



# Metatranscriptome Profiling of a Specialized Microbial Consortium during the Degradation of Nixtamalized Maize Pericarp

José Germán Serrano-Gamboa,<sup>a</sup> Carlos Abraham Fócil-Espinosa,<sup>a</sup> Pedro J. Cabello-Yeves,<sup>b</sup> Felipe H. Coutinho,<sup>b</sup> Rafael Antonio Rojas-Herrera,<sup>a</sup>  Mónica Noel Sánchez-González<sup>a</sup>

<sup>a</sup>Facultad de Ingeniería Química, Universidad Autónoma de Yucatán, Mérida, México

<sup>b</sup>Evolutionary Genomics Group, Departamento de Producción Vegetal y Microbiología, Universidad Miguel Hernández, Alicante, Spain

**ABSTRACT** Lignocellulose degradation by microbial consortia is multifactorial; hence, it must be analyzed from a holistic perspective. In this study, the temporal transcriptional activity of consortium PM-06, a nixtamalized maize pericarp (NMP) degrader, was determined and related to structural and physicochemical data to give insights into the mechanism used to degrade this substrate. Transcripts were described in terms of metabolic profile, carbohydrate-active enzyme (CAZyme) annotation, and taxonomic affiliation. The PM-06 gene expression pattern was closely related to the differential rates of degradation. The environmental and physiological conditions preceding high-degradation periods were crucial for CAZyme expression. The onset of degradation preceded the period with the highest degradation rate in the whole process, and in this time, several CAZymes were upregulated. Functional analysis of expressed CAZymes indicated that PM-06 overcomes NMP recalcitrance through modular enzymes operating at the proximity of the insoluble substrate. Increments in the diversity of expressed modular CAZymes occurred in the last stages of degradation where the substrate is more recalcitrant and environmental conditions are stressing. Taxonomic affiliation of CAZyme transcripts indicated that *Paenibacillus macerans* was fundamental for degradation. This microorganism established synergistic relationships with *Bacillus thuringiensis* for the degradation of cellulose and hemicellulose and with *Microbacterium*, *Leifsonia*, and *Nocardia* for the saccharification of oligosaccharides.

**IMPORTANCE** Nixtamalized maize pericarp is an abundant residue of the tortilla industry. Consortium PM-06 efficiently degraded this substrate in 192 h. In this work, the temporal transcriptional profile of PM-06 was determined. Findings indicated that differential degradation rates are important sample selection criteria since they were closely related to the expression of carbohydrate-active enzymes (CAZymes). The initial times of degradation were crucial for the consumption of nixtamalized pericarp. A transcriptional profile at the onset of degradation is reported for the first time. Diverse CAZyme genes were rapidly transcribed after inoculation to produce different enzymes that participated in the stage with the highest degradation rate in the whole process. This study provides information about the regulation of gene expression and mechanisms used by PM-06 to overcome recalcitrance. These findings are useful in the design of processes and enzyme cocktails for the degradation of this abundant substrate.

**KEYWORDS** lignocellulose degradation, microbial consortium, metatranscriptome, nixtamalized maize pericarp, carbohydrate-active enzyme expression

Lignocellulosic residues derived from agroindustry are abundant and low-cost resources with the potential to produce second-generation biofuels and other value-added products (1, 2). Tortillas, a staple food in the Mexican diet, are produced

**Editor** Jeffrey A. Gralnick, University of Minnesota

**Copyright** © 2022 Serrano-Gamboa et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Mónica Noel Sánchez-González, [monica.sanchez@correo.uady.mx](mailto:monica.sanchez@correo.uady.mx).

The authors declare no conflict of interest.

**Received** 22 November 2021

**Accepted** 3 December 2021

**Published** 5 January 2022

in a process called nixtamalization, where maize grains are cooked under alkaline conditions. The tortilla industry annually generates approximately 14,400,000 m<sup>3</sup> of wastewater containing about 300,000 tons of nixtamalized maize pericarp (NMP) (3). Although some pericarp components are lost during nixtamalization, NMP is an important source of fermentable sugars and fine chemicals (4).

Transformation of agro-industrial residues is limited by the recalcitrance of lignocellulose structures. In nature, microbial communities are involved in the decomposition of lignocellulosic substrates as part of the carbon cycle. Through the enrichment of environmental samples, it is possible to obtain specialized microbial consortia with desired degradative properties (5, 6).

The use of specialized lignocellulose-degrading consortia is a promising approach because selection enhances the carbohydrate-active enzyme (CAZyme) repertoire, achieving efficient decomposition of recalcitrant substrates (7, 8). During degradation, notable changes occur under environmental conditions, such as pH, cell and protein concentration, nutrient availability, and substrate composition among others (9–11).

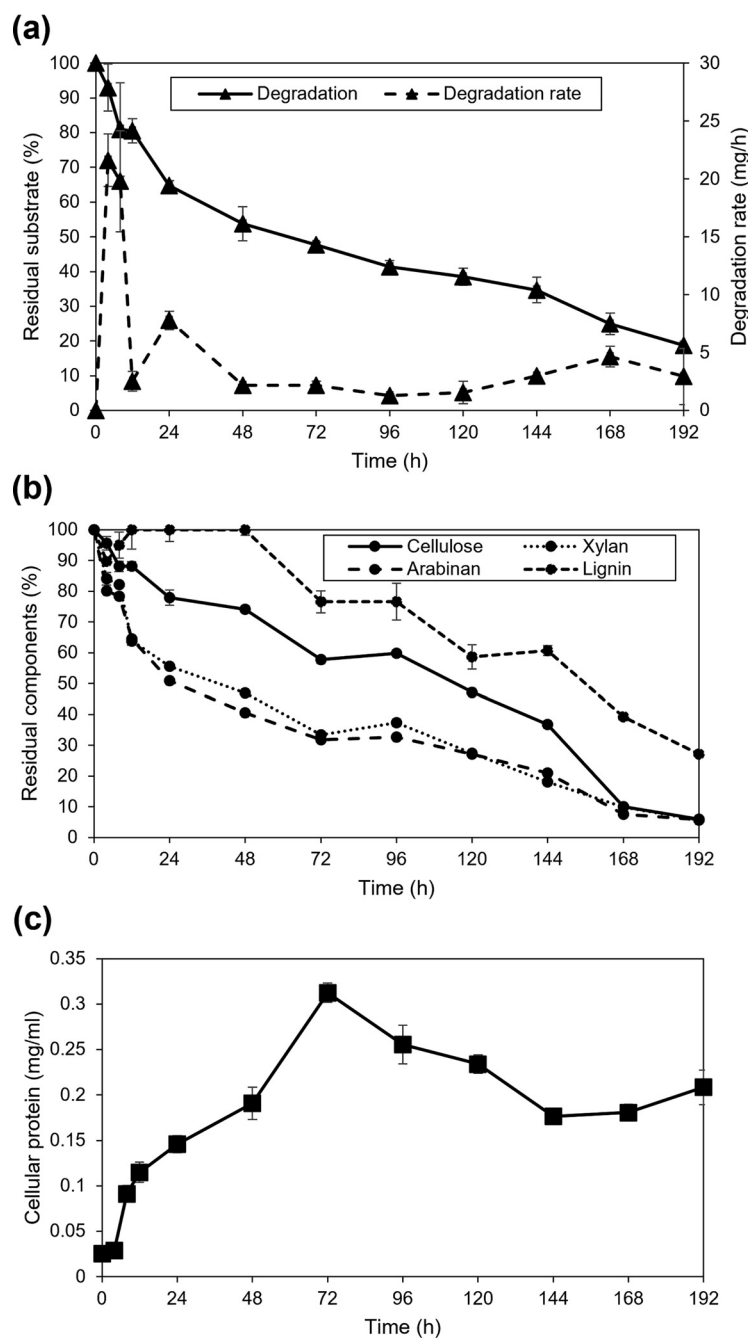
Temporal metagenomic and metatranscriptomic analyses contribute to the understanding of microbial interactions, enzyme mechanisms, and functions under particular environmental conditions along a degradation process (7, 12–15). To obtain representative samples, the selection criteria are an important task of transcriptomic analysis. In temporal studies, the selection of samples is frequently based on degradation stages (12, 13) or enzyme activities (14). The stages of degradation are usually defined based on substrate consumption rates and calculated using the inoculation time as a reference (integral rates) (7, 12). However, if rates are calculated using the preceding time as a reference (differential rates), even small changes in the speed of substrate consumption can be determined.

Consortium PM-06, a specialized community enriched from NMP endogenous microbiota, degrades 87% of this tortilla residue in 192 h. *Aneurinibacillus migulanus*, *Paenibacillus macerans*, *Bacillus coagulans*, *Microbacterium* sp. LCT-H2, and *Bacillus thuringiensis* are the most abundant species in this community, comprising 86% of the population (16). Although PM-06 is an efficient NMP degrader, there is scarce information about the role of CAZymes in the process.

In this study, a temporal analysis of the transcriptional activity of consortium PM-06 during the degradation of NMP was performed. Samples were selected based on differential degradation rates under the hypothesis that the expression of CAZymes is up-regulated in times preceding a high degradation rate period. Transcripts of selected samples were described in terms of metabolic profile, CAZyme annotation, and taxonomic affiliation. Moreover, the transcriptional CAZyme profile was related to structural and physicochemical data to give insights into the mechanism used for the degradation of NMP.

## RESULTS

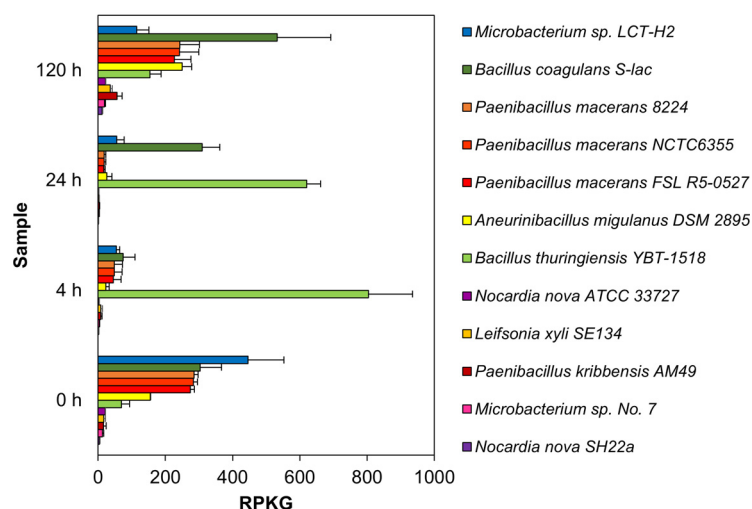
**Microbial growth and degradation kinetics.** Consortium PM-06 degraded 81% of NMP solids after 192 h (Fig. 1a). Differential degradation rates varied along the process, obtaining the highest values during the first 8 h (~20 mg/h). Two additional increments were determined at 24 h (7.7 mg/h) and at 168 h (4.6 mg/h) (Fig. 1a). At the beginning of the process, NMP contained 15.3% cellulose, 23.7% xylan, 12.6% arabinose substituents, 33% starch, 4.4% lignin, and 11% of other components (ashes, extractives, uronic acids, and protein). Hemicellulose (arabinan and xylan) was consumed faster than cellulose and lignin. The fastest cellulose consumption was determined from 144 to 168 h, where concentration decreased 26.7% (Fig. 1b). At the end of the process, NMP residual solids contained more lignin than the initial substrate. Cell concentration exponentially increased after 4 h of incubation, reaching a maximum (0.31 mg/mL) at 72 h (Fig. 1c). After this time, the number of cells decayed. Selection of samples for metatranscriptomic analysis was based on differential degradation rates. Therefore,



**FIG 1** NMP degradation and microbial growth kinetics. (a) NMP solids consumption and differential degradation rates. (b) Degradation profile of NMP components. (c) PM-06 microbial growth in terms of cellular protein concentration.

samples from high differential degradation rate periods (4 and 24 h) and from times preceding a high-rate period (0 h and 120 h) were selected.

**Reference metatranscriptome overview.** A reference metatranscriptome containing a total of 14,257 contigs was obtained by *de novo* cross-assembly of 308 million high-quality reads from all samples. The reference metatranscriptome contained 48,516 transcripts encoding 44,758 predicted proteins. Taxonomic annotation showed that transcripts were assigned mainly to the Bacteria domain (97.34%). *Firmicutes* and *Actinobacteria* were the dominant phyla (50.9 and 45.8%, respectively), and at the species level (sequence identity of  $\geq 99\%$ ), the predicted proteins were largely associated



**FIG 2** Transcriptional activity of consortium members estimated by recruitment analysis of metatranscriptomic reads against 12 bacterial genomes. Genome coverage in each sample is expressed in reads per kilobase of genome per gigabase of metatranscriptome (RPKG).

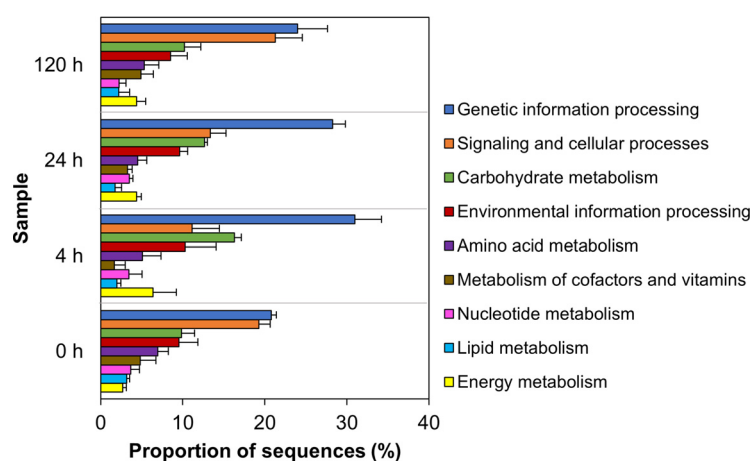
with *Aneurinibacillus migulanus* (13.26%), *Paenibacillus macerans* (12.89%), *Nocardia nova* (9.98%), *Bacillus coagulans* (6.90%), *Microbacterium* sp. LCT-H2 (4.63%), *Bacillus thuringiensis* (2.42%), and *Leifsonia xyli* (2.09%). All identified taxa were consistent with those previously identified in the PM-06 metagenome (16).

**Transcriptional activity of consortium members.** The transcriptional activity of microorganisms was determined through a recruitment analysis (Fig. 2). For this purpose, the parameter reads per kilobase of genome per megabase of metatranscriptome (RPKG) was estimated by calculating the coverage of the mRNA reads against genomes of the most abundant species and according to taxonomic annotation of the reference metatranscriptomic and metagenomic data.

*Microbacterium*, *Paenibacillus*, *Bacillus*, and *Aneurinibacillus* were the most transcriptionally active genera with RPKG values greater than 50. At 0 h, the genome of *Microbacterium* sp. LCT-H2 recruited a larger set of mRNAs (445.62 RPKG), followed by *Bacillus coagulans* S-lac (303.07 RPKG) and *Paenibacillus macerans* 8224 (285.09 RPKG). In contrast, the transcriptional activity at 4 h was predominantly attributed to *Bacillus thuringiensis* YBT 518 (803.61 RPKG). At 24 h, genomes of *Bacillus thuringiensis* YBT 518 and *Bacillus coagulans* S-lac recruited most of the metatranscriptome reads (619.98 RPKG and 309.34 RPKG, respectively). Finally, at 120 h *Bacillus*, *Paenibacillus*, and *Aneurinibacillus* were the most transcriptionally active genera.

**Metabolic profile.** Functional analysis using KEGG orthology annotation identified the most represented pathways (Fig. 3). Metabolism of carbohydrates, genetic information processing, and environmental information processing were the most representative functions in all samples. The distribution of KEGG categories changed with degradation, indicating metabolic shifts. At 4 h, the proportion of transcripts associated with carbohydrate metabolism, genetic information processing, and energy metabolism increased and reached the highest values (16.29%, 21.75%, and 6.37%, respectively) in the whole process, indicating a stage of large uptake of carbon sources and preparation for exponential growth. In contrast, functions associated with metabolism of cofactors and vitamins decreased by one-third compared to 0 h. Lipid, nucleotide, and amino acid metabolism showed similar proportions in all samples.

**Global pattern of CAZyme genes.** A total of 1,221 coding sequences (2.7% of predicted proteins) were identified as putative CAZymes; however, only 555 were consistently annotated by the three tools provided by the automated CAZy annotation database (dbCAN; Data S1A in the supplemental material). From these sets, glycosyl hydrolase (GH) was the most abundant group (56.58%), followed by glycosyl transferases



**FIG 3** Metabolic profile of transcripts according to KEGG orthology functional categories. Percentages correspond to the mean of retrieved mapped sequences in each sample ( $n = 3$ ). The average proportion of the categories shown was greater than 1%. Error bars represent the standard deviation.

(GT) (27.93%) and carbohydrate esterases (CE) (11.17%). Coding sequences for carbohydrate-binding modules (CBM), polysaccharide lyases (PL), and auxiliary activities (AA) were identified in smaller proportions (Data S1A). The taxonomic affiliation of CAZymes was dominated by *Firmicutes* (62%) and *Actinobacteria* (33%) phyla. At the genus level, *Paenibacillus*, *Bacillus*, *Microbacterium*, and *Leifsonia* were the major contributors to the expression of different CAZyme families.

**Modular CAZymes in NMP degradation.** PM-06 exhibited transcripts encoding CAZymes with multidomain architecture involved in the decomposition of the structural components of NMP (Table S1). Identified catalytic domains associated with CBMs included several hemicellulases (GH10, GH26, GH43, GH44, and GH81), cellulases (GH5, GH6, GH9, GH48, and GH74), and a lytic polysaccharide monooxygenase (LPMO) belonging to the AA10 family. Except for AA10, which is taxonomically affiliated with *Bacillus cereus*, all modular CAZymes were expressed by *Paenibacillus macerans* (Table 1).

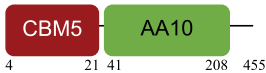
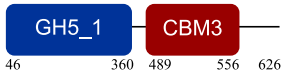
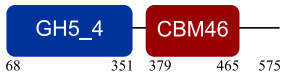
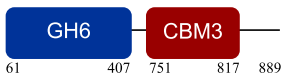
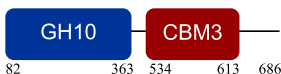









CBM3 was the most frequent among all expressed CBMs. The catalytic domains associated with this module included GH5 and GH6 (endoglucanases), GH48 (exoglucanase), and GH10 (endoxyylanase). CBM59 (mannan/xylan/cellulose-binding) was associated with GH10, suggesting a versatile role of this enzyme. The architecture of LPMO included CBM5 with affinity to crystalline cellulose and complex polysaccharides (17, 18).

CBM6, CBM13, and CBM66 (involved in the recognition of substituted xylan [19]) were identified in two different multimodular enzymes containing GH43 (xylosidase/arabinofuranosidase) as the catalytic domain. CBM32 and CBM35 participated in different domain architectures targeting hemicellulose in association with GH26, GH44, or GH81.

**Taxonomic affiliation of differentially expressed transcripts.** The differential expression analysis performed to determine variations in transcriptional activity identified 21,000 differentially expressed (DE) transcripts (Fig. S2). A total of 407 genes encoding putative CAZymes were DE. From this set, all genes annotated as GT and CBM (unassociated) and/or affiliated with low-abundant microbial genera were manually removed. Finally, 181 CAZymes were selected for further analysis (Data S1B in the supplemental material).

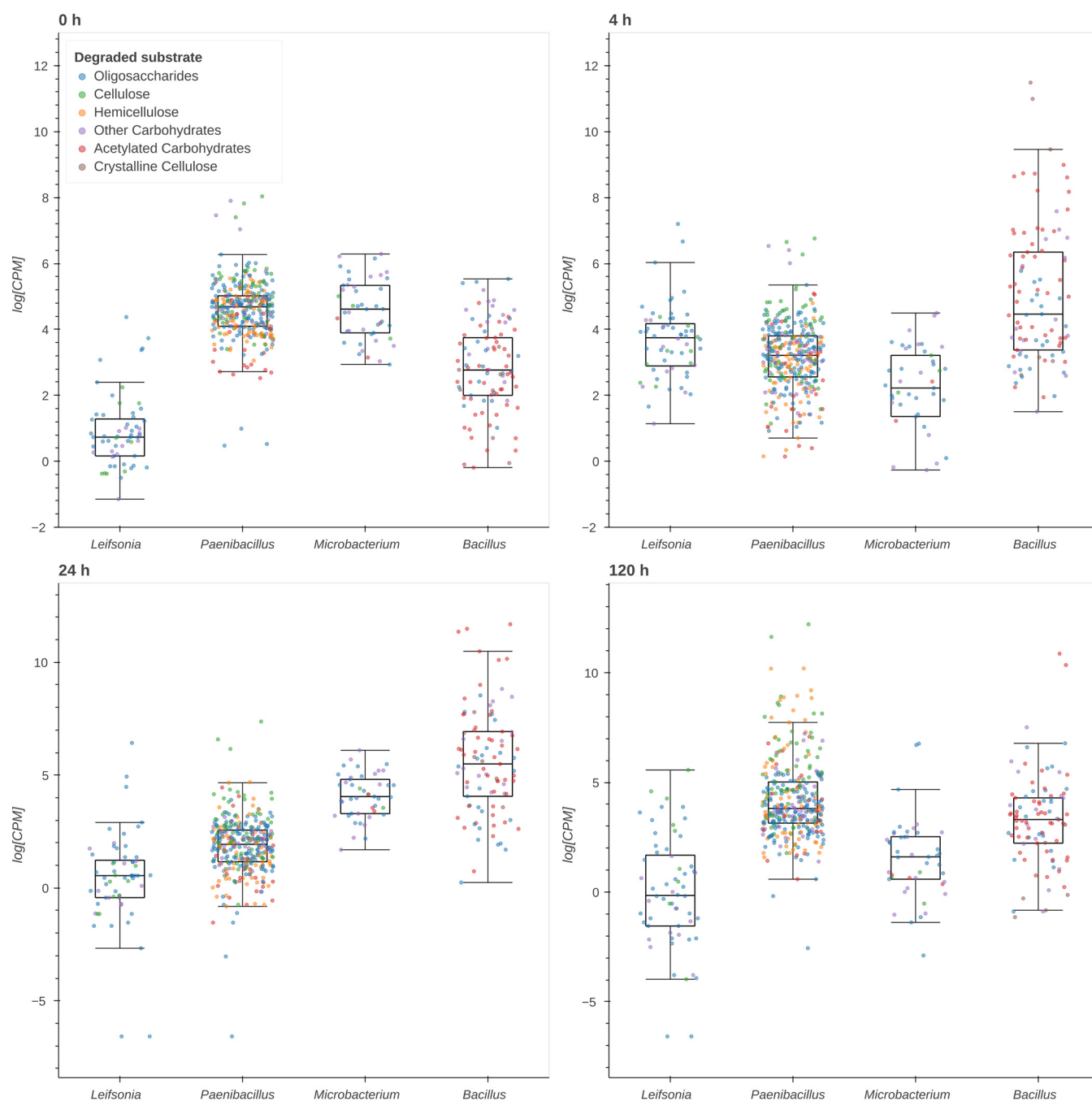
DE transcripts were taxonomically affiliated and grouped according to preference over oligosaccharides, cellulose, hemicellulose, other carbohydrates, acetylated carbohydrates, and crystalline cellulose (Fig. 4) (DE transcript groups are defined in Data S1B). *Paenibacillus* and *Microbacterium* transcripts presented the highest expression levels ( $\sim 10^5$  counts per million [CPM]) at 0 h. *Paenibacillus* expressed a diverse set of CAZymes, including cellulases, hemicellulases, and oligosaccharidases. *Microbacterium*, *Leifsonia*, and *Nocardia* expressed a smaller and less diverse set of CAZymes, which was dominated by oligosaccharidases and other glycanases. At 4 and 24 h, *Bacillus* showed

**TABLE 1** Predicted structure and function of modular CAZymes identified in the PM-06 metatranscriptome

Taxonomic affiliation and time of major expression	CAZyme annotation and architecture <sup>a</sup>	Functional characterization (reference)
<i>Bacillus cereus</i> 4 h		Lytic polysaccharide monoxygenase (59)
<i>Paenibacillus macerans</i> 120 h		Endoglucanase (60)
<i>Paenibacillus macerans</i> 0 h, 120 h		$\beta$ -1,3-1,4-Endoglucanase (61)
<i>Paenibacillus macerans</i> 0 h, 120 h		Cellobiohydrolase (60)
<i>Paenibacillus macerans</i> 0 h		Endo-1,4- $\beta$ -xylanase (62)
<i>Paenibacillus macerans</i> 120 h		Endo-1,4- $\beta$ -xylanase (63)
<i>Paenibacillus macerans</i> 120 h		Cellobiohydrolase (64)
<i>Paenibacillus macerans</i> 0 h		Xyloglucanase (60, 64)
<i>Paenibacillus macerans</i> 120 h		Endoglucanase (65)
<i>Paenibacillus macerans</i> 0 h, 120 h		Mannanase (60)
<i>Paenibacillus macerans</i> 120 h		$\alpha$ -L-Arabinofuranosidase/ $\beta$ -xylosidase (19)
<i>Paenibacillus macerans</i> 0 h		$\alpha$ -L-Arabinofuranosidase (64)
<i>Paenibacillus macerans</i> 120 h		Endoglucanase/xylanase (60)
<i>Paenibacillus macerans</i> 0 h		$\beta$ -1,3-Glucanase (66)

<sup>a</sup>Enzyme architectures are schematically (not to scale) represented. Numbers indicate amino acid positions of predicted domains.

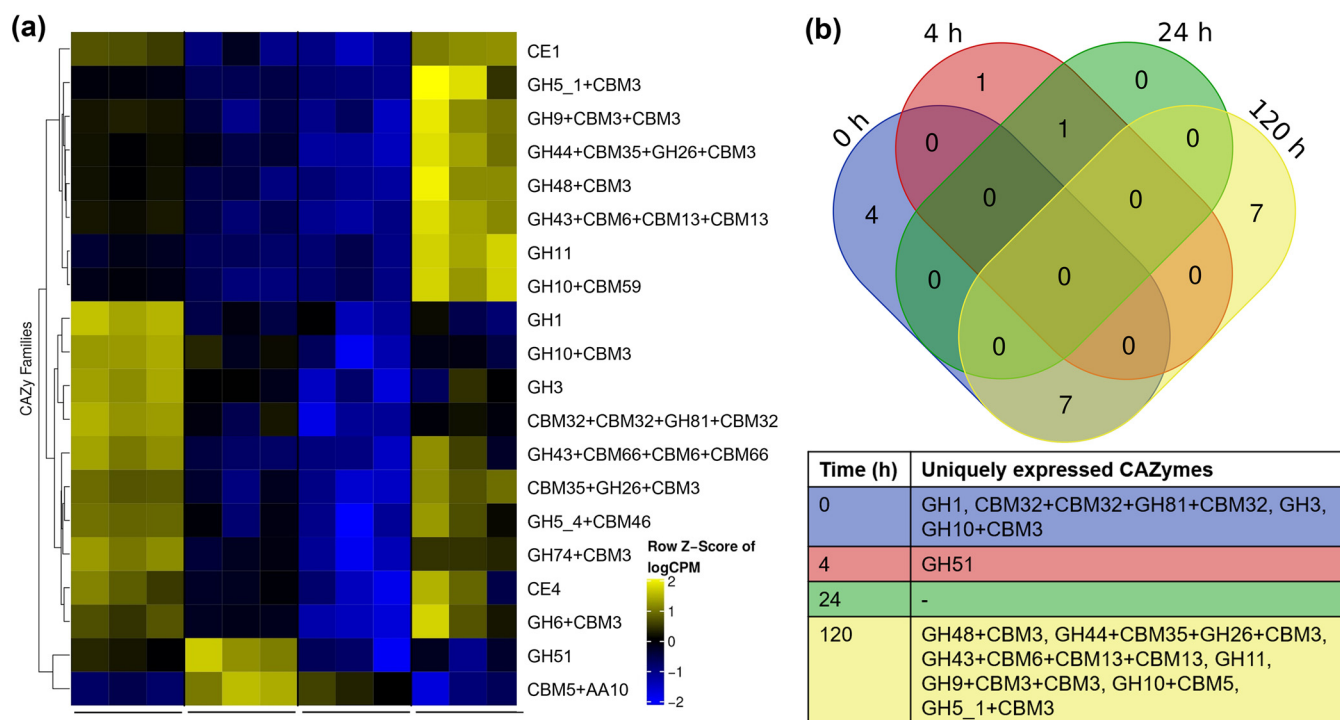




**FIG 4** Taxonomic affiliation of differentially expressed transcripts grouped according to substrate preference. Expression values were estimated as log CPM.

high levels of expression ( $\sim 10^8$  CPM) of a carbohydrate esterase associated with hemicellulose deacetylation, an LPMO, and other CAZymes not directly related to the degradation of lignocellulose (i.e., amylases). Although *Paenibacillus* expression levels were low in these times ( $\sim 10^3$  CPM), they included a broad set of enzyme functionalities. At 120 h, *Paenibacillus* CAZymes targeting cellulose and hemicellulose reached  $10^{10}$  CPM, the maximum expression value of the whole process. The expression levels of *Microbacterium* and *Leifsonia* were lower than those determined at 0 h.

**Differentially expressed CAZymes involved in NMP degradation.** A heatmap was generated to visualize gene expression pattern and clustering (Fig. 5a). The selection of CAZyme transcripts was based on the activity over NMP components, expression values (transcript abundance above  $1 \times 10^4$  CPM), and functional domains. The



**FIG 5** Differential expression analysis. (a) Heat map of differentially expressed CAZyme genes (row Z-score of normalized CPM values) involved in NMP degradation. (b) Venn diagram of upregulated CAZymes (average Z-score of  $\geq 0.4$ ), indicating shared and unique families between time samples.

hierarchical clustering of families was performed based on the Z-score computed across samples. Clustering revealed two groups with similar expression profiles. The first contained 0- and 120-h samples, and the second included 4- and 24-h transcripts. This result correlated with the pattern observed in the recruitment analysis (Fig. 2).

A total of 11 and 14 CAZymes were overexpressed at 0 and 120 h with only 7 coincidences. Conversely, at 4 and 24 h, only two and one CAZyme, respectively, were upregulated. (Fig. 5b). CAZymes with CE1, CE4, GH3, GH10, GH26, and GH43 domains, related to hemicellulose decomposition (20, 21) and saccharification of solubilized oligosaccharides (22, 23), were upregulated at the onset of the degradation process. Likewise, significant expression of GH5 (average of  $5.8 \times 10^4$  CPM) and GH6 ( $4.2 \times 10^4$  CPM) families, associated with cellulose hydrolysis, was observed. (Fig. 5a).

At 4 h, the transcription of CAZymes decreased, and only GH51 (arabinofuranosidase) and AA10 (LPMO) were upregulated. All CAZyme families were downregulated at 24 h. The upregulated set of CAZyme genes at 120 h included GH5, GH9, GH10, and GH11 genes, which hydrolyze internal bonds of cellulose and xylan (24, 25). A notable expression of GH48 ( $1.32 \times 10^5$  CPM), an exoglucanase with a prominent role in the decomposition of cellulose, was determined at this time (26). The high expression of CE1 ( $1.6 \times 10^5$  CPM) and CE4 ( $7.4 \times 10^4$  CPM) genes indicated the deacetylation of hemicellulose.

## DISCUSSION

In nature, lignocellulose degradation is a complex and dynamic process that involves the participation of different enzymes working in constant interaction. The amount and diversity of the enzymes produced depend on the source of the consortium, enrichment and selection conditions, and substrate structure and composition. In addition to CAZymes, the transformation of complex polysaccharides involves biochemical and physiological responses, many of them still unknown. In this study, the temporal expression of CAZyme genes by consortium PM-06 was analyzed and associated with environmental conditions and substrate composition changes to get insight into the mechanism used for the degradation of NMP.



PM-06 degraded NMP with different rates, obtaining the highest rate during the initial 8 h. The results indicated a relationship between the differential degradation rates and the transcriptional activity. In times preceding high degradation rate periods (0 and 120 h), the transcriptional activity was divided between different microbial species. The participation of distinct microorganisms suggests a synergistic behavior between degraders, cooperating for the production of enzymes with complementary functions in the so-called niche partition (27). Although at 0 and 120 h, several CAZyme genes were upregulated, there were few coincidences, suggesting differences in the regulation process.

Interestingly, at 0 h, the lowest concentration of cells was determined; however, a high number of CAZymes was overexpressed. *Paenibacillus* and *Microbacterium* were the main contributors to the expression of CAZymes related to the saccharification of polysaccharides and oligosaccharides, respectively. The initial stages of microbial growth have been associated with dramatic changes in the proteome and transcriptome for the production of digestive enzymes, biomass building, and preparation for cell division (28). In some lignocellulose degradation processes, the initial times have been recognized as important for degradation (14); however, there is scarce information about the CAZymes expressed, mainly because there are technical difficulties associated with the small number of cells (29). Consortia stabilization includes sequential transfers that select the population and accelerate the adaptation of bacteria to new environments (30). In this rapid adaptation, bacteria use strategies such as the “poised state” of the RNA polymerase, participation of the transcription initiation factor  $\sigma 70$ , or the two-component system to transduce information from the environment between other mechanisms (28, 31, 32). The complex control regulatory networks responsible for predictable and sustainable responses to predefined inputs are conceptually related to biological memory (33, 34). Biological memory can be defined as “enduring changes in the mechanism of behavior based on prior experiences with environmental inputs” (35). Therefore, the high transcriptional levels at 0 h could be the result of biological memory restoration mechanisms. More studies are necessary to determine the regulation mechanisms involved in CAZyme transcription at the onset of degradation.

The population density at 120 h was higher than at 0 h; however, microorganisms were in the decay stage, probably because of a low concentration of soluble sugars. The high degradation rates obtained during the initial 48 h generated soluble sugars that supported exponential growth of the community (Fig. 1c). After this period, degradation rates decreased, probably because of the presence of a more recalcitrant substrate (increased amount of lignin), causing starvation. These conditions may induce the expression of a high number of CAZymes. In this time, like at 0 h, CAZymes were expressed mainly by *Paenibacillus* and *Microbacterium*.

At 4 and 24 h, degradation rates increased, and the expression of lignocellulose-targeted CAZymes was downregulated. The degradative activity incremented the concentration of soluble sugars, reducing the expression of CAZymes probably by carbon catabolite repression (36). Moreover, the presence of easily assimilable sugars could decrease the metabolic cooperation between consortia members (27). Therefore, in these times, the transcriptional activity was mainly devoted to *Bacillus thuringiensis*. At 4 h, the transcriptional activity of *Bacillus* CAZymes was associated with carbohydrate deacetylation and saccharification of cellulose and complex polysaccharides by LPMO (AA10). Cellulose is more recalcitrant than hemicellulose and showed lower degradation rates. LPMO activity may facilitate the action of hydrolytic enzymes, increasing the efficiency of cellulose degradation (37, 38). At 24 h, the expression of all CAZymes was downregulated. In this time, the concentration of soluble sugars could be incrementing the competition between members, benefiting microorganisms that do not effectively contribute to lignocellulose degradation (39).

In bacteria, the association of CBMs with different catalytic domains increments the functional diversity of CAZymes (40). The major proportion of CAZymes expressed by PM-06 exhibited modular architectures, which included one or more catalytic domains and CBMs. CBMs improve the catalytic activity of CAZymes by increasing enzyme

proximity to substrates, especially to insoluble ones (19, 40). In PM-06, the number of modular CAZymes and the architecture complexity increased at 120 h. CBMs could be enhancing the activity and increasing the penetration of enzymes into the NMP's more recalcitrant structures. Moreover, CBMs are also involved in the modulation of enzyme stability, function, and specificity (41). PM-06 expressed GH43, GH26, GH10, and GH5 at 0 and 120 h; however, these catalytic domains were associated with different CBMs, generating enzymes toward different substrates. These findings suggest that PM-06 is overcoming recalcitrance with a diversity of modular CAZymes operating in the proximity of insoluble substrates.

Microorganisms in PM-06 execute different but complementary metabolic tasks. *Paenibacillus* expressed modular enzymes with activity toward insoluble substrates; *Bacillus* expressed an LPMO with a CBM5 active on complex polysaccharides, such as crystalline cellulose and xylan (17), and CE that allow the access of enzymes into hemicellulose. *Microbacterium*, *Leifsonia*, and *Nocardia* expressed genes related to the hydrolysis of oligosaccharides. Moreover, the synergy between degraders and nondegraders is probably required to maintain microbial viability. Although the main focus of this study is lignocellulose degradation, the transcriptional activity of other abundant microorganisms in PM-06, such as *Aneurinibacillus*, could be related to the expression of genes with still unknown beneficial functions for the community. A metaproteomic study is needed to obtain more information about the proteins and enzymes produced by the consortium during degradation of NMP.

## MATERIALS AND METHODS

**PM-06 culture and sampling.** The consortium PM-06 was obtained by the dilution-to-stimulation approach (42) from the endogenous microbiota of the NMP. The initial culture (seed) was incubated for 7 days at 37°C with shaking at 125 rpm in flasks containing 25 mL of liquid medium (40 g/L NMP and 5 g/L yeast extract). At the end of the cultivation cycle (day 7), the consortium was inoculated in fresh medium under the same conditions. The metatranscriptomic analysis was carried out in the batch corresponding to the 10th transfer after evaluating its degrading stability and viability (data not shown). To obtain the genetic material, 5 mL of cell suspensions containing between 0.1 and 10 mg/mL cellular protein was taken directly from the fermentation broth and preserved in RNAprotect cell reagent (Qiagen) at −20°C before shipping. The samples were taken in triplicate (biological replicates) at the time of inoculation (0 h) and at 4, 24, and 120 h after inoculation.

**NMP degradation kinetics.** Degradation kinetics of NMP were performed as previously described (16). Differential degradation rates were calculated using the formula

$$\Delta W / \Delta t$$

where  $W$  is the weight of solids consumed in the lapse of time defined by  $t_2 - t_1$  ( $\Delta t$ ).

**RNA isolation and sequencing.** Preserved PM-06 cell suspensions were processed for total RNA isolation and DNase treatment. Reverse transcription optimized for low-input RNA using SMARTer stranded total RNA-seq kit v2 (TaKaRa Bio USA, Inc.) was performed using treated RNA from all samples. Adapters were ligated for sequencing on the Illumina HiSeq2500 platform at a depth of 50 million paired-end reads (25 million per side). All procedures were carried out by Admera Health (NJ, USA).

Once the raw reads were obtained, adapters used in the sequencing were removed as well as sequences belonging to *Zea mays* (host organism). Filtered reads were processed with the FastQC tool, and all reads with a Phred score of <25 were discarded. Additionally, low-complexity sequences were identified and removed with Prinseq (43) to ensure the initial quality before proceeding with the assembly.

**Metatranscriptomic data processing and annotation.** Cross-assembly of post-quality control (post-QC) reads was performed with metaSPAdes (44) with default settings. High-quality reads were *de novo* assembled in 14,257 contigs. Open reading frame (ORF) calling was performed using Prodigal 2.60 with the meta procedure (45). Protein sequences were locally aligned by DIAMOND (46) against the NCBI nr database for taxonomic and functional annotation.

Metabolic annotation of predicted proteins was conducted using the GhostKOALA web platform (47). For CAZyme annotation, the protein sequences were processed with the automated CAZy annotation database (dbCAN) (48) in the dbCAN2 meta server (<https://bcb.unl.edu/dbCAN2>) using all search tools provided (HMMER, DIAMOND, and Hotpep).

**Genome recruitment.** Genome coverage of PM-06 transcripts in each stage of the NMP degradation process were estimated by a recruitment analysis. For the above, 12 publicly available complete microbial genomes (Table 2) were selected according to the most abundant species obtained from the taxonomic classification of protein-coding sequences. Metatranscriptomic reads in each sample were mapped against the genomes using the BLAT algorithm (49). Genome coverage of recruited reads was expressed in terms of the number of reads recruited per kilobase of genome per gigabase of

**TABLE 2** Complete bacterial genomes selected to be recruited by metatranscriptomic reads

Organism	Genome size (bp)	Accession no. <sup>a</sup>
<i>Aneurinibacillus migulanus</i> DSM 2895	6,348,994	LGUG01000000
<i>Bacillus coagulans</i> S-lac	3,694,837	CP011939.1
<i>Bacillus thuringiensis</i> YBT-1518	6,002,284	CP005935.1
<i>Leifsonia xyli</i> SE134	3,596,761	CP014761.1
<i>Microbacterium</i> sp. No. 7	4,599,046	CP012697.1
<i>Microbacterium</i> sp. LCT-H2	3,356,231	MODW01000000
<i>Nocardia nova</i> ATCC 33727	7,553,834	PYHS01000000
<i>Nocardia nova</i> SH22a	8,348,532	NZ_CP006850
<i>Paenibacillus kribbensis</i> AM49	5,778,702	CP020028.1
<i>Paenibacillus macerans</i> 8224	7,331,450	JMQA01000000
<i>Paenibacillus macerans</i> FSL R5-0527	8,378,081	MRTR01000000
<i>Paenibacillus macerans</i> NCTC6355	7,390,297	UGSI01000000

<sup>a</sup>All the genomes used are publicly available in the NCBI databases (<https://www.ncbi.nlm.nih.gov>).

metatranscriptome (RPKG) by counting the matching transcripts (hits) with a length of  $\geq 50$  bp, an identity of  $\geq 95\%$ , and an E value of  $\leq 1E-5$ . Recruitment plots were generated to visualize genome coverage of strains (i.e., *Paenibacillus macerans* 8224 [Fig. S1 in the supplemental material]).

**Transcript abundance estimation and differential expression analysis.** All reads were mapped to the reference metatranscriptome and quantified using kallisto with default parameters (50). Raw counts were used as input for edgeR to perform pairwise statistical tests for differential expression between samples (51, 52). Transcripts that presented at least a  $\log_2$  fold change of 2 at a false-discovery rate (FDR)-corrected *P* value cutoff of 0.001 in any of the pairwise statistical tests were considered differentially expressed. To visualize the contribution of microbial members to the degradation process, transcripts associated with CAZyme families from highly abundant microbial genera along with families classified as GH, AA, and PL were retrieved from the set of differentially expressed transcripts. Families were manually annotated according to the substrate in which they act.

To understand the role of specific CAZymes associated with lignocellulose degradation, families reported as relevant for this process (13, 14, 53–57) were retrieved from the previously filtered set. Transcripts were grouped according to their families, summing their corresponding expression values and obtaining the total log CPM for each CAZyme family. Finally, a Z-score was computed for each family across samples, followed by hierarchical clustering and visualization using the ComplexHeatmap library from Bioconductor (58). A Z-score of  $\geq 0.4$  was used as a threshold value for upregulation within the set of CAZyme transcripts analyzed.

**Data availability.** The reference metatranscriptome and raw reads for each sample were deposited in the Sequence Read Archive (SRA) database under BioProject accession number [PRJNA735556](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA735556).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**SUPPLEMENTAL FILE 1**, XLSX file, 0.1 MB.

**SUPPLEMENTAL FILE 2**, PDF file, 0.7 MB.

## ACKNOWLEDGMENTS

This research was supported by CONACyT project CB 242952. J.G.S.-G. received a CONACyT doctoral fellowship. We thank Harinera de Yucatán and S.A. de C.V. for providing the nixtamalized maize pericarp. In addition, J.G.S.-G. thanks Francisco Rodríguez Valera for allowing a stay in the Evolutionary Genomics Group lab.

J.G.S.-G. developed the methodology, performed the formal analysis, and wrote, reviewed, and edited the paper. C.A.F.-E. developed the methodology and performed the formal analysis. P.J.C.-Y. performed the formal analysis, methodology, and data curation. F.H.C. contributed to data curation. R.A.R.-H. conceptualized the project and developed the methodology. M.N.S.-G. acquired funding, developed the methodology, conceptualized and supervised the project, performed the formal analysis, and reviewed and edited the paper.

## REFERENCES

- Gopinath KP, Sankaranarayanan AR, Nivedhitha L. 2016. Platform chemical biorefinery and agroindustrial waste management, p 379–391. In Brar SK, Sarma SJ, Pakshirajan K (ed), Platform chemical biorefinery. Elsevier, Amsterdam, the Netherlands.
- Marzo C, Diaz AB, Caro I, Blandino A. 2019. Valorization of agro-industrial wastes to produce hydrolytic enzymes by fungal solid-state fermentation. Waste Manag Res 37:149–156. <https://doi.org/10.1177/0734242X18798699>.

3. Castro-Muñoz R, Fila V, Durán-Páramo E. 2019. A review of the primary by-product (nejayote) of the nixtamalization during maize processing: potential reuses. *Waste Biomass Valorization* 10:13–22. <https://doi.org/10.1007/s12649-017-0029-4>.
4. Rostro M, Sánchez-González M, Rivas S, Moure A, Domínguez H, Parajó JC. 2014. Non-isothermal autohydrolysis of nixtamalized maize pericarp: production of nutraceutical extracts. *LWT* 58:550–556. <https://doi.org/10.1016/j.lwt.2014.03.043>.
5. de Lima Bossi MJ, Jimenez DJ, Cortes-Totalpa L, van Elsas JD. 2016. Soil-derived microbial consortia enriched with different plant biomass reveal distinct players acting in lignocellulose degradation. *Microb Ecol* 71: 616–627. <https://doi.org/10.1007/s00248-015-0683-7>.
6. Puentes-Tellez PE, Falcao Salles J. 2018. Construction of effective minimal active microbial consortia for lignocellulose degradation. *Microb Ecol* 76: 419–429. <https://doi.org/10.1007/s00248-017-1141-5>.
7. Lazuka A, Auer L, O'Donohue S, Nikolaus M, Ince B, Ince O. 2018. Anaerobic lignocellulolytic microbial consortium derived from termite gut: enrichment, lignocellulose degradation and community dynamics. *Biotechnol Biofuels* 11:284. <https://doi.org/10.1186/s13068-018-1282-x>.
8. Ozbayram EG, Kleinsteuber S, Nikolaus M, Ince B, Ince O. 2018. Enrichment of lignocellulose-degrading microbial communities from natural and engineered methanogenic environments. *Appl Microbiol Biotechnol* 102:1035–1043. <https://doi.org/10.1007/s00253-017-8632-7>.
9. Bohra V, Dafale NA, Purohit HJ. 2019. Understanding the alteration in rumen microbiome and CAZymes profile with diet and host through comparative metagenomic approach. *Arch Microbiol* 201:1385–1397. <https://doi.org/10.1007/s00203-019-01706-z>.
10. Hinsu AT, Tulsani NJ, Panchal KJ, Pandit RJ, Jyotsana B, Dafale NA, Patil NV, Purohit HJ, Joshi CG, Jakhesara SJ. 2021. Characterizing rumen microbiota and CAZyme profile of Indian dromedary camel (*Camelus dromedarius*) in response to different roughages. *Sci Rep* 11:9400. <https://doi.org/10.1038/s41598-021-88943-9>.
11. van Munster JM, Dobruchowska JM, Veloo R, Dijkhuizen L, van der Maarel MJ. 2015. Characterization of the starvation-induced chitinase CfcA and  $\alpha$ -1,3-glucanase AgnB of *Aspergillus niger*. *Appl Microbiol Biotechnol* 99: 2209–2223. <https://doi.org/10.1007/s00253-014-6062-3>.
12. Alessi AM, Bird SM, Oates NC, Li Y, Dowle AA, Novotny EH, deAzevedo ER, Bennett JP, Polikarpov I, Young JPW, McQueen-Mason SJ, Bruce NC. 2018. Defining functional diversity for lignocellulose degradation in a microbial community using multi-omics studies. *Biotechnol Biofuels* 11:166. <https://doi.org/10.1186/s13068-018-1164-2>.
13. Antunes LP, Martins LF, Pereira RV, Thomas AM, Barbosa D, Lemos LN, Silva GM, Moura LM, Epaminio GW, Digiamietri LA, Lombardi KC, Ramos PL, Quaggio RB, de Oliveira JC, Pascon RC, Cruz JB, da Silva AM, Setubal JC. 2016. Microbial community structure and dynamics in thermophilic composting viewed through metagenomics and metatranscriptomics. *Sci Rep* 6:38915. <https://doi.org/10.1038/srep38915>.
14. Jimenez DJ, Chaib De Mares M, Salles JF. 2018. Temporal expression dynamics of plant biomass-degrading enzymes by a synthetic bacterial consortium growing on sugarcane bagasse. *Front Microbiol* 9:299. <https://doi.org/10.3389/fmicb.2018.00299>.
15. Wei H, Wang L, Hassan M, Xie B. 2018. Succession of the functional microbial communities and the metabolic functions in maize straw composting process. *Bioresour Technol* 256:333–341. <https://doi.org/10.1016/j.biortech.2018.02.050>.
16. Serrano-Gamboa JG, Rojas-Herrera RA, Gonzalez-Burgos A, Folch-Mallol JL, Jimenez DJ, Sanchez-Gonzalez MN. 2019. Degradation profile of nixtamalized maize pericarp by the action of the microbial consortium PM-06. *AMB Express* 9:85. <https://doi.org/10.1186/s13568-019-0812-7>.
17. Limsakul P, Phitsuan P, Waeonukul R, Pason P, Tachaapaikoon C, Poomputsa K, Kosugi A, Sakka M, Sakka K, Ratanakhanokchai K. 2020. A novel AAI10 from *Paenibacillus curdlandolyticus* and its synergistic action on crystalline and complex polysaccharides. *Appl Microbiol Biotechnol* 104:7533–7550. <https://doi.org/10.1007/s00253-020-10758-x>.
18. Manjeet K, Madhuprakash J, Mormann M, Moerschbacher BM, Podile AR. 2019. A carbohydrate binding module-5 is essential for oxidative cleavage of chitin by a multi-modular lytic polysaccharide monooxygenase from *Bacillus thuringiensis* serovar *kurstaki*. *Int J Biol Macromol* 127: 649–656. <https://doi.org/10.1016/j.ijbiomac.2019.01.183>.
19. Cuskin F, Flint JE, Gloster TM, Morland C, Basle A, Henrissat B, Coutinho PM, Strazzulli A, Solovyova AS, Davies GJ, Gilbert HJ. 2012. How nature can exploit nonspecific catalytic and carbohydrate binding modules to create enzymatic specificity. *Proc Natl Acad Sci U S A* 109:20889–20894. <https://doi.org/10.1073/pnas.1212034109>.
20. Mewis K, Lenfant N, Lombard V, Henrissat B, Nojiri H. 2016. Dividing the large glycoside hydrolase family 43 into subfamilies: a motivation for detailed enzyme characterization. *Appl Environ Microbiol* 82:1686–1692. <https://doi.org/10.1128/AEM.03453-15>.
21. Razeq FM, Jurak E, Stogios PJ, Yan R, Tenkanen M, Kabel MA, Wang W, Master ER. 2018. A novel acetyl xylan esterase enabling complete deacetylation of substituted xylans. *Biotechnol Biofuels* 11:74. <https://doi.org/10.1186/s13068-018-1074-3>.
22. Biely P, Singh S, Puchart V. 2016. Towards enzymatic breakdown of complex plant xylan structures: state of the art. *Biotechnol Adv* 34:1260–1274. <https://doi.org/10.1016/j.biotechadv.2016.09.001>.
23. Comtet-Marre S, Parisot N, Lepercq P, Chaucheyras-Durand F, Mosoni P, Peyretilade E, Bayat AR, Shingfield KJ, Peyret P, Forano E. 2017. Metatranscriptomics reveals the active bacterial and eukaryotic fibrolytic communities in the rumen of dairy cow fed a mixed diet. *Front Microbiol* 8:67. <https://doi.org/10.3389/fmicb.2017.00067>.
24. Ghio S, Ontañón O, Piccinni FE, Marrero Díaz de Villegas R, Talia P, Grasso DH, Campos E. 2018. *Paenibacillus* sp. A59 GH10 and GH11 extracellular endoxylanases: application in biomass bioconversion. *Bioenerg Res* 11: 174–190. <https://doi.org/10.1007/s12155-017-9887-7>.
25. Sukharnikov LO, Alahuhta M, Brunecky R, Upadhyay A, Himmel ME, Lunin VV, Zhulin IB. 2012. Sequence, structure, and evolution of cellulases in glycoside hydrolase family 48. *J Biol Chem* 287:41068–41077. <https://doi.org/10.1074/jbc.M112.405720>.
26. Brunecky R, Alahuhta M, Sammond DW, Xu Q, Chen M, Wilson DB, Brady JW, Himmel ME, Bomble YJ, Lunin VV. 2017. Natural diversity of glycoside hydrolase family 48 exoglucanases: insights from structure. *Biotechnol Biofuels* 10:274. <https://doi.org/10.1186/s13068-017-0951-5>.
27. Cortes-Totalpa L, Salles JF, van Elsas JD. 2017. Bacterial synergism in lignocellulose biomass degradation—complementary roles of degraders as influenced by complexity of the carbon source. *Front Microbiol* 8:1628. <https://doi.org/10.3389/fmicb.2017.01628>.
28. Bertrand RL. 2019. Lag phase is a dynamic, organized, adaptive, and evolvable period that prepares bacteria for cell division. *J Bacteriol* 201: e00697-18. <https://doi.org/10.1128/JB.00697-18>.
29. Abot A, Arnal G, Auer L, Lazuka A, Labourdette D, Lamarre S, Trouilh L, Laville E, Lombard V, Potocki-Veronese G, Henrissat B, O'Donohue M, Hernandez-Raquet G, Dumon C, Leberre VA. 2016. CAZyChip: dynamic assessment of exploration of glycoside hydrolases in microbial ecosystems. *BMC Genomics* 17:671. <https://doi.org/10.1186/s12864-016-2988-4>.
30. Lambert G, Kussell E. 2014. Memory and fitness optimization of bacteria under fluctuating environments. *PLoS Genet* 10:e1004556. <https://doi.org/10.1371/journal.pgen.1004556>.
31. Grainger DC, Hurd D, Harrison M, Holdstock J, Busby SJ. 2005. Studies of the distribution of *Escherichia coli* cAMP-receptor protein and RNA polymerase along the *E. coli* chromosome. *Proc Natl Acad Sci U S A* 102: 17693–17698. <https://doi.org/10.1073/pnas.0506687102>.
32. Tollerson R, Il, Ibba M. 2020. Translational regulation of environmental adaptation in bacteria. *J Biol Chem* 295:10434–10445. <https://doi.org/10.1074/jbc.REV120.012742>.
33. Urrios A, Macia J, Manzoni R, Conde N, Bonforti A, de Nadal E, Posas F, Sole R. 2016. A synthetic multicellular memory device. *ACS Synth Biol* 5: 862–873. <https://doi.org/10.1021/acssynbio.5b00252>.
34. Wolf DM, Fontaine-Bodin L, Bischofs I, Price G, Keasling J, Arkin AP. 2008. Memory in microbes: quantifying history-dependent behavior in a bacterium. *PLoS One* 3:e1700. <https://doi.org/10.1371/journal.pone.0001700>.
35. Crystal JD, Glanzman DL. 2013. A biological perspective on memory. *Curr Biol* 23:R728–R731. <https://doi.org/10.1016/j.cub.2013.07.082>.
36. Rombouts JL, Mos G, Weissbrodt DG, Kleerebezem R, Van Loosdrecht MCM. 2019. The impact of mixtures of xylose and glucose on the microbial diversity and fermentative metabolism of sequencing-batch or continuous enrichment cultures. *FEMS Microbiol Ecol* 95:fiz112. <https://doi.org/10.1093/femsec/fiz112>.
37. Eibinger M, Ganner T, Bubner P, Rosker S, Kracher D, Haltrich D, Ludwig R, Plank H, Nidetzky B. 2014. Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. *J Biol Chem* 289:35929–35938. <https://doi.org/10.1074/jbc.M114.602227>.
38. Song B, Li B, Wang X, Shen W, Park S, Collings C, Feng A, Smith SJ, Walton JD, Ding SY. 2018. Real-time imaging reveals that lytic polysaccharide monooxygenase promotes cellulase activity by increasing cellulose accessibility. *Biotechnol Biofuels* 11:41. <https://doi.org/10.1186/s13068-018-1023-1>.



39. Jimenez DJ, Dini-Andreote F, DeAngelis KM, Singer SW, Salles JF, van Elsas JD. 2017. Ecological insights into the dynamics of plant biomass-degrading microbial consortia. *Trends Microbiol* 25:788–796. <https://doi.org/10.1016/j.tim.2017.05.012>.
40. Sidar A, Albuquerque ED, Voshol GP, Ram AFJ, Vijgenboom E, Punt PJ. 2020. Carbohydrate binding modules: diversity of domain architecture in amylases and cellulases from filamentous microorganisms. *Front Bioeng Biotechnol* 8:871. <https://doi.org/10.3389/fbioe.2020.00871>.
41. Walker JA, Takasuka TE, Deng K, Bianchetti CM, Udell HS, Prom BM, Kim H, Adams PD, Northen TR, Fox BG. 2015. Multifunctional cellulase catalysis targeted by fusion to different carbohydrate-binding modules. *Biotechnol Biofuels* 8:220. <https://doi.org/10.1186/s13068-015-0402-0>.
42. Lee DJ, Show KY, Wang A. 2013. Unconventional approaches to isolation and enrichment of functional microbial consortium—a review. *Bioresour Technol* 136:697–706. <https://doi.org/10.1016/j.biortech.2013.02.075>.
43. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
44. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>.
45. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
46. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12:59–60. <https://doi.org/10.1038/nmeth.3176>.
47. Kanehisa M, Sato Y, Morishima K. 2016. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* 428:726–731. <https://doi.org/10.1016/j.jmb.2015.11.006>.
48. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 40:W445–W451. <https://doi.org/10.1093/nar/gks479>.
49. Kent WJ. 2002. BLAT—the BLAST-like alignment tool. *Genome Res* 12: 656–664. <https://doi.org/10.1101/gr.229202>.
50. Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 34:525–527. <https://doi.org/10.1038/nbt.3519>.
51. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. <https://doi.org/10.1093/bioinformatics/btp616>.
52. Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11:R25. <https://doi.org/10.1186/gb-2010-11-3-r25>.
53. Berlemont R, Martiny AC. 2016. Glycoside hydrolases across environmental microbial communities. *PLoS Comput Biol* 12:e1005300. <https://doi.org/10.1371/journal.pcbi.1005300>.
54. Mello BL, Alessi AM, Riano-Pachon DM, deAzevedo ER, Guimaraes FEG, Espirito Santo MC, McQueen-Mason S, Bruce NC, Polikarpov I. 2017. Targeted metatranscriptomics of compost-derived consortia reveals a GH11 exerting an unusual exo-1,4- $\beta$ -xylanase activity. *Biotechnol Biofuels* 10: 254. <https://doi.org/10.1186/s13068-017-0944-4>.
55. Montella S, Ventorino V, Lombard V, Henrissat B, Pepe O, Faraco V. 2017. Discovery of genes coding for carbohydrate-active enzyme by metagenomic analysis of lignocellulosic biomasses. *Sci Rep* 7:42623. <https://doi.org/10.1038/srep42623>.
56. Nguyen STC, Freund HL, Kasanjian J, Berlemont R. 2018. Function, distribution, and annotation of characterized cellulases, xylanases, and chitinases from CAZy. *Appl Microbiol Biotechnol* 102:1629–1637. <https://doi.org/10.1007/s00253-018-8778-y>.
57. Ransom-Jones E, McCarthy AJ, Haldenby S, Doonan J, McDonald JE, Glass NL. 2017. Lignocellulose-degrading microbial communities in landfill sites represent a repository of unexplored biomass-degrading diversity. *mSphere* 2:e00300-17. <https://doi.org/10.1128/mSphere.00300-17>.
58. Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32:2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>.
59. Mutahir Z, Mekasha S, Loose JSM, Abbas F, Vaaje-Kolstad G, Eijsink VGH, Forsberg Z. 2018. Characterization and synergistic action of a tetra-modular lytic polysaccharide monooxygenase from *Bacillus cereus*. *FEBS Lett* 592:2562–2571. <https://doi.org/10.1002/1873-3468.13189>.
60. Pasari N, Adlakha N, Gupta M, Bashir Z, Rajacharya GH, Verma G, Munde M, Bhatnagar R, Yazdani SS. 2017. Impact of module-X2 and carbohydrate binding module-3 on the catalytic activity of associated glycoside hydrolases towards plant biomass. *Sci Rep* 7:3700. <https://doi.org/10.1038/s41598-017-03927-y>.
61. Venditto I, Najmudin S, Luis AS, Ferreira LM, Sakka K, Knox JP, Gilbert HJ, Fontes CM. 2015. Family 46 carbohydrate-binding modules contribute to the enzymatic hydrolysis of xyloglucan and  $\beta$ -1,3-1,4-glucans through distinct mechanisms. *J Biol Chem* 290:10572–10586. <https://doi.org/10.1074/jbc.M115.637827>.
62. Sakka M, Higashi Y, Kimura T, Ratanakhanokchai K, Sakka K. 2011. Characterization of *Paenibacillus curdlanolyticus* B-6 Xyn10D, a xylanase that contains a family 3 carbohydrate-binding module. *Appl Environ Microbiol* 77:4260–4263. <https://doi.org/10.1128/AEM.00226-11>.
63. Hagiwara Y, Mihara Y, Sakagami K, Sagara R, Bat-Erdene U, Yatsunami R, Nakamura S. 2020. Isolation of four xylanases capable of hydrolyzing corn fiber xylan from *Paenibacillus* sp. H2C. *Biosci Biotechnol Biochem* 84: 640–650. <https://doi.org/10.1080/09168451.2019.1693253>.
64. López-Mondéjar R, Zühlke D, Větrovský T, Becher D, Riedel K, Baldrian P. 2016. Decoding the complete arsenal for cellulose and hemicellulose deconstruction in the highly efficient cellulose decomposer *Paenibacillus* O199. *Biotechnol Biofuels* 9:104. <https://doi.org/10.1186/s13068-016-0518-x>.
65. Duan CJ, Huang MY, Pang H, Zhao J, Wu CX, Feng JX. 2017. Characterization of a novel theme C glycoside hydrolase family 9 cellulase and its CBM-chimeric enzymes. *Appl Microbiol Biotechnol* 101:5723–5737. <https://doi.org/10.1007/s00253-017-8320-7>.
66. Pluvinage B, Fillo A, Massel P, Boraston AB. 2017. Structural analysis of a family 81 glycoside hydrolase implicates its recognition of  $\beta$ -1,3-glucan quaternary structure. *Structure* 25:1348–1359.