

Glycosyl Hydrolase Genes and Enzymes of *Neurospora crassa*

Alan Radford

Faculty of Biological Sciences, The University of Leeds, Leeds LS2 9JT, U K;

Fungal Genetics Newsletter 53:12-14

An analysis of the genome of *Neurospora crassa* has identified genes encoding 84 putative glycosyl hydrolases, representing 24 different families in Henrissat's classification. Functionally, they include enzymes for the degradation of all major polysaccharides (including cellulase, hemicellulase, chitinase and pectinase). There is evidence of high levels of gene amplification, despite the presence of RIP, as there are eight representatives of family 3 (beta-glucosidases and xylosidases), five of family 7 (endo- and exo-glucanases), six of family 13 (amylases and maltase), nine of family 18 (chitinase), eight of family 47 (ER alpha-mannosidases), eleven of family 61 (endoglucanases) and seven of family 76 (alpha-mannanases).

In *The Mycota III: Biochemistry and Molecular Biology*, 2nd edn. (ed. Brambl & Marzluf), p 243, I included a preliminary list of the glycosyl hydrolases of *Neurospora crassa* identified in assembly 3 of the *Neurospora* genome (Radford, 2004). This listed putative NCU identities by glycosyl hydrolase family (Henrissat and Romeu 1995), and the different enzyme activities found in each family; *n.b.* in some families, different members may have different enzyme activities. These data have been reassessed in genome assembly 7, and only those confirmed by current annotation, Pfam glycosyl hydrolase family predictions and close Blast hits to homologs with identified enzyme activities are included in the table below.

Glycosyl hydrolases are what might be termed "modular" enzymes. They are normally under carbon catabolite regulation, and typically have one or more copies of the C-regulation motif (consensus 5'-G/CPyGGGG-3') upstream of the ORF. The table includes the number of such motifs identified for each gene. Some glycosyl hydrolases also have a substrate-binding domain, located at either the N- or the C-terminus of the mature enzyme. Only the unique glucoamylase has an identifiable starch-binding domain (SBD: PF00686). Three chitinases have a chitin-recognition region. Some cellulases and xylanases have what is commonly called a fungal cellulose-binding domain (CBD: PF00734). Finally, most glycosyl hydrolases are targeted for secretion, to the cell wall, etc., and the presence of signal peptides has been identified by the SignalP v 3.0 web server (Bendtsen *et al.* 2004).

Because of the level of confidence in the identification of the glycosyl hydrolases, gene symbols have been assigned, as shown in the final column of Table 1. Some of these genes, *cbh-1*, *gla-1*, *gla-2*, *inv* and *tre-2* have already been described and named (Perkins *et al.* 2001). I have chosen to base the nomenclature of the new genes on the glycosyl hydrolase family number rather than the specific enzyme activity, *gh1* (glycosyl hydrolase family 1), *gh2* (glycosyl hydrolase family 2), etc. as the symbols and names. Where there is more than one gene in a family, hyphenated numbering is used, e.g. *gh3-1*, *gh3-2*, *gh3-3*.

Table 1 The glycosyl hydrolases of *Neurospora crassa*

<i>GH family</i>	<i>MIPS predicted function</i>	<i>C-regul motif</i>	<i>signal sequence</i>	<i>binding domain</i>	<i>identity</i>	<i>gene</i>	<i>EC</i>
1	beta-glucosidase	3	no		ncu00130	<i>gh1-1</i>	3.2.1.21
2	beta-mannosidase	3	no		ncu00890	<i>gh2-1</i>	3.2.1.25
2	beta-galactosidase	2	no		ncu05956	<i>gh2-2</i>	3.2.1.23
2	beta-galactosidase	4	no		ncu00810	<i>gh2-3</i>	3.2.1.23
3	beta-glucosidase	3	yes		ncu03641	<i>gh3-1</i>	3.2.1.21
3	beta-glucosidase	3	no		ncu08054	<i>gh3-2</i>	3.2.1.21
3	beta-glucosidase	0	yes		ncu08755	<i>gh3-3</i>	3.2.1.21
3	beta-glucosidase	2	yes		ncu04952	<i>gh3-4</i>	3.2.1.21
3	beta-glucosidase	2	no		ncu05577	<i>gh3-5</i>	3.2.1.21
3	beta-glucosidase	2	no		ncu07487	<i>gh3-6</i>	3.2.1.21
3	xylan 1,4-beta-xylosidase	3	no		ncu09923	<i>gh3-7</i>	3.2.1.37
3	xylan 1,4-beta-xylosidase	2	yes		ncu00709	<i>gh3-8</i>	3.2.1.37
5	endoglucanase II	2	yes	CBD	ncu00762	<i>gh5-1</i>	3.2.1.4
6	cellobiohydrolase II	1	yes		ncu03996	<i>gh6-1</i>	3.2.1.91
6	cellobiohydrolase II	1	yes	CBD	ncu09680	<i>gh6-2</i>	3.2.1.91
6	cellobiohydrolase II	3	yes		ncu07190	<i>gh6-3</i>	3.2.1.91
7	endoglucanase I	2	yes		ncu05057	<i>gh7-1</i>	3.2.1.4
7	cellobiohydrolase I	1	yes		ncu04854	<i>gh7-2</i>	3.2.1.91

7	endoglucanase I	2	yes		ncu04027	gh7-3	B.2.1.4
7	cellobiohydrolase I	2	yes		ncu05104	gh7-4	B.2.1.91
7	cellobiohydrolase I	3	yes	CBD	ncu07340	cbh-1	B.2.1.91
10	endo-1,4-beta-xylanase	2	yes		ncu05924	gh10-1	B.2.1.8
10	endo-1,4-beta-xylanase	2	yes		ncu08189	gh10-2	B.2.1.8
10	endo-1,4-beta-xylanase	0	yes	CBD	ncu04997	gh10-3	B.2.1.8
10	endo-1,4-beta-xylanase	3	yes		ncu07130	gh10-4	B.2.1.8
11	endo-1,4-beta-xylanase	3	yes		ncu02855	gh11-1	B.2.1.8
11	endo-1,4-beta-xylanase	4	yes	CBD	ncu07225	gh11-2	B.2.1.8
13	alpha-amylase	1	yes		ncu08131	gh13-1	B.2.1.1
13	alpha-amylase	3	yes		ncu09805	gh13-2	B.2.1.1
13	alpha-amylase	2	no		ncu09486	gh13-3	B.2.1.1
13	maltase	1	no		ncu06523	gh13-4	B.2.1.20
13	maltase	1	no		ncu07860	gh13-5	B.2.1.20
13	alpha-amylase	4	no		ncu05873	gh13-6	B.2.1.1
15	glucoamylase	0	yes	SBD	ncu01517	gla-1	B.2.1.3
18	chitinase	2	yes	CBD	ncu04500	gh18-1	B.2.1.14
18	chitinase	0	yes		ncu03209	gh18-2	B.2.1.14
18	chitinase	2	no		ncu03026	gh18-3	B.2.1.14
18	chitinase	0	yes		ncu04883	gh18-4	B.2.1.14
18	chitinase	1	no		ncu04554	gh18-5	B.2.1.14
18	chitinase	3	yes	ChBD	ncu05317	gh18-6	B.2.1.14
18	chitinase	4	yes		ncu06020	gh18-7	B.2.1.14
18	chitinase	0	yes	ChBD	ncu07484	gh18-8	B.2.1.14
18	chitinase	1	yes	ChBD	ncu07035	gh18-9	B.2.1.14
28	polygalacturonase II	1	yes		ncu02369	gh28-1	B.2.1.15
28	exopolygalacturonase	9	yes		ncu06961	gh28-2	B.2.1.67
30	glucosylceramidase	1	yes		ncu04395	gh30-1	B.2.1.45
31	alpha-glucosidase	1	yes		ncu02583	gla-2	B.2.1.20
31	alpha-glucosidase	0	yes		ncu09281	gh31-1	B.2.1.20
31	alpha-glucosidase	1	yes		ncu04203	gh31-2	B.2.1.20
31	alpha-glucosidase	2	yes		ncu04674	gh31-3	B.2.1.20
32	invertase	2	yes		ncu04265	inv	B.2.1.26
35	beta-galactosidase	0	yes		ncu00642	gh35-1	B.2.1.23
35	beta-galactosidase	1	yes		ncu04623	gh35-2	B.2.1.23
37	alpha, alpha-trehalase	5	no		ncu04221	tre-2	B.2.1.28
38	alpha-mannosidase	1	no		ncu07404	gh38-1	B.2.1.24
45	endoglucanase V	2	yes	CBD	ncu05121	gh45-1	B.2.1.4
47	ER alpha-mannosidase	1	no		ncu02778	gh47-1	B.2.1.24
47	ER alpha-mannosidase	0	yes		ncu03134	gh47-2	B.2.1.24
47	ER alpha-mannosidase	1	yes		ncu01059	gh47-3	B.2.1.24
47	ER alpha-mannosidase	2	yes		ncu05836	gh47-4	B.2.1.24
47	ER alpha-mannosidase	5	yes		ncu07067	gh47-5	B.2.1.24
47	ER alpha-mannosidase	1	yes		ncu02235	gh47-6	B.2.1.24
47	ER alpha-mannosidase	3	yes		ncu02091	gh47-7	B.2.1.24
47	ER alpha-mannosidase	1	yes		ncu09028	gh47-8	B.2.1.24
61	endoglucanase IV	1	yes	CBD	ncu02240	gh61-1	B.2.1.4
61	endoglucanase IV	1	yes	CBD	ncu07760	gh61-2	B.2.1.4
61	endoglucanase IV	5	yes	CBD	ncu02916	gh61-3	B.2.1.4
61	endoglucanase IV	5	yes		ncu01050	gh61-4	B.2.1.4
61	endoglucanase IV	4	yes	CBD	ncu08760	gh61-5	B.2.1.4
61	endoglucanase IV	2	yes		ncu03328	gh61-6	B.2.1.4
61	endoglucanase IV	5	yes	CBD	ncu00836	gh61-7	B.2.1.4

61	endoglucanase IV	1	no		ncu03000	<i>gh61-8</i>	β.2.1.4
61	endoglucanase IV	2	yes		ncu05969	<i>gh61-9</i>	β.2.1.4
61	endoglucanase IV	1	yes	CBD	ncu01867	<i>gh61-10</i>	β.2.1.4
61	endoglucanase IV	0	yes		ncu07520	<i>gh61-11</i>	β.2.1.4
63	ER mannosyl- oligosaccharide glucosidase I	2	yes		ncu03657	<i>gh63-1</i>	β.2.1.106
76	alpha-1,6-mannanase	1	yes		ncu02032	<i>gh76-1</i>	β.2.1.101
76	alpha-1,6-mannanase	2	yes		ncu04262	<i>gh76-2</i>	β.2.1.101
76	alpha-1,6-mannanase	1	yes		ncu08127	<i>gh76-3</i>	β.2.1.101
76	alpha-1,6-mannanase	3	no		ncu06319	<i>gh76-4</i>	β.2.1.101
76	alpha-1,6-mannanase	3	yes		ncu09937	<i>gh76-5</i>	β.2.1.101
76	alpha-1,6-mannanase	1	yes		ncu02216	<i>gh76-6</i>	β.2.1.101
76	alpha-1,6-mannanase	1	yes		ncu03770	<i>gh76-7</i>	β.2.1.101
81	beta-1,3-exoglucanase	2	no		ncu07076	<i>gh81-1</i>	β.2.1.58

References

Bendtsen, J. D., H. Nielsen, G. Von Heijne and S. Brunak, 2004. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340: 783-795.

Henrissat, B., and A. Romeu, 1995. Families, superfamilies and subfamilies of glycosyl hydrolases. *Biochem. J.* 311: 350-1351.

Perkins, D. D., A. Radford and M. S. Sachs, 2001. *The Neurospora Compendium: Chromosomal loci.* Academic Press, San Diego, CA.

Radford, A., 2004. Polysaccharidases in *The Mycota: a comprehensive treatise on fungi as experimental systems for basic and applied research*, edited by R. Brambl and G. A. Marzluf. Springer.