Reconstructing evolutionary history

Can I infer the past from present day observations?

Assumption 1: Genes are passed on vertically from ancestor to offspring

Assumption 2: A gene is not invented twice

Analysis

In this exercise, we will extract the sequences of a subset of orthologs for two PCD gene families from our interactive webtool. We will align these sequences and calculate a phylogenetic tree. This tree will then help us to better understand the evolutionary history of the gene family

Task 1: GH27_QRW17738.1

1. Extract the sequences of GH27_QRW17738.1 orthologs for the following taxa:

   GH27_QRW17738.1_taxa.txt

   ```
   Rhizoctonia solani
   Aspergillus oryzae
   Trichomonas vaginalis
   Cordyceps militaris
   Bradysia coprophila
   Hydra vulgaris
   Notolabrus celidotus
   Homo sapiens
   ```

2. To get the sequences click on the dot in the columns of the respective species in the GH27_QRW17738.1 row (use the search boxes on the left to find them!).
   - In the bottom left corner, you can click on “Detailed plot”, opening up a window with more information that also contains the sequence.
   - If no sequence is shown, look at the barcharts on the top of the page. Click on the highest barchart to select one gene from the available options. This will give you the sequence
   - Save the sequences in a FASTA file (use the species name in the FASTA header!)
3. Align the sequences using ClustalOmega

4. Inspect the distance tree generated by ClustalOmega. What can you learn?

GH27_QRW17738.fa

```
>GH27_QRW17738.1|RHISO@456999@016906535_1|XP_043177975.1|1
MIGIPVATFIFLGCAIGIVDAQQPLYTQCGGMVGRSPHHYECCASAIGCWTQSPVGKTPALGWNSWNAYQG
NINEAKIRAAADALVSLGLKDAGYNYVNIDDCWSNITGRDVGTSRIKPDLNKFPNGIKSVADYIHSLGLLV
GIYSDAGTLTCAGFPGSLGYENIDAQTFAEWGIDYLKYDNCNVPSNWTDTYGGPDFDWYNSNSGIRYRQMT
SALASVASIRPIQYGICNWGNADVWTWGSRVGHSWRMSGDSSASWSYITSIIARNTNYLNYVEFYGHNDMD
MMEIGNGALTINEQRTHFAMWVALKSPILLGTDLSKLNATQLAIIKNAELLAFSQDTTVGAPAKPYKSGTT
TPPEYYTGKSAKGTHVFIMNVGSAPASKIVTFSEVTGLTCTSCKVHDMWTGKGKCIFSHTTTIVFNCGLL
>GH27_QRW17738.1|ASPOR@510516@000184455_2|XP_001820370.1|1
MQRYISLSVSLSLLSGANALVRPDGVGRLPALGWNTWNAFGCDIDASKVLTAAEETINLGLKDAGYEYINI
DDCWSVKSGRDPNTKRIIPDSAKFPDGISGVASKIHDLGLKVGIYSSAGTETCAGYPASLGYEKIDAESFA
EWGIDYLKYDNCGVPTNWTDTYTHCVPDNSNGSKFPNGTCPDISNPAPTAYDWSSSNTAQRYNAMRDALLG
VNRTILYSLCEWGQADVNTWGNGTGNSWRTTGDITPDWSRIVEIANENSFLMNYADFWGYPDPDMLEVGNG
NLTLEENRAHFALWAAMKSPLIIGTALDSINEEHLAILKNKPLLSFHQDPVIGRPAYPYKWGYNPDWTFDP
AHPAEYWSGPSSTLGGTLVLMFNSEDSAKHRTAVWSEIPELKDSAEKGSGYRVTEIWTGEDLGCVKDQYDV
ELQSHDIAALVVGESC
>GH27_QRW17738.1|TRIVA@412133@000002825_2|XP_001315888.1|1
MLISFLQLATSTEYETPTMGWSSWNTYRVHISEELIKKQATAMAEKGLKEKGYKYINIDDGYFGGRNETTG
QLLINPERFPNGLEPVVTHIHNLGFKAGIYSDGGRNTCGNYYDADVLGKGVGLYAHDDQDADLFFNTLQFD
FIKLDYCGGDPNQNSERLNLDEQTRYTSIRKALNKITRSIRVNVCRWAYPGSWVTKLATSWRISSDIENNF
DAIKRLINVNMFYSAYAVEGHYNDMDMLEVGRGMSDDEDKTHFGMWCIMSSPLLIGCDMTTIKDSTLELLK
NEELIALNQDPLALQAQLVETQSNTFIFVKDILEYHSTTRAFAVYNPSDSEQTVNINFESLGLSGNIKVRD
VYAKSDVGTYPSSQPFSVKVNKHGTRIFTISGETRLEQVRYEAENSWIERGSAITSAQVARVDSDSLCSLG
HKVTYIGGNENDNYIEWRDVYSQNGGKYNMKIAFLSGEDRILRVIVNGKTNVLKLNAGSWTAVGTKTIEIE
LQKGSNVIRMDNPISFAPDIDYFEITPV
>GH27_QRW17738.1|CORMI@983644@000225605_1|XP_006673409.1|1
MGFLAALSSASPYSSILDLASTMKSQLVLASAAGLATSTHVDLPLPPMGFNNWARFTTHINQSIFTDAATS
MAANGMLAAGYNRLNLDDAWSTHERAANGSMVVDSVKFPQGLSWLAAFIKDKGFIPGIYTDSGTLSCGQYP
AAYGHEAIDLQDFSDWGYEYLKLDGCNVPGEDEAAYRRVYFDVWRELLLTFPKPVVFSNSAPAYFCGADNL
TDWYSIMGWAQQMGQLARHSYDIINYDSKGSAWDSLMTNYNQHVRLARYQRPNFFNDPDFLNTDHPNYTLE
EKKTHMALWSSFSAPLIVSADIPNLKKEELEILLNKDLIAIDQDPLVQQATLVSSSPDWDVLSKNLANGDR
VLTILNKAGKAGDYTVPWARIGIQTDRISPDTALSVKDLWTGKTEQVKISAGGITAKNVGPHGTFVYRISG
RMGGSSHPIVLPTGLIFNTFSLKCLSDDKSGKVAWKTCDGSDGQTWTVSDEGRVHSLLRPNECLADVQGKL
LSRHSGCRTDAWKYYNSGNLVNGNSNRCLTEGAAGAASVSDCGYATNEQVIGLPVGAKVIEVDYYANK
>GH27_QRW17738.1|BRACO@38358@014529535_1|XP_037033872.1|1
MFSASVIIVCVLVTSQVYALENGLARTPPMGWMTWQRFRCTTDCNIFPNDCISENLIKRTVDALVRDGYRD
AGYEYVIIDDCWPEMERSSEGLLVPDKNRFPSGMKALSDYIHARGLKFGIYQDYGTKTCAGYPGIIGYQEL
DAILFASWGVDYVKLDGCYSEPESMDEGYVSFGQLLNKTGRPMVYSCSWPVYQEEKGIEPNFTLISEHCNL
WRNWDDINDSWSSLASIADYFAKNQERIQRFAGPGHWNDPDMLLIGNFGLSLEQSKVQMSLWAILAAPLLM
SNDLDTIRPEFKDILLNREILSINQDPLGIQGTRVKVIDNIEIWTRPILPRTDSSHSFAVAYVSLRTDGAP
HRIQLNLVDIGLTNLGGYVEMDVFEDTVSPTVKYPETLSSVIVNPTSCVLHKFVAL
>GH27_QRW17738.1|BRACO@38358@014529535_1|XP_037037659.1|0
MIYPRSPPHIVLIILSLLAVAYHINGLDNGLARTPPMGWLSWERFRCNTDCEGDPDNCISENLFRTMADLV
VSEGYAALGYEYINVDDCWLEKSRGPQGQLVADRKRFPRGMRSLSDYVHSKGLKFGIYEDYGNYTCAGYPG
IIGYGRSDAQQFADWNVDYVKLDGCYSLPLDMDQGYPEFGRHLNSTGRQMIYSCSWPVYQIYAGISPNFSS
IIEHCNLWRNYDDIQDSWASLETIIDYYGNNQDAIVPNAGPGHWNDPDMLIIGNFGLSYEQSKTQFALWAI
MAAPLLMSVDLRTIRPEFKAILQNQKIIAVDQDPLGIQGRRIYKHKGIEIWSRPITPIYQTYYSYAVAFVN
```

```
RRTDGTPSDVAVTLRELGLQSPTGYRVEDLYEDVNYGILSPNTKIKVKVNPSGVVILSAHVQPERITRTTY
NPLNYYHVK
>GH27_QRW17738.1|BRACO@38358@014529535_1|XP_037037657.1|0
MIMIYPRSPPHIVLIILSLLAVAYHINGLDNGLARTPPMGWLSWERFRCNTDCEGDPDNCISENLFRTMAD
LVVSEGYAALGYEYINVDDCWLEKSRGPQGQLVADRKRFPRGMRSLSDYVHSKGLKFGIYEDYGNYTCAGY
PGIIGYGRSDAQQFADWNVDYVKLDGCYSLPLDMDQGYPEFGRHLNSTGRQMIYSCSWPVYQIYAGISPNF
SSIIEHCNLWRNYDDIQDSWASLETIIDYYGNNQDAIVPNAGPGHWNDPDMLIIGNFGLSYEQSKTQFALW
AIMAAPLLMSVDLRTIRPEFKAILQNQKIIAVDQDPLGIQGRRIYKHKGIEIWSRPITPIYQTYYSYAVAF
VNRRTDGTPSDVAVTLRELGLQSPTGYRVEDLYEDVNYGILSPNTKIKVKVNPSGVVILSAHVQPERITRT
TYNPLNYYHVK
>GH27_QRW17738.1|BRACO@38358@014529535_1|XP_037050710.1|0
MWSKILSFVLLLWQQNKCLVDALDNGLALTPPMGWMSWQRFRCIIDCDLYPDECISERLFKTAADVIASEG
YKDVGYEYVIIDDCWLESERDNVTKELVPDRKRFPNGMKALADYVHSKGLKFGIYENYGAKTCEGYPGIID
HMELDAKTFAKWDVDYIKVDGCYADYLDYDVGYPEFGRLLNKTGRPMVYSCSWPVYQEIKGLTPDYEALKQ
HCNLWRNWKDIQDSYESLRNITEYFAVNQSRIQPHAGRGHWNDADMLLIGNYGLSIDQSKTQMAIWAVIAS
PLIMSNDLKNIRPEFKEILLNKEVIDVSQDELGVQGIRKRFEARIETWMRPIKPIVNGELSYAIAWSSKRI
DGAPFTVSANVGSLKLTNANGYRVTDLFNKNRFVMDILPAQNITVRVNPSGVVFYKFTAK
>GH27_QRW17738.1|HYDVU@6087@000004095_1|XP_002159485.2|1
MLYLLIISQIVGLYTLDNGLARTPPMGWLQWERFRCIIDCETYPDQCISENLFKRTADKLVSDGYAKLGYK
YVNIDDCWPELQRDLDGNLVADKKRFPNGIKALADYIHNHGLKMGIYTDFGTLTCGGYPASIFHMQKDTQI
FSNWGVDMVKLDGCYACESLYADGYKAFGWYMNQTERPMLYSCSWPAYVEKNKFDXLLIGDYSLSESQSKV
QFALWSILAAPLFISADLDTMKQWQKDILLNEEIIAVNQDVLGKMGRRVYKDNNQQNWVRYLDNNAAAVAI
MSLRTDIPVYMEVTFSSINMKGPFHVRDLFLHEDLGVFNTSFKAIVDPVGVVMVKLTPMA
>GH27_QRW17738.1|NOTCE@1203425@009762535_1|XP_034555657.1|1
MHLAVFLLASTLSAITFALDNGEMRTPPMGWLAWERFRCDIDCKNDPKNCISENLFIDMADRLAEDGWREL
GYTYVNIDDCWASMERDENGRLQADPKRFPGGIPRLSRYMHDRGLKLGIYGDMGTHTCGGYPGTPLDKIEV
DAQTFADWEVDMLKFDGCYSNASEQEKGYPLMSKALNATGRPIGYSCSWPAYQGGLPPKVNYTQLGEICNL
WRNYGDIQDSWGSVLSIADWFFENQDILIPAAGPGRWNDPDMLVVGDFGLSMDQSRSQMALWAIMAAPLFM
SNDLRTISSGARNILQNKMIISINQDPMGIQGRRIVKEKSHIEVFWRSLSNGASALVFFSRRTDMPYRYQT
SLSKLNYTTGSYKIYDVYTDQTTMLKDTTNFVVSVNPTGVVMWYVSAPAKLHERHFYKSGNLMESVHRGDE
NTIPRVFL
>GH27_QRW17738.1|HUMAN@9606@000001405_39|NP_000160.1|1
MQLRNPELHLGCALALRFLALVSWDIPGARALDNGLARTPTMGWLHWERFMCNLDCQEEPDSCISEKLFME
MAELMVSEGWKDAGYEYLCIDDCWMAPQRDSEGRLQADPQRFPHGIRQLANYVHSKGLKLGIYADVGNKTC
AGFPGSFGYYDIDAQTFADWGVDLLKFDGCYCDSLENLADGYKHMSLALNRTGRSIVYSCEWPLYMWPFQK
PNYTEIRQYCNHWRNFADIDDSWKSIKSILDWTSFNQERIVDVAGPGGWNDPDMLVIGNFGLSWNQQVTQM
ALWAIMAAPLFMSNDLRHISPQAKALLQDKDVIAINQDPLGKQGYQLRQGDNFEVWERPLSGLAWAVAMIN
RQEIGGPRSYTIAVASLGKGVACNPACFITQLLPVKRKLGFYEWTSRLRSHINPTGTVLLQLENTMQMSLK
DLL
>GH27_QRW17738.1|HUMAN@9606@000001405_39|NP_000253.1|0
MLLKTVLLLGHVAQVLMLDNGLLQTPPMGWLAWERFRCNINCDEDPKNCISEQLFMEMADRMAQDGWRDMG
YTYLNIDDCWIGGRDASGRLMPDPKRFPHGIPFLADYVHSLGLKLGIYADMGNFTCMGYPGTTLDKVVQDA
QTFAEWKVDMLKLDGCFSTPEERAQGYPKMAAALNATGRPIAFSCSWPAYEGGLPPRVNYSLLADICNLWR
NYDDIQDSWWSVLSILNWFVEHQDILQPVAGPGHWNDPDMLLIGNFGLSLEQSRAQMALWTVLAAPLLMST
DLRTISAQNMDILQNPLMIKINQDPLGIQGRRIHKEKSLIEVYMRPLSNKASALVFFSCRTDMPYRYHSSL
GQLNFTGSVIYEAQDVYSGDIISGLRDETNFTVIINPSGVVMWYLYPIKNLEMSQQ
>GH27_QRW17738.1|HUMAN@9606@000001405_39|NP_001349777.1|0
MLLKTVLLLGHVAQVLMLDNGLLQTPPMGWLAWERFRCNINCDEDPKNCISEQLFMEMADRMAQDGWRDMG
YTYLNIDDCWIGGRDASGRLMPDPKRFPHGIPFLADYVHSLGLKLGIYADMGNFTCMGYPGTTLDKVVQDA
QTFAEWKVDMLKLDGCFSTPEERAQGYPKMAAALNATGRPIAFSCSWPAYEGGLPPRVNYSLLADICNLWR
NYDDIQDSWWSVLSILNWFVEHQDILQPVAGPGHWNDPDMLLIGNFGLSLEQSRAQMALWTVLAAPLLMST
```

```
DLRTISAQNMDILQNPLMIKINQDPLGIQGRRIHKEKSLIEVYMRPLSNKASALVFFSCRTDMPYRYHSSL
GQLNFTGSVIYEAQDVYSGDIISGLRDETNFTVIINPSGVVMWYLYPIKNLEMSQQ
>GH27_QRW17738.1|HUMAN@9606@000001405_39|NP_001349779.1|0
MLLKTVLLLGHVAQVLMLDNGLLQTPPMGWLAWERFRCNINCDEDPKNCISEQLFMEMADRMAQDGWRDMG
YTYLNIDDCWIGGRDASGRLMPDPKRFPHGIPFLADYVHSLGLKLGIYADMGNFTCMGYPGTTLDKVVQDA
QTFAEWKVDMLKLDGCFSTPEERAQGYPKMAAALNATGRPIAFSCSWPAYEGGLPPRVNYSLLADICNLWR
NYDDIQDSWWSVLSILNWFVEHQDILQPVAGPGHWNDPDMLLIGNFGLSLEQSRAQMALWTVLAAPLLMST
DLRTISAQNMDILQNPLMIKINQDPLGIQGRRIHKEKSLIEVYMRPLSNKASALVFFSCRTDMPYRYHSSL
GQLNFTGSVIYEAQDVYSGDIISGLRDETNFTVIINPSGVVMWYLYPIKNLEMSQQ
```

Task 2: GH28_QRW23661.1

1. Repeat the analysis using GH28_QRW23661.1 and the following taxon set:

   GH28_QRW23661.1_taxa.txt

   ```
   Spodoptera litura
   Rhizoctonia solani
   Rhagoletis zephyria
   Bradysia coprophila (XP_037032572.1)
   Protomyces lactucae-debilis
   Pyricularia oryzae 70-15
   Arcicela rosea (WP_367284634.1)
   ```

```
>WP_367284634.1 glycoside hydrolase family 28 protein [Arcicella rosea]
MKKRISLLIAISSAMFCNSLVAQKTPTYSWKNLPKIVQPTFAKDTFNVTKFGAKPDGITLNTQAINDAIT
ACSKKGGGVVLVPNGMWLTGPIVLKSNVNLHIRKAATLLFTEDKSQYPLVEGSYEGKSAARNQSPISATN
QENIAITGQGIIDGSGDVWRAVNKAQLTESEWNAKKASGGVLKANGLTWYPSEQFMKASVENRSMLLKDG
VSLQSFADMKDFLRPNLLVITKCKKVLLEGVTFQNSPAWCLHPLMSENLTLRNLTIKNPEYAHNGDGMDI
ESCKNFLVDGCTIDVGDDAICIKSGKDEEGRKRGMPTENGIIRNCIVYNGHGGFVVGSEMSGGARNIFVY
DCTFAGTDKGLRFKSVRGRGGIVENIYAKNIFMKDIAQEAIFFDMYYFVKFATDSPRDERPVVNEGTPIF
RNMKFENIVCKGANKGIFVRGLSEMPIQNIQMSNIVLDTKIGAEFIDASNITLEKVTLISENTKPVISVN
NSDGLTFNTIQYKANAELLFAIAGERSKAIQILQTDSSKAQKQIEFTEGATKEAITVSTGK
>GH28_QRW23661.1|PYROR@242507@000002495_2|XP_003719232.1|1
MKLSSLLAMLGVATTATAFMPADRPRNAQEFRAKHPVARRADSGCRRRFTPRASTHDLDDVSAEFEQAVRDANNG
GTVHLPKDQLFVIGKPLDLTFLNDIHVKLEGTIRFTNDTPYWQANAFYHPFQRSLMFWKWGGKDIKIYGEGVLDG
NGQRWWNEFSGLEILDPDNPYLRPVLFYAENTTNLHVEGIFMKDSPVWHNFVVTSKDVTYKDVIIEAISNNATSP
PKNSDFFNSLNVDGVTVERVWVNIGDDCFSPKSNTSNVHVNTMYCNGTHGQSMGSLGQYEGEVSIVENVLIENVA
LLNGDNGARLKIWAGESVGSGWIRNVTFRNFYAANLDFVARLDSCYFNIPSETCNRFPSKMSIQDITFENFSGTT
SGKNGDAVARLTCSTSPDAVCENVVFKNFNITSPCGGAPVVICDGITGLQHGCVAFDSAEAKAAMDNKCKAPVAS
IEPPWPVRDWRNSK
>GH28_QRW23661.1|RHISO@456999@016906535_1|XP_043183898.1|1
MLSAILAVALGAAVALASKTGVKTQTCIVPSHGNVNISDTPAVHATFKKCGKGGHIIFSENTNYTLRELTTMTPC
IGCTVQLEGTIQMADNITYWLKNETTNTPNITAETFPHLVYYPFQDTVAYLILKDWSHSTLVSKTGKGLIDGLGQ
LWWDAAVGQQILLPGTLRRPVLFTLDGANNVTVDNVTMRNPANWFNWVTDSSNVVYKNIRLSALSANKNPPANAD
```

```
GWDTYRTSHFELRDSHIVSGDDCFAFKPNSTYITIENVYCQNSHGVSVGSLAQYPGVLDIVEHVKVKNVTFVGNG
DSSSNGARIKIWSGPVGSAIVNDIHYEDLRVENVTNPLVVDSCYFSSAYCATGKPVATITNVTVTNITGTSTGKV
VSSIICPEGSTCDIKLKNVNIVPKTGVAPVYRCFSVASEDLGVNCTYPTIVNGAFKWPA
>GH28_QRW23661.1|PROLA@2754530@002105105_1|XP_040722516.1|1
MLPQHCAIAAAWLQLLSLGYATAQGLSISSSHLLTSNAAVKPNPTEANVAPKVCSIKANMNGISKDLLQAAQACK
SNGHIIIEAGDSLIDNVVVMADLKDVTISIQGTLHLKADPVFWAQNAYKHDLVNFQNSSAAILFRDCEGLKFGGA
EVFKTGSTIIGYGAPFWSAYLADNTIMRPNLVTLQRCTNCEISNLRFLDTPKWAMYVTDSDHVNIHDMMIQSIGG
SVMAMNTDGIDIRNSTNVEFHNNYIDNTDDCVAIKGMCSNIYVHDIICGSKTAGVAIGSLGNIVGVNEFVKDVVI
QDVLISGTPRGGIRIKIWPGNRVKAKELLGGGGLADVHNITASRIHTSDNSQALYIDTCYSLGRTNSNCYGYPSL
GTITDVTIDGIYGTSPTGSFGHVLCSNPSKCQNIKINNVIMATAVDTASGIGTTAVLNTVGVPPSILSGAPADIG
LFDLSLPNGCYLERAAYATDSTPVPYFAPKPRDRPLGYLDGVPKTPADCLQPGSKWWLAQGQPSDFTLSTTSYKP
PVYPKVHLSQIPASRVKVSPAFPKNG
>GH28_QRW23661.1|RHAZE@28612@001687245_1|XP_017491101.1|1
LNPVVVQINDSSHIIVSDLTFINSPWFTVAPYRSEYVTIDHVTVKNPSNAPNTDGIHPEACHHVVISNCFVSTGD
DGIVITSETDKNTNAKYSSENITVNHCTVHSGHGGIVIGSAISGDVRDIHAHNLHFEGTLRGVRLKSTRDHGGIV
ENIYIKDITMKNISDEGITISAFYDVKNFDPHNIPSKTFDVSKTPTFRNIHLTNVTGDSKLGLQIVGLPEKHFDK
IELKQVHLVAKKEEIIVNADHVLKENFIYKIDKHLHDN
>GH28_QRW23661.1|BRACO@38358@014529535_1|XP_037032572.1|0
MKNLITIATFIGSVLTSQLLPLPTQFFSSRLGDSAPLDKKFIVENCDYNGRLLAFSIENCTDTEGICNIVTKREY
NVNITFEPSVSTSNLIWRVIFNINGQDQVLVERPIYEHVQPGVTYTLFNTFAFGFENEGLSFPATFQIIDASTER
AEICHSAVLDVNSWHMVPEILSRIKAPIFPDRDFDITTYGAVPDGETDNTEAFSRAIVHCHLLGGGRVVVPPGVF
LSSAITLLSNVNLHLKEGSTILFTQNTTAYPNVFTRLGGLELINFSPFIYAFGAENIALTGSGVLNGNADCEHWW
PWKGRNNNLELLCGIIEGFPTEEADVAALTEMAERNVPVEERIFGEGHFMRPVFVQPYNSKNILIEGVTFLRSPN
WILNPVLCENVIVRGVTINSTGPNSDGCNPESSKDVLIENVKFITGDDCIAVKSGRNADGRRINVKSENIVIQNC
EMENGHGGFTIGSEISGGAQNIFCQNCSMNSPQLEQGLRFKNNAVRGGLIEDIYIRNIHIPELYTGTSASRGMVL
SIDFFYEEGPNGNYPPVVRNVDIRNVTALKSNYALYLRGFPTDQITNVRLYDCHFNGVVRGSVIEHVENLGLFNV
TVNGDVIEVPAV
>GH28_QRW23661.1|SPOLI@69820@002706865_1|XP_022836218.1|1
MFCNSLVAQKTPTYSWKNLPKIVQPTFAKDTFNVTKFGAKPDGITLNTQAINDAITACSKKGGGVVLVPNGMWLT
GPIVLKSNVNLHIRKAATLLFTEDKSQYPLVEGSYEGKSAARNQSPISATNQENIAITGQGIIDGSGDVWRAVNK
AQLTESEWNAKKASGGVLKANGLTWYPSEQFMKASVENRSMLLKDGVSLQSFADMKDFLRPNLLVITKCKKVLLE
GVTFQNSPAWCLHPLMSENLTLRNLTIKNPEYAHNGDGMDIESCKNFLVDGCTIDVGDDAICIKSGKDEEGRKRG
MPTENGIIRNCIVYNGHGGFVVGSEMSGGARNIFVYDCTFAGTDKGLRFKSVRGRGGIVENIYAKNIFMKDIAQE
AIFFDMYYFVKFATDSPRDERPVVNEGTPIFRNMKFENIVCKGANKGIFVRGLSEMPIQNIQMSNIVLDTKIGAE
FIDASNITLEKVTLISENTKPVISVNNSDGLTFNTIQYKANAELLFAIAGERSKAIQILQTDSSKAQKQIEFTEG
ATKEAITVSTGK
```

Summary and Discussion

- Why are there plant cell wall degrading enzymes in animals?
- Why is GH28 of *Spodoptera litura* identical to a bacterial sequence? What are the consequences given the assumptions stated above?

cellulase_all.extended.fa.gz

Back to main