12-2018

Biochemical and transcriptomic characterization of glycoside hydrolases in *Thermobia domestica* and *Ctenolepisma longicaudata*

Ratnasri Mallipeddi
University of Tennessee, rmallipe@vols.utk.edu

Recommended Citation

Mallipeddi, Ratnasri, "Biochemical and transcriptomic characterization of glycoside hydrolases in *Thermobia domestica* and *Ctenolepisma longicaudata*. " PhD diss., University of Tennessee, 2018.
https://trace.tennessee.edu/utk_graddiss/5290

To the Graduate Council:

I am submitting herewith a dissertation written by Ratnasri Mallipeddi entitled "Biochemical and transcriptomic characterization of glycoside hydrolases in *Thermobia domestica* and *Ctenolepisma longicaudata*." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Entomology, Plant Pathology and Nematology.

Juan Luis Jurat-Fuentes, Major Professor

We have read this dissertation and recommend its acceptance:

Parwinder S. Grewal, William E. Klingeman, Margaret E. Staton

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Biochemical and transcriptomic characterization of glycoside hydrolases in *Thermobia domestica* and *Ctenolepisma longicaudata*

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Ratnasri Mallipeddi
December 2018

Dedication

This dissertation is dedicated to all that have supported me:

My husband and Daughter: Satyendra and Riya Pothula

My in-laws: Sesharatnam and Satyanarayana Murthy Pothula

My Mom: Ammaji Mallipeddi

My Brothers-in-law, his wife and their son: Subhash, Viharika and Yuvan Pothula

My Brothers and their Wives: Hareen, Hemanth Mallipeddi and and their wives

My Professors and Friends:

Dr. Juan Luis Jurat-Fuentes, Dr. Parwinder S. Grewal, Dr. William E. Klingeman, Dr. Brian R. Johnson, Dr. Margaret E. Staton, Dr. Gary Phillips, Dr. Heba Abdelgaffar, Robert J. Pivar, Sunny, Claire, James and LaVerne Phillips, Jagadish Cherukuri, Samatha, Srikanth and Shanthi Earpina, Bakkareddy, Sarala, Manaswini, Gyandeep and Manasa Kankanala.

Acknowledgements

I am greatly indebted to Dr. Brian Johnson for his guidance and support on bioinformatic analyses. I would like to thank Dr. Ernest C. Bernard for helping with the microscopy images and Dr. Liesel Schneider for advising on statistical analyses. I appreciate Dr. Heba Abdelgaffar for helping with histology technique and Matthew Huff for helping with bioinformatic analyses. This project was funded by grant number 1456662 from the Division of Integrative Organismal Systems of the National Science Foundation (NSF).

Abstract

The digestive system of phytophagous insects is considered a relevant prospecting resource for identification of novel cellulolytic enzymes that may improve industrial processes of cellulose degradation. While much is known from insect models of cellulose digestion, such as termites and roaches (Blattodea), there is a dearth of information on insects belonging to basal hexapod groups. As part of a screening effort to identify insects with highly active cellulolytic systems, we detected species of Zygentoma as displaying the highest relative cellulolytic activity levels compared to all other tested insect orders, including Blattodea. The goals of the present study were to provide a morphohistological and biochemical characterization of the digestive system of Zygentoma, to annotate and screen for plant cell wall degrading enzyme (PCWDE) genes present in *Thermobia domestica* and *Ctenolepisma longicaudata* genomes, and to study their differential expression depending on diets with varying degree of cellulosic content. Morphohistological characterization supported no relevant differences in the digestive tube of *T. domestica* and *C. longicaudata*. Quantitative and qualitative cellulase assays identified the foregut as the region with the highest cellulolytic activity in both the tested insects, yet *T. domestica* was found to have higher endoglucanase, xylanase and pectinase activities compared to *C. longicaudata*. Annotation of predicted coding sequences from genomes of *T. domestica* and *C. longicaudata* reported numerous genes encoding for endoglucanases, glucosidases, β-1,3-glucanases, maltases, amylases, mannosidases, glucuronidases and lytic polysaccharide monoxygenases (LPMOs). Differential gene expression analysis of both species revealed that cellulase gene expression is primarily driven by type of tissue rather than diet. However, within each tissue of *T. domestica* and *C. longicaudata*, a higher number of PCWDEs and LPMOs were significantly up-regulated in paper, the most highly cellulosic diet, compared to all other tested diets. The annotation of coding sequences and differentially expressed PCWDE genes revealed highest identity to insect homologs, which suggests the potential conservation of PCWDEs through evolution and an ancient origin for cellulases in insects. These findings advance our understanding of cellulose digestion in a basal hexapod group and the identification of novel cellulolytic enzymes with potential application in the biofuel industry.

Table of contents

List of tables

List of figures

Chapter 1

General introduction

Introduction

The feeding preferences of an organism are reflected in the digestive enzymes produced in its gut. For instance, organisms feeding on fungal cell walls and arthropods express chitinases to degrade chitin; organisms feeding on fungal cell contents express trehalases to digest sugars; and omnivorous feeders express a complex of enzymes to digest different substrates. Similarly, organisms feeding on plant materials, algae, paper and other cellulose-rich materials produce a complex of plant cell wall degrading enzymes (PCWDEs) (Calderón-Cortés et al., 2012; Watanabe and Tokuda, 2010).

Photosynthetic organisms such as plants, algae and few species of bacteria are producing more than 10^{11} tons of organic matter annually by fixing CO_2. More than half of this organic matter is made up of a linear polymer of D-glucopyranosyl units linked by β-1,4 linkages, cellulose (Leschine, 1995). Additionally, cellulose is widely present in municipal, paper and wood waste (Li et al., 2009). Therefore, cellulose is the most available and renewable energy source on earth.

Cellulose is degraded to glucose monomers by the combined action of three types of enzymes, based on their mode of action and substrate specificities. Endoglucanases (EC 3.2.1.4) cut at random internal points in cellulose chains, while exoglucanases (EC 3.2.1.91) cleave at the non-reducing ends releasing cellobiose units that are digested to glucose by β-glucosidases (EC 3.2.1.21) (Watanabe and Tokuda, 2010). Apart from cellulases, PCWDEs include hemicellulases and pectinases. Hemicellulases and pectinases digest hemicellulose and pectin polysaccharides, respectively, which are interlocked with cellulose fibers in the plant cell wall (Gilbert, 2010). The potential of cellulases in degrading cellulose as the most available and renewable energy resource on earth motivated the exploration to identify highly efficient cellulases in a variety of organisms.

Initially, it was believed that only microorganisms could produce cellulases, therefore cellulolytic research was confined to bacteria, fungi and other microorganisms (Tomme et al., 1995). Later, cellulolytic activity was found for the first time in animals in the land snail, *Helix pomatia* (Biedermann and Moritz, 1898). Since then, cellulolytic activity has been found in many wood feeding insects, especially termites and cockroaches. In the early years of cellulolytic research, cellulose digestion in termites was believed to merely occur by virtue of protozoans that lived in their hindgut (Cleveland, 1924). Later, higher termites lacking protozoans were also found to digest cellulose, yet this ability was misattributed to the resident bacteria in their hindgut (Martin, 1991). Substantial evidence has been reported supporting the acquisition of cellulases by termites and other wood feeding insects from symbiotic fungi (Abo-Khatwa, 1978; Martin and Martin, 1978; Rouland et al., 1991, 1988). Based on this growing body of evidence it was generalized that the acquisition of cellulase genes was probably common in insects (Martin, 1991). However, in 1998 termites were proved to produce cellulases endogenously (Watanabe et al., 1998). Subsequently, endogenous insect cellulases were reported from different insect orders including Blattodea, Coleoptera, Orthoptera, Phthiraptera, Hemiptera, Phasmida, Lepidoptera, Diptera, and Hymenoptera (Calderón-Cortés et al., 2012; Chatterjee et al., 2015; Fischer et al., 2013; Watanabe and Tokuda, 2010).

Although research to identify cellulases in hexapods started in the 1900s, much of the characterization of these cellulolytic systems was confined to a few groups such as Isoptera, Blattodea, Coleoptera, Lepidoptera and Diptera, probably due to the availability of sequenced genomes and other metagenomic resources (Davison and Blaxter, 2005; Watanabe and Tokuda, 2010). Insects belonging to 16 taxonomic orders were reported to have endogenous production of one or more PCWDE genes (Calderón-Cortés et al., 2012). However, molecular evidence confirming the presence of PCWDE genes including endoglucanases belonging to glycoside hydrolase family (GH) 9 and 45 , β-glucosidases of GH 5, hemicellulases such as xyloglucanases of GH 5 and GH 11, β-1,3-glucanases of GH 16 and pectinases of GH 28, has only been obtained from insects belonging to 8 taxonomic orders (Calderón-Cortés et al., 2012). In contrast, cellulolytic systems in other insect orders that contain species specialized to feed on plant material and highly cellulosic materials are understudied, especially primitive hexapods (Terry and Whiting, 2005). In this regard, we identified species of Zygentoma as displaying the highest relative cellulolytic activity levels compared to all other tested insect orders (Oppert et al., 2010 and Chapter 2). Many species of Zygentoma, including *Ctenolepisma longicaudata*, *Lepisma saccharina*, *Thermobia domestica* and *Acrotelsa collaris* (Laibach, 1952; Lindsay, 1940; Modder, 1975; Sahrhage, 1953) were reported to feed on paper and other cellulose-rich substrates. Degraded cellulose fibers and avicel were microscopically observed in the guts of *C. longicaudata* and *T. domestica*, respectively (Lindsay, 1940; Sabbadin et al., 2018). Gut fluids of Zygentoma were characterized as including endoglucanase, β-glucosidase, amylase, maltase, sucrase and lactase activities (Lasker and Giese, 1956; Zinkler and Götze, 1987). Although five fungal and four bacterial species were found in the gut of *T. domestica*, only the fungus *Mycotypha microspore* displayed cellulolytic activity (Woodbury and Gries, 2013a). However, axenic *C. lineata* and defaunated *T. domestica* supported endogenous cellulase production, suggesting the existence of endogenous cellulases (Lasker and Giese, 1956; Treves and Martin, 1994). In fact, *T. domestica* can digest cellulose in a symbiont-independent manner, but needs the aid of microbes in accessing cellulosic substrates (Woodbury and Gries, 2013b). More recently, Sabbadin et al (2018) investigated the digestive proteome of *T. domestica* and identified carbohydrate degrading enzymes including lytic polysaccharide monooxygenases (LPMOs), which weaken cellulose fibers making them more accessible to cellulose degradation. Even though members of Zygentoma were found to have diverse cellulolytic activities for efficient cellulose digestion (Lasker and Giese, 1956; Zinkler and Götze, 1987), molecular evidence confirming endogenous production of these cellulases is still lacking.

The goals of the present study were to provide a morphohistological characterization of digestive system, biochemical characterization of highly active cellulase enzymes in digestive fluids, to annotate and screen for PCWDEs genes present in *T. domestica* and *C. longicaudata* genomes and to study their differential expression in foregut and rest of the body samples when *T. domestica* and *C. longicaudata* were fed on four diets with varying degree of cellulosic content.

List of references

Abo-Khatwa, N., 1978. Cellulase of fungus-growing termites: A new hypothesis on its origin. Experientia 34, 559–560. https://doi.org/10.1007/BF01936956

Biedermann, W., Moritz, P., 1898. Beitr¨age zur vergeleichenden Physiologie der Verdauung. II. Uver ein cellulosel ¨ odendes Enzyme im Leversecret der Schnecke (Helix pomatia). Pfludg Arch Gem 73, 219–287.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H., Oyama, K., 2012. Endogenous Plant Cell Wall Digestion: A Key Mechanism in Insect Evolution. Annu. Rev. Ecol. Evol. Syst. 43, 45–71. https://doi.org/10.1146/annurev-ecolsys-110411-160312

Chatterjee, S., Sharma, S., Prasad, R.K., Datta, S., Dubey, D., Meghvansi, M.K., Vairale, M.G., Veer, V., 2015. Cellulase enzyme based biodegradation of cellulosic materials: an overview. South Asian J. Exp. Biol. 5, 271–282.

Cleveland, L.R., 1924. The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to reticulitermes flavipes kollar. Biol. Bull. 46, 178–201. https://doi.org/10.2307/1536507

Davison, A., Blaxter, M., 2005. Ancient Origin of Glycosyl Hydrolase Family 9 Cellulase Genes. Mol. Biol. Evol. 22, 1273–1284. https://doi.org/10.1093/molbev/msi107

Fischer, R., Ostafe, R., Twyman, R.M., 2013. Cellulases from Insects, in: Vilcinskas, A. (Ed.), Yellow Biotechnology II: Insect Biotechnology in Plant Protection and Industry, Advances in Biochemical Engineering/Biotechnology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 51–64. https://doi.org/10.1007/10_2013_206

Gilbert, H.J., 2010. The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. Plant Physiol. 153, 444–455. https://doi.org/10.1104/pp.110.156646

Kim, N., Choo, Y.M., Lee, K.S., Hong, S.J., Seol, K.Y., Je, Y.H., Sohn, H.D., Jin, B.R., 2008. Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket Teleogryllus emma. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 150, 368–376. https://doi.org/10.1016/j.cbpb.2008.04.005

Laibach, H., 1952. Lepisma saccharina L., das Silberflschchen. Z F Hyg Zool U Schddlingsbekdmpfung 40, 321–370.

Lasker, R., Giese, A.C., 1956. Cellulose Digestion By The Silverfish Ctenolepisma Lineata. J. Exp. Biol. 33, 542–553.

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Dae Sohn, H., Jin, B.R., 2004. cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, Apriona germari. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 139, 107–116. https://doi.org/10.1016/j.cbpc.2004.06.015

Lindsay, E., 1940. The Biology of the Silverfish, Ctenolepisma longicaudata Esch. with particular Reference to its feeding Habits. Proc. R. Soc. Vic. 52.

Martin, M.M., 1991. The evolution of cellulose digestion in insects. Phil Trans R Soc Lond B 333, 281–288. https://doi.org/10.1098/rstb.1991.0078

Martin, M.M., Martin, J.S., 1978. Cellulose Digestion in the Midgut of the Fungus-Growing Termite Macrotermes natalensis: The Role of Acquired Digestive Enzymes. Science 199, 1453–1455. https://doi.org/10.1126/science.199.4336.1453

Modder, W.W.D., 1975. Feeding and growth of Acrotelsa collaris (Fabricius) (Thysanura, Lepismatidae) on different types of paper. J. Stored Prod. Res. 11, 71–74. https://doi.org/10.1016/0022-474X(75)90042-9

Oppert, C., Klingeman, W.E., Willis, J.D., Oppert, B., Jurat-Fuentes, J.L., 2010. Prospecting for cellulolytic activity in insect digestive fluids. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 155, 145–154. https://doi.org/10.1016/j.cbpb.2009.10.014

Rouland, C., Civas, A., Renoux, J., Petek, F., 1988. Purification and properties of cellulases from the termite Macrotermes mülleri (termitidae, macrotermitinae) and its symbiotic fungus termitomyces sp. Comp. Biochem. Physiol. Part B Comp. Biochem. 91, 449–458. https://doi.org/10.1016/0305-0491(88)90004-1

Rouland, C., Lenoir, F., Lepage, M., 1991. The role of the symbiotic fungus in the digestive metabolism of several species of fungus-growing termites. Comp. Biochem. Physiol. A Physiol. 99, 657–663. https://doi.org/10.1016/0300-9629(91)90146-4

Sabbadin, F., Hemsworth, G.R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T., Marques, R.D.S., Sweeney, S.T., Besser, K., Elias, L., Pesante, G., Li, Y., Dowle, A.A., Bates, R., Gomez, L.D., Simister, R., Davies, G.J., Walton, P.H., Bruce, N.C., McQueen-Mason, S.J., 2018. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. Nat. Commun. 9, 756. https://doi.org/10.1038/s41467-018-03142-x

Sahrhage, D., 1953. Okologische Untersuchengen an Thermobia domestica (Packard) und Lepisma saccharina L. Z Wiss Zool 157, 77–168.

Terry, M.D., Whiting, M.F., 2005. Mantophasmatodea and phylogeny of the lower neopterous insects. Cladistics 21, 240–257. https://doi.org/10.1111/j.1096-0031.2005.00062.x

Tomme, P., Warren, R.A.J., Gilkes, N.R., 1995. Cellulose Hydrolysis by Bacteria and Fungi, in: Poole, R.K. (Ed.), Advances in Microbial Physiology. Academic Press, pp. 1–81. https://doi.org/10.1016/S0065-2911(08)60143-5

Treves, D.S., Martin, M.M., 1994. Cellulose digestion in primitive hexapods: Effect of ingested antibiotics on gut microbial populations and gut cellulase levels in the firebrat,Thermobia domestica (Zygentoma, Lepismatidae). J. Chem. Ecol. 20, 2003–2020. https://doi.org/10.1007/BF02066239

Watanabe, H., Noda, H., Tokuda, G., Lo, N., 1998. A cellulase gene of termite origin. Nat. Lond. 394, 330–1. http://dx.doi.org/10.1038/28527

Watanabe, H., Tokuda, G., 2010. Cellulolytic Systems in Insects. Annu. Rev. Entomol. 55, 609–632. https://doi.org/10.1146/annurev-ento-112408-085319

Woodbury, N., Gries, G., 2013a. Firebrats, Thermobia domestica, aggregate in response to the microbes Enterobacter cloacae and Mycotypha microspora. Entomol. Exp. Appl. 147, 154–159. https://doi.org/10.1111/eea.12054

Woodbury, N., Gries, G., 2013b. How Firebrats (Thysanura: Lepismatidae) Detect and Nutritionally Benefit From Their Microbial Symbionts Enterobacter cloacae and Mycotypha microspora. Environ. Entomol. 42, 860–867. https://doi.org/10.1603/EN13104

Zinkler, D., Götze, M., 1987. Cellulose digestion by the firebrat Thermobia domestica. Comp. Biochem. Physiol. Part B Comp. Biochem. 88, 661–666. https://doi.org/10.1016/0305-0491(87)90360-9

Chapter 2

The digestive system in Zygentoma as a model for high cellulase activity

Ratnasri Pothula; O. P. Perera, William E. Klingeman; Heba M. Y. Abd-Elgaffar; Brian R. Johnson and Juan Luis Jurat-Fuentes

My contributions included: (1) planning and performing experiments, (2) data collection and analysis, (3) writing the manuscript and making figures. O. P. Perera helped with 2 and 3, Brian R. Johnson assisted with 3, William E. Klingeman provided insects and assisted with 3, Heba M. Y. Abd-Elgaffar helped with histology and Juan Luis Jurat-Fuentes assisted with 1, 2 and 3.

Abstract

The digestive system of phytophagous insects is expected to include novel cellulolytic enzymes that may improve industrial cellulose degradation. While much research has been performed on model insects such as termites and roaches, there is dearth of information on cellulolytic systems in insects belonging to basal hexapod groups. As part of a screening effort to identify insects with highly active cellulolytic systems, we detected species of Zygentoma as displaying the highest relative cellulolytic activity levels compared to all other tested insect orders including Blattodea. The goal of the present study was to provide a morphohistological characterization of cellulose digestion and to identify highly active cellulase enzymes in digestive fluids in two species of Zygentoma, the firebrat (*Thermobia domestica*) and the gray silverfish (*Ctenolepisma longicaudata*). Morphohistological characterization supported no relevant differences in the digestive tube of *T. domestica* and *C. longicaudata*. Quantitative and qualitative cellulase assays identified the foregut as the region with the highest cellulolytic activity in both the tested insects, yet *T. domestica* was found to have higher endoglucanase, xylanase and pectinase

activities compared to *C. longicaudata*. In summary, we identify that the firebrat is a member of zygentoma displaying highest relative cellulolytic activity compared to other insect orders including model insects for cellulolytic research. Additionally, digestive fluids of firebrat displayed higher cellulase, xylanase and pectinase activities, which are necessary for efficient plant cell wall degradation. These findings advance our understanding of cellulose digestion in a basal hexapod group and identify novel cellulolytic enzymes with potential application in industrial cellulose digestion.

Introduction

The digestive system of phytophagous insects is considered a relevant prospecting resource for identification of new cellulolytic enzymes to improve cellulose degradation to glucose, a step accounting for >40% of production costs of ethanol biofuel from plant biomass (Zhu et al., 2009; Bekmuradov et al., 2014). Cellulose is a linear polymer of D-glucopyranosyl units linked by β-1,4 linkages that is degraded to glucose subunits by the combined action of three types of enzymes, based on their mode of action and substrate specificities. Endoglucanases (EC 3.2.1.4) cut at random internal points in cellulose chains, while exoglucanases (EC 3.2.1.91) cleave at the non-reducing ends releasing cellobiose units that are digested to glucose by β-glucosidases (EC 3.2.1.21) (Watanabe and Tokuda, 2010). Although research on cellulolytic systems in insects was initially confined to symbiotic microorganisms (Cleveland, 1924), in the last decade insect endogenous plant cell wall degrading enzymes (PCWDEs) have been described in Isoptera, Blattaria, Coleoptera, Orthoptera, Pthiraptera, Hemiptera, Phasmatodea, Lepidoptera, Diptera and Hymenoptera (Watanabe and Tokuda, 2010, Calderón-Cortés et al., 2012, Fischer et al., 2013, and Chatterjee et al., 2015). Much of the research on insect cellulolytic enzymes has concentrated on Isoptera, Blattodea, Coleoptera, Lepidoptera and Diptera, probably due to the availability of sequenced genomes and other metagenomic resources (Davison and Blaxter, 2005; Watanabe and Tokuda, 2010). In contrast, cellulolytic systems in other insect orders that contain species specialized to feed on plant material are understudied (Terry and Whiting, 2005). As part of a screening effort to identify insects with highly active cellulolytic systems (Oppert et al., 2010), we detected species of Zygentoma as displaying the highest relative cellulolytic activity levels compared to all other tested insect orders.

Members of Zygentoma are known to feed and digest highly cellulosic materials such as paper, cardboard, flour and insulation (Berger 1945 and Sahrhage 1953). Description of the digestive system in *Ctenolepisma campbelli* and *Lepisma saccharinum* supported similarities with Orthoptera, including slightly longer than body length, differentiated into foregut, midgut and hindgut, and presence of a muscular proventriculus with sclerotized teeth like structures (Barnhart, 1961). Production of endogenous cellulases was previously reported in *Ctenolepisma lineata* and the firebrat, *Thermobia domestica* (Lasker and Giese 1956; Zinkler and Götze 1987; Treves and Martin 1994). Additionally, the crop was characterized as displaying the highest cellulolytic activity compared to other digestive regions in *T. domestica* (Zinkler and Götze, 1987). More recently, Sabbadin et al (2018) investigated the digestive proteome of *T. domestica* and identified carbohydrate degrading enzymes including lytic polysaccharide monooxygenases (LPMOs), which weaken cellulose fibers making them more accessible to cellulose degradation.

The goal of the present study was to provide a morphohistological and biochemical characterization of cellulose digestion in Zygentoma. Initial characterization supported no relevant morphological differences in the digestive tube of the firebrat (*T. domestica*) and the gray silverfish (*Ctenolepisma longicaudata*). Quantitative and qualitative cellulase assays identified the foregut as the region with the highest cellulolytic activity and *T. domestica* as displaying higher endoglucanase, xylanase and pectinase activities compared to *C. longicaudata*. These findings advance our understanding of cellulose digestion in a basal hexapod group and the identification of novel cellulolytic enzymes with potential application in industrial cellulose digestion.

Materials and methods

Insects

Adult silverfish (*Ctenolepisma longicaudata*) and firebrat (*Thermobia domestica*) were used for different objectives in this study. Nymphs and adults of *C. longicaudata* were hand-collected using lid-less plastic dishes (15 cm) (Pioneer Plastics Inc., Dixon, KY) baited with whole-grain oat flakes (Quick 1-minute Oats, Quaker Oats Co., Chicago, IL) or swept into collection containers using a 3” long, soft-bristle make-up brush in buildings and reared in the laboratory at room temperature on rolled oats, paper, and dry dog food (Pedigree adult complete nutrition, Mars, Inc., Mount Olive, NJ) as protein source. An established culture of *T. domestica* was generously provided by Patrick Stanley and Eric Snell (Snell Scientific, LLC, Meansville, GA) from a colony derived from a culture at the

Department of Entomology at Ohio State University (Columbus, OH). This colony has been maintained at 34^{0}C in a dark incubator in the Department of Entomology and Plant Pathology at the University of Tennessee for >8 years using printer paper and NatureWise chick starter grower feed (Nutrena, Minneapolis, MN) as carbohydrate and protein source, respectively.

Gut morphology and histology

For gut morphological studies, adult *C. longicaudata* and *T. domestica* were anesthetized for 10 min at 4^{0}C and dissected under a Zeiss Stemi 2000-C stereo microscope (Carl Zeiss Microscopy, LLC, Thornwood, NY). The gut was carefully dissected from the rest of the body and images were taken with a Canon DS126311 camera (Canon, Ota, Tokyo, Japan) mounted on the stereo microscope. Adult *C. longicaudata* and *T. domestica* for histological studies were sacrificed by incubation at -20°C for ten minutes and then fixed in Carnoy's (60% ethanol, 30% chloroform, and 10% glacial acetic acid) for four hours at 4^{0}C. After fixing, whole insects were transferred to 70% ethyl alcohol and sent to the Biomedical and Diagnostic Services, at the University of Tennessee College of Veterinary Medicine (Knoxville, TN) for sectioning and staining with hematoxylin and eosin. Histological sections were examined and documented using an Olympus BX63F upright microscope (Olympus Corporation, Shinjuku, Tokyo, Japan).

Preparation of samples for biochemical tests

Adult *C. longicaudata* and *T. domestica* were starved individually in Petri dishes for five days, and then fed with high cellulosic standard 92 multipurpose printing paper (Georgia-Pacific, Atlanta, GA, USA) or control (bovine serum albumin) diets and allowed to feed for five days under culture conditions. A piece of 1% agar was provided and changed every other day as a water source. On the 11th day, anesthesia and dissections were carried out on ice. The digestive tract along with the head were separated from the rest of the body, and the gut was further divided into the foregut, midgut and hindgut regions using a sterile scalpel. The head was separated from the foregut, and the midgut region was identified from the origin of gastric caeca to the origin of Malpighian tubules, which was considered as the start for the hindgut region. Tissues pooled from six individual insects were placed in microfuge tubes containing 100 µl of sterile water and homogenized with disposable pellet pestles, vortexed briefly and then centrifuged at 21,000 x *g* for 3 min at room temperature. The supernatant was collected into microfuge tubes and stored at -80^{0}C until used.

Zymography

Cellulolytic activity in tissues of *C. longicaudata* and *T. domestica* was tested using zymography on SDS-12%PAGE gels containing 0.2% carboxymethyl cellulose (CMC) as substrate (Willis et al., 2010). The sample fluids were allowed to thaw on ice and protein concentrations were estimated using the Protein Quantification kit in a Qubit fluorometer (Invitrogen, Carlsbad, CA). Samples (100 µg) were mixed with an equal volume of 2X sample buffer (50 mM Tris–HCl, pH 6.8, 2% SDS, 10% glycerol, 1% β-mercaptoethanol, 0.01% bromophenol blue) and the mixture was heated at 72^{0}C for 15 min to partially denature proteins. Commercial cellulase from *Aspergillus niger* (Tokyo Chemical Industry Co., Ltd., Portland, OR) was used as positive control. Samples were resolved by electrophoresis at 100 V until the dye reached the bottom of the gel and the gels were washed in 50 ml of 0.1 M Sodium succinate buffer (pH 5.8) containing 10 mM Dithiothreitol (DTT) for five washes of 30 min with constant shaking. Gels were then incubated in 0.1 M Sodium succinate buffer (pH 5.8) with no DTT for 30 min at 60^0C and then stained with 0.1% of Congo red (Acros Organics, Waltham, MA) for 10 min. Gels were destained by incubating in 50 ml of 1 M NaCl until the cellulolytic activity bands were clearly visible as clear bands on a red background. Glacial acetic acid (100 µl) was added to shift the background gel color to dark-purple for more clear observation of activity bands. Gel images were taken with a Versadoc 1000 Imager (Bio-Rad, Hercules, CA).

Quantification of cellulolytic, xylanase and pectinase activities

Quantitative activity against carboxymethylcellulose (CMC) in gut fluids of diverse insects (shown in Fig. 2.1A) was determined as described in Oppert et al. (2010). Assays with gut fluids from *C. longicaudata* were performed concomitantly but were not originally included in Oppert et al (2010). In the present study, the protein content in dissected head and foregut samples was quantified using the Protein Quantification kit in a Qubit fluorometer (Invitrogen, Carlsbad, CA). Cellulose degrading activity in head and foregut tissues of *C. longicaudata* and *T. domestica* was quantified using a cellulase assay kit (Megazyme, Ireland) to quantify endoglucanase activity against 4-nitrophenyl-β-D-cellopentaoside (BPNPG5) as substrate, 4-nitrophenyl β-D-cellobioside (pNPC) (Sigma-Aldrich, St. Louis, MO) as substrate for β-glucosidase, 4-nitrophenyl β-D-xylopyranoside (pNPX) (Sigma-Aldrich, St. Louis, MO) as substrate to quantify xylanase activity, and pectin from citrus peel (Sigma-Aldrich, St. Louis, MO) to measure pectinase activity.

Briefly, β-glucosidase and xylanase activities were measured in samples (20 µl containing 10 µg of protein) mixed with 130 µl of 10 mM substrate in 50 mM sodium acetate buffer (pH 5.0), and incubated at 50^{0}C for 30 min. To a 50 µl aliquot of reaction mixture, 50 µl of 2 M Na_2CO_3 was added to stop the reaction and absorbance was measured at 405 nm in a Synergy HT microplate reader (BioTek, Winooski, VT) using the Gen5 software (v. 2.0, BioTek, Winooski, VT). A 4-nitrophenol standard curve (0-1 mM) was used to quantify specific activity and background activity was corrected by subtracting final values from initial values. Specific activity was expressed in U/mg of protein, with 1 U defined as the amount of enzyme resulting in production of 1 µmol of 4-nitrophenol per min at pH 5.0 and 50^{0}C.

Endoglucanase activity was measured in samples (5 µg of protein in 25 µl) mixed with 50 µl of substrate and incubated at 40^{0}C for 10 min. Reactions were terminated by adding alkaline solution (125 µl of Tris buffer solution pH 9.0) and absorbance was measured at 405 nm as above. Activity was calculated according to the Mega-Calc method from the manufacturer (https://secure.megazyme.com/Cellulose-Assay-Kit-CELLG5-Method).

Pectinase activity was determined in samples (10 µg of protein in 20 µl) mixed with 15 µl of 1% pectin and 115 µl of 50 mM sodium acetate buffer (pH 5.0). The mixture was incubated at 50^{0}C for 1 h, and then 50 µl of 3,5-Dinitrosalicylic acid (DNSA) reagent was added and absorbance measured at 540 nm as above. The specific activity was calculated using a glucose standard curve (0-20 mM) and background activity was corrected by subtracting final values from initial values. Specific activity was expressed in U/mg of protein, with 1 U defined as the amount of enzyme resulting in production of 1 μmol of glucose per min at pH 5.0 and 50^0C.

All activity assays were carried out using at least three biological and three technical replicates. The statistical design for each activity assay was a completely randomized design with a 2x2x2 factorial. Statistical analyses were performed through SAS (SAS Institute, Inc., Cary, NC) using a mixed model analysis of variance. Prior to analysis, data that failed to pass the Shapiro-Wilk normality test were log transformed. Least square means were separated using Tukey's option and significant differences were considered at $P < 0.05$.

Results

High cellulase activity and gut morphology and histology in Zygentoma

As a part of a quantitative prospecting effort to identify insects with high cellulolytic activity (Oppert et al., 2010), we detected species in Zygentoma as having the highest relative cellulolytic (endoglucanase, CMCase) activity among all taxonomic orders tested (Fig. 2.1A).

The digestive systems of two species of Zygentoma, *T. domestica* and *C. longicaudata* had similar morphology and histology, although the digestive tube in *C. longicaudata* was longer and larger than in *T. domestica* (Fig. 2.1B). Consequently, we focused on *C. longicaudata* for further characterization of the digestive tube due to its relatively bigger size. This digestive tube was longer than the insect body length and could be divided into foregut, midgut and hindgut regions. The foregut was the largest part of the digestive system and included an enlarged crop extending throughout the thoracic region and making up half of the digestive tube (Fig. 2.1C). It was observed in most dissections that among the three gut compartments, the food bolus was always found in the crop.

Histological observations of the crop wall in *C. longicaudata* identified a monolayer of epidermal cells supported by circular muscle cells (Fig. 2.2 A). The crop opened posteriorly into the proventriculus, which was highly muscular and had six sclerotized teeth-like structures (Fig. 2.2 B). The midgut was the second longest part of digestive system and appeared as a simple tube-like structure with gastric caecae at the anterior region. The midgut wall was characterized by the presence of a single layer of columnar cells with apical brush border membrane, and nidi of stem cells appeared interspersed in the epithelium (Fig. 2.2 C). The connection between midgut and hindgut was traced by the presence of Malpighian tubules, which were numerous in number and longer than the insect body length (Fig. 2.1 B). The hindgut was also short and simple tube-like structure with a monolayer of epidermal cells and ended in rectal pads (Fig. 2.2 D).

Qualitative and quantitative detection of cellulolytic activity in the digestive system of *T. domestica* and *C. longicaudata*

Zymograms of *T. domestica* gut fluids had more and brighter bands of activity against CMC compared to *C. longicaudata* (Fig. 2.3). When comparing among gut regions, higher cellulolytic activity was found in samples from the head and foregut compared to midgut and hindgut tissues (Fig. 2.3). Consequently, head and foregut tissues were selected for quantitative enzymatic assays for plant cell wall degrading enzyme (PCWDE) activities.

Activities tested quantitatively included endoglucanase, β-glucosidase, xylanase and pectinase in digestive fluids obtained from head and foregut samples of both *T. domestica* and *C. longicaudata* (Fig. 2.4). As observed in the qualitative zymograms, digestive fluids from both head and foregut tissues of *T. domestica* had significantly higher endoglucanase activity compared to *C. longicaudata* ($P < 0.05$). Within *T. domestica*, the digestive fluids from the foregut had significantly higher endoglucanase activity than fluids from head tissue ($P < 0.05$), while significant differences were not observed between samples from foregut and head tissues of *C. longicaudata* (Fig 2.4 A). Both *T. domestica* and *C. longicaudata* had no β-glucosidase activity in head fluids, however similar levels of β-glucosidase activity were found in the digestive fluids from foregut tissues of both species (Fig 2.4 B). Xylanase activity was significantly higher (about six-fold, $P < 0.05$) in the foregut fluids of *T. domestica* compared to *C. longicaudata*, and very small levels of xylanase activity were detected in head fluids from both insects (Fig. 2.4 C). Pectinase activity was absent from *C. longicaudata* and present in both head and foregut tissues of *T. domestica* (Fig. 2.4 D). Feeding both insects on a protein-rich (BSA) or a cellulose-rich (paper) diet did not result in significant differences in any of the tested enzyme activities ($P > 0.05$) (Fig. 2.4).

Discussion

Members of Zygentoma displayed significantly higher (>4-fold) cellulase activity compared to species in taxonomic orders traditionally considered as insect models for cellulase research, such as Coleoptera, Blattodea and Isoptera. Zygentoma is a basal hexapod group known to feed on highly lignocellulosic substrates (Berger, 1945) and to produce endogenous cellulases (Lasker and Giese, 1956; Zinkler and Götze, 1987; Treves and Martin, 1994). Recently, *T. domestica* has been identified to endogenously produce lytic polysaccharide monooxygenases (LPMOs) in addition to carbohydrate degrading enzymes (Sabbadin et al., 2018), which may explain the relatively highest cellulolytic activity in this group.

Although both *T. domestica* and *C. longicaudata* belong to Zygentoma and have similar digestive system morphology, we detected significant differences in cellulose activity between these species. Within a species, we also detected significant differences among fluids from different digestive regions in their ability to degrade different plant cell wall substrates. Nevertheless, in both the insect species the highest levels of enzymatic activity against the tested substrates were detected for digestive fluids from foregut compared to any other tissue in the digestive system. This observation is also supported by previous reports documenting higher endoglucanase and β-glucosidase activities in the foregut compared to other gut tissues in *T. domestica* (Zinkler and Gotze, 1987) and *Acrotelsa collaris* (Modder, 1964). In addition, cellulose fibers were reported to be digested in the crop of *C. longicaudata* (Lindsay, 1940). Localization of the main cellulolytic activity (endoglucanase, CMCase) to the foregut has also been reported in other arthropod groups, such as millipedes (*Chicobulus sp.*), desert locust (*Schistocerca gregaria*) and a longhorn beetle (*Hylotrupes bajules*) (Cazemier et al., 1997). In contrast, Lasker and Giese (1956) reported no cellulolytic activity in the fluids from the crop of *Ctenolepisma lineata*. Interestingly, in our histological sections the gut secretory columnar cells were only found in the midgut epithelium and were absent from the foregut. Recently, expression of LPMOs was localized to salivary glands, crop and midgut tissues of *T. domestica*, with relatively higher expression in the midgut tissue (Sabbadin et al., 2018). All these observations may suggest that both LPMOs and cellulases may be produced in the midgut and foregut (crop), although it is also possible that in some insects the enzymes may be secreted from midgut cells but flow towards the foregut (Terra, 1990; Terra and Ferreira, 1994). Taken together, the current evidence supports that the foregut is the most important tissue for plant cell wall digestion in Zygentoma.

Diverse PCWDEs, including cellulases such as endoglucanases and β-glucosidases; and hemicellulases like xylanases, were found in the digestive fluids of both *T. domestica* and *C. longicaudata*. In addition, *T. domestica* displayed pectinase activity, which indicates that *T. domestica* has all the necessary plant cell wall degrading enzymes to digest complex cellulolytic substrates. Comparatively, *T. domestica* had significantly higher levels of endoglucanase, xylanase and pectinase activities, which suggests a more efficient and complex cellulolytic system compared to *C. longicaudata*.

Feeding *T. domestica* and *C. longicaudata* a cellulose-rich diet did not result in increased production of cellulases, which suggests that cellulase production in these insects is not driven by diet. Similar results were reported from a gut proteome analysis of *T. domestica* fed on different cellulosic substrates, which did not alter production of carbohydrate digesting enzymes but increased abundance of LPMOs when fed on crystalline cellulose (Sabbadin et al., 2018). Consequently, it is plausible that in Zygentoma the production of cellulases remains constant while the production of LPMOs could be driven by the content of cellulose in the diet.

Identification of proteins in chromatographic fractions with CMCase activity in gut fluids of *T. domestica* (Pothula *et al*, submitted) revealed the presence of endoglucanases with similarity to enzymes from termites, beetles, and the herbivorous crustacean *Daphnia pulex*. Firebrat was reported to associate with five fungal species (*Mycotypha microspore*, *Aspergillus ochraceus*, *Aspergillus niger* and two species of *Penicillium*) and four bacterial species (*Enterobacter cloacae*, *Bacillus sps*., *Micrococcus sps*., and *Klebsiella sps*.) (Woodbury and Gries, 2013a). However, firebrats exhibit aggregation behavior only in the presence of *Enterobacter cloacae* and *Mycotypha microspore* due to the presence of D-glucose (Woodbury and Gries, 2013a). Of these two microbes, only the fungus *Mycotypha microspore* was able to degrade cellulose into glucose and the bacterium *Enterobacter cloacae* had D-glucose as a constituent of thick polysaccharide surface coating (Woodbury and Gries, 2013b). However, feeding firebrats with antibiotics resulted in significant reduction of gut microbial load but did not alter the cellulolytic activity of gut fluids (Treves and Martin, 1994). These results suggest that firebrats can digest cellulose in a symbiont-independent manner, but may need the aid of microbes in accessing cellulosic substrates. Accordingly, most of the identified proteins with CMCase activity in firebrat gut fluids were matched to insect genes (Pothula *et al*, submitted) indicating the endogenous origin of cellulases. Exceptions to this observation included two glucan endo-1,6-beta-glucosidases matching to *Haloplasma contractile* and *Paenibacillus* sp. JDR-2, which probably aid in cellulose digestion in firebrats but further research is needed to confirm their role.

Overall, our work suggests that members of Zygentoma express a repertoire of PCWDEs, including cellulases, xylanases and pectinases. Digestive fluids of *T. domestica* appeared significantly more active than in *C. longicaudata*, in both insects the highest levels of digestion were detected in the foregut. Considering the results in this work and the dearth of information on Zygentoma, we propose the need for further research to learn more on the evolution of PCWDE in these insects and test their capacity in prospecting for new enzymes for use in production of industrial cellulose digestion.

List of references

Abo-Khatwa, N., 1978. Cellulase of fungus-growing termites: A new hypothesis on its origin. Experientia 34, 559–560. https://doi.org/10.1007/BF01936956

Ali, S.S., Wu, J., Xie, R., Zhou, F., Sun, J., Huang, M., 2017. Screening and characterizing of xylanolytic and xylose-fermenting yeasts isolated from the wood-feeding termite, Reticulitermes chinensis. PLOS ONE 12, e0181141. https://doi.org/10.1371/journal.pone.0181141

Barnhart, C.S., 1961. The Internal Anatomy of the Silverfish Ctenolepisma campbelli and Lepisma saccharinum (Thysanura: Lepismatidae). Ann. Entomol. Soc. Am. 54, 177–196. https://doi.org/10.1093/aesa/54.2.177

Berger, B., 1945. Food preferences of the firebrat. J. Econ. Entomol. 38, 577–582. https://doi.org/10.1093/jee/38.5.577

Biedermann, W., Moritz, P., 1898. Beitr¨age zur vergeleichenden Physiologie der Verdauung. II. Uver ein cellulosel ¨ odendes Enzyme im Leversecret der Schnecke (Helix pomatia). Pfludg Arch Gem 73, 219–287.

Bones, A.M., Rossiter, J.T., 1996. The myrosinase-glucosinolate system, its organisation and biochemistry. Physiol. Plant. 97, 194–208. https://doi.org/10.1111/j.1399-3054.1996.tb00497.x

Brand, P., Robertson, H.M., Lin, W., Pothula, R., Klingeman, W.E., Jurat-Fuentes, J.L., Johnson, B.R., 2018. The origin of the odorant receptor gene family in insects [WWW Document]. eLife. https://doi.org/10.7554/eLife.38340

Brennan, Y., Callen, W.N., Christoffersen, L., Dupree, P., Goubet, F., Healey, S., Hernández, M., Keller, M.C., Li, K., Palackal, N., Sittenfeld, A., Tamayo, G., Wells, S., Hazlewood, G.P., Mathur, E.J., Short, J.M., Robertson, D.E., Steer, B.A., 2004. Unusual microbial xylanases from insect guts. Appl. Environ. Microbiol. 70, 3609–3617.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H., Oyama, K., 2012. Endogenous Plant Cell Wall Digestion: A Key Mechanism in Insect Evolution. Annu. Rev. Ecol. Evol. Syst. 43, 45–71. https://doi.org/10.1146/annurev-ecolsys-110411-160312

Cazemier, A.E., den Camp, H.J.M.O., Hackstein, J.H.P., Vogels, G.D., 1997. Fibre Digestion in Arthropods. Comp. Biochem. Physiol. A Physiol. 118, 101–109. https://doi.org/10.1016/S0300-9629(96)00443-4

Chatterjee, S., Sharma, S., Prasad, R.K., Datta, S., Dubey, D., Meghvansi, M.K., Vairale, M.G., Veer, V., 2015. Cellulase enzyme based biodegradation of cellulosic materials: an overview. South Asian J. Exp. Biol. 5, 271–282.

Cleveland, L.R., 1924. The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to reticulitermes flavipes kollar. Biol. Bull. 46, 178–201. https://doi.org/10.2307/1536507

Davison, A., Blaxter, M., 2005. Ancient Origin of Glycosyl Hydrolase Family 9 Cellulase Genes. Mol. Biol. Evol. 22, 1273–1284. https://doi.org/10.1093/molbev/msi107

Fischer, R., Ostafe, R., Twyman, R.M., 2013. Cellulases from Insects, in: Vilcinskas, A. (Ed.), Yellow Biotechnology II: Insect Biotechnology in Plant Protection and Industry, Advances in Biochemical Engineering/Biotechnology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 51–64. https://doi.org/10.1007/10_2013_206

Gilbert, H.J., 2010. The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. Plant Physiol. 153, 444–455. https://doi.org/10.1104/pp.110.156646

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883

Husebye, H., Arzt, S., Burmeister, W.P., Härtel, F.V., Brandt, A., Rossiter, J.T., Bones, A.M., 2005. Crystal structure at 1.1Å resolution of an insect myrosinase from Brevicoryne brassicae shows its close relationship to β-glucosidases. Insect Biochem. Mol. Biol. 35, 1311–1320. https://doi.org/10.1016/j.ibmb.2005.07.004

Kim, N., Choo, Y.M., Lee, K.S., Hong, S.J., Seol, K.Y., Je, Y.H., Sohn, H.D., Jin, B.R., 2008. Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket Teleogryllus emma. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 150, 368–376. https://doi.org/10.1016/j.cbpb.2008.04.005

Laibach, H., 1952. Lepisma saccharina L., das Silberflschchen. Z F Hyg Zool U Schddlingsbekdmpfung 40, 321–370.

Lasker, R., Giese, A.C., 1956. Cellulose Digestion By The Silverfish Ctenolepisma Lineata. J. Exp. Biol. 33, 542–553.

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Dae Sohn, H., Jin, B.R., 2004. cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, Apriona germari. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 139, 107–116. https://doi.org/10.1016/j.cbpc.2004.06.015

Lindsay, E., 1940. The Biology of the Silverfish, Ctenolepisma longicaudata Esch. with particular Reference to its feeding Habits. Proc. R. Soc. Vic. 52.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lynd, L.R., Cushman, J.H., Nichols, R.J., Wyman, C.E., 1991. Fuel Ethanol from Cellulosic Biomass. Science 251, 1318–1323. https://doi.org/10.1126/science.251.4999.1318

Martin, M.M., 1991. The evolution of cellulose digestion in insects. Phil Trans R Soc Lond B 333, 281–288. https://doi.org/10.1098/rstb.1991.0078

Martin, M.M., Martin, J.S., 1978. Cellulose Digestion in the Midgut of the Fungus-Growing Termite Macrotermes natalensis: The Role of Acquired Digestive Enzymes. Science 199, 1453–1455. https://doi.org/10.1126/science.199.4336.1453

Modder, W.W.D., 1975. Feeding and growth of Acrotelsa collaris (Fabricius) (Thysanura, Lepismatidae) on different types of paper. J. Stored Prod. Res. 11, 71–74. https://doi.org/10.1016/0022-474X(75)90042-9

Oppert, C., Klingeman, W.E., Willis, J.D., Oppert, B., Jurat-Fuentes, J.L., 2010. Prospecting for cellulolytic activity in insect digestive fluids. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 155, 145–154. https://doi.org/10.1016/j.cbpb.2009.10.014

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419. https://doi.org/10.1038/nmeth.4197

Pauchet, Y., Heckel, D.G., 2013. The genome of the mustard leaf beetle encodes two active xylanases originally acquired from bacteria through horizontal gene transfer. Proc. R. Soc. Lond. B Biol. Sci. 280, 20131021. https://doi.org/10.1098/rspb.2013.1021

Rouland, C., Civas, A., Renoux, J., Petek, F., 1988. Purification and properties of cellulases from the termite Macrotermes mülleri (termitidae, macrotermitinae) and its symbiotic fungus termitomyces sp. Comp. Biochem. Physiol. Part B Comp. Biochem. 91, 449–458. https://doi.org/10.1016/0305-0491(88)90004-1

Rouland, C., Lenoir, F., Lepage, M., 1991. The role of the symbiotic fungus in the digestive metabolism of several species of fungus-growing termites. Comp. Biochem. Physiol. A Physiol. 99, 657–663. https://doi.org/10.1016/0300-9629(91)90146-4

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K., Jones, T.A., 1990. Three-dimensional structure of cellobiohydrolase II from Trichoderma reesei. Science 249, 380–386. https://doi.org/10.1126/science.2377893

Sabbadin, F., Hemsworth, G.R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T., Marques, R.D.S., Sweeney, S.T., Besser, K., Elias, L., Pesante, G., Li, Y., Dowle, A.A., Bates, R., Gomez, L.D., Simister, R., Davies, G.J., Walton, P.H., Bruce, N.C., McQueen-Mason, S.J., 2018. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. Nat. Commun. 9, 756. https://doi.org/10.1038/s41467-018-03142-x

Sahrhage, D., 1953. Okologische Untersuchengen an Thermobia domestica (Packard) und Lepisma saccharina L. Z Wiss Zool 157, 77–168.

Scrivener, A.M., Slaytor, M., 1994. Properties of the endogenous cellulase from Panesthia cribrata saussure and purification of major endo-β-1,4-glucanase components. Insect Biochem. Mol. Biol. 24, 223–231. https://doi.org/10.1016/0965-1748(94)90001-9

Shi, W., Ding, S.-Y., Yuan, J.S., 2011. Comparison of Insect Gut Cellulase and Xylanase Activity Across Different Insect Species with Distinct Food Sources. BioEnergy Res. 4, 1–10. https://doi.org/10.1007/s12155-010-9096-0

Terra, W.R., 1990. Evolution of Digestive Systems of Insects. Annu. Rev. Entomol. 35, 181–200. https://doi.org/10.1146/annurev.en.35.010190.001145

Terra, W.R., Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. Part B Comp. Biochem. 109, 1–62. https://doi.org/10.1016/0305-0491(94)90141-4

Terry, M.D., Whiting, M.F., 2005. Mantophasmatodea and phylogeny of the lower neopterous insects. Cladistics 21, 240–257. https://doi.org/10.1111/j.1096-0031.2005.00062.x

Tomme, P., Warren, R.A.J., Gilkes, N.R., 1995. Cellulose Hydrolysis by Bacteria and Fungi, in: Poole, R.K. (Ed.), Advances in Microbial Physiology. Academic Press, pp. 1–81. https://doi.org/10.1016/S0065-2911(08)60143-5

Treves, D.S., Martin, M.M., 1994. Cellulose digestion in primitive hexapods: Effect of ingested antibiotics on gut microbial populations and gut cellulase levels in the

firebrat, Thermobia domestica (Zygentoma, Lepismatidae). J. Chem. Ecol. 20, 2003–2020. https://doi.org/10.1007/BF02066239

Ververis, C., Georghiou, K., Christodoulakis, N., Santas, P., Santas, R., 2004. Fiber dimensions, lignin and cellulose content of various plant materials and their suitability for paper production. Ind. Crops Prod. 19, 245–254. https://doi.org/10.1016/j.indcrop.2003.10.006

Villares, A., Moreau, C., Bennati-Granier, C., Garajova, S., Foucat, L., Falourd, X., Saake, B., Berrin, J.-G., Cathala, B., 2017. Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. Sci. Rep. 7, 40262. https://doi.org/10.1038/srep40262

Watanabe, H., Noda, H., Tokuda, G., Lo, N., 1998. A cellulase gene of termite origin. Nat. Lond. 394, 330–1. http://dx.doi.org/10.1038/28527

Watanabe, H., Tokuda, G., 2010. Cellulolytic Systems in Insects. Annu. Rev. Entomol. 55, 609–632. https://doi.org/10.1146/annurev-ento-112408-085319

Willis, J.D., Oppert, B., Oppert, C., Klingeman, W.E., Jurat-Fuentes, J.L., 2011. Identification, cloning, and expression of a GHF9 cellulase from Tribolium castaneum (Coleoptera: Tenebrionidae). J. Insect Physiol. 57, 300–306. https://doi.org/10.1016/j.jinsphys.2010.11.019

Willis, J.D., Oppert, C., Jurat-Fuentes, J.L., 2010. Methods for discovery and characterization of cellulolytic enzymes from insects. Insect Sci. 17, 184–198. https://doi.org/10.1111/j.1744-7917.2010.01322.x

Woodbury, N., Gries, G., 2013a. Firebrats, Thermobia domestica, aggregate in response to the microbes Enterobacter cloacae and Mycotypha microspora. Entomol. Exp. Appl. 147, 154–159. https://doi.org/10.1111/eea.12054

Woodbury, N., Gries, G., 2013b. How Firebrats (Thysanura: Lepismatidae) Detect and Nutritionally Benefit From Their Microbial Symbionts Enterobacter cloacae and Mycotypha microspora. Environ. Entomol. 42, 860–867. https://doi.org/10.1603/EN13104

Zhao, H., Baker, G.A., Song, Z., Olubajo, O., Crittle, T., Peters, D., 2008. Designing enzyme-compatible ionic liquids that can dissolve carbohydrates. Green Chem. 10, 696. https://doi.org/10.1039/b801489b

Zhao, H., Jones, C.L., Baker, G.A., Xia, S., Olubajo, O., Person, V.N., 2009. Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. J. Biotechnol. 139, 47–54. https://doi.org/10.1016/j.jbiotec.2008.08.009

Zhu, Z., Sathitsuksanoh, N., Vinzant, T., Schell, D.J., McMillan, J.D., Zhang, Y.-H.P., 2009. Comparative study of corn stover pretreated by dilute acid and cellulose solvent-based lignocellulose fractionation: Enzymatic hydrolysis, supramolecular structure, and substrate accessibility. Biotechnol. Bioeng. 103, 715–724. https://doi.org/10.1002/bit.22307

Zinkler, D., Götze, M., 1987. Cellulose digestion by the firebrat Thermobia domestica. Comp. Biochem. Physiol. Part B Comp. Biochem. 88, 661–666. https://doi.org/10.1016/0305-0491(87)90360-9

Appendix 2

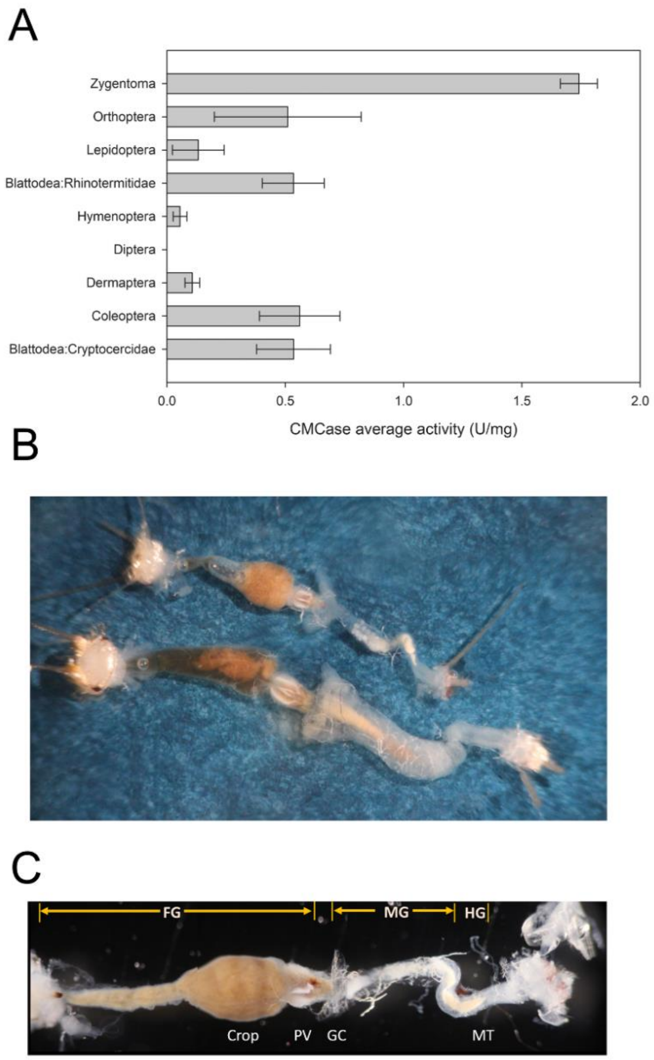

Figure 2.1. Relative CMCase activity in Zygentoma compared to other insect orders and structural comparison of the digestive tract in firebrat (*Thermobia domestica*) and silverfish (*Ctenolepisma longicaudata*). A) Activity (U/mg) of gut digestive fluids against CMC substrate in the most active samples from species of Zygentoma (*C. longicaudata*), and representative species from Orthoptera (*Conocephalus strictus*), Lepidoptera (*Halysidota tessellaris*), Blattodea (formerly Isoptera): Rhinotermitidae (*Reticulitermes flavipes*), Hymenoptera (*Neodiprion lecontei*), Diptera (*Monarthropalpus flavus*), Dermaptera (*Forficula auricularia*), Coleoptera (*Scolytinae spp*.) and Blattodea (formerly Blattaria): Cryptocercidae (*Cryptocercus spp*.). Shown are the average activity and corresponding standard error from at least three biological replicates performed in triplicate for each species. All activity assay experiments were concurrent, but all activities except for the Zygentoma sample were taken from Oppert et al. (2010). B) Dissected digestive tracts of firebrat (top) and silverfish (bottom). Note the relatively larger size of the tract in silverfish compared to firebrat. C) Morphological parts of the digestive tract of firebrat. FG, foregut; MG, midgut; HG, hindgut; GC, gastric caecae; MT, Malpighian tubules.

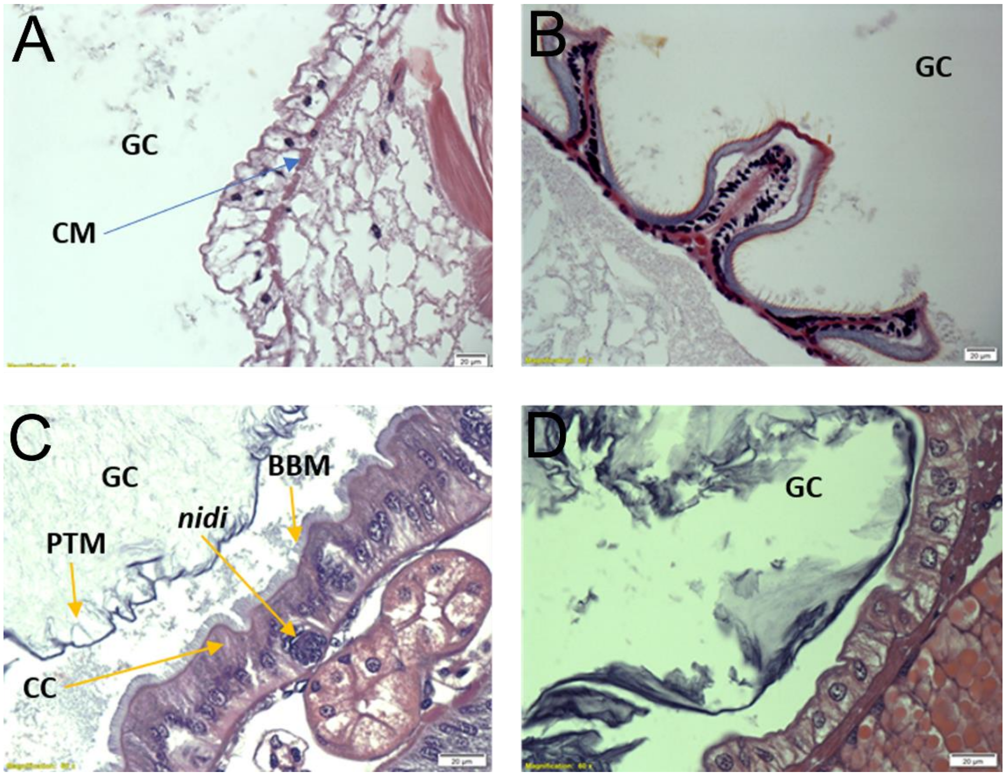

Figure 2.2. Histology of the digestive system regions in *Ctenolepisma longicaudata*. A) Longitudinal section of the crop showing the monolayer of epidermal cells and the underlying circular muscle cells. B) Longitudinal section of proventriculus. C) Longitudinal section of midgut showing peritrophic membrane, columnar cells lined with brush border membrane and intermitted by a group of nidi cells at the bottom. D) Longitudinal section of hindgut wall showing the monolayer of epidermal cells. All sections were stained with hematoxylin and eosin stain. GC, gut cavity; CM, circular muscle cells; PTM, peritrophic matrix; CC, columnar cells; BBM, brush border membrane.

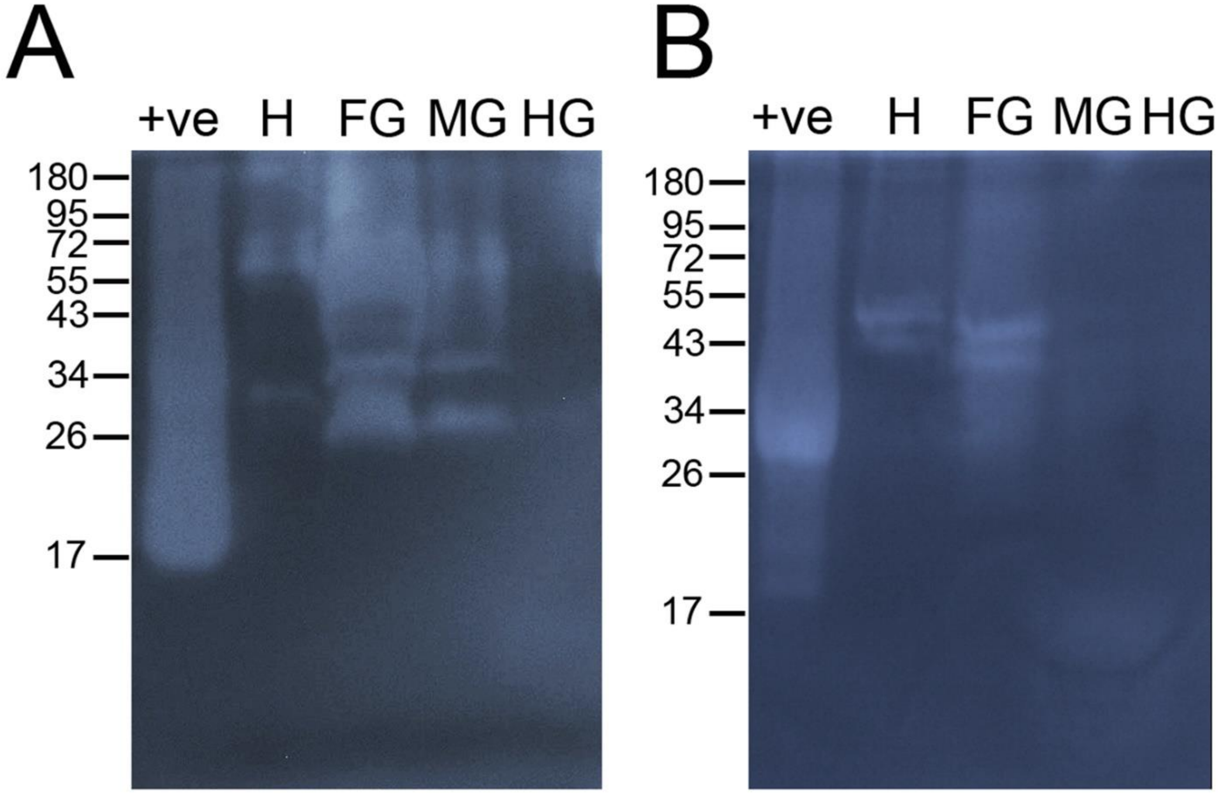

Figure 2.3. Detection of cellulolytic activity in digestive fluids of *Thermobia domestica* and *Ctenolepisma longicaudata*. Zymograms with 0.2% carboxymethyl cellulose were used to detect the cellulolytic (CMCase) activity in head, foregut, midgut and hindgut tissues of *T. domestica* (left) and *C. longicaudata* (right). Numbers indicate pre-stained protein molecular marker; +ve, commercial cellulase used as positive control; H, head; FG, foregut; MG, midgut; HG, hindgut.

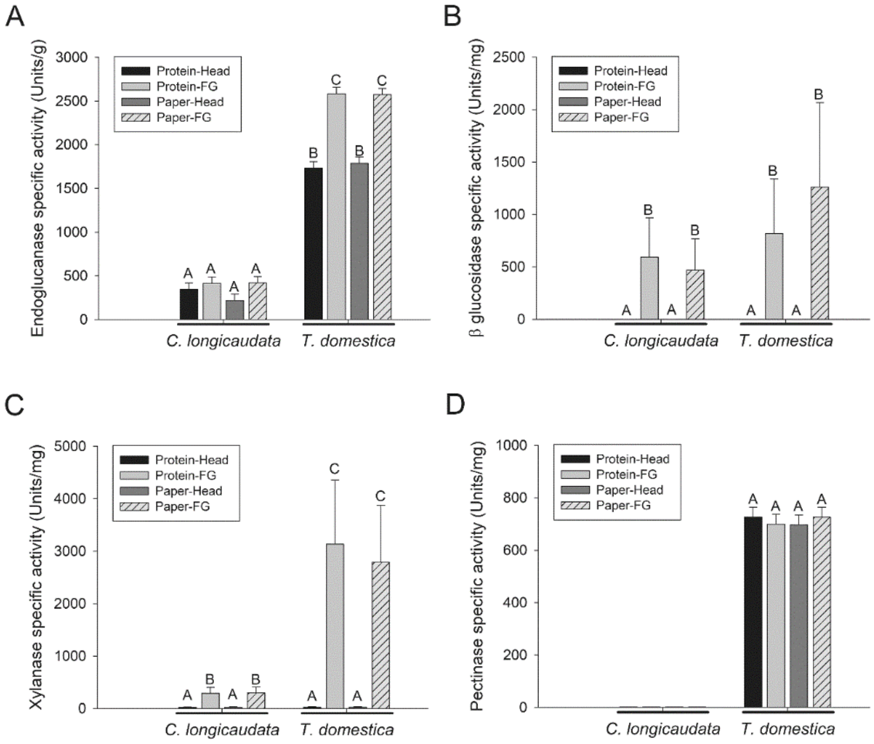

Figure 2.4. Quantification of plant cell wall degrading enzyme activities in the fluids derived from head and foregut tissues of *Ctenolepisma longicaudata* and *Thermobia domestica*. Fluids from head and foregut (FG) tissues of *C. longicaudata* and *T. domestica* fed on protein (BSA) or paper diet (see Materials and Methods) were used in assays to detect A) endoglucanase activity against 4-nitrophenyl-β-D-cellopentaoside (BPNPG5), B) β-glucosidase activity against 4-nitrophenyl β-D-cellobioside (pNPC), C) xylanase activity against 4-nitrophenyl β-D-xylpyranoside (pNPX), and D) pectinase activity against pectin from citrus peel. Shown are the means and corresponding standard errors calculated from three biological and three technical replicates. Different letters above the bars indicate significant differences in the mean activity ($P < 0.05$). Units of specific enzyme activity are per mg of protein in all the graphs except in graph A) where it is expressed per g of protein. One unit of enzyme activity was defined as the amount of enzyme required to release 1 μmol of 4-nitrophenol from the respective substrate in all the graphs except in D), where it is 1 μmol of glucose.

Chapter 3

Annotation of plant cell wall degrading enzymes (PCWDEs) among coding sequences from genomes of *Thermobia domestica* and *Ctenolepisma longicaudata*

Pothula, R.; Johnson, B.R.; Klingeman, W.E. and J.L. Jurat-Fuentes. (2018).

My contributions included: (1) planning and performing experiments, (2) data collection and analysis, (3) writing the manuscript and making figures. Brian R. Johnson assisted with 2, William E. Klingeman provided insects and Juan Luis Jurat-Fuentes assisted with (1 and 2).

Abstract

In the last decade, insects have emerged as a prospecting resource for new plant cell wall degrading enzymes (PCWDEs) given their endogenous production of cellulases, hemicellulases and pectinases. Although insects belonging to 16 taxonomic orders have been reported to have endogenous production of one or more of PCWDEs, molecular evidence has confirmed the presence of PCWDEs in insects from only eight taxonomic orders. We have identified members of Zygentoma as having the highest relative cellulolytic activity against carboxymethylcellulose compared to other insect groups, including termites and cockroaches. Consequently, in the present work our goal was to find PCWDEs genes present in the *Thermobia domestica* and *Ctenolepisma longicaudata* genomes. Annotation of predicted coding sequences from genomes of *T. domestica* and *C. longicaudata* reported numerous genes encoding for endoglucanases, glucosidases, β-1,3-glucanases, maltases, amylases, mannosidases, glucuronidases and lytic polysaccharide monoxygenases (LPMOs), which may help explain the relatively high cellulolytic activity displayed by these compared to other insects. Additionally, except LPMOs, the majority of the coding sequences encoding for different glycoside hydrolases were most similar to

Blattodea, which indicates the conservation of PCWDE genes through evolution in insects. Our study contributes to enhance the availability of annotated genetic information on insect PCWDEs in general, and especially in a primitive insect group.

Introduction

Plant cell walls are composed of cellulose, hemicellulose, pectin and lignin. Cellulose is considered the most available renewable energy source on earth (Lynd et al., 1991). Many organisms, including insects, are able to feed and digest plant material to obtain energy. In the last decade, insects have emerged as a prospecting resource for new plant cell wall degrading enzymes (PCWDs) given their endogenous production of cellulases, hemicellulases and pectinases (Calderón-Cortés et al., 2012; Watanabe and Tokuda, 2010).

Cellulases are a group of glycosyl hydrolase enzymes that aid in complete digestion of cellulose to glucose, which in biorefineries can be fermented by yeast to generate bioethanol. This group of enzymes includes endoglucanases, which cleave the cellulose chain internally at random locations, exoglucanases that cleave the cellulose chain from the ends releasing two molecules of glucose (cellobiose) and β-glucosidases, which degrade cellobiose to glucose subunits (Watanabe and Tokuda, 2010). Hemicellulases and pectinases are involved in the breakdown of hemicellulose and pectin polysaccharides, respectively, which are interlocked with cellulose in plant material (Gilbert, 2010).

Insects belonging to 16 taxonomic orders have been reported to have endogenous production of one or more of PCWDEs (Calderón-Cortés et al., 2012). However, molecular evidence has confirmed the presence and characterized PCWDEs from species in only eight taxonomic orders. These characterized enzymes include endoglucanases belonging to glycoside hydrolase (GH) families 9 and 45, β-glucosidases of GH 5, hemicellulases such as xyloglucanases of GH 5 and GH 11, β-1,3-glucanases of GH 16, and pectinases of GHF 28 (Calderón-Cortés et al., 2012). Therefore, there is still a dearth of molecular data for the identification of endogenous cellulases in insects, especially from basal hexapod groups.

We have recently identified members of Zygentoma as having the highest relative cellulolytic activity against carboxymethylcellulose compared to termites and cockroaches (Chapter 2).

Zygentoma is a basal hexapod group, with members feeding on highly cellulosic materials such as starch, paper and cardboard (Berger, 1945; Sahrhage, 1953). A defaunation study on *Ctenolepisma lineata* and *Thermobia domestica* supported endogenous production of cellulases in these species (Lasker and Giese, 1956; Treves and Martin, 1994). Biochemical evidence indicates that digestive fluids, especially from the foregut of *T. domestica*, contain endoglucanase, β-glucosidase, xylanase, pectinase, amylase, maltase, sucrase and lactase activities (Zinkler and Götze, 1987; Chapter 2). In addition to cellulases, investigation of the digestive proteome of *T. domestica* revealed the production of lytic polysaccharide monoxygenases (LPMOs), which are predicted to soften the cellulose fibers and make them more tractable to cellulases (Sabbadin et al., 2018). Even though the firebrat (*T. domestica*) and the silverfish (*Ctenolepisma longicaudata*) have similar gut morphohistology, *T. domestica* displayed higher cellulolytic activity than *C. longicaudata* (Chapter 2). Consequently, in the present work our goal was to find endogenous PCWDE genes in the genome and confirm their expression from transcriptomes of *T. domestica* and *C. longicaudata*. Our analyses revealed that both *T.*

domestica and *C. longicaudata* contain and express numerous genes encoding for endoglucanases, glucosidases, β-1,3-glucanases, maltases, amylases, mannosidases, glucuronidases and LPMOs, which may help explain the relatively high cellulolytic activity displayed by these compared to other insects. Interestingly, pectinases were not found in both species and only three contigs containing xylanase genes were found in *C. longicaudata*. Except LPMOs, the majority of the contig sequences with genes encoding for different glycoside hydrolases had highest sequence similarity to Blattodea, which indicates conservation of PCWDE genes in insects through evolution. Our study contributes to enhance the availability of annotated genetic information on insect PCWDEs in general, especially in a primitive insect group.

Materials and Methods

Insect rearing

Adult silverfish (*Ctenolepisma longicaudata*) and firebrat (*Thermobia domestica*) were used to prepare the genetic materials used for sequencing of genomes used in this study. Several batches of nymphs and adults of *C. longicaudata* were hand-collected by sweeping into collection containers with a 3” long soft bristled brush or collected into lid-less plastic dishes (Pioneer Plastics Inc., Dixon, KY) baited with whole-grain oat flakes (Quick 1-minute Oats, Quaker Oats Co., Chicago, IL) in buildings and reared in the laboratory on rolled oats, paper and Pedigree adult complete nutrition dry dog food (Mars, Inc., Mount Olive, NJ), at room temperature. An established *T. domestica* colony which was originally derived from the culture being maintained at the Department of Entomology at Ohio State University (Columbus, OH) was kindly provided by Patrick Stanley and Eric Snell

(Snell Scientific, LLC, Meansville, GA). This *T. domestica* colony had been maintained in the laboratory for > 8 years at the Department of Entomology and Plant Pathology at the University of Tennessee on printing paper as a carbohydrate resource and NatureWise chick starter grower feed (Nutrena, Minneapolis, MN) as protein source in a dark incubator set at 34^0C.

Genome assembly and annotation

The draft genome of *T. domestica* was sequenced and assembled as described elsewhere (Brand et al., 2018). The genome of *C. longicaudata* was assembled by following the same bioinformatics pipeline as that of *T. domestica* (B. Johnson, unpublished). Since annotated genomes of Zygentoma members or closely related insects were not available, all the predicted coding sequences from each genome file were pooled into a Fasta file and annotated using Blast2GO (https://www.blast2go.com/) against the NCBI translated nucleotide, nr (arthropoda) database by setting the expected E value to 0.001. Sequences were also annotated using Interproscan in Blast2GO and then merged with the blast annotations. Blast, project and enzyme code statistics of the annotated coding sequences of *T. domestica* and *C. longicaudata* were generated using Blast2GO. Annotated coding sequences of *T. domestica* and *C. longicaudata* were manually screened to identify PCWDEs and LPMOs. Transcriptomes of *T. domestica* and *C. longicaudata* were blasted against resulted coding sequences of PCWDEs and LPMOs to find the expressed PCWDE and LPMO genes with an expected E value set at 10^{-6}.

Results

T. domestica annotation

The genome of *T. domestica* had 159,864 coding sequences, of which only 17% (27,751) had Blast2Go annotations, 32% (50,359) had gene ontology (GO) mapping, 20% (32,716) had BLAST hits and 31% (32,716) returned no BLAST hits (Fig. 3.1). The majority of BLAST hits were to homologs from insects and other arthropods. The species with the highest number and similarity of matching coding sequences was *Cryptotermes secundus* (Fig. 3.2).

PCWDEs in *T. domestica*

Based on the enzyme codes assigned by Blast2GO, approximately 3,300 coding sequences were hydrolases and 700 sequences were oxidoreductases, which include LPMOs (Fig. 3.3). Among the 3,300 hydrolases, only approximately 240 sequences were identified as glycosylases (Fig. 3.4). Interestingly, in *T. domestica*, 501 predicted coding sequences encoded enzymes belonging to diverse glycoside hydrolase families involved in cellulose digestion. Among them, endoglucanases were dominant with 85 genes encoding predicted endoglucanases belonging to GH 9 (Fig. 3.5). Other enzymes identified included α-glucosidases of GH 31 (32 sequences), β-glucosidases of GH 1 (19 sequences), β-1,3-glucanases of GH 55 (9 sequences), β-galactosidases of GH 1 (21 sequences), α-mannosidases of GH 31 (73 sequences), β-mannosidases of GH 5 (26 sequences), β-glucuronidase of GH 2 (39 sequences), maltase, maltase-glucoamylase (36 sequences) and sucrase-isomaltase (3 sequences) of GH 31, amylases of GH 13 (24 sequences), and myrosinase of GH 1 (56 sequences). Along with these hydrolases, 59 predicted coding sequences matched to LPMOs of auxillary activity family 15, and 88 sequences matched to chitinases of GH 18 (Fig. 3.5). Among 501 predicted coding sequences from the genome, 393 were expressed in the transcriptome of *T. domestica*. Majority of the expressed genes included mannosidases (71) followed by endoglucanases (62), chitinases (60), myrosinases (39), LPMOs (32), glucosidases (31), β-glucuronidase (29), maltases (23), amylases (19), galactosidases (13) and β-1,3-glucanases (5).

Except for LPMOs which matched to genes in *T. domestica*, all the sequences encoding for different PCWDEs were most similar to homologs in other insects and ono-insect arthropods. Most of the endoglucanases (51), glucosidases (35), β-1,3-glucanases (5), mannosidases (29), and maltases (20) matched with highest identity to enzymes from members of Blattodea, which includes termites and cockroaches (Table 3.1, 3.2, 3.3, 3.4 and 3.6). On the other hand, the majority of amylases (10) matched to Hymenoptera (Table 3.7) and most of the myrosinases (18) were most similar to Lepidoptera (Table 3.9). The majority of glucuronidases and galactosidases were most similar to enzymes in other arthropods excluding insects (Table 3.5 and 3.8). All but nine LPMOs, were most similar to other arthropod members and to a coleopteran, *Anoplophora glabripennis* (Table 3.10).

***C. longicaudata* annotation**

The genome of *C. longicaudata* had 171,652 coding sequences, which is 6.9% more, and a higher percentage of coding sequences (30%=52,254) had Blast2GO annotations compared to *T. domestica*. However, only 20% of the predicted coding sequences had GO mapping and 18% had BLAST hits, which were lower compared to *T. domestica*. About 27% of coding sequences did not have any blast hits (Fig. 3.6). Similarly to *T. domestica*, the majority of sequences in *C. longicaudata* had homologs in insects from other orders and other arthropods, with most of the sequences having highest identity to sequences from *Cryptotermes secundus* (Fig. 3.7).

PCWDEs in *C. longicaudata*

Similarly to *T. domestica*, the enzyme code statistics of *C. longicaudata* also revealed that the majority of coding sequences (approximately 4,000) were encoding proteins that belong to hydrolases followed by transferases (approximately 1,600) and oxidoreductases (approximately 800) enzyme groups (Fig. 3.8). However, among all hydrolases, only 240 coding sequences were found matching to glycosylases (Fig. 3.9). A total of 580 coding sequences were found in *C. longicaudata* encoding for enzymes belonging to diverse glycoside hydrolase families, among which chitinases of GH 18 (108 coding sequences) were dominant (Fig. 3.10). Other hydrolases included α-mannosidases of GH 31 (75 coding sequences) and β-mannosidases of GH 1 (32 coding sequences), myrosinases of GH 1 (96 coding sequences), glucosidases including α-glucosidases of GH 31 (50 coding sequences) and β-glucosidases (22 coding sequences) of GH 1, endoglucanases of GH 9 (69 coding sequences), maltases of GH 31 (33 coding sequences), β-glucuronidases of GH 2 (30 coding sequences), β-galactosidases of GH 1 (26 coding sequences), amylases of GH 13 (23 coding sequences), β-1,3-glucanases of GH 55 (16 coding sequences) and endo-1,4-beta-xylanases of GH 10 (3 coding sequences) (Fig. 3.10). Out of 580 predicted coding sequences from the genome, 517 were expressed in the transcriptome of *C. longicaudata*. The majority of the expressed genes included chitinases (89), followed by mannosidases (87), endoglucanases (62), myrosinases (60), glucosidases (56), LPMOs (50), maltases (29), β-glucuronidase

(24), galactosidases (22), amylases (20), β-1,3-glucanases (15) and endo-1,4-beta-xylanases (3).

Among oxidoreductases found in *C. longicaudata* 67 genes encoded for LPMOs (Fig. 3.10). Similar to *T. domestica*, the majority of LPMOs (55) in *C. longicaudata* closely matched to *T. domestica* genes, with the remaining (12) coding sequences being most similar to Hemiptera, Hymenoptera and other arthropods (Table 3.21). Except LPMOs, all other coding sequences encoding for diverse glycoside hydrolases were most closely matched to enzymes from other insects and arthropods members outside of Zygentoma. The majority of endoglucanases, mannosidases, glucosidases, amylases, galactosidases, myrosinases and glucuronidases were most closely matched to enzymes from termites and cockroaches (Blattodea) (Table 3.11, 3.12, 3.14, 3.15, 3.17, 3.18 and 3.20). On the other hand, the majority of β-1,3-glucanases and maltases were most similar to Blattodea and non-insect arthropods (Table 3.13 and 3.16). Only three xylanases were found in *C. londicauadata*, two of which matched to non-isect arthropods and one matched to a hemipteran (Table 3.19).

Discussion

In our previous work (Chapter 2), we found that members of Zygentoma, especially *T. domestica*, display highest relative cellulolytic activity compared to other insects, including termites. In the present work, we present the annotation and identification of PCWDEs from the genome of *T. domestica* and *C. longicaudata* as representative members of Zygentoma. Annotation of all coding sequences for both *T. domestica* and *C. longicaudata* revealed that about 1/3 of coding sequences in both species did not yield any blast hits, which indicates the dearth of genetic information on primitive insect groups. Additionally, most of the sequences encoding for PCWDEs had highest identity to genes in Blattodea (termites and cockroaches), which suggests the conservation of genes encoding PCWDEs through evolution.

Both *T. domestica* and *C. longicaudata* had numerous coding sequences encoding for diverse glycoside hydrolases. However, *T. domestica* had 85 sequences encoding for endoglucanases (Table 3.1) while *C. longicaudata* had 69 (Table 3.11), which may explain the higher endoglucanase activity reported in *T. domestica* compared to *C. longicaudata* (Chapter 2). On the other hand, genomes of both species yielded nearly equal number of β-glucosidase genes (Table 3.2 and 3.12), which was reflected in similar enzyme activity levels (Chapter 2).

In contrast to cellulases, xylanases are rarely described as endogenously produced in insects (Calderón-Cortés et al., 2012). In agreement with this observation, only *C. longicaudata* had three sequences matching to xylanases, while *T. domestica* had no coding sequences encoding for xylanases. However, it is possible that hemicellulose could be digested in these insects by other enzymes, such as mannanases, α-glucuronidases, endoglucanases and β-1,3-glucanases, which were present in both species (Calderón-Cortés et al., 2012). Similarly, pectinases, which are not commonly found in insects, were absent from *T. domestica* and *C. longicaudata*. One possibility to explain the lack of pectinases may be that they may be produced by microorganisms present in the gut fluids. In addition to cellulases, several enzymes involved in digestion of starch and other polysaccharides, such as maltases, amylases, and mannosidases, were found in both species. The highest number of sequences among these enzyme groups was found for myrosinases, which play an important role in plant defense against herbivores (Husebye et al., 2005). Myrosinases have also been reported from aphids and the crystal structure of a myrosinase from *Brevicoryne brassicae* revealed its highest similarity with β-glucosidase (Bones and Rossiter, 1996; Husebye et al., 2005). In our analysis, although the sequence descriptions matched to myrosinases, in most instances the corresponding blast-hit description identified the sequence as a β-glucosidase (Table 3.9 and 3.20).

Both *T. domestica* and *C. longicaudata* were found to encode LPMOs, with 59 and 67 sequences identified, respectively. In comparison, a previous study on the digestive proteome and transcriptome of *T. domestica* was able to identify and annotate 21 LPMO genes (Sabbadin et al., 2018). Moreover, in our study we found 9 and 12 coding sequences of *T. domestica* and *C. longicaudata*, respectively, matching to LPMOs in other insect orders and to other arthropod groups, which probably indicates the depth of our genome coverage.

Overall, the genomes of *T. domestica* and *C. longicaudata* yielded a diverse array of PCWDEs, which indicate their ability to breakdown and digest cellulose completely. The presence of a high number of endoglucanases and β-glucosidases in both insect species could explain their relatively higher cellulolytic activity compared to other insect groups (Table 3.22). Although most of the PCWDE encoding sequences had highest identity to genes from Blattodea, numerous sequences from all enzyme classes matched to other arthropod groups including primitive collembola as well as highly and recently evolved insects such as hymenopterans. The extent of PCWDE homology within Inecta and other arthropod groups may suggest the evolution of cellulases in insects from a common ancestor rather than through horizontal transfer from microbes. This work contributes to increase the availability of insect glycosyl hydrolase annotated sequences in general, and especially for a primitive insect group.

List of references

Abo-Khatwa, N., 1978. Cellulase of fungus-growing termites: A new hypothesis on its origin. Experientia 34, 559–560. https://doi.org/10.1007/BF01936956

Ali, S.S., Wu, J., Xie, R., Zhou, F., Sun, J., Huang, M., 2017. Screening and characterizing of xylanolytic and xylose-fermenting yeasts isolated from the wood-feeding termite, Reticulitermes chinensis. PLOS ONE 12, e0181141. https://doi.org/10.1371/journal.pone.0181141

Barnhart, C.S., 1961. The Internal Anatomy of the Silverfish Ctenolepisma campbelli and Lepisma saccharinum (Thysanura: Lepismatidae). Ann. Entomol. Soc. Am. 54, 177–196. https://doi.org/10.1093/aesa/54.2.177

Berger, B., 1945. Food preferences of the firebrat. J. Econ. Entomol. 38, 577–582. https://doi.org/10.1093/jee/38.5.577

Biedermann, W., Moritz, P., 1898. Beitr¨age zur vergeleichenden Physiologie der Verdauung. II. Uver ein cellulosel ¨ odendes Enzyme im Leversecret der Schnecke (Helix pomatia). Pfludg Arch Gem 73, 219–287.

Bones, A.M., Rossiter, J.T., 1996. The myrosinase-glucosinolate system, its organisation and biochemistry. Physiol. Plant. 97, 194–208. https://doi.org/10.1111/j.1399-3054.1996.tb00497.x

Brand, P., Robertson, H.M., Lin, W., Pothula, R., Klingeman, W.E., Jurat-Fuentes, J.L., Johnson, B.R., 2018. The origin of the odorant receptor gene family in insects [WWW Document]. eLife. https://doi.org/10.7554/eLife.38340

Brennan, Y., Callen, W.N., Christoffersen, L., Dupree, P., Goubet, F., Healey, S., Hernández, M., Keller, M.C., Li, K., Palackal, N., Sittenfeld, A., Tamayo, G., Wells, S., Hazlewood, G.P., Mathur, E.J., Short, J.M., Robertson, D.E., Steer, B.A., 2004. Unusual microbial xylanases from insect guts. Appl. Environ. Microbiol. 70, 3609–3617.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H., Oyama, K., 2012. Endogenous Plant Cell Wall Digestion: A Key Mechanism in Insect Evolution. Annu. Rev. Ecol. Evol. Syst. 43, 45–71. https://doi.org/10.1146/annurev-ecolsys-110411-160312

Cazemier, A.E., den Camp, H.J.M.O., Hackstein, J.H.P., Vogels, G.D., 1997. Fibre Digestion in Arthropods. Comp. Biochem. Physiol. A Physiol. 118, 101–109. https://doi.org/10.1016/S0300-9629(96)00443-4

Chatterjee, S., Sharma, S., Prasad, R.K., Datta, S., Dubey, D., Meghvansi, M.K., Vairale, M.G., Veer, V., 2015. Cellulase enzyme based biodegradation of cellulosic materials: an overview. South Asian J. Exp. Biol. 5, 271–282.

Cleveland, L.R., 1924. The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to reticulitermes flavipes kollar. Biol. Bull. 46, 178–201. https://doi.org/10.2307/1536507

Davison, A., Blaxter, M., 2005. Ancient Origin of Glycosyl Hydrolase Family 9 Cellulase Genes. Mol. Biol. Evol. 22, 1273–1284. https://doi.org/10.1093/molbev/msi107

Fischer, R., Ostafe, R., Twyman, R.M., 2013. Cellulases from Insects, in: Vilcinskas, A. (Ed.), Yellow Biotechnology II: Insect Biotechnology in Plant Protection and Industry, Advances in Biochemical Engineering/Biotechnology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 51–64. https://doi.org/10.1007/10_2013_206

Gilbert, H.J., 2010. The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. Plant Physiol. 153, 444–455. https://doi.org/10.1104/pp.110.156646

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883

Husebye, H., Arzt, S., Burmeister, W.P., Härtel, F.V., Brandt, A., Rossiter, J.T., Bones, A.M., 2005. Crystal structure at 1.1Å resolution of an insect myrosinase from Brevicoryne brassicae shows its close relationship to β-glucosidases. Insect Biochem. Mol. Biol. 35, 1311–1320. https://doi.org/10.1016/j.ibmb.2005.07.004

Kim, N., Choo, Y.M., Lee, K.S., Hong, S.J., Seol, K.Y., Je, Y.H., Sohn, H.D., Jin, B.R., 2008. Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket Teleogryllus emma. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 150, 368–376. https://doi.org/10.1016/j.cbpb.2008.04.005

Laibach, H., 1952. Lepisma saccharina L., das Silberflschchen. Z F Hyg Zool U Schddlingsbekdmpfung 40, 321–370.

Lasker, R., Giese, A.C., 1956. Cellulose Digestion By The Silverfish Ctenolepisma Lineata. J. Exp. Biol. 33, 542–553.

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Dae Sohn, H., Jin, B.R., 2004. cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, Apriona germari. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 139, 107–116. https://doi.org/10.1016/j.cbpc.2004.06.015

Lindsay, E., 1940. The Biology of the Silverfish, Ctenolepisma longicaudata Esch. with particular Reference to its feeding Habits. Proc. R. Soc. Vic. 52.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lynd, L.R., Cushman, J.H., Nichols, R.J., Wyman, C.E., 1991. Fuel Ethanol from Cellulosic Biomass. Science 251, 1318–1323. https://doi.org/10.1126/science.251.4999.1318

Martin, M.M., 1991. The evolution of cellulose digestion in insects. Phil Trans R Soc Lond B 333, 281–288. https://doi.org/10.1098/rstb.1991.0078

Martin, M.M., Martin, J.S., 1978. Cellulose Digestion in the Midgut of the Fungus-Growing Termite Macrotermes natalensis: The Role of Acquired Digestive Enzymes. Science 199, 1453–1455. https://doi.org/10.1126/science.199.4336.1453

Modder, W.W.D., 1975. Feeding and growth of Acrotelsa collaris (Fabricius) (Thysanura, Lepismatidae) on different types of paper. J. Stored Prod. Res. 11, 71–74. https://doi.org/10.1016/0022-474X(75)90042-9

Oppert, C., Klingeman, W.E., Willis, J.D., Oppert, B., Jurat-Fuentes, J.L., 2010. Prospecting for cellulolytic activity in insect digestive fluids. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 155, 145–154. https://doi.org/10.1016/j.cbpb.2009.10.014

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419. https://doi.org/10.1038/nmeth.4197

Pauchet, Y., Heckel, D.G., 2013. The genome of the mustard leaf beetle encodes two active xylanases originally acquired from bacteria through horizontal gene transfer. Proc. R. Soc. Lond. B Biol. Sci. 280, 20131021. https://doi.org/10.1098/rspb.2013.1021

Rouland, C., Civas, A., Renoux, J., Petek, F., 1988. Purification and properties of cellulases from the termite Macrotermes mülleri (termitidae, macrotermitinae) and its symbiotic fungus termitomyces sp. Comp. Biochem. Physiol. Part B Comp. Biochem. 91, 449–458. https://doi.org/10.1016/0305-0491(88)90004-1

Rouland, C., Lenoir, F., Lepage, M., 1991. The role of the symbiotic fungus in the digestive metabolism of several species of fungus-growing termites. Comp. Biochem. Physiol. A Physiol. 99, 657–663. https://doi.org/10.1016/0300-9629(91)90146-4

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K., Jones, T.A., 1990. Three-dimensional structure of cellobiohydrolase II from Trichoderma reesei. Science 249, 380–386. https://doi.org/10.1126/science.2377893

Sabbadin, F., Hemsworth, G.R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T., Marques, R.D.S., Sweeney, S.T., Besser, K., Elias, L., Pesante, G., Li, Y., Dowle, A.A., Bates, R., Gomez, L.D., Simister, R., Davies, G.J., Walton, P.H., Bruce, N.C., McQueen-Mason, S.J., 2018. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. Nat. Commun. 9, 756. https://doi.org/10.1038/s41467-018-03142-x

Sahrhage, D., 1953. Okologische Untersuchengen an Thermobia domestica (Packard) und Lepisma saccharina L. Z Wiss Zool 157, 77–168.

Scrivener, A.M., Slaytor, M., 1994. Properties of the endogenous cellulase from Panesthia cribrata saussure and purification of major endo-β-1,4-glucanase components. Insect Biochem. Mol. Biol. 24, 223–231. https://doi.org/10.1016/0965-1748(94)90001-9

Shi, W., Ding, S.-Y., Yuan, J.S., 2011. Comparison of Insect Gut Cellulase and Xylanase Activity Across Different Insect Species with Distinct Food Sources. BioEnergy Res. 4, 1–10. https://doi.org/10.1007/s12155-010-9096-0

Terra, W.R., 1990. Evolution of Digestive Systems of Insects. Annu. Rev. Entomol. 35, 181–200. https://doi.org/10.1146/annurev.en.35.010190.001145

Terra, W.R., Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. Part B Comp. Biochem. 109, 1–62. https://doi.org/10.1016/0305-0491(94)90141-4

Terry, M.D., Whiting, M.F., 2005. Mantophasmatodea and phylogeny of the lower neopterous insects. Cladistics 21, 240–257. https://doi.org/10.1111/j.1096-0031.2005.00062.x

Tomme, P., Warren, R.A.J., Gilkes, N.R., 1995. Cellulose Hydrolysis by Bacteria and Fungi, in: Poole, R.K. (Ed.), Advances in Microbial Physiology. Academic Press, pp. 1–81. https://doi.org/10.1016/S0065-2911(08)60143-5

Treves, D.S., Martin, M.M., 1994. Cellulose digestion in primitive hexapods: Effect of ingested antibiotics on gut microbial populations and gut cellulase levels in the

firebrat, Thermobia domestica (Zygentoma, Lepismatidae). J. Chem. Ecol. 20, 2003–2020. https://doi.org/10.1007/BF02066239

Ververis, C., Georghiou, K., Christodoulakis, N., Santas, P., Santas, R., 2004. Fiber dimensions, lignin and cellulose content of various plant materials and their suitability for paper production. Ind. Crops Prod. 19, 245–254. https://doi.org/10.1016/j.indcrop.2003.10.006

Villares, A., Moreau, C., Bennati-Granier, C., Garajova, S., Foucat, L., Falourd, X., Saake, B., Berrin, J.-G., Cathala, B., 2017. Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. Sci. Rep. 7, 40262. https://doi.org/10.1038/srep40262

Watanabe, H., Noda, H., Tokuda, G., Lo, N., 1998. A cellulase gene of termite origin. Nat. Lond. 394, 330–1. http://dx.doi.org/10.1038/28527

Watanabe, H., Tokuda, G., 2010. Cellulolytic Systems in Insects. Annu. Rev. Entomol. 55, 609–632. https://doi.org/10.1146/annurev-ento-112408-085319

Willis, J.D., Oppert, B., Oppert, C., Klingeman, W.E., Jurat-Fuentes, J.L., 2011. Identification, cloning, and expression of a GHF9 cellulase from Tribolium castaneum (Coleoptera: Tenebrionidae). J. Insect Physiol. 57, 300–306. https://doi.org/10.1016/j.jinsphys.2010.11.019

Willis, J.D., Oppert, C., Jurat-Fuentes, J.L., 2010. Methods for discovery and characterization of cellulolytic enzymes from insects. Insect Sci. 17, 184–198. https://doi.org/10.1111/j.1744-7917.2010.01322.x

Woodbury, N., Gries, G., 2013a. Firebrats, Thermobia domestica, aggregate in response to the microbes Enterobacter cloacae and Mycotypha microspora. Entomol. Exp. Appl. 147, 154–159. https://doi.org/10.1111/eea.12054

Woodbury, N., Gries, G., 2013b. How Firebrats (Thysanura: Lepismatidae) Detect and Nutritionally Benefit From Their Microbial Symbionts Enterobacter cloacae and Mycotypha microspora. Environ. Entomol. 42, 860–867. https://doi.org/10.1603/EN13104

Zhao, H., Baker, G.A., Song, Z., Olubajo, O., Crittle, T., Peters, D., 2008. Designing enzyme-compatible ionic liquids that can dissolve carbohydrates. Green Chem. 10, 696. https://doi.org/10.1039/b801489b

Zhao, H., Jones, C.L., Baker, G.A., Xia, S., Olubajo, O., Person, V.N., 2009. Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. J. Biotechnol. 139, 47–54. https://doi.org/10.1016/j.jbiotec.2008.08.009

Zhu, Z., Sathitsuksanoh, N., Vinzant, T., Schell, D.J., McMillan, J.D., Zhang, Y.-H.P., 2009. Comparative study of corn stover pretreated by dilute acid and cellulose solvent-based lignocellulose fractionation: Enzymatic hydrolysis, supramolecular structure, and substrate accessibility. Biotechnol. Bioeng. 103, 715–724. https://doi.org/10.1002/bit.22307

Zinkler, D., Götze, M., 1987. Cellulose digestion by the firebrat Thermobia domestica. Comp. Biochem. Physiol. Part B Comp. Biochem. 88, 661–666. https://doi.org/10.1016/0305-0491(87)90360-9

Appendix 3

Table 3.1. Endoglucanases in *Thermobia domestica:* Coding sequences in *T. domestica* genome encoding for endoglucanases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00057401-RA	endoglucanase E-4-like	468	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	4.22E-21	83.01887	90.1225	53	44
Th_d_00113252-RA	endoglucanase E-4-like	288	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	1.43E-25	67.5	99.3673	80	54
Th_d_00036852-RA	Endoglucanase E-4	612	gi|695189911|gb|AIT11911.1|endo-beta-1,4-glucanase [Parasesarma erythrodactyla]	Crustacea	AIT11911	1.32E-15	68.57143	80.4925	70	48
Th_d_00085723-RA	endoglucanase 15-like	390	gi|805808256|ref|XP_012146513.1|PREDICTED: endoglucanase 15-like [Megachile rotundata]gi|805808259|ref|XP_012146514.1|PREDICTED: endoglucanase 15-like [Megachile rotundata]	Hymenoptera	XP_012146513, XP_012146514	1.66E-22	90.56604	97.4413	53	48

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00090601-RA	endoglucanase E-4-like	573	gi|1000740642|ref|XP_015592927.1|PREDICTED: endoglucanase E-4-like [Cephus cinctus]gi|1000740644|ref|XP_015592928.1|PREDICTED: endoglucanase E-4-like [Cephus cinctus]gi|1000740646|ref|XP_015592929.1|PREDICTED: endoglucanase E-4-like [Cephus cinctus]	Hymenoptera	XP_015592927, XP_015592928, XP_015592929	8.06E-59	70.2381	196.823	168	118
Th_d_00118343-RA	endoglucanase 15-like	243	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	1.02E-23	78.68852	93.9745	61	48
Th_d_00004757-RA	endoglucanase E-4-like	522	gi|646689408|gb|KDR06579.1|Endoglucanase 1 [Zootermopsis nevadensis]	Blattodea	KDR06579	3.23E-44	78.81356	148.288	118	93
Th_d_00004759-RA	Endoglucanase E-4	381	gi|197691951|dbj|BAG70027.1|endo-beta-1,4-glucanase, partial [Salganea esakii]	Blattodea	BAG70027	5.19E-18	73.01587	83.5741	63	46

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00004756-RA	endoglucanase 15-like	735	gi|1080053219|ref|XP_018568194.1|uncharacterized protein LOC108908596 isoform X1 [Anoplophora glabripennis]gi|1080053221|ref|XP_018568195.1|uncharacterized protein LOC108908596 isoform X1 [Anoplophora glabripennis]gi|1325348753|ref|XP_023310346.1|uncharacterized protein LOC108908596 isoform X2 [Anoplophora glabripennis]	Coleoptera	XP_018568194, XP_018568195, XP_023310346	1.31E-19	86.79245	93.5893	53	46
Th_d_00029458-RA	Endoglucanase F	569	gi|646711640|gb|KDR16731.1|Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR16731	2.2E-47	76.2963	166.777	135	103
Th_d_00029459-RA	Endoglucanase E-4 precursor, putative	432	gi|646711640|gb|KDR16731.1|Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR16731	1.6E-31	90.47619	122.479	63	57

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00089623-RA	Endoglucanase E-4 precursor, putative	288	gi|646711640|gb|KDR16731.1|Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR16731	8.97E-33	91.80328	124.02	61	56
Th_d_00089451-RA	Endoglucanase E-4	258	gi|1061478177|gb|ODM95820.1|Endoglucanase E-4 [Orchesella cincta]	Collembola	ODM95820	1.71E-17	76.27119	81.2629	59	45
Th_d_00078629-RA	endoglucanase E-4-like	1143	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	6.72E-12	94.87179	72.0182	39	37

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00041503-RA	AChain A, The Structure Of Endoglucanase From Termite, Nasutitermes Takasagoensis, At Ph 2.5.	1160	gi|28373491|pdb|1KS8|AChain A, The Structure Of Endoglucanase From Termite, Nasutitermes Takasagoensis, At Ph 2.5.gi|28373492|pdb|1KSC|AChain A, The Structure Of Endoglucanase From Termite, Nasutitermes Takasagoensis, At Ph 5.6.gi|28373493|pdb|1KSD|AChain A, The Structure Of Endoglucanase From Termite, Nasutitermes Takasagoensis, At Ph 6.5.	Blattodea	1KS8_A, 1KSC_A, 1KSD_A	5.6E-143	73.5376	418.698	359	264

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00068682-RA	endoglucanase E-4-like	1007	gi|1339056265|ref|XP_023716596.1|uncharacterized protein LOC111869358 [Cryptotermes secundus]gi|1339056267|ref|XP_023716597.1|uncharacterized protein LOC111869358 [Cryptotermes secundus]gi|1330895261|gb|PNF24408.1|hypothetical protein B7P43_G09674 [Cryptotermes secundus]	Blattodea	XP_023716596, XP_023716597, PNF24408	1.9E-121	90.55794	362.844	233	211
Th_d_00000353-RA	Endoglucanase A	1836	gi|1339087696|ref|XP_023704929.1|uncharacterized protein LOC111863126 [Cryptotermes secundus]gi|1330920323|gb|PNF36365.1|Endoglucanase A [Cryptotermes secundus]	Blattodea	XP_023704929, PNF36365	3.77E-95	67.09677	304.679	310	208
Th_d_00000350-RA	endoglucanase 15-like	270	gi|769842984|ref|XP_011632703.1|PREDICTED: endoglucanase 15-like [Pogonomyrmex barbatus]gi|769842986|ref|XP_011632704.1|PREDICTED: endoglucanase 15-like [Pogonomyrmex barbatus]	Hymenoptera	XP_011632703, XP_011632704	4.89E-22	67.94872	94.3597	78	53

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00000354-RA	endoglucanase E-4-like	984	gi|985388348|ref|XP_015378728.1| PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	5.16E-21	70.37037	94.7449	81	57
Th_d_00121658-RA	Endoglucanase A	267	gi|1078570861|gb|AOV94255.1|cellulase [Antipaluria urichi]	Embioptera	AOV94255	6.75E-09	84.375	56.9954	32	27
Th_d_00107645-RA	Endoglucanase E-4 precursor, putative	369	gi|1325291025|ref|XP_023329807.1|uncharacterized protein LOC111702374 isoform X2 [Eurytemora affinis]	Crustacea	XP_023329807	1.34E-39	72.80702	135.191	114	83
Th_d_00096060-RA	endoglucanase E-4-like	402	gi|952540008|gb|KRT85487.1|hypothetical protein AMK59_2656 [Oryctes borbonicus]	Coleoptera	KRT85487	6.43E-26	71.54472	104.76	123	88
Th_d_00023147-RA	Endoglucanase E-4	849	gi|44885844|dbj|BAD12011.1|putative endo-beta-1,4-glucanase NtEG2, partial [Nasutitermes takasagoensis]	Blattodea	BAD12011	2.1E-36	72.35772	139.043	123	89

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00113210-RA	endoglucanase 15-like	225	gi|805808256|ref|XP_012146513.1|PREDICTED: endoglucanase 15-like [Megachile rotundata]gi|805808259|ref|XP_012146514.1|PREDICTED: endoglucanase 15-like [Megachile rotundata]	Hymenoptera	XP_012146513, XP_012146514	4.29E-22	88.46154	93.5893	52	46
Th_d_00045439-RA	Endoglucanase E-4	906	gi|1080053219|ref|XP_018568194.1|uncharacterized protein LOC108908596 isoform X1 [Anoplophora glabripennis]gi|1080053221|ref|XP_018568195.1|uncharacterized protein LOC108908596 isoform X1 [Anoplophora glabripennis]gi|1325348753|ref|XP_023310346.1|uncharacterized protein LOC108908596 isoform X2 [Anoplophora glabripennis]	Coleoptera	XP_018568194, XP_018568195, XP_023310346	4.02E-23	71.18644	105.145	118	84
Th_d_00034989-RA	endoglucanase E-4-like	1350	gi|1022761055|gb|KZS07093.1|Endoglucanase [Daphnia magna]	Crustacea	KZS07093	1.2E-119	68.36158	370.548	354	242

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00031500-RA	Endoglucanase A	1326	gi|1339087696|ref|XP_023704929.1|uncharacterized protein LOC111863126 [Cryptotermes secundus]gi|1330920323|gb|PNF36365.1|Endoglucanase A [Cryptotermes secundus]	Blattodea	XP_023704929, PNF36365	1.5E-126	63.70023	379.407	427	272
Th_d_00040844-RA	Endoglucanase A	1860	gi|1339087696|ref|XP_023704929.1|uncharacterized protein LOC111863126 [Cryptotermes secundus]gi|1330920323|gb|PNF36365.1|Endoglucanase A [Cryptotermes secundus]	Blattodea	XP_023704929, PNF36365	3.25E-07	61.01695	60.077	59	36
Th_d_00097963-RA	endoglucanase E-4-like	408	gi|1059424762|ref|XP_017785110.1|PREDICTED: endoglucanase E-4-like [Nicrophorus vespilloides]	Coleoptera	XP_017785110	4.64E-23	84.74576	98.9821	59	50
Th_d_00060602-RA	endoglucanase 15-like	771	gi|1228018665|ref|XP_021941322.1|endoglucanase 7-like, partial [Zootermopsis nevadensis]	Blattodea	XP_021941322	3.53E-64	73.68421	204.527	152	112

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00085360-RA	Endoglucanase E-4 precursor, putative	360	gi|197691959|dbj|BAG70031.1|endo-beta-1,4-glucanase, partial [Panesthia angustipennis]	Blattodea	BAG70031	5.4E-24	63.73626	99.3673	91	58
Th_d_00007226-RA	AF220596_1beta-1,4-endoglucanase 1	2020	gi|8886827|gb|AAF80584.1|AF220596_1beta-1,4-endoglucanase 1 [Panesthia cribrata]	Blattodea	AAF80584	1.5E-173	69.14894	508.449	470	325
Th_d_00062804-RA	endoglucanase 15-like	432	gi|1233161131|ref|XP_022200004.1|uncharacterized protein LOC111056901 [Nilaparvata lugens]gi|830997584|gb|AKL90411.1|endo-beta-1,4-glucanase [Nilaparvata lugens]	Hemiptera	XP_022200004, AKL90411	1.38E-21	90.38462	95.5153	52	47
Th_d_00044871-RA	endoglucanase E-4-like	864	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	4.06E-19	69.5122	88.5817	82	57

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00048316-RA	endoglucanase E-4-like	1194	gi|646711640|gb|KDR16731.1|Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR16731	4.8E-153	87.95987	446.047	299	263
Th_d_00081071-RA	putative endo-beta-1,4-glucanase HsEG2	849	gi|1228007791|ref|XP_021935807.1|uncharacterized protein LOC110837698, partial [Zootermopsis nevadensis]	Blattodea	XP_021935807	2.43E-51	73.28767	177.178	146	107
Th_d_00036851-RA	AF220594_1beta-1,4-glucanase 2	507	gi|7546878|gb|AAF63725.1|AF220594_1beta-1,4-glucanase 2, partial [Mastotermes darwiniensis]	Blattodea	AAF63725	1.33E-34	72.72727	127.872	121	88
Th_d_00091320-RA	endo-beta-1,4-glucanase 1	441	gi|1228018669|ref|XP_021941325.1|LOW QUALITY PROTEIN: uncharacterized protein LOC110840537 [Zootermopsis nevadensis]	Blattodea	XP_021941325	3.31E-29	65.83333	115.161	120	79

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00079350-RA	AF220583_1beta-1,4-glucanase 3	450	gi|7546856|gb|AAF63714.1|AF220583_1beta-1,4-glucanase 3, partial [Polyphaga aegyptiaca]	Blattodea	AAF63714	1.83E-40	74.79675	142.124	123	92
Th_d_00010317-RA	AF220588_1beta-1,4-glucanase 3	309	gi|197691951|dbj|BAG70027.1|endo-beta-1,4-glucanase, partial [Salganea esakii]	Blattodea	BAG70027	4.63E-29	72.63158	112.079	95	69
Th_d_00000351-RA	AF220588_1beta-1,4-glucanase 3	1464	gi|197691947|dbj|BAG70025.1|endo-beta-1,4-glucanase, partial [Salganea esakii]	Blattodea	BAG70025	1.55E-59	79.42857	204.527	175	139
Th_d_00089290-RA	endo-beta-1,4-glucanase	309	gi|375151369|dbj|BAL60587.1|cellulase [Neomysis intermedia]	Crustacea	BAL60587	2.93E-24	80.64516	100.908	62	50

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00075942-RA	1,4-alpha-glucan-branching enzyme	225	gi|1339092712|ref|XP_023707569.1|1,4-alpha-glucan-branching enzyme [Cryptotermes secundus]gi|1330914862|gb|PNF33674.1|1,4-alpha-glucan-branching enzyme [Cryptotermes secundus]	Blattodea	XP_023707569, PNF33674	9.67E-25	91.22807	101.679	57	52
Th_d_00049037-RA	endo-beta-1,4-glucanase	997	gi|13537532|dbj|BAB40693.1|endo-b-1,4-glucanase [Coptotermes formosanus]gi|13537534|dbj|BAB40694.1|endo-b-1,4-glucanase [Coptotermes formosanus]	Blattodea	BAB40693, BAB40694	6.7E-114	78.87324	343.199	284	224
Th_d_00013433-RA	1,4-alpha-glucan-branching enzyme	618	gi|795023516|ref|XP_011860965.1|PREDICTED: 1,4-alpha-glucan-branching enzyme [Vollenhovia emeryi]	Hymenoptera	XP_011860965	8.83E-51	56.37255	180.259	204	115
Th_d_00035324-RA	endo-b-1,4-glucanase	1332	gi|13095576|gb|AAK12339.1|cellulase [Coptotermes acinaciformis]	Blattodea	AAK12339	1.3E-164	73.25843	476.478	445	326

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00104889-RA	endo-beta-1,4-glucanase 1	474	gi|7546856|gb|AAF63714.1|AF220583_1beta-1,4-glucanase 3, partial [Polyphaga aegyptiaca]	Coleoptera	AAF63714	2.06E-44	82.30088	152.525	113	93
Th_d_00113063-RA	glycoside hydrolase family 9	228	gi|521313293|gb|AGP76419.1|endo-beta-1,4-glucanase 5 [Odontotermes formosanus]	Blattodea	AGP76419	1.53E-17	68.75	80.8777	64	44
Th_d_00113037-RA	glycoside hydrolase family 9	216	gi|992051711|gb|AMH40365.1|glycoside hydrolase family 9 [Extatosoma tiaratum]	Phamatodea	AMH40365	1.35E-15	80.76923	75.0998	52	42
Th_d_00071154-RA	glycoside hydrolase family 9	822	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	7.71E-21	85.96491	95.5153	57	49

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00035331-RA	glycoside hydrolase family 9	1546	gi|24940551|emb|CAD54729.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	0	79.53488	561.222	430	342
Th_d_00115973-RA	glycoside hydrolase family 9	279	gi|992051866|gb|AMH40374.1|glycoside hydrolase family 9 [Peruphasma schultei]	Phamatodea	AMH40374	8.96E-17	68.11594	79.7221	69	47
Th_d_00029874-RA	glycoside hydrolase family 9	1827	gi|1330724231|gb|PNE09439.1|hypothetical protein B7P43_G00116 [Cryptotermes secundus]	Blattodea	PNE09439	2.47E-67	57.24382	227.254	283	162
Th_d_00010318-RA	glycoside hydrolase family 9	1538	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	0	73.67206	533.102	433	319

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00010316-RA	glycoside hydrolase family 9	375	gi|646689408|gb|KDR06579.1| Endoglucanase 1 [Zootermopsis nevadensis]	Blattodea	KDR06579	8.71E-43	68.64407	142.51	118	81
Th_d_00007347-RA	glycoside hydrolase family 9	1482	gi|24940551|emb|CAD54729.1| beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	0	78.4897	563.148	437	343
Th_d_00007348-RA	glycoside hydrolase family 9	1555	gi|24940553|emb|CAD54730.1| beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54730	0	79.52941	556.214	425	338
Th_d_00000352-RA	glycoside hydrolase family 9	465	gi|13095576|gb|AAK12339.1|cellulase [Coptotermes acinaciformis]	Blattodea	AAK12339	6.41E-47	77.3913	163.696	115	89

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00093510-RA	glycoside hydrolase family 9	282	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	1.69E-39	75.4717	137.887	106	80
Th_d_00018984-RA	glycoside hydrolase family 9	1937	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	1.08E-92	80.82192	293.893	219	177
Th_d_00089430-RA	glycoside hydrolase family 9	495	gi|646711640|gb|KDR16731.1|Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR16731	1.39E-26	84.41558	110.153	77	65

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00135584-RA	glycoside hydrolase family protein 5	144	gi|315570658|gb|ADU33333.1|glycoside hydrolase family protein 5 [Gastrophysa viridula]	Coleoptera	ADU33333	0.000749	71.42857	40.817	35	25
Th_d_00015659-RA	glycoside hydrolase family 9	2150	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	9.34E-92	80.66038	298.901	212	171
Th_d_00079172-RA	glycoside hydrolase family 9	606	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	1.17E-90	82	272.707	200	164
Th_d_00090714-RA	glycoside hydrolase family 9	360	gi|992052122|gb|AMH40384.1|glycoside hydrolase family 9 [Sipyloidea sipylus]	Phamatodea	AMH40384	2.15E-22	79.66102	96.6709	59	47
Th_d_00078933-RA	glycoside hydrolase family 9	723	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	2.13E-21	84.48276	95.9005	58	49
Th_d_00074140-RA	glycoside hydrolase family 9	657	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phamatodea	AMH40360	6.36E-70	68.51064	226.098	235	161

Table 3.1. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00100655-RA	glycoside hydrolase family 9	498	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phamatodea	AMH40360	3.79E-32	65	124.79	120	78
Th_d_00096112-RA	glycoside hydrolase family 9	612	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phamatodea	AMH40360	4.82E-50	76.66667	173.711	120	92
Th_d_00019623-RA	glycoside hydrolase family 9	1554	gi|24940551|emb|CAD54729.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	0	79.76744	561.222	430	343
Th_d_00035252-RA	glycoside hydrolase family 9	1463	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	2.4E-180	72.32558	518.464	430	311
Th_d_00111221-RA	glycoside hydrolase family 9	567	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	6.4E-37	76.14679	135.191	109	83
Th_d_00061520-RA	glycoside hydrolase family 9	351	gi|646689408|gb|KDR06579.1|Endoglucanase 1 [Zootermopsis nevadensis]	Blattodea	KDR06579	1.89E-42	71.05263	141.354	114	81
Th_d_00005705-RA	glycoside hydrolase family 9	1593	gi|24940551|emb|CAD54729.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	0	79.76744	559.681	430	343

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00094204-RA	glycoside hydrolase family 9	399	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phamatodea	AMH40360	3.17E-51	77.11864	173.711	118	91
Th_d_00104189-RA	glycoside hydrolase family 9	514	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phamatodea	AMH40360	1.6E-47	82.35294	166.007	102	84
Th_d_00046715-RA	glycoside hydrolase family 9	1068	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	6.6E-144	76.06061	420.239	330	251
Th_d_00028499-RA	glycoside hydrolase family 9	1779	gi|24940551|emb|CAD54729.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	0	78.42697	566.614	445	349
Th_d_00038486-RA	glycoside hydrolase family 9	1307	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phamatodea	AMH40360	6E-173	78.63014	497.664	365	287
Th_d_00091520-RA	glycoside hydrolase family 9	471	gi|8886829|gb|AAF80585.1|AF220597_1beta-1,4-endoglucanase 2 [Panesthia cribrata]	Blattodea	AAF80585	1.22E-21	56.05096	95.9005	157	88
Th_d_00034779-RA	glycoside hydrolase family 9	1436	gi|24940551|emb|CAD54729.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	0	79.81221	556.599	426	340

Table 3.1. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00014865-RA	glycoside hydrolase family 9	1574	gi|13095576|gb|AAK12339.1|cellulase [Coptotermes acinaciformis]	Blattodea	AAK12339	0	72.57019	545.428	463	336
Th_d_00075629-RA	glycoside hydrolase family 9	606	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	1.17E-90	82	272.707	200	164
Th_d_00053691-RA	glycoside hydrolase family 9	1006	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phamatodea	AMH40360	1.1E-127	78.94737	378.252	285	225
Th_d_00130530-RA	PREDICTED: uncharacterized protein LOC107165552	321	gi|985403057|ref|XP_015369326.1|PREDICTED: uncharacterized protein LOC107165552 [Diuraphis noxia]	Hemiptera	XP_015369326	6.8E-09	56.52174	57.7658	92	52

Table 3.2. Glucosidases in *Thermobia domestica:* Coding sequences in *T. domestica* genome encoding for glucosidases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00054985-RA	uncharacterized family 31 glucosidase KIAA1161-like	2754	gi|1228365487|ref|XP_021943634.1|uncharacterized family 31 glucosidase KIAA1161-like [Folsomia candida]gi|1228365489|ref|XP_021943636.1|uncharacterized family 31 glucosidase KIAA1161-like [Folsomia candida]	Collembola	XP_021943634, XP_021943636	4.7E-107	58.159	350.903	478	278
Th_d_00105563-RA	uncharacterized family 31 glucosidase KIAA1161-like	288	gi|1101351520|ref|XP_018901738.1|PREDICTED: uncharacterized family 31 glucosidase KIAA1161-like [Bemisia tabaci]	Hemiptera	XP_018901738	2.21E-25	87.30159	104.375	63	55
Th_d_00032903-RA	uncharacterized family 31 glucosidase KIAA1161-like	2130	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	1.4E-139	62.47379	428.328	477	298

Table 3.2. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00062466-RA	glucosidase 2 subunit beta	633	gi|1339077375|ref|XP_023726071.1|glucosidase 2 subunit beta isoform X1 [Cryptotermes secundus]	Blattodea	XP_023726071	1.55E-21	74.64789	93.2041	71	53
Th_d_00050423-RA	uncharacterized family 31 glucosidase KIAA1161-like	2034	gi|577754846|gb|AHH86052.1|glycoside hydrolase family 31 [Callosobruchus maculatus]	Coleoptera	AHH86052	3E-159	58.69565	478.404	598	351
Th_d_00142013-RA	uncharacterized family 31 glucosidase KIAA1161-like	1404	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	1.4E-145	62.76596	434.876	470	295

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00061435-RA	uncharacterized family 31 glucosidase KIAA1161-like	1950	gi|1199397963|ref|XP_021194260.1|uncharacterized family 31 glucosidase KIAA1161-like [Helicoverpa armigera]	Lepidoptera	XP_021194260	2.5E-169	56.64557	503.827	632	358
Th_d_00114590-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	246	gi|1330878954|gb|PNF16838.1|Myrosinase 1, partial [Cryptotermes secundus]	Blattodea	PNF16838	1.45E-36	86.41975	133.265	81	70

Table 3.2. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00025773-RA	glucosidase 2 subunit beta	453	gi|1228013272|ref|XP_021938600.1|glucosidase 2 subunit beta [Zootermopsis nevadensis]	Blattodea	XP_021938600	1.37E-24	75.30864	104.375	81	61
Th_d_00025772-RA	glucosidase 2 subunit beta	243	gi|1228013272|ref|XP_021938600.1|glucosidase 2 subunit beta [Zootermopsis nevadensis]	Blattodea	XP_021938600	4.17E-17	76	80.1073	50	38
Th_d_00149067-RA	uncharacterized family 31 glucosidase KIAA1161-like	468	gi|1101351520|ref|XP_018901738.1|PREDICTED: uncharacterized family 31 glucosidase KIAA1161-like [Bemisia tabaci]	Hemiptera	XP_018901738	5.7E-24	86.15385	103.219	65	56
Th_d_00098342-RA	uncharacterized family 31 glucosidase KIAA1161-like	396	gi|646703149|gb|KDR11965.1|hypothetical protein L798_13618, partial [Zootermopsis nevadensis]	Blattodea	KDR11965	1.63E-49	81.41593	170.629	113	92

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00018632-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	1317	gi|303324839|pdb|3AHZ|AChain A, Crystal Structure Of Beta-Glucosidase From Termite Neotermes Koshunensis In Complex With Trisgi|393715252|pdb|3VIF|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Gluconolactonegi|393715253|pdb|3VIG|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With 1-deoxynojirimycingi|393715254|pdb|3VIH|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glycerolgi|393715255|pdb|3VII|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Bis-tris	Blattodea	3AHZ_A, 3VIF_A, 3VIG_A, 3VIH_A, 3VII_A	4.1E-180	75.23148	517.309	432	325

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00057277-RA	uncharacterized family 31 glucosidase KIAA1161-like	426	gi|646703149|gb|KDR11965.1|hypothetical protein L798_13618, partial [Zootermopsis nevadensis]	Blattodea	KDR11965	4.78E-49	80.70175	169.859	114	92
Th_d_00069820-RA	lysosomal alpha-glucosidase-like	1173	gi|1032766152|ref|XP_016841489.1|PREDICTED: lysosomal alpha-glucosidase-like [Nasonia vitripennis]gi|1032766154|ref|XP_016841490.1|PREDICTED: lysosomal alpha-glucosidase-like [Nasonia vitripennis]	Hymenoptera	XP_016841489, XP_016841490	6.08E-08	55.30303	61.6178	132	73
Th_d_00081027-RA	beta-glucosidase	471	gi|1227984161|ref|XP_021923730.1|myrosinase 1-like isoform X1 [Zootermopsis nevadensis]gi|646712919|gb|KDR17465.1|Lactase-phlorizin hydrolase [Zootermopsis nevadensis]	Blattodea	XP_021923730, KDR17465	2.56E-48	63.15789	169.474	152	96
Th_d_00106209-RA	glucosidase 2 subunit beta	633	gi|1339077375|ref|XP_023726071.1|glucosidase 2 subunit beta isoform X1 [Cryptotermes secundus]	Blattodea	XP_023726071	1.74E-21	57.84314	93.2041	102	59

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00099763-RA	uncharacterized family 31 glucosidase KIAA1161	1347	gi|1228003695|ref|XP_021933733.1|uncharacterized family 31 glucosidase KIAA1161-like isoform X1 [Zootermopsis nevadensis]gi|646703150|gb|KDR11966.1|putative family 31 glucosidase [Zootermopsis nevadensis]	Blattodea	XP_021933733, KDR11966	1.7E-121	54.18251	374.785	526	285
Th_d_00000860-RA	neutral alpha-glucosidase AB	414	gi|1339041694|ref|XP_023710162.1|neutral alpha-glucosidase AB [Cryptotermes secundus]gi|1330933432|gb|PNF42738.1|Neutral alpha-glucosidase AB [Cryptotermes secundus]	Blattodea	XP_023710162, PNF42738	3.75E-55	82.44275	191.045	131	108
Th_d_00000862-RA	Neutral alpha-glucosidase AB	435	gi|1227991790|ref|XP_021927674.1|neutral alpha-glucosidase AB isoform X2 [Zootermopsis nevadensis]	Blattodea	XP_021927674	1.85E-46	85.59322	166.777	118	101
Th_d_00000861-RA	Neutral alpha-glucosidase AB	315	gi|1227991792|ref|XP_021927675.1|neutral alpha-glucosidase AB isoform X3 [Zootermopsis nevadensis]gi|646709565|gb|KDR15365.1|Neutral alpha-glucosidase AB [Zootermopsis nevadensis]	Blattodea	XP_021927675, KDR15365	7.97E-23	94.11765	97.4413	51	48

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00000863-RA	Neutral alpha-glucosidase AB	519	gi|930677078|gb|KPJ17206.1|Neutral alpha-glucosidase AB [Papilio machaon]	Lepidoptera	KPJ17206	1.18E-53	80.53097	173.711	113	91
Th_d_00000865-RA	neutral alpha-glucosidase AB	1599	gi|1339041694|ref|XP_023710162.1|neutral alpha-glucosidase AB [Cryptotermes secundus]gi|1330933432|gb|PNF42738.1|Neutral alpha-glucosidase AB [Cryptotermes secundus]	Blattodea	XP_023710162, PNF42738	1.4E-103	70.22059	337.421	272	191
Th_d_00000864-RA	Neutral alpha-glucosidase AB	633	gi|242019253|ref|XP_002430076.1|Neutral alpha-glucosidase AB precursor, putative [Pediculus humanus corporis]gi|212515157|gb|EEB17338.1|Neutral alpha-glucosidase AB precursor, putative [Pediculus humanus corporis]	Phthiraptera	XP_002430076, EEB17338	2.81E-42	74.25743	157.147	101	75
Th_d_00000866-RA	Neutral alpha-glucosidase AB	780	gi|1108484091|emb|CRK87777.1|CLUMA_CG001536, isoform A [Clunio marinus]	Diptera	CRK87777	2.15E-18	75.36232	90.8929	69	52

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00017412-RA	mannosyl-oligosaccharide glucosidase isoform X2	6713	gi|1228005382|ref|XP_021934582.1|mannosyl-oligosaccharide glucosidase isoform X2 [Zootermopsis nevadensis]gi|1228005384|ref|XP_021934583.1|mannosyl-oligosaccharide glucosidase isoform X2 [Zootermopsis nevadensis]gi|1228005386|ref|XP_021934584.1|mannosyl-oligosaccharide glucosidase isoform X2 [Zootermopsis nevadensis]gi|1228005388|ref|XP_021934585.1|mannosyl-oligosaccharide glucosidase isoform X2 [Zootermopsis nevadensis]gi|1228005390|ref|XP_021934586.1|mannosyl-oligosaccharide glucosidase isoform X2 [Zootermopsis nevadensis]gi|1228005392|ref|XP_021934587.1|mannosyl-oligosaccharide glucosidase isoform X2 [Zootermopsis nevadensis]gi|1228005394|ref|XP_021934588.1|mannosyl-oligosaccharide glucosidase isoform X2 [Zootermopsis nevadensis]gi|646701956|gb|KDR11423.1|Mannosyl-oligosaccharide glucosidase [Zootermopsis nevadensis]	Blattodea	XP_021934582, XP_021934583, XP_021934584, XP_021934585, XP_021934586, XP_021934587, XP_021934588, KDR11423	0	73.50917	1048.88	872	641

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00050585-RA	lysosomal alpha-glucosidase-like	954	gi|1227993488|ref|XP_021928525.1|lysosomal alpha-glucosidase-like [Zootermopsis nevadensis]gi|646708767|gb|KDR14932.1|Lysosomal alpha-glucosidase [Zootermopsis nevadensis]	Blattodea	XP_021928525, KDR14932	3.6E-145	86.31579	440.654	285	246
Th_d_00079942-RA	uncharacterized family 31 glucosidase KIAA1161-like	2136	gi|1228365487|ref|XP_021943634.1|uncharacterized family 31 glucosidase KIAA1161-like [Folsomia candida]gi|1228365489|ref|XP_021943636.1|uncharacterized family 31 glucosidase KIAA1161-like [Folsomia candida]	Collembola	XP_021943634, XP_021943636	6.9E-139	57.80347	427.943	519	300
Th_d_00081056-RA	lysosomal alpha-glucosidase	519	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	3.49E-12	44.44444	70.0922	126	56

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00041545-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	487	gi|1061476505|gb|ODM94351.1|Lactase-phlorizin hydrolase [Orchesella cincta]	Collembola	ODM94351	3.44E-53	79.83871	181.415	124	99
Th_d_00055895-RA	uncharacterized family 31 glucosidase KIAA1161-like	261	gi|646703149|gb|KDR11965.1|hypothetical protein L798_13618, partial [Zootermopsis nevadensis]	Blattodea	KDR11965	6.34E-19	78.33333	85.5001	60	47
Th_d_00055894-RA	uncharacterized family 31 glucosidase KIAA1161-like	390	gi|1228003699|ref|XP_021933735.1|uncharacterized family 31 glucosidase KIAA1161-like [Zootermopsis nevadensis]	Blattodea	XP_021933735	1.06E-22	69.41176	98.2117	85	59
Th_d_00014523-RA	beta-glucosidase	423	gi|1090651209|gb|AOY34571.1|beta-glucosidase [Coptotermes formosanus]	Blattodea	AOY34571	1.22E-35	77.35849	134.035	106	82

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00011270-RA	uncharacterized family 31 glucosidase KIAA1161-like	2460	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	5E-166	72.50608	499.975	411	298
Th_d_00011271-RA	uncharacterized family 31 glucosidase KIAA1161-like	309	gi|1228003699|ref|XP_021933735.1|uncharacterized family 31 glucosidase KIAA1161-like [Zootermopsis nevadensis]	Blattodea	XP_021933735	6.06E-18	81.13208	83.5741	53	43
Th_d_00011273-RA	uncharacterized family 31 glucosidase KIAA1161-like	366	gi|1228003699|ref|XP_021933735.1|uncharacterized family 31 glucosidase KIAA1161-like [Zootermopsis nevadensis]	Blattodea	XP_021933735	3.35E-21	80	93.5893	60	48
Th_d_00154707-RA	uncharacterized family 31 glucosidase KIAA1161-like	1431	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	4.3E-161	71.6707	474.937	413	296
Th_d_00105015-RA	uncharacterized family 31 glucosidase KIAA1161-like	357	gi|1339040796|ref|XP_023708840.1|myogenesis-regulating glycosidase [Cryptotermes secundus]gi|1330911883|gb|PNF32422.1|hypothetical protein B7P43_G04891 [Cryptotermes secundus]	Blattodea	XP_023708840, PNF32422	7.38E-49	84.0708	171.014	113	95

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00141763-RA	uncharacterized family 31 glucosidase KIAA1161-like	1434	gi|636630780|gb|AIA09350.1| alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	1.6E-145	64.22414	435.261	464	298
Th_d_00099575-RA	beta-glucosidase	417	gi|1152526136|gb|AQW43010.1|beta-glucosidase [Nilaparvata lugens]	Hemiptera	AQW43010	2.22E-46	75.80645	154.836	124	94
Th_d_00032476-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	270	gi|393715256|pdb|3VIJ|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glucose	Blattodea	3VIJ_A	7.05E-38	86.74699	137.502	83	72
Th_d_00035285-RA	Lysosomal alpha-glucosidase	1962	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	1.28E-51	56.63717	199.519	226	128

Table 3.2. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_000115658-RA	Lysosomal alpha-glucosidase	441	gi|1022742921|gb|KZR99685.1|Lysosomal alpha-glucosidase, partial [Daphnia magna]	Crustacea	KZR99685	6.63E-63	74.14966	204.142	147	109
Th_d_000049005-RA	lysosomal alpha-glucosidase-like	1122	gi|1061468644|gb|ODM88819.1|Maltase-glucoamylase, intestinal [Orchesella cincta]	Collembola	ODM88819	4.51E-13	64.89362	77.411	94	61
Th_d_000013779-RA	lysosomal alpha-glucosidase-like	324	gi|1009533976|ref|XP_015903978.1|maltase-glucoamylase, intestinal [Parasteatoda tepidariorum]	Arachnida	XP_015903978	8.34E-29	78.48101	114.775	79	62
Th_d_000089527-RA	glucosidase 2 subunit beta	420	gi|1228013272|ref|XP_021938600.1|glucosidase 2 subunit beta [Zootermopsis nevadensis]	Blattodea	XP_021938600	5.08E-65	84.78261	211.846	138	117

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00025279-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	1536	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.63E-13	73.33333	77.411	60	44
Th_d_00048538-RA	Uncharacterized family 31 glucosidase KIAA1161	2302	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	6.9E-176	69.02287	523.472	481	332

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00159760-RA	lysosomal alpha-glucosidase-like	903	gi|1330905640|gb|PNF29605.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905643|gb|PNF29608.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905644|gb|PNF29609.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]	Blattodea	PNF29605, PNF29608, PNF29609	9.96E-54	59.93266	187.193	297	178
Th_d_00127607-RA	lysosomal alpha-glucosidase-like	903	gi|1330905640|gb|PNF29605.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905643|gb|PNF29608.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905644|gb|PNF29609.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]	Blattodea	PNF29605, PNF29608, PNF29609	1.02E-52	59.59596	184.882	297	177

Table 3.2. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00018760-RA	glucoside xylosyltransferase 1	1251	gi|1227970630|ref|XP_021916719.1|glucoside xylosyltransferase 2 isoform X2 [Zootermopsis nevadensis]	Blattodea	XP_021916719	1.5E-159	87.21805	460.685	266	232
Th_d_00076537-RA	uncharacterized family 31 glucosidase KIAA1161	1368	gi|1228003695|ref|XP_021933733.1|uncharacterized family 31 glucosidase KIAA1161-like isoform X1 [Zootermopsis nevadensis]gi|646703150|gb|KDR11966.1|putative family 31 glucosidase [Zootermopsis nevadensis]	Blattodea	XP_021933733, KDR11966	1.2E-121	54.18251	375.555	526	285

Table 3.3. β-1,3-glucanases in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for β-1,3-glucanases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00107538-RA	beta-1,3-glucan-binding protein-like	351	gi|1330905524|gb|PNF29497.1|hypothetical protein B7P43_G04546 [Cryptotermes secundus]	Blattodea	PNF29497	4.38E-35	67.54386	129.798	114	77
Th_d_00117669-RA	beta-1,3-glucan-binding protein-like	450	gi|913297561|ref|XP_013188774.1|PREDICTED: beta-1,3-glucan-binding protein-like [Amyelois transitella]	Lepidoptera	XP_013188774	3.33E-30	57.93103	117.857	145	84
Th_d_00017253-RA	beta-1,3-glucan-binding protein precursor	4398	gi|646713430|gb|KDR17776.1|Apolipophorin [Zootermopsis nevadensis]	Blattodea	KDR17776	0	63.85705	1267.68	1483	947
Th_d_00007520-RA	Beta-1,3-glucan-binding protein	258	gi|506968285|gb|AGM32486.1|gram negative bacteria binding protein 2, partial [Coptotermes formosanus]	Blattodea	AGM32486	3.3E-10	70.37037	60.4622	54	38

Table 3.3. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00007517-RA	beta-1,3-glucan-binding protein-like	923	gi|156535746|gb|ABU80005.1|GNBP-B1 [Anopheles merus]	Diptera	ABU80005	5.25E-08	88.57143	60.4622	35	31
Th_d_00103657-RA	beta-1,3-glucan-binding protein-like	419	gi|913297561|ref|XP_013188774.1|PREDICTED: beta-1,3-glucan-binding protein-like [Amyelois transitella]	Lepidoptera	XP_013188774	2.46E-29	60.15038	115.161	133	80
Th_d_00059385-RA	Beta-1,3-glucan-binding protein	696	gi|1227967090|ref|XP_021914900.1|beta-1,3-glucan-binding protein-like [Zootermopsis nevadensis]	Blattodea	XP_021914900	3.55E-45	54.22886	159.844	201	109
Th_d_00012339-RA	Beta-1,3-glucan-binding protein	792	gi|1022754729|gb|KZS02483.1|Beta-1,3-glucan-binding protein [Daphnia magna]	Crustacea	KZS02483	3.41E-50	72.85714	174.096	140	102
Th_d_00117568-RA	Beta-1,3-glucan-binding protein	285	gi|646720035|gb|KDR21902.1|Beta-1,3-glucan-binding protein 1 [Zootermopsis nevadensis]	Blattodea	KDR21902	4.49E-22	73.9726	92.8189	73	54

Table 3.4. Mannanases in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for mannanases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00048327-RA	endo-beta-1,4-mannanase	1087	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	3.9E-116	68	347.436	325	221
Th_d_00037769-RA	Endo-beta-1,4-mannanase	1229	gi|1022765557|gb|KZS10752.1|Endo-beta-1,4-mannanase [Daphnia magna]	Crustacea	KZS10752	6.9E-129	67.41573	381.719	356	240
Th_d_00117927-RA	endo-beta-1,4-mannanase	552	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	2.26E-45	57.22543	159.073	173	99
Th_d_00114892-RA	endo-beta-1,4-mannanase	540	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	7.1E-45	57.80347	157.532	173	100
Th_d_00022705-RA	Endo-beta-1,4-mannanase	614	gi|1022765557|gb|KZS10752.1|Endo-beta-1,4-mannanase [Daphnia magna]	Crustacea	KZS10752	3.03E-47	68.42105	164.851	152	104
Th_d_00022704-RA	endo-beta-1,4-mannanase	838	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	3.74E-75	71.5	239.58	200	143
Th_d_00040234-RA	lysosomal alpha-mannosidase	296	gi|1101377119|ref|XP_018908486.1|PREDICTED: lysosomal alpha-mannosidase [Bemisia tabaci]	Hemiptera	XP_018908486	1.88E-07	52.5	53.5286	80	42

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00119102-RA	beta-mannosidase isoform X1	447	gi|1192752750|ref|XP_015929642.2|beta-mannosidase [Parasteatoda tepidariorum]	Arachnida	XP_015929642	1.01E-61	84.73282	209.534	131	111
Th_d_00080655-RA	alpha-mannosidase 2	552	gi|1022777739|gb|KZS21240.1|Alpha-mannosidase 2x [Daphnia magna]	Crustacea	KZS21240	1.15E-20	90	94.7449	50	45
Th_d_00105670-RA	lysosomal alpha-mannosidase isoform X1	264	gi|195377741|ref|XP_002047646.1|uncharacterized protein Dvir_GJ11812 [Drosophila virilis]gi|194154804|gb|EDW69988.1|uncharacterized protein Dvir_GJ11812 [Drosophila virilis]	Diptera	XP_002047646, EDW69988	8.58E-16	97.05882	76.6406	34	33
Th_d_00080493-RA	lysosomal alpha-mannosidase-like	1059	gi|642940243|ref|XP_008199468.1|PREDICTED: RNA-directed DNA polymerase from mobile element jockey-like, partial [Tribolium castaneum]	Coleoptera	XP_008199468	6.87E-22	55.42169	100.523	166	92

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00121568-RA	ER degradation-enhancing alpha-mannosidase-like protein 3	402	gi|1227972165|ref|XP_021917507.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X1 [Zootermopsis nevadensis]gi|1227972167|ref|XP_021917508.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X1 [Zootermopsis nevadensis]gi|646718115|gb|KDR20717.1|ER degradation-enhancing alpha-mannosidase-like 3 [Zootermopsis nevadensis]	Blattodea	XP_021917507, XP_021917508, KDR20717	1.82E-54	90.90909	189.119	110	100
Th_d_00003139-RA	beta-mannosidase	1500	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	5E-163	65.13026	490.345	499	325

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00003143-RA	Beta-mannosidase	345	gi|1061485315|gb|ODN02454.1|Beta-mannosidase [Orchesella cincta]	Collembola	ODN02454	2.93E-31	76.53061	122.094	98	75
Th_d_00003142-RA	beta-mannosidase isoform X1	273	gi|766933223|ref|XP_011498870.1| PREDICTED: beta-mannosidase [Ceratosolen solmsi marchali]	Hymenoptera	XP_011498870	4.33E-23	75	97.4413	68	51
Th_d_00003141-RA	beta-mannosidase isoform X1	1077	gi|1192752750|ref|XP_015929642.2|beta-mannosidase [Parasteatoda tepidariorum]	Arachnida	XP_015929642	1.58E-09	72.54902	66.2402	51	37
Th_d_00003140-RA	beta-mannosidase, putative	387	gi|1192752750|ref|XP_015929642.2|beta-mannosidase [Parasteatoda tepidariorum]	Arachnida	XP_015929642	3.06E-28	88.0597	114.005	67	59
Th_d_00036630-RA	Mannan endo-1,4-beta-mannosidase	1341	gi|1061473607|gb|ODM91993.1|Mannan endo-1,4-beta-mannosidase [Orchesella cincta]	Collembola	ODM91993	3.97E-77	57.05128	251.136	312	178

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00032183-RA	Mannosyl-oligosaccharide alpha-1,2-mannosidase isoform, putative	843	gi|242010114|ref|XP_002425821.1|Mannosyl-oligosaccharide alpha-1,2-mannosidase isoform, putative [Pediculus humanus corporis]gi|212509754|gb|EEB13083.1|Mannosyl-oligosaccharide alpha-1,2-mannosidase isoform, putative [Pediculus humanus corporis]	Phthiraptera	XP_002425821, EEB13083	4.88E-73	91.85185	238.81	135	124
Th_d_00032182-RA	mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like	501	gi|1121144990|ref|XP_019559624.1|PREDICTED: mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like [Aedes albopictus]gi|1121144992|ref|XP_019559625.1|PREDICTED: mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like [Aedes albopictus]gi|1121144994|ref|XP_019559626.1|PREDICTED: mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like [Aedes albopictus]gi|1121144996|ref|XP_019559627.1|PREDICTED: mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like [Aedes albopictus]	Diptera	XP_019559624, XP_019559625, XP_019559626, XP_019559627	7.38E-36	90.41096	137.117	73	66

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00095658-RA	alpha-mannosidase 2	600	gi|1330899240|gb|PNF26420.1|hypothetical protein B7P43_G16606, partial [Cryptotermes secundus]	Blattodea	PNF26420	1.23E-59	69.23077	201.83	195	135
Th_d_00000733-RA	lysosomal alpha-mannosidase isoform X1	3979	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	0	79.41176	823.157	578	459
Th_d_00063258-RA	lysosomal alpha-mannosidase-like	615	gi|157111150|ref|XP_001651410.1|lysosomal alpha-mannosidase [Aedes aegypti]gi|108878513|gb|EAT42738.1|AAEL005749-PA [Aedes aegypti]	Diptera	XP_001651410, EAT42738	1.1E-08	70.21277	60.8474	47	33
Th_d_00031859-RA	Endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase	399	gi|1339074955|ref|XP_023724784.1|endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase isoform X2 [Cryptotermes secundus]	Blattodea	XP_023724784	1.95E-46	78.86179	163.31	123	97

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00031860-RA	endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase	777	gi|1218211818|ref|XP_021712668.1|endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase-like [Aedes aegypti]gi|1218211820|ref|XP_021712669.1|endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase-like [Aedes aegypti]gi|1218211822|ref|XP_021712671.1|endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase-like [Aedes aegypti]	Diptera	XP_021712668, XP_021712669, XP_021712671	1.7E-99	78.89908	301.982	218	172
Th_d_00002247-RA	lysosomal alpha-mannosidase isoform X1	591	gi|929380069|ref|XP_014100402.1|PREDICTED: lysosomal alpha-mannosidase-like, partial [Bactrocera oleae]	Diptera	XP_014100402	3.18E-98	80.72917	289.271	192	155
Th_d_00002245-RA	lysosomal alpha-mannosidase isoform X2	528	gi|557780736|ref|XP_005189989.1|PREDICTED: lysosomal alpha-mannosidase isoform X2 [Musca domestica]	Diptera	XP_005189989	4.11E-56	70.58824	196.052	170	120
Th_d_00002244-RA	lysosomal alpha-mannosidase isoform X1	492	gi|1228005950|ref|XP_021934868.1|lysosomal alpha-mannosidase isoform X1 [Zootermopsis nevadensis]	Blattodea	XP_021934868	7.1E-41	70.71429	151.754	140	99

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00002242-RA	Lysosomal alpha-mannosidase	249	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	4E-06	60.3448 3	48.906 2	58	35
Th_d_00002243-RA	lysosomal alpha-mannosidase	1497	gi|907620158|ref|XP_013097338.1|PREDICTED: lysosomal alpha-mannosidase isoform X2 [Stomoxys calcitrans]	Diptera	XP_01309733 8	2.17E-50	47.0997 7	191.81 5	431	203
Th_d_00094171-RA	alpha-mannosidase 2	558	gi|1009600604|ref|XP_015928616.1|alpha-mannosidase 2 isoform X2 [Parasteatoda tepidariorum]gi|1009600606|ref|XP_015928617.1|alpha-mannosidase 2 isoform X2 [Parasteatoda tepidariorum]	Arachnida	XP_015928616, XP_015928617	2.03E-39	65.2694 6	148.67 3	167	109
Th_d_00012219-RA	Mannosyl-oligosaccharide alpha-1,2-mannosidase isoform, putative	273	gi|242010114|ref|XP_002425821.1|Mannosyl-oligosaccharide alpha-1,2-mannosidase isoform, putative [Pediculus humanus corporis]gi|212509754|gb|EEB13083.1|Mannosyl-oligosaccharide alpha-1,2-mannosidase isoform, putative [Pediculus humanus corporis]	Phthiraptera	XP_002425821, EEB13083	4.37E-41	92.4050 6	146.74 7	79	73

Table 3.4. Continued.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Th_d_00012218-RA</td><td>mannosyl-oligosaccharide alpha-1,2-mannosidase IA</td><td>261</td><td>gi|1022764168|gb|KZS09649.1|Alpha-1,2-Mannosidase [Daphnia magna]</td><td>Crustacea</td><td>KZS09649</td><td>4.13E-28</td><td>91.80328</td><td>109.768</td><td>61</td><td>56</td></tr>
<tr><td>Th_d_00012222-RA</td><td>mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like</td><td>492</td><td>gi|1279729672|ref|XP_022909489.1|mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like [Onthophagus taurus]</td><td>Coleoptera</td><td>XP_022909489</td><td>2.69E-49</td><td>85.29412</td><td>171.014</td><td>102</td><td>87</td></tr>
<tr><td>Th_d_00012221-RA</td><td>mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like</td><td>771</td><td>gi|1233182429|ref|XP_022192805.1|mannosyl-oligosaccharide alpha-1,2-mannosidase IA [Nilaparvata lugens]</td><td>Hemiptera</td><td>XP_022192805</td><td>3.42E-34</td><td>82.75862</td><td>135.961</td><td>87</td><td>72</td></tr>
<tr><td>Th_d_00121172-RA</td><td>alpha-mannosidase 2</td><td>354</td><td>gi|1009600604|ref|XP_015928616.1|alpha-mannosidase 2 isoform X2 [Parasteatoda tepidariorum]gi|1009600606|ref|XP_015928617.1|alpha-mannosidase 2 isoform X2 [Parasteatoda tepidariorum]</td><td>Arachnida</td><td>XP_015928616, XP_015928617</td><td>1.79E-50</td><td>89</td><td>177.178</td><td>100</td><td>89</td></tr>
</table>

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00073548-RA	lysosomal alpha-mannosidase-like	837	gi|1233202594|ref|XP_022200324.1|lysosomal alpha-mannosidase-like [Nilaparvata lugens]	Hemiptera	XP_022200324	2.81E-46	60.75949	172.17	237	144
Th_d_00091843-RA	lysosomal alpha-mannosidase-like	663	gi|1351647487|ref|XP_024086113.1|lysosomal alpha-mannosidase-like [Cimex lectularius]	Hemiptera	XP_024086113	2.38E-18	64.04494	89.3521	89	57
Th_d_00022617-RA	Beta-mannosidase-like Protein	264	gi|1192752750|ref|XP_015929642.2|beta-mannosidase [Parasteatoda tepidariorum]	Arachnida	XP_015929642	7.8E-12	78.43137	65.4698	51	40
Th_d_00022616-RA	beta-mannosidase isoform X2	990	gi|675378716|gb|KFM71618.1|Beta-mannosidase, partial [Stegodyphus mimosarum]	Arachnida	KFM71618	3.13E-76	63.36996	251.521	273	173

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00116423-RA	beta-mannosidase	366	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	2.92E-16	65.88235	79.7221	85	56
Th_d_00087751-RA	ER degradation-enhancing alpha-mannosidase-like protein 2	573	gi|1247029483|gb|PCG66925.1|hypothetical protein B5V51_7083, partial [Heliothis virescens]	Lepidoptera	PCG66925	3.7E-111	92.22222	321.627	180	166
Th_d_00046444-RA	alpha-mannosidase 2	438	gi|321444696|gb|EFX60486.1|hypothetical protein DAPPUDRAFT_343157, partial [Daphnia pulex]	Crustacea	EFX60486	1.95E-59	84.44444	192.971	135	114

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00075487-RA	lysosomal alpha-mannosidase	444	gi|1098664167|ref|XP_018794803.1|PREDICTED: lysosomal alpha-mannosidase [Bactrocera latifrons]	Diptera	XP_018794803	3.48E-13	76.78571	72.0182	56	43
Th_d_00058303-RA	lysosomal alpha-mannosidase-like	1106	gi|970882260|ref|XP_015121072.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Diachasma alloeum]	Hymenoptera	XP_015121072	2.8E-49	54.39189	184.111	296	161
Th_d_00062699-RA	alpha-mannosidase 2	1077	gi|1330899240|gb|PNF26420.1|hypothetical protein B7P43_G16606, partial [Cryptotermes secundus]	Blattodea	PNF26420	3.46E-33	62.83784	135.961	148	93
Th_d_00023682-RA	Lysosomal alpha-mannosidase	1545	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	5.9E-24	59.72222	112.849	144	86
Th_d_00086716-RA	Mannan endo-1,4-beta-mannosidase	240	gi|683427464|gb|AIN40245.1|mannanase, partial [Cherax quadricarinatus]	Crustacea	AIN40245	1.34E-23	66.66667	92.0485	78	52

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00038326-RA	lysosomal alpha-mannosidase-like	735	gi|662196129|ref|XP_008471080.1|PREDICTED: lysosomal alpha-mannosidase-like [Diaphorina citri]	Hemiptera	XP_008471080	2.82E-90	82.68156	278.1	179	148
Th_d_00038325-RA	lysosomal alpha-mannosidase-like	279	gi|1101337044|ref|XP_018914717.1|PREDICTED: lysosomal alpha-mannosidase-like [Bemisia tabaci]	Hemiptera	XP_018914717	4.25E-17	78.26087	80.4925	46	36
Th_d_00074895-RA	mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like	882	gi|1227996327|ref|XP_021929979.1|mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X1 [Zootermopsis nevadensis]	Blattodea	XP_021929979	2.96E-45	88.88889	168.318	144	128
Th_d_00097964-RA	lysosomal alpha-mannosidase	267	gi|193580067|ref|XP_001946748.1|PREDICTED: lysosomal alpha-mannosidase [Acyrthosiphon pisum]	Hemiptera	XP_001946748	6.53E-13	70.3125	68.5514	64	45
Th_d_00043225-RA	lysosomal alpha-mannosidase isoform X1	975	gi|929380069|ref|XP_014100402.1|PREDICTED: lysosomal alpha-mannosidase-like, partial [Bactrocera oleae]	Diptera	XP_014100402	6.43E-51	60.79545	173.711	176	107
Th_d_00059073-RA	alpha-mannosidase 2	711	gi|1330899240|gb|PNF26420.1|hypothetical protein B7P43_G16606, partial [Cryptotermes secundus]	Blattodea	PNF26420	7.08E-44	68.96552	161.384	174	120

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00115822-RA	mannosyl-oligosaccharide 1,2-alpha-mannosidase IA isoform X1	1500	gi|91091790|ref|XP_970226.1|PREDICTED: mannosyl-oligosaccharide 1,2-alpha-mannosidase IA isoform X1 [Tribolium castaneum]gi|1004395369|gb|KYB24855.1|Mannosyl-oligosaccharide alpha-1,2-mannosidase isoform A-like Protein [Tribolium castaneum]	Coleoptera	XP_970226, KYB24855	9.3E-116	55.8822	360.533	501	280
Th_d_00071728-RA	beta-mannosidase	486	gi|1041544405|ref|XP_008476446.2|PREDICTED: LOW QUALITY PROTEIN: beta-mannosidase [Diaphorina citri]	Hemiptera	XP_008476446	7.33E-16	79.5918	80.1073	49	39
Th_d_00120508-RA	beta-mannosidase	387	gi|1191620180|gb|OTF82415.1|Beta-mannosidase-like protein, partial [Euroglyphus maynei]	Arachnida	OTF82415	3.3E-24	78.125	100.138	64	50
Th_d_00013923-RA	lysosomal alpha-mannosidase isoform X1	3389	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	9.9E-169	64.0575	530.791	626	401

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00025482-RA	ER degradation-enhancing alpha-mannosidase-like protein 3	873	gi|1339066460|ref|XP_023725333.1| ER degradation-enhancing alpha-mannosidase-like protein 3 [Cryptotermes secundus]	Blattodea	XP_023725333	2.77E-94	96.79487	305.449	156	151
Th_d_00025483-RA	ER degradation-enhancing alpha-mannosidase-like protein 3	540	gi|1339066460|ref|XP_023725333.1| ER degradation-enhancing alpha-mannosidase-like protein 3 [Cryptotermes secundus]	Blattodea	XP_023725333	1.6E-54	71.01449	191.43	138	98
Th_d_00086879-RA	beta-mannosidase	654	gi|1279738204|ref|XP_022913917.1| beta-mannosidase [Onthophagus taurus]	Coleoptera	XP_022913917	3.56E-36	57.277	140.969	213	122
Th_d_00087558-RA	alpha-mannosidase 2	525	gi|1022777739|gb|KZS21240.1|Alpha-mannosidase 2x [Daphnia magna]	Crustacea	KZS21240	5.49E-21	70.73171	95.5153	82	58
Th_d_00012668-RA	lysosomal alpha-mannosidase	228	gi|158293330|ref|XP_557731.3|AGAP008584-PA [Anopheles gambiae str. PEST]gi|157016653|gb|EAL40242.3|AGAP008584-PA [Anopheles gambiae str. PEST]	Diptera	XP_557731, EAL40242	3.67E-12	85.71429	65.855	35	30

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00012667-RA	lysosomal alpha-mannosidase isoform X1	480	gi|1233202594|ref|XP_022200324.1|lysosomal alpha-mannosidase-like [Nilaparvata lugens]	Hemiptera	XP_022200324	1.43E-23	63.0137	102.064	146	92
Th_d_00012666-RA	lysosomal alpha-mannosidase	1071	gi|1096257085|gb|APA33853.1|seminal fluid protein [Nilaparvata lugens]	Hemiptera	APA33853	3.71E-88	71.48289	291.967	263	188
Th_d_00012665-RA	lysosomal alpha-mannosidase-like	177	gi|1233202594|ref|XP_022200324.1|lysosomal alpha-mannosidase-like [Nilaparvata lugens]	Hemiptera	XP_022200324	4.24E-20	90.69767	87.4261	43	39
Th_d_00022076-RA	Lysosomal alpha-mannosidase	756	gi|662196129|ref|XP_008471080.1|PREDICTED: lysosomal alpha-mannosidase-like [Diaphorina citri]	Hemiptera	XP_008471080	4.87E-57	58.95197	192.971	229	135
Th_d_00101434-RA	Mannan endo-1,4-beta-mannosidase	363	gi|683427464|gb|AIN40245.1|mannanase, partial [Cherax quadricarinatus]	Crustacea	AIN40245	2.09E-27	74.68354	103.219	79	59

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00038541-RA	ER degradation-enhancing alpha-mannosidase-like protein 2	1428	gi|1339057577|ref|XP_023717292.1|ER degradation-enhancing alpha-mannosidase-like protein 2 [Cryptotermes secundus]gi|1330893438|gb|PNF23714.1|ER degradation-enhancing alpha-mannosidase-like protein 2 [Cryptotermes secundus]	Blattodea	XP_023717292, PNF23714	0	92.06349	686.026	378	348
Th_d_00057594-RA	lysosomal alpha-mannosidase-like	1073	gi|1238836368|ref|XP_022237398.1|lysosomal alpha-mannosidase-like [Limulus polyphemus]	Arthropoda-Chelicerata	XP_022237398	1.13E-87	65.68915	288.115	341	224
Th_d_00075466-RA	beta-mannosidase	495	gi|998509058|ref|XP_015514937.1|PREDICTED: beta-mannosidase [Neodiprion lecontei]	Hymenoptera	XP_015514937	3.24E-23	87.93103	101.293	58	51
Th_d_00096336-RA	lysosomal alpha-mannosidase	249	gi|193580067|ref|XP_001946748.1|PREDICTED: lysosomal alpha-mannosidase [Acyrthosiphon pisum]	Hemiptera	XP_001946748	4.12E-13	70.3125	68.9366	64	45
Th_d_00047111-RA	Beta-mannosidase	417	gi|1061485315|gb|ODN02454.1|Beta-mannosidase [Orchesella cincta]	Collembola	ODN02454	1.79E-11	74.54545	66.6254	55	41

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00047110-RA	Beta-mannosidase	258	gi|1228363654|ref|XP_021968444.1|beta-mannosidase-like [Folsomia candida]gi|1215278402|gb|OXA64711.1|Beta-mannosidase [Folsomia candida]	Collembola	XP_021968444, OXA64711	1.86E-12	67.3076 9	67.0106	52	35
Th_d_00053244-RA	lysosomal alpha-mannosidase isoform X2	831	gi|242016705|ref|XP_002428888.1|predicted protein [Pediculus humanus corporis]gi|212513656|gb|EEB16150.1|predicted protein [Pediculus humanus corporis]	Phthiraptera	XP_002428888, EEB16150	5.24E-48	74.63768	176.792	138	103
Th_d_00078892-RA	lysosomal alpha-mannosidase	432	gi|195377741|ref|XP_002047646.1|uncharacterized protein Dvir_GJ11812 [Drosophila virilis]gi|194154804|gb|EDW69988.1|uncharacterized protein Dvir_GJ11812 [Drosophila virilis]	Diptera	XP_002047646, EDW69988	1.64E-16	75	81.2629	60	45

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00070856-RA	alpha-mannosidase 2	615	gi|1339052711|ref|XP_023714691.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052713|ref|XP_023714692.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052715|ref|XP_023714693.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052717|ref|XP_023714694.1|alpha-mannosidase 2 [Cryptotermes secundus]	Blattodea	XP_023714691, XP_023714692, XP_023714693, XP_023714694	9.43E-97	84.73684	309.686	190	161
Th_d_00087276-RA	lysosomal alpha-mannosidase-like	528	gi|662196129|ref|XP_008471080.1|PREDICTED: lysosomal alpha-mannosidase-like [Diaphorina citri]	Hemiptera	XP_008471080	3.48E-67	85.82677	216.083	127	109
Th_d_00003340-RA	beta-mannosidase	435	gi|1325354705|ref|XP_023311948.1|beta-mannosidase-like [Anoplophora glabripennis]	Coleoptera	XP_023311948	4.96E-14	73.01587	72.0182	63	46
Th_d_00003341-RA	beta-mannosidase	435	gi|1325354705|ref|XP_023311948.1|beta-mannosidase-like [Anoplophora glabripennis]	Coleoptera	XP_023311948	4.96E-14	73.01587	72.0182	63	46

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00069730-RA	beta-mannosidase	648	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	9.92E-37	76.52174	142.124	115	88
Th_d_00027880-RA	mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like	537	gi|642937291|ref|XP_008198773.1|PREDICTED: mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X2 [Tribolium castaneum]	Coleoptera	XP_008198773	5.22E-87	88.20225	270.396	178	157

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_000083699-RA	alpha-mannosidase 2	629	gi|1227998244|ref|XP_021930961.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998246|ref|XP_021930962.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998248|ref|XP_021930963.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998250|ref|XP_021930964.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998252|ref|XP_021930965.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998254|ref|XP_021930966.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|646706309|gb|KDR13614.1|Alpha-mannosidase 2 [Zootermopsis nevadensis]	Blattodea	XP_021930961, XP_021930962, XP_021930963, XP_021930964, XP_021930965, XP_021930966, KDR13614	1.31E-29	56.14973	121.709	187	105

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00043179-RA	beta-mannosidase	1224	gi|242009114|ref|XP_002425337.1|beta-mannosidase precursor, putative [Pediculus humanus corporis]gi|212509122|gb|EEB12599.1|beta-mannosidase precursor, putative [Pediculus humanus corporis]	Phthiraptera	XP_002425337, EEB12599	1.6E-84	77.51196	281.567	209	162
Th_d_00059913-RA	alpha-mannosidase 2	696	gi|1330899240|gb|PNF26420.1|hypothetical protein B7P43_G16606, partial [Cryptotermes secundus]	Blattodea	PNF26420	5.14E-44	68.96552	161.77	174	120
Th_d_00048797-RA	beta-mannosidase	1293	gi|1316169311|ref|XP_023230069.1|beta-mannosidase-like isoform X1 [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023230069	2.83E-31	52.81385	133.265	231	122
Th_d_00094873-RA	ER degradation-enhancing alpha-mannosidase-like protein 2	708	gi|1339057577|ref|XP_023717292.1|ER degradation-enhancing alpha-mannosidase-like protein 2 [Cryptotermes secundus]gi|1330893438|gb|PNF23714.1|ER degradation-enhancing alpha-mannosidase-like protein 2 [Cryptotermes secundus]	Blattodea	XP_023717292, PNF23714	7.7E-114	84.03756	350.132	213	179

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_000113270-RA	ER degradation-enhancing alpha-mannosidase-like protein 2	1272	gi|1227987439|ref|XP_021925413.1|ER degradation-enhancing alpha-mannosidase-like protein 2, partial [Zootermopsis nevadensis]	Blattodea	XP_021925413	5.28E-43	92.7108	166.777	83	77
Th_d_00004556-RA	beta-mannosidase	474	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	1.22E-19	48.07692	90.8929	156	75
Th_d_00058265-RA	lysosomal alpha-mannosidase isoform X2	782	gi|242016705|ref|XP_002428888.1|predicted protein [Pediculus humanus corporis]gi|212513656|gb|EEB16150.1|predicted protein [Pediculus humanus corporis]	Phthiraptera	XP_002428888, EEB16150	7.14E-49	73.28767	178.333	146	107

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00084290-RA	Lysosomal alpha-mannosidase	434	gi|1080066641|ref|XP_018575534.1|lysosomal alpha-mannosidase isoform X2 [Anoplophora glabripennis]	Coleoptera	XP_018575534	1.31E-18	53.23741	87.4261	139	74
Th_d_00038457-RA	ER degradation-enhancing alpha-mannosidase-like protein 3	1560	gi|1227972165|ref|XP_021917507.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X1 [Zootermopsis nevadensis]gi|1227972167|ref|XP_021917508.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X1 [Zootermopsis nevadensis]gi|646718115|gb|KDR20717.1|ER degradation-enhancing alpha-mannosidase-like 3 [Zootermopsis nevadensis]	Blattodea	XP_021917507, XP_021917508, KDR20717	2.58E-58	54.18994	215.312	358	194
Th_d_00008943-RA	lysosomal alpha-mannosidase	1628	gi|1000733841|ref|XP_015589368.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Cephus cinctus]	Hymenoptera	XP_015589368	6.09E-50	70.94972	191.43	179	127
Th_d_00008944-RA	lysosomal alpha-mannosidase-like	432	gi|1035439623|gb|ANJ04662.1|lysosomal alpha-mannosidase-like protein [Nilaparvata lugens]	Hemiptera	ANJ04662	1.89E-33	72.32143	129.798	112	81

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_000049153-RA	lysosomal alpha-mannosidase-like	711	gi|668446735|gb|KFB36711.1|lysosomal alpha-mannosidase (mannosidase alpha class 2b member 1) [Anopheles sinensis]	Diptera	KFB36711	1.29E-13	54.80769	76.2554	104	57
Th_d_000012242-RA	ER degradation-enhancing alpha-mannosidase-like protein 1	534	gi|1227982315|ref|XP_021922781.1|ER degradation-enhancing alpha-mannosidase-like protein 1 [Zootermopsis nevadensis]gi|646713767|gb|KDR17988.1|ER degradation-enhancing alpha-mannosidase-like 1 [Zootermopsis nevadensis]	Blattodea	XP_021922781, KDR17988	6.29E-98	92.07317	299.671	164	151
Th_d_000012243-RA	ER degradation-enhancing alpha-mannosidase-like protein 1	384	gi|1339070655|ref|XP_023722504.1|ER degradation-enhancing alpha-mannosidase-like protein 1, partial [Cryptotermes secundus]	Blattodea	XP_023722504	1.42E-64	92.56198	206.068	121	112
Th_d_000012244-RA	ER degradation-enhancing alpha-mannosidase-like protein 1	870	gi|1330882322|gb|PNF18187.1|ER degradation-enhancing alpha-mannosidase-like protein 1, partial [Cryptotermes secundus]	Blattodea	PNF18187	1.1E-25	90.90909	110.538	66	60

Table 3.4. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00015392-RA	beta-mannosidase	753	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	4.88E-48	67.68293	176.022	164	111
Th_d_00015391-RA	beta-mannosidase	564	gi|242009114|ref|XP_002425337.1|beta-mannosidase precursor, putative [Pediculus humanus corporis]gi|212509122|gb|EEB12599.1|beta-mannosidase precursor, putative [Pediculus humanus corporis]	Phthiraptera	XP_002425337, EEB12599	3.55E-34	78.30189	133.65	106	83
Th_d_00042295-RA	lysosomal alpha-mannosidase-like	1251	gi|321463091|gb|EFX74109.1|hypothetical protein DAPPUDRAFT_324591 [Daphnia pulex]	Crustacea	EFX74109	1.35E-72	81.76101	251.136	159	130

Table 3.4. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00071433-RA	alpha-mannosidase 2	2781	gi|1339078211|ref|XP_023726521.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339078213|ref|XP_023726522.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339078215|ref|XP_023726523.1|alpha-mannosidase 2 [Cryptotermes secundus]	Blattodea	XP_023726521, XP_023726522, XP_023726523	0	76.78571	1333.55	1008	774
Th_d_00091792-RA	lysosomal alpha-mannosidase isoform X1	531	gi|929380069|ref|XP_014100402.1|PREDICTED: lysosomal alpha-mannosidase-like, partial [Bactrocera oleae]	Diptera	XP_014100402	5.57E-61	79.23077	194.126	130	103
Th_d_00089230-RA	alpha-mannosidase 2x-like	372	gi|241567773|ref|XP_002402365.1|lysosomal alpha-mannosidase, putative [Ixodes scapularis]gi|215501997|gb|EEC11491.1|lysosomal alpha-mannosidase, putative, partial [Ixodes scapularis]	Arthropoda-Chelicerata	XP_002402365, EEC11491	9.37E-17	71.875	76.2554	64	46
Th_d_00042901-RA	lysosomal alpha-mannosidase-like	573	gi|662196129|ref|XP_008471080.1|PREDICTED: lysosomal alpha-mannosidase-like [Diaphorina citri]	Hemiptera	XP_008471080	8.16E-45	70.99237	158.688	131	93

Table 3.5. Glucuronidases in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for glucuronidases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00016998-RA	beta-glucuronidase isoform X1	306	gi|1067103892|ref|XP_018013304.1|PREDICTED: beta-glucuronidase-like [Hyalella azteca]	Crustacea	XP_018013304	1.56E-21	82.8125	93.5893	64	53
Th_d_00066936-RA	beta-glucuronidase-like isoform X1	855	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023221887	1.44E-75	61.50943	244.588	265	163
Th_d_00011907-RA	beta-glucuronidase	465	gi|909578964|ref|XP_013145358.1|PREDICTED: beta-glucuronidase [Papilio polytes]	Lepidoptera	XP_013145358	3.2E-57	73.8255	192.971	149	110
Th_d_00011908-RA	beta-glucuronidase isoform X3	539	gi|1108476517|emb|CRK95286.1|CLUMA_CG008644, isoform B [Clunio marinus]	Diptera	CRK95286	6.19E-12	62.66667	69.3218	75	47
Th_d_00057287-RA	beta-glucuronidase-like isoform X1	634	gi|1330889853|gb|PNF22051.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	PNF22051	9.67E-83	78.16092	253.832	174	136

Table 3.5. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00033402-RA	beta-glucuronidase isoform X3	1014	gi|194756422|ref|XP_001960477.1|uncharacterized protein Dana_GF11493 [Drosophila ananassae]gi|190621775|gb|EDV37299.1|uncharacterized protein Dana_GF11493 [Drosophila ananassae]	Diptera	XP_001960477, EDV37299	2.44E-47	52.33333	175.637	300	157
Th_d_00068821-RA	beta-glucuronidase-like	521	gi|1022772830|gb|KZS16987.1|Beta-glucuronidase [Daphnia magna]	Crustacea	KZS16987	4.99E-56	75.53957	178.718	139	105
Th_d_00098822-RA	beta-glucuronidase-like	528	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023221887	5.6E-64	79.08497	209.92	153	121

Table 3.5. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00020056-RA	beta-glucuronidase isoform X1	1260	gi|1227991600|ref|XP_021927576.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991602|ref|XP_021927578.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991604|ref|XP_021927579.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991606|ref|XP_021927580.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991608|ref|XP_021927581.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991610|ref|XP_021927582.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991612|ref|XP_021927583.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991614|ref|XP_021927584.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|646709721|gb|KDR15442.1|Beta-glucuronidase [Zootermopsis nevadensis]	Blattodea	XP_021927576, XP_021927578, XP_021927579, XP_021927580, XP_021927581, XP_021927582, XP_021927583, XP_021927584, KDR15442	7.26E-95	57.66017	303.523	359	207
Th_d_00105741-RA	beta-glucuronidase isoform X3	273	gi|939251322|ref|XP_014245262.1|beta-glucuronidase isoform X5 [Cimex lectularius]	Hemiptera	XP_014245262	4.79E-16	82.97872	77.411	47	39

Table 3.5. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00031354-RA	Beta-glucuronidase	684	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1 |hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	2.29E-49	52.94118	175.252	221	117
Th_d_00031355-RA	beta-glucuronidase isoform X1	441	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023221887	8.16E-23	72.97297	98.9821	74	54
Th_d_00017720-RA	beta-glucuronidase isoform X3	882	gi|1121168380|ref|XP_019529786.1|PREDICTED: beta-glucuronidase-like isoform X1 [Aedes albopictus]	Diptera	XP_019529786	1.26E-59	58.00712	207.608	281	163
Th_d_00017721-RA	beta-glucuronidase-like	462	gi|325303090|tpg|DAA34282.1|TPA_inf: beta-glucuronidase GUSB [Amblyomma variegatum]	Arachnida	DAA34282	9.22E-48	74.60317	159.073	126	94
Th_d_00029861-RA	beta-glucuronidase isoform X3	1027	gi|557767812|ref|XP_005183593.1|PREDICTED: beta-glucuronidase isoform X2 [Musca domestica]	Diptera	XP_005183593	5.35E-78	69.23077	257.299	247	171

Table 3.5. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00094572-RA	beta-glucuronidase	303	gi|1229725314|ref|XP_022125377.1|beta-glucuronidase isoform X3 [Pieris rapae]gi|1229725316|ref|XP_022125378.1|beta-glucuronidase isoform X3 [Pieris rapae]	Lepidoptera	XP_022125377, XP_022125378	1.59E-15	66.17647	76.6406	68	45
Th_d_00060577-RA	Beta-glucuronidase	819	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	2.87E-57	66.85714	197.978	175	117
Th_d_00103092-RA	beta-glucuronidase isoform X3	356	gi|939667591|ref|XP_014280878.1|PREDICTED: beta-glucuronidase-like isoform X2 [Halyomorpha halys]	Hemiptera	XP_014280878	1.76E-26	78.08219	108.612	73	57
Th_d_00085133-RA	beta-glucuronidase isoform X2	516	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023221887	5.38E-48	81.19658	167.933	117	95

Table 3.5. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00018098-RA	beta-glucuronidase isoform X3	993	gi|646696603|gb|KDR08779.1|Beta-glucuronidase [Zootermopsis nevadensis]	Blattodea	KDR08779	2.5E-147	78.36991	434.491	319	250
Th_d_00119947-RA	beta-glucuronidase isoform X2	318	gi|951547932|ref|XP_014474109.1|PREDICTED: beta-glucuronidase isoform X1 [Dinoponera quadriceps]	Hymenoptera	XP_014474109	1.61E-14	57.69231	73.9442	78	45
Th_d_00005617-RA	beta-glucuronidase-like isoform X1	1886	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	0	62.9174	559.681	569	358
Th_d_00005619-RA	beta-glucuronidase-like isoform X1	597	gi|241309940|ref|XP_002407819.1|beta-glucuronidase (GusB), putative [Ixodes scapularis]gi|215497226|gb|EEC06720.1|beta-glucuronidase (GusB), putative [Ixodes scapularis]	Arachnida	XP_002407819, EEC06720	2.9E-45	74.60317	155.606	126	94
Th_d_00056689-RA	beta-glucuronidase isoform X3	866	gi|1022772830|gb|KZS16987.1|Beta-glucuronidase [Daphnia magna]	Crustacea	KZS16987	1.62E-58	74.63768	189.119	138	103

Table 3.5. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00098010-RA	Beta-glucuronidase	432	gi|1339045451|ref|XP_023710854.1|beta-glucuronidase-like [Cryptotermes secundus]gi|1339045453|ref|XP_023710855.1|beta-glucuronidase-like [Cryptotermes secundus]gi|1330907803|gb|PNF30365.1|Beta-glucuronidase [Cryptotermes secundus]	Blattodea	XP_023710854, XP_023710855, PNF30365	4.82E-13	62.66667	71.2478	75	47
Th_d_00094268-RA	Beta-glucuronidase	309	gi|662210289|ref|XP_008478777.1|PREDICTED: beta-glucuronidase-like [Diaphorina citri]	Hemiptera	XP_008478777	4.17E-31	78.40909	115.546	88	69
Th_d_00049078-RA	beta-glucuronidase-like isoform X3	916	gi|1330889853|gb|PNF22051.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	PNF22051	4.5E-115	76.55678	339.732	273	209
Th_d_00077520-RA	Beta-glucuronidase	801	gi|1339045451|ref|XP_023710854.1|beta-glucuronidase-like [Cryptotermes secundus]gi|1339045453|ref|XP_023710855.1|beta-glucuronidase-like [Cryptotermes secundus]gi|1330907803|gb|PNF30365.1|Beta-glucuronidase [Cryptotermes secundus]	Blattodea	XP_023710854, XP_023710855, PNF30365	2.48E-32	76.78571	130.954	112	86

Table 3.5. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00068165-RA	beta-glucuronidase-like isoform X2	618	gi|1022772830|gb|KZS16987.1|Beta-glucuronidase [Daphnia magna]	Crustacea	KZS16987	7.51E-55	74.10072	176.792	139	103
Th_d_00063635-RA	beta-glucuronidase-like isoform X3	819	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	1.63E-42	67.68293	158.688	164	111
Th_d_00075299-RA	beta-glucuronidase	498	gi|985391249|ref|XP_015380321.1|PREDICTED: beta-glucuronidase [Diuraphis noxia]	Crustacea	XP_015380321	3.58E-51	69.93464	175.252	153	107
Th_d_00070918-RA	beta-glucuronidase isoform X2	591	gi|91089481|ref|XP_969353.1|PREDICTED: beta-glucuronidase [Tribolium castaneum]gi|270011394|gb|EFA07842.1|Beta-glucuronidase-like Protein [Tribolium castaneum]	Coleoptera	XP_969353, EFA07842	1.4E-12	59.25926	71.633	81	48

Table 3.5. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00080262-RA	beta-glucuronidase-like isoform X1	708	gi|1339045451|ref|XP_023710854.1|beta-glucuronidase-like [Cryptotermes secundus]gi|1339045453|ref|XP_023710855.1|beta-glucuronidase-like [Cryptotermes secundus]gi|1330907803|gb|PNF30365.1|Beta-glucuronidase [Cryptotermes secundus]	Blattodea	XP_023710854, XP_023710855, PNF30365	3.33E-50	78.26087	162.54	115	90
Th_d_00100552-RA	beta-glucuronidase isoform X2	354	gi|91089481|ref|XP_969353.1|PREDICTED: beta-glucuronidase [Tribolium castaneum]gi|270011394|gb|EFA07842.1|Beta-glucuronidase-like Protein [Tribolium castaneum]	Coleoptera	XP_969353, EFA07842	7.51E-12	64.61538	66.6254	65	42
Th_d_00070705-RA	beta-glucuronidase-like	552	gi|1316206353|ref|XP_023213588.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023213588	4.75E-55	77.53623	180.644	138	107
Th_d_00071727-RA	beta-glucuronidase-like	819	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023221887	7.65E-46	59.15493	166.392	213	126

Table 3.5. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00043633-RA	beta-glucuronidase-like isoform X1	480	gi|1022772830|gb|KZS16987.1|Beta-glucuronidase [Daphnia magna]	Crustacea	KZS16987	7.75E-56	74.10072	177.563	139	103
Th_d_00043632-RA	beta-glucuronidase isoform X1	276	gi|662210289|ref|XP_008478777.1|PREDICTED: beta-glucuronidase-like [Diaphorina citri]	Hemiptera	XP_008478777	2.15E-22	69.73684	92.4337	76	53
Th_d_00017000-RA	beta-glucuronidase isoform X3	849	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	4.1E-103	72.5	317.39	280	203

Table 3.6. Maltases in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for maltases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00074079-RA	maltase 2-like	678	gi|1228382183|ref|XP_021952619.1|alpha-glucosidase-like [Folsomia candida]	Collembola	XP_021952619	1.68E-21	87.5	98.5969	56	49
Th_d_00040803-RA	maltase-glucoamylase, intestinal-like	697	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	2.31E-56	64.02116	199.519	189	121
Th_d_00083166-RA	Maltase-glucoamylase, intestinal	580	gi|1022742921|gb|KZR99685.1|Lysosomal alpha-glucosidase, partial [Daphnia magna]	Crustacea	KZR99685	3.96E-57	64.58333	191.43	192	124
Th_d_00152677-RA	Maltase 1	177	gi|646696837|gb|KDR08865.1|Maltase 1 [Zootermopsis nevadensis]	Blattodea	KDR08865	6.92E-16	85.71429	75.485	42	36
Th_d_00066308-RA	maltase 2-like	732	gi|1228013734|ref|XP_021938828.1|maltase 2-like [Zootermopsis nevadensis]	Blattodea	XP_021938828	5.68E-54	67.42857	187.578	175	118

Table 3.6. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00031469-RA	Maltase 1	594	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	2.17E-61	72	204.527	175	126
Th_d_00031468-RA	maltase A3-like isoform X2	513	gi|636630766|gb|AIA09343.1|alpha-glucosidase [Periplaneta americana]	Blattodea	AIA09343	4.88E-11	66.15385	66.6254	65	43
Th_d_00112867-RA	maltase 1-like isoform X2	483	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	1.09E-51	68.87417	177.563	151	104

Table 3.6. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00056395-RA	maltase 1-like isoform X2	1182	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	4.5E-112	64.76965	342.813	369	239
Th_d_00065847-RA	Maltase 1	420	gi|1330891185|gb|PNF22670.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	PNF22670	3.52E-40	71.96262	139.813	107	77
Th_d_00108497-RA	Maltase 1	420	gi|1061480896|gb|ODM98289.1|Maltase 1 [Orchesella cincta]	Collembola	ODM98289	3.7E-40	77.3913	147.132	115	89
Th_d_00059940-RA	maltase 1-like isoform X2	1074	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	2.8E-118	64.72222	357.066	360	233

Table 3.6. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00081100-RA	Maltase 1	600	gi|1228013734|ref|XP_021938828.1|maltase 2-like [Zootermopsis nevadensis]	Blattodea	XP_021938828	5.61E-67	67.66169	219.935	201	136
Th_d_00061180-RA	maltase-glucoamylase, intestinal-like	1203	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	2.46E-34	73	141.739	100	73
Th_d_00084554-RA	Maltase-glucoamylase, intestinal	598	gi|1022742921|gb|KZR99685.1|Lysosomal alpha-glucosidase, partial [Daphnia magna]	Crustacea	KZR99685	4.25E-57	64.58333	191.43	192	124
Th_d_00033664-RA	Maltase-glucoamylase, intestinal	924	gi|1022768805|gb|KZS13426.1|Uncharacterized protein APZ42_021387 [Daphnia magna]	Crustacea	KZS13426	3.76E-53	75.17241	192.971	145	109
Th_d_00033663-RA	maltase-glucoamylase, intestinal-like	402	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	6.99E-06	60	50.447	55	33
Th_d_00108505-RA	Maltase 1	222	gi|1228013734|ref|XP_021938828.1|maltase 2-like [Zootermopsis nevadensis]	Blattodea	XP_021938828	1.9E-24	85	100.523	60	51

Table 3.6. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00010925-RA	maltase-glucoamylase, intestinal-like	360	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	7.79E-17	64	81.2629	75	48
Th_d_00010928-RA	sucrase-isomaltase, intestinal-like	651	gi|1238867838|ref|XP_022249218.1|sucrase-isomaltase, intestinal-like [Limulus polyphemus]	Arthropoda-atlantic horse-shoe crab	XP_022249218	7.1E-28	71.56863	116.701	102	73
Th_d_00110401-RA	Maltase 1	267	gi|1061480896|gb|ODM98289.1|Maltase 1 [Orchesella cincta]	Collembola	ODM98289	7.7E-22	83.05085	93.9745	59	49
Th_d_00096117-RA	maltase 1	555	gi|1227980746|ref|XP_021921963.1|uncharacterized protein KIAA0513 isoform X3 [Zootermopsis nevadensis]	Blattodea	XP_021921963	1.5E-115	93.25843	341.658	178	166

Table 3.6. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00073223-RA	Maltase 1	927	gi|1339067252|ref|XP_023720702.1|uncharacterized protein KIAA0513 [Cryptotermes secundus]gi|1330886365|gb|PNF20093.1|hypothetical protein B7P43_G05254 [Cryptotermes secundus]gi|1330886366|gb|PNF20094.1|hypothetical protein B7P43_G05254 [Cryptotermes secundus]	Blattodea	XP_023720702, PNF20093, PNF20094	1.78E-66	71.91781	221.09	292	210
Th_d_00008617-RA	Maltase 1	1668	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	0	68.47826	608.216	552	378
Th_d_00008616-RA	Maltase 1	335	gi|1061473040|gb|ODM91568.1|Maltase 1 [Orchesella cincta]	Collembola	ODM91568	3.54E-09	65.625	58.9214	64	42
Th_d_00037981-RA	sucrase-isomaltase, intestinal-like	1347	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	1.17E-90	55.1963	305.449	433	239

Table 3.6. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00106192-RA	maltase 2-like	228	gi|1233179883|ref|XP_022191429.1|maltase A3-like [Nilaparvata lugens]	Hemiptera	XP_022191429	8.92E-16	61.97183	75.8702	71	44
Th_d_00094575-RA	Maltase 1	432	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	7.63E-36	65.21739	134.806	138	90
Th_d_00029984-RA	maltase-glucoamylase, intestinal	1509	gi|321476730|gb|EFX87690.1|hypothetical protein DAPPUDRAFT_312136 [Daphnia pulex]	Crustacea	EFX87690	4E-99	71.05263	325.094	266	189
Th_d_00098066-RA	maltase 2-like	477	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	4.21E-45	73.48485	161.384	132	97

Table 3.6. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_000091760-RA	Maltase 1	819	gi|1339067252|ref|XP_023720702.1|uncharacterized protein KIAA0513 [Cryptotermes secundus]gi|1330886365|gb|PNF20093.1|hypothetical protein B7P43_G05254 [Cryptotermes secundus]gi|1330886366|gb|PNF20094.1|hypothetical protein B7P43_G05254 [Cryptotermes secundus]	Blattodea	XP_023720702, PNF20093, PNF20094	5.25E-71	76.83824	231.106	272	209
Th_d_000092131-RA	Maltase 1	453	gi|1061480896|gb|ODM98289.1|Maltase 1 [Orchesella cincta]	Collembola	ODM98289	1.88E-22	85.9375	98.5969	64	55
Th_d_000086494-RA	maltase 2-like	600	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	5.83E-71	74.58564	231.106	181	135
Th_d_000105557-RA	sucrase-isomaltase, intestinal-like	507	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	4.86E-41	65.66265	152.525	166	109

Table 3.6. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00041901-RA	Maltase 1	201	gi|1153707331|ref|XP_020295361.1|alpha-glucosidase-like [Pseudomyrmex gracilis]	Hymenoptera	XP_020295361	1.64E-16	75.75758	77.7962	66	50
Th_d_00093457-RA	Maltase 1	438	gi|1228013734|ref|XP_021938828.1|maltase 2-like [Zootermopsis nevadensis]	Blattodea	XP_021938828	1.07E-23	85	101.679	60	51
Th_d_00075301-RA	maltase 1	690	gi|1227980746|ref|XP_021921963.1|uncharacterized protein KIAA0513 isoform X3 [Zootermopsis nevadensis]	Blattodea	XP_021921963	6.5E-118	96.51163	349.747	172	166
Th_d_00013780-RA	Maltase-glucoamylase, intestinal	525	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	4.78E-30	61.9403	121.324	134	83
Th_d_00093020-RA	maltase 2-like	627	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	8.8E-70	75.70621	228.409	177	134

Table 3.7. Amylases in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for amylases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00029486-RA	alpha-amylase 1-like	558	gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605017	3.02E-20	83.60656	93.5893	61	51
Th_d_00029487-RA	Alpha-amylase 1	837	gi|1339094654|ref|XP_023708592.1|alpha-amylase 1-like [Cryptotermes secundus]gi|1330912880|gb|PNF32627.1|Alpha-amylase 1 [Cryptotermes secundus]	Blattodea	XP_023708592, PNF32627	6.11E-29	83.56164	120.939	73	61
Th_d_00082350-RA	alpha-amylase 1-like	538	gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605017	4.65E-09	81.57895	60.8474	38	31
Th_d_00112407-RA	alpha-amylase 1-like	246	gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605017	1.27E-21	90.38462	92.8189	52	47
Th_d_00058535-RA	alpha-amylase-like isoform X2	402	gi|854959469|gb|AKN79742.1|amylase, partial [Procambarus clarkii]	Crustacea	AKN79742	1.85E-10	64.0625	63.5438	64	41

Table 3.7. Continued.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Th_d_00049613-RA</td><td>alpha-amylase-like isoform X2</td><td>732</td><td>gi|646713944|gb|KDR18093.1|hypothetical protein L798_07788, partial [Zootermopsis nevadensis]</td><td>Blattodea</td><td>KDR18093</td><td>6.29E-37</td><td>85.55556</td><td>141.354</td><td>90</td><td>77</td></tr>
<tr><td>Th_d_00147613-RA</td><td>alpha-amylase-like isoform X2</td><td>156</td><td>gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]</td><td>Hymenoptera</td><td>XP_015605017</td><td>7.59E-19</td><td>84.61538</td><td>83.5741</td><td>52</td><td>44</td></tr>
<tr><td>Th_d_00086652-RA</td><td>Alpha-amylase 1</td><td>447</td><td>gi|540849793|gb|AGV15452.1|alpha-amylase, partial [Blattella germanica]</td><td>Blattodea</td><td>AGV15452</td><td>1.84E-66</td><td>78.52349</td><td>215.312</td><td>149</td><td>117</td></tr>
<tr><td>Th_d_00132371-RA</td><td>alpha-amylase 1-like</td><td>171</td><td>gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]</td><td>Hymenoptera</td><td>XP_015605017</td><td>3.22E-19</td><td>82.14286</td><td>84.7297</td><td>56</td><td>46</td></tr>
<tr><td>Th_d_00074258-RA</td><td>alpha-amylase 1-like</td><td>225</td><td>gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]</td><td>Hymenoptera</td><td>XP_015605017</td><td>4E-19</td><td>72.85714</td><td>85.5001</td><td>70</td><td>51</td></tr>
</table>

Table 3.7. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00050952-RA	alpha-amylase-like isoform X2	378	gi|1000764249|ref|XP_015605210.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605210	7.65E-42	88.88889	149.828	90	80
Th_d_00120835-RA	alpha-amylase A-like	276	gi|68266167|gb|AAY88846.1|alpha-amylase, partial [Musca domestica]	Diptera	AAY88846	1.03E-31	73.91304	114.005	92	68
Th_d_00095053-RA	Alpha-amylase 1	249	gi|62955866|gb|AAY23288.1|1,4-alpha-D-glucan glucanohydrolase precursor [Blattella germanica]	Blattodea	AAY23288	6.52E-16	77.08333	76.6406	48	37
Th_d_00106074-RA	Pancreatic alpha-amylase	294	gi|1061479081|gb|ODM96628.1|Pancreatic alpha-amylase [Orchesella cincta]	Collembola	ODM96628	1.59E-15	55.55556	76.2554	90	50
Th_d_00026691-RA	alpha-amylase-like isoform X2	360	gi|1000764249|ref|XP_015605210.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605210	5.74E-42	88.88889	149.828	90	80

Table 3.7. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00026690-RA	Alpha-amylase 1	669	gi|1339094654|ref|XP_023708592.1|alpha-amylase 1-like [Cryptotermes secundus]gi|1330912880|gb|PNF32627.1|Alpha-amylase 1 [Cryptotermes secundus]	Blattodea	XP_023708592, PNF32627	3.96E-31	83.33333	125.176	78	65
Th_d_00060374-RA	Pancreatic alpha-amylase	726	gi|817060103|ref|XP_012251588.1|alpha-amylase A-like [Athalia rosae]	Hymenoptera	XP_012251588	9.71E-25	52.12121	107.842	165	86
Th_d_00013282-RA	alpha-amylase 1-like	429	gi|1339094654|ref|XP_023708592.1|alpha-amylase 1-like [Cryptotermes secundus]gi|1330912880|gb|PNF32627.1|Alpha-amylase 1 [Cryptotermes secundus]	Blattodea	XP_023708592, PNF32627	5.94E-31	83.56164	121.324	73	61
Th_d_00013283-RA	Pancreatic alpha-amylase	219	gi|755948639|ref|XP_011300849.1|PREDICTED: alpha-amylase-like [Fopius arisanus]	Hymenoptera	XP_011300849	2.28E-15	94.59459	74.7146	37	35
Th_d_00013284-RA	alpha-amylase-like isoform X2	735	gi|62955866|gb|AAY23288.1|1,4-alpha-D-glucan glucanohydrolase precursor [Blattella germanica]	Blattodea	AAY23288	1.69E-10	82.92683	67.0106	41	34

Table 3.7. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00048517-RA	alpha-amylase	563	gi|1022776232|gb|KZS19963.1|Alpha-amylase 1 [Daphnia magna]	Crustacea	KZS19963	2.18E-74	72.34043	238.81	188	136
Th_d_00117725-RA	alpha-amylase 1-like	240	gi|62955866|gb|AAY23288.1|1,4-alpha-D-glucan glucanohydrolase precursor [Blattella germanica]	Blattodea	AAY23288	1.05E-15	85.71429	75.8702	42	36
Th_d_00090295-RA	alpha amylase	414	gi|390429576|gb|AFL90691.1|alpha amylase, partial [Ochlerotatus theobaldi]	Diptera	AFL90691	9.56E-18	73.77049	82.8037	61	45
Th_d_00013285-RA	alpha amylase	843	gi|1341138358|gb|AVA17418.1|putative Per a 11 allergen variant, partial [Periplaneta americana]	Blattodea	AVA17418	5.03E-16	63.41463	82.4185	82	52

Table 3.8. Galactosidases in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for galactosidases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00001121-RA	beta-galactosidase-1-like protein 2	2986	gi|1067065262|ref|XP_018016606.1|PREDICTED: beta-galactosidase-1-like protein 2, partial [Hyalella azteca]	Crustacea	XP_018016606	5.7E-22	69.79167	108.227	96	67
Th_d_00001119-RA	Beta-galactosidase-1-like protein 2	1191	gi|1330893432|gb|PNF23708.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattodea	PNF23708	1.54E-39	64.47368	152.525	152	98
Th_d_00001120-RA	beta-galactosidase-1-like protein 2	1119	gi|1330893428|gb|PNF23704.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattodea	PNF23704	3.74E-31	66.66667	127.872	114	76
Th_d_00001122-RA	Beta-galactosidase-1-like protein 2	1887	gi|1000751438|ref|XP_015598573.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X1 [Cephus cinctus]	Hymenoptera	XP_015598573	0	70.07874	648.277	635	445

Table 3.8. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00020110-RA	beta-galactosidase-like isoform X1	511	gi|1191621300|gb|OTF82605.1|beta-galactosidase-like protein, partial [Euroglyphus maynei]	Arachnida	OTF82605	2.46E-18	68.35443	65.855	79	54
Th_d_00020108-RA	beta-galactosidase, putative	294	gi|240999711|ref|XP_002404775.1|beta-galactosidase, putative [Ixodes scapularis]gi|215491658|gb|EEC01299.1|beta-galactosidase, putative, partial [Ixodes scapularis]	Arachnida	XP_002404775, EEC01299	1.37E-17	68.83117	82.0333	77	53
Th_d_00017291-RA	beta-galactosidase	483	gi|1238876097|ref|XP_022251593.1|beta-galactosidase-like isoform X2 [Limulus polyphemus]	Arthropoda-Atlantic horseshoe crab	XP_022251593	4.84E-17	83.33333	83.1889	54	45
Th_d_00017287-RA	beta-galactosidase precursor, putative	333	gi|1316154457|ref|XP_023222257.1|beta-galactosidase-like [Centruroides sculpturatus]	Arthropoda-Scorpiones	XP_023222257	1.22E-20	81.81818	91.6633	55	45

Table 3.8. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00003374-RA	beta-galactosidase isoform X2	1720	gi|321478650|gb|EFX89607.1|hypothetical protein DAPPUDRAFT_303198 [Daphnia pulex]	Crustacean	EFX89607	1.3E-115	66.50124	363.229	403	268
Th_d_00117214-RA	Beta-galactosidase	300	gi|1325320316|ref|XP_023341954.1|beta-galactosidase-like isoform X3 [Eurytemora affinis]	Phasmatodea	XP_023341954	1.54E-30	81.57895	119.013	76	62
Th_d_00019277-RA	Beta-galactosidase-1-like protein 2	1269	gi|1330893432|gb|PNF23708.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattodea	PNF23708	1.4E-154	71.42857	450.284	406	290
Th_d_00019278-RA	beta-galactosidase-1-like protein 2	765	gi|998516749|ref|XP_015519174.1|PREDICTED: beta-galactosidase-1-like protein 2 [Neodiprion lecontei]gi|998516751|ref|XP_015519175.1|PREDICTED: beta-galactosidase-1-like protein 2 [Neodiprion lecontei]	Hymenoptera	XP_015519174, XP_015519175	8E-67	68.24034	224.172	233	159

Table 3.8. Continued.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Th_d_00006652-RA</td><td>Beta-galactosidase</td><td>1249</td><td>gi|321478650|gb|EFX89607.1|hypothetical protein DAPPUDRAFT_303198 [Daphnia pulex]</td><td>Crustacean</td><td>EFX89607</td><td>4.9E-116</td><td>62.83619</td><td>358.607</td><td>409</td><td>257</td></tr>
<tr><td>Th_d_00006654-RA</td><td>beta-galactosidase</td><td>291</td><td>gi|1233169119|ref|XP_022186636.1|beta-galactosidase [Nilaparvata lugens]</td><td>Hemiptera</td><td>XP_022186636</td><td>2.39E-21</td><td>81.35593</td><td>92.8189</td><td>59</td><td>48</td></tr>
<tr><td>Th_d_00006653-RA</td><td>beta-galactosidase</td><td>930</td><td>gi|1233169119|ref|XP_022186636.1|beta-galactosidase [Nilaparvata lugens]</td><td>Hemiptera</td><td>XP_022186636</td><td>2.06E-35</td><td>74.03846</td><td>141.354</td><td>104</td><td>77</td></tr>
<tr><td>Th_d_00038143-RA</td><td>Beta-galactosidase</td><td>1074</td><td>gi|1330925577|gb|PNF39271.1|hypothetical protein B7P43_G16708 [Cryptotermes secundus]</td><td>Blattodea</td><td>PNF39271</td><td>3.78E-50</td><td>56.74419</td><td>181.8</td><td>215</td><td>122</td></tr>
<tr><td>Th_d_00102809-RA</td><td>Beta-galactosidase-1-like protein 2</td><td>358</td><td>gi|1330893429|gb|PNF23705.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]</td><td>Blattodea</td><td>PNF23705</td><td>9.88E-25</td><td>72.72727</td><td>102.834</td><td>77</td><td>56</td></tr>
<tr><td>Th_d_00032922-RA</td><td>Beta-galactosidase-1-like protein 2</td><td>855</td><td>gi|1330893432|gb|PNF23708.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]</td><td>Blattodea</td><td>PNF23708</td><td>2.6E-113</td><td>74.63235</td><td>339.732</td><td>272</td><td>203</td></tr>
</table>

Table 3.8. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00037252-RA	Beta-galactosidase	1416	gi|321478650|gb|EFX89607.1|hypothetical protein DAPPUDRAFT_303198 [Daphnia pulex]	Crustacea	EFX89607	1.1E-116	63.18408	362.459	402	254
Th_d_00052174-RA	beta-galactosidase-1-like protein 2	1143	gi|1000751440|ref|XP_015598574.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751442|ref|XP_015598576.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751444|ref|XP_015598577.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751446|ref|XP_015598578.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751448|ref|XP_015598579.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751450|ref|XP_015598580.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]	Hymenoptera	XP_015598574, XP_015598576, XP_015598577, XP_015598578, XP_015598579, XP_015598580	1.3E-110	65.36313	342.813	358	234

Table 3.8. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00079338-RA	beta-galactosidase-like isoform X1	513	gi|1233169119|ref|XP_022186636.1|beta-galactosidase [Nilaparvata lugens]	Hemiptera	XP_022186636	2.11E-45	76.78571	163.696	112	86

Table 3.9. Myrosinases in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for mirosinases and their blast description.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Th_d_00144418-RA</td><td>myrosinase 1-like</td><td>306</td><td>gi|768427825|ref|XP_011555023.1|PREDICTED: uncharacterized protein LOC105386209 [Plutella xylostella]</td><td>Lepidoptera</td><td>XP_011555023</td><td>1.77E-30</td><td>73.68421</td><td>119.013</td><td>95</td><td>70</td></tr>
<tr><td>Th_d_00030415-RA</td><td>myrosinase 1-like</td><td>1170</td><td>gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]</td><td>Blattodea</td><td>XP_023721112, XP_023721113, PNF19653, PNF19654</td><td>8.9E-118</td><td>65.81197</td><td>358.992</td><td>351</td><td>231</td></tr>
<tr><td>Th_d_00055165-RA</td><td>myrosinase 1-like</td><td>243</td><td>gi|636630776|gb|AIA09348.1|beta-glucosidase [Periplaneta americana]</td><td>Blattodea</td><td>AIA09348</td><td>9.26E-29</td><td>79.22078</td><td>112.464</td><td>77</td><td>61</td></tr>
<tr><td>Th_d_00046945-RA</td><td>myrosinase 1-like</td><td>426</td><td>gi|910313450|ref|XP_013175517.1|PREDICTED: uncharacterized protein LOC106123648 [Papilio xuthus]</td><td>Lepidoptera</td><td>XP_013175517</td><td>6.81E-15</td><td>83.33333</td><td>76.6406</td><td>48</td><td>40</td></tr>
</table>

Table 3.9. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00063063-RA	myrosinase 1-like	636	gi|1114635916|gb|APM84101.1|GH 1 beta-glucosidase [Microcerotermes annandalei]	Blattodea	APM84101	2.94E-30	78.04878	122.094	82	64
Th_d_00012572-RA	myrosinase 1-like	1560	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	6.6E-109	62.6943	340.887	386	242
Th_d_00135824-RA	myrosinase 1-like	321	gi|910324408|ref|XP_013165281.1|PREDICTED: myrosinase 1-like [Papilio xuthus]	Lepidoptera	XP_013165281	2.82E-26	64.15094	107.457	106	68
Th_d_00113027-RA	Myrosinase 1	372	gi|1061486758|gb|ODN03847.1|Myrosinase 1 [Orchesella cincta]	Collembola	ODN03847	3.72E-33	59.67742	127.487	124	74
Th_d_00117962-RA	Myrosinase 1	321	gi|746857492|ref|XP_011059478.1|PREDICTED: myrosinase 1 isoform X2 [Acromyrmex echinatior]	Hymenoptera	XP_011059478	1.61E-33	82.22222	126.716	90	74

Table 3.9. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00105850-RA	Myrosinase 1	246	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	4.92E-33	84.14634	120.168	82	69
Th_d_00058602-RA	myrosinase 1-like	678	gi|269965728|dbj|BAI50023.1|beta-glucosidase [Nasutitermes takasagoensis]	Blattodea	BAI50023	2.49E-89	73.85321	277.33	218	161
Th_d_00056050-RA	Myrosinase 1	798	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	2.9E-104	77.69231	319.316	260	202
Th_d_00093969-RA	myrosinase 1-like	441	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	1.79E-54	75.69444	181.8	144	109
Th_d_00058955-RA	Myrosinase 1	579	gi|574959088|gb|AHG54239.1|beta-glucosidase 3 [Lygus lineolaris]	Hemiptera	AHG54239	1.88E-34	77.17391	132.494	92	71

Table 3.9. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00100160-RA	Myrosinase 1	369	gi|1070157753|ref|XP_018359559.1|PREDICTED: myrosinase 1-like [Trachymyrmex cornetzi]gi|1009403499|gb|KYN23032.1|Myrosinase 1 [Trachymyrmex cornetzi]	Hymenoptera	XP_018359559, KYN23032	7.4E-34	82.6087	128.257	92	76
Th_d_00141241-RA	myrosinase 1-like	411	gi|910324408|ref|XP_013165281.1|PREDICTED: myrosinase 1-like [Papilio xuthus]	Lepidoptera	XP_013165281	1.11E-25	79.72973	107.071	74	59
Th_d_00089076-RA	myrosinase 1-like	366	gi|364023587|gb|AEW46868.1|seminal fluid protein CSSFP018, partial [Chilo suppressalis]	Lepidoptera	AEW46868	2.5E-40	74.5098	139.813	102	76
Th_d_00076743-RA	myrosinase 1-like	450	gi|1080062831|ref|XP_018573476.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573476	3.39E-50	81.95489	172.94	133	109
Th_d_00111777-RA	myrosinase 1-like	192	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	2.21E-13	84.09091	67.781	44	37
Th_d_00074352-RA	myrosinase 1-like	486	gi|913323605|ref|XP_013193075.1|PREDICTED: myrosinase 1-like [Amyelois transitella]	Lepidoptera	XP_013193075	2.95E-12	77.77778	69.707	45	35

Table 3.9. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00064915-RA	myrosinase 1-like	531	gi|1101349013|ref|XP_018900400.1|PREDICTED: myrosinase 1-like isoform X2 [Bemisia tabaci]	Hemiptera	XP_018900400	8.6E-34	79.41176	130.568	102	81
Th_d_00053277-RA	myrosinase 1	786	gi|748995286|gb|AJE75665.1|putative glycosyl hydrolase [Chrysomela lapponica]	Coleoptera	AJE75665	7.01E-56	64.48087	192.2	183	118
Th_d_00074186-RA	myrosinase 1-like	936	gi|930669849|gb|KPJ10712.1|hypothetical protein RR48_07718 [Papilio machaon]	Lepidoptera	KPJ10712	3.09E-24	72.28916	105.145	83	60
Th_d_00126328-RA	myrosinase 1-like	687	gi|768432251|ref|XP_011557441.1|PREDICTED: uncharacterized protein LOC105388261 [Plutella xylostella]	Lepidoptera	XP_011557441	8.38E-35	67.22689	136.346	119	80
Th_d_00052665-RA	myrosinase 1-like	741	gi|1101347123|ref|XP_018899386.1|PREDICTED: myrosinase 1 [Bemisia tabaci]	Hemiptera	XP_018899386	2.86E-69	69.66825	225.713	211	147
Th_d_00141011-RA	myrosinase 1-like	258	gi|910324408|ref|XP_013165281.1|PREDICTED: myrosinase 1-like [Papilio xuthus]	Lepidoptera	XP_013165281	8.52E-34	80.23256	127.872	86	69

Table 3.9. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00043623-RA	myrosinase 1-like	1014	gi|795032132|ref|XP_011863828.1|PREDICTED: myrosinase 1-like [Vollenhovia emeryi]gi|795032135|ref|XP_011863829.1|PREDICTED: myrosinase 1-like [Vollenhovia emeryi]	Hymenoptera	XP_011863828, XP_011863829	7.62E-16	49.12281	84.7297	114	56
Th_d_00014522-RA	Myrosinase 1	1410	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	1.2E-118	62.80788	364.385	406	255
Th_d_00082198-RA	Myrosinase 1	678	gi|636630776|gb|AIA09348.1|beta-glucosidase [Periplaneta americana]	Blattodea	AIA09348	3.66E-08	62.35294	59.6918	85	53
Th_d_00084786-RA	myrosinase 1-like	417	gi|795043525|ref|XP_011867620.1|PREDICTED: myrosinase 1-like [Vollenhovia emeryi]	Hymenoptera	XP_011867620	2.65E-17	87.23404	83.1889	47	41
Th_d_00066036-RA	Myrosinase 1	744	gi|930652287|gb|KPI94627.1|Myrosinase 1 [Papilio xuthus]	Lepidoptera	KPI94627	8.92E-60	56.45161	202.216	248	140

Table 3.9. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00062853-RA	myrosinase 1-like	228	gi|910324408|ref|XP_013165281.1| PREDICTED: myrosinase 1-like [Papilio xuthus]	Lepidoptera	XP_013165281	1.38E-10	78.04878	61.2326	41	32
Th_d_00041776-RA	myrosinase 1-like	840	gi|1080062338|ref|XP_018573207.1|myrosinase 1-like isoform X2 [Anoplophora glabripennis]	Coleoptera	XP_018573207	5.58E-85	64.28571	268.47	252	162
Th_d_00052134-RA	myrosinase 1-like	693	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	6.99E-74	72.9064	231.106	203	148
Th_d_00101375-RA	myrosinase 1-like	543	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.39E-14	86.36364	75.485	44	38
Th_d_00038374-RA	myrosinase 1-like	705	gi|1325349814|ref|XP_023310604.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_023310604	2.9E-21	83.87097	97.8265	62	52
Th_d_00038373-RA	myrosinase 1-like	294	gi|636630776|gb|AIA09348.1|beta-glucosidase [Periplaneta americana]	Blattodea	AIA09348	9.32E-29	86.36364	113.235	88	76

Table 3.9. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00031000-RA	myrosinase 1-like	1119	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	8E-129	67.03601	386.726	361	242
Th_d_00033250-RA	myrosinase 1-like	1290	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	1.1E-160	75.44757	469.929	391	295
Th_d_00032603-RA	myrosinase 1-like	255	gi|939630683|ref|XP_014278711.1|PREDICTED: myrosinase 1-like [Halyomorpha halys]	Hemiptera	XP_014278711	3.67E-13	78.26087	68.9366	46	36
Th_d_00032602-RA	Myrosinase 1	717	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	4.97E-62	75.81699	201.06	153	116

Table 3.9. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00053701-RA	Myrosinase 1	762	gi|364023593|gb|AEW46871.1|seminal fluid protein CSSFP021 [Chilo suppressalis]	Lepidoptera	AEW46871	5.44E-29	84.21053	120.168	76	64
Th_d_00104679-RA	Myrosinase 1	303	gi|930661725|gb|KPJ03612.1|Myrosinase 1 [Papilio xuthus]	Lepidoptera	KPJ03612	2.28E-40	77.22772	147.132	101	78
Th_d_00103233-RA	myrosinase 1-like	341	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	5.14E-26	72.41379	105.531	87	63
Th_d_00032475-RA	myrosinase 1-like	900	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	3.18E-89	64.49275	277.33	276	178
Th_d_00015265-RA	myrosinase 1-like	402	gi|768432251|ref|XP_011557441.1|PREDICTED: uncharacterized protein LOC105388261 [Plutella xylostella]	Lepidoptera	XP_011557441	1.73E-40	65.64885	147.902	131	86
Th_d_00056242-RA	myrosinase 1-like	891	gi|568256136|gb|ETN64735.1|lactase-phlorizin hydrolase [Anopheles darlingi]	Diptera	ETN64735	2.23E-31	53.125	128.642	192	102
Th_d_00068261-RA	myrosinase 1-like	616	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	3.63E-44	67.37589	157.532	141	95

Table 3.9. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00033820-RA	myrosinase 1-like	1359	gi|1121153069|ref|XP_019563383.1|PREDICTED: lactase-phlorizin hydrolase-like [Aedes albopictus]	Diptera	XP_019563383	6.31E-26	69.36937	117.857	111	77
Th_d_00078948-RA	myrosinase 1-like	336	gi|768427825|ref|XP_011555023.1|PREDICTED: uncharacterized protein LOC105386209 [Plutella xylostella]	Lepidoptera	XP_011555023	7.56E-34	75.72816	129.028	103	78
Th_d_00085145-RA	Myrosinase 1	495	gi|1070157753|ref|XP_018359559.1|PREDICTED: myrosinase 1-like [Trachymyrmex cornetzi]gi|1009403499|gb|KYN23032.1|Myrosinase 1 [Trachymyrmex cornetzi]	Hymenoptera	XP_018359559, KYN23032	1.27E-40	71.64179	148.288	134	96
Th_d_00150104-RA	myrosinase 1-like	414	gi|768427825|ref|XP_011555023.1|PREDICTED: uncharacterized protein LOC105386209 [Plutella xylostella]	Lepidoptera	XP_011555023	1.23E-31	60.14493	124.02	138	83
Th_d_00094424-RA	myrosinase 1-like	399	gi|364023587|gb|AEW46868.1|seminal fluid protein CSSFP018, partial [Chilo suppressalis]	Lepidoptera	AEW46868	2.38E-39	74.5098	137.502	102	76
Th_d_00084437-RA	myrosinase 1-like	613	gi|636630776|gb|AIA09348.1|beta-glucosidase [Periplaneta americana]	Blattodea	AIA09348	2.8E-17	56.89655	85.5001	116	66

Table 3.9. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00088179-RA	myrosinase 1	483	gi|574959088|gb|AHG54239.1|beta-glucosidase 3 [Lygus lineolaris]	Hemiptera	AHG54239	9.24E-30	80.51948	118.627	77	62
Th_d_00095744-RA	Myrosinase 1	357	gi|930661725|gb|KPJ03612.1|Myrosinase 1 [Papilio xuthus]	Lepidoptera	KPJ03612	1.02E-40	76.47059	149.058	102	78

Table 3.10. Lytic polysaccharide monooxygenases (LPMOs) in *Thermobia domestica*: Coding sequences in *T. domestica* genome encoding for LPMOs and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00069199-RA	Lytic polysaccharide monooxygenase	789	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.14E-56	77.69231	185.652	130	101
Th_d_00098297-RA	Lytic polysaccharide monooxygenase	645	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2E-86	74.07407	259.996	216	160
Th_d_00119312-RA	Lytic polysaccharide monooxygenase	645	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	6E-109	84.34343	317.005	198	167
Th_d_00120034-RA	Lytic polysaccharide monooxygenase	270	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.99E-30	73.86364	112.464	88	65
Th_d_00110473-RA	Lytic polysaccharide monooxygenase	570	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.61E-78	78.82353	238.81	170	134
Th_d_00083543-RA	Lytic polysaccharide monooxygenase	684	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.46E-96	78.60697	286.189	201	158
Th_d_00052762-RA	Lytic polysaccharide monooxygenase	936	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	4.36E-80	63.56275	247.669	247	157

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00007079-RA	Lytic polysaccharide monooxygenase	697	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.1E-160	100	447.203	213	213
Th_d_00007078-RA	Lytic polysaccharide monooxygenase	645	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	8.5E-109	84.34343	316.62	198	167
Th_d_00007075-RA	Lytic polysaccharide monooxygenase	999	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.1E-77	78.82353	242.276	170	134
Th_d_00007076-RA	Lytic polysaccharide monooxygenase	507	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	3.91E-74	78.39506	226.868	162	127
Th_d_00056739-RA	Lytic polysaccharide monooxygenase	519	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.88E-51	72.51908	168.703	131	95
Th_d_00120998-RA	Lytic polysaccharide monooxygenase	477	gi|926610136|ref|XP_013787168.1|uncharacterized protein LOC106471123 [Limulus polyphemus]	Arthropoda-chelicerata-atlantic horseshoe crab	XP_013787168	5.85E-69	73.41772	213.001	158	116

Table 3.10. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Th_d_00058386-RA	Lytic polysaccharide monooxygenase	936	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	5.5E-102	80.69307	303.523	202	163
Th_d_00114368-RA	Lytic polysaccharide monooxygenase	693	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	4.45E-71	75.5814	221.861	172	130
Th_d_00075908-RA	Lytic polysaccharide monooxygenase	462	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	8.07E-50	73.4375	164.466	128	94
Th_d_00058359-RA	Lytic polysaccharide monooxygenase	657	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	4.37E-93	75.58685	276.944	213	161
Th_d_00056072-RA	Lytic polysaccharide monooxygenase	765	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.04E-73	63.17829	229.565	258	163
Th_d_00064235-RA	Lytic polysaccharide monooxygenase	849	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	4.19E-69	74.86911	218.779	191	143
Th_d_00056727-RA	Lytic polysaccharide monooxygenase	758	gi|1009534315|ref|XP_015904140.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Spider	XP_015904140	5.7E-88	71.56398	266.159	211	151

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00083583-RA	Lytic polysaccharide monooxygenase	865	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	9.69E-67	75.79618	213.001	157	119
Th_d_00003122-RA	Lytic polysaccharide monooxygenase	817	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.3E-61	66.66667	199.134	195	130
Th_d_00023793-RA	Lytic polysaccharide monooxygenase	708	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.9E-101	81.68317	298.901	202	165
Th_d_00023791-RA	Lytic polysaccharide monooxygenase	507	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	3.41E-75	79.01235	229.565	162	128
Th_d_00023792-RA	Lytic polysaccharide monooxygenase	297	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	8.8E-17	77.19298	77.411	57	44
Th_d_00058774-RA	Lytic polysaccharide monooxygenase	882	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.5E-159	100	447.203	213	213
Th_d_00057442-RA	Lytic polysaccharide monooxygenase	895	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	9.18E-91	78.23834	274.248	193	151

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00118174-RA	Lytic polysaccharide monooxygenase	312	gi|1080063909|ref|XP_018574059.1|uncharacterized protein LOC108913075 [Anoplophora glabripennis]	Beetle	XP_018574059	5.39E-15	75	74.7146	52	39
Th_d_00109123-RA	Lytic polysaccharide monooxygenase	633	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.1E-125	99.45946	358.607	185	184
Th_d_00014400-RA	Lytic polysaccharide monooxygenase	834	gi|1009534315|ref|XP_015904140.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Spider	XP_015904140	2.3E-84	73.11321	258.07	212	155
Th_d_00075450-RA	Lytic polysaccharide monooxygenase	870	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	3.78E-96	78.60697	287.73	201	158
Th_d_00072326-RA	Lytic polysaccharide monooxygenase	894	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	5.13E-97	77.83019	290.041	212	165
Th_d_00049111-RA	Lytic polysaccharide monooxygenase	799	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.79E-93	75.9434	279.641	212	161

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00104344-RA	Lytic polysaccharide monooxygenase	460	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.41E-62	76.19048	196.438	147	112
Th_d_00061982-RA	Lytic polysaccharide monooxygenase	850	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.33E-69	75.39267	219.55	191	144
Th_d_00102441-RA	Lytic polysaccharide monooxygenase	460	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.41E-62	76.19048	196.438	147	112
Th_d_00140949-RA	Lytic polysaccharide monooxygenase	402	gi|1067063705|ref|XP_018015775.1|PREDICTED: uncharacterized protein LOC108672588 [Hyalella azteca]	Crustacean	XP_018015775	2.47E-53	74.80315	184.882	127	95
Th_d_00084210-RA	Lytic polysaccharide monooxygenase	758	gi|1009534315|ref|XP_015904140.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Spider	XP_015904140	4.11E-88	72.03791	266.544	211	152
Th_d_00095724-RA	Lytic polysaccharide monooxygenase	645	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.7E-149	100	419.468	214	214

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00064622-RA	Lytic polysaccharide monooxygenase	865	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.16E-93	76.16822	280.796	214	163
Th_d_00158168-RA	Lytic polysaccharide monooxygenase	183	gi|1238888528|ref|XP_022256420.1|uncharacterized protein LOC106472108 [Limulus polyphemus]	Arthropoda-chelicerata-atlantic horseshoe crab	XP_022256420	3.97E-14	87.87879	68.9366	33	29
Th_d_00047815-RA	Lytic polysaccharide monooxygenase	384	gi|1061482501|gb|ODM99783.1|hypothetical protein Ocin01_06883 [Orchesella cincta]	Springtail	ODM99783	3.04E-44	75	149.443	124	93
Th_d_00054826-RA	Lytic polysaccharide monooxygenase	630	gi|1061482501|gb|ODM99783.1|hypothetical protein Ocin01_06883 [Orchesella cincta]		ODM99783	1.12E-42	74.79675	148.673	123	92
Th_d_00108306-RA	Lytic polysaccharide monooxygenase	555	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.32E-72	72.72727	223.787	176	128

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00075363-RA	Lytic polysaccharide monooxygenase	786	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.31E-55	60.65574	182.956	183	111
Th_d_00058638-RA	Lytic polysaccharide monooxygenase	882	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.7E-158	100	445.662	213	213
Th_d_00076106-RA	Lytic polysaccharide monooxygenase	894	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.04E-97	77.83019	290.812	212	165
Th_d_00092785-RA	Lytic polysaccharide monooxygenase	459	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2E-15	74.19355	75.8702	62	46
Th_d_00068335-RA	Lytic polysaccharide monooxygenase	777	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	5.2E-103	80.69307	303.908	202	163
Th_d_00112031-RA	Lytic polysaccharide monooxygenase	570	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.98E-76	77.64706	233.802	170	132
Th_d_00032707-RA	Lytic polysaccharide monooxygenase	861	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.15E-73	75.40984	229.95	183	138

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00044687-RA	Lytic polysaccharide monooxygenase	600	gi|1355978318|gb|PRD27424.1|hypothetical protein NCL1_35386 [Nephila clavipes]	Spider	PRD27424	2.93E-50	73.68421	165.622	133	98
Th_d_00068347-RA	Lytic polysaccharide monooxygenase	862	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	3.95E-91	78.23834	275.018	193	151
Th_d_00071731-RA	Lytic polysaccharide monooxygenase	821	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.6E-159	100	447.203	213	213
Th_d_00095665-RA	Lytic polysaccharide monooxygenase	615	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	5.17E-77	75.97765	235.728	179	136
Th_d_00085144-RA	Lytic polysaccharide monooxygenase	480	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.56E-65	74.375	204.527	160	119
Th_d_00071020-RA	Lytic polysaccharide monooxygenase	837	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	7.67E-49	71.9697	166.777	132	95
Th_d_00037467-RA	Lytic polysaccharide monooxygenase	1077	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.23E-67	75.75758	217.624	165	125

Table 3.10. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Th_d_00037466-RA	Lytic polysaccharide monooxygenase	567	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.16E-75	75	230.335	180	135

Table 3.11. Endoglucanases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for endoglucanases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00020160-RA	endoglucanase 13-like	513	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	2.62E-22	86.88525	96.2857	61	53
Lep_00115213-RA	endoglucanase 7-like	561	gi|1228018665|ref|XP_021941322.1|endoglucanase 7-like, partial [Zootermopsis nevadensis]	Blattodea	XP_021941322	3.32E-82	83.01887	247.284	159	132
Lep_00014665-RA	putative endo-beta-1,4-glucanase NkEG3	618	gi|197691949|dbj|BAG70026.1|endo-beta-1,4-glucanase, partial [Salganea esakii]	Blattodea	BAG70026	8.42E-17	80	83.1889	60	48
Lep_00071461-RA	Endoglucanase A	692	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	4.2E-107	84.40367	321.627	218	184
Lep_00045067-RA	putative endo-beta-1,4-glucanase OfEG3	726	gi|24940551|emb|CAD54729.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	1.47E-69	70.61856	225.713	194	137
Lep_00016310-RA	endoglucanase 7-like	519	gi|1228018665|ref|XP_021941322.1|endoglucanase 7-like, partial [Zootermopsis nevadensis]	Blattodea	XP_021941322	2.5E-76	82.19178	231.876	146	120

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00016312-RA	Endoglucanase 1	837	gi|157313367|gb|ABV32557.1|cellulase [Teleogryllus emma]gi|167859905|gb|ACA04897.1|cellulase [Teleogryllus emma]	Orthoptera	ABV32557, ACA04897	1.08E-42	69.7479	157.532	119	83
Lep_00105874-RA	AF220593_1beta-1,4-glucanase 1	522	gi|7546876|gb|AAF63724.1|AF220593_1beta-1,4-glucanase 1, partial [Mastotermes darwiniensis]	Blattodea	AAF63724	9.2E-66	77.5	207.608	160	124
Lep_00112904-RA	endoglucanase 11-like	528	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	1.56E-22	86.53846	94.7449	52	45
Lep_00052326-RA	endo-beta-1,4-glucanase 1	1168	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	3.1E-100	63.10976	310.457	328	207

Table 3.11. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00013083-RA	Endoglucanase E-4 precursor, putative	1834	gi|1339056265|ref|XP_023716596.1|uncharacterized protein LOC111869358 [Cryptotermes secundus]gi|1339056267|ref|XP_023716597.1|uncharacterized protein LOC111869358 [Cryptotermes secundus]gi|1330895261|gb|PNF24408.1|hypothetical protein B7P43_G09674 [Cryptotermes secundus]	Blattodea	XP_023716596, XP_023716597, PNF24408	1.42E-40	78.47222	159.844	144	113
Lep_00013082-RA	endoglucanase E-4-like	1314	gi|646711640|gb|KDR16731.1|Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR16731	5.94E-83	73.21429	268.47	224	164
Lep_00125033-RA	endoglucanase E-4-like	300	gi|380452610|gb|AFD33365.1|endoglucanase [Macrotermes barneyi]	Blattodea	AFD33365	1.16E-20	83.33333	90.8929	60	50
Lep_00014982-RA	putative endo-beta-1,4-glucanase OfEG1	819	gi|44885838|dbj|BAD12008.1|putative endo-beta-1,4-glucanase OfEG1, partial [Odontotermes formosanus]	Blattodea	BAD12008	1.65E-78	70.48458	246.899	227	160

Table 3.11. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00051684-RA	AF220596_1beta-1,4-endoglucanase 1	294	gi|8886827|gb|AAF80584.1|AF220596_1beta-1,4-endoglucanase 1 [Panesthia cribrata]	Blattodea	AAF80584	2.32E-37	78.49462	135.961	93	73
Lep_00095655-RA	AF220593_1beta-1,4-glucanase 1	627	gi|7546876|gb|AAF63724.1|AF220593_1beta-1,4-glucanase 1, partial [Mastotermes darwiniensis]	Blattodea	AAF63724	4.39E-83	75.12438	253.447	201	151
Lep_00104311-RA	AF220597_1beta-1,4-endoglucanase 2	480	gi|197691959|dbj|BAG70031.1|endo-beta-1,4-glucanase, partial [Panesthia angustipennis]	Blattodea	BAG70031	8.35E-57	73.54839	186.037	155	114
Lep_00142934-RA	Endoglucanase E-4 precursor, putative	225	gi|1022772340|gb|KZS16514.1|Endoglucanase [Daphnia magna]	Crustacea	KZS16514	4.63E-19	67.56757	85.5001	74	50
Lep_00095117-RA	endo-beta-1,4-glucanase 1	531	gi|44885846|dbj|BAD12012.1|putative endo-beta-1,4-glucanase SmEG1, partial [Sinocapritermes mushae]	Blattodea	BAD12012	3.51E-49	67.04545	168.318	176	118
Lep_00100398-RA	Endoglucanase F	504	gi|646711640|gb|KDR16731.1|Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR16731	3.35E-25	87.32394	106.301	71	62

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00053282-RA	Endoglucanase E-4 precursor, putative	279	gi|1022772340|gb|KZS16514.1|Endoglucanase [Daphnia magna]	Crustacea	KZS16514	9.09E-24	74.24242	99.7525	66	49
Lep_00029521-RA	AF220597_1beta-1,4-endoglucanase 2	1666	gi|197691959|dbj|BAG70031.1|endo-beta-1,4-glucanase, partial [Panesthia angustipennis]	Blattodea	BAG70031	1.35E-53	64.79592	190.66	196	127
Lep_00107688-RA	endoglucanase 15-like	399	gi|1078570857|gb|AOV94253.1|cellulase [Antipaluria urichi]	Embioptera	AOV94253	1.5E-24	79.03226	103.219	62	49
Lep_00003639-RA	endoglucanase E-4-like	282	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	2E-23	79.03226	93.9745	62	49
Lep_00032008-RA	1,4-alpha-glucan-branching enzyme	876	gi|1339092712|ref|XP_023707569.1|1,4-alpha-glucan-branching enzyme [Cryptotermes secundus]gi|1330914862|gb|PNF33674.1|1,4-alpha-glucan-branching enzyme [Cryptotermes secundus]	Blattodea	XP_023707569, PNF33674	1.19E-73	87.41259	244.588	143	125

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00032826-RA	1,4-alpha-glucan-branching enzyme	1779	gi|1339092712|ref|XP_023707569.1|1,4-alpha-glucan-branching enzyme [Cryptotermes secundus]gi|1330914862|gb|PNF33674.1|1,4-alpha-glucan-branching enzyme [Cryptotermes secundus]	Blattodea	XP_023707569, PNF33674	0	89.77636	541.576	313	281
Lep_00026154-RA	endo-beta-1,4-glucanase 1	456	gi|44885830|dbj|BAD12004.1|putative endo-beta-1,4-glucanase HsEG4, partial [Hodotermopsis sjostedti]	Blattodea	BAD12004	2.41E-44	82.6087	155.221	115	95
Lep_00026155-RA	endo-beta-1,4-glucanase 1	339	gi|521313323|gb|AGP76424.1|endo-beta-1,4-glucanase 1 [Pericapritermes sp. PpEG1]	Blattodea	AGP76424	6.06E-15	79.62963	75.485	54	43
Lep_00038056-RA	AF220597_1beta-1,4-endoglucanase 2	693	gi|8886829|gb|AAF80585.1|AF220597_1beta-1,4-endoglucanase 2 [Panesthia cribrata]	Blattodea	AAF80585	1.69E-43	55.55556	157.532	207	115
Lep_00006772-RA	AF220594_1beta-1,4-glucanase 2	372	gi|646700778|gb|KDR10771.1|hypothetical protein L798_15338, partial [Zootermopsis nevadensis]	Blattodea	KDR10771	1.61E-53	86.84211	176.407	114	99

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00008343-RA	endoglucanase E-4-like	291	gi|1000740642|ref|XP_015592927.1|PREDICTED: endoglucanase E-4-like [Cephus cinctus]gi|1000740644|ref|XP_015592928.1|PREDICTED: endoglucanase E-4-like [Cephus cinctus]gi|1000740646|ref|XP_015592929.1|PREDICTED: endoglucanase E-4-like [Cephus cinctus]	Hymenoptera	XP_015592927, XP_015592928, XP_015592929	4.68E-30	74.74747	116.701	99	74
Lep_00008346-RA	Endoglucanase E-4 precursor, putative	519	gi|985425232|ref|XP_015378087.1|PREDICTED: endoglucanase 4-like [Diuraphis noxia]	Hemiptera	XP_015378087	2.46E-40	74.07407	144.05	135	100
Lep_00008347-RA	endoglucanase E-4-like	720	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	1.3E-21	70.83333	94.3597	72	51
Lep_00008344-RA	Endoglucanase E-4	360	gi|992052372|gb|AMH40395.1|glycoside hydrolase family 9 [Timema cristinae]	Phasmatodea	AMH40395	6.7E-13	62.5	69.707	80	50
Lep_00044891-RA	glycoside hydrolase family 9	1194	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	0	79.36508	528.094	378	300

Table 3.11. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00020161-RA	glycoside hydrolase family 9	1350	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	3E-159	78.15126	463.381	357	279
Lep_00054377-RA	glycoside hydrolase family 9	1059	gi|992052372|gb|AMH40395.1|glycoside hydrolase family 9 [Timema cristinae]	Phasmatodea	AMH40395	2.25E-66	56.07143	222.246	280	157
Lep_00014666-RA	glycoside hydrolase family 9	463	gi|985388348|ref|XP_015378728.1|PREDICTED: endoglucanase E-4-like [Diuraphis noxia]	Hemiptera	XP_015378728	6.6E-26	67.77778	102.449	90	61
Lep_00071754-RA	glycoside hydrolase family 9	838	gi|992051866|gb|AMH40374.1|glycoside hydrolase family 9 [Peruphasma schultei]	Phasmatodea	AMH40374	3.65E-80	75.58685	254.603	213	161
Lep_00016311-RA	glycoside hydrolase family 9	834	gi|13095576|gb|AAK12339.1|cellulase [Coptotermes acinaciformis]	Blattodea	AAK12339	9.36E-57	71.07843	194.126	204	145
Lep_00100613-RA	glycoside hydrolase family 9	438	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	1.65E-37	78.49462	134.806	93	73
Lep_00034744-RA	glycoside hydrolase family 9	1742	gi|992051866|gb|AMH40374.1|glycoside hydrolase family 9 [Peruphasma schultei]	Phasmatodea	AMH40374	4.5E-149	63.34661	442.58	502	318

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00014981-RA	glycoside hydrolase family 9	1605	gi|1078570851|gb|AOV94250.1|cellulase [Antipaluria urichi]	Embioptera	AOV94250	1.37E-59	71.42857	209.534	168	120
Lep_00014983-RA	glycoside hydrolase family 9	558	gi|1080053223|ref|XP_018568196.1|uncharacterized protein LOC108908596 isoform X3 [Anoplophora glabripennis]gi|1325348755|ref|XP_023310347.1|uncharacterized protein LOC108908596 isoform X4 [Anoplophora glabripennis]	Coleoptera	XP_018568196, XP_023310347	1.48E-23	59.29204	102.449	113	67
Lep_00071132-RA	glycoside hydrolase family 9	685	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	1E-105	83.48624	317.775	218	182
Lep_00051683-RA	glycoside hydrolase family 9	244	gi|992052372|gb|AMH40395.1|glycoside hydrolase family 9 [Timema cristinae]	Phasmatodea	AMH40395	4.36E-25	79.72973	102.064	74	59
Lep_00038648-RA	glycoside hydrolase family 9	580	gi|992052101|gb|AMH40383.1|glycoside hydrolase family 9 [Ramulus artemis]	Phasmatodea	AMH40383	1.45E-23	76.47059	102.449	68	52
Lep_00060202-RA	glycoside hydrolase family 9	852	gi|13095576|gb|AAK12339.1|cellulase [Coptotermes acinaciformis]	Blattodea	AAK12339	3.03E-92	76.19048	285.804	210	160

Table 3.11. Continued.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Lep_00022398-RA</td><td>glycoside hydrolase family 9</td><td>1453</td><td>gi|992052326|gb|AMH40392.1|glycoside hydrolase family 9 [Timema cristinae]</td><td>Phasmatodea</td><td>AMH40392</td><td>1.5E-177</td><td>74.01392</td><td>511.146</td><td>431</td><td>319</td></tr>
<tr><td>Lep_00036184-RA</td><td>glycoside hydrolase family 9</td><td>1518</td><td>gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]</td><td>Blattodea</td><td>XP_021941324</td><td>0</td><td>79.35035</td><td>570.852</td><td>431</td><td>342</td></tr>
<tr><td>Lep_00026826-RA</td><td>glycoside hydrolase family 9</td><td>1573</td><td>gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]</td><td>Blattodea</td><td>XP_021941324</td><td>0</td><td>78.19026</td><td>578.941</td><td>431</td><td>337</td></tr>
<tr><td>Lep_00061557-RA</td><td>glycoside hydrolase family 9</td><td>1095</td><td>gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]</td><td>Blattodea</td><td>PNF24409</td><td>3.9E-35</td><td>75</td><td>135.961</td><td>100</td><td>75</td></tr>
<tr><td>Lep_00034832-RA</td><td>glycoside hydrolase family 9</td><td>537</td><td>gi|7546876|gb|AAF63724.1|AF220593_1beta-1,4-glucanase 1, partial [Mastotermes darwiniensis]</td><td>Blattodea</td><td>AAF63724</td><td>6.15E-47</td><td>71.65354</td><td>159.844</td><td>127</td><td>91</td></tr>
<tr><td>Lep_00078550-RA</td><td>glycoside hydrolase family 9</td><td>753</td><td>gi|992051866|gb|AMH40374.1|glycoside hydrolase family 9 [Peruphasma schultei]</td><td>Phasmatodea</td><td>AMH40374</td><td>4.17E-82</td><td>70.12987</td><td>258.455</td><td>231</td><td>162</td></tr>
</table>

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00081122-RA	glycoside hydrolase family 9	693	gi|992051866|gb|AMH40374.1|glycoside hydrolase family 9 [Peruphasma schultei]	Phasmatodea	AMH40374	9.56E-79	72.12389	249.21	226	163
Lep_00135466-RA	glycoside hydrolase family 9	285	gi|992051866|gb|AMH40374.1|glycoside hydrolase family 9 [Peruphasma schultei]	Phasmatodea	AMH40374	1.97E-25	80.95238	103.605	63	51
Lep_00127920-RA	glycoside hydrolase family 9	231	gi|1022761055|gb|KZS07093.1|Endoglucanase [Daphnia magna]	Crustacea	KZS07093	1.77E-22	75.67568	95.1301	74	56
Lep_00096399-RA	glycoside hydrolase family 9	1155	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	4.5E-49	74.49664	172.94	149	111
Lep_00003635-RA	glycoside hydrolase family 9	975	gi|1228018667|ref|XP_021941324.1|uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattodea	XP_021941324	8.2E-126	75.48387	373.244	310	234
Lep_00003637-RA	glycoside hydrolase family 9	312	gi|1022772340|gb|KZS16514.1|Endoglucanase [Daphnia magna]	Crustacea	KZS16514	7.52E-21	77.9661	91.6633	59	46
Lep_00003640-RA	glycoside hydrolase family 9	234	gi|695941540|gb|AIT38270.1|endo-beta-1,4-glucanase, partial [Pseudohelice subquadrata]	Arthropoda-crab	AIT38270	1.39E-08	82.85714	51.9878	35	29

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00026156-RA	glycoside hydrolase family 9	1451	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phasmatodea	AMH40360	6.24E-65	75	222.246	172	129
Lep_00038055-RA	glycoside hydrolase family 9	852	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	1.47E-88	75.90909	270.781	220	167
Lep_00094505-RA	glycoside hydrolase family 9	912	gi|1330895262|gb|PNF24409.1|hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattodea	PNF24409	2.13E-51	72.90323	176.407	155	113
Lep_00006775-RA	glycoside hydrolase family 9	1508	gi|24940551|emb|CAD54729.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54729	0	72.24576	537.339	472	341
Lep_00006773-RA	glycoside hydrolase family 9	441	gi|992051639|gb|AMH40360.1|glycoside hydrolase family 9 [Aretaon asperrimus]	Phasmatodea	AMH40360	1.54E-18	84.61538	62.7734	39	33
Lep_00006774-RA	glycoside hydrolase family 9	1143	gi|992051683|gb|AMH40363.1|glycoside hydrolase family 9 [Extatosoma tiaratum]	Phasmatodea	AMH40363	1.99E-20	82.69231	98.5969	52	43
Lep_00042852-RA	glycoside hydrolase family 9	834	gi|24940553|emb|CAD54730.1|beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD54730	5.1E-109	76.57993	328.561	269	206

Table 3.11. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00042851-RA	glycoside hydrolase family 9	444	gi|1078570851|gb|AOV94250.1|cellulase [Antipaluria urichi]	Crustacea	AOV94250	5.06E-12	66.66667	68.5514	57	38

Table 3.12. Glucosidases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for glucosidases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0009587 8-RA	glucosidase 2 subunit beta	633	gi|1339080613|ref|XP_023728054.1|glucosidase 2 subunit beta [Cryptotermes secundus]gi|1330927816|gb|PNF40060.1|Glucosidase 2 subunit beta [Cryptotermes secundus]	Blattodea	XP_023728054, PNF40060	2.28E-36	84.61538	139.428	78	66
Lep_0005546 2-RA	Glucosidase 2 subunit beta	708	gi|1339077375|ref|XP_023726071.1|glucosidase 2 subunit beta isoform X1 [Cryptotermes secundus]	Blattodea	XP_023726071	4.64E-26	51.63399	105.916	153	79
Lep_0005546 0-RA	glucosidase 2 subunit beta-like	213	gi|1228015572|ref|XP_021939758.1|glucosidase 2 subunit beta [Zootermopsis nevadensis]	Blattodea	XP_021939758	5.55E-08	80.64516	51.9878	31	25
Lep_0009887 5-RA	protein-glucosylgalactosylhydroxylysine glucosidase	669	gi|512925622|ref|XP_004930884.1|protein-glucosylgalactosylhydroxylysine glucosidase [Bombyx mori]	Lepidoptera	XP_004930884	3.41E-23	60.83333	103.605	120	73

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00052604-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	957	gi|303324839|pdb|3AHZ|AChain A, Crystal Structure Of Beta-Glucosidase From Termite Neotermes Koshunensis In Complex With Trisgi|393715252|pdb|3VIF|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Gluconolactonegi|393715253|pdb|3VIG|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With 1-deoxynojirimycingi|393715254|pdb|3VIH|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glycerolgi|393715255|pdb|3VII|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Bis-tris	Blattodea	3AHZ_A, 3VIF_A, 3VIG_A, 3VIH_A, 3VII_A	1.6E-122	75.52448	365.925	286	216

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00036380-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	1194	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	6.7E-106	62.95265	328.946	359	226

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00004872-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	1844	gi|303324839|pdb|3AHZ|AChain A, Crystal Structure Of Beta-Glucosidase From Termite Neotermes Koshunensis In Complex With Trisgi|393715252|pdb|3VIF|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Gluconolactonegi|393715253|pdb|3VIG|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With 1-deoxynojirimycingi|393715254|pdb|3VIH|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glycerolgi|393715255|pdb|3VII|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Bis-tris	Blattodea	3AHZ_A, 3VIF_A, 3VIG_A, 3VIH_A, 3VII_A	9E-147	71.8498	368.622	313	225

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00004873-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	375	gi|1339073035|ref|XP_023723774.1|myrosinase 1-like [Cryptotermes secundus]gi|1330878957|gb|PNF16841.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023723774, PNF16841	2.43E-51	78.81356	174.866	118	93

Table 3.12. Continued.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Lep_00021305-RA</td><td>AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product</td><td>843</td><td>gi|303324839|pdb|3AHZ|AChain A, Crystal Structure Of Beta-Glucosidase From Termite Neotermes Koshunensis In Complex With Trisgi|393715252|pdb|3VIF|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Gluconolactonegi|393715253|pdb|3VIG|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With 1-deoxynojirimycingi|393715254|pdb|3VIH|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glycerolgi|393715255|pdb|3VII|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Bis-tris</td><td>Blattodea</td><td>3AHZ_A, 3VIF_A, 3VIG_A, 3VIH_A, 3VII_A</td><td>8.07E-94</td><td>73.30508</td><td>290.812</td><td>236</td><td>173</td></tr>
</table>

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00016622-RA	glucosidase 2 subunit beta	831	gi|1339080613|ref|XP_023728054.1|glucosidase 2 subunit beta [Cryptotermes secundus]gi|1330927816|gb|PNF40060.1|Glucosidase 2 subunit beta [Cryptotermes secundus]	Blattodea	XP_023728054, PNF40060	1.66E-24	66.35945	108.612	217	144
Lep_00016621-RA	glucosidase 2 subunit beta	459	gi|1228013272|ref|XP_021938600.1|glucosidase 2 subunit beta [Zootermopsis nevadensis]	Blattodea	XP_021938600	1.04E-17	80.95238	85.1149	63	51
Lep_00016623-RA	glucosidase 2 subunit beta	678	gi|1339080613|ref|XP_023728054.1|glucosidase 2 subunit beta [Cryptotermes secundus]gi|1330927816|gb|PNF40060.1|Glucosidase 2 subunit beta [Cryptotermes secundus]	Blattodea	XP_023728054, PNF40060	1.1E-34	60.71429	135.191	140	85

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00061060-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	926	gi|1339073027|ref|XP_023723770.1|myrosinase 1-like isoform X1 [Cryptotermes secundus]	Blattodea	XP_023723770	3.28E-82	67.5	262.692	240	162
Lep_00022254-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	291	gi|393715257|pdb|3VIK|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Cellobiosegi|393715258|pdb|3VIL|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Salicin	Blattodea	3VIK_A, 3VIL_A	5.81E-37	83.90805	135.191	87	73

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00004127-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	837	gi|1339073029|ref|XP_023723771.1|myrosinase 1-like isoform X2 [Cryptotermes secundus]	Blattodea	XP_023723771	1.61E-89	62.72401	280.411	279	175

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00004131-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	1032	gi|303324839|pdb|3AHZ|AChain A, Crystal Structure Of Beta-Glucosidase From Termite Neotermes Koshunensis In Complex With Trisgi|393715252|pdb|3VIF|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Gluconolactonegi|393715253|pdb|3VIG|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With 1-deoxynojirimycingi|393715254|pdb|3VIH|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glycerolgi|393715255|pdb|3VII|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Bis-tris	Blattodea	3AHZ_A, 3VIF_A, 3VIG_A, 3VIH_A, 3VII_A	2.7E-102	63.69427	315.079	314	200

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_000045500-RA	bile acid beta-glucosidase	618	gi|568254064|gb|ETN63073.1|bile acid beta-glucosidase [Anopheles darlingi]	Diptera	ETN63073	1.75E-39	70.55215	149.828	163	115
Lep_000077160-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	792	gi|303324839|pdb|3AHZ|AChain A, Crystal Structure Of Beta-Glucosidase From Termite Neotermes Koshunensis In Complex With Trisgi|393715252|pdb|3VIF|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Gluconolactonegi|393715253|pdb|3VIG|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With 1-deoxynojirimycingi|393715254|pdb|3VIH|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glycerolgi|393715255|pdb|3VII|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Bis-tris	Blattodea	3AHZ_A, 3VIF_A, 3VIG_A, 3VIH_A, 3VII_A	4.2E-76	66.39344	244.588	244	162

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0013137 3-RA	glucosidase 2 subunit beta	165	gi|1228013272|ref|XP_021938600.1|glucosidase 2 subunit beta [Zootermopsis nevadensis]	Blattodea	XP_021938600	3.42E-17	80.8510 6	78.951 8	47	38
Lep_0007334 5-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	897	gi|380452608|gb|AFD33364.1|beta-glucosidase [Macrotermes barneyi]	Blattodea	AFD33364	3.41E-87	71.4912 3	275.01 8	228	163

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0002084 6-RA	mannosyl-oligosaccharide glucosidase isoform X2	2442	gi|1339080389|ref|XP_023727931.1|mannosyl-oligosaccharide glucosidase [Cryptotermes secundus]gi|1339080391|ref|XP_023727934.1|mannosyl-oligosaccharide glucosidase [Cryptotermes secundus]gi|1339080393|ref|XP_023727935.1|mannosyl-oligosaccharide glucosidase [Cryptotermes secundus]gi|1330927938|gb|PNF40173.1|Mannosyl-oligosaccharide glucosidase GCS1 [Cryptotermes secundus]gi|1330927939|gb|PNF40174.1|Mannosyl-oligosaccharide glucosidase GCS1 [Cryptotermes secundus]	Blattodea	XP_023727931, XP_023727934, XP_023727935, PNF40173, PNF40174	0	76.94175	1047.73	824	634

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0005205 1-RA	AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With A New Glucopyranosidic Product	954	gi|303324839|pdb|3AHZ|AChain A, Crystal Structure Of Beta-Glucosidase From Termite Neotermes Koshunensis In Complex With Trisgi|393715252|pdb|3VIF|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Gluconolactonegi|393715253|pdb|3VIG|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With 1-deoxynojirimycingi|393715254|pdb|3VIH|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Glycerolgi|393715255|pdb|3VII|AChain A, Crystal Structure Of Beta-glucosidase From Termite Neotermes Koshunensis In Complex With Bis-tris	Blattodea	3AHZ_A, 3VIF_A, 3VIG_A, 3VIH_A, 3VII_A	1.5E-122	75.5244 8	365.925	286	216

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0004125 2-RA	lysosomal alpha-glucosidase-like	1458	gi|769856746|ref|XP_011639981.1|PREDICTED: lysosomal alpha-glucosidase-like [Pogonomyrmex barbatus]gi|769856748|ref|XP_011639982.1|PREDICTED: lysosomal alpha-glucosidase-like [Pogonomyrmex barbatus]gi|769856750|ref|XP_011639983.1|PREDICTED: lysosomal alpha-glucosidase-like [Pogonomyrmex barbatus]gi|769856752|ref|XP_011639984.1|PREDICTED: lysosomal alpha-glucosidase-like [Pogonomyrmex barbatus]gi|769856754|ref|XP_011639985.1|PREDICTED: lysosomal alpha-glucosidase-like [Pogonomyrmex barbatus]	Hymenoptera	XP_011639981, XP_011639982, XP_011639983, XP_011639984, XP_011639985	6.37E-37	87.8048 8	151.754	82	72
Lep_0008481 0-RA	uncharacterized family 31 glucosidase KIAA1161-like	1857	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	0	71.7625 9	665.611	556	399

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0011960 7-RA	uncharacterized family 31 glucosidase KIAA1161	370	gi|1316140871|ref|XP_023215385.1|uncharacterized protein LOC111618155 isoform X1 [Centruroides sculpturatus]	Arthropoda	XP_023215385	1.31E-15	75.8064 5	77.7962	62	47
Lep_0015999 4-RA	uncharacterized family 31 glucosidase KIAA1161-like	1107	gi|1067066097|ref|XP_018017064.1|PREDICTED: uncharacterized family 31 glucosidase KIAA1161-like [Hyalella azteca]	Crustacea	XP_018017064	2.01E-86	55.6149 7	278.1	374	208
Lep_0010032 1-RA	lysosomal alpha-glucosidase-like	2250	gi|1228004020|ref|XP_021933895.1|probable maltase-glucoamylase 2 [Zootermopsis nevadensis]gi|1228004022|ref|XP_021933896.1|probable maltase-glucoamylase 2 [Zootermopsis nevadensis]	Blattodea	XP_021933895, XP_021933896	5.4E-96	50.1906	325.094	787	395
Lep_0006089 9-RA	uncharacterized family 31 glucosidase KIAA1161-like	501	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	6.85E-45	82.7586 2	161.384	116	96

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00060900-RA	uncharacterized family 31 glucosidase KIAA1161-like	405	gi|1101398930|ref|XP_018910486.1|PREDICTED: uncharacterized family 31 glucosidase KIAA1161-like [Bemisia tabaci]	Hemiptera	XP_018910486	3.4E-19	80.95238	88.5817	63	51
Lep_00061350-RA	Lysosomal alpha-glucosidase	1163	gi|321476729|gb|EFX87689.1|hypothetical protein DAPPUDRAFT_312137 [Daphnia pulex]	Crustacea	EFX87689	1.3E-101	71.33956	325.865	321	229
Lep_00077538-RA	uncharacterized family 31 glucosidase KIAA1161-like	2040	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	0	65.625	646.736	608	399
Lep_00039192-RA	Lysosomal alpha-glucosidase	861	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	1.73E-83	65.23297	278.87	279	182
Lep_00044754-RA	Lysosomal alpha-glucosidase	717	gi|1022768804|gb|KZS13425.1|Lysosomal alpha-glucosidase [Daphnia magna]	Crustacea	KZS13425	1.16E-84	73.60671	276.944	224	165

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0004475 5-RA	lysosomal alpha-glucosidase-like	561	gi|1067072134|ref|XP_018019496.1|PREDICTED: LOW QUALITY PROTEIN: probable maltase-glucoamylase 2, partial [Hyalella azteca]	Crustacea	XP_018019496	1.61E-11	49.6296 3	68.551 4	135	67
Lep_0008251 8-RA	Neutral alpha-glucosidase AB	381	gi|242019253|ref|XP_002430076.1|Neutral alpha-glucosidase AB precursor, putative [Pediculus humanus corporis]gi|212515157|gb|EEB17338.1|Neutral alpha-glucosidase AB precursor, putative [Pediculus humanus corporis]	Phthiraptera	XP_002430076, EEB17338	1.31E-60	81.7460 3	203.37 1	126	103
Lep_0008251 9-RA	neutral alpha-glucosidase AB	270	gi|321473948|gb|EFX84914.1|hypothetical protein DAPPUDRAFT_300811 [Daphnia pulex]	Crustacea	EFX84914	6.12E-24	67.5	100.13 8	80	54
Lep_0010666 8-RA	uncharacterized family 31 glucosidase KIAA1161	1305	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	2.4E-129	57.9957 4	392.11 9	469	272
Lep_0005768 7-RA	lysosomal alpha-glucosidase-like	1101	gi|1233197498|ref|XP_022198397.1|lysosomal alpha-glucosidase-like, partial [Nilaparvata lugens]	Hemiptera	XP_022198397	2.66E-64	54.3543 5	226.48 3	333	181

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00131971-RA	lysosomal alpha-glucosidase-like	291	gi|1238867836|ref|XP_022249217.1|lysosomal alpha-glucosidase-like [Limulus polyphemus]	Arthropoda	XP_022249217	3.33E-32	68	114.39	100	68
Lep_00133481-RA	lysosomal alpha-glucosidase-like	303	gi|1080050544|ref|XP_018566730.1|lysosomal alpha-glucosidase-like [Anoplophora glabripennis]gi|1080050546|ref|XP_018566731.1|lysosomal alpha-glucosidase-like [Anoplophora glabripennis]	Coleoptera	XP_018566730, XP_018566731	1.2E-50	92	176.792	100	92
Lep_00057296-RA	Lysosomal alpha-glucosidase	1214	gi|321476729|gb|EFX87689.1|hypothetical protein DAPPUDRAFT_312137 [Daphnia pulex]	Crustacea	EFX87689	1.3E-102	69.52663	329.331	338	235
Lep_00106368-RA	Neutral alpha-glucosidase AB	390	gi|1227991792|ref|XP_021927675.1|neutral alpha-glucosidase AB isoform X3 [Zootermopsis nevadensis]gi|646709565|gb|KDR15365.1|Neutral alpha-glucosidase AB [Zootermopsis nevadensis]	Blattodea	XP_021927675, KDR15365	3.47E-43	93.02326	156.762	86	80
Lep_00027362-RA	lysosomal alpha-glucosidase-like	402	gi|952513507|gb|KRT80074.1|glycoside hydrolase, partial [Oryctes borbonicus]	Coleoptera	KRT80074	2.41E-38	83.33333	138.272	84	70

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0002736 1-RA	Lysosomal alpha-glucosidase	891	gi|1022773228|gb|KZS17340.1|Lysosomal alpha-glucosidase [Daphnia magna]	Crustacea	KZS17340	5.29E-39	92.5	152.14	80	74
Lep_0016092 0-RA	uncharacterized family 31 glucosidase KIAA1161-like	186	gi|1228003699|ref|XP_021933735.1|uncharacterized family 31 glucosidase KIAA1161-like [Zootermopsis nevadensis]	Blattodea	XP_021933735	2.82E-22	83.0508 5	93.974 5	59	49
Lep_0008518 2-RA	uncharacterized family 31 glucosidase KIAA1161 isoform X1	2409	gi|242004884|ref|XP_002423306.1|conserved hypothetical protein [Pediculus humanus corporis]gi|212506315|gb|EEB10568.1|conserved hypothetical protein [Pediculus humanus corporis]	Phthiraptera	XP_002423306, EEB10568	0	72.9938 3	707.98 3	648	473
Lep_0007601 3-RA	uncharacterized family 31 glucosidase KIAA1161-like	180	gi|1101361447|ref|XP_018913028.1|PREDICTED: uncharacterized family 31 glucosidase KIAA1161-like [Bemisia tabaci]	Hemiptera	XP_018913028	2.48E-20	83.0188 7	88.196 5	53	44

Table 3.12. Continued.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Lep_00076014-RA</td><td>uncharacterized family 31 glucosidase KIAA1161-like</td><td>405</td><td>gi|1101398930|ref|XP_018910486.1|PREDICTED: uncharacterized family 31 glucosidase KIAA1161-like [Bemisia tabaci]</td><td>Hemiptera</td><td>XP_018910486</td><td>3.47E-19</td><td>80.95238</td><td>88.5817</td><td>63</td><td>51</td></tr>
<tr><td>Lep_00089518-RA</td><td>uncharacterized family 31 glucosidase KIAA1161-like isoform X1</td><td>1839</td><td>gi|1330911882|gb|PNF32421.1|hypothetical protein B7P43_G04891 [Cryptotermes secundus]</td><td>Blattodea</td><td>PNF32421</td><td>3.8E-172</td><td>60</td><td>509.22</td><td>585</td><td>351</td></tr>
<tr><td>Lep_00071777-RA</td><td>Uncharacterized family 31 glucosidase KIAA1161</td><td>2259</td><td>gi|1228003699|ref|XP_021933735.1|uncharacterized family 31 glucosidase KIAA1161-like [Zootermopsis nevadensis]</td><td>Blattodea</td><td>XP_021933735</td><td>2.5E-150</td><td>67.98144</td><td>458.373</td><td>431</td><td>293</td></tr>
<tr><td>Lep_00161546-RA</td><td>uncharacterized family 31 glucosidase KIAA1161-like</td><td>1098</td><td>gi|1067066097|ref|XP_018017064.1|PREDICTED: uncharacterized family 31 glucosidase KIAA1161-like [Hyalella azteca]</td><td>Crustacea</td><td>XP_018017064</td><td>3.56E-86</td><td>55.82656</td><td>277.715</td><td>369</td><td>206</td></tr>
<tr><td>Lep_00074517-RA</td><td>uncharacterized family 31 glucosidase KIAA1161</td><td>156</td><td>gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]</td><td>Blattodea</td><td>AIA09350</td><td>3.74E-14</td><td>81.25</td><td>70.0922</td><td>48</td><td>39</td></tr>
</table>

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0004284 6-RA	neutral alpha-glucosidase AB	789	gi|1339041694|ref|XP_023710162.1|neutral alpha-glucosidase AB [Cryptotermes secundus]gi|1330933432|gb|PNF42738.1|Neutral alpha-glucosidase AB [Cryptotermes secundus]	Blattodea	XP_023710162, PNF42738	2.4E-54	70.45455	194.126	176	124
Lep_0009708 4-RA	Lysosomal alpha-glucosidase	717	gi|321476729|gb|EFX87689.1|hypothetical protein DAPPUDRAFT_312137 [Daphnia pulex]	Crustacea	EFX87689	1.87E-84	72.76786	274.633	224	163
Lep_0013618 4-RA	uncharacterized family 31 glucosidase KIAA1161-like	1527	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	2E-163	71.60194	481.871	412	295
Lep_0007907 3-RA	uncharacterized family 31 glucosidase KIAA1161-like	336	gi|1228003699|ref|XP_021933735.1|uncharacterized family 31 glucosidase KIAA1161-like [Zootermopsis nevadensis]	Blattodea	XP_021933735	2.34E-21	78.78788	93.5893	66	52
Lep_0007907 4-RA	uncharacterized family 31 glucosidase KIAA1161-like	315	gi|1330911882|gb|PNF32421.1|hypothetical protein B7P43_G04891 [Cryptotermes secundus]	Blattodea	PNF32421	8.01E-18	77.27273	83.1889	66	51

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00128081-RA	lysosomal alpha-glucosidase-like	444	gi|1080050544|ref|XP_018566730.1|lysosomal alpha-glucosidase-like [Anoplophora glabripennis]gi|1080050546|ref|XP_018566731.1|lysosomal alpha-glucosidase-like [Anoplophora glabripennis]	Coleoptera	XP_018566730, XP_018566731	1.1E-45	91.39785	164.851	93	85
Lep_00119531-RA	uncharacterized family 31 glucosidase KIAA1161-like isoform X2	368	gi|1316140871|ref|XP_023215385.1|uncharacterized protein LOC111618155 isoform X1 [Centruroides sculpturatus]	Arthropoda	XP_023215385	1.27E-15	75.80645	77.7962	62	47
Lep_00126255-RA	uncharacterized family 31 glucosidase KIAA1161-like	285	gi|1228003699|ref|XP_021933735.1|uncharacterized family 31 glucosidase KIAA1161-like [Zootermopsis nevadensis]	Blattodea	XP_021933735	1.99E-22	83.33333	95.9005	60	50
Lep_00113055-RA	alpha-glucosidase isoform X2	393	gi|1227106192|gb|OXU28826.1|hypothetical protein TSAR_016821 [Trichomalopsis sarcophagae]	Hymenoptera	OXU28826	5.61E-25	67.1875	70.4774	64	43
Lep_00068885-RA	lysosomal alpha-glucosidase-like isoform X1	900	gi|1022742921|gb|KZR99685.1|Lysosomal alpha-glucosidase, partial [Daphnia magna]	Crustacea	KZR99685	1.55E-53	52.68456	186.037	298	157

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00076643-RA	lysosomal alpha-glucosidase-like	1260	gi|1330905640|gb|PNF29605.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905643|gb|PNF29608.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905644|gb|PNF29609.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]	Blattodea	PNF29605, PNF29608, PNF29609	3.67E-92	66.89189	291.197	296	198
Lep_00135198-RA	uncharacterized family 31 glucosidase KIAA1161-like	597	gi|646703149|gb|KDR11965.1|hypothetical protein L798_13618, partial [Zootermopsis nevadensis]	Blattodea	KDR11965	1.14E-67	71.42857	221.09	189	135
Lep_00136521-RA	uncharacterized family 31 glucosidase KIAA1161-like	666	gi|646703149|gb|KDR11965.1|hypothetical protein L798_13618, partial [Zootermopsis nevadensis]	Blattodea	KDR11965	2.69E-75	70.95238	241.891	210	149
Lep_00094746-RA	uncharacterized family 31 glucosidase KIAA1161-like	1476	gi|636630780|gb|AIA09350.1|alpha-glucosidase family 31, partial [Periplaneta americana]	Blattodea	AIA09350	1.7E-162	71.35922	479.174	412	294

Table 3.12. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00102111-RA	uncharacterized family 31 glucosidase KIAA1161-like	519	gi|646703149|gb|KDR11965.1|hypothetical protein L798_13618, partial [Zootermopsis nevadensis]	Blattodea	KDR11965	7.91E-49	80.17241	170.629	116	93
Lep_00054791-RA	lysosomal alpha-glucosidase-like	291	gi|1330905640|gb|PNF29605.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905643|gb|PNF29608.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905644|gb|PNF29609.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]	Blattodea	PNF29605, PNF29608, PNF29609	4.26E-44	89.77273	153.68	88	79
Lep_00054790-RA	lysosomal alpha-glucosidase-like	957	gi|1330905640|gb|PNF29605.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905643|gb|PNF29608.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]gi|1330905644|gb|PNF29609.1|hypothetical protein B7P43_G01635 [Cryptotermes secundus]	Blattodea	PNF29605, PNF29608, PNF29609	6.66E-66	62.1875	219.935	320	199

Table 3.12. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0011734 9-RA	lysosomal alpha-glucosidase-like	459	gi|952513507|gb|KRT80074.1|glycoside hydrolase, partial [Oryctes borbonicus]	Coleoptera	KRT80074	1.23E-38	83.33333	139.813	84	70
Lep_0011244 5-RA	Neutral alpha-glucosidase AB	597	gi|1227991790|ref|XP_021927674.1|neutral alpha-glucosidase AB isoform X2 [Zootermopsis nevadensis]	Blattodea	XP_021927674	2.41E-35	83.52941	137.502	85	71
Lep_0001146 6-RA	lysosomal alpha-glucosidase-like	324	gi|1022768805|gb|KZS13426.1|Uncharacterized protein APZ42_021387 [Daphnia magna]	Crustacea	KZS13426	7.46E-17	76.5625	80.4925	64	49

Table 3.13. β-1,3-glucanases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for β-1,3-glucanases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00027153-RA	beta-1,3-glucan-binding protein precursor	7782	gi|1339068387|ref|XP_023721291.1|uncharacterized protein LOC111872033 isoform X1 [Cryptotermes secundus]	Blattodea	XP_023721291	0	53.22396	1382.08	2621	1395
Lep_00050787-RA	Beta-1,3-glucan-binding protein	1287	gi|149777474|gb|ABR28480.1|beta-1,3(4)-glucanase LIC1, partial [Periplaneta americana]	Blattodea	ABR28480	7.09E-42	68.38235	157.147	136	93
Lep_00039162-RA	Beta-1,3-glucan-binding protein	389	gi|506968285|gb|AGM32486.1|gram negative bacteria binding protein 2, partial [Coptotermes formosanus]	Blattodea	AGM32486	3.49E-30	67.22689	115.161	119	80
Lep_00010402-RA	beta-1,3-glucan-binding protein precursor	4791	gi|1339068387|ref|XP_023721291.1|uncharacterized protein LOC111872033 isoform X1 [Cryptotermes secundus]	Blattodea	XP_023721291	0	62.48331	1239.17	1498	936

Table 3.13. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00052055-RA	Beta-1,3-glucan-binding protein	975	gi|1227967090|ref|XP_021914900.1|beta-1,3-glucan-binding protein-like [Zootermopsis nevadensis]	Blattodea	XP_021914900	5.18E-25	74.71264	109.383	87	65
Lep_00039163-RA	beta-1,3-glucan-binding protein-like	393	gi|546684219|gb|ERL93924.1|hypothetical protein D910_11210 [Dendroctonus ponderosae]	Coleoptera	ERL93924	4E-11	53.84615	65.0846	91	49
Lep_00045091-RA	beta-1,3-glucan-binding protein-like	393	gi|546684203|gb|ERL93908.1|hypothetical protein D910_11194, partial [Dendroctonus ponderosae]	Coleoptera	ERL93908	3.77E-11	84.21053	65.4698	38	32
Lep_00036322-RA	beta-1,3-glucan-binding protein-like	783	gi|1022754729|gb|KZS02483.1|Beta-1,3-glucan-binding protein [Daphnia magna]	Crustacea	KZS02483	3.91E-15	45	80.1073	140	63

Table 3.13. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00071308-RA	Beta-1,3-glucan-binding protein	952	gi|321457960|gb|EFX69036.1|hypothetical protein DAPPUDRAFT_203138 [Daphnia pulex]	Crustacea	EFX69036	3.84E-74	69.63563	237.654	247	172
Lep_00120361-RA	Beta-1,3-glucan-binding protein	420	gi|1022753497|gb|KZS02148.1|Beta-1,3-glucan-binding protein, partial [Daphnia magna]	Crustacea	KZS02148	6.93E-07	66.66667	51.6026	45	30
Lep_00045092-RA	Beta-1,3-glucan-binding protein	837	gi|321461266|gb|EFX72300.1|hypothetical protein DAPPUDRAFT_201160 [Daphnia pulex]	Crustacea	EFX72300	4.45E-91	69.35484	279.256	248	172
Lep_00031099-RA	beta-1,3-glucan-binding protein-like	1094	gi|321457960|gb|EFX69036.1|hypothetical protein DAPPUDRAFT_203138 [Daphnia pulex]	Crustacea	EFX69036	3.02E-52	55.88235	182.57	238	133
Lep_00029894-RA	beta-1,3-glucan-binding protein-like	321	gi|1121175018|ref|XP_019532058.1|PREDICTED: beta-1,3-glucan-binding protein [Aedes albopictus]	Diptera	XP_019532058	2.14E-08	60.34483	56.6102	58	35

Table 3.13. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00029895-RA	beta-1,3-glucan-binding protein-like	534	gi|820836955|ref|XP_003690362.2|PREDICTED: beta-1,3-glucan-binding protein 1 [Apis florea]	Hymenoptera	XP_003690362	5.8E-13	66.66667	72.4034	75	50
Lep_00029896-RA	beta-1,3-glucan-binding protein-like	921	gi|914569839|gb|KOB73371.1| Beta-1,3-glucanase, partial [Operophtera brumata]	Lepidoptera	KOB73371	4.41E-11	54.90196	67.781	102	56
Lep_00028853-RA	beta-1,3-glucan-binding protein-like	974	gi|208972541|gb|ACI32831.1|beta-1,3-glucanase [Colias eurytheme]	Lepidoptera	ACI32831	9.7E-30	61.90476	122.479	126	78

Table 3.14. Mannanases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for mannanases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00137319-RA	Endo-beta-1,4-mannanase	483	gi|1022765559|gb|KZS10754.1|Endo-beta-1,4-mannanase [Daphnia magna]	Crustacea	KZS10754	3.07E-51	71.63121	173.326	141	101
Lep_00047049-RA	endo-beta-1,4-mannanase	1575	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	2.49E-24	68.47826	110.923	92	63
Lep_00042299-RA	Endo-beta-1,4-mannanase	1022	gi|1022765557|gb|KZS10752.1|Endo-beta-1,4-mannanase [Daphnia magna]	Crustacea	KZS10752	3.33E-92	66.66667	285.419	288	192
Lep_00015029-RA	Endo-beta-1,4-mannanase	1044	gi|1022765557|gb|KZS10752.1|Endo-beta-1,4-mannanase [Daphnia magna]	Crustacea	KZS10752	6.22E-77	59.60265	246.899	302	180
Lep_00012530-RA	Endo-beta-1,4-mannanase	1041	gi|1022765557|gb|KZS10752.1|Endo-beta-1,4-mannanase [Daphnia magna]	Crustacea	KZS10752	9E-119	70.12579	353.599	318	223
Lep_00057963-RA	Endo-beta-1,4-mannanase	996	gi|1022765557|gb|KZS10752.1|Endo-beta-1,4-mannanase [Daphnia magna]	Crustacea	KZS10752	1.88E-44	60.98901	162.155	182	111

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00015051-RA	endo-beta-1,4-mannanase	1047	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	7.01E-71	60.4811	231.491	291	176
Lep_00015050-RA	endo-beta-1,4-mannanase	502	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	7.81E-30	69.09091	117.857	110	76
Lep_00015049-RA	endo-beta-1,4-mannanase	1005	gi|321460555|gb|EFX71596.1|endo-beta-1,4-mannanase [Daphnia pulex]	Crustacea	EFX71596	4.75E-63	54.41696	210.69	283	154
Lep_00122231-RA	Beta-mannosidase	297	gi|1339072353|ref|XP_023723402.1|beta-mannosidase isoform X2 [Cryptotermes secundus]gi|1339072355|ref|XP_023723403.1|beta-mannosidase isoform X2 [Cryptotermes secundus]gi|1330879524|gb|PNF17252.1|Beta-mannosidase [Cryptotermes secundus]	Blattodea	XP_023723402, XP_023723403, PNF17252	6.21E-13	61.29032	68.9366	93	57

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00067016-RA	beta-mannosidase isoform X2	636	gi|815919755|ref|XP_012244839.1|beta-mannosidase [Bombus impatiens]	Hymenoptera	XP_012244839	2.47E-47	61.16505	172.555	206	126
Lep_00065194-RA	beta-mannosidase precursor, putative	795	gi|998509058|ref|XP_015514937.1|PREDICTED: beta-mannosidase [Neodiprion lecontei]	Hymenoptera	XP_015514937	2.23E-26	83.58209	114.39	67	56
Lep_00114547-RA	Beta-mannosidase	465	gi|1061485315|gb|ODN02454.1|Beta-mannosidase [Orchesella cincta]	Collembola	ODN02454	1.81E-34	74.28571	133.265	105	78
Lep_00101167-RA	beta-mannosidase	546	gi|1279738204|ref|XP_022913917.1|beta-mannosidase [Onthophagus taurus]	Coleoptera	XP_022913917	4.68E-36	63.94558	139.043	147	94
Lep_00015028-RA	Mannan endo-1,4-beta-mannosidase	744	gi|1061469313|gb|ODM89178.1|Mannan endo-1,4-beta-mannosidase, partial [Orchesella cincta]	Collembola	ODM89178	7.01E-10	78.04878	63.5438	41	32

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00113998-RA	beta-mannosidase, putative	585	gi|1325354705|ref|XP_023311948.1|beta-mannosidase-like [Anoplophora glabripennis]	Coleoptera	XP_023311948	5.87E-14	64.70588	73.1738	102	66
Lep_00003857-RA	beta-mannosidase	1761	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	1.6E-146	60.3352	451.055	537	324

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_000038 54-RA	beta-mannosidase	819	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	8.09 E-68	70.33493	231.491	209	147
Lep_000038 56-RA	Beta-mannosidase-like protein	372	gi|1192752750|ref|XP_015929642.2|beta-mannosidase [Parasteatoda tepidariorum]	Arachnida	XP_015929642	4.61 E-27	83.82353	110.538	68	57
Lep_000765 39-RA	Beta-mannosidase	651	gi|1339072353|ref|XP_023723402.1|beta-mannosidase isoform X2 [Cryptotermes secundus]gi|1339072355|ref|XP_023723403.1|beta-mannosidase isoform X2 [Cryptotermes secundus]gi|1330879524|gb|PNF17252.1|Beta-mannosidase [Cryptotermes secundus]	Blattodea	XP_023723402, XP_023723403, PNF17252	4.1E-10	73.46939	65.4698	49	36

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_000038836-RA	beta-mannosidase isoform X1	576	gi|998509058|ref|XP_015514937.1|PREDICTED: beta-mannosidase [Neodiprion lecontei]	Hymenoptera	XP_015514937	5.68E-30	78.31325	121.709	83	65
Lep_00124705-RA	beta-mannosidase	261	gi|1325354705|ref|XP_023311948.1|beta-mannosidase-like [Anoplophora glabripennis]	Coleoptera	XP_023311948	6.52E-13	74.60317	66.6254	63	47
Lep_00070663-RA	Beta-mannosidase	533	gi|1061485315|gb|ODN02454.1|Beta-mannosidase [Orchesella cincta]	Collembola	ODN02454	3.15E-32	75.72816	127.872	103	78
Lep_000406649-RA	Beta-mannosidase	909	gi|1339072353|ref|XP_023723402.1|beta-mannosidase isoform X2 [Cryptotermes secundus]gi|1339072355|ref|XP_023723403.1|beta-mannosidase isoform X2 [Cryptotermes secundus]gi|1330879524|gb|PNF17252.1|Beta-mannosidase [Cryptotermes secundus]	Blattodea	XP_023723402, XP_023723403, PNF17252	2.4E-110	75.09294	346.28	269	202
Lep_000406650-RA	beta-mannosidase	639	gi|1035608711|ref|XP_016911936.1|PREDICTED: beta-mannosidase isoform X2 [Apis cerana]	Hymenoptera	XP_016911936	1.03E-40	72.86822	150.984	129	94

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0001253 1-RA	Mannan endo-1,4-beta-mannosidase	375	gi|1061469313|gb|ODM89178.1|Mannan endo-1,4-beta-mannosidase, partial [Orchesella cincta]	Collembola	ODM89178	6.21 E-28	75.94937	107.457	79	60
Lep_0000536 1-RA	beta-mannosidase	1290	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	2.1E-164	74.07407	491.115	405	300

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_000005360-RA	beta-mannosidase	867	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	1.76E-38	58.03571	150.214	224	130
Lep_000016184-RA	Beta-mannosidase	564	gi|675378716|gb|KFM71618.1|Beta-mannosidase, partial [Stegodyphus mimosarum]	Arachnida	KFM71618	4.76E-74	75	239.965	176	132
Lep_000016185-RA	beta-mannosidase	546	gi|1279738204|ref|XP_022913917.1|beta-mannosidase [Onthophagus taurus]	Coleoptera	XP_022913917	5.38E-32	62.41135	127.102	141	88
Lep_000016183-RA	beta-mannosidase	1071	gi|926610764|ref|XP_013790438.1|beta-mannosidase-like [Limulus polyphemus]	Arthropoda	XP_013790438	8.68E-54	59.18367	196.438	245	145

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00053636-RA	beta-mannosidase	1464	gi|1227973104|ref|XP_021917998.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973106|ref|XP_021917999.1|beta-mannosidase [Zootermopsis nevadensis]gi|1227973108|ref|XP_021918000.1|beta-mannosidase [Zootermopsis nevadensis]gi|646717625|gb|KDR20416.1|Beta-mannosidase [Zootermopsis nevadensis]	Blattodea	XP_021917998, XP_021917999, XP_021918000, KDR20416	3.03E-86	75.77093	289.271	227	172
Lep_00102654-RA	alpha-mannosidase 2	381	gi|1330899240|gb|PNF26420.1|hypothetical protein B7P43_G16606, partial [Cryptotermes secundus]	Blattodea	PNF26420	1.93E-15	76.19048	77.411	63	48
Lep_00022533-RA	Lysosomal alpha-mannosidase	1704	gi|1339064620|ref|XP_023715606.1|lysosomal alpha-mannosidase isoform X2 [Cryptotermes secundus]gi|1330931495|gb|PNF42229.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715606, PNF42229	1.1E-117	63.84298	378.637	484	309

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0002253 2-RA	lysosomal alpha-mannosidase isoform X1	588	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	6.28E-35	63.30935	136.346	139	88
Lep_0006144 5-RA	mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X2	1386	gi|1227996329|ref|XP_021929980.1|mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X2 [Zootermopsis nevadensis]	Blattodea	XP_021929980	0	93.10345	534.643	319	297
Lep_0000169 3-RA	lysosomal alpha-mannosidase-like	1128	gi|985414995|ref|XP_015372896.1|PREDICTED: lysosomal alpha-mannosidase-like [Diuraphis noxia]	Hemiptera	XP_015372896	8.4E-79	65.78947	266.929	304	200

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00001694-RA	lysosomal alpha-mannosidase-like	1206	gi|1233202594|ref|XP_022200324.1|lysosomal alpha-mannosidase-like [Nilaparvata lugens]	Hemiptera	XP_022200324	1.81E-44	63.84977	171.4	213	136
Lep_00001695-RA	lysosomal alpha-mannosidase-like	1440	gi|662196129|ref|XP_008471080.1|PREDICTED: lysosomal alpha-mannosidase-like [Diaphorina citri]	Hemiptera	XP_008471080	2.07E-58	86.29032	204.142	124	107
Lep_00001697-RA	lysosomal alpha-mannosidase (mannosidase alpha class 2b member 1)	438	gi|668446735|gb|KFB36711.1|lysosomal alpha-mannosidase (mannosidase alpha class 2b member 1) [Anopheles sinensis]	Diptera	KFB36711	3.39E-16	86.95652	80.4925	46	40

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0000169 2-RA	lysosomal alpha-mannosidase isoform X1	351	gi|817191866|ref|XP_012271344.1|lysosomal alpha-mannosidase isoform X1 [Orussus abietinus]	Hymenoptera	XP_012271344	1.93E-06	60.31746	51.2174	63	38
Lep_0000169 6-RA	lysosomal alpha-mannosidase-like	345	gi|1229882973|ref|XP_022161630.1|lysosomal alpha-mannosidase-like [Myzus persicae]	Hemiptera	XP_022161630	3.83E-05	75	47.7506	40	30
Lep_0006766 3-RA	lysosomal alpha-mannosidase-like	591	gi|1048052005|ref|XP_017460755.1|PREDICTED: lysosomal alpha-mannosidase-like, partial [Rhagoletis zephyria]	Diptera	XP_017460755	3.85E-35	72.17391	133.65	115	83
Lep_0006766 4-RA	lysosomal alpha-mannosidase-like	318	gi|1133436203|ref|XP_019874318.1|PREDICTED: lysosomal alpha-mannosidase-like [Aethina tumida]	Coleoptera	XP_019874318	1.81E-05	64.58333	48.1358	48	31
Lep_0000152 8-RA	lysosomal alpha-mannosidase-like	312	gi|1080066641|ref|XP_018575534.1|lysosomal alpha-mannosidase isoform X2 [Anoplophora glabripennis]	Coleoptera	XP_018575534	6.07E-10	73.4375	60.8474	64	47

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0011751 6-RA	mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X1	1014	gi|1227996327|ref|XP_021929979.1|mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X1 [Zootermopsis nevadensis]	Blattodea	XP_021929979	1.9E-123	76.94805	374.785	308	237

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_000920000-RA	alpha-mannosidase 2	735	gi|1227998244|ref|XP_021930961.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998246|ref|XP_021930962.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998248|ref|XP_021930963.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998250|ref|XP_021930964.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998252|ref|XP_021930965.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998254|ref|XP_021930966.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|646706309|gb|KDR13614.1|Alpha-mannosidase 2 [Zootermopsis nevadensis]	Blattodea	XP_021930961, XP_021930962, XP_021930963, XP_021930964, XP_021930965, XP_021930966, KDR13614	8.41E-30	58.41121	123.635	214	125

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0001493 2-RA	lysosomal alpha-mannosidase isoform X1	2208	gi|1339064620|ref|XP_023715606.1|lysosomal alpha-mannosidase isoform X2 [Cryptotermes secundus]gi|1330931495|gb|PNF42229.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715606, PNF42229	3E-139	59.7756 4	441.039	624	373
Lep_0001493 3-RA	Lysosomal alpha-mannosidase	490	gi|759077107|ref|XP_011348312.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Ooceraea biroi]gi|607353605|gb|EZA48346.1|Lysosomal alpha-mannosidase [Ooceraea biroi]	Hymenoptera	XP_011348312, EZA48346	3.4E-20	66.9354 8	92.8189	124	83
Lep_0014866 4-RA	mannosyl-oligosaccharide alpha-1,2-mannosidase IA-like	735	gi|1227996327|ref|XP_021929979.1|mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X1 [Zootermopsis nevadensis]	Blattodea	XP_021929979	1.21E-62	70.6976 7	213.001	215	152

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_000050917-RA	lysosomal alpha-mannosidase	516	gi|766932221|ref|XP_011498322.1|PREDICTED: lysosomal alpha-mannosidase [Ceratosolen solmsi marchali]	Hymenoptera	XP_011498322	9.21E-32	88.73239	126.331	71	63
Lep_000050916-RA	Lysosomal alpha-mannosidase	666	gi|1059382461|ref|XP_017786052.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Nicrophorus vespilloides]	Coleoptera	XP_017786052	6.15E-19	74.32432	91.2781	74	55
Lep_000116113-RA	alpha-mannosidase 2	315	gi|1339052711|ref|XP_023714691.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052713|ref|XP_023714692.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052715|ref|XP_023714693.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052717|ref|XP_023714694.1|alpha-mannosidase 2 [Cryptotermes secundus]	Blattodea	XP_023714691, XP_023714692, XP_023714693, XP_023714694	1.37E-14	73.68421	73.9442	57	42
Lep_000021753-RA	lysosomal alpha-mannosidase	927	gi|1000733841|ref|XP_015589368.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Cephus cinctus]	Hymenoptera	XP_015589368	3.01E-48	62.84585	179.104	253	159

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00021751-RA	lysosomal alpha-mannosidase	459	gi|985414995|ref|XP_015372896.1|PREDICTED: lysosomal alpha-mannosidase-like [Diuraphis noxia]	Hemiptera	XP_015372896	1.11E-28	57.97101	116.701	138	80
Lep_00039375-RA	ER degradation-enhancing alpha-mannosidase-like protein 2	1389	gi|1339057577|ref|XP_023717292.1|ER degradation-enhancing alpha-mannosidase-like protein 2 [Cryptotermes secundus]gi|1330893438|gb|PNF23714.1|ER degradation-enhancing alpha-mannosidase-like protein 2 [Cryptotermes secundus]	Blattodea	XP_023717292, PNF23714	0	91.20879	814.683	455	415
Lep_00030657-RA	lysosomal alpha-mannosidase-like	546	gi|321449456|gb|EFX61897.1|hypothetical protein DAPPUDRAFT_337859, partial [Daphnia pulex]	Crustacea	EFX61897	8.22E-79	80.9816	254.603	163	132
Lep_00030656-RA	lysosomal alpha-mannosidase isoform X1	150	gi|751453449|ref|XP_011181145.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Zeugodacus cucurbitae]	Diptera	XP_011181145	1.99E-17	92.30769	79.337	39	36

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0003065 8-RA	lysosomal alpha-mannosidase isoform X2	585	gi|1321314822|ref|XP_023299647.1|lysosomal alpha-mannosidase isoform X2 [Lucilia cuprina]	Diptera	XP_023 299647	4.77E -46	79.8319 3	168. 318	119	95
Lep_0008673 7-RA	alpha-mannosidase 2-like isoform X1	720	gi|1325316915|ref|XP_023340186.1|alpha-mannosidase 2-like isoform X1 [Eurytemora affinis]	Phasmatodea	XP_023 340186	1.4E-05	45.6521 7	52.3 73	138	63
Lep_0007609 7-RA	ER degradation-enhancing alpha-mannosidase-like protein 3	813	gi|1339066460|ref|XP_023725333.1|ER degradation-enhancing alpha-mannosidase-like protein 3 [Cryptotermes secundus]	Blattodea	XP_023 725333	1.85E -35	72.1739 1	140. 969	115	83
Lep_0007272 4-RA	alpha-mannosidase 2	720	gi|1330899240|gb|PNF26420.1|hypothetical protein B7P43_G16606, partial [Cryptotermes secundus]	Blattodea	PNF264 20	1.8E-13	75.4098 4	75.4 85	61	46

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00004798-RA	lysosomal alpha-mannosidase	625	gi|170061340|ref|XP_001866193.1|lysosomal alpha-mannosidase [Culex quinquefasciatus]gi|167879594|gb|EDS42977.1|lysosomal alpha-mannosidase [Culex quinquefasciatus]	Diptera	XP_001866193, EDS42977	2.19E-19	60.36036	92.0485	111	67
Lep_00084806-RA	ER degradation-enhancing alpha-mannosidase-like protein 3	417	gi|1227972165|ref|XP_021917507.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X1 [Zootermopsis nevadensis]gi|1227972167|ref|XP_021917508.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X1 [Zootermopsis nevadensis]gi|646718115|gb|KDR20717.1|ER degradation-enhancing alpha-mannosidase-like 3 [Zootermopsis nevadensis]	Blattodea	XP_021917507, XP_021917508, KDR20717	2.05E-31	63.63636	123.635	132	84
Lep_00029715-RA	Endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase	516	gi|1339074955|ref|XP_023724784.1|endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase isoform X2 [Cryptotermes secundus]	Blattodea	XP_023724784	3.36E-75	86.2069	240.35	145	125

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_000029716-RA	endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase	966	gi|1339074955|ref|XP_023724784.1|endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase isoform X2 [Cryptotermes secundus]	Blattodea	XP_023724784	1.66E-88	88.82353	280.796	170	151
Lep_000040597-RA	mannosyl-oligosaccharide 1,2-alpha-mannosidase IA isoform X1	1698	gi|1227996327|ref|XP_021929979.1|mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X1 [Zootermopsis nevadensis]	Blattodea	XP_021929979	6.9E-142	64.87603	431.409	484	314
Lep_000074953-RA	lysosomal alpha-mannosidase isoform X1	549	gi|675365773|gb|KFM58675.1|Lysosomal alpha-mannosidase, partial [Stegodyphus mimosarum]	Arachnida	KFM58675	1.12E-61	79.54545	206.068	132	105
Lep_000008575-RA	Lysosomal alpha-mannosidase	1755	gi|1061483689|gb|ODN00899.1|Lysosomal alpha-mannosidase [Orchesella cincta]	Collembola	ODN00899	2.2E-100	56.91244	332.798	434	247
Lep_000008576-RA	lysosomal alpha-mannosidase-like	1730	gi|321463090|gb|EFX74108.1|hypothetical protein DAPPUDRAFT_252189 [Daphnia pulex]	Crustacea	EFX74108	5.11E-78	61.5942	271.166	276	170

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_000087043-RA	lysosomal alpha-mannosidase isoform X1	822	gi|332029389|gb|EGI69344.1|Lysosomal alpha-mannosidase [Acromyrmex echinatior]	Hymenoptera	EGI69344	1.02E-67	64.36782	229.18	261	168
Lep_000077283-RA	lysosomal alpha-mannosidase-like	339	gi|1199407503|ref|XP_021199415.1|lysosomal alpha-mannosidase-like isoform X4 [Helicoverpa armigera]gi|1199407505|ref|XP_021199416.1|lysosomal alpha-mannosidase-like isoform X5 [Helicoverpa armigera]	Lepidoptera	XP_021199415, XP_021199416	1.9E-21	58.8785	93.9745	107	63
Lep_000012477-RA	Lysosomal alpha-mannosidase	1635	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	1.9E-124	66.66667	395.586	435	290
Lep_000012478-RA	lysosomal alpha-mannosidase isoform X2	345	gi|1351647487|ref|XP_024086113.1|lysosomal alpha-mannosidase-like [Cimex lectularius]	Hemiptera	XP_024086113	1.85E-18	58.65385	85.5001	104	61

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00012479-RA	lysosomal alpha-mannosidase-like	828	gi|1351647487|ref|XP_024086113.1|lysosomal alpha-mannosidase-like [Cimex lectularius]	Hemiptera	XP_024086113	7.12E-07	75.40984	56.9954	61	46
Lep_00051181-RA	ER degradation-enhancing alpha-mannosidase-like protein 1	1284	gi|1227982315|ref|XP_021922781.1|ER degradation-enhancing alpha-mannosidase-like protein 1 [Zootermopsis nevadensis]gi|646713767|gb|KDR17988.1|ER degradation-enhancing alpha-mannosidase-like 1 [Zootermopsis nevadensis]	Blattodea	XP_021922781, KDR17988	5.05E-91	91.875	291.967	160	147
Lep_00062728-RA	ER degradation-enhancing alpha-mannosidase-like protein 3	327	gi|751228804|ref|XP_011167829.1|PREDICTED: ER degradation-enhancing alpha-mannosidase-like protein 3, partial [Solenopsis invicta]	Hymenoptera	XP_011167829	1.73E-51	97.67442	164.466	86	84

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00062729-RA	ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X2	390	gi|1227972169|ref|XP_021917509.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X2 [Zootermopsis nevadensis]gi|1227972171|ref|XP_021917511.1|ER degradation-enhancing alpha-mannosidase-like protein 3 isoform X2 [Zootermopsis nevadensis]	Blattodea	XP_021917509, XP_021917511	2.03E-23	100	100.523	43	43
Lep_00064816-RA	lysosomal alpha-mannosidase isoform X2	588	gi|1096257085|gb|APA33853.1|seminal fluid protein [Nilaparvata lugens]	Hemiptera	APA33853	3.78E-12	75	70.4774	56	42

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00059016-RA	alpha-mannosidase 2	252	gi|1339052711|ref|XP_023714691.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052713|ref|XP_023714692.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052715|ref|XP_023714693.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052717|ref|XP_023714694.1|alpha-mannosidase 2 [Cryptotermes secundus]	Blattodea	XP_023714691, XP_023714692, XP_023714693, XP_023714694	7.81E-17	73.77049	79.337	61	45
Lep_00060769-RA	lysosomal alpha-mannosidase isoform X2	899	gi|817191868|ref|XP_012271345.1|lysosomal alpha-mannosidase isoform X2 [Orussus abietinus]	Hymenoptera	XP_012271345	2.4E-69	69.02985	238.039	268	185
Lep_00039690-RA	lysosomal alpha-mannosidase isoform X1	778	gi|1351647487|ref|XP_024086113.1|lysosomal alpha-mannosidase-like [Cimex lectularius]	Hemiptera	XP_024086113	5.36E-38	60.52632	147.517	190	115

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0039689-RA	lysosomal alpha-mannosidase	867	gi|1059382461|ref|XP_017786052.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Nicrophorus vespilloides]	Coleoptera	XP_017786052	1.29E-47	81.67939	176.407	131	107
Lep_0033060-RA	lysosomal alpha-mannosidase-like	711	gi|662196129|ref|XP_008471080.1|PREDICTED: lysosomal alpha-mannosidase-like [Diaphorina citri]	Hemiptera	XP_008471080	9.03E-81	70.76271	253.447	236	167
Lep_0033059-RA	lysosomal alpha-mannosidase-like	447	gi|929380069|ref|XP_014100402.1|PREDICTED: lysosomal alpha-mannosidase-like, partial [Bactrocera oleae]	Diptera	XP_014100402	1.74E-59	79.84496	189.119	129	103

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00077642-RA	alpha-mannosidase 2	735	gi|1227998244|ref|XP_021930961.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998246|ref|XP_021930962.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998248|ref|XP_021930963.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998250|ref|XP_021930964.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998252|ref|XP_021930965.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998254|ref|XP_021930966.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|646706309|gb|KDR13614.1|Alpha-mannosidase 2 [Zootermopsis nevadensis]	Blattodea	XP_021930961, XP_021930962, XP_021930963, XP_021930964, XP_021930965, XP_021930966, KDR13614	1.22E-57	57.55102	203.756	245	141
Lep_00034060-RA	lysosomal alpha-mannosidase	465	gi|1130264682|ref|XP_019770786.1|PREDICTED: lysosomal alpha-mannosidase-like [Dendroctonus ponderosae]	Coleoptera	XP_019770786	1.7E-20	54.67626	93.2041	139	76

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0001189 3-RA	lysosomal alpha-mannosidase isoform X1	2730	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	0	64.50428	756.518	817	527
Lep_0012411 0-RA	alpha-mannosidase 2	336	gi|1009600602|ref|XP_015928615.1|alpha-mannosidase 2 isoform X1 [Parasteatoda tepidariorum]	Arachnida	XP_015928615	2.28E-45	80.18868	162.155	106	85
Lep_0004690 2-RA	lysosomal alpha-mannosidase isoform X1	1275	gi|1048012429|ref|XP_017467120.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Rhagoletis zephyria]gi|1048012433|ref|XP_017467123.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Rhagoletis zephyria]	Diptera	XP_017467120, XP_017467123	9.16E-97	60.59783	319.316	368	223
Lep_0011570 1-RA	ER degradation-enhancing alpha-mannosidase-like protein 1	399	gi|1330882322|gb|PNF18187.1|ER degradation-enhancing alpha-mannosidase-like protein 1, partial [Cryptotermes secundus]	Blattodea	PNF18187	6.99E-64	91.59664	204.527	119	109

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00098856-RA	ER degradation-enhancing alpha-mannosidase-like protein 1	336	gi|1227982315|ref|XP_021922781.1|ER degradation-enhancing alpha-mannosidase-like protein 1 [Zootermopsis nevadensis]gi|646713767|gb|KDR17988.1|ER degradation-enhancing alpha-mannosidase-like 1 [Zootermopsis nevadensis]	Blattodea	XP_021922781, KDR17988	7.98E-37	81.81818	136.732	88	72
Lep_00127203-RA	Lysosomal alpha-mannosidase	231	gi|1339064620|ref|XP_023715606.1|lysosomal alpha-mannosidase isoform X2 [Cryptotermes secundus]gi|1330931495|gb|PNF42229.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715606, PNF42229	1.65E-12	62.66667	66.6254	75	47
Lep_00136882-RA	lysosomal alpha-mannosidase isoform X2	504	gi|1000733841|ref|XP_015589368.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Cephus cinctus]	Hymenoptera	XP_015589368	3.38E-37	63.95349	141.739	172	110
Lep_00065533-RA	lysosomal alpha-mannosidase isoform X1	825	gi|929380069|ref|XP_014100402.1|PREDICTED: lysosomal alpha-mannosidase-like, partial [Bactrocera oleae]	Diptera	XP_014100402	2.66E-67	78.66667	214.157	150	118

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00082857-RA	alpha-mannosidase 2	326	gi|1009600602|ref|XP_015928615.1|alpha-mannosidase 2 isoform X1 [Parasteatoda tepidariorum]	Arachnida	XP_015928615	5.75E-45	80.95238	160.999	105	85
Lep_00099228-RA	lysosomal alpha-mannosidase isoform X2	719	gi|1351647487|ref|XP_024086113.1|lysosomal alpha-mannosidase-like [Cimex lectularius]	Hemiptera	XP_024086113	1.15E-48	64.35644	177.563	202	130
Lep_00101070-RA	lysosomal alpha-mannosidase	462	gi|985414995|ref|XP_015372896.1|PREDICTED: lysosomal alpha-mannosidase-like [Diuraphis noxia]	Hemiptera	XP_015372896	3.49E-29	55.7047	117.857	149	83
Lep_00069567-RA	alpha-mannosidase 2	3048	gi|1339078211|ref|XP_023726521.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339078213|ref|XP_023726522.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339078215|ref|XP_023726523.1|alpha-mannosidase 2 [Cryptotermes secundus]	Blattodea	XP_023726521, XP_023726522, XP_023726523	0	79.20605	1444.1	1058	838
Lep_00062235-RA	lysosomal alpha-mannosidase-like	435	gi|1096257085|gb|APA33853.1|seminal fluid protein [Nilaparvata lugens]	Hemiptera	APA33853	2.72E-33	68.10345	129.413	116	79

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0000587 0-RA	lysosomal alpha-mannosidase isoform X1	873	gi|1048012429|ref|XP_017467120.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Rhagoletis zephyria]gi|1048012433|ref|XP_017467123.1|PREDICTED: lysosomal alpha-mannosidase isoform X1 [Rhagoletis zephyria]	Diptera	XP_017467120, XP_017467123	2E-110	78.92562	350.132	242	191
Lep_0000586 7-RA	Lysosomal alpha-mannosidase	1357	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	1.75E-78	63.71951	269.24	328	209
Lep_0000586 8-RA	lysosomal alpha-mannosidase-like	612	gi|1247040342|gb|PCG75287.1|hypothetical protein B5V51_11959 [Heliothis virescens]	Lepidoptera	PCG75287	9.48E-10	73.07692	63.929	52	38

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00001127-RA	alpha-mannosidase 2	2151	gi|1339052711|ref|XP_023714691.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052713|ref|XP_023714692.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052715|ref|XP_023714693.1|alpha-mannosidase 2 [Cryptotermes secundus]gi|1339052717|ref|XP_023714694.1|alpha-mannosidase 2 [Cryptotermes secundus]	Blattodea	XP_023714691, XP_023714692, XP_023714693, XP_023714694	0	78.50953	791.186	577	453

Table 3.14. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00001128-RA	alpha-mannosidase 2	990	gi|1227998244|ref|XP_021930961.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998246|ref|XP_021930962.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998248|ref|XP_021930963.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998250|ref|XP_021930964.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998252|ref|XP_021930965.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|1227998254|ref|XP_021930966.1|alpha-mannosidase 2 [Zootermopsis nevadensis]gi|646706309|gb|KDR13614.1|Alpha-mannosidase 2 [Zootermopsis nevadensis]	Blattodea	XP_021930961, XP_021930962, XP_021930963, XP_021930964, XP_021930965, XP_021930966, KDR13614	5.46E-66	66.42599	230.72	277	184
Lep_00001129-RA	alpha-mannosidase 2	918	gi|1330899240|gb|PNF26420.1|hypothetical protein B7P43_G16606, partial [Cryptotermes secundus]	Blattodea	PNF26420	5.88E-82	72.60274	264.618	219	159

Table 3.14. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00006052-RA	Lysosomal alpha-mannosidase	2921	gi|1339064618|ref|XP_023715597.1|lysosomal alpha-mannosidase isoform X1 [Cryptotermes secundus]gi|1330931496|gb|PNF42230.1|Lysosomal alpha-mannosidase [Cryptotermes secundus]	Blattodea	XP_023715597, PNF42230	0	63.4434	783.867	848	538

Table 3.15. Glucuronidases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for glucuronidases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00101143-RA	beta-glucuronidase-like isoform X1	414	gi|1330907784|gb|PNF30346.1|hypothetical protein B7P43_G13405 [Cryptotermes secundus]	Blattodea	PNF30346	1.24E-30	78.57143	116.316	84	66
Lep_00022238-RA	beta-glucuronidase-like	918	gi|241309940|ref|XP_002407819.1|beta-glucuronidase (GusB), putative [Ixodes scapularis]gi|215497226|gb|EEC06720.1|beta-glucuronidase (GusB), putative [Ixodes scapularis]	Arachnida	XP_002407819, EEC06720	1.97E-87	70.51282	267.314	234	165
Lep_00022239-RA	beta-glucuronidase	576	gi|1344820106|ref|XP_023953356.1|beta-glucuronidase-like, partial [Bicyclus anynana]	Lepidoptera	XP_023953356	3.75E-31	65.21739	122.865	115	75
Lep_00000156-RA	beta-glucuronidase isoform X3	944	gi|939667583|ref|XP_014280874.1|PREDICTED: beta-glucuronidase-like isoform X1 [Halyomorpha halys]gi|939667586|ref|XP_014280876.1|PREDICTED: beta-glucuronidase-like isoform X1 [Halyomorpha halys]gi|939667589|ref|XP_014280877.1|PREDICTED: beta-glucuronidase-like isoform X1 [Halyomorpha halys]	Hemiptera	XP_014280874, XP_014280876, XP_014280877	7.49E-68	77.05882	229.18	170	131

Table 3.15. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0000158-RA	beta-glucuronidase-like isoform X1	942	gi|641675506|ref|XP_008187067.1|PREDICTED: beta-glucuronidase isoform X2 [Acyrthosiphon pisum]	Hemiptera	XP_008187067	3.27E-95	75.96567	300.827	233	177
Lep_0006534 6-RA	beta-glucuronidase-like isoform X1	285	gi|1233178675|ref|XP_022190762.1|beta-glucuronidase-like [Nilaparvata lugens]	Hemiptera	XP_022190762	8.47E-23	87.30159	97.0561	63	55
Lep_0006534 5-RA	beta-glucuronidase-like	372	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda	XP_023221887	7.47E-13	67.21311	69.707	61	41
Lep_0007599 9-RA	beta-glucuronidase-like isoform X1	672	gi|1022772830|gb|KZS16987.1|Beta-glucuronidase [Daphnia magna]	Crustacea	KZS16987	8.94E-61	77.53623	192.586	138	107
Lep_0005373 8-RA	beta-glucuronidase-like	510	gi|1330889853|gb|PNF22051.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	PNF22051	3.6E-69	75.49669	217.624	151	114
Lep_0005585 5-RA	beta-glucuronidase-like	468	gi|1330889853|gb|PNF22051.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	PNF22051	2.73E-69	75.49669	217.238	151	114

Table 3.15. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0011856 3-RA	Beta-glucuronidase	657	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Crustacea	XP_023221887	6.22E-44	76.22951	159.073	122	93
Lep_0011833 3-RA	Beta-glucuronidase	657	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Crustacea	XP_023221887	4.16E-43	76.22951	157.147	122	93
Lep_0004714 6-RA	beta-glucuronidase-like isoform X1	1386	gi|1339060577|ref|XP_023718876.1|beta-glucuronidase-like isoform X2 [Cryptotermes secundus]	Blattodea	XP_023718876	2.8E-163	70.29478	480.33	441	310
Lep_0003971 3-RA	beta-glucuronidase-like	414	gi|1228013990|ref|XP_021938962.1|beta-glucuronidase [Zootermopsis nevadensis]gi|1228013992|ref|XP_021938963.1|beta-glucuronidase [Zootermopsis nevadensis]gi|1228013994|ref|XP_021938964.1|beta-glucuronidase [Zootermopsis nevadensis]	Blattodea	XP_021938962, XP_021938963, XP_021938964	3.24E-11	79.06977	65.855	43	34
Lep_0003971 4-RA	beta-glucuronidase (GusB), putative	258	gi|241309940|ref|XP_002407819.1|beta-glucuronidase (GusB), putative [Ixodes scapularis]gi|215497226|gb|EEC06720.1|beta-glucuronidase (GusB), putative [Ixodes scapularis]	Arachnida	XP_002407819, EEC06720	7.11E-24	79.03226	95.9005	62	49

Table 3.15. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00039715-RA	beta-glucuronidase isoform X3	621	gi|646696603|gb|KDR08779.1|Beta-glucuronidase [Zootermopsis nevadensis]	Blattodea	KDR08779	8.4E-19	75.38462	90.1225	65	49
Lep_00022144-RA	beta-glucuronidase-like	2202	gi|1339060575|ref|XP_023718875.1|beta-glucuronidase-like isoform X1 [Cryptotermes secundus]gi|1330889851|gb|PNF22049.1|Beta-glucuronidase [Cryptotermes secundus]	Blattodea	XP_023718875, PNF22049	0	65.75758	657.136	660	434
Lep_00046431-RA	beta-glucuronidase-like isoform X3	1161	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	1E-124	73.65269	377.096	334	246
Lep_00028614-RA	beta-glucuronidase-like	738	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	4.1E-71	69.77778	233.032	225	157

Table 3.15. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00028613-RA	beta-glucuronidase-like isoform X3	750	gi|1316206353|ref|XP_023213588.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda	XP_023213588	3.42E-47	85.84906	162.925	106	91
Lep_00130536-RA	beta-glucuronidase-like isoform X1	195	gi|1233178675|ref|XP_022190762.1|beta-glucuronidase-like [Nilaparvata lugens]	Hemiptera	XP_022190762	3.27E-23	87.30159	96.6709	63	55
Lep_00021005-RA	beta-glucuronidase-like isoform X1	564	gi|1339060579|ref|XP_023718877.1|beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi|1330889852|gb|PNF22050.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	XP_023718877, PNF22050	1.89E-48	67.58242	171.4	182	123
Lep_00021003-RA	beta-glucuronidase-like isoform X3	726	gi|1316206353|ref|XP_023213588.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda	XP_023213588	6.62E-79	75	243.817	192	144
Lep_00021004-RA	beta-glucuronidase isoform X1	375	gi|968019329|ref|XP_015011863.1|uncharacterized protein Dere_GG20922, isoform C [Drosophila erecta]gi|945202794|gb|KQS62122.1|uncharacterized protein Dere_GG20922, isoform C [Drosophila erecta]	Diptera	XP_015011863, KQS62122	2.97E-13	72.88136	71.2478	59	43

Table 3.15. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00022123-RA	beta-glucuronidase-like	1026	gi|1227991600|ref|XP_021927576.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991602|ref|XP_021927578.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991604|ref|XP_021927579.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991606|ref|XP_021927580.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991608|ref|XP_021927581.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991610|ref|XP_021927582.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991612|ref|XP_021927583.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|1227991614|ref|XP_021927584.1|beta-glucuronidase-like [Zootermopsis nevadensis]gi|646709721|gb|KDR15442.1|Beta-glucuronidase [Zootermopsis nevadensis]	Blattodea	XP_021927576, XP_021927578, XP_021927579, XP_021927580, XP_021927581, XP_021927582, XP_021927583, XP_021927584, KDR15442	2.8E-100	64.50617	314.309	324	209
Lep_00022122-RA	Beta-glucuronidase	1276	gi|1228362729|ref|XP_021967946.1|beta-glucuronidase-like isoform X1 [Folsomia candida]	Collembola	XP_021967946	1.1E-122	76.97842	376.711	278	214

Table 3.15. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00022813-RA	beta-glucuronidase isoform X3	609	gi|1330889853|gb|PNF22051.1|hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattodea	PNF22051	9.44E-64	69.8324	204.912	179	125
Lep_00022814-RA	beta-glucuronidase-like isoform X3	612	gi|641675506|ref|XP_008187067.1|PREDICTED: beta-glucuronidase isoform X2 [Acyrthosiphon pisum]	Hemiptera	XP_008187067	4.85E-48	72.60274	172.17	146	106
Lep_00024190-RA	beta-glucuronidase-like	1134	gi|1316153756|ref|XP_023221887.1|beta-glucuronidase-like [Centruroides sculpturatus]	Arthropoda	XP_023221887	1.14E-94	71.98582	297.36	282	203
Lep_00024189-RA	beta-glucuronidase isoform X1	621	gi|1321309229|ref|XP_023296676.1|beta-glucuronidase isoform X2 [Lucilia cuprina]	Diptera	XP_023296676	7.51E-66	76.50602	219.935	166	127

Table 3.16. Maltases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for maltases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_0005745-RA	Maltase 1	1110	gi|1061480896|gb|ODM98289.1|Maltase 1 [Orchesella cincta]	Collembola	ODM98289	4.28E-20	79.16667	98.2117	72	57
Lep_0011065 9-RA	Maltase 1	651	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	2.88E-87	75.46296	273.863	216	163
Lep_0009100 6-RA	maltase 2-like	522	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	6.51E-43	76.03306	155.992	121	92
Lep_0009671 1-RA	Maltase 1	648	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	1.48E-56	78.83212	192.586	137	108

Table 3.16. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00077259-RA	maltase 2-like	834	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	5.56E-70	70.93596	231.876	203	144
Lep_00035421-RA	sucrase-isomaltase, intestinal-like	1707	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	2.66E-59	61.37072	220.32	321	197
Lep_00034930-RA	Maltase 1	1410	gi|1330891183|gb|PNF22668.1|Maltase 1 [Cryptotermes secundus]	Blattodea	PNF22668	1.1E-122	60.63348	375.941	442	268
Lep_00113997-RA	Maltase 1	615	gi|1228019306|ref|XP_021941640.1|maltase 2-like isoform X1 [Zootermopsis nevadensis]	Blattodea	XP_021941640	5.24E-25	64.28571	107.842	98	63
Lep_00062357-RA	Maltase 1	759	gi|1339059590|ref|XP_023718348.1|maltase 2-like [Cryptotermes secundus]gi|1330891147|gb|PNF22635.1|Maltase 2 [Cryptotermes secundus]	Blattodea	XP_023718348, PNF22635	4.13E-89	75.79909	280.411	219	166

Table 3.16. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00033214-RA	Maltase-glucoamylase, intestinal	570	gi|1067100065|ref|XP_018012117.1|PREDICTED: sucrase-isomaltase, intestinal-like [Hyalella azteca]	Crustacea	XP_018012117	1.45E-12	42.01681	71.633	119	50
Lep_00033215-RA	sucrase-isomaltase, intestinal-like	852	gi|675371953|gb|KFM64855.1|Maltase-glucoamylase, intestinal, partial [Stegodyphus mimosarum]	Arachnida	KFM64855	6.34E-60	71.5847	202.216	183	131
Lep_01404 43-RA	Maltase 1	193	gi|1339059395|ref|XP_023718247.1|probable maltase isoform X1 [Cryptotermes secundus]	Blattodea	XP_023718247	1.17E-16	73.4375	77.7962	64	47
Lep_0112138-RA	maltase 2-like	378	gi|1101349818|ref|XP_018900831.1|PREDICTED: maltase A1-like isoform X3 [Bemisia tabaci]	Hemiptera	XP_018900831	1.58E-20	87.03704	92.0485	54	47

Table 3.16. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_000069757-RA	maltase-glucoamylase, intestinal-like	969	gi|1022768805|gb|KZS13426.1|Uncharacterized protein APZ42_021387 [Daphnia magna]	Crustacea	KZS13426	9.17E-36	60.36585	143.665	164	99
Lep_00056135-RA	maltase 1-like isoform X2	627	gi|1153833895|ref|XP_020293571.1|maltase 1-like isoform X2 [Pseudomyrmex gracilis]	Hymenoptera	XP_020293571	0.000312	71.42857	47.7506	42	30
Lep_00056136-RA	maltase 1-like isoform X2	387	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	7.88E-44	71.09375	155.221	128	91

Table 3.16. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_000082089-RA	maltase-glucoamylase, intestinal-like	666	gi|1238867836|ref|XP_022249217.1|lysosomal alpha-glucosidase-like [Limulus polyphemus]	Arthropoda	XP_022249217	4.65E-35	70.70707	126.331	99	70
Lep_000063763-RA	Maltase-glucoamylase, intestinal	312	gi|1067072134|ref|XP_018019496.1|PREDICTED: LOW QUALITY PROTEIN: probable maltase-glucoamylase 2, partial [Hyalella azteca]	Crustacea	XP_018019496	1.12E-12	58.57143	68.5514	70	41
Lep_000063764-RA	maltase-glucoamylase, intestinal-like	507	gi|1048057520|ref|XP_017462059.1|PREDICTED: maltase-glucoamylase, intestinal-like, partial [Rhagoletis zephyria]	Diptera	XP_017462059	6.34E-49	66.44737	166.007	152	101
Lep_000056080-RA	maltase 1	279	gi|1036973924|ref|XP_016963459.1|PREDICTED: maltase 1 [Drosophila biarmipes]	Diptera	XP_016963459	2.52E-12	59.67742	67.0106	62	37
Lep_000056081-RA	Maltase 1	699	gi|1061480896|gb|ODM98289.1|Maltase 1 [Orchesella cincta]	Collembola	ODM98289	2.82E-22	78.37838	100.908	74	58
Lep_000086862-RA	sucrase-isomaltase, intestinal-like	723	gi|675371953|gb|KFM64855.1|Maltase-glucoamylase, intestinal, partial [Stegodyphus mimosarum]	Arachnida	KFM64855	2.49E-72	64.68254	232.646	252	163

Table 3.16. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00124732-RA	maltase 2-like	234	gi|1185574477|ref|XP_003395914.2|maltase 1 [Bombus terrestris]	Hymenoptera	XP_003395914	6.11E-17	77.35849	79.337	53	41
Lep_00005124-RA	Maltase 1	678	gi|1227980746|ref|XP_021921963.1|uncharacterized protein KIAA0513 isoform X3 [Zootermopsis nevadensis]	Blattodea	XP_021921963	2.51E-42	66.83938	154.836	193	129
Lep_00043003-RA	maltase 2 isoform X1	918	gi|1317983107|ref|XP_012271316.2|uncharacterized protein LOC105694836 [Orussus abietinus]	Hymenoptera	XP_012271316	3.48E-45	79.2	170.244	125	99
Lep_00043002-RA	maltase 2-like	567	gi|1101338755|ref|XP_018917587.1|PREDICTED: maltase 2-like [Bemisia tabaci]gi|1101338757|ref|XP_018917588.1|PREDICTED: maltase 2-like [Bemisia tabaci]gi|1101338759|ref|XP_018917589.1|PREDICTED: maltase 2-like [Bemisia tabaci]	Hemiptera	XP_018917587, XP_018917588, XP_018917589	1.08E-17	75.75758	86.2705	66	50

Table 3.16. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00006058-RA	maltase 2-like	2673	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	2.51E-96	70.84871	317.39	271	192
Lep_00068867-RA	Maltase-glucoamylase, intestinal	915	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	3.4E-105	67.32026	339.347	306	206
Lep_00011465-RA	sucrase-isomaltase, intestinal-like	1767	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	0	67.08408	596.66	559	375
Lep_00045479-RA	maltase-glucoamylase, intestinal	1600	gi|321476731|gb|EFX87691.1|hypothetical protein DAPPUDRAFT_306567 [Daphnia pulex]	Crustacea	EFX87691	1.6E-133	57.28346	410.223	508	291
Lep_00108134-RA	Maltase 1	441	gi|646696837|gb|KDR08865.1|Maltase 1 [Zootermopsis nevadensis]	Blattodea	KDR08865	3.54E-23	81.96721	100.523	61	50

Table 3.16. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00036231-RA	sucrase-isomaltase, intestinal-like	1596	gi|1067098296|ref|XP_018011173.1|PREDICTED: maltase-glucoamylase, intestinal-like [Hyalella azteca]	Crustacea	XP_018011173	1.5E-130	63.63636	416.772	451	287
Lep_00071068-RA	Maltase 1	672	gi|1228013734|ref|XP_021938828.1|maltase 2-like [Zootermopsis nevadensis]	Blattodea	XP_021938828	2.71E-21	84.74576	97.8265	59	50

Table 3.17. Amylases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for amylases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00082293-RA	alpha amylase	573	gi|1339059399|ref|XP_023718249.1|maltase 2-like isoform X3 [Cryptotermes secundus]gi|1330891182|gb|PNF22667.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]gi|1330891184|gb|PNF22669.1|hypothetical protein B7P43_G07129 [Cryptotermes secundus]	Blattodea	XP_023718249, PNF22667, PNF22669	8.61E-67	70.58824	218.0009	187	132
Lep_00018357-RA	alpha-amylase-like isoform X2	423	gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605017	8.16E-39	80.19802	143.665	101	81
Lep_00018359-RA	alpha-amylase 2-like	1356	gi|540849793|gb|AGV15452.1|alpha-amylase, partial [Blattella germanica]	Blattodea	AGV15452	2.75E-28	93.84615	123.25	65	61
Lep_00081749-RA	alpha-amylase-like isoform X2	567	gi|62955866|gb|AAY23288.1|1,4-alpha-D-glucan glucanohydrolase precursor [Blattella germanica]	Blattodea	AAY23288	2.05E-11	76.74419	68.1662	43	33

Table 3.17. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00035375-RA	alpha-amylase A-like	372	gi|998506683|ref|XP_015513623.1|PREDICTED: alpha-amylase 2-like [Neodiprion lecontei]	Hymenoptera	XP_015513623	2.84E-44	76.47059	155.992	102	78
Lep_00035373-RA	alpha-amylase 4N-like	561	gi|68266131|gb|AAY88833.1|alpha-amylase, partial [Coelopa frigida]	Diptera	AAY88833	1.39E-14	86.36364	76.6406	44	38
Lep_00082269-RA	Alpha-amylase 1	567	gi|540849793|gb|AGV15452.1|alpha-amylase, partial [Blattella germanica]	Blattodea	AGV15452	1.13E-71	79.87013	230.335	154	123
Lep_00131724-RA	alpha-amylase 1-like	315	gi|1000763879|ref|XP_015605017.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605017	1.38E-18	81.48148	85.5001	54	44
Lep_00075793-RA	Pancreatic alpha-amylase	642	gi|62955866|gb|AAY23288.1|1,4-alpha-D-glucan glucanohydrolase precursor [Blattella germanica]	Blattodea	AAY23288	8.52E-10	75.60976	63.929	41	31
Lep_00062602-RA	Pancreatic alpha-amylase	663	gi|1061479081|gb|ODM96628.1|Pancreatic alpha-amylase [Orchesella cincta]	Collembola	ODM96628	3.48E-26	69.30693	111.309	101	70

Table 3.17. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_000091367-RA	alpha-amylase 1-like	252	gi|602182726|gb|AHN91843.1|alpha-amylase [Marsupenaeus japonicus]gi|602182728|gb|AHN91844.1|alpha-amylase [Marsupenaeus japonicus]	Crustacea	AHN91843, AHN91844	2.09E-21	86	92.4337	50	43
Lep_000082183-RA	alpha-amylase 1-like isoform X1	702	gi|1067066079|ref|XP_018017054.1|PREDICTED: alpha-amylase 1-like isoform X1 [Hyalella azteca]	Crustacea	XP_018017054	1.77E-35	82.47423	137.502	97	80
Lep_000111509-RA	alpha-amylase 1-like	600	gi|540849793|gb|AGV15452.1|alpha-amylase, partial [Blattella germanica]	Blattodea	AGV15452	3.72E-47	81.65138	167.162	109	89
Lep_000083303-RA	Alpha-amylase 1	687	gi|85002763|gb|ABC68516.1|alpha-amylase [Blattella germanica]	Blattodea	ABC68516	2.3E-25	77.77778	109.383	81	63
Lep_000099194-RA	alpha-amylase A-like	519	gi|998506681|ref|XP_015513622.1|PREDICTED: alpha-amylase A-like [Neodiprion lecontei]	Hymenoptera	XP_015513622	6.71E-42	81.52174	147.902	92	75
Lep_000114471-RA	alpha-amylase 1-like	366	gi|602182726|gb|AHN91843.1|alpha-amylase [Marsupenaeus japonicus]gi|602182728|gb|AHN91844.1|alpha-amylase [Marsupenaeus japonicus]	Crustacea	AHN91843, AHN91844	2.16E-20	86	91.2781	50	43

Table 3.17. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00103549-RA	alpha-amylase 1-like	354	gi|1000764249|ref|XP_015605210.1|PREDICTED: alpha-amylase 1-like [Cephus cinctus]	Hymenoptera	XP_015605210	7.69E-14	78	72.4034	50	39
Lep_00113794-RA	alpha-amylase A-like	378	gi|68266167|gb|AAY88846.1|alpha-amylase, partial [Musca domestica]	Diptera	AAY88846	3.88E-45	76.10619	149.828	113	86
Lep_00093647-RA	alpha-amylase 1-like	738	gi|1339092808|ref|XP_023707620.1|alpha-amylase 1-like [Cryptotermes secundus]gi|1330914763|gb|PNF33597.1|Alpha-amylase 1 [Cryptotermes secundus]	Blattodea	XP_023707620, PNF33597	1.83E-28	78.87324	118.627	71	56
Lep_00133347-RA	alpha-amylase 1-like	261	gi|168830277|gb|ACA34383.1|alpha-amylase, partial [Lepisma saccharina]	Zygentoma	ACA34383	9.86E-26	88.46154	96.2857	52	46
Lep_00131390-RA	alpha-amylase 1-like	354	gi|168830254|gb|ACA34371.1|alpha-amylase, partial [Periplaneta americana]	Blattodea	ACA34371	1.87E-47	73.91304	154.451	115	85
Lep_00031081-RA	alpha-amylase 1-like	1584	gi|1227981869|ref|XP_021922548.1|alpha-amylase 2-like [Zootermopsis nevadensis]	Blattodea	XP_021922548	1.42E-84	71.36564	276.559	227	162
Lep_00112687-RA	alpha-amylase A-like	513	gi|139478961|gb|ABO77639.1|alpha-amylase, partial [Musca domestica]	Diptera	ABO77639	1.73E-64	78.16901	202.601	142	111

Table 3.18. Galactosidases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for galactosidases and their blast description.

<table>
<tr><th>Sequence name</th><th>Sequence desc.</th><th>Sequence length</th><th>Hit desc.</th><th>Order</th><th>Hit ACC</th><th>E-Value</th><th>Similarity</th><th>Bit-Score</th><th>Alignment length</th><th>Positives</th></tr>
<tr><td>Lep_00086810-RA</td><td>alpha-galactosidase A-like</td><td>609</td><td>gi|1227958457|ref|XP_021926901.1|alpha-N-acetylgalactosaminidase-like [Zootermopsis nevadensis]gi|1227958459|ref|XP_021926910.1|alpha-N-acetylgalactosaminidase-like [Zootermopsis nevadensis]gi|646723001|gb|KDR23773.1|Alpha-N-acetylgalactosaminidase [Zootermopsis nevadensis]</td><td>Blattodea</td><td>XP_021926901, XP_021926910, KDR23773</td><td>3.54E-48</td><td>80</td><td>167.933</td><td>110</td><td>88</td></tr>
<tr><td>Lep_00127369-RA</td><td>beta-galactosidase-1-like protein 2</td><td>324</td><td>gi|1228405415|ref|XP_021961831.1|beta-galactosidase-1-like protein 2 [Folsomia candida]</td><td>Collembola</td><td>XP_021961831</td><td>1.22E-23</td><td>73.9726</td><td>92.8189</td><td>73</td><td>54</td></tr>
<tr><td>Lep_00026440-RA</td><td>Beta-galactosidase</td><td>2145</td><td>gi|1339081983|ref|XP_023702009.1|LOW QUALITY PROTEIN: beta-galactosidase-like [Cryptotermes secundus]</td><td>Blattodea</td><td>XP_023702009</td><td>0</td><td>70.03106</td><td>685.256</td><td>644</td><td>451</td></tr>
</table>

Table 3.18. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00059101-RA	beta-galactosidase-1-like protein 2	1416	gi|1000751440|ref|XP_015598574.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751442|ref|XP_015598576.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751444|ref|XP_015598577.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751446|ref|XP_015598578.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751448|ref|XP_015598579.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751450|ref|XP_015598580.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]	Hymenoptera	XP_015598574, XP_015598576, XP_015598577, XP_015598578, XP_015598579, XP_015598580	2.76E-66	67.69912	230.335	226	153

Table 3.18. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00123361-RA	beta-galactosidase-1-like protein 2 isoform X2	240	gi|1330893432|gb|PNF23708.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattodea	PNF23708	1.52E-27	86.88525	108.612	61	53
Lep_00011817-RA	Beta-galactosidase	1926	gi|1227962476|ref|XP_021942674.1|beta-galactosidase isoform X2 [Zootermopsis nevadensis]	Blattodea	XP_021942674	1.2E-146	65.08621	445.662	464	302
Lep_00011816-RA	beta-galactosidase isoform X3	738	gi|1227962478|ref|XP_021942675.1|beta-galactosidase isoform X3 [Zootermopsis nevadensis]	Blattodea	XP_021942675	3.56E-46	84.61538	167.548	104	88
Lep_00166804-RA	beta-galactosidase-1-like protein 2 isoform X2	189	gi|1000751438|ref|XP_015598573.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X1 [Cephus cinctus]	Hymenoptera	XP_015598573	5.02E-27	86.88525	107.457	61	53

Table 3.18. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00006759-RA	Beta-galactosidase-1-like protein 2	814	gi|1215258369|gb|OXA44696.1|Beta-galactosidase-1-like protein 2 [Folsomia candida]	Collembola	OXA44696	1.77E-39	55	144.05	200	110

Table 3.18. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00006757-RA	beta-galactosidase-1-like protein 2	2064	gi|1000751440|ref|XP_015598574.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751442|ref|XP_015598576.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751444|ref|XP_015598577.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751446|ref|XP_015598578.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751448|ref|XP_015598579.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751450|ref|XP_015598580.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]	Hymenoptera	XP_015598574, XP_015598576, XP_015598577, XP_015598578, XP_015598579, XP_015598580	1.19E-63	72.12389	229.18	226	163

Table 3.18. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00028717-RA	beta-galactosidase isoform X3	1014	gi|1330925577|gb|PNF39271.1|hypothetical protein B7P43_G16708 [Cryptotermes secundus]	Blattodea	PNF39271	3.01E-98	56.89655	306.605	348	198
Lep_00028716-RA	Beta-galactosidase	729	gi|1022758129|gb|KZS04700.1|Beta-galactosidase [Daphnia magna]	Crustacea	KZS04700	3.54E-36	61.07784	140.969	167	102
Lep_00117981-RA	beta-galactosidase-like	1149	gi|675376294|gb|KFM69196.1|Beta-galactosidase, partial [Stegodyphus mimosarum]	Arachnida	KFM69196	3.48E-24	62.39316	106.686	117	73
Lep_00067905-RA	beta-galactosidase	699	gi|970886870|ref|XP_015108515.1|PREDICTED: beta-galactosidase [Diachasma alloeum]	Hymenoptera	XP_015108515	3.98E-18	82.35294	88.9669	51	42
Lep_00126791-RA	beta-galactosidase-1-like protein 2 isoform X2	393	gi|1330893432|gb|PNF23708.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattodea	PNF23708	1.25E-62	83.46457	203.371	127	106
Lep_00026029-RA	Beta-galactosidase	852	gi|1330925577|gb|PNF39271.1|hypothetical protein B7P43_G16708 [Cryptotermes secundus]	Blattodea	PNF39271	2.03E-66	73.57513	222.246	193	142

Table 3.18. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00026030-RA	Beta-galactosidase	798	gi|1227962474|ref|XP_021942673.1|beta-galactosidase isoform X1 [Zootermopsis nevadensis]gi|646721601|gb|KDR22879.1|Beta-galactosidase [Zootermopsis nevadensis]	Blattodea	XP_021942673, KDR22879	3.23E-29	61.19403	122.094	134	82
Lep_00008899-RA	Beta-galactosidase-1-like protein 2	2361	gi|1339057562|ref|XP_023717284.1|beta-galactosidase-1-like protein 2 isoform X2 [Cryptotermes secundus]	Blattodea	XP_023717284	0	56.35649	579.711	763	430

Table 3.18. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00036877-RA	beta-galactosidase-1-like protein 2	954	gi|1000751440|ref|XP_015598574.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751442|ref|XP_015598576.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751444|ref|XP_015598577.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751446|ref|XP_015598578.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751448|ref|XP_015598579.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]gi|1000751450|ref|XP_015598580.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X2 [Cephus cinctus]	Hymenoptera	XP_015598574, XP_015598576, XP_015598577, XP_015598578, XP_015598579, XP_015598580	8.9E-62	56.6529	213.386	293	166

Table 3.18. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00036876-RA	beta-galactosidase-1-like protein 2	405	gi|1000751438|ref|XP_015598573.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X1 [Cephus cinctus]	Hymenoptera	XP_015598573	5.28E-31	72.32143	122.094	112	81
Lep_00002666-RA	Beta-galactosidase	1782	gi|1339081983|ref|XP_023702009.1|LOW QUALITY PROTEIN: beta-galactosidase-like [Cryptotermes secundus]	Blattodea	XP_023702009	6.8E-173	68.33013	511.531	521	356
Lep_00002668-RA	beta-galactosidase-like isoform X1	474	gi|1227962478|ref|XP_021942675.1|beta-galactosidase isoform X3 [Zootermopsis nevadensis]	Blattodea	XP_021942675	1.05E-47	84.61538	167.933	104	88
Lep_00002498-RA	beta-galactosidase-1-like protein 2	2398	gi|1000751438|ref|XP_015598573.1|PREDICTED: beta-galactosidase-1-like protein 2 isoform X1 [Cephus cinctus]	Hymenoptera	XP_015598573	0	68.42105	659.833	684	468
Lep_00002499-RA	beta-galactosidase-1-like protein 2	2103	gi|998516729|ref|XP_015519163.1|PREDICTED: beta-galactosidase-1-like protein 2 [Neodiprion lecontei]	Hymenoptera	XP_015519163	2.9E-131	57.50916	407.527	546	314

Table 3.18. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00002500-RA	beta-galactosidase-1-like protein 2 isoform X2	786	gi|1330893432|gb|PNF23708.1|hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattodea	PNF23708	5.46E-60	84.12698	201.83	126	106
Lep_00118693-RA	beta-galactosidase-like isoform X1	601	gi|675376294|gb|KFM69196.1|Beta-galactosidase, partial [Stegodyphus mimosarum]	Arachnida	KFM69196	3.34E-25	71.91011	104.76	89	64

Table 3.19. Xylanases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for xylanases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00045817-RA	Endo-1,4-beta-xylanase A	231	gi|1229883103|ref|XP_022161694.1|uncharacterized protein LOC111027604 [Myzus persicae]	Hemiptera	XP_022161694	1.74E-10	70.3125	60.8474	64	45
Lep_00097333-RA	Endo-1,4-beta-xylanase A	696	gi|391327465|ref|XP_003738220.1|PREDICTED: uncharacterized protein LOC100898192 [Galendromus occidentalis]	Arachnida	XP_003738220	1.62E-11	58.18182	69.707	110	64
Lep_00130240-RA	Endo-1,4-beta-xylanase A	546	gi|391327465|ref|XP_003738220.1|PREDICTED: uncharacterized protein LOC100898192 [Galendromus occidentalis]	Arachnida	XP_003738220	5.2E-13	52.14286	72.4034	140	73

Table 3.20. Myrosinases in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for myrosinases and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00117435-RA	myrosinase 1-like	294	gi|998489397|gb|AMK48629.1| glycoside hydrolase family 1, partial [Rhynchophorus ferrugineus]	Coleoptera	AMK48629	7.01E-33	70.78652	116.701	89	63
Lep_00004871-RA	myrosinase 1-like	411	gi|1042851809|gb|ANQ90707.1|beta-glucosidase 2, partial [Manduca quinquemaculata]	Lepidoptera	ANQ90707	3.12E-29	75	108.612	84	63
Lep_00004870-RA	myrosinase 1-like	537	gi|1080062829|ref|XP_018573475.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573475	1.26E-54	78.19549	185.652	133	104
Lep_00034695-RA	myrosinase 1-like	390	gi|506967929|gb|AGM32308.1| beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	7.17E-14	82.97872	72.0182	47	39
Lep_00034694-RA	myrosinase 1-like	717	gi|506965870|gb|AGM32287.1| beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32287	4.1E-31	83.90805	121.709	87	73
Lep_00134073-RA	myrosinase 1-like	309	gi|1069789897|ref|XP_018320188.1|PREDICTED: myrosinase 1-like [Agrilus planipennis]	Coleoptera	XP_018320188	2.62E-31	88.57143	119.398	70	62
Lep_00019550-RA	myrosinase 1-like	1608	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	6.91E-42	68.08511	160.999	141	96

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00019549-RA	myrosinase 1-like isoform X1	279	gi|1229709445|ref|XP_022116625.1|myrosinase 1-like [Pieris rapae]	Lepidoptera	XP_022116625	1.04E-09	81.08108	59.6918	37	30
Lep_00021306-RA	myrosinase 1-like	564	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	9.13E-60	80.45113	193.356	133	107
Lep_00021307-RA	myrosinase 1-like	288	gi|1228411798|ref|XP_021965263.1|myrosinase 1-like [Folsomia candida]gi|1215254226|gb|OXA40567.1|Lactase-phlorizin hydrolase [Folsomia candida]	Collembola	XP_021965263, OXA40567	7.19E-08	70.17544	54.299	57	40
Lep_00042970-RA	myrosinase 1	1422	gi|1325357344|ref|XP_023312869.1|myrosinase 1 [Anoplophora glabripennis]	Coleoptera	XP_023312869	1.5E-115	73.22034	354.369	295	216
Lep_00048639-RA	myrosinase 1-like	456	gi|636793015|dbj|BAO85048.1|beta-glucosidase, partial [Salganea esakii]	Blattodea	BAO85048	2.17E-25	69.14894	103.99	94	65
Lep_00048640-RA	myrosinase 1-like	324	gi|939630683|ref|XP_014278711.1|PREDICTED: myrosinase 1-like [Halyomorpha halys]	Hemiptera	XP_014278711	1.34E-16	76	79.7221	50	38

Table 3.20. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00124670-RA	myrosinase 1-like	261	gi|1069790020|ref|XP_018320243.1|PREDICTED: myrosinase 1-like isoform X2 [Agrilus planipennis]gi|1069790022|ref|XP_018320244.1|PREDICTED: myrosinase 1-like isoform X3 [Agrilus planipennis]	Coleoptera	XP_018320243, XP_018320244	3.14E-15	66.66667	73.9442	90	60
Lep_00039278-RA	myrosinase 1-like	1776	gi|1339073035|ref|XP_023723774.1|myrosinase 1-like [Cryptotermes secundus]gi|1330878957|gb|PNF16841.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023723774, PNF16841	1E-118	59.77011	366.696	435	260
Lep_00152982-RA	myrosinase 1	387	gi|1080048148|ref|XP_018565412.1|myrosinase 1 [Anoplophora glabripennis]	Coleoptera	XP_018565412	0.00094	62.7907	43.8986	43	27
Lep_00114845-RA	myrosinase 1-like	537	gi|1080062831|ref|XP_018573476.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573476	2.13E-51	68.3871	177.178	155	106
Lep_00041124-RA	myrosinase 1-like	1305	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	2.28E-75	85.25641	242.276	156	133
Lep_00122267-RA	myrosinase 1-like	243	gi|1133421334|ref|XP_019866289.1|PREDICTED: myrosinase 1-like [Aethina tumida]	Coleoptera	XP_019866289	1.82E-34	86.41975	127.872	81	70

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00065284-RA	myrosinase 1-like	735	gi|506965870|gb|AGM32287.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32287	3.18E-62	58.8235 3	202.216	204	120
Lep_00068929-RA	myrosinase 1-like	726	gi|998493172|ref|XP_015513162.1|PREDICTED: myrosinase 1-like [Neodiprion lecontei]	Hymenoptera	XP_015513162	1.6E-11	63.3802 8	69.707	71	45
Lep_00120227-RA	myrosinase 1	204	gi|1008445131|ref|XP_975665.2|PREDICTED: myrosinase 1 isoform X2 [Tribolium castaneum]	Coleoptera	XP_975665	3.3E-07	83.6734 7	51.2174	49	41
Lep_00028153-RA	Myrosinase 1	1590	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	2.9E-119	64.0540 5	368.237	370	237
Lep_00133665-RA	myrosinase 1-like	321	gi|1069792344|ref|XP_018321256.1|PREDICTED: uncharacterized protein LOC108734274 [Agrilus planipennis]	Coleoptera	XP_018321256	3.91E-17	56.1904 8	81.2629	105	59

Table 3.20. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00022255-RA	Myrosinase 1	735	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	7.19E-88	73.7069	271.552	232	171
Lep_00022256-RA	myrosinase 1-like	396	gi|1042851809|gb|ANQ90707.1|beta-glucosidase 2, partial [Manduca quinquemaculata]	Lepidoptera	ANQ90707	7.93E-29	81.1594 2	107.457	69	56
Lep_00062260-RA	Myrosinase 1	324	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.8E-37	82.4175 8	132.88	91	75
Lep_00062259-RA	myrosinase 1	225	gi|1325352588|ref|XP_023311344.1|lactase-phlorizin hydrolase-like [Anoplophora glabripennis]	Coleoptera	XP_023311344	6.3E-23	80	96.2857	65	52
Lep_00004129-RA	myrosinase 1	435	gi|1042851809|gb|ANQ90707.1|beta-glucosidase 2, partial [Manduca quinquemaculata]	Lepidoptera	ANQ90707	4.49E-48	83.0357 1	156.762	112	93
Lep_00004128-RA	myrosinase 1-like	750	gi|1061476505|gb|ODM94351.1|Lactase-phlorizin hydrolase [Orchesella cincta]	Collembola	ODM94351	1.22E-26	80.8219 2	113.62	73	59
Lep_00004130-RA	myrosinase 1-like	456	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.3E-16	82.9787 2	80.1073	47	39

Table 3.20. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00106154-RA	Myrosinase 1	324	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.69E-37	83.51648	132.88	91	76
Lep_00022729-RA	myrosinase 1-like	1096	gi|1279707180|ref|XP_022915768.1|myrosinase 1-like [Onthophagus taurus]	Coleoptera	XP_022915768	9.87E-93	72.01646	291.967	243	175
Lep_00022730-RA	myrosinase 1-like	486	gi|1080058576|ref|XP_018571131.1|myrosinase 1 [Anoplophora glabripennis]	Coleoptera	XP_018571131	4.89E-21	73.01587	94.7449	63	46
Lep_00047597-RA	myrosinase 1-like	558	gi|636793009|dbj|BAO85045.1|beta-glucosidase, partial [Zootermopsis nevadensis]	Blattodea	BAO85045	4.43E-51	73.28767	171.785	146	107
Lep_00047598-RA	myrosinase 1-like isoform X1	624	gi|1069782941|ref|XP_018335102.1|PREDICTED: myrosinase 1-like [Agrilus planipennis]	Coleoptera	XP_018335102	4.66E-30	77.06422	121.709	109	84

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00015101-RA	Myrosinase 1	1457	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	1.2E-148	65.05263	441.81	475	309
Lep_00106870-RA	myrosinase 1	303	gi|1101347123|ref|XP_018899386.1|PREDICTED: myrosinase 1 [Bemisia tabaci]	Hemiptera	XP_018899386	1.73E-15	77.19298	76.2554	57	44
Lep_00111910-RA	myrosinase 1-like	501	gi|748995286|gb|AJE75665.1|putative glycosyl hydrolase [Chrysomela lapponica]	Coleoptera	AJE75665	6.21E-53	72.02797	180.644	143	103
Lep_00105473-RA	myrosinase 1-like	327	gi|218749835|ref|NP_001136332.1|glycoside hydrolase-like protein [Nasonia vitripennis]	Hymenoptera	NP_001136332	4.12E-14	76.59574	72.7886	47	36
Lep_00096192-RA	Myrosinase 1	165	gi|1339057501|ref|XP_023717250.1|myrosinase 1-like isoform X2 [Cryptotermes secundus]gi|1330893380|gb|PNF23656.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023717250, PNF23656	1.37E-16	86.66667	77.0258	45	39

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00020465-RA	myrosinase 1-like isoform X2	501	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	1.09E-59	79.02098	196.052	143	113
Lep_00020464-RA	myrosinase 1-like	885	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.56E-49	76.81159	171.014	138	106
Lep_00020463-RA	myrosinase 1-like	330	gi|1228381282|ref|XP_021952145.1|myrosinase 1-like [Folsomia candida]gi|1215269331|gb|OXA55643.1|Lactase-phlorizin hydrolase [Folsomia candida]	Collembola	XP_021952145, OXA55643	1.64E-10	84.84848	62.7734	33	28
Lep_00041138-RA	Myrosinase 1	882	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	2.61E-51	66.09195	182.956	174	115

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00041137-RA	myrosinase 1-like	408	gi|1080062338|ref|XP_018573207.1|myrosinase 1-like isoform X2 [Anoplophora glabripennis]	Coleoptera	XP_018573207	1.27E-20	72.72727	92.4337	66	48
Lep_00134507-RA	myrosinase 1-like	294	gi|998489397|gb|AMK48629.1|glycoside hydrolase family 1, partial [Rhynchophorus ferrugineus]	Coleoptera	AMK48629	1E-33	71.42857	118.627	91	65
Lep_00027729-RA	myrosinase 1-like	744	gi|1080062829|ref|XP_018573475.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573475	1.97E-57	78.91156	195.667	147	116
Lep_00027731-RA	myrosinase 1-like	276	gi|1228381282|ref|XP_021952145.1|myrosinase 1-like [Folsomia candida]gi|1215269331|gb|OXA55643.1|Lactase-phlorizin hydrolase [Folsomia candida]	Collembola	XP_021952145, OXA55643	3.37E-10	87.5	60.8474	32	28
Lep_00027730-RA	Myrosinase 1	303	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	6.27E-15	84.44444	73.559	45	38
Lep_00060394-RA	myrosinase 1	463	gi|1080062823|ref|XP_018573472.1|myrosinase 1, partial [Anoplophora glabripennis]	Coleoptera	XP_018573472	2.42E-29	73.40426	116.701	94	69

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00060393-RA	Myrosinase 1	423	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.51E-56	86.95652	182.956	115	100
Lep_00136832-RA	myrosinase 1-like	243	gi|1069789897|ref|XP_018320188.1|PREDICTED: myrosinase 1-like [Agrilus planipennis]	Coleoptera	XP_018320188	1.01E-16	84	78.9518	50	42
Lep_00101061-RA	myrosinase 1-like	393	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	2.29E-29	78.66667	112.849	75	59
Lep_00074338-RA	myrosinase 1	330	gi|1008445131|ref|XP_975665.2|PREDICTED: myrosinase 1 isoform X2 [Tribolium castaneum]	Coleoptera	XP_975665	1.15E-28	81.91489	113.62	94	77
Lep_00074339-RA	myrosinase 1-like	291	gi|1080062829|ref|XP_018573475.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573475	8.33E-22	76.1194	93.9745	67	51
Lep_00073232-RA	Myrosinase 1	621	gi|1090651209|gb|AOY34571.1|beta-glucosidase [Coptotermes formosanus]	Blattodea	AOY34571	1.77E-85	79.29293	266.929	198	157
Lep_00143227-RA	Myrosinase 1	168	gi|1233173996|ref|XP_022188302.1|myrosinase 1-like [Nilaparvata lugens]	Hemiptera	XP_022188302	1.33E-13	91.30435	68.9366	46	42

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00078589-RA	myrosinase 1-like	690	gi|1080062831|ref|XP_018573476.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573476	2.55E-54	78.19549	187.193	133	104
Lep_00042919-RA	myrosinase 1	1194	gi|1325357344|ref|XP_023312869.1|myrosinase 1 [Anoplophora glabripennis]	Coleoptera	XP_023312869	1.1E-116	72.60726	354.369	303	220
Lep_00045854-RA	myrosinase 1-like	480	gi|1228404666|ref|XP_021961442.1|myrosinase 1-like [Folsomia candida]	Collembola	XP_021961442	1.2E-15	41.13924	77.0258	158	65
Lep_00060792-RA	myrosinase 1-like	426	gi|668445331|gb|KFB35692.1|AGAP006424-PA-like protein [Anopheles sinensis]	Diptera	KFB35692	1.86E-15	84	77.7962	50	42
Lep_00020108-RA	myrosinase 1	526	gi|1247040849|gb|PCG75727.1|hypothetical protein B5V51_11090 [Heliothis virescens]	Lepidoptera	PCG75727	3.61E-12	49.0566	69.707	159	78
Lep_00020109-RA	myrosinase 1-like	1104	gi|939630683|ref|XP_014278711.1|PREDICTED: myrosinase 1-like [Halyomorpha halys]	Hemiptera	XP_014278711	5.59E-13	65.07937	76.6406	63	41
Lep_00113513-RA	myrosinase 1-like	477	gi|1339073035|ref|XP_023723774.1|myrosinase 1-like [Cryptotermes secundus]gi|1330878957|gb|PNF16841.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023723774, PNF16841	2.89E-39	80.64516	144.436	93	75

Table 3.20. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00058059-RA	myrosinase 1-like	318	gi|1008445131|ref|XP_975665.2|PREDICTED: myrosinase 1 isoform X2 [Tribolium castaneum]	Coleoptera	XP_975665	8.9E-29	81.91489	113.62	94	77
Lep_00058058-RA	myrosinase 1-like	513	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.35E-33	73.78641	125.561	103	76
Lep_00042495-RA	myrosinase 1-like	708	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	5.89E-73	74.30168	228.794	179	133
Lep_00042496-RA	myrosinase 1-like	480	gi|1080062831|ref|XP_018573476.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573476	4.63E-22	82.53968	97.4413	63	52
Lep_00054176-RA	myrosinase 1-like	1080	gi|1233170929|ref|XP_022187264.1|myrosinase 1-like isoform X2 [Nilaparvata lugens]	Hemiptera	XP_022187264	2.54E-51	76.87075	181.415	147	113
Lep_00093809-RA	myrosinase 1-like	465	gi|636793015|dbj|BAO85048.1|beta-glucosidase, partial [Salganea esakii]	Blattodea	BAO85048	1.26E-21	74.32432	93.9745	74	55
Lep_00092612-RA	myrosinase 1	717	gi|668445331|gb|KFB35692.1|AGAP006424-PA-like protein [Anopheles sinensis]	Diptera	KFB35692	4.36E-14	84	77.0258	50	42

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00042140-RA	myrosinase 1-like	234	gi|1227954946|ref|XP_021937620.1|myrosinase 1-like isoform X1 [Zootermopsis nevadensis]	Blattodea	XP_021937620	1.65E-19	85.41667	86.6557	48	41
Lep_00042138-RA	myrosinase 1	351	gi|1080062831|ref|XP_018573476.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_018573476	1.68E-21	81.96721	93.9745	61	50
Lep_00113948-RA	myrosinase 1-like	417	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	9.31E-31	70.4	116.701	125	88
Lep_00082283-RA	Myrosinase 1	612	gi|506965870|gb|AGM32287.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32287	4.89E-30	87.2093	117.857	86	75
Lep_00150133-RA	myrosinase 1-like	294	gi|1114635916|gb|APM84101.1|GH1 beta-glucosidase [Microcerotermes annandalei]	Blattodea	APM84101	8.15E-30	76.25	115.931	80	61
Lep_00106211-RA	myrosinase 1	438	gi|1042851809|gb|ANQ90707.1|beta-glucosidase 2, partial [Manduca quinquemaculata]	Lepidoptera	ANQ90707	4.14E-46	75.80645	151.754	124	94
Lep_00089193-RA	Myrosinase 1	780	gi|506965870|gb|AGM32287.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32287	1.83E-18	74.28571	88.5817	70	52

Table 3.20. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00033948-RA	Myrosinase 1	1259	gi|1339068038|ref|XP_023721112.1|myrosinase 1-like [Cryptotermes secundus]gi|1339068040|ref|XP_023721113.1|myrosinase 1-like [Cryptotermes secundus]gi|1330885855|gb|PNF19653.1|Myrosinase 1 [Cryptotermes secundus]gi|1330885856|gb|PNF19654.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023721112, XP_023721113, PNF19653, PNF19654	8.2E-136	67.45283	406.371	424	286
Lep_00024904-RA	Myrosinase 1	1413	gi|1233170926|ref|XP_022187263.1|myrosinase 1-like isoform X1 [Nilaparvata lugens]	Hemiptera	XP_022187263	4.3E-147	70.40816	434.106	392	276
Lep_00024905-RA	myrosinase 1-like	228	gi|1215275228|gb|OXA61537.1|Lactase-phlorizin hydrolase [Folsomia candida]	Collembola	OXA61537	3.85E-12	77.55102	65.855	49	38
Lep_00013204-RA	myrosinase 1-like	1312	gi|1279707180|ref|XP_022915768.1|myrosinase 1-like [Onthophagus taurus]	Coleoptera	XP_022915768	1.55E-97	72.42798	306.99	243	176
Lep_00013205-RA	myrosinase 1-like	372	gi|1080062338|ref|XP_018573207.1|myrosinase 1-like isoform X2 [Anoplophora glabripennis]	Coleoptera	XP_018573207	1.84E-20	75	91.2781	64	48
Lep_00097528-RA	myrosinase 1-like	761	gi|1090651209|gb|AOY34571.1|beta-glucosidase [Coptotermes formosanus]	Blattodea	AOY34571	3.36E-47	76.47059	169.474	119	91

Table 3.20. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00009457-RA	myrosinase 1-like	831	gi|1325349814|ref|XP_023310604.1|myrosinase 1-like [Anoplophora glabripennis]	Coleoptera	XP_023310604	6.7E-27	74.19355	114.775	93	69
Lep_00009460-RA	Myrosinase 1	912	gi|1330893407|gb|PNF23683.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	PNF23683	4.35E-29	58.20896	122.479	134	78
Lep_00072701-RA	myrosinase 1-like	921	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	7.56E-32	82	125.561	100	82
Lep_00011410-RA	myrosinase 1-like isoform X2	339	gi|506967929|gb|AGM32308.1|beta-glucosidase, partial [Coptotermes formosanus]	Blattodea	AGM32308	1.49E-37	70.53571	133.265	112	79
Lep_00011407-RA	myrosinase 1-like	618	gi|1090651209|gb|AOY34571.1|beta-glucosidase [Coptotermes formosanus]	Blattodea	AOY34571	8.8E-22	59.18367	98.5969	98	58
Lep_00011408-RA	myrosinase 1-like	318	gi|1061483168|gb|ODN00405.1|Lactase-phlorizin hydrolase [Orchesella cincta]	Collembola	ODN00405	2.78E-09	93.33333	58.9214	30	28

Table 3.20. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00011409-RA	myrosinase 1-like	423	gi|746857492|ref|XP_011059478.1|PREDICTED: myrosinase 1 isoform X2 [Acromyrmex echinatior]	Hymenoptera	XP_011059478	1.08E-41	81.52174	150.214	92	75
Lep_00011406-RA	myrosinase 1-like	474	gi|1339073035|ref|XP_023723774.1|myrosinase 1-like [Cryptotermes secundus]gi|1330878957|gb|PNF16841.1|Myrosinase 1 [Cryptotermes secundus]	Blattodea	XP_023723774, PNF16841	1.33E-08	68.18182	58.9214	44	30
Lep_00113088-RA	myrosinase 1-like	447	gi|1227984163|ref|XP_021923731.1|lactase-phlorizin hydrolase-like isoform X2 [Zootermopsis nevadensis]	Blattodea	XP_021923731	8.45E-67	81.63265	215.698	147	120
Lep_00102915-RA	myrosinase 1	318	gi|1339057505|ref|XP_023717252.1|myrosinase 1-like [Cryptotermes secundus]	Blattodea	XP_023717252	2.18E-06	92.30769	50.8322	26	24
Lep_00039189-RA	myrosinase 1	579	gi|1080052826|ref|XP_018567975.1|lactase-like protein [Anoplophora glabripennis]	Coleoptera	XP_018567975	1.85E-35	62.73292	135.576	161	101

Table 3.21. Lytic polysaccharide monooxygenases (LPMOs) in *Ctenolepisma longicaudata*: Coding sequences in *C. longicaudata* genome encoding for LPMOs and their blast description.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00097915-RA	Lytic polysaccharide monooxygenase	841	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.58E-87	75.72816	265.388	206	156
Lep_00074511-RA	Lytic polysaccharide monooxygenase	888	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	3.8E-133	92.05607	381.719	214	197
Lep_00127363-RA	Lytic polysaccharide monooxygenase	408	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.27E-12	66.66667	67.3958	57	38
Lep_00119247-RA	Lytic polysaccharide monooxygenase	383	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.08E-72	90.83333	221.476	120	109
Lep_00130110-RA	Lytic polysaccharide monooxygenase	237	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.9E-18	66.23377	80.4925	77	51
Lep_00039537-RA	Lytic polysaccharide monooxygenase	842	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.06E-85	70.8134	260.766	209	148

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0003953 8-RA	Lytic polysaccharide monooxygenase	834	gi|1349059415|emb|SIW61372 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 372	2.9E-67	69.61 326	213.7 72	181	126
Lep_0006835 4-RA	Lytic polysaccharide monooxygenase	1281	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 373	6.59E-51	68.87 417	176.7 92	151	104
Lep_0015107 8-RA	Lytic polysaccharide monooxygenase	168	gi|939644305|ref|XP_0142723 00.1|PREDICTED: uncharacterized protein LOC106678350 [Halyomorpha halys]	Hemiptera	XP_01 42723 00	2.78E-15	74.50 98	73.17 38	51	38
Lep_0009782 6-RA	Lytic polysaccharide monooxygenase	716	gi|1009534315|ref|XP_015904 140.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Arachnida	XP_01 59041 40	2.39E-83	66.96 429	253.8 32	224	150
Lep_0007768 0-RA	Lytic polysaccharide monooxygenase	876	gi|1349059415|emb|SIW61372 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 372	1.14E-96	75.81 395	289.2 71	215	163

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0008180 4-RA	Lytic polysaccharide monooxygenase	889	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 373	3.46E-62	67.33 668	201.4 45	199	134
Lep_0009778 8-RA	Lytic polysaccharide monooxygenase	466	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 373	2.08E-35	78.35 052	127.8 72	97	76
Lep_0009909 3-RA	Lytic polysaccharide monooxygenase	684	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 373	4.34E-77	71	236.8 84	200	142
Lep_0011857 7-RA	Lytic polysaccharide monooxygenase	597	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 373	1.64E-82	77.84 431	249.5 95	167	130
Lep_0002939 6-RA	Lytic polysaccharide monooxygenase	703	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 373	1.51E-81	80.74 534	248.4 4	161	130
Lep_0002939 5-RA	Lytic polysaccharide monooxygenase	228	gi|1349059415|emb|SIW61372 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 372	3.77E-11	91.17 647	61.61 78	34	31

Table 3.21. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_000029394-RA	Lytic polysaccharide monooxygenase	648	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	7.28E-54	77.3913	177.178	115	89
Lep_000061662-RA	Lytic polysaccharide monooxygenase	703	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.06E-81	80.74534	248.054	161	130
Lep_000061663-RA	Lytic polysaccharide monooxygenase	300	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	5.54E-16	90.47619	75.485	42	38
Lep_000043371-RA	Lytic polysaccharide monooxygenase	867	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	4.55E-80	73.09645	246.899	197	144
Lep_000043372-RA	Lytic polysaccharide monooxygenase	573	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.07E-66	73.05389	209.149	167	122
Lep_001220017-RA	Lytic polysaccharide monooxygenase	558	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.2E-75	75.5814	231.491	172	130

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0009384 5-RA	Lytic polysaccharide monooxygenase	848	gi|1349059415|emb|SIW61372 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.71E-98	79.24528	292.738	212	168
Lep_0006722 6-RA	Lytic polysaccharide monooxygenase	938	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.38E-62	67.33668	202.601	199	134
Lep_0009761 4-RA	Lytic polysaccharide monooxygenase	425	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.25E-35	78.35052	127.872	97	76
Lep_0009840 8-RA	Lytic polysaccharide monooxygenase	684	gi|1349059417|emb|SIW61373 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	6.3E-101	93.67089	297.36	158	148
Lep_0009931 8-RA	Lytic polysaccharide monooxygenase	660	gi|1349059415|emb|SIW61372 .1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	8.61E-97	76.27907	286.574	215	164

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00038947-RA	Lytic polysaccharide monooxygenase	876	gi|1316145279|ref|XP_023217685.1|uncharacterized protein LOC111620072 [Centruroides sculpturatus]gi|1316145289|ref|XP_023217690.1|uncharacterized protein LOC111620076 [Centruroides sculpturatus]	Arthropoda - Scorpiones	XP_023217685, XP_023217690	2.1E-78	70.95238	242.662	210	149
Lep_00038948-RA	Lytic polysaccharide monooxygenase	768	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.36E-74	62.24066	231.106	241	150
Lep_00106678-RA	Lytic polysaccharide monooxygenase	507	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	7.73E-54	77.3913	175.252	115	89
Lep_00066650-RA	Lytic polysaccharide monooxygenase	534	gi|1067063705|ref|XP_018015775.1|PREDICTED: uncharacterized protein LOC108672588 [Hyalella azteca]	Crustacean	XP_018015775	9.93E-55	72.46377	190.66	138	100
Lep_00018140-RA	Lytic polysaccharide monooxygenase	824	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	7.5E-91	70.75472	273.478	212	150

Table 3.21. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00136402-RA	Lytic polysaccharide monooxygenase	384	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	3.45E-54	77.3913	174.481	115	89
Lep_00109934-RA	Lytic polysaccharide monooxygenase	942	gi|1009534315|ref|XP_015904140.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Arachnida	XP_015904140	1.08E-68	70.55556	219.55	180	127
Lep_00121885-RA	Lytic polysaccharide monooxygenase	549	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.61E-51	74.4186	170.244	129	96
Lep_00072322-RA	Lytic polysaccharide monooxygenase	961	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.7E-100	78.30189	300.056	212	166
Lep_00119656-RA	Lytic polysaccharide monooxygenase	617	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	6.69E-94	81.08108	278.485	185	150
Lep_00071838-RA	Lytic polysaccharide monooxygenase	796	gi|1009534315|ref|XP_015904140.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Arachnida	XP_015904140	1.14E-65	73.56322	209.92	174	128

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0008301 6-RA	Lytic polysaccharide monooxygenase	619	gi|1349059417|emb|SIW61373.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	5.03E-62	60.39604	197.593	202	122
Lep_0003755 6-RA	Lytic polysaccharide monooxygenase	1029	gi|1349059415|emb|SIW61372.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.11E-49	68.66667	169.859	150	103
Lep_0003755 7-RA	Lytic polysaccharide monooxygenase	537	gi|1349059417|emb|SIW61373.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.98E-06	78.78788	52.373	33	26
Lep_0008328 3-RA	Lytic polysaccharide monooxygenase	853	gi|1349059415|emb|SIW61372.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	8.76E-70	74.4186	220.32	172	128
Lep_0003092 2-RA	Lytic polysaccharide monooxygenase	390	gi|1351638095|pdb|5MSZ|AChain A, Lytic Polysaccharide Monooxygenase AA15 from Thermobia domestica in the Cu(I) State		5MSZ_A	5.65E-10	80.4878	60.4622	41	33

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00030924-RA	Lytic polysaccharide monooxygenase	306	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	8.91E-39	79.41176	134.035	102	81
Lep_00028253-RA	Lytic polysaccharide monooxygenase	987	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	2.2E-100	78.77358	300.056	212	167
Lep_00028252-RA	Lytic polysaccharide monooxygenase	615	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.75E-75	73.51351	231.106	185	136
Lep_00028251-RA	Lytic polysaccharide monooxygenase	219	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.32E-11	93.54839	62.7734	31	29
Lep_00103234-RA	Lytic polysaccharide monooxygenase	753	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.33E-69	66.83673	218.009	196	131

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0006615 4-RA	Lytic polysaccharide monooxygenase	228	gi|1349059417|emb|SIW61373.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61 373	3.7E-11	87.87 879	61.61 78	33	29
Lep_0004517 1-RA	Lytic polysaccharide monooxygenase	835	gi|759036658|ref|XP_011345707 .1|PREDICTED: uncharacterized protein LOC105284120 [Ooceraea biroi]gi|607367294|gb|EZA6144 1.1|hypothetical protein X777_07774 [Ooceraea biroi]	Hymenoptera	XP_01 13457 07, EZA61 441	3.24E-48	71.97 452	165.2 36	157	113
Lep_0004517 2-RA	Lytic polysaccharide monooxygenase	357	gi|759036658|ref|XP_011345707 .1|PREDICTED: uncharacterized protein LOC105284120 [Ooceraea biroi]gi|607367294|gb|EZA6144 1.1|hypothetical protein X777_07774 [Ooceraea biroi]	Hymenoptera	XP_01 13457 07, EZA61 441	4.63E-35	66.08 696	125.5 61	115	76
Lep_0009944 1-RA	Lytic polysaccharide monooxygenase	724	gi|1009534315|ref|XP_01590414 0.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Arachnida	XP_01 59041 40	1.1E-66	73.56 322	211.4 6	174	128

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_00084324-RA	Lytic polysaccharide monooxygenase	883	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.61E-73	77.84091	230.335	176	137
Lep_00059868-RA	Lytic polysaccharide monooxygenase	618	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.75E-76	72.97297	234.572	185	135
Lep_00059867-RA	Lytic polysaccharide monooxygenase	295	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	1.49E-10	93.33333	60.8474	30	28
Lep_00156681-RA	Lytic polysaccharide monooxygenase	171	gi|939644305|ref|XP_014272300.1|PREDICTED: uncharacterized protein LOC106678350 [Halyomorpha halys]	Hemiptera	XP_014272300	7.34E-16	73.58491	74.7146	53	39
Lep_00072794-RA	Lytic polysaccharide monooxygenase	384	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	2.21E-54	78.26087	174.866	115	90

Table 3.21. Continued.

Sequence name	Sequence desc.	Sequence length	Hit desc.	Order	Hit ACC	E-Value	Similarity	Bit-Score	Alignment length	Positives
Lep_0011844 7-RA	Lytic polysaccharide monooxygenase	380	gi|1349059415|emb|SIW61372.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.53E-73	91.59664	223.402	119	109
Lep_0007870 7-RA	Lytic polysaccharide monooxygenase	588	gi|1349059415|emb|SIW61372.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	3.4E-99	88.04348	291.582	184	162
Lep_0007870 8-RA	Lytic polysaccharide monooxygenase	594	gi|1349059415|emb|SIW61372.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	4.25E-81	88.05031	245.743	159	140
Lep_0013784 4-RA	Lytic polysaccharide monooxygenase	471	gi|1349059415|emb|SIW61372.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.15E-49	75.75758	164.081	132	100
Lep_0006057 3-RA	Lytic polysaccharide monooxygenase	1038	gi|1349059415|emb|SIW61372.1 |Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	4.49E-74	73.51351	233.802	185	136

Table 3.21. Continued.

Sequence name	**Sequence desc.**	**Sequence length**	**Hit desc.**	**Order**	**Hit ACC**	**E-Value**	**Similarity**	**Bit-Score**	**Alignment length**	**Positives**
Lep_00006587-RA	Lytic polysaccharide monooxygenase	870	gi|1349059415|emb|SIW61372.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61372	1.05E-91	72.85714	276.559	210	153
Lep_00006590-RA	Lytic polysaccharide monooxygenase	941	gi|1349059417|emb|SIW61373.1|Lytic polysaccharide monooxygenase, partial [Thermobia domestica]		SIW61373	4.21E-55	56.52174	184.111	207	117
Lep_00006588-RA	Lytic polysaccharide monooxygenase	1363	gi|1238881740|ref|XP_022253834.1|uncharacterized protein LOC106469570 isoform X1 [Limulus polyphemus]gi|1238881742|ref|XP_022253835.1|uncharacterized protein LOC106469570 isoform X1 [Limulus polyphemus]	Arthropoda	XP_022253834, XP_022253835	6.63E-60	55.55556	204.912	288	160
Lep_00006589-RA	Lytic polysaccharide monooxygenase	335	gi|1009534315|ref|XP_015904140.1|uncharacterized protein LOC107436837 [Parasteatoda tepidariorum]	Arachnida	XP_015904140	1.22E-15	65.82278	75.485	79	52

Table 3.22. Number of predicted endogenously produced endoglucanases, β-glucosidases and β-glucuronidases in genomes of insects (highlighted indicates our work). Data on *Cryptocercus punctulatus* and *Reticulitermes flavipes* were collected from identical protein groups database of NCBI.

Insect species	**Endoglucanase**	**β-glucosidase**	**β-glucuronidase**
Blattodea			
Blatella germanica	2	0	4
Cryptocercus punctulatus	0	1	0
Reticulitermes flavipes	5	1	0
Zootermopsis nevadensis	5	2	4
Zygentoma			
Thermobia domestica	**85**	**19**	**39**
Ctenolepisma longicaudata	**69**	**22**	**30**
Coleoptera			
Tribolium castaneum	1	0	0
Dendroctonus ponderosae	10	5	3
Lepidoptera			
Bombyx mori	0	4	1
Manduca sexta	1	9	10
Plutella xylostella	2	14	1
Diptera			
Drosophila melanogaster	1	3	7
Bactrocera cucurbitae	0	0	8

Table 3.22. Continued.

Insect species	**Endoglucanase**	**β-glucosidase**	**β-glucuronidase**
Hemiptera			
Acyrthosiphon pisum	4	20	4
Nilaparvta lugens	7	13	1
Hymenoptera			
Apis dorsata	1	0	3
Megachile rotundata	3	2	4
Anoplura			
Pediculus humanus	1	1	1

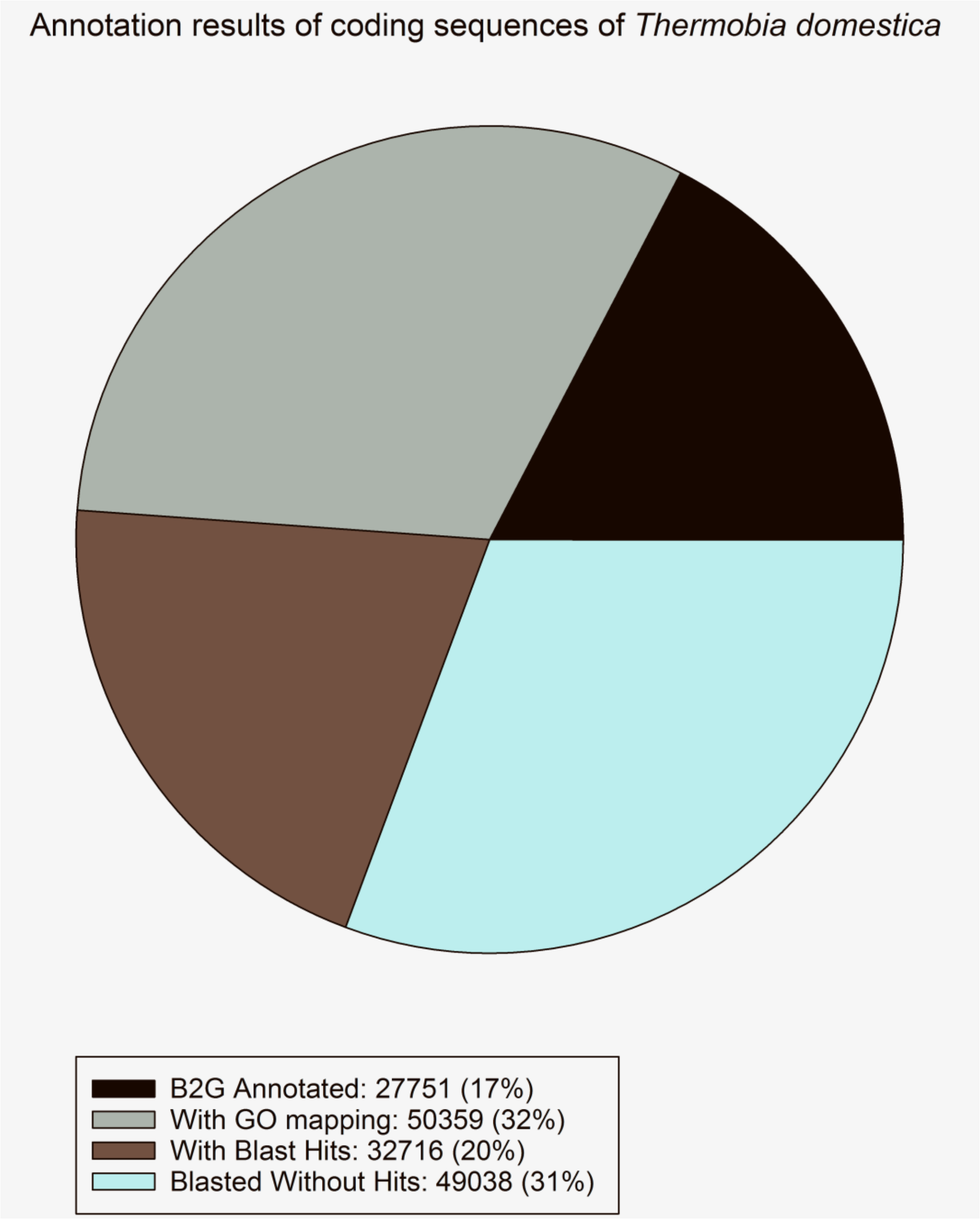

Figure 3.1. Annotation of *Thermobia domestica* coding sequences: Number and relative percentage of *T. domestica* coding sequences with blast hits, with no blast hits, annotated and mapped.

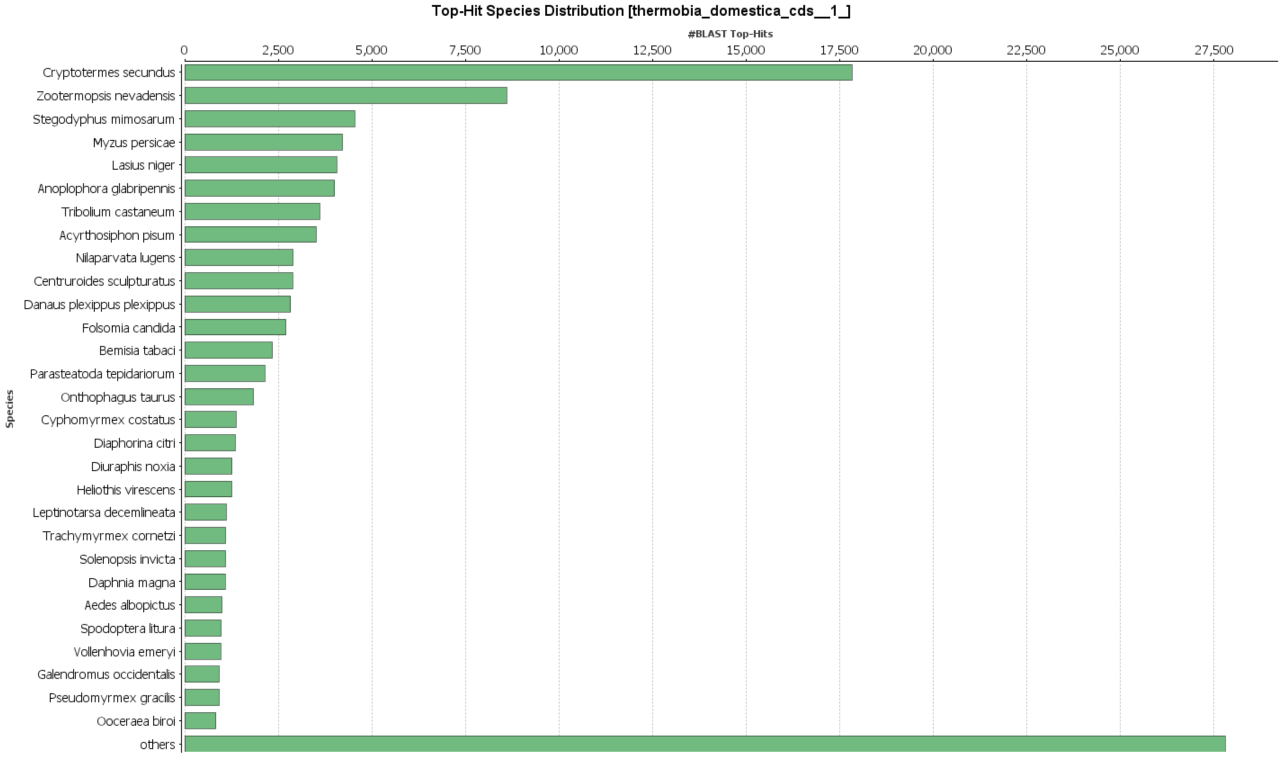

Figure 3.2. Top-hit species distribution for *Thermobia domestica*: Number of blast hits matching with highest identity to different species.

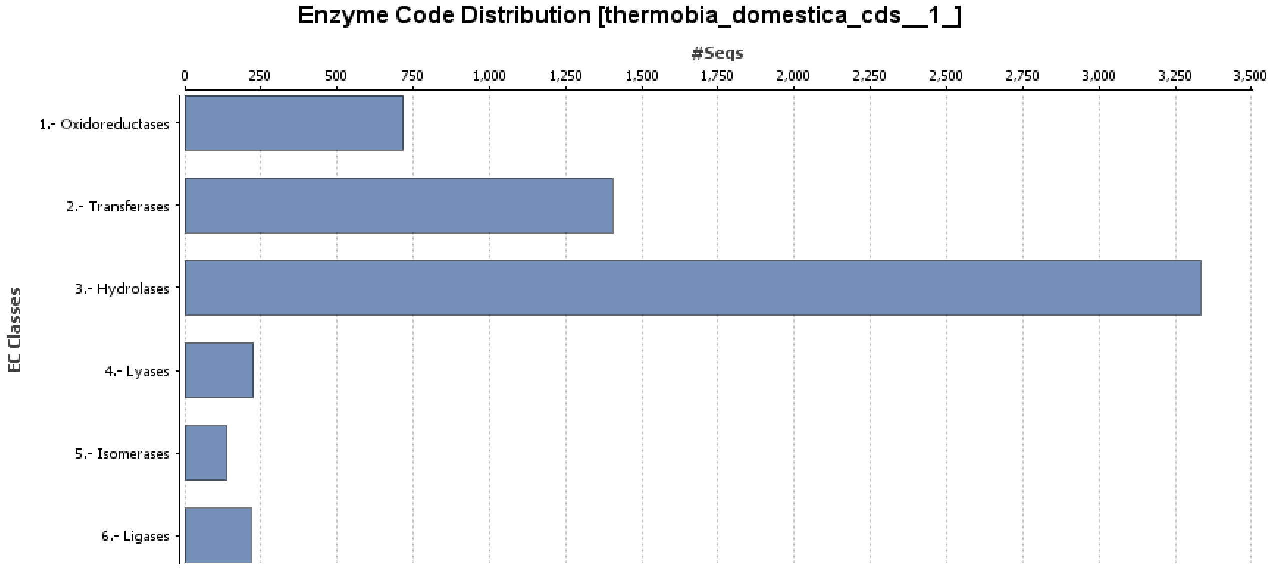

Figure 3.3. Enzyme Code distribution in *Thermobia domestica*: Number of coding sequences of *T. domestica* encoding for enzymes belonging to different enzyme commission classes (EC).

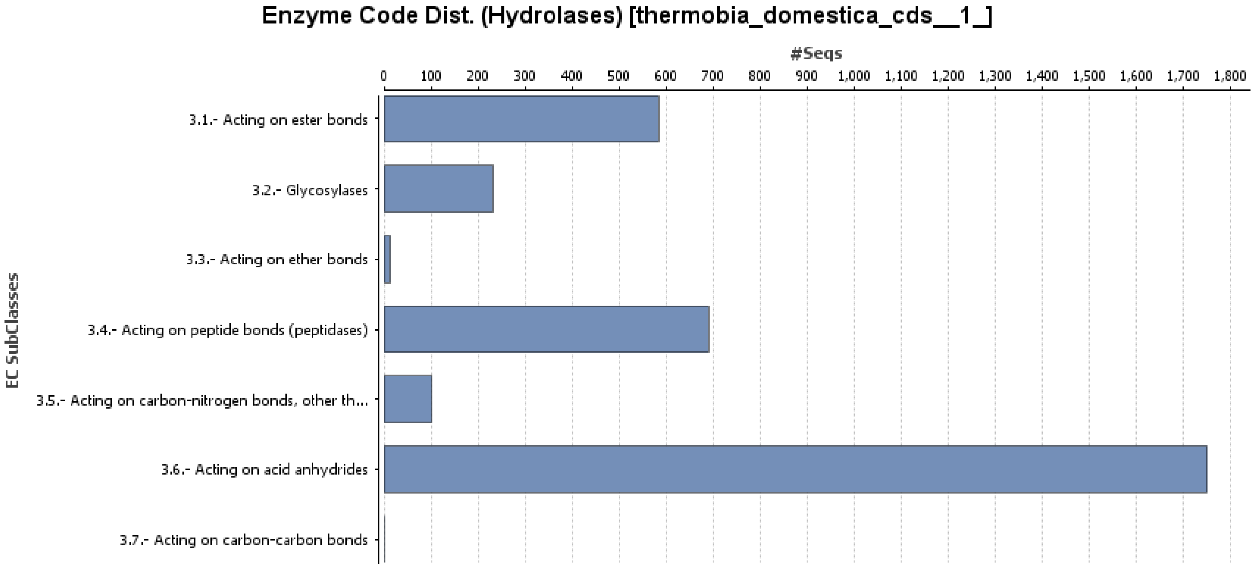

Figure 3.4. Enzyme Code distribution of hydrolases in *Thermobia domestica*: Number of coding sequences of *T. domestica* encoding for enzymes belonging to different enzyme commission (EC) subclasses of hydrolases.

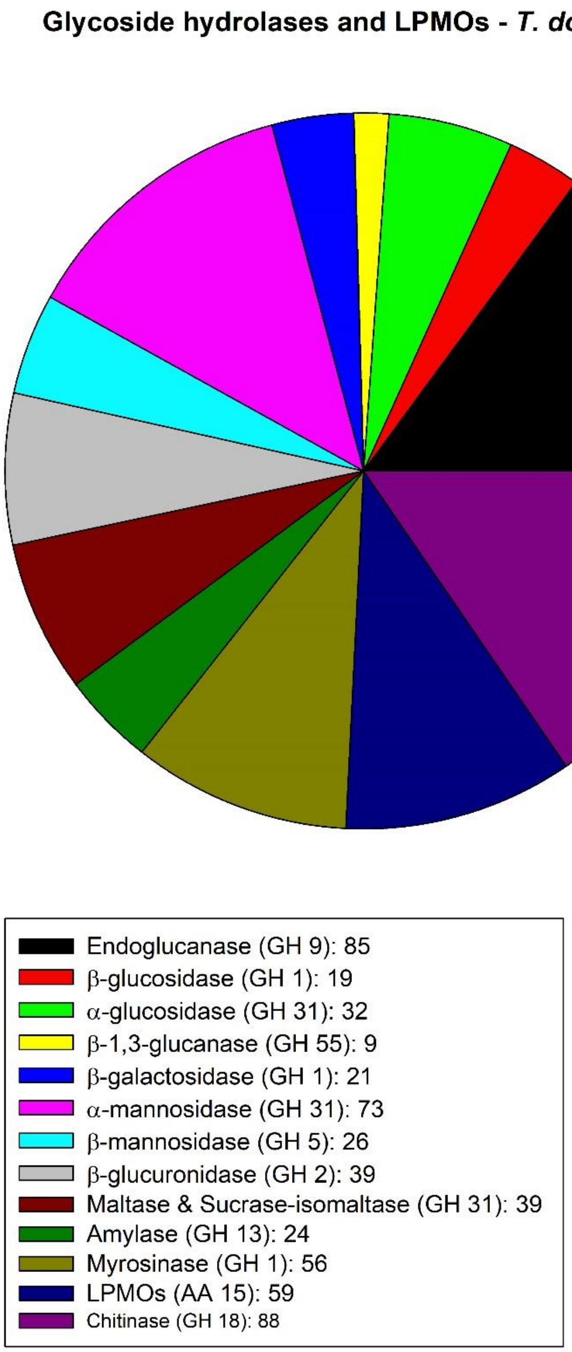

Figure 3.5. Glycoside hydrolases (GH) and lytic polysaccharide monoxygenases (LPMOs) in *Thermobia domestica*: Distribution of enzymes encoding for different glycoside hydrolases and LPMOs in *T. domestica*.

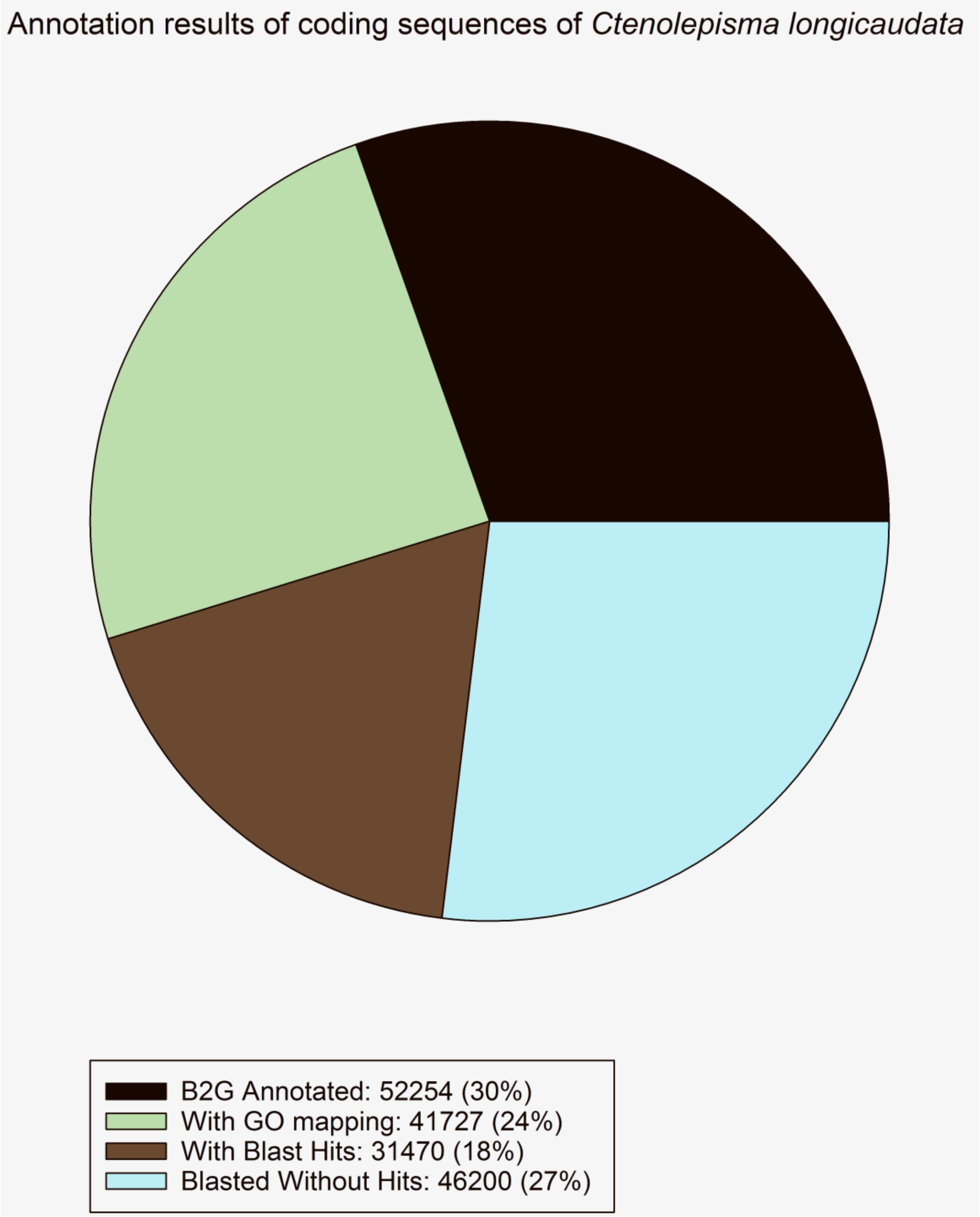

Figure 3.6. Annotation of *Ctenolepisma longicaudata* coding sequences: Number and percentage of *C. longicaudata* coding sequences with blast hits, with no blast hits, annotated and mapped.

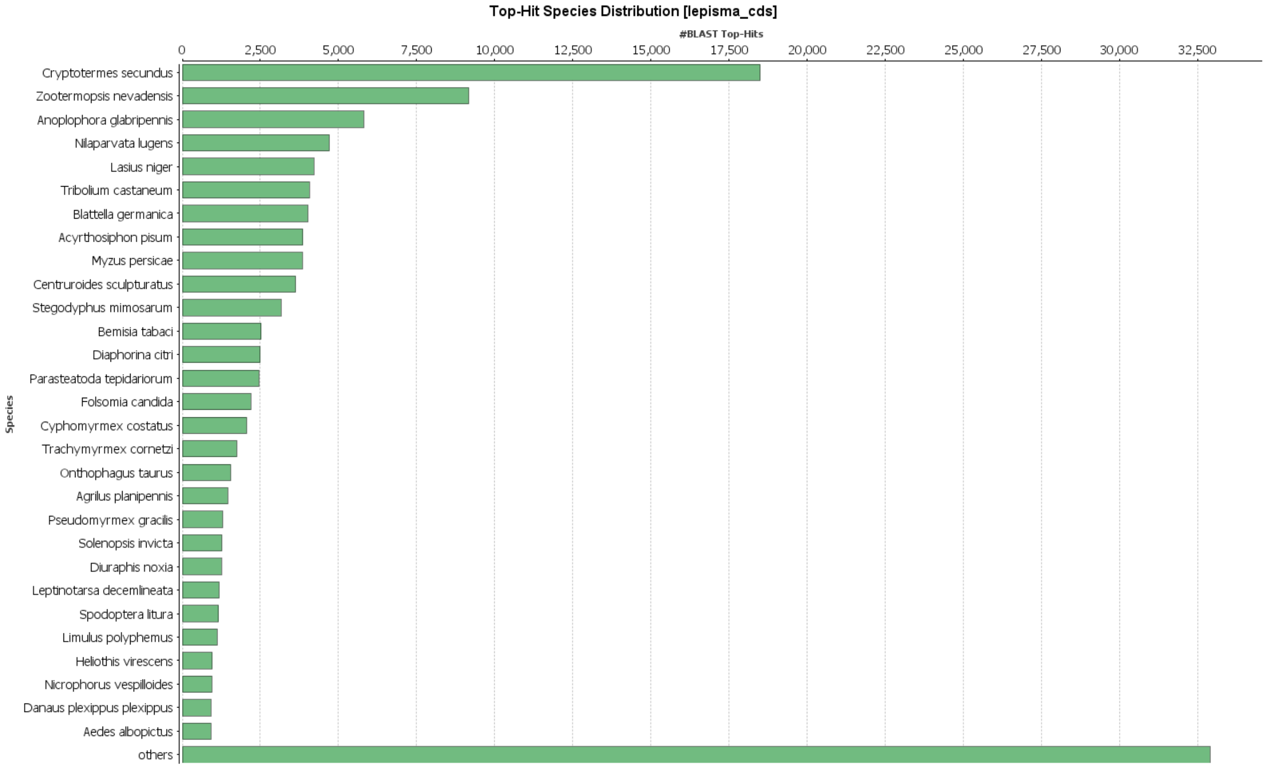

Figure 3.7. Top-hit species distribution for *Ctenolepisma longicaudata*: Number of blast hits with highest identity to different species.

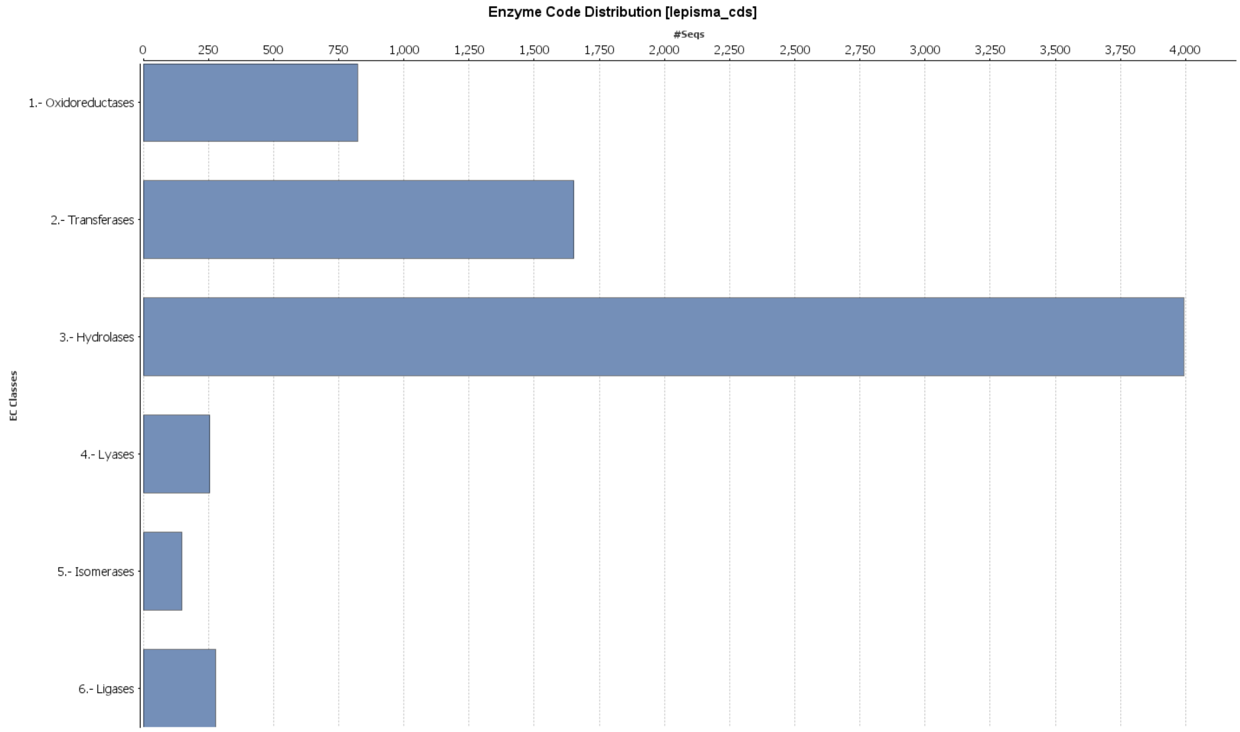

Figure 3.8. Enzyme Code distribution in *Ctenolepisma longicaudata*: Number of coding sequences of *C. longicaudata* encoding for enzymes belonging to different enzyme commission classes (EC).

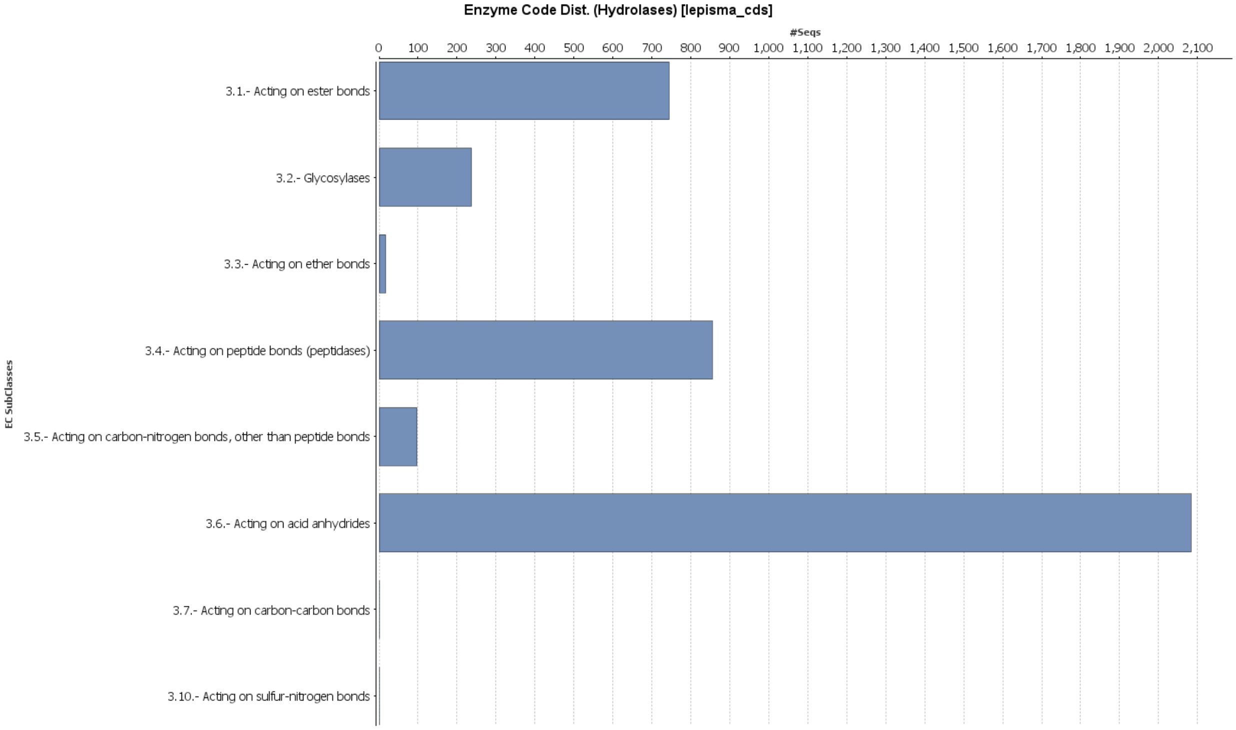

Figure 3.9. Enzyme Code distribution of hydrolases in *Ctenolepisma longicaudata*: Number of coding sequences of *C. longicaudata* encoding for enzymes belonging to different enzyme commission (EC) subclasses of hydrolases.

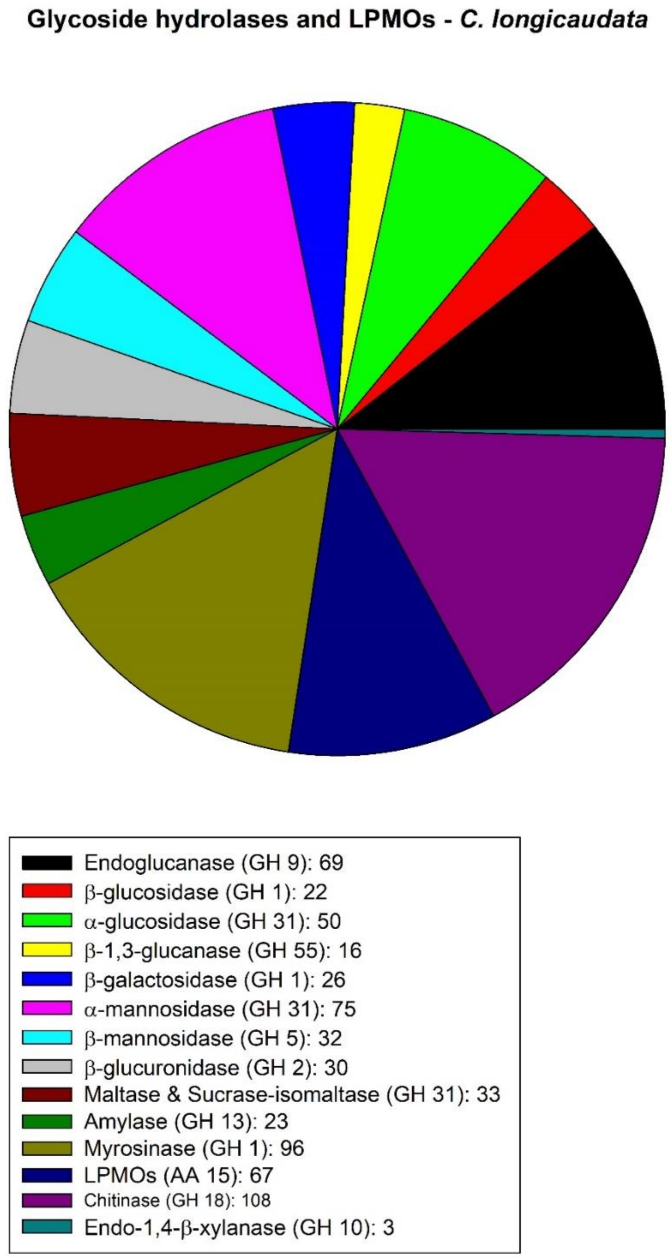

Figure 3.10. Glycoside hydrolases (GH) and lytic polysaccharide monoxygenases (LPMOs) in *Ctenolepisma longicaudata*: Distribution of enzymes encoding belong to different glycoside hydrolases and LPMOs in *C. longicaudata*.

Chapter 4

Differential expression of cellulose degrading enzyme genes in *Thermobia domestica* and *Ctenolepisma longicaudata* in response to diets of different cellulosic content

Pothula, R.; Johnson, B.R.; Klingeman, W.E.; Huff, M.; Staton, M.E. and J.L. Jurat-Fuentes. (2018).

My contributions included: (1) planning and performing experiments, (2) data collection and analysis, (3) writing the manuscript and making figures. Brian R. Johnson, Matthew Huff and Margaret Staton helped with 2 and Juan Luis Jurat-Fuentes assisted with (1, 2 and 3).

Abstract

Members of Zygentoma have been characterized as having the highest relative cellulolytic activity compared to traditional model insects used for biofuel research such as termites, cockroaches and beetles (Chapter 2). However, not much-information is available on cellulolytic genes in these non-model organisms. In the present work, our goal was to study the differential expression of cellulase genes in the foregut and other body tissues when *Thermobia domestica* and *Ctenolepisma longicaudata* were fed on four diets with varying degree of cellulosic content. Using an RNASeq approach, differential gene expression analysis of both species revealed that cellulase gene expression is primarily driven by type of tissue rather than diet. However, within each tissue of *T. domestica* and *C. longicaudata*, a higher number of plant cell wall degrading enzymes (PCWDEs) and lytic polysaccharide monoxygenases were significantly up-regulated in the paper diet treatment, which is highly cellulosic, compared to all other tested diets. The annotation of differentially expressed PCWDE genes revealed highest identity to insect homologs, which suggests the potential conservation of PCWDEs through evolution and the ancient origin of cellulases in insects. Overall, our research contributes to increasing the amount of

information available on functional PCWDE genes and lytic polysaccharide monoxygenases (LPMOs) from a primitive hexapod group with potential for industrial biofuel applications.

Introduction

Cellulases are a group of enzymes that completely digest plant cellulose into glucose, which in the biofuel industry can be fermented by yeast to generate bioethanol. Enzymatic activities within cellulases include endoglucanase, which cleaves the cellulose chain internally at random, exoglucanase that cleaves the cellulose chain at the ends releasing cellobiose, and β-glucosidase that degrades cellobiose into glucose units (Watanabe and Tokuda, 2010). These plant cell wall degrading enzymes (PCWDEs) have been traditionally described from microorganisms, yet in the last decade or so insects have also been considered a prospecting resource for endogenously produced PCWDEs, including cellulases, hemicellulases and pectinases (Calderón-Cortés et al., 2012; Watanabe and Tokuda, 2010). Endoglucanases and β-glucosidases are commonly found in insects, yet to date insect exoglucanases have not been reported (Martin, 1983; Scrivener and Slaytor, 1994). This may be due to the function of exoglucanases in insects being compensated by the presence of a higher number of endoglucanases with dual endo/exo activity (Scrivener and Slaytor, 1994), and/or physical processing by mandibles and proventriculus and/or a long digestive tube that allows for increased length of digestion in the gut (Calderón-Cortés et al., 2012; Watanabe and Tokuda, 2010). Hemicellulases and pectinases digest hemicellulose and pectin polysaccharides, respectively, which are interlocked with cellulose fibers in the plant cell wall (Gilbert, 2010). Insects belonging to 16 taxonomic orders were reported to have endogenous production of one or more of PCWDEs

(Calderón-Cortés et al., 2012). However, molecular evidence confirming the presence of PCWDE genes including endoglucanases belonging to glycoside hydrolase family (GH) 9 and 45 , β-glucosidases of GH 5, hemicellulases such as xyloglucanases of GH 5 and GH 11, β-1,3-glucanases of GH 16 and pectinases of GH 28, has only been obtained from insects belonging to only 8 taxonomic orders (Calderón-Cortés et al., 2012).

Whole transcriptome shotgun sequencing, or RNA-Seq, is a next generation sequencing technique allowing for quantitative determination of total transcripts present as a proxy for the level of expression of the corresponding gene in a cell, tissue or whole organism (Wang et al., 2009). Unlike microarrays and traditional sequencing technologies, RNA-Seq can be conveniently used to find functional genes, such as cellulases, in non-model organisms (Shelomi et al., 2014; Vera et al., 2008). Additionally, RNA-Seq studies can identify differentially expressed genes when comparing between different tissues and/or under distinct conditions.

Members of Zygentoma have been characterized as having the highest relative cellulolytic activity against carboxymethylcellulose compared to traditional model insects used for biofuel research such as termites, cockroaches and beetles (Chapter 2). However, there is a lack of information on the molecular characterization of cellulolytic activity in these non-model organisms. Previous work detected endoglucanase, β-glucosidase, xylanase, pectinase, amylase, maltase, sucrase and lactase activities in foregut fluids from *Thermobia domestica* and *Ctenolepisma longicaudata*through biochemical tests (Zinkler and Götze, 1987; Chapter 2). In addition to cellulase and xylanase activities, investigation of the digestive proteome of *T. domestica* revealed the production of lytic polysaccharide monoxygenases (LPMOs), which are enzymes that form nicks in cellulose fibers thereby making them more accessible to cellulases (Sabbadin et al., 2018).

In the present work, our goal was to find cellulase genes that are responsible for the high relative cellulolytic activity in *T. domestica* and *C. longicaudata* compared to other insects. In addition, we used RNAseq to study their differential expression in foregut and rest of the body samples when *T. domestica* and *C. longicaudata* were fed on four diets with varying degree of cellulosic content. Differential gene expression analysis of both species revealed that cellulase gene expression is primarily driven by type of tissue rather than diet. However, within each tissue of *T. domestica* and *C. longicaudata*, higher number of PCWDEs were significantly up-regulated in insects fed on the paper diet treatment, which is highly cellulosic compared to all other tested diets. Additionally, more LPMOs were up-regulated in the foregut tissue of paper-fed *T. domestica* than other diets, yet LPMO up-regulation was not as prominent in *C. longicaudata*, which may help explain higher cellulolytic activity in *T. domestica* compared to *C. longicaudata*. The annotation of differentially expressed PCWDE genes revealed highest identity to insect homologs, which suggests the potential conservation of PCWDEs through evolution. Overall, our research contributes to increasing the amount of information available on functional PCWDE genes and LPMOs from a primitive hexapod group with potential for industrial biofuel applications.

Materials and Methods

Insect rearing

Adult silverfish (*Ctenolepisma longicaudata*) and firebrat (*Thermobia domestica*) were used in this study. Several batches of nymphs and adults of *C. longicaudata* were hand-collected by sweeping into collection containers or collected into lid-less plastic dishes (Pioneer Plastics Inc., Dixon, KY) baited with whole-grain oat flakes (Quick 1-minute Oats, Quaker Oats Co., Chicago, IL) in buildings. These insects were then reared in the laboratory on rolled oats, paper and Pedigree adult complete nutrition dry dog food (Mars, Inc., Mount Olive, NJ) at room temperature. An established *T. domestica* colony which was originally derived from the culture being maintained at the Department of Entomology at Ohio State University (Columbus, OH) was kindly provided by Patrick Stanley and Eric Snell (Snell Scientific, LLC, Meansville, GA). This *T. domestica* colony had been maintained in the Department of Entomology and Plant Pathology at the University of Tennessee for > 8 years on printing paper as a carbohydrate resource and NatureWise chick starter grower feed (Nutrena, Minneapolis, MN) as protein source in a dark incubator set at 34^0C.

Diet treatments

Six insects per treatment were starved for five days in individual Petri dishes. On the sixth day, six randomly selected insects were provided with one of four diet treatments and allowed to feed for five days. Treatments included 0.2% carboxymethylcellulose (CMC) dissolved in 1% agar, standard 92 multipurpose printing paper (Georgia-Pacific, Atlanta, GA, USA), pulverized switchgrass (generously provided by Dr. Nicole Labbe, Center for Renewable Carbon, University of Tennessee), and a protein-balanced diet (NatureWise chick starter grower feed) as a control. A block of 1% agar was provided as a water source to all insects. Diet and agar blocks were changed every other day. Each treatment was replicated thrice.

Dissections, RNA extraction and sequencing

Three replicates of six fed insects per each treatment were anesthesized by incubation at 4^0C for 10 min. The foregut was quickly dissected from the rest of the body by pulling from the head with forceps and then cutting the head from the foregut with a disposable scalpel. Individual foreguts and rest of the body (carcass including head) were collected into separate microcentrifuge tubes, snap frozen in liquid nitrogen and stored at -80^0C. Total RNA was extracted from a pool of six foreguts or rest of the body samples using Direct-zol RNA MiniPrep Plus (Zymo Research, Irvine, CA, USA) and RNeasy Maxi kits (Qiagen, Hilden, Germany), respectively. The use of different kits depending on tissue was due to differences in tissue weights. Purity of total RNA was checked using NanoDrop and then submitted to the Genomic Sequencing and Analysis Facility at The University of Texas at Austin (Texas) for library preparation and sequencing. Total RNA quality was analyzed using a Bioanalyzer (Agilent Technologies, Santa Clara, CA), and then total RNA was enriched for poly-A mRNA and the resulting RNA was paired-end sequenced (three hundred base pairs) on a Hi-Seq 4000 (Illumina, Inc., San Diego, CA).

Differential gene expression analysis

A draft genome of *T. domestica* was previously sequenced and assembled (Brand et al., 2018). Similarly, a draft genome of *C. longicaudata* was assembled following the same bioinformatic pipeline as that used for *T. domestica* (B. Johnson, unpublished). Coding sequences from the draft genome of each insect were provided as fasta files by B. Johnson (University of California at Davis) and indexed using Salmon v0.8.2 (Patro et al., 2017). The abundance of RNA-seq reads in each insect was quantified using Salmon due to similar accuracy but increased speed when compared to similar quantification programs. The resulting read raw counts per transcript in each replicate were used as input to detect the differentially expressed genes between treatments using DESeq2 (Love et al., 2014). The DESeq2 program is an R package which uses negative binomial generalized linear models to test for differentially expressed genes (DEG) between treatments. Among the detected DEGs, significance was established at an adjusted *P*-value (P_{adj}) < 0.05. Principal component analysis plots and heatmaps were developed to visualize the differential expression of cellulase genes and LPMO's across treatments and tissues using "plotPCA" and "pheatmap" options in RStudio version 1.1.423 (RStudio, Inc. Boston, MA).

Annotation

To find the differentially expressed cellulase genes in *T. domestica* and *C. longicaudata*, all the DEGs ($P_{\mathrm{adj}} < 0.05$) between pair-wise treatment comparisons were annotated using a custom-made database in Blast2GO (https://www.blast2go.com/). This custom database included all protein sets from NCBI (accessed on January 4, 2018) with known cellulase activity in insects, bacteria, fungi and protozoa. All the resulting DEG PCWDEs and LPMOs were manually blasted in the NCBI website against the non-redundant (nr) protein database to get the most specific match (accessed on September 6, 2018).

Results

de novo transcriptome assemblies of *T. domestica* and *C. longicaudata*

Sequencing of RNA extracted from either pooled foregut or the remaining body tissues (rest of the body sample) of *T. domestica* and *C. longicaudata* resulted in 580,047,044 and 523,554,706 raw reads respectively. The *de novo* assembly of all reads of *T. domestica* and *C. longicaudata* by Trinity (Grabherr et al., 2011) yielded 1,368,850 and 666,881 contigs, respectively. Trinity *de novo* assembly of *T. domestica* and *C. longicaudata* also resulted in 1,121,220 and 542,714 genes, respectively.

Differential cellulase gene expression analysis in *T. domestica*

Comparisons of foregut tissue from *T. domestica* fed on CMC or switchgrass to protein diet did not detect up-regulation of any cellulase genes, yet PCWDEs were found to be up-regulated in the protein diet (Table 4.1). However, one endoglucanase (GH 9) and one β-galactosidase-1,3were significantly up-regulated in *T. domestica* foregut tissue fed on paper compared to protein diet (Table 4.1). In addition, five endoglucanases, one β-galactosidase, one β-glucuronidase and one glucosidase were significantly up-regulated in *T. domestica* foregut tissue fed on paper compared to switchgrass, while one endoglucanase was up-regulated in paper compared to CMC (Table 4.1).

In contrast to cellulase gene expression, differential gene expression analyses revealed significant up-regulation of LPMOs in foregut tissue of *T. domestica* fed on paper compared to other diets (Table 4.2). Notably, paper feeding *T. domestica* resulted in up-regulated 21, 17 and 3 transcripts encoding for LPMOs in the foregut tissue compared to *T. domestica* fed protein, switchgrass and CMC diets, respectively (Table 4.2). Similarly, seven transcripts encoding for LPMOs were up-regulated in the foregut tissue of *T. domestica* fed on CMC compared to protein and switchgrass diets (Table 4.2). However, LPMOs were not differentially expressed in the foregut tissue of *T. domestica* fed on switchgrass compared to protein diet (Table 4.2).

Unlike observations in the foregut tissue, cellulase gene expression in the rest of the body sample was increased as cellulosic content increased in the diet. The most striking difference was detected when feeding *T. domestica* on paper compared to protein diet, which resulted in up-regulation of 15 endoglucanase genes (Table 4.3). The rest of the body sample of *T. domestica* fed on paper diet had 14 endoglucanases that were up-regulated compared to *T. domestica* fed switchgrass, and 9 endoglucanases, one β-glucosidase, and one mannanase significantly up-regulated compared to a CMC diet. However, CMC and switchgrass diets resulted in significant up-regulation of only two endoglucanases compared to protein diet (Table 4.3).

In contrast to foregut tissue, the differential gene expression analyses of the rest of the body sample of *T. domestica* fed on different diets revealed that very few LPMOs were differentially expressed (Table 4.4). Feeding *T. domestica* on paper and switchgrass diets did not result in differential expression of LPMOs compared to protein diet. Nevertheless, CMC diet resulted in up-regulation of one and three LPMOs compared to protein and paper diets, respectively (Table 4.4). Additionally, feeding on a paper diet resulted in up-regulation of five transcripts encoding for LPMOs compared to switchgrass diet in rest of the body sample of *T. domestica* (Table 4.4). Among all significantly differentially expressed genes in different pair-wise comparisons, seven endoglucanases and seven LPMOs were commonly expressed in both foregut and rest of the body samples. Overall, differential gene expression analysis detected 35 significantly up-regulated cellulase genes in *T. domestica* fed on different diets (Table 4.9). In the same analyses, 26 and 9 transcripts encoding for LPMOs were significantly differentially expressed in different pair-wise comparisons among treatments in foregut and rest of the body samples, respectively (Table 4.2 and 4.4).

Annotation of differentially expressed PCWDE genes in *T. domestica*

Among the 35 genes differentially expressed, endoglucanases (26 coding sequences) were the most abundant cellulases, 22 of which matched endoglucanases of insect origin. Among the other four endoglucanases, two were most similar to molluscan cellulases, one matched to a copepod (*Eurytemora affinis*) cellulase and another was most identical to an enzyme from a cnidarian (*Orbicella faveolata*) (Table 4.9). In addition to endoglucanases, cellulases differentially expressed included four β-galactosidases matching to different organisms such as sawfly (*Cephus cinctus*), water flea (*Daphnia pulex*), fish (*Austrofundulus limnaeus*) and opossum (*Monodelphis domestica*); two β-glucuronidases matching to termites (*Zootermopsis nevadensis* and *Neotermes koshunensis*); two mannanases, one matching to a springtail (*Orchesella cincta*) and another matching to freshwater crayfish (*Cherax quadricarinatus*); and one transcript of a β-glucosidase in GH 31 matching to the cockroach, *Periplaneta americana* (Table 4.9).

Differential cellulase gene expression analysis in *C. longicaudata*

Similar to results from *T. domestica*, differential gene expression analysis in *C. longicaudata* foregut tissue fed on CMC diet did not result in detection of differential expression of any cellulase genes compared to the insects fed on protein diet (Table 4.5). However, in contrast to *T. domestica*, feeding *C. longicaudata* on switchgrass resulted in up-regulation of one endoglucanase (GH 9), one β-galactosidase, and two β-glucuronidases in the foregut tissue compared to protein diet (Table 4.5). Additionally, 13 endoglucanases were significantly up-regulated in the foregut tissue of *C. longicaudata* fed on paper diet compared to protein diet. Similarly, 12 genes encoding for endoglucases were up-regulated in *C. longicaudata* foregut tissue fed on paper compared to CMC diet and one mannosidase encoding gene was up-regulated in CMC diet fed *C. longicaudata* compared to protein diet (Table 4.5). Furthermore, feeding *C. longicaudata* on paper resulted in up-regulation of five endoglucanase encoding genes and one β-1,3-glucanase compared to the switchgrass treatment (Table 4.5).

Similar to cellulases, differential gene expression analysis of LPMO genes revealed that foregut tissue of *C. longicaudata* fed on CMC did not have any differentially expressed LPMO genes compared to that fed on protein diet. However, 11 and three LPMO encoding genes were significantly up-regulated in the foregut tissue of *C. longicaudata* fed on paper and switchgrass, respectively, when compared to protein diet (Table 4.6). Feeding of *C. longicaudata* on CMC and switchgrass resulted in up-regulation of only one LPMO gene in the foregut tissue compared to feeding on paper and CMC, respectively (Table 4.6).

Unlike in foregut tissue, differential gene expression analysis on the rest of the body sample of *C. longicaudata* revealed that feeding *C. longicaudata* on CMC resulted in differential cellulase gene expression compared to protein diet. One glucosidase was up-regulated in the rest of the body sample of *C. longicaudata* fed on CMC compared to protein diet, while two endoglucanase encoding genes, two β-galactosidases and nine glucosidases were up-regulated in the reciprocal comparison (Table 4.7). Similarly, seven genes encoding for endoglucanases and one glucosidase were up-regulated on paper diet compared to protein. When comparing protein to paper diet, two endoglucanases, two glucosidases, and one galactosidase were up-regulated (Table 4.7). Likewise, four endoglucanase encoding genes, three β-glucosidases, one β-1,3-glucanase, and one gene encoding for an enzyme in GH 65 were up-regulated in *C. longicaudata* fed on switchgrass compared to protein. In the reciprocal comparison, we detected four glucosidases, two endoglucanases, and one β-galactosidase as up-regulated (Table 4.7). Feeding *C. longicaudata* on switchgrass up-regulated the expression of 15 glucosidases, nine endoglucanases, five β-glucuronidases, one β-1,3-glucanase, one β-galactosidase, one β-1,6-glucanase, and one GH 65 enzyme in the rest of the body sample when compared to feeding on CMC. In contrast, CMC up-regulated the expression of only three endoglucanases compared to switchgrass. However, in contrast to other pair-wise diet comparisons, *C. longicaudata* fed a paper diet up-regulated all cellulase encoding genes, 16 endoglucanases, two β-glucosidases, and one β-1,6-glucanase compared to feeding on CMC (Table 4.7). Additionally, 16 endoglucanases were up-regulated in rest of the body sample of paper fed *C. longicaudata* compared to feeding on switchgrass. In the reciprocal comparison, only one mannosidase, one glucosidase and one β-galactosidase were up-regulated in switchgrass fed *C. longicaudata* rest of the body sample compared to paper diet (Table 4.7).

LPMOs were not significantly differentially expressed in *C. longicaudata* rest of the body sample fed on CMC and paper compared to protein diet. However, six LPMO encoding genes were significantly up-regulated in paper diet compared to CMC (Table 4.8). Similarly, the expression of 18, six and five LPMOs were significantly up-regulated in *C. longicaudata* rest of the body sample in insects fed on switchgrass compared to CMC, protein and paper diets, respectively (Table 4.8). Overall, 16 genes encoding cellulases and six LPMO encoding genes were commonly expressed in both foregut and rest of the body samples of *C. longicaudata*.

Annotation of differentially expressed PCWDE genes in *C. longicaudata*

Annotation of all significantly differentially expressed cellulase genes across all fedding treatments in *C. longicaudata* yielded 70 genes encoding for cellulases. Similar to *T. domestica*, the majority (31 sequences) of the differentially expressed cellulase genes in *C. longicaudata* encoded for β-1,4-endoglucanases, 30 of which matched to proteins of insect origin and one was most similar to a sea anemone (*Nematostella vectensis*) (Table 4.10). After endoglucanases, the most common differentially expressed cellulases were β-glucosidases (18 genes), 13 of which matched to insect enzymes, while 3 were most similar to molluscans, one was most similar to a fish (*Acanthochromis polyacanthus*), and another was most similar to the crown-of-thorns starfish (*Acanthaster planci*) (Table 4.10). In addition to β-1,4-endoglucanases, two sequences encoding for β-1,6-glucanases, which were most similar to molluscan genes, and two β-1,3-glucanases, one similar to a coleopteran (*Dendroctonus ponderosae*) and one most similar to a crustacean (*Daphnia pulex*) gene, were found. Additionally, one mannanase most similar to a crustacean (*Daphnia magna*) genes; six α-glucosidases which matched to other insect homologs; six β-glucuronidases matching to insect or other arthropod genes; and five β-galactosidases matching to insects were found (Table 4.10). Of the 84 LPMO encoding sequences annotated in the *C. longicaudata* genome, only 15 and 25 LPMO encoding sequences were significantly differentially expressed in foregut and rest of the body samples across different pair-wise treatment comparisons.

Discussion

In our previous work (Chapter 2), we found Zygentoma displayed highest relative cellulase activity compared to other model insects for biofuel research such as termites, cockroaches and beetles. In addition, digestive fluids from both *T. domestica* and *C. longicaudata* were found to have endoglucanase, xylanase, β-glucosidase and pectinase activities, which are responsible for efficient digestion of cellulose into glucose. Additionally, higher CMCase activity was found in foregut compared to other regions of the digestive system in both species. In the present work, we investigated the endogenous PCWDEs and their differential expression in *T. domestica* and *C. longicaudata* foregut and rest of the body samples in response to diets with varying cellulose content through RNA-Seq.

Differential expression of cellulase genes in foregut and rest of the body samples of *T. domestica* and *C. longicaudata* fed on protein, paper, CMC and switchgrass diets was primarily driven by tissue type rather than diet (Fig. 4.1 and 4.2). Thus, in most cases cellulase gene expression was localized to a particular tissue (foregut versus rest of the body), independently of the diet. However, almost all of the up-regulated genes in foregut tissue were encoding for endoglucanases, while the majority of the up-regulated genes in the rest of the body sample encoded for β-glucosidases. This observation indicates the compartmentalization of cellulose digestion in both *T. domestica* and *C. longicaudata*, as proposed for other insects (Fischer et al., 2013) (Fig. 4.1 and 4.2 & Table 4.9 and 4.10). According to this model, the long cellulose chain is broken down into smaller pieces by endoglucanases up-regulated in the foregut tissue, which is followed by further digestion of intermediate products into glucose by β-glucosidases, which were up-regulated in the rest of the body samples that include midgut and hindgut tissues.

Similar to cellulases, differential LPMO gene expression in *T. domestica* and *C. longicaudata* was also mostly driven by the type of tissue rather than diet, although differences in expression were detected also when comparing between some diets (Fig. 4.3 and 4.5). Nevertheless, almost all significantly differentially expressed LPMOs were up-regulated in foregut tissue of *T. domestica* and *C. longicaudata* compared to the rest of the body samples (Figs. 4.4 and 4.6). The up-regulation of LPMOs in foregut tissue is in agreement with the function of LPMOs in attacking long cellulose polymers to form nicks in cellulose polymers and make them more accessible to cellulases (Villares et al., 2017).

Although cellulase gene expression was primarily driven by tissue type, within each tissue the cellulase gene expression was regulated by type of diet. Differential expression of cellulase genes among different pair-wise treatment combinations within the foregut tissue revealed that cellulase encoding genes were up-regulated in *T. domestica* fed paper compared to diets with protein, switchgrass and CMC (Table 4.1). This result indicates a correlation between cellulase gene up-regulation and highly cellulosic paper diet. Compared to foregut tissue, the differential expression of cellulases in the rest of the body was more firmly correlated with type of diet. For example, more PCWDEs were up-regulated in the rest of the body sample of *T. domestica* fed on paper compared to other diets (Table 4.3), which indicates the more recalcitrant nature of cellulose present in the paper compared to other diets.

Similarly, more LPMOs were up-regulated in foregut tissue of *T. domestica* fed on paper compared to all other diets (Table 4.2). This result also indicates the need for production of more enzymes to digest recalcitrant paper diet. Additionally, LPMOs were also up-regulated *T. domestica* fed CMC compared to protein and switchgrass diets, and no LPMOs were differentially expressed between *T. domestica* fed protein and switchgrass (Table 4.2). These observations may be explained by the protein and switchgrass diets being a mixture of carbohydrate and protein, whereas the CMC diet is composed of only carbohydrate whereby insects may need to consume and digest more CMC to meet their energy demands. Similar up-regulation of LPMOs was reported in gut of *T. domestica* fed on avicel versus less cellulosic diets (Sabbadin et al., 2018). Compared to foregut tissue, very few LPMOs were significantly up-regulated in different pair-wise treatment comparisons of the rest of the body sample (Table 4.4), which affirms the need for LPMO production within foregut tissue versus other regions of the digestive system.

Unlike in *T. domestica*, cellulase gene expression in *C. longicaudata* foregut tissue was correlated with type of diet. Among all pair-wise treatment combinations, more PCWDEs were up-regulated in the foregut tissue of *C. longicaudata* fed on paper compared to other diets (Table 4.5), which indicates the up-regulation of cellulase genes in response to the most recalcitrant form of cellulose among the tested diets. Similarly, switchgrass diet also resulted in up-regulated expression of a few cellulases compared to protein diet (Table

4.5), which again may be related to the more recalcitrant nature of switchgrass versus a protein diet. In contrast to *T. domestica* and foregut tissue of *C. longicaudata*, the cellulase gene expression in the rest of the body of *C. longicaudata* was regulated by both diets within a pair-wise treatment comparison (Table 4.7). This observation may indicate that cellulase gene expression in the rest of the body sample, which includes midgut and hindgut, is controlled by the type of diet. Conversely, very few LPMOs were up-regulated in foregut tissue of *C. longicaudata* (Tables 4.2 and 4.6), which could potentially help explain lower relative cellulolytic activity of *C. longicaudata* compared to *T. domestica* (Chapter 2).

Overall, the annotation of significantly differentially expressed cellulase genes identified 35 and 70 coding sequences encoding for cellulases in *T. dom*estica and *C. longicaudata*, respectively, across tissues and treatments (Table 4.9 and 4.10). Both *T. domestica* and *C. longicaudata* were found expressing numerous endoglucanases and β-glucosidases, which are considered as the main cellulase complex in insects for breakdown of complex cellulose into glucose (Calderón-Cortés et al., 2012). Additionally, *T. domestica* and *C. longicaudata* express 35 and 40 LPMOs, respectively, which probably are involved in the higher cellulolytic activity in this group compared to other insects (Chapter 2). However, *C. longicaudata* was found to express a higher number and also a more diverse pool of PCWDEs and LPMOs compared to *T. domestica*. In contrast, *T. domestica* was found to display higher endoglucanase, xylanase and pectinase activities than *C. longicaudata* (Chapter 2). The higher activity in *T. domestica* with fewer enzyme genes expressed may be explained by the correlation of type of diet and the expression of LPMOs in foregut tissue, which was more prominent in *T. domestica*.

Most of the annotated PCWDEs in these Zygentoma species matched with highest identity to proteins from termites and cockroaches, which may be indicative of close evolutionary relationships between these groups. Moreover, some of the PCWDEs in *T. domestica* and *C. longicaudata* were most identical to the most primitive hexapod groups, such as Collembola (*Orchesella cincta*), or to highly evolved insect groups such as hymenopterans, which may suggest the conservation of PCWDEs in insects through evolution rather than frequent horizontal transfer from microorganisms (Calderón-Cortés et al., 2012). However, a considerable number of PCWDEs matched to organisms from other phyla, emphasizing the need for in-depth molecular characterization of PCWDEs in all insect groups to shed insight on the evolution of cellulases in insects. Overall, our research provides additional information on functional PCWDEs and LPMOs in *T. domestica* and *C. longicaudata* as two representative members of a primitive hexapod group with high relative cellulase activity.

List of references

Abo-Khatwa, N., 1978. Cellulase of fungus-growing termites: A new hypothesis on its origin. Experientia 34, 559–560. https://doi.org/10.1007/BF01936956

Ali, S.S., Wu, J., Xie, R., Zhou, F., Sun, J., Huang, M., 2017. Screening and characterizing of xylanolytic and xylose-fermenting yeasts isolated from the wood-feeding termite, Reticulitermes chinensis. PLOS ONE 12, e0181141. https://doi.org/10.1371/journal.pone.0181141

Barnhart, C.S., 1961. The Internal Anatomy of the Silverfish Ctenolepisma campbelli and Lepisma saccharinum (Thysanura: Lepismatidae). Ann. Entomol. Soc. Am. 54, 177–196. https://doi.org/10.1093/aesa/54.2.177

Berger, B., 1945. Food preferences of the firebrat. J. Econ. Entomol. 38, 577–582. https://doi.org/10.1093/jee/38.5.577

Biedermann, W., Moritz, P., 1898. Beitr¨age zur vergeleichenden Physiologie der Verdauung. II. Uver ein cellulosel ¨ odendes Enzyme im Leversecret der Schnecke (Helix pomatia). Pfludg Arch Gem 73, 219–287.

Bones, A.M., Rossiter, J.T., 1996. The myrosinase-glucosinolate system, its organisation and biochemistry. Physiol. Plant. 97, 194–208. https://doi.org/10.1111/j.1399-3054.1996.tb00497.x

Brand, P., Robertson, H.M., Lin, W., Pothula, R., Klingeman, W.E., Jurat-Fuentes, J.L., Johnson, B.R., 2018. The origin of the odorant receptor gene family in insects [WWW Document]. eLife. https://doi.org/10.7554/eLife.38340

Brennan, Y., Callen, W.N., Christoffersen, L., Dupree, P., Goubet, F., Healey, S., Hernández, M., Keller, M.C., Li, K., Palackal, N., Sittenfeld, A., Tamayo, G., Wells, S., Hazlewood, G.P., Mathur, E.J., Short, J.M., Robertson, D.E., Steer, B.A., 2004. Unusual microbial xylanases from insect guts. Appl. Environ. Microbiol. 70, 3609–3617.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H., Oyama, K., 2012. Endogenous Plant Cell Wall Digestion: A Key Mechanism in Insect Evolution. Annu. Rev. Ecol. Evol. Syst. 43, 45–71. https://doi.org/10.1146/annurev-ecolsys-110411-160312

Cazemier, A.E., den Camp, H.J.M.O., Hackstein, J.H.P., Vogels, G.D., 1997. Fibre Digestion in Arthropods. Comp. Biochem. Physiol. A Physiol. 118, 101–109. https://doi.org/10.1016/S0300-9629(96)00443-4

Chatterjee, S., Sharma, S., Prasad, R.K., Datta, S., Dubey, D., Meghvansi, M.K., Vairale, M.G., Veer, V., 2015. Cellulase enzyme based biodegradation of cellulosic materials: an overview. South Asian J. Exp. Biol. 5, 271–282.

Cleveland, L.R., 1924. The physiological and symbiotic relationships between the intestinal protozoa of termites and their host, with special reference to reticulitermes flavipes kollar. Biol. Bull. 46, 178–201. https://doi.org/10.2307/1536507

Davison, A., Blaxter, M., 2005. Ancient Origin of Glycosyl Hydrolase Family 9 Cellulase Genes. Mol. Biol. Evol. 22, 1273–1284. https://doi.org/10.1093/molbev/msi107

Fischer, R., Ostafe, R., Twyman, R.M., 2013. Cellulases from Insects, in: Vilcinskas, A. (Ed.), Yellow Biotechnology II: Insect Biotechnology in Plant Protection and Industry, Advances in Biochemical Engineering/Biotechnology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 51–64. https://doi.org/10.1007/10_2013_206

Gilbert, H.J., 2010. The Biochemistry and Structural Biology of Plant Cell Wall Deconstruction. Plant Physiol. 153, 444–455. https://doi.org/10.1104/pp.110.156646

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883

Husebye, H., Arzt, S., Burmeister, W.P., Härtel, F.V., Brandt, A., Rossiter, J.T., Bones, A.M., 2005. Crystal structure at 1.1Å resolution of an insect myrosinase from Brevicoryne brassicae shows its close relationship to β-glucosidases. Insect Biochem. Mol. Biol. 35, 1311–1320. https://doi.org/10.1016/j.ibmb.2005.07.004

Kim, N., Choo, Y.M., Lee, K.S., Hong, S.J., Seol, K.Y., Je, Y.H., Sohn, H.D., Jin, B.R., 2008. Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket Teleogryllus emma. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 150, 368–376. https://doi.org/10.1016/j.cbpb.2008.04.005

Laibach, H., 1952. Lepisma saccharina L., das Silberflschchen. Z F Hyg Zool U Schddlingsbekdmpfung 40, 321–370.

Lasker, R., Giese, A.C., 1956. Cellulose Digestion By The Silverfish Ctenolepisma Lineata. J. Exp. Biol. 33, 542–553.

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Dae Sohn, H., Jin, B.R., 2004. cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, Apriona germari. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 139, 107–116. https://doi.org/10.1016/j.cbpc.2004.06.015

Lindsay, E., 1940. The Biology of the Silverfish, Ctenolepisma longicaudata Esch. with particular Reference to its feeding Habits. Proc. R. Soc. Vic. 52.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Lynd, L.R., Cushman, J.H., Nichols, R.J., Wyman, C.E., 1991. Fuel Ethanol from Cellulosic Biomass. Science 251, 1318–1323. https://doi.org/10.1126/science.251.4999.1318

Martin, M.M., 1991. The evolution of cellulose digestion in insects. Phil Trans R Soc Lond B 333, 281–288. https://doi.org/10.1098/rstb.1991.0078

Martin, M.M., Martin, J.S., 1978. Cellulose Digestion in the Midgut of the Fungus-Growing Termite Macrotermes natalensis: The Role of Acquired Digestive Enzymes. Science 199, 1453–1455. https://doi.org/10.1126/science.199.4336.1453

Modder, W.W.D., 1975. Feeding and growth of Acrotelsa collaris (Fabricius) (Thysanura, Lepismatidae) on different types of paper. J. Stored Prod. Res. 11, 71–74. https://doi.org/10.1016/0022-474X(75)90042-9

Oppert, C., Klingeman, W.E., Willis, J.D., Oppert, B., Jurat-Fuentes, J.L., 2010. Prospecting for cellulolytic activity in insect digestive fluids. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 155, 145–154. https://doi.org/10.1016/j.cbpb.2009.10.014

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., Kingsford, C., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419. https://doi.org/10.1038/nmeth.4197

Pauchet, Y., Heckel, D.G., 2013. The genome of the mustard leaf beetle encodes two active xylanases originally acquired from bacteria through horizontal gene transfer. Proc. R. Soc. Lond. B Biol. Sci. 280, 20131021. https://doi.org/10.1098/rspb.2013.1021

Rouland, C., Civas, A., Renoux, J., Petek, F., 1988. Purification and properties of cellulases from the termite Macrotermes mülleri (termitidae, macrotermitinae) and its symbiotic fungus termitomyces sp. Comp. Biochem. Physiol. Part B Comp. Biochem. 91, 449–458. https://doi.org/10.1016/0305-0491(88)90004-1

Rouland, C., Lenoir, F., Lepage, M., 1991. The role of the symbiotic fungus in the digestive metabolism of several species of fungus-growing termites. Comp. Biochem. Physiol. A Physiol. 99, 657–663. https://doi.org/10.1016/0300-9629(91)90146-4

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K., Jones, T.A., 1990. Three-dimensional structure of cellobiohydrolase II from Trichoderma reesei. Science 249, 380–386. https://doi.org/10.1126/science.2377893

Sabbadin, F., Hemsworth, G.R., Ciano, L., Henrissat, B., Dupree, P., Tryfona, T., Marques, R.D.S., Sweeney, S.T., Besser, K., Elias, L., Pesante, G., Li, Y., Dowle, A.A., Bates, R., Gomez, L.D., Simister, R., Davies, G.J., Walton, P.H., Bruce, N.C., McQueen-Mason, S.J., 2018. An ancient family of lytic polysaccharide monooxygenases with roles in arthropod development and biomass digestion. Nat. Commun. 9, 756. https://doi.org/10.1038/s41467-018-03142-x

Sahrhage, D., 1953. Okologische Untersuchengen an Thermobia domestica (Packard) und Lepisma saccharina L. Z Wiss Zool 157, 77–168.

Scrivener, A.M., Slaytor, M., 1994. Properties of the endogenous cellulase from Panesthia cribrata saussure and purification of major endo-β-1,4-glucanase components. Insect Biochem. Mol. Biol. 24, 223–231. https://doi.org/10.1016/0965-1748(94)90001-9

Shi, W., Ding, S.-Y., Yuan, J.S., 2011. Comparison of Insect Gut Cellulase and Xylanase Activity Across Different Insect Species with Distinct Food Sources. BioEnergy Res. 4, 1–10. https://doi.org/10.1007/s12155-010-9096-0

Terra, W.R., 1990. Evolution of Digestive Systems of Insects. Annu. Rev. Entomol. 35, 181–200. https://doi.org/10.1146/annurev.en.35.010190.001145

Terra, W.R., Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. Comp. Biochem. Physiol. Part B Comp. Biochem. 109, 1–62. https://doi.org/10.1016/0305-0491(94)90141-4

Terry, M.D., Whiting, M.F., 2005. Mantophasmatodea and phylogeny of the lower neopterous insects. Cladistics 21, 240–257. https://doi.org/10.1111/j.1096-0031.2005.00062.x

Tomme, P., Warren, R.A.J., Gilkes, N.R., 1995. Cellulose Hydrolysis by Bacteria and Fungi, in: Poole, R.K. (Ed.), Advances in Microbial Physiology. Academic Press, pp. 1–81. https://doi.org/10.1016/S0065-2911(08)60143-5

Treves, D.S., Martin, M.M., 1994. Cellulose digestion in primitive hexapods: Effect of ingested antibiotics on gut microbial populations and gut cellulase levels in the

firebrat, Thermobia domestica (Zygentoma, Lepismatidae). J. Chem. Ecol. 20, 2003–2020. https://doi.org/10.1007/BF02066239

Ververis, C., Georghiou, K., Christodoulakis, N., Santas, P., Santas, R., 2004. Fiber dimensions, lignin and cellulose content of various plant materials and their suitability for paper production. Ind. Crops Prod. 19, 245–254. https://doi.org/10.1016/j.indcrop.2003.10.006

Villares, A., Moreau, C., Bennati-Granier, C., Garajova, S., Foucat, L., Falourd, X., Saake, B., Berrin, J.-G., Cathala, B., 2017. Lytic polysaccharide monooxygenases disrupt the cellulose fibers structure. Sci. Rep. 7, 40262. https://doi.org/10.1038/srep40262

Watanabe, H., Noda, H., Tokuda, G., Lo, N., 1998. A cellulase gene of termite origin. Nat. Lond. 394, 330–1. http://dx.doi.org/10.1038/28527

Watanabe, H., Tokuda, G., 2010. Cellulolytic Systems in Insects. Annu. Rev. Entomol. 55, 609–632. https://doi.org/10.1146/annurev-ento-112408-085319

Willis, J.D., Oppert, B., Oppert, C., Klingeman, W.E., Jurat-Fuentes, J.L., 2011. Identification, cloning, and expression of a GHF9 cellulase from Tribolium castaneum (Coleoptera: Tenebrionidae). J. Insect Physiol. 57, 300–306. https://doi.org/10.1016/j.jinsphys.2010.11.019

Willis, J.D., Oppert, C., Jurat-Fuentes, J.L., 2010. Methods for discovery and characterization of cellulolytic enzymes from insects. Insect Sci. 17, 184–198. https://doi.org/10.1111/j.1744-7917.2010.01322.x

Woodbury, N., Gries, G., 2013a. Firebrats, Thermobia domestica, aggregate in response to the microbes Enterobacter cloacae and Mycotypha microspora. Entomol. Exp. Appl. 147, 154–159. https://doi.org/10.1111/eea.12054

Woodbury, N., Gries, G., 2013b. How Firebrats (Thysanura: Lepismatidae) Detect and Nutritionally Benefit From Their Microbial Symbionts Enterobacter cloacae and Mycotypha microspora. Environ. Entomol. 42, 860–867. https://doi.org/10.1603/EN13104

Zhao, H., Baker, G.A., Song, Z., Olubajo, O., Crittle, T., Peters, D., 2008. Designing enzyme-compatible ionic liquids that can dissolve carbohydrates. Green Chem. 10, 696. https://doi.org/10.1039/b801489b

Zhao, H., Jones, C.L., Baker, G.A., Xia, S., Olubajo, O., Person, V.N., 2009. Regenerating cellulose from ionic liquids for an accelerated enzymatic hydrolysis. J. Biotechnol. 139, 47–54. https://doi.org/10.1016/j.jbiotec.2008.08.009

Zhu, Z., Sathitsuksanoh, N., Vinzant, T., Schell, D.J., McMillan, J.D., Zhang, Y.-H.P., 2009. Comparative study of corn stover pretreated by dilute acid and cellulose solvent-based lignocellulose fractionation: Enzymatic hydrolysis, supramolecular structure, and substrate accessibility. Biotechnol. Bioeng. 103, 715–724. https://doi.org/10.1002/bit.22307

Zinkler, D., Götze, M., 1987. Cellulose digestion by the firebrat Thermobia domestica. Comp. Biochem. Physiol. Part B Comp. Biochem. 88, 661–666. https://doi.org/10.1016/0305-0491(87)90360-9

Appendix 4

Table 4.1. Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in foregut tissue of *Thermobia domestica.* Differentially expressed PCWDE genes in different pair-wise treatment comparisons of foregut tissue of *T. domestica* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{adj} < 0.05$).

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
CMC vs Protein			
Coding sequence			
Th_d_00035294-RA	hydrolase family 9	Protein	0.0003
Th_d_00006652-RA	β-galactosidase	Protein	0.0003
Th_d_00006654-RA	β-galactosidase	Protein	0.03
Th_d_00006653-RA	β galactosidase	Protein	0.01
Paper vs Protein			
Coding sequence			
Th_d_00001122-RA	β-galactosidase-1 3	Paper	0.02
Th_d_00093510-RA	hydrolase family 9	Paper	0.03
Th_d_00048538-RA	uncharacterized family 31 glucosidase KIAA1161-like	Paper	0.005
Switchgrass vs Protein			
Coding sequence			
Th_d_00000352-RA	hydrolase family 9	Protein	<0.001
Th_d_00046715-RA	hydrolase family 9	Protein	0.02
Th_d_00028499-RA	hydrolase family 9	Protein	0.0001
Th_d_00000351-RA	endo- β-1,4-glucanase	Protein	0.001
Th_d_00036630-RA	Man5-K	Protein	0.02
Th_d_00006653-RA	β-galactosidase	Protein	0.04
Th_d_00006652-RA	β-galactosidase	Protein	0.03
CMC vs Paper			
Coding sequence			
Th_d_00031500-RA	Endoglucanase	Paper	0.03
Switchgrass vs Paper			
Coding sequence			
Th_d_00001122-RA	β-galactosidase-1 3	Paper	0.05
Th_d_00071154-RA	hydrolase family 9	Paper	0.03
Th_d_00093510-RA	hydrolase family 9	Paper	0.001
Th_d_00061520-RA	hydrolase family 9	Paper	0.02

Table 4.1. Continued.

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Th_d_00005705-RA	hydrolase family 9	Paper	0.001
Th_d_00018098-RA	β-glucuronidase-like	Paper	0.04
Th_d_00048538-RA	uncharacterized family 31 glucosidase KIAA1161-like	Paper	0.02
Th_d_00034779-RA	hydrolase family 9	Paper	0.01

Table 4.2. Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in foregut tissue of *Thermobia domestica.* Differentially expressed of LPMO genes in different pair-wise treatment comparisons of foregut tissue of *T. domestica* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{adj} < 0.05$).

Diets tested	**Up-regulated diet**	**P_{adj} value**
CMC vs Protein		
Coding sequence		
Th_d_00014400-RA	CMC	< 0.001
Th_d_00056727-RA	CMC	0.03
Th_d_00064622-RA	CMC	0.02
Th_d_00068335-RA	CMC	0.02
Th_d_00109123-RA	CMC	0.002
Th_d_00110473-RA	CMC	0.04
Th_d_00120034-RA	CMC	0.03
Paper vs Protein		
Coding sequence		
Th_d_00014400-RA	Paper	0.005
Th_d_00023793-RA	Paper	< 0.001
Th_d_00032707-RA	Paper	0.02
Th_d_00056072-RA	Paper	0.01
Th_d_00056739-RA	Paper	0.03
Th_d_00056740-RA	Paper	0.003
Th_d_00058359-RA	Paper	0.04
Th_d_00064622-RA	Paper	< 0.001
Th_d_00068335-RA	Paper	< 0.001
Th_d_00068347-RA	Paper	0.03
Th_d_00071020-RA	Paper	0.002
Th_d_00083583-RA	Paper	< 0.001
Th_d_00084210-RA	Paper	0.005
Th_d_00092785-RA	Paper	0.04
Th_d_00095724-RA	Paper	0.03
Th_d_00108306-RA	Paper	0.005
Th_d_00109123-RA	Paper	0.01
Th_d_00110473-RA	Paper	0.01
Th_d_00119312-RA	Paper	0.02
Th_d_00120034-RA	Paper	0.04
Th_d_00120998-RA	Paper	<0.001
Paper vs CMC		
Coding sequence		
Th_d_00007075-RA	Paper	0.04

Table 4.2. Continued.

Diets tested	**Up-regulated diet**	**P_{adj} value**
Th_d_00084210-RA	Paper	0.04
Th_d_00120998-RA	Paper	0.03
Switchgrass vs CMC		
Coding sequence		
Th_d_00014400-RA	CMC	< 0.001
Th_d_00064622-RA	CMC	0.001
Th_d_00098297-RA	CMC	0.04
Th_d_00104344-RA	CMC	0.005
Th_d_00109123-RA	CMC	0.003
Th_d_00110473-RA	CMC	0.03
Th_d_00120034-RA	CMC	0.03
Paper vs Switchgrass		
Coding sequence		
Th_d_00014400-RA	Paper	0.04
Th_d_00037466-RA	Paper	<0.001
Th_d_00056739-RA	Paper	<0.001
Th_d_00056740-RA	Paper	0.01
Th_d_00064622-RA	Paper	0.002
Th_d_00071020-RA	Paper	<0.001
Th_d_00083583-RA	Paper	0.04
Th_d_00092785-RA	Paper	0.002
Th_d_00095665-RA	Paper	0.003
Th_d_00095724-RA	Paper	0.02
Th_d_00098297-RA	Paper	0.03
Th_d_00108306-RA	Paper	<0.001
Th_d_00109123-RA	Paper	0.004
Th_d_00110473-RA	Paper	0.04
Th_d_00119312-RA	Paper	0.03
Th_d_00120034-RA	Paper	0.006
Th_d_00120998-RA	Paper	0.01

Table 4.3. Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in rest of the body sample of *Thermobia domestica.* Differentially expressed PCWDE genes in different pair-wise treatment comparisons of rest of the body sample of *T. domestica* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{adj} < 0.05$).

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
CMC vs Protein			
Coding sequence			
Th_d_00005705-RA	hydrolase family 9	CMC	0.02
Th_d_00104189-RA	hydrolase family 9	CMC	0.02
Paper vs Protein			
Coding sequence			
Th_d_00071154-RA	hydrolase family 9	Paper	< 0.001
Th_d_00029874-RA	hydrolase family 9	Paper	0.003
Th_d_00041503-RA	Endoglucanase	Paper	0.006
Th_d_00000348-RA	Endoglucanase	Paper	0.002
Th_d_00018984-RA	hydrolase family 9	Paper	0.04
Th_d_00045439-RA	β-1,4-glucanase	Paper	0.008
Th_d_00031500-RA	Endoglucanase	Paper	0.003
Th_d_00015659-RA	hydrolase family 9	Paper	0.04
Th_d_00007226-RA	β-1,4-endoglucanase 1	Paper	0.02
Th_d_00111221-RA	hydrolase family 9	Paper	0.02
Th_d_00005705-RA	hydrolase family 9	Paper	< 0.001
Th_d_00104189-RA	hydrolase family 9	Paper	< 0.001
Th_d_00046715-RA	hydrolase family 9	Paper	< 0.001
Th_d_00028499-RA	hydrolase family 9	Paper	< 0.001

Table 4.3. Continued

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Th_d_00034779-RA	hydrolase family 9	Paper	< 0.001
Switchgrass vs Protein			
Coding sequence			
Th_d_00104189-RA	hydrolase family 9	Switchgrass	0.02
Th_d_00035324-RA	Endoglucanase	Switchgrass	0.001
Paper vs CMC			
Coding sequence			
Th_d_00035294-RA	hydrolase family 9	Paper	0.04
Th_d_00071154-RA	hydrolase family 9	Paper	0.002
Th_d_00029874-RA	hydrolase family 9	Paper	0.01
Th_d_00018632-RA	β-glucosidase	Paper	0.04
Th_d_00000353-RA	hydrolase family 9	Paper	< 0.001
Th_d_00107645-RA	hydrolase family 9	Paper	< 0.001
Th_d_00018984-RA	hydrolase family 9	Paper	0.004
Th_d_00045439-RA	β-1,4-glucanase	Paper	< 0.001
Th_d_00101434-RA	Mannanase	Paper	0.01
Th_d_00007226-RA	β-1,4-endoglucanase 1	CMC	0.003
Th_d_00034779-RA	hydrolase family 9	Paper	0.007
Switchgrass vs CMC			
Coding sequence			
Th_d_00107645-RA	hydrolase family 9	Switchgrass	0.02

Table 4.3. Continued.

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Th_d_00035324-RA	Endoglucanase	Switchgrass	< 0.001
Paper vs Switchgrass			
Coding sequence			
Th_d_00071154-RA	hydrolase family 9	Paper	< 0.001
Th_d_00029874-RA	hydrolase family 9	Paper	< 0.001
Th_d_00118343-RA	hydrolase family 9	Paper	0.04
Th_d_00041503-RA	Endoglucanase	Paper	< 0.001
Th_d_00000348-RA	Endoglucanase	Paper	< 0.001
Th_d_00000353-RA	hydrolase family 9	Paper	< 0.001
Th_d_00000354-RA	hydrolase family 9	Paper	0.04
Th_d_00018984-RA	hydrolase family 9	Paper	0.01
Th_d_00007226-RA	β-1,4-endoglucanase 1	Paper	< 0.001
Th_d_00111221-RA	hydrolase family 9	Paper	0.004
Th_d_00005705-RA	hydrolase family 9	Paper	0.01
Th_d_00028499-RA	hydrolase family 9	Paper	< 0.001
Th_d_00038486-RA	hydrolase family 9	Paper	0.002
Th_d_00034779-RA	hydrolase family 9	Paper	< 0.001

Table 4.4. Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in rest of the body sample of *Thermobia domestica.* Differentially expressed LPMO genes in different pair-wise treatment comparisons of rest of the body sample of *T. domestica* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{\text{adj}} < 0.05$).

Diets tested	**Up-regulated diet**	**P_{adj} value**
CMC vs Protein		
Coding sequence		
Th_d_00037467-RA	CMC	0.03
Paper vs CMC		
Coding sequence		
Th_d_00056739-RA	CMC	< 0.001
Th_d_00068347-RA	Paper	0.04
Th_d_00083583-RA	CMC	0.04
Th_d_00120998-RA	CMC	0.02
Paper vs switchgrass		
Coding sequence		
Th_d_00056739-RA	Switchgrass	< 0.001
Th_d_00056740-RA	Paper	0.001
Th_d_00058638-RA	Paper	0.02
Th_d_00068347-RA	Paper	0.005
Th_d_00083583-RA	Switchgrass	0.02
Th_d_00092785-RA	Switchgrass	0.03
Th_d_00112968-RA	Switchgrass	0.004
Th_d_00120998-RA	Switchgrass	0.001

Table 4.5. Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in foregut tissue of *Ctenolepisma longicaudata.* Differentially expressed PCWDE genes in different pair-wise treatment comparisons of foregut tissue of *C. longicaudata* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{adj} < 0.05$).

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Paper vs Protein			
Coding sequence			
Lep_00006775-RA	hydrolase family 9	Paper	< 0.001
Lep_00022398-RA	hydrolase family 9	Paper	0.02
Lep_00034744-RA	hydrolase family 9	Paper	0.04
Lep_00036184-RA	hydrolase family 9	Paper	0.002
Lep_00044891-RA	hydrolase family 9	Paper	0.001
Lep_00051683-RA	hydrolase family 9	Paper	0.005
Lep_00051684-RA	hydrolase family 9	Paper	0.002
Lep_00052326-RA	hydrolase family 9	Paper	0.003
Lep_00060202-RA	hydrolase family 9	Paper	0.001
Lep_00071132-RA	hydrolase family 9	Paper	< 0.001
Lep_00071754-RA	hydrolase family 9	Paper	0.005
Lep_00115213-RA	hydrolase family 9	Paper	0.008
Lep_00135466-RA	hydrolase family 9	Paper	< 0.001
Switchgrass vs Protein			
Coding sequence			
Lep_00006757-RA	β-galactosidase-1 3	Switchgrass	0.04
Lep_00028614-RA	β-glucuronidase	Switchgrass	0.02
Lep_00055855-RA	β-glucuronidase	Switchgrass	0.001
Lep_00096399-RA	hydrolase family 9	Protein	0.02
Paper vs CMC			
Coding sequence			
Lep_00006775-RA	hydrolase family 9	Paper	< 0.001
Lep_00012530-RA	Mannanase	CMC	0.04
Lep_00022398-RA	hydrolase family 9	Paper	0.002
Lep_00034832-RA	endo- β-1,4-glucanase	Paper	0.02
Lep_00036184-RA	hydrolase family 9	Paper	< 0.001
Lep_00044891-RA	hydrolase family 9	Paper	< 0.001
Lep_00060202-RA	hydrolase family 9	Paper	< 0.001
Lep_00071132-RA	hydrolase family 9	Paper	0.001
Lep_00071461-RA	hydrolase family 9	Paper	< 0.001
Lep_00078550-RA	hydrolase family 9	Paper	< 0.001

Table 4.5. Continued.

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Lep_00104311-RA	endo- β-1,4-glucanase	Paper	0.04
Lep_00115213-RA	hydrolase family 9	Paper	< 0.001
Lep_00135466-RA	hydrolase family 9	Paper	< 0.001
Switchgrass vs Paper			
Coding sequence			
Lep_00029521-RA	endo- β-1,4-glucanase	Paper	0.005
Lep_00036184-RA	hydrolase family 9	Paper	0.005
Lep_00039163-RA	β-1,3-glucan-binding -like	Paper	< 0.001
Lep_00071461-RA	hydrolase family 9	Paper	< 0.001
Lep_00078550-RA	hydrolase family 9	Paper	< 0.001
Lep_00135466-RA	hydrolase family 9	paper	< 0.001

Table 4.6. Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in foregut tissue of *Ctenolepisma longicaudata*. Differentially expressed lytic polysaccharide monoxygenases (LPMOs) genes in different pair-wise treatment comparisons of foregut tissue of *C. longicaudata* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{adj} < 0.05$).

Diets tested	**Up-regulated diet**	**P_{adj} value**
Paper vs Protein		
Coding sequence		
Lep_00018140-RA	Paper	0.04
Lep_00029394-RA	Paper	0.03
Lep_00037556-RA	Paper	< 0.001
Lep_00039537-RA	Paper	0.006
Lep_00083016-RA	Paper	0.02
Lep_00084324-RA	Paper	0.02
Lep_00097915-RA	Paper	0.03
Lep_00103234-RA	Paper	0.02
Lep_00106678-RA	Paper	< 0.001
Lep_00118447-RA	Paper	0.005
Lep_00136402-RA	Paper	0.04
Switchgrass vs Protein		
Coding sequence		
Lep_00083016-RA	Switchgrass	0.005
Lep_00097788-RA	Switchgrass	0.04
Lep_00119656-RA	Switchgrass	0.04
Paper vs CMC		
Coding sequence		
Lep_00110038-RA	CMC	0.04
Switchgrass vs CMC		
Coding sequence		
Lep_00121885-RA	Switchgrass	0.02

Table 4.7. Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in rest of the body sample of *Ctenolepisma longicaudata*. Differentially expressed PCWDE genes in different pair-wise treatment comparisons of rest of the body sample of *C. longicaudata* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{\text{adj}} < 0.05$).

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
CMC vs Protein			
Coding sequence			
Lep_00004873-RA	Glucosidase	CMC	0.04
Lep_00011407-RA	Glucosidase	Protein	0.04
Lep_00015101-RA	Glucosidase	CMC	0.02
Lep_00020108-RA	Glucosidase	CMC	0.006
Lep_00026156-RA	hydrolase family 9	CMC	0.03
Lep_00036877-RA	β-galactosidase-1 2	CMC	0.03
Lep_00038648-RA	hydrolase family 9	CMC	0.04
Lep_00052604-RA	Glucosidase	CMC	0.002
Lep_00059101-RA	β-galactosidase-1 2	CMC	0.006
Lep_00079073-RA	uncharacterized family 31 glucosidase KIAA1161-like	CMC	0.001
Lep_00094746-RA	uncharacterized family 31 glucosidase KIAA1161-like	CMC	< 0.001
Lep_00097528-RA	Glucosidase	CMC	0.02
Lep_00098705-RA	Glucosidase	CMC	0.001
Lep_00102111-RA	uncharacterized family 31 glucosidase KIAA1161-like	CMC	< 0.001
Paper vs Protein			
Coding sequence			
Lep_00004872-RA	Glucosidase	Protein	0.002
Lep_00006775-RA	hydrolase family 9	Paper	< 0.001
Lep_00026154-RA	Endoglucanase	Protein	0.04
Lep_00026156-RA	hydrolase family 9	Protein	< 0.001
Lep_00036184-RA	hydrolase family 9	Paper	< 0.001
Lep_00042852-RA	hydrolase family 9	Paper	< 0.001
Lep_00059101-RA	β-galactosidase-1 2	Protein	0.04
Lep_00060202-RA	hydrolase family 9	Paper	0.005
Lep_00071461-RA	hydrolase family 9	Paper	< 0.001
Lep_00078550-RA	hydrolase family 9	Paper	< 0.001

Table 4.7. Continued.

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Lep_00102111-RA	uncharacterized family 31 glucosidase KIAA1161-like	Protein	0.04
Lep_00113088-RA	Glucosidase	Paper	0.008
Lep_00135466-RA	hydrolase family 9	Paper	< 0.001
Switchgrass vs Protein			
Coding sequence			
Lep_00004872-RA	Glucosidase	Protein	0.04
Lep_00014981-RA	hydrolase family 9	Protein	0.004
Lep_00014982-RA	β-1,4-glucanase 5	Protein	0.003
Lep_00014983-RA	hydrolase family 9	Protein	0.03
Lep_00021305-RA	Glucosidase	Switchgrass	0.04
Lep_00034316-RA	β-1,3-glucan-binding -like	Switchgrass	0.04
Lep_00036877-RA	β-galactosidase-1 2	Protein	0.03
Lep_00042852-RA	hydrolase family 9	Switchgrass	< 0.001
Lep_00050000-RA	hydrolase family 65	Switchgrass	0.003
Lep_00065284-RA	Glucosidase	Switchgrass	0.003
Lep_00077538-RA	uncharacterized family 31 glucosidase KIAA1161-like	Protein	0.04
Lep_00079073-RA	uncharacterized family 31 glucosidase KIAA1161-like	Protein	0.03
Lep_00095117-RA	endo- β-1,4-glucanase	Switchgrass	0.005
Lep_00096399-RA	hydrolase family 9	Protein	<0.001
Lep_00102111-RA	uncharacterized family 31 glucosidase KIAA1161-like	Protein	< 0.001
Lep_00113088-RA	Glucosidase	Protein	< 0.001
Paper vs CMC			
Coding sequence			
Lep_00006775-RA	hydrolase family 9	Paper	< 0.001
Lep_00015101-RA	Glucosidase	Paper	< 0.001
Lep_00022398-RA	hydrolase family 9	Paper	0.001
Lep_00024218-RA	β-1,6-glucanase	Paper	0.04
Lep_00036184-RA	hydrolase family 9	Paper	< 0.001

Table 4.7. Continued.

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Lep_00038648-RA	hydrolase family 9	Paper	< 0.001
Lep_00042851-RA	hydrolase family 9	Paper	0.04
Lep_00044891-RA	hydrolase family 9	Paper	< 0.001
Lep_00051683-RA	hydrolase family 9	Paper	0.004
Lep_00051684-RA	hydrolase family 9	Paper	< 0.001
Lep_00052326-RA	hydrolase family 9	Paper	< 0.001
Lep_00060202-RA	hydrolase family 9	Paper	< 0.001
Lep_00071132-RA	hydrolase family 9	Paper	< 0.001
Lep_00071461-RA	hydrolase family 9	Paper	< 0.001
Lep_00071754-RA	hydrolase family 9	Paper	< 0.001
Lep_00078550-RA	hydrolase family 9	Paper	< 0.001
Lep_00082758-RA	Glucosidase	Paper	0.004
Lep_00115213-RA	hydrolase family 9	Paper	< 0.001
Lep_00135466-RA	hydrolase family 9	Paper	< 0.001
Switchgrass vs CMC			
Coding sequence			
Lep_00004873-RA	Glucosidase	Switchgrass	0.04
Lep_00015101-RA	Glucosidase	Switchgrass	< 0.001
Lep_00016310-RA	hydrolase family 9	CMC	< 0.001
Lep_00016311-RA	endo- β-1,4-glucanase	CMC	< 0.001
Lep_00016312-RA	hydrolase family 9	CMC	< 0.001
Lep_00019550-RA	Glucosidase	Switchgrass	0.03
Lep_00020108-RA	Glucosidase	Switchgrass	0.04
Lep_00021305-RA	Glucosidase	Switchgrass	0.04
Lep_00022398-RA	hydrolase family 9	Switchgrass	0.002
Lep_00024904-RA	Glucosidase	Switchgrass	0.001
Lep_00026029-RA	β-galactosidase	Switchgrass	0.03
Lep_00028613-RA	β-glucuronidase-like	Switchgrass	0.004
Lep_00028614-RA	β-glucuronidase	Switchgrass	< 0.001
Lep_00034316-RA	β-1,3-glucan-binding -like	Switchgrass	0.04
Lep_00036380-RA	Glucosidase	Switchgrass	0.03
Lep_00038648-RA	hydrolase family 9	Switchgrass	< 0.001
Lep_00042851-RA	hydrolase family 9	Switchgrass	0.03
Lep_00042852-RA	hydrolase family 9	Switchgrass	0.002
Lep_00046431-RA	β-glucuronidase	Switchgrass	0.04
Lep_00048639-RA	Glucosidase	Switchgrass	0.03
Lep_00050000-RA	hydrolase family 65	Switchgrass	0.004

Table 4.7. Continued.

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Lep_00051683-RA	hydrolase family 9	Switchgrass	0.001
Lep_00051684-RA	hydrolase family 9	Switchgrass	< 0.001
Lep_00052326-RA	hydrolase family 9	Switchgrass	0.001
Lep_00052604-RA	Glucosidase	Switchgrass	< 0.001
Lep_00053738-RA	β-glucuronidase	Switchgrass	0.004
Lep_00060393-RA	Glucosidase	Switchgrass	0.01
Lep_00065284-RA	Glucosidase	Switchgrass	0.003
Lep_00065345-RA	β-glucuronidase	Switchgrass	0.005
Lep_00071461-RA	hydrolase family 9	Switchgrass	0.01
Lep_00074516-RA	uncharacterized family 31 glucosidase KIAA1161-like	Switchgrass	0.001
Lep_00082758-RA	Glucosidase	Switchgrass	0.04
Lep_00087818-RA	β-1,6-glucanase	Switchgrass	0.001
Lep_00094746-RA	uncharacterized family 31 glucosidase KIAA1161-like	Switchgrass	0.03
Lep_00095117-RA	endo- β-1,4-glucanase	Switchgrass	0.007
Lep_00097528-RA	Glucosidase		0.04
Paper vs Switchgrass			
Coding sequence			
Lep_00006775-RA	hydrolase family 9	Paper	< 0.001
Lep_00008899-RA	β-galactosidase-1 3	Switchgrass	0.03
Lep_00014981-RA	hydrolase family 9	Paper	0.007
Lep_00014982-RA	β-1,4-glucanase 5	Paper	0.002
Lep_00014983-RA	hydrolase family 9	Paper	0.03
Lep_00015051-RA	Mannanase	Switchgrass	0.007
Lep_00016310-RA	hydrolase family 9	Paper	0.003
Lep_00016311-RA	endo- β-1,4-glucanase	Paper	0.001
Lep_00016312-RA	hydrolase family 9	Paper	< 0.001
Lep_00036184-RA	hydrolase family 9	Paper	< 0.001
Lep_00044891-RA	hydrolase family 9	Paper	0.003
Lep_00060202-RA	hydrolase family 9	Paper	< 0.001
Lep_00071132-RA	hydrolase family 9	Paper	0.01
Lep_00071461-RA	hydrolase family 9	Paper	< 0.001
Lep_00074516-RA	uncharacterized family 31 glucosidase KIAA1161-like	Switchgrass	0.04

Table 4.7. Continued.

Diets tested	**Enzyme**	**Up-regulated diet**	**P_{adj} value**
Lep_00078550-RA	hydrolase family 9	Paper	< 0.001
Lep_00096399-RA	hydrolase family 9	Paper	< 0.001
Lep_00115213-RA	hydrolase family 9	Paper	0.002
Lep_00135466-RA	hydrolase family 9	Paper	0.008

Table 4.8. Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in rest of the body sample of *Ctenolepisma longicaudata*. Differentially expressed LPMO genes in different pair-wise treatment comparisons of rest of the body sample of *C. longicaudata* fed on carboxymethyl cellulose (CMC), paper, protein and switchgrass diets ($P_{\mathrm{adj}} < 0.05$).

Diets tested	**Up-regulated diet**	**P_{adj} value**
Switchgrass vs Protein		
Coding sequence		
Lep_00006589-RA	Switchgrass	0.004
Lep_00037556-RA	Switchgrass	< 0.001
Lep_00071838-RA	Switchgrass	< 0.001
Lep_00078707-RA	Switchgrass	0.01
Lep_00099093-RA	Switchgrass	0.006
Lep_00127363-RA	Switchgrass	0.02
Paper vs CMC		
Coding sequence		
Lep_00018140-RA	Paper	< 0.001
Lep_00030922-RA	Paper	< 0.001
Lep_00030924-RA	Paper	< 0.001
Lep_00037556-RA	Paper	0.04
Lep_00072794-RA	Paper	0.02
Lep_00083283-RA	Paper	0.02
Switchgrass vs CMC		
Coding sequence		
Lep_00006587-RA	Switchgrass	0.04
Lep_00006588-RA	Switchgrass	0.04
Lep_00006589-RA	Switchgrass	0.02
Lep_00006590-RA	Switchgrass	0.03
Lep_00018140-RA	Switchgrass	0.005
Lep_00029394-RA	Switchgrass	0.04
Lep_00037556-RA	Switchgrass	0.01
Lep_00039537-RA	Switchgrass	< 0.001
Lep_00043371-RA	Switchgrass	0.03
Lep_00060573-RA	Switchgrass	0.03
Lep_00072794-RA	Switchgrass	< 0.001
Lep_00083283-RA	Switchgrass	< 0.001
Lep_00093845-RA	Switchgrass	0.04
Lep_00099441-RA	Switchgrass	0.03
Lep_00103234-RA	Switchgrass	0.001
Lep_00109934-RA	Switchgrass	0.002
Lep_00119656-RA	Switchgrass	0.01

Table 4.8. Continued.

Diets tested	**Up-regulated diet**	**P_{adj} value**
Lep_00127363-RA	Switchgrass	0.02
Paper vs Switchgrass		
Coding sequence		
Lep_00006587-RA	Switchgrass	0.001
Lep_00006588-RA	Switchgrass	0.03
Lep_00006589-RA	Switchgrass	0.02
Lep_00039538-RA	Switchgrass	0.02
Lep_00127363-RA	Switchgrass	0.04

Table 4.9. Differentially expressed plant cell wall degrading enzyme (PCWDE) gene in *Thermobia domestica* and their top-blast hit. Significantly differentially expressed PCWDE genes ($P_{\mathrm{adj}} < 0.05$) were blasted against the NCBInr database, and the matching protein and the tissue in *T. domestica* are listed.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Th_d_00000348	Cellulase	*Coptotermes acinaciformis*	90%	6e-180	59%	AAK12339.1	Rest of the body
Th_d_00000351	Hypothetical proteinB7P43 G18040 (β-1,4-glucanase)	*Cryptotermes secundus*	55%	4e-61	44%	PNF36367.1	Foregut
Th_d_00000352	Cellulase	*Coptotermes acinaciformis*	74%	2e-45	65%	AAK12339.1	Foregut
Th_d_00000353	β-1,4-endoglucanase	*Cryptotermes secundus*	77%	1e-93	52%	XP_023704929.1	Rest of the body
Th_d_00000354	Endoglucanase E-4-like	*Diuraphis noxia*	24%	2e-19	57%	XP_015378728.1	Rest of the body
Th_d_00001122	β-galactosidase	*Cephus cinctus*	98%	0	54%	XP_015598547.1	Foregut
Th_d_00005705	β-1,4-endoglucanase	*Mastotermes darwiniensis*	80%	0	66%	CAD54729.1	Both
Th_d_00006652	β-galactosidase	*Daphnia pulex*	96%	2e-114	46%	EFX89607.1	Foregut
Th_d_00006653	β-galactosidase	*Austrofundulus limnaeus*	79%	6e-48	42%	XP_013863168.1	Foregut
Th_d_00006654	β-galactosidase	*Monodelphis domestica*	58%	2e-20	70%	XP_007505096.1	Foregut

Table 4.9. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Th_d_00007226	β-1,4-endoglucanase	*Panesthia cribrata*	69%	5e-172	55%	AAF80584.1	Rest of the body
Th_d_00015659	β-1,4-endoglucanase	*Zootermopsis nevadensis*	63%	3e-90	69%	XP_021941324.1	Rest of the body
Th_d_00018098	β-glucuronidase	*Zootermopsis nevadensis*	96%	8e-146	62%	KDR08779.1	Fore gut
Th_d_00018632	β-glucosidase	*Neotermes koshunensis*	98%	1e-178	61%	3AHZ_A	Rest of the body
Th_d_00018984	β-1,4-endoglucanase	*Cryptotermes secundus*	42%	3e-91	66%	PNF24409.1	Rest of the body
Th_d_00028499	β-1,4-endoglucanase	*Mastotermes darwiniensis*	75%	0	65%	CAD54729.1	Both
Th_d_00029874	β-1,4-endoglucanase	*Blattella germanica*	71%	3e-76	48%	PSN31180.1	Rest of the body
Th_d_00031500	β-1,4-endoglucanase	*Cryptotermes secundus*	96%	5e-125	49%	XP_023704929.1	Both
Th_d_00034779	β-1,4-endoglucanase	*Mastotermes darwiniensis*	88%	0	67%	CAD54729.1	Both
Th_d_00035294	β-1,4-endoglucanase	*Zootermopsis nevadensis*	87%	8e-179	59%	XP_021941324.1	Both
Th_d_00035324	β-1,4-endoglucanase	*Coptotermes acinaciformis*	99%	4e-163	58%	AAK12339.1	Rest of the body
Th_d_00036630	Endo- β-1,4-mannanase	*Orchesella cincta*	63%	1e-75	44%	ODM91993.1	Fore gut

Table 4.9. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Th_d_00038486	Glycoside hydrolase family 9	*Aretaon asperrimus*	83%	2e-171	65%	AMH40360.1	Rest of the body
Th_d_00041503	β-1,4-endoglucanase	*Nasutitermes takasagoensis*	92%	2e-141	56%	1KS8_A	Rest of the body
Th_d_00045439	β-1,4-endoglucanase	*Anoplophora glabripennis*	38%	1e-21	53%	XP_018568194.1	Rest of the body
Th_d_00046715	β-1,4-endoglucanase	*Zootermopsis nevadensis*	92%	2e-142	63%	XP_021941324.1	Both
Th_d_00048538	Glucosidase family 31	*Periplaneta americana*	70%	3e-174	53%	AIA09350.1	Foregut
Th_d_00061520	β-1,4-endoglucanase	*Zootermopsis nevadensis*	96%	6e-41	58%	KDR06579.1	Foregut
Th_d_00071154	Cellulase	*Haliotis kamtschatkana*	19%	1e-21	77%	ACS15347.1	Both
Th_d_00093510	Cellulase	*Cryptotermes secundus*	100%	5e-38	62%	PNF24409.1	Foregut
Th_d_00101434	Endo- β-1,4-mannanase	*Cherax quadricarinatus*	65%	7e-26	56%	AIN40245.1	Rest of the body
Th_d_00104189	Endoglucanase E-4-like	*Blattella germanica*	56%	2e-47	74%	PSN33998.1	Rest of the body
Th_d_00107645	β-1,4-endoglucanase	*Eurytemora affinis*	91%	4e-38	60%	XP_023329807.1	Rest of the body
Th_d_00111221	Endoglucanase 4-like	*Orbicella faveolata*	57%	1e-36	60%	XP_020614867.1	Rest of the body

Table 4.9. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Th_d_00118343	Endoglucanase 4-like	*Pomacea canaliculata*	76%	4e-23	69%	PVD35381.1	Rest of the body

Table 4.10. Differentially expressed plant cell wall degrading enzyme (PCWDE) genes in *Ctenolepisma longicaudata* and their top-blast hit. Significantly differentially expressed PCWDEs ($P_{\text{adj}} < 0.05$) were blasted against NCBInr database, and the matching protein and tissue in *C. longicaudata* are listed.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Lep_00044891-RA	Uncharacterized protein LOC110840536 (endo-β-1,4-glucanase)	*Zootermopsis nevadensis*	94	0.0	66	XP_021941324.1	Both
Lep_00098705-RA	β-glucosidase	*Acanthaster planci*	25	5e-07	55	XP_022090546.1	Rest of the body
Lep_00077538-RA	α-glucosidase family 31	*Periplaneta americana*	89	0.0	53	AIA09350.1	Rest of the body
Lep_00087818-RA	Hypothetical protein LOTGIDRAFT 101222 (endo-β-1,6-glucanase)	*Lottia gigantea*	70	3e-122	49	XP_009054280.1	Rest of the body
Lep_00036380-RA	Myrosinase 1-like (β-glucosidase)	*Cryptotermes secundus*	88	2e-104	50	XP_023721112.1	Rest of the body
Lep_00115213-RA	Endoglucanase 7-like	*Zootermopsis nevadensis*	84	1e-80	72	XP_021941322.1	Both
Lep_00004872-RA	β-glucosidase	*Coptotermes formosanus*	76	2e-145	56	AOY34571.1	Rest of the body

Table 4.10. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Lep_00004873-RA	Myrosinase 1-like (β-glucosidase)	*Cryptotermes secundus*	94	8e-50	68	XP_023723774.1	Rest of the body
Lep_00065345-RA	β-glucuronidase-like	*Centruroides sculpturatus*	49	2e-11	51	XP_023221887.1	Rest of the body
Lep_00071461-RA	Uncharacterized protein LOC110840536 (endo-β-1,4-glucanase)	*Zootermopsis nevadensis*	94	1e-105	72	XP_021941324.1	Both
Lep_00019550-RA	Hypothetical protein B5V51_11090 (β-glucosidase)	*Heliothis virescens*	74	8e-50	34	PCG75727.1	Rest of the body
Lep_00021305-RA	Chain A, Crystal Structure of β-Glucosidase From Termite Neotermes Koshunensis in complex with Tris	*Neotermes koshunensis*	83	3e-92	59	3AHZ_A	Rest of the body
Lep_00048639-RA	β-glucosidase	*Salganea esakii*	59	7e-24	55	BAO85048.1	Rest of the body

Table 4.10. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Lep_00053738-RA	Hypothetical protein B7P43_G09739 (β-glucuronidase)	*Cryptotermes secundus*	88	1e-67	62	PNF22051.1	Rest of the body
Lep_00059101-RA	β -galactosidase-1-like protein 2	*Cephus cinctus*	74	9e-65	51	XP_015598574.1	Rest of the body
Lep_00071754-RA	Endoglucanase E-4	*Blattella germanica*	77	6e-79	60	PSN33998.1	Both
Lep_00055855-RA	Hypothetical protein B7P43_G09739 (β-glucuronidase)	*Cryptotermes secundus*	96	9e-68	62	PNF22051.1	Fore gut
Lep_00016310-RA	Endoglucanase 7-like	*Zootermopsis nevadensis*	84	8e-75	71	XP_021941322.1	Rest of the body
Lep_00016311-RA	Cellulase	*Coptotermes acinaciformis*	71	3e-55	58	AAK12339.1	Rest of the body
Lep_00016312-RA	Endoglucanase E-4	*Blattella germanica*	38	8e-44	65	PSN33998.1	Rest of the body
Lep_00074516-RA	Hypothetical protein B7P43_G04891 (Uncharacterized family 31 glucosidase)	*Cryptotermes secundus*	89	1e-147	47	PNF32421.1	Rest of the body
Lep_00065284-RA	β-glucosidase	*Coptotermes formosanus*	83	1e-60	50	AGM32287.1	Rest of the body

Table 4.10. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Lep_00052326-RA	Uncharacterized protein LOC110840536 (endo-β-1,4-glucanase)	*Zootermopsis nevadensis*	83	1e-98	50	XP_021941324.1	Both
Lep_00006757-RA	β -galactosidase-1-like protein 2	*Cephus cinctus*	32	4e-62	52	XP_015598574.1	Fore gut
Lep_00034744-RA	Glycoside hydrolase family 9	*Peruphasma schultei*	86	1e-147	50	AMH40374.1	Fore gut
Lep_00015101-RA	Myrosinase 1-like (β-glucosidase)	*Cryptotermes secundus*	96	4e-147	50	XP_023721112.1	Rest of the body
Lep_00014981-RA	Cellulase	*Antipaluria urichi*	31	4e-58	57	AOV94250.1	Rest of the body
Lep_00014982-RA	Putative endo-β-1,4-glucanase of EG1	*Odontotermes formosanus*	82	5e-77	56	BAD12008.1	Rest of the body
Lep_00014983-RA	Predicted protein (endo-β-1,4-glucanase)	*Nematostella vectensis*	31	7e-23	78	XP_001640311.1	Rest of the body
Lep_00071132-RA	Uncharacterized protein LOC110840536 (endo-β-1,4-glucanase)	*Zootermopsis nevadensis*	95	3e-104	72	XP_021941324.1	Both
Lep_00051683-RA	Glycoside hydrolase family 9	*Timema cristinae*	90	1e-23	64	AMH40395.1	Both

Table 4.10. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Lep_00051684-RA	β-1,4-endoglucanase 1	*Panesthia cribrate*	94	8e-36	63	AAF80584.1	Both
Lep_00079073-RA	Hypothetical protein C0J52_14633 (Uncharacterized family 31 glucosidase)	*Blatella germanica*	60	2e-20	65	PSN31308.1	Rest of the body
Lep_00104311-RA	Hypothetical protein C0J52_21511 (endo-β-1,4-glucanase	*Blatella germanica*	99	2e-55	57	PSN31180.1	Fore gut
Lep_00046431-RA	β-glucuronidase-like isoform X3	*Cryptotermes secundus*	86	3e-123	57	XP_023718877.1	Rest of the body
Lep_00028614-RA	β-glucuronidase-like isoform X3	*Cryptotermes secundus*	91	1e-69	52	XP_023718877.1	Both
Lep_00028613-RA	β-glucuronidase-like	*Halyomorpha halys*	42	6e-46	72	XP_024217164.1	Rest of the body
Lep_00095117-RA	Putative endo- β-1,4-glucanase SmEG1	*Sinocapritermes mushae*	99	1e-47	53	BAD12012.1	Rest of the body
Lep_00060393-RA	β-glucosidase	*Coptotermes formosanus*	81	5e-55	70	AGM32308.1	Rest of the body
Lep_00039163-RA	Hypothetical protein D910_11210 (β-1,3-glucanase)	*Dendroctonus ponderosae*	68	1e-09	44	ERL93924.1	Fore gut

Table 4.10. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Lep_00082758-RA	Lactase-phlorizin hydrolase-like (β-glucosidase)	*Acanthochromis polyacanthus*	90	6e-09	30	XP_022078461.1	Rest of the body
Lep_00038648-RA	Glycoside hydrolase family 9	*Ramulus artemis*	35	5e-22	65	AMH40383.1	Rest of the body
Lep_00034316-RA	Hypothetical protein DAPPUDRAFT_203138 (β-1,3-glucanase)	*Daphnia pulex*	49	3e-36	42	EFX69036.1	Rest of the body
Lep_00060202-RA	Endoglucanase E-4	*Blattella germanica*	73	2e-91	62	PSN33998.1	Both
Lep_00022398-RA	Glycoside hydrolase family 9	*Timema cristinae*	88	5e-176	61	AMH40392.1	Both
Lep_00026029-RA	Hypothetical protein B7P43_G16708 (β-galactosidase)	*Cryptotermes secundus*	66	7e-65	58	PNF39271.1	Rest of the body
Lep_00036184-RA	Uncharacterized protein LOC110840536 (endo- β-1,4-glucanase)	*Zootermopsis nevadensis*	84	0.0	66	XP_021941324.1	Both
Lep_00012530-RA	Endo- β-1,4-mannanase	*Daphnia magna*	90	3e-117	54	KZS10752.1	Fore gut

Table 4.10. Continued.

Sequence_Id	Identified protein	Organism	Query cover	E value	% identity	Accession	Sample
Lep_00050000-RA	Protein-glucosylgalactosylhydroxylysine glucosidase-like isoform X2	*Spodoptera litura*	90	3e-74	52	XP_022821190.1	Rest of the body
Lep_00008899-RA	β-galactosidase-1-like protein 2 isoform X2	*Cryptotermes secundus*	95	0.0	43	XP_023717284.1	Rest of the body
Lep_00094746-RA	α-glucosidase family 31	*Periplaneta Americana*	83	5e-161	56	AIA09350.1	Rest of the body
Lep_00102111-RA	Hypothetical protein L798_13618 (uncharacterized family 31 glucosidase)	*Zootermopsis nevadensis*	67	3e-47	66	KDR11965.1	Rest of the body
Lep_00020108-RA	Hypothetical protein LOTGIDRAFT_207250 (β-glucosidase)	*Lottia gigantea*	20	3e-11	83	XP_009064666.1	Rest of the body
Lep_00034832-RA	β -1,4-endoglucanase 1	*Mastotermes darwiniensis*	70	2e-45	64	AAF63724.1	Fore gut
Lep_00024218-RA	Hypothetical protein LOTGIDRAFT_101222 (β-1,6-glucanase)	*Lottia gigantea*	72	1e-19	43	XP_009054280.1	Rest of the body

Table 4.10. Continued.

Sequence_Id	Identified protein	Organism	Query cover	E value	% identity	Accession	Sample
Lep_00036877-RA	β-galactosidase-1-like protein 2	*Cephus cinctus*	88	3e-60	41	XP_015598574.1	Rest of the body
Lep_00029521-RA	Endo- β-1,4-glucanase	*Panesthia angustipennis*	56	4e-52	51	BAG70031.1	Foregut
Lep_00078550-RA	Glycoside hydrolase family 9	*Peruphasma schultei*	92	1e-80	59	AMH40374.1	Both
Lep_00135466-RA	Glycoside hydrolase family 9	*Peruphasma schultei*	66	6e-24	73	AMH40374.1	Both
Lep_00052604-RA	Chain A, Crystal Structure Of β-glucosidase From Termite Neotermes Koshunensis in Complex with tris	*Neotermes koshunensis*	89	5e-121	59	3AHZ_A	Rest of the body
Lep_00096399-RA	Hypothetical protein B7P43_G09674 (Endo- β-1,4-glucanase)	*Cryptotermes secundus*	38	1e-47	62	XP_021924915.1	Both
Lep_00024904-RA	Myrosinase 1-like isoform X1 (β-glucosidase)	*Nilaparvata lugens*	82	1e-145	56	XP_022187263.1	Rest of the body
Lep_00026156-RA	Endoglucanase E-4	*Blattella germanica*	35	2e-70	62	PSN33998.1	Rest of the body

Table 4.10. Continued.

Sequence_Id	**Identified protein**	**Organism**	**Query cover**	**E value**	**% identity**	**Accession**	**Sample**
Lep_00026154-RA	Putative endo-β-1,4-glucanase HsEG4	*Hodotermopsis sjostedti*	75	8e-43	70	BAD12004.1	Rest of the body
Lep_00097528-RA	Lactase-phlorizin hydrolase-like (β-glucosidase)	*Pomacea canaliculata*	46	8e-47	64	XP_025110416.1	Rest of the body
Lep_00006775-RA	Uncharacterized protein LOC110840536 (Endo- β-1,4-glucanase)	*Zootermopsis nevadensis*	93	0.0	60	XP_021941324.1	Both
Lep_00011407-RA	Lactase-phlorizin hydrolase-like (β-glucosidase)	*Pomacea canaliculata*	47	4e-21	47	XP_025110416.1	Rest of the body
Lep_00113088-RA	Lactase-phlorizin hydrolase-like isoform X2 (β-glucosidase)	*Zootermopsis nevadensis*	97	3e-65	66	XP_021923731.1	Rest of the body
Lep_00042852-RA	β-1,4-endoglucanase	*Mastotermes darwiniensis*	96	2e-107	61	CAD54730.1	Rest of the body
Lep_00042851-RA	Cellulase	*Antipaluria urichi*	38	2e-10	49	AOV94250.1	Rest of the body

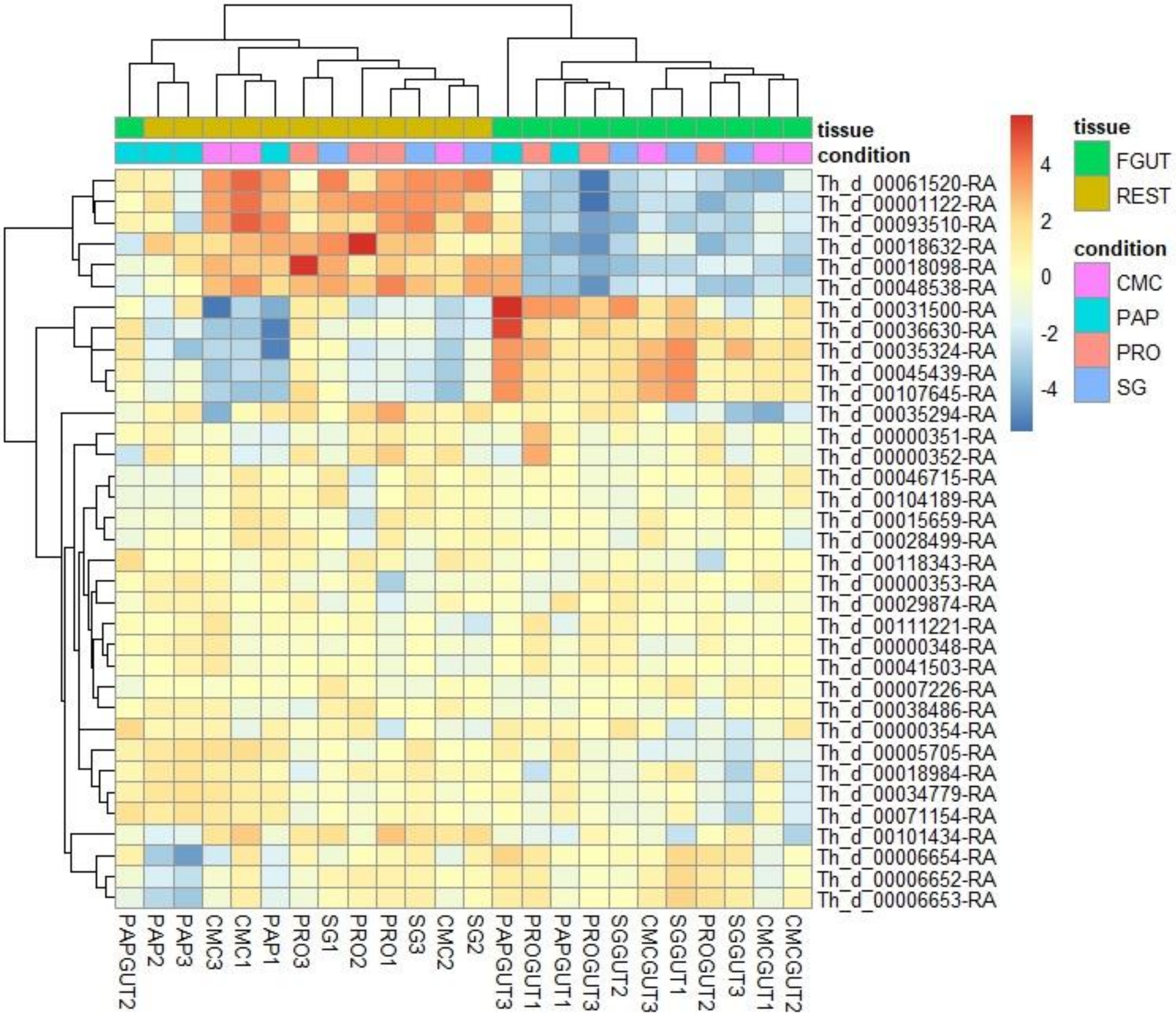

Figure 4.1. Heatmap of expression of of all differentially expressed PCWDE genes in *Thermobia domestica*: heatmap showing the significantly differentially expressed PCWDE genes in *T. domestica* foregut (FGUT) and rest of the body samples (REST) fed on carboxymethyl cellulose (CMC), paper (PAP), protein (PRO) and switchgrass (SG) diets ($P_{\text{adj}} < 0.05$).

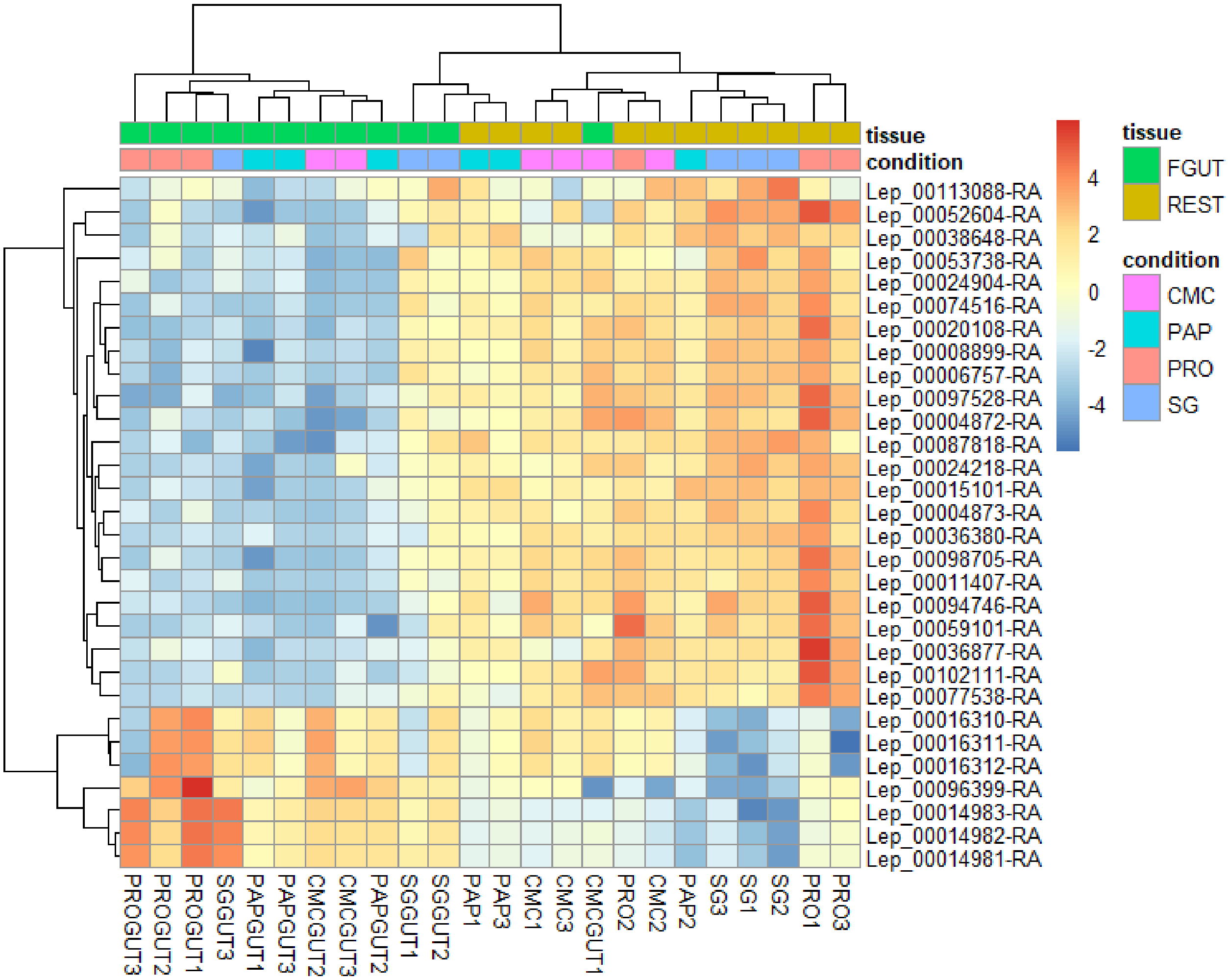

Figure 4.2. Heatmap of expression of all differentially expressed PCWDE genes in *Ctenolepisma longicaudata*: heatmap showing the significantly differentially expressed PCWDEs in *C. longicaudata* foregut (FGUT) and rest of the body samples (REST) fed on carboxymethyl cellulose (CMC), paper (PAP), protein (PRO) and switchgrass (SG) diets ($P_{adj} < 0.05$).

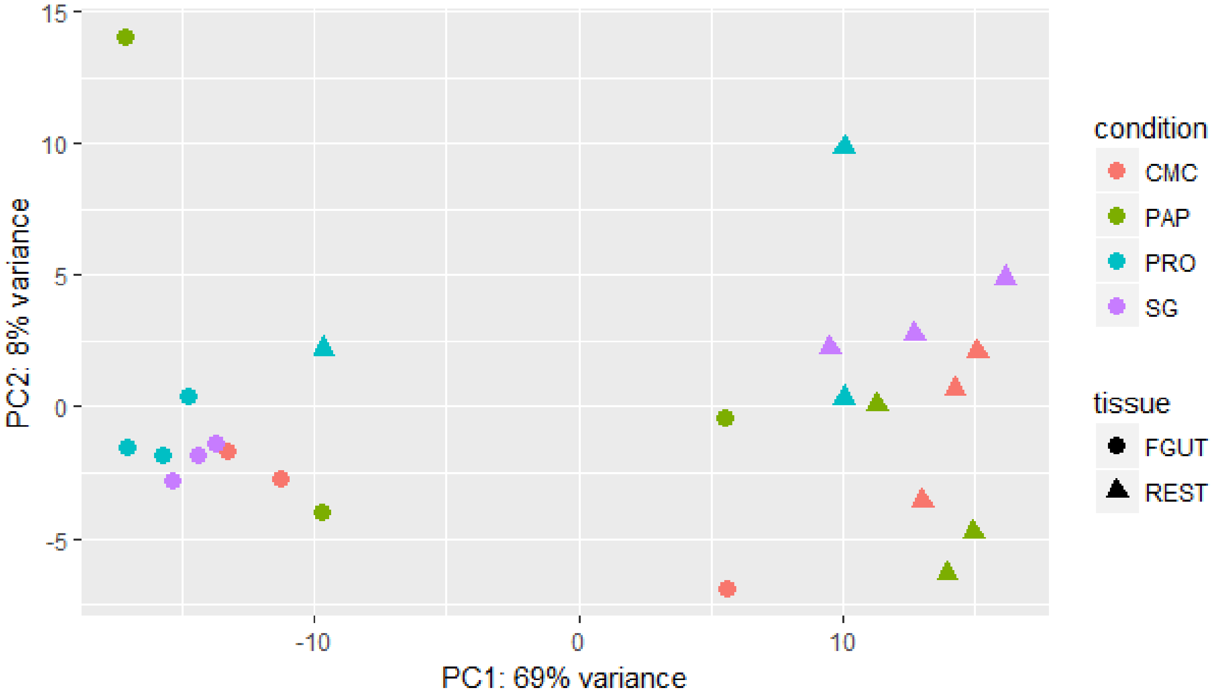

Figure 4.3. Overall effect of tissue (PC1) and condition (diet, PC2) on gene expression of lytic polysaccharide monoxygenases (LPMOs) in *Thermobia domestica*: PCA plot showing the strong effect of tissue compared to diet on the gene expression of significantly differentially expressed LPMOs in foregut (FGUT) and rest of the body (REST) tissues of *T. domestica* fed on carboxymethyl cellulose (CMC), paper (PAP), protein (PRO) and switchgrass (SG) diets ($P_{adj} < 0.05$).

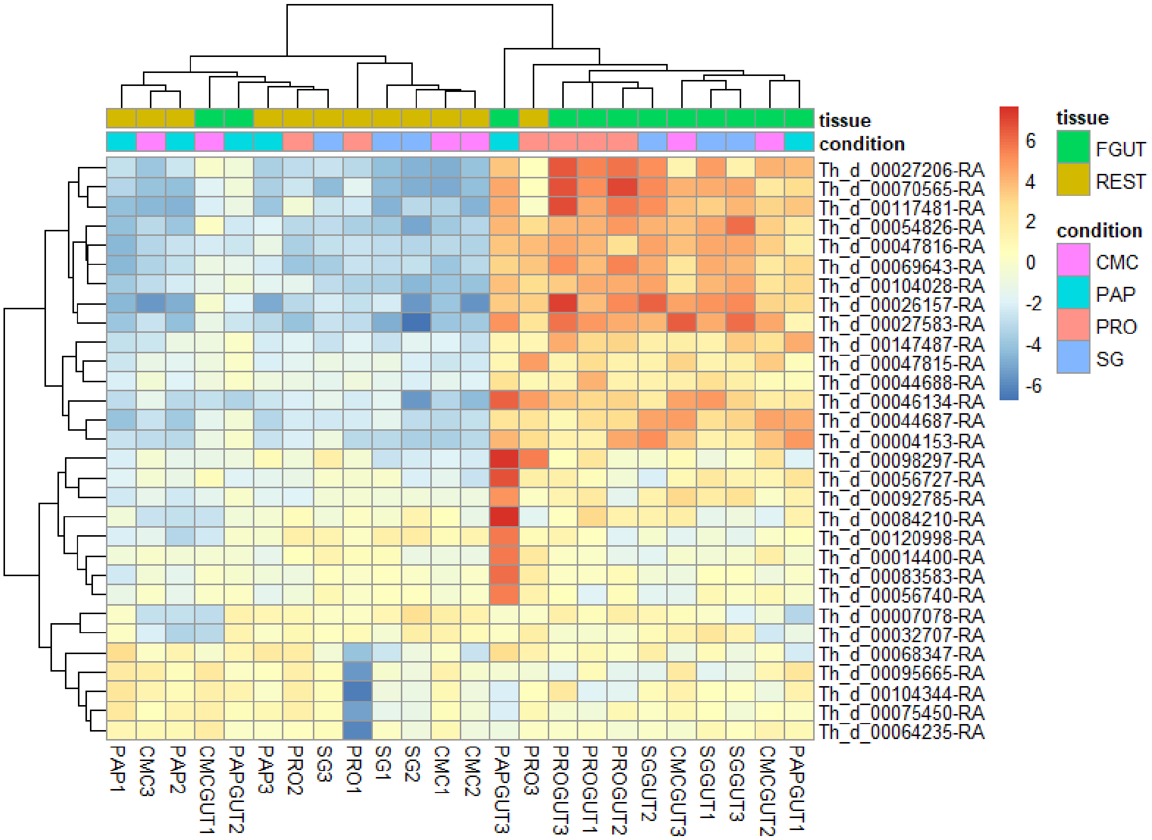

Figure 4.4. Differential expression of all lytic polysaccharide momoxygenases (LPMOs) in *Thermobia domestica*: heatmap showing the significantly differentially expressed LPMOs in *T. domestica* foregut (FGUT) and rest of the body samples (REST) fed on carboxymethyl cellulose (CMC), paper (PAP), protein (PRO) and switchgrass (SG) diets ($P_{\text{adj}} < 0.05$).

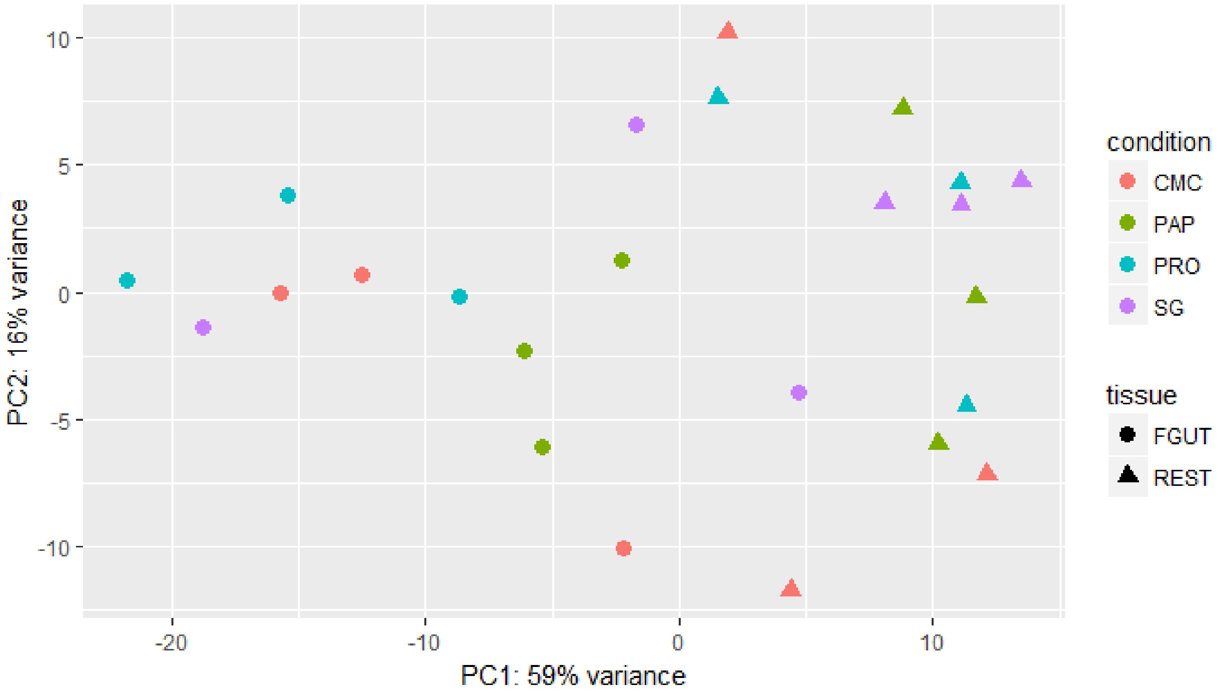

Figure 4.5. Overall effect of tissue (PC1) and condition (diet, PC2) on gene expression of lytic polysaccharide monoxygenases (LPMOs) in *Ctenolepisma longicaudata*: PCA plot showing the strong effect of tissue compared to diet on the gene expression of significantly differentially expressed LPMOs in foregut (FGUT) and rest of the body (REST) tissues of *C. longicaudata* fed on carboxymethyl cellulose (CMC), paper (PAP), protein (PRO) and switchgrass (SG) diets ($P_{adj} < 0.05$).

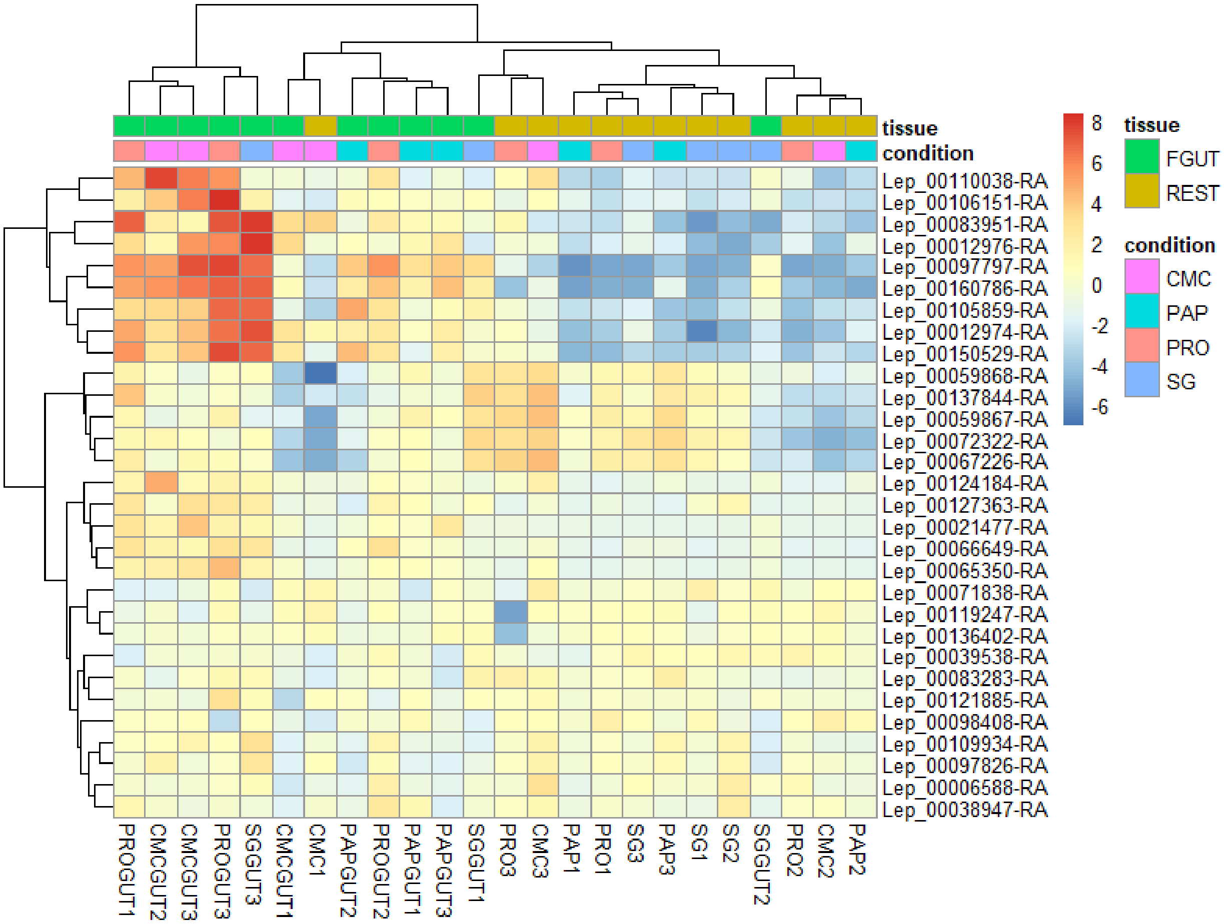

Figure 4.6. Differential expression of all lytic polysaccharide momoxygenases (LPMOs) in *Ctenolepisma longicaudata*: heatmap showing the significantly differentially expressed LPMOs in *C. longicaudata* foregut (FGUT) and rest of the body samples (REST) fed on carboxymethyl cellulose (CMC), paper (PAP), protein (PRO) and switchgrass (SG) diets ($P_{adj} < 0.05$).

Chapter 5

General conclusions

Conclusions

Our research focused on *Thermobia domestica* and *Ctenolepisma longicaudata* and unraveled many interesting observations about these two species. Initial morphohistological characterization supported no relevant morphological and histological adaptations to house symbionts in the digestive system of firebrat (*T. domestica*) and the gray silverfish (*C. longicaudata*), which may suggest the endogenous production of cellulases in these insects. Previous studies supported the endogenous digestion of cellulose in the firebrat (Treves and Martin, 1994; Zinkler and Götze, 1987). Additionally, no morphohistological differences were found in the digestive tube of both the tested species.

Significant differences were observed in cellulase activities between species. Quantitative and qualitative cellulase assays identified the foregut as the region with the highest cellulolytic activity compared to other digestive regions in both the species. This observation is also supported by previous reports documenting higher endoglucanase and β-glucosidase activities in the foregut compared to other gut tissues in *T. domestica* (Zinkler and Gotze, 1987). Additionally, *T. domestica* was found displaying higher endoglucanase, xylanase activities compared to *C. longicaudata* and pectinase activity was only observed in *T. domestica*. However, pectinase genes were not detected in the corresponding *T. domestica* genome. A possible explanation for this discrepancy could be that pectinase activity in *T. domestica* may be provided by microorganisms living in its digestive system, which were not included in the genome sequencing. Metatranscriptomics on the microbiome of *T. domestica* will help in determing the contribution of microbial enzymes to pectinase digestion in *T. domestica*. Alternatively, it is possible that pectinase genes may be included among the sequences not returing relevant BLAST matches in our Blast2Go analysis. Searches using conserved pectinase catalytic domains may allow identification of pectinase-like sequences in Zygentoma genomes. On the other hand, the lack of pectinase activity in *C. longicauadata* may suggest that this polysaccharide is not relevant to its nutrition.

T. domestica also displayed significantly higher xylanase activity than *C. longicaudata*. However, genes encoding for xylanases were not detected in *T. domestica* genome and only three xylanase encoding genes were found in the genome of *C. longicauadata*. Many insects including *T. domestica* were previously found to display xylanase activity (Sabbadin et al., 2018; Shi et al., 2011; Terra and Ferreira, 1994), yet xylanases are rarely described as endogenously produced in insects (Calderón-Cortés et al., 2012) and in most instances expected to come from symbiotic microbiota (Ali et al., 2017; Brennan et al., 2004) or through horizontal gene transfer from symbionts (Pauchet and Heckel, 2013). Sequencing the metatranscriptome of *T. domestica* and *C. longicaudata* gut microbiota would help in understanding the source of xylanase activity in these insects. It is also possible that hemicellulose could be digested in insects by other enzymes, such as mannanases, α-glucuronidases, endoglucanases and β-1,3-glucanases(Calderón-Cortés et al., 2012), which were present in the genomes of both species.

Zygentoma displayed relatively high cellulolytic activity compared to other insects (Pothula *et al*, submitted), which may be explained by the detected genes encoding for diverse glycosyl hydrolases in their genomes. Consequently, annotation of coding sequences from the genomes of *T. domestica* and *C. longicaudata* reported numerous genes encoding for endoglucanases, glucosidases, β-1,3-glucanases, maltases, amylases, mannosidases and glucuronidases. Compared to *C. longicaudata*, *T. domestica* had more sequences encoding for endoglucanases, which may explain the higher endoglucanase activity reported in *T. domestica* than *C. longicaudata*. Additionally, both species yielded nearly an equal number of β-glucosidase genes, which was reflected in similar enzyme activity levels. Apart from glycosyl hydrolases, lytic polysaccharide monoxygenases (LPMOs), which were shown to enhance the activity of glycoside hydrolases synergistically (Sabbadin et al., 2018) were abundantly reported in the genomes of both species. The presence of high number of genes encoding LPMOs may also be responsible for higher enzyme activities in Zygentoma compaed to other tested insects.

Differential gene expression analysis was conducted to see the influence of diet on the gene expression of glycoside hydrolases and LPMOs in both foregut and rest of the body samples of both *T. domestica* and *C. longicaudata*. PCWDE gene expression was primarily driven by type of tissue rather than diet, yet within each tissue higher number of PCWDEs were significantly up-regulated in paper-fed insects, which is more cellulosic compared to all other tested diets. In addition, more LPMOs were up-regulated in the foregut tissue of paper-fed *T. domestica* than other diets, yet LPMO up-regulation was not as prominent in *C.*

longicaudata. The paper diet used majorly consists of recalcitrant cellulose and traces of hemicellulose and lignin, while switchgrass diet was composed of a variety of components such as recalcitrant cellulose, hemicellulose, lignin along with easily digestible starch, vitamins and minerals (Ververis et al., 2004). The availability of only recalcitrant cellulose in paper may be responsible for up-regulation of majority of PCWDEs to digest more cellulose in meeting energy requirements by insects.

The annotation of differentially expressed PCWDE and LPMO encoding genes revealed highest identity to insect homologs, which suggests the potential conservation of PCWDEs through evolution.

Overall, our work reports that members of Zygentoma display cellulase, xylanase and pectinase activities. Digestive fluids of *T. domestica* appeared significantly more active than in *C. longicaudata*, although in both insects the highest levels of digestion were detected in the foregut. Additionally, both species were found as containing repertoires of numerous and diverse PCWDE and LPMO genes. However, cellulase gene expression and LPMOs was strongly driven by tissue in *T. domestica* and *C. longicaudata*. We contribute to increasing the amount of information available on functional PCWDE genes and LPMOs from a primitive hexapod group, which will help in characterizing more efficient cellulases. In contrast to existing commercial cellulases, insect cellulases were reported to retain their highest activity at alkaline pH (Willis et al., 2011). This unique trait makes it possible to combine these enzymes with ionic liquids used for lignin digestion in biorefineries (Zhao et al., 2009, 2008). In the future, cloning and expression of these PCWDE and LPMO genes in heterologous systems such as yeast and testing their activity under different temperature and pH regimes may help in identifying efficient cellulases with potential for industrial biofuel applications.

Vita

Ratnasri Mallipeddi is originally from India and graduated with a B.Sc. degree in Agricultural sciences form the Acharya N. G. Ranga Agricultural University, Hyderabad, India. She later obtained her M.Sc. degree in Entomology from the University of Agricultural Sciences, Dharwad, India. Upon graduation, she worked as a Research Associate at the Maharastra Hybrid Coorpation (MAHYCO) in India. Later, she was accepted into the Entomology, Plant Pathology and Nematology PhD program at the University of Tennessee at Knoxville under the guidance of Dr. Juan Luis Juarat-Fuentes.