

## Supplementary information

Table S1. Bacterial species and strains containing GH18-genes.

Table S2. Summary of the number GH18s examined in the current work.

Table S3. Domain-structure of all GH18s examined in the current work.

## Supplementary tables

Table S1

Bacterial species and strains containing GH18-genes

Phylum	Class/Order	Species/Strain
Actinobacteria 12/41 (29%)		Acidothermus cellulolyticus 11B
		Arthrobacter sp. FB24
		Frankia sp. EAN1pec
		Mycobacterium flavescensPYR-GCK
		Mycobacterium vanbaalenii PYR-1
		Nocardioides sp. JS614
		Salinispora tropica CNB-440
		Streptomyces avermitilis MA-4680
		Streptomyces coelicolor A(3)2
		Symbiobacterium thermophilum IAM 14863
		Thermobifida fusca YX
		Marine actinobacterium PHSC20C1
Bacteroidetes/Chlorobi 4/27 (15%)		Bacteroides thetaiotaomicron VPI-5482
		Flavobacteriales bacterium HTCC2170
		Flavobacterium johnsoniae UW101
		Salinibacter ruber DSM 13855
Cyanobacteria 4/27 (15%)		Nostoc punctiforme PCC 73102
		Synechococcus sp. RS9916
		Synechococcus sp. RS9917
		Synechococcus sp. WH 7805
Firmicutes 40/126 (32%)	Bacillales 23/42 (55%)	Bacillus anthracis str. A2012

Bacillus anthracis str. Ames  
Bacillus cereus ATCC 10987  
Bacillus cereus ATCC 14579  
Bacillus cereus E33L  
Bacillus cereus G9241  
Bacillus cereus subsp. cytotoxis NVH 391-98  
Bacillus clausii KSM-K16  
Bacillus halodurans C-125  
Bacillus licheniformis ATCC 14580  
Bacillus sp. NRRL B-14911  
Bacillus subtilis subsp. subtilis str. 168  
Bacillus thuringiensis serovar israelensis ATCC 35646  
Bacillus thuringiensis serovar konkukian str. 97-27  
Bacillus weihenstephanensis KBAB4  
Geobacillus kaustophilus HTA426  
Listeria innocua Clip11262  
Listeria monocytogenes EGD-e  
Listeria monocytogenes str. 1/2a F6854  
Listeria monocytogenes str. 4b F2365  
Listeria monocytogenes str. 4b H7858  
Listeria welshimeri serovar 6b str. SLCC5334  
Oceanobacillus iheyensis HTE831

Clostridiales 13/22 (59%)

Clostridium acetobutylicum ATCC 824  
Clostridium beijerincki NCIMB 8052  
Clostridium difficile QCD-32g58  
Clostridium phytofermentans ISDg  
Clostridium sp. OhILAs  
Clostridium thermocellum ATCC 27405  
Desulfitobacterium hafniense DCB-2  
Desulfitobacterium hafniense Y51  
Desulfotomaculum reducens MI-1

	Halothermothrix orenii H 168
	Syntrophomonas wolfei subsp. wolfei str. Goettingen
	Thermoanaerobacter ethanolicus X514
	Thermoanaerobacter tengcongensis
Lactobacillales 3/46 (7%)	
	Enterococcus faecalis V583
	Lactococcus lactis subsp. cremoris SK11
	Lactococcus lactis subsp. lactis I11403
<u>Mollicutes</u> 1/16 (6%)	
	Mesoplasma florum L1
Proteobacteria 67/350 (19%)	
Alpha 4/89 (4%)	
Alpha/Rickettsiales 3/25 (12%)	
	Rickettsia bellii OSU 85-389
	Rickettsia bellii RML369-C
	Rickettsia felis URRWXC12
Alpha/Others 1/59 (2%)	
	Xanthobacter autotrophicus Py2
Beta 14/55 (25%)	
Beta/Burkholderiales 13/36 (36%)	
	Burkholderia cenocepacia AU 1054
	Burkholderia cepacia AMMD
	Burkholderia dolosa AUO158
	Burkholderia mallei ATCC 23344
	Burkholderia pseudomallei 1710b
	Burkholderia pseudomallei 668
	Burkholderia pseudomallei K96243
	Burkholderia sp. 383
	Burkholderia thailandensis E264
	Burkholderia xenovorans LB400
	Ralstonia solanacearum GMI1000

	<i>Ralstonia solanacearum</i> UW551
	<i>Acidovorax avenae</i> subsp. <i>Citrulli</i> AAC00-1
Beta/Neisseriales 1/4 (25%)	
	<i>Chromobacterium violaceum</i> ATCC 12472
Delta 2/19 (11%)	
	<i>Myxococcus xanthus</i> DK 1622
	<i>Stigmatella aurantiaca</i> DW4/3-1
Gamma 67/161 (42%)	
Gamma/Enterobacteriales	
17/52 (33%)	
	<i>Escherichia coli</i> 101-1
	<i>Escherichia coli</i> 536
	<i>Escherichia coli</i> 53638
	<i>Escherichia coli</i> CFT073
	<i>Escherichia coli</i> F11
	<i>Escherichia coli</i> HS
	<i>Escherichia coli</i> K12
	<i>Escherichia coli</i> UTI89
	<i>Escherichia coli</i> W3110
	<i>Photobacterium luminescens</i> subsp. <i>laumondii</i> TTO1
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi Ty2
	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi str. CT18
	<i>Salmonella typhimurium</i> LT2
	<i>Sodalis glossinidius</i> str. 'morsitans'
	<i>Yersinia bercovieri</i> ATCC 43970
	<i>Yersinia pseudotuberculosis</i> IP 31758
	<i>Yersinia pseudotuberculosis</i> IP 32953
Gamma/Pseudomonadales	
12/15 (80%)	
	<i>Pseudomonas aeruginosa</i> 2192
	<i>Pseudomonas aeruginosa</i> C3719
	<i>Pseudomonas aeruginosa</i> PA7

	Pseudomonas aeruginosa PACS2
	Pseudomonas aeruginosa PAO1
	Pseudomonas aeruginosa UCBPP-PA14
	Pseudomonas entomophila L48
	Pseudomonas fluorescens Pf-5
	Pseudomonas fluorescens PfO-1
	Pseudomonas syringae pv. phaseolicola 1448A
	Pseudomonas syringae pv. syringae B728a
	Pseudomonas syringae pv. tomato str. DC3000
Gamma/Vibrionales 18/18 (100%)	
	Photobacterium profundum 3TCK
	Photobacterium profundum SS9
	Photobacterium sp. SKA34
	Vibrio alginolyticus 12G01
	Vibrio angustum S14
	Vibrio cholerae MO10
	Vibrio cholerae O1 biovar eltor str. N16961
	Vibrio cholerae O395
	Vibrio cholerae RC385
	Vibrio cholerae V51
	Vibrio cholerae V52
	Vibrio fisheri ES114
	Vibrio parahaemolyticus RIMD 2210633
	Vibrio sp. Ex25
	Vibrio sp. MED222
	Vibrio splendidus 12B01
	Vibrio vulnificus CMCP6
	Vibrio vulnificus YJ016
Gamma/Others 20/55 (36%)	
	Coxiella burnetii Dugway 7E9-12
	Coxiella burnetii RSA 331

	Coxiella burnetii RSA 493
	Francisella tularensis subsp. holarctica
	Francisella tularensis subsp. holarctica OSU18
	Francisella tularensis subsp. tularensis SCHU S4
	Hahella chejuensis KCTC 2396
	Legionella pneumophila str. Lens
	Legionella pneumophila str. Paris
	Legionella pneumophila subsp. pneumophila str. Philadelphia 1
	Pseudoalteromonas tunicata D2
	Reinekea sp. MED297
	Saccharophagus degradans 2-40
	Shewanella baltica OS155
	Shewanella baltica OS195
	Shewanella denitrificans OS217
	Shewanella oneidensis MR-1
	Shewanella sp. ANA-3
	Shewanella sp. MR-4
	Shewanella sp. MR-7
<u>Spirochaetes</u> 2/9 (22%)	
	Leptospira interrogans serovar Copenhageni str. Fiocruz L1-130
	Leptospira interrogans serovar Lai str. 56601
Acidobacteria 2/2 (100%)	
	Acidobacteria bacterium Ellin345
	Solibacter usitatus Ellin6076
Chloroflexi 3/6 (50%)	
	Chloroflexus aurantiacus J-10-fl
	Herpetosiphon aurantiacus ATCC 23779
	Roseiflexus sp. RS-1

---

Table S2

Summary of the number of GH18s examined in the current work.

Organism	Group	Gene no.	GH18 domain no.
Virus			
<i>Orgyia pseudotsugata</i> MNPV <sup>a</sup>	Baculoviridae	1	1
<i>Agrotis segetum</i> GV <sup>a</sup>	Baculoviridae	1	1
<i>Agrotis segetum</i> GV <sup>a</sup>	Baculoviridae	1	1
<i>Bombyx mori</i> NPV <sup>a</sup>	Baculoviridae	1	1
<i>Choristoneura fumiferana</i> MNPV <sup>a</sup>	Baculoviridae	1	1
<i>Spodoptera exigua</i> NPV <sup>a</sup>	Baculoviridae	1	1
<i>Helicoverpa zea</i> SNPV <sup>a</sup>	Baculoviridae	1	1
Animal			
<i>Caenorhabditis elegans</i>	Rhabditida	33	35
<i>Drosophila melanogaster</i>	Diptera	18	22

Plant

<i>Arabidopsis thaliana</i>	Brassicales	10	10
<i>Malus domestica</i> <sup>a</sup>	<b>Rosales</b>	<b>1</b>	<b>1</b>
<i>Citrullus lanatus</i> <sup>a</sup>	<b>Cucurbitales</b>	<b>1</b>	<b>1</b>
<i>Coffea arabica</i> <sup>a</sup>	<b>Gentianales</b>	<b>1</b>	<b>1</b>
<i>Hevea brasiliensis</i> <sup>a</sup>	<b>Malpighiales</b>	<b>1</b>	<b>1</b>
<i>Triticum aestivum</i> <sup>a</sup>	<b>Poales</b>	<b>1</b>	<b>1</b>
<i>Oryza sativa</i> <sup>a</sup>	<b>Poales</b>	<b>1</b>	<b>1</b>
<i>Canavalia ensiformis</i> <sup>a</sup>	<b>Fabales</b>	<b>1</b>	<b>1</b>

Bacteria

<i>Rickettsia bellii</i> OSU 85-389	Alphaproteobacteria	1	1
<i>Rickettsia bellii</i> RML369-C	Alphaproteobacteria	1	1
<i>Xanthobacter autotrophicus</i> Py2	Alphaproteobacteria	2	2
<i>Burkholderia cenocepacia</i> AU1054	Betaproteobacteria	1	1
<i>Burkholderia pseudomallei</i> S13	Betaproteobacteria	1	1



<i>Chromobacterium violaceum</i> ATCC12472	Betaproteobacteria	5	5
<i>Escherichia coli</i> 101-1	Gammaproteobacteria	1	1
<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi Ty2	Gammaproteobacteria	1	1
<i>Yersinia pseudotuberculosis</i> IP32953	Gammaproteobacteria	1	1
<b><i>Serratia marcescens</i><sup>a</sup></b>	<b>Gammaproteobacteria</b>	<b>2</b>	<b>2</b>
<i>Hahella chejuensis</i> KCTC2396	Gammaproteobacteria	6	7
<i>Legionella pneumophila</i> Paris	Gammaproteobacteria	2	2
<i>Shewanella oneidensis</i> MR-1	Gammaproteobacteria	1	1
<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	Gammaproteobacteria	1	1
<i>Pseudomonas fluorescens</i> PfO-1	Gammaproteobacteria	1	1
<i>Pseudomonas entomophila</i> L48	Gammaproteobacteria	1	1
<i>Saccharophagus degradans</i> 2-40	Gammaproteobacteria	4	5
<i>Photobacterium</i> sp. SKA34	Gammaproteobacteria	8	8
<i>Vibrio angustum</i> S14	Gammaproteobacteria	8	8
<i>Vibrio vulnificus</i> CMCP6	Gammaproteobacteria	5	5
<i>Vibrio cholerae</i> N16961	Gammaproteobacteria	4	4

<i>Myxococcus xanthus</i> DK1622	Deltaproteobacteria	4	4
<i>Streptomyces coelicolor</i> A3(2)	Actinobacteria	10	10
<i>Symbiobacterium thermophilum</i> IAM14863	Actinobacteria	4	4
<i>Acidothermus cellulolyticus</i> 11B	Actinobacteria	5	5
<i>Salinispora tropica</i> CNB-440	Actinobacteria	3	3
<i>Frankia</i> sp. EAN1pec	Actinobacteria	1	1
<i>Mycobacterium flavescens</i> PYR-GCK	Actinobacteria	1	1
<i>Streptococcus pyogenes</i> MGAS9429	Bacilli	1	1
<i>Bacillus licheniformis</i> ATCC14580	Bacilli	5	5
<i>Bacillus cereus</i> ATCC10987	Bacilli	3	3
<i>Bacillus thuringiensis</i> serovar <i>israelensis</i> ATCC 35646	Bacilli	5	5
<i>Listeria monocytogenes</i> 4b F2365	Bacilli	2	2
<i>Listeria innocua</i> Clip11262	Bacilli	2	2
<i>Enterococcus faecalis</i> V583	Bacilli	3	3
<i>Clostridium phytofermentans</i> ISDg	Clostridia	5	5
<i>Mesoplasma florum</i> L1	Mollicutes	1	1

<i>Acidobacteria sp.</i> Ellin345	Acidobacteria	2	2
<i>Solibacter usitatus</i> Ellin6076	Solibacteres	2	2
<i>Flavobacterium johnsoniae</i> UW101	Flavobacteria	4	4
<i>Salinibacter ruber</i> DSM 13855	Sphingobacteria	1	1
<i>Nostoc punctiforme</i> PCC 73102	Cyanobacteria	1	1
<i>Synechococcus sp.</i> RS9916	Cyanobacteria	1	1
Fungi			
<i>Rhizopus oryzae</i>	Mucormycetes	15	15
<i>Saccharomyces cerevisiae</i>	Saccharomycetes	2	2
<i>Kluyveromyces lactis</i>	Saccharomycetes	2	2
<i>Candida albicans</i>	Saccharomycetes	5	5
<i>Lodderomyces elongisporus</i>	Saccharomycetes	4	4
<i>Schizosaccharomyces pombe</i>	Schizosaccharomycetes	1	1
<i>Gibberella zeae</i>	Sordariomycetes	19	19
<i>Hypocrea jecorina</i>	Sordariomycetes	20	20

<i>Neurospora crassa</i>	Sordariomycetes	12	12
<i>Magnaporthe grisea</i>	Sordariomycetes	15	15
<b><i>Metarhizium anisopliae</i><sup>a</sup></b>	<b>Sordariomycetes</b>	<b>2</b>	<b>2</b>
<i>Botryotinia fuckeliana</i>	Leotiomycetes	10	10
<i>Sclerotinia sclerotiorum</i>	Leotiomycetes	14	14
<i>Emericella nidulans</i>	Eurotiomycetes	20	20
<i>Coccidioides immitis</i>	Eurotiomycetes	9	9
<i>Ajellomyces capsulatus</i>	Eurotiomycetes	9	9
<i>Phaeosphaeria nodorum</i>	Dothideomycetes	18	18
<i>Phanerochaete chrysosporium</i>	Homobasidiomycetes	10	10
<i>Coprinopsis cinerea</i>	Homobasidiomycetes	8	8
<i>Laccaria bicolor</i>	Homobasidiomycetes	11	11
<i>Filobasidiella neoformans</i>	Heterobasidiomycetes	4	4
<i>Ustilago maydis</i>	Ustilaginomycetes	3	3

Archaea

<i>Pyrococcus furiosus</i>	Euryarchaeota	2	2
<i>Pyrococcus kodakaraensis</i>	Euryarchaeota	1	2
<i>Halobacterium sp.</i>	Euryarchaeota	2	2

---

<sup>a</sup> Bold text indicate single proteins retrieved from GenBank.

Table S3

Domain-structure of all GH18s examined in the current work, as predicted by InterProScan (Zdobnov and Apweiler, 2001), Conserved Domain Database (Marchler-Bauer et al., 2005) or Pfam (Finn et al., 2006). Numbers refers to the domain-order, from the amino-terminal start towards the carboxy-terminal end of the proteins.

## Group A-II

Protein and species	Taxonomic group	Glycoside hydrolase, family 18, catalytic domain	Endoplasmic reticulum targeting sequence	Predicted hydrolase domain	Signal peptide	Chitin-binding type 1 domain
02650 <i>A. capsulatum</i>	Eurotiomycetes	1 <sup>1</sup>				
AN0221 <i>E. nidulans</i>	Eurotiomycetes	1 <sup>1</sup>	2			
06020 <i>N. crassa</i>	Sordariomycete	1 <sup>1</sup>		2		
09277 <i>P. nodorum</i>	Dothideomycete	1 <sup>1</sup>	2			
EAA76014 <i>G. zeae</i>	Sordariomycete	1 <sup>1</sup>	2			
Chi18-4 <i>H. jecorina</i>	Sordariomycete	1 <sup>1</sup>	2			
05838 <i>A. capsulatum</i>	Eurotiomycetes	2			1	
02644 <i>P. nodorum</i>	Dothideomycete	1 <sup>1</sup>				2
AN11233 <i>E. nidulans</i>	Eurotiomycetes	2			1	3
EAA69039 <i>G. zeae</i>	Sordariomycete	1 <sup>1</sup>				2
MGG 06594 <i>M. grisea</i>	Sordariomycete	2 <sup>1</sup>			1	3
Chi18-11 <sup>2</sup> <i>H. jecorina</i>	Sordariomycete	1 <sup>1</sup>				

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

<sup>2</sup>Protein sequence is not full-length.

## Group A-III

Protein and species	Taxonomic group	Signal peptide	Glycoside hydrolase, family 18, catalytic	Phox-domain	Carbohydrate-binding family V/XII
---------------------	-----------------	----------------	---	-------------	-----------------------------------

			domain		
01315.1 <i>C. cinerea</i>	Homobasidiomycete	1	2		
03252.1 <i>C. cinerea</i>	Homobasidiomycete	1	2		
246840 <i>L. bicolor</i>	Homobasidiomycete	1	2		
39872 <i>P. chrysosporium</i>	Homobasidiomycete		1		
231399 <i>L. bicolor</i>	Homobasidiomycete		1		
05285.1 <i>C. cinerea</i>	Homobasidiomycete	1	2		
05291.1 <i>C. cinerea</i>	Homobasidiomycete	1	2		
180319 <i>L. bicolor</i>	Homobasidiomycete	1	2		
191848 <i>L. bicolor</i>	Homobasidiomycete	1	2		
310136 <i>L. bicolor</i>	Homobasidiomycete	1	2		
128098 <i>P. chrysosporium</i>	Homobasidiomycete		1		
138098 <i>P. chrysosporium</i>	Homobasidiomycete	1	2		
EAA71245 <i>G. zeae</i>	Sordariomycete		1		
MGG 04732 <i>M. grisea</i>	Sordariomycete	1	2		
01586.1 <i>C. cinerea</i>	Homobasidiomycete		2	1	
317003 <i>L. bicolor</i>	Homobasidiomycete		2	1	
UM06190 <i>U. maydis</i>	Ustilaginomycete		1		
129436 <i>P. chrysosporium</i>	Homobasidiomycete		1		
144643 <sup>2</sup> <i>L. bicolor</i>	Homobasidiomycete		1		
04245 <i>F. neoformans</i>	Heterobasidiomycete	1	2		
16170 <i>R. oryzae</i>	Mucormucete	1	2		
01334 <i>R. oryzae</i>	Mucormucete	1	2		
13934 <i>R. oryzae</i>	Mucormucete	1	2		
07611 <i>R. oryzae</i>	Mucormucete	1	2		3
10252 <i>R. oryzae</i>	Mucormucete		1		
14659 <i>R. oryzae</i>	Mucormucete	1	2		

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

<sup>2</sup> Protein sequence is not full-length.

#### Group A-IV

Protein and species	Taxonomic group	Glycoside	2-S globulin
---------------------	-----------------	-----------	--------------

		hydrolase, family 18, catalytic domain	
EAA69503 <i>G. zeae</i>	Sordariomycete	1 <sup>1</sup>	2, 3
Chi18-3 <i>H. jecorina</i>	Sordariomycete	1 <sup>1</sup>	2, 3
03026 <i>N. crassa</i>	Sordariomycete	1 <sup>1</sup>	
01594 <i>B. fuckeliana</i>	Leotiomycetes	1	2, 3, 4
05897 <i>S. sclerotiorum</i>	Leotiomycetes	1	
04742 <i>A. capsulatum</i>	Eurotiomycetes	1	
04750 <i>C. immitis</i>	Eurotiomycetes	1	
04719 <i>P. nodorum</i>	Dothideomycetes	1	
EAA70860 <i>G. zeae</i>	Sordariomycete	1	2, 3
Chi18-2 <i>H. jecorina</i>	Sordariomycete	1	2, 3
MGG 08458 <i>M. grisea</i>	Sordariomycete	1	

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

#### Group A-V

Protein and species	Taxonomic group	Signal peptide	Glycoside hydrolase, family 18, catalytic domain	Calcium-binding EF-hand	Peptidoglycan-binding LysM	Carbohydrate-binding family V/XII
EAA67655 <i>G. zeae</i>	Sordariomycete		1			
MGG 05533 <i>M. grisea</i>	Sordariomycete	1	2	3		
Chi18-6 <i>H. jecorina</i>	Sordariomycete	1	2			
MGG 01247 <i>M. grisea</i>	Sordariomycete		1			



03209 N. crassa	Sordariomycete	1	2			
EAA73155 G. zeae	Sordariomycete		1			
Chi18-5 H. jecorina	Sordariomycete	1	2			
EAA74986 G. zeae	Sordariomycete		1			
MGG 00086 M. grisea	Sordariomycete	1	2			
04883 N. crassa	Sordariomycete		1			
EAA74223 G. zeae	Sordariomycete		1			
Chi18-7 H. jecorina	Sordariomycete		1			
MGG 07927 M. grisea	Sordariomycete		1			
04554 N. crassa	Sordariomycete		1			
05350 B. fuckeliana	Leotiomycetes		1			
11304 S. sclerotiorum	Leotiomycetes		1			
11407 B. fuckeliana	Leotiomycetes	1	2			
11700 S. sclerotiorum	Leotiomycetes	1	2			
08020 S. sclerotiorum	Leotiomycetes		1			
01250 A. capsulatum	Eurotiomycetes		2		1	

05250 C. immitis	Eurotiomycetes		2		1	
05886 A. capsulatum	Eurotiomycetes	1	2			
06565 A. capsulatum	Eurotiomycetes	1	2			
03822 C. immitis	Eurotiomycetes		1			
02795 C. immitis	Eurotiomycetes		1			
AN5454 E. nidulans	Eurotiomycetes		1			
AN4871 E. nidulans	Eurotiomycetes	1	2			
00269 P. nodorum	Dothideomycetes		1			
12247 P. nodorum	Dothideomycetes		1			
14474 P. nodorum	Dothideomycetes	1	2			
04870.1 C. cinereus	Homobasidiomycetes	1	2			
124149 P. chrysosporium	Homobasidiomycetes		1			
03412 F. neoformans	Heterobasidiomycetes	1	2			
09921.1 C. cinereus	Homobasidiomycetes	1	2			4
134311 P. chrysosporium	Homobasidiomycetes		1			2
13635 R. oryzae	Mucormycetes		1			

UM04261 U. maydis	Ustilaginomycetes		1			
CHT4 C. albicans	Saccharomycetes		1			
01119 L. elongisporum	Saccharomycetes		1			
XP 454053 K. lactis	Saccharomycetes		1			
CTS2 S. cerevisiae	Saccharomycetes		1			

#### Group C-I

Protein and species	Taxonomic group	Chitin-binding type 1 domain	Glycoside hydrolase, family 18, catalytic domain	Signal peptide	Ankyrin repeats	Endoplasmic reticulum targeting	Glu/Leu/Phe/Val dehydrogenase	Inorganic pyrophosphatase	Cysteine peptidase active site	Homeo domain-like, DNA-binding	Calcium-binding EF-hand	DNA-binding spore protein	Sas10/Utp3
AN0549 E. nidulans	Eurotiomycetes	1, 2, 3	4										
AN8481 E. nidulans	Eurotiomycetes	1	2										
12510 S. sclerotiorum	Leotiomyces	1	2										
EAA74768 G. zeae	Sordariomycete	1, 2	3										
07035 N. crassa	Sordariomycete		2	1									
EAA78168 G. zeae	Sordariomycete	1	2 <sup>1</sup>										
00773 S. sclerotiorum	Leotiomyces	2	3 <sup>1</sup>	1	4, 5, 6, 7, 8, 9, 10, 11, 12, 13								
AN9390 E. nidulans	Eurotiomycetes	1	2										
MGG 04534 M. grisea	Sordariomycete	2	3	1		4							
AN0517 E. nidulans	Eurotiomycetes	1, 2	3				4						

AN0541 <i>E. nidulans</i>	Eurotiomycetes		1					2					
AN5077 <i>E. nidulans</i>	Eurotiomycetes	1, 2	3						4	5			
05454 <i>S. sclerotiorum</i>	Leotiomycetes	2, 3	4	1							5		
04761 <i>C. immitis</i>	Eurotiomycetes	1	2									3	
14527 <i>P. nodorum</i>	Dothideomycetes		1										
04323 <i>A. capsulatum</i>	Eurotiomycetes		1										
EAA75711 <i>G. zeae</i>	Sordariomycete	1, 2	3 <sup>1</sup>										
15411 <i>P. nodorum</i>	Dothideomycetes	1	2 <sup>1</sup>										
MGG 01336 <i>M. grisea</i>	Sordariomycete	2	3 <sup>1</sup>	1									
00677 <i>S. sclerotiorum</i>	Leotiomycetes	2	3 <sup>1</sup>	1									
00533 <i>B. fuckeliana</i>	Leotiomycetes	2	3 <sup>1</sup>										1
02214 <i>P. nodorum</i>	Dothideomycetes	1	2 <sup>1</sup>										

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

#### Group C-II

Protein and species	Taxonomic group	Signal peptide	Peptidoglycan-binding LysM	Chitin-binding type 1 domain	Glycoside hydrolase, family 18, catalytic domain	X8-carbohydrate binding	Transmembrane regions	EGF-like region	Neisseria TspB
AN0509 <i>E. nidulans</i>	Eurotiomycetes			1	2				
07484 <i>N. crassa</i>	Sordariomycete	1	2, 3	3	4	5	6, 7		
05317 <i>N. crassa</i>	Sordariomycete		1	2	3				
EAA68447 <i>G. zeae</i>	Sordariomycete		1, 2	3	4				
02128 <i>S. sclerotiorum</i>	Leotiomycetes		1	2	3				
Chi18-1 <i>H. jecorina</i>	Sordariomycete	1	2, 3	4	5 <sup>1</sup>				
AN7613 <i>E. nidulans</i>	Eurotiomycetes		1	2	3				
EAA77156 <i>G. zeae</i>	Sordariomycete		1, 2	3	4				
Chi18-10 <i>H. jecorina</i>	Sordariomycete	1	2, 3	4	5			6	
Chi18-8 <i>H. jecorina</i>	Sordariomycete	1	2, 3	4	5				
Chi18-9 <i>H. jecorina</i>	Sordariomycete	1	2, 3	4	5				6

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

Group I

Protein and species	Taxonomic group	Immunoglobulin-like fold <sup>1</sup>	Glycoside hydrolase, family 18, catalytic domain	Endoplasmic reticulum targeting sequence	PKD-domain	Carbohydrate-binding family V/XII domain <sup>2</sup>	Gonadotropin-releasing hormone domain	Lipid attachment site	Kringle-like fold
O10363 O. pseudotsugata MNPV	Baculoviridae	1	2	3					
AAP29899 C. fumiferana MNPV	Baculoviridae	1	2	3					
AAV73807 B. mori NPV	Baculoviridae	1	2	3					
AAL56186 H. zea SNPV	Baculoviridae	1	2	3					
AAF33549 S. exigua NPV	Baculoviridae	1	2 <sup>3</sup>	3					
AAZ38189 A. segetum NPV	Baculoviridae	1	2						
AAS82706 A. segetum GV	Baculoviridae	1	2						
ZP 01161077 Photobacterium sp.	Gammaproteobacteria	1	2		3	4, 5	6		
ZP 01236284 V. angustum	Gammaproteobacteria	1	2		3	4, 5	6		
ZP 01159013 Photobacterium sp.	Gammaproteobacteria	1	2		3, 4	5			
ZP 01233025 V. angustum	Gammaproteobacteria	1	2		3, 4	5			
AAO08114 V. vulnificus	Gammaproteobacteria	1	2		3, 4	5			
NP 232428	Gammaproteobacteria	1	2		3, 4	5			

V. cholerae									
YP 435912 H. chejuensis	Gammaproteobacteria		3			1, 2			
YP 529337.1 S. degradans	Gammaproteobacteria	4	2 <sup>4</sup> , 5					1	3
ZP 01159325 Photobacterium sp.	Gammaproteobacteria		1						
ZP 01235109 V. angustum	Gammaproteobacteria		1						
AAO07498 V. vulnificus	Gammaproteobacteria		1						
ZP 01160869 Photobacterium sp.	Gammaproteobacteria		1 <sup>3</sup>						
ZP 01235465 V. angustum	Gammaproteobacteria		1 <sup>3</sup>						
YP 432951 H. chejuensis	Gammaproteobacteria		4			1, 2, 3			

<sup>1</sup> Also classified as PKD-domain and Chitinase A N-terminal-domain.

<sup>2</sup> Also classified as Chitin-binding domain, type 3.

<sup>3</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

<sup>4</sup> The catalytic domain originates from the cluster A.

## Group II

Protein and species	Taxonomic group	Glycoside hydrolase, family 18, catalytic domain	Carbohydrate-binding family V/XII	Carbohydrate-binding, CenC-like / Galactose-binding like	Cellulose-binding, family II, bacterial type	Fibronectin, type III	PKD-domain	Signal peptide	Calcium-binding EF-hand
ZP 01137509 A. cellulolyticus	Actinobacteria	1	2	3					
ZP 01136287 A. cellulolyticus	Actinobacteria	1							
ChiE	Actinobacteria	5			1	2, 3, 4			

S. coelicolor									
NP 903910 C. violaceum	Betaproteobacteria	2	1						
ChiC S. coelicolor	Actinobacteria	4			1	2, 3			
ChiD S. coelicolor	Actinobacteria	1							
YP 529072.1 S. degradans	Gammaproteobacteria	6	1, 3				2, 4, 5		
ZP 01431615 S. tropica	Actinobacteria	3			2			1	4

### Group III

Protein and species	Taxonomic group	Glycoside hydrolase, family 18, catalytic domain	Fibronectin, type III	Cellulose-binding, family II, bacterial type	Aminoacyl-tRNA synthetase	Lipid attachment site	Carbohydrate-binding family V/XII
ZP 00739282 B. thuringiensis	Bacilli	1	2	3	4		
NP 976824 B. cereus	Bacilli	1	2	3	4		
YP 077581 B. licheniformis	Bacilli	1	2	3			
631138 M. xanthus	Deltaproteobacteria	3				1	2

### Group V

Protein and species	Taxonomic group	Glycoside hydrolase, family 18, catalytic domain	Carbohydrate binding family 6	Lipid attachment site	Carbohydrate-binding family V/XII	Orn/DAP/Arg decarboxylase 2
ZP 01247029 F.	Flavobacteria	1	2			

johnsoniae						
ZP 01245634 F. johnsoniae	Flavobacteria	2		1		
ZP 01234549 V. angustum	Gammaproteobacteria	2		1		
ZP 01160697 Photobacterium sp.	Gammaproteobacteria	2		1		
ZP 01158584 Photobacterium sp.	Gammaproteobacteria	1				
YP 434282 H. chejuensis	Gammaproteobacteria	1	2			
YP 436543 H. chejuensis	Gammaproteobacteria	1, 4 <sup>1</sup>			2, 3	5

<sup>1</sup> The catalytic domain originates from the cluster B, group IVa.

#### Group VI

Protein and species	Taxonomic group	PKD-domain	Carbohydrate-binding family V/XII	Glycoside hydrolase, family 18, catalytic domain	Chitinase, C-terminal-domain	DNA-directed DNA polymerase B	Chitin-binding domain type 3	Immunoglobulin-like fold	Lipid attachment site	Kring-like fold	Signal peptide
YP 432439 H. chejuensis	Gammaproteobacteria		1, 3	2							
YP 529449 S. degradans	Gammaproteobacteria	1 <sup>1</sup> , 2, 3, 4	7	5	6						
ChiI S. coelicolor	Actinobacteria		3	1	2						
YP 077582 B. licheniformis	Bacilli			1	2						
NP 902986 C. violaceum	Betaproteobacteria		1	2	3						
ZP 01159608 Photobacterium sp.	Gammaproteobacteria		1, 5	3	4	2					



ZP 01234708 <i>V. angustum</i>	Gammaproteobacteria		1, 5	3	4	2					
NP 233087 <i>V. cholerae</i>	Gammaproteobacteria	2 <sup>2</sup>	1	3	4		5				
AAO07185 <i>V. vulnificus</i>	Gammaproteobacteria		1	2	3		4				
NP 230718 <i>V. cholerae</i>	Gammaproteobacteria		5 <sup>3</sup>	2	3			1			
YP 529337.1 <i>S. degradans</i>	Gammaproteobacteria			2, 5 <sup>4</sup>				4 <sup>1</sup>	1	3	
ChiB <i>S. marcescens</i>	Gammaproteobacteria		3 <sup>3</sup>	2							1

<sup>1</sup> Also classified as Chitinase A N-terminal-domain

<sup>2</sup> Also classified as Immunoglobulin-like fold

<sup>3</sup> Also classified as Chitin-binding domain type 3-domain

<sup>4</sup> The catalytic domain originates from the cluster A, group I.

#### Group VII

Protein and species	Taxonomic group	Peptidoglycan-binding LysM	Glycoside hydrolase, family 18, catalytic domain	Bacterial SH3-like region
NP 979924 <i>B. cereus</i>	Bacilli	1, 2	3 <sup>1</sup>	
YP 077302 <i>B. licheniformis</i>	Bacilli	1, 2	3 <sup>1</sup>	
YP 074366 <i>S. thermophilum</i>	Actinobacteria	1, 2	3 <sup>1</sup>	
YP 075811 <i>S. thermophilum</i>	Actinobacteria	1, 2	3 <sup>1</sup>	
ZP 01354047 <i>C. phytofermentans</i>	Clostridia	1, 2	3 <sup>1</sup>	
YP 079121 <i>B.</i>	Bacilli	1, 2	3 <sup>1</sup>	

licheniformis				
YP 091539 B. licheniformis	Bacilli	1, 2	3 <sup>1</sup>	
YP 075619 S. thermophilum	Actinobacteria		1 <sup>1</sup>	
NP 902406 C. violaceum	Betaproteobacteria		1 <sup>1</sup>	
ZP 01136704 A. cellulolyticus	Actinobacteria		1 <sup>1</sup>	
ZP 00526542 S. usitatus	Solibacteres		1 <sup>1</sup>	
ZP 01354926 C. phytofermentans	Clostridia		2	1

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

#### Group VIII

Protein and species	Taxonomic group	Glycoside hydrolase, family 18, catalytic domain	Fibronectin, type III	Carbohydrate-binding family V/XII
NP 469498 L. innocua	Bacilli	1 <sup>1</sup>	2	3
YP 012734 L. monocytogenes	Bacilli	1 <sup>1</sup>	2	3
ZP 01246933 F. johnsoniae	Flavobacteria	1		
ZP 01354150 C. phytofermentans	Clostridia	1		

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

Group Archaeal (Cluster A)

Protein and species	Taxonomic group	Fibronectin type III	Carbohydrate-binding family V/XII	Glycoside hydrolase, family 18, catalytic domain	Proline rich extensin signature	Cellulose-binding, family II, bacterial type	PKD-domain	Haem peroxidase, plant/fungal/bacterial
ChiA P. kodakaraensis	Euryarchaeota	1, 3	2	4, 7 <sup>1</sup>		5, 6		
Q8U1H4 P. furiosus	Euryarchaeota	1, 3	2	4				
Q8U1H5 P. furiosus	Euryarchaeota			1	2, 3, 4, 6	5		
AAG19274 Halobacterium sp.	Euryarchaeota		1	3			2	
AAG19275 Halobacterium sp.	Euryarchaeota		1	4			2	3

<sup>1</sup> The catalytic domain originates from cluster C.

Group *C. elegans* (Cluster A)

Protein and species	Taxonomic group	Glycoside hydrolase, family 18, catalytic domain	Transmembrane region	Chitin-binding, type 1-domain	Peptidoglycan-binding LysM	Plant self-incompatibility S1
CE06282 C. elegans	Rhabditida	1 <sup>1</sup>				
CE23907 C. elegans	Rhabditida	2 <sup>1</sup>	1			
CE40270 C. elegans	Rhabditida	1 <sup>1</sup>				
CE03547 C. elegans	Rhabditida	1 <sup>1</sup>				
CE25073 C. elegans	Rhabditida	1 <sup>1</sup>				

CE03548 C. elegans	Rhabditida	1 <sup>1</sup>				
CE32917 C. elegans	Rhabditida	1 <sup>1</sup>				
CE34579 C. elegans	Rhabditida	1 <sup>1</sup>				
CE03542 C. elegans	Rhabditida	1 <sup>1</sup>				
CE03543 C. elegans	Rhabditida	1 <sup>1</sup>				
CE03544 C. elegans	Rhabditida	1 <sup>1</sup>				
CE12474 C. elegans	Rhabditida	1 <sup>1</sup>				
CE15845 C. elegans	Rhabditida	1 <sup>1</sup> , 2 <sup>1</sup>				
CE32863 C. elegans	Rhabditida	1 <sup>1</sup> , 2 <sup>1</sup>				
CE34961 C. elegans	Rhabditida	1 <sup>1</sup>				
CE20043 C. elegans	Rhabditida	1 <sup>1</sup>				
CE08731 C. elegans	Rhabditida	1 <sup>1</sup>				
CE08730 C. elegans	Rhabditida	1 <sup>1</sup>				
CE28638 C. elegans	Rhabditida	2 <sup>1</sup>	1			
CE01673 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02997 C. elegans	Rhabditida	1 <sup>1</sup>				

CE40296 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02990 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02995 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02993 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02996 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02992 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02999 C. elegans	Rhabditida	1 <sup>1</sup>				
CE02998 C. elegans	Rhabditida	1 <sup>1</sup>				
CE36953 C. elegans	Rhabditida	1		2, 3, 4		
CE09229 C. elegans	Rhabditida	15		16, 17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13	14
CE12928 C. elegans	Rhabditida	7		8, 9	1, 2, 3, 4, 5	6

<sup>1</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

## **Cluster B**

Group B-I through B-IV

Protein	Taxonomic group	Signal peptide	Glycoside hydrolase, family 18,	Fungal cellulose binding	Chitin-binding, type 1	MANSC seven cysteine	Glycoside hydrolase family 18,	Carbohydrate-binding family	Subtilase family, serine
---------	-----------------	----------------	---------------------------------	--------------------------	------------------------	----------------------	--------------------------------	-----------------------------	--------------------------

			catalytic domain	region	domain	domain	carbohydrate-binding domain	V/XII <sup>3</sup>	active site
05455 <i>B. fuckeliana</i>	Leotiomyce	1	2	3					
11212 <i>S. sclerotiorum</i>	Leotiomyce		1	2					
14944 <i>B. fuckeliana</i>	Leotiomyce	1	2	3					
08695 <i>S. sclerotiorum</i>	Leotiomyce	1	2	3					
04500 <i>N. crassa</i>	Sordariomyce		1	2					
MGG 01333 <i>M. grisea</i>	Sordariomyce	1	2	3					
AN11059 <i>E. nidulans</i>	Eurotiomyce	1 <sup>1</sup>	2						
07348 <i>C. immitis</i>	Eurotiomyce		1						
Chi18-12 <i>H. jecorina</i>	Sordariomyce		1						
Chi18-17 <i>H. jecorina</i>	Sordariomyce		1	2					
08680 <i>P. nodorum</i>	Dothideomycetes		1						
13155 <i>S. sclerotiorum</i>	Leotiomyce		1						
EAA72615 <i>G. zeae</i>	Sordariomyce		1						
Chi18-18 <i>H. jecorina</i>	Sordariomyce		1						
02184 <i>N. crassa</i>	Sordariomyce		1						
MGG 03599 <i>M. grisea</i>	Sordariomyce	1 <sup>1</sup>	2		3 <sup>2</sup>				
16099 <i>R. oryzae</i>	Mucormycetes		1						
10112 <i>R. oryzae</i>	Mucormycetes	1	2			3	4		
04523 <i>R. oryzae</i>	Mucormycetes		1				2		
11152 <i>R. oryzae</i>	Mucormycetes	1	2				3		
08153 <i>R. oryzae</i>	Mucormycetes		1					2, 3	
04407.1 <i>C. cinereus</i>	Homobasidiomyce	1	2					3	
40899 <sup>4</sup> <i>P. chrysosporium</i>	Homobasidiomyce	1	2						
6412 <i>P. chrysosporium</i>	Homobasidiomyce	1	2					3, 4	
2991 <i>P. chrysosporium</i>	Homobasidiomyce	1	2					3	
301747 <i>L. bicolor</i>	Homobasidiomyce	1	2					3, 4	
NP 594130 <i>S. pombe</i>	Schizosaccharomyce		1 <sup>5</sup>						

	te								
00864 P. nodorum	Dothideomycetes		1		2				
AN8241 E. nidulans	Eurotiomycete		1						
Chi18-14 H. jecorina	Sordariomycete		1	3					2
Chi18-16 H. jecorina	Sordariomycete		1	2					
CAC07216 M. anisopliae	Sordariomycete		1	2					
Chi18-13 <sup>4</sup> H. jecorina	Sordariomycete	1	2						
Chi30 M. anisopliae	Sordariomycete		1						
XP 452410 K. lactis	Saccharomycete		1				2		
CTS1 S. cerevisiae	Saccharomycete		1				2		
CHT1 C. albicans	Saccharomycete		1						
CHT2 C. albicans	Saccharomycete	1	2						
03890 L. elongisporum	Saccharomycete	1	2						
CHT3 C. albicans	Saccharomycete	1	2						
05283 L. elongisporum	Saccharomycete	1	2						

<sup>1</sup> Also classified as a Prokaryotic membrane lipoprotein lipid attachment site.

<sup>2</sup> Also classified as a Glycoside hydrolase family 19 catalytic domain.

<sup>3</sup> Also classified as a Chitin-binding domain type 3.

<sup>4</sup> Partial sequence.

<sup>5</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

### Plant class III

Protein	Taxonomic group	Signal peptide	Glycoside hydrolase, family 18, catalytic domain
BAA21861.1 A. thaliana	Brassicales	1 <sup>1</sup>	2
AAG25709 M. domestica	Rosales		1
ABA26457 C. lanatus	Cucurbitales		1

CAJ43737 <i>C. arabica</i>	Gentianales		1
Hevamine H. <i>brasiliensis</i>	Malpighiales		1
Q8L5C6 <i>T. aestivum</i>	Poales		1 <sup>2</sup>
Q53NL5 <i>O. sativa</i>	Poales	1	2 <sup>2</sup>
ConcanavalinB <i>C. ensiformis</i>	Fabales		1 <sup>2</sup>

<sup>1</sup> Also classified as a Prokaryotic membrane lipoprotein lipid attachment site.

<sup>2</sup> Consensus sequence of catalytic domain, (LIVMFY)-(DN)-G-(LIVMF)-(DN)-(LIVMF)-(DN)-X-E is not intact.

#### Group B-V

Protein	Taxonomic group	Signal peptide	Glycoside hydrolase, family 18, catalytic domain	Alkaline phosphatase domain	Subtilase serine protease
Chi18-20 <i>H. jecorina</i>	Sordariomycete	1	2		
08394 <i>N. crassa</i>	Sordariomycete		1		
MGG 01876 <i>M. grisea</i>	Sordariomycete	1	2		
Chi18-19 <i>H. jecorina</i>	Sordariomycete		1		
01393 <i>N. crassa</i>	Sordariomycete		1		
XP 386145 <i>G. zeae</i>	Sordariomycete		1		
01513 <i>S. sclerotiorum</i>	Leotiomycete	1	2		
05559 <i>B. fuckeliana</i>	Leotiomycete	1	2		
MGG 04073 <i>M. grisea</i>	Sordariomycete		1		
06595 <i>C. immitis</i>	Eurotiomycete		1		
AN8245 <i>E. nidulans</i>	Eurotiomycete		1	2	
XP 382346 <i>G. zeae</i>	Sordariomycete		1		
03420 <i>S. sclerotiorum</i>	Leotiomycete	1	2		
07854 <i>B. fuckeliana</i>	Leotiomycete	1	2		
236861 <i>L. bicolor</i>	Homobasidiomycete		1		



185397 <i>L. bicolor</i>	Homobasidiomycete		1		
UM02758 <i>U. maydis</i>	Ustilaginomycete		2		1
02351 <i>F. neoformans</i>	Heterobasidiomycete	1	2		

#### Group IVa

Protein and species	Taxonomic group	Carbohydrate-binding family V/XII <sup>1</sup>	Immunoglobulin-like, PKD domain	Glycoside hydrolase, family 18, catalytic domain	Signal peptide	Fibronectin type III domain	Carbohydrate-binding, CenC-like domain	Bacterial cellulose-binding family II	Carbohydrate-binding fold	Carboxypeptidase regulatory region
ZP 01159542 <i>Photobacterium</i> sp.	Gammaproteobacteria	1, 3	2	4						
ZP 01234779 <i>V. angustum</i>	Gammaproteobacteria	1, 4	2, 3, 5	6						
NP 231586 <i>V. cholerae</i>	Gammaproteobacteria	2, 3		4	1					
AAO10239 <i>V. vulnificus</i>	Gammaproteobacteria	1, 2	3	4						
YP 527176.1 <i>S. degradans</i>	Gammaproteobacteria	1, 2		3						
NP 902605 <i>C. violaceum</i>	Betaproteobacteria	2		3	1 <sup>2</sup>					
NP 980050 <i>B. cereus</i>	Bacilli			1						
ZP 00742457 <i>B. thuringiensis</i>	Bacilli			1						
634143 <i>M. xanthus</i>	Deltaproteobacteria	1		4		2, 3				
YP 432187 <i>H. chejuensis</i>	Gammaproteobacteria	1, 3	2, 4, 5	6						
YP 436543 <i>H. chejuensis</i>	Gammaproteobacteria	2, 3		1 <sup>3</sup> , 4						
ChiB <i>S. coelicolor</i>	Actinobacteria			3		2	1			
ChiA <i>S. coelicolor</i>	Actinobacteria			3		2	1			

ZP 01353924 C. phytofermentans	Clostridia	1		2						
ZP 01430879 S. tropica	Actinobacteria			2				1		
ZP 01245639 F. johnsoniae	Flavobacteria			1 <sup>4</sup> , 2					2, 3, 4, 6, 7	5
ChiJ S. coelicolor	Actinobacteria			1						
Chi18-15 H. jecorina	Sordariomycete			1						

<sup>1</sup> Also classified as a Chitin-binding domain type 3.

<sup>2</sup> Also classified as a prokaryotic lipoprotein attachment site.

<sup>3</sup> The catalytic domain originates from cluster A, group V.

<sup>4</sup> The catalytic domain originates from cluster A.

#### Group IVb

Protein and species	Taxonomic group	Thioredoxin related fold	Carbohydrate-binding family V/XII <sup>1</sup>	Glycoside hydrolase, family 18, catalytic domain	Fibronectin type III domain	Blue type 1 copper domain	Signal peptide	Bacterial Ig fold
YP 349297 P. fluorescens	Gammaproteobacteria	1, 3		2				
YP 608169 P. entomophila	Gammaproteobacteria			1				
ChiC S. marcescens	Gammaproteobacteria		3	1	2			
YP 237842 P. syringae	Gammaproteobacteria		6	1	2, 3, 4, 5	7		
NP 814153 E. faecalis	Bacilli			1				
YP 014504 L. monocytogenes	Bacilli	1, 2		3				
NP 471330 L. innocua	Bacilli	1, 2		3				
NP 719615 S.	Gammaproteobacteria		2, 3	1				

oneidensis	ia							
ZP 01236292 V. angustum	Gammaproteobacteria		4, 5	2			1 <sup>2</sup>	3
YP 071855 Y. pseudotuberculosis	Gammaproteobacteria		3	1	2			
YP 053588 M. florum	Mollicutes			2			1 <sup>2</sup>	

<sup>1</sup> Also classified as a Chitin-binding domain type 3.

<sup>2</sup> Also classified as a prokaryotic lipoprotein attachment site.

## **Cluster C**

Protein and Species	Taxonomic group	Signal peptide	Transmembrane-region	Glycoside hydrolase, family 18 catalytic domain	Carbohydrate-binding family V/XII	Fibronectin type III	Membrane lipoprotein/lipid attachment site	Carbohydrate-binding family II	Serralysin-like metalloprotease	Chitin-binding domain type 3	Na-Ca exchange domain	Glycoside hydrolase, family 5 (cellulase)	Carbohydrate-binding family IV/IX
621070 B. cenocepacia	Betaproteobacteria	1	2	3	4, 5								
ZP 01331104 B. pseudomallei	Betaproteobacteria	1	2	3	4, 5								
NP 901110 C. violaceum	Betaproteobacteria			1	2, 3								
ZP 00739603 B. thuringiensis	Bacilli			2	1, 5	3, 4							
YP 124481 L. pneumophila	Gammaproteobacteria			1									
ChiK S. coelicolor	Actinobacteria			2			1						
ZP 00570566 Frankia sp.	Actinobacteria			2				1					
ZP 01199781 X.	Alphaproteobacteria			4	2	3			1				

autotrophicus													
ChiA P. kodakaraensis	Euryarchaeot a			4 <sup>1</sup> , 7	2	1, 3		5, 6					
ZP 00923356 E. coli	Gammaprotei bacteria	1		8						1, 2, 3, 4, 5, 6, 7			
ZP 01191795 M. flavescens	Actinobacteri a			4				6, 8	1		2, 3, 5, 7, 9, 11	10	
ChiH S. coelicolor	Actinobacteri a			1	2								
ZP 01137245 A. cellulolyticus	Actinobacteri a			1	2								3

<sup>1</sup> The catalytic domain originates from the Archaea-group, cluster A.