UNIVERSIDADE FEDERAL DE UBERLÂNDIA
INSTITUTO DE BIOTECNOLOGIA
PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

ANA CARLA MARTINS DA SILVA

IDENTIFICAÇÃO E ANÁLISE DA DIVERSIDADE DE GLICOSIL HIDROLASES DA FAMÍLIA GH18 EM ALGAS, BRIÓFITAS E PTERIDÓFITAS

PATOS DE MINAS - MG
MARÇO DE 2021

ANA CARLA MARTINS DA SILVA

IDENTIFICAÇÃO E ANÁLISE DA DIVERSIDADE DE GLICOSIL HIDROLASES DA FAMÍLIA GH18 EM ALGAS, BRIÓFITAS E PTERIDÓFITAS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Biotecnologia como requisito parcial para obtenção do título de mestre em Biotecnologia.

Orientador: Prof. Dr. Aulus Estevão Anjos de Deus Barbosa

PATOS DE MINAS - MG
MARÇO DE 2021

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

S586i
2021

Silva, Ana Carla Martins da, 1992-
Identificação e análise da diversidade de glicosil hidrolases da família GH18 em algas, briófitas e pteridófitas [recurso eletrônico] / AnaCarla Martins da Silva. - 2021.

Orientador: Aulus Estevão Anjos de Deus Barbosa.
Dissertação (mestrado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Biotecnologia.
Modo de acesso: Internet.
Disponível em: http://doi.org/10.14393/ufu.di.2021.6036
Inclui bibliografia.
Inclui ilustrações.

1. Biotecnologia. I. Barbosa, Aulus Estevão Anjos de Deus, 1979-, (Orient.). II. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Biotecnologia. III. Título.

CDU: 57.08

Rejâne Maria da Silva – CRB6/1925
Bibliotecária

IDENTIFICAÇÃO E ANÁLISE DA DIVERSIDADE DE GLICOSIL HIDROLASES DA FAMÍLIA GH18 EM ALGAS, BRIÓFITAS E PTERIDÓFITAS

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Biotecnologia como requisito parcial para obtenção do título de mestre em Biotecnologia.

Aprovado em 09/03/2021

BANCA EXAMINADORA

Aulus Barbosa

Prof. Dr. Aulus Estevão Anjos de Deus Barbosa

Leonardo Lima Pepino de Macedo

Prof. Dr. Leonardo Lima Pepino de Macedo

Prof. Dr. Matheus de Souza Gomes

Patos de Minas – MG

2021

AGRADECIMENTOS

Ao meu pai Valmir Jacó e à minha mãe Elizabeth pela confiança e incentivo, e por sempre apoiarem a realização dos meus projetos. Vocês são meu exemplo de vida!

À minha irmã Ana Flávia e minha cunhada Mila por sempre seguirem ao meu lado me incentivando e me apoiando.

Ao Prof. Dr Aulus, pelas orientações, compreensão e disponibilidade na elaboração deste trabalho. Obrigada!

Aos professores e colaboradores da UFU-Campus Patos de Minas que ao longo dessa jornada contribuíram com importantes ensinamentos para minha formação.

Gratidão a todos que diretamente ou indiretamente contribuíram para a realização deste trabalho.

RESUMO

O Reino Plantae é um grupo monofilético composto por de organismos eucariontes e fotossintetizantes. Em termos evolutivos, as algas verdes são consideradas ancestrais das plantas terrestres, pois este grupo compartilha características importantes, como a capacidade de realizar fotossíntese e os mesmos tipos de clorofilas. Ao longo da evolução, ao se adaptarem ao ambiente terrestre, as plantas também desenvolveram importantes mecanismos fisiológicos de defesa ao ataque de predadores. Neste sentido, as quitinases vegetais são importantes proteínas que produzem resposta de defesa ao ataque de patógenos e artrópodes que contém quitina, atuando também no desenvolvimento das plantas. As quitinases pertencem à família GH18 e GH19 das glicosil hidrolases e até o momento foram amplamente descritas em gimnospermas e angiospermas, porém pouco descritas em espécies vegetais mais primitivas. Assim, este trabalho tem por objetivo, através de um conjunto de ferramentas de bioinformática, identificar e realizar anotação das quitinases da família GH18 em algas, briófitas e pteridófitas. Inicialmente foi realizada uma busca por sequências referências de quitinases vegetais no site CAZy e a criação de um banco de dados de proteínas traduzidas a partir dos transcriptomas de 388 espécies de algas, briófitas e pteridófitas, fornecidas pela plataforma OneKP. A busca por sequências de possíveis quitinases da família GH18 neste banco de dados foi feita por meio de alinhamentos no programa BLASTp. Então, foi realizada uma análise de domínios conservados no programa Blast2GO das possíveis quitinases encontradas. Foram realizados alinhamentos múltiplos no programa MUSCLE para selecionar apenas as sequências com o domínio catalítico conservado. Para classificar as quitinases identificadas, foi realizada a construção de árvores filogenéticas utilizando o método Neighbor-Joining. No total foram encontradas 648 prováveis quitinases da família GH18, sendo 162 em algas verdes e 23 em outros filos de algas, 238 em briófitas e 225 em pteridófitas. Para comparar as relações estruturais, foi realizada a modelagem de estruturas tridimensionais no servidor Swiss-Model. Os resultados desta pesquisa são importantes para uma melhor compreensão das origens das quitinases em grupos vegetais primitivos, mostrando que estas enzimas podem ser encontradas desde as algas mais antigas e possibilitar a comparação com quitinases de plantas vasculares superiores.

Palavras-chave: Quitinases. Evolução. Transcriptômica. Defesa vegetal.

ABSTRACT

Kingdom Plantae is a monophyletic group composed of eukaryotic and photosynthetic organisms. Green algae are considered ancestors of terrestrial plants, as these groups share important characteristics, such as the ability to perform photosynthesis and the same chlorophylls types. Throughout evolution, when adapting to the terrestrial environment, plants have also developed important physiological mechanisms to defend against predators. In this sense, plant chitinases are important proteins that produce a defense response against pathogens and arthropods that contain chitin, also acting on plant development. Chitinases belong to the GH18 and GH19 family of glycosyl hydrolases and have so far been widely described in gymnosperms and angiosperms, but little described in more primitive plant species. Thus, this work aims, through a set of bioinformatics tools, to identify and annotate chitinases of the GH18 family in algae, bryophytes, and pteridophytes. Initially, a search for plant chitinases reference sequences was carried out on the CAZy website, and the creation of protein database translated from transcriptomes of 388 algae, bryophytes, and pteridophytes species, provided by the OneKP platform. Search for sequences of putative chitinases of the GH18 family in this database was done through alignments in the BLASTp program. Then, was performed a conserved domains analysis, using Blast2GO, of the putative chitinases found. Multiple alignments were performed with MUSCLE method to select only the sequences with the conserved catalytic domain. To classify identified chitinases, phylogenetic trees were constructed using Neighbor-Joining method. In total, 648 putative chitinases from the GH18 family were found, 162 in green algae and 23 in other algae phyla, 238 in bryophytes, and 225 in pteridophytes. To compare structural relationships, three-dimensional structures were modeled on the Swiss-Model server. The results of this research are important for a better understanding of chitinases origins in primitive plant groups, showing that these enzymes can be found from the oldest algae and enable the comparison with chitinases of superior vascular plants.

Keywords: Chitinases. Evolution. Transcriptomic. Plant defense.

LISTA DE ABREVIATURAS E SIGLAS

GH – Glicosil hidrolases
GH18 – Glicosil hidrolases da família 18
GH19 – Glicosil hidrolases da família 19
BLASTp – Basic Local Alignment Search Tool protein
MEGA – Molecular Evolutionary Genetics Analysis
MUSCLE – Multiple Sequence Comparison by Log-Expectation
NCBI – National Center for Biotechnology
PDB – Protein Data Bank
PR – Pathogenesis Related
JTT – Jones Taylor Thornton

SUMÁRIO

CAPÍTULO I – REFERENCIAL TEÓRICO

1 INTRODUÇÃO

1.1 Problema de pesquisa

O Reino Vegetal é formado por organismos fotossintetizantes organizados em diferentes filos de algas, briófitas e plantas vasculares, incluindo pteridófitas, gimnospermas e angiospermas. Ao longo da evolução, as plantas foram adquirindo tecidos especializados, o que favoreceu a sua adaptação e ocupação do ambiente terrestre (INGROUILLE; EDDIE, 2006). Evidências apontam que as plantas vasculares evoluíram a partir de algas verdes, devido às semelhanças compartilhadas entre esses grupos, como por exemplo, a presença de pigmentos fotossintetizantes armazenados em plastídios e parede celular constituída por celulose (CASANOVA; NAIRN, 2016).

Ao longo do processo evolutivo, desde as algas, foram selecionados uma série de mecanismos fisiológicos de defesa contra o ataque de predadores. Sabe-se que em plantas vasculares as proteínas quitinases estão relacionadas a mecanismos de defesa contra o ataque de patógenos e artrópodes que contém quitina, pois essas proteínas agem promovendo a hidrólise do polímero de quitina (KESARI et al., 2015). Além dessa importante função, tem sido relatado que as quitinases também desempenham papel no desenvolvimento e crescimento das plantas terrestres (EDREVA, 2005; HAMID et al., 2013). Uma grande diversidade de quitinases são atualmente estudadas em plantas gimnospermas e angiospermas, mas até o momento, pouco relatadas em algas, briófitas e pteridófitas.

Diversos estudos sugerem que as quitinases passaram por um rápido processo evolutivo, uma vez que muitas quitinases, até então estudadas, demonstraram grandes diferenças em suas sequências de aminoácidos, conformação estrutural e até mesmo mecanismos enzimáticos diferentes (FUKAMIZO, 2000). Há uma grande diversidade de quitinases relatadas em plantas vasculares superiores, mas será que os grupos mais ancestrais das plantas vasculares mais recentes apresentam também essa grande variedade de quitinases? As quitinases em algas, briófitas e pteridófitas divergem muito em similaridade de sequências e estruturas tridimensionais quando comparadas com as quitinases de plantas vasculares superiores?

1.1 **Hipótese**

Até o momento, estudos têm relatado a identificação de quitinases em diversas plantas vasculares superiores (gimnospermas e angiospermas), mas os estudos de quitinases em algas, briófitas e pteridófitas ainda são limitados. Pressupõe-se que nesses grupos ancestrais há também uma grande diversidade de quitinases, que possam ter diferenças significativas com as quitinases de plantas vasculares superiores, uma vez que é sugerido que as quitinases vem sofrendo rápidas adaptações evolutivas.

1.2 **Objetivos**

1.2.1 Objetivo geral

Realizar investigação e anotação de proteínas quitinases nos grupos ancestrais das plantas terrestres, desde algas, briófitas e pteridófitas, por meio de um conjunto de ferramentas de bioinformática.

1.2.2 Objetivos específicos

- Identificar e analisar as sequências de aminoácidos de prováveis proteínas de espécies de algas, briófitas e pteridófitas para verificar se há presença de proteínas quitinases;
- Analisar se as proteínas identificadas são quitinases, verificando a presença do domínio catalítico característico da família GH18;
- Classificar as quitinases identificadas da família GH18 por meio de análise filogenética;
- Modelar estruturas tridimensionais de quitinases de cada grupo analisado para comparar as relações estruturais.

1.3 **Justificativa**

Sabe-se que as quitinases da família GH18 são proteínas presentes em plantas e em diversos organismos. Possuem funções importantes para o desenvolvimento e crescimento das plantas, principalmente relacionadas a defesa contra o ataque de patógenos e artrópodes que contém quitina, uma vez que essas enzimas promovem a hidrólise do polímero de quitina. A maioria dos estudos relatam análises de quitinases em gimnospermas e angiospermas, no entanto, existem poucos dados sobre quitinases de algas, briófitas e pteridófitas disponíveis. Devido à importância dessas proteínas no Reino vegetal e para compreender melhor a dinâmica dessas proteínas em algas, briófitas e pteridófitas, este estudo busca investigar e realizar anotação da diversidade de quitinases da família GH18 nesses grupos vegetais distintos e mais primitivos. A anotação de dados poderá ser útil para o fornecimento de informações para estudos futuros e auxiliar na compreensão das relações evolutivas entre quitinases de grupos vegetais ancestrais com as plantas vasculares superiores.

2 REFERENCIAL TEÓRICO

2.1 **Origem e evolução das plantas**

O Reino Plantae é um grupo monofilético composto por mais de 370.000 espécies de organismos eucariontes, geralmente autotróficos, e adaptados ao ambiente terrestre (CHRISTENHUSZ; BYNG, 2016). Este grande grupo abrange alguns tipos de algas unicelulares e multicelulares, briófitas, pteridófitas, gimnospermas e angiospermas (BENNICI, 2008). Estão incluídos neste grupo organismos eucariontes que possuem cloroplastos, realizam fotossíntese através da clorofila e armazenam os produtos da fotossíntese, como o amido, além de possuírem parede celular constituída por celulose (BLACKWELL, 2003; NISHIO, 2000). As plantas são seres pluricelulares com elevado nível de organização e especialização, enquanto em algas não existe diferenciação de tecidos, mas sim um conjunto de células, algumas podendo apresentar certa especialização (NISHIO, 2000). Atualmente as plantas terrestres são agrupadas no clado Embriófitas, que incluem briófitas, pteridófitas, gimnospermas e angiospermas (EVERT; EICHHORN, 2014).

Evidências evolutivas indicam que as algas verdes são ancestrais das plantas terrestres, com registros fósseis relatados de 900 milhões de anos (INGROUILLE; EDDIE, 2006). Enquanto, que a expansão das plantas no ambiente terrestre ocorreu por volta de 470 milhões de anos atrás (ISHIZAKI, 2017). Estudos indicam que a origem das plantas é baseada a partir de uma cianobactéria de vida livre (ancestral procariótico comum) que são eubactérias gram-negativas referidas como "algas azuis-verdes", das quais derivaram os plastídios; organelas que acumulam pigmentos fotossintetizantes e possibilitam assim, a realização da fotossíntese, e essa teoria é apoiada pela semelhança do material genético que os constituem (MARTIN; GARG; ZIMORSKI, 2015; MCFADDEN, 2014). Algumas cianobactérias apresentam clorofilas *a* e *b*, pigmentos carotenoides e membranas tilacoides, assim como plastídios; outras apresentam ficobilina, que são pigmentos encontrados em algas vermelhas, e que reforça ainda mais essa teoria (BHATTACHARYA; YOON; HACKETT, 2004).

A teoria do surgimento de plastídios em algas verdes e plantas envolve um processo de endossimbiose primária de uma cianobactéria por protistas não fotossintéticos, enquanto a diferença de complexidade entre os outros tipos de algas é descrito por um evento de endossimbiose secundária que envolve o englobamento de um plastídio primário (como é o caso de algumas algas euglenoides que se tornaram fotossintéticas por englobar o plastídio de uma alga verde fotossintética), e até mesmo por endossimbioses terciarias por englobamento de plastídios secundários (observado em algas que apresentam de três a quatro membranas envolvendo as organelas fotossintéticas) (BHATTACHARYA; YOON; HACKETT, 2004; MARTIN; GARG; ZIMORSKI, 2015). Ao longo do processo evolutivo das plantas, diversas alterações significantes ocorreram no ambiente, o que favoreceu o desenvolvimento de caracteres adaptativos as novas condições (ANDERSON; WILLIS; MITCHELL-OLDS, 2011). Assim, as plantas terrestres desenvolveram estruturas adaptativas mais complexas, como vasos condutores de nutrientes e água, presença de embrião, além de apresentarem órgãos reprodutivos complexos, como o órgão reprodutor feminino, denominado de arquegônio, e o órgão reprodutor masculino, conhecido como anterídio (GRAHAM; COOK; BUSSE, 2000; ISHIZAKI, 2017).

2.2 **Algas**

Assim como as plantas terrestres, as algas realizam papel importante no habitat em que vivem, podendo ser encontradas em regiões rochosas, em água doce e águas marinhas, e até mesmo em regiões desérticas (PANAWALA, 2017a). Uma característica marcante em algas, assim como em plantas terrestres, é a capacidade de realizarem fotossíntese, atuando, portanto, no ciclo do carbono com a transformação de dióxido de carbono (CO_2) em carboidratos e em carbonato de cálcio (RAVEN et al., 2012). Além da capacidade de consumir dióxido de carbono, que é um dos principais gases do efeito estufa, e produzir oxigênio, as algas são a base essencial na cadeia alimentar de muitos organismos marinhos, alguns de água doce e em menor parte cadeias alimentares terrestres (PANAWALA, 2017a).

Existem vários filos de algas conhecidos, além do importante grupo de algas verdes, abrangendo um número de aproximadamente 70.000 espécies vivas e extintas, com diferenças significativas, que foram classificadas de acordo com sua fisiologia (INGROUILLE; EDDIE, 2006). Entre elas, encontram-se as algas euglenoides (filo Euglenophyta), que são consideradas uma das linhagens mais antigas; as criptófitas (filo Cryptophyta); haptófitas (filo Haptophyta); dinoflagelados (filo Dinophyta); diatomáceas (filo Bacillariophyta); algas douradas (filo Chrysophyta); algas verde-amarelas ou xantofíceas (filo Heterokontophyta); algas pardas (filo Ochrophyta); algas vermelhas (filo Rhodophyta) e algas verdes (filo Chlorophyta) (RAVEN; GIORDANO, 2014). A diferença evolutiva nos filos de algas está relacionada principalmente ao tipo de pigmento que constituem seus plastídios. Algumas algas apresentam clorofilas *a* e *b* ou *c* em suas estruturas, podem armazenar pigmentos carotenoides, enquanto outras podem armazenar o pigmento ficobilina (CAVALIER-SMITH, 1982; LEWIS; MCOURT, 2004).

2.2.1 Algas verdes

As algas verdes constituem um grande grupo diversificado de aproximadamente 8,000 espécies, que variam de organismos simples unicelulares até organismos multicelulares complexos. Podem ser encontradas, em grande parte, em regiões aquáticas, principalmente água doce, com exceções de alguns grupos que podem se desenvolver em regiões extremas, como a superfície de neve, desertos, tronco de árvores, solo e também em associações simbióticas com liquens, fungos, protozoários de água doce, esponjas e celenterados (GUIRY, 2012; LEWIS; MCCOURT, 2004).

As algas verdes apresentam semelhanças com briófitas e plantas vasculares, uma vez em que ambas armazenam amido nos plastos, possuem parede celular constituída por celulose e sintetizam clorofilas *a* e *b* (LEWIS; MCCOURT, 2004). O grupo de algas verdes (filo Chlorophyta) é dividido em três principais classes: Chlorophyceae, Charophyceae e Ulvophyceae (COOPER, 2014). A classe Chlorophyceae (conhecidas por clorófitas) abrange algas verdes com características de morfologia e ecologia diferentes das demais algas deste grupo, podendo haver clorófitas unicelulares e multicelulares, de complexidade diferente (FANG et al., 2017). A classe Charophyceae (carófitas) é composta principalmente por algas verdes de água doce, com algumas espécies habitando regiões de águas salinas e de vida livre em terra, enquanto Ulvophyceae (ulvócitas) é uma classe de menor dimensão, composta principalmente por algas marinhas, filamentosas ou lâminas achatadas (RAVEN; GIORDANO, 2014).

No filo de algas verdes, é evidenciado que algas da classe Charophyceae são as que possuem maior grau de parentesco com briófitas e plantas vasculares, devido a similaridades morfológicas, bioquímicas e genéticas observadas entre esses grupos (BENNICI, 2008; ISHIZAKI, 2017). Carófitas e briófitas armazenam clorofilas *a* e *b* e betacaroteno como pigmentos fotossintetizantes; armazenam alimento na forma de amido e possuem parede celular constituída por celulose (BLACKWELL, 2003). Devido a preservação dessas características no decorrer da evolução, as algas verdes são consideradas ancestrais das plantas terrestres (denominadas embriófitas), e por isso são muito estudadas para compreender a origem e evolução das plantas (CASANOVA; NAIRN, 2016).

2.3 **Briófitas**

O grupo das briófitas é constituído por aproximadamente 25,000 espécies de plantas que possuem crescimento regulado por células apicais, tecidos pouco diferenciados, além de não possuírem vasos condutores especializados em transporte de água e nutrientes (FRAHM et al., 2003; PITTERMANN, 2010). As briófitas são plantas de pequeno porte, justamente por não possuírem tecido vascular, e habitam geralmente regiões úmidas e sombreadas, junto ao solo, sobre rochas ou sobre outras plantas; algumas espécies podem ainda se desenvolver em regiões desérticas, sob condições extremas de calor (PANAWALA; 2017a).

As briófitas apresentam tecidos de maior especialização e muitas características que não estão presentes em algas verdes, mas que são compartilhadas com pteridófitas, como os gametângios masculinos e femininos, denominados anterídios e arquegônios, e tecidos produzidos por um meristema apical (EVERT; EICHHORN, 2014). Assim, devido às semelhanças compartilhadas entre esses grupos, estudos evolutivos apontam as briófitas como as primeiras plantas terrestres simples, derivadas diretamente de algas verdes, sendo consideradas, portanto, os ancestrais das pteridófitas, auxiliando na compreensão do surgimento e evolução das plantas na Terra (BENNICI, 2008; PANAWALA; 2017a).

As briófitas são divididas em três filos: Marchantiophyta (hepáticas), Bryophyta (musgos) e Anthocerophyta (antóceros) (CASANOVA; NAIRN, 2016; PUTTICK et al., 2018). As briófitas hepáticas, são geralmente talosas ou folhosas; os musgos são as briófitas mais conhecidas e espalhadas pelo mundo; enquanto que o grupo dos antóceros é considerado o menos diversificado de briófitas (FRANGEDAKIS et al., 2020; HE-NYGRÉN et al., 2006).

2.4 **Pteridófitas**

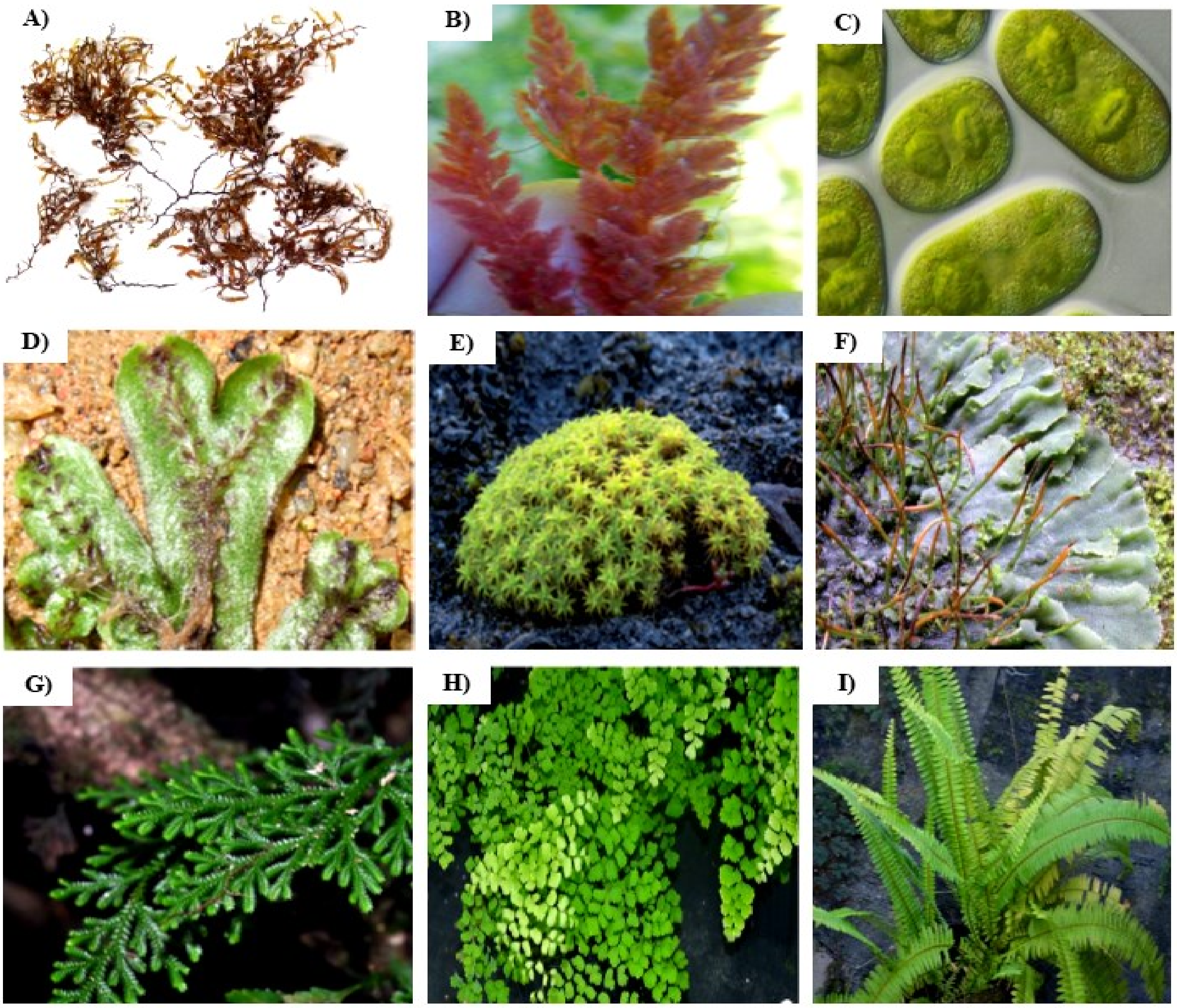
As primeiras pteridófitas surgiram a cerca de 400 milhões de anos durante a Era Paleozoica, dos períodos Siluriano ao Devoniano. As pteridófitas são plantas vasculares sem sementes, com aproximadamente 10.000 espécies, encontradas em diferentes habitats que incluem rochas, solo, ambientes escuros e úmidos, predominantemente em regiões tropicais (TRYON, R. M; TRYON, A. F, 1982).

Um importante estágio inicialmente observado em pteridófitas, no decorrer da evolução, foi o estabelecimento de um esporófito livre e independente, o que contribuiu para o processo de adaptação ao ambiente terrestre, uma vez que possibilitou maior resistência e um maior crescimento vegetal (BLACKWELL, 2003). Outra característica importante que surgiu nas pteridófitas foi a presença de vasos condutores, xilema e floema, o que contribuiu para a sobrevivência das plantas, pois facilitou o processo de transporte de água e nutrição para plantas maiores (PANAWALA, 2017b). O surgimento de heteroesporia ocorreu também primeiramente em pteridófitas, nos quais as plantas passaram a produzir esporos diferentes por meiose, sendo estes denominados micrósporos e megásporos, originando respectivamente, os gametófitos masculinos e femininos (WILLIS; MCELWAIN, 2002).

Atualmente, o grupo de plantas pteridófitas compreende as divisões Lycophyta e Monilophyta (samambaias e cavalinhas) (CHRISTENHUSZ; CHASE, 2014; PRYER et al., 2004). Este grupo também era composto pelos filos Rhyniophyta, Zosterophyllophyta e Trimerophytophyta que entraram em extinção no final do Devoniano. O estudo de pteridófitas é importante para a melhor compressão do processo evolutivo das plantas, com ênfase na evolução de gimnospermas e angiospermas.

A figura 1 contém exemplos de espécies das principais divisões dos filos de algas, briófitas e pteridófitas.

Figura 1 - Exemplos de espécies de algas, briófitas e pteridófitas. **A)** Alga marrom do filo Ochrophyta (espécie *Sargassum natans*). **B)** Alga vermelha do filo Rhodophyta (espécie *Asparagopsis armata*). **C)** Alga verde unicelular carófita (espécie *Cylindrocystis crassa*). **D)** Briófita do filo Marchantiophyta (espécie *Marchantia emarginata).* **E)** Musgo da espécie *Syntrichia ruralis.* **F)** Briófita do filo Anthocerophyta (espécie *Phaeomegaceros hirticalyx)*. **G)** Pteridófita do clado Lycophyta (espécie *Selaginella involvens*). **H)** Pteridófita do clado Monilophyta conhecida como avenca (espécie *Adiantum capillus-veneris).* **I)** Pteridófita do clado Monilophyta conhecida como samambaia (espécie *Nephrolepis cordifolia*).



Fonte: Baskaran et al. (2018); Cotas et al. (2020); Ho (2013); Huttunen; Bell e Hedenäs (2018); Mendez-Tejeda e Rosado (2019); Samolov et al. (2020); Villarreal et al. (2014).

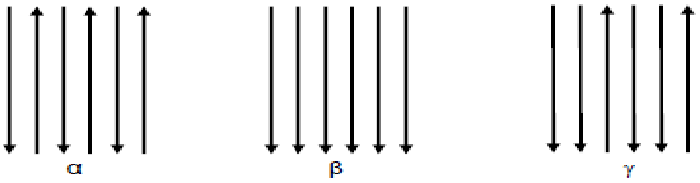
2.5 Quitina

A quitina foi isolada pela primeira vez a partir de fungos em um estudo realizado por Braconnot (1811). A quitina é considerada o segundo polissacarídeo mais abundante da natureza (MARTÍNEZ-CABALLERO et al., 2014), com a celulose se destacando em primeiro lugar (CAMPANA-FILHO et al., 2007). É um polímero linear, insolúvel, composto por uma longa cadeia de unidades repetitivas de 2-acetamido-2-desoxi-D-glucose, geralmente designado também de N-acetil-D-glucosamina (GlcNAc), que são unidos por ligação glicosídicas β-(1-4), (SEIDL, 2008).

A quitina é um biopolímero que compõe a matriz da estrutura esquelética de muitos invertebrados, estando presente no exoesqueleto de artrópodes, moluscos e celenterados, em algas diatomáceas, e também nas paredes celulares de determinados fungos, desempenhando função de revestimento e materiais de suporte aos organismos (CAMPANA-FILHO et al., 2007).

Existem na natureza três formas diferentes de quitina as quais diferem na disposição das suas cadeias. Na forma α as cadeias estão orientadas no sentido antiparalelo, na forma β as cadeias estão dispostas no sentido paralelo e em γ ocorre um misto de cadeias paralelas e antiparalelas (Figura 2) (SEIDL, 2008). Essas diferenças estruturais conferem a quitina funções diferentes nos organismos, na qual a forma α fornece maior rigidez, enquanto as formas β e γ conferem maior flexibilidade e resistência aos organismos (KAYA et al., 2017).

Figura 2 - Orientação das cadeias poliméricas de quitina. Cadeias α são posicionadas no sentido antiparalelo, na forma β as cadeias são dispostas no sentido paralelo, enquanto na forma γ ocorre um misto de cadeias α e β.



Fonte: Matsiu (2007).

A quitina é um biopolímero com alta versatilidade e é muito utilizada na indústria de alimentos como agente floculante, no tratamento de efluentes, como adsorvente na clarificação de óleos, e em grande parte na produção de quitosana (MOURA et al., 2006).

2.6 Quitinases em plantas

As plantas não contém quitina em seus constituintes, porém, possuem genes que expressam quitinases de forma constitutiva ou induzível (OHNUMA et al., 2011a). As plantas geralmente expressam níveis relativamente baixos de quitinases e só sintetizam níveis mais elevados da enzima quando entram em contato com fitopatógenos que contem quitina. Além de estar presente em plantas, as quitinases também são amplamente encontradas em crustáceos, bactérias e fungos, as quais podem desempenhar diferentes funções. Em plantas, as quitinases podem ser expressas em caules, sementes, tubérculos e flores (KHAN et al., 2015).

As plantas produzem uma série de mecanismos de defesa, e entre eles destaca-se a expressão de um complexo de proteínas relacionadas a patogênese, denominadas proteínas PR *(Pathogenesis-Related)* (NEUHAUS et al., 1996). As quitinases estão incluídas neste grupo e são vistas como proteínas de defesa quando submetidas a estresses bióticos e abióticos. Essas proteínas apresentam a capacidade de induzir resistência local ou sistêmica ao ataque de fitopatógenos, como fungos patogênicos, por agir diretamente no polímero de quitina, que é um dos principais componentes da parede celular da maioria dos fungos (KESARI et al., 2015). Algumas quitinases relacionadas a patogênese exibem também atividades antibacteriana, nematicidas e antivirais (CHEN et al., 2015; MEDEIROS et al., 2018; KAUR et al., 2019)

As quitinases de plantas podem também apresentar um nível maior de expressão quando induzidas por fatores abióticos (CAO; TAN, 2019). Determinadas condições ambientais extremas, como elevadas concentrações salinas, longos períodos de frio e estresse hídrico podem promover a transcrição de alguns genes de quitinase, sugerindo assim, que essas proteínas possam estar envolvidas com mecanismos de resposta ao estresse em plantas (BADARIOTTI et al., 2006; CAO; TAN, 2019). Também é sugerido a atuação de quitinases no crescimento e desenvolvimento vegetal, e podem ser principalmente detectadas nos estágios iniciais de crescimento (EDREVA, 2005; HAMID et al., 2013).

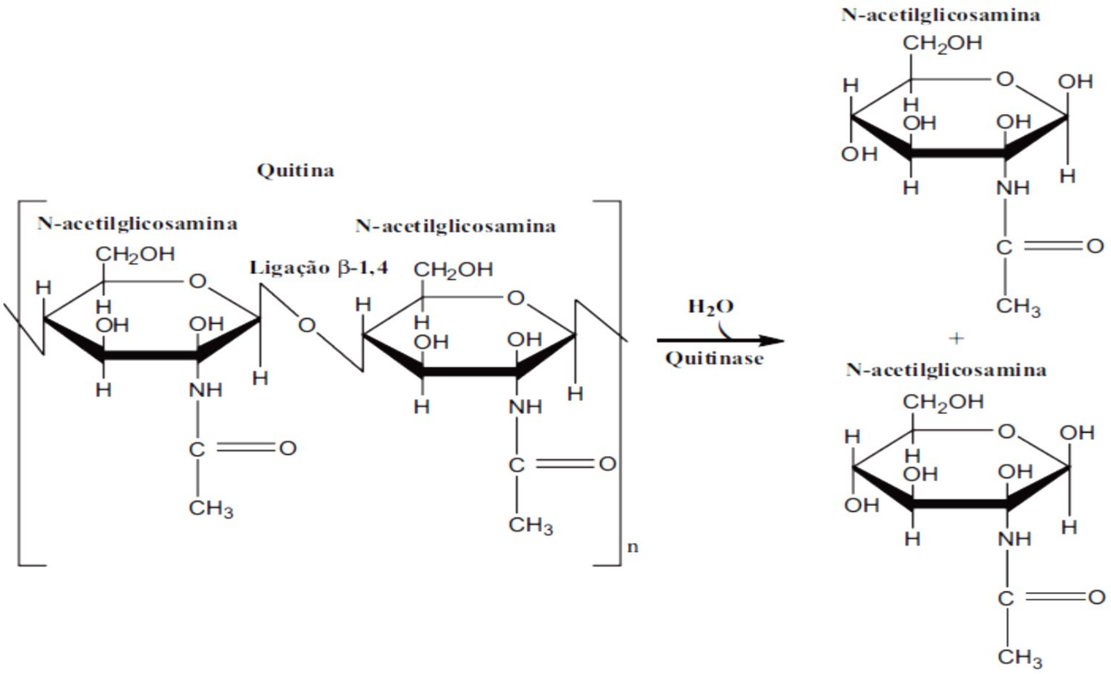
2.7 Características de quitinases

As quitinases são enzimas que catalisam a hidrólise das ligações glicosídicas β-1,4 entre os resíduos de N-acetilglucosamina (GlcNAc) que compõem o polímero de quitina (Figura 3) (MARTÍNEZ-CABALLERO et al., 2014; SEIDL et al., 2005).

Dependendo de seus padrões de clivagem, que estão relacionados ao tipo de produto formado após a hidrólise do substrato, as quitinases podem ser divididas em duas categorias: endoquitinases e exoquitinases. De acordo com o Comitê de Nomenclatura da União Internacional de Bioquímica e Biologia Molecular, as quitinases podem ser classificadas em três grupos de enzimas com três números distintos (números E.C., do inglês Enzyme Commission): (i) endoquitinases (E.C. 3.2.1.14), que clivam aleatoriamente ligações glicosídicas internas ao longo do polímero de quitina e geram multímeros de GlcNAc com alto peso molecular, (ii) exoquitinases que clivam na extremidade não redutora da cadeia polimérica, levando a formação de monômeros de N-acetilglucosamina (E.C. 3.2.1.200) e (iii) exoquitinases que catalisam a liberação de N-acetilglucosamina a partir da extremidade redutora da quitina (E.C. 3.2.1.201) (BRZEZINSKA et al., 2014; SEIDL, 2008). Por estarem mais amplamente distribuídas entre os organismos, as endoquitinases são mais estudadas do que as exoquitinases (HAN et al., 2016).

As quitinases vegetais ainda podem apresentar isoformas básicas e ácidas, de acordo com seus pontos isoelétricos, e são provenientes de regiões diferentes das plantas. Quitinases básicas são encontradas no vacúolo de plantas, enquanto quitinases ácidas são frequentemente secretadas para regiões extracelulares (LEGER et al., 1996).

Figura 3 - Hidrólise de quitina por proteínas quitinases.



Fonte: Fleuri e Sato (2005).

2.8 Classificação das quitinases

As quitinases pertencem a superfamília das glicosil hidrolases (GH), uma família de enzimas que hidrolisam ligações glicosídicas em carboidratos (NEUHAUS et al., 1996). Baseando-se na semelhança na sequência de aminoácidos do domínio catalítico de quitinases e no mecanismo de catálise, as quitinases vegetais, e foram divididas em famílias GH18 e GH19 das glicosil hidrolases (KESARI et al., 2015; LOMBARD et al., 2014). A família GH18 possui sequências de aminoácidos com baixa similaridade entre si, abrangendo quitinases de diversos organismos como vírus, bactérias, fungos, animais e algumas plantas. Já a família GH19 possuem sequências de alta similaridade entre si, e é composta principalmente por quitinases vegetais, podendo incluir também algumas quitinases bacterianas (PATIL; GHORMADE,; DESHPANDE, 2000).

Baseando-se em análises evolutivas, acredita-se que as quitinases das famílias GH18 e GH19 tenham evoluído de ancestrais diferentes, pois não compartilham similaridade em suas sequências de aminoácidos, e apresentam estruturas tridimensionais e mecanismos enzimáticos diferentes (FUKAMIZO, 2000). Apesar de quitinases da família GH18 apresentarem baixa similaridade com as quitinases da família GH19, suas sequências se assemelham às de quitinases de fungos e bactérias (KESARI et al., 2015).

Inicialmente, as quitinases vegetais foram divididas em seis classes (I-VI) (NEUHAUS et al., 1996). Posteriormente, foi identificada uma nova classe de quitinases, com características diferentes das quitinases anteriormente caracterizadas, as quais foram designadas de classe VII (JI; JINYUAN, 2002). As classes III e V pertencem à família GH18, enquanto as classes I, II, IV, VI e VII são integrantes da família GH19 (CAO; TAN, 2019). A maioria das quitinases vegetais estão distribuídas entre as classes I, II, III, IV e V.

As quitinases de classe I apresentam um domínio de ligação a quitina rico em cisteína na região N-terminal, contendo aproximadamente 40 aminoácidos, e um domínio catalítico C-terminal, altamente conservado, e geralmente são localizadas no vacúolo das plantas (ARIE et al., 2000; SARMA et al., 2012). As quitinases de classe II, são semelhantes às quitinases de classe I, porém apresentam apenas o domínio catalítico e não apresentam o domínio de ligação a quitina, e são frequentemente secretadas para o espaço extracelular (apoplasto). A classe III é composta por quitinases que apresentam baixa similaridade com as classes I e II, não apresentam domínio de ligação a quitina e são geralmente secretadas para o apoplasto (ARIE et al., 2000). Na classe IV encontram-se quitinases semelhantes a classe I, com a presença de domínios catalítico e de ligação a quitina, porém são de tamanhos relativamente menores, apresentando algumas deleções nos domínios catalíticos (ARIE et al., 2000; MARTÍNEZ-CABALLERO et al., 2014). Quitinases dessa classe tem sido identificadas, principalmente, em plantas dicotiledôneas (EL-KATATNY et al., 2001). Já a classe V, compreende um grupo de quitinases que geralmente possuem dois domínios de ligação a quitina e um domínio catalítico, de pouca semelhança com as demais classes (EL-KATATNY et al., 2001). As quitinases de classe VI possuem os domínios de ligação a quitina duplicados na região N-terminal, enquanto as quitinases de classe VII não possuem domínios de ligação a quitina (SU et al., 2015).

As proteínas relacionadas a patogênese (PR) também são divididas em famílias, de acordo com o grau de identidade de sequências e com a atividade biológica. Entre as 17 famílias atualmente descritas de proteínas PR, as quitinases estão distribuídas nas famílias PR-3, 4, 8 e 11 (SELS et al., 2008).

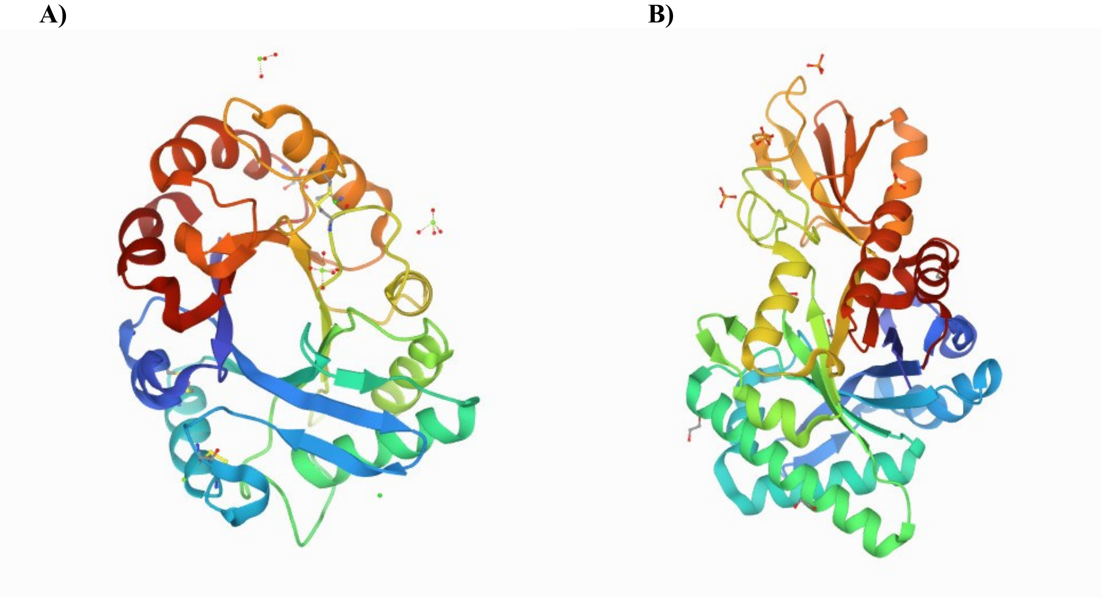
As plantas, em geral, apresentam uma grande quantidade de proteínas semelhantes à quitinases (CLPs) que também fazem parte das famílias GH18 e GH19 das glicosil hidrolases e, assim como as quitinases, também apresentam o domínio catalítico, porém, devido a modificações de resíduos de aminoácidos, ocorre a perda da função catalítica (MARTÍNEZ-CABALLERO et al., 2014). Dessa forma, as CLPs têm a capacidade de se ligarem a quitina, mas não conseguem promover a hidrólise das ligações glicosídicas, e, portanto, não realizam catálise.

2.9 **Família GH18**

A família GH18 é composta por endoquitinases e exoquitinases que possuem um domínio catalítico (C-terminal) caracterizado por uma estrutura de barril (α / β) constituído por 8 folhas-β e 8 α-hélices (Figura 4) (BADARIOTTI et al., 2006; CHUANG; LIN, H. Y.; LIN, F. P, 2008). Além do domínio catalítico, essas proteínas também podem apresentar domínios de ligação a quitina, situados na região N-terminal, que são regiões que não realizam catálise, mas que intensificam a afinidade da quitinase ao substrato, o que pode proporcionar um aumento da eficiência da hidrólise da quitina (YAN; FONG, 2015). Os domínios de ligação a quitina geralmente são conectados ao domínio catalítico por ligantes, o que contribui para que os domínios funcionem e evoluam de uma forma independente (KARLSSON; STENLID, 2009). Muitas quitinases GH18 possuem também um peptídeo sinal, que as direciona para serem secretadas para o espaço extracelular ou para uma localização subcelular (SEIDL et al., 2005).

As quitinases da família GH18 realizam reações de hidrólise das ligações glicosídicas pelo mecanismo de retenção da configuração anomérica, formando produtos com conformação β (BRAMELD et al., 1998; SEIDL et al., 2005).

Figura 4 - Estruturas tridimensionais de quitinases da família GH18 constituídas por_α-hélices e folhas-β. **A)** Quitinase de classe III de *Punica granatum.* **B)** Quitinase de classe V de *Nicotiana Tabaccum.*



Fonte: Masuda; Zhao; Mikami (2015) e Ohnuma et al., (2011b).

A família GH18 é, portanto, constituída de quitinases de classes III e V. As enzimas da classe III apresentam semelhanças com quitinases bacterianas, pois compartilham do mesmo mecanismo de hidrólise e similaridade na estrutura tridimensional (ISELI et al., 1996). A classe III tem uma característica peculiar, e são consideradas enzimas bifuncionais, pois além de hidrolisar quitina, também possuem atividade de lisozima, podendo assim, contribuir para a proteção contra infeções bacterianas (PARK et al., 2002). A função biológica principal da família GH18 em plantas está relacionada a degradação da quitina, mas em outros organismos elas podem desempenhar outras funções. Em fungos filamentosos a família GH18 está relacionada aos processos de degradação e modificação da parede celular, germinação de esporos, ramificação de hifas, diferenciação de esporos, autólise e micoparasitismo (ADAMS, 2004). Muitas espécies de bactérias secretam quitinases para utilizar quitina como uma fonte nutritiva de carbono e nitrogênio ou para defesa contra outros microrganismos (BHATTACHARYA; YOON; HACKETT, 2004). Por exemplo, em *Crassostrea gigas* (ostra-do-pacífico) as quitinases da família GH18 estão envolvidas no desenvolvimento embrionário inicial e funcionam como efetores da imunidade (BADARIOTTI, 2006)

CAPÍTULO II – ARTIGO CIENTÍFICO

Identification and analysis of GH18 glycosyl hydrolases in algae, bryophytes, and pteridophytes

Author names and affiliations:

Ana Carla Martins da Silva[a], Aulus Estevão Anjos de Deus Barbosa[a,*]

[a] Instituto de Biotecnologia, Universidade Federal de Uberlândia, Patos de Minas MG, Brazil.

Authors' e-mail address: anamartins@ufu.br
aulus@ufu.br

***Corresponding author:**

Aulus Estevão Anjos de Deus Barbosa
Instituto de Biotecnologia
Universidade Federal de Uberlândia
Avenida Getúlio Vargas, 230, 38700-103. Patos de Minas, MG, Brazil
Tel: +55 (34) 9 8400-8934
E-mail address: aulus@ufu.br

ABSTRACT

The Plantae Kingdom is a monophyletic group composed of eukaryotic and photosynthetic organisms. Throughout evolution, to adapt to the terrestrial environment, plants have developed important physiological defense mechanisms against predators and pathogens. Plant chitinases are important proteins that produce a defense response to pathogens and arthropods that contain chitin, also acting in plant development. Chitinases belong to a large protein family called glycosyl hydrolases (GH) and until now have been widely described in gymnosperms and angiosperms, however, little is reported in more ancient plant species. Plant chitinases are part of the GH18 and GH19 glycosyl hydrolases families. In this study, it was performed searching and annotation of GH18 chitinases in algae, bryophytes, and pteridophytes species, using a set of bioinformatics tools. In total, 648 putative chitinases from the GH18 family were found. The most complete sequences (270 sequences) were selected to additional analyzes. Using multiple sequence alignments, it was possible to observe the high catalytic domain conservation in all analyzed groups. However, some sequences showed substituted amino acids in catalytic residues, which can compromise the chitinolytic function. Phylogenetic analyzes enabled identified chitinases classification in GH18 classes III and V. Structural comparisons of some identified chitinases demonstrated high conservation between three-dimensional chitinase structures in algae, bryophytes and pteridophytes. The results of this study are important to assist in understanding the chitinases relationship between the most primitive plant groups with superior vascular plants, in addition to providing data for future studies related to GH18 chitinases.

Keywords: Chitinases, GH18, plant evolution, bioinformatics, transcriptome.

1 Introduction

The Plantae Kingdom is composed of approximately 370,000 eukaryotic organisms, usually autotrophic and adapted to terrestrial life (Christenhusz and Byng, 2016). This large group includes algae, bryophytes, pteridophytes, gymnosperms and angiosperms (Bennici, 2008). Like terrestrial plants, algae have great ecological importance because they perform photosynthesis and are responsible for much of the oxygen production in the terrestrial environment (Vecchi et al., 2020).

Green algae constitute a diverse group of approximately 8,000 species, being simple single-celled organisms and complex multicellular organisms (Guiry, 2012). Green algae are similar to bryophytes and vascular plants, since both stock starch in the plastids and have chlorophylls *a* and *b* (Nishio, 2000). Green algae group (phylum Chlorophyta) is divided into three main classes: Chlorophyceae, Charophyceae (also known as Streptophyta), and Ulvophyceae (Endymion D Cooper, 2014). Many studies report that Charophyceae algae have highest degree of relationship with bryophytes and vascular plants, due to morphological, biochemical, and genetic similarities (Bennici, 2008). Bryophytes group consists of approximately 25,000 species that show growth regulated by apical cells, poorly differentiated tissues, but do not have conduct vessels specialized in transporting water and nutrients (Frahm et al., 2003; Pittermann, 2010). They are, therefore, small plants, precisely because they do not have vascular tissue. Pteridophytes are seedless vascular plants, with approximately 10,000 species found in different habitats that include rocks, soil, dark and humid environments, predominantly in tropical regions (R M Tryon and Tryon, 1982). These plants have genes that allow them to adapt to different environments, allowing survival in the face of biotic and abiotic challenges.

Chitinases (EC 3.2.1.14) are enzymes that hydrolyze β-1,4 glycosidic bonds between N-acetylglucosamine (GlcNAc) residues that make up the chitin polymer (Martínez-Caballero et al., 2014; Seidl, 2008). Chitin is a linear, insoluble polymer, composed of a long chain of repetitive units of 2-acetamido-2-deoxy-D-glucose, generally also called N-acetyl-D-glucosamine, which are linked by β- (1-4) glycosidic bonds. Depending on their cleavage patterns, which are related to product type formed after substrate hydrolysis, chitinases can be divided into two categories: endo-chitinases and exo-chitinases (M S Brzezinska et al., 2014; Seidl, 2008).

Chitinases are widely found in plants, crustaceans, bacteria, and fungi, which can perform different functions. In plants, chitinases can be located in stems, seeds, turtles, and flowers (Khan et al., 2015). Plants do not contain chitin in their constituents, however, they have genes capable of producing chitinases in a constitutive or inducible way (Ohnuma et al., 2011). Plants generally express relatively low chitinases levels and only synthesize in higher levels when attacked by phytopathogens. Chitinases are included in a protein group related to pathogenesis called PR proteins *(Related to pathogenesis)* because their expression is induced when in contact with certain pathogens (Neuhaus et al., 1996). Chitinases can induce local or systemic resistance to phytopathogens attack, such as pathogenic fungi, by acting directly on chitin polymer, which is a major component of fungi cell wall (Kesari et al., 2015). Some pathogenesis-related chitinases may also exhibit potential antibacterial, nematicide, and antiviral activities (Chen et al., 2015; Medeiros et al., 2018; Kaur et al., 2019). Plant chitinases can also show a higher expression level when induced by abiotic factors (Cao and Tan, 2019). Certain extreme environmental conditions, such as high salt concentrations, long cold periods and water stress can promote chitinase genes transcription, thus suggesting these proteins may be involved in plants stress response mechanisms (Badariotti et al., 2006; Cao and Tan, 2019). Studies suggest chitinases also act on plant growth and development since chitinase accumulation was observed in the plants transition during flowering and senescence (Edreva, 2005).

Chitinases belong to glycosyl hydrolases (GH) superfamily that hydrolyze glycosidic bonds in carbohydrates (Neuhaus et al., 1996). Based on the similarity of chitinases catalytic domain sequence and in catalysis mechanism, plant chitinases were classified into the GH18 and GH19 glycosyl hydrolases families (Kesari et al., 2015; LOMBARD et al., 2014). Based on evolutionary analysis, it is suggested that chitinases of the GH18 and GH19 families have evolved from different ancestors, as they share low similarity between their amino acid sequences, and have different three-dimensional structures and enzymatic mechanisms (Fukamizo, 2000). Although GH18 chitinases show low similarity with GH19 chitinases, their sequences are similar to those fungi and bacteria chitinases (Kesari et al., 2015).

Plant chitinases were divided into seven classes (I-VII) (Li and Liu, 2003; Neuhaus et al., 1996). Classes III and V belong to the GH18 family, while classes I, II, IV, VI, and VII are members of the GH19 family (Cao and Tan, 2019). Class III is composed of chitinases that do not show similarity with classes I and II, do not show the chitin-binding domain and are generally located in the apoplast (Arie et al., 2000). Class V, on the other hand, comprises a group of chitinases that have two chitin-binding domains and a catalytic domain, of little resemblance to the other classes (El-Katatny et al., 2001).

GH18 chitinases perform glycosidic bonds hydrolysis by retention mechanism of the anomeric configuration, forming products with β conformation (Ken A Brameld et al., 1998; Seidl et al., 2005). They have a catalytic domain (C-terminal) characterized by a barrel-shaped structure consisting of 8 α-helices and 8 β-sheets (Chuang et al., 2008). In addition to the catalytic domain, these enzymes may present chitin-binding domains in the N-terminal region, which are regions that do not perform catalysis, but that enhances chitinase affinity to substrate, which can provide an increase in the efficiency of chitin hydrolysis (Yan and Fong, 2015b).

In this study, were investigated the presence of GH18 family chitinases in transcriptomes of several algae, bryophytes, and pteridophytes species, performed annotation and comparative analysis of these proteins diversity, and explored the phylogenetic relationships and structural characterization. As most chitinase studies are related to superior vascular plants (gymnosperms and angiosperms), the data from this study may be important to help in understanding the relationships of these proteins in the most primitive groups with the most current plants groups.

2 Materials and methods

2.1 *Chitinases proteins identification in plant species*

Initially, a search for plant chitinases amino acid sequences was performed. Selected sequences were used as references in subsequent analyzes. These sequences were obtained from the CAZy database - Carbohydrate Active Enzymes (http://www.cazy.org/Glycoside-Hydrolases.html). Six amino acid GH18 family chitinases sequences (classes III and V) from different plant species were selected. Class III sequences from *Acacia Koa* (AFY08286.1), *Bambusa oldhamii* (ABW75910.1), and *Punica granatum* (PDB: 4TOQ_1) were obtained. From class V, *Arabidopsis thaliana* (PDB: 3AQU_1), *Lotus japonicus* (ANS10044.1), and *Momordica charantia* (AAM18075.1) were obtained.

2.2 *Protein database from algae, bryophytes, and pteridophytes*

Protein sequences from different algae, bryophytes, and pteridophytes phyla were obtained on the OneKP platform (https://db.cngb.org/onekp/) and used to build a database. This platform provides predicted amino acid sequence data from transcriptome of plant species. In addition to green algae phylum, algae species from the phyla Euglenophyta, Dinophyta, Glaucophyta, Haptophyta, Chrysophyta, Ochrophyta, and Rhodophyta were also analyzed. In this work, 196 families and 397 species provided by the OneKP platform were analyzed (Table 1).

Table 1. Number of families and species analyzed in different phyla of algae, bryophytes, and pteridophytes.

Groups analyzed	Families	Species
Euglenophyta	1	1
Dinophyta	1	1
Glaucophyta	1	4
Haptophyta	3	3
Chrysophyta	4	5
Ochrophyta	11	24
Rodophyta	16	28
Green Algae	63	156
Bryophytes	58	77
Pteridophytes	38	98
TOTAL	196	397

To identify the putative chitinases present in that bank, an alignment by BLASTp was performed out in the BLAST program using the reference sequences of GH18 plant chitinases.

2.3 *Conserved domains analysis*

Subsequently, a conserved domains analysis was performed on putative chitinases found in algae, bryophytes, and pteridophytes, through alignments by BLASTp and InterProScan using Blast2GO 5.2 program (Conesa et al., 2005). The conserved domains analyzed at this step can be accessed in data bank Pfam (Mistry et al., 2021).For the next stage, only chitinases with the GH18 catalytic domains were analyzed.

2.4 *Chitinase multiple alignments*

Chitinase multiple alignments of each investigated group were performed against reference chitinases. Alignments were performed using MUSCLE method in MEGA 7.0.26 software (Kumar et al., 2016). In this step, conserved domains were also analyzed to verify whether chitinases selected in the previous step contained GH18 chitinase catalytic domain amino acid sequence. Proteins that presented incomplete catalytic domain were discarded from the subsequent analyzes.

2.5 *Phylogenetic Analysis*

Phylogenetic analyzes were performed to classify these proteins into classes III and V of the GH18 family. Phylogenetic trees were built using the MEGA 7.0.26 program (Kumar et al., 2016) using the Neighbor-Joining method (Saitou and Nei, 1987) and Jones-Taylor-Thornton (JTT) model, with phylogeny test by the “bootstrap” method with 500 replicates, (Jones et al., 1992). A phylogenetic tree was elaborated for each analyzed plant group (Supplementary figures). From each group we selected some chitinases sequences and a phylogenetic tree was created to represent the chitinases identified in this study.

2.6 *Modeling of three-dimensional structures*

Three-dimensional structures modeling was performed using the Swiss-Model server (https://swissmodel.expasy.org/). For structures prediction, two amino acid sequences were selected from each analyzed plant group, one from class III and the other from class V.

3 Results and discussion

Using the data provided by the OneKP platform were identified 648 putative GH18 family chitinases with catalytic domain PFAM-IPR001223 (Table 2). Many of the sequences indicated as chitinases were incomplete or even had large deletions in the conserved domain, and therefore we did not use these to perform alignments and phylogenetic trees. Thus, we selected 270 complete chitinase sequences, which were used for classification and phylogenetic analysis. Similarity and E-value data for the amino acid sequences provided by Blast2GO are available in Supplementary Table S1.

Table 2. Number of chitinases found in each plant group, including families, species, and variation in transcripts number by species.

Groups analyzed	Transcripts found	Families	Species	Variation in transcripts number by specie
Euglenophyta	0	0	0	0
Dinophyta	1	1	1	1
Glaucophyta	1	1	1	1
Haptophyta	0	0	0	0
Chrysophyta	1	1	1	1
Ochrophyta	12	2	2	1 – 8
Rhodophyta	8	4	6	1 – 2

Green Algae	162	31	80	1 – 9
Bryophytes	238	51	70	1 – 13
Pteridophytes	225	16	44	1 – 15
TOTAL	648	107	205	

In algae phyla Euglenophyta and Haptophyta, putative chitinases were not identified. In the phyla Dinophyta, Glaucophyta and Chrysophyta, only one chitinase was identified in each phylum. In Rhodophyta (phylum of red algae) 8 chitinases were found and in the phylum Ochrophyta (brown algae) 12 chitinases were found. However, there are few species of these algae groups in database (Table 1), and this may have influenced the low number of identified chitinases.

162 chitinases were found in green algae, 102 of them in charophytes. Although chitinases were identified in the Chlorophyceae and Ulvophyceae classes, as well as in other smaller classes, the Charophyceae class stood out with the highest chitinases presence. Charophyceae (charophytes) is composed mainly of green freshwater algae and is considered the group that has the highest degree of phylogenetic relationships with bryophytes and vascular plants, due to similarities in photosynthetic pigments (chlorophylls *a* and *b* and beta-carotene), in the energy storage in the form of starch and the cell wall constituted by cellulose (Will. H Blackwell, 2003; Nishio, 2000). *Desmediceae* family, from charophytes group, presented 53 chitinases, being the green algae family with the highest amount of chitinases identified. This family consists of microscopic green algae that are found inhabiting all aquatic freshwater environments (Gontcharov and Melkonian, 2008). Within this family, *Bambusina borreri* stood out with the presence of 9 chitinases.

It is likely that green algae, as well as terrestrial plants, use chitinases as a form of defense against arthropod attack and pathogenic fungi that contain chitin. There are many fungi that make symbiotic associations with algae (interaction beneficial to both), however, some aquatic pathogenic fungi can cause disease and affect algae survival (Raghukumar, 2006; Rasconi et al., 2011). Chytridiomycete fungi are algae parasites, and have vegetative and reproductive structures constituted of chitinaceous walls (Rasconi et al., 2011). In addition to being widely present in freshwater regions, there are also many species of fungi inhabiting marine waters and many algae are prevalent in these regions (Amend et al., 2019; Ji and Wang, 2016), this may further reinforce the fact that the presence of chitinases in algae may be related to defense mechanisms against these microorganisms. Aquatic insects also feed on algae, and chitinases can protect against these organisms (Tuno et al., 2018).

Bryophytes are divided into three phyla: Liverworts, Bryophyta and Hornworts (Michelle T Casanova and Nairn, 2016). 238 chitinases were identified in bryophytes, the majority belong to the Liverworts phylum, reaching 125 chitinases. In Hornworts 22 chitinases were found, while in Mosses 87 chitinases were found. In the genus *Marchantia* (Liverworts) the highest chitinases amount was found, reaching 20 chitinases. However, the species *Phaeoceros carolinianus* (Hornworts) presented 15 chitinases, being, therefore, the bryophyte with the highest chitinase amount.

Currently, the pteridophytes group comprises two phyla with existing representatives: Lycophytes and Monilophytes (known as ferns) (Peterson, 1994; Pryer et al., 2001). In total, 225 chitinases were found in pteridophytes. In the Lycophytes phylum 126 chitinases were identified, while in Monilophytes 99 chitinases were identified. *Selaginella* and *Isoestes* genera (Lycophytes) were the ones that presented the highest amount of chitinases, especially *Selaginella wallacei* and *Isoestes tegetiformans* with 15 chitinases each.

In angiosperm plants, *Arabidopsis thaliana* and tomato (*Solanum Lycopersicum*), 10 and 11 GH18 chitinases were identified, respectively (Cao and Tan, 2019; Passarinho and de Vries, 2002). Wang et al. (2015) identified 12 genes from GH18 chitinases in blackberry (Morus L.). In *Eucalyptus grandis*, 36 GH18 chitinases genes were identified (Tobias et al., 2017). In this study, the maximum number of chitinases found in a species of green alga is relatively less than the number of chitinases identified in species of bryophytes and pteridophytes. These numbers are also close to the number of chitinases described in angiosperms in the literature.

3.1 *Chitinases conserved domains analysis*

Although the amino acid sequence similarity between class III and V chitinases is relatively low (approximately 15% identity), both have a fundamental catalytic motive for chitinolytic activity, formed by two aspartate residues and one of glutamate (DxDxE), which is conserved in chitinases of various organisms (Ohnuma et al., 2011). Substitutions or deletions of these conserved residues can interfere with chitinolytic function (Tyler et al., 2010). This catalytic motif was also evidenced in class III and V chitinases identified in this study, however, in some sequences these residues are substituted by other amino acids, assuming that some of the identified chitinases may have lost their chitinolytic function.

Enzymes catalytic domains are highly conserved regions and, therefore, these regions were used as a parameter for chitinases selection in analyzed plant species. Chitinase-binding domains are common to class V chitinases, but it is not present in all classes, because despite being important in the interaction between the chitin polymer and chitinase, catalysis does not have to occur. Chitin-binding domains interact selectively and non-covalently with chitin and can occur in one or more copies, usually having 43 amino acid residues (Lerner and Raikhel, 1992).

Evolutionary studies indicate that chitinases have undergone a rapid evolutionary process. Chitinases acquired evolutionary adaptations, presenting different isoforms, new functions, coming from different organisms (Tyler et al., 2010). Also, several chitinases maintained the chitin bind ability while losing the hydrolysis capacity, which can affect various biological processes (Oyeleye and Normi, 2018). An example is enzymes without conserved catalytic residues, such as nodulins, which are proteins involved in interactions with symbiotic bacteria, as well as narbonins and concanavalin B (seed proteins), which are also grouped within the GH18 family (Tyler et al., 2010).

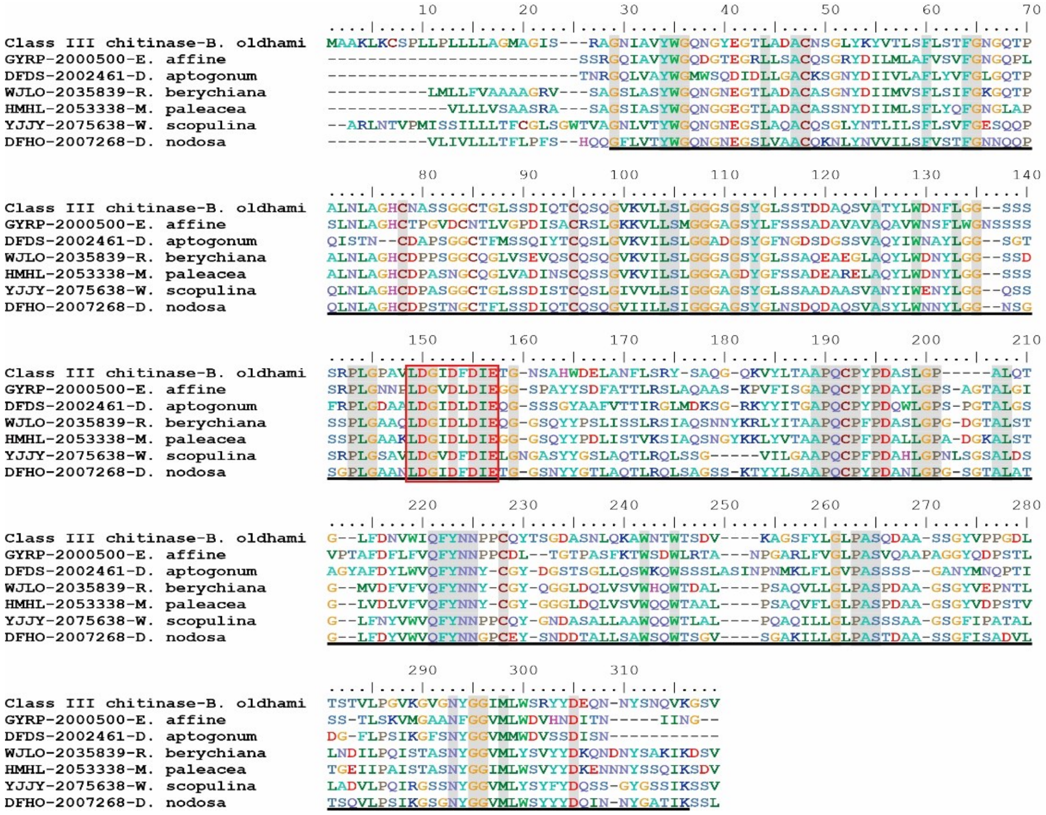
3.1.1 Class III

Class III enzymes have similarities with bacterial chitinases, as they share the same hydrolysis mechanism and similarity in the three-dimensional structure (Iseli et al., 1996). Class III has a peculiar characteristic, and they are considered bifunctional enzymes, because, in addition to hydrolyzing chitin, they also have lysozyme activity, thus being able to contribute against bacterial infections (Park et al., 2002).

Although certain chitinases have highly conserved domains, they can present different substrate specificities. This fact was observed in class III chitinases in rice, in which two chitinases with highly similar sequences had different expression patterns, where one expressed in response to pathogens attack, while the other had its expression induced in response to environmental stresses (Park et al., 2002). Therefore, small changes in the sequences can infer the GH18 chitinases function. Analysis and comparison of chitinase sequences are very important to assist in predicting these proteins function, as well as in the new chitinases groups discovery.

Although GH18 chitinases are divided into classes III and V, they show low similarity to each other (Ebrahim et al., 2011). Thus, the GH18 chitinase alignments were performed separately for each class, thus making it possible to observe differences in the catalytic domains between these classes (Supplementary figures). In this work, 124 complete class III chitinases sequences were identified. Among them, we selected six sequences for the representative class III alignment, two of green algae, two of bryophytes, and two pteridophytes sequences (Fig. 1).

It was possible to observe a high similarity between the class III chitinases even with the *Bambusa oldhamii* chitinase, which was used as a reference in the alignment. Even green algae chitinase (considered the terrestrial plants ancestral group) has shown its sequence very similar to chitinases identified in bryophytes and pteridophytes. In class III, in green algae and mainly in bryophytes, some sequences with the catalytic motif DxDxE were observed, presenting residues replaced by other amino acids. However, in pteridophytes this important region for chitinolytic activity is highly conserved in all analyzed species.



```
                                       10        20        30        40        50        60        70
                               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Class III chitinase-B. oldhami MAAKLKCSPLLPLLLLAGMAGIS---RAGNIAVYWGQNGYEGTLADACNSGLYKYVTLSFLSTFGNGQTP
GYRP-2000500-E. affine         -------------------------SSRGQIAVYWGQDGTEGRLLSACQSGRYDILMLAFVSVFGNGQPL
DFDS-2002461-D. aptogonum      -------------------------TNRGQLVAYWGMWSQDIDLLGACKSGNYDIIVLAFLYVFGLGQTP
WJLO-2035839-R. berychiana     ----------LMLLFVAAAAGRV---SAGSLASYWGQNGNEGTLADACASGNYDIIMVSFLSIFGKGQTP
HMHL-2053338-M. paleacea       ------------VLLLVSAASRA---SAGSIASYWGQNGGEGTLADACASSNYDIIMLSFLYQFGNGLAP
YJJY-2075638-W. scopulina      --ARLNTVPMISSILLLTFCGLSGWTVAGNLVTYWGQNGNEGSLAQACQSGLYNTLILSFLSVFGESQQP
DFHO-2007268-D. nodosa         ----------VLIVLLLTFLPFS--HQQGFLVTYWGQNGNEGSLVAACQKNLYNVVILSFVSTFGNNQQP

                                       80        90       100       110       120       130       140
                               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Class III chitinase-B. oldhami ALNLAGHCNASSGGCTGLSSDIQTCQSQGVKVLLSLGGGSGSYGLSSTDDAQSVATYLWDNFLGG--SSS
GYRP-2000500-E. affine         SLNLAGHCTPGVDCNTLVGPDISACRSLGKKVLLSMGGGAGSYLFSSSADAVAVAQAVWNSFLWGNSSSS
DFDS-2002461-D. aptogonum      QISTN--CDAPSGGCTFMSSQIYTCQSLGVKVILSLGGADGSYGFNGDSDGSSVAQYIWNAYLGG--SGT
WJLO-2035839-R. berychiana     ALNLAGHCDPPSGGCQGLVSEVQSCQSQGVKVILSLGGGSGSYGLSSAQEAEGLAQYLWDNYLGG--SSD
HMHL-2053338-M. paleacea       ALNLAGHCDPASNGCQGLVADINSCQSSGVKVILSLGGGAGDYGFSSADEARELAQYLWDNYLGG--SSS
YJJY-2075638-W. scopulina      QLNLAGHCDPASGGCTGLSSDISTCQSLGIVVLLSIGGGAGSYGLSSAADAASVANYIWENYLGG--QSS
DFHO-2007268-D. nodosa         QLNLAGHCDPSTNGCTFLSSDIQTCQSQGVIILLSIGGGAGSYGLNSDQDAQSVASYLWNNYLGG--NSG

                                      150       160       170       180       190       200       210
                               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Class III chitinase-B. oldhami SRPLGPAVLDGIDFDIETG-NSAHWDELANFLSRY-SAQG-QKVYLTAAPQCPYPDASLGP-----ALQT
GYRP-2000500-E. affine         SRPLGNNELDGVDLDIEGG-SPAYYSDFATTLRSLAQAAS-KPVFISGAPQCPYPDAYLGPS-AGTALGI
DFDS-2002461-D. aptogonum      FRPLGDAALDGIDLDIEQG-SSSGYAAFVTTIRGLMDKSG-RKYYITGAPQCPYPDQWLGPS-PGTALGS
WJLO-2035839-R. berychiana     SSPLGAAQLDGIDLDIEQG-GSQYYPSLISSLRSIAQSNNYKRLYITAAPQCPFPDASLGPG-DGTALST
HMHL-2053338-M. paleacea       SSPLGAAKLDGIDLDIEGG-GSQYYPDLISTVKSIAQSNGYKKLYVTAAPQCPFPDALLGPA-DGKALST
YJJY-2075638-W. scopulina      SRPLGSAVLDGVDFDIELGNGASYYGSLAQTLRQLSSG-----VILGAAPQCPFPDAHLGPNLSGSALDS
DFHO-2007268-D. nodosa         SGPLGAANLDGIDFDIETG-GSNYYGTLAQTLRQLSAGSS-KTYYLSAAPQCPYPDANLGPG-SGTALAT

                                      220       230       240       250       260       270       280
                               ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Class III chitinase-B. oldhami G--LFDNVWIQFYNNPPCQYTSGDASNLQKAWNTWTSDV----KAGSFYLGLPASQDAA-SSGYVPPGDL
GYRP-2000500-E. affine         VPTAFDFLFVQFYNNPPCDL--TGTPASFKTWSDWLRTA---NPGARLFVGLPASVQAAPAGGYQDPSTL
DFDS-2002461-D. aptogonum      AGYAFDYLWVQFYNNY-CGY-DGSTSGLLQSWKQWSSSLASINPNMKLFLGVPASSSS--GANYMNQPTI
WJLO-2035839-R. berychiana     G--MVDFVFVQFYNNY-CGY-QGGLDQLVSVWHQWTDAL----PSAQVLLGLPASPDAA-GSGYVEPNTL
HMHL-2053338-M. paleacea       G--LVDLVFVQFYNNY-CGY-GGGLDQLVSVWQQWTAAL----PSAQVFLGLPASPDAA-GSGYVDPSTV
YJJY-2075638-W. scopulina      G--LFNYVWVQFYNNPPCQY-GNDASALLAAWQQWTLAL----PQAQILLGLPASSSAA-GSGFIPATAL
DFHO-2007268-D. nodosa         G--LFDYVWVQFYNNGPCEY-SNDDTALLSAWSQWTSGV----SGAKILLGLPASTDAA-SSGFISADVL

                                      290       300       310
                               ....|....|....|....|....|....|....|....
Class III chitinase-B. oldhami TSTVLPGVKGVGNYGGIMLWSRYYDEQN-NYSNQVKGSV
GYRP-2000500-E. affine         SS-TLSKVMGAANFGGVMLWDVHNDITN-----IING--
DFDS-2002461-D. aptogonum      DG-FLPSIKGFSNYGGVMMWDVSSDISN-----------
WJLO-2035839-R. berychiana     LNDILPQISTASNYGGVMLYSVYYDKQNDNYSAKIKDSV
HMHL-2053338-M. paleacea       TGEIIPAISTASNYGGIMLWSVYYDKENNNYSSQIKSDV
YJJY-2075638-W. scopulina      LADVLPQIRGSSNYGGVMLYSYFYDQSS-GYGSSIKSSV
DFHO-2007268-D. nodosa         TSQVLPSIKGSGNYGGVMLWSYYYDQIN-NYGATIKSSL
```

Fig. 1. Class III chitinase multiple alignments of sequences identified in green algae, bryophytes, and pteridophytes. Shaded regions indicate conserved residues in the species. Region marked by a red rectangle indicates the conserved DxDxE catalytic motif that is fundamental for chitinolytic activity. Regions highlighted in black at the bottom of the alignment indicate the conserved catalytic domains. The identification code GYRP-2000500 indicates a green algae chitinase from *Euastrum affine* and DFDS-2002461 indicates a green algae chitinase from *Desmidium aptogonum*. Codes WJLO-2035839 and HMHL-2053338 indicate chitinases of bryophytes *Riccia berychiana* and *Marchantia paleacea*, respectively. Codes YJJY-2075638 and DFHO-2007268 indicate chitinases identified in pteridophytes *Woodsia scopulina* and *Danaea nodosa*, respectively.

3.1.2 Class V

Ohnuma et al. (2011) observed in *Arabidopsis thaliana* the presence of a conserved motif in the catalytic class V domain (constituted, consecutively, by the amino acids aspartate, leucine, aspartate, tryptophan, and glutamate - DLDWE). This region proved to be important in the process of catalyzing these enzymes and was also observed in class V chitinases of the GH18 family analyzed in this study (Fig. 2). Class V chitinases identified in this study also demonstrated high sequence similarity. In total, were found 140 complete sequences from GH18 class V chitinases. Of these, we selected six sequences to elaborate a representative alignment, two of green algae, two of bryophytes, and two sequences of pteridophytes (Fig. 2). In class V, only one green algae sequence presented the catalytic motif DxDxE with residues replaced by other amino acids. In the other analyzed algal phyla (Rhodophyta and Ochrophyta), in bryophytes and pteridophytes, this region proved to be highly conserved in all analyzed species.

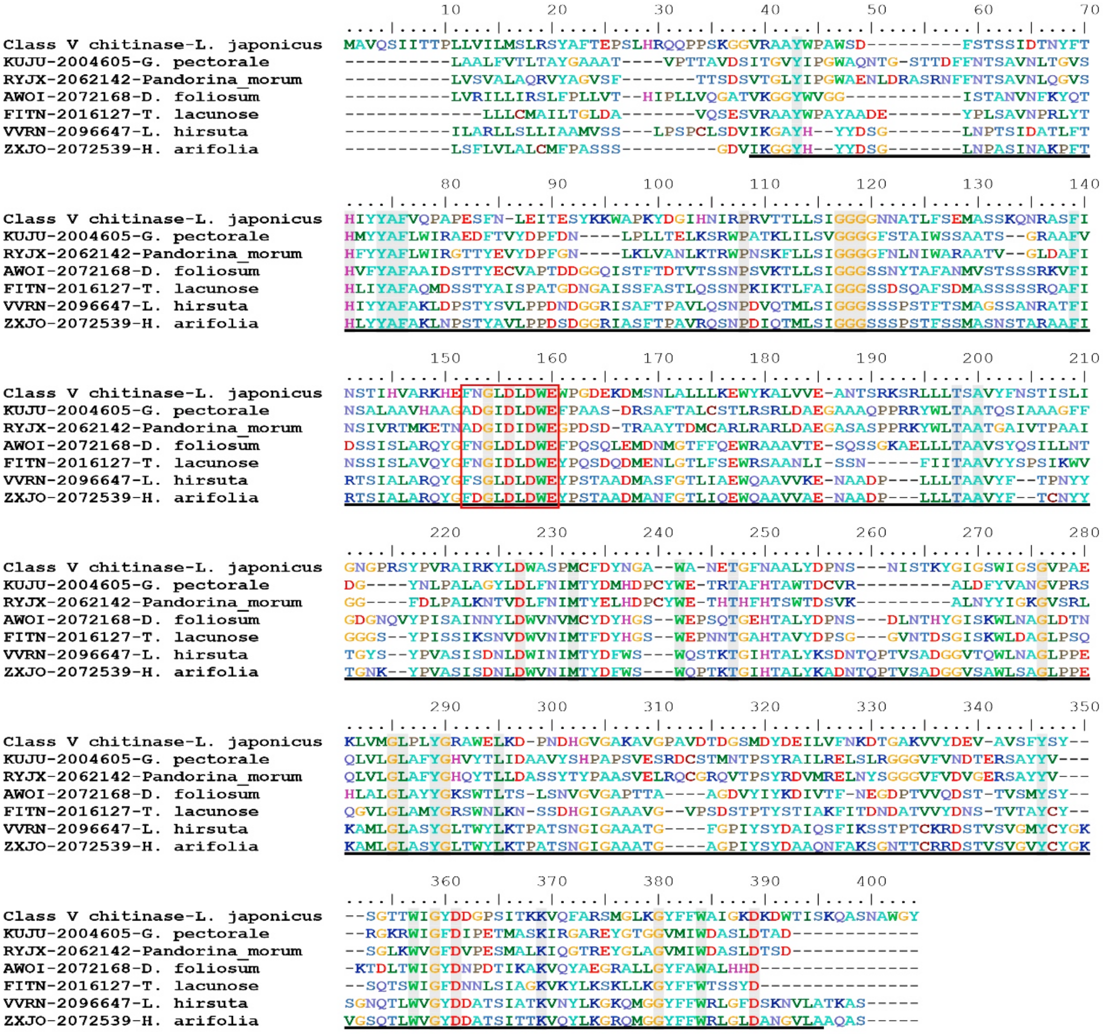


Fig. 2. Class V chitinases multiple alignments of sequences identified in green algae, bryophytes, and pteridophytes. Shaded regions indicate conserved residues. Region marked by a red rectangle indicates the conserved DxDxE catalytic motif that is fundamental for chitinolytic activity. Regions highlighted in black at the bottom of the alignment indicate the conserved catalytic domains. Codes KUJU-2004605 and RYJX-206214 represent green algae chitinases from *Gonium pectorale* and *Pandorina morum*. Codes AWOI-2072168 and FITN-2016127 indicate bryophytes chitinases from *Diphyscium foliosum* and *Treubia lacunose*, respectively. While in pteridophytes code VVRN-2096647 indicates chitinase from *Lonchitis hirsuta* and ZXJO-2072539 indicates *Hemionitis arifolia* chitinase.

3.2 *Phylogenetic relationships and classification of identified chitinases*

270 algae, bryophytes and pteridophytes chitinases sequences were evaluated by phylogenetic analysis. To summarize this result, a phylogenetic tree was made up with 32 complete chitinase sequences, together with the six reference sequences (Fig. 3). Phylogenetic trees with complete sequences allows that all the information contained in the protein sequences can contribute to the phylogenetic positioning (Tyler et al., 2010). Class III chitinases from all groups analyzed were grouped in a different clade from Class V, precisely because they share low sequence similarity. In Fig. 3, it is noted that most sequences were grouped with phylogenetic relationships of the different plant clades (algae, green algae, bryophytes, and pteridophytes). Phylogenetic trees with all the complete sequences identified in algae, green algae, bryophytes, and pteridophytes used for classification are available in Supplementary figures.

We identified three complete chitinases in algae from different phyla, two chitinases in *Sargassum thunbergii* (phylum Ochrophyta), and one chitinase in *Symphyocladia latiuscula* (phylum Rhodophyta). These chitinases were grouped in class V and none were grouped in class III. In the analyzed Chrysophyta, Dinophyta, and Glaucophyta phyla, complete chitinases were not found. This shows that chitinase genes are not present or are not expressed in most primitive algal species, this question would be answered only with complete genomes sequencing of these species.

In the green algae phylum, a much larger chitinases number from GH18 family was identified when compared with the other analyzed algae phyla, with a total of 81 complete chitinases. Through phylogenetic analysis, it was possible to verify that, from these proteins identified in green algae, 34 were grouped in the class III chitinases, while 47 are class V members.

In the phylogenetic analysis of different bryophytes species were found 99 complete chitinases of the GH18 family, greater than the chitinases number identified in green algae. It was possible to observe 42 chitinases were grouped in the class III chitinase branch and 57 were class V chitinases in phylogenetic analysis. In pteridophytes, 87 complete chitinases were identified, 48 were grouped in the class III chitinases and 39 are class V members.

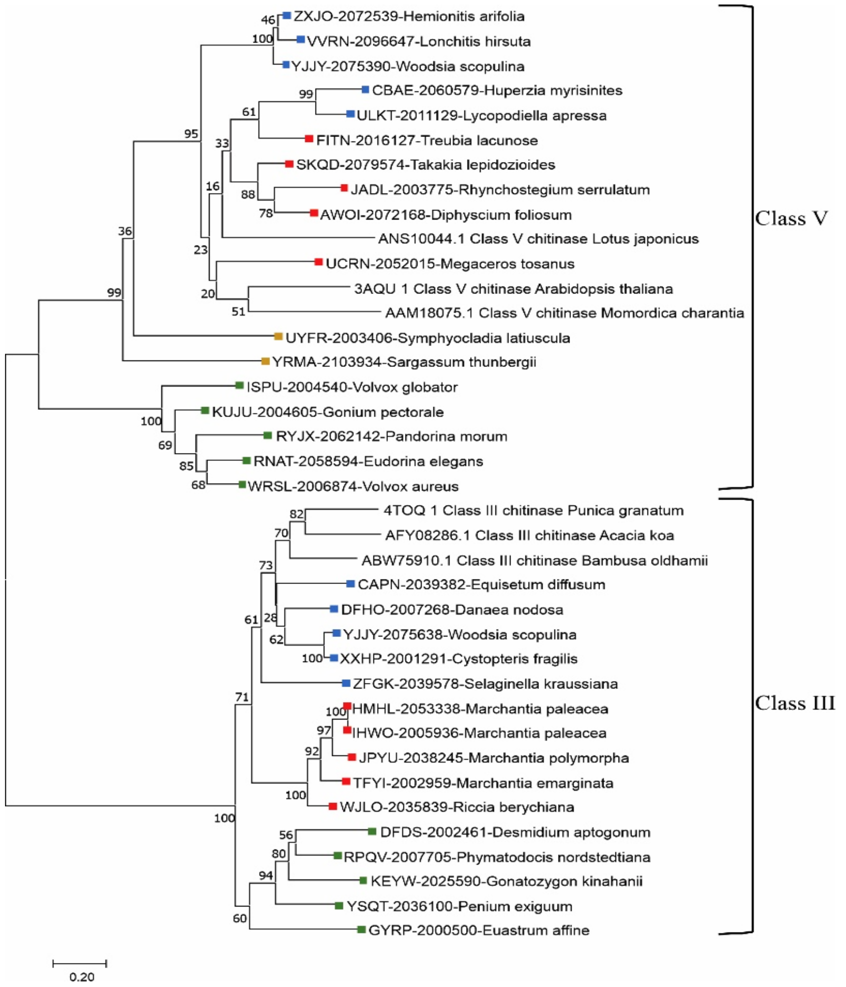


Fig. 3. Phylogenetic relationships of class III and V GH18 family chitinases identified in algae, bryophytes, and pteridophytes. 32 chitinases sequences were selected from the groups analyzed for the development of the phylogenetic tree. Blue squares represent chitinases identified in bryophytes, red squares indicate chitinases identified in pteridophytes, yellow squares indicate chitinases identified in algae while the green squares indicate chitinases in green algae. Angiosperm reference sequences are not indicated by squares.

3.3 *Structural characterization*

Catalytic domains showed a high structure similarity in all analyzed groups (Fig. 4 and 5). According to Fukamizo (2000), structural differences may be associated with the functional chitinases diversity, therefore related to the different physiological functions that these enzymes can perform. Thus, three-dimensional structures can provide important information for chitinases catalytic mechanisms studies, since the enzyme plays its role when the arrangement of the three-dimensional structure occurs properly.

Comparing class III with class V, it is possible to notice a low similarity in the 3D structures analyzed, precisely because classes III and V consist of catalytic domains with low similarity to each other (also, the catalytic domain comprises a large part of the protein chitinase). In both class III and class V structures, it is possible to observe the α-helices are located in the outer region, involving the β- sheets, leading to the barrel shape. Analyzing the class III three-dimensional structures of green algae (Fig. 4A), bryophytes (Fig. 4B) and pteridophytes (Fig. 4C), we observed that these structures are high similarity to each other.

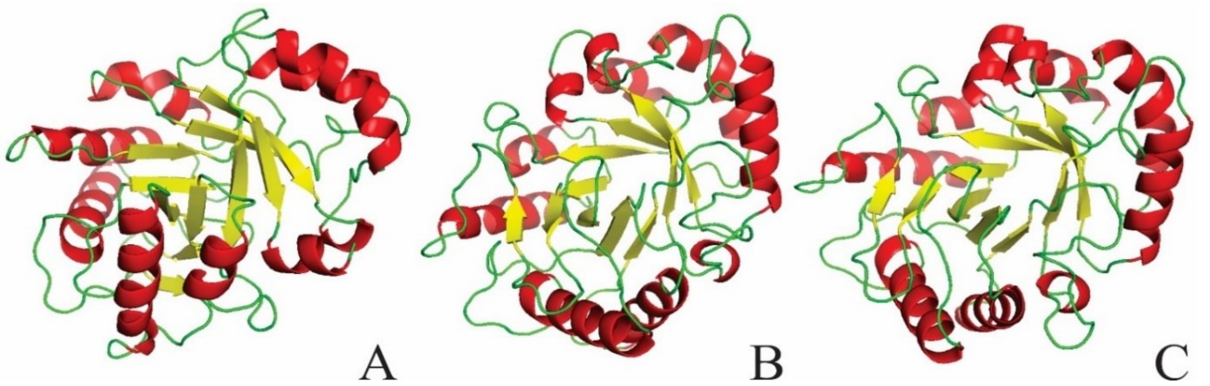


Fig. 4. Structural comparisons of class III GH18 family chitinases. (A) Three-dimensional structure of green algae chitinase *Desmidium aptogonum* DFDS-2002461. (B) Three-dimensional structure of bryophyte chitinase *Riccia berychiana* WJLO-2035839. (C) Three-dimensional structure of pteridophyte chitinase *Woodsia scopulina* (YJJY-2075638). Structures were modeled using the Swiss-Model server, highlighting the yellow β- sheets and red α-helices.

In class V (Fig. 5), three-dimensional structures of green algae (Fig. 5A) bryophytes (Fig. 5B) and pteridophytes (Fig. 5C) also are highly similar to each other. The arrangements of α-helices and β-sheets proved to be similar in bryophytes and pteridophytes class V chitinases.

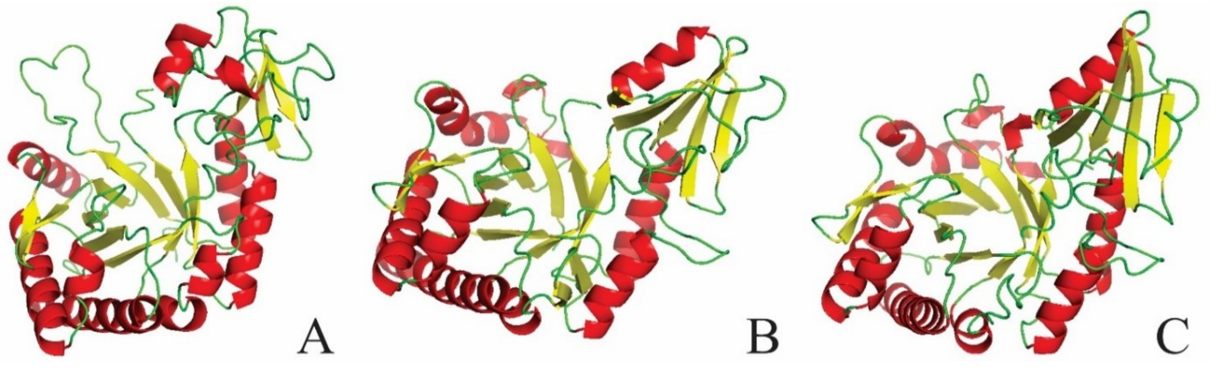


Fig. 5. Structural comparisons of class V GH18 family chitinases. (A) Three-dimensional structure of green algae chitinase *Gonium pectorale* KUJU-2004605. (B) Three-dimensional structure of bryophyte chitinase *Diphyscium foliosum* AWOI-2072168. (C) Three-dimensional structure of pteridophyte chitinase *Lonchitis hirsuta* VVRN-2096647. Structures were modeled using the Swiss-Model server, highlighting the yellow β- sheets and red α-helices.

4 Conclusions

In this study, we investigated chitinase presence in the most ancient terrestrial plants groups, which are algae, bryophytes, and pteridophytes, since most studies of chitinases are reported in higher vascular plants. A large chitinases number from the GH18 family were found in green algae, bryophytes, and pteridophytes species. However, only three chitinases were identified in two algae phyla, Rhodophyta and Ochrophyta. In the green algae, bryophytes and pteridophytes, a high number of chitinases distributed in different species was identified. Chitinases with highly conserved catalytic domains were identified in all groups analyzed, but some sequences showed low similarity and even changes in the catalytic residues essential for catalysis. Phylogenetic tree separated class III from class V because despite being part of the same family, these classes have low similarity to each other. Three-dimensional structures analyzed of green algae, bryophytes and pteridophytes showed high similarity to each other. Chitinases identification and characterization are important to understand the relationships of these proteins with current plant groups since studies suggest these proteins have undergone a rapid evolutionary process, leading to new functions emergence or even the loss of chitinolytic function. The results of this study can also serve to provide data for future studies related to GH18 family chitinases.

Acknowledgment

This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES - Brazil (CAPES). The authors are grateful to LBAM/UFU for providing the infrastructure for the development of this work.

This research did not receive any specific subsidies from funding agencies in the public, commercial, or non-profit sectors.

Declaration of competing interest: The authors declare no conflict of interest.

References

Amend, A., Burgaud, G., Cunliffe, M., Edgcomb, V.P., Ettinger, C.L., Gutiérrez, M.H., Heitman, J., Hom, E.F.Y., Ianiri, G., Jones, A.C., Kagami, M., Picard, K.T., Quandt, C.A., Raghukumar, S., Riquelme, M., Stajich, J., Vargas-Muñiz, J., Walker, A.K., Yarden, O., Gladfelter, A.S., 2019. Fungi in the marine environment: Open questions and unsolved problems. MBio 10, 1–15. https://doi.org/10.1128/mBio.01189-18

Arie, M., Hikichi, K., Takahashi, K., Esaka, M., 2000. Characterization of a basic chitinase which is secreted by cultured pumpkin cells. Physiol. Plant. 110, 232–239. https://doi.org/10.1034/j.1399-3054.2000.110213.x

Badariotti, F., Kypriotou, M., Lelong, C., Dubos, M.P., Renard, E., Galera, P., Favrel, P., 2006. The phylogenetically conserved molluscan chitinase-like protein 1 (Cg-Clp1), homologue of human HC-gp39, stimulates proliferation and regulates synthesis of extracellular matrix components of mammalian chondrocytes. J. Biol. Chem. 281, 29583–29596. https://doi.org/10.1074/jbc.M605687200

Bennici, A., 2008. Origin and early evolution of land plants. Commun. Integr. Biol. 1, 212–218. https://doi.org/10.4161/cib.1.2.6987

Blackwell, Will. H, 2003. Two Theories of Origin of the Land-Plant Sporophyte: Which Is Left Sanding? Bot. Rev. 69, 125–148. https://doi.org/https://doi.org/10.1663/0006-8101(2003)069[0125:TTOOOT]2.0.CO;2

Brameld, Ken A., Shrader, W.D., Imperiali, B., Goddard, W.A., 1998. Substrate assistance in the mechanism of family 18 chitinases: Theoretical studies of potential intermediates and inhibitors. J. Mol. Biol. 280, 913–923. https://doi.org/10.1006/jmbi.1998.1890

Brzezinska, M. S., Jankiewicz, U., Burkowska, A., Walczak, M., 2014. Chitinolytic microorganisms and their possible application in environmental protection. Curr. Microbiol. 68, 71–81. https://doi.org/10.1007/s00284-013-0440-4

Cao, J., Tan, X., 2019. Comprehensive analysis of the chitinase family genes in tomato (Solanum Lycopersicum). Plants 8. https://doi.org/10.3390/plants8030052

Casanova, Michelle T., Nairn, L.C., 2016. Macroalgae, charophytes and bryophytes. Veg. Aust. Riverine Landscapes.

Chen, L., Jiang, H., Cheng, Q., Chen, J., Wu, G., Kumar, A., Sun, M., Liu, Z., 2015. Enhanced nematicidal potential of the chitinase pachi from Pseudomonas aeruginosa in association with Cry21Aa. Sci. Rep. 5, 1–11. https://doi.org/10.1038/srep14395

Christenhusz, M.J.M., Byng, J.W., 2016. Phytotaxa. Phytotaxa 261, 201–217.

https://doi.org/10.11646/phytotaxa.261.3.1

Chuang, H.H., Lin, H.Y., Lin, F.P., 2008. Biochemical characteristics of C-terminal region of recombinant chitinase from Bacillus licheniformis- implication of necessity for enzyme properties. FEBS J. 275, 2240–2254. https://doi.org/10.1111/j.1742-4658.2008.06376.x

Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21, 3674–3676. https://doi.org/10.1093/bioinformatics/bti610

Cooper, Endymion D., 2014. Overly simplistic substitution models obscure green plant phylogeny. Trends Plant Sci. 19, 576–582. https://doi.org/10.1016/j.tplants.2014.06.006

Medeiros, S.C., Monteiro-Júnior, J.E., Passos Sales, G.W., Grangeiro, T.B., Pinto Nogueira, N.A., 2018. Chitinases as antibacterial proteins: A systematic review. J. Young Pharm. 10, 144–148. https://doi.org/10.5530/jyp.2018.10.33

Ebrahim, S., Usha, K., Singh, B., 2011. Pathogenesis-related (PR)-proteins: Chitinase and β-1,3-glucanase in defense mechanism against malformation in mango (Mangifera indica L.). Sci. Hortic. (Amsterdam). 130, 847–852. https://doi.org/10.1016/j.scienta.2011.09.014

Edreva, A., 2005. Pathogenesis-related proteins: research progress in the last 15 years. Gen Appl Plant Physiol 31, 105–24.

El-Katatny, M.H., Gudelj, M., Robra, K.H., Elnaghy, M.A., Gübitz, G.M., 2001. Characterization of a chitinase and an endo-β-1,3-glucanase from Trichoderma harzianum Rifai T24 involved in control of the phytopathogen Sclerotium rolfsii. Appl. Microbiol. Biotechnol. 56, 137–143. https://doi.org/10.1007/s002530100646

Frahm, J.P., Pócs, T., O'Shea, B., Koponen, T., Piipo, S., Enroth, J., Rao, P., Fang, Y., 2003. Manual of tropical bryology. Trop. Bryol. 23, 1–196. https://doi.org/10.11646/bde.23.1.1

Fukamizo, T., 2000. Chitinolytic Enzymes: Catalysis, Substrate Binding, and their Application. Curr. Protein Pept. Sci. 1, 105–124. https://doi.org/10.2174/1389203003381450

Gontcharov, A.A., Melkonian, M., 2008. In search of monophyletic taxa in the family Desmidiaceae (Zygnematophyceae, Viridiplantae): The genus Cosmarium. Am. J. Bot. 95, 1079–1095. https://doi.org/10.3732/ajb.0800046

Guiry, M.D., 2012. How many species of algae are there? J. Phycol. 48, 1057–1063. https://doi.org/10.1111/j.1529-8817.2012.01222.x

Hamid, R., Khan, M.., Ahmed, M., Ahmad, M.M., Abdin, M.Z., Musarrat, J., Javed, S., 2013. Chitinases: An update. J Pharm Bioallied Sci 5, 21–29. https://doi.org/10.4103/0975-

7406.106559

Iseli, B., Armand, S., Boller, T., Neuhaus, J.M., Henrissat, B., 1996. Plant chitinases use two different hydrolytic mechanisms. FEBS Lett. 382, 186–188. https://doi.org/10.1016/0014-5793(96)00174-3

Ji, N.Y., Wang, B.G., 2016. Mycochemistry of marine algicolous fungi. Fungal Divers. 80, 301–342. https://doi.org/10.1007/s13225-016-0358-9

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Bioinformatics 8, 275–282. https://doi.org/https://doi.org/10.1093/bioinformatics/8.3.275

Kaur, R., Neetu, Mudgal, R., Jose, J., Kumar, P., Tomar, S., 2019. Glycan-dependent chikungunya viral infection divulged by antiviral activity of NAG specific chi-like lectin. Virology 526, 91–98. https://doi.org/10.1016/j.virol.2018.10.009

Kesari, P., Patil, D.N., Kumar, Pramod, Tomar, S., Sharma, A.K., Kumar, Pravindra, 2015. Structural and functional evolution of chitinase-like proteins from plants. Proteomics 15, 1693–1705. https://doi.org/10.1002/pmic.201400421

Khan, F.I., Bisetty, K., Singh, S., Permaul, K., Hassan, M.I., 2015. Chitinase from Thermomyces lanuginosus SSBP and its biotechnological applications. Extremophiles 19, 1055–1066. https://doi.org/10.1007/s00792-015-0792-8

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054

Lerner, D.R., Raikhel, N. V, 1992. The gene for stinging nettle lectin (Urtica dioica agglutinin) encodes both a lectin and a chitinase. J. Biol. Chem. 267, 11085–11091.

Li, J., Liu, J., 2003. A Novel cotton gene encoding a new class of chitinase. Acta Bot. Sin. 45, 1489–1496.

Martínez-Caballero, S., Cano-Sánchez, P., Mares-Mejía, I., Díaz-Sánchez, A.G., Macías-Rubalcava, M.L., Hermoso, J.A., Rodríguez-Romero, A., 2014a. Comparative study of two GH19 chitinase-like proteins from Hevea brasiliensis, one exhibiting a novel carbohydrate-binding domain. FEBS J. 281, 4535–4554. https://doi.org/10.1111/febs.12962

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., Bateman, A., 2021. Pfam: The protein families database in 2021. Nucleic Acids Res. 49, D412–D419. https://doi.org/10.1093/nar/gkaa913

Neuhaus, J.-M., Fritig, B., Linthorst, H.J.., Meins, J., Mikkelsen, J.., Ryals, J., 1996. A Revised Nomenclature for Chitinase Genes. Plant Mol. Biol. Report. 14, 102–104. https://doi.org/https://doi.org/10.1007/BF02684897

Nishio, J.N., 2000. Why are higher plants green? Evolution of the higher plant photosynthetic pigment complement. Plant, Cell Environ. 23, 539–548. https://doi.org/10.1046/j.1365-3040.2000.00563.x

Ohnuma, T., Numata, T., Osawa, T., Mizuhara, M., Lampela, O., Juffer, A.H., Skriver, K., Fukamizo, T., 2011. A class V chitinase from Arabidopsis thaliana: Gene responses, enzymatic properties, and crystallographic analysis. Planta 234, 123–137. https://doi.org/10.1007/s00425-011-1390-3

Oyeleye, A., Normi, Y.M., 2018. Chitinase: Diversity, limitations, and trends in Engineering for suitable applications. Biosci. Rep. 38, 1–21. https://doi.org/10.1042/BSR20180323

Park, S.M., Kim, D.H., Truong, N.H., Itoh, Y., 2002. Heterologous expression and characterization of class III chitinases from rice (Oryza sativa L.). Enzyme Microb. Technol. 30, 697–702. https://doi.org/10.1016/S0141-0229(02)00042-X

Passarinho, P.A., de Vries, S.C., 2002. Arabidopsis Chitinases : a Genomic Survey . Arab. B. 1, e0023. https://doi.org/10.1199/tab.0023

Peterson, K.J., 1994. The Origin and Early Evolution of the Craniata. Short Courses Paleontol. 7, 14–37. https://doi.org/10.1017/s2475263000001240

Pittermann, J., 2010. The evolution of water transport in plants: An integrated approach. Geobiology 8, 112–139. https://doi.org/10.1111/j.1472-4669.2010.00232.x

Pryer, K.M., Schneider, H., Smith, A.R., Cranfill, R., Wolf, P.G., Hunt, J.S., Sipes, S.D., 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409, 618–621. https://doi.org/10.1038/35054555

Raghukumar, C., 2006. Algal-fungal interactions in the marine ecosystem: Symbiosis to parasitism. Recent Adv. Appl. Asp. Indian Mar. Algae with Ref. to Glob. Scenar. 1, 366–385.

Rasconi, S., Jobard, M., Sime-Ngando, T., 2011. Parasitic fungi of phytoplankton: Ecological roles and implications for microbial food webs. Aquat. Microb. Ecol. 62, 123–137. https://doi.org/10.3354/ame01448

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Seidl, V., 2008. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. Fungal Biol. Rev. 22, 36–42. https://doi.org/10.1016/j.fbr.2008.03.002

Seidl, V., Huemer, B., Seiboth, B., Kubicek, C.P., 2005. A complete survey of Trichoderma chitinases reveals three distinct subgroups of family 18 chitinases. FEBS J. 272, 5923–5939. https://doi.org/10.1111/j.1742-4658.2005.04994.x

Tobias, P.A., Christie, N., Naidoo, S., Guest, D.I., Külheim, C., 2017. Identification of the Eucalyptus grandis chitinase gene family and expression characterization under different biotic stress challenges. Tree Physiol. 37, 565–582. https://doi.org/10.1093/treephys/tpx010

Tryon, R. M., Tryon, A.F., 1982. Ferns and allied Plants with Special reference to Tropical America. Springer-Verlag, New York. https://doi.org/10.2307/1546858

Tuno, N., Kohzu, A., Tayasu, I., Nakayama, T., Githeko, A., Yan, G., 2018. An Algal Diet Accelerates Larval Growth of Anopheles gambiae (Diptera: Culicidae) and Anopheles arabiensis (Diptera: Culicidae). J. Med. Entomol. https://doi.org/10.1093/jme/tjx244

Tyler, L., Bragg, J.N., Wu, J., Yang, X., Tuskan, G.A., Vogel, J.P., 2010. Annotation and comparative analysis of the glycoside hydrolase genes in Brachypodium distachyon. BMC Genomics 11, 600. https://doi.org/10.1186/1471-2164-11-600

Vecchi, V., Barera, S., Bassi, R., Dall'osto, L., 2020. Potential and challenges of improving photosynthesis in algae. Plants 9. https://doi.org/10.3390/plants9010067

Wang, X., He, N., Zeng, Q., Xiang, Z., 2015. Identification and expression analyses of chitinase genes in mulberry (Morus L.) plants. Plant Omics 8, 183–189.

Yan, Q., Fong, S.S., 2015. Bacterial chitinase: Nature and perspectives for sustainable bioproduction. Bioresour. Bioprocess. 2. https://doi.org/10.1186/s40643-015-0057-5

Supplementary material

Table S1. Blast2GO results of algae phyla, bryophytes and pteridophytes putative GH18 chitinases

ALGAE – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-EBWI-2035542-Ochromonas_sp	Chrysophyta	Chromulinaceae	glycoside hydrolase	1,59292E-18	71.42	IPR001223 (PFAM)
scaffold-IEHF-2003601-Dumontia_simplex	Rodophyta	Dumontiaceae	Chitinase	5,62296E-74	58.65	IPR001223 (PFAM)
scaffold-IKIZ-2012657-Grateloupia_livida	Rodophyta	Halymeniaceae	Chitinase	9,18487E-100	73.3	IPR001223 (PFAM)
scaffold-JEBK-2004603-Eucheuma_denticulatum	Rodophyta	Solieriaceae	glycoside hydrolase family 18 protein	1,61585E-78	72.99	IPR001223 (PFAM)
scaffold-POOW-2046963-Glaucocystis_cf._nostochinearum	Glaucophyta	Glaucocystaceae	chitotriosidase-1 precursor	1,80205E-27	68.94	IPR001223 (PFAM)
scaffold-PYDB-2002012-Sinotubimorpha_guangdongensis	Rodophyta	Halymeniaceae	probable chitinase 10	1,34808E-42	72.51	IPR001223 (PFAM)
scaffold-PYDB-2020269-Sinotubimorpha_guangdongensis	Rodophyta	Halymeniaceae	acidic mammalian chitinase-like	7,6859E-46	83.85	IPR001223 (PFAM)
scaffold-RAPY-2001279-Kjellmaniella_crassifolia	Ochrophyta	Laminariaceae	chitinase 4	6,70733E-48	80.81	IPR001223 (PFAM)
scaffold-RAPY-2073914-Kjellmaniella_crassifolia	Ochrophyta	Laminariaceae	chitinase 4 precursor	5,90828E-32	71.96	IPR001223 (PFAM)
scaffold-RAPY-2075238-Kjellmaniella_crassifolia	Ochrophyta	Laminariaceae	acidic mammalian chitinase-like	1,67502E-73	66.99	IPR001223 (PFAM)
scaffold-RAPY-2078410-Kjellmaniella_crassifolia	Ochrophyta	Laminariaceae	Chitinase	1,24398E-42	65.31	IPR001223 (PFAM)
scaffold-TZJQ-2121870-Prorocentrum_micans	Dinophyta	Prorocentraceae	Acidic mammalian chitinase	5,74682E-73	57.66	IPR001223 (PFAM)
scaffold-UYFR-2003406-Symphyocladia_latiuscula	Rodophyta	Rhodomelaceae	acidic mammalian chitinase-like	8,55986E-137	71.38	IPR001223 (PFAM)
scaffold-UYFR-2004651-Symphyocladia_latiuscula	Rodophyta	Rhodomelaceae	chitotriosidase-1	5,84526E-47	75.74	IPR001223 (PFAM)
scaffold-YRMA-2002440-Sargassum_thunbergii	Ochrophyta	Sargassaceae	AChain A	1,77261E-38	72.75	IPR001223 (PFAM)
scaffold-YRMA-2005642-Sargassum_thunbergii	Ochrophyta	Sargassaceae	acidic mammalian chitinase-like	4,17761E-16	77.42	IPR001223 (PFAM)
scaffold-YRMA-2017852-Sargassum_thunbergii	Ochrophyta	Sargassaceae	chitinase-3-like protein 1	2,84134E-165	75.78	IPR001223 (PFAM)
scaffold-YRMA-2097681-Sargassum_thunbergii	Ochrophyta	Sargassaceae	probable chitinase 10	1,19293E-53	81.17	IPR001223 (PFAM)
scaffold-YRMA-2102570-Sargassum_thunbergii	Ochrophyta	Sargassaceae	chitinase-3-like protein 1	3,97527E-56	59.96	IPR001223 (PFAM)
scaffold-YRMA-2103020-Sargassum_thunbergii	Ochrophyta	Sargassaceae	AChain A	4,37781E-68	64.43	IPR001223 (PFAM)
scaffold-YRMA-2103934-Sargassum_thunbergii	Ochrophyta	Sargassaceae	chitotriosidase-1 precursor	2,11977E-109	62.96	IPR001223 (PFAM)
scaffold-YRMA-2105318-Sargassum_thunbergii	Ochrophyta	Sargassaceae	acidic mammalian chitinase isoform x2	2,49986E-76	57.43	IPR001223 (PFAM)
scaffold-ZJOJ-2054710-Grateloupia_filicina	Rodophyta	Halymeniaceae	Chitinase	1,63683E-101	73.14	IPR001223 (PFAM)

GREEN ALGAE – GH18					
SeqName	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-ACRY-2005832-Pteromonas_sp	Phacotaceae	acidic mammalian chitinase-like	2,8735E-30	58.2	IPR001223 (PFAM)
scaffold-AEKF-2009302-Penium_margaritaceum	Peniaceae	class V chitinase	1,05668E-43	54.91	IPR001223 (PFAM)
scaffold-AEKF-2044086-Penium_margaritaceum	Peniaceae	chitinase-like protein 4	3,20758E-10	59.12	IPR001223 (PFAM)
scaffold-BAZF-2007777-Chaetopeltis_orbicularis	Chaetopeltidaceae	Chitinase	1,5718E-64	54.95	IPR001223 (PFAM)
scaffold-BAZF-2016149-Chaetopeltis_orbicularis	Chaetopeltidaceae	chitotriosidase-1	2,15302E-59	5574	IPR001223 (PFAM)
scaffold-BFIK-2029987-Entransia_fimbriat	Klebsormidiaceae	acidic mammalian chitinase-like	1,18834E-39	52.89	IPR001223 (PFAM)
scaffold-BHBK-2048640-Cosmarium_tinctum	Desmidiaceae	brain chitinase and chia	9,84079E-102	52.91	IPR001223 (PFAM)
scaffold-DFDS-2000769-Desmidium_aptogonum	Desmidiaceae	hevamine-A-like	2,44497E-40	69.4	IPR001223 (PFAM)
scaffold-DFDS-2002461-Desmidium_aptogonum	Desmidiaceae	acidic endochitinase	1,09525E-62	61.1	IPR001223 (PFAM)
scaffold-DFDS-2002462-Desmidium_aptogonum	Desmidiaceae	hevamine-A-like	4,3324E-57	62.79	IPR001223 (PFAM)
scaffold-DRFX-2014366-Closterium_lunula	Closteriaceae	acidic mammalian chitinase-like	8,84728E-36	56.07	IPR001223 (PFAM)
scaffold-DZPJ-2054002-Cylindrocapsa_geminella	Mesotaeniaceae	Chitinase	2,88943E-86	55.57	IPR001223 (PFAM)
scaffold-EEJO-2039453-Neochloris_oleoabundans	incertae sedis	brain chitinase and chia	5,77463E-142	58.81	IPR001223 (PFAM)
scaffold-ETGN-2057676-Botryococcus_braunii	Botryococcaceae	Glycoside hydrolase	4,01752E-85	88.76	IPR001223 (PFAM)
scaffold-FFGR-2004388-Netrium_digitus	Mesotaeniaceae	acidic endochitinase	1,59149E-41	63.63	IPR001223 (PFAM)
scaffold-FMVB-2005958-Scherffelia_dubia	Chlorodendraceae	Chitinase	3,79911E-59	70.48	IPR001223 (PFAM)
scaffold-FOYQ-2002254-Microspora_cf._tumidula	Microsporaceae	glycoside hydrolase	6,12419E-37	52.66	IPR001223 (PFAM)
scaffold-FXHG-2079786-Hafniomonas_reticulata	Dunaliellaceae	chitinase	9,13289E-47	48.39	IPR001223 (PFAM)
scaffold-GBGT-2000381-Xanthidium_antilopaeum	Desmidiaceae	hevamine-A-like	1,44545E-20	67.86	IPR001223 (PFAM)
scaffold-GBGT-2002008-Xanthidium_antilopaeum	Desmidiaceae	acidic mammalian chitinase-like	2,64665E-18	61.55	IPR001223 (PFAM)
scaffold-GBGT-2002009-Xanthidium_antilopaeum	Desmidiaceae	acidic mammalian chitinase-like	2,64665E-18	61.55	IPR001223 (PFAM)
scaffold-GBGT-2024109-Xanthidium_antilopaeum	Desmidiaceae	acidic mammalian chitinase-like	1,86839E-31	56.96	IPR001223 (PFAM)
scaffold-GBGT-2103934-Xanthidium_antilopaeum	Desmidiaceae	class V chitinase-like	1,57782E-27	60.2	IPR001223 (PFAM)
scaffold-GFUR-2016925-Chloromonas_subdivisa	Chlamydomonadaceae	acidic mammalian chitinase	1,33954E-68	49.85	IPR001223 (PFAM)
scaffold-GGWH-2009903-Onychonema_laeve	Desmidiaceae	Acidic endochitinase	3,07289E-76	63.74	IPR001223 (PFAM)
scaffold-GGWH-2048506-Onychonema_laeve	Desmidiaceae	acidic endochitinase-like	2,64921E-76	64.83	IPR001223 (PFAM)
scaffold-GYRP-2000499-Euastrum_affine	Desmidiaceae	acidic endochitinase-like	1,22857E-63	63.15	IPR001223 (PFAM)
scaffold-GYRP-2000500-Euastrum_affine	Desmidiaceae	acidic endochitinase-like	7,68057E-65	63.71	IPR001223 (PFAM)
scaffold-GYRP-2000501-Euastrum_affine	Desmidiaceae	acidic endochitinase-like	1,22857E-63	63.15	IPR001223 (PFAM)
scaffold-GYRP-2000502-Euastrum_affine	Desmidiaceae	acidic endochitinase-like	2,9316E-43	63.14	IPR001223 (PFAM)
scaffold-GYRP-2000503-Euastrum_affine	Desmidiaceae	acidic endochitinase-like	9,96149E-60	57.95	IPR001223 (PFAM)
scaffold-HAOX-2002128-Spirogyra_sp	Zygnemataceae	acidic endochitinase	1,39013E-74	64.29	IPR001223 (PFAM)
scaffold-HAOX-2024957-Spirogyra_sp	Zygnemataceae	acidic endochitinase	2,31867E-74	64.45	IPR001223 (PFAM)
scaffold-HHXJ-2002324-Tetraselmis_striata	Chlorodendraceae	chitinase	1,77379E-46	67.76	IPR001223 (PFAM)
scaffold-HIDG-2009902-Cosmarium_broomei	Desmidiaceae	AChain A	1,58101E-64	57.68	IPR001223 (PFAM)

GREEN ALGAE – GH18					
SeqName	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-HIDG-2009903-Cosmarium_broomei	Desmidiaceae	AChain A	9,26631E-65	57.67	IPR001223 (PFAM)
scaffold-HIDG-2009904-Cosmarium_broomei	Desmidiaceae	acidic mammalian chitinase-like	2,52567E-52	59.5	IPR001223 (PFAM)
scaffold-HIDG-2018628-Cosmarium_broomei	Desmidiaceae	hevamine-A-like	3,17674E-37	63.65	IPR001223 (PFAM)
scaffold-HJVM-2001428-Cosmarium_ochthodes	Desmidiaceae	Acidic endochitinase	3,37497E-83	66.04	IPR001223 (PFAM)
scaffold-HJVM-2105402-Cosmarium_ochthodes	Desmidiaceae	acidic mammalian chitinase-like	8,89255E-17	60.86	IPR001223 (PFAM)
scaffold-HJVM-2110782-Cosmarium_ochthodes	Desmidiaceae	acidic endochitinase-like	2,68544E-77	61.92	IPR001223 (PFAM)
scaffold-HKZW-2012228-Mesotaenium_caldariorum	Mesotaeniaceae	acidic mammalian chitinase	4,5789E-54	57.22	IPR001223 (PFAM)
scaffold-HVNO-2013597-Tetraselmis_chui	Chlorodendraceae	chitinase	4,62445E-61	67.87	IPR001223 (PFAM)
scaffold-IRYH-2007489-Heterochlamydomonas_inaequalis	Chlamydomonadaceae	probable chitinase 3	6,19725E-103	52.09	IPR001223 (PFAM)
scaffold-IRYH-2025398-Heterochlamydomonas_inaequalis	Chlamydomonadaceae	acidic mammalian chitinase-like	1,42798E-21	61.8	IPR001223 (PFAM)
scaffold-ISHC-2004481-Staurastrum_sebaldi	Desmidiaceae	hevamine-A-like	1,01518E-65	64.11	IPR001223 (PFAM)
scaffold-ISHC-2004482-Staurastrum_sebaldi	Desmidiaceae	acidic endochitinase-like	3,14706E-67	63.55	IPR001223 (PFAM)
scaffold-SIM-2003838-Nephroselmis_pyriformis	Nephroselmidaceae	acidic mammalian chitinase-like	5,78804E-60	62.12	IPR001223 (PFAM)
scaffold-ISPU-2004539-Volvox_globator	Volvocaceae	chitotriosidase-1 isoform X2	1,95846E-137	53.24	IPR001223 (PFAM)
scaffold-ISPU-2004540-Volvox_globator	Volvocaceae	chitinase	2,16629E-125	54.03	IPR001223 (PFAM)
scaffold-JKKI-2008476-Lobomonas_rostrata	Chlamydomonadaceae	glycoside hydrolase family 18	5,491E-28	54.53	IPR001223 (PFAM)
scaffold-JMTE-2014447-Pseudoscourfieldia_marina	Pycnococcaceae	chitotriosidase-1-like	3,6133E-39	74.11	IPR001223 (PFAM)
scaffold-JMTE-2014781-Pseudoscourfieldia_marina	Pycnococcaceae	chitotriosidase-1-like isoform X2	1,10422E-33	80.8	IPR001223 (PFAM)
scaffold-JMTE-2015199-Pseudoscourfieldia_marina	Pycnococcaceae	acidic endochitinase-like	1,2356E-47	85.06	IPR001223 (PFAM)
scaffold-JOJQ-2005804-Cylindrocystis_cushleckae	Zygnemataceae	acidic endochitinase-like	7,69732E-85	66.84	IPR001223 (PFAM)
scaffold-JOJQ-2040355-Cylindrocystis_cushleckae	Zygnemataceae	acidic mammalian chitinase-like	2,2866E-59	54.32	IPR001223 (PFAM)
scaffold-KADG-2007209-Ignatius_tetrasporus	Ulotrichaceae	related to chitinase	5,96352E-14	57.08	IPR001223 (PFAM)
scaffold-KEYW-2025590-Gonatozygon_kinahanii	Gonatozygaceae	Acidic endochitinase	1,58417E-80	65.86	IPR001223 (PFAM)
scaffold-KEYW-2026154-Gonatozygon_kinahanii	Gonatozygaceae	acidic mammalian chitinase-like	2,14258E-66	56.52	IPR001223 (PFAM)
scaffold-KMNX-2031577-Nucleotaenium_eifelense	Mesotaeniaceae	class V chitinase	8,25201E-12	68.51	IPR001223 (PFAM)
scaffold-KMNX-2035459-Nucleotaenium_eifelense	Mesotaeniaceae	AChain A	5,65275E-18	63.27	IPR001223 (PFAM)
scaffold-KUJU-2004605-Gonium_pectorale	Goniaceae	chitinase A1	1,88791E-170	54.47	IPR001223 (PFAM)
scaffold-LFVP-2004510-Marchantia_paleacea-2_samples_combined	Marchantiaceae	hevamine-A-like	4,58662E-151	75.07	IPR001223 (PFAM)
scaffold-LFVP-2009172- Marchantia_paleacea-2_samples_combined	Marchantiaceae	acidic mammalian chitinase-like	0.0	61.69	IPR001223 (PFAM)
scaffold-LFVP-2009173- Marchantia_paleacea-2_samples_combined	Marchantiaceae	acidic mammalian chitinase-like	0.0	61.69	IPR001223 (PFAM)
scaffold-LFVP-2081766-Marchantia_paleacea-2_samples_combined	Marchantiaceae	acidic endochitinase	4,00518E-135	75.0	IPR001223 (PFAM)
scaffold-LFVP-2083179- Marchantia_paleacea-2_samples_combined	Marchantiaceae	acidic endochitinase-like	7,04516E-93	67.43	IPR001223 (PFAM)

GREEN ALGAE – GH18					
SeqName	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-LFVP-2083339-Marchantia_paleacea-2_samples_combined	Marchantiaceae	acidic endochitinase-like	1,12303E-84	66.22	IPR001223 (PFAM)
scaffold-LFVP-2083413-Marchantia_paleacea-2_samples_combined	Marchantiaceae	hevamine A precursor-like protein	3,41875E-157	67.63	IPR001223 (PFAM)
scaffold-LFVP-2083566- Marchantia_paleacea-2_samples_combined	Marchantiaceae	acidic endochitinase-like	2,3841E-173	69.82	IPR001223 (PFAM)
scaffold-LNIL-2052393-Pteromonas_angulosa	Phacotaceae	chitotriosidase-1 isoform X1	4,53761E-56	51.64	IPR001223 (PFAM)
scaffold-MCHJ-2000297-Micrasterias_fimbriata	Desmidiaceae	acidic mammalian chitinase-like	8,43705E-40	50.59	IPR001223 (PFAM)
scaffold-MCHJ-2001441-Micrasterias_fimbriata	Desmidiaceae	AChain A	2,3481E-57	57.02	IPR001223 (PFAM)
scaffold-MFZO-2017215-Zygnemopsis_sp	Zygnemataceae	glycoside hydrolase family 18	1,94157E-42	74.89	IPR001223 (PFAM)
scaffold-MFZO-2022650-Zygnemopsis_sp	Zygnemataceae	AChain A	1,97128E-49	53.98	IPR001223 (PFAM)
scaffold-MMKU-2006934-Nephroselmis_olivace	Nephroselmidaceae	acidic mammalian chitinase-like	4,55345E-92	62.13	IPR001223 (PFAM)
scaffold-MNCB-2024465-Eremosphaera_viridi	Oocystaceae	AChain A	4,90316E-56	55.57	IPR001223 (PFAM)
scaffold-MNCB-2024466-Eremosphaera_viridi	Oocystaceae	chitotriosidase-1 precursor	1,25264E-56	55.85	IPR001223 (PFAM)
scaffold-MNCB-2024468-Eremosphaera_viridi	Oocystaceae	AChain A	2,18804E-47	53.45	IPR001223 (PFAM)
scaffold-MNNM-2000121-Cosmarium_granatum	Desmidiaceae	Glycoside hydrolase	1,28289E-76	66.58	IPR001223 (PFAM)
scaffold-MNNM-2037700-Cosmarium_granatum	Desmidiaceae	acidic endochitinase-like	3,30384E-45	68.26	IPR001223 (PFAM)
scaffold-MOYY-2009074-Pleurotaenium_trabecul	Desmidiaceae	AChain A	3,69547E-12	60.82	IPR001223 (PFAM)
scaffold-MOYY-2018456-Pleurotaenium_trabecul	Desmidiaceae	acidic mammalian chitinase-like	2,23137E-63	56.61	IPR001223 (PFAM)
scaffold-MOYY-2018457-Pleurotaenium_trabecul	Desmidiaceae	AChain A	1,39958E-65	56.34	IPR001223 (PFAM)
scaffold-MULF-2041340-Chlamydomonas_bilatus-A	Chlamydomonadaceae	related to chitinase	1,48889E-30	55.7	IPR001223 (PFAM)
scaffold-NBYP-2006024-Mesotaenium_kramstei	Mesotaeniaceae	acidic mammalian chitinase-like	2,4348E-64	53.56	IPR001223 (PFAM)
scaffold-OFUE-2010323-Lobochlamys_segnis	Chlamydomonadaceae	bacteriodes thetaiotaomicron symbiotic chitinase	1,24825E-57	50.79	IPR001223 (PFAM)
scaffold-OFUE-2047377-Lobochlamys_segnis	Chlamydomonadaceae	Chitinase 18-5	1,70246E-47	52.93	IPR001223 (PFAM)
scaffold-OTQG-2001226-Ankistrodesmus_sp	Ankistrodesmaceae	acidic mammalian chitinase-like	8,86151E-122	52.83	IPR001223 (PFAM)
scaffold-OTQG-2001227-Ankistrodesmus_sp	Ankistrodesmaceae	acidic mammalian chitinase-like	8,86151E-122	52.83	IPR001223 (PFAM)
scaffold-PFUD-2002294-Geminella_sp	Chlorellaceae	class V chitinase	9,05005E-28	48.85	IPR001223 (PFAM)
scaffold-PFUD-2034986-Geminella_sp	Chlorellaceae	glycoside hydrolase	5,22709E-48	71.92	IPR001223 (PFAM)
scaffold-POIR-2027547-Volvox_aureus-2_samples_combined	Volvocaceae	probable chitinase 3	1,73593E-117	54.25	IPR001223 (PFAM)
scaffold-PZBH-2032995-Oltmannsiellopsis_viridis-A	Oltmannsiellopsidaceae	acidic mammalian chitinase-like	4,77263E-33	69.06	IPR001223 (PFAM)
scaffold-PZIF-2015162-Scenedesmus_dimorphus	Scenedesmaceae	chitotriosidase-1 isoform X1	3,04739E-64	52.23	IPR001223 (PFAM)
scaffold-QJYX-2001962-Oltmannsiellopsis_viridis-B	Oltmannsiellopsidaceae	AChain A	3,31738E-74	56.94	IPR001223 (PFAM)
scaffold-QPDY-2002748-Coleochaete_irregularis	Coleochaetaceae	Acidic mammalian chitinase	1,57104E-07	47.14	IPR001223 (PFAM)
scaffold-QPDY-2030822-Coleochaete_irregularis	Coleochaetaceae	acidic mammalian chitinase-like	8,39575E-20	44.73	IPR001223 (PFAM)
scaffold-QPDY-2031501-Coleochaete_irregularis	Coleochaetaceae	chitotriosidase-1 isoform X1	4,91999E-71	53.67	IPR001223 (PFAM)
scaffold-QWFV-2003026-Bambusina_borreri	Desmidiaceae	acidic mammalian chitinase-like	2,3847E-51	55.21	IPR001223 (PFAM)
scaffold-QWFV-2003027-Bambusina_borreri	Desmidiaceae	acidic mammalian chitinase-like	1,00078E-13	55.45	IPR001223 (PFAM)
scaffold-QWFV-2003029-Bambusina_borreri	Desmidiaceae	acidic mammalian chitinase-like	4,65254E-35	54.4	IPR001223 (PFAM)

GREEN ALGAE – GH18					
SeqName	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-QWFV-2003030-Bambusina_borreri	Desmidiaceae	acidic mammalian chitinase-like	2,3847E-51	55.21	IPR001223 (PFAM)
scaffold-QWFV-2005300-Bambusina_borreri	Desmidiaceae	acidic endochitinase	1,30932E-67	64.06	IPR001223 (PFAM)
scaffold-QWFV-2005301-Bambusina_borreri	Desmidiaceae	acidic endochitinase	1,30932E-67	64.06	IPR001223 (PFAM)
scaffold-QWFV-2012344-Bambusina_borreri	Desmidiaceae	hevamine-A-like	3,7377E-53	64.37	IPR001223 (PFAM)
scaffold-WFV-2020226-Bambusina_borreri	Desmidiaceae	acidic endochitinase	1,25489E-43	73.67	IPR001223 (PFAM)
scaffold-QWFV-2020306-Bambusina_borreri	Desmidiaceae	hevamine-A-like	1,17698E-68	63.88	IPR001223 (PFAM)
scaffold-QYXY-2008419-Botryococcus_terribilis	Botryococcaceae	acidic endochitinase	1,50253E-20	64.57	IPR001223 (PFAM)
scaffold-QYXY-2040497-Botryococcus_terribilis	Botryococcaceae	glycoside hydrolase	6,12575E-21	64.09	IPR001223 (PFAM)
scaffold-RNAT-2058594-Eudorina_elegans	Volvocaceae	acidic mammalian chitinase-like	1,1601E-176	52.66	IPR001223 (PFAM)
scaffold-RPGL-2014644-Cylindrocystis_brebissonii-M2853	Mesotaeniaceae	acidic mammalian chitinase-like	9,71076E-19	54.29	IPR001223 (PFAM)
scaffold-RPGL-2027497-Cylindrocystis_brebissonii-M2853	Mesotaeniaceae	acidic endochitinase	7,87281E-60	65.83	IPR001223 (PFAM)
scaffold-RPQV-2005631-Phymatodocis_nordstedtiana	Desmidiaceae	hevamine-A-like	5,78316E-71	61.06	IPR001223 (PFAM)
scaffold-RPQV-2005632-Phymatodocis_nordstedtiana	Desmidiaceae	hevamine-A-like	2,62744E-63	62.28	IPR001223 (PFAM)
scaffold-RPQV-2006998-Phymatodocis_nordstedtiana	Desmidiaceae	hevamine-A-like	8,04738E-64	62.21	IPR001223 (PFAM)
scaffold-RPQV-2007705-Phymatodocis_nordstedtiana	Desmidiaceae	acidic endochitinase-like	4,18562E-97	66.9	IPR001223 (PFAM)
scaffold-RPQV-2036136-Phymatodocis_nordstedtiana	Desmidiaceae	AChain A	9,46972E-69	57.99	IPR001223 (PFAM)
scaffold-RPRU-2009659-Staurodesmus_omearii	Desmidiaceae	hevamine-A-like	9,12319E-52	60.64	IPR001223 (PFAM)
scaffold-RPRU-2039277-Staurodesmus_omearii	Desmidiaceae	acidic endochitinase-like	5,75826E-89	66.85	IPR001223 (PFAM)
scaffold-RQFE-2000106-Cosmocladium_cf._constrictum	Desmidiaceae	acidic mammalian chitinase-like	8,02421E-46	55.18	IPR001223 (PFAM)
scaffold-RQFE-2000107-Cosmocladium_cf._constrictum	Desmidiaceae	acidic mammalian chitinase	3,60652E-41	55.07	IPR001223 (PFAM)
scaffold-RQFE-2007301-Cosmocladium_cf._constrictum	Desmidiaceae	Acidic endochitinase	5,25319E-84	67.06	IPR001223 (PFAM)
scaffold-RYJX-2062142-Pandorina_morum	Volvocaceae	acidic mammalian chitinase-like	8,3005E-150	52.87	IPR001223 (PFAM)
scaffold-SNOX-2004982-Planotaenium_ohtanii	Mesotaeniaceae	probable chitinase 3	1,47872E-60	58.17	IPR001223 (PFAM)
scaffold-STKJ-2099024-Zygnema_sp.-A	Zygnemataceae	acidic endochitinase-like	8,78342E-76	61.66	IPR001223 (PFAM)
scaffold-STKJ-2099840-Zygnema_sp.-A	Zygnemataceae	Acidic endochitinase	3,37497E-83	66.04	IPR001223 (PFAM)
scaffold-UKUC-2105707-Dunaliella_salina-UTEX_LB_200	Dunaliellaceae	Acidic endochitinase	2,0939E-88	88.48	IPR001223 (PFAM)
scaffold-UKUC-2108261-Dunaliella_salina-UTEX_LB_200	Dunaliellaceae	probable chitinase 2	8,97051E-32	61.25	IPR001223 (PFAM)
scaffold-UTRE-2015935-Chloromonas_tughillensi	Chlamydomonadaceae	acidic mammalian chitinase-like	1,45728E-87	51.26	IPR001223 (PFAM)
scaffold-UTRE-2015936-Chloromonas_tughillensi	Chlamydomonadaceae	acidic mammalian chitinase-like	1,48853E-94	52.02	IPR001223 (PFAM)
scaffold-UTRE-2084885-Chloromonas_tughillensi	Chlamydomonadaceae	carbohydrate-binding module family 5	3,26042E-36	68.76	IPR001223 (PFAM)
scaffold-VAZE-2000533-Cylindrocystis_sp	Mesotaeniaceae	Acidic endochitinase	1,33575E-18	75.66	IPR001223 (PFAM)
scaffold-VAZE-2009036-Cylindrocystis_sp	Mesotaeniaceae	acidic mammalian chitinase-like	7,76265E-56	55.78	IPR001223 (PFAM)
scaffold-VBLH-2007028-Cladophora_glomerata	Cladophoraceae	acidic mammalian chitinase-like	7,36388E-32	53.6	IPR001223 (PFAM)
scaffold-VHIJ-2008486-Blastophysa_cf._rhizopus	Chaetosiphonaceae	chitotriosidase-1 precursor	3,16866E-21	52.23	IPR001223 (PFAM)l
scaffold-VHIJ-2008487-Blastophysa_cf._rhizopus	Chaetosiphonaceae	chitotriosidase-1 precursor	3,16866E-21	52.23	IPR001223 (PFAM)
scaffold-VIAU-2009074-Carteria_crucifera	Carteriaceae	acidic mammalian chitinase-like	1,65806E-21	54.78	IPR001223 (PFAM)
scaffold-VIAU-2009075-Carteria_crucifera	Carteriaceae	acidic mammalian chitinase-like	1,74084E-21	54.78	IPR001223 (PFAM)
scaffold-VIAU-2053932-Carteria_crucifera	Carteriaceae	acidic mammalian chitinase-like	1,47776E-08	66.93	IPR001223 (PFAM)

GREEN ALGAE – GH18					
SeqName	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-VQBJ-2003884-Coleochaete_scutata	Coleochaetaceae	AChain A	1,49241E-06	50.52	IPR001223 (PFAM)
scaffold-VQBJ-2004216-Coleochaete_scutata	Coleochaetaceae	acidic mammalian chitinase-like	1,28633E-12	51.83	IPR001223 (PFAM)
scaffold-VQBJ-2009052-Coleochaete_scutata	Coleochaetaceae	acidic mammalian chitinase-like	2,69157E-19	58.73	IPR001223 (PFAM)
scaffold-WCQU-2007350-Staurodesmus_convergens	Desmidiaceae	acidic mammalian chitinase-like	6,58893E-50	57.27	IPR001223 (PFAM)
scaffold-WDCW-2045137-Mesotaenium_endlicherianum	Mesotaeniaceae	acidic mammalian chitinase-like	2,18734E-39	60.97	IPR001223 (PFAM)
scaffold-WDCW-2046043-Mesotaenium_endlicherianum	Mesotaeniaceae	acidic endochitinase-like	1,14709E-78	62.84	IPR001223 (PFAM)
scaffold-WDGV-2049195-Cosmarium_subtumidum	Desmidiaceae	acidic mammalian chitinase-like	1,98274E-38	57.85	IPR001223 (PFAM)
scaffold-WDWX-2034378-Dunaliella_primolecta-UTEX_LB_1000	Dunaliellaceae	Acidic endochitinase	1,25342E-58	88.85	IPR001223 (PFAM)
scaffold-WGMD-3000448-Zygnema_sp.-B	Zygnemataceae	acidic endochitinase-like	6,87803E-77	65.33	IPR001223 (PFAM)
scaffold-WGMD-3000449-Zygnema_sp.-B	Zygnemataceae	acidic endochitinase-like	6,87803E-77	65.33	IPR001223 (PFAM)
scaffold-WGMD-3000450-Zygnema_sp.-B	Zygnemataceae	acidic endochitinase-like	6,87803E-77	65.33	IPR001223 (PFAM)
scaffold-WGMD-3000451-Zygnema_sp.-B	Zygnemataceae	Acidic endochitinase	6,2214E-39	73.96	IPR001223 (PFAM)
scaffold-WGMD-3003070-Zygnema_sp.-B	Zygnemataceae	chitotriosidase-1 isoform X1	9,17075E-47	59.96	IPR001223 (PFAM)
scaffold-WGMD-3003071-Zygnema_sp.-B	Zygnemataceae	chitotriosidase-1 isoform X1	9,17075E-47	59.96	IPR001223 (PFAM)
scaffold-WGMD-3014691-Zygnema_sp.-B	Zygnemataceae	acidic mammalian chitinase-like	1,28285E-15	56.35	IPR001223 (PFAM)
scaffold-WRSL-2006874-Volvox_aureus-M2242	Volvocaceae	chitinase C	5,35671E-171	55.68	IPR001223 (PFAM)
scaffold-WSJO-2006870-Mesotaenium_braunii	Mesotaeniaceae	Acidic mammalian chitinase	1,59936E-25	58.12	IPR001223 (PFAM)
scaffold-WSJO-2048613-Mesotaenium_braunii	Mesotaeniaceae	Chitinase	6,36077E-60	50.14	IPR001223 (PFAM)
scaffold-XDLL-2006213-Oogamochlamys_gigantea	Chlamydomonadaceae	Chitinase 18-5	1,70246E-47	52.93	IPR001223 (PFAM)
scaffold-XDLL-2006214-Oogamochlamys_gigantea	Chlamydomonadaceae	Chitinase	4,44353E-63	52.81	IPR001223 (PFAM)
scaffold-XIVI-2090327-Cymbomonas_sp	Pyramimonadaceae	glycoside hydrolase family 18	2,64631E-13	52.64	IPR001223 (PFAM)
scaffold-XRTZ-2004862-Roya_obtusa	Mesotaeniaceae	acidic mammalian chitinase-like	1,50556E-57	59.26	IPR001223 (PFAM)
scaffold-YOXI-2002515-Cylindrocystis_brebissonii-M2213	Mesotaeniaceae	acidic mammalian chitinase	6,46442E-53	57.26	IPR001223 (PFAM)
scaffold-YOXI-2007181-Cylindrocystis_brebissonii-M2213	Mesotaeniaceae	hevamine-A-like	1,67289E-71	66.98	IPR001223 (PFAM)
scaffold-YSQT-2005657-Penium_exiguum	Peniaceae	AChain A	1,2438E-63	62.69	IPR001223 (PFAM)
scaffold-YSQT-2032479-Penium_exiguum	Peniaceae	glycoside hydrolase family 18	8,51608E-36	62.69	IPR001223 (PFAM)
scaffold-YSQT-2036100-Penium_exiguum	Peniaceae	hevamine-A-like	1,8449E-82	66.93	IPR001223 (PFAM)
scaffold-ZNUM-2002100-Leptosira_obovata	Leptosiraceae	glycoside hydrolase	2,42083E-39	48.42	IPR001223 (PFAM)
scaffold-ZNUM-2002101-Leptosira_obovata	Leptosiraceae	glycoside hydrolase	8,06833E-37	49.16	IPR001223 (PFAM)
scaffold-ZRMT-2003702-Mougeotia_sp	Zygnemataceae	acidic endochitinase-like	3,38161E-80	65.93	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-ABCD-2061358-Niphotrichum_elongatum	Mosses	Grimmiaceae	AChain A	9,20213E-39	60.27	IPR001223 (PFAM)
scaffold-AEXY-2001889-Blasia_sp	Liverworts	Blasiaceae	Chitinase	1,50485E-69	80.81	IPR001223 (PFAM)
scaffold-AEXY-2002225-Blasia_sp	Liverworts	Blasiaceae	glycoside hydrolase family 18	1,13991E-49	91.29	IPR001223 (PFAM)
scaffold-AEXY-2002415-Blasia_sp	Liverworts	Blasiaceae	glycosyl hydrolase family 18	1,10586E-55	69.44	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-AEXY-2010785-Blasia_sp	Liverworts	Blasiaceae	Chitinase	2,60338E-91	90.89	IPR001223 (PFAM)
scaffold-AEXY-2015542-Blasia_sp	Liverworts	Blasiaceae	chitinase 18-7	3,77538E-127	85.09	IPR001223 (PFAM)
scaffold-AEXY-2142432-Blasia_sp	Liverworts	Blasiaceae	class V chitinase	1,02141E-106	85.45	IPR001223 (PFAM)
scaffold-AEXY-2149299-Blasia_sp	Liverworts	Blasiaceae	class V chitinase ChiB1	7,35261E-150	71.57	IPR001223 (PFAM)
scaffold-AEXY-2151689-Blasia_sp	Liverworts	Blasiaceae	Chitinase	2,4465E-91	100.0	IPR001223 (PFAM)
scaffold-AEXY-2152298-Blasia_sp	Liverworts	Blasiaceae	brain chitinase and chia	2,88425E-15	68.61	IPR001223 (PFAM)
scaffold-AKXB-2067103-Phaeomegaceros_coriaceus	Hornworts	Dendrocerotaceae	acidic mammalian chitinase-like	2,14273E-98	62.56	IPR001223 (PFAM)
scaffold-ANON-2000173-Leiosporoceros_dussii-B	Hornworts	Leiosporocerotaceae	acidic mammalian chitinase-like	4,90038E-76	62.39	IPR001223 (PFAM)
scaffold-AWOI-2003923-Diphyscium_foliosum	Mosses	Diphysciaceae	glycoside hydrolase	1,68908E-39	69.15	IPR001223 (PFAM)
scaffold-AWOI-2069936-Diphyscium_foliosum	Mosses	Diphysciaceae	class V chitinase-like	2,13322E-111	63.27	IPR001223 (PFAM)
scaffold-AWOI-2072168-Diphyscium_foliosum	Mosses	Diphysciaceae	acidic mammalian chitinase-like	5,97531E-162	62.52	IPR001223 (PFAM)
scaffold-BGXB-2075632-Plagiomnium_insigne	Mosses	Mielichhoferiaceae	acidic mammalian chitinase-like	1,62616E-86	65.59	IPR001223 (PFAM)
scaffold-BPSG-2088193-Scouleria_aquatica	Mosses	Scouleriaceae	class V chitinase-like	3,56822E-171	64.32	IPR001223 (PFAM)
scaffold-BSNI-2050540-Anthoceros_agrestis-B	Hornworts	Anthoceroceae	acidic mammalian chitinase-like	9,8284E-93	64.35	IPR001223 (PFAM)
scaffold-CHJJ-2002473-Lejeuneaceae_sp	Liverworts	Frullaniaceae	acidic mammalian chitinase	5,44865E-15	64.43	IPR001223 (PFAM)
scaffold-CHJJ-2019067-Lejeuneaceae_sp	Liverworts	Frullaniaceae	acidic mammalian chitinase-like	1,63216E-52	61.5	IPR001223 (PFAM)
scaffold-CHJJ-2135670-Lejeuneaceae_sp	Liverworts	Frullaniaceae	glycoside hydrolase	6,58039E-72	63.36	IPR001223 (PFAM)
scaffold-CHJJ-2137152-Lejeuneaceae_sp	Liverworts	Frullaniaceae	class V chitinase-like	1,35375E-124	58.36	IPR001223 (PFAM)
scaffold-CMEQ-2010358-Orthotrichum_lyellii	Mosses	Orthotrichaceae	acidic mammalian chitinase-like	0.0	62.91	IPR001223 (PFAM)
scaffold-CMEQ-2084544-Orthotrichum_lyellii	Mosses	Orthotrichaceae	Chitinase 1	1,41156E-151	83.41	IPR001223 (PFAM)
scaffold-DHWX-2004311-Fontinalis_antipyretica	Mosses	Fontinalaceae	acidic mammalian chitinase-like	4,93069E-172	61.95	IPR001223 (PFAM)
scaffold-DXOU-2037702-Nothoceros_aenigmaticus	Hornworts	Dendrocerotaceae	acidic mammalian chitinase-like	6,54902E-94	62.44	IPR001223 (PFAM)
scaffold-EEMJ-2001989-Thuidium_delicatulum	Mosses	Thuidiaceae	class V chitinase	3,84084E-152	61.02	IPR001223 (PFAM)
scaffold-EEMJ-2001990-Thuidium_delicatulum	Mosses	Thuidiaceae	class V chitinase	2,89724E-56	60.49	IPR001223 (PFAM)
scaffold-FFPD-2013361-Ceratodon_purpureus	Mosses	Ditrichaceae	acidic mammalian chitinase	1,35708E-151	63.94	IPR001223 (PFAM)
scaffold-FITN-2000573-Treubia_lacunose	Liverworts	Treubiaceae	acidic mammalian chitinase-like	1,89337E-125	67.86	IPR001223 (PFAM)
scaffold-FITN-2000574-Treubia_lacunose	Liverworts	Treubiaceae	acidic mammalian chitinase-like	3,27295E-32	66.86	IPR001223 (PFAM)
scaffold-FITN-2000575-Treubia_lacunose	Liverworts	Treubiaceae	class V chitinase	6,29708E-125	67.04	IPR001223 (PFAM)
scaffold-FITN-2007298-Treubia_lacunose	Liverworts	Treubiaceae	acidic endochitinase	1,42066E-102	75.86	IPR001223 (PFAM)
scaffold-FITN-2007299-Treubia_lacunose	Liverworts	Treubiaceae	acidic endochitinase	3,43287E-87	71.08	IPR001223 (PFAM)
scaffold-FITN-2007300-Treubia_lacunose	Liverworts	Treubiaceae	acidic endochitinase	1,97106E-100	77.93	IPR001223 (PFAM)
scaffold-FITN-2007301-Treubia_lacunose	Liverworts	Treubiaceae	acidic endochitinase	3,68093E-101	78.23	IPR001223 (PFAM)
scaffold-FITN-2013117-Treubia_lacunose	Liverworts	Treubiaceae	class V chitinase	4,77637E-84	57.67	IPR001223 (PFAM)
scaffold-FITN-2013118-Treubia_lacunose	Liverworts	Treubiaceae	class V chitinase	2,46802E-74	59.6	IPR001223 (PFAM)
scaffold-FITN-2016127-Treubia_lacunose	Liverworts	Treubiaceae	class V chitinase	2,20471E-103	65.88	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-FITN-2017805-Treubia_lacunose	Liverworts	Treubiaceae	Glycosyl hydrolase family protein with chitinase insertion domain-containing	2,91922E-111	69.16	IPR001223 (PFAM)
scaffold-FITN-2091255-Treubia_lacunose	Liverworts	Treubiaceae	Glycosyl hydrolase family protein with chitinase insertion domain-containing	1,80846E-128	66.44	IPR001223 (PFAM)
scaffold-GOWD-2009968-Sphagnum_lescurii	Mosses	Sphagnaceae	chitotriosidase-1 isoform X1	6,0112E-91	58.66	IPR001223 (PFAM)
scaffold-GOWD-2019400-Sphagnum_lescurii	Mosses	Sphagnaceae	Chitinase-like protein Idgf4	1,01141E-85	80.2	IPR001223 (PFAM)
scaffold-GOWD-2019795-Sphagnum_lescurii	Mosses	Sphagnaceae	Chitinase-like protein Idgf4	3,9289E-133	68.83	IPR001223 (PFAM)
scaffold-GOWD-2079062-Sphagnum_lescurii	Mosses	Sphagnaceae	acidic mammalian chitinase-like	1,601E-57	72.68	IPR001223 (PFAM)
scaffold-GOWD-2082285-Sphagnum_lescurii	Mosses	Sphagnaceae	chitinase-like protein Idgf4	1,60351E-119	79.31	IPR001223 (PFAM)
scaffold-GOWD-2083251-Sphagnum_lescurii	Mosses	Sphagnaceae	glycoside hydrolase family 18	8,89648E-103	75.44	IPR001223 (PFAM)
scaffold-GRKU-2005548-Syntrichia_princeps	Mosses	Pottiaceae	acidic mammalian chitinase	1,04373E-132	66.43	IPR001223 (PFAM)
scaffold-HERT-2006371-Sphaerocarpos_texanus	Liverworts	Sphaerocarpaceae	acidic mammalian chitinase	1,12565E-48	63.46	IPR001223 (PFAM)
scaffold-HERT-2038302-Sphaerocarpos_texanus	Liverworts	Sphaerocarpaceae	acidic mammalian chitinase	5,30488E-43	63.79	IPR001223 (PFAM)
scaffold-HMHL-2009920-Marchantia_paleacea-non_mycorrizal	Liverworts	Marchantiaceae	acidic mammalian chitinase-like	0.0	61.69	IPR001223 (PFAM)
scaffold-HMHL-2052653-Marchantia_paleacea-non_mycorrizal	Liverworts	Marchantiaceae	hevamine A precursor-like protein	3,41875E-157	67.63	IPR001223 (PFAM)
scaffold-HMHL-2053338-Marchantia_paleacea-non_mycorrizal	Liverworts	Marchantiaceae	hevamine-A-like	4,58662E-151	75.07	IPR001223 (PFAM)
scaffold-HPXA-2012251-Ptilidium_pulcherrimum	Liverworts	Ptilidiaceae	AChain A	6,82885E-74	61.11	IPR001223 (PFAM)
scaffold-HPXA-2018253-Ptilidium_pulcherrimum	Liverworts	Ptilidiaceae	glycoside hydrolase family 18	2,26627E-46	52.85	IPR001223 (PFAM)
scaffold-HPXA-2018545-Ptilidium_pulcherrimum	Liverworts	Ptilidiaceae	Endochitinase 1	6,26944E-98	78.46	IPR001223 (PFAM)
scaffold-HPXA-21 scaffold-22978-Ptilidium_pulcherrimum	Liverworts	Ptilidiaceae	Chitinase 1	1,47149E-26	62.48	IPR001223 (PFAM)
scaffold-HPXA-2133118-Ptilidium_pulcherrimum	Liverworts	Ptilidiaceae	acidic mammalian chitinase-like	3,22663E-97	64.81	IPR001223 (PFAM)
scaffold-HVBQ-2126934-Tetraphis_pellucida	Mosses	Tetraphidaceae	acidic mammalian chitinase-like	1,36213E-64	66.28	IPR001223 (PFAM)
scaffold-IGUH-2002948-Schwetschkeopsis_fabronia	Mosses	Leucodontaceae	class V chitinase-like	0.0	63.17	IPR001223 (PFAM)
scaffold-IGUH-2002949-Schwetschkeopsis_fabronia	Mosses	Leucodontaceae	class V chitinase-like	0.0	63.17	IPR001223 (PFAM)
scaffold-IGUH-2158382-Schwetschkeopsis_fabronia	Mosses	Leucodontaceae	glycoside hydrolase family 18 protein	9,78462E-33	66.58	IPR001223 (PFAM)
scaffold-IGUH-2163943-Schwetschkeopsis_fabronia	Mosses	Leucodontaceae	Chitinase	1,23696E-65	69.43	IPR001223 (PFAM)
scaffold-IGUH-2164561-Schwetschkeopsis_fabronia	Mosses	Leucodontaceae	acidic mammalian chitinase-like	2,81279E-39	62.16	IPR001223 (PFAM)
scaffold-IHWO-2005936-Marchantia_paleacea-mycorrizal	Liverworts	Marchantiaceae	hevamine-A-like	4,58662E-151	75.07	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-IHWO-2011148-Marchantia_paleacea-mycorrizal	Liverworts	Marchantiaceae	acidic mammalian chitinase-like	0.0	61.69	IPR001223 (PFAM)
scaffold-IHWO-2065167-Marchantia_paleacea-mycorrizal	Liverworts	Marchantiaceae	acidic endochitinase-like	7,04516E-93	67.43	IPR001223 (PFAM)
scaffold-IHWO-2065195-Marchantia_paleacea-mycorrizal	Liverworts	Marchantiaceae	hevamine A precursor-like protein	3,41875E-157	67.63	IPR001223 (PFAM)
scaffold-IHWO-2065225-Marchantia_paleacea-mycorrizal	Liverworts	Marchantiaceae	acidic endochitinase-like	2,3841E-173	69.82	IPR001223 (PFAM)
scaffold-IHWO-2065363-Marchantia_paleacea-mycorrizal	Liverworts	Marchantiaceae	acidic endochitinase-like	1,12303E-84	66.22	IPR001223 (PFAM)
scaffold-ILBQ-2043168-Conocephalum_conicum	Liverworts	Conocephalaceae	acidic endochitinase-like	1,52664E-137	73.27	IPR001223 (PFAM)
scaffold-ILBQ-2045545-Conocephalum_conicum	Liverworts	Conocephalaceae	acidic mammalian chitinase-like	4,39239E-174	66.09	IPR001223 (PFAM)
scaffold-IQJU-2000727-Anthoceros_formosae	Hornworts	Anthoceroceae	AChain A	3,31658E-63	57.36	PR001223 (PFAM)
scaffold-IRBN-2013604-Scapania_nemorosa	Liverworts	Scapaniaceae	Chitinase	3,02276E-75	78.75	IPR001223 (PFAM)
scaffold-IRBN-2136833-Scapania_nemorosa	Liverworts	Scapaniaceae	acidic mammalian chitinase-like	7,31856E-49	72.14	IPR001223 (PFAM)
scaffold-IRBN-2151386-Scapania_nemorosa	Liverworts	Scapaniaceae	glycoside hydrolase family 18	1,09715E-58	63.64	IPR001223 (PFAM)
scaffold-IRBN-2152154-Scapania_nemorosa	Liverworts	Scapaniaceae	glycoside hydrolase	8,92567E-60	66.68	IPR001223 (PFAM)
scaffold-IRBN-2161616-Scapania_nemorosa	Liverworts	Scapaniaceae	putative glycoside hydrolase family 18 protein	1,50676E-86	71.44	IPR001223 (PFAM)
scaffold-IRBN-2162788-Scapania_nemorosa	Liverworts	Scapaniaceae	glycoside hydrolase family 18	1,4566E-68	61.46	IPR001223 (PFAM)
scaffold-JADL-2003775-Rhynchostegium_serrulatum	Mosses	Brachytheciaceae	acidic mammalian chitinase-like	0.0	62.26	IPR001223 (PFAM)
scaffold-JHFI-2004290-Pellia_neesiana	Liverworts	Pelliaceae	hevamine-A-like	3,71414E-65	73.13	IPR001223 (PFAM)
scaffold-JHFI-2012415-Pellia_neesiana	Liverworts	Pelliaceae	acidic endochitinase	1,74126E-87	69.23	IPR001223 (PFAM)
scaffold-JHFI-2012990-Pellia_neesiana	Liverworts	Pelliaceae	acidic endochitinase	2,83176E-90	68.52	IPR001223 (PFAM)
scaffold-JHFI-2013319-Pellia_neesiana	Liverworts	Pelliaceae	Acidic endochitinase	1,93413E-24	66.0	IPR001223 (PFAM)
scaffold-JHFI-2013440-Pellia_neesiana	Liverworts	Pelliaceae	acidic endochitinase	1,52649E-107	73.35	IPR001223 (PFAM)
scaffold-JHFI-2095885-Pellia_neesiana	Liverworts	Pelliaceae	class V chitinase	9,55848E-44	65.22	IPR001223 (PFAM)
scaffold-JHFI-2109525-Pellia_neesiana	Liverworts	Pelliaceae	glycoside hydrolase family 18	1,98464E-116	76.47	IPR001223 (PFAM)
scaffold-JHFI-2112286-Pellia_neesiana	Liverworts	Pelliaceae	acidic endochitinase	3,01572E-96	66.99	IPR001223 (PFAM)
scaffold-JMXW-2011418-Bryum_argenteum	Mosses	Bryaceae	acidic mammalian chitinase-like	5,44726E-166	62.49	IPR001223 (PFAM)
scaffold-JPYU-2007981-Marchantia_polymorpha	Liverworts	Marchantiaceae	acidic endochitinase-like	5,88343E-171	76.32	IPR001223 (PFAM)
scaffold-JPYU-2038245-Marchantia_polymorpha	Liverworts	Marchantiaceae	acidic endochitinase-like	0.0	73.7	IPR001223 (PFAM)
scaffold-JPYU-2038298-Marchantia_polymorpha	Liverworts	Marchantiaceae	acidic mammalian chitinase-like	0.0	65.15	IPR001223 (PFAM)
scaffold-KEFD-2052193-Encalypta_streptocarpa	Mosses	Encalyptaceae	acidic mammalian chitinase-like	2,76823E-11	65.38	IPR001223 (PFAM)
scaffold-KEFD-2060547-Encalypta_streptocarpa	Mosses	Encalyptaceae	class V chitinase	0.0	63.16	IPR001223 (PFAM)
scaffold-KRUQ-2106675-Porella_navicularis	Liverworts	Porellaceae	Chitinase	7,28084E-40	81.53	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-KRUQ-2106844-Porella_navicularis	Liverworts	Porellaceae	acidic mammalian chitinase isoform X2	2,96939E-33	70.75	IPR001223 (PFAM)
scaffold-KRUQ-2111018-Porella_navicularis	Liverworts	Porellaceae	glycoside hydrolase	2,93169E-72	62.19	IPR001223 (PFAM)
scaffold-KRUQ-2111802-Porella_navicularis	Liverworts	Porellaceae	glycoside hydrolase	2,28783E-84	65.94	IPR001223 (PFAM)
scaffold-LGOW-2003062-Schistochila_sp	Liverworts	Schistochilaceae	carbohydrate-binding module family 1 protein	1,09654E-158	79.9	IPR001223 (PFAM)
scaffold-LGOW-2012088-Schistochila_sp	Liverworts	Schistochilaceae	putative chitinase	7,98664E-34	78.02	IPR001223 (PFAM)
scaffold-LGOW-2016622-Schistochila_sp	Liverworts	Schistochilaceae	class III chitinase ChiA1	2,77662E-96	64.34	IPR001223 (PFAM)
scaffold-LGOW-2017619-Schistochila_sp	Liverworts	Schistochilaceae	chitinase 18-5	3,96304E-114	85.43	IPR001223 (PFAM)
scaffold-LGOW-2087427-Schistochila_sp	Liverworts	Schistochilaceae	class V chitinase	1,01574E-27	71.7	IPR001223 (PFAM)
scaffold-LGOW-2102714-Schistochila_sp	Liverworts	Schistochilaceae	glycoside hydrolase family 18	1,00455E-142	69.23	IPR001223 (PFAM)
scaffold-LGOW-2103657-Schistochila_sp	Liverworts	Schistochilaceae	hevamine-A-like	7,99851E-86	7054	IPR001223 (PFAM)
scaffold-LNSF-2002884-Hypnum_subimponens	Mosses	Pylaisiaceae	class V chitinase-like	1,07601E-167	62.93	IPR001223 (PFAM)
scaffold-LNSF-2009955-Hypnum_subimponens	Mosses	Pylaisiaceae	glycoside hydrolase family 18	6,86338E-59	68.23	IPR001223 (PFAM)
scaffold-LNSF-2011653-Hypnum_subimponens	Mosses	Pylaisiaceae	glycoside hydrolase family 18	1,38533E-54	63.07	IPR001223 (PFAM)
scaffold-LNSF-2067677-Hypnum_subimponens	Mosses	Pylaisiaceae	class V chitinase	2,81196E-109	59.18	IPR001223 (PFAM)
scaffold-LNSF-2068240-Hypnum_subimponens	Mosses	Pylaisiaceae	glycoside hydrolase family 18	7,96472E-46	63.16	IPR001223 (PFAM)
scaffold-MIRS-2086783-Climacium_dendroides	Mosses	Climaciaceae	class V chitinase-like	5,22273E-162	64.28	IPR001223 (PFAM)
scaffold-NGTD-2089690-Dicranum_scoparium	Mosses	Dicranaceae	glycoside hydrolase	6,22035E-41	70.24	IPR001223 (PFAM)
scaffold-NGTD-2090084-Dicranum_scoparium	Mosses	Dicranaceae	glycoside hydrolase family	6,5183E-19	58.84	IPR001223 (PFAM)
scaffold-NGTD-2094229-Dicranum_scoparium	Mosses	Dicranaceae	glycoside hydrolase	3,19773E-66	69.03	IPR001223 (PFAM)
scaffold-NGTD-2098818-Dicranum_scoparium	Mosses	Dicranaceae	class V chitinase	2,73451E-144	64.23	IPR001223 (PFAM)
scaffold-NRWZ-2005174-Metzgeria_crassipilis	Liverworts	Metzgeriaceae	acidic mammalian chitinase-like	1,75861E-114	59.88	IPR001223 (PFAM)
scaffold-OFTV-2008694-Barbilophozia_barbata	Liverworts	Scapaniaceae	glycoside hydrolase family 18	1,48933E-52	74.63	IPR001223 (PFAM)
scaffold-OFTV-2010947-Barbilophozia_barbata	Liverworts	Scapaniaceae	glycoside hydrolase family 18	4,45478E-60	68.02	IPR001223 (PFAM)
scaffold-OFTV-2010948-Barbilophozia_barbata	Liverworts	Scapaniaceae	glycoside hydrolase family 18	4,45478E-60	68.02	IPR001223 (PFAM)
scaffold-OFTV-2128372-Barbilophozia_barbata	Liverworts	Scapaniaceae	chitinase-3-like protein 1	3,49963E-78	73.16	IPR001223 (PFAM)
scaffold-OFTV-2131944-Barbilophozia_barbata	Liverworts	Scapaniaceae	glycoside hydrolase family 18	1,83953E-135	81.0	IPR001223 (PFAM)
scaffold-ORKS-2003937-Philonotis_fontana	Mosses	Bartramiaceae	probable chitinase 10	7,79284E-37	90.32	IPR001223 (PFAM)
scaffold-ORKS-2003938-Philonotis_fontana	Mosses	Bartramiaceae	probable chitinase 10	2,52236E-44	89.73	IPR001223 (PFAM)
scaffold-ORKS-2055632-Philonotis_fontana	Mosses	Bartramiaceae	probable chitinase 10	9,59857E-59	91.97	IPR001223 (PFAM)
scaffold-ORKS-2056670-Philonotis_fontana	Mosses	Bartramiaceae	chitinase-like protein Idgf4	1,7006E-117	74.39	IPR001223 (PFAM)
scaffold-ORKS-2056948-Philonotis_fontana	Mosses	Bartramiaceae	acidic mammalian chitinase-like	0.0	62.35	IPR001223 (PFAM)
scaffold-PIUF-2010922-Pellia_cf._Epiphylla	Liverworts	Pelliaceae	hevamine-A-like	2,69193E-64	71.95	IPR001223 (PFAM)
scaffold-PIUF-2012805-Pellia_cf._Epiphylla	Liverworts	Pelliaceae	acidic endochitinase	2,71444E-105	72.24	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-PIUF-2082270-Pellia_cf._Epiphylla	Liverworts	Pelliaceae	acidic endochitinase	6,01528E-40	63.53	IPR001223 (PFAM)
scaffold-PIUF-2090613-Pellia_cf._Epiphylla	Liverworts	Pelliaceae	acidic endochitinase	3,53021E-95	67.35	IPR001223 (PFAM)
scaffold-PIUF-2092517-Pellia_cf._Epiphylla	Liverworts	Pelliaceae	acidic mammalian chitinase-like	5,72258E-125	61.84	IPR001223 (PFAM)
scaffold-QKQO-2049619-Pseudotaxiphyllum_elegans	Mosses	Plagiotheciaceae	class V chitinase-like	1,18021E-175	63.03	IPR001223 (PFAM)
scaffold-QMWB-2012715-Anomodon_attenuatus	Mosses	Myuriaceae	acidic mammalian chitinase-like	2,1312E-170	62.93	IPR001223 (PFAM)
scaffold-QMWB-2060103-Anomodon_attenuatus	Mosses	Myuriaceae	putative Chitinase	1,74272E-59	60.94	IPR001223 (PFAM)
scaffold-RCBT-2024866-Sphagnum_palustre	Mosses	Sphagnaceae	chitinase-3-like protein 1	1,23818E-133	70.93	IPR001223 (PFAM)
scaffold-RCBT-2029889-Sphagnum_palustre	Mosses	Sphagnaceae	chitinase-3-like protein 1	1,1818E-130	70.69	IPR001223 (PFAM)
scaffold-RCBT-2175611-Sphagnum_palustre	Mosses	Sphagnaceae	class III chitinase	5,65486E-26	75.36	IPR001223 (PFAM)
scaffold-RCBT-2182330-Sphagnum_palustre	Mosses	Sphagnaceae	acidic mammalian chitinase-like	7,42305E-25	57.22	IPR001223 (PFAM)
scaffold-RCBT-2185854-Sphagnum_palustre	Mosses	Sphagnaceae	glycoside hydrolase family 18	6,11726E-102	75.53	IPR001223 (PFAM)
scaffold-RDOO-2001513-Racomitrium_varium	Mosses	Grimmiaceae	related to endochitinase 2 precursor	9,73557E-76	63.44	IPR001223 (PFAM)
scaffold-RGKI-2011057-Leucobryum_glaucum	Mosses	Leucobryaceae	acidic mammalian chitinase-like	5,10235E-127	63.31	IPR001223 (PFAM)
scaffold-RTMU-2003364-Calypogeia_fissa	Liverworts	Calypogeiaceae	acidic mammalian chitinase-like	1,17979E-109	66.64	IPR001223 (PFAM)
scaffold-RTMU-2006799-Calypogeia_fissa	Liverworts	Calypogeiaceae	acidic mammalian chitinase-like	6,42373E-91	68.38	IPR001223 (PFAM)
scaffold-RTMU-2010105-Calypogeia_fissa	Liverworts	Calypogeiaceae	glycoside hydrolase family 18	1,04185E-104	82.17	IPR001223 (PFAM)
scaffold-RTMU-2010106-Calypogeia_fissa	Liverworts	Calypogeiaceae	glycoside hydrolase family 18	1,04185E-104	82.17	IPR001223 (PFAM)
scaffold-RTMU-2019205-Calypogeia_fissa	Liverworts	Calypogeiaceae	chitin recognition protein	1,73439E-88	66.03	IPR001223 (PFAM)
scaffold-RTMU-2143976-Calypogeia_fissa	Liverworts	Calypogeiaceae	acidic mammalian chitinase	2,94563E-36	66.37	IPR001223 (PFAM)
scaffold-RTMU-2175404-Calypogeia_fissa	Liverworts	Calypogeiaceae	glycosyl hydrolase family 18	6,6491E-79	82.89	IPR001223 (PFAM)
scaffold-RTMU-2175514-Calypogeia_fissa	Liverworts	Calypogeiaceae	acidic mammalian chitinase isoform X4	5,65258E-73	71.23	IPR001223 (PFAM)
scaffold-RTMU-2179298-Calypogeia_fissa	Liverworts	Calypogeiaceae	acidic mammalian chitinase-like	2,77308E-82	67.06	IPR001223 (PFAM)
scaffold-RTMU-2179446-Calypogeia_fissa	Liverworts	Calypogeiaceae	chitinase-3-like protein 1	8,94988E-100	69.82	IPR001223 (PFAM)
scaffold-RTMU-2179721-Calypogeia_fissa	Liverworts	Calypogeiaceae	chitinase-like protein Idgf4	1,94877E-124	75.38	IPR001223 (PFAM)
scaffold-RTMU-2182337-Calypogeia_fissa	Liverworts	Calypogeiaceae	glycoside hydrolase	4,39345E-104	70.41	IPR001223 (PFAM)
scaffold-RTMU-2183105-Calypogeia_fissa	Liverworts	Calypogeiaceae	class V chitinase-like	8,4533E-110	61.16	IPR001223 (PFAM)
scaffold-RXRQ-2121621-Phaeoceros_carolinianus-sporophyte	Hornworts	Notothyladaceae	acidic mammalian chitinase-like isoform X1	1,86137E-29	67.61	IPR001223 (PFAM)
scaffold-RXRQ-2133493-Phaeoceros_carolinianus-sporophyte	Hornworts	Notothyladaceae	glycoside hydrolase family 18	1,3225E-145	83.4	IPR001223 (PFAM)
scaffold-RXRQ-2133873-Phaeoceros_carolinianus-sporophyte	Hornworts	Notothyladaceae	endochitinase 42	1,11909E-116	84.86	IPR001223 (PFAM)
scaffold-RXRQ-2137557-Phaeoceros_carolinianus-sporophyte	Hornworts	Notothyladaceae	chitinase-like protein CG5210	9,53502E-124	89.95	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-RXRQ-2137809-Phaeoceros_carolinianus-sporophyte	Hornworts	Notothyladaceae	chitinase-like protein Idgf4	7,23016E-53	94.66	IPR001223 (PFAM)
scaffold-RXRQ-2139890-Phaeoceros_carolinianus-sporophyte	Mosses	Takakiaceae	class V chitinase-like	2,86542E-80	61.24	IPR001223 (PFAM)
scaffold-SKQD-2073562-Takakia_lepidozioides	Mosses	Takakiaceae	class V chitinase	1,21297E-77	66.48	IPR001223 (PFAM)
scaffold-SKQD-2079574-Takakia_lepidozioides	Mosses	Takakiaceae	acidic mammalian chitinase	6,88966E-137	63.61	IPR001223 (PFAM)
scaffold-SZYG-2012109-Polytrichum_commune-gametophyte	Mosses	Polytrichaceae	class V chitinase-like	1,68946E-116	60.85	IPR001223 (PFAM)
scaffold-SZYG-2040465-Polytrichum_commune-gametophyte	Mosses	Polytrichaceae	acidic mammalian chitinase-like	5,72987E-54	66.79	IPR001223 (PFAM)
scaffold-TAVP-2019781-Calliergon_cordifolium	Mosses	Calliergonaceae	acidic mammalian chitinase-like	1,17084E-164	62.68	IPR001223 (PFAM)
scaffold-TCBC-2013303-Megaceros_vincentianus	Hornworts	Dendrocerotaceae	class V chitinase-like	1,30311E-45	65.2	IPR001223 (PFAM)
scaffold-TFDQ-2006119-Monoclea_gottschei	Liverworts	Monocleaceae	related to endochitinase 2 precursor	1,2314E-98	73.44	IPR001223 (PFAM)
scaffold-TFDQ-2008991-Monoclea_gottschei	Liverworts	Monocleaceae	glycosyl hydrolase family 18	1,10586E-55	69.44	IPR001223 (PFAM)
scaffold-TFDQ-2009959-Monoclea_gottschei	Liverworts	Monocleaceae	glycoside hydrolase family 18	4,19293E-65	79.15	IPR001223 (PFAM)
scaffold-TFDQ-2120835-Monoclea_gottschei	Liverworts	Monocleaceae	Endochitinase 1	0.0	92.51	IPR001223 (PFAM)
scaffold-TFDQ-2120916-Monoclea_gottschei	Liverworts	Monocleaceae	Chitinase	1,16384E-147	87.77	IPR001223 (PFAM)
scaffold-TFDQ-2121018-Monoclea_gottschei	Liverworts	Monocleaceae	related to chitinase	2,88006E-159	93.81	IPR001223 (PFAM)
scaffold-TFDQ-2123325-Monoclea_gottschei	Liverworts	Monocleaceae	related to RF2 protein	2,4972E-131	83.2	IPR001223 (PFAM)
scaffold-TFYI-2000164-Marchantia_emarginata	Liverworts	Marchantiaceae	chitinase-3-like protein 1	2,72805E-55	63.4	IPR001223 (PFAM)
scaffold-TFYI-2002021-Marchantia_emarginata	Liverworts	Marchantiaceae	acidic endochitinase	6,06757E-55	79.36	IPR001223 (PFAM)
scaffold-TFYI-2002959-Marchantia_emarginata	Liverworts	Marchantiaceae	hevamine-A-like	1,24185E-125	74.5	IPR001223 (PFAM)
scaffold-TFYI-2066800-Marchantia_emarginata	Liverworts	Marchantiaceae	acidic mammalian chitinase-like	6,07755E-72	72.82	IPR001223 (PFAM)
scaffold-TFYI-2066998-Marchantia_emarginata	Liverworts	Marchantiaceae	acidic endochitinase	3,08156E-69	80.8	IPR001223 (PFAM)
scaffold-TFYI-2068733-Marchantia_emarginata	Liverworts	Marchantiaceae	Acidic endochitinase	8,881E-44	73.83	IPR001223 (PFAM)
scaffold-TFYI-2073368-Marchantia_emarginata	Liverworts	Marchantiaceae	acidic endochitinase	5,49591E-115	71.76	IPR001223 (PFAM)
scaffold-TFYI-2073779-Marchantia_emarginata	Liverworts	Marchantiaceae	hevamine-A-like	1,27042E-89	68.23	IPR001223 (PFAM)
scaffold-TGKW-2013377-Frullania	Liverworts	Frullaniaceae	acidic mammalian chitinase-like	5,43486E-122	57.48	IPR001223 (PFAM)
scaffold-TGKW-2013378-Frullania	Liverworts	Frullaniaceae	acidic mammalian chitinase-like	5,43486E-122	57.48	IPR001223 (PFAM)
scaffold-TGKW-2125559-Frullania	Liverworts	Frullaniaceae	acidic mammalian chitinase-like	1,86899E-39	71.31	IPR001223 (PFAM)
scaffold-TMAJ-2010269-Neckera_douglasii	Mosses	Neckeraceae	class V chitinase-like	1,41027E-171	63.33	IPR001223 (PFAM)
scaffold-TMAJ-2010270-Neckera_douglasii	Mosses	Neckeraceae	class V chitinase-like	1,41027E-171	63.33	IPR001223 (PFAM)
scaffold-TMAJ-2010271-Neckera_douglasii	Mosses	Neckeraceae	class V chitinase-like	1,41027E-171	63.33	IPR001223 (PFAM)
scaffold-TMAJ-2021574-Neckera_douglasii	Mosses	Neckeraceae	family 18 glycoside hydrolase	2,44549E-45	62.72	IPR001223 (PFAM)
scaffold-TMAJ-2023280-Neckera_douglasii	Mosses	Neckeraceae	class V chitinase-like	3,54382E-83	62.71	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-TMAJ-2166782-Neckera_douglasii	Mosses	Neckeraceae	glycoside hydrolase	3,73592E-35	65.55	IPR001223 (PFAM)
scaffold-TMAJ-2174094-Neckera_douglasii	Mosses	Neckeraceae	glycoside hydrolase family 18	4,96823E-64	78.0	IPR001223 (PFAM)
scaffold-TWUW-2049624-Anthoceros_agrestis-A	Hornworts	Anthoceroceae	acidic mammalian chitinase-like	5,54997E-70	65.44	IPR001223 (PFAM)
scaffold-TXVB-2013493-Lunularia_cruciata	Liverworts	Lunulariaceae	Acidic endochitinase	4,32158E-132	71.89	IPR001223 (PFAM)
scaffold-TXVB-2063412-Lunularia_cruciata	Liverworts	Lunulariaceae	acidic endochitinase	4,40415E-35	74.58	IPR001223 (PFAM)
scaffold-TXVB-2070336-Lunularia_cruciata	Liverworts	Lunulariaceae	acidic endochitinase	1,61757E-68	75.65	IPR001223 (PFAM)
scaffold-TXVB-2072783-Lunularia_cruciata	Liverworts	Lunulariaceae	chitinase family protein	5,44521E-92	67.78	IPR001223 (PFAM)
scaffold-TXVB-2074185-Lunularia_cruciata	Liverworts	Lunulariaceae	acidic mammalian chitinase-like	1,77591E-143	61.89	IPR001223 (PFAM)
scaffold-UCRN-2052015-Megaceros_tosanus	Hornworts	Dendrocerotaceae	class V chitinase-like	1,85221E-95	64.11	IPR001223 (PFAM)
scaffold-UUHD-2149435-Porella_pinnata	Liverworts	Porellaceae	acidic mammalian chitinase-like	1,62164E-125	61.68	IPR001223 (PFAM)
scaffold-VBMM-2011897-Anomodon_rostratus	Mosses	Brachytheciaceae	class V chitinase-like	7,63365E-169	62.94	IPR001223 (PFAM)
scaffold-VMXJ-2127708-Leucobryum_albidum	Mosses	Leucobryaceae	acidic mammalian chitinase-like	1,09517E-101	65.44	IPR001223 (PFAM)
scaffold-WCZB-2004735-Phaeoceros_carolinianus-gametophyte	Hornworts	Notothyladaceae	glycoside hydrolase family 18 protein	1,29928E-119	80.06	IPR001223 (PFAM)
scaffold-WCZB-2004736-Phaeoceros_carolinianus-gametophyte	Hornworts	Notothyladaceae	glycoside hydrolase family 18 protein	1,29928E-119	80.06	IPR001223 (PFAM)
scaffold-WCZB-2104761-Phaeoceros_carolinianus-gametophyte	Hornworts	Notothyladaceae	AChain A	3,27928E-68	96.78	IPR001223 (PFAM)
scaffold-WCZB-2114497-Phaeoceros_carolinianus-gametophyte	Hornworts	Notothyladaceae	chitinase-like protein Idgf2	6,73286E-43	87.2	IPR001223 (PFAM)
scaffold-WCZB-2116905-Phaeoceros_carolinianus-gametophyte	Hornworts	Notothyladaceae	chitinase-like protein Idgf4	7,23016E-53	94.66	IPR001223 (PFAM)
scaffold-WCZB-2116970-Phaeoceros_carolinianus-gametophyte	Hornworts	Notothyladaceae	chitinase-3-like protein 1	6,12448E-87	63.7	IPR001223 (PFAM)
scaffold-WCZB-2119571-Phaeoceros_carolinianus-gametophyte	Hornworts	Notothyladaceae	class V chitinase-like	3,01785E-81	62.26	IPR001223 (PFAM)
scaffold-WEEQ-2089607-Phaeoceros_carolinianus-gametophyte_1575	Hornworts	Notothyladaceae	class V chitinase-like	4,39733E-84	60.93	IPR001223 (PFAM)
scaffold-WJLO-2007388-Riccia_berychiana	Liverworts	Ricciaceae	Acidic endochitinase	6,85753E-112	70.28	IPR001223 (PFAM)
scaffold-WJLO-2007710-Riccia_berychiana	Liverworts	Ricciaceae	acidic mammalian chitinase-like	3,69849E-171	63.68	IPR001223 (PFAM)
scaffold-WJLO-2035839-Riccia_berychiana	Liverworts	Ricciaceae	acidic endochitinase-like	2,31588E-136	74.4	IPR001223 (PFAM)
scaffold-WNGH-2008645-Aulacomnium_heterostichum	Mosses	Aulacomniaceae	class V chitinase-like	1,03798E-171	62.82	IPR001223 (PFAM)
scaffold-WNGH-2008646-Aulacomnium_heterostichum	Mosses	Aulacomniaceae	class V chitinase-like	2,00165E-66	63.09	IPR001223 (PFAM)
scaffold-WNGH-2014912-Aulacomnium_heterostichum	Mosses	Aulacomniaceae	putative Chitinase	1,28673E-63	60.72	IPR001223 (PFAM)
scaffold-WNGH-2088662-Aulacomnium_heterostichum	Mosses	Aulacomniaceae	glycoside hydrolase	6,85689E-75	63.46	IPR001223 (PFAM)
scaffold-WSPM-2000719-Rhytidiadelphus_loreus	Mosses	Hylocomiaceae	class V chitinase-like	7,8804E-118	65.49	IPR001223 (PFAM)

BRYOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-WSPM-2000720-Rhytidiadelphus_loreus	Mosses	Hylocomiaceae	class V chitinase-like	2,97353E-69	69.62	IPR001223 (PFAM)
scaffold-WZYK-2086114-Bazzania_trilobata	Liverworts	Lepidoziaceae	acidic mammalian chitinase-like	5,46974E-67	65.81	IPR001223 (PFAM)
scaffold-XWHK-2008595-Funaria	Mosses	Bryaceae	acidic mammalian chitinase-like	1,84147E-119	67.12	IPR001223 (PFAM)
scaffold-XWHK-2045122-Funaria	Mosses	Bryaceae	acidic mammalian chitinase-like	4,20745E-65	66.69	IPR001223 (PFAM)
scaffold-YBQN-2014678-Odontoschisma_prostratum	Liverworts	Cephaloziaceae	glycoside hydrolase family 18	2,40124E-108	63.95	IPR001223 (PFAM)
scaffold-YBQN-2119580-Odontoschisma_prostratum	Liverworts	Cephaloziaceae	probable chitinase 3	1,33636E-19	49.95	IPR001223 (PFAM)
scaffold-YBQN-2120139-Odontoschisma_prostratum	Liverworts	Cephaloziaceae	putative chitinase 3	3,97672E-26	62.49	IPR001223 (PFAM)
scaffold-YBQN-2120337-Odontoschisma_prostratum	Liverworts	Cephaloziaceae	acidic mammalian chitinase-like	2,74304E-15	52.44	IPR001223 (PFAM)
scaffold-YBQN-2121296-Odontoschisma_prostratum	Liverworts	Cephaloziaceae	glycoside hydrolase family 18	1,98189E-98	84.9	IPR001223 (PFAM)
scaffold-YBQN-2123041-Odontoschisma_prostratum	Liverworts	Cephaloziaceae	glycoside hydrolase family 18	3,86459E-79	84.15	IPR001223 (PFAM)
scaffold-YEPO-2005117-cf._Physcomicromitrium_sp	Mosses	Funariaceae	oviductal secretory glycoprotein	7,16217E-129	54.3	IPR001223 (PFAM)
scaffold-YEPO-2005679-cf._Physcomicromitrium_sp	Mosses	Funariaceae	class V chitinase	0.0	64.27	IPR001223 (PFAM)
scaffold-YWNF-2048821-Hedwigia_ciliata	Mosses	Hedwigiaceae	glycoside hydrolase family 18	5,44232E-35	69.65	IPR001223 (PFAM)
scaffold-YWNF-2051820-Hedwigia_ciliata	Mosses	Hedwigiaceae	class V chitinase	7,40555E-170	62.66	IPR001223 (PFAM)
scaffold-ZACW-2001860-Leucodon_sciuroides	Mosses	Leucodontaceae	acidic mammalian chitinase-like	2,2733E-173	62.5	IPR001223 (PFAM)
scaffold-ZFRE-2107597-Phaeoceros_carolinianus-sporophyte_1575	Hornworts	Notothyladaceae	acidic mammalian chitinase-like	3,53694E-42	66.32	IPR001223 (PFAM)
scaffold-ZQRI-2018149-Timmia_austriaca	Mosses	Timmiaceae	acidic mammalian chitinase	2,85841E-146	61.03	IPR001223 (PFAM)
scaffold-ZTHV-2012050-Atrichum_angustatum	Mosses	Polytrichaceae	Glycosyl hydrolase family protein with chitinase insertion domain-containing protein	3,53703E-65	62.56	IPR001223 (PFAM)
scaffold-ZTHV-2017150-Atrichum_angustatum	Mosses	Polytrichaceae	acidic mammalian chitinase-like	1,07991E-139	62.18	IPR001223 (PFAM)
scaffold-ZTHV-2070247-Atrichum_angustatum	Mosses	Polytrichaceae	chitinase-3-like protein 1	8,77938E-33	69.83	IPR001223 (PFAM)
scaffold-ZTHV-2075062-Atrichum_angustatum	Mosses	Polytrichaceae	class V chitinase-like	9,85696E-32	61.28	IPR001223 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-ABIJ-2006436-Selaginella_lepidophylla	Lycophytes	Selaginellaceae	acidic endochitinase	7,37915E-133	81.16	IPR001223 (PFAM)
scaffold-ABIJ-2008792-Selaginella_lepidophylla	Lycophytes	Selaginellaceae	class V chitinase	2,20284E-44	73.88	IPR001223 (PFAM)
scaffold-ABIJ-2008890-Selaginella_lepidophylla	Lycophytes	Selaginellaceae	class V chitinase-like	1,47507E-107	68.38	IPR001223 (PFAM)
scaffold-ABIJ-2009255-Selaginella_lepidophylla	Lycophytes	Selaginellaceae	acidic endochitinase	2,06505E-129	81.3	IPR001223 (PFAM)
scaffold-ABIJ-2038774-Selaginella_lepidophylla	Lycophytes	Selaginellaceae	acidic endochitinase	2,56057E-127	79.24	IPR001223 (PFAM)
scaffold-ABIJ-2038934-Selaginella_lepidophylla	Lycophytes	Selaginellaceae	class V chitinase-like	3,18407E-174	67.99	IPR001223 (PFAM)
scaffold-ALVQ-2001326-Tmesipteris_parva	Eusporangiate Monilophytes	Psilotaceae	acidic mammalian chitinase-like	2,48305E-06	64.34	IPR001223 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-ALVQ-2001327-Tmesipteris_parva	Eusporangiate Monilophytes	Psilotaceae	acidic mammalian chitinase-like	3,47019E-36	58.48	IPR001223 (PFAM)
scaffold-ALVQ-2016777-Tmesipteris_parva	Eusporangiate Monilophytes	Psilotaceae	class V chitinase	6,52882E-78	63.56	IPR001223 (PFAM)
scaffold-BEGM-2017282-Botrypus_virginianus	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase	5,28469E-36	60.59	IPR001223 (PFAM)
scaffold-BEGM-2102286-Botrypus_virginianus	Eusporangiate Monilophytes	Ophioglossaceae	acidic endochitinase-like	4,10295E-113	76.97	IPR001223 (PFAM)
scaffold-BEGM-2102306-Botrypus_virginianus	Eusporangiate Monilophytes	Ophioglossaceae	acidic mammalian chitinase-like	1,61725E-74	61.74	IPR001223 (PFAM)
scaffold-CAPN-2006447-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	chitotriosidase-1	6,87591E-41	68.83	IPR001223 (PFAM)
scaffold-CAPN-2006448-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	class V chitinase	1,36146E-34	69.07	IPR001223 (PFAM)
scaffold-CAPN-2009848-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	class V chitinase	3,15691E-81	61.44	IPR001223 (PFAM)
scaffold-CAPN-2037054-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	acidic mammalian chitinase-like	2,46257E-12	54.61	IPR001223 (PFAM)
scaffold-CAPN-2037262-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	AChain A	2,51567E-12	61.46	IPR001223 (PFAM)
scaffold-CAPN-2037836-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	acidic mammalian chitinase-like	1,12237E-23	59.72	IPR001223 (PFAM)
scaffold-CAPN-2038165-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	class V chitinase-like	2,25621E-31	59.92	IPR001223 (PFAM)
scaffold-CAPN-2039382-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	Acidic endochitinase	2,18968E-101	73.68	IPR001223 (PFAM)
scaffold-CAPN-2040486-Equisetum_diffusum	Eusporangiate Monilophytes	Equisetaceae	class V chitinase	1,03037E-75	63.15	IPR001223 (PFAM)
scaffold-CBAE-2001838-Huperzia_myrisinites	Lycophytes	Huperziaceae	acidic endochitinase-like	2,05668E-41	69.81	IPR001223 (PFAM)
scaffold-CBAE-2001839-Huperzia_myrisinites	Lycophytes	Huperziaceae	acidic endochitinase	2,51333E-36	74.34	IPR001223 (PFAM)
scaffold-CBAE-2005523-Huperzia_myrisinites	Lycophytes	Huperziaceae	acidic endochitinase	5,70093E-79	66.78	IPR001223 (PFAM)
scaffold-CBAE-2005524-Huperzia_myrisinites	Lycophytes	Huperziaceae	acidic endochitinase	1,00309E-88	71.26	IPR001223 (PFAM)
scaffold-CBAE-2051921-Huperzia_myrisinites	Lycophytes	Huperziaceae	acidic mammalian chitinase-like	1,46167E-45	70.45	IPR001223 (PFAM)
scaffold-CBAE-2060579-Huperzia_myrisinites	Lycophytes	Huperziaceae	class V chitinase	2,92073E-110	66.64	IPR001223 (PFAM)
scaffold-DFHO-2001788-Danaea_nodosa	Eusporangiate Monilophytes	Marattiaceae	acidic mammalian chitinase-like	4,10254E-29	76.1	IPR001223 (PFAM)
scaffold-DFHO-2007268-Danaea_nodosa	Eusporangiate Monilophytes	Marattiaceae	acidic endochitinase	1,63091E-99	72.69	IPR001223 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-DFHO-2044420-Danaea_nodosa	Eusporangiate Monilophytes	Marattiaceae	acidic mammalian chitinase-like	1,47876E-35	70.86	IPR001223 (PFAM)
scaffold-DFHO-2055341-Danaea_nodosa	Eusporangiate Monilophytes	Marattiaceae	class V chitinase-like	1,20678E-91	63.44	IPR001223 (PFAM)
scaffold-EEAQ-2005608-Sceptridium_dissectum	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase	1,3358E-88	61.98	IPR001223 (PFAM)
scaffold-EEAQ-2005609-Sceptridium_dissectum	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase	1,3358E-88	61.98	IPR001223 (PFAM)
scaffold-EEAQ-2008990-Sceptridium_dissectum	Eusporangiate Monilophytes	Ophioglossaceae	acidic endochitinase	1,45893E-59	72.41	IPR001223 (PFAM)
scaffold-EEAQ-2008991-Sceptridium_dissectum	Eusporangiate Monilophytes	Ophioglossaceae	acidic endochitinase	4,58592E-44	71.83	IPR001223 (PFAM)
scaffold-EEAQ-2008992-Sceptridium_dissectum	Eusporangiate Monilophytes	Ophioglossaceae	Acidic endochitinase	9,00319E-71	76.08	IPR001223 (PFAM)
scaffold-EEAQ-2082693-Sceptridium_dissectum	Eusporangiate Monilophytes	Ophioglossaceae	acidic mammalian chitinase-like	1,05958E-15	59.68	IPR001223 (PFAM)
scaffold-EEAQ-2086905-Sceptridium_dissectum	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase	3,78098E-87	62.0	IPR001223 (PFAM)
scaffold-ENQF-2003741-Lycopodium_annotinum	Lycophytes	Lycopodiaceae	class V chitinase	5,81446E-96	69.33	IPR001223 (PFAM)
scaffold-ENQF-2005526-Lycopodium_annotinum	Lycophytes	Lycopodiaceae	Acidic endochitinase	4,58236E-102	7348	IPR001223 (PFAM)
scaffold-ENQF-2012682-Lycopodium_annotinum	Lycophytes	Lycopodiaceae	class V chitinase	4,39319E-109	67.04	IPR001223 (PFAM)
scaffold-ENQF-2015939-Lycopodium_annotinum	Lycophytes	Lycopodiaceae	class V chitinase-like	1,78278E-63	68.18	IPR001223 (PFAM)
scaffold-ENQF-2017435-Lycopodium_annotinum	Lycophytes	Lycopodiaceae	class V chitinase	2,11436E-90	66.8	IPR001223 (PFAM)
scaffold-GAON-2001240-Huperzia_squarrosa	Lycophytes	Lycopodiaceae	acidic endochitinase	2,47711E-42	79.16	IPR001223 (PFAM)
scaffold-GAON-2001241-Huperzia_squarrosa	Lycophytes	Lycopodiaceae	acidic endochitinase	8,43922E-43	74.67	IPR001223 (PFAM)
scaffold-GAON-2007131-Huperzia_squarrosa	Lycophytes	Lycopodiaceae	acidic endochitinase	7,98915E-88	71.16	IPR001223 (PFAM)
scaffold-GAON-2016742-Huperzia_squarrosa	Lycophytes	Lycopodiaceae	class V chitinase	8,04208E-104	66.99	IPR001223 (PFAM)
scaffold-GAON-2055613-Huperzia_squarrosa	Lycophytes	Lycopodiaceae	class V chitinase	6,34137E-63	69.89	IPR001223 (PFAM)
scaffold-GKAG-2011802-Huperzia_lucidula	Lycophytes	Lycopodiaceae	acidic endochitinase	2,93548E-97	72.27	IPR001223 (PFAM)
scaffold-GKAG-2011803-Huperzia_lucidula	Lycophytes	Lycopodiaceae	acidic endochitinase	9,6752E-98	73.45	IPR001223 (PFAM)
scaffold-GKAG-2017355-Huperzia_lucidula	Lycophytes	Lycopodiaceae	class V chitinase	7,49927E-101	64.52	IPR001223 (PFAM)
scaffold-GKAG-2022275-Huperzia_lucidula	Lycophytes	Lycopodiaceae	class V chitinase	1,66475E-96	68.81	IPR001223 (PFAM)
scaffold-GKAG-2097278-Huperzia_lucidula	Lycophytes	Lycopodiaceae	bacteriodes thetaiotaomicron symbiotic chitinase	1,9125E-131	64.36	IPR001223 (PFAM)
scaffold-GTUO-2001847-Huperzia_selago	Lycophytes	Lycopodiaceae	acidic endochitinase-like	9,47804E-34	74.35	IPR001223 (PFAM)
scaffold-GTUO-2025296-Huperzia_selago	Lycophytes	Lycopodiaceae	acidic mammalian chitinase	7,38789E-103	66.63	IPR001223 (PFAM)
scaffold-GTUO-2026196-Huperzia_selago	Lycophytes	Lycopodiaceae	class V chitinase	1,07813E-57	69.46	IPR001223 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-JKAA-2002507-Selaginella_wallacei	Lycophytes	Selaginellaceae	class V chitinase-like	1,00998E-95	621.4	IPR001223 (PFAM)
scaffold-JKAA-2002508-Selaginella_wallacei	Lycophytes	Selaginellaceae	class V chitinase-like	5,93482E-83	65.9	IPR001223 (PFAM)
scaffold-JKAA-2008532-Selaginella_wallacei	Lycophytes	Selaginellaceae	acidic mammalian chitinase-like	5,55276E-18	61.15	IPR001223 (PFAM)
scaffold-JKAA-2013110-Selaginella_wallacei	Lycophytes	Selaginellaceae	acidic endochitinase	8,65916E-94	77.2	IPR001223 (PFAM)
scaffold-JKAA-2158634-Selaginella_wallacei	Lycophytes	Selaginellaceae	class V chitinase	5,47248E-36	65.98	IPR001223 (PFAM)
scaffold-JKAA-2159729-Selaginella_wallacei	Lycophytes	Selaginellaceae	putative class V chitinase	5,17795E-37	77.41	IPR001223 (PFAM)
scaffold-JKAA-2169170-Selaginella_wallacei	Lycophytes	Selaginellaceae	glycoside hydrolase	1,65812E-39	73.95	IPR001223 (PFAM)
scaffold-JKAA-2173170-Selaginella_wallacei	Lycophytes	Selaginellaceae	glycoside hydrolase family 18	1,39492E-60	89.3	IPR001223 (PFAM)
scaffold-JKAA-2174276-Selaginella_wallacei	Lycophytes	Selaginellaceae	glycoside hydrolase family 18	5,40609E-32	63.2	IPR001223 (PFAM)
scaffold-JKAA-2177540-Selaginella_wallacei	Lycophytes	Selaginellaceae	acidic endochitinase-like	1,42583E-123	77.19	IPR001223 (PFAM)
scaffold-JKAA-2178165-Selaginella_wallacei	Lycophytes	Selaginellaceae	glycoside hydrolase	1,57212E-70	83.97	IPR001223 (PFAM)
scaffold-JKAA-2178880-Selaginella_wallacei	Lycophytes	Selaginellaceae	acidic endochitinase	7,27741E-115	78.35	IPR001223 (PFAM)
scaffold-JKAA-2180047-Selaginella_wallacei	Lycophytes	Selaginellaceae	glycoside hydrolase family 18	1,29131E-143	85.45	IPR001223 (PFAM)
scaffold-JKAA-2180911-Selaginella_wallacei	Lycophytes	Selaginellaceae	glycoside hydrolase	7,29076E-72	63.35	IPR001223 (PFAM)
scaffold-JKAA-2181250-Selaginella_wallacei	Lycophytes	Selaginellaceae	glycoside hydrolase	1,30916E-69	64.19	IPR001223 (PFAM)
scaffold-JVSZ-2002787-Equisetum_hymale	Eusporangiate Monilophytes	Equisetaceae	acidic mammalian chitinase-like	2,55616E-78	61.42	IPR001223 (PFAM)
scaffold-JVSZ-2002788-Equisetum_hymale	Eusporangiate Monilophytes	Equisetaceae	class V chitinase	7,93305E-78	61.15	IPR001223 (PFAM)
scaffold-JVSZ-2002789-Equisetum_hymale	Eusporangiate Monilophytes	Equisetaceae	class V chitinase-like	1,12384E-09	64.97	IPR001223 (PFAM)
scaffold-JVSZ-2119489-Equisetum_hymale	Eusporangiate Monilophytes	Equisetaceae	AChain A	5,8541E-15	59.24	IPR001223 (PFAM)
scaffold-KJYC-2002529-Selaginella_willdenowii	Lycophytes	Selaginellaceae	class V chitinase	5,52651E-46	72.41	IPR001223 (PFAM)
scaffold-KJYC-2006273-Selaginella_willdenowii	Lycophytes	Selaginellaceae	class V chitinase	2,39536E-46	69.56	IPR001223 (PFAM)
scaffold-KJYC-2006274-Selaginella_willdenowii	Lycophytes	Selaginellaceae	class V chitinase	2,39536E-46	69.56	IPR001223 (PFAM)
scaffold-KJYC-2007922-Selaginella_willdenowii	Lycophytes	Selaginellaceae	class V chitinase	5,31759E-65	71.76	IPR001223 (PFAM)
scaffold-KJYC-2013372-Selaginella_willdenowii	Lycophytes	Selaginellaceae	acidic endochitinase	1,28264E-86	67.36	IPR001223 (PFAM)
scaffold-KJYC-2044359-Selaginella_willdenowii	Lycophytes	Selaginellaceae	acidic mammalian chitinase-like	1,35468E-36	68.81	IPR001223 (PFAM)
scaffold-KJYC-2047116-Selaginella_willdenowii	Lycophytes	Selaginellaceae	acidic endochitinase	1,38735E-56	77.88	IPR001223 (PFAM)
scaffold-KJYC-2047493-Selaginella_willdenowii	Lycophytes	Selaginellaceae	acidic endochitinase	2,15134E-62	78.68	IPR001223 (PFAM)
scaffold-KJYC-2048999-Selaginella_willdenowii	Lycophytes	Selaginellaceae	class V chitinase-like	4,43578E-93	70.79	IPR001223 (PFAM)
scaffold-KUXM-2001058-Selaginella_selaginoides	Lycophytes	Selaginellaceae	class V chitinase-like	1,09246E-13	67.59	IPR001223 (PFAM)
scaffold-KUXM-2009202-Selaginella_selaginoides	Lycophytes	Selaginellaceae	class V chitinase	2,17918E-48	63.68	IPR001223 (PFAM)
scaffold-KUXM-2043003-Selaginella_selaginoides	Lycophytes	Selaginellaceae	acidic endochitinase	6,68209E-136	77.97	IPR001223 (PFAM)
scaffold-LGDQ-2008640-Selaginella_apoda	Lycophytes	Selaginellaceae	class V chitinase-like	0.0	69.48	IPR001223 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-LGDQ-2008641-Selaginella_apoda	Lycophytes	Selaginellaceae	class V chitinase-like	0.0	69.48	IPR001223 (PFAM)
scaffold-LGDQ-2011793-Selaginella_apoda	Lycophytes	Selaginellaceae	acidic mammalian chitinase	6,76004E-158	65.8	IPR001223 (PFAM)
scaffold-LGDQ-2040164-Selaginella_apoda	Lycophytes	Selaginellaceae	acidic endochitinase	3,67656E-82	80.34	IPR001223 (PFAM)
scaffold-LGDQ-2041235-Selaginella_apoda	Lycophytes	Selaginellaceae	acidic endochitinase	1,43414E-78	77.87	IPR001223 (PFAM)
scaffold-NYBX-2005847-Huperzia_selago-A	Lycophytes	Lycopodiaceae	chitinase-3-like protein 1	5,43986E-38	52.6	IPR001223 (PFAM)
scaffold-NYBX-2006529-Huperzia_selago-A	Lycophytes	Lycopodiaceae	acidic endochitinase	1,02835E-101	74.97	IPR001223 (PFAM)
scaffold-NYBX-2006530-Huperzia_selago-A	Lycophytes	Lycopodiaceae	acidic endochitinase	1,02835E-101	74.97	IPR001223 (PFAM)
scaffold-NYBX-2006531-Huperzia_selago-A	Lycophytes	Lycopodiaceae	acidic endochitinase	1,02835E-101	74.97	IPR001223 (PFAM)
scaffold-NYBX-2021614-Huperzia_selago-A	Lycophytes	Lycopodiaceae	chitinase-3-like protein 1	1,62726E-126	67.53	IPR001223 (PFAM)
scaffold-NYBX-2021615-Huperzia_selago-A	Lycophytes	Lycopodiaceae	chitinase-3-like protein 1	8,32011E-105	68.91	IPR001223 (PFAM)
scaffold-NYBX-2023892-Huperzia_selago-A	Lycophytes	Lycopodiaceae	class V chitinase	3,84357E-42	64.89	IPR001223 (PFAM)
scaffold-NYBX-2120238-Huperzia_selago-A	Lycophytes	Lycopodiaceae	probable chitinase 10	2,3095E-27	83.1	IPR001223 (PFAM)
scaffold-NYBX-2126520-Huperzia_selago-A	Lycophytes	Lycopodiaceae	probable chitinase 10	2,91486E-77	75.64	IPR001223 (PFAM)
scaffold-NYBX-2129604-Huperzia_selago-A	Lycophytes	Lycopodiaceae	class V chitinase-like	4,85702E-71	71.44	IPR001223 (PFAM)
scaffold-NYBX-2130744-Huperzia_selago-A	Lycophytes	Lycopodiaceae	chitinase-3-like protein 1	6,06156E-143	70.8	IPR001223 (PFAM)
scaffold-PKOX-2000311-Isoetes_tegetiformans	Lycophytes	Isoetaceae	acidic endochitinase	1,69475E-100	78.92	IPR001223 (PFAM)
scaffold-NWWI-2025427-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	chitotriosidase-1-like	3,83769E-66	62.39	IPR001223 (PFAM)
scaffold-NWWI-2025428-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	acidic mammalian chitinase-like	3,37917E-66	62.86	IPR001223 (PFAM)
scaffold-NWWI-2005288-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	class IV chitinase	1,4924E-77	67.87	IPR000726 (PFAM)
scaffold-NWWI-2005289-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	class IV chitinase	1,81351E-78	67.22	IPR000726 (PFAM)
scaffold-NWWI-2005290-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	class IV chitinase	2,08408E-74	66.77	IPR000726 (PFAM)
scaffold-NWWI-2005291-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	class IV chitinase	1,14732E-75	69.2	IPR000726 (PFAM)
scaffold-NWWI-2008181-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	endochitinase EP3-like	1,03823E-68	57.63	IPR000726 (PFAM)
scaffold-NWWI-2008182-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	putative class I chitinase	1,16154E-50	65.41	IPR000726 (PFAM)
scaffold-NWWI-2008183-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	endochitinase EP3-like	9,25697E-64	69.17	IPR000726 (PFAM)
scaffold-NWWI-2011801-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	putative class I chitinase	2,45443E-35	69.66	IPR000726 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-NWWI-2011802-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	putative class I chitinase	2,45443E-35	69.66	IPR000726 (PFAM)
scaffold-NWWI-2014583-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	Glycoside hydrolase	1,04812E-27	56.31	IPR000726 (PFAM)
scaffold-NWWI-2126383-Nephrolepis_exaltata	Leptosporangiate Monilophytes	Lomariopsidaceae	endochitinase EP3-like	1,22529E-54	72.41	IPR000726 (PFAM)
scaffold-PKOX-2093371-Isoetes_tegetiformans	Lycophytes	Isoetaceae	chitinase-3-like protein 1	1,44362E-35	60.98	IPR001223 (PFAM)
scaffold-PKOX-2097106-Isoetes_tegetiformans	Lycophytes	Isoetaceae	class V chitinase	2,30735E-103	67.04	IPR001223 (PFAM)
scaffold-PKOX-2098297-Isoetes_tegetiformans	Lycophytes	Isoetaceae	acidic mammalian chitinase	9,02551E-120	68.13	IPR001223 (PFAM)
scaffold-PQTO-2012178-Lycopodium_deuterodensum	Lycophytes	Lycopodiaceae	acidic mammalian chitinase-like	2,35836E-112	66.97	IPR001223 (PFAM)
scaffold-PQTO-2076779-Lycopodium_deuterodensum	Lycophytes	Lycopodiaceae	class V chitinase	4,35967E-68	67.06	IPR001223 (PFAM)
scaffold-PQTO-2079202-Lycopodium_deuterodensum	Lycophytes	Lycopodiaceae	acidic endochitinase	1,68223E-101	73.76	IPR001223 (PFAM)
scaffold-PYHZ-2003251-Isoetes_sp	Lycophytes	Isoetaceae	acidic mammalian chitinase-like	1,91385E-125	67.86	IPR001223 (PFAM)
scaffold-PYHZ-2003252-Isoetes_sp	Lycophytes	Isoetaceae	class V chitinase	4,08874E-29	71.38	IPR001223 (PFAM)
scaffold-PYHZ-2003253-Isoetes_sp	Lycophytes	Isoetaceae	class V chitinase	1,6321E-112	67.3	IPR001223 (PFAM)
scaffold-PYHZ-2004504-Isoetes_sp	Lycophytes	Isoetaceae	acidic endochitinase	1,25255E-99	75.15	IPR001223 (PFAM)
scaffold-PYHZ-2007052-Isoetes_sp	Lycophytes	Isoetaceae	class V chitinase	2,76274E-78	57.02	IPR001223 (PFAM)
scaffold-PYHZ-2007053-Isoetes_sp	Lycophytes	Isoetaceae	class V chitinase	5,77231E-75	58.66	IPR001223 (PFAM)
scaffold-PYHZ-2007870-Isoetes_sp	Lycophytes	Isoetaceae	acidic endochitinase	1,99239E-100	77.93	IPR001223 (PFAM)
scaffold-PYHZ-2007871-Isoetes_sp	Lycophytes	Isoetaceae	hevamine-A-like	7,80435E-109	78.57	IPR001223 (PFAM)
scaffold-PYHZ-2007872-Isoetes_sp	Lycophytes	Isoetaceae	acidic endochitinase	2,79998E-93	78.34	IPR001223 (PFAM)
scaffold-PYHZ-2009287-Isoetes_sp	Lycophytes	Isoetaceae	class V chitinase-like	7,33748E-64	58.66	IPR001223 (PFAM)
scaffold-PYHZ-2071770-Isoetes_sp	Lycophytes	Isoetaceae	Glycosyl hydrolase family protein with chitinase insertion domain-containing	4,87098E-114	68.97	IPR001223 (PFAM)
scaffold-PYHZ-2073439-Isoetes_sp	Lycophytes	Isoetaceae	class V chitinase	5,71777E-103	66.54	IPR001223 (PFAM)
scaffold-PYHZ-2075391-Isoetes_sp	Lycophytes	Isoetaceae	Glycosyl hydrolase family protein with chitinase insertion domain-containing	1,82802E-128	66.44	IPR001223 (PFAM)
scaffold-QVMR-2015945-Psilotum_nudum	Eusporangiate Monilophytes	Psilotaceae	class V chitinase	1,30096E-43	66.92	IPR001223 (PFAM)
scaffold-QVMR-2017494-Psilotum_nudum	Eusporangiate Monilophytes	Psilotaceae	class V chitinase	1,68136E-91	62.54	IPR001223 (PFAM)
scaffold-UGNK-2002386-Marattia_attenuata	Eusporangiate Monilophytes	Marattiaceae	class V chitinase-like	6,23691E-97	63.35	IPR001223 (PFAM)
scaffold-UGNK-2002387-Marattia_attenuata	Eusporangiate Monilophytes	Marattiaceae	class V chitinase-like	2,0814E-78	65.28	IPR001223 (PFAM)
scaffold-UGNK-2016335-Marattia_attenuata	Eusporangiate Monilophytes	Marattiaceae	acidic mammalian chitinase-like	3,21432E-93	64.19	IPR001223 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-UGNK-2019481-Marattia_attenuata	Eusporangiate Monilophytes	Marattiaceae	class V chitinase-like	4,26556E-53	62.32	IPR001223 (PFAM)
scaffold-ULKT-2008784-Lycopodiella_apressa	Lycophytes	Lycopodiaceae	acidic endochitinase	2,01495E-94	72.16	IPR001223 (PFAM)
scaffold-ULKT-2011128-Lycopodiella_apressa	Lycophytes	Lycopodiaceae	class V chitinase	1,52648E-71	67.32	IPR001223 (PFAM)
scaffold-ULKT-2011129-Lycopodiella_apressa	Lycophytes	Lycopodiaceae	acidic mammalian chitinase	1,77605E-111	67.96	IPR001223 (PFAM)
scaffold-ULKT-2015752-Lycopodiella_apressa	Lycophytes	Lycopodiaceae	acidic endochitinase	7,00327E-107	72.7	IPR001223 (PFAM)
scaffold-ULKT-2016810-Lycopodiella_apressa	Lycophytes	Lycopodiaceae	class V chitinase	4,36749E-32	73.85	IPR001223 (PFAM)
scaffold-UPMJ-2005627-Pseudolycopodiella_caroliniana	Lycophytes	Lycopodiaceae	Acidic endochitinase	6,80861E-104	73.56	IPR001223 (PFAM)
scaffold-UPMJ-2009660-Pseudolycopodiella_caroliniana	Lycophytes	Lycopodiaceae	class V chitinase	7,30261E-75	68.9	IPR001223 (PFAM)
scaffold-UPMJ-2018058-Pseudolycopodiella_caroliniana	Lycophytes	Lycopodiaceae	acidic endochitinase	1,28806E-105	72.48	IPR001223 (PFAM)
scaffold-UPMJ-2077280-Pseudolycopodiella_caroliniana	Lycophytes	Lycopodiaceae	acidic endochitinase	2,28796E-24	71.55	IPR001223 (PFAM)
scaffold-UPMJ-2082325-Pseudolycopodiella_caroliniana	Lycophytes	Lycopodiaceae	class V chitinase	1,0466E-114	68.74	IPR001223 (PFAM)
scaffold-UXCS-2007028-Marattia_sp	Eusporangiate Monilophytes	Marattiaceae	acidic endochitinase	3,11625E-47	71.83	IPR001223 (PFAM)
scaffold-UXCS-2007029-Marattia_sp	Eusporangiate Monilophytes	Marattiaceae	acidic endochitinase	5,46914E-90	75.27	IPR001223 (PFAM)
scaffold-UXCS-2016006-Marattia_sp	Eusporangiate Monilophytes	Marattiaceae	class V chitinase-like	2,84033E-57	63.13	IPR001223 (PFAM)
scaffold-UXCS-2120386-Marattia_sp	Eusporangiate Monilophytes	Marattiaceae	class V chitinase	1,70031E-07	59.85	IPR001223 (PFAM)
scaffold-VVRN-2010946-Lonchitis_hirsuta	Leptosporangiate Monilophytes	Lonchitidaceae	class V chitinase	1,02648E-17	52.99	IPR001223 (PFAM)
scaffold-VVRN-2013897-Lonchitis_hirsuta	Leptosporangiate Monilophytes	Lonchitidaceae	class V chitinase	1,58384E-80	59.8	IPR001223 (PFAM)
scaffold-VVRN-2022226-Lonchitis_hirsuta	Leptosporangiate Monilophytes	Lonchitidaceae	class V chitinase	9,75699E-67	57.1	IPR001223 (PFAM)
scaffold-VVRN-2095349-Lonchitis_hirsuta	Leptosporangiate Monilophytes	Lonchitidaceae	Acidic mammalian chitinase	2,83484E-28	52.52	IPR001223 (PFAM)
scaffold-VVRN-2096647-Lonchitis_hirsuta	Leptosporangiate Monilophytes	Lonchitidaceae	chitotriosidase-1-like isoform X2	1,71199E-70	60.5	IPR001223 (PFAM)
scaffold-WAFT-2001766-Diphasiastrum_digitatum	Lycophytes	Lycopodiaceae	class V chitinase	1,1233E-105	66.31	IPR001223 (PFAM)
scaffold-WAFT-2002690-Diphasiastrum_digitatum	Lycophytes	Lycopodiaceae	Acidic endochitinase	1,78929E-93	72.89	IPR001223 (PFAM)
scaffold-WAFT-2014739-Diphasiastrum_digitatum	Lycophytes	Lycopodiaceae	class V chitinase	1,48598E-120	66.2	IPR001223 (PFAM)

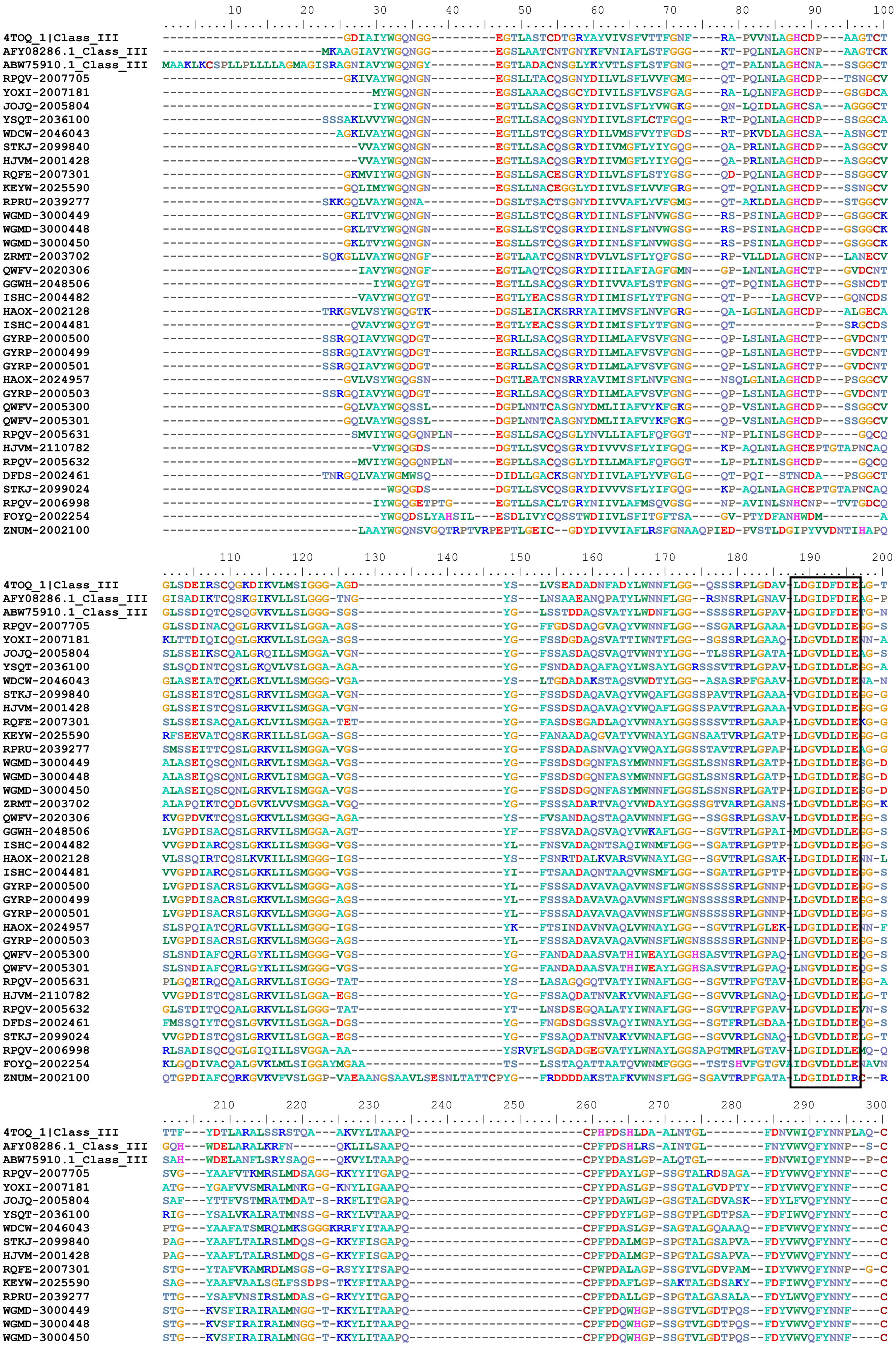
PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-WAFT-2068592-Diphasiastrum_digitatum	Lycophytes	Lycopodiaceae	class V chitinase	5,66453E-80	65.2	IPR001223 (PFAM)
scaffold-WCLG-2004660-Adiantum_aleuticum	Leptosporangiate Monilophytes	Pteridaceae	acidic mammalian chitinase-like	4,03347E-19	57.62	IPR001223 (PFAM)
scaffold-WGTU-2061504-Leucostegia_immersa	Leptosporangiate Monilophytes	Hypodematiaceae	acidic mammalian chitinase-like isoform X1	3,14229E-15	66.76	IPR001223 (PFAM)
scaffold-WGTU-2061726-Leucostegia_immersa	Leptosporangiate Monilophytes	Hypodematiaceae	class V chitinase	8,15691E-30	62.6	IPR001223 (PFAM)
scaffold-WGTU-2071532-Leucostegia_immersa	Leptosporangiate Monilophytes	Hypodematiaceae	acidic endochitinase	1,28058E-91	72.31	IPR001223 (PFAM)
scaffold-WGTU-2072225-Leucostegia_immersa	Leptosporangiate Monilophytes	Hypodematiaceae	class V chitinase	3,17767E-54	62.4	IPR001223 (PFAM)
scaffold-WQML-2003973-Cryptogramma_acrostichoides	Leptosporangiate Monilophytes	Pteridaceae	acidic endochitinase	2,26435E-87	75.49	IPR001223 (PFAM)
scaffold-WQML-2007180-Cryptogramma_acrostichoides	Leptosporangiate Monilophytes	Pteridaceae	chitotriosidase-1-like	2,56956E-65	59.85	IPR001223 (PFAM)
scaffold-WTJG-2002514-Ophioglossum_vulgatum	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase	1,03948E-37	59.52	IPR001223 (PFAM)
scaffold-WTJG-2002515-Ophioglossum_vulgatum	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase	3,84283E-37	59.94	IPR001223 (PFAM)
scaffold-WTJG-2016964-Ophioglossum_vulgatum	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase-like	5,74628E-31	70.09	IPR001223 (PFAM)
scaffold-WTJG-2095237-Ophioglossum_vulgatum	Eusporangiate Monilophytes	Ophioglossaceae	class V chitinase	3,25002E-53	63.44	IPR001223 (PFAM)
scaffold-XDDT-2003802-Argyrochosma_nivea	Leptosporangiate Monilophytes	Pteridaceae	chitotriosidase-1-like isoform X2	2,59369E-68	59.2	IPR001223 (PFAM)
scaffold-XDDT-2009479-Argyrochosma_nivea	Leptosporangiate Monilophytes	Pteridaceae	acidic endochitinase	1,50649E-100	74.39	IPR001223 (PFAM)
scaffold-XDDT-2009480-Argyrochosma_nivea	Leptosporangiate Monilophytes	Pteridaceae	acidic endochitinase	2,54611E-100	75.56	IPR001223 (PFAM)
scaffold-XDDT-2073722-Argyrochosma_nivea	Leptosporangiate Monilophytes	Pteridaceae	class V chitinase-like	1,46976E-67	63.43	IPR001223 (PFAM)
scaffold-XNXF-2066461-Dendrolycopodium_obscurum	Lycophytes	Lycopodiaceae	class V chitinase-like	1,70055E-63	70.04	IPR001223 (PFAM)
scaffold-XXHP-2001291-Cystopteris_fragilis-Stewart	Leptosporangiate Monilophytes	Cystopteridaceae	acidic endochitinase-like	1,2233E-83	74.28	IPR001223 (PFAM)
scaffold-XXHP-2013082-Cystopteris_fragilis-Stewart	Leptosporangiate Monilophytes	Cystopteridaceae	class V chitinase	1,55666E-59	60.12	IPR001223 (PFAM)
scaffold-XXHP-2015468-Cystopteris_fragilis-Stewart	Leptosporangiate Monilophytes	Cystopteridaceae	chitinase-like protein EN03	6,81729E-151	66.53	IPR001223 (PFAM)

PTERIDOPHYTES – GH18

SeqName	Phylum	Family	Description	e-Value	Similarity	InterPro GO IDs
scaffold-XXHP-2015710-Cystopteris_fragilis-Stewart	Leptosporangiate Monilophytes	Cystopteridaceae	acidic mammalian chitinase-like	1,00029E-24	48.85	IPR001223 (PFAM)
scaffold-XXHP-2015711-Cystopteris_fragilis-Stewart	Leptosporangiate Monilophytes	Cystopteridaceae	acidic mammalian chitinase-like	1,32015E-24	48.92	IPR001223 (PFAM)
scaffold-XXHP-2080057-Cystopteris_fragilis-Stewart	Leptosporangiate Monilophytes	Cystopteridaceae	class V chitinase	2,18375E-26	65.19	IPR001223 (PFAM)
scaffold-XXHP-2083486-Cystopteris_fragilis-Stewart	Leptosporangiate Monilophytes	Cystopteridaceae	acidic mammalian chitinase-like	3,22284E-32	57.6	IPR001223 (PFAM)
scaffold-YCKE-2009010-Notholaena_montieliae	Leptosporangiate Monilophytes	Pteridaceae	chitotriosidase-1-like	2,68089E-43	68.81	IPR001223 (PFAM)
scaffold-YCKE-2054510-Notholaena_montieliae	Leptosporangiate Monilophytes	Pteridaceae	class V chitinase-like	4,3756E-18	63.5	IPR001223 (PFAM)
scaffold-YCKE-2060473-Notholaena_montieliae	Leptosporangiate Monilophytes	Pteridaceae	acidic endochitinase	2,77449E-65	73.28	IPR001223 (PFAM)
scaffold-YIXP-2008121-Lindsaea_microphylla	Leptosporangiate Monilophytes	Lindsaeaceae	class V chitinase	9,70482E-28	55.15	IPR001223 (PFAM)
scaffold-YJJY-2011496-Woodsia_scopulina	Leptosporangiate Monilophytes	Woodsiaceae	class V chitinase	1,94271E-80	58.92	IPR001223 (PFAM)
scaffold-YJJY-2071484-Woodsia_scopulina	Leptosporangiate Monilophytes	Woodsiaceae	class V chitinase	7,33428E-35	69.06	IPR001223 (PFAM)
scaffold-YJJY-2075390-Woodsia_scopulina	Leptosporangiate Monilophytes	Woodsiaceae	chitotriosidase-1-like	5,81232E-66	60.35	IPR001223 (PFAM)
scaffold-YJJY-2075638-Woodsia_scopulina	Leptosporangiate Monilophytes	Woodsiaceae	Acidic endochitinase	1,63225E-71	7221	IPR001223 (PFAM)
scaffold-YJJY-2076688-Woodsia_scopulina	Leptosporangiate Monilophytes	Woodsiaceae	class V chitinase	1,12501E-66	50.03	IPR001223 (PFAM)
scaffold-YLJA-2065344-Polypodium_amorphum	Leptosporangiate Monilophytes	Polypodiaceae	acidic mammalian chitinase	2,60911E-29	66.38	IPR001223 (PFAM)
scaffold-YLJA-2066129-Polypodium_amorphum	Leptosporangiate Monilophytes	Polypodiaceae	acidic endochitinase-like	2,96103E-46	72.11	IPR001223 (PFAM)
scaffold-YLJA-2066868-Polypodium_amorphum	Leptosporangiate Monilophytes	Polypodiaceae	class V chitinase	1,01873E-37	58.89	IPR001223 (PFAM)
scaffold-YOWV-2000206-Cystopteris_protrusa	Leptosporangiate Monilophytes	Cystopteridaceae	class V chitinase	7,95557E-52	59.57	IPR001223 (PFAM)
scaffold-YOWV-2004880-Cystopteris_protrusa	Leptosporangiate Monilophytes	Cystopteridaceae	class V chitinase	5,30312E-31	56.08	IPR001223 (PFAM)
scaffold-YOWV-2008552-Cystopteris_protrusa	Leptosporangiate Monilophytes	Cystopteridaceae	acidic mammalian chitinase-like	1,25271E-20	50.25	IPR001223 (PFAM)
YOWV-2057388-Cystopteris_protrusa	Leptosporangiate Monilophytes	Cystopteridaceae	class V chitinase	3,46563E-12	65.03	IPR001223 (PFAM)

PTERIDOPHYTES – GH18						
SeqName	**Phylum**	**Family**	**Description**	**e-Value**	**Similarity**	**InterPro GO IDs**
scaffold-YOWV-2066886-Cystopteris_protrusa	Leptosporangiate Monilophytes	Cystopteridaceae	acidic endochitinase	1,11201E-81	74.29	IPR001223 (PFAM)
scaffold-YQEC-2002428-Woodsia_ilvensis	Leptosporangiate Monilophytes	Woodsiaceae	Glycosyl hydrolase family protein with chitinase insertion domain-containing	2,21718E-37	69.64	IPR001223 (PFAM)
scaffold-YQEC-2012943-Woodsia_ilvensis	Leptosporangiate Monilophytes	Woodsiaceae	acidic mammalian chitinase-like	2,82211E-41	65.08	IPR001223 (PFAM)
scaffold-YQEC-2074942-Woodsia_ilvensis	Leptosporangiate Monilophytes	Woodsiaceae	class V chitinase	5,55586E-80	59.74	IPR001223 (PFAM)
scaffold-YQEC-2075825-Woodsia_ilvensis	Leptosporangiate Monilophytes	Woodsiaceae	class V chitinase	2,99339E-67	60.44	IPR001223 (PFAM)
scaffold-ZFGK-2039578-Selaginella_kraussiana	Lycophytes	Selaginellaceae	acidic endochitinase	6,75082E-145	78.43	IPR001223 (PFAM)
scaffold-ZFGK-2039644-Selaginella_kraussiana	Lycophytes	Selaginellaceae	class V chitinase	2,60541E-73	63.17	IPR001223 (PFAM)
scaffold-ZFGK-2040681-Selaginella_kraussiana	Lycophytes	Selaginellaceae	acidic endochitinase	3,32101E-124	76.55	IPR001223 (PFAM)
scaffold-ZFGK-2041140-Selaginella_kraussiana	Lycophytes	Selaginellaceae	class V chitinase	2,11352E-125	63.66	IPR001223 (PFAM)
scaffold-ZQYU-2108391-Polypodium_plectolens	Leptosporangiate Monilophytes	Polypodiaceae	Acidic mammalian chitinase	1,63673E-19	56.31	IPR001223 (PFAM)
scaffold-ZQYU-2112073-Polypodium_plectolens	Leptosporangiate Monilophytes	Polypodiaceae	acidic endochitinase	2,46249E-95	73.84	IPR001223 (PFAM)
scaffold-ZQYU-2112352-Polypodium_plectolens	Leptosporangiate Monilophytes	Polypodiaceae	Chitinase	1,76715E-76	68.5	IPR001223 (PFAM)
scaffold-ZXJO-2006790-Hemionitis_arifolia	Leptosporangiate Monilophytes	Pteridaceae	chitinase-like protein EN03	1,18525E-62	72.39	IPR001223 (PFAM)
scaffold-ZXJO-2006791-Hemionitis_arifolia	Leptosporangiate Monilophytes	Pteridaceae	chitinase-like protein EN03	4,45541E-163	69.01	IPR001223 (PFAM)
scaffold-ZXJO-2015647-Hemionitis_arifolia	Leptosporangiate Monilophytes	Pteridaceae	acidic endochitinase	4,4529E-78	72.95	IPR001223 (PFAM)
scaffold-ZXJO-2070208-Hemionitis_arifolia	Leptosporangiate Monilophytes	Pteridaceae	acidic mammalian chitinase	8,46351E-40	64.78	IPR001223 (PFAM)
scaffold-ZXJO-2070689-Hemionitis_arifolia	Leptosporangiate Monilophytes	Pteridaceae	acidic mammalian chitinase-like	6,31667E-19	57.44	IPR001223 (PFAM)
scaffold-ZXJO-2072539-Hemionitis_arifolia	Leptosporangiate Monilophytes	Pteridaceae	chitotriosidase-1-like isoform X2	7,3608E-67	59.87	IPR001223 (PFAM)
scaffold-ZZOL-2048644-Selaginella_stauntoniana	Lycophytes	Selaginellaceae	acidic endochitinase	3,81683E-57	83.65	IPR001223 (PFAM)
scaffold-ZZOL-2050389-Selaginella_stauntoniana	Lycophytes	Selaginellaceae	class V chitinase	8,68342E-73	72.54	IPR001223 (PFAM)
scaffold-ZZOL-2053046-Selaginella_stauntoniana	Lycophytes	Selaginellaceae	acidic endochitinase	8,78561E-159	80.56	IPR001223 (PFAM)

Figure S1. Class III chitinases alignments of complete sequences identified in green algae



```
                              10        20        30        40        50        60        70        80        90       100
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      -------------------------GDIAIYWGQNGG---------EGTLASTCDTGRYAYVIVSFVTTFGNF----RA-PVVNLAGHCDP---AAGTCT
AFY08286.1_Class_III  ----------------------MKAAGIAVYWGQNGG---------EGSLAATCNTGNYKFVNIAFLSTFGGG----KT-PQLNLAGHCNP---AAGTCK
ABW75910.1_Class_III  MAAKLKCSPLLPLLLLAGMAGISRAGNIAVYWGQNGY---------EGTLADACNSGLYKYVTLSFLSTFGNG----QT-PALNLAGHCNA---SSGGCT
RPQV-2007705          -------------------------GKIVAYWGQNGN---------EGSLLTACQSGNYDILVLSFLVVFGMG----QT-PQLNLAGHCDP---TSNGCV
YOXI-2007181          -----------------------------MYWGQNGN---------EGSLAAACQSGCYDIVILSFLVSFGAG----RA-LQLNFAGHCDP---GSGDCA
JOJQ-2005804          -----------------------------IYWGQNGN---------EGTLLSACQSGRYDIIVLSFLYVWGKG----QN-LQIDLAGHCSA---AGGGCT
YSQT-2036100          ----------------------SSSAKLVVYWGQNGN---------EGTLLSACQSGNYDIIVLSFLCTFGQG----RT-PQLNLAGHCDP---SSGGCA
WDCW-2046043          ------------------------AGKLVAYWGQNGN---------EGTLLSTCQSGRYDILVMSFVYTFGDS----RT-PKVDLAGHCSA---ASNGCT
STKJ-2099840          ---------------------------VVAYWGQNGN---------EGTLLSACQSGRYDIIVMGFLYIYGQG----QA-PRLNLAGHCDP---ASGGCV
HJVM-2001428          ---------------------------VVAYWGQNGN---------EGTLLSACQSGRYDIIVMGFLYIYGQG----QA-PRLNLAGHCDP---ASGGCV
RQFE-2007301          -------------------------GKMVIYWGQNGN---------EGSLLSACESGRYDILVLSFLSTYGSG----QD-PQLNLAGHCDP---SSGGCV
KEYW-2025590          -------------------------GQLIMYWGQNGN---------EGSLLNACEGGLYDIIVLSFLVVFGRG----QT-PQLNLAGHCDP---SSNGCV
RPRU-2039277          ----------------------SKKGQLVAYWGQNA----------DGSLTSACTSGNYDIIVVAFLYVFGMG----QT-AKLDLAGHCDP---STGGCV
WGMD-3000449          -------------------------GKLTVYWGQNGN---------EGSLLSTCQSGRYDIINLSFLNVWGSG----RS-PSINLAGHCDP---GSGGCK
WGMD-3000448          -------------------------GKLTVYWGQNGN---------EGSLLSTCQSGRYDIINLSFLNVWGSG----RS-PSINLAGHCDP---GSGGCK
WGMD-3000450          -------------------------GKLTVYWGQNGN---------EGSLLSTCQSGRYDIINLSFLNVWGSG----RS-PSINLAGHCDP---GSGGCK
ZRMT-2003702          ----------------------SQKGLLVAYWGQNGF---------EGTLAATCQSNRYDVLVLSFLYQFGSG----RP-VLLDLAGHCNP---LANECV
QWFV-2020306          ---------------------------IAVYWGQNGF---------EGTLAQTCQSGRYDIIILAFIAGFGMN----GP-LNLNLAGHCTP---GVDCNT
GGWH-2048506          -----------------------------IYWGQYGT---------EGTLLSACQSGRYDIIVVAFLSTFGNG----QT-PQINLAGHCTP---GSNCDT
ISHC-2004482          ---------------------------VAVYWGQYGT---------EGTLYEACSSGRYDIIMISFLYTFGNG----QT-P---LAGHCVP---GQNCDS
HAOX-2002128          ----------------------TRKGVLVSYWGQGTK---------DGSLEIACKSRRYAIIMVSFLNVFGRG----QA-LGLNLAGHCDP---ALGECA
ISHC-2004481          --------------------------QVAVYWGQYGT---------EGTLYEACSSGRYDIIMISFLYTFGNG----QT-----------P---SRGCDS
GYRP-2000500          ----------------------SSRGQIAVYWGQDGT---------EGRLLSACQSGRYDILMLAFVSVFGNG----QP-LSLNLAGHCTP---GVDCNT
GYRP-2000499          ----------------------SSRGQIAVYWGQDGT---------EGRLLSACQSGRYDILMLAFVSVFGNG----QP-LSLNLAGHCTP---GVDCNT
GYRP-2000501          ----------------------SSRGQIAVYWGQDGT---------EGRLLSACQSGRYDILMLAFVSVFGNG----QP-LSLNLAGHCTP---GVDCNT
HAOX-2024957          -------------------------GVLVSYWGQGSN---------DGTLEATCNSRRYAVIMISFLNVFGNG----NSQLGLNLAGHCDP---PSGGCV
GYRP-2000503          ----------------------SSRGQIAVYWGQDGT---------EGRLLSACQSGRYDILMLAFVSVFGNG----QP-LSLNLAGHCTP---GVDCNT
QWFV-2005300          -------------------------GQLVAYWGQSSL---------DGPLNNTCASGNYDMLIIAFVYKFGKG----QP-VSLNLAGHCDP---SSGGCV
QWFV-2005301          -------------------------GQLVAYWGQSSL---------DGPLNNTCASGNYDMLIIAFVYKFGKG----QP-VSLNLAGHCDP---SSGGCV
RPQV-2005631          --------------------------SMVIYWGQGQNPLN------EGSLLSACQSGLYNVLLIAFLFQFGGT----NP-PLINLSGHCDP-----GQCQ
HJVM-2110782          -----------------------------VYWGQGDS---------DGTLLSVCQSGRYDIVVVSFLYIFGQG----KP-AQLNLAGHCEPTGTAPNCAQ
RPQV-2005632          ---------------------------MVIYWGQGQNPLN------EGPLLSACQSGLYDILLMAFLFQFGGT----LP-PLINLSGHCDP-----GQCQ
DFDS-2002461          ----------------------TNRGQLVAYWGMWSQ---------DIDLLGACKSGNYDIIVLAFLYVFGLG----QT-PQI--STNCDA---PSGGCT
STKJ-2099024          -------------------------------WGQGDS---------DGTLLSVCQSGRYDIVVVSFLYIFGQG----KP-AQLNLAGHCEPTGTAPNCAQ
RPQV-2006998          -----------------------------IYWGQGETPTG------EGTLLSACLTGRYNIIVLAFMSQVGSG----NP-AVINLSNHCNP--TVTGDCQ
FOYQ-2002254          ------------------------------YWGQDSLYAHSIL---ESDLIVYCQSSTWDIIVLSFITGFTSA----GV-PTYDFANHWDM--------A
ZNUM-2002100          ---------------------------LAAYWGQNSVGQTRPTVRPEPTLGEIC--GDYDIVVIAFLRSFGNAAQPIED-PVSTLDGIPYVVDNTIHAPQ

                             110       120       130       140       150       160       170       180       190       200
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      GLSDEIRSCQGKDIKVLMSIGGG-AGD--------------------YS---LVSEADADNFADYLWNNFLGG--QSSSRPLGDAV-LDGIDFDIELG-T
AFY08286.1_Class_III  GISADIKTCQSKGIKVLLSLGGG-TNG--------------------YS---LNSAAEANQPATYLWNNFLGG--RSNSRPLGNAV-LDGIDFDIEAG-P
ABW75910.1_Class_III  GLSSDIQTCQSQGVKVLLSLGGG-SGS--------------------YG---LSSTDDAQSVATYLWDNFLGG--SSSSRPLGPAV-LDGIDFDIETG-N
RPQV-2007705          GLSSDINACQGLGRKVILSLGGA-AGS--------------------YG---FFGDSDAQGVAQYVWNNFLGG--SSGARPLGAAA-LDGVDLDIEGG-S
YOXI-2007181          KLTTDIQICQGLGKKVLLSLGGA-SGS--------------------YG---FSSDGDAQSVATTIWNTFLGG--SGSSRPLGAAQ-LDGVDLDIENN-A
JOJQ-2005804          SLSSEIKSCQALGRQILLSMGGA-VGN--------------------YG---FSSASDAQSVAQTVWNTYLGG--TLSSRPLGATA-LDGVDLDIEAG-S
YSQT-2036100          SLSQDINTCQSLGKQVLVSLGGA-AGA--------------------YG---FSNDADAQAFAQYLWSAYLGGRSSSVTRPLGPAV-LDGIDLDLEGG-A
WDCW-2046043          GLASEIATCQKLGKLVLLSMGGG-VGA--------------------YS---LTGDADAKSTAQSVWDTYLGG--ASASRPFGAAV-LDGVDLDIENA-N
STKJ-2099840          GLSSEISTCQSLGRKVILSMGGA-VGN--------------------YG---FSSDSDAQAVAQYVWQAFLGGSSPAVTRPLGAAA-VDGIDLDIEGG-G
HJVM-2001428          GLSSEISTCQSLGRKVILSMGGA-VGN--------------------YG---FSSDSDAQAVAQYVWQAFLGGSSPAVTRPLGAAA-VDGIDLDIEGG-G
RQFE-2007301          SLSSEISACQALGKLVILSMGGA-TET--------------------YG---FASDSEGADLAQYVWNAYLGGSSSSVTRPLGAAP-LDGVDLDIEKG-G
KEYW-2025590          RFSEEVATCQSKGRKILLSLGGA-SGS--------------------YG---FANAADAQGVATYVWNAYLGGNSAATVRPLGATP-LDGVDLDIEGG-G
RPRU-2039277          SMSSEITTCQSLGRKVILSMGGA-VGS--------------------YG---FSSDADASNVAQYVWQAYLGGSSTAVTRPLGPAP-LDGVDLDIEAG-G
WGMD-3000449          ALASEIQSCQNLGRKVLISMGGA-VGS--------------------YG---FSSDSDGQNFASYMWNNFLGGSLSSNSRPLGATP-LDGIDLDIESG-D
WGMD-3000448          ALASEIQSCQNLGRKVLISMGGA-VGS--------------------YG---FSSDSDGQNFASYMWNNFLGGSLSSNSRPLGATP-LDGIDLDIESG-D
WGMD-3000450          ALASEIQSCQNLGRKVLISMGGA-VGS--------------------YG---FSSDSDGQNFASYMWNNFLGGSLSSNSRPLGATP-LDGIDLDIESG-D
ZRMT-2003702          ALAPQIKTCQDLGVKLVVSMGGA-VGQ--------------------YG---FSSSADARTVAQYVWDAYLGGSSGTVARPLGANS-LDGVDLDLEGG-K
QWFV-2020306          KVGPDVKTCQSLGKKVLLSMGGG-AGA--------------------YS---FVSANDAQSTAQAVWNNFLGG--SSGSRPLGSAV-LDGVDLDIEGG-S
GGWH-2048506          LVGPDISACQSLGRKVILSMGGA-AGT--------------------YF---FSSVADAQSVAQYVWKAFLGG--SGVTRPLGPAI-MDGVDLDLEGG-S
ISHC-2004482          VVGPDIARCQSLGKKLILSMGGG-VGS--------------------YL---FNSVADAQNTSAQIWNMFLGG--SGATRPLGPTP-LDGVDLDIEGG-S
HAOX-2002128          VLSSQIRTCQSLKVKILLSMGGG-IGS--------------------YS---FSNRTDALKVARSVWNAYLGG--SGVTRPLGSAK-LDGIDLDIENN-L
ISHC-2004481          VVGPDIARCQSLGKKLILSMGGG-VGS--------------------YI---FTSAADAQNTAAQVWSMFLGG--SGATRPLGPTP-LDGVDLDIEGG-S
GYRP-2000500          LVGPDISACRSLGKKVLLSMGGG-AGS--------------------YL---FSSSADAVAVAQAVWNSFLWGNSSSSSRPLGNNP-LDGVDLDIEGG-S
GYRP-2000499          LVGPDISACRSLGKKVLLSMGGG-AGS--------------------YL---FSSSADAVAVAQAVWNSFLWGNSSSSSRPLGNNP-LDGVDLDIEGG-S
GYRP-2000501          LVGPDISACRSLGKKVLLSMGGG-AGS--------------------YL---FSSSADAVAVAQAVWNSFLWGNSSSSSRPLGNNP-LDGVDLDIEGG-S
HAOX-2024957          SLSPQIATCQRLGVKLLLSMGGG-IGS--------------------YK---FTSINDAVNVAQLVWNAYLGG--SGVTRPLGLEK-LDGIDLDIENN-F
GYRP-2000503          LVGPDISACRSLGKKVLLSMGGG-AGS--------------------YL---FSSSADAVAVAQAVWNSFLWGNSSSSSRPLGNNP-LDGVDLDIEGG-S
QWFV-2005300          SLSNDIAFCQRLGYKLILSMGGG-VGS--------------------YG---FANDADAASVATHIWEAYLGGHSASVTRPLGPAQ-LNGVDLDIEQG-S
QWFV-2005301          SLSNDIAFCQRLGYKLILSMGGG-VGS--------------------YG---FANDADAASVATHIWEAYLGGHSASVTRPLGPAQ-LNGVDLDIEQG-S
RPQV-2005631          PLGQEIRQCQALGRKVLLSIGGG-TAT--------------------YS---LASAGQGQTVATYIWNAFLGG--SGVTRPFGTAV-LDGVDLDIEGG-A
HJVM-2110782          VVGPDISTCQSLGRKVILSLGGA-EGS--------------------YG---FSSAQDATNVAKYVWNAFLGG--SGVVRPLGNAQ-LDGVDLDIELG-T
RPQV-2005632          GLSTDITQCQALGRKVLLSLGGG-TAT--------------------YT---LNSDSEGQALATYIWNAFLGG--SGVTRPFGPAV-LDGVDLDIEVN-S
DFDS-2002461          FMSSQIYTCQSLGVKVILSLGGA-DGS--------------------YG---FNGDSDGSSVAQYIWNAYLGG--SGTFRPLGDAA-LDGIDLDIEQG-S
STKJ-2099024          VVGPDISTCQSLGRKVILSLGGA-EGS--------------------YG---FSSAQDATNVAKYVWNAFLGG--SGVVRPLGNAQ-LDGVDLDIELG-T
RPQV-2006998          RLSADISQCQGLGIQILLSVGGA-AA---------------------YSRVFLSGDADGEGVATYLWNAYLGGSAPGTMRPLGTAV-LDGIDLDLEMQ-Q
FOYQ-2002254          KLGQDIVACQALGVKLMLSIGGAYMGAA-------------------TS---LSSTAQATTAATQVWNMFGGG--TSTSHVFGTGVAIDGVDLDLENAVN
ZNUM-2002100          QTGPDIAFCQRKGVKVFVSLGGP-VAEAANGSAAVLSESNLTATTCPYG---FRDDDDAKSTAFKVWNSFLGG-SGAVTRPFGATA-LDGIDLDIRC--R

                             210       220       230       240       250       260       270       280       290       300
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      TTF---YDTLARALSSRSTQA---AKVYLTAAPQ------------------------CPHPDSHLDA-ALNTGL--------FDNVWIQFYNNPLAQ-C
AFY08286.1_Class_III  GQH---WDELARALKRFN------QKLILSAAPQ------------------------CPFPDSHLRS-AINTGL--------FNYVWVQFYNNP--S-C
ABW75910.1_Class_III  SAH---WDELANFLSRYSAQG---QKVYLTAAPQ------------------------CPYPDASLGP-ALQTGL--------FDNVWIQFYNNP--P-C
RPQV-2007705          SVG---YAAFVTKMRSLMDSAGG-KKYYITGAPQ------------------------CPFPDAYLGP-SSGTALRDSAGA--FDYVWVQFYNNF----C
YOXI-2007181          ATG---YGAFVVSMRALMNKG-G-KNYLIGAAPQ------------------------CPYPDASLGP-SSGTALGVDPTY--FDYVWVQFYNNF----C
JOJQ-2005804          SAF---YTTFVSTMRATMDAT-S-RKFLITGAPQ------------------------CPYPDAWLGP-GSGTALGDVASK--FDYLFVQFYNNY----C
YSQT-2036100          RIG---YSALVKALRATMNSS-G-RKYLVTAAPQ------------------------CPFPDYFLGP-SSGTPLGDTPSA--FDFIWVQFYNNY----C
WDCW-2046043          PTG---YAAFATSMRQLMKSGGGKRRFYITAAPQ------------------------CPFPDASLGP-SAGTALGQAAAQ--FDFVWVQFYNNY----C
STKJ-2099840          PAG---YAAFLTALRSLMDQS-G-KKYFISGAPQ------------------------CPFPDALMGP-SPGTALGSAPVA--FDYVWVQFYNNY----C
HJVM-2001428          PAG---YAAFLTALRSLMDQS-G-KKYFISGAPQ------------------------CPFPDALMGP-SPGTALGSAPVA--FDYVWVQFYNNY----C
RQFE-2007301          STG---YTAFVKAMRDLMSGS-G-RSYYITSAPQ------------------------CPWPDALAGP-SSGTVLGDVPAM--IDYVWVQFYNNP--G-C
KEYW-2025590          SAG---YAAFVAALSGLFSSDPS-TKYFITAAPQ------------------------CPFPDAFLGP-SAKTALGDSAKY--FDFIWVQFYNNY----C
RPRU-2039277          TTG---YSAFVNSIRSLMDAS-G-RKYYITGAPQ------------------------CPFPDALLGP-SPGTALGASALA--FDYLWVQFYNNY----C
WGMD-3000449          STG---KVSFIRAIRALMNGG-T-KKYLITAAPQ------------------------CPFPDQWHGP-SSGTVLGDTPQS--FDYVWVQFYNNF----C
WGMD-3000448          STG---KVSFIRAIRALMNGG-T-KKYLITAAPQ------------------------CPFPDQWHGP-SSGTVLGDTPQS--FDYVWVQFYNNF----C
WGMD-3000450          STG---KVSFIRAIRALMNGG-T-KKYLITAAPQ------------------------CPFPDQWHGP-SSGTVLGDTPQS--FDYVWVQFYNNF----C
```

```
ZRMT-2003702          SAF---YADFVVALRALMNKT-S-RKFLITAAPQ------------------------CPFPDLWHGP-GAGTAFGDAISN--FDYIWVQFYNNF----C
QWFV-2020306          PNY---YNTFVTTLKNLGNSA-G-KHVYISGAPQ------------------------CPYPDAFLGP-SSNAVLGQVPQY--FDFLWVQFYNNP--P-C
GGWH-2048506          PAY---YSDFVTAIRSLMDAS-G-RKYYITAAPQ------------------------CVYPDYYLGP-KAGSALGVVPQL--FDYIWVQFYNNP--P-C
ISHC-2004482          PAF---YSDFITRLKFLAASA-G-NRVYITGAPQ------------------------CPFPDAYLGP-SPGSALGTVAPS--FDFIFVQFYNNPS---C
HAOX-2002128          ASH---YSDFLVEIRRLMATDSS-KSYYIAAAPQ------------------------CPFPDGSLGP-NFPSAFQDSPTL--FDFIFVQFYNNY----C
ISHC-2004481          RAF---YSDFVTRLKSLASAA-G-RRVYITAAPQ------------------------CPYPDAYVGP-SPASALGTVAAS--FDFIFVQFYNNPA---C
GYRP-2000500          PAY---YSDFATTLRSLAQAA-S-KPVFISGAPQ------------------------CPYPDAYLGP-SAGTALGIVPTA--FDFLFVQFYNNP--P-C
GYRP-2000499          PAY---YSDFATTLRSLAQAA-S-KPVFISGAPQ------------------------CPYPDAYLGP-SAGTALGIVPTA--FDFLFVQFYNNP--P-C
GYRP-2000501          PAY---YSDFATTLRSLAQAA-S-KPVFISGAPQ------------------------CPYPDAYLGP-SAGTALGIVPTA--FDFLFVQFYNNP--P-C
HAOX-2024957          PLY---YSNFLVEMRRLMATDKS-KSYYIGSAPQ------------------------CPFPDGSLGP-NFPSAFQDAPNE--FDFIFVQFYNND----C
GYRP-2000503          PAY---YSDFATTLRSLAQAS-A-KSVYISGAPQCPYPDAYLGPSPGTALGTLSGAPQCPYPDAYLGP-SAGTALGIVPTA--FDFLFVQFYNNP--P-C
QWFV-2005300          SEY---YATFVTTMRGLMDSS-G-SHYVITAAPQ------------------------CPFPDTYLGP-GAHTPLGSVAHA--FDYIWVQFYNNP----C
QWFV-2005301          SEY---YATFVTTMRGLMDSS-G-SHYVITAAPQ------------------------CPFPDTYLGP-GAHTPLGSVAHA--FDYIWVQFYNNP----C
RPQV-2005631          AIG---YTALIRTLYGYASSS-G-RKYYFSAAPQ------------------------CPYPDAMVGP-RPGGPLGDVPTM--IDFVWVQFYNNI--P-C
HJVM-2110782          KFG---YSTFVTSIRALMDGG-A-KKYYISAAPQ------------------------CPKPDALLGP-SSGSAFGEKDSILGFDYIWVQFYNNY----C
RPQV-2005632          LAG---FPALVRTLYGYAAAS-S-RKFYFSAAPQ------------------------CPFPDANVGP-KPGTPLGDVPTM--FDFVWVQFYNNR--Q-C
DFDS-2002461          SSG---YAAFVTTIRGLMDKS-G-RKYYITGAPQ------------------------CPYPDQWLGP-SPGTALGSAGYA--FDYLWVQFYNNY----C
STKJ-2099024          KFG---YSTFVTSIRALMDGG-A-KKYYISAAPQ------------------------CPKPDALLGP-SSGSAFGEKDSILGFDYIWVQFYNNY----C
RPQV-2006998          PAG---YAALVKKLYGYMSAS-G-RKFYISAAPQ------------------------CPYPDLSVGP-GTGTPLGDAAPM--FDYVWVQFYNNR--NYC
FOYQ-2002254          GAY---YSTFVQQFRSLAASA-P-KPYYVSAAPQ------------------------CPYPDASLGP-AAGKPLPATGASL-ADWVWVQFYNNPY---C
ZNUM-2002100          PSGRSSYAAFISQLRKLMDADSS-KPYYITARVD------------------------CTFPDPILGPVNTSQALTVSAYR--FDYLWLQFYGSG--Q-C

                               310       320       330       340       350       360       370       380       390       400
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      QYSS---GNTN-----DILSSWNTWTS----STTAGKIFLGLPAAPEAA-GS----GYIPPDVLTGQILPQIKTSAK--YGGVMLYSKFYDTT---YSTT
AFY08286.1_Class_III  QYTG---GNIN-----NLVNSYKQWTT-----VKASQVFLGVPASTAAA-GS----GFIPVNVMNSQVLPAIKGSTK--YGGVMLYDRFNDLKNG-YSKA
ABW75910.1_Class_III  QYTS---GDAS-----NLQKAWNTWTS----DVKAGSFYLGLPASQDAA-SS----GYVPPGDLTSTVLPGVKGVGN--YGGIMLWSRYYDEQNN-YSNQ
RPQV-2007705          GNNP---GGQS-----ALSQSWNQWSGWLATANRKAKLVLGVPASQSAA-GS----GYIDPST-MRTFLSSIKTSSN--YGGVMMWDVFHDISNGPYSAA
YOXI-2007181          QYSSS--SGVS-----GLIQSWKQWNSWMASASPKTKLVMGLPASTSAA-GS----GFIPQSTVINTVLPAIKSSSN--YGGIMMWSVYYD---------
JOJQ-2005804          QYN----GGTT-----ALLQSWNQWSTWLASAAPSCQLVLGIPASPQAA-GS----GSIPQGTLIATVLPAIKMAAN--YGGVMMWSVYYD---------
YSQT-2036100          QYQG---GSAT-----ALLSSWQQWSSWLTGASPGTRLVLGVPASPSAA-GS----GFMDQPT-LVSIMKQISSSPN--YGGVMMWSVFYD---------
WDCW-2046043          GFTTA--SQTT-----QLLQSWNQWSAWLAGANSNATLVLGLPASSTASTGSFSADYYMAPSA-LSTFLAKIRSSKN--YGGVMLWSVYYDVTN------
STKJ-2099840          QFN----GNAN-----NLIQAWNQWSSWMAGANPSMRLFLGVPGSASAA-GS----GYMNQTT-LRWFLGTIKGYAN--YGGVMIWDVYNDINNS-DAN-
HJVM-2001428          QFN----GNAN-----NLIQAWNQWSSWMAGANPSMRLFLGVPGSASAA-GS----GYMNQTT-LRWFLGTIKGYAN--YGGVMIWDVYNDINNS-DAN-
RQFE-2007301          MYNADT-SSTA-----DLLASWNQWSSWMAEASPSTKLMMGVPASTSAA-SS----GFLPRDT-LISFLPSIKGSSN--YGGVMMWSAYDD---------
KEYW-2025590          GYQR---SGPD-----ALLQSWDTWSKWLATANSNAKLVLGVPASPLA--GA----NFMDQST-LTSLLPVLKAYKN--YGGVMMWSTYYDIRN------
RPRU-2039277          EYN----GNTV-----GLLQSWQQWSSWMAGANPSMKLFLGVPASSTAA-GS----GYMNQTT-IKTFLPSIKGYSN--YGGVMMWDVYED---------
WGMD-3000449          GFSP---SSST-----ALTESWKQWSDWMAQAAPNCKLSLGLPASPSAA-GS----GFIPQNTLISTILPNFKKSSN--FGGVMLW--------------
WGMD-3000448          GFSP---SSST-----ALTESWKQWSDWMAQAAPNCKLSLGLPASPSAA-GS----GFIPQNTLISTILPNFKKSSN--FGGVMLW--------------
WGMD-3000450          GFSP---SSST-----ALTESWKQWSDWMAQAAPNCKLSLGLPASPSAA-GS----GFIPQNTLISTILPNFKKSSN--FGGVMLW--------------
ZRMT-2003702          QYN----GGTA-----GLFQSWNEWSSWLASANPAAKLVLGIPASASAA-GS----GFIPQSNVISSVLPFIKNSNN--YGGIMMWSVFYD---------
QWFV-2020306          M------GSSS-------PQSFKQWASYLSQKSPTTRLLLGLPASRNAAPSG----GYQDPST-LASTLSQIKGLSN--YGGVMLWDVYED---------
GGWH-2048506          D------GANC-------ISSFNQWSSWLAGVAPNTKLFLGLPASTAAAPAG----GYLDQPT-MINTINQLKSIPN--FGGVMLWDVHNDVLN------
ISHC-2004482          ALT----NSPA---------SFRQWSTWLTSANPNGTLFLGVPASPAAA-GS----GYQDLTL-LASTLSQIRDLPN--FGGVMMWDVHNDRENR-AFN-
HAOX-2002128          GYSSWDHSGTT-----QLLASWKIWSQWMIKKMKHAKLFLGIPASPAAA-GS----GYVKLAI-LKTLLPKIKVSSN--YGGVMMWS-------------
ISHC-2004481          ALQ----NSPA---------SFRQWSTWLASANPNGKLFLGVPASPAAA-GS----GYQDLAL-LASTLGQIRGLSN--FGGVMMWDVHND---------
GYRP-2000500          DLT----GTPA---------SFKTWSDWLRTANPGARLFVGLPASVQAAPAG----GYQDPST-LSSTLSKVMGAAN--FGGVMLWDVHNDITNI-ING-
GYRP-2000499          SLT----GTPA---------SFRTWSDWLRTHNPGARLFVGLPASAQAAPAG----GYQDPPS-LSSTLSQVMGAGN--FGGVMLWDVHNDINNL-ING-
GYRP-2000501          SLT----GTPA---------SFRTWSDWLRTHNPGARLFVGLPASAQAAPAG----GYQDPPS-LSSTLSQVMGAGN--FGGVMLWDVHNDINNL-ING-
HAOX-2024957          GYSSA--SGTT-----QLLASWKTWSDWMIQ-KTNAKLFLGIPASPNAA-GS----GYMTPAQ-VNAILPKIKASSN--YGGVMMWS-------------
GYRP-2000503          SLT----GTPA---------SFRTWSDWLRTHNPGARLFVGLPASAQAAPAG----GYQDPPS-LSSTLSQVMGAGN--FGGVMLWDVHNDINNL-ING-
QWFV-2005300          GYS----GSTA-----PILSSWQQWSSSLGALASNVKLFLGLPASPQIA-------NYMHPET-VAALLPSLKNFFN--YGGVMIW--------------
QWFV-2005301          GYS----GSTA-----PILSSWQQWSSSLGALASNVKLFLGLPASPQIA-------NYMHPET-VAALLPSLKNFFN--YGGVMIW--------------
RPQV-2005631          DIR----SGVS-----GVGQAWSTWSSWLARANPKAKLFLGIAASSPAATAG----GYMSVQE-VKSVLAVIKRSPN--YGGVMMWDVAEDLNNQ-ING-
HJVM-2110782          DWK----GGNA-----VFTSVWNDWSSWLAGANPKARLFLGVPASPAAA-GS----GYMTQST-LATVLGQVKKSTN--YGGVMLWDVSNDVKNQ-ING-
RPQV-2005632          DIF----SGVT-----GLVSSWTTWSSWLAQANPAAKLFLGMPASPTA--GW----DYMSAAQ-AASVLAAIKGSSN--YGGVMVWN-------------
DFDS-2002461          GYD----GSTS-----GLLQSWKQWSSSLASINPNMKLFLGVPASSSS--GA----NYMNQPT-IDGFLPSIKGFSN--YGGVMMWDVSSDISN------
STKJ-2099024          DWK----GGNA-----VFTSVWNDWSSWLAGANPKARLFLGVPASPAAA-GS----GYMTQST-LATVLGQVKKSTN--YGGVMLWDVSNDVKNQ-ING-
RPQV-2006998          DFT----AG-A-----GLSKYWTQWSSWLATANPNAKLMLGIPASANPTQFD---SGYMTVAE-VSSVLAAIKGSSN--YGGVMMW--------------
FOYQ-2002254          QP-----GSAQ-----YNIASWAAWAQ-----GAGAKLFIGMPGDPVAA-GS----GYLTATQ-LRTYIDQAAADYPTVFAGVMTW--------------
ZNUM-2002100          AWAN---GGKNFTANDTLIIAFNTWARWAGSVDPPPLLFVNLPAAPGA--GT----GFIPSDA-AASLINFLNTRAN--YGGVMLWDVGWD---------

                      ....|
4TOQ_1|Class_III      IKDQV
AFY08286.1_Class_III  IIGSV
ABW75910.1_Class_III  VKGSV
RPQV-2007705          IRSAV
YOXI-2007181          -----
JOJQ-2005804          -----
YSQT-2036100          -----
WDCW-2046043          -----
STKJ-2099840          -----
HJVM-2001428          -----
RQFE-2007301          -----
KEYW-2025590          -----
RPRU-2039277          -----
WGMD-3000449          -----
WGMD-3000448          -----
WGMD-3000450          -----
ZRMT-2003702          -----
QWFV-2020306          -----
GGWH-2048506          -----
ISHC-2004482          -----
HAOX-2002128          -----
ISHC-2004481          -----
GYRP-2000500          -----
GYRP-2000499          -----
GYRP-2000501          -----
HAOX-2024957          -----
GYRP-2000503          -----
QWFV-2005300          -----
QWFV-2005301          -----
RPQV-2005631          -----
HJVM-2110782          -----
RPQV-2005632          -----
DFDS-2002461          -----
STKJ-2099024          -----
RPQV-2006998          -----
FOYQ-2002254          -----
ZNUM-2002100          -----
```

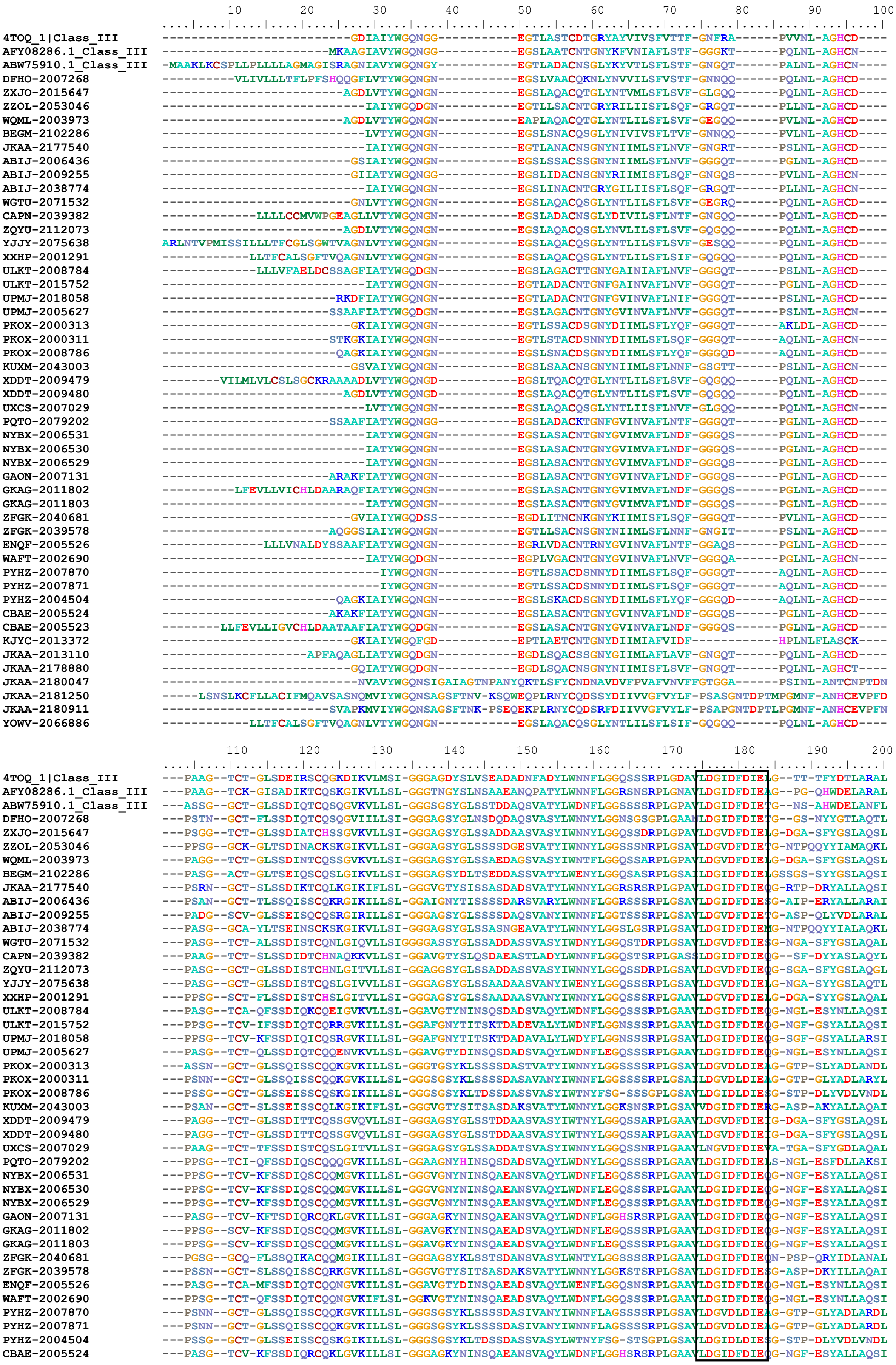
Figure S2. Class III chitinases alignments of complete sequences identified in bryophytes

```
                              10        20        30        40        50        60        70        80        90       100
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      -------------------------GDIAIYWGQN--GG---E-----GTLASTC-DT----GRYAYVIVSFVTTF----GNF-RA-PVVNLAG---HC-
AFY08286.1_Class_III  ----------------------MKAAGIAVYWGQN--GG---E-----GSLAATC-NT----GNYKFVNIAFLSTF----GGG-KT-PQLNLAG---HC-
ABW75910.1_Class_III  MAAKLKCSPLLPLLLLAGMAGISRAGNIAVYWGQN--GY---E-----GTLADAC-NS----GLYKYVTLSFLSTF----GNG-QT-PALNLAG---HC-
LFVP-2004510          ------------VLLLVSAASRASAGSIASYWGQN--GG---E-----GTLADAC-AS----SNYDIIMLSFLYQF----GNG-LA-PALNLAG---HC-
LFVP-2083339          ---------------------------IVVYWGQY--DE---------GTLSSAC-DT----GNYDMIVLSFLNVF----GKG-QK-PQINLSS---HC-
LFVP-2083179          ---------------------------VVTYWGQF--DE---------GSLADAC-DT----GDYGIVILSFLSTF----GRG-NE-PKLDLSG---HC-
LFVP-2083413          ----------------------------AAYWGQN--GN---E-----GTLQEAC-SS----GRYAVVLISFLTHF----GYN-GV-HRLNLAG---HC-
LFVP-2083566          -----------VLLISNQAILTSAAGLVAGYWGQH--GNDSTE-----GSLASLC-TS----GNYGIVIIAYLSRF----GRN-QS-PVLNLGG---HC-
HMHL-2053338          ------------VLLLVSAASRASAGSIASYWGQN--GG---E-----GTLADAC-AS----SNYDIIMLSFLYQF----GNG-LA-PALNLAG---HC-
IHWO-2065363          ---------------------------IVVYWGQY--DE---------GTLSSAC-DT----GNYDMIVLSFLNVF----GKG-QK-PQINLSS---HC-
HMHL-2052653          ----------------------------AAYWGQN--GN---E-----GTLQEAC-SS----GRYAVVLISFLTHF----GYN-GV-HRLNLAG---HC-
IHWO-2065195          ----------------------------AAYWGQN--GN---E-----GTLQEAC-SS----GRYAVVLISFLTHF----GYN-GV-HRLNLAG---HC-
IHWO-2065225          -----------VLLISNQAILTSAAGLVAGYWGQH--GNDSTE-----GSLASLC-TS----GNYGIVIIAYLSRF----GRN-QS-PVLNLGG---HC-
IHWO-2065167          ---------------------------VVTYWGQF--DE---------GSLADAC-DT----GDYGIVILSFLSTF----GRG-NE-PKLDLSG---HC-
IHWO-2005936          ------------VLLLVSAASRASAGSIASYWGQN--GG---E-----GTLADAC-AS----SNYDIIMLSFLYQF----GNG-LA-PALNLAG---HC-
WNGH-2088662          ------------LLVCMSIQTVLSANQMVIYWGQNSAGSFTTVASQQEKPLRNYC-QD----SSYDIIVVGFVYLFPSADGNTDPTLPGMNFAN---HCE
KRUQ-2111018          ------------LLVCISIQTVLSSNQMVIYWGQNSAGSITTVPSQQEKPLRNYC-QD----SSYDIIVVGFVYLFPSADGNTDPTLPGMNFAN---HCE
CHJJ-2135670          ---------------------------MVIYWGQNSAGSFTNIKSQWEQPLRNYC-ED----SSYDIIVVGFVYLFPSASGNTDPTLPGMNFAN---HCE
KRUQ-2111802          ---------------------------VVVYWGQSSSGN---E-----GDLREYC-AS----PSIDIVVVSFVYEFPTAPGS--QY-PGLNFAG---HCY
YBQN-2014678          --------------------------SLAVYWGQG--SG---Q-----KSLGSFC-AD----STIDIIPIGFMNVFPDQNGTY-GY-PGTNFGN---ACG
RTMU-2182337          -------------ISLANAFDASANSNMAVYWGQG--LN---Q-----PRLSHFC-AQ----SAIDIIPIGFVNTFPSG-GST-NY-PGSNFGS---ACG
RDOO-2001513          ---------------------------LNLYWGQS--GM---------DRLSDYC-GR----EHVKILTLAFVNNS----PDSNGV-PGTNFGG---HCG
LGOW-2016622          ---------------LTGVAANAPVGKFNGYWGQT--GY---------DRLKDRC-DE----GA-DFISLSFVNVSPEN-GNGTGY-PGTNFAG---HCA
HPXA-2018253          ----------------------SSKTNLALYWGAD--GD---------LSLSAYC-SNRAKYGNFNVALVGFLTGI--------AT-PTLNIAG------
LGOW-2003062          -----------LVPFALAEASVSASSNIAVYWGQN--SYDQSSGPLEQQRLSYYC-AS----SEIDIIPIAFLVTI--------TN-PELNFANAGNNC-
JPYU-2038245          ------------VLVLASVSSQASAGLIASYWGQN--GG---E-----GTLADAC-AS----SNYDIIMLAFLYRF----GNG-LT-PALNLAG---HC-
WJLO-2035839          ----------LMLLFVAAAAGRVSAGSLASYWGQN--GN---E-----GTLADAC-AS----GNYDIIMVSFLSIF----GKG-QT-PALNLAG---HC-
TXVB-2013493          -------------------------GTIATYWGQN--GN---E-----GTLASAC-SS----TNYGIIMLSFLSEF----GR--NP-PVLNLAG---HC-
TFYI-2002959          ------------VLLLVSAASRVSAGSIAVYWGQN--GG---E-----GTLADAC-SS----GNYDIVMLSFLSSF----GNG-VT-PTLNLAG---HC-
ILBQ-2043168          ----------ILVLLLVAGANGVSAGSLASYWGQN--GN---E-----GTLANAC-AS----GNYDIIMLAFLSSF----GSGITT-PVLNLAG---HC-
TFYI-2073368          ----------MVVLVMAPSASRVSAGSIAAYWGQN--GN---E-----GSLADACLRP----NNYGIVLIAFLSQF----GKG-QS-PVLNLAG---HC-
WJLO-2007388          ---------------------------IAAYWGQ---GSRQVE-----GSLDKLC-NS----GNYGIVLLAFLSQF----GRG-KA-PTLNLAG---HC-
JPYU-2007981          ----------LIILVVLSAVNVSSAGQIATYWGQS--GN---DDAGSEGSLADLC-NS----NNYGIVMIAFLSAF----GRG-QT-PVLNLAG---HC-
TXVB-2072783          ---------------------------IARYWGQK--TN---E-----GRLSSVC-KS----GNYNIVLISFLSRF----GLG-KS-PELNLAG---HC-
LGOW-2103657          ----------LVVIQLAISSRSVQAGFITVYWGQG--DSS--E-----GTLASAC-NS----GHYKIIQVGFLSVF----GNN-QT-PQLNLAG---HC-
FITN-2007300          -----------------------------IYWGQN--GN---E-----GTLSSAC-DS----NNYDIIMLSFLSQF----GGG-QT-AQLNLAG---HC-
FITN-2007298          -----------------------QAGKIAIYWGQN--GN---E-----GSLSKAC-DS----GNYDIIMLSFLYQF----GGG-QD-AQLNLAG---HC-
FITN-2007299          -----------------------------IYWGQN--GN---E-----GTLSSAC-DS----NNYDIIMLSFLSQF----GGG-QT-AQLNLAG---HC-
TFYI-2073779          ---------------------------IVVYWGQY--DE---------GSLSSAC-DT----GNYDIIVLSFLNVF----GND-QK-PQIDLSK---HC-
PIUF-2012805          -------------VLLASSASVQGKGSIAAYWGQG--ADSSVE-----GTLAQAC-QS----GNYGIVIVAFIAVF----GNG-QT-PELNLAG---HC-
JHFI-2013440          ---------LASLVLLAPSGSVQGKGSIAAYWGQG--ADSSVE-----GTLAQAC-QS----GNYGIVIVAFIAVF----GNG-QT-PQLNLAG---HC-
PIUF-2090613          -------------------------GFIVTYWGQD--GT---E-----GTLSEAC-LT----GNYHIIILAFISSF----GGG-IT-PVLNLAG---HC-
JHFI-2112286          -------------------------GFIVTYWGQD--GT---E-----GTLTEAC-LT----GNYHIIILAFISSF----GGG-IT-PVLNLAG---HC-
JHFI-2012415          RSSLMLAVLVLILSFSINSAEARKKGGVVVYWGVY--DG---E-----GGLAEAC-NT----GNYEIVIIAFIDSF----GHG-KP-PTLDLA----HC-
JHFI-2012990          -------------------------GTVVIYWGVY--DG---E-----GGLAEAC-NT----GNYDIVIIAFIDSF----GHG-QP-PTLDLA----HC-

                             110       120       130       140       150       160       170       180       190       200
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      ------DPAA-----GTCTG-LSDEIRSCQGKDIK-VLMSIGG-----------GAGDYSLVSEADADNFADYLWNNF-LGG----QSSS--RPLG----
AFY08286.1_Class_III  ------NPAA-----GTCKG-ISADIKTCQSKGIK-VLLSLGG-----------GTNGYSLNSAAEANQPATYLWNNF-LGG----RSNS--RPLG----
ABW75910.1_Class_III  ------NASS-----GGCTG-LSSDIQTCQSQGVK-VLLSLGG-----------GSGSYGLSSTDDAQSVATYLWDNF-LGG----SSSS--RPLG----
LFVP-2004510          ------DPAS-----NGCQG-LVADINSCQSSGVK-VILSLGG-----------GAGDYGFSSADEARELAQYLWDNY-LGG----SSSS--SPLG----
LFVP-2083339          ------DTSPA----GNCTG-LTADITNCQSKGIS-ILLSLGG-----------AVGDYGIASPSDAVDVAAYLWNSY-LGGSEEFSNNSDGRPLG----
LFVP-2083179          ------DAGS-----GECKG-LDSQIEHCQNNGVK-ILLSLGG-----------ADGNYGLDSKDDANNVAQYLWNTY-LGG----QDGS--GPLG----
LFVP-2083413          -----PNIDS-----SGCKM-LESQVAECQKLKVK-VLLSLGGS----------GDAYYGFNTQAEAENTANYLWNNF-FGG----SSSS--SPFG----
LFVP-2083566          ------DVLS-----GSCQR-YASEVATCQRRGIR-VLLSLGG-----------GAGGYGFDTADEARLLAQDIWDNY-MGG----SSSN--SPLG----
HMHL-2053338          ------DPAS-----NGCQG-LVADINSCQSSGVK-VILSLGG-----------GAGDYGFSSADEARELAQYLWDNY-LGG----SSSS--SPLG----
IHWO-2065363          ------DTSPA----GNCTG-LTADITNCQSKGIS-ILLSLGG-----------AVGDYGIASPSDAVDVAAYLWNSY-LGGSEEFSNNSDGRPLG----
HMHL-2052653          -----PNIDS-----SGCKM-LESQVAECQKLKVK-VLLSLGGS----------GDAYYGFNTQAEAENTANYLWNNF-FGG----SSSS--SPFG----
IHWO-2065195          -----PNIDS-----SGCKM-LESQVAECQKLKVK-VLLSLGGS----------GDAYYGFNTQAEAENTANYLWNNF-FGG----SSSS--SPFG----
IHWO-2065225          ------DVLS-----GSCQR-YASEVATCQRRGIR-VLLSLGG-----------GAGGYGFDTADEARLLAQDIWDNY-MGG----SSSN--SPLG----
IHWO-2065167          ------DAGS-----GECKG-LDSQIEHCQNNGVK-ILLSLGG-----------ADGNYGLDSKDDANNVAQYLWNTY-LGG----QDGS--GPLG----
IHWO-2005936          ------DPAS-----NGCQG-LVADINSCQSSGVK-VILSLGG-----------GAGDYGFSSADEARELAQYLWDNY-LGG----SSSS--SPLG----
WNGH-2088662          VPFDNSNPYV-----LNCTQNVAPDIQYCQSRGKK-IFLSFGG-----------GVGRYGFNSDAQGTAFATTVWNMF-LGG----NGAI--RPFG----
KRUQ-2111018          VPFDDRNPFV-----LNCTQNIAPDIQFCQSRGKK-IFLSFGG-----------GVGRYGFSSDAQATTFATTVWNMF-LGG----NGAI--RPFG----
CHJJ-2135670          VPFDDANPFV-----LNCTGNVAPDITYCQQRGKK-IILSFGG-----------GVGNYGFTSDSQAITFANTVWNMF-LGG----SGAI--RPFG----
KRUQ-2111802          KTFSNANPLL-----LDCSDSIGPDITYCQSLGKQ-VLLSFGG-----------ADGNYGFSSDTQAQTFATTVWNMF-LGG----TSSV--RPFG----
YBQN-2014678          GYYVDSKNETTQLL-GSCPQ-IGADITYCQSQGKK-ILLSLGG-----------GSTGVFVASDDSAVAFGDFLHGAF---GPPS-SSWTGPRPFG----
RTMU-2182337          GAMATNPDGTTSGFLGSCDE-IGQDIKTCQAAGKK-ILISLGGA----------VPETYKIDSESTANQLADTLWNAF---GPKQ-ASYTGARPFG----
RDOO-2001513          ADVTYVDNGQTTGLLMSCDT-VKEDIKYCQHKGVK-IMLSIGGEFNA-------TTSNYSLNDDASGVQFADFLVKAF---GPYD-PTSSIPRPFDLDDG
LGOW-2016622          GEYYTKDGRKSRLL-SQCTY-IAEDIPYCQSKGIK-MLLSIGGVWSDTEDANGFGYSDYTISSQQNGVAFAQFLSGAF---GKPS-ADWEGVRPFG----
HPXA-2018253          ------NQLK------------AADLQTCHENGVN-VLVSMGG--------------PTGITDTKVNASMAQGIWDYI---GPDKADNMSPTRLWG----
LGOW-2003062          --TAINGSSL-----FSCPQ-LEADIAECQTTYKKTILLSIGGA----------TYSEGGFTSSDAAVSAANNIWSIFGPEG----SNSSVPRPFG----
JPYU-2038245          ------DPAS-----GGCQG-LVSEINSCQSRGVK-VILSLGG-----------GVGDYGFSSADEALELAQYLWDNY-LGG----SSSS--SPLG----
WJLO-2035839          ------DPPS-----GGCQG-LVSEVQSCQSQGVK-VILSLGG-----------GSGSYGLSSAQEAEGLAQYLWDNY-LGG----SSDS--SPLG----
TXVB-2013493          ------DPAS-----GGCQG-QVTDVNSCQGQGVK-IILSLGG-----------ASGAYGFDTADQARDLAQYLWDNF-FGG----STAT--SPLG----
TFYI-2002959          ------DPGS-----GGCQG-LVTEVNSCQNQNVK-VILSLGG-----------GAGSYGFSSADDARNLAQYLWDNF-LGG----SSSS--SPLG----
ILBQ-2043168          ------DPGS-----GGCQS-LVSEVNSCQNAGKK-VILSLGG-----------SAGTYGFSSSQEAEDLAQYLWDNF-LGG----SSDS--SPLG----
TFYI-2073368          ------DPGS-----GGCRG-LVSDIQTCQNNGRK-IILSLGG-----------GAGSYGFNNANEARALAQYLWDNY-FGG----SSGS--SPLG----
WJLO-2007388          ------DPPS-----GSCRV-YASQIQSCQS-----KLLSLGG-----------GGGDYGFNSAAEARSLAQHIWDNY-MGG----NSGN--RPLG----
JPYU-2007981          --DPHCDPAS-----GTCAT-YASDIAACQSLGIK-VLLSLGG-----------GAGGYGFDSADEARQLAQYIWDNY-LGG----SNAN--SPLG----
TXVB-2072783          ------NPSA-----GDCTE-FENDVKTCQDSGVK-VLLSLGG-----------GVEGYGFKSADEAKRLANHIWHNY-LGG----NSGT--RPLG----
LGOW-2103657          ------DPSS-----NGCQG-LTSDINTCHNKGIK-ILLSIGG-----------GAGSYKLVSTQDAKNVAAYLWDGY-LGG----KSSSVTRPLG----
FITN-2007300          ------DPSN-----NGCTG-LSSQISSCQQKGVK-ILLSLGG-----------GSGSYKLSSSSDASIVANYIWNNF-LAG----SSSS--RPLG----
FITN-2007298          ------DPSS-----GGCTG-LSSEISSCQSKGIK-ILLSLGG-----------GSGSYKLTDSSDASSVASYLWTNY-FSG----STS---GPLG----
FITN-2007299          ------DPSN-----NGCTG-LSSQISSCQQKGVK-ILLSLGG-----------GSGSYKLSSSSDASIVANYIWNNF-LAG----SSSS--RPLG----
TFYI-2073779          ------DTTLA----GNCTG-LTVDITHCQSKGIS-ILLSLGG-----------AIGNYGIASPSDAAAVAAYLWDNF-LGGTPNSSNLSDARPLG----
PIUF-2012805          ------DPSS-----GGCVG-YSSDIDTCHANGVK-VLLSLGG-----------GAGAYELVSAQDAEDVANYLWNNY-LGG----HSDS--RPLG----
JHFI-2013440          ------DPSS-----GGCTG-ISSDILTCHNNGVK-VLLSLGG-----------GAGSYELVSVQDAENVANYLWNNY-LGG----QSNS--RPLG----
PIUF-2090613          ------SPYD-----NGCVG-LTTEITTCQANGVK-VGIALGG-----------GGGGYGLTSPSDAEQLGTYLWNFF-LGG----SDNTSSRPLG----
JHFI-2112286          ------SPYD-----NGCAG-LTTDITTCQAHGVK-VGIALGG-----------GGGGYGLTSPSDAEQLGTYLWNFY-LGG----SDNTSSRPLG----
JHFI-2012415          ------DASS-----GGCQG-LSKDISQCQNQGVK-VMVSLGG-----------AAGDYGLNSDYDAKKVANDLWNNF-LGG----QSGN--RPLG----
JHFI-2012990          ------DASN-----GGCQG-LSKDISHCQSQAVK-VMISLGG-----------ADGNYGLDSDDDAKNVANDLWNDF-LGG----QSDN--RPLG----
                             210       220       230       240       250       260       270       280       290       300
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      -DAVLDGIDFDIE-LGTT-------T-FYDTLARALSSRST----Q--A---AKVYLTAAPQCPHPDSHL-----DAALNTGLFDNVWIQFYNNPLAQCQ
AFY08286.1_Class_III  -NAVLDGIDFDIE-AGPG-------Q-HWDELARALKRFN------------QKLILSAAPQCPFPDSHL-----RSAINTGLFNYVWVQFYNNP--SCQ
```

```
ABW75910.1_Class_III  -PAVLDGIDFDIE-TGNS-------A-HWDELANFLSRYSA----Q--G---QKVYLTAAPQCPYPDASLGP-----ALQTGLFDNVWIQFYNNP--PCQ
LFVP-2004510          -AAKLDGIDLDIE-GGGS-------Q-YYPDLISTVKSIAQ----S--NGY-KKLYVTAAPQCPFPDALLGPAD-GKALSTGLVDLVFVQFYNN---YCG
LFVP-2083339          -GAVLDGIDFDIE-NGSS-------A-HYADLAQSLRSLMR----T--GD--KQYFLTAAPQCPFPDEWTNP-----ALQSGLFDYVWVQFYNNP--GCD
LFVP-2083179          -QSNLDGIDLDIE-RGSP-------K-HYNTLVQSLHDLMQ----D--GD--KKFYLGAAPQCPFPDEWAGPGG-DSALQTGLFDYVWVQFYNND--QCQ
LFVP-2083413          -AVKLDGIDLDIE-SGGS-------S-YYPDLIRKLLAVSE----E--KNYGKKVYITAAPQCPKPDTYLGPNGEKTALGTGLIDYVFVQFYNN---YCD
LFVP-2083566          -EVKLDGIDLRIY-SGGF-------R-FYADLGGRLREIAM----D--NGY-KYVYFGVSPQCPCPDPSLGPDSAGKFMETELVDLVFVQFYNNP--GCD
HMHL-2053338          -AAKLDGIDLDIE-GGGS-------Q-YYPDLISTVKSIAQ----S--NGY-KKLYVTAAPQCPFPDALLGPAD-GKALSTGLVDLVFVQFYNN---YCG
IHWO-2065363          -GAVLDGIDFDIE-NGSS-------A-HYADLAQSLRSLMR----T--GD--KQYFLTAAPQCPFPDEWTNP-----ALQSGLFDYVWVQFYNNP--GCD
HMHL-2052653          -AVKLDGIDLDIE-SGGS-------S-YYPDLIRKLLAVSE----E--KNYGKKVYITAAPQCPKPDTYLGPNGEKTALGTGLIDYVFVQFYNN---YCD
IHWO-2065195          -AVKLDGIDLDIE-SGGS-------S-YYPDLIRKLLAVSE----E--KNYGKKVYITAAPQCPKPDTYLGPNGEKTALGTGLIDYVFVQFYNN---YCD
IHWO-2065225          -EVKLDGIDLRIY-SGGF-------R-FYADLGGRLREIAM----D--NGY-KYVYFGVSPQCPCPDPSLGPDSAGKFMETELVDLVFVQFYNNP--GCD
IHWO-2065167          -QSNLDGIDLDIE-RGSP-------K-HYNTLVQSLHDLMQ----D--GD--KKFYLGAAPQCPFPDEWAGPGG-DSALQTGLFDYVWVQFYNND--QCQ
IHWO-2005936          -AAKLDGIDLDIE-GGGS-------Q-YYPDLISTVKSIAQ----S--NGY-KKLYVTAAPQCPFPDALLGPAD-GKALSTGLVDLVFVQFYNN---YCG
WNGH-2088662          -NAIFDGIDLDIE-GGSS-------T-GYVPFVATLRSYFA----A--QTS-RSYYISAAPQCFYPDR-IGPGA-GTPLTEAWFDYIWIQFYNN---HCG
KRUQ-2111018          -NAVFDGVDLDIE-GGAQ-------T-GYIQFIATLRSYFA----A--QSS-RAYYISSAPQCFYPDR-LGPGA-GTPLTEAWHDYIWIQFYNN---HCG
CHJJ-2135670          -NAIFDGVDLDIE-GGAQ-------T-GYIAFINTLRSKFA----A--SG--RQYLITAAPQCFYPDR-LGPGS-GTPLTEAWFDYIWIQFYNN---HCG
KRUQ-2111802          -NAILDGIDLDIE-HGGT-------A-GYASFITTLRTYFT----A--QNA-RTYYISGAPQCVYPDAMMGPGA-GTALGNAWFDYVWVQFYNN---YCG
YBQN-2014678          -DACVDGFDFDIE-ATASAIPVPLGA-HYATMVNHLRTL-------------GTMMITASPQCIIPDANM-----GDALTNAYFDAIFVQFYNTP--QCS
RTMU-2182337          -DAVIDGFDLDIE-HNGT-------N-GWATLINRLRTHYK----T--DSS-KEYYISGAPQCVLPDAQL-----SDAIQNSWFDFIWVQFYNTP--GCS
RDOO-2001513          THNSVDGFDVAIQTMFNS-------QTPWIKFIETLRAGI------------PNVLISGAPACTTNPKYFFM---SDIIKQAKFDWINVLFYNDP--SCD
LGOW-2016622          -DSEVDGFDIDVEKDLEN-------K-PWISLIDTLRSLN------------PSYFISGAPQCPLQNG-QGPYYMNEMIQEAQFDALFIQFYNNP--GCD
HPXA-2018253          -TESFDGIDIDTKTTYLG-------K-GLIEFASALRLLMD----DATEKTGKTYYMTAAPQCPKRDADI-----TDLISKVSLDWINLQFFDN---YCN
LGOW-2003062          -ASVVDGFDFDFE--STV-------Q-NMPAFANQLRTLMD----A--DTS-KTYYLSAAPQCPYPDA----------------DFVWVQFYNN---FCG
JPYU-2038245          -AAKLDGIDLDIE-QGGS-------Q-YYPDLIMAVKTMAQ----N--NAY-KRLYVTAAPQCPFPDASLGPGG-GTALSTGLVDLVFVQFYNN---YCG
WJLO-2035839          -AAQLDGIDLDIE-QGGS-------Q-YYPSLISSLRSIAQ----S--NNY-KRLYITAAPQCPFPDASLGPGD-GTALSTGMVDFVFVQFYNN---YCG
TXVB-2013493          -AAKLDGIDLDIE-SGGS-------Q-FYPDLIGRLKEIAN----T--NAY-KELLITAAPQCPIPDASLGPGD-GTALGTGLVNYVFVQFYNN---YCG
TFYI-2002959          -EAKLDGIDLDIE-GGGS-------Q-YYPDLIGRVKEIAQ----N--NNY-KQLYVTAAPQCPYPDAYLGPED-GKALSTGLVDYVFVQFYNN---YCG
ILBQ-2043168          -AAQLDGIDLDIE-MGGS-------Q-YYPNLISSLRSLAE----S--NNY-KKLYITAAPQCPYPDAKLGPGD-GTALGTGSVDYVFIQFYNN---YCG
TFYI-2073368          -AAKLDGIDLDLE-GGGY-------L-YYPDLMGRLKEIAQ----N--NGY-KRLYFTAAPQCPFPDYYLGPAS-GKSLSSGLVDYVFVQFYNN---YCQ
WJLO-2007388          -AAQLDGIDLDIE-HGGS-------A-FYADLAGRLREIAQ----N--NGY-KRVYFSAAPQCPFPDASLGPAP-GKFLDAKLADFVFVQFYNNP--GCD
JPYU-2007981          -AAQLDGIDLDIE-SGGP-------S-YYPDLGGRLKEIAQ----N--NNY-KQLYMSAAPQCPYPDASLGPAT-GTFLGASLADFVFVQFYNNP--DCD
TXVB-2072783          -KVKLDGIDLDIE-VGRG-------D-YYVDLGRRLRDIVK----E--KGYSKRLFLSAAPQCPFPDKRLGLP--GQFLRAGLADYVFVQFYNNP--DCD
LGOW-2103657          -SAVLDGIDFDIE-QGSD-------S-YYATLATQLKNYAS----S--SG--TKVYLSAAPQCPYPDAYDGPGS-GTMLQTGDFDYVYIQFYNN---YCE
FITN-2007300          -SAVLDGVDLDIE-AGGT-------PGLYADLARDLKSLAG----S------RKIYISAAPQCPYPDANLG----SALLTSGLFDYVYVQFYNNP--DCQ
FITN-2007298          -SAVLDGIDLDIE-SGST-------PDLYVDLVNDLKSLAG----N------TKLYISAAPQCPYPDSNLG----SALLTSGLFDYVYVQFYNNQ--DCQ
FITN-2007299          -SAVLDGVDLDIE-AGGT-------PGLYADLARDLKSLAG----S------RKIYISAAPQCPYPDAN----------------------YNGA-----
TFYI-2073779          -DAVLDGIDFDIE-NGSS-------A-HYADLAVSLRSLMR----T--GD--KQYYLSAAPQCPFPDQWANP-----ALQSGLFDYVWVQFYNNN--PCD
PIUF-2012805          -PAILDGIDLDIE-GGSP-------L-YYSNLAQTLKSLAK----N--AG--LTLYLSAAPQCPFPDAYDGPGA-GTALATGLFDYVWVQFYNN---YCQ
JHFI-2013440          -PAILDGIDLDIE-GGSP-------L-YYSNLAQKLKSLAK----T--AG--LTLYLSAAPQCPFPDAYDGPGA-GTALATGLFDYVWVQFYNN---YCQ
PIUF-2090613          -SAILDGVDFDIE-AGSS-------Q-YYANLAETLKKFYASAKLDCEEPSAKDYFISAAPQCPFPDAWDGPGN-GTALATGLFDYVWVQFYNNP--GCQ
JHFI-2112286          -SAILDGIDFDIE-GGSS-------Q-YYANLAETLKKFYAIAKLDCEEPSAKDYFISAAPQCPFPDAWDGPGN-GTALATGLFDYVWVQFYNNP--DCQ
JHFI-2012415          -NAVLDGVDFDIE-KGSS-------H-YYATLATTLYDLFQ----Q--TD--KKFYLGAAPQCPFPDEWDGPGG-HTALATGLFDYVWIQFYNND--QCQ
JHFI-2012990          -EAILDGVDFDLE-SGSS-------N-HYATLATTLHNLFQ----Q--GD--KQFFLGAAPQCPFPDQWDGPGG-HTALATGLFDYVWVQFYNND--ECQ

                              310       320       330       340       350       360       370       380       390       400
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      YS------SGN-----TNDILSS--WNTWTSSTTA----GKIFLGLP--AAPEAA------GSG-YIPPDVL--TGQILPQIKTSAK-YGGVMLYSKFYD
AFY08286.1_Class_III  YT------GGN-----INNLVNS--YKQWTTVKA-----SQVFLGVP--ASTAAA------GSG-FIPVNVM--NSQVLPAIKGSTK-YGGVMLYDRFND
ABW75910.1_Class_III  YT------SGD-----ASNLQKA--WNTWTSDVKA----GSFYLGLP--ASQDAA------SSG-YVPPGDL--TSTVLPGVKGVGN-YGGIMLWSRYYD
LFVP-2004510          YG------GG------LDQLVSV--WQQWTAALPS----AQVFLGLP--ASPDAA------GSG-YVDPSTV--TGEIIPAISTASN-YGGIMLWSVYYD
LFVP-2083339          YNP----VSGD-----VSGIIKA--WDIWTSGLPPG---SSVLLGLP--ASPDAATDEWSLEHG-YLPVNIL--TSLVLPRIVVHPR-FGGVMLWSYFYD
LFVP-2083179          YHR----NGNH-----LDDLFSS--WDQWTSSVPN----AKILMGLP--ASPDAATDNYDLDSG-WLPSDVM--VSDVLPHIRGYNN-FGGVMLWSYYYD
LFVP-2083413          YR------D-------FPTMLQV--WKQWTGDSALSN--AKIFLGLP--AAVGAA------GSG-YVPPEDL--LNKVLPQIQNDPL-YGGVMLWSVYYD
LFVP-2083566          IR------GG------INDLVSS--WRKWITTITASK--PAIFISLP--ASPSA-------GEG-FMDPNTL--KDKVLPSIKRIGN-YSGVALWAPSTD
HMHL-2053338          YG------GG------LDQLVSV--WQQWTAALPS----AQVFLGLP--ASPDAA------GSG-YVDPSTV--TGEIIPAISTASN-YGGIMLWSVYYD
IHWO-2065363          YNP----VSGD-----VSGIIKA--WDIWTSGLPPG---SSVLLGLP--ASPDAATDEWSLEHG-YLPVNIL--TSLVLPRIVVHPR-FGGVMLWSYFYD
HMHL-2052653          YR------D-------FPTMLQV--WKQWTGDSALSN--AKIFLGLP--AAVGAA------GSG-YVPPEDL--LNKVLPQIQNDPL-YGGVMLWSVYYD
IHWO-2065195          YR------D-------FPTMLQV--WKQWTGDSALSN--AKIFLGLP--AAVGAA------GSG-YVPPEDL--LNKVLPQIQNDPL-YGGVMLWSVYYD
IHWO-2065225          IR------GG------INDLVSS--WRKWITTITASK--PAIFISLP--ASPSA-------GEG-FMDPNTL--KDKVLPSIKRIGN-YSGVALWAPSTD
IHWO-2065167          YHR----NGNH-----LDDLFSS--WDQWTSSVPN----AKILMGLP--ASPDAATDNYDLDSG-WLPSDVM--VSDVLPHIRGYNN-FGGVMLWSYYYD
IHWO-2005936          YG------GG------LDQLVSV--WQQWTAALPS----AQVFLGLP--ASPDAA------GSG-YVDPSTV--TGEIIPAISTASN-YGGIMLWSVYYD
WNGH-2088662          LD------EYP-----TNFNFAK--WVAGMANSANPN--VKLFIGAP--AAPNAA------ASG-FVSLATL--QTIATAISQAYPTKYGGLMLWDASTS
KRUQ-2111018          LD------EP------TNFNFNK--WAAGMANAANPN--VKLFIGAP--ASASAA------ASG-FVSLSAL--QAIATSVSQAYPAQYGGVMVWDASTS
CHJJ-2135670          LD------EFP-----NNFNFAT--WAAAMTKAANPN--VKLFIGAP--ASSLAA------ASG-FVPLATL--ETIATAIQNTFPNLYGGLMLWDASTS
KRUQ-2111802          VN------YYP-----SEFNFNV--WAAWAKTSLNPN--VKVFIGAP--GSKNAA------GQG-YVTSATLEIISNAVSA--TYPTYYGGIMLW-----
YBQN-2014678          ARAFFDNTYGAYGGPATNISYDS--WISWMQSNSKNPN-TLVYLGLP--AAPGAT-----YGEM-YLNTTEA--QEILTYFQCKYPAQFGGAMLWEATF-
RTMU-2182337          ARDFISNTGH------MN--FDA--WVAAVQQSNGNKN-AKIFIGLPGDATPATA------NAGFYLNPTEV--RSLVSSFQSKFPNQFGGLMLW-----
RDOO-2001513          FV------NHN-----SGFNYDQ--WESILEDSNRSHK-AKIYIGIP--GAPTAA------ASG-YLTPDEV---QELVCTYGNHSS-FGGINVWDAY--
LGOW-2016622          LI------VGNSQFEEEGFNYDQ--WEKFLAQTKKSQG-AKIYIGLP--A--DSA------DNA-FVSPAPL---KNVVCKYSKRAS-FGGISLW-----
HPXA-2018253          AN---------------TELYNYANWDEWIKDDAKNSD-MKVLVGLP--ASSTAA------NSG-YIASSYQ--LTSVFRTTEAFSS-FAGAMLWDTLH-
LGOW-2003062          LQNFVVGASSQ-----TNFNFNT--WDTWAKN-------TKVFLGIP--ANQGAA------GSG-YESGSTL---SDIITYTKTFSS-FGGVMMW-----
JPYU-2038245          YG------GG------LDQLVTV--WKQWAGALPG----SQIFLGLP--ASPDAA------GSG-YVDPSTV--TDQIIPAISTVPN-YGGVMLFSVFYD
WJLO-2035839          YQ------GG------LDQLVSV--WHQWTDALPS----AQVLLGLP--ASPDAA------GSG-YVEPNTL--LNDILPQISTASN-YGGVMLYSVYYD
TXVB-2013493          FQ------GG------LDQMVDV--WTQWTDALPS----AQIFLGLP--ASPDAA------GSG-YIDPTTL--TGTVLPPIKAAGN-YGGVMLWNVFND
TFYI-2002959          YQ------GG------LDQMVGV--WEQWTSGVPS----AKIFLGLP--ASQDAA------GSG-YVDPSVV--TDTILPAINGASN-YGGVMLWSVFYD
ILBQ-2043168          YP------GG------LDQMVSV--WNQWSTALPS----AQVFLGMP--ASADAA------GSG-YVEASIL--TSEILPQISTASN-YGGVMLWNVYYD
TFYI-2073368          YS------GG------LSQMLSV--WQQWTNALPN----CKIFLGLP--ASPNAA------GSG-YVSPSTL--INNILPSIKAAGN-YGGVMLWNVYED
WJLO-2007388          IR------GG------VNAVVNA--WRQWTSGLPDPK--PKVFIGLL--ASQSA-------GDG-YMSPNTL--KNNVLPQIKKIGN-YGGVMLW-----
JPYU-2007981          FR------AG------VSGMVSS--WTQWTSNIPGT---PQVFLGLP--ASATAA------GGG-YIDSSTLIEGGGVLEQIKMIGN-YGGVMLWA----
TXVB-2072783          FR------KG------ADAMVNA--WKKWAAEMPMN---TRIFLGIP--ESTAIP--------G-FISSSTL--KTAVLPRIKKIGK-YGGVMLWPDNSD
LGOW-2103657          YS------GG------FSQMTSV--WNEWSSGISSVSKATKIFLGLP--ASSAAA------GSG-YTSSSVV--VNTILPKFKKSTN-YGGVMLYDVQYD
FITN-2007300          YN------GD------ASNLISS--WNNWNSKLPS----TKIFMGLP--AASAAA------GSG-YMSSDVL--KSKVLPTIKSSSK-YGGVMFWSVYYD
FITN-2007298          YN------GD------TSNLISS--WNEWDSKLPD----TSIFMGLP--ASSDAA------GSG-YMSSDVL--KSDMLPTIKGSSK-YGGVMLWSVYYD
FITN-2007299          ----------------TTNLISS--WNEWNSKLPS----TKIFMGLP--AASAAA------GSG-YMSSAVL--KSQVLPTIKSSSK-YGGVMLWSVYYD
TFYI-2073779          YNP----VSGD-----VSGLLKA--WTTWTSGLPSG---SLVFLGLP--ASPEAATDEWSLEHG-YLPVYIL--ESLVLPRIGVDPR-FGGVMLWSYFHD
PIUF-2012805          YQ------GG------LAQELSV--WDQWTAALPN----AKIFMGLP--ASTSAA------GTG-FVAGDVV--TSDILPQIKKSSN-YGGVMLWSVYYD
JHFI-2013440          YQ------GG------LAQELSV--WNQWTAALPN----AKIFMGLP--ASTSAA------GTG-FVASNVV--TSNILPQIKKSSN-YGGVMLWSVYYD
PIUF-2090613          YHP----EGND-----TSGLLAA--WGTWTSALPW----SKIFLGLP--ASPSASTGVYSLEKG-FIPSDVA--VSDVLPHLKKSSN-YGGVMLWSFKYD
JHFI-2112286          YHP----EGND-----TSGLLAA--WGTWTTSLPG----SKIFLGLP--ASPSASTGVYSLDKG-FIPSDVT--VSDVLPHLKKSIN-YGGVMLWSFKYD
JHFI-2012415          YHR----KGNH-----YSDLYSS--YEQWSQALPN----SQIFLGLP--SSPTAANDSYDLDLG-YIPSSKV--VSNVLPHIRTYPN-YGGVMLWCYYYD
JHFI-2012990          YHR----NGNH-----YSDLYSS--YEQWSQALPN----SLIFMGLP--SSPTAADDSYNLDLG-YIPSNVV--ISNVLPHIRTYPN-YGGVMLWCYYYD

                              410
                      ....|....|....
4TOQ_1|Class_III      T---TYSTTIKDQV
AFY08286.1_Class_III  LKN-GYSKAIIGSV
ABW75910.1_Class_III  EQN-NYSNQVKGSV
LFVP-2004510          KENNNYSSQIKSDV
LFVP-2083339          KKS-HFSATIRHKV
LFVP-2083179          KNP-SYSAAIK---
LFVP-2083413          NQN-SYSNQISNGI
LFVP-2083566          R---TYAASIKSVV
HMHL-2053338          KENNNYSSQIKSDV
```

```
IHWO-2065363          KKS-HFSATIRHKV
HMHL-2052653          NQN-SYSNQISNGI
IHWO-2065195          NQN-SYSNQISNGI
IHWO-2065225          R---TYAASIKSVV
IHWO-2065167          KNP-SYSAAIK---
IHWO-2005936          KENNNYSSQIKSDV
WNGH-2088662          DLNN----------
KRUQ-2111018          DLNN----------
CHJJ-2135670          DMN-----------
KRUQ-2111802          --------------
YBQN-2014678          --------------
RTMU-2182337          --------------
RDOO-2001513          --------------
LGOW-2016622          --------------
HPXA-2018253          --------------
LGOW-2003062          --------------
JPYU-2038245          KQNNNYSAQIRSGV
WJLO-2035839          KQNDNYSAKIKDSV
TXVB-2013493          IQN-NYSAQIKSAV
TFYI-2002959          RENNNFSAQIRSAV
ILBQ-2043168          QQNNNYSAQIKGSV
TFYI-2073368          RQN-GYSNQIKSHV
WJLO-2007388          --------------
JPYU-2007981          --------------
TXVB-2072783          T---GYASAIKGAV
LGOW-2103657          KQT-GFSSAIASAV
FITN-2007300          KQT-NYAASIRSSI
FITN-2007298          GQT-SYASSIRSSV
FITN-2007299          KQT-TYASSIRSSV
TFYI-2073779          KTS-QFSATIR---
PIUF-2012805          QIT-SFSSSI----
JHFI-2013440          RIT-SFSSSI----
PIUF-2090613          QNP-SYSGEI----
JHFI-2112286          QNP-SYSGEILRAV
JHFI-2012415          --------------
JHFI-2012990          --------------
```

Figure S3. Class III chitinases alignments of complete sequences identified in pteridophytes

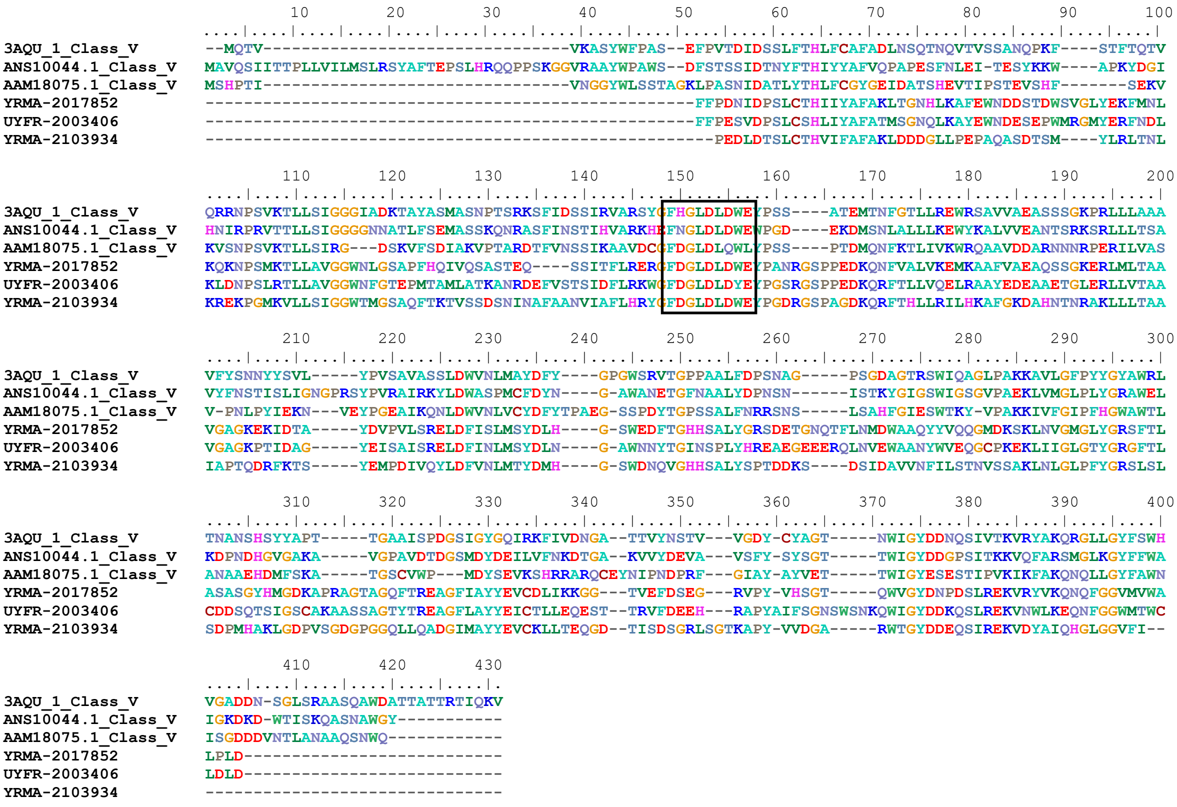


```
                               10        20        30        40        50        60        70        80        90       100
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      --------------------------GDIAIYWGQNGG-----------EGTLASTCDTGRYAYVIVSFVTTF-GNFRA------PVVNL-AGHCD----
AFY08286.1_Class_III  -----------------------MKAAGIAVYWGQNGG-----------EGSLAATCNTGNYKFVNIAFLSTF-GGGKT------PQLNL-AGHCN----
ABW75910.1_Class_III  -MAAKLKCSPLLPLLLLAGMAGISRAGNIAVYWGQNGY-----------EGTLADACNSGLYKYVTLSFLSTF-GNGQT------PALNL-AGHCN----
DFHO-2007268          ----------VLIVLLLTFLPFSHQQGFLVTYWGQNGN-----------EGSLVAACQKNLYNVVILSFVSTF-GNNQQ------PQLNL-AGHCD----
ZXJO-2015647          -------------------------AGDLVTYWGQNGN-----------EGSLAQACQTGLYNTVMLSFLSVF-GLGQQ------PQLNL-AGHCD----
ZZOL-2053046          ----------------------------IAIYWGQDGN-----------EGTLLSACNTGRYRILIISFLSQF-GRGQT------PLLNL-AGHCD----
WQML-2003973          -------------------------AGDLVTYWGQNGN-----------EAPLAQACQTGLYNTLILSFLSVF-GEGQQ------PVLNL-AGHCD----
BEGM-2102286          ----------------------------LVTYWGQNGN-----------EGSLSNACQSGLYNIVIVSFLTVF-GNNQQ------PVLNL-AGHCD----
JKAA-2177540          ----------------------------IAIYWGQNGN-----------EGTLANACNSGNYNIIMLSFLNVF-GNGRT------PSLNL-AGHCD----
ABIJ-2006436          --------------------------GSIAIYWGQNGN-----------EGSLSSACSSGNYNIIMLSFLNVF-GGGQT------PGLNL-AGHCD----
ABIJ-2009255          --------------------------GIIATYWGQNGG-----------EGSLIDACNSGNYRIIMISFLSQF-GNGQS------PVLNL-AGHCN----
ABIJ-2038774          ----------------------------IAIYWGQNGN-----------EGSLINACNTGRYGILIISFLSQF-GRGQT------PLLNL-AGHCN----
WGTU-2071532          --------------------------GNLVTYWGQNGN-----------EGSLAQACQSGLYNTLILSFLSVF-GEGRQ------PQLNL-AGHCD----
CAPN-2039382          -------------LLLLCCMVWPGEAGLLVTYWGQNGN-----------EGSLADACNSGLYDIVILSFLNTF-GNGQQ------PQLNL-AGHCD----
ZQYU-2112073          -------------------------AGDLVTYWGQNGN-----------EGSLAQACQSGLYNVLILSFLSVF-GQGQQ------PQLNL-AGHCD----
YJJY-2075638          ARLNTVPMISSILLLTFCGLSGWTVAGNLVTYWGQNGN-----------EGSLAQACQSGLYNTLILSFLSVF-GESQQ------PQLNL-AGHCD----
XXHP-2001291          ------------LLTFCALSGFTVQAGNLVTYWGQNGN-----------EGSLAQACQSGLYNTLILSFLSIF-GQGQQ------PQLNL-AGHCD----
ULKT-2008784          -------------LLLVFAELDCSSAGFIATYWGQDGN-----------EGSLAGACTTGNYGAINIAFLNVF-GGGQT------PSLNL-AGHCD----
ULKT-2015752          ----------------------------IATYWGQNGN-----------EGTLADACNTGNFGAINVAFLNVF-GGGQT------PGLNL-AGHCD----
UPMJ-2018058          ------------------------RKDFIATYWGQNGN-----------EGTLADACNTGNFGVINVAFLNIF-GGGQT------PSLNL-AGHCD----
UPMJ-2005627          -----------------------SSAAFIATYWGQDGN-----------EGSLAGACNTGNYGVINVAFLNVF-GGGQT------PSLNL-AGHCN----
PKOX-2000313          --------------------------GKIAIYWGQNGN-----------EGTLSSACDSGNYDIIMLSFLYQF-GGGQT------AKLDL-AGHCD----
PKOX-2000311          -----------------------STKGKIAIYWGQNGN-----------EGTLSTACDSNNYDIIMLSFLSQF-GGGQT------AQLNL-AGHCN----
PKOX-2008786          ------------------------QAGKIAIYWGQNGN-----------EGSLSNACDSGNYDIIMLSFLYQF-GGGQD------AQLNL-AGHCD----
KUXM-2043003          --------------------------GSVAIYWGQNGN-----------EGSLSAACNSGNYNIIMLSFLNNF-GSGTT------PSLNL-AGHCN----
XDDT-2009479          --------VILMLVLCSLSGCKRAAAADLVTYWGQNGD-----------EGSLTQACQTGLYNTLILSFLSVF-GQGQQ------PQLNL-AGHCD----
XDDT-2009480          -------------------------AGDLVTYWGQNGD-----------EGSLAQACQTGLYNTLILSFLSVF-GQGQQ------PQLNL-AGHCD----
UXCS-2007029          ----------------------------LVTYWGQNGN-----------EGSLAQACQSGLYNTLIISFLNVF-GLGQQ------PQLNL-AGHCN----
PQTO-2079202          -----------------------SSAAFIATYWGQNGG-----------EGSLADACKTGNFGVINVAFLNTF-GGGQT------PGLNL-AGHCD----
NYBX-2006531          ----------------------------IATYWGQNGN-----------EGSLASACNTGNYGVIMVAFLNDF-GGGQS------PGLNL-AGHCD----
NYBX-2006530          ----------------------------IATYWGQNGN-----------EGSLASACNTGNYGVIMVAFLNDF-GGGQS------PGLNL-AGHCD----
NYBX-2006529          ----------------------------IATYWGQNGN-----------EGSLASACNTGNYGVIMVAFLNDF-GGGQS------PGLNL-AGHCD----
GAON-2007131          -----------------------ARAKFIATYWGQNGN-----------EGSLASACNTGNYGVINVAFLNDF-GGGQS------PGLNL-AGHCD----
GKAG-2011802          ----------LFEVLLVICHLDAARAQFIATYWGQNGN-----------EGSLASACNTGNYGVIMVAFLNDF-GGGQS------PGLNL-AGHCD----
GKAG-2011803          ----------------------------IATYWGQNGN-----------EGSLASACNTGNYGVIMVAFLNDF-GGGQS------PGLNL-AGHCD----
ZFGK-2040681          --------------------------GVIAIYWGQDSS-----------EGDLITNCNKGNYKIIMISFLSQF-GGGQT------PVLNL-AGHCD----
ZFGK-2039578          -----------------------AQGGSIAIYWGQNGN-----------EGTLLSACNSGNYNIIMLSFLNNF-GNGIT------PSLNL-AGHCD----
ENQF-2005526          --------------LLLVNALDYSSAAFIATYWGQNGN-----------EGRLVDACNTRNYGVINVAFLNTF-GGAQS------PGLNL-AGHCD----
WAFT-2002690          ----------------------------IATYWGQDGN-----------EGPLVGACNTGNYGVINVAFLNVF-GGGQA------PGLNL-AGHCN----
PYHZ-2007870          ------------------------------IYWGQNGN-----------EGTLSSACDSNNYDIIMLSFLSQF-GGGQT------AQLNL-AGHCD----
PYHZ-2007871          ------------------------------IYWGQNGN-----------EGTLSSACDSNNYDIIMLSFLSQF-GGGQT------AQLNL-AGHCD----
PYHZ-2004504          ------------------------QAGKIAIYWGQNGN-----------EGSLSKACDSGNYDIIMLSFLYQF-GGGQD------AQLNL-AGHCD----
CBAE-2005524          -----------------------AKAKFIATYWGQNGN-----------EGSLASACNTGNYGVINVAFLNDF-GGGQS------PGLNL-AGHCD----
CBAE-2005523          --------LLFEVLLIGVCHLDAATAAFIATYWGQDGN-----------EGSLASACNTGNYGVINVAFLNDF-GGGQS------PGLNL-AGHCD----
KJYC-2013372          --------------------------GKIAIYWGQFGD-----------EPTLAETCNTGNYDIIMIAFVIDF------------HPLNLFLASCK----
JKAA-2013110          --------------------APFAQAGLIATYWGQDGN-----------EGDLSQACSSGNYGIIMLAFLAVF-GNGQT------PQLNL-AGHCD----
JKAA-2178880          --------------------------GQIATYWGQDGN-----------EGDLSQACNSGNYNIIMISFLSVF-GNGQT------PQLNL-AGHCT----
JKAA-2180047          ---------------------------NVAVYWGQNSIGAIAGTNPANYQKTLSFYCNDNAVDVFPVAFVNVFFGTGGA------PSINL-ANTCNPTDN
JKAA-2181250          -----LSNSLKCFLLACIFMQAVSASNQMVIYWGQNSAGSFTNV-KSQWEQPLRNYCQDSSYDIIVVGFVYLF-PSASGNTDPTMPGMNF-ANHCEVPFD
JKAA-2180911          -----------------------SVAPKMVIYWGQNSAGSFTNK-PSEQEKPLRNYCQDSRFDIIVVGFVYLF-PSAPGNTDPTLPGMNF-ANHCEVPFN
YOWV-2066886          ------------LLTFCALSGFTVQAGNLVTYWGQNGN-----------EGSLAQACQSGLYNTLILSFLSIF-GQGQQ------PQLNL-AGHCD----

                              110       120       130       140       150       160       170       180       190       200
                      ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III      ---PAAG--TCT-GLSDEIRSCQGKDIKVLMSI-GGGAGDYSLVSEADADNFADYLWNNFLGGQSSSRPLGDAVLDGIDFDIELG--TT-TFYDTLARAL
AFY08286.1_Class_III  ---PAAG--TCK-GISADIKTCQSKGIKVLLSL-GGGTNGYSLNSAAEANQPATYLWNNFLGGRSNSRPLGNAVLDGIDFDIEAG--PG-QHWDELARAL
ABW75910.1_Class_III  ---ASSG--GCT-GLSSDIQTCQSQGVKVLLSL-GGGSGSYGLSSTDDAQSVATYLWDNFLGGSSSSRPLGPAVLDGIDFDIETG--NS-AHWDELANFL
DFHO-2007268          ---PSTN--GCT-FLSSDIQTCQSQGVIILLSI-GGGAGSYGLNSDQDAQSVASYLWNNYLGGNSGSGPLGAANLDGIDFDIETG--GS-NYYGTLAQTL
ZXJO-2015647          ---PSGG--TCT-GLSSDIATCHSSGVKVLLSI-GGGAGSYGLSSADDAASVASYIWNNYLGGQSSDRPLGPAVLDGVDFDIELG-DGA-SFYGSLAQSL
ZZOL-2053046          ---PPSG--GCK-GLTSDINACKSKGIKVLLSI-GGGAGSYGLSSSSDGESVATYIWNNYLGGSSSNRPLGSAVLDGVDFDIETG-NTPQQYYIAMAQKL
WQML-2003973          ---PAGG--TCT-GLSSDINTCQSSGVKVLLSI-GGGAGSYGLSSAEDAGSVASYIWNTFLGGQSSARPLGPAVLDGVDFDIELG-DGA-SFYGSLAQSL
BEGM-2102286          ---PASG--ACT-GLTSEIQSCQSLGIKVLLSI-GGGAGSYDLTSEDDASSVATYLWENYLGGQSASRPLGSAILDGIDFDIELGSSGS-SYYGSLAQSI
JKAA-2177540          ---PSRN--GCT-SLSSDIKTCQLKGIKIFLSL-GGGVGTYSISSASDADSVATYLWNNYLGGRSRSRPLGPAVLDGIDFDIEQG-RTP-DRYALLAQSI
ABIJ-2006436          ---PSAN--GCT-TLSSQISSCQKRGIKILLSL-GGAIGNYTISSSSDARSVARYLWNNFLGGRSSSRPLGSAVLDGIDFDIESG-AIP-ERYALLARAI
ABIJ-2009255          ---PADG--SCV-GLSSEISQCQSRGIRILLSI-GGGAGSYGLSSSSDAQSVANYIWNNFLGGTSSSRPLGSAVLDGVDFDIETG-ASP-QLYVDLARAL
ABIJ-2038774          ---PASG--GCA-YLTSEINSCKSKGIKVLLSI-GGGAGSYGLSSASNGEAVATYLWNNYLGGSLGSRPLGSAVLDGIDFDIEMG-NTPQQYYIALAQKL
WGTU-2071532          ---PASG--TCT-ALSSDISTCQNLGIQVLLSIGGGGASSYGLSSADDASSVASYIWDNYLGGQSTDRPLGSAVLDGVDFDIESG-NGA-SFYGSLAQAL
CAPN-2039382          ---PAAG--TCT-SLSSDIDTCHNAQKKVLLSI-GGAVGTYSLQSDAEASTLADYLWNNFLGGQSTSRPLGASSLDGIDFDIEQG--SF-DYYASLAQYL
ZQYU-2112073          ---PASG--GCT-GLSSDISTCHNLGITVLLSI-GGAGGSYGLSSADDASSVASYIWNNYLGGQSSDRPLGSAVLDGVDFDIEQG-SGA-SFYGSLAQGL
YJJY-2075638          ---PASG--GCT-GLSSDISTCQSLGIVVLLSI-GGGAGSYGLSSAADAASVANYIWENYLGGQSSSRPLGSAVLDGVDFDIELG-NGA-SYYGSLAQTL
XXHP-2001291          ---PPSG--SCT-FLSSDISTCHSLGITVLLSI-GGGAGSYGLSSAADAASVANYIWNNYLGGQSSSRPLGAAVLDGVDFDIELG-DGA-SYYGSLAQAL
ULKT-2008784          ---PASG--TCA-QFSSDIQKCQEIGVKVLLSL-GGAVGTYNINSQSDADSVAQYLWDNFLGGQSSSRPLGAAVLDGIDFDIEQG-NGL-ESYNLLAQSI
ULKT-2015752          ---PPSG--TCV-IFSSDIQTCQRRGVKILLSL-GGAFGNYTITSKTDADEVALYLWDNFLGGNSSSRPLGAAVLDGIDFDIEQG-NGF-GSYALLAQSI
UPMJ-2018058          ---PPSG--TCV-KFSSDIQICQSRGVKILLSL-GGAFGNYTITSKTDADAVALYLWDYFLGGNSSSRPLGSAVLDGIDFDIEQG-SGF-GSYALLARSI
UPMJ-2005627          ---PASG--TCT-QLSSDIQTCQQENVKVLLSL-GGAVGTYDINSQSDADSVAQYLWDNFLEGQSSSRPLGAAVLDGIDFDIEQG-NGL-ESYNLLAQSI
PKOX-2000313          ---ASSN--GCT-GLSSQISSCQQKGVKILLSL-GGGTGSYKLSSSSDASTVATYIWNNYLGGSSSSRPLGSAVLDGVDLDIEAG-GTP-SLYADLANDL
PKOX-2000311          ---PSNN--GCT-GLSSQISSCQQKGVKILLSL-GGGSGSYKLSSSSDASAVANYIWNNFLGGSSSSRPLGSAILDGVDLDIEAG-GTP-GLYADLARYL
PKOX-2008786          ---PSSG--GCT-GLSSEISSCQSKGIKILLSL-GGGSGSYKLTDSSDASSVASYIWTNYFSG-SSSGPLGSAVLDGIDLDIESG-STP-DLYVDLVNDL
KUXM-2043003          ---PSAN--GCT-SLSSEISSCQLKGIKIFLSL-GGGVGTYSITSASDAKSVATYLWNNYLGGKSNSRPLGSAVVDGIDFDIERG-ASP-AKYALLAQAI
XDDT-2009479          ---PAGG--TCT-GLSSDITTCQSSGVQVLLSI-GGGAGSYGLSSTDDAASVASYIWTNYLGGQSSARPLGAAVLDGVDFDIEIG-DGA-SFYGSLAQSL
XDDT-2009480          ---PAGG--TCT-GLSSDITTCQSSGVQVLLSI-GGGAGSYGLSSTDDAASVASYIWTNYLGGQSSARPLGAAVLDGVDFDIEIG-DGA-SFYGSLAQSL
UXCS-2007029          ---PAAG--TCT-TFSSDISTCQSLGITVLLSI-GGGAGSYGLSSADDATSVASYIWNNYLGGQSSNRPLGAAVLNGVDFDIEVA-TGA-SFYGDLAQAL
PQTO-2079202          ---PPSG--TCI-QFSSDIQSCQQQGVKILLSL-GGAAGNYHINSQSDADSVAQYLWDNYLGGQSSSRPLGAAVLDGIDFDIELS-NGL-ESFDLLAKSI
NYBX-2006531          ---PPSG--TCV-KFSSDIQSCQQMGVKILLSL-GGGVGNYNINSQAEANSVAQYLWDNFLEGQSSSRPLGAAVLDGIDFDIEQG-NGF-ESYALLAQSI
NYBX-2006530          ---PPSG--TCV-KFSSDIQSCQQMGVKILLSL-GGGVGNYNINSQAEANSVAQYLWDNFLEGQSSSRPLGAAVLDGIDFDIEQG-NGF-ESYALLAQSI
NYBX-2006529          ---PPSG--TCV-KFSSDIQSCQQMGVKILLSL-GGGVGNYNINSQAEANSVAQYLWDNFLEGQSSSRPLGAAVLDGIDFDIEQG-NGF-ESYALLAQSI
GAON-2007131          ---PASG--TCV-KFTSDIQRCQKLGVKILLSI-GGGAGKYNINSQAEANSVAQYLWDNFLGGHSRSRPLGAAVLDGIDFDIEQG-NGF-ESYALLAQSI
GKAG-2011802          ---PPSG--SCV-KFSSDIQSCQQMGVKILLSL-GGAVGKYNINSQAEADSVAQYLWDNFLEGQSSSRPLGAAVLDGIDFDIEQG-NGF-ESYALLAQSI
GKAG-2011803          ---PPSG--SCV-KFSSDIQSCQQMGVKILLSL-GGAVGKYNINSQAEADSVAQYLWDNFLEGQSSSRPLGAAVLDGIDFDIEQG-NGF-ESYALLAQSI
ZFGK-2040681          ---PGSG--GCQ-FLSSQIKACQQMGIKILLSI-GGGAGSYKLSSTSDANSVASYLWNTYLGGSSSSRPLGAAVLDGIDFDIEQN-PSP-QRYIDLANAL
ZFGK-2039578          ---PSSN--GCT-SLSSQISSCQRKGVKILLSL-GGGVGTYSITSASDAKSVATYLWNNYLGGKSTSRPLGSAVLDGIDFDIESG-ASP-DKYILLAQAI
ENQF-2005526          ---PASG--TCA-MFSSDIQTCQQNGVKILLSL-GGAVGTYDINSQAEADSVAQYLWENFLGGQSNSRPLGAAVLDGIDFDIEQG-NGL-ESYNLLAQSI
WAFT-2002690          ---PPSG--TCT-QFSSDIQTCQQNGVKIFLSL-GGKVGTYNINSQAEADSVAQYLWDNFLGGQSSSRPLGAAVLDGIDFDIEQG-NGL-ESYNLLAQSI
PYHZ-2007870          ---PSNN--GCT-GLSSQISSCQQKGVKILLSL-GGGSGSYKLSSSSDASIVANYIWNNFLAGSSSSRPLGSAVLDGVDLDIEAG-GTP-GLYADLARDL
PYHZ-2007871          ---PSNN--GCT-GLSSQISSCQQKGVKILLSL-GGGSGSYKLSSSSDASIVANYIWNNFLAGSSSSRPLGSAVLDGVDLDIEAG-GTP-GLYADLARDL
PYHZ-2004504          ---PSSG--GCT-GLSSEISSCQSKGIKILLSL-GGGSGSYKLTDSSDASSVASYLWTNYFSG-STSGPLGSAVLDGIDLDIESG-STP-DLYVDLVNDL
CBAE-2005524          ---PASG--TCV-KFSSDIQRCQKLGVKILLSI-GGGAGKYNINSQAEANSVAQYLWDNFLGGHSRSRPLGAAVLDGIDFDIEQG-NGF-ESYALLAQSI
```

```
CBAE-2005523         ---PASG--TCV-KFSSDIQRCQKLGVKILLSI-GGGAGKYNINSQAEANSVAQYLWDNFLGGHSRSRPLGAAVLDGIDFDIEQG-NGF-ESYALLANP-
KJYC-2013372         --DPPSG--GCI-ALGPDIDACHEKGVKLLLSL-GGAAGTYHLNDKDDAQRVAADLFNNFLGGNSSSRPFGPAKLDGIDLDIEYL-ITP-DLYAELAMEL
JKAA-2013110         ---PPSG--GCA-RFGSQIQACQQQGIKVLLSI-GGGAGGYGLSSSGDADSLADHIWNNYLGGNSGSGPLGSAVLDGVDFDIETG-GTP-QLYVDLANGL
JKAA-2178880         ---PSSG--GCT-FLSSQIKSCQQRGILILLSI-GGGAGSYGLSSTADANSVATYLWNNYLGGSSPNRPLGAAVLDGVDFDIEQG-STP-QRYVDLANAL
JKAA-2180047         ATFPGTALPNCS-KLASDITTCQSKGKLVTISL-GGATGSVGFASDAQGTAFAQTIWNLFLGGTSSTRPFGTAVLDGVDLDIEGG--GG-VGYAAFVNTI
JKAA-2181250         DANPFVL--NCTANIAPDITYCQQRGKKIILSF-GGGVGNYGFNSDSQAITFANTVWNMFLGGSGSVRPFGNAIFDGVDLDIEGG--TS-TGFIAFINTL
JKAA-2180911         DANPFVL--NCTANVAPDIKYCQQQGKTILLSF-GGGIGPYGFSSDNQGVTFAETVWNMFLGGSGSIRPFGDAIFDGVDLDIEGG--TT-AGWIPFVKTL
YOWV-2066886         ---PPSG--SCT-FLSSDISTCHSLGITVLLSI-GGGAGSYGLSSAADAASVATYIWNNYLGGQSSSRPLGAAVLDGVDFDIELG-DGA-SYYGSLAQAL

                             210       220       230       240       250       260       270       280       290       300
                     ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III     SSRST-------QAAKVYLTAAPQCPHPDSHLDA-----ALNTGLFDNVWIQFYNNPLAQCQYS---SGNTNDILSSWNTWTSST---TAGKIFLGLPAA
AFY08286.1_Class_III KRFNQ----------KLILSAAPQCPFPDSHLRS-----AINTGLFNYVWVQFYNNPS--CQYT---GGNINNLVNSYKQWTTV----KASQVFLGVPAS
ABW75910.1_Class_III SRYSA-------QGQKVYLTAAPQCPYPDASLGP-----ALQTGLFDNVWIQFYNNPP--CQYT---SGDASNLQKAWNTWTSDV---KAGSFYLGLPAS
DFHO-2007268         RQLSA-GS-----SKTYYLSAAPQCPYPDANLGPG-SGTALATGLFDYVWVQFYNNGP--CEY----SNDDTALLSAWSQWTSGV---SGAKILLGLPAS
ZXJO-2015647         RQLSS----------QVILGAAPQCPFPDAHLGPNLSGSALDTSLFTYVWVQFYNNPS--CQY----GSDASALLAAWQQWTSAL---PSAKVLLGLPAS
ZZOL-2053046         KSFNC----------NLILSAAPQCPIPDASLGS----VIKVPGLFTYVFVQFYNNPQ--CQYN---NNGVSGILSSWNNWVSNSVTPSTVKVFMGLPAS
WQML-2003973         RQFSS----------QVILGAAPQCPFPDAHLGPSLSGSALDSGLFNYVWVQFYNNPS--CQY----GSDASGLLAAWQQWTSAL---PSARVLLGLPAS
BEGM-2102286         SQLSG---------GQALLTAAPQCPFPDAHLGP-----ALNTGLFNYVWVQFYNNPQ--CQYS---GGDATALLSAWSQWTSSL---SGAKVLLGLPAS
JKAA-2177540         KRISK------GGSRPVLLSAAPQCPFPDANLGS----AISQPGLFDYVFVQFYNNPQ--CQYD---GSSASKLLSSWSQWTSQL---RHVKLFMGLPAA
ABIJ-2006436         RDVSR--------GSKVLISAAPQCPFPDANLGS----AIAKPGLFDYIYVQFYNNPP--CQYD---GNSTSNLLDSWREWSSKT---PDLKLFMGLPAA
ABIJ-2009255         RSFNG----------GMILSAAPQCPIPDANLGS----TIQVSGLFNYIFVQFYNNPA--CQFD---GNSVSNILASWNSWVSDSVTPSSATVFMGLPAS
ABIJ-2038774         KSYNP----------NLILSAAPQCPIPDASLGN----VIKPPGLFTYIFVQFYNNPP--CQY----NGGVSQIISSWNRWVSSSVTPSTVKVFMGLPAS
WGTU-2071532         RQLSS----------GVILGAAPQCPFPDAHLGPNLSGSALDSGLFNYVWVQFYNNPP--CQY----GSDASALLAAWQQWTSAL---PQAQILLGLPAS
CAPN-2039382         HNYDS----------GMILSAAPQCPYPDAYMGP-----ALSTGLFQYVWVQFYNNPP--CQY----SGDDSALVTSWNTWTSSL---SSSLILLGLPAS
ZQYU-2112073         RQLSP----------GVILGAAPQCPFPDAHLGPNLSGSALDSGLFNYVWVQFYNNPP--CQY----GSDASSLLAAWQQWTSAL---PQAQILLGLPAS
YJJY-2075638         RQLSS----------GVILGAAPQCPFPDAHLGPNLSGSALDSGLFNYVWVQFYNNPP--CQY----GNDASALLAAWQQWTLAL---PQAQILLGLPAS
XXHP-2001291         RQLSS----------GVILGAAPQCPFPDAHLGPNLSGSALDSGLFNYVWVQFYNNPP--CQY----GSDASSLLAAWQQWTSNL---PQAQILLGLPAA
ULKT-2008784         TSLTN-GG-----TKQYYLSAAPQCPFPDAHLGPG-SGSALDTGLFDYVWVQFYNNPS--CNFNPSSSGDQSSALSAWSNWVQNT---QAQNIFLGLPAS
ULKT-2015752         RHITESKSPASFNKKNYYLSAAPQCPFPDAILGPG-GSSALATGLFDYVWVQFYNNPQ--CEFNATANGDQSTGLAAWKTWIHQT---PAEKIFLGLPAS
UPMJ-2018058         RKITIAKGLLCPEKKTYYLSAAPQCPFPDAHLGPG-GSSALQTGLFDYVWVQFYNNPP--CNFNSTPDGDQSSALATWKAWIKQT---PARKIFLGLPAS
UPMJ-2005627         RSLTN-GG-----SKQYYLSAAPQCPFPDAHLGPG-AGSALDTGLFDYVWVQFYNNPA--CEFNSSANGDQSSALSAWNTWVQNT---QAKDIFVGLPAS
PKOX-2000313         RSLAG--------SKKIYISAAPQCPYPDANL------------LFDYVYVQFYNNPD--CQY----NGDTSNLISSWNTWNSKL---PSTRIFMGLPAA
PKOX-2000311         KSLAG--------SRKIYISAAPQCPYPDANLGS----ALLNSGLFDYVYVQFYNNPG--CQY----NGATTNLISSWNEWNSKL---PSTKIFMGLPAA
PKOX-2008786         KSLAG--------NTKLYI--APQCPYPDSNLGS----ALLTSGLFDYVYVQFYNNQD--CQY----NGDTSNLISSWNEWDSKL---PDTSIFMGLPAS
KUXM-2043003         KDISG---------SQVLLSAAPQCPFPDANLGS----AISKPGLFDYIYVQFYNNPS--CQYD---GSSVDKLIASWNQWNSKL---PNVQIFMGLPAS
XDDT-2009479         RQFSS----------QVILGAAPQCPFPDAHLGPNLSGSALDTGLFNYVWVQFYNNPS--CQY----GSDASGLLAAWQQWTSAL---PSARILLGLPAS
XDDT-2009480         RQFSS----------QVILGAAPQCPFPDAHLGPNLSGSALDTGLFNYVWVQFYNNPS--CQY----GSDASGLLAAWQQWTSAL---PSARILLGLPAS
UXCS-2007029         RQFSS----------SVILSAAPQCPFPDANLGPNLSGSALDSNLFNYVWVQFYNNPP--CQY----GTDASALLAAWQQWTTAL---PQAQILLGLPAS
PQTO-2079202         RQRTN-GG-----SKQYFLSAAPQCIYPDAHLGPG-AGSALQTGLFDYVWVQFYNNPP--CNFNRNSNADQSSGLAAWNSWLQTPA--QNTMVFLGLPAS
NYBX-2006531         RKLTGLRG-----SKRYYLSAAPQCPIPDAHLGPG-DGSALQTKLFDYVWVQFYNNPP--CNFNPTSGGDQSSALSSWNKWIQQT---PAQKIFLGLPAS
NYBX-2006530         RKLTGLRG-----SKRYYLSAAPQCPIPDAHLGPG-DGSALQTKLFDYVWVQFYNNPP--CNFNPTSGGDQSSALSSWNKWIQQT---PAQKIFLGLPAS
NYBX-2006529         RKLTGLRG-----SKRYYLSAAPQCPIPDAHLGPG-DGSALQTKLFDYVWVQFYNNPP--CNFNPTSGGDQSSALSSWNKWIQQT---PAQKIFLGLPAS
GAON-2007131         RKLTN-KG-----SKKYYLSAAPQCPFPDAHLGPG-DGSALHTGLFDYVWVQFYNNPP--CNFNPNSHGDQSRGLSSWNKWIQKT---PAQKIFLGLPAS
GKAG-2011802         RKLTG--------------LRGSKCPIPDAHLGPG-DGSALQTGLFDYVWVQFYNNPP--CNFNPTSGGDQSSALSSWNKWIQQT---PAQKIFLGLPAS
GKAG-2011803         RKLTG--------------LRGSKCPIPDAHLGPG-DGSALQTGLFDYVWVQFYNNPP--CNFNPTSGGDQSSALSSWNKWIQQT---PAQKIFLGLPAS
ZFGK-2040681         RKFGS-----------MILSAAPQCPIPDAQLGS-----TIQAGVFNYVYVQFYNNQP--CEY----NGGVTNLLNSWANWVSNSVTPSNTKVFMGLPAS
ZFGK-2039578         RDISR--------GSRVLLSAAPQCPFPDANLGS----AISKPGLFDYVYVQFYNNPP--CQYD---GSSVDKLIASWRQWNSKL---PNVKLFMGLPAS
ENQF-2005526         RKLTN-GG-----SKQYYLSAAPQCPFPDAHLGPG-DGSALQTGLFDYVWVQFYNNPP--CNFNPNSNGDQSSGLAAWNNWVTT---PAQKIFLGLPAS
WAFT-2002690         RKLTN-GG-----SKQYYLSAAPQCIFPDAHLGPG-SGSALDTGLFDYVWVQFYNNPP--CNFNPNSNGDQSGGLAAWNNWVQA----QAQKIFLGLPAS
PYHZ-2007870         KSLAG--------SRKIYISAAPQCPYPDANLGS----ALLTSGLFDYVYVQFYNNPD--CQY----NGDASNLISSWNNWNSKL---PSTKIFMGLPAA
PYHZ-2007871         KSLAG--------SRKIYISAAPQCPYPDANLGS----ALLNSGLFDYVYVQFYNNPG--CQY----NGATTNLISSWNEWNSKL---PSTKIFMGLPAA
PYHZ-2004504         KSLAG--------NTKLYI--APQCPYPDSNLGS----ALLTSGLFDYVYVQFYNNQD--CQY----NGDTSNLISSWNEWDSKL---PDTSIFMGLPAS
CBAE-2005524         RKLTN-KG-----SKKYYLSAAPQCPFPDAHLGPG-DGSALHTGLFDYVWVQFYNNPP--CNFNPNSHGDQSRALSSWNKWIQQT---PARQIFLGLPAS
CBAE-2005523         -----------------FASSLTRDLFPDAHLGPG-DGSALHTGLFDYVWVQFYNNPP--CNFNPNSHGDQSRALSSWNKWIQQT---PARQIFLGLPAS
KJYC-2013372         RKLDP----------NVIISAAPQCPIPDVYIGS----ALKIDGLFDYIWVQFYNNPP--CMYK---SDVGADLLASWDNWVKKEVS-ASSKIFVGLPAA
JKAA-2013110         RSRGA-----------SLLSAAPQCPIPDANLGS----AVQTPGLFNYVFVQFYNNPP--CQY----SGDTTNLISSWNNWVSDSTTSSGTQVFMGLPAS
JKAA-2178880         RRFNA----------KLKLSAAPQCPIPDANLGS----AIQVANTFDMVFVQFYNNNP--CQYD---GSSTTNLLNSWNQWVSNSVTPGNTKVFMGLPAA
JKAA-2180047         RSLAN-GA-----SKRYYITAAPQCVYPDAALGS-----VLNTANFDAIYVQFYNNP---CSLP-NFPGASTWNFGIWDQWARTLSPNPNVKVYIGAPAS
JKAA-2181250         RNYFA-SS-----GRKYYITSAPQCFFPD-RLGPG-TGTPLTEAWFDYIWIQFYNNH---CGLD---EFPTNFNFAKWAAAMTTAAN-PDVKLFIGAPAS
JKAA-2180911         RAHFA-TS-----SKTYYISSAPQCFYPD-RLGPG-AGTPLTEAWFDYLWIQFYNNH---CGLD-AFPGAQFN-FAQWASSFSNAAN-PNVKLFIGAPAS
YOWV-2066886         RQLSS----------GVILGAAPQCPFPDAHLGPNLSGSALDSGLFNYVWVQFYNNPP--CQY----GSDASSLLAAWQQWTSNL---PQAQILLGLPAA

                             310       320       330       340       350
                     ....|....|....|....|....|....|....|....|....|....|....|
4TOQ_1|Class_III     PEAAGS-GYIPPDVLTGQILPQIKTSA--KYGGVMLYSKFYD--TTYSTTIKDQV
AFY08286.1_Class_III TAAAGS-GFIPVNVMNSQVLPAIKGST--KYGGVMLYDRFNDLKNGYSKAIIGSV
ABW75910.1_Class_III QDAASS-GYVPPGDLTSTVLPGVKGVG--NYGGIMLWSRYYDEQNNYSNQVKGSV
DFHO-2007268         TDAASS-GFISADVLTSQVLPSIKGSG--NYGGVMLWSYYYDQINNYGATIKSSL
ZXJO-2015647         PAAAGS-GYLPPDTLLSDVLPQIRSTS--NYGGVMLYSYFYDQSSGYGSAIKASV
ZZOL-2053046         PQAAGS-GYMPPNVLTSQVLPTIKGSS--KYGGVMLYSVAYD--KSYSSAIVSSV
WQML-2003973         SSAAGS-GFLPSAILLSDVLPQIRGSS--NYGGVMLYSYFYDQTSGYGSAIKSSV
BEGM-2102286         ADAAPSGGFITSDALTSQVLPQIKETS--NYGGVMLYSYFYDGTSGYGASIKSSL
JKAA-2177540         SEAAGS-GFIPSSVFRSSVLPVIKRER--NYGGVMFWSVFYDRQTSYTNSIRAS-
ABIJ-2006436         PAAAGS-GFLPSQVLRDRVLPRIR-------------------------------
ABIJ-2009255         SSAAGS-GFMPSSTLTSQVLPTIRSSP--KYGGVMLYSVFYDRQ-----------
ABIJ-2038774         TQAAGS-GFMPANVLTSQVLPTIKSSS--KYGGVMLYSVAYD--QGYSQAIVSSV
WGTU-2071532         SDAAGS-GFIPSNTLLSDVLPQIRDSS--NYGGVMLWSYFYDQSSGYGSAIKASV
CAPN-2039382         TDAAPSGGFIASSDLINTVLPDIKGTS--NYGGVMIYSRYYDGMSGYGTAIRSSV
ZQYU-2112073         TGAAGS-GFIPASTLLSDVLPQIHGSS--NYGGVMLYSYFYDQSSGYGASIKSSV
YJJY-2075638         SSAAGS-GFIPATALLADVLPQIRGSS--NYGGVMLYSYFYDQSSGYGSSIKSSV
XXHP-2001291         SGAAGS-GFIPASTLISEVLPQIRSTS--NYGGVMLYSYYYDETSGYGSTIKSSV
ULKT-2008784         SSAAGS-GFIPSDVLTSTILPQVKTAD--KYGGVMVWNYAQD--NGYSAAIISSV
ULKT-2015752         PSAAGS-GFIPSSVLISDILPRIKSSS--KFGGVMIWDFFHD-QGNYSSAIKASV
UPMJ-2018058         PSAAGS-GFIPPGVLISDILPRLKISSYGKFGGVMLWDYFHD-QGNYSSAIKSKV
UPMJ-2005627         SSAATS-GFIPSDVLKSTILP--------KYGGVMVWNYAQD--TGYSSAIISSI
PKOX-2000313         SAAAGS-GYMSSDVLKSKVLPTIKSSS--KYGGVMFWSVYYDKQTNYAASIRSSI
PKOX-2000311         SAAAGS-GYMSSAVLKSQVLPTIKSSS--KYGGVMLWSVYYDKQTTYASSIRSSV
PKOX-2008786         SDAAGS-GYMSSDVLRSDVLPTIKGSS--KYGGVMLWSVYYDGQTSYASSIRSSV
KUXM-2043003         SSAAGS-GYITSQTFKSNVLPKIKNSS--KYAGVMFWSVFHDRQTSYAQSIKAYV
XDDT-2009479         PSAAGS-GFLPSATLLSDVLPQIRSSS--NY------------------------
XDDT-2009480         PSAAGS-GFLPSATLLSDVLPQIRSSS--NY------------------------
UXCS-2007029         ADAAGS-GFISSDTLLSDVLPQIHGSS--NYGGVMLYSYFYD-------------
PQTO-2079202         DSAAGS-GFIPSDVVISDILPKIKGST--KYGGVMVWDYSHD-VNGYSSAIKPAI
NYBX-2006531         SSAAGS-GFIPSDVMTSTILPRIRNSP--KFGGVMVWDYSHD--NGYSSSIK---
NYBX-2006530         SSAAGS-GFIPSDVMTSTILPRIRNSP--KFGGVMVWDYSHD--NGYSSSIK---
NYBX-2006529         SSAAGS-GFIPSDVMTSTILPRIRNSP--KFGGVMVWDYSHD--NGYSSSIK---
GAON-2007131         SSAAGS-GFIPSGVLTSTILPRIKNSP--KFGGVMLWDYSHD--KGYSSSIK---
GKAG-2011802         SSAAGS-GFIPSDVLTSTILPRIRNSP--KFGGVMVWDYSHD--NGYSSSIK---
GKAG-2011803         SSAAGS-GFIPSDVLTSTILPRIRNSP--KFGGVMVWDYSHD--NGYSSSIK---
ZFGK-2040681         TAAAGS-GFIPSNVLVSSVLPTIKSSS--KYGGVMLYDCAQD--RTYSTAIKSSV
ZFGK-2039578         SAAAGS-GYINSQTFKSQVLPKIKGSN--KYAGVMFWSVFHDRQTSYAQSIRASV
ENQF-2005526         TSAAGS-GFIPSDVLISSILPQIKSST--KYGGVMVWDYSHD-------------
```

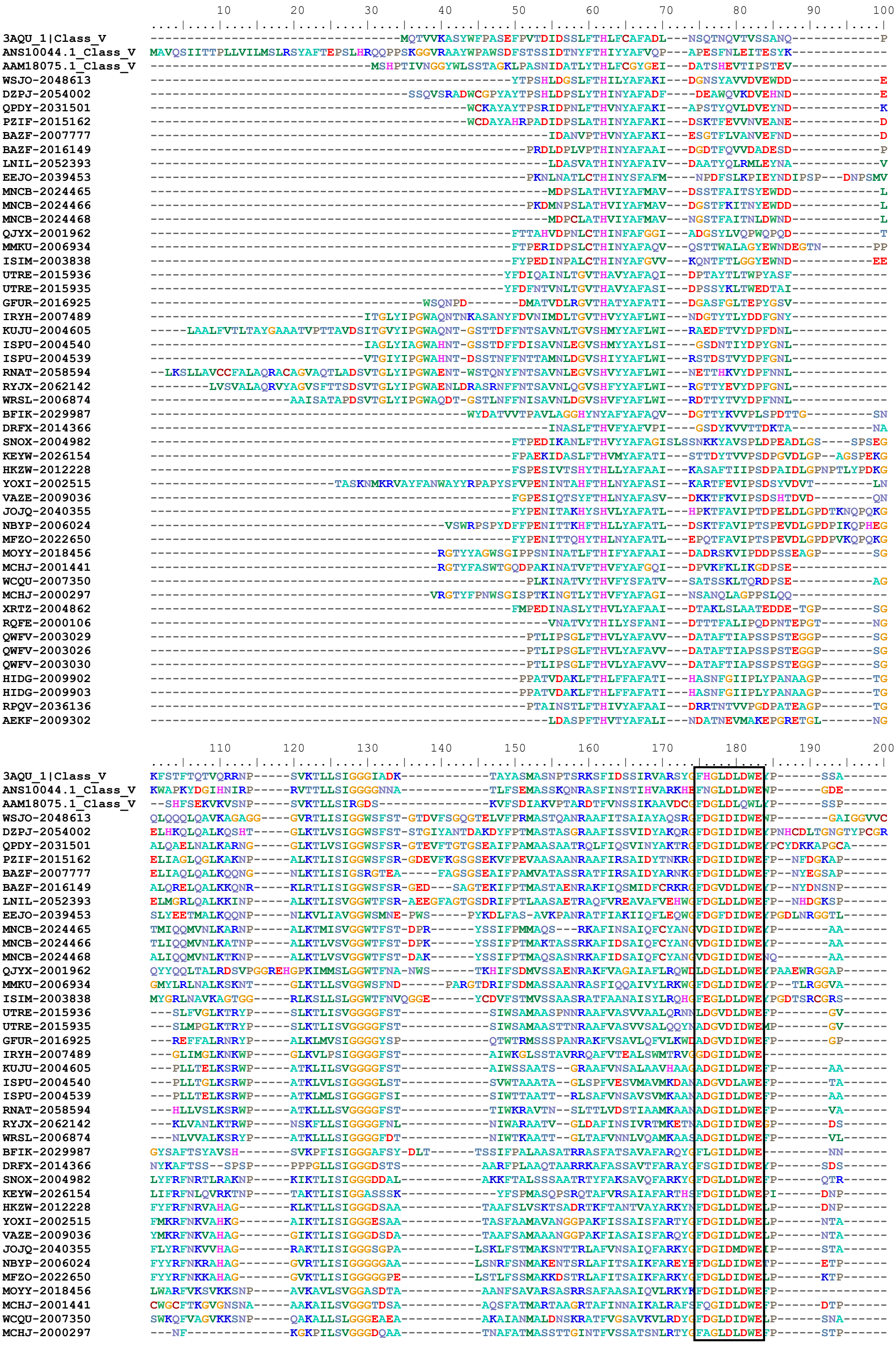
```
WAFT-2002690          TSAAGS-GFIPSDVLISSILPQIKSST--KYGGVMVWDYSHD--NGYSSAIK---
PYHZ-2007870          SAAAGS-GYMSSDVLKSKVLPTIKSSS--KYGGVMFWSVYYDKQTNYAASIRSSI
PYHZ-2007871          SAAAGS-GYMSSAVLKSQVLPTIKSSS--KYGGVMLWSVYYDKQTTYASSIRSSV
PYHZ-2004504          SDAAGS-GYMSSDVLKSDMLPTIKGSS--KYGGVMLWSVYYDGQTSYASSIRSSV
CBAE-2005524          SSAAGS-GFIPSGVLTSTILPRIKNSP--KFGGVMLWDYSHD--KGYSSSIK---
CBAE-2005523          SSAAGS-GFIPSGVLTSTILPRIKNSP--KFGGVMLWDYSHD--KGYSSSIK---
KJYC-2013372          TEAAPSGGYIEPDKFISEVLPHASLSP--KYGGVMLWAVYFDLQTKYSSKIRPYV
JKAA-2013110          TAAAGS-GFIPSDVLVSQVLPQIKSSS--KYGGVMLWSVAQD--SGYSAAIKSSV
JKAA-2178880          PTAAPSGGYIPPSVLTSRVLPAIKTSP--KYGGIMLYSVAYD-------------
JKAA-2180047          SSAAGS-GYVDINTLSKIAVQMRKSFS--SFGGVMLW------------------
JKAA-2181250          SLAAAS-GFVSLSTLQT-ISTAIQNTFPNLYGGMMLW------------------
JKAA-2180911          SMAAAS-GFVPISTLKT-ISTTIANTFPNIYGGMMVW------------------
YOWV-2066886          PGAAGS-GFIPASTLISEVLPQIRSTS--NYGGVMLYSYYYDETSGYGSTIKSSV
```

Figure S4. Class V chitinases alignments of complete sequences identified in algae phyla Rhodophyta and Ochrophyta



```
                            10        20        30        40        50        60        70        80        90       100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1_Class_V      --MQTV--------------------------------VKASYWFPAS--EFPVTDIDSSLFTHLFCAFADLNSQTNQVTVSSANQPKF----STFTQTV
ANS10044.1_Class_V  MAVQSIITTPLLVILMSLRSYAFTEPSLHRQQPPSKGGVRAAYWPAWS--DFSTSSIDTNYFTHIYYAFVQPAPESFNLEI-TESYKKW----APKYDGI
AAM18075.1_Class_V  MSHPTI--------------------------------VNGGYWLSSTAGKLPASNIDATLYTHLFCGYGEIDATSHEVTIPSTEVSHF-------SEKV
YRMA-2017852        ---------------------------------------------------FFPDNIDPSLCTHIIYAFAKLTGNHLKAFEWNDDSTDWSVGLYEKFMNL
UYFR-2003406        ---------------------------------------------------FFPESVDPSLCSHLIYAFATMSGNQLKAYEWNDESEPWMRGMYERFNDL
YRMA-2103934        -----------------------------------------------------PEDLDTSLCTHVIFAFAKLDDDGLLPEPAQASDTSM----YLRLTNL

                           110       120       130       140       150       160       170       180       190       200
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1_Class_V      QRRNPSVKTLLSIGGGIADKTAYASMASNPTSRKSFIDSSIRVARSYGFHGLDLDWEYPSS----ATEMTNFGTLLREWRSAVVAEASSSGKPRLLLAAA
ANS10044.1_Class_V  HNIRPRVTTLLSIGGGGNNATLFSEMASSKQNRASFINSTIHVARKHEFNGLDLDWEWPGD----EKDMSNLALLLKEWYKALVVEANTSRKSRLLLTSA
AAM18075.1_Class_V  KVSNPSVKTLLSIRG---DSKVFSDIAKVPTARDTFVNSSIKAAVDCGFDGLDLQWLYPSS----PTDMQNFKTLIVKWRQAAVDDARNNNRPERILVAS
YRMA-2017852        KQKNPSMKTLLAVGGWNLGSAPFHQIVQSASTEQ----SSITFLRERGFDGLDLDWEYPANRGSPPEDKQNFVALVKEMKAAFVAEAQSSGKERLMLTAA
UYFR-2003406        KLDNPSLRTLLAVGGWNFGTEPMTAMLATKANRDEFVSTSIDFLRKWGFDGLDLDYEYPGSRGSPPEDKQRFTLLVQELRAAYEDEAAETGLERLLVTAA
YRMA-2103934        KREKPGMKVLLSIGGWTMGSAQFTKTVSSDSNINAFAANVIAFLHRYGFDGLDLDWEYPGDRGSPAGDKQRFTHLLRILHKAFGKDAHNTNRAKLLLTAA

                           210       220       230       240       250       260       270       280       290       300
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1_Class_V      VFYSNNYYSVL-----YPVSAVASSLDWVNLMAYDFY----GPGWSRVTGPPAALFDPSNAG-----PSGDAGTRSWIQAGLPAKKAVLGFPYYGYAWRL
ANS10044.1_Class_V  VYFNSTISLIGNGPRSYPVRAIRKYLDWASPMCFDYN----G-AWANETGFNAALYDPNSN------ISTKYGIGSWIGSGVPAEKLVMGLPLYGRAWEL
AAM18075.1_Class_V  V-PNLPYIEKN---VEYPGEAIKQNLDWVNLVCYDFYTPAEG-SSPDYTGPSSALFNRRSNS-----LSAHFGIESWTKY-VPAKKIVFGIPFHGWAWTL
YRMA-2017852        VGAGKEKIDTA-----YDVPVLSRELDFISLMSYDLH----G-SWEDFTGHHSALYGRSDETGNQTFLNMDWAAQYYVQQGMDKSKLNVGMGLYGRSFTL
UYFR-2003406        VGAGKPTIDAG-----YEISAISRELDFINLMSYDLN----G-AWNNYTGINSPLYHREAEGEEERQLNVEWAANYWVEQGCPKEKLIIGLGTYGRGFTL
YRMA-2103934        IAPTQDRFKTS-----YEMPDIVQYLDFVNLMTYDMH----G-SWDNQVGHHSALYSPTDDKS----DSIDAVVNFILSTNVSSAKLNLGLPFYGRSLSL

                           310       320       330       340       350       360       370       380       390       400
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1_Class_V      TNANSHSYYAPT-----TGAAISPDGSIGYGQIRKFIVDNGA--TTVYNSTV---VGDY-CYAGT-----NWIGYDDNQSIVTKVRYAKQRGLLGYFSWH
ANS10044.1_Class_V  KDPNDHGVGAKA-----VGPAVDTDGSMDYDEILVFNKDTGA--KVVYDEVA---VSFY-SYSGT-----TWIGYDDGPSITKKVQFARSMGLKGYFFWA
AAM18075.1_Class_V  ANAAEHDMFSKA-----TGSCVWP---MDYSEVKSHRRARQCEYNIPNDPRF---GIAY-AYVET-----TWIGYESESTIPVKIKFAKQNQLLGYFAWN
YRMA-2017852        ASASGYHMGDKAPRAGTAGQFTREAGFIAYYEVCDLIKKGG---TVEFDSEG---RVPY-VHSGT-----QWVGYDNPDSLREKVRYVKQNQFGGVMVWA
UYFR-2003406        CDDSQTSIGSCAKAASSAGTYTREAGFLAYYEICTLLEQEST--TRVFDEEH---RAPYAIFSGNSWSNKQWIGYDDKQSLREKVNWLKEQNFGGWMTWC
YRMA-2103934        SDPMHAKLGDPVSGDGPGGQLLQADGIMAYYEVCKLLTEQGD--TISDSGRLSGTKAPY-VVDGA-----RWTGYDDEQSIREKVDYAIQHGLGGVFI--

                           410       420       430
                    ....|....|....|....|....|....|.
3AQU_1_Class_V      VGADDN-SGLSRAASQAWDATTATTRTIQKV
ANS10044.1_Class_V  IGKDKD-WTISKQASNAWGY-----------
AAM18075.1_Class_V  ISGDDDVNTLANAAQSNWQ------------
YRMA-2017852        LPLD---------------------------
UYFR-2003406        LDLD---------------------------
YRMA-2103934        -------------------------------
```

Figure S5. Class V chitinases alignments of complete sequences chitinases identified in green algae



```
                             10        20        30        40        50        60        70        80        90       100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      ----------------------------------MQTVVKASYWFPASEFPVTDIDSSLFTHLFCAFADL---NSQTNQVTVSSANQ------------P
ANS10044.1_Class_V  MAVQSIITTPLLVILMSLRSYAFTEPSLHRQQPPSKGGVRAAYWPAWSDFSTSSIDTNYFTHIYYAFVQP---APESFNLEITESYK-------------
AAM18075.1_Class_V  ------------------------------MSHPTIVNGGYWLSSTAGKLPASNIDATLYTHLFCGYGEI---DATSHEVTIPSTEV-------------
WSJO-2048613        -------------------------------------------------YTPSHLDGSLFTHILYAFAKI---DGNSYAVVDVEWDD------------E
DZPJ-2054002        -----------------------------------SSQVSRADWCGPYAYTPSHLDPSLYTHINYAFADF----DEAWQVKDVEHND------------E
QPDY-2031501        -------------------------------------------WCKAYAYTPSRIDPNLFTHVNYAFAKI---APSTYQVLDVEYND------------K
PZIF-2015162        -------------------------------------------WCDAYAHRPADIDPSLATHINYAFAKI---DSKTFEVVNVEANE------------D
BAZF-2007777        ------------------------------------------------------IDANVPTHVNYAFAKI---ESGTFLVANVEFND------------D
BAZF-2016149        ---------------------------------------------------PRDLDPLVPTHINYAFAAI---DGDTFQVVDADESD------------P
LNIL-2052393        ------------------------------------------------------LDASVATHINYAFAIV---DAATYQLRMLEYNA------------V
EEJO-2039453        ---------------------------------------------------PKNLNATLCTHINYSFAFM----NPDFSLKPIEYNDIPSP---DNPSMV
MNCB-2024465        ------------------------------------------------------MDPSLATHVIYAFMAV---DSSTFAITSYEWDD------------L
MNCB-2024466        ---------------------------------------------------PKDMNPSLATHVIYAFMAV---DGSTFKITNYEWDD------------L
MNCB-2024468        ------------------------------------------------------MDPCLATHVIYAFMAV---NGSTFAITNLDWND------------L
QJYX-2001962        -------------------------------------------------FTTAHVDPNLCTHINFAFGGI---ADGSYLVQPWQPQD------------T
MMKU-2006934        -------------------------------------------------FTPERIDPSLCTHINYAFAQV---QSTTWALAGYEWNDEGTN-------PP
ISIM-2003838        -------------------------------------------------FYPEDINPALCTHINYAFGVV---KQNTFTLGGYEWND-----------EE
UTRE-2015936        ------------------------------------------------YFDIQAINLTGVTHAVYAFAQI---DPTAYTLTWPYASF-------------
UTRE-2015935        ------------------------------------------------YFDFNTVNLTGVTHAVYAFASI---DPSSYKLTWEDTAI-------------
GFUR-2016925        -------------------------------------WSQNPD-------DMATVDLRGVTHATYAFATI---DGASFGLTEPYGSV-------------
IRYH-2007489        -----------------------------ITGLYIPGWAQNTNKASANYFDVNIMDLTGVTHVYYAFLWI---NDGTYTLYDDFGNY-------------
KUJU-2004605        -----LAALFVTLTAYGAAATVPTTAVDSITGVYIPGWAQNT-GSTTDFFNTSAVNLTGVSHMYYAFLWI---RAEDFTVYDPFDNL-------------
ISPU-2004540        -----------------------------IAGLYIAGWAHNT-GSSTDFFDISAVNLEGVSHMYYAYLSI----GSDNTIYDPYGNL-------------
ISPU-2004539        -----------------------------VTGIYIPGWAHNT-DSSTNFFNTTAMNLDGVSHIYYAFLWI---RSTDSTVYDPFGNL-------------
RNAT-2058594        --LKSLLAVCCFALAQRACAGVAQTLADSVTGLYIPGWAENT-WSTQNYFNTSAVNLEGVSHVYYAFLWI---NETTHKVYDPFNNL-------------
RYJX-2062142        --------LVSVALAQRVYAGVSFTTSDSVTGLYIPGWAENLDRASRNFFNTSAVNLQGVSHFYYAFLWI---RGTTYEVYDPFGNL-------------
WRSL-2006874        -------------------AAISATAPDSVTGLYIPGWAQDT-GSTLNFFNISAVNLDGVSHVFYAFLWI---RDTTYTVYDPFNNL-------------
BFIK-2029987        -------------------------------------------WYDATVVTPAVLAGGHYNYAFYAFAQV---DGTTYKVVPLSPDTTG---------SN
DRFX-2014366        ------------------------------------------------------INASLFTHVFYAFVPI----GSDYKVVTTDKTA-----------NA
SNOX-2004982        -------------------------------------------------FTPEDIKANLFTHVYYAFAGISLSSNKKYAVSPLDPEADLGS----SPSEG
KEYW-2026154        -------------------------------------------------FPAEKIDASLFTHVMYAFATI---STTDYTVVPSDPGVDLGP--AGSPEKG
HKZW-2012228        -------------------------------------------------FSPESIVTSHYTHLLYAFAAI---KASAFTIIPSDPAIDLGPNPTLYPDKG
YOXI-2002515        -------------------------TASKNMKRVAYFANWAYYRPAPYSFVPENINTAHFTHLNYAFASI---KARTFEVIPSDSYVDVT-------LN
VAZE-2009036        -------------------------------------------------FGPESIQTSYFTHLNYAFASV---DKKTFKVIPSDSHTDVD-------QN
JOJQ-2040355        -------------------------------------------------FYPENITAKHYSHVLYAFATL---HPKTFAVIPTDPELDLGPDTKNQPQKG
NBYP-2006024        ----------------------------------------VSWRPSPYDFFPENITTKHFTHLLYAFATL---DSKTFAVIPTSPEVDLGPDPIKQPHEG
MFZO-2022650        -------------------------------------------------FYPENITTQHYTHLNYAFATL---EPQTFAVIPTSPEVDLGPDPVKQPQKG
MOYY-2018456        ---------------------------------------RGTYYAGWSGIPPSNINATLFTHIFYAFAAI---DADRSKVIPDDPSSEAGP-------SG
MCHJ-2001441        ---------------------------------------RGTYFASWTGQDPAKINATVFTHVFYAFGQI---DPVKFKLIKGDPSE-------------
WCQU-2007350        ---------------------------------------------------PLKINATVYTHVFYSFATV---SATSSKLTQRDPSE-----------AG
MCHJ-2000297        --------------------------------------VRGTYFPNWSGISPTKINGTLYTHVFYAFAGI---NSANQLAGPPSLQQ-------------
XRTZ-2004862        -------------------------------------------------FMPEDINASLYTHVLYAFAAI---DTAKLSLAATEDDE-TGP-------SG
RQFE-2000106        ------------------------------------------------------VNATVYTHILYSFANI---DTTTFALIPQDPNTEPGT-------NG
QWFV-2003029        ---------------------------------------------------PTLIPSGLFTHVLYAFAVV---DATAFTIAPSSPSTEGGP-------SG
QWFV-2003026        ---------------------------------------------------PTLIPSGLFTHVLYAFAVV---DATAFTIAPSSPSTEGGP-------SG
QWFV-2003030        ---------------------------------------------------PTLIPSGLFTHVLYAFAVV---DATAFTIAPSSPSTEGGP-------SG
HIDG-2009902        --------------------------------------------------PPATVDAKLFTHLFFAFATI---HASNFGIIPLYPANAAGP-------TG
HIDG-2009903        --------------------------------------------------PPATVDAKLFTHLFFAFATI---HASNFGIIPLYPANAAGP-------TG
RPQV-2036136        ---------------------------------------------------PTAINSTLFTHIVYAFAAI---DRRTNTVVPGDPATEAGP-------TG
AEKF-2009302        ------------------------------------------------------LDASPFTHVTYAFALI---NDATNEVMAKEPGRETGL-------NG

                            110       120       130       140       150       160       170       180       190       200
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      KFSTFTQTVQRRNP-----SVKTLLSIGGGIADK------------TAYASMASNPTSRKSFIDSSIRVARSYGFHGLDLDWEYP------SSA------
ANS10044.1_Class_V  KWAPKYDGIHNIRP-----RVTTLLSIGGGGNNA------------TLFSEMASSKQNRASFINSTIHVARKHEFNGLDLDWEWP------GDE------
AAM18075.1_Class_V  --SHFSEKVKVSNP-----SVKTLLSIRGDS---------------KVFSDIAKVPTARDTFVNSSIKAAVDCGFDGLDLQWLYP------SSP------
WSJO-2048613        QLQQQLQAVKAGAGG----GVRTLISIGGWSFST-GTDVFSGQGTELVFPRMASTQANRAAFITSAIAYAQSRGFDGIDIDWEWP-------GAIGGVVC
DZPJ-2054002        ELHKQLQALKQSHT-----GLKTLVSIGGWSFST-STGIYANTDAKDYFPTMASTASGRAAFISSVIDYAKQRGFDGIDIDWEYPNHCDLTGNGTYPCGR
QPDY-2031501        ALQAELNALKARNG-----GLKTLVSIGGWSFSR-GTEVFTGTGSEAIFPAMAASAATRQLFIQSVINYAKTRGFDGIDLDWEYPCYDKKAPGCA-----
PZIF-2015162        ELIAGLQGLKAKNP-----ALKTLISIGGWSFSR-GDEVFKGSGSEKVFPEVAASAANRAAFIRSAIDYTNKRGFDGIDIDWEFP--NFDGKAP------
BAZF-2007777        ELIAQLQALKQQNG-----NLKTLISIGSRGTEA-----FAGSGSEAIFPAMVATASSRATFIRSAIDYARNKGFDGIDLDWEFP--NYEGSAP------
BAZF-2016149        ALQRELQALKKQNR-----KLRTLISIGGWSFSR-GED---SAGTEKIFPTMASTAENRAKFIQSMIDFCRKRGFDGVDLDWEYP--NYDNSNP------
LNIL-2052393        ELMGRLQALKKINP-----ALKTLISVGGWTFSR-AEEGFAGTGSDRIFPTLAASAETRAQFVREAVAFVEHWGFDGLDLDWEFP--NHDGKSP------
EEJO-2039453        SLYEETMALKQQNP-----NLKVLIAVGGWSMNE-PWS-----PYKDLFAS-AVKPANRATFIAKIIQFLEQWGFDGFDIDWEYPGDLNRGGTL------
MNCB-2024465        TMIQQMVNLKARNP-----ALKTMISVGGWTFST-DPR------YSSIFPMMAQS---RKAFINSAIQFCYANGVDGIDIDWEYP-------AA------
MNCB-2024466        TLIQQMVNLKATNP-----ALKTLVSVGGWTFST-DPK------YSSIFPTMAKTASSRKAFIDSAIQFCYANGVDGIDIDWEYP-------AA------
MNCB-2024468        ALIQQMVNLKTKNP-----ALKTLVSVGGWTFST-DAK------YSSIFPTMAQSASNRKAFIDSAIQFCYANGVDGIDIDWENQ-------AA------
QJYX-2001962        QYYQQLTALRDSVPGGREHGPKIMMSLGGWTFNA-NWS------TKHIFSDMVSSAENRAKFVAGAIAFLRQWDLDGLDLDWEYPAAEWRGGAP------
MMKU-2006934        GMYLRLNALKSKNT-----GLKTLLSVGGWSFND------PARGTDRIFSDMASSAANRASFIQQAIVYLRKWGFDGLDIDWEYP--TLRGGVA------
ISIM-2003838        MYGRLNAVKAGTGG-----RLKSLLSLGGWTFNVQGGE------YCDVFSTMVSSAASRATFAANAISYLRQHGFEGLDLDWEYPGDTSRCGRS------
UTRE-2015936        ---SLFVGLKTRYP-----SLKTLISVGGGGFST------------SIWSAMAASPNNRAAFVASVVAALQRNNLDGVDLDWEFP-------GV------
UTRE-2015935        ---SLMPGLKTRYP-----SLKTLISVGGGGFST------------SIWSAMAASTTNRAAFVASVVSALQQYNADGVDIDWEMP-------GV------
GFUR-2016925        ---REFFALRNRYP-----ALKLMVSIGGGGYSP------------QTWTRMSSSPANRAKFVSAVLQFVLKWDADGVDIDWEFP-------GP------
IRYH-2007489        ---GLIMGLKNKWP-----GLKVLPSIGGGGFST------------AIWKGLSSTAVRRQAFVTEALSWMTRVGGDGIDLDWE-----------------
KUJU-2004605        ---PLLTELKSRWP-----ATKLILSVGGGGFST------------AIWSSAATS--GRAAFVNSALAAVHAAGADGIDLDWEFP-------AA------
ISPU-2004540        ---PLLTGLKSRWP-----ATKLVLSIGGGGLST------------SVWTAAATA--GLSPFVESVMAVMKDANADGVDLAWEFP-------TA------
ISPU-2004539        ---PLLTELKSRWP-----ATKLMLSIGGGGFSI------------SIWTTAATT--RLSAFVNSAVSVMKAANADGIDLDWEFP-------AA------
RNAT-2058594        ---HLLVSLKSRWP-----ATKLLLSVGGGGFST------------TIWKRAVTN--SLTTLVDSTIAAMKAANADGIDLDWEFP-------VA------
RYJX-2062142        ---KLVANLKTRWP-----NSKFLLSIGGGGFNL------------NIWARAATV--GLDAFINSIVRTMKETNADGIDIDWEGP-------DS------
WRSL-2006874        ---NLVVALKSRYP-----ATKLLLSIGGGGFDT------------NIWTKAATT--GLTAFVNNLVQAMKAASADGIDLDWEFP-------VL------
BFIK-2029987        GYSAFTSYAVSH-------SVKPFISIGGGAFSY-DLT------TSSIFPALAASATRRASFATSAVAFARQYGFLGIDLDWE---------NN------
DRFX-2014366        NYKAFTSS--SPSP-----PPPGLLSIGGGDSTS-----------AARFPLAAQTAARRKAFASSAVTFARAYGFSGIDIDWEYP------SDS------
SNOX-2004982        LYFRFNRTLRAKNP-----KIKTLISIGGGDDAL-----------AKKFTALSSSAATRTYFAKSAVQFARKYGFDGLDLDWEFP------QTR------
KEYW-2026154        LIFRFNLQVRKTNP-----TAKTLISIGGASSSK-------------YFSPMASQPSRQTAFVRSAIAFARTHSFDGIDLDWEPI------DNP------
HKZW-2012228        FYFRFNRVAHAG-------KLKTLLSIGGGDSAA-----------TAAFSLVSKTSADRTKFTANTVAYARKYNFDGLDLDWELP------DNP------
YOXI-2002515        FMKRFNKVAHKG-------AIKTLISIGGGESAA-----------TASFAAMAVANGGPAKFISSAISFARTYGFDGIDLDWELP------NTA------
VAZE-2009036        YMKRFNKVAHAG-------GIKTLISIGGGDSDA-----------TAAFSAMAAANGGPAKFIASAISFARQYGFDGIDLDWELP------NTA------
JOJQ-2040355        FLYRFNKVVHAG-------RAKTLISIGGGSGPA----------LSKLFSTMAKSNTTRLAFVNSAIQFARKYGFDGIDMDWEIP------STA------
NBYP-2006024        FYYRFNKRAHAG-------GVRTLISIGGGGGAA----------LSNRFSNMAKENTSRLAFITSAIKFAREYEFDGLDIDWETP------ETP------
MFZO-2022650        FYYRFNKKAHAG-------GVKTLISIGGGGGPE----------LSTLFSSMAKKDSTRLAFITSAIKFARKYGFDGLDIDWELP------KTP------
MOYY-2018456        LWARFVKSVKKSNP-----AVKAVLSVGGASDTA-----------AANFSAVARSASRRSAFAASAIQVLRKYKFDGIDLDWEFP---------------
MCHJ-2001441        CWGCFTKGVGNSNA-----AAKAILSVGGGTDSA-----------AQSFATMARTAAGRTAFINNAIKALRAFSFQGIDLDWEFP------DTP------
WCQU-2007350        SWKQFVAGVKKSNP-----QAKALLSLGGGEAEA-----------AKAIANMALDNSKRATFVGSAVKVLRDYGFDGLDIDWELP------SNA------
MCHJ-2000297        ---NF--------------KGKPILSVGGGDQAA-----------TNAFATMASSTTGINTFVSSATSNLRTYGFAGLDLDWEFP------STP------
```

```
XRTZ-2004862        LYARFVKHIKTTNP-----AAKPLLSLGGGDSVA-----------AAAFASVSSTNAKRVTFARNAVAFARQFGFEGIDIDWEFP------NGA------
RQFE-2000106        LWSRFTKAIKASNP-----TCKAILSVGGGGGSY---------DAAAPFAIMSRTKAGRDAFAANAVALLRQYGFDGLDLDWEFP------SNA------
QWFV-2003029        VWTKFSSRVHSV-------KAKAILSLGGGSPGV-----------AQNFSAVASSASRRALFAANAVRVLRAYSFDGLDVDWEFP------DNA------
QWFV-2003026        VWTKFSSRVHSV-------KAKAILSLGGGSPGV-----------AQNFSAVASSASRRALFAANAVRVLRAYSFDGLDVDWEFP------DNA------
QWFV-2003030        VWTKFSSRVHSV-------KAKAILSLGGGSPGV-----------AQNFSAVASSASRRALFAANAVRVLRAYSFDGLDVDWEFP------DNA------
HIDG-2009902        AWVQFNAAVKKSNP-----SVKTLLSIGGGEASA-----------TAAFVAMASAPARRQAFVNSAVKVAREHNFDGIDLDWEFP------KNA------
HIDG-2009903        AWVQFNAAVKKSNP-----SVKTLLSIGGGEASA-----------TAAFVAMASAPARRQAFVNSAVKVAREHNFDGIDLDWEFP------KNA------
RPQV-2036136        KWAAFTAGVKQRNR-----KVKALLSIGGGDPNA-----------TAAFRSMASSAARRQAFISSGISTARSYGFDGLDLDWELP------NGP------
AEKF-2009302        LYFRFTQAVKQSNP-----EAKTLLSIGGAGEAV-----------AAAFSRMSSSATARRSFIASCIKMARTYGFDGLDVDWEYP------KSP------

                           210       220       230       240       250       260       270       280       290       300
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      ------------TEMTNFGTLLREWRSAVVAEASSS--GKPRLLLAAAVFYSNNYYSV------LY-PVSAVASSLDWVNLMAYDFY-GPGWS-------
ANS10044.1_Class_V  ------------KDMSNLALLLKEWYKALVVEANTS--RKSRLLLTSAV-YFNSTISLIGNGPRSY-PVRAIRKYLDWASPMCFDYN-GA-WA-------
AAM18075.1_Class_V  ------------TDMQNFKTLIVKWRQAAVDDARNN--NRPERILVASVPNLPYIEKNV-----EY-PGEAIKQNLDWVNLVCYDFY-TP-AEG---SSP
WSJO-2048613        TDASSGGCAACTQSIADFSSLLQELRAA----------MGPSMILTTALRA-TPTPEI------HY-NLKDVHPSVDFINLMTYDYHGGS-FGA---SEG
DZPJ-2054002        D-----------QEVADFTQLLRELSDA----------ARPRLLVSAALRA-SPNPQR------HY-QLSEIHQYLDFINLMAYDFHGGS-FE------T
QPDY-2031501        ------------GEKLAFTSLVQELRNA----------MPQGMLLTAALAA-SPTPEQ------HY-EVGKVSSVLDMVNLMSYDFHGGS-FE------M
PZIF-2015162        ------------QEVADFTSLMAELGTAT---------KAAGMLLTTAMRATPDPQQH------A--NVKAVAAAVDFINLMTYDYH-GA-FDFG--SGY
BAZF-2007777        ------------QEIDQFSALMSELWDAVNQESQSS--GRTRLLVTAALRASPTPAD-------HV-NIKAIMDKVDWVNLMTYDFH-GA-FDVPSGGGR
BAZF-2016149        ------------REVDDFTILLAEARAAFNKEACA---GRRRLLLTAAIPADPRT---------SF-NLTAIKASLDFINLMSYDFA-GP-WSLT--DGQ
LNIL-2052393        ------------KEAEDFVLLLQDLRAAT---------SARGLLLTAAVRGFDAAGPIP--------DAGVVTDLLDFVNLMSYDYH-GS-FDLPA-SGV
EEJO-2039453        ------------QDKANLAAFCKEFKEEVAR-------RGKSYLLTMATGAGPSGYQ-------GL-DLPQLGQSLDFINIMTYDFH-GA-WD-------
MNCB-2024465        ------------TDTVNYSNLLMEFRAAITTAASA---GKPQLLLSHAAPAGPQNYQN------T--NLGLVHQYLDFINLMAYDFH-GA-FE-----TN
MNCB-2024466        ------------TDTDNYSNLLKEFRAAITTAASA---GRQQLLLSHAAPAGPQNYQN------T--NLGVVHQYLDFINLMAYDFH-GA-FE-----TN
MNCB-2024468        ------------ADAVNYSNLLKEFRAAITAAASA---GKPQLLLSQAAPAYLDT------------NLAPVHQYLDFINLMTYDFH-GA-FE-----TN
QJYX-2001962        ------------TDKANFVLLLEELLLAFRDEAVAT--QRPRLQLTLAVGIGKPTADA------GY-DVLGIAGAADFVNLMTYDMH-GA-WE-------
MMKU-2006934        ------------ADRANFSLLLQETLAAFKSETRAP--GVPRLLLTIAVGVGKDTADA------AY-QASGVHPFVDWINLMTYDLH-GA-WE-------
ISIM-2003838        ------------ADRANFTLLLRELRAAFEEEAAAS--GRERLLLTIAAGVGKGTADA------AY-EYEQIHESLDWINLMTYDLH-GG-WE-------
UTRE-2015936        ------------ADKVAFISLLQALRTAFTAMGR----ATPPLITAAVSGIVTGGVMD------GL-NLPAAAQYLDMWNVMTYEMHDPCYWE-------
UTRE-2015935        ------------SDKASYVALLQALRTALTGMKR----SSPGLLTAAVPGVVTGGVMN------GF-NMPAVTQYLDMWNVMTYDMHDPCYWE-------
GFUR-2016925        ------------SDRDNFILLLRELRTAL---------GLGRLLTAAVPATSGSPSDEDAM---GM-DIPTAAGILDFINVMAYDLHDPCYWE-------
IRYH-2007489        --------------------LLKEFRTQLDALPRAA--GAPKYLLTAALTSLPTDAA-------GL-DVPGMNMYMDLFNVMTYDMHDPCYWE-------
KUJU-2004605        ------------SDRSAFTALCSTLRSRLDAEGAAAQPPRRYWLTAATQSIAAAGFFD------GY-NLPALAGYLDLFNIMTYDMHDPCYWE-------
ISPU-2004540        ------------ADRASFTRMAATLRTRFDEEEAST--GRHYWLTAVV----APGPSN------TY-DLTALQGSLDLFNIKTYDMHDPCYGE-------
ISPU-2004539        ------------ADRATFTNLAATLRARFDKEEASS--GRHYWLTAATQSIVATGTFD------GY-DLVALKNSLDLFNIMTYDMHDPCYWE-------
RNAT-2058594        ------------ATRTNFTRLCAALRARLDAEGAAASPGRHYLLTAATQSIVSSDSFA------GY-DLAVLRNHIDLFNIMTYTMHDPCYWE-------
RYJX-2062142        ------------DTRAAYTDMCARLRARLDAEGASASPPRKYWLTAATGAIVTPAAIG------GF-DLPALKNTVDLFNIMTYELHDPCYWE-------
WRSL-2006874        ------------ATRSVFTQLCATLRARLDAEGAAATPPRNYVLTAATQAIVSTGSFA------GY-DLTALKSYLDLFNIMTYDMHDPCYWE-------
BFIK-2029987        ------------VVAADFTKLMQALRTAITSDAAAS--GKAPLLLSFAAQVD----IT------LY-QVSVVSS-VDFINLMTYDFH-GS-WEC---TPN
DRFX-2014366        ------------TQAQNHGLLLQALRAVVDADGTAS--GKPKLLLTAAVSG----IQS------AY-VASAVIAAVDFIGVMAYDLH-GS-WE-------
SNOX-2004982        ------------SDKRNFVLLLRDIRTEIAKESRANKRGQSRLLLTCAVSSYKPQIET------SY-DVTGIAKYVDWVGVMCYEFH-GE-WE-------
KEYW-2026154        ------------TDKANFETLLKLLRAEVTKEAAGSV-GRVPLLLTAAVAAWKPSVVA------SY-NVPVMNQVLDLVGVMAYDLH-GP-WE-------
HKZW-2012228        ------------TDKTNFALLLANFRATITANAKKL--NKTALLLTAAVAAWKPNVVA------SY-DVPKLNKYLDWANVMTYDFH-GS-WE-------
YOXI-2002515        ------------TLKAHFSDLLINFHSAIAAESTRT--KRTALLLTAAVAAWSPDVDL------SY-NVPVLNKYLNWANVMCYDFH-GS-WE-------
VAZE-2009036        ------------TKKAQFSALLASFYSSMVAEAKLT--KRTRLLLTAAVAAYAPDVDL------SY-NVPVLNKYLDWANVMTYDLH-GS-WE-------
JOJQ-2040355        ------------ADKANFVSLLRTFRSQISADRQKT--KKAALLLTAAVAAWKPDVLK------AY-DVPSLNKYLDWANVMTYDFH-GS-WD-------
NBYP-2006024        ------------QDRSNFASLLIMFRSQITAESRRS--KKSALLLTAAVVAWKPHVIT------AY-VVPTLNKYLDWANVMAYDLH-GS-WD-------
MFZO-2022650        ------------QDKANFAALLVSFRSKITAEAKSS--KKAALLLTAAVSAYKPDIET------AY-VVPTLNSQLDWANVMTYDFH-GA-WE-------
MOYY-2018456        ------------ENKSDFRALVLSLRSAINAEARRT--RKPRLLLTAAVSADKPVVEA------SY-DVGNLNSFLDLVLVMCYDFH-GS-WD-------
MCHJ-2001441        ------------GNKSNFAALLTAFRWAIAREAKTS--SRPPLLLTAAVSAYGPAIAA------AY-DIKALSANLDLALVMCYDFH-GE-WE-------
WCQU-2007350        ------------QQRAALSKLLQDLRSAFTAEARRT--GKPPLLLSAAVSAAKGDICK------SY-DVAALNQALDWALVMCYDFH-GE-WE-------
MCHJ-2000297        ------------TTRANFIKLLKALRAGFNSEAQTS--KKTPLLLTAAVSA-GGSATS------YY-DFPTLNANLDLVVLMCYDFH-GAGWE------P
XRTZ-2004862        ------------SQRANFMLLASSLRSAIQADATSS--GRPPLLLSAAVAVLADDIKS------YYGDVGSLGKLLDWVGVMAYELH-GT-WE-------
RQFE-2000106        ------------TEKAQFTALMLALRTAFDADVSSS--AKPKLLLSAAVTPVNYQVAT------SY-NITACNAYLAWVFIMCYEFH-GS-WD-------
QWFV-2003029        ------------KDLAYCSAFLTALRSAFNTEKLST--GQQRLLLSAAVSAVPSTIQA------AY-NISTLNQALDWANLMCYDFH-GS-WE-------
QWFV-2003026        ------------KDLAYCSAFLTALRSAFNTEKLST--GQQRLLLSAAVSAVPSTIQA------AY-NISTLNQALDWANLMCYDFH-GS-WE-------
QWFV-2003030        ------------KDLAYCSAFLTALRSAFNTEKLST--GQQRLLLSAAVSAVPSTIQA------AY-NISTLNQALDWANLMCYDFH-GS-WE-------
HIDG-2009902        ------------QEKKNFTELIVSLRRAFAGEKR----GKRAVLLLTAAGGYRPTITT------AY-DIAALKTNLDFAVVMCYDYH-GS-WE-------
HIDG-2009903        ------------QEKKNFTELIVSLRRAFAGEKR----GKRAVLLLTAAGGYRPTITT------AY-DIAALKTNLDFAVVMCYDYH-GS-WE-------
RPQV-2036136        ------------GDRTNFVLLLGSFRAAILQEAKSS--GRPALLLSAAVSAYGPDIAL------SY-VIPAVDANLDWVDVMCYDFH-GS-WD-------
AEKF-2009302        ------------TDKVNFATLMRELRGAIDSEAS----GAESLLLSAAVASWEPRIAE------GY-DVPALSATVDWLNVMCYDFC-GS-WD-------

                           310       320       330       340       350       360       370       380       390       400
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      RVTGPPAALF-DPSNA-----GPSGDAGTRSWIQAG-LPAKKAVLGFPYYGYAWRLTNANSHS-YY----AP---TTGA--AISPD-------GSIGYGQ
ANS10044.1_Class_V  NETGFNAALY-DPNS------NISTKYGIGSWIGSG-VPAEKLVMGLPLYGRAWELKDPNDHG-------------VGA--KAVGPAVDTD--GSMDYDE
AAM18075.1_Class_V  DYTGPSSALF-NRRSN-----SLSAHFGIESWTKY--VPAKKIVFGIPFHGWAWTLANAAEHD-MF----SK---ATGS-------CV-----WPMDYSE
WSJO-2048613        GVANVNTPVV-DCTARTP---NFDIAGALDAYLAAG-VPSSKILLGLANYGRTFALTGPGDTP-GV----VS---ADGAPHAQQSTCTYED--GVLSYYE
DZPJ-2054002        SPANVHTPVL-DCTRQAD---KYDISGAVAAYLADG-VPAEKINLGLGTYGRTFALTSTDATP-GS----AP---ANGA--APRGECTLEP--GTISWYE
QPDY-2031501        TPAHVHTPVL-DCRSTK----HFDISGAVEAYLAQG-VDPRKVNLGLGTYGRTFRRTSSDSTP-GS----AP---ANGA--SPKGACSQTE--GFLALYE
PZIF-2015162        RKVNTHTPVL-SCVNPLKK--GWDIAGAIDAYEKQG-APLNKLVFGLGTYGRGFRFAGGEMRPGLG----SALNSDEGQ--PLQGECTQTK--YTLAWYE
BAZF-2007777        GKLGHNAPVV-DCKSP-N---NFDIRGAIDGYIEAG-LDMAKATIGLATYGRGVTTTPGDANP-GD----EV---GTGP--NTQGVCTQTP--GSLAWYE
BAZF-2016149        GPLALNAPML-DCAG------GFDIRSSIDLYLAAG-VSPHQINQGLATYGRGFATTPGSSIL-PG---AAS---AVGV--GVAGECTQEE--GSLAWFE
LNIL-2052393        GIVNSHTPVL-DCSPSGKGK-GKDMVSALEDYAALG-ADLSKFNLGLASYGRSFTLAVPGGQ--FGELRPGH---AHGVAPGSRGDCTMAP--GVLAFYE
EEJO-2039453        TAVNFHTPWK-DPLGG-----TLDIQSALNHYLIAGQVPPSKLNLGIGAYGRSWTLADRTKST-LG----SA---AYSP--GAAGSCTGEA--GYIGWFE
MNCB-2024465        VAPKPHAILK-DCSMS-----NWDIDSAVQAYLRAG-TPAAKINLGLATYGRAFQLVNGL---------------VAGA--APAGKCTLEA--GMLAAYE
MNCB-2024466        VKPKPHAILK-DCSTT-----NWDIDSAVQAYLRAG-TPAAKI--GLATYGRSYQLVNGL---------------VAGA--APAGMCTLEA--GLLATYE
MNCB-2024468        VKPKPHAILK-DCSTT-----NWDIDSAVQAYLRAG-TPAAKI--GLATYGRSYQLVNGL---------------VAGA--APAGMCTLEA--GLLATYE
QJYX-2001962        TRTGHNAPLD-DPGNGGGAGYPLGVKWAVQYWLDRG-CPAYKLVLGMAAYGRSFKLANAADTA-VG----SPTAWAGAG--GEPGPVTGEA--GFLAYHE
MMKU-2006934        STTSFNAPLKWVAGDFQSY--PLSVEWAVSYWLGQG-VPGSKLTLGMASYGRGFQLASSSNHA-VN----AP---AVGG--SQAGQYTQEV--GFMAYYE
ISIM-2003838        SATGLNAPLDYVPGDLPGDPYPLSVRWAANYWIGKG-TPASKLVLGMGIYGRGFTLSDPGSYG-IG----AS---ASGA--SLPGPYTREA--GFLASYE
UTRE-2015936        TQTACHTDWP-------------DCVSAINFYTGQG-VPAIKLNLGLAFYGHGYTLLSASQYA-------FP----SPS--VELKDCSNEI---TPSYRQ
UTRE-2015935        TQTAFHTDWT-------------DCVAAITWYTAQG-VPAANINLGLAFYGHGYTLLNASHFT-------YP----SPS--VELKDCSNQI---EPSNRQ
GFUR-2016925        SSTAFHTSWS-------------DITRSLDYFASQN-VPPHKILLGLAFYGHQYNLTNKKQYA-------YP----SPS--VMSKDCSTMI---TRDYRD
IRYH-2007489        DHTNFHVAWS-------------SAAVGLDYYVGRG-VPASKIILGLAFYGHTYTLLNAAVYQ-------HP----APS--VELRDCSKMN---TPSYRA
KUJU-2004605        TRTAFHTAWT-------------DCVRALDFYVANG-VPRSQLVLGLAFYGHVYTLIDAAVYS-------HP----APS--VESRDCSTMN---TPSYRA
ISPU-2004540        TRTNFHSAWS-------------EVRRALEFYISKG-VPNSQLVVGLALYGHTYTLIDPANCV-------FP----APS--VASKDCSTAE---KVSYRN
ISPU-2004539        TTTNFHAPWS-------------GAATALDFYISQG-VPRSQIVLGLAFYGHTYTLIDPANYG-------YL----APS--VESRNCSTME---TASYRS
RNAT-2058594        TETHFHTAWS-------------ECVTVLNYYIKAG-VNRTQLVLGLAFYGHEYKLLYPSRYS-------YP---GAPS--VEGRDCSIMT---TRSYRE
RYJX-2062142        THTHFHTSWT-------------DSVKALNYYIGKG-VSRLQLVLGLAFYGHQYTLLDASSYT-------YP----AAS--VELRQCGRQV---TPSYRD
WRSL-2006874        TETHFHTAWS-------------ECVTALDYYVQQG-VPRTQLVLGLAFYGHVYTLLNPEQYG--------P---SSPS--VESRDCSTMR---TVPYRG
BFIK-2029987        CATAELAQLQ-NANEA-----SNTVAGYVQAWLTAG-VPPAKLIVGLAAYGNSYTL-STSARG-VG----AP---ASSF--G-----------ASLGYKD
DRFX-2014366        AVTGLHTALD-DSASP-----DLSVKGGVNAWMNLG-MPASKIVLGLAAYGRSWTLKSATSNG-VG----AA---ASGP--GAAGPTSQEA--GVLSYAE
SNOX-2004982        TVTGFHTALS-DPTYP-----TLSTKGAASAWIDMG-LPLSKLTLGLASYGHTWTLRDKTKHG-VG----AS---AVGV--GKAGKYTQEA--GTLSYFE
KEYW-2026154        TTTGLHTALE-DATNP-----DLSVKGAVSAWLGLG-LSPTKLVLGLASYGHSWTLANPQVHG-VN----AP---APKP--GDPAPYTQES--GYIAYFE
HKZW-2012228        KKTGQHSAFT-DSVDP-----SLSIKGAVQSWLALG-LSRRKLVVGLAAYGRSWTLKTGSIHG-EG----AA---AVGV--GYGGPITQDSSGGSWAYFE
YOXI-2002515        SVTGEHTALA-DPSDA-----RVSIEGAVAAWLKRG-LSAKKLVVGLATYGHSWTLKQATVHG-VG----AA---TSAV--GAAGPITQDKSGGTWAYFE
VAZE-2009036        STTGEHTALA-DPSDQ-----NVSIEGAVAKWLSRG-LAANKLVVGLATYGHSWTLKTATSHG-VG----AA---ATGV--GYEGPITKDTSERSWAYYE
JOJQ-2040355        DVTGEHTALA-DPSDS-----SVSTKGAVKAWLDLG-LSRKKLVVGLAAYGHSWTLANGKKTG-VG----AP---TTGK--GYAGPITKDTEGGSWAFFE
NBYP-2006024        KSTGAHTALE-DPAAP-----ALSIKGAVKAWIELG-LSPKKLVVGMAAYGRSWTLASQQKRG-VG----AP---ATGV--GYAGPVVKDMEAGTWAYFE
MFZO-2022650        QTTGQHTALA-DPANP-----DASIKGAVKAWIHLG-LSPKKLVVGLAAYGHSWTLANQQQRW-VG----AP---AKSV--GYGGPITRDAEGGTWAYFE
MOYY-2018456        DVTGMHTALF-DAQDP-----DVSCTGALQAWTAAG-MSPSKILMGLAAYGKTWTLKSRREHG-VG----AR---ATGP--GVAGKTTQEP--GTLSYAE
```

```
MCHJ-2001441        SVTGMHSALY-DPNDP-----AVSCVGGVQAWTQGG-FAPAKLLMGLASYGKSWTLANRNKHG-VG----AK---ATGP--GAAGKFTQEP--GTLSFAE
WCQU-2007350        DATGIHSALY-DAAFP-----DQSCSAALQTWISGG-LDPKKVIIGLASYGHFWTLQSAKSNG-VG----VP---ATDP-----------R--GIWSYPQ
MCHJ-2000297        TITGLHTALY-DTSNP-----TASCSQAVQTFTAGG-VAASKLVMGLASYGNSWTLKYTTQTG-VGAWVGAA---ATGP--GAAGAITRTA--GILSYPE
XRTZ-2004862        SQTGEHTALQ-DKNDP-----PLSIAGAIDAWKSAG-MPANKILAGLAFYGHTWTLKSPALHG-VG----AP---ATGA--GAVGAVSQDP--GTLAYFE
RQFE-2000106        TTTGAHTALE-DLAYP-----AASCKGAISYWASGG-LSASKTVMGLASYGVGFTLTNPNITG-MG----AP---AKGG--------------ADMSYAA
QWFV-2003029        SVTGEHTALY-DPSDP-----GLSCVGGISVWRRGG-MAASKILMGLAAYGRTWTLVSAGSHG-IG----AP---ASGP--GAPGSQTNAP--GIFSYGE
QWFV-2003026        SVTGEHTALY-DPSDP-----GLSCVGGISVWRRGG-MAASKILMGLAAYGRTWTLVSAGSHG-IG----AA---ASGP--GAPGLQTNAP--GVFSYGE
QWFV-2003030        SVTGEHTALY-DPSDP-----GLSCVGGISVWRRGG-MAASKILMGLAAYGRTWTLVSAGSHG-IG----AA---ASGP--GAPGLQTNAP--GVFSYGE
HIDG-2009902        QHTGLHAALE-DPANP-----DASCSGAVQALTGGG-MPASKLLMGLAAYGRSWTLKDAADNG-VG----AP---TTGV--GLPGPQTKDP--GLLSYAE
HIDG-2009903        QHTGLHAALE-DPANP-----DASCSGAVQALTGGG-MPASKLLMGLAAYGRSWTLKDAADNG-VG----AP---TTGV--GLPGPQTEDP--GLLSYAE
RPQV-2036136        DVTGEPTALA-DPSNV-----DTSCTGALQAWLGGG-LNPSRFVVGLASYGATWTLASARRYG-IG----AK---TVGP--GTPGPATQEA--GILSYGE
AEKF-2009302        KITGPHTALE-NPARP-----DLSCKNSIRAWLDRG-LAASKCVLGLAAYGRSWTLDSEWQTG-IG----AP---ASQP--GSPGPLSQEP--GYLAYFE

                            410       420       430       440       450       460       470       480       490
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|..
3AQU_1|Class_V      IRKFI--VDNGA----TTVYNSTV-VGDYCY-AGT-NWIGYDDNQSIVTKVRYAKQR-GLLGYFSWHVGADDNSGLSRAASQAWDATTATTRTIQKV
ANS10044.1_Class_V  ILVFN--KDTGA----KVVYDEVA-VSFYSY-SGT-TWIGYDDGPSITKKVQFARSM-GLKGYFFWAIGKDKDWTISKQASNAWGY-----------
AAM18075.1_Class_V  VKSHRRARQCEY----NIPNDPRF-GIAYAY-VET-TWIGYESESTIPVKIKFAKQN-QLLGYFAWNISGDDDVNTLANAAQSNWQ-----------
WSJO-2048613        IERFV----DSG----AIKVDPVS-MAAYAR-YNESSWVGFDNAETHRRKMCYARYR-QLGGVFLW-------------------------------
DZPJ-2054002        LKRKL----APD----SIKVDSKSLLAAYATYDDGQSWAGFDNQESHRLKMCWARQQ-KLGGVMAW-------------------------------
QPDY-2031501        INRLV----NKV----DIREDTASLMAAWAP-YGADGWVGFDTENTHRLKTCYARYK-RLGGIMVW-------------------------------
PZIF-2015162        IKKVV------A----KSIIDPVQ-MAAYGV-TTSGHWVGFDTPETHRMKMCYARAR-GISGVMVWDSDTD--------------------------
BAZF-2007777        IKQLV----PND----AIKVDSVA-MAAYASYKSGTEWVSFDTEETLRLKLCYARSR-GITGAMVWDTNMD--------------------------
BAZF-2016149        IKRKV----PTA----DIRIDEKA-LAAYASFDGGRQWIGFDNVESHRLKMCYARSL-GLGGVMVWEASLD--------------------------
LNIL-2052393        IQRAI--QAEGL----AVREDKHH-MAAYVVMRGGREWVGYDTPSTMQQKVCWARAM-GVGGVFLWTADED--------------------------
EEJO-2039453        IKRLM---EQGA----QVTIDEGA-MAAYLV-SGS-TWVAFDLPQTIYMKILAARQL-GLGGLMVW-------------------------------
MNCB-2024465        IKDTL----GSI----QPSVNTLW-ASAYAV-YGSSYWVGYDTEATLGLKVCYARSY-NLGG-----------------------------------
MNCB-2024466        IKNAL----GSI----QPSVDTVW-ASAYAV-YGSSYWVGFDTEATLALKVCYAKYY-KLGG-----------------------------------
MNCB-2024468        IKNAL----GSI----QPSVDTVW-ASAYAV-YGSSYWVGFDTEATLALKVCYAKYY-KLGG-----------------------------------
QJYX-2001962        IRAIL--ERDDP----YVYRDAPY-GVPHFV-YGDGNWVSYDDVASIAAKTDFLREK-ALRGGMVWALDLDA-------------------------
MMKU-2006934        IAAKV--AAGTA----TAVYDVAR-EAPYAY-WSN-QWVGYDNAQSIQRKVDLANQN-GLLGAMLWALDLD--------------------------
ISIM-2003838        VAALV--ASGAA----TRVFDEAR-GAPYAH-WGD-QWVGYDDEESLARKAELVRAL-GLRGGMVWALDLD--------------------------
UTRE-2015936        IMAEM--AALGG----AVYKSDAQ-RSAYYV-RGT-LWVGFDVQETFYEKMQGAAAF-GLGGVMIWDLSLDSPDQRL--------------------
UTRE-2015935        LMAQM--AAQGG----SVYKDNTQ-RSAYYV-LGT-LWVGFVVQETLYAKMQGAAAY-GLGGVMVWDLSQDTTDQRL--------------------
GFUR-2016925        IMAAR--GSAAGSTKAGIFQSTTE-RSAYFV-DGT-SWVGFDVEETLWAKIQGATER-GLGGVMIWEVSLDTF------------------------
IRYH-2007489        LMKEQ--VANGG----TVFVDVDE-KSAYYV-KGT-RWVGFDTPETMWLKIQAAQQR-GLGGVMIWDASLDTVNQQL--------------------
KUJU-2004605        ILREL--SLRGG----GVFVNDTE-RSAYYV-RGK-RWIGFDIPETMASKIRGAREY-GTGGVMIWDASLDTAD-----------------------
ISPU-2004540        IVREF--EVYGG----KILVDATE-RSSYYV-RGT-RWIGFDTERTLAAKLQGAREY-GIGGMMIWEASQDT-------------------------
ISPU-2004539        IMSEL--LTKGG----GIFVEGAE-RSAYYV-RGT-RWIGFDIEESFAAKIQGSRSY-GIGGVMIWDASLD--------------------------
RNAT-2058594        ILRDL--ATTGG----GVFVDTQQ-RSAYYV-YDT-RWIGFDIPETMALKMQGSRSF-GVGGVMIWDASLDTPD-----------------------
RYJX-2062142        VMREL--NYSGG----GVFVDVGE-RSAYYV-SGL-KWVGFDVPESMALKIQGTREY-GLAGVMIWDASLDTSD-----------------------
WRSL-2006874        ILHEL--ATIGG----SVFMDVPQ-RSAYYV-RGT-RWIGFDIPETMALKIQGSRTF-GVGGVMIWDASLDVRE-----------------------
BFIK-2029987        IKALVTSAKNPW----TSLFDSNT-QCQYAY-GGGNQWVGYDNPTTIGLKASFAKSK-GLGGYFFWEVSLDS-------------------------
DRFX-2014366        IKSNV--QSGGW----TAKYYSPA-TAMYAY-KGK-QWVGYDN------------------------------------------------------
SNOX-2004982        ILSML--SLPGT----ITVFDNTT-SSVYSV-NGL-QWVGYDNPQSIRTKVKYFKNSIKALGYFFWAIDLD--------------------------
KEYW-2026154        ILNFI---QGGA----TVVFDKAS-TSMYAY-KGT-QWVGYDDVSTIQTKAKYAQSN-GLGGYMFWAMDLD--------------------------
HKZW-2012228        INSFI---RSGA----TVVFNQSK-ASVYAY-KGS-HWIGYDDTRTISLKAKYAQSQ-GFGGYMFWAMDLD--------------------------
YOXI-2002515        ITDKF--LKSGA----TAMFDNTT-SSMYAY-KGR-QWIGYDNEKSIGIKASYAQKH-GLGGYMFWNVDMD--------------------------
VAZE-2009036        IVNFF--LKGGA----TAVFDKTT-SSEYAY-KGR-QWIGYDGEKSIGIKAKYAQSH-GLGGYMFWTVDMD--------------------------
JOJQ-2040355        IQEYF--VNKGA----TSVFDKTT-SSVYAF-KGL-QWIGYDNTKSIALKAKYAQVQ-GLAGYMFW-------------------------------
NBYP-2006024        IEKFF--VKAGA----TVVFDKTT-SSMYAY-KRL-QWIGYDNTKTIAIKAKYVKTQ-GLAGYMFWTLDLD--------------------------
MFZO-2022650        IKRYF--NKGGA----TEVFHSTT-SSAYAY-KGP-LWIGYDNTKSIGIKAKY--------------------------------------------
MOYY-2018456        INLFL--RGRGA----VSVFDRKT-TCQYAF-KGN-QWVGYDNTLTITKKVRFAKMR-GLSGYAFWA------------------------------
MCHJ-2001441        INLML---RSGG----VSVFNATT-TCQYAY-KGH-QWVSYDNTNTISLKVQFARQR-GLAGYGFWA------------------------------
WCQU-2007350        IKAFL---KKGA----VSNYSAAY-VGQYAY-KGT-QWVGYDNENTIAVKVKYAQSR-GIGGYGFW-------------------------------
MCHJ-2000297        ILTFL--RRAQA-------TDSGR-VGQYAW-NGY-QWVGFDNPNTIATKVRFAAGK-GLAGYTFWAYNLDSND-----------------------
XRTZ-2004862        IMGFL---HSGA----TSAFDAPS-SSMYAY-KGN-QWVGYDNRATIKIKAGYTKSQ-KLGGYMIWALDLD--------------------------
RQFE-2000106        IVQFV---KNGA----TSVFHAPS-TSQYAY-SGT-KWIGFDNP-----------------------------------------------------
QWFV-2003029        IVAFV---KGGA----QSMFDAAT-SSMYA-------------------------------------------------------------------
QWFV-2003026        IVTFI---KGGA----QSVFDAAT-SSTYAF-KGK-QWVGYDNTSTIQMKVLYAEAS-GLGGYAFWDFGLD--------------------------
QWFV-2003030        IVTFI---KGGA----QSVFDAAT-SSTYAF-KGK-QWVGYDNTSTIQMKVLYAEAS-GLGGYAFWDFGLD--------------------------
HIDG-2009902        IAGYV---KNGA----VAKLDSAT-SSMYAH-WGT-QWVGYDNPATIALKVQFASKR-GLGGYAFW-------------------------------
HIDG-2009903        IAAYI---KNGA----VAKLDSAT-SSMYAR-WGT-QWVGYDDPATISLKVRFASNR-GLGGYAFWSWDLD--------------------------
RPQV-2036136        IKKFI---ASGA----AAVFDVTT-SSMYAH-KGR-QWVGYDNPKTITLKAKFARQY-GLGGYAFWAHNLD--------------------------
AEKF-2009302        IVQLL---AAGA----NCVRDPAS-SSVYAF-KNR-KWVGYDDEQSITMKVAFARQH-GLAGYAFWAVDQD--------------------------
```

Figure S6. Class V chitinases alignments of complete sequences identified in bryophytes

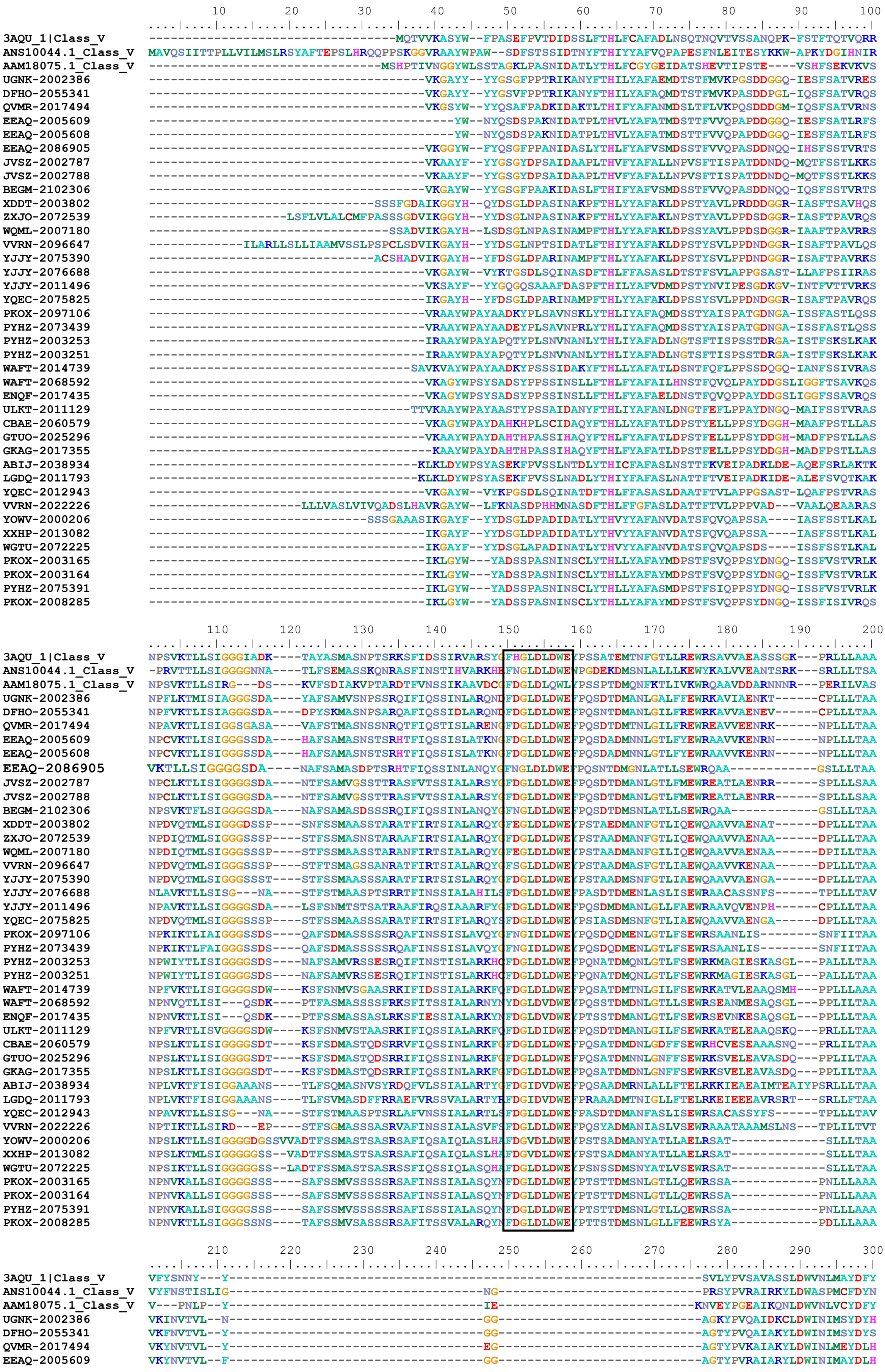
```
                            10        20        30        40        50        60        70        80        90       100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      ------------------------------MQTVVKAS----YWFPASE--FPVTDIDSSLFTHLFCAFADLNSQTNQVTV-----SSANQPKFSTFTQT
ANS10044.1_Class_V  MAVQSIITTPLLVILMSLRSYAFTEPSLHRQQPPSKGGVRAAYWPAWSD--FSTSSIDTNYFTHIYYAFVQPAPESFNLEI-----TESY-KKWAPKYDG
AAM18075.1_Class_V  ----------------------------MSHPTIVNGG----YWLSSTAGKLPASNIDATLYTHLFCGYGEIDATSHEVTI--PS------TEVSHFSEK
LFVP-2009172        ------------------------------SAATVKGG----YWQSGSL-RIPASSINATLFTHIFYAFAHIEENTSLLKP-----HAND-ADIETFSQI
LFVP-2009173        ------------------------------SAATVKGG----YWQSGSL-RIPASSINATLFTHIFYAFAHIEENTSLLKP-----HAND-ADIETFSQI
IHWO-2011148        ------------------------------SAATVKGG----YWQSGSL-RIPASSINATLFTHIFYAFAHIEENTSLLKP-----HAND-ADIETFSQI
HMHL-2009920        ------------------------------SAATVKGG----YWQSGSL-RIPASSINATLFTHIFYAFAHIEENTSLLKP-----HAND-ADIETFSQI
UCRN-2052015        ----------------------------------VKGA----YWSMDSA--TPPASIPSRLYTHLFYAFVELNSSTFQVDQ------GSH-NLLVTFSAN
FITN-2000573        ----------------------------------IRAA----YWPAYAPQTYPLSNVNANLYTHLIYAFADLNGTSFTISP--S--STDR-GAISTFSKS
AKXB-2067103        ---------------LLGVMVGLSCGSAQTDVSPVKAA----YWNLDDT--TPPESIPSQLYTHLFFAFVGLNSSTFQVVP----GSHKLAEMLATFSAT
FITN-2000575        ----------------------------------IRAA----YWPAYSAQTYPLSKVNANLYTHLFYAFADLNGTTFTISP--S--TGDS-DAINTFSKT
RTMU-2183105        ------------------LLVLIMTPTFAAGQSIVKGG----YYVGGTSKLLPTSGLNASLFTHLFYAFANLADNSPKVVM--P--SNNA-QEVSTFSSN
WJLO-2007710        -----------------------------SSTTKVKGA----YWPTGSK-RVPASSINATLFTHLFYAFARIENNTSLLKP-----DPDDDGIIERFSAT
DXOU-2037702        ----------------------------------VKGA----YWFMDDT---TPESIPSQLYTHLFFAFADLNSSSFQVVP----GSHDVAEALATFSAT
ANON-2000173        -----------AMVVLGMLLLSSCLGAAQTGDLAVRGA----YWAAGLASMLPAASIPSNLFTHLFYAFVGFDSATYQVIP-----DAGD-AVLSTFSAT
BSNI-2050540        ----------------------------------VKGA----YWNSGSASLQPPASIASNLFTHLFFAFVKVDNVTFQVVP--SDAVTSQ--QLSGFSQA
JPYU-2038298        --------------------------------TTVKGG----YWQEGST-RVPASAINASLFTHIFYAFAFIESNTSLLKP-----SAND-VDIETFSQT
WCZB-2119571        ----------------------------------VKGG----YWTVDFETVFPPKSIPSNLYTHLFFAFVGFNSSTFQVTP---GRHAVA-ELLASFSAT
SKQD-2079574        ----------------------------------VKGG----YWEAGVS--ASTSQFNANFNTHILYAFAVMDSTSFNVGA--P--TDDN-GQISSFTSI
FITN-2016127        ----------------LLLCMAILTGLDAVQSESVRAA----YWPAYAADEYPLSAVNPRLYTHLIYAFAQMDSSTYAISP--A--TGDN-GAISSFAST
TXVB-2074185        --------------------------------SKVKGG----YWQRGSV-LVPARAINATLFTHLFYAFVEIESTTSFVIP-----AEDDNGGLETFSGI
AWOI-2072168        --------LVRILLIRSLFPLLVTHIPLLVQGATVKGG----YWVGGIS----TANVNFKYQTHVFYAFAAIDSTTYECVA--P--TDDG-GQISTFTDT
TGKW-2013377        --------------------LAPGFAAASGQGSKVKGG----YYTGGTSHLFPSSGLNATLFTHLFYAFATVSQENLTVVA--D--PDDV-EEMKTFSTD
TGKW-2013378        --------------------LAPGFAAASGQGSKVKGG----YYTGGTSHLFPSSGLNATLFTHLFYAFATVSQENLTVVA--D--PDDV-EEMKTFSTD
JMXW-2011418        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYELVA--P--SDDN-GQIATFVAV
CHJJ-2137152        ----------------------------------VKGG----YYTAGSSHLLPSSGLNATLFTHLFYAFATVSQENFTVVA--P--ADDV-EEMRTFSKG
RXRQ-2139890        -----------------LGLLGLSCEAALTDESPVKGG----YWTVDFETVFPPKSIPSNLYTHLFFAFVGFNSSTFQVTP---GRHAVA-ELLASFSAT
KEFD-2060547        ----------------------------------VKGG----YWLGDVA----LENINFNYQTHVFYAFAGVVNTSYELVP--P--TGDN-GQYATFVAT
ILBQ-2045545        ----------------------------------VKGG----YWPSGSK-RVPASSINATLFTHLFYAFAGIESNTSFLKP-----NPDDGGIIETFSKT
JADL-2003775        ----------------------------------VKGG----YWLGGES----LATINFNFQTHVFYAFAGVDPTTYQLVP--PNATEDN-GQIATFVAS
UUHD-2149435        ----------------------------------VKGG----YYQAGTSGRLPSNGIKATLFTHLFYAFAHIDSTSFEVAA--S--PENV-VEMGAFSVN
CMEQ-2010358        ----------------------------------VKGG----YWLGGVS----LTNINFNFQTHVFYAFAGVDPTTYQLVP--PDVSDDD-GQIATFVAT
WEEQ-2089607        -----------------LGLLGLSCEAALTDESPVKGG----YWTVDFETVFPPKSIPSNLYTHLFFAFVGFNSSTFQVTP---GRHAVA-ELLASFSAT
IGUH-2002948        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNATEDN-GQITTFVGS
IGUH-2002949        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNATEDN-GQITTFVGS
YEPO-2005679        ----------------------------------VRGG----YWLQDLP----LVNINFNYETHVYYAFAGLDPSSYQVVA--P--TTDN-GQYATFVAT
NRWZ-2005174        ----------------------------------VKGG----YYTAGSSSLLPSGQIEAGYFTHLFFAFVHIDPTSYDVVV--P--VGQT-PEMQGFSAS
QMWB-2012715        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNATEDN-GQIATFVVS
DHWX-2004311        ----------------------------------VKGG----YWQGGES----LAHINFNFQTHVFYAFAGVDPSTYQLVP--PNATEDQ-GQIATFVVA
ORKS-2056948        ----------------------------------VKGG----YWLGDLV----LENVNFNFQTHVFYAFAGVDSTTYQLVP--PNPTEDN-GQIATFVAV
TMAJ-2010271        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNAADDN-GQIATFVVS
TMAJ-2010269        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNAADDN-GQIATFVVS
TMAJ-2010270        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNAADDN-GQIATFVVS
TAVP-2019781        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--P--TEDN-GQIATFVVS
ZACW-2001860        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNATE----QIATFVGS
WNGH-2008645        ----------------------------------VKGG----YWLGGLS----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNATEDN-GQIATFVAT
VBMM-2011897        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNVTDDN-GQIATFVVS
QKQO-2049619        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPTTYQLVP--PNATEDN-GQIATFVVS
PIUF-2092517        --------------------------------TNVKGG----YYSSST-----LTKVDFTQFTHLFYAFARVDNTSTVLAG-----PHDA-GGLEKFPFF
BPSG-2088193        ----------------------------------VKGG----YWGGGAS---MLESVNFNYETHVYYAFAGVDPTSYQLVP--PNATEDD-GQIATFVAI
YWNF-2051820        ----------------------------------VKGG----YWLGDVE----LSNVNFNFQTHVFYAFAGVDPTTYQLVP--PNATEDH-GQIATFVAM
ZTHV-2017150        ----------------------------------VKGG----YWVGGTS----TMHLNLNFETHVWYAFAGIDPTTYACIP--PAAADDG-GQYKSFVAT
LNSF-2002884        ----------------------------------VKGG----YWFGGES----LANINFNFQTHVFYAFAGVDPTTYQVVP--PSATEDN-GQIAKFVIS
SZYG-2012109        ----------------------------------VKGG----YWPGYIY--IPDGNLNLEFQTHVWYAFAGIN-SSYACVP--P--IFDDYGQYEMFVAT
ZQRI-2018149        ----------------------------------VKGG----YWLGGAV----LENINFNFQTHVFYAFAGVDPSTYELVA--PNATEDY-GQIATFVAI
NWWI-2025427        ----------------------------------IKGA----YFYYDSG--LDAADIDATLYTHVYYAFADVDATSFQVQA--PS------GTIASFSST
EEMJ-2001989        ----------------------------------VKGG----YWLGGES----LANINFNFQTHVFYAFAGVDPKTYQLVP--PNVTDDS-GQIATFVVS
IQJU-2000727        -----------------------------------RGS----YFPGGTG--IAPATVDAKLFTHLFFAFATIHASNFGIIPLYPANAAGPTGAWVQFNAA
NWWI-2025428        ----------------------------------IKGA----YFYYDSG--LDAADIDATLYTHVYYAFADVDATSFQVQA--PS------GTIASFSST
FITN-2091255        ----------------------------------IKLG----YWYADSS---PASNINSCLYTHLLYAFASMDPSTFSIQP--P--SYDN-GQISSFVST

                           110       120       130       140       150       160       170       180       190       200
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      VQRRNPSVKTLLSIGGGIADK-TAYASMASNPTSRKSFIDSSIRVARSYGFHGLDLDWEYPSSATEMTNFGTLLREWRSAVVAE---------ASSSGKP
ANS10044.1_Class_V  IHNIRPRVTTLLSIGGGGNNA-TLFSEMASSKQNRASFINSTIHVARKHEFNGLDLDWEWPGDEKDMSNLALLLKEWYKALVVE---------ANTSRKS
AAM18075.1_Class_V  VKVSNPSVKTLLSIRGDS----KVFSDIAKVPTARDTFVNSSIKAAVDCGFDGLDLQWLYPSSPTDMQNFKTLIVKWRQAAVDD---------ARNNNRP
LFVP-2009172        VKNINPEIKTLISISASGPD----LQEVMHTEQSRQAFINSSISLARSNGFDGLDLDYEFAQSQGEMDNFSSFLSEWRLAVNAE------------STTS
LFVP-2009173        VKNINPEIKTLISISASGPD----LQEVMHTEQSRQAFINSSISLARSNGFDGLDLDYEFAQSQGEMDNFSSFLSEWRLAVNAE------------STTS
IHWO-2011148        VKNINPEIKTLISISASGPD----LQEVMHTEQSRQAFINSSISLARSNGFDGLDLDYEFAQSQGEMDNFSSFLSEWRLAVNAE------------STTS
HMHL-2009920        VKNINPEIKTLISISASGPD----LQEVMHTEQSRQAFINSSISLARSNGFDGLDLDYEFAQSQGEMDNFSSFLSEWRLAVNAE------------STTS
UCRN-2052015        VKVANPFVKTLISIGGGGSDP-LAFANMTSNPARRSAFINSSIALARENGFDGLDLDYELPDSDVQMANFGALLDDWRTGITAE---------ANATSRT
FITN-2000573        LKAKNPWIYTLISIGGGGSDS-NAFSAMVRSSESRQIFINSTISLARKHCFDGLDLDWEFPQNATDMQNLGTLFSEWRKMAGIE---------SKASGLP
AKXB-2067103        VKVANPFVKTSISIGGGGFDP-GAFANMTSKPALRSAFINSSIALARENGFDGLDLDYEFPNSDEQMANFGHLLDEWRDGITAE---------ANATSRT
FITN-2000575        LKSKNPWIYTLISIGGGGSDS-NAFSAMVRSSESRQIFINSTISLARKHCFDGLDLDWEFPQNATDMQNLGTLFSEWRKMAGIE---------SKASGLP
RTMU-2183105        MKSKNPDVLTLLSIQATGAE----LAVVLQNNQTRGNFISSSITLARSYQFDGIDLDWEFARSQTEMNAFGVLIAEWRHAVSVE---------GQNTNRT
WJLO-2007710        VKSINPNVKTLISISASGAE----FQQVMQSEQSRRAFITSSINLARSNGFDGLDLDYEFPQSQAEMDDFSSFLSQWREAVDAD------------SQSS
DXOU-2037702        VKVANPFVKTLISIGGGDSDR-EAFANMTSKPALRSVFINSSIALARENGFDGLDLDYEFPDSDERMANFSDFLDDWRNGVTAE---------AKATSRT
ANON-2000173        VKASHPSVKTLLSIGGAGSDV-QGFAIMTATPALRAIFTNSSIALAREYGFDGLDLDYEFPASVEDTINLGFFLQEWRKGVEAE---------AAATNRT
BSNI-2050540        VKANNPAVKTLISIGGGDADP-DVLGNMTASAGLRSVFINSSIALARENGFDGLDLDYEFPKTDEYMKLYGEFLEEWRAAVVAE---------SVATNRS
JPYU-2038298        VKKINPEVKTLISISASGPE----FQEVMQTEQSRQAFINSSISLARSNGFDGLDLDYEFAQSQAEMDNFSSFIAEWRVAVNTE------------STTA
WCZB-2119571        VKLANPSVKTLISIGGGGYGE-VHFANMTSQPAYRSAFIKSSIALAREHGFDGVDLDYEFPESVENMALLGTFLDEWRAGVTAE---------AQETSRA
SKQD-2079574        ASASNPAVKTLLSIGGGNSNS-TAFGNMVSTAANRKVFINSSISLARQYGFNGLDLDWEYPNSDTDTANLATFLEEWRAAVVSE---------SQSSGKS
FITN-2016127        LQSSNPKIKTLFAIGGGSSDS-QAFSDMASSSSSRQAFINSSISLAVQYGFNGIDLDWEYPQSDQDMENLGTLFSEWRSAANLI--------------SS
TXVB-2074185        VKINNPDVKTLISIFAKGDE----FLHVMNSSLSRQAFIDSSIVLARTNGFDGLDLDYEFPESQADMDKFASFLAEWRQAVIAD---------SRNTSKS
AWOI-2072168        VTSSNPSVKTLLSIGGGSSNY-TAFANMVSTSSSRKVFIDSSISLARQYGFNGLDLDWEFPQSQLEMDNMGTFFQEWRAAAVTE---------SQSSGKA
TGKW-2013377        MKSSNPDVKTLISISVSGAE----FASVLQSNASRKAFINSSIALARAFQFDGLDLDWEFSRSQDEMDEFGRFIGEWRQAVDEE---AAARQQSSSSHPS
TGKW-2013378        MKSSNPDVKTLISISVSGAE----FASVLQSNASRKAFINSSIALARAFQFDGLDLDWEFSRSQDEMDEFGRFIGEWRQAVDEE---AAARQQSSSSHPS
JMXW-2011418        AKASNPDVVTLLSIGGAAAPF-ATFGAMVSTKARRKVFVDSSIKLARKYGFEGLDLDWESPQTQLEMTNLATLLKEWRAAVKKE---------ARASKKA
CHJJ-2137152        MKSSNPDVKTLISISISGAE----FATVLQTNASQRAFIDSSIALARAYQFDGLDLDWEFSRSQNEMDQFGGFIGKWRRAVDKEAEARQRQQSSSSSHPS
RXRQ-2139890        VKLVNPSVKTLISIGGGGYSE-VDFANMTSQPAYRSAFIKSSIALARENGFDGVDLDYEYPDSVENMALWGTFLDEWRAGVTAE---------AQETSRA
KEFD-2060547        AKNSNPSVVTLLSIGGAEAPF-ATFGAMVSESSSRQTFIDSSIALARQHGFEGLDLDWESPQTQQEMDNLALLLKEWRAAAVAE---------SQSSGKP
ILBQ-2045545        VKSINPDVKTLISISASGTE----FSQVMRNEQTRQAFITSSISLARSNGFDGLDLDYEYPQSQAEMDSFSSFLSEWREAVNAE------------SKAS
JADL-2003775        AKNSNPAVVPLLSIGGAAAPY-ATFGDMVSTSARRKVFIDTSIEVARQYGFEGLDLDWESPQTQQEMENLALLLKEWRAATTSE---------SEATGNT
UUHD-2149435        QKAVNPAVLTLISISAGGNE----FSNVLKSSASQQAFIESSIALARTNKFDGLDLDWEYPQSQADMDAFGLFLSDWRQAVVAE---------ATQSSSS
CMEQ-2010358        AKGSNPAVVPLLSIGGAGAPI-ATFSAMVSTSARRKVFVDSSIALARQHGFEGLDLDWESPQSQQEMANLALLLQEWRAAAAQE---------SQTSGKT
WEEQ-2089607        VKLVNPSVKTLISIGGGDYSE-VDFANMTSQPAYRSAFIKSSIALARENGFDGVDLDYEYPDSVENMALWGTFLDEWRAGVTAE---------AQETSRA
IGUH-2002948        AKNSNPAVVPLLSIGGAAAPY-ATFGDMVSTSARRKVFIDTSIVLARQYGFEGLDLDWESPQTQQEMEKLALLLKEWRAATASE---------SQASGKT
```

```
IGUH-2002949        AKNSNPAVVPLLSIGGAAAPY-ATFGDMVSTSARRKVFIDTSIVLARQYGFEGLDLDWESPQTQQEMEKLALLLKEWRAATASE---------SQASGKT
YEPO-2005679        AKSSNPSVVTLLSIGGGAADF-TTFGEMVSTSTNRQAFIDSSISLARQYSYEGLDLDWESPQSQTEMDNLALLLQEWRAAAKTE---------AQSSGNT
NRWZ-2005174        VKTINSAVLTLISIRTDGPG----FAPVLQSKTAQAAFINSSISLARSLNFDGLDLDWETSQSQDQMTAFGSLIGDWRNAVVQE---------AEQSGKP
QMWB-2012715        AKNSNPAVVPLLSIGGATAPY-ATFGDMVSTSARRKVFIDTSIAVARQYGFEGLDLDWESPQTQQEMEKLALLLQEWRAATTLE---------SQSSGKT
DHWX-2004311        AKNSNPAVVPLLSIGGATAPY-ATFGDMVSTEARRRVFIDTSIALARQHGFEGLDLDWESPQTQQEMEMLALLLKEWRAATTLE---------AQASGKT
ORKS-2056948        AKRSNPAVVPLLSIGGAAAPY-ATFGAMVSTSARRQVFIDSSIALARLYGYEGLDLDWESPQTQQEMENLAVLLKEWRAAAAQE---------SQVSGKT
TMAJ-2010271        AKNSNPAVVPLLSIGGGNATL-ATFSDMVSTSARRKVFIDTSIVLARQYGFEGLDLDWETPQSQQEMEKLALLLKEWRAATALE---------SQASGNT
TMAJ-2010269        AKNSNPAVVPLLSIGGGNATL-ATFSDMVSTSARRKVFIDTSIVLARQYGFEGLDLDWETPQSQQEMEKLALLLKEWRAATALE---------SQASGNT
TMAJ-2010270        AKNSNPAVVPLLSIGGGNATL-ATFSDMVSTSARRKVFIDTSIVLARQYGFEGLDLDWETPQSQQEMEKLALLLKEWRAATALE---------SQASGNT
TAVP-2019781        AKNSNPAVVPLLSIGGAAAPY-ATFGDMVSTSARRKVFTDTSIELARQYGFEGLDLDWESPQTQQEMENLALLLKEWRAATA-------------ASGKT
ZACW-2001860        AKNSNLAVVPLLFIGGVAAPY-ATFGDMVSTSARRKVFIDTSIVLARQYGFEGLDLDWESPQTQQEMEKLALLLKEWRAATASE---------SQASGKT
WNGH-2008645        AKSSNAAVVPLLSIGGAAAPY-ATFGAMVSTSARRKVFIDSSIVLARQYGYEGLDLDWESPQTQQEMENLALLLKEWRAATVLE---------SQASGNT
VBMM-2011897        AKNSNSAVVPLLSIGGAAAPY-ATFGEMVSTSARRKVFIDTSIQVARQYGFEGLDLDWESPQTQQEMENLALLLKEWRAATALE---------SEATGKT
QKQO-2049619        AKNSNPAVVPLLSIGGAAAPY-ATFGDMVSTSARRKVFIDTSIALARQYGFEGLDLDWESPQTQQEMEKLALLLKEWRATTALE---------SQVSGKT
PIUF-2092517        AKESNPSVFTLISISLTNDE----FTRVVATDDSRRTFINSSISLARSSGYDGLDMDWEFAQSQNDMDAFALFIQQWRQAVVSE---------SQLSSGA
BPSG-2088193        AKKSNPSVVPLLSIGGAAAPF-STFGEMVSTSSRRQVFIDSSIALARQHGFEGLDLDWESPQTQLEMDNLATLLKEWRAAAVTE---------SQSSGKT
YWNF-2051820        AKSSNPAVVPLLSIGGAAAPF-ATFGDMVSTSARRKVFIDSSIALAREYGYEGLDLDWESPQTQKEMEKLALLLKEWRAAAVAE---------SKSSGKA
ZTHV-2017150        ALASNPHVKTLLSVGGASSNQ-SDFATMVSKASTRKVFIDSSISLARSHGFHGLDLDWEFPQNQSEMNNMGIFFKEWRAAAV-E---------CKASKEP
LNSF-2002884        AKNSNPAVVPLLSIGGGKAPF-ATFGNMVSTFARRKVFINTSIEVARKFGFEGLDLDWESPQTQQETKNLALLLKEWRAATALE---------SRASGKT
SZYG-2012109        ALGSNPNIKVLLSIGGGTSNA-SLFASMVSTAADRKVFIDTSIALARQYNYSGLDLDWEFPASQTEMTNLGTFFQEWNAATVAE---------SQSSGKA
ZQRI-2018149        AKKSNPKVVTLLSIGGASAPQ-ATFGAMVSTSSRRAIFISSSIALARQHGFEGLDLDWESPQNQAQMVNLATLLKEWRAATKAE---------SLVSGKS
NWWI-2025427        LKALNPSVKTVLSIGGGSSSA-DTFSSMAATSASRSSFIQSAIQLASQNAFDGLDLDWEYPSTSSDMSNLAILLSEWRSA-------------------T
EEMJ-2001989        AKNSNPAVVPLLSIGGAAAPY-ATFGEMVSTSA--KVLIDSSIELARQHGFEGLDLDWESPQTQQEMENLALLLKP-TSSTLVE---------SQASGKT
IQJU-2000727        VKKSNPSVKTLLSIGGGEASATAAFVAMASAPARRQAFVNSAVKVAREHNFDGIDLDWEFPKNAQEKKNFTELIVSLRRAFAGE----------KRGKRA
NWWI-2025428        LKALNPSVKTVLSIGGGSSSA-DTFSSMAATSASRSSFIQSAIQLASQNAFDGLDLDWEYPSTSSDMSNLAILLSEWRSA-------------------T
FITN-2091255        VRLKNPNVKALLSIGGGGSSS-SAFSSMVSSSSSRSAFISSSIALASQYNFDGLDLDWEYPTSTTDMSNLGTLLQEWR------------------SSAP

                              210       220       230       240       250       260       270       280       290       300
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      RLLLAAAVF-YSNNY--YS----VLYPVSAVASSLDWVNLMAYDFY----GPGWS-RVTGPPAALFDPS--NAGPSGDAGTRSWIQAGLPAKKAVLGFPY
ANS10044.1_Class_V  RLLLTSAVY-FNSTI-SLIGNGPRSYPVRAIRKYLDWASPMCFDYN----GA-WA-NETGFNAALYDPN---SNISTKYGIGSWIGSGVPAEKLVMGLPL
AAM18075.1_Class_V  ERILVASVP-NLPYI-----EKNVEYPGEAIKQNLDWVNLVCYDFYTPAEGS--SPDYTGPSSALFNRR--SNSLSAHFGIESWTKY-VPAKKIVFGIPF
LFVP-2009172        PLLLTAAVY-YKPVL-ANVGVGT--YSVQAISENLDFVNIMCYDFH----GS-WDVTATGANANLYDSA---SNLSVDFGITSWLEAGLPSQKAVLGLTL
LFVP-2009173        PLLLTAAVY-YKPVL-ANVGVGT--YSVQAISENLDFVNIMCYDFH----GS-WDVTATGANANLYDSA---SNLSVDFGITSWLEAGLPSQKAVLGLTL
IHWO-2011148        PLLLTAAVY-YKPVL-ANVGVGT--YSVQAISENLDFVNIMCYDFH----GS-WDVTATGANANLYDSA---SNLSVDFGITSWLEAGLPSQKAVLGLTL
HMHL-2009920        PLLLTAAVY-YKPVL-ANVGVGT--YSVQAISENLDFVNIMCYDFH----GS-WDVTATGANANLYDSA---SNLSVDFGITSWLEAGLPSQKAVLGLTL
UCRN-2052015        PLLLTAAVY-YASSLIFYSSNAS--YPIPSIIKNLDFVNLMLYDLH----GP-WDKTQVNYNAPLFAPS---STLSASYGVGTWLTAGLPPSKAVFGIPA
FITN-2000573        ALLLTAAVY-YSATL-QYVGGGT--YPISSIKANLNFVNIMAFDYH----GS-WEPNKTGVHTAVFDPS---SGVNTDSGITKWLNAGLPGSKAVMGLAM
AKXB-2067103        PLLLTAAVY-YASSLIFYPSNAS--YPIPSIIKNLDFVNVMLYDLH----GP-WDKTQVNYNGPLFDPS---SVLSGSYGVRTWLTAGLPPSKAVFGIPA
FITN-2000575        ALLLTAAVY-YSATL-QYVGGGT--YPISSIKANLNFVNIMAFDYH----GS-WEPNKTGVHTAVFDPS---SGVNTDSGITKWLNAGLPGSKAVMGLAM
RTMU-2183105        TLLLTAAVY-FNSTM-IYAGKGS--YPIGAIADNLDFVNVMSYDYH----GS-WQPSETGANTNLYDNS---SLVSTDGGITSWLQAGMPPKKLVLGLTL
WJLO-2007710        PLLLTAAVY-YKPVL-KWVGEGA--YPIDSIKQNLDFVNVMCYDFH----GS-WDVTATGANANLYDPD---SDLSADFGITSWLKAGLPSEKAVLGLTL
DXOU-2037702        PLLLTAAVY-YASSL-IFSVANAS-FPVSSIIKNLDFVNLMLYDLH----GR-WDETQVNYNAPLFAPS---SVLSGSYGVRTWLTAGLPPSKAVFGIPA
ANON-2000173        PLLLTAAVY-YSSSL-LLSSGAT--YPVASMSKNLDFVNLMEYDLH----GS-WNKTEANYNAPLHGTG----VLSVDYGVQSWLRSGLPASKAVLGLPA
BSNI-2050540        ALLLTAAVY-FAPSF-IY-GDGRA-YPVSSMASSLDFVNLMLYDLH----GS-WEKERIDYNAPLFAFG----DLSASYGVQSWLGAGFPPSKLVFGIPS
JPYU-2038298        PLLLTAAVY-YKPVL-ANVGVGT--YPVKSISENLDFVNIMCYDFH----GS-WDVTQTGANANLYDPD---SNLSADFGITSWLEAGLPSEKAVLGLTL
WCZB-2119571        PLLLTAAVY-YSSSR-EFSSSNAS-YPIPSITRNLDFVNLMLYDLH----GP-WDRTRAEYNAPLFATG----SLSLAYGVQKWLTAGLPPSKAVFGLPS
SKQD-2079574        ELLLTAAVY-YLPTL-IYTGPGTQKYPISAINSYLDWVNMMAYDYH----GG-WESSQTGEHTALYDPN---SNLNTDYGVTEWLNAGLNASRLVMGLAF
FITN-2016127        NFIITAAVY-YSPSI-KWVGGGS--YPISSIKSNVDWVNIMTFDYH----GS-WEPNNTGAHTAVYDPS---GGVNTDSGISKWLDAGLPSQQGVLGLAM
TXVB-2074185        PLLLTAAVY-YQNEV-KNVGAGV--YPIESISENLDFVNIMCYDFH----GS-WDVTATGANTQLYDNS---STLSADHGISSWLNAGLPSVKAVLGMTL
AWOI-2072168        ELLLTAAVS-YQSIL-LNTGDGNQVYPISAINNYLDWVNVMCYDYH----GS-WEPSQTGEHTALYDPN---SDLNTHYGISKWLNAGLDTNHLALGLAY
TGKW-2013377        PLLLTAAVY-FASTL-LYTGKGT--YPIASIVQNLDFVNIMCYDFH----GS-WDVSHTGANANLYDNH---SVLSADHGITSWLDAGLPPEKASLGLTL
TGKW-2013378        PLLLTAAVY-FASTL-LYTGKGT--YPIASIVQNLDFVNIMCYDFH----GS-WDVSHTGANANLYDNH---SVLSADHGITSWLDAGLPPEKASLGLTL
JMXW-2011418        ELLLTAAVS-FQSIL-LYTGDGNQTWPIKAFNSYLNWANIMNYDYH----GS-WEPTITGEHTALYDPT---SIISTSYGISNWLQAGLKPKKLCLGLAY
CHJJ-2137152        PLLLTAAVY-FASTI-LYTGKGT--YPIASIAQNLDFVNIMCYDFH----GS-WDVSHTGANANLYDNR---SVISADHGITSWLDAGLPPEKATLGLTL
RXRQ-2139890        PLLLTAAVY-YSSSR-LFSSSNAS-YPIPSITRNLDFVNLMLYDLH----GP-WDTTRAEYNAPLFATG----SLSLDYGVQKWLTAGLPPSKAVFGLPA
KEFD-2060547        ELLLTAAVS-YQSIL-LYTGDMNQTWPVGAFNTYLNWVNVMNYDYH----GS-WEPTITGEHTALYDPD---SDIDTDYGISNWLAAGLQPDKLAMGLAY
ILBQ-2045545        PLLLTAAVY-FKPVL-KWVGEGD--YPIESISKNLDFVNIMCYDFH----GS-WDVTATGANANLYDPD---SDLSADYGITSWLKAGLPSEKAVLGLTL
JADL-2003775        ELLLTTAVS-FQSIL-LHTGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPTITGEHTALYDPS---SNISTDYGISNWLQAGLQKEKLALGLAY
UUHD-2149435        PLLLTAAVY-FSTPI-RYNGNGA--YPRLSMQENLDFVNIMCYDFH----GK-WELN-TGAGANLYDPN---STTSADNGINSWLDAGLPSSKAVLGLTL
CMEQ-2010358        ELLLTAAVS-YQAIL-LYTGDGNQTWPVDAFNTYLDWANVMAYDYH----GS-WEPTVTGEHTALYDPS---STISTDYGISNWLQAGLHQDKLALGLAY
WEEQ-2089607        PLLLTAAVY-YSSSR-LFISSNAS-YPIPSITRNLDFVNLMLYDLH----GR-WDTTRAEYNAPLFATG----SLSLDYGVQKWLTAGLPPSKAVFGLPA
IGUH-2002948        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPNITGEHTALYDPS---SNISTDYGISNWLQAGLQKEKLALGLAY
IGUH-2002949        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPNITGEHTALYDPS---SNISTDYGISNWLQAGLQKEKLALGLAY
YEPO-2005679        ELLLTAAVS-YQSIL-LYTGVGNQVWPITAFNNYLDWVNVMTYDYH----GS-WEPTTTGEHTALYDPN---SDVDTDYGINNWVSAGMQADKMCLGLAF
NRWZ-2005174        PLLLTAAVY-FASTF-PFPGTGNGTYPIAAITENLDFVNIMCYDFH----GS-WDKTATGANANLYAPNTTTTVLSADHGITTWLDAGLPLDKAVLGLTM
QMWB-2012715        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNSYLDWANVMTYDYH----GS-WEPTITGEHTALYDPS---SDVSTDYGISNWLQAGLQKQKLALGLAY
DHWX-2004311        ELLLTAAVS-YQPLQ-LYAGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPTLTGEHTALYDPS---SNISTDYGISNWLQAGLRKEKLALGLAY
ORKS-2056948        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNTYLDWANVMNYDYH----GS-WEPTITGEHTALYDPN---SNIDTDYGISNWLQAGLHADKLALGLAY
TMAJ-2010271        ELLLTAAVS-YLSIL-LYDGDGTRTWPIDAFNSYLDWANVMNYDYH----GS-WESTITGEHTALYDPS---YNVSTDYGISNWLQAGLPKEKLALGLAY
TMAJ-2010269        ELLLTAAVS-YLSIL-LYDGDGTRTWPIDAFNSYLDWANVMNYDYH----GS-WESTITGEHTALYDPS---YNVSTDYGISNWLQAGLPKEKLALGLAY
TMAJ-2010270        ELLLTAAVS-YLSIL-LYDGDGTRTWPIDAFNSYLDWANVMNYDYH----GS-WESTITGEHTALYDPS---YNVSTDYGISNWLQAGLPKEKLALGLAY
TAVP-2019781        ELLLTAAVS-YQSIQ-LYAGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPNITGEHTALYDPS---SDLSTDHGISNWLQAGLKKEKLALGLAY
ZACW-2001860        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPNITGEHTALYDPS---SNISTDYGISNWLQAGLQKEKLALGLAY
WNGH-2008645        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPNITGEHTALYDPN---SNISTSYGISNWLQAGLLKDKLALGLAY
VBMM-2011897        ELLLTAAVS-SQSIL-LFTGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WEPTITGEHTALYDPS---SNISTDYGISNWLQAGLQKEKLALGLAY
QKQO-2049619        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFKSYLDWANVMNYDYH----GS-WEPNITGEHTALYDPS---SNISTDYGISNWLKAGLQKEKLALGLAY
PIUF-2092517        PLLLTAAVY-FSSTI-LFPGKGS--YPITAIAENLDFVNIMCYDFH----GS-WNVNNTGANSNLYAPH---SVLSADYGISSWLSAGLPPDKAVLGLTL
BPSG-2088193        ELLLTAAVS-YQSIL-LYTGDMNQTWPIKAFNKYLDWANVMNYDYH----GS-WEPNITGEHTALYDPN---SNISTDYGISNWLQAGLHADKLCLGLAY
YWNF-2051820        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNNYLDWANVMNYDYH----GS-WEPTLTGEHTALYDPS---SNISTDYGISNWLQAGLHADKLALGLAY
ZTHV-2017150        ELLLTAAVS-YQSIL-LYTGDGNQTYPIEAINKYLDWVNVMNYDYH----GS-WECCNTGEHTALYDPT---SNLNTNYGISNWLKAGLRRKKCALGLAY
LNSF-2002884        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNSYLDWANVMNYDYH----GS-WPPNITAEHTALYDPN---SHISTDYGISNWLQAGLRKEKLALGLAY
SZYG-2012109        KLLLTAAVY-YQTTL-IYEQNQS--YNITAINSYLDWVNVMTYDYH----GS-WE-SQTGEHTTLYDPS--RPNLNTDYGVSNWLSSGLDATKLLMGLAF
ZQRI-2018149        ELLLTAAVS-YQSIL-LYTGDGNQTWPIDAFNKYLDWANVMNYEYH----GS-WEPTITGEHTALYDPN---STIDTDYGISNWLQAGL---KAQLGLAY
NWWI-2025427        SLLLTGAVY-FTLLL-----DGRD-YPVQSMAATLSWVNIMTYDYW-----T-WQSDKTGIHTALYESDSTQPTISTNGGVTAWLSAGLPSSQAALGLAS
EEMJ-2001989        EILLTAAVS-YQSIL-LFTGDGNQIWPIDAFNSYLDWANVMNYDYH----GS-WEPNITGEPTALYDPS---SNISTDYGISNWLQAGLRKEKLALGLAY
IQJU-2000727        VLLLTAAVAGYRPPA----------YDIAALKTNLDFAVVMCYDYH----GS-WE-QHTGLHAALEDPA--NPDASCSGAVQALTGGGMPASKLLMGLAA
NWWI-2025428        SLLLTGAVY-FTPYY--WTGAT---YPVQSMAATLSWVNIMTYDYW-----T-WQSDKTGIHTALYESDSTQPTISTNGGVTAWLSAGLPSSQAALGLAS
FITN-2091255        NLLLAAAVY-YTNPLPSWMGGAS--YPTSSIQSNLDWVLVMSYDLH----GS-WE-SYTNNPAPLYDPS--NPTVNGDYGFSQWLQY-LAMRKLLYGFPA

                              310       320       330       340       350       360       370       380       390       400
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      YGYAWRLTNANS-HSYYAPTTGAAISPDGSIG------YGQIRKFI-VDNGA--TTVYNS-TVVGDYCY----AGTNWIGYDDN---QSIVTKVRYAKQR
ANS10044.1_Class_V  YGRAWELKDPND-HGVGAKAV---GPAVD-TDGSMD--YDEILVFN-KDTGA--KVVYDE-VAVSFYSY----SGTTWIGYDDG---PSITKKVQFARSM
AAM18075.1_Class_V  HGWAWTLANAAE-HDMFSKATGSCVWPMD---------YSEVKSHR-RARQCEYNIPNDP-RFGIAYAY----VETTWIGYESE---STIPVKIKFAKQN
LFVP-2009172        YGRSWLLKNQTEAAGLNAPAVA-AGPAQPMSQSEGVFLYAEVTTFI-KEKTA--TVVHDL-EFVTAYCY----ADNVWVSYEDP---SIITAKVNYANSK
LFVP-2009173        YGRSWLLKNQTEAAGLNAPAVA-AGPAQPMSQSEGVFLYAEVTTFI-KEKTA--TVVHDL-EFVTAYCY----ADNVWVSYEDP---SIITAKVNYANSK
IHWO-2011148        YGRSWLLKNQTEAAGLNAPAVA-AGPAQPMSQSEGVFLYAEVTTFI-KEKTA--TVVHDL-EFVTAYCY----ADNVWVSYEDP---SIITAKVNYANSK
HMHL-2009920        YGRSWLLKNQTEAAGLNAPAVA-AGPAQPMSQSEGVFLYAEVTTFI-KEKTA--TVVHDL-EFVTAYCY----ADNVWVSYEDP---SIITAKVNYANSK
UCRN-2052015        YGYSWILADVSN-HTIGSAAIG-LGPQGGVLI------YKEVAEFI-TQEGV--TAVYDD-DSVSSYAY--N-ATGLWVGYDGP---QSVTAKVLYAKEF
FITN-2000573        YGKEWNLKNSSD-HGIGSPAVGVPSDSTPI--------FSSIVKFI-KQNHA--ALVFDN-TTLCSYCY----SGTSWIGFDDR---ISIAAKVGFLKSK
AKXB-2067103        YGYSWILADISK-NTIGSAAIG-LGPPQNVSGHEGLFLYKEVAEFI-TREGV--TAVYDD-ESVSSYAY--D-ATGLWVGYDGP---QSVTAKVLYAKEF
FITN-2000575        YGKEWNLKNSSD-HGIGSPAVGVPSDSTPI--------FSSIVRFI-KQKHA--SLVFDN-TTLCSYCY----SGTSWIGFDDRVSIPSIAAKVGFLKSR
```

```
RTMU-2183105        YGRSWTLKDENTANGLGAAAVG-LGPAQPVSNTAGIFYYSEVVSFI-ANDNA--TVVHDM-TEVTGYCY----AQNIWVSYEDP---FIIGNKVLYSKAK
WJLO-2007710        YGRSWLLKNQSEAAGLNAPAVA-AGPPQPISESEGVFFYAEVKMFI-QEKNA--AVVHDE-KLVTAYCY----ANNVWVTYEDP---IIIDTKVKYANSK
DXOU-2037702        YGYSWILADKSK-NTIGSAAIG-LGPHNG------IFTYKEVAEFI-TQEGV--TAVYDN-ENVSSYAY--N-ATGVWVGYDGP---QSVSAKVLYAKQL
ANON-2000173        YGYSWILANVSS-SSVGSAAIG-LGPAQNLSKAAGLFLYREVVAFT-KKPGV--KTVYDS-LNVSSYAY--D-ASGLWVGYDDP---RSVSTKVLYAKSQ
BSNI-2050540        YGVSWILAHVAN-HTIGSPAIG-LGTPQNVSGEAGLFLYREVVEFS-KKEGV--TTVYDN-ESVSSFAY--D-TAGLWVGFDSP---ESVRDKVKYAKRF
JPYU-2038298        YGRSWVLKDQNEATGLNAPAVA-AGPAQPISQEKGVFFYAEVVTFI-KDKSA--SVVTDG-TLVTGYCY----ADNLWVSYESQ---IVITRKVFYAQSK
WCZB-2119571        YGYSWILKDKSI-NAPGSATTG-LAPAQNVSGAPGLFVYSEIAEFI-KREDV--VSVYDA-KVVSSYAY--D-ATGVWVGYDSP---ESVAAKVLYSKQW
SKQD-2079574        YGHSWTLSNSSQ-TGVGAPTSA-AGSVYI---------YSDIVTFN---QDA--TVVEDS-TTVSEYSY--N-SSLTWIGYDNP---NTIGSKVQYAYGK
FITN-2016127        YGRSWNLKNSSD-HGIGAAAVGVPSDSTPT--------YSTIAKFI-TDNDA--TVVYDN-STVTAYCY----SQTSWIGFDNN---LSIAGKVKYLKSK
TXVB-2074185        YGRSWTLKNQSEAAGLGAPAVA-AGPPQPVSKAEGVFLYKEVKAFI-QEFNA--TVVHDN-EKVTAYCY----ANNTWVSYEDP---IVTTTKVDYAKNK
AWOI-2072168        YGKSWTLTSLSN-VGVGAPTTA-AGDVYI---------YKDIVTFN-E-GDP--TVVQDS-TTVSMYSY--K-TDLTWIGYDNP---DTIKAKVQYAEGR
TGKW-2013377        YGRSWVLKDQKAAAGLNAPATA-AGPAQPVSDAAGIFYYFEVANWI-EAKNA--TVVHND-TQVTGYCY----ADDVWVSYEDP---TIISAKVLYAKMK
TGKW-2013378        YGRSWVLKDQKAAAGLNAPATA-AGPAQPVSDAAGIFYYFEVANWI-EAKNA--TVVHND-TQVTGYCY----ADDVWVSYEDP---TIISAKVLYAKMK
JMXW-2011418        YGKIWNLTSLSN-TSVGAPATS-GGDPII---------YSDIVAYI-QQGGA--TVVYDN-TTVSMYSY--K-SDLTWIGYDNP---DTIAAKVRYAKSR
CHJJ-2137152        YGRSWVLKDQKAAAGLNAPATA-AGPAQPVSHAAGVFYYFEVANWI-QAKNA--TVVHDD-AQVTGYCY----ADNVWISYEDP---IIISAKVLYAKMK
RXRQ-2139890        YGYSWILEDKSK-TAPGSAATG-LAPAQNVSGAPGLFVYSEIAEFI-KREDV--VSVYDA-EVVSSYAY--D-ATGVWVGYDSP---ESIAAKVLYSKQW
KEFD-2060547        YGKQWNLTSLDD-VGVGAPATS-GGVPIT---------YEEIVAYN-Q-GGA--TVVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSK
ILBQ-2045545        YGRSWTLKNQSAAAGLNAAAVA-AGPPQPISQSKGVFFYAEVKNFI-QEKNA--TVVHDY-NFVTAYCY----ANNVWVTYEDP---VIISAKVNYANSK
JADL-2003775        YGKEWNLTSLNN-TGVGAPATS-GGDPIS---------YKDIVVYN-DAGGA--TVVQDA-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVQYAKGK
UUHD-2149435        YGPSWTLTDQSTAAGLNAPAIG-PGPVQGLSNEAGIFLYSEVVNFI-AANQATVTVVHDP-VEVTDYCY----ANNIWVSYEGP---VTITTKVQYAKSK
CMEQ-2010358        YGKEWNLTSVSN-TGVGAPATS-GGDPVV---------YSDIVVYN-EGGGA--TVVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTITAKVQYAKSR
WEEQ-2089607        YGYSWILEDKSK-TAPGSAATG-LAPAQNVSGAPGLFVYSEIAEFI-KREDV--VSVYDA-EVVSSYAY--D-ATGVWVGYDSP---ESIAAKVLYSKQW
IGUH-2002948        YGKEWNLTSLNN-TGVGAPATS-GGDPIF---------YKDIVVYN-E-GGA--TVVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
IGUH-2002949        YGKEWNLTSLNN-TGVGAPATS-GGDPIF---------YKDIVVYN-E-GGA--TVVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
YEPO-2005679        YGKKWVLASVAN-TGVGAPATS-GGDPIT---------YADIVTYN-NAGDA--TVVQDS-TTVSMYSY--K-SDLTWIGYDNP---DTIAAKVQYAQSK
NRWZ-2005174        YGRSWTLQNQNLASGLGAAATA-AGPAQPVSKTPGIFYYSEIPNWI-DTNNA--SVVRDL-RNVTAYCY----ADNVWISYEDP---IVITIKVFYAKAR
QMWB-2012715        YGKEWNLTSLNN-TGVGAPATS-GGDPIA---------YKDIVVYN-D-EGA--TVVQDN-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
DHWX-2004311        YGKEWNLTSLAN-TGIGAPATS-GGDPVL---------YKDIVAYN-EAGGA--TVVQDP-TTVSMYSY--K-SDLTWIGYDNP---DTIAAKVRYAKSR
ORKS-2056948        YGKEWNLTSLDN-VGVGAPATS-GGDPIV---------YSDIVVYN-E-GGA--TVVQDP-TTVCMYSY--K-SDLTWIGYDNP---DTIAAKVRYAKSK
TMAJ-2010271        YGKEWNLTSLDN-TGIGAPATS-GGDPIS---------YKDIVVYN-D-GGA--TVVQDS-TTVSMYSY--K-PDLTWIGYDNP---DTVAAKVRYAKSR
TMAJ-2010269        YGKEWNLTSLDN-TGIGAPATS-GGDPIS---------YKDIVVYN-D-GGA--TVVQDS-TTVSMYSY--K-PDLTWIGYDNP---DTVAAKVRYAKSR
TMAJ-2010270        YGKEWNLTSLDN-TGIGAPATS-GGDPIS---------YKDIVVYN-D-GGA--TVVQDS-TTVSMYSY--K-PDLTWIGYDNP---DTVAAKVRYAKSR
TAVP-2019781        YGKEWNLTSLNN-TGVGAPATS-GGDPIS---------YKDIVVYN-D-GGA--TVVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
ZACW-2001860        YGKEWNLTSLNN-TGVGAPATS-GGDPIF---------YKDIVVYN-E-GGA--TVVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
WNGH-2008645        YGKEWNLTSLNN-TGVGAPATS-GGDPII---------YSDIVVYN-E-GGA--TIVQDK-TTVCMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
VBMM-2011897        YGKEWNLTSLNN-TGVGAPATS-GGDPIS---------YKDIVVYN-D-GGA--TVVQDN-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
QKQO-2049619        YGKEWNLTSLNN-TGIGAPATS-GGDPII---------YKDIVVYN-D-GGA--TIVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
PIUF-2092517        YGRSWTLKNQTLDAGLNASATA-AGPAQPFSQTAGVFYYYEVANWI-EANNA--TVVHDQ-EEVTAYCY----ANDIWVSYEDP---EIVSAKVLYAMSK
BPSG-2088193        YGKEWNLTSLDN-VGVGAPARG-GGDPIS---------YKDIVVYN-E-GGA--TVVYDP-TTVSMYSY--KSSELTWIGYDNP---DTIAAKVRYAKSR
YWNF-2051820        YGKEWNLTSLAD-VGVGAPATS-GGDPIV---------YSDIVVYN-E-GGA--TVVQDP-TTVCMYSY--K-PDLTWVGYDNP---DTIAAKVRYAKGR
ZTHV-2017150        YGKTWTLVSLKN-ISVGAPTSK-AGDYLL---------YRDIVTFNKEAPKS--TVFMDK-ATVSMISV--K-SNLTWVGYDNP---STIARKVKYAKRR
LNSF-2002884        YGKEWKLTSLQN-TRIGAPTKS-GGDPIS---------YKDIVVYN-D-GGA--TVVQDP-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAKSR
SZYG-2012109        YGKSWVLSSLSN-TSIGAPTSS-YYEDIS---------YSDIVTFN-TNSSV--IVVQDN-TTVCMYSY---------NTLDSP---TTITSKVQYAKNK
ZQRI-2018149        FGKQWNLTSLDN-VGVGAPATS-GGDPIF---------YRDIVVYN-Q-GGA--TVVQDP-VTVSMYSY--K-SDLTWNGYDNP---DTITAKVQYAKSK
NWWI-2025427        YGVTWTLESASN-NGIGAPATA-FGSPMTLG-------YDQIQSFI-QSSST--TCQRDSTTTVASYCYGSSGSGVLWVGFDDA---TSIAAKVEYLKSK
EEMJ-2001989        YGKEWNLTSLSN-TGVGAPAAPLGGVPIR---------YEDIVVYN-E-GGA--TVVQDS-TTVSMYSY--K-PDLTWIGYDNP---DTIAAKVRYAQSR
IQJU-2000727        YGRSWTLKDAAD-NGVGAPTTG-VGLPGPQTKDPGLLSYAEIAGYV--KNGA--VAKLDS-ATSSMYAH----WGTQWVGYDNP---ATIALKVQFASKR
NWWI-2025428        YGVTWTLESASN-NGIGAPATA-FGSPMTLG-------YDQIQSFI-QSSST--TCQRDSTTTVASYCYGSSGSGVLWVGFDDA---TSIAAKVEYLKSK
FITN-2091255        YGYTWNLSNPSN-HGTGAPASG-SYTSLGETPS-----YARIATFL-TSTGA--TIVHNA-TYVSTYAY----SGRAWVGYDDA---WSITRKVNYLKSK

                             410       420       430
                    ....|....|....|....|....|....|....|....
3AQU_1|Class_V      GLLGYFSWHVGADDNSGLSRAASQAWDATTATTRTIQKV
ANS10044.1_Class_V  GLKGYFFWAIGKDKD----------WTISKQASNAWGY-
AAM18075.1_Class_V  QLLGYFAWNISGDDDV---------NTLANAAQSNWQ--
LFVP-2009172        NLRGYFFWAITQDSL----------WEVSSAASVA----
LFVP-2009173        NLRGYFFWAITQDSL----------WEVSSAASVA----
IHWO-2011148        NLRGYFFWAITQDSL----------WEVSSAASVA----
HMHL-2009920        NLRGYFFWAITQDSL----------WEVSSAASVA----
UCRN-2052015        NLRGYFFWTVGYDDSN---------WSLSESASRA----
FITN-2000573        GLQGYFFWTASYDDKK---------YTLSTTAAQS----
AKXB-2067103        KLRGYFFWAVGYDDSK---------WSLSESASRA----
FITN-2000575        GLQGYFCWTASYDDKK---------NTLSKTAAQA----
RTMU-2183105        DLRGYFFWQVSEDNN----------FQVSAAASQ-----
WJLO-2007710        NLRGYFFWAITQDSD----------WEVSTAASQ-----
DXOU-2037702        NLRGYFFWAVGFDDSK---------WSLSESASR-----
ANON-2000173        QMRGYFFWAVNYDDSN---------WSLSTAA-------
BSNI-2050540        NMRGYFFWAVNYNDAQ---------WTISKAASNA----
JPYU-2038298        NLRGYFFWAITQDDDL---------WSISYAASVA----
WCZB-2119571        KLRGYFFWAVSYDDTE---------SSLGEAASRA----
SKQD-2079574        SLLGYFAWALHQDDSN---------WSMA----------
FITN-2016127        LLKGYFFWTSSYD--------------------------
TXVB-2074185        NLGGYFFWAITQDNN----------WEVSTAASQA----
AWOI-2072168        ALLGYFAWALHHD--------------------------
TGKW-2013377        NLRGYFFWEVSYDRN----------WEVSTAASQ-----
TGKW-2013378        NLRGYFFWEVSYDRN----------WEVSTAASQ-----
JMXW-2011418        KLRGFFAWALHHDTLN---------WALQSAA-------
CHJJ-2137152        NLRGYFFWEVSYDRN----------WEISSAASQ-----
RXRQ-2139890        KLRGYFFWSVSYD--------------------------
KEFD-2060547        NLLGYFAWALHQDTNT---------WSLASAA-------
ILBQ-2045545        NLRGYFFWAVTQDSD----------WEVSTTASQ-----
JADL-2003775        DLLGFFAWALYHDTAN---------WSLASAACN-----
UUHD-2149435        NLLGYFFWAVSDDSN----------WEISTAASQA----
CMEQ-2010358        DLLGFFAWALHQDTST---------WSLAS---------
WEEQ-2089607        KLRGYFFWSVSYD--------------------------
IGUH-2002948        DLLGYFAWALHHDTGN---------WSLASAASN-----
IGUH-2002949        DLLGYFAWALHHDTGN---------WSLASAASN-----
YEPO-2005679        SLLGYFAWALHQDDAN---------FSLASAAQN-----
NRWZ-2005174        NLRGYFFWQVGQDLN----------WQLSSAASQ-----
QMWB-2012715        DLLGFFAWALHHDTGN---------WSLASAAS------
DHWX-2004311        DLRGFFAWALYHDTEN---------WSLASAAS------
ORKS-2056948        NLLGYFAWALHQDTKN---------WSLASAA-------
TMAJ-2010271        DLLGFFAWALYQDTEN---------WSLASAASN-----
TMAJ-2010269        DLLGFFAWALYQDTEN---------WSLASAASN-----
TMAJ-2010270        DLLGFFAWALYQDTEN---------WSLASAASN-----
TAVP-2019781        DLLGFFAWALYHDTAN---------WSLASAASN-----
ZACW-2001860        DLLGYFAWALHHDTGN---------WSLASAASN-----
WNGH-2008645        DLLGFFAWALHQDTGN---------WSLASAASN-----
VBMM-2011897        DLLGFFAWALYHDTGN---------WSLASAASN-----
QKQO-2049619        DLLGFFAWALHHDTRN---------WSLASAAS------
PIUF-2092517        GLRGYFFWAVTFDRD----------AEVSTAASQ-----
BPSG-2088193        NLLGYFAWALYHDT-------------------------
YWNF-2051820        DLLGFFAWALHQDTNS---------WSLASAA-------
```

```
ZTHV-2017150        GLLGYFAWALHHDDDK---------WSLA----------
LNSF-2002884        DLLGFFAWALSQDTKN---------WSLAS---------
SZYG-2012109        SVGGYFAWNLAQDDSN---------WSLAKAASN-----
ZQRI-2018149        KLLGFFAWALHQDDDA---------WSLAS---------
NWWI-2025427        QLAGYFFWRLGLD--------------------------
EEMJ-2001989        ELLGFFAWALYHDTGN---------WSLASAASN-----
IQJU-2000727        GLGGYAFW-------------------------------
NWWI-2025428        QLAGYFFWRLGLD--------------------------
FITN-2091255        GLRGYGFWT------------------------------
```

Figure S7. Class V chitinases alignments of complete sequences identified in pteridophytes



```
                            10        20        30        40        50        60        70        80        90       100
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      ----------------------------------MQTVVKASYW--FPASEFPVTDIDSSLFTHLFCAFADLNSQTNQVTVSSANQPK-FSTFTQTVQRR
ANS10044.1_Class_V  MAVQSIITTPLLVILMSLRSYAFTEPSLHRQQPPSKGGVRAAYWPAW--SDFSTSSIDTNYFTHIYYAFVQPAPESFNLEITESYKKW-APKYDGIHNIR
AAM18075.1_Class_V  --------------------------------MSHPTIVNGGYWLSSTAGKLPASNIDATLYTHLFCGYGEIDATSHEVTIPSTE----VSHFSEKVKVS
UGNK-2002386        --------------------------------------VKGAYY--YYGSGFPPTRIKANYFTHILYAFAEMDTSTFMVKPGSDDGGQ-IESFSATVRES
DFHO-2055341        --------------------------------------VKGAYY--YYGSVFPPTRIKANYFTHILYAFAQMDTSTFMVKPASDDPGL-IQSFSATVRQS
QVMR-2017494        --------------------------------------VKGSYW--YYQSAFPADKIDAKTLTHIFYAFANMDSLTFLVKPQSDDDGM-IQSFSATVRNS
EEAQ-2005609        ------------------------------------------YW--NYQSDSPAKNIDATPLTHVLYAFATMDSTTFVVQPAPDDGGQ-IESFSATLRFS
EEAQ-2005608        ------------------------------------------YW--NYQSDSPAKNIDATPLTHVLYAFATMDSTTFVVQPAPDDGGQ-IESFSATLRFS
EEAQ-2086905        --------------------------------------VKGGYW--FYQSGFPPANIDASLYTHLFYAFVSMDSSTFVVQPASDDNQQ-IHSFSSTVRTS
JVSZ-2002787        --------------------------------------VKAAYF--YYGSGYDPSAIDAAPLTHVFYAFALLNPVSFTISPATDDNDQ-MQTFSSTLKKS
JVSZ-2002788        --------------------------------------VKAAYF--YYGSGYDPSAIDAAPLTHVFYAFALLNPVSFTISPATDDNDQ-MQTFSSTLKKS
BEGM-2102306        --------------------------------------VKGAYW--YYGSGFPAAKIDASLFTHIFYAFVSMDSSTFVVQPASDDNQQ-IQSFSSTVRTS
XDDT-2003802        -------------------------------SSSFGDAIKGGYH--QYDSGLDPASINAKPFTHLYYAFAKLDPSTYAVLPRDDDGGR-IASFTSAVHQS
ZXJO-2072539        -------------------LSFLVLALCMFPASSSGDVIKGGYH--YYDSGLNPASINAKPFTHLYYAFAKLNPSTYAVLPPDSDGGR-IASFTPAVRQS
WQML-2007180        ---------------------------------SSADVIKGAYH--LSDSGLNPASINAMPFTHLYYAFAKLDPSSYAVLPPDSDGGR-IAAFTPAVRRS
VVRN-2096647        -------------ILARLLSLLIAAMVSSLPSPCLSDVIKGAYH--YYDSGLNPTSIDATLFTHIYYAFAKLDPSTYSVLPPDNDGGR-ISAFTPAVLQS
YJJY-2075390        -------------------------------ACSHADVIKGAYH--YFDSGLDPARINAMPFTHLYYAFAKLDPSTYSVLPPDNDGGR-ISAFTPAVRKS
YJJY-2076688        --------------------------------------VKGAYW--VYKTGSDLSQINASDFTHLFFASASLDTSTFSVLAPPGSAST-LLAFPSIIRAS
YJJY-2011496        --------------------------------------VKSAYF--YYGQGQSAAAFDASPFTHILYAFVDMDPSTYNVIPESGDKGV-INTFVTTVRKS
YQEC-2075825        --------------------------------------IKGAYH--YFDSGLDPARINAMPFTHLYYAFAKLDPSSYSVLPPDNDGGR-ISAFTPAVRQS
PKOX-2097106        --------------------------------------VRAAYWPAYAADKYPLSAVNSKLYTHLIYAFAQMDSSTYAISPATGDNGA-ISSFASTLQSS
PYHZ-2073439        --------------------------------------VRAAYWPAYAADEYPLSAVNPRLYTHLIYAFAQMDSSTYAISPATGDNGA-ISSFASTLQSS
PYHZ-2003253        --------------------------------------IRAAYWPAYAPQTYPLSNVNANLYTHLIYAFADLNGTSFTISPSSTDRGA-ISTFSKSLKAK
PYHZ-2003251        --------------------------------------IRAAYWPAYAPQTYPLSNVNANLYTHLIYAFADLNGTSFTISPSSTDRGA-ISTFSKSLKAK
WAFT-2014739        ------------------------------------SAVKVAYWPAYAADKYPSSSIDAKYFTHLLYAFATLDSNTFQFLPPSSDQGQ-IANFSSIVRAS
WAFT-2068592        --------------------------------------VKAGYWPSYSADSYPPSSINSLLFTHLFYAFAILHNSTFQVQLPAYDDGSLIGGFTSAVKQS
ENQF-2017435        --------------------------------------VKAGYWPSYSADSYPPSSINSLLFTHLFYAFAELDNSTFQVQPPAYDDGSLIGGFSSAVRQS
ULKT-2011129        ------------------------------------TTVKAAYWPAYAASTYPSSAIDANYFTHLIYAFANLDNGTFEFLPPAYDNGQ-MAIFSSTVRAS
CBAE-2060579        --------------------------------------VKAGYWPAYDAHKHPLSCIDAQYFTHLLYAFATLDPSTYELLPPSYDGGH-MAAFPSTLLAS
GTUO-2025296        --------------------------------------VKAAYWPAYDAHTHPASSIHAQYFTHLFYAFATLDPSTFELLPPSYDDGH-MADFPSTLLAS
GKAG-2017355        ---------------------------------------KAAYWPAYDAHTHPASSIHAQYFTHLFYAFATLDPSTFELLPPSYDDGH-MADFPSTLLAS
ABIJ-2038934        -------------------------------------KLKLDYWPSYASEKFPVSSLNTDLYTHICFAFASLNSTTFKVEIPADKLDE-AQEFSRLAKTK
LGDQ-2011793        -------------------------------------KLKLDYWPSYASEKFPVSSLNADLYTHIYFAFASLNATTFTVEIPADKIDE-ALEFSVQTKAK
YQEC-2012943        --------------------------------------VKGAYW--VYKPGSDLSQINATDFTHLFFASASLDAATFTVLAPPGSAST-LQAFPSTVRAS
VVRN-2022226        ---------------------LLLVASLVIVQADSLHAVRGAYW--LFKNASDPHHMNASDFTHLFFGFASLDATTFTVLPPPVAD---VAALQEAARAS
YOWV-2000206        ------------------------------SSSGAAASIKGAYF--YYDSGLDPADIDATLYTHVYYAFANVDATSFQVQAPSSA----IASFSSTLKAL
XXHP-2013082        --------------------------------------IKGAYF--YYDSGLDPADIDATLYTHVYYAFANVDATSFQVQAPSSA----IASFSSTLKAL
WGTU-2072225        --------------------------------------IKGAYF--YYDSGLAPADINATLYTHVYYAFANVDATSFQVQAPSDS----ISSFSSTLKAL
PKOX-2003165        --------------------------------------IKLGYW---YADSSPASNINSCLYTHLLYAFAYMDPSTFSVQPPSYDNGQ-ISSFVSTVRLK
PKOX-2003164        --------------------------------------IKLGYW---YADSSPASNINSCLYTHLLYAFAYMDPSTFSVQPPSYDNGQ-ISSFVSTVRLK
PYHZ-2075391        --------------------------------------IKLGYW---YADSSPASNINSCLYTHLLYAFASMDPSTFSIQPPSYDNGQ-ISSFVSTVRLK
PKOX-2008285        --------------------------------------IKLGYW---YADSSPASNINSCLYTHLLYAFAYMDPSTFSVQPPSYDNGQ-ISSFISIVRQS

                           110       120       130       140       150       160       170       180       190       200
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      NPSVKTLLSIGGGIADK----TAYASMASNPTSRKSFIDSSIRVARSYGFHGLDLDWEYPSSATEMTNFGTLLREWRSAVVAEASSSGK---PRLLLAAA
ANS10044.1_Class_V  -PRVTTLLSIGGGGNNA----TLFSEMASSKQNRASFINSTIHVARKHEFNGLDLDWEWPGDEKDMSNLALLLKEWYKALVVEANTSRK---SRLLLTSA
AAM18075.1_Class_V  NPSVKTLLSIRG---DS----KVFSDIAKVPTARDTFVNSSIKAAVDCGFDGLDLQWLYPSSPTDMQNFKTLIVKWRQAAVDDARNNNR---PERILVAS
UGNK-2002386        NPFLKTMISIAGGGSDA----YAFSAMVSNPSSRQIFIQSSINLARQNDFDGLDLDWEFPQSDTDMANLGALFFEWRKAVIAENKT------CPLLLTAA
DFHO-2055341        NPFVKTLISIAGGGSDA----DPYSKMASNPSARQAFIQSSIDLARQNDFDGLDLDWEFPQSNTDMANLGILFREWRKAVVAENEV------CPLLLTAA
QVMR-2017494        NPAVKTLISIGGSGASA----VAFSTMASNSSNRQTFIQSTISLARENGFDGLDLDWEFPQSDTDMTNLGILFREWREAVVEENRK------NPLLLTAA
EEAQ-2005609        NPCVKTLISIGGGSSDA----HAFSAMASNSTSRHTFIQSSISLATKNGFDGLDLDWEFPQSDADMNNLGTLFYEWRAAVVKENRN------NPLLLTAA
EEAQ-2005608        NPCVKTLISIGGGSSDA----HAFSAMASNSTSRHTFIQSSISLATKNGFDGLDLDWEFPQSDADMNNLGTLFYEWRAAVVKENRN------NPLLLTAA
EEAQ-2086905        VKTLLSIGGGGSDA----NAFSAMASDPTSRHTFIQSSINLANQYGFNGLDLDWEFPQSNTDMGNLATLLSEWRQAA-----------GSLLLTAA
JVSZ-2002787        NPCLKTLISIGGGGSDA----NTFSAMVGSSTTRASFVTSSIALARSYGFDGLDLDWEFPQSDTDMANLGTLFMEWREATLAENRR------SPLLLSAA
JVSZ-2002788        NPCLKTLISIGGGGSDA----NTFSAMVGSSTTRASFVTSSIALARSYGFDGLDLDWEFPQSDTDMANLGTLFMEWREATLAENRR------SPLLLSAA
BEGM-2102306        NPSVKTFLSIGGGGSDA----NAFSAMASDSSSRQIFIQSSINLANQYGFNGLDLDWEFPQSDTDMSNLATLLSEWRQAA------------GSLLLTAA
XDDT-2003802        NPDVQTMLSIGGGDSSP----SNFSSMAASSTARATFIRTSIALARQYGFEGLDLDWEYPSTAEDMANFGTLIQEWQAAVVAENAT------DPLLLTAA
ZXJO-2072539        NPDIQTMLSIGGGSSSP----STFSSMASNSTARAAFIRTSIALARQYGFDGLDLDWEYPSTAADMANFGTLIQEWQAAVVAENAA------DPLLLTAA
WQML-2007180        NPDIQTMLSIGGGSSSP----STFSSMAASSTARANFIRTSIALARQYGFEGLDLDWEYPSTAADMANFGILIQEWQAAVVAENAA------DPILLTAA
VVRN-2096647        NPDVQTMLSIGGGSSSP----STFTSMAGSSANRATFIRTSIALARQYGFSGLDLDWEYPSTAADMASFGTLIAEWQAAVVKENAA------DPLLLTAA
YJJY-2075390        NPDVQTMLSIGGGSSST----STFSSMAASSSARATFIRTSISLARQYGFDGLDLDWEYPSTAADMSNFGTLIAEWQAAVVAENGA------DPLLLTAA
YJJY-2076688        NLAVKTLLSISG---NA----STFSTMAASPTSRRTFINSSIALAHILSFDGLDLDWEFPASDTDMENLASLISEWRAACASSNFS------TPLLLTAV
YJJY-2011496        NPAVKTLLSIGGGGSDA----LSFSNMTSTSATRAAFIRQSIAAARFYGFDGLDLDWEFPQSDMDMANLGLLFAEWRAAVQVENPH------CPLLLTAA
YQEC-2075825        NPDVQTMLSIGGGSSSP----STFSSMAASSSARATFIRTSIFLARQYSFDGLDLDWEYPSIASDMSNFGTLIAEWQAAVVAENGA------DPLLLTAA
PKOX-2097106        NPKIKTLIAIGGGSSDS----QAFSDMASSSSSRQAFINSSISLAVQYGFNGIDLDWEYPQSDQDMENLGTLFSEWRSAANLIS--------SNFIITAA
PYHZ-2073439        NPKIKTLFAIGGGSSDS----QAFSDMASSSSSRQAFINSSISLAVQYGFNGIDLDWEYPQSDQDMENLGTLFSEWRSAANLIS--------SNFIITAA
PYHZ-2003253        NPWIYTLISIGGGGSDS----NAFSAMVRSSESRQIFINSTISLARKHCFDGLDLDWEFPQNATDMQNLGTLFSEWRKMAGIESKASGL---PALLLTAA
PYHZ-2003251        NPWIYTLISIGGGGSDS----NAFSAMVRSSESRQIFINSTISLARKHCFDGLDLDWEFPQNATDMQNLGTLFSEWRKMAGIESKASGL---PALLLTAA
WAFT-2014739        NPFVKTLISIGGGGSDW----KSFSNMVSGAASRKIFIDSSIALARKFQFDGLDLDWEFPQSATDMTNLGILFSEWRKATVLEAAQSMH---PPLLLAAA
WAFT-2068592        NPNVQTLISI---QSDK----PTFASMASSSSFRKSFITSSIALARNYNYDGLDVDWEYPQSSTDMDNLGTLLSEWRSEANMESAQSGL---PPLILTAA
ENQF-2017435        NPNVKTLISI---QSDK----PTFSSMASSASLRKSFIESSIALARKYNFDGLDVDWEYPTSSTDMANLGTLFSEWRSEVNKESAQSGL---PPLLLTAA
ULKT-2011129        NPFVRTLISVGGGGSDW----KSFSNMVSTAASRKIFIQSSIALARKFQFDGLDIDWEFPQSDTDMANLGILFSEWRKATELEAAQSKQ---PRLLLTAA
CBAE-2060579        NPSLKTLISIGGGGSDT----KSFSDMASTQDSRRVFIQSSINLARKFGFDGLDLDWEFPQSATDMDNLGDFFSEWRHCVESEAAASNQ---PRLILTAA
GTUO-2025296        NPSLKTLISIGGGGSDT----KSFSDMASTQDSRRIFIQSSINLARKFGFDGLDLDWEFPQSATDMNNLGNFFSEWRKSVELEAVASDQ---PPLILTAA
GKAG-2017355        NPSLKTLISIGGGGSDT----KSFSDMASTQDSRRIFIQSSINLARKFGFDGLDLDWEFPQSATDMDNLGNFFSEWRKSVELEAVASDQ---PPLILTAA
ABIJ-2038934        NPLVKTFISIGGAAANS----TLFSQMASNVSYRDQFVLSSIALARTYGFDGIDVDWEFPQSAADMRNLALLFTELRKKIEAEAIMTEAIYPSRLLLTAA
LGDQ-2011793        NPLVKTFISIGGAAANS----TLFSVMASDFFRRAEFVRSSVALARTYRFDGIDVDWEFPRAAADMTNIGLLFTELRKEIEEEAVRSRT---SRLLFTAA
YQEC-2012943        NPAVKTLLSISG---NA----STFSTMAASPTSRLAFVNSSIALARTLSFDGLDLDWEFPASDTDMANFASLISEWRSACASSYFS------TPLLLTAV
VVRN-2022226        NPTIKTLLSIRD---EP----STFSGMASSSASRVAFINSSIALASVFSFDGLDLDWEFPQSYADMANIASLVSEWRAAATAAAMSLNS---TPLILTVT
YOWV-2000206        NPSLKTLLSIGGGGDGSSVVADTFSSMASTSASRSAFIQSAIQLASLHAFDGVDLDWEYPSTSADMANYATLLAELRSAT-------------SLLLTAA
XXHP-2013082        NPSLKTMLSIGGGGGGSS--VADTFSSMASTSASRSAFIQSAIQLASLHAFDGVDLDWEYPSTSADMANYATLLAELRSAT-------------SLLLTAA
WGTU-2072225        NPSIKTLLSIGGGGGGSS--LADTFSSMASTSASRSSFIQSSIQLASQHAFDGLDLDWEYPSNSSDMSNYATLVSEWRSAT-------------SLLLTAA
PKOX-2003165        NPNVKALLSIGGGGSSS----SAFSSMVSSSSSRSAFINSSIALASQYNFDGLDLDWEYPTSTTDMSNLGTLLQEWRSSA------------PNLLLAAA
PKOX-2003164        NPNVKALLSIGGGGSSS----SAFSSMVSSSSSRSAFINSSIALASQYNFDGLDLDWEYPTSTTDMSNLGTLLQEWRSSA------------PNLLLAAA
PYHZ-2075391        NPNVKALLSIGGGGSSS----SAFSSMVSSSSSRSAFISSSIALASQYNFDGLDLDWEYPTSTTDMSNLGTLLQEWRSSA------------PNLLLAAA
PKOX-2008285        NPNVKTLLSIGGGSSNS----TAFSSMVSASSSRSAFITSSVALARQYNFDGLDLDWEYPTTSTDMSNLGLLFEEWRSYA------------PDLLLAAA

                           210       220       230       240       250       260       270       280       290       300
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      VFYSNNY---Y-----------------------------------------------------------------SVLYPVSAVASSLDWVNLMAYDFY
ANS10044.1_Class_V  VYFNSTISLIG-----------------------------------NG----------------------------PRSYPVRAIRKYLDWASPMCFDYN
AAM18075.1_Class_V  V---PNLP--Y-----------------------------------IE---------------------------KNVEYPGEAIKQNLDWVNLVCYDFY
UGNK-2002386        VKINVTVL--N-----------------------------------GG----------------------------AGKYPVQAIDKCLDWINIMSYDYH
DFHO-2055341        VKFNVTVL--Y-----------------------------------GG----------------------------AGTYPVQAIAKYLDWINIMSYDYS
QVMR-2017494        VKYNVTVL--Y-----------------------------------EG----------------------------AGTYPVKAIAKYLDWINLMEYDLH
EEAQ-2005609        VKYNVTVL--F-----------------------------------GG----------------------------AGTYPVRAIARYLDWINIMAYDLH
```

```
EEAQ-2005608        VKYNVTVL--F-----------------------------------GG----------------------------AGTYPVRAIARYLDWINIMAYDLH
EEAQ-2086905        VHFNATRE--D-----------------------------------IS----------------------------VTTYPVAAIAQNLDWVNLMEYDLH
JVSZ-2002787        VKYNVTIL--Y-----------------------------------GG----------------------------AGTYPAAIMSKYLDFVGLMEFDLH
JVSZ-2002788        VKYNVTIL--Y-----------------------------------GG----------------------------AGTYPAAIMSKYLDFVGLMEFDLH
BEGM-2102306        VHFNVTRV--D-----------------------------------TE----------------------------VTTYPVVAIAQNLDWVNLMEYDLH
XDDT-2003802        VYYTCNY---Y-----------------------------------TG-----------------------------NTYPVASISDNLDWVNIMTYDFW
ZXJO-2072539        VYFTCNY---Y-----------------------------------TG-----------------------------NKYPVASISDNLDWVNIMTYDFW
WQML-2007180        VYFTCNY---W-----------------------------------TG-----------------------------NTYPVASISDNLDWINIMAYDFW
VVRN-2096647        VYFTPNY---Y-----------------------------------TG-----------------------------YSYPVASISDNLDWINIMTYDFW
YJJY-2075390        VYFTCNY---Y-----------------------------------TG-----------------------------FTYPVPSISDNLDWVNIMAYDFW
YJJY-2076688        VKYNVTVMSSSANAIVGLSSHINTTITFQDQNGIVSTHSINATVSSGG--------TNATLSSGRSNISVPTSSSIRATYPVQAIAKSLDWVNIKAFDYH
YJJY-2011496        VKYNVTVV--F-----------------------------------GG----------------------------AGTYPVKSIAQNLDWINIMSYDLH
YQEC-2075825        VYFTCNY---Y-----------------------------------TG-----------------------------FTYPVASISDNLDWVNIMAYDFW
PKOX-2097106        VYYSPSIK--W-----------------------------------VG----------------------------GGSYPISSIKSNVDWVNIMTFDYH
PYHZ-2073439        VYYSPSIK--W-----------------------------------VG----------------------------GGSYPISSIKSNVDWVNIMTFDYH
PYHZ-2003253        VYYSATLQ--Y-----------------------------------VG----------------------------GGTYPISSIKANLNFVNIMAFDYH
PYHZ-2003251        VYYSATLQ--Y-----------------------------------VG----------------------------GGTYPISSIKANLNFVNIMAFDYH
WAFT-2014739        VYYSSKLQ--Y-----------------------------------VG----------------------------EGSYPIASITANLDWVNIMAFDYH
WAFT-2068592        VYYAATIV--YD----------------------------------VG---------------------------IQPTYPIDSIAQNLDWVNIMSYDYH
ENQF-2017435        VYYASSII--YD----------------------------------VG---------------------------NEPAYPIDSIASNLDWVNIMTYDYH
ULKT-2011129        VYYSSKLQ--F-----------------------------------VG----------------------------TGSYPIDSIASNLDWVNIMAFDYH
CBAE-2060579        VYYSARLG--Y-----------------------------------VG----------------------------IRSYPIGSIVANLDWVSAMAFDYH
GTUO-2025296        VYYSARLE--Y-----------------------------------IG----------------------------EGSYPIGSIAANLDWVSIMAFDYH
GKAG-2017355        VYYSARLE--Y-----------------------------------IG----------------------------EGSYPIGSIAANLDWVSIMAFDYH
ABIJ-2038934        VYYSSSVVDLLT----------------------------------AG---------------------------LSYSYPVATLAATLDWANIMCYDYH
LGDQ-2011793        VYYSSAVVD-WTL---------------------------------PG---------------------------TSYSYPVATLTATLDWANIMCYDYH
YQEC-2012943        VKYNVTVMSSSANAIVWPSSHINNTINLHDQNGTVSSHGINASIPSGGTNATSSSGRLINTGILPSNDSAPASSSTPATYPVQAIAKSLDWVNIKAFDYH
VVRN-2022226        AKYNVTV---S-----------------------------------SA----------------------------TATYPVAAMARNLDWINIKAMDYH
YOWV-2000206        VYFTPLY---W-----------------------------------TG-----------------------------ATYPVQSISATLNWVNIMTYDYW
XXHP-2013082        VYFTPLY---W-----------------------------------TG-----------------------------ATYPVQSISATLNWVNIMTYDYW
WGTU-2072225        VYFTPNY---W-----------------------------------TG-----------------------------ATYPVQSIAASLNWVNIMTYDYW
PKOX-2003165        VYYTNPLPS-W-----------------------------------MG----------------------------GASYPTSSIQSNLDWVLVMSYDLH
PKOX-2003164        VYYTNPLPS-W-----------------------------------MG----------------------------GASYPTSSIQSNLDWVLVMSYDLH
PYHZ-2075391        VYYTNPLPS-W-----------------------------------MG----------------------------GASYPTSSIQSNLDWVLVMSYDLH
PKOX-2008285        VYYSNPLAS-W-----------------------------------LG----------------------------GASYPTSSIQSSLDWVLVMSYDFH

                            310       320       330       340       350       360       370       380       390       400
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
3AQU_1|Class_V      --GPGWS--RVTGPPAALF--DPSNAGPSGDAGTRSWIQAGLPAKKAVLGFPYYGYAWRLTN-ANSHSYYAPTTGAAISP--D----GSIGYGQIRKFIV
ANS10044.1_Class_V  --G-AWA--NETGFNAALY--DPNSN-ISTKYGIGSWIGSGVPAEKLVMGLPLYGRAWELKD-PNDHGVGAKAVGPAVDT--D----GSMDYDEILVFNK
AAM18075.1_Class_V  TPAEGSSP-DYTGPSSALF--NRRSNSLSAHFGIESWTKY-VPAKKIVFGIPFHGWAWTLAN-AAEHDMFSKATGSCV---------WPMDYSEVKSHRR
UGNK-2002386        --G-AWDP-TQTGEHTALYEVDPSAE-MNTNYGITQWLGAGLPPAKAVMGLAMYGRNWLLKD-ASDHGVGAPAISAGPTTTVG---AGVYTYDEVRQFVR
DFHO-2055341        --G-AWDP-NETGEHTALYKSDPSAT-INTNYGITQWLAAGLPPAKAVMGLAMYGRTWFLKD-ASNHGVGAPASSAGPTTAEG---VGVYTYDDIQQFVH
QVMR-2017494        --G-SWDP-TQTGEHTALYASDPSDP-LNINYGVNHWLAAGLPPSRATMGLAMYGRSWYLKN-ASENGVGAPAIAGGPN---D----GTYTYDKIQQFKK
EEAQ-2005609        --G-SWEP-SKTGEHTALYDSDPAGV-LNVNYGITQWLAAGLPPVRAALGLAMYGHSWYLLN-ASETGVDAPARIGGT----------TYTFDEIQQFKT
EEAQ-2005608        --G-SWEP-SKTGEHTALYDSDPAGV-LNVNYGITQWLAAGLPPVRAALGLAMYGHSWYLLN-ASETGVDAPARIGGT----------TYTFDEIQQFKT
EEAQ-2086905        --G-SWEP-QKTGEHTALY--DTGSQ--SIDYGVTNWLAAGLPTEKAVVGLAMYGRTWYLQD-ASQNGVGAAATGAGPN---G----GEYTYDEIQQFNK
JVSZ-2002787        --G-SWE--SVTGIHTAIRSSNPSDV-LIADYGIRQWLTAGLSASRALFGMAMYGRTWILQNPSTDNGIGAAAVGAGPN---G----GVLTYDEIQQFKT
JVSZ-2002788        --G-SWE--SVTGIHTAIRSSNPSDV-LIADYGIRQWLTAGLSASRALFGMAMYGRTWILQNPSTDNGIGAAAVGAGPN---G----GVLTYDEIQQFKT
BEGM-2102306        --G-SWEP-QKTGEHTALY--DTGA--LSIDYGVTNWLAAELPTDKAVMGLAMYGRTWYLQD-ASENGVGAVANGAGPN---G----GEYTYDQIKQFNV
XDDT-2003802        ----SWPPHNKTGIHTALYKSDDSQPNVSADGGVSAWLSTGLPPEKAMLGLASYGLTWYLATPATSNGIGAAATGAGP----------VYSYDAAQNFAK
ZXJO-2072539        ----SWQP-TKTGIHTALYKADNTQPTVSADGGVSAWLSAGLPPEKAMLGLASYGLTWYLKTPATSNGIGAAATGAGP----------IYSYDAAQNFAK
WQML-2007180        ----SWQP-TKTGVHTALYKSDNTQPTVSADGGVSQWLSAGLPPDKAMLGLATYGLTWYLKTPATSNGIGAAATGAGP----------IYSYDAVQSFIN
VVRN-2096647        ----SWQS-TKTGIHTALYKSDNTQPTVSADGGVTQWLNAGLPPEKAMLGLASYGLTWYLKTPATSNGIGAAATGFGP----------IYSYDAIQSFIK
YJJY-2075390        ----SWQP-TKTGIHTALYKSDNTQPTVSADGGVTQWLSAGLPPEKAMLGLATYGLTWYLKTPATSNGIGATATGAGP----------IYSYDAVQSFIK
YJJY-2076688        --G-TWEP-TKTGEHTSLFDTNPNQT-ATTDYGVRQWLAAGLPLKQAVLGLATYGRTWILQD-PTQNIVGVPAMSAGPG---D----GIYTYDEIKAFVN
YJJY-2011496        --G-GWE--AETGEHTALFDSTELQT-FTANFGIQHWLAAGLPPHRAALGLASYGHLWYLSD-PTKNGVGAPSSGGAP----------TSTFTEIQLAVQ
YQEC-2075825        ----SWQP-MKTGIHTALYKSDNTQPTVSADGGVTQWLSAGLPPDKAMLGLATYGLTWYLKTPATSNGIGATATGAGP----------IYSYDAVQSFIK
PKOX-2097106        --G-SWEP-NNTGAHTAVY--DPSGG-VSTDSGISKWLDAGLPSQQGVLGLAMYGRSWNLKN-SSDHGIGAAAVGV-PS---D----STPTYSTIAKFIT
PYHZ-2073439        --G-SWEP-NNTGAHTAVY--DPSGG-VNTDSGISKWLDAGLPSQQGVLGLAMYGRSWNLKN-SSDHGIGAAAVGV-PS---D---STPTYSTIAKFIT
PYHZ-2003253        --G-SWEP-NKTGVHTAVF--DPSSG-VNTDSGITKWLNAGLPGSKAVMGLAMYGKEWNLKN-SSDHGIGSPAVGV-PS---D----STPIFSSIVKLLD
PYHZ-2003251        --G-SWEP-NKTGVHTAVF--DPSSG-VNTDSGITKWLNAGLPGSKAVMGLAMYGKEWNLKN-SSDHGIGSPAVGV-PS---D----STPIFSSIVKFIK
WAFT-2014739        --G-SWEP-HTTGVHTALF--DAHSP-INTDHGIATWIGAGLPPEKAVLGLAMYGHSWILQN-ESDNGIGAPAVGPGPG---D----ATPVFSDIVKFIL
WAFT-2068592        --G-SWEP-TKTGEHTALY--DPNSD-LSTSFGVASWLNAGLSSKQAVLGLAMYGRSWILKNPSSDTGVGAAAVGGGID---G----GTPVYYQIVDLIN
ENQF-2017435        --G-SWEP-NKTGEHTALY--DPNSD-ISTSFGVGSWLNAGLPSKQAVLGLATYGRSWILKS-TSDTGVGAAAVGGGID---D----GTPVYYQIVDLIN
ULKT-2011129        --G-SWEP-HITGVHTALF--DPKSP-INTDRGIASWIGAGLPSEKGVLGLAMYGHSWILQN-ESDNGIGAPAAGPGPG---D----ATPIFSDIVKFIL
CBAE-2060579        --G-SWEP-HRTGAHTALF--DPCSK-INTDCGITTWINAGLPPEKVVLGLAMYGRSWILQN-ELDNGIGAPALGPGPG---N----ATKIFLDIVEFIC
GTUO-2025296        --G-SWEP-HTTGVHTALF--DPRSK-INTDRGIATWINAGLPPEKVVLGLAMYGRSWILQN-ELDNGIGAPALGPGPG---T----ATTIFLEIAKFIW
GKAG-2017355        --G-SWEP-HTTGVHTALF--DPRSK-INTDRGIATWINAGLPPEKVVLGLAMYGRSWILQN-ELDNGIGAPALGPGPG---T----ATTIFLEIAKFIW
ABIJ-2038934        --G-SWEP-SITGEHSAL---NDSSTQLSSSYGIQSWLNAGLPPDKAVMGLASYGRTWLLANSSLDHRVGAAARGPGPAMAGGSGEVGVLFFDEIEEFVR
LGDQ-2011793        --G-GWEP-SITGAHSAL---DDSSTRLSSSYGIQSWLDAGLPPHKAVMGLALYGRTWLLANSSRDHGIGAAARGPGPAMAGGSAEAGVLFFDEIEAFVR
YQEC-2012943        --G-TWEP-TKTGEHTSLFDPNPNQT-ATTDYGVRRWLAAGLPQTQAVLGLAMYGRTWILQD-PTQNIVGVPAMSAGPG---D----GIYTYDEIKAFVN
VVRN-2022226        --G-AWEP-RMTGEHTSLFSTDPN---LTTDYGVRQWLVAGLPPAQVVLGLAMYGRTWYLQD-SMQNSVGEAAVRAGPS---D----GIYTYDEIKAFVN
YOWV-2000206        ----TWQP-SKTGIHTALYESDSTQSSISTDSGVSAWLTAGLPPSQAALGLASYGVTWNLQS-ASSNGIGAPATGFGSS--------MTLSYDQIQSFIQ
XXHP-2013082        ----TWQP-SKTGIHTALYESDSTQSSISTDSGVSAWLTAGLPPSQAALGLASYGVTWNLQS-ASSNGIGAPATGFGSS--------MTLSYDQIQSFIQ
WGTU-2072225        ----TWQP-TKTGIHTALYESDSTQPTVSTDNGVTAWLNAGLPSTQAALGLTSYGVTWILQS-ASSNGIGAPATDFGSP--------MTLSYDQIQSFIQ
PKOX-2003165        --G-SWE--NYTNNPAPLY--DPSNPTINGDYGFSQWLQY-LAMRKLVYGFPAYGYTWNLSN-PSNHGTGAPASGSYTSL--G----ETPTYASIATFLT
PKOX-2003164        --G-SWE--NYTNNPAPLY--DPSNPTINGDYGFSQWLQY-LAMRKLVYGFPAYGYTWNLSN-PSNHGTGAPASGSYTSL--G----ETPTYASIATFLT
PYHZ-2075391        --G-SWE--SYTNNPAPLY--DPSNPTVNGDYGFSQWLQY-LAMRKLLYGFPAYGYTWNLSN-PSNHGTGAPASGSYTSL--G----ETPSYARIATFLT
PKOX-2008285        --G-SWE--SYTNPPAALY--DSSNPTVNGNYGFTHWLQS-LPMGKLLYGIPAYAHAWNLSN-PSNHGTGAPAIGSSTSL--G----DTPTYAQIVSFLS

                            410       420       430       440       450       460       470       480       490
                    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|..
3AQU_1|Class_V      -DNGATT-VYNSTVVGD-YCYAG--------TNWIGYDDNQSIVTKVRYAKQRGLLGYFSWHVGADDNSGLSRAASQAWDATTATTRTIQKV
ANS10044.1_Class_V  -DTGAKV-VYDEVAVSF-YSYSG--------TTWIGYDDGPSITKKVQFARSMGLKGYFFWAIGKDKDWTISKQASNAWGY----------
AAM18075.1_Class_V  -ARQCEYNIPNDPRFGIAYAYVE--------TTWIGYESESTIPVKIKFAKQNQLLGYFAWNISGDDDVNTLANAAQSNWQ----------
UGNK-2002386        -NRNATC-EDDAITKSG-YCYGD-TEGG---VLWVGYDDKHTIDIKVKYAKSKNLRGYFFWSLGSD--------------------------
DFHO-2055341        -ESGATC-EDDAITKSA-YCYGQ-TEGG---VLWVGYDDKHTIDTKVKYAKSKNLRGYFFW-------------------------------
QVMR-2017494        -ETNAVC-KDDLTTKSA-YCYGK-MTNG---TLWVGYDDTNTIAMKVDYLKSKKLRGYAFW-------------------------------
EEAQ-2005609        -GSNATC-RDDAVTKSA-YCYGN-TKDG---ILWVGYDDTNTIAMKVEYLKSRNLRGYSFW-------------------------------
EEAQ-2005608        -GSNATC-RDDAVTKSA-YCYGN-TKDG---ILWVGYDDTNTIAMKVEYLKSRNLRGYSFW-------------------------------
EEAQ-2086905        -ESNAVC-QVDGTTKCA-YCYGS-TSNG---ILWVGYDNTDTIATKVSYLKEKNLRGYFFWSLGADQNGDLSKQASQS--------------
JVSZ-2002787        -ENGAIC-QDDSVTMAS-YCYAT-TATG---TAWVGYDETTAISMKVSYLKSMALRGYFFW-------------------------------
JVSZ-2002788        -ENGAIC-QDDSVTMAS-YCYAT-TATG---TAWVGYDETTAIAMKVKYLKRKGLRGYFFW-------------------------------
BEGM-2102306        -ESNAVC-QVDDTTKCA-YCYGS-TSNG---TLWVGYDNTDTIATKVSYLKDKS-RGYFFWSLSGDQNGDLSKQASQS--------------
XDDT-2003802        -SSGVTCQRDSSVSVGV-YCYGE-VGSQ---TLWVGYDDATSIAAKVNYLKGRQMGGYFFWRLGLDANGVLAAQAS----------------
ZXJO-2072539        -SGNTTCRRDSTVSVGV-YCYGK-VGSQ---TLWVGYDDATSITTKVQYLKGRQMGGYFFWRLGLDANGVLAAQAS----------------
WQML-2007180        -SSGVTCRRDSTVTVAV-YCYGK-VGSQ---TLWVGYDDATSIAAKVNYLKGKHMGGYTFWRLGLDANGVLAAQAS----------------
VVRN-2096647        -SSTPTCKRDSTVSVGM-YCYGK-SGNQ---TLWVGYDDATSIATKVNYLKGKQMGGYFFWRLGFDSKNVLATKAS----------------
YJJY-2075390        -SSSATCRRDSSVSVGV-YCYGK-VGSQ---TLWVGYDDATSIAAKVNYLKGKHMGGYTFWRLGFDSNNVLAAQAS----------------
YJJY-2076688        -SSKAYC-RDDVVTKAA-FCYGKLSSGSSSNCLWVGFDDPRTIATKVRYLKAKKLRGYFLWSLGFD--------------------------
YJJY-2011496        -TNNGTC-AHDAVTRSA-YCFWT-TATSPPQTLWAGFDDTVTIAAKVRYLKSLHLRGYSFWSVGGDT-------------------------
YQEC-2075825        -SSSATCRRDSSVTVAV-YCYGK-VGSQ---TLWVGYDDATSIAAKVNYLKGKHMGGYTFWRLGFDSNNVLAAQAS----------------
```

```
PKOX-2097106        -DNDATV-VYDNSTVTA-YCYSQ--------TSWIGFDNNLSIAGKVKYLKSKLLKGYFFWTSSYDDNKFTLSRTAAYA-------------
PYHZ-2073439        -DNDATV-VYDNSTVTA-YCYSQ--------TSWIGFDNNLSIAGKVKYLKSKLLKGYFFWTSSYDDNKFTLSRTAAYA-------------
PYHZ-2003253        SSSHASL-VFDNTTLCS-YCYSG--------TSWIGFDDRVSIAAKV---------------------------------------------
PYHZ-2003251        -QNHAAL-VFDNTTLCS-YCYSG--------TSWIGFDDRISIAAKVGFLKSKGLQGYFFWTASYDDKKYTLSTTAAQS-------------
WAFT-2014739        -DKHATV-VFDKATVTT-YCYAD--------NLWIGFDDPVSIFKKVRYAKQKGLKGYFFWTASFDS-------------------------
WAFT-2068592        -SSNATV-VFDDTTASV-YTYSG--------TIWIGYDNPRSISSKVAFLNSKGLLGYFFWTVSFD--------------------------
ENQF-2017435        -SSNATV-VFDETTSSV-YTYSG--------TIWIGYDDPRSISSKVAFLKSKGLLGYFFWTVSFD--------------------------
ULKT-2011129        -EKNATV-VFDRTTVTT-YCYAD--------KLWIGFDDPVSIFEKVRYAKEKGLKGYFFWTSSFDSKWTLSSSAYTA--------------
CBAE-2060579        -KKHARV-VFDPTTVTT-YCYAH--------NLWVGFDDSVSIFEKVRYAKHKGLKGYFFWTVSFD--------------------------
GTUO-2025296        -KKHARV-VFDPTTVTT-YCYAH--------NVWIGFDDPVSIFEKVRYAKQKGLKGYFFWTVSFD--------------------------
GKAG-2017355        -KKHARV-VFDPTTVTT-YCYAH--------NLW-QFQ-----SEKVRYAKQKGLKGYFFWTVSFD--------------------------
ABIJ-2038934        -KNNATV-VDDAVTGSA-YAFAG--------DLWVGFDNKVSIEKKVTYLKSKEMRGYFFWTSSFD--------------------------
LGDQ-2011793        -EKNATV-VDDAVTGCA-YAFAG--------DFWVGFDNKVSIEKKVTFLKSKDMRGYFFWTSSFD--------------------------
YQEC-2012943        -NSKAYC-RDDVVTKAA-FCYGKLRSGSSSNCLWVGFDNPRTIATKVRYLKAKKLRGYFLWSLGFDKDGDLSKQAS----------------
VVRN-2022226        -SSKAYC-RDDAVTKAS-FCYGRQLSSGNSMVLWAGFDSPRTIATKVRYLNTKKLRGYCFWSLNFD--------------------------
YOWV-2000206        -ANSATCQRDSTITVAS-YCYGS-SGSD---ILWVGFDDAISIAAKVEYLKSKQLGGYFFWRLGFDQNNVLASQASTS--------------
XXHP-2013082        -ANSATCQRDSTITVAS-YCYGS-SGSD---ILWVGFDDAISIAAKVEYLKSKQLGGYFFWRLGFDQNNVLASQASTS--------------
WGTU-2072225        -SSSTTCQRDSTVTVAS-YCYGN-SGSN---VLWVGFDDATSIAAKVKYLKSKQLAGYFFWRLGFD--------------------------
PKOX-2003165        -STGATI-VHNATSVST-YAYSG--------RAWVGYDDAWSITRKVNYLKSKGLRGYGFWT------------------------------
PKOX-2003164        -STGATI-VHNATSVST-YAYSG--------RAWVGYDDAWSITRKVNYLKSKGLRGYGFWT------------------------------
PYHZ-2075391        -STGATI-VHNATYVST-YAYSG--------RAWVGYDDAWSITRKVNYLKSKGLRGYGFWT------------------------------
PKOX-2008285        -SSGATV-VHNTSSVST-YAYSG--------ITWVGYDDSWSITRKVNYLKSKG-RGYGFWTVLSDDSNFTLA-------------------
```

Figure S8. Phylogenetic tree of GH18 family chitinases identified in algae phyla Rhodophyta and Ochrophyta

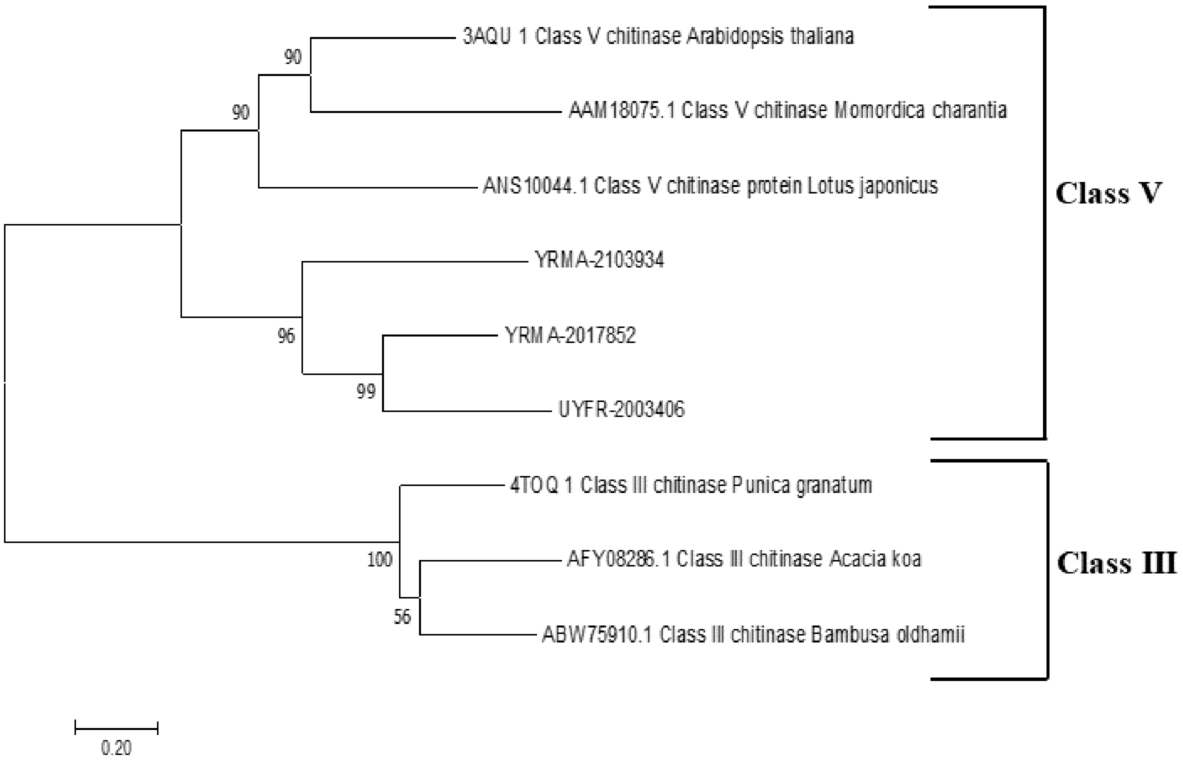


Figure S9. Phylogenetic tree of GH18 family chitinases identified in green algae.

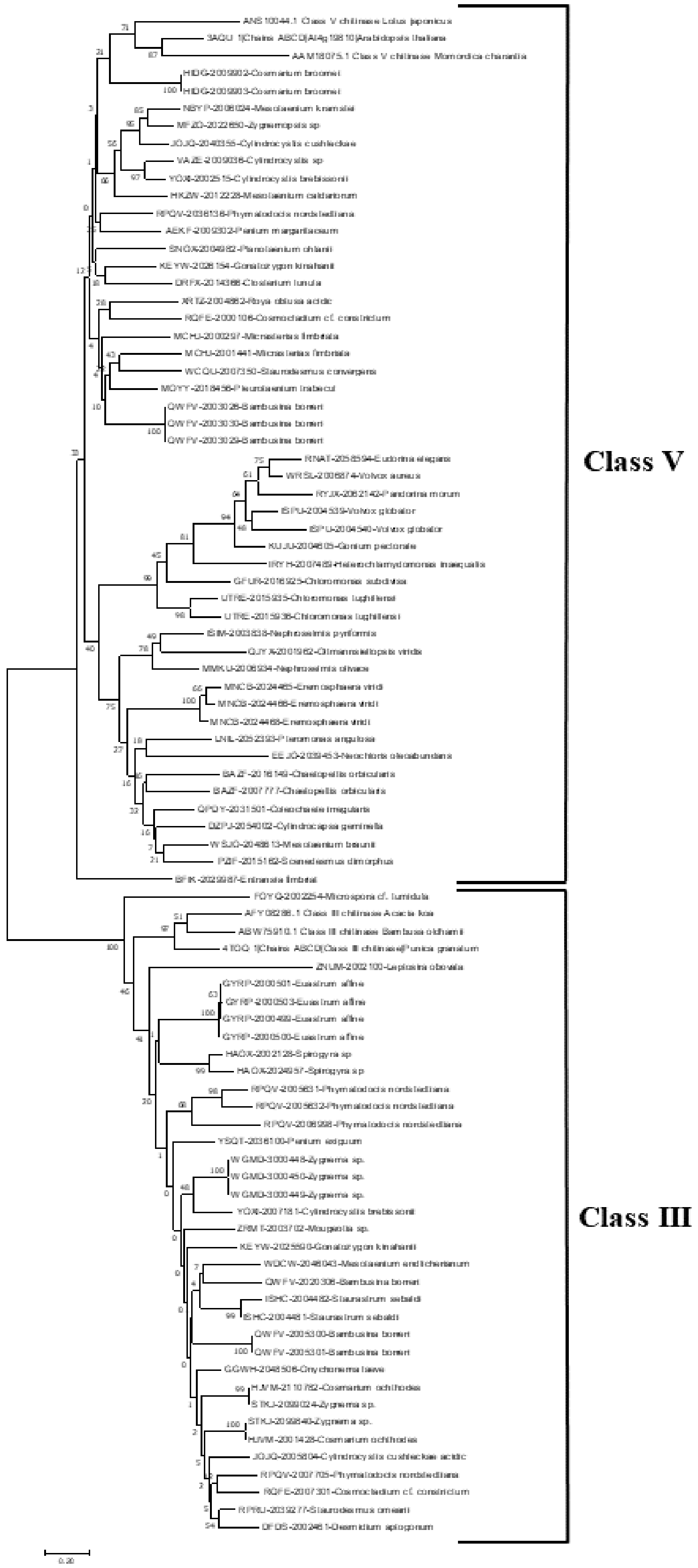


Figure S10. Phylogenetic tree of GH18 family chitinases complete sequences identified in bryophytes

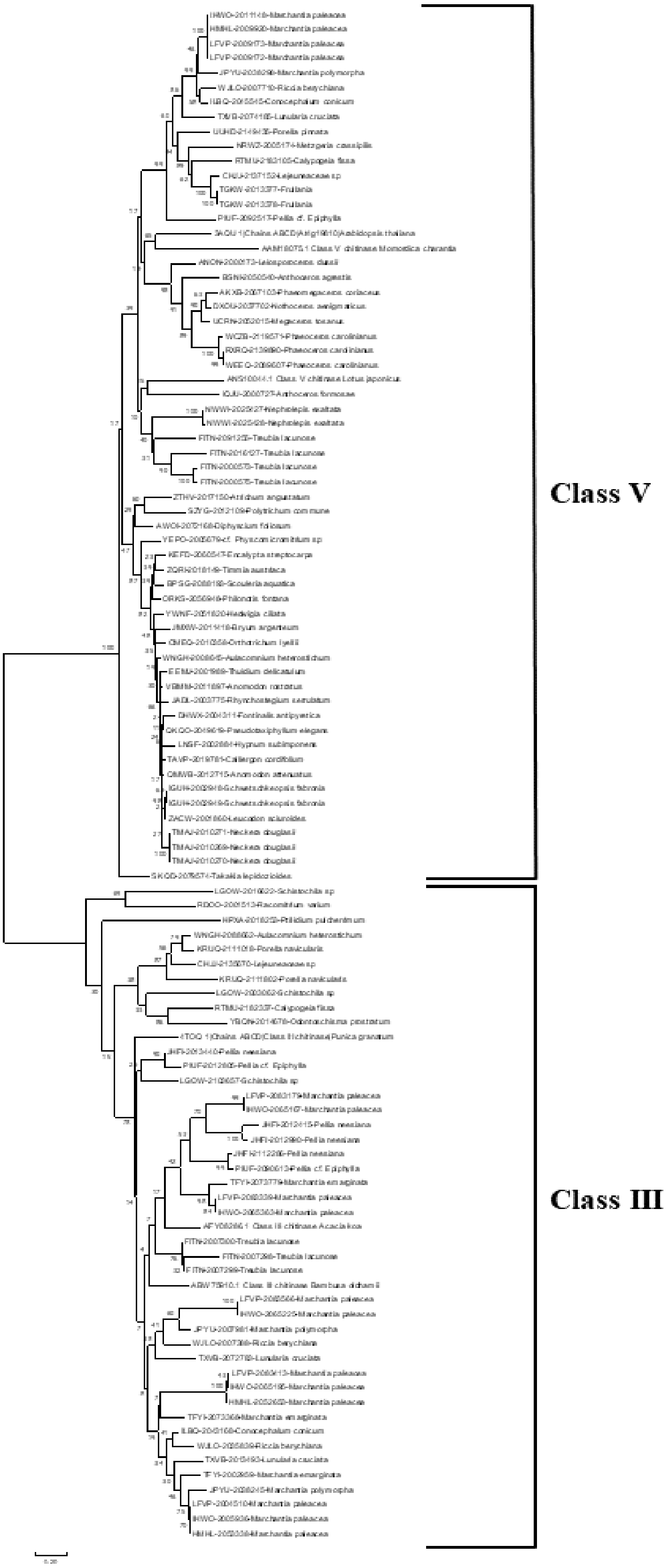
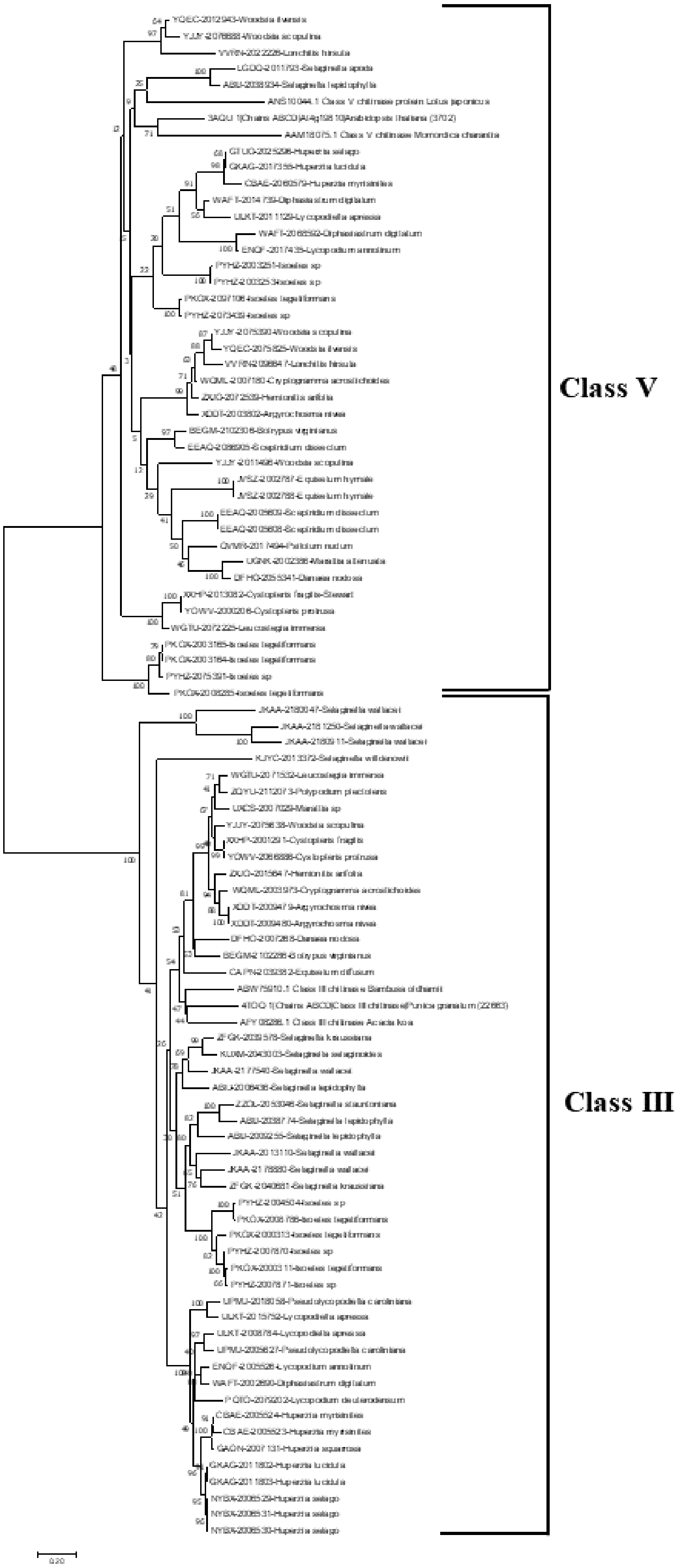


Figure S11. Phylogenetic tree of GH18 family chitinases complete sequences identified in pteridophytes



CONCLUSÃO

As análises conduzidas neste trabalho revelaram a elevada diversidade de quitinases da família GH18 em algas, briófitas e pteridófitas. No total, foram encontradas 648 possíveis quitinases da família GH18, distribuídas em algas, briófitas e pteridófitas. De um modo geral, as quitinases demonstraram ter os domínios catalíticos altamente conservados em todos os grupos analisados. Mas algumas quitinases apresentaram deleções ou substituições em resíduos de aminoácidos fundamentais para a catálise, o que pode interferir na função quitinolítica. A árvore filogenética separou claramente a classe III da classe V, pois apesar de pertencerem a mesma família, essas classes apresentam baixa similaridade entre si. As estruturas tridimensionais de algas verdes, briófitas e pteridófitas demonstraram ter arranjos semelhantes entre si. A identificação e caracterização de quitinases em grupos mais ancestrais das plantas terrestres é importante para compreender a dinâmica dessas proteínas nesses grupos e possibilitar a comparação com quitinases de gimnospermas e angiospermas. Os resultados deste estudo podem também ser úteis para o fornecimento de dados para estudos futuros relacionados a quitinases da família GH18 em plantas.

REFERÊNCIAS

ADAMS, D. J. Fungal cell wall chitinases and glucanases. **Microbiology**, v. 150, n. 7, p. 2029–2035, 2004. https://doi.org/10.1099/mic.0.26980-0

AMEND, A. et al. Fungi in the marine environment: Open questions and unsolved problems. **mBio**, v. 10, n. 2, p. 1–15, 2019. https://doi.org/10.1128/mBio.01189-18

ANDERSON, J. T.; WILLIS, J. H.; MITCHELL-OLDS, T. Evolutionary genetics of plant adaptation. **Trends Genet**, v. 27, n. 8, p. 258–266, 2011. https://doi.org/10.1016/j.tig.2011.04.001

ARIE, M. et al. Characterization of a basic chitinase which is secreted by cultured pumpkin cells. **Physiologia Plantarum**, v. 110, n. 2, p. 232–239, 2000. https://doi.org/10.1034/j.1399-3054.2000.110213.x

BADARIOTTI, F. et al. The phylogenetically conserved molluscan chitinase-like protein 1 (Cg-Clp1), homologue of human HC-gp39, stimulates proliferation and regulates synthesis of extracellular matrix components of mammalian chondrocytes. **Journal of Biological Chemistry**, v. 281, n. 40, p. 29583–29596, 2006. https://doi.org/10.1074/jbc.M605687200

BASKARAN, X-R, et al. A review of the use of pteridophytes for treating human ailments. **Journal of Zhejiang University: Science B**, v. 19, n. 2, 2018. https://doi.org/10.1631/jzus.B1600344

BENNICI, A. Origin and early evolution of land plants. **Communicative & Integrative Biology**, v. 1, n. 2, p. 212–218, 2008. https://doi.org/10.4161/cib.1.2.6987

BHATTACHARYA, D.; YOON, H. S.; HACKETT, J. D. Photosynthetic eukaryotes unite: Endosymbiosis connects the dots. **BioEssays**, v. 26, n. 1, p. 50–60, 2004. https://doi.org/10.1002/bies.10376

BLACKWELL, W. H. Two Theories of origin of the land-plant sporophyte: Which Is Left Sanding? **The Botanical Review**, v. 69, n. 2, p. 125–148, 2003. Diponível em: link.gale.com/apps/doc/A112542838/AONE?u=capes&sid=AONE&xid=004c651d. Acesso em: 15 nov. 2020.

BRAMELD, K. A. et al. Substrate assistance in the mechanism of family 18 chitinases: Theoretical studies of potential intermediates and inhibitors. **Journal of Molecular Biology**, v. 280, n. 5, p. 913–923, 1998. https://doi.org/10.1006/jmbi.1998.1890

BRZEZINSKA, M. S. et al. Chitinolytic microorganisms and their possible application in environmental protection. **Current Microbiology**, v. 68, n. 1, p. 71–81, 2014. https://doi.org/10.1007/s00284-013-0440-4

CAMPANA-FILHO, S. P. et al. Extração, estruturas e propriedades de α- e β-quitina. **Quimica Nova**, v. 30, n. 3, p. 644–650, 2007. https://doi.org/10.1590/S0100-40422007000300026

CAO, J.; TAN, X. Comprehensive analysis of the chitinase family genes in tomato (*Solanum Lycopersicum*). **Plants**, v. 8, n. 3, 2019. https://doi.org/10.3390/plants8030052

CASANOVA, M. T.; NAIRN, L. C. Macroalgae, charophytes and bryophytes. *In*: CAPON, S., JAMES, A., REID, M. **Vegetation of Australian Riverine Landscapes**, Austrália: CSIRO, 2016. cap. 5, p. 67–103. Disponível em: https://www.researchgate.net/publication/296396211. Acesso em: 15 nov. 2020.

CAVALIER-SMITH, T. The origins of plastids. **Biological Journal of the Linnean Society**, v. 17, n. 3, p. 289–306, 1982. https://doi.org/10.1111/j.1095-8312.1982.tb02023.x

CHEN, L. et al. Enhanced nematicidal potential of the chitinase pachi from *Pseudomonas aeruginosa* in association with Cry21Aa. **Scientific Reports**, v. 5, p. 1–11, 2015. https://doi.org/10.1038/srep14395

CHRISTENHUSZ, M. J. M.; BYNG, J. W. The number of known plants species in the world and its annual increase. **Phytotaxa**, v. 261, n. 3, p. 201–217, 2016. https://doi.org/10.11646/phytotaxa.261.3.1

CHRISTENHUSZ, M. J. M.; CHASE, M. W. Trends and concepts in fern classification. **Annals of Botany**, v. 113, n. 4, p. 571–594, 2014. https://doi.org/10.1093/aob/mct299

CHUANG, H. H.; LIN, H. Y.; LIN, F. P. Biochemical characteristics of C-terminal region of recombinant chitinase from *Bacillus licheniformis* - implication of necessity for enzyme properties. **FEBS Journal**, v. 275, n. 9, p. 2240–2254, 2008. https://doi.org/10.1111/j.1742-4658.2008.06376.x

COOPER, E. D. Overly simplistic substitution models obscure green plant phylogeny. **Trends in Plant Science**, v. 19, n. 9, p. 576–582, 2014. https://doi.org/10.1016/j.tplants.2014.06.006

COTAS, J. et al. A comprehensive review of the nutraceutical and therapeutic applications of red seaweeds (Rhodophyta). **Life**, v. 10, n. 3, 2020. https://doi.org/10.3390/life10030019

EDREVA, A. Pathogenesis-related proteins: Research progress in the last 15 years. **Gen Appl Plant Physiol**, v. 31, n. 1–2, p. 105–24, 2005. Disponível em: http://www.bio21.bas.bg/ipp/gapbfiles/v-31/05_1-2_105-124.pdf. Acesso em: 12 nov. 2020.

EL-KATATNY, M. H. et al. Characterization of a chitinase and an endo-β-1,3-glucanase from *Trichoderma harzianum* Rifai T24 involved in control of the phytopathogen Sclerotium rolfsii. **Applied Microbiology and Biotechnology**, v. 56, n. 1–2, p. 137–143, 2001. https://doi.org/10.1007/s002530100646

EVERT, R. F.; EICHHORN, S. E. **Raven Biologia Vegetal**. 8ª ed. Rio de Janeiro: Guanabara Koogan, 2014. Dispsonível em: file:///D:/DADOS%20USER/Downloads/LIVRO_2014%20-%20Raven%20Biologia%20Vegetal%208%C3%82%C2%AAEd._%20Fungos%20558.pdf. Acesso em: 10 ago. 2020.

FANG, L. et al. Evolution of the chlorophyta: Insights from chloroplast phylogenomic analyses. **Journal of Systematics and Evolution**, v. 55, n. 4, p. 322–332, 2017.

https://doi.org/10.1111/jse.12248

FRANGEDAKIS et al., 2020. The hornworts: Morphology, evolution and development. **New Phytol**, v. 229, n. 2, p. 735–754. https://doi.org/10.1111/nph.16874

FLEURI, L. F.; SATO, H. H. Produção, purificação, clonagem e aplicação de enzimas líticas. **Química Nova**, v. 28, n. 5, p. 871–879, 2005. https://doi.org/10.1590/S0100-40422005000500026

FRAHM, J.-P. et al. Manual of tropical bryology. **Bryophyte Diversity and Evolution**, v. 23, n. 1, p. 1–200, 2003. https://doi.org/10.11646/bde.23.1.1

FUKAMIZO, T. Chitinolytic enzymes: Catalysis, substrate binding, and their application. **Current Protein & Peptide Science**, v. 1, n. 1, p. 105–124, 2000. https://doi.org/10.2174/1389203003381450

GRAHAM, L. E.; COOK, M. E.; BUSSE, J. S. The origin of plants: Body plan changes contributing to a major evolutionary radiation. **Proceedings of the National Academy of Sciences of the United States of America**, v. 97, n. 9, p. 4535–4540, 2000. https://doi.org/10.1073/pnas.97.9.4535

GUIRY, M. D. How many species of algae are there? **Journal of Phycology**, v. 48, n. 5, p. 1057–1063, 2012. https://doi.org/10.1111/j.1529-8817.2012.01222.x

HAMID, R. et al. Chitinases: An update. **J Pharm Bioallied Sci**, v. 5, n. 1, p. 21–29, 2013. https://doi.org/10.4103/0975-7406.106559

HAN, P. et al. Identification and characterization of a novel chitinase with antifungal activity from 'Baozhu' pear (*Pyrus ussuriensis* Maxim.). **Food Chemistry**, v. 196, p. 808–814, 2016. https://doi.org/10.1016/j.foodchem.2015.10.006

HE-NYGRÉN, X. et al. Illuminating the evolutionary history of liverworts (Marchantiophyta) - Towards a natural classification. **Cladistics**, v. 22, n. 1, p. 1–31, 2006. https://doi.org/10.1111/j.1096-0031.2006.00089.x

HO, B. The liverwort genus Marchantia L.(Marchantiophyta: Marchantiopsida) in Singapore, with a new species record. **Nature in Singapore**, v. 6, p. 187–190, 2013. Disponível em: https://www.researchgate.net/publication/260851139. Acesso em: 10 nov. 2020.

HUTTUNEN, S.; BELL, N.; HEDENÄS, L. The evolutionary diversity of mosses–Taxonomic heterogeneity and its ecological drivers. **Critical Reviews in Plant Sciences**, v. 37, n. 2–3, p. 128–174, 2018. https://doi.org/10.1080/07352689.2018.1482434

INGROUILLE, M.; EDDIE, B. **Plants: Evolution and diversity**. New York: Cambridge University Press, 2006. https://doi.org/10.1017/CBO9780511812972

ISELI, B. et al. Plant chitinases use two different hydrolytic mechanisms. **FEBS Letters**, v. 382, n. 1–2, p. 186–188, 1996. https://doi.org/10.1016/0014-5793(96)00174-3

ISHIZAKI, K. Evolution of land plants: Insights from molecular studies on basal lineages.

Bioscience, Biotechnology and Biochemistry, v. 81, n. 1, p. 73–80, 2017. https://doi.org/10.1080/09168451.2016.1224641

KARLSSON, M.; STENLID, J. Evolution of family 18 glycoside hydrolases: Diversity, domain structures and phylogenetic relationships. **Journal of Molecular Microbiology and Biotechnology**, v. 16, n. 3–4, p. 208–2023, 2009. https://doi.org/10.1159/000151220

KAUR, R. et al. Glycan-dependent chikungunya viral infection divulged by antiviral activity of NAG specific chi-like lectin. **Virology**, v. 526, p. 91–98, 2019. https://doi.org/10.1016/j.virol.2018.10.009

KAYA, M. et al. On chemistry of γ-chitin. **Carbohydrate Polymers**, v. 176, p. 177–186, 2017. https://doi.org/10.1016/j.carbpol.2017.08.076

KESARI, P. et al. Structural and functional evolution of chitinase-like proteins from plants. **Proteomics**, v. 15, n. 10, p. 1693–1705, 2015. https://doi.org/10.1002/pmic.201400421

KHAN, F. I. et al. Chitinase from *Thermomyces lanuginosus* SSBP and its biotechnological applications. **Extremophiles**, v. 19, n. 6, p. 1055–1066, 2015. https://doi.org/10.1007/s00792-015-0792-8

LEGER, R. J. S. et al. Characterization and ultrastructural localization of chitinases from *Metarhizium anisoplia*e, *M. flavoviride*, and *Beauveria bassiana* during fungal invasion of host (*Manduca sext*a) cuticle. **Applied and Environmental Microbiology**, v. 62, n. 3, p. 907–912, 1996. https://doi.org/10.1128/AEM.62.3.907-912.1996

LEWIS, L. A.; MCCOURT, R. M. Green algae and the origin of land plants. **American Journal of Botany**, v. 91, n. 10, p. 1535–1556, 2004. https://doi.org/10.3732/ajb.91.10.1535

LOMBARD, V. et al. The carbohydrate-active enzymes database (CAZy) in 2013. **Nucleic Acids Research**, v. 42, n. D1, p. 490–495, 2014. https://doi.org/10.1093/nar/gkt1178

MARTIN, W. F.; GARG, S.; ZIMORSKI, V. Endosymbiotic theories for eukaryote origin. **Philosophical transactions of the Royal Society of London, Series B, Biological sciences**, v. 370, n. 1678, p. 1–18, 2015. https://doi.org/10.1098/rstb.2014.0330

MARTÍNEZ-CABALLERO, S. et al. Comparative study of two GH19 chitinase-like proteins from Hevea brasiliensis, one exhibiting a novel carbohydrate-binding domain. **FEBS Journal**, v. 281, n. 19, p. 4535–4554, 2014. https://doi.org/10.1111/febs.12962

MASUDA, T.; ZHAO, G.; MIKAMI, B. Crystal structure of class III chitinase from pomegranate provides the insight into its metal storage capacity. **Bioscience, Biotechnology and Biochemistry**, v. 79, n. 1, p. 45–50, 2015. https://doi.org/10.1080/09168451.2014.962475

MATSUI, M. **Correlações entre estrutura química, superestrutura macromolecular e morfologia das blendas e redes poliméricas à base de quitina e poliuretano**. 2007. Tese (Doutorado em Engenharia) - Programa de Pós-Graduação em Engenharia, Universidade Federal do Paraná, Curitiba, 2007. Disponível em: https://acervodigital.ufpr.br/bitstream/handle/1884/16198/MITSUKA%20MATSUI%20.pdf?s equence=1&isAllowed=y. Acesso em: 14 ago. 2020.

MCFADDEN, G. I. Origin and evolution of plastids and photosynthesis in eukaryotes. **Cold Spring Harbor Perspectives in Biology**, v. 6, n. 4, p. 1–9, 2014. https://doi.org/10.1101/cshperspect.a016105

MEDEIROS, S. C. et al. Chitinases as antibacterial proteins: A systematic review. **Journal of Young Pharmacists**, v. 10, n. 2, p. 144–148, 2018. https://doi.org/10.5530/jyp.2018.10.33

MENDEZ-TEJEDA, R.; ROSADO, J. G. A. Influence of climatic factors on *Sargassum arrivals* to the coasts of the Dominican Republic. **Journal of Oceanography and Marine Science**, v. 10, n. 2, p. 22–32, 2019. https://doi.org/10.5897/JOMS2019.0156

MOURA, C. et al. Quitina e quitosana produzidas a partir de resíduos de camarão e siri: avaliação do processo em escala piloto. **Vetor**, v. 16, n. 1–2, p. 37–45, 2006. https://periodicos.furg.br/vetor/article/view/294/85

NEUHAUS, J.-M. et al. A revised nomenclature for chitinase genes. **Plant Molecular Biology Reporter**, v. 14, p. 102–104, 1996. https://doi.org/10.1007/BF02684897

NISHIO, J. N. Why are higher plants green? Evolution of the higher plant photosynthetic pigment complement. **Plant, Cell and Environment**, v. 23, n. 6, p. 539–548, 2000. https://doi.org/10.1046/j.1365-3040.2000.00563.x

OHNUMA, T. et al. A class V chitinase from *Arabidopsis thaliana*: Gene responses, enzymatic properties, and crystallographic analysis. **Planta**, v. 234, n. 1, p. 123–137, 2011a. https://doi.org/10.1007/s00425-011-1390-3

OHNUMA, T. et al. Crystal structure and mode of action of a class V chitinase from *Nicotiana tabacum*. **Plant Molecular Biology**, v. 75, n. 3, p. 291–304, 2011b. https://doi.org/10.1007/s11103-010-9727-z

PANAWALA, L. **Difference Between Aerobic and Anaerobic Respiration**, 2017a. Disponível em: https://www.researchgate.net/publication/316583908_Difference_Between_Aerobic_and_Anaerobic_respiration. Acesso em: 02 dez. 2020.

PANAWALA, L. **Difference Between Bryophytes and Pteridophytes**, 2017b. Disponível em: https://www.researchgate.net/publication/318305486_Difference_Between_Bryophytes_and_Pteridophytes. Acesso em: 02 dez. 2020.

PARK, S. M. et al. Heterologous expression and characterization of class III chitinases from rice (Oryza sativa L.). **Enzyme and Microbial Technology**, v. 30, n. 6, p. 697–702, 2002. https://doi.org/10.1016/S0141-0229(02)00042-X

PATIL, R.; GHORMADE, V.; DESHPANDE, M. Chitinolytic enzymes: An exploration. **Enzyme and Microbial Technology**, v. 26, n. 7, p. 473–483, 2000. https://doi.org/10.1016/S0141-0229(00)00134-4

PITTERMANN, J. The evolution of water transport in plants: An integrated approach. **Geobiology**, v. 8, n. 2, p. 112–139, 2010. https://doi.org/10.1111/j.1472-4669.2010.00232.x

PRYER, K. M. et al. Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. **American Journal of Botany**, v. 91, n. 10, p. 1582–1598, 2004. https://doi.org/10.3732/ajb.91.10.1582

PUTTICK, M. N. et al. The interrelationships of land plants and the nature of the ancestral embryophyte. **Current Biology**, v. 28, n. 5, p. 733- 745, 2018. https://doi.org/10.1016/j.cub.2018.01.063

RAVEN, J. A. et al. Algal evolution in relation to atmospheric CO2: Carboxylases, carbon-concentrating mechanisms and carbon oxidation cycles. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 367, n. 1588, p. 493–507, 2012. https://doi.org/10.1098/rstb.2011.0212

RAVEN, J. A.; GIORDANO, M. Algae. **Current Biology**, v. 24, n. 13, p. 590–595, 2014. https://doi.org/10.1016/j.cub.2014.05.039

SAMOLOV, E. et al. Biodiversity of algae and cyanobacteria in biological soil crusts collected along a climatic gradient in chile using an integrative approach. **Microorganisms**, v. 8, n. 7, p. 1–28, 2020. https://doi.org/10.3390/microorganisms8071047

SARMA, K. et al. A comparative proteomic approach to analyse structure, function and evolution of rice chitinases: A step towards increasing plant fungal resistance. **Journal of Molecular Modeling**, v. 18, n. 11, p. 4761–4780, 2012. https://doi.org/10.1007/s00894-012-1470-8

SEIDL, V. et al. A complete survey of Trichoderma chitinases reveals three distinct subgroups of family 18 chitinases. **FEBS Journal**, v. 272, n. 22, p. 5923–5939, 2005. https://doi.org/10.1111/j.1742-4658.2005.04994.x

SEIDL, V. Chitinases of filamentous fungi: A large group of diverse proteins with multiple physiological functions. **Fungal Biology Reviews**, v. 22, n. 1, p. 36–42, 2008. https://doi.org/10.1016/j.fbr.2008.03.002

SELS, J. et al. Plant pathogenesis-related (PR) proteins: A focus on PR peptides. **Plant Physiology and Biochemistry**, v. 46, n. 11, p. 941–950, 2008. https://doi.org/10.1016/j.plaphy.2008.06.011

SU, Y. et al. Identification, phylogeny, and transcript of chitinase family genes in sugarcane. **Scientific Reports**, v. 5, p. 1–15, 2015. https://doi.org/10.1038/srep10708

TRYON, R. M.; TRYON, A. F. **Ferns and allied plants with Special reference to Tropical America**. New Yor: Springer-Verlag, 1982. https://doi.org/10.2307/1546858

VILLARREAL, J. C. et al. A synthesis of hornwort diversity: Patterns, causes and future work. **Phytotaxa**, v. 9, n. 1, p. 150, 2014. https://doi.org/10.11646/phytotaxa.9.1.8

WILLIS, K. J.; MCELWAIN, J. C. The evolution of plants. **Annals of Botany**, v. 90, n. 5, p. 678–679, 2002. https://doi.org/10.1093/aob/mcf232

YAN, Q.; FONG, S. S. Bacterial chitinase: Nature and perspectives for sustainable bioproduction. **Bioresources and Bioprocessing**, v. 2, n. 1, 2015. https://doi.org/10.1186/s40643-015-0057-5

ANEXOS

ANEXO A - Normas da Revista *Plant Gene*

GUIDE FOR AUTHORS

Your Paper Your Way

We now differentiate between the requirements for new and revised submissions. You may choose to submit your manuscript as a single Word or PDF file to be used in the refereeing process. Only when your paper is at the revision stage, will you be requested to put your paper in to a 'correct format' for acceptance and provide the items required for the publication of your article. **To find out more, please visit the Preparation section below.**

Plant Gene publishes papers that focus on the **regulation, expression, function** and **evolution** of **genes** in plants, algae and other photosynthesizing organisms (e.g., cyanobacteria), and plant-associated microorganisms.

Plant Gene strives to be a diverse plant journal and topics in multiple fields will be considered for publication. Although not limited to the following, some general topics include: Gene discovery and characterization Gene regulation in response to environmental stress (e.g., salinity, drought, etc.) Genetic effects of transposable elements Genetic control of secondary metabolic pathways and metabolic enzymes. Herbal Medicine - regulation and medicinal properties of plant products. Plant hormonal signaling. Plant evolutionary genetics, molecular evolution, population genetics, and phylogenetics. Profiling of plant gene expression and genetic variation Plant-microbe interactions (e.g., influence of endophytes on gene expression; horizontal gene transfer studies; etc.) Agricultural genetics - biotechnology and crop improvement.

Plant Gene encourages submission of novel manuscripts that present a reasonable level of analysis, functional relevance and/or mechanistic insight. Plant Gene also welcomes papers that have predominantly a descriptive component but improve the essential basis of knowledge for subsequent functional studies, or provide important confirmation of recently published discoveries.

Types of article

Full length articles

Organization should be Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgments, and References.

Minireviews

The journal encourages the submission of concise, highly focused review articles summarizing recent progress in very active areas of research involving the analysis of genetics.

Short Communications are brief papers that make a specific well-documented point. In general, a Communication should include no more than four figures and tables. The text will be continuous, with technical and methodological detail printed in the legend to the tables and figures.

Manuscripts that do not meet the general criteria or standards for publication in Plant Gene will be immediately returned to the authors, without detailed review.

Contact details for submission

Articles for Plant Gene should be submitted via the journal's online submission and editorial system at https://ees.elsevier.com/plantgene.

Customer support is available 24/7:

Please use our help site at: https://service.elsevier.com/. Here you will be able to learn more about the online submission and editorial system via interactive tutorials, explore a range of problem solutions via our knowledgebase, and find answers to frequently asked questions. You will also find our 24/7 support contact details should you need any assistance from one of our customer service representatives.

For questions on the submission and reviewing process, please contact the Editorial Office at plantgene@elsevier.com.

Pre-Submission Enquiries

Please note that we do not do pre-submission check of the manuscripts before they are submitted to the Editorial System. Please check if the content of the paper is according to the Aims and Scope of the journal and submit the same in our Editorial System. We will review and will get back to you with the comments.

Submission checklist

You can use this list to carry out a final check of your submission before you send it to the journal for review. Please check the relevant section in this Guide for Authors for more details.

Ensure that the following items are present:

One author has been designated as the corresponding author with contact details:
- E-mail address
- Full postal address

All necessary files have been uploaded:

Manuscript:

- Include keywords
- All figures (include relevant captions)
- All tables (including titles, description, footnotes)
- Ensure all figure and table citations in the text match the files provided
- Indicate clearly if color should be used for any figures in print

Graphical Abstracts / Highlights files (where applicable)
Supplemental files (where applicable)

Further considerations:
- Manuscript has been 'spell checked' and 'grammar checked'
- All references mentioned in the Reference List are cited in the text, and vice versa

- Permission has been obtained for use of copyrighted material from other sources (including the Internet)
- A competing interests statement is provided, even if the authors have no competing interests to declare
- Journal policies detailed in this guide have been reviewed
- Referee suggestions and contact details provided, based on journal requirements

For further information, visit our Support Center.

BEFORE YOU BEGIN

Ethics in publishing

Please see our information pages on Ethics in publishing and Ethical guidelines for journal publication.

Studies in humans and animals

If the work involves the use of human subjects, the author should ensure that the work described has been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. The manuscript should be in line with the Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals and aim for the inclusion of representative human populations (sex, age and ethnicity) as per those recommendations. The terms sex and gender should be used correctly.

Authors should include a statement in the manuscript that informed consent was obtained for experimentation with human subjects. The privacy rights of human subjects must always be observed.

All animal experiments should comply with the ARRIVE guidelines and should be carried out in accordance with the U.K. Animals (Scientific Procedures) Act, 1986 and associated guidelines, EU Directive 2010/63/EU for animal experiments, or the National Institutes of Health guide for the care and use of Laboratory animals (NIH Publications No. 8023, revised

1978) and the authors should clearly indicate in the manuscript that such guidelines have been followed. The sex of animals must be indicated, and where appropriate, the influence (or association) of sex on the results of the study.

Informed consent and patient details

Studies on patients or volunteers require ethics committee approval and informed consent, which should be documented in the paper. Appropriate consents, permissions and releases must be obtained where an author wishes to include case details or other personal information or images of patient and any other individuals in an Elsevier publication. Written consents must be retained by the author but copies should not be provided to the journal. Only if specifically requested by the journal in exceptional circumstances (for example if a legal issue arises) the author must provide copies of the consents or evidence that such consents have been obtained. For more information, please review the Elsevier Policy on the Use of Images or Personal Information of Patients or other Individuals. Unless you have written permission from the patient (or, where applicable, the next of kin), the personal details of any patient included in any part of the article and in any supplementary materials (including all illustrations and videos) must be removed before submission.

Declaration of interest

All authors must disclose any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work. Examples of potential competing interests include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding. Authors must disclose any interests in two places: 1. A summary declaration of interest statement in the title page file (if double-blind) or the manuscript file (if single-blind). If there are no interests to declare then please state this: 'Declarations of interest: none'. This summary statement will be ultimately published if the article is accepted. 2. Detailed disclosures as part of a separate Declaration of Interest form, which forms part of the journal's official records. It is important for potential interests to be declared in both places and that the information matches. More information.

Submission declaration and verification

Submission of an article implies that the work described has not been published previously (except in the form of an abstract, a published lecture or academic thesis, see 'Multiple, redundant or concurrent publication' for more information), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. To verify originality, your article may be checked by the originality detection service Crossref Similarity Check.

Preprints

Please note that preprints can be shared anywhere at any time, in line with Elsevier's sharing policy. Sharing your preprints e.g. on a preprint server will not count as prior publication (see 'Multiple, redundant or concurrent publication' for more information).

Use of inclusive language

Inclusive language acknowledges diversity, conveys respect to all people, is sensitive to differences, and promotes equal opportunities. Content should make no assumptions about the beliefs or commitments of any reader; contain nothing which might imply that one individual is superior to another on the grounds of age, gender, race, ethnicity, culture, sexual orientation, disability or health condition; and use inclusive language throughout. Authors should ensure that writing is free from bias, stereotypes, slang, reference to dominant culture and/or cultural assumptions. We advise to seek gender neutrality by using plural nouns ("clinicians, patients/clients") as default/wherever possible to avoid using "he, she," or "he/she." We recommend avoiding the use of descriptors that refer to personal attributes such as age, gender, race, ethnicity, culture, sexual orientation, disability or health condition unless they are relevant and valid. These guidelines are meant as a point of reference to help identify appropriate language but are by no means exhaustive or definitive.

Author contributions

For transparency, we encourage authors to submit an author statement file outlining their individual contributions to the paper using the relevant CRediT roles: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Roles/Writing - original draft; Writing - review & editing. Authorship statements should be formatted with the names of authors first and CRediT role(s) following. More details and an example

Changes to authorship

Authors are expected to consider carefully the list and order of authors **before** submitting their manuscript and provide the definitive list of authors at the time of the original submission. Any addition, deletion or rearrangement of author names in the authorship list should be made only **before** the manuscript has been accepted and only if approved by the journal Editor. To request such a change, the Editor must receive the following from the **corresponding author:** (a) the reason for the change in author list and (b) written confirmation (e-mail, letter) from all authors that they agree with the addition, removal or rearrangement. In the case of addition or removal of authors, this includes confirmation from the author being added or removed.
Only in exceptional circumstances will the Editor consider the addition, deletion or rearrangement of authors **after** the manuscript has been accepted. While the Editor considers the request, publication of the manuscript will be suspended. If the manuscript has already been published in an online issue, any requests approved by the Editor will result in a corrigendum.

Article transfer service

This journal is part of our Article Transfer Service. This means that if the Editor feels your article is more suitable in one of our other participating journals, then you may be asked to consider transferring the article to one of those. If you agree, your article will be transferred automatically on your behalf with no need to reformat. Please note that your article will be reviewed again by the new journal. More information.

Copyright

Upon acceptance of an article, authors will be asked to complete a 'Journal Publishing Agreement' (see more information on this). An e-mail will be sent to the corresponding author confirming receipt of the manuscript together with a 'Journal Publishing Agreement' form or a link to the online version of this agreement.

Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution and for all other derivative works, including compilations and translations. If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases.

For gold open access articles: Upon acceptance of an article, authors will be asked to complete an 'Exclusive License Agreement' (more information). Permitted third party reuse of gold open access articles is determined by the author's choice of user license.

Author rights

As an author you (or your employer or institution) have certain rights to reuse your work. More information.

Elsevier supports responsible sharing

Find out how you can share your research published in Elsevier journals.

Role of the funding source

You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement then this should be stated.

Open access

Please visit our Open Access page for more information.

Elsevier Researcher Academy

Researcher Academy is a free e-learning platform designed to support early and mid-career researchers throughout their research journey. The "Learn" environment at Researcher Academy offers several interactive modules, webinars, downloadable guides and resources to guide you through the process of writing for research and going through peer review. Feel free to use these free resources to improve your submission and navigate the publication process with ease.

Language (usage and editing services)

Please write your text in good English (American or British usage is accepted, but not a mixture of these). Authors who feel their English language manuscript may require editing to eliminate possible grammatical or spelling errors and to conform to correct scientific English may wish to use the English Language Editing service available from Elsevier's Author Services.

Submission

Our online submission system guides you stepwise through the process of entering your article details and uploading your files. The system converts your article files to a single PDF file used in the peer-review process. Editable files (e.g., Word, LaTeX) are required to typeset your article for final publication. All correspondence, including notification of the Editor's decision and requests for revision, is sent by e-mail.

Submit your article

Please submit your article via https://ees.elsevier.com/plantgene

Referees

Please submit the names and institutional e-mail addresses of several potential referees. For more details, visit our Support site. Note that the editor retains the sole right to decide whether or not the suggested reviewers are used.

PREPARATION

Peer review

This journal operates a single blind review process. All contributions will be initially assessed by the editor for suitability for the journal. Papers deemed suitable are then typically sent to a minimum of two independent expert reviewers to assess the scientific quality of the paper. The Editor is responsible for the final decision regarding acceptance or rejection of articles. The Editor's decision is final. More information on types of peer review.

Article structure

Subdivision - numbered sections

Divide your article into clearly defined and numbered sections. Subsections should be numbered 1.1 (then 1.1.1, 1.1.2, ...), 1.2, etc. (the abstract is not included in section numbering). Use this numbering also for internal cross-referencing: do not just refer to 'the text'. Any subsection may be given a brief heading. Each heading should appear on its own separate line.

Introduction

State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Material and methods

Provide sufficient details to allow the work to be reproduced by an independent researcher. Methods that are already published should be summarized, and indicated by a reference. If quoting directly from a previously published method, use quotation marks and also cite the source. Any modifications to existing methods should also be described.

Theory/calculation

A Theory section should extend, not repeat, the background to the article already dealt with in the Introduction and lay the foundation for further work. In contrast, a Calculation section represents a practical development from a theoretical basis.

Results

Results should be clear and concise.

Discussion

This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Conclusions

The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.

Appendices

If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.

Essential title page information

- **Title.** Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.
- **Author names and affiliations.** Please clearly indicate the given name(s) and family name(s) of each author and check that all names are accurately spelled. You can add your name between parentheses in your own script behind the English transliteration. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower- case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.

- **Corresponding author.** Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. This responsibility includes answering any future queries about Methodology and Materials. **Ensure that the e-mail address is given and that contact details are kept up to date by the corresponding author**.
- **Present/permanent address.** If an author has moved since the work described in the article was done, or was visiting at the time, a 'Present address' (or 'Permanent address') may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Highlights

Highlights are optional yet highly encouraged for this journal, as they increase the discoverability of your article via search engines. They consist of a short collection of bullet points that capture the novel results of your research as well as new methods that were used during the study (if any). Please have a look at the examples here: example Highlights.
Highlights should be submitted in a separate editable file in the online submission system. Please use 'Highlights' in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point).

Abstract

A concise and factual abstract is required. The abstract should state briefly the purpose of the research, the principal results and major conclusions. An abstract is often presented separately from the article, so it must be able to stand alone. For this reason, References should be avoided, but if essential, then cite the author(s) and year(s). Also, non-standard or uncommon abbreviations should be avoided, but if essential they must be defined at their first mention in the abstract itself.

Keywords

Immediately after the abstract, provide a maximum of 6 keywords, using American spelling and avoiding general and plural terms and multiple concepts (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible. These keywords will be used for indexing purposes.

Abbreviations

Define abbreviations that are not standard in this field in a footnote to be placed on the first page of the article. Such abbreviations that are unavoidable in the abstract must be defined at their first mention there, as well as in the footnote. Ensure consistency of abbreviations throughout the article.

Acknowledgements

Collate acknowledgements in a separate section at the end of the article before the references and do not, therefore, include them on the title page, as a footnote to the title or otherwise. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

Formatting of funding sources

List funding sources in this standard way to facilitate compliance to funder's requirements:
Funding: This work was supported by the National Institutes of Health [grant numbers xxxx, yyyy]; the Bill & Melinda Gates Foundation, Seattle, WA [grant number zzzz]; and the United States Institutes of Peace [grant number aaaa].
It is not necessary to include detailed descriptions on the program or type of grants and awards. When funding is from a block grant or other resources available to a university, college, or other research institution, submit the name of the institute or organization that provided the funding.
If no funding has been provided for the research, please include the following sentence:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Units

Follow internationally accepted rules and conventions: use the international system of units (SI). If other units are mentioned, please give their equivalent in SI.

Footnotes

Footnotes should be used sparingly. Number them consecutively throughout the article. Many word processors can build footnotes into the text, and this feature may be used. Otherwise, please indicate the position of footnotes in the text and list the footnotes themselves separately at the end of the article. Do not include footnotes in the Reference list.

Artwork

Image manipulation

Whilst it is accepted that authors sometimes need to manipulate images for clarity, manipulation for purposes of deception or fraud will be seen as scientific ethical abuse and will be dealt with accordingly.
For graphical images, this journal is applying the following policy: no specific feature within an image may be enhanced, obscured, moved, removed, or introduced. Adjustments of brightness, contrast, or color balance are acceptable if and as long as they do not obscure or eliminate any information present in the original. Nonlinear adjustments (e.g. changes to gamma settings) must be disclosed in the figure legend.

Electronic artwork

General points

- Make sure you use uniform lettering and sizing of your original artwork.
- Embed the used fonts if the application provides that option.
- Aim to use the following fonts in your illustrations: Arial, Courier, Times New Roman, Symbol, or use fonts that look similar.
- Number the illustrations according to their sequence in the text.
- Use a logical naming convention for your artwork files.
- Provide captions to illustrations separately.
- Size the illustrations close to the desired dimensions of the published version.
- Submit each illustration as a separate file.
- Ensure that color images are accessible to all, including those with impaired color vision.

A detailed guide on electronic artwork is available.

You are urged to visit this site; some excerpts from the detailed information are given here.

Formats If your electronic artwork is created in a Microsoft Office application (Word, PowerPoint, Excel) then please supply 'as is' in the native document format.

Regardless of the application used other than Microsoft Office, when your electronic artwork is finalized, please 'Save as' or convert the images to one of the following formats (note the resolution requirements for line drawings, halftones, and line/halftone combinations given below):

- EPS (or PDF): Vector drawings, embed all used fonts.
- TIFF (or JPEG): Color or grayscale photographs (halftones), keep to a minimum of 300 dpi.
- TIFF (or JPEG): Bitmapped (pure black & white pixels) line drawings, keep to a minimum of 1000 dpi.
- TIFF (or JPEG): Combinations bitmapped line/half-tone (color or grayscale), keep to a minimum of 500 dpi.

Please do not:

- Supply files that are optimized for screen use (e.g., GIF, BMP, PICT, WPG); these typically have a low number of pixels and limited set of colors;
- Supply files that are too low in resolution;
- Submit graphics that are disproportionately large for the content.

Color artwork

Please make sure that artwork files are in an acceptable format (TIFF (or JPEG), EPS (or PDF) or MS Office files) and with the correct resolution. If, together with your accepted article, you submit usable color figures then Elsevier will ensure, at no additional charge, that these figures will appear in color online (e.g., ScienceDirect and other sites). Further information on the preparation of electronic artwork.

Illustration services

Elsevier's Author Services offers Illustration Services to authors preparing to submit a manuscript but concerned about the quality of the images accompanying their article. Elsevier's expert illustrators can produce scientific, technical and medical-style images, as well as a full range of charts, tables and graphs. Image 'polishing' is also available, where our illustrators take your image(s) and improve them to a professional standard. Please visit the website to find out more.

Figure captions

Ensure that each illustration has a caption. Supply captions separately, not attached to the figure. A caption should comprise a brief title (**not** on the figure itself) and a description of the illustration. Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.

Tables

Please submit tables as editable text and not as images. Tables can be placed either next to the relevant text in the article, or on separate page(s) at the end. Number tables consecutively in accordance with their appearance in the text and place any table notes below the table body. Be sparing in the use of tables and ensure that the data presented in them do not duplicate results described elsewhere in the article. Please avoid using vertical rules and shading in table cells.

References

Citation in text

Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list they should follow the standard reference style of the journal and should include a substitution of the publication date with either 'Unpublished results' or 'Personal communication'. Citation of a reference as 'in press' implies that the item has been accepted for publication.

Reference links

Increased discoverability of research and high quality peer review are ensured by online links to the sources cited. In order to allow us to create links to abstracting and indexing services, such as Scopus, CrossRef and PubMed, please ensure that data provided in the references are correct. Please note that incorrect surnames, journal/book titles, publication year and pagination may prevent link creation. When copying references, please be careful as they may already contain errors. Use of the DOI is highly encouraged.

A DOI is guaranteed never to change, so you can use it as a permanent link to any electronic article. An example of a citation using DOI for an article not yet in an issue is: VanDecar J.C., Russo R.M., James D.E., Ambeh W.B., Franke M. (2003). Aseismic continuation of the Lesser Antilles slab beneath northeastern Venezuela. Journal of Geophysical Research, https://doi.org/10.1029/2001JB000884. Please note the format of such citations should be in the same style as all other references in the paper.

Web references

As a minimum, the full URL should be given and the date when the reference was last accessed. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

Data references

This journal encourages you to cite underlying or relevant datasets in your manuscript by citing them in your text and including a data reference in your Reference List. Data references should include the following elements: author name(s), dataset title, data repository, version (where available), year, and global persistent identifier. Add [dataset] immediately before the reference so we can properly identify it as a data reference. The [dataset] identifier will not appear in your published article.

References in a special issue

Please ensure that the words 'this issue' are added to any references in the list (and any citations in the text) to other articles in the same Special Issue.

Reference management software

Most Elsevier journals have their reference template available in many of the most popular reference management software products. These include all products that support Citation Style Language styles, such as Mendeley. Using citation plug-ins from these products, authors only need to select the appropriate journal template when preparing their article, after which citations and bibliographies will be automatically formatted in the journal's style. If no template is yet available for this journal, please follow the format of the sample references and citations as shown in this Guide. If you use reference management software, please ensure that you remove all field codes before submitting the electronic manuscript. More information on how to remove field codes from different reference management software.

Users of Mendeley Desktop can easily install the reference style for this journal by clicking the following link: http://open.mendeley.com/use-citation-style/plant-gene

When preparing your manuscript, you will then be able to select this style using the Mendeley plug- ins for Microsoft Word or LibreOffice.

Reference style

Text: All citations in the text should refer to:

1. Single author: the author's name (without initials, unless there is ambiguity) and the year of publication;
2. Two authors: both authors' names and the year of publication;
3. Three or more authors: first author's name followed by 'et al.' and the year of publication.

Citations may be made directly (or parenthetically). Groups of references can be listed either first alphabetically, then chronologically, or vice versa.

Examples: 'as demonstrated (Allan, 2000a, 2000b, 1999; Allan and Jones, 1999).... Or, as demonstrated (Jones, 1999; Allan, 2000)... Kramer et al. (2010) have recently shown ...'

List: References should be arranged first alphabetically and then further sorted chronologically if necessary. More than one reference from the same author(s) in the same year must be identified by the letters 'a', 'b', 'c', etc., placed after the year of publication.

Examples:

- Reference to a journal publication:

Van der Geer, J., Hanraads, J.A.J., Lupton, R.A., 2010. The art of writing a scientific article. J. Sci. Commun. 163, 51–59. https://doi.org/10.1016/j.Sc.2010.00372.

- Reference to a journal publication with an article number:

Van der Geer, J., Hanraads, J.A.J., Lupton, R.A., 2018. The art of writing a scientific article. Heliyon. 19, e00205. https://doi.org/10.1016/j.heliyon.2018.e00205.

- Reference to a book:

Strunk Jr., W., White, E.B., 2000. The Elements of Style, fourth ed. Longman, New York.

- Reference to a chapter in an edited book:

Mettam, G.R., Adams, L.B., 2009. How to prepare an electronic version of your article, in: Jones, B.S., Smith , R.Z. (Eds.), Introduction to the Electronic Age. E-Publishing Inc., New York, pp. 281–304.

- Reference to a website:

Cancer Research UK, 1975. Cancer statistics reports for the UK. http://www.cancerresearchuk.org/aboutcancer/statistics/cancerstatsreport/ (accessed 13 March 2003).

- Reference to a dataset:

[dataset] Oguro, M., Imahiro, S., Saito, S., Nakashizuka, T., 2015. Mortality data for Japanese oak wilt disease and surrounding forest compositions. Mendeley Data, v1. https://doi.org/10.17632/ xwj98nb39r.1.

Journal abbreviations source

Journal names should be abbreviated according to the List of Title Word Abbreviations

Video

Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include links to these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. In order to ensure that your video or animation material is directly usable, please provide the file in one of our recommended file formats with a preferred maximum size of 150 MB per file, 1 GB in total. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages. Note: since video and animation cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

Data visualization

Include interactive data visualizations in your publication and let your readers interact and engage more closely with your research. Follow the instructions here to find out about available data visualization options and how to include them with your article.

Supplementary material

Supplementary material such as applications, images and sound clips, can be published with your article to enhance it. Submitted supplementary items are published exactly as they are received (Excel or PowerPoint files will appear as such online). Please submit your material together with the article and supply a concise, descriptive caption for each supplementary file. If you wish to make changes to supplementary material during any stage of the process, please make sure to provide an updated file.

Do not annotate any corrections on a previous version. Please switch off the 'Track Changes' option in Microsoft Office files as these will appear in the published version.

Research data

This journal encourages and enables you to share data that supports your research publication where appropriate, and enables you to interlink the data with your published articles. Research data refers to the results of observations or experimentation that validate research findings. To facilitate reproducibility and data reuse, this journal also encourages you to share your software, code, models, algorithms, protocols, methods and other useful materials related to the project.
Below are a number of ways in which you can associate data with your article or make a statement about the availability of your data when submitting your manuscript. If you are sharing data in one of these ways, you are encouraged to cite the data in your manuscript and reference list. Please refer to the "References" section for more information about data citation. For more information on depositing, sharing and using research data and other relevant research materials, visit the research data page.

Data linking

If you have made your research data available in a data repository, you can link your article directly to the dataset. Elsevier collaborates with a number of repositories to link articles on ScienceDirect with relevant repositories, giving readers access to underlying data that gives them a better understanding of the research described.
There are different ways to link your datasets to your article. When available, you can directly link your dataset to your article by providing the relevant information in the submission system. For more information, visit the database linking page.
For supported data repositories a repository banner will automatically appear next to your published article on ScienceDirect.
In addition, you can link to relevant data or entities through identifiers within the text of your manuscript, using the following format: Database: xxxx (e.g., TAIR: AT1G01020; CCDC: 734053; PDB: 1XFN).

Mendeley Data

This journal supports Mendeley Data, enabling you to deposit any research data (including raw and processed data, video, code, software, algorithms, protocols, and methods) associated with your manuscript in a free-to-use, open access repository. During the submission process, after uploading your manuscript, you will have the opportunity to upload your relevant datasets

directly to Mendeley Data. The datasets will be listed and directly accessible to readers next to your published article online.
For more information, visit the Mendeley Data for journals page.

Data in Brief

You have the option of converting any or all parts of your supplementary or additional raw data into one or multiple data articles, a new kind of article that houses and describes your data. Data articles ensure that your data is actively reviewed, curated, formatted, indexed, given a DOI and publicly available to all upon publication. You are encouraged to submit your article for Data in Brief as an additional item directly alongside the revised version of your manuscript. If your research article is accepted, your data article will automatically be transferred over to Data in Brief where it will be editorially reviewed and published in the open access data journal, Data in Brief. Please note an open access fee of 600 USD is payable for publication in Data in Brief. Full details can be found on the Data in Brief website. Please use this template to write your Data in Brief.

Data statement

To foster transparency, we encourage you to state the availability of your data in your submission. This may be a requirement of your funding body or institution. If your data is unavailable to access or unsuitable to post, you will have the opportunity to indicate why during the submission process, for example by stating that the research data is confidential. The statement will appear with your published article on ScienceDirect

AFTER ACCEPTANCE

Online proof correction

To ensure a fast publication process of the article, we kindly ask authors to provide us with their proof corrections within two days. Corresponding authors will receive an e-mail with a link to our online proofing system, allowing annotation and correction of proofs online. The environment is similar to MS Word: in addition to editing text, you can also comment on figures/tables and answer questions from the Copy Editor. Web-based proofing provides a faster and less error-prone process by allowing you to directly type your corrections, eliminating the potential introduction of errors.

If preferred, you can still choose to annotate and upload your edits on the PDF version. All instructions for proofing will be given in the e-mail we send to authors, including alternative methods to the online version and PDF.

We will do everything possible to get your article published quickly and accurately. Please use this proof only for checking the typesetting, editing, completeness and correctness of the text, tables and figures. Significant changes to the article as accepted for publication will only be considered at this stage with permission from the Editor. It is important to ensure that all corrections are sent back to us in one communication. Please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed. Proofreading is solely your responsibility.

We will do everything possible to get your article published quickly and accurately. Please use this proof only for checking the typesetting, editing, completeness and correctness of the text, tables and figures. Significant changes to the article as accepted for publication will only be considered at this stage with permission from the Editor. It is important to ensure that all corrections are sent back to us in one communication. Please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed. Proofreading is solely your responsibility.

Offprints

The corresponding author will, at no cost, receive a customized Share Link providing 50 days free access to the final published version of the article on ScienceDirect. The Share Link can be used for sharing the article via any communication channel, including email and social media. For an extra charge, paper offprints can be ordered via the offprint order form which is sent once the article is accepted for publication. Both corresponding and co-authors may order offprints at any time via Elsevier's Author Services. Corresponding authors who have published their article gold open access do not receive a Share Link as their final published version of the article is available open access on ScienceDirect and can be shared through the article DOI link.

AUTHOR INQUIRIES

Visit the Elsevier Support Center to find the answers you need. Here you will find everything from Frequently Asked Questions to ways to get in touch.
You can also check the status of your submitted article or find out when your accepted article will be published.