



**HAL**  
open science

## Fibrochip, a functional DNA microarray to monitor cellulolytic and hemicellulolytic activities of rumen microbiota

Sophie Marre, Frédérique Durand, Ourdia Bouzid, Pascale Mosoni, Ali R. Bayat, Pierre Peyret, Evelyne Forano

### ► To cite this version:

Sophie Marre, Frédérique Durand, Ourdia Bouzid, Pascale Mosoni, Ali R. Bayat, et al.. Fibrochip, a functional DNA microarray to monitor cellulolytic and hemicellulolytic activities of rumen microbiota. *Frontiers in Microbiology*, 2018, 9, 10.3389/fmicb.2018.00215 . hal-02623146

**HAL Id: hal-02623146**

**<https://hal.inrae.fr/hal-02623146>**

Submitted on 26 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# FibroChip, a Functional DNA Microarray to Monitor Cellulolytic and Hemicellulolytic Activities of Rumen Microbiota

Sophie Comtet-Marre<sup>1,2</sup>, Frédérique Chaucheyras-Durand<sup>1,2</sup>, Ourdia Bouzid<sup>1†</sup>, Pascale Mosoni<sup>1</sup>, Ali R. Bayat<sup>3</sup>, Pierre Peyret<sup>1</sup> and Evelyne Forano<sup>1\*</sup>

<sup>1</sup> UMR 454 MEDIS, INRA, Université Clermont Auvergne, Clermont-Ferrand, France, <sup>2</sup> R&D Animal Nutrition, Lallemand, Blagnac, France, <sup>3</sup> Milk Production Solutions, Green Technology, Natural Resources Institute Finland (Luke), Helsinki, Finland

## OPEN ACCESS

### Edited by:

Stuart Edward Denman,  
Commonwealth Scientific  
and Industrial Research Organisation  
(CSIRO), Australia

### Reviewed by:

Rosalind Ann Gilbert,  
Department of Agriculture  
and Fisheries, Queensland  
Government, Australia  
Robert John Forster,  
Agriculture and Agri-Food Canada  
(AAFC), Canada

### \*Correspondence:

Evelyne Forano  
evelyne.forano@inra.fr

### † Present address:

Ourdia Bouzid,  
Division Bioénergie et Environnement,  
Centre de Développement des  
Energies Renouvelables, Alger, Algeria

### Specialty section:

This article was submitted to  
Microbial Symbioses,  
a section of the journal  
Frontiers in Microbiology

Received: 30 November 2017

Accepted: 30 January 2018

Published: 13 February 2018

### Citation:

Comtet-Marre S,  
Chaucheyras-Durand F, Bouzid O,  
Mosoni P, Bayat AR, Peyret P and  
Forano E (2018) FibroChip,  
a Functional DNA Microarray  
to Monitor Cellulolytic  
and Hemicellulolytic Activities  
of Rumen Microbiota.  
Front. Microbiol. 9:215.  
doi: 10.3389/fmicb.2018.00215

Ruminants fulfill their energy needs for growth primarily through microbial breakdown of plant biomass in the rumen. Several biotic and abiotic factors influence the efficiency of fiber degradation, which can ultimately impact animal productivity and health. To provide more insight into mechanisms involved in the modulation of fibrolytic activity, a functional DNA microarray targeting genes encoding key enzymes involved in cellulose and hemicellulose degradation by rumen microbiota was designed. Eight carbohydrate-active enzyme (CAZyme) families (GH5, GH9, GH10, GH11, GH43, GH48, CE1, and CE6) were selected which represented 392 genes from bacteria, protozoa, and fungi. The DNA microarray, designated as FibroChip, was validated using targets of increasing complexity and demonstrated sensitivity and specificity. In addition, FibroChip was evaluated for its explorative and semi-quantitative potential. Differential expression of CAZyme genes was evidenced in the rumen bacterium *Fibrobacter succinogenes* S85 grown on wheat straw or cellobiose. FibroChip was used to identify the expressed CAZyme genes from the targeted families in the rumen of a cow fed a mixed diet based on grass silage. Among expressed genes, those encoding GH43, GH5, and GH10 families were the most represented. Most of the *F. succinogenes* genes detected by the FibroChip were also detected following RNA-seq analysis of RNA transcripts obtained from the rumen fluid sample. Use of the FibroChip also indicated that transcripts of fiber degrading enzymes derived from eukaryotes (protozoa and anaerobic fungi) represented a significant proportion of the total microbial mRNA pool. FibroChip represents a reliable and high-throughput tool that enables researchers to monitor active members of fiber degradation in the rumen.

**Keywords:** rumen, cellulolysis, hemicellulolysis, glycoside hydrolases, carbohydrate esterases, functional DNA microarray, metatranscriptomics

## INTRODUCTION

Ruminants are among the most efficient herbivorous animals to convert plant biomass into edible products, principally due to a symbiotic relationship with microorganisms inhabiting the rumen. The rumen microbiota is composed of a very diverse and complex population of bacteria, archaea, protozoa, and fungi. This microbial community has a remarkable ability to degrade and ferment

complex plant polymers and thereby provides its host with protein and volatile fatty acids which represent an important nitrogen and energy source for the animal. The efficiency of the plant biomass degradation depends on the activity of the microbial enzymes produced by the rumen microorganisms. These enzymes include glycoside hydrolases (GH) active against the main plant structural polysaccharides (cellulose and hemicelluloses) and carbohydrate esterases (CE) which cleave polysaccharide substituents (Flint et al., 2012). Most of these enzymes are organized in a modular structure including one or several catalytic domain(s) bearing the hydrolytic activity and accessory domains involved in various functions such as substrate binding [carbohydrate binding module (CBM)] or cellulosome assembly (Gilbert, 2010; White et al., 2014; Artzi et al., 2017). Fibrolytic activity of rumen microorganisms can be affected by diverse biotic or abiotic factors, such as chemical composition and physical structure of the plant material, use of dietary additives, or microbiota dysbiosis leading to unfavorable ruminal conditions (Chaucheyras-Durand et al., 2012).

A few cellulolytic and hemicellulolytic bacterial species are considered as major biomass degraders in the rumen (White et al., 2014). Sequencing of their genome has revealed a high number of genes which could be involved in plant fiber degradation. Also, metagenomic approaches have given an insight into the genes collectively present in the rumen and potentially implicated in this process (Brulc et al., 2009; Hess et al., 2011; Wang et al., 2013; Güllert et al., 2016; Pitta et al., 2016). However, these approaches only identify a degradation potential of the rumen microbiota and do not identify the enzymes actually produced *in vivo* or the regulation of their expression. Recently, metatranscriptomic approaches were applied to the rumen microbiota to identify the key actors in the plant fiber deconstruction (Dai et al., 2015; Shinkai et al., 2016; Comtet-Marre et al., 2017). These studies described the main carbohydrate-active enzyme (CAZyme) families involved in this process under their experimental conditions, but also underlined the lack of gene catalogs to give a complete view of the rumen microbiota at the taxonomic and functional level. At a larger scale, knowledge gaps remain on how the activity of rumen fibrolytic microorganism is modulated by various factors at the transcriptomic level.

DNA microarrays have been extensively used as very sensitive and high-throughput tools for exploring the structure of complex microbial ecosystems (Kang et al., 2013; Tottey et al., 2013) as well as investigating functional gene sets (Dugat-Bony et al., 2012a; Tu et al., 2014). However, their use has been drastically decreased since the rise of new generation sequencing technologies applied to metagenomics and metatranscriptomics. Indeed, these non-targeted approaches enable the detection of millions of genes or transcripts present in the studied ecosystem, that have been already identified or not, and that appear of interest or not. On the contrary, microarrays generally focus on a determined number of known genes and also on their genetic variants in the case of explorative microarrays (Dugat-Bony et al., 2012b). Functional gene arrays (FGA), which are microarrays containing probes for key genes involved in microbial functional processes (He et al., 2012; Abot et al., 2016), present several advantages for specific

applications, including (i) they target functions or metabolic pathways of interest; (ii) they allow to monitor in parallel and at low cost a high number of samples such as those generated by studies on large number of animals or kinetics, and (iii) up to now, analysis of results has been also easier and quicker than for metatranscriptomic data.

In the present work, we developed a FGA targeting fiber-degrading activities in the rumen ecosystem, designated as FibroChip. The microarray was designed to target genes belonging to eight CAZyme families known to contain very efficient cellulases and hemicellulases (i.e., GH5, GH9, GH10, GH11, GH43, GH48, CE1, and CE6) which are present in the genome of the major rumen fibrolytic microorganisms including bacteria, fungi, and protozoa. The DNA microarray allows the detection of 392 genes and is composed of 4249 probes. We present here the validation of this FGA using targets of increasing complexity and its use to analyze the transcriptome of one of the major rumen cellulolytic bacterium *F. succinogenes*. We also applied this new tool to characterize the expression of the CAZyme gene repertoire of the rumen microbiota of a dairy cow and compared the results with a RNA-sequencing (RNA-seq) approach. The FibroChip demonstrated sensitivity and specificity as well as explorative potential.

## MATERIALS AND METHODS

This work was undertaken according to the relevant biosecurity and INRA safety procedures.

### Sequence Retrieval, Probe Design, and Microarray Features

For all rumen microorganisms known to degrade cellulose, hemicelluloses, and oligosaccharides, nucleic sequences encoding the CAZyme families GH5, GH9, GH10, GH11, GH43, GH48, CE1, and CE6 were extracted from the CAZy database (RRID:SCR\_012909<sup>1</sup>; Lombard et al., 2014) when available (Supplementary Table S1). Sequences from environmental anaerobic microorganisms not yet known to be part of the rumen microbiota were also included. Probes were designed specifically in the coding sequence of the catalytic domain of each gene. Consensus sequences were determined for sequences with more than 96% of identity that should not be distinguished with the microarray. In this particular case, probes were designed in conserved regions and targeted a set of genes instead of a unique gene. Sense oligonucleotide probes were designed using the HiSpOD software (RRID:SCR\_014403<sup>2</sup>; Dugat-Bony et al., 2011). Parameters used were: probe length = 25-mers, T<sub>m</sub> min = 46°C, T<sub>m</sub> max = 64°C, and Complexity = 10. Specificity of deduced probes was tested against the EnvExBase database implemented in the HiSpOD software, and probes with more than 85% of identity and more than 15 consecutive bases (similarity stretch) with a non-targeted sequence were discarded by following the criteria of Kane et al. (2000) and He et al.

<sup>1</sup><http://www.cazy.org>

<sup>2</sup><http://g2im.u-clermont1.fr/hispod/index.php>



## RNA Hybridization Experiment

### Strain and Culture Conditions

The strain *F. succinogenes* S85 was grown with either 0.2% of Clb (Sigma–Aldrich, Saint-Louis, MO, United States) under static conditions or 0.5% of WS under shaking conditions. The strain was cultured in triplicate on the same medium. Cultures were harvested at the end of the exponential growth after 6 and 50 h, respectively (Béra-Maillet et al., 2000; Nouaille et al., 2009).

### Animal Experiment, Diet, and Sampling of Rumen Contents

All experimental procedures were approved by the National Ethics Committee (Hämeenlinna, Finland) in accordance with the guidelines established by the European Community Council Directive 86/609/EEC (Louhimies, 2012). The analyzed rumen sample was obtained from a rumen cannulated lactating dairy cow fed a total mixed ration based on grass silage (forage:concentrate ratio 50:50, on a dry matter basis). Details of the diet fed have been reported previously (Bayat et al., 2015). In brief, the total mixed ration contained (gram per kilogram dry matter) neutral detergent fiber (401), acid detergent fiber (228), and starch (120). Samples were taken before the morning feeding from five different sites in the rumen, composited, and mixed thoroughly to obtain a representative sample of rumen contents. For molecular analysis, rumen content was subsampled (50 g) and mixed with 100 ml of RNeasy (Thermo Fisher Scientific, Waltham, MA, United States) to prevent RNA degradation. The mixture was maintained overnight at +4°C and then stored at –80°C until nucleic acid extraction.

### RNA Isolation and mRNA Enrichment

Total RNA from *F. succinogenes* S85 was extracted in duplicate from 10 ml of culture for both substrate conditions (Clb and WS). RNA isolation was performed using Trizol reagent (Invitrogen, Thermo Fisher Scientific, Waltham, MA, United States) with slight modifications of the manufacturer's instructions. Cultures were centrifuged for 15 min at 12,000 × *g* at +4°C. Then, 10 ml of Trizol was added to the pellets and the mix was vigorously vortexed. Cells were disrupted with ~100 mg of 0.1 mm Zirconia Beads (BioSpec Products, Bartlesville, OK, United States) in a 2 ml-microtube using the FastPrep Instrument (MP Biomedicals, Irvine, CA, United States), set to 6.5 m/s, during four cycles of 30 s interspaced with 5 min of incubation on ice. Following steps were those recommended by the Trizol protocol.

Total RNA from rumen content was extracted in duplicate from 400 mg of centrifuged rumen content using Trizol as previously described (Comtet-Marre et al., 2017). The integrity of RNA was assessed with the Agilent 2100 Bioanalyzer using RNA Nano Chip (Agilent Technologies, Santa Clara, CA, United States) and quantity was assessed with the Nanodrop 1000 Spectrophotometer (Thermo Fisher Scientific, Illkirch-Graffenstaden, France). Duplicates of total RNA were pooled and the resulting RNA solutions were stored at –80°C until further processing. Total RNA from *F. succinogenes* S85 and rumen contents were submitted to mRNA enrichment using an in-house-procedure based on rRNA subtractive hybridization principle as previously described (Comtet-Marre et al., 2017).

### Target Labeling and Hybridization

Fractions of mRNA-enriched RNA (200 ng) were amplified with the MessageAmp II-Bacteria RNA Amplification Kit (Ambion, Thermo Fisher Scientific, Waltham, MA, United States) according to the modified protocol of Dugat-Bony et al. (2012a) to incorporate aminoallyl UTP during a 14 h-*in vitro* transcription (Ambion, Thermo Fisher Scientific, Waltham, MA, United States). For each sample, 10 µg of amplified RNA (aRNA) was labeled with Cy3 or Cy5 cyanines (Pack Amersham CyDye Post-Labeling Reactive Dye, GE Healthcare, Little Chalfont, United Kingdom), following the manufacturer's recommendations. Labeled aRNA was finally purified with NucleoSpin RNA Clean-Up Kit (Macherey-Nagel, Düren, Germany). Labeling density was acceptable when values were between 30 and 60 (one dye per every 30–60 unlabeled nucleotides). RNA hybridization (1 µg for the strain *F. succinogenes* S85 and 5 µg for rumen samples) was performed at 65°C during 17 h following the Agilent protocol “Two-Color Microarray-Based Gene Expression” with few modifications for the use of a supplementary blocking agent (Kreablock, Kreatech Diagnostics, Amsterdam, Netherlands). Briefly, RNA fragmentation was performed at 60°C during 30 min in a mix composed of 12.5 µl of Kreablock buffer (Kreatech Diagnostics, Amsterdam, Netherlands), 6.5 µl of aRNA, 5 µl of 10× Gene Expression Blocking Agent (Agilent Technologies, Santa Clara, CA, United States), and 1 µl of 25× fragmentation buffer (Agilent Technologies, Santa Clara, CA, United States). Next steps of hybridization and slide washing were performed according to Agilent protocol.

### Imaging, Data Preprocessing, and Analysis

Immediately after washing steps, microarray slides were scanned using an Agilent G2505B microarray scanner (Agilent Technologies, Santa Clara, CA, United States) with a resolution of 5 µm and the XDR mode (Extended Dynamic Range) set up at 10 and 100% PhotoMultiplier Tubes (PMT) gain. Fluorescence intensities of oligonucleotide spots were extracted from the scanned images using the Feature Extraction software (V11.0, Agilent Technologies, Santa Clara, CA, United States). Data are available under NCBI GEO (accession number GSE107550). Data were processed with dedicated scripts based on C++ and Delphi languages. For each probe, median intensity value of the three replicates was conserved and used as the probe signal value. The signal-to-noise ratio (SNR), similar to the detection threshold response and corresponding to the probe signal value divided by the local background intensity value, was calculated for each probe. Results for 25- and 54-mer probes were treated independently and for each probe type, a gene was considered as expressed when 65% of probes were positive. The SNR value of a detected transcript was then calculated by the mean of the SNR of all 25- or 54-mer probes targeting this gene. For PCR product hybridizations, gDNA hybridizations, and *F. succinogenes* RNA hybridizations, the SNR thresholds were set to 3 for 25-mer probes and 6 for 54-mer probes. For rumen RNA hybridizations, the SNR threshold was set at 2 such that a positive response was

considered when the signal intensity of a probe was two times superior to the signal of the background and negative control probes (He and Zhou, 2008).

## Quantification of the Expression of Genes from *Fibrobacter succinogenes* S85

For *in vitro* cultures, *F. succinogenes* S85 total RNA (100 ng) was reverse-transcribed into cDNA using Superscript II reverse transcriptase (Invitrogen, Thermo Fisher Scientific, Waltham, MA, United States) according to the manufacturer recommendations. Three replicates were used for each substrate condition. For each RNA sample, a negative control (no addition of reverse transcriptase) was performed to assess DNA contamination.

For qPCR quantification of *F. succinogenes* S85 gene expression in the cow rumen content, mRNA enriched RNA (180 ng in duplicate) from total rumen content was reverse-transcribed in cDNA with the polymerase M-MLV RT RNase H Minus DNA (Euromedex, Souffelweyersheim, France) following manufacturer recommendations.

The different sets of cDNA were then used for qPCR quantification of the expression of *F. succinogenes* S85 genes in the *in vitro* cultures and in the rumen content. Primers used for the qPCR quantifications of *endB* (GH9 ACX73672), *endA* (GH9 ACX75948), and *xynD* (GH10 ACX75384) were those described in Béra-Maillet et al. (2009). Primers used for *axe6B* (CE6 ACX75360) amplification were Forward(5′-3′)-TGTGAGCATCCGCGCCTTCG and Reverse(5′-3′)-TTGCGCCAGTTGGCATCGGT, and primers used for *celE* (GH9 ACX 75451) cDNA quantification were Forward(5′-3′)-CAGGCTGCAGATGGTGGCGT and Reverse(5′-3′)-TTGGCCGCCTCAAGGCACTG. qPCR reactions were achieved mostly as described in Béra-Maillet et al. (2009). They were carried out in duplicate in 96-well plates using a total volume of 20  $\mu$ l with 10  $\mu$ l of 2 $\times$  Takara SYBR Premix Ex Taq (Lonza, Bâle, Switzerland), 1 ng of cDNA, and 0.5  $\mu$ M of each primer. For each cDNA quantification, standard curves were used in order to determine the absolute number of copies as described in Béra-Maillet et al. (2009). The gene copy numbers were calculated from three biological replicates and their two technical replicates.

## Isolation of a GH11 Transcript of *Piromyces communis* from Total Rumen Content

Messenger RNA enriched RNA (180 ng) from total rumen content was reverse-transcribed in cDNA as described above and using a mix of primers deduced from probes targeted the gene encoding a GH11 hemicellulase from *Piromyces communis* (protein accession number AAG18439). Primers were eliminated by column-based purification using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). One-tenth of the reverse transcription reaction was used to perform a primary PCR. Then a secondary nested PCR was achieved with 1.5  $\mu$ l of the first PCR reaction used as PCR template.

PCR was performed using the polymerase GoTaq G2 DNA (Promega, Madison, WI, United States) following supplier recommendations. Amplification steps were: initial denaturation (94°C, 2 min), 35 cycles comprising denaturation (95°C, 30 s), hybridization (50°C, 30 s), elongation (72°C, 60 s), and final elongation (72°C, 5 min). Primers used are reported in Supplementary Table S5. Several secondary PCR reactions were realized in order to obtain enough PCR products for sequencing. PCR products were gel-purified using the QIAquick Gel extraction Kit (Qiagen, Hilden, Germany) according to the supplier instructions. Finally purified PCR products were concentrated with the MinElute PCR Purification Kit (Qiagen, Hilden, Germany) and sequenced by the MWG Company (Mulhouse, France) using primers of the secondary PCR reaction.

## RNA-Seq Data

Microarray data from rumen RNA hybridization were compared with RNA-seq data (available on Sequence Read Archive, SRP070140) obtained in a previous work (Comtet-Marre et al., 2017). Briefly, translated non-rRNA were submitted to a local version of dbCAN (RRID:SCR\_013208) (Yin et al., 2012) to detect and annotate sequences for CAZymes (*e*-value < 1e−05). dbCAN positive sequences were subjected to sequence similarity searches using BLASTX (BLAST+ 2.5.0; RRID:SCR\_001653) (Camacho et al., 2008) against the NCBI NR database protein (RRID:SCR\_003257) (*e*-value < 1e−05). BLAST results (25 best hits for each read) were used for taxonomic mining of putative CAZyme transcripts using the MEGAN software version 6 (RRID:SCR\_011942) (Huson et al., 2011). The taxonomic level was determined using the lowest common ancestor-based algorithm (LCA) implemented in the software. LCA parameters were “Top percent” set to 10, “Min support percent” 1e−05, and “Min score” set to 90 (Comtet-Marre et al., 2017).

## RESULTS AND DISCUSSION

### Overall Description of FibroChip Features

The FibroChip was designed with the objective to propose a high-throughput tool providing a rapid picture of the capacity of rumen microorganisms to degrade cellulose and hemicelluloses based on a targeted metatranscriptomic approach. We chose to focus on a few number of genes by targeting main ruminal fibrolytic microorganisms and selected CAZyme families that may have a pivotal role in cellulose and hemicellulose degradation. The microarray targets the coding sequence of catalytic domains from eight CAZyme families involved in cellulose and hemicellulose degradation (i.e., GH GH5, GH9, GH10, GH11, GH43, and GH48, and CE CE1 and CE6). Taken together, these families present complementary activities needed for the complete degradation of cellulose and hemicelluloses (Table 1).

### Selection of CAZyme Families

Carbohydrate-active enzyme families were principally selected based on known mechanisms of enzymes and genomic data from

fibrolytic microorganisms. Families GH5, GH9, and GH43 are present in large numbers in the genome of fibrolytic bacteria as well as families GH9 and GH48 in fungal genomes (Berg-Miller et al., 2009; Purushe et al., 2010; Suen et al., 2011; Pajon et al., 2013, Unpublished; Youssef et al., 2013). Additionally, cellulases from GH9 and GH48 are key enzymes of cellulosomes of *R. albus* and fungi (Devillard et al., 2004; Arai et al., 2006). These enzymes have potentially a significant role in the degradation of crystalline cellulose (Davison and Blaxter, 2005). Furthermore, endoglucanases from GH5 and GH9 have shown synergistic activity with GH48 cellobiohydrolases (Irwin et al., 2000). Regarding hemicellulose degradation, families GH5 (members of the subfamily 21), GH10, GH11, and GH43 are key enzymes of this process. Families GH10 and GH11 are almost exclusively constituted of endoxylanases that hydrolyse  $\beta(1,4)$  linkages of the xylan backbone. They have also been identified in xylanase utilization systems (XUS) of xylanolytic Bacteroidetes (Dodd et al., 2010). Endoxylanases from family GH5 are present in XUS systems suggesting a substantial role in xylan degradation (Aspeborg et al., 2012). Family GH43, retrieved in large number of copies in genomes of *Bacteroides* spp., is polyspecific. Although some xylanases are found in this family, it contains mainly xylosidase and debranching enzymes like arabinofuranosidases<sup>4</sup> (Lombard et al., 2014). Finally, two families of CE (CE1 and CE6) were also selected. The family CE1, for which a high number of copies is retrieved from genomes of bacteria and fungi, is mainly composed of acetyl-, feruloyl-, and cinnamoyl-esterases, but other enzymes such as carboxylesterases that are not involved in fiber degradation, are also found in this family (Lombard et al., 2014)<sup>4</sup>. Feruloyl esterases, owing to their ability to dissociate lignin from hemicelluloses, have a pivotal role in xylan degradation by making polysaccharides

accessible to other enzymes (Girio et al., 2010). The family CE6 is exclusively constituted of acetyl esterases that release acetyl groups from the xylan backbone. Their activity is also necessary to allow xylanolytic enzymes to access to their substrate (Faulds et al., 2006; Gottschalk et al., 2010; Zhang et al., 2011).

These eight families have been found among the most transcribed in the published rumen metatranscriptomes (Dai et al., 2015; Li and Guan, 2016; Shinkai et al., 2016; Comtet-Marre et al., 2017). Moreover, transcriptomic data obtained from *Prevotella bryantii* B<sub>14</sub> (Dodd et al., 2010), *Ruminococcus flavefaciens* FD-1 (Berg-Miller et al., 2009), *B. xylanisolvens* XB1A (Despres et al., 2016), and *Orpinomyces* sp. C1A (Youssef et al., 2013) have identified these families as up-regulated in pure culture with polysaccharide compared to simple sugar substrates. The targeted GH families were also among the 20 most expressed GH families detected by the CAZyChip, a DNA microarray targeting more than 55,000 bacterial GH genes, after 5 days of fermentation of WS by a simplified rumen microbial community (Abot et al., 2016). Finally, four out of the eight targeted CAZyme families (GH5, 43, 48, and CE1) on a total of 16 GH families identified with a metaproteomic approach were found in abundance in the rumen of cows fed corn silage, grass silage, or hay (Deusch et al., 2017). Hence, induction of genes encoding these families and their high expression in rumen metatranscriptomes suggests their strong contribution to the cellulose and hemicellulose degradation process.

### Genes Targeted and Probe Design

In total, 392 nucleic sequences encoding 363 GH and 29 CE were kept for the microarray design. These sequences originated from 41 bacterial species, 4 protozoal species, and 10 fungal species (Supplementary Table S1). The main fibrolytic bacterial genera targeted were *Fibrobacter*, *Ruminococcus*, *Prevotella*, *Bacteroides*,

<sup>4</sup><http://www.cazy.org>

**TABLE 1 |** Substrate and known activities of the CAZyme families targeted with the FibroChip.

Substrate	CAZy family	Known activities
Cellulose	GH5 (except subfamily 21)	Endo- $\beta$ -1,4-glucanase/cellulase (EC 3.2.1.4); $\beta$ -glucosidase (EC 3.2.1.21); $\beta$ -mannosidase (EC 3.2.1.25); glucan $\beta$ -1,3-glucosidase (EC 3.2.1.58); licheninase (EC 3.2.1.73); exo- $\beta$ -1,4-glucanase/cellobiohydrolase (EC 3.2.1.74); glucan endo-1,6- $\beta$ -glucosidase (EC 3.2.1.75); mannan endo- $\beta$ -1,4-mannosidase (EC 3.2.1.78); cellulose $\beta$ -1,4-cellobiosidase (EC 3.2.1.91); chitosanase (EC 3.2.1.132); endo- $\beta$ -1,6-galactanase (EC 3.2.1.164); $\beta$ -1,3-mannanase (EC 3.2.1.-); mannan transglycosylase (EC 2.4.1.-)
	GH9	Endoglucanase (EC 3.2.1.4); endo- $\beta$ -1,3(4)-glucanase/lichenase-laminarinase (EC 3.2.1.6); $\beta$ -glucosidase (EC 3.2.1.21); lichenase/endo- $\beta$ -1,3-1,4-glucanase (EC 3.2.1.73); exo- $\beta$ -1,4-glucanase/cellobiohydrolase (EC 3.2.1.74); cellobiohydrolase (EC 3.2.1.91); xyloglucan-specific endo- $\beta$ -1,4-glucanase/endo-xyloglucanase (EC 3.2.1.151); exo- $\beta$ -glucosaminidase (EC 3.2.1.165)
	GH48	Reducing end-acting cellobiohydrolase (EC 3.2.1.176); endo- $\beta$ -1,4-glucanase (EC 3.2.1.4); chitinase (EC 3.2.1.14)
Hemicellulose	GH5 (subfamily 21)	Endo- $\beta$ -1,4-xylanase (EC 3.2.1.8); xyloglucan-specific endo- $\beta$ -1,4-glucanase (EC 3.2.1.151); arabinoxylan-specific endo- $\beta$ -1,4-xylanase (EC 3.2.1.-)
	GH10	Endo-1,4- $\beta$ -xylanase (EC 3.2.1.8); endo-1,3- $\beta$ -xylanase (EC 3.2.1.32)
	GH11	Endo- $\beta$ -1,4-xylanase (EC 3.2.1.8); endo- $\beta$ -1,3-xylanase (EC 3.2.1.32)
	GH43	$\beta$ -Xylosidase (EC 3.2.1.37); $\alpha$ -L-arabinofuranosidase (EC 3.2.1.55); arabinanase (EC 3.2.1.99); xylanase (EC 3.2.1.8); galactan 1,3- $\beta$ -galactosidase (EC 3.2.1.145); $\alpha$ -1,2-L-arabinofuranosidase (EC 3.2.1.-); exo- $\alpha$ -1,5-L-arabinofuranosidase (EC 3.2.1.-); [inverting] exo- $\alpha$ -1,5-L-arabinanase (EC 3.2.1.-); $\beta$ -1,3-xylosidase (EC 3.2.1.-)
	CE1	Acetyl xylan esterase (EC 3.1.1.72); cinnamoyl esterase (EC 3.1.1.-); feruloyl esterase (EC 3.1.1.73)
	CE6	Acetyl xylan esterase (EC 3.1.1.72)

*Butyrivibrio*, and *Cellulosilyticum*. Other gut bacterial genera known to degrade oligosaccharides and soluble polysaccharides, i.e., *Bifidobacterium*, were also targeted as well as some anaerobic fibrolytic bacteria from the genera *Clostridium* and *Cellulomonas* since they are likely present in the digestive ecosystems even if they are not essential to plant cell wall degradation (Li et al., 2013; Ziemer, 2014). Regarding commensal rumen eukaryotes, the fungal genera *Piromyces*, *Neocallimastix*, and *Orpinomyces*, and the protozoal genera *Epidinium*, *Eudiplodinium*, and *Polyplastron* were targeted (Supplementary Table S1). The microarray was composed of 1631 25-mer probes, ensuring the specific recognition of either unique sequences or groups of very closely related genes. Indeed, for sequences with more than 96% of identity, mostly originated from the same microbial species, consensus sequences were determined leading to 54 sets of 2–19 sequences for probe design. In summary, an average of five 25-mer probes per target gene was obtained (up to 14 probes per gene when a unique gene was targeted and up to 10 probes in case of a group of genes). Additionally, the GoArrays strategy, consisting in associating two probes targeting the same gene (Rimour et al., 2005), led to the determination of 2618 composite probes of 54-mers (GoArrays probes) with an average of 9 probes per target gene (up to 22 probes per unique gene and up to 37 probes in case of a group of genes). Finally, the FibroChip was constituted by a total of 4249 probes (Supplementary Table S2). All the probes targeted a gene region coding for the catalytic domain of the enzyme, which carries the activity (Supplementary Table S3). We did not target regions coding for accessory domains (CBM and dockerins), as these modules are sometimes associated with other proteins than CAZymes (Peer et al., 2009).

## Evaluation of the FibroChip Sensitivity and Specificity

The sensitivity of the probes defined by their ability to accurately recognize genes involved in the fiber degradation was assessed using seven genes amplified by PCR. These genes were selected in order to represent the diversity of targeted CAZyme families and microorganisms (Table 2), and were cloned to ensure that a unique sequence was hybridized. Around 50 ng of each gene, representing  $\sim 10^{10}$  copies, was hybridized. Additionally, 1  $\mu$ g of

gDNA ( $\sim 10^{11}$  copies of genome) from three strains of fibrolytic bacteria (two rumen strains *F. succinogenes* S85 and *R. albus* 20, and one strain from the human gut *B. xylanisolvens* XB1A) was used (Table 3). The hybridization results show that all the probes gave a strong positive signal (Tables 2, 3). GoArrays probes presented a better sensitivity than short probes, with a SNR of 5- to 17-fold higher for PCR products (not shown for gDNA), confirming the high sensitivity.

The genomes of rumen fungi harbor multicopies of GH48 genes of which nucleic sequences could reach 99% of identity (Youssef et al., 2013). This observation was confirmed for the strain *Piromyces* sp. M4 for which several GH48 genes were found (Forano et al., Unpublished data). Thereby, detection of sequence variants of genes used for probe design can be particularly useful (Dugat-Bony et al., 2012b). The capacity of FibroChip probes to detect these sequences was evaluated by hybridizing a GH48 clone from *Piromyces* sp. M4 (GenBank accession number MF773967), containing a sequence with 95% identity to the ones from other strains of *Piromyces* sp. used for probe design. The gene hybridized with 25-mer probes with perfect match, but no hybridization occurred with probes containing mismatches (Supplementary Table S6) confirming the high specificity expected for the short probes. GoArrays probes constituted by two perfectly matching probes and by one perfectly matching probe combined with one probe containing mismatches gave also a signal. The strongest signal was measured with the two perfectly matching probes (Supplementary Table S6). Besides, the GoArrays probes constituted by the two probes with mismatches did not produce any hybridization signal (not shown). This result demonstrates the exploratory potential of the FibroChip, which could be particularly useful to detect genes that are present in high number of variants such as GH48 and GH9 genes in rumen fungi (Youssef et al., 2013).

## Application of the FibroChip to Targeted Analysis of the Transcriptome of *Fibrobacter succinogenes* S85

Transcriptomic and proteomic data have shown that gut fibrolytic microorganisms are able to adapt the expression and

**TABLE 2** | Microarray results obtained by hybridization of seven cloned genes encoding different CAZymes from diverse microorganisms targeted by the FibroChip.

Target genes	25-mer probes			GoArrays probes			Other detected genes
	Positive probes <sup>a</sup> /total probes	Gene SNR <sup>b</sup>	Detection	Positive probes <sup>a</sup> /total probes	Gene SNR <sup>b</sup>	Detection	
GH5 <i>F. succinogenes</i> S85	5/5	312	+	9/9	3225	+	0
GH9 <i>R. albus</i> 20	6/6	474	+	8/8	2403	+	0
GH10 <i>B. xylanisolvens</i> XB1A	4/4	157	+	4/4	2166	+	0
GH11 <i>R. albus</i> 7	11/11	166	+	11/11	2860	+	0
GH48 <i>R. albus</i> 20	6/6	175	+	6/6	1939	+	0
GH10 <i>P. multivesiculatum</i>	6/6	152	+	8/8	2573	+	0
GH9 <i>Piromyces</i> sp. M4	6/6	272	+	14/14	2433	+	0

<sup>a</sup>Number of probes that gave a positive signal (positive probes). <sup>b</sup>SNR: signal of hybridization defined as a signal-to-noise ratio.



production of CAZymes according to the substrate available (Béra-Maillet et al., 2000; Berg-Miller et al., 2009; Despres et al., 2016). To test the ability of the FibroChip to detect the differential expression of genes, the microarray was used to analyze the transcriptome of the rumen bacterium *F. succinogenes* S85 cultivated in two media differing by their carbon source, i.e., the simple sugar Clb or the complex substrate WS. Indeed, *F. succinogenes* S85 is known to degrade efficiently cellulose and xylans even if it cannot use xylose (Matulova et al., 2008). The substrate WS was essentially composed of cellulose, hemicelluloses (xylans), and lignin, and comprised 81% carbohydrates, with a monomer composition of 58% glucose and 34% xylose. WS was assumed to induce the expression of the *F. succinogenes* cellulase and hemicellulase genes targeted by the microarray. On both substrates, all the 31 genes targeted by the microarray were detected as expressed (GoArrays results, SNR threshold of 6 and at least 65% of probes positive). The SNR threshold value of 6 ensured a very specific response of the FibroChip.

During growth in WS relative to Clb, nine genes were considered to be up-regulated based on the  $\log_2$  ratio  $> 2$  (Figure 1). They belonged to all the targeted CAZyme families, except GH10. The most induced genes ( $\log_2$  ratio  $> 2.5$ ) encoded a GH11 endo- $\beta$ -1,4-xylanase (protein accession number ACX76028), the GH5 cellodextrinase Ceda (ACX75179), a GH5 not yet characterized (ACX74338), and a GH9 endo- $\beta$ -1,4-glucanase (EGB, Cel9G; ACX73672). The genes encoding EGB and Ceda were also previously found as the most expressed genes by *F. succinogenes* S85 grown on cellulose *in vitro*, and were also among the most *F. succinogenes* expressed genes in the rumen of sheep (Béra-Maillet et al., 2000, 2009). The transcripts of a GH9 endoglucanase (EGE or CelE; ACX75451) were the only ones with a negative  $\log_2$  ratio value. The expression of *celE* was then assessed by RT-qPCR. Expression of this gene was quantified at  $1.4 \times 10^5 \pm 2.8 \times 10^3$  copies/ng of RNA on Clb and at  $7.4 \times 10^4 \pm 1.0 \times 10^4$  copies/ng of RNA on WS, confirming the results of the DNA microarray. However, CelE (ACX75451) has been identified as a major cellulase EG1 produced by *F. succinogenes* S85 cultivated on cellulose, accounting for 32% of the endoglucanase activity released in the culture medium of the bacterium grown in a chemostat (McGavin and Forsberg, 1988; Malburg et al., 1996). Several hypotheses could be proposed to explain our results, such as the regulation of expression of this gene according to the growth stage of the bacteria, higher induction on Clb or WS structure, and composition resulting in more or less release of repressor compounds from WS. This gene should undergo complementary studies as well as other genes encoding enzymes not characterized yet, and which were over-expressed on WS substrate, such as the genes coding for GH5 ACX74338 and ACX76661. The sequencing of the genome of *F. succinogenes* S85 (Lucas et al., 2009, Unpublished) has identified a large number of GH (102) and CE (16) encoding genes<sup>5</sup> (Lombard et al., 2014), but biochemical characterization of the proteins on different substrates remains necessary to confirm the enzymatic activities predicted by

bioinformatics approaches. Targeted transcriptomics by DNA microarray may give opportunities to enhance our knowledge on how *F. succinogenes* S85 acts and identify which genes or proteins deserve further characterization even though the FibroChip microarray does not target all CAZyme genes present on its genome.

## Application of the FibroChip to a Targeted Metatranscriptomic Analysis of the Rumen Microbiota

The FibroChip was then used to analyze the metatranscriptome of a cow rumen content. The rumen sample was from a dairy cow fed a mixed diet and was previously analyzed for fibrolytic activity by a RNA-seq metatranscriptomic approach (Comtet-Marre et al., 2017). For complex ecosystems, the SNR threshold is usually set at lower values than for pure cultures (Dugat-Bony et al., 2012a). Therefore, for rumen RNA hybridizations, the SNR threshold was set at 2 (He and Zhou, 2008). Only results from GoArrays were used because long probes were more adapted to complex samples. The results obtained by the GoArrays probes are given in Figure 2. The most numerous transcripts identified by the FibroChip coded for GH43, GH5, and GH10 enzymes (Figure 2A). *Bacteroides*, *Fibrobacter*, and *Ruminococcus* were the major contributors to the GH families targeted by the FibroChip, while fungi and protozoa represented 8.1 and 6.7% of total activity, respectively (Figure 2B). Among the 31 *F. succinogenes* S85 genes targeted by the FibroChip, 29 were detected by the GoArrays probes as expressed in the cow rumen. The transcripts not detected were those of a CE6 (ACX75901) and a GH9 (ACX75620) gene. Interestingly, transcripts from the gene *celE* (ACX75451), which was found down-regulated on WS compared to Clb in our *in vitro* experiment, were detected in the cow rumen, indicating that this gene can be expressed *in vivo* on complex substrates. The expression of four *F. succinogenes* S85 genes (coding for GH10 ACX75384, GH9 ACX73672 and ACX75948, and CE6 ACX75360) in the cow rumen was confirmed using RT-qPCR (not shown).

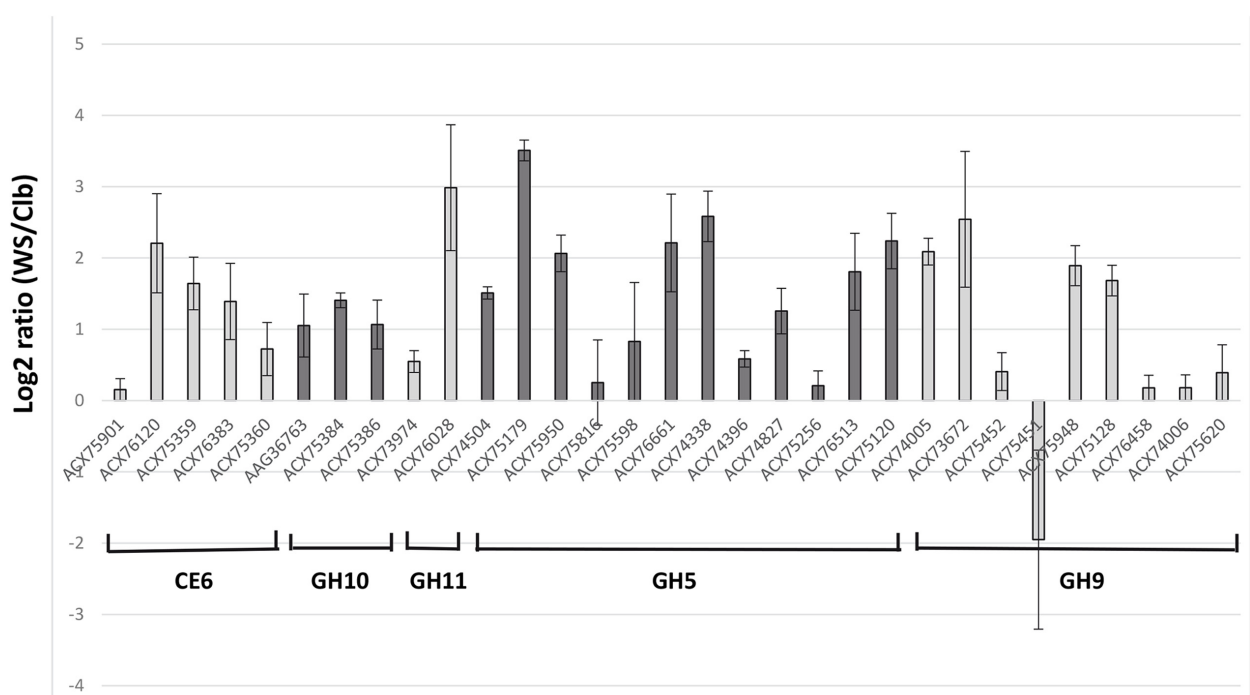
The contribution of eukaryotes was also analyzed further (Figure 3). Fungi produced a diversified array of GH transcripts, covering all but one targeted GH families (Figure 3A), while protozoa produced mainly GH5 and to a lesser extent GH10 and GH11 transcripts (Figure 3C). The genus *Epidinium* appeared

**TABLE 3** | Microarray results obtained after hybridization of genomic DNA from selected bacterial strains.

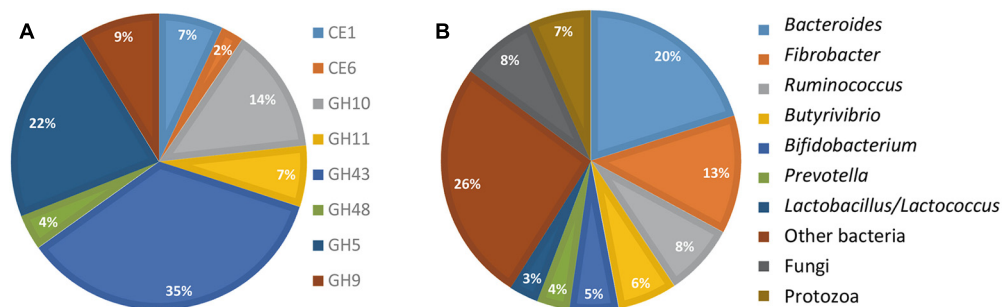
Bacterial strain	Detected genes <sup>a</sup> /targeted genes <sup>b</sup>	Cross-hybridizations <sup>c</sup>	Other detected genes (origin)
<i>F. succinogenes</i> S85	31/31	2	1 ( <i>Fibrobacter</i> sp.)
<i>B. xylanisolvans</i> XBIA	20/20	0	0
<i>R. albus</i> 20	2/2	44	3 ( <i>R. albus</i> 8)
<i>E. coli</i> K12	0/0	0	0

<sup>a</sup>Number of genes detected with the two probe types. <sup>b</sup>Number of genes from the specific strain targeted by the FibroChip. <sup>c</sup>Number of probes (25-mers and GoArrays) that led to unexpected hybridization.

<sup>5</sup><http://www.cazy.org>



**FIGURE 1** | Microarray gene expression analysis of the targeted genes from *Fibrobacter succinogenes* S85 cultivated with cellobiose (Clb) or wheat straw (WS) as a sole substrate. Relative expression of expressed genes on WS versus Clb is expressed as log<sub>2</sub> ratio (WS/Clb) ± SEM. Genes are denominated by the accession number of the protein encoded.



**FIGURE 2** | FibroChip results obtained after hybridization of RNA from rumen contents of a cow. **(A)** Relative contribution of the CAZyme transcripts detected by the FibroChip. **(B)** Taxonomic origin of the CAZyme transcripts detected by the FibroChip. Results presented are obtained with the GoArrays probes. A transcript was considered as detected when its SNR was >2 and at least 65% of the probes were positive. The relative SNR value was not taken into account for the graph representation. CAZyme families **(A)** and microorganism origin **(B)** were picked from CAZy (<http://www.cazy.org>) and taxonomic data of each detected transcript.

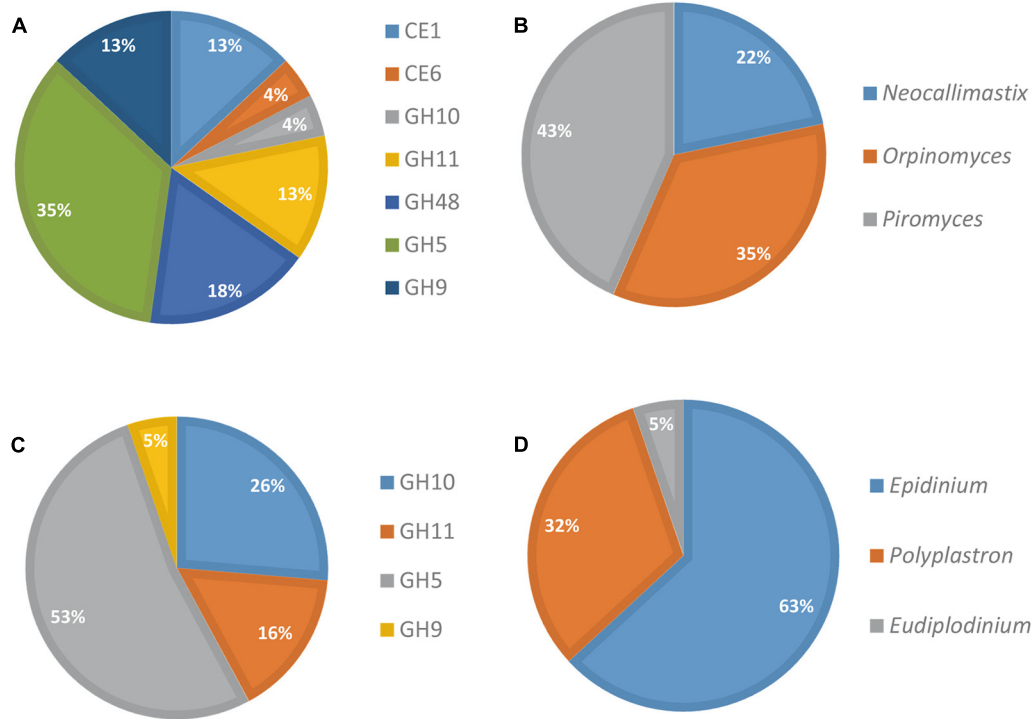
as a major contributor to the production of transcripts from the targeted CAZyme families. However, it should be noted that sequences from rumen fungi and protozoa are still very scarce in databases, and the FibroChip was designed according to the sequences known at the time of the design. The snapshot given by the FibroChip may be biased, particularly for the rumen eukaryotes.

In order to further assess the specificity of the GoArrays probes and confirm the FibroChip data for the rumen sample, transcripts of a *P. communis* GH11 gene (AAG18439) detected by the FibroChip were amplified by RT-PCR and isolated by cloning. Isolated sequence showed 92% identity with the

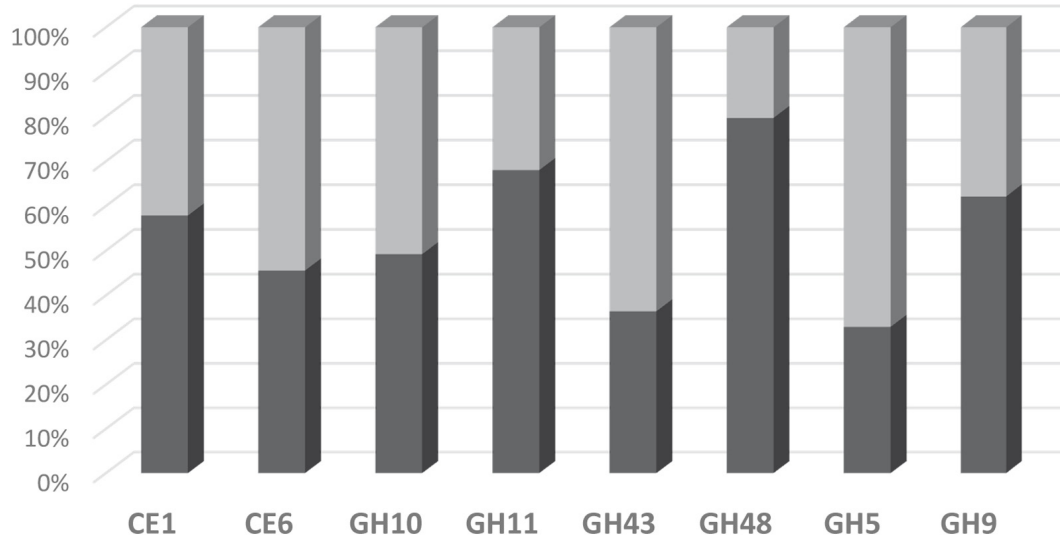
targeted GH11 gene (AAG18439), suggesting that it was a variant of this gene. Such results illustrated the exploratory potential of the FibroChip microarray with a complex microbial community.

## Comparison of Microarray and RNA-Seq Data

The suitability of the FibroChip microarray was then evaluated against RNA-seq analysis performed on the same rumen sample (Comtet-Marre et al., 2017). The distribution of transcripts into the eight CAZyme families as assessed by the two high-throughput approaches was compared (Figure 4). The two



**FIGURE 3** | Distribution of fungi and protozoa to CAZyme transcripts in rumen contents of a cow. Relative contribution of fungi (A) and protozoa (C) CAZyme mRNA detected by the FibroChip; taxonomic origin of fungi (B) and protozoa (D) transcripts detected by the FibroChip. Results presented are obtained with the GoArrays probes. A transcript was considered as detected when its SNR was >2 and at least 65% of the probes were positive. The relative SNR value was not taken into account for the graph representation. CAZyme families (A,C) and microorganism origin (B,D) were picked from CAZy (<http://www.cazy.org>) and taxonomic data of each detected transcript.



**FIGURE 4** | Distribution of detected transcripts into CE and GH families as assessed by the FibroChip (light gray) or the RNA-seq analysis (dark gray). Fifty percent indicates that the contribution of one family to the total of the targeted families is equally estimated by both methods.

methods gave quite similar results for transcripts assigned to families CE1, CE6, GH9, and GH10. However, a discrepancy was observed for the other four families GH11, GH43, GH48, and

GH5, which were either under-estimated (GH11, GH48) or over-estimated (GH43, GH5) by the DNA microarray. In the RNA-seq dataset, there was a high proportion (15%) of GH11 sequences

affiliated to “uncultured bovine ciliate from environmental samples,” which were probably not trapped by the FibroChip, possibly explaining the divergences between the two results. Concerning GH48, there was a clear under-estimation of bacterial GH48 transcripts by the microarray. Indeed, bacteria accounted for 89.4 and 36.4% of total GH48 transcripts in the RNA-seq and FibroChip data, respectively. These results indicated that the FibroChip must be completed by integrating new sets of probes targeting new sequences from GH11 and GH48 families which are now available in the databases. Regarding GH5 the differences between FibroChip and RNA-seq data were mainly due to sequences affiliated to fungal GH5. Indeed, the fungal GH5 transcripts represented 12.7 and 0.5% of the total GH5 sequences as measured by the FibroChip and RNA-seq analysis, respectively. For GH43, the differences between the two approaches were distributed on several taxons. Finally, the over-estimation of transcripts affiliated to families GH5 and GH43 by the FibroChip might result from the under-estimation of other families with the microarray (GH11 and GH48), which would thus lead to different proportions of CAZyme families. In addition, the methodologies used for RNA-seq annotation and taxonomic affiliation of CAZymes transcripts (described in Comtet-Marre et al., 2017) may also introduce some bias in the results and thereby in the comparison.

The FibroChip data highlighted the important, probably still under-estimated, role of eukaryotes in rumen fiber degradation which was in line with the RNA-seq metatranscriptomic analysis. The RNA-seq data indicated that rumen fungi sequences accounted for 6% of the sequences from families targeted by the FibroChip which is close to what was found with the microarray (8%). Qi et al. (2011) showed a higher discovery rate of fungal CAZymes in the rumen of the muskox by metatranscriptomics compared to previous metagenomic studies; however, in other studies, a substantial number of fungal transcripts was not always identified (Edwards et al., 2017). This may depend on how the rumen sample was prepared, in particular if it included the solid phase to which rumen fungi are mostly associated, but also on the procedure of mRNA isolation (Elekwachi et al., 2017). It should also be noted that the FibroChip enabled detection of gene transcripts, which may not totally reflect the actual cellulolytic and hemicellulolytic corresponding activities. This highlights the needs to use complementary methods, such as metabolomics or metaproteomics, to understand more in detail the microbiome activity of the rumen ecosystem (Deusch et al., 2017).

## CONCLUSION

The new FGA, FibroChip, represents a high-throughput tool that would enable monitoring the gene expression of key

## REFERENCES

Abot, A., Arnal, G., Auer, L., Lazuka, A., Labourdette, D., Lamarre, S., et al. (2016). CAZyChip: dynamic assessment of exploration of glycoside hydrolases in microbial ecosystems. *BMC Genomics* 17:671. doi: 10.1186/s12864-016-2988-4

rumen prokaryotic and eukaryotic microorganisms involved in cellulolysis and hemicellulolysis. It would be applicable to study temporal variations in CAZyme gene expression profiles in the rumen according to diet composition, environmental changes, supplementation of feed additives, etc. This monitoring could be applied in nutritional experiments involving a large number of animals, and would allow to better understand the mechanisms of control or modulation of the fiber-degrading activity of rumen microbiota. With the constant increase in available information in the genomic, metagenomic, and specific databases such as CAZy, it will be possible to update and enrich the current version of the FibroChip with new target genes, in particular for rumen eukaryotes.

## AUTHOR CONTRIBUTIONS

SC-M, EF, FC-D, PM, and PP conceived and designed the experiments. AB conceived and performed the animal experiment. SC-M and OB designed the chip. SC-M performed the molecular analyses. SC-M, FC-D, and EF performed the data analysis and wrote the manuscript. All authors have read and approved the final manuscript.

## FUNDING

The work of SC-M was supported by grants from the Association Nationale de la Recherche et de la Technologie (ANRT, Contract 783/2010), the Région Auvergne (Bourse Innovation et Transfert de Technologie, Program 1145), and funding from Lallemand SAS, Blagnac, France (Contract 23000695).

## ACKNOWLEDGMENTS

The authors thank Faouzi Jaziri, Pascale Lepercq, Aurélie Ameilbonne, and Réjane Beugnot for their technical contribution to this work and Christel Béra-Maillet and Caroline Mirande for the gift of plasmids with cloned genes. They acknowledge the contribution of late Kevin J. Shingfield (deceased September 11, 2016) in planning and performing the animal experimentation.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2018.00215/full#supplementary-material>

Arai, T., Kosugi, A., Chan, H., Koukiekolo, R., Yukawa, H., Inui, M., et al. (2006). Properties of cellulosomal family 9 cellulases from *Clostridium cellulovorans*. *Appl. Microbiol. Biotechnol.* 71, 654–660. doi: 10.1007/s00253-005-0249-6

Artzi, L., Bayer, E. A., and Morais, S. (2017). Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* 15, 83–95. doi: 10.1038/nrmicro.2016.164

- Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H., and Henriissat, B. (2012). Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol. Biol.* 12:186. doi: 10.1186/1471-2148-12-186
- Bayat, A. R., Kairenius, P., Stefański, T., Leskinen, H., Comtet-Marre, S., Forano, E., et al. (2015). Effect of camelina oil or live yeasts (*Saccharomyces cerevisiae*) on ruminal methane production, rumen fermentation, and milk fatty acid composition in lactating cows fed grass silage diets. *J. Dairy Sci.* 98, 3166–3181. doi: 10.3168/jds.2014-7976
- Béra-Maillet, C., Gaudet, G., and Forano, E. (2000). Endoglucanase activity and relative expression of glycoside hydrolase genes of *Fibrobacter succinogenes* S85 grown on different substrates. *Biochim. Biophys. Acta* 1543, 77–85. doi: 10.1016/S0167-4838(00)00194-1
- Béra-Maillet, C., Mosoni, P., Kwasiborski, A., Suau, F., Ribot, Y., and Forano, E. (2009). Development of a RT-qPCR method for the quantification of *Fibrobacter succinogenes* S85 glycoside hydrolase transcripts in the rumen content of gnotobiotic and conventional sheep. *J. Microbiol. Methods* 77, 8–16. doi: 10.1016/j.mimet.2008.11.009
- Berg-Miller, M. E., Antonopoulos, D. A., Rincon, M. T., Band, M., Bari, A., Akraiko, T., et al. (2009). Diversity and strain specificity of plant cell wall degrading enzymes revealed by the draft genome of *Ruminococcus flavefaciens* FD-1. *PLoS ONE* 4:e6650. doi: 10.1371/journal.pone.0006650
- Bruc, J. M., Antonopoulos, D. A., Miller, M. E., Wilson, M. K., Yannarell, A. C., Dinsdale, E. A., et al. (2009). Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc. Natl. Acad. Sci. U.S.A* 106, 1948–1953. doi: 10.1073/pnas.0806191105
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2008). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421
- Chaucheyras-Durand, F., Chevaux, E., Martin, C., and Forano, E. (2012). “Use of yeast probiotics in ruminants: Effects and mechanisms of action on rumen pH, fibre degradation, and microbiota according to the diet,” in *Probiotic in Animals*, ed. E. C. Rigobello (Rijeka: Intech), 119–152.
- Chaucheyras-Durand, F., Maseglia, S., Fonty, G., and Forano, E. (2010). Influence of the composition of the cellulolytic flora on the development of hydrogenotrophic microorganisms, hydrogen utilization and methane production in the rumen of gnotobiotically-reared lambs. *Appl. Environ. Microbiol.* 76, 7931–7937. doi: 10.1128/AEM.01784-10
- Comtet-Marre, S., Parisot, N., Lepercq, P., Chaucheyras-Durand, F., Mosoni, P., Peyretailade, E., et al. (2017). Metatranscriptomics reveals the active bacterial and eukaryotic fibrolytic communities in the rumen of dairy cow fed a mixed diet. *Front. Microbiol.* 8:67. doi: 10.3389/fmicb.2017.00067
- Dai, X., Tian, Y., Li, J., Su, X., Wang, X., Zhao, S., et al. (2015). Metatranscriptomic analyses of plant cell wall polysaccharide degradation by microorganisms in the cow rumen. *Appl. Environ. Microbiol.* 81, 1375–1386. doi: 10.1128/AEM.03682-14
- Davison, A., and Blaxter, M. (2005). Ancient origin of glycosyl hydrolase family 9 cellulase genes. *Mol. Biol. Evol.* 22, 1273–1284. doi: 10.1093/molbev/msi107
- Despres, J., Forano, E., Lepercq, P., Comtet-Marre, S., Jubelin, G., Chambon, C., et al. (2016). Xylan degradation by the human gut *Bacteroides xylanisolvens* XB1A<sup>T</sup> involves two distinct gene clusters that are linked at the transcriptional level. *BMC Genomics* 17:326. doi: 10.1186/s12864-016-2680-8
- Deusch, S., Camarinha-Silva, A., Conrad, J., Beifuss, U., Rodehuscord, M., and Siefert, J. (2017). A structural and functional elucidation of the rumen microbiome influenced by various diets and microenvironments. *Front. Microbiol.* 8:1605. doi: 10.3389/fmicb.2017.01605
- Devillard, E., Goodheart, D. B., Karnati, S. K. R., Bayer, E. A., Lamed, R., Miron, J., et al. (2004). *Ruminococcus albus* 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel modular architecture. *J. Bacteriol.* 186, 136–145. doi: 10.1128/JB.186.1.136-145.2004
- Dodd, D., Moon, Y.-H., Swaminathan, K., Mackie, R. I., and Cann, I. K. O. (2010). Transcriptomic analyses of xylan degradation by *Prevotella bryantii* and insights into energy acquisition by xylanolytic bacteroidetes. *J. Biol. Chem.* 285, 30261–30273. doi: 10.1074/jbc.M110.141788
- Dugat-Bony, E., Biderre-Petit, C., Jaziri, F., David, M. M., Denonfoux, J., Lyon, D. Y., et al. (2012a). In situ TCE degradation mediated by complex dehalorespiring communities during biostimulation processes. *Microb. Biotechnol.* 5, 642–653. doi: 10.1111/j.1751-7915.2012.00339.x
- Dugat-Bony, E., Missaoui, M., Peyretailade, E., Biderre-Petit, C., Bouzid, O., Gouinaud, C., et al. (2011). HiSPoD: probe design for functional DNA microarrays. *Bioinformatics* 27, 641–648. doi: 10.1093/bioinformatics/btq712
- Dugat-Bony, E., Peyretailade, E., Parisot, N., Biderre-Petit, C., Jaziri, F., Hill, D., et al. (2012b). Detecting unknown sequences with DNA microarrays: explorative probe design strategies. *Environ. Microbiol.* 14, 356–371. doi: 10.1111/j.1462-2920.2011.02559.x
- Edwards, J. E., Forster, R. J., Callaghan, T. M., Dollhofer, V., Dagar, S. S., Cheng, Y., et al. (2017). PCR and omics based techniques to study the diversity, ecology and biology of anaerobic fungi: insights, challenges and opportunities. *Front. Microbiol.* 8:1657. doi: 10.3389/fmicb.2017.01657
- Elekwach, C. O., Wang, Z., Wu, X., Rabee, A., and Forster, R. J. (2017). Total rRNA-Seq analysis gives insight into bacterial, fungal, protozoal and archaeal communities in the rumen using an optimized RNA isolation method. *Front. Microbiol.* 8:1814. doi: 10.3389/fmicb.2017.01814
- Faulds, C. B., Mandalari, G., Lo Curto, R. B., Bisignano, G., Christakopoulos, P., and Waldron, K. W. (2006). Synergy between xylanases from glycoside hydrolase family 10 and family 11 and a feruloyl esterase in the release of phenolic acids from cereal arabinoxylan. *Appl. Microbiol. Biotechnol.* 71, 622–629. doi: 10.1007/s00253-005-0184-6
- Flint, H. J., Scott, K. P., Duncan, S. H., Louis, P., and Forano, E. (2012). Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* 3, 289–306. doi: 10.4161/gmic.19897
- Gilbert, H. J. (2010). The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiol.* 153, 444–455. doi: 10.1104/pp.110.156646
- Girio, F. M., Fonseca, C., Carvalheiro, F., Duarte, L. C., Marques, S., and Bogel-Lukasik, R. (2010). Hemicelluloses for fuel ethanol: a review. *Bioresour. Technol.* 101, 4775–4800. doi: 10.1016/j.biortech.2010.01.088
- Gottschalk, L. M. F., Oliveira, R. A., da Silva Bon, E. P., Maria, L., Gottschalk, F., and Pinto, E. (2010). Cellulases, xylanases,  $\beta$ -glucosidase and ferulic acid esterase produced by *Trichoderma* and *Aspergillus* act synergistically in the hydrolysis of sugarcane bagasse. *Biochem. Eng. J.* 51, 72–78. doi: 10.1016/j.bej.2010.05.003
- Güllert, S., Fischer, M. A., Turaev, D., Noebauer, B., Ilmberger, N., Wemheuer, B., et al. (2016). Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal major differences in carbohydrate hydrolysis strategies. *Biotechnol. Biofuels* 9:121. doi: 10.1186/s13068-016-0534-x
- He, Z., Deng, Y., and Zhou, J. (2012). Development of functional gene microarrays for microbial community analysis. *Curr. Opin. Biotechnol.* 23, 49–55. doi: 10.1016/j.copbio.2011.11.001
- He, Z., Wu, L., Li, X., Fields, M. W., and Zhou, J. (2005). Empirical establishment of oligonucleotide probe design criteria. *Appl. Environ. Microbiol.* 71, 3753–3760. doi: 10.1128/AEM.71.7.3753-3760.2005
- He, Z., and Zhou, J. (2008). Empirical evaluation of a new method for calculating signal-to-noise ratio for microarray data analysis. *Appl. Environ. Microbiol.* 74, 2957–2966. doi: 10.1128/AEM.02536-07
- Hess, M., Sczyrba, A., Egan, R., Kim, T.-W., Chokhawala, H., Schroth, G., et al. (2011). Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* 331, 463–467. doi: 10.1126/science.1200387
- Hungate, R. E. (1950). The anaerobic mesophilic cellulolytic bacteria. *Bacteriol. Rev.* 14, 1–49.
- Huson, D. H., Mitra, S., Ruscheweyh, H.-J., Weber, N., and Schuster, S. C. (2011). Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* 21, 1552–1560. doi: 10.1101/gr.120618.111
- Irwin, D. C., Zhang, S., and Wilson, D. B. (2000). Cloning, expression and characterization of a family 48 exocellulase, Cel48A, from *Thermobifida fusca*. *Eur. J. Biochem.* 267, 4988–4997. doi: 10.1046/j.1432-1327.2000.01546.x
- Kane, M. D., Jatko, T. A., Stumpf, C. R., Lu, J., Thomas, J. D., and Madore, S. J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res.* 28, 4552–4557. doi: 10.1093/nar/28.22.4552
- Kang, S. H., Evans, P., Morrison, M., and McSweeney, C. (2013). Identification of metabolically active proteobacterial and archaeal communities in the rumen by DNA- and RNA-derived 16S rRNA gene. *J. Appl. Microbiol.* 115, 644–653. doi: 10.1111/jam.12270

- Li, F., and Guan, L. L. (2016). Metatranscriptomic profiling reveals linkages between the active rumen microbiome and feed efficiency in beef cattle. *Appl. Environ. Microbiol.* 83, e00361-17. doi: 10.1128/AEM.00061-17
- Li, Z. P., Liu, H. L., Li, G. Y., Bao, K., Wang, K. Y., Xu, C., et al. (2013). Molecular diversity of rumen bacterial communities from tannin-rich and fiber-rich forage fed domestic Sika deer (*Cervus nippon*) in China. *BMC Microbiol.* 13:151. doi: 10.1186/1471-2180-13-151
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Louhimies, S. (2012). Directive 86/609/EEC on the protection of animals used for experimental and other scientific purposes. *Altern. Lab. Anim.* 2, 217–219.
- Malburg, L. M. Jr., Iyo, A. H., and Forsberg, C. W. (1996). A novel family 9 endoglucanase gene (celD), whose product cleaves substrates mainly to glucose, and its adjacent upstream homolog (celE) from *Fibrobacter succinogenes* S85. *Appl. Environ. Microbiol.* 62, 898–906.
- Matulova, M., Nouaille, R., Capek, P., Péan, M., Delort, A.-M., and Forano, E. (2008). NMR study of cellulose and wheat straw degradation by *Ruminococcus albus* 20. *FEBS J.* 275, 3503–3511. doi: 10.1111/j.1742-4658.2008.06497.x
- McGavin, M., and Forsberg, C. W. (1988). Isolation and characterization of endoglucanases 1 and 2 from *Bacteroides succinogenes* S85. *J. Bacteriol.* 170, 2914–2922. doi: 10.1128/jb.170.7.2914-2922.1988
- Nouaille, R., Matulova, M., Pátoprstý, V., Delort, A.-M., and Forano, E. (2009). Production of oligosaccharides and cellobionic acid by *Fibrobacter succinogenes* S85 growing on sugars, cellulose and wheat straw. *Appl. Microbiol. Biotechnol.* 83, 425–433. doi: 10.1007/s00253-009-1884-0
- Orpin, C. G. (1975). Studies on the rumen flagellate *Neocallimastix frontalis*. *J. Gen. Microbiol.* 91, 249–262. doi: 10.1099/00221287-91-2-249
- Peer, A., Smith, S. P., Bayer, E. A., Lamed, R., and Borovok, I. (2009). Noncellulosomal cohesin- and dockerin-like modules in the three domains of life. *FEMS Microbiol. Lett.* 291, 1–16. doi: 10.1111/j.1574-6968.2008.01420.x
- Pitta, D. W., Pinchak, W. E., Indugu, N., Vecchiarelli, B., Sinha, R., and Fulford, J. D. (2016). Metagenomic analysis of the rumen microbiome of steers with wheat-induced frothy bloat. *Front. Microbiol.* 7:689. doi: 10.3389/fmicb.2016.00689
- Purushe, J., Fouts, D. E., Morrison, M., White, B. A., Mackie, R. I., Coutinho, P. M., et al. (2010). Comparative genome analysis of *Prevotella ruminicola* and *Prevotella bryantii*: insights into their environmental niche. *Microb. Ecol.* 60, 721–729. doi: 10.1007/s00248-010-9692-8
- Qi, M., Wang, P., O'Toole, N., Barboza, P. S., Ungerfeld, E., Leigh, M. B., et al. (2011). Snapshot of the eukaryotic gene expression in muskoxen rumen - a metatranscriptomic approach. *PLOS ONE* 6:e20251. doi: 10.1371/journal.pone.0020521
- Rimour, S., Hill, D., Milton, C., and Peyret, P. (2005). GoArrays: highly dynamic and efficient microarray probe design. *Bioinformatics* 21, 1094–1103. doi: 10.1093/bioinformatics/bti112
- Shinkai, T., Mitsumori, M., Sofyan, A., Kanamori, H., Sasaki, H., Katayose, Y., et al. (2016). Comprehensive detection of bacterial carbohydrate-active enzyme coding genes expressed in cow rumen. *Anim. Sci. J.* 87, 1363–1370. doi: 10.1111/asj.12585
- Suen, G., Weimer, P. J., Stevenson, D. M., Aylward, F. O., Boyum, J., Deneke, J., et al. (2011). The complete genome sequence of *Fibrobacter succinogenes* S85 reveals a cellulolytic and metabolic specialist. *PLOS ONE* 6:e18814. doi: 10.1371/journal.pone.0018814
- Tottey, W., Denonfoux, J., Jaziri, F., Parisot, N., Missaoui, M., Hill, D., et al. (2013). The human gut chip “HuGChip”, an explorative phylogenetic microarray for determining gut microbiome diversity at family level. *PLOS ONE* 8:e62544. doi: 10.1371/journal.pone.0062544
- Tu, Q., Yu, H., He, Z., Deng, Y., Wu, L., Van Nostrand, J. D., et al. (2014). GeoChip 4: a functional gene-array-based high-throughput environmental technology for microbial community analysis. *Mol. Ecol. Resour.* 14, 914–928. doi: 10.1111/1755-0998.12239
- Wang, L., Hatem, A., Catalyurek, U. V., Morrison, M., and Yu, Z. (2013). Metagenomic insights into the carbohydrate-active enzymes carried by the microorganisms adhering to solid digesta in the rumen of cows. *PLOS ONE* 8:e78507. doi: 10.1371/journal.pone.0078507
- White, B. A., Lamed, R., Bayer, E. A., and Flint, H. J. (2014). Biomass utilization by gut microbiomes. *Annu. Rev. Microbiol.* 68, 279–296. doi: 10.1146/annurev-micro-092412-155618
- Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–W451. doi: 10.1093/nar/gks479
- Youssef, N. H., Couger, M., Struchtemeyer, C. G., Ligginstoffer, A. S., Prade, R. A., Najjar, F. Z., et al. (2013). Genome of the anaerobic fungus *Orpinomyces* sp. C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. *Appl. Environ. Microbiol.* 79, 4620–4634. doi: 10.1128/AEM.00821-13
- Zhang, J., Siika-Aho, M., Tenkanen, M., and Viikari, L. (2011). The role of acetyl xylan esterase in the solubilization of xylan and enzymatic hydrolysis of wheat straw and giant reed. *Biotechnol. Biofuels* 4:60. doi: 10.1186/1754-6834-4-60
- Ziemer, C. J. (2014). Newly cultured bacteria with broad diversity isolated from eight-week continuous culture enrichments of cow feces on complex polysaccharides. *Appl. Environ. Microbiol.* 80, 574–585. doi: 10.1128/AEM.03016-13

**Conflict of Interest Statement:** FC-D is an employee of Lallemand SAS Company. SC-M was also employed by Lallemand Company during part of this work. Lallemand contributed to the funding of this work.

The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Comtet-Marre, Chaucheyras-Durand, Bouzid, Mosoni, Bayat, Peyret and Forano. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.