

AEM03803-12: Classification correlating to function of carbohydrate-active enzymes by recognition of short, conserved peptide motifs.

Supplemental Material.

Accession numbers of eukaryotic, characterized GH5 proteins analyzed by Peptide Pattern Recognition

ACR82487.1, AAK21882.1, ABY28340.1, BAB62319.1, BAB86867.1, AAC63094.1, AAC33848.1, BAA36216.1, BAD90558.1, CAA61740.1, AAE60102.1, AAC49731.1, ABF50872.1, AAC60541.1, AAL33630.1, BAA29030.1, BAB62317.1, AAR29981.1, AAB69348.1, ACD36972.1, AAN03647.1, CAH69214.1, AAE59927.1, AAK21881.1, AAB51451.1, AAX18655.1, BAD72778.1, ACB06750.1, CAB92326.1, AAK21883.2, AAM21469.1, ABF50886.1, CAA86948.1, CAA11018.1, AAM08614.1, AAF65310.1, AAV05307.1, CAY69081.1, CAA86950.1, AAT97707.1, CAB50968.1, ABK27195.1, AAL84696.1, AAM26920.1, CAK96471.1, AAL01213.1, EAL85463.1, ACH56965.1, ABF50878.1, AAB87859.2, ABV68808.1, ABG79370.1, ACH58411.1, ABC59553.1, AAA67426.1, ACH58410.1, CAA90423.1, ACU52526.1, BAD67544.1, AAG59832.1, AAC05164.1, AAC15708.1, BAA12676.1, ACH63253.1, AAA34213.1, ADH51728.1, AAG50051.1, ABY52965.1, CAA53631.1, ACN43345.1, BAD01163.1, AAD04193.1, BAA90480.1, ACE06751.1, AAC15707.1, AAC48341.1, AAN03645.1, ABF50848.1, AAL16412.1, AAD43818.1, AAC08587.1, BAA76394.1, AAD45868.1, ABZ29259.1, AAC48325.1, AAK85303.1, AAB03889.1, AAB69347.1, CAA86949.1, AAU09785.1, BAE58099.1, CAA63536.1, BAD97445.1, AAA34599.1, CAA86951.1, ACP74152.1, CAA86952.1, ABK27199.1, CAA21163.1, ABF50867.1, ACE81431.1, AAN04103.1, ABF50861.1, CAC08442.1, ADO14134.1, ADN93457.1, CAC51690.3, AAG00315.1, ABB88954.1, ABJ41267.1, BAE78456.1, CAC81056.1, AAX01860.1, ABF50863.1, AAA34208.1, AAK56557.1, ACJ06979.1, ADF28533.1

Accession numbers of 743 GH61 proteins analyzed by Peptide Pattern Recognition

Q5BCX8.1, ABT35335.1, ACE10235.1, ACE10233.1, ABW56451.1, ACE10231.1, A1C4H2.1, ACE10232.1, ACE10234.1, AAA53434.1, EEH05325.1, EGC46978.1, XP_001543425.1, EGE82146.1, EGE79257.1, EEQ84968.1, EEQ89326.1, XP_002626386.1, XP_002628516.1, XP_003015133.1, XP_003174002.1, XP_002848955.1, XP_001268194.1, XP_001271679.1, XP_001269010.1, EAW13459.1, XP_001273555.1, EAW14268.1, XP_001274726.1, XP_002380189.1, XP_002382772.1, XP_002381515.1

,XP_002376462.1,EDP54951.1,EDP48657.1,EDP47167.1,EAL88805.1,XP_748707.1,EAL93749.1,XP_747482.1,XP_748715.1,XP_752040.1,CAF31975.1,BAB62318.1,EAA59072.1,XP_868906.1,CBF76006.1,CBF73467.1,XP_660650.1,EAA64499.1,EAA58450.1,EAA59125.1,CBF74037.1,XP_658645.1,EAA59545.1,XP_001402213.2,XP_001397023.2,XP_001395335.2,XP_001392640.1,XP_001400893.2,XP_001402213.1,CAK41095.1,XP_001400893.1,XP_001395530.1,CAK42466.1,CAK97324.1,XP_001820737.1,BAE56764.1,XP_001826694.1,BAE55582.1,BAE61530.1,XP_001821292.1,XP_001822663.2,BAE58643.1,XP_001820645.2,XP_001817584.2,BAE64395.1,BAE60320.1,EAU31074.1,XP_001208821.1,XP_001209496.1,XP_001216541.1,XP_001210534.1,XP_001216411.1,EAU33838.1,XP_001214594.1,XP_001213388.1,XP_001208400.1,EAU34150.1,XP_001216734.1,EDW20470.1,ABV60986.1,CAJ81216.1,CAJ81218.1,CAJ81217.1,CAJ81215.1,XP_001546774.1,XP_001553240.1,XP_001546354.1,XP_001560175.1,XP_001552214.1,XP_001558279.1,XP_001560894.1,CCD50451.1,CCD50139.1,CCD47228.1,CCD52645.1,CCD34368.1,CCD51504.1,CCD50144.1,CCD48549.1,CCD49290.1,XP_001225249.1,XP_001230041.1,XP_001230192.1,XP_001222998.1,XP_001225243.1,XP_001227602.1,XP_001223598.1,XP_001222154.1,EAQ86198.1,XP_001225332.1,EAQ85056.1,XP_001226408.1,XP_001227307.1,XP_001224385.1,XP_001229931.1,EAQ91757.1,XP_001223312.1,XP_001227193.1,EAQ87022.1,XP_001226548.1,EAQ92448.1,XP_001222797.1,XP_001224181.1,XP_001225324.1,XP_001229971.1,XP_001224238.1,XP_001220412.1,XP_001226220.1,XP_001222217.1,XP_001225692.1,XP_001219301.1,XP_001225412.1,EAQ88640.1,EAQ85642.1,EAQ83639.1,XP_001225930.1,XP_001227732.1,XP_001219583.1,XP_001228820.1,EAQ89525.1,XP_001219444.1,EAQ86619.1,XP_001227508.1,XP_001223687.1,EGS17840.1,EGS20384.1,EGS21628.1,EGS21609.1,EGS20667.1,EGS22278.1,EGS17558.1,EGS21088.1,EGS20367.1,EGS19579.1,EGS23404.1,EGS23697.1,EGS21285.1,EGS24042.1,EGS19451.1,EGS18626.1,EGS17857.1,XP_001242606.1,EER27647.1,AAM76663.1,XP_002910838.1,XP_001832515.2,XP_001841257.2,EAU80944.1,XP_001841005.1,XP_001832374.2,EAU87138.1,XP_001838602.1,XP_001832512.1,EAU80738.1,XP_001840230.2,EAU89754.1,XP_001840824.2,EAU80891.2,XP_001832511.1,XP_001834453.2,XP_001832364.2,XP_001834810.1,XP_001832633.2,XP_002911925.1,XP_001835420.2,EAU81728.2,EAU80304.1,XP_001839611.1,XP_001832109.2,EAU86100.1,XP_001841073.2,EAU91503.2,XP_001834458.1,XP_001835421.1,EAU87436.2,EAU83857.1,EAU85243.2,XP_002910575.1,XP_001834768.1,CAG27578.1,ADV19810.1,XP_777868.1,AAW41121.1,ADX07320.1,EGU85487.1,EGU84223.1,EGU75413.1,EGU75302.1,EGU72891.1,EGU85465.1,EGU84205.1,EGU77464.1,EGU80176.1,EGU85118.1,EGU86650.1,EGU74313.1,EGU79199.1,EGU87361.1,XP_386573.1,XP_386263.1,XP_388187.1,XP_388088.1,XP_386286.1,XP_382378.1,XP_384949.1,XP_383871.1,XP_386384.1,XP_384857.1,XP_384144.1,XP_3

91664.1,AEJ35168.1,EFQ25679.1,EFQ34021.1,EFQ36570.1,EFQ25116.1,EFQ31502.1,EFQ35295
.1,EFQ35966.1,EFQ34537.1,EFQ28963.1,EFQ34071.1,EFQ35170.1,EFQ36324.1,EFQ31130.1,EF
Q30363.1,EFQ27332.1,EFQ34935.1,EFQ34588.1,EFQ36036.1,EFQ27394.1,EFQ30762.1,EFQ364
23.1,EFQ29051.1,EFQ34827.1,EFQ25762.1,EFQ34070.1,EFQ33926.1,EFQ33676.1,EFQ35556.1,
CAQ16206.1,CAQ16278.1,CAQ16208.1,CAQ16217.1,EFX04021.1,EFX04819.1,CAG27577.1,C
AA71999.1,AAP57753.1,XP_001883971.1,XP_001883194.1,XP_001883867.1,XP_001877259.1,
EDR01287.1,XP_001882292.1,XP_001885903.1,EDR07117.1,EDR06657.1,EDR15540.1,XP_001
881959.1,CBX98254.1,CBX96476.1,CBX96949.1,CBY01256.1,CBY02242.1,CBX96550.1,CBX9
4572.1,CBY01974.1,CBX91667.1,CBX98126.1,CBY00196.1,CBX91886.1,CBX91313.1,CBX939
65.1,CBY01257.1,CBX95655.1,CBX94224.1,CBX97718.1,CBY01204.1,CBX94532.1,CBX93546.
1,EAA51422.1,EDK04335.1,EAA56945.1,EAA54572.1,EAA56258.1,EAA53354.1,EAA53298.1,
XP_370106.1,EAA50788.1,EAA53409.1,EAA54517.1,EAA57051.1,XP_001404656.1,XP_001408
162.1,XP_369715.2,XP_362640.1,XP_361583.1,XP_370106.2,XP_362483.1,XP_367775.2,XP_00
1410346.1,XP_364864.1,XP_001522073.1,XP_001408450.1,XP_001405918.1,XP_369395.1,XP_
364682.1,XP_001406958.1,XP_367664.2,EGG03285.1,EGG03321.1,EGG08525.1,EGG05780.1,E
FY92676.1,EFY91792.1,EFZ04469.1,XP_002392372.1,XP_002391587.1,XP_002389015.1,XP_00
2388995.1,EEB93545.1,XP_002391588.1,XP_002392321.1,XP_002386353.1,XP_002389005.1,X
P_002386746.1,XP_002393652.1,XP_002391345.1,XP_002389159.1,XP_002399025.1,XP_00239
8849.1,XP_002386478.1,XP_002392498.1,XP_002390217.1,XP_002399040.1,AEO59955.1,AEO
60271.1,AEO58412.1,AEO58921.1,AEO54509.1,AEO56016.1,AEO55776.1,AEO58169.1,AEO56
542.1,AEO56498.1,AEO56547.1,AEO56416.1,AEO59482.1,AEO61257.1,AEO61305.1,AEO5983
6.1,AEO55082.1,AEO59823.1,AEO61304.1,AEO56642.1,AEO55652.1,AEO56665.1,EGP90215.1
,EGP91468.1,XP_003042567.1,XP_003043105.1,XP_003046131.1,XP_003041587.1,XP_0030459
40.1,EEU39531.1,XP_003041666.1,XP_003040612.1,XP_003043569.1,XP_003044062.1,XP_003
050482.1,XP_003050495.1,XP_003045670.1,XP_001260913.1,XP_001267517.1,XP_001266231.1
,XP_001267236.1,XP_001264507.1,EAW16173.1,XP_001259147.1,XP_001262213.1,XP_001260
907.1,EAA29018.1,EAA33408.1,XP_959499.1,XP_958583.1,EAA36150.2,XP_962414.1,XP_963
702.1,EAA36362.1,EAA26873.1,EAA29132.1,CAD21296.1,CAB97283.2,CAE81966.1,CAD7034
7.1,CAF05857.1,EGO58503.1,EGO54705.1,EGO56889.1,EGO52064.1,EGO53412.1,EGO58845.1
,EGO55727.1,EGO61029.1,EGO53245.1,EGO55619.1,EGO61608.1,EGO53330.1,EGO55637.1,E
GO59472.1,EGO58321.1,XP_002559170.1,CAP92380.1,XP_002558168.1,CAP86439.1,XP_0021
46302.1,EAT81259.2,XP_001791800.1,XP_001792546.1,EAT89418.1,EAT77790.2,XP_00179082
6.1,EAT83900.1,XP_001802947.1,XP_001791868.1,XP_001802011.1,XP_001790904.1,XP_0017

96146.1,EAT81480.1,XP_001791556.1,EAT82406.1,XP_001796313.1,XP_001795846.1,XP_001804665.1,EAT77212.1,XP_001792455.1,XP_001802269.1,XP_001793524.1,EAT92310.2,XP_001797837.1,XP_001806210.1,XP_001802359.1,XP_001796133.1,XP_001801858.1,XP_001791546.1,AAM22493.1,CCA74246.1,CCA72192.1,CCA68244.1,CCA72221.1,CCA70703.1,CCA67659.1,CCA73144.1,CCA75037.1,CCA74449.1,CCA72183.1,CCA76671.1,CCA72220.1,CCA73151.1,CCA67656.1,CCA66803.1,CCA74814.1,CCA70418.1,CCA71764.1,CCA70417.1,CCA77877.1,CCA76320.1,CCA67658.1,CCA67657.1,CCA70035.1,CCA72182.1,CAP64732.1,CAP59702.1,XP_001909024.1,XP_001906078.1,XP_001907502.1,CAP68352.1,XP_001906822.1,CAP65855.1,XP_001907069.1,XP_001906810.1,CAP67176.1,CAP68375.1,CAP73072.1,XP_001904712.1,CAP68309.1,XP_001906532.1,CAP67190.1,CAP61048.1,XP_001911495.1,CAP71839.1,CAP73311.1,CAP71532.1,CAP72740.1,CAP65971.1,XP_001904958.1,CAP67466.1,CAP65111.1,CAP61650.1,CAP65866.1,XP_001903620.1,CAP61476.1,XP_001909116.1,CAP73254.1,XP_002474489.1,XP_002476952.1,XP_002474815.1,EED85774.1,XP_003336129.1,XP_003335269.1,XP_003325408.1,XP_003299575.1,XP_003296866.1,XP_003297809.1,XP_003302923.1,XP_003301020.1,XP_003300728.1,XP_003299166.1,XP_003303895.1,XP_003300926.1,XP_003299671.1,XP_003304335.1,XP_003296033.1,XP_003306650.1,XP_003298432.1,XP_003302575.1,XP_003298498.1,XP_003305323.1,XP_003300983.1,XP_003299949.1,XP_003301611.1,XP_003303817.1,XP_003305915.1,XP_003299155.1,XP_003300990.1,XP_003306203.1,EDU50404.1,XP_001935377.1,XP_001931371.1,XP_001933351.1,EDU46492.1,XP_001932393.1,XP_001934008.1,XP_001940528.1,EDU49193.1,XP_001937216.1,XP_001932895.1,XP_001934763.1,XP_001939876.1,EDU48860.1,XP_001933274.1,XP_001935843.1,XP_001936296.1,XP_001930525.1,XP_001939155.1,XP_001936652.1,XP_001932318.1,XP_001941054.1,XP_001930436.1,XP_001936366.1,XP_001933725.1,EGU12834.1,XP_003031059.1,XP_003031883.1,XP_003034283.1,XP_003026941.1,XP_003032377.1,XP_003031503.1,EFI94694.1,XP_003033157.1,EFJ00004.1,XP_003029539.1,XP_003029596.1,XP_003028465.1,EFI92060.1,EFI92963.1,XP_003037795.1,XP_003027612.1,EFI92604.1,XP_003026867.1,XP_003026965.1,EFI95045.1,XP_003027567.1,EFJ03083.1,XP_003032165.1,EFI91671.1,XP_003029697.1,EFI92673.1,XP_001590753.1,XP_001589530.1,EDN93954.1,XP_001597642.1,XP_001584877.1,EDO00181.1,XP_001587030.1,XP_001591031.1,EDO00288.1,EGN99889.1,EGO03895.1,EGN93952.1,EGN94279.1,XP_003349473.1,XP_003351982.1,XP_003352382.1,XP_003352682.1,XP_003345666.1,XP_003352294.1,XP_003351237.1,XP_003348785.1,XP_003348222.1,CBI54781.1,CBI52679.1,CBI58321.1,CBI56951.1,CBI57737.1,CBI58198.1,CBI54870.1,CBI52566.1,CBI54738.1,CBI56992.1,CBI56993.1,YP_001260935.1,XP_002478590.1,AEO65580.1,AEO71031.1,AEO68023.1,AEO64605.1,AEO69043.1,AEO67396.1,AEO68157.1,AEO64177.1,AEO67395.1

,AEO68763.1,AEO66274.1,AEO68577.1,AEO62422.1,CAG27576.1,EGR49412.1,EGR49821.1,ADB89217.1,ACH92573.1,ACD36971.1,ADJ57703.1,EGE04098.1,XP_003235451.1,EGD92942.1,XP_003019976.1,XP_002838651.1,CAZ84655.1,XP_002544405.1,XP_003001381.1,XP_003007169.1,XP_003006393.1,EEY23940.1,XP_003002463.1,XP_003000738.1,XP_003005258.1,XP_003002719.1,XP_003002930.1,XP_003000747.1,XP_003002008.1,XP_003002068.1,XP_003008052.1,XP_003009235.1,XP_003004661.1,XP_003000578.1,EEY22380.1,XP_003001513.1,EEY16323.1,XP_003007405.1,XP_003000367.1,XP_003003091.1,AAT64005.1,NP_001130665.1,ACF86151.1,,JGI_PC15_v2.0_4,JGI_PC15_v2.0_13,JGI_PC15_v2.0_16,JGI_PC15_v2.0_6,JGI_PC15_v2.0_5,EGO19398.1,JGI_PC15_v2.0_9,JGI_PC15_v2.0_8,JGI_PC15_v2.0_7,JGI_PC15_v2.0_1,JGI_PC15_v2.0_14,JGI_PC15_v2.0_11,JGI_PC15_v2.0_2,JGI_PC15_v2.0_15,JGI_PC15_v2.0_10,EGN94280.1,JGI_PC15_v2.0_3,JGI_PC15_v2.0_12,JGI_AB_v2.0_1,JGI_AB_v2.0_2,JGI_AB_v2.0_3,JGI_AB_v2.0_4,JGI_AB_v2.0_5,JGI_AB_v2.0_6,JGI_AB_v2.0_7,JGI_AB_v2.0_8,JGI_AB_v2.0_9,GWW2.10.277.1,AAM22493.1,GWW2.2.628.1,GWW2.10.287.1,FGENESH1_PG.C_SCAFFOLD_8000142,E_GWH2.5.359.1,E_GWH2.2.646.1,EGO19770.1

Accession numbers are from Genbank except JGI_AB_v2.0_1 - JGI_AB_v2.0_9 from Joint Genome Institute *Agaricus bisporus* genome project and JGI_PC15_v2.0_1 - JGI_PC15_v2.0_16 from Joint Genome Institute *Pleurotus ostreatus* genome project. These sequence data were produced by the US Department of Energy Joint Genome Institute <http://www.jgi.doe.gov/> in collaboration with the user community. GWW2.10.277.1, GWW2.10.287.1, GWW2.2.628.1, E_GWH2.2.646.1, E_GWH2.5.359.1 and FGENESH1_PG.C_SCAFFOLD_8000142 were found in the *Phanerochaete chrysosporium* genome (Genbank accession number AADS00000000.1).

DNA and amino acid sequences of GH61 PCR products amplified with degenerated primers

Chaetomium senegalense

DNA:

ctggagttttcaagatcgacgaggtcggcctggcgcacggcgccaacgcgccccggttctggggctcggaccagctgatcgccaacaacgcc
gcgtggatggtccagatccccgaggacatcgtccccgc

Amino acid:

LEFFKIDEVGLVDGANAPGFWGSDQLIANNAAWMVQIPEDIAP

Chaetomium thermophilum

DNA:

cacattgtcacaagtcggccacgaacgccaaagggccatgccgtcatcaaggccgggtgactctgtctacatccagtgggacacctggcccag
tcgccaccagggccggtcacgactacctgccagctgggcagcggcggtgcgagacggtcgacaagaccagctcgagttctcaagat
cgccgagggcggctgattgacggctcccaggtcccggaaagtgggctgccgatcagctcgcgccagaacaactcgtggctggcacca
tccccgagaatatcaagccgct

Amino acid:

HICHKSATNAKGHAVIKAGDSVYIQWDTWPESHGHPVIDYLASCGSAGCETVDKTQLEFF
KIAEAGLIDGSQAPGKWAADQLIAQNNSWLVTIPENIKP

Corynascus thermophilus

DNA:

gtcggccaccaacgccggcggcatgccgttgtgccggcgacaagattccatccagtgggacacctggcccagtcgcaccacggtc
cggtcacgactacctgccgactggcgacggcggtgcgagaaggtcgacaagaccacgctcgagttctcaagatcagcgagaaggg
cctgatcgacggcagcagcgcgccggcaggtggcggtccgacgagctgatcgccaacaacaactcgtggctgggtccagatcccggcca
catcgcccc

Amino acid:

ICHKSATNAGGHAVVAAGDKISIQWDTWPESHGHPVIDYLADCGDAGCEKVDKTTLEFFK
ISEKGLIDGSSAPGRWASDELIANNNSWL VQIPPDIAPA

Melanocarpus albomyces

DNA:

tatttgcacaagagcgcgacccccggcgggcgccacgccacggtcgctgccggcgacaagatcagcctggctggacccccgagtggccc
gagagccacattggccccgtcatcgactacatggcggtgcaacggcgactgcgagacggtaacaaggagtcgctcggtgggtcaagat
cgacggcgccggctacgacagcagcaagggccagtgggccgccgacgcgctgcgtgagaacggcaacagctggctgggtgcagatccct
cggacctgccccggcaactacgtgctccggcacgagatcatgctgca

Amino acid:

ICHKSATPGGGHATVAAGDKISLVWTPPEWPEHIGHPVIDYMAACNGDCETVNKESLRWFKI
DGAGYDSSKGQWAADALRENGNSWL VQIPSDLAPGNYVLRHEIML

Talaromyces byssochlamydoides

DNA:

caccatgggcccgggtcatcaactacctggccaactgtggcgacagctgcgagacggtcgacaagaccacctcaagttcttaagatcgacggc
gtcggctcttgggacgacaccactgtgccgggcacctggggcgccgaccagctgatcagcaacaacaacagctggctggtcgagatcccccc
gacctgcgccggaact

Amino acid:

HHGPVINYLANCGDSCETVDKTTLKFFKIDGVGLVDDTTVPGTWGADQLISNNNSWLVEIP
PTLRRN

Talaromyces leycettanus

DNA:

ccatcacgggcccgggtcatcagctacctggccaactgtggcgccagctgtgagacagtcgacaaaacgacgttgcaattcttaagatcgacaac
atcggtttcatagatgactctccccccaggcatctgggcagccgatcaattggaagcaacaacaacacctggctcgtggagatccccggac
catcgtcca

Amino acid:

HHGPVISYLANCGASCETVDKTTLQFFKIDNIGFIDDSSPPGIWAADQLEANNTWLVEIPR
PSL

Talaromyces emersonii

DNA:

catcacgggcccgggtcatcacgtacctggccaactgcaacggcaactgctcgaccgttgacaagacgcagctggagttcttaagatcgaccag
tcgggcctgatcaacgacctgacccgccgggcacctgggcgtcgacaacctcatcgccaacaacaacagctggaccgtgacctcccca
gcaccctcgagccgggcaactacgtgctgcgccacgagacatc

Amino acid:

HHGPVITYLANCNGNCSTVDKTQLEFFKIDQSGLINDTDPPGTWASDNLIANNNSWTVTIPS
TLEPGNYVLRHET

Talaromyces thermophilus

DNA:

gacgacattattgtcacaaggaggccaagccggcccggccatgccaccgtgaaagccggcgacaagatctacctggacgacattattgt
cacaagaatgccaacctgctcccaacaagcccaaatcaagcgggcgacaaggtgctgctcagtgctcctcatggcccggcctcctga
ccaccaggggcctatcatcaactacctagcgagctgcaacggacctgctcaagtgtgcagaaggagagcttaagtgggcaagattgacga
gactgggctcttccgaacggaacgtgggcaacggacgtgctccgctccaatggaacacatgggatgtgaagatccatcagacctgctcct

Amino acid:

DDIICHKEAKPARGHATVKAGDKIYLDDIICHKNAKPAPNKAQIQAGDKVRLEWSPWPGPP
DHQGPIINYLASNGPCSSVQKESLKWAKIDETGLFPNGTWATDVLRSNGNTWDVKIPSDL
L

Thermoascus aurantiacus

DNA:

tgccacaaggggtgctgagcctggtgccctgagcgccaaggtggctgctggaggaccgtcgagctgcagtgaggcggattggcctgagagtc
acaagggcccggcattgactacctgccgctgtaacgggaactgctcgactgtcgacaagaccaaactagagttctcaagattgatgagag
tggcctaattgacggcagcagcgcgccagggcacatgggcctctgacaacttgattgccaataacaacagctggaccgtcaccatcccagacac
gatgctaccga

Amino acid:

CHKGAEPGALSAKVAAGGTVELQWTDWPESHKGPVIDYLAACNGNCSTVDKTKLEFFKI
DEGLIDGSSAPGTWASDNLIANNNSWTVTIPSTMLP

Malbranchea cinnamomea

DNA:

cacggtgactgctcgtccgtcgacaagacctcgctcaagttctcaagatcagcgaggccggcctaaacgacggctccaacccccggcca
gtgggctccgacgatctcattgccaacaacaacagctggactgtgaccatccccaaagtcgatcggccgggcaactacgtgctgcccacga
gatcatcgccctgcacagcggcggcaaccagaatggcgcccaaaactaccgtccagaagacattgg

Amino acid:

HGDCSSVDKTSLKFFKISEAGLNDGSNAPGQWASDDLIANNNSWTVTIPKSIAPGNYVLRH
EIIALHSAGNQNGAQNYPSRRHW

Remersonia thermophila

DNA:

gacattattgtcaacagtcggccaccaacgccggcgggtacgccgtggtggcggccggcgacaaggctcatcaccagtgaggaccagtggcc
cgagagccaccacggccccgtcatcgactacctgccagctcgccgagcagcgggctcgacgccgtcaacaaggccgacctcagttgttc
aagatcggcgaggtcggcctgatcgacggccggcagggccccggcttctggggctccgaccagctcatcgccaacaacgccggctggctcg
tccagatccccctccgacctcgcgtgatcatctctcgagct

Amino acid:

DIICQQSATNAGGYAVVAAGDKVYIQWDQWPESHGHPVIDYLASCGSTGCDAVNKADLE
LFKIGEVLIDGRQAPGFWGSDQLIANNAGWLQIPSDLASIISRA

Scytalidium indonesiacum

DNA:

tgtcataaccaggccaccaacgcggcgccacgccgtggctcggcgccggcgacaagatctggatccagtgaggaccaatggcccagagac
caccacggccccgtcctcgactacctcgctcctgcggcagctcgggctgcgagtcggtaacaagctcgacctcaagttctcaagatggc
gaaaagggcctgatcgacggctcctccgcggcgccgggtggcgctggacgagctgatcgtaacaacgctggctggctggccagatcc
ccgccgacatcgcgctc

Amino acid:

CHNQATNAGGHAVVAAGDKIWIQWDQWPESHHPVLDYLASCGSSGCESVNKLDLKFFK
IGEKGGLIDGSSAPGRWASDELIANNAGWLVQIPADIAP

Scytalidium thermophilum

DNA:

ctggagttttcaagatcgatgccgctggcttcgaggacggcaagtgggcttctgacaagctcatcgccaacaacaacacctggaccgtgaccg
tccccgacagcattgcccttgtaatacgtcctccgtcacgagatcatgctctccac

Amino acid:

LEFFKIDAAGFEDGKWASDKLIANNNTWTVTVPDSIAPGQYVLRHEIMLS

Thermomyces lanuginosus

DNA:

gacattatctgtcacaaggaagccacaccagccccgggtcatgtctccgtgaaggccgggtgacaagatctacatccaatggcagccgaatccat
ggccggattcccaccacggtgccccgtcctggactatctggccccttgaacgggcccctgtgagtcctcgacaagaccagcttgcgcttctt
caagatcgacggagtggtcttatcgacggctcttctcctccgggcaagtgggcccagacgaactccatgcgaacggcaacgggtggctgg
ttcagatccccgaggacatcaagccg

Amino acid:

DIICHKEATPARGHVSVKAGDKIYIQWQPNPWPDSHHG
GPVLDYLAPCNGPCESVDKTSRFFKIDGVGLIDGSSPPGKWADDEL
HANGNGWLVQIPEDIKP

Fig S1: Flow scheme for Peptide Pattern Recognition.

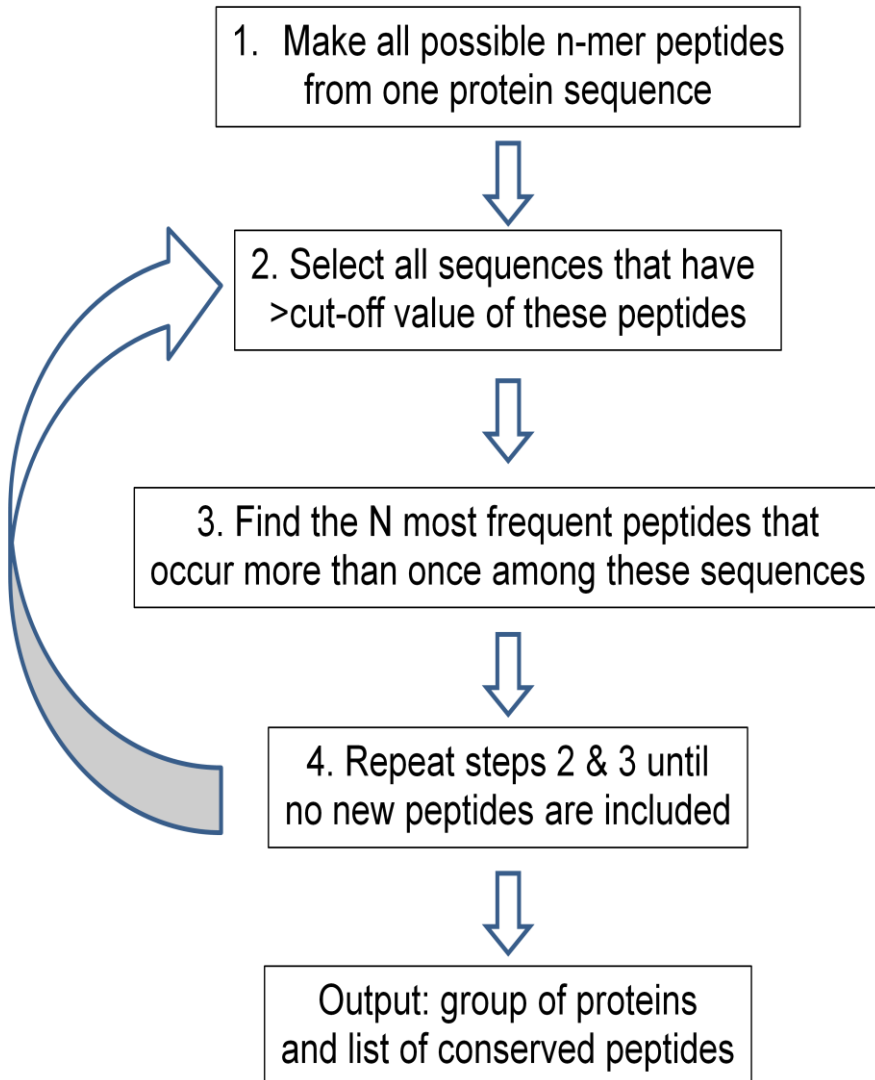




Fig S2: Phylogenetic tree of 118 eukaryotic GH5 proteins made with MUSCLE, PhyML and TreeDyn (phylogeny.fr). Proteins with different EC numbers are labeled in different colors as indicated.

% correct predictions							
limit\pep length	2	3	4	5	6	8	10
10	44	74	96	97	98	99	100
20	46	77	96	96	98	99	100
30	47	83	93	94	96	98	100
50	48	87	92	96	96	97	100
70	49	87	93	96	97	99	100
100	55	84	89	96	98	99	100
150	69	85	90	98	98	98	100
200	66	84	88	97	99	99	100
400	39	87	91	97	98	98	100
800		88	91	97	99	98	100
1600		87	92	98	98	98	100
3200		91	92	97	98	99	100

% of proteins predicted							
limit\pep length	2	3	4	5	6	8	10
10	67	92	67	48	40	34	26
20	83	96	69	61	48	41	31
30	89	97	73	60	53	44	34
50	94	98	77	66	58	43	37
70	96	98	86	73	58	51	38
100	98	98	87	75	62	52	43
150	98	98	92	79	69	54	40
200	98	98	94	79	73	53	45
400	98	98	95	85	77	63	46
800		98	97	90	80	63	47
1600		98	98	91	87	73	51
3200		98	98	92	86	74	57

correct predictions * proteins predicted							
limit\pep length	2	3	4	5	6	8	10
10	29	68	64	46	39	33	26
20	38	73	66	58	47	40	31
30	41	80	67	56	50	43	34
50	45	85	70	63	55	41	37
70	47	85	79	70	56	50	38
100	53	82	77	72	60	51	43
150	67	83	82	77	67	52	40
200	64	82	82	76	72	52	45
400	38	85	86	82	75	61	46
800		86	88	87	79	61	47
1600		85	90	89	85	71	51
3200		89	90	89	84	73	57

Fig S3: Use of function-specific peptide lists for prediction of the function of GH5 proteins.

For each enzyme class of characterized GH5 proteins half of the proteins were chosen at random and used to make Peptide Patterns with the indicated parameters. These peptide patterns were used to predict the function of the rest of the characterized GH5 proteins and the percent correct prediction (EC 3.2.1.4, 3.2.1.58, 3.2.1.75 or 3.2.1.78) was measured with the function reported in CAZy as reference. The peptide patterns consisted of 10 – 3200 peptides with a length of 2 – 10 amino acids. No peptide cut-off was used. No data were calculated for 800 – 3200 dimer peptides as there are only 400 possible dimers. Each number represents the average correct prediction rate of ten repetitions.

Table S1 Sequence of primers for PCR of group 1.

Peptide	Primer name	Primer sequence ^a
DIICHK	61.1	CTGGACGAYATHATHHTGYCAYAA
EIHALH	61.2	CTGGACTGIAGIGCDATDATYTC
HHGPV	61.3	CTGGACCAYCAYGGICCIGT
GAQNYP	61.4	CTGGACGGRTARTTYTGIGCICC
LEFFKI	61.5	CTGGACCTIGARTTYTTYAARAT

^aI = Inosine, Y = C and T, H = A, C and T, D = A, G and T, R = A and G.

Table S2 Distribution of 540 functionally characterized enzymes in the GH13 subfamilies.

Subfamily	Assigned function	Other functions in the subfamily
1	10 EC5.4.99.16	9 EC3.2.1.10; 7 EC3.2.1.20; 3 EC3.2.1.70; 1 EC2.4.1.-
2	16 EC3.2.1.68	
3	18 EC2.4.1.18	1 EC3.2.1.1
4	67 EC3.2.1.1	
5	17 EC3.2.1.54	10 EC3.2.1.41; 10 EC3.2.1.133; 4 EC3.2.1.135; 2 EC3.2.1.1; 2 EC3.2.1.-; 1 EC3.2.1.33
6	14 EC3.2.1.141	
7	12 EC5.4.99.15	
8	3 EC3.2.1.20	1 EC3.2.1.28; 1 EC3.2.1.10
9	16 EC3.2.1.41	1 EC3.2.1.68
10	none described	
11	29 EC3.2.1.1	2 EC3.2.1.60; 2 EC3.2.1.-; 1 EC3.2.1.98; 1 EC3.2.1.116
12	25 EC2.4.1.18	
13	18 EC3.2.1.1	
14	4 EC3.2.1.98	2 EC3.2.1.1; 1 EC3.2.1.60; 1 EC3.2.1.-
15	none described	
16	4 EC2.4.1.4	2 EC3.2.1.-
17	31 EC3.2.1.1	2 EC2.4.1.25; 1 EC3.2.1.2; 1 EC3.2.1.116; 1 EC3.2.1.-
18	8 EC3.2.1.41	
19	5 EC2.4.1.7	
20	1 EC3.2.1.54	
21	25 EC2.4.1.19	3 EC2.4.1.-; 2 EC3.2.1.1; 1 EC3.2.1.133
22	5 EC3.2.1.93	
23	3 EC3.2.1.54	1 EC3.2.1.135
24	16 EC3.2.1.1	2 EC3.2.1.116
25	13 EC3.2.1.20	10 EC5.4.99.11
26	10 EC3.2.1.20	2 EC3.2.1.10
27	none described	
28	1 EC5.4.99.16	
29	8 EC2.4.1.183	
30	2 EC3.2.1.33	
31	5 EC3.2.1.1	1 EC2.4.1.19
32	15 EC3.2.1.1	
33	1 EC3.2.1.1	
34	none described	
35	none described	
36	3 EC3.2.1.1	
37	5 EC2.4.1.7	
38	10 EC3.2.1.1	
39	10 EC3.2.1.1	2 EC3.2.1.54
40	none described	
41	1 EC5.4.99.16	1 EC3.2.1.54; 1 EC3.2.1.1
42	1 EC3.2.1.1	
43	none described	
44	6 EC3.2.1.1	1 EC3.2.1.-
45	2 EC3.2.1.54	
46	1 EC3.2.1.133	1 EC3.2.1.1
47	none described	
48	none described	
49	1 EC2.4.1.19	
50	none described	
Total	442 enzymes	98 enzymes

Table S3 Accession numbers of proteins in the 16 GH61 subfamilies

Subfamily	Accession numbers ^a
1	AEO64177.1, AEO58921.1, XP_002478590.1, EAQ85056.1, XP_958583.1, ACH92573.1, EGO54705.1, ADB89217.1, EGE79257.1, XP_002626386.1, ADJ57703.1, EEQ84968.1, CAA71999.1, XP_003050495.1, CCD50144.1, XP_001553240.1, XP_001213388.1, CBI54781.1, EGU75413.1, XP_002146302.1, EGO61608.1, EAA29018.1, XP_003043569.1, EGS19579.1, CAP92380.1, XP_868906.1, XP_001216541.1, XP_001210534.1, EFX04021.1, CAF31975.1, XP_001392640.1, XP_752040.1, XP_001597642.1, XP_001264507.1, NP_001130665.1, XP_658645.1, XP_001906078.1, ABW56451.1, XP_388088.1, CCD48549.1, XP_386286.1, XP_748707.1, XP_001259147.1, CBF73467.1, XP_003044062.1, EAA59545.1, CBI56951.1, EDO00288.1, EDP47167.1, EFQ36324.1, XP_001395530.1, XP_001267517.1, CCD50139.1, XP_002559170.1, XP_382378.1, XP_001400893.1, XP_001269010.1, XP_001404656.1, XP_001267236.1, XP_001400893.2, XP_003043105.1, XP_001402213.1, EGU85487.1, XP_001225249.1, XP_001402213.2, XP_002558168.1, BAE64395.1, BAE56764.1, AEO56542.1, XP_002381515.1, XP_001821292.1, ACF86151.1, EGU80176.1, EEY23940.1, EFQ35170.1, CAP73254.1, XP_002380189.1, XP_001934008.1, EAA59072.1, XP_003303895.1, XP_001214594.1, CBF76006.1, EFQ28963.1, EAA59125.1, CAQ16278.1, EGP90215.1, XP_003002008.1, BAE60320.1, XP_003007405.1, EAQ83639.1, XP_001230192.1, XP_001584877.1, XP_001208821.1, XP_383871.1, EFQ34935.1, XP_001271679.1, EGS20667.1, EGU85118.1, XP_001273555.1, EFQ31130.1, EGU84205.1, EFQ30363.1, XP_003298498.1, XP_001941054.1, XP_001792546.1, XP_384949.1, XP_001791868.1, XP_001546774.1, XP_003002930.1, EDO00181.1
2	EFQ34021.1, EAQ86198.1, AEO59955.1, CAD21296.1, EGO61029.1, EAA53409.1, XP_367775.2, CBI52679.1, XP_959499.1, XP_003352294.1, EGO55727.1, XP_003003091.1, AAT64005.1, AEO68023.1, CCA67656.1, AEO56642.1, CCA67657.1, EFQ27332.1, CCA67658.1, XP_001937216.1, CCA67659.1, CAP73072.1, CCA72182.1, XP_001790826.1, EGS17558.1, EAA36362.1, XP_001589530.1, XP_003306203.1, EFQ34537.1, CCA72183.1, EAU86100.1, CCD49290.1, XP_003351982.1, XP_002391345.1, XP_001560175.1, EGO53330.1, XP_001841073.2, CAG27577.1, JGI_PC15_v2.0_16, XP_003029697.1, JGI_AB_v2.0_6, AEO56416.1, XP_001834453.2, CCA72192.1, CAP72740.1, CBY00196.1, CAP61476.1, EGS23404.1, XP_001225324.1, JGI_PC15_v2.0_10, ADX07320.1, XP_001225412.1, CCA70035.1, AEO56016.1, CCA70703.1, EAQ92448.1, CCA72220.1, CBI57737.1, CCA72221.1, XP_001226408.1, XP_001936652.1, XP_003300926.1, EGS18626.1, AEO62422.1, XP_001904712.1
3	AEO59836.1, EAT92310.2, AEO68157.1, XP_001227732.1, CCA74246.1, XP_001904958.1, EGS24042.1, EGS19451.1, XP_003351237.1, AEO60271.1, ACE10234.1, XP_003305915.1, CAP65866.1, CBX94532.1, EDU50404.1, CAG27576.1, CAD70347.1, XP_001906810.1, EGO56889.1, EGO55619.1, AEO61304.1, CBI56992.1, EAA58450.1, XP_001219583.1, CAF05857.1, CBX94572.1, EAQ87022.1, XP_001791556.1, EGS22278.1, EDK04335.1, CCA76320.1, EFQ25679.1, XP_002392372.1, CBI52566.1, EAU33838.1, AEO58169.1, XP_003302575.1, EFI92673.1, CCA71764.1, EFQ34070.1, XP_003300728.1, JGI_PC15_v2.0_13, EAT77790.2, XP_001935377.1, XP_003002719.1, EFQ27394.1, EEY22380.1, CCD51504.1, XP_003034283.1, XP_362483.1
4	AEO61257.1, EGO58845.1, XP_962414.1, CAP71839.1, XP_003349473.1, EFQ31502.1, EEU39531.1, AEO66274.1, EGU77464.1, XP_003296866.1, XP_003006393.1, XP_001223312.1, XP_001932318.1, EDN93954.1, XP_003004661.1, EFQ35966.1, CAQ16208.1, CCD50451.1, XP_001546354.1,

	JGI_PC15_v2.0_9, CBY01974.1, XP_001793524.1, CCA73151.1, JGI_PC15_v2.0_11, EFJ00004.1, AEJ35168.1, XP_001907502.1, JGI_PC15_v2.0_5, JGI_AB_v2.0_3, EFJ03083.1, CBX91313.1
5	XP_001222797.1, EGO53245.1, AEO67395.1, CBI58321.1, CAE81966.1, EAA26873.1, CBI54870.1, EGO58503.1, XP_001791546.1, XP_001229931.1, ACE10231.1, CAP67190.1, EFQ34588.1, EGO03895.1, XP_003000747.1, JGI_PC15_v2.0_7, CCA70418.1, XP_001933274.1, EGS20384.1, CAJ81215.1, CAP68375.1, XP_001408450.1, EFI92604.1, XP_003026941.1, XP_001560894.1, AEO55652.1, XP_001230041.1, CAP68309.1, EAA57051.1, EFQ36570.1, XP_001832511.1
6	EGG03285.1, XP_388187.1, AEO54509.1, XP_003352682.1, EGS23697.1, XP_001222154.1, CAJ81218.1, ACE10235.1, XP_001907069.1, AEO64605.1, XP_963702.1, EGO52064.1, CAP61650.1, EFQ34071.1, CBI54738.1, XP_003299155.1, XP_001932393.1, XP_003002463.1, EAA50788.1, CCD47228.1, AEO59482.1, EGS21609.1, XP_001558279.1, CBY02242.1, XP_001226548.1, AEO55082.1, XP_001591031.1, CAP71532.1, XP_001223687.1, EAT89418.1
7	EER27647.1, XP_002544405.1, XP_001242606.1, XP_003015133.1, XP_002848955.1, EGD92942.1, XP_003019976.1, XP_003235451.1, EGE04098.1, XP_003174002.1, XP_002382772.1, BAE61530.1, XP_001822663.2, EEH05325.1, EGC46978.1, XP_002628516.1, XP_001543425.1, EEQ89326.1, EGE82146.1, XP_001266231.1, XP_386263.1, EGU85465.1, EGR49412.1, XP_003348222.1, EFZ04469.1, XP_003050482.1, EFY91792.1, CBF74037.1, XP_364682.1
8	XP_003352382.1, , EGO53412.1, CAB97283.2, EAA36150.2, AEO65580.1, ACE10232.1, XP_001410346.1, XP_001522073.1, EDU48860.1, XP_003304335.1, XP_001806210.1, CBX96476.1, EAQ89525.1, CAP73311.1, EGS21628.1, XP_001838602.1, ACE10233.1, JGI_PC15_v2.0_15, EFI94694.1, XP_001220412.1, AEO56665.1, CCA74449.1, XP_003029596.1
9	XP_001820645.2, BAE58643.1, XP_001260907.1, EDP54951.1, XP_001208400.1, EAW14268.1, XP_369395.1, EAA64499.1, EFX04819.1, XP_001590753.1, CCD52645.1, XP_386384.1, XP_001552214.1, XP_001939155.1, XP_003302923.1, EGU84223.1, EGR49821.1, EAT83900.1, EFY92676.1, CBX93965.1, AEO55776.1, EGS21285.1, EFQ30762.1, XP_003045940.1
10	JGI_PC15_v2.0_8, XP_001834768.1, JGI_PC15_v2.0_14, XP_001883194.1, JGI_PC15_v2.0_4, JGI_PC15_v2.0_12, EDR15540.1, EAU87138.1, JGI_AB_v2.0_5, XP_003026965.1, XP_002474815.1, EGN94280.1, JGI_AB_v2.0_4, EED85774.1, EDR07117.1, XP_002389005.1, EFI92060.1, EGN94279.1, EGO19770.1, GWW2.10.277.1, FGENESH1_PG.C_SCAFFOLD_8000142, XP_001802947.1, XP_001228820.1
11	XP_001826694.1, EAL88805.1, EAA56258.1, EAW16173.1, EAU34150.1, XP_369715.2, EFQ36423.1, CAP59702.1, XP_001903620.1, EGP91468.1, XP_362640.1, EFQ33676.1, EAT81480.1, EAQ85642.1, XP_003305323.1, XP_001223598.1, XP_003303817.1, XP_001939876.1
12	XP_003001513.1, EAL93749.1, CAK41095.1, XP_001395335.2, CAP86439.1, XP_002376462.1, XP_001820737.1, EAW13459.1, XP_001260913.1, XP_660650.1, XP_001274726.1, EFQ25762.1, CAQ16206.1, EFQ35295.1
13	EDP48657.1, XP_747482.1, XP_001216411.1, XP_001262213.1, BAE55582.1, XP_001817584.2, XP_001397023.2, CAK42466.1, BAB62318.1, A1C4H2.1, Q5BCX8.1, CBX98254.1
14	XP_001835420.2, EAU80304.1, XP_001835421.1, CAG27578.1, XP_002392498.1, XP_003031503.1, JGI_AB_v2.0_1, CCA68244.1, XP_002398849.1, XP_001885903.1, XP_001839611.1, XP_002399040.1
15	XP_001841005.1, EGO19398.1, EAU80738.1, EGN93952.1, EAU80891.2, XP_003027567.1, JGI_AB_v2.0_2, XP_001840824.2, EAU89754.1, XP_002910838.1

16 EGO58321.1, EAA29132.1, CBI58198.1, AEO58412.1, AEO68763.1, EAA53298.1, XP_367664.2, CAP68352.1, EGS20367.1, XP_001229971.1

^aAccession numbers are from Genbank except JGI_AB_v2.0_1 - JGI_AB_v2.0_9 from Joint Genome Institute *Agaricus bisporus* genome project and JGI_PC15_v2.0_1 - JGI_PC15_v2.0_16 from Joint Genome Institute *Pleurotus ostreatus* genome project. GWW2.10.277.1 and FGENESH1_PG.C_SCAFFOLD_8000142 were found in the *Phanerochaete chrysosporium* genome (Genbank accession number AADS00000000.1)

Table S4 Sequence and frequency of the conserved peptides for the 16 GH61 subfamilies

Subfamily	Peptides (sequence, frequency)
1	(YVLRHE,0.81) (GAQNYP,0.79) (RHEIIA,0.75) (EIIALH,0.75) (HEIIAL,0.75) (QNYPQC,0.74) (AQNYPQ,0.74) (VLRHEI,0.7) (LRHEII,0.68) (NYVLRH,0.55) (GNYVLR,0.55) (IIALHS,0.49) (LEFFKI,0.49) (IALHSA,0.48) (YPQCFN,0.46) (NGAQNY,0.45) (SHHGVP,0.41) (NYPQCF,0.41) (DIICFK,0.35) (PQCFNL,0.35) (ALHSAG,0.34) (PGNYVL,0.32) (EFFKID,0.31) (HHGPVI,0.31) (PDIICH,0.31) (GPVIDY,0.3) (TDPGIL,0.3) (APGNYV,0.3) (ESHHGP,0.29) (PESHHG,0.29) (WPESHH,0.29) (YPQCIN,0.29) (AAHGHV,0.29) (VAAHGH,0.28) (IPGPAL,0.27) (LANCNG,0.27) (DGAQNY,0.27) (TWPESH,0.26) (VIDYLA,0.26) (PVIDYL,0.26) (CETVDK,0.26) (YLANCN,0.25) (FFKIDG,0.25) (IALHGA,0.25) (LHSAGQ,0.23) (ETVDKT,0.23) (ILVNIY,0.23) (GILVNI,0.23) (SHKGPV,0.22) (PQCINL,0.22) (VTGSGT,0.22) (TGSGTA,0.21) (ANCNGD,0.21) (HKGPVI,0.21) (VDKTTL,0.21) (LIANNN,0.21) (NCNGDC,0.21) (DIICHR,0.21) (IAPGNY,0.2) (DPGILV,0.2) (SLEFFK,0.2)
2	(SFPGAY,0.71) (VSFPGA,0.69) (YMECAQ,0.68) (TVSFPG,0.63) (MECAQI,0.54) (GQYLLR,0.51) (FYMECA,0.48) (QFYMEC,0.48) (GAQFYM,0.46) (AQFYME,0.46) (FPGAYS,0.42) (CAQINV,0.42) (APGQYL,0.38) (SAHTIF,0.38) (ECAQIN,0.38) (YLLRAE,0.37) (QYLLRA,0.35) (YLAKVD,0.34) (YTIPGP,0.34) (ASHKGP,0.34) (PGQYLL,0.32) (KVDNAA,0.32) (ASAHTI,0.32) (IAPGQY,0.31) (ELIALH,0.31) (CIAPGQ,0.31) (AKVDNA,0.31) (AYLKKV,0.31) (LAKVDN,0.29) (STVSFP,0.29) (LLRAEM,0.29) (YPGAQF,0.28) (LYGQCG,0.28) (YGQCGG,0.28) (SHKGPI,0.28) (PGAQFY,0.28) (SHKGPT,0.26) (GWFKIQ,0.26) (IALHAA,0.25) (QCGGQG,0.25) (PITDVT,0.25) (GPITDV,0.25) (SYPGAQ,0.25) (GPTTCA,0.25) (GGWFK,0.23) (TGPTTC,0.23) (GGWFKI,0.23) (ITDVTS,0.23) (VDNAAT,0.22) (EMIALH,0.22) (VYLAKV,0.22) (QLYMEC,0.22) (AEMIAL,0.22) (AYLAKV,0.22) (FPGAYK,0.22) (GLKWFK,0.22) (GAQLYM,0.22) (AQLYME,0.22) (LRAEMI,0.22) (CGGQGW,0.22) (DVTSND,0.22) (RAEMIA,0.22) (YLLRVE,0.22) (GQCGGQ,0.2) (LYMECA,0.2) (CAQINI,0.2)
3	(GGAQFY,0.68) (EHIALH,0.58) (SFPGAY,0.54) (VSFPGA,0.52) (YISCAQ,0.48) (AQFYIS,0.46) (HIALHS,0.46) (QFYISC,0.46) (TDPGIL,0.44) (PGAYKA,0.44) (GAQFYI,0.44) (ATDPGI,0.44) (FPGAYK,0.44) (IALHSA,0.44) (FYISCA,0.44) (PGILIN,0.4) (PVPTS Y,0.36) (GILINI,0.36) (YLLRAE,0.36) (DPGILI,0.36) (GAVWFK,0.34) (AHYTFP,0.34) (YPVPTS,0.34) (KATDPG,0.32) (YKATDP,0.32) (YYPVPT,0.32) (GAYKAT,0.32) (AYKATD,0.32) (NIYYPV,0.3) (IYYPVP,0.3) (VWFKIY,0.28) (LLRAEH,0.28) (GDYLLR,0.26) (FPGAYS,0.26) (ALHSAS,0.26)

(ISCAQI,0.26) (LVSFPG,0.26) (LHSASS,0.26) (LRAEHI,0.26) (LRVEHI,0.24)
(VEHIAL,0.24) (GEYLLR,0.24) (VAFPGA,0.24) (RVEHIA,0.24) (AFPGAY,0.24)
(INIYYP,0.24) (PGAYSA,0.22) (GQTAAT,0.22) (YTPPGP,0.22) (LLRVEH,0.22)
(GPVTDV,0.22) (AVWFKI,0.22) (VPTSyt,0.22) (GAQFYL,0.22) (ISCAQL,0.22)
(ILINIY,0.22) (YLLRVE,0.22) (VTGGGS,0.2) (SYTPPG,0.2) (GHPGPL,0.2)
(DWQYVR,0.2) (IGHPPG,0.2)

4 (HEIIAL,0.68) (EIIALH,0.68) (RHEIIA,0.68) (YLVRHE,0.65) (GAQFYP,0.61)
(PGAQFY,0.58) (YPGAQF,0.58) (LVRHEI,0.55) (GPVITY,0.52) (VRHEII,0.52)
(PGNGPV,0.52) (IPGNGP,0.48) (IPGPAV,0.48) (PVITYM,0.45) (WTLWPD,0.45)
(VITYMA,0.45) (ITYMAR,0.45) (GNGPVE,0.42) (FYPGCH,0.42) (VSFPGA,0.42)
(LWPDSH,0.42) (DSHVGP,0.42) (PDSHVG,0.42) (WPDSHV,0.42) (NGPVED,0.42)
(SFPGAY,0.42) (TYMARC,0.42) (YYLVRH,0.42) (GYLVR,0.42) (PGCHQL,0.42)
(TLWPDS,0.42) (YMARCP,0.42) (YPGCHQ,0.42) (AQFYPG,0.39)
(AAHGYV,0.39) (FPGAYK,0.39) (QFYPGC,0.39) (IALHAA,0.39) (MARCPD,0.39)
(SHVGPV,0.35) (AHGYVD,0.35) (SYPGAQ,0.35) (LVSFPG,0.32) (AVWFKV,0.32)
(HVGPVI,0.32) (YQDPYM,0.32) (IIALHA,0.32) (YTIPGP,0.32) (AAGSTV,0.32)
(VAAHGY,0.32) (VGPVIT,0.32) (PYQDPY,0.32) (PGPAVF,0.29) (GPVEDV,0.29)
(YTIPSC,0.29) (VTGSGT,0.29) (ITYMAK,0.29) (TYMAKC,0.29) (QVTGSG,0.29)
(HGYVDN,0.29) (LVAAHG,0.29) (YPIPGP,0.29)

5 (YLLRAE,0.77) (VTVEMH,0.68) (TVEMHQ,0.68) (GDYLLR,0.61)
(DYLLRA,0.61) (VEMHQQ,0.58) (GGAQFY,0.55) (GAQFYM,0.55)
(SNSPVT,0.55) (AIGGAH,0.55) (EAIGGA,0.55) (NEAIGG,0.52) (TVTTEM,0.45)
(EMHQQP,0.45) (AQFYMS,0.42) (QFYMSC,0.42) (TCKVGS,0.39)
(KCPVKA,0.39) (GGAHYG,0.39) (GAHYGP,0.39) (YMSCYQ,0.39)
(IGGAHY,0.39) (CPVKAG,0.39) (AHYGPV,0.39) (LLRAEV,0.35) (AVSPPY,0.35)
(KFPGAY,0.35) (PYYSQC,0.35) (VKFPGA,0.35) (FYMSCY,0.35) (SPPYYS,0.35)
(YMTCYQ,0.35) (APGPAV,0.35) (FPGAYK,0.35) (AHATFQ,0.35) (VSPPYY,0.35)
(PPYYSQ,0.35) (GWFKIF,0.35) (ALHTAG,0.32) (EVIALH,0.32) (LRAEVI,0.32)
(PGPAVY,0.32) (PGAYSA,0.32) (AGGTVT,0.32) (SGDYLL,0.32) (AIGGNH,0.32)
(GYTGCT,0.32) (TVKFPG,0.32) (FQQLWV,0.29) (IALHTA,0.29) (MHQQPG,0.29)
(ADGSSG,0.29) (RAEVIA,0.29) (STCKVG,0.29) (YGQCGG,0.29) (NSPVTD,0.29)
(VYMSKV,0.29) (EAIGGN,0.29) (PGDRSC,0.29) (QPGDRS,0.29) (GGTVTV,0.29)
(GCESTC,0.29) (LLRAEA,0.29) (GTVTVE,0.29) (AEVIAL,0.29)

6 (VYHQGP,0.87) (TDPGYT,0.8) (YLLRIQ,0.7) (LRCNVG,0.67) (HYIFQQ,0.63)
(AHYIFQ,0.6) (GEYLLR,0.6) (AVYHQG,0.6) (DLRCNV,0.6) (QFYISC,0.6)
(IPGAFK,0.6) (YISCAQ,0.6) (PQFYIS,0.6) (FYISCA,0.6) (GYTANI,0.57)
(YTANIY,0.57) (YTVPGP,0.53) (SNDLRC,0.53) (EYLLRI,0.53) (SAHYIF,0.53)
(PGYTAN,0.53) (NDLRCN,0.53) (DPGYTA,0.53) (NYNSPV,0.53) (YMSKAP,0.5)
(YNSPVT,0.47) (SPVTDL,0.43) (NSPVTD,0.43) (NIYSNF,0.43) (RCNVGG,0.43)
(MSKAPG,0.4) (DYDGSG,0.4) (LLRIQS,0.37) (WFKIKD,0.37) (IHNPWP,0.37)
(VKAGDS,0.37) (GIPQFY,0.37) (PGAFKE,0.33) (KETDPG,0.33) (SIPGAF,0.33)
(GIHNPW,0.33) (YHQGPI,0.33) (LRIQSL,0.33) (HNPWPA,0.33) (ANIYSN,0.33)
(FKETDP,0.33) (AFKETD,0.33) (ETDPGY,0.33) (TNYNSP,0.33) (NYTVPG,0.33)
(LGIHNP,0.33) (DTAVYH,0.33) (LLRIQQ,0.33) (TANIYS,0.33) (LAIHNP,0.33)
(HQGPIS,0.33) (GAFKET,0.33) (NPWPAG,0.33) (IYSNFN,0.3) (IPQFYI,0.3)
(VSIPGA,0.3) (ISCAQI,0.3) (ISCAQV,0.3) (YIFQQL,0.3) (ASAHYI,0.3)

7 (GGGSCQ,0.97) (HGGGSC,0.97) (YMNCAP,0.93) (VHGGGS,0.86)
(AVHGGG,0.83) (GSAVHG,0.76) (SAVHGG,0.76) (SIEGGC,0.72) (GGSCQV,0.72)
(RIGNRE,0.72) (QRCASG,0.72) (NRIGNR,0.72) (IEGGCP,0.69) (YTLAWT,0.69)
(GSCQVS,0.69) (TLAWTW,0.69) (MNCAPI,0.66) (KVIHSI,0.62) (GNREMY,0.59)
(REMYMN,0.59) (NREMYM,0.59) (AWTWFN,0.59) (LAWTWF,0.59)
(MYMNCA,0.59) (EMYMNC,0.59) (CASGQW,0.55) (IAPGEY,0.55)
(VSLTTD,0.55) (SCQVSL,0.55) (CQVSLT,0.55) (SVWKVI,0.55) (APGEYT,0.55)
(SIAPGE,0.55) (RCASGQ,0.55) (QVSLTT,0.55) (WNCIGG,0.52) (EGGCPA,0.52)
(CPANVD,0.48) (GGCPAN,0.48) (FQRCAS,0.48) (GCPANV,0.48) (ANVDGN,0.48)
(PANVDG,0.48) (HRRRGH,0.45) (IGNREM,0.45) (VDGNLP,0.45)
(NVDGNL,0.45) (ASGQWT,0.41) (FYMNCA,0.41) (EFYMNC,0.41)
(REFYMN,0.41) (MFVANI,0.38) (VWKVIH,0.38) (IGNREF,0.38) (NREFYM,0.38)
(EYTLAW,0.38) (TWFNRI,0.38) (HSIEGG,0.38) (IHSIEG,0.38) (GEYTLA,0.38)
(IGSAVH,0.38) (FNRIGN,0.38) (WTFNRI,0.38) (WKVIHS,0.38) (VIHSIE,0.38)
(GNREFY,0.38) (WFNRIG,0.38) (FIGSAV,0.38)

8 (GAVTSY,0.79) (HGAVTS,0.79) (RHELLA,0.75) (APGNYL,0.71) (AHGAVT,0.71)
(ALHQAN,0.71) (ELLALH,0.67) (LALHQA,0.67) (GWFKID,0.67) (HELLAL,0.67)
(LLALHQ,0.67) (WFKIDE,0.63) (LIRHEL,0.63) (YLIRHE,0.63) (HQANTP,0.58)
(ANTPQF,0.58) (NTPQFY,0.58) (IRHELL,0.58) (QANTPQ,0.58) (LAPGNY,0.58)
(GNYLIR,0.54) (CNGGTS,0.54) (AAHGAV,0.54) (NYLIRH,0.54) (YAECAQ,0.54)

(FYAECA,0.54) (QFYAEC,0.54) (NGGTSA,0.54) (PQFYAE,0.54) (PGYEGF,0.5)
(SPASSP,0.5) (TYPGYE,0.5) (GFSPAS,0.5) (FSPASS,0.5) (YPGYEG,0.5) (LHQANT,0.5)
(WKQWTH,0.5) (EGFSPA,0.46) (PGNYLI,0.46) (YEGFSP,0.46) (GYEGFS,0.46)
(QFYPEC,0.46) (SPKTIQ,0.46) (TPQFYA,0.46) (VTSYII,0.46) (FYPECA,0.46)
(RCNGGT,0.46) (PQFYPE,0.46) (AVTSYI,0.42) (PASSPK,0.42) (SSPKTI,0.42)
(ASSPKT,0.42) (GDGKGW,0.38) (KGWFKI,0.38) (QGPVMV,0.38) (DGKGWF,0.38)
(THEQGP,0.38) (YPECAQ,0.38) (HEQGPV,0.38) (WTHEQG,0.38) (GKGWFK,0.38)
(WPDYNP,0.38) (QWPDYN,0.38) (VWMYKC,0.38) (QWTHEQ,0.38)

9 (GPCAVY,0.75) (HKGPCA,0.67) (KGPCAV,0.67) (QGDGTC,0.58) (GDGTCV,0.54)
(GDGWFK,0.54) (DGTCVR,0.54) (GWFKIW,0.54) (AGDGWF,0.5) (AAGDGW,0.5)
(CAVYMK,0.46) (YVGCAQ,0.46) (ELLALH,0.46) (PCAVYM,0.46) (GYLVR,0.42)
(NQGDGT,0.42) (LTFEFR,0.42) (SHKGPC,0.42) (DGWFKI,0.42) (NANWCG,0.38)
(DPQFYV,0.38) (GTCVRM,0.38) (YLVRPE,0.38) (QFYVGC,0.38) (CTEKLI,0.38)
(WCTEKL,0.38) (PQFYVG,0.38) (FYVGCA,0.38) (GCAQIF,0.33) (KWCTEK,0.33)
(VYMKKV,0.33) (LALHAA,0.33) (YPIPGP,0.33) (YYLVRP,0.33) (TTLYVD,0.29)
(NAAGDG,0.29) (KNANWC,0.29) (RPEILA,0.29) (RPELLA,0.29) (PELLAL,0.29)
(EILALH,0.29) (AHTVFT,0.29) (TTFTTL,0.29) (MKKVDD,0.29) (TCVRMP,0.29)
(HTTFTT,0.29) (YMKKVD,0.29) (GPPNKG,0.29) (PPNKGK,0.29) (PEILAL,0.29)
(TVSIPG,0.29)

10 (GGAEFY,0.78) (GAEFYP,0.65) (EFYPSC,0.57) (VSFPGA,0.57) (SFPGAY,0.57)
(AEFYPS,0.57) (HEIAL,0.52) (RHEIA,0.52) (EIALH,0.52) (LVSFPG,0.52)
(APGNYL,0.52) (FYPSCT,0.43) (IALHL,0.43) (YLIRHE,0.43) (YPSCTQ,0.43)
(IRHEI,0.43) (LIRHEI,0.39) (ELVSFP,0.39) (AKWFKI,0.39) (IALHLA,0.39)
(GNYLIR,0.3) (TYMASC,0.3) (PGAYSD,0.3) (LAPGNY,0.3) (ASCGST,0.3)
(WPHNTG,0.3) (GPMLTY,0.3) (YMASC,0.3) (PHNTGP,0.3) (NYLIRH,0.26)
(WFKIDQ,0.26) (YVFPGP,0.26) (NLAPGN,0.26) (NIAPGN,0.26) (YLLRHE,0.26)
(IAPGNY,0.26) (HNTGPM,0.26) (FPGAYS,0.26) (PGNYLI,0.26) (VFPGBP,0.26)
(TGPMLT,0.26) (FPGAYK,0.22) (NTGPML,0.22) (TFPGGY,0.22) (EGGAEF,0.22)
(MLTYLA,0.22) (PSCTQI,0.22) (KWFKID,0.22) (PMLTYL,0.22) (GGSGTG,0.22)
(TYLASC,0.22) (YLASC,0.22)

11 (LVRVEH,0.94) (YLVRVE,0.94) (VRVEHI,0.89) (GDWFKV,0.61) (YDGSGD,0.56)
(MSKAPG,0.56) (DGSGDW,0.56) (VKIPGL,0.5) (GSGDWF,0.5) (RVEHIA,0.5)
(SKAPGD,0.5) (SGDWFK,0.5) (CAQIEV,0.5) (KIPGLY,0.5) (PGPLVK,0.44)
(YFTCAQ,0.44) (FYFTCA,0.44) (FTCAQI,0.44) (YMSKAP,0.44) (TCAQIE,0.44)
(LVKIPG,0.44) (GPLVKI,0.44) (EFYFTC,0.44) (KAPGDV,0.44) (PLVKIP,0.44)
(EHIAIH,0.39) (GQYLVR,0.39) (QFYMEC,0.39) (QYLVRV,0.39) (GEYLVR,0.39)

	(RVEHIG,0.39) (PSFTSE,0.33) (SYDGSG,0.33) (IEVTGS,0.33) (VEHIGL,0.33) (QIEVTG,0.33) (EVTGSG,0.33) (EYLVRV,0.33) (EHIGLH,0.33) (AQFYME,0.33) (DGDWFK,0.33) (WYSQCV,0.33) (AQIEVT,0.33) (VEHIAI,0.33) (GDGDWF,0.33) (YMPSFT,0.33)
12	(SAHYFF,1.0) (AHYFFD,1.0) (FRCNKG,0.86) (YMSKAP,0.86) (YNPTKW,0.86) (DFRCNK,0.86) (KYNPTK,0.86) (EHIGVH,0.86) (YPMPGP,0.86) (DGEYLI,0.79) (GEYLIR,0.79) (VKAGSK,0.79) (VYMSKA,0.79) (LVYMSK,0.79) (VRSNTR,0.79) (EVKAGS,0.79) (PSFNFS,0.71) (PDGEYL,0.71) (VSAHYF,0.71) (DDMTPD,0.71) (TAEVKA,0.71) (IGVHGA,0.71) (RSNTRP,0.71) (AKYNPT,0.71) (HIGVHG,0.71) (DPSFNF,0.71) (AEVKAG,0.71) (RDDMTP,0.71) (GVHGAH,0.64) (GDGDWF,0.64) (GQAEFY,0.64) (GDWFKI,0.64) (AGSKLA,0.64) (GTAEVK,0.64) (KDYPMP,0.64) (CAQVKV,0.64) (TGTAEV,0.64) (KAGSKL,0.64) (DYPMPG,0.64) (DGDWFK,0.64) (TRPAKY,0.57) (RPAKYN,0.57) (GPALVY,0.57) (YLIRPE,0.57) (RPEHIG,0.57) (PGPAVW,0.57) (PALVYM,0.57) (IRPEHI,0.57) (GTPGPT,0.57) (HPGPAL,0.57) (TFAGQT,0.57) (LIRPEH,0.57) (SNTRPA,0.57) (PAKYNP,0.57) (FTFAGQ,0.57) (MQHPGP,0.57) (ALVYMS,0.57) (NTRPAK,0.57) (PGPALV,0.57) (EYLIRP,0.57)
13	(RPEIIA,1.0) (PEIIAL,1.0) (ALHEGN,1.0) (EIIALH,1.0) (IALHEG,1.0) (IIALHE,1.0) (LHEGNR,1.0) (DMTCNV,0.92) (VYMAPT,0.92) (YMECVQ,0.92) (TFEWHH,0.92) (LALAAA,0.92) (SDDIIA,0.92) (IIASSH,0.83) (DDIIAS,0.83) (GVSIPG,0.83) (GAQFYM,0.83) (AQFYME,0.83) (DIIASS,0.83) (MTCNVN,0.83) (QFYMEC,0.83) (FYMECV,0.83) (WVKIAE,0.83) (PDVPAG,0.75) (YIRSPP,0.75) (YPIPGP,0.75) (LLALAA,0.75) (SYPIPG,0.75) (GYIRSP,0.75) (GGAQFY,0.75) (TCNVNG,0.75) (VPDVPA,0.75) (PGPAVW,0.67) (GILFDI,0.67) (ILFDIY,0.67) (FRPEII,0.67) (TVPDVP,0.67) (YLFIRPE,0.67) (LFRPEI,0.67) (WDGSSS,0.58) (ECVQVK,0.58) (MVYMAP,0.58) (FDIYNS,0.58) (DGSSSG,0.58) (DIYNSF,0.58) (GPAVWD,0.58) (SFDSYP,0.58) (GSSSGS,0.58) (DSDDII,0.58) (VWDGSS,0.58) (VQVKVT,0.58) (GWVKIA,0.58) (QVKVTS,0.58) (PAVWDG,0.58) (YMAPTE,0.58) (LAAAAK,0.58) (SDSDDI,0.58) (ADGYIR,0.58) (VKAGDK,0.58) (RSDSDD,0.58) (MECVQV,0.58) (ALAAAA,0.58) (CVQVKV,0.58) (AVWIND,0.58) (ITVPDV,0.58) (SSSGSS,0.58)
14	(PGLTVN,0.67) (DPGLTV,0.67) (KVSIPG,0.67) (GLTVNI,0.67) (GLHVAG,0.67) (ISCAQI,0.58) (AAIYLG,0.58) (VSIPGY,0.58) (PAAIYL,0.58) (IGLHVA,0.58) (SIPGYV,0.58) (GPAAIY,0.58) (LHVAGA,0.58) (YTVPGP,0.58) (SWDGSG,0.58) (HVAGAP,0.58) (IPGYVS,0.58) (VAGAPQ,0.58) (YISCAQ,0.58) (SYTVPG,0.5) (ASWDGS,0.5) (PTSYTV,0.5) (LTVNIY,0.5) (WFKIAE,0.5) (FKIAEW,0.5) (KIAEWG,0.5) (WGATFN,0.42) (AIYLGQ,0.42) (IAEWGA,0.42) (HQPAA,0.42) (YNPVPT,0.42) (QIGLHV,0.42) (EQIGLH,0.42) (GSGANW,0.42) (RQPLNN,0.42) (NPPKVS,0.42)

	(SAHYIW,0.42) (GATFNP,0.42) (PPKVSI,0.42) (IYNPVP,0.42) (WDGSGA,0.42) (TSYTV,0.42) (SGANWF,0.42) (NIYNPV,0.42) (GNPPKV,0.42) (ATFNPF,0.42) (EWGATF,0.42) (DGSGAN,0.42) (PKVSIP,0.42) (GANWFK,0.42) (YHQGPA,0.42)
15	(SCAQIN,0.8) (LVSIPG,0.7) (PLVSIP,0.7) (PGPAVW,0.7) (CAQINV,0.7) (SIPGVY,0.7) (VSIPGV,0.7) (WFKVHE,0.6) (GTPGPL,0.6) (GPLVSI,0.6) (TPGPLV,0.6) (PGPLVS,0.6) (VWFKVH,0.6) (GNGTPG,0.6) (NGTPGP,0.6) (LSCAQI,0.6) (GGNGTP,0.5) (QPGPAV,0.5) (PSGQYL,0.5) (LPSGQY,0.5) (SLPSGQ,0.5) (GQVWFK,0.5) (VRRTDN,0.5) (QYVRR,0.4) (PIPANY,0.4) (GVYNGR,0.4) (PGVYNG,0.4) (IPGVYN,0.4) (FKVHEI,0.4) (NGREPG,0.4) (YNGREP,0.4) (QFFLSC,0.4) (YVRRTD,0.4) (VYNGRE,0.4) (GGAQFY,0.4) (VTDVTS,0.4) (YLSCAQ,0.4) (AQINVQ,0.4) (PVTDTV,0.4) (WQYVRR,0.4) (GILINI,0.4) (EPGILI,0.4) (PGILIN,0.4)
16	(PDIICH,0.9) (GAQNYP,0.9) (QWNGWP,0.9) (ALAHSH,0.7) (RWATDV,0.6) (GTPVVV,0.6) (DIICHR,0.6) (AHAPVR,0.6) (IQWNGW,0.6) (GPPGRW,0.5) (WATDVL,0.5) (PYVLRH,0.5) (APVRAG,0.5) (VLSYLA,0.5) (GGAQNY,0.5) (DDGFVT,0.5) (GPVLSY,0.5) (HELIAL,0.5) (GPYVLR,0.5) (PPGRWA,0.5) (ATDVLI,0.5) (YVLRHE,0.5) (ELIALH,0.5) (HAPVRA,0.5) (RHELIA,0.5) (PVLSYL,0.5) (DGFVTP,0.5) (TPVVVT,0.4) (SQALAH,0.4) (VQWNGW,0.4) (CVNLWV,0.4) (TDVLIA,0.4) (IALHYA,0.4) (VNLWVE,0.4) (QALAH,0.4) (GCAGVD,0.4) (PVRAGD,0.4) (GFDPRP,0.4) (AQNYPL,0.4) (IICHRD,0.4) (VDDGFV,0.4) (LAHSHI,0.4) (SYLAPC,0.4) (WTKIDD,0.4) (QNYPLC,0.4) (NNSWQV,0.4) (ICHRDG,0.4) (LSYLAP,0.4) (CAGVDK,0.4) (AVDDGF,0.4) (TKIDDS,0.4) (IALHFA,0.4) (PGPYVL,0.4) (HIQWNG,0.4)

Table S5: Number of conserved peptides from each GH61 subfamily and sum of their frequencies in GH61E (Genbank accession number XP_003657366).

Subfamily	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Peptides	1	2	26	2	3	3	0	0	0	1	0	0	1	4	3	0
Σ Frequencies	0.3	0.79	9.12	0.61	1.64	1.8	0	0	0	0.26	0	0	0.75	2	1.2	0

Table S6: Classification of the enzymatically characterized GH61 into PPR subfamilies.

Accession	Species	Type	Ref ^a	Subfamily	Freq ^b	Peps ^c	P-value
CAP73254.1	<i>Podospora anserina</i>	nd ^d	6	1	8.3	18	1.3×10 ⁻⁸⁰
CAP68375.1	<i>Podospora anserina</i>	nd ^d	6	5	7.7	18	1.3×10 ⁻⁸⁰
NCU02916	<i>Neurospora crassa</i>	nd ^d	24	2	8.0	22	1.9×10 ⁻⁹⁹
NCU01867	<i>Neurospora crassa</i>	nd ^d	24	5	15.8	39	1.8×10 ⁻¹⁸⁹
NCU03328	<i>Neurospora crassa</i>	nd ^d	24	3	8.1	22	1.4×10 ⁻¹⁰⁴
NCU08760	<i>Neurospora crassa</i>	1	31	5	15.6	39	1.4×10 ⁻¹⁸⁹
NCU01050	<i>Neurospora crassa</i>	2	31	2	11.9	32	6.8×10 ⁻¹⁵⁷
NCU07898	<i>Neurospora crassa</i>	2	31	4	20.7	48	2.7×10 ⁻²⁵⁰
ABW56451.1	<i>Thermoascus aurantiacus</i>	1	15	1	12.4	25	9.3×10 ⁻¹²⁰
ACE10231.1	<i>Thielavia terrestris</i>	nd ^d	15	5	12.7	32	1.9×10 ⁻¹⁵²
ACE10234.1	<i>Thielavia terrestris</i>	nd ^d	15	3	9.1	26	8.9×10 ⁻¹¹⁸
BAL43430.1	<i>Phanerochaete chrysosporium</i>	nd ^d	39	13	11.5	21	1.9×10 ⁻¹⁰³

^aReference, ^bSum of frequencies of peptides, ^cnumber of peptides, ^dnot determined