

## **Characterization of PdCP1, a serine carboxypeptidase from *Pseudogymnoascus destructans*, the causal agent of white-nose syndrome**

Chapman Beekman, Zhenze Jiang, Brian M. Suzuki, Jonathan M. Palmer, Daniel L. Lindner, Anthony J. O'Donoghue, Giselle M. Knudsen and Richard J. Bennett

### **Supplementary material**

**Supplementary Tables 1-5.** Proteins identified by LC-MS/MS in *P. destructans* (Table S1), *P. verrucosus* (S2), 03VT05 (S3), 05NY08 (S4), and WSF3629 (S5). N=3 replicates per species. MS/MS-related metrics from Protein Prospector software are included in table columns (for each replicate in order from left to right: the number of uniquely mapped peptides, the number of observed spectra for a given protein identified—peptide count, the percent sequence coverage, the best individual peptide score, the best discriminant score, and the best expectation value for an individual peptide in each secretome analysis). The accession number of each sequence within the corresponding proteome database is also provided (first column) as well as the protein predicted molecular weight and name (last two columns).

**Supplementary Tables 6-10.** Proteins identified by LC-MS/MS with functional annotations provided by PANNZER software (Koskinen et al. 2015) (webtool) for *P. destructans* (Table S6), *P. verrucosus* (S7), 03VT05 (S8), 05NY08 (S9), and WSF3629 (S10). Annotation statistics and descriptions are included in table columns (from left to right: Gene ID, cluster\_GSZ, cluster\_RM1sum, cluster\_size, cluster\_desccount, RM2, val\_avg, jac\_avg, description, genename). See PANNZER publication for more information on associated statistical analyses (Koskinen et al. 2015).

CLUSTAL O(1.2.4) multiple sequence alignment

```

MER0002010      MKAFTSLLCGLGLSTTLAKAISLQRPLGLD---KDVLLQAAEKFG---LDLDDLHLLKEL 54
MER0375986      -----0
MER0120077      -----0
MER0350236      -----0
PdCP1           -MRFPSIFTAVLFAA---SSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPF 55
MER0944165      -----0

MER0002010      DS-----NVLDA-----WA---QIEHLYPNQVMSLE 77
MER0375986      -----MHFAVIVITLLVAS---SAVA---LENPHRKAVRPIKLD-HGHLKPRAVTVAD 46
MER0120077      -----MRGLNSLALLL---AAS-ATSVVAI---LPHHEAHAKYAKRSLPKQILARQPVVRS 49
MER0350236      -----MRGLNSLAVLL---AAS-VTSVVAV---LPHHEAHAKYAKRAQPKHVLPRQPALKP 49
PdCP1           SNSTNNGLLFINTTIIASIAAKEEGVSLEKRPAQNPHKEAKAKYVQKRSAPV--SKHASRA 114
MER0944165      -MQTVRALLLL-----L-----AASVVTAAQNPHDKAAKYVQKRSAPV--SQHASRA 46
                                     :

MER0002010      TSTKPKFPKPAIKTKKDWFVVKNDAIENYQLRVNLIKDKPKILGIDPNVTQYTGYLE--V 134
MER0375986      DG-----GYKYLNKQTRFLVNGTGIPEVDFDV-----GESYAGTLPNTPA 87
MER0120077      RS-----TSKYLTNSTASFAVNGTALPEVDFDI-----GESYAGTLPVSTN 90
MER0350236      RS-----ISKYLTNSTASFAVNGTALPEIDFDI-----GESYAGTLPISTN 90
PdCP1           AN-----TAHFLNNKTAKFAINGSALPEVDFNI-----GESYAGTLPISSK 155
MER0944165      AD-----TAHFLNNKTAKFAVNGSALPEVDFNI-----GESYAGTLPISNK 87
                                     .:. * : . : : : : . : * * *

MER0002010      EDEDKHFFFWTFESRNDPAKDPVILWLNGGPGCSSLTGLFFELGPSSIGP-DLKPIGNPY 193
MER0375986      --GNSSLFFWFFPSQNPKAHDEITIWLNNGGPGCSSLDGLLQENGPFLLWQSGTYKPIRNPY 145
MER0120077      ASDTNRLWFWFFPTDNPAAEKEITIWLNNGGPGCSSLDGFFQENGPLSWQSGTYAPILNPY 150
MER0350236      ASDTNRLWFWFFPSNPLAEKEITIWLNNGGPGCSSLDGLFQENGPFWSQSGTYAPIPNPY 150
PdCP1           KDDENRLWFWFFPSSNPAAKEITIWLNNGGPGCSSLNGLFQENGPFLLWQPGTYAPFANPY 215
MER0944165      HDDKNRLWFWFFPSSNPAAKEITIWLNNGGPGCSSLDGLFQENGPFLLWQPGTYAPFANPY 147
                                     . : * * * : * * . : : * * * * * : * * *

MER0002010      SWNSNATVIFLDQPVNVGFSYSGSSGVSNVAAGKDVYNFLELFFDQFPEYVNKGQDFHI 253
MER0375986      SWTNLTNMVYVDQPVSTGFSPG-PSTVNNNEEDVARQFKSWFKHFVD---FNLHGRKVYI 201
MER0120077      SWTNLTNMIWIDQPVSTGFSPG-DILVDDEIDVGNQFAAFWKNFIDT---FSMQGYKYIYI 206
MER0350236      SWTNLTNMIWIDQPVSTGFSPG-DILVDDEIDVGNQFAAFWKNFIDT---FSMQGYKYIYI 206
PdCP1           SWVNLTNMIYIDQPVSTGFSPG-TVLVDDENDVAEQFMGFWKNFIDT---FSMQGYKYIYI 271
MER0944165      SWVNLTNMIYIDQPVSTGFSPG-DIKVDNEDDVAEQFMGFWKNFIDT---FSMQGYKYIYI 203
                                     ** . : : : : * * . . . * . * . : . . . : : * * : : * . : *

MER0002010      AGEYAGHYIPVFASEILSHKDRNF-NLTSVLIGNGLTDP-----LTQYNYEPMA 303
MER0375986      TGEYAGQYIPYIASAMLDEKDKKYFNVKGIQINDPSINDDSVMIYAPAVSHLNQYLNVF 261
MER0120077      TGEYAGQYIPYIASNFLDRNDTIYYNLKGIQINDPSINEFDTMGSAVPTAAALYYQNIL 266
MER0350236      TGEYAGQYIPYIASNFLDRNDTIYYNLKGIQINDPSINDFDTMGSAVPTAAALYYQNVL 266
PdCP1           TGEYAGQYIPYIASGMLDTKDNKYFNVKGVQINDPSINTDDVLLHTPIVPALNYFN NVI 331
MER0944165      TGEYAGQYIPYIASGMLDTKDEKYFNVKGVQINDPSINTDDVLLHAPVVPALKYNNVI 263
                                     : * * * : * * : * . : * : : * . : : : : : : :

MER0002010      CGEGG-EPS-VLPSEECSAMEDSLERCLGLIESCYD-----SQSVWSCVPATYICNNAQ 355
MER0375986      SLNDTFVKHINKRAEECGYNK-FL----DE-AITYPPPKEFPVAPDPSKNNCALWDDIVE 315
MER0120077      NLNDTYIANITARAKSCGYTE-FL----EYATVFPPAGPIPTAPSSEYGCGLYDDIYN 320
MER0350236      NLNDTYIANITARAKSCGYTD-FL----EYASVFPPAGPIPTAPSSEYGCGLYDDIYN 320
PdCP1           NLNESFIANITARADSCGYTD-FF----NKWTTEFPSSKIPTAPSWKEPGCDIFDEVYN 386
MER0944165      NLNESYIANISARADSCGYTD-FF----NKWTSEFPSSKIPTAPSWKEPGCDIYDDVYN 318

MER0002010      LAPYQRTGRNVYDIRKDCEG--GNLCYPTLQ-DIDDYLNQDYVKEAVGAEV-DHYESC NF 411
MER0375986      AAYYVNPCFNIYHLTDFCPYLWDEMGPFLAGGPNNYFNRSVQKALHVPP-TDYSVCGE 374
MER0120077      AAYYVNPCFNIYHLTDYCPYLWDELGFPSLGGGPNNYFNRTDVKEVIHAPVDTDYFVCTG 380
MER0350236      AVYYVNPCFNIYHLTDFCPFLWDELGFPSLGGGPNNYFNRTDVKEVLHAPVDVDYFVCTG 380

```

```

PdCP1      AAYENPCFNLYHLTDYCPYQWNLLGFPSMAGGPNNYFNRTDVQKAINAPP-TNYMVC GG 445
MER0944165 AAYENPCFNLYHLTDYCPYQWNLLGFPSLAGGPNNYFNRTDVQKAINAPP-TNYMVC GG 377
      . * . * . * . : : * : : . : * : : * : : . * *

MER0002010 DINRNLFLFAGDWMKPY-HTAVTDLLNQDLPILVYAGDK FICNWLGNKAWTDVLPWKYDE 470
MER0375986 -T--TIFAKGDQSVPSALGPLPSVIERTNNVLIGHGWL D YLLFVNGSLATIQNMTWNGAQ 431
MER0120077 GP--NLFPNGDKSIDALGPLPSVIERTNNVIIGHGLL D FLLFANGSLITIQNMTWNGLQ 438
MER0350236 GP--NLFPNGDKSIDALGPLPSVIERTNNVIIGHGLL D FLLFANGSLITIQNMTWNGLQ 438
PdCP1      GH--NLFPNGDKSIESSLGPLPSVIERTNNTIIGHGLL D FLLFANGSLISIQNMTWNGAQ 503
MER0944165 GD--NLFPNGDKSIESSLGPLPSVIERTNNTIIGHGLL D FLLFANGSLISIQNMTWNGAQ 435
      . : : * * : : : : : : : * * : : * . : : * : :

MER0002010 EFASQKV-----RNWTASITDEVAGEVKSYPKHTYLRV 503
MER0375986 GFQHPV--EPLYVPYHYGLAELVTST----APNPYTLNAGAGYLGTAHTERGLTFSTV 484
MER0120077 GFQEPSSSTQNLVVPYHQSLGTILTIANAAIPNTPPQNDVAGAGFQGTWHTERGLTFATV 498
MER0350236 GFQEPSSSTQNFYVPYHQSIGEILSIVNDAIPNSPPQYDTAGAGYQGTWHTERGLTFSTV 498
PdCP1      GFQTSPPFKKQNFYVPYHQTNGEILQYANGI--NTHVFTDTAGGGFQGVTHTERGLTYVTV 561
MER0944165 GFQTSPPFKKQNFYVPYHQTNGEILQYANGI--DTHVFTDTAGGGFQGVTHTERGLTYVTV 493
      * : : : : * : : : :

MER0002010 FNGG M VPFDPENALSMVNEWIHGGFSL----- 532
MER0375986 YMAG H GKENHLNPAVH-----AWSSLPSIGIPA---RPY----- 515
MER0120077 NLAG H EIPQYTPGVGYRQLE-FLLGRIANLSVVGDTYTQT-GNFTGVSAPLKMGRLEY 554
MER0350236 NLAG H EIPQYTPGAGYRHLE-FLLGRIANLSVVGDTYTQT-GNFTGVSAPLKMGRLEY 553
PdCP1      NLAG H MIPQYVPGAAYRQLE-FLLGRVSSLEQRGDFTTGPQGNYTGGTWHHHHHH*-- 615
MER0944165 NLAG H MIPQYVPGAAYRQLE-FLLGRVSSLEQRGDFTTGPQGNYTAGTSRRV----- 544
      . * * :

>MER0002010 - carboxypeptidase Y [S10.001] peptidase unit: 112-532 (active site
residue(s): 257,449,508) (Saccharomyces cerevisiae) (Source: UniProt P00729)

>MER0375986 - carboxypeptidase O [S10.014] peptidase unit: 68-504 (active site
residue(s): 205,410,489) (Aspergillus fumigatus) (Source: EMBL nucleotide XM_726431)

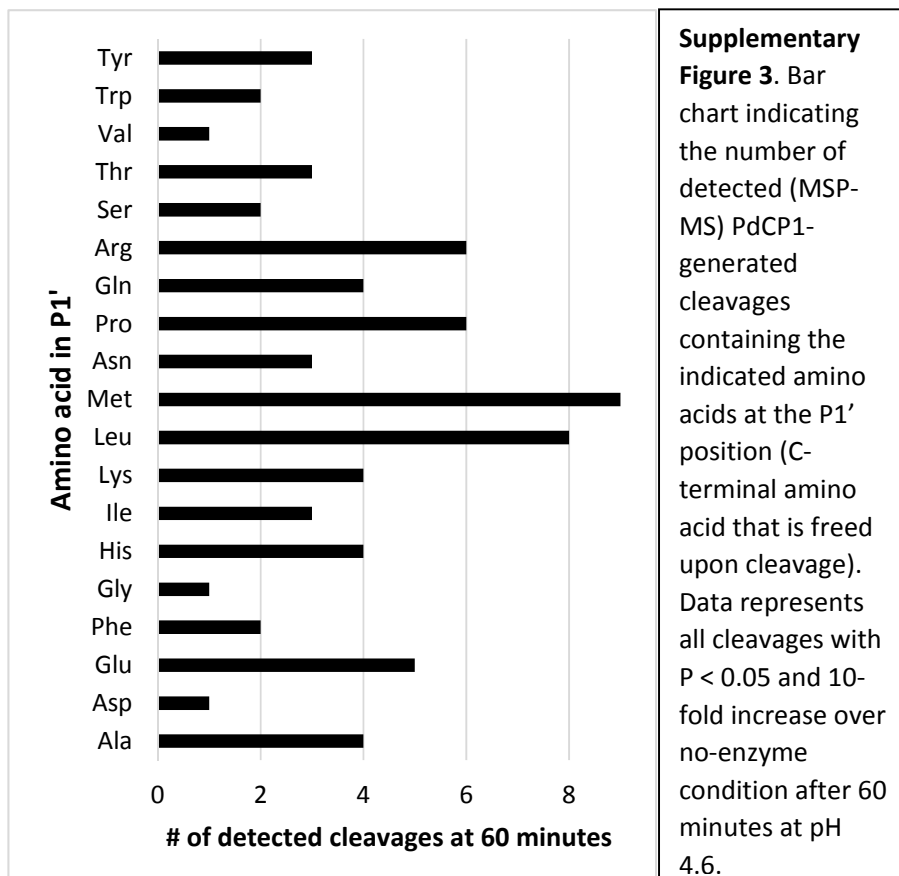
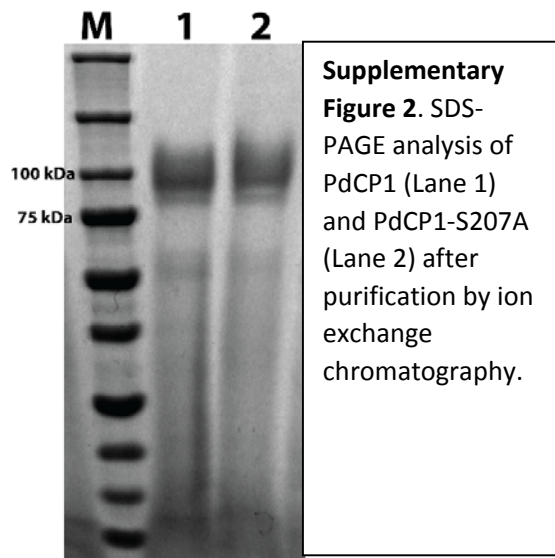
>MER0120077 - carboxypeptidase O [S10.014] peptidase unit: 71-527 (active site
residue(s): 210,417,503) (Sclerotinia sclerotiorum) (Source: EMBL nucleotide
XM_001589703)

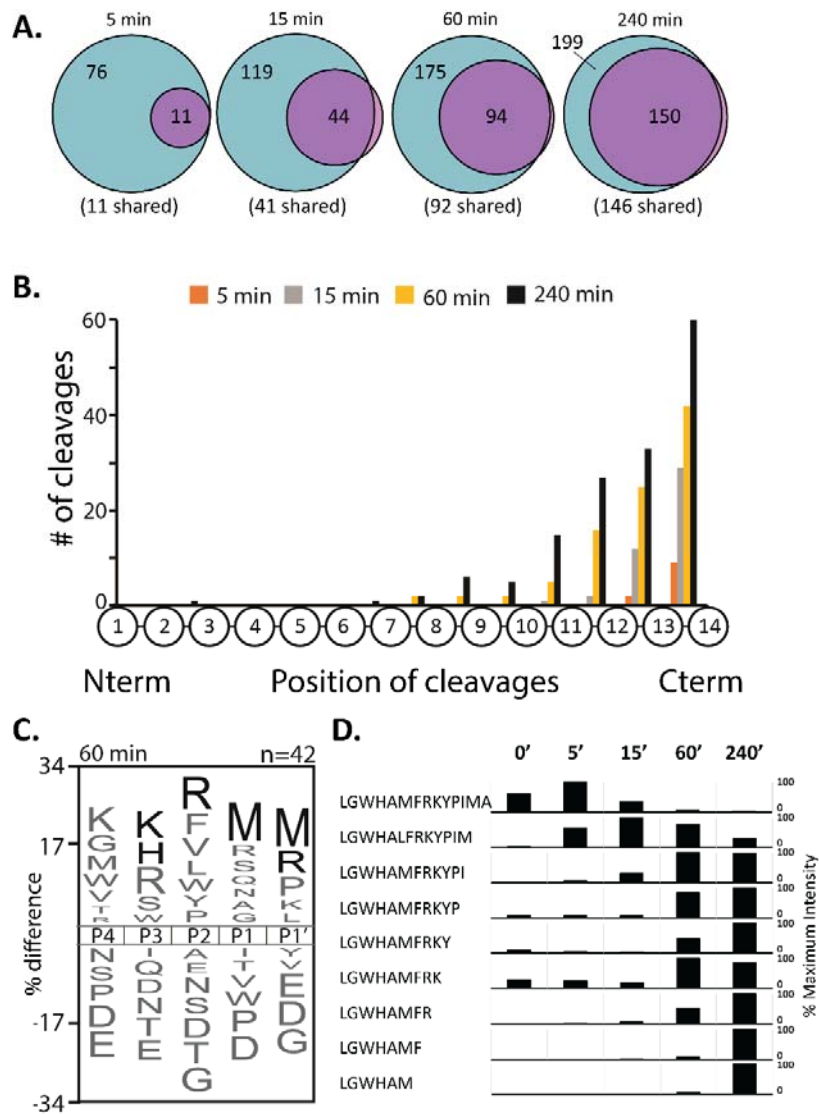
>MER0350236 - carboxypeptidase O [S10.014] peptidase unit: 71-527 (active site
residue(s): 210,417,503) (Botryotinia fuckeliana) (Source: EMBL nucleotide FQ790300)

>MER0944165 - carboxypeptidase O [S10.014] peptidase unit: 68-522 (active site
residue(s): 207,414,498) (Pseudogymnoascus pannorum) (Source: EMBL nucleotide
JPJU01000881)

```

**Supplementary Figure 1.** Amino acid alignment of PdCP1 and related fungal carboxypeptidases generated with Clustal Omega v1.2.4 (Sievers et al. 2011). Highlighted amino acids (in green) indicate the location of catalytic residues (Ser, Asp, His).





**Supplementary Figure 4. MSP-MS analysis of PdCP1-mediated cleavage events at pH 6.2.** **A.** Venn diagrams representing total number of PdCP1-dependent cleavage events detected in MSP library at each time point at pH 4.6 (green) and pH 6.2 (purple). **B.** Histogram indicating total number and position of all PdCP1-generated cleavages detected at each time point. **C.** Substrate specificity signature, based on C-terminal PdCP1-generated cleavages detected after 60 minutes at pH 6.2 using IceLogo (Colaert et al. 2009). Favored residues are indicated above the horizontal line and disfavored residues below. Horizontal positions of residues indicate their distance from the cleaved peptide bond (between P1 and P1'). **D.** bar charts indicating measured intensities of fragments generated by PdCP1 cleavage of 2 representative MSP-MS substrates across each time point.

## References

- Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. *Nat Methods* 6; 786-7.
- Koskinen, P., Toronen, P., Nokso-Koivisto, J., and Holm, L. (2015). PANNZER: high-throughput functional annotation of uncharacterized proteins in an error-prone environment. *Bioinformatics* 31; 1544-52.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J. D., and Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7; 539.