

Additional materials to the article

Dotsenko G.S., Dotsenko A.S. *Mathematical Biology and Bioinformatics*. 2020. V. 15. № 2. P. 429–440

doi: [10.17537/2020.15.429](https://doi.org/10.17537/2020.15.429)

ACCESSION NUMBERS OF HYPOTHETICAL LPMO AMINO ACID SEQUENCES FOUND BY CPRENN AND HMMER PROGRAMS

Table S1. Accession numbers (UniProtKB/TrEMBL unless otherwise stated) of hypothetical LPMO amino acid sequences found in 19 ascomycete proteomes by CPRENN and HMMER programs. CAZy families are indicated in **bold**

Ascomycete	Proteome accession number (UniProt)	Accession numbers of hypothetical LPMO amino acid sequences	
		CPRENN	HMMER
<i>Arthroderma gypseum</i>	UP000002669	AA11: E4UPP1, E4USU7, E4UV55, E4V726	AA10: E4UPB3; AA11: E4UPP1, E4USU7, E4UV55, E4V2H4, E4V726; AA13: E4V1A1
<i>Arthroderma otae</i>	UP000002035	AA11: C5FI90, C5FPY9, C5G141	AA10: C5FSS3; AA11: C5FI90, C5FPY9, C5G141, C5G191; AA13: C5FUG4
<i>Aspergillus nidulans</i>	UP000000560	AA9: C8V4I9, C8V530, Q5AQA6, Q5AZ52, Q5B6H0, Q5B8T4, Q5BAP2, Q5BEI9, sp Q5BCX8; AA11: Q5AU55, Q5BFS8; AA13: Q5B027, Q5B1W7; AA16: Q5BFA2	AA9: C8V4I9, C8V528, C8V530, Q5AQA6, Q5AZ52, Q5B6H0, Q5B8T4, Q5BAP2, Q5BEI9, sp Q5BCX8; AA11: Q5AU55, Q5BFS8; AA13: C8VLL3, G5EAT0, Q5AWC8, Q5B027, Q5B1W7; AA15: Q5B027, Q5B1W7; AA16: Q5BFA2
<i>Aspergillus niger</i>	UP000006706	AA9: A2QJX0, A2QR94, A2QYU6, A2QZE1, A2R313, A2R5J9, sp A2R5N0; AA11: A2QSQ7, A2QUC6, A2R516; AA16: A2QAR2	AA9: A2QJX0, A2QR94, A2QYU6, A2QZE1, A2R313, A2R5J9, sp A2R5N0; AA11: A2QSQ7, A2QUC6, A2R516; AA13: A2QHE1; AA16: A2QAR2
<i>Ceratocystis fimbriata</i>	UP000222788	AA9: A0A2C5X1Y4, A0A2C5X264, A0A2C5XES6, A0A2C5XGQ0; AA11: A0A2C5WV73, A0A2C5XCJ6, A0A2C5XK11	AA9: A0A2C5X1Y4, A0A2C5X264, A0A2C5XES6, A0A2C5XGQ0; AA11: A0A2C5WV73, A0A2C5XCJ6, A0A2C5XK11; AA13: A0A2C5X2Z5, A0A2C5XD69; AA16: A0A2C5XCI2
<i>Chaetomium globosum</i>	UP000001056	AA9: Q2GPR1, Q2GQE9, Q2GRM4, Q2GRY8, Q2GSI4, Q2GTT3, Q2GU73, Q2GUS9, Q2GUT0, Q2GVG8, Q2GVY2, Q2GW98, Q2GWH8, Q2GWI6, Q2GWR1, Q2GWR7, Q2GX53, Q2GYQ9, Q2GY77, Q2GZ75, Q2GZM2, Q2GZS9, Q2H173, Q2H1G2, Q2H362, Q2H3R3, Q2H5C1, Q2H5E3, Q2H5K6, Q2H7W2, Q2H7X8,	AA9: Q2GM51, Q2GPR1, Q2GQE9, Q2GQS9, Q2GR23, Q2GRM4, Q2GRY8, Q2GSI4, Q2GTT3, Q2GU73, Q2GUS9, Q2GUT0, Q2GVG8, Q2GVY2, Q2GW98, Q2GWH8, Q2GWI6, Q2GWR1, Q2GWR7, Q2GX53, Q2GYQ9, Q2GY77, Q2GZ75, Q2GZM2, Q2GZS9, Q2H173, Q2H1G2, Q2H298, Q2H362, Q2H3R3, Q2H5C1, Q2H5E3, Q2H5K6, Q2H7W2, Q2H7X8, Q2H8C9, Q2H8N9, Q2HBV0, Q2HF13, Q2HGH1, Q2HHE2,

Ascomycete	Proteome accession number (UniProt)	Accession numbers of hypothetical LPMO amino acid sequences	
		CPRENN	HMMER
		Q2H8C9, Q2H8N9, Q2HBV0, Q2HF13, Q2HGH1, Q2HHE2, Q2HHT1, Q2HI74; AA11: Q2GNJ6, Q2GRG8, Q2GWY6, Q2HBF7, Q2HET7; AA13: Q2GMU5, Q2GZT1; AA16: Q2GUN2	Q2HHT1, Q2HI74; AA11: Q2GNJ6, Q2GRG8, Q2GWH5, Q2GWY6, Q2HBF7, Q2HET7; AA13: Q2GMU5, Q2GVN7, Q2GZT1; AA15: Q2GZT1; AA16: Q2GUN2, Q2HEM2, Q2HHP0
<i>Chaetomium thermophilum</i>	UP000008066	AA9: G0RYQ6, G0RYX6, G0RZR9, G0S0S9, G0S2P5, G0S3Z1, G0S408, G0S4L0, G0S5P8, G0S6F4, G0S6H3, G0SA07, G0SB71, G0SDM5, G0SFU2, G0SH69; AA11: G0S225, G0S841	AA9: G0RYQ6, G0RYX6, G0RZR9, G0S0S9, G0S2P5, G0S3Z1, G0S408, G0S4L0, G0S5P8, G0S6F4, G0S6H3, G0SA07, G0SB71, G0SDM5, G0SFU2, G0SH69; AA11: G0S225, G0S841; AA13: G0RZA2; AA14: G0S074, G0SBL0
<i>Coccidioides immitis</i>	UP000054565	AA11: A0A0J7B1F1, A0A0J7B2J0	AA11: A0A0J6YEW0, A0A0J7B1F1, A0A0J7B2J0; AA13: A0A0J6XZ06
<i>Cyphellophora europaea</i>	UP000030752	AA9: W2RZZ2, W2S3Q1, W2S7D9, W2SA85, W2SCD4, W2SDR8; AA11: W2S9C7	AA9: W2RZZ2, W2S3Q1, W2S7D9, W2SA85, W2SCD4, W2SDR8; AA11: W2S208, W2S9C7, W2SEN5; AA13: W2RNE4; AA14: W2RVU8, W2S6P0, W2SDJ0
<i>Hypocrea jecorina</i>	UP000024376	AA9: A0A024SDR7, A0A024SFJ2, A0A024SM10; AA11: A0A024SBU8, A0A024SK73, A0A024SL74	AA9: A0A024SDR7, A0A024SFJ2, A0A024SM10; AA10: A0A024SC98; AA11: A0A024SBU8, A0A024SK73, A0A024SL74; AA13: A0A024SN40; AA14: A0A024RZI4, A0A024SGF0
<i>Metarhizium anisopliae</i>	UP000054544	AA9: A0A0D9P2I3; AA11: A0A0D9NR53, A0A0D9NRB7, A0A0D9NWB9, A0A0D9P120, A0A0D9P671, A0A0D9P927	AA9: A0A0D9P2I3; AA10: A0A0D9NQP0; AA11: A0A0D9NR53, A0A0D9NRB7, A0A0D9NWB9, A0A0D9P120, A0A0D9P671, A0A0D9P927; AA13: A0A0D9PB9; AA14: A0A0D9NVH7, A0A0D9NYG3, A0A0D9P8Q5
<i>Myceliophthora thermophila</i>	UP000007322	AA9: G2Q2F7, G2Q4M0, G2Q6N6, G2Q774, G2Q7A5, G2Q9F7, G2Q9N9, G2Q9T8, G2QA92, G2QAB5, G2QCJ3, G2QCS9, G2QFY5, G2QHR1, G2QI82, G2QJR7, G2QJT0, G2QK49, G2QLE4, G2QNT0, G2QNY4, sp G2Q9T3; AA11: G2PZJ4, G2Q4Y4, G2Q8W8, G2QNB8; AA13: G2QP40; AA16: G2QH80, G2QJ05, G2QNN7	AA9: G2Q2F7, G2Q4M0, G2Q6N6, G2Q774, G2Q7A5, G2Q9F7, G2Q9N9, G2Q9T8, G2QA92, G2QAB5, G2QCJ3, G2QCS9, G2QFY5, G2QHR1, G2QI82, G2QJR7, G2QJT0, G2QK49, G2QLE4, G2QNT0, G2QNY4, sp G2Q9T3; AA11: G2PZJ4, G2Q4Y4, G2Q8W8, G2QNB8; AA13: G2Q1W3, G2QP40, G2QPS0; AA14: G2Q335; AA16: G2QH80, G2QJ05, G2QNN7

Ascomycete	Proteome accession number (UniProt)	Accession numbers of hypothetical LPMO amino acid sequences	
		CPRENN	HMMER
<i>Neurospora crassa</i>	UP000001805	AA9: Q1K4Q1, Q1K8B6, Q7RV41, Q7RWN7, Q7S111, Q7S1A0, Q7S1V2, Q7S411, Q7S439, Q7SA19, Q7SAR4, Q7SCJ5, Q7SHD9, Q7SHI8; AA11: Q1K703, Q1K8R8, Q7S0P1, Q7SEB2; AA13: Q7SCE9	AA9: Q1K4Q1, Q1K8B6, Q7RV41, Q7RWN7, Q7S111, Q7S1A0, Q7S1V2, Q7S411, Q7S439, Q7SA19, Q7SAR4, Q7SCJ5, Q7SHD9, Q7SHI8; AA11: Q1K703, Q1K8R8, Q7S0P1, Q7SEB2; AA13: Q7SCE9, Q7SDJ6, sp P14804; AA14: Q7SF43
<i>Sordaria macrospora</i>	UP000001881	AA9: F7VLD7, F7VQ67, F7VR19, F7VRX4, F7VSL3, F7VUC2, F7VVG9, F7VVJ2, F7VVV6, F7VXA8, F7VZC6, F7VZC7, F7W1P4, F7W2U5, F7W4L9, F7W4V7, F7W4Z9, F7WAD8, F7WC55; AA11: F7VR04, F7VZH1, F7W0R2; AA13: F7VQ11; AA16: F7VJY8, F7VMX5	AA9: F7VLD7, F7VQ67, F7VR19, F7VRX4, F7VSL3, F7VUC2, F7VVG9, F7VVJ2, F7VVV6, F7VXA8, F7VZC6, F7VZC7, F7W1P4, F7W2U5, F7W4L9, F7W4V7, F7W4Z9, F7WAD8, F7WC55; AA11: F7VR04, F7VZH1, F7W0R2, F7W4W2; AA13: F7VP41, F7VQ11, F7VRR8, F7WA61; AA14: F7VLT0; AA15: F7VQ11; AA16: F7VJY8, F7VMX5
<i>Talaromyces marneffei</i>	UP000029285	AA9: A0A093UTW6; AA11: A0A093VS10	AA9: A0A093UTW6; AA11: A0A093VS10; AA13: A0A093V3C8, A0A093V3D6, A0A093V4S3
<i>Talaromyces stipitatus</i>	UP000001745	AA9: B8M2G3; AA11: B8LX66, B8MBC2, B8MBJ5, B8MQN3	AA9: B8M2G3; AA11: B8LX66, B8MBC2, B8MBJ5, B8MQN3; AA13: B8M0W0, B8M188, B8MKW4, B8ML80
<i>Thielavia terrestris</i>	UP000008181	AA9: G2QQL2, G2QRC5, G2QUZ5, G2QV07, G2QYT2, G2QZK6, G2QZU1, G2R009, G2R1U0, G2R6N0, G2R7W4, G2R898, G2R8B3, G2R9S5, G2RB72, G2RB73, G2RGE5, G2RGE6; AA11: G2QTD0, G2QVT1, G2QW76, G2RCI4, G2RG22; AA16: G2RGB5	AA9: G2QQL2, G2QRC5, G2QUZ5, G2QV07, G2QYT2, G2QZK6, G2QZU1, G2R009, G2R1U0, G2R286, G2R6N0, G2R7W4, G2R898, G2R8B3, G2R9S5, G2RB72, G2RB73, G2RCT3, G2RGE5, G2RGE6; AA11: G2QTD0, G2QVT1, G2QW76, G2RCI4, G2RG22; AA13: G2QUZ8, G2R2D7, G2RHN1; AA14: G2RF87; AA16: G2RGB5
<i>Trichophyton rubrum</i>	UP000008864	AA11: F2SK00, F2SKG2, F2SNJ5, F2SZ47	AA10: F2SWK9; AA11: F2SK00, F2SKG2, F2SNJ5, F2SZ47, F2T040; AA13: F2SXE3
<i>Uncinocarpus reesii</i>	UP000002058	AA11: C4JKV1, C4JM64, C4JPQ2	AA11: C4JKV1, C4JM64, C4JPQ2; AA13: C4JS39

Table S2. Accession numbers (UniProtKB/TrEMBL unless otherwise stated) of hypothetical LPMO amino acid sequences found in 18 basidiomycete proteomes by CPRENN and HMMER programs. CAZy families are indicated in **bold**

Basidiomycete	Proteome accession number (UniProt)	Accession numbers of hypothetical LPMO amino acid sequences	
		CPRENN	HMMER
<i>Botryobasidium botryosum</i>	UP000027195	AA9: A0A067LUS6, A0A067LX50, A0A067LYP7, A0A067M0B2, A0A067M1H2, A0A067M2W6, A0A067M2Y0, A0A067M4U4, A0A067M4W5, A0A067M8V9, A0A067M9A9, A0A067M9Q4, A0A067MAD5, A0A067MAJ4, A0A067MC82, A0A067MDF4, A0A067MEW5, A0A067MGW7, A0A067MHC2, A0A067MIV4, A0A067MJE3, A0A067ML63, A0A067MRF6, A0A067MTZ4, A0A067MU04, A0A067MUJ1, A0A067MYB6, A0A067N321, A0A067N827, A0A067N9L8, A0A067NAJ9, A0A067NCM0; AA10: A0A067MNS7	AA9: A0A067LUS6, A0A067LX50, A0A067LYP7, A0A067M0B2, A0A067M1H2, A0A067M2W6, A0A067M2Y0, A0A067M4U4, A0A067M4W5, A0A067M8V9, A0A067M9A9, A0A067M9Q4, A0A067MA70, A0A067MAD5, A0A067MAJ4, A0A067MC82, A0A067MDF4, A0A067MEW5, A0A067MGW7, A0A067MHC2, A0A067MIV4, A0A067MJE3, A0A067ML63, A0A067MRF6, A0A067MTZ4, A0A067MU04, A0A067MUJ1, A0A067MYB6, A0A067N321, A0A067N827, A0A067N9L8, A0A067NAJ9, A0A067NCM0; AA10: A0A067MNS7; AA13: A0A067LUS3, A0A067MYQ6; AA14: A0A067LZ47, A0A067LZE3, A0A067M1W2, A0A067M720, A0A067M9Y1, A0A067MLD3, A0A067MTX7, A0A067MVP4; AA16: A0A067M7V3, A0A067MFP1
<i>Ceriporiopsis subvermispora</i>	UP000016930	AA9: M2PHI6, M2PI15, M2QEW2, M2QVZ7, M2QX02, M2QYQ9, M2R9L3, M2RAI8, M2RL70; AA14: M2QZB0	AA9: M2PHI6, M2PI15, M2QEW2, M2QVZ7, M2QX02, M2QYQ9, M2R9L3, M2RAI8, M2RL70; AA13: M2QWA3, M2RLN3, M2RLU2, M2RNS1; AA14: M2QZB0, M2R6U9
<i>Coprinopsis cinerea</i>	UP000001861	AA9: A8NBA9, A8NBB0, A8NC30, A8NCG7, A8NCG8, A8NCH1, A8NKH7, A8NKI3, A8NKI4, A8NLH8, A8NLT0, A8NLZ6, A8NPW5, A8NPW7, A8NRC9, A8NUU4, A8P0V5, A8P3S3, A8P8P6, A8PBN8, A8PER1, A8PET0, A8PFJ0, A8PFS9, A8PFT1, A8PI41	AA9: A8N681, A8NBA9, A8NBB0, A8NC20, A8NC30, A8NCG7, A8NCG8, A8NCH1, A8NCV0, A8NKH7, A8NKI3, A8NKI4, A8NLH8, A8NLT0, A8NLZ6, A8NPW5, A8NPW7, A8NRC9, A8NUU4, A8P0V5, A8P3S3, A8P8P6, A8PBN8, A8PER1, A8PET0, A8PFJ0, A8PFS9, A8PFT1, A8PGQ9, A8PI41, D6RKI0, D6RNP5; AA13: A8NRI4, A8NSG1; AA14: A8N3C6, A8NQ92, A8NQA1, A8NV23, A8PAN7
<i>Cryptococcus neoformans</i>	UP000002149	AA9: F5HH24; AA11: Q5KE44; AA14: A0A0S2LHR4	AA9: F5HH24; AA11: Q5KE44; AA14: A0A0S2LHR4, Q5KEF3
<i>Fibroporia radiculosa</i>	UP000006352	AA9: J4HU60, J4I7Z1	AA9: J4HU60, J4I7Z1; AA13: J4I862; AA14: J4G816, J4GP13, J4GSV2
<i>Fomitopsis pinicola</i>	UP000015241	AA9: S8DTC0, S8EBE2, S8EE26, S8F7D2	AA9: S8DTC0, S8EBE2, S8EE26, S8F7D2; AA13: S8DMF8, S8DWK5; AA14: S8DVX7, S8EKB3, S8FWT6, S8G5V6

Basidiomycete	Proteome accession number (UniProt)	Accession numbers of hypothetical LPMO amino acid sequences	
		CPRENN	HMMER
<i>Galerina marginata</i>	UP000027222	AA9: A0A067SEZ5, A0A067SUS0, A0A067SV93, A0A067SZQ0, A0A067SZR5, A0A067T140, A0A067T399, A0A067T410, A0A067T6I0, A0A067TDB4, A0A067TDQ2, A0A067TJD3, A0A067TJT4, A0A067TM36, A0A067TPI0, A0A067TRR1, A0A067TUC4, A0A067TUE6; AA10: A0A067SN17	AA9: A0A067SEZ5, A0A067SUS0, A0A067SV93, A0A067SZQ0, A0A067SZR5, A0A067T140, A0A067T399, A0A067T410, A0A067T6I0, A0A067TDB4, A0A067TDQ2, A0A067TJD3, A0A067TJT4, A0A067TM36, A0A067TPI0, A0A067TRR1, A0A067TUC4, A0A067TUE6; AA10: A0A067SN17; AA13: A0A067T4C3, A0A067TJS5, A0A067TUD9; AA14: A0A067SQ32, A0A067T677, A0A067TSY6; AA16: A0A067SXJ4
<i>Gloeophyllum trabeum</i>	UP000030669	AA9: S7Q0E7, S7QKE2, S7QP81, S7RK00; AA14: S7Q411	AA9: S7Q0E7, S7QKE2, S7QP81, S7RK00; AA13: S7Q3A2, S7Q4V9; AA14: S7Q3B1, S7Q411
<i>Heterobasidion irregulare</i>	UP000030671	AA9: W4JR58, W4JUX2, W4JX02, W4JZ18, W4K048, W4K498, W4K8M0, W4K9D6, W4KIK0, W4KMP1; AA14: W4KAJ4	AA9: W4JR58, W4JUX2, W4JX02, W4JZ18, W4K048, W4K498, W4K8M0, W4K9D6, W4KIK0, W4KMP1; AA13: W4JNY4, W4JS51, W4K832; AA14: W4JSK5, W4KAG0, W4KAJ4; AA16: W4JYN3
<i>Jaapia argillacea</i>	UP000027265	AA9: A0A067P992, A0A067PCP6, A0A067PD40, A0A067PDQ4, A0A067PE53, A0A067PFS2, A0A067PKN5, A0A067PN15, A0A067PN21, A0A067PQV6, A0A067PT15, A0A067PZX3, A0A067Q4K5, A0A067Q7C0	AA9: A0A067P6H1, A0A067P992, A0A067PCP6, A0A067PD40, A0A067PDQ4, A0A067PE53, A0A067PFS2, A0A067PKN5, A0A067PN15, A0A067PN21, A0A067PQV6, A0A067PT15, A0A067PZX3, A0A067Q4K5, A0A067Q7C0; AA13: A0A067PTB3, A0A067PW10, A0A067Q6R3; AA14: A0A067Q6I8, A0A067Q911; AA16: A0A067PN45, A0A067PN58, A0A067PY66, A0A067Q0U4
<i>Laccaria bicolor</i>	UP000001194	AA9: B0CQE5, B0DGU6, B0DPG3	AA9: B0CQE5, B0CZ36, B0DEE6, B0DGU6, B0DIW0, B0DPG3, B0DVP6; AA13: B0CVJ1; AA14: B0CYY7, B0D2M4, B0DH69, B0DN14
<i>Phanerochaete carnosa</i>	UP000008370	AA9: K5UN12, K5VKD6, K5VN09, K5VXP2, K5W115, K5WEP2, K5X3Y3, K5X412, K5X7Y6, K5XB47	AA9: K5UN12, K5VKD6, K5VN09, K5VWY7, K5VXP2, K5W115, K5WEP2, K5X3Y3, K5X412, K5X7Y6, K5XB47; AA13: K5W2K6, K5W345, K5WMZ0; AA14: K5WFH2, K5WGL0, K5WMA2
<i>Pleurotus ostreatus</i>	UP000027073	AA9: A0A067N594, A0A067N5S2, A0A067NBB9, A0A067NEB7, A0A067NHA4, A0A067NIQ6, A0A067NIV2, A0A067NJX7, A0A067NLK7, A0A067NNU6, A0A067NPI2, A0A067NTB5, A0A067NVR6, A0A067NWT8, A0A067NYI6, A0A067NYR2, A0A067NZ88,	AA9: A0A067N594, A0A067N5S2, A0A067NBB9, A0A067NEB7, A0A067NHA4, A0A067NIQ6, A0A067NIV2, A0A067NJX7, A0A067NLK7, A0A067NNU6, A0A067NPI2, A0A067NTB5, A0A067NVR6, A0A067NWT8, A0A067NYC9, A0A067NYI6, A0A067NYR2, A0A067NZ88, A0A067P022, A0A067P4T8, A0A067P559, A0A067P6A2, A0A067P6D0, A0A067P815, A0A067PAD5, A0A067PAJ1, A0A067PBW4, A0A067PCB2;

Basidiomycete	Proteome accession number (UniProt)	Accession numbers of hypothetical LPMO amino acid sequences	
		CPRENN	HMMER
		A0A067P022, A0A067P4T8, A0A067P559, A0A067P6A2, A0A067P6D0, A0A067P815, A0A067PAD5, A0A067PAJ1, A0A067PBW4, A0A067PCB2	AA13: A0A067NJI9, A0A067NPC1, A0A067NPJ8; AA14: A0A067NG07, A0A067NNI0, A0A067P1E5; AA16: A0A067N4F9
<i>Postia placenta</i>	UP000194127	AA9: A0A1X6NIH6; AA14: A0A1X6MR03	AA9: A0A1X6N2G5, A0A1X6NIH6; AA13: A0A1X6MTY1; AA14: A0A1X6MR03, A0A1X6MR05, A0A1X6N3G9
<i>Pycnoporus cinnabarinus</i>	UP000029665	AA9: A0A060S4D7, A0A060S5J9, A0A060S8F5, A0A060S9V0, A0A060SA61, A0A060SB47, A0A060SH23, A0A060SHL9, A0A060SJX3, A0A060SPE8, A0A060SPU4, A0A060SQW9, A0A060SU08, A0A060SU90, A0A060SY86; AA14: A0A060SMC0, A0A060SRI5, A0A060SS34, A0A060SWU0	AA9: A0A060S4D7, A0A060S5J9, A0A060S8F5, A0A060S9V0, A0A060SA61, A0A060SB47, A0A060SH23, A0A060SHL9, A0A060SJX3, A0A060SP53, A0A060SPE8, A0A060SPU4, A0A060SQW9, A0A060SU08, A0A060SU90, A0A060SY86; AA13: A0A060S3R1; AA14: A0A060SMC0, A0A060SRI5, A0A060SS34, A0A060SWU0
<i>Schizophyllum commune</i>	UP000007431	AA9: D8PNG1, D8PQJ8, D8PVB3, D8PZ63, D8Q364, D8Q696, D8Q8A3, D8QB45, D8QB46, D8QBP8, D8QGA8, D8QH48, D8QHF7, D8QHH2, D8QHN3, D8QHN6, D8QJ75, D8QJB9, D8QJC1; AA11: D8PYP3, D8Q4M7, D8QAT3, D8QAT9, D8QB70, D8QK94, D8QLA8, D8QMF2; AA16: D8PQ52	AA9: D8PNG1, D8PQJ8, D8PVB3, D8PZ63, D8Q364, D8Q5J9, D8Q5V3, D8Q696, D8Q8A3, D8QB45, D8QB46, D8QBP8, D8QGA8, D8QH48, D8QHF7, D8QHH2, D8QHN3, D8QHN6, D8QJ75, D8QJB6, D8QJB9, D8QJC1; AA11: D8PYP3, D8Q4M7, D8QAT3, D8QAT9, D8QB70, D8QK94, D8QLA8, D8QMF2; AA13: sp D8Q9M3; AA14: D8PQI8, D8PUM5, D8PWG2; AA16: D8PQ52, D8Q6G1
<i>Ustilago maydis</i>	UP000000561	AA10: A0A0D1CHL0	AA10: A0A0D1CHL0
<i>Wolfiporia cocos</i>	UP000218811	AA9: A0A2H3ISM5, A0A2H3JAV7	AA9: A0A2H3ISM5, A0A2H3JAV7; AA13: A0A2H3JCA1; AA14: A0A2H3JF26, A0A2H3JKL4, A0A2H3K018

Table S3. Accession numbers (UniProtKB/TrEMBL unless otherwise stated) of hypothetical LPMO amino acid sequences found in 18 bacterial proteomes by CPRENN and HMMER programs. CAZy families are indicated in **bold**

Bacterium	Proteome accession number (UniProt)	Accession numbers of hypothetical LPMO amino acid sequences	
		CPRENN	HMMER
<i>Actinoplanes missouriensis</i>	UP000007882	AA10: I0H648, I0H7A5, I0HBF4, I0HBN9, I0HBU0, I0HCK5, I0HJD3	AA10: I0H3D8, I0H648, I0H7A5, I0H883, I0H8C5, I0HAU5, I0HBF4, I0HBN9, I0HBU0, I0HCK5, I0HHI7, I0HJD3; AA13: I0H3G0, I0HDC7
<i>Bacillus thuringiensis</i>	UP000011719	AA10: M1QAC4, M1QMP1, M1QYJ5	AA10: M1QAC4, M1QMP1, M1QYJ5, M1QZ54
<i>Brevibacillus laterosporus</i>	UP000005850	AA10: A0A075R1J7, A0A075R6S4	AA10: A0A075R1J7, A0A075R6S4
<i>Enterobacter cloacae</i>	UP000007838	–	–
<i>Herpetosiphon aurantiacus</i>	UP000000787	AA10: A9B107	AA10: A9B107; AA13: A9AVU7
<i>Klebsiella oxytoca</i>	UP000236461	AA10: A0A2I8WKD6	AA10: A0A2I8WKD6
<i>Lactococcus lactis</i>	UP000002196	AA10: Q9CE94	AA10: Q9CE94, Q9CE95
<i>Listeria monocytogenes</i>	UP000000817	AA10: Q8Y4H4	AA10: Q8Y4H4, Q8Y619
<i>Nocardiopsis dassonvillei</i>	UP000002219	AA10: D7AW10, D7AWR9, D7B4C4, D7B925	AA10: D7AW10, D7AWR9, D7B4C4, D7B5S5, D7B925
<i>Photorhabdus asymbiotica</i>	UP000002747	AA10: C7BK22	AA10: C7BK22
<i>Pseudomonas aeruginosa</i>	UP000002438	AA10: sp Q9I589	AA10: Q9I1H5, sp Q9I589
<i>Saccharophagus degradans</i>	UP000001947	AA10: Q21N34	AA10: Q21N11, Q21N34, Q21PQ1, Q7PC52
<i>Salinispora arenicola</i>	UP000001153	AA10: A8LUM8, A8LVC0, A8LVQ6, A8LVV7	AA10: A8LUM8, A8LV01, A8LVC0, A8LVQ6, A8LVV7; AA13: A8M6K2
<i>Serratia marcescens</i>	UP000050507	AA10: A0A1C0CF47, A0A1C0DMQ2, A0A1C0P555	AA10: A0A1C0CF47, A0A1C0DMQ2, A0A1C0E185, A0A1C0P555
<i>Stenotrophomonas maltophilia</i>	UP000006955	AA10: J7USI5, J7V8J6	AA10: J7USI5, J7V8J6
<i>Streptomyces viridochromogenes</i>	UP000011205	AA10: L8P3K0, L8PAZ5, L8PB92, L8PL89, L8PLS9, L8PMS0	AA10: L8NZP1, L8P3K0, L8PAZ5, L8PB92, L8PIH2, L8PL89, L8PLS9, L8PMS0
<i>Vibrio cholerae</i>	UP000000584	AA10: Q9KN26, sp Q9KLD5	AA10: Q9KN26, Q9KQP6, sp Q9KLD5
<i>Xylanimonas cellulosilytica</i>	UP000002255	AA10: D1BV23, D1BWE0	AA10: D1BV23, D1BW40, D1BWE0