Systematic and Applied Microbiology 33 (2010) 165–173

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

Seasonal and regional diversity of maple sap microbiota revealed using community PCR fingerprinting and 16S rRNA gene clone libraries

Marie Filteau [a], Luc Lagacé [b], Gisèle LaPointe [a], Denis Roy [a,*]

[a] Institut des nutraceutiques et aliments fonctionnels (INAF), Département des Sciences des aliments et de nutrition, Université Laval, QC, Canada G1K 7P4
[b] Centre de recherche, de développement et de transfert technologique acéricole inc. (ACER) St-Norbert d'Arthabaska, QC, Canada G0P 1B0

ARTICLE INFO

Article history:
Received 12 November 2009

Keywords:
Maple sap contamination
Community PCR fingerprinting method
Isolate typing
16S rRNA gene clone library
Community membership
Community structure
Operational taxonomic units
Pseudomonas
Rahnella

ABSTRACT

An arbitrary primed community PCR fingerprinting technique based on capillary electrophoresis was developed to study maple sap microbial community characteristics among 19 production sites in Québec over the tapping season. Presumptive fragment identification was made with corresponding fingerprint profiles of bacterial isolate cultures. Maple sap microbial communities were subsequently compared using a representative subset of 13 16S rRNA gene clone libraries followed by gene sequence analysis. Results from both methods indicated that all maple sap production sites and flow periods shared common microbiota members, but distinctive features also existed. Changes over the season in relative abundance of predominant populations showed evidence of a common pattern. *Pseudomonas* (64%) and *Rahnella* (8%) were the most abundantly and frequently represented genera of the 2239 sequences analyzed. *Janthinobacterium*, *Leuconostoc*, *Lactococcus*, *Weissella*, *Epilithonimonas* and *Sphingomonas* were revealed as occasional contaminants in maple sap. Maple sap microbiota showed a low level of deep diversity along with a high variation of similar 16S rRNA gene sequences within the *Pseudomonas* genus. Predominance of *Pseudomonas* is suggested as a typical feature of maple sap microbiota across geographical regions, production sites, and sap flow periods.

Introduction

Sap collected from maple trees (*Acer saccharum*) during spring has traditionally been used to produce maple syrup in the province of Québec. Sap inside the xylem of a healthy maple tree is virtually sterile but is contaminated during tapping by bacteria, yeast and molds [16,23]. In contrast to other aqueous environments, maple sap is rich in sucrose and other nutrients, making it a perfect growth medium for psychrophilic microorganisms that can exceed 10^7 CFU mL^{-1} by the end of the season [16,19]. Complex microbial communities also form biofilms inside the tubing collection system that alter sap properties as it is collected. Microbial contamination of maple sap is considered a source of quality variation for maple syrup [16,23,24]. The microbial community of maple sap could be important for flavor development [24]. However, its exact role has yet not been elucidated. The latest hypothesis on this topic states that stable members of the maple sap microbiota contribute to characteristic maple product properties [17], such as flavor and color. Dark color and off-flavors, such as burnt or sour, are often associated with an increased microbial load, but exceptions do occur [16]. These defects could be the result of unstable microbial communities that induce sucrose inversion or sap acidification. The first step to confirm this hypothesis is to demonstrate the existence of such a stable microbiota that would be similar across maple sap collection sites. The predominant bacteria previously found over two consecutive seasons in maple sap and biofilms at one experimental site have been identified as belonging to the genera *Pseudomonas* and *Rahnella* [17]. A comprehensive characterization of maple sap microbiota on a large scale is still needed in order to discriminate common versus non-stable or occasional members that may contribute to regional differences in product characteristics. Such knowledge is essential for determining the role that these community members play in maple product quality in order to develop new contamination control strategies.

So far, culture-independent techniques, such as denaturing gradient gel electrophoresis (DGGE), have advanced our understanding of maple sap tubing biofilm communities [17]. Other culture-independent techniques used to study microbial communities include 16S ribosomal RNA gene clone library analysis, a powerful approach to identify community members and measure relative abundance [26]. Additionally, new sequence analysis tools such as DOTUR and SONS are available for comparing

Abbreviations: RAPD, Random amplification of polymorphic DNA; DGGE, Denaturing gradient gel electrophoresis; T-RFLP, Terminal restriction fragment length polymorphisms; OTU, Operational taxonomic unit

* Correspondence to: Institut des nutraceutiques et des aliments fonctionnels (INAF), Université Laval, 2440, Boul. Hochelaga, QC, Canada G1V 0A6
E-mail address: denis.roy@inaf.ulaval.ca (D. Roy).

0723-2020/$ - see front matter © 2010 Elsevier GmbH. All rights reserved.
doi:10.1016/j.syapm.2010.02.003

community structure and membership [28,29]. Random amplification of polymorphic DNA (RAPD) is a PCR fingerprinting technique that can be adapted to environmental samples, with the advantage of treating large numbers of samples in parallel [33,34]. Moreover, it involves non-specific targets that require no prior knowledge of DNA sequences [33]. The downside of the method is that reference organisms must be analyzed with the same method to affiliate an organism to a phylogenetic group. Nonetheless, it can be useful to compare a large number of samples and select a representative subset before engaging in more laborious experiments such as clone libraries.

To assess maple sap stable microbiota on a genetic level and to identify stable and non-stable members, a new community PCR fingerprinting technique and 16S rRNA gene clone libraries were used as a combined approach. This study monitors the genetic variation across 19 production sites in six geographical regions at five sampling points over the sap flow period. To our knowledge, this is the first large-scale study on maple sap microbial communities. Furthermore, a representative subset of 13 samples was used to compare microbial community composition, diversity, structure and membership.

Material and methods

Maple sap samples. Maple sap osmosis concentrates (4×, corresponding to approximately 8 Brix) were obtained from 19 Québec production sites at 0%, 25%, 50%, 75% and 100% cumulative sap flow during the 2005 spring season. Sap collection systems were not sanitized during the course of the sap flow season. Concentrated maple sap samples were immediately frozen (–20 °C) after sampling until analysis.

Microbial counts and isolate selection. Bacterial counts were performed as previously described [17]. Fungal counts were obtained on acidified potato dextrose agar (Difco) incubated for 5 days at 23 °C. Morphologically different isolates were selected on plate count or tryptic soy agar incubated at 7 °C for 3 to 10 days.

DNA extraction. Maple sap osmosis concentrates (500 mL) were centrifuged for 50 min at 4 °C, 12,000*g*. Pellets were frozen at –80 °C until DNA extraction. Maple sap biomass pellets were thawed on ice and suspended in 400 μL of buffer containing 12% sucrose, 25 mM Tris–HCl, pH 8.0 and 16 mg lysozyme (Boehringer Mannheim, Laval, QC, Canada), and subsequently handled following a protocol previously used to extract DNA from maple sap [17]. Bacterial isolate DNA was extracted with the DNeasy Blood and Tissue Kit (Qiagen, Mississauga, ON, Canada), according to the protocol provided for Gram-positive bacteria. Twenty-five microliters of DNA were treated with 1 μL RNase (Roche Applied Sciences, Indianapolis, IN, USA) and the concentration of purified DNA was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA).

PCR fingerprinting. Two long degenerate primers (L: 5′-6-FAM-ATYTGYGGBTRYGCSCGSCTSCTSCC-3′ and A: 5′-6-FAM-CTSGTSAAYGGBGCSATGGCSACSAC-3′) were designed to amplify unspecific PCR fragments from maple sap samples and isolate cultures. DNA samples (100 ng each) and negative controls were deposited into a 96-well PCR plate. Each PCR contained 1.25 U Taq DNA polymerase (Biolab, Dorval, QC, Canada), 1X supplied polymerase buffer, 100 pmol primer and 40 nmol total dNTP in a 50 μL total reaction volume. PCR amplifications were performed in a Tgradient thermocycler (Biometra, Goettingen, Germany) with a program consisting of an initial 2 min denaturation step at 94 °C, then 10 cycles of 30 s at 94 °C, a 30 s annealing step with a touchdown from 60 to 50 °C (–1 °C/cycle), a 2 min elongation step at 72 °C, followed by 30 additional cycles of 30 s at 94 °C, a 30 s annealing step at 50 °C, a 2 min elongation step at 72 °C and a final 5 min elongation step at 72 °C.

PCR products were then purified with an Omega-30 K 96-well plate (Pall Corporation, Mississauga, ON, Canada) as follows: 10 μL of PCR product were loaded with 50 μL Tris buffer (10 mM pH 7.4), centrifuged for 5 min at 3000*g*, then washed with 200 μL buffer. The purified product was washed and eluted with 30 μL buffer. The loading mix for electrophoresis, consisting of 2 μL purified PCR product with 0.3 μL MapMarker 1000 (Bioventures, Murfreesboro, TN, USA) and 10 μL formamide, was denatured at 95 °C for 5 min. Each replicate PCR plate was purified and electrophoresed on a 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Sequencer output files were analyzed using GeneMapper® v3.7 software (Applied Biosystems) with the following settings: analysis range=250–1000 bp, bin size=1.3 bp, peak half width=3, polynomial degree=3, peak detection window=11, baseline=35 rfu, and using the light smoothing option. Peak height values were then imported to a spreadsheet (Microsoft® Excel®) for transformation. To account for method variation, a mean consensus profile was constructed for each sample (see supplementary material Fig. S1 for example of replicate PCR fingerprint profiles). To eliminate artifacts while accounting for potential binning errors, average peak height in at least two replicates was reported in the consensus profiles.

Consensus fingerprint data were transformed with the Hellinger transformation, which corresponds to the square root of relative peak heights [2,22], and per sample data were normalized to a mean of 0 and a standard deviation of 1. The statistical software JMP 7 (SAS Institute, Cary, NC, USA) was used to perform clustering and principal component analysis to find natural groups of samples. Two-way analysis of similarity (ANOSIM) was performed with PAST software (http://folk.uio.no/ohammer/past/) to test the null hypothesis of no difference between regions and sap flow periods. ANOSIM reports the level of dissimilarity between sample groups where R values > 0.5 and P values < 0.05 are significant.

16S rRNA gene amplification. Primers 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 788R (5′-GGACTACCAGGGTATCTAA-3′) were used to amplify the 16S rRNA gene. PCR consisted of 1.5 U Taq DNA polymerase (Biolab), 1X supplied polymerase buffer, 10 pmol of each primer, 40 nmol total dNTP and 100 ng genomic DNA in a 50 μL total reaction volume. PCR was performed in a Tgradient thermocycler (Biometra). The program comprised an initial 1 min denaturation step at 94 °C followed by 33 cycles of 30 s at 94 °C, a 30 s annealing step at 53 °C, a 1 min 30 s elongation step at 72 °C and a final 5 min elongation step at 72 °C. Bacterial isolate PCR products were purified with Exosap-it® (Biolynx, ON, CA), which degrades excess primers and nucleotides.

Clone library construction. Negative control reactions were performed for every amplification set. Three PCRs were performed on each DNA sample, as previously described. Products were then purified with a QIAquick PCR Purification Kit (Qiagen) and pooled. Purified products were cloned into the pCR2.1-TOPO vector (Invitrogen, Carlsbad, CA, USA). A total of 192 colonies were randomly selected and plasmid DNA was purified with a Montage™ Plasmid MiniprepHTS Kit (Milipore, Billerica, MA, USA).

DNA sequencing. Purified PCR products or plasmid inserts were sequenced bidirectionally with the 27F and 788R primers using an ABI PRISM® dGTP BigDyeTM Terminator v.3.0 Ready Reaction Cycle Sequencing Kit (Applied Biosystems) and a 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Sequence traces were edited manually using Sequencing Analysis 5.1.1 (Applied Biosystems), then assembled using the contig

assembly program (CAP) in BioEdit Sequence Alignment Editor version 7.0.5.2. [13].

Alignment and phylogenetic analysis. Phylogenetic affiliation of each sequence was attributed using BLASTN (http://www.ncbi.nlm.nih.gov/) and the ribosomal database project (RDP) classifier (http://rdp.cme.msu.edu/). A suite of tools available at the Greengenes website (http://greengenes.lbl.gov) was used to analyze the set of sequences. First, sequences were screened for chimeras using the Bellerophon tool and standard settings [14]. A total of 26 putative chimeric sequences were removed from the dataset (see supplementary material, Table S1). The NAST alignment tool [4] was then used to align the sequences according to a core set of alignment templates and the distance matrix tool was used to determine 16S rRNA gene sequence similarities.

Operational taxonomic unit (OTU) determination and community analysis. The resulting distance matrix was used as input to DOTUR [28]. Similarity cut-offs of 97% and 99% were used to determine $OTUs_{0.03}$ and $OTUs_{0.01}$, respectively. A neighbor-joining tree of a representative sequence of each $OTU_{0.03}$ was constructed using MEGA4 [31]. Shannon diversity indices and Chao1 richness estimates were calculated by DOTUR to assess microbial diversity and richness. The SONS package [29] was used to calculate J_{abund} and θ similarity indices and standard errors to compare community membership and structure.

Nucleotide sequence accession numbers. Clone library and pure culture 16S rRNA gene sequences are available in GenBank under accession numbers FJ933491–FJ935729 and GU068618–GU068661.

Results

Microbial counts

As expected, total aerobic counts increased over the tapping season to reach an average of 6.7 log CFU/mL (Fig. 1). All average counts were similar between geographical regions, except for the average anaerobe count that was higher at 0% sap flow in the Lower St. Lawrence region (data not shown).

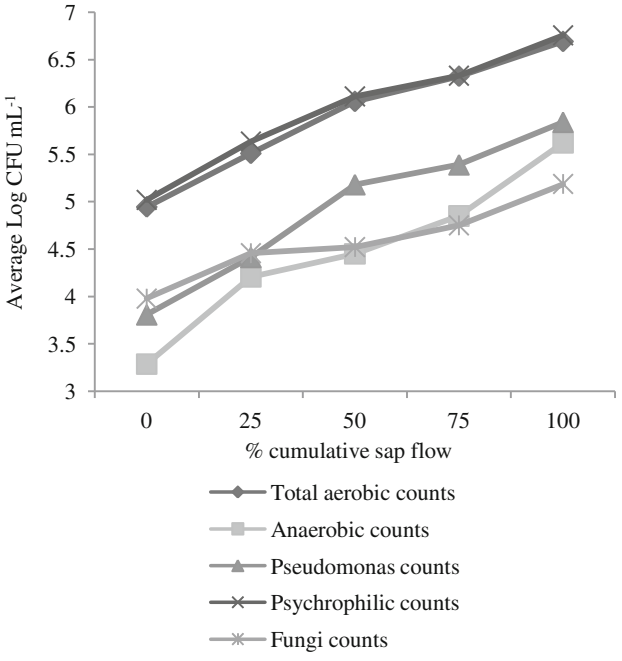

Fig. 1. Average microbial counts of 19 production sites in Québec for the 2005 season. Sampling points are evenly distributed in terms of sap flow percentage.

Community PCR fingerprints of maple sap

The microbial community genetic pool of maple sap samples from 19 production sites in five regions of Québec at five periods over the tapping season was captured by community PCR fingerprinting. In an effort to find a phylogenetic affiliation for relevant peaks representing DNA fragments, 44 morphologically different bacterial isolates from maple sap were typed using the same fingerprinting method (Fig. 2). Individual profiles yielded variable peak numbers and few peaks were shared across genera. Thus, it was presumed that most peaks present in community profiles were genus or species specific, but the number of peaks did not reflect microbial richness.

In all, 87 out of 95 maple sap DNA samples were positive for both primers, i.e., they had detectable DNA fragments from 250 to 1000 bp. A total of 636 possible peak positions were analyzed of which 41.2% were presumptively identified. No peak was shared by any of the samples. However, five peaks (0.7%) were found across all production sites and 152 peaks (22.3%), including these five, were found throughout the sap flow season.

Among the five peaks found across all production sites, peaks L10, L125 and A63 were present in both *Pseudomonas* and *Rahnella* isolate profiles. Peak L324 was present in one *Pseudomonas* isolate (75-p15_P4) and peak L300 was not found among isolate profiles. Peaks L69, L142 and A41 were found in 17, 17 and 18 production site profiles, respectively, as well as in several isolate profiles of the genus *Pseudomonas* only. Peaks associated only with *Rahnella* isolates were individually found in 14 production sites at most, but altogether they were present from all 19 production sites. Similarly, *Janthinobacterium*-, *Chryseobacterium*- and *Epilithonimonas*-associated peaks were found in 13, 9 and 5 production sites, respectively. Sample consensus profiles with some of the presumptively identified peaks are shown in Fig. 3.

To explore gene pool variation among microbial communities, PCR profiles were classified using Ward's method, which is known to be a suitable clustering method for finding natural groups [2] (Supplementary material Fig. S2). Collection site was the most significant clustering criterion observed, as profiles from the same production site were generally grouped together. However, some 0% and 25% sap flow samples clustered apart. Samples from Chaudière-Appalaches and Estrie clustered more closely than samples from other regions. To confirm that these observations were not caused by the clustering method, principal component analysis (Supplementary material, Fig. S3) was used as an alternative method. Score plots showed no separate cluster, although 0% and 25% flow samples, as well as Chaudière-Appalaches samples, were mostly positive for principal component 3. As a third method, two-way ANOSIM confirmed that profiles were not significantly influenced by geographical region ($R = -0.049$, $P = 0.827$) or sap flow period ($R = 0.0446$, $P = 0.188$).

Maple sap 16S rRNA gene clone libraries

To complement the PCR-based community fingerprints, 13 clone libraries were constructed with partial 16S rRNA gene amplification of sap samples. Four production sites in the 0%, 50% and 100% sap flow periods were selected to evaluate variation between geographical sites over the season. An additional 100% sample from a fifth production site was added to assess inter-site variation within a region. A total of 2239 sequences was retained after removal of chimeras and grouped into operational taxonomic units (OTUs) using the furthest neighbor method in DOTUR. A species-level OTU is generally defined as containing sequences that are $\geq 97\%$ identical, since it is the historically accepted level for species definition when the complete 16S rRNA

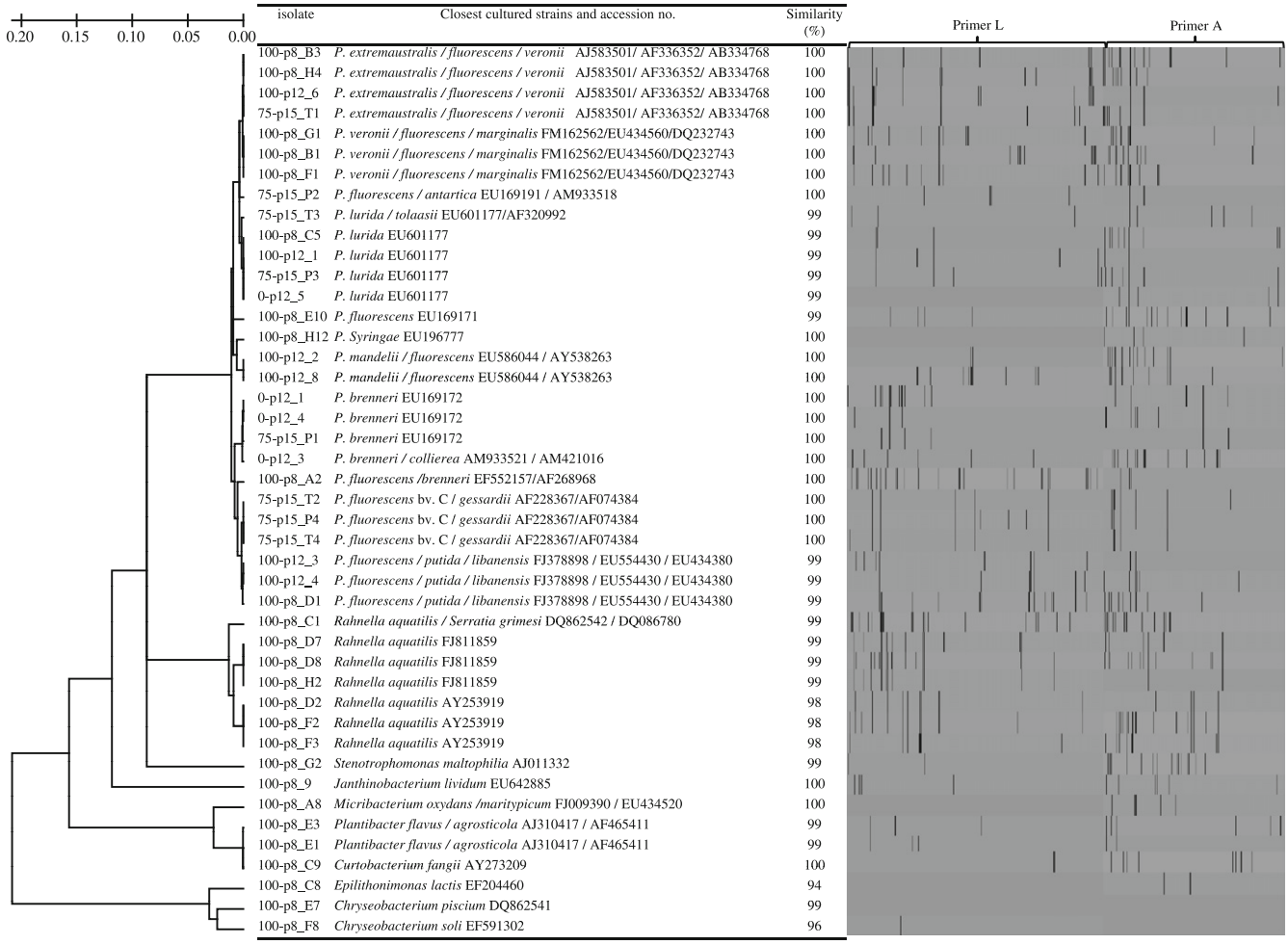

Fig. 2. Phylogenetic affiliation and typing of 44 maple sap isolates. The left axis is a neighbor-joining tree (1000 bootstrap replicates) composed of partial 16S rRNA gene sequences from each isolate. Distance units represent the number of base substitutions per site. There were a total of 714 positions in the final dataset. BLAST results for closest cultured strains are identified in the table. The image on the right represents the joint consensus PCR fingerprint profiles obtained with primers L and A for each isolate.

sequence is taken into account [10,28]. However, for environments with low deep diversity, a more stringent analysis using a 99% identity cut-off can allow a better understanding of the microbial community. Therefore, both similarity thresholds of 97% ($OTU_{0.03}$) and 99% ($OTU_{0.01}$) were applied and they yielded 85 $OTUs_{0.03}$ and 167 $OTUs_{0.01}$, respectively.

The phylogenetic architecture and relative abundance of $OTUs_{0.03}$ are shown in Fig. 4 (See also supplementary material, Fig. S4). The largest $OTU_{0.03}$, belonging to the *Gammaproteobacteria* class and the *Pseudomonas* genus, contained 1427 sequences (64% of the total). The second largest $OTU_{0.03}$ also belonged to the *Gammaproteobacteria* class, but was affiliated with the *Rahnella* genus and contained only 173 sequences (8%). Other common $OTUs_{0.03}$ belonged to the genus *Janthinobacterium*, *Chryseobacterium*, *Epilithonimonas* and *Sphingomonas*. Some less frequent but significantly abundant $OTUs_{0.03}$ were members of the lactic acid bacteria group belonging to the genera *Leuconostoc*, *Lactococcus* and *Weissella*.

A total of 38 $OTUs_{0.03}$ (74 $OTUs_{0.01}$) contained only one sequence each. Overall, 1834 (82%) sequences were 100% identical to at least one other sequence and 405 (18%) sequences were unique. Only 17 of the 85 identified $OTUs_{0.03}$ were identified at all sap flow periods, although they represented 90% of the total of 2239 sequences. Unique $OTUs_{0.03}$ were identified for each sap flow period, and they were most abundant at the beginning of the season (26) and minimal at mid-season (8).

The indices calculated by SONS were used to further describe the relationships between microbial communities while taking into account species coverage and undetected species. The J_{abund} value estimates the community membership overlap of two samples, with probability between 0 and 1 that a randomly selected OTU in one community is present in the other [29]. The J_{abund} value and its standard error reported suggested that community membership at 97% similarity was similar among all sap flow periods tested (Table 1). However, the J_{abund} index was lower at 99% similarity, showing that differences were detected especially between the community of the 0% sap flow and those of the other periods.

Community structure was further characterized by taking into account OTU abundance using the similarity index θ (Table 1). Values ranged from 0 (completely dissimilar community structure) to 1 (identical community structure) [29]. $OTUs_{0.01}$ indicated differences in community structure between all flow periods whereas $OTUs_{0.03}$ indicated a difference only for the 100% flow period. These results suggested that shifts in bacterial community structure were largely a result of changing abundance patterns of organisms, rather than the appearance of novel organisms.

Of the 85 $OTUs_{0.03}$ identified, 7 $OTUs_{0.03}$ (81% of all sequences) were present at all production sites, whereas 49 $OTUs_{0.03}$ (4% of all sequences) were specific to one production site. Microbial communities were compared pairwise using SONS (Table 2). The J_{abund} value and its standard error suggested that community

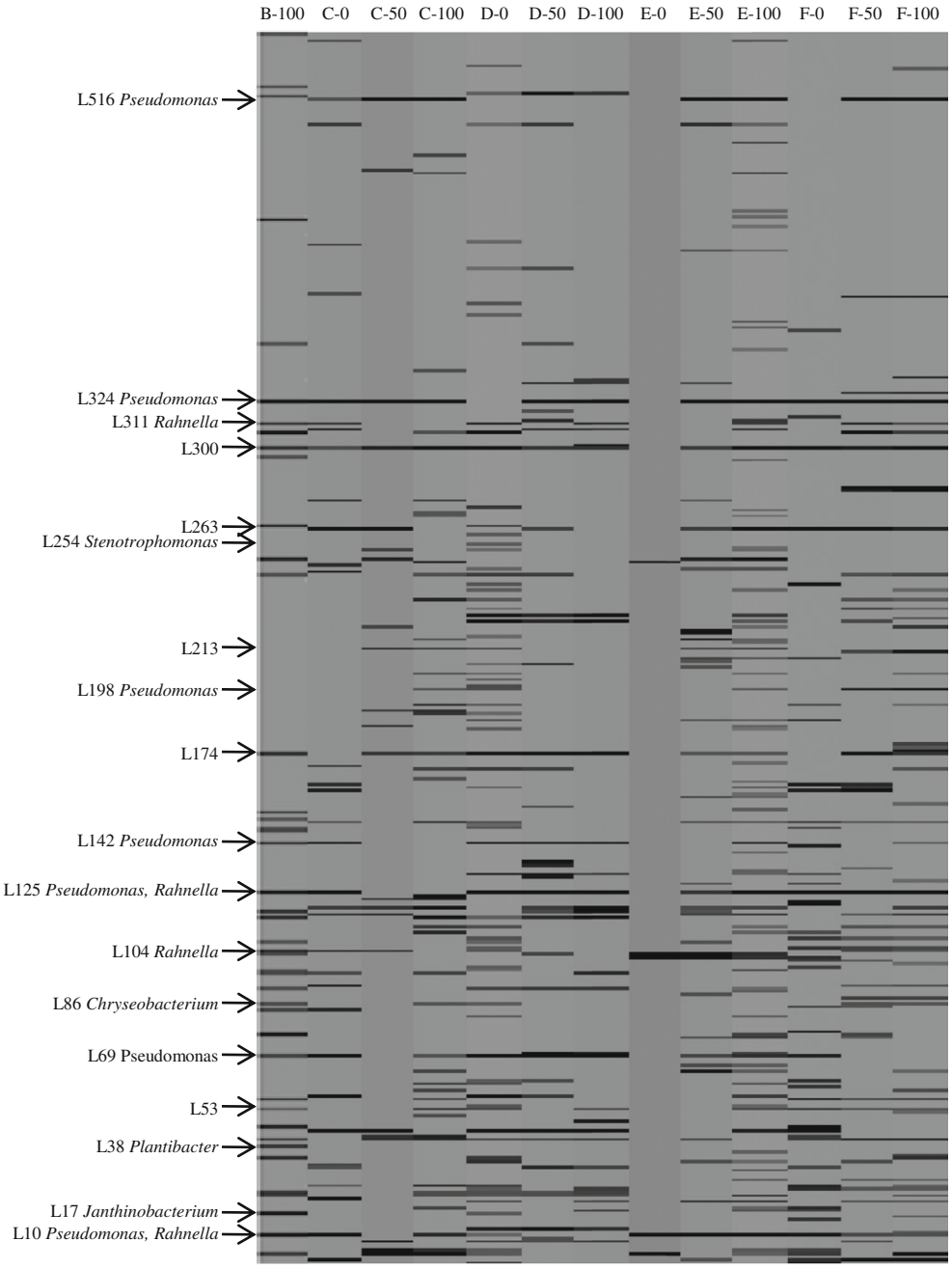

Fig. 3. Examples of PCR fingerprint consensus profiles obtained with primer L. Columns correspond to results from producers B, C, D, E and F at 0%, 50% or 100% sap flow period. Presumptive phylogenetic affiliations of peaks based on maple sap bacterial isolate profiles are identified on the left.

memberships for $OTUs_{0.03}$ were quite similar. However, greater differences were detected for $OTUs_{0.01}$. With respect to community structure, although all sites showed only small differences at 97% similarity, a drastic decrease in θ values was seen at 99% similarity (Table 2).

The Shannon diversity index was used as an estimate of the relative diversity captured by each sampling period for each production site. Results indicated no common pattern in diversity change over the season (data not shown). However, the Shannon index was highest for the 100% sap flow samples, suggesting highest diversity at the end of the flow season.

Sequences (32 out of 44) from the isolates were identical to sequences (321) from 10 different $OTUs_{(0.01)}$, which totalled 66.7% of the 2239 sequences retrieved. Furthermore, 12 isolate 16S rRNA gene sequences had no identical sequence in the clone libraries, meaning that these isolates represented a rare fraction of the microbiota.

Discussion

Maple syrup producers are already advised to sanitize their tapping equipment to minimize the microbial contamination of tap holes [1]. Nevertheless, maple sap microbial contamination remains a critical issue for maple syrup quality. Each year, an average of 11% maple syrups is rejected because of flavor defects and 20% are classified with slight off-flavors [9]. Some causes of defects are cleaning agent residue, excess of defoaming agents and microbial spoilage of sap. For instance, *Aerobacter aerogenes* causes ropiness in maple syrup, and is the only reported case of a

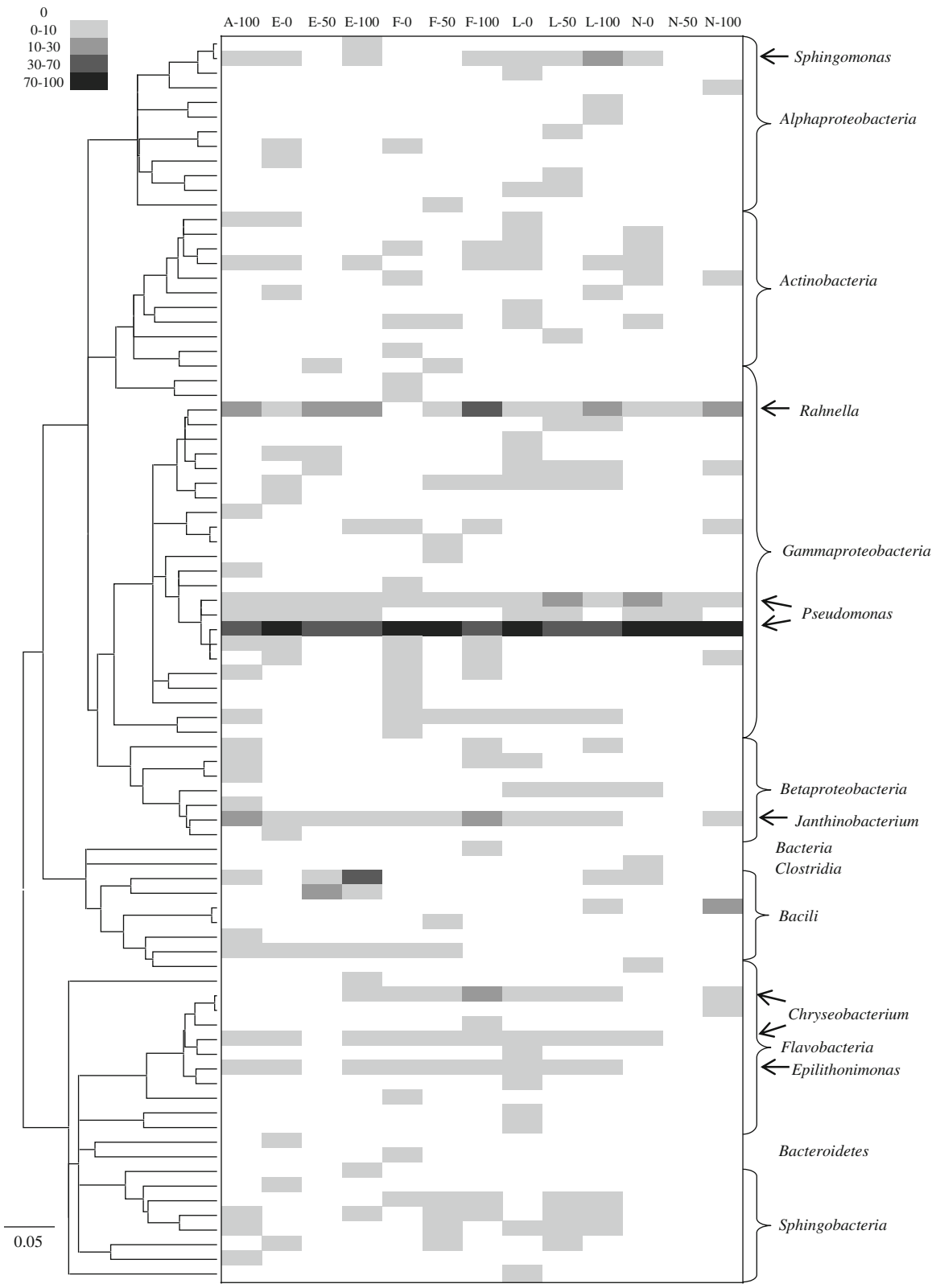

Fig. 4. Phylogenetic architecture and relative abundance of $OTUs_{0.03}$ from five maple sap producers. The left axis is a neighbor-joining tree (1000 bootstrap replicates) composed of representative sequences from each of the 85 $OTUs_{0.03}$. Distance units represent the number of base substitutions per site. There were a total of 562 positions in the final dataset. Relative abundance of $OTUs_{0.03}$ in each sample in the heat plot is indicated by a percentage greyscale value. Columns correspond to clone library results from producers B, C, D, E and F at 0%, 50% or 100% sap flow period. Phylogenetic classes and genera are identified on the right.

spoilage agent in maple sap [8]. To date, no study has extensively explored the microbial communities and their potential role in maple sap and syrup quality. This study focused on providing a comprehensive portrait of maple sap microbial communities representative of the Québec industry. Results show evidence of stable microbial community members in maple sap along with a pattern in the evolution of their relative abundance over the season.

Table 1
Pairwise comparison of maple sap microbial communities by flow period.

$OTU_{0.03}$ (%)	Chao1 shared richness (OTUs)	J_{abund}	J_{abund} SE	θ	θ SE
0–100	45	0.94	0.10	0.85	0.02
0–50	48	0.90	0.15	0.99	< 0.01
50–100	22	0.90	0.10	0.89	0.02
$OTU_{0.01}$ (%)					
0–100	68	0.83	0.08	0.33	0.03
0–50	63	0.84	0.08	0.43	0.04
50–100	176	0.97	0.07	0.83	0.03

Table 2
Variation in community membership and structure between five production sites.

$OTU_{0.03}$	Chao1 shared richness (OTUs)	J_{abund}	J_{abund} SE	θ	θ SE
E–F	59	1.00	0.06	0.82	0.03
E–C	27	0.85	0.13	0.93	0.01
E–D	21	0.84	0.13	0.94	0.01
E–B	21	0.86	0.12	0.92	0.02
F–C	50	0.94	0.20	0.85	0.03
F–D	26	0.94	0.21	0.87	0.03
F–B*	25	0.93	0.23	0.80	0.05
C–D	23	0.78	0.20	0.96	0.01
C–B	25	0.90	0.13	0.94	0.02
D–B	21	0.87	0.13	0.97	0.02
$OTU_{0.01}$					
E–F	100	0.99	0.10	0.31	0.04
E–C	92	0.81	0.10	0.48	0.05
E–D	46	0.75	0.09	0.60	0.04
E–B	26	0.69	0.10	0.41	0.04
F–C	31	0.93	0.10	0.21	0.03
F–D	37	0.90	0.12	0.34	0.04
F–B*	30	0.89	0.11	0.52	0.07
C–D	44	0.68	0.12	0.72	0.04
C–B	22	0.79	0.10	0.51	0.06
D–B	38	0.74	0.09	0.63	0.06

* Producing sites F and B belong to the same geographical region.

A new community PCR fingerprinting method using RAPD with a combination of previously reported modifications, such as the use of degenerate primers, long primers, touchdown PCR [11,27,32,34] and capillary electrophoresis combined with fluorescent primers [3], was used to survey the genetic variation in maple sap microbial communities from six geographical regions of Québec at five different flow periods evenly distributed over the season by sap volume percentage. This new method was rapid, and avoided the digestion time needed by other community fingerprinting methods, such as terminal restriction fragment length polymorphisms (T-RFLP). The profiles were reproducible, and did not give a higher number of classification errors than T-RFLP [2]. The number of classification errors for replicate PCR profiles (one out of 38 possible grouping errors, see supplementary material Fig. S1) was comparable to T-RFLP (two out of 32 possible grouping errors) [2]. Maple sap community fingerprint clustering and PCA agreed with clone library $OTU_{0.01}$ structure analysis, indicating a more important difference between 0% and other flow periods. This confirms that the method is useful to fingerprint communities composed of closely related species sharing high 16S rRNA sequence similarity. Moreover, PCR fingerprints of maple sap isolates demonstrate a high level of genetic diversity even for isolates with similar partial 16S rRNA gene sequences. Although fragments cannot be used directly to identify members of a community, they can be presumptively attributed to isolates typed with the same fingerprinting method. Alternatively, they can be cloned and used as genetic markers [34].

To complete the PCR fingerprinting results, this study analyzed the microbial communities identified in a subset of 13 representative samples with 16S rRNA gene clone libraries using two OTU cut-off values (97% and 99% similarity). The predominance of *Pseudomonas* in the Lower St-Lawrence, Mauricie, Estrie and Center of Québec regions suggests that this is not a unique feature for only one region of Québec [17,19]. Furthermore, the presence of *Rahnella,* first reported by a previous DGGE study in only the Center of Québec region [17], was also found in each of the five production sites analyzed.

For the first time, the 16S rRNA gene clone libraries revealed the presence of *Janthinobacterium*, *Leuconostoc*, *Lactococcus*, *Weissella*, *Epilithonimonas* and *Sphingomonas* in maple sap. These new findings could be attributed to the increased resolution of the method compared to DGGE or conventional cultivation methods, but also to the large-scale sampling. Distinctive microbiota members that occur sporadically may not have enough support to be used as geographical biomarkers, but these potential contaminants may be correlated with other types of flavor defects. Among others, lactic acid bacteria were occasional but relatively abundant contaminants that could be responsible for sap acidification, which is sometimes associated with flavor defects [5]. *Janthinobacterium* is also known as a psychrotrophic food spoilage agent [15] that can form biofilms [25] and some strains of *Janthinobacterium lividum* can utilise sucrose as the sole source of carbon [30]. The genus *Ralstonia*, although it was previously reported as an important contaminant of maple tree tap holes [19], was not found in maple sap sampled with clone libraries, suggesting that it is only an occasional contaminant when taking into account multiple production sites.

Maple sap microbial community membership

This study demonstrates that maple sap microbial communities share common members. All production sites were contaminated by at least 10^4 CFU mL^{-1} of *Pseudomonas* by the end of the season, and all PCR fingerprints contained presumptive *Pseudomonas* peaks. One peak common to all 19 production sites was presumptively attributed to a single *Pseudomonas* isolate whose 16S rRNA gene sequence corresponded to the most abundant $OTU_{0.03}$ (and $OTU_{0.01}$). This $OTU_{0.03}$, along with 6 others, was found in the five production sites analyzed with clone libraries. They were affiliated to *Pseudomonas*, *Rahnella*, *Janthinobacterium*, *Sphingomonas*, *Chryseobacterium* and *Frigoribacterium.* Peaks associated with *Rahnella* isolates were also found in the 19 production site PCR fingerprints. Five peaks could not be presumptively attributed to an isolate, perhaps because only a few isolates of *Janthinobacterium* and *Chryseobacterium* and no isolates of *Sphingomonas* and *Frigoribacterium* were recovered for typing.

Other evidence supporting shared membership of maple sap communities was provided by the J_{abund} values that were extremely high, even for $OTUs_{0.01}$. ANOSIM and principal component analysis of PCR fingerprint data also reflected the close membership of the communities, as no significant dissimilarity was found and no separate cluster of samples was formed. Such a strong common membership between maple sap communities, across all producing regions of Québec, is most certainly not random. It could mean that there are common selection factors determining the major members of the microbiota in maple sap. For instance, the ability to adhere and form biofilms

inside the polyethylene collection tubing system is a selective advantage of *Pseudomonas* [18], an extremely versatile microorganism that can flourish in nearly all moist environments [12]. Moreover, microorganisms adapted to grow at low temperature and able to resist the freeze-thaw cycles required for maple sap to flow are more likely to be selected in this environment. Producers also store unprocessed sap in closed tanks which could favor microorganisms capable of growth in the absence of oxygen, such as the enteric bacterium *Rahnella* that produces acid from sucrose [7].

The flow period comparison indicated that the 0% flow period was the most different in terms of membership. This may reflect a common source of initial contamination, perhaps containing microorganisms that are not able to grow in maple sap. The origin of contamination in maple sap has not yet been studied, but it has been hypothesized that bacterial contamination most likely comes from the surrounding environment [19,23]. A study using culture methods to survey bacteria in maple tap holes found some degree of analogy with forest soil microbial communities [19]. Lamarche et al. [20] showed that forest floor bacterial community composition varies across the southern boreal landscape of Québec as a function of stand type, stand age and geological parent material. Given that sugar bushes all have the same dominant stand type (*Acer saccharum*), it is possible that common contaminants of maple sap are also a common feature of the surrounding soils. However, the maple sap microbial community only has a low level of deep diversity compared to other environmental samples, such as soil, which can harbor more that 20 bacterial divisions [6]. In maple sap, members of only four divisions were recovered: Proteobacteria, Actinobacteria, Bacteroidetes and Firmicutes. Interestingly, these are the same divisions as those reported for human skin microbiota, where *Pseudomonas* and *Janthinobacterium* are also predominant members [12].

Maple sap microbial community structure

Community structure, reflected by relative abundance, evolved over the flow period. Structural changes occurred mostly in the second half of the season, as higher *Pseudomonas* than anaerobe counts occurred at mid-season whereas similar counts occurred at the end. Meanwhile, *Pseudomonas* relative abundance decreased towards 100% sap flow, as *Rahnella* and other lactic acid bacteria increased. However, even greater structural change occurred between 0% and 50% sap flow for $OTU_{0.01}$. Interestingly, the relative abundance of several *Pseudomonas* $OTUs_{0.01}$ dropped between the 0% and 50% flow period sampling points. This is probably why structural changes were not reflected by microbial counts for these sampling points. Also, the higher diversity for 100% sap flow samples is consistent with a decrease of the relative abundance of the predominant OTU toward the end of the season, as it is generally minimal at 50% and maximal at 100% sap flow volume, although this may differ between sampling sites. For instance, Lagacé et al. [17] found less DGGE band diversity at the end of the season. Our findings indicate that structural differences occurred between production sites, underlining the importance of sampling different regions.

This overall microbial community variation pattern was not expected and could be partially responsible for certain observed seasonal changes in maple syrup quality, such as darkening color. In fact, *Pseudomonas* spp. use an intracellular phosphorylase instead of an extracellular invertase enzyme, meaning that they do not hydrolyse sucrose to glucose and fructose [21], which further react in the Maillard reaction during the heating process, leading to dark pigment production. Future work could validate the hypothesis that, when present in sufficient amounts in maple sap, *Pseudomonas* may prevent darkening of maple syrup by inhibiting or competing with invertase producing microorganisms.

In conclusion, predominance of *Pseudomonas* is suggested as a general characteristic of the maple sap microbial community across geographical regions, production sites, and sap flow periods. It has been hypothesized that a stable microbiota present in the biofilms could be associated with the development of characteristic maple syrup color and flavor [17]. *Pseudomonas* would undoubtedly contribute to this stable microbiota [17,19]. *Rahnella* could also be considered a stable member of the maple sap microbiota, although it was not always a predominant member. Additionally, the variation in relative abundance observed within the microbial population might be associated with maple sap composition changes in sugars, organic acids and phenolic compounds that in turn may affect syrup quality. Further work is necessary to validate the association between particular microbiota members and flavor development. Quantitative real time PCR associated with corresponding syrup quality data could be useful to answer the ultimate question: which organisms are beneficial and which are detrimental to maple sap and syrup quality?

Acknowledgements

This work was supported by the NSERC Canada Research Chair awarded to D. Roy and an NSERC Strategic Project grant awarded to D. Roy, G. LaPointe and L. Lagacé. The authors acknowledge the support of the Centre ACER Research Fund (St-Norbert d'Arthabaska, Québec, Canada).

We are grateful to C. Charron and R. Desruisseaux for their technical help. We thank Éric Rasolofo (Laval University) for providing some pure cultures isolated from maple sap biofilm.

Appendix. Supporting materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.syapm.2010.02.003.

References

[1] G.B. Allard, L'entaillage des érables. Info fiche acéricole 140a0599, Centre Acer inc., 1999, 10 pp.
[2] Blackwood, C.B., Marsh, T., Kim, S.H., Paul, E.A., 2003. Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. Appl. Environ. Microbiol. 69, 926–932.
[3] Corley-Smith, G.E., Lim, C.J., Kalmar, G.B., Brandhorst, B.P., 1997. Efficient detection of DNA polymorphisms by fluorescent RAPD analysis. Biotechniques 22, 690–692 694, 696.
[4] DeSantis Jr., T.Z., Hugenholtz, P., Keller, K., Brodie, E.L., Larsen, N., Piceno, Y.M., Phan, R., Andersen, G.L., 2006. NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. Nucleic Acids Res. 34, W394–399.
[5] J. Dumont, Mise au point d'un outil de mesure rapide permettant d'évaluer la tendance d'une sève à donner un sirop ayant un défaut de goût majeur. 322-FIN-0299, Centre Acer Inc., 1999, 18 pp.
[6] Dunbar, J., Barns, S.M., Ticknor, L.O., Kuske, C.R., 2002. Empirical and theoretical bacterial diversity in four Arizona soils. Appl. Environ. Microbiol. 68, 3035–3045.
[7] El-Hendawy, H.H., Osman, M.E., Sorour, N.M., 2003. Characterization of two antagonistic strains of *Rahnella aquatilis* isolated from soil in Egypt. Folia Microbiol. 48, 799–804.
[8] Fabian, F.W., Buskirk, H.H., 1935. *Aerobacter aerogenes* as a cause of ropiness in maple sirup. J. Ind. Eng. Chem. 27, 349–350.
[9] FPAQ, Dossier économique 2008. Fédération des producteurs acéricoles du Québec, 2008, 17 pp.
[10] Gevers, D., Cohan, F.M., Lawrence, J.G., Spratt, B.G., Coenye, T., Feil, E.J., Stackebrandt, E., Van de Peer, Y., Vandamme, P., Thompson, F.L., Swings, J.,

2005. Opinion: re-evaluating prokaryotic species. Nat. Rev. Microbiol. 3, 733–739.
[11] Gillings, M., Holley, M., 1997. Amplification of anonymous DNA fragments using pairs of long primers generates reproducible DNA fingerprints that are sensitive to genetic variation. Electrophoresis 18, 1512–1518.
[12] Grice, E.A., Kong, H.H., Renaud, G., Young, A.C., Bouffard, G.G., Blakesley, R.W., Wolfsberg, T.G., Turner, M.L., Segre, J.A., 2008. A diversity profile of the human skin microbiota. Genome Res. 18, 1043–1050.
[13] Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.
[14] Huber, T., Faulkner, G., Hugenholtz, P., 2004. Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. Bioinformatics 20, 2317–2319.
[15] Jay, J.M., Loessner, M.J., Golden, D.A., 2005. Modern Food Microbiology seventh ed. Springer, New-York, NY, USA 790 pp.
[16] Lagacé, L., Girouard, C., Dumont, J., Fortin, J., Roy, D., 2002. Rapid prediction of maple syrup grade and sensory quality by estimation of microbial quality of maple sap using ATP bioluminescence. J. Food Sci. 67, 1851–1854.
[17] Lagacé, L., Jacques, M., Mafu, A.A., Roy, D., 2006. Compositions of maple sap microflora and collection system biofilms evaluated by scanning electron microscopy and denaturing gradient gel electrophoresis. Int. J. Food Microbiol. 109, 9–18.
[18] Lagace, L., Jacques, M., Mafu, A.A., Roy, D., 2006. Biofilm formation and biocides sensitivity of *Pseudomonas marginalis* isolated from a maple sap collection system. J. Food Prot. 69, 2411–2416.
[19] Lagacé, L., Pitre, M., Jacques, M., Roy, D., 2004. Identification of the bacterial community of maple sap by using amplified ribosomal DNA (rDNA) restriction analysis and rDNA sequencing. Appl. Environ. Microbiol. 70, 2052–2060.
[20] Lamarche, J., Bradley, R.L., Hooper, E., Shipley, B., Simao Beaunoir, A.M., Beaulieu, C., 2007. Forest floor bacterial community composition and catabolic profiles in relation to landscape features in Quebec's southern boreal forest. Microb. Ecol. 54, 10–20.
[21] Latour, X., Lemanceau, P., 1997. Métabolisme carboné et énergétique des *Pseudomonas* spp. fluorescents saprophytes à oxydase positive. Agronomie 17, 427–443.
[22] Legendre, P., Gallagher, E.D., 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129, 271–280.
[23] Morselli, M.F., Whalen, M.L., 1991. Aseptic tapping of sugar maple (*Acer saccharum*) results in light color grade syrup. Can. J. For. Res. 21, 999–1005.
[24] Naghski, J., Reed, L.L., Willits, C.O., 1957. Maple sirup X. Effect of controled fermentation of maple sap on color and flavor of maple sirup. Food Res. 22, 176–181.
[25] Pantanella, F., Berlutti, F., Passariello, C., Sarli, S., Morea, C., Schippa, S., 2007. Violacein and biofilm production in *Janthinobacterium lividum*. J. Appl. Microbiol. 102, 992–999.
[26] Röling, W.F.M., Head, I.M., 2005. Prokaryotic systematics: PCR and sequence analysis of amplified 16S rRNA genes. In: Osborn, A.M., Smith, J.C. (Eds.), Molecular Microbial Ecology. Taylor & Francis Group, New-York, NY, pp. 25–64.
[27] Sakallah, S.A., Lanning, R.W., Cooper, D.L., 1995. DNA fingerprinting of crude bacterial lysates using degenerate RAPD primers. PCR Meth. Appl. 4, 265–268.
[28] Schloss, P.D., Handelsman, J., 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. 71, 1501–1506.
[29] Schloss, P.D., Handelsman, J., 2006. Introducing SONS, a tool for operational taxonomic unit-based comparisons of microbial community memberships and structures. Appl. Environ. Microbiol. 72, 6773–6779.
[30] Shivaji, S., Ray, M.K., Seshu Kumar, G., Reddy, G.S.N., Saisree, L., Wynn-Williams, D.D., 1991. Identification of *Janthinobacterium lividum* from the soils of the islands of Scotia Ridge and from Antarctic peninsula. Polar Biol. 11, 267–271.
[31] Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
[32] Trevino-Castellano, M., Rodriguez-Novoa, S., Llovo-Taboada, J., Garcia-Zabarte, A., Garcia-Riestra, C., Regueiro-Garcia, B.J., 2003. Combined used of RAPD and touchdown PCR for epidemiological studies of *Aspergillus fumigatus*. Enferm Infecc. Microbiol. Clin. 21, 472–476.
[33] Winget, D.M., Wommack, K.E., 2008. Randomly amplified polymorphic DNA PCR as a tool for assessment of marine viral richness. Appl. Environ. Microbiol. 74, 2612–2618.
[34] Yan, X., Xu, Z., Feng, X., Liu, Y., Liu, B., Zhang, X., Zhu, C., Zhao, L., 2007. Cloning of environmental genomic fragments as physical markers for monitoring microbial populations in coking wastewater treatment system. Microb. Ecol. 53, 163–172.