAACA CATGCAAGTC-3') and 1387r (5'-CRTGTGTGGCGG GCA-3') (Marchesi et al. 1998). One pair of primers was used to amplify archaeal rRNA: 21f (5'-TTCCGGTTGATC CYGCCGGA-3') and 958r (5'-YCCGGCGTTGAMTCC AATT-3') (Bano et al. 2004). PCR amplifications were performed using an Eppendorf thermocycler in a final volume of 50  $\mu\text{l}$ , containing 5  $\mu\text{l}$  of extracted DNA (approximately 500 ng), 0.2 mM of each dATP, dCTP, dGTP, and dTTP, 0.4  $\mu\text{M}$  of each primer, 1.25 IU of *Taq*

**Table 1** Identification of bacterial species in the different cocoa bean fermentation samples by 16 S rDNA gene clone library sequencing (referred to as rDNA), expressed as the number of times the sequence was retrieved in each sample (number between brackets); by 16 S rRNA-PCR-DGGE (referred to as DGGE) with: I: LAC1-LAC2 primers in 35–60% denaturant gradients, II: 357f–518r universal primers in 35–70% denaturant gradients, and/or III: WBAC1–WBAC2 primers in 45–70% denaturant gradients; and by culturing on selective media (referred to IS) where *x* indicates the presence of a strain of the corresponding species among the isolates [identified through gene sequencing (16 S rRNA gene, *pheS*, *atpA*, and/or *rpoA*) for LAB or 16 S rRNA gene sequencing and/or DNA-DNA hybridization for AAB, both after classification by (GTG)<sub>5</sub>-PCR fingerprinting]

	Ghana 2005 (H5) 60h			Brazil 2006 (F1,F2) 84h			Brazil 2007 (B1,B2) 48h			Brazil 2007 (B1,B2) 84h			Mixed Brazil (F1,F2,B1,B2) 48h						
	rDNA (70)	DGGE	IS	rDNA (57)	DGGE	IS	rDNA (45)	DGGE	IS	rDNA (62)	DGGE	IS	rDNA (39)	DGGE	IS	rDNA (109)	DGGE	IS	
	<i>T<sub>i</sub></i> : 42.0°C; <i>T<sub>E</sub></i> : 23.0°C			<i>T<sub>i</sub></i> : 44.5°C; <i>T<sub>E</sub></i> : 27.6°C			<i>T<sub>i</sub></i> : 46.3°C; <i>T<sub>E</sub></i> : 22.2°C			<i>T<sub>i</sub></i> : 37.0°C; <i>T<sub>E</sub></i> : 23.1°C			<i>T<sub>i</sub></i> : 48.2°C; <i>T<sub>E</sub></i> : 21.9°C						
Acetic acid bacteria																			
<i>A. ghanensis/syzygii</i>	12	III	x	9	III	N/A	0	–	N/A	0	–	–	0	–	–	7	N/A	N/A	
<i>A. pasteurianus/pomorum</i>	4	–	–	1	–	N/A	41	III	N/A	1	–	x	8	III	x	0	N/A	N/A	
<i>A. lovaniensis/fabarium</i>	0	–	–	7	III	N/A	0	–	N/A	0	–	–	0	–	–	2	N/A	N/A	
<i>G. europaeus/swingsii</i>	1	–	–	2	–	N/A	1	–	N/A	0	–	–	0	–	–	1	N/A	N/A	
<i>Gluconacetobacter entanii</i>	0	–	–	1	–	N/A	0	–	N/A	0	–	–	0	–	–	0	N/A	N/A	
<i>Gluconacetobacter saccharivorans</i>	0	–	–	0	–	N/A	1	–	N/A	0	–	–	0	–	–	0	N/A	N/A	
<i>Gluconacetobacter persimomis</i>	0	–	–	0	–	N/A	1	–	N/A	0	–	–	0	–	–	0	N/A	N/A	
<i>A. senegalensis</i>	0	–	–	0	–	N/A	0	–	N/A	0	–	x	0	–	–	0	N/A	N/A	
Lactic acid bacteria																			
<i>L. fermentum</i>	51	I,III	x	18	I	N/A	0	I	N/A	61	I,II,III	x	28	I,II,III	x	89	N/A	N/A	
<i>L. cacaonum</i>	0	–	–	2	–	N/A	0	–	N/A	0	–	–	0	–	–	3	N/A	N/A	
<i>L. mali</i>	0	–	–	2	I	N/A	0	I	N/A	0	–	–	0	–	–	0	N/A	N/A	
<i>L. nagellii</i>	0	–	–	7	I, II	N/A	1	I	N/A	0	–	–	0	–	–	2	N/A	N/A	
<i>Lactobacillus brevis</i>	0	–	–	2	–	N/A	0	–	N/A	0	–	–	0	–	–	1	N/A	N/A	
<i>L. plantarum</i>	0	–	x	4	I, II	N/A	0	I	N/A	0	I	x	0	–	–	0	N/A	N/A	
<i>L. vaginalis</i>	0	–	–	0	–	N/A	0	–	N/A	0	–	–	3	–	–	1	N/A	N/A	
<i>L. durianis</i>	0	–	–	1	I, II	N/A	0	I	N/A	0	–	–	0	–	–	0	N/A	N/A	
<i>V. carniphilus</i>	0	–	–	1	–	N/A	0	–	N/A	0	–	–	0	–	–	0	N/A	N/A	
<i>Lc. pseudomesenteroides</i>	0	–	–	0	I, II	N/A	0	–	N/A	0	–	–	0	–	–	0	N/A	N/A	
<i>Weissella</i> sp.	0	–	–	0	–	N/A	0	–	N/A	0	I	–	0	–	–	0	N/A	N/A	
<i>Pediococcus acidilactici</i>	0	–	–	0	–	N/A	0	I	N/A	0	–	–	0	–	–	0	N/A	N/A	
Enterobacteriaceae																			
<i>Tatumella</i> sp.	1	–	N/A	1	–	N/A	0	–	N/A	0	–	N/A	0	–	N/A	2	N/A	N/A	
<i>Pantoea punctata</i>	1	–	N/A	0	–	N/A	0	–	N/A	0	–	N/A	0	–	N/A	0	N/A	N/A	
<i>E. tasmaniensis</i>	0	–	N/A	1	–	N/A	0	–	N/A	0	–	N/A	0	–	N/A	1	N/A	N/A	

For each sample, the following environmental parameters are listed: sampling time of the Ghanaian heap (H5) and Brazilian box (F1, F2, B1, B2) fermentations, pH of the pulp-bean mass, temperature [*T<sub>i</sub>*; either internal temperature of the pulp-bean mass (*T<sub>i</sub>*) or external temperature (*T<sub>E</sub>*) of the environment]. Hyphen (–) indicates that the species was not retrieved in the sample N/A not available

DNA polymerase (Qiagen, Hilden, Germany), 10 mM Tris–HCl, 50 mM KCl, and 1.5 mM MgCl<sub>2</sub>. One single PCR core program was used for all primer pairs: initial denaturation at 95 °C for 5 min; 20 cycles of denaturation at 95 °C for 1 min, annealing (using a gradient of 5 °C around the primer-specific temperature) for 1 min, and extension at 72 °C for 2 min; and a final extension at 72 °C for 10 min, followed by cooling to 4 °C.

#### Library construction and sequencing

After PCR amplification, unpurified 16 S rRNA PCR amplicons were cloned, using a TOPO TA Cloning Kit for sequencing (Invitrogen, Carlsbad, CA, USA), with the pCR4-TOPO vector into *Escherichia coli* (TOP 10) chemically competent cells, according to the manufacturer's instructions. The clones obtained were arrayed in 96-well plates, which were conserved as glycerol stocks for further analysis. The cloned inserts were reamplified using the vector primers M13 forward and reverse (25 cycles of 94 °C for 1 min, 48 °C for 1 min, and 72 °C for 2 min); the presence and size of the PCR amplicons were checked by agarose gel (0.8%, w/v) electrophoresis. PCR products were purified using a PCR Clean-Up kit (Qiagen) according to the manufacturer's instructions and sequenced in a commercial DNA sequencing facility (VIB Genetic Service Facility, Antwerp, Belgium) using the forward M13 primer and applying capillary sequencing technology. Sequences were manually trimmed to remove vector sequences and low-quality sequences and checked for chimera (CHECK\_CHIMERA online analysis of Ribosomal Database Project II; <http://rdp8.cme.msu.edu/cgis/chimera.cgi?su=SSU>).

#### Population diversity analysis

The BLAST (Basic Local Alignment Search Tool) program was used to search in the GenBank database (<http://www.ncbi.nlm.nih.gov/BLAST>) for the closest known relatives of the partial 16 S rRNA gene sequences obtained. Sequences were assigned as belonging to the same species, using a 98% sequence identity cutoff level. Rarefaction curves and richness estimators (Chao et al. 2005) were generated using EstimateS (Version 7; for each 16 S rRNA gene clone library as well as for the total of the libraries R. K. Colwell, <http://purl.oclc.org/estimates>). Estimations of the frequency of the different species, expressed as percentages of all sequences obtained for each rDNA clone library analyzed, were calculated as well.

#### 16 S rRNA-PCR-DGGE

For detection of LAB, a pair of primers (LAC1, LAC2) that targets the V3–V4 region of the 16 S rRNA gene of LAB

species was used (Walter et al. 2001), while a pair of primers (WBAC1, WBAC2) that targets the V7–V8 region of the 16 S rRNA gene of both AAB and LAB species was applied to detect AAB and LAB simultaneously (Lopez et al. 2003). A third pair of universal primers (357f–518r; Ercolini et al. 2001) was used to amplify DNA of the V3 region within the 16 S rRNA gene of all bacteria present. To facilitate separation by DGGE, a GC-rich sequence was attached to the LAC2, WBAC2, and 357f primers. PCR conditions and DGGE analyses were as reported previously (Camu et al. 2007; Vasilopoulos et al. 2008). To optimize the concentration of the DNA (1 µl) that was used as template for the PCR reactions, three different dilutions of purified DNA (500, 100, and 50 ng µl<sup>-1</sup>), extracted from cocoa bean fermentation samples of the same fermentation (B2, Brazil), were used to amplify the 16 S rRNA gene with universal primers, and the PCR amplicons were subjected to DGGE. Denaturant gradients from 35% to 60% and from 35% to 70% for the PCR products obtained with the LAC1–LAC2 and universal primers, respectively, were applied. For the amplicons resulting from PCR with the WBAC primers, an optimization of the denaturing gradients (35–60%, 35–70%, 45–70%, 50–70%, and 55–70%) was performed, under the same conditions of electrophoresis, during preliminary experiments. A 45–70% denaturing gradient was optimal (data not shown). Gel processing, DGGE band pattern cluster analysis (with the band-based Dice coefficient), DNA sequencing of excised bands, and BLAST analysis were performed as described previously (Camu et al. 2007).

## Results

#### Optimization of DNA extraction, PCR amplification, and 16 S rDNA clone library construction

To reduce PCR biases and to have a better idea of the relative abundance of the sequences retrieved, different aspects had to be considered in setting up the PCR amplification reactions: an adequate DNA extraction procedure had to be developed, the choice of the primers had to be optimized, and a reduced number of cycles and the lowest annealing temperature possible had to be used during the PCR reactions. These conditions were optimized during preliminary experiments, as mentioned in “Materials and methods.”

In all the samples, from the three pairs of universal primers used to amplify the bacterial 16 S rRNA gene, only the combination of 27f–1492r resulted in a proper amplification. With these primers, a single band of approximately 1,500 bp was obtained in all samples and for all tested annealing temperatures (48 °C, 49 °C, 50 °C, 51 °C, 52 °C, and 53 °C) after agarose gel electrophoresis. With the



archeal primers, a positive control (DNA from *Sulfolobus solfataricus* P2) allowed us to confirm the suitability of the primers and the PCR program, but no band was obtained from the cocoa samples, suggesting the absence of Archae in the samples.

For the construction of the libraries, a mix of two separate PCR reaction tubes, carried out at the lowest tested annealing temperature (48 °C), was used. From the total amount of clones obtained in each library (always more than 800), 400 clones were arrayed in 96-well microtiter plates for storage. From each clone, the inserts were amplified using vector primers and PCR products were sequenced until the rarefaction curves reached saturation, indicating a sufficient sample size.

#### Bacterial population diversity analysis through 16 S rRNA gene clone library sequencing

A total of 382 16 S rRNA gene clones from six separated libraries were sequenced (Table 1). Although the clone libraries were rather small, they were considered sufficient to represent a low-diversity ecosystem as shown below.

No chimera was detected. After trimming and BLAST analysis, all alignments presented identities higher than 98% with sequences retrieved in the database, allowing the identification of 19 different phylotypes. The three major bacterial families represented were the Acetobacteraceae (seven species), the Lactobacillaceae (eight species), and the Enterobacteriaceae that were represented by three genera, namely, *Tatumella*, *Pantoea*, and *Erwinia*.

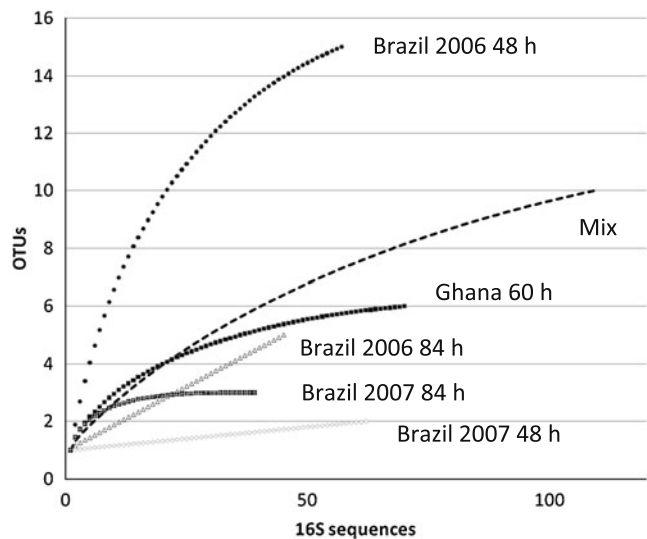
The overall dominant species was *L. fermentum*, which represented 60% of all the sequences, followed by *A. pasteurianus/pomorum* (15% of all sequences; no difference could be made between these two species based on the 16 S rRNA gene fragment sequenced). Species of the *Gluconacetobacter* genus as well as the *Enterobacteriaceae* species represented less than 1% of all the sequences.

A restricted bacterial diversity was found in the Ghanaian cocoa bean heap fermentation sample, with *L. fermentum* being the predominating species and the occurrence of enterobacterial species. The Brazilian samples of the box fermentations of 2006 harbored a wide LAB and AAB diversity (both 48- and 84-h samples) plus the presence of *Enterobacteriaceae* in the 48-h sample. *L. fermentum* and *A. ghanensis/syzygii* were the dominant species in the 48-h sample and *A. pasteurianus/pomorum* was by far the main bacterial representative in the 84-h sample, indicating the dominance of these fermentations by AAB. However, in the Brazilian samples of 2007, a more restricted bacterial diversity was observed. No *Enterobacteriaceae* were found in the 48-h fermentation sample and *L. fermentum* dominated the 84-h sample, followed by *A. pasteurianus*.

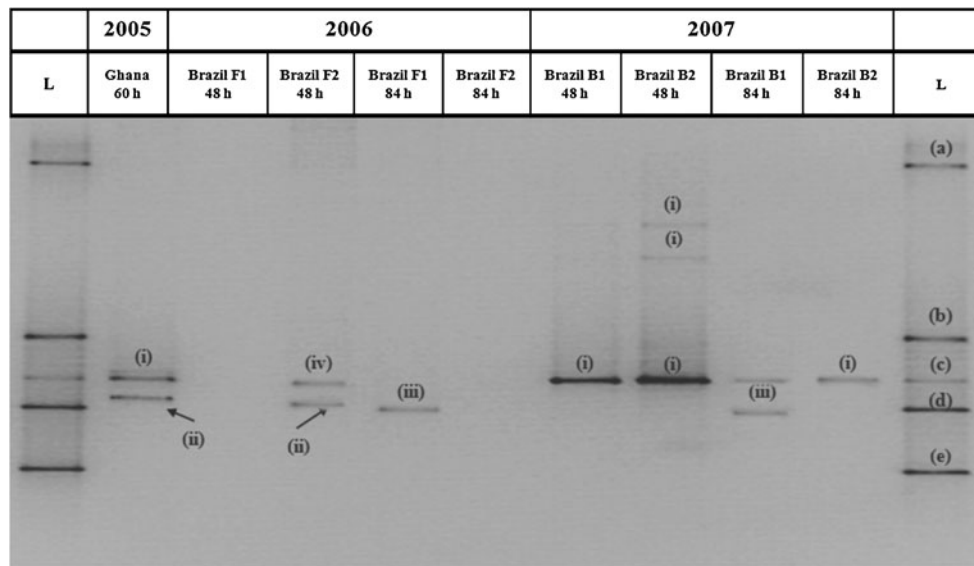
The six rarefaction curves presented in Fig. 1 correspond to the six different samples of this study. The sample of Brazil 2006 at 48 h was from far the one with the most diverse bacterial population. All curves showed saturation or started to saturate, except for the Brazil 2006 at 84 h, probably because this sample, despite being for 90% dominated by a single species (*L. fermentum*), contained four other phylotypes (operational taxonomic units [OTUs]) that appeared only once. Overall, these rarefaction curves suggest that the analysis of the 16 S rDNA library sequences gave a good estimation of the bacterial diversity in the cocoa bean fermentation samples. When analyzing a mixed sample (Brazil 2006 and Brazil 2007 48 h), we expected to find 17 phylotypes (as shown in Table 1) and the saturation curve indicates a minimum of 10 phylotypes (Fig. 1), but saturation was not reached, even after analyzing more than 100 sequences, demonstrating that dominance by one species can result in an underestimation of the population diversity.

#### 16 S rRNA-PCR-DGGE of cocoa bean fermentation samples

In contrast with what has been reported before for Ghanaian cocoa bean fermentation samples (Camu et al. 2007), WBAC primers allowed the detection of AAB species in the samples used for 16 S rDNA library construction (Fig. 2), thanks to the subjection of DNA of low concentration to PCR amplification (50 ng/μl). Indeed, an influence on and inhibition of the PCR amplification by highly concentrated DNA (as used in Camu et al. 2007) was seen (Fig. 3). Yet, some DGGE bands were of low intensity, still indicating the more difficult detection of



**Fig. 1** Separated rarefaction curves for the six 16 S rRNA gene clone library sequences mentioned in Table 1. The curve corresponding to “mixed” samples corresponds to a mix of samples coming from two 48-h fermentations done in Brazil in 2006 and 2007



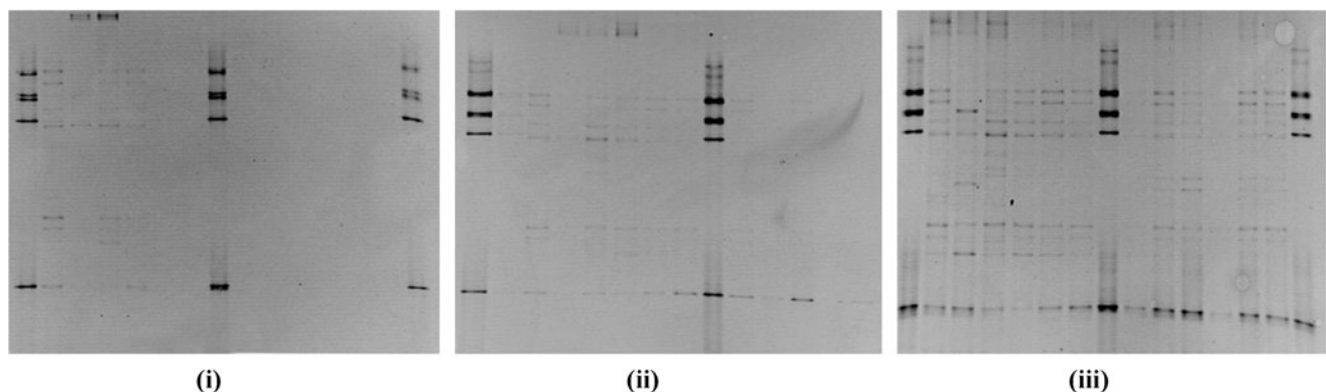
**Fig. 2** 16S rRNA-PCR-DGGE profiles with WBAC1–WBAC2 primers (45–70% denaturing gradient), representing LAB and AAB (sequenced bands) from the fermented cocoa bean samples of Ghana and Brazil. On top of each lane, the year, fermentation sample code, and sampling time are indicated. The reference ladder (L) consisted of **(a)** *L. plantarum* LMG 6907 T, **(b)** *Lc. pseudomesenteroides* 274 (Camu et al. 2007), **(c)** *L. fermentum* LMG 6902 T, **(d)** *A.*

*pasteurianus* LMG 1262 T, and **(e)** *G. europaeus* LMG 18890 T. The closest relatives of the gene fragments sequenced (the percentages of nucleotides corresponding with sequences retrieved from the GenBank database are shown in parentheses) were as follows: (i) *L. fermentum* (100%), (ii) *A. ghanensis/syzygii* (100%), (iii) *A. pasteurianus/pomorum* (100%), and (iv) *A. lovaniensis/fabarum* (100%)

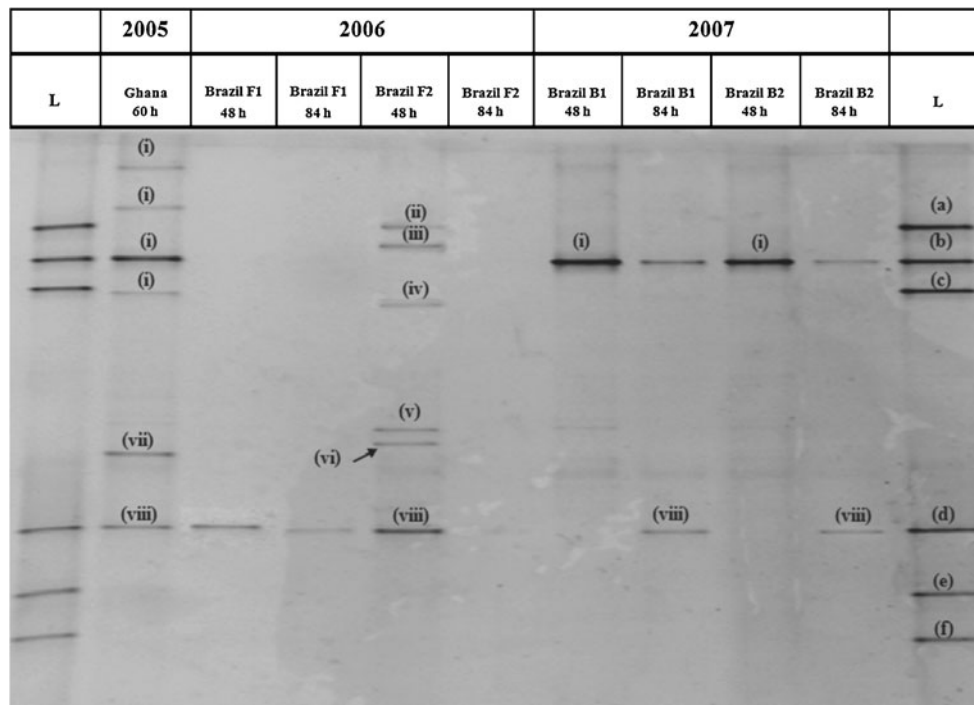
AAB of low population density by 16 S rRNA-PCR-DGGE (Fig. 2). Spiking of an Ivorian cocoa bean fermentation sample, obviously free from AAB but with or without *L. fermentum* as dominating species, with different cell densities (107, 105, and 103 CFU/ml) of a reference strain of *A. pasteurianus* (LMG 1262 T), followed by DNA extraction of the sample and 16 S rRNA-PCR-DGGE analysis with universal primers, allowed recovery of this species, even at the lowest concentration (data not shown). In addition, universal primers allowed the detection of AAB in the original samples, but only at genus level, as the nucleotide sequence that was amplified with this primer pair seems to be present in all AAB species (Fig. 4).

16 S rRNA gene clone library sequencing analysis in comparison with 16 S rRNA-PCR-DGGE and culturing

For the AAB, the 16 S rDNA library revealed a higher diversity compared to 16 S rRNA-PCR-DGGE and culturing. For instance, in the 60-h sample of Ghana 2005, only *A. ghanensis* was found through 16 S rRNA-PCR-DGGE and culturing, while several *Acetobacter* species plus a *Gluconacetobacter* species were found through sequencing of the 16 S rDNA library (Table 1). Moreover, the latter showed the presence of four species of *Gluconacetobacter*, in samples from Ghana and Brazil 2006, which were not detected by the DGGE analysis or by cultivation.



**Fig. 3** Influence of the concentration of template DNA used for PCR amplification with universal primers (357f–518r), carried out on samples of the cocoa bean box (B2) fermentation in Brazil, on the visualization of DGGE bands: (i) 500 ng  $\mu\text{l}^{-1}$ , (ii) 100 ng  $\mu\text{l}^{-1}$ , and (iii) 50 ng  $\mu\text{l}^{-1}$



**Fig. 4** 16S rRNA-PCR-DGGE profiles with universal primers (35–70% denaturing gradient), representing bacteria (sequenced bands) from the fermented cocoa bean samples of Ghana and Brazil. On top of each lane, the year, fermentation sample code, and sampling time are indicated. The reference ladder (L) consisted of **(a)** *L. plantarum* LMG 6907 T, **(b)** *L. fermentum* LMG 6902 T, **(c)** *Lc. pseudomesenteroides* 274 (Camu et al. 2007), **(d)** *A. pasteurianus* LMG 1262 T, **(e)** *G. europaeus* LMG 18890 T, and **(f)** *A. senegalensis* LMG

23690 T. The closest relatives of the gene fragments sequenced (the percentages of nucleotides corresponding with sequences retrieved from the GenBank database are shown in parentheses) were as follows: (i) *L. fermentum* (100%), (ii) *L. plantarum* (100%), (iii) *L. durianis* (100%), (iv) *Lc. pseudomesenteroides* (100%), (v) *L. nagelii* (98%), (vi) *L. nagelii* (100%), (vii) chloroplast DNA, and (viii) AAB (100%)

For the LAB, when analyzing all the samples as a single pool, the results obtained by 16 S rDNA library sequencing did not differ substantially from 16 S rRNA-PCR-DGGE or traditional cultivation methods (Table 1). However, some interesting discrepancies were observed, as exemplified by *Lc. pseudomesenteroides*, *Pediococcus acidilactici*, and *Weissella* sp. that were detected only by the 16 S rRNA-PCR-DGGE method (both LAC and universal primers in some of the samples), while *Lactobacillus vaginalis* was only found by 16 S rDNA library sequencing analysis in the Brazil 2007 sample at 84 h of fermentation. Likewise, *Vagococcus carniphilus* was only detected by 16 S rDNA library sequencing in the 48-h sample from Brazil 2006.

Large differences could be observed when analyzing the fermentation samples separately (Table 1). In the 48-h sample from Brazil 2006, 16 S rDNA library sequencing analysis revealed species that were found neither by 16 S rRNA-PCR-DGGE nor through culture analysis. However, in the 84-h sample of Brazil 2006, where LAB represented only 2% of all sequences, 16 S rRNA-PCR-DGGE analysis revealed the presence of six different LAB species (*L. plantarum*, *Lactobacillus mali*, *Lactobacillus durianis*, *L. fermentum*, *Pediococcus acidilactici*, *Lactobacillus nagelii*),

while 16 S rDNA library sequencing analysis only detected one (*L. nagelii*). However, it has to be mentioned that the DGGE bands of the LAB species present in this sample were of low intensity, except for the band of *L. nagelii*. Hence, it is likely that 16 S rDNA library sequencing analysis allowed the detection of the predominant LAB species solely, as this particular fermentation was dominated by AAB, indicating that 16 S rDNA library sequencing gave an idea of the relative abundance of the different species. Also, in the Brazilian 2006 samples, an interesting shift between the species of AAB present at 48 and 84 h could be observed. *A. ghanensis* and *Acetobacter lovaniensis*, which dominated the sample at 48 h, disappeared at 84 h into fermentation, while *A. pasteurianus* increased from 2% to 91%. Brazilian samples from the 2007 experimental fermentation presented a very different composition, since at 48 h 99% of the 16 S rDNA library sequences corresponded with *L. fermentum*. This result was consistent with the results found by 16 S rRNA-PCR-DGGE and through molecular identification of the isolates. Likewise, only one species of AAB, *A. pasteurianus*, was detected in the same sample with an increasing abundance from 1% at 48 h to about 20% at 84 h into fermentation.



Few *Enterobacteriaceae* sequences (16 S rDNA library), corresponding to *Tatumella* sp. and *Pantoea punctata* [now reclassified as *Tatumella punctata* (Brady et al. 2009)] were present in the Ghanaian sample, while the same *Tatumella* sp. and *Erwinia tasmaniensis* were present in one sample (48 h) of Brazil 2006. No *Enterobacteriaceae* could be detected by the 16 S rRNA-PCR-DGGE method (universal primers), probably because of their low abundance.

The rarefaction curves in Fig. 1 show the differences in the species diversity of the cocoa bean fermentation samples. Biodiversity estimators (Chao 2) illustrate these differences, even in the same country at the same time point; for instance, in the 48- and 84-h samples of Brazil 2007, the maximum numbers of expected phylotypes (Chao 2, upper bound of the 95% confidence interval) were two and seven, respectively, while in the sample of Brazil 2006, a maximum of 42 different phylotypes (OTUs) could be expected at 48 h and 33 phylotypes at 84 h into fermentation, revealing a higher bacterial diversity, albeit still confined to LAB and AAB.

## Discussion

In recent years, both culture-dependent and culture-independent approaches have been used to study the bacterial composition of cocoa bean fermentation samples (Camu et al. 2007, 2008b; Nielsen et al. 2005, 2007b). To the best of our knowledge, none of the previous studies on the cocoa bean fermentation microbiota were based on the sequencing of a 16 S rDNA library. Despite the recent development of 454 pyrosequencing, the classic approach used here, which consists in cloning and sequencing, remains the ‘gold standard’ for identification of lineages (Tringe and Hugenholtz 2008). Sequencing of a 16 S rDNA library from six different samples (one from a Ghanaian heap fermentation carried out in 2005, two from Brazilian box fermentations carried out in 2006, two from Brazilian box fermentations carried out in 2007, and a mixed sample of the Brazilian box fermentations 2006–2007), using universal primers for bacteria, globally confirmed the results obtained by 16 S rRNA-PCR-DGGE and molecular identification of isolates. Although the 27f–1492r primers have been shown to amplify the 16 S rRNA genes of most bacterial groups (Lane 1991), only LAB and AAB were detected in the cocoa bean fermentation samples analyzed, except for few sequences corresponding to *Enterobacteriaceae*. Other authors (Tringe et al. 2005) using the same pair of primers detected 847 distinct ribotypes from more than a dozen phyla in soil samples, indicating that the primers used allow the amplification of very diverse bacterial 16 S rRNA genes. Therefore, the present study

can in confidence confirm that the cocoa bean fermentation process is almost exclusively mediated by LAB and AAB as far as the bacterial players are concerned. Indeed, the primers used in the present study only aimed at the 16 S rRNA genes of bacterial species, which means that eukaryotic microorganisms such as yeasts, which are known to play an important role in the beginning of the cocoa bean fermentation process (ethanol production and pectin degradation), are missed (Jespersen et al. 2005; Nielsen et al. 2005; Daniel et al. 2009). In particular, it turned out that *L. fermentum* was by far the most dominant species, followed by *A. pasteurianus*. Taking the metabolite analysis data into account (Camu et al. 2007, 2008b), the former species is responsible for heterolactate fermentation of glucose and the production of mannitol by using fructose as alternative external electron acceptor, while the latter species oxidizes ethanol—formed by the yeasts—into acetic acid. Although there was a good agreement between 16 S rRNA-PCR-DGGE and 16 S rDNA library sequencing analysis, some interesting differences have to be noticed. Notably, only the 16 S rDNA library sequencing approach allowed the detection of *Gluconacetobacter* species that were not found when using the primers WBAC1–WBAC2 or 357f–518r in the 16 S rRNA-PCR-DGGE protocol. Interestingly, this group of AAB has not been reported in any of the recent studies on cocoa bean fermentation, neither by cultivation or culture-independent methods, which always revealed the predominance of *Acetobacter* species (Ardhana and Fleet 2003; Camu et al. 2007; Nielsen et al. 2007b; Camu et al. 2008b). The presence of another representative of AAB, *Gluconobacter oxydans*, previously reported in Ghanaian cocoa bean fermentation (Nielsen et al. 2007b), could not be confirmed in the present study. Therefore, it is likely that different fermentation methods under different conditions and in different countries could slightly influence the bacterial composition. However, whereas *Glucono(aceto)bacter* prefers glucose as the energy source, which has been depleted by both yeasts and LAB in the early phases of cocoa bean fermentation, *Acetobacter* preferentially oxidizes ethanol into acetic acid, a key metabolic trait during cocoa bean fermentation (Camu et al. 2008b). Thus, the presence of (unfermented) glucose is necessary for later growth of *Glucono(aceto)bacter* compared to LAB, which may reflect suboptimal fermentation conditions of the “pulp-bean mass, as indicated by metabolite analysis data (Zoi Papalexandratou, Gino Vrancken, and Luc De Vuyst; unpublished results).

Some of the novel species, recently isolated from cocoa bean fermentation samples, were retrieved as well in the present study, such as *L. cacaonum* (De Bruyne et al. 2009a) and *A. ghanensis* (Cleenwerck et al. 2007). Their relatively high prevalence indicates a possible functional role during the cocoa bean fermentation process. Their

detection will depend on the sensitivity of the techniques used. Particularly in the case of AAB, a 16 S rDNA library seems to represent a better approach to unravel bacterial diversity, as the higher length of the sequences obtained (800 bp on average after trimming) allowed a more accurate identification of bacteria than the 16 S rRNA-PCR-DGGE method (bands of 380, 310, and 217 bp, respectively, for the LAC, WBAC, and universal primers, respectively). Furthermore, when comparing the 16 S rDNA library sequences of all AAB retrieved in these samples, it can be noticed that differences are slight in general, and in some cases only few nucleotides are different among the sequences. For instance, between the reference 16 S rRNA gene sequences of *Gluconacetobacter europaeus* and *Gluconacetobacter Swingsii*, a single nucleotide substitution occurs. Likewise, between *A. pasteurianus* and *A. pomorum*, there are only two substitutions, while between *A. ghanensis* and *A. syzygii*, only four nucleotides are different, explaining why they are presented together in Table 1.

Large differences in species diversity could be noticed between the different cocoa bean fermentation samples as well, illustrating the importance of the fermentation conditions on the bacterial composition. The Brazilian 2006 samples showed the highest bacterial diversity, but corresponded with a fermentation carried out with less care (Zoi Papalexandratou, Nicholas Camu and Luc De Vuyst, unpublished results). Indeed, for the 2006 fermentation in Brazil, the pods were not carefully selected and hence included both healthy and infected beans, the placenta was not removed, and there was apparently no turning of the beans during fermentation. The high biodiversity of the 2006 Brazilian samples can thus be explained by a high—rather irrelevant—inoculation from the environment. The dominance of *A. pasteurianus* in this sample was probably due to the fact that this AAB species is better adapted to the cocoa bean fermentation ecosystem, for instance to the increase of the temperature upon fermentation (Camu et al. 2007). Dominance of AAB during cocoa bean fermentation often results in sour beans, as the citrate and sugar fermentation by the LAB in the pulp-bean mass is taken over too early by ethanol oxidation by the AAB (Schwan 1998; Camu et al. 2008b). The restricted species diversity of the Brazilian 2007 samples seemed to be the result of more controlled practices, resulting in excellent fermentation courses and metabolite concentrations (Zoi Papalexandratou, Gino Vrancken, and Luc De Vuyst, unpublished results), as was the case for the Ghanaian fermentation sample (Camu et al. 2007), Ghana usually representing good fermentation practices (Leiter and Harding 2004).

*Enterobacteriaceae* sequences, in particular *Tatumella* sp., *Pantoea punctata*, and *Erwinia tasmaniensis*, were

only found in samples from Ghana 2005 and Brazil 2006. Some enterobacterial species are present in soil (Kageyama et al. 1992; Pujol and Kado 2000) and this could be the source of these bacteria in the pulp-bean mass. These species, especially *Tatumella* species, could occur as phytopathogens in cocoa pods, as a *Tatumella* sp. has been reported as a phytopathogen in pineapple fields in Mexico (Marin-Cevada et al. 2007). Alternatively, *E. tasmaniensis* has been described as a non-phytopathogenic *Erwinia*, the genome of which has recently been published (Kube et al. 2008). Given the fact that *Pantoea* species have already been reported in cocoa bean fermentation processes in Ghana before (Camu et al. 2008b), *Pantoea* and other *Enterobacteriaceae* may play a role in citrate fermentation at the initial stages of the cocoa bean fermentation process (Zoi Papalexandratou, Gino Vrancken, and Luc De Vuyst, unpublished results). Alternatively, two species of *Pantoea* have been described as phytopathogens as well, *P. agglomerans* and *P. ananatis*, infecting pineapple (Barash and Manulis-Sasson 2009; Coutinho and Venter 2009). Therefore, these *Enterobacteriaceae* could be colonizers of the cocoa fruit and persist without being involved during the fermentation process.

To conclude, the present study suggests that most of the bacterial species implicated in the cocoa bean fermentation process have previously been well identified and that almost all are cultivable. 16 S rDNA library sequencing analysis allowed us to reveal the presence of the *Gluconacetobacter* species in these fermentations and further confirmed the results of both 16 S rRNA-PCR-DGGE and isolate identifications, namely that *L. fermentum* is by far the dominant species, followed by *A. pasteurianus*. As the fermentation samples displayed limited bacterial species diversity, metagenomic 16 S rRNA gene sequencing could not change the overall known picture about cocoa bean fermentation, but it helped to reveal the possible interference of *Gluconacetobacter* species and highlighted the presence of *Enterobacteriaceae*. The finding that some species detected by sequencing were not isolated by plating can be important because these members may play an important role in the fermentation process and fermented food quality in general, yet would have been ignored if only cultivation-based methods were used. Meta-transcriptomic analysis could be an additional approach, as it will allow detecting the genes at work in each step of the fermentation process. Furthermore, as the implicated species are well known and as almost all are cultivable, it will be possible to make the link to know for each time point which species are doing what, in particular with respect to the involvement of the novel LAB and AAB species described recently, *Gluconacetobacter* species among the other AAB, and whether *Enterobacteriaceae* are involved in fermentation or plant disease.

**Acknowledgements** This research was funded by a Concerted Research Action (Geconcerteerde Onderzoeksactie GOA47) from the Research Council of the Vrije Universiteit Brussel. Also, financial contributions from the Fund for Scientific Research-Flanders and Barry-Callebaut are acknowledged. Daniel Charlier is thanked for providing *Sulfolobus solfataricus* DNA.

## References

- Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL, Polz MF (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430:551–554
- Ardhana MM, Fleet GH (2003) The microbial ecology of cocoa bean fermentations in Indonesia. *Int J Food Microbiol* 86:87–99
- Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16 S primers. *J Microbiol Meth* 55:541–555
- Bano N, Ruffin S, Ransom B, Hollibaugh JT (2004) Phylogenetic composition of Arctic Ocean archaeal assemblages and comparison with Antarctic assemblages. *Appl Environ Microbiol* 70:781–789
- Barash I, Manulis-Sasson S (2009) Recent evolution of bacterial pathogens: the gall-forming *Pantoea agglomerans* case. *Annu Rev Phytopathol* 47:133–152
- Beckett ST (2009) Industrial chocolate manufacture and use, 4th edn. John Wiley & Sons Ltd., Chichester
- Brady CL, Venter SN, Cleenwerck I, Vandemeulebroecke K, de Vos P, Coutinho TA (2009) Transfer of *Pantoea citrea*, *Pantoea punctata* and *Pantoea terrea* to the genus *Tatumella* emend. as *Tatumella citrea* comb. nov., *Tatumella punctata* comb. nov. and *Tatumella terrea* (Kageyama et al. 1992) comb. nov. and description of *Tatumella morbirosei* sp. nov. *Int J Syst Evol Microbiol* 60:484–494
- Camu N, De Winter T, Addo SK, Takrama JS, Bernaert H, De Vuyst L (2008a) Fermentation of cocoa beans: influence of microbial activities and polyphenol concentrations on the flavour of chocolate. *J Sci Food Agric* 88:2288–2297
- Camu N, De Winter T, Verbrughe K, Cleenwerck I, Vandamme P, Takrama JS, Vancanneyt M, De Vuyst L (2007) Dynamics and biodiversity of populations of lactic acid bacteria and acetic acid bacteria involved in spontaneous heap fermentation of cocoa beans in Ghana. *Appl Environ Microbiol* 73:1809–1824
- Camu N, Gonzalez A, De Winter T, Van Schoor A, De Bruyne K, Vandamme P, Takrama JS, Addo SK, De Vuyst L (2008b) Influence of turning and environmental contamination on the dynamics of populations of lactic acid and acetic acid bacteria involved in spontaneous cocoa bean heap fermentation in Ghana. *Appl Environ Microbiol* 74:86–98
- Chao A, Chazdon RL, Colwell RK, Shen TJ (2005) A new statistical approach for assessing similarity of species composition with incidence and composition data. *Ecol Lett* 8:148–159
- Cleenwerck I, Camu N, Engelbeen K, De Winter T, Vandemeulebroecke K, De Vos P, De Vuyst L (2007) *Acetobacter ghanensis* sp. nov., a novel acetic acid bacterium isolated from traditional heap fermentations of Ghanaian cocoa beans. *Int J Syst Evol Microbiol* 57:1647–1652
- Cleenwerck I, Gonzalez A, Camu N, Engelbeen K, De Vos P, De Vuyst L (2008) *Acetobacter fabarum* sp. nov., an acetic acid bacterium from a Ghanaian cocoa bean heap fermentation. *Int J Syst Evol Microbiol* 58:2180–2185
- Coutinho TA, Venter SN (2009) *Pantoea ananatis*: an unconventional plant pathogen. *Mol Plant Pathol* 10:325–335
- Daniel HM, Vrancken G, Takrama JF, Camu N, De Vos P, De Vuyst L (2009) Yeast diversity of Ghanaian cocoa bean heap fermentations. *FEMS Yeast Res* 9:774–783
- De Bruyne K, Camu N, De Vuyst L, Vandamme P (2009a) *Lactobacillus fabifermentans* sp. nov. and *Lactobacillus cacao* sp. nov., isolated from Ghanaian cocoa fermentations. *Int J Syst Evol Microbiol* 59:7–12
- De Bruyne K, Camu N, De Vuyst L, Vandamme P (2009b) *Weissella fabaria* sp. nov. from Ghanaian cocoa fermentation. *Int J Syst Evol Microbiol*. doi:10.1099/ijs.1090.019323-019320
- De Bruyne K, Camu N, Lefebvre K, De Vuyst L, Vandamme P (2008) *Weissella ghanensis* sp. nov., isolated from a Ghanaian cocoa fermentation. *Int J Syst Evol Microbiol* 58:2721–2725
- De Vuyst L, Camu N, De Winter T, Vandemeulebroecke K, De Perre VV, Vancanneyt M, De Vos P, Cleenwerck I (2008) Validation of the (GTG)(5)-rep-PCR fingerprinting technique for rapid classification and identification of acetic acid bacteria, with a focus on isolates from Ghanaian fermented cocoa beans. *Int J Food Microbiol* 125:79–90
- Ercolini D (2004) PCR-DGGE fingerprinting: novel strategies for detection of microbes in food. *J Microbiol Meth* 56:297–314
- Ercolini D, Moschetti G, Blaiotta G, Coppola S (2001) Behavior of variable V3 region from 16 S rDNA of lactic acid bacteria in denaturing gradient gel electrophoresis. *Curr Microbiol* 42:199–202
- Ercolini D, Russo F, Torrieri E, Masi P, Villani F (2006) Changes in the spoilage-related microbiota of beef during refrigerated storage under different packaging conditions. *Appl Environ Microbiol* 72:4663–4671
- Fontana C, Coconcelli PS, Vignolo G (2006) Direct molecular approach to monitoring bacterial colonization on vacuum-packaged beef. *Appl Environ Microbiol* 72:5618–5622
- Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ (2008) Critical evaluation of two primers commonly used for amplification of bacterial 16 S rRNA genes. *Appl Environ Microbiol* 74:2461–2470
- Giraffa G, Neviani E (2001) DNA-based, culture-independent strategies for evaluating microbial communities in food-associated ecosystems. *Int J Food Microbiol* 67:19–34
- Hansen CE, del Olmo M, Burri C (1998) Enzyme activities in cocoa beans during fermentation. *J Sci Food Agric* 77:273–281
- Jany JL, Barbier G (2008) Culture-independent methods for identifying microbial communities in cheese. *Food Microbiol* 25:839–848
- Jespersen L, Nielsen DS, Honholt S, Jakobsen M (2005) Occurrence and diversity of yeasts involved in fermentation of West African cocoa beans. *FEMS Yeast Res* 5:441–453
- Kageyama B, Nakae M, Yagi S, Sonoyama T (1992) *Pantoea punctata* sp. nov., *Pantoea citrea* sp. nov., and *Pantoea terrea* sp. nov. isolated from fruit and soil samples. *Int J Syst Bacteriol* 42:203–210
- Kostinek M, Ban-Koffi L, Ottah-Atikpo M, Teniola D, Schillinger U, Holzapfel WH, Franz CMAP (2008) Diversity of predominant lactic acid bacteria associated with cocoa fermentation in Nigeria. *Curr Microbiol* 56:306–314
- Kube M, Migdoll AM, Muller I, Kuhl H, Beck A, Reinhardt R, Geider K (2008) The genome of *Erwinia tasmaniensis* strain Et1/99, a non-pathogenic bacterium in the genus *Erwinia*. *Environ Microbiol* 10:2211–2222
- Lane DJ (1991) 16S/23S rRNA sequencing. In: Stackebrandt EaG M (ed) Sequencing techniques in bacterial systematics. Wiley, London, pp 115–174
- Leiter J, Harding S (2004) Trinidad, Brazil, and Ghana: three melting moments in the history of cocoa. *J Rural Stud* 20:113–130
- Lopez I, Ruiz-Larrea F, Coccolin L, Orr E, Phister T, Marshall M, VanderGheynst J, Mills DA (2003) Design and evaluation of PCR primers for analysis of bacterial populations in wine by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* 69:6801–6807

- Marchesi JR, Sato T, Weightman AJ, Martin TA, Fry JC, Hiom SJ, Dymock D, Wade WG (1998) Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16 S rRNA. *Appl Environ Microbiol* 64:795–799
- Marin-Cevada V, Castaneda-Lucio M, Caballero-Mellado J, Bustillos-Cristales R, Munoz-Rojas J, Mascarua-Esparza MA, Fuentes-Ramirez LE (2007) *Tatumella* sp Enterobacteria involved in pink disease of pineapple. *Phytopathology* 97:S70–S70
- Ndoye B, Cleenwerck I, Engelbeen K, Dubois-Dauphin R, Guiro AT, Van Trappen S, Willems A, Thonart P (2007) *Acetobacter senegalensis* sp. nov., a thermotolerant acetic acid bacterium isolated in Senegal (sub-Saharan Africa) from mango fruit (*Mangifera indica* L.). *Int J Syst Evol Microbiol* 57:1576–1581
- Nielsen DS, Honholt S, Tano-Debrah K, Jespersen L (2005) Yeast populations associated with Ghanaian cocoa fermentations analysed using denaturing gradient gel electrophoresis (DGGE). *Yeast* 22:271–284
- Nielsen DS, Schillinger U, Franz CMAP, Bresciani J, Amoa-Awua W, Holzapfel WH, Jakobsen M (2007a) *Lactobacillus ghanensis* sp nov., a motile lactic acid bacterium isolated from Ghanaian cocoa fermentations. *Int J Syst Evol Microbiol* 57:1468–1472
- Nielsen DS, Snitkjaer P, van den Berg F (2008) Investigating the fermentation of cocoa by correlating Denaturing Gradient Gel Electrophoresis profiles and Near Infrared spectra. *Int J Food Microbiol* 125:133–140
- Nielsen DS, Teniola OD, Ban-Koffi L, Owusu M, Andersson TS, Holzapfel WH (2007b) The microbiology of Ghanaian cocoa fermentations analysed using culture-dependent and culture-independent methods. *Int J Food Microbiol* 114:168–186
- Pace NR, Stahl DA, Lane DJ, Olsen GJ (1986) The analysis of natural microbial-populations by ribosomal-RNA sequences. *Adv Microb Ecol* 9:1–55
- Pujol CJ, Kado CI (2000) Genetic and biochemical characterization of the pathway in *Pantoea citrea* leading to pink disease of pineapple. *J Bacteriol* 182:2230–2237
- Schwan RF (1998) Cocoa fermentations conducted with a defined microbial cocktail inoculum. *Appl Environ Microbiol* 64:1477–1483
- Schwan RF, Wheals AE (2004) The microbiology of cocoa fermentation and its role in chocolate quality. *Crit Rev Food Sci Nutr* 44:205–221
- Thompson SS, Miller KB, Lopez AS (2001) Cocoa and coffee. In: Doyle MJ, Beuchat LR, Montville TJ (eds) *Food Microbiology, fundamentals and frontiers*. ASM Press, Washington DC, pp 721–733
- Tringe SG, Hugenholtz P (2008) A renaissance for the pioneering 16 S rRNA gene. *Curr Opin Microbiol* 11:442–446
- Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, Bork P, Hugenholtz P, Rubin EM (2005) Comparative metagenomics of microbial communities. *Science* 308:554–557
- Vasilopoulos C, Ravyts F, De Maere H, De Mey E, Paelinck H, De Vuyst L, Leroy F (2008) Evaluation of the spoilage lactic acid bacteria in modified-atmosphere-packaged artisan-type cooked ham using culture-dependent and culture-independent approaches. *J Appl Microbiol* 104:1341–1353
- von Mering C, Hugenholtz P, Raes J, Tringe SG, Doerks T, Jensen LJ, Ward N, Bork P (2007) Quantitative phylogenetic assessment of microbial communities in diverse environments. *Science* 315:1126–1130
- Walter J, Hertel C, Tannock GW, Lis CM, Munro K, Hammes WP (2001) Detection of *Lactobacillus*, *Pediococcus*, *Leuconostoc*, and *Weissella* species in human feces by using group-specific PCR primers and denaturing gradient gel electrophoresis. *Appl Environ Microbiol* 67:2578–2585