Lignocellulose-Degrading Microbial Communities in Landfill Sites Represent a Repository of Unexplored Biomass-Degrading Diversity

Emma Ransom-Jones,[a] Alan J. McCarthy,[b] Sam Haldenby,[c] James Doonan,[a]
James E. McDonald[a]

School of Biological Sciences, Bangor University, Bangor, Gwynedd, United Kingdom[a]; Microbiology Research Group, Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom[b]; Centre for Genomics Research, Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom[c]

ABSTRACT The microbial conversion of lignocellulosic biomass for biofuel production represents a renewable alternative to fossil fuels. However, the discovery of new microbial enzymes with high activity is critical for improving biomass conversion processes. While attempts to identify superior lignocellulose-degrading enzymes have focused predominantly on the animal gut, biomass-degrading communities in landfill sites represent an unexplored resource of hydrolytic enzymes for biomass conversion. Here, to address the paucity of information on biomass-degrading microbial diversity beyond the gastrointestinal tract, cellulose (cotton) "baits" were incubated in landfill leachate microcosms to enrich the landfill cellulolytic microbial community for taxonomic and functional characterization. Metagenome and 16S rRNA gene amplicon sequencing demonstrated the dominance of *Firmicutes*, *Bacteroidetes*, *Spirochaetes*, and *Fibrobacteres* in the landfill cellulolytic community. Functional metagenome analysis revealed 8,371 carbohydrate active enzymes (CAZymes) belonging to 244 CAZyme families. In addition to observing biomass-degrading enzymes of anaerobic bacterial "cellulosome" systems of members of the *Firmicutes*, we report the first detection of the *Fibrobacter* cellulase system and the *Bacteroidetes* polysaccharide utilization locus (PUL) in landfill sites. These data provide evidence for the presence of multiple mechanisms of biomass degradation in the landfill microbiome and highlight the extraordinary functional diversity of landfill microorganisms as a rich source of biomass-degrading enzymes of potential biotechnological significance.

IMPORTANCE The microbial conversion of lignocellulosic biomass for biofuel production represents a renewable alternative to fossil fuels. However, the discovery of new microbial enzymes with high activity is critical for improving biomass conversion processes. While attempts to identify superior lignocellulose-degrading enzymes have focused predominantly on the animal gut, biomass-degrading communities in landfill sites represent an unexplored resource of hydrolytic enzymes for biomass conversion. Here, we identified *Firmicutes*, *Spirochaetes*, and *Fibrobacteres* as key phyla in the landfill cellulolytic community, detecting 8,371 carbohydrate active enzymes (CAZymes) that represent at least three of the recognized strategies for cellulose decomposition. These data highlight substantial hydrolytic enzyme diversity in landfill sites as a source of new enzymes for biomass conversion.

KEYWORDS *Bacteroidetes*, biomass, CAZymes, *Fibrobacter*, *Firmicutes*, landfill, cellulose degradation, cultivation, genomics, metagenomics, microbial ecology, *Spirochaetes*

Biomass conversion and biofuel production from the microbial decomposition of lignocellulosic substrates are an attractive and sustainable alternative to fossil fuels (1). However, progress in this area has been hampered by the recalcitrance of lignocellulose, which requires expensive pretreatment (2, 3). Recent attempts to identify superior lignocellulose-degrading enzymes and microorganisms have focused predominantly on anaerobic gut environments such as the bovine rumen (4), elephant gut (5), and termite gut (6), in addition to biogas reactors (7). These environments harbor microbial communities that have evolved to attack lignocellulosic biomass without the pretreatments currently used in commercial processes (8).

Anaerobic plant biomass conversion is best studied in the rumen, where *Fibrobacter succinogenes*, *Ruminococcus albus*, and *Ruminococcus flavefaciens* are the predominant cellulolytic bacterial species (9). Recent advances in DNA sequencing technology, particularly for metagenome sequencing of microbial communities, have transformed our ability to characterize unexplored biomass-degrading diversity in anoxic environments (4, 6, 7). For example, Hess and colleagues (2011) performed deep metagenomic sequencing of the rumen cellulolytic community, generating 446 draft genomes (including 15 genomes from uncultivated species), and identified 27,755 putative carbohydrate active enzymes (CAZymes) (4), revealing significant potential of rumen microorganisms for biomass conversion. In addition, metagenome and metaproteome inventories of the termite hindgut also revealed a variety of cellulolytic bacteria and CAZymes, many of which are related to those found in the rumen (6). More recently, metagenomic studies have utilized taxonomic binning to characterize the organisms involved in the anaerobic fermentation that occurs in biogas plants (10–12). However, there is still a paucity of information on the diversity and function of microorganisms in non-gut environments where cellulose hydrolysis occurs.

Landfill sites are highly heterogeneous, comprise mainly lignocellulosic material (13), and are therefore ideal environments for studying biomass conversion. In landfill, the decomposition of lignocellulosic biomass to methane is mediated by syntrophic interactions between hydrolytic (including cellulolytic) and fermentative bacteria, acetogenic bacteria, and methanogenic archaea (14). Members of the *Firmicutes* phylum associated with biomass conversion are consistently abundant in both culture-based (15, 16), and 16S rRNA gene (17–20) inventories of landfill sites, leading to the suggestion that they are the predominant degraders of biomass in landfill (15). However, members of the *Chlamydiae/Verrucomicrobia* group (17, 18), the *Cytophaga–Flexibacter–Bacteroides* group (17, 18), and phyla *Planctomycetes* (17), TM6 (19), *Chloroflexi* (19, 20), *Actinobacteria* (19), *Proteobacteria* (19, 20), *Lentisphaerae* (20), *Spirochaetes* (20), *Synergistetes* (20), *Thermotogae* (20), and *Fibrobacteres* (21) are also detected.

Cellulolytic clostridia (*Firmicutes* phylum), are more readily isolated from landfill sites and more amenable to PCR amplification, where they can represent as much as 100% of the sequencing output (13). However, it is clear that important cellulolytic functional groups in landfill sites have evaded detection by general 16S rRNA gene sequencing inventories. For example, novel *Fibrobacter* spp. (phylum *Fibrobacteres*, associated with cellulose hydrolysis in the rumen) were detected in landfill sites via the use of genus-specific 16S rRNA gene PCR primers (21), despite their absence from both 16S rRNA gene clone libraries (17, 18, 22) and studies utilizing 16S rRNA gene pyrosequencing (19, 20), with the exception of a single pyrosequencing study of landfill leachate (23). However, genus-specific quantitative PCR (qPCR) of landfill cDNA revealed that *Fibrobacter* spp. are abundant members of the landfill community, comprising ≤40% of the total 16S rRNA molecules in landfill (21). qPCR analysis of microbial DNA from heavily degraded cellulose (cotton) as studied here demonstrated that fibrobacters accounted for 29% of the total bacterial 16S rRNA gene copies, in comparison to members of the *Clostridia*, for which the highest relative abundance was that of *Clostridium* cluster III (17%) (24).

These taxa possess contrasting mechanisms for cellulose hydrolysis: *Firmicutes* utilize the cellulosomal method of cellulose decomposition (25), whereas *Fibrobacteres*

possess fibro-slime proteins and pili for biomass attachment, followed by secretion of hydrolytic enzymes (26). Members of the *Bacteroidetes* are present in landfill sites (17–20), but their function is unknown; however, rumen *Bacteroidetes* possessing polysaccharide utilization loci (PUL) have recently been implicated in cellulose hydrolysis (27), suggesting a potential role in landfill cellulose decomposition. Whole-community metagenome sequencing studies can therefore obviate the biases associated with PCR and enable the simultaneous assessment of bacterial, eukaryotic, archaeal, and viral diversity, in addition to assigning function and generating taxonomic bins, informing follow-up attempts to isolate novel taxa (28, 29), and enabling the reconstruction of genomes (4). Hess et al. (4) utilized ultradeep metagenomic sequencing of a single pooled sample from the cow rumen in order to maximize the opportunity for genome reconstruction of individual members of that community, which resulted in the assembly of 15 genomes of uncultured microbial species. A similar approach has also been utilized to reconstruct genomes belonging to members of the *Fibrobacteres* phylum from metagenomic data derived from the termite gut, anaerobic digesters, and the ovine rumen (30).

Here, to address our hypothesis that landfill sites represent a repository of unexplored biomass-degrading diversity, we utilized a combination of 16S rRNA gene amplicon sequencing and shotgun metagenomics with taxonomic binning of reads to characterize the taxonomic and functional diversity of hydrolytic microbial communities on cotton (cellulose) baits in landfill leachate microcosms. The aims of our study were to (i) utilize 16S rRNA gene amplicon sequencing of replicated raw leachate and cellulose enrichment microcosm samples to identify members of the landfill microbiome that are significantly enriched with cellulose amendment and (ii) to identify the functional diversity and taxonomic identity of the landfill biomass-degrading microbiome using deep metagenomic sequencing and taxonomic binning. This study provides the first descriptions of functional diversity of landfill biomass-degrading communities, demonstrating the significant potential of landfill sites for the provision of novel CAZymes of ecological and biotechnological significance.

RESULTS

Bacterial community composition of raw landfill leachate and cellulose enrichment microcosms. The community composition of DNA extracted from three raw leachate samples and nine replicate landfill leachate microcosms containing 1% (wt/vol) Avicel was determined using 16S rRNA gene amplicon sequencing on the Ion Torrent PGM platform. On average, the raw leachate samples contained 26 phyla, whereas the Avicel enrichment microcosms contained an average of 23 phyla (Fig. 1). At the phylum level, the raw leachate and enrichment microcosm samples contained similar taxa; however, read counts for members of the *Firmicutes* (38.0 to 46.4%; $P = 0.04$), *Bacteroidetes* (15.2 to 20.0%; $P = 0.06$), *Fibrobacteres* (0.2 to 0.8%; $P = 0.26$), and *Spirochaetes* (1.4 to 6.8%; $P = <0.01$) increased in the Avicel enrichment microcosms, while members of the *Proteobacteria* (28.4 to 13.2%; $P = <0.01$) decreased (Fig. 1). At the family level, the raw leachate samples contained on average 90 families, in contrast to the Avicel-enriched microcosm samples, which contained an average of 68 families. Within the Avicel enrichment microcosms, members of the *Ruminococcaceae* (2.5 to 11.8%, *Firmicutes* phylum), *Clostridiaceae* (11.9 to 14.1%, *Firmicutes* phylum), *Bacillaceae* (1.8 to 6.2%, *Firmicutes* phylum), *Fibrobacteraceae* (0.1 to 0.7%, *Fibrobacteres* phylum), and *Spirochaetaceae* (1.3 to 6.5%, *Spirochaetes* phylum) were enriched in comparison to the raw leachate sample (see Table S1 in the supplemental material).

Microbiome analysis of colonized cotton biofilms from landfill leachate microcosms. A dewaxed cotton string bait was incubated in a landfill leachate microcosm and retrieved for 16S rRNA gene amplicon and shotgun metagenome analysis after 6 weeks; visually, the cotton sample was heavily degraded and had little remaining structure when picked up with forceps. The colonized cotton biofilm, also analyzed by qPCR in a previous study (24), was subjected to DNA extraction and 454 pyrosequencing of 16S rRNA gene PCR amplicons (V1-to-V3 region) and shotgun metagenome
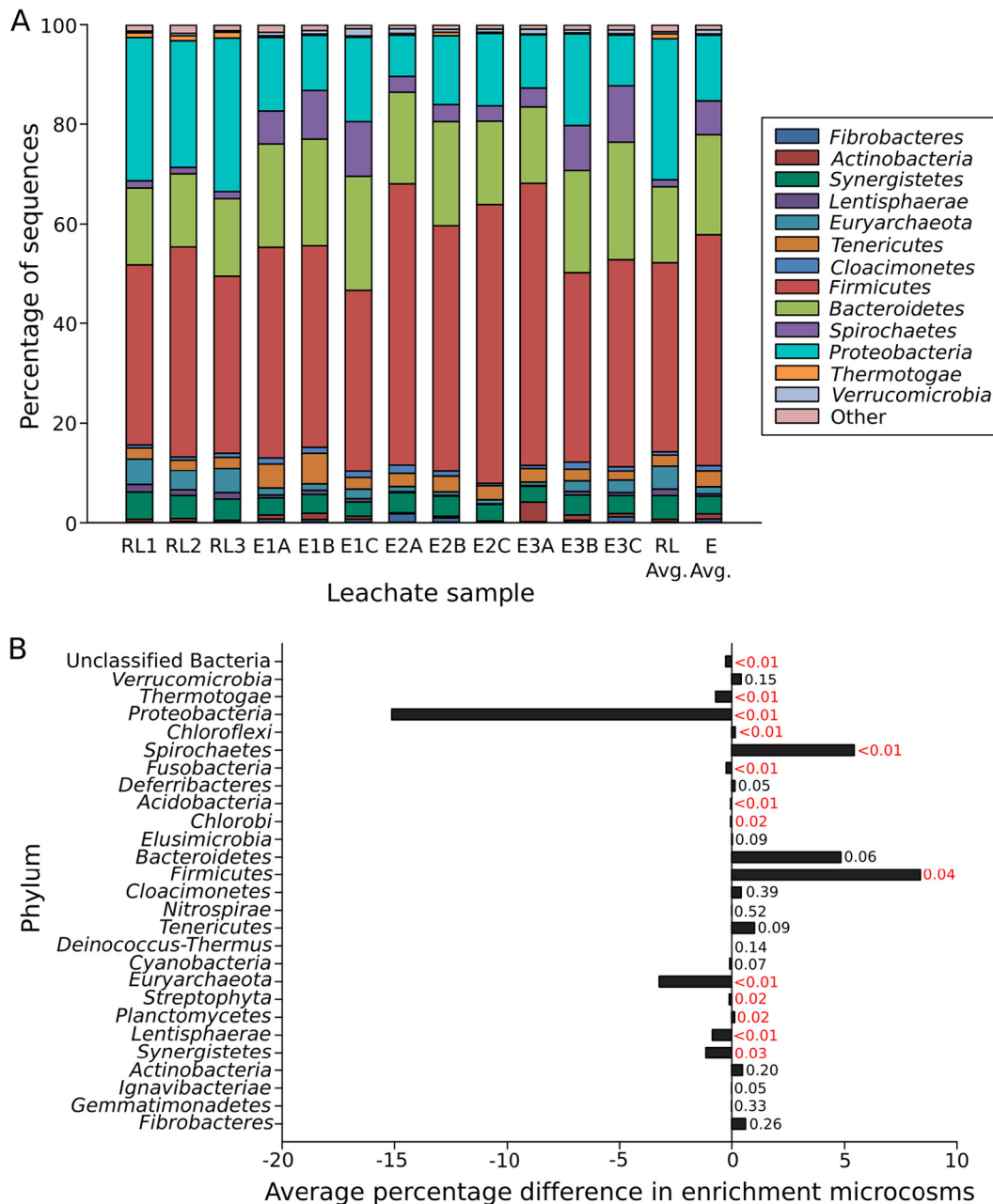
FIG 1 (A) Taxonomic identity at the phylum level of raw leachate (RL) and enrichment microcosms (E) as determined via Ion Torrent 16S rRNA gene amplicon sequencing and classification against the NCBI nucleotide database. The phyla shown are those that were >1% of the community. (B) Average percentage difference of phyla detected in Avicel enrichment microcosms in comparison to the raw leachate samples. Analysis of variance (ANOVA) was used to determine the significance between the samples, P values are shown on the bar plot; P values indicating a significant difference are shown in red.

sequencing. A total of 6,690 16S rRNA gene sequence reads were generated from the cotton biofilm and taxonomically assigned using the EzTaxon database. A rarefaction curve demonstrated that the majority of operational taxonomic units (OTUs) had been sampled (see Fig. S1 in the supplemental material), and the Shannon diversity index was 4.48. Nineteen phyla were detected in the 16S rRNA gene data set, with *Firmicutes* (37.4%), *Bacteroidetes* (20.9%), *Spirochaetes* (14.8%), and *Fibrobacteres* (14.2%) dominating the sequence reads (Fig. 2). These data are congruent with the major taxa associated with microbiome shifts in the replicated cellulose enrichment cultures described above. At the family level, *Ruminococcaceae* (24.1%, *Firmicutes* phylum), *Spirochaetaceae* (14.8%, *Spirochaetes* phylum), and *Fibrobacteraceae* (14.2%, *Fibrobac-*
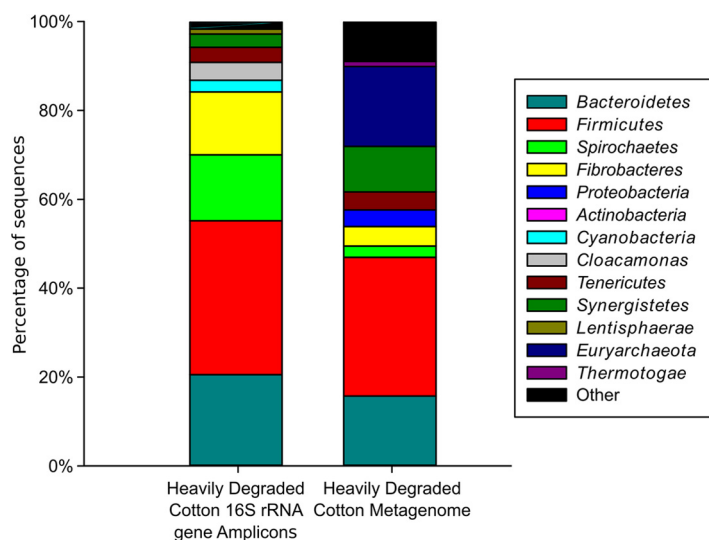
FIG 2 Phylum-level taxonomic distribution of the community of a heavily degraded cotton biofilm from landfill leachate microcosms as determined via 16S rRNA gene amplicon and metagenome sequencing. The phyla shown are those that were >1% of the community.

teres phylum) (see Table S2 in the supplemental material) were the dominant taxa. The taxonomy of metagenome contigs assembled via Ray Meta (31) from the heavily degraded cotton sample was determined via comparison using One Codex (32) and classified against the One Codex database. A total of 63 phyla were identified in the metagenome data set, with the predominant phyla determined as *Firmicutes* (31.2%), *Euryarchaeota* (18.0%), *Bacteroidetes* (15.7%), *Synergistetes* (10.2%), and *Fibrobacteres* (4.4%) (Fig. 2).

Identification of CAZymes in the landfill microbiome. CAZyme prediction on phylum-level binned metagenome contigs revealed 4,223 CAZymes in the *Bacteroidetes* bin, 3,385 in the *Firmicutes* bin, 604 in the *Spirochaetes* bin, 133 in the *Proteobacteria* bin, and 26 in the *Fibrobacteres* bin (Fig. 3). Hydrolytic enzyme systems are often modular in nature: for example, cellulosomes are multicomponent, multienzyme complexes found on the surface of cellulolytic bacteria and comprise a combination of catalytic enzymes, scaffold molecules, and carbohydrate binding molecules (carbohydrate binding modules [CBMs]) to maintain close contact with the substrate (25). The majority of CAZymes detected were glycoside hydrolases (GHs) (2,049 in the metagenome data set), glycosyl transferases (1,151), and CBMs (1,110), with auxiliary activity enzymes (113), carbohydrate esterases (634), cohesins (34), dockerins (85), polysaccharide lyases (107), and S-layer homology domains (320) also detected (Fig. 3). These enzymes included a number of GH families involved in lignocellulose degradation such as GH3, GH5, GH8, GH9, GH30, GH48, GH51, GH74, and GH94 (see Table S3 in the supplemental material), in addition to enzymes involved in the degradation of other polysaccharides (Table 1).

Spirochaetes **CAZymes.** Six CAZyme families containing cellulolytic enzymes (GH5, GH9, GH30, GH51, GH74, and GH94), five families containing hemicellulases (GH2, GH10, GH26, GH31, and GH43), and two cellulose-binding CBM families (CBM30 and CBM37) were detected in the *Spirochaetes* phylum bin (Table 1). Annotation of this bin via PROKKA also identified 5 TonB-dependent receptor (SusC) coding domains (Table 1), which are predominantly involved in the polysaccharide utilization locus (PUL) exhibited by members of the *Bacteroidetes* (33), although they have also been identified in *Spirochaetes* (34).

Firmicutes **CAZymes.** Eight CAZyme families associated with cellulose degradation (GH5, GH9, GH30, GH44, GH48, GH51, GH74, and GH94) and 10 families associated with hemicellulase activity (GH2, GH8, GH10, GH11, GH26, GH31, GH39, GH42, GH43, and

	Bacteroidetes	Fibrobacteres	Firmicutes	Proteobacteria	Spirochaetes
Auxiliary Activities	51	0	78	4	35
Carbohydrate Binding Modules	831	9	638	21	90
Carbohydrate Esterases	538	0	370	13	82
Cohesin	11	0	37	0	0
Dockerin	5	0	129	0	0
Glycoside Hydrolases	1665	13	938	46	286
Glycosyl Transferases	1006	2	584	44	105
Polysaccharide Lyases	115	2	28	3	6
S-layer Homology Domain	1	0	583	2	0

FIG 3 Number of carbohydrate active enzymes (CAZymes) detected in the *Bacteriodetes*, *Firmicutes*, *Spirochaetes*, *Proteobacteria*, and *Fibrobacteres* bins as determined via the dbCAN server with an E value of <1e–3.

GH53) were detected in the *Firmicutes* phylum bin, as well as eight CBM families associated with binding cellulose (CBM2, CBM3, CBM16, CBM30, CBM37, CBM44, CBM46, and CBM49) (Table 1). In addition, 37 cohesin and 129 dockerin coding domains were also detected (Table 1), which are required for the assembly of cellulosomes that are utilized by the majority of cellulolytic clostridia for hydrolysis (35).

***Fibrobacteres* CAZymes.** Four CAZyme families associated with cellulase activity (GH5, GH9, GH12, and GH45), three hemicellulases (GH8, GH26, and GH43), and a carbohydrate binding module 30 (CBM30) associated with cellulases were detected in the *Fibrobacteres* phylum bin (Table 1). Of particular interest is the detection of the GH45 cellulase exclusively in the landfill *Fibrobacteres* bin, as this CAZyme family has been found in all studied members of the *Fibrobacteres* and is thought to be distinctive to this group (30). In addition, 84 coding domains in the metagenome were identified as the *Fibrobacter succinogenes* major protein, a putative extracytoplasmic cellulose binding protein thought to be a cohesin analog (6, 30). All of these CAZyme families and associated proteins have been previously detected in the genomes of *Fibrobacteres* derived from cellulolytic environments of the rumen, termite gut, and anaerobic digesters and are now described for the first time in landfill site *Fibrobacteres*.

***Bacteroidetes* CAZymes.** Nine CAZyme families associated with cellulase activity (GH5, GH9, GH12, GH30, GH44, GH48, GH51, GH74, and GH94) and 10 associated with hemicellulase activity (GH2, GH8, GH10, GH11, GH26, GH31, GH39, GH42, GH43, and GH53) were detected in the *Bacteroidetes* phylum bin (Table 1). Nine CBM families associated with cellulases (CBM3, CBM8, CBM16, CBM30, CBM37, CBM44, CBM46, CBM59, and CBM69) were detected, in addition to 11 cohesin and 5 dockerin coding domains (Table 1). Annotation of the metagenome via PROKKA also revealed the presence of a substantial number of coding domains corresponding to SusD family proteins ($n = 334$) and TonB-dependent receptors (SusC) ($n = 511$ [Table 1]), which are involved in the polysaccharide utilization locus (PUL) (33) and were recently hypothesized as being part of a novel mechanism for cellulose decomposition within the *Bacteroidetes* (27), alongside GH5, GH9, and GH94, which were also detected in this

TABLE 1 Numbers of CAZymes and proteins associated with cellulose degradation identified in phylum-level bins constructed from the metagenome of a landfill leachate cellulose enrichment microcosm

CAZyme family	No. of CAZymes, proteins, or predicted genes				
	Bacteroidetes	*Fibrobacteres*	*Firmicutes*	*Proteobacteria*	*Spirochaetes*
Cellulases					
GH5	67	3	46	5	5
GH9	28	3	37		1
GH12	1	1			
GH30	13		6		1
GH44	5		2		
GH45		1			
GH48	1		3		
GH51	27		24		3
GH74	76		19		1
GH94	11		34		5
Hemicellulases					
GH2	108		36		17
GH8	15	2	6		
GH10	29		28		6
GH11	4		2		
GH26	30	2	21		1
GH31	30		26	1	10
GH39	4		8		
GH42	10		7	3	
GH43	184	2	42		13
GH53	16		5		
CBMs associated with cellulases					
CBM2			12		
CBM3	1		39		
CBM8	4				
CBM16	19		8	2	
CBM30	12	1	14		2
CBM37	7		2		8
CBM44	149		18		
CBM46	2		3		
CBM49			2		
CBM59	1				
CBM69	1				
Other cellulose-binding proteins					
Cohesin	11		37		
Dockerin	5		129		
SusC like (TonB-dependent receptor)	511			1	5
SusD like	334			1	
Total no. of predicted genes	62,632	204	82,534	4,945	19,832

study. To our knowledge, this is the first evidence for cellulose degradation by *Bacteroidetes* in landfill sites.

DISCUSSION

Previously, the composition of cellulolytic microbial communities in landfill sites had only been inferred through isolation studies (15), PCR inventories of specific taxa (21, 24, 34, 36), and 16S rRNA gene amplicon studies with "universal" primer sets (22). Here, we applied a "hook-bait" approach to enrich cellulolytic microorganisms from landfill leachate samples for taxonomic and functional analysis using 16S rRNA amplicon sequencing and shotgun metagenomics with taxonomic binning of reads. The use of a single metagenome sample to enable the reconstruction of genomes, as previously described by Hess et al. (4), resulted in a total of 371 individual genomes with low coverage due to the lack of available sequence data (see Table S4 in the supplemental material). However, this approach enabled binning of sequence reads from the phylum to species level and functional predictions of the role that key taxa play in cellulose hydrolysis in the landfill environment. This study therefore represents the first description of functional diversity in landfill biomass-degrading communities.

Here, a combined 16S rRNA gene amplicon and metagenome sequencing approach has demonstrated that *Firmicutes*, *Bacteroidetes*, *Spirochaetes*, and *Fibrobacteres* dominate the cellulolytic microbial community in landfill sites (Fig. 1). Raw leachate samples contained an average of 26 phyla, in comparison to the Avicel enrichment microcosms, which contained 23 phyla on average, with members of the *Firmicutes* (38.0 to 46.4%), *Bacteroidetes* (15.2 to 20.0%), *Fibrobacteres* (0.2 to 0.8%), and *Spirochaetes* (1.4 to 6.8%) enriched in the Avicel microcosms (Fig. 1). Members of the *Bacteroidetes* have previously been identified in landfill sites both via general bacterial 16S rRNA gene clone libraries (17, 18) and 454 pyrosequencing of 16S rRNA gene PCR amplicons (19, 20) and are known to occupy a variety of ecological niches, including activated sludge, decaying plant material, and compost (37). However, in addition to the decomposition of a range of polysaccharides, we provide the first detection of the major components of a *Bacteroidetes* cellulase system in landfill sites, suggesting a key role for *Bacteroidetes* in landfill cellulose decomposition.

Historically, *Firmicutes* have been considered the major degraders of cellulosic biomass in landfill sites (22), comprising 100% and 90% of 16S rRNA gene clones in libraries derived from solid cellulosic material and mixed cellulosic/leachate material, respectively, from a bioreactor treating landfill leachate (22). 454 pyrosequencing studies targeting the 16S rRNA gene have also detected both *Firmicutes* and, more specifically, *Clostridia* within an anaerobic bioreactor (19) and a lab-scale reactor treating landfill leachate (20), with *Clostridia* identified as the most abundant class within the *Firmicutes* (19, 20). Historically, anoxic environments are expected to contain large populations of clostridia, which are generally easier to isolate and cultivate than other obligate anaerobes of the *Bacteroidetes* and *Fibrobacteres*. Identification of members of the *Clostridia* as major components of landfill cellulolytic community supports previous qPCR analysis of this heavily degraded colonized cotton sample, where *Clostridium* clusters III, IV, and XIV totaled 21% of the 16S rRNA gene copies; however, it is significant that a greater proportion of 16S rRNA gene copies detected belonged to members of the genus *Fibrobacter* (29%) (24). Of the 19 recognized *Clostridium* clusters, four (I, III, IV, and XIVab) contain cellulolytic species (16), and it is likely that members of these clusters have played a role in the degradation of the cotton sampled here, with members of clusters III and IV most commonly identified in landfills (22, 24, 34, 38), in addition to cluster XIV (22, 24).

Metagenome analysis identified members of the *Firmicutes* (31.2%), *Euryarchaeota* (18.0%), *Bacteroidetes* (15.7%), *Synergistetes* (10.2%), and *Fibrobacteres* (4.4%) as the most abundant phyla in the cellulolytic biofilm (Fig. 2). While the presence of members of the *Firmicutes* and *Bacteroidetes* was largely consistent between the 16S rRNA gene amplicon and metagenome of the heavily degraded colonized cotton and the Ion Torrent sequence data, the distributions of other phyla differed between the two data sets (Fig. 2). Reads classified as *Spirochaetes* and *Fibrobacteres* were more prevalent in the 16S rRNA gene data set (14.8 and 14.2%, respectively) than the metagenome (2.6 and 4.4%, respectively) (Fig. 2) or Ion Torrent data (1.4 to 6.8% and 0.2 to 0.8% on average, respectively) (Fig. 1). The disparity between these results may be explained by the nature of the different sequencing approaches used and the underrepresentation of these phyla in genome databases. Despite the detection of *Fibrobacteres* in landfill sites via genus-specific 16S rRNA gene PCR primers (21, 24), they have remained undetected in this environment via either 16S rRNA gene clone libraries (17, 18, 22) or 454 pyrosequencing approaches (19, 20), resulting in a limited representation of the members of this phylum in sequence databases. Previous analysis of this data set against an earlier version of the One Codex database identified members of the *Fibrobacteres* as 0.1% of the total metagenome reads (data not shown). However, since this analysis was performed, Rahman et al. (30) utilized taxonomic binning to construct *Fibrobacter* genomes from metagenome data sets, resulting in the addition of seven new genomes to the One Codex database (*Fibrobacteria* bacterium genomes AD111, AD312, AD80, GUT221, GUT307, GUT31, and GUT77), and reanalysis of our metagenome data set increased the percentage of contigs assigned to the *Fibrobacteres* to 4.4%

(Fig. 2), demonstrating that the lower relative abundance of poorly studied members of cellulolytic communities could potentially be due to underrepresentation in sequence databases, rather than these organisms playing a limited role in this environment. This phenomenon highlights the importance of further studies to enhance the representation of these taxa in the public databases.

Despite this, *Fibrobacteres* CAZymes associated with cellulase and hemicellulase activity, carbohydrate binding, and the *Fibrobacter succinogenes* major protein associated with cellulose binding were detected in the landfill *Fibrobacteres* phylum bin, demonstrating marked similarity to the repertoire of enzymes and proteins associated with cellulolytic members of the *Fibrobacteres* studied in the rumen, termite gut, and anaerobic digesters (30). These data add to the growing body of evidence that cellulose hydrolysis is a unifying feature of the *Fibrobacteres* phylum (30) and extend the ecological range of detection of the *Fibrobacter* cellulase system to include landfill sites. Given the absence of *Fibrobacter* spp. in other landfill 16S rRNA gene inventories (17–20, 22), due to their apparent underrepresentation by general bacterial primers, their detection here as the fourth most abundant phylum on highly degraded cotton is significant (14.2% of 16S rRNA gene sequences) (Fig. 2) and supports the assertion that fibrobacters are prevalent members of the landfill hydrolytic community (24). This is supported by previous qPCR analysis of the same cotton biofilm analyzed in this study, which determined that fibrobacters represented 29% of the total bacterial 16S rRNA gene copies (24). The abundance of cellulolytic fibrobacters in the landfill community is intriguing and potentially important, as Gullert et al. (7) reported a decreased richness of lignocellulolytic enzymes in biogas fermenters (compared with gut environments) due to low abundances of *Bacteriodetes* and *Fibrobacteres* and suggested that increasing the proportion of these taxa could potentially enhance hydrolytic performance. The abundance of landfill fibrobacters in this study therefore suggests that there is the potential to enhance future biomass conversion processes by using landfill-derived fibrobacters and *Bacteroidetes* as inocula.

Spirochaetes were the third most dominant phylum as determined via 16S rRNA gene amplicon sequencing (14.8%), and they were also abundant in the metagenome (2.6%) and Ion Torrent data set (average of 1.4 to 6.8%). Members of the *Spirochaetes* have been identified in 16S rRNA gene clone libraries (17, 18) and 16S rRNA gene pyrosequencing inventories (19, 20) of landfill sites, but their function is currently unknown. Spirochetes have also been isolated from the bovine rumen (39), and although those strains were not cellulolytic, they are capable of utilizing polymers such as xylan, pectin, starch, and cellobiose and may act in a symbiotic manner with cellulolytic organisms in order to improve the hydrolysis of cellulose (25). Kudo et al. (40) tested the cellulolytic capabilities of two rumen bacteria, *F. succinogenes* and *Ruminococcus albus*, both in pure culture and in coculture with *Treponema bryantii*, a spirochete. When grown in coculture with *T. bryantii*, both strains showed an increase in barley straw degradation and volatile fatty acid production compared to the pure cultures, despite the fact that *T. bryantii* is not capable of degrading cellulose, suggesting a symbiotic relationship between the organisms. In addition, transmission electron microscopy of the colonized barley straw showed that *T. bryantii* was closely associated with both *F. succinogenes* and the cellulose fibers (40). The importance of *Spirochaetes* in the rumen environment, their detection both here and in previous studies, and their close association with other members of the microbial community and the degraded cotton suggest that they are important symbiotic members of the anaerobic cellulose-degrading community in landfill sites.

Developments in taxonomic binning of metagenome data sets have transformed our ability to assign functional attributes of mixed microbial communities to specific taxa (41). Here, we utilized Taxator-tk to generate phylum-level taxonomic bins containing metagenome contigs for gene annotation and CAZyme profiling of each phylum. Phylum-level bins derived from the metagenome data set revealed that the *Bacteroidetes* bin contained the most CAZymes (4,223), compared to the *Firmicutes* (3,385), *Spirochaetes* (604), *Proteobacteria* (133), and *Fibrobacteres* (26) (Fig. 3), despite

the fact that the *Firmicutes* were the dominant phylum within the metagenome (31.2% in comparison to 15.7% *Bacteroidetes*). While the detection of specific CAZyme groups and other elements of known cellulase systems in the phylum-level metagenome bins was largely congruent with the known composition of these systems described in other environments, it should be noted that for several reasons, including the potential lack of representation of landfill microbial genomes in the public databases, a proportion of sequences may have been incorrectly assigned to taxonomic bins. As discussed above, certain taxa are almost certainly underrepresented in the current publicly available databases—specifically those from landfill sites—and therefore it is likely that the CAZymes detected here represent only a fraction of the total present in these populations. This has been highlighted in this study, where the addition of eight additional *Fibrobacteres* genomes by Rahman et al. (30), increased the read-level composition of *Fibrobacteres* in the metagenome data set from 0.1 to 4.4%. This demonstrates the need for attempts to isolate and cultivate novel taxa from landfill sites and for the application of emerging technologies such as cell sorting and single-cell genomics, which could generate a step change in our knowledge of the landfill biomass-degrading community.

In total, 244 CAZyme families were identified in this study, including the families GH3, GH5, GH8, GH9, GH30, GH48, GH51, GH74, and GH94, which are associated with lignocellulose hydrolysis. The detection of these GH families is unsurprising given the high cellulosic content of landfill sites (13) and the fact that members of all of these GH families have been detected in similar studies on the bovine rumen (4), elephant gut (5), and a biogas reactor (7), and all but families GH30 and GH48 have also been identified in the hindgut of wood-feeding termites (6). In addition, cohesin and dockerin domains that mediate the assembly of cellulosomes in cellulolytic bacteria were also detected in the metagenome (Fig. 3), and the largest number of these genes was found in the *Firmicutes* bin; this is expected given that cohesins and dockerins are major components of cellulosomes that are exhibited by the majority of cellulolytic clostridia (42). Additionally, within the *Bacteroidetes* bin, both SusD family proteins and TonB-dependent receptors (SusC) were detected, which form part of the PUL that enables *Bacteroidetes* to degrade a variety of substrates (33). We have therefore demonstrated the presence of genes for at least three of the recognized microbial strategies for cellulose decomposition in the biosphere within landfill sites (the cellulosomal mechanism, the *Fibrobacteres* fibro-slime strategy, and *Bacteroidetes* PUL genes), highlighting the importance of landfill as an environment for the study of biomass decomposition.

Conclusions. Due to the recalcitrant nature of lignocellulosic substrates, understanding the diversity of microbial biomass conversion is a fundamental step toward unlocking their potential as a source for biofuel production. Recently, anaerobic environments such as the bovine rumen (4), elephant gut (5), termite gut (6), and a biogas reactor (7), where the microbial community has evolved to hydrolyze lignocellulose, have been identified as potential sources of novel enzymes. Here we established landfill sites as an unexplored and important source of novel hydrolytic diversity. We utilized a combination of molecular methods to characterize the cellulolytic biofilm of a heavily degraded cotton sample from a landfill leachate microcosm. These data demonstrated that members of the *Firmicutes*, *Bacteroidetes*, *Spirochaetes*, and *Fibrobacteres* are abundant in the landfill cellulolytic microbiome and possess an array of CAZymes that suggest an important role in the cellulose degradation that occurs in landfill sites. Additional Ion Torrent sequencing of 16S rRNA gene amplicons derived from raw leachate and Avicel enrichment micrososms also demonstrated enrichment of members of the *Firmicutes*, *Bacteroidetes*, *Fibrobacteres*, and *Spirochaetes* in the Avicel enrichment microcosms. This was further supported by metagenome sequencing of the heavily degraded cotton sample, which demonstrated the presence of members of the *Firmicutes* (31.2%), *Euryarchaeota* (18.0%), *Bacteroidetes* (15.7%), *Synergistetes* (10.2%), and *Fibrobacteres* (4.4%). Functional annotation of the total metagenome and phylum-level bins detected 244 CAZyme families, including members of families GH3,

GH5, GH8, GH9, GH30, GH48, GH51, GH74, and GH94, which are known to be involved in cellulose degradation. Here, we report the first detection of the *Fibrobacter* cellulase system and the *Bacteroidetes* polysaccharide utilization locus (PUL) in landfill sites, providing evidence for the presence of multiple mechanisms of biomass degradation in the landfill microbiome. These data highlight landfill sites as a repository of unexplored biomass-degrading enzyme diversity, with potential application in the effective breakdown of recalcitrant lignocellulosic plant biomass for alternative fuel production and biotechnological processes.

MATERIALS AND METHODS

Construction of landfill leachate microcosms containing dewaxed cotton string. The samples used in this study were obtained from microcosms previously described by McDonald et al. (24). Briefly, each microcosm was constructed in a sterile Nalgene carboy (10 liters) containing dewaxed cotton string (43) suspended in a nylon mesh bag. Microcosm 1 contained leachate from risers 3 and 4 of the Bromborough Dock landfill site (Wirral, United Kingdom), and microcosm 2 contained leachate from Bromborough Dock riser 5. The dewaxed cotton string was removed after 6 weeks of static incubation at ambient temperature and stored at −80°C prior to use as the source material for cultivation and DNA extraction.

Sampling of landfill leachate and construction of microcosms containing Avicel. Leachate samples were collected from the Hafod landfill site, Wrexham, United Kingdom. Three samples of landfill leachate were collected and transported to the laboratory, where they were stored at 4°C prior to analysis. A total of nine landfill leachate microcosms were established (three technical replicates for each of the three landfill leachate samples) in sterile 100-ml Nalgene bottles, each containing 1% (wt/vol) Avicel (Sigma). Each microcosm was inoculated with 100 ml of landfill leachate immediately after sampling and incubated for 2 weeks at 41°C.

DNA extraction of leachate and microcosms containing Avicel. Raw leachate samples were shaken gently, and 50 ml from each sample was removed and centrifuged at 5,000 rpm for 20 min. Subsequently, the supernatant was removed and DNA was extracted from the pellet using the PowerSoil DNA isolation kit (MoBio) according to the manufacturer's instructions. After a 2-week incubation with cellulose enrichment, the microcosms were shaken to ensure the Avicel was suspended and 25 ml from each sample was removed ($n = 9$) and centrifuged at 5,000 rpm for 20 min prior to the removal of the supernatant. DNA was extracted from the cell/biomass pellets using the PowerSoil DNA isolation kit (MoBio) according to the manufacturer's protocol.

DNA extraction of colonized cotton from landfill leachate microcosms and bacterial cultures. DNA was extracted from 1.5 ml of broth culture. Cell cultures were centrifuged at 14,000 rpm for 5 min, the supernatant was removed, and the cell pellet was resuspended in nuclease-free water (Bioline) to a final volume of 500 μl. DNA extraction from 0.5 g of colonized cotton was performed using the phenol-chloroform method of Griffiths et al. (44). DNA was visualized on a 1% agarose (Bioline) gel with HyperLadder 1kb (Bioline) as a marker. DNA concentrations were determined using the Qubit double-stranded DNA (dsDNA) BR assay kit (Life Technologies, Inc.) and a Qubit fluorometer (Life Technologies, Inc.).

Metagenome sequencing and analysis of heavily degraded colonized cotton. Total DNA extracted from the heavily degraded colonized cotton from microcosm 1 was utilized to generate three Nextera sequencing libraries with insert sizes of 300, 400, and 600 bp. DNA libraries were then sequenced on one lane of an Illumina HiSeq, generating paired-end libraries (2×100 bp), by the Centre for Genomic Research, Liverpool, United Kingdom. For the 300-, 400-, and 600-bp libraries, 135,007,994, 103,519,620 and 93,776,958 reads were obtained, respectively, representing a total of 84.6 Gbp of metagenome sequence data (see Table S5 in the supplemental material).

Adapter sequences were removed using Cutadapt (version 1.2.1) (45) and trimmed via Sickle (version 1.2) (46) with a minimum window quality score of 20 and reads shorter than 10 bp removed. The three sequence libraries were combined and assembled via Ray Meta (31) (version 2.3.1, k-mer = 31) using the HPC Wales computing network. Raw reads and assembled contigs were uploaded as separate data sets to One Codex and classified against the One Codex database (32).

The assembled metagenome contigs were subjected to taxonomic assignment at the phylum level using the Taxator-tk (version 1.2.2) script binning-blast (41). Gene prediction was performed on both the whole-metagenome data set and phylum-level bins belonging to the *Bacteroidetes*, *Fibrobacteres*, *Firmicutes*, *Proteobacteria*, and *Spirochaetes* via Prodigal v2.6 (47) and annotated via the dbCAN server (48) for the presence of CAZymes with a cutoff E value of <1e−3. Additional annotation was performed using Prokka v1.11 (49).

454 pyrosequencing and analysis of general bacterial 16S rRNA gene PCR amplicons generated from a heavily degraded cotton biofilm. DNA extracted from the heavily degraded cotton was subjected to PCR with barcoded general bacterial primers that targeted the V1-to-V3 region of the 16S rRNA gene designed by Chunlab, Inc. (Republic of Korea): forward primer B16S-F (5′→3′ sequence GAGTTTGATCMTGGCTCAG) and reverse primer B16 (5′→3′ sequence WTTACCGCGGCTGCTGG) by Chunlab, Inc., Republic of Korea. These PCR amplicons were then purified via the QIAquick PCR purification kit (Qiagen) before sequencing with the 454-GS FLX Titanium sequencing system by Chunlab, Inc.

The 16S rRNA gene sequences were processed to separate the samples via the barcodes before removal of the barcode, linker, and PCR primer sequences, quality filtering, and chimera detection and

removal and clustered into OTUs at 97% sequence similarity via Chunlab, Inc. Sequences were classified via CLcommunity against the ExTaxon database using the default parameters. Rarefaction analysis was performed and the Shannon diversity index was calculated using the Ribosomal Database Project pipeline.

Ion Torrent PGM sequencing of raw leachate and Avicel enrichment microcosms. DNA from the three raw leachate samples and nine microcosms was subjected to PCR with 16S rRNA gene PCR primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′) using the HotStar-Taq Plus master mix kit (Qiagen, USA). The PCR cycling conditions were 94°C for 3 min, followed by 30 cycles of 94°C for 30 s, 53°C for 40 s, and 72°C for 1 min, followed by a final elongation step at 72°C for 5 min. Sequencing was performed on an Ion Torrent PGM following the manufacturer's procedure, and data were processed to remove barcodes, primers, sequences of <150 bp, sequences with ambiguous base calls and with homopolymer runs exceeding 6 bp, and chimeras by MR DNA (Shallowater, TX). Operational taxonomic units (OTUs) were defined by clustering at 97% similarity and taxonomically classified using BLAST search against the NCBI nucleotide database. Sequences with homology to chloroplast 16S rRNA genes, plant nuclear and mitochondrial 18S rRNA genes, and arthropod 18S rRNA genes were removed from the data set. Analysis of variance (ANOVA) was used to determine the effect of enrichment with Avicel on the phyla present in leachate microcosms.

Accession number(s). The sequence data from this study have been deposited under the NCBI BioProject no. PRJNA351238. Metagenome assembly and phylum-level taxonomic bins are available at https://github.com/emmarj/Metagenome.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mSphere.00300-17.

FIG S1, PDF file, 0.1 MB.
TABLE S1, PDF file, 0.1 MB.
TABLE S2, PDF file, 0.05 MB.
TABLE S3, PDF file, 0.1 MB.
TABLE S4, PDF file, 0.1 MB.
TABLE S5, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

REFERENCES

1. Sanderson K. 2006. US biofuels: a field in ferment. Nature 444:673–676. https://doi.org/10.1038/444673a.

2. Agbor VB, Cicek N, Sparling R, Berlin A, Levin DB. 2011. Biomass pretreatment: fundamentals toward application. Biotechnol Adv 29:675–685. https://doi.org/10.1016/j.biotechadv.2011.05.005.

3. Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ. 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. Bioresour Technol 101:4851–4861. https://doi.org/10.1016/j.biortech.2009.11.093.

4. Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G, Luo S, Clark DS, Chen F, Zhang T, Mackie RI, Pennacchio LA, Tringe SG, Visel A, Woyke T, Wang Z, Rubin EM. 2011. Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. Science 331:463–467. https://doi.org/10.1126/science.1200387.

5. Ilmberger N, Güllert S, Dannenberg J, Rabausch U, Torres J, Wemheuer B, Alawi M, Poehlein A, Chow J, Turaev D, Rattei T, Schmeisser C, Salomon J, Olsen PB, Daniel R, Grundhoff A, Borchert MS, Streit WR. 2014. A comparative metagenome survey of the fecal microbiota of a breast- and a plant-fed Asian elephant reveals an unexpectedly high diversity of glycoside hydrolase family enzymes. PLoS One 9:e106707. https://doi.org/10.1371/journal.pone.0106707.

6. Warnecke F, Luginbühl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M, McHardy AC, Djordjevic G, Aboushadi N, Sorek R, Tringe SG, Podar M, Martin HG, Kunin V, Dalevi D, Madejska J, Kirton E, Platt D, Szeto E, Salamov A, Barry K, Mikhailova N, Kyrpides NC, Matson EG, Ottesen EA, Zhang X, Hernández M, Murillo C, Acosta LG, Rigoutsos I, Tamayo G, Green BD, Chang C, Rubin EM, Mathur EJ, Robertson DE, Hugenholtz P, Leadbetter JR. 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature 450:560–565. https://doi.org/10.1038/nature06269.

7. Güllert S, Fischer MA, Turaev D, Noebauer B, Ilmberger N, Wemheuer B, Alawi M, Rattei T, Daniel R, Schmitz RA, Grundhoff A, Streit WR. 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. https://doi.org/10.1186/s13068-016-0534-x.

8. Flint HJ. 1997. The rumen microbial ecosystem—some recent developments. Trends Microbiol 5:483–488. https://doi.org/10.1016/S0966-842X(97)01159-1.

9. Krause DO, Denman SE, Mackie RI, Morrison M, Rae AL, Attwood GT, McSweeney CS. 2003. Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics. FEMS Microbiol Rev 27:663–693. https://doi.org/10.1016/S0168-6445(03)00072-X.

10. Maus I, Koeck DE, Cibis KG, Hahnke S, Kim YS, Langer T, Kreubel J, Erhard M, Bremges A, Off S, Stolze Y, Jaenicke S, Goesmann A, Sczyrba A,

Scherer P, König H, Schwarz WH, Zverlov VV, Liebl W, Pühler A, Schlüter A, Klocke M. 2016. Unraveling the microbiome of a thermophilic biogas plant by metagenome and metatranscriptome analysis complemented by characterization of bacterial and archaeal isolates. Biotechnol Biofuels 9:171. https://doi.org/10.1186/s13068-016-0581-3.

11. Stolze Y, Bremges A, Rumming M, Henke C, Maus I, Pühler A, Sczyrba A, Schlüter A. 2016. Identification and genome reconstruction of abundant distinct taxa in microbiomes from one thermophilic and three mesophilic production-scale biogas plants. Biotechnol Biofuels 9:156. https://doi.org/10.1186/s13068-016-0565-3.

12. Stolze Y, Zakrzewski M, Maus I, Eikmeyer F, Jaenicke S, Rottmann N, Siebner C, Pühler A, Schlüter A. 2015. Comparative metagenomics of biogas-producing microbial communities from production-scale biogas plants operating under wet or dry fermentation conditions. Biotechnol Biofuels 8:14–32. https://doi.org/10.1186/s13068-014-0193-8.

13. Bookter TJ, Ham RK. 1982. Stabilization of solid-waste in landfills. J Environ Eng Div ASCE 108:1089–1100.

14. Barlaz M, Schaefer D, Ham R. 1989. Bacterial population development and chemical characteristics of refuse decomposition in a simulated sanitary landfill. Appl Environ Microbiol 55:55–65.

15. Westlake K, Archer DB, Boone DR. 1995. Diversity of cellulolytic bacteria in landfill. J Appl Bacteriol 79:73–78. https://doi.org/10.1111/j.1365-2672.1995.tb03126.x.

16. Benoit L, Cailliez C, Petitdemange E, Gitton J. 1992. Isolation of cellulolytic mesophilic clostridia from a municipal solid waste digester. Microb Ecol 23:117–125. https://doi.org/10.1007/BF00172634.

17. Huang LN, Zhou H, Zhu S, Qu LH. 2004. Phylogenetic diversity of bacteria in the leachate of a full-scale recirculating landfill. FEMS Microbiol Ecol 50:175–183. https://doi.org/10.1016/j.femsec.2004.06.008.

18. Huang LN, Zhu S, Zhou H, Qu LH. 2005. Molecular phylogenetic diversity of bacteria associated with the leachate of a closed municipal solid waste landfill. FEMS Microbiol Lett 242:297–303. https://doi.org/10.1016/j.femsle.2004.11.021.

19. Xie Z, Wang Z, Wang Q, Zhu C, Wu Z. 2014. An anaerobic dynamic membrane bioreactor (AnDMBR) for landfill leachate treatment: performance and microbial community identification. Bioresour Technol 161:29–39. https://doi.org/10.1016/j.biortech.2014.03.014.

20. Bareither CA, Wolfe GL, McMahon KD, Benson CH. 2013. Microbial diversity and dynamics during methane production from municipal solid waste. Waste Manag 33:1982–1992. https://doi.org/10.1016/j.wasman.2012.12.013.

21. McDonald JE, Lockhart RJ, Cox MJ, Allison HE, McCarthy AJ. 2008. Detection of novel Fibrobacter populations in landfill sites and determination of their relative abundance via quantitative PCR. Environ Microbiol 10:1310–1319. https://doi.org/10.1111/j.1462-2920.2007.01544.x.

22. Burrell PC, O'Sullivan C, Song H, Clarke WP, Blackall LL. 2004. Identification, detection, and spatial resolution of Clostridium populations responsible for cellulose degradation in a methanogenic landfill leachate bioreactor. Appl Environ Microbiol 70:2414–2419. https://doi.org/10.1128/AEM.70.4.2414-2419.2004.

23. Köchling T, Sanz JL, Gavazza S, Florencio L. 2015. Analysis of microbial community structure and composition in leachates from a young landfill by 454 pyrosequencing. Appl Microbiol Biotechnol 99:5657–5668. https://doi.org/10.1007/s00253-015-6409-4.

24. McDonald JE, Houghton JNI, Rooks DJ, Allison HE, McCarthy AJ. 2012. The microbial ecology of anaerobic cellulose degradation in municipal waste landfill sites: evidence of a role for fibrobacters. Environ Microbiol 14:1077–1087. https://doi.org/10.1111/j.1462-2920.2011.02688.x.

25. Leschine SB. 1995. Cellulose degradation in anaerobic environments. Annu Rev Microbiol 49:399–426. https://doi.org/10.1146/annurev.mi.49.100195.002151.

26. Burnet MC, Dohnalkova AC, Neumann AP, Lipton MS, Smith RD, Suen G, Callister SJ. 2015. Evaluating models of cellulose degradation by Fibrobacter succinogenes S85. PLoS One 10:e0143809. https://doi.org/10.1371/journal.pone.0143809.

27. Naas AE, Mackenzie AK, Mravec J, Schückel J, Willats WG, Eijsink VG, Pope PB. 2014. Do rumen Bacteroidetes utilize an alternative mechanism for cellulose degradation? mBio 5:e01401-14. https://doi.org/10.1128/mBio.01401-14.

28. Pope PB, Smith W, Denman SE, Tringe SG, Barry K, Hugenholtz P, McSweeney CS, McHardy AC, Morrison M. 2011. Isolation of Succinivibrionaceae implicated in low methane emissions from Tammar wallabies. Science 333:646–648. https://doi.org/10.1126/science.1205760.

29. Renesto P, Crapoulet N, Ogata H, La Scola B, Vestris G, Claverie JM, Raoult D. 2003. Genome-based design of a cell-free culture medium for Tropheryma whipplei. Lancet 362:447–449. https://doi.org/10.1016/S0140-6736(03)14071-8.

30. Rahman N, Parks D, Vanwonterghem I, Morrison M, Tyson G, Hugenholtz P. 2016. A phylogenomic analysis of the bacterial phylum Fibrobacteres. Front Microbiol 6:1469. https://doi.org/10.3389/fmicb.2015.01469.

31. Boisvert S, Raymond F, Godzaridis E, Laviolette F, Corbeil J. 2012. Ray Meta: scalable de novo metagenome assembly and profiling. Genome Biol 13:R122. https://doi.org/10.1186/gb-2012-13-12-r122.

32. Minot SS, Krumm N, Greenfield NB. 2015. One Codex: a sensitive and accurate data platform for genomic microbial identification. bioRxiv https://doi.org/10.1101/027607.

33. Terrapon N, Lombard V, Gilbert HJ, Henrissat B. 2015. Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. Bioinformatics 31:647–655. https://doi.org/10.1093/bioinformatics/btu716.

34. Louvel H, Saint Girons IS, Picardeau M. 2005. Isolation and characterization of FecA- and FeoB-mediated iron acquisition systems of the spirochete Leptospira biflexa by random insertional mutagenesis. J Bacteriol 187:3249–3254. https://doi.org/10.1128/JB.187.9.3249-3254.2005.

35. Lockhart RJ, Van Dyke MI, Beadle IR, Humphreys P, McCarthy AJ. 2006. Molecular biological detection of anaerobic gut fungi (Neocallimastigales) from landfill sites. Appl Environ Microbiol 72:5659–5661. https://doi.org/10.1128/AEM.01057-06.

36. Van Dyke MI, McCarthy AJ. 2002. Molecular biological detection and characterization of Clostridium populations in municipal landfill sites. Appl Environ Microbiol 68:2049–2053. https://doi.org/10.1128/AEM.68.4.2049-2053.2002.

37. Thomas F, Hehemann JH, Rebuffet E, Czjzek M, Michel G. 2011. Environmental and gut Bacteroidetes: the food connection. Front Microbiol 2:93. https://doi.org/10.3389/fmicb.2011.00093.

38. Li TL, Mazéas L, Sghir A, Leblon G, Bouchez T. 2009. Insights into networks of functional microbes catalysing methanization of cellulose under mesophilic conditions. Environ Microbiol 11:889–904. https://doi.org/10.1111/j.1462-2920.2008.01810.x.

39. Stanton TB, Canale-Parola E. 1979. Enumeration and selective isolation of rumen spirochetes. Appl Environ Microbiol 38:965–973.

40. Kudo H, Cheng KJ, Costerton JW. 1987. Interactions between Treponema bryantii and cellulolytic bacteria in the in vitro degradation of straw cellulose. Can J Microbiol 33:244–248. https://doi.org/10.1139/m87-041.

41. Dröge J, Gregor I, McHardy AC. 2015. Taxator-tk: precise taxonomic assignment of metagenomes by fast approximation of evolutionary neighborhoods. Bioinformatics 31:817–824. https://doi.org/10.1093/bioinformatics/btu745.

42. Schwarz WH. 2001. The cellulosome and cellulose degradation by anaerobic bacteria. Appl Microbiol Biotechnol 56:634–649. https://doi.org/10.1007/s002530100710.

43. Wood TM. 1988. Preparation of crystalline, amorphous, and dyed cellulase substrates. Methods Enzymol 160:19–25. https://doi.org/10.1016/0076-6879(88)60103-0.

44. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. 2000. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. Appl Environ Microbiol 66:5488–5491. https://doi.org/10.1128/AEM.66.12.5488-5491.2000.

45. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.

46. Joshi N. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. https://github.com/najoshi/sickle.

47. 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.

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–WW451. https://doi.org/10.1093/nar/gks479.

49. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.