

A

GH1-1 M-----KTILI----VSICCSVI-----
GH1-2 M-----EYTFC----LLFLWVGR-----
GH1-3 M-----KLFIL----ICLIALCF-----
GH1-4 M-----KVFLF----ALLLGVCL-----
GH1-5 M-----KLVGCFLVFLS----VLALGAAD-----
GH1-6 M-----DKVL----VIAVFSFF-----
GH1-7 M-----TANHFFYFLFLV-----
GH1-8 M-----AYQLLRKCEDTMCKVIYFLVLLNYL-----
GH1-9 M-----MSPKHLTLLRYLIYICLTFNL-----
GH1-10 M-----
GH1-11 M-----VRALHILATLLALKLSYGNPQPGYPVDPHCPPEDCEPYPVHCPNPIPEPVCEIPVIPPPHYPHPIPDCDYPPGYPHG
GH1-12 M-----TSPVLI----FILVHLSN-----
GH1-13 M-----LLNFRVLIL----TVSFSNIL-----
GH1-14 M-----LCIKF---IAVIFILNL-----
GH1-15 M-----KVSLL----LLLVVSSV-----
GH1-16 M-----KVILI----SLLGLSL-----
GH1-17 M-----RSDLQTMAPVPKIFFL----IFVVVSSY-----
GH1-18 M-----FRHRVWV----LCVLVISA-----
GH1-19 M-----ISKELW----LLCIVLST-----
GH1-20 M-----KIYV----LLLIVCTS-----
GH1-21 M-----TLKYMLI----TFLLVGIT-----
GH1-22 MIGSFIYKSFGFELKFEFKSFKMMWI----YMVFGFLL-----
GH1-23 -----
GH1-24 -----
GH1-25 -----
GH1-26 M-----KMKFTIFTSLVAVFA-----
GH1-27 -----
GH1-28 -----

GH1-1 -----
GH1-2 -----
GH1-3 -----
GH1-4 -----
GH1-5 -----
GH1-6 -----
GH1-7 -----
GH1-8 -----
GH1-9 -----
GH1-10 -----
GH1-11 PINPGPLPPhLPLPPhLPiPKPCDGLPGAWPNEWPCYFPCGWPDDWPCDFPLGLPNNWPIDFPHGLPHDWPNNFPHGWPVPHGYPNEWPCWPDNWQEYW
GH1-12 -----
GH1-13 -----
GH1-14 -----
GH1-15 -----
GH1-16 -----
GH1-17 -----
GH1-18 -----
GH1-19 -----
GH1-20 -----
GH1-21 -----
GH1-22 -----
GH1-23 -----
GH1-24 -----
GH1-25 -----
GH1-26 -----
GH1-27 -----
GH1-28 -----

GH1-1 -----VLR----TS--TAKEEFKPKDFLFGIASSAYQIEGGYNADGKGESIYDYTKTKHARFYNNNGNIACDSYHKWRADVRLLKDLG
GH1-2 -----TTS----AK--LPMYDFPKKFSFGVATSAFQIEGGWNASGKGVSIWDHLVHDTNEIIVDGSNADVACDSYNKVDYDILLLKALG
GH1-3 -----SAN----AKDPINNRRFPKTFKVGVANAAQIEGAWNTDGKGETIWDHFAHTQPEKFIIDRSTPDVACNSYQYKEDIAMVKEMG
GH1-4 -----AQN----AD--TNPRRFPASFKLGVANAAAQIEGAWNEEDGKGEHIWDRWSKVRPEMISDRSTPAVAADSYHKYKEDVRLIKEMG
GH1-5 -----GDDEFANYT--ITNRKFPKGFTFGVATAAYQIEGAWNLGKGEQVWDTFFHEHPDRVDDGTNGDVATDSYHLYKEDVRCMEEVG
GH1-6 -----TTC----FA--DNDRTFPETFSGAATAAFQIEGAWNENKGESIWDRFIHTNQSRVVDGSNADIACDSYHKYKEDVALAAMKMG
GH1-7 -----TYA----QA--KNGSFPPKDFLFGVATSAFQIEGAWNEEDGKGENIWDHMHVQVPSPIKNNDTGDIACDSYHKYKEDVDILEDLG
GH1-8 -----RFS----DA--KNGTFPANFLFGVATSSQIEGAWNEEDGKGESMWDEYNHRVPSPIRNGDTADIACDSYHKYKEDVKLVADLG
GH1-9 -----VAA----ESQELHYNRKFPPDFLFGAATSAYQIEGAWNVGKGESIWDTYVHRVPSPIKNGDTGDIACDSYKYKEDVKLAADLG
GH1-10 -----ENRGENIWDHAIKANPKIIEDGSSNASSADESYTHYEDDFKALAQMG
GH1-11 PDCWPNYWDGCDYPYPAGSVPAIPPPVESCDVFPKESFGVASGAYHIEGAWDQDCRSESIWDRWLHTCPSVLDGSGNDISADAYNHKGDIELLNLG
GH1-12 -----NVA----TS--INYNRFPDFLWGAATSFAFQIEGAWNEEDGKAPSVWDYFLHNQPYFVDPDGNADVANSANAYHMVETDVELAKSLN
GH1-13 -----AED----YV--INNKRFPADFMFGVATSAFQIEGGWNEEDGKGVNIWDYFTHRNSSVIRDKNNGDIACDSYHRIKEDIAILKDLG
GH1-14 -----GFL----DA--LNNKKFPDFIFGCATAAFQIEGAWNEEDGKGPSVWDTTNNKIPSVIEDNSTADVSCDSYHKYKEDVALLKELG
GH1-15 -----YST-----PDSLFPPEGLFPGVSTAAYQIEGGWQDQKGENIWDHQLHTNSNITIDHSNGDVACDSYHRIKRDVILNLD
GH1-16 -----AED----P--INNKKFPDFLFGVASSAYQIEGAWNEEDGKGESLWDHRVHTDPNFIEDNSTGDVACDSYHKIKEDVAILKELG
GH1-17 -----ADD----YQ--INNKRFPNFMFGVGTSSLQIEGAPNEEDGKTEMTWERYIADHPGYVADGSDVSVACDSYHQYKRDVAMVKELG
GH1-18 -----TCG----EE--LNNKTFPGYFIFGTATASYQIEGAWNEEDGKQNIWDNFTHEDPSRIFGNANGDVACDSYHKYKEDVALLKDLG
GH1-19 -----SLA----TT--INVSFKPKDFMFGTATASYQIEGAWNEEDGKGENTWDEVTNRVPSPIVDNSTGDVACDSYHKYKEDVAMLKHLG
GH1-20 -----KLV----EP--LNNRRFPAHFMIGAATASYQIEGAWNKDGKGENIWDRLTHSPNSPVLGNQNGNIACDSYHKYKEDVAMLKHLG
GH1-21 -----VCD----TE--LNKKRFPKDFMFGVATSSYQIEGAWNEEDGKGESTWDHFVHQTPSPIVNNYTGDVACDSYHKMKEDVALLKYL
GH1-22 -----SNA----IA--DDNTRKFPPEGTFFGVATSAFQIEGAWNEEDGKGENNDRVAHTIPNFIYHNSGDVTCDSYHQWEEDIRILKELK
GH1-23 -----
GH1-24 -----MS----NS--EDQKRFPEGLFPGTASSAYQIEGGYNVDGKGESVYDYTHKTPPEKFNHNGSDVACDSYHKWQEDVHILKDLG
GH1-25 -----
GH1-26 -----KTT----QS--LNTKKFPKFMFGTATAAYQIEGGWNEEDGKGENNWDHLSHSPSPIDHHTGDVACDSYHKYKEDVQLLKLK
GH1-27 -----
GH1-28 -----

GH1-1 VQFYRFSISWSRVLPGY-SNKVNPDP--GLRYNDLINELKNGIEPMVTLYHNDLPMPLQE-IG-GWTNPYTAYFYEDYARIMFSYFGDRVKYWITM--
GH1-2 VSHYRFSIAWTRILPNGF-RNHVSED--GVKYNYNLINKLLENGIEPVWTIFHFDPYGMFL-IG-GMFNEKFSEYLSDYADLLFNLYGDRVKTWVTINE
GH1-3 LDHYRFSIAWSRILPTGY-LDSLNLQ--GVQYKKNVQELKKNNIEPLVTLYHWDLPQLQDQMN-GWINETIIDIFADYCKLFCFLFGDDVKWITINE
GH1-4 LDVYRFSISWARIYPTGYPTAEPNAA--GVQYKKNLAECKQNGIQMATLYHWDLPQLQDDFG-GWLNETVVDLFAAFADVSFKLFGDDVTWITINE
GH1-5 VDYRFSIAWSRILPDGT-LDNVSE--GVNYVNLFLKELRAKGIKSFVTLYHWDIPTALEK-QG-GWRNPKVVDWFQDYAEFCFKTFGEYVDSWVTINE
GH1-6 LTHYRFSISWSRLLPTGF-ISPVNLI--AISYQDLIRELEKYKIEPVVTIYHWDLPQMLQE-LG-GWENSKIYDYFVDYADLVFASF-PSVKHWITFNE
GH1-7 VNFYRFSISWSRILPTGY-PSVTKA--GLEYYKNLTKELVSRNITPVATIFHWDLQPQLSE-ID-GWSNPRADLFEVYSRIVIQGL-PDVGWVITFNE
GH1-8 ADFYRFSISWPRIPTGY-TDNINQK--GIQYQNLVKEILKYNMVPATLYHWELPQTLVA-SNLDWVNPDIVDIFVNYTRVVIQNL-PDVGWVITVNE
GH1-9 IKLYRFSISWPRLPTGK-ADNISEA--GLNYKKNLVEIRNYGMIPVATIIYHWDLPQKLYD-EGIHWSNVSLVPIYDVARIVIKAL-PDVGWVITINE
GH1-10 ITHYKLSISWSRMIPTGY-TDNIDMSNGTTHYQALISELVKQKITPVVLYNWDLPQPIQD-YG-GWTNESNIDAFVYARKVFK-YCSEAKYFITFDN
GH1-11 VRHYVFSISWPRIYDYG--CGEVNPNY--GIEHYRKLKLLKLNHIEPIVVLYYGDLPQALQD-QG-GFLCPSFVEWYTEYARVCFEFCGDDVKYWITFHK
GH1-12 LKVFRTSVAWTRILPQGN-EGKPNQI--AIDHYKNLFRLLKENDIEPLVTMHSFDLPQKLSL-YG-GFLNESLVWEFTNYARICFDNFGDDVKQWATFNE
GH1-13 VSHYRFSISWPRIPTGY-RDEVNQK--GIDYKKNLTKELKDNNIEPLVTLYHWDLPQLQD-LG-GFFSSSLEVWFAEYARICFEAFGKDVKYWVITFNE
GH1-14 ATHYRFSISWPRIPTGY-NDYINPL--GLAYYKNLIAELKANNIEPLVTISHGDLQVQLGD-IG-GYNNESFPDWRDFARVAFDEFEGDDVKYWFTINE
GH1-15 VDHYRFSISWSRILPTGL-DDNINEA--GVNYKKNLKLKIVIQPMVTLHFDLPQKQLD-MG-GFLNESIVDFWGNYSRVAYRFLFGDDVKFWITINE
GH1-16 VNHYRFSLSWTRILPTGL-ADEINVG--GVIYKKNLSELKANNIIPLVTIYHWDLPQVQD-YG-GFLNDSIVDWYSNYARVVFSLFGDDVKHWITINE
GH1-17 VTHYRFSIAWTRILPNGV-NDYVNPL--GVQYKKNLTKELKANNIEPVVSIHWDIPQVQLD-MD-GFLNDSIVDWYTDYARVCFEEFGDDVKYWATFNE
GH1-18 VDHYRFSISWSRVLPTGY-SNEVNEA--GVNYKKNLTKELNDNIFPLVTFMHWDTQPQLD-IG-GWTNDTIIDRFVQYQFLFETFGDDVQMWITFNE
GH1-19 VTHYRFSLSWSRILPTGF-NNKVNLS--GIAYYKKNLTKELRANNIEPLVTIFHWDTQPQLD-LG-GWTNELIVDRFVDFAKVAFENFGDDVYWLTFNE
GH1-20 VNHYRFSISWSRILPTGF-TNRINQA--GITYYKKNLTKELKANNIEPVVSIHWDIPQVQLD-IG-GWTNELIVDKFRDYAKVLFDFNFGDDVKYWVITFNE
GH1-21 VKHYRFSLSWSRILPTGY-TNKINQA--AIKYKKNLIRELTDNGIEPVVTLHCDTPQALED-MG-GWTTDWIIDRYEYAAVCFENFGADVKYWLTFNE
GH1-22 VTHYRFSISWSRLLPHGF-NFEVQA--GVDYTNLIKALKAAANIEPVVTLFHWDLQPFLA-LG-GWPEREIVQWYSEYARLFCFLFGDDVKHWMTFNE
GH1-23 -----MNEWVDRFVEYVEVCFKNGEDVQHWITFNE
GH1-24 VDFYRFSISWSRILPTGY-SNKINEE--GVKYENLIDGLLKNGIQPMITLYHWDLPMPIQE-LG-GWTNIKTAEFYEDYAEIVFERFNGKVKYWITIN-
GH1-25 -----
GH1-26 VNHYRFSISWSRILPSGF-DNKINPL--GIQYKKNLTKELKDNIEPLVTMHHWDTPQPLEN-IG-GWTNEKIVDMFADYARILYENFGDDIKYWITFNE
GH1-27 -----
GH1-28 -----GY-TDNINQK--GIQYQNLVKECLKYNLIPVATLYHWEIPQSLFA-AKLDWTNPAFIDIFVNYTRVVIQNL-PDVGWVITINE

GH1-1 --MSSNGYGD-DKYPPFINQ-PGIANYICQHVMLLAHAKAYHLYDAEFRKNQQGKIGVAVHATWYIPATLS-KDVEAERAREFKIGS IMNPIFHTDGN
GH1-2 PYSVCSVSGTRVLGEYRP-----NGIVEYLCGHTFLKAHAAVYDLYQRKYKSQQNGKISMAFNLQCNYPETNSAEDIEAERANQFALGWVHVHPV--FGN
GH1-3 PKQVCQAGYGS-GAFAPGIVS-NGIMDYVCTKNVLLAHAKAWHIYDEQFREKNKGQVAMVIDATWYEPGSDNQEDKDAERAQQFDIGIYGNPLF--NGD
GH1-4 PKQVCHAGYGD-GYFAPGVVS-SGVGEYTCARNVLLAHAKAYHNYNDNYKASQGGKYSMVIDTLWFEPGSDSDADKDAERLLQFQFGLYGNPIF--NGD
GH1-5 PKQICHAGYGV-GGYAPGVVS-PGIGEYVCARHVLLAHAKTWRLYDRSYRQQLKSRTIVIDSDWYEPETESEAETAEEHQRQFVYGYMANPIV--YGN
GH1-6 PKQVCRSGYGK-GNMAGIAR-SGTADYMCYHVIKHAHAAVHYQDRYK-SLGGKITMALDGVWSIPYWDNEEDRQAERQLNFEFLYAHPIF--FGD
GH1-7 PKQICHYGYGV-GTVAPAIHS-SGLREYWCAYTVLNAHAKTYHMYKS---LNLKAPMGITIDCEWYEPDTSVQDKQAARLNLFECGLYAHPIF--IGD
GH1-8 PKQICHHGIGE-GKFAPGRHN-SGVDDYQCDYVVLKAHAKAYHMYKEEF-PHYKALMSLTIDCEWYEPDTSKEDMNAARRNIDFECGLYSNPVY--NGD
GH1-9 PKQVCRIGYGA-TMLAPALNS-SGLLEYDCAYVLVKSHAAIYRMYKKEF-PDYKAKMSIVIDCQWIEPQTNA PKDKDAERLQFECGLYFNPIFN--NGD
GH1-10 PRAICKNGYGD-GSLAPFSSN-SGLDDYECARVLLKAHAAVYKVFKEEFASKNGAKISIALDASWSIPASNNEADKEAAERLNLFECGLYADPIFK--TGN
GH1-11 PTVICDKDGYGQ-GTFAPGINDGPNYEICGHNLLKAHAAVHYDDYRRCQNGKISITLNSDWYEPQSCCDTDLCASETKLQFQIGWFAHPIY--LGD
GH1-12 P YVWCVGAYDV-GYFAHSETR-PGFWGYTCGHNLLKAHAAVWHMYDKYFRPTQKGTVSLTIAGDWHIPKTSSEIDKAAERLQFQKVGWVHVHPLY--HGN
GH1-13 PSSICNNGYGK-GKFAPGIYN-PGIGEYLCIHNLLNAHAAVHYVYDKEFRSTYNGKVTTVHSISYCPKTNTEDEI EADQKMQFNLGWMADPIY--FGD
GH1-14 PHEVC-----IYD-GAEKVYDCAANLLRAHAKIWHMYDEEYRSKQHGKVMVNLNLYWYEPETDKTDDIIAAETKMQFWSGLFGHALY--HGD
GH1-15 PLTICYGYGT-GAVPPFIQS-EGILEYQCLRNLLKAHARAYHYVAEEFRNKQGGKLSMSFNTNAFLPGSDNPKDIEATERVQQFNLGVYAHPIF--VGD
GH1-16 PQQICSGGGYGY-GYFPPTIKS-EDLLEFQCIHNLVKAHAKAWHIYDDEFRSTQKGVISITLDTSSYIPASNKSEDEI EADRMQHFTLGVYANPIY--VGD
GH1-17 PKTYCNGGYGY-AYNPPQIPS-QGLLEYVCSHNLLRAHAKYRMYQEYFRPQGGKIAIVLDTYWYRGATDSEEDKTAATAALHFDLWYANPVF--NGD
GH1-18 PKQTCNGGYGN-GILAPGIFS-PGVGEYLCIHNLVKAHAAVWHLYDEKYRRRQNGHLGITLDCFFGEPDSDSQEDKDAERLQFNQYGYANAIV--NGN
GH1-19 PKQICNEGYGN-KQKAPLVS-PGIGEYLCIHNLLKAHAKAWHLYDEQFRSTQKGFVGTITDITAWMEPDTSDTEDVDAERLQFNHGWIYARPLL--LGD
GH1-20 PKQTCQEGYGS-GQKAPAIKS-HGIGEYLCIHNLVKSHAKVWHLYDRYRKEQRMVGI VLDTMWMEPDTESELDKVAERLQFIFGWYANPII--NGD
GH1-21 AFYMCTLG YGT-GTQAPGRKA-SGIGDYICGHNLIKAHAKAFHLYNDYRVPVQKGSVGISLVS IYAEPKSDSDEDEKAAERMIYFNFGWFASPIM--FGS
GH1-22 PKQCTNGYGS-GSYAPAIKS-PGVGEYMC IHNLIKAHAAVHYIYDTEFRDKQNGKVGIVDSAWWEP-YNATDSDASERLLHFTFGMYANPIF--YGD
GH1-23 PVLMI DNGYGI-GVIAPGRKT-GGVGEYLCIHNLVIRAHAKVPHKYQEAAYNPQKGI G I SLVTAWNDPRTDKPEDIAANRMLLFTFGWFAHPLV--FGS
GH1-24 ---TALLGYAD-SEFPPYINQ-PGISNYMCIYVTL LAHAKAFRLYHNEF-KQYKQVGI VVDGRWYQPGSDSDKDVEATNRALEFEIGIWLHPLLSPKGD
GH1-25 -----SGIGEYLCIHNLVKAHAKTWHLYDKHFRSTQKGVGTITDITAWLEPDTNSTRDVAERQQFVHGWIYTRPLT--LGD
GH1-26 PKQTCQQGYGS-GLRAPAIKS-HGIGEYLCIHNLLKAHAKAWHMYDDEYRQRQHGFGVGTITLDSMWMELDTSELDKAAAERA IQFTFGWFANPII--KGD
GH1-27 -----
GH1-28 PKQICHQGYGE-GTLAPGLLL-SGLEDYQCDYIVLKAHAKTYNMYKKEF-PHYQAHMSLTIDCEWYQPYTNSKEDI DAARRNLDFECGMYSHPLY--YGD

GH1-1 YPALVKEMVDNRS-ISEGYLQSRMPTLTFQEADYIQGTDFAGLNI FTSYRVKHDA-KGT--ANTTLEK DANVTLHQDTEWEKSGSDWLRVVP I GVRRLV
GH1-2 YPQVMIDSIANIS-KTQGFLESRLPVLTAEDQKLLKGSFDFI GLNNYD TYLVSD---DRKVDVDP SYINDQGVNFTYFPNWTVDYS--IEMRSKGLGIIL
GH1-3 WPEIVKERVAYS-RKLEGFTE SRLPAFTEDEIKYIKGTS DYLG LNHYSTLLTNHTA-DAPI-GTSPFDNDKSVLLWRDWDW TQGSADWFFDVPWGLRKF
GH1-4 WPKLKDRVAMRS-EKEGFKSRLPAWTFQEEIDYIKGTDFVALNHYATHMANGTN-EAPI-GNPSFGSDISAEWARPEWPKGDGDFWFSI VWPWGLRKL
GH1-5 WPQIMIDNVADFS-KKQGFKESRLPAFSEEEI ELIKGTDYDLAVNHYTSLMVKARA-EPTL--IVSWDDDAGVDAYQKSTWETAAGIWFKKVPWGFGLL
GH1-6 WPQVVKDRVNYRS-KMENY P ESRLPEFTAEMKYINRTADYVAFNFYNTKLIKIDID-EASF-DITSFNDL RVKDDVDPRWTI-AMDGNTIYPQGLRSYL
GH1-7 WPEDVKMRIYERS-MAENYNTSRLPELTLAQSYIYGTDFEGLNYYVTFVTADGK-EGPT-NVTSYDNDVRI VLG NVSN-ADIDVTGFP I VPLGLTKVL
GH1-8 WPSSVKERVKLRS-QLQGYNRSRLPEFTTEEIN YIKGTS DLYLLNVYFAYLVKDVP-EAPS-NVTSFRSDVKAE LIRFPG-TSVGINGMPIVPWSVENVL
GH1-9 WPKVVKDRVDFRS-LKANLTT SRLRAFTEEEKIYKGTDFLGNHYTMLARDEAE EAPF-NETNYENDVGTVDSD FDKWVVE SNGLFVI VPGGVFKML
GH1-10 WPQVIDRTDLRS-KGEQAQSR LPIFNQTEKDAIKGTFDFMAIEFFDTKLVADDT-EAKF-ETSSYDNDL RVKTSNDPNWTV-DADEHAIAPGSLRNL
GH1-11 YPEVMIERIRLNC-DYNGY NARLPVFSHAEI EYIKGTHDFSLAYDSAFYVTACV-DNYW-GNICYESDVGVLV GCEG-----EFLTCGIRSL
GH1-12 YPSAMIEIKNISINEGRNSSRLPEFTTEEIEYIKGTNDV FYLNTYH VYHVS NVI-KK---LSPSFDNDIRGTIEEIPDG-----AAWGIRSFI
GH1-13 YPEIMKTRIAIRS-KLEGFIESRLPEFTTEE QERLKN TSDYFAVNSYDSYFAFAIP-EPAIVLPPSLDADSGVGGTKEG-----IGDGSFGMGKLL
GH1-14 WPPIMKSRIGLRS-KLEGYPKSR LPEFTKEEIEYIKGTNDFLAVNSYTT SII RAID-EPEI-GDPTLEKDIGVYEQRDWDSSGSIWLVKVPWGM RKL
GH1-15 WPEIVKTRVAIRS-KLEGRNESRLPAFTPEEIDYIKGTADFLGLN TYTAQLVLDTEGEPPLGDHPTKEHDALITTFQDDKWE D TNVPTVKVVPQSIKPL
GH1-16 YPEIVKTRVAERS-KAEGREQSRLPFTTQEEIEYIKGTHDFGLNVTYGLSLVADLGYDPPVQNPSPSKWQDVAVNNYYPPEWENTTILYTKVPW SARYLV
GH1-17 YPEIMKSRIAARS-AAEGRAQSRLPEFTDAEKAALKGSADYFALNTYTG NLVKSIP-DPPIFNIPSRYQDVGLTYTDPKWESSSLDWLKVAPWGI VDLL
GH1-18 YPKVMRNVIARRS-AAQGLSKSRLPDFTSKEIQFIKGSFDFLGLNYYTVNMVKARG-SEVP-VVSSWEKDMEVDIYLKEEWPKSSSWLRIVPWGM RKM
GH1-19 YPEAMKKTIAERS-ALQGFQSRLPEFTQDEIEYLN GTVDYLG LNYTTVMATNAA-DKRI-DVVSWEADVEVNTYQKEEWP TASSWLRVVPWGLRKT
GH1-20 YPAQMKQFI AKRS-IAQGFQSR L PQTPEEVKYIKGSDIFLGVNYSMLVRNMD-KYK-TEVAWEADSETYIYQPDNWESTATD WLKVPWALRKL
GH1-21 YPYQMEITINERS-VVQGFQSR LPSFTEEEKAYIAGTVDYFGVNYSSVAENLP-DFKD-LDVSWEADVEVIPSQDPHWENTTTPQFVYDPGLYKLL
GH1-22 YPEVVKSSVARRS-QSQGFPSRLPEFTAEEQNTIKGTIDFLGLNQYSSFYVTDKT-DDK--DGMGYDADA EVK TWADES WQQGAADWLNKAPFGIRNML
GH1-23 YPQEMERTVSNRS-DAQGYSKSRHTGFTTEESAYIKGTVDYFGINYYT SVLISHYD-DLHQ-NESGWAADSQVIYQYPPDWEDTVVNMFKVYPHGIYKLV
GH1-24 FPEIVKERIAAIS-MAEGYRQSRLPKFSTTDIDLKGSVDFVGLNLYTSHLVEHVE-DEHF-KGTSTIRDEGIRLYQDASWKGQS T WLVKHPGARKVL
GH1-25 YPELMKKTIAKRS-SLQGFQSRLPQFTKDEIEYLGKTLDYLG LNYTTFMARDSD-DEKI-DDISFEADAQISAYQKDEWPKSATPWL RVVPWGLRKL
GH1-26 YPWQMKQFIARRS-LAQGFSESRLPEFTAEEIKYIKGTIDYLG VNYTAGLVRHVE-DKNR-TKISWEADQET YNYQPEHWEKTAANWLKSV PWSMRKLV
GH1-27 -----LIRFPG-TSVGINGFP IVPWSIENML
GH1-28 WPPSVKARVKFRS-ELEGYNKSR LPEFTTEEIN YI-----

GH1-1 KWKVKDYD-DPEIFITNGFPD-KGE---INDVKRIQYLQKYLKEVSI SIQEDGVKLGKGTVWSFLDSFQWT-----EGYREKFGLYHVNFSKDKNRKRT
GH1-2 NWIKKTYN-DPEIRILNQGFPE-TST--DLYDHNHRKYIVASLQQLWKS INEDNVNVTAYTFWSLMDNFEWD-----KGYSVRYGLNYVDFNDPNRIRI
GH1-3 NWLKNTYG-NPEIVITNGFAD-TNG--TLEDDNRITYLEGHVSACLDAIYEDNVNLTAYTFWS IMDDWEWT-----GGYTSFLGMYKVD FNDPNRARI
GH1-4 VWLKDTYN-DQEI IITNGLSD-NTG--VLEDDHRVDYFRDYISNCLDAIYEDNVNLTSTYIAWS IIDDWEWG-----GGYKSFLGMYKVD FNDPERPRI
GH1-5 RWIKKTYG-DIEIVISNGVSD-RSG--TLQDEHRIKFIQAYMSHMLDAIH-DGVNVTAYTLWS IIDNFWEWT-----QGFNGKLG IYVNVQSDPSP LPRI
GH1-6 KWI SENYN-SPEI IITNGIAD-NGT--SLEDSE R ISYLS DY LNAVLD SIYEDKVNVTGYT MWSL LDNF EWT-----SGYSMRFGLYS VDFEDSNRTRA
GH1-7 NHIKTEYN-VPKILITNGISD-NG--TLEDESRIKTLKNYFSAILD AVYDHDVNVIGVTIWSLMDNLEWL-----KGYSAHFGLYS VDFTD DNRTRT
GH1-8 KFIKDEYK-DPIILISVGTSE-DGT--SLEDDMRIDFYEKAFNAMLEAMYNVDVKGITVTVWSL LDNLEWT-----DGYAAHFGLYHVD FNHPNRTRT
GH1-9 KWIKQNYG-DQDI IITNGMSD-NGT--YLNDKDRIDFYTEYFCNILEAM--DEVKVKGI ILSLMDNFEWN-----SGYTAHFVGYHVD FQPNRTRT
GH1-10 KYIRDNYN-NPEILITASGISD-DNG--TYNDDHRVAFLADQLSDVADAINEDKVN VFYTYWSL L DSFEWT-----KGMKPHFGLYAVNL-ELETSRV
GH1-11 GWISRNYY-NPSVFITNDGYAS-CSR--DCYDCDRIEYIKAILRNVRLSMIEDYTRVYGYTYYSYIDGF EWNLV DGR LGGYTLKYGLYDVEFD CPQKRTRT
GH1-12 SWIRKNCN-NVPIFITNGKGSTDYT--GLQDDYRIDFMQKMFSSIRDSMEIDNATILGYTFWSLYDGFVQ-----YGYRKKMGLAYIDFNSNKRTRI
GH1-13 LWIKDRYN-NPDMIITNGLMT-QTE--TLEDDDRISMVKNCLSRFRDVM--DQVNL LGYTLWSVMDNF EWT-----RGYTAKYGIYHVD FNNSNRTRT
GH1-14 NWFKQEYF-NPEILITNGYS D-FTG--QLNDPNRTHYLTEYLSA IKDAMDYDGVNVI GYNVWSLIDNF EWN-----AGYTVKFGLVNVDMNDPNRTRT
GH1-15 KWIKDQYN-NPGVFITNGYPS-PAGDMY LKDDRRVHYIESYLR Y IRESMEEDEVSM LGYTVWSLMDNF EWI-----FGYTVKYGLYD VDFTS PDLTRA
GH1-16 NWIKDKYN-NPGIIVTNGYPS-YGD---KDDRRIYIRGYSALRDAMEKDNVSI LGYTVWSIMDNFEWT-----SGYGQKFGLYEVD FSSSNRTRT
GH1-17 NFIKNNYN-DPDI MVTNGLCD-KGG--LDDPRRVNYHKLYLSYIHDAMINQVKVFGYTAWSLMDNF EWA-----SGYTQKFGLYS VDFSSPNRTRT
GH1-18 NWMKKTYGDDTEIFITNGVSD-DGS--SLDDQIRIDYQQHISGVRDALD-DGVNVI GYTAWSLMDNF EWI-----RGYSECFGLYQVNF TDADRPT
GH1-19 NWIKNTYG-DIPI IITNGISE-DGT--SLEDDTRIDYQQHLSAVKDAMD-DGVNVI GF TAWSLMDNF EWL-----RGYSEDFGLYHVD FSSPD RTRT
GH1-20 NWIKQTYG-NIPQIITNGYS D-NGT--TLEDDGRINYYKQHL SNIRDAID-DGVNVI GYTTWSLMDNF EWF-----RGYSEKFGLYHVD FTS PQRTRT
GH1-21 KWISNNYN-NISIMITNGYPD-KGE--TLQDDERISYIQRHLSKVRDALD-EGVQVLGYTFWSLMDNF EWL-----MGYDTKYGLFEVD FNSPDRTRT
GH1-22 KWIKKTYP-DIPVI IITNGWSD-FDG--TLDDPRIDYISSYLSNVKASMEDEKVN VVGYTVWSLIDNLEWL-----FGFTNKFGLYRVD FNDPNRTRT
GH1-23 KWVSKTYN-NIPI IITNGYPD-DGS--TLDDNKRI SYIKRHLSKLRDALD-EGVNLIGYTYVWSFMDNF EWL-----SGYSVKFGVI AVDFNSPNRTRT
GH1-24 NWWKETYN-NPVVFITNGFSD-DGE--IEDNNRIEYLQGYLTN I LKAIYEDKVDVKGYAVWSLMDNF-----
GH1-25 -----
GH1-26 NWISKTYG-NIPQFITNGYAD-NGT--TLEDDRRIHYYQETLSNLRDAMD-DGANVIGYTAWSLMDNF-----
GH1-27 KFIKDEYK-DPILVITNGTSE-DGS--TLEDDMRIDYLEKTYNAILEAMYNVDVKGVTVWSL LDNLEWM-----DGYTARFGLYHVD FNDPNRTRT
GH1-28 -----

GH1-1 PKLSATWYKKVIKERRV-----DWKEVTSNLEEDL*-----
GH1-2 PRLSAITYNHVSTKKT LR-----GLL*
GH1-3 KRKSADYFTNVVKNRCLV-----EESQCV*-----
GH1-4 KRKSADYFTQIVKNRCLV-----DADKCV*-----
GH1-5 AKDSSKYANVIKTKCLV-----DSCVDN*-----
GH1-6 AKTSVDYYSQIIKNRKIPE*-----
GH1-7 AKASAEYKKNVKTKEIDNSHSD-----GGQSITKVN GYVVGFIIVS FYFFK L W*
GH1-8 PKKSATYYKKNVINTRLLRAD-----NGSNINKISGFMTIVLLFPILWNTLM*
GH1-9 PKESAGFVKKLSKDKT LSCGNP SKYRSVPMALPKWTPANANVSIWNNGTGNRNHTNP GTTPTPLRNNNNGNGNGNRGKNKDC*
GH1-10 EKKSQAQYALAIKERKVPDADVETTTTSTTTEGH-----QGGAGNRYQPVPVIF I I SLISYVLQ*
GH1-11 ARSSAAYYQHLASTGCIDNPC-----PIPYCYDYN DYHY*-----
GH1-12 PRKSAYYSNVSR TGCSISNRCVYVSSCD-----GYTFPMSIMTAITVHILAILTL CRS*
GH1-13 PKKSAEWYKHVIQTRCLT-----DICVD*-----
GH1-14 KKDSFYYYQKVCQTHCIV-----DQCID*-----
GH1-15 AKPSAQYYKEVILTRCAT-----IVGCFND*-----
GH1-16 PRPSAMFYKNVIATRCLV-----ETCDD*-----
GH1-17 IKQSGKFYQNVCHTRCLL-----DEDQCADY*-----
GH1-18 PKRSVEFLKVMTRCVT-----DVCE DY*-----
GH1-19 PKKSVEYLKNVIQYRCVL-----GNSTCDD*-----
GH1-20 PKKSVS YFKHVAETKCI V-----DVKDCVIDVDEIVRYIYTNVDKNVLA*
GH1-21 PKKSADYVKKVYESRCLLE-----DSSECYD*-----
GH1-22 AKKSAKWYTNVIENRCLG-----TCE*-----
GH1-23 PRKS AKYLKNVYATRCIME-----DNSQCSD*-----
GH1-24 -----
GH1-25 -----
GH1-26 -----
GH1-27 PKKSVGYYQYVASTKRLLSAD-----NGSNLINISGFMTIVLLFPILWNTLM*
GH1-28 -----

B

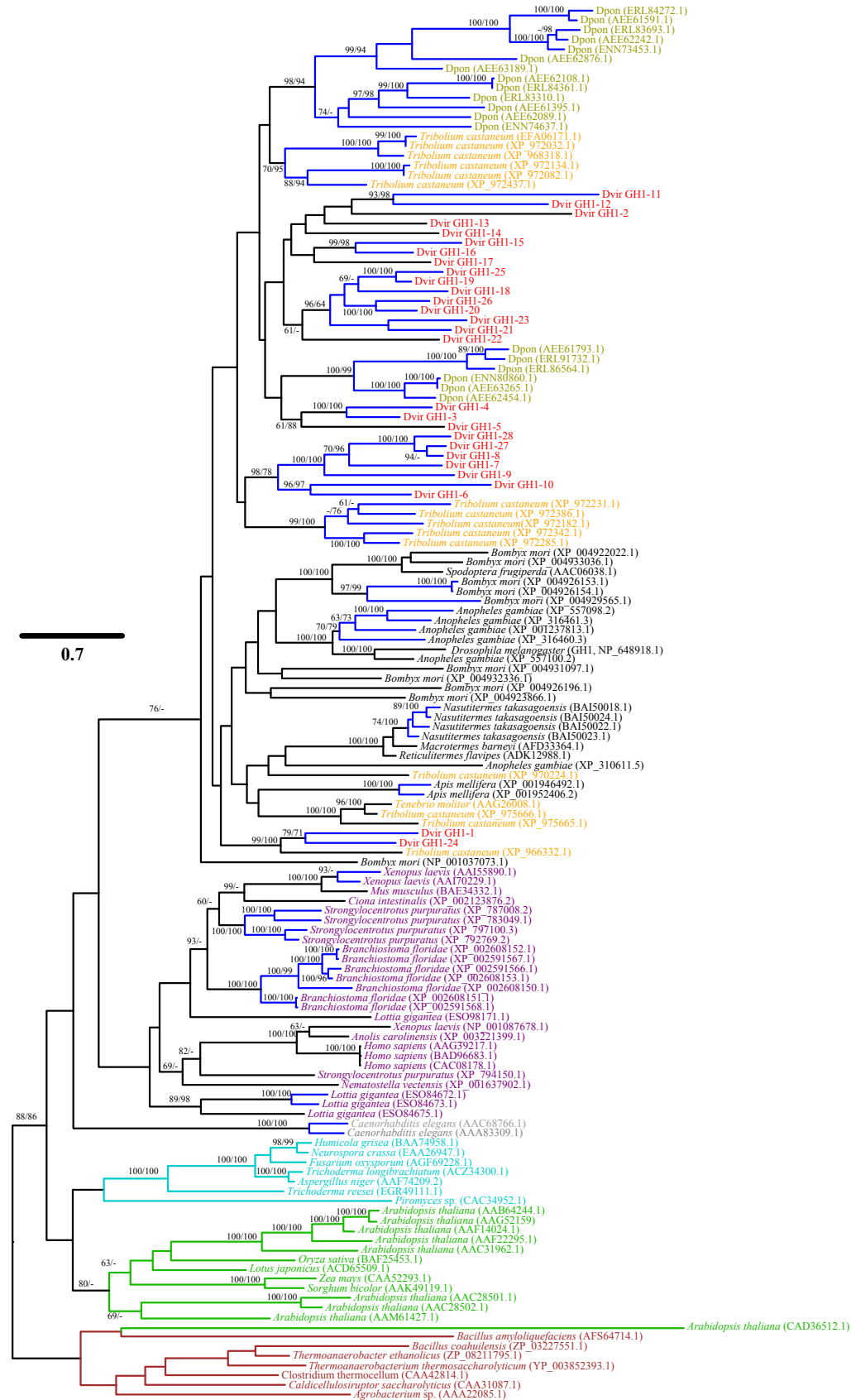


Figure S5. GH1 family gene sequences identified from the *D. v. virgifera* transcriptome (A) and the maximum-likelihood phylogeny of GH1 family proteins (B). In the alignment, the labels for partial sequences are shown in italics. Potential residues for the catalytic nucleophile and the proton are highlighted with magenta and green, respectively (based on Marana *et al.*, 2001, *Biochim Biophys Acta* **1545**: 41; Scharf *et al.* 2010, *Insect Biochem Mol Biol* **40**: 611). In the phylogeny, labels for the coleopteran species belonging to the superfamily Curculionoidea, *D. v. virgifera*, and other beetle sequences are shown in olive, red, and orange, respectively. Their species abbreviations are found in Table S5. Arthropod, other metazoan, nematode, fungal, plant, and bacterial sequences are indicated by black, purple, grey, cyan, green, and brown, respectively. The numbers at internal branches show the bootstrap support values (%) for the maximum-likelihood and neighbor-joining phylogenies in this order. Supporting values are shown only when higher than 60%. The scale bar represents the number of amino acid substitutions per site.