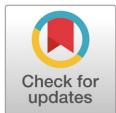


Nucleotide and protein researches on anaerobic fungi during four decades

Jongsoo Chang* and Hyunjin Park

Department of Agricultural Sciences, Korea National Open University, Seoul 03087, Korea



Received: Feb 20, 2020
 Revised: Mar 7, 2020
 Accepted: Mar 7, 2020

*Corresponding author

Jongsoo Chang
 Department of Agricultural Sciences,
 Korea National Open University,
 Seoul 03087, Korea.
 Tel: +82-2-3668-4636
 E-mail: jschang@kno.u.ac.kr

Copyright © 2020 Korean Society of Animal Sciences and Technology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID

Jongsoo Chang
<https://orcid.org/0000-0003-4118-120X>
 Hyunjin Park
<https://orcid.org/0000-0002-8098-2619>

Competing interests

No potential conflict of interest relevant to this article was reported.

Funding sources

This research was supported by Korea National Open University Research Fund.

Acknowledgements

The authors appreciate KNOU's financial support.

Availability of data and material

Upon reasonable request, the datasets of this study can be available from the corresponding author.

Authors' contributions

Conceptualization: Chang JS.
 Data curation: Park H.

Abstract

Anaerobic fungi habitat in the gastrointestinal tract of foregut fermenters or hindgut fermenters and degrade fibrous plant biomass through the hydrolysis reactions with a wide variety of cellulolytic enzymes and physical penetration through fiber matrix with their rhizoids. To date, seventeen genera have been described in family *Neocallimasticeae*, class *Neocallimastigomycetes*, phylum *Neocallimastigomycota* and one genus has been described in phylum *Neocallimastigomycota*. In National Center for Biotechnology Information (NCBI) database (DB), 23,830 nucleotide sequences and 59,512 protein sequences have been deposited and most of them were originated from *Piromyces*, *Neocallimastix* and *Anaeromyces*. Most of protein sequences (44,025) were acquired with PacBio next generation sequencing system. The whole genome sequences of *Anaeromyces robustus*, *Neocallimastix californiae*, *Pecoramyces ruminantium*, *Piromyces finnis* and *Piromyces* sp. E2 are available in Joint Genome Institute (JGI) database. According to the results of protein prediction, average Isoelectric points (pIs) were ranged from 5.88 (*Anaeromyces*) to 6.57 (*Piromyces*) and average molecular weights were ranged from 38.7 kDa (*Orpinomyces*) to 56.6 kDa (*Piromyces*). In Carbohydrate-Active enZymes (CAZY) database, glycoside hydrolases (36), carbohydrate binding module (11), carbohydrate esterases (8), glycosyltransferase (5) and polysaccharide lyases (3) from anaerobic fungi were registered. During four decades, 1,031 research articles about anaerobic fungi were published and 444 and 719 articles were available in PubMed (PM) and PubMed Central (PMC) DB.

Keywords: Anaerobic fungi, *Neocallimastigomycota*, *Neocallimasticeae*, Nucleotide, Protein

INTRODUCTION

Herbivorous animals can utilize fibrous carbohydrates as feed resources through the actions of various carbohydrate hydrolyzing enzymes secreted by microorganisms in their gastrointestinal tracts (GIT) including rumen for foregut fermenters and cecum for hindgut fermenters. Produced volatile fatty acids and microbial cells are absorbed and used as energy and nitrogen sources by host animals. The performance and production of herbivorous animals are closely related to the microbial ecosystem and individual microorganisms in GIT and this has made scientists isolate, identify and characterize microorganisms from GIT. In the GIT of herbivores, bacteria, protozoa and fungi are the major microorganisms which degrade fibrous feed and bacteria and protozoa degrade fiber only with the reaction

Formal analysis: Chang JS.
 Software: Park H.
 Validation: Chang JS.
 Writing - original draft: Chang JS.
 Writing - review & editing: Chang JS.

Ethics approval and consent to participate
 Not applicable.

of fiber degrading enzymes [1,2]. However, the anaerobic fungi (AF) degrade fiber through their physical and biochemical reactions. During fiber degradation, the anaerobic fungi penetrate their rhizoid filaments into the fiber matrix [3] and secrete a wide variety of both carbohydrate hydrolyzing enzymes and protein hydrolyzing enzymes [4]. The carbohydrate hydrolyzing enzymes secreted from the anaerobic fungi during fiber degradation are both glucanases [5] and xylanases [6] and the former hydrolyzes cellulose or starch to hexose glucose and the later hydrolyzes hemicellulose to pentose xylose.

More than four decades, the carbohydrate degrading enzymes from the anaerobic fungi were considered as promising resources both for animal production and other industries including food, paper, fabric and bioenergy. For this reason, many research teams have tried to isolate and identify the anaerobic fungi either from the foregut or hindgut of herbivores. However, only limited numbers of anaerobic fungi have been isolated and identified from the gastrointestinal tract of herbivores due to the difficulties in culture conditions. Besides, the sequence analysis of anaerobic fungi had been hindered with their high adenine (A) and thymine (T) contents and high A-T rich repeats in nucleotide sequences. Recently, the development of high throughput sequencing using the next generation sequencer (NGS) made it possible to analyze the fungal genome and transcriptome with or without culturing technique. In this paper, the researches on the anaerobic fungi during the last four decades were reviewed based on publicly available DB.

ABOUT ANAEROBIC FUNGI

Anaerobic fungi were classified as flagellated protozoa due to their flagellated zoospore [7,8] before Orpin's report (1975) and they were reclassified as family *Neocallimastigaceae* [9], order *Neocallimastigales* [10], and phylum *Neocallimastigomycota* [11]. Finally, the anaerobic fungi are classified in the phylum *Neocallimastigomycota* containing one class; *Neocallimastigomycetes*, one order; *Neocallimastigales*, one family; *Neocallimastigaceae* and 18 genera (Fig. 1). The genera in the family *Neocallimastigaceae* have been classified based on the morphological characteristics including zoospore flagellation (uniflagellate vs. polyflagellate), the sporangia development (monocentric vs. polycentric) and the thallus morphology (filamentous vs. bulbous) [12]. The number of flagella on zoospores is less than 4 in uniflagellate fungi and uniflagellate fungi are *Agriosomyces*, *Akiloshbomyces*, *Anaeromyces*, *Burwchfawromyces*, *Caecomyces*, *Capellomyces*, *Joblinomyces*, *Khoyollomyces*, *Liebetanzomyces*, *Oontomyces*, *Pecoromyces* and *Piromyces*. The polyflagellate fungi including *Cyllamyces*, *Feramyces*, *Ghzallomyces*, *Neocallimastix*, and *Orpinomyces* have more than 4 flagella on their zoospores. In monocentric fungi including *Agriosomyces*, *Akiloshbomyces*, *Burwchfawromyces*, *Caecomyces*, *Capellomyces*, *Cyllamyces*, *Feramyces*, *Ghzallomyces*, *Joblinomyces*, *Khoyollomyces*, *Liebetanzomyces*, *Neocallimastix*, *Oontomyces*, *Pecoromyces* and *Piromyces*, single sporangium containing a nucleus develops on its thallus and their rhizoidal system are anucleate. On the other hand, several sporangia develop on their thallus and the nucleus migrate through rhizoid with repeated division in polycentric fungi including *Anaeromyces* and *Orpinomyces*. The genera *Caecomyces* and *Cyllamyces* are bulbous type thallus with holdfast sporangia [13,14]. However, there have been many difficulties in morphology based identification due to variations in the number of flagellate and shape and size of sporangia under different nutritional environments [13]. The development of molecular technique led to the application of gene sequence for fungal identification and several different barcode markers were suggested. Dore and Stahl [15] compared the small subunit (SSU) rRNA genes of different anaerobic fungi for the first time, however, inter-relationship between genera was not clear due to highly conserved sequences in SSU. The internal transcribed spacer 1 region (ITS1) of rRNA was suggested by Li and Heath [16] and established [17]. Later several regions within the ITS1 were suggested for identification

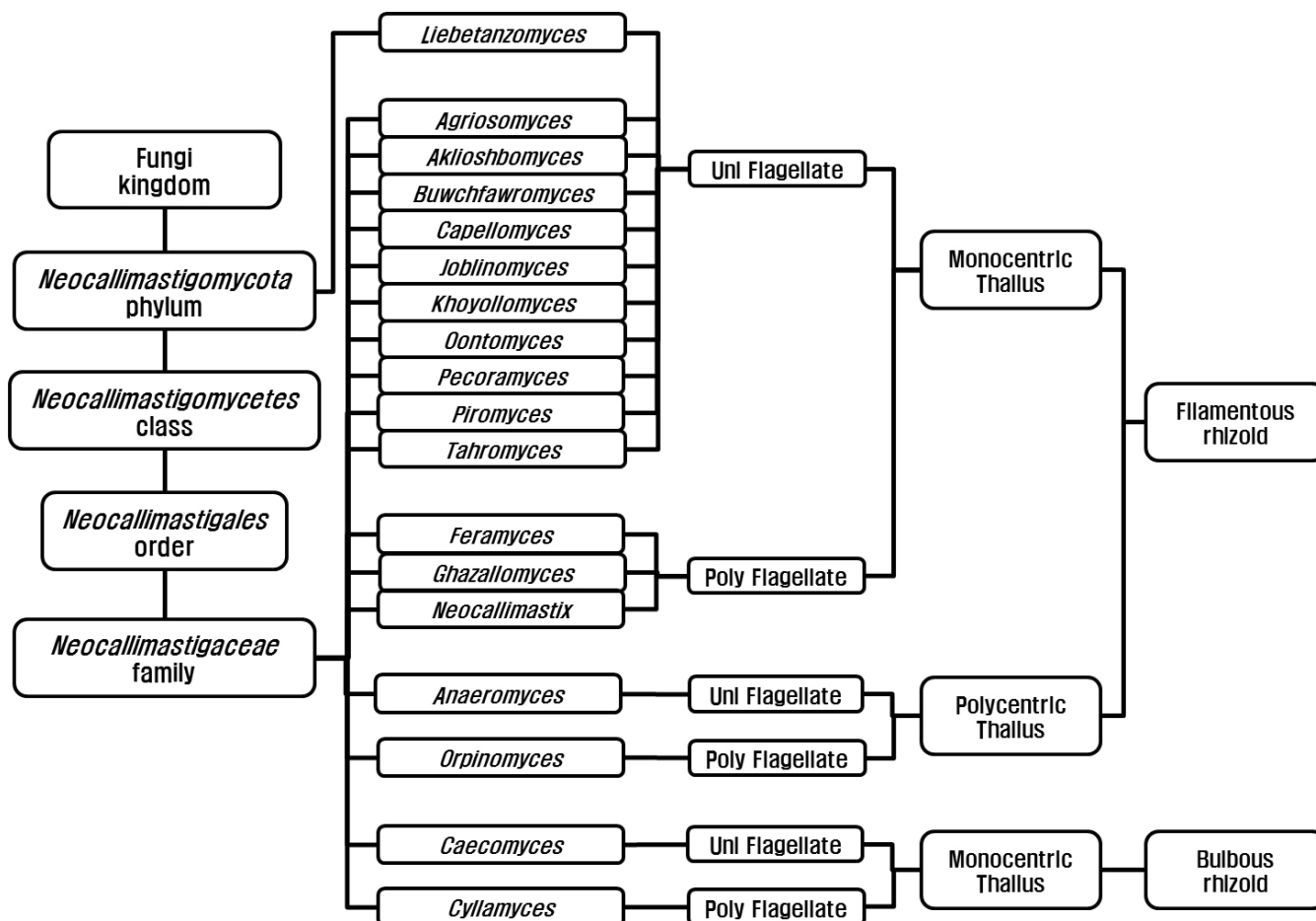


Fig. 1. Taxonomy of anaerobic fungi.

and quantification of anaerobic fungi [18,19]. The use of large subunit (LSU) rRNA gene was reported by Hausner and colleagues [20] for the first time and practical application methods using D1/D2 region within LSU for identification were reported [21,22]. Currently, both ITS1 and LSU sequences are analyzed for the identification of anaerobic fungi to compensate for the limitations of each marker [23].

The identified 35 species of anaerobic fungi were isolated from either ruminant (cattle, buffalo, sheep, wild goat, and camel) or non-ruminant herbivores (horse, donkey, and elephants) (Table 1). Among 35 identified species, 19 species were isolated from rumen contents and 16 species were isolated from fecal samples. During the 1980s, three filamentous type fungi including *Neocallimastix*, *Orpinomyces* and *Piromyces* and one bulbous type fungus, *Caecomyces*, with a total of 5 species were isolated and described from New Zealand (1), Canada (2) and United Kingdom (3). The new genus *Anaeromyces* and 3 genera with 11 species were isolated and described from Australia (1), Canada (1), United Kingdom (1), France (2), New Zealand (2) and Malaysia (4) during the 1990s and Ho and her colleagues [42] isolated 4 genera with 5 species during this period. The bulbous type fungi *Caecomyces* and *Cyllamyces* and filamentous type fungus *Piromyces* with a total of 3 species were isolated and described from United Kingdom (1) and Taiwan (2) during the 2000s. After the report on the use of barcode markers for identification of anaerobic fungi [15–17], *Caecomyces sympodialis* was the first described anaerobic fungi based on both barcode marker and morphological characteristics

Table 1. The isolation of anaerobic fungi from different host animals

Genus	Species	Host animal	Sample	Year	Ref.
<i>Agriosomyces</i>	<i>longus</i>	<i>Ovis orientalis</i>	Feces	2020	[24]
<i>Aklioshbomyces</i>	<i>papillarum</i>	<i>Odocoileus virginianus</i>	Feces	2020	[24]
<i>Anaeromyces</i>	<i>contortus</i>	<i>Bos taurus</i>	Feces	2018	[25]
	<i>elegans</i>	<i>Bos taurus</i>	Rumen	1993	[13]
	<i>mucronatus</i>	<i>Bos taurus</i>	Rumen	1990	[26]
	<i>polycephalus</i>	<i>Babalus befullus</i>	Rumen	2012	[27]
	<i>robustus</i>	<i>Ovis aries</i>	Feces	2016	[28]
<i>Buwchfawromyces</i>	<i>eastonii</i>	<i>Bulbalus bubalis</i>	Feces	2015	[29]
<i>Caecomycetes</i>	<i>churrovis</i>	<i>Ovis aries</i>	Feces	2017	[30]
	<i>communis</i>	<i>Ovis aries</i>	Feces	1988	[31]
	<i>equi</i>	<i>Equus ferus</i>	Feces	1988	[31]
	<i>hurleyensis</i>	<i>Ovis aries</i>	Rumen	2012	[27]
	<i>sympodialis</i>	<i>Bos indicus</i>	Rumen	2007	[32]
<i>Capellomyces</i>	<i>elongates</i>	<i>Capra aegagrus</i>	Feces	2020	[24]
	<i>foraminis</i>	<i>Capra aegagrus</i>	Feces	2020	[24]
<i>Cyllamyces</i>	<i>aberensis</i>	<i>Bos taurus</i>	Feces	2001	[33]
<i>Feramyces</i>	<i>austinii</i>	<i>Ammotragus lervia</i>	Rumen	2018	[34]
<i>Ghazallomyces</i>	<i>constrictus</i>	<i>Axis axis</i>	Feces	2020	[24]
<i>Joblinomyces</i>	<i>apicalis</i>	<i>Capra aegagrus</i>	Feces	2020	[24]
<i>Khoyollomyes</i>	<i>ramosus</i>	<i>Eauus grevyi</i>	Feces	2020	[24]
<i>Liebetanzomyces</i>	<i>polymorphus</i>	<i>Capra aegarrus</i>	Rumen	2018	[35]
<i>Neocallimastix</i>	<i>californiae</i>	<i>Capra aegarrus</i>	Feces	2016	[28]
	<i>cameroonii</i>	<i>Ovis aries</i>	Feces	2015	[36]
	<i>frontalis</i>	<i>Ovis aries</i>	Rumen	1983	[9]
	<i>hurleyensis</i>	<i>Ovis aries</i>	Rumen	1991	[37]
	<i>patriciarum</i>	<i>Ovis aries</i>	Rumen	1986	[38]
	<i>variabilis</i>	<i>Bos indicus</i>		1993	[39]
	<i>ankstri</i>	<i>Damelus dromedarius</i>	Rumen	2015	[40]
<i>Orpinomyces</i>	<i>bovis</i>	<i>Bos taurus</i>	Rumen	1989	[41]
	<i>intercalaris</i>	<i>Bos indicus</i>	Rumen	1994	[42]
	<i>joyonii</i>	<i>Bos taurus</i>	Rumen	1991	[43]
<i>Pecoromyces</i>	<i>ruminantium</i>	<i>Bos taurus</i>	Feces	2017	[44]
<i>Piromyces</i>	<i>communis</i>	<i>Ovis aries</i>	Feces	1988	[31]
	<i>cryptodigmaticus</i>	<i>Bos taurus</i>	Feces	2012	[27]
	<i>dumbonicus</i>	<i>Elephas maximus</i>	Feces	1990	[45]
	<i>finnis</i>	<i>Equus ferus</i>	Feces	2016	[28]
	<i>irregularis</i>	<i>Bos taurus</i>	Rumen	2015	[36]
	<i>mae</i>	<i>Equus ferus</i>	Feces	1990	[45]
	<i>minutus</i>	<i>Cervus nippon</i>	Rumen	1993	[46]
	<i>polycephalus</i>	<i>Babalus befullus</i>	Rumen	2002	[47]
	<i>rhizinflatus</i>	<i>Equus africanus</i>	Feces	1991	[48]
	<i>spiralis</i>	<i>Capra aegarrus</i>	Rumen	1993	[49]
<i>Tahromyces</i>	<i>munnaensis</i>	<i>Nilgiritragus hylocrius</i>	Feces	2020	[24]

Ref, reference.

[49]. During the last decade, 12 new genera with a total 23 species were identified due to advanced barcode marker techniques and increased needs for new enzyme resources required for bioenergy production.

NUCLEOTIDES OF ANAEROBIC FUNGI

After the first deposition of genomic DNA sequence from *Neocallimastix* spp., by Brownlee [50], 23,830 nucleotide sequences were available in NCBI database (Table 1) and the numbers were in the order of *Piromyces* (82.63%), *Neocallimastix* (9.25%), *Anaeromyces* (4.87%), *Orpinomyces* (1.65%), *Cyllamyces* (0.54%), *Caecomyces* (0.38%), *Pecoromyces* (0.29%), and *Feramyces* (0.29%). In the early 90s, the applications of cDNA library construction method using λ -ZAP II vector system were reported [51,52] and the needs of potent carbohydrate degrading enzyme for bioenergy production triggered researches on the functional genes originated from the anaerobic fungi during. Also, genome projects of JGI supported by the U.S. Department of Energy (DOE) accelerated the researches on the anaerobic fungi and resulted in large amounts of nucleotide sequences information about the anaerobic fungi during the third decade. During this period, about 17,100 out of 18,771 nucleotide sequences were acquired with transcriptome analysis method using constructed cDNA library in efforts to get the sequence information about carbohydrate degrading enzymes. Currently, whole genome sequences of *Anaeromyces robustus*, *Neocallimastix californiae*, *Pecoromyces ruminantium*, *Piromyces finnis* and *Piromyces* sp. E2 are available in JGI (<https://mycocosm.jgi.doe.gov/index.html>). The development of NGS technologies also affected the nucleotide researches on AF during the fourth decade and total 3,865 nucleotides sequenced with NGS including Illumina system (2016), PacBio system (1833) and Roch 454 GS-FLX system (16) were deposited in NCBI DB. The number of deposited nucleotide sequences was smaller than expected because the developments of NGS technologies and bioinformatics tools made it possible to translate acquired nucleotide sequences to amino acid sequences for protein prediction.

After the first report of ITS1 and 18S rRNA sequences (AF170188.10 [17]), 1,160 nucleotide sequences including 210 partial sequences from *Anaeromyces* were deposited in NCBI DB and most of them (925) were acquired with whole genome sequence analyses. The sequencing methods were PacBio sequencing system (925), Sanger method (115) and Illumina sequencing system (19) and. The number of reported ribosomal RNA sequences from *Anaeromyces* was 194 including ITS1 (186), 5.8S rRNA (149) and 28S rRNA (68) sequences and those were used for identification of anaerobic fungi. Only 16 mRNA sequences including carbohydrate hydrolyzing enzymes, polysaccharide lyase (PL), esterases, phosphatase, and acyltransferase were reported. The carbohydrate hydrolyzing enzyme sequences including cellulose (cel 1, GI 33327793), lichenase (licB, GI 33327791), glycosyl hydrolase family (GH) 1 (MH043785.1), GH10 (MH043796.1), GH32 (MH043815.1), and GH 67 (MH043829.1) were reported. In addition, three carboxyl esterases (CE) (CE7; MH043852.1, CE 12; MH043857.1, MH043856.1) and two polysaccharide lyases (PL) (PL9; MH043867.1, PL11; MH043868.1) sequences were also reported.

After the first report about 18S ribosomal RNA (M62707.1) originated from *Caecomyces communis* in 1993, 91 nucleotide sequences including 77 partial sequences were available in NCBI DB and the sequences were acquired with Sanger method (45) and Illumina system (23). Most of the nucleotides (68) from *Caecomyces* were related to fungal identification including ITS1 (61), 5.8S rRNA (61), 18S rRNA (22) and 28S rRNA (9) and only 20 nucleotide sequences were related to the information about the fungal protein. The sequences related to glycosyl hydrolase families (GH3; MH043789.1, GH8; MH043793.1, GH11; MH043803.1, GH18; MH043803.1, GH43; MH043822.1, and GH115; MH043838.1), carboxyl esterases (CE2; MH043845.1, CE3;

MH043846.1, and CE15; MH043860.1) and polysaccharide lyase (PL9; MH043865.1) from *Caecomyces* were reported. One interesting sequence was the immunity protein 51 which was not reported from other anaerobic fungi but bacteria as a hypothetical protein.

The number of partial nucleotide sequences originated from genus *Cyllamyces* in NCBI DB was 122 out of total 130 nucleotide sequences and 113 nucleotide sequences were from uncultured *Cyllamyces*. The sequences related to ribosomal RNA were 122 including ITS1 (64), 5.8S rRNA (64), 18S rRNA (118), and 28S rRNA (4) and 8 sequences were related to the patent of xylose isomerase production. However, no mRNA sequences related to protein were reported yet.

The number of partial nucleotide sequences originated from genus *Feramyces* in NCBI DB was 65 out of total 69 nucleotide sequences and the sequences were acquired with Sanger method (46) and Illumina system (23). The number of ribosomal RNA sequences were 44 including ITS1 (28), 5.8S rRNA (25) and 28S rRNA (16) sequences. The nucleotide sequences related to carbohydrate degrading proteins including glycosyl hydrolase (GH2; MH043788.1, GH5; MH043790.1, GH10; MH043795.1, GH30; MH043812.1, GH43; MH043823.1, GH47; MH043824.1, GH48; MH043825.1, GH76; MH043830.1, MH043831.1 GH97; MH043835.1, and GH130; MH043840.1), carboxyl esterase (CE1; MH043842.1, CE2; MH043843.1, CE4; MH043847.1, and CE12; MH043854.1), α -amylase (MH044722.1) and galactoside-O-acetyltransferase (MH043885.1) were deposited by Murphy and coworkers (<http://ncbi.nlm.nih.gov/nuccore>).

After the first report on AT-rich region of *Neocallimastix* LM-2 DNA (X14665.1) [50], 2,111 nucleotide sequences including 1,152 mRNAs and 163 rRNAs including ITS1 (140), 5.8S (110), 18S (98) and 28S (35) have been deposited in NCBI DB. In addition, 45 nucleotides were patent sequences related with AXEs, xylanases or enzyme production. The genes related with 11 glycosyl hydrolases (GH1, GH16, GH17, GH19, GH28, GH32, GH35, GH36, GH43, Gh64, and GH108), 4 carboxyl esterases (CE2, CE4, CE13, and CE16), glucanases (CelA, CelB, CelD, and Cel48) and cellobiohydrolases (CBH6, CBH20). Most nucleotide sequences were acquired with either ZAP II cDNA libraries (997) during transcriptome analyses or shotgun assembly (713) during whole genome analyses and used sequencing technologies were PacBio sequencing system (705), Sanger dideoxy sequencing method (78) or Illumina system (26).

The evolutionary closeness of cyclophilin originated from *Orpinomyces* sp. PC-2 and human were the first report of the nucleotide information originated from the genus *Orpinomyces* [53] and total 394 nucleotide sequences including 195 rRNAs and 55 mRNAs in NCBI DB. In addition, 12 nucleotide sequences were related to information about patent and 47 ITS1, 115 5.8S, 141 18S, and 77 28S rRNA sequences were related to fungal classification. The nucleotide sequences of 4 cellulases, 9 cellobiohydrolases, 8 glycosyl hydrolases, 2 glucanohydrolases, beta-glucosidase and endo-glucanase were deposited. Besides, the nucleotide sequences of other 3 hexose-hydrolases including α -amylase, lichenase, mannase and 3 pentose-hydrolases including xylanase and xylose isomerase were deposited in NCBI DB.

The whole genome sequences of *Orpinomyces* sp. strain C1A was reported by Yousseff and her colleagues in 2013 [54] and *Orpinomyces* sp. strain C1A was reclassified as *Pecoramyces ruminatum* strain C1A in 2017 [44]. Among available 70 nucleotide sequences, 32 mRNAs sequences and 21 rRNA sequences including ITS1 (15), 5.8S (7), 18S (3), and 28S (6) rRNA sequences originated from the genus *Pecoramyces* were deposited in NCBI DB. In addition, 16 nucleotide sequences related with carbohydrate degradation including α -amylase, β -glucosidase (BGL1 and BGL3), cellulase (Cel6A, Cel6B, and Cel48), glycosyl hydrolase (GH18, GH31, GH39, GH53, GH78, and GH88), polysaccharide lyase (PL1), and xylanase (XYL11 and xylan 1,4- β -xylosidase) originated from *Orpinomyces* were available.

Most of *Piromyces* nucleotide sequences were acquired from expressed sequence tag (EST) se-

quences and 9 nucleotide sequences were from genome sequences including shotgun sequences, and chromosome sequences. Interestingly, 258 nucleotide sequences were from patents and 94 nucleotide sequences were ribosomal RNA sequences. In *Piromyces*, 10 glycosyl hydrolases (GH1, GH5, GH6, GH8, GH25, GH26, GH31, GH43, GH57, and GH 127), 3 carboxyl esterases (CE4, CE12 and CE15), 12 glucanases (Cel1B, Cel1C, Cel1D, Cel3A, Cel5, Cel6A, Cel6B, Cel6C, Cel6E, Cel6G, Cel9, and Cel48), 7 cellobiohydrolases (CBHB, CBH6, CBH29, CBH120, CBHYW23-1, CBHYW23-2, and CBHYW23-4), 2 polysaccharide lyase (PL4 and PL9), xy-lanases and xylose isomerases were reported.

Only a limited number (12) of nucleotide sequences from genus *Burwchfawromyces* were available on NCBI DB and no information about functional genes from *Burwchfawromyces* has been reported yet. In addition, the numbers of nucleotide sequence from genus *Liebetanzomyces* and genus *Oontomyces* were 4 and 3, respectively, and all of them were ribosomal RNAs acquired with Sanger dideoxy sequencing method.

PROTEINS OF ANAEROBIC FUNGI

Until the end of the third decade, the total number of deposited protein amino acid sequences in NCBI DB was 189 even with the application of EST methods (Table 2). The length of sequence

Table 2. The nuclei and proteins deposited in NCBI database during 4 decades

Fungi	DB	Year				Total
		1980s	1990s	2000s	2010s	
<i>Anaeromyces</i>	Nucleus	0	2	43	1,115	1,160
	Protein	0	0	2	12,851	12,853
<i>Burwchfawromyces</i>	Nucleus	0	0	5	7	12
	Protein	0	0	0	0	0
<i>Caecomyces</i>	Nucleus	0	1	13	77	91
	Protein	0	0	0	20	20
<i>Cyllamyces</i>	Nucleus	0	0	68	62	130
	Protein	0	0	0	0	0
<i>Feramyces</i>	Nucleus	0	0	0	69	69
	Protein	0	0	0	21	21
<i>Liebetanzomyces</i>	Nucleus	0	0	0	4	4
	Protein	0	0	0	0	0
<i>Neocallimastix</i>	Nucleus	1	39	1,220	946	2,206
	Protein	0	2	59	20,404	20,465
<i>Oontomyces</i>	Nucleus	0	0	0	3	3
	Protein	0	0	0	0	0
<i>Orpinomyces</i>	Nucleus	0	21	164	209	394
	Protein	0	1	72	69	142
<i>Pecoromyces</i>	Nucleus	0	0	0	70	70
	Protein	0	0	0	41	41
<i>Piromyces</i>	Nucleus	0	52	17,258	2,381	19,691
	Protein	0	7	46	25,917	25,970
Total	Nucleus	1	115	18,771	4,943	23,830
	Protein	0	10	179	59,323	59,512

NCBI, National Center for Biotechnology Information; DB, database.

acquired from EST was ranged approximately from 500 to 800 nucleotides and this was not long enough to predict protein under the limited number of protein information. In addition, other working horse steps such as rapid amplification of cDNA ends (RACE) PCR method should be performed to get a complete amino acid sequence of the target protein [55]. It could be easy to get complete sequence information of expressed proteins with λ -ZAP II vector system, however either enough protein database or proper substrates for enzyme reactions to get the functional information about expressed proteins. The development of NGS technologies and their application to cDNA library of AF resulted in remarkable increase in sequence information of fungal proteome during the last two decades (Table 2). The number of sequence information acquired with PacBio system was 44,026 protein sequences from three genera including *Neocallimastix* (20,214 proteins), *Anaeromyces* (12,833 proteins) and *Piromyces* (10,979 proteins) and that with Illumina system was 14,618 proteins from three genera including *Piromyces* (14,613 proteins), *Neocallimastix* (4 proteins) and *Orpinomyces* (1 proteins). On the other hand, the classical Sanger dideoxy sequencing method was also used for sequence analyses of 14,618 proteins.

Protein separation using two dimensional gel electrophoresis (2D-GE), sequence analysis using mass spectrometry (MS) and protein identification are three key steps for classic proteome analysis technique and this could be a quite powerful technique to collect and compare each protein information (Fig. 2) [56]. However, it requires long training time to get high quality results of 2D-GE due to lots of tricky steps and the repeatability of 2D-GE is comparatively low. Technical advances in MS, protein separation and sequence analyses could be performed with a single run of LC-MS and this method was regarded as more effective one than 2D-GE.

In this paper, the protein virtual gels of individual fungus based on publically available amino acid sequences were constructed using JVirGel version 2.0 software (jvirgel.de). Among 12,853 protein sequences from *Anaeromyces*, 1,383 and 2,874 proteins were predicted as secretory proteins and membrane proteins and 5,452 proteins were predicted as remaining proteins that were not predicted either of them (Fig. 3, Supplementary Table S1). For the total proteome of *Anaeromyces*, the pI values were ranged from 3.00 (hypothetical protein, ORX83422.1) to 11.10 (hypothetical protein, ORX75296.1) and molecular weights were ranged from 200 kDa (Ketoacyl-synt-domain-containing protein, ORX78811.1) to 12.7 kDa (hypothetical protein, ORX75296.1). Among known functional proteomes, the domain of unknown function (DUF) 6 containing protein (ORX84229.1) with pI 9.82 had the highest pI and cellulose-domain containing protein (ORX78609.1) with pI

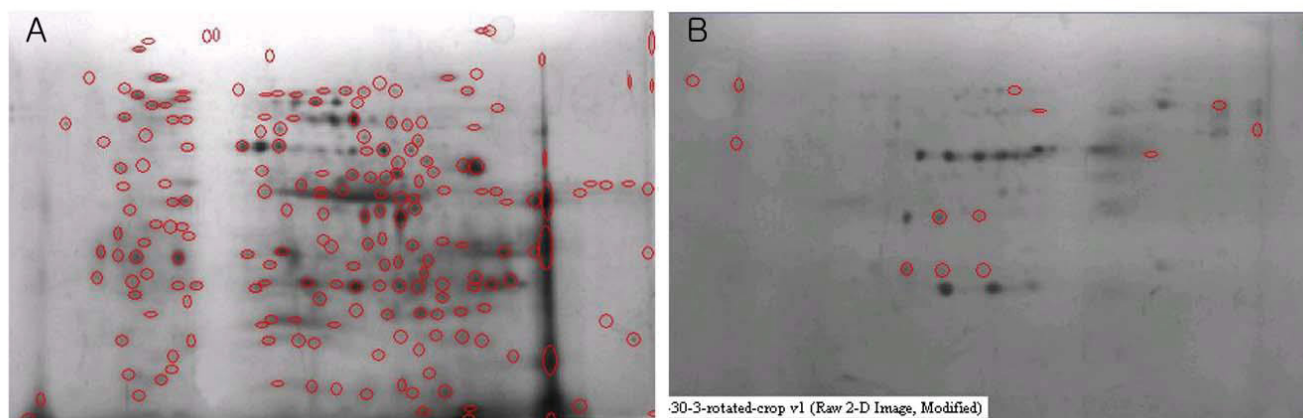


Fig. 2. Comparison of fungal mat proteome (A) and culture supernatant proteome (B) of anaerobic fungus *Neocallimastix frontalis* PMA02 on 2D-GE.

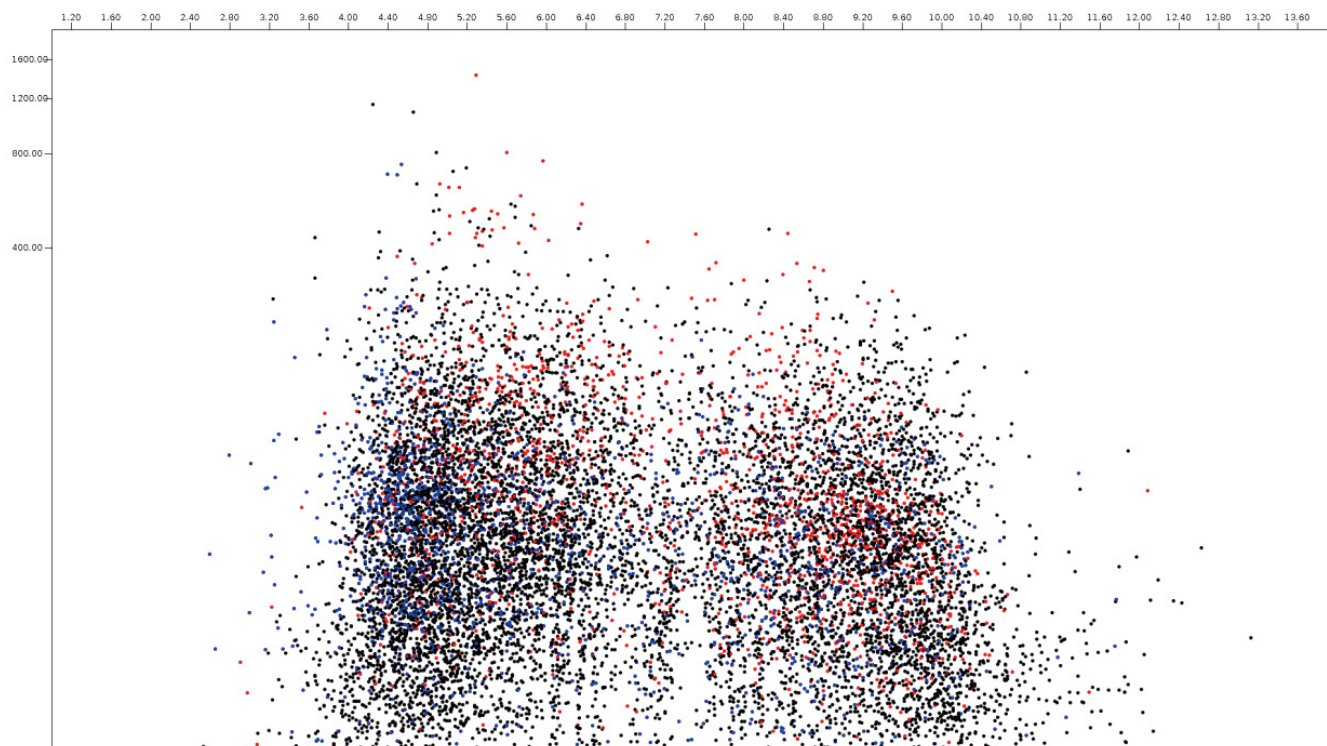


Fig. 3. Virtual 2D gel of *Anaeromyces* proteome. Blue, predicted secretome; red, predicted membrane proteome; black, remaining proteins.

3.80 had the lowest pI among functionally known secretomes. The abi-domain containing protein (ORX55221.1) with 15 kDa was the smallest and P-loop containing nucleoside triphosphate hydrolase protein (ORX86651.1) with 182.4 kDa was the largest among functionally known secretome. The average pI and molecular weight of *Anaeromyces* secretome were 5.88 and 55.4 kDa, respectively. The MFS (major facilitator superfamily) transporter (ORX87586.1) with pI 9.99 had the highest pI and WD40 (β -transducin) repeat-like protein (ORX61822.1) with pI 3.88 had the lowest pI among functionally known membrane proteome of *Anaeromyces*. The keotacyl-synthetase domain containing protein (ORX78811.1) with 199.8 kDa was the largest and RER1 protein (ORX55221.1) which involved in the retrieval of endoplasmic reticulum membrane protein with 16 kDa was the smallest among functionally known predicted membrane proteome. The average pI and molecular weight of membrane proteome were 6.88 and 69.4 kDa, respectively.

Among 15,745 protein sequences from *Neocallimastix*, 2,095 and 4,335 protein sequences were predicted as secretome and membrane proteome, respectively, and 9,315 protein sequences were predicted as remaining proteome (Fig. 4, Supplementary Table S2). The molecular weight of *Neocallimastix* proteome was a range from 200 kDa (hypothetical protein LY90DRAFT_666922, ORY72815.1) to 12.1 kDa (partial sequence of hypothetical protein LY90DRAFT_664663 from *Neocallimastix californiae* ORY79598.1) with an average of 58.8 kDa and hypothetical protein LY90DRAFT_499033 (ORY85919.1) with 13.8 kDa was the smallest among complete proteomes. One of hypothetical protein (ORY54617.1) with pI 10.0 had the highest pI, and another hypothetical protein (ORY24039.1) with pI 2.97 had the lowest pI among predicted secretomes of *Neocallimastix*. However, RNI-like protein (ORY54484.1) which involve in protein binding with pI 9.95 had the highest pI and invertase (ORY74006.1) with pI 3.84 had the lowest pI among functionally known secretomes of *Neocallimastix*. Scaffolding (ORY55229.1) which involved signaling

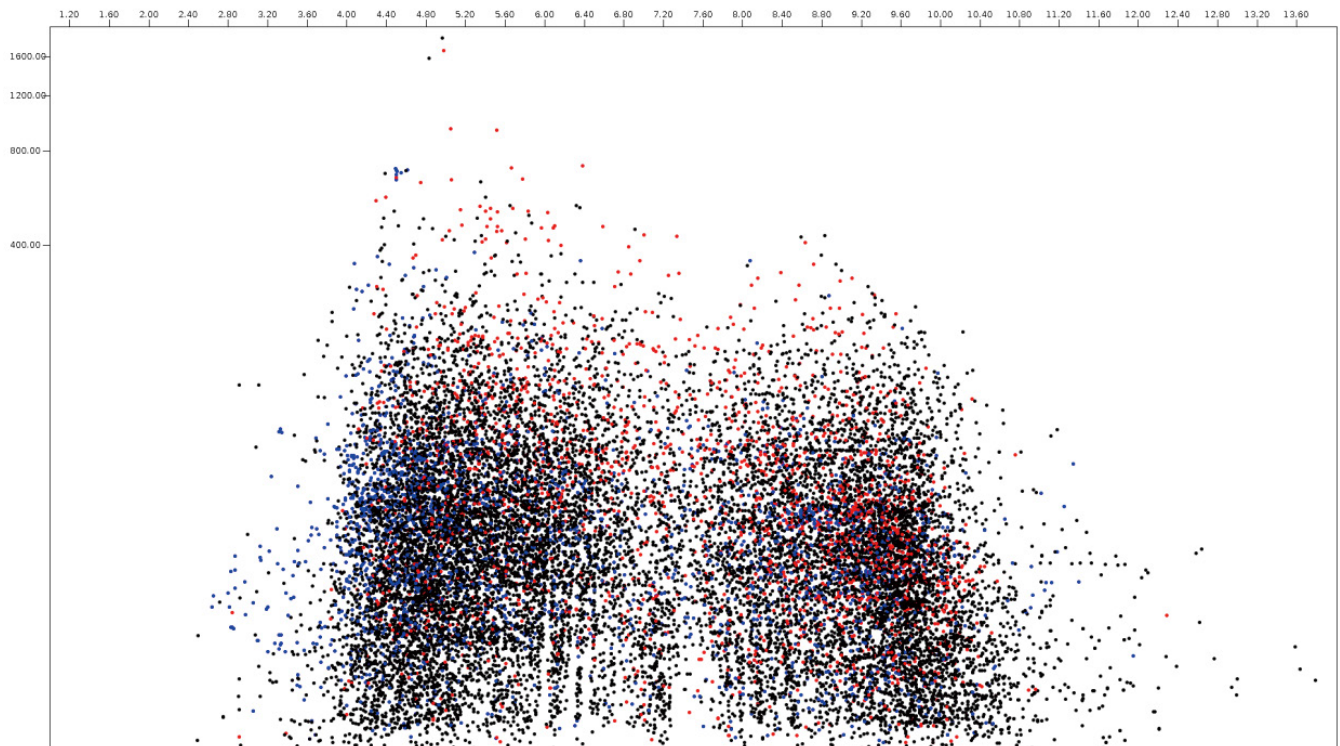


Fig. 4. Virtual 2D gel of *Neocallimastix* proteome. Blue, predicted secretome; red, predicted membrane proteome; black, remaining proteins.

pathway with 196 kDa was the largest and MFS general substrate transporter (ORY80567.1) with 14.3 kDa the smallest functionally known predicted secretome of *Neocallimastix*. The average pI and molecular weight of *Neocallimastix* secretome were 6.06 and 53.6 kDa, respectively. The pI 9.99 of essential protein for acyl-CoA-dependent ceramide synthesis, LAG1-domain containing protein (ORY33646.1) was the highest and the pI 3.88 of glycoside hydrolase/deacetylase (ORY79242.1) was the lowest among functionally known predicted membrane proteome of *Neocallimastix*. The protein related to the retrieval of early ER protein, Rer1 (ORT22199.1) with 15.7 kDa was the smallest and P-loop containing nucleoside triphosphate hydrolase protein (ORY43884.1) was the largest among functionally known predicted membrane proteomes of *Neocallimastix*. The average pI and molecular weight of *Neocallimastix* membrane proteome were 6.96 and 66.3 kDa, respectively.

Among 126 *Orpinomyces* proteomes, 29 protein sequences were predicted as membrane proteome and 97 protein sequences were predicted as remaining proteins (Fig. 5, Supplementary Table S3). The molecular weight of reported *Orpinomyces* proteome was ranged from 133.1 kDa (partial sequence of hypothetical protein from *Orpinomyces* sp. OUS1, CAI29548.1) to 10.1 kDa (partial sequence of the hypothetical protein from *Orpinomyces* sp. OUS1, CAI29549.1) with average 38.7 kDa. In addition, the pI of reported *Orpinomyces* proteome was ranged from 12.35 (partial sequence of the hypothetical protein from *Orpinomyces* sp OUS1, CAI29565.1) to 3.23 (partial sequence of the hypothetical protein from *Orpinomyces* sp OUS1, CAI29558.1) with an average of 6.47. The GH 95 from *Orpinomyces jayonii* (AWI66988.1) with 117.2 kDa was the largest and Cyclophilin B precursor from *Orpinomyces jayonii* (ABC47329.1) with 22.0 kDa was the smallest *Orpinomyces* membrane protein. In addition, pI value of reported *Orpinomyces* membrane proteins was ranged from 9.29 (GH114 from *Orpinomyces jayonii*, AWI66991.1) to 4.08 (partial sequence of putative 5-azacytidine resistance protein from *Orpinomyces* sp. OUS1, CAI11365.1).

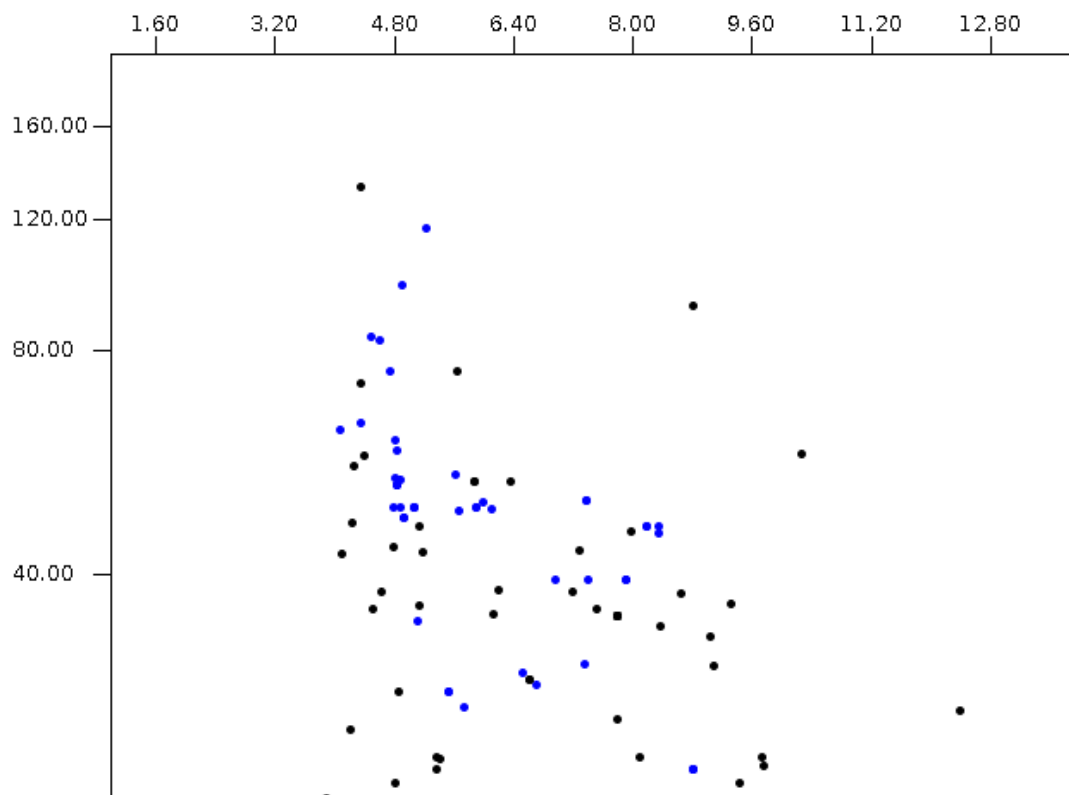


Fig. 5. Virtual 2D gel of *Orpinomyces* proteome. Blue, predicted secretome; black, remaining proteins.

Among 18,822 *Piromyces* proteomes, 2,270 and 4,863 protein sequences were predicted as secretome and membrane proteome, respectively, and 11,689 protein sequences were predicted as remaining proteins (Fig. 6, Supplementary Table S4). The molecular weight of *Piromyces* proteome was ranged from 200.0 kDa (Hypothetical protein from *Piromyces* sp. E2, OUM64425.1) to 12.9 kDa (hypothetical protein from *Piromyces* sp. E2, OUM60062.1) with an average of 56.9 kDa and the pIs were ranged from 11.00 (hypothetical protein from *Piromyces finnis*, ORX60159.1) to 2.92 (hypothetical protein from *Piromyces* sp. E2, OUM67877.1) with average of 6.57. The pI 10.06 of DUF6-domain-containing protein (ORX43281.1) was the highest and the pI 2.95 of non-catalytic module family DOC2 (OUM64499.1) was the lowest among functionally known predicted secretomes of *Piromyces*. The enzyme endoglucanase 5A with 192.9 kDa was the largest and putative terbinafine resistance locus protein (ORX49356.1) with 13.1 kDa was the smallest among functionally known predicted secretomes of *Piromyces*. The average pI and molecular weight of *Piromyces* secretome were 6.03 and 49.9 kDa, respectively. The glycosyltransferase family 2 protein (OUM61397.1) with 193.3 kDa was the largest and putative terbinafine resistance locus protein (ORX49356.1) with 13.1 kDa was the smallest among functionally known predicted membrane proteomes of *Piromyces*.

CARBOHYDRATE DEGRADING PROTEINS

The information about the proteins related to carbohydrate degradation including glycoside hydrolase (GH), glycosyl transferase (GT), polysaccharide lyase (PL), carbohydrate esterase (CE) and carbohydrate binding module (CBM) as an associated module is publically available on the

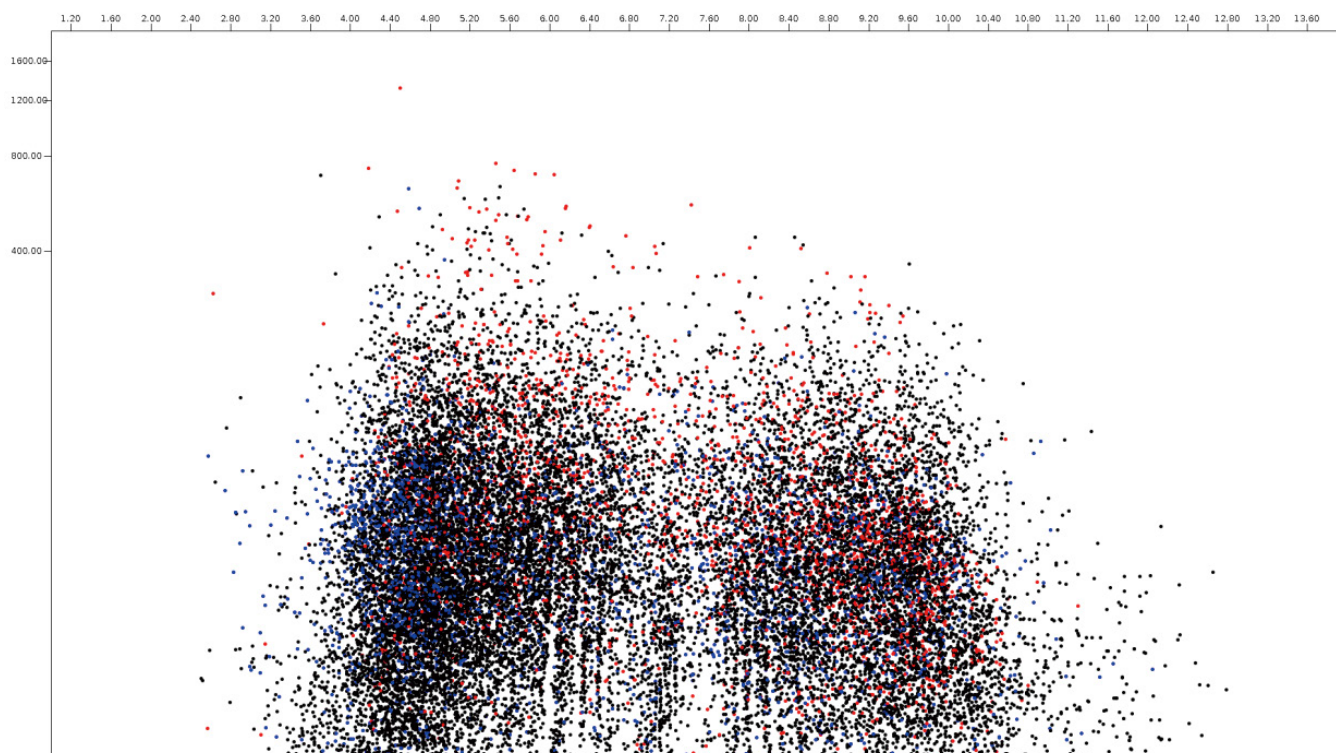


Fig. 6. Virtual 2D gel of *Piromyces* proteome. Blue, predicted secretome; red, predicted membrane proteome; black, remaining proteins.

CAZY DB (www.cazy.org). The GH family proteins hydrolyze the glycosidic bonds between sugar or bond between sugar and non-sugar moiety and 165 GH families have been registered to CAZY database. From anaerobic fungi, 36 GH family were registered and 5 different GH clans including (α/α)6, (β/α)8, 5-fold β -propeller and β -jelly roll were reported. The hexose hydrolyzing enzymes including α -glucosidase (EC 3.2.1.20; GH31), β -glucosidase (EC 3.2.1.21, GH1 and GH3) which hydrolyze non-reducing glucosyl residue terminal were detected from *Caecomyces*, *Neocallimastix*, *Orpinomyces*, *Pecoromyces* and *Piromyces*. The enzyme endoglucanases (EC 3.2.1.4; GH5, GH6, GH9, GH45, and GH 48) which hydrolyze glucosidic linkage of cellulose, lichenin and glucans were detected from *Anaeromyces*, *Neocallimastix*, *Orpinomyces* and *Piromyces*. The enzyme α -amylases (EC 3.2.1.1; GH13 and GH57) which hydrolyze α -D-glycosidic linkages of polysaccharide were detected from *Orpinomyces*, *Pecoromyces* and *Piromyces*. The enzyme β -mannase (EC 3.2.1.78; GH26) is a representative enzyme of GH 26 family proteins and the glucanase (EC 3.2.1.39) is the representative enzyme of GH 17 and GH 64 family proteins. The enzyme α -galactosidase (EC 3.2.1.22) and β -galactosidase (EC 3.2.1.23) were the representative enzymes of GH 36 and GH 35 family proteins, respectively. The enzymes involved in the hydrolyses of pentose sugar including endo-1,4- β -xylanase (EC 3.2.1.8, GH10 and GH11), β -xylosidase (EC 3.2.1.37, GH43 and GH 120), α -L-fucosidase (EC 3.2.1.51, GH 95 and GH 141), α -L-rhamnosidase (EC 3.2.1.40, GH 28 and GH 78), β -L-arabinofuranosidase (EC 3.2.1.185, GH 127) were detected from anaerobic fungi. In addition, chitinase (EC 3.2.1.14; GH18 and GH19) and chitosanase (EC 3.2.1.132; GH8) were also detected from anaerobic fungi.

The glycosyl transferases transfer sugar during glycosidic bond synthesis and 5 glycosyl transferase were registered and 3 polysaccharide lyase (PL), pectate lyase (EC 4.2.2.2, PL1), rhamnogalacturona endolyase (EC4.2.2.23, PL4) and rhamnogalacturona exolyase (EC 4.2.2.24, PL11) were

detected from anaerobic fungi. The carbohydrate esterases (CE) hydrolyze esters into an acid and an alcohol and 8 CE families were detected from anaerobic fungi. The enzyme acetyl xylan esterase (EC 3.1.1.72), feruloyl esterase (EC3.1.1.73), cinnamoyl esterase (EC 3.1.1.-) and carboxylesterase belong to CE 1 family, the existence of only acetyl xylan esterase and feruloyl esterase were reported. The acetyl xylan esterase from *Neocallimastix frontalis* PMA02 cleaves ester bond between acetyl side group and xylan or xylo-oligosaccharides [57] and feruloyl esterase from *Anaeromyces mucronatus* cleaves ester bond between ferulate and polysaccharide [58]. The representative enzyme of CE 8 is pectin methylesterase (EC 3.1.1.11) and the partial sequence of CE 8 from *Orpinomyces joyonii* D3B (AWI67007.1) could be methylesterase. Likewise, partial sequence of CE 13 from *Neocallimastix cameroonii* G3 (AWI67012.1) and CE 15 from *Piromyces* sp could be pectin acylesterase (EC 3.1.1.-) and 4-o-methyl-glucuronoyl methylesterase (EC3.1.1.-), respectively. The protein 3D structure of CE 1, CE3, CE6, and CE13 family were ($\alpha / \beta / \alpha$)-sandwich type and those of CE 4, CE 8 family were (β / α)-barrel and (β)-helix type, respectively. The carbohydrate esterases do not directly hydrolyze glucosidic linkage of polysaccharides, however, glucose hydrolysis is promoted with the removal of non-sugar side groups by CE [57].

RESEARCHES ON ANAEROBIC FUNGI

The research articles about anaerobic fungi were searched on PubMed (PM) and PubMed Central (PMC) DB using individual genus name as keywords and the numbers of hits were 1,138 in PM and 1,344 in PMC during the period from 1980 to 2019 (Table 3). Before 1980, 5 journal articles were detected in PM and one was detected in PMC. After filtration by deleting duplicated one, the numbers were reduced to 444 in PM and 719 in PMC and the articles overlapped both PM and PMC were 132 (Supplementary Table S5). Most of the results acquired from PM were directly related to anaerobic fungi, however, some results from PMC were not directly related to keyword or irrelevant to it. According to PM results, *Neocallimastix* had been the most popular topic in the anaerobic fungal research until 1990 because it was the earliest genus of anaerobic fungi. In addition, the numbers of research articles about *Neocallimastix* were the highest both in PM and PMC database and those about *Orpinomyces* and *Piromyces* were the next. The researches on *Piromyces* were continuously reported last four decades. Even with their early discovery, researches on *Anaeromyces*, and *Caecomyces* were actively reported during the 2000s. The research on newly discovered anaerobic fungi including *Burwchfarwomyces*, *Feromyces*, *Liebetanzomyces*, *Oontomyces*, and *Pecoromyces* were reported during 2010s.

The AF does not possess mitochondria but hydrogenosome for ATP production [59] and the hydrogenosome is surrounded by the double membrane [60]. The hydrogen can be produced through enzyme cascade reactions in hydrogenosome [61] and the biochemical characteristics of hydrogenosomal enzymes including pyruvate format-lyase [62], hydrogenase [63] and malic enzyme [64] were reported.

The first report on the physicochemical mechanism of AF in fiber degradation [3] suggested the importance of AF during fiber digestion in the rumen and it triggered the researches on the possible use of AF themselves or their enzymes as feed additives to increase the fiber digestibility in the rumen. The glycosidase activities in the culture supernatant of *Neocallimastix frontalis* (*N. frontalis*) using filter paper or avicel as substrate were compared [65] and the substrate conditions were expanded to natural fiber including wheat straw [6], Italian ryegrass [66] and maize stem [67]. The β -glucosidase [68] and xylanases [69] from *N. frontalis* were purified and characterized through the biochemical procedure using chromatograms and cellulases (celA, celB, and celC) from *Neocallimastix patriciarum* (*N. patriciarum*) were produced through molecular procedure using cDNA clon-

Table 3. The published research papers about anaerobic fungi during 4 decades in PubMed and PMC DB

Fungi	DB	Year				Total
		1980s	1990s	2000s	2010s	
<i>Anaeromyces</i>	PubMed	1	18	95	63	177
	PMC	0	2	27	69	98
<i>Buwchfawromyces</i>	PubMed	0	0	0	2	2
	PMC	0	0	0	11	11
<i>Caecomyces</i>	PubMed	1	25	96	63	185
	PMC	5	6	26	65	102
<i>Cyllamyces</i>	PubMed		16	96	58	170
	PMC	0	2	25	54	81
<i>Feromyces</i>	PubMed				1	1
	PMC	0	0	0	4	4
<i>Liebetanzomyces</i>	PubMed				1	1
	PMC	0	0	0	2	2
<i>Neocallimastix</i>	PubMed	40	90	52	50	232
	PMC	36	68	84	263	451
<i>Oontomyces</i>	PubMed				3	3
	PMC	0	0	0	9	9
<i>Orpinomyces</i>	PubMed		35	102	81	218
	PMC	0	14	48	182	244
<i>Pecoromyces</i>	PubMed				4	4
	PMC	0	0	0	12	12
<i>Piromyces</i>	PubMed	3	39	51	52	145
	PMC	5	22	80	223	330
Fungi	PubMed	3		1		4
	PMC	10	1	2	5	18
Order	PubMed		65	79	186	330
	PMC	0	2	24	37	63
Phylum	PubMed		16	92	94	202
	PMC	0	2	21	180	203
Total	PubMed	48	304	664	658	1,674
	PMC	56	119	337	1,116	1,628
	Total	104	423	1,001	1,774	3,302

PMC, PubMed Central DB; DB, database.

ing and consecutive heterologous expression in *Escherichia coli* [70]. Later, xylanase and mannanase from *Piromyces* [51], cellulase and xylanases from *Orpinomyces* [71], xylanases from *Anaeromyces* [72] and cellulase from *Pecoromyces* [73] were reported.

Unlike other eukaryotic genes, no introns were detected from most of endoglucanases [74,75] originated from AF and the horizontal transfer of bacterial genes to AF could be the reason for the intronless endoglucanases in AF [75,76]. The horizontal gene transfer in AF is crucial for the survival of AF in the rumen [77] and it could be beneficial to researchers for the production of AF protein. The use of cDNA library constructed with vector system was considered a promising tool for the production of enzymes from AF and has been used for three decades. The development NGS system with bioinformatics enormously increased the amount of protein sequence information with a single run and assembled 27,560 transcripts were acquired with 70.2 Gb reading with

Illumina HiSeq 2,000 platform [78]. The Illumina systems cover their short read length (up to 300 bp) with a huge amount of read (up to 6,000 Gb for NovaSeq), however, there might be some possibilities for misleading due to the characteristics of anaerobic fungal proteins. Recently, Pacific Biosciences developed PacBio system which armed with long read length (15–20 kb), however solid reports have not been published yet.

Natural celluloses were classified as cellulose I, II, III, and IV based upon its crystalline allomorphs [79] and the possible effects of cellulose crystallinity index on cellulose degradability was proposed [80]. The adsorption and activity of cellulolytic enzymes from AF were affected by the cellulose microcrystallinity [81] and the protein, cellulose binding module (CBM) in cellulases could be responsible for ligand binding action [82]. After molecular and biochemical characterization of CBM29 from *Piromyces equi* [83], 11 CBM including CBM1, 6, 10, 13, 18, 22, 26, 29, 35, 52, and 66 from *Anaeromyces*, *Caecomyces*, *Neocallimastix*, *Orpinomyces*, *Pecoromyces* and *Piromyces* are available in CAZY database (www.cazy.org). Anaerobic fungi improved their survivability under various substrate conditions with a wide variety of carbohydrate degrading enzymes which consisted of combinations of GH and CBM domains and this could be the reason why AFs maintain a wide range of enzyme systems for fiber digestion. The use of enzyme cocktail has been used in the glycosylation process of fibrous biomass for biofuel production and the similar concept was proposed using enzyme cocktail form AF [71]. The research on designing chimeric enzymes using AF genes in efforts to make multipurpose with thermostability was reported [84].

SUMMARY

Anaerobic fungi produce a wide variety of powerful carbohydrate hydrolyzing enzymes that can be used for animal production, biofuel production, food production and other purposes. The difficulties in the cultivation of AF under strictly anaerobic conditions caused hesitation in AF researches, however, the social requirement of substitutional energy source to petroleum and the development of molecular techniques including NGS stimulated researches on AF. During the last decades, development in culture medium led the discovery of new 12 genera of AF, however, there might be possibilities for more hidden AF in untouched area. In addition, large amounts of protein sequences have been produced during the last two decades, however, the information was skewed to three genera including *Piromyces*, *Neocallimastix* and *Anaeromyces*. The researches on AF should be expanded not only to newly described genera but the target itself. With a massive amount of sequence data, it might be difficult to scrutinize each protein, and researchers normally screen target sequences mechanically. However, sometimes AFs bear unexpected gift which researchers can easily miss.

SUPPLEMENTARY MATERIALS

Supplementary materials are only available online from: <https://doi.org/10.5187/jast.2020.62.2.121>.

REFERENCES

1. Hungate RE. The rumen and its microbes. New York, NY: Academic Press; 1966.
2. Orpin CG. Studies on the rumen flagellate *Neocallimastix frontalis*. J Gen Microbiol. 1975; 91:249-62.
3. Orpin CG. Invasion of plant tissue in the rumen by the flagellate *Neocallimastix frontalis*. J Gen Microbiol. 1977;98:423-30.
4. Ho YW, Bauchop T. Morphology of three polycentric rumen fungi and description of a proce-

- ture for the induction of zoosporogenesis and release of zoospores in cultures. *J Gen Microbiol.* 1991;137:213-7.
5. Mountfort DO, Asher RA. Production and regulation of cellulase by two strains of the rumen anaerobic fungus *Neocallimastix frontalis*. *Appl Environ Microbiol.* 1985;49:1314-22.
 6. Lowe SE, Theodorou MK, Trinci AP. Cellulases and xylanase of an anaerobic rumen fungus grown on wheat straw, wheat straw holocellulose, cellulose, and xylan. *Appl Environ Microbiol.* 1987;53:1216-23.
 7. Braune R. Untersuchungen über die in Wiederkauermagen vorkommenden Protozoen. *Arch Prot.* 1913;32:111-70.
 8. Liebetanz E. Die parasitischen Protozoen der Wiederkauermagens. *Arch Prot.* 1910;19:19-90.
 9. Heath IB, Bauchop T, Skipp RA. Assignment of the rumen anaerobe *Neocallimastix frontalis* to the Spizellomycetales (Chytridiomycetes) on the basis of its polyflagellate zoospore ultrastructure. *Can J Bot.* 1983;61:295-307.
 10. Li JL, Heath IB, Packer LN. The phylogenetic relationships of the anaerobic chytridiomycetous gut fungi (Neocallimasticaceae) and the Chytridiomycota. II. cladistic analysis of structural data and description of Neocallimasticales ord. nov. *Can J Bot.* 1993;17:393-407.
 11. Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, et al. A higher-level phylogenetic classification of the fungi. *Mycol Res.* 2007;111:509-47.
 12. Ho YW, Abdullah N, Jalaludin S. The diversity and taxonomy of anaerobic gut fungi. *Fungal Divers.* 2000;4:37-51.
 13. Ho YW, Barr DJS, Abdullah N, Jalaludin S. *Anaeromyces*, an earlier name for *Ruminomyces*. *Mycotaxon.* 1993;47:283-4.
 14. Gruninger RJ, Puniya AK, Callaghan TM, Edwards JE, Youssef N, Dagar SS, et al. Anaerobic fungi (phylum Neocallimastigomycota): advances in understanding their taxonomy, life cycle, ecology, role and biotechnological potential. *FEMS Microbiol Ecol.* 2014;90:1-17.
 15. Dore J, Stahl DA. Phylogeny of anaerobic rumen Chytridiomycetes inferred from small subunit ribosomal RNA sequence comparisons. *Can J Bot.* 1991;69:1964-71.
 16. Li J, Heath IB. The phylogenetic relationships of the anaerobic Chytridiomycetous gut fungi (Neocallimasticaceae) and the Chytridiomycota. I. cladistic analysis of rRNA sequences. *Can J Bot.* 1992;70:1738-46.
 17. Brookman JL, Mennim G, Trinci APJ, Theodorou MK, Tuckwell D. Identification and characterization of anaerobic gut fungi using molecular methodologies based on ribosomal ITS1 and 18S rRNA. *Microbiology.* 2000;146:393-403.
 18. Tuckwell DS, Nicholson MJ, McSweeney CS, Theodorou MK, Brookman JL. The rapid assignment of ruminal fungi to presumptive genera using ITS1 and ITS2 RNA secondary structures to produce group-specific fingerprints. *Microbiology.* 2005;151:1557-67.
 19. Koetschan C, Kittelmann S, Lu J, Al-Halbouni D, Jarvis GN, Müller T, et al. Internal transcribed spacer 1 secondary structure analysis reveals a common core throughout the anaerobic fungi (Neocallimastigomycota). *PLoS ONE.* 2014;9:e91928.
 20. Hausner G, Inglis GD, Yanke LJ, Kawchuk LM, McAllister TA. Analysis of restriction fragment length polymorphisms in the ribosomal DNA of a selection of anaerobic chytrids. *Can J Bot.* 2000;78:917-27.
 21. Fliegerová, K, Mrázek J, Voigt K. Differentiation of anaerobic polycentric fungi by rDNA PCR-RFLP. *Folia Microbiol.* 2006;51:273-7.
 22. Dagar SS, Kumar S, Mudgil P, Singh R, Puniya AK. D1/D2 domain of large-subunit ribosomal DNA for differentiation of *Orpinomyces* spp. *Appl Environ Microbiol.* 2011;77:6722-5.
 23. Edwards JE, Forster RJ, Callaghan TM, Dollhofer V, Dagar SS, Cheng Y, et al. PCR and

- omics based techniques to study the diversity, ecology and biology of anaerobic fungi: insights, challenges and opportunities. *Front Microbiol.* 2017;8:1657.
24. Hanafy RA, Lanjekar VB, Dhakephalkar PK, Callaghan TM, Dagar SS, Griffith GW, et al. Seven new Neocallimastigomycota genera from wild, zoo-housed, and domesticated herbivores greatly expand the taxonomic diversity of the phylum. *Mycologia.* 2020:1-28.
 25. Hanafy RA, Johnson B, Elshahed MS, Youssef H. *Anaeromyces contortus*, sp. nov., a new anaerobic gut fungal species (Neocallimastigomycota) isolated from the feces of cow and goat. *Mycologia.* 2018;110:502-12.
 26. Breton A, Bernalier A, Dusser M, Fonty G, Gaillard-Martinie B, Guillot J. *Anaeromyces mucronatus* nov. gen., nov. sp. A new strictly anaerobic rumen fungus with polycentric thallus. *FEMS Microbiol Lett.* 1990;58:177-82.
 27. Kirk PM. Nomenclatural novelties. [Internet]. Index Fungorum. 2012. <http://www.indexfungorum.org/Publications/Index%20Fungorum%20no.1.pdf>. Accessed 25 February 2020.
 28. Li GJ, Hyde KD, Zhao RL, Hongsanan S, Abdel-Aziz FA, Abdel-Wahab MA, et al. Fungal Diversity notes 253-366: taxonomic and phylogenetic contributions to fungal taxa. *Fungal Divers.* 2016;78:1-237.
 29. Callaghan TM, Podmirseg SM, Hohlweck D, Edwards JE, Puniya AK, Dagar SS, et al. *Buwchfawromyces eastonii* gen. nov., sp. nov.: a new anaerobic fungus (Neocallimastigomycota) isolated from buffalo faeces. *MycoKeys.* 2015;9:11-28.
 30. Henske JK, Gilmore SP, Knop D, Cunningham FJ, Sexton JA, Smallwood CR, et al. Transcriptomic characterization of *Caecomycetes churrovis*: a novel, non-rhizoid-forming lignocellulolytic anaerobic fungus. *Biotechnol Biofuels.* 2017;10:305.
 31. Gold JJ, Heath IB, Bauchop T. Ultrastructural description of a new chytrid genus of caecum anaerobe, *Caecomycetes equi* gen. nov., sp. nov., assigned to the Neocallimasticeae. *BioSystems.* 1988;21:403-15.
 32. Chen YC, Tsai SD, Cheng HL, Chien CY, Hu CY, Cheng TY. *Caecomycetes sympodialis* sp. nov., a new rumen fungus isolated from *Bos indicus*. *Mycologia.* 2007;99:125-30.
 33. Ozkose E, Thomas BJ, Davies DR, Griffith GW, Theodorou MK. *Cyllamyces aberensis* gen. nov. sp. nov., a new anaerobic gut fungus with branched sporangiophores isolated from cattle. *Can J Bot.* 2001;79:666-73.
 34. Hanafy RA, Elshahed MS, Youssef NH. *Feramyces austinii*, gen. nov., sp. nov., an anaerobic gut fungus from rumen and fecal samples of wild Barbary sheep and fallow deer. *Mycologia.* 2018;110:513-25.
 35. Joshi A, Lanjekar VB, Dhakephalkar PK, Callaghan TM, Griffith GW, Dagar SS. *Liebetanzomyces polymorphus* gen. et sp. nov., a new anaerobic fungus (Neocallimastigomycota) isolated from the rumen of a goat. *MycoKeys.* 2018;40:89-110.
 36. Ariyawansa HA, Hyde KD, Jayasiri SC, Buyck B, Chethana KWT, Dai DQ, et al. Fungal diversity notes 111-252-taxonomic and phylogenetic contributions to fungal taxa. *Fungal Divers.* 2015;75:27-274.
 37. Webb J, Theodorou MK. *Neocallimastix hurleyensis* sp. nov., an anaerobic fungus from the ovine rumen. *Can J Bot.* 1991;69:1220-4.
 38. Orpin CG, Munn EA. *Neocallimastix patriciarum* sp. nov., a new member of the Neocallimasticeae inhabiting the rumen of sheep. *Trans Br Mycol Soc.* 1986;86:178-81.
 39. Ho YW, Barr DJS, Abdullah N, Jalaludin S, Kudo H. *Neocallimastix variabilis*, a new species of anaerobic fungus from the rumen of cattle. *Mycotaxon.* 1993;46:241-58.
 40. Dagar SS, Kumar S, Griffith GW, Edwards JE, Callaghan TM, Singh R, et al. A new anaerobic fungus (*Oontomyces anksri* gen. nov., sp. nov.) from the digestive tract of the Indian camel

- (*Camelus dromedarius*). *Fungal Biol.* 2015;119:731-7.
41. Barr DJS, Kudo H, Jakober KD, Cheng KJ. Morphology and development of rumen fungi: *Neocallimastix* sp., *Piromyces communis*, and *Orpinomyces bovis* gen.nov., sp.nov. *Can J Bot.* 1989;67:2815-24.
 42. Ho YW, Abdullah N, Jalaludin S. *Orpinomyces intercalaris*, a new species of polycentric anaerobic rumen fungus from cattle. *Mycotaxon.* 1994;50:139-50.
 43. Li J, Heath IB, Cheng KJ. The development and zoospore ultrastructure of a polycentric chytridiomycete gut fungus, *Orpinomyces joyonii* comb.nov. *Can J Bot.* 1991;69:580-9.
 44. Hanafy RA, Elshahed MS, Liggenstoffer AS, Griffith GW, Youssef NH. *Pecoramycetes ruminantium*, gen. nov., sp. nov., an anaerobic gut fungus from the feces of cattle and sheep. *Mycologia.* 2017;109:231-43.
 45. Li J, Heath IB, Bauchop T. *Piromyces mae* and *Piromyces dumbonica*, two new species of uniflagellate anaerobic chytridiomycete fungi from the hindgut of the horse and elephant. *Can J Bot.* 1990;68:1021-33.
 46. Ho YW, Barr DJS, Abdullah N, Jalaludin S, Kudo H. A new species of *Piromyces* from the rumen of deer in Malaysia. *Mycotaxon.* 1993;47:285-93.
 47. Chen YC, Hseu RS, Chien CY. *Piromyces polycephalus* (Neocallimastigaceae), a new rumen fungus. *Nova Hedwigia.* 2002;75:409-14.
 48. Breton A, Dusser M, Gaillard-Martine B, Guillot J, Millet L, Prensier G. *Piromyces rhizinflata* nov. sp., a strictly anaerobic fungus from faeces of the Saharian ass: a morphological, metabolic and ultrastructural study. *FEMS Microbiol Lett.* 1991;82:1-8.
 49. Ho YW, Barr DJS, Abdullah N, Jalaludin S, Kudo H. *Piromyces spiralis*, a new species of anaerobic fungus from the rumen of goat. *Mycotaxon.* 1993;48:59-68.
 50. Brownlee AG. Remarkably AT-rich genomic DNA from the anaerobic fungus *Neocallimastix*. *Nucleic Acids Res.* 1989;17:1327-35.
 51. Xue GP, Gobius KS, Orpin CG. A novel polysaccharide hydrolase cDNA (*celD*) from *Neocallimastix patriciarum* encoding three multi-functional catalytic domains with high endoglucanase, cellobiohydrolase and xylanase activities. *J Gen Microbiol.* 1992;138:2397-403.
 52. Fanutti C, Ponyi T, Black GW, Hazlewood GP, Gilbert HJ. The conserved noncatalytic 40-residue sequence in cellulases and hemicellulases from anaerobic fungi functions as a protein docking domain. *J Biol Chem.* 1995;270:29314-22.
 53. Chen H, Li XL, Ljungdahl LG. A cyclophilin from the polycentric anaerobic rumen fungus *Orpinomyces* sp. strain PC-2 is highly homologous to vertebrate cyclophilin B. *Proc Natl Acad Sci USA.* 1995;92:2587-91.
 54. Youssef NH, Couger MB, Struchtemeyer CG, Liggenstoffer AS, Prade RA, Najjar FZ, et al. The genome of the anaerobic fungus *Orpinomyces* sp. strain C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. *Appl Environ Microbiol.* 2013;79:4620-34.
 55. Kwon M, Song J, Ha JK, Park HS, Chang J. Analysis of functional genes in carbohydrate metabolic pathway of anaerobic rumen fungus *Neocallimastix frontalis* PMA02. *Asian-Australas J Anim Sci.* 2009; 22:1555-65.
 56. Park MA, Son HJ, Kim JN, Choi NJ, Jeong HY, Ha JK, et al. Analysis of protein expression from the cultures of anaerobic rumen fungus, *Neocallimastix* sp NLRI 003. In: Proceedings of the 11th Animal Science Congress; 2004. Serdang, Malaysia. p. 305-6.
 57. Kwon M, Song J, Park HS, Park H, Chang J. Characterization of heterologously expressed acetyl xylan esterase1 isolated from the anaerobic rumen fungus *Neocallimastix frontalis* PMA02. *Asian-Australas J Anim Sci.* 2016;29:1576-84.
 58. Qi M, Wang P, Selinger LB, Yanke LJ, Forster RJ, McAllister TA. Isolation and characteriza-

- tion of a ferulic acid esterase (Fae1A) from the rumen fungus *Anaeromyces mucronatus*. *J Appl Microbiol*. 2011;110:1341-50.
59. Yarlett N, Orpin CG, Munn EA, Yarlett NC, Greenwood CA. Hydrogenosomes in the rumen fungus *Neocallimastix patriciarum*. *Biochem J*. 1986;236:729-39.
 60. Benchimol M, Durand R, Almeida JC. A double membrane surrounds the hydrogenosomes of the anaerobic fungus *Neocallimastix frontalis*. *FEMS Microbiol Lett*. 1997;154:277-82.
 61. Marvin-Sikkema FD, Pedro Gomes TM, Grivet JP, Gottschal JC, Prins RA. Characterization of hydrogenosomes and their role in glucose metabolism of *Neocallimastix* sp. L2. *Arch Microbiol*. 1993;160:388-96.
 62. Akhmanova A, Voncken FGJ, Hosea KM, Harhangi H, Keltjens JT, Camp HJMOD, et al. A hydrogenosome with pyruvate formate-lyase: anaerobic chytrid fungi use an alternative route for pyruvate catabolism. *Mol Microbiol*. 1999;32:1103-14.
 63. Davidson EA, van der Giezen M, Horner DS, Embley TM, Howe CJ. An [Fe] hydrogenase from the anaerobic hydrogenosome-containing fungus *Neocallimastix frontalis* L2. *Gene*. 2002; 296:45-52.
 64. Park MA. Genomic and proteomic approaches for elucidation of carbohydrate metabolism in anaerobic rumen fungus *Neocallimastix frontalis* PMA02 [Ph.D. dissertation]. Seoul, Korea: Seoul National University; 2011.
 65. Pearce PD, Bauchop T. Glycosidases of the rumen anaerobic fungus *Neocallimastix frontalis* grown on cellulosic substrates. *Appl Environ Microbiol*. 1985;49:1265-9.
 66. Theodorou MK, Longland AC, Dhanoa MS, Lowe SE, Trinci APJ. Growth of *Neocallimastix* sp. strain R1 on Italian ryegrass hay: removal of neutral sugars from plant cell walls. *Appl Environ Microbiol*. 1989;55:1363-7.
 67. Roger V, Grenet E, Jamot J, Bernalier A, Fonty G, Gouet P. Degradation of maize stem by two rumen fungal species, *Piromyces communis* and *Caecomyces communis*, in pure cultures or in association with cellulolytic bacteria. *Reprod Nutr Dev*. 1992;32:321-9.
 68. Li X, Calza RE. Fractionation of cellulases from the ruminal fungus *Neocallimastix frontalis* EB188. *Appl Environ Microbiol*. 1991;57:3331-6.
 69. de Segura BG, Fevre M. Purification and characterization of two 1,4-beta-xylan endohydrolases from the rumen fungus *Neocallimastix frontalis*. *Appl Environ Microbiol*. 1993;59:3654-60.
 70. Xue GP, Orpin CG, Gobius KS, Aylward JH, Simpson GD. Cloning and expression of multiple cellulase cDNAs from the anaerobic rumen fungus *Neocallimastix patriciarum* in *Escherichia coli*. *J Gen Microbiol*. 1992;138:1413-20.
 71. Li XL, Chen H, Ljungdahl LG. Monocentric and polycentric anaerobic fungi produce structurally related cellulases and xylanases. *Appl Environ Microbiol*. 1997;63:628-35.
 72. Novotná Z, Procházka J, Šimůnek J, Fliegerová K. Xylanases of anaerobic fungus *Anaeromyces mucronatus*. *Folia Microbiol*. 2010;55:363-7.
 73. Morrison JM, Elshahed MS, Youssef NH. Defined enzyme cocktail from the anaerobic fungus *Orpinomyces* sp. strain C1A effectively releases sugars from pretreated corn stover and switchgrass. *Sci Rep*. 2016;6:29217.
 74. Zhou L, Xue GP, Orpin CG, Black GW, Gilbert HJ, Hazlewood GP. Intronless celB from the anaerobic fungus *Neocallimastix patriciarum* encodes a modular family A endoglucanase. *Biochem J*. 1994;297:359-64.
 75. Liu JH, Selinger LB, Hu YJ, Cheng KJ, Beauchemin KA, Moloney MM. An endoglucanase from the anaerobic fungus *Orpinomyces joyonii*: characterization of the gene and its product. *Can J Microbiol*. 1997;43:477-85.

76. Garcia-Vallvé S, Romeu A, Palau J. Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. *Mol Biol Evol.* 2000;17:352-61.
77. Murphy CL, Youssef NH, Hanafy RA, Couger MB, Stajich JE, Wang Y, et al. Horizontal gene transfer as an indispensable driver for evolution of Neocallimastigomycota into a distinct gut-dwelling fungal lineage. *Appl Environ Microbiol.* 2019;85:e00988-19.
78. Couger MB, Youssef NH, Struchtemeyer CG, Liggenstoffer AS, Elshahed MS. Transcriptomic analysis of lignocellulosic biomass degradation by the anaerobic fungal isolate *Orpinomyces* sp. strain C1A. *Biotechnol Biofuels.* 2015;8:208.
79. Hayashi J, Sufoka A, Ohkita J, Watanabe S. The confirmation of existences of cellulose III, IIII, IVI, and IVII by the X-ray method. *J Pol Sci Lett.* 1975;13:23-27.
80. Park S, Baker JO, Himmel ME, Parilla PA, Johnson DK. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol Biofuels.* 2010;3:10.
81. Dijkerman R, Vervuren MBF, Camp HJMOD, van der Drift C. Adsorption characteristics of cellulolytic enzymes from the anaerobic fungus *Piromyces* sp. strain E2 on microcrystalline cellulose. *Appl Environ Microbiol.* 1996;62:20-5.
82. Steenbakkens PJM, Li XL, Ximenes EA, Arts JG, Chen H, Ljungdahl LG, Camp HJMOD. Noncatalytic docking domains of cellulosomes of anaerobic fungi. *J Bacteriol.* 2001;183:5325-33.
83. Freelove ACJ, Bolam DN, White P, Hazlewood GP, Gilbert HJ. A novel carbohydrate-binding protein is a component of the plant cell wall-degrading complex of *Piromyces equi*. *J Biol Chem.* 2001;276:43010-7.
84. Gilmore SP, Lillington SP, Haitjema CH, de Groot R, O'Malley MA. Designing chimeric enzymes inspired by fungal cellulosomes. *Synth Syst Biotechnol.* 2020;5:23-32.