2013

Transcriptomic and metatranscriptomic approaches to characterizing genes coding for fiber digestion within the rumen ecosystem

Wang, Pan

Lethbridge, Alta. : University of Lethbridge, Dept. of Biological Sciences

http://hdl.handle.net/10133/3459

Downloaded from University of Lethbridge Research Repository, OPUS
To my children, Eric and Alan
Abstract

The rumen microbiome constitutes a unique genetic resource of plant fiber degrading microbial enzymes that could be used for agricultural and industrial purposes. *Anaeromyces mucronatus* is a poorly characterized anaerobic lignocellulolytic fungus in the rumen. This thesis aimed at better understanding *A. mucronatus* YE505 and the particle associated rumen microbiota based on transcriptomic and metatranscriptomic approaches. High quality RNA was isolated from the fiber-associated rumen sample based on an improved RNA extraction method. A transcriptomic study was performed to investigate the expression of the fiber degrading system of *A. mucronatus* YE505, and the functional diversity of the fiber-associated eukaryotes from the rumen of muskoxen (*Ovibos moschatus*) was explored by a metatranscriptomic study. Much carbohydrate degradation related protein modules were detected. This study established effective approaches to characterizing the functional contents of rumen eukaryotic microbiome as well as rumen fungi, and identified several candidate genes that merit further investigation.
Acknowledgements

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, I would like to express my utmost gratitude to Dr. Tim McAllister and Dr. Brent Selinger for all their guidance and understanding, who not only took a chance on me and give me an incredible opportunity, but were most supportive to this project. Both Dr. McAllister and Dr. Selinger are always there to encourage and inspire me as I overcame all the obstacles in the completion this research work.

I would like to express my sincere acknowledgement to the members of my supervisory committee Dr. Robert Forster, Dr. Tony Russell and Dr. Steven Mosimann for their many suggestions and support over the years, and to my external examiner Dr. Mostafa Elshahed for his helpful comments and discussions.

I would like to thank the members of the McAllister and Selinger lab both past and present at the Department of Biological Sciences, University of Lethbridge, and Lethbridge Research Centre, Agriculture and Agri-Food Canada, for their kindness and support. I would like to thank Dr. Adrian Tsang and many dedicated collaborators at Centre for Structural and Functional Genomics, Concordia University and Genome Quebec Innovation Centre. I would like to thank the collaborators at Department of Biology and Wildlife, Institute of Arctic Biology, University of Alaska. Without their collaboration, the completion of this project would not have been possible.

Finally I would like to thank my husband, children and parents who have accompanied me along the long journey always offering support and love.
Table of contents

Approval page ................................................................. ii
Dedication ........................................................................ iii
Abstract ............................................................................. iv
Acknowledgements ........................................................... v
Table of contents .............................................................. vi
List of Tables ....................................................................... x
List of Figures ..................................................................... xi
List of Abbreviations ........................................................ xiii
Chapter 1  Literature Review .............................................. 1
  1.1  Introduction................................................................... 1
  1.2  The Rumen................................................................. 1
  1.3  The rumen ecosystem .................................................. 2
    1.3.1  Rumen microorganisms ........................................... 4
    1.3.1.1  Rumen bacteria .................................................. 4
    1.3.1.2  Rumen fungi ...................................................... 5
    1.3.1.3  Rumen protozoa ................................................. 7
    1.3.1.4  Other rumen microorganisms .............................. 8
  1.3.2  Microbe-microbe Interactions ...................................... 10
  1.4  Plant cell wall degradation by rumen microorganisms ...... 13
  1.4.1  Structure of plant cell walls ....................................... 14
  1.4.2  Carbohydrate active enzymes .................................... 16
  1.4.3  Cellulose biodegradation by microorganisms ............... 22
  1.5  Current advances in omics studies ............................... 24
    1.5.1  Next generation sequencing .................................... 24
    1.5.2  Transcriptomics, metatranscriptomics and RNA-Seq ..... 28
    1.5.3  Limitations and Challenges related to NGS and RNA-Seq ................................................................ 32
    1.5.4  Current advances of omics studies related to rumen fungi and microbiomes focusing on plant fiber degradation ................................................................ 33
  1.6  Research objectives ..................................................... 36
  1.7  Tables and Figures ....................................................... 38

Chapter 2  Isolation of high-quality total RNA from rumen anaerobic bacteria and fungi, and subsequent detection of glycoside hydrolases .................................................. 43
  2.1  Introduction................................................................... 43
  2.2  Methods ....................................................................... 45
2.2.1 Animals and rumen sampling ................................................................. 45
2.2.2 Total RNA isolation ................................................................................. 46
  2.2.2.1 Liquid phase rumen sample RNA isolation ........................................... 46
  2.2.2.2 Solid phase rumen sample RNA isolation .............................................. 47
  2.2.2.3 Effects of RNAProtect Bacteria Reagent on RNA isolation from solid phase rumen contents ................................................................. 48
2.2.3 Effects of RNA clean-up kits on RNA quality ......................................... 49
2.2.4 Evaluation of RNA quantity and quality .................................................. 49
2.2.5 Reverse transcription and PCR ............................................................... 49
2.3 Results ........................................................................................................... 50
  2.3.1 Optimization of the Liquid Phase RNA isolation method ......................... 50
  2.3.2 Total RNA isolation from rumen solids .................................................. 51
  2.3.3 Effect of RNAProtect bacteria reagent on RNA isolation from rumen solids .................................................................................................................. 53
  2.3.4 RT-PCR and detection of glycoside hydrolases ........................................ 53
2.4 Discussion ..................................................................................................... 54
2.5 Tables and Figures .......................................................................................... 58

Chapter 3 Snapshot of the eukaryotic gene expression in muskoxen rumen – a metatranscriptomic approach ................................................................. 65
3.1 Introduction.................................................................................................... 65
3.2 Materials and Methods ................................................................................ 68
  3.2.1 Ethics statement .................................................................................... 68
  3.2.2 Muskoxen rumen sampling .................................................................... 68
  3.2.3 RNA extraction and purification ............................................................ 69
  3.2.4 RNA sequencing and sequence assembly .............................................. 69
  3.2.5 Sequence analysis .................................................................................. 70
  3.2.6 rRNA sequence identification ............................................................... 70
  3.2.7 Binning ................................................................................................... 71
  3.2.8 Functional annotation of the coding RNA sequences ............................. 72
  3.2.9 Carbohydrate active protein annotation .............................................. 72
  3.2.10 Glycoside hydrolase cluster analysis ................................................... 73
  3.2.11 Sequence Data Availability ................................................................ 74
3.3 Results .......................................................................................................... 74
  3.3.1 General sequence statistics .................................................................... 74
  3.3.2 Analysis of sequencing coverage ............................................................ 76
3.3.3 Functional analysis of the putative protein-coding reads .................................. 77
3.3.4 Lignocellulolytic enzymes .................................................................................. 78
   3.3.4.1 Catalytic modules ......................................................................................... 79
   3.3.4.2 Accessory modules ..................................................................................... 80
   3.3.4.3 Lignocellulolytic gene expression ................................................................. 81
3.4 Discussion .................................................................................................................... 82
3.5 Tables and Figures ....................................................................................................... 88
Chapter 4 Comparative transcriptomic analysis of *Anaeromyces mucronatus* YE505... 124
4.1 Introduction .................................................................................................................... 124
4.2 Material and methods ................................................................................................... 126
   4.2.1 Chemicals .......................................................................................................... 126
   4.2.2 Fungal strain and culturing conditions ................................................................. 126
   4.2.3 Enzyme assays .................................................................................................... 127
   4.2.4 Isolation of total RNA ....................................................................................... 128
   4.2.5 Sequencing and sequence assembly ................................................................. 128
   4.2.6 Bioinformatic sequence analysis ........................................................................ 129
   4.2.7 Extracellular protein preparation .................................................................... 130
   4.2.8 LC-MS/MS and peptide identification ............................................................. 131
4.3 Results.......................................................................................................................... 131
   4.3.1 Growth of *Anaeromyces mucronatus* YE505 on different carbon sources and enzyme activity in extracellular culture fluid .................................................. 131
   4.3.2 Sequencing summary, assembly and BLAST analysis ...................................... 134
   4.3.3 Functional analyses of ORFs identified ............................................................ 135
   4.3.4 CAZymes prediction ......................................................................................... 137
      4.3.4.1 Catalytic modules ......................................................................................... 137
      4.3.4.2 Accessory modules ..................................................................................... 139
      4.3.4.3 Effect of carbon source on CAZyme gene transcription in *Anaeromyces mucronatus* YE505 ............................................................ 139
   4.3.5 Secretomic analysis of *Anaeromyces mucronatus* YE505 grown on different carbon sources ........................................................................................................ 142
   4.3.6 ORFs without predicted function ...................................................................... 142
4.4 Discussion .................................................................................................................... 143
4.5 Tables and Figures ....................................................................................................... 154
Chapter 5 Conclusions and future directions .................................................................. 170
5.1 Accomplishments ......................................................................................................... 170
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2 Future perspectives</td>
<td>174</td>
</tr>
<tr>
<td>5.3 Conclusions</td>
<td>177</td>
</tr>
<tr>
<td>References</td>
<td>179</td>
</tr>
<tr>
<td>Appendices</td>
<td>196</td>
</tr>
<tr>
<td>Supplementary Tables</td>
<td>196</td>
</tr>
</tbody>
</table>
List of Tables

Table 1.1 Currently classified anaerobic fungal species isolated from the gut of herbivores (Borneman and Akin, 1994; Breton et al., 1991; Ozkose et al., 2001; Nagpal et al., 2009). ............................................................38
Table 1.2 Comparison of next-generation sequencing platforms (Loman et al., 2012) (reprinted with permission). ................................................................................39
Table 2.1 Primers used for RT-PCR ..................................................................58
Table 2.2 Comparison of RNA yield and quality isolated from the rumen fluid (RF) by using different isolation procedures. ..............................................................59
Table 3.1 Feed composition. ........................................................................88
Table 3.2 Primers used for validating lignocellulolytic enzyme related contigs. ..........89
Table 3.3 Metabolic related KOG/COG groups represented by 5000 or more reads in the metatranscriptome from muskoxen rumen eukaryotes..............................................92
Table 3.4 List of putative carbohydrate esterases, polysaccharide lyases and carbohydrate binding related modules in the muskoxen rumen eukaryotic metatranscriptome (Muskoxen MT), and comparison of their abundance of selected CAZy modules with those of three other foregut microbiomes: wallaby (Macropod); bovine rumen (Bovine 2009 and Bovine 2011) and the hindgut of termite (Termite). .................................................94
Table 3.5 The abundance of contigs coding lignocellulolytic enzymes [glycoside hydrolases (GHs), carbohydrate esterases (CEs), polysaccharide lyases (PLs), carbohydrate-binding modules (CBMs), and other related modules] in the muskoxen eukaryotic metatranscriptome (Muskoxen MT) and a comparison of their abundance in the databases of rumen fungal (Ru. fungi) and rumen protozoal genes (Ru. prot.) as well as different anaerobic bacteria, including Bacteroides fragilis (Bfra), Butyrivibrio proteoclasticus (Bpro), Clostridium thermocellum (Cthe), Fibrobacter succinogenes (Fsuc), Prevotella ruminicola (Prum), Ruminococcus florofaciens (Rfla), and the rumen fungus Piromyces sp. E2 ESTs (Pir. ESTs). ..........................................................96
Table 3.6 CAZyme contigs (≥500 bp) identified with 10,000 or more reads in the metatranscriptome from the muskoxen rumen..........................................................107
Table 4.1 Matches of secretomic peptide detected by LC-MS/MS to predicted CAZy ORFs from RNA-Seq results when Aneromyces mucronatus YE505 grown on five different carbon sources. .................................................................154
Table 4.2 ORFs containing CBM10-like domain in association with predicted non-CAZy functional domains. .........................................................................................157
Table S.1 Muskoxen rumen metatranscriptome contigs (≥500 bp) that have one putative CAZY module .........................................................................................196
Table S.2 Muskoxen rumen metatranscriptome contigs (≥500 bp) that have two or more distinct putative CAZy modules. ..................................................................278
Table S.3 ORFs with predicted CAZy modules and their expression level represented by FPKMs for Aneromyces mucronatus grown on four different carbon sources and the clade they were grouped into as illustrated in Figure 4.12. ........................................303
List of Figures

Figure 1.1 Composition of lignocellulosic residues (Sánchez, 2009) (reprinted with permission) ................................................................. 41
Figure 1.2 Simplified schematic diagram illustrating plant cell wall components and enzymes involved in plant cell wall degradation. ........................................ 42
Figure 2.1 RNA isolation experimental flowchart ........................................... 60
Figure 2.2 Analysis of total RNA integrity extracted from liquid phase and solid phase of ruminal contents. ................................................................. 61
Figure 2.3 Comparison of column purification of RNA ........................................ 62
Figure 2.4 Effect of different RNAprotect treatments as described in section 2.2.2.3, on total RNA yield and quality from rumen solids (RS), and comparison with SRCl (Solid Ruminal Contents Isolation) results. ......................................................... 63
Figure 2.5 RT-PCR amplification of three Fibrobacter succinogenes S85 transcripts. 64
Figure 3.1 Summary statistics for the muskoxen rumen eukaryotic metatranscriptome. 108
Figure 3.2 Length distribution of muskoxen rumen metatranscriptome contigs. ........ 109
Figure 3.3 GC content analysis of the muskox rumen microbial community metatranscriptome. ........................................................................... 110
Figure 3.4 Phylogenetic distribution of muskoxen rumen metatranscriptome putative protein encoding reads (a total of 21.1 million) based on MEGAN analysis of top BLASTX hits against the NRMO database. The percentages of the major phylogenetic groups are indicated ........................................................................ 111
Figure 3.5 Top 30 phylogenetic bins of the muskoxen rumen metatranscriptome as determined by comparison against NCBI's non-redundant amino acid (nr) database. ... 112
Figure 3.6 Collector’s curve of richness as a function of reads analyzed. ...................... 113
Figure 3.7 Category distribution of the muskoxen rumen metatranscriptome as annotated using Eukaryotic Orthologous Groups (KOGs, for reads showing top BLASTX match to eukaryotic genes; solid bar) and clusters of orthologous groups (COGs, reads showing top BLASTX match to bacterial genes; dotted bar). ...................................................... 114
Figure 3.8 Top 30 KOG bins of the muskoxen rumen metatranscriptome as determined by comparison against KOG database .......................................................... 115
Figure 3.9 Schematic representation of plant cell wall polysaccharide and energy metabolism by the muskoxen rumen eukaryotic population. ........................................ 116
Figure 3.10 Comparison of the carbohydrate active enzymes identified from muskoxen rumen metatranscriptome (using all assembled contigs) with those of three other foregut metagenomes, the termite hindgut and five rumen/anaerobic microorganisms. ............ 117
Figure 3.11 Amino acid sequence similarities of carbohydrate active enzymes identified from the muskoxen rumen metatranscriptome (using all assembled contigs) versus top BLASTX match to the Genbank nr database. .......................................................... 119
Figure 3.12 Top 30 phylogenetic bins of the bacterial reads of muskoxen rumen metatranscriptome as determined by comparison against NCBI's non-redundant amino acid (nr) database ........................................................................... 120
Figure 3.13 Evidence of eukaryotic origin of the metatranscriptome sequences based on BLASTX searches. .................................................................122

Figure 3.14 Phylogenetic distribution of muskoxen rumen metatranscriptome putative carbohydrate active enzymes based on MEGAN analysis of top BLASTX hits of the contigs against the Genbank non-redundant amino acid (nr) database. .........................123

Figure 4.1 The total gas volume generated by Aneromyces mucronatus YE505 grown on different carbon sources. .................................................................158

Figure 4.2 The extracellular crude Avicel degrading enzyme activities of Aneromyces mucronatus YE505 grown on different carbon sources. ........................................159

Figure 4.3 The extracellular crude carboxymethylcellulose (CMC) degrading enzyme activities of Aneromyces mucronatus YE505 grown on different carbon sources. ..........160

Figure 4.4 The extracellular crude lichenan degrading enzyme activities of Aneromyces mucronatus YE505 grown on different carbon sources. ........................................161

Figure 4.5 The extracellular crude starch degrading enzyme activities of Aneromyces mucronatus YE505 grown on different carbon sources. ........................................162

Figure 4.6 The extracellular crude oat spelt xylan degrading enzyme activities of Aneromyces mucronatus YE505 grown on different carbon sources. .................................163

Figure 4.7 The extracellular crude p-nitrophenyl-β-d-cellobioside (pNPC) degrading enzyme activities of Aneromyces mucronatus YE505 grown on different carbon sources. ...............................................164

Figure 4.8 The extracellular crude p-nitrophenyl-β-d-glucoside (pNPG) degrading enzyme activities of Aneromyces mucronatus YE505 grown on different carbon sources. ........................................165

Figure 4.9 The extracellular crude carbohydrate esterase activities toward substrate α-Naphthyl acetate (α-NA) of Aneromyces mucronatus YE505 grown on different carbon sources. ........................................166

Figure 4.10 Size distribution of ORFs that were identified from Aneromyces mucronatus YE505. ..............................................................................................................167

Figure 4.11 ORFs with predicted KOG/COG functions and their total expression in four different carbon sources. .................................................................168

Figure 4.12 Comparison of expression level of predicted CAZymes from Aneromyces mucronatus YE505 grown on four different carbon sources. .................................169
List of Abbreviations

AGPC acid guanidinium-phenol-chloroform
BAC Bacterial Artificial Chromosome
bp Base pair
CAZy Carbohydrate active enzyme database
CAZymes Carbohydrate-active enzymes
CBM Carbohydrate binding module
CD Catalytic domain
cDNA complementary DNA
CE Carbohydrate Esterase
CelD_N cellulase N-terminal immuno-globulin domain
CMC carboxymethylcellulose
COG Cluster of Orthologous Groups
Da Dalton
DGGE Denaturing gradient gel electrophoresis
ECCF Extra-cellular cultural fluid
EF1 Elongation factor 1
EPS Extracellular polymeric substances
EST Expressed sequence tag
fn3 fibronectin-3
FPKM fragments per kilobase per million fragments mapped
Gbp Giga base pair
GCS Glucose-cellobiose-starch
GH Glycoside Hydrolase
GI tract Gastro-intestinal tract
GT Glycosyl Transferase
HGT Horizontal gene transfer
ITS Internal transcribed spacer
KOG euKaryotic Orthologous Groups
LC-MS/MS liquid chromatography coupled with tandem mass spectrometry
LRCI Liquid Ruminal Contents Isolation
LSU Ribosomal large subunit
Mbp Million base pair
mRNA Messenger RNA
MW Molecular weight
NCBI National Center for Biotechnology Information
ncRNA Non-coding RNA
NGS Next-generation sequencing
nr database Non-redundant amino acid database
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRMO database</td>
<td>Trimmed down non-redundant muskoxen amino acid database</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>OTU</td>
<td>Operational taxonomic unit</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PEP</td>
<td>phosphor enol pyruvate</td>
</tr>
<tr>
<td>PL</td>
<td>Polysaccharide Lyase</td>
</tr>
<tr>
<td>pNPC</td>
<td>$p$-nitrophenyl-$\beta$-d-cellobioside</td>
</tr>
<tr>
<td>pNPG</td>
<td>$p$-nitrophenyl-$\beta$-d-glucoside</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative PCR</td>
</tr>
<tr>
<td>RF</td>
<td>rumen fluid</td>
</tr>
<tr>
<td>RIN</td>
<td>RNA integrity number</td>
</tr>
<tr>
<td>RNA-Seq</td>
<td>RNA sequencing</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>RS</td>
<td>Rumen solids</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>quantitative reverse transcriptase-PCR</td>
</tr>
<tr>
<td>SLH</td>
<td>S-layer homolog domain</td>
</tr>
<tr>
<td>SRCI</td>
<td>Solid Ruminal Contents Isolation</td>
</tr>
<tr>
<td>SSU</td>
<td>Ribosomal small subunit</td>
</tr>
<tr>
<td>T-RFLP</td>
<td>Terminal restriction fragment length polymorphism</td>
</tr>
<tr>
<td>TGS</td>
<td>Third-generation sequencing</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
<tr>
<td>U</td>
<td>Unit</td>
</tr>
<tr>
<td>UPGMA</td>
<td>Unweighted Pair Group Method with Arithmetic Mean</td>
</tr>
<tr>
<td>VFAs</td>
<td>Volatile fatty acids</td>
</tr>
<tr>
<td>$\alpha$-NA</td>
<td>$\alpha$-Naphthyl acetate</td>
</tr>
</tbody>
</table>
Chapter 1 Literature Review

1.1 Introduction

Ruminants are unique in terms of their ability to use high fiber feeds. These animals (e.g., cattle, sheep, goats, deer) are able to obtain nutrients through the transformation of complex polysaccharides in plant cell walls into microbial protein and volatile fatty acids (VFAs) to support their own growth (Russell and Rychlik, 2001). However, ruminants themselves do not produce the enzymes required for the degradation of complex plant cell wall polysaccharides. Rather, they have developed a symbiotic relationship with a wide range of anaerobic microorganisms including bacteria, fungi and protozoa that colonize the digestive tract within a few weeks after birth. These microorganisms ferment plant carbohydrates ingested by the ruminant host yielding VFAs, vitamins and microbial protein as end products. VFAs are in turn used by the ruminant as an energy source (Wallace, 1994). As the microbial populations are able to produce vitamins and serve as a source of protein for the host, ruminants can adapt to nutrient-sparse environments. This property has led to their near global distribution, occupying environments ranging from the equator to the high arctic.

1.2 The Rumen

Ruminants are called foregut fermenters. The uniqueness of the ruminant digestive tract in anatomy is the stomach, which is composed of four compartments: reticulum, rumen, omasum and abomasum. Compared to some other monogastric herbivores, which are hindgut fermenters (e.g., rabbits, horses), fermentation in ruminants occurs in the forestomach comprised of the reticulum and rumen. The rumen is the major site of feed
digestion in the digestive tract, accounting for approximately 70% of the total digestive tract volume (Hobson, 1997). Its capacity varies greatly in adult ruminants, ranging from about 10 L in sheep to about 200 L in cattle, and the range is even greater if wild ruminants such as moose and mouse deer are considered. Forestomach fermentation offers several nutritional advantages over hindgut fermentation, as it allows the opportunity for fermentation end products to be digested and absorbed within the host’s lower digestive tract. Fermentation products that are not absorbed through the rumen wall as well as undigested feeds flow from the rumen to the omasum, where omasal leaves provide a large surface area for efficient water and mineral absorption. The omasum also acts as a muscular pump, moving the digesta from the reticulorumen to the abomasum. The abomasum functions as a ‘true’ stomach, producing enzymes and hydrochloric acid which hydrolyses proteins in a manner similar to the monogastric stomach. The high quality microbial protein derived in the reticulo-rumen from low quality recalcitrant plant sources can be digested efficiently in the abomasum; thereby, meeting a large proportion of the protein requirements of the host (Wallace, 1994). Rumination enables repeated mastication of feed, which enhances the ability of the microbial population to ferment it.

1.3 The rumen ecosystem

The rumen maintains a fastidious anaerobic environment, with relatively constant pH and temperature and mixing of microbes with substrate through rumen contractions. It represents the most active fibrolytic fermentation system currently known (Selinger et al., 1996). A diverse population of obligate anaerobic microorganisms exists in the rumen. Bacteria dominate this ecosystem in number and can reach levels as high as $10^{11}$ cells per
mL. In contrast, protozoa (10^5 to 10^7 per mL), anaerobic fungi (zoospores up to 10^5 per mL) and archaea (0.3 - 3% of the biomass) are less abundant (Li and Heath, 1992). These microorganisms actively interact with each other to form a symbiotic community.

The rumen microbiota is stable, but also dynamic in nature. Temperature in the rumen is held almost constant at about 39 °C with the pH typically ranging between 5.2 and 6.8 (Flint, 1997), but it can decline below 5.0 when cattle are fed high-starch diets (Nagaraja and Titgemeyer, 2007). Strictly anaerobic conditions are required for efficient microbial fermentation in the rumen. Trace amounts of oxygen may enter the rumen either with ingested feeds or via diffusion across the rumen wall from the blood stream, but it is quickly consumed by facultative anaerobic bacteria residing on the rumen epithelium.

Overall, rather than geographical location or even species of ruminant, diet has been found to be the main factor to influence the types and numbers of the predominant rumen microbes in adult ruminants (Hobson, 1997; Stewart et al., 1997). The microbial population changes considerably with changes in diet composition, as well as with the quantity and frequency of consumption. For example, nutrient composition, texture of diet and the presence of additives such as plant secondary metabolites and essential oils can affect the distribution of microbial species within the rumen, and overall digestive activity of the ruminal microbiota (McAllister and Newbold, 2008; McGinn et al., 2004).

Microbial species that occupy the rumen may also be isolated from other environments including in the caeca and large intestines of non-ruminant herbivores and omnivores, and the digestive tract of some insects such as termites. Certain species also
exist in the soil microflora and contribute to the anaerobic decomposition of plant debris (Hobson, 1997).

In a normally functioning rumen, proteins and polymeric carbohydrates, which usually make up the largest part of incoming feed, are fermented by the microbiota to VFAs, ammonia, carbon dioxide and hydrogen. The hydrogen is utilized by methanogens to reduce carbon dioxide to methane. The VFAs are absorbed across the rumen wall and serve as major carbon and energy sources for the host. A portion of the VFAs, undigested feed components, and microbial cells pass from the rumen and enter the lower digestive tract where they can also be absorbed or if undigested, excreted in the feces.

1.3.1 Rumen microorganisms

Currently over a thousand microbial species or operational taxonomic units (OTUs) have been identified in the rumen (Hess et al., 2011), and it has been estimated that only a fraction of these (less than 10%) have been cultivated in the laboratory (Flint et al., 2008; Kim et al., 2011). Consequently, the majority of microorganisms in the rumen have been identified strictly based on molecular techniques. Our present knowledge of the microbial community is primarily based on information gained from the culture of only a small portion of the microbial species present in this unique environment.

1.3.1.1 Rumen bacteria

The rumen bacteria account for the largest portion of microbial biomass in the rumen, exhibit the richest species diversity and are responsible for the majority of ruminal feed degradation (Stewart et al., 1997). The majority of ruminal bacteria are Gram-negative, obligate anaerobes. The rumen bacteria are roughly divided into groups
based on digestive activity or preference for feed components and include species with cellulolytic, hemicellulolytic, amylolytic, pectinolytic, proteolytic, ammonia-producing, sugar-utilizing, acid-utilizing and lipid-utilizing activities (Kamra, 2005).

Microorganisms within the rumen have evolved the capability to efficiently utilize plant cell wall fiber through the synergistic activities of the microbial enzymes. Rumen bacteria and their abilities to degrade plant fiber has been studied for over a century (Hungate, 1947). Fibrolytic bacteria such as *Fibrobacter succinogenes, Ruminococcus flavefaciens*, and *Ruminococcus albus* have received much attention during the past several decades, owing to their strong capability to utilize plant fiber and the comparative easiness whereby they can be cultured in the laboratory. Many fiber degrading enzymes have been isolated from these bacteria and their catalytic activities characterized in detail (Krause et al., 2003).

### 1.3.1.2 Rumen fungi

In contrast to bacteria, the anaerobic fungi were relatively recently discovered in the rumen (Orpin, 1975) in the mid-1970’s, even though the flagellated zoospore stage of their life cycle was identified as early as in 1910 and mistakenly classified as a flagellated protozoa. Based on mycelium/zoospore morphological characteristics, as well as molecular markers such as internal transcribed spacer (ITS) sequences, the rumen anaerobic fungi have been grouped into six genera, including polycentric fungi (*Anaeromyces, Cyllamyces* and *Orpinomyces*) and monocentric fungi (*Caecomyces, Neocallimastix* and *Piromyces*) (Kittelmann et al., 2012; Li and Heath, 1992; Liggenstoffer et al., 2010). All species are placed currently in the family of Neocallimataigales, which is the sole family in the newly erected phylum
Neocallimastigomycota (Table 1.1). Rumen fungi have a life cycle and morphology that is typical of the chytridiomycetes and possess chitin in their cell wall (Orpin and Joblin, 1997), but unlike typical chytrids, they are strict anaerobes.

Anaerobic rumen fungi can ferment a variety of plant cell wall polysaccharides to a number of fermentation end products including formate, acetate, lactate, carbon dioxide and hydrogen. The genetic details of the metabolic pathways that they employ in fermentation are largely unknown. However, instead of the mitochondria that are present in the cells of aerobic organisms, rumen fungi possess hydrogenosomes, a membrane bound mitochondria-like organelle that generates ATP and hydrogen (Akhmanova et al., 1999; Boxma et al., 2004). The hydrogen generated within the hydrogenosomes supports the formation of a symbiotic relationship between anaerobic fungi and methanogens (Boxma et al., 2005).

When ruminants are fed fiber rich forage diets, rumen fungi account for about 8 - 20% of the microbial biomass within the rumen (Orpin and Joblin, 1997; Rezaeian et al., 2004). The fungal mycelia penetrate plant tissue as a result of their filamentous growth. Consequently, anaerobic fungi in the rumen as well as in other regions of the gastrointestinal tract (GI tract) of herbivores are believed to play an active role in fiber degradation (Krause et al., 2003). Recently this group has attracted more research attention due to their distinct characteristics and potential to serve as a source of active fibrolytic enzymes for commercial purposes (Wang et al., 2011). Unlike rumen bacteria, they have the unique capacity to penetrate the cuticular surface and the lignified tissues of plant cell walls, and digest the fiber found in recalcitrant forages such as cereal straw (Orpin and Joblin, 1997). Fungi have been shown to produce a wide range of highly
active plant cell-wall degrading enzymes and are most abundant in the rumen of ruminants fed recalcitrant high-fiber diets (Orpin and Joblin, 1997). Genomic information on rumen fungi is still very limited, mainly due to difficulties in analyzing the extreme AT rich (80 – 85% mol%) coding and non-coding regions of their genome (Brownlee, 1989; Chen et al., 2006; Nicholson et al., 2005).

Some researchers believe the rumen fungi play a pivotal role as the initial/primary colonizers of plant fiber in the rumen (Joblin et al., 2002; Orpin and Joblin, 1997), while others consider their role in the process to be negligible owing to their low population density (Orpin and Joblin, 1997; Tuckwell et al., 2005). Increases in fungal biomass in the rumen of hosts fed poor quality high-fiber diets suggest that fungi may play a prominent role in feed digestion under these conditions (Orpin and Joblin, 1997).

1.3.1.3 Rumen protozoa

Based on cell counts, the amount of protozoa in the rumen is relatively low (10⁵ to 10⁷ per ml), but due to their large size these unicellular eukaryotes can account for up to 40% of rumen microbial biomass (Flint, 1997). Protozoa were detected in domestic ruminants as early as the nineteenth century (Kamra, 2005). Over 100 species of rumen protozoa have been identified, representing over 25 genera. Based on their morphology, ciliate protozoa have been classified into two groups, i.e., holotrichs and entodiniomorphids. Among the holotrichs, *Isotricha, Dasytricha, Buetschlia* and *Charonina* are widely distributed in the rumen and GI tract of non-rumen herbivores (Williams and Coleman, 1997).

The ciliate protozoa are generally considered as predators within the rumen ecosystem, preying on bacterial cells and fungal zoospores; thereby contributing to
nitrogen recycling in the rumen. Some are able to digest starch, pectin, soluble sugars and other plant particles as energy sources (Williams and Coleman, 1997). Enzymes responsible for cellulose and hemicellulose degradation have also been reported and it has been estimated that protozoa may account for up to 30% of ruminal fiber digestion (Russell and Rychlik, 2001). Many protozoa engulf and store starch granules, thereby modulating the rate of starch fermentation in the rumen (Russell and Rychlik, 2001).

Although protozoa float freely in the rumen fluid, large numbers may also attach to the surface of feed particles. Because they are predators of rumen bacteria, the number of protozoa in the rumen fluctuates inversely with the number of bacteria. In fact, ruminants can survive periods of complete defaunation where no protozoa are detected in the rumen, therefore, unlike bacteria, protozoa are not absolutely essential for rumen fermentation. Considerable effort has been devoted to the development of technologies to eliminate or alter the species composition of the ruminal protozoal population (Firkins et al., 2007). This defaunation process was reported to increase bacterial density, stimulate starch degradation and propionate production, and decrease methanogenesis, but a reduction in fiber digestion has also been reported (Morgavi et al., 2010; Mosoni et al., 2011). At present no defaunation technologies are routinely employed in commercial livestock production (McAllister and Newbold, 2008).

1.3.1.4 Other rumen microorganisms

In addition to the previously described major microbial groups, other organisms exist in the rumen, including methanogens, mycoplasmas and bacteriophages.

Bacteriophages are observed in the rumen in concentrations of $10^{10}$ per mL of ruminal fluid. Diurnal fluctuations in bacteriophage numbers in ruminal fluid have been
observed, and likely reflect changes in the number of available host bacterial cells, which is influenced by the feeding cycle of the host animal. Bacteriophages have been found in association with cellulolytic, amylolytic, methanogenic and acetogenic rumen bacteria (Klieve et al., 2004). Infection of rumen bacteria by lytic phage may account in part for the high levels of bacterial cell lysis, frequently reported in studies using ruminal fluid (Wells and Russell, 1996). In the future, bacteriophage therapy administered by inoculating the rumen with phage targeted against undesirable bacterial species (e.g., *Streptococcus bovis* or *Escherichia coli* O157:H7) may serve as a means of preventing some digestive diseases or pathogen transmission (Herrera et al., 2009; Rivas et al., 2010).

Although anaerobic mycoplasmas were identified in the rumen about 40 years ago, they are the least studied of the rumen microorganisms (Stewart et al., 1997). Mycoplasmas are commonly co-isolated with protozoa, fungi and methanogens, likely because they lack a cell wall and are insensitive to the antibiotics frequently used in the isolation of these organisms of interest.

Methanogens belong to the domain *Archaea*. Methanogens provide anaerobic ecosystems with a route of hydrogen disposal, enabling reduced cofactors such as NADH to be reoxidized. Thus, they play a critical metabolic role in the recycling of reducing equivalents, enabling rumen microbes to derive energy from the fermentation of carbohydrates, proteins and lipids (Whitford et al., 2001). The greenhouse gas – methane is formed when hydrogen is used to reduce carbon dioxide. This is the case for most species of methanogens, but other substrates including formate, acetate, methylamine and methanol, can also be used as substrates by some species of methanogens (Stewart et al., 1997).
To date, over 100 species of methanogens have been identified, but only seven of them have been isolated and cultivated from the rumen (Joblin, 2005). Methanogens are estimated to comprise approximately 0.3 – 3% of the rumen microbial biomass (Janssen and Kirs, 2008). Although they make up only a small portion of the rumen microbial biomass, methanogens play a crucial role in rumen function and animal nutrition. Efficient \( \text{H}_2 \) removal eliminates the inhibitory effect of hydrogen accumulation on microbial fermentation and leads to a more favourable pattern of VFA formation nutritionally and to an increased rate of fermentation (McAllister and Newbold, 2008).

The population densities of methanogens in the rumen appear to be influenced by diet, with emissions per unit of feed digested increasing when ruminants are fed high fiber diets (McAllister and Newbold, 2008). The possibility of negative consequences of climate change has led a major international effort by researchers to explore strategies to lessen ruminal methane emissions (McAllister and Newbold, 2008; Morgavi et al., 2010).

### 1.3.2 Microbe-microbe Interactions

The rumen microbiota is not a random mixture of hundreds of species of microorganisms; rather it is a structured and dynamic ecosystem. Over 80% of rumen microbial cells are attached to solid feed particles and thus establish and function in the form of biofilms (Cheng and McAllister, 1997; Costerton et al., 1987). The members of the rumen microbial community interact extensively. Examples of both synergistic and antagonistic relationships between rumen microorganism have been observed (Orpin and Joblin, 1997; Stewart et al., 1997; Williams and Coleman, 1997). The examples include both synergistic and antagonistic relationships among bacterial species, predation of
ruminal bacteria and fungi by ciliate protozoa and initial plant cell wall invasion by fungal hyphae providing bacteria with access to the interior of plant cells.

Synergism is more than general protocooperation, and there are several examples of this most common relationship identified in the rumen microbiota. Not only do both synergists benefit from collaboration, the resultant substrate consumption or product formation is substantially higher than the sum of activity of the individuals (Nikolaev and Plakunov, 2007). For example, although facultative anaerobic microorganisms only exist in the rumen ecosystem in low abundance, their consumption of oxygen facilitates the growth of their strict anaerobic ‘roommates’ (Wolin et al., 1997).

Cellulolytic microorganisms establish synergistic relationships with non-cellulolytic species, an interaction that accelerates the rate of cellulose degradation. The specific adhesion of bacteria to plant fibers is the essential first step in plant cell wall digestion, but optimal rates of cellulose digestion are not achieved unless the organisms are combined with coworkers such as bacteria or fungi (Costerton et al., 1987). As soluble nutrients arising from cellulose digestion accumulate within the biofilm, they become available to the cellulolytic bacteria themselves as well as to heterotrophic secondary colonizers which are attracted by chemotaxis and stimulated to divide and to form structured consortia. For example, cells of *Treponema bryantii* and *Butyrivibrio fibrisolvens* are commonly found in association with adherent cellulolytic bacteria, especially *F. succinogenes* (Costerton, 2007; Costerton et al., 1987; Kudo et al., 1987). While they have no cellulolytic enzymes, *T. bryantii* exhibits a chemotactic response towards butyrate, a fermentation end product generated by cellulolytic bacteria that also inhibits cellulolytic enzymes if it accumulates. Thus by utilizing butyrate for their own
growth, *T. bryantii* enhances the activity of the cellulolytic bacteria by preventing end product accumulation (Costerton, 2007; Kudo et al., 1987).

As stated previously, methanogenic archaea establish synergistic relationships with cellulolytic bacteria, anaerobic fungi and even protozoa, by utilizing hydrogen and formate produced in the course of cellulose fermentation. This prevents the accumulation of reduced coenzyme NADH and stimulates ATP synthesis within the microbial community (McAllister et al., 1994; Nikolaev and Plakunov, 2007).

Competition between various species of ruminal bacteria is also common. For example, *R. albus* 7 produces a bacteriocin with activity against *R. flavefaciens* FD-1 (Chen et al., 2004; Nikolaev and Plakunov, 2007; Odenyo et al., 1994), most probably because these two species compete for the same nutritional source as they are both cellulose degraders and occupy the same niche.

The specific order in which various species colonize the digesta surface is thought to influence the spatial organisation of the rumen microbiota (McAllister et al., 1994), although the initial rate of attachment of rumen bacteria to forage is thought to be similar among colonizing species (Edwards et al., 2007; Edwards et al., 2008a). Fungal spores, on the contrary, are thought to attach to the forage slower than bacteria (Edwards et al., 2008b). But fungal zoospores are able to colonize the lignified tissues preferentially, and the vegetative thalli are better at penetrating plant tissue than are bacteria and protozoa, and provide new attachment sites for the latter groups (Nagpal et al., 2009). Current results infer that utilization of nutrients by primary and secondary colonizers promotes further proliferation and stimulates subsequent development and maturation of the biofilm into a structured consortium (Edwards et al., 2008a).
Organisation of ruminal microorganisms into biofilms has several advantages. Firstly, the self-produced extracellular polymeric substances (EPS) that coat biofilm communities trap nutrients and concentrate enzymes adjacent to their targeted substrate (McAllister et al., 1994). Secondly, competing microbes are excluded from the digestion site. Formation of EPS protects the cellulytic enzymes on the cell surface from degradation by ruminal proteases (Miron et al., 2001). Thirdly, mature ‘stable’ multi-species biofilms are retained in the rumen as much as three times longer than planktonic cells, and are resistant to detachment (Edwards et al., 2008a; McAllister et al., 1994), thus this arrangement increases their opportunity to thoroughly digest plant fibers (Miron et al., 2001). Fourthly, adherent microbes are protected from a range of antimicrobials including antibodies, antibiotics and bacteriophages (Costerton et al., 1987). It has also been proposed that biofilms offer protection from predation (Costerton et al., 1987), although many protozoa attach to feed particles (Williams and Coleman, 1997) and some researchers have argued that they have mechanisms to predate attached bacterial communities (Edwards et al., 2008a). Moreover, horizontal gene transfer (HGT) has been documented among rumen microbial members (Keeling and Palmer, 2008; Ricard et al., 2006), which may also be facilitated by the high density of microorganisms associated with biofilms (Flemming and Wingender, 2010).

1.4 Plant cell wall degradation by rumen microorganisms

In nature, the hydrolysis of plant cell wall fiber is carried out by fiber-degrading microorganisms, which include both aerobic and anaerobic fungi and bacteria present in soil and the guts of animals. These microorganisms synthesize a complex collection of
cellulases, hemicellulases and ligninases. The microbial consortium in the rumen is unique and amongst the most active fiber degrading system known (Selinger et al., 1996).

1.4.1 Structure of plant cell walls

The plant cell wall is composed primarily of a group of polymers known as lignocellulose, which comprises about half of the plant biomass and stores a large portion of the solar energy captured through photosynthesis (Sánchez, 2009). It represents the most abundant renewable organic resource on earth. Lignocellulose consists of three major components: cellulose, hemicellulose and lignin, which are strongly intermeshed and chemically bound by non-covalent forces and covalent linkages (Figure 1.1) (Kumar et al., 2008). Cellulose and hemicellulose are macromolecules constructed from different sugar residues; whereas lignin is composed of various polyphenolics.

Cellulose is a major constituent of plants and is a linear biopolymer of D-glucose subunits linked through β – 1, 4 glycosidic bonds. The elemental fibrils are linked together by hydrogen bond and van der Waals forces (Sánchez, 2009). Depending on the degree of hydrogen bonding within and between cellulose molecules, this polysaccharide is found in crystalline or paracrystalline (amorphous) forms. In the latter conformation, cellulose is more susceptible to enzymatic degradation (Krause et al., 2003). In nature, cellulose is associated with other plant polymers, primarily hemicellulose and lignin, and this association may affect its biodegradation (Lynd et al., 2002). Cellulose-hydrolyzing enzymes (i.e. cellulases) are divided into three major groups: cellobiohydrolase (exoglucanase), endoglucanase, and β-glucosidase (Figure 1.2) (Lynd et al., 2002).
Hemicellulose, the second most abundant component of lignocellulosic biomass, is a group of branched heterogeneous polysaccharide composed of pentoses (D-xylose, D-arabinose), hexoses (D-mannose, D-glucose, D-galactose) and sugar acids (4-O-methylglucuronic, D-galacturonic and D-glucuronic acid) and typically has a lower molecular weight than cellulose (Sánchez, 2009). The subunits are generally linked together by β – 1, 4 –, and sometimes β – 1, 3 – glycosidic bonds. Hemicelluloses link cellulose fibers into microfibrils and cross-link with lignins, creating a complex network that provides structural strength.

Xylan is a common hemicellulose, and is composed of β – 1, 4 glycosidic bond linked D-xylose backbone. The xylose residues can be substituted with acetic acid at the C2 and/or C3 positions, 4-O-methylglucuronic acid at the C2 position, and arabinose at the C3 position. The arabinose may be further esterified by phenolic acids, which crosslink xylan and lignin within the matrix (Christov and Prior, 1993). Other common forms of hemicelluloses include mannan and glucomannan, xyloglucan, and β-glucan, based on the composition of the backbone sugar residues (Scheller and Ulvskov, 2010). The backbone of mannan consists entirely of mannose residues, and that of glucomannan is formed by D-glucose and D-mannose residues. Both xyloglucan and β-glucan have a backbone composed of D-glucose residues. In the former, most of the glucose residues are substituted with α – 1, 6 – linked xylose residues, but in latter, the backbone is linked through either β – 1, 4 or β – 1, 3, and in some cases β – 1, 6 glycosidic bonds. Similar to xylan, the backbones of these polymers are commonly acetylated or substituted by sugar/sugar acid residues. The type and degree of substitution depends on the plant species and tissues to which it is composed.
Due to their heterogeneity, the degradation of hemicelluloses involves many kinds of glycoside hydrolases and carbohydrate esterases (Dashtban et al., 2009; Sánchez, 2009). Major enzymes involved in xylan degradation include xylanase, glucuronidase, arabinofuranosidase, acetylxylan esterase and ferulic acid esterase (Figure 1.2). Enzymes such as endomannase, galactosidase, β–mannosidase and β–glucosidase are also involved in hemicellulose degradation.

Lignin is typically a complex polyphenolic 3-dimensional framework containing thousands of phenolic units (Dashtban et al., 2009). It is an amorphous heteropolymer, insoluble in water and optically inactive. It is formed from phenylpropane units joined together by non-hydrolyzable linkages including C-C and aryl-ether linkages (Sánchez, 2009). Oxidation of lignin is catalyzed by ligninases including lignin peroxidase, manganese peroxidase, versatile peroxidase and laccase, and their activities require oxygen. The basidiomycetes “white-rot” fungi are currently the only known efficient lignin degraders (Martínez et al., 2009a). The anaerobic microorganisms in the rumen are not capable of degrading lignin (Weimer et al., 2009).

1.4.2 Carbohydrate active enzymes

As shown in Figure 1.2, many different kinds of enzymes are involved in lignocellulose degradation. Combined, enzymes involved in plant cell wall carbohydrate digestion are called carbohydrate active enzymes (CAZymes). A specialized database, CAZy – the Carbohydrate-Active enZYmes Database, is dedicated to the display and analysis of genomic, structural and biochemical information on carbohydrate active enzymes including catalytic domains (CDs) and carbohydrate-binding modules (CBMs) that degrade, modify, or create glycosidic bonds (Cantarel et al., 2009). This
classification system was first introduced in late 1980s (Henrissat et al., 1989), and now it has become the gold standard for the classification of these kinds of enzymes (Cantarel et al., 2009). This complete classification system groups enzymes into families based on primary structure comparisons of their catalytic domains (Collins et al., 2005). The classification groups continue to grow as new CAZy sequences are identified. As the structure and molecular mechanisms of an enzyme are related to its primary sequence, the CAZy system reflects both structural and mechanistic features. Enzymes within a particular family have a similar three-dimensional structure and similar catalytic mechanism; however, members within one classification family may be very diverse in their substrate specificity.

Currently CAZymes are classified into the following major classes: glycoside hydrolases (GHs), glycosyl transferases (GTs), polysaccharide lyases (PLs), carbohydrate esterases (CEs) and carbohydrate-binding modules (CBMs). Glycoside hydrolases hydrolyze and/or rearrange glycosidic bonds. This is the largest CAZyme group, consisting of over 100 different families, depending on their sequence and enzyme characteristics. It contains many important enzymes involved in polysaccharide degradation, such as cellulases, xylanases and many other sugar hydrolases. Glycosyl transferases are the enzymes which form glycosidic bonds between sugar residues. Many enzymes involved in polysaccharide synthesis are assigned to this class. Polysaccharide lyases cleave glycosidic bonds through non-hydrolytic mechanisms. Carbohydrate esterases hydrolyze carbohydrate esters, and are actively involved in modifying and removing sugar residues such as during the debranching of xylan molecules. Carbohydrate-binding modules are not catalytic modules, but contiguous amino acid
sequences that exhibit carbohydrate-binding activity. Usually CBMs are found within CAZymes and usually function as the substrate binding domain. Some exceptions are CBMs in cellulosomal scaffoldin proteins and rare instances where putative independent CBMs have been reported (http://www.cazy.org).

1.4.2.1 Enzymes involved in cellulose degradation

Cellulases are very diverse in their sequences, structures and mechanisms, even though cellulose is a structurally simple homopolymer of glucose. All cellulases use one of two mechanisms to degrade cellulose: hydrolysis with retention or hydrolysis with inversion mechanism (Wilson, 2008). The former maintains the stereochemistry of the anomeric hydroxyl group of the sugar subunit, in this case, the glucose residue; while the latter inverts the stereochemistry of the anomeric hydroxyl group. There are two functionally different types of cellulases: endoglucanases (also called endocellulases) and cellobiohydrolases (also called exocellulases). A third enzyme, β-glucosidase, which cleaves di- and oligosaccharides, the products of cellulases, is required to completely hydrolyze cellulose to glucose. Generally speaking, exocellulases are processive enzymes that remain attached to the cellulose chain until it is completely hydrolyzed, whereas endocellulases can be either processive or non-processive (Kurasin and Valjamae, 2011; Sukharnikov et al., 2011).

Exocellulases sequentially cleave di- and oligosaccharides (usually 2 – 4 residues) from the end of the cellulose chain and accordingly, their active sites are in the shape of a tunnel (Sukharnikov et al., 2011). There are two classes of exocellulases: the first attacks the non-reducing end whereas the second attacks the reducing end of cellulose (Lynd et al., 2002). All known fungal and the majority of bacterial exocellulases that are active on
the non-reducing end of the chain are classified in family GH6, but several from bacteria are classified in GH5 and GH9. All exocellulases from aerobic fungi with activity against the reducing end are classified in family GH7, while the bacterial members are in family GH48. It is interesting that the anaerobic fungal members of this class fit in family GH48 instead of GH7 (Wilson, 2008). Actually, recent studies have shown that there is a strong similarity between the CAZymes of the rumen anaerobic fungi and those from rumen bacteria, rather than those from aerobic fungi, suggesting that anaerobic fungi exchanged genes coding for these enzymes through horizontal gene transfer with bacteria (Garcia-Vallve et al., 2000).

On the contrary, endoglucanases randomly bind to the interior of long cellulose chains and cleave the glycosidic bond between sugar residues. They belong to GH families 5, 6, 7, 9 and to over a dozen other families. All known structures of endocellulase CDs have an open active site, a prerequisite for their ability to bind the interior region of the cellulose chain (Wilson, 2008).

Processive endocellulases have interesting properties and were first discovered in late 1990s (Irwin et al., 1998; Reverbel-Leroy et al., 1997; Sakon et al., 1997). These enzymes initially bind to an interior point along a cellulose molecule, but instead of releasing from the cellulose fiber after the first cleavage, the non-reducing end of the cellulose chain is shifted to the enzyme’s empty -4 to -1 subsites, enabling processive cleavage of cellotetraose from the non-reducing end of the cellulose chain (Wilson, 2012). These enzymes are currently assigned to either GH9 or GH48 families. Because most anaerobic cellulolytic bacteria do not produce GH6 exocellulases (those working on the non-reducing end), it is believed that the processive endocellulase appears to replace
their activity and play an important role in enabling anaerobic bacteria to degrade cellulose (Wilson, 2012).

Cellulases often act synergistically on crystalline cellulose, with the specific activity of some cellulase mixtures containing four to six enzymes showing an activity that is over 10 times higher than that of any single cellulase in the mixture (Lynd et al., 2002). It seems that synergism only occurs when two cellulases attack different regions of the cellulose molecule, with each cellulase creating new attack sites for other enzymes within the mixture (Sánchez, 2009).

### 1.4.2.2 Enzymes involved in xylan degradation

With many kinds of side chain modifications, xylan contains a variety of chemical linkages, and thus its degradation requires a number of different enzymatic activities. The breakage of xylan backbone $\beta-1,4-\beta$ linkage requires only a single enzyme: xylanase (aka endo $1,4-\beta$ xylanase). Like endoglucanases, endoxylanases randomly cleave the glycosidic bond at the interior of xylan molecules, by either a retaining or inverting mechanism in terms of the anomeric configuration of the reactant xylose residue (Collins et al., 2005). The major xylanase families are GH10 and GH11, while GH5, 7, 8, 43 and a few other families also possess some members that exhibit this activity. Families GH5, 7, 10 and 11 contain enzymes which carry out hydrolysis with a retaining mechanism (Collins et al., 2005). In contrast, enzymes in families GH8 and 43 typically utilize an inverting mechanism (Collins et al., 2005). In general, GH10 xylanases have broader substrate specificity than those of GH11. Specifically, GH10 enzymes not only degrade linear chains of $\beta-1,4-\beta$ linked xylose residues, but also xylan backbones that are highly substituted as well as smaller xylo-oligosaccharides (van den
Brink and de Vries, 2011). The xylo-oligosaccharides released by endoxylanases are further degraded by β-xylosidases, mostly belonging to the GH3 or GH43 families (van den Brink and de Vries, 2011).

To completely degrade xylan, all substitutions on the backbones have to be released. This requires several different enzymes divided over many GH and CE families. Acetylxylan and ferulic acid esterases are two major kinds of carbohydrate esterases (Dashtban et al., 2009). Acetylxylan esterases release acetyl residues from xylan chains. They are distributed into at least eight CE families, including CE1, 4, 5, and 16 (Biely, 2012). The presence of acetylxylan esterases is essential for efficient degradation of the xylan backbone by endoxylanases (van den Brink and de Vries, 2011). The major difference between the CE families lies in the degree to which they hydrolyze different O-linked acetyl groups. Families CE1, 4, and 5 have a strong preference for side chains linked to hydroxyl group at C-2 position (2-O-linked) of the xylose residue, which is the most common linkage in hemicellulose, while CE16 prefers 3-O- and 4-O-linked residues (Biely, 2012; Li et al., 2008).

Ferulic acid esterases remove p-coumaric acid and ferulic acid, the two cinnamic acids present in xylan. Some of these esterases belong to CE1, while a considerable number have yet to be assigned to a CE family. Ferulic acid esterases have been divided into five types based on substrate specificity, the nature of the product released and the degree of similarity in amino acid sequences (Qi et al., 2011). Some particular groups of esterases show preference for substrates with methoxy substituents such as ferulic acid, while others prefer substrates containing one or two hydroxyl substitutions, such as p-coumaric acid (van den Brink and de Vries, 2011).
1.4.3 Cellulose biodegradation by microorganisms

Fungi are predominantly responsible for lignocellulose degradation in the environment with the most rapid degraders belonging to the basidiomycetes (Sánchez, 2009). The aerobic fungus *Hypocreajecorina*, originally called *Trichoderma reesei*, is the most studied aerobic cellulolytic microorganism (Wilson, 2008).

Since many microorganisms are unable to transport insoluble materials across the cell membrane, the enzymatic degradation of cellulose and hemicellulose needs to occur extracellularly. The produced soluble sugars are then transported inside the cell for further metabolism. Currently, two strategies for plant cell wall digestion by lignocellulolytic microorganisms have been described: (i) free cellulose mechanism: secretion of “free enzymes” extracellularly and (ii) cellulosomal mechanism: enzymes that are maintained in close association with the outer cell envelope layer (Wilson, 2009). Some researchers believe that there is a third strategy as discussed below (Wilson, 2008).

Most aerobic cellulolytic microbes, including bacteria and fungi, secrete sets of individual enzymes, which act synergistically on native cellulose. Many of these enzymes contain one or more CBMs, joined by a flexible linker peptide to the CD. The CBMs may be found on the C- or N- terminus, depending on the enzyme, but the location of a CBM normally does not affect enzyme activity (Wilson, 2012). Processive exocellulases and endocellulases are believed to be important components of the cellulose degrading enzyme complex and often account for more than half of the total cellulose degrading proteins (Wilson, 2009).

Most anaerobic microorganisms utilize a different strategy for cellulose degradation in the form of large multienzyme complexes termed cellulosomes, with molecular
weights of over 1 million Dalton (Da) (Bayer et al., 2008). Only a few of the enzymes in cellulosomes contain CBMs, but the scaffoldin protein to which the enzymes attach contains a CBM that binds to cellulose (Bayer et al., 2008). Processive cellulases are also important cellulosomal components (Wilson, 2009).

A strategy used by two cellulolytic bacteria does not seem to conform to either of the two methods of plant cell wall degradation described above (Wilson, 2008). The Gram-negative anaerobe *F. succinogenes* is one of the major cellulose degraders within the rumen and has been extensively studied (Jun et al., 2007). It grows very rapidly when utilizing cellulose as the sole carbon and energy source, owing to its efficient cellulose degrading mechanism (Fields et al., 2000). Another Gram-negative bacterium *Cytophaga hutchinsonii* is an aerobic cellulolytic bacterium. Both the genomic sequences of *C. hutchinsonii* (Xie et al., 2007) and *F. succinogenes* (Qi et al., 2005; Ransom-Jones et al., 2012) provide strong evidence that their mechanism of cellulose digestion differs from the two strategies previously described. Most of the cellulase genes do not encode for a CBM, nor a dockerin domain or scaffoldin gene. All of the cellulase genes appear to code for endoglucanases, and there are no genes that code for any known exocellulases or processive endocellulases (Qi et al., 2005; Ransom-Jones et al., 2012). These aspects suggest that these organisms do not utilize the free cellulase mechanism or the cellulosome mechanism for the degradation of cellulose. One possible mechanism proposed was that individual cellulose molecules are transported into the periplasmic space where they are degraded by endoglucanases (Ransom-Jones et al., 2012; Wilson, 2008).
1.5 Current advances in omics studies

In the past, investigations focusing on rumen microbial communities were usually directed at describing diversity and richness. Studies investigating the functionality of the rumen ecosystem were directed at isolating the dominant species. In the last decade, studies began to investigate the functionality of the complete microbial consortium by using methods such as real time quantitative PCR (qPCR), cDNA (complementary DNA) libraries, microarrays and more recently, meta-omic approaches (i.e., metagenomics, metatranscriptomics and metaproteomics). These approaches have become possible as a result of recent advances in high-throughput sequencing and mass spectrometry based peptide sequencing as well as computational analyses. Combining these techniques has allowed researchers to determine the microbial genes involved and gene expression by natural communities without the need for cultivation in the laboratory (Rosen et al., 2009).

1.5.1 Next generation sequencing

Prior to the introduction of next-generation sequencing (NGS), the automated Sanger method dominated the DNA sequencing market for almost 20 years, and led to a number of historically outstanding accomplishments, including the sequencing of the human genome. Sanger sequencing is based on the chain-termination method, in which a series of different-sized fragments of DNA are generated from numerous identical copies of one DNA molecule starting at the same location, but ending at different locations with a chain terminating dideoxynucleotide labelled with one of four fluorescent dyes. All the fragments are then resolved in order of the length via capillary electrophoresis and the
original sequence is determined through sequentially “reading” the last chain terminating fluorescently labelled dideoxynucleotide of each fragment (Schadt et al., 2010).

The limitations of automated Sanger sequencing — chiefly low throughput and high cost, created a demand for new and improved technologies for mass sequencing of genomes. In comparison to automated Sanger sequencing, newer methods are referred to as NGS. The major advantage offered by NGS is the ability to produce an enormous volume of data at a very affordable price. At the same time, even newer techniques are emerging and are referred to as third-generation sequencing (TGS), the most promising of which is the single molecular real-time sequencing technology developed by Pacific Biosciences (Pac-Bio) (Schadt et al., 2010) which has just been newly introduced to the market.

Currently several NGS technologies are commercially available including Roche/454, Illumina/Solexa, Life Technologies/SOLiD, and Ion Torrent (Loman et al., 2012). Although these platforms are quite diverse in sequencing biochemistry, conceptually their work flows are similar (Shendure and Ji, 2008). They all rely on a three-stage workflow of library preparation, template amplification and sequencing (Loman et al., 2012).

Generally speaking, an initial fragmentation step is required to generate random, overlapping DNA fragments ranging from 150 base pair (bp) to 800 bp in length by either mechanical or enzymatic fragmentation. Specific adaptors can be ligated to the ends of the fragmented molecules to serve as primer-binding sites for the subsequent template amplification reaction (Loman et al., 2012). Mate pair sequencing is supported by all platforms. In this method, the ends of DNA fragments (typically several kilobases)
are joined together to form circular molecules and subjected to a second fragmentation. Fragments flanking the joint position are then selected and adaptors added (Metzker, 2010). Paired-end sequencing is similar to mate pair sequencing, but DNA fragments are directly sequenced from each end without the need for additional preparation steps (Metzker, 2010). The Illumina platform has direct support for paired-end sequencing. Mate pair and paired-end sequencing provide valuable information about the location of sequences distributed across the genome, facilitating assembly (Loman et al., 2012).

Current commercial NGS platforms immobilize and spatially separate millions or billions of template molecules on a solid surface where they are amplified. Simultaneous solid-phase amplification of the immobilized singular template fragments enables massive parallel sequencing (Shendure and Ji, 2008). Pac-Bio’s TGS technology only requires unamplified single DNA molecule templates, with no amplification step required (Schadt et al., 2010).

The actual sequencing procedure is often described as “sequencing by synthesis” and relies on imaging-based data acquisition (Shendure and Ji, 2008). Several biochemistry mechanisms are applied by different platforms (Table 1.2) (Loman et al., 2012), and the enzyme driving the synthesis can be either a DNA polymerase or a DNA ligase. Data are acquired each cycle by imaging fluorescent signals of the full array, which are generated when fluorescently labelled nucleotides are incorporated. Each sequencing run generates millions of short sequences called reads. The detailed sequencing and imaging mechanisms are clearly explained and illustrated in a number of excellent review papers (Mardis, 2008; Metzker, 2010; Shendure and Ji, 2008).
It is likely that multiple platforms will coexist in the marketplace with some exhibiting advantages for particular applications over others, since there is considerable variation in performance including throughput, read length, error rate, as well as, cost and run time (Table 1.2) (Loman et al., 2012). The efficiency of these technologies is rapidly advancing as NGS companies are constantly improving their platforms to enable more rapid and comprehensive sequencing at lower cost.

The NGS technologies dramatically outperform older Sanger-sequencing technologies by a factor of 100 - 1,000 in daily throughput, and reduce the cost of sequencing one million nucleotides to as low as 0.1% of that associated with Sanger sequencing (Kircher and Kelso, 2010). This dramatic improvement in sequencing efficiency at reduced cost has opened up new approaches on how sequencing based technologies can be applied. As a result there has been an exponential increase in publications in which NGS is applied for a vast variety of research purposes. Important applications include: (I) full-genome sequencing, more targeting discovery of mutations or polymorphisms or large-scale comparative and evolutionary studies by sequencing many related organisms or strains within one species (pangenomics) (Metzker, 2010); (II) metagenomics, which targets the whole microorganism ecosystem directly obtained from environmental samples, instead of depending on cultivation of individual microbial species (Mardis, 2008; Turnbaugh et al., 2009); (III) mapping of structural rearrangements, including copy number variation, balanced translocation breakpoints and chromosomal inversions (Shendure and Ji, 2008); (IV) large-scale analysis of DNA methylation (Meissner et al., 2008); (V) ‘Chip-Seq’—genome wide mapping of DNA-protein interactions, by deep sequencing of DNA fragments which are isolated through
chromatin immunoprecipitation (Mardis, 2008; Park, 2009); and (VI) ‘RNA-Seq’—sequencing of RNA molecules, which is covered in detail in section 1.5.2.

1.5.2 Transcriptomics, metatranscriptomics and RNA-Seq

The transcriptome is the complete set of transcripts in a cell, and the quantity that has been synthesized for a specific developmental stage or physiological condition (Wang et al., 2009). Transcriptomics, the study of the transcriptome, is essential for interpreting the expressed functional elements of the genome. Metatranscriptomics is a branch of transcriptomics, which studies and correlates the transcriptomes of a group of interacting organisms or species. Since its inception, transcriptomics has quickly become an important and promising tool for ecological studies, especially those focusing on complex communities (Warnecke and Hess, 2009). Although DNA-based genomics and metagenomics provide abundant information on the metabolic and functional capacity of an organism or a microbial community, they cannot differentiate between expressed and non-expressed genes, and thus are not a true reflection of metabolic activities (Sorek and Cossart, 2010). On the contrary, transcriptomics and metatranscriptomics retrieve and sequence RNAs from a species or environmental samples. Thus, they provide the most unbiased perspective on gene transcription in situ (Su et al., 2012).

Before the wide application of NGS technologies, hybridization-based technologies such as microarray and Sanger sequencing were applied to assess the transcriptome. Microarrays were usually preferable, as it was not practical to use Sanger sequencing to sequence such a large volume of genetic material (Conway and Schoolnik, 2003; Ozsolak and Milos, 2011). In these studies, only a portion of the transcript was analysed and isoforms were generally indistinguishable from each other. These disadvantages limit the
capability of annotating the structure of transcriptomes. These studies also faced several challenges including the low recovery of high-quality mRNA from environmental samples, the short half-lives of mRNA species, and the need for separation of mRNA from other RNA species (Simon and Daniel, 2011). These limitations have been overcome to a great extent with the improvement of RNA isolation techniques in the past decade, together with the NGS-based RNA-Seq (RNA sequencing) technique (Wang et al., 2009). In contrast to microarray methods, NGS-based approaches directly determine the cDNA, or even RNA sequences. RNA-Seq provides a powerful method for both mapping and quantifying transcriptomes. The experimental procedure of RNA-Seq is similar to other NGS applications. Depending on the purpose of the study, a fraction of the total RNAs are isolated from a species or environmental sample and serves as the starting material for library construction. Several fractionation methods are used, based on the length/size or the traits of the target molecules. For example, studies targeting expression profiles of eukaryotes enrich messenger RNAs (mRNAs) by taking advantage of the fact that mature eukaryotic mRNAs are modified with the addition of polyadenylic acids to the 3’ end of the mRNA molecules (poly-A tailed). Oligo-dT primers hybridize to the poly-A tailed RNA fraction and thus selectively enrich the mRNA molecules that typically only account for 5-10% of total RNAs. With this approach, the highly abundant ribosomal RNAs and transfer RNAs are largely eliminated from the sample. Afterwards the RNAs are converted to a library of cDNA fragments with adaptors attached to the ends (Costa et al., 2010). The constructed library is sequenced in a high-throughput manner to obtain short sequences from one end (single-end sequencing) or both ends (paired-end sequencing) as described in section 1.5.1. Following sequencing, the
resulting reads are either aligned to a reference genome or reference transcriptome, or assembled \textit{de novo} without a reference genome. This generates a genome-scale transcription map that consists of both the transcriptional structure and level of expression of each gene (Martin and Wang, 2011).

RNA-Seq has provided the most promising approach for mapping and quantifying transcriptomes, especially metatranscriptomes (Wang et al., 2009), and offers several key advantages. First, unlike hybridization-based methods, RNA-Seq does not require prior knowledge of what genes might exist or be expressed, and thus is not limited to detecting transcripts that correspond to an existing genomic sequence. After sequencing, the resulting reads can not only be aligned to an existing reference genomic sequence or reference transcripts, but assembled \textit{de novo} without the genomic sequence (Martin and Wang, 2011). This makes it possible to identify novel gene sequences, and to quantify rare transcripts without prior knowledge of a particular gene. Consequently RNA-Seq is particularly attractive for non-model organisms and complex environmental samples where limited existing sequence information is available. Metatranscriptomics have been used to analyze many microbial communities including ocean surface waters (Frias-Lopez et al., 2008), coastal waters (Gilbert et al., 2008), soil samples (Urich et al., 2008) and the human gut (Gosalbes et al., 2011).

A further strength of RNA-Seq is its ability to detect and quantify individual transcript isoforms. Alternative splicing is known to contribute to functional diversity in eukaryotes, but it has not been well studied at the level of the transcriptome, principally because of the difficulty of measuring expression for each isoform (Malone and Oliver, 2011). RNA-Seq approaches provide direct sequence information that spans exon/exon
boundaries and makes it possible to study the expression, diversity and abundance of different isoforms of a gene (Malone and Oliver, 2011). Precise location of transcription boundaries and other RNA processing events can also be obtained (Metzker, 2010; Nagalakshmi et al., 2010).

Additionally, RNA-Seq is able to detect a very large dynamic range of expression level of transcripts (Gilbert and Hughes, 2011). In theory, the sequencing depth used is the only restriction on quantification limit. Deeper sequencing will detect sequences expressed at lower levels and quantify expression levels more accurately (Malone and Oliver, 2011). In contrast, microarrays are not sensitive enough to quantify genes expressed at very low levels whereas those expressed at very high levels can saturate the array (Costa et al., 2010). RNA-Seq has also shown consistent results when compared with qPCR results, with high levels of reproducibility (Wang et al., 2011).

Finally, RNA-Seq usually requires lower amounts of RNA sample compared to Sanger sequencing, because there are no cloning steps involved in library construction (Ozsolak and Milos, 2011). This is a huge advantage, especially for projects where limited amounts of RNA can be isolated.

Taking all of these advantages into account, RNA-Seq is the first sequencing method that allows the entire transcriptome to be surveyed in a very high-throughput manner. At a reasonable cost, RNA-Seq offers single-base resolution for annotation and quantification of gene expression levels at the genome scale.
1.5.3 Limitations and Challenges related to NGS and RNA-Seq

Next generation sequencing technologies have been extensively improved since the introduction of the first commercial platform in 2005 (Loman et al., 2012). Currently in addition to reducing the per-base cost of sequencing by several orders of magnitude, NGS instruments also have fewer infrastructure requirements. However, when compared to Sanger sequencing, NGS is still limited in terms of read-length and accuracy.

Compared to Sanger sequencing’s $10^{-4} – 10^{-5}$ error rate, the error rates of NGS technologies are extremely high at $10^{-2} – 10^{-3}$ (Kircher and Kelso, 2010). Consequently, even though deep sequencing provides abundant sampling depth compared to traditional approaches, such as DGGE (Denaturing gradient gel electrophoresis), T-RFLP (Terminal restriction fragment length polymorphism), or 16S ribosomal RNA (rRNA) gene clone libraries, the high error rate may result in the overestimation of rare phylotypes (Su et al., 2012). Despite that, direct sequencing of metagenomic DNA is still proposed to be the most accurate approach currently available for assessment of taxonomic composition as it avoids the bias introduced by polymerase chain reaction (PCR) amplification of DNA in approaches such as DGGE and 16S rRNA clone libraries (Su et al., 2012).

Read length for NGS technologies also remains limited. Sanger sequencing can normally reach over 1,000 bp; however, Roche/454 has the longest average read-length of up to 500 bp, while the other NGS technologies only read 100 – 200 bp (Loman et al., 2012). The relatively short read lengths consequently raise bioinformatic challenges as to how best and most efficiently extract biologically meaningful insights from the very large datasets produced (Kircher and Kelso, 2010; Shendure and Ji, 2008).
Downstream data management and bioinformatic analysis are the principal challenges with NGS (Pop and Salzberg, 2008). The massive data sets produced place substantial demand on information technology in terms of data storage, tracking, quality control as well as statistical analysis (Datta et al., 2010; Pop and Salzberg, 2008). Assembling millions of short reads into contigs before alignment to the reference genomic sequence, or mapping the short reads directly to the reference genome raises considerable bioinformatic challenges (Pop and Salzberg, 2008). Special attention must be paid to exon-exon junctions and polyA tails (Costa et al., 2010). Repetitive DNA or extremely AT or GC rich sequences present technical challenges as these regions are ambiguous for sequence alignment (Treangen and Salzberg, 2012).

Although small-scale projects in the kilobase-to-megabase range will still likely use conventional Sanger sequencing, future large-scale projects will likely rely entirely on NGS. As shown in Table 1.2, there are important differences among the NGS platforms that result in advantages with respect to specific applications. Some applications may be more tolerant of short read-lengths than others or differ in their overall accuracy and source of errors such as the rate of insertion-deletion vs substitution errors. Other considerations include the availability of the platform and sequencing costs.

1.5.4 Current advances of omics studies related to rumen fungi and microbiomes focusing on plant fiber degradation

Previous to the wide application of NGS technologies, an excellent study defining the nature of the termite hindgut microbiome generate over 71 million base pairs (Mbps) by applying Sanger sequencing method (Warnecke et al., 2007). This work set a new
standard for metagenomic research that focused on microbial consortia involved in plant fiber degradation. Since, as mentioned, the rumen is a unique resource for the discovery of novel plant cell wall degrading enzymes, a research group applied NGS to analyze the bovine rumen metagenome (Brulc et al., 2009). This study using 454 sequencing technology generated 103 Mbp of sequences from three fiber-adherent and one pooled liquid sample obtained from the rumens of three Angus Simmental Cross steers. Shortly after, the great power of Illumina NGS technologies was demonstrated in a cow rumen metagenomic project aiming at searching for potential novel lignocellulosic degrading enzymes suitable for the cellulosic biofuel industry (Hess et al., 2011). This project is the deepest and most complete rumen metagenomic study currently available and greatly expanded our understanding of rumen microbiota to a new level. From a total of 268 giga base pairs (Gbps), a little over 2.5 million open reading frames (ORFs) were assembled, with over half predicted to represent full-length genes. Over 1,500 OTUs and 27,000 putative CAZymes were identified, a level that was much higher than those identified from previous studies (Brulc et al., 2009; Pope et al., 2010; Warnecke et al., 2007). The functions of the CAZymes were also investigated. Ninety candidate proteins were expressed and 57% were found to be enzymatically active.

In an attempt to identify more unique CAZymes, studies have also been aimed at microbiomes from animals living in unique environments. Since macropods evolved in geographical isolation of other herbivores, they show a wide range of unique adaptations to diets and were proposed to harbour a different microbiome in their foregut than those existing in the bovine rumen. Pope et al. (2010) studied the Tammar wallaby foregut microbiome using Sanger sequencing to characterize 16S rRNA gene clone libraries and
shotgun libraries, with 454 sequencing performed on selected fosmids. The sequencing obtained a total of over 600 Mbp sequences.

When the OTUs detected from the termite (Warnecke et al., 2007), bovine rumen (Brulc et al., 2009) and wallaby were compared, clear host-specificity was obvious, with only a small number of OTUs shared between the bovine and wallaby microbiome, with those in termite microbial community being unique (Pope et al., 2010). This was expected as each unique host has a highly adapted microbiota. At the same time, the microbiomes from three individual bovine rumens fed on the same diet also exhibited considerable diversity (Brulc et al., 2009).

Dai et al. (2012) recently explored the rumen cellulolytic microbiome of Tibetan yak rumen by screening bacterial artificial chromosome (BAC) library for fibrolytic enzyme activities and 223 positive BAC clones were pyrosequenced on Roche/454 platform. About 150 glycoside hydrolase genes were identified with the majority occurring in gene clusters.

Pope et al. (2012) also investigated the rumen microbiome of arctic reindeer. Multiple polysaccharide utilization loci-like systems were found, as well as about 5,000 putative GH gene fragments from over 20 CAZy families, by analyzing the sequences as well as metabolic reconstruction of the Bacteriodales-related clade. A number of cohesin/dockerin modules, which were rarely reported in previous rumen metagenomic studies, were also identified, suggesting that cellulosomes may play an important role in cellulose digestion within these arctic ruminants.

Most of the previous studies have directly targeted the fiber degrading microbiota, and as expected, resulted in most of the sequences being of bacterial origin, as bacteria
represent the majority of the biomass in the rumen microbiota. A newly published paper aiming at transcriptomics and secretomics of Neocallimastix patriciarum W5 improved our knowledge towards the anaerobic fungi, a group of active fiber degraders (Wang et al., 2011). Other than sequencing the genome of this fungus, the researchers focused on functional characterization by sequencing transcriptomes expressed under various growth conditions, as well as secretomic analysis based on mass spectroscopic analysis. Both 454 and Illumina sequencing platforms were applied. This study helped to gain a better global understanding of the GHs produced by N. patriciarum. A total of 219 putative GHs were classified into 25 GH families. Some highly expressed or potentially full length contig candidates were expressed. At least five novel cellulases displayed activities, and one β-glucosidase and one exocellulase demonstrated high enzyme activities.

1.6 Research objectives

The rumen microbial ecosystem has now been investigated in detail for over half a century. Over the past two decades, great improvements have been made towards understanding the dynamic nature of this unique microbiota and it is one of the most accessible and understood microbial ecosystems (Flint, 1997). The rumen is widely recognized as one of the most unique fibrolytic microbial ecosystem that is second to none in its ability to convert plant cell wall polysaccharide to fermentable sugars.

The overall goal of this thesis project is to increase our understanding of the transcriptome of a representative rumen anaerobic fungus Anaeromyces mucronatus YE505 and the metatranscriptome of particle associated microbiota from muskoxen (Ovibos moschatus) rumen, with an emphasis on CAZyme coding sequences. Based on
the current incomplete understanding of the rumen microbiota, it is reasonable to propose
the hypothesis that a broad range of fibrolytic degrading enzymes yet unknown will be
discovered by this project by achieving the following objectives:

1) Establish a fast and reliable RNA isolation method for extracting total RNA
from rumen samples, especially from feed particle associated microbiota.

2) Elucidate the transcriptomes from the rumen fungus *A. mucronatus* YE505
grown on various carbon sources.

3) Elucidate the metatranscriptome of the rumen solid associated eukaryotes from
muskoxen and study the gene expression profile of these feed particle associated
microorganisms.

Metatranscriptomics is a rapidly emerging field and has shown considerable
potential as a means of identifying novel biocatalysts (Sorek and Cossart, 2010;
Warnecke and Hess, 2009). By applying (meta)transcriptomic analysis, this research will
provide a broader and deeper picture of the rumen fungi and the rumen ecosystem of
muskoxen, and enable sequence-based approaches to identify genes coding for novel
enzymes. It will enhance our understanding of the molecular mechanism of lignocellulose
bioconversion. Furthermore, the methods and procedures established in this study will
enable more detailed investigation of the impact of the host, diet and other conditions on
rumen function at the gene expression level. Knowledge generated by this study will also
aid in the industrial conversion of renewable plant fiber biomass to value added
economically significant products.
1.7 Tables and Figures

Table 1.1 Currently classified anaerobic fungal species isolated from the gut of herbivores (Borneman and Akin, 1994; Breton et al., 1991; Ozkose et al., 2001; Nagpal et al., 2009).

<table>
<thead>
<tr>
<th>Genus</th>
<th>Species</th>
<th>Source of isolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neocallimastix</td>
<td>N. frontalis</td>
<td>Sheep</td>
</tr>
<tr>
<td></td>
<td>N. patriciarum</td>
<td>Sheep</td>
</tr>
<tr>
<td></td>
<td>N. hurleyensis</td>
<td>Sheep</td>
</tr>
<tr>
<td></td>
<td>N. variabilis</td>
<td>Cattle</td>
</tr>
<tr>
<td>Piromyces</td>
<td>P. communis</td>
<td>Sheep</td>
</tr>
<tr>
<td></td>
<td>P. mae</td>
<td>Horse</td>
</tr>
<tr>
<td></td>
<td>P. dumbonica</td>
<td>Elephant</td>
</tr>
<tr>
<td></td>
<td>P. rhizinflata</td>
<td>Ass</td>
</tr>
<tr>
<td></td>
<td>P. minutus</td>
<td>Deer</td>
</tr>
<tr>
<td></td>
<td>P. spiralis</td>
<td>Goat</td>
</tr>
<tr>
<td></td>
<td>P. citronii</td>
<td>Horse</td>
</tr>
<tr>
<td></td>
<td>P. polycephalus</td>
<td>Water buffalo</td>
</tr>
<tr>
<td>Caecomyces</td>
<td>C. communis</td>
<td>Sheep</td>
</tr>
<tr>
<td></td>
<td>C. equi</td>
<td>Horse</td>
</tr>
<tr>
<td>Orpinomyces</td>
<td>O. intercalaris</td>
<td>Cattle</td>
</tr>
<tr>
<td></td>
<td>O. joyonii</td>
<td>Sheep</td>
</tr>
<tr>
<td>Anaeromyces</td>
<td>A. elegans</td>
<td>Cattle</td>
</tr>
<tr>
<td></td>
<td>A. mucronatus</td>
<td>Cattle</td>
</tr>
<tr>
<td>Cyllamyces</td>
<td>C. aberensis</td>
<td>Cattle</td>
</tr>
</tbody>
</table>
Table 1.2 Comparison of next-generation sequencing platforms (Loman et al., 2012) (reprinted with permission).

<table>
<thead>
<tr>
<th>Machine (Manufacturer)</th>
<th>Chemistry</th>
<th>Model read length (base)</th>
<th>Run time</th>
<th>Gb per run</th>
<th>Current approx. cost (US $)</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High-end instruments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 454GS FLX+ (Roche)     | Pyrosequencing | 700-800                  | 23 h     | 0.7        | 500k                        | • Long read lengths | • Appreciable hands-on time  
  • High reagent costs  
  • High error rate in homopolymers |
| HiSeq 2000/2500 (Illumina) | Reversible terminator | 2 × 100  
  (regular mode) or 2 d (rapid run mode)§ | 11 d      | 600 (regular mode) or 120 (rapid run mode)§ | 750k | • Cost-effectiveness  
  • Steadily improving read lengths  
  • Massive throughput  
  • Minimal hands-on time | • Long run time  
  • Short read lengths |
| 5500xl SOLiD (Life Technologies) | Ligation | 75 + 35  
  (max 15,000) | 8 d      | 150       | 350k | • Low error rate  
  • Massive throughput | • Very short read lengths  
  • Long run times |
| PacBio RS (Pacific Biosciences) | Real-time sequencing | 3,000 (max 15,000) | 20 min | 3 per day | 750k | • Simple sample preparation  
  • Low reagent costs  
  • Very long read lengths | • High error rate  
  • Expensive system  
  • Difficult installation |
| **Bench-top instruments** |           |                          |          |            |                             |            |               |
| 454 GS Junior (Roche)   | Pyrosequencing | 500                     | 8 h      | 0.035      | 100k                        | • Long read lengths | • Appreciable hands-on time  
  • High reagent costs  
  • High error rate in homopolymers |
| Ion Personal Genome Machine (Life Technologies) | Proton detection | 100 or 200  
  (316 chip) or up to 1 (318 chip) | 3 h | 0.01-0.1 (314 chip), 0.1-0.5 (316 chip) | 80k (including OneTouch and server) | • Short run times  
  • Appropriate throughput for microbial application | • Appreciable hands-on time  
  • High error rate in homopolymers |
| Ion Proton (Life Technologies) | Proton detection | Up to 200 | 2 h | Up to 10 (Proton I chip) or up to 100 (Proton II chip) | 145k + 75k for compulsory server | • Short run times
• Flexible chip reagents | • Instrument not available at time of writing |
| MiSeq (Illumina) | Reversible terminator | $2 \times 150$ | 27 h | 1.5 | 125k | • Cost-effectiveness
• Short run times
• Appropriate throughput for microbial applications
• Minimal hands-on time | • Cost-Read lengths too short for efficient assembly |

* Average read length for a fragment-based run.

# Approximate cost per machine plus additional instrumentation and service contract.

§ Available only on the HiSeq 250.
Figure 1.1 Composition of lignocellulosic residues (Sánchez, 2009) (reprinted with permission).
**Figure 1.2** Simplified schematic diagram illustrating plant cell wall components and enzymes involved in plant cell wall degradation.
Chapter 2 Isolation of high-quality total RNA from rumen anaerobic bacteria and fungi, and subsequent detection of glycoside hydrolases*

2.1 Introduction

The rumen is a highly specialized fermentation chamber containing an array of unique fibrolytic enzymes produced by bacteria, protozoa and fungi with potential broad application in both research and industry (Selinger et al., 1996). Microbial populations within the rumen have commonly been studied from the perspective of the liquid or solid fraction of rumen contents (Cheng and McAllister, 1997). Solid-associated microorganisms represent the major proportion of total rumen microbes (McAllister et al., 1994; Yu and Forster, 2005), and are estimated to produce up to 90% of the endoglucanase and xylanase activities in the rumen (Miron et al., 2001). Consequently, the study of the solid phase microbial community is likely to yield the most information about rumen microbial function.

Until recently, research on rumen microbial communities was mainly targeted at describing diversity, while research on function was limited to a relatively small number


43
of culturable microbial species. Recently, developments in next generation sequencing have made the culture-independent study of complex natural microbial habitats possible (Warnecke and Hess, 2009). However, to describe the genes present and their expression, reliable and repeatable nucleic acid isolation techniques are required.

Although total RNA has been isolated from environmental microbial communities in soil and the human intestinal tract (Peršoh et al., 2008; Sessitsch et al., 2002; Zoetendal et al., 2006), obtaining representative extracts of RNA from rumen contents, especially solid phase, remains challenging. Previous attempts to isolate solid-attached ruminal microorganisms relied on detachment of cells from feed particles. However, attached microorganisms grow in the form of multi-species biofilms that often lyse during the separation step and result in quick RNA degradation. Since rumen fungi penetrate feed particles, they are generally inaccessible to isolation methods that rely on biofilm detachment. It has repeatedly been shown that the recovery of attached microorganisms using a variety of detachment methods is incomplete (Martín-Orúe et al., 1998; Ranilla and Carro, 2003; Trabalza-Marinucci et al., 2006; Whitehouse et al., 1994), and that the detachment method affects the profile of the microbiota recovered from solid particles (Martínez et al., 2009b; Ramos et al., 2009). Furthermore, feed particles are rich in phenolic acids, polysaccharides and proteoglycans that readily form complexes with nucleic acids and inhibit reverse transcription and/or PCR reactions (Monteiro et al., 2001; Sharma et al., 2003). To date, only two RNA isolation procedures from rumen liquids have been reported, and resulted in modest yields of partially intact RNA, with no evidence of RNA contributions from rumen eukaryotes (Béra-Maillet et al., 2009; Kang et al., 2009). The isolation of intact total RNA from rumen solids has yet to be reported.
Here we report a successful isolation of total RNA from both solid and liquid phases of rumen contents. This procedure should facilitate the identification of actively transcribed genes from a variety of feed-associated microbes, including rumen fungi, enabling more representative gene expression profiles of the rumen ecosystem to be compared among individual hosts and with changes in diet and other conditions. Additionally, the technique may facilitate the isolation of intact, high quality total RNA from a variety of environmental sources.

2.2 Methods

2.2.1 Animals and rumen sampling

The overall experimental flowchart used for this study is illustrated in Figure 2.1. For isolation of total RNA from liquid phase rumen samples and for comparison of methods developed in this study and previous studies (Béra-Maillet et al., 2009; Kang et al., 2009), samples of ruminal content were obtained from a ruminally cannulated Holstein cow fed a 40% barley grain - mixed grass hay diet immediately prior to feeding. The cow was housed in a tie-stall barn at the Agriculture and Agri-Food Canada Lethbridge Research Centre in Lethbridge, Alberta, Canada and was cared for in accordance with the guidelines set by the Canadian Council on Animal Care (CCAC, 1993). Immediately after withdrawal of a sample, a liquid phase sample was separated by squeezing through four layers of cheesecloth and the liquid was transported to the laboratory in an insulated vessel. Any remaining large particulate fragments were then separated using a Bodum coffee filter plunger (Bodum Inc., Triengen, Switzerland). Fluid
phase aliquots of 0.2 mL each were placed in 2 mL microfuge tubes and stored at -80 °C until processed further (Figure 2.1, pane A).

Solid phase rumen contents were obtained from ruminally cannulated muskoxen at the R. G. White Large Animal Research Station, University of Alaska, Fairbanks, Alaska (Figure 2.1, pane B). All procedures with muskoxen were approved under protocol No. #139821-2 by the Institutional Animal Care and Use Committee at the University of Alaska Fairbanks. Solid phase samples were obtained by placing whole ruminal contents in a heavy walled 250 mL beaker and by separating the particulate and liquid phases using a Bodum coffee filter as described above. Subsamples of solid digesta (~ 2.5 g) were immediately flash-frozen in liquid nitrogen (Figure 2.1, pane B, branch B1). To examine the effectiveness of RNAprotect bacteria reagent (Qiagen, Mississauga, Ontario, Canada), solid phase samples were separated into six portions (~2.5 g each), with 5 mL of RNAprotect immediately stirred into each of the six samples just prior to being flash-frozen in liquid nitrogen. All samples were frozen within 5 min of the sample being withdrawn from the animal (Figure 2.1, pane B, branch B2). Samples were immediately transferred to the laboratory, and stored at -80 °C until processed further.

2.2.2 Total RNA isolation

2.2.2.1 Liquid phase rumen sample RNA isolation

The acid guanidinium-phenol-chloroform (AGPC) method (Chomczynski and Sacchi, 1987) served as the standard RNA isolation procedure (“control method”). Two published procedures for rumen sample RNA isolation (Béra-Maillet et al., 2009; Kang et
al., 2009) were tested and they were referred to as method K (Kang et al., 2009) and method B (Béra-Maillet et al., 2009).

An optimized isolation procedure, designated LRCI (Liquid Ruminal Contents Isolation) was developed. All experimental procedures were performed on duplicate subsamples as follows. Microfuge tubes containing 0.2 mL liquid phase rumen subsamples (Figure 2.1, pane A) were withdrawn from the freezer and 1.5 mL of TRIzol reagent was immediately added into each tube. The samples were then allowed to thaw at room temperature. Cells were disrupted by bead beating for 3 min at 300 revolutions per second with 0.2 g of glass beads of size range 0.7–1.1 mm (Sigma-Aldrich, Oakville, Ontario, Canada) at room temperature on TissueLyser (Qiagen). The homogenized sample was allowed to stand at room temperature for 5 min and RNA was isolated following the AGPC method (Chomczynski and Sacchi, 1987). The air-dried RNA pellet was re-dissolved in 100 µL of nuclease-free water (Qiagen). The RNA cleanup was performed by using either RNeasy mini kit or MEGAclear kit according to the manufacturer’s instructions.

2.2.2.2 Solid phase rumen sample RNA isolation

A procedure designated SRCI (Solid Ruminal Contents Isolation) was developed. First, rumen solids (RS) were manually ground to a fine powder in liquid nitrogen using a mortar and pestle, and then further ground for 5 min in liquid nitrogen using a Retsch RM100 grinder (Retsch GmbH, Haan, Germany). Ground samples (~200 mg) were placed in 2 mL microfuge tubes and each was mixed with 1.5 mL of TRIzol reagent. The samples were thawed, incubated at room temperature for 5 min and subsequently the
RNA was extracted using the AGPC method (Chomczynski and Sacchi, 1987), as described in the LRCI procedure.

The LCRI procedure with minor modification was also tested on rumen solids, by putting a small piece of frozen sample (~0.2 g) into a 2 mL microfuge tube and 1.5 mL of TRIzol reagent was immediately added. Cell disruption by bead beating used glass beads and the same equipment settings as liquid-phase samples but was carried out twice instead of once, with 5 min of incubation on ice between intervals.

2.2.2.3 **Effects of RNAprotect Bacteria Reagent on RNA isolation from solid phase rumen contents**

To test the efficacy of RNAprotect to preserve solid-phase rumen RNA, we employed RNAprotect in three different ways during the extraction of RNA from rumen solids and compared the results with the RNAprotect-free SRCI procedure. RNAprotect was added to rumen solids as described above in the sampling section. The six samples were divided into three groups in duplicate for the following three treatments (Figure 2.1, pane B, branch B2). In the first method (treatment I), the ~ 2.5 g sample was ground in liquid nitrogen in the presence of 5 mL of RNAprotect. Subsequently, TRIzol reagent was added and RNA was extracted according to the above SRCI procedure without removing RNAprotect. In the second method (treatment II), the sample was thawed on ice and centrifuged at 4 °C for 10 min at 5000 × g. Excess RNAprotect was removed from the sample as recommended in the RNAprotect manual. The pellets were then ground in liquid nitrogen and RNA was subsequently extracted. In the third method (treatment III), the sample was ground in liquid nitrogen in the presence of RNAprotect, but the sample was allowed to completely thaw and was centrifuged at 4 °C for 10 min at
5000 × g to remove the supernatant, which contained most of the RNAprotect reagent, and then TRIzol reagent was added to the pellets and RNA was subsequently extracted.

2.2.3 Effects of RNA clean-up kits on RNA quality

Three laboratory kits: RNeasy mini kit (Qiagen), RiboPure kit (Applied Biosystems/Ambion, Streetsville, Ontario, Canada), and MEGAclear kit (Applied Biosystems/Ambion), were tested for their ability to purify SRCl-extracted RNA from solid ruminal contents. The purification procedures were performed according to the manufacturer’s instructions.

2.2.4 Evaluation of RNA quantity and quality

RNA purity was estimated by measuring the absorbance ratio at A260/A280 and A260/A230 using a NanoDrop 2000 Spectrophotometer (Thermo Scientific, Waltham, Massachusetts). RNA concentration and integrity were estimated using an Agilent 2100 bioanalyzer (Agilent Technologies, Mississauga, Ontario, Canada) and RNA 6000 Nano kit (Agilent Technologies) according to the manufacturer’s recommendations. The prokaryotic total RNA nano assay protocol was used since prokaryotes account for the majority of RNA in rumen contents (Yu and Forster, 2005).

Large subunit/small subunit (LSU/SSU) rRNA peak area ratio and RNA integrity number (RIN) analyses were performed for each RNA sample using the 2100 Expert software version B.02.07 (Agilent Technologies).

2.2.5 Reverse transcription and PCR

Total RNA isolated from solid-phase muskoxen rumen sample was treated with a DNA-free kit (Applied Biosystems/Ambion) following the recommended procedures.
After treatment, total RNA (1 μg) was reverse-transcribed into cDNA using a SuperScript III first-strand synthesis kit (Invitrogen) and random hexamer primers according to the manufacturer’s instructions. A reverse transcriptase negative control was also included in all runs and the generated products were used in subsequent PCRs.

Fragments of three glycoside hydrolase genes celF, xynD, and cel3 from *Fibrobacter succinogenes* S85 were selected to detect transcript production. Two sets of primers for each gene were chosen as described by Béra-Maillet et al. (2009) (Table 2.1).

Subsequent PCRs were performed by using Platinum *Taq* DNA polymerase High Fidelity (Invitrogen). To each 25 μL PCR reaction system, 1 μL of first-strand cDNA was added. The PCR conditions for whole length primers (designated W, Table 2.1) consisted of an initial denaturation step for 1 min at 94 °C, followed by 40 amplification cycles of 15 s at 94 °C, 30 s at 55 °C, and 3 min at 68 °C. The final cycle included elongation at 68 °C for 5 min. The PCR conditions for internal primers (designated I, Table 2.1) consisted of an initial denaturation step for 1 min at 94 °C, followed by 40 amplification cycles of 15 s at 94 °C, 30 s at 55 °C, and 45 s at 68 °C. The final cycle included elongation at 68 °C for 2 min. A PCR negative control (no addition of first-strand cDNA template) was included with each PCR procedure. The PCR products were visualized on a 1% agarose gel stained with ethidium bromide.

### 2.3 Results

#### 2.3.1 Optimization of the Liquid Phase RNA isolation method

The extraction efficiency of three AGPC reagents—TRIzol (Invitrogen), TriPure (Roche Diagnostics, Laval, Quebec, Canada) and Tri Reagent (Applied
Biosystems/Ambion), was compared and the RNA obtained showed no apparent
differences with respect to RNA yield or quality (data not shown). Therefore these three
reagents were considered equivalent in our method, and we elected to use TRIzol in
further extractions owing to its widespread use.

The three methods examined here resulted in similar RNA purity, with an
A260/A280 ratio of 1.8 – 2.0 and an A260/A230 ratio of approximately 2.0. Among them,
method K was the most complicated as it involved cell dissociation, both enzymatic and
mechanical lyses and a column purification step (Kang et al., 2009). However, method K
yielded <40 µg of RNA per mL of rumen fluid (RF), with a LSU/SSU rRNA ratio of <1
and a RIN of <5, both of which indicated considerable degradation (Table 2.2). In
contrast with method K, fewer isolation steps yielded higher amounts of RNA from
Method B (152 ± 16 µg·(mL RF)^−1) (Béra-Maillet et al., 2009) and the LRCI method
(172 ± 14 µg·(mL RF)^−1) (Table 2.2). The RNA generated from method B were of higher
quality than that from method K, with an rRNA ratio of >1.2 and a RIN of >8. The LRCI
procedure isolated RNA with the highest quality and quantity, with rRNA ratios as high
as 1.8 and a RIN of >9.4 (Table 2.2). When a column cleanup step was included, the
RNA quality was further improved by removing 5S region fragments, without
compromising RNA yield. A typical RNA sample isolated from rumen liquids by
following this LRCI procedure is shown in Figure 2.2, part A.

2.3.2 Total RNA isolation from rumen solids

The application of our optimized method (Figure 2.1, pane B, branch B1) yielded
high-quality RNA from solid rumen samples. A typical electropherogram result is
illustrated in Figure 2.2, part B, with an average yield of approximately 110 µg·(g RS)^−1,
a LSU/SSU rRNA ratio of 1.9, and a RIN of 9.8. We compared the effect of bead beading and grinding under liquid nitrogen on RNA isolation. Disruption of solid contents by bead beading for 3 min two times at room temperature resulted in a low yield of poor-quality RNA (60 μg·(g RS)\(^{-1}\) with a LSU/SSU rRNA ratio of only 0.6. Therefore, grinding of samples in liquid nitrogen was selected for further RNA extraction from rumen solids.

RNA extracted from solid-phase samples from a muskoxen showed a higher complexity than that from the liquid-phase sample from a cow, as is evidenced by the two major ribosome RNA peaks (Figure 2.2). Each peak showed a combination of slightly different-sized fragments, which was expected, since the sizes of rRNAs differ between bacteria, archaea, and eukaryotes. A band was noticeable in the 5S RNA region and may arise from small RNA fragments or other impurities in rumen contents (Figure 2.3, lane 1). Column purification using either RiboPure or MEGAclear kits reduced the presence of these small fragments more effectively than did the RNeasy mini kit (Figure 2.3, lanes 2 to 4). Brownish material, possibly arising from plant phenolics, remained visible after isopropanol precipitation but was removed after column purification. The MEGAclear kit was chosen and used in further RNA extractions from rumen solids.

Isolated total RNA was stable when stored in nuclease-free water at -80 °C for at least 3 months. Up to four freeze-thaw cycles were applied without changing the RNA concentration, LSU/SSU rRNA ratio, or RIN, as indicated by the Bioanalyzer analysis on the stored samples (data not shown).
2.3.3 Effect of RNAprotect bacteria reagent on RNA isolation from rumen solids

RNAprotect was added in preliminary investigations to obtain total RNA from rumen solids, but this approach failed to yield high quality RNA. Subsequently, we investigated three approaches for inclusion as well as exclusion of RNAprotect from our procedure for the extraction of RNA from rumen solids. Samples that were extracted without RNAprotect yielded high quality total RNA, whereas those that included RNAprotect yielded similar quantity, but poorer quality RNA (Figure 2.4). All three methods that included RNAprotect resulted in an rRNA ratio of <1.0. Treatment III (complete thaw of sample after sample disruption, followed by removal of RNAprotect before the addition of TRIzol) seemed to be particularly unfavourable, as one of the samples was almost completely degraded, and had an rRNA ratio of only 0.2 (Figure 2.4, lane 8). This may have occurred because the disrupted cells were thawed without sufficient protection before the addition of TRIzol.

2.3.4 RT-PCR and detection of glycoside hydrolases

Internal fragments of glycoside hydrolase genes (celF, xynD and cel3) were amplified by RT-PCR with the fragment size corresponding to the predicted amplicon length. The 3 kb fragment of celF was also detected (Figure 2.5). However, the nearly full length amplicons of xynD and cel3 were not detected. All the negative controls that lacked reverse transcriptase did not show bands corresponding to the amplicons of interest, confirming that positive amplicons were reverse transcribed from mRNA. However, negative RT-PCR reactions from the template before DNase I treatment...
generated amplicons of all three gene fragments, suggesting that trace amounts of DNA were still present in the RNA sample (data not shown).

2.4 Discussion

The present study describes a rapid and effective method to isolate high quality, highly representative total RNA from both liquid and solid ruminal contents.

The LRCI and SRCl methods were simple as no separation step of microorganisms from feed particles was necessary. This reduced the amount of time the samples were exposed to oxygen (which damages the anaerobic microbes) and decreased the time before the samples were frozen, thus reducing the extent of RNA degradation during the separation procedure. Procedures for the extraction of RNA from rumen solid and liquid contents were virtually identical, with the exception that a more vigorous cell disruption process was employed for solid rumen contents. Microbial cells within the liquid phase of rumen contents were efficiently lysed by simple bead beating at room temperature under the protection of TRIzol reagent. However, the quality and quantity of RNA isolated from solid rumen content by the bead beating procedure was far lower. Thus, it was necessary to disrupt microbial cells within the biofilms and particles of solid contents by grinding in liquid nitrogen, as the solids contained large undigested feed particles. These particles could not be efficiently disrupted by bead beating, and prevented TRIzol reagent from penetrating rapidly for sufficient protection. It is important to efficiently break down the microbial cells that are attached to the surface as well as those that have penetrated into the interior of feed particles, since most fibrolytic microbes colonize in these locations, and it has been shown that the largest diversity of bacteria in the rumen exists in the residual particulate fraction (Kong et al., 2010). Grinding of the whole solid
contents in liquid nitrogen not only allows more RNA of greater quality to be obtained, but would also eliminate to a great extent the bias in the microbial populations recovered from solid particles by detachment procedures (Martínez et al., 2009b; Ramos et al., 2009).

The RIN was initially designed to help estimate the integrity of total eukaryotic RNA. It is determined not only by the ratio of the ribosomal bands alone, but also by the entire electrophoretic trace of the RNA sample (Schroeder et al., 2006). A RIN of 10 represents a perfectly intact RNA sample. Although the Agilent software also gives a calculated RIN value for prokaryotic RNA, this approach has not been extensively validated in terms of its value as an indicator of the integrity of prokaryotic RNA. In our study, we noticed that the 2100 Expert software was occasionally unable to discern the rRNA peak areas accurately, making it necessary to manually adjust the peak recognition. In these cases, the RINs did not always correspond to the rRNA ratio or apparent RNA integrity. For example, two SRCI-isolated RNA samples (Figure 2.4, lane 1 and 2) showed similar concentrations, with the same rRNA ratio of 1.7. But under default software parameters, the software only gave a RIN of 5.9 for the sample in lane 2, as opposed to a RIN of 8.2 for the sample in lane 1. However, the RNAprotect samples (Figure 2.4, lane 3 to 8) had rRNA ratios of < 1.0, but had higher RINs (8.4 to 8.6). As the Agilent Bioanalyzer software does not possess a module for mixed rRNA samples, the generated RINs should be regarded as approximate.

According to the manufacturer, RNAprotect bacteria reagent was designed to prevent both degradation of RNA transcripts and induction of genes, and thus provide immediate stabilization of the gene expression profile of bacteria. However, at least in
this study, this reagent reduced both RNA yield and quality. Immediate freezing and grinding of samples in liquid nitrogen followed by mixing in TRIzol reagent yielded RNA of high quality and stability. The addition of RNAprotect to the procedure increased the complexity of the extraction. It resulted in lower RNA yield and quality, whether or not it was removed prior to the addition of TRIzol. Therefore, in the present study we found no value to including RNAprotect in the RNA extraction procedure for either the liquid or solid fraction of rumen contents.

RT-PCR showed that three typical GH genes from *F. succinogenes* S85 were amplifiable from the isolated total RNA. As was found by Béra-Maillet et al. (2009), the three approximately 200 bp internal fragments of *celF*, *xynD* and *cel3* were all detected. A 3000 bp length of the *celF* gene was amplified by using W series primers (i.e., whole-length primers), but no full length amplicons of *xynD* and *cel3* were generated. A possible explanation for this may be related to the fact that random hexamer primers were used for reverse transcription. These reaction conditions are unavoidably biased towards the generation of fragmented, short length first strand cDNA products as opposed to whole length sequences. It was possible that compared with the other two genes, a high transcription level of *celF* existed in the isolated total RNA and resulted in enough whole length cDNA product to be detectable by PCR amplification. We selected these three genes based upon the previous quantitative reverse transcriptase-PCR (RT-qPCR) results by Béra-Maillet et al. (2009), which showed that these genes are highly expressed in rumen contents. According to these authors, *celF* had the lowest transcript level of the three highly transcribed genes under their experiment conditions. However, our study analyzed solid ruminal contents from muskoxen, in which the distribution of, and relation
to known *Fibrobacter* species is unknown. CelF (previously named EGF) was reported to be one of the major cellulose binding proteins identified in *F. succinogenes* (McGavin and Forsberg, 1988; Mitsumori and Minato, 1995). Highly similar gene fragments have also been found in strains of *Fibrobacter intestinalis* (Béra-Maillet et al., 2004; Qi et al., 2005). Considering that the *Fibrobacter* genus is a major contributor to fibrolytic activity within the rumen (Stewart et al., 1997), the detection of *celF* gene expression would seem probable in most rumen systems. The detection of the 3 kb *celF* fragment attests to the integrity of the isolated RNA.

Using the SRCI procedure described in this manuscript, sequencing of eukaryotic polyadenylated mRNA isolated from rumen of muskoxen was carried out using the Illumina Genome Analyzer (Illumina, San Diego, California). Detailed sequence analysis is described in Chapter 3.

Metatranscriptomics is a rapidly emerging field and has shown considerable potential as a means of identifying novel biocatalysts (Sorek and Cossart, 2010; Warnecke and Hess, 2009). Our method makes it practical to obtain large quantities of high-quality total RNA, enabling sequence-based approaches to identify genes coding for novel enzymes from environmental samples. The procedure could be easily adapted to other environments, such as compost, leaf cutter ant gardens, or soil, with little difficulty. For the rumen environment, the method enables the investigation of the impact of the host, diet, and other conditions that affect ruminal function and gene expression within this unique ecosystem.
### 2.5 Tables and Figures

**Table 2.1** Primers used for RT-PCR.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Gene</th>
<th>Primer</th>
<th>Sequence</th>
<th>Product size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EGF</td>
<td>celF</td>
<td>celFWF</td>
<td>GTCCGCATCTGGCTGTGTA</td>
<td>3053</td>
</tr>
<tr>
<td></td>
<td></td>
<td>celFWR</td>
<td>CTTGCCGACCTTGATACCC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>celFIF</td>
<td>CAAGAACGTTGGCGAATC</td>
<td>186</td>
</tr>
<tr>
<td></td>
<td></td>
<td>celFIR</td>
<td>CGGTGTGTGTCACAGAG</td>
<td></td>
</tr>
<tr>
<td>XynD</td>
<td>xynD</td>
<td>xynDWF</td>
<td>GCCGCATGACGTACTTT</td>
<td>2505</td>
</tr>
<tr>
<td></td>
<td></td>
<td>xynDWR</td>
<td>GTGCAGGAGCCAAATACCT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>xynDIF</td>
<td>GGCAAGAAGTGGACCTT</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td></td>
<td>xynDIR</td>
<td>TGTCCTGGTCGTAGTCCT</td>
<td></td>
</tr>
<tr>
<td>Cel3</td>
<td>cel3</td>
<td>cel3WF</td>
<td>CATAAAAACGACCTCCCAAAT</td>
<td>2156</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cel3WR</td>
<td>ATTGCAGCATTCCCTGTACT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>cel3IF</td>
<td>AGCGATGGTAAGGTACTGC</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td></td>
<td>cel3IR</td>
<td>GTGATGGTGCGTAGTCC</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.2 Comparison of RNA yield and quality isolated from the rumen fluid (RF) by using different isolation procedures.

<table>
<thead>
<tr>
<th>Protocols</th>
<th>Total RNA yield (µg (ml RF)^{-1})</th>
<th>Agilent bioanalyzer analysis</th>
<th>Absorbance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LSU/SSU rRNA ratio</td>
<td>RIN value</td>
</tr>
<tr>
<td>Method K</td>
<td>38.0</td>
<td>0.7</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>39.0</td>
<td>0.8</td>
<td>4.9</td>
</tr>
<tr>
<td>Method B</td>
<td>140</td>
<td>1.4</td>
<td>8.1</td>
</tr>
<tr>
<td></td>
<td>163</td>
<td>1.2</td>
<td>8.2</td>
</tr>
<tr>
<td>LRCI without column</td>
<td>232</td>
<td>1.6</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td>201</td>
<td>1.7</td>
<td>9.4</td>
</tr>
<tr>
<td>LRCI</td>
<td>182</td>
<td>1.8</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>162</td>
<td>1.8</td>
<td>9.5</td>
</tr>
</tbody>
</table>
**Figure 2.1** RNA isolation experimental flowchart.

Bold indicates the optimal procedural steps established in this study.

LRCI, Liquid Ruminal Contents Isolation;

SRCI, Solid Ruminal Contents Isolation.
Figure 2.2 Analysis of total RNA integrity extracted from liquid phase and solid phase of ruminal contents.

A: Total RNA extracted from liquid phase rumen contents from a cow using LRCI (Liquid Ruminal Contents Isolation, LSU/SSU ratio: 1.8; RIN (RNA integrity number): 9.5)

B: Total RNA extracted from solid phase rumen contents from a muskoxen fed triticale straw using SRCI (Solid Ruminal Contents Isolation, LSU/SSU ratio: 1.9; RIN: 9.8)
Figure 2.3 Comparison of column purification of RNA.

Total RNA was isolated from solid rumen contents either without column purification or purified using three different commercial kits as described in section 2.2.3. RNAprotect bacteria reagent was not included in the extraction. RNA was analyzed using an Agilent 2100 bioanalyzer and RNA6000 nano chip. The gel pherogram was generated with the 2100 Expert software. The arrow corresponds to the 5S RNA region.

L: RNA ladder (RNA 6000 Nano kit);

Lane 1: Total RNA without column purification;

Lane 2: Total RNA purified with Ribopure kit (Ambion);

Lane 3: Total RNA purified with MEGAclear kit (Ambion), i.e., SRCI (Solid Ruminal Contents Isolation) procedure;

Lane 4: Total RNA purified with RNeasy mini kit (Qiagen).
<table>
<thead>
<tr>
<th></th>
<th>SRCI (without RNAprotect)</th>
<th>RNAprotect added</th>
<th>Treatment I</th>
<th>Treatment II</th>
<th>Treatment III</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA conc. (μg ml⁻¹)</td>
<td>210 214</td>
<td></td>
<td>74</td>
<td>95</td>
<td>53</td>
</tr>
<tr>
<td>RNA yield (μg (g RC)⁻¹)</td>
<td>84.0 85.6</td>
<td></td>
<td>59.2</td>
<td>76.0</td>
<td>42.4</td>
</tr>
<tr>
<td>LS/SS rRNA ratio</td>
<td>1.7 1.7</td>
<td></td>
<td>0.8</td>
<td>0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>RIN</td>
<td>8.2 5.9</td>
<td></td>
<td>8.5</td>
<td>8.6</td>
<td>8.5</td>
</tr>
</tbody>
</table>

**Figure 2.4** Effect of different RNAprotect treatments as described in section 2.2.2.3, on total RNA yield and quality from rumen solids (RS), and comparison with SRCI (Solid Ruminal Contents Isolation) results.

For each treatment there were two samples, hence two lanes per treatment.

Lane Ladder: the RNA ladder (RNA 6000 Nano kit).

NA: the RIN (RNA integrity number) was not provided by the software.
Figure 2.5 RT-PCR amplification of three *Fibrobacter succinogenes* S85 transcripts.

L: DNA ladder;

—: gene-specific negative controls (without addition of reverse transcriptase);

PCR—: PCR negative control.
Chapter 3 Snapshot of the eukaryotic gene expression in muskoxen rumen – a metatranscriptomic approach†

3.1 Introduction

Within the gastrointestinal tract of herbivores a diverse group of anaerobic microorganisms, including bacteria, archaea and eukaryotes, produces a vast array of lignocellulolytic enzymes that in turn digest complex plant cell wall polysaccharides and ferment the released simple sugars. The resulting volatile fatty acids and microbial protein are a source of carbon, nitrogen and energy for the host (Flint, 1997; Russell and Rychlik, 2001). Substantial efforts have been made to understand polysaccharide digestion within the rumen through isolation and identification of cellulolytic species, characterization of their enzymes (Krause et al., 2003), and sequencing the genomes of the major culturable rumen bacteria (Berg Miller et al., 2009; Cai et al., 2010; Flint et al., 2008; Kelly et al., 2010; Morrison et al., 2010; Purushe et al., 2010). The recent introduction of massively parallel sequencing technologies has enabled the sequencing of herbivore gut microbiomes, including the foreguts of cattle and wallabies (Brulc et al., 2009; Hess et al., 2011; Pope et al., 2010) and the hindgut of termites (Warnecke et al., 2011).


1 These authors contributed equally to this manuscript.
These studies have led to the identification of novel cellulolytic enzymes, many of which quite likely arise from microbes that are unculturable in the laboratory (Flint et al., 2008).

Despite the prolific activity directed at understanding the rumen microbiome, there is a distinct lack of information about the eukaryotic component of the rumen metagenome and no rumen fungal or protozoal genomes have been reported. Only a small portion of putative genes described in previous metagenomic studies were attributed to eukaryotes (Brulc et al., 2009; Hess et al., 2011; Pope et al., 2010), although the role of anaerobic fungi (Neocallimastigomycota) and rumen protozoa (Litostomatea) in rumen cellulose digestion is widely recognized (Orpin and Joblin, 1997; Williams and Coleman, 1997). The paucity of genomic information about anaerobic eukaryotes in the rumen is likely a consequence of 1) the low abundance of eukaryotic DNA in the rumen metagenome; 2) the inadvertent exclusion of eukaryotic species by sample preparation methods; and 3) the failure of bioinformatic analyses to annotate novel eukaryotic gene sequences.

Rumen anaerobic fungi not only produce highly active fibrolytic enzymes targeting the plant cell walls, but they also physically disrupt plant cell walls including the cuticle via penetrating rhizoids. Zoospores, the mobile phase of the fungal life cycle, also preferentially colonize lignin-rich regions of the plant cell wall and upon germination, solubilize these regions to a greater extent than rumen bacteria. Studies have shown that rumen fungi may account for up to 8~20% of the total rumen microbial biomass in ruminants consuming forage (Orpin and Joblin, 1997; Rezaeian et al., 2004). A recent
study demonstrated that anaerobic fungi are widely distributed in both ruminant and non-ruminant herbivores (Liggenstoffer et al., 2010).

The effects of rumen protozoa on fiber digestion are less clear. Previous studies are inconsistent and reports on the effects of protozoa range from a 50% increase to a 15% decrease in fiber digestion (for review, see Williams and Coleman, 1997). Studies on the contribution of rumen protozoa to fiber degradation have also been hampered by the difficulty in cultivating protozoa in vitro without the presence of associated bacteria. However, glycoside hydrolase activities and genes have been identified from these organisms (Béra-Maillet et al., 2005; Devillard et al., 1999).

Identification of potent lignocellulolytic and other carbohydrate active enzymes are of great interest for industrial processes, such as cellulosic ethanol production (Percival Zhang et al., 2006; Tilman et al., 2009). In the present study we applied mRNA-Seq technology (Wang et al., 2009) to target the polyadenylated eukaryotic mRNA from the muskoxen (Ovibos moschatus) rumen microbial consortium. Muskoxen are arctic ruminants that live primarily in northern Canada, Alaska and Greenland. For the majority of the year, their food sources are limited to forages high in lignocellulose content, due to the very short arctic summer (Barboza et al., 2006). Consequently, we speculated that muskoxen have evolved to harbour a microbial community that efficiently degrades plant cell wall fiber (Crater et al., 2007). The sampled muskoxen were maintained at an isolated wildlife research facility (R.G. White Large Animal Research Station, University of Alaska Fairbanks, Fairbanks, AK) and had not been in contact with domestic ruminants. It was hypothesized that the metatranscriptome approach would lead to the identification of genes coding for novel enzymes and also, for the first time, provide
information on the expression of carbohydrate active enzymes by the eukaryotic community in the rumen.

3.2 Materials and Methods

3.2.1 Ethics statement

The animals were cared for according to procedures that were approved under protocol No. #139821-2 by the Institutional Animal Care and Use Committee at the University of Alaska, Fairbanks, AK.

3.2.2 Muskoxen rumen sampling

Ruminal samples were obtained from two cannulated muskoxen at the University of Alaska (Fairbanks, AK) in September, 2009. The muskoxen were mature male castrates with mean body mass between 245 and 271 kg. During a 1-month period, the muskoxen were fed a triticale (× Triticosecale) straw or a brome grass hay (Bromus sp.) based high fiber diet, offered twice daily, plus a small amount of protein and mineral supplement once in the morning (335g·d⁻¹; Table 3.1). During the last week of the period, rumen samples were obtained in the morning, before the muskoxen were fed. The ruminal contents were transferred to a heavy walled 250 ml beaker and the solid and liquid phases were separated using a Bodum coffee filter plunger (Bodum Inc., Triengen, Switzerland). Subsamples of solid digesta (~ 2.5 g) were immediately flash-frozen in liquid nitrogen. All samples were frozen within 5 min of the sample being withdrawn from the animal. Samples were immediately transferred to the lab, and stored at -80 °C until further processing.
3.2.3 RNA extraction and purification

Total RNA was isolated from rumen solids according to the method established in Chapter 2. The quality of total RNA was estimated by running the samples on RNA 6000 nano chip on an Agilent 2100 BioAnalyzer.

3.2.4 RNA sequencing and sequence assembly

mRNA-Seq libraries were constructed from 100 µg of total RNA using the Illumina mRNA-Seq sample preparation kit according to the manufacturer’s instructions (Illumina Inc, San Diego, USA). Two samples from two individual muskoxen (one fed triticale straw, one fed brome grass hay) were combined and sequenced using the Illumina Genome Analyzer II system at the McGill University/Genome Quebec Innovation Centre. Obtained sequencing reads were assembled de novo using a combination of Velvet (available at http://www.ebi.ac.uk/~zerbino/velvet/) (Zerbino and Birney, 2008) and CAP3 (available at http://pbil.univ-lyon1.fr/cap3.php) (Huang and Madan, 1999) programs. The initial RNA-seq data set in fastq format was split into 14 separate data sets of 2 million reads each. On each of these split data sets, the Velvet suite of programs was run with three different k-mer parameters; k=37, k=45 and k=53. The resultant files were then concatenated into a single contigs file. The program CAP3 was then run on this file, with default parameters. The files with extensions .contigs and .singlets were concatenated into a single file which represented the assembled transcript contigs in the present study. This method of transcript assembly was selected following extensive experimentation and produces more long contigs containing full length transcripts when compared with traditional assembly methods which are more suited to DNA reads. The
contigs were then clustered together at 95% sequence identity over 80% of their lengths using the DNA version of CD-HIT in the Rammcap package (Li, 2009).

To validate contig assemblies, 20 contigs containing different putative carbohydrate active enzyme genes were amplified by reverse transcription PCR using primers designed specifically for each contig. The target contigs and primer sequences are listed in Table 3.2. Briefly, the muskoxen total RNA samples were reverse transcribed using a Superscript III kit and oligo-dT primers (Invitrogen). PCRs were carried out using Platinum Taq polymerase Hi Fidelity (Invitrogen) using the conditions recommended by the manufacturer.

3.2.5 Sequence analysis

All sequence analyses, unless otherwise specified were performed using both the reads and the assembled contigs. The databases employed for this analysis were the latest versions available during the analysis period (Jun 2010 to Dec 2010). The genome sequence of the rumen bacterium *Fibrobacter succinogenes* S85 (Accession No. NC_013410), expressed sequence tags (ESTs) of rumen protozoa (Ricard et al., 2006) and rumen fungi *Pyromyces* sp. E2 were retrieved from National Center for Biotechnology Information (NCBI) Genbank databases (http://www.ncbi.nlm.nih.gov/).

3.2.6 rRNA sequence identification

Ribosomal RNA sequences were firstly identified by BLASTN searches against LSU and SSU ribosomal RNA databases (Version 104) from the ARB-Silva database (Pruesse et al., 2007). Subsequently, all the sequences were further analyzed by the rRNA-hmm (Huang et al., 2009) and tRNA (transfer RNA)-scan (Lowe and Eddy, 1997)
programs in the Rammcap package (Li, 2009) using the default settings. The sequences that had E-values equal or less than $10^{-5}$ (bit score ≥ 52) and overlap ≥ 50bp to entries in the SSU/LSU database, as well as those identified by rRNA-hmm program (SSU rRNA, LSU rRNA and tRNA) are referred to as non-coding RNA (ncRNA) in this paper.

### 3.2.7 Binning

The functional based taxonomic assignment was constructed by the Metagenomic Analyzer (MEGAN) software (Huson et al., 2007) based upon the best BLASTX hit to an in-house database named as NRMO database (trimmed down non-redundant muskoxen amino acid database). The NRMO database contained all protein sequences in the Genbank non-redundant amino acid (nr) database that had a match to any of our assembled contigs, with an E-value no greater than $10^{-5}$. There were about 230,000 entries in the NRMO database. To validate the NRMO database, 20,000 reads were randomly picked and BLASTed against both the nr database and NRMO database respectively, and compared. The results, especially for the taxonomy distribution at genus level, were very similar to each other (data not shown). Collector’s curves were produced from an *ad hoc* Perl script and plotted in Microsoft Excel version 2003.

For the taxonomic community composition analysis based on the amino acid sequences of actin and elongation factor 1 (EF1), related sequences were retrieved from the nr database. Because related protein sequences from rumen fungal and/or rumen protozoa species were not present in the nr database, actin sequences of *Piromyces* E2 ESTs (Genbank Accession numbers: GT912769, GT909886, and GT912949) and EF1 sequences from Rumen protozoa ESTs (Genbank Accession numbers AM051945, AM053457, AM051946, AM054620, AM051759, and AM053167) were retrieved. The
translated EST sequences were merged into Genbank nr derived sequences. BLASTX searches against each of the two protein family databases were carried out using the non-rRNA reads as the query. The results were analyzed by MEGAN with a bit score cut-off of 52. The taxonomic composition was also estimated by running the software MLTreeMap on the assembled contigs as described (Stark et al., 2010).

Putative full-length ORFs were identified as follows: The assembled contigs were BLASTX-ed against the UniProt database. Contigs were then translated in the proper reading frame based on the BLASTX hits. The resulting amino acid sequences were searched for all full length ORFs of at least 70 amino acids which fully encompass the alignment of the BLAST Hits.

3.2.8 Functional annotation of the coding RNA sequences

The coding RNA sequences were searched using RPS-BLAST against both the KOG and the COG databases and the Genbank nr database. Bacteria-like reads identified by nr BLASTX were further searched against the COG database. The functional roles of the sequences were assigned based on the KOG and COG searches. The matches that had E-values equal or less than $10^{-5}$ were considered significant.

3.2.9 Carbohydrate active protein annotation

Lignocellulolytic gene containing reads and contigs were identified and classified based on CAZy database (Cantarel et al., 2009) as described by Warnecke and colleagues (Warnecke et al., 2007), with the following modifications. Both HMMER3 (Eddy, 2009) and BLASTX searches were carried out as follows: Step A) Glycoside hydrolase and carbohydrate binding module (CBM) families that have associated Pfam HMMs (v24.0)
(Pope et al., 2010) were used directly for HMMER hmmsearch. Step B) In an attempt to associate Pfam HMMs to CAZy families without such models, all members of these CAZy families were searched against the Pfam-A and Pfam-B databases (v24.0). Results were manually checked and Pfam HMMs were conservatively chosen for a CAZy family only when the following two criteria were met: i) all hits to that Pfam group were from the same CAZy family; ii) At least 80% of group members were identified to conform to the conserved Pfam model. In instances where one Pfam HMM model represented members from two or more closely related CAZy groups, a class of combined CAZy groups was assigned. Step C) For those CAZy families that currently are not represented by a Pfam model, the representative sequences as described by Warnecke et al. (2007) were used in BLASTX searches with a score cut-off of 52. Step D) For CAZy families with neither a Pfam accession nor representative amino acid sequences, an HMM profile was built based on T-coffee (Notredame et al., 2000) alignment of representative members selected from the CAZy web site and used for HMMER3.

3.2.10 Glycoside hydrolase cluster analysis

To compare putative genes coding for carbohydrate active proteins identified in the muskoxen rumen metatranscriptome with other genomes/metagenomes, the percentages of glycoside hydrolase families were calculated. A two-dimensional matrix was constructed, consisting of the GHs that were identified from genomes or metagenomes, wherein each cell in the matrix indicated how often a GH family was seen within a particular sample. Pearson correlation coefficients of each two samples were calculated and transformed into distances and clustered by using the unweighted pair group method
with the arithmetic mean algorithm as previously described (Garcia-Vallve et al., 2000; Qi et al., 2005).

### 3.2.11 Sequence Data Availability

The sequencing reads are available from the NCBI short read archive under Accession number SRA030623.1.

### 3.3 Results

#### 3.3.1 General sequence statistics

In the present study, we adopted a metatranscriptomic approach to identify enzymes from the muskoxen rumen microbial community. We used samples from muskoxen fed triticale straw and brome hay for deep sequencing with the goal to obtain transcripts encoding carbohydrate-active enzymes. Total RNA was extracted from rumen solids to ensure isolation of the cellulolytic microbial biofilm as well as RNA from microbes deeply embedded in the plant fiber. After purification, the eukaryotic mRNA was sequenced using the Illumina Genome Analyzer II platform. This approach resulted in a total of 25,900,806 reads, with an average read length of 108 nt, generating a total of 2.8 gigabases of sequence data (Figure 3.1).

Raw sequencing reads were assembled into a total of 59,129 contigs with an average length of 310 bp, for a total of 19 M base pairs. The maximum length of the contigs was 13,498 bp, which contained a single open reading frame of 13,083 bp. It encoded for a 4,354-amino-acid protein that showed 28% identity to a hypothetical protein from a strain of *Orpinomyces* (Nicholson et al., 2005). The size distribution of all the contigs is illustrated (Figure 3.2). Over 12,000 contigs were represented by 100 or
more reads, including 2,551 contigs represented by more than 1000 reads and 545 represented by more than 5,000 reads. To validate the contig assembly, twenty glycoside hydrolase related contigs (≥500 bp) were chosen and primers designed to amplify the target (Table 3.2). Nineteen of the twenty contigs were successfully amplified using at least one set of primers.

Using BLASTN searches against the Silva LSU and SSU ribosomal RNA databases coupled with rRNA-HMM and tRNA-scan, we identified 4.77 million nc RNA reads or 18.4% of the total (Figure 3.1). During the RNA sample preparation step, oligo-dT coated magnetic beads were used to remove a large proportion of rRNA. Assuming ncRNA account for 95% of the original total RNA (Neidhardt and Umbarger, 1996), approximately 99% of this amount was successfully removed as indicated by the percentage of the ncRNA reads identified.

The average GC content of ncRNA reads was 51%. In contrast, the GC content of all potential protein encoding RNA reads was 39.2%, a value that was significantly lower than the ncRNA reads and much lower than those reported in metagenomic studies of other microbial communities associated with ocean, fresh water and various animal environments (average GC% is 49.56 ± 4.9% (Willner et al., 2009)). The average GC content of the assembled contigs was 37.9%, which were also lower than other metagenomic studies (Figure 3.3).

BLASTX searches using the 21 million potential protein encoding reads against the NRMO database indicated 7.8 million reads had at least one significant match (E-value ≤ 10^{-5}) (Figure 3.1). RPS-BLAST searches against the euKaryotic Orthologous Groups (KOG) and the Cluster of Orthologous Groups (COG) identified 6.0 million reads that
may have conserved domains, of which 0.9 million reads were not identified by BLASTX searches. The remaining reads (48% of all reads) did not show any similarity to any of the above mentioned databases.

The first BLASTX match was used to estimate the possible origin of each putative protein coding RNA reads, according to MEGAN analysis (Figure 3.4). About 6.6 million reads (31%) showed highest similarity to proteins from the Eukaryota kingdom. Among these, rumen anaerobic fungi, which belong to the phylum Neocallimastigomycota (Hibbett et al., 2007), were represented by 1.4 million reads. Rumen ciliate protozoa, which belong to the Litostomatea class of the Alveolata group, were represented by 1.1 million reads. At the genus level, the most represented genera were known rumen/anaerobic species, including *Entodinium*, *Piromyces*, *Neocallimastix*, *Trichomonas*, *Orpinomyces*, *Entamoeba*, and *Epidinium*, that were represented by over 100,000 reads each (Figure 3.5). A total of 0.7 million reads (3.4%) and 0.02 million reads (0.1%) were binned to Bacteria and Archaea kingdoms, respectively.

### 3.3.2 Analysis of sequencing coverage

The sequencing coverage was first assessed by looking at the matches to rumen eukaryotic proteins that were present in the nr database. Rumen anaerobic fungal protein sequences (total of 257) were obtained from the Genbank database as of July, 2010. TBLASTN searches were performed using all of these proteins as queries, and matches to 220 of them (with Expect value ≤ 1E-20) were found in the muskoxen contigs. Another similar search identified 104 of 107 published rumen protozoal protein sequences in our dataset. The identification of almost all known rumen fungal and protozoal sequences demonstrated the comprehensive coverage of the current sequencing approach.
The coverage was further evaluated by plotting collector’s curves based on the number of functional gene categories (matched to the KOG database) and gene accession numbers identified (matched to the NRMO database), as a function of the number of reads (Figure 3.6). Saturating coverage was found for both curves, as roughly 80% of the total richness was found at a point of less than 14% of the sequencing effort.

3.3.3 Functional analysis of the putative protein-coding reads

Based on the RPS-BLAST search results, 6.0 million reads could be assigned to a cluster of the KOG/COG databases. Most of the assignable sequences belonged to the “Translation, ribosomal structure and biogenesis” cluster (45% of all the assigned sequences), while Cytoskeleton (16%) was the second largest cluster (Figure 3.7). Correspondingly, 9 of the top 10 KOG/COGs also belonged to these two categories, with KOG0676 (Actin and related proteins, which were represented by 551,126 reads) and KOG0052 (Translation elongation factor one, 230,087 reads) as the first two abundant KOG groups (Figure 3.8). These results indicate that these groups of genes were actively transcribed. About 18% of the reads that matched the KOG/COG databases were involved in metabolism, including metabolism of carbohydrate (7.9%), energy conversion (4.2%) and metabolism of amino acids (2.6%) (Figure 3.7). KOG/COG groups involved in glycolysis and pyruvate metabolism accounted for the highest read numbers (Table 3.3, Figure 3.9), demonstrating genes belonging to those clusters had a central role in metabolism.

Hydrogenosomes are membrane-bound organelles present in anaerobic eukaryotes including rumen fungi and protozoa (Boxma et al., 2005; Williams, 1986). Hydrogenosomes are distantly related to mitochondria and are the centre of ATP and
hydrogen generation in these microorganisms (Boxma et al., 2005). In the present rumen metatranscriptome dataset, KOG/COGs linking with hydrogenosomes were also identified, including iron-only hydrogenase, malic enzyme, pyruvate:formate lyase, succinyl-CoA synthetase and acetate:succinate CoA-transferase. Combining the highly expressed KOG/COGs, a representative energy metabolism pathway of muskoxen rumen eukaryotes was reconstructed (Figure 3.9).

3.3.4 Lignocellulolytic enzymes

Sequence homology based enzyme annotation was biased toward the identification of known enzymes that were already present in the database. To minimize this bias, we used a more sensitive Pfam-HMM search to identify CAZyme modules (Table 3.4, Table 3.5, Figure 3.10). In all, these analyses identified over 400,000 reads potentially encoding lignocellulolytic enzyme modules, spanning about 110 CAZy families. The read number in each family gives an indication of the expression status of that group. However, it needs to be pointed out that the number of reads that matched to lignocellulolytic enzymes was likely underestimated as the short reads (108 nt, translated into 36 amino acid residues maximum) result in a less robust database search score than that obtained using the full sequence of complete genes.

To circumvent this problem, similar analyses were also performed on the assembled contigs. In total, we identified over 2,500 contigs with a significant match to at least one CAZy module (Table 3.5). Since the use of short contigs may overestimate the total number of enzymes, we further restricted our targets to those contigs longer than 500 base pairs (1,082 in total, Table 3.5, Table S.1, Table S.2). These contigs, especially those assembled from larger numbers of reads, could serve as good candidates of
potentially useful lignocellulolytic enzymes. Only 17% of these contigs were more than 70% identical to sequences in the nr database, while 46% of them had less than 50% identity (Figure 3.11). Seventeen percent of the CAZy module coding sequences identified were most similar to nr database sequences annotated as “(conserved) hypothetical protein”, “predicted protein” or “unnamed protein product”. These results indicate that rumen eukaryotes produce a large variety of CAZymes with many of them remaining uncharacterized. There were 242 contigs that had two or more distinct CAZy domains. The proposed fungal dockerin CBM10 modules were predominant in these predicted multi domain enzymes, and were found in 190 (78.5%) of these contigs. Glycoside hydrolases from families GH6, GH45, and GH48 were found in 25, 25 and 23 multi-domain contigs respectively, most of which were linked to CBM10 modules.

3.3.4.1 Catalytic modules

Most of the muskoxen rumen microbiome cellulases identified were classified as families GH5, 6, 9, 45 and 48, which were represented by 30 to 51 contigs. Similar to other rumen metagenomic studies (Brulc et al., 2009; Hess et al., 2011; Pope et al., 2010), no contigs showed similarity to family GH7 or GH44. In this study, putative swollenin modules were detected in 16 contigs. Swollenin was reported to dissociate cellulose fiber with no hydrolytic activity (Saloheimo et al., 2002) and has not been previously reported to be associated with anaerobic microorganisms. Hemicellulose degrading enzymes from GH8, GH10, GH11, GH26 and GH53 were also identified. GH10 and GH11 were the predominant families that were represented by 29 and 33 contigs, respectively. Carbohydrate esterases remove the ester bond on the xylan backbone, exposing it to glycoside hydrolases. There were 111 contigs showing similarity to carbohydrate
esterases in the CAZy database. The Family CE4 family was the largest family with 50 contigs. The CE1 family, encoding feruloyl esterases important for lignin solubilization in the rumen, was represented by 13 contigs. Twenty seven contigs were annotated to be polysaccharide lyases (Table 3.5).

GH family profiles recovered from the metatranscriptome of muskoxen rumen were compared to the metagenomes of the termite hindgut (Warnecke et al., 2007), wallaby foregut (Pope et al., 2010) and bovine rumen (Brulc et al., 2009; Hess et al., 2011), as well as the genomes of several anaerobic bacteria and expressed sequence tags (ESTs) of the rumen fungi *Piromyces* sp. E2 (Table 3.5, Figure 3.10). Clustering analysis of the GH family distribution (Figure 3.10) showed that the muskoxen metatranscriptome was most closely related to anaerobic cellulolytic bacteria and the rumen fungus *Piromyces*. There were some major differences between putative genes identified by our study and previous metagenome sequencing studies. Approximately 28% of identified GHs were cellulases in the muskoxen rumen transcriptome, as compared to 8.5% or less in reported rumen metagenome studies. A large number of genes potentially encoding GH6, GH48 and swollenin enzymes identified in the present study were rarely found in other studies of rumen metagenomes. Conversely GH29, 35, 39 and 42 family members described in other rumen metagenome studies (Brulc et al., 2009; Hess et al., 2011; Pope et al., 2010) were not identified in the rumen metatranscriptome of muskoxen (Figure 3.10).

### 3.3.4.2 Accessory modules

The most abundant accessory module was CBM10, which was identified from 403 contigs longer than 500 bp (Table 3.4). CBM10 has been shown to be associated with many rumen fungal glycoside hydrolases. It was once proposed to be a eukaryotic
counterpart of bacterial cellulosomal dockerin (Steenbakkers et al., 2001), but a recent study (Nagy et al., 2007) suggested that the structure of this domain differs from bacterial dockerins. Dockerin containing cellulases are known to interact with cellulosome scaffolding proteins in bacteria to form the cellulosome structure. However, a cellulosome scaffolding protein has yet to be identified from rumen fungi. The exact function of CBM10 modules remain to be explored.

CBM18, known as a chitin binding module, was the second largest CBM group, and was identified from 108 contigs. Other major CBMs identified included CBM1, CBM6 and CBM13, that presented in 33, 27 and 31 contigs, respectively. Accessory modules that are commonly found in bacterial cellulases, such as bacterial cellulosome dockerin, cohesin, S-layer homolog domain (SLH), cellulase N-terminal immunoglobulin domain (CelD_N) and fibronectin-3 (fn3) modules, were not found in the contigs or reads.

3.3.4.3 Lignocellulolytic gene expression

Both gene diversity and gene expression information can be obtained from metatranscriptomic sequencing analysis. To demonstrate the latter, we summarized the read numbers associated with contigs/genes (Table 3.6, Table S.1, Table S.2). Cellulase and xylanase coding sequences in GH families 6, 11, 45 and 48 were over-represented compared to other families (Table 3.6). Putative GH1 genes, which encode oligosaccharide degrading enzymes, were also over-represented. In addition, a putative family 6 polysaccharide lyase, which has never been reported from eukaryotes, was represented by over 20 k reads.
3.4 Discussion

Cellulase activities are well known to be present in microbial communities found in soil, compost, and herbivore digestive tracts including the rumen. Metagenomic technology coupled with high throughput sequencing has enabled the identification of thousands of sequences encoding for enzymes that degrade plant cell walls (Allgaier et al., 2010; Brulc et al., 2009; Hess et al., 2011). In the present study, we used rumen solid samples from muskoxen fed a highly lignified diet of triticale straw or brome grass hay. Rumen solids are reportedly responsible for 80~90% of the endoglucanase and xylanase activities in the rumen (McAllister et al., 1994; Williams and Strachan, 1984), as attachment and the formation of digestive microbial biofilms is a prerequisite for the ruminal digestion of feed (McAllister et al., 1994). By applying an improved isolation method, high quality RNA was extracted from rumen solids (Chapter 2). Combined with an efficient method of de novo assembly of short sequence reads, the present study has provided a comprehensive catalogue of eukaryotic cellulolytic enzymes in the muskoxen rumen, many of which are supported by full-length cDNA information.

To our knowledge, we are the first group to report the whole eukaryotic transcriptome of a rumen microbial community. Metatranscriptomic studies have been previously carried out in various microbiomes focused on marine microbial communities (Frias-Lopez et al., 2008; Gifford et al., 2011; Gilbert et al., 2008; Poretsky et al., 2009), and the gut microbial population of the piglet (Poroyko et al., 2010). Most of these studies used Roche 454 pyrosequencing technology and although the read lengths were longer than that obtained with Illumina sequencing, far fewer total sequences were produced. In fact most of these studies produced less than 500 M bp of sequences and
sample-sequencing depth was low (Gifford et al., 2011). Illumina sequencing in the present study generated 2.8 gigabases of sequencing data, which is at least 6 times that of previous metatranscriptomic studies (Frias-Lopez et al., 2008; Gifford et al., 2011; Gilbert et al., 2008; Poretsky et al., 2009; Poroyko et al., 2010). The deep sequencing coverage obtained in this study is attested by the observation that the slopes of the collector’s curves reached a plateau and that most of the rumen eukaryotic sequences reported in Genbank were identified in the metatranscriptome (331 out of 364 sequences).

Metatranscriptomics has distinct properties when compared to metagenomics, the first being that metatranscriptomic analysis identifies most extensively transcribed genes while metagenomic sequencing identifies the most numerically dominant genes. For example, *Prevotella* is a group of predominant rumen bacteria that at times account for as much as 60% of the total bacteria in the rumen (Kong et al., 2010; Stevenson and Weimer, 2007), even though they play no active role in the digestion of recalcitrant cellulose (Purushe et al., 2010). Indeed, the GH profiles of *Prevotella ruminicola* clustered closely to those of the wallaby and bovine rumen metagenomes, with GHs involved in the degradation of oligosaccharides and hemicelluloses being highly represented, whereas the proportion of GHs related to cellulase were much lower (Figure 3.10). If a gene encoding a GH was present in a microbial species of low rumen abundance, even if highly expressed, it would be unlikely to be recovered by metagenomic sequencing. Such a scenario may be applicable to the rumen anaerobic fungi as they are reported to account for no more than 8% of the total rumen microbial biomass (Orpin and Joblin, 1997).

Secondly, metatranscriptomic analysis may provide insight into the degree of gene expression (Table 3.6), which could help focus gene mining towards those enzymes that
are most active in plant cell wall degradation. Consequently, the yield of potential gene
targets for further development may be far higher with metatranscriptomic than
metagenomic approaches. A recent rumen metagenomics study using Illumina
sequencing technology generated 268 gigabases of metagenomic DNA with about 103
putative carbohydrate active enzymes identified per gigabase (Hess et al., 2011). In our
present study, we were able to identify 2500 candidates in 2.8 gigabases of RNA
sequence, translating to 893 putative carbohydrate active enzymes per gigabase.

When muskoxen are fed on high fiber diets, cellulolytic microorganisms attach to
and penetrate the fiber and expression of many of their cellulolytic enzyme genes is
induced. Genes that are highly expressed generate more sequencing reads, increasing
sequence coverage, resulting in the assemblage of longer contigs and thereby increases
the likelihood of regenerating full length genes or modules. Indeed, of the 59,129 contigs
in the present study, there were over 2,500 contigs with lengths over 1.0 kb and 6,022
with lengths between 0.5 to 1 kb (Figure 3.2). Among the contigs longer than 500 bp,
over 1,000 of these harboured sequences putatively encoding for carbohydrases. Most
putative CAZy genes longer than 500 bp (96%) were represented by 100 or more reads
(Table S.1, Table S.2), corresponding to an average sequence coverage of about 139
times. These highly expressed putative CAZy genes are likely to play critical roles in the
breakdown of plant fiber by rumen eukaryotes and have potential for use in cellulosic
biofuel production as well as other industrial processes.

Cellulases from different families often have different substrate specificity and
reaction mechanisms. Efficient degradation of the plant cell wall requires synergistic
interactions between enzymes with high activity for different substrates (Lynd et al.,
2002). Not surprisingly, we identified cellulases from a wide range of CAZy families. The range of glycoside hydrolases identified showed remarkable differences as compared to previous rumen metagenomic studies. For example, a large number of putative GH6, GH48, and Swollenin genes were identified. All three GH families are important for the degradation of crystalline cellulose, which is the most recalcitrant part of plant cell walls. However, these GH families were not described by earlier metagenomic approaches in the bovine rumen or wallaby foregut (Brulc et al., 2009; Pope et al., 2010). Even in the most recent deep metagenome sequencing study (Hess et al., 2011), only three putative GH48 genes were identified (Hess et al., 2011), while GH6 and Swollenin were not found. In contrast, our study identified 31 GH6, 33 GH48 and 16 Swollenin sequences from rumen eukaryotes and GH6 and GH48 were among the most actively transcribed families (Table 3.5). These data clearly suggest that rumen eukaryotes play an important role in crystalline cellulose digestion through expression of a large amount of exo-glucanases, which were not detected in other rumen metagenomic studies.

The different CAZy families identified by rumen metagenomic studies and our metatranscriptomic study could be due not only to different nucleic acid sequencing targets, but also to the differences in the samples. Muskoxen could have developed a different plant cell wall degrading rumen microflora for survival in the arctic as compared to domesticated cows and sheep. Indeed, our preliminary analysis has identified differences in the microbial population between muskoxen and domesticated cattle (Forster, RJ, personal communication).

Assessment of the combined rumen/gut microbiome sequencing information across studies would suggest that the rumen seems to lack cellulases from GH7 and GH12
families. So far, all members of GH7, a family of exo-glucanases, have been isolated from aerobic fungi. The GH7 activity of releasing cellobiose from the reducing end of the cellulose chain may be undertaken by GH48 cellulases within the rumen. Enzymes in family GH12 have been shown to have endoglucanase and xyloglucases activities (Gloster et al., 2007). Functional aspects of these enzymes may arise from GH74 enzymes in the rumen (Yaoi et al., 2007).

A total of 3.4% reads showed top BLASTX matches to bacteria (Figure 3.4) and over half of these reads exhibited the highest similarity to proteins from bacteria that are known as predominant rumen/gut residents, such as families of *Fibrobacteraceae, Clostridiaceae, Ruminococcaceae, Prevotellaceae, Bacteroidaceae*, and *Lachnospiraceae* (Figure 3.12). However, these bacteria-like reads are very unlikely to come from bacterial mRNA because only enriched polyA RNAs were sequenced, which were rarely found in bacteria mature mRNA. Phylogenetic binning using protein markers also confirmed the absence of bacteria-derived genes in our dataset (Figure 3.13). In addition, the GC content of the “bacteria-like” genes identified from muskoxen transcriptome were about 40.1%, which is also within the range of rumen eukaryotic coding sequences identified, but much lower than the average GC content of the metagenomic studies which represent the bacterial population (Figure 3.3). These findings imply that bacteria-like coding sequences, most of which are involved in a variety of metabolic functions (Figure 3.7) (including 35% of all putative CAZy genes identified), may have been horizontally transferred into the genome of rumen eukaryotes, most likely from rumen bacteria, a possibility that has been previously raised (Garcia-Vallve et al., 2000; Ricard et al., 2006).
Anaerobic fungi and protozoa are the two major groups of eukaryotes in the rumen (Hungate et al., 1964; Orpin, 1975). In the present study, functional based phylogenetic binning (Figure 3.4), top BLAST matches of the CAZy enzymes (Figure 3.14) and phylogenetic analysis based on SSU ribosome RNA sequence (data not shown), all indicated the presence of both groups. There were significantly ($\chi^2=348, p<0.0001$) more CAZy enzymes matching to rumen fungi than to rumen protozoa (Figure 3.14), indicating that rumen fungi may play a more significant role in fiber digestion in the muskoxen rumen. However, since there are more CAZy genes from rumen fungi than from protozoa in the nr database (101 vs 28, Table 3.5), the differences could be due in part to the number of homologues currently present in the database.

Eukaryotic anaerobic microbes are poorly understood, especially from a molecular perspective. Although this study focused primarily on genes encoding enzymes involved in plant cell wall degradation, the data presented greatly expands our current knowledge of these unique eukaryotes and should provide further insight into their co-evolution, metabolism and function within the rumen microbial community.
### 3.5 Tables and Figures

**Table 3.1 Feed composition.**

<table>
<thead>
<tr>
<th>Feed</th>
<th>DM (^1) (g·g(^{-1}))</th>
<th>NDF (^2) (g·(g DM(^{-1})))</th>
<th>ADF (^3) (g·(g DM(^{-1})))</th>
<th>Cellulose (g·(g DM(^{-1})))</th>
<th>Hemicellulose (g·(g DM(^{-1})))</th>
<th>Lignin (g·(g DM(^{-1})))</th>
<th>Nitrogen (g·(g DM(^{-1})))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hay</td>
<td>0.875</td>
<td>0.699</td>
<td>0.378</td>
<td>0.308</td>
<td>0.321</td>
<td>0.066</td>
<td>0.0112</td>
</tr>
<tr>
<td>Straw</td>
<td>0.878</td>
<td>0.837</td>
<td>0.522</td>
<td>0.437</td>
<td>0.315</td>
<td>0.074</td>
<td>0.00673</td>
</tr>
<tr>
<td>SEM</td>
<td>1.82 ( \times ) 10(^{-3})</td>
<td>0.0143</td>
<td>0.0150</td>
<td>0.0175</td>
<td>2.40 ( \times ) 10(^{-3})</td>
<td>0.0135</td>
<td>0.0464</td>
</tr>
<tr>
<td>P =</td>
<td>0.49</td>
<td>0.02</td>
<td>0.02</td>
<td>0.007</td>
<td>0.19</td>
<td>0.71</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

\(^1\) DM: dry matter.

\(^2\) NDF: neutral detergent fiber.

\(^3\) ADF: acid detergent fiber.
### Table 3.2 Primers used for validating lignocellulolytic enzyme related contigs.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer sequence</th>
<th>Target Contig</th>
<th>Contig Length</th>
<th>Position</th>
<th>Theoretical Product Size</th>
<th>CAZY Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>30088_62_outerF</td>
<td>ATGGTGTTGATACCAACACTCTTG</td>
<td>Contig30088</td>
<td>1072</td>
<td>62</td>
<td>941</td>
<td>CE1</td>
</tr>
<tr>
<td>30088_70_outerR</td>
<td>CCCATTCTACCGTCACCTTC</td>
<td></td>
<td></td>
<td>1003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30088_265_innerF</td>
<td>AGTCTTAAGGATACACCAACCC</td>
<td></td>
<td></td>
<td>265</td>
<td>423</td>
<td></td>
</tr>
<tr>
<td>30088_385_innerR</td>
<td>AGATCAAAGGCTGATGGAGCAG</td>
<td></td>
<td></td>
<td>688</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29149_62_outerF</td>
<td>TTACCATTACCTTACCGTGACCTC</td>
<td>Contig29149</td>
<td>1132</td>
<td>62</td>
<td>967</td>
<td>CE1</td>
</tr>
<tr>
<td>29149_104_outerR</td>
<td>TTTCCCAGGCGGCAGATGG</td>
<td></td>
<td></td>
<td></td>
<td>1029</td>
<td></td>
</tr>
<tr>
<td>29149_200_innerF</td>
<td>GCCCATTCAGCTAAGTTACCC</td>
<td></td>
<td></td>
<td>200</td>
<td>718</td>
<td></td>
</tr>
<tr>
<td>29149_215_innerR</td>
<td>GGGTGGATTCACTCAAGATGA</td>
<td></td>
<td></td>
<td>918</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2424_106_outerF</td>
<td>GAGCACCAACACAAAGCAGTAG</td>
<td>Contig2424</td>
<td>1033</td>
<td>106</td>
<td>832</td>
<td>CE3</td>
</tr>
<tr>
<td>2424_96_outerR</td>
<td>GTGGTATGGTGTTATGTTG</td>
<td></td>
<td></td>
<td>938</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2424_280_innerF</td>
<td>CCTGATCAGTACCCATACAGC</td>
<td></td>
<td></td>
<td>280</td>
<td></td>
<td></td>
</tr>
<tr>
<td>111_106_outerF</td>
<td>GGGTGTACAGTGAATACACCC</td>
<td>Contig111</td>
<td>1476</td>
<td>106</td>
<td>1255</td>
<td>GH6</td>
</tr>
<tr>
<td>111_116_outerR</td>
<td>CCAAGATCTCGGACACCTTG</td>
<td></td>
<td></td>
<td></td>
<td>1361</td>
<td></td>
</tr>
<tr>
<td>111_315_innerF</td>
<td>GGGTGATCAGAACAAATGTG</td>
<td></td>
<td></td>
<td>315</td>
<td>856</td>
<td></td>
</tr>
<tr>
<td>111_300_innerR</td>
<td>ACCAGGTACAGAAGCGTTG</td>
<td></td>
<td></td>
<td>1171</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22047_114_outerF</td>
<td>GGCATGGATAGCAAAAGATTG</td>
<td>Contig22047</td>
<td>1921</td>
<td>114</td>
<td>1694</td>
<td>GH6</td>
</tr>
<tr>
<td>22047_114_outerR</td>
<td>TGGAGCCTGTGTTGTATGCG</td>
<td></td>
<td></td>
<td></td>
<td>1808</td>
<td></td>
</tr>
<tr>
<td>22047_450_innerF</td>
<td>TGCTTTGCCGTAAGGCTTC</td>
<td></td>
<td></td>
<td>450</td>
<td>1172</td>
<td></td>
</tr>
<tr>
<td>22047_300_innerR</td>
<td>ACCCTGGTGCTACGGTGTC</td>
<td></td>
<td></td>
<td>1622</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30327-O1</td>
<td>GGAATTGTTTCTTGGGGTGTTG</td>
<td>Contig30327</td>
<td>1485</td>
<td>98</td>
<td>1252</td>
<td>CE1</td>
</tr>
<tr>
<td>30327-O2C</td>
<td>CCGGAGTGTGTTTCTTTTG</td>
<td></td>
<td></td>
<td></td>
<td>1350</td>
<td></td>
</tr>
<tr>
<td>30327-I1</td>
<td>GGGTGTTACAAACAAACAACC</td>
<td></td>
<td></td>
<td>420</td>
<td>651</td>
<td></td>
</tr>
<tr>
<td>30327-12c</td>
<td>GACATACCACCCACCCATT</td>
<td></td>
<td></td>
<td></td>
<td>1071</td>
<td></td>
</tr>
<tr>
<td>30515-O1</td>
<td>GACAGCATGGGTAACCTA</td>
<td>Contig30515</td>
<td>1129</td>
<td>93</td>
<td>911</td>
<td>CE6</td>
</tr>
<tr>
<td>30515-O2c</td>
<td>CTGCTATCCACCCACTTG</td>
<td></td>
<td></td>
<td></td>
<td>1004</td>
<td></td>
</tr>
<tr>
<td>30515-I1</td>
<td>TTGGCCTCGTGCTAGTACCTTTCC</td>
<td></td>
<td></td>
<td>391</td>
<td>503</td>
<td></td>
</tr>
<tr>
<td>30515-12c</td>
<td>GCCGTTCCAGGTTGTAGATTCC</td>
<td></td>
<td></td>
<td></td>
<td>894</td>
<td></td>
</tr>
<tr>
<td>Node</td>
<td>Sequence</td>
<td>Contig</td>
<td>Start</td>
<td>End</td>
<td>Gene</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------------</td>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>29571-O1</td>
<td>CTGTCCCACCACCTTGCAATTG</td>
<td>Contig29571</td>
<td>1198</td>
<td>233</td>
<td>912</td>
<td>CE6</td>
</tr>
<tr>
<td>29571-O2c</td>
<td>TCGACATCCACCTTACCATCAG</td>
<td>Contig29571</td>
<td>1145</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29571-I1</td>
<td>TGTTGCTGTGCTGTTGTTGTG</td>
<td>Contig29571</td>
<td>339</td>
<td>745</td>
<td></td>
<td>GH45</td>
</tr>
<tr>
<td>29571-I2c</td>
<td>AACATGGGTAGCCAAGATCC</td>
<td>Contig29571</td>
<td>1084</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1733-O1</td>
<td>CCAAAACCACCCACCTGGAAT</td>
<td>Contig1733</td>
<td>1399</td>
<td>256</td>
<td>1018</td>
<td>GH45</td>
</tr>
<tr>
<td>1733-O2c</td>
<td>GGTGAATGGGGTGTTGAAAACG</td>
<td>Contig1733</td>
<td>1274</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1733-I1</td>
<td>AGCAGCAGCGTAAACCATAGGA</td>
<td>Contig1733</td>
<td>538</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1733-I2c</td>
<td>CGGATTGTGTGAGCTTGGGA</td>
<td>Contig1733</td>
<td>958</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32-O1</td>
<td>GGGCAACAAGCCAAGGTTAC</td>
<td>Contig32</td>
<td>998</td>
<td>998</td>
<td>998</td>
<td>GH45</td>
</tr>
<tr>
<td>32-O2c</td>
<td>GCTTTAAGGCAGCTGGAAGG</td>
<td>Contig32</td>
<td>1111</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32-I1</td>
<td>TGATGGTAACTGGGCCATTG</td>
<td>Contig32</td>
<td>315</td>
<td>635</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32-I2c</td>
<td>CCACCAGTGTGTTGTTGTTTG</td>
<td>Contig32</td>
<td>950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29098-O1</td>
<td>GGCTGGTAAAGGTCTGTAAGG</td>
<td>Contig29098</td>
<td>1484</td>
<td>71</td>
<td>1228</td>
<td>GH48</td>
</tr>
<tr>
<td>29098-O2c</td>
<td>CATGTACCCACCCCGAAAC</td>
<td>Contig29098</td>
<td>1299</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29098-I1</td>
<td>GGGACAGTGGAGCTGGAAGCACC</td>
<td>Contig29098</td>
<td>852</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29098-I2c</td>
<td>TGTTGGATGGAGCTTGGTG</td>
<td>Contig29098</td>
<td>852</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Node 3576-O1</td>
<td>TCCATCTTCACCAGCTACCTATGC</td>
<td>Contig_Node3576</td>
<td>2067</td>
<td>123</td>
<td>1534</td>
<td>GH48</td>
</tr>
<tr>
<td>NODE3576-O2c</td>
<td>CCCAGAATCTGTGTTGTAACGGAATC</td>
<td>Contig Node3576</td>
<td>1657</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE3576-I1</td>
<td>GCTCGTGTATTTGCTCTTAC</td>
<td>Contig Node3576</td>
<td>718</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE3576-I2c</td>
<td>CTACCGTCTCAAGGAAGTTTGTG</td>
<td>Contig Node3576</td>
<td>1226</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2118-I1</td>
<td>TGGGTCTTGGTTGCTGATA</td>
<td>Contig2118</td>
<td>2396</td>
<td>866</td>
<td>1026</td>
<td>GH74</td>
</tr>
<tr>
<td>2118-I2c</td>
<td>TGCTGAGGTGATGTTTAC</td>
<td>Contig2118</td>
<td>1892</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2118-O1</td>
<td>TTTCTGGGGCTCCTCATATC</td>
<td>Contig2118</td>
<td>295</td>
<td>1959</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2118-O2c</td>
<td>GTCTCGGTTATTGTTGTC</td>
<td>Contig2118</td>
<td>2254</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24305-O1</td>
<td>CAAATGGGCTACATGGACTGACC</td>
<td>Contig24305</td>
<td>1399</td>
<td>262</td>
<td>1012</td>
<td>GH74</td>
</tr>
<tr>
<td>24305-O2c</td>
<td>CACCTGTACCCACCTGCTCTCTC</td>
<td>Contig24305</td>
<td>1274</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24305-I1</td>
<td>TTGCTGAGATTGCGCTTGGAGAT</td>
<td>Contig24305</td>
<td>516</td>
<td>436</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24305-I2c</td>
<td>GGCGAAACACCACCATCTTCTTG</td>
<td>Contig24305</td>
<td>952</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23421-O1</td>
<td>CTCCAGCTATTGCCCCAATTTCG</td>
<td>Contig23421</td>
<td>1355</td>
<td>65</td>
<td>1127</td>
<td>CE3</td>
</tr>
<tr>
<td>23421-O2c</td>
<td>CAGCAGCTCCATTACCACAAC</td>
<td>Contig23421</td>
<td>1192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Sequence</td>
<td>Length</td>
<td>Start</td>
<td>End</td>
<td>Annotation</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------</td>
<td>--------</td>
<td>-------</td>
<td>-----</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>23421-I1</td>
<td>GGTATGTTCCGTTGGAGGTCAAAG</td>
<td>235</td>
<td>1</td>
<td>235</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23421-I2C</td>
<td>AGCAGTAGTTGGCTGGTGAGATTAG</td>
<td>1059</td>
<td>1</td>
<td>1059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>900-O1</td>
<td>TGGCCCAACAGGAAGAATCAAC</td>
<td>Contig900</td>
<td>1425</td>
<td>108</td>
<td>1218 CBM10/CE15</td>
<td></td>
</tr>
<tr>
<td>900-O2C</td>
<td>TTAGCCGATTGGGAATGCAGAC</td>
<td></td>
<td></td>
<td></td>
<td>1326</td>
<td></td>
</tr>
<tr>
<td>900-I1</td>
<td>ACCGAATCTGTCTTCCACTC</td>
<td></td>
<td></td>
<td></td>
<td>490 662</td>
<td></td>
</tr>
<tr>
<td>900-I1c</td>
<td>CCTAGTGGTAAATGGTCCATTCC</td>
<td></td>
<td></td>
<td></td>
<td>1152</td>
<td></td>
</tr>
<tr>
<td>28807-O1</td>
<td>GTGGAGCGACTAAAGACGTAAG</td>
<td>Contig28807</td>
<td>1617</td>
<td>327</td>
<td>1194 CBM10/CE15</td>
<td></td>
</tr>
<tr>
<td>28807-O2c</td>
<td>CTTGGCTACCCATGTTGTG</td>
<td></td>
<td></td>
<td></td>
<td>1521</td>
<td></td>
</tr>
<tr>
<td>28807-I1</td>
<td>CTAAACGGGAACAAACCAG</td>
<td></td>
<td></td>
<td></td>
<td>582 625</td>
<td></td>
</tr>
<tr>
<td>28807-I2c</td>
<td>TGCTCCAACTCCAGATAC</td>
<td></td>
<td></td>
<td></td>
<td>1207</td>
<td></td>
</tr>
<tr>
<td>node7061-O2</td>
<td>CTTGCCTTCCAGCTGTTAATGC</td>
<td>Contig Node7061</td>
<td>960</td>
<td>44</td>
<td>651 DUF297/CBM10</td>
<td></td>
</tr>
<tr>
<td>node7061-O2c</td>
<td>TCGGTTGGGTAACCCGTAAAGAG</td>
<td></td>
<td></td>
<td></td>
<td>695</td>
<td></td>
</tr>
<tr>
<td>node7061-I1</td>
<td>CTTATAACGGTGCTACCAAGG</td>
<td></td>
<td></td>
<td></td>
<td>168 373</td>
<td></td>
</tr>
<tr>
<td>Node7061-I2c</td>
<td>AACCAGCAGCCATGTAGTTAC</td>
<td></td>
<td></td>
<td></td>
<td>541</td>
<td></td>
</tr>
<tr>
<td>260-I1</td>
<td>GCCAGAGAAGAAGCTAAAGGG</td>
<td>Contig260</td>
<td>1229</td>
<td>301</td>
<td>581 DUF297</td>
<td></td>
</tr>
<tr>
<td>260-I2c</td>
<td>AGTAGAGGCCAGAACCAGAAG</td>
<td></td>
<td></td>
<td></td>
<td>882</td>
<td></td>
</tr>
<tr>
<td>260-O1</td>
<td>CTTGGGCTTGGTACTAAGGG</td>
<td></td>
<td></td>
<td></td>
<td>141 763</td>
<td></td>
</tr>
<tr>
<td>260-O2c</td>
<td>TAGCAGTGGTAGTTGGTTCTTC</td>
<td></td>
<td></td>
<td></td>
<td>904</td>
<td></td>
</tr>
<tr>
<td>7694-O1</td>
<td>GTGATGCTCGCAATCTCTAC</td>
<td>Contig7694</td>
<td>1435</td>
<td>88</td>
<td>587 PL09</td>
<td></td>
</tr>
<tr>
<td>7694-O2C</td>
<td>CCATCAGCATAGGCAACCATAAG</td>
<td></td>
<td></td>
<td></td>
<td>675</td>
<td></td>
</tr>
<tr>
<td>7694-I1</td>
<td>CCCCATTCCATGGCTGAAATGC</td>
<td></td>
<td></td>
<td></td>
<td>114 419</td>
<td></td>
</tr>
<tr>
<td>7694-I2C</td>
<td>TTGCTCCACTAACCCAGATACCC</td>
<td></td>
<td></td>
<td></td>
<td>533</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.3 Metabolic related KOG/COG groups represented by 5000 or more reads in the metatranscriptome from muskoxen rumen eukaryotes.

<table>
<thead>
<tr>
<th>KOG/COG Description</th>
<th>Reads Number</th>
<th>KOG/COG Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>KOG2670, Enolase</td>
<td>60668</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG0657, Glyceraldehyde 3-phosphate dehydrogenase</td>
<td>58954</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG0626, Beta-glucosidase, lactase phlorizinhydrolase, and related proteins</td>
<td>45908</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>COG0574, PpsA, Phosphoenolpyruvate synthase/pyruvate phosphate dikinase</td>
<td>41292</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG2099, Glycogen phosphorylase</td>
<td>28858</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG4153, Fructose 1,6-bisphosphate aldolase</td>
<td>24624</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG1367, 3-phosphoglycerate kinase</td>
<td>24163</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>COG2211, MelB, Na+/melibiose symporter and related transporters</td>
<td>17403</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG0625, Phosphoglucomutase</td>
<td>10338</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG1643, Triosephosphate isomerase</td>
<td>8573</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG1369, Hexokinase</td>
<td>7768</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>COG0057, GapA, Glyceraldehyde-3-phosphate dehydrogenase/erythrose-4-phosphate</td>
<td>6101</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>dehydrogenase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COG2115, XylA, Xylose isomerase</td>
<td>5791</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>KOG0372, Serine/threonine specific protein phosphatase involved in glycogen</td>
<td>5660</td>
<td>Carbohydrate transport and metabolism</td>
</tr>
<tr>
<td>accumulation, PP2A-related</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KOG3749, Phosphoenolpyruvate carboxykinase</td>
<td>66252</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG1494, NAD-dependent malate dehydrogenase</td>
<td>29529</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG1255, Succinyl-CoA synthetase, alpha subunit</td>
<td>22008</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>COG0281, SfcA, Malic enzyme</td>
<td>17260</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG2799, Succinyl-CoA synthetase, beta subunit</td>
<td>17208</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG0749, Mitochondrial ADP/ATP carrier proteins</td>
<td>14650</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG1257, NADP+-dependent malic enzyme</td>
<td>14255</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>COG1882, PfdD, Pyruvate-formate lyase</td>
<td>12898</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG0232, Vacuolar H+-ATPase V0 sector, subunits c/c'</td>
<td>12765</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG1352, Vacuolar H+-ATPase V1 sector, subunit A</td>
<td>10826</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG0453, Aconitase/homoaconitase (aconitase superfamily)</td>
<td>9878</td>
<td>Energy production and conversion</td>
</tr>
<tr>
<td>KOG</td>
<td>Description</td>
<td>COGID</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>KOG1351</td>
<td>Vacuolar H+-ATPase V1 sector, subunit B</td>
<td>8407</td>
</tr>
<tr>
<td>KOG1447</td>
<td>GTP-specific succinyl-CoA synthetase, beta subunit</td>
<td>6336</td>
</tr>
<tr>
<td>KOG2189</td>
<td>Vacuolar H+-ATPase V0 sector, subunit a</td>
<td>5029</td>
</tr>
<tr>
<td>KOG2250</td>
<td>Glutamate/leucine/phenylalanine/valine dehydrogenases</td>
<td>15125</td>
</tr>
<tr>
<td>KOG2467</td>
<td>Glycine/serine hydroxymethyltransferase</td>
<td>8025</td>
</tr>
<tr>
<td>KOG2448</td>
<td>Dihydroxy-acid dehydratase</td>
<td>7053</td>
</tr>
<tr>
<td>KOG2263</td>
<td>Methionine synthase II (cobalamin-independent)</td>
<td>6888</td>
</tr>
<tr>
<td>KOG0053</td>
<td>Cystathionine beta-lyases/cystathionine gamma-synthases</td>
<td>6162</td>
</tr>
<tr>
<td>KOG2790</td>
<td>Phosphoserine aminotransferase</td>
<td>6060</td>
</tr>
<tr>
<td>KOG1268</td>
<td>Glucosamine 6-phosphate synthetases, contain amidotransferase and phosphosugar isomerase domains</td>
<td>7607</td>
</tr>
<tr>
<td>KOG1506</td>
<td>S-adenosylmethionine synthetase</td>
<td>24216</td>
</tr>
<tr>
<td>KOG1370</td>
<td>S-adenosylhomocysteine hydrodase</td>
<td>8926</td>
</tr>
<tr>
<td>KOG0204</td>
<td>Calcium transporting ATPase</td>
<td>19494</td>
</tr>
<tr>
<td>KOG0693</td>
<td>Myo-inositol-1-phosphate synthase</td>
<td>25686</td>
</tr>
<tr>
<td>KOG059</td>
<td>Lipid exporter ABCA1 and related proteins, ABC superfamily</td>
<td>9000</td>
</tr>
<tr>
<td>KOG0888</td>
<td>Nucleoside diphosphate kinase</td>
<td>8411</td>
</tr>
<tr>
<td>KOG055</td>
<td>Multidrug/pheromone exporter, ABC superfamily</td>
<td>5171</td>
</tr>
</tbody>
</table>
Table 3.4 List of putative carbohydrate esterases, polysaccharide lyases and carbohydrate binding related modules in the muskoxen rumen eukaryotic metatranscriptome (Muskoxen MT), and comparison of their abundance of selected CAZy modules with those of three other foregut microbiomes: wallaby (Macropod); bovine rumen (Bovine 2009 and Bovine 2011) and the hindgut of termite (Termite).

Domains identified from muskoxen MT contigs that significantly differ from the rest data are indicated in bold.

<table>
<thead>
<tr>
<th></th>
<th>MuskOxen Contigs (≥500bp)</th>
<th>Macropod</th>
<th>Bovine 2009</th>
<th>Bovine 2011</th>
<th>Termite</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Carbohydrate Esterases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE1</td>
<td>13</td>
<td>0</td>
<td>11</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE2/CE3</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE4</td>
<td>50</td>
<td>24</td>
<td>4</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE7</td>
<td>19</td>
<td>3</td>
<td>2</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE8</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE9</td>
<td>3</td>
<td>14</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE10</td>
<td>3</td>
<td>18</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE11</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE12</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE13</td>
<td>5</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE14</td>
<td>0</td>
<td>1</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE15</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>CE16</td>
<td>8</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td><strong>Polysaccharide lyases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL1</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>PL2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>PL3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>PL5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>5</td>
</tr>
<tr>
<td>PL6</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>PL9</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>PL10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>PL11</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>5</td>
</tr>
<tr>
<td>PL12+15+17+21</td>
<td>0</td>
<td>0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td><strong>Carbohydrate Binding Modules</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBM1</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>CBM</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>CBM4/9/16/22</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>417</td>
<td>5</td>
</tr>
<tr>
<td>CBM6</td>
<td>27</td>
<td>3</td>
<td>0</td>
<td>932</td>
<td>13</td>
</tr>
<tr>
<td><strong>CBM10</strong></td>
<td><strong>403</strong></td>
<td>0</td>
<td>0</td>
<td><strong>1</strong></td>
<td>0</td>
</tr>
<tr>
<td>CBM11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>3</td>
</tr>
<tr>
<td>CBM13</td>
<td>31</td>
<td>2</td>
<td>1</td>
<td>118</td>
<td>0</td>
</tr>
<tr>
<td><strong>CBM18</strong></td>
<td><strong>108</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM20</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>CBM21</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CBM26</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>0</td>
</tr>
<tr>
<td><strong>CBM29</strong></td>
<td><strong>11</strong></td>
<td>0</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>0</td>
</tr>
<tr>
<td>CBM30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>1~8</td>
</tr>
<tr>
<td>CBM32/47</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>747</td>
<td>4</td>
</tr>
<tr>
<td>CBM34</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>72</td>
<td>0</td>
</tr>
<tr>
<td>CBM35</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>0~1</td>
</tr>
<tr>
<td>CBM36</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>2~13</td>
</tr>
<tr>
<td>CBM37</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>1</td>
</tr>
<tr>
<td>CBM48</td>
<td>3</td>
<td>22</td>
<td>0</td>
<td><strong>NR</strong></td>
<td>0~1</td>
</tr>
<tr>
<td>CBM50</td>
<td>11</td>
<td>47</td>
<td><strong>NR</strong></td>
<td>1887</td>
<td><strong>NR</strong></td>
</tr>
<tr>
<td>CBM51</td>
<td>0</td>
<td>1</td>
<td><strong>NR</strong></td>
<td>173</td>
<td><strong>NR</strong></td>
</tr>
</tbody>
</table>

^ NR: not reported.
Table 3.5 The abundance of contigs coding lignocellulolytic enzymes [glycoside hydrolases (GHs), carbohydrate esterases (CEs), polysaccharide lyases (PLs), carbohydrate-binding modules (CBMs), and other related modules] in the muskoxen eukaryotic metatranscriptome (Muskoxen MT) and a comparison of their abundance in the databases of rumen fungal (Ru. fungi) and rumen protozoal genes (Ru. prot.) as well as different anaerobic bacteria, including *Bacteroides fragilis* (Bfra), *Butyrivibrio proteoclasticus* (Bpro), *Clostridium thermocellum* (Cthe), *Fibrobacter succinogenes* (Fsuc), *Prevotella ruminicola* (Prum), *Ruminococcus flavocantiens* (Rfla), and the rumen fungus *Piromyces* sp. E2 ESTs (Pir. ESTs).

<table>
<thead>
<tr>
<th>CAZY family</th>
<th>Pfam Accession</th>
<th>Muskoxen MT Contigs all sizes</th>
<th>Muskoxen MT Contigs ≥500bp</th>
<th>Muskoxen MT Reads</th>
<th>Pir. ESTs</th>
<th>Ru. fungi</th>
<th>Ru. prot</th>
<th>Bfra</th>
<th>Bpro</th>
<th>Cthe</th>
<th>Fsuc</th>
<th>Prum</th>
<th>Rfla</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycoside hydrolase catalytic domains</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH1</td>
<td>PF00232.11</td>
<td>24</td>
<td>17</td>
<td>32628</td>
<td>32</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH2</td>
<td>PF02836.10</td>
<td>12</td>
<td>10</td>
<td>186</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>8</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>GH3</td>
<td>PF00933.14</td>
<td>50</td>
<td>22</td>
<td>4967</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>2</td>
<td>3</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>GH4</td>
<td>PF02056.9</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH5</td>
<td>PF00150.11</td>
<td>75</td>
<td>46</td>
<td>2423</td>
<td>9</td>
<td>13</td>
<td>13</td>
<td>0</td>
<td>5</td>
<td>11</td>
<td>12</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>GH6</td>
<td>PF01341.10</td>
<td>55</td>
<td>31</td>
<td>26950</td>
<td>20</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH7</td>
<td>PF00840.13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH8</td>
<td>PF02770.10</td>
<td>6</td>
<td>6</td>
<td>496</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH9</td>
<td>PF00759.12</td>
<td>52</td>
<td>42</td>
<td>10894</td>
<td>22</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>16</td>
<td>9</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>GH10</td>
<td>PF00331.13</td>
<td>71</td>
<td>29</td>
<td>5658</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>GH11</td>
<td>PF00457.10</td>
<td>45</td>
<td>33</td>
<td>22294</td>
<td>13</td>
<td>30</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>GH12</td>
<td>PF01670.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH13</td>
<td>PF00128.17</td>
<td>77</td>
<td>47</td>
<td>4564</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>6</td>
<td>14</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>GH14</td>
<td>PF01373.10</td>
<td>5</td>
<td>0</td>
<td>144</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CAZY family</td>
<td>Pfam Accession</td>
<td>Muskoxen MT Contigs all sizes</td>
<td>Muskoxen MT Contigs ≥500bp</td>
<td>Muskoxen MT Reads</td>
<td>Pir. ESTs</td>
<td>Ru. fungi</td>
<td>Ru. prot</td>
<td>Bfra</td>
<td>Bpro</td>
<td>Cthe</td>
<td>Fsuc</td>
<td>Prum</td>
<td>Rfla</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>GH15</td>
<td>PF00723.14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH16</td>
<td>PF00722.14</td>
<td>17</td>
<td>10</td>
<td>1797</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>GH17</td>
<td>PF00322.10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH18</td>
<td>PF00704.21</td>
<td>30</td>
<td>20</td>
<td>1905</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GH19</td>
<td>PF00182.12</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH20</td>
<td>PF00728.15</td>
<td>1</td>
<td>0</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>GH21</td>
<td>deleted family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH22</td>
<td>PF00062.13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH23</td>
<td>PF01464.13</td>
<td>4</td>
<td>1</td>
<td>105</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>GH24</td>
<td>PF00959.12</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GH25</td>
<td>PF01183.13</td>
<td>31</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>GH26</td>
<td>PF02156.8</td>
<td>13</td>
<td>8</td>
<td>1737</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>GH27</td>
<td>PF02065.11</td>
<td>9</td>
<td>6</td>
<td>85</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GH28</td>
<td>PF00295.10</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>GH29</td>
<td>PF01120.10</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>GH30</td>
<td>PF02055.9</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>GH31</td>
<td>PF01055.19</td>
<td>35</td>
<td>4</td>
<td>1411</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>GH32</td>
<td>PF00251.13</td>
<td>9</td>
<td>3</td>
<td>221</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>GH33</td>
<td>PF02012.13</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GH34</td>
<td>PF00064.11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH35</td>
<td>PF01301.12</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>GH36</td>
<td>PF02065</td>
<td>9</td>
<td>6</td>
<td>85</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GH37</td>
<td>PF01204.11</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH38</td>
<td>PF01074.15</td>
<td>1</td>
<td>0</td>
<td>80</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH39</td>
<td>PF01229.10</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CAZY family</td>
<td>Pfam Accession</td>
<td>Muskoxen MT Contigs all sizes</td>
<td>Muskoxen MT Contigs ≥500bp</td>
<td>Muskoxen MT Reads</td>
<td>Pir. ESTs</td>
<td>Ru. fungi</td>
<td>Ru. prot</td>
<td>Bfra</td>
<td>Bpro</td>
<td>Cthe</td>
<td>Fsuc</td>
<td>Prum</td>
<td>Rfla</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>------------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>GH40</td>
<td>deleted family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH41</td>
<td>deleted family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH42</td>
<td>PF02449.8</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GH43</td>
<td>PF04616.7</td>
<td>81</td>
<td>54</td>
<td>139</td>
<td>75</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>13</td>
<td>5</td>
<td>13</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>GH44</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GH45</td>
<td>PF02015.9</td>
<td>58</td>
<td>51</td>
<td>14034</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH46</td>
<td>PF01374.11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH47</td>
<td>PF01532.13</td>
<td>0</td>
<td>0</td>
<td>51</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH48</td>
<td>PF02011.8</td>
<td>54</td>
<td>33</td>
<td>106034</td>
<td>37</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GH49</td>
<td>PF03718.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH50</td>
<td>PB001653</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH51</td>
<td>PF06964</td>
<td>1</td>
<td>1</td>
<td>112</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>GH52</td>
<td>PF03512.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH53</td>
<td>PF07745.6</td>
<td>4</td>
<td>0</td>
<td>83</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>GH54</td>
<td>PF09206.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH55</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH56</td>
<td>PF01630.11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH57</td>
<td>PF03065.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH58</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH59</td>
<td>PF02057.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH60</td>
<td>Deleted family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH61</td>
<td>PF03443.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH62</td>
<td>PF03664.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH63</td>
<td>PF03200.9</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CAZY family</td>
<td>Pfam Accession</td>
<td>Muskoxen MT Contigs all sizes</td>
<td>Muskoxen MT Contigs ≥500bp</td>
<td>Muskoxen MT Reads</td>
<td>Pir. ESTs</td>
<td>Ru. fungi</td>
<td>Ru. prot</td>
<td>Bfra</td>
<td>Bpro</td>
<td>Cthe</td>
<td>Fsuc</td>
<td>Prum</td>
<td>Rfla</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>GH64</td>
<td>PB001434</td>
<td>4</td>
<td>3</td>
<td>593</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH65</td>
<td>PF03632.8</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH66</td>
<td>PB003959</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH67</td>
<td>PF07488.5</td>
<td>1</td>
<td>1</td>
<td>639</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH68</td>
<td>PF02435.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH69</td>
<td>Deleted: now PL16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GH70</td>
<td>PF02324.9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH71</td>
<td>PF03659.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH72</td>
<td>PF03198.7</td>
<td>5</td>
<td>0</td>
<td>129</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH73</td>
<td>PF01832.13</td>
<td>2</td>
<td>2</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH74</td>
<td>BLAST</td>
<td>22</td>
<td>6</td>
<td>171</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH75</td>
<td>PF07335.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH76</td>
<td>PF03663.7</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH77</td>
<td>PF02446.10</td>
<td>22</td>
<td>8</td>
<td>2643</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GH78</td>
<td>PF05592.4</td>
<td>5</td>
<td>0</td>
<td>181</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH79</td>
<td>PF03662.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH80</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH81</td>
<td>PF03639.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH82</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH83</td>
<td>PF00423.12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH84</td>
<td>PF07555.6</td>
<td>3</td>
<td>1</td>
<td>182</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH85</td>
<td>PF03644.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH86</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH87</td>
<td>BLAST</td>
<td>7</td>
<td>3</td>
<td>433</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH88</td>
<td>PF07470.6</td>
<td>9</td>
<td>1</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CAZY family</td>
<td>Pfam Accession</td>
<td>Muskoxen MT Contigs all sizes</td>
<td>Muskoxen MT Contigs ≥500bp</td>
<td>Muskoxen MT Reads</td>
<td>Pir. ESTs</td>
<td>Ru. fungi</td>
<td>Ru. prot</td>
<td>Bfra</td>
<td>Bpro</td>
<td>Cthe</td>
<td>Fsuc</td>
<td>Prum</td>
<td>Rfla</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>----------------------------</td>
<td>-------------------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>GH89</td>
<td>PF05089.5</td>
<td>15</td>
<td>6</td>
<td>1191</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>GH90</td>
<td>PF09251</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH91</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH92</td>
<td>PF07971.5</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>GH93</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH94</td>
<td>BLAST</td>
<td>2</td>
<td>1</td>
<td>4308</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>GH95</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH96</td>
<td>PF10566.2</td>
<td>2</td>
<td>2</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>GH97</td>
<td>PF08306.4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH98</td>
<td>PB000173</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH99</td>
<td>PF04853.5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH100</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH101</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH102</td>
<td>PF03562.10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH103</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH104</td>
<td>PF00959.12</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH105</td>
<td>PF07470.6</td>
<td>9</td>
<td>1</td>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GH106</td>
<td>PB000239</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH107</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH108</td>
<td>PF05838.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH109</td>
<td>PF01408</td>
<td>2</td>
<td>0</td>
<td>51</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>GH110</td>
<td>PB004422</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH111</td>
<td>PWGH111</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH112</td>
<td>PF09508</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH113</td>
<td>PB000085</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH114</td>
<td>PF03537</td>
<td>23</td>
<td>11</td>
<td>343</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CAZY family</td>
<td>Pfam Accession</td>
<td>Muskoxen MT Contigs all sizes</td>
<td>Muskoxen MT Contigs ≥500bp</td>
<td>Muskoxen MT Reads</td>
<td>Pir. ESTs</td>
<td>Ru. fungi</td>
<td>Ru. prot</td>
<td>Bfra</td>
<td>Bpro</td>
<td>Cthe</td>
<td>Fsuc</td>
<td>Prum</td>
<td>Rfla</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>GH115</td>
<td>PB000184</td>
<td>13</td>
<td>10</td>
<td>1541</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH116</td>
<td>PF12215</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GH117</td>
<td>PWGH117</td>
<td>16</td>
<td>14</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>GH118</td>
<td>PWGH118</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**CBMs**

<p>| CBM1        | PF00734.11      | 85                            | 33                          | 3697              | 20        | 11        | 0        | 0    | 0    | 7    | 0    | 0    | 0    |
| CBM2        | PF00553.12      | 0                             | 0                           | 2                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM3        | PF00942.11      | 3                             | 3                           | 1                 | 0         | 0         | 2        | 0    | 1    | 20   | 0    | 0    | 3    |
| CBM4        | PF02018.10      | 5                             | 3                           | 0                 | 0         | 0         | 3        | 2    | 0    | 8    | 5    | 4    | 19   |
| CBM5        | PF02839.7       | 0                             | 0                           | 1                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM6        | PF03422.8       | 40                            | 27                          | 84                | 0         | 1         | 0        | 3    | 3    | 15   | 30   | 2    | 6    |
| CBM7        | deleted entry   |                               |                             |                   |           |           |          |      |      |      |      |      |      |
| CBM8        | BLAST           | 0                             | 0                           | 0                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM9        | PF02018.10      | 5                             | 3                           | 0                 | 0         | 0         | 3        | 2    | 0    | 8    | 5    | 4    | 19   |
| CBM10       | PF02013.9       | 908                           | 403                         | 100192            | 183       | 84        | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM11       | PF03425.6       | 0                             | 0                           | 0                 | 0         | 1         | 0        | 0    | 0    | 1    | 6    | 0    | 0    |
| CBM12       | PF02839.7       | 0                             | 0                           | 1                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM13       | PF00652.15      | 40                            | 31                          | 5                 | 2         | 0         | 1        | 0    | 3    | 1    | 0    | 0    | 8    |
| CBM14       | PF01607.17      | 0                             | 0                           | 0                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM15       | PF03426.7       | 0                             | 0                           | 0                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM16       | PF02018.10      | 5                             | 3                           | 0                 | 0         | 0         | 3        | 2    | 0    | 8    | 5    | 4    | 19   |
| CBM17       | PF03424.7       | 0                             | 0                           | 0                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM18       | PF00187.12      | 370                           | 108                         | 22914             | 88        | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM19       | PF03427.6       | 0                             | 0                           | 0                 | 0         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |
| CBM20       | PF00686.12      | 6                             | 5                           | 84                | 0         | 0         | 1        | 0    | 0    | 0    | 0    | 0    | 2    |
| CBM21       | PF03370.6       | 1                             | 0                           | 52                | 3         | 0         | 0        | 0    | 0    | 0    | 0    | 0    | 0    |</p>
<table>
<thead>
<tr>
<th>CAZY family</th>
<th>Pfam Accession</th>
<th>Muskoxen MT Contigs all sizes</th>
<th>Muskoxen MT Contigs ≥500bp</th>
<th>Muskoxen MT Reads</th>
<th>Pir. ESTs</th>
<th>Ru. fungi</th>
<th>Ru. prot</th>
<th>Bfra</th>
<th>Bpro</th>
<th>Cthe</th>
<th>Fsuc</th>
<th>Prum</th>
<th>Rfla</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBM22</td>
<td>PF02018.10</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>CBM23</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM24</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM25</td>
<td>PF03423.6</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM26</td>
<td>PB013554</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM27</td>
<td>PF09212.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM28</td>
<td>PF03424.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM29</td>
<td>BLAST</td>
<td>13</td>
<td>11</td>
<td>2318</td>
<td>9</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM30</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM31</td>
<td>PF11606</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM32</td>
<td>PF00754.18</td>
<td>6</td>
<td>2</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>CBM33</td>
<td>PF03067.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM34</td>
<td>PF02903.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM35</td>
<td>BLAST</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>CBM36</td>
<td>BLAST</td>
<td>21</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CBM37</td>
<td>BLAST</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>CBM38</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM39</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM40</td>
<td>PF02973.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM41</td>
<td>PF03714.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM42</td>
<td>PF05270.6</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM43</td>
<td>PF07983.6</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM44</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM45</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM46</td>
<td>PF03442</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM47</td>
<td>PF00754.18</td>
<td>6</td>
<td>2</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CAZY family</td>
<td>Pfam Accession</td>
<td>Muskoxen MT Contigs all sizes</td>
<td>Muskoxen MT Contigs ≥500bp</td>
<td>Muskoxen MT Reads</td>
<td>Pir. ESTs</td>
<td>Ru. fungi</td>
<td>Ru. prot</td>
<td>Bfra</td>
<td>Bpro</td>
<td>Cthe</td>
<td>Fsuc</td>
<td>Prum</td>
<td>Rfla</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>------------------------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>CBM48</td>
<td>PF02922.11</td>
<td>6</td>
<td>3</td>
<td>98</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>CBM49</td>
<td>PF09478.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM50</td>
<td>PF01476.13</td>
<td>18</td>
<td>11</td>
<td>1520</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>10</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM51</td>
<td>PF08305.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CBM52</td>
<td>PF10645</td>
<td>4</td>
<td>1</td>
<td>101</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM53</td>
<td>BLAST</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM54</td>
<td>PWCBM54</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM55</td>
<td>PWCBM55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM56</td>
<td>PWCBM56</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM57</td>
<td>PF11721</td>
<td>2</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM58</td>
<td>PWCBM58</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CBM59</td>
<td>PWCBM59</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Carbohydrate esterases**

<p>| CE1         | PF00756        | 17                           | 13                       | 86               | 0         | 2       | 0       | 4   | 6   | 3   | 1    | 7    | 9   |
| CE2+CE3     | PB002673       | 18                           | 6                        | 85               | 0         | 3       | 0       | 3   | 2   | 1   | 1    | 2    | 3   |
| CE4         | PF01522.14     | 78                           | 50                       | 983              | 4         | 0       | 2       | 5   | 5   | 3   | 2    | 6    |     |
| CE5         | PF01083.15     | 0                            | 0                        | 0                | 0         | 0       | 0       | 0   | 0   | 0   | 0    | 0    |     |
| CE6         | PF03629        | 19                           | 19                       | 17               | 0         | 2       | 0       | 2   | 1   | 0   | 5    | 3    | 0   |
| CE7         | PF05448        | 1                            | 0                        | 6                | 0         | 0       | 2       | 1   | 1   | 1   | 3    | 2    |     |
| CE8         | PF01095.12     | 10                           | 3                        | 96               | 0         | 0       | 0       | 0   | 2   | 1   | 1    | 2    | 1   |
| CE9         | PF01979.13     | 7                            | 3                        | 31               | 0         | 0       | 0       | 5   | 3   | 6   | 2    | 1    | 3   |
| CE10        | PF00135        | 15                           | 4                        | 745              | 10        | 0       | 0       | 1   | 5   | 1   | 1    | 3    | 0   |
| CE11        | PF03331.6      | 0                            | 0                        | 4                | 0         | 0       | 1       | 0   | 0   | 1   | 1    | 1    | 0   |
| CE12        | PB008046       | 10                           | 5                        | 553              | 0         | 0       | 1       | 3   | 1   | 5   | 4    | 3    |     |
| CE13        | PF03283        | 0                            | 0                        | 157              | 0         | 0       | 0       | 1   | 0   | 0   | 0    | 0    | 1   |
| CE14        | PF02585.10     | 0                            | 0                        | 5                | 0         | 0       | 1       | 1   | 1   | 0   | 1    | 0    |     |</p>
<table>
<thead>
<tr>
<th>CAZY family</th>
<th>Pfam Accession</th>
<th>Muskoxen MT Contigs all sizes</th>
<th>Muskoxen MT Contigs ≥500bp</th>
<th>Muskoxen MT Reads</th>
<th>Pir. ESTs</th>
<th>Ru. fungi</th>
<th>Ru. prot</th>
<th>Bfra</th>
<th>Bpro</th>
<th>Cthe</th>
<th>Fsuc</th>
<th>Prum</th>
<th>Rfla</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE15</td>
<td>PWCE015</td>
<td>9</td>
<td>8</td>
<td>4231</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>CE16</td>
<td>PWCE016</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**Polysaccharide lyases**

| PL1         | PF00544.12     | 10                          | 6                         | 244              | 0        | 0        | 0      | 1   | 2   | 6   | 1    | 4    |       |
| PL2         | PF06917        | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL3         | PF03211.6      | 10                          | 4                         | 3660             | 2        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL4         | PF06045        | 1                           | 1                         | 0                | 0        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL5         | PF09284        | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL6         | PWPL006        | 8                           | 7                         | 0                | 0        | 0        | 0      | 0   | 1   | 0   | 0    | 0    |       |
| PL7+PL18    | PF08787        | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL8         | PF02278.11     | 0                           | 0                         | 8                | 0        | 0        | 0      | 1   | 0   | 0   | 0    | 0    |       |
| PL9         | QMPL09         | 15                          | 8                         | 325              | 0        | 0        | 0      | 2   | 4   | 2   | 2    | 0    | 2    |
| PL10        | PF09492        | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 1   | 1   | 0    | 1    |       |
| PL11        | PWPL011        | 1                           | 1                         | 113              | 0        | 0        | 0      | 1   | 1   | 2   | 1    | 7    |       |
| PL13        | PWPL013        | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL14        | PB002765       | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 1   | 0   | 0    | 0    |       |
| PL16        | PF07212        | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL19        | Deleted:No wGH91 |                   |                           |                  |          |          |        |     |     |     |      |      |       |
| PL20        | PWPL020        | 0                           | 0                         | 0                | 0        | 0        | 0      | 0   | 0   | 0   | 0    | 0    |       |
| PL22        | PB009195       | 0                           | 0                         | 0                | 0        | 0        | 0      | 5   | 0   | 0   | 2    | 3    | 0    |

**Other domains associated with GH catalytic or carbohydrate binding domains**

<p>| AXE1        | PF05448.5      | 1                           | 0                         | 6                | 0        | 0        | 0      | 2   | 1   | 1   | 1    | 3    | 2    |</p>
<table>
<thead>
<tr>
<th>CAZY family</th>
<th>Pfam Accession</th>
<th>Musken MT Contigs all sizes</th>
<th>Muskeen MT Contigs ≥500bp</th>
<th>Musken MT Reads</th>
<th>Pir. ESTs</th>
<th>Ru. fungi</th>
<th>Ru. prot</th>
<th>Bfra</th>
<th>Bpro</th>
<th>Cthe</th>
<th>Fsuc</th>
<th>Prum</th>
<th>Rfla</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha-amylase_C</td>
<td>PF02806.11</td>
<td>2</td>
<td>2</td>
<td>143</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Alpha-L-AE_C</td>
<td>PF06964.5</td>
<td>1</td>
<td>1</td>
<td>112</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Alpha-mann mid</td>
<td>PF09261.4</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Bac_rhamnoid_n</td>
<td>PF08531.3</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Big_1</td>
<td>PF02369.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Big_2</td>
<td>PF02368.11</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Big_3</td>
<td>PF07523.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Big_4</td>
<td>PF07532.4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Bgal_small_N</td>
<td>PF02929.10</td>
<td>3</td>
<td>3</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>CBMX</td>
<td>PF06204.4</td>
<td>1</td>
<td>1</td>
<td>793</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>CelD_N</td>
<td>PF02927.7</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>CHB_Hex</td>
<td>PF03173.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CHB_Hex_C</td>
<td>PF03174.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ChiC</td>
<td>PF06483.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ChitinaseA_N</td>
<td>PF08329.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cohesin</td>
<td>PF00963.11</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Dockerin_1</td>
<td>PF00404.11</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>72</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>fn3</td>
<td>PF00041.14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>GDE_C</td>
<td>PF06202.7</td>
<td>3</td>
<td>3</td>
<td>357</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Glucodextran_B</td>
<td>PF09136.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Glucodextran_N</td>
<td>PF09137.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CAZY family</td>
<td>Pfam Accession</td>
<td>Muskoxen MT Contigs all sizes</td>
<td>Muskoxen MT Contigs ≥500bp</td>
<td>Muskoxen MT Reads</td>
<td>Pir. ESTs</td>
<td>Ru. fungi</td>
<td>Ru. prot</td>
<td>Bfra</td>
<td>Bpro</td>
<td>Cthe</td>
<td>Fsuc</td>
<td>Prum</td>
<td>Rfla</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>---------</td>
<td>--------</td>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>GH2N</td>
<td>PF02837.11</td>
<td>9</td>
<td>7</td>
<td>463</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>2</td>
<td>2</td>
<td>13</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>GH2</td>
<td>PF00703.14</td>
<td>6</td>
<td>4</td>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>GH3C</td>
<td>PF01915.15</td>
<td>37</td>
<td>23</td>
<td>2035</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>GH20b</td>
<td>PF02838.8</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH32C</td>
<td>PF08244.5</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GH38C</td>
<td>PF07748.6</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GH42M</td>
<td>PF08532.3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GH42C</td>
<td>PF08533.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH65N</td>
<td>PF03636.8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH65C</td>
<td>PF03633.8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH67N</td>
<td>PF03648.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>GH67C</td>
<td>PF07477.5</td>
<td>1</td>
<td>1</td>
<td>406</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>GH98C</td>
<td>PF08307.4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Glyco_transf_36</td>
<td>PF06165.4</td>
<td>1</td>
<td>1</td>
<td>333</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>GT36_AF</td>
<td>PF06205.4</td>
<td>2</td>
<td>1</td>
<td>588</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>He_PIG</td>
<td>PF05345.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>PKD</td>
<td>PF00801.13</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TIG</td>
<td>PF01833.17</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SusD</td>
<td>PF07980.4</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>59</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>TonB_dep_r</td>
<td>PF00593.17</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>82</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>41</td>
<td>0</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLH</td>
<td>PF00395</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Swollenin</td>
<td>SWOLLEN</td>
<td>22</td>
<td>16</td>
<td>7065</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.6 CAZyme contigs (≥500 bp) identified with 10,000 or more reads in the metatranscriptome from the muskoxen rumen

<table>
<thead>
<tr>
<th>Contig</th>
<th>Domains</th>
<th>Length</th>
<th>No. of Reads</th>
<th>E-value</th>
<th>Id%</th>
<th>BLASTX Hit Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig21589</td>
<td>GH48; CBM10</td>
<td>2292</td>
<td>52778</td>
<td>0</td>
<td>78.2</td>
<td>cellulase Cel48A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig28863</td>
<td>GH1</td>
<td>1148</td>
<td>44380</td>
<td>0</td>
<td>86.8</td>
<td>xylanase [Neocallimastix frontalis]</td>
</tr>
<tr>
<td>Contig22047</td>
<td>GH6; CBM10</td>
<td>1921</td>
<td>44342</td>
<td>0</td>
<td>83.1</td>
<td>cellobiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig28684</td>
<td>GH48</td>
<td>898</td>
<td>33408</td>
<td>1.00E-154</td>
<td>81.1</td>
<td>cellulase Cel48A precursor [Piromyces equi]</td>
</tr>
<tr>
<td>Contig627</td>
<td>CBM20; PL6</td>
<td>1777</td>
<td>20162</td>
<td>7.00E-14</td>
<td>24.5</td>
<td>Parallel beta-helix repeat protein [Planctomyces maris DSM 8797]</td>
</tr>
<tr>
<td>Contig27493</td>
<td>GH16</td>
<td>878</td>
<td>19963</td>
<td>9.00E-74</td>
<td>57.1</td>
<td>Licheninase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig28658</td>
<td>GH48; CBM10</td>
<td>1093</td>
<td>16001</td>
<td>1.00E-145</td>
<td>74.3</td>
<td>cellulase Cel48A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig21206</td>
<td>GH45</td>
<td>925</td>
<td>15508</td>
<td>1.00E-68</td>
<td>52.4</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig21311</td>
<td>GH94; GT36_AF</td>
<td>2521</td>
<td>15361</td>
<td>0</td>
<td>72.5</td>
<td>celllobiose phosphorylase [Prevotella ruminicola 23]</td>
</tr>
<tr>
<td>Contig21506</td>
<td>GH1</td>
<td>821</td>
<td>14352</td>
<td>1.00E-110</td>
<td>88.6</td>
<td>beta-glucosidase Cel1C [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig29533</td>
<td>GH1</td>
<td>819</td>
<td>14351</td>
<td>1.00E-138</td>
<td>88.2</td>
<td>beta-glucosidase [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig3078</td>
<td>Glucosaminidase</td>
<td>601</td>
<td>13803</td>
<td>1.00E-26</td>
<td>72.5</td>
<td>Muramidase (flagellum-specific) [Eubacterium rectale A1-86 (DSM 17629)]</td>
</tr>
<tr>
<td>Contig29741</td>
<td>GH1</td>
<td>586</td>
<td>13272</td>
<td>7.00E-91</td>
<td>84.7</td>
<td>beta-glucosidase Cel1C [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig29986</td>
<td>CBM1</td>
<td>854</td>
<td>13133</td>
<td>6.00E-33</td>
<td>37.8</td>
<td>Xylanase B: XYL B [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig29325</td>
<td>GH43</td>
<td>1409</td>
<td>12451</td>
<td>1.00E-113</td>
<td>59.0</td>
<td>beta-xylanosidase/alpha-L-arabinofuranosidase, putative, gly43F [Cellvibrio japonicus Ueda107]</td>
</tr>
<tr>
<td>Contig26982</td>
<td>GH1</td>
<td>1879</td>
<td>11902</td>
<td>0</td>
<td>82.3</td>
<td>beta-glucosidase [Piromyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig12311</td>
<td>GH1</td>
<td>1369</td>
<td>11613</td>
<td>0</td>
<td>83.7</td>
<td>beta-glucosidase [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig30005</td>
<td>GH48</td>
<td>1203</td>
<td>11562</td>
<td>0</td>
<td>83.1</td>
<td>cellulase Cel48A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>NODE_3576</td>
<td>GH48; CBM10</td>
<td>2067</td>
<td>11488</td>
<td>0</td>
<td>79.8</td>
<td>cellulase Cel48A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig29850</td>
<td>GH48; CBM10</td>
<td>1092</td>
<td>11311</td>
<td>1.00E-152</td>
<td>73.8</td>
<td>cellulase Cel48A precursor [Piromyces equi]</td>
</tr>
<tr>
<td>Contig30163</td>
<td>GH1</td>
<td>954</td>
<td>11088</td>
<td>1.00E-146</td>
<td>74.1</td>
<td>beta-glucosidase [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig30798</td>
<td>GH45</td>
<td>523</td>
<td>10896</td>
<td>3.00E-48</td>
<td>63.6</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig29098</td>
<td>GH48; CBM10</td>
<td>1484</td>
<td>10522</td>
<td>1.00E-145</td>
<td>71.3</td>
<td>cellulase Cel48A precursor [Piromyces equi]</td>
</tr>
<tr>
<td>Contig1597</td>
<td>GH6; CBM10</td>
<td>1020</td>
<td>10347</td>
<td>1.00E-117</td>
<td>65.4</td>
<td>exoglucanase Cel6A [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig27001</td>
<td>GH43</td>
<td>1127</td>
<td>10074</td>
<td>1.00E-115</td>
<td>60.1</td>
<td>beta-xylanosidase/alpha-L-arabinofuranosidase, putative, gly43F [Cellvibrio japonicus Ueda107]</td>
</tr>
</tbody>
</table>
Figure 3.1 Summary statistics for the muskoxen rumen eukaryotic metatranscriptome.
**Figure 3.2** Length distribution of muskoxen rumen metatranscriptome contigs.

The number of contigs is indicated on the right side of the bar.
Figure 3.3 GC content analysis of the muskox rumen microbial community metatranscriptome.

The GC molar % of each contig was calculated. Number shown on the column indicates number of contigs with a certain GC range. The data of ocean microbiome metatranscriptome (Poretsky et al., 2009), bovine rumen metagenome (Brulc et al., 2009) and termite gut metagenome (Warnecke et al., 2007) are also shown.
Figure 3.4 Phylogenetic distribution of muskoxen rumen metatranscriptome putative protein encoding reads (a total of 21.1 million) based on MEGAN analysis of top BLASTX hits against the NRMO database. The percentages of the major phylogenetic groups are indicated.
Figure 3.5 Top 30 phylogenetic bins of the muskoxen rumen metatranscriptome as determined by comparison against NCBI's non-redundant amino acid (nr) database. Ranks are determined by the highest total reads number at the genus level.
Figure 3.6 Collector’s curve of richness as a function of reads analyzed.
Figure 3.7 Category distribution of the muskoxen rumen metatranscriptome as annotated using Eukaryotic Orthologous Groups (KOGs, for reads showing top BLASTX match to eukaryotic genes; solid bar) and clusters of orthologous groups (COGs, reads showing top BLASTX match to bacterial genes; dotted bar).

The assigned letters are based on KOG/COG classifications (Tatusov et al., 2003).

A total of 5.7 million out of 21.1 million putative protein encoding sequences in the muskoxen rumen eukaryotic metatranscriptome were annotated to a KOG category or COG category. The percentage of annotated ORFs for each KOG/COG category is shown.
Figure 3.8 Top 30 KOG bins of the muskoxen rumen metatranscriptome as determined by comparison against KOG database.

Ranks are determined by the highest number of total reads for each KOG category.
Figure 3.9 Schematic representation of plant cell wall polysaccharide and energy metabolism by the muskoxen rumen eukaryotic population.

The inner box represents the hydrogenosome present in anaerobic fungi and possibly the rumen protozoa. The number after each enzyme represents the read number identified by KOG/COG searches. Abbreviations: ASCT, Acetate: Succinate CoA-transferase; CAZY, carbohydrate active enzymes; GAPDH, glyceraldehydes-3-phosphate dehydrogenase; PEP, phosphoenolpyruvate; PFL, Pyruvate: Formate lyase; PFO, Pyruvate: ferredoxin oxidoreductase; PGK, Phosphoglycerate kinase; PGM, Phosphoglycerate mutase.
**Figure 3.10** Comparison of the carbohydrate active enzymes identified from muskoxen rumen metatranscriptome (using all assembled contigs) with those of three other foregut metagenomes, the termite hindgut and five rumen/anaerobic microorganisms.

The percentages of each enzyme family were shown in the cells. Refer to Table 3.5 for a complete comparison. Dendrogram on the top indicates the relationship of the GHs identified based on similar percentage distribution. Muskox MT: Muskoxen rumen metatranscriptome; Fibrobacter: Genome of *Fibrobacter succinogenes* S85; Ruminococcus: Genome of *Ruminococcus flavefaciens*; Clostridium: Genome of
*Clostridium thermocellum*; Piromyces: EST sequence of *Piromyces* sp. E2; Macropod MG: Macropod foregut microbiome (Pope et al., 2010); Termite MG: Termite hindgut microbiome (Warnecke et al., 2007); Bovine MG-Hess: Bovine Rumen microbiome by Hess et al. (2011); Bovine MG-Brulc: Bovine Rumen microbiome by Brulc et al. (2009); and Prevotella: Genome of *Prevotella ruminicola*. 

118
Figure 3.11 Amino acid sequence similarities of carbohydrate active enzymes identified from the muskoxen rumen metatranscriptome (using all assembled contigs) versus top BLASTX match to the Genbank nr database.
Figure 3.12 Top 30 phylogenetic bins of the bacterial reads of muskoxen rumen metatranscriptome as determined by comparison against NCBI's non-redundant amino acid (nr) database.

Ranks are determined by the highest number of total reads at the family level.
Figure 3.13 See page 122 for caption.
Figure 3.13 Evidence of eukaryotic origin of the metatranscriptome sequences based on BLASTX searches.

A): reads that were assigned to actin (KOG0676).

B): reads that were assigned to translation elongation factor EF1 (KOG0052).

C): MLTreeMap analysis of all the contigs. Eukaryotes are colored pink, archaea green and bacteria blue.

Number of reads that matched to each node are indicated in A and B.
Figure 3.14 Phylogenetic distribution of muskoxen rumen metatranscriptome putative carbohydrate active enzymes based on MEGAN analysis of top BLASTX hits of the contigs against the Genbank non-redundant amino acid (nr) database.

The number of contigs (≥500 bp) that matched to each node is indicated.
Chapter 4 Comparative transcriptomic analysis of

*Anaeromyces mucronatus YE505*

4.1 Introduction

Ruminants rely primarily on the microbial community in the rumen to break down plant polysaccharides and ferment released soluble sugars. Bacteria, protozoa and fungi are the key participants involved in cellulose degradation in the rumen. Within the past ten years, genomes of a number of rumen bacteria have been sequenced, and analyses of these genomes have indicated that rumen lignocellulolytic bacteria have many unique features in their cellulase system as compared to their aerobic counterparts (Berg Miller et al., 2009; Purushe et al., 2010; Suen et al., 2011a; Suen et al., 2011b). Many more previously unknown cellulolytic enzymes have been identified via bioinformatic analysis, highlighting the potential of this ecosystem to provide unique carbohydrate active enzymes for agricultural and commercial usage.

In contrast, although some enzymes have been detected or isolated from rumen anaerobic fungi, genomic information on these poorly researched microorganisms is lacking. Rumen fungi are strict anaerobes and widely distributed in the gastrointestinal tracts of many domestic and wild ruminant and non-ruminant herbivores (Gordon and Phillips, 1998; Liggenstoffer et al., 2010). Their high efficiency in plant cell wall degradation has been documented and it is undeniable that they play an important role in rumen function (Borneman and Akin, 1994). It is now generally known that the degradation of plant cell wall polysaccharides by rumen fungi accelerates digestion by decreasing the particle size of plant tissues, as not only do they produce highly active
fibrolytic enzymes, but are also capable of physically disrupting the plant cell walls including the cuticle via penetrating hyphae of vegetative thalli (Nagpal et al., 2009; Orpin and Joblin, 1988). Zoospores, the mobile phase of the fungal life cycle, also preferentially colonize lignin-rich regions of the plant cell wall, establish colonies localized on sclerenchyma and xylem cells and solubilize these regions to a greater extent than rumen bacteria (Akin and Borneman, 1990). As lignin is not degraded under anaerobic conditions, hydrolysis of both the ester linkages and sugar residue branches existing in hemicellulose plays a key role in the solubilization of lignin and exposure of hemicellulose to microbial xylanases. Thus, fungi play a key role in facilitating plant cell wall degradation by other members of the rumen microbial community.

*Anaeromyces* is one of six genera of anaerobic fungi currently identified (Li et al., 1993; Liggenstoffer et al., 2010). It has two defined species originally isolated from the ovine rumen: *Anaeromyces mucronatus* (Breton et al., 1990) and *Anaeromyces elegans* (Ho et al., 1993). The genus *Anaeromyces* is characterized by a polycentric thallus, a polynuclear rhizomycelium and uniflagellated zoospores. *A. mucronatus* is known to produce a broad range of intracellular and extracellular enzymes involved in the degradation of plant structural and storage polysaccharides including cellulase, xylanase, β-glucosidase, mannosidase, amylase and chitinase (Novotná et al., 2010; Yang and Yue, 2012). Like other rumen fungi, the genome of *A. mucronatus* has an extremely rich AT content (GC% approx. 20%), and this property greatly increased the difficulties associated with molecular manipulation and sequence analysis (Chen et al., 2006; Nicholson et al., 2005). Based on molecular analysis, *Anaeromyces* clusters separately from *Orpinomyces, Neocallimastix,* and *Piromyces* (Li and Heath, 1992) suggesting that
it may possess some unique attributes. Presently very few studies have characterized the genomic nature of this unique microorganism (Qi et al., 2011).

In this chapter, the strain *Anaeromyces mucronatus* YE505 was investigated for its ability to utilize various carbon sources, and the transcriptome produced on each carbon source was characterized using an Illumina sequencing platform. Subsequently, the transcriptomes were analyzed comparatively through assembly of the RNA-Seq short reads into full length ORFs to generate a comprehensive view of the plant cell wall degrading enzymes and their regulation. By applying tandem mass spectrometry analysis a number of putative CAZy proteins in the extracellular culture fraction were also identified.

### 4.2 Material and methods

#### 4.2.1 Chemicals

Unless otherwise noted, all the chemicals utilized were reagent grade or higher, purchased from either Sigma Aldrich (Oakville, Ontario, Canada) or Fisher Scientific (Ottawa, Ontario, Canada).

#### 4.2.2 Fungal strain and culturing conditions

*A. mucronatus* YE505 was originally isolated from an elk (Hausner et al., 2000) and was grown anaerobically at 39 °C in modified semi defined Lowe’s medium B (Lowe et al., 1985) with 0.67% (wt·vol⁻¹) of one of the following carbon sources added: 1) Glucose, 2) Cellobiose, 3) Glucose-cellobiose-starch (GCS, weight ratio 1:1:1), 4) Avicel cellulose, 5) Oat spelt xylan, 6) Barley straw, or 7) Alfalfa hay.
For growth curve and extracellular enzyme activity assays, gas volumes produced were measured and liquid cultural samples were taken from three cultural replicates at the same time point every 24 hours for a period of 10 days. Samples were frozen immediately at -20 °C prior to further analysis.

4.2.3 Enzyme assays

Enzyme assays were carried out in 50 mmol·L⁻¹ sodium phosphate buffer, pH 6.5 at 37 °C unless otherwise stated. One unit (U) of enzyme activity was defined as one µmol of product released (glucose equivalent, p-nitrophenol or α-naphthol) per minute. Background corrections were made by subtracting the readings for assays conducted with heat-inactivated enzymes.

Glycoside hydrolase activities on polysaccharides were assayed by incubating appropriately diluted enzyme samples in an assay mixture containing one of the following substrates: 1% (wt·vol⁻¹) low viscosity carboxymethylcellulose (CMC), 1% Avicel cellulose, 1% oat spelt xylan, 1% starch or 1% lichenan. A standard incubation period of 1 h was used. Released reducing sugars were detected by the p-hydroxybenzoic acid hydrazide (PAHBAH) method as described by Lever (1972). The absorbance of each assay mixture was read at 420 nm using a Synergy – HT microtitre plate reader (BioTek, Winooski, VT). The amount of reducing sugar produced by the enzyme was calculated by reference to a glucose standard curve.

Glycoside hydrolase activities on arylglycosides were determined by incubation of appropriately diluted enzymes in an assay mixture containing 5 mmol·L⁻¹ of the substrate: p-nitrophenyl-β-D-cellobioside (pNPC) or p-nitrophenyl-β-D-glucoside (pNPG). The reaction was stopped by addition of an equal volume of 1 mol·L⁻¹ Na₂CO₃ and released
p-nitrophenol was determined by measuring the absorbance at 405 nm with a Synergy – HT microtitre plate reader using p-nitrophenol as the reference standard.

The acetyl esterase assay procedure described by Qi et. al. (2011) was used with the following minor modifications. In a 100 µl reaction mixture, 1 mmol·L⁻¹ α-naphthyl acetate (α-NA) was used as the substrate. After incubation at 37 °C for 30 min, 50 µl of 2.5 mg·ml⁻¹ (wt·vol⁻¹) fast garnet GBC in 10% SDS was added. Absorbance at 560 nm was measured in a Synergy – HT microtitre plate reader with α-naphthol as the reference standard.

4.2.4 Isolation of total RNA

For isolation of mRNA for deep sequencing, A.mucronatus YE505 was grown anaerobically at 39 °C for 72 h in Lowe’s medium B with 0.67% (wt·vol⁻¹) of the various carbon sources described in section 4.2.2. The mycelia were harvested from the culture medium by vacuum filtration through four layers of cheesecloth and immediately frozen in liquid nitrogen. Total RNA was subsequently isolated from the mycelia by following the improved total RNA isolation procedure for solid rumen samples established in Chapter 2.

4.2.5 Sequencing and sequence assembly

Equal amounts of RNA extracted from fungal cultures grown on GCS, Avicel, xylan, barley straw and alfalfa hay were pooled together to constitute a mixed sample and sequenced. The RNA samples from fungal cultures grown on GCS, xylan, barley straw and alfalfa hay were also sequenced individually. Using the Illumina mRNA-Seq sample preparation kit according to the manufacturer’s instructions (Illumina Inc, San Diego,
mRNA libraries were constructed. High throughput sequencing was performed on an Illumina HiSeq 2000 sequencer system at the McGill University/Génomique Québec Innovation Centre.

Obtained sequencing reads were assembled *de novo* with two assemblers: the Trinity assembler (Grabherr et al., 2011) using the "jellyfish" method fork-mer counting, and the Velvet assembler (Zerbino and Birney, 2008). The assembly with Velvet was done on a split dataset comprised of 45 sets of 2 million reads, then reassembled with CAP3 (Huang and Madan, 1999). Because the contigs from the Trinity and Velvet assembly are highly redundant, results from the two analyses were combined and a dataset with 95% sequence identity was obtained with the CD-HIT program (http://www.bioinformatics.org/cd-hit/) (Li, 2009). These contigs were translated in all six reading frames and those that comprised at least 150 amino acids and possessed a start and stop codon were considered to be a potential open reading frame (ORF).

### 4.2.6 Bioinformatic sequence analysis

All sequence analyses, unless otherwise specified were performed using the assembled full length ORFs. The databases employed for this analysis were the latest versions available during the analysis period (May 2012 to Sep 2012).

The predicted ORF sequences were searched using RPS-BLAST against both the KOG and the COG databases (Tatusov et al., 2003) and the GenBank non-redundant amino acid (nr) database. The functional roles of the sequences were assigned based on KOG and COG searches. Matches that had E-values less than or equal to $10^{-5}$ were considered significant. CAZy protein annotation was performed by both HMMER3 (Eddy, 2009) and BLASTX searches following the procedures described in Chapter 3.
Relative transcript expression levels were calculated by the FPKM (Fragments Per Kilobase of transcript per Million mapped) method using Cufflinks software (Trapnell et al., 2010). Expression data analysis was performed using Spotfire Software (Spotfire Inc., Somerville, MA, USA). Cluster analysis of the log$_2$-transformed transcript expression data obtained from CAZymes was carried out using UPGMA (Unweighted Pair Group Method with Arithmetic Mean) analysis, with an ordering function based on the input rank. The resulting dendrograms were generated with the closest branches of the tree representing samples with similar gene expression patterns.

4.2.7 Extracellular protein preparation

Extracellular proteins were collected at the same time as mycelia were collected for total RNA isolation by collecting the liquid fractions, which were filtered through cheesecloth as described in section 4.2.4. The liquid fraction was centrifuged at 10,000 x g for 30 min and supernatants were concentrated ~100-fold by ultrafiltration using an Amicon concentrator with an Ultracel YM-10 membrane (Millipore, Billerica, MA, USA). The samples were further cleared of particulate matter by centrifugation at 3200 x g for 30 min at 4 °C. Proteins were precipitated by adding 1 equal volume of precipitating solution (0.2% dithiothreitol, 20 % trichloroacetic acid in acetone) followed by incubation for 1 h on ice. The pellet was subsequently washed with prechilled 20 mmol·L$^{-1}$ DTT/80% acetone (-20 °C).

Pellets were resuspended in 2 x Laemmli buffer (~ 50 µL) and quantified using the RCDC Protein Assay Kit (Bio-Rad, Mississauga Ontario, Canada). Isolated protein (15 µg) was loaded in each lane of a 10% SDS-PAGE gel. After electrophoresis and visualization by staining, the whole lane was excised, proteins were destained, reduced,
cysteine-alkylated and in-gel digested with trypsin over-night as previously described (Wasiak et al., 2002).

4.2.8 LC-MS/MS and peptide identification

The peptide extracts of each corresponding lane were subjected to LC-MS/MS on a Velos LTQ-Orbitrap (ThermoFisher Scientific, San Jose, CA, USA) instrument at McGill University and Génomique Québec Innovation Centre.

Mass spectrometric data were acquired by employing the Data Dependent Scans from the Xcalibur 2.1 software (Thermo Fisher Scientific, San Jose, CA, USA). Raw data from LC-MS/MS were processed with the Proteome Discoverer 1.3 software (Thermo Fisher Scientific). The peaklist files were searched against the generated database of the de novo assembled RNA-Seq contigs. SEQUEST (Thermo Fisher Scientific) was used to carry out peptide identification using a fragment ion mass tolerance of 0.80 Da and a parent ion tolerance of 10.0 ppm. An iodoaceticamide derivative of cysteine was set as static modification and oxidation of methionine was specified as a variable modification. Peptide identifications were accepted if they could be established at high probability and with a SEQUEST score greater than 2.5. Protein identifications were accepted if they could be established at high probability and at least two unique peptides assigned.

4.3 Results

4.3.1 Growth of Anaeromyces mucronatus YE505 on different carbon sources and enzyme activity in extracellular culture fluid

The ability of A. mucronatus to grow on a number of carbon sources in a chemically semi-defined medium designed by Lowe (Lowe et al., 1985) was tested using
glucose, cellobiose, GCS, xylan, avicel, barley straw and alfalfa hay. Rumen fungi produce both hydrogen and carbon dioxide during fermentation so the volume of gas produced can be used as a crude indicator of active growth. As demonstrated by the gas production curves in Figure 4.1, YE505 yielded approximately 2 fold higher gas volume on GCS or xylan in comparison to alfalfa hay or barley straw. Growth on Avicel had an extended lag phase lasting nearly 96 h; however, culture yields approached those achieved on GCS and xylan after 8 – 10 d growth. Suprisingly, little growth was detected on media containing only glucose or cellobiose as the sole carbon source. On most substrates, the highest rate of gas production occurred between day 2 and day 5. This period of rapid growth was shifted by 3 d when YE505 was grown on Avicel (Figure 4.1). These results were also verified by visual observation of mycelial growth inside the culture container. Thus, thereafter only the five carbon sources (GCS, Avicel, xylan, barley straw and alfalfa hay) were selected as growth substrates for futher studies with *A. mucronatus* YE505.

As the first step to define the expression level of carbohydrate degrading enzymes under different culture conditions, the extracellular cultural fluid (ECCF) of *A. mucronatus* grown on five carbon sources was sampled over a period of 10 days, and the samples’ activities toward eight enzyme substrates were measured as described in section 4.2.3 (Figure 4.2 - Figure 4.9).

Enzyme activities detected in the ECCF against the substrate Avicel were barely detectable regardless of the carbon source used to support fungal growth (Figure 4.2). The activities on CMC (Figure 4.3) - endoglucanase activities, on lichenan (Figure 4.4), on starch (Figure 4.5) and on xylan (Figure 4.6) were found to increase rapidly after 24 h,
and were generally higher when the fungus was grown on alfalfa hay and barley straw as compared to other carbon sources. Interestingly, *A. mucronatus* grown on Avicel after 8 days showed comparable CMCase, xylanase and lichenase activities as those grown on alfalfa hay and barley straw. This was consistent with the gas production (Figure 4.1), in that when *A. mucronatus* was grown on Avicel it required a longer time to reach the active growth phase. Amylase activities remained low in ECCF from Avicel even after 8 days. An interesting aspect was that when grown on GCS *A. mucronatus* did not exhibit high amylase activity in ECCF even though the medium contained starch. Growth on xylan generated xylanase activities that were higher than GCS, comparable to Avicel (for the first 6 days), but lower than alfalfa and barley straw. The xylanase activities in GCS medium were barely detectable over the 10 day experiment.

For the synthetic substrates tested, the same trends were observed: enzyme activities were observed sooner in the incubation period and were higher when *A. mucronatus* was grown on alfalfa hay or barley straw. Compared to other synthetic substrates, activities to pNPC remained low throughout 10 day incubation, regardless of substrate (Figure 4.7). β-glycosidase activities detected by the pNPG assay were comparably high in alfalfa and barley straw, with that from Avicel reaching similar levels after 7 d of incubation (Figure 4.8). *A. mucronatus* grown on alfalfa exhibited the highest esterase activity to α-NA, being two fold higher than that grown on barley straw (Figure 4.9), but growth as reflected by gas production was comparable on these two substrates (Figure 4.1).

Generally speaking, the enzyme activity levels started to increase on day 2 and reached the highest level after day 4 or day 5, which was in agreement with the active
growth period demonstrated by gas accumulation. *A. mucronatus* grown on alfalfa produced the highest detectable ECCF activities to five substrates (CMC, lichenan, starch, xylan and α-NA); whereas activity against pNPG was higher in barley straw.

Based on the growth curve and enzyme activity profiles, we chose 96 h for Avicel and 72 h for the rest of carbon sources as the time points to harvest mycelia for RNA extraction, as well as ECCF collection. By choosing these time points, we were able to obtain samples during the most active growing period that yielded sufficient levels of RNA and extracellular protein for analysis.

### 4.3.2 Sequencing summary, assembly and BLAST analysis

In the present study, the RNA-Seq approach was applied to identify ORFs coding for proteins and to define the level of transcript expression in *A. mucronatus* YE505 grown on different carbon sources. Total RNA was extracted from mycelia and after purification, the mRNA samples produced on four different carbon sources (i.e., GCS, xylan, alfalfa hay and barley straw) were sequenced using an Illumina HiSeq 2000 sequencer. Unfortunately, the mRNA sample obtained from *A. mucronatus* grown on Avicel was not sequenced individually; however, it contributed 20% of the RNA in the mixed samples, and consequently contributed to the assembled sequences from this sample.

A total of 89,241,369 paired end 100 x 100 base reads were obtained. The raw sequencing reads were assembled with both the Trinity and Velvet assemblers and the resultant contigs were combined. Potential full length ORFs were predicted as described in section 4.2.5. A total of 6,670 full length ORFs were obtained, with an average length of 1,427 bp, for a total of 9.52 Mb. Size distribution of the ORFs is illustrated in Figure
with the longest ORFs being 16.5 kb encoding for a hypothetical protein with approximately 5,500 amino acid residues. This protein showed high similarity to a hypothetical protein from another rumen fungi, *Orpinomyces* sp. OUS1 (Nicholson et al., 2005). A large hypothetical protein showing similarity to this protein was also detected within the muskoxen rumen eukaryotic metatranscriptome described in Chapter 3. The average GC content of all the ORFs was 28.8%, which was expected owing to the AT-rich genomes of rumen fungi.

A BLAST search was performed using all the potential ORF sequences against the nr database from GenBank. A total of 1,200 sequences from *A. mucronatus* were most similar to sequences from the chytrid *Batrachochytrium dendrobatidis*. BLAST searches also identified *A. mucronatus* homologues to over 600 sequences from *Rhizoporusoryzae* and ~160 from *Myceliophthora thermophila*. Interestingly, over 600 ORFs showed high similarity to bacterial derived genes as opposed to those originating from eukaryotic species.

### 4.3.3 Functional analyses of ORFs identified

Based on RPS-BLAST search results, 3,808 ORFs could be assigned to a cluster of the KOG database (Figure 4.11, pane A). From the rest non-KOG matched candidates, further RPS-BLAST identified 422 ORFs that matched to the bacterial genome derived COG database (Figure 4.11, pane B).

The KOG/COG databases classified each ORF into one of 20 groups (Tatusov et al., 2003). Proteins involved in “translation, ribosomal structure and biogenesis” had the highest total FPKM number (Figure 4.11), indicating that the fungal mycelia were harvested during active growth. Particularly, AmuTC4, an ORF of 1,359 bp encoding for
α unit of the elongation factor 1 that delivers aminoacyl-tRNAs to the ribosome during protein translation, was represented by a total FPKM of 136,764, or 5.3% of all the sequencing reads. Many putative genes coding for roles in “carbohydrate transport and metabolism”, and “energy production and conversion” were also highly expressed.

Most of the Embden-Meyerhof-Parnas pathway enzyme sequences were recovered from our dataset. Some of these ORFs were transcribed at very high levels. Among these, ORFs coding for enolase, glyceraldehyde-3-phosphate dehydrogenase and aldolase were the most highly expressed. Genes involved in the partial TCA cycle were also highly expressed including a phosphoenolpyruvate carboxykinase and an NAD-dependent malate dehydrogenase.

Many ORFs with a very high FPKM were not members of the above-mentioned categories. One group included genes involved in rhizoidal growth. Among these, ORFs coding for actin and tubulin, important cytoskeleton components in actively growing fungi, were represented by FPKMs of 36,853 and 21,510, respectively. Several putative genes that function in amino acid transport and metabolism were also highly expressed, including those encoding ketol-acid reductoisomerase (which is known to be involved in the biosynthesis of leucine/isoleucine/valine), glycine/serine hydroxymethyltransferase, lysine-ketoglutarate reductase-saccharopine dehydrogenase, and argininosuccinate synthase.

The function of some highly transcribed ORFs, like AmuTC14, could not be accurately predicted by KOG/COG analysis. This putative gene was annotated as coding for a sugar transporter, but without biochemical characterization, its substrate preference could not be predicted.
4.3.4 CAZymes prediction

Like other rumen anaerobic fungi, *A. mucronatus* degrades a number of plant cell wall polysaccharides including cellulose, xylan, mannan and pectin (Novotná et al., 2010). From the RNA-Seq data, a total of 344 ORFs containing at least one CAZy domain were identified, including glycoside hydrolases (GHs), carbohydrate esterases (CEs), pectin lyases (PLs) and carbohydrate binding modules (CBMs) (Figure 4.12, Table S.3). The length of the ORFs for CAZy genes ranged from 491 bp to 5,916 bp, with an average coding sequence length of 1,592 bp.

Translated BLAST searches against the GenBank nr database showed that only 12% of the products coded by these ORFs were more than 70% identical to proteins in the nr database, while 59% of them showed an identity of less than 50%. Altogether 121 ORFs coded for two or more distinct CAZy domains, among which, CBM10 like domains were in 93 ORFs, accounting for 7% of over 6,000 total full length ORFs obtained.

4.3.4.1 Catalytic modules

Most of the putative cellulases identified from *A. mucronatus* transcriptomes were classified as members of families GH5, 6, 9, 45 and 48, representing a total of 35 ORFs. In addition, for the first time from an anaerobic fungus, five ORFs were predicted to code for swollenin, a protein without hydrolytic activity that has been reported to dissociate cellulose fibers (Brotman et al., 2008). Open reading frames matching GH families containing β-glycosidases, including GH1 and GH3, were also identified. The ORF AmuTC51 harbouring a potential GH1 module was highly expressed on alfalfa, xylan and barley straw. The encoded protein was 72% identical to a β-glycoside hydrolase from
*Piromyces* sp. E2 that was also highly expressed when this fungus was grown on fructose (Harhangi et al., 2002).

Xylan and other hemicellulose degrading enzymes were identified from 39 ORFs. They were mostly from GH families 10, 11, 26, 39 and 43. In this study, seven GH10 and six GH11 ORFs were identified. All the GH10 enzymes were associated with at least one CBM. Sixteen ORFs that contained GH43 domain were identified. Two ORFs encoding putative xylosidases from GH39 were identified, both of which contained a CBM13 and a CBM10 module. AmuVC11847, an ORF coding for a GH8 module coupled with two tandem CBM10 modules, was also identified. The predicted gene product was 74% identical to a putative xylanase in *Fibrobacter succinogenes* (Suen et al., 2011b). To our knowledge, this is the first GH8 family member reported for a rumen fungus.

There were 59 ORFs showing similarity to carbohydrate esterases in the CAZy database. The majority of these (20 ORFs) were associated with family CE4. The CE1 family, which contains feruloyl esterases, was represented by 10 ORFs. A total of 13 ORFs coded for family CE10 CAZymes, which includes pectin acetyl esterases. Two CE15 ORFs encoding for 4-O-methyl-glucuronoyl methylesterases were also identified. As with feruloyl esterases, CE15 family enzymes cleave cross linkages between lignin and the xylan backbone.

Other catalytic modules identified included pectin degrading enzymes from GH53 (endo-β-1,4-galactanase) as well as from the polysaccharide lyase (PL) family 1, 2, 4 and 9.
4.3.4.2 Accessory modules

The most prominent accessory module was CBM10. A total of 350 CBM10 modules were identified in 162 ORFs. Most of the ORFs have 2~3 tandem CBM10 domains. A total of 91 CBM18 domains were identified in 40 ORFs. The number of CBM18 coded by one ORF ranged from one to 15. Other major carbohydrate binding modules include CBM1 (35 ORFs), CBM13 (20 ORFs), CBM6 (6 ORFs) and CBM29 (2 ORFs).

4.3.4.3 Effect of carbon source on CAZyme gene transcription in

*Anaeromyces mucronatus* YE505

The expression of plant cell wall degrading enzymes in *A. mucronatus* YE505 was influenced by carbon source (Figure 4.12). For all carbohydrate active enzyme transcripts identified, the average FPKMs ranged from 71 in GCS to 206 in barley straw, 209 in oat spelt xylan and 321 in alfalfa hay. Generally speaking, the majority of predicted CAZymes were most induced when *A. mucronatus* was grown on alfalfa hay, followed by barley straw and xylan, and the lowest when grown on GCS.

The transcript expression levels of individual ORFs differed considerably, regardless whether or not they belonged to the same CAZy family. For example, ORFs *Amu*TC352 and *Amu*VC1295 both belong to GH11. Expression of *Amu*TC352 was higher when *A. mucronatus* YE505 was grown on GCS or xylan, while expression of *Amu*VC1295 was 25 and 125 times higher on alfalfa hay and barley straw as compared to expression on GCS and xylan, respectively. This strongly suggests that there is no one-strategy-fits-all for regulation of gene expression within GH families. On the contrary, individual genes are likely subject to differential regulation individually. The same
situation was observed for CAZymes with the CBM10 modules, where the ratio of the FPKM for growth on alfalfa, straw and xylan as compared to GCS ranged broadly from 0.04 (repressed) to 1,260 (induced).

To better illustrate the complex expression patterns of different CAZymes, ORFs with similar expression patterns were grouped using the UPGMA method and normalized expression levels (i.e., log2FPKM; Figure 4.12). The gene transcription profile for A. mucronatus YE505 grown on GCS differed substantially from those grown on xylan, barley straw or alfalfa hay. Expression patterns were most similar for cultures grown on barley straw and alfalfa hay. The CAZyme transcripts grouped into seven clades according to expression levels and patterns when A. mucronatus YE505 was grown on different substrates.

Clade A, C and E generally showed less variation among the four carbon sources. In clade A, 10 ORFs were expressed at high levels regardless of substrate, with average FPKM ranging from 1,227 for GCS to 4,246 for alfalfa hay. This group included AmuTC99, an ORF encoding a putative GH43 β-xylosidase that cleaves xylo-oligosaccharides into xylose. The transcript expression level for this ORF was very high, especially when grown on xylan (Figure 4.12). Within this clade, there were two CAZymes belonging to GH6 and GH48 with putative exoglucanase activity. Intriguingly, ORF AmuTC72 contained two CBM29 and three CBM10 domains, but no catalytic module. Two CE4 esterases and two CAZymes harbouring a CBM13 and a CBM10 module were also identified within this group of 10 enzymes.
Clade C was a large group that contained 62 CAZymes, which were expressed at lower levels than clade A. Average FPKMs within this group ranged from 118 to 439 and were comparable across substrates.

Clade E also represented a large group containing 72 ORFs that were moderately expressed with average FPKMs ranging from 47 to 119. Interestingly, some of the ORFs (e.g., AmuVC2993 and AmuVC4852) in this clade were down regulated when *A. mucronatus* was grown on alfalfa hay or barley straw.

Eight CAZymes belonging to Clade B were induced when *A. mucronatus* was grown on barley straw or alfalfa hay, as illustrated by the dramatic increase in average FPKMs 17 for GCS to 994 in straw and 2,030 in alfalfa. ORFs in this clade included a GH6, a GH10 and a GH11 member. Within this clade, two esterases from CE6 and CE15 were each linked to a CBM10 module. Three CBM10 harbouring ORFs without known catalytic domains were also members of this clade.

Similarly, the 11 CAZymes in clade D were expressed at low levels with GCS, compared to FPKM values 100 fold higher for barley straw and alfalfa hay. The ORFs in this clade were more difficult to annotate with only two catalytic domains (a GH8 and a CE4) being identified. The remainder of CAZy proteins in this clad appeared to be CBMs, which appeared to be upregulated when *A. mucronatus* was grown on complex plant fibers.

Similar to clades B and D, Clade F contained 56 ORFs that were induced by xylan, straw and alfalfa. Their expression levels as well as their regulation tended to be lower and less stringent with less variation between different carbon sources, indicating they may play a less important role in fiber degradation.
The remainders of the putative CAZy genes were expressed at low levels with FPKMs averaging from 8 to 35, and were grouped into clade G. Two of these CAZymes, a GH26 member and a CE15 member were expressed at much higher levels when \textit{A. mucronatus} was grown on alfalfa hay than on other carbon sources.

4.3.5 Secretomic analysis of \textit{Anaeromyces mucronatus} YE505 grown on different carbon sources

To further characterize the plant cell wall degrading enzyme system present in the ECCF, a secretomic study was performed to analyze the ECCF protein profiles produced on five different carbon sources. Together with the other four samples, the Avicel grown ECCF protein sample was also subjected to LC-MS/MS and secretomic analysis. In total, over 3,400 peptide spectra were identified, which matched to 341 ORFs, including 103 ORFs belonging to predicted CAZy proteins (Table 4.1). Transcripts with higher FPKM values as determined by RNA-Seq were more likely to have a protein counterpart identified by protein mass spectroscopy, such as AmuTC51, AmuTC352 and AmuTC99. Interestingly, alfalfa hay or barley straw grown samples generated the fewest number of transcript.peptide matches. Growth on Avicel yielded a peptide subset of CAZy proteins that were not detected in the other samples (Table 4.1).

4.3.6 ORFs without predicted function

After combined search of COG/KOG, pFAM, CAZy and nr databases, about 2,200 ORFs still could not be assigned a function. This accounts for approximately one third of the total ORFs obtained. BLAST searches against the nr database showed about 550 of these coded for conserved hypothetical proteins, which matched to proteins with
unknown function from other organisms. The remaining 1,650 ORFs showed no match to any reported sequence. As these ORFs were assembled from mRNA, they are likely to represent functional genes. Some of these ORFs, such as AmuVC11385 and AmuTC42 were expressed at higher levels when *A. mucronatus* was grown on alfalfa hay, barley straw and xylan. These genes may be involved in some unknown pathway of plant cell wall digestion and warrant future investigation.

### 4.4 Discussion

This study identified a number of CAZymes from different families with differing substrate specificity. CAZymes act synergistically to enable rumen fungi to degrade recalcitrant plant cell wall polysaccharides (Blum et al., 1999). Difference in expression levels with different carbon sources was clearly demonstrated by the grouped log$_2$FPKM patterns in Figure 4.12. Logarithm scale (log$_2$) analysis is an accepted and widely applied approach for normalizing relative expression levels obtained through microarray and RNA-Seq analysis of microorganisms grown on different substrates (Brooks et al., 2011; Marioni et al., 2008).

The extracellular enzyme activities were calculated based on the ECCF volume instead of the protein mass, since the protein concentrations in the ECCF were too low to be accurately measured, especially when components in the medium could interfere with the protein concentration assay. Apparently, growth rate would influence the amount of proteins secreted, and consequently enzyme activities detected. But on the contrary, despite the relatively moderate growth rates, the enzyme activities on alfalfa hay and barley straw were generally the highest, being more than 10 fold higher than those on GCS in the case of xylan and α-NA activities (Figure 4.6, Figure 4.9) during the active
growing period. Meanwhile, the esterase activities were ~ 2 fold higher with alfalfa hay than with barley straw while the gas production of *A. mucronatus* on these two substrates was comparable. This higher level of induction was probably because alfalfa cell wall contains more ester linked branches compared to barley straw (Varel et al., 1989). Therefore, when considering the growth rate influence, differences in enzyme activities between complex carbon sources and GCS would likely have been even greater than detected and were most probably due to differences in gene expression. This observation is in agreement with the expression patterns identified from RNA-Seq, as generally CAZymes were more highly expressed when grown on alfalfa hay or barley straw.

The LC-MS/MS procedure applied in this study is not considered quantitative, but still there was a weak correlation that the ORFs with high FPKMs were more likely detected by LC-MS/MS. Occassionally more peptide spectra from Avicel, xylan or GCS grown samples were matched to one ORF than those from alfalfa hay or barley straw, regardless whether it was up-regulated by the latter carbon sources (Table 4.1, Figure 4.12). This was probably because the protein samples grown on simple carbon sources (Avicel, xylan and GCS) contained fewer impurities and was more suitable for mass spectral detection in preparations.

It has been previously reported that xylanases were predominant among the glycoside hydrolases in *A. mucronatus* strain KF8 (Novotná et al., 2010). Our findings through activity assays and transcriptomic sequencing agree with this report. Among all the predicted CAZy families, GH43 enzymes coding mainly for xylanases or xylosidases had the highest FPKMs, together with some GH10 and GH11 xylanase candidates. But contrary to enzyme assays, which showed that xylanase activities were highest with
alfalfa hay followed by barley straw (Figure 4.6), the RNA-Seq analysis showed that fungi grown on xylan generated the highest xylanase FPKM values (Figure 4.12). And very surprisingly, xylanase activity was almost undetectable in the ECCF from A. mucronatus grown on GCS (Figure 4.6), even though considerable xylanase expression was detected in GCS grown transcriptome. One possible explanation is that direct enzyme assays only reflected those xylanases that were secreted into the medium, whereas RNA-Seq detected the overall expression including those from intracellular, cell-associated and extracellular xylanases. This may well reflect the possibility that a large portion of xylanases are tightly associated with the fungal cell surface. Another aspect that is worth pointing out is that the existence of highly expressed enzymes does not always equal to high enzyme activities, since specific activities can vary considerably between different enzymes. It is also possible that genes may not be effectively translated to proteins even when they are highly transcribed. All these factors contributed to the difficulties for clearly elucidating and comparing all the details in transcriptomic studies.

Xylanases from GH10 and GH11 families have been previously cloned and characterized from rumen fungi (Black et al., 1994; Gilbert et al., 1992). But to date, other than a report from Neocallimastix patriciarum (Wang et al., 2011), no GH43 enzymes have been identified from rumen fungi. The enzyme assay also detected very high esterase activities when A. mucronatus YE505 was grown on alfalfa hay as well as on barley straw (Figure 4.9). Indeed, transcriptomic sequencing revealed that CE4 members were highly expressed, together with CE6 and CE15 members also contributing activities (Figure 4.12). These highly expressed ORFs such as AmuTC99 (GH43, with 38 Mass Spec peptide matches), AmuTC468 (CE4) and AmuVC9821 (CE15, which was
highly induced in alfalfa hay) suggest that they have an important function in plant cell wall degradation and are good candidates for future detailed biochemical analysis considering their uniqueness and high level of expression.

Other than the ordinary CAZy members involved in different stages of plant cell wall degradation, a particular group of hypothetical swollenins was of great interest to us. Swollenin was firstly identified from *Trichoderma reesei* (Saloheimo et al., 2002) and was subsequently identified in several other aerobic fungal species (Chen et al., 2010; Yao et al., 2008), and exhibits sequence similarity to plant expansins. Swollenin is believed to disrupt the structure of crystallized cellulose by breaking hydrogen bonds between cellulose fibers, without detectable formation of reducing sugars. Pretreatment of cellulose with swollenin decreased the particle size of cellulosic substrates as well as cellulose crystallinity and increased cellulose hydrolysis rates by cellulases (Jäger et al., 2011). In the present study, five putative swollenin genes were identified, all of which were also attached to CBM10-like modules. Four of these rumen fungal putative swollenin genes showed higher FPKM values when the fungus was grown on alfalfa hay, barley straw or xylan as compared to GCS medium. Swollenin like sequences were also identified from muskoxen rumen eukaryotes metatranscriptome as outlined in Chapter 3. All of these results suggest swollenin plays an important role in rumen fungal cellulose digestion and is worthy of further characterization.

A reason for anaerobic fungi’s high cellulolytic capability may be attributed to the possible presence of cellulosomes. High-molecular-mass enzyme complexes (>700 kDa) have been described in *Piromyces*, *Orpinomyces*, and *Neocalimastix*, which contain as many as 15 protein components (Ali et al., 1995; Li et al., 1997; Wilson and Wood,
This high-molecular-mass structure resembles the cellulosomes produced by several anaerobic bacteria such as *Clostridium* and *Ruminococcus* spp. in which various cellulases are attached to a protein called scaffoldin through a dockrin-cohesin interaction, resulting in a complex that is very efficient at cellulose degradation (Bayer et al., 2008).

A CBM10 like, 40-amino-acid cysteine-rich, non-catalytic domain has been shown to be associated with many rumen fungal glycoside hydrolases, and was proposed to be a fungal dockerin by some researchers (Nagy et al., 2007; Raghothama et al., 2001; Steenbakkers et al., 2001). But to date, no cellulosome scaffoldin that mimics the bacterial counterpart has been identified from rumen fungi, although it has been shown that a CBM10 domain can interact with a GH3 enzyme in *Piromyces equi* (Nagy et al., 2007). This finding is surprising since unlike other scaffoldins in bacteria, this protein possesses only a catalytic domain and no identifiable cohesion domain. Although no promising candidate has been detected yet, transcriptome sequencing data generated from this study may facilitate the identification of the potential dockerin-interacting partners in the future as more evidence is accumulated.

Interestingly, it is still debatable whether or not the fungal CBM10 module is a true fungal dockerin. A number of previously identified anaerobic fungal genes have been shown to have one or more CBM10 like domains (Blum et al., 1999; Nagy et al., 2007; Steenbakkers et al., 2008; Steenbakkers et al., 2001). In the present study 350 CBM10 domains were identified in 162 ORFs (Figure 4.12). As expected, many of these ORFs had CAZy catalytic domains, including members with cellulase or xylanase activity from families GH1, 3, 5, 6, 9, 10, 11, 43, 45 and 48. Carbohydrate esterase family CE15 members were also associated with CBM10. But the fungal CBM10 module showed no
sequence homology to bacterial dockerins, and NMR structures of these modules from *P. equi* showed no structural similarity to bacterial dockerins (Nagy et al., 2007; Raghothama et al., 2001). In addition, the CBM10 module did not bind to either xylan or cellulose (Fanutti et al., 1995). Rather this module from *P. equi* recognized and bond to a glycosylated β-glucosidase via its oligosaccharide components (Nagy et al., 2007). Therefore the function of this module as dockerin remains to be further verified, with some researchers considering this module to be merely a fungal CBM with unknown carbohydrate preference (Peer et al., 2009).

A little surprisingly, no CAZy catalytic domains were identified from 78 of the CBM10 containing ORFs. The size of polypeptides encoded by these ORFs range from 161 to 1,750 residues, suggesting that it is likely that other non-CAZy catalytic functional domains may be present. In order to identify potential non-CAZy domains, BLASTP searches were performed using the 78 ORFs against the Genbank nr database. Over a half of them did not exhibit catalytic domains. Among them, the ORF AmuTC72 was highly expressed when YE505 was grown on xylan, barley straw or alfalfa hay and contained two CBM29 domains other than three CBM10 domains (Figure 4.12). The protein product of this transcript was also detected by LC-MS/MS (Table 4.1). Amino acid sequence of AmuTC72 showed 59% identity to NCP1 protein from *P. equi*, which was proposed to be a non catalytic protein that anchored the fungal enzyme complex onto the plant cell wall through CBM29-polysaccharide interaction (Freelove et al., 2001).

The other 31 ORFs possessed domains which shared similarities to serpin, CotH and Ser/Thr protein phosphatase, as well as swollenin as discussed above (Table 4.2). In seven ORFs, these CBM10-like domains were found to attach to a serpin-like protease
inhibitor. Serpins have been previously identified in the anaerobic fungus *Piromyces* sp. E2 (Steenbakkers et al., 2008). Fungal serpins are believed to be involved in protection of the cellulosome against proteases produced by plants or other microbes within the plant cell wall degrading community (Meguro et al., 2011; Steenbakkers et al., 2008). Similar protease inhibitors have also been identified in cellulosome producing bacteria such as *Clostridium thermocellum* (Zverlov et al., 2005), *Clostridium cellulolyticum* (Fendri et al., 2009) and *Clostridium cellulovorans* (Meguro et al., 2011). These proteins may play a critical role in conferring functionality to the cellulosome for the period of time that is required for plant cell wall degradation.

Another group of 13 ORFs contained a module loosely related to spore coat assembly protein (CotH) domain. CotH protein was first identified and characterized from *Bacillus subtilis* (Naclerio et al., 1996), which played a role in the assembly of at least nine other coat components into endospore protein shell called spore coat (Kim et al., 2006). A similar protein has also been identified in *Clostridium* sp. (Henriques and Moran, 2007). A CotH containing ORFs was also detected to be associated with a putative cellulase from *N. patriciarum* W5 (Wang et al., 2011). The role of the CotH-like domain in rumen fungi is unknown. It is logical to conjecture that it guides the formation of certain macromolecular protein structures. Our bold guess would be that the fungal CotH-like protein may function similar to the bacterial counterpart, directing the assembly of a subset of the zoospore’s spore coat proteins that help to protect the zoospores in the resistant stage, and with a CBM10 modules attached, may also facilitate the zoospores to attach to fiber rich substrates prior to germination and growth. Or rather cautiously, the CotH-like protein may play a role in the assembly of the fungal
cellulosome through interaction with other CAZy proteins, in conjunction with CBM10 modules (Nagy et al., 2007). Thus, CotH like protein may lead to a greater understanding of the cellulosome-like complexes that may be produced by rumen fungi.

Another six of the CBM10 containing ORFs showed similarity to Ser/Thr protein phosphatases. It is commonly known that protein phosphatases are very important regulators of variety of physiological processes, such as cell cycle control, regulation of cell growth and division (Depaoli-Roach et al., 1994; Ingebritsen and Cohen, 1983). However, the role of Ser/Thr protein phosphatase domains associated with potential cellulosome CBM10 like domain will require further detailed biochemical analysis.

Although the exact function of the fungal CBM10 module may currently remain unclear, its high prevalence suggests that it plays an important role. Besides the possibility to bind to some unknown plant cell wall components, a reasonable hypothesis is that it probably anchor a large variety of proteins onto the fungal cell wall, functioning as the fungal counterpart of the bacterial SLH (Fontes and Gilbert, 2010). Detailed binding assay will be needed to determine the components that this module interacts with.

It is unlikely that the association of CBM10 with a variety of different domains could be an artifact of the de novo assembly process. As the assembly pipeline used in this chapter was also applied to the muskoxen metagenomic study reported in Chapter 3, and artifact assembly was demonstrated to be minimum, if any. It is also in agreement with other researchers who frequently found the CBM10 module associated with putative GH proteins in the N. patriciarum transcriptome (Wang et al., 2011). On the other hand, because of the short length of CBM10 (only ~100 bp), the Illumina Hi-Seq sequencing reads supported the abundant expression of CBM10 domain, and subsequently indicated
its importance in the metabolism, with minimum assembly and regardless of whether
and/or which catalytic domains CBM10 was associated to.

The CBM18 domain was the second most abundant CAZy member predicted. From
40 ORFs, a total of 91 CBM18 domains were identified. CBM18 is known as a chitin
binding domain, which has been identified exclusively in eukaryotes, including fungi,
plants and arthropods (Suetake et al., 2000). The function of the CBM18 within rumen
fungi is unknown, but it has been characterized in other organisms. The domain binds to
fungal cell wall chitin, which in turn protects the fungi from the environmental chitinases.
This may be important for rumen fungi to maintain a stable biomass within the rumen
ecosystem, as several rumen microorganisms, including bacteria (Kopecny et al., 1996)
and ciliate protozoa (Morgavi et al., 1994) produce chitinases. The expression of CBM18
modules by rumen fungi may serve to block chitinase activity in the rumen, through
protective binding to fungal chitin. Multiple CBM18 modules existing in one ORF may
enhance the binding affinity to chitin. Three CBM18 modules were associated with
polysaccharide deacetylases (CE4). This enzyme removes the acetyl group from chitin
and converts it to chitosan. This function may also serve as a protection for the fungi as
chitosan is a poor substrate for chitinases (El Gueddari et al., 2002). Two ORFs
harbouring the CBM18 modules were associated with one CBM13 module. Since the
CBM13 can bind to xylan, these two ORFs may serve as a bridge that also promotes
fungal attachment to xylan.

To my knowledge, the study on *N. patriciarum* was the first and only published
transcriptomic study targeting an anaerobic fungus (Wang et al., 2011), displaying
several similar findings with my research reported here. In the absence of genomic
information, we both combined NGS based transcriptomic and LC-MS/MS based secretomic studies, with a focus on plant cell wall degrading enzymes. The *Neocallimastix* study identified a similar number of total 219 putative GH proteins, including the major induced cellulases from GH families 1, 3, 6, 10, 11, 43 and 48 with most of them possessing a CBM10-like domain. However, unlike the study of Wang et al. (2011), GH9 and GH45 members were also detected to be highly expressed in my study, possibly a reflection of genetic differences between two fungal genera as well as the use of different carbon sources for growth. My study also provided gene expression comparison among several carbon sources, and besides GH modules, CE, PL and CBM modules were also predicted. Unfortunately as no sequence data from the other study is publicly available, no further sequence comparison could be made.

With the continued growth of the human population and the demand for animal products increasing, it is likely that livestock agriculture will have to shift from low fiber/high concentrate diets to a greater reliance on high fiber feedstocks. This could lead to an increased interest in defining the importance of anaerobic fungi in rumen function (Nagpal et al., 2011). The results reported here demonstrated many previously unknown features of the rumen fungus *A. mucronatus*, with many potential candidate genes predicted. Putative swollenins were identified from a rumen fungus for the first time. It may be a good candidate as a feed additive or feed pretreatment agent, aiding in improving the digestibility of poor quality fibrous feeds for livestock production. The high abundance of CBM10 modules in association with various functional domains raise interests for more detailed characterization, and may facilitate to elucidate the nature of rumen fungal cellulosome-like complex. Many CAZyme candidates, such as members
from GH43, CE4 and CE15 may possess potential properties suitable for cellulosic biofuel industry or agricultural application. All the information will open up brand new avenues to illustrate the full potential of the anaerobic fungi in the future.
### 4.5 Tables and Figures

**Table 4.1** Matches of secretomic peptide detected by LC-MS/MS to predicted CAZy ORFs from RNA-Seq results when *Aneromyces mucronatus* YE505 grown on five different carbon sources.

<table>
<thead>
<tr>
<th>ORF</th>
<th>CAZy Domains</th>
<th>GCS</th>
<th>Avicel</th>
<th>Oat</th>
<th>Spelt</th>
<th>Xylan</th>
<th>Barley</th>
<th>Straw</th>
<th>Alfalfa</th>
<th>Hay</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AmuTC51</td>
<td>GH1</td>
<td>11</td>
<td>6</td>
<td>21</td>
<td>12</td>
<td>2</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC306</td>
<td>CBM26/GH31</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>28</td>
<td>47</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC352</td>
<td>GH11</td>
<td>12</td>
<td>4</td>
<td>16</td>
<td>8</td>
<td>3</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC99</td>
<td>GH43</td>
<td>11</td>
<td>3</td>
<td>11</td>
<td>7</td>
<td>6</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC2165</td>
<td>GH43</td>
<td>11</td>
<td>2</td>
<td>15</td>
<td>4</td>
<td>1</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC686_seq3</td>
<td>CE4/CBM10</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>10</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC11828</td>
<td>CBM10/GH3</td>
<td>13</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC9826</td>
<td>CBM13</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>12</td>
<td>2</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC72</td>
<td>CBM10/CBM29_Blast</td>
<td>5</td>
<td>3</td>
<td>12</td>
<td>6</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC10375</td>
<td>GH11</td>
<td>12</td>
<td>7</td>
<td>2</td>
<td>4</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC1910</td>
<td>CBM10</td>
<td>16</td>
<td>6</td>
<td>1</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC2498</td>
<td>CBM18</td>
<td>12</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC8</td>
<td>GH6</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC858</td>
<td>CBM13/CBM10/GH39</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC12166</td>
<td>CBM10/GH48</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC149</td>
<td>GH117/CBM6/GH43</td>
<td>14</td>
<td>3</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC290</td>
<td>CBM10/GH9</td>
<td>5</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC1002</td>
<td>CBM10</td>
<td>6</td>
<td>7</td>
<td>3</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC219</td>
<td>CBM10/GH45</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC557</td>
<td>CBM13/CBM10</td>
<td>6</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC12040</td>
<td>CBM10</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC468</td>
<td>CE4</td>
<td>8</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC1082</td>
<td>GH117/CBM6/GH43</td>
<td>13</td>
<td>1</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC459</td>
<td>CE15/CBM10</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC1982</td>
<td>GH115</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC10744</td>
<td>CBM10/GH48</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC585</td>
<td>GH3 /CBM10/GH6/GH3</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC2417</td>
<td>CBM10/GH26</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC12088</td>
<td>CBM13/CBM10/GH39</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC10088</td>
<td>CE1</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC623</td>
<td>CE4</td>
<td>7</td>
<td>3</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC732</td>
<td>CBM10/GH6</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORF</td>
<td>CAZy Domains</td>
<td>GCS</td>
<td>Avicel</td>
<td>Oat Spelt</td>
<td>Barley</td>
<td>Alfalfa</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------</td>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC169</td>
<td>CBM10</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC58</td>
<td>CBM10/CBM1</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC9724</td>
<td>CBM10/GH9</td>
<td>7</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC5285</td>
<td>CE10</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC2993</td>
<td>CBM10/GH11</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC2262</td>
<td>CE6/CBM10</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC12041</td>
<td>CE4</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC1325</td>
<td>CBM10/GH11</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC9760</td>
<td>CBM1/GH6</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC9693</td>
<td>CBM13/CBM10</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC3045</td>
<td>CE2 or CE3</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC12184</td>
<td>CBM10</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC1464</td>
<td>Swollenin/CBM10</td>
<td>6</td>
<td>2</td>
<td>6</td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC10470</td>
<td>CBM10/GH53</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC9781</td>
<td>CBM26/GH31</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC10313</td>
<td>CBM10</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC11854</td>
<td>CE6/CBM10</td>
<td>3</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC11861</td>
<td>CBM13/CBM10/CBM6/GH43/CBM36_Blast</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC309</td>
<td>Swollenin/CBM10</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC9956</td>
<td>Swollenin/CBM10</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC2143</td>
<td>CBM10</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC3998</td>
<td>CBM1</td>
<td>6</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC3673</td>
<td>CBM10/CBM6/GH43</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC10435</td>
<td>GH2/Bgal_small_N/CBM10</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC375</td>
<td>CBM18</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC10015</td>
<td>CE1</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC11110</td>
<td>GH117/CBM6/GH43/CBM36_Blast</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC12245</td>
<td>CBM10/GH48</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC5124</td>
<td>CBM10/GH10</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC3418</td>
<td>CBM10</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC1304</td>
<td>CBM10/GH18</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC1201</td>
<td>CBM6/GH43</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC172</td>
<td>CE4</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC2171</td>
<td>GH67</td>
<td>1</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC11161</td>
<td>CBM13/CBM10</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC3782</td>
<td>CBM10/GH5</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuTC2294_seq 2</td>
<td>CBM10/CBM6</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AmuVC2158</td>
<td>CBM10</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORF</td>
<td>CAZy Domains</td>
<td>GCS</td>
<td>Avicel</td>
<td>Oat</td>
<td>Spelt</td>
<td>Xylan</td>
<td>Barley</td>
<td>Alfalfa</td>
<td>Hay</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------</td>
<td>-----</td>
<td>--------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>-----</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>AmuTC832</td>
<td>CBM13/CBM10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuTC1644</td>
<td>CE4</td>
<td></td>
<td>1</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC1308</td>
<td>CE4</td>
<td></td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC34</td>
<td>Swollenin/CBM10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC8468</td>
<td>CBM10/GH9</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC11847</td>
<td>CBM10/GH8</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC852</td>
<td>CBM10/CBM1</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC10350</td>
<td>CE1</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC12163</td>
<td>CBM10/GH5</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuTC1512</td>
<td>GH3 _/GH3</td>
<td>1</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuTC1541Seq2</td>
<td>PL1/CBM1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>AmuVC10033</td>
<td>CBM10</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuVC11850</td>
<td>GH1</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuVC2325</td>
<td>CBM10/GH6</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuVC8984</td>
<td>CBM10/CBM1</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuTC4153</td>
<td>GH18</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuVC1073</td>
<td>CBM18</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuTC3711</td>
<td>GH5</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuVC537</td>
<td>CBM10</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>AmuTC290</td>
<td>CBM10</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuTC910</td>
<td>GH117/CBM10/GH43</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC11071</td>
<td>CBM10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC2112</td>
<td>CBM10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC408</td>
<td>CE1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC6446</td>
<td>CBM10/CBM1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC923</td>
<td>CBM18/CBM13</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC10050</td>
<td>CBM13/CBM10/GH10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC10153</td>
<td>CBM10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC10637</td>
<td>GH18</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC1718</td>
<td>CBM10/GH10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC10672</td>
<td>CE1/CBM10</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuVC2278</td>
<td>CBM10/CBM1</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AmuTC2525</td>
<td>GH13/CBM26</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td><strong>Total peptide matched</strong></td>
<td>385</td>
<td>306</td>
<td>192</td>
<td>59</td>
<td>115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1057</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2 ORFs containing CBM10-like domain in association with predicted non-CAZy functional domains.

<table>
<thead>
<tr>
<th>Domain</th>
<th>ORF</th>
</tr>
</thead>
<tbody>
<tr>
<td>CotH, Spore coat assembly protein</td>
<td>AmanuTC10428</td>
</tr>
<tr>
<td></td>
<td>AmanuTC20741</td>
</tr>
<tr>
<td></td>
<td>AmanuTC2143</td>
</tr>
<tr>
<td></td>
<td>AmanuTC4860</td>
</tr>
<tr>
<td></td>
<td>AmanuTC87421</td>
</tr>
<tr>
<td></td>
<td>AmanuVC10133</td>
</tr>
<tr>
<td></td>
<td>AmanuVC10313</td>
</tr>
<tr>
<td></td>
<td>AmanuVC1910</td>
</tr>
<tr>
<td></td>
<td>AmanuVC3496</td>
</tr>
<tr>
<td></td>
<td>AmanuVC2112</td>
</tr>
<tr>
<td></td>
<td>AmanuVN2600</td>
</tr>
<tr>
<td></td>
<td>AmanuVN13467</td>
</tr>
<tr>
<td></td>
<td>AmanuVC2278</td>
</tr>
<tr>
<td>Serpin</td>
<td>AmanuTC169</td>
</tr>
<tr>
<td></td>
<td>AmanuVC11071</td>
</tr>
<tr>
<td></td>
<td>AmanuVC12225</td>
</tr>
<tr>
<td></td>
<td>AmanuVC1590</td>
</tr>
<tr>
<td></td>
<td>AmanuVC3541</td>
</tr>
<tr>
<td></td>
<td>AmanuVC3669</td>
</tr>
<tr>
<td></td>
<td>AmanuVC523</td>
</tr>
<tr>
<td>Ser/Thr protein phosphatase</td>
<td>AmanuVC808</td>
</tr>
<tr>
<td></td>
<td>AmanuVC10249</td>
</tr>
<tr>
<td></td>
<td>AmanuVC10681</td>
</tr>
<tr>
<td></td>
<td>AmanuVC3357</td>
</tr>
<tr>
<td></td>
<td>AmanuVC6446</td>
</tr>
<tr>
<td></td>
<td>AmanuVC8984</td>
</tr>
<tr>
<td>Swollenin</td>
<td>AmanuTC4</td>
</tr>
<tr>
<td></td>
<td>AmanuVC1464</td>
</tr>
<tr>
<td></td>
<td>AmanuVC309</td>
</tr>
<tr>
<td></td>
<td>AmanuVC34</td>
</tr>
<tr>
<td></td>
<td>AmanuVC9956</td>
</tr>
</tbody>
</table>
Figure 4.1 The total gas volume generated by *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.2 The extracellular crude Avicel degrading enzyme activities of *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.3 The extracellular crude carboxymethylcellulose (CMC) degrading enzyme activities of *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.4 The extracellular crude lichenan degrading enzyme activities of *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.5 The extracellular crude starch degrading enzyme activities of *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.6 The extracellular crude oat spelt xylan degrading enzyme activities of *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.7 The extracellular crude p-nitrophenyl-β-d-cellobioside (pNPC) degrading enzyme activities of *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.8 The extracellular crude $p$-nitrophenyl-β-d-glucoside (pNPG) degrading enzyme activities of *Aneromyces mucronatus* YE505 grown on different carbon sources.
Figure 4.9 The extracellular crude carbohydrate esterase activities toward substrate α-Naphthyl acetate (α-NA) of Aneromyces mucronatus YE505 grown on different carbon sources.
Figure 4.10 Size distribution of ORFs that were identified from *Aneromyces mucronatus* YE505.
Figure 4.11 ORFs with predicted KOG/COG functions and their total expression in four different carbon sources.

The assigned letters are based on KOG/COG classifications (Tatusov et al., 2003).

(I): ORFs with predicted KOG functions;

(II): ORFs with predicted COG (but without KOG) functions.
Figure 4.12 Comparison of expression level of predicted CAZymes from *Aneromyces mucronatus* YE505 grown on four different carbon sources.

The CAZymes were grouped into seven clades (A-G) according to their expression patterns and color illustrated based on log$_2$FPKM value. Refer to Table S.3 for complete list of the predicted CAZymes and their expression.

169
Chapter 5 Conclusions and future directions

5.1 Accomplishments

This dissertation established an improved RNA isolation procedure to extract excellent quality RNA from particle associated rumen contents, described the current advances of the eukaryotic metatranscriptome of the muskoxen rumen microbiome, as well as the comparative transcriptomic and secretomic analyses of the rumen anaerobic fungus, *Anaeromyces mucronatus* YE505 grown on different substrates.

(Meta)transcriptomics has distinct properties when compared to (meta)genomics. The (Meta)genome of an organism or ecosystem is relatively stable; in contrast (meta)transcriptomes are dynamic and in a continuous state of change with alterations in environmental conditions. While (meta)genomic sequencing identifies the most numerically dominant genes, metatranscriptomic analysis identifies those genes that are most extensively transcribed, and provides more direct and rational evidence for selecting active gene candidates for future studies. On the other hand, genome sequencing can provide useful information on the structure of gene clusters and possible regulatory mechanisms of gene expression, information that is difficult to obtain from mRNA sequencing.

Prior to my research project, several studies focused on the metagenomics of the plant fiber digesting gut microbiota from termite (Warnecke et al., 2007), wallaby (Pope et al., 2010) and the bovine rumen (Brulc et al., 2009; Hess et al., 2011). However no metatranscriptomic study of the mammalian digestive tract was described in the literature. Compared to the well established DNA isolation from gut samples, RNA isolation is
more challenging because it is dynamically changing and readily degraded by the myriad of RNases that are present in microbial dense environments. In order to remove this obstacle to the study of gene expression in the rumen ecosystem, I improved the RNA isolation method for rumen samples, and in particular developed a procedure that was optimal for rumen solid samples as described in Chapter 2. Subsequently, this method made it possible to isolate the RNA required for the metatranscriptome of the muskoxen rumen and the transcriptome of the anaerobic fungus *Aneromyces mucronatus*.

In Chapter 3, a metatranscriptomic analysis focusing specifically on feed particle-associated rumen eukaryotic microorganisms was carried out. As particle-associated microorganisms represent the major proportion of total rumen microbes (McAllister et al., 1994), and account for up to 90% of the endoglucanase and xylanase activities in the rumen (Miron et al., 2001), I selected to study the solid phase microbial community as it was likely to yield the most information about the function of rumen microbes. Although rumen cellulolytic bacteria are generally believed to play a major role in ruminal plant cell wall biomass degradation, anaerobic fungi are thought to play a significant role in the degradation of low quality forages in part due to their ability to physically disrupt plant particles through mycelia growth (Nagpal et al., 2009). Furthermore, it was more practical to target expressed eukaryotic genes, since the nature of the polyadenylated mature eukaryotic mRNAs enables effective mRNA enrichment by hybridizing to immobilized oligo(dT) for subsequent sequencing, substantially increasing sensitivity through the removal of the most abundant non-coding RNA as well as bacterial mRNA. The resultant Illumina sequencing dataset was analyzed, with a focus on plant cell wall polysaccharide degrading enzymes. The putative genes were found mainly from rumen
eukaryotes especially anaerobic fungi. A total of over 1,000 CAZy proteins were identified in muskoxen rumen samples with the majority from rumen eukaryotes, including anaerobic fungi, protozoa and possibly transient yeast. Compared to the previous gut metagenomics studies which were based solely on DNA sequencing and the genes present (Brulc et al., 2009; Hess et al., 2011; Pope et al., 2010), my study directly elucidated the actual expressed eukaryotic genes of the rumen sample, and was the first report of the rumen eukaryotic metatranscriptome.

In the research described in Chapter 4, transcriptomic sequencing and secretomic analysis were executed upon the anaerobic fungus A. mucronatus YE505, one of the least characterized of the cultured anaerobic fungi. Prior to this study, a number of glycoside hydrolases and carbohydrate esterases have been isolated from rumen fungi. However, due to the limitation of classical methods based on activity, cDNA library screening, or microarray, genes that show low activities on substrates used or low sequence similarity to previously reported genes would have been overlooked. The genomic sequencing of A. mucronatus YE505 was attempted on various next generation sequencing platforms, including Illumina Hi-Seq and PacBio genetic analyzer (unpublished results). However, assembly of the derived sequence into a draft genome proved to be extremely difficult owing to the AT-rich nature of the genome and the relatively short reads and high error rates associated with NGS technologies. Here my research demonstrated that under such limitations, the transcriptomic studies served to be a practical approach to circumvent these obstacles to explore the interesting anaerobic fungi for novel and potential useful genes, by targeting transcribed sequences directly and assembling sequencing reads into full-length ORFs. Should improvements in the accuracy of NGS or the predictive power
of bioinformatic techniques occur in the future, the established transcriptomes may serve as a solid blueprint for facilitating future genomic assembly.

Thus similar to the experimental design described in Chapter 3, RNA-Seq was performed using an Illumina sequencing platform for A. mucronatus YE505. For the first time, a comprehensive insight to the physiological system of A. mucronatus was elucidated. Over 300 putative proteins containing CAZy modules were identified. By comparing transcriptomes from four culture conditions on different carbon sources, the actual gene transcription profiles were obtained, and the potentially important highly expressed enzymes were identified by comparing the FPKM values. A number of putative CAZyme genes were induced when A. mucronatus was grown on alfalfa hay, xylan or barley straw, suggesting their important roles in the degradation of respective substrates. According to the FPKM, members from GH43, GH6, GH1, GH48, GH45 and CE4 were the predominant CAZymes in A. mucronatus. Secretomic study by applying LC-MS/MS and subsequent analysis complemented and provided more proof of the genes detected from the above transcriptomic studies.

When the CAZy proteins detected from the muskoxen rumen and those from A. mucronatus were compared, they shared relatively similar CAZy sets responsible for plant fiber digestion. As expected, a larger CAZy set existed in the rumen sample with over 1,000 members (>500 bp) covering 92 CAZy families, compared to over 300 members in 75 families from A. mucronatus. A total of 67 families were shared by both datasets, including most CAZymes involved in cellulose and xylan degradations. Over 80% of CAZy members from each dataset shared sequence similarities with members from the other dataset. Unsurprisingly, there were several differences in the two characterized
datasets. For example, glycoside hydrolases from GH39 and GH74, which potentially encode xylosidase and endoglucanase activities respectively, were not identified in the *A. mucronatus* YE505 transcriptome, suggesting either these sequences were absent from strain YE505 or not sufficiently expressed to be detected. The presence of these hydrolases in the muskoxen metatranscriptome suggests that these GHs may exist in other fungal species or protozoa.

### 5.2 Future perspectives

Based on the proven effective strategies established in my studies, more experiments are currently underway. Comparative transcriptomic studies of three other rumen fungal species (*Neocallimastix patriciarum* 27, *Orpinomyces joyonii* SG4, *Piromyces rhizinflata* YM600) from the Agriculture and Agri-Food Canada anaerobic culture are currently being performed in a manner similar to that described in Chapter 4. Total RNA was isolated from the fungal strains grown on the same carbon sources as *A. mucronatus* YE505, and RNA-Seq sequencing is currently underway. As the sequence assembly and analysis pipeline have already been established for *A. mucronatus* YE505, we anticipate that this will expedite the analysis of these additional genera. The secreted proteins were collected in a similar way as described in Chapter 4, and will be subject to LC-MS/MS. Our primary study has already shown that transcriptome analysis enables full length cDNA assembly and the secretome can facilitate the identification of regions coding for proteins of interest.

However, bioinformatics alone is not able to verify the actual function of predicted proteins and classical biochemical based studies and protein structural studies will be necessary to assign functions to these predicted proteins. The work in this thesis has
identified several candidate proteins that are worthy of further study. Anaerobic fungi are believed to produce cellulosome-like (see Sections 1.4.3 and 4.4) structures, but to date the foundational scaffolding protein required for these structures has not been identified. It has been proposed that the anaerobic fungal CBM10 domain may act as a dockerin module that facilitates the binding of catalytic proteins to the cellulosome-like complex (Nagy et al., 2007; Raghothama et al., 2001; Steenbakkers et al., 2001). In the present study, many CBM10-like modules were detected, highly transcribed and associated with many sequences coding for protein modules from A. mucronatus, as well as the fiber degrading rumen eukaryotic consortia from muskoxen (see Chapters 3 and 4). On the other hand, all the CBM10 containing enzymes previously identified were involved in plant cell wall degradation, but there were exceptions to this pattern in my study. The CBM10 like module was found to be associated with four non-CAZy domains including swollenin, CotH, Serpin and phosphatase. This raised the possibility that the CBM10 domain may have a function other than those directly related to fiber degradation, and could provide integral information towards defining the function of the rumen fungal cellulosome.

Wild type and truncated mutations from a GH43 gene associated with two tandem CBM10 modules from A. mucronatus YE505 have been introduced into expression vectors. However, no activity has been detected when these recombinant proteins were expressed in Escherichia coli (unpublished data). This may reflect that the E. coli expression system is not suitable to express these genes from anaerobic fungi as a result of different codon usage (unpublished data) or possibly the lack of appropriate post translational modification such as glycosylation. In the future, satisfactory expression
may be obtained using other expression systems such as *Saccharomyces cerevisiae*, *Pichia pastoris* or *Aspergillus niger*. By doing so, studies will be performed to investigate the function of CBM10, for example, whether it can bind certain types of polysaccharides, or whether the existence of CBM10 influences enzyme activities.

Swollenin is another module worthy of further study. This project identified mRNA coding for swollenin both within the muskoxen rumen and from pure cultures of rumen fungus (Chapters 3 and 4). This module was reported to destabilize the cellulose structure with no hydrolytic activity (Brotman et al., 2008). Most swollenin proteins detected in this study were associated with CBM10 modules, again raising interest in the function of this domain.

Rumen fungi are a novel group of microorganisms and their importance in fiber digestion within the rumen community has probably been long underestimated. In recent years, the desire to create a cellulosic biofuel industry has increased the demand for novel lignocellulolysic enzymes. The new trend to use more high fiber diets for agricultural livestock production is likely on the horizon, a development that will make it even more imperative that a better understanding of the process of ruminal fiber digestion be established. My studies described in this dissertation, and a recent study (Wang et al., 2011) have demonstrated the unique CAZyme gene pool harboured and actively utilized by some anaerobic fungi. Enzymatic functional analysis will further target the potential candidates for industrial or agricultural usage. Hypothetical xylanases from GH43, carbohydrate esterases from CE4 and CE15 are good candidates to start with, as there are few studies on these putative enzymes and these enzymes were found to be expressed at a high level when *A. mucronatus* was grown on alfalfa hay and barley straw (Chapter 4).
My thesis and the combined rumen microbiome and genome sequencing information across studies (Brulc et al., 2009; Dai et al., 2012; Hess et al., 2011; Pope et al., 2010; Pope et al., 2012) suggest that the rumen microbiota seems to lack cellulases from GH7 and GH12 families. So far, all members of GH7, a family of exoglucanases, have been isolated from aerobic fungi. Enzymes in family GH12 have been shown to have endoglucanase and xyloglucan hydrolase activities. Although functional aspects of these enzymes may be complemented by other enzymes in the rumen, it would be worth investigating if the addition of the fungal enzymes from GH7 and GH12 families results in significant improvements in rumen fiber digestion.

5.3 Conclusions

Although lots of effort has been made, many aspects of the complex rumen system remain in a black box, with many of the microbial species present and their interactions remaining undefined. My studies provided an overview of gene expression information pertaining to the active eukaryotic lignocellulolytic degradation system existing in rumen fungus *A. mucronatus* and the rumen of muskoxen, and elucidated the potential power of these poorly characterized rumen eukaryotic microorganisms. My dissertation from the perspective of transcriptomic and metatranscriptomic studies shed light on a corner of the black box, demonstrating that sequencing results obtained from high-throughput RNA-Seq and *de novo* assembly were able to provide excellent comprehensive overview of the metabolic activities of the rumen eukaryotic population, as well as rumen fungi in pure culture. By applying various bioinformatic tools, a unique set of hypothetical carbohydrate active enzymes and binding modules were identified. This provided a
powerful source for discovering enzymes that may have significance to both agricultural and biofuel industries.
References


integrity number for assigning integrity values to RNA measurements. BMC Mol. Biol. 7: 3.


## Appendices

### Supplementary Tables

**Table S.1** Muskoxen rumen metatranscriptome contigs (≥500 bp) that have one putative CAZY module

<table>
<thead>
<tr>
<th>Contig number</th>
<th>Domains</th>
<th>Length (bp)</th>
<th>Number of Reads</th>
<th>Accession</th>
<th>GI</th>
<th>E-value</th>
<th>HSP Length</th>
<th>HSP Id%</th>
<th>Hit Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig26982</td>
<td>GH1</td>
<td>1879</td>
<td>11902</td>
<td>AAD45834</td>
<td>gi</td>
<td>5639612</td>
<td>gb</td>
<td>AAD45834.1</td>
<td>AF016864_1</td>
</tr>
<tr>
<td>Contig27636</td>
<td>GH1</td>
<td>734</td>
<td>1195</td>
<td>AAD45834</td>
<td>gi</td>
<td>5639612</td>
<td>gb</td>
<td>AAD45834.1</td>
<td>AF016864_1</td>
</tr>
<tr>
<td>Contig30163</td>
<td>GH1</td>
<td>954</td>
<td>11088</td>
<td>AAD45834</td>
<td>gi</td>
<td>5639612</td>
<td>gb</td>
<td>AAD45834.1</td>
<td>AF016864_1</td>
</tr>
<tr>
<td>Contig22426</td>
<td>GH1</td>
<td>846</td>
<td>2011</td>
<td>AAD45834</td>
<td>gi</td>
<td>5639612</td>
<td>gb</td>
<td>AAD45834.1</td>
<td>AF016864_1</td>
</tr>
<tr>
<td>Contig22325</td>
<td>GH1</td>
<td>895</td>
<td>2868</td>
<td>AAD45834</td>
<td>gi</td>
<td>5639612</td>
<td>gb</td>
<td>AAD45834.1</td>
<td>AF016864_1</td>
</tr>
<tr>
<td>Contig3834</td>
<td>GH1</td>
<td>691</td>
<td>2289</td>
<td>AAP30745</td>
<td>gi</td>
<td>30315031</td>
<td>gb</td>
<td>AAP30745.1</td>
<td></td>
</tr>
<tr>
<td>Contig29365</td>
<td>GH1</td>
<td>683</td>
<td>3236</td>
<td>AAD45834</td>
<td>gi</td>
<td>5639612</td>
<td>gb</td>
<td>AAD45834.1</td>
<td>AF016864_1</td>
</tr>
<tr>
<td>Contig12311</td>
<td>GH1</td>
<td>1369</td>
<td>11613</td>
<td>CAC34952</td>
<td>gi</td>
<td>3445202</td>
<td>gb</td>
<td>CAC34952.1</td>
<td></td>
</tr>
<tr>
<td>Contig21506</td>
<td>GH1</td>
<td>821</td>
<td>14352</td>
<td>AAP30745</td>
<td>gi</td>
<td>30315031</td>
<td>gb</td>
<td>AAP30745.1</td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>Function</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----</td>
<td>-------</td>
<td>-----</td>
<td>---------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29533</td>
<td>GH1</td>
<td>819</td>
<td>14351</td>
<td>beta-glucosidase</td>
<td>[Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29741</td>
<td>GH1</td>
<td>586</td>
<td>13272</td>
<td>beta-glucosidase</td>
<td>Cel1C [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29793</td>
<td>GH1</td>
<td>574</td>
<td>2659</td>
<td>beta-glucosidase</td>
<td>[Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig26250</td>
<td>GH1</td>
<td>982</td>
<td>2876</td>
<td>beta-glucosidase</td>
<td>Cel1C [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30783</td>
<td>GH1</td>
<td>699</td>
<td>5751</td>
<td>beta-glucosidase</td>
<td>[Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21061</td>
<td>GH1</td>
<td>605</td>
<td>3796</td>
<td>beta-glucosidase</td>
<td>Cel1C [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig8436</td>
<td>GH1</td>
<td>541</td>
<td>193</td>
<td>beta-glucosidase</td>
<td>[Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig20700</td>
<td>GH2C</td>
<td>776</td>
<td>254</td>
<td>beta-galactosidase</td>
<td>[Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2681</td>
<td>GH2C</td>
<td>730</td>
<td>248</td>
<td>beta-galactosidase</td>
<td>[Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5538</td>
<td>GH2C</td>
<td>663</td>
<td>168</td>
<td>beta-galactosidase</td>
<td>[Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig290</td>
<td>GH2C</td>
<td>1528</td>
<td>89</td>
<td>glycoside hydrolase</td>
<td>family protein [Clostridium phytofermentans ISDg]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1357</td>
<td>GH2C</td>
<td>1203</td>
<td>347</td>
<td>Beta-galactosidase/beta-glucuronidase</td>
<td>[Eubacterium siraeum V10Sc8a]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig8041</td>
<td>GH2C</td>
<td>567</td>
<td>216</td>
<td>glycosyl hydrolase, family 2</td>
<td>[Prevotella ruminicola 23]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH Type</td>
<td>Start</td>
<td>End</td>
<td>Ref</td>
<td>Description</td>
<td>Accession/ID</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-------</td>
<td>------</td>
<td>-------</td>
<td>----------------------------------------------------------------------------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig7865</td>
<td>GH2C</td>
<td>633</td>
<td>150</td>
<td>YP_001560254</td>
<td>glycoside hydrolase family protein [Clostridium phytfermentans ISDg]</td>
<td>gi</td>
<td>160881286</td>
<td>ref</td>
<td>YP_001560254.1</td>
</tr>
<tr>
<td>Contig801</td>
<td>GH2C</td>
<td>844</td>
<td>496</td>
<td>ZP_03477286</td>
<td>glycoside hydrolase family 2 sugar binding [Clostridium phytfermentans ISDg]</td>
<td>gi</td>
<td>218263041</td>
<td>ref</td>
<td>ZP_03477286.1</td>
</tr>
<tr>
<td>Contig28997</td>
<td>GH2C</td>
<td>1736</td>
<td>3250</td>
<td>ZP_03477286</td>
<td>hypothetical protein PRABACTJOHN_02967 [Parabacteroides johnsonii DSM 18315]</td>
<td>gi</td>
<td>160881286</td>
<td>ref</td>
<td>YP_001560254.1</td>
</tr>
<tr>
<td>Contig671</td>
<td>GH2C</td>
<td>2005</td>
<td>37</td>
<td>ZP_06617065</td>
<td>hypothetical protein PRABACTJOHN_02967 [Parabacteroides johnsonii DSM 18315]</td>
<td>gi</td>
<td>218263041</td>
<td>ref</td>
<td>ZP_03477286.1</td>
</tr>
<tr>
<td>Contig21714</td>
<td>GH3</td>
<td>2343</td>
<td>1925</td>
<td>AAO41704</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
</tr>
<tr>
<td>Contig9134</td>
<td>GH3</td>
<td>1201</td>
<td>746</td>
<td>AAO41704</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
</tr>
<tr>
<td>Contig23162</td>
<td>GH3</td>
<td>918</td>
<td>678</td>
<td>AAO41704</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
</tr>
<tr>
<td>Contig4265</td>
<td>GH3</td>
<td>2203</td>
<td>1437</td>
<td>AAO41704</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
</tr>
<tr>
<td>Contig5433</td>
<td>GH3</td>
<td>2446</td>
<td>4116</td>
<td>AAO41704</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
</tr>
<tr>
<td>Contig24191</td>
<td>GH3</td>
<td>957</td>
<td>857</td>
<td>AAO41704</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
</tr>
<tr>
<td>Contig29507</td>
<td>GH3</td>
<td>3032</td>
<td>4266</td>
<td>ZP_04744384</td>
<td>beta-glucosidase A [Roseburia intestinalis L1-82]</td>
<td>gi</td>
<td>240145783</td>
<td>ref</td>
<td>ZP_047443</td>
</tr>
<tr>
<td>Contig</td>
<td>Type</td>
<td>Length</td>
<td>Sequence</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>--------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig463</td>
<td>GH3</td>
<td>384</td>
<td>gi</td>
<td>gi291538316</td>
<td>glucosidase A [Roseburia intestinalis L1-82]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig14426</td>
<td>GH3</td>
<td>360</td>
<td>gi</td>
<td>gi291538316</td>
<td>Beta-glucosidase-related glycosidases [Roseburia intestinalis XB6B4]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3004</td>
<td>GH3</td>
<td>2108</td>
<td>gi</td>
<td>gi291538316</td>
<td>Beta-glucosidase-related glycosidases [Roseburia intestinalis XB6B4]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1098</td>
<td>GH3</td>
<td>567</td>
<td>gi</td>
<td>gi280977789</td>
<td>glucosidase [Cellulosilyticum ruminicola]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig23195</td>
<td>GH3</td>
<td>1055</td>
<td>gi</td>
<td>gi291538316</td>
<td>Beta-glucosidase-related glycosidases [Roseburia intestinalis L1-82]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3679</td>
<td>GH3</td>
<td>2061</td>
<td>gi</td>
<td>gi291538316</td>
<td>Beta-glucosidase-related glycosidases [Roseburia intestinalis L1-82]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig13027</td>
<td>GH3</td>
<td>1034</td>
<td>gi</td>
<td>gi28557461</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig25386</td>
<td>GH3</td>
<td>572</td>
<td>gi</td>
<td>gi291538316</td>
<td>Beta-glucosidase-related glycosidases [Roseburia intestinalis XB6B4]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6513</td>
<td>GH3</td>
<td>1140</td>
<td>gi</td>
<td>gi293402355</td>
<td>beta-N-acetylhexosaminidase [Erysipelotrichaceae bacterium _5_2_54FAA]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21869</td>
<td>GH3</td>
<td>1046</td>
<td>gi</td>
<td>gi291538316</td>
<td>beta-glucosidase-related glycosidases [Roseburia intestinalis L1-82]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5802</td>
<td>GH3</td>
<td>524</td>
<td>gi</td>
<td>gi28557461</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22281</td>
<td>GH3</td>
<td>860</td>
<td>gi</td>
<td>gi28557461</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21835</td>
<td>Cellulase</td>
<td>2428</td>
<td>gi</td>
<td>gi51090374</td>
<td>cellulase celA [Polyplastron multivesiculatum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Cellulase</td>
<td>Start</td>
<td>End</td>
<td>GI</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Contig21651 | Cellulase | 2258  | 4303  | BAC57896 | gi|51090374|d
| | | | | | bj|BAC57896.2| 0 | 639  | 69.2 | cellulase celA [Polyplastron multivesiculatum] |
| Contig29732 | Cellulase | 1569  | 1435  | AAC06321 | gi|2981484|gb| AAC06321.1| 0 | 513  | 76.0 | cellulase CelID [Neocallimastix patriciarum] |
| Contig792 | Cellulase | 1400  | 145   | AAC06321 | gi|2981484|gb| AAC06321.1| 0 | 472  | 76.5 | cellulase CelID [Neocallimastix patriciarum] |
| Contig21888 | *Cellulase | 1388  | 2071  | ZP_06720041 | gi|294642164|ref|ZP_06720041.1| 1.00E-148 | 441  | 58.5 | cellulase (glycosyl hydrolase family 5) [Ruminococcus albus 8]|
| Contig4250 | Cellulase | 1297  | 2251  | CAL91968 | gi|218081332|emb|CAL91968.1| 0 | 480  | 67.9 | cellulase [Epidinium caudatum] |
| Contig22088 | *Cellulase | 1229  | 2727  | CAL91974 | gi|218081346|emb|CAL91974.1| 1.00E-166 | 354  | 75.4 | cellulase [Epidinium caudatum] |
| Contig28894 | Cellulase | 1256  | 1373  | CAL91974 | gi|218081346|emb|CAL91974.1| 1.00E-136 | 356  | 64.6 | cellulase [Epidinium caudatum] |
| Contig98 | Cellulase | 2087  | 64    | BAC57896 | gi|51090374|d
| | | | | | bj|BAC57896.2| 1.00E-142 | 397  | 59.4 | cellulase celA [Polyplastron multivesiculatum] |
| Contig1999 | Cellulase | 1192  | 63    | BAC57893 | gi|28569970|d
| | | | | | bj|BAC57893.1| 1.00E-150 | 323  | 69.0 | endoglucanase epi2 [Epidinium caudatum] |
| Contig4398 | Cellulase | 1097  | 353   | CAH69214 | gi|59932919|emb|CAH69214.1| 0 | 364  | 94.0 | cellulase family 5 protein [Epidinium caudatum] |
| Contig28882 | Cellulase | 1250  | 4749  | AAD30364 | gi|4836186|gb| AAD30364.1| 1.00E-141 | 302  | 60.6 | CelB [Caldicellulosiruptor sp. Tok7B.1] |
| Contig28881 | Cellulase | 1118  | 951   | BAC57896 | gi|51090374|d
<p>| | | | | | bj|BAC57896.2| 1.00E-122 | 375  | 56.0 | cellulase celA [Polyplastron multivesiculatum] |
| Contig21436 | Cellulase | 1585  | 4195  | CAH69214 | gi|59932919|emb|CAH69214.1| 0 | 485  | 84.1 | cellulase family 5 protein [Epidinium caudatum] |
| Contig1110 | Cellulase | 1120  | 76    | CBK96759 | gi|291531174| 1.00E-359 | 359  | 58.2 | Endoglucanase [Eubacterium siraeum 70/3] |</p>
<table>
<thead>
<tr>
<th>Contig ID</th>
<th>Description</th>
<th>Start</th>
<th>End</th>
<th>Length</th>
<th>Description</th>
<th>Identity</th>
<th>Accession</th>
<th>Description</th>
<th>E-value</th>
<th>Length</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig29277</td>
<td>Cellulase</td>
<td>1165</td>
<td>2707</td>
<td>1542</td>
<td>Cellulase</td>
<td>1.00E-158</td>
<td>gi</td>
<td>51090374</td>
<td>d</td>
<td>BAC57896.2</td>
<td>104.0</td>
</tr>
<tr>
<td>Contig2730</td>
<td>Cellulase</td>
<td>597</td>
<td>78</td>
<td>519</td>
<td>Cellulase</td>
<td>1.00E-115</td>
<td>gi</td>
<td>2981484</td>
<td>gb</td>
<td>AAC06321.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Contig6521</td>
<td>Cellulase</td>
<td>874</td>
<td>441</td>
<td>433</td>
<td>Cellulase</td>
<td>1.00E-155</td>
<td>gi</td>
<td>4586414</td>
<td>db</td>
<td>BAA76394.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Contig2292</td>
<td>Cellulase</td>
<td>922</td>
<td>231</td>
<td>692</td>
<td>Cellulase</td>
<td>1.00E-104</td>
<td>gi</td>
<td>218081334</td>
<td>emb</td>
<td>CAL91969.1</td>
<td>100.0</td>
</tr>
<tr>
<td>Contig21902</td>
<td>Cellulase</td>
<td>827</td>
<td>301</td>
<td>526</td>
<td>Cellulase</td>
<td>5.00E-89</td>
<td>gi</td>
<td>268612028</td>
<td>ref</td>
<td>ZP_06145755.1</td>
<td>65.8</td>
</tr>
<tr>
<td>Contig26471</td>
<td>Cellulase</td>
<td>660</td>
<td>185</td>
<td>475</td>
<td>Cellulase</td>
<td>1.00E-60</td>
<td>gi</td>
<td>1827681</td>
<td>pd</td>
<td>1ECEA</td>
<td>Chain A, Acidothermus Cellulolyticus Endocellulase E1 Catalytic Domain In Complex With A Cellotetraosegi</td>
</tr>
<tr>
<td>Contig5076</td>
<td>Cellulase</td>
<td>790</td>
<td>334</td>
<td>456</td>
<td>Cellulase</td>
<td>1.00E-77</td>
<td>gi</td>
<td>51090374</td>
<td>d</td>
<td>BAC57896.2</td>
<td>104.0</td>
</tr>
<tr>
<td>Contig22970</td>
<td>Cellulase</td>
<td>800</td>
<td>379</td>
<td>421</td>
<td>Cellulase</td>
<td>2.00E-71</td>
<td>gi</td>
<td>51090374</td>
<td>d</td>
<td>BAC57896.2</td>
<td>104.0</td>
</tr>
<tr>
<td>Contig22908</td>
<td>Cellulase</td>
<td>927</td>
<td>540</td>
<td>387</td>
<td>Cellulase</td>
<td>9.00E-72</td>
<td>gi</td>
<td>268609198</td>
<td>ref</td>
<td>ZP_06142925.1</td>
<td>46.0</td>
</tr>
<tr>
<td>NODE_26010_length_577_cov_5.098787</td>
<td>Cellulase</td>
<td>613</td>
<td>236</td>
<td>377</td>
<td>Cellulase</td>
<td>2.00E-66</td>
<td>gi</td>
<td>268612028</td>
<td>ref</td>
<td>ZP_06145755.1</td>
<td>65.8</td>
</tr>
<tr>
<td>Contig519</td>
<td>Cellulase</td>
<td>2045</td>
<td>127</td>
<td>2118</td>
<td>Cellulase</td>
<td>5.00E-64</td>
<td>gi</td>
<td>294642514</td>
<td>ref</td>
<td>ZP_06720388.1</td>
<td>42.3</td>
</tr>
<tr>
<td>Contig29835</td>
<td>Cellulase</td>
<td>504</td>
<td>582</td>
<td>80</td>
<td>Cellulase</td>
<td>1.00E-54.7</td>
<td>gi</td>
<td>290959511</td>
<td></td>
<td>54.7</td>
<td>148</td>
</tr>
<tr>
<td>Contig</td>
<td>Cellulase</td>
<td>1645</td>
<td>737</td>
<td>ABY52798</td>
<td>gi</td>
<td>164375385</td>
<td></td>
<td>1.00E</td>
<td>349</td>
<td>72.8</td>
<td>1,4-beta-D-glucan-cellubiohydrodrolase</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>NODE_18303</td>
<td>GH6</td>
<td>1445</td>
<td>737</td>
<td>ABY52798</td>
<td>gi</td>
<td>164375385</td>
<td></td>
<td>1.00E</td>
<td>349</td>
<td>72.8</td>
<td>1,4-beta-D-glucan-cellubiohydrodrolase</td>
</tr>
<tr>
<td>Contig21599</td>
<td>Cellulase</td>
<td>537</td>
<td>325</td>
<td>Q12647</td>
<td>gi</td>
<td>2494328</td>
<td>sp</td>
<td>Q12647.1</td>
<td></td>
<td>GU</td>
<td>NB_NEOPA</td>
</tr>
<tr>
<td>Contig27963</td>
<td>Cellulase</td>
<td>511</td>
<td>180</td>
<td>ACU30843</td>
<td>gi</td>
<td>255710036</td>
<td>gb</td>
<td>ACU30843 .1</td>
<td></td>
<td>8.00E</td>
<td>-30</td>
</tr>
</tbody>
</table>
| Contig8299 | Cellulase | 501 | 220 | XP_001430 | gi|145488428|ref|XP_00143 0218.1| | 1.00E | -14 | 166 | 32.5 | hypothetical protein [Paramecium tetraurelia strain d4-2]gi|124397314|emb|CAK62820.1| unnamed protein product [Paramecium tetr
| Contig4151 | Cellulase | 719 | 276 | XP_002679 | gi|290993512|ref|XP_00267 9377.1| | 2.00E | -33 | 280 | 32.1 | predicted protein [Naegleria gruberi]gi|284092993|gb|EFC46633.1| |
| CONTig9951 | Cellulase | 846 | 405 | XP_002679 | gi|290993512|ref|XP_00267 9377.1| | 2.00E | -33 | 280 | 32.1 | predicted protein [Naegleria gruberi]gi|284092993|gb|EFC46633.1| |
| Contig3875 | Cellulase | 518 | 142 | CAL91968 | gi|218081332|emb|CAL919 68.1| | 8.00E | -60 | 170 | 62.4 | cellulase celA [Epidinium ecaudatum] |
| Contig30151 | Cellulase | 563 | 485 | BAC57896 | gi|51090374|db|BAC57896. 2| | 7.00E | -94 | 188 | 79.8 | cellulase celA [Polyplastron multivesiculatum] |
| Contig15879 | Cellulase | 691 | 452 | BAC57895 | gi|28569974|db|BAC57895. 1| | 4.00E | -66 | 213 | 53.1 | cellulase celA [Epidinium caudatum] |
| Contig5107 | Cellulase | 1307 | 463 | ZP_020251 | gi|154482720|ref|ZP_02025 168.1| | 1.00E | -145 | 409 | 58.7 | hypothetical protein EUBVEN_00397 [Eubacterium ventriosum ATCC 27560]gi|149736496|gb|EDM52382.1| |
| Contig15712 | Cellulase | 1864 | 7135 | ZP_020251 | gi|154482720|ref|ZP_02025 168.1| | 1.00E | -175 | 502 | 58.4 | hypothetical protein EUBVEN_00397 [Eubacterium ventriosum ATCC 27560]gi|149736496|gb|EDM52382.1| |

<p>| ref|YP_00349 0693.1| | -41 | 87.22 | putative secreted cellulase [Streptomyces scabiei 87.22] |</p>
<table>
<thead>
<tr>
<th>Contig</th>
<th>GH</th>
<th>Length</th>
<th>Amino Acid</th>
<th>Database Reference</th>
<th>E-value</th>
<th>Identity</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>23192</td>
<td>GH6</td>
<td>1085</td>
<td>ACX32999</td>
<td>gi</td>
<td>260169862</td>
<td>gb</td>
<td>ACX32999</td>
</tr>
<tr>
<td>29702</td>
<td>GH6</td>
<td>676</td>
<td>ABY52799</td>
<td>gi</td>
<td>164375387</td>
<td>gb</td>
<td>ABY52799</td>
</tr>
<tr>
<td>1051</td>
<td>GH6</td>
<td>775</td>
<td>AAP33843</td>
<td>gi</td>
<td>32395719</td>
<td>gb</td>
<td>AAP33843</td>
</tr>
<tr>
<td>346</td>
<td>GH6</td>
<td>863</td>
<td>AAL01211</td>
<td>gi</td>
<td>15529294</td>
<td>gb</td>
<td>AAL01211</td>
</tr>
<tr>
<td>31943</td>
<td>GH8</td>
<td>1253</td>
<td>YP_003248</td>
<td>ref</td>
<td>YP_003248</td>
<td>8565.1</td>
<td>0 379 84.4 glycoside hydrolase family 8 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>3115</td>
<td>GH8</td>
<td>904</td>
<td>YP_003248</td>
<td>ref</td>
<td>YP_003248</td>
<td>8565.1</td>
<td>1.00E-109 214 82.2 glycoside hydrolase family 8 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>925</td>
<td>GH9</td>
<td>1836</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>gb</td>
<td>AAM81966</td>
</tr>
<tr>
<td>1316</td>
<td>GH9</td>
<td>1868</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>gb</td>
<td>AAM81966</td>
</tr>
<tr>
<td>21367</td>
<td>GH9</td>
<td>1881</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>gb</td>
<td>AAM81966</td>
</tr>
<tr>
<td>29001</td>
<td>GH9</td>
<td>1810</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>gb</td>
<td>AAM81966</td>
</tr>
<tr>
<td>Contig</td>
<td>GH9</td>
<td></td>
<td>AAM81967</td>
<td>gi</td>
<td>21929669</td>
<td>g b</td>
<td>AAM81967. l</td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>---</td>
<td>----------</td>
<td>---</td>
<td>--------</td>
<td>---</td>
<td>-------------</td>
</tr>
<tr>
<td>Contig595</td>
<td>GH9</td>
<td>1813</td>
<td>1215</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
</tr>
<tr>
<td>Contig1790′</td>
<td>GH9</td>
<td>1984</td>
<td>4432</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
</tr>
<tr>
<td>Contig30085</td>
<td>GH9</td>
<td>1824</td>
<td>1505</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
</tr>
<tr>
<td>Contig17971</td>
<td>GH9</td>
<td>1900</td>
<td>1159</td>
<td>AAM81967</td>
<td>gi</td>
<td>21929669</td>
<td>g b</td>
</tr>
<tr>
<td>Contig29004′</td>
<td>GH9</td>
<td>1944</td>
<td>2673</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
</tr>
<tr>
<td>Contig28899</td>
<td>GH9</td>
<td>1861</td>
<td>4329</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
</tr>
<tr>
<td>Contig9942</td>
<td>GH9</td>
<td>1538</td>
<td>1128</td>
<td>XP_002603873</td>
<td>gi</td>
<td>260818001</td>
<td>ref</td>
</tr>
<tr>
<td>Contig29228</td>
<td>GH9</td>
<td>1164</td>
<td>1002</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
</tr>
<tr>
<td>Contig1989</td>
<td>GH9</td>
<td>1667</td>
<td>44</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>RefseqID</td>
<td>gi</td>
<td>Refseq</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>--------</td>
<td>------</td>
<td>------------</td>
<td>---</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Contig4681</td>
<td>GH9</td>
<td>1091</td>
<td>600</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig26087</td>
<td>GH9</td>
<td>755</td>
<td>410</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig265</td>
<td>GH9</td>
<td>807</td>
<td>118</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig29128</td>
<td>GH9</td>
<td>560</td>
<td>549</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig13687</td>
<td>GH9</td>
<td>971</td>
<td>473</td>
<td>CAL91976</td>
<td>gi</td>
<td>218081351</td>
<td>cellulase [Epidinium ecaudatum]</td>
</tr>
<tr>
<td>Contig29291</td>
<td>GH9</td>
<td>889</td>
<td>1049</td>
<td>CAL91976</td>
<td>gi</td>
<td>218081351</td>
<td>cellulase [Epidinium ecaudatum]</td>
</tr>
<tr>
<td>Contig4260</td>
<td>GH9</td>
<td>515</td>
<td>153</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig25347</td>
<td>GH9</td>
<td>639</td>
<td>628</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig30906</td>
<td>GH9</td>
<td>923</td>
<td>553</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig31652</td>
<td>GH9</td>
<td>642</td>
<td>209</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig9838</td>
<td>GH9</td>
<td>564</td>
<td>228</td>
<td>EFA81213</td>
<td>gi</td>
<td>21807029</td>
<td>xylanase 270-6 [Polysphondylium pallidum PN500]</td>
</tr>
<tr>
<td>Contig806</td>
<td>GH10</td>
<td>1258</td>
<td>688</td>
<td>AAB30669</td>
<td>gi</td>
<td>560649</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig405</td>
<td>GH10</td>
<td>1102</td>
<td>497</td>
<td>CAL91982</td>
<td>gi</td>
<td>218081365</td>
<td>xylanase [Eudiplodioum maggi]</td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Length</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig306</td>
<td>1000</td>
<td>979</td>
<td>CAL91981</td>
<td>xylanase [Epidinium ecaudatum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21293</td>
<td>1111</td>
<td>3483</td>
<td>CAL91982</td>
<td>xylanase [Eudiplodinium maggii]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig196</td>
<td>1246</td>
<td>782</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig28861</td>
<td>1016</td>
<td>617</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6672</td>
<td>816</td>
<td>277</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig13683</td>
<td>866</td>
<td>454</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6829</td>
<td>825</td>
<td>342</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5473</td>
<td>651</td>
<td>209</td>
<td>CAL91979</td>
<td>Xylanase [Epidinium ecaudatum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5361</td>
<td>807</td>
<td>460</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig20523</td>
<td>745</td>
<td>545</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig9398</td>
<td>547</td>
<td>139</td>
<td>YP_003506085</td>
<td>Xylanase B_BUTFI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig9325</td>
<td>707</td>
<td>380</td>
<td>P26223</td>
<td>RecName: Full=Endo-1,4-beta-xylanase [Meiothermus ruber DSM 1279]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig4858</td>
<td>691</td>
<td>199</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig10190</td>
<td>516</td>
<td>167</td>
<td>AAB30669</td>
<td>Xylanase B; XYLB [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3258</td>
<td>805</td>
<td>4643</td>
<td>CAD56867</td>
<td>xylanase 11D [Polyplastron multivesiculatum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>gi/emb/gb</td>
<td>E-value</td>
<td>P-value</td>
</tr>
<tr>
<td>------------</td>
<td>----</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Contig31636</td>
<td>GH11</td>
<td>2035</td>
<td>3284</td>
<td>YP_003248456</td>
<td>gi</td>
<td>261414773</td>
<td>1.00E-106</td>
</tr>
<tr>
<td>Contig16126</td>
<td>GH11</td>
<td>615</td>
<td>150</td>
<td>CAD56867</td>
<td>gi</td>
<td>38343952</td>
<td>2.00E-87</td>
</tr>
<tr>
<td>Contig3557</td>
<td>GH11</td>
<td>706</td>
<td>358</td>
<td>CAD56867</td>
<td>gi</td>
<td>38343952</td>
<td>2.00E-90</td>
</tr>
<tr>
<td>Contig21286</td>
<td>GH11</td>
<td>782</td>
<td>5803</td>
<td>CAD56867</td>
<td>gi</td>
<td>38343952</td>
<td>2.00E-86</td>
</tr>
<tr>
<td>Contig28863</td>
<td>GH11</td>
<td>1148</td>
<td>44380</td>
<td>AAT99015</td>
<td>gi</td>
<td>51236716</td>
<td>0</td>
</tr>
<tr>
<td>Contig30226</td>
<td>GH11</td>
<td>919</td>
<td>1309</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
<td>1.00E-103</td>
</tr>
<tr>
<td>Contig13018</td>
<td>GH11</td>
<td>871</td>
<td>6499</td>
<td>ACL68347</td>
<td>gi</td>
<td>21996451</td>
<td>1.00E-101</td>
</tr>
<tr>
<td>Contig1400</td>
<td>GH11</td>
<td>814</td>
<td>743</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
<td>1.00E-102</td>
</tr>
<tr>
<td>Contig12222</td>
<td>GH11</td>
<td>1624</td>
<td>1671</td>
<td>AAG18439</td>
<td>gi</td>
<td>10505338</td>
<td>1.00E-104</td>
</tr>
<tr>
<td>Contig26517</td>
<td>GH11</td>
<td>684</td>
<td>396</td>
<td>CAA57820</td>
<td>gi</td>
<td>565626</td>
<td>2.00E-90</td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>Length</td>
<td>Accession</td>
<td>Identity</td>
<td>E-value</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Contig1997</td>
<td>GH11</td>
<td>1089</td>
<td>87</td>
<td>ABW04217</td>
<td>gi</td>
<td>157930095</td>
<td>gb</td>
</tr>
<tr>
<td>Contig23170</td>
<td>GH11</td>
<td>1106</td>
<td>732</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
<td>sp</td>
</tr>
<tr>
<td>Contig5033</td>
<td>GH11</td>
<td>972</td>
<td>662</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
<td>sp</td>
</tr>
<tr>
<td>Contig8774</td>
<td>GH11</td>
<td>802</td>
<td>379</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
<td>sp</td>
</tr>
<tr>
<td>Contig4808</td>
<td>GH11</td>
<td>639</td>
<td>388</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
<td>sp</td>
</tr>
<tr>
<td>Contig29382</td>
<td>GH11</td>
<td>612</td>
<td>715</td>
<td>YP_00324 456</td>
<td>gi</td>
<td>261414773</td>
<td>ref</td>
</tr>
<tr>
<td>Contig23899</td>
<td>Alpha-amylase</td>
<td>1319</td>
<td>1344</td>
<td>XP_00146 315</td>
<td>gi</td>
<td>145553281</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>Protein</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>GI</td>
<td>Ref Accession</td>
<td>Ref GI</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>---------</td>
<td>---------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Contig22533</td>
<td>Alpha-amylase</td>
<td>1325</td>
<td>1567</td>
<td>XP_640516</td>
<td>gi</td>
<td>66812674</td>
<td>ef</td>
</tr>
<tr>
<td>Contig2108</td>
<td>Alpha-amylase</td>
<td>1065</td>
<td>294</td>
<td>XP_002114</td>
<td>911</td>
<td>gi</td>
<td>19601009</td>
</tr>
<tr>
<td>Contig30641</td>
<td>Alpha-amylase</td>
<td>1246</td>
<td>995</td>
<td>XP_001441</td>
<td>176</td>
<td>gi</td>
<td>14551048</td>
</tr>
<tr>
<td>Contig28650</td>
<td>Alpha-amylase</td>
<td>1004</td>
<td>747</td>
<td>XP_001441</td>
<td>508</td>
<td>gi</td>
<td>14551116</td>
</tr>
<tr>
<td>Contig22809</td>
<td>Alpha-amylase</td>
<td>1294</td>
<td>1251</td>
<td>XP_001462</td>
<td>315</td>
<td>gi</td>
<td>14555328</td>
</tr>
<tr>
<td>NODE_11739</td>
<td>Alpha-amylase</td>
<td>1385</td>
<td>4489</td>
<td>XP_001441</td>
<td>508</td>
<td>gi</td>
<td>14551116</td>
</tr>
<tr>
<td>Contig310</td>
<td>Alpha-amylase</td>
<td>1417</td>
<td>259</td>
<td>XP_001462</td>
<td>315</td>
<td>gi</td>
<td>14555328</td>
</tr>
<tr>
<td>Contig6668</td>
<td>Alpha-amylase</td>
<td>1381</td>
<td>3445</td>
<td>XP_001462</td>
<td>315</td>
<td>gi</td>
<td>14555328</td>
</tr>
<tr>
<td>Contig5037</td>
<td>Alpha-amylase</td>
<td>1088</td>
<td>526</td>
<td>XP_001462</td>
<td>315</td>
<td>gi</td>
<td>14555328</td>
</tr>
<tr>
<td>Contig21244</td>
<td>Alpha-amylase</td>
<td>1326</td>
<td>3086</td>
<td>XP_001462</td>
<td>315</td>
<td>gi</td>
<td>14555328</td>
</tr>
<tr>
<td>Contig</td>
<td>Alpha-amylase</td>
<td>Start</td>
<td>End</td>
<td>GI</td>
<td>Ref</td>
<td>E-value</td>
<td>Bit-score</td>
</tr>
<tr>
<td>--------</td>
<td>--------------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>Contig3803</td>
<td>Alpha-amylase</td>
<td>1382</td>
<td>1860</td>
<td>XP_001462</td>
<td>315</td>
<td>1.00E-105</td>
<td>449 43.2</td>
</tr>
<tr>
<td>Contig3478</td>
<td>Alpha-amylase</td>
<td>1799</td>
<td>2970</td>
<td>XP_001462</td>
<td>315</td>
<td>1.00E-101</td>
<td>421 44.7</td>
</tr>
<tr>
<td>Contig29368</td>
<td>Alpha-amylase</td>
<td>746</td>
<td>2874</td>
<td>XP_001441</td>
<td>176</td>
<td>1.00E-69</td>
<td>230 55.2</td>
</tr>
<tr>
<td>Contig22555</td>
<td>Alpha-amylase</td>
<td>889</td>
<td>357</td>
<td>ZP_057584</td>
<td>82</td>
<td>2.00E-59</td>
<td>303 43.2</td>
</tr>
<tr>
<td>Contig448</td>
<td>Alpha-amylase</td>
<td>625</td>
<td>108</td>
<td>XP_001441</td>
<td>176</td>
<td>1.00E-69</td>
<td>212 56.6</td>
</tr>
<tr>
<td>Contig6348</td>
<td>Alpha-amylase</td>
<td>630</td>
<td>441</td>
<td>XP_001460</td>
<td>601</td>
<td>1.00E-66</td>
<td>208 52.9</td>
</tr>
<tr>
<td>Contig30877</td>
<td>Alpha-amylase</td>
<td>705</td>
<td>1199</td>
<td>YP_001193</td>
<td>561</td>
<td>1.00E-38</td>
<td>215 38.1</td>
</tr>
<tr>
<td>Contig6120</td>
<td>Alpha-amylase</td>
<td>1267</td>
<td>814</td>
<td>ACD93218</td>
<td></td>
<td>1.00E-108</td>
<td>386 49.0</td>
</tr>
<tr>
<td>Contig30966</td>
<td>Alpha-amylase</td>
<td>549</td>
<td>469</td>
<td>XP_002114</td>
<td>911</td>
<td>2.00E-56</td>
<td>182 52.2</td>
</tr>
<tr>
<td>Contig17090</td>
<td>Alpha-amylase</td>
<td>588</td>
<td>139</td>
<td>CAI59813</td>
<td></td>
<td>6.00E-59</td>
<td>169 62.7</td>
</tr>
<tr>
<td>Contig</td>
<td>Alpha-amylase</td>
<td>1154</td>
<td>486</td>
<td>ACD93218</td>
<td>gi</td>
<td>288915565</td>
<td>gb</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Contig25421</td>
<td>Alpha-amylase</td>
<td>696</td>
<td>1511</td>
<td>ZP_05758482</td>
<td>gi</td>
<td>260172070</td>
<td>ref[ZP_05758482.1]</td>
</tr>
<tr>
<td>Contig21016</td>
<td>Alpha-amylase</td>
<td>520</td>
<td>132</td>
<td>XP_002114911</td>
<td>gi</td>
<td>196010093</td>
<td>ref[XP_002114911.1]</td>
</tr>
<tr>
<td>Contig4922</td>
<td>Alpha-amylase</td>
<td>1483</td>
<td>734</td>
<td>XP_001584271</td>
<td>gi</td>
<td>154422518</td>
<td>ref[XP_001584271.1]</td>
</tr>
<tr>
<td>Contig24372</td>
<td>Alpha-amylase</td>
<td>1918</td>
<td>3528</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411087</td>
<td>emb[CAL92192.1]</td>
</tr>
<tr>
<td>Contig22216'</td>
<td>Alpha-amylase</td>
<td>1896</td>
<td>8564</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411097</td>
<td>emb[CAL92192.1]</td>
</tr>
<tr>
<td>Contig2327</td>
<td>Alpha-amylase</td>
<td>1466</td>
<td>115</td>
<td>CAL92191</td>
<td>gi</td>
<td>218411085</td>
<td>emb[CAL92191.1]</td>
</tr>
<tr>
<td>Contig29562</td>
<td>Alpha-amylase</td>
<td>1523</td>
<td>2226</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411097</td>
<td>emb[CAL92192.1]</td>
</tr>
<tr>
<td>Contig2471</td>
<td>Alpha-amylase</td>
<td>1857</td>
<td>149</td>
<td>XP_001584271</td>
<td>gi</td>
<td>154422518</td>
<td>ref[XP_001584271.1]</td>
</tr>
<tr>
<td>Contig</td>
<td>Alpha-amyrase</td>
<td>Start</td>
<td>End</td>
<td>GI</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------</td>
<td>-------</td>
<td>-----</td>
<td>----</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19126</td>
<td>Alpha-amyrase</td>
<td>1992</td>
<td>4174</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411097</td>
<td>emb</td>
</tr>
<tr>
<td>7354</td>
<td>Alpha-amyrase</td>
<td>1112</td>
<td>756</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411097</td>
<td>emb</td>
</tr>
<tr>
<td>640</td>
<td>Alpha-amyrase</td>
<td>883</td>
<td>1084</td>
<td>CAL92191</td>
<td>gi</td>
<td>218411095</td>
<td>emb</td>
</tr>
<tr>
<td>17739</td>
<td>Alpha-amyrase</td>
<td>745</td>
<td>311</td>
<td>CAL92191</td>
<td>gi</td>
<td>218411095</td>
<td>emb</td>
</tr>
<tr>
<td>29090</td>
<td>Alpha-amyrase</td>
<td>1979</td>
<td>1857</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411097</td>
<td>emb</td>
</tr>
<tr>
<td>8432</td>
<td>Alpha-amyrase</td>
<td>601</td>
<td>204</td>
<td>XP_001462 315</td>
<td>gi</td>
<td>14553281</td>
<td>ref</td>
</tr>
<tr>
<td>29102</td>
<td>Alpha-amyrase</td>
<td>774</td>
<td>89</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411097</td>
<td>emb</td>
</tr>
<tr>
<td>24480</td>
<td>Alpha-amyrase</td>
<td>653</td>
<td>287</td>
<td>XP_001298 742</td>
<td>gi</td>
<td>23382892</td>
<td>ref</td>
</tr>
<tr>
<td>17242</td>
<td>Alpha-amyrase</td>
<td>1023</td>
<td>2315</td>
<td>CAL92192</td>
<td>gi</td>
<td>218411097</td>
<td>emb</td>
</tr>
<tr>
<td>376</td>
<td>GH16</td>
<td>804</td>
<td>208</td>
<td>YP_003250 162</td>
<td>gi</td>
<td>261416479</td>
<td>ref</td>
</tr>
<tr>
<td>27493</td>
<td>GH16</td>
<td>878</td>
<td>19963</td>
<td>YP_003250 162</td>
<td>gi</td>
<td>261416479</td>
<td>ref</td>
</tr>
<tr>
<td>contig</td>
<td>GH</td>
<td>accession</td>
<td>gi</td>
<td>length</td>
<td>e-value</td>
<td>identity</td>
<td>classification</td>
</tr>
<tr>
<td>---------</td>
<td>----</td>
<td>-------------</td>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>Contig22042</td>
<td>GH16</td>
<td>YP_003250</td>
<td>gi</td>
<td>261416479</td>
<td>ref</td>
<td>YP_003250 0162.1</td>
<td>7.00E-84</td>
</tr>
<tr>
<td>Contig24068</td>
<td>GH16</td>
<td>YP_003250</td>
<td>gi</td>
<td>261416479</td>
<td>ref</td>
<td>YP_003250 0162.1</td>
<td>3.00E-74</td>
</tr>
<tr>
<td>Contig12909</td>
<td>GH16</td>
<td>YP_003250</td>
<td>gi</td>
<td>261416479</td>
<td>ref</td>
<td>YP_003250 0162.1</td>
<td>6.00E-55</td>
</tr>
<tr>
<td>Contig16779</td>
<td>GH16</td>
<td>YP_003250</td>
<td>gi</td>
<td>261416479</td>
<td>ref</td>
<td>YP_003250 0162.1</td>
<td>4.00E-50</td>
</tr>
<tr>
<td>Contig4908</td>
<td>GH16</td>
<td>ZP_055122</td>
<td>gi</td>
<td>256773773</td>
<td>ref</td>
<td>ZP_05512 236.1</td>
<td>4.00E-85</td>
</tr>
<tr>
<td>Contig30582</td>
<td>GH18</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>ref</td>
<td>ZP_02027 367.1</td>
<td>1.00E-140</td>
</tr>
<tr>
<td>Contig2533</td>
<td>GH18</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>ref</td>
<td>ZP_02027 367.1</td>
<td>1.00E-158</td>
</tr>
<tr>
<td>Contig385</td>
<td>GH18</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>ref</td>
<td>ZP_02027 367.1</td>
<td>1.00E-134</td>
</tr>
<tr>
<td>Contig21608</td>
<td>GH18</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>ref</td>
<td>ZP_02027 367.1</td>
<td>1.00E-130</td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>Code</td>
<td>Accession</td>
<td>E-value</td>
<td>Bit-score</td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
<td>-------</td>
<td>-----</td>
<td>------</td>
<td>-----------</td>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>Contig21754</td>
<td>GH18</td>
<td>1313</td>
<td>734</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>1.00E-131</td>
</tr>
<tr>
<td>Contig20742</td>
<td>GH18</td>
<td>1217</td>
<td>565</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>1.00E-110</td>
</tr>
<tr>
<td>Contig22956</td>
<td>GH18</td>
<td>1236</td>
<td>722</td>
<td>ZP_042602</td>
<td>gi</td>
<td>229131382</td>
<td>6.00E-44</td>
</tr>
<tr>
<td>Contig5831</td>
<td>GH18</td>
<td>1008</td>
<td>371</td>
<td>CAC35202</td>
<td>gi</td>
<td>13508934</td>
<td>6.00E-43</td>
</tr>
<tr>
<td>Contig7444</td>
<td>GH18</td>
<td>1105</td>
<td>388</td>
<td>ZP_048509</td>
<td>gi</td>
<td>253573651</td>
<td>1.00E-35</td>
</tr>
<tr>
<td>Contig23448</td>
<td>GH18</td>
<td>780</td>
<td>358</td>
<td>YP_001643</td>
<td>gi</td>
<td>163938376</td>
<td>6.00E-39</td>
</tr>
<tr>
<td>Contig5897</td>
<td>GH18</td>
<td>780</td>
<td>173</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>1.00E-67</td>
</tr>
<tr>
<td>Contig13273</td>
<td>GH18</td>
<td>1120</td>
<td>686</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>4.00E-81</td>
</tr>
<tr>
<td>Contig3673</td>
<td>GH18</td>
<td>642</td>
<td>176</td>
<td>YP_001643</td>
<td>gi</td>
<td>163938376</td>
<td>2.00E-30</td>
</tr>
<tr>
<td>Contig ID</td>
<td>Source</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2109</td>
<td>GH18</td>
<td>981</td>
<td>1036</td>
<td>ZP_020273</td>
<td>gi</td>
<td>154484919</td>
<td>ref</td>
</tr>
<tr>
<td>Contig6553</td>
<td>SLT</td>
<td>638</td>
<td>981</td>
<td>AAB61345</td>
<td>gi</td>
<td>2198832</td>
<td>gb</td>
</tr>
<tr>
<td>Contig1850</td>
<td>Phage_ly sozyme</td>
<td>837</td>
<td>1065</td>
<td>XP_001840</td>
<td>847</td>
<td>gi</td>
<td>169868552</td>
</tr>
<tr>
<td>Contig2228</td>
<td>Phage_ly sozyme</td>
<td>965</td>
<td>556</td>
<td>ZP_046561</td>
<td>41</td>
<td>gi</td>
<td>238912304</td>
</tr>
<tr>
<td>Contig2496</td>
<td>GH25</td>
<td>758</td>
<td>926</td>
<td>ZP_064722</td>
<td>67</td>
<td>gi</td>
<td>289640042</td>
</tr>
<tr>
<td>Contig29697</td>
<td>GH25</td>
<td>626</td>
<td>423</td>
<td>ZP_067182</td>
<td>32</td>
<td>gi</td>
<td>294640234</td>
</tr>
<tr>
<td>Contig21944</td>
<td>GH25</td>
<td>844</td>
<td>516</td>
<td>ZP_064722</td>
<td>67</td>
<td>gi</td>
<td>289640042</td>
</tr>
<tr>
<td>Contig22378</td>
<td>GH25</td>
<td>1512</td>
<td>969</td>
<td>ZP_061444</td>
<td>24</td>
<td>gi</td>
<td>268610697</td>
</tr>
<tr>
<td>Contig1638</td>
<td>GH25</td>
<td>628</td>
<td>10</td>
<td>ZP_064722</td>
<td>67</td>
<td>gi</td>
<td>289640042</td>
</tr>
<tr>
<td>Contig1351</td>
<td>GH25</td>
<td>618</td>
<td>170</td>
<td>ZP_064722</td>
<td>67</td>
<td>gi</td>
<td>289640042</td>
</tr>
<tr>
<td>Contig</td>
<td>GH25</td>
<td>Node</td>
<td>GI</td>
<td>Ref</td>
<td>E-Value</td>
<td>Length (bp)</td>
<td>Protein Description</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>---------</td>
<td>-------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>NODE_22897</td>
<td>620</td>
<td>396</td>
<td>ZP_064722</td>
<td>9.00E-45</td>
<td>196</td>
<td>47.4</td>
<td>glycoside hydrolase family 25 [Ethanoligenens harbinense YUAN-3]</td>
</tr>
<tr>
<td>Contig699</td>
<td>1021</td>
<td>172</td>
<td>ZP_020922</td>
<td>4.00E-29</td>
<td>204</td>
<td>38.7</td>
<td>hypothetical protein FAEPRAM212_02501 [Faecalibacterium prausnitzii M21/2]</td>
</tr>
<tr>
<td>Contig5156</td>
<td>627</td>
<td>243</td>
<td>ZP_064722</td>
<td>1.00E-36</td>
<td>189</td>
<td>43.9</td>
<td>glycoside hydrolase family 25 [Ethanoligenens harbinense YUAN-3]</td>
</tr>
<tr>
<td>Contig24106</td>
<td>646</td>
<td>273</td>
<td>ZP_064722</td>
<td>1.00E-47</td>
<td>203</td>
<td>48.8</td>
<td>glycoside hydrolase family 25 [Ethanoligenens harbinense YUAN-3]</td>
</tr>
<tr>
<td>Contig22298</td>
<td>699</td>
<td>545</td>
<td>ZP_064722</td>
<td>1.00E-46</td>
<td>203</td>
<td>47.3</td>
<td>glycoside hydrolase family 25 [Ethanoligenens harbinense YUAN-3]</td>
</tr>
<tr>
<td>Contig3961</td>
<td>1118</td>
<td>447</td>
<td>ZP_034636</td>
<td>4.00E-25</td>
<td>203</td>
<td>39.4</td>
<td>hypothetical protein BACPEC_02724 [Bacteroides pectinophilus ATCC 43243]</td>
</tr>
<tr>
<td>Contig4078</td>
<td>628</td>
<td>291</td>
<td>ZP_061425</td>
<td>2.00E-36</td>
<td>200</td>
<td>41.5</td>
<td>Lysozyme M1 (1,4-beta-N-acetylmuramidase) [Ruminococcus flavefaciens FD-1]</td>
</tr>
<tr>
<td>Contig27175</td>
<td>661</td>
<td>1598</td>
<td>ZP_064722</td>
<td>3.00E-39</td>
<td>191</td>
<td>45.0</td>
<td>glycoside hydrolase family 25 [Ethanoligenens harbinense YUAN-3]</td>
</tr>
<tr>
<td>Contig4426</td>
<td>889</td>
<td>735</td>
<td>XP_001749</td>
<td>3.00E-201</td>
<td>201</td>
<td>43.3</td>
<td>hypothetical protein [Monosiga brevicollis]</td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>Length</td>
<td>Ref</td>
<td>GI</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>----</td>
<td>-------</td>
<td>-----</td>
<td>--------</td>
<td>-----</td>
<td>----</td>
<td>-------------</td>
</tr>
<tr>
<td>Contig12360</td>
<td>GH25</td>
<td>658</td>
<td>469</td>
<td>191</td>
<td>ref[XP_001749 813]</td>
<td>gi</td>
<td>167536282</td>
</tr>
<tr>
<td>Contig31052</td>
<td>GH26</td>
<td>1689</td>
<td>2993</td>
<td>1305</td>
<td>ref[ZP_061440 05]</td>
<td>gi</td>
<td>268610278</td>
</tr>
<tr>
<td>Contig23792</td>
<td>GH26</td>
<td>895</td>
<td>1346</td>
<td>451</td>
<td>ref[YP_001559 233]</td>
<td>gi</td>
<td>160880265</td>
</tr>
<tr>
<td>Contig1939</td>
<td>GH26</td>
<td>1054</td>
<td>44</td>
<td>1010</td>
<td>ref[ADE83298]</td>
<td>gb</td>
<td>ADE83298.1</td>
</tr>
<tr>
<td>Contig29593</td>
<td>GH26</td>
<td>511</td>
<td>4477</td>
<td>3966</td>
<td>ref[ZP_061440 05]</td>
<td>gi</td>
<td>268610278</td>
</tr>
<tr>
<td>Contig1519</td>
<td>Melibias</td>
<td>1066</td>
<td>68</td>
<td>1028</td>
<td>ref[XP_002513 158]</td>
<td>gi</td>
<td>255544191</td>
</tr>
<tr>
<td>Contig31797</td>
<td>Melibias</td>
<td>1026</td>
<td>509</td>
<td>527</td>
<td>ref[XP_001014 570]</td>
<td>gi</td>
<td>118362944</td>
</tr>
<tr>
<td>Contig19135</td>
<td>Melibias</td>
<td>619</td>
<td>274</td>
<td>345</td>
<td>ref[AAA73964]</td>
<td>gi</td>
<td>927577</td>
</tr>
<tr>
<td>Contig4109</td>
<td>Melibias</td>
<td>686</td>
<td>466</td>
<td>220</td>
<td>ref[XP_002325 481]</td>
<td>gi</td>
<td>224144974</td>
</tr>
<tr>
<td>Contig21614</td>
<td>GH30</td>
<td>1217</td>
<td>4832</td>
<td>3616</td>
<td>ref[YP_001039 401]</td>
<td>gi</td>
<td>25975491</td>
</tr>
</tbody>
</table>


<p>| Contig  | GH  | 1167 | 1561 | CBL17903 | gi|291544794|emb|CBL17903.3.1| 1.00E-120 | 366 | 59.6 | O-Glycosyl hydrolase [Ruminococcus sp. 18P13] |
|--------|-----|------|------|----------|--------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Contig  | GH  | 1282 | 2110 | CBL17903 | gi|291544794|emb|CBL17903.3.1| 1.00E-127 | 388 | 58.2 | O-Glycosyl hydrolase [Ruminococcus sp. 18P13] |
| Contig  | GH  | 815  | 422  | ZP_030096 | ref|ZP_030096.674.1| 1.00E-104 | 267 | 64.8 | hypothetical protein BACCOP_01536 [Bacteroides coprocola DSM 17136] |
| Contig  | GH  | 681  | 286  | ZP_038528 | ref|ZP_038528.821.1| 2.00E-51 | 240 | 42.1 | alpha-glycosidase [Chryseobacterium gleum ATCC 35910] |
| Contig  | GH  | 975  | 464  | ZP_030656 | ref|ZP_030656.660.1| 1.00E-121 | 307 | 62.5 | hypothetical protein BACOVA_02646 [Bacteroides ovatus ATCC 8483] |
| Contig  | GH  | 635  | 347  | XP_002676 | ref|XP_002676.411.1| 1.00E-28 | 209 | 37.3 | glycoside hydrolase [Naegleria gruberi] |
| Contig  | GH  | 563  | 185  | ZP_036441 | ref|ZP_036441.168.1| 3.00E-30 | 108 | 58.3 | hypothetical protein BACCOPRO_02544 [Bacteroides coprophilus DSM 18228] |
| Contig  | GH  | 643  | 175  | YP_003180 | ref|YP_003180.274| 1.00E-93 | 214 | 72.4 | LPXTG-motif cell wall anchor domain protein [Atopobium parvulum DSM] |</p>
<table>
<thead>
<tr>
<th>Contig</th>
<th>GH</th>
<th>Start</th>
<th>End</th>
<th>Name</th>
<th>Accession</th>
<th>GO</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig1404 GH32N</td>
<td>610</td>
<td>518</td>
<td>0274.1</td>
<td></td>
<td></td>
<td></td>
<td>LPXTG-motif cell wall anchor domain protein [Atopobium parvulum DSM 20469]</td>
</tr>
<tr>
<td>Contig29808 GH43</td>
<td>972</td>
<td>1181</td>
<td>0274.1</td>
<td></td>
<td></td>
<td></td>
<td>hypothetical protein BACINT_01279 [Bacteroides intestinalis DSM 17393]</td>
</tr>
<tr>
<td>Contig29797 GH43</td>
<td>993</td>
<td>2602</td>
<td>0274.1</td>
<td></td>
<td></td>
<td></td>
<td>hypothetical protein BACINT_01279 [Bacteroides intestinalis DSM 17393]</td>
</tr>
<tr>
<td>Contig4855 GH43</td>
<td>1011</td>
<td>334</td>
<td>0274.1</td>
<td></td>
<td></td>
<td></td>
<td>carbohydrate-binding family 6 protein [Clostridium thermocellum ATCC 27405]</td>
</tr>
<tr>
<td>Contig2691 GH43</td>
<td>1633</td>
<td>677</td>
<td>0274.1</td>
<td></td>
<td></td>
<td></td>
<td>glycoside hydrolase family 43 [Clostridium thermocellum JW20]</td>
</tr>
<tr>
<td>Contig10710 GH43</td>
<td>687</td>
<td>506</td>
<td>0274.1</td>
<td></td>
<td></td>
<td></td>
<td>carbohydrate-binding family 6 protein [Clostridium thermocellum ATCC 27405]</td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Length</td>
<td>Accession</td>
<td>E-value</td>
<td>Identity</td>
<td>Function Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>--------</td>
<td>-----------------</td>
<td>---------</td>
<td>----------</td>
<td>--------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Contig1312</td>
<td>GH43</td>
<td>565</td>
<td>ZP_02065668</td>
<td>1.00E-64</td>
<td>194</td>
<td>hypothetical protein BACOVA_02654 [Bacteroides ovatus ATCC 8483]</td>
<td></td>
</tr>
<tr>
<td>Contig31679</td>
<td>GH43</td>
<td>1635</td>
<td>ACZ98594</td>
<td>1.00E-122</td>
<td>493</td>
<td>hypothetically protein BACOVA_02654 [Bacteroides ovatus ATCC 8483]</td>
<td></td>
</tr>
<tr>
<td>Contig8161</td>
<td>GH43</td>
<td>1227</td>
<td>CBL17231</td>
<td>1.00E-176</td>
<td>396</td>
<td>Beta-xyllosidase [Ruminococcus sp. 18P13]</td>
<td></td>
</tr>
<tr>
<td>Contig27001</td>
<td>GH43</td>
<td>1127</td>
<td>YP_001981807</td>
<td>1.00E-115</td>
<td>311</td>
<td>beta-xyllosidase/alpha-L-arabinofuranosidase, putative, gly43F [Cellvibrio japonicus Ueda107]</td>
<td></td>
</tr>
<tr>
<td>Contig4795</td>
<td>GH43</td>
<td>600</td>
<td>ADE82665</td>
<td>5.00E-74</td>
<td>202</td>
<td>glycosyl hydrolase, family 43 [Prevotella ruminicola 23]</td>
<td></td>
</tr>
<tr>
<td>Contig22716</td>
<td>GH43</td>
<td>686</td>
<td>ACZ98594</td>
<td>7.00E-85</td>
<td>223</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
<td></td>
</tr>
<tr>
<td>Contig29325</td>
<td>GH43</td>
<td>1409</td>
<td>YP_001981807</td>
<td>1.00E-113</td>
<td>315</td>
<td>beta-xyllosidase/alpha-L-arabinofuranosidase, putative, gly43F [Cellvibrio japonicus Ueda107]</td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>ID</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21412</td>
<td>GH43</td>
<td>760</td>
<td>389</td>
<td>CBK75020</td>
<td>Cellobiohydrolase A (1,4-beta-cellobiosidase A) [Butyrivibrio fibrisolvens 16/4]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22944</td>
<td>GH43</td>
<td>713</td>
<td>323</td>
<td>YP_001981807</td>
<td>beta-xylosidase/alpha-L-arabinofuranosidase, putative, gly43F [Cellvibrio japonicus Ueda107]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30609</td>
<td>GH43</td>
<td>574</td>
<td>3277</td>
<td>YP_001981807</td>
<td>beta-xylosidase/alpha-L-arabinofuranosidase, putative, gly43F [Cellvibrio japonicus Ueda107]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig24702</td>
<td>GH43</td>
<td>824</td>
<td>352</td>
<td>YP_001981807</td>
<td>beta-xylosidase/alpha-L-arabinofuranosidase, putative, gly43F [Cellvibrio japonicus Ueda107]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21677</td>
<td>GH43</td>
<td>1649</td>
<td>989</td>
<td>ADE83588</td>
<td>glycosyl hydrolase, family 43 [Prevotella ruminicola 23]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig13452</td>
<td>GH43</td>
<td>668</td>
<td>495</td>
<td>CBK75020</td>
<td>Cellobiohydrolase A (1,4-beta-cellobiosidase A) [Butyrivibrio fibrisolvens 16/4]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig988</td>
<td>GH43</td>
<td>1584</td>
<td>417</td>
<td>ADE83588</td>
<td>glycosyl hydrolase, family 43 [Prevotella ruminicola 23]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29906</td>
<td>GH43</td>
<td>1383</td>
<td>877</td>
<td>ADE83588</td>
<td>glycosyl hydrolase, family 43 [Prevotella ruminicola 23]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig13831</td>
<td>GH43</td>
<td>1430</td>
<td>690</td>
<td>ADE83588</td>
<td>glycosyl hydrolase, family 43 [Prevotella ruminicola 23]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig28788</td>
<td>GH43</td>
<td>791</td>
<td>458</td>
<td>ACZ98594</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig13688</td>
<td>GH43</td>
<td>763</td>
<td>199</td>
<td>ACZ98594</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>GI</td>
<td>Reference</td>
<td>Feature</td>
<td>Similarity</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>Contig202</td>
<td>GH43</td>
<td>980</td>
<td>43</td>
<td>CBK73885</td>
<td>gi</td>
<td>291518664</td>
<td>-80</td>
</tr>
<tr>
<td>Contig23129</td>
<td>GH43</td>
<td>572</td>
<td>603</td>
<td>CBL17682</td>
<td>gi</td>
<td>291544573</td>
<td>1.00E-126</td>
</tr>
<tr>
<td>Contig3262</td>
<td>GH43</td>
<td>1176</td>
<td>631</td>
<td>ZP_059218 16</td>
<td>gi</td>
<td>261207127</td>
<td>9.00E-92</td>
</tr>
<tr>
<td>Contig869</td>
<td>GH43</td>
<td>1222</td>
<td>317</td>
<td>CBK73885</td>
<td>gi</td>
<td>291518664</td>
<td>1.00E-140</td>
</tr>
<tr>
<td>Contig25381</td>
<td>GH43</td>
<td>518</td>
<td>178</td>
<td>ACZ98594</td>
<td>gb</td>
<td>ACZ98594 .1</td>
<td>7.00E-43</td>
</tr>
<tr>
<td>Contig26483</td>
<td>GH43</td>
<td>695</td>
<td>353</td>
<td>CBK73780</td>
<td>gi</td>
<td>291518559</td>
<td>1.00E-105</td>
</tr>
<tr>
<td>Contig11742</td>
<td>GH43</td>
<td>507</td>
<td>159</td>
<td>ZP_061423 38</td>
<td>gi</td>
<td>268608611</td>
<td>7.00E-63</td>
</tr>
<tr>
<td>Contig14307</td>
<td>GH43</td>
<td>640</td>
<td>166</td>
<td>ZP_055916 60</td>
<td>gi</td>
<td>257413461</td>
<td>3.00E-70</td>
</tr>
<tr>
<td>Contig4054</td>
<td>GH43</td>
<td>1207</td>
<td>538</td>
<td>ZP_047431 77</td>
<td>gi</td>
<td>240144576</td>
<td>1.00E-155</td>
</tr>
<tr>
<td>Contig</td>
<td>GH</td>
<td>Length</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6586</td>
<td>GH3</td>
<td>783</td>
<td>CBL17103</td>
<td>1.00E-109 62.9  Glycosyl hydrolases family 43./Dockerin type I repeat. [Ruminococcus sp. 18P13]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1414</td>
<td>GH3</td>
<td>1051</td>
<td>ZP_059218</td>
<td>7.00E-98 52.3  predicted protein [Enterococcus faecium TC6]gi</td>
<td>289162906</td>
<td>ref</td>
<td>ZP_06445701.1</td>
</tr>
<tr>
<td>Contig6617</td>
<td>GH3</td>
<td>1763</td>
<td>YP_003486</td>
<td>6.00E-51 38.4  putative hydrolase [Streptomyces scabiei 87.22]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig28666</td>
<td>GH4</td>
<td>1199</td>
<td>BAC53956</td>
<td>2.00E-74 62.0  endo-beta-1,4-D-glucanase [Rhizopus oryzae]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_15203</td>
<td>GH4</td>
<td>788</td>
<td>BAC53956</td>
<td>7.00E-80 63.9  endo-beta-1,4-D-glucanase [Rhizopus oryzae]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5478</td>
<td>GH4</td>
<td>651</td>
<td>CAB92325</td>
<td>1.00E-76 60.0  endoglucanase 45A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig31867</td>
<td>GH4</td>
<td>561</td>
<td>BAC53956</td>
<td>3.00E-66 58.9  endo-beta-1,4-D-glucanase [Rhizopus oryzae]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29608</td>
<td>GH4</td>
<td>1062</td>
<td>ABU49185</td>
<td>4.00E-55 61.2  endoglucanase [Syncephalastrum racemosum]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_50541</td>
<td>GH4</td>
<td>688</td>
<td>BAC53987</td>
<td>7.00E-42 66.1  endo-glucanase RCE2 [Rhizopus oryzae]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Contig22615  | GH4 | 775    | YP_003249 | 5.00E-22 56.8  Cellulase [Fibrobacter succinogenes subsp. 
<table>
<thead>
<tr>
<th>Contig ID</th>
<th>GH45</th>
<th>Length</th>
<th>YP Accession</th>
<th>Ref Accession</th>
<th>E-value</th>
<th>Score</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig16557</td>
<td>GH45</td>
<td>577</td>
<td>247</td>
<td>BAA98045</td>
<td>1.00E-36</td>
<td>127</td>
<td>family 45 cellulase homologue [Reticulitermes speratus hindgut protist]</td>
</tr>
<tr>
<td>NODE_9170_1</td>
<td>GH45</td>
<td>717</td>
<td>1764</td>
<td>YP_003249</td>
<td>1.00E-76</td>
<td>232</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig28909</td>
<td>GH45</td>
<td>713</td>
<td>1341</td>
<td>YP_003249</td>
<td>1.00E-69</td>
<td>225</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig30798</td>
<td>GH45</td>
<td>523</td>
<td>10896</td>
<td>YP_003249</td>
<td>3.00E-48</td>
<td>151</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig28966</td>
<td>GH45</td>
<td>784</td>
<td>1076</td>
<td>YP_003249</td>
<td>6.00E-66</td>
<td>231</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig29058</td>
<td>GH45</td>
<td>765</td>
<td>2269</td>
<td>YP_003249</td>
<td>3.00E-60</td>
<td>232</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig914</td>
<td>GH45</td>
<td>719</td>
<td>100</td>
<td>YP_003249</td>
<td>2.00E-67</td>
<td>236</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig21206</td>
<td>GH45</td>
<td>925</td>
<td>15508</td>
<td>YP_003249</td>
<td>1.00E-68</td>
<td>248</td>
<td>Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig ID</td>
<td>GH</td>
<td>Start</td>
<td>End</td>
<td>YP</td>
<td>gi/GI/Ref</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Contig21772</td>
<td>GH45</td>
<td>836</td>
<td>607</td>
<td>YP_003249505</td>
<td>gi[261415822]ref[YP_0032499505.1]</td>
<td>6.00E-64</td>
<td>49.8</td>
</tr>
<tr>
<td>Contig24962</td>
<td>GH45</td>
<td>830</td>
<td>8742</td>
<td>YP_003249505</td>
<td>gi[261415822]ref[YP_0032499505.1]</td>
<td>1.00E-62</td>
<td>50.2</td>
</tr>
<tr>
<td>Contig916</td>
<td>GH45</td>
<td>782</td>
<td>48</td>
<td>YP_003249505</td>
<td>gi[261415823]ref[YP_0032499506.1]</td>
<td>2.00E-58</td>
<td>47.1</td>
</tr>
<tr>
<td>Contig28726</td>
<td>GH45</td>
<td>711</td>
<td>1452</td>
<td>YP_003249506</td>
<td>gi[261415823]ref[YP_0032499506.1]</td>
<td>6.00E-48</td>
<td>53.8</td>
</tr>
<tr>
<td>Contig1946</td>
<td>GH45</td>
<td>546</td>
<td>240</td>
<td>YP_003249506</td>
<td>gi[261415823]ref[YP_0032499506.1]</td>
<td>4.00E-37</td>
<td>56.8</td>
</tr>
<tr>
<td>Contig22616</td>
<td>GH45</td>
<td>840</td>
<td>492</td>
<td>YP_003249505</td>
<td>gi[261415822]ref[YP_0032499505.1]</td>
<td>1.00E-44</td>
<td>41.1</td>
</tr>
<tr>
<td>Contig21346</td>
<td>GH45</td>
<td>725</td>
<td>5791</td>
<td>YP_003249506</td>
<td>gi[261415823]ref[YP_0032499506.1]</td>
<td>7.00E-46</td>
<td>51.5</td>
</tr>
<tr>
<td>Contig764</td>
<td>GH45</td>
<td>630</td>
<td>188</td>
<td>YP_003249505</td>
<td>gi[261415822]ref[YP_0032499505.1]</td>
<td>1.00E-54</td>
<td>52.6</td>
</tr>
</tbody>
</table>

Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]gi|261372278|gb|ACX75023.1] Cellulase [Fibrobacter succinogenes subsp. succinogenes S85]
<p>| Contig  | GH48 | 1203 | 11562 | AAN76734 | gi|25990955|gb|AAN76734.1|AF449412_1 | 0 | 372  | 83.1 | cellulase Cel48A precursor [Piromyces sp. E2] |
|--------|------|------|-------|----------|-----------------|-----------------|---|-----|-----|----------------------------------|
| Contig  | GH48 | 1134 | 225   | AAN76735 | gi|25990957|gb|AAN76735.1|AF449413_1 | 0   | 374  | 81.8 | cellulase Cel48A precursor [Piromyces equi] |
| Contig  | GH48 | 898  | 33408 | AAN76735 | gi|25990957|gb|AAN76735.1|AF449413_1 | 1.00E-154 | 302 | 81.1 | cellulase Cel48A precursor [Piromyces equi] |
| Contig  | GH48 | 773  | 418   | AAN76734 | gi|25990955|gb|AAN76734.1|AF449412_1 | 1.00E-120 | 229 | 86.0 | cellulase Cel48A precursor [Piromyces sp. E2] |
| Contig  | GH48 | 671  | 7160  | AAN76734 | gi|25990955|gb|AAN76734.1|AF449412_1 | 2.00E-98  | 214 | 76.2 | cellulase Cel48A precursor [Piromyces sp. E2] |
| Contig  | GH48 | 550  | 4168  | AAN76734 | gi|25990955|gb|AAN76734.1|AF449412_1 | 2.00E-75  | 169 | 78.1 | cellulase Cel48A precursor [Piromyces sp. E2] |
| Contig  | GH48 | 552  | 396   | AAN76735 | gi|25990957|gb|AAN76735.1|AF449413_1 | 2.00E-75  | 190 | 68.9 | cellulase Cel48A precursor [Piromyces equi] |
| Contig  | GH48 | 586  | 3032  | AAN76734 | gi|25990955|gb|AAN76734.1|AF449412_1 | 3.00E-73  | 168 | 76.8 | cellulase Cel48A precursor [Piromyces sp. E2] |
| Contig  | GH48 | 549  | 436   | AAN76734 | gi|25990955|gb|AAN76734.1|AF449412_1 | 2.00E-80  | 182 | 74.2 | cellulase Cel48A precursor [Piromyces sp. E2] |
| Contig  | GH48 | 559  | 7121  | AAN76734 | gi|25990955|gb|AAN76734.1|AF449412_1 | 3.00E-64  | 124 | 81.5 | cellulase Cel48A precursor [Piromyces sp. E2] |
| Contig  | GH67M| 1745 | 2054  | YP_003096| gi|255535865|gb|YP_003096.1|AF449412_1 | 1.00E533 | 533 | 57.0 | Alpha-glucuronidase [Flavobacteriaceae] |</p>
<table>
<thead>
<tr>
<th>Contig</th>
<th>Accession</th>
<th>Length</th>
<th>Identity</th>
<th>Evalue</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig9529</td>
<td>Glucosa minidase</td>
<td>236</td>
<td>ref[YP_00309 6236.1]</td>
<td>-179</td>
<td>bacterium 3519-10[gi</td>
</tr>
<tr>
<td>Contig3078</td>
<td>Glucosa minidase</td>
<td>601</td>
<td>gi</td>
<td>291524997</td>
<td>emb</td>
</tr>
<tr>
<td>Contig21748</td>
<td>GH76</td>
<td>1093</td>
<td>gi</td>
<td>223940201</td>
<td>ref</td>
</tr>
<tr>
<td>Contig1443</td>
<td>GH77</td>
<td>1591</td>
<td>gi</td>
<td>260889168</td>
<td>ref</td>
</tr>
<tr>
<td>Contig30043</td>
<td>GH77</td>
<td>1574</td>
<td>gi</td>
<td>260889168</td>
<td>ref</td>
</tr>
<tr>
<td>Contig21437</td>
<td>GH77</td>
<td>1010</td>
<td>gi</td>
<td>253681277</td>
<td>ref</td>
</tr>
<tr>
<td>Contig21520</td>
<td>GH77</td>
<td>1129</td>
<td>gi</td>
<td>253681277</td>
<td>ref</td>
</tr>
<tr>
<td>Contig23542</td>
<td>GH77</td>
<td>919</td>
<td>gi</td>
<td>168186989</td>
<td>ref</td>
</tr>
<tr>
<td>Contig30017</td>
<td>GH77</td>
<td>978</td>
<td>gi</td>
<td>291561123</td>
<td>emb</td>
</tr>
<tr>
<td>Contig</td>
<td>Accession</td>
<td>Length</td>
<td>Start</td>
<td>End</td>
<td>E-value</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>--------</td>
<td>-------</td>
<td>-----</td>
<td>---------</td>
</tr>
<tr>
<td>Contig22344</td>
<td>GH77</td>
<td>907</td>
<td>620</td>
<td>ZP_048620 75</td>
<td>gi</td>
</tr>
<tr>
<td>Contig9224</td>
<td>GH77</td>
<td>595</td>
<td>334</td>
<td>ZP_037068 25</td>
<td>gi</td>
</tr>
<tr>
<td>Contig2949</td>
<td>NAGidas e</td>
<td>1209</td>
<td>1199</td>
<td>ZP_054037 93</td>
<td>gi</td>
</tr>
<tr>
<td>Contig5988</td>
<td>GH88</td>
<td>620</td>
<td>285</td>
<td>YP_001197 274</td>
<td>gi</td>
</tr>
<tr>
<td>Contig8574</td>
<td>NAGLU</td>
<td>1069</td>
<td>579</td>
<td>XP_001638 539</td>
<td>gi</td>
</tr>
<tr>
<td>Contig2412</td>
<td>NAGLU</td>
<td>788</td>
<td>178</td>
<td>XP_001638 539</td>
<td>gi</td>
</tr>
<tr>
<td>Contig18975</td>
<td>NAGLU</td>
<td>723</td>
<td>569</td>
<td>XP_001638 539</td>
<td>gi</td>
</tr>
<tr>
<td>Contig7502</td>
<td>NAGLU</td>
<td>527</td>
<td>187</td>
<td>XP_002156 234</td>
<td>gi</td>
</tr>
<tr>
<td>Contig7175</td>
<td>NAGLU</td>
<td>534</td>
<td>224</td>
<td>YP_003122 794</td>
<td>gi</td>
</tr>
<tr>
<td>Contig8333</td>
<td>NAGLU</td>
<td>525</td>
<td>228</td>
<td>XP_002156 234</td>
<td>gi</td>
</tr>
<tr>
<td>Contig15910</td>
<td>GH97</td>
<td>1464</td>
<td>671</td>
<td>ADE82389</td>
<td>gi</td>
</tr>
<tr>
<td>Contig</td>
<td>GH97</td>
<td>853</td>
<td>207</td>
<td>ADE82389</td>
<td>ruminicola 23</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>----------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>294473000</td>
<td>gb</td>
<td>ADE82389</td>
<td>1.00E-116</td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>45233787</td>
<td>ref</td>
<td>XP_001400266</td>
<td>3.00E-66</td>
</tr>
<tr>
<td></td>
<td>gb</td>
<td>ADE82389</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>294473000</td>
<td>gb</td>
<td>ADE82389</td>
<td>1.00E-116</td>
</tr>
<tr>
<td></td>
<td>gi</td>
<td>45233787</td>
<td>ref</td>
<td>XP_001400266</td>
<td>3.00E-66</td>
</tr>
<tr>
<td></td>
<td>gb</td>
<td>ADE82389</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The table continues with similar entries.
<table>
<thead>
<tr>
<th>Contig</th>
<th>CBM</th>
<th>Start</th>
<th>End</th>
<th>Accession</th>
<th>Type</th>
<th>E-value</th>
<th>P-value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig1687</td>
<td>CBM4_9</td>
<td>572</td>
<td>41</td>
<td>CAL91978</td>
<td>xylanase [Polyplastron multivesiculatum]</td>
<td>2.00E-76</td>
<td>95.8</td>
<td>gi</td>
</tr>
<tr>
<td>Contig5630</td>
<td>CBM6</td>
<td>586</td>
<td>146</td>
<td>ACZ98594</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
<td>2.00E-40</td>
<td>41.5</td>
<td>gb</td>
</tr>
<tr>
<td>Contig29224</td>
<td>CBM6</td>
<td>830</td>
<td>483</td>
<td>ACZ98594</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
<td>4.00E-47</td>
<td>42.1</td>
<td>gb</td>
</tr>
<tr>
<td>Contig8109</td>
<td>CBM6</td>
<td>590</td>
<td>174</td>
<td>ADE82665</td>
<td>glycosyl hydrolase, family 43 [Prevotella ruminicola 23]</td>
<td>2.00E-40</td>
<td>56.2</td>
<td>gb</td>
</tr>
<tr>
<td>Contig22052</td>
<td>CBM6</td>
<td>824</td>
<td>287</td>
<td>ACZ98594</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
<td>2.00E-56</td>
<td>42.8</td>
<td>gb</td>
</tr>
<tr>
<td>Contig21969</td>
<td>CBM10</td>
<td>659</td>
<td>248</td>
<td>AAL01212</td>
<td>cellbiohydrolase II-like cellulase CellI [Orpinomyces sp. PC-2]</td>
<td>1.00E-25</td>
<td>59.5</td>
<td>gi</td>
</tr>
<tr>
<td>Contig22570</td>
<td>CBM10</td>
<td>554</td>
<td>671</td>
<td>AAL01212</td>
<td>cellbiohydrolase II-like cellulase CellI [Orpinomyces sp. PC-2]</td>
<td>1.00E-47</td>
<td>66.7</td>
<td>gi</td>
</tr>
<tr>
<td>Contig31340</td>
<td>CBM10</td>
<td>571</td>
<td>535</td>
<td>AAL01212</td>
<td>cellbiohydrolase II-like cellulase CellI [Orpinomyces sp. PC-2]</td>
<td>7.00E-48</td>
<td>67.4</td>
<td>gi</td>
</tr>
<tr>
<td>Contig21298</td>
<td>CBM10</td>
<td>522</td>
<td>175</td>
<td>AAK20910</td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
<td>3.00E-32</td>
<td>47.4</td>
<td>gi</td>
</tr>
<tr>
<td>Contig29014</td>
<td>CBM10</td>
<td>1048</td>
<td>1072</td>
<td>AAL01212</td>
<td>cellbiohydrolase II-like cellulase CellI [Orpinomyces sp. PC-2]</td>
<td>2.00E-26</td>
<td>45.5</td>
<td>gi</td>
</tr>
<tr>
<td>Contig6731'</td>
<td>CBM10</td>
<td>557</td>
<td>215</td>
<td>CAB92325</td>
<td>endoglucanase 45A [Piromyces equi]</td>
<td>2.00E-31</td>
<td>44.6</td>
<td>gi</td>
</tr>
<tr>
<td>Contig30127</td>
<td>CBM10</td>
<td>643</td>
<td>401</td>
<td>AAM81966</td>
<td>cellulose Cel9A precursor [Piromyces sp. E2]</td>
<td>7.00E-62</td>
<td>54.2</td>
<td>gi</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM10</td>
<td>Length</td>
<td>Accession 1</td>
<td>Accession 2</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-------</td>
<td>--------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_4_length_806_cov_7.171216</td>
<td>CBM10</td>
<td>842</td>
<td>P55296</td>
<td>gi</td>
<td>1708917</td>
<td>sp</td>
<td>P55296.1</td>
<td>5894</td>
</tr>
<tr>
<td>Contig92</td>
<td>CBM10</td>
<td>605</td>
<td>43</td>
<td>AAR97891</td>
<td>4.00E-28 cellulosomal serpin mannase [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1167</td>
<td>CBM10</td>
<td>844</td>
<td>1037</td>
<td>AAM81966</td>
<td>6.00E-27 cellulase C9A precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_51456_length_483_cov_2997930</td>
<td>CBM10</td>
<td>519</td>
<td>161</td>
<td>AAK20910</td>
<td>1.00E-44 non-catalytic protein 1 [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29345</td>
<td>CBM10</td>
<td>537</td>
<td>370</td>
<td>ACL68347</td>
<td>2.00E-30 endo-1,4-beta-xylanase [Neocallimastix patriciae]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig11735</td>
<td>CBM10</td>
<td>554</td>
<td>146</td>
<td>ACL68347</td>
<td>2.00E-44 endo-1,4-beta-xylanase [Neocallimastix patriciae]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5673</td>
<td>CBM10</td>
<td>2133</td>
<td>1569</td>
<td>ZP_061453.32</td>
<td>0 557 55.7 hypothetical protein RfaF_19133 [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3456</td>
<td>CBM10</td>
<td>1100</td>
<td>394</td>
<td>ZP_061454.04</td>
<td>9.00E-38 cellulose enzyme, dockerin type I [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29963</td>
<td>CBM10</td>
<td>864</td>
<td>805</td>
<td>CAB92325</td>
<td>3.00E-28 endoglucanase 45A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig4815</td>
<td>CBM10</td>
<td>549</td>
<td>259</td>
<td>AAL01214</td>
<td>2.00E-51 endoglucanase CelJ [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig8213</td>
<td>CBM10</td>
<td>608</td>
<td>325</td>
<td>AAK20910</td>
<td>7.00E-29 non-catalytic protein 1 [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>CBM10</td>
<td>635</td>
<td>449</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>e mb</td>
<td>CAB92325</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td>----------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>NODE_4349</td>
<td>CBM10</td>
<td>704</td>
<td>336</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>e mb</td>
<td>CAB92325</td>
</tr>
<tr>
<td>Contig5770</td>
<td>CBM10</td>
<td>3281</td>
<td>1617</td>
<td>ABY52795</td>
<td>gi</td>
<td>164375379</td>
<td>e mb</td>
<td>ABY52795</td>
</tr>
<tr>
<td>Contig15680</td>
<td>CBM10</td>
<td>650</td>
<td>305</td>
<td>AAK20910</td>
<td>gi</td>
<td>13446353</td>
<td>g b</td>
<td>AAK20910</td>
</tr>
<tr>
<td>NODE_26024</td>
<td>CBM10</td>
<td>631</td>
<td>381</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>g b</td>
<td>AAM81966</td>
</tr>
<tr>
<td>Contig2021</td>
<td>CBM10</td>
<td>665</td>
<td>264</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>e mb</td>
<td>CAB92325</td>
</tr>
<tr>
<td>Contig2340</td>
<td>CBM10</td>
<td>542</td>
<td>239</td>
<td>AAL01214</td>
<td>gi</td>
<td>15529300</td>
<td>g b</td>
<td>AAL01214</td>
</tr>
<tr>
<td>Contig22135</td>
<td>CBM10</td>
<td>520</td>
<td>175</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>e mb</td>
<td>CAB92325</td>
</tr>
<tr>
<td>Contig9042</td>
<td>CBM10</td>
<td>679</td>
<td>191</td>
<td>P55296</td>
<td>gi</td>
<td>1708917</td>
<td>sp</td>
<td>P55296</td>
</tr>
<tr>
<td>Contig6361</td>
<td>CBM10</td>
<td>543</td>
<td>219</td>
<td>AAL01214</td>
<td>gi</td>
<td>15529300</td>
<td>g b</td>
<td>AAL01214</td>
</tr>
<tr>
<td>Contig31078</td>
<td>CBM10</td>
<td>1075</td>
<td>551</td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>g b</td>
<td>AAO41704</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM</td>
<td>Length</td>
<td>Start</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26995</td>
<td>CBM10</td>
<td>645</td>
<td>270</td>
<td>endoglucanase 5A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>750</td>
<td>CBM10</td>
<td>1467</td>
<td>121</td>
<td>hypothetical protein RfIAF_19133 [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2191</td>
<td>CBM10</td>
<td>1573</td>
<td>192</td>
<td>endoglucanase 45A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3568</td>
<td>CBM10</td>
<td>621</td>
<td>199</td>
<td>endo-1,4-beta-xylanase [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3684</td>
<td>CBM10</td>
<td>608</td>
<td>263</td>
<td>mannann endo-1,4-bet-mannosidase; mannannase A [Piromyces sp.]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6963</td>
<td>CBM10</td>
<td>525</td>
<td>211</td>
<td>mannannase ManA [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7410</td>
<td>CBM10</td>
<td>802</td>
<td>514</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21510</td>
<td>CBM10</td>
<td>522</td>
<td>1608</td>
<td>cellbiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5819</td>
<td>CBM10</td>
<td>757</td>
<td>565</td>
<td>endoglucanase B [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29862</td>
<td>CBM10</td>
<td>526</td>
<td>903</td>
<td>cellbiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>CBM</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13907</td>
<td>CBM10</td>
<td>1148</td>
<td>930</td>
<td>YP_003250</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31197</td>
<td>CBM10</td>
<td>596</td>
<td>197</td>
<td>Q9Y871</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_4207</td>
<td>CBM10</td>
<td>623</td>
<td>1016</td>
<td>AAP30750</td>
<td>cellbiohydrolase Cel6C [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3862</td>
<td>CBM10</td>
<td>534</td>
<td>6927</td>
<td>AAL01211</td>
<td>cellbiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29489</td>
<td>CBM10</td>
<td>1181</td>
<td>624</td>
<td>CAB92325</td>
<td>endoglucanase 45A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31049</td>
<td>CBM10</td>
<td>730</td>
<td>431</td>
<td>Q9Y871</td>
<td>S.3821548&lt;gi&gt;gb</td>
<td>AAL01211.1</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>23116</td>
<td>CBM10</td>
<td>564</td>
<td>324</td>
<td>AAL92497</td>
<td>exoglucanase Cel6A [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6783</td>
<td>CBM10</td>
<td>590</td>
<td>396</td>
<td>AAL01211</td>
<td>cellbiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22741</td>
<td>CBM10</td>
<td>583</td>
<td>326</td>
<td>AAL92497</td>
<td>exoglucanase Cel6A [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>CBM</td>
<td>pI</td>
<td>w</td>
<td>gi</td>
<td>gi</td>
<td>Ref</td>
<td>Score</td>
<td>E-value</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>----</td>
<td>---</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>Contig894</td>
<td>CBM10</td>
<td>928</td>
<td>98</td>
<td>P55296</td>
<td>gi</td>
<td>1708917</td>
<td>sp</td>
<td>P55296.1</td>
</tr>
<tr>
<td>Contig28640</td>
<td>CBM10</td>
<td>617</td>
<td>1105</td>
<td>AAM94167</td>
<td>gi</td>
<td>33620325</td>
<td>gb</td>
<td>AAM94167.1</td>
</tr>
<tr>
<td>NODE_8335_1length_1156_cov_10.518167</td>
<td>CBM10</td>
<td>1208</td>
<td>3559</td>
<td>Q9Y871</td>
<td>gi</td>
<td>23821548</td>
<td>sp</td>
<td>Q9Y871.1</td>
</tr>
<tr>
<td>Contig2286</td>
<td>CBM10</td>
<td>500</td>
<td>414</td>
<td>AAL01212</td>
<td>gi</td>
<td>15529296</td>
<td>gb</td>
<td>AAL01212.1</td>
</tr>
<tr>
<td>Contig23820</td>
<td>CBM10</td>
<td>644</td>
<td>344</td>
<td>Q9Y871</td>
<td>gi</td>
<td>23821548</td>
<td>sp</td>
<td>Q9Y871.1</td>
</tr>
<tr>
<td>Contig26006</td>
<td>CBM10</td>
<td>1246</td>
<td>793</td>
<td>YP_003250020</td>
<td>gi</td>
<td>261416337</td>
<td>ref</td>
<td>YP_003250020.1</td>
</tr>
<tr>
<td>Contig21644</td>
<td>CBM10</td>
<td>534</td>
<td>12</td>
<td>AAP30750</td>
<td>gi</td>
<td>30315041</td>
<td>gb</td>
<td>AAP30750.1</td>
</tr>
<tr>
<td>Contig17552</td>
<td>CBM10</td>
<td>524</td>
<td>263</td>
<td>AAL92497</td>
<td>gi</td>
<td>29465670</td>
<td>gb</td>
<td>AAL92497.1</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM10</td>
<td>1026</td>
<td>859</td>
<td>AAL01214</td>
<td>gi</td>
<td>15529300</td>
<td>gb</td>
<td>AAL01214.1</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>------</td>
<td>-----</td>
<td>----------</td>
<td>-----------------</td>
<td>--</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Contig28743</td>
<td>CBM10</td>
<td>666</td>
<td>475</td>
<td>AAP30750</td>
<td>gi</td>
<td>30315041</td>
<td>gb</td>
<td>AAP30750.1</td>
</tr>
<tr>
<td>NODE_37676_length_519Conditional_4.001927</td>
<td>CBM10</td>
<td>555</td>
<td>178</td>
<td>AAL01212</td>
<td>gi</td>
<td>15529296</td>
<td>gb</td>
<td>AAL01212.1</td>
</tr>
<tr>
<td>Contig29725</td>
<td>CBM10</td>
<td>1174</td>
<td>2896</td>
<td>YP_003250020</td>
<td>gi</td>
<td>261416337</td>
<td>ref</td>
<td>YP_003250020.1</td>
</tr>
<tr>
<td>Contig29013</td>
<td>CBM10</td>
<td>940</td>
<td>707</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>gb</td>
<td>AAM81966.1</td>
</tr>
<tr>
<td>Contig26433*</td>
<td>CBM10</td>
<td>654</td>
<td>831</td>
<td>AAP30750</td>
<td>gi</td>
<td>30315041</td>
<td>gb</td>
<td>AAP30750.1</td>
</tr>
<tr>
<td>Contig27699</td>
<td>CBM10</td>
<td>663</td>
<td>230</td>
<td>AAB69090</td>
<td>gi</td>
<td>2231243</td>
<td>gb</td>
<td>AAB69090.1</td>
</tr>
<tr>
<td>Contig22604*</td>
<td>CBM10</td>
<td>1261</td>
<td>3053</td>
<td>AAM94167</td>
<td>gi</td>
<td>33620325</td>
<td>gb</td>
<td>AAM94167.1</td>
</tr>
<tr>
<td>Contig2685*</td>
<td>CBM10</td>
<td>1520</td>
<td>1526</td>
<td>ZP_06145404</td>
<td>gi</td>
<td>26861167</td>
<td>ref</td>
<td>ZP_06145404.1</td>
</tr>
<tr>
<td>Contig13044*</td>
<td>CBM10</td>
<td>1277</td>
<td>2431</td>
<td>Q9Y871</td>
<td>gi</td>
<td>23821548</td>
<td>gb</td>
<td>Q9Y871.1</td>
</tr>
<tr>
<td>Contig28799*</td>
<td>CBM10</td>
<td>848</td>
<td>5681</td>
<td>AAM94167</td>
<td>gi</td>
<td>33620325</td>
<td>gb</td>
<td>AAM94167.1</td>
</tr>
</tbody>
</table>
| Contig8961 | CBM10 | 879 | 232 | P55296 | gi|1708917|sp|P55296.1|MA NA_PIRSP | 3.00E-22 | 127 | 40.9 | RecName: Full=Mannan endo-1,4-beta-mannosidase A; AltName: Full=Beta-mannanase A; AltName: Full=1,4-beta-D-mannan mannanohydrolase A; Flags: Precursor|gi|1197368|emb|CAA62968.1|mannan endo-1,4-beta-mannosidase; mannanase A [Piromyces sp.]
| Contig28756 | CBM10 | 720 | 4490 | AAL92497 | gi|29465670|gb|AAL92497.1|3.00E-52 | 179 | 56.4 | exoglucanase Cel6A [Piromyces sp. E2]
| Contig15299 | CBM10 | 1098 | 298 | ZP_06145332 | gi|26861605|ref|ZP_06145332.1|1.00E-67 | 230 | 55.7 | hypothetical protein RflaF_19133 [Ruminococcus flavefaciens FD-1]
| NODE_1804_length_2041 | CBM10 | 2085 | 1159 | AAP30750 | gi|30315041|gb|AAP30750.1|3.00E-24 | 87 | 55.2 | cellobiohydrolase Cel6C [Piromyces sp. E2]
| Contig2507 | CBM10 | 717 | 768 | CAB92325 | gi|8053234|emb|CAB92325.1|7.00E-33 | 214 | 36.4 | endoglucanase 45A [Piromyces equi]
| Contig668 | CBM10 | 656 | 78 | AAM94167 | gi|33620325|gb|AAM94167.1|2.00E-26 | 99 | 55.6 | cellulosomal glycoside hydrolase family 6 exoglucanase Cel6A [Piromyces equi]
| Contig31387 | CBM10 | 530 | 253 | Q12647 | gi|2494328|sp|Q12647.1|GU NB_NEOPA|1.00E-24 | 97 | 53.6 | RecName: Full=Endoglucanase B; AltName: Full=Endo-1,4-beta-glucanase B; AltName: Full=Cellulase B; Flags: Precursor|gi|467687|emb|CAA83238.1|endoglucanase B [Neocallimastix patriciarum] mannanase ManA [Orpinomyces sp. PC-2]
<p>| Contig2880 | CBM10 | 704 | 432 | AAL01213 | gi|15529298|gb|AAL01213.1|AF177206.1|9.00E-30 | 106 | 55.7 |
| Contig30437 | CBM10 | 738 | 460 | AAM94167 | gi|33620325|gb|AAM94167.1|2.00E-25 | 99 | 54.5 |
| Contig24987 | CBM10 | 757 | 377 | AAP30750 | gi|30315041|gb|AAP30750.1|6.00E-22 | 86 | 51.2 |
| Contig29774 | CBM10 | 954 | 780 | AAP30750 | gi|30315041|gb|AAP30750.1|5.00E-25 | 82 | 57.3 |</p>
<table>
<thead>
<tr>
<th>Contig</th>
<th>CBM</th>
<th>Start</th>
<th>End</th>
<th>Evalue</th>
<th>Identity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig1262</td>
<td>CBM10</td>
<td>2048</td>
<td>87</td>
<td>AAP30750</td>
<td>1.00E-24</td>
<td>celllobiohydrolase Cel6C [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig28841</td>
<td>CBM10</td>
<td>540</td>
<td>6915</td>
<td>AAD04194</td>
<td>9.00E-52</td>
<td>xylanase [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig29794</td>
<td>CBM10</td>
<td>568</td>
<td>430</td>
<td>AAL01213</td>
<td>5.00E-25</td>
<td>mannanase ManA [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig29726</td>
<td>CBM10</td>
<td>574</td>
<td>1415</td>
<td>AAP30750</td>
<td>2.00E-25</td>
<td>celllobiohydrolase Cel6C [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig204</td>
<td>CBM10</td>
<td>842</td>
<td>4025</td>
<td>AAB92679</td>
<td>2.00E-27</td>
<td>cellulase C [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig29063</td>
<td>CBM10</td>
<td>1299</td>
<td>957</td>
<td>ZP_061453</td>
<td>1.00E-113</td>
<td>hypothetical protein RflaF_19133 [Ruminococcus flavefaciens FD-1]</td>
</tr>
<tr>
<td>Contig28951</td>
<td>CBM10</td>
<td>534</td>
<td>1177</td>
<td>AAB69092</td>
<td>2.00E-23</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig21348</td>
<td>CBM10</td>
<td>787</td>
<td>4478</td>
<td>AAP30750</td>
<td>1.00E-25</td>
<td>celllobiohydrolase Cel6C [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig539</td>
<td>CBM10</td>
<td>1386</td>
<td>5913</td>
<td>AAL01214</td>
<td>2.00E-27</td>
<td>endo-glucanase CelJ [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig29505</td>
<td>CBM10</td>
<td>2076</td>
<td>2721</td>
<td>ZP_046681</td>
<td>4.00E-52</td>
<td>methionine adenosyltransferase [Clostridiales bacterium 1_7_47_FAA]</td>
</tr>
<tr>
<td>Contig29865</td>
<td>CBM10</td>
<td>604</td>
<td>210</td>
<td>AAL01214</td>
<td>4.00E-26</td>
<td>endo-glucanase CelJ [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig21376</td>
<td>CBM10</td>
<td>1283</td>
<td>1725</td>
<td>AAB69092</td>
<td>0</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig3265</td>
<td>CBM10</td>
<td>618</td>
<td>512</td>
<td>AAM94167</td>
<td>3.00E-23</td>
<td>cellulosomal glycoside hydrolase family 6 exoglucanase Cel6A [Piromyces equi]</td>
</tr>
<tr>
<td>Contig86922</td>
<td>CBM10</td>
<td>860</td>
<td>866</td>
<td>AAF70241</td>
<td>1.00E-182</td>
<td>feruloyl esterase A [Orpinomyces sp. PC-2]</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM10</td>
<td>1355</td>
<td>2867</td>
<td>AAB69092</td>
<td>AAF70241.1</td>
<td>AF164351.1</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>----------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>Contig24133</td>
<td>CBM10</td>
<td>895</td>
<td>500</td>
<td>AAB69090</td>
<td>gi</td>
<td>2231243</td>
</tr>
<tr>
<td>Contig36</td>
<td>CBM10</td>
<td>663</td>
<td>1752</td>
<td>AAN76734</td>
<td>gi</td>
<td>2599095</td>
</tr>
<tr>
<td>Contig17825</td>
<td>CBM10</td>
<td>2062</td>
<td>1161</td>
<td>ZP_061453</td>
<td>gi</td>
<td>268611605</td>
</tr>
<tr>
<td>Contig325</td>
<td>CBM10</td>
<td>791</td>
<td>60</td>
<td>YP_003250</td>
<td>gi</td>
<td>261416337</td>
</tr>
<tr>
<td>Contig4100</td>
<td>CBM10</td>
<td>679</td>
<td>174</td>
<td>CAA57717</td>
<td>gi</td>
<td>561530</td>
</tr>
<tr>
<td>Contig4551</td>
<td>CBM10</td>
<td>1122</td>
<td>613</td>
<td>AAQ10006</td>
<td>gi</td>
<td>33329212</td>
</tr>
<tr>
<td>Contig2401</td>
<td>CBM10</td>
<td>1237</td>
<td>488</td>
<td>YP_003250</td>
<td>gi</td>
<td>261417274</td>
</tr>
<tr>
<td>Contig29858</td>
<td>CBM10</td>
<td>1137</td>
<td>989</td>
<td>AAQ10006</td>
<td>gi</td>
<td>33329212</td>
</tr>
<tr>
<td>Contig1365</td>
<td>CBM10</td>
<td>1603</td>
<td>107</td>
<td>ZP_054964</td>
<td>gi</td>
<td>256755664</td>
</tr>
<tr>
<td>Contig7493</td>
<td>CBM10</td>
<td>977</td>
<td>382</td>
<td>CBK75021</td>
<td>gi</td>
<td>291519800</td>
</tr>
<tr>
<td>Contig1319</td>
<td>CBM10</td>
<td>766</td>
<td>91</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM10</td>
<td>Length</td>
<td>name</td>
<td>gi</td>
<td>gb</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>------</td>
<td>---</td>
<td>---</td>
<td>---------</td>
</tr>
<tr>
<td>NODE_21249</td>
<td>CBM10</td>
<td>567</td>
<td>231</td>
<td>AAD43818</td>
<td>gi</td>
<td>5457159</td>
</tr>
<tr>
<td>Contig6232</td>
<td>CBM10</td>
<td>540</td>
<td>218</td>
<td>AAM94167</td>
<td>gi</td>
<td>33620325</td>
</tr>
<tr>
<td>Contig28921</td>
<td>CBM10</td>
<td>1122</td>
<td>1289</td>
<td>ACL68347</td>
<td>gi</td>
<td>219964511</td>
</tr>
<tr>
<td>Contig24540</td>
<td>CBM10</td>
<td>741</td>
<td>232</td>
<td>AAB69090</td>
<td>gi</td>
<td>2231243</td>
</tr>
<tr>
<td>Contig3018'</td>
<td>CBM10</td>
<td>2298</td>
<td>470</td>
<td>AAL01211</td>
<td>gi</td>
<td>15529294</td>
</tr>
<tr>
<td>Contig30041</td>
<td>CBM10</td>
<td>1158</td>
<td>886</td>
<td>ZP_06142046</td>
<td>gi</td>
<td>268608319</td>
</tr>
<tr>
<td>Contig28851</td>
<td>CBM10</td>
<td>863</td>
<td>388</td>
<td>YP_003250020</td>
<td>gi</td>
<td>261416337</td>
</tr>
<tr>
<td>Contig5636</td>
<td>CBM10</td>
<td>2418</td>
<td>2391</td>
<td>ZP_04668137</td>
<td>gi</td>
<td>239625106</td>
</tr>
<tr>
<td>Contig1477</td>
<td>CBM10</td>
<td>1160</td>
<td>667</td>
<td>ACL68347</td>
<td>gi</td>
<td>219964511</td>
</tr>
<tr>
<td>Contig28836</td>
<td>CBM10</td>
<td>994</td>
<td>874</td>
<td>ACL68347</td>
<td>gi</td>
<td>219964511</td>
</tr>
<tr>
<td>Contig28653</td>
<td>CBM10</td>
<td>554</td>
<td>758</td>
<td>AAB69092</td>
<td>gi</td>
<td>2231247</td>
</tr>
<tr>
<td>Contig582</td>
<td>CBM10</td>
<td>1150</td>
<td>495</td>
<td>AAQ10005</td>
<td>gi</td>
<td>33329210</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM</td>
<td>1270</td>
<td>683</td>
<td>AAP30750</td>
<td>gi</td>
<td>30315041</td>
</tr>
<tr>
<td>Contig225</td>
<td>CBM10</td>
<td>1135</td>
<td>300</td>
<td>ZP_061420</td>
<td>46</td>
<td>gi</td>
</tr>
<tr>
<td>Contig31183</td>
<td>CBM10</td>
<td>644</td>
<td>241</td>
<td>YP_003250</td>
<td>045</td>
<td>gi</td>
</tr>
<tr>
<td>Contig772</td>
<td>CBM10</td>
<td>988</td>
<td>325</td>
<td>ABY52795</td>
<td>gi</td>
<td>164375379</td>
</tr>
<tr>
<td>Contig7593</td>
<td>CBM10</td>
<td>1084</td>
<td>1070</td>
<td>YP_003250</td>
<td>020</td>
<td>gi</td>
</tr>
<tr>
<td>Contig7347</td>
<td>CBM10</td>
<td>2022</td>
<td>1576</td>
<td>ZP_046681</td>
<td>37</td>
<td>gi</td>
</tr>
<tr>
<td>Contig31237</td>
<td>CBM10</td>
<td>876</td>
<td>470</td>
<td>ACL68347</td>
<td>gi</td>
<td>219964511</td>
</tr>
<tr>
<td>Contig30406</td>
<td>CBM10</td>
<td>753</td>
<td>533</td>
<td>AAB69092</td>
<td>gi</td>
<td>2231247</td>
</tr>
<tr>
<td>Contig29362</td>
<td>CBM10</td>
<td>566</td>
<td>681</td>
<td>AAP30750</td>
<td>gi</td>
<td>30315041</td>
</tr>
<tr>
<td>Contig1116</td>
<td>CBM10</td>
<td>507</td>
<td>109</td>
<td>AAP30747</td>
<td>gi</td>
<td>30315035</td>
</tr>
<tr>
<td>Contig5325</td>
<td>CBM10</td>
<td>1015</td>
<td>424</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
</tr>
<tr>
<td>Contig22393</td>
<td>CBM10</td>
<td>687</td>
<td>601</td>
<td>AAP30750</td>
<td>gi</td>
<td>30315041</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM10</td>
<td>Length</td>
<td>Start</td>
<td>End</td>
<td>gi</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>--------</td>
<td>-------</td>
<td>-----</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Contig606</td>
<td>CBM10</td>
<td>790</td>
<td>234</td>
<td>Q9Y871</td>
<td>4.00E-56</td>
<td>121</td>
</tr>
<tr>
<td>Contig21916</td>
<td>CBM10</td>
<td>708</td>
<td>256</td>
<td>AAM81966</td>
<td>2.00E-47</td>
<td>237</td>
</tr>
<tr>
<td>Contig1411</td>
<td>CBM10</td>
<td>1515</td>
<td>831</td>
<td>ZP_0614504</td>
<td>2.00E-50</td>
<td>251</td>
</tr>
<tr>
<td>Contig12252</td>
<td>CBM10</td>
<td>572</td>
<td>153</td>
<td>AAP30750</td>
<td>1.00E-21</td>
<td>85</td>
</tr>
<tr>
<td>Contig26764</td>
<td>CBM10</td>
<td>1454</td>
<td>594</td>
<td>ZP_0549636</td>
<td>1.00E-73</td>
<td>361</td>
</tr>
<tr>
<td>Contig30162</td>
<td>CBM10</td>
<td>785</td>
<td>513</td>
<td>YP_003248565</td>
<td>6.00E-66</td>
<td>173</td>
</tr>
<tr>
<td>Contig29219</td>
<td>CBM10</td>
<td>1180</td>
<td>4114</td>
<td>YP_003250045</td>
<td>1.00E-116</td>
<td>331</td>
</tr>
<tr>
<td>Contig11932</td>
<td>CBM10</td>
<td>1037</td>
<td>1156</td>
<td>AAM81966</td>
<td>2.00E-23</td>
<td>82</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM</td>
<td>Start</td>
<td>Length</td>
<td>Description</td>
<td>E-value</td>
<td>Expect</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>-------------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td>4621</td>
<td>CBM10</td>
<td>901</td>
<td>355</td>
<td>BAB39493</td>
<td>2.00E-49</td>
<td>145</td>
</tr>
<tr>
<td>25171</td>
<td>CBM10</td>
<td>1883</td>
<td>2210</td>
<td>ZP_054004_07</td>
<td>2.00E-24</td>
<td>405</td>
</tr>
<tr>
<td>28821</td>
<td>CBM10</td>
<td>1843</td>
<td>931</td>
<td>ZP_054004_07</td>
<td>3.00E-20</td>
<td>390</td>
</tr>
<tr>
<td>6286</td>
<td>CBM10</td>
<td>1015</td>
<td>642</td>
<td>Q12647</td>
<td>1.00E-21</td>
<td>87</td>
</tr>
<tr>
<td>10643</td>
<td>CBM10</td>
<td>731</td>
<td>270</td>
<td>NP_149281</td>
<td>1.00E-21</td>
<td>115</td>
</tr>
<tr>
<td>4497</td>
<td>CBM10</td>
<td>825</td>
<td>327</td>
<td>AAN76734</td>
<td>2.00E-20</td>
<td>79</td>
</tr>
<tr>
<td>22537</td>
<td>CBM10</td>
<td>1628</td>
<td>977</td>
<td>ZP_061420_46</td>
<td>4.00E-52</td>
<td>243</td>
</tr>
<tr>
<td>1836</td>
<td>CBM10</td>
<td>1232</td>
<td>175</td>
<td>YP_003250_045</td>
<td>1.00E-121</td>
<td>266</td>
</tr>
<tr>
<td>3112</td>
<td>CBM10</td>
<td>630</td>
<td>368</td>
<td>ACL68347</td>
<td>6.00E-50</td>
<td>163</td>
</tr>
<tr>
<td>269</td>
<td>CBM10</td>
<td>582</td>
<td>2554</td>
<td>AAB92679</td>
<td>4.00E-18</td>
<td>64</td>
</tr>
<tr>
<td>2420</td>
<td>CBM10</td>
<td>1091</td>
<td>420</td>
<td>CAB92325</td>
<td>1.00E-20</td>
<td>132</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Contig29549</td>
<td>CBM10</td>
<td>525</td>
<td>494</td>
<td>AAM81966</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
<td></td>
</tr>
<tr>
<td>Contig1228</td>
<td>CBM10</td>
<td>978</td>
<td>45</td>
<td>ZP_061428 857</td>
<td>glycoside hydrolase family protein [Ruminococcus flavefaciens FD-1]</td>
<td></td>
</tr>
<tr>
<td>Contig688</td>
<td>CBM10</td>
<td>1403</td>
<td>642</td>
<td>ZP_054964 36</td>
<td>lipolytic protein G-D-S-L family [Clostridium papyrosolvens DSM 2782]</td>
<td></td>
</tr>
<tr>
<td>Contig1503</td>
<td>CBM10</td>
<td>2230</td>
<td>2105</td>
<td>ZP_061420 046</td>
<td>hypothetical protein RflaF_02313 [Ruminococcus flavefaciens FD-1]</td>
<td></td>
</tr>
<tr>
<td>Contig16443</td>
<td>CBM10</td>
<td>775</td>
<td>629</td>
<td>P55296</td>
<td>carbohydrate-binding family 6 protein [Clostridium thermocellum ATCC 27405]</td>
<td></td>
</tr>
<tr>
<td>Contig6381</td>
<td>CBM10</td>
<td>701</td>
<td>232</td>
<td>AAL01212</td>
<td>celllobiohydrolase II-like cellulase Cell [Orpinomyces sp. PC-2]</td>
<td></td>
</tr>
<tr>
<td>Contig22989</td>
<td>CBM10</td>
<td>663</td>
<td>473</td>
<td>YP_001038 591</td>
<td>carbohydrate-binding family 6 protein [Clostridium thermocellum ATCC 27405]</td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Chr</td>
<td>Start</td>
<td>End</td>
<td>GI</td>
<td>Sp</td>
<td>GenBank ID</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>------------</td>
</tr>
<tr>
<td>Contig30121'</td>
<td>CBM10</td>
<td>712</td>
<td>588</td>
<td>AAF70241</td>
<td>gi</td>
<td>7839348</td>
</tr>
<tr>
<td>Contig22174</td>
<td>CBM10</td>
<td>1175</td>
<td>718</td>
<td>AAD43818</td>
<td>gi</td>
<td>5457159</td>
</tr>
<tr>
<td>Contig29181'</td>
<td>CBM10</td>
<td>972</td>
<td>1684</td>
<td>AAD43818</td>
<td>gi</td>
<td>5457159</td>
</tr>
<tr>
<td>Contig22301</td>
<td>CBM10</td>
<td>1693</td>
<td>1197</td>
<td>AAF70241</td>
<td>gi</td>
<td>7839348</td>
</tr>
<tr>
<td>Contig2744</td>
<td>CBM10</td>
<td>841</td>
<td>740</td>
<td>P55296</td>
<td>gi</td>
<td>1708917</td>
</tr>
<tr>
<td>Contig30404</td>
<td>CBM10</td>
<td>516</td>
<td>287</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
</tr>
<tr>
<td>Contig21698</td>
<td>CBM10</td>
<td>556</td>
<td>261</td>
<td>P55297</td>
<td>gi</td>
<td>1708920</td>
</tr>
<tr>
<td>Contig18696</td>
<td>CBM10</td>
<td>930</td>
<td>585</td>
<td>AAB69092</td>
<td>gi</td>
<td>2231247</td>
</tr>
<tr>
<td>Contig29445</td>
<td>CBM10</td>
<td>1015</td>
<td>671</td>
<td>ZP_05496436</td>
<td>gi</td>
<td>256755664</td>
</tr>
<tr>
<td>NODE_9175_length_815_cov_8.579141</td>
<td>CBM10</td>
<td>859</td>
<td>657</td>
<td>YP_003099042</td>
<td>gi</td>
<td>256375382</td>
</tr>
<tr>
<td>Contig22783</td>
<td>CBM10</td>
<td>1101</td>
<td>3835</td>
<td>AAD43818</td>
<td>gi</td>
<td>5457159</td>
</tr>
<tr>
<td>Contig22403</td>
<td>CBM10</td>
<td>678</td>
<td>774</td>
<td>AAB92679</td>
<td>gi</td>
<td>1813486</td>
</tr>
<tr>
<td>Contig9577</td>
<td>CBM10</td>
<td>598</td>
<td>259</td>
<td>CBL17363</td>
<td>gi</td>
<td>291544254</td>
</tr>
<tr>
<td>Contig21635</td>
<td>CBM10</td>
<td>1309</td>
<td>761</td>
<td>YP_002504781</td>
<td>gi</td>
<td>220927872</td>
</tr>
<tr>
<td>NODE_74290_length_492_cov_4.481707</td>
<td>CBM10</td>
<td>528</td>
<td>205</td>
<td>AAD43818</td>
<td>gi</td>
<td>5457159</td>
</tr>
<tr>
<td>Contig24467</td>
<td>CBM10</td>
<td>676</td>
<td>365</td>
<td>AAC06321</td>
<td>gi</td>
<td>2981484</td>
</tr>
<tr>
<td>Contig8953</td>
<td>CBM10</td>
<td>742</td>
<td>479</td>
<td>CAB92326</td>
<td>gi</td>
<td>8052316</td>
</tr>
<tr>
<td>Contig5038</td>
<td>CBM10</td>
<td>1395</td>
<td>938</td>
<td>ZP_04668137</td>
<td>gi</td>
<td>239625106</td>
</tr>
<tr>
<td>Contig23376</td>
<td>CBM10</td>
<td>501</td>
<td>165</td>
<td>AAC05164</td>
<td>gi</td>
<td>2935581</td>
</tr>
<tr>
<td>Contig19</td>
<td>CBM10</td>
<td>572</td>
<td>24</td>
<td>AAM94167</td>
<td>gi</td>
<td>33620325</td>
</tr>
<tr>
<td>Contig29925</td>
<td>CBM10</td>
<td>944</td>
<td>573</td>
<td>P55296</td>
<td>gi</td>
<td>1708917</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM</td>
<td>Accession</td>
<td>gi</td>
<td>gb</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
<td>-----------</td>
<td>----</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Contig23136</td>
<td>CBM10</td>
<td>521</td>
<td>243</td>
<td>ACL68347</td>
<td>mannan endo-1,4-beta-mannosidase; mannanase A [Piromyces sp.] endo-1,4-beta-xylanase [Neocallimastix patriciarum]</td>
<td></td>
</tr>
<tr>
<td>Contig19516</td>
<td>CBM10</td>
<td>1437</td>
<td>3285</td>
<td>YP_003250 020</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
</tr>
<tr>
<td>Contig1217</td>
<td>CBM10</td>
<td>573</td>
<td>252</td>
<td>AAL01211</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
</tr>
<tr>
<td>Contig29965</td>
<td>CBM10</td>
<td>1143</td>
<td>1000</td>
<td>YP_003250 020</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
</tr>
<tr>
<td>Contig23266</td>
<td>CBM10</td>
<td>1413</td>
<td>1003</td>
<td>YP_003250 020</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
</tr>
<tr>
<td>Contig30390</td>
<td>CBM10</td>
<td>1022</td>
<td>5033</td>
<td>YP_003250 020</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
</tr>
<tr>
<td>Contig29158</td>
<td>CBM10</td>
<td>1144</td>
<td>1521</td>
<td>YP_003250 020</td>
<td>hypothetical protein Fisuc_1948 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
</tr>
<tr>
<td>Contig23209</td>
<td>CBM10</td>
<td>943</td>
<td>686</td>
<td>Q9Y871</td>
<td>RecName: Full=Feruloyl esterase B; AltName: Full=Ferulic acid esterase B; AltName: Full=Esterase A; Short=EstA; AltName: Full=Cinnamoyl ester hydrolase; Flags: Precursor</td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Domain</td>
<td>Length</td>
<td>Accession</td>
<td>Accession 2</td>
<td>Accession 3</td>
<td>[Piromyces equi]</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>--------</td>
<td>-----------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>NODE_48706</td>
<td>CBM10</td>
<td>567</td>
<td>gi</td>
<td>158428896</td>
<td>pdb</td>
<td>J4M[A</td>
</tr>
<tr>
<td>Contig29630</td>
<td>CBM10</td>
<td>605</td>
<td>gi</td>
<td>2231243</td>
<td>gb</td>
<td>AAB69090.1</td>
</tr>
<tr>
<td>Contig4303</td>
<td>CBM10</td>
<td>1438</td>
<td>gi</td>
<td>30315035</td>
<td>gb</td>
<td>AAP30747.1</td>
</tr>
<tr>
<td>Contig2424</td>
<td>CBM10</td>
<td>1033</td>
<td>gi</td>
<td>2231247</td>
<td>gb</td>
<td>AAB69092.1</td>
</tr>
<tr>
<td>Contig30327</td>
<td>CBM10</td>
<td>1485</td>
<td>gi</td>
<td>23821548</td>
<td>gb</td>
<td>Q9Y871.1</td>
</tr>
<tr>
<td>Contig23562</td>
<td>CBM10</td>
<td>599</td>
<td>gi</td>
<td>2353007</td>
<td>gb</td>
<td>AAB69348.1</td>
</tr>
<tr>
<td>Contig10196</td>
<td>CBM10</td>
<td>657</td>
<td>gi</td>
<td>219964511</td>
<td>gb</td>
<td>ACL68347.1</td>
</tr>
<tr>
<td>Contig8050</td>
<td>CBM10</td>
<td>517</td>
<td>gi</td>
<td>219964511</td>
<td>gb</td>
<td>ACL68347.1</td>
</tr>
<tr>
<td>Contig21712</td>
<td>CBM10</td>
<td>2367</td>
<td>gi</td>
<td>229085040</td>
<td>ref</td>
<td>ZP_04217 292.1</td>
</tr>
<tr>
<td>NODE_1486</td>
<td>CBM10</td>
<td>934</td>
<td>gi</td>
<td>219964511</td>
<td>gb</td>
<td>ACL68347.1</td>
</tr>
<tr>
<td>Contig22573</td>
<td>CBM10</td>
<td>1373</td>
<td>gi</td>
<td>256755664</td>
<td>ref</td>
<td>ZP_05496436.1</td>
</tr>
<tr>
<td>Contig10776</td>
<td>CBM10</td>
<td>2066</td>
<td>gi</td>
<td>261868885</td>
<td>gb</td>
<td>ACY02060</td>
</tr>
<tr>
<td>Contig</td>
<td>Accession</td>
<td>gi</td>
<td>Length</td>
<td>Protein</td>
<td>Function</td>
<td>Identity</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>---</td>
<td>--------</td>
<td>---------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>NODE_39361</td>
<td>CBM10</td>
<td>505</td>
<td>135</td>
<td>AAL01212</td>
<td>celllobiohydrolase II-like cellulase [Orpinomyces sp. PC-2]</td>
<td>8.00E-14</td>
</tr>
<tr>
<td>Contig3198</td>
<td>CBM10</td>
<td>728</td>
<td>293</td>
<td>Q12667</td>
<td>RecName: Full=Endo-1,4-beta-xylanase A; AltName: Full=1,4-beta-D-xylan xylanohydrolase; AltName: Full=Xylanase A; Short=XYLA; Flags: Precursor gi</td>
<td>1197372</td>
</tr>
<tr>
<td>Contig21960</td>
<td>CBM10</td>
<td>1037</td>
<td>580</td>
<td>AAD43818</td>
<td>endoglucanase precursor [Piromyces rhizinflatus]</td>
<td>2.00E-08</td>
</tr>
<tr>
<td>Contig25183</td>
<td>CBM10</td>
<td>581</td>
<td>228</td>
<td>AAN76734</td>
<td>cellulase Cel48A precursor [Piromyces sp. E2]</td>
<td>2.00E-08</td>
</tr>
<tr>
<td>Contig24413</td>
<td>CBM10</td>
<td>582</td>
<td>168</td>
<td>AAN76734</td>
<td>cellulase Cel48A precursor [Piromyces sp. E2]</td>
<td>2.00E-08</td>
</tr>
<tr>
<td>Contig8015</td>
<td>CBM10</td>
<td>636</td>
<td>315</td>
<td>AAB92678</td>
<td>cellulase A [Orpinomyces sp. PC-2]</td>
<td>9.00E-11</td>
</tr>
<tr>
<td>Contig6151</td>
<td>CBM10</td>
<td>546</td>
<td>222</td>
<td>XP_00237</td>
<td>beta-galactosidase, putative [Aspergillus flavus NRRL3357]</td>
<td>0.46</td>
</tr>
<tr>
<td>Contig1878</td>
<td>Ricin_B_lectin</td>
<td>1344</td>
<td>181</td>
<td>CAJ81035</td>
<td>putative galactocerebrosidase [Actinoplanes sp. SE50/110]</td>
<td>4.00E-11</td>
</tr>
<tr>
<td>Contig24564</td>
<td>Ricin_B_lectin</td>
<td>1005</td>
<td>552</td>
<td>ZP_048055</td>
<td>Ricin B lectin [Clostridium cellulovorans 743B]</td>
<td>4.00E-08</td>
</tr>
<tr>
<td>Contig30729</td>
<td>Ricin_B_lectin</td>
<td>1230</td>
<td>512</td>
<td>CBL16867</td>
<td>Pectate lyase [Ruminococcus sp. 18P13]</td>
<td>1.00E-06</td>
</tr>
<tr>
<td>Contig</td>
<td>Protein</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Contig3006</td>
<td>Ricin B lectin</td>
<td>1660</td>
<td>298</td>
<td>ZP_04805535</td>
<td>gi</td>
<td>242260820</td>
</tr>
<tr>
<td>Contig12042</td>
<td>Ricin B lectin</td>
<td>1096</td>
<td>1009</td>
<td>YP_003115585</td>
<td>gi</td>
<td>256394021</td>
</tr>
<tr>
<td>Contig2473</td>
<td>Ricin B lectin</td>
<td>1163</td>
<td>414</td>
<td>YP_001195010</td>
<td>gi</td>
<td>146300419</td>
</tr>
<tr>
<td>Contig2316</td>
<td>Ricin B lectin</td>
<td>894</td>
<td>388</td>
<td>XP_002520852</td>
<td>gi</td>
<td>255556965</td>
</tr>
<tr>
<td>Contig7830</td>
<td>Ricin B lectin</td>
<td>811</td>
<td>286</td>
<td>XP_002372370</td>
<td>gi</td>
<td>238482263</td>
</tr>
<tr>
<td>Contig399</td>
<td>Ricin B lectin</td>
<td>1350</td>
<td>1239</td>
<td>YP_003250020</td>
<td>gi</td>
<td>261416337</td>
</tr>
<tr>
<td>Contig13671</td>
<td>Ricin B lectin</td>
<td>638</td>
<td>176</td>
<td>CBL1686</td>
<td>gi</td>
<td>291543758</td>
</tr>
<tr>
<td>Contig24952</td>
<td>Ricin B lectin</td>
<td>576</td>
<td>424</td>
<td>YP_590752</td>
<td>gi</td>
<td>94968704</td>
</tr>
<tr>
<td>Contig</td>
<td>Protein</td>
<td>P</td>
<td>Q</td>
<td>GI</td>
<td>Ref</td>
<td>E-value</td>
</tr>
<tr>
<td>----------</td>
<td>-----------</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>31172</td>
<td>Ricin B_</td>
<td>552</td>
<td>306</td>
<td>XP_001458591</td>
<td>gi</td>
<td>145545815</td>
</tr>
<tr>
<td>4385</td>
<td>Ricin B_</td>
<td>537</td>
<td>177</td>
<td>XP_002372370</td>
<td>gi</td>
<td>238482263</td>
</tr>
<tr>
<td>23124</td>
<td>Chitin_b_</td>
<td>1060</td>
<td>918</td>
<td>Q9AVB0</td>
<td>gi</td>
<td>564046599</td>
</tr>
<tr>
<td>27417</td>
<td>Chitin_b_</td>
<td>857</td>
<td>474</td>
<td>XP_384927</td>
<td>gi</td>
<td>46119101</td>
</tr>
<tr>
<td>21647</td>
<td>Chitin_b_</td>
<td>799</td>
<td>150</td>
<td>XP_384927</td>
<td>gi</td>
<td>46119101</td>
</tr>
<tr>
<td>27332</td>
<td>Chitin_b_</td>
<td>716</td>
<td>606</td>
<td>EEY22269</td>
<td>gi</td>
<td>261359841</td>
</tr>
<tr>
<td>590</td>
<td>Chitin_b_</td>
<td>1113</td>
<td>780</td>
<td>AAR97890</td>
<td>gi</td>
<td>40950521</td>
</tr>
<tr>
<td>1173</td>
<td>Chitin_b_</td>
<td>988</td>
<td>33</td>
<td>EEY22269</td>
<td>gi</td>
<td>261359841</td>
</tr>
</tbody>
</table>

**Contig 31172**
Ricin B_lectin
552 306 XP_001458591

**Contig 4385**
Ricin_B_lectin
537 177 XP_002372370

**Contig 23124**
Chitin_bind_1
1060 918 Q9AVB0

**Contig 5823**
Chitin_bind_1
1002 402 EEY22269

**Contig 27417**
Chitin_bind_1
857 474 XP_384927

**Contig 4939**
Chitin_bind_1
893 509 Q9AVB0

**Contig 7224**
Chitin_bind_1
781 322 EEY22269

**Contig 72332**
Chitin_bind_1
716 606 EEY22269

**Contig 21647**
Chitin_bind_1
799 150 XP_384927

**Contig 590**
Chitin_bind_1
1113 780 AAR97890

**Contig 1173**
Chitin_bind_1
988 33 EEY22269
<p>| Contig28903 | Chitin_bind_1 | 944 | 1849 | EEY22269 | gi|261359841|gb|EEY22269 .1| 1.00E-41 | 275 | 37.1 | lectin [Verticillium albo-atrum VaMs.102] |
| Contig29161 | Chitin_bind_1 | 829 | 742 | XP_384927 | gi|46119101|ef|XP_384927 .1| 4.00E-27 | 281 | 33.5 | hypothetical protein FG04751.1 [Gibberella zeae PH-1] |
| NODE_6978_length_664_cov_3.314759 | Chitin_bind_1 | 700 | 124 | XP_001905 632 | gi|171681377|ref|XP_001905 5632.1| 3.00E-30 | 218 | 34.9 | unnamed protein product [Podospora anserina] |
| Contig241 | Chitin_bind_1 | 1220 | 843 | EEY22269 | gi|261359841|gb|EEY22269 .1| 6.00E-41 | 269 | 33.1 | lectin [Verticillium albo-atrum VaMs.102] |
| Contig10585 | Chitin_bind_1 | 2014 | 963 | EEY22269 | gi|261359841|gb|EEY22269 .1| 4.00E-35 | 393 | 26.0 | lectin [Verticillium albo-atrum VaMs.102] |
| Contig23027 | Chitin_bind_1 | 632 | 429 | EEY22269 | gi|261359841|gb|EEY22269 .1| 6.00E-34 | 213 | 34.3 | lectin [Verticillium albo-atrum VaMs.102] |
| Contig23154 | Chitin_bind_1 | 707 | 385 | EEY22269 | gi|261359841|gb|EEY22269 .1| 4.00E-33 | 191 | 39.8 | lectin [Verticillium albo-atrum VaMs.102] |
| Contig10002 | Chitin_bind_1 | 809 | 245 | BAI44118 | gi|260279095|db|BAI44118 .1| 7.00E-22 | 211 | 36.0 | chitin binding protein 4 [Magnaporthe oryzae] |
| Contig6491 | Chitin_bind_1 | 1998 | 2815 | EEY22269 | gi|261359841|gb|EEY22269 .1| 2.00E-34 | 229 | 35.4 | lectin [Verticillium albo-atrum VaMs.102] |
| Contig20734 | Chitin_bind_1 | 938 | 620 | XP_384927 | gi|46119101|ef|XP_384927 .1| 2.00E-25 | 236 | 37.7 | hypothetical protein FG04751.1 [Gibberella zeae PH-1] |
| Contig24365 | Chitin_bind_1 | 1037 | 780 | XP_384927 | gi|46119101|ef|XP_384927 .1| 2.00E-25 | 204 | 35.8 | hypothetical protein FG04751.1 [Gibberella zeae PH-1] |
| Contig29503 | Chitin_bind_1 | 630 | 664 | BAI44115 | gi|260279089|db|BAI44115 .1| 9.00E-09 | 152 | 25.7 | chitin binding protein 1 [Magnaporthe oryzae] |
| Contig15249 | Chitin_bind_1 | 512 | 518 | EEY22269 | gi|261359841|gb|EEY22269 .1| 8.00E-31 | 170 | 39.4 | lectin [Verticillium albo-atrum VaMs.102] |</p>
<table>
<thead>
<tr>
<th>Contig</th>
<th>Status</th>
<th>Start</th>
<th>End</th>
<th>Accession</th>
<th>Accession</th>
<th>Description</th>
<th>E-values</th>
<th>Scores</th>
<th>Name</th>
<th>Metadata</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig5044</td>
<td>253</td>
<td>875</td>
<td>533</td>
<td>XP_384927</td>
<td>gi</td>
<td>46119101</td>
<td>r ef[XP_384927 .1]</td>
<td>2.00E-25</td>
<td>272</td>
<td>32.0</td>
</tr>
<tr>
<td>Contig360</td>
<td>567</td>
<td>1752</td>
<td>EEY22269</td>
<td>gi</td>
<td>261359841</td>
<td>gb</td>
<td>EEY22269 .1]</td>
<td>3.00E-34</td>
<td>178</td>
<td>41.6</td>
</tr>
<tr>
<td>Contig5276</td>
<td>657</td>
<td>399</td>
<td>BAI44126</td>
<td>gi</td>
<td>260279111</td>
<td>dbj</td>
<td>BAI44126 .1]</td>
<td>8.00E-19</td>
<td>167</td>
<td>31.7</td>
</tr>
<tr>
<td>Contig14080</td>
<td>653</td>
<td>562</td>
<td>BAI44118</td>
<td>gi</td>
<td>260279095</td>
<td>dbj</td>
<td>BAI44118 .1]</td>
<td>3.00E-24</td>
<td>187</td>
<td>33.7</td>
</tr>
<tr>
<td>Contig22883</td>
<td>1040</td>
<td>724</td>
<td>XP_001905632</td>
<td>gi</td>
<td>171681377</td>
<td>ref[XP_001905632.l]</td>
<td>4.00E-27</td>
<td>220</td>
<td>33.2</td>
<td>unnamed protein product [Podospora anserina]</td>
</tr>
<tr>
<td>Contig26354</td>
<td>613</td>
<td>768</td>
<td>BAI44115</td>
<td>gi</td>
<td>260279089</td>
<td>dbj</td>
<td>BAI44115 .1]</td>
<td>7.00E-24</td>
<td>139</td>
<td>41.7</td>
</tr>
<tr>
<td>Contig5345</td>
<td>650</td>
<td>452</td>
<td>BAI44118</td>
<td>gi</td>
<td>260279095</td>
<td>dbj</td>
<td>BAI44118 .1]</td>
<td>3.00E-26</td>
<td>178</td>
<td>39.9</td>
</tr>
<tr>
<td>Contig3168</td>
<td>530</td>
<td>188</td>
<td>BAI44118</td>
<td>gi</td>
<td>260279095</td>
<td>dbj</td>
<td>BAI44118 .1]</td>
<td>5.00E-20</td>
<td>129</td>
<td>37.2</td>
</tr>
<tr>
<td>Contig4429</td>
<td>630</td>
<td>1156</td>
<td>XP_384927</td>
<td>gi</td>
<td>46119101</td>
<td>r ef[XP_384927 .1]</td>
<td>7.00E-33</td>
<td>216</td>
<td>38.9</td>
<td>hypothetical protein FG04751.1 [Gibberella zeae PH-1]</td>
</tr>
<tr>
<td>Contig8326</td>
<td>686</td>
<td>343</td>
<td>EEY22269</td>
<td>gi</td>
<td>261359841</td>
<td>gb</td>
<td>EEY22269 .1]</td>
<td>8.00E-25</td>
<td>110</td>
<td>49.1</td>
</tr>
<tr>
<td>Contig2280</td>
<td>2155</td>
<td>2020</td>
<td>EEY22269</td>
<td>gi</td>
<td>261359841</td>
<td>gb</td>
<td>EEY22269 .1]</td>
<td>8.00E-26</td>
<td>201</td>
<td>33.3</td>
</tr>
<tr>
<td>Contig12064</td>
<td>506</td>
<td>243</td>
<td>XP_384927</td>
<td>gi</td>
<td>46119101</td>
<td>r ef[XP_384927 .1]</td>
<td>1.00E-29</td>
<td>157</td>
<td>43.9</td>
<td>hypothetical protein FG04751.1 [Gibberella zeae PH-1]</td>
</tr>
<tr>
<td>Contig29245</td>
<td>1200</td>
<td>1544</td>
<td>Q9AVB0</td>
<td>gi</td>
<td>56404659</td>
<td>sp</td>
<td>Q9AVB0.1] LECB_PHY A M</td>
<td>5.00E-23</td>
<td>175</td>
<td>34.9</td>
</tr>
<tr>
<td>Contig</td>
<td>Chitin_binding_1</td>
<td>Accession</td>
<td>GI</td>
<td>Accession</td>
<td>DBJ</td>
<td>ID</td>
<td>Score</td>
<td>E-value</td>
<td>Length</td>
<td>Identity</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------</td>
<td>-----------</td>
<td>----</td>
<td>-----------</td>
<td>-----</td>
<td>----</td>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>Contig8652</td>
<td>Chitin_binding_1</td>
<td>709</td>
<td>256</td>
<td>BAI44126</td>
<td></td>
<td></td>
<td>2.00E-17</td>
<td>174</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td>Contig8855</td>
<td>Chitin_binding_1</td>
<td>597</td>
<td>213</td>
<td>EEY22269</td>
<td></td>
<td></td>
<td>2.00E-27</td>
<td>226</td>
<td>34.5</td>
<td></td>
</tr>
<tr>
<td>Contig21263</td>
<td>Chitin_binding_1</td>
<td>746</td>
<td>813</td>
<td>BAI44115</td>
<td></td>
<td></td>
<td>3.00E-22</td>
<td>202</td>
<td>35.6</td>
<td></td>
</tr>
<tr>
<td>Contig23746</td>
<td>Chitin_binding_1</td>
<td>919</td>
<td>1194</td>
<td>BAI44118</td>
<td></td>
<td></td>
<td>2.00E-17</td>
<td>117</td>
<td>44.4</td>
<td></td>
</tr>
<tr>
<td>Contig23001</td>
<td>Chitin_binding_1</td>
<td>993</td>
<td>1104</td>
<td>CAP65615</td>
<td></td>
<td></td>
<td>1.00E-25</td>
<td>267</td>
<td>28.8</td>
<td></td>
</tr>
<tr>
<td>Contig907</td>
<td>Chitin_binding_1</td>
<td>2282</td>
<td>3332</td>
<td>EEY22269</td>
<td></td>
<td></td>
<td>9.00E-21</td>
<td>113</td>
<td>46.9</td>
<td></td>
</tr>
<tr>
<td>Contig5885</td>
<td>Chitin_binding_1</td>
<td>639</td>
<td>151</td>
<td>XP_001558</td>
<td>442</td>
<td></td>
<td>8.00E-10</td>
<td>93</td>
<td>35.5</td>
<td></td>
</tr>
<tr>
<td>Contig2347</td>
<td>Chitin_binding_1</td>
<td>644</td>
<td>568</td>
<td>XP_001904</td>
<td>882</td>
<td></td>
<td>8.00E-26</td>
<td>177</td>
<td>38.4</td>
<td></td>
</tr>
<tr>
<td>Contig628</td>
<td>Chitin_binding_1</td>
<td>2145</td>
<td>2223</td>
<td>EEY21089</td>
<td></td>
<td></td>
<td>1.00E-19</td>
<td>124</td>
<td>44.4</td>
<td></td>
</tr>
<tr>
<td>Contig26500</td>
<td>Chitin_binding_1</td>
<td>551</td>
<td>253</td>
<td>XP_001560</td>
<td>867</td>
<td></td>
<td>2.00E-14</td>
<td>86</td>
<td>45.3</td>
<td></td>
</tr>
<tr>
<td>Contig13714</td>
<td>Chitin_binding_1</td>
<td>789</td>
<td>286</td>
<td>ABK58015</td>
<td></td>
<td></td>
<td>2.00E-19</td>
<td>133</td>
<td>39.8</td>
<td></td>
</tr>
<tr>
<td>Contig22529</td>
<td>Chitin_binding_1</td>
<td>1224</td>
<td>701</td>
<td>BAI44126</td>
<td></td>
<td></td>
<td>3.00E-25</td>
<td>204</td>
<td>30.4</td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Chitin_bind_1</td>
<td>Start</td>
<td>End</td>
<td>GI</td>
<td>Ref</td>
<td>E-value</td>
<td>Score</td>
<td>Similarity</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------------</td>
<td>-------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>---------</td>
<td>-------</td>
<td>------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Contig21374</td>
<td>Chitin_bind_1</td>
<td>622</td>
<td>234</td>
<td>XP_001904882</td>
<td>gi</td>
<td>171679872</td>
<td>ref</td>
<td>XP_001904882.1</td>
<td>2.00E-14</td>
<td>202 30.7 unnamed protein product [Podospora anserina]</td>
</tr>
<tr>
<td>Contig2590</td>
<td>Chitin_bind_1</td>
<td>712</td>
<td>556</td>
<td>EEY22269</td>
<td>gi</td>
<td>261359841</td>
<td>gb</td>
<td>EEY22269.1</td>
<td>4.00E-26</td>
<td>188 35.6 lectin [Verticillium albo-atrum VaMs.102]</td>
</tr>
<tr>
<td>Contig2837</td>
<td>Chitin_bind_1</td>
<td>968</td>
<td>102</td>
<td>XP_001598553</td>
<td>gi</td>
<td>156065263</td>
<td>ref</td>
<td>XP_001598553.1</td>
<td>5.00E-38</td>
<td>347 32.0 hypothetical protein SS1G_00642 [Sclerotinia sclerotiorum 1980]</td>
</tr>
<tr>
<td>Contig4943</td>
<td>Chitin_bind_1</td>
<td>640</td>
<td>220</td>
<td>XP_384927</td>
<td>gi</td>
<td>46119101</td>
<td>ref</td>
<td>XP_384927.1</td>
<td>1.00E-33</td>
<td>184 42.4 hypothetical protein FG04751.1 [Gibberella zeae PH-1]</td>
</tr>
<tr>
<td>Contig4943</td>
<td>Chitin_bind_1</td>
<td>640</td>
<td>220</td>
<td>XP_384927</td>
<td>gi</td>
<td>46119101</td>
<td>ref</td>
<td>XP_384927.1</td>
<td>1.00E-28</td>
<td>165 41.8 hypothetical protein FG04751.1 [Gibberella zeae PH-1]</td>
</tr>
<tr>
<td>Contig22543</td>
<td>Chitin_bind_1</td>
<td>2138</td>
<td>1547</td>
<td>ZP_02181401</td>
<td>gi</td>
<td>163786953</td>
<td>ref</td>
<td>ZP_02181401.1</td>
<td>4.00E-17</td>
<td>432 22.9 hypothetical protein FBALC1_17247 [Flavobacteriales bacterium ALC-1]</td>
</tr>
<tr>
<td>Contig29996</td>
<td>Chitin_bind_1</td>
<td>720</td>
<td>318</td>
<td>EEY22269</td>
<td>gi</td>
<td>261359841</td>
<td>gb</td>
<td>EEY22269.1</td>
<td>2.00E-15</td>
<td>113 33.6 lectin [Verticillium albo-atrum VaMs.102]</td>
</tr>
<tr>
<td>Contig5477</td>
<td>Chitin_bind_1</td>
<td>826</td>
<td>659</td>
<td>XP_001933394</td>
<td>gi</td>
<td>189194111</td>
<td>ref</td>
<td>XP_001933394.1</td>
<td>3.00E-20</td>
<td>118 45.8 chitin binding protein [Pyrenophora tritici-repentis Pt-1C-BFP]</td>
</tr>
<tr>
<td>Contig6430</td>
<td>Chitin_bind_1</td>
<td>555</td>
<td>288</td>
<td>EEY21089</td>
<td>gi</td>
<td>261358661</td>
<td>gb</td>
<td>EEY21089.1</td>
<td>1.00E-125</td>
<td>40.8 chitinase [Verticillium albo-atrum VaMs.102]</td>
</tr>
<tr>
<td>Contig</td>
<td>Chitin_bind_1</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>Begins</td>
<td>Ends</td>
<td>Identity</td>
<td>E-value</td>
<td>Score</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>--------------</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>--------</td>
<td>------</td>
<td>----------</td>
<td>---------</td>
<td>-------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Contig30348</td>
<td>Chitin_bind_1</td>
<td>521</td>
<td>295</td>
<td>XP_384927</td>
<td>gb</td>
<td>EY21089 .1</td>
<td>-20</td>
<td>9.00E-14</td>
<td>176</td>
<td>29.5</td>
</tr>
<tr>
<td>Contig1125</td>
<td>Chitin_bind_1</td>
<td>707</td>
<td>765</td>
<td>ABK58015</td>
<td>gi</td>
<td>17937826</td>
<td>gb</td>
<td>ABK58015 .1</td>
<td>2.00E-20</td>
<td>103</td>
</tr>
<tr>
<td>Contig2064</td>
<td>Chitin_bind_1</td>
<td>1619</td>
<td>32</td>
<td>EY22269</td>
<td>gi</td>
<td>261359841</td>
<td>gb</td>
<td>EY22269 .1</td>
<td>3.00E-14</td>
<td>218</td>
</tr>
<tr>
<td>Contig21567</td>
<td>Chitin_bind_1</td>
<td>504</td>
<td>220</td>
<td>XP_001912</td>
<td>gi</td>
<td>171695512</td>
<td>ref</td>
<td>XP_001912680.1</td>
<td>4.00E-10</td>
<td>37</td>
</tr>
<tr>
<td>NODE_7530.1_length_476_cov_9.714286</td>
<td>Chitin_bind_1</td>
<td>520</td>
<td>447</td>
<td>EEU39765</td>
<td>gi</td>
<td>256726404</td>
<td>gb</td>
<td>EEU39765 .1</td>
<td>4.00E-18</td>
<td>114</td>
</tr>
<tr>
<td>Contig6241</td>
<td>Chitin_bind_1</td>
<td>739</td>
<td>306</td>
<td>BAH4126</td>
<td>gi</td>
<td>26027911</td>
<td>db</td>
<td>BAH4126 .1</td>
<td>5.00E-17</td>
<td>123</td>
</tr>
<tr>
<td>Contig3097</td>
<td>Chitin_bind_1</td>
<td>893</td>
<td>720</td>
<td>EY21089</td>
<td>gi</td>
<td>261358661</td>
<td>gb</td>
<td>EY21089 .1</td>
<td>3.00E-17</td>
<td>117</td>
</tr>
<tr>
<td>Contig3790</td>
<td>Chitin_bind_1</td>
<td>581</td>
<td>590</td>
<td>1ULK</td>
<td>gi</td>
<td>40898776</td>
<td>pdb</td>
<td>1ULK[A</td>
<td>9.00E-15</td>
<td>113</td>
</tr>
<tr>
<td>Contig23623</td>
<td>Chitin_bind_1</td>
<td>539</td>
<td>426</td>
<td>XP_002384</td>
<td>gi</td>
<td>238505728</td>
<td>ref</td>
<td>XP_002384073.1</td>
<td>4.00E-09</td>
<td>34</td>
</tr>
<tr>
<td>Contig23088</td>
<td>Chitin_bind_1</td>
<td>616</td>
<td>503</td>
<td>EY21089</td>
<td>gi</td>
<td>261358661</td>
<td>gb</td>
<td>EY21089 .1</td>
<td>6.00E-15</td>
<td>107</td>
</tr>
<tr>
<td>Contig4227</td>
<td>Chitin_bind_1</td>
<td>665</td>
<td>263</td>
<td>XP_361310</td>
<td>gi</td>
<td>39943546</td>
<td>ref</td>
<td>XP_361310 .1</td>
<td>2.00E-12</td>
<td>109</td>
</tr>
<tr>
<td>Contig</td>
<td>Chitin-binding_1</td>
<td>Start</td>
<td>End</td>
<td>Length</td>
<td>E-Value</td>
<td>Score</td>
<td>Length_1</td>
<td>Cover_1</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td>----------</td>
<td>---------</td>
<td>----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Contig29832</td>
<td>Chitin-binding_1</td>
<td>534</td>
<td>867</td>
<td>333</td>
<td>3.00E-16</td>
<td>121</td>
<td>40.5</td>
<td>15</td>
<td>lectin [Verticillium albo-atrum VaM.s.102]</td>
<td></td>
</tr>
<tr>
<td>Contig17079</td>
<td>Chitin-binding_1</td>
<td>523</td>
<td>332</td>
<td>191</td>
<td>4.00E-19</td>
<td>171</td>
<td>30.4</td>
<td>15</td>
<td>lectin [Verticillium albo-atrum VaM.s.102]</td>
<td></td>
</tr>
<tr>
<td>Contig7906</td>
<td>Chitin-binding_1</td>
<td>521</td>
<td>234</td>
<td>287</td>
<td>4.00E-08</td>
<td>42</td>
<td>54.8</td>
<td>15</td>
<td>hypothetical protein MGG_12939 [Magnaporthe grisea 70-15]</td>
<td></td>
</tr>
<tr>
<td>Contig31423</td>
<td>Chitin-binding_1</td>
<td>1042</td>
<td>651</td>
<td>391</td>
<td>3.00E-16</td>
<td>191</td>
<td>30.4</td>
<td>15</td>
<td>hypothetical protein MGG_05351 [Magnaporthe grisea 70-15]</td>
<td></td>
</tr>
<tr>
<td>NODE_12804_length_1122_cov_11.425134</td>
<td>Chitin-binding_1</td>
<td>1166</td>
<td>3175</td>
<td>1209</td>
<td>1.00E-13</td>
<td>119</td>
<td>36.1</td>
<td>15</td>
<td>putative chitin deacetylase [Batrachochytrium dendrobatidis]</td>
<td></td>
</tr>
<tr>
<td>NODE_4383_length_1229_cov_11.374288</td>
<td>Chitin-binding_1</td>
<td>1281</td>
<td>2636</td>
<td>1355</td>
<td>3.00E-12</td>
<td>119</td>
<td>34.5</td>
<td>15</td>
<td>putative chitin deacetylase [Batrachochytrium dendrobatidis]</td>
<td></td>
</tr>
<tr>
<td>Contig1224</td>
<td>Chitin-binding_1</td>
<td>537</td>
<td>251</td>
<td>286</td>
<td>3.00E-11</td>
<td>137</td>
<td>30.7</td>
<td>15</td>
<td>hypothetical protein SNOG_07620 [Phaeosphaeria nodorum SN15]</td>
<td></td>
</tr>
<tr>
<td>Contig21417</td>
<td>Chitin-binding_1</td>
<td>540</td>
<td>676</td>
<td>136</td>
<td>6.00E-20</td>
<td>171</td>
<td>31.0</td>
<td>15</td>
<td>chitin binding protein 4 [Magnaporthe oryzae]</td>
<td></td>
</tr>
<tr>
<td>Contig1644</td>
<td>Chitin-binding_1</td>
<td>1065</td>
<td>64</td>
<td>1001</td>
<td>1.00E-10</td>
<td>277</td>
<td>23.8</td>
<td>15</td>
<td>AAR069Wp [Ashbya gossypii ATCC 10895] [AAR069Wp [Ashbya gossypii ATCC 10895]</td>
<td></td>
</tr>
<tr>
<td>Contig6589</td>
<td>Chitin-binding_1</td>
<td>581</td>
<td>308</td>
<td>273</td>
<td>9.00E-12</td>
<td>40</td>
<td>72.5</td>
<td>15</td>
<td>unnamed protein product [Podospora anserina]</td>
<td></td>
</tr>
<tr>
<td>Contig29186</td>
<td>Chitin-binding_1</td>
<td>1991</td>
<td>2982</td>
<td>991</td>
<td>2.00E-21</td>
<td>431</td>
<td>27.1</td>
<td>15</td>
<td>predicted protein [Mollicutes bacterium D7] [predicted</td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Chitin_bind_1</td>
<td>506</td>
<td>166</td>
<td>Q01MB6</td>
<td>gi</td>
<td>52013345</td>
<td>sp</td>
<td>Q01MB6.2</td>
<td>AGI ORYSI</td>
<td>6.00E-20</td>
</tr>
<tr>
<td>----------</td>
<td>---------------</td>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>----------</td>
<td>-------------------</td>
<td>--------</td>
<td>-----</td>
</tr>
<tr>
<td>Contig244</td>
<td>Chitin_bind_1</td>
<td>641</td>
<td>355</td>
<td>ABK58015</td>
<td>gi</td>
<td>17937826</td>
<td>gb</td>
<td>ABK58015 .1</td>
<td></td>
<td>4.00E-20</td>
</tr>
<tr>
<td>Contig20861</td>
<td>Chitin_bind_1</td>
<td>652</td>
<td>424</td>
<td>XP_001792 460</td>
<td>gi</td>
<td>169598073</td>
<td>ref</td>
<td>XP_001792 2460.1</td>
<td></td>
<td>2.00E-14</td>
</tr>
<tr>
<td>Contig2203</td>
<td>Chitin_bind_1</td>
<td>559</td>
<td>147</td>
<td>BAH44115</td>
<td>gi</td>
<td>260279809</td>
<td>dbj</td>
<td>BAH44115 .1</td>
<td></td>
<td>5.00E-12</td>
</tr>
<tr>
<td>Contig5517</td>
<td>Chitin_bind_1</td>
<td>539</td>
<td>353</td>
<td>XP_002384 073</td>
<td>gi</td>
<td>238505728</td>
<td>ref</td>
<td>XP_002384 4073.1</td>
<td></td>
<td>2.00E-10</td>
</tr>
<tr>
<td>Contig29243</td>
<td>Chitin_bind_1</td>
<td>2286</td>
<td>2254</td>
<td>EDP52952</td>
<td>gi</td>
<td>159127837</td>
<td>gb</td>
<td>EDP52952 .1</td>
<td></td>
<td>1.00E-07</td>
</tr>
<tr>
<td>Contig2638</td>
<td>Chitin_bind_1</td>
<td>904</td>
<td>72</td>
<td>XP_001912 680</td>
<td>gi</td>
<td>171695512</td>
<td>ref</td>
<td>XP_001912 2680.1</td>
<td></td>
<td>2.00E-09</td>
</tr>
<tr>
<td>Contig29854</td>
<td>CBM20</td>
<td>1425</td>
<td>1653</td>
<td>ZP_018529 68</td>
<td>gi</td>
<td>49174341</td>
<td>ref</td>
<td>ZP_01852 968.1</td>
<td></td>
<td>4.00E-09</td>
</tr>
<tr>
<td>Contig ID</td>
<td>Feature Type</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------------</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig4424</td>
<td>F5_F8_type_C</td>
<td>595</td>
<td>660</td>
<td>YP_003123</td>
<td>468 glycoside hydrolase family 18 [Chitinophaga pinensis DSM 2588] gi</td>
<td>256422815</td>
<td>ref</td>
<td>YP_003123468.1</td>
<td>6.00E -17 117 41.9 glycoside hydrolase family 18 [Chitinophaga pinensis DSM 2588]</td>
<td></td>
</tr>
<tr>
<td>Contig9433</td>
<td>CBM48</td>
<td>550</td>
<td>267</td>
<td>EEQ33998</td>
<td>1.00E -39 142 54.9 1,4-alpha-glucan branching enzyme [Microsporum canis CBS 113480]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig8017</td>
<td>LysM</td>
<td>762</td>
<td>301</td>
<td>ZP_06250496</td>
<td>ErfK/YhiS/Ycfs/YuhG family protein [Clostridium thermocellum JW20] gi</td>
<td>281419482</td>
<td>ref</td>
<td>ZP_06250496.1</td>
<td>0.001 175 29.1 ErfK/YhiS/Ycfs/YuhG family protein [Clostridium thermocellum JW20]</td>
<td></td>
</tr>
<tr>
<td>Contig731</td>
<td>LysM</td>
<td>827</td>
<td>286</td>
<td>ZP_05853124</td>
<td>N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583] gi</td>
<td>260587211</td>
<td>ref</td>
<td>ZP_05853124.1</td>
<td>7.00E -33 165 44.8 N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583]</td>
<td></td>
</tr>
<tr>
<td>Contig29178</td>
<td>LysM</td>
<td>1118</td>
<td>2115</td>
<td>YP_003432889</td>
<td>lipoprotein [Hydrogenobacter thermophilus TK-6] gi</td>
<td>288818541</td>
<td>ref</td>
<td>YP_003432889.1</td>
<td>0.026 112 26.8 lipoprotein [Hydrogenobacter thermophilus TK-6]</td>
<td></td>
</tr>
<tr>
<td>Contig23026</td>
<td>LysM</td>
<td>712</td>
<td>538</td>
<td>XP_001310466</td>
<td>hypothetical protein [Trichomonas vaginalis G3] gi</td>
<td>12343988</td>
<td>ref</td>
<td>XP_001310466.1</td>
<td>1.00E -63 168 66.7 hypothetical protein [Trichomonas vaginalis G3]</td>
<td></td>
</tr>
<tr>
<td>Contig29982</td>
<td>LysM</td>
<td>805</td>
<td>2273</td>
<td>ZP_05853124</td>
<td>N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583] gi</td>
<td>260587211</td>
<td>ref</td>
<td>ZP_05853124.1</td>
<td>2.00E -33 166 45.2 N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583]</td>
<td></td>
</tr>
<tr>
<td>Contig21424</td>
<td>LysM</td>
<td>672</td>
<td>6275</td>
<td>ZP_05853124</td>
<td>N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583] gi</td>
<td>260587211</td>
<td>ref</td>
<td>ZP_05853124.1</td>
<td>7.00E -23 139 45.3 N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583]</td>
<td></td>
</tr>
<tr>
<td>Contig2890</td>
<td>LysM</td>
<td>752</td>
<td>1330</td>
<td>ZP_05853124</td>
<td>N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583] gi</td>
<td>260587211</td>
<td>ref</td>
<td>ZP_05853124.1</td>
<td>1.00E -35 165 49.7 N-acetylmuramoyl-L-alanine amidase [Blaeutia hanseni DSM 20583]</td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>LysM</td>
<td>567</td>
<td>694</td>
<td>EFE40484</td>
<td>gi</td>
<td>291184982</td>
<td>gb</td>
<td>EFE40484.1</td>
<td>6.00E-04</td>
<td>99 28.3</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-----</td>
<td>-----</td>
<td>----------</td>
<td>--------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1569</td>
<td>LysM</td>
<td>751</td>
<td>1094</td>
<td>ZP_058531</td>
<td>24</td>
<td>gi</td>
<td>260587211</td>
<td>ref</td>
<td>ZP_05853 124.1</td>
<td>1.00E-34</td>
</tr>
<tr>
<td>Contig30786</td>
<td>LysM</td>
<td>621</td>
<td>3701</td>
<td>ZP_037163</td>
<td>06</td>
<td>gi</td>
<td>225027114</td>
<td>ref</td>
<td>ZP_03716 306.1</td>
<td>0.058</td>
</tr>
<tr>
<td>Contig8327</td>
<td>LysM</td>
<td>516</td>
<td>245</td>
<td>XP_001883</td>
<td>065</td>
<td>gi</td>
<td>70103701</td>
<td>ref</td>
<td>XP_00188 306.1</td>
<td>3.00E-04</td>
</tr>
<tr>
<td>NODE_10482</td>
<td>Polysacc_deac_1</td>
<td>752</td>
<td>371</td>
<td>ZP_053490</td>
<td>62</td>
<td>gi</td>
<td>255284507</td>
<td>ref</td>
<td>ZP_05349 062.1</td>
<td>5.00E-27</td>
</tr>
<tr>
<td>Contig30430</td>
<td>Polysacc_deac_1</td>
<td>628</td>
<td>412</td>
<td>EFE44609</td>
<td>gi</td>
<td>291189315</td>
<td>gb</td>
<td>EFE44609.1</td>
<td>1.00E-24</td>
<td>212 34.0</td>
</tr>
<tr>
<td>Contig14819</td>
<td>Polysacc_deac_1</td>
<td>641</td>
<td>273</td>
<td>ZP_063479</td>
<td>96</td>
<td>gi</td>
<td>283798843</td>
<td>ref</td>
<td>ZP_06347 996.1</td>
<td>9.00E-22</td>
</tr>
<tr>
<td>Contig21423</td>
<td>Polysacc_deac_1</td>
<td>1445</td>
<td>2355</td>
<td>XP_359754</td>
<td>47</td>
<td>gi</td>
<td>399404344</td>
<td>ref</td>
<td>XP_359754 996.1</td>
<td>4.00E-26</td>
</tr>
<tr>
<td>Contig15114</td>
<td>Polysacc_deac_1</td>
<td>988</td>
<td>632</td>
<td>ACF22100</td>
<td>gi</td>
<td>193804915</td>
<td>gb</td>
<td>ACF22100.1</td>
<td>2.00E-28</td>
<td>222 37.4</td>
</tr>
<tr>
<td>Contig31473</td>
<td>Polysacc_deac_1</td>
<td>651</td>
<td>452</td>
<td>CAQ16203</td>
<td>gi</td>
<td>209570274</td>
<td>emb</td>
<td>CAQ162 03.1</td>
<td>5.00E-26</td>
<td>203 36.0</td>
</tr>
<tr>
<td>Contig</td>
<td>Polysacc_deac_1</td>
<td>Length</td>
<td>E-values</td>
<td>gi</td>
<td>Ref</td>
<td>Product Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>--------</td>
<td>----------</td>
<td>----</td>
<td>-----</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7842</td>
<td></td>
<td>570</td>
<td>167</td>
<td>gi [261357640] gb [EEY20068]</td>
<td></td>
<td>chitin deacetylase [Verticillium albo-atrum VaMs.102]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23048</td>
<td></td>
<td>955</td>
<td>893</td>
<td>gi [261358935] gb [EEY21363]</td>
<td></td>
<td>chitin deacetylase [Verticillium albo-atrum VaMs.102]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_45052_length_553_cov_5.039783</td>
<td>Polysacc_deac_1</td>
<td>589</td>
<td>217</td>
<td>gi [209570274] emb [CAQ16203]</td>
<td></td>
<td>hypothetical protein [Glomerella graminicola]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_7121_length_901_cov_10.762486</td>
<td>Polysacc_deac_1</td>
<td>945</td>
<td>1563</td>
<td>gi [46124343] ref [XP_386725]</td>
<td></td>
<td>hypothetical protein FG06549.1 [Gibberella zeae PH-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1097</td>
<td></td>
<td>668</td>
<td>250</td>
<td>gi [112491417] pdb [2IW0]</td>
<td></td>
<td>Chain A, Structure Of The Chitin Deacetylase From The Fungal Pathogen Colletotrichum Lindemuthianum</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22718</td>
<td></td>
<td>941</td>
<td>330</td>
<td>gi [169850453] ref [XP_001831922]</td>
<td></td>
<td>hypothetical protein CC1G_12897 [Coprinopsis cinerea okayama7#130]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Polysacc_deac_1</td>
<td>Start</td>
<td>End</td>
<td>gi/strain</td>
<td>gb/strain</td>
<td>E-value</td>
<td>Bit Score</td>
<td>Function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>-----------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig25338</td>
<td>Polysacc_deac_1</td>
<td>761</td>
<td>469</td>
<td>gi[261357640]</td>
<td>gb[EEY20068.1]</td>
<td>7.00E-26</td>
<td>228</td>
<td>35.1</td>
<td>chitin binding protein [Pyrenophora tritici-repentis Pt-1C-BFP]</td>
<td></td>
</tr>
<tr>
<td>Contig25066</td>
<td>Polysacc_deac_1</td>
<td>804</td>
<td>379</td>
<td>XP_363190</td>
<td>gi[39949397]</td>
<td>5.00E-20</td>
<td>211</td>
<td>31.8</td>
<td>hypothetical protein MGG_08774 [Magnaporthe grisea 70-15]</td>
<td></td>
</tr>
<tr>
<td>Contig22541</td>
<td>Polysacc_deac_1</td>
<td>836</td>
<td>596</td>
<td>EEU39765</td>
<td>gi[256726404]</td>
<td>4.00E-17</td>
<td>248</td>
<td>32.3</td>
<td>hypothetical protein NECHADRAFT_93052 [Nectria haematococca mpVI 77-13-4]</td>
<td></td>
</tr>
<tr>
<td>Contig29997</td>
<td>Polysacc_deac_1</td>
<td>977</td>
<td>587</td>
<td>XP_001268141</td>
<td>gi[121699752]</td>
<td>2.00E-17</td>
<td>165</td>
<td>34.5</td>
<td>chitin deacetylase, putative [Aspergillus clavatus NRRL 1]</td>
<td></td>
</tr>
<tr>
<td>Contig28783</td>
<td>Polysacc_deac_1</td>
<td>1109</td>
<td>1716</td>
<td>XP_958437</td>
<td>gi[164422949]</td>
<td>8.00E-25</td>
<td>232</td>
<td>36.2</td>
<td>hypothetical protein NCU09508 [Neurospora crassa OR74A]</td>
<td></td>
</tr>
<tr>
<td>Contig23626</td>
<td>Polysacc_deac_1</td>
<td>889</td>
<td>813</td>
<td>EEY20068</td>
<td>gi[261357640]</td>
<td>4.00E-24</td>
<td>220</td>
<td>35.0</td>
<td>chitin deacetylase [Verticillium albo-atrum VaMs.102]</td>
<td></td>
</tr>
<tr>
<td>Contig931</td>
<td>Polysacc_deac_1</td>
<td>805</td>
<td>830</td>
<td>EEU19684</td>
<td>gi[261357256]</td>
<td>5.00E-17</td>
<td>238</td>
<td>29.8</td>
<td>polysaccharide deacetylase family protein [Verticillium albo-atrum VaMs.102]</td>
<td></td>
</tr>
<tr>
<td>Contig7835</td>
<td>Polysacc_deac_1</td>
<td>1253</td>
<td>522</td>
<td>XP_002374100</td>
<td>gi[238485724]</td>
<td>3.00E-18</td>
<td>253</td>
<td>29.2</td>
<td>chitin deacetylase, putative [Aspergillus flavus NRRL3357]</td>
<td></td>
</tr>
<tr>
<td>Contig21498</td>
<td>Polysacc_deac_1</td>
<td>981</td>
<td>2424</td>
<td>YP_677852</td>
<td>gi[110637645]</td>
<td>3.00E-55</td>
<td>313</td>
<td>42.2</td>
<td>bifunctional xylanase/esterase; CBM9 module, glycoside hydrolase family 8 protein and carbohydrate esterase family 4 protein [Cytophaga hutchinsonii ATCC 33406]</td>
<td></td>
</tr>
<tr>
<td>Contig2002</td>
<td>Polysacc_deac_1</td>
<td>619</td>
<td>595</td>
<td>XP_000223740</td>
<td>gi[110637645]</td>
<td>3.00E-55</td>
<td>313</td>
<td>42.2</td>
<td>bifunctional xylanase/esterase; CBM9 module, glycoside hydrolase family 8 protein and carbohydrate esterase family 4 protein [Cytophaga hutchinsonii ATCC 33406]</td>
<td></td>
</tr>
<tr>
<td>Contig25462</td>
<td>Polysacc_deac_1</td>
<td>918</td>
<td>621</td>
<td>XP_005096500</td>
<td>gi[1564422949]</td>
<td>8.00E-25</td>
<td>232</td>
<td>36.2</td>
<td>hypothetical protein NCU09508 [Neurospora crassa OR74A]</td>
<td></td>
</tr>
<tr>
<td>Contig2002</td>
<td>Polysacc_deac_1</td>
<td>619</td>
<td>595</td>
<td>XP_000223740</td>
<td>gi[110637645]</td>
<td>3.00E-55</td>
<td>313</td>
<td>42.2</td>
<td>bifunctional xylanase/esterase; CBM9 module, glycoside hydrolase family 8 protein and carbohydrate esterase family 4 protein [Cytophaga hutchinsonii ATCC 33406]</td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Polysacc_deac_1</td>
<td>Start</td>
<td>End</td>
<td>gi/Ref</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>-------</td>
<td>-------</td>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8991</td>
<td></td>
<td>940</td>
<td>1944</td>
<td>gi[261357640] gb[EEY20068.1]</td>
<td>2.00E-24 227 33.0 chitin deacetylase [Verticillum albo-atrum VaMs.102]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1781</td>
<td></td>
<td>990</td>
<td>1108</td>
<td>gi[110637645] ref[YP_67785 2.1]</td>
<td>1.00E-52 312 40.1 bifunctional xylanase/esterase; CBM9 module, glycoside hydrolase family 8 protein and carbohydrate esterase family 4 protein [Cytophaga hutchinsonii ATCC 33406]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4101</td>
<td></td>
<td>800</td>
<td>337</td>
<td>gi[164422983] ref[XP_95969 3.2]</td>
<td>2.00E-18 160 32.5 hypothetical protein NCU09582 [Neurospora crassa OR74A]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10063</td>
<td></td>
<td>679</td>
<td>324</td>
<td>gi[261357640] gb[EEY20068.1]</td>
<td>1.00E-12 230 27.0 chitin deacetylase [Verticillum albo-atrum VaMs.102]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24262</td>
<td></td>
<td>536</td>
<td>186</td>
<td>gi[284800701] ref[YP_003412 566.1]</td>
<td>1.00E-08 133 29.3 hypothetical protein LM5578_0448 [Listeria monocytogenes 08-5578]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4916</td>
<td></td>
<td>967</td>
<td>612</td>
<td>gi[261416362] ref[YP_003250 0045.1]</td>
<td>1.00E-155 327 79.5 polysaccharide deacetylase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22653</td>
<td></td>
<td>643</td>
<td>386</td>
<td>gi[169617349] ref[XP_001802 089]</td>
<td>1.00E-24 225 32.0 hypothetical protein SNOG_11852 [Phaeosphaeria nodorum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Description</td>
<td>Accession</td>
<td>Length</td>
<td>Identity</td>
<td>E-value</td>
<td>P-value</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------------</td>
<td>-------------</td>
<td>--------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig426</td>
<td>Pectinesterase</td>
<td>ZP_067188 52</td>
<td>1124</td>
<td>63.2</td>
<td>1.00E-129</td>
<td>367</td>
<td>hypothetical protein SNOG_11852 [Phaeosphaeria nodorum SN15]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21578</td>
<td>Pectinesterase</td>
<td>ZP_061438 61</td>
<td>1218</td>
<td>63.2</td>
<td>1.00E-113</td>
<td>372</td>
<td>Pectate lyase/Amb allergen [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig7215</td>
<td>Pectinesterase</td>
<td>ZP_061438 61</td>
<td>1074</td>
<td>65.4</td>
<td>1.00E-133</td>
<td>355</td>
<td>Pectate lyase/Amb allergen [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2434</td>
<td>Amidohydrolase 1</td>
<td>XP_001580 549</td>
<td>1371</td>
<td>70.7</td>
<td>1.00E-67</td>
<td>451</td>
<td>Amidohydrolase family protein [Trichomonas vaginalis G3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig25221</td>
<td>Amidohydrolase 1</td>
<td>XP_001580 549</td>
<td>777</td>
<td>51.8</td>
<td>1.00E-125</td>
<td>255</td>
<td>Amidohydrolase family protein [Trichomonas vaginalis G3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig9584</td>
<td>Amidohydrolase 1</td>
<td>XP_001580 549</td>
<td>501</td>
<td>81.2</td>
<td>4.00E-73</td>
<td>165</td>
<td>Amidohydrolase family protein [Trichomonas vaginalis G3]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22339*</td>
<td>Pectate lyase C</td>
<td>ZP_061422 60</td>
<td>1515</td>
<td>51.0</td>
<td>1.00E-120</td>
<td>404</td>
<td>Pectate lyase/Amb allergen [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig828</td>
<td>Pectate lyase C</td>
<td>ZP_067178 64</td>
<td>1837</td>
<td>51.9</td>
<td>1.00E-156</td>
<td>613</td>
<td>pectate lyase [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3060</td>
<td>Pectate lyase C</td>
<td>YP_002487 258</td>
<td>671</td>
<td>37.6</td>
<td>2.00E-26</td>
<td>197</td>
<td>Pectate lyase/Amb allergen [Arthrobacter chlorophenolicus A6]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig16158</td>
<td>Pectate lyase C</td>
<td>YP_003243 122</td>
<td>547</td>
<td>48.9</td>
<td>6.00E-39</td>
<td>178</td>
<td>Pectate lyase/Amb allergen [Geobacillus sp. Y412MC10]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6104</td>
<td>Pectate lyase C</td>
<td>CBL17651</td>
<td>604</td>
<td>37.2</td>
<td>4.00E-29</td>
<td>191</td>
<td>Pectate lyase [Ruminococcus sp. 18P13]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Protein Name</td>
<td>Length</td>
<td>Start Position</td>
<td>Accession</td>
<td>E-Value</td>
<td>ID</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
<td>--------</td>
<td>----------------</td>
<td>-----------</td>
<td>---------</td>
<td>----</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30471</td>
<td>Pec_lyase e_C</td>
<td>518</td>
<td>285</td>
<td>YP_003243 122</td>
<td>1.00E -14</td>
<td>110 36.4</td>
<td>Pectate lyase/Amb allergen [Geobacillus sp. Y412MC10]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2859</td>
<td>Pectate_1yase</td>
<td>795</td>
<td>534</td>
<td>XP_001554 464</td>
<td>8.00E -52</td>
<td>219 51.1</td>
<td>hypothetical protein BC1G_07052 [Botryotinia fuckeliana B05.10]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig28811</td>
<td>Pectate_1yase</td>
<td>828</td>
<td>4010</td>
<td>XP_001554 464</td>
<td>3.00E -48</td>
<td>233 48.1</td>
<td>hypothetical protein BC1G_07052 [Botryotinia fuckeliana B05.10]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig20357</td>
<td>Pectate_1yase</td>
<td>800</td>
<td>5711</td>
<td>XP_001554 464</td>
<td>3.00E -50</td>
<td>223 50.2</td>
<td>hypothetical protein BC1G_07052 [Botryotinia fuckeliana B05.10]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2492</td>
<td>Pectate_1yase</td>
<td>526</td>
<td>585</td>
<td>XP_001215 463</td>
<td>9.00E -30</td>
<td>146 50.0</td>
<td>hypothetical protein ATEG_06285 [Aspergillus terreus NIH2624]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21247</td>
<td>Alpha-L-ARF_C</td>
<td>1382</td>
<td>873</td>
<td>CBL090557</td>
<td>1.00E -127</td>
<td>466 49.1</td>
<td>Alpha-L-arabinofuranosidase [Roseburia intestinalis M50/1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig8415</td>
<td>Bgal_small_N</td>
<td>721</td>
<td>219</td>
<td>ZP_061423 36</td>
<td>1.00E -91</td>
<td>236 68.2</td>
<td>beta-galactosidase [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3630</td>
<td>Bgal_small_N</td>
<td>832</td>
<td>356</td>
<td>ZP_061423 36</td>
<td>1.00E -96</td>
<td>264 64.4</td>
<td>beta-galactosidase [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig10124</td>
<td>Bgal_small_N</td>
<td>703</td>
<td>261</td>
<td>ZP_061423 36</td>
<td>2.00E -70</td>
<td>236 55.9</td>
<td>beta-galactosidase [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig10613</td>
<td>GDE_C</td>
<td>3641</td>
<td>3096</td>
<td>XP_001912 837</td>
<td>0</td>
<td>1238 42.5</td>
<td>unnamed protein product [Podospora anserina]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Contig</th>
<th>Gene Family</th>
<th>Accession</th>
<th>% Identity</th>
<th>E Value</th>
<th>Coverage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig24123</td>
<td>GDE_C</td>
<td>XP_001646218</td>
<td>1.00E-126</td>
<td>463</td>
<td>48.6</td>
<td>hypothetical protein Kpol_1013p31 [Vanderwaltozyma polyspora DSM 70294]</td>
</tr>
<tr>
<td>Contig22750</td>
<td>GDE_C</td>
<td>XP_567389</td>
<td>2.00E-49</td>
<td>203</td>
<td>48.8</td>
<td>Glycogen debranching enzyme [Cryptococcus neoformans var. neoformans JEC21]</td>
</tr>
<tr>
<td>Contig864</td>
<td>GH2N</td>
<td>ZP_06142336</td>
<td>3.00E-91</td>
<td>216</td>
<td>72.7</td>
<td>beta-galactosidase [Ruminococcus flavefaciens FD-1]</td>
</tr>
<tr>
<td>Contig29590</td>
<td>GH2N</td>
<td>ZP_06721189</td>
<td>4.00E-70</td>
<td>216</td>
<td>57.9</td>
<td>glycosyl hydrolase family 2, sugar binding domain protein [Bacteroides ovatus SD CC 2a]</td>
</tr>
<tr>
<td>Contig9801</td>
<td>GH2N</td>
<td>ZP_06142336</td>
<td>4.00E-90</td>
<td>228</td>
<td>69.7</td>
<td>beta-galactosidase [Ruminococcus flavefaciens FD-1]</td>
</tr>
<tr>
<td>Contig4594</td>
<td>GH2N</td>
<td>YP_206064</td>
<td>1.00E-123</td>
<td>558</td>
<td>43.2</td>
<td>beta-mannosidase [Vibrio fischeri ES114]</td>
</tr>
<tr>
<td>Contig23341</td>
<td>GH3C</td>
<td>AAO41704</td>
<td>1.00E-90</td>
<td>302</td>
<td>56.0</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig1978</td>
<td>GH3C</td>
<td>AAO41704</td>
<td>5.00E-50</td>
<td>251</td>
<td>42.6</td>
<td>beta-glucosidase precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>----</td>
<td>-----------</td>
<td>gil</td>
<td>28557461</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>291538316</td>
<td>emb</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>76365700</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>GH3C</td>
<td></td>
<td>AAO41704</td>
<td>gi</td>
<td>291538316</td>
<td>emb</td>
</tr>
<tr>
<td>Contig</td>
<td>TIG</td>
<td></td>
<td>XP_002293728</td>
<td>gi</td>
<td>224009540</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>TIG</td>
<td></td>
<td>XP_002293728</td>
<td>gi</td>
<td>224009540</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>TIG</td>
<td></td>
<td>EEY69233</td>
<td>gi</td>
<td>262111181</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>TIG</td>
<td></td>
<td>EER11220</td>
<td>gi</td>
<td>239888533</td>
<td>gb</td>
</tr>
<tr>
<td>Contig</td>
<td>DUF297</td>
<td>667</td>
<td>51</td>
<td>EDK40790</td>
<td>gi</td>
<td>190348350</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>-----</td>
<td>----</td>
<td>----------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Contig11687</td>
<td>DUF297</td>
<td>627</td>
<td>243</td>
<td>EDK40790</td>
<td>gi</td>
<td>190348350</td>
</tr>
<tr>
<td>Contig21765</td>
<td>DUF297</td>
<td>1250</td>
<td>511</td>
<td>ZP_011133 06</td>
<td>gi</td>
<td>88797718</td>
</tr>
<tr>
<td>Contig260</td>
<td>DUF297</td>
<td>1229</td>
<td>562</td>
<td>ZP_012165 73</td>
<td>gi</td>
<td>90408411</td>
</tr>
<tr>
<td>Contig5652</td>
<td>DUF297</td>
<td>783</td>
<td>342</td>
<td>YP_003503 987</td>
<td>gi</td>
<td>291287171</td>
</tr>
<tr>
<td>Contig3727</td>
<td>DUF297</td>
<td>513</td>
<td>186</td>
<td>YP_340939</td>
<td>gi</td>
<td>77361364</td>
</tr>
<tr>
<td>Contig8184</td>
<td>DUF297</td>
<td>938</td>
<td>361</td>
<td>YP_156978</td>
<td>gi</td>
<td>56461697</td>
</tr>
<tr>
<td>Contig21383</td>
<td>DUF297</td>
<td>517</td>
<td>169</td>
<td>ZP_012165 73</td>
<td>gi</td>
<td>90408411</td>
</tr>
<tr>
<td>Contig12037</td>
<td>DUF297</td>
<td>673</td>
<td>570</td>
<td>YP_156978</td>
<td>gi</td>
<td>56461697</td>
</tr>
<tr>
<td>Contig16432</td>
<td>DUF297</td>
<td>543</td>
<td>199</td>
<td>YP_003503 987</td>
<td>gi</td>
<td>291287171</td>
</tr>
<tr>
<td>Contig17499</td>
<td>Esterase</td>
<td>1143</td>
<td>674</td>
<td>ACZ98648</td>
<td>gi</td>
<td>280977861</td>
</tr>
<tr>
<td>Contig1579</td>
<td>Esterase</td>
<td>906</td>
<td>27</td>
<td>ZP_061423 43</td>
<td>gi</td>
<td>268608616</td>
</tr>
<tr>
<td>Contig8173</td>
<td>Esterase</td>
<td>695</td>
<td>317</td>
<td>AAF70241</td>
<td>gi</td>
<td>7839348</td>
</tr>
<tr>
<td>Contig18676</td>
<td>Esterase</td>
<td>1525</td>
<td>1044</td>
<td>AAF70241</td>
<td>gi</td>
<td>7839348</td>
</tr>
<tr>
<td>Contig9000</td>
<td>Esterase</td>
<td>1201</td>
<td>727</td>
<td>AAF70241</td>
<td>gi</td>
<td>7839348</td>
</tr>
<tr>
<td>Contig2667</td>
<td>Esterase</td>
<td>1461</td>
<td>213</td>
<td>AAF70241</td>
<td>gi</td>
<td>7839348</td>
</tr>
<tr>
<td>Contig</td>
<td>ID</td>
<td>Length</td>
<td>Accession</td>
<td>Description</td>
<td>E-value</td>
<td>Score</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>-----------</td>
<td>--------------------------------------------------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>Contig3971</td>
<td>Esterase</td>
<td>622</td>
<td>ACZ98648</td>
<td>gi</td>
<td>280977861</td>
<td>gb</td>
</tr>
<tr>
<td>Contig21493</td>
<td>DUF303</td>
<td>1232</td>
<td>ZP_034585 34</td>
<td>gi</td>
<td>218129730</td>
<td>ref</td>
</tr>
<tr>
<td>Contig22188</td>
<td>DUF303</td>
<td>592</td>
<td>ZP_045458 78</td>
<td>gi</td>
<td>237715397</td>
<td>ref</td>
</tr>
<tr>
<td>Contig21894</td>
<td>DUF303</td>
<td>708</td>
<td>AAB69090</td>
<td>gi</td>
<td>2231243</td>
<td>gb</td>
</tr>
<tr>
<td>Contig541</td>
<td>DUF303</td>
<td>867</td>
<td>ZP_030166 04</td>
<td>gi</td>
<td>189467819</td>
<td>ref</td>
</tr>
<tr>
<td>Contig25491</td>
<td>DUF303</td>
<td>1118</td>
<td>AAB69090</td>
<td>gi</td>
<td>2231243</td>
<td>gb</td>
</tr>
<tr>
<td>Contig685</td>
<td>DUF303</td>
<td>801</td>
<td>ZP_045458 78</td>
<td>gi</td>
<td>237715397</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>Domain</td>
<td>Length</td>
<td>Identity</td>
<td>Accession</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>--------</td>
<td>----------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>30565</td>
<td>DUF303</td>
<td>865</td>
<td>1.00E-102</td>
<td>AAC14690</td>
<td>acetyl xylan esterase A [Orpinomyces sp. PC-2]</td>
<td></td>
</tr>
<tr>
<td>12482</td>
<td>DUF303</td>
<td>1063</td>
<td>3.00E-88</td>
<td>AAC14690</td>
<td>acetyl xylan esterase A [Orpinomyces sp. PC-2]</td>
<td></td>
</tr>
<tr>
<td>29194</td>
<td>DUF303</td>
<td>906</td>
<td>1.00E-116</td>
<td>AAC14690</td>
<td>acetyl xylan esterase A [Orpinomyces sp. PC-2]</td>
<td></td>
</tr>
<tr>
<td>22752</td>
<td>DUF303</td>
<td>972</td>
<td>1.00E-75</td>
<td>AAB69090</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
<td></td>
</tr>
<tr>
<td>3819</td>
<td>DUF303</td>
<td>605</td>
<td>3.00E-65</td>
<td>AAB69090</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
<td></td>
</tr>
<tr>
<td>29175</td>
<td>DUF303</td>
<td>544</td>
<td>2.00E-67</td>
<td>AAB69090</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
<td></td>
</tr>
<tr>
<td>840</td>
<td>COesterase</td>
<td>1670</td>
<td>1.00E-158</td>
<td>CBK73419</td>
<td>Carboxylesterase type B [Butyrivibrio fibrisolvens 16/4]</td>
<td></td>
</tr>
<tr>
<td>1868</td>
<td>COesterase</td>
<td>1029</td>
<td>1.00E-158</td>
<td>NP_149214</td>
<td>Para-nitrobenzyl esterase, a/b hydrolase [Clostridium acetobutylicum ATCC 824]</td>
<td></td>
</tr>
<tr>
<td>1174</td>
<td>COesterase</td>
<td>557</td>
<td>8.00E-22</td>
<td>ZP_054978</td>
<td>Alpha/beta hydrolase fold-3 domain protein [Clostridium papyrosolvens DSM 2782]</td>
<td></td>
</tr>
<tr>
<td>257</td>
<td>COesterase</td>
<td>743</td>
<td>2.00E-17</td>
<td>YP_079964</td>
<td>para-nitrobenzyl esterase (intracellular esterase B) [Bacillus licheniformis ATCC 14580]</td>
<td></td>
</tr>
</tbody>
</table>

*Note: The table lists various hypothetical proteins and conserved domains found in different bacterial species, along with their accession numbers and identities.*
<table>
<thead>
<tr>
<th>Contig</th>
<th>Pfam-B_184</th>
<th>Start</th>
<th>End</th>
<th>gi</th>
<th>ref</th>
<th>Protein Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig28300</td>
<td>Pfam-B_184</td>
<td>2733</td>
<td>1434</td>
<td>gi</td>
<td>ref</td>
<td>putative protein [Bacillus licheniformis ATCC 14580]</td>
</tr>
<tr>
<td>Contig7752</td>
<td>Pfam-B_184</td>
<td>2512</td>
<td>2505</td>
<td>gi</td>
<td>ref</td>
<td>conserved hypothetical protein [Roseburia intestinalis L1-82]</td>
</tr>
<tr>
<td>Contig29412</td>
<td>Pfam-B_184</td>
<td>1336</td>
<td>1514</td>
<td>gi</td>
<td>ref</td>
<td>conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]</td>
</tr>
<tr>
<td>Contig19696</td>
<td>Pfam-B_184</td>
<td>1048</td>
<td>476</td>
<td>gi</td>
<td>ref</td>
<td>conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]</td>
</tr>
<tr>
<td>Contig1848</td>
<td>Pfam-B_184</td>
<td>1400</td>
<td>62</td>
<td>gi</td>
<td>ref</td>
<td>conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]</td>
</tr>
<tr>
<td>Contig21105</td>
<td>Pfam-B_184</td>
<td>1501</td>
<td>635</td>
<td>gi</td>
<td>ref</td>
<td>conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]</td>
</tr>
<tr>
<td>Contig7431</td>
<td>Pfam-B_184</td>
<td>524</td>
<td>245</td>
<td>gi</td>
<td>ref</td>
<td>conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]</td>
</tr>
<tr>
<td>Contig31355</td>
<td>Pfam-B_184</td>
<td>541</td>
<td>392</td>
<td>gi</td>
<td>ref</td>
<td>conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]</td>
</tr>
<tr>
<td>Contig</td>
<td>Pfam-B</td>
<td>Start</td>
<td>End</td>
<td>ZP_ID</td>
<td>Evalue</td>
<td>BitScore</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>Contig25356 Pfam-B_184</td>
<td>516</td>
<td>177</td>
<td>ZP_054981</td>
<td>09</td>
<td>1.00E-33</td>
<td>170</td>
</tr>
<tr>
<td>Contig21626 Pfam-B_1434</td>
<td>1213</td>
<td>5338</td>
<td>ZP_020248</td>
<td>95</td>
<td>1.00E-124</td>
<td>374</td>
</tr>
<tr>
<td>Contig10848 Pfam-B_1434</td>
<td>1065</td>
<td>356</td>
<td>ZP_020248</td>
<td>95</td>
<td>1.00E-116</td>
<td>355</td>
</tr>
<tr>
<td>Contig5554 Pfam-B_2673</td>
<td>2133</td>
<td>873</td>
<td>AAB69091</td>
<td></td>
<td>2.00E-84</td>
<td>361</td>
</tr>
<tr>
<td>Contig1179 Pfam-B_2673</td>
<td>1395</td>
<td>64</td>
<td>ZP_061426</td>
<td>63</td>
<td>1.00E-124</td>
<td>462</td>
</tr>
<tr>
<td>Contig21582 Pfam-B_2673</td>
<td>817</td>
<td>503</td>
<td>AAB69091</td>
<td></td>
<td>2.00E-55</td>
<td>262</td>
</tr>
<tr>
<td>Contig23673 Pfam-B_2673</td>
<td>1259</td>
<td>734</td>
<td>AAB69091</td>
<td></td>
<td>1.00E-178</td>
<td>344</td>
</tr>
<tr>
<td>Contig25553 Pfam-B_2673</td>
<td>770</td>
<td>359</td>
<td>AAB69091</td>
<td></td>
<td>5.00E-80</td>
<td>357</td>
</tr>
<tr>
<td>Contig22864 Pfam-B_8046</td>
<td>761</td>
<td>712</td>
<td>ADE81126</td>
<td></td>
<td>3.00E-58</td>
<td>254</td>
</tr>
<tr>
<td>Contig11936 Pfam-B_8046</td>
<td>508</td>
<td>237</td>
<td>ADE81126</td>
<td></td>
<td>5.00E-40</td>
<td>194</td>
</tr>
<tr>
<td>Contig26846 Pfam-B_8046</td>
<td>700</td>
<td>554</td>
<td>XP_001829</td>
<td>175</td>
<td>2.00E-36</td>
<td>214</td>
</tr>
</tbody>
</table>

Conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]
<table>
<thead>
<tr>
<th>Contig</th>
<th>Pfam</th>
<th>Length</th>
<th>Contig</th>
<th>CE15</th>
<th>Length</th>
<th>Contig</th>
<th>PL6</th>
<th>Length</th>
<th>NODE</th>
<th>Length</th>
<th>Contig</th>
<th>PL11_C</th>
<th>Length</th>
<th>Contig</th>
<th>PL09</th>
<th>Length</th>
<th>Contig</th>
<th>PL09</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>285</td>
<td>Pfam-B_8046</td>
<td>522</td>
<td>252</td>
<td>XP_002394825</td>
<td>1.00E-28</td>
<td>175</td>
<td>41.1</td>
<td>hypothetical protein MPER_05224 [Moniliophthora perniciosa FA553] gi</td>
<td>215464484</td>
<td>gb</td>
<td>EEB95755.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7303</td>
<td></td>
<td></td>
<td>277</td>
<td>YP_67785</td>
<td>2.00E-86</td>
<td>259</td>
<td>60.6</td>
<td>beta-xylosidase/endoglucanase [Cytophaga hutchinsonii ATCC 33406] gi</td>
<td>110280324</td>
<td>gb</td>
<td>ABG58510.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11541</td>
<td></td>
<td></td>
<td>277</td>
<td>YP_67785</td>
<td>2.00E-79</td>
<td>252</td>
<td>56.7</td>
<td>beta-xylosidase/endoglucanase [Cytophaga hutchinsonii ATCC 33406] gi</td>
<td>110280324</td>
<td>gb</td>
<td>ABG58510.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29100</td>
<td>PL6</td>
<td></td>
<td>805</td>
<td>ZP_02862960</td>
<td>3.00E-08</td>
<td>194</td>
<td>27.8</td>
<td>hypothetical protein ANASTE_02192 [Anaerofustis stercoriomnins DSM 17244] gi</td>
<td>169258505</td>
<td>gb</td>
<td>EDS72471.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Parallel beta-helix repeat protein [Planctomyces maris DSM 8797] gi</td>
<td>148846886</td>
<td>gb</td>
<td>EDL61222.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>cellulosome protein dockerin type I [Clostridium papyrosolvens DSM 2782] gi</td>
<td>256744398</td>
<td>gb</td>
<td>EUL57552.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pectate lyase [Clostridium cellulovorans 743B] gi</td>
<td>24228465</td>
<td>gb</td>
<td>EUL31701.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Putative pectate lyase related protein, polysaccharide lyase family [Streptococcus galolyticus UCN34] gi</td>
<td>288732497</td>
<td>emb</td>
<td>CBI14069.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Putative pectate lyase related protein, polysaccharide lyase family [Streptococcus galolyticus UCN34] gi</td>
<td>288732497</td>
<td>emb</td>
<td>CBI14069.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Location</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig8180</td>
<td>PL09</td>
<td>1249 514 YP_003430 993</td>
<td>refYP_003430_993.1</td>
<td>-76</td>
<td>polysaccharide lyase family [Streptococcus gallolyticus UCN34]gi</td>
<td>288732497</td>
<td>emb</td>
<td>CBI14069.1</td>
<td>Putative pectate lyase related protein, polysaccharide lyase family [Streptococcus gallolyticus UCN34]gi</td>
<td>288905771</td>
<td>ref</td>
<td>YP_003430_993.1</td>
<td>2.00E-77</td>
<td>416 40.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6102</td>
<td>PL09</td>
<td>1275 723 YP_003430 993</td>
<td>refYP_003430_993.1</td>
<td>8.00E-75</td>
<td>400 39.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig13767</td>
<td>PL09</td>
<td>702 338 YP_003430 993</td>
<td>refYP_003430_993.1</td>
<td>3.00E-48</td>
<td>232 47.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22095</td>
<td>PL09</td>
<td>732 456 YP_003430 993</td>
<td>refYP_003430_993.1</td>
<td>1.00E-60</td>
<td>256 48.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig18967</td>
<td>PL09</td>
<td>523 174 YP_002505 335</td>
<td>refYP_002505_335.1</td>
<td>2.00E-30</td>
<td>136 55.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29157</td>
<td>Swollenin</td>
<td>1221 1032 CAB92328 535</td>
<td>gi</td>
<td>8052455</td>
<td>e mb</td>
<td>CAB92328</td>
<td>1.00E-131</td>
<td>388 56.7</td>
<td>swollenin [Hypocreajecorina]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Protein</td>
<td>Length</td>
<td>Identity</td>
<td>E-Value</td>
<td>Score</td>
<td>Prediction</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>--------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig4246</td>
<td>swollenin</td>
<td>722</td>
<td>641</td>
<td>BA83433</td>
<td>3.00E-77</td>
<td>231</td>
<td>59.3 swollenin like protein [Aspergillus fumigatus]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig31130</td>
<td>swollenin</td>
<td>689</td>
<td>371</td>
<td>ABV57767</td>
<td>2.00E-73</td>
<td>218</td>
<td>57.8 swollenin [Hypocrea pseudokoningii]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5172</td>
<td>swollenin</td>
<td>595</td>
<td>225</td>
<td>ABV57767</td>
<td>4.00E-59</td>
<td>183</td>
<td>59.0 swollenin [Hypocrea pseudokoningii]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22476</td>
<td>swollenin</td>
<td>509</td>
<td>365</td>
<td>ABV57767</td>
<td>6.00E-33</td>
<td>160</td>
<td>47.5 swollenin [Hypocrea pseudokoningii]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig14372</td>
<td>GH95</td>
<td>817</td>
<td>162</td>
<td>ZP_03016604</td>
<td>6.00E-97</td>
<td>263</td>
<td>63.9 hypothetical protein BACINT_04211 [Bacteroides intestinalis DSM 17393]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig10206</td>
<td>GH95</td>
<td>702</td>
<td>116</td>
<td>ZP_0345834</td>
<td>1.00E-82</td>
<td>229</td>
<td>60.7 hypothetical protein BACEGG_01309 [Bacteroides eggerthii DSM 20697]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig23822</td>
<td>GH87</td>
<td>2301</td>
<td>2654</td>
<td>YP_001558424</td>
<td>0</td>
<td>593</td>
<td>60.9 APHP domain-containing protein [Clostridium phytofermentans ISDg]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6769</td>
<td>GH87</td>
<td>891</td>
<td>255</td>
<td>YP_001558424</td>
<td>1.00E-102</td>
<td>180</td>
<td>70.6 APHP domain-containing protein [Clostridium phytofermentans ISDg]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig13875</td>
<td>GH87</td>
<td>1351</td>
<td>1222</td>
<td>YP_001558424</td>
<td>1.00E-130</td>
<td>356</td>
<td>65.4 APHP domain-containing protein [Clostridium phytofermentans ISDg]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2118</td>
<td>GH74</td>
<td>2396</td>
<td>3004</td>
<td>ZP_06145570</td>
<td>0</td>
<td>763</td>
<td>55.0 endo-1,4-beta-glucanase/xylolucanase, putative, gly74A [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Carbohydrate family</td>
<td>Start (Contig)</td>
<td>End (Contig)</td>
<td>Accession</td>
<td>gi</td>
<td>Reference</td>
<td>E-value</td>
<td>Score</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------------</td>
<td>---------------</td>
<td>-------------</td>
<td>-----------</td>
<td>----</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig28721</td>
<td>GH74</td>
<td>1428</td>
<td>894</td>
<td>ZP_061455_70</td>
<td>gi</td>
<td>268611843</td>
<td>ref</td>
<td>ZP_061455_70</td>
<td>1.00E-128</td>
<td>520</td>
<td>47.3</td>
<td>endo-1,4-beta-glucanase/xyloglucanase, putative, gly74A [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig24305</td>
<td>GH74</td>
<td>1399</td>
<td>232</td>
<td>ZP_061455_70</td>
<td>gi</td>
<td>268611843</td>
<td>ref</td>
<td>ZP_061455_70</td>
<td>1.00E-146</td>
<td>486</td>
<td>55.1</td>
<td>endo-1,4-beta-glucanase/xyloglucanase, putative, gly74A [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3464</td>
<td>GH74</td>
<td>769</td>
<td>380</td>
<td>ZP_061455_70</td>
<td>gi</td>
<td>268611843</td>
<td>ref</td>
<td>ZP_061455_70</td>
<td>2.00E-64</td>
<td>237</td>
<td>54.4</td>
<td>endo-1,4-beta-glucanase/xyloglucanase, putative, gly74A [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig23418</td>
<td>GH74</td>
<td>538</td>
<td>108</td>
<td>ZP_061455_70</td>
<td>gi</td>
<td>268611843</td>
<td>ref</td>
<td>ZP_061455_70</td>
<td>1.00E-40</td>
<td>185</td>
<td>48.1</td>
<td>endo-1,4-beta-glucanase/xyloglucanase, putative, gly74A [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5455</td>
<td>GH74</td>
<td>597</td>
<td>197</td>
<td>ZP_061455_70</td>
<td>gi</td>
<td>268611843</td>
<td>ref</td>
<td>ZP_061455_70</td>
<td>2.00E-45</td>
<td>203</td>
<td>45.8</td>
<td>endo-1,4-beta-glucanase/xyloglucanase, putative, gly74A [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22447</td>
<td>CBM29</td>
<td>1580</td>
<td>368</td>
<td>AAK20910</td>
<td>gi</td>
<td>13446353</td>
<td>gb</td>
<td>AAK20910.1</td>
<td>4.00E-08</td>
<td>347</td>
<td>24.5</td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29519</td>
<td>CBM29</td>
<td>553</td>
<td>857</td>
<td>AAK20910</td>
<td>gi</td>
<td>13446353</td>
<td>gb</td>
<td>AAK20910.1</td>
<td>2.00E-57</td>
<td>185</td>
<td>60.0</td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig7989</td>
<td>CBM29</td>
<td>549</td>
<td>77</td>
<td>AAK20910</td>
<td>gi</td>
<td>13446353</td>
<td>gb</td>
<td>AAK20910.1</td>
<td>4.00E-46</td>
<td>190</td>
<td>48.9</td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2640</td>
<td>CBM36</td>
<td>1288</td>
<td>497</td>
<td>YP_001038_591</td>
<td>gi</td>
<td>125974681</td>
<td>ref</td>
<td>YP_001038_8591.1</td>
<td>1.00E-156</td>
<td>425</td>
<td>62.1</td>
<td>carbohydrate-binding family 6 protein [Clostridium thermocellum ATCC 27405]gi</td>
<td>256004120</td>
<td>ref</td>
<td>ZP_05429104.1</td>
<td>Carbohydrate binding family 6 [Clostridium thermocellum DSM 2360]gi</td>
<td>281418847</td>
<td>ref</td>
<td>ZP_06249866.1</td>
</tr>
</tbody>
</table>
Contig22009 | CBM36 | 1443 | 1537 | ACZ98594 | gi|280977753| gb|ACZ98594 .1| 1.00E-136 | 462 | 51.9 | endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]

Note: * Possible Full Length Gene

**Table S.2** Muskoxen rumen metatranscriptome contigs (≥500 bp) that have two or more distinct putative CAZy modules.

<table>
<thead>
<tr>
<th>Contig number</th>
<th>Domains</th>
<th>Length</th>
<th>Number of Reads</th>
<th>Accession</th>
<th>GI</th>
<th>E-value</th>
<th>HSP Length</th>
<th>HSP Id%</th>
<th>Hit Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig3930</td>
<td>GH3; GH6; CBM10</td>
<td>3927</td>
<td>6647</td>
<td>AAD02028</td>
<td>gi</td>
<td>4104400</td>
<td>gb</td>
<td>AAD02028.1</td>
<td>0</td>
</tr>
<tr>
<td>Contig22419</td>
<td>GH3; CBM10</td>
<td>2647</td>
<td>2360</td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
<td>0</td>
</tr>
<tr>
<td>Contig28965</td>
<td>GH3; CBM10</td>
<td>2585</td>
<td>3385</td>
<td>AAO41704</td>
<td>gi</td>
<td>28557461</td>
<td>gb</td>
<td>AAO41704.1</td>
<td>0</td>
</tr>
<tr>
<td>Contig1220</td>
<td>Cellulase; CBM10</td>
<td>2239</td>
<td>417</td>
<td>CAB92326</td>
<td>gi</td>
<td>8052316</td>
<td>emb</td>
<td>CAB92326.1</td>
<td>0</td>
</tr>
<tr>
<td>Contig1475</td>
<td>Cellulase; CBM10</td>
<td>1612</td>
<td>1800</td>
<td>AAL83749</td>
<td>gi</td>
<td>19070186</td>
<td>gb</td>
<td>AAL83749.1</td>
<td>1.00E-107</td>
</tr>
<tr>
<td>Contig5338</td>
<td>Cellulase; CBM10</td>
<td>1737</td>
<td>2114</td>
<td>AAL83749</td>
<td>gi</td>
<td>19070186</td>
<td>gb</td>
<td>AAL83749.1</td>
<td>1.00E-106</td>
</tr>
<tr>
<td>Contig23559</td>
<td>Cellulase; CBM10</td>
<td>1658</td>
<td>1110</td>
<td>NP_521723</td>
<td>gi</td>
<td>17548383</td>
<td>ref</td>
<td>NP_521723 .1</td>
<td>2.00E-66</td>
</tr>
<tr>
<td>Contig31560</td>
<td>Cellulase; CBM10</td>
<td>1206</td>
<td>881</td>
<td>AAL83749</td>
<td>gi</td>
<td>19070186</td>
<td>gb</td>
<td>AAL83749.1</td>
<td>2.00E-59</td>
</tr>
<tr>
<td>Contig23055</td>
<td>Cellulase; CBM10</td>
<td>787</td>
<td>422</td>
<td>CBL17363</td>
<td>gi</td>
<td>291544254</td>
<td>emb</td>
<td>CBL17363 3.1</td>
<td>7.00E-45</td>
</tr>
<tr>
<td>Contig</td>
<td>Type</td>
<td>Accession</td>
<td>E-value</td>
<td>Score</td>
<td>Identities</td>
<td>RecName</td>
<td>AltName</td>
<td>Other Information</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
<td>------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Contig30802</td>
<td>Cellulase; CBM10</td>
<td>P23548</td>
<td>2.00E-28</td>
<td>137</td>
<td>43.1</td>
<td>RecName: Full=Endoglucanase; AltName: Full=Endo-1,4-beta-glucanase; AltName: Full=Cellulasegi</td>
<td>143271</td>
<td>gb</td>
<td>AAA22631.1</td>
</tr>
<tr>
<td>Contig30795</td>
<td>Cellulase; CBM10</td>
<td>CAB92326</td>
<td>3.00E-77</td>
<td>203</td>
<td>62.6</td>
<td>endoglucanase 5A [Pirimyces equi]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22047</td>
<td>GH6; CBM10</td>
<td>AAL01211</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>celllobiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
<td></td>
</tr>
<tr>
<td>Contig111</td>
<td>GH6; CBM10</td>
<td>AAB92678</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>cellulase A [Orpinomyces sp. PC-2]</td>
<td></td>
</tr>
<tr>
<td>Contig1523</td>
<td>GH6; CBM1</td>
<td>AAR08200</td>
<td>1.00E-178</td>
<td>319</td>
<td>96.6</td>
<td>CeL [Neocallimastix frontalis]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29791</td>
<td>GH6; CBM10</td>
<td>ABY52799</td>
<td>1.00E-125</td>
<td>287</td>
<td>74.9</td>
<td>1,4-beta-D-glucan-celllobiohydrolase [Pirimyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3027</td>
<td>GH6; CBM10</td>
<td>AAD51054</td>
<td>1.00E-174</td>
<td>433</td>
<td>69.1</td>
<td>exocelllobiohydrolase Chb120 [Pirimyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1794</td>
<td>GH6; CBM10</td>
<td>ACX32999</td>
<td>1.00E-115</td>
<td>372</td>
<td>58.6</td>
<td>1,4-beta-glucanase [Pirimyces sp. BTrP1]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig7415</td>
<td>GH6; CBM10</td>
<td>AAB92678</td>
<td>1.00E-120</td>
<td>308</td>
<td>70.1</td>
<td>cellulase A [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig28654</td>
<td>GH6; GH3C</td>
<td>AAD02028</td>
<td>1.00E-113</td>
<td>252</td>
<td>78.2</td>
<td>exocelllobiohydrolase precursor [Pirimyces rhizinflatus]</td>
<td>exoglucanase Cel6A [Pirimyces sp. E2]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1597</td>
<td>GH6; CBM10</td>
<td>AAL92497</td>
<td>1.00E-117</td>
<td>309</td>
<td>65.4</td>
<td>exoglucanase Cel6A [Pirimyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30772</td>
<td>GH6; CBM10</td>
<td>AAD51054</td>
<td>1.00E-129</td>
<td>350</td>
<td>66.3</td>
<td>exocelllobiohydrolase Chb120 [Pirimyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2418</td>
<td>GH6; CBM10</td>
<td>AAD02028</td>
<td>1.00E-125</td>
<td>335</td>
<td>66.0</td>
<td>exocelllobiohydrolase precursor [Pirimyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig9323</td>
<td>GH6; CBM10</td>
<td>AAP33843</td>
<td>1.00E-52</td>
<td>125</td>
<td>80.0</td>
<td>hybrid 1,4-beta-glucanase [synthetic construct]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM</td>
<td>Length</td>
<td>Identity</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>--------</td>
<td>----------</td>
<td>-----------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig339</td>
<td>GH; CBM10</td>
<td>784</td>
<td>519</td>
<td>AAL01212</td>
<td>cellobiohydrolase II-like cellulase CelI [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21750</td>
<td>GH; CBM10</td>
<td>641</td>
<td>264</td>
<td>AAM94167</td>
<td>cellulosomal glycoside hydrolase family 6 exoglucanase Cel6A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22729</td>
<td>GH; CBM10</td>
<td>632</td>
<td>1925</td>
<td>AAM94167</td>
<td>cellulosomal glycoside hydrolase family 6 exoglucanase Cel6A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29107</td>
<td>GH; CBM10</td>
<td>719</td>
<td>2591</td>
<td>AAL01212</td>
<td>cellobiohydrolase II-like cellulase CelI [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6887</td>
<td>GH; CBM10</td>
<td>657</td>
<td>542</td>
<td>AAL01211</td>
<td>cellobiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig18293</td>
<td>GH; CBM10</td>
<td>714</td>
<td>360</td>
<td>AAD02028</td>
<td>exocelllobiohydrolase precursor [Piromyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig9</td>
<td>GH; CBM10</td>
<td>1423</td>
<td>159</td>
<td>AAL92497</td>
<td>exoglucanase Cel6A [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig24462</td>
<td>GH; CBM10</td>
<td>532</td>
<td>240</td>
<td>AAL92497</td>
<td>exoglucanase Cel6A [Piromyces sp. E2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29682</td>
<td>GH; CBM10</td>
<td>695</td>
<td>821</td>
<td>AAM94167</td>
<td>cellulosomal glycoside hydrolase family 6 exoglucanase Cel6A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30051</td>
<td>GH; CBM10</td>
<td>573</td>
<td>1216</td>
<td>AAD02028</td>
<td>exocelllobiohydrolase precursor [Piromyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig31138</td>
<td>GH; CBM10</td>
<td>637</td>
<td>333</td>
<td>AAL01211</td>
<td>cellobiohydrolase II-like cellulase CelH [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29495</td>
<td>GH; CBM10</td>
<td>728</td>
<td>3299</td>
<td>AAD02028</td>
<td>exocelllobiohydrolase precursor [Piromyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21518</td>
<td>GH; CBM10</td>
<td>1799</td>
<td>4366</td>
<td>YP_003248565</td>
<td>glycoside hydrolase family 8 [Fibrobacter succinogenes subsp. succinogenes]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM</td>
<td>Length</td>
<td>Molec. Wt.</td>
<td>Accession</td>
<td>Identity</td>
<td>E-value</td>
<td>Function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>--------</td>
<td>------------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22627</td>
<td>GH8; GH11; CBM10</td>
<td>3510 5526 YP_003248 565</td>
<td>0 432 74.5</td>
<td>gi</td>
<td>261414882</td>
<td>ref[YP_003248 8565.1]</td>
<td>glycoside hydrolase family 8 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29700</td>
<td>GH8; CBM10</td>
<td>1784 3464 YP_003248 565</td>
<td>0 429 76.5</td>
<td>gi</td>
<td>261414882</td>
<td>ref[YP_003248 8565.1]</td>
<td>glycoside hydrolase family 8 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig24688</td>
<td>GH8; CBM10</td>
<td>2005 6086 YP_003248 565</td>
<td>0 434 75.3</td>
<td>gi</td>
<td>261414882</td>
<td>ref[YP_003248 8565.1]</td>
<td>glycoside hydrolase family 8 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21257</td>
<td>GH9; CBM10</td>
<td>2365 1114 AAM81967 565</td>
<td>0 783 63.7</td>
<td>gi</td>
<td>21929669</td>
<td>gb[AAM81967.1]</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21102</td>
<td>GH9; CBM10</td>
<td>2268 956 AAM81966 565</td>
<td>0 765 62.6</td>
<td>gi</td>
<td>21929667</td>
<td>gb[AAM81966.1]</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig21532</td>
<td>GH9; CBM10</td>
<td>2325 2223 AAM81966 565</td>
<td>0 771 63.0</td>
<td>gi</td>
<td>21929667</td>
<td>gb[AAM81966.1]</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29363</td>
<td>GH9; CBM10</td>
<td>2220 2267 AAM81966 565</td>
<td>0 749 78.0</td>
<td>gi</td>
<td>21929667</td>
<td>gb[AAM81966.1]</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>gi</td>
<td>Accession</td>
<td>Identity</td>
<td>E-value</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>---------</td>
<td>-----------</td>
<td>----------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>Contig1941</td>
<td>GH9; CBM10</td>
<td>2386</td>
<td>2616</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>663</td>
<td>81.1</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig3953</td>
<td>GH9; CBM10</td>
<td>1867</td>
<td>1195</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>475</td>
<td>45.9</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig8450</td>
<td>GH9; CBM10</td>
<td>2358</td>
<td>1807</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>666</td>
<td>59.5</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig22760</td>
<td>GH9; CBM10</td>
<td>1655</td>
<td>724</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>425</td>
<td>44.0</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig4358</td>
<td>GH9; CBM10</td>
<td>2239</td>
<td>1281</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>507</td>
<td>39.8</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig436</td>
<td>GH9; CBM3</td>
<td>3099</td>
<td>161</td>
<td>CAL91976</td>
<td>gi</td>
<td>21808135</td>
<td>621</td>
<td>75.2</td>
<td>cellulase [Epidinium ecaudatum]</td>
</tr>
<tr>
<td>Contig3929</td>
<td>GH9; CBM10</td>
<td>2170</td>
<td>1186</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>607</td>
<td>52.9</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig21739</td>
<td>GH9; CBM10</td>
<td>2073</td>
<td>1407</td>
<td>AAM81966</td>
<td>gi</td>
<td>21929667</td>
<td>685</td>
<td>73.9</td>
<td>cellulase Cel9A precursor [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig7660</td>
<td>GH9; CBM10</td>
<td>1425</td>
<td>1524</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>373</td>
<td>75.1</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig8733</td>
<td>GH9; CBM3</td>
<td>855</td>
<td>898</td>
<td>CAL91976</td>
<td>gi</td>
<td>21808135</td>
<td>285</td>
<td>88.1</td>
<td>cellulase [Epidinium ecaudatum]</td>
</tr>
<tr>
<td>Contig1701</td>
<td>GH9; CBM3</td>
<td>797</td>
<td>274</td>
<td>CAL91976</td>
<td>gi</td>
<td>21808135</td>
<td>264</td>
<td>95.8</td>
<td>cellulase [Epidinium ecaudatum]</td>
</tr>
<tr>
<td>Contig30777</td>
<td>GH9; CBM10</td>
<td>989</td>
<td>566</td>
<td>AAP30753</td>
<td>gi</td>
<td>30315047</td>
<td>289</td>
<td>41.5</td>
<td>cellulosomal glycoside hydrolase family 9 endoglucanase Cel9B [Piromyces sp. E2]</td>
</tr>
<tr>
<td>Contig240</td>
<td>GH10; CBM4_9</td>
<td>1618</td>
<td>248</td>
<td>CAB65753</td>
<td>gi</td>
<td>6692066</td>
<td>514</td>
<td>94.9</td>
<td>xylanase 10B precursor</td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM</td>
<td>3168</td>
<td>2175</td>
<td>AAB30669</td>
<td>gi</td>
<td>560649</td>
<td>gb</td>
<td>AAB30669.1</td>
<td>1.00E-157</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>------</td>
<td>------</td>
<td>-----------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Contig21350</td>
<td>GH10; Ricin_B_lectin</td>
<td>1439</td>
<td>2576</td>
<td>BAC57894</td>
<td>gi</td>
<td>28569972</td>
<td>db</td>
<td>BAC57894.1</td>
<td>0</td>
</tr>
<tr>
<td>Contig29712</td>
<td>GH10; CBM1</td>
<td>1493</td>
<td>1531</td>
<td>AAB30669</td>
<td>gi</td>
<td>560649</td>
<td>gb</td>
<td>AAB30669.1</td>
<td>1.00E-129</td>
</tr>
<tr>
<td>Contig24573</td>
<td>GH10; CBM10</td>
<td>1138</td>
<td>548</td>
<td>AAB30669</td>
<td>gi</td>
<td>560649</td>
<td>gb</td>
<td>AAB30669.1</td>
<td>1.00E-120</td>
</tr>
<tr>
<td>Contig85</td>
<td>GH10; CBM10</td>
<td>1125</td>
<td>484</td>
<td>AAB30669</td>
<td>gi</td>
<td>560649</td>
<td>gb</td>
<td>AAB30669.1</td>
<td>7.00E-96</td>
</tr>
<tr>
<td>Contig29849</td>
<td>GH10; CBM29</td>
<td>2295</td>
<td>AAB30669</td>
<td>gi</td>
<td>560649</td>
<td>gb</td>
<td>AAB30669.1</td>
<td>6.00E-43</td>
<td>318</td>
</tr>
<tr>
<td>Contig28853</td>
<td>GH10; Ricin_B_lectin</td>
<td>1071</td>
<td>2539</td>
<td>BAC57894</td>
<td>gi</td>
<td>28569972</td>
<td>db</td>
<td>BAC57894.1</td>
<td>1.00E-171</td>
</tr>
<tr>
<td>Contig222047</td>
<td>GH10; CBM10</td>
<td>653</td>
<td>352</td>
<td>AAB30669</td>
<td>gi</td>
<td>560649</td>
<td>gb</td>
<td>AAB30669.1</td>
<td>1.00E-31</td>
</tr>
<tr>
<td>Contig28800</td>
<td>GH10; CBM10</td>
<td>751</td>
<td>622</td>
<td>AAB69092</td>
<td>gi</td>
<td>2231247</td>
<td>gb</td>
<td>AAB69092.1</td>
<td>3.00E-28</td>
</tr>
<tr>
<td>Contig3245</td>
<td>GH10; CBM10</td>
<td>703</td>
<td>579</td>
<td>AAB69092</td>
<td>gi</td>
<td>2231247</td>
<td>gb</td>
<td>AAB69092.1</td>
<td>2.00E-28</td>
</tr>
<tr>
<td>Contig11078</td>
<td>GH10; CBM10</td>
<td>1305</td>
<td>562</td>
<td>AAQ10005</td>
<td>gi</td>
<td>333929210</td>
<td>gb</td>
<td>AAQ10005.1</td>
<td>1.00E-115</td>
</tr>
<tr>
<td>Contig19309</td>
<td>GH11; CBM10</td>
<td>1830</td>
<td>1583</td>
<td>Q12667</td>
<td>gi</td>
<td>2494347</td>
<td>sp</td>
<td>Q12667.1</td>
<td>XYNA_PIRSP</td>
</tr>
<tr>
<td>Contig22239</td>
<td>GH11; CBM10</td>
<td>2106</td>
<td>6698</td>
<td>ACL68347</td>
<td>gi</td>
<td>219964511</td>
<td>gb</td>
<td>ACL68347.1</td>
<td>1.00E-107</td>
</tr>
<tr>
<td>Contig22400</td>
<td>GH11; Polysacc_deac_1</td>
<td>1388</td>
<td>1774</td>
<td>AAF14365</td>
<td>gi</td>
<td>650258</td>
<td>gb</td>
<td>AAF14365.1</td>
<td>AFI13252</td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM</td>
<td>Start</td>
<td>End</td>
<td>Accession 1</td>
<td>Accession 2</td>
<td>E-value</td>
<td>Identity</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>-------------</td>
<td>---------</td>
<td>----------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Contig30387</td>
<td>GH11; CBM10</td>
<td>2084</td>
<td>7333</td>
<td>ACL68347</td>
<td>gb</td>
<td>ACL68347</td>
<td>1.00E-139</td>
<td>68.9</td>
<td>endo-1,4-beta-xylanase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig30545</td>
<td>GH11; CBM1</td>
<td>1032</td>
<td>1115</td>
<td>ABW04217</td>
<td>gb</td>
<td>ABW0421</td>
<td>1.00E-118</td>
<td>87.0</td>
<td>endo-1,4-beta-xylanase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig6117</td>
<td>GH11; Polysacc_deac_1</td>
<td>1744</td>
<td>1843</td>
<td>ABW04217</td>
<td>gb</td>
<td>ABW0421</td>
<td>1.00E-108</td>
<td>78.4</td>
<td>endo-1,4-beta-xylanase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig23659</td>
<td>GH11; CBM10</td>
<td>1804</td>
<td>9175</td>
<td>ACL68347</td>
<td>gb</td>
<td>ACL68347</td>
<td>0</td>
<td>79.2</td>
<td>endo-1,4-beta-xylanase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig22195</td>
<td>GH11; CBM10</td>
<td>1036</td>
<td>4354</td>
<td>Q12667</td>
<td>gb</td>
<td>Q12667</td>
<td>1.00E-112</td>
<td>61.6</td>
<td>RecName: Full=Endo-1,4-beta-xylanase A; AltName: Full=1,4-beta-D-xylan xylanohydrolase; AltName: Full=Xylanase A; Short=XYLA; Flags: Precursor</td>
</tr>
<tr>
<td>Contig31461</td>
<td>GH11; CBM1</td>
<td>548</td>
<td>267</td>
<td>Q12667</td>
<td>gb</td>
<td>Q12667</td>
<td>4.00E-46</td>
<td>74.8</td>
<td>RecName: Full=Endo-1,4-beta-xylanase A; AltName: Full=1,4-beta-D-xylan xylanohydrolase; AltName: Full=Xylanase A; Short=XYLA; Flags: Precursor</td>
</tr>
<tr>
<td>Contig21648</td>
<td>GH11; CBM10</td>
<td>1187</td>
<td>576</td>
<td>Q12667</td>
<td>gb</td>
<td>Q12667</td>
<td>2.00E-90</td>
<td>68.8</td>
<td>RecName: Full=Endo-1,4-beta-xylanase A; AltName: Full=1,4-beta-D-xylan xylanohydrolase; AltName: Full=Xylanase A; Short=XYLA; Flags: Precursor</td>
</tr>
<tr>
<td>Contig30887</td>
<td>GH11; CBM10</td>
<td>1082</td>
<td>811</td>
<td>Q12667</td>
<td>gb</td>
<td>Q12667</td>
<td>1.00E-86</td>
<td>77.1</td>
<td>RecName: Full=Endo-1,4-beta-xylanase A; AltName: Full=1,4-beta-D-xylan xylanohydrolase; AltName: Full=Xylanase A; Short=XYLA; Flags:</td>
</tr>
<tr>
<td>Contig</td>
<td>GH1; CBM12</td>
<td>916</td>
<td>266</td>
<td>Q12667</td>
<td>gi</td>
<td>2494337</td>
<td>sp</td>
<td>Q12667.1</td>
<td>XY NA_PIRSP</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>----</td>
<td>--------</td>
<td>--------</td>
<td>---------------</td>
<td>----------</td>
</tr>
<tr>
<td>Contig2264</td>
<td>GH11; CBM10</td>
<td>541</td>
<td>2541</td>
<td>AAD04194</td>
<td>gi</td>
<td>1655815</td>
<td>gb</td>
<td>AAD04194.1</td>
<td>6.00E-84</td>
</tr>
<tr>
<td>Contig22817</td>
<td>GH16; CBM10</td>
<td>1146</td>
<td>1105</td>
<td>AAB69347</td>
<td>gi</td>
<td>2353005</td>
<td>gb</td>
<td>AAB69347.1</td>
<td>1.00E-18</td>
</tr>
<tr>
<td>Contig29248</td>
<td>GH11; CBM10</td>
<td>541</td>
<td>2541</td>
<td>AAD04194</td>
<td>gi</td>
<td>1655815</td>
<td>gb</td>
<td>AAD04194.1</td>
<td>6.00E-84</td>
</tr>
<tr>
<td>Contig29583</td>
<td>GH16; CBM10</td>
<td>1039</td>
<td>820</td>
<td>Q12647</td>
<td>gi</td>
<td>2494328</td>
<td>sp</td>
<td>Q12647.1</td>
<td>GU NB_NEOPA</td>
</tr>
<tr>
<td>Contig25795</td>
<td>GH16; CBM10</td>
<td>1080</td>
<td>911</td>
<td>ZP_055122 36</td>
<td>gi</td>
<td>256773773</td>
<td>ref</td>
<td>ZP_055122 236.1</td>
<td>4.00E-90</td>
</tr>
<tr>
<td>Contig1050</td>
<td>GH18; Chitin_bind_1</td>
<td>2417</td>
<td>3248</td>
<td>YP_001643 260</td>
<td>gi</td>
<td>163938376</td>
<td>ref</td>
<td>YP_001643 3260.1</td>
<td>2.00E-56</td>
</tr>
<tr>
<td>Contig</td>
<td>GH18; Chitin_bind_1</td>
<td>2273</td>
<td>1527</td>
<td>YP_001643 260</td>
<td>gi</td>
<td>163938376</td>
<td>ref</td>
<td>YP_001643 3260.1</td>
<td>3.00E-57</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Contig8720</td>
<td>GH18; Chitin_bind_1</td>
<td>1439</td>
<td>567</td>
<td>Q01MB6</td>
<td>gi</td>
<td>152013345</td>
<td>sp</td>
<td>Q01MB6.2</td>
<td>AGI_ORYSI</td>
</tr>
<tr>
<td>Contig12835</td>
<td>GH18; Chitin_bind_1</td>
<td>1118</td>
<td>480</td>
<td>Q01MB6</td>
<td>gi</td>
<td>152013345</td>
<td>sp</td>
<td>Q01MB6.2</td>
<td>AGI_ORYSI</td>
</tr>
<tr>
<td>Contig30315</td>
<td>GH26; CBM10</td>
<td>1355</td>
<td>1527</td>
<td>YP_003249 349</td>
<td>gi</td>
<td>261415666</td>
<td>ref</td>
<td>YP_003249 9349.1</td>
<td>2.00E-91</td>
</tr>
<tr>
<td>Contig30263</td>
<td>GH26; CBM10</td>
<td>1272</td>
<td>781</td>
<td>YP_003249 349</td>
<td>gi</td>
<td>261415666</td>
<td>ref</td>
<td>YP_003249 9349.1</td>
<td>1.00E-103</td>
</tr>
<tr>
<td>Contig</td>
<td>Gene Description</td>
<td>E-value</td>
<td>Accession</td>
<td>Bit Score</td>
<td>Percent Identity</td>
<td>Classification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------------------</td>
<td>---------</td>
<td>---------------</td>
<td>-----------</td>
<td>------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22949</td>
<td><strong>Fibrobacter succinogenes subsp. succinogenes S85</strong> GH26; CBM10 1278 865 YP_003249349</td>
<td>4.00E-96</td>
<td>gi</td>
<td>261415666</td>
<td>ref</td>
<td>YP_0032499349.1</td>
<td>281 57.3 Beta-mannanase-like protein [Fibrobacter succinogenes subsp. succinogenes S85]gi</td>
<td>261372122</td>
<td>gb</td>
</tr>
<tr>
<td>Contig21972</td>
<td><strong>Fibrobacter succinogenes s</strong>GH26; CBM35 688 ZP_06144005</td>
<td>3.00E-51</td>
<td>gi</td>
<td>268610278</td>
<td>ref</td>
<td>ZP_06144005.1</td>
<td>205 52.2 mannann endo-1,4-beta-mannosidase [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig4027</td>
<td>Melibiase; Ricin_B_lectin 1309 1006 BAC66445</td>
<td>1.00E-112</td>
<td>gi</td>
<td>29335747</td>
<td>d</td>
<td>BAC66445.1</td>
<td>281 64.1 alpha-galactosidase [Helianthus annuus]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2025</td>
<td>Melibiase; Ricin_B_lectin 1082 31 ABK25009</td>
<td>1.00E-102</td>
<td>gi</td>
<td>116788809</td>
<td>gb</td>
<td>ABK25009.1</td>
<td>264 63.6 unknown [Picea sitchensis]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig7911</td>
<td>GH43; CBM10 1247 1162 YP_001038591</td>
<td>1.00E-101</td>
<td>gi</td>
<td>125974681</td>
<td>ref</td>
<td>YP_001038591.1</td>
<td>297 59.6 carbohydrate-binding family 6 protein [Clostridium thermocellum ATCC 27405]gi</td>
<td>256004120</td>
<td>ref</td>
</tr>
<tr>
<td>Contig2640</td>
<td>GH43; CBM6 1288 705 YP_001038591</td>
<td>1.00E-156</td>
<td>gi</td>
<td>125974681</td>
<td>ref</td>
<td>YP_001038591.1</td>
<td>425 62.1 carbohydrate-binding family 6 protein [Clostridium thermocellum ATCC 27405]gi</td>
<td>256004120</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>GH43; CBM</td>
<td>1215</td>
<td>1515</td>
<td>YP_001038</td>
<td>591</td>
<td>gi</td>
<td>25974681</td>
<td>ref</td>
<td>YP_001038591.1</td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>------</td>
<td>------</td>
<td>-------------</td>
<td>-----</td>
<td>-----</td>
<td>----------------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>Contig2588</td>
<td>GH43; CBM10</td>
<td>1215</td>
<td>1515</td>
<td>YP_001038</td>
<td>591</td>
<td>gi</td>
<td>25974681</td>
<td>ref</td>
<td>YP_001038591.1</td>
</tr>
<tr>
<td>Contig22546</td>
<td>GH43; CBM6; CBM10</td>
<td>1686</td>
<td>1303</td>
<td>ACZ98594</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594</td>
<td>1.00E-140</td>
</tr>
<tr>
<td>Contig22546</td>
<td>GH43; CBM10</td>
<td>2028</td>
<td>183</td>
<td>ACZ98594</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594</td>
<td>1.00E-125</td>
</tr>
<tr>
<td>Contig717</td>
<td>GH43; CBM1; CBM6</td>
<td>2028</td>
<td>183</td>
<td>ACZ98594</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594</td>
<td>1.00E-125</td>
</tr>
<tr>
<td>Contig22009</td>
<td>GH43; CBM6</td>
<td>1443</td>
<td>2151</td>
<td>ACZ98594</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594</td>
<td>1.00E-136</td>
</tr>
<tr>
<td>Contig110</td>
<td>GH43; CBM6</td>
<td>1448</td>
<td>3325</td>
<td>ACZ98594</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594</td>
<td>1.00E-140</td>
</tr>
<tr>
<td>ContigID</td>
<td>Genotype</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>Identity</td>
<td>Expect</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>-------</td>
<td>-----</td>
<td>-----------</td>
<td>----------</td>
<td>--------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig12288</td>
<td>GH43; CBM6</td>
<td>1266</td>
<td>761</td>
<td>ACZ98594</td>
<td>1.00E-124</td>
<td>416</td>
<td>54.3 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig23444</td>
<td>GH43; CBM6</td>
<td>1567</td>
<td>552</td>
<td>ACZ98594</td>
<td>1.00E-123</td>
<td>470</td>
<td>48.1 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29841</td>
<td>GH43; CBM6</td>
<td>2231</td>
<td>2528</td>
<td>ACZ98594</td>
<td>1.00E-129</td>
<td>442</td>
<td>52.7 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30950</td>
<td>GH43; CBM6</td>
<td>1662</td>
<td>5701</td>
<td>ACZ98594</td>
<td>1.00E-114</td>
<td>476</td>
<td>46.0 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_1061_length_1544_cov_11.881476</td>
<td>GH43; CBM6</td>
<td>1596</td>
<td>4307</td>
<td>ACZ98594</td>
<td>1.00E-1121</td>
<td>474</td>
<td>48.7 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3221</td>
<td>CBM10/CBM6/ GH43/Ricin_B_lectin</td>
<td>2333</td>
<td>1609</td>
<td>CBK75020</td>
<td>1.00E-162</td>
<td>661</td>
<td>46.9 Cellubiohydrolase A (1,4-beta-cellobiosidase A) [Butyrivibrio fibrisolvens 16/4]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig880</td>
<td>GH43; CBM6</td>
<td>1573</td>
<td>680</td>
<td>ACZ98594</td>
<td>1.00E-115</td>
<td>430</td>
<td>50.7 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig14064</td>
<td>GH43; Ricin_B_lectin</td>
<td>1463</td>
<td>646</td>
<td>ZP_06142338</td>
<td>1.00E-129</td>
<td>558</td>
<td>47.3 Alpha-N-arabinofuranosidase [Ruminococcus flavefaciens FD-1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig9591</td>
<td>GH43; CBM6</td>
<td>929</td>
<td>559</td>
<td>ADE82665</td>
<td>1.00E-108</td>
<td>291</td>
<td>64.6 glycosyl hydrolase, family 43 [Prevotella ruminicola 23]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30960</td>
<td>GH43; CBM10/ Ricin_B_lectin</td>
<td>1712</td>
<td>377</td>
<td>ACZ98594</td>
<td>4.00E-91</td>
<td>477</td>
<td>39.8 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1720</td>
<td>GH43; CBM6; CBM10</td>
<td>1206</td>
<td>390</td>
<td>ACZ98594</td>
<td>5.00E-77</td>
<td>298</td>
<td>49.7 endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1733</td>
<td>GH45; CBM10</td>
<td>1399</td>
<td>1630</td>
<td>BAC53956</td>
<td>1.00E-79</td>
<td>325</td>
<td>46.8 endo-beta-1,4-D-glucanase [Rhizopus oryzae]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22152</td>
<td>GH45; CBM10</td>
<td>1130</td>
<td>1019</td>
<td>CAB92325</td>
<td>3.00E-91</td>
<td>402</td>
<td>43.8 endoglucanase 45A [Piromyces equi]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM</td>
<td>1094</td>
<td>900</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-121</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>------</td>
<td>------</td>
<td>----------</td>
<td>gb</td>
<td>182149135</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-103</td>
</tr>
<tr>
<td>Contig31575</td>
<td>GH; CBM</td>
<td>1425</td>
<td>1422</td>
<td>BAD95809</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-78</td>
</tr>
<tr>
<td>Contig929</td>
<td>GH; CBM</td>
<td>1296</td>
<td>116</td>
<td>ABU49185</td>
<td>gi</td>
<td>158138919</td>
<td>gb</td>
<td>ABU49185 2</td>
<td>3.00E-73</td>
</tr>
<tr>
<td>Contig10492</td>
<td>GH; CBM</td>
<td>1102</td>
<td>984</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-101</td>
</tr>
<tr>
<td>Contig30272</td>
<td>GH; CBM</td>
<td>861</td>
<td>474</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-79</td>
</tr>
<tr>
<td>Contig32'</td>
<td>GH; CBM</td>
<td>1284</td>
<td>214</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>0</td>
</tr>
<tr>
<td>Contig22584</td>
<td>GH; CBM</td>
<td>1142</td>
<td>4309</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-102</td>
</tr>
<tr>
<td>Contig3753'</td>
<td>GH; CBM</td>
<td>1440</td>
<td>1409</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-103</td>
</tr>
<tr>
<td>Contig17541</td>
<td>GH; CBM</td>
<td>1420</td>
<td>433</td>
<td>ABU49185</td>
<td>gi</td>
<td>158138919</td>
<td>gb</td>
<td>ABU49185 2</td>
<td>8.00E-71</td>
</tr>
<tr>
<td>Contig1445'</td>
<td>GH; CBM</td>
<td>1513</td>
<td>3881</td>
<td>BAC53988</td>
<td>gi</td>
<td>27530617</td>
<td>db</td>
<td>BAC53988 2</td>
<td>5.00E-74</td>
</tr>
<tr>
<td>Contig30384'</td>
<td>GH; CBM</td>
<td>1251</td>
<td>620</td>
<td>CAB92325</td>
<td>gi</td>
<td>8052314</td>
<td>emb</td>
<td>CAB92325 5.1</td>
<td>1.00E-177</td>
</tr>
<tr>
<td>Contig22532</td>
<td>GH; CBM</td>
<td>1275</td>
<td>1091</td>
<td>ABU49185</td>
<td>gi</td>
<td>158138919</td>
<td>gb</td>
<td>ABU49185 2</td>
<td>5.00E-53</td>
</tr>
<tr>
<td>Contig25134</td>
<td>GH; CBM</td>
<td>765</td>
<td>240</td>
<td>BAC53956</td>
<td>gi</td>
<td>27530542</td>
<td>db</td>
<td>BAC53956 2</td>
<td>7.00E-128</td>
</tr>
<tr>
<td>Contig</td>
<td>GH45; CBM1; CBM10</td>
<td>Locus</td>
<td>Accession</td>
<td>gi</td>
<td></td>
<td>Description</td>
<td>[Species]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-------------------</td>
<td>-------</td>
<td>-----------</td>
<td>---------</td>
<td>-------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6490</td>
<td></td>
<td>1252</td>
<td>BAC53956</td>
<td></td>
<td>45</td>
<td>endo-beta-1,4-D-glucanase</td>
<td>Rhizopus oryzae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29633</td>
<td></td>
<td>859</td>
<td>CAB92325</td>
<td></td>
<td>243</td>
<td>endoglucanase 45A</td>
<td>Piromyces equi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22037</td>
<td></td>
<td>1010</td>
<td>BAC53988</td>
<td></td>
<td>201</td>
<td>endo-glucanase RCE3</td>
<td>Rhizopus oryzae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23383</td>
<td></td>
<td>1094</td>
<td>BAC53988</td>
<td></td>
<td>206</td>
<td>endo-glucanase RCE3</td>
<td>Rhizopus oryzae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29736</td>
<td></td>
<td>706</td>
<td>CAB92325</td>
<td></td>
<td>246</td>
<td>endoglucanase 45A</td>
<td>Piromyces equi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6087</td>
<td></td>
<td>685</td>
<td>CAB92325</td>
<td></td>
<td>220</td>
<td>endoglucanase 45A</td>
<td>Piromyces equi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29807</td>
<td></td>
<td>774</td>
<td>CAB92325</td>
<td></td>
<td>259</td>
<td>endoglucanase 45A</td>
<td>Piromyces equi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21589</td>
<td></td>
<td>2292</td>
<td>AAN76734</td>
<td></td>
<td>746</td>
<td>cellulase Cel48A</td>
<td>Piromyces sp. E2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NODE_3576_length_2015_cov_12.189578</td>
<td>GH45; CBM10</td>
<td>2067</td>
<td>AAN76734</td>
<td></td>
<td>673</td>
<td>cellulase Cel48A</td>
<td>Piromyces sp. E2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22697</td>
<td></td>
<td>1232</td>
<td>AAN76734</td>
<td></td>
<td>409</td>
<td>cellulase Cel48A</td>
<td>Piromyces sp. E2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29557</td>
<td></td>
<td>954</td>
<td>AAN76734</td>
<td></td>
<td>318</td>
<td>cellulase Cel48A</td>
<td>Piromyces sp. E2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29850</td>
<td>GH45; CBM10</td>
<td>1092</td>
<td>AAN76735</td>
<td></td>
<td>340</td>
<td>cellulase Cel48A</td>
<td>Piromyces equi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM</td>
<td>Length</td>
<td>Start</td>
<td>End</td>
<td>Accession</td>
<td>E-value</td>
<td>Score</td>
<td>Identity</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
<td>----------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Contig22790</td>
<td>GH48; CBM10</td>
<td>1029</td>
<td>641</td>
<td>AAN76734</td>
<td>[gi</td>
<td>25990955</td>
<td>gb</td>
<td>AAN76734.1</td>
<td>AF449413_1]</td>
</tr>
<tr>
<td>Contig29259</td>
<td>GH48; CBM10</td>
<td>1037</td>
<td>4832</td>
<td>AAN76734</td>
<td>[gi</td>
<td>25990955</td>
<td>gb</td>
<td>AAN76734.1</td>
<td>AF449412_1]</td>
</tr>
<tr>
<td>Contig274</td>
<td>GH48; CBM10</td>
<td>1054</td>
<td>8504</td>
<td>AAN76734</td>
<td>[gi</td>
<td>25990955</td>
<td>gb</td>
<td>AAN76734.1</td>
<td>AF449412_1]</td>
</tr>
<tr>
<td>Contig30113</td>
<td>GH48; CBM10</td>
<td>1009</td>
<td>1234</td>
<td>AAN76734</td>
<td>[gi</td>
<td>25990955</td>
<td>gb</td>
<td>AAN76734.1</td>
<td>AF449412_1]</td>
</tr>
<tr>
<td>Contig28658</td>
<td>GH48; CBM10</td>
<td>1093</td>
<td>16001</td>
<td>AAN76734</td>
<td>[gi</td>
<td>25990955</td>
<td>gb</td>
<td>AAN76734.1</td>
<td>AF449412_1]</td>
</tr>
<tr>
<td>Contig28843</td>
<td>GH48; CBM10</td>
<td>980</td>
<td>2910</td>
<td>AAN76735</td>
<td>[gi</td>
<td>25990957</td>
<td>gb</td>
<td>AAN76735.1</td>
<td>AF449413_1]</td>
</tr>
<tr>
<td>Contig22913</td>
<td>GH48; CBM10</td>
<td>1134</td>
<td>496</td>
<td>AAN76734</td>
<td>[gi</td>
<td>25990955</td>
<td>gb</td>
<td>AAN76734.1</td>
<td>AF449412_1]</td>
</tr>
<tr>
<td>Contig29098</td>
<td>GH48; CBM10</td>
<td>1484</td>
<td>10522</td>
<td>AAN76735</td>
<td>[gi</td>
<td>25990957</td>
<td>gb</td>
<td>AAN76735.1</td>
<td>AF449413_1]</td>
</tr>
<tr>
<td>Contig19530</td>
<td>GH48; CBM10</td>
<td>939</td>
<td>1085</td>
<td>AAN76735</td>
<td>[gi</td>
<td>25990957</td>
<td>gb</td>
<td>AAN76735.1</td>
<td>AF449413_1]</td>
</tr>
<tr>
<td>Contig29792</td>
<td>GH48; CBM10</td>
<td>993</td>
<td>721</td>
<td>AAN76735</td>
<td>[gi</td>
<td>25990957</td>
<td>gb</td>
<td>AAN76735.1</td>
<td>AF449413_1]</td>
</tr>
<tr>
<td>Contig</td>
<td>GH; CBM10</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Value 5</td>
<td>Value 6</td>
<td>Value 7</td>
<td>Value 8</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Contig22124</td>
<td>GH48; CBM10</td>
<td>975</td>
<td>2451</td>
<td>AAN76735</td>
<td></td>
<td></td>
<td>1.00E-129</td>
<td></td>
<td>324</td>
</tr>
<tr>
<td>Contig26288</td>
<td>GH48; CBM10</td>
<td>1029</td>
<td>351</td>
<td>AAN76735</td>
<td></td>
<td></td>
<td>1.00E-116</td>
<td></td>
<td>306</td>
</tr>
<tr>
<td>Contig29641</td>
<td>GH48; CBM10</td>
<td>505</td>
<td>1065</td>
<td>AAN76734</td>
<td></td>
<td></td>
<td>5.00E-56</td>
<td></td>
<td>164</td>
</tr>
<tr>
<td>Contig28825</td>
<td>GH48; CBM10</td>
<td>518</td>
<td>2213</td>
<td>AAN76735</td>
<td></td>
<td></td>
<td>3.00E-47</td>
<td></td>
<td>172</td>
</tr>
<tr>
<td>Contig28681</td>
<td>GH48; CBM10</td>
<td>595</td>
<td>6599</td>
<td>AAN76735</td>
<td></td>
<td></td>
<td>2.00E-63</td>
<td></td>
<td>165</td>
</tr>
<tr>
<td>Contig713</td>
<td>CBM1; CBM10</td>
<td>1882</td>
<td>73</td>
<td>NP_593986</td>
<td></td>
<td></td>
<td>1.00E-08</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM Type</td>
<td>Length</td>
<td>Identity</td>
<td>Accession</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>--------</td>
<td>----------</td>
<td>-------------</td>
<td>----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig56</td>
<td>CBM1; CBM10</td>
<td>2097</td>
<td>97</td>
<td>AAL01212</td>
<td>Cellulbiohydrolase II-like cellulase CelI [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig10481</td>
<td>CBM1; CBM10</td>
<td>529</td>
<td>85</td>
<td>AAL01212</td>
<td>Cellulbiohydrolase II-like cellulase CelI [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1742</td>
<td>CBM1; CBM10</td>
<td>2042</td>
<td>115</td>
<td>CAB92325</td>
<td>Endoglucanase 45A [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig30371</td>
<td>CBM1; Chitin_bind_1</td>
<td>829</td>
<td>36</td>
<td>ABY52793</td>
<td>Cellulbiohydrolase [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2052</td>
<td>CBM1; Pfam_B_8046</td>
<td>994</td>
<td>200</td>
<td>XP_001829</td>
<td>Hypothetical protein CC1G_01855 [Coproinopsis cinerea okayama7#130]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig12713</td>
<td>CBM1; CBM10</td>
<td>1642</td>
<td>90</td>
<td>AAL01212</td>
<td>Cellulbiohydrolase II-like cellulase CelI [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig17415</td>
<td>CBM1; CBM10</td>
<td>2081</td>
<td>91</td>
<td>AAL01213</td>
<td>Mannanase ManA [Orpinomyces sp. PC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig1019</td>
<td>CBM1; CBM10</td>
<td>1686</td>
<td>90</td>
<td>AAD43818</td>
<td>Endoglucanase precursor [Piromyces rhizinflatus]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29787</td>
<td>CBM6; Ricin_B_lectin</td>
<td>1281</td>
<td>417</td>
<td>NP_149283</td>
<td>Xylan degradation protein [Clostridium acetobutylicum ATCC 824]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig6460</td>
<td>CBM6; CBM36</td>
<td>520</td>
<td>172</td>
<td>ACZ98594</td>
<td>Endo-1,4-beta-xylanase</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>CBM6; CBM10</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594.1</td>
<td>2.00E-25</td>
<td>147</td>
<td>40.8</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>------------------------------------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>----------</td>
<td>--------</td>
<td>-------------------</td>
<td>-------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>NODE_4815_length_842_contig_8_675772</td>
<td>CBM6; CBM10</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594.1</td>
<td>9.00E-34</td>
<td>178</td>
<td>42.1</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
</tr>
<tr>
<td>Contig29852</td>
<td>CBM6; CBM10</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594.1</td>
<td>2.00E-35</td>
<td>186</td>
<td>39.8</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
</tr>
<tr>
<td>Contig30353</td>
<td>CBM6; CBM10; Ricin_B_lectin</td>
<td>gi</td>
<td>280977753</td>
<td>gb</td>
<td>ACZ98594.1</td>
<td>3.00E-55</td>
<td>346</td>
<td>37.0</td>
<td>endo-1,4-beta-xylanase [Cellulosilyticum ruminicola]</td>
</tr>
<tr>
<td>Contig24142</td>
<td>CBM10; Rhamnogal_lys e</td>
<td>gi</td>
<td>186517294</td>
<td>ref</td>
<td>NP_195516.2</td>
<td>3.00E-69</td>
<td>531</td>
<td>33.1</td>
<td>lyase [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>Contig345</td>
<td>CBM10; Swollenin</td>
<td>gi</td>
<td>159122532</td>
<td>gb</td>
<td>EDP47653.1</td>
<td>1.00E-127</td>
<td>432</td>
<td>55.1</td>
<td>swollenin, putative [Aspergillus fumigatus A1163]</td>
</tr>
<tr>
<td>Contig28884</td>
<td>CBM10; CBM29</td>
<td>gi</td>
<td>13446353</td>
<td>gb</td>
<td>AAK20910.1</td>
<td>2.00E-70</td>
<td>314</td>
<td>44.9</td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
</tr>
<tr>
<td>Contig536</td>
<td>CBM10; CBM29</td>
<td>gi</td>
<td>13446353</td>
<td>gb</td>
<td>AAK20910.1</td>
<td>0</td>
<td>478</td>
<td>63.0</td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
</tr>
<tr>
<td>Contig30769</td>
<td>CBM10; Swollenin</td>
<td>gi</td>
<td>1708917</td>
<td>sp</td>
<td>P55296.1</td>
<td>MA NA_PIRSP</td>
<td>4.00E-32</td>
<td>131</td>
<td>52.7</td>
</tr>
<tr>
<td>Contig3979</td>
<td>CBM10; Swollenin</td>
<td>gi</td>
<td>157488002</td>
<td>gb</td>
<td>ABV57767.1</td>
<td>1.00E-119</td>
<td>392</td>
<td>56.6</td>
<td>swollenin [Hypocrea pseudokoningii]</td>
</tr>
<tr>
<td>Contig28713</td>
<td>CBM10; CBM29</td>
<td>gi</td>
<td>13446353</td>
<td>gb</td>
<td>AAK20910.1</td>
<td>1.00E-131</td>
<td>490</td>
<td>49.2</td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM;</td>
<td>CBM;</td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>----------</td>
<td>---------------</td>
<td>---------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-31</td>
<td>151</td>
<td>44.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Beta-1,4-xylanase [Butyrivibrio fibrisolvens 16/4]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig28612</td>
<td>CBM; CBM29</td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-121</td>
<td>474</td>
<td>46.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>non-catalytic protein 1 [Piromyces equi]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig22156</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-131</td>
<td>485</td>
<td>49.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>swollenin [Hypocrea pseudokoningii]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig2177</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-138</td>
<td>347</td>
<td>66.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29571</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>394</td>
<td>86.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig9091</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-128</td>
<td>385</td>
<td>57.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>swollenin [Trichoderma asperellum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig900</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-140</td>
<td>380</td>
<td>61.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>beta-xyllosidase/endoglucanase [Cytophaga hutchinsonii ATCC 33406](gi</td>
<td>110280324)[gb</td>
<td>ABG58510.1] CHU large protein; beta-xyllosidase/endoglucanase [Cytophaga hutchinsonii ATCC 33406]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3933</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-21</td>
<td>166</td>
<td>39.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>swollenin [Hypocrea jecorina]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig5849</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.00E-26</td>
<td>104</td>
<td>54.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>celllobiohydrolase II-like cellulase Cell [Orpinomyces sp. FC-2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig29074</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-174</td>
<td>393</td>
<td>75.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3715</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.00E-81</td>
<td>303</td>
<td>52.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>swollenin [Hypocrea jecorina]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig25059</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-83</td>
<td>227</td>
<td>68.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cip2 [Fibrobacter succinogenes subsp. succinogenes S85](gi</td>
<td>261373189)[gb</td>
<td>ACX75934.1] Cip2 [Fibrobacter succinogenes subsp. succinogenes S85]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig3876</td>
<td></td>
<td></td>
<td></td>
<td>gi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00E-400</td>
<td>52.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>swollenin [Trichoderma asperellum]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contig</td>
<td>Accession</td>
<td>Description</td>
<td>Start</td>
<td>End</td>
<td>Gene ID</td>
<td>E-value</td>
<td>Identity</td>
<td>Function</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>-------------</td>
<td>-------</td>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>Contig30515</td>
<td>CBM10; DUF303</td>
<td><em>Swollenin</em></td>
<td>1129</td>
<td>614</td>
<td>AAB69090</td>
<td>4.00E-94</td>
<td>377</td>
<td>48.0</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig5283</td>
<td>CBM10; CE15</td>
<td>DUF303</td>
<td>815</td>
<td>322</td>
<td>YP_677850</td>
<td>2.00E-60</td>
<td>189</td>
<td>58.2</td>
<td>beta-xilosidase/endoglucanase [Cytophaga hutchinsonii ATCC 33406]</td>
</tr>
<tr>
<td>Contig21613</td>
<td>CBM10; Swollenin</td>
<td>DUF303</td>
<td>1773</td>
<td>2760</td>
<td>CAB92328</td>
<td>1.00E-109</td>
<td>451</td>
<td>48.1</td>
<td>swollenin [Hypocrea jecorina]</td>
</tr>
<tr>
<td>Contig26668</td>
<td>CBM10; Ricin_B_lectin</td>
<td>DUF303</td>
<td>1134</td>
<td>607</td>
<td>CBK75021</td>
<td>1.00E-39</td>
<td>151</td>
<td>51.7</td>
<td>Beta-1,4-xylanase [Butyrivibrio fibrisolvens 16/4]</td>
</tr>
<tr>
<td>Contig23671</td>
<td>CBM10; Swollenin</td>
<td>DUF303</td>
<td>1094</td>
<td>549</td>
<td>CAB92328</td>
<td>3.00E-39</td>
<td>218</td>
<td>45.9</td>
<td>swollenin [Hypocrea jecorina]</td>
</tr>
<tr>
<td>Contig11249</td>
<td>CBM10; Polysacc_deac_1</td>
<td>DUF303</td>
<td>1357</td>
<td>686</td>
<td>YP_003250045</td>
<td>1.00E-142</td>
<td>332</td>
<td>73.5</td>
<td>polysaccharide deacetylase [Fibrobacter succinogenes subsp. succinogenes S85]</td>
</tr>
<tr>
<td>Contig14319</td>
<td>CBM10; Ricin_B_lectin</td>
<td>DUF303</td>
<td>791</td>
<td>299</td>
<td>CBK74914</td>
<td>5.00E-37</td>
<td>154</td>
<td>53.9</td>
<td>Enterochelin esterase and related enzymes [Butyrivibrio fibrisolvens 16/4]</td>
</tr>
<tr>
<td>Contig29530</td>
<td>CBM10; DUF303</td>
<td>DUF303</td>
<td>1201</td>
<td>1925</td>
<td>AAB69090</td>
<td>1.00E-141</td>
<td>374</td>
<td>64.4</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig11158</td>
<td>CBM10; Ricin_B_lectin</td>
<td>DUF303</td>
<td>683</td>
<td>161</td>
<td>ABY52795</td>
<td>1.00E-28</td>
<td>88</td>
<td>61.4</td>
<td>endo-1,4-beta-xylanase [Piromyces communis]</td>
</tr>
<tr>
<td>Contig22082</td>
<td>CBM10; DUF303</td>
<td>DUF303</td>
<td>1055</td>
<td>758</td>
<td>AAB69090</td>
<td>1.00E-111</td>
<td>373</td>
<td>53.6</td>
<td>acetylxylan esterase [Neocallimastix patriciarum]</td>
</tr>
<tr>
<td>Contig3201</td>
<td>CBM10; Pfam-B_184</td>
<td>DUF303</td>
<td>2196</td>
<td>1316</td>
<td>ZP_05498108</td>
<td>1.00E-84</td>
<td>502</td>
<td>37.1</td>
<td>conserved hypothetical protein [Clostridium papyrosolvens DSM 2782]</td>
</tr>
</tbody>
</table>

*gi* indicates GenBank accession numbers; *ref* indicates reference accession numbers.
<table>
<thead>
<tr>
<th>Contig</th>
<th>CBM/CE</th>
<th>AAF/YP</th>
<th>gi/gb</th>
<th>Ref</th>
<th>E-value</th>
<th>Length</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contig1549</td>
<td>CBM10; Esterase</td>
<td>1738 35</td>
<td>AAF70241</td>
<td>gi</td>
<td>7839348[gb] AAF70241.1 AF164351.1</td>
<td>1.00E-175</td>
<td>450 64.2</td>
</tr>
<tr>
<td>Contig29391</td>
<td>CBM10; CE15</td>
<td>1623 6258 416</td>
<td>YP_003250</td>
<td>gi</td>
<td>261416733 ref[YP_003250 0416.1]</td>
<td>1.00E-170</td>
<td>381 76.6</td>
</tr>
<tr>
<td>Contig15160</td>
<td>CBM10; Polysacc_deac_1</td>
<td>1409 2791 416</td>
<td>YP_003250 045</td>
<td>gi</td>
<td>261416362 ref[YP_003250 0045.1]</td>
<td>1.00E-152</td>
<td>401 67.3</td>
</tr>
<tr>
<td>Contig7471</td>
<td>CBM10; CE15</td>
<td>1565 3404 416</td>
<td>YP_003250 045</td>
<td>gi</td>
<td>261416733 ref[YP_003250 0416.1]</td>
<td>1.00E-168</td>
<td>381 75.9</td>
</tr>
<tr>
<td>Contig23065</td>
<td>CBM10; Polysacc_deac_1</td>
<td>1544 999 416</td>
<td>YP_003250 045</td>
<td>gi</td>
<td>261416362 ref[YP_003250 0045.1]</td>
<td>1.00E-141</td>
<td>333 73.6</td>
</tr>
<tr>
<td>Contig3750</td>
<td>CBM10; Polysacc_deac_1</td>
<td>1487 9636 416</td>
<td>YP_003250 045</td>
<td>gi</td>
<td>261416362 ref[YP_003250 0045.1]</td>
<td>1.00E-152</td>
<td>331 76.7</td>
</tr>
<tr>
<td>Contig28807</td>
<td>CBM10; CE15</td>
<td>1617 4605 416</td>
<td>YP_003250 045</td>
<td>gi</td>
<td>261416733 ref[YP_003250 0416.1]</td>
<td>1.00E-156</td>
<td>380 71.6</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM10;</td>
<td>Start</td>
<td>End</td>
<td>gi/Ref</td>
<td>E-value</td>
<td>Length</td>
<td>Homology</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------------</td>
<td>-------</td>
<td>------</td>
<td>----------------------------</td>
<td>---------</td>
<td>--------</td>
<td>-----------</td>
</tr>
<tr>
<td>5391</td>
<td>CBM10; Ricin_B_lectin</td>
<td>1086</td>
<td>580</td>
<td>YP_003250020</td>
<td>1.00E-48</td>
<td>137</td>
<td>65.7</td>
</tr>
<tr>
<td>21252</td>
<td>CBM10;</td>
<td>1456</td>
<td>1459</td>
<td>AAC14690</td>
<td>1.00E-114</td>
<td>273</td>
<td>72.9</td>
</tr>
<tr>
<td>24180</td>
<td>CBM10; Chitin_bind_1</td>
<td>1682</td>
<td>1144</td>
<td>XP_001560867</td>
<td>7.00E-19</td>
<td>102</td>
<td>47.1</td>
</tr>
<tr>
<td>12787</td>
<td>CBM10; Esterase</td>
<td>1804</td>
<td>3399</td>
<td>AAF70241</td>
<td>1.00E-165</td>
<td>455</td>
<td>61.3</td>
</tr>
<tr>
<td>4845</td>
<td>CBM10; Ricin_B_lectin</td>
<td>613</td>
<td>187</td>
<td>NP_149235</td>
<td>3.00E-31</td>
<td>142</td>
<td>47.9</td>
</tr>
<tr>
<td>8656</td>
<td>CBM10; Ricin_B_lectin</td>
<td>680</td>
<td>186</td>
<td>NP_149235</td>
<td>9.00E-40</td>
<td>148</td>
<td>55.4</td>
</tr>
<tr>
<td>156</td>
<td>CBM10; Ricin_B_lectin</td>
<td>612</td>
<td>182</td>
<td>CBK75021</td>
<td>4.00E-34</td>
<td>154</td>
<td>47.4</td>
</tr>
<tr>
<td>4849</td>
<td>CBM10; Swollenin</td>
<td>605</td>
<td>340</td>
<td>CAB92328</td>
<td>7.00E-18</td>
<td>116</td>
<td>42.2</td>
</tr>
<tr>
<td>1517</td>
<td>CBM10; Esterase</td>
<td>1037</td>
<td>83</td>
<td>ACZ98648</td>
<td>3.00E-276</td>
<td>47.8</td>
<td>esterase [Cellulosilyticum]</td>
</tr>
<tr>
<td>Contig</td>
<td>Description</td>
<td>Accession</td>
<td>Start</td>
<td>End</td>
<td>Length</td>
<td>Bit Score</td>
<td>E Value</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------</td>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>Contig771</td>
<td>CBM10; Esterase</td>
<td>gi</td>
<td>280977861</td>
<td>gb</td>
<td>ACZ98648 .1</td>
<td>65</td>
<td>6.00E-59</td>
</tr>
<tr>
<td>Contig21459</td>
<td>Chitin bind_1; Polysacc deac_1</td>
<td>gi</td>
<td>54323105</td>
<td>ref</td>
<td>XP_00156 867</td>
<td>3.00E-40</td>
<td>371</td>
</tr>
<tr>
<td>Contig27125</td>
<td>Chitin bind_1; Polysacc deac_1</td>
<td>gi</td>
<td>156064515</td>
<td>ref</td>
<td>XP_00159 179</td>
<td>7.00E-47</td>
<td>378</td>
</tr>
<tr>
<td>Contig6919</td>
<td>Chitin bind_1; Polysacc deac_1</td>
<td>gi</td>
<td>156064515</td>
<td>ref</td>
<td>XP_00159 179</td>
<td>6.00E-44</td>
<td>312</td>
</tr>
<tr>
<td>Contig442</td>
<td>Chitin bind_1; Polysacc deac_1</td>
<td>gi</td>
<td>156064515</td>
<td>ref</td>
<td>XP_00159 179</td>
<td>3.00E-45</td>
<td>486</td>
</tr>
<tr>
<td>Contig1675</td>
<td>Chitin bind_1; Polysacc deac_1</td>
<td>gi</td>
<td>46116608</td>
<td>r</td>
<td>XP_384322 1</td>
<td>3.00E-51</td>
<td>372</td>
</tr>
<tr>
<td>Contig28694</td>
<td>Chitin bind_1; Polysacc deac_1</td>
<td>gi</td>
<td>156064515</td>
<td>ref</td>
<td>XP_00159 179</td>
<td>5.00E-39</td>
<td>397</td>
</tr>
<tr>
<td>Contig29559</td>
<td>Chitin bind_1; Polysacc deac_1</td>
<td>gi</td>
<td>154323105</td>
<td>ref</td>
<td>XP_00156 867</td>
<td>5.00E-41</td>
<td>376</td>
</tr>
<tr>
<td>Contig</td>
<td>Chitin_bind_1; Polysacc_deac_1</td>
<td>921</td>
<td>348</td>
<td>XP_001912 680</td>
<td>gi</td>
<td>171695512</td>
<td>ref</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------</td>
<td>------</td>
<td>-------</td>
<td>---------------</td>
<td>------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Contig</td>
<td>Chitin_bind_1; Polysacc_deac_1</td>
<td>585</td>
<td>158</td>
<td>XP_001394 100</td>
<td>gi</td>
<td>145243114</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM20; PL6</td>
<td>1772</td>
<td>1446</td>
<td>ZP_018529 68</td>
<td>gi</td>
<td>149174341</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM20; PL6</td>
<td>1733</td>
<td>5701</td>
<td>ZP_018529 68</td>
<td>gi</td>
<td>149174341</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM20; PL6</td>
<td>1777</td>
<td>20162</td>
<td>ZP_018529 68</td>
<td>gi</td>
<td>149174341</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM20; PL6</td>
<td>1676</td>
<td>5416</td>
<td>ZP_018529 68</td>
<td>gi</td>
<td>149174341</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM48; Alpha-amylase_C</td>
<td>2217</td>
<td>2906</td>
<td>XP_002470 098</td>
<td>gi</td>
<td>242208495</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>Carb_bind; Pfam-B_1434</td>
<td>1589</td>
<td>1775</td>
<td>ZP_020248 95</td>
<td>gi</td>
<td>154482447</td>
<td>ref</td>
</tr>
<tr>
<td>Contig</td>
<td>GH94; GT36_AF</td>
<td>2521</td>
<td>15361</td>
<td>ADE82788</td>
<td>gi</td>
<td>294473399</td>
<td>0</td>
</tr>
<tr>
<td>---------</td>
<td>---------------</td>
<td>------</td>
<td>-------</td>
<td>----------</td>
<td>----------------</td>
<td>---</td>
<td>-----</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM29; CBM10</td>
<td>1338</td>
<td>1764</td>
<td>P55296</td>
<td>gi</td>
<td>1708917</td>
<td>1.00E-25</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM29; CBM10</td>
<td>969</td>
<td>784</td>
<td>AAK20910</td>
<td>gi</td>
<td>13446353</td>
<td>1.00E-78</td>
</tr>
<tr>
<td>Contig</td>
<td>CBM29; CBM1</td>
<td>1014</td>
<td>576</td>
<td>1W90</td>
<td>gi</td>
<td>62738220</td>
<td>8.00E-05</td>
</tr>
</tbody>
</table>

Note: * Possible Full Length ORFs.
Table S.3 ORFs with predicted CAZy modules and their expression level represented by FPKMs for *Aneromyces mucronatus* grown on four different carbon sources and the clade they were grouped into as illustrated in Figure 4.12.

<table>
<thead>
<tr>
<th>ORF</th>
<th>Length (bp)</th>
<th>CAZy Domains</th>
<th>Clade</th>
<th>FPKM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GCS</td>
</tr>
<tr>
<td>AmuTC623</td>
<td>957</td>
<td>CE4</td>
<td>A</td>
<td>1412.9</td>
</tr>
<tr>
<td>AmuTC72</td>
<td>1428</td>
<td>CBM10/CBM29_Blast</td>
<td>A</td>
<td>141.7</td>
</tr>
<tr>
<td>AmuTC8</td>
<td>1020</td>
<td>GH6</td>
<td>A</td>
<td>1035.6</td>
</tr>
<tr>
<td>AmuTC832</td>
<td>1638</td>
<td>CBM13/CBM10</td>
<td>A</td>
<td>337.0</td>
</tr>
<tr>
<td>AmuTC99</td>
<td>984</td>
<td>GH43</td>
<td>A</td>
<td>6738.8</td>
</tr>
<tr>
<td>AmuVC10744</td>
<td>708</td>
<td>CBM10/GH48</td>
<td>A</td>
<td>259.9</td>
</tr>
<tr>
<td>AmuVC219</td>
<td>1101</td>
<td>CBM10/GH45</td>
<td>A</td>
<td>375.5</td>
</tr>
<tr>
<td>AmuVC9693</td>
<td>1629</td>
<td>CBM13/CBM10</td>
<td>A</td>
<td>199.3</td>
</tr>
<tr>
<td>AmuTC459</td>
<td>1539</td>
<td>CE15/CBM10</td>
<td>B</td>
<td>13.1</td>
</tr>
<tr>
<td>AmuTC686_seq4</td>
<td>1182</td>
<td>CBM10</td>
<td>B</td>
<td>19.6</td>
</tr>
<tr>
<td>AmuVC11851</td>
<td>1314</td>
<td>CBM10/GH10</td>
<td>B</td>
<td>17.7</td>
</tr>
<tr>
<td>AmuVC11854</td>
<td>1164</td>
<td>CE6/CBM10</td>
<td>B</td>
<td>26.5</td>
</tr>
<tr>
<td>AmuVC1295</td>
<td>1215</td>
<td>CBM10/GH11</td>
<td>B</td>
<td>6.9</td>
</tr>
<tr>
<td>AmuVC1489</td>
<td>1185</td>
<td>CBM10</td>
<td>B</td>
<td>7.0</td>
</tr>
<tr>
<td>AmuVC732</td>
<td>1377</td>
<td>CBM10/GH6</td>
<td>B</td>
<td>23.7</td>
</tr>
<tr>
<td>AmuVN325</td>
<td>1164</td>
<td>CBM10</td>
<td>B</td>
<td>21.8</td>
</tr>
<tr>
<td>AmuTC1006</td>
<td>1059</td>
<td>CBM10</td>
<td>C</td>
<td>14.9</td>
</tr>
<tr>
<td>AmuTC149</td>
<td>1635</td>
<td>GH117/CBM6/GH43</td>
<td>C</td>
<td>330.5</td>
</tr>
<tr>
<td>AmuTC1541Seq1</td>
<td>5034</td>
<td>GDE_C</td>
<td>C</td>
<td>95.8</td>
</tr>
<tr>
<td>AmuTC1644</td>
<td>1113</td>
<td>CE4</td>
<td>C</td>
<td>272.2</td>
</tr>
<tr>
<td>AmuTC169</td>
<td>1506</td>
<td>CBM10</td>
<td>C</td>
<td>48.1</td>
</tr>
<tr>
<td>AmuTC169_seq3</td>
<td>1521</td>
<td>CBM10</td>
<td>C</td>
<td>25.8</td>
</tr>
<tr>
<td>AmuTC1949</td>
<td>3165</td>
<td>CBM13/CBM10/GH95_Blast</td>
<td>C</td>
<td>12.8</td>
</tr>
<tr>
<td>AmuTC2165</td>
<td>1671</td>
<td>GH43</td>
<td>C</td>
<td>161.0</td>
</tr>
<tr>
<td>AmuTC223</td>
<td>1476</td>
<td>GH13</td>
<td>C</td>
<td>168.1</td>
</tr>
<tr>
<td>AmuTC290</td>
<td>5250</td>
<td>CBM10</td>
<td>C</td>
<td>45.7</td>
</tr>
<tr>
<td>AmuTC3141</td>
<td>666</td>
<td>CBM13/CBM10</td>
<td>C</td>
<td>51.1</td>
</tr>
<tr>
<td>AmuTC352</td>
<td>822</td>
<td>GH11</td>
<td>C</td>
<td>473.9</td>
</tr>
<tr>
<td>AmuTC4</td>
<td>1926</td>
<td>Swollenin/CBM13/CBM10</td>
<td>C</td>
<td>123.6</td>
</tr>
<tr>
<td>AmuTC533</td>
<td>2943</td>
<td>CBM10/GH10</td>
<td>C</td>
<td>75.1</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>GCS Spelt</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>AmuTC585</td>
<td>3891</td>
<td>GH3_C/CBM10/GH6/GH3</td>
<td>C</td>
<td>121.2</td>
</tr>
<tr>
<td>AmuTC66</td>
<td>1701</td>
<td>CBM10/GH5</td>
<td>C</td>
<td>148.1</td>
</tr>
<tr>
<td>AmuTC686_seq3</td>
<td>1338</td>
<td>CE4/CBM10</td>
<td>C</td>
<td>73.6</td>
</tr>
<tr>
<td>AmuTC858</td>
<td>2106</td>
<td>CBM13/CBM10/GH39</td>
<td>C</td>
<td>26.0</td>
</tr>
<tr>
<td>AmuTC910</td>
<td>1209</td>
<td>GH117/CBM10/GH43</td>
<td>C</td>
<td>98.0</td>
</tr>
<tr>
<td>AmuVC10013</td>
<td>840</td>
<td>CBM10</td>
<td>C</td>
<td>52.7</td>
</tr>
<tr>
<td>AmuVC10015</td>
<td>1692</td>
<td>CE1</td>
<td>C</td>
<td>173.4</td>
</tr>
<tr>
<td>AmuVC10033</td>
<td>1557</td>
<td>CBM10</td>
<td>C</td>
<td>66.5</td>
</tr>
<tr>
<td>AmuVC10141</td>
<td>1014</td>
<td>CBM10</td>
<td>C</td>
<td>48.2</td>
</tr>
<tr>
<td>AmuVC10249</td>
<td>2664</td>
<td>CBM10/CBM1</td>
<td>C</td>
<td>184.1</td>
</tr>
<tr>
<td>AmuVC10358</td>
<td>945</td>
<td>CBM10</td>
<td>C</td>
<td>23.2</td>
</tr>
<tr>
<td>AmuVC11828</td>
<td>2589</td>
<td>CBM10/GH3</td>
<td>C</td>
<td>149.1</td>
</tr>
<tr>
<td>AmuVC11850</td>
<td>885</td>
<td>GH1</td>
<td>C</td>
<td>483.8</td>
</tr>
<tr>
<td>AmuVC11861</td>
<td>2121</td>
<td>CBM13/CBM10/CBM6/GH43/CBM36_Blast</td>
<td>C</td>
<td>93.0</td>
</tr>
<tr>
<td>AmuVC11927</td>
<td>936</td>
<td>CE6</td>
<td>C</td>
<td>21.8</td>
</tr>
<tr>
<td>AmuVC1201</td>
<td>1596</td>
<td>CBM6/GH43</td>
<td>C</td>
<td>215.1</td>
</tr>
<tr>
<td>AmuVC12062</td>
<td>1590</td>
<td>GH114</td>
<td>C</td>
<td>150.8</td>
</tr>
<tr>
<td>AmuVC12088</td>
<td>2103</td>
<td>CBM13/CBM10/GH39</td>
<td>C</td>
<td>30.3</td>
</tr>
<tr>
<td>AmuVC12166</td>
<td>1035</td>
<td>CBM10/GH48</td>
<td>C</td>
<td>56.9</td>
</tr>
<tr>
<td>AmuVC126</td>
<td>1248</td>
<td>GH117/CBM10/GH43</td>
<td>C</td>
<td>43.3</td>
</tr>
<tr>
<td>AmuVC1308</td>
<td>876</td>
<td>CE4</td>
<td>C</td>
<td>175.6</td>
</tr>
<tr>
<td>AmuVC1394</td>
<td>1269</td>
<td>CE4/CBM18</td>
<td>C</td>
<td>73.0</td>
</tr>
<tr>
<td>AmuVC1590</td>
<td>1464</td>
<td>CBM10</td>
<td>C</td>
<td>14.6</td>
</tr>
<tr>
<td>AmuVC172</td>
<td>939</td>
<td>CE4</td>
<td>C</td>
<td>103.0</td>
</tr>
<tr>
<td>AmuVC1910</td>
<td>1899</td>
<td>CBM10</td>
<td>C</td>
<td>293.0</td>
</tr>
<tr>
<td>AmuVC2077</td>
<td>1275</td>
<td>CBM10</td>
<td>C</td>
<td>68.8</td>
</tr>
<tr>
<td>AmuVC226</td>
<td>645</td>
<td>CBM10</td>
<td>C</td>
<td>267.6</td>
</tr>
<tr>
<td>AmuVC2304</td>
<td>1443</td>
<td>CBM10/CBM10/GH45</td>
<td>C</td>
<td>72.1</td>
</tr>
<tr>
<td>AmuVC2325</td>
<td>1380</td>
<td>CBM10/GH6</td>
<td>C</td>
<td>89.6</td>
</tr>
<tr>
<td>AmuVC2417</td>
<td>1263</td>
<td>CBM10/GH26</td>
<td>C</td>
<td>44.7</td>
</tr>
<tr>
<td>AmuVC306</td>
<td>3843</td>
<td>CBM26/GH31</td>
<td>C</td>
<td>53.6</td>
</tr>
<tr>
<td>AmuVC309</td>
<td>1725</td>
<td>Swollenin/CBM10</td>
<td>C</td>
<td>167.7</td>
</tr>
<tr>
<td>AmuVC3374</td>
<td>1527</td>
<td>CBM1</td>
<td>C</td>
<td>136.4</td>
</tr>
<tr>
<td>AmuVC34</td>
<td>1734</td>
<td>Swollenin/CBM10</td>
<td>C</td>
<td>145.9</td>
</tr>
<tr>
<td>AmuVC427</td>
<td>1059</td>
<td>CBM10/GH45</td>
<td>C</td>
<td>131.2</td>
</tr>
<tr>
<td>AmuVC4389</td>
<td>1122</td>
<td>CBM10/GH11</td>
<td>C</td>
<td>14.8</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>GCS</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>------------------------</td>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td>AmuVC557</td>
<td>1614</td>
<td>CBM13/CBM10</td>
<td>C</td>
<td>28.7</td>
</tr>
<tr>
<td>AmuVC567</td>
<td>1470</td>
<td>CBM10/GH45</td>
<td>C</td>
<td>57.9</td>
</tr>
<tr>
<td>AmuVC58</td>
<td>2079</td>
<td>CBM10/CBM1</td>
<td>C</td>
<td>202.7</td>
</tr>
<tr>
<td>AmuVC8269</td>
<td>1494</td>
<td>CBM10/CBM1/GH45</td>
<td>C</td>
<td>33.5</td>
</tr>
<tr>
<td>AmuVC8468</td>
<td>2355</td>
<td>CBM10/GH9</td>
<td>C</td>
<td>114.7</td>
</tr>
<tr>
<td>AmuVC88</td>
<td>594</td>
<td>CBM13/CBM10</td>
<td>C</td>
<td>16.3</td>
</tr>
<tr>
<td>AmuVC9698</td>
<td>1263</td>
<td>GH18</td>
<td>C</td>
<td>280.0</td>
</tr>
<tr>
<td>AmuVC9715</td>
<td>2697</td>
<td>CBM10/GH9</td>
<td>C</td>
<td>102.7</td>
</tr>
<tr>
<td>AmuVC9724</td>
<td>2337</td>
<td>CBM10/GH9</td>
<td>C</td>
<td>105.3</td>
</tr>
<tr>
<td>AmuVC9753</td>
<td>2094</td>
<td>CBM10</td>
<td>C</td>
<td>43.0</td>
</tr>
<tr>
<td>AmuVC9826</td>
<td>1599</td>
<td>CBM13</td>
<td>C</td>
<td>84.3</td>
</tr>
<tr>
<td>AmuVC9956</td>
<td>1698</td>
<td>Swollenin/CBM10</td>
<td>C</td>
<td>64.1</td>
</tr>
<tr>
<td>AmuTC1887</td>
<td>3573</td>
<td>CBM13</td>
<td>D</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuTC991</td>
<td>525</td>
<td>CE4</td>
<td>D</td>
<td>2.8</td>
</tr>
<tr>
<td>AmuVC1058</td>
<td>555</td>
<td>CBM1</td>
<td>D</td>
<td>6.3</td>
</tr>
<tr>
<td>AmuVC11071</td>
<td>1398</td>
<td>CBM10</td>
<td>D</td>
<td>0.9</td>
</tr>
<tr>
<td>AmuVC11847</td>
<td>1791</td>
<td>CBM10/GH8</td>
<td>D</td>
<td>2.9</td>
</tr>
<tr>
<td>AmuVC12024</td>
<td>1170</td>
<td>CBM18</td>
<td>D</td>
<td>0.6</td>
</tr>
<tr>
<td>AmuVC12049</td>
<td>1209</td>
<td>CBM10</td>
<td>D</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuVC3541</td>
<td>1485</td>
<td>CBM10</td>
<td>D</td>
<td>2.2</td>
</tr>
<tr>
<td>AmuVC3669</td>
<td>1575</td>
<td>CBM10</td>
<td>D</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuVC523</td>
<td>1530</td>
<td>CBM10</td>
<td>D</td>
<td>1.9</td>
</tr>
<tr>
<td>AmuVC852</td>
<td>1428</td>
<td>CBM10/CBM1</td>
<td>D</td>
<td>2.9</td>
</tr>
<tr>
<td>AmuTC11219</td>
<td>492</td>
<td>CBM50</td>
<td>E</td>
<td>6.7</td>
</tr>
<tr>
<td>AmuTC11272</td>
<td>954</td>
<td>CE10</td>
<td>E</td>
<td>22.9</td>
</tr>
<tr>
<td>AmuTC1198</td>
<td>1434</td>
<td>CBM10/CBM1/GH45</td>
<td>E</td>
<td>41.4</td>
</tr>
<tr>
<td>AmuTC1213</td>
<td>3591</td>
<td>CBM21</td>
<td>E</td>
<td>30.7</td>
</tr>
<tr>
<td>AmuTC1303</td>
<td>654</td>
<td>CE10</td>
<td>E</td>
<td>32.2</td>
</tr>
<tr>
<td>AmuTC17890</td>
<td>744</td>
<td>CE4</td>
<td>E</td>
<td>44.8</td>
</tr>
<tr>
<td>AmuTC2143</td>
<td>1830</td>
<td>CBM10</td>
<td>E</td>
<td>74.7</td>
</tr>
<tr>
<td>AmuTC2171</td>
<td>1761</td>
<td>GH67</td>
<td>E</td>
<td>125.8</td>
</tr>
<tr>
<td>AmuTC290</td>
<td>2376</td>
<td>CBM10/GH9</td>
<td>E</td>
<td>220.6</td>
</tr>
<tr>
<td>AmuTC2911</td>
<td>1452</td>
<td>GH43</td>
<td>E</td>
<td>24.6</td>
</tr>
<tr>
<td>AmuTC3045</td>
<td>1464</td>
<td>CE2 or CE3</td>
<td>E</td>
<td>14.5</td>
</tr>
<tr>
<td>AmuTC3201</td>
<td>1011</td>
<td>CE12/CBM1</td>
<td>E</td>
<td>19.1</td>
</tr>
<tr>
<td>AmuTC3825</td>
<td>957</td>
<td>CE4</td>
<td>E</td>
<td>15.1</td>
</tr>
<tr>
<td>AmuTC3998</td>
<td>456</td>
<td>CBM1</td>
<td>E</td>
<td>92.0</td>
</tr>
<tr>
<td>AmuTC408</td>
<td>681</td>
<td>CBM10</td>
<td>E</td>
<td>20.3</td>
</tr>
<tr>
<td>AmuTC408_seq2</td>
<td>642</td>
<td>CBM10</td>
<td>E</td>
<td>22.1</td>
</tr>
<tr>
<td>AmuTC4473_seq2</td>
<td>771</td>
<td>GH16</td>
<td>E</td>
<td>20.7</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>FPKM</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------</td>
<td>--------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GCS</td>
</tr>
<tr>
<td>AmuTC4587</td>
<td>1311</td>
<td>CBM10</td>
<td>E</td>
<td>43.9</td>
</tr>
<tr>
<td>AmuTC471</td>
<td>2940</td>
<td>GH31</td>
<td>E</td>
<td>21.5</td>
</tr>
<tr>
<td>AmuTC4860</td>
<td>2175</td>
<td>CBM10</td>
<td>E</td>
<td>32.6</td>
</tr>
<tr>
<td>AmuTC5285</td>
<td>963</td>
<td>CE10</td>
<td>E</td>
<td>61.0</td>
</tr>
<tr>
<td>AmuTC7281</td>
<td>2238</td>
<td>PL3</td>
<td>E</td>
<td>29.6</td>
</tr>
<tr>
<td>AmuTC9476</td>
<td>1122</td>
<td>CE4</td>
<td>E</td>
<td>37.1</td>
</tr>
<tr>
<td>AmuVC10002</td>
<td>2334</td>
<td>CBM10/GH9</td>
<td>E</td>
<td>12.9</td>
</tr>
<tr>
<td>AmuVC1002</td>
<td>2454</td>
<td>CBM10</td>
<td>E</td>
<td>53.0</td>
</tr>
<tr>
<td>AmuVC10088</td>
<td>1611</td>
<td>CE1</td>
<td>E</td>
<td>15.4</td>
</tr>
<tr>
<td>AmuVC10228</td>
<td>1452</td>
<td>CBM10/GH6</td>
<td>E</td>
<td>23.1</td>
</tr>
<tr>
<td>AmuVC10296</td>
<td>1068</td>
<td>GH114</td>
<td>E</td>
<td>18.6</td>
</tr>
<tr>
<td>AmuVC1032</td>
<td>1407</td>
<td>CBM10/GH5</td>
<td>E</td>
<td>34.0</td>
</tr>
<tr>
<td>AmuVC10350</td>
<td>1548</td>
<td>CE1</td>
<td>E</td>
<td>29.5</td>
</tr>
<tr>
<td>AmuVC10375</td>
<td>1506</td>
<td>GH11</td>
<td>E</td>
<td>17.6</td>
</tr>
<tr>
<td>AmuVC10459</td>
<td>3141</td>
<td>CBM21</td>
<td>E</td>
<td>229.0</td>
</tr>
<tr>
<td>AmuVC10543</td>
<td>1065</td>
<td>CBM10</td>
<td>E</td>
<td>51.8</td>
</tr>
<tr>
<td>AmuVC10681</td>
<td>2484</td>
<td>CBM10/CBM1</td>
<td>E</td>
<td>20.6</td>
</tr>
<tr>
<td>AmuVC1082</td>
<td>1596</td>
<td>GH117/CBM6/GH4/3</td>
<td>E</td>
<td>100.2</td>
</tr>
<tr>
<td>AmuVC1106</td>
<td>1341</td>
<td>GH114</td>
<td>E</td>
<td>114.1</td>
</tr>
<tr>
<td>AmuVC1110</td>
<td>1467</td>
<td>GH117/CBM6/GH4/3/CBM36_Blast</td>
<td>E</td>
<td>52.7</td>
</tr>
<tr>
<td>AmuVC11161</td>
<td>1641</td>
<td>CBM13/CBM10</td>
<td>E</td>
<td>49.0</td>
</tr>
<tr>
<td>AmuVC11366</td>
<td>1974</td>
<td>CBM18</td>
<td>E</td>
<td>293.7</td>
</tr>
<tr>
<td>AmuVC1148</td>
<td>2760</td>
<td>CBM21</td>
<td>E</td>
<td>28.0</td>
</tr>
<tr>
<td>AmuVC12040</td>
<td>2004</td>
<td>CBM10</td>
<td>E</td>
<td>14.1</td>
</tr>
<tr>
<td>AmuVC12041</td>
<td>933</td>
<td>CE4</td>
<td>E</td>
<td>66.1</td>
</tr>
<tr>
<td>AmuVC12163</td>
<td>1587</td>
<td>CBM10/GH5</td>
<td>E</td>
<td>90.4</td>
</tr>
<tr>
<td>AmuVC12184</td>
<td>1548</td>
<td>CBM10</td>
<td>E</td>
<td>47.3</td>
</tr>
<tr>
<td>AmuVC12245</td>
<td>1407</td>
<td>CBM10/GH4</td>
<td>E</td>
<td>29.4</td>
</tr>
<tr>
<td>AmuVC1268</td>
<td>1041</td>
<td>CBM18</td>
<td>E</td>
<td>21.4</td>
</tr>
<tr>
<td>AmuVC1464</td>
<td>1689</td>
<td>Swollenin/CBM10</td>
<td>E</td>
<td>29.5</td>
</tr>
<tr>
<td>AmuVC1601</td>
<td>1485</td>
<td>CE4/CBM18</td>
<td>E</td>
<td>56.5</td>
</tr>
<tr>
<td>AmuVC2112</td>
<td>1911</td>
<td>CBM10</td>
<td>E</td>
<td>141.7</td>
</tr>
<tr>
<td>AmuVC2440</td>
<td>1869</td>
<td>CBM10</td>
<td>E</td>
<td>25.2</td>
</tr>
<tr>
<td>AmuVC2668</td>
<td>1488</td>
<td>GH114</td>
<td>E</td>
<td>17.3</td>
</tr>
<tr>
<td>AmuVC2684</td>
<td>2055</td>
<td>GH13_C/CBM48/GH13</td>
<td>E</td>
<td>59.5</td>
</tr>
<tr>
<td>AmuVC279</td>
<td>1401</td>
<td>CBM10/CBM1</td>
<td>E</td>
<td>30.1</td>
</tr>
<tr>
<td>AmuVC2993</td>
<td>1554</td>
<td>CBM10/GH11</td>
<td>E</td>
<td>128.0</td>
</tr>
<tr>
<td>AmuVC338</td>
<td>1128</td>
<td>CBM10</td>
<td>E</td>
<td>63.1</td>
</tr>
<tr>
<td>AmuVC3496</td>
<td>1938</td>
<td>CBM10</td>
<td>E</td>
<td>30.4</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>GCS</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>AmuVC3673</td>
<td>1767</td>
<td>CBM10/CBM6/GH43</td>
<td>E</td>
<td>38.2</td>
</tr>
<tr>
<td>AmuVC408</td>
<td>1563</td>
<td>CE1</td>
<td>E</td>
<td>6.1</td>
</tr>
<tr>
<td>AmuVC420</td>
<td>1743</td>
<td>CBM10</td>
<td>E</td>
<td>28.3</td>
</tr>
<tr>
<td>AmuVC4510</td>
<td>744</td>
<td>CBM18</td>
<td>E</td>
<td>12.4</td>
</tr>
<tr>
<td>AmuVC4852</td>
<td>1095</td>
<td>CBM10/CBM1</td>
<td>E</td>
<td>48.9</td>
</tr>
<tr>
<td>AmuVC514</td>
<td>5130</td>
<td>GH33</td>
<td>E</td>
<td>21.1</td>
</tr>
<tr>
<td>AmuVC6446</td>
<td>2439</td>
<td>CBM10/CBM1</td>
<td>E</td>
<td>57.4</td>
</tr>
<tr>
<td>AmuVC677</td>
<td>2511</td>
<td>CBM21</td>
<td>E</td>
<td>50.2</td>
</tr>
<tr>
<td>AmuVC7056</td>
<td>1437</td>
<td>CBM18</td>
<td>E</td>
<td>69.2</td>
</tr>
<tr>
<td>AmuVC7831</td>
<td>2091</td>
<td>CBM18</td>
<td>E</td>
<td>13.0</td>
</tr>
<tr>
<td>AmuVC784</td>
<td>1731</td>
<td>CBM10/GH5</td>
<td>E</td>
<td>23.8</td>
</tr>
<tr>
<td>AmuVC8984</td>
<td>4731</td>
<td>CBM10/CBM1</td>
<td>E</td>
<td>28.5</td>
</tr>
<tr>
<td>AmuVC923</td>
<td>3573</td>
<td>CBM18/CBM13</td>
<td>E</td>
<td>32.1</td>
</tr>
<tr>
<td>AmuVC9648</td>
<td>1083</td>
<td>CBM18</td>
<td>E</td>
<td>24.6</td>
</tr>
<tr>
<td>AmuVN2600</td>
<td>1938</td>
<td>CBM10</td>
<td>E</td>
<td>35.9</td>
</tr>
<tr>
<td>AmuVN3221</td>
<td>1365</td>
<td>CBM10/GH45</td>
<td>E</td>
<td>31.2</td>
</tr>
<tr>
<td>AmuTC10428</td>
<td>1947</td>
<td>CBM10</td>
<td>F</td>
<td>0.4</td>
</tr>
<tr>
<td>AmuTC1512</td>
<td>1935</td>
<td>GH3_C/GH3</td>
<td>F</td>
<td>14.6</td>
</tr>
<tr>
<td>AmuTC17</td>
<td>2439</td>
<td>CBM10/CBM4-9/CBM4-9/GH43/CBM37_Blast</td>
<td>F</td>
<td>5.6</td>
</tr>
<tr>
<td>AmuTC1982</td>
<td>3372</td>
<td>GH115</td>
<td>F</td>
<td>6.4</td>
</tr>
<tr>
<td>AmuTC26300</td>
<td>1185</td>
<td>GH5</td>
<td>F</td>
<td>1.1</td>
</tr>
<tr>
<td>AmuTC3377</td>
<td>1269</td>
<td>CBM18</td>
<td>F</td>
<td>0.6</td>
</tr>
<tr>
<td>AmuTC3752</td>
<td>2307</td>
<td>PL4/CBM10</td>
<td>F</td>
<td>4.5</td>
</tr>
<tr>
<td>AmuTC3767_seq2</td>
<td>1014</td>
<td>CE12/CBM1</td>
<td>F</td>
<td>6.0</td>
</tr>
<tr>
<td>AmuTC4153</td>
<td>1017</td>
<td>GH18</td>
<td>F</td>
<td>2.9</td>
</tr>
<tr>
<td>AmuTC5124</td>
<td>1926</td>
<td>CBM10/GH10</td>
<td>F</td>
<td>5.1</td>
</tr>
<tr>
<td>AmuTC5594</td>
<td>2247</td>
<td>CBM48/GH13</td>
<td>F</td>
<td>5.4</td>
</tr>
<tr>
<td>AmuTC6167</td>
<td>1008</td>
<td>CBM10/CBM1</td>
<td>F</td>
<td>6.7</td>
</tr>
<tr>
<td>AmuTC9091</td>
<td>2622</td>
<td>GH27</td>
<td>F</td>
<td>9.6</td>
</tr>
<tr>
<td>AmuVC10042</td>
<td>1098</td>
<td>CE6/CBM10</td>
<td>F</td>
<td>11.6</td>
</tr>
<tr>
<td>AmuVC10050</td>
<td>1821</td>
<td>CBM13/CBM10/GH10</td>
<td>F</td>
<td>13.3</td>
</tr>
<tr>
<td>AmuVC10133</td>
<td>1974</td>
<td>CBM10</td>
<td>F</td>
<td>1.7</td>
</tr>
<tr>
<td>AmuVC10153</td>
<td>1890</td>
<td>CBM10</td>
<td>F</td>
<td>8.6</td>
</tr>
<tr>
<td>AmuVC10238</td>
<td>1635</td>
<td>CE6/CBM10</td>
<td>F</td>
<td>10.1</td>
</tr>
<tr>
<td>AmuVC10423</td>
<td>1794</td>
<td>CBM1/GH10</td>
<td>F</td>
<td>18.1</td>
</tr>
<tr>
<td>AmuVC10435</td>
<td>4830</td>
<td>GH2/Bgal_small_N/CBM10/</td>
<td>F</td>
<td>3.8</td>
</tr>
<tr>
<td>AmuVC10470</td>
<td>1626</td>
<td>CBM10/GH53</td>
<td>F</td>
<td>4.2</td>
</tr>
<tr>
<td>AmuVC105</td>
<td>1857</td>
<td>CBM10</td>
<td>F</td>
<td>4.1</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>GCS</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>AmuVC10545</td>
<td>1860</td>
<td>CE1</td>
<td>F</td>
<td>0.9</td>
</tr>
<tr>
<td>AmuVC10637</td>
<td>1755</td>
<td>GH18</td>
<td>F</td>
<td>4.5</td>
</tr>
<tr>
<td>AmuVC1073</td>
<td>1323</td>
<td>CBM18</td>
<td>F</td>
<td>2.5</td>
</tr>
<tr>
<td>AmuVC1090</td>
<td>1839</td>
<td>CBM10</td>
<td>F</td>
<td>0.6</td>
</tr>
<tr>
<td>AmuVC1177</td>
<td>1905</td>
<td>CBM18</td>
<td>F</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuVC11891</td>
<td>1878</td>
<td>CBM10</td>
<td>F</td>
<td>2.1</td>
</tr>
<tr>
<td>AmuVC12009</td>
<td>2175</td>
<td>CBM10</td>
<td>F</td>
<td>7.3</td>
</tr>
<tr>
<td>AmuVC12225</td>
<td>1488</td>
<td>CBM10</td>
<td>F</td>
<td>0.4</td>
</tr>
<tr>
<td>AmuVC1325</td>
<td>1125</td>
<td>CBM10/GH11</td>
<td>F</td>
<td>7.0</td>
</tr>
<tr>
<td>AmuVC1405</td>
<td>2478</td>
<td>GH3_C/GH3</td>
<td>F</td>
<td>3.6</td>
</tr>
<tr>
<td>AmuVC1718</td>
<td>1800</td>
<td>CBM10/GH10</td>
<td>F</td>
<td>11.0</td>
</tr>
<tr>
<td>AmuVC1852</td>
<td>1320</td>
<td>CBM18</td>
<td>F</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuVC2037</td>
<td>1125</td>
<td>CBM10</td>
<td>F</td>
<td>1.3</td>
</tr>
<tr>
<td>AmuVC2117</td>
<td>1104</td>
<td>GH45</td>
<td>F</td>
<td>3.0</td>
</tr>
<tr>
<td>AmuVC2220</td>
<td>1401</td>
<td>CBM10/GH10</td>
<td>F</td>
<td>9.0</td>
</tr>
<tr>
<td>AmuVC2424</td>
<td>1626</td>
<td>CBM1/GH43</td>
<td>F</td>
<td>0.8</td>
</tr>
<tr>
<td>AmuVC2605</td>
<td>1134</td>
<td>CBM18</td>
<td>F</td>
<td>7.4</td>
</tr>
<tr>
<td>AmuVC3357</td>
<td>3558</td>
<td>CBM10/CBM1</td>
<td>F</td>
<td>7.7</td>
</tr>
<tr>
<td>AmuVC3418</td>
<td>1389</td>
<td>CBM10</td>
<td>F</td>
<td>7.3</td>
</tr>
<tr>
<td>AmuVC3661</td>
<td>1707</td>
<td>CE1/CBM10</td>
<td>F</td>
<td>2.3</td>
</tr>
<tr>
<td>AmuVC3782</td>
<td>1668</td>
<td>CBM10/GH5</td>
<td>F</td>
<td>2.9</td>
</tr>
<tr>
<td>AmuVC4171</td>
<td>1116</td>
<td>CBM10</td>
<td>F</td>
<td>0.7</td>
</tr>
<tr>
<td>AmuVC4623</td>
<td>993</td>
<td>CBM10</td>
<td>F</td>
<td>0.1</td>
</tr>
<tr>
<td>AmuVC479</td>
<td>1554</td>
<td>CBM18</td>
<td>F</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuVC486</td>
<td>2772</td>
<td>CBM10/GH5</td>
<td>F</td>
<td>3.7</td>
</tr>
<tr>
<td>AmuVC606</td>
<td>1716</td>
<td>GH117/CBM13/CBM10/GH43</td>
<td>F</td>
<td>3.9</td>
</tr>
<tr>
<td>AmuVC808</td>
<td>2184</td>
<td>CBM10</td>
<td>F</td>
<td>4.5</td>
</tr>
<tr>
<td>AmuVC900</td>
<td>1008</td>
<td>CBM10/GH16</td>
<td>F</td>
<td>8.8</td>
</tr>
<tr>
<td>AmuVC9708</td>
<td>2307</td>
<td>CBM10/GH9</td>
<td>F</td>
<td>2.7</td>
</tr>
<tr>
<td>AmuVC9725</td>
<td>963</td>
<td>CE12/CBM1</td>
<td>F</td>
<td>5.5</td>
</tr>
<tr>
<td>AmuVC9760</td>
<td>1428</td>
<td>CBM1/GH6</td>
<td>F</td>
<td>9.9</td>
</tr>
<tr>
<td>AmuVC9781</td>
<td>4044</td>
<td>CBM26/GH31</td>
<td>F</td>
<td>7.3</td>
</tr>
<tr>
<td>AmuVN3685</td>
<td>1536</td>
<td>CBM18</td>
<td>F</td>
<td>3.4</td>
</tr>
<tr>
<td>AmuTC10803</td>
<td>1098</td>
<td>CE10</td>
<td>G</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuTC1101</td>
<td>1980</td>
<td>CBM18/CBM10/GH18</td>
<td>G</td>
<td>47.6</td>
</tr>
<tr>
<td>AmuTC110785</td>
<td>1323</td>
<td>GH64</td>
<td>G</td>
<td>0.5</td>
</tr>
<tr>
<td>AmuTC134187</td>
<td>681</td>
<td>GH37</td>
<td>G</td>
<td>8.5</td>
</tr>
<tr>
<td>AmuTC14421</td>
<td>942</td>
<td>CE10</td>
<td>G</td>
<td>0.7</td>
</tr>
<tr>
<td>AmuTC144673</td>
<td>906</td>
<td>GH24or104</td>
<td>G</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuTC14860</td>
<td>2031</td>
<td>CBM1</td>
<td>G</td>
<td>6.1</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>GCS</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------</td>
<td>-----</td>
</tr>
<tr>
<td>AmuTC1541Seq2</td>
<td>1320</td>
<td>PL1/CBM1</td>
<td>G</td>
<td>1.0</td>
</tr>
<tr>
<td>AmuTC17634</td>
<td>501</td>
<td>CE4</td>
<td>G</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuTC17666</td>
<td>804</td>
<td>CBM18</td>
<td>G</td>
<td>6.7</td>
</tr>
<tr>
<td>AmuTC17966</td>
<td>1299</td>
<td>CBM13</td>
<td>G</td>
<td>7.4</td>
</tr>
<tr>
<td>AmuTC18125</td>
<td>957</td>
<td>CE10</td>
<td>G</td>
<td>1.0</td>
</tr>
<tr>
<td>AmuTC19625</td>
<td>735</td>
<td>GH16</td>
<td>G</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuTC196606</td>
<td>549</td>
<td>CBM1/GH5</td>
<td>G</td>
<td>0.1</td>
</tr>
<tr>
<td>AmuTC20089</td>
<td>498</td>
<td>CE1</td>
<td>G</td>
<td>4.2</td>
</tr>
<tr>
<td>AmuTC20259</td>
<td>1569</td>
<td>CBM1</td>
<td>G</td>
<td>2.2</td>
</tr>
<tr>
<td>AmuTC20360_seq2</td>
<td>954</td>
<td>CBM18</td>
<td>G</td>
<td>3.2</td>
</tr>
<tr>
<td>AmuTC20682</td>
<td>1317</td>
<td>CE10</td>
<td>G</td>
<td>9.5</td>
</tr>
<tr>
<td>AmuTC20741</td>
<td>1722</td>
<td>CBM10</td>
<td>G</td>
<td>8.3</td>
</tr>
<tr>
<td>AmuTC20928</td>
<td>921</td>
<td>CE10</td>
<td>G</td>
<td>10.0</td>
</tr>
<tr>
<td>AmuTC21521</td>
<td>1101</td>
<td>GH109</td>
<td>G</td>
<td>12.6</td>
</tr>
<tr>
<td>AmuTC21906</td>
<td>1392</td>
<td>CE4</td>
<td>G</td>
<td>4.7</td>
</tr>
<tr>
<td>AmuTC22233</td>
<td>1491</td>
<td>CE10</td>
<td>G</td>
<td>9.8</td>
</tr>
<tr>
<td>AmuTC2294</td>
<td>1605</td>
<td>CBM10/GH9</td>
<td>G</td>
<td>1.5</td>
</tr>
<tr>
<td>AmuTC2294_seq2</td>
<td>1266</td>
<td>CBM10/CBM6</td>
<td>G</td>
<td>19.8</td>
</tr>
<tr>
<td>AmuTC2294_seq4</td>
<td>1437</td>
<td>CBM10/GH9</td>
<td>G</td>
<td>1.7</td>
</tr>
<tr>
<td>AmuTC23754</td>
<td>846</td>
<td>GH31</td>
<td>G</td>
<td>8.4</td>
</tr>
<tr>
<td>AmuTC24443</td>
<td>1440</td>
<td>CBM1</td>
<td>G</td>
<td>0.1</td>
</tr>
<tr>
<td>AmuTC24689</td>
<td>963</td>
<td>GH3</td>
<td>G</td>
<td>5.1</td>
</tr>
<tr>
<td>AmuTC24777</td>
<td>582</td>
<td>CE10</td>
<td>G</td>
<td>22.5</td>
</tr>
<tr>
<td>AmuTC25985</td>
<td>2688</td>
<td>GH47</td>
<td>G</td>
<td>10.4</td>
</tr>
<tr>
<td>AmuTC28640</td>
<td>498</td>
<td>CE4</td>
<td>G</td>
<td>0.4</td>
</tr>
<tr>
<td>AmuTC28884</td>
<td>483</td>
<td>CBM10</td>
<td>G</td>
<td>1.1</td>
</tr>
<tr>
<td>AmuTC29801</td>
<td>1026</td>
<td>CE10</td>
<td>G</td>
<td>11.9</td>
</tr>
<tr>
<td>AmuTC29844</td>
<td>1110</td>
<td>GH25</td>
<td>G</td>
<td>20.1</td>
</tr>
<tr>
<td>AmuTC30267</td>
<td>1617</td>
<td>GH37</td>
<td>G</td>
<td>13.6</td>
</tr>
<tr>
<td>AmuTC30342</td>
<td>729</td>
<td>CBM21</td>
<td>G</td>
<td>3.6</td>
</tr>
<tr>
<td>AmuTC30990</td>
<td>2040</td>
<td>GH20</td>
<td>G</td>
<td>6.9</td>
</tr>
<tr>
<td>AmuTC31964</td>
<td>915</td>
<td>CE4</td>
<td>G</td>
<td>6.7</td>
</tr>
<tr>
<td>AmuTC33256</td>
<td>1164</td>
<td>GH16</td>
<td>G</td>
<td>0.0</td>
</tr>
<tr>
<td>AmuTC34112</td>
<td>1047</td>
<td>GH109</td>
<td>G</td>
<td>2.8</td>
</tr>
<tr>
<td>AmuTC34284</td>
<td>459</td>
<td>CBM13</td>
<td>G</td>
<td>4.3</td>
</tr>
<tr>
<td>AmuTC3711</td>
<td>1098</td>
<td>GH5</td>
<td>G</td>
<td>8.4</td>
</tr>
<tr>
<td>AmuTC37304</td>
<td>780</td>
<td>CBM1</td>
<td>G</td>
<td>0.0</td>
</tr>
<tr>
<td>AmuTC3759</td>
<td>927</td>
<td>GH24or104</td>
<td>G</td>
<td>2.5</td>
</tr>
<tr>
<td>AmuTC37632</td>
<td>927</td>
<td>CE10</td>
<td>G</td>
<td>10.0</td>
</tr>
<tr>
<td>AmuTC3767</td>
<td>1077</td>
<td>CE12/CBM1</td>
<td>G</td>
<td>11.0</td>
</tr>
<tr>
<td>AmuTC39485</td>
<td>1464</td>
<td>GH117/CBM10/GH</td>
<td>G</td>
<td>0.4</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>GCS</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>AmuTC41693</td>
<td>954</td>
<td>CE16</td>
<td>G</td>
<td>3.5</td>
</tr>
<tr>
<td>AmuTC4220_seq2</td>
<td>777</td>
<td>GH114</td>
<td>G</td>
<td>12.2</td>
</tr>
<tr>
<td>AmuTC42946</td>
<td>1365</td>
<td>CBM18</td>
<td>G</td>
<td>8.2</td>
</tr>
<tr>
<td>AmuTC45468</td>
<td>843</td>
<td>CBM10</td>
<td>G</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuTC45752</td>
<td>795</td>
<td>CE12</td>
<td>G</td>
<td>1.3</td>
</tr>
<tr>
<td>AmuTC46314</td>
<td>2097</td>
<td>CBM13</td>
<td>G</td>
<td>12.1</td>
</tr>
<tr>
<td>AmuTC55544</td>
<td>1242</td>
<td>GH109</td>
<td>G</td>
<td>5.7</td>
</tr>
<tr>
<td>AmuTC61070</td>
<td>1956</td>
<td>CBM26/CBM25/GH13</td>
<td>G</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuTC62238</td>
<td>1449</td>
<td>CE10</td>
<td>G</td>
<td>8.6</td>
</tr>
<tr>
<td>AmuTC62584</td>
<td>831</td>
<td>GH114</td>
<td>G</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuTC64800</td>
<td>462</td>
<td>CBM50</td>
<td>G</td>
<td>2.6</td>
</tr>
<tr>
<td>AmuTC7633</td>
<td>612</td>
<td>CE4</td>
<td>G</td>
<td>0.1</td>
</tr>
<tr>
<td>AmuTC7767</td>
<td>987</td>
<td>GH114</td>
<td>G</td>
<td>44.3</td>
</tr>
<tr>
<td>AmuTC79685</td>
<td>453</td>
<td>CE2 or CE3</td>
<td>G</td>
<td>5.6</td>
</tr>
<tr>
<td>AmuTC79765</td>
<td>870</td>
<td>CE1</td>
<td>G</td>
<td>9.2</td>
</tr>
<tr>
<td>AmuTC8236</td>
<td>1500</td>
<td>GH114</td>
<td>G</td>
<td>3.8</td>
</tr>
<tr>
<td>AmuTC84759</td>
<td>1116</td>
<td>GH76</td>
<td>G</td>
<td>0.0</td>
</tr>
<tr>
<td>AmuTC8654</td>
<td>3639</td>
<td>TIG</td>
<td>G</td>
<td>21.2</td>
</tr>
<tr>
<td>AmuTC87312</td>
<td>1188</td>
<td>CBM18</td>
<td>G</td>
<td>0.1</td>
</tr>
<tr>
<td>AmuTC87421</td>
<td>825</td>
<td>CBM10</td>
<td>G</td>
<td>1.2</td>
</tr>
<tr>
<td>AmuTC8966</td>
<td>1626</td>
<td>CBM1</td>
<td>G</td>
<td>2.5</td>
</tr>
<tr>
<td>AmuTC8982</td>
<td>984</td>
<td>CE7</td>
<td>G</td>
<td>9.7</td>
</tr>
<tr>
<td>AmuTC915</td>
<td>648</td>
<td>CBM10/CBM1</td>
<td>G</td>
<td>2.3</td>
</tr>
<tr>
<td>AmuTC9150</td>
<td>2064</td>
<td>PL09</td>
<td>G</td>
<td>10.7</td>
</tr>
<tr>
<td>AmuVC10313</td>
<td>1920</td>
<td>CBM10</td>
<td>G</td>
<td>17.7</td>
</tr>
<tr>
<td>AmuVC10379</td>
<td>1305</td>
<td>CBM1</td>
<td>G</td>
<td>8.0</td>
</tr>
<tr>
<td>AmuVC10419</td>
<td>1263</td>
<td>CE4/CBM18</td>
<td>G</td>
<td>17.0</td>
</tr>
<tr>
<td>AmuVC10672</td>
<td>1293</td>
<td>CE1/CBM10</td>
<td>G</td>
<td>25.3</td>
</tr>
<tr>
<td>AmuVC1070</td>
<td>1245</td>
<td>CBM1</td>
<td>G</td>
<td>8.5</td>
</tr>
<tr>
<td>AmuVC10773</td>
<td>1560</td>
<td>CBM10</td>
<td>G</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuVC10781</td>
<td>1299</td>
<td>CBM10</td>
<td>G</td>
<td>26.0</td>
</tr>
<tr>
<td>AmuVC10945</td>
<td>1236</td>
<td>CBM10</td>
<td>G</td>
<td>4.1</td>
</tr>
<tr>
<td>AmuVC12239</td>
<td>1146</td>
<td>CBM1</td>
<td>G</td>
<td>0.5</td>
</tr>
<tr>
<td>AmuVC12453</td>
<td>2154</td>
<td>CBM10/GH9</td>
<td>G</td>
<td>7.2</td>
</tr>
<tr>
<td>AmuVC1304</td>
<td>1608</td>
<td>CBM10/GH18</td>
<td>G</td>
<td>20.3</td>
</tr>
<tr>
<td>AmuVC1484</td>
<td>1329</td>
<td>CE4/CBM50</td>
<td>G</td>
<td>25.0</td>
</tr>
<tr>
<td>AmuVC1613</td>
<td>1686</td>
<td>CBM18</td>
<td>G</td>
<td>9.0</td>
</tr>
<tr>
<td>AmuVC1727</td>
<td>1065</td>
<td>CE1/CBM10</td>
<td>G</td>
<td>10.5</td>
</tr>
<tr>
<td>AmuVC1775</td>
<td>1716</td>
<td>CBM18</td>
<td>G</td>
<td>16.2</td>
</tr>
<tr>
<td>ORF</td>
<td>Length (bp)</td>
<td>CAZy Domains</td>
<td>Clade</td>
<td>GCS</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>AmuVC1871</td>
<td>1404</td>
<td>GH3_C</td>
<td>G</td>
<td>6.0</td>
</tr>
<tr>
<td>AmuVC2008</td>
<td>930</td>
<td>CBM10</td>
<td>G</td>
<td>46.7</td>
</tr>
<tr>
<td>AmuVC21</td>
<td>2562</td>
<td>GH3_C/GH3</td>
<td>G</td>
<td>5.4</td>
</tr>
<tr>
<td>AmuVC2158</td>
<td>1263</td>
<td>CBM10</td>
<td>G</td>
<td>14.5</td>
</tr>
<tr>
<td>AmuVC2212</td>
<td>1110</td>
<td>CBM18</td>
<td>G</td>
<td>1.1</td>
</tr>
<tr>
<td>AmuVC2262</td>
<td>1578</td>
<td>CE6/CBM10</td>
<td>G</td>
<td>31.1</td>
</tr>
<tr>
<td>AmuVC2278</td>
<td>2982</td>
<td>CBM10/CBM1</td>
<td>G</td>
<td>4.8</td>
</tr>
<tr>
<td>AmuVC2376</td>
<td>825</td>
<td>CBM10</td>
<td>G</td>
<td>5.5</td>
</tr>
<tr>
<td>AmuVC2498</td>
<td>1344</td>
<td>CBM18</td>
<td>G</td>
<td>16.3</td>
</tr>
<tr>
<td>AmuVC2604</td>
<td>891</td>
<td>GH3</td>
<td>G</td>
<td>0.5</td>
</tr>
<tr>
<td>AmuVC262</td>
<td>2775</td>
<td>CBM18/CBM10</td>
<td>G</td>
<td>13.0</td>
</tr>
<tr>
<td>AmuVC2712</td>
<td>1350</td>
<td>CBM52/GH16</td>
<td>G</td>
<td>13.8</td>
</tr>
<tr>
<td>AmuVC2926</td>
<td>957</td>
<td>CBM18</td>
<td>G</td>
<td>0.3</td>
</tr>
<tr>
<td>AmuVC3592</td>
<td>1779</td>
<td>CBM10</td>
<td>G</td>
<td>5.5</td>
</tr>
<tr>
<td>AmuVC3612</td>
<td>2025</td>
<td>CBM18</td>
<td>G</td>
<td>0.6</td>
</tr>
<tr>
<td>AmuVC363</td>
<td>1833</td>
<td>CBM10/GH43</td>
<td>G</td>
<td>9.3</td>
</tr>
<tr>
<td>AmuVC3671</td>
<td>2172</td>
<td>CBM48/GH13</td>
<td>G</td>
<td>12.0</td>
</tr>
<tr>
<td>AmuVC375</td>
<td>5916</td>
<td>CBM18</td>
<td>G</td>
<td>8.5</td>
</tr>
<tr>
<td>AmuVC3846</td>
<td>2172</td>
<td>CBM18</td>
<td>G</td>
<td>29.7</td>
</tr>
<tr>
<td>AmuVC4088</td>
<td>1647</td>
<td>GH5</td>
<td>G</td>
<td>14.4</td>
</tr>
<tr>
<td>AmuVC4244</td>
<td>1227</td>
<td>CBM18</td>
<td>G</td>
<td>2.0</td>
</tr>
<tr>
<td>AmuVC4325</td>
<td>1641</td>
<td>CBM18</td>
<td>G</td>
<td>3.4</td>
</tr>
<tr>
<td>AmuVC4403</td>
<td>1389</td>
<td>CBM18</td>
<td>G</td>
<td>0.1</td>
</tr>
<tr>
<td>AmuVC4589</td>
<td>1296</td>
<td>CBM10</td>
<td>G</td>
<td>3.2</td>
</tr>
<tr>
<td>AmuVC508</td>
<td>1683</td>
<td>CBM10/CBM6/GH26/CBM35</td>
<td>G</td>
<td>3.9</td>
</tr>
<tr>
<td>AmuVC5227</td>
<td>2697</td>
<td>GH32C/GH32N</td>
<td>G</td>
<td>4.2</td>
</tr>
<tr>
<td>AmuVC537</td>
<td>1617</td>
<td>CBM10</td>
<td>G</td>
<td>18.2</td>
</tr>
<tr>
<td>AmuVC568</td>
<td>1038</td>
<td>CBM10</td>
<td>G</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuVC5814</td>
<td>1011</td>
<td>GH114</td>
<td>G</td>
<td>23.0</td>
</tr>
<tr>
<td>AmuVC6325</td>
<td>981</td>
<td>CBM1</td>
<td>G</td>
<td>2.7</td>
</tr>
<tr>
<td>AmuVC728</td>
<td>1953</td>
<td>CBM18</td>
<td>G</td>
<td>11.2</td>
</tr>
<tr>
<td>AmuVC893</td>
<td>1350</td>
<td>PL3/CBM1</td>
<td>G</td>
<td>9.2</td>
</tr>
<tr>
<td>AmuVC9126</td>
<td>1119</td>
<td>GH114</td>
<td>G</td>
<td>0.2</td>
</tr>
<tr>
<td>AmuVC9538</td>
<td>3726</td>
<td>CBM18/CBM13</td>
<td>G</td>
<td>0.5</td>
</tr>
<tr>
<td>AmuVC9660</td>
<td>5346</td>
<td>CBM18</td>
<td>G</td>
<td>15.5</td>
</tr>
<tr>
<td>AmuVC9821</td>
<td>1539</td>
<td>CE15/CBM10</td>
<td>G</td>
<td>2.2</td>
</tr>
<tr>
<td>AmuVN11467</td>
<td>1647</td>
<td>CBM18</td>
<td>G</td>
<td>8.8</td>
</tr>
<tr>
<td>AmuVN13467</td>
<td>2139</td>
<td>CBM10</td>
<td>G</td>
<td>4.5</td>
</tr>
<tr>
<td>AmuTC2525</td>
<td>2178</td>
<td>GH13/CBM26</td>
<td>NA</td>
<td>0.0</td>
</tr>
</tbody>
</table>