

University of Tennessee, Knoxville TRACE: Tennessee Research and Creative Exchange

Doctoral Dissertations

Graduate School

12-2018

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

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

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

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

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.

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

Copyright © 2018 by Ratnasri Mallipeddi and Juan Luis Jurat-Fuentes. "Biochemical and transcriptomic characterization of glycoside hydrolases in *Thermobia domestica* and *Ctenolepisma longicaudata*."

All rights reserved.

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

v

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

Chapter 1: General introduction1
Introduction2
List of references6
Chapter 2: The digestive system in Zygentoma as a model for high cellulase activity9
Abstract10
Introduction11
Materials and methods13
Insects13
Gut morphology and histology14
Preparation of samples for biochemical tests14
Zymography15
Quantification of cellulase, xylanase and pectinase activities16
Results18
High cellulase activity and gut morphology in Zygentoma18
Qualitative and quantitative location of cellulolytic activity in the digestive
system of <i>T. domestica</i> and <i>C. longicaudata</i> 19
Discussion20
List of references25
Appendix 231

Chapter 3: Annotation of plant cell wall degrading enzymes (PCWDEs) among coding

Insect rearing
Diet treatments
Dissections, RNA extraction and sequencing
Differential gene expression analysis
Annotation
Results
de novo transcriptome assemblies of T. domestica and C. longicaudata329
Differential cellulase gene expression analysis in <i>T. domestica</i>
Annotation of differentially expressed PCWDE genes in <i>T. domestica</i> 331
Differential cellulase gene expression analysis in <i>C. longicaudata</i>
Annotation of differentially expressed PCWDE genes in <i>C. longicaudata</i> 334
Discussion
List of references
Appendix 4346
Chapter 5: General conclusions
Conclusion
Vita

List of tables

Table 3.1 Endoglucanases in <i>Thermobia domestica</i> : Coding sequences in <i>T. domestica</i>
genome encoding for endoglucanases and their blast description56
Table 3.2 Glucosidases in Thermobia domestica: Coding sequences in T. domestica genome
encoding for glucosidases and their blast description77
Table 3.3 β -1,3-glucanases in <i>Thermobia domestica</i> : Coding sequences in <i>T. domestica</i>
genome encoding for β -1,3-glucanases and their blast description
Table 3.4 Mannanases in Thermobia domestica: Coding sequences in T. domestica genome
encoding for mannanases and their blast description96
Table 3.5 Glucuronidases in Thermobia domestica: Coding sequences in T. domestica
genome encoding for glucuronidases and their blast description123
Table 3.6 Maltases in Thermobia domestica: Coding sequences in T. domestica genome
encoding for maltases and their blast description133
Table 3.7 Amylases in Thermobia domestica: Coding sequences in T. domestica genome
encoding for amylases and their blast description142
Table 3.8 Galactosidases in <i>Thermobia domestica</i> : Coding sequences in <i>T. domestica</i> genome
encoding for galactosidases and their blast description147
Table 3.9 Myrosinases in Thermobia domestica: Coding sequences in T. domestica genome
encoding for mirosinases and their blast description153
Table 3.10 Lytic polysaccharide monooxygenases (LPMOs) in Thermobia domestica: Coding
sequences in <i>T. domestica</i> genome encoding for LPMOs and their blast
description164

Table 3.11 Endoglucanases in *Ctenolepisma longicaudata*: Coding sequences in *C.longicaudata* genome encoding for endoglucanases and their blast description.....174

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...213

- Table 3.15 Glucuronidases in *Ctenolepisma longicaudata*: Coding sequences in *C.longicaudata* genome encoding for glucuronidases and their blast description......250

Table 3.18 Galactosidases in *Ctenolepisma longicaudata*: Coding sequences in *C.longicaudata* genome encoding for galactosidases and their blast description......269

Table 3.21 Lytic polysaccharide monooxygenases (LPMOs) in <i>Ctenolepisma longicaudata</i> :
Coding sequences in <i>C. longicaudata</i> genome encoding for LPMOs and their blast
description296
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
dentical protein groups database of NCBI
Table 4.1 Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in
foregut tissue of <i>Thermobia domestica</i> 347
Table 4.2 Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in
foregut tissue of <i>Thermobia domestica</i> 349
Table 4.3 Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in
rest of the body sample of <i>Thermobia domestica</i> 351
Table 4.4 Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in
rest of the body sample of <i>Thermobia domestica</i> 354
Table 4.5 Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in
foregut tissue of <i>Ctenolepisma longicaudata</i> 355
Table 4.6 Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in
foregut tissue of <i>Ctenolepisma longicaudata</i> 357
Table 4.7 Effect of diet on plant cell wall degrading enzyme (PCWDE) gene expression in
rest of the body sample of <i>Ctenolepisma longicaudata</i>

Table 4.8 Effect of diet on lytic polysaccharide monoxygenase (LPMO) gene expression in	1
rest of the body sample of <i>Ctenolepisma longicaudata</i> 3	63
Table 4.9 Differentially expressed plant cell wall degrading enzyme (PCWDE) genes in	
<i>Thermobia domestica</i> and their top-blast hit3	65
Table 4.10 Differentially expressed plant cell wall degrading enzyme (PCWDE) genes in	
Ctenolepisma longicaudata and their top-blast hit	69

List of figures

Figure 2.1 Structural comparison of the digestive tract in firebrat (<i>Thermobia domestica</i>) and
silverfish (Ctenolepisma longicaudata), and relative CMCase activity in Zygentoma
compared to other insect orders32
Figure 2.2 Histology of the digestive system regions in <i>Ctenolepisma longicaudata</i> 33
Figure 2.3 Detection of cellulolytic activity in digestive fluids of <i>Thermobia domestica</i> and
Ctenolepisma longicaudata34
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
domestica35
Figure 3.1 Annotation of <i>Thermobia domestica</i> coding sequences
Figure 3.2 Top-hit species distribution for <i>Thermobia domestica</i> 312
Figure 3.3 Enzyme Code distribution in <i>Thermobia domestica</i>
Figure 3.4 Enzyme Code distribution of hydrolases in <i>Thermobia domestica</i> 314
Figure 3.5 Glycoside hydrolases (GH) and lytic polysaccharide monoxygenases (LPMOs) in
Thermobia domestica315
Figure 3.6 Annotation of <i>Ctenolepisma longicaudata</i> coding sequences
Figure 3.7 Top-hit species distribution for <i>Ctenolepisma longicaudata</i>
Figure 3.8 Enzyme Code distribution in <i>Ctenolepisma longicaudata</i> 318
Figure 3.9 Enzyme Code distribution of hydrolases in <i>Ctenolepisma longicaudata</i> 319
Figure 3.10 Glycoside hydrolases (GH) and lytic polysaccharide monoxygenases (LPMOs) in
Ctenolepisma longicaudata320

Figure 4.1 Heatmap of expression of all differentially expressed PCWDE genes in <i>Thermobia</i>
domestica378
Figure 4.2 Heatmap of expression of all differentially expressed PCWDE genes in
Ctenolepisma longicaudata379
Figure 4.3 Overall effect of tissue (PC1) and condition (diet, PC2) on gene expression of lytic
polysaccharide monoxygenases (LPMOs) in <i>Thermobia domestica</i>
Figure 4.4 Differential expression of all lytic polysaccharide momoxygenases (LPMOs) in
Thermobia domestica
Figure 4.5 Overall effect of tissue (PC1) and condition (diet, PC2) on gene expression of lytic
polysaccharide monoxygenases (LPMOs) in <i>Ctenolepisma longicaudata</i>
Figure 4.6 Differential expression of all lytic polysaccharide momoxygenases (LPMOs) in
Ctenolepisma longicaudata

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 CO₂. 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°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°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°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 11^{th} 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°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°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° C 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 -Dcellopentaoside (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°C for 30 min. To a 50 μl aliquot of reaction mixture, 50 μl of 2 M Na₂CO₃ 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°C.

Endoglucanase activity was measured in samples (5 µg of protein in 25 µl) mixed with 50 µl of substrate and incubated at 40°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 μ g of protein in 20 μ l) mixed with 15 μ l of 1% pectin and 115 μ l of 50 mM sodium acetate buffer (pH 5.0). The mixture was incubated at 50°C for 1 h, and then 50 μ l 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 μ mol of glucose per min at pH 5.0 and 50°C.

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 20 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 symbiontindependent 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 enzymecompatible 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 solventbased 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-cellopisma 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 μ mol of 4-nitrophenol from the respective substrate in all the graphs except in D), where it is 1 μ mol 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 -1,3glucanases, 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 lidless 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^oC.

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

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 -glactosidases of GH 1 (19 sequences), β -1,3glucanases 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 onoinsect 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 43 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-betaxylanases (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 βglucosidase 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 ^a 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 enzymecompatible 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 solventbased 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
		Segu							Align	
	Seque	ence				E-	Simi	Bit-	ment	Posi
Sequenc	nce	leng				Valu	larit	Scor	lengt	tive
e name	desc.	th	Hit desc.	Order	Hit ACC	e	v	e	h	s
	endogl	•				-	9			
	ucanas		gil9853883481refIXP_015378728.1				83.0			
Th d 000	е Е-4-		IPREDICTED: endoglucanase E-4-	Heminter	XP 015	4 22	188	90.1		
57401-RA	like	468	like [Diuranhis noxia]	a	378728	E-21	7	225	53	44
	endogl	100		u	0/0/20				00	
	ucanas		gil985388348lreflXP_0153787281							
Th d 001	е Е-4-		IPREDICTED: endoglucanase E-4-	Heminter	XP 015	143		993		
13252-RA	like	288	like [Diuranhis novia]	a	378728	E-25	675	673	80	54
10202 101	Fndogl	200	gil695189911 gb AIT11911 1 endo	ŭ	0/0/20	1 20	68.5	070	00	
Th d 000	ucanas		-heta-1 4-glucanase [Parasesarma		AIT119	1 32	714	80.4		
36852-RA	e E-4	612	erythrodactylal	Crustacea	11	E-15	3	925	70	48
50052 101		012	gil805808256lreflXP_012146513.1	Grustacea	11		5	725	70	10
			IPREDICTED: endoglucanase 15-							
			like [Megachile							
	endogl		rotundatalgil8058082591refIXP_01		XP 012					
	ucanas		2146514 1 PREDICTED		146512		90.5			
Th d 000	o 15-		endoglucanase 15-like [Megachile	Hymenon	XP 012	1 66	660	974		1
85723-PA	liko	390	rotundatal	tera	146514	F-22	4	413	53	48
05725-NA	IIKC	570	Totunuata	luia	110314	6-22	т	TIJ	55	-10

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

		Sequ							Align	
	Seque	ence				Е-	Simi	Bit-	ment	Posi
Sequenc	nce	leng				Valu	larit	Scor	lengt	tive
e name	desc.	th	Hit desc.	Order	Hit ACC	e	У	е	h	S
			gi 1000740642 ref XP_015592927.							
			1 PREDICTED: endoglucanase E-4-							
			like [Cephus							
			cinctus]gi 1000740644 ref XP_015							
			592928.1 PREDICTED:		XP_015					
			endoglucanase E-4-like [Cephus		592927,					
	endogl		cinctus]gi 1000740646 ref XP_015		XP_015					
	ucanas		592929.1 PREDICTED:		592928,					
Th_d_000	e E-4-		endoglucanase E-4-like [Cephus	Hymenop	XP_015	8.06	70.2	196.		
90601-RA	like	573	cinctus]	tera	592929	E-59	381	823	168	118
	endogl									
	ucanas		gi 985388348 ref XP_015378728.1				78.6			
Th_d_001	e 15-		PREDICTED: endoglucanase E-4-	Hemipter	XP_015	1.02	885	93.9		
18343-RA	like	243	like [Diuraphis noxia]	а	378728	E-23	2	745	61	48
	endogl									
	ucanas		gi 646689408 gb KDR06579.1 End				78.8			
Th_d_000	e E-4-		oglucanase 1 [Zootermopsis		KDR065	3.23	135	148.		
04757-RA	like	522	nevadensis]	Blattodea	79	E-44	6	288	118	93
	Endogl		gi 197691951 dbj BAG70027.1 end				73.0			
Th_d_000	ucanas		o-beta-1,4-glucanase, partial		BAG700	5.19	158	83.5		
04759-RA	e E-4	381	[Salganea esakii]	Blattodea	27	E-18	7	741	63	46

Sequenc	Seque nce	Sequenc				E- Valu	Simi larit	Bit- Scor	Align ment lengt	Posi tive
e name	desc.	e length	Hit desc.	Order	Hit ACC	е	У	е	h	S
Th d 000	endogl ucanas		gi 1080053219 ref XP_018568 194.1 uncharacterized protein LOC108908596 isoform X1 [Anoplophora glabripennis]gi 1080053221 r ef XP_018568195.1 uncharact erized protein LOC108908596 isoform X1 [Anoplophora glabripennis]gi 1325348753 r ef XP_023310346.1 uncharact erized protein LOC108908596 isoform X2 [Anoplophora	Coloonton	XP_018 568194, XP_018 568195,	1 21	86.7	02 5		
04756-RA	like	735	glabripennisl	a	310346	E-19	5	893	53	46
Th_d_000 29458-RA	Endogl ucanas e F	569	gi 646711640 gb KDR16731.1 Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR167 31	2.2E -47	76.2 963	166. 777	135	103
Th_d_000 29459-RA	Endogl ucanas e E-4 precur sor, putativ e	432	gi 646711640 gb KDR16731.1 Endoglucanase F [Zootermopsis nevadensis]	Blattodea	KDR167 31	1.6E -31	90.4 761 9	122. 479	63	57

Sequenc e name	Seque nce desc.	Sequenc e length	Hit desc.	Order	Hit ACC	E- Valu e	Simi larit y	Bit- Scor e	Align ment lengt h	Posi tive s
	Endogl									
	ucanas									
	e E-4									
	precur									
	sor,		gi 646711640 gb KDR16731.1				91.8			
Th_d_000	putativ		Endoglucanase F		KDR167	8.97	032	124.		
89623-RA	e	288	[Zootermopsis nevadensis]	Blattodea	31	E-33	8	02	61	56
	Endogl		gi 1061478177 gb 0DM95820				76.2			
Th_d_000	ucanas		.1 Endoglucanase E-4	Collembol	ODM95	1.71	711	81.2		
89451-RA	e E-4	258	[Orchesella cincta]	а	820	E-17	9	629	59	45
	endogl		gi 1330895262 gb PNF24409.							
	ucanas		1 hypothetical protein				94.8			
Th_d_000	e E-4-		B7P43_G09674, partial		PNF244	6.72	717	72.0		
78629-RA	like	1143	[Cryptotermes secundus]	Blattodea	09	E-12	9	182	39	37

Sequenc e name	Seque nce desc.	Seq uen ce leng th	Hit desc.	Order	Hit ACC	E- Valu e	Simi larit v	Bit- Scor e	Align ment lengt h	Posi tive s
	AChain						5			
	A, The									
	Structu									
	re Of		gi 28373491 pdb 1KS8 AChain A,							
	Endogl		The Structure Of Endoglucanase							
	ucanas		From Termite, Nasutitermes							
	e From		Takasagoensis, At Ph							
	Termit		2.5.gi 28373492 pdb 1KSC AChain							
	е,		A, The Structure Of Endoglucanase							
	Nasutit		From Termite, Nasutitermes							
	ermes		Takasagoensis, At Ph							
	Takasa		5.6.gi 28373493 pdb 1KSD AChain							
	goensis		A, The Structure Of Endoglucanase		1KS8_A,					
Th_d_000	, At Ph	116	From Termite, Nasutitermes		1KSC_A,	5.6E	73.5	418.		
41503-RA	2.5.	0	Takasagoensis, At Ph 6.5.	Blattodea	1KSD_A	-143	376	698	359	264

	Seque	Seq uen ce				E-	Simi	Bit-	Align ment	Posi
Sequenc	nce	leng				Valu	larit	Scor	lengt	tive
e name	desc.	th	Hit desc.	Order	Hit ACC	e	У	е	h	S
			gi 1339056265 ref XP_023716596.							
			1 uncharacterized protein							
			LOC111869358 [Cryptotermes							
			secundus]gi 1339056267 ref XP_02							
			3716597.1 uncharacterized protein		XP_023					
			LOC111869358 [Cryptotermes		716596,					
	endogl		secundus]gi 1330895261 gb PNF24		XP_023					
	ucanas		408.1 hypothetical protein		716597,		90.5			
Th_d_000	e E-4-	100	B7P43_G09674 [Cryptotermes		PNF244	1.9E	579	362.		
68682-RA	like	7	secundus]	Blattodea	08	-121	4	844	233	211
			gi 1339087696 ref XP_023704929.							
			1 uncharacterized protein							
			LOC111863126 [Cryptotermes		XP_023					
	Endogl		secundus]gi 1330920323 gb PNF36		704929,		67.0			
Th_d_000	ucanas	183	365.1 Endoglucanase A		PNF363	3.77	967	304.		
00353-RA	e A	6	[Cryptotermes secundus]	Blattodea	65	E-95	7	679	310	208
			gi 769842984 ref XP_011632703.1							
			PREDICTED: endoglucanase 15-like							
			[Pogonomyrmex							
	endogl		barbatus]gi 769842986 ref XP_011		XP_011					
	ucanas		632704.1 PREDICTED:		632703,		67.9			
Th_d_000	e 15-		endoglucanase 15-like	Hymenop	XP_011	4.89	487	94.3		
00350-RA	like	270	[Pogonomyrmex barbatus]	tera	632704	E-22	2	597	78	53

		Seq							A 11	
	Soquo	uen				F-	Simi	Dit.	Align	Doci
Sequenc	nce	leng				Valu	larit	Scor	lengt	tive
e name	desc.	th	Hit desc.	Order	Hit ACC	e	v	e	h	s
	endogl					_	5	_		
	ucanas		gi 985388348 ref XP_015378728.1				70.3			
Th_d_000	e E-4-		PREDICTED: endoglucanase E-4-like	Hemipter	XP_015	5.16	703	94.7		
00354-RA	like	984	[Diuraphis noxia]	a	378728	E-21	7	449	81	57
	Endogl									
Th_d_001	ucanas		gi 1078570861 gb A0V94255.1 cell	Embiopte	A0V942	6.75	84.3	56.9		
21658-RA	e A	267	ulase [Antipaluria urichi]	ra	55	E-09	75	954	32	27
	Endogl									
	ucanas									
	e E-4									
	precur		gi 1325291025 ref XP_023329807.							
	sor,		1 uncharacterized protein				72.8			
Th_d_001	putativ		LOC111702374 isoform X2	_	XP_023	1.34	070	135.		
07645-RA	е	369	[Eurytemora affinis]	Crustacea	329807	E-39	2	191	114	83
	endogl									
T I 1000	ucanas		gi 952540008 gb KRT85487.1 hypo			6.40	71.5	101		
Th_d_000	e E-4-	100	thetical protein AMK59_2656	Coleopter	KRT854	6.43	447	104.	100	
96060-RA	like	402	[Oryctes borbonicus]	а	87	E-26	2	76	123	88
			gi 44885844 dbj BAD12011.1 putat				72.2			
	Endogl		Ive endo-beta-1,4-glucanase NtEG2,		DAD120	2.15	/2.3	120		
In_a_000	ucanas	040	partial [Nasutitermes	Dlatta de -	BAD120	2.1E	5//	139.	122	00
2314/-KA	е Е-4	049	takasagoensis	ыаттоцеа	11	-30	L	043	123	89

		Seq							Align	
	Seque	ce				E-	Simi	Bit-	ment	Posi
Sequenc	nce	leng				Valu	larit	Scor	lengt	tive
e name	desc.	th	Hit desc.	Order	Hit ACC	e	У	е	h	S
			gi 805808256 ref XP_012146513.1							
			PREDICTED: endoglucanase 15-like							
			[Megachile							
	endogl		rotundata]gi 805808259 ref XP_01		XP_012					
	ucanas		2146514.1 PREDICTED:		146513,		88.4			
Th_d_001	e 15-		endoglucanase 15-like [Megachile	Hymenop	XP_012	4.29	615	93.5		
13210-RA	like	225	rotundata]	tera	146514	E-22	4	893	52	46
			gi 1080053219 ref XP_018568194.							
			1 uncharacterized protein							
			LOC108908596 isoform X1							
			[Anoplophora							
			glabripennis]gi 1080053221 ref XP							
			_018568195.1 uncharacterized							
			protein LOC108908596 isoform X1		XP_018					
			[Anoplophora		568194,					
			glabripennis]gi 1325348753 ref XP		XP_018					
	Endogl		_023310346.1 uncharacterized		568195,		71.1			
Th_d_000	ucanas		protein LOC108908596 isoform X2	Coleopter	XP_023	4.02	864	105.		
45439-RA	e E-4	906	[Anoplophora glabripennis]	а	310346	E-23	4	145	118	84
	endogl									
	ucanas						68.3			
Th_d_000	e E-4-	135	gi 1022761055 gb KZS07093.1 End		KZS070	1.2E	615	370.		
34989-RA	like	0	oglucanase [Daphnia magna]	Crustacea	93	-119	8	548	354	242

	C	Seq uen				F	Circui	Dit	Align	Deel
Sequenc	nce	ce leng				E- Valu	larit	Bit- Scor	lengt	tive
e name	desc.	th	Hit desc.	Order	Hit ACC	e	у	е	h	S
			gi 1339087696 ref XP_023704929.							
			1 uncharacterized protein							
			LOC111863126 [Cryptotermes		XP_023					
	Endogl		secundus]gi 1330920323 gb PNF36		704929,		63.7			
Th_d_000	ucanas	132	365.1 Endoglucanase A		PNF363	1.5E	002	379.		
31500-RA	e A	6	[Cryptotermes secundus]	Blattodea	65	-126	3	407	427	272
			gi 1339087696 ref XP_023704929.							
			1 uncharacterized protein							
			LOC111863126 [Cryptotermes		XP_023					
	Endogl		secundus]gi 1330920323 gb PNF36		704929,		61.0			
Th_d_000	ucanas	186	365.1 Endoglucanase A		PNF363	3.25	169	60.0		
40844-RA	e A	0	[Cryptotermes secundus]	Blattodea	65	E-07	5	77	59	36
	endogl									
	ucanas		gi 1059424762 ref XP_017785110.				84.7			
Th_d_000	e E-4-		1 PREDICTED: endoglucanase E-4-	Coleopter	XP_017	4.64	457	98.9		
97963-RA	like	408	like [Nicrophorus vespilloides]	а	785110	E-23	6	821	59	50
	endogl									
	ucanas		gi 1228018665 ref XP_021941322.				73.6			
Th_d_000	e 15-		1 endoglucanase 7-like, partial		XP_021	3.53	842	204.		
60602-RA	like	771	[Zootermopsis nevadensis]	Blattodea	941322	E-64	1	527	152	112

Coquene	Seque	Seque				E-	Simi	Bit-	Align ment	Posi
e name	desc.	length	Hit desc.	Order	Hit ACC	e valu	lai lu V	e	h	s
	Endogl	Tengen				-	y	-		5
	ucanas									
	e E-4									
	precur									
	sor,		gi 197691959 dbj BAG70031.1 e				63.7			
Th_d_000	putativ		ndo-beta-1,4-glucanase, partial		BAG700	5.4E	362	99.3		
85360-RA	е	360	[Panesthia angustipennis]	Blattodea	31	-24	6	673	91	58
	AF220									
	596_1b									
	eta-									
	1,4-						60.4			
T I 1000	endogl		gi 8886827 gb AAF80584.1 AF2			1 60	69.1	500		
In_d_000	ucanas	2020	20596_1beta-1,4-endoglucanase	Distribut	AAF805	1.5E	489	508.	470	225
07226-RA	eı	2020	I [Panestnia cribrata]	Blattodea	84	-1/3	4	449	4/0	325
			gi 1233161131 ref XP_02220000							
	ondogl		LOC111056001 [Nilaparvata		VD 022					
	ucanas		lugonelgil820907584lgblAKI 904		200004		00.3			
Th d 000	$\rho 15$		11 1lendo-beta-1 4-glucanase	Heminter	ΔΚΙ 904	1 38	90.3 846	95 5		
62804-RA	like	432	[Nilanarvata lugens]	a	11	F-21	2	153	52	47
02001101	endogl	152		a	11		2	155	52	17
	ucanas		gil9853883481reflXP 015378728							
Th d 000	e E-4-		.1 PREDICTED: endoglucanase E-	Hemipter	XP 015	4.06	69.5	88.5		
44871-RA	like	864	4-like [Diuraphis noxia]	a	378728	E-19	122	817	82	57

Table 3	3.1. C	ontin	ued.
---------	--------	-------	------

									Align	
	Seque	Seque				E-	Simi	Bit-	ment	Posi
Sequenc	nce	nce				Valu	larit	Scor	lengt	tive
e name	desc.	length	Hit desc.	Order	Hit ACC	е	У	е	h	S
	endogl									
	ucanas		gi 646711640 gb KDR16731.1 E				87.9			
Th_d_000	e E-4-		ndoglucanase F [Zootermopsis		KDR167	4.8E	598	446.		
48316-RA	like	1194	nevadensis]	Blattodea	31	-153	7	047	299	263
	putativ									
	e endo-									
	beta-									
	1,4-		gi 1228007791 ref XP_02193580							
	glucan		7.1 uncharacterized protein				73.2			
Th_d_000	ase		LOC110837698, partial		XP_021	2.43	876	177.		
81071-RA	HsEG2	849	[Zootermopsis nevadensis]	Blattodea	935807	E-51	7	178	146	107
	AF220									
	594_1b									
	eta-		gi 7546878 gb AAF63725.1 AF2							
	1,4-		20594_1beta-1,4-glucanase 2,				72.7			
Th_d_000	glucan		partial [Mastotermes		AAF637	1.33	272	127.		
36851-RA	ase 2	507	darwiniensis]	Blattodea	25	E-34	7	872	121	88
	endo-		gi 1228018669 ref XP_02194132							
	beta-		5.1 LOW QUALITY PROTEIN:							
	1,4-		uncharacterized protein				65.8			
Th_d_000	glucan		LOC110840537 [Zootermopsis		XP_021	3.31	333	115.		
91320-RA	ase 1	441	nevadensis]	Blattodea	941325	E-29	3	161	120	79

Table 3	3.1. C	ontin	ued.
---------	--------	-------	------

	6	Comme				г	Circl	D'4	Align	Deel
Sequenc	seque nce	seque nce				E- Valu	Simi larit	Bit-	lengt	tive
e name	desc.	length	Hit desc.	Order	Hit ACC	e	y	e	h	S
	AF220									
	583_1b									
	eta-									
	1,4-		gi 7546856 gb AAF63714.1 AF2				74.7			
Th_d_000	glucan		20583_1beta-1,4-glucanase 3,		AAF637	1.83	967	142.		
79350-RA	ase 3	450	partial [Polyphaga aegyptiaca]	Blattodea	14	E-40	5	124	123	92
	AF220									
	588_1b									
	eta-						5 0 (
	1,4-		gi 197691951 dbj BAG70027.1 e		D I G B G G G G G G G G G G		72.6			
Th_d_000	glucan	200	ndo-beta-1,4-glucanase, partial		BAG700	4.63	315	112.	05	(0)
10317-RA	ase 3	309	[Salganea esakii]	Blattodea	27	E-29	8	079	95	69
	AF220									
	588_10									
			~:1107(010471db:104C7002511a				70.4			
Th d 000	1,4- alucan		gi[19/09194/[ub]]DAG/0025.1[e		PAC700	1 55	79.4 20E	204		
00251 DA	giucan	1464	[Salganoa osakii]	Plattodaa	DAG700 25	1.55 E 50	205	204. 527	175	120
00331-KA	ase s	1404		Diattouea	23	E-39	/	527	1/5	139
	hota-									
	1 4-						80.6			
Th d 000	glucan		gil375151369ldbilBAL60587 11c		BAL605	2.93	451	100		
89290-RA	ase	309	ellulase [Neomysis intermedia]	Crustacea	87	E-24	6	908	62	50

	Seque	Seque				E-	Simi	Bit-	Align ment	Posi
Sequenc	nce	nce				Valu	larit	Scor	lengt	tive
e name	desc.	length	Hit desc.	Order	Hit ACC	е	У	е	h	S
	1,4-		gi 1339092712 ref XP_02370756							
	alpha-		9.1 1,4-alpha-glucan-branching							
	glucan-		enzyme [Cryptotermes							
	branch		secundus]gi 1330914862 gb PNF		XP_023					
	ing		33674.1 1,4-alpha-glucan-		707569,		91.2			
Th_d_000	enzym		branching enzyme [Cryptotermes		PNF336	9.67	280	101.		
75942-RA	е	225	secundus]	Blattodea	74	E-25	7	679	57	52
	endo-		gi 13537532 dbj BAB40693.1 en							
	beta-		do-b-1,4-glucanase [Coptotermes		BAB406					
	1,4-		formosanus]gi 13537534 dbj BA		93,		78.8			
Th_d_000	glucan		B40694.1 endo-b-1,4-glucanase		BAB406	6.7E	732	343.		
49037-RA	ase	997	[Coptotermes formosanus]	Blattodea	94	-114	4	199	284	224
	1,4-									
	alpha-									
	glucan-									
	branch		gi 795023516 ref XP_011860965				-			
	ing		.1 PREDICTED: 1,4-alpha-glucan-			0.00	56.3	100		
Th_d_000	enzym	(10	branching enzyme [Vollenhovia	Hymenop	XP_011	8.83	725	180.		
13433-RA	e	618	emeryi	tera	860965	E-51	5	259	204	115
	endo-									
m 1 1 0 0 0	b-1,4-		gi[13095576]gb]AAK12339.1 cell		4.4.144.000	1.05	73.2	1.5		
Th_d_000	glucan	4000	ulase [Coptotermes		AAK123	1.3E	584	476.		
35324-RA	ase	1332	acinaciformis	Blattodea	39	-164	3	478	445	326

	Seque	Seque				E-	Simi	Bit-	Align ment	Posi
Sequenc	nce	nce				Valu	larit	Scor	lengt	tive
e name	desc.	length	Hit desc.	Order	Hit ACC	е	У	е	h	S
	endo-									
	beta-									
	1,4-		gi 7546856 gb AAF63714.1 AF2	_			82.3			
Th_d_001	glucan		20583_1beta-1,4-glucanase 3,	Coleopter	AAF637	2.06	008	152.		
04889-RA	ase 1	474	partial [Polyphaga aegyptiaca]	а	14	E-44	8	525	113	93
	glycosi									
	de									
	nyaroi		a:15212122021ab14CD7(410.11ap							
Th d 001	ase		gi[521313293]gb]AGP76419.1]en		ACD764	1 52	607	00.0		
111_U_UUI 12062 DA		220	[Odontotormos formosonus]	Plattadaa	AGP/04	1.55 E 17	00.7 E	00.0 777	61	4.4
13003-KA	9 glycosi	220		Diattouea	19	L-17	5	///	04	44
	de									
	hvdrol									
	ase		gil992051711lgblAMH40365.1lg				80.7			
Th d 001	family		lycoside hydrolase family 9	Phamato	AMH40	1.35	692	75.0		
13037-RA	9	216	[Extatosoma tiaratum]	dea	365	E-15	3	998	52	42
	glycosi									
	de									
	hydrol		gi 1330895262 gb PNF24409.1							
	ase		hypothetical protein				85.9			
Th_d_000	family		B7P43_G09674, partial		PNF244	7.71	649	95.5		
71154-RA	9	822	[Cryptotermes secundus]	Blattodea	09	E-21	1	153	57	49

Sequenc	Seque nce	Sequenc				E- Valu	Simi larit	Bit- Scor	Align ment lengt	Posi tive
e name	desc.	e length	Hit desc.	Order	Hit ACC	е	У	е	h	S
Th_d_000	glycosi de hydrol ase family		gi 24940551 emb CAD54729.1 beta-1,4-endoglucanase		CAD547		79.5 348	561.		
35331-RA	9	1546	[Mastotermes darwiniensis]	Blattodea	29	0	8	222	430	342
Th d 001	glycosi de hydrol ase family		gi 992051866 gb AMH40374.1	Dhamato	АМН4О	8 96	68.1	70 7		
15973-RA	9	279	[Peruphasma schultei]	dea	374	E-17	4	221	69	47
Th_d_000 29874-RA	glycosi de hydrol ase family 9	1827	gi 1330724231 gb PNE09439.1 hypothetical protein B7P43_G00116 [Cryptotermes secundus]	Blattodea	PNE094 39	2.47 E-67	57.2 438 2	227.	283	162
Th d 000	glycosi de hydrol ase family		gi 1228018667 ref XP_021941 324.1 uncharacterized protein LOC110840536 [Zootermonsis		XP 021		73.6	533.		
10318-RA	9	1538	nevadensis]	Blattodea	941324	0	6	102	433	319

	Seque	6				E-	Simi	Bit-	Align ment	Posi
Sequenc e name	nce desc.	Sequenc e length	Hit desc.	Order	Hit ACC	valu e	larit v	Scor e	lengt h	tive S
Th d 000	glycosi de hydrol ase family	e rengen	gi 646689408 gb KDR06579.1 Endoglucanase 1 [Zootermopsis		KDR065	8.71	68.6 440	142.		5
10316-RA	9	375	nevadensis]	Blattodea	79	E-43	7	51	118	81
Th_d_000	glycosi de hydrol ase family		gi 24940551 emb CAD54729.1 beta-1,4-endoglucanase		CAD547		78.4	563.		
07347-RA	9	1482	[Mastotermes darwiniensis]	Blattodea	29	0	897	148	437	343
Th_d_000 07348-RA	glycosi de hydrol ase family 9	1555	gi 24940553 emb CAD54730.1 beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattodea	CAD547 30	0	79.5 294 1	556. 214	425	338
Th_d_000	glycosi de hydrol ase family		gi 13095576 gb AAK12339.1 c ellulase [Coptotermes		AAK123	6.41	77.3	163.		
00352-RA	9	465	acinaciformis]	Blattodea	39	E-47	913	696	115	89

Sequenc	Seque nce dosc	Sequenc	Hitdosc	Ordor	Hit ACC	E- Valu	Simi larit	Bit- Scor	Align ment lengt	Posi tive
ename	alvcosi	elengui		order	IIIt ACC	C	У	C	11	3
	de									
	hvdrol		gil1330895262lgblPNF24409.1							
	ase		hypothetical protein							
Th_d_000	family		B7P43_G09674, partial		PNF244	1.69	75.4	137.		
93510-RA	9	282	[Cryptotermes secundus]	Blattodea	09	E-39	717	887	106	80
	glycosi									
	de									
	hydrol		gi 1330895262 gb PNF24409.1							
	ase		hypothetical protein				80.8			
Th_d_000	family		B7P43_G09674, partial		PNF244	1.08	219	293.		
18984-RA	9	1937	[Cryptotermes secundus]	Blattodea	09	E-92	2	893	219	177
	glycosi									
	de									
	hydrol									
	ase		gi 646711640 gb KDR16731.1				84.4			
Th_d_000	family		Endoglucanase F [Zootermopsis		KDR167	1.39	155	110.		
89430-RA	9	495	nevadensis]	Blattodea	31	E-26	8	153	77	65

		Sequ							Align	
		ence				Е-	Simi	Bit-	ment	Posi
Sequenc	Sequence	lengt				Valu	larit	Scor	lengt	tive
e name	desc.	h	Hit desc.	Order	Hit ACC	е	У	e	h	S
	glycoside		gi 315570658 gb ADU33333.1							
	hydrolase		glycoside hydrolase family			0.00	71.4			
Th_d_001	family		protein 5 [Gastrophysa	Coleopter	ADU333	074	285	40.8		
35584-RA	protein 5	144	viridula]	а	33	9	7	17	35	25
			gi 1228018667 ref XP_021941							
	glycoside		324.1 uncharacterized protein				80.6			
Th_d_000	hydrolase		LOC110840536 [Zootermopsis		XP_021	9.34	603	298.		
15659-RA	family 9	2150	nevadensis]	Blattodea	941324	E-92	8	901	212	171
			gi 1330895262 gb PNF24409.1							
	glycoside		hypothetical protein							
Th_d_000	hydrolase		B7P43_G09674, partial		PNF244	1.17		272.		
79172-RA	family 9	606	[Cryptotermes secundus]	Blattodea	09	E-90	82	707	200	164
	glycoside		gi 992052122 gb AMH40384.1				79.6			
Th_d_000	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	2.15	610	96.6		
90714-RA	family 9	360	[Sipyloidea sipylus]	dea	384	E-22	2	709	59	47
			gi 1330895262 gb PNF24409.1							
	glycoside		hypothetical protein				84.4			
Th_d_000	hydrolase		B7P43_G09674, partial		PNF244	2.13	827	95.9		
78933-RA	family 9	723	[Cryptotermes secundus]	Blattodea	09	E-21	6	005	58	49
	glycoside		gi 992051639 gb AMH40360.1				68.5			
Th_d_000	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	6.36	106	226.		
74140-RA	family 9	657	[Aretaon asperrimus]	dea	360	E-70	4	098	235	161

		Sequ							Align	
		ence				E-	Simi	Bit-	ment	Posi
Sequenc	Sequence	lengt				Valu	larit	Scor	lengt	tive
e name	desc.	h	Hit desc.	Order	Hit ACC	е	У	е	h	S
	glycoside		gi 992051639 gb AMH40360.1							
Th_d_001	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	3.79		124.		
00655-RA	family 9	498	[Aretaon asperrimus]	dea	360	E-32	65	79	120	78
	glycoside		gi 992051639 gb AMH40360.1				76.6			
Th_d_000	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	4.82	666	173.		
96112-RA	family 9	612	[Aretaon asperrimus]	dea	360	E-50	7	711	120	92
	glycoside		gi 24940551 emb CAD54729.1				79.7			
Th_d_000	hydrolase		beta-1,4-endoglucanase		CAD547		674	561.		
19623-RA	family 9	1554	[Mastotermes darwiniensis]	Blattodea	29	0	4	222	430	343
			gi 1228018667 ref XP_021941							
	glycoside		324.1 uncharacterized protein				72.3			
Th_d_000	hydrolase		LOC110840536 [Zootermopsis		XP_021	2.4E	255	518.		
35252-RA	family 9	1463	nevadensis]	Blattodea	941324	-180	8	464	430	311
			gi 1330895262 gb PNF24409.1							
	glycoside		hypothetical protein				76.1			
Th_d_001	hydrolase		B7P43_G09674, partial		PNF244	6.4E	467	135.		
11221-RA	family 9	567	[Cryptotermes secundus]	Blattodea	09	-37	9	191	109	83
	glycoside		gi 646689408 gb KDR06579.1				71.0			
Th_d_000	hydrolase		Endoglucanase 1 [Zootermopsis		KDR065	1.89	526	141.		
61520-RA	family 9	351	nevadensis]	Blattodea	79	E-42	3	354	114	81
	glycoside		gi 24940551 emb CAD54729.1				79.7			
Th_d_000	hydrolase		beta-1,4-endoglucanase		CAD547		674	559.		
05705-RA	family 9	1593	[Mastotermes darwiniensis]	Blattodea	29	0	4	681	430	343

Table 3.1.	Continued.
------------	------------

		Sequ							Align	
		ence				E-	Simi	Bit-	ment	Posi
Sequenc	Sequence	lengt				Valu	larit	Scor	lengt	tive
e name	desc.	h	Hit desc.	Order	Hit ACC	е	У	e	h	S
	glycoside		gi 992051639 gb AMH40360.1				77.1			
Th_d_000	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	3.17	186	173.		
94204-RA	family 9	399	[Aretaon asperrimus]	dea	360	E-51	4	711	118	91
	glycoside		gi 992051639 gb AMH40360.1				82.3			
Th_d_001	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	1.6E	529	166.		
04189-RA	family 9	514	[Aretaon asperrimus]	dea	360	-47	4	007	102	84
			gi 1228018667 ref XP_021941							
	glycoside		324.1 uncharacterized protein				76.0			
Th_d_000	hydrolase		LOC110840536 [Zootermopsis		XP_021	6.6E	606	420.		
46715-RA	family 9	1068	nevadensis]	Blattodea	941324	-144	1	239	330	251
	glycoside		gi 24940551 emb CAD54729.1				78.4			
Th_d_000	hydrolase		beta-1,4-endoglucanase		CAD547		269	566.		
28499-RA	family 9	1779	[Mastotermes darwiniensis]	Blattodea	29	0	7	614	445	349
	glycoside		gi 992051639 gb AMH40360.1				78.6			
Th_d_000	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	6E-	301	497.		
38486-RA	family 9	1307	[Aretaon asperrimus]	dea	360	173	4	664	365	287
			gi 8886829 gb AAF80585.1 AF							
	glycoside		220597_1beta-1,4-				56.0			
Th_d_000	hydrolase		endoglucanase 2 [Panesthia		AAF805	1.22	509	95.9		
91520-RA	family 9	471	cribrata]	Blattodea	85	E-21	6	005	157	88
	glycoside		gi 24940551 emb CAD54729.1				79.8			
Th_d_000	hydrolase		beta-1,4-endoglucanase		CAD547		122	556.		
34779-RA	family 9	1436	[Mastotermes darwiniensis]	Blattodea	29	0	1	599	426	340

Table 3.1.	Continued.
------------	------------

		Sequ							Align	
		ence				E-	Simi	Bit-	ment	Posi
Sequenc	Sequence	lengt				Valu	larit	Scor	lengt	tive
e name	desc.	h	Hit desc.	Order	Hit ACC	e	У	е	h	S
	glycoside		gi 13095576 gb AAK12339.1 c				72.5			
Th_d_000	hydrolase		ellulase [Coptotermes		AAK123		701	545.		
14865-RA	family 9	1574	acinaciformis]	Blattodea	39	0	9	428	463	336
			gi 1330895262 gb PNF24409.1							
	glycoside		hypothetical protein							
Th_d_000	hydrolase		B7P43_G09674, partial		PNF244	1.17		272.		
75629-RA	family 9	606	[Cryptotermes secundus]	Blattodea	09	E-90	82	707	200	164
	glycoside		gi 992051639 gb AMH40360.1				78.9			
Th_d_000	hydrolase		glycoside hydrolase family 9	Phamato	AMH40	1.1E	473	378.		
53691-RA	family 9	1006	[Aretaon asperrimus]	dea	360	-127	7	252	285	225
	PREDICTE									
	D:									
	uncharact		gi 985403057 ref XP_0153693							
	erized		26.1 PREDICTED:							
	protein		uncharacterized protein				56.5			
Th_d_001	LOC10716		LOC107165552 Diuraphis	Hemipter	XP_015	6.8E	217	57.7		
30530-RA	5552	321	noxia]	a	369326	-09	4	658	92	52

blabe acber	perom									
		Sequ ence				E-		Bit-	Align ment	Posi
Sequenc	Sequenc	lengt				Valu	Simil	Scor	lengt	tive
e name	e desc.	h	Hit desc.	Order	Hit ACC	е	arity	е	h	S
	uncharac		gi 1228365487 ref XP_021943634.							
	terized		1 uncharacterized family 31							
	family		glucosidase KIAA1161-like							
	31		[Folsomia							
	glucosid		candida]gi 1228365489 ref XP_02		XP_021					
Th_d_00	ase		1943636.1 uncharacterized family		943634,					
054985-	KIAA116		31 glucosidase KIAA1161-like	Collem	XP_021	4.7E	58.15	350.9		
RA	1-like	2754	[Folsomia candida]	bola	943636	-107	9	03	478	278
	uncharac									
	terized									
	family									
	31									
	glucosid		gi 1101351520 ref XP_018901738.							
Th_d_00	ase		1 PREDICTED: uncharacterized							
105563-	KIAA116		family 31 glucosidase KIAA1161-	Hemipt	XP_018	2.21	87.30	104.3		
RA	1-like	288	like [Bemisia tabaci]	era	901738	E-25	159	75	63	55
	uncharac									
	terized									
	family									
	31									
	glucosid									
Th_d_00	ase		gi 636630780 gb AIA09350.1 alph					1005		
032903-	KIAA116		a-glucosidase family 31, partial	Blattod	AIA093	1.4E	62.47	428.3		
RA	1-like	2130	[Periplaneta americana]	ea	50	-139	379	28	477	298

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

Table 3.2. Continued	l.
----------------------	----

		Sequ ence				E-		Bit-	Align ment	Posi
Sequenc	Sequenc	lengt				Valu	Simil	Scor	lengt	tive
e name	e desc.	h	Hit desc.	Order	Hit ACC	e	arity	е	h	S
	glucosid		gi 1339077375 ref XP_023726071.							
Th_d_00	ase 2		1 glucosidase 2 subunit beta							
062466-	subunit		isoform X1 [Cryptotermes	Blattod	XP_023	1.55	74.64	93.20		
RA	beta	633	secundus]	ea	726071	E-21	789	41	71	53
	uncharac									
	terized									
	family									
	31									
	glucosid									
Th_d_00	ase		gi 577754846 gb AHH86052.1 gly							
050423-	KIAA116		coside hydrolase family 31	Coleop	AHH86	3E-	58.69	478.4		
RA	1-like	2034	[Callosobruchus maculatus]	tera	052	159	565	04	598	351
	uncharac									
	terized									
	family									
	31									
	glucosid									
Th_d_00	ase		gi 636630780 gb AIA09350.1 alph							
142013-	KIAA116		a-glucosidase family 31, partial	Blattod	AIA093	1.4E	62.76	434.8		
RA	1-like	1404	[Periplaneta americana]	ea	50	-145	596	76	470	295

		Seque				F_		Dit-	Align	Posi
Sequence	Sequenc	lengt				Valu	Simil	Scor	lengt	tive
name	e desc.	h	Hit desc.	Order	Hit ACC	е	arity	е	h	S
	uncharact									
	erized									
	family 31		gi 1199397963 ref XP_0211							
	glucosida		94260.1 uncharacterized							
	se		family 31 glucosidase							
Th_d_000614	KIAA116		KIAA1161-like [Helicoverpa	Lepido	XP_021	2.5E	56.64	503.8		
35-RA	1-like	1950	armigera]	ptera	194260	-169	557	27	632	358
	AChain A,									
	Crystal									
	Structure									
	Of Beta-									
	glucosida									
	se From									
	Termite									
	Neoterme									
	S									
	Koshunen									
	sis In									
	Complex									
	With A									
	New									
	Glucopyr		gi 1330878954 gb PNF1683							
Th_d_001145	anosidic		8.1 Myrosinase 1, partial	Blattod	PNF168	1.45	86.41	133.2		
90-RA	Product	246	[Cryptotermes secundus]	ea	38	E-36	975	65	81	70

		Seque							Align	
	_	nce				E-		Bit-	ment	Posi
Sequence	Sequenc	lengt				Valu	Simil	Scor	lengt	tive
name	e desc.	h	Hit desc.	Order	Hit ACC	e	arity	е	h	S
	glucosida		gi 1228013272 ref XP_0219							
	se 2		38600.1 glucosidase 2							
Th_d_000257	subunit		subunit beta [Zootermopsis	Blattod	XP_021	1.37	75.30	104.3		
73-RA	beta	453	nevadensis]	ea	938600	E-24	864	75	81	61
	glucosida		gi 1228013272 ref XP_0219							
	se 2		38600.1 glucosidase 2							
Th_d_000257	subunit		subunit beta [Zootermopsis	Blattod	XP_021	4.17		80.10		
72-RA	beta	243	nevadensis]	ea	938600	E-17	76	73	50	38
	uncharact									
	erized									
	family 31		gi 1101351520 ref XP_0189							
	glucosida		01738.1 PREDICTED:							
	se		uncharacterized family 31							
Th_d_001490	KIAA116		glucosidase KIAA1161-like	Hemipt	XP_018	5.7E	86.15	103.2		
67-RA	1-like	468	[Bemisia tabaci]	era	901738	-24	385	19	65	56
	uncharact									
	erized									
	family 31									
	glucosida		gi 646703149 gb KDR11965							
	se		.1 hypothetical protein							
Th_d_000983	KIAA116		L798_13618, partial	Blattod	KDR119	1.63	81.41	170.6		
42-RA	1-like	396	[Zootermopsis nevadensis]	ea	65	E-49	593	29	113	92

		Seq							Ali gn m en t	
Seque		nce				E-		Bit-	le	Posi
nce	Sequenc	len			Hit	Valu	Simil	Scor	ng	tive
name	e desc.	gth	Hit desc.	Order	ACC	e	arity	е	th	S
			gi 303324839 pdb 3AHZ AChain A,							
	AChain		Crystal Structure Of Beta-Glucosidase							
	А,		From Termite Neotermes Koshunensis							
	Crystal		In Complex With							
	Structur		Trisgi 393715252 pdb 3VIF AChain A,							
	e Of		Crystal Structure Of Beta-glucosidase							
	Beta-		From Termite Neotermes Koshunensis							
	glucosid		In Complex With							
	ase From		Gluconolactonegi 393715253 pdb 3VIG							
	Termite		AChain A, Crystal Structure Of Beta-							
	Neoterm		glucosidase From Termite Neotermes							
	es		Koshunensis In Complex With 1-							
	Koshune		deoxynojirimycingi 393715254 pdb 3VI							
	nsis In		H AChain A, Crystal Structure Of Beta-		3AHZ_					
	Complex		glucosidase From Termite Neotermes		А,					
	With A		Koshunensis In Complex With		3VIF_A,					
Th_d_	New		Glycerolgi 393715255 pdb 3VII AChain		3VIG_A					
00018	Glucopyr		A, Crystal Structure Of Beta-glucosidase		,					
632-	anosidic	131	From Termite Neotermes Koshunensis		3VIH_A	4.1E	75.23	517.3	43	
RA	Product	7	In Complex With Bis-tris	Blattodea	, 3VII_A	-180	148	09	2	325

Sequen		Seque				E-		Bit-	Align ment	Posi
ce	Sequence	nce		Orde		- Valu	Simil	Sco	lengt	tive
name	desc.	length	Hit desc.	r	Hit ACC	е	arity	re	h	s
	uncharacteri									
	zed family									
	31		gi 646703149 gb KDR11965.1 h							
Th_d_00	glucosidase		ypothetical protein L798_13618,					169		
057277	KIAA1161-		partial [Zootermopsis	Blatt	KDR1196	4.78	80.70	.85		
-RA	like	426	nevadensis]	odea	5	E-49	175	9	114	92
			gi 1032766152 ref XP_0168414							
			89.1 PREDICTED: lysosomal							
			alpha-glucosidase-like [Nasonia							
	lysosomal		vitripennis]gi 1032766154 ref X		XP_01684					
Th_d_00	alpha-		P_016841490.1 PREDICTED:	Hym	1489,			61.		
069820	glucosidase-		lysosomal alpha-glucosidase-like	enopt	XP_01684	6.08	55.30	617		
-RA	like	1173	[Nasonia vitripennis]	era	1490	E-08	303	8	132	73
			gi 1227984161 ref XP_0219237							
			30.1 myrosinase 1-like isoform							
			X1 [Zootermopsis							
			nevadensis]gi 646712919 gb K		XP_02192					
Th_d_00			DR17465.1 Lactase-phlorizin		3730,			169		
081027	beta-		hydrolase [Zootermopsis	Blatt	KDR1746	2.56	63.15	.47		
-RA	glucosidase	471	nevadensis]	odea	5	E-48	789	4	152	96
			gi 1339077375 ref XP_0237260							
Th_d_00	glucosidase		71.1 glucosidase 2 subunit beta					93.		
106209	2 subunit		isoform X1 [Cryptotermes	Blatt	XP_02372	1.74	57.84	204		
-RA	beta	633	secundus]	odea	6071	E-21	314	1	102	59

6		C				F		D:4	Align	Deel
Sequen	Sequence	Seque		Ordo	U;+	E- Volu	Cimil	BIT-	ment	POSI
le namo	dosc	longth	Hit dosc	r		valu	ority	ro	h	uve
name	uest.	length	m(1229002605)rof(VD 021022722)	1	ALL	e	anty	Te	11	5
			1 Juncharactorized family 21							
			alucosidase KIA A1161-like isoform							
	uncharacteri		X1 [7ootermonsis		XP 021					
	zed family		nevadensis]gi 646703150 gh KDR		93373					
Th d 00	31		11966 1 Inutative family 31		3			374		
099763	glucosidase		glucosidase [Zootermonsis	Blatt	KDR11	1.7E	54.18	.78		
-RA	KIAA1161	1347	nevadensisl	odea	966	-121	251	5	526	285
		2017	gi 1339041694 ref XP 023710162.	0000	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,			<u> </u>	020	
			1 neutral alpha-glucosidase AB		XP 023					
	neutral		Cryptotermes		71016					
Th_d_00	alpha-		secundus]gi 1330933432 gb PNF4		2,			191		
000860	glucosidase		2738.1 Neutral alpha-glucosidase	Blatt	PNF42	3.75	82.44	.04		
-RA	AB	414	AB [Cryptotermes secundus]	odea	738	E-55	275	5	131	108
	Neutral		gi 1227991790 ref XP_021927674.							
Th_d_00	alpha-		1 neutral alpha-glucosidase AB		XP_021			166		
000862	glucosidase		isoform X2 [Zootermopsis	Blatt	92767	1.85	85.59	.77		
-RA	AB	435	nevadensis]	odea	4	E-46	322	7	118	101
			gi 1227991792 ref XP_021927675.							
			1 neutral alpha-glucosidase AB		XP_021					
	Neutral		isoform X3 [Zootermopsis		92767					
Th_d_00	alpha-		nevadensis]gi 646709565 gb KDR		5,			97.		
000861	glucosidase		15365.1 Neutral alpha-glucosidase	Blatt	KDR15	7.97	94.11	441		
-RA	AB	315	AB [Zootermopsis nevadensis]	odea	365	E-23	765	3	51	48

									Align	
Sequen		Seque				E-		Bit-	ment	Posi
се	Sequence	nce		Orde	Hit	Valu	Simil	Sco	lengt	tive
name	desc.	length	Hit desc.	r	ACC	е	arity	re	h	S
	Neutral									
Th_d_00	alpha-		gi 930677078 gb KPJ17206.1 Neut	Lepid				173		
000863	glucosidase		ral alpha-glucosidase AB [Papilio	opter	KPJ172	1.18	80.53	.71		
-RA	AB	519	machaon]	а	06	E-53	097	1	113	91
			gi 1339041694 ref XP_023710162.							
			1 neutral alpha-glucosidase AB		XP_023					
	neutral		[Cryptotermes		71016					
Th_d_00	alpha-		secundus]gi 1330933432 gb PNF4		2,			337		
000865	glucosidase		2738.1 Neutral alpha-glucosidase	Blatt	PNF42	1.4E	70.22	.42		
-RA	AB	1599	AB [Cryptotermes secundus]	odea	738	-103	059	1	272	191
			gi 242019253 ref XP_002430076.1							
			Neutral alpha-glucosidase AB							
			precursor, putative [Pediculus							
			humanus		XP_002					
	Neutral		corporis]gi 212515157 gb EEB173		43007					
Th_d_00	alpha-		38.1 Neutral alpha-glucosidase AB	Phthi	6,			157		
000864	glucosidase		precursor, putative [Pediculus	rapte	EEB17	2.81	74.25	.14		
-RA	AB	633	humanus corporis]	ra	338	E-42	743	7	101	75
	Neutral									
Th_d_00	alpha-		gi 1108484091 emb CRK87777.1					90.		
000866	glucosidase		CLUMA_CG001536, isoform A	Dipte	CRK87	2.15	75.36	892		
-RA	AB	780	[Clunio marinus]	ra	777	E-18	232	9	69	52

Table 3.2. Continued	l.
----------------------	----

		Seque							Align	
Sequen	Seque	nce				E-		Bit-	ment	Posi
ce	nce	lengt			Hit	Val	Simil	Sco	lengt	tive
name	desc.	h	Hit desc.	Order	ACC	ue	arity	re	h	S
			gi 1228005382 ref XP_021934582.1 ma							
			nnosyl-oligosaccharide glucosidase		XP_021					
			isoform X2 [Zootermopsis		93458					
			nevadensis]gi 1228005384 ref XP_0219		2,					
			34583.1 mannosyl-oligosaccharide		XP_021					
			glucosidase isoform X2 [Zootermopsis		93458					
			nevadensis]gi 1228005386 ref XP_0219		3,					
			34584.1 mannosyl-oligosaccharide		XP_021					
			glucosidase isoform X2 [Zootermopsis		93458					
			nevadensis]gi 1228005388 ref XP_0219		4,					
			34585.1 mannosyl-oligosaccharide		XP_021					
			glucosidase isoform X2 [Zootermopsis		93458					
			nevadensis]gi 1228005390 ref XP_0219		5,					
			34586.1 mannosyl-oligosaccharide		XP_021					
			glucosidase isoform X2 [Zootermopsis		93458					
	manno		nevadensis]gi 1228005392 ref XP_0219		6,					
	syl-		34587.1 mannosyl-oligosaccharide		XP_021					
	oligosa		glucosidase isoform X2 [Zootermopsis		93458					
	ccharid		nevadensis]gi 1228005394 ref XP_0219		7,					
	e		34588.1 mannosyl-oligosaccharide		XP_021					
	glucosi		glucosidase isoform X2 [Zootermopsis		93458					
Th_d_00	dase		nevadensis]gi 646701956 gb KDR11423.		8,			104		
017412	isofor		1 Mannosyl-oligosaccharide glucosidase	Blatto	KDR11		73.50	8.8		
-RA	m X2	6713	[Zootermopsis nevadensis]	dea	423	0	917	8	872	641

									Alig nme	
Sequen						E-		Bit-	nt	Posi
ce		Sequenc		Orde	Hit	Valu	Simil	Scor	leng	tive
name	Sequence desc.	e length	Hit desc.	r	ACC	е	arity	e	th	S
			gi 1227993488 ref XP_021							
			928525.1 lysosomal alpha-							
			glucosidase-like							
			[Zootermopsis		XP_021					
			nevadensis]gi 646708767		92852					
Th_d_00			gb KDR14932.1 Lysosomal		5,					
050585	lysosomal alpha-		alpha-glucosidase	Blatt	KDR14	3.6E	86.31	440.		
-RA	glucosidase-like	954	[Zootermopsis nevadensis]	odea	932	-145	579	654	285	246
			gi 1228365487 ref XP_021							
			943634.1 uncharacterized							
			family 31 glucosidase							
			KIAA1161-like [Folsomia		XP_021					
			candida]gi 1228365489 re		94363					
	uncharacterized		f XP_021943636.1 unchara		4,					
Th_d_00	family 31		cterized family 31	Colle	XP_021					
079942	glucosidase		glucosidase KIAA1161-like	mbol	94363	6.9E	57.80	427.		
-RA	KIAA1161-like	2136	[Folsomia candida]	а	6	-139	347	943	519	300
			gi 1067098296 ref XP_018							
			011173.1 PREDICTED:							
Th_d_00			maltase-glucoamylase,		XP_018					
081056	lysosomal alpha-		intestinal-like [Hyalella	Crust	01117	3.49	44.44	70.0		
-RA	glucosidase	519	azteca]	acea	3	E-12	444	922	126	56

Sequen ce		Sequ ence lengt		Orde	Hit	E- Valu	Simil	Bit- Scor	Alig nme nt leng	Posi tive
name	Sequence desc.	h	Hit desc.	r	ACC	е	arity	е	th	S
	AChain A, Crystal									
	Structure Of Beta-									
	glucosidase From									
	Termite									
	Neotermes									
	Koshunensis In									
	Complex With A									
Th_d_00	New		gi 1061476505 gb 0DM94351	Colle						
041545	Glucopyranosidic		.1 Lactase-phlorizin hydrolase	mbol	ODM9	3.44	79.83	181.		
-RA	Product	487	[Orchesella cincta]	а	4351	E-53	871	415	124	99
	uncharacterized		gi 646703149 gb KDR11965.1							
Th_d_00	family 31		hypothetical protein							
055895	glucosidase		L798_13618, partial	Blatt	KDR11	6.34	78.33	85.5		
-RA	KIAA1161-like	261	[Zootermopsis nevadensis]	odea	965	E-19	333	001	60	47
	uncharacterized		gi 1228003699 ref XP_021933							
Th_d_00	family 31		735.1 uncharacterized family		XP_021					
055894	glucosidase		31 glucosidase KIAA1161-like	Blatt	93373	1.06	69.41	98.2		
-RA	KIAA1161-like	390	[Zootermopsis nevadensis]	odea	5	E-22	176	117	85	59
Th_d_00			gi 1090651209 gb AOY34571.							
014523			1 beta-glucosidase	Blatt	AOY34	1.22	77.35	134.		
-RA	beta-glucosidase	423	[Coptotermes formosanus]	odea	571	E-35	849	035	106	82

		Sequ				_			Align	_
Sequen		ence		Ordo	U;+	E- Volu	Simi	Bit-	ment	Posi
ce namo	Soguanca dosc	h	Hit dosc	r		valu		Scor	h	uve
name	uncharactorized	11	ril6266207901ablAIA002E0.11	1	ALL	e	У	е	11	3
	family 21		gilosoosu/oulgulAIA09550.1				72 5			
In_a_00	family 31		alpha-glucosidase family 31,	חות	41400	FD	/2.5	400		
011270	glucosidase	0.4.60	partial [Periplaneta	Blatt	AIA09	5E-	060	499.		200
-RA	KIAA1161-like	2460	americana	odea	350	166	8	975	411	298
	uncharacterized		gi 1228003699 ref XP_021933							
Th_d_00	family 31		735.1 uncharacterized family		XP_02		81.1			
011271	glucosidase		31 glucosidase KIAA1161-like	Blatt	19337	6.06	320	83.5		
-RA	KIAA1161-like	309	[Zootermopsis nevadensis]	odea	35	E-18	8	741	53	43
	uncharacterized		gi 1228003699 ref XP_021933							
Th_d_00	family 31		735.1 uncharacterized family		XP_02					
011273	glucosidase		31 glucosidase KIAA1161-like	Blatt	19337	3.35		93.5		
-RA	KIAA1161-like	366	[Zootermopsis nevadensis]	odea	35	E-21	80	893	60	48
	uncharacterized		gi 636630780 gb AIA09350.1							
Th_d_00	family 31		alpha-glucosidase family 31,							
154707	glucosidase		partial [Periplaneta	Blatt	AIA09	4.3E	71.6	474.		
-RA	KIAA1161-like	1431	americana]	odea	350	-161	707	937	413	296
			gi 1339040796 ref XP_023708							
			840.1 myogenesis-regulating							
			glycosidase [Cryptotermes		XP_02					
	uncharacterized		secundus]gi 1330911883 gb P		37088					
Th_d_00	family 31		NF32422.1 hypothetical		40,					
105015	glucosidase		protein B7P43 G04891	Blatt	PNF3	7.38	84.0	171.		
-RA	KIAA1161-like	357	[Cryptotermes secundus]	odea	2422	E-49	708	014	113	95

Table 3.2.	Continued.
------------	------------

									Align	
Sequen		Sequen				E-	Simi	Bit-	ment	Posi
ce		ce		Orde	Hit	Valu	larit	Scor	lengt	tive
name	Sequence desc.	length	Hit desc.	r	ACC	е	У	е	h	S
	uncharacterized		gi 636630780 gb AIA09350.1							
Th_d_00	family 31		alpha-glucosidase family 31,				64.2			
141763	glucosidase		partial [Periplaneta	Blatt	AIA09	1.6E	241	435.		
-RA	KIAA1161-like	1434	americana]	odea	350	-145	4	261	464	298
Th_d_00			gi 1152526136 gb AQW43010				75.8			
099575	beta-		.1 beta-glucosidase	Hemi	AQW4	2.22	064	154.		
-RA	glucosidase	417	[Nilaparvata lugens]	ptera	3010	E-46	5	836	124	94
	AChain A,									
	Crystal									
	Structure Of									
	Beta-									
	glucosidase									
	From Termite									
	Neotermes									
	Koshunensis In		gi 393715256 pdb 3VIJ AChai							
	Complex With A		n A, Crystal Structure Of Beta-							
Th_d_00	New		glucosidase From Termite				86.7			
032476	Glucopyranosid		Neotermes Koshunensis In	Blatt		7.05	469	137.		
-RA	ic Product	270	Complex With Glucose	odea	3VIJ_A	E-38	9	502	83	72
			gi 1067098296 ref XP_018011							
Th_d_00	Lysosomal		173.1 PREDICTED: maltase-		XP_01		56.6			
035285	alpha-		glucoamylase, intestinal-like	Crust	80111	1.28	371	199.		
-RA	glucosidase	1962	[Hyalella azteca]	acea	73	E-51	7	519	226	128

									Align	
Seque		Sequen						Bit-	ment	Posi
nce		се			Hit	E-	Simil	Scor	lengt	tive
name	Sequence desc.	length	Hit desc.	Order	ACC	Value	arity	e	h	S
			gi 1022742921 gb KZR							
			99685.1 Lysosomal							
Th_d_0	Lysosomal		alpha-glucosidase,							
01156	alpha-		partial [Daphnia	Crustace	KZR99	6.63E-	74.14	204.		
58-RA	glucosidase	441	magna]	а	685	63	966	142	147	109
			gi 1061468644 gb 0D							
Th_d_0	lysosomal		M88819.1 Maltase-							
00490	alpha-		glucoamylase, intestinal	Collemb	ODM8	4.51E-	64.89	77.4		
05-RA	glucosidase-like	1122	[Orchesella cincta]	ola	8819	13	362	11	94	61
			gi 1009533976 ref XP_							
			015903978.1 maltase-							
Th_d_0	lysosomal		glucoamylase, intestinal		XP_015					
00137	alpha-		[Parasteatoda	Arachnid	90397	8.34E-	78.48	114.		
79-RA	glucosidase-like	324	tepidariorum]	а	8	29	101	775	79	62
			gi 1228013272 ref XP_							
			021938600.1 glucosida							
Th_d_0			se 2 subunit beta		XP_021					
00895	glucosidase 2		[Zootermopsis	Blattode	93860	5.08E-	84.78	211.		
27-RA	subunit beta	420	nevadensis]	а	0	65	261	846	138	117

Seque		Sequen						Bit-	Alignm	
nce	Sequence	ce				Е-	Simil	Scor	ent	Positiv
name	desc.	length	Hit desc.	Order	Hit ACC	Value	arity	е	length	es
	AChain A,									
	Crystal									
	Structure Of									
	Beta-									
	glucosidase									
	From Termite									
	Neotermes		gi 506967929 gb							
	Koshunensis		AGM32308.1 beta-							
	In Complex		glucosidase,							
Th_d_0	With A New		partial							
00252	Glucopyranosi		[Coptotermes		AGM3230	1.63E-	73.33	77.4		
79-RA	dic Product	1536	formosanus]	Blattodea	8	13	333	11	60	44
			gi 636630780 gb							
			AIA09350.1 alpha							
	Uncharacteriz		-glucosidase							
Th_d_0	ed family 31		family 31, partial							
00485	glucosidase		[Periplaneta			6.9E-	69.02	523.		
38-RA	KIAA1161	2302	americana]	Blattodea	AIA09350	176	287	472	481	332
Seque		Sequen				E-		Bit-	Align	Posi
--------	--------------	--------	---------------------------	-----------	-------	-----	-------	------	--------	------
nce	Sequence	се			Hit	Val	Simil	Sco	ment	tive
name	desc.	length	Hit desc.	Order	ACC	ue	arity	re	length	S
			gi 1330905640 gb PNF2960							
			5.1 hypothetical protein							
			B7P43_G01635							
			[Cryptotermes							
			secundus]gi 1330905643 gb							
			PNF29608.1 hypothetical							
			protein B7P43_G01635		PNF2					
			[Cryptotermes		9605,					
	lysosomal		secundus]gi 1330905644 gb		PNF2					
Th_d_0	alpha-		PNF29609.1 hypothetical		9608,	9.9		187		
01597	glucosidase-		protein B7P43_G01635		PNF2	6E-	59.93	.19		
60-RA	like	903	[Cryptotermes secundus]	Blattodea	9609	54	266	3	297	178
			gi 1330905640 gb PNF2960							
			5.1 hypothetical protein							
			B7P43_G01635							
			[Cryptotermes							
			secundus]gi 1330905643 gb							
			PNF29608.1 hypothetical							
			protein B7P43_G01635		PNF2					
			[Cryptotermes		9605,					
	lysosomal		secundus]gi 1330905644 gb		PNF2					
Th_d_0	alpha-		PNF29609.1 hypothetical		9608,	1.0		184		
01276	glucosidase-		protein B7P43_G01635		PNF2	2E-	59.59	.88		
07-RA	like	903	[Cryptotermes secundus]	Blattodea	9609	52	596	2	297	177

Seque		Sequen						Bit-	Align ment	Posi
nce	Sequence	се		Orde			Simil	Scor	lengt	tive
name	desc.	length	Hit desc.	r	Hit ACC	E-Value	arity	e	h	S
			gi 1227970630 ref XP_0							
			21916719.1 glucoside							
			xylosyltransferase 2							
Th_d_0	glucoside		isoform X2							
00187	xylosyltransfe		[Zootermopsis	Blatt	XP_0219	1.5E-	87.21	460.6		
60-RA	rase 1	1251	nevadensis]	odea	16719	159	805	85	266	232
			gi 1228003695 ref XP_0							
			21933733.1 uncharacter							
			ized family 31							
			glucosidase KIAA1161-							
			like isoform X1							
			[Zootermopsis							
			nevadensis]gi 64670315							
	uncharacteriz		0 gb KDR11966.1 putati		XP_0219					
Th_d_0	ed family 31		ve family 31 glucosidase		33733,					
00765	glucosidase		[Zootermopsis	Blatt	KDR119	1.2E-	54.18	375.5		
37-RA	KIAA1161	1368	nevadensis]	odea	66	121	251	55	526	285

								Bit-	Align ment	
Sequenc	Sequenc	Sequenc			Hit		Similari	Scor	lengt	Posit
e name	e desc.	e length	Hit desc.	Order	ACC	E-Value	ty	е	h	ives
	beta-1,3-		gi 1330905524 gb PNF							
	glucan-		29497.1 hypothetical							
Th_d_00	binding		protein B7P43_G04546							
107538-	protein-		[Cryptotermes	Blatto	PNF29		67.5438	129.7		
RA	like	351	secundus]	dea	497	4.38E-35	6	98	114	77
	beta-1,3-		gi 913297561 ref XP_0							
	glucan-		13188774.1 PREDICTE							
Th_d_00	binding		D: beta-1,3-glucan-		XP_013					
117669-	protein-		binding protein-like	Lepid	18877		57.9310	117.8		
RA	like	450	[Amyelois transitella]	optera	4	3.33E-30	3	57	145	84
	beta-1,3-									
	glucan-									
	binding		gi 646713430 gb KDR1							
Th_d_00	protein		7776.1 Apolipophorin							
017253-	precurso		[Zootermopsis	Blatto	KDR17		63.8570	1267.		
RA	r	4398	nevadensis]	dea	776	0	5	68	1483	947
			gi 506968285 gb AGM							
			32486.1 gram negative							
	Beta-1,3-		bacteria binding							
Th_d_00	glucan-		protein 2, partial							
007520-	binding		[Coptotermes	Blatto	AGM32		70.3703	60.46		
RA	protein	258	formosanus]	dea	486	3.3E-10	7	22	54	38

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.

								Bit-	Align ment	
Sequenc	Sequenc	Sequenc		0	Hit	E Valaa	Similari	Scor	lengt	Posit
e name	e desc.	e length	Hit desc.	Urder	ALL	E-value	ty	e	n	ives
	beta-1,3-									
	glucan-									
Th_d_00	binding		gi 156535746 gb ABU8							
007517-	protein-		0005.1 GNBP-B1	Dipter	ABU80		88.5714	60.46		
RA	like	923	[Anopheles merus]	а	005	5.25E-08	3	22	35	31
	beta-1,3-		gi 913297561 ref XP_0							
	glucan-		13188774.1 PREDICTE							
Th_d_00	binding		D: beta-1,3-glucan-		XP_013					
103657-	protein-		binding protein-like	Lepid	18877		60.1503	115.1		
RA	like	419	[Amyelois transitella]	optera	4	2.46E-29	8	61	133	80
			gi 1227967090 ref XP_							
	Beta-1,3-		021914900.1 beta-1,3-							
Th_d_00	glucan-		glucan-binding protein-		XP_021					
059385-	binding		like [Zootermopsis	Blatto	91490		54.2288	159.8		
RA	protein	696	nevadensis]	dea	0	3.55E-45	6	44	201	109
	Beta-1,3-		gi 1022754729 gb KZS							
Th_d_00	glucan-		02483.1 Beta-1,3-							
012339-	binding		glucan-binding protein	Crusta	KZS02		72.8571	174.0		
RA	protein	792	[Daphnia magna]	cea	483	3.41E-50	4	96	140	102
			gi 646720035 gb KDR2							
	Beta-1,3-		1902.1 Beta-1,3-							
Th_d_00	glucan-		glucan-binding protein							
117568-	binding		1 [Zootermopsis	Blatto	KDR21			92.81		
RA	protein	285	nevadensis]	dea	902	4.49E-22	73.9726	89	73	54

Sequenc	Sequence	Seque	Hit desc.	Order	Hit	E- Value	Simil	Bit- Scor	Align	Po sit
ename	uest.	longth			ACC	value	arity	0	longt	iv
		length						C	h	es
Th d 00	endo-beta-		gil321460555lgblEFX71596.1len							0.5
048327-	1.4-		do-beta-1.4-mannanase [Daphnia	Crusta	EFX71	3.9E-		347.		22
RA	mannanase	1087	pulex]	cea	596	116	68	436	325	1
Th d 00	Endo-beta-		gi 1022765557 gb KZS10752.1 E							
037769-	1,4-		ndo-beta-1,4-mannanase	Crusta	KZS10	6.9E-	67.41	381.		24
RA	mannanase	1229	[Daphnia magna]	cea	752	129	573	719	356	0
Th_d_00	endo-beta-		gi 321460555 gb EFX71596.1 en							
117927-	1,4-		do-beta-1,4-mannanase [Daphnia	Crusta	EFX71	2.26E-	57.22	159.		
RA	mannanase	552	pulex]	cea	596	45	543	073	173	99
Th_d_00	endo-beta-		gi 321460555 gb EFX71596.1 en							
114892-	1,4-		do-beta-1,4-mannanase [Daphnia	Crusta	EFX71	7.1E-	57.80	157.		10
RA	mannanase	540	pulex]	cea	596	45	347	532	173	0
Th_d_00	Endo-beta-		gi 1022765557 gb KZS10752.1 E							
022705-	1,4-		ndo-beta-1,4-mannanase	Crusta	KZS10	3.03E-	68.42	164.		10
RA	mannanase	614	[Daphnia magna]	cea	752	47	105	851	152	4
Th_d_00	endo-beta-		gi 321460555 gb EFX71596.1 en							
022704-	1,4-		do-beta-1,4-mannanase [Daphnia	Crusta	EFX71	3.74E-		239.		14
RA	mannanase	838	pulex]	cea	596	75	71.5	58	200	3
	lysosomal									
Th_d_00	alpha-		gi 1101377119 ref XP_01890848		XP_018					
040234-	mannosida		6.1 PREDICTED: lysosomal alpha-	Hemip	90848	1.88E-		53.5		
RA	se	296	mannosidase [Bemisia tabaci]	tera	6	07	52.5	286	80	42

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

Sequenc	Sequence	Seque	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Pos
e name	desc.	nce				Value	arity	Scor	ment	itiv
		length						е	length	es
			gi 1192752750 ref XP_01							
	beta-		5929642.2 beta-							
Th_d_00	mannosida		mannosidase							
119102-	se isoform		[Parasteatoda	Arachnid	XP_015	1.01E-	84.73	209.5		
RA	X1	447	tepidariorum]	а	929642	61	282	34	131	111
Th_d_00	alpha-		gi 1022777739 gb KZS21							
080655-	mannosida		240.1 Alpha-mannosidase	Crustace	KZS212	1.15E-		94.74		
RA	se 2	552	2x [Daphnia magna]	а	40	20	90	49	50	45
			gi 195377741 ref XP_002							
			047646.1 uncharacterized							
			protein Dvir_GJ11812							
	lysosomal		[Drosophila							
	alpha-		virilis]gi 194154804 gb E		XP_002					
Th_d_00	mannosida		DW69988.1 uncharacteriz		047646,					
105670-	se isoform		ed protein Dvir_GJ11812		EDW69	8.58E-	97.05	76.64		
RA	X1	264	[Drosophila virilis]	Diptera	988	16	882	06	34	33
			gi 642940243 ref XP_008							
			199468.1 PREDICTED:							
			RNA-directed DNA							
	lysosomal		polymerase from mobile							
Th_d_00	alpha-		element jockey-like,							
080493-	mannosida		partial [Tribolium	Coleopte	XP_008	6.87E-	55.42	100.5		
RA	se-like	1059	castaneum]	ra	199468	22	169	23	166	92

Sequenc	Sequen	Sequen	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Po
e name	ce desc.	ce			ACC	Value	arity	Scor	ment	sit
		length						е	lengt	iv
									h	es
			gi 1227972165 ref XP_021917507.							
	ER		1 ER degradation-enhancing alpha-							
	degrada		mannosidase-like protein 3 isoform							
	tion-		X1 [Zootermopsis							
	enhanci		nevadensis]gi 1227972167 ref XP_		XP_021					
	ng		021917508.1 ER degradation-		91750					
	alpha-		enhancing alpha-mannosidase-like		7,					
	mannos		protein 3 isoform X1 [Zootermopsis		XP_021					
	idase-		nevadensis]gi 646718115 gb KDR2		91750					
Th_d_00	like		0717.1 ER degradation-enhancing		8,					
121568-	protein		alpha-mannosidase-like 3	Blatto	KDR20	1.82E-	90.90	189.		10
RA	3	402	[Zootermopsis nevadensis]	dea	717	54	909	119	110	0
			gi 1227973104 ref XP_021917998.		XP_021					
			1 beta-mannosidase [Zootermopsis		91799					
			nevadensis]gi 1227973106 ref XP_		8,					
			021917999.1 beta-mannosidase		XP_021					
			[Zootermopsis		91799					
			nevadensis]gi 1227973108 ref XP_		9,					
			021918000.1 beta-mannosidase		XP_021					
			[Zootermopsis		91800					
Th_d_00	beta-		nevadensis]gi 646717625 gb KDR2		0,					
003139-	mannos		0416.1 Beta-mannosidase	Blatto	KDR20	5E-	65.13	490.		32
RA	idase	1500	[Zootermopsis nevadensis]	dea	416	163	026	345	499	5

Sequenc	Sequen	Sequen	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Po
e name	ce desc.	ce			ACC	Value	arity	Scor	ment	sit
		length						е	lengt	iv
Th d 00	Rota-								- 11	es
003143-	mannos		gil1061485315lgbl0DN02454 1lBet	Colle		2 93F-	76 53	122		
RA	idase	345	a-mannosidase [Orchesella cincta]	mhola	454	31	061	122. 094	98	75
	heta-	515		mbola	151	51	001	071		75
	mannos									
Th d 00	idase		gi 766933223 ref XP 011498870.1	Hvme	XP 011					
003142-	isoform		PREDICTED: beta-mannosidase	nopte	49887	4.33E-		97.4		
RA	X1	273	[Ceratosolen solmsi marchali]	ra	0	23	75	413	68	51
	beta-									
	mannos									
Th_d_00	idase		gi 1192752750 ref XP_015929642.		XP_015					
003141-	isoform		2 beta-mannosidase [Parasteatoda	Arach	92964	1.58E-	72.54	66.2		
RA	X1	1077	tepidariorum]	nida	2	09	902	402	51	37
	beta-									
Th_d_00	mannos		gi 1192752750 ref XP_015929642.		XP_015					
003140-	idase,		2 beta-mannosidase [Parasteatoda	Arach	92964	3.06E-	88.05	114.		
RA	putative	387	tepidariorum]	nida	2	28	97	005	67	59
	Mannan									
	endo-									
	1,4-									
Th_d_00	beta-		gi 1061473607 gb 0DM91993.1 Ma	A 11	0.0.1/0	0.075		054		4 1
036630-	mannos	1011	nnan endo-1,4-beta-mannosidase	Colle	UDM9	3.97E-	57.05	251.	212	17
KA	idase	1341	[Urchesella cincta]	mbola	1993	77	128	136	312	8

Seque	Sequenc	Sequ	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Po
nce	e desc.	ence			ACC	Value	arity	Scor	ment	sit
name		lengt						е	lengt	iv
		h							h	es
	Mannosyl									
	-		gi 242010114 ref XP_002425821.1 Ma							
	oligosacc		nnosyl-oligosaccharide alpha-1,2-							
	haride		mannosidase isoform, putative							
	alpha-1,2-		[Pediculus humanus		XP_002					
Th_d_	mannosid		corporis]gi 212509754 gb EEB13083.1		42582					
00032	ase		Mannosyl-oligosaccharide alpha-1,2-		1,					
183-	isoform,		mannosidase isoform, putative	Phthir	EEB13	4.88E-	91.85	238.		12
RA	putative	843	[Pediculus humanus corporis]	aptera	083	73	185	81	135	4
			gi 1121144990 ref XP_019559624.1 P							
			REDICTED: mannosyl-oligosaccharide							
			alpha-1,2-mannosidase IA-like [Aedes							
			albopictus]gi 1121144992 ref XP_0195		XP_019					
			59625.1 PREDICTED: mannosyl-		55962					
			oligosaccharide alpha-1,2-mannosidase		4,					
			IA-like [Aedes		XP_019					
	mannosyl		albopictus]gi 1121144994 ref XP_0195		55962					
	-		59626.1 PREDICTED: mannosyl-		5,					
	oligosacc		oligosaccharide alpha-1,2-mannosidase		XP_019					
	haride		IA-like [Aedes		55962					
Th_d_	alpha-1,2-		albopictus]gi 1121144996 ref XP_0195		6,					
00032	mannosid		59627.1 PREDICTED: mannosyl-		XP_019					
182-	ase IA-		oligosaccharide alpha-1,2-mannosidase	Dipter	55962	7.38E-	90.41	137.		
RA	like	501	IA-like [Aedes albopictus]	а	7	36	096	117	73	66

Seque	Sequence	Sequ	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Po
nce name	desc.	lengt			ALL	value	arity	<u> 5сог</u>	ment lengt	sit iv
nume		h						C	h	es
Th_d_										
00095	alpha-		gi 1330899240 gb PNF26420.1 hyp							
658-	mannosidase		othetical protein B7P43_G16606,	Blatto	PNF26	1.23E-	69.23	201.		13
RA	2	600	partial [Cryptotermes secundus]	dea	420	59	077	83	195	5
			gi 1339064618 ref XP_023715597.							
			1 lysosomal alpha-mannosidase							
			isoform X1 [Cryptotermes		XP_023					
Th_d_	lysosomal		secundus]gi 1330931496 gb PNF42		71559					
00000	alpha-		230.1 Lysosomal alpha-		7,					
733-	mannosidase		mannosidase [Cryptotermes	Blatto	PNF42		79.41	823.		45
RA	isoform X1	3979	secundus]	dea	230	0	176	157	578	9
			gi 157111150 ref XP_001651410.1							
			lysosomal alpha-mannosidase		XP_001					
Th_d_	lysosomal		[Aedes		65141					
00063	alpha-		aegypti]gi 108878513 gb EAT4273	_	0,					
258-	mannosidase		8.1 AAEL005749-PA [Aedes	Dipter	EAT42	1.1E-	70.21	60.8		
RA	-like	615	aegyptij	а	738	08	277	474	47	33
	Endoplasmic									
	reticulum		gi 1339074955 ref XP_023724784.							
Th_d_	mannosyl-		1 endoplasmic reticulum mannosyl-							
00031	oligosacchari		oligosaccharide 1,2-alpha-		XP_023	1055	-0.01	1.60		
859-	de 1,2-alpha-		mannosidase isoform X2	Blatto	72478	1.95E-	78.86	163.		
RA	mannosidase	399	Cryptotermes secundus	dea	4	46	179	31	123	97

Seque	Sequence	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Po
nce	desc.	ence			ACC	Valu	larit	Sco	ment	sit
name		lengt				е	У	re	lengt	iv
		h							h	es
			gi 1218211818 ref XP_021712668.1 e							
			ndoplasmic reticulum mannosyl-							
			oligosaccharide 1,2-alpha-							
			mannosidase-like [Aedes		XP_02					
			aegypti]gi 1218211820 ref XP_02171		17126					
			2669.1 endoplasmic reticulum		68,					
	endoplasmic		mannosyl-oligosaccharide 1,2-alpha-		XP_02					
	reticulum		mannosidase-like [Aedes		17126					
Th_d_	mannosyl-		aegypti]gi 1218211822 ref XP_02171		69,					
00031	oligosacchari		2671.1 endoplasmic reticulum		XP_02		78.8	301		
860-	de 1,2-alpha-		mannosyl-oligosaccharide 1,2-alpha-		17126	1.7E-	990	.98		17
RA	mannosidase	777	mannosidase-like [Aedes aegypti]	Diptera	71	99	8	2	218	2
Th_d_	lysosomal		gi 929380069 ref XP_014100402.1 PR							
00002	alpha-		EDICTED: lysosomal alpha-		XP_01		80.7	289		
247-	mannosidase		mannosidase-like, partial [Bactrocera		41004	3.18E	291	.27		15
RA	isoform X1	591	oleae]	Diptera	02	-98	7	1	192	5
Th_d_	lysosomal		gi 557780736 ref XP_005189989.1 PR							
00002	alpha-		EDICTED: lysosomal alpha-		XP_00		70.5	196		
245-	mannosidase		mannosidase isoform X2 [Musca		51899	4.11E	882	.05		12
RA	isoform X2	528	domestica]	Diptera	89	-56	4	2	170	0
Th_d_	lysosomal									
00002	alpha-		gi 1228005950 ref XP_021934868.1 l		XP_02		70.7	151		
244-	mannosidase		ysosomal alpha-mannosidase isoform	Blattod	19348	7.1E-	142	.75		
RA	isoform X1	492	X1 [Zootermopsis nevadensis]	ea	68	41	9	4	140	99

Seque	Sequence	Sequ	Hit desc.	Ord	Hit	E-	Simi	Bit-	Align	Po
nce	desc.	ence		er	ACC	Valu	larit	Sco	ment	sit
name		lengt				е	У	re	lengt	iv
		h							h	es
			gi 1339064618 ref XP_023715597.1 lyso							
			somal alpha-mannosidase isoform X1		XP_02					
Th_d_			[Cryptotermes		37155					
00002	Lysosomal		secundus]gi 1330931496 gb PNF42230.		97,		60.3	48.		
242-	alpha-		1 Lysosomal alpha-mannosidase	Blatt	PNF4	4E-	448	906		
RA	mannosidase	249	[Cryptotermes secundus]	odea	2230	06	3	2	58	35
Th_d_										
00002	lysosomal		gi 907620158 ref XP_013097338.1 PRE		XP_01		47.0	191		
243-	alpha-		DICTED: lysosomal alpha-mannosidase	Dipt	30973	2.17E	997	.81		20
RA	mannosidase	1497	isoform X2 [Stomoxys calcitrans]	era	38	-50	7	5	431	3
			gi 1009600604 ref XP_015928616.1 alp		XP_01					
			ha-mannosidase 2 isoform X2		59286					
Th_d_			[Parasteatoda		16,					
00094	alpha-		tepidariorum]gi 1009600606 ref XP_015	Arac	XP_01		65.2	148		
171-	mannosidase		928617.1 alpha-mannosidase 2 isoform	hnid	59286	2.03E	694	.67		10
RA	2	558	X2 [Parasteatoda tepidariorum]	а	17	-39	6	3	167	9
			gi 242010114 ref XP_002425821.1 Man							
			nosyl-oligosaccharide alpha-1,2-							
	Mannosyl-		mannosidase isoform, putative [Pediculus							
	oligosacchari		humanus		XP_00					
Th_d_	de alpha-1,2-		corporis]gi 212509754 gb EEB13083.1		24258					
00012	mannosidase		Mannosyl-oligosaccharide alpha-1,2-	Phth	21,		92.4	146		
219-	isoform,		mannosidase isoform, putative [Pediculus	irapt	EEB1	4.37E	050	.74		
RA	putative	273	humanus corporis]	era	3083	-41	6	7	79	73

Table 3.4.	Continued.
------------	------------

Sequen	Sequence	Seque	Hit desc.	Ord	Hit	E-	Simi	Bit-	Align	Po
ce	desc.	nce		er	ACC	Valu	larit	Sco	ment	sit
name		length				е	у	re	lengt	iv
		_					-		h	es
	mannosyl-									
	oligosacchari									
Th_d_00	de alpha-1,2-			Crus			91.8	109		
012218	mannosidase		gi 1022764168 gb KZS09649.1 Alpha	tace	KZS09	4.13E	032	.76		
-RA	IA	261	-1,2-Mannosidase [Daphnia magna]	а	649	-28	8	8	61	56
	mannosyl-									
	oligosacchari		gi 1279729672 ref XP_022909489.1							
Th_d_00	de alpha-1,2-		mannosyl-oligosaccharide alpha-1,2-	Cole	XP_02		85.2	171		
012222	mannosidase		mannosidase IA-like [Onthophagus	opte	29094	2.69E	941	.01		
-RA	IA-like	492	taurus]	ra	89	-49	2	4	102	87
	mannosyl-									
	oligosacchari									
Th_d_00	de alpha-1,2-		gi 1233182429 ref XP_022192805.1	Hem	XP_02		82.7	135		
012221	mannosidase		mannosyl-oligosaccharide alpha-1,2-	ipter	21928	3.42E	586	.96		
-RA	IA-like	771	mannosidase IA [Nilaparvata lugens]	а	05	-34	2	1	87	72
			gi 1009600604 ref XP_015928616.1							
			alpha-mannosidase 2 isoform X2		XP_01					
			[Parasteatoda		59286					
			tepidariorum]gi 1009600606 ref XP_		16,					
Th_d_00	alpha-		015928617.1 alpha-mannosidase 2	Arac	XP_01			177		
121172	mannosidase		isoform X2 [Parasteatoda	hnid	59286	1.79E		.17		
-RA	2	354	tepidariorum]	а	17	-50	89	8	100	89

Sequen	Sequenc	Seq	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posi
се	e desc.	uen				Value	arity	Scor	ment	tive
name		се						е	lengt	S
		leng							h	
		th								
	lysosomal		gi 1233202594 ref XP_02							
Th_d_00	alpha-		2200324.1 lysosomal							
073548	mannosid		alpha-mannosidase-like	Hemipte	XP_02220032	2.81E-	60.75	172.		
-RA	ase-like	837	[Nilaparvata lugens]	ra	4	46	949	17	237	144
	lysosomal		gi 1351647487 ref XP_02							
Th_d_00	alpha-		4086113.1 lysosomal							
091843	mannosid		alpha-mannosidase-like	Hemipte	XP_02408611	2.38E-	64.04	89.3		
-RA	ase-like	663	[Cimex lectularius]	ra	3	18	494	521	89	57
			gi 1192752750 ref XP_01							
	Beta-		5929642.2 beta-							
Th_d_00	mannosid		mannosidase							
022617	ase-like		[Parasteatoda	Arachnid	XP_01592964	7.8E-	78.43	65.4		
-RA	Protein	264	tepidariorum]	а	2	12	137	698	51	40
	beta-									
	mannosid		gi 675378716 gb KFM716							
Th_d_00	ase		18.1 Beta-mannosidase,							
022616	isoform		partial [Stegodyphus	Arachnid		3.13E-	63.36	251.		
-RA	X2	990	mimosarum]	а	KFM71618	76	996	521	273	173

Sequen	Sequenc	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Po
се	e desc.	nce		er		Val	larit	Sco	ment	sit
name		length				ue	У	re	lengt	iv
									h	es
			gi 1227973104 ref XP_021917998.1 bet							
			a-mannosidase [Zootermopsis							
			nevadensis]gi 1227973106 ref XP_0219		VD 001					
			1/999.1 beta-mannosidase		XP_021					
			2001er mopsis novadangislail 227072109 rof VD 0210		91/996, VD 021					
			18000 1lbeta-mannosidase		AF_021 917999					
			[Zootermonsis		XP 021					
Th d 00	beta-		nevadensis]gil646717625lgblKDR20416.		918000.	2.9	65.8	79.		
116423	mannosid		1 Beta-mannosidase [Zootermopsis	Blatt	KDR204	2E-	823	722		
-RA	ase	366	nevadensis]	odea	16	16	5	1	85	56
	ER									
	degradati									
	on-									
	enhancin									
T I 1 0 0	g alpha-					~ -				
Th_d_00	mannosid		gi 1247029483 gb PCG66925.1 hypothet	Lepi	Daacco	3.7	92.2	321		10
087751	ase-like	570	ical protein B5V51_7083, partial	dopt	PCG669	E-	222	.62	100	16
-RA	protein 2	573	[Heliothis virescens]	era	25	111	2	7	180	6
Th_d_00	alpha-		gi 321444696 gb EFX60486.1 hypotheti	Crus		1.9	84.4	192		
046444	mannosid		cal protein DAPPUDRAFT_343157,	tace	EFX604	5E-	444	.97		11
-RA	ase 2	438	partial [Daphnia pulex]	а	86	59	4	1	135	4

Table 3.4	4. Conti	inued.
-----------	----------	--------

Sequen	Sequenc	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Po
ce	e desc.	nce		er		Val	larit	Sco	ment	sit
name		length				ue	У	re	lengt	iv
									h	es
	lysosomal									
Th_d_00	alpha-		gi 1098664167 ref XP_018794803.1 PRE			3.4	76.7	72.		
075487	mannosid		DICTED: lysosomal alpha-mannosidase	Dipt	XP_018	8E-	857	018		
-RA	ase	444	[Bactrocera latifrons]	era	794803	13	1	2	56	43
	lysosomal			Hym						
Th_d_00	alpha-		gi 970882260 ref XP_015121072.1 PRE	eno		2.8	54.3	184		
058303	mannosid		DICTED: lysosomal alpha-mannosidase	pter	XP_015	E-	918	.11		16
-RA	ase-like	1106	isoform X1 [Diachasma alloeum]	а	121072	49	9	1	296	1
Th_d_00	alpha-		gi 1330899240 gb PNF26420.1 hypothet			3.4	62.8	135		
062699	mannosid		ical protein B7P43_G16606, partial	Blatt	PNF264	6E-	378	.96		
-RA	ase 2	1077	[Cryptotermes secundus]	odea	20	33	4	1	148	93
			gi 1339064618 ref XP_023715597.1 lyso							
			somal alpha-mannosidase isoform X1							
	Lysosoma		[Cryptotermes		XP_023					
Th_d_00	l alpha-		secundus]gi 1330931496 gb PNF42230.		715597,	5.9	59.7	112		
023682	mannosid		1 Lysosomal alpha-mannosidase	Blatt	PNF422	E-	222	.84		
-RA	ase	1545	[Cryptotermes secundus]	odea	30	24	2	9	144	86
	Mannan									
	endo-1,4-									
Th_d_00	beta-			Crus		1.3	66.6	92.		
086716	mannosid		gi 683427464 gb AIN40245.1 mannanas	tace	AIN402	4E-	666	048		
-RA	ase	240	e, partial [Cherax quadricarinatus]	а	45	23	7	5	78	52

Sequen	Sequence	Sequ	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Po
ce	desc.	ence		er		Val	larit	Sco	ment	sit
name		lengt				ue	У	re	lengt	iv
		h							h	es
	lysosomal									
Th_d_00	alpha-		gi 662196129 ref XP_008471080.1 PR	Hem		2.8	82.6			
038326	mannosidase		EDICTED: lysosomal alpha-	ipter	XP_008	2E-	815	278		14
-RA	-like	735	mannosidase-like [Diaphorina citri]	а	471080	90	6	.1	179	8
	lysosomal									
Th_d_00	alpha-		gi 1101337044 ref XP_018914717.1 P	Hem		4.2	78.2	80.		
038325	mannosidase		REDICTED: lysosomal alpha-	ipter	XP_018	5E-	608	492		
-RA	-like	279	mannosidase-like [Bemisia tabaci]	а	914717	17	7	5	46	36
	mannosyl-									
	oligosacchari		gi 1227996327 ref XP_021929979.1 m							
Th_d_00	de alpha-1,2-		annosyl-oligosaccharide alpha-1,2-			2.9	88.8	168		
074895	mannosidase		mannosidase IA isoform X1	Blatt	XP_021	6E-	888	.31		12
-RA	IA-like	882	[Zootermopsis nevadensis]	odea	929979	45	9	8	144	8
Th_d_00	lysosomal		gi 193580067 ref XP_001946748.1 PR	Hem		6.5		68.		
097964	alpha-		EDICTED: lysosomal alpha-	ipter	XP_001	3E-	70.3	551		
-RA	mannosidase	267	mannosidase [Acyrthosiphon pisum]	а	946748	13	125	4	64	45
	lysosomal		gi 929380069 ref XP_014100402.1 PR							
Th_d_00	alpha-		EDICTED: lysosomal alpha-			6.4	60.7	173		
043225	mannosidase		mannosidase-like, partial [Bactrocera	Dipt	XP_014	3E-	954	.71		10
-RA	isoform X1	975	oleae]	era	100402	51	5	1	176	7
Th_d_00	alpha-		gi 1330899240 gb PNF26420.1 hypoth			7.0	68.9	161		
059073	mannosidase		etical protein B7P43_G16606, partial	Blatt	PNF264	8E-	655	.38		12
-RA	2	711	[Cryptotermes secundus]	odea	20	44	2	4	174	0

Sequen	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Po
ce	desc.	ence				Val	larit	Sco	ment	sit
name		lengt				ue	У	re	lengt	iv
		h							h	es
			gi 91091790 ref XP_970226.1 PR							
			EDICTED: mannosyl-							
			oligosaccharide 1,2-alpha-							
			mannosidase IA isoform X1							
			[Tribolium							
	mannosyl-		castaneum]gi 1004395369 gb KY							
	oligosaccharid		B24855.1 Mannosyl-		XP_970					
Th_d_00	e 1,2-alpha-		oligosaccharide alpha-1,2-		226,	9.3	55.8	360		
115822	mannosidase		mannosidase isoform A-like	Coleopt	KYB248	E-	882	.53		28
-RA	IA isoform X1	1500	Protein [Tribolium castaneum]	era	55	116	2	3	501	0
			gi 1041544405 ref XP_00847644							
Th_d_00			6.2 PREDICTED: LOW QUALITY			7.3	79.5	80.		
071728	beta-		PROTEIN: beta-mannosidase	Hemipt	XP_008	3E-	918	107		
-RA	mannosidase	486	[Diaphorina citri]	era	476446	16	4	3	49	39
Th_d_00			gi 1191620180 gb 0TF82415.1 B			3.3		100		
120508	beta-		eta-mannosidase-like protein,	Arachni	OTF824	E-	78.1	.13		
-RA	mannosidase	387	partial [Euroglyphus maynei]	da	15	24	25	8	64	50
			gi 1339064618 ref XP_02371559							
			7.1 lysosomal alpha-mannosidase							
			isoform X1 [Cryptotermes							
	lysosomal		secundus]gi 1330931496 gb PNF		XP_023					
Th_d_00	alpha-		42230.1 Lysosomal alpha-		715597,	9.9	64.0	530		
013923	mannosidase		mannosidase [Cryptotermes	Blattod	PNF422	E-	575	.79		40
-RA	isoform X1	3389	secundus]	ea	30	169	1	1	626	1

Sequen	Sequence	Sequ	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Po
ce	desc.	ence		er		Val	larit	Sco	ment	sit
name		lengt				ue	У	re	lengt	iv
		h							h	es
	ER									
	degradation-									
	enhancing		gi 1339066460 ref XP_023725333.1							
Th_d_00	alpha-		ER degradation-enhancing alpha-			2.7	96.7	305		
025482	mannosidase-		mannosidase-like protein 3	Blatt	XP_023	7E-	948	.44		15
-RA	like protein 3	873	[Cryptotermes secundus]	odea	725333	94	7	9	156	1
	ER									
	degradation-									
	enhancing		gi 1339066460 ref XP_023725333.1							
Th_d_00	alpha-		ER degradation-enhancing alpha-			1.6	71.0			
025483	mannosidase-		mannosidase-like protein 3	Blatt	XP_023	E-	144	191		
-RA	like protein 3	540	[Cryptotermes secundus]	odea	725333	54	9	.43	138	98
Th_d_00			gi 1279738204 ref XP_022913917.1	Cole		3.5		140		
086879	beta-		beta-mannosidase [Onthophagus	opte	XP_022	6E-	57.2	.96		12
-RA	mannosidase	654	taurus]	ra	913917	36	77	9	213	2
Th_d_00				Crus		5.4	70.7	95.		
087558	alpha-		gi 1022777739 gb KZS21240.1 Alpha	tace	KZS212	9E-	317	515		
-RA	mannosidase 2	525	-mannosidase 2x [Daphnia magna]	а	40	21	1	3	82	58
			gi 158293330 ref XP_557731.3 AGAP							
			008584-PA [Anopheles gambiae str.		XP_557					
Th_d_00	lysosomal		PEST]gi 157016653 gb EAL40242.3		731,	3.6	85.7			
012668	alpha-		AGAP008584-PA [Anopheles gambiae	Dipt	EAL402	7E-	142	65.		
-RA	mannosidase	228	str. PEST]	era	42	12	9	855	35	30

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit	Е-	Simi	Bit-	Alig	Po
e name	desc.	ence			ACC	Value	larit	Sco	nme	sit
		lengt					У	re	nt	iv
		h							leng	es
									th	
	lysosomal									
	alpha-									
Th_d_000	mannosida		gi 1233202594 ref XP_02220032		XP_022			102		
12667-	se isoform		4.1 lysosomal alpha-mannosidase-	Hemipte	20032	1.43E-	63.0	.06		
RA	X1	480	like [Nilaparvata lugens]	ra	4	23	137	4	146	92
	lysosomal									
Th_d_000	alpha-		gi 1096257085 gb APA33853.1 s				71.4	291		
12666-	mannosida		eminal fluid protein [Nilaparvata	Hemipte	APA33	3.71E-	828	.96		18
RA	se	1071	lugens]	ra	853	88	9	7	263	8
	lysosomal									
Th_d_000	alpha-		gi 1233202594 ref XP_02220032		XP_022		90.6	87.		
12665-	mannosida		4.1 lysosomal alpha-mannosidase-	Hemipte	20032	4.24E-	976	426		
RA	se-like	177	like [Nilaparvata lugens]	ra	4	20	7	1	43	39
	Lysosomal		gi 662196129 ref XP_008471080.							
Th_d_000	alpha-		1 PREDICTED: lysosomal alpha-		XP_008		58.9	192		
22076-	mannosida		mannosidase-like [Diaphorina	Hemipte	47108	4.87E-	519	.97		13
RA	se	756	citri]	ra	0	57	7	1	229	5
	Mannan									
	endo-1,4-									
Th_d_001	beta-		gi 683427464 gb AIN40245.1 ma				74.6	103		
01434-	mannosida		nnanase, partial [Cherax	Crustace	AIN40	2.09E-	835	.21		
RA	se	363	quadricarinatus]	а	245	27	4	9	79	59

Seque	Sequence	Sequ	Hit desc.	Orde	Hit	Е-	Simi	Bit-	Alig	Po
nce	desc.	ence		r	ACC	Val	larit	Sco	nme	sit
name		lengt				ue	У	re	nt	iv
		h							leng	es
									th	
			gi 1339057577 ref XP_023717292.1 ER							
			degradation-enhancing alpha-							
	ER		mannosidase-like protein 2							
	degradation-		[Cryptotermes		XP_02					
	enhancing		secundus]gi 1330893438 gb PNF23714.		37172					
Th_d_0	alpha-		1 ER degradation-enhancing alpha-		92,		92.0	686		
00385	mannosidase-		mannosidase-like protein 2	Blatt	PNF2		634	.02		34
41-RA	like protein 2	1428	[Cryptotermes secundus]	odea	3714	0	9	6	378	8
				Arthr						
				opod						
	lysosomal			a-						
Th_d_0	alpha-		gi 1238836368 ref XP_022237398.1 lyso	Cheli	XP_02	1.1	65.6	288		
00575	mannosidase-		somal alpha-mannosidase-like [Limulus	cerat	22373	3E-	891	.11		22
94-RA	like	1073	polyphemus]	а	98	87	5	5	341	4
Th_d_0			gi 998509058 ref XP_015514937.1 PRE	Hym	XP_01	3.2	87.9	101		
00754	beta-		DICTED: beta-mannosidase [Neodiprion	enopt	55149	4E-	310	.29		
66-RA	mannosidase	495	lecontei]	era	37	23	3	3	58	51
Th_d_0	lysosomal		gi 193580067 ref XP_001946748.1 PRE		XP_00	4.1		68.		
00963	alpha-		DICTED: lysosomal alpha-mannosidase	Hemi	19467	2E-	70.3	936		
36-RA	mannosidase	249	[Acyrthosiphon pisum]	ptera	48	13	125	6	64	45
Th_d_0				Colle		1.7	74.5	66.		
00471	Beta-		gi 1061485315 gb 0DN02454.1 Beta-	mbol	ODN0	9E-	454	625		
11-RA	mannosidase	417	mannosidase [Orchesella cincta]	а	2454	11	5	4	55	41

Seque	Sequence	Sequence	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Po
nce	desc.	length			ACC	Value	larit	Sco	nme	sit
name							У	re	nt	iv
									leng	es
									th	
			gi 1228363654 ref XP_02196							
			8444.1 beta-mannosidase-		XP_021					
			like [Folsomia		96844					
Th_d_0			candida]gi 1215278402 gb 0		4,		67.3	67.		
00471	Beta-		XA64711.1 Beta-mannosidase	Collemb	OXA64	1.86E-	076	010		
10-RA	mannosidase	258	[Folsomia candida]	ola	711	12	9	6	52	35
			gi 242016705 ref XP_002428							
			888.1 predicted protein		XP 002					
	lysosomal		[Pediculus humanus		42888					
Th d 0	alpha-		corporis]gi 212513656 gb EE		8,		74.6	176		
00532	mannosidase		B16150.1 predicted protein	Phthira	EEB16	5.24E-	376	.79		10
44-RA	isoform X2	831	[Pediculus humanus corporis]	ptera	150	48	8	2	138	3
			gil195377741 ref XP_002047							
			646.1 Juncharacterized							
			protein Dvir GI11812							
			[Drosophila		XP 002					
			virilislgi 194154804 gb EDW		04764					
Th d 0	lvsosomal		69988.1 Juncharacterized		6.			81.		
00788	alpha-		protein Dvir GI11812		EDW6	1.64E-		262		
92-RA	mannosidase	432	[Drosophila virilis]	Diptera	9988	16	75	9	60	45

Seque	Sequence	Sequ	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Po
nce	desc.	ence		r	ACC	Val	larit	Sco	nme	sit
name		lengt				ue	У	re	nt	iv
		h							leng	es
									th	
					XP_02					
			gi 1339052711 ref XP_023714691.1 alp		37146					
			ha-mannosidase 2 [Cryptotermes		91,					
			secundus]gi 1339052713 ref XP_023714		XP_02					
			692.1 alpha-mannosidase 2		37146					
			[Cryptotermes		92,					
			secundus]gi 1339052715 ref XP_023714		XP_02					
			693.1 alpha-mannosidase 2		37146					
			[Cryptotermes		93,					
Th_d_0	alpha-		secundus]gi 1339052717 ref XP_023714		XP_02	9.4	84.7	309		
00708	mannosidase		694.1 alpha-mannosidase 2	Blatt	37146	3E-	368	.68		16
56-RA	2	615	[Cryptotermes secundus]	odea	94	97	4	6	190	1
	lysosomal									
Th_d_0	alpha-		gi 662196129 ref XP_008471080.1 PRE		XP_00	3.4	85.8	216		
00872	mannosidase-		DICTED: lysosomal alpha-mannosidase-	Hemi	84710	8E-	267	.08		10
76-RA	like	528	like [Diaphorina citri]	ptera	80	67	7	3	127	9
Th_d_0			gi 1325354705 ref XP_023311948.1 bet		XP_02	4.9	73.0	72.		
00033	beta-		a-mannosidase-like [Anoplophora	Coleo	33119	6E-	158	018		
40-RA	mannosidase	435	glabripennis]	ptera	48	14	7	2	63	46
Th_d_0			gi 1325354705 ref XP_023311948.1 bet		XP_02	4.9	73.0	72.		
00033	beta-		a-mannosidase-like [Anoplophora	Coleo	33119	6E-	158	018		
41-RA	mannosidase	435	glabripennis]	ptera	48	14	7	2	63	46

Table 3	.4. Con	tinued.
---------	---------	---------

Seque	Sequence	Sequence	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Po
nce	desc.	length			ACC	Valu	arity	Scor	ment	sit
name						е		е	lengt	iv
									h	es
			gi 1227973104 ref XP_0219							
			17998.1 beta-mannosidase							
			[Zootermopsis		XP_021					
			nevadensis]gi 1227973106 r		91799					
			ef XP_021917999.1 beta-		8,					
			mannosidase [Zootermopsis		XP_021					
			nevadensis]gi 1227973108 r		91799					
			ef XP_021918000.1 beta-		9,					
			mannosidase [Zootermopsis		XP_021					
			nevadensis]gi 646717625 gb		91800					
Th_d_0			KDR20416.1 Beta-		0,					
00697	beta-		mannosidase [Zootermopsis	Blatto	KDR20	9.92	76.52	142.1		
30-RA	mannosidase	648	nevadensis]	dea	416	E-37	174	24	115	88
			gi 642937291 ref XP_00819							
	mannosyl-		8773.1 PREDICTED:							
	oligosaccharid		mannosyl-oligosaccharide							
Th_d_0	e alpha-1,2-		alpha-1,2-mannosidase IA		XP_008					
00278	mannosidase		isoform X2 [Tribolium	Coleo	19877	5.22	88.20	270.3		15
80-RA	IA-like	537	castaneum]	ptera	3	E-87	225	96	178	7

Seque	Sequence	Sequenc	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Alig	Po
nce	desc.	e length				Val	larit	Sco	nme	sit
name						ue	У	re	nt	iv
									leng	es
									th	
			gi 1227998244 ref XP_0219309							
			61.1 alpha-mannosidase 2							
			[Zootermopsis							
			nevadensis]gi 1227998246 ref							
			XP_021930962.1 alpha-							
			mannosidase 2 [Zootermopsis							
			nevadensis]gi 1227998248 ref							
			XP_021930963.1 alpha-		XP_02193					
			mannosidase 2 [Zootermopsis		0961,					
			nevadensis]gi 1227998250 ref		XP_02193					
			XP_021930964.1 alpha-		0962,					
			mannosidase 2 [Zootermopsis		XP_02193					
			nevadensis]gi 1227998252 ref		0963,					
			XP_021930965.1 alpha-		XP_02193					
			mannosidase 2 [Zootermopsis		0964,					
			nevadensis]gi 1227998254 ref		XP_02193					
			XP_021930966.1 alpha-		0965,					
			mannosidase 2 [Zootermopsis		XP_02193					
Th_d_0	alpha-		nevadensis]gi 646706309 gb K		0966,	1.3	56.1	121		
00836	mannosidase		DR13614.1 Alpha-mannosidase	Blatto	KDR1361	1E-	497	.70		10
99-RA	2	629	2 [Zootermopsis nevadensis]	dea	4	29	3	9	187	5

Seque	Sequence	Sequ	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Po
nce	desc.	ence		r	ACC	Val	larit	Sco	nme	sit
name		lengt				ue	У	re	nt	iv
		h							leng	es
									th	
			gi 242009114 ref XP_002425337.1 beta-							
			mannosidase precursor, putative		XP_00					
			[Pediculus humanus		24253					
Th_d_0			corporis]gi 212509122 gb EEB12599.1 b	Phthi	37,	1.6	77.5	281		
00431	beta-		eta-mannosidase precursor, putative	rapte	EEB1	E-	119	.56		16
79-RA	mannosidase	1224	[Pediculus humanus corporis]	ra	2599	84	6	7	209	2
Th_d_0	alpha-		gi 1330899240 gb PNF26420.1 hypothet			5.1	68.9			
00599	mannosidase		ical protein B7P43_G16606, partial	Blatt	PNF2	4E-	655	161		12
13-RA	2	696	[Cryptotermes secundus]	odea	6420	44	2	.77	174	0
				Arthr						
				opod						
Th_d_0			gi 1316169311 ref XP_023230069.1 bet	a-	XP_02	2.8	52.8	133		
00487	beta-		a-mannosidase-like isoform X1	Scorp	32300	3E-	138	.26		12
97-RA	mannosidase	1293	[Centruroides sculpturatus]	iones	69	31	5	5	231	2
			gi 1339057577 ref XP_023717292.1 ER							
			degradation-enhancing alpha-							
	ER		mannosidase-like protein 2							
	degradation-		[Cryptotermes		XP_02					
	enhancing		secundus]gi 1330893438 gb PNF23714.		37172					
Th_d_0	alpha-		1 ER degradation-enhancing alpha-		92,	7.7	84.0	350		
00948	mannosidase-		mannosidase-like protein 2	Blatt	PNF2	E-	375	.13		17
73-RA	like protein 2	708	[Cryptotermes secundus]	odea	3714	114	6	2	213	9

Seque	Sequence	Sequ	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Po
nce	desc.	ence		r	ACC	Val	larit	Sco	nme	sit
name		lengt				ue	У	re	nt	iv
		h							leng	es
									th	
	ER									
	degradation-									
	enhancing		gi 1227987439 ref XP_021925413.1 ER							
Th_d_0	alpha-		degradation-enhancing alpha-		XP_02	5.2	92.7	166		
01132	mannosidase-		mannosidase-like protein 2, partial	Blatt	19254	8E-	710	.77		
70-RA	like protein 2	1272	[Zootermopsis nevadensis]	odea	13	43	8	7	83	77
			gi 1227973104 ref XP_021917998.1 bet		XP_02					
			a-mannosidase [Zootermopsis		19179					
			nevadensis]gi 1227973106 ref XP_0219		98,					
			17999.1 beta-mannosidase		XP_02					
			[Zootermopsis		19179					
			nevadensis]gi 1227973108 ref XP_0219		99,					
			18000.1 beta-mannosidase		XP_02					
			[Zootermopsis		19180					
Th_d_0			nevadensis]gi 646717625 gb KDR20416.		00,	1.2	48.0	90.		
00045	beta-		1 Beta-mannosidase [Zootermopsis	Blatt	KDR2	2E-	769	892		
56-RA	mannosidase	474	nevadensis]	odea	0416	19	2	9	156	75
			gi 242016705 ref XP_002428888.1 predi		XP_00					
	lysosomal		cted protein [Pediculus humanus		24288					
Th_d_0	alpha-		corporis]gi 212513656 gb EEB16150.1	Phthi	88,	7.1	73.2	178		
00582	mannosidase		predicted protein [Pediculus humanus	rapte	EEB1	4E-	876	.33		10
65-RA	isoform X2	782	corporis]	ra	6150	49	7	3	146	7

Seque	Sequence	Sequ	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Po
nce	desc.	ence		r	ACC	Val	larit	Sco	nme	sit
name		lengt				ue	У	re	nt	iv
		h							leng	es
									th	
Th_d_0	Lysosomal		gi 1080066641 ref XP_018575534.1 lyso		XP_01	1.3	53.2	87.		
00842	alpha-		somal alpha-mannosidase isoform X2	Coleo	85755	1E-	374	426		
90-RA	mannosidase	434	[Anoplophora glabripennis]	ptera	34	18	1	1	139	74
			gi 1227972165 ref XP_021917507.1 ER							
			degradation-enhancing alpha-							
			mannosidase-like protein 3 isoform X1							
			[Zootermopsis							
			nevadensis]gi 1227972167 ref XP_0219		XP_02					
			17508.1 ER degradation-enhancing		19175					
	ER		alpha-mannosidase-like protein 3		07,					
	degradation-		isoform X1 [Zootermopsis		XP_02					
	enhancing		nevadensis]gi 646718115 gb KDR20717.		19175					
Th_d_0	alpha-		1 ER degradation-enhancing alpha-		08,	2.5	54.1	215		
00384	mannosidase-		mannosidase-like 3 [Zootermopsis	Blatt	KDR2	8E-	899	.31		19
57-RA	like protein 3	1560	nevadensis]	odea	0717	58	4	2	358	4
Th_d_0	lysosomal		gi 1000733841 ref XP_015589368.1 PRE	Hym	XP_01	6.0	70.9			
00089	alpha-		DICTED: lysosomal alpha-mannosidase	enopt	55893	9E-	497	191		12
43-RA	mannosidase	1628	isoform X1 [Cephus cinctus]	era	68	50	2	.43	179	7
	lysosomal									
Th_d_0	alpha-		gi 1035439623 gb ANJ04662.1 lysosoma			1.8	72.3	129		
00089	mannosidase-		l alpha-mannosidase-like protein	Hemi	ANJ04	9E-	214	.79		
44-RA	like	432	[Nilaparvata lugens]	ptera	662	33	3	8	112	81

Seque	Sequence desc.	Sequenc	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Po
nce		e length		r	ACC	Val	larit	Sco	nme	sit
name						ue	У	re	nt	iv
									leng	es
									th	
			gi 668446735 gb KFB36711.1 ly							
Th_d_0			sosomal alpha-mannosidase			1.2	54.8	76.		
00491	lysosomal alpha-		(mannosidase alpha class 2b	Dipte	KFB3	9E-	076	255		
53-RA	mannosidase-like	711	member 1) [Anopheles sinensis]	ra	6711	13	9	4	104	57
			gi 1227982315 ref XP_02192278							
			1.1 ER degradation-enhancing							
			alpha-mannosidase-like protein 1							
			[Zootermopsis		XP_02					
	ER degradation-		nevadensis]gi 646713767 gb KD		19227					
Th_d_0	enhancing alpha-		R17988.1 ER degradation-		81,	6.2	92.0	299		
00122	mannosidase-like		enhancing alpha-mannosidase-	Blatt	KDR1	9E-	731	.67		15
42-RA	protein 1	534	like 1 [Zootermopsis nevadensis]	odea	7988	98	7	1	164	1
			gi 1339070655 ref XP_02372250							
	ER degradation-		4.1 ER degradation-enhancing							
Th_d_0	enhancing alpha-		alpha-mannosidase-like protein		XP_02	1.4	92.5	206		
00122	mannosidase-like		1, partial [Cryptotermes	Blatt	37225	2E-	619	.06		11
43-RA	protein 1	384	secundus	odea	04	64	8	8	121	2
	ER degradation-		gi 1330882322 gb PNF18187.1							
Th_d_0	enhancing alpha-		ER degradation-enhancing alpha-			1.1	90.9	110		
00122	mannosidase-like		mannosidase-like protein 1,	Blatt	PNF1	E-	090	.53		
44-RA	protein 1	870	partial [Cryptotermes secundus]	odea	8187	25	9	8	66	60

Seque	Sequence desc.	Sequ	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Po
nce		ence		r	ACC	Val	larit	Sco	nme	sit
name		lengt				ue	У	re	nt	iv
		h							leng	es
									th	
			gi 1227973104 ref XP_021917998.1		XP_02					
			beta-mannosidase [Zootermopsis		19179					
			nevadensis]gi 1227973106 ref XP_02		98,					
			1917999.1 beta-mannosidase		XP_02					
			[Zootermopsis		19179					
			nevadensis]gi 1227973108 ref XP_02		99,					
			1918000.1 beta-mannosidase		XP_02					
			[Zootermopsis		19180					
Th_d_0			nevadensis]gi 646717625 gb KDR20		00,	4.8	67.6	176		
00153	beta-		416.1 Beta-mannosidase	Blatt	KDR2	8E-	829	.02		11
92-RA	mannosidase	753	[Zootermopsis nevadensis]	odea	0416	48	3	2	164	1
			gi 242009114 ref XP_002425337.1 b							
			eta-mannosidase precursor, putative							
			[Pediculus humanus		XP_00					
			corporis]gi 212509122 gb EEB12599		24253					
Th_d_0			.1 beta-mannosidase precursor,	Phthi	37,	3.5	78.3			
00153	beta-		putative [Pediculus humanus	rapte	EEB1	5E-	018	133		
91-RA	mannosidase	564	corporis]	ra	2599	34	9	.65	106	83
Th_d_0			gi 321463091 gb EFX74109.1 hypoth			1.3	81.7	251		
00422	lysosomal alpha-		etical protein DAPPUDRAFT_324591	Crust	EFX74	5E-	610	.13		13
95-RA	mannosidase-like	1251	[Daphnia pulex]	acea	109	72	1	6	159	0

Seque	Sequence desc.	Sequ	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Po
nce		ence		r	ACC	Val	larit	Sco	nme	sit
name		lengt				ue	У	re	nt	iv
		h							leng	es
									th	
					XP_02					
			gi 1339078211 ref XP_023726521.1		37265					
			alpha-mannosidase 2 [Cryptotermes		21,					
			secundus]gi 1339078213 ref XP_023		XP_02					
			726522.1 alpha-mannosidase 2		37265					
			[Cryptotermes		22,					
Th_d_0			secundus]gi 1339078215 ref XP_023		XP_02		76.7	133		
00714	alpha-		726523.1 alpha-mannosidase 2	Blatt	37265		857	3.5	100	77
33-RA	mannosidase 2	2781	[Cryptotermes secundus]	odea	23	0	1	5	8	4
			gi 929380069 ref XP_014100402.1 P							
Th_d_0	lysosomal alpha-		REDICTED: lysosomal alpha-		XP_01	5.5	79.2	194		
00917	mannosidase		mannosidase-like, partial [Bactrocera	Dipte	41004	7E-	307	.12		10
92-RA	isoform X1	531	oleae]	ra	02	61	7	6	130	3
			gi 241567773 ref XP_002402365.1 ly	Arthr						
			sosomal alpha-mannosidase, putative	opod	XP_00					
			[Ixodes	a-	24023					
Th_d_0	alpha-		scapularis]gi 215501997 gb EEC114	Cheli	65,	9.3		76.		
00892	mannosidase 2x-		91.1 lysosomal alpha-mannosidase,	cerat	EEC11	7E-	71.8	255		
30-RA	like	372	putative, partial [Ixodes scapularis]	а	491	17	75	4	64	46
Th_d_0			gi 662196129 ref XP_008471080.1 P		XP_00	8.1	70.9	158		
00429	lysosomal alpha-		REDICTED: lysosomal alpha-	Hemi	84710	6E-	923	.68		
01-RA	mannosidase-like	573	mannosidase-like [Diaphorina citri]	ptera	80	45	7	8	131	93

Sequenc e name	Sequenc e desc.	Sequen ce length	Hit desc.	Order	Hit ACC	E- Value	Simi larit y	Bit- Scor e	Align ment lengt	Posi tive s
									h	
Th_d_00 016998- RA	beta- glucuroni dase isoform X1	306	gi 1067103892 ref XP_01801 3304.1 PREDICTED: beta- glucuronidase-like [Hyalella azteca]	Crustace a	XP_018 01330 4	1.56E- 21	82.8 125	93.5 893	64	53
Th_d_00 066936- RA	beta- glucuroni dase-like isoform X1	855	gi 1316153756 ref XP_02322 1887.1 beta-glucuronidase- like [Centruroides sculpturatus]	Arthropo da- Scorpion es	XP_023 22188 7	1.44E- 75	61.5 094 3	244. 588	265	163
Th_d_00 011907- RA	beta- glucuroni dase	465	gi 909578964 ref XP_013145 358.1 PREDICTED: beta- glucuronidase [Papilio polytes]	Lepidopt era	XP_013 14535 8	3.2E- 57	73.8 255	192. 971	149	110
Th_d_00 011908- RA	beta- glucuroni dase isoform X3	539	gi 1108476517 emb CRK952 86.1 CLUMA_CG008644, isoform B [Clunio marinus]	Diptera	CRK95 286	6.19E- 12	62.6 666 7	69.3 218	75	47
Th_d_00 057287- RA	beta- glucuroni dase-like isoform X1	634	gi 1330889853 gb PNF22051 .1 hypothetical protein B7P43_G09739 [Cryptotermes secundus]	Blattode	PNF22 051	9.67E- 83	78.1 609 2	253. 832	174	136

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

Sequenc	Sequenc	Sequence	Hit desc.	Order	Hit	Е-	Simi	Bit-	Align	Posi
e name	e desc.	length			ACC	Value	larit	Scor	ment	tive
							У	e	lengt	S
									h	
			gi 194756422 ref XP_00							
			1960477.1 uncharacteri							
			zed protein							
			Dana_GF11493							
	_		[Drosophila							
	beta-		ananassae]gi 19062177		XP_001					
	glucuroni		5 gb EDV37299.1 uncha		96047					
Th_d_00	dase		racterized protein		7,		52.3			
033402-	isoform		Dana_GF11493		EDV37	2.44E-	333	175.		
RA	X3	1014	[Drosophila ananassae]	Diptera	299	47	3	637	300	157
			gi 1022772830 gb KZS1							
Th_d_00	beta-		6987.1 Beta-				75.5			
068821-	glucuroni		glucuronidase [Daphnia	Crustace	KZS16	4.99E-	395	178.		
RA	dase-like	521	magna]	а	987	56	7	718	139	105
			gi 1316153756 ref XP_0							
			23221887.1 beta-	Arthropo						
Th_d_00	beta-		glucuronidase-like	da-	XP_023		79.0			
098822-	glucuroni		[Centruroides	Scorpion	22188	5.6E-	849	209.		
RA	dase-like	528	sculpturatus]	es	7	64	7	92	153	121

Sequ	Seque	Sequ	Hit desc.	Ord	Hit ACC	E-	Si	Bit-	Align	Po
ence	nce	ence		er		Valu	mi	Sco	ment	sit
nam	desc.	lengt				е	lar	re	lengt	iv
e		h					ity		h	es
			gi 1227991600 ref XP_021927576.1 beta-							
			glucuronidase-like [Zootermopsis		XP_02192					
			nevadensis]gi 1227991602 ref XP_021927578		7576,					
			.1 beta-glucuronidase-like [Zootermopsis		XP_02192					
			nevadensis]gi 1227991604 ref XP_021927579		7578,					
			.1 beta-glucuronidase-like [Zootermopsis		XP_02192					
			nevadensis]gi 1227991606 ref XP_021927580		7579,					
			.1 beta-glucuronidase-like [Zootermopsis		XP_02192					
			nevadensis]gi 1227991608 ref XP_021927581		7580,					
			.1 beta-glucuronidase-like [Zootermopsis		XP_02192					
			nevadensis]gi 1227991610 ref XP_021927582		7581,					
			.1 beta-glucuronidase-like [Zootermopsis		XP_02192					
	beta-		nevadensis]gi 1227991612 ref XP_021927583		7582,					
	glucur		.1 beta-glucuronidase-like [Zootermopsis		XP_02192					
Th_d_	onidas		nevadensis]gi 1227991614 ref XP_021927584		7583,		57.			
0002	e		.1 beta-glucuronidase-like [Zootermopsis		XP_02192		66	303		
0056	isofor		nevadensis]gi 646709721 gb KDR15442.1 Bet	Blatt	7584,	7.26	01	.52		20
-RA	m X1	1260	a-glucuronidase [Zootermopsis nevadensis]	odea	KDR15442	E-95	7	3	359	7
	beta-									
	glucur									
Th_d_	onidas						82.			
0010	е			Hem			97			
5741	isofor		gi 939251322 ref XP_014245262.1 beta-	ipter	XP_01424	4.79	87	77.		
-RA	m X3	273	glucuronidase isoform X5 [Cimex lectularius]	a	5262	E-16	2	411	47	39

Sequ	Sequence	Seque	Hit desc.	Ord	Hit ACC	E-	Si	Bit-	Align	Po
ence	desc.	nce		er		Valu	mi	Sco	ment	sit
nam		lengt				е	lar	re	lengt	iv
e		h					ity		h	es
			gi 1339060579 ref XP_023718877.1 beta-							
			glucuronidase-like isoform X3							
Th_d_			[Cryptotermes				52.			
0003	Beta-		secundus]gi 1330889852 gb PNF22050.1		XP_02371		94	175		
1354	glucuronid		hypothetical protein B7P43_G09739	Blatt	8877,	2.29	11	.25		11
-RA	ase	684	[Cryptotermes secundus]	odea	PNF22050	E-49	8	2	221	7
				Arth						
				ropo						
Th_d_	beta-			da-			72.			
0003	glucuronid		gi 1316153756 ref XP_023221887.1 beta-	Scor			97	98.		
1355	ase		glucuronidase-like [Centruroides	pion	XP_02322	8.16	29	982		
-RA	isoform X1	441	sculpturatus]	es	1887	E-23	7	1	74	54
Th_d_	beta-						58.			
0001	glucuronid		gi 1121168380 ref XP_019529786.1 PRE				00	207		
7720	ase		DICTED: beta-glucuronidase-like isoform	Dipt	XP_01952	1.26	71	.60		16
-RA	isoform X3	882	X1 [Aedes albopictus]	era	9786	E-59	2	8	281	3
Th_d_							74.			
0001	beta-		gi 325303090 tpg DAA34282.1 TPA_inf:	Arac			60	159		
7721	glucuronid		beta-glucuronidase GUSB [Amblyomma	hnid		9.22	31	.07		
-RA	ase-like	462	variegatum]	а	DAA34282	E-48	7	3	126	94
Th_d_	beta-						69.			
0002	glucuronid		gi 557767812 ref XP_005183593.1 PRED				23	257		
9861	ase		ICTED: beta-glucuronidase isoform X2	Dipt	XP_00518	5.35	07	.29		17
-RA	isoform X3	1027	[Musca domestica]	era	3593	E-78	7	9	247	1

Sequen	Sequence	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Po
ce	desc.	nce		er		Valu	larit	Sco	ment	sit
name		length				е	У	re	lengt	iv
									h	es
			gi 1229725314 ref XP_022125377.1							
			beta-glucuronidase isoform X3							
			[Pieris		XP_0221					
Th_d_00	beta-		rapae]gi 1229725316 ref XP_02212	Lepi	25377,		66.1	76.		
094572	glucuronida		5378.1 beta-glucuronidase isoform	dopt	XP_0221	1.59	764	640		
-RA	se	303	X3 [Pieris rapae]	era	25378	E-15	7	6	68	45
			gi 1339060579 ref XP_023718877.1							
			beta-glucuronidase-like isoform X3							
			[Cryptotermes							
			secundus]gi 1330889852 gb PNF22		XP_0237					
Th_d_00	Beta-		050.1 hypothetical protein		18877,		66.8	197		
060577	glucuronida		B7P43_G09739 [Cryptotermes	Blatt	PNF220	2.87	571	.97		11
-RA	se	819	secundus]	odea	50	E-57	4	8	175	7
	beta-									
Th_d_00	glucuronida		gi 939667591 ref XP_014280878.1	Hem			78.0	108		
103092	se isoform		PREDICTED: beta-glucuronidase-like	ipter	XP_0142	1.76	821	.61		
-RA	X3	356	isoform X2 [Halyomorpha halys]	a	80878	E-26	9	2	73	57
				Arth						
				ropo						
	beta-			da-						
Th_d_00	glucuronida		gi 1316153756 ref XP_023221887.1	Scor			81.1	167		
085133	se isoform		beta-glucuronidase-like	pion	XP_0232	5.38	965	.93		
-RA	X2	516	[Centruroides sculpturatus]	es	21887	E-48	8	3	117	95
Sequen	Sequence	Sequ	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Po
---------	---------------	-------	-------------------------------------	-------	---------	------	-------	------	-------	-----
ce	desc.	ence		er		Valu	larit	Sco	ment	sit
name		lengt				е	У	re	lengt	iv
		h							h	es
Th_d_00	beta-		gi 646696603 gb KDR08779.1 Beta-				78.3	434		
018098	glucuronidase		glucuronidase [Zootermopsis	Blatt	KDR087	2.5E	699	.49		25
-RA	isoform X3	993	nevadensis]	odea	79	-147	1	1	319	0
				Hym						
Th_d_00	beta-		gi 951547932 ref XP_014474109.1	eno			57.6	73.		
119947	glucuronidase		PREDICTED: beta-glucuronidase	pter	XP_0144	1.61	923	944		
-RA	isoform X2	318	isoform X1 [Dinoponera quadriceps]	а	74109	E-14	1	2	78	45
			gi 1339060579 ref XP_023718877.1							
			beta-glucuronidase-like isoform X3							
			[Cryptotermes							
	beta-		secundus]gi 1330889852 gb PNF22		XP_0237					
Th_d_00	glucuronidase		050.1 hypothetical protein		18877,			559		
005617	-like isoform		B7P43_G09739 [Cryptotermes	Blatt	PNF220		62.9	.68		35
-RA	X1	1886	secundus]	odea	50	0	174	1	569	8
			gi 241309940 ref XP_002407819.1							
			beta-glucuronidase (GusB), putative							
	beta-		[Ixodes		XP_0024					
Th_d_00	glucuronidase		scapularis]gi 215497226 gb EEC067	Arac	07819,		74.6	155		
005619	-like isoform		20.1 beta-glucuronidase (GusB),	hnid	EEC0672	2.9E	031	.60		
-RA	X1	597	putative [Ixodes scapularis]	а	0	-45	7	6	126	94
Th_d_00	beta-			Crus			74.6	189		
056689	glucuronidase		gi 1022772830 gb KZS16987.1 Beta	tace	KZS1698	1.62	376	.11		10
-RA	isoform X3	866	-glucuronidase [Daphnia magna]	а	7	E-58	8	9	138	3

Sequen	Sequence	Sequ	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Po
ce	desc.	ence		er		Valu	larit	Sco	ment	sit
name		lengt				е	У	re	lengt	iv
		h							h	es
			gi 1339045451 ref XP_023710854.1 bet							
			a-glucuronidase-like [Cryptotermes							
			secundus]gi 1339045453 ref XP_02371		XP_0237					
			0855.1 beta-glucuronidase-like		10854,					
			[Cryptotermes		XP_0237					
Th_d_00	Beta-		secundus]gi 1330907803 gb PNF30365.		10855,		62.6	71.		
098010	glucuronid		1 Beta-glucuronidase [Cryptotermes	Blatt	PNF303	4.82	666	247		
-RA	ase	432	secundus]	odea	65	E-13	7	8	75	47
Th_d_00	Beta-		gi 662210289 ref XP_008478777.1 PRE	Hem			78.4	115		
094268	glucuronid		DICTED: beta-glucuronidase-like	ipter	XP_0084	4.17	090	.54		
-RA	ase	309	[Diaphorina citri]	а	78777	E-31	9	6	88	69
	beta-									
Th_d_00	glucuronid		gi 1330889853 gb PNF22051.1 hypothe				76.5	339		
049078	ase-like		tical protein B7P43_G09739	Blatt	PNF220	4.5E	567	.73		20
-RA	isoform X3	916	[Cryptotermes secundus]	odea	51	-115	8	2	273	9
			gi 1339045451 ref XP_023710854.1 bet							
			a-glucuronidase-like [Cryptotermes							
			secundus]gi 1339045453 ref XP_02371		XP_0237					
			0855.1 beta-glucuronidase-like		10854,					
			[Cryptotermes		XP_0237					
Th_d_00	Beta-		secundus]gi 1330907803 gb PNF30365.		10855,		76.7	130		
077520	glucuronid		1 Beta-glucuronidase [Cryptotermes	Blatt	PNF303	2.48	857	.95		
-RA	ase	801	secundus]	odea	65	E-32	1	4	112	86

Table 3.5. Continued.

Sequen	Sequence	Sequence	Hit desc.	Orde	Hit ACC	E-	Simil	Bit-	Align	Po
се	desc.	length		r		Valu	arity	Scor	ment	sit
name						е		е	lengt	iv
									h	es
	beta-									
Th_d_00	glucuronida		gi 1022772830 gb KZS16987.							
068165	se-like		1 Beta-glucuronidase [Daphnia	Crust	KZS169	7.51	74.10	176.		10
-RA	isoform X2	618	magna]	acea	87	E-55	072	792	139	3
			gi 1339060579 ref XP_023718							
			877.1 beta-glucuronidase-like							
			isoform X3 [Cryptotermes							
	beta-		secundus]gi 1330889852 gb P		XP_023					
Th_d_00	glucuronida		NF22050.1 hypothetical		718877,					
063635	se-like		protein B7P43_G09739	Blatt	PNF220	1.63	67.68	158.		11
-RA	isoform X3	819	[Cryptotermes secundus]	odea	50	E-42	293	688	164	1
			gi 985391249 ref XP_0153803							
Th_d_00	beta-		21.1 PREDICTED: beta-							
075299	glucuronida		glucuronidase [Diuraphis	Crust	XP_015	3.58	69.93	175.		10
-RA	se	498	noxia]	acea	380321	E-51	464	252	153	7
			gi 91089481 ref XP_969353.1							
			PREDICTED: beta-							
			glucuronidase [Tribolium							
	beta-		castaneum]gi 270011394 gb E		XP_969					
Th_d_00	glucuronida		FA07842.1 Beta-		353,					
070918	se isoform		glucuronidase-like Protein	Coleo	EFA078	1.4E	59.25	71.6		
-RA	X2	591	[Tribolium castaneum]	ptera	42	-12	926	33	81	48

Table 3.5. (Continued.
--------------	------------

Sequen	Sequenc	Seque	Hit desc.	Order	Hit	E-Value	Simi	Bit-	Align	Po
ce	e desc.	nce			ACC		larit	Sco	ment	sit
name		lengt					У	re	lengt	iv
		h							h	es
			gi 1339045451 ref XP_02371085							
			4.1 beta-glucuronidase-like		XP_023					
			[Cryptotermes		71085					
			secundus]gi 1339045453 ref XP_		4,					
	beta-		023710855.1 beta-glucuronidase-		XP_023					
	glucuroni		like [Cryptotermes		71085					
Th_d_00	dase-like		secundus]gi 1330907803 gb PNF		5,		78.2			
080262	isoform		30365.1 Beta-glucuronidase	Blattod	PNF30		608	162		
-RA	X1	708	[Cryptotermes secundus]	ea	365	3.33E-50	7	.54	115	90
			gi 91089481 ref XP_969353.1 PR							
	beta-		EDICTED: beta-glucuronidase							
	glucuroni		[Tribolium		XP_969					
Th_d_00	dase		castaneum]gi 270011394 gb EFA		353,		64.6	66.		
100552	isoform		07842.1 Beta-glucuronidase-like	Coleopt	EFA07		153	625		
-RA	X2	354	Protein [Tribolium castaneum]	era	842	7.51E-12	8	4	65	42
				Arthrop						
Th_d_00	beta-		gi 1316206353 ref XP_02321358	oda-	XP_023		77.5	180		
070705	glucuroni		8.1 beta-glucuronidase-like	Scorpio	21358		362	.64		10
-RA	dase-like	552	[Centruroides sculpturatus]	nes	8	4.75E-55	3	4	138	7
				Arthrop						
Th_d_00	beta-		gi 1316153756 ref XP_02322188	oda-	XP_023		59.1	166		
071727	glucuroni		7.1 beta-glucuronidase-like	Scorpio	22188		549	.39		12
-RA	dase-like	819	[Centruroides sculpturatus]	nes	7	7.65E-46	3	2	213	6

Table 3.5.	Continued.
------------	------------

Sequen	Sequenc	Seque	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posi
се	e desc.	nce				Valu	arity	Scor	ment	tive
name		lengt h				e		е	lengt h	S
Th d 00	beta- glucuroni dase-like		gi 1022772830 gb KZS16987.1							
043633	isoform		Beta-glucuronidase [Daphnia	Crustace	KZS169	7.75E	74.10	177.		
-RA	X1	480	magna]	a	87	-56	072	563	139	103
Th_d_00 043632	beta- glucuroni dase isoform X1	276	gi 662210289 ref XP_00847877 7.1 PREDICTED: beta- glucuronidase-like [Diaphorina citril	Hemipte	XP_008	2.15E	69.73 684	92.4	76	53
Th_d_00 017000	beta- glucuroni dase isoform	270	gi 1339060579 ref XP_0237188 77.1 beta-glucuronidase-like isoform X3 [Cryptotermes secundus]gi 1330889852 gb PN F22050.1 hypothetical protein B7P43_G09739 [Cryptotermes	Blattode	XP_023 718877, PNF220	4.1E-		317.		55
-RA	X3	849	secundus]	а	50	103	72.5	39	280	203

Sequenc	Sequenc	Sequenc	Hit desc.	Order	Hit ACC	E- Valu	Simil	Bit-	Align	Posit
e name	e desc.	eiength				valu	arity	SCOP	longt	ives
						e		e	h	
			gil12283821831reflXP_021						11	
Th d 00			952619.1 alpha-							
074079-	maltase		glucosidase-like [Folsomia	Collemb	XP 0219	1.68E		98.59		
RA	2-like	678	candida]	ola	52619	-21	87.5	69	56	49
	maltase-		gi 1067098296 ref XP_018							
	glucoam		011173.1 PREDICTED:							
Th_d_00	ylase,		maltase-glucoamylase,							
040803-	intestina		intestinal-like [Hyalella	Crustac	XP_0180	2.31E	64.02	199.5		
RA	l-like	697	azteca]	ea	11173	-56	116	19	189	121
	Maltase-									
	glucoam		gi 1022742921 gb KZR99							
Th_d_00	ylase,		685.1 Lysosomal alpha-							
083166-	intestina		glucosidase, partial	Crustac	KZR996	3.96E	64.58	191.4		
RA	1	580	[Daphnia magna]	ea	85	-57	333	3	192	124
Th_d_00			gi 646696837 gb KDR088							
152677-	Maltase		65.1 Maltase 1	Blattod	KDR088	6.92E	85.71	75.48		
RA	1	177	[Zootermopsis nevadensis]	ea	65	-16	429	5	42	36
Th_d_00			gi 1228013734 ref XP_021							
066308-	maltase		938828.1 maltase 2-like	Blattod	XP_0219	5.68E	67.42	187.5		
RA	2-like	732	[Zootermopsis nevadensis]	ea	38828	-54	857	78	175	118

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

Table 3.6.	Continued.
------------	------------

Seque	Seque	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posi
nce	nce	nce			ACC	Valu	arity	Scor	ment	tive
name	desc.	length				е		е	length	S
			gi 1339059399 ref XP_023718249.1							
			maltase 2-like isoform X3							
			[Cryptotermes							
			secundus]gi 1330891182 gb PNF22		XP_023					
			667.1 hypothetical protein		71824					
			B7P43_G07129 [Cryptotermes		9,					
Th_d_			secundus]gi 1330891184 gb PNF22		PNF22					
00031			669.1 hypothetical protein		667,					
469-	Maltas		B7P43_G07129 [Cryptotermes	Blattode	PNF22	2.17E		204.		
RA	e 1	594	secundus]	а	669	-61	72	527	175	126
	maltas									
Th_d_	e A3-									
00031	like									
468-	isofor		gi 636630766 gb AIA09343.1 alpha-	Blattode	AIA09	4.88E	66.15	66.6		
RA	m X2	513	glucosidase [Periplaneta americana]	а	343	-11	385	254	65	43
			gi 1339059399 ref XP_023718249.1							
			maltase 2-like isoform X3							
			[Cryptotermes							
			secundus]gi 1330891182 gb PNF22		XP_023					
			667.1 hypothetical protein		71824					
	maltas		B7P43_G07129 [Cryptotermes		9,					
Th_d_	e 1-		secundus]gi 1330891184 gb PNF22		PNF22					
00112	like		669.1 hypothetical protein		667,					
867-	isofor		B7P43_G07129 [Cryptotermes	Blattode	PNF22	1.09E	68.87	177.		
RA	m X2	483	secundus]	а	669	-51	417	563	151	104

Seque	Seque	Sequ	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
nce	nce	ence			ACC	Val	arity	Scor	ment	ives
name	desc.	lengt				ue		е	lengt	
		h							h	
			gi 1339059399 ref XP_023718249.1 m							
			altase 2-like isoform X3 [Cryptotermes							
			secundus]gi 1330891182 gb PNF2266		XP_02					
			7.1 hypothetical protein		37182					
			B7P43_G07129 [Cryptotermes		49,					
	maltas		secundus]gi 1330891184 gb PNF2266		PNF22					
Th_d_0	e 1-like		9.1 hypothetical protein		667,	4.5				
00563	isofor		B7P43_G07129 [Cryptotermes	Blattod	PNF22	E-	64.76	342.		
95-RA	m X2	1182	secundus]	ea	669	112	965	813	369	239
Th_d_0			gi 1330891185 gb PNF22670.1 hypoth			3.5				
00658	Maltas		etical protein B7P43_G07129	Blattod	PNF22	2E-	71.96	139.		
47-RA	e 1	420	[Cryptotermes secundus]	ea	670	40	262	813	107	77
Th_d_0						3.7				
01084	Maltas		gi 1061480896 gb 0DM98289.1 Malta	Collem	ODM9	E-	77.39	147.		
97-RA	e 1	420	se 1 [Orchesella cincta]	bola	8289	40	13	132	115	89
			gi 1339059399 ref XP_023718249.1 m							
			altase 2-like isoform X3 [Cryptotermes							
			secundus]gi 1330891182 gb PNF2266		XP_02					
			7.1 hypothetical protein		37182					
			B7P43_G07129 [Cryptotermes		49,					
	maltas		secundus]gi 1330891184 gb PNF2266		PNF22					
Th_d_0	e 1-like		9.1 hypothetical protein		667,	2.8				
00599	isofor		B7P43_G07129 [Cryptotermes	Blattod	PNF22	E-	64.72	357.		
40-RA	m X2	1074	secundus]	ea	669	118	222	066	360	233

Seque	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
nce	e desc.	nce longt			ACC	Val	arity	Scor	ment	ives
name		h				ue		C	h	
Th_d_0			gi 1228013734 ref XP_021938828.		XP_02	5.6				
00811			1 maltase 2-like [Zootermopsis	Blattod	19388	1E-	67.66	219.		
00-RA	Maltase 1	600	nevadensis]	ea	28	67	169	935	201	136
	maltase-									
	glucoamy		gi 1067098296 ref XP_018011173.							
Th_d_0	lase,		1 PREDICTED: maltase-		XP_01	2.4				
00611	intestinal		glucoamylase, intestinal-like	Crusta	80111	6E-		141.		
80-RA	-like	1203	[Hyalella azteca]	cea	73	34	73	739	100	73
	Maltase-									
Th_d_0	glucoamy		gi 1022742921 gb KZR99685.1 Lys			4.2				
00845	lase,		osomal alpha-glucosidase, partial	Crusta	KZR99	5E-	64.58	191.		
54-RA	intestinal	598	[Daphnia magna]	cea	685	57	333	43	192	124
	Maltase-									
Th_d_0	glucoamy		gi 1022768805 gb KZS13426.1 Unc			3.7				
00336	lase,		haracterized protein APZ42_021387	Crusta	KZS13	6E-	75.17	192.		
64-RA	intestinal	924	[Daphnia magna]	cea	426	53	241	971	145	109
	maltase-									
	glucoamy		gi 1067098296 ref XP_018011173.							
Th_d_0	lase,		1 PREDICTED: maltase-		XP_01	6.9				
00336	intestinal		glucoamylase, intestinal-like	Crusta	80111	9E-		50.4		
63-RA	-like	402	[Hyalella azteca]	cea	73	06	60	47	55	33
Th_d_0			gi 1228013734 ref XP_021938828.		XP_02	1.9				
01085			1 maltase 2-like [Zootermopsis	Blattod	19388	E-		100.		
05-RA	Maltase 1	222	nevadensis]	ea	28	24	85	523	60	51

Sequenc	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
e name	e desc.	nce			ACC	Val	arity	Scor	ment	ives
		length				ue		е	lengt	
	maltase-								11	
	glucoamy		gil1067098296lreflXP 0180111							
Th d 000	lase,		73.1 PREDICTED: maltase-		XP 01	7.7				
10925-	intestinal		glucoamylase, intestinal-like	Crusta	80111	9E-		81.2		
RA	-like	360	[Hyalella azteca]	cea	73	17	64	629	75	48
				Arthro						
	sucrase-			poda-						
	isomaltas		gi 1238867838 ref XP_0222492	atlantic						
Th_d_000	е,		18.1 sucrase-isomaltase,	horse-	XP_02	7.1				
10928-	intestinal		intestinal-like [Limulus	shoe	22492	E-	71.56	116.		
RA	-like	651	polyphemus]	crab	18	28	863	701	102	73
Th_d_001						7.7				
10401-			gi 1061480896 gb 0DM98289.1	Collem	ODM9	E-	83.05	93.9		
RA	Maltase 1	267	Maltase 1 [Orchesella cincta]	bola	8289	22	085	745	59	49
			gi 1227980746 ref XP_0219219							
Th_d_000			63.1 uncharacterized protein		XP_02	1.5				
96117-			KIAA0513 isoform X3	Blattod	19219	E-	93.25	341.		
RA	maltase 1	555	[Zootermopsis nevadensis]	ea	63	115	843	658	178	166

Seque	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
nce	e desc.	nce			ACC	Val	arity	Scor	ment	ives
name		lengt				ue		е	lengt	
		h							h	
			gi 1339067252 ref XP_023720702.							
			1 uncharacterized protein							
			KIAA0513 [Cryptotermes							
			secundus]gi 1330886365 gb PNF20		XP_02					
			093.1 hypothetical protein		37207					
			B7P43_G05254 [Cryptotermes		02,					
			secundus]gi 1330886366 gb PNF20		PNF20					
Th_d_0			094.1 hypothetical protein		093,	1.7				
00732			B7P43_G05254 [Cryptotermes	Blattod	PNF20	8E-	71.91	221.		
23-RA	Maltase 1	927	secundus]	ea	094	66	781	09	292	210
			gi 1339059590 ref XP_023718348.		XP_02					
			1 maltase 2-like [Cryptotermes		37183					
Th_d_0			secundus]gi 1330891147 gb PNF22		48,					
00086			635.1 Maltase 2 [Cryptotermes	Blattod	PNF22		68.47	608.		
17-RA	Maltase 1	1668	secundus]	ea	635	0	826	216	552	378
Th_d_0						3.5				
00086			gi 1061473040 gb 0DM91568.1 Ma	Collem	ODM9	4E-	65.62	58.9		
16-RA	Maltase 1	335	ltase 1 [Orchesella cincta]	bola	1568	09	5	214	64	42
	sucrase-									
	isomaltas		gi 1067098296 ref XP_018011173.							
Th_d_0	e,		1 PREDICTED: maltase-		XP_01	1.1				
00379	intestinal		glucoamylase, intestinal-like	Crusta	80111	7E-	55.19	305.		
81-RA	-like	1347	[Hyalella azteca]	cea	73	90	63	449	433	239

Table 3.6.	Continued.
------------	------------

Seque	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
nce	e desc.	nce			ACC	Val	arity	Scor	ment	ives
name		lengt				ue		е	lengt	
		h							h	
Th_d_0			gi 1233179883 ref XP_022191429.		XP_02	8.9				
01061	maltase		1 maltase A3-like [Nilaparvata	Hemipt	21914	2E-	61.97	75.8		
92-RA	2-like	228	lugens]	era	29	16	183	702	71	44
			gi 1339059399 ref XP_023718249.							
			1 maltase 2-like isoform X3							
			[Cryptotermes							
			secundus]gi 1330891182 gb PNF22		XP_02					
			667.1 hypothetical protein		37182					
			B7P43_G07129 [Cryptotermes		49,					
			secundus]gi 1330891184 gb PNF22		PNF22					
Th_d_0			669.1 hypothetical protein		667,	7.6				
00945			B7P43_G07129 [Cryptotermes	Blattod	PNF22	3E-	65.21	134.		
75-RA	Maltase 1	432	secundus]	ea	669	36	739	806	138	90
	maltase-		gi 321476730 gb EFX87690.1 hypo							
Th_d_0	glucoamy		thetical protein							
00299	lase,		DAPPUDRAFT_312136 [Daphnia	Crusta	EFX87	4E-	71.05	325.		
84-RA	intestinal	1509	pulex]	cea	690	99	263	094	266	189
			gi 1339059590 ref XP_023718348.		XP_02					
			1 maltase 2-like [Cryptotermes		37183					
Th_d_0			secundus]gi 1330891147 gb PNF22		48,	4.2				
00980	maltase		635.1 Maltase 2 [Cryptotermes	Blattod	PNF22	1E-	73.48	161.		
66-RA	2-like	477	secundus]	ea	635	45	485	384	132	97

Seque	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
nce	e desc.	nce			ACC	Val	arity	Scor	ment	ives
name		lengt				ue		e	lengt	
		h							h	-
			gi 1339067252 ref XP_023720702.							
			1 uncharacterized protein							
			KIAA0513 [Cryptotermes							
			secundus]gi 1330886365 gb PNF20		XP_02					
			093.1 hypothetical protein		37207					
			B7P43_G05254 [Cryptotermes		02,					
			secundus]gi 1330886366 gb PNF20		PNF20					
Th_d_0			094.1 hypothetical protein		093,	5.2				
00917			B7P43_G05254 [Cryptotermes	Blattod	PNF20	5E-	76.83	231.		
60-RA	Maltase 1	819	secundus]	ea	094	71	824	106	272	209
Th_d_0						1.8				
00921			gi 1061480896 gb 0DM98289.1 Ma	Collem	ODM9	8E-	85.93	98.5		
31-RA	Maltase 1	453	ltase 1 [Orchesella cincta]	bola	8289	22	75	969	64	55
			gi 1339059590 ref XP_023718348.		XP_02					
			1 maltase 2-like [Cryptotermes		37183					
Th_d_0			secundus]gi 1330891147 gb PNF22		48,	5.8				
00864	maltase		635.1 Maltase 2 [Cryptotermes	Blattod	PNF22	3E-	74.58	231.		
94-RA	2-like	600	secundus]	ea	635	71	564	106	181	135
	sucrase-									
	isomaltas		gi 1067098296 ref XP_018011173.							
Th_d_0	e,		1 PREDICTED: maltase-		XP_01	4.8				
01055	intestinal		glucoamylase, intestinal-like	Crusta	80111	6E-	65.66	152.		
57-RA	-like	507	[Hyalella azteca]	cea	73	41	265	525	166	109

Table 3.6.	Continued.
------------	------------

Seque	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
nce	e desc.	nce			ACC	Val	arity	Scor	ment	ives
name		lengt				ue		е	lengt	
		h							h	
Th_d_0			gi 1153707331 ref XP_020295361.		XP_02	1.6				
00419			1 alpha-glucosidase-like	Hymen	02953	4E-	75.75	77.7		
01-RA	Maltase 1	201	[Pseudomyrmex gracilis]	optera	61	16	758	962	66	50
Th_d_0			gi 1228013734 ref XP_021938828.		XP_02	1.0				
00934			1 maltase 2-like [Zootermopsis	Blattod	19388	7E-		101.		
57-RA	Maltase 1	438	nevadensis]	ea	28	23	85	679	60	51
			gi 1227980746 ref XP_021921963.							
Th_d_0			1 uncharacterized protein		XP_02	6.5				
00753			KIAA0513 isoform X3	Blattod	19219	E-	96.51	349.		
01-RA	maltase 1	690	[Zootermopsis nevadensis]	ea	63	118	163	747	172	166
	Maltase-		gi 1067098296 ref XP_018011173.							
Th_d_0	glucoamy		1 PREDICTED: maltase-		XP_01	4.7				
00137	lase,		glucoamylase, intestinal-like	Crusta	80111	8E-	61.94	121.		
80-RA	intestinal	525	[Hyalella azteca]	cea	73	30	03	324	134	83
			gi 1339059590 ref XP_023718348.		XP_02					
			1 maltase 2-like [Cryptotermes		37183					
Th_d_0			secundus]gi 1330891147 gb PNF22		48,	8.8				
00930	maltase		635.1 Maltase 2 [Cryptotermes	Blattod	PNF22	E-	75.70	228.		
20-RA	2-like	627	secundus]	ea	635	70	621	409	177	134

Sequence name	Sequ ence dosc	Sequ ence	Hit desc.	Order	Hit ACC	E- Val	Simil arity	Bit- Scor	Align ment	Posit ives
	uesc.	h				ue		C	length	
	alpha-									
	amyla		gi 1000763879 ref XP_015605			3.0				
Th_d_000	se 1-	==0	017.1 PREDICTED: alpha-	Hymeno	XP_0156	2E-	83.60	93.5	(1	= 1
29486-RA	like	558	amylase 1-like [Cephus cinctus]	ptera	05017	20	656	893	61	51
			gi 1339094654 ref XP_023708							
	Alpha		[Cruntotormos		VD 0227					
			secundus]gil1330912880lghlP		08592	61				
Th d 000	amvla		NF32627 1 Alpha-amylase 1	Blattode	PNF3262	1E-	83 56	120		
29487-RA	se 1	837	[Cryptotermes secundus]	a	7	29	164	939	73	61
	alpha-									
	amyla		gi 1000763879 ref XP_015605			4.6				
Th_d_000	se 1-		017.1 PREDICTED: alpha-	Hymeno	XP_0156	5E-	81.57	60.8		
82350-RA	like	538	amylase 1-like [Cephus cinctus]	ptera	05017	09	895	474	38	31
	alpha-									
	amyla		gi 1000763879 ref XP_015605			1.2				
Th_d_001	se 1-	246	017.1 PREDICTED: alpha-	Hymeno	XP_0156	7E-	90.38	92.8	=0	
12407-RA	like	246	amylase 1-like [Cephus cinctus]	ptera	05017	21	462	189	52	47
	alpha-									
	amyia									
	like		gil854959469lgblAKN79742 11			18				
Th d 000	isofor		amylase nartial [Procambarus	Crustace	AKN7974	5E-	64.06	63 5		
58535-RA	m X2	402	clarkii]	a	2	10	25	438	64	41

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

 description.

Sequenc	Sequenc	Seq	Hit desc.	Ord	Hit	Е-	Simi	Bit-	Alig	Posi
e name	e desc.	uen		er	ACC	Valu	larit	Scor	nme	tive
		ce				е	У	e	nt	S
		leng							leng	
		th							th	
	alpha-									
	amylase-									
Th_d_00	like		gi 646713944 gb KDR18093.1 hypotheti				85.5			
049613-	isoform		cal protein L798_07788, partial	Blatt	KDR18	6.29	555	141.3		
RA	X2	732	[Zootermopsis nevadensis]	odea	093	E-37	6	54	90	77
	alpha-									
	amylase-			Hym						
Th_d_00	like		gi 1000763879 ref XP_015605017.1 PRE	eno	XP_015		84.6			
147613-	isoform		DICTED: alpha-amylase 1-like [Cephus	pter	60501	7.59	153	83.57		
RA	X2	156	cinctus]	а	7	E-19	8	41	52	44
Th_d_00	Alpha-						78.5			
086652-	amylase		gi 540849793 gb AGV15452.1 alpha-	Blatt	AGV15	1.84	234	215.3		
RA	1	447	amylase, partial [Blattella germanica]	odea	452	E-66	9	12	149	117
				Hym						
Th_d_00	alpha-		gi 1000763879 ref XP_015605017.1 PRE	eno	XP_015		82.1			
132371-	amylase		DICTED: alpha-amylase 1-like [Cephus	pter	60501	3.22	428	84.72		
RA	1-like	171	cinctus]	а	7	E-19	6	97	56	46
				Hym						
Th_d_00	alpha-		gi 1000763879 ref XP_015605017.1 PRE	eno	XP_015		72.8			
074258-	amylase		DICTED: alpha-amylase 1-like [Cephus	pter	60501	4E-	571	85.50		
RA	1-like	225	cinctus]	а	7	19	4	01	70	51

Sequenc	Sequenc	Seq	Hit desc.	Order	Hit	E-	Similar	Bit-	Alig	Posi
e name	e desc.	uen			ACC	Valu	ity	Scor	nme	tive
		ce				е		е	nt	S
		leng							leng	
		th							th	
	alpha-									
	amylase-									
Th_d_00	like		gi 1000764249 ref XP_01560521		XP_01					
050952-	isoform		0.1 PREDICTED: alpha-amylase 1-	Hymeno	56052	7.65E	88.8888	149.8		
RA	X2	378	like [Cephus cinctus]	ptera	10	-42	9	28	90	80
Th_d_00	alpha-		gi 68266167 gb AAY88846.1 alph							
120835-	amylase		a-amylase, partial [Musca		AAY8	1.03E	73.9130	114.0		
RA	A-like	276	domestica]	Diptera	8846	-31	4	05	92	68
Th_d_00	Alpha-		gi 62955866 gb AAY23288.1 1,4-							
095053-	amylase		alpha-D-glucan glucanohydrolase	Blattod	AAY2	6.52E	77.0833	76.64		
RA	1	249	precursor [Blattella germanica]	ea	3288	-16	3	06	48	37
Th_d_00	Pancreat		gi 1061479081 gb 0DM96628.1 P							
106074-	ic alpha-		ancreatic alpha-amylase	Collemb	ODM9	1.59E	55.5555	76.25		
RA	amylase	294	[Orchesella cincta]	ola	6628	-15	6	54	90	50
	alpha-									
	amylase-									
Th_d_00	like		gi 1000764249 ref XP_01560521		XP_01					
026691-	isoform		0.1 PREDICTED: alpha-amylase 1-	Hymeno	56052	5.74E	88.8888	149.8		
RA	X2	360	like [Cephus cinctus]	ptera	10	-42	9	28	90	80

Sequenc	Sequenc	Seq	Hit desc.	Ord	Hit	Е-	Simi	Bit-	Alig	Posi
e name	e desc.	uen		er	ACC	Valu	larit	Scor	nme	tive
		ce				е	У	е	nt	S
		leng							leng	
		th							th	
			gi 1339094654 ref XP_023708592.1 alp		XP_023					
			ha-amylase 1-like [Cryptotermes		70859					
Th_d_00	Alpha-		secundus]gi 1330912880 gb PNF32627.		2,		83.3			
026690-	amylase		1 Alpha-amylase 1 [Cryptotermes	Blatt	PNF32	3.96	333	125.1		
RA	1	669	secundus]	odea	627	E-31	3	76	78	65
				Hym						
Th_d_00	Pancreat			eno	XP_012		52.1			
060374-	ic alpha-		gi 817060103 ref XP_012251588.1 alpha	pter	25158	9.71	212	107.8		
RA	amylase	726	-amylase A-like [Athalia rosae]	a	8	E-25	1	42	165	86
			gi 1339094654 ref XP_023708592.1 alp		XP_023					
			ha-amylase 1-like [Cryptotermes		70859					
Th_d_00	alpha-		secundus]gi 1330912880 gb PNF32627.		2,		83.5			
013282-	amylase		1 Alpha-amylase 1 [Cryptotermes	Blatt	PNF32	5.94	616	121.3		
RA	1-like	429	secundus]	odea	627	E-31	4	24	73	61
				Hym						
Th_d_00	Pancreat		gi 755948639 ref XP_011300849.1 PRE	eno	XP_011		94.5			
013283-	ic alpha-		DICTED: alpha-amylase-like [Fopius	pter	30084	2.28	945	74.71		
RA	amylase	219	arisanus]	a	9	E-15	9	46	37	35
	alpha-									
	amylase-									
Th_d_00	like		gi 62955866 gb AAY23288.1 1,4-alpha-				82.9			
013284-	isoform		D-glucan glucanohydrolase precursor	Blatt	AAY23	1.69	268	67.01		
RA	X2	735	[Blattella germanica]	odea	288	E-10	3	06	41	34

Sequenc	Sequenc	Seq	Hit desc.	Ord	Hit	E-	Simi	Bit-	Alig	Posi
e name	e desc.	uen		er	ACC	Valu	larit	Scor	nme	tive
		ce				е	У	е	nt	S
		leng							leng	
		th							th	
Th_d_00				Crus			72.3			
048517-	alpha-		gi 1022776232 gb KZS19963.1 Alpha-	tace	KZS19	2.18	404	238.8		
RA	amylase	563	amylase 1 [Daphnia magna]	а	963	E-74	3	1	188	136
Th_d_00	alpha-		gi 62955866 gb AAY23288.1 1,4-alpha-				85.7			
117725-	amylase		D-glucan glucanohydrolase precursor	Blatt	AAY23	1.05	142	75.87		
RA	1-like	240	[Blattella germanica]	odea	288	E-15	9	02	42	36
Th_d_00							73.7			
090295-	alpha		gi 390429576 gb AFL90691.1 alpha	Dipt	AFL90	9.56	704	82.80		
RA	amylase	414	amylase, partial [Ochlerotatus theobaldi]	era	691	E-18	9	37	61	45
Th_d_00			gi 1341138358 gb AVA17418.1 putative				63.4			
013285-	alpha		Per a 11 allergen variant, partial	Blatt	AVA17	5.03	146	82.41		
RA	amylase	843	[Periplaneta americana]	odea	418	E-16	3	85	82	52

Sequence name	Sequen ce desc.	Sequ ence	Hit desc.	Order	Hit ACC	E- Val	Simil arity	Bit- Scor	Align ment	Posit ives
		h				ue		e	length	
Th_d_000 01121-RA	beta- galactos idase-1- like protein 2	2986	gi 1067065262 ref XP_01801 6606.1 PREDICTED: beta- galactosidase-1-like protein 2, partial [Hyalella azteca]	Crustac ea	XP_0180 16606	5.7 E- 22	69.79 167	108. 227	96	67
Th_d_000 01119-RA	Beta- galactos idase-1- like protein 2	1191	gi 1330893432 gb PNF23708. 1 hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattode a	PNF2370 8	1.5 4E- 39	64.47 368	152. 525	152	98
Th_d_000 01120-RA	beta- galactos idase-1- like protein 2	1119	gi 1330893428 gb PNF23704. 1 hypothetical protein B7P43_G02480 [Cryptotermes secundus]	Blattode	PNF2370 4	3.7 4E- 31	66.66 667	127. 872	114	76
Th_d_000 01122-RA	Beta- galactos idase-1- like protein 2	1887	gi 1000751438 ref XP_01559 8573.1 PREDICTED: beta- galactosidase-1-like protein 2 isoform X1 [Cephus cinctus]	Hymeno ptera	XP_0155 98573	0	70.07	648. 277	635	445

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

e namedesc.nce lengthnce lengthweVal uearity ueScor ement lengtives ivesbeta- galactosid Th_d_000 20110-jeil1191621300 gb 0TF82605. 1 beta-galactosidase-like protein, partial [Euroglyphusarity arityScor uement lengt hives eVal uearity eScor ement lengt hives e	sit
IengthIengthueelengthbeta- galactosidgi 1191621300 gb 0TF82605Th_d_000 20110-ase-like1 beta-galactosidase-like-2.4protein, partial [EuroglyphusArachni0TF8266E-68.3565.8-	2S
beta- galactosidgi 1191621300 gb OTF82605.IITh_d_000 20110-ase-like1 beta-galactosidase-like2.4protein, partial [Euroglyphus]ArachniOTF8266E-68.3565.8	
galactosidgi 1191621300 gb OTF82605.Th_d_000ase-like20110-isoformprotein, partial [Euroglyphus]ArachniOTF8266E-68.3565.8	
Th_d_000ase-like1 beta-galactosidase-like2.420110-isoformprotein, partial [EuroglyphusArachniOTF8266E-68.3565.8	
20110- isoform protein, partial [Euroglyphus Arachni OTF826 6E- 68.35 65.8	
RA X1 511 maynei] da 05 18 443 55 79 54	
gi 240999711 ref XP_002404	
775.1 beta-galactosidase,	
putative [Ixodes	
beta- scapularis]gi 215491658 gb E XP_0024	
Th_d_000 galactosid EC01299.1 beta- 04775, 1.3	
20108- ase, galactosidase, putative, partial Arachni EEC0129 7E- 68.83 82.0	
RA putative 294 [Ixodes scapularis] da 9 17 117 333 77 53	
Arthrop	
gi 1238876097 ref XP_02225 oda-	
Th_d_000 beta- 1593.1 beta-galactosidase- Atlantic 4.8	
17291- galactosid like isoform X2 [Limulus horsech XP_0222 4E- 83.33 83.1	
RA ase 483 polyphemus] oe crab 51593 17 333 889 54 45	
galactosia gi 131615445/jref XP_02322 Arthrop 2257 1lbeta galactosidaga	
$\begin{bmatrix} 111_u_000 \\ 35e \end{bmatrix} = \begin{bmatrix} 225/.1 \\ be Contrarregides \end{bmatrix} = \begin{bmatrix} 00a- \\ 1.2 \\ 0101 \\ 0101 \end{bmatrix} = \begin{bmatrix} 1.2 \\ 0101 \\ 0101 \end{bmatrix}$	
$\begin{bmatrix} 1/20/2 \\ pitculsoi \end{bmatrix}$ Inte [centruroides $\begin{bmatrix} 5corpio \\ AP_0232 \\ 22257 \\ 20 \\ 919 \\ 622 \\ 55 \\ 45 \end{bmatrix}$	

Sequenc	Sequence	Seque	Hit desc.	Order	Hit ACC	Е-	Simil	Bit-	Align	Posit
e name	desc.	nce				Val	arity	Scor	ment	ives
		length				ue		е	lengt	
									h	
	beta-									
	galactosid		gi 321478650 gb EFX89607.1			1.3				
Th_d_000	ase		hypothetical protein			E-				
03374-	isoform		DAPPUDRAFT_303198	Crustac	EFX8960	11	66.50	363.		
RA	X2	1720	[Daphnia pulex]	ean	7	5	124	229	403	268
			gi 1325320316 ref XP_02334							
Th_d_001	Beta-		1954.1 beta-galactosidase-			1.5				
17214-	galactosid		like isoform X3 [Eurytemora	Phasma	XP_0233	4E-	81.57	119.		
RA	ase	300	affinis]	todea	41954	30	895	013	76	62
	Beta-		gi 1330893432 gb PNF23708.			1.4				
Th_d_000	galactosid		1 hypothetical protein			E-				
19277-	ase-1-like		B7P43_G02480 [Cryptotermes	Blattod	PNF237	15	71.42	450.		
RA	protein 2	1269	secundus]	ea	08	4	857	284	406	290
			gi 998516749 ref XP_015519							
			174.1 PREDICTED: beta-							
			galactosidase-1-like protein 2							
			Neodiprion							
			lecontei]gi 998516751 ref XP							
	beta-		_015519175.1 PREDICTED:		XP_0155					
Th_d_000	galactosid		beta-galactosidase-1-like		19174,					
19278-	ase-1-like		protein 2 [Neodiprion	Hymen	XP_0155	8E-	68.24	224.		
RA	protein 2	765	lecontei	optera	19175	67	034	172	233	159

Sequenc	Sequence	Seque	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posit
e name	desc.	nce				Val	arity	Scor	ment	ives
		length				ue		е	lengt	
									h	
			gi 321478650 gb EFX89607.1			4.9				
Th_d_000	Beta-		hypothetical protein			E-				
06652-	galactosid		DAPPUDRAFT_303198	Crustac	EFX8960	11	62.83	358.		
RA	ase	1249	[Daphnia pulex]	ean	7	6	619	607	409	257
Th_d_000	beta-		gi 1233169119 ref XP_02218			2.3				
06654-	galactosid		6636.1 beta-galactosidase	Hemipt	XP_0221	9E-	81.35	92.8		
RA	ase	291	[Nilaparvata lugens]	era	86636	21	593	189	59	48
Th_d_000	beta-		gi 1233169119 ref XP_02218			2.0				
06653-	galactosid		6636.1 beta-galactosidase	Hemipt	XP_0221	6E-	74.03	141.		
RA	ase	930	[Nilaparvata lugens]	era	86636	35	846	354	104	77
			gi 1330925577 gb PNF39271.							
Th_d_000	Beta-		1 hypothetical protein			3.7				
38143-	galactosid		B7P43_G16708 [Cryptotermes	Blattod	PNF392	8E-	56.74	181.		
RA	ase	1074	secundus]	ea	71	50	419	8	215	122
	Beta-		gi 1330893429 gb PNF23705.							
Th_d_001	galactosid		1 hypothetical protein			9.8				
02809-	ase-1-like		B7P43_G02480 [Cryptotermes	Blattod	PNF237	8E-	72.72	102.		
RA	protein 2	358	secundus]	ea	05	25	727	834	77	56
	Beta-		gi 1330893432 gb PNF23708.			2.6				
Th_d_000	galactosid		1 hypothetical protein			E-				
32922-	ase-1-like		B7P43_G02480 [Cryptotermes	Blattod	PNF237	11	74.63	339.		
RA	protein 2	855	secundus]	ea	08	3	235	732	272	203

Seque	Sequen	Sequenc	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Pos
nce	ce desc.	e length				Valu	larit	Scor	ment	itiv
name						е	У	е	lengt	es
									h	
Th_d_0	Beta-		gi 321478650 gb EFX89607.1 hypoth				63.1			
00372	galactos		etical protein DAPPUDRAFT_303198	Crusta	EFX8960	1.1E-	840	362.		
52-RA	idase	1416	[Daphnia pulex]	cea	7	116	8	459	402	254
			gi 1000751440 ref XP_015598574.1 P							
			REDICTED: beta-galactosidase-1-like							
			protein 2 isoform X2 [Cephus							
			cinctus]gi 1000751442 ref XP_015598							
			576.1 PREDICTED: beta-galactosidase-							
			1-like protein 2 isoform X2 [Cephus							
			cinctus]gi 1000751444 ref XP_015598							
			577.1 PREDICTED: beta-galactosidase-		XP_0155					
			1-like protein 2 isoform X2 [Cephus		98574,					
			cinctus]gi 1000751446 ref XP_015598		XP_0155					
			578.1 PREDICTED: beta-galactosidase-		98576,					
			1-like protein 2 isoform X2 [Cephus		XP_0155					
			cinctus]gi 1000751448 ref XP_015598		98577,					
	beta-		579.1 PREDICTED: beta-galactosidase-		XP_0155					
	galactos		1-like protein 2 isoform X2 [Cephus		98578,					
	idase-1-		cinctus]gi 1000751450 ref XP_015598		XP_0155					
Th_d_0	like		580.1 PREDICTED: beta-galactosidase-	Hyme	98579,		65.3			
00521	protein		1-like protein 2 isoform X2 [Cephus	nopte	XP_0155	1.3E-	631	342.		
74-RA	2	1143	cinctus]	ra	98580	110	3	813	358	234

Seque nce name	Sequen ce desc.	Sequenc e length	Hit desc.	Order	Hit ACC	E- Valu e	Simi larit y	Bit- Scor e	Align ment lengt h	Pos itiv es
	beta- galactos idase-									
Th_d_0	like						76.7			
00793	isoform		gi 1233169119 ref XP_022186636.1 b	Hemip	XP_0221	2.11	857	163.		
38-RA	X1	513	eta-galactosidase [Nilaparvata lugens]	tera	86636	E-45	1	696	112	86

Sequenc	Seque	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posit
e name	nce	ence				Val	arity	Scor	ment	ives
	uest.	h				ue		e	h	
			gi 768427825 ref XP_011555023.1							
Th_d_001	myros		PREDICTED: uncharacterized			1.7				
44418-	inase		protein LOC105386209 [Plutella	Lepido	XP_0115	7E-	73.68	119.		
RA	1-like	306	xylostella]	ptera	55023	30	421	013	95	70
			gi 1339068038 ref XP_023721112.							
			1 myrosinase 1-like [Cryptotermes							
			secundus]gi 1339068040 ref XP_0							
			23721113.1 myrosinase 1-like		XP_0237					
			[Cryptotermes		21112,					
			secundus]gi 1330885855 gb PNF1		XP_0237					
			9653.1 Myrosinase 1		21113,					
			[Cryptotermes		PNF1965					
Th_d_000	myros		secundus]gi 1330885856 gb PNF1		3,	8.9				
30415-	inase		9654.1 Myrosinase 1	Blatto	PNF1965	E-	65.81	358.		
RA	1-like	1170	[Cryptotermes secundus]	dea	4	118	197	992	351	231
Th_d_000	myros		gi 636630776 gb AIA09348.1 beta			9.2				
55165-	inase		-glucosidase [Periplaneta	Blatto	AIA0934	6E-	79.22	112.		
RA	1-like	243	americana]	dea	8	29	078	464	77	61
			gi 910313450 ref XP_013175517.1							
Th_d_000	myros		PREDICTED: uncharacterized			6.8				
46945-	inase		protein LOC106123648 [Papilio	Lepido	XP_0131	1E-	83.33	76.6		
RA	1-like	426	xuthus]	ptera	75517	15	333	406	48	40

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

Table 3.9. Continued.

Sequen	Seque	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
се	nce	nce			ACC	Valu	arity	Scor	ment	itiv
name	desc.	length				e		e	lengt	es
									h	
Th_d_00	myrosi		gi 1114635916 gb APM84101.1 GH							
063063	nase 1-		1 beta-glucosidase [Microcerotermes	Blattod	APM84	2.94E	78.04	122.0		
-RA	like	636	annandalei]	ea	101	-30	878	94	82	64
			gi 1339068038 ref XP_023721112.1		XP_023					
			myrosinase 1-like [Cryptotermes		72111					
			secundus]gi 1339068040 ref XP_02		2,					
			3721113.1 myrosinase 1-like		XP_023					
			[Cryptotermes		72111					
			secundus]gi 1330885855 gb PNF19		3,					
			653.1 Myrosinase 1 [Cryptotermes		PNF19					
Th_d_00	myrosi		secundus]gi 1330885856 gb PNF19		653,					
012572	nase 1-		654.1 Myrosinase 1 [Cryptotermes	Blattod	PNF19	6.6E-	62.69	340.8		
-RA	like	1560	secundus]	ea	654	109	43	87	386	242
Th_d_00	myrosi		gi 910324408 ref XP_013165281.1		XP_013					
135824	nase 1-		PREDICTED: myrosinase 1-like	Lepido	16528	2.82E	64.15	107.4		
-RA	like	321	[Papilio xuthus]	ptera	1	-26	094	57	106	68
Th_d_00										
113027	Myrosi		gi 1061486758 gb 0DN03847.1 Myr	Collem	ODN03	3.72E	59.67	127.4		
-RA	nase 1	372	osinase 1 [Orchesella cincta]	bola	847	-33	742	87	124	74
Th_d_00			gi 746857492 ref XP_011059478.1		XP_011					
117962	Myrosi		PREDICTED: myrosinase 1 isoform	Hymen	05947	1.61E	82.22	126.7		
-RA	nase 1	321	X2 [Acromyrmex echinatior]	optera	8	-33	222	16	90	74

Table 3.9.	Continued.
------------	------------

Sequen	Seque	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
се	nce	nce			ACC	Valu	arity	Scor	ment	itiv
name	desc.	length				е		е	lengt	es
									h	
Th_d_00			gi 506967929 gb AGM32308.1 beta-							
105850	Myrosi		glucosidase, partial [Coptotermes	Blattod	AGM32	4.92E	84.14	120.1		
-RA	nase 1	246	formosanus]	ea	308	-33	634	68	82	69
Th_d_00	myrosi		gi 269965728 dbj BAI50023.1 beta-							
058602	nase 1-		glucosidase [Nasutitermes	Blattod	BAI500	2.49E	73.85	277.3		
-RA	like	678	takasagoensis]	ea	23	-89	321	3	218	161
			gi 1339068038 ref XP_023721112.1		XP_023					
			myrosinase 1-like [Cryptotermes		72111					
			secundus]gi 1339068040 ref XP_02		2,					
			3721113.1 myrosinase 1-like		XP_023					
			[Cryptotermes		72111					
			secundus]gi 1330885855 gb PNF19		3,					
			653.1 Myrosinase 1 [Cryptotermes		PNF19					
Th_d_00			secundus]gi 1330885856 gb PNF19		653,					
056050	Myrosi		654.1 Myrosinase 1 [Cryptotermes	Blattod	PNF19	2.9E-	77.69	319.3		
-RA	nase 1	798	secundus]	ea	654	104	231	16	260	202
Th_d_00	myrosi		gi 1233170929 ref XP_022187264.1		XP_022					
093969	nase 1-		myrosinase 1-like isoform X2	Hemipt	18726	1.79E	75.69			
-RA	like	441	[Nilaparvata lugens]	era	4	-54	444	181.8	144	109
Th_d_00										
058955	Myrosi		gi 574959088 gb AHG54239.1 beta-	Hemipt	AHG54	1.88E	77.17	132.4		
-RA	nase 1	579	glucosidase 3 [Lygus lineolaris]	era	239	-34	391	94	92	71

Table 3.9.	Continued.
------------	------------

Sequen	Seque	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
се	nce	nce			ACC	Valu	arity	Scor	ment	itiv
name	desc.	length				е		е	lengt	es
									h	
			gi 1070157753 ref XP_018359559.1							
			PREDICTED: myrosinase 1-like		XP_018					
			[Trachymyrmex		35955					
Th_d_00			cornetzi]gi 1009403499 gb KYN230		9,					
100160	Myrosi		32.1 Myrosinase 1 [Trachymyrmex	Hymen	KYN23	7.4E-	82.60	128.2		
-RA	nase 1	369	cornetzi]	optera	032	34	87	57	92	76
Th_d_00	myrosi		gi 910324408 ref XP_013165281.1		XP_013					
141241	nase 1-		PREDICTED: myrosinase 1-like	Lepido	16528	1.11E	79.72	107.0		
-RA	like	411	[Papilio xuthus]	ptera	1	-25	973	71	74	59
Th_d_00	myrosi		gi 364023587 gb AEW46868.1 semi							
089076	nase 1-		nal fluid protein CSSFP018, partial	Lepido	AEW4	2.5E-	74.50	139.8		
-RA	like	366	[Chilo suppressalis]	ptera	6868	40	98	13	102	76
Th_d_00	myrosi		gi 1080062831 ref XP_018573476.1		XP_018					
076743	nase 1-		myrosinase 1-like [Anoplophora	Coleop	57347	3.39E	81.95	172.9		
-RA	like	450	glabripennis]	tera	6	-50	489	4	133	109
Th_d_00	myrosi		gi 506967929 gb AGM32308.1 beta-							
111777	nase 1-		glucosidase, partial [Coptotermes	Blattod	AGM32	2.21E	84.09	67.78		
-RA	like	192	formosanus]	ea	308	-13	091	1	44	37
Th_d_00	myrosi		gi 913323605 ref XP_013193075.1		XP_013					
074352	nase 1-		PREDICTED: myrosinase 1-like	Lepido	19307	2.95E	77.77	69.70		
-RA	like	486	[Amyelois transitella]	ptera	5	-12	778	7	45	35

Sequen	Sequ	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Alig	Po
се	ence	nce		er		Valu	larit	Scor	nme	sit
name	desc.	lengt				e	У	е	nt	iv
		h							leng	es
									th	
	myro						70.4			
In_d_00	sinas		gi 1101349013 ref XP_018900400.1 PREDI	Hem		0.65	/9.4	100		
064915	e 1-	504	CTED: myrosinase 1-like isoform X2	ipter	XP_0189	8.6E-	117	130.	100	0.1
-RA	like	531	[Bemisia tabaci]	a	00400	34	6	568	102	81
Th_d_00	myro			Cole		5 045	64.4	100		
053277	sinas	-06	gi 748995286 gb A]E75665.1 putative	opte	AJE7566	7.01E	808	192.	100	11
-RA	e 1	786	glycosyl hydrolase [Chrysomela lapponica]	ra	5	-56	7	2	183	8
	myro									
Th_d_00	sinas			Lepi			72.2			
074186	e 1-		gi 930669849 gb KPJ10712.1 hypothetical	dopt	KPJ1071	3.09E	891	105.		
-RA	like	936	protein RR48_07718 [Papilio machaon]	era	2	-24	6	145	83	60
	myro									
Th_d_00	sinas		gi 768432251 ref XP_011557441.1 PREDIC	Lepi			67.2			
126328	e 1-		TED: uncharacterized protein	dopt	XP_0115	8.38E	268	136.		
-RA	like	687	LOC105388261 [Plutella xylostella]	era	57441	-35	9	346	119	80
	myro									
Th_d_00	sinas			Hem			69.6			
052665	e 1-		gi 1101347123 ref XP_018899386.1 PREDI	ipter	XP_0188	2.86E	682	225.		14
-RA	like	741	CTED: myrosinase 1 [Bemisia tabaci]	а	99386	-69	5	713	211	7
	myro									
Th_d_00	sinas			Lepi			80.2			
141011	e 1-		gi 910324408 ref XP_013165281.1 PREDIC	dopt	XP_0131	8.52E	325	127.		
-RA	like	258	TED: myrosinase 1-like [Papilio xuthus]	era	65281	-34	6	872	86	69

Sequen	Sequ	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Alig	Po
се	ence	nce		er		Valu	larit	Scor	nme	sit
name	desc.	lengt				е	У	е	nt	iv
		h							leng	es
									th	
			gi 795032132 ref XP_011863828.1 PREDIC							
	myro		TED: myrosinase 1-like [Vollenhovia	Hym	XP_0118					
Th_d_00	sinas		emeryi]gi 795032135 ref XP_011863829.1	eno	63828,		49.1			
043623	e 1-		PREDICTED: myrosinase 1-like [Vollenhovia	pter	XP_0118	7.62E	228	84.7		
-RA	like	1014	emeryi]	а	63829	-16	1	297	114	56
			gi 1339068038 ref XP_023721112.1 myros		XP_0237					
			inase 1-like [Cryptotermes		21112,					
			secundus]gi 1339068040 ref XP_02372111		XP_0237					
			3.1 myrosinase 1-like [Cryptotermes		21113,					
			secundus]gi 1330885855 gb PNF19653.1		PNF196					
Th_d_00	Myro		Myrosinase 1 [Cryptotermes		53,		62.8			
014522	sinas		secundus]gi 1330885856 gb PNF19654.1	Blatt	PNF196	1.2E-	078	364.		25
-RA	e 1	1410	Myrosinase 1 [Cryptotermes secundus]	odea	54	118	8	385	406	5
Th_d_00	Myro						62.3			
082198	sinas		gi 636630776 gb AIA09348.1 beta-	Blatt	AIA0934	3.66E	529	59.6		
-RA	e 1	678	glucosidase [Periplaneta americana]	odea	8	-08	4	918	85	53
	myro			Hym						
Th_d_00	sinas		gi 795043525 ref XP_011867620.1 PREDIC	eno			87.2			
084786	e 1-		TED: myrosinase 1-like [Vollenhovia	pter	XP_0118	2.65E	340	83.1		
-RA	like	417	emeryi]	а	67620	-17	4	889	47	41
Th_d_00	Myro			Lepi			56.4			
066036	sinas		gi 930652287 gb KPI94627.1 Myrosinase 1	dopt	KPI9462	8.92E	516	202.		14
-RA	e 1	744	[Papilio xuthus]	era	7	-60	1	216	248	0

Table 3	.9. Con	tinued.
---------	---------	---------

Sequen	Seque	Sequenc	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Alig	Po
ce	nce	e length				Valu	larit	Scor	nme	sit
name	desc.					е	У	е	nt	iv
									leng	es
									th	
Th_d_00	myrosi		gi 910324408 ref XP_013165281.1				78.0			
062853	nase 1-		PREDICTED: myrosinase 1-like	Lepido	XP_0131	1.38E	487	61.2		
-RA	like	228	[Papilio xuthus]	ptera	65281	-10	8	326	41	32
Th_d_00	myrosi		gi 1080062338 ref XP_018573207.1				64.2			
041776	nase 1-		myrosinase 1-like isoform X2	Coleop	XP_0185	5.58E	857	268.		16
-RA	like	840	[Anoplophora glabripennis]	tera	73207	-85	1	47	252	2
Th_d_00	myrosi		gi 506967929 gb AGM32308.1 beta-							
052134	nase 1-		glucosidase, partial [Coptotermes	Blattod	AGM323	6.99E	72.9	231.		14
-RA	like	693	formosanus]	ea	08	-74	064	106	203	8
Th_d_00	myrosi		gi 506967929 gb AGM32308.1 beta-				86.3			
101375	nase 1-		glucosidase, partial [Coptotermes	Blattod	AGM323	1.39E	636	75.4		
-RA	like	543	formosanus]	ea	08	-14	4	85	44	38
Th_d_00	myrosi		gi 1325349814 ref XP_023310604.1				83.8			
038374	nase 1-		myrosinase 1-like [Anoplophora	Coleop	XP_0233	2.9E-	709	97.8		
-RA	like	705	glabripennis]	tera	10604	21	7	265	62	52
Th_d_00	myrosi						86.3			
038373	nase 1-		gi 636630776 gb AIA09348.1 beta-	Blattod	AIA0934	9.32E	636	113.		
-RA	like	294	glucosidase [Periplaneta americana]	ea	8	-29	4	235	88	76

Sequen	Seque	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Alig	Po
се	nce	nce		er		Valu	larit	Scor	nme	sit
name	desc.	lengt				е	У	е	nt	iv
		h							leng	es
									th	
			gi 1339068038 ref XP_023721112.1 myr		XP_0237					
			osinase 1-like [Cryptotermes		21112,					
			secundus]gi 1339068040 ref XP_0237211		XP_0237					
			13.1 myrosinase 1-like [Cryptotermes		21113,					
			secundus]gi 1330885855 gb PNF19653.1		PNF196					
Th_d_00	myrosi		Myrosinase 1 [Cryptotermes		53,		67.0			
031000	nase 1-		secundus]gi 1330885856 gb PNF19654.1	Blatt	PNF196	8E-	360	386.		24
-RA	like	1119	[Myrosinase 1 [Cryptotermes secundus]	odea	54	129	1	726	361	2
			gi 1339068038 ref XP_023721112.1 myr		XP_0237					
			osinase 1-like [Cryptotermes		21112,					
			secundus]gi 1339068040 ref XP_0237211		XP_0237					
			13.1 myrosinase 1-like [Cryptotermes		21113,					
			secundus]gi 1330885855 gb PNF19653.1		PNF196					
Th_d_00	myrosi		Myrosinase 1 [Cryptotermes		53,		75.4			
033250	nase 1-		secundus]gi 1330885856 gb PNF19654.1	Blatt	PNF196	1.1E-	475	469.		29
-RA	like	1290	[Myrosinase 1 [Cryptotermes secundus]	odea	54	160	7	929	391	5
Th_d_00	myrosi		gi 939630683 ref XP_014278711.1 PRED	Hem			78.2			
032603	nase 1-		ICTED: myrosinase 1-like [Halyomorpha	ipter	XP_0142	3.67E	608	68.9		
-RA	like	255	halys]	а	78711	-13	7	366	46	36
Th_d_00			gi 506967929 gb AGM32308.1 beta-				75.8			
032602	Myrosi		glucosidase, partial [Coptotermes	Blatt	AGM323	4.97E	169	201.		11
-RA	nase 1	717	formosanus]	odea	08	-62	9	06	153	6

Table 3.9.	Continued.
------------	------------

Sequen	Seque	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Alig	Po
ce	nce	nce		er		Valu	larit	Scor	nme	sit
name	desc.	lengt				е	У	е	nt	iv
		h							leng	es
									th	
Th_d_00			gi 364023593 gb AEW46871.1 seminal	Lepi			84.2			
053701	Myrosi		fluid protein CSSFP021 [Chilo	dopt	AEW468	5.44E	105	120.		
-RA	nase 1	762	suppressalis]	era	71	-29	3	168	76	64
Th_d_00				Lepi			77.2			
104679	Myrosi		gi 930661725 gb KPJ03612.1 Myrosinase	dopt	KPJ0361	2.28E	277	147.		
-RA	nase 1	303	1 [Papilio xuthus]	era	2	-40	2	132	101	78
Th_d_00	myrosi		gi 1233170929 ref XP_022187264.1 myr	Hem			72.4			
103233	nase 1-		osinase 1-like isoform X2 [Nilaparvata	ipter	XP_0221	5.14E	137	105.		
-RA	like	341	lugens]	а	87264	-26	9	531	87	63
Th_d_00	myrosi		gi 1233170929 ref XP_022187264.1 myr	Hem			64.4			
032475	nase 1-		osinase 1-like isoform X2 [Nilaparvata	ipter	XP_0221	3.18E	927	277.		17
-RA	like	900	lugens]	а	87264	-89	5	33	276	8
Th_d_00	myrosi		gi 768432251 ref XP_011557441.1 PRED	Lepi			65.6			
015265	nase 1-		ICTED: uncharacterized protein	dopt	XP_0115	1.73E	488	147.		
-RA	like	402	LOC105388261 [Plutella xylostella]	era	57441	-40	5	902	131	86
Th_d_00	myrosi									
056242	nase 1-		gi 568256136 gb ETN64735.1 lactase-	Dipt	ETN647	2.23E	53.1	128.		10
-RA	like	891	phlorizin hydrolase [Anopheles darlingi]	era	35	-31	25	642	192	2
Th_d_00	myrosi		gi 1233170929 ref XP_022187264.1 myr	Hem			67.3			
068261	nase 1-		osinase 1-like isoform X2 [Nilaparvata	ipter	XP_0221	3.63E	758	157.		
-RA	like	616	lugens]	а	87264	-44	9	532	141	95

Table 3.9. Continued.

Sequen	Seque	Seque	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Alig	Po
се	nce	nce		er		Valu	larit	Scor	nme	sit
name	desc.	lengt				е	У	e	nt	iv
		h							leng	es
									th	
Th_d_00	myrosi		gi 1121153069 ref XP_019563383.1 PRE				69.3			
033820	nase 1-		DICTED: lactase-phlorizin hydrolase-like	Dipt	XP_0195	6.31E	693	117.		
-RA	like	1359	[Aedes albopictus]	era	63383	-26	7	857	111	77
Th_d_00	myrosi		gi 768427825 ref XP_011555023.1 PRED	Lepi			75.7			
078948	nase 1-		ICTED: uncharacterized protein	dopt	XP_0115	7.56E	281	129.		
-RA	like	336	LOC105386209 [Plutella xylostella]	era	55023	-34	6	028	103	78
			gi 1070157753 ref XP_018359559.1 PRE							
			DICTED: myrosinase 1-like	Hym	XP_0183					
Th_d_00			[Trachymyrmex	eno	59559,		71.6			
085145	Myrosi		cornetzi]gi 1009403499 gb KYN23032.1	pter	KYN230	1.27E	417	148.		
-RA	nase 1	495	Myrosinase 1 [Trachymyrmex cornetzi]	а	32	-40	9	288	134	96
Th_d_00	myrosi		gi 768427825 ref XP_011555023.1 PRED	Lepi			60.1			
150104	nase 1-		ICTED: uncharacterized protein	dopt	XP_0115	1.23E	449	124.		
-RA	like	414	LOC105386209 [Plutella xylostella]	era	55023	-31	3	02	138	83
Th_d_00	myrosi		gi 364023587 gb AEW46868.1 seminal	Lepi						
094424	nase 1-		fluid protein CSSFP018, partial [Chilo	dopt	AEW468	2.38E	74.5	137.		
-RA	like	399	suppressalis]	era	68	-39	098	502	102	76
Th_d_00	myrosi						56.8			
084437	nase 1-		gi 636630776 gb AIA09348.1 beta-	Blatt	AIA0934	2.8E-	965	85.5		
-RA	like	613	glucosidase [Periplaneta americana]	odea	8	17	5	001	116	66

Table 3.9.	Continued.
------------	------------

Sequen ce name	Seque nce desc.	Seque nce lengt h	Hit desc.	Ord er	Hit ACC	E- Valu e	Simi larit y	Bit- Scor e	Alig nme nt leng	Po sit iv es
									th	
Th_d_00				Hem			80.5			
088179	myrosi		gi 574959088 gb AHG54239.1 beta-	ipter	AHG542	9.24E	194	118.		
-RA	nase 1	483	glucosidase 3 [Lygus lineolaris]	а	39	-30	8	627	77	62
Th_d_00				Lepi			76.4			
095744	Myrosi		gi 930661725 gb KPJ03612.1 Myrosinase	dopt	KPJ0361	1.02E	705	149.		
-RA	nase 1	357	1 [Papilio xuthus]	era	2	-40	9	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.

Sequen	Sequence desc.	Sequ	Hit desc.	Orde	Hit	E-	Simi	Bit-	Alig	Pos
ce		ence		r	ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		n							leng	
Th d 00	Lutio		cil12400E04171cmblSUM61272111				776	105	ui	
111_U_00	Lytic		gi 1549059417 jellib jsi w 01575.1 jL		SIMC	214	022	105		
D 4	polysaccilariue	700	ytic polysaccilariue monocygenase,		1272	2.14 E E 6	923	.05	120	101
-KA	monooxygenase	789	partial [Thermobia domestica]		13/3	E-20	1	2	130	101
_, , , , , ,										
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				74.0	259		
098297	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2E-	740	.99		
-RA	monooxygenase	645	partial [Thermobia domestica]		1373	86	7	6	216	160
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				84.3	317		
119312	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	6E-	434	.00		
-RA	monooxygenase	645	partial [Thermobia domestica]		1372	109	3	5	198	167
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				73.8	112		
120034	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.99	636	.46		
-RA	monooxygenase	270	partial [Thermobia domestica]		1373	E-30	4	4	88	65
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				78.8			
110473	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.61	235	238		
-RA	monooxygenase	570	partial [Thermobia domestica]		1373	E-78	3	.81	170	134
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				78.6	286		
083543	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.46	069	.18		
-RA	monooxygenase	684	partial [Thermobia domestica]		1373	E-96	7	9	201	158
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				63.5	247		
052762	polysaccharide		ytic polysaccharide monooxygenase.		SIW6	4.36	627	.66		
-RA	monooxygenase	<u>93</u> 6	partial [Thermobia domestica]		1372	E-80	5	9	247	157

Sequen	Sequence desc.	Seq	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
се		uen			ACC	Valu	larit	Sco	nme	itiv
name		се				е	У	re	nt	es
		leng							leng	
		th							th	
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					447		
007079	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.1E		.20		
-RA	monooxygenase	697	partial [Thermobia domestica]		1372	-160	100	3	213	213
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				84.3			
007078	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	8.5E	434	316		
-RA	monooxygenase	645	partial [Thermobia domestica]		1372	-109	3	.62	198	167
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				78.8	242		
007075	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.1E	235	.27		
-RA	monooxygenase	999	partial [Thermobia domestica]		1373	-77	3	6	170	134
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				78.3	226		
007076	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.91	950	.86		
-RA	monooxygenase	507	partial [Thermobia domestica]		1373	E-74	6	8	162	127
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				72.5	168		
056739	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.88	190	.70		
-RA	monooxygenase	519	partial [Thermobia domestica]		1372	E-51	8	3	131	95
				Arthrop						
				oda-						
				chelicer						
			gi 926610136 ref XP_013787168.1	ata-						
Th_d_00	Lytic		uncharacterized protein	atlantic	XP_0		73.4	213		
120998	polysaccharide		LOC106471123 [Limulus	horsesh	1378	5.85	177	.00		
-RA	monooxygenase	477	polyphemus]	oe crab	7168	E-69	2	1	158	116

Table 3.1	0. Cont	tinued.
-----------	---------	---------

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
ce		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				80.6	303		
058386	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	5.5E	930	.52		
-RA	monooxygenase	936	partial [Thermobia domestica]		1373	-102	7	3	202	163
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L					221		
114368	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	4.45	75.5	.86		
-RA	monooxygenase	693	partial [Thermobia domestica]		1373	E-71	814	1	172	130
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					164		
075908	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	8.07	73.4	.46		
-RA	monooxygenase	462	partial [Thermobia domestica]		1372	E-50	375	6	128	94
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				75.5	276		
058359	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	4.37	868	.94		
-RA	monooxygenase	657	partial [Thermobia domestica]		1372	E-93	5	4	213	161
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				63.1	229		
056072	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.04	782	.56		
-RA	monooxygenase	765	partial [Thermobia domestica]		1373	E-73	9	5	258	163
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				74.8	218		
064235	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	4.19	691	.77		
-RA	monooxygenase	849	partial [Thermobia domestica]		1373	E-69	1	9	191	143
			gi 1009534315 ref XP_015904140.							
Th_d_00	Lytic		1 uncharacterized protein		XP_0		71.5	266		
056727	polysaccharide		LOC107436837 [Parasteatoda		1590	5.7E	639	.15		
-RA	monooxygenase	758	tepidariorum]	Spider	4140	-88	8	9	211	151

Table 3.1	0. Cont	tinued.
-----------	---------	---------

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
се		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				75.7	213		
083583	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	9.69	961	.00		
-RA	monooxygenase	865	partial [Thermobia domestica]		1373	E-67	8	1	157	119
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				66.6	199		
003122	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.3E	666	.13		
-RA	monooxygenase	817	partial [Thermobia domestica]		1372	-61	7	4	195	130
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				81.6	298		
023793	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.9E	831	.90		
-RA	monooxygenase	708	partial [Thermobia domestica]		1373	-101	7	1	202	165
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				79.0	229		
023791	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.41	123	.56		
-RA	monooxygenase	507	partial [Thermobia domestica]		1373	E-75	5	5	162	128
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				77.1			
023792	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	8.8E	929	77.		
-RA	monooxygenase	297	partial [Thermobia domestica]		1372	-17	8	411	57	44
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					447		
058774	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.5E		.20		
-RA	monooxygenase	882	partial [Thermobia domestica]		1372	-159	100	3	213	213
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				78.2	274		
057442	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	9.18	383	.24		
-RA	monooxygenase	895	partial [Thermobia domestica]		1373	E-91	4	8	193	151

Table 3.10. Continu	ed.
---------------------	-----

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
се		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
			gi 1080063909 ref XP_018574059.							
Th_d_00	Lytic		1 uncharacterized protein		XP_0			74.		
118174	polysaccharide		LOC108913075 [Anoplophora		1857	5.39		714		
-RA	monooxygenase	312	glabripennis]	Beetle	4059	E-15	75	6	52	39
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				99.4	358		
109123	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.1E	594	.60		
-RA	monooxygenase	633	partial [Thermobia domestica]		1373	-125	6	7	185	184
			gi 1009534315 ref XP_015904140.							
Th_d_00	Lytic		1 uncharacterized protein		XP_0		73.1			
014400	polysaccharide		LOC107436837 [Parasteatoda		1590	2.3E	132	258		
-RA	monooxygenase	834	tepidariorum]	Spider	4140	-84	1	.07	212	155
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				78.6			
075450	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.78	069	287		
-RA	monooxygenase	870	partial [Thermobia domestica]		1373	E-96	7	.73	201	158
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				77.8	290		
072326	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	5.13	301	.04		
-RA	monooxygenase	894	partial [Thermobia domestica]		1372	E-97	9	1	212	165
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					279		
049111	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.79	75.9	.64		
-RA	monooxygenase	799	partial [Thermobia domestica]		1372	E-93	434	1	212	161

Table 3.1	0. Cont	tinued.
-----------	---------	---------

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
се		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				76.1	196		
104344	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.41	904	.43		
-RA	monooxygenase	460	partial [Thermobia domestica]		1372	E-62	8	8	147	112
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				75.3			
061982	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.33	926	219		
-RA	monooxygenase	850	partial [Thermobia domestica]		1373	E-69	7	.55	191	144
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				76.1	196		
102441	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.41	904	.43		
-RA	monooxygenase	460	partial [Thermobia domestica]		1372	E-62	8	8	147	112
			gi 1067063705 ref XP_018015775.							
Th_d_00	Lytic		1 PREDICTED: uncharacterized		XP_0		74.8	184		
140949	polysaccharide		protein LOC108672588 [Hyalella	Crusta	1801	2.47	031	.88		
-RA	monooxygenase	402	azteca]	cean	5775	E-53	5	2	127	95
			gi 1009534315 ref XP_015904140.							
Th_d_00	Lytic		1 uncharacterized protein		XP_0		72.0	266		
084210	polysaccharide		LOC107436837 [Parasteatoda		1590	4.11	379	.54		
-RA	monooxygenase	758	tepidariorum]	Spider	4140	E-88	1	4	211	152
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L					419		
095724	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.7E		.46		
-RA	monooxygenase	645	partial [Thermobia domestica]		1373	-149	100	8	214	214

Table 3.10. Cont	inued.
------------------	--------

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
ce		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				76.1	280		
064622	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.16	682	.79		
-RA	monooxygenase	865	partial [Thermobia domestica]		1373	E-93	2	6	214	163
				Arthr						
				opoda						
				-						
				chelic						
				erata-						
				atlanti						
			gi 1238888528 ref XP_022256420.	с						
Th_d_00	Lytic		1 uncharacterized protein	horses	XP_0		87.8	68.		
158168	polysaccharide		LOC106472108 [Limulus	hoe	2225	3.97	787	936		
-RA	monooxygenase	183	polyphemus]	crab	6420	E-14	9	6	33	29
Th_d_00	Lytic		gi 1061482501 gb 0DM99783.1 hy		ODM			149		
047815	polysaccharide		pothetical protein Ocin01_06883	Spring	9978	3.04		.44		
-RA	monooxygenase	384	[Orchesella cincta]	tail	3	E-44	75	3	124	93
Th_d_00	Lytic		gi 1061482501 gb 0DM99783.1 hy		ODM		74.7	148		
054826	polysaccharide		pothetical protein Ocin01_06883		9978	1.12	967	.67		
-RA	monooxygenase	630	[Orchesella cincta]		3	E-42	5	3	123	92
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				72.7	223		
108306	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.32	272	.78		
-RA	monooxygenase	555	partial [Thermobia domestica]		1372	E-72	7	7	176	128

Table 3.1	0. Cont	tinued.
-----------	---------	---------

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
се		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				60.6	182		
075363	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.31	557	.95		
-RA	monooxygenase	786	partial [Thermobia domestica]		1372	E-55	4	6	183	111
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					445		
058638	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.7E		.66		
-RA	monooxygenase	882	partial [Thermobia domestica]		1372	-158	100	2	213	213
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				77.8	290		
076106	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.04	301	.81		
-RA	monooxygenase	894	partial [Thermobia domestica]		1372	E-97	9	2	212	165
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				74.1	75.		
092785	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2E-	935	870		
-RA	monooxygenase	459	partial [Thermobia domestica]		1373	15	5	2	62	46
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				80.6	303		
068335	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	5.2E	930	.90		
-RA	monooxygenase	777	partial [Thermobia domestica]		1373	-103	7	8	202	163
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				77.6	233		
112031	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.98	470	.80		
-RA	monooxygenase	570	partial [Thermobia domestica]		1373	E-76	6	2	170	132
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				75.4			
032707	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	2.15	098	229		
-RA	monooxygenase	861	partial [Thermobia domestica]		1372	E-73	4	.95	183	138

Table 3.1	0. Cont	tinued.
-----------	---------	---------

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
се		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
Th_d_00	Lytic		gi 1355978318 gb PRD27424.1 hy				73.6	165		
044687	polysaccharide		pothetical protein NCL1_35386		PRD2	2.93	842	.62		
-RA	monooxygenase	600	[Nephila clavipes]	Spider	7424	E-50	1	2	133	98
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				78.2	275		
068347	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.95	383	.01		
-RA	monooxygenase	862	partial [Thermobia domestica]		1373	E-91	4	8	193	151
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					447		
071731	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.6E		.20		
-RA	monooxygenase	821	partial [Thermobia domestica]		1372	-159	100	3	213	213
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L				75.9	235		
095665	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	5.17	776	.72		
-RA	monooxygenase	615	partial [Thermobia domestica]		1372	E-77	5	8	179	136
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L					204		
085144	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.56	74.3	.52		
-RA	monooxygenase	480	partial [Thermobia domestica]		1373	E-65	75	7	160	119
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					166		
071020	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	7.67	71.9	.77		
-RA	monooxygenase	837	partial [Thermobia domestica]		1372	E-49	697	7	132	95
Th_d_00	Lytic		gi 1349059417 emb SIW61373.1 L				75.7	217		
037467	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	1.23	575	.62		
-RA	monooxygenase	1077	partial [Thermobia domestica]		1373	E-67	8	4	165	125

Sequen	Sequence desc.	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Alig	Pos
ce		ence			ACC	Valu	larit	Sco	nme	itiv
name		lengt				е	У	re	nt	es
		h							leng	
									th	
Th_d_00	Lytic		gi 1349059415 emb SIW61372.1 L					230		
037466	polysaccharide		ytic polysaccharide monooxygenase,		SIW6	3.16		.33		
-RA	monooxygenase	567	partial [Thermobia domestica]		1372	E-75	75	5	180	135

Seque	Sequence	Sequen	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
nce name	desc.	ce length			ACC	Valu e	arity	Sco re	ment lengt h	ives
Lep_00 02016 0-RA	endoglucanas e 13-like	513	gi 1330895262 gb PNF24409 .1 hypothetical protein B7P43_G09674, partial [Cryptotermes secundus]	Blattod ea	PNF24 409	2.62E -22	86.88 525	96. 285 7	61	53
Lep_00 11521 3-RA	endoglucanas e 7-like	561	gi 1228018665 ref XP_02194 1322.1 endoglucanase 7-like, partial [Zootermopsis nevadensis]	Blattod ea	XP_021 94132 2	3.32E -82	83.01 887	247 .28 4	159	132
Lep_00 01466 5-RA	putative endo-beta- 1,4-glucanase NkEG3	618	gi 197691949 dbj BAG70026. 1 endo-beta-1,4-glucanase, partial [Salganea esakii]	Blattod ea	BAG70 026	8.42E -17	80	83. 188 9	60	48
Lep_00 07146 1-RA	Endoglucanas e A	692	gi 1228018667 ref XP_02194 1324.1 uncharacterized protein LOC110840536 [Zootermopsis nevadensis]	Blattod ea	XP_021 94132 4	4.2E- 107	84.40 367	321 .62 7	218	184
Lep_00 04506 7-RA	putative endo-beta- 1,4-glucanase OfEG3	726	gi 24940551 emb CAD54729. 1 beta-1,4-endoglucanase [Mastotermes darwiniensis]	Blattod ea	CAD54 729	1.47E -69	70.61 856	225 .71 3	194	137
Lep_00 01631 0-RA	endoglucanas e 7-like	519	gi 1228018665 ref XP_02194 1322.1 endoglucanase 7-like, partial [Zootermopsis nevadensis]	Blattod ea	XP_021 94132 2	2.5E- 76	82.19 178	231 .87 6	146	120

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

Sequence	Sequenc	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Alig	Po
name	e desc.	ence				Val	arity	Sco	nme	siti
		leng				ue		re	nt	ves
		th							leng	
									th	
			gi 157313367 gb ABV32557							
			.1 cellulase [Teleogryllus		ABV3255					
			emma]gi 167859905 gb AC		7,	1.0				
Lep_000163	Endoglu		A04897.1 cellulase		ACA0489	8E-	69.7	157.		
12-RA	canase 1	837	[Teleogryllus emma]	Orthoptera	7	42	479	532	119	83
	AF2205									
	93_1bet		gi 7546876 gb AAF63724.1							
	a-1,4-		AF220593_1beta-1,4-			9.2				
Lep_001058	glucanas		glucanase 1, partial			E-		207.		12
74-RA	e 1	522	[Mastotermes darwiniensis]	Blattodea	AAF63724	66	77.5	608	160	4
			gi 985388348 ref XP_01537							
	endogluc		8728.1 PREDICTED:			1.5				
Lep_001129	anase		endoglucanase E-4-like		XP_01537	6E-	86.5	94.7		
04-RA	11-like	528	[Diuraphis noxia]	Hemiptera	8728	22	3846	449	52	45
	endo-		gi 1228018667 ref XP_0219							
	beta-1,4-		41324.1 uncharacterized			3.1				
Lep_000523	glucanas		protein LOC110840536		XP_02194	E-	63.1	310.		20
26-RA	e 1	1168	[Zootermopsis nevadensis]	Blattodea	1324	100	0976	457	328	7

Sequence	Sequenc	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Po
name	e desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
			gi 1339056265 ref XP_02371659							
			6.1 uncharacterized protein							
			LOC111869358 [Cryptotermes							
			secundus]gi 1339056267 ref XP_							
			023716597.1 uncharacterized							
	Endoglu		protein LOC111869358							
	canase		[Cryptotermes		XP_02371					
	E-4		secundus]gi 1330895261 gb PNF		6596,					
	precurso		24408.1 hypothetical protein		XP_02371	1.4		15		
Lep_000130	r,		B7P43_G09674 [Cryptotermes	Blattod	6597,	2E-	78.4	9.8		11
83-RA	putative	1834	secundus]	ea	PNF24408	40	7222	44	144	3
	endogluc		gi 646711640 gb KDR16731.1 E			5.9		26		
Lep_000130	anase E-		ndoglucanase F [Zootermopsis	Blattod	KDR1673	4E-	73.2	8.4		16
82-RA	4-like	1314	nevadensis]	ea	1	83	1429	7	224	4
	endogluc		gi 380452610 gb AFD33365.1 en			1.1		90.		
Lep_001250	anase E-		doglucanase [Macrotermes	Blattod	AFD3336	6E-	83.3	89		
33-RA	4-like	300	barneyi]	ea	5	20	3333	29	60	50
	putative									
	endo-		gi 44885838 dbj BAD12008.1 pu							
	beta-1,4-		tative endo-beta-1,4-glucanase			1.6		24		
Lep_000149	glucanas		OfEG1, partial [Odontotermes	Blattod	BAD1200	5E-	70.4	6.8		16
82-RA	e OfEG1	819	formosanus]	ea	8	78	8458	99	227	0

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Alig	Po
e name	desc.	ence				Val	arity	Sco	nme	siti
		leng				ue		re	nt	ves
		th							leng	
	AE220E06								un	
	$\frac{AF220590}{1 \text{ hota } 1.4}$		gil00060271gb144E00504114E22			22				
Lop 0005	iDeta-1,4-		0506 1 hota 1 4 ondoglucanaso 1	Plattad		2.3 2E	70 /	125		
1684-BA		291	[Panesthia cribrata]	Diattou	AAF80584	211-	9462	135. 961	93	73
1004-104		274	ratestina cribiatagil7546876lgblAAF63724 11AF22	ca	AAPOUJUT	57	7402	701	,,	75
	AF220593		0593 1beta-1 4-glucanase 1			43				
$I_{en} 0009$	$1het_{2} 1 4$		nartial [Mastotermes	Rlattod		9F.	75 1	253		15
5655-RA	glucanase 1	627	darwiniensisl	ea	AAF63724	83	2438	233. 447	201	1
5055 101	AF220597	027		cu	1111 05721	00	2150	117	201	1
	1beta-1.4-		gil197691959ldbilBAG70031.1le			8.3				
Lep 0010	endoglucan		ndo-beta-1.4-glucanase, partial	Blattod	BAG7003	5E-	73.5	186.		11
4311-RA	ase 2	480	[Panesthia angustipennis]	ea	1	57	4839	037	155	4
	Endoglucan					_				
	ase E-4					4.6				
Lep_0014	precursor,		gi 1022772340 gb KZS16514.1 E	Crustac		3E-	67.5	85.5		
2934-RA	putative	225	ndoglucanase [Daphnia magna]	ea	KZS16514	19	6757	001	74	50
			gi 44885846 dbj BAD12012.1 pu							
	endo-beta-		tative endo-beta-1,4-glucanase			3.5				
Lep_0009	1,4-		SmEG1, partial [Sinocapritermes	Blattod	BAD1201	1E-	67.0	168.		11
5117-RA	glucanase 1	531	mushae]	ea	2	49	4545	318	176	8
			gi 646711640 gb KDR16731.1 E			3.3				
Lep_0010	Endoglucan		ndoglucanase F [Zootermopsis	Blattod	KDR1673	5E-	87.3	106.		
0398-RA	ase F	504	nevadensis]	ea	1	25	2394	301	71	62

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	Е-	Simil	Bit-	Alig	Po
e name	desc.	ence				Val	arity	Sco	nme	siti
		leng				ue		re	nt	ves
		th							leng	
									th	
	Endoglucan									
	ase E-4					9.0				
Lep_0005	precursor,		gi 1022772340 gb KZS16514.1 E	Crustac		9E-	74.2	99.7		
3282-RA	putative	279	ndoglucanase [Daphnia magna]	ea	KZS16514	24	4242	525	66	49
	AF220597_									
	1beta-1,4-		gi 197691959 dbj BAG70031.1 e			1.3				
Lep_0002	endoglucan		ndo-beta-1,4-glucanase, partial	Blattod	BAG7003	5E-	64.7	190.		12
9521-RA	ase 2	1666	[Panesthia angustipennis]	ea	1	53	9592	66	196	7
						1.5				
Lep_0010	endoglucan		gi 1078570857 gb A0V94253.1 c	Embio	A0V9425	E-	79.0	103.		
7688-RA	ase 15-like	399	ellulase [Antipaluria urichi]	ptera	3	24	3226	219	62	49
			gi 985388348 ref XP_015378728							
Lep_0000	endoglucan		.1 PREDICTED: endoglucanase E-	Hemipt	XP_01537	2E-	79.0	93.9		
3639-RA	ase E-4-like	282	4-like [Diuraphis noxia]	era	8728	23	3226	745	62	49
			gi 1339092712 ref XP_02370756							
			9.1 1,4-alpha-glucan-branching							
			enzyme [Cryptotermes							
	1,4-alpha-		secundus]gi 1330914862 gb PNF							
	glucan-		33674.1 1,4-alpha-glucan-		XP_02370	1.1				
Lep_0003	branching		branching enzyme [Cryptotermes	Blattod	7569,	9E-	87.4	244.		12
2008-RA	enzyme	876	secundus	ea	PNF33674	73	1259	588	143	5

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Po
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
			gi 1339092712 ref XP_02370756							
			9.1 1,4-alpha-glucan-branching							
			enzyme [Cryptotermes							
	1,4-alpha-		secundus]gi 1330914862 gb PNF							
	glucan-		33674.1 1,4-alpha-glucan-		XP_02370			54		
Lep_0003	branching		branching enzyme [Cryptotermes	Blattod	7569,		89.7	1.5		28
2826-RA	enzyme	1779	secundus]	ea	PNF33674	0	7636	76	313	1
			gi 44885830 dbj BAD12004.1 pu							
	endo-beta-		tative endo-beta-1,4-glucanase			2.4		15		
Lep_0002	1,4-		HsEG4, partial [Hodotermopsis	Blattod	BAD1200	1E-	82.6	5.2		
6154-RA	glucanase 1	456	sjostedti]	ea	4	44	087	21	115	95
	endo-beta-		gi 521313323 gb AGP76424.1 en			6.0		75.		
Lep_0002	1,4-		do-beta-1,4-glucanase 1	Blattod	AGP7642	6E-	79.6	48		
6155-RA	glucanase 1	339	[Pericapritermes sp. PpEG1]	ea	4	15	2963	5	54	43
	AF220597_									
	1beta-1,4-		gi 8886829 gb AAF80585.1 AF22			1.6		15		
Lep_0003	endoglucan		0597_1beta-1,4-endoglucanase 2	Blattod		9E-	55.5	7.5		11
8056-RA	ase 2	693	[Panesthia cribrata]	ea	AAF80585	43	5556	32	207	5
			gi 646700778 gb KDR10771.1 hy							
	AF220594_		pothetical protein L798_15338,			1.6		17		
Lep_0000	1beta-1,4-		partial [Zootermopsis	Blattod	KDR1077	1E-	86.8	6.4		
6772-RA	glucanase 2	372	nevadensis]	ea	1	53	4211	07	114	99

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Po
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
			gi 1000740642 ref XP_01559292							
			7.1 PREDICTED: endoglucanase							
			E-4-like [Cephus							
			cinctus]gi 1000740644 ref XP_01							
			5592928.1 PREDICTED:		XP_01559					
			endoglucanase E-4-like [Cephus		2927,					
			cinctus]gi 1000740646 ref XP_01		XP_01559					
			5592929.1 PREDICTED:		2928,	4.6		11		
Lep_0000	endoglucan		endoglucanase E-4-like [Cephus	Hymen	XP_01559	8E-	74.7	6.7		
8343-RA	ase E-4-like	291	cinctus]	optera	2929	30	4747	01	99	74
	Endoglucan									
	ase E-4		gi 985425232 ref XP_015378087			2.4		14		
Lep_0000	precursor,		.1 PREDICTED: endoglucanase 4-	Hemipt	XP_01537	6E-	74.0	4.0		10
8346-RA	putative	519	like [Diuraphis noxia]	era	8087	40	7407	5	135	0
			gi 985388348 ref XP_015378728			1.3		94.		
Lep_0000	endoglucan		.1 PREDICTED: endoglucanase E-	Hemipt	XP_01537	E-	70.8	35		
8347-RA	ase E-4-like	720	4-like [Diuraphis noxia]	era	8728	21	3333	97	72	51
			gi 992052372 gb AMH40395.1 gl			6.7		69.		
Lep_0000	Endoglucan		ycoside hydrolase family 9	Phasm	AMH4039	E-		70		
8344-RA	ase E-4	360	[Timema cristinae]	atodea	5	13	62.5	7	80	50
			gi 1228018667 ref XP_02194132							
	glycoside		4.1 uncharacterized protein					52		
Lep_0004	hydrolase		LOC110840536 [Zootermopsis	Blattod	XP_02194		79.3	8.0		30
4891-RA	family 9	1194	nevadensis]	ea	1324	0	6508	94	378	0

Table 3.11. Continued	l.
-----------------------	----

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Po
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
			gi 1228018667 ref XP_02194132							
	glycoside		4.1 uncharacterized protein					46		
Lep_0002	hydrolase		LOC110840536 [Zootermopsis	Blattod	XP_02194	3E-	78.1	3.3		27
0161-RA	family 9	1350	nevadensis]	ea	1324	159	5126	81	357	9
	glycoside		gi 992052372 gb AMH40395.1 gl			2.2		22		
Lep_0005	hydrolase		ycoside hydrolase family 9	Phasm	AMH4039	5E-	56.0	2.2		15
4377-RA	family 9	1059	[Timema cristinae]	atodea	5	66	7143	46	280	7
	glycoside		gi 985388348 ref XP_015378728			6.6		10		
Lep_0001	hydrolase		.1 PREDICTED: endoglucanase E-	Hemipt	XP_01537	E-	67.7	2.4		
4666-RA	family 9	463	4-like [Diuraphis noxia]	era	8728	26	7778	49	90	61
	glycoside		gi 992051866 gb AMH40374.1 gl			3.6		25		
Lep_0007	hydrolase		ycoside hydrolase family 9	Phasm	AMH4037	5E-	75.5	4.6		16
1754-RA	family 9	838	[Peruphasma schultei]	atodea	4	80	8685	03	213	1
	glycoside		gi 13095576 gb AAK12339.1 cell			9.3		19		
Lep_0001	hydrolase		ulase [Coptotermes	Blattod	AAK1233	6E-	71.0	4.1		14
6311-RA	family 9	834	acinaciformis]	ea	9	57	7843	26	204	5
			gi 1330895262 gb PNF24409.1 h							
	glycoside		ypothetical protein			1.6		13		
Lep_0010	hydrolase		B7P43_G09674, partial	Blattod		5E-	78.4	4.8		
0613-RA	family 9	438	[Cryptotermes secundus]	ea	PNF24409	37	9462	06	93	73
	glycoside		gi 992051866 gb AMH40374.1 gl			4.5		44		
Lep_0003	hydrolase		ycoside hydrolase family 9	Phasm	AMH4037	E-	63.3	2.5		31
4744-RA	family 9	1742	[Peruphasma schultei]	atodea	4	149	4661	8	502	8

Table 3.11. Continued	l.
-----------------------	----

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Po
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
	glycoside					1.3		20		
Lep_0001	hydrolase		gi 1078570851 gb A0V94250.1 c	Embio	A0V9425	7E-	71.4	9.5		12
4981-RA	family 9	1605	ellulase [Antipaluria urichi]	ptera	0	59	2857	34	168	0
			gi 1080053223 ref XP_01856819							
			6.1 uncharacterized protein							
			LOC108908596 isoform X3							
			[Anoplophora							
			glabripennis]gi 1325348755 ref		XP_01856					
	glycoside		XP_023310347.1 uncharacterized		8196,	1.4		10		
Lep_0001	hydrolase		protein LOC108908596 isoform	Coleop	XP_02331	8E-	59.2	2.4		
4983-RA	family 9	558	X4 [Anoplophora glabripennis]	tera	0347	23	9204	49	113	67
			gi 1228018667 ref XP_02194132							
	glycoside		4.1 uncharacterized protein					31		
Lep_0007	hydrolase		LOC110840536 [Zootermopsis	Blattod	XP_02194	1E-	83.4	7.7		18
1132-RA	family 9	685	nevadensis]	ea	1324	105	8624	75	218	2
	glycoside		gi 992052372 gb AMH40395.1 gl			4.3		10		
Lep_0005	hydrolase		ycoside hydrolase family 9	Phasm	AMH4039	6E-	79.7	2.0		
1683-RA	family 9	244	[Timema cristinae]	atodea	5	25	2973	64	74	59
	glycoside		gi 992052101 gb AMH40383.1 gl			1.4		10		
Lep_0003	hydrolase		ycoside hydrolase family 9	Phasm	AMH4038	5E-	76.4	2.4		
8648-RA	family 9	580	[Ramulus artemis]	atodea	3	23	7059	49	68	52
	glycoside		gi 13095576 gb AAK12339.1 cell			3.0		28		
Lep_0006	hydrolase		ulase [Coptotermes	Blattod	AAK1233	3E-	76.1	5.8		16
0202-RA	family 9	852	acinaciformis]	ea	9	92	9048	04	210	0

Table 3.1	1. Con	tinued.
-----------	--------	---------

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Po
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
	glycoside		gi 992052326 gb AMH40392.1 gl			1.5		51		
Lep_0002	hydrolase		ycoside hydrolase family 9	Phasm	AMH4039	E-	74.0	1.1		31
2398-RA	family 9	1453	[Timema cristinae]	atodea	2	177	1392	46	431	9
			gi 1228018667 ref XP_02194132							
	glycoside		4.1 uncharacterized protein					57		
Lep_0003	hydrolase		LOC110840536 [Zootermopsis	Blattod	XP_02194		79.3	0.8		34
6184-RA	family 9	1518	nevadensis]	ea	1324	0	5035	52	431	2
			gi 1228018667 ref XP_02194132							
	glycoside		4.1 uncharacterized protein					57		
Lep_0002	hydrolase		LOC110840536 [Zootermopsis	Blattod	XP_02194		78.1	8.9		33
6826-RA	family 9	1573	nevadensis]	ea	1324	0	9026	41	431	7
			gi 1330895262 gb PNF24409.1 h							
	glycoside		ypothetical protein			3.9		13		
Lep_0006	hydrolase		B7P43_G09674, partial	Blattod		E-		5.9		
1557-RA	family 9	1095	[Cryptotermes secundus]	ea	PNF24409	35	75	61	100	75
			gi 7546876 gb AAF63724.1 AF22							
	glycoside		0593_1beta-1,4-glucanase 1,			6.1		15		
Lep_0003	hydrolase		partial [Mastotermes	Blattod		5E-	71.6	9.8		
4832-RA	family 9	537	darwiniensis]	ea	AAF63724	47	5354	44	127	91
	glycoside		gi 992051866 gb AMH40374.1 gl			4.1		25		
Lep_0007	hydrolase		ycoside hydrolase family 9	Phasm	AMH4037	7E-	70.1	8.4		16
8550-RA	family 9	753	[Peruphasma schultei]	atodea	4	82	2987	55	231	2

Table	3.11.	Continued.

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	Е-	Simil	Bit	Align	Po
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
	glycoside		gi 992051866 gb AMH40374.1 gl			9.5		24		
Lep_0008	hydrolase		ycoside hydrolase family 9	Phasm	AMH4037	6E-	72.1	9.2		16
1122-RA	family 9	693	[Peruphasma schultei]	atodea	4	79	2389	1	226	3
	glycoside		gi 992051866 gb AMH40374.1 gl			1.9		10		
Lep_0013	hydrolase		ycoside hydrolase family 9	Phasm	AMH4037	7E-	80.9	3.6		
5466-RA	family 9	285	[Peruphasma schultei]	atodea	4	25	5238	05	63	51
	glycoside					1.7		95.		
Lep_0012	hydrolase		gi 1022761055 gb KZS07093.1 E	Crustac		7E-	75.6	13		
7920-RA	family 9	231	ndoglucanase [Daphnia magna]	ea	KZS07093	22	7568	01	74	56
			gi 1330895262 gb PNF24409.1 h							
	glycoside		ypothetical protein			4.5		17		
Lep_0009	hydrolase		B7P43_G09674, partial	Blattod		E-	74.4	2.9		11
6399-RA	family 9	1155	[Cryptotermes secundus]	ea	PNF24409	49	9664	4	149	1
			gi 1228018667 ref XP_02194132							
	glycoside		4.1 uncharacterized protein			8.2		37		
Lep_0000	hydrolase		LOC110840536 [Zootermopsis	Blattod	XP_02194	E-	75.4	3.2		23
3635-RA	family 9	975	nevadensis]	ea	1324	126	8387	44	310	4
	glycoside					7.5		91.		
Lep_0000	hydrolase		gi 1022772340 gb KZS16514.1 E	Crustac		2E-	77.9	66		
3637-RA	family 9	312	ndoglucanase [Daphnia magna]	ea	KZS16514	21	661	33	59	46
	glycoside		gi 695941540 gb AIT38270.1 en	Arthro		1.3		51.		
Lep_0000	hydrolase		do-beta-1,4-glucanase, partial	poda-		9E-	82.8	98		
3640-RA	family 9	234	[Pseudohelice subquadrata]	crab	AIT38270	08	5714	78	35	29

Table 3.1	1. Con	tinued.
-----------	--------	---------

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Po
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
	glycoside		gi 992051639 gb AMH40360.1 gl			6.2		22		
Lep_0002	hydrolase		ycoside hydrolase family 9	Phasm	AMH4036	4E-		2.2		12
6156-RA	family 9	1451	[Aretaon asperrimus]	atodea	0	65	75	46	172	9
			gi 1330895262 gb PNF24409.1 h							
	glycoside		ypothetical protein			1.4		27		
Lep_0003	hydrolase		B7P43_G09674, partial	Blattod		7E-	75.9	0.7		16
8055-RA	family 9	852	[Cryptotermes secundus]	ea	PNF24409	88	0909	81	220	7
			gi 1330895262 gb PNF24409.1 h							
	glycoside		ypothetical protein			2.1		17		
Lep_0009	hydrolase		B7P43_G09674, partial	Blattod		3E-	72.9	6.4		11
4505-RA	family 9	912	[Cryptotermes secundus]	ea	PNF24409	51	0323	07	155	3
	glycoside		gi 24940551 emb CAD54729.1 b					53		
Lep_0000	hydrolase		eta-1,4-endoglucanase	Blattod	CAD5472		72.2	7.3		34
6775-RA	family 9	1508	[Mastotermes darwiniensis]	ea	9	0	4576	39	472	1
	glycoside		gi 992051639 gb AMH40360.1 gl			1.5		62.		
Lep_0000	hydrolase		ycoside hydrolase family 9	Phasm	AMH4036	4E-	84.6	77		
6773-RA	family 9	441	[Aretaon asperrimus]	atodea	0	18	1538	34	39	33
	glycoside		gi 992051683 gb AMH40363.1 gl			1.9		98.		
Lep_0000	hydrolase		ycoside hydrolase family 9	Phasm	AMH4036	9E-	82.6	59		
6774-RA	family 9	1143	[Extatosoma tiaratum]	atodea	3	20	9231	69	52	43
	glycoside		gi 24940553 emb CAD54730.1 b			5.1		32		
Lep_0004	hydrolase		eta-1,4-endoglucanase	Blattod	CAD5473	E-	76.5	8.5		20
2852-RA	family 9	834	[Mastotermes darwiniensis]	ea	0	109	7993	61	269	6

Table 3.11. Continue	J.
----------------------	----

Sequenc	Sequence	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit	Align	Ро
e name	desc.	ence				Val	arity	-	ment	siti
		leng				ue		Sco	lengt	ves
		th						re	h	
	glycoside					5.0		68.		
Lep_0004	hydrolase		gi 1078570851 gb A0V94250.1 c	Crustac	AOV9425	6E-	66.6	55		
2851-RA	family 9	444	ellulase [Antipaluria urichi]	ea	0	12	6667	14	57	38

Seque	Sequence	Sequen	Hit desc.	Order	Hit ACC	E-Value	Simil	Bit-	Align	Pos
nce	desc.	ce					arity	Scor	ment	itiv
name		length						е	length	es
			gi 1339080613 ref XP_0							
			23728054.1 glucosidase							
			2 subunit beta							
			[Cryptotermes							
			secundus]gi 133092781							
			6 gb PNF40060.1 Glucos		XP_02372					
Lep_00			idase 2 subunit beta		8054,					
09587	glucosidase 2		[Cryptotermes	Blatto	PNF4006	2.28E-	84.61	139.4		
8-RA	subunit beta	633	secundus]	dea	0	36	538	28	78	66
			gi 1339077375 ref XP_0							
			23726071.1 glucosidase							
Lep_00			2 subunit beta isoform							
05546	Glucosidase 2		X1 [Cryptotermes	Blatto	XP_02372	4.64E-	51.63	105.9		
2-RA	subunit beta	708	secundus]	dea	6071	26	399	16	153	79
			gi 1228015572 ref XP_0							
			21939758.1 glucosidase							
Lep_00	glucosidase 2		2 subunit beta							
05546	subunit beta-		[Zootermopsis	Blatto	XP_02193	5.55E-	80.64	51.98		
0-RA	like	213	nevadensis]	dea	9758	08	516	78	31	25
	protein-		gi 512925622 ref XP_00							
	glucosylgalac		4930884.1 protein-							
Lep_00	tosylhydroxy		glucosylgalactosylhydro							
09887	lysine		xylysine glucosidase	Lepid	XP_00493	3.41E-	60.83	103.6		
5-RA	glucosidase	669	[Bombyx mori]	optera	0884	23	333	05	120	73

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

Table 3.12. Continue	d.
----------------------	----

Sequenc	Sequence	Sequ	Hit desc.	0r	Hit	E-	Simi	Bit-	Align	Pos
e name	desc.	ence		der	ACC	Valu	larit	Scor	ment	itiv
		lengt				e	У	e	lengt	es
		h							h	
			gi 303324839 pdb 3AHZ AChain A,							
			Crystal Structure Of Beta-Glucosidase							
			From Termite Neotermes Koshunensis							
	AChain A,		In Complex With							
	Crystal		Trisgi 393715252 pdb 3VIF AChain A,							
	Structure		Crystal Structure Of Beta-glucosidase							
	Of Beta-		From Termite Neotermes Koshunensis							
	glucosidas		In Complex With							
	e From		Gluconolactonegi 393715253 pdb 3VIG							
	Termite		AChain A, Crystal Structure Of Beta-							
	Neoterme		glucosidase From Termite Neotermes		3AHZ					
	S		Koshunensis In Complex With 1-		_A,					
	Koshunen		deoxynojirimycingi 393715254 pdb 3V		3VIF_					
	sis In		IH AChain A, Crystal Structure Of Beta-		А,					
	Complex		glucosidase From Termite Neotermes		3VIG_					
	With A		Koshunensis In Complex With		А,					
	New		Glycerolgi 393715255 pdb 3VII AChain		3VIH					
Lep_000	Glucopyra		A, Crystal Structure Of Beta-glucosidase	Blat	_A,		75.5			
52604-	nosidic		From Termite Neotermes Koshunensis	tod	3VII_	1.6E-	244	365.9		
RA	Product	957	In Complex With Bis-tris	ea	А	122	8	25	286	216

Sequenc	Sequenc	Seque	Hit desc.	0r	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	nce		der	ACC	Valu	larit	Scor	ment	itiv
		length				е	У	е	lengt	es
									h	
	AChain									
	А,									
	Crystal									
	Structur									
	e Of									
	Beta-									
	glucosid									
	ase From									
	Termite									
	Neoterm		gi 1339068038 ref XP_023721112.1 m		XP_0					
	es		yrosinase 1-like [Cryptotermes		2372					
	Koshune		secundus]gi 1339068040 ref XP_02372		1112,					
	nsis In		1113.1 myrosinase 1-like		XP_0					
	Complex		[Cryptotermes		2372					
	With A		secundus]gi 1330885855 gb PNF1965		1113,					
	New		3.1 Myrosinase 1 [Cryptotermes		PNF1					
Lep_000	Glucopyr		secundus]gi 1330885856 gb PNF1965	Blat	9653,		62.9			
36380-	anosidic		4.1 Myrosinase 1 [Cryptotermes	tod	PNF1	6.7E-	526	328.9		
RA	Product	1194	secundus]	ea	9654	106	5	46	359	226

Table 3.12. Continued.

Sequenc	Sequenc	Seque	Hit desc.	0r	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	nce		der	ACC	Valu	larit	Scor	ment	itiv
		length				е	У	е	lengt	es
									h	
			gi 303324839 pdb 3AHZ AChain A,							
	AChain		Crystal Structure Of Beta-Glucosidase							
	А,		From Termite Neotermes Koshunensis							
	Crystal		In Complex With							
	Structur		Trisgi 393715252 pdb 3VIF AChain A,							
	e Of		Crystal Structure Of Beta-glucosidase							
	Beta-		From Termite Neotermes Koshunensis							
	glucosid		In Complex With							
	ase From		Gluconolactonegi 393715253 pdb 3VIG							
	Termite		AChain A, Crystal Structure Of Beta-							
	Neoterm		glucosidase From Termite Neotermes		3AHZ					
	es		Koshunensis In Complex With 1-		_A,					
	Koshune		deoxynojirimycingi 393715254 pdb 3V		3VIF_					
	nsis In		IH AChain A, Crystal Structure Of Beta-		А,					
	Complex		glucosidase From Termite Neotermes		3VIG_					
	With A		Koshunensis In Complex With		А,					
	New		Glycerolgi 393715255 pdb 3VII AChain		3VIH					
Lep_000	Glucopyr		A, Crystal Structure Of Beta-glucosidase	Blat	_A,		71.8			
04872-	anosidic		From Termite Neotermes Koshunensis	tod	3VII_	9E-	849	368.6		
RA	Product	1844	In Complex With Bis-tris	ea	А	147	8	22	313	225

Table 3.12. (Continued.
---------------	------------

Sequenc	Sequenc	Seque	Hit desc.	Or	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	nce length		aer	ALL	e valu	y	Scor e	ment lengt h	es
	AChain									
	А,									
	Crystal									
	Structur									
	e Of									
	Beta-									
	glucosid									
	ase From									
	Termite									
	Neoterm									
	es									
	Koshune									
	nsis In									
	Complex									
	With A		gi 1339073035 ref XP_023723774.1 m		XP_0					
	New		yrosinase 1-like [Cryptotermes		2372					
Lep_000	Glucopyr		secundus]gi 1330878957 gb PNF1684	Blat	3774,		78.8			
04873-	anosidic		1.1 Myrosinase 1 [Cryptotermes	tod	PNF1	2.43E	135	174.8		
RA	Product	375	secundus]	ea	6841	-51	6	66	118	93

Table 3.12. Continued

Sequenc	Sequenc	Seque	Hit desc.	0r	Hit ACC	E-	Simil	Bit-	Alig	Pos
e name	e desc.	nce		der		Val	arity	Scor	nme	itiv
		length				ue		е	nt	es
									leng	
									th	
			gi 303324839 pdb 3AHZ AChain A,							
			Crystal Structure Of Beta-Glucosidase							
			From Termite Neotermes							
	AChain		Koshunensis In Complex With							
	А,		Trisgi 393715252 pdb 3VIF AChain							
	Crystal		A, Crystal Structure Of Beta-							
	Structur		glucosidase From Termite Neotermes							
	e Of		Koshunensis In Complex With							
	Beta-		Gluconolactonegi 393715253 pdb 3V							
	glucosid		IG AChain A, Crystal Structure Of							
	ase From		Beta-glucosidase From Termite							
	Termite		Neotermes Koshunensis In Complex							
	Neoterm		With 1-							
	es		deoxynojirimycingi 393715254 pdb							
	Koshune		3VIH AChain A, Crystal Structure Of							
	nsis In		Beta-glucosidase From Termite							
	Complex		Neotermes Koshunensis In Complex							
	With A		With		3AHZ_A,					
	New		Glycerolgi 393715255 pdb 3VII ACha		3VIF_A,					
Lep_000	Glucopyr		in A, Crystal Structure Of Beta-	Blat	3VIG_A,	8.0				
21305-	anosidic		glucosidase From Termite Neotermes	tod	3VIH_A,	7E-	73.30	290.		
RA	Product	843	Koshunensis In Complex With Bis-tris	ea	3VII_A	94	508	812	236	173

Table 3.12. Continue	d.
----------------------	----

Sequen	Sequence desc.	Seque	Hit desc.	0r	Hit ACC	E-	Simil	Bit-	Alig	Pos
се		nce		der		Val	arity	Scor	nme	itiv
name		lengt				ue		е	nt	es
		h							leng	
									th	
			gi 1339080613 ref XP_023728							
			054.1 glucosidase 2 subunit							
			beta [Cryptotermes							
			secundus]gi 1330927816 gb P							
Lep_00			NF40060.1 Glucosidase 2	Blat	XP_02372	1.6				
016622	glucosidase 2		subunit beta [Cryptotermes	tod	8054,	6E-	66.35	108.		
-RA	subunit beta	831	secundus]	ea	PNF40060	24	945	612	217	144
			gi 1228013272 ref XP_021938							
Lep_00			600.1 glucosidase 2 subunit	Blat		1.0				
016621	glucosidase 2		beta [Zootermopsis	tod	XP_02193	4E-	80.95	85.1		
-RA	subunit beta	459	nevadensis]	ea	8600	17	238	149	63	51
			gi 1339080613 ref XP_023728							
			054.1 glucosidase 2 subunit							
			beta [Cryptotermes							
			secundus]gi 1330927816 gb P							
Lep_00			NF40060.1 Glucosidase 2	Blat	XP_02372	1.1				
016623	glucosidase 2		subunit beta [Cryptotermes	tod	8054,	E-	60.71	135.		
-RA	subunit beta	678	secundus]	ea	PNF40060	34	429	191	140	85

Table 3.12. Continued

Sequen	Sequence desc.	Seque	Hit desc.	0r	Hit ACC	Е-	Simil	Bit-	Alig	Pos
се		nce		der		Val	arity	Scor	nme	itiv
name		lengt				ue		е	nt	es
		h							leng	
									th	
	AChain A, Crystal									
	Structure Of Beta-									
	glucosidase From									
	Termite									
	Neotermes									
	Koshunensis In									
	Complex With A		gi 1339073027 ref XP_023723							
Lep_00	New		770.1 myrosinase 1-like	Blat		3.2				
061060	Glucopyranosidic		isoform X1 [Cryptotermes	tod	XP_02372	8E-		262.		
-RA	Product	926	secundus]	ea	3770	82	67.5	692	240	162
			gi 393715257 pdb 3VIK AChai							
	AChain A, Crystal		n A, Crystal Structure Of Beta-							
	Structure Of Beta-		glucosidase From Termite							
	glucosidase From		Neotermes Koshunensis In							
	Termite		Complex With							
	Neotermes		Cellobiosegi 393715258 pdb 3							
	Koshunensis In		VIL AChain A, Crystal							
	Complex With A		Structure Of Beta-glucosidase							
Lep_00	New		From Termite Neotermes	Blat		5.8				
022254	Glucopyranosidic		Koshunensis In Complex With	tod	3VIK_A,	1E-	83.90	135.		
-RA	Product	291	Salicin	ea	3VIL_A	37	805	191	87	73

Table 3.12. (Continued.
---------------	------------

Sequen ce name	Sequence desc.	Seque nce lengt h	Hit desc.	Order	Hit ACC	E- Val ue	Simil arity	Bit- Scor e	Alig nme nt leng th	Pos itiv es
Lep_00	AChain A, Crystal Structure Of Beta- glucosidase From Termite Neotermes Koshunensis In Complex With A New		gi 1339073029 ref XP_0 23723771.1 myrosinase 1-like isoform X2		VD 02272	1.6	(2.72)	200		
004127 -RA	Glucopyranosidic Product	837	[Cryptotermes secundus]	Blattodea	XP_02372 3771	1E- 89	62.72 401	280. 411	279	175

Table 3.12. Continued

Sequen	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Alig	Pos
ce	e desc.	nce			ACC	Val	arity	Scor	nme	itiv
name		length				ue		е	nt	es
									leng	
									th	
			gi 303324839 pdb 3AHZ AChain A,							
			Crystal Structure Of Beta-Glucosidase							
			From Termite Neotermes							
			Koshunensis In Complex With							
			Trisgi 393715252 pdb 3VIF AChain							
	AChain A,		A, Crystal Structure Of Beta-							
	Crystal		glucosidase From Termite Neotermes							
	Structure		Koshunensis In Complex With							
	Of Beta-		Gluconolactonegi 393715253 pdb 3V							
	glucosida		IG AChain A, Crystal Structure Of							
	se From		Beta-glucosidase From Termite							
	Termite		Neotermes Koshunensis In Complex							
	Neoterme		With 1-							
	S		deoxynojirimycingi 393715254 pdb		3AHZ_					
	Koshunen		3VIH AChain A, Crystal Structure Of		А,					
	sis In		Beta-glucosidase From Termite		3VIF_					
	Complex		Neotermes Koshunensis In Complex		А,					
	With A		With		3VIG_					
	New		Glycerolgi 393715255 pdb 3VII ACha		А,					
Lep_00	Glucopyr		in A, Crystal Structure Of Beta-		3VIH_	2.7				
004131	anosidic		glucosidase From Termite Neotermes	Blattod	А,	E-	63.69	315.		
-RA	Product	1032	Koshunensis In Complex With Bis-tris	ea	3VII_A	102	427	079	314	200

Table 3.12. Continued	
-----------------------	--

Seque	Sequenc	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	e desc.	ue		r		Val	larit	Sco	nme	itiv
name		nce				ue	У	re	nt	es
		len							leng	
		gth							th	
	bile acid									
Lep_00	beta-					1.7	70.5	149		
04550	glucosida		gi 568254064 gb ETN63073.1 bile acid	Dipte	ETN630	5E-	521	.82		
0-RA	se	618	beta-glucosidase [Anopheles darlingi]	ra	73	39	5	8	163	115
			gi 303324839 pdb 3AHZ AChain A, Crystal							
			Structure Of Beta-Glucosidase From Termite							
	AChain A,		Neotermes Koshunensis In Complex With							
	Crystal		Trisgi 393715252 pdb 3VIF AChain A,							
	Structure		Crystal Structure Of Beta-glucosidase From							
	Of Beta-		Termite Neotermes Koshunensis In Complex							
	glucosida		With							
	se From		Gluconolactonegi 393715253 pdb 3VIG ACh							
	Termite		ain A, Crystal Structure Of Beta-glucosidase							
	Neoterme		From Termite Neotermes Koshunensis In							
	S		Complex With 1-							
	Koshunen		deoxynojirimycingi 393715254 pdb 3VIH A							
	sis In		Chain A, Crystal Structure Of Beta-							
	Complex		glucosidase From Termite Neotermes							
	With A		Koshunensis In Complex With		3AHZ_A,					
	New		Glycerolgi 393715255 pdb 3VII AChain A,		3VIF_A,					
Lep_00	Glucopyr		Crystal Structure Of Beta-glucosidase From		3VIG_A,	4.2	66.3	244		
07716	anosidic		Termite Neotermes Koshunensis In Complex	Blatt	3VIH_A,	E-	934	.58		
0-RA	Product	792	With Bis-tris	odea	3VII_A	76	4	8	244	162

Table 3.12. Continued

Seque	Sequenc	Seq	Hit desc.	Orde	Hit ACC	Е-	Simi	Bit-	Alig	Pos
nce	e desc.	ue		r		Val	larit	Sco	nme	itiv
name		nce				ue	У	re	nt	es
		len							leng	
		gth							th	
	glucosida									
Lep_00	se 2		gi 1228013272 ref XP_021938600.1 glucosi			3.4	80.8	78.		
13137	subunit		dase 2 subunit beta [Zootermopsis	Blatt	XP_021	2E-	510	951		
3-RA	beta	165	nevadensis]	odea	938600	17	6	8	47	38
	AChain A,									
	Crystal									
	Structure									
	Of Beta-									
	glucosida									
	se From									
	Termite									
	Neoterme									
	S									
	Koshunen									
	sis In									
	Complex									
	With A									
	New									
Lep_00	Glucopyr					3.4	71.4	275		
07334	anosidic		gi 380452608 gb AFD33364.1 beta-	Blatt	AFD333	1E-	912	.01		
5-RA	Product	897	glucosidase [Macrotermes barneyi]	odea	64	87	3	8	228	163

Seque	Sequenc	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	e desc.	ue		r		Val	larit	Sco	nme	itiv
name		nce				ue	У	re	nt	es
		len							leng	
		gth							th	
			gi 1339080389 ref XP_023727931.1 manno							
			syl-oligosaccharide glucosidase							
			[Cryptotermes							
			secundus]gi 1339080391 ref XP_023727934							
			.1 mannosyl-oligosaccharide glucosidase							
			[Cryptotermes		XP_023					
			secundus]gi 1339080393 ref XP_023727935		727931,					
	mannosyl		.1 mannosyl-oligosaccharide glucosidase		XP_023					
	-		[Cryptotermes		727934,					
	oligosacc		secundus]gi 1330927938 gb PNF40173.1 M		XP_023					
	haride		annosyl-oligosaccharide glucosidase GCS1		727935,					
	glucosida		[Cryptotermes		PNF401					
Lep_00	se		secundus]gi 1330927939 gb PNF40174.1 M		73,		76.9	104		
02084	isoform	244	annosyl-oligosaccharide glucosidase GCS1	Blatt	PNF401		417	7.7		
6-RA	X2	2	[Cryptotermes secundus]	odea	74	0	5	3	824	634
Table 3.12. (Continued.									
---------------	------------									
---------------	------------									

Seque	Sequenc	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	e desc.	ue		r		Val	larit	Sco	nme	itiv
name		nce				ue	У	re	nt	es
		len							leng	
		gth							th	
			gi 303324839 pdb 3AHZ AChain A, Crystal							
			Structure Of Beta-Glucosidase From Termite							
	AChain A,		Neotermes Koshunensis In Complex With							
	Crystal		Trisgi 393715252 pdb 3VIF AChain A,							
	Structure		Crystal Structure Of Beta-glucosidase From							
	Of Beta-		Termite Neotermes Koshunensis In Complex							
	glucosida		With							
	se From		Gluconolactonegi 393715253 pdb 3VIG ACh							
	Termite		ain A, Crystal Structure Of Beta-glucosidase							
	Neoterme		From Termite Neotermes Koshunensis In							
	S		Complex With 1-							
	Koshunen		deoxynojirimycingi 393715254 pdb 3VIH A							
	sis In		Chain A, Crystal Structure Of Beta-							
	Complex		glucosidase From Termite Neotermes							
	With A		Koshunensis In Complex With		3AHZ_A,					
	New		Glycerolgi 393715255 pdb 3VII AChain A,		3VIF_A,					
Lep_00	Glucopyr		Crystal Structure Of Beta-glucosidase From		3VIG_A,	1.5	75.5	365		
05205	anosidic		Termite Neotermes Koshunensis In Complex	Blatt	3VIH_A,	E-	244	.92		
1-RA	Product	954	With Bis-tris	odea	3VII_A	122	8	5	286	216

Seque	Sequenc	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	e desc.	ue		r		Val	larit	Sco	nme	itiv
name		nce				ue	У	re	nt	es
		len							leng	
		gth							th	
			gi 769856746 ref XP_011639981.1 PREDIC							
			TED: lysosomal alpha-glucosidase-like							
			[Pogonomyrmex							
			barbatus]gi 769856748 ref XP_011639982.1							
			PREDICTED: lysosomal alpha-glucosidase-							
			like [Pogonomyrmex		XP_011					
			barbatus]gi 769856750 ref XP_011639983.1		639981,					
			PREDICTED: lysosomal alpha-glucosidase-		XP_011					
			like [Pogonomyrmex		639982,					
			barbatus]gi 769856752 ref XP_011639984.1		XP_011					
			PREDICTED: lysosomal alpha-glucosidase-		639983,					
	lysosomal		like [Pogonomyrmex		XP_011					
Lep_00	alpha-		barbatus]gi 769856754 ref XP_011639985.1	Hym	639984,	6.3	87.8	151		
04125	glucosida	145	PREDICTED: lysosomal alpha-glucosidase-	enopt	XP_011	7E-	048	.75		
2-RA	se-like	8	like [Pogonomyrmex barbatus]	era	639985	37	8	4	82	72
	uncharact									
	erized									
	family 31									
	glucosida									
Lep_00	se		gi 636630780 gb AIA09350.1 alpha-				71.7	665		
08481	KIAA116	185	glucosidase family 31, partial [Periplaneta	Blatt	AIA093		625	.61		
0-RA	1-like	7	americana]	odea	50	0	9	1	556	399

Seque	Sequence	Seq	Hit desc.	Orde	Hit ACC	Е-	Simi	Bit-	Alig	Pos
nce	desc.	uen		r		Val	larit	Sco	nme	itiv
name		се				ue	У	re	nt	es
		leng							leng	
		th							th	
	uncharacte									
	rized family									
Lep_00	31		gi 1316140871 ref XP_023215385.1 unch	Arthr		1.3	75.8	77.		
11960	glucosidase		aracterized protein LOC111618155	opod	XP_023	1E-	064	796		
7-RA	KIAA1161	370	isoform X1 [Centruroides sculpturatus]	а	215385	15	5	2	62	47
	uncharacte									
	rized family									
	31		gi 1067066097 ref XP_018017064.1 PRE							
Lep_00	glucosidase		DICTED: uncharacterized family 31			2.0	55.6			
15999	KIAA1161-	110	glucosidase KIAA1161-like [Hyalella	Crust	XP_018	1E-	149	278		
4-RA	like	7	azteca]	acea	017064	86	7	.1	374	208
			gi 1228004020 ref XP_021933895.1 prob							
			able maltase-glucoamylase 2							
	lysosomal		[Zootermopsis		XP_021					
Lep_00	alpha-		nevadensis]gi 1228004022 ref XP_02193		933895,	5.4		325		
10032	glucosidase	225	3896.1 probable maltase-glucoamylase 2	Blatt	XP_021	E-	50.1	.09		
1-RA	-like	0	[Zootermopsis nevadensis]	odea	933896	96	906	4	787	395
	uncharacte									
	rized family									
	31									
Lep_00	glucosidase		gi 636630780 gb AIA09350.1 alpha-			6.8	82.7	161		
06089	KIAA1161-		glucosidase family 31, partial [Periplaneta	Blatt	AIA093	5E-	586	.38		
9-RA	like	501	americana]	odea	50	45	2	4	116	96

Seque	Sequence	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	desc.	uen		r		Val	larit	Sco	nme	itiv
name		се				ue	У	re	nt	es
		leng							leng	
		th							th	
	uncharacte									
	rized family									
	31		gi 1101398930 ref XP_018910486.1 PRE							
Lep_00	glucosidase		DICTED: uncharacterized family 31			3.4	80.9	88.		
06090	KIAA1161-		glucosidase KIAA1161-like [Bemisia	Hemi	XP_018	E-	523	581		
0-RA	like	405	tabaci]	ptera	910486	19	8	7	63	51
Lep_00	Lysosomal		gi 321476729 gb EFX87689.1 hypothetic			1.3	71.3	325		
06135	alpha-	116	al protein DAPPUDRAFT_312137	Crust	EFX876	E-	395	.86		
0-RA	glucosidase	3	[Daphnia pulex]	acea	89	101	6	5	321	229
	uncharacte									
	rized family									
	31									
Lep_00	glucosidase		gi 636630780 gb AIA09350.1 alpha-					646		
07753	KIAA1161-	204	glucosidase family 31, partial [Periplaneta	Blatt	AIA093		65.6	.73		
8-RA	like	0	americana]	odea	50	0	25	6	608	399
Lep_00	Lysosomal		gi 1067098296 ref XP_018011173.1 PRE			1.7	65.2			
03919	alpha-		DICTED: maltase-glucoamylase, intestinal-	Crust	XP_018	3E-	329	278		
2-RA	glucosidase	861	like [Hyalella azteca]	acea	011173	83	7	.87	279	182
Lep_00	Lysosomal					1.1	73.6	276		
04475	alpha-		gi 1022768804 gb KZS13425.1 Lysosoma	Crust	KZS134	6E-	607	.94		
4-RA	glucosidase	717	l alpha-glucosidase [Daphnia magna]	acea	25	84	1	4	224	165

Seque	Sequence	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	desc.	uen		r		Val	larit	Sco	nme	itiv
name		се				ue	У	re	nt	es
		leng							leng	
		th							th	
	lysosomal		gi 1067072134 ref XP_018019496.1 PRE							
Lep_00	alpha-		DICTED: LOW QUALITY PROTEIN:			1.6	49.6	68.		
04475	glucosidase		probable maltase-glucoamylase 2, partial	Crust	XP_018	1E-	296	551		
5-RA	-like	561	[Hyalella azteca]	acea	019496	11	3	4	135	67
			gi 242019253 ref XP_002430076.1 Neutr							
			al alpha-glucosidase AB precursor,							
	Neutral		putative [Pediculus humanus		XP_002					
Lep_00	alpha-		corporis]gi 212515157 gb EEB17338.1 N	Phthi	430076,	1.3	81.7	203		
08251	glucosidase		eutral alpha-glucosidase AB precursor,	rapte	EEB173	1E-	460	.37		
8-RA	AB	381	putative [Pediculus humanus corporis]	ra	38	60	3	1	126	103
	neutral									
Lep_00	alpha-		gi 321473948 gb EFX84914.1 hypothetic			6.1		100		
08251	glucosidase		al protein DAPPUDRAFT_300811	Crust	EFX849	2E-		.13		
9-RA	AB	270	[Daphnia pulex]	acea	14	24	67.5	8	80	54
	uncharacte									
	rized family									
Lep_00	31		gi 636630780 gb AIA09350.1 alpha-			2.4	57.9	392		
10666	glucosidase	130	glucosidase family 31, partial [Periplaneta	Blatt	AIA093	E-	957	.11		
8-RA	KIAA1161	5	americana]	odea	50	129	4	9	469	272
	lysosomal									
Lep_00	alpha-		gi 1233197498 ref XP_022198397.1 lysos			2.6	54.3	226		
05768	glucosidase	110	omal alpha-glucosidase-like, partial	Hemi	XP_022	6E-	543	.48		
7-RA	-like	1	[Nilaparvata lugens]	ptera	198397	64	5	3	333	181

Seque	Sequence	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	desc.	uen		r		Val	larit	Sco	nme	itiv
name		се				ue	У	re	nt	es
		leng							leng	
		th							th	
	lysosomal									
Lep_00	alpha-		gi 1238867836 ref XP_022249217.1 lysos	Arthr		3.3				
13197	glucosidase		omal alpha-glucosidase-like [Limulus	opod	XP_022	3E-		114		
1-RA	-like	291	polyphemus]	а	249217	32	68	.39	100	68
			gi 1080050544 ref XP_018566730.1 lysos							
	lysosomal		omal alpha-glucosidase-like [Anoplophora		XP_018					
Lep_00	alpha-		glabripennis]gi 1080050546 ref XP_0185		566730,	1.2		176		
13348	glucosidase		66731.1 lysosomal alpha-glucosidase-like	Coleo	XP_018	E-		.79		
1-RA	-like	303	[Anoplophora glabripennis]	ptera	566731	50	92	2	100	92
Lep_00	Lysosomal		gi 321476729 gb EFX87689.1 hypothetic			1.3	69.5	329		
05729	alpha-	121	al protein DAPPUDRAFT_312137	Crust	EFX876	E-	266	.33		
6-RA	glucosidase	4	[Daphnia pulex]	acea	89	102	3	1	338	235
			gi 1227991792 ref XP_021927675.1 neut							
			ral alpha-glucosidase AB isoform X3							
	Neutral		[Zootermopsis		XP_021					
Lep_00	alpha-		nevadensis]gi 646709565 gb KDR15365.		927675,	3.4	93.0	156		
10636	glucosidase		1 Neutral alpha-glucosidase AB	Blatt	KDR153	7E-	232	.76		
8-RA	AB	390	[Zootermopsis nevadensis]	odea	65	43	6	2	86	80
	lysosomal									
Lep_00	alpha-					2.4	83.3	138		
02736	glucosidase		gi 952513507 gb KRT80074.1 glycoside	Coleo	KRT800	1E-	333	.27		
2-RA	-like	402	hydrolase, partial [Oryctes borbonicus]	ptera	74	38	3	2	84	70

Table 3.12. Continued

Seque	Sequence	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	desc.	uen		r		Val	larit	Sco	nme	itiv
name		ce				ue	У	re	nt	es
		leng							leng	
		th							th	
Lep_00	Lysosomal					5.2				
02736	alpha-		gi 1022773228 gb KZS17340.1 Lysosoma	Crust	KZS173	9E-		152		
1-RA	glucosidase	891	l alpha-glucosidase [Daphnia magna]	acea	40	39	92.5	.14	80	74
	uncharacte									
	rized family									
	31									
Lep_00	glucosidase		gi 1228003699 ref XP_021933735.1 unch			2.8	83.0	93.		
16092	KIAA1161-		aracterized family 31 glucosidase	Blatt	XP_021	2E-	508	974		
0-RA	like	186	KIAA1161-like [Zootermopsis nevadensis]	odea	933735	22	5	5	59	49
	uncharacte		gi 242004884 ref XP_002423306.1 conse							
	rized family		rved hypothetical protein [Pediculus							
	31		humanus		XP_002					
Lep_00	glucosidase		corporis]gi 212506315 gb EEB10568.1 c	Phthi	423306,		72.9	707		
08518	KIAA1161	240	onserved hypothetical protein [Pediculus	rapte	EEB105		938	.98		
2-RA	isoform X1	9	humanus corporis]	ra	68	0	3	3	648	473
	uncharacte									
	rized family									
	31		gi 1101361447 ref XP_018913028.1 PRE							
Lep_00	glucosidase		DICTED: uncharacterized family 31			2.4	83.0	88.		
07601	KIAA1161-		glucosidase KIAA1161-like [Bemisia	Hemi	XP_018	8E-	188	196		
3-RA	like	180	tabaci]	ptera	913028	20	7	5	53	44

Table 3.12. Continued	
-----------------------	--

Seque	Sequence desc.	Seque	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce		nce		r		Val	larit	Sco	nme	itiv
name		length				ue	У	re	nt	es
									leng	
									th	
	uncharacterized		gi 1101398930 ref XP_018910486.							
Lep_00	family 31		1 PREDICTED: uncharacterized			3.4	80.9	88.		
07601	glucosidase		family 31 glucosidase KIAA1161-	Hemi	XP_018	7E-	523	581		
4-RA	KIAA1161-like	405	like [Bemisia tabaci]	ptera	910486	19	8	7	63	51
	uncharacterized									
	family 31									
Lep_00	glucosidase		gi 1330911882 gb PNF32421.1 hy			3.8				
08951	KIAA1161-like		pothetical protein B7P43_G04891	Blatt	PNF324	E-		509		
8-RA	isoform X1	1839	[Cryptotermes secundus]	odea	21	172	60	.22	585	351
	Uncharacterize		gi 1228003699 ref XP_021933735.							
Lep_00	d family 31		1 uncharacterized family 31			2.5	67.9	458		
07177	glucosidase		glucosidase KIAA1161-like	Blatt	XP_021	E-	814	.37		
7-RA	KIAA1161	2259	[Zootermopsis nevadensis]	odea	933735	150	4	3	431	293
	uncharacterized		gi 1067066097 ref XP_018017064.							
Lep_00	family 31		1 PREDICTED: uncharacterized			3.5	55.8	277		
16154	glucosidase		family 31 glucosidase KIAA1161-	Crust	XP_018	6E-	265	.71		
6-RA	KIAA1161-like	1098	like [Hyalella azteca]	acea	017064	86	6	5	369	206
	uncharacterized									
Lep_00	family 31		gi 636630780 gb AIA09350.1 alph			3.7		70.		
07451	glucosidase		a-glucosidase family 31, partial	Blatt	AIA093	4E-	81.2	092		
7-RA	KIAA1161	156	[Periplaneta americana]	odea	50	14	5	2	48	39

Table 3.12. Continued	
-----------------------	--

Seque	Sequence desc.	Seque	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	_	nce		r		Val	larit	Sco	nme	itiv
name		length				ue	У	re	nt	es
									leng	
									th	
			gi 1339041694 ref XP_023710162.							
			1 neutral alpha-glucosidase AB							
			[Cryptotermes		XP_023					
Lep_00			secundus]gi 1330933432 gb PNF4		710162,	2.4	70.4	194		
04284	neutral alpha-		2738.1 Neutral alpha-glucosidase	Blatt	PNF427	E-	545	.12		
6-RA	glucosidase AB	789	AB [Cryptotermes secundus]	odea	38	54	5	6	176	124
			gi 321476729 gb EFX87689.1 hyp							
Lep_00	Lysosomal		othetical protein			1.8	72.7	274		
09708	alpha-		DAPPUDRAFT_312137 [Daphnia	Crust	EFX876	7E-	678	.63		
4-RA	glucosidase	717	pulex]	acea	89	84	6	3	224	163
	uncharacterized									
Lep_00	family 31		gi 636630780 gb AIA09350.1 alph				71.6	481		
13618	glucosidase		a-glucosidase family 31, partial	Blatt	AIA093	2E-	019	.87		
4-RA	KIAA1161-like	1527	[Periplaneta americana]	odea	50	163	4	1	412	295
	uncharacterized		gi 1228003699 ref XP_021933735.							
Lep_00	family 31		1 uncharacterized family 31			2.3	78.7	93.		
07907	glucosidase		glucosidase KIAA1161-like	Blatt	XP_021	4E-	878	589		
3-RA	KIAA1161-like	336	[Zootermopsis nevadensis]	odea	933735	21	8	3	66	52
	uncharacterized									
Lep_00	family 31		gi 1330911882 gb PNF32421.1 hy			8.0	77.2	83.		
07907	glucosidase		pothetical protein B7P43_G04891	Blatt	PNF324	1E-	727	188		
4-RA	KIAA1161-like	315	[Cryptotermes secundus]	odea	21	18	3	9	66	51

Table 3.12. Continue	d.
----------------------	----

Seque	Sequence desc.	Seque	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce		nce		r		Val	larit	Sco	nme	itiv
name		length				ue	У	re	nt	es
									leng	
									th	
			gi 1080050544 ref XP_018566730.							
			1 lysosomal alpha-glucosidase-like							
			[Anoplophora							
			glabripennis]gi 1080050546 ref X		XP_018					
Lep_00	lysosomal		P_018566731.1 lysosomal alpha-	_	566730,	1.1	91.3	164		
12808	alpha-		glucosidase-like [Anoplophora	Coleo	XP_018	E-	978	.85		
1-RA	glucosidase-like	444	glabripennis]	ptera	566731	45	5	1	93	85
	uncharacterized									
	family 31		gi 1316140871 ref XP_023215385.							
Lep_00	glucosidase		1 uncharacterized protein	Arthr		1.2	75.8	77.		
11953	KIAA1161-like		LOC111618155 isoform X1	opod	XP_023	7E-	064	796		
1-RA	isoform X2	368	[Centruroides sculpturatus]	а	215385	15	5	2	62	47
	uncharacterized		gi 1228003699 ref XP_021933735.							
Lep_00	family 31		1 uncharacterized family 31			1.9	83.3	95.		
12625	glucosidase		glucosidase KIAA1161-like	Blatt	XP_021	9E-	333	900		
5-RA	KIAA1161-like	285	[Zootermopsis nevadensis]	odea	933735	22	3	5	60	50
Lep_00	alpha-		gi 1227106192 gb 0XU28826.1 hy	Hym		5.6		70.		
11305	glucosidase		pothetical protein TSAR_016821	enopt	OXU288	1E-	67.1	477		
5-RA	isoform X2	393	[Trichomalopsis sarcophagae]	era	26	25	875	4	64	43
	lysosomal									
Lep_00	alpha-		gi 1022742921 gb KZR99685.1 Ly			1.5	52.6	186		
06888	glucosidase-like		sosomal alpha-glucosidase, partial	Crust	KZR996	5E-	845	.03		
5-RA	isoform X1	900	[Daphnia magna]	acea	85	53	6	7	298	157

Table 3.12. Continued	
-----------------------	--

Seque	Sequence desc.	Seque	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Alig	Pos
nce	_	nce		r		Val	larit	Sco	nme	itiv
name		length				ue	У	re	nt	es
									leng	
									th	
			gi 1330905640 gb PNF29605.1 hy							
			pothetical protein B7P43_G01635							
			[Cryptotermes							
			secundus]gi 1330905643 gb PNF2							
			9608.1 hypothetical protein		PNF296					
			B7P43_G01635 [Cryptotermes		05,					
			secundus]gi 1330905644 gb PNF2		PNF296					
Lep_00	lysosomal		9609.1 hypothetical protein		08,	3.6	66.8	291		
07664	alpha-		B7P43_G01635 [Cryptotermes	Blatt	PNF296	7E-	918	.19		
3-RA	glucosidase-like	1260	secundus]	odea	09	92	9	7	296	198
	uncharacterized									
Lep_00	family 31		gi 646703149 gb KDR11965.1 hyp			1.1	71.4			
13519	glucosidase		othetical protein L798_13618,	Blatt	KDR119	4E-	285	221		
8-RA	KIAA1161-like	597	partial [Zootermopsis nevadensis]	odea	65	67	7	.09	189	135
	uncharacterized									
Lep_00	family 31		gi 646703149 gb KDR11965.1 hyp			2.6	70.9	241		
13652	glucosidase		othetical protein L798_13618,	Blatt	KDR119	9E-	523	.89		
1-RA	KIAA1161-like	666	partial [Zootermopsis nevadensis]	odea	65	75	8	1	210	149
	uncharacterized									
Lep_00	family 31		gi 636630780 gb AIA09350.1 alph			1.7	71.3	479		
09474	glucosidase		a-glucosidase family 31, partial	Blatt	AIA093	E-	592	.17		
6-RA	KIAA1161-like	1476	[Periplaneta americana]	odea	50	162	2	4	412	294

Seque	Sequence	Seque	Hit desc.	Ord	Hit	E-	Simi	Bit-	Alig	Pos
nce	desc.	nce		er	ACC	Val	larit	Sco	nme	itiv
name		lengt				ue	У	re	nt	es
		h							leng	
									th	
	uncharacte									
	rized family									
	31									
Lep_00	glucosidase		gi 646703149 gb KDR11965.1 hypothetica			7.9	80.1	170		
10211	KIAA1161-		l protein L798_13618, partial	Blatt	KDR1	1E-	724	.62		
1-RA	like	519	[Zootermopsis nevadensis]	odea	1965	49	1	9	116	93
			gi 1330905640 gb PNF29605.1 hypothetic							
			al protein B7P43_G01635 [Cryptotermes							
			secundus]gi 1330905643 gb PNF29608.1		PNF2					
			hypothetical protein B7P43_G01635		9605,					
	lysosomal		[Cryptotermes		PNF2					
Lep_00	alpha-		secundus]gi 1330905644 gb PNF29609.1		9608,	4.2	89.7			
05479	glucosidase		hypothetical protein B7P43_G01635	Blatt	PNF2	6E-	727	153		
1-RA	-like	291	[Cryptotermes secundus]	odea	9609	44	3	.68	88	79
			gi 1330905640 gb PNF29605.1 hypothetic							
			al protein B7P43_G01635 [Cryptotermes							
			secundus]gi 1330905643 gb PNF29608.1		PNF2					
			hypothetical protein B7P43_G01635		9605,					
	lysosomal		[Cryptotermes		PNF2					
Lep_00	alpha-		secundus]gi 1330905644 gb PNF29609.1		9608,	6.6		219		
05479	glucosidase		hypothetical protein B7P43_G01635	Blatt	PNF2	6E-	62.1	.93		
0-RA	-like	957	[Cryptotermes secundus]	odea	9609	66	875	5	320	199

Seque	Sequence	Seque	Hit desc.	Ord	Hit	E-	Simi	Bit-	Alig	Pos
nce	desc.	nce		er	ACC	Val	larit	Sco	nme	itiv
name		lengt				ue	У	re	nt	es
		h							leng	
									th	
	lysosomal									
Lep_00	alpha-			Cole		1.2	83.3	139		
11734	glucosidase		gi 952513507 gb KRT80074.1 glycoside	opte	KRT8	3E-	333	.81		
9-RA	-like	459	hydrolase, partial [Oryctes borbonicus]	ra	0074	38	3	3	84	70
	Neutral									
Lep_00	alpha-		gi 1227991790 ref XP_021927674.1 neutr		XP_02	2.4	83.5	137		
11244	glucosidase		al alpha-glucosidase AB isoform X2	Blatt	19276	1E-	294	.50		
5-RA	AB	597	[Zootermopsis nevadensis]	odea	74	35	1	2	85	71
	lysosomal									
Lep_00	alpha-		gi 1022768805 gb KZS13426.1 Uncharact	Crus		7.4		80.		
01146	glucosidase		erized protein APZ42_021387 [Daphnia	tace	KZS13	6E-	76.5	492		
6-RA	-like	324	magna]	а	426	17	625	5	64	49

Sequenc	Seque	Seque	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posit
e name	nce	nce				Val	arity	Scor	ment	ives
	desc.	length				ue		e	length	
	beta-									
	1,3-									
	glucan-		gi 1339068387 ref XP_02372							
	binding		1291.1 uncharacterized							
	protein		protein LOC111872033							
Lep_0002	precurs		isoform X1 [Cryptotermes	Blattode	XP_0237		53.22	138		
7153-RA	or	7782	secundus]	а	21291	0	396	2.08	2621	1395
	Beta-									
	1,3-		gi 149777474 gb ABR28480.1							
	glucan-		beta-1,3(4)-glucanase LIC1,			7.0				
Lep_0005	binding		partial [Periplaneta	Blattode	ABR284	9E-	68.38	157.		
0787-RA	protein	1287	americana]	а	80	42	235	147	136	93
	Beta-									
	1,3-		gi 506968285 gb AGM32486.							
	glucan-		1 gram negative bacteria			3.4				
Lep_0003	binding		binding protein 2, partial	Blattode	AGM324	9E-	67.22	115.		
9162-RA	protein	389	[Coptotermes formosanus]	а	86	30	689	161	119	80
	beta-									
	1,3-									
	glucan-		gi 1339068387 ref XP_02372							
	binding		1291.1 uncharacterized							
	protein		protein LOC111872033							
Lep_0001	precurs		isoform X1 [Cryptotermes	Blattode	XP_0237		62.48	123		
0402-RA	or	4791	secundus]	а	21291	0	331	9.17	1498	936

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.

Table 3.13. Continue	d.
----------------------	----

Sequenc	Seque	Seque	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posit
e name	nce	nce				Val	arity	Scor	ment	ives
	desc.	length				ue		e	length	
	Beta-									
	1,3-		gi 1227967090 ref XP_02191							
	glucan-		4900.1 beta-1,3-glucan-			5.1				
Lep_0005	binding		binding protein-like	Blattode	XP_0219	8E-	74.71	109.		
2055-RA	protein	975	[Zootermopsis nevadensis]	а	14900	25	264	383	87	65
	beta-									
	1,3-									
	glucan-		gi 546684219 gb ERL93924.1							
	binding		hypothetical protein							
Lep_0003	protein		D910_11210 [Dendroctonus	Coleopt	ERL9392	4E-	53.84	65.0		
9163-RA	-like	393	ponderosae]	era	4	11	615	846	91	49
	beta-									
	1,3-									
	glucan-		gi 546684203 gb ERL93908.1							
	binding		hypothetical protein			3.7				
Lep_0004	protein		D910_11194, partial	Coleopt	ERL9390	7E-	84.21	65.4		
5091-RA	-like	393	[Dendroctonus ponderosae]	era	8	11	053	698	38	32
	beta-									
	1,3-									
	glucan-									
	binding		gi 1022754729 gb KZS02483.			3.9				
Lep_0003	protein		1 Beta-1,3-glucan-binding	Crustac	KZS0248	1E-		80.1		
6322-RA	-like	783	protein [Daphnia magna]	ea	3	15	45	073	140	63

Table 3.13. Continued	l.
-----------------------	----

Sequenc	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Posit
e name	desc.	nce			ACC	Val	arity	Scor	ment	ives
		lengt				ue		е	length	
		h								
	Beta-1,3-		gi 321457960 gb EFX69036.1							
	glucan-		hypothetical protein			3.8				
Lep_0007	binding		DAPPUDRAFT_203138	Crustac	EFX69	4E-	69.63	237.		
1308-RA	protein	952	[Daphnia pulex]	ea	036	74	563	654	247	172
	Beta-1,3-		gi 1022753497 gb KZS02148.1							
	glucan-		Beta-1,3-glucan-binding			6.9				
Lep_0012	binding		protein, partial [Daphnia	Crustac	KZS02	3E-	66.66	51.6		
0361-RA	protein	420	magna]	ea	148	07	667	026	45	30
	Beta-1,3-		gi 321461266 gb EFX72300.1							
	glucan-		hypothetical protein			4.4				
Lep_0004	binding		DAPPUDRAFT_201160	Crustac	EFX72	5E-	69.35	279.		
5092-RA	protein	837	[Daphnia pulex]	ea	300	91	484	256	248	172
	beta-1,3-									
	glucan-		gi 321457960 gb EFX69036.1							
	binding		hypothetical protein			3.0				
Lep_0003	protein-		DAPPUDRAFT_203138	Crustac	EFX69	2E-	55.88	182.		
1099-RA	like	1094	[Daphnia pulex]	ea	036	52	235	57	238	133
	beta-1,3-									
	glucan-		gi 1121175018 ref XP_019532							
	binding		058.1 PREDICTED: beta-1,3-		XP_01	2.1				
Lep_0002	protein-		glucan-binding protein [Aedes		95320	4E-	60.34	56.6		
9894-RA	like	321	albopictus]	Diptera	58	08	483	102	58	35

Sequenc e name	Sequence desc.	Seque nce lengt	Hit desc.	Order	Hit ACC	E- Val ue	Simil arity	Bit- Scor e	Align ment length	Posit ives
Lep_0002 9895-RA	beta-1,3- glucan- binding protein- like	534	gi 820836955 ref XP_0036903 62.2 PREDICTED: beta-1,3- glucan-binding protein 1 [Apis florea]	Hymeno ptera	XP_00 36903 62	5.8 E- 13	66.66 667	72.4 034	75	50
Lep_0002 9896-RA	beta-1,3- glucan- binding protein- like	921	gi 914569839 gb KOB73371.1 Beta-1,3-glucanase, partial [Operophtera brumata]	Lepidop tera	KOB7 3371	4.4 1E- 11	54.90 196	67.7 81	102	56
Lep_0002 8853-RA	beta-1,3- glucan- binding protein- like	974	gi 208972541 gb ACI32831.1 b eta-1,3-glucanase [Colias eurytheme]	Lepidop tera	ACI32 831	9.7 E- 30	61.90 476	122. 479	126	78

Hit desc. Hit Simil Bit-Seque Sequence Seque Order E-Alignme Positiv desc. nce ACC Va arity Score nt es nce length lu length name е 3.0 gi|1022765559|gb|KZS10 Endo-beta-754.1|Endo-beta-1,4-7E Lep_00 137319 mannanase [Daphnia KZS10 1,4-Crusta 71.63 -483 51 121 173.326 141 101 -RA magna] 754 mannanase cea gi|321460555|gb|EFX715 2.4 endo-beta-96.1|endo-beta-1,4-9E Lep 00 047049 1,4mannanase [Daphnia EFX71 68.47 Crusta 1575 826 110.923 92 63 -RA mannanase pulex] 596 24 cea gi|1022765557|gb|KZS10 3.3 Endo-beta-752.1|Endo-beta-1,4-3E Lep 00 042299 1,4mannanase [Daphnia KZS10 66.66 Crusta --RA 1022 magnal cea 752 92 667 285.419 288 192 mannanase gi|1022765557|gb|KZS10 6.2 Endo-beta-752.1|Endo-beta-1,4-2E Lep 00 015029 1,4mannanase [Daphnia KZS10 59.60 Crusta --RA 1044 752 77 265 246.899 302 180 magna] mannanase cea gi|1022765557|gb|KZS10 9E Lep_00 Endo-beta-752.1|Endo-beta-1,4-70.12 012530 1.4mannanase [Daphnia KZS10 11 Crusta 579 353.599 318 223 1041 magnal 752 9 -RA mannanase cea gi|1022765557|gb|KZS10 1.8 752.1|Endo-beta-1,4-Lep_00 Endo-beta-8E 057963 1,4mannanase [Daphnia KZS10 60.98 Crusta -996 752 44 901 162.155 182 -RA 111 mannanase magna] cea

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

Sequenc	Sequence	Seq	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Ali	Posi
e name	desc.	uen				Valu	arity	Scor	gn	tive
		leng				e		e	nt	3
		th							len	
									gth	
	endo-beta-									
	1,4-		gi 321460555 gb EFX71596.1							
Lep_0001	mannanas	104	endo-beta-1,4-mannanase	Crustac		7.01E	60.48	231.	29	
5051-RA	е	7	[Daphnia pulex]	ea	EFX71596	-71	11	491	1	176
	endo-beta-									
1 0004	1,4-		gi 321460555 gb EFX71596.1	a .		5 045	60.00	445		
Lep_0001	mannanas	500	endo-beta-1,4-mannanase	Crustac		7.81E	69.09	117.	11	-
5050-RA	e	502	[Daphnia pulex]	ea	EFX71596	-30	091	857	0	76
	endo-beta-									
1 0001	1,4-	100	g1 321460555 gb EFX/1596.1	<u> </u>			F 4 4 4	24.0	20	
	mannanas	100	endo-beta-1,4-mannanase	Crustac		4./5E	54.41	210.	28	1 5 4
5049-RA	e	5	[Daphnia pulex]	ea	EFX71596	-63	696	69	3	154
			gi 1339072353 ref XP_023723							
			402.1 beta-mannosidase							
			Isoform X2 [Cryptotermes							
			secundus]gl[13390/2355]ref[X							
			$P_023/23403.1$ [beta-		VD 000700					
			mannosidase isoform X2		XP_023723					
	Data		[Uryptotermes		402, VD 022722					
L 0012	Beta-		secundusjgi 13308/9524 gD P	Dlasta J	AP_023/23	(215	(1.20	(0.0		
	mannosid	207	NF1/252.1 Beta-mannosidase	Blattod	403, DNE17252	0.21E	01.29	08.9	02	F 7
2231-KA	ase	297	[Lryptotermes secundus]	ea	PNF1/252	-13	032	366	93	5/

Sequenc	Sequenc	Sequen	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Ali	Posi
e name	e desc.	ce				Valu	arity	Scor	gn	tive
		length				е		е	me	S
									nt	
									len oth	
	beta-								gui	
	mannosi									
	dase		gi 815919755 ref XP 012244							
Lep 0006	isoform		839.1 beta-mannosidase	Hymen	XP 012244	2.47E	61.16	172.	20	
7016-RA	X2	636	[Bombus impatiens]	optera	839	-47	505	555	6	126
	beta-			1						
	mannosi									
	dase		gi 998509058 ref XP_015514							
	precurso		937.1 PREDICTED: beta-							
Lep_0006	r,		mannosidase [Neodiprion	Hymen	XP_015514	2.23E	83.58	114.		
5194-RA	putative	795	lecontei]	optera	937	-26	209	39	67	56
	Beta-		gi 1061485315 gb 0DN0245							
Lep_0011	mannosi		4.1 Beta-mannosidase	Collem		1.81E	74.28	133.	10	
4547-RA	dase	465	[Orchesella cincta]	bola	ODN02454	-34	571	265	5	78
	beta-		gi 1279738204 ref XP_02291							
Lep_0010	mannosi		3917.1 beta-mannosidase	Coleopt	XP_022913	4.68E	63.94	139.	14	
1167-RA	dase	546	[Onthophagus taurus]	era	917	-36	558	043	7	94
	Mannan									
	endo-		gi 1061469313 gb 0DM8917							
	1,4-beta-		8.1 Mannan endo-1,4-beta-							
Lep_0001	mannosi		mannosidase, partial	Collem		7.01E	78.04	63.5		
5028-RA	dase	744	[Orchesella cincta]	bola	ODM89178	-10	878	438	41	32

Table 3.14. Continue	d.
----------------------	----

Sequenc	Sequenc	Sequence	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Posi
e name	e desc.	length				Valu	larit	Scor	ment	tive
						е	У	е	lengt	S
									h	
			gi 1325354705 ref XP_0							
	beta-		23311948.1 beta-							
	mannosi		mannosidase-like				64.7			
Lep_0011	dase,		[Anoplophora	Coleopt	XP_023311	5.87	058	73.1		
3998-RA	putative	585	glabripennis]	era	948	E-14	8	738	102	66
			gi 1227973104 ref XP_0							
			21917998.1 beta-							
			mannosidase							
			[Zootermopsis							
			nevadensis]gi 12279731							
			06 ref XP_021917999.1							
			beta-mannosidase							
			[Zootermopsis							
			nevadensis]gi 12279731							
			08 ref XP_021918000.1							
			beta-mannosidase		XP_021917					
			[Zootermopsis		998,					
			nevadensis]gi 64671762		XP_021917					
			5 gb KDR20416.1 Beta-		999,					
	beta-		mannosidase		XP_021918					
Lep_0000	mannosi		[Zootermopsis	Blattod	000,	1.6E-	60.3	451.		
3857-RA	dase	1761	nevadensis]	ea	KDR20416	146	352	055	537	324

Table	3.14.	Continued.
-------	-------	------------

Seque	Seque	Sequ	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	nce	ence		r		Valu	larit	Scor	ment	tive
name	desc.	lengt				е	У	е	lengt	S
		h							h	
			gi 1227973104 ref XP_021917998.1 bet							
			a-mannosidase [Zootermopsis							
			nevadensis]gi 1227973106 ref XP_0219							
			17999.1 beta-mannosidase		XP_0219					
			[Zootermopsis		17998,					
			nevadensis]gi 1227973108 ref XP_0219		XP_0219					
			18000.1 beta-mannosidase		17999,					
			[Zootermopsis		XP_0219					
Lep_00	beta-		nevadensis]gi 646717625 gb KDR20416.		18000,		70.3			
00385	manno		1 Beta-mannosidase [Zootermopsis	Blatt	KDR204	8.09	349	231.		
4-RA	sidase	819	nevadensis]	odea	16	E-68	3	491	209	147
	Beta-									
	manno									
Lep_00	sidase-		gi 1192752750 ref XP_015929642.2 bet				83.8			
00385	like		a-mannosidase [Parasteatoda	Arac	XP_0159	4.61	235	110.		
6-RA	protein	372	tepidariorum]	hnida	29642	E-27	3	538	68	57
			gi 1339072353 ref XP_023723402.1 bet							
			a-mannosidase isoform X2 [Cryptotermes							
			secundus]gi 1339072355 ref XP_023723		XP_0237					
			403.1 beta-mannosidase isoform X2		23402,					
			[Cryptotermes		XP_0237					
Lep_00	Beta-		secundus]gi 1330879524 gb PNF17252.		23403,		73.4			
07653	manno		1 Beta-mannosidase [Cryptotermes	Blatt	PNF172	4.1E-	693	65.4		
9-RA	sidase	651	secundus]	odea	52	10	9	698	49	36

Table	3.14.	Continued.
-------	-------	------------

Seque	Seque	Sequ	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	nce	ence		r		Valu	larit	Scor	ment	tive
name	desc.	lengt				е	У	е	lengt	S
		h							h	
	beta-									
	manno									
Lep_00	sidase		gi 998509058 ref XP_015514937.1 PRE	Hym			78.3			
03883	isofor		DICTED: beta-mannosidase [Neodiprion	enopt	XP_0155	5.68	132	121.		
6-RA	m X1	576	lecontei]	era	14937	E-30	5	709	83	65
Lep_00	beta-		gi 1325354705 ref XP_023311948.1 bet				74.6			
12470	manno		a-mannosidase-like [Anoplophora	Coleo	XP_0233	6.52	031	66.6		
5-RA	sidase	261	glabripennis]	ptera	11948	E-13	7	254	63	47
Lep_00	Beta-			Colle			75.7			
07066	manno		gi 1061485315 gb 0DN02454.1 Beta-	mbol	0DN024	3.15	281	127.		
3-RA	sidase	533	mannosidase [Orchesella cincta]	а	54	E-32	6	872	103	78
			gi 1339072353 ref XP_023723402.1 bet							
			a-mannosidase isoform X2 [Cryptotermes							
			secundus]gi 1339072355 ref XP_023723		XP_0237					
			403.1 beta-mannosidase isoform X2		23402,					
			[Cryptotermes		XP_0237					
Lep_00	Beta-		secundus]gi 1330879524 gb PNF17252.		23403,		75.0			
04064	manno		1 Beta-mannosidase [Cryptotermes	Blatt	PNF172	2.4E-	929	346.		
9-RA	sidase	909	secundus]	odea	52	110	4	28	269	202
Lep_00	beta-		gi 1035608711 ref XP_016911936.1 PRE	Hym			72.8			
04065	manno		DICTED: beta-mannosidase isoform X2	enopt	XP_0169	1.03	682	150.		
0-RA	sidase	639	[Apis cerana]	era	11936	E-40	2	984	129	94

Table 3.14. Continued.

Seque	Sequence	Sequence	Hit desc.	Order	Hit ACC	Е-	Simi	Bit-	Align	Posi
nce	desc.	length				Valu	larit	Scor	ment	tive
name						e	У	е	lengt	S
	Mannan								n	
	endo-1,4-		gi 1061469313 gb 0DM89							
Lep_00	beta-		178.1 Mannan endo-1,4-				75.9			
01253	mannosidas		beta-mannosidase, partial	Collemb	ODM891	6.21	493	107.		
1-RA	е	375	[Orchesella cincta]	ola	78	E-28	7	457	79	60
			gi 1227973104 ref XP_021							
			917998.1 beta-							
			mannosidase							
			[Zootermopsis							
			nevadensis]gi 1227973106							
			ref XP_021917999.1 beta-							
			mannosidase							
			[Zootermopsis							
			nevadensis]gi 1227973108		XP_0219					
			ref XP_021918000.1 beta-		17998,					
			mannosidase		XP_0219					
			[Zootermopsis		17999,					
			nevadensis]gi 646717625		XP_0219					
Lep_00	beta-		gb KDR20416.1 Beta-		18000,		74.0			
00536	mannosidas		mannosidase	Blattod	KDR204	2.1E-	740	491.		
1-RA	e	1290	[Zootermopsis nevadensis]	ea	16	164	7	115	405	300

Table	3.14.	Continued.
-------	-------	------------

Seque	Seque	Seque	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	nce	nce		r		Valu	larit	Scor	ment	tive
name	desc.	lengt				е	У	е	lengt	S
		h							h	
			gi 1227973104 ref XP_021917998.1 bet							
			a-mannosidase [Zootermopsis							
			nevadensis]gi 1227973106 ref XP_0219							
			17999.1 beta-mannosidase		XP_0219					
			[Zootermopsis		17998,					
			nevadensis]gi 1227973108 ref XP_0219		XP_0219					
			18000.1 beta-mannosidase		17999,					
			[Zootermopsis		XP_0219					
Lep_00	beta-		nevadensis]gi 646717625 gb KDR2041		18000,		58.0			
00536	manno		6.1 Beta-mannosidase [Zootermopsis	Blatt	KDR204	1.76	357	150.		
0-RA	sidase	867	nevadensis	odea	16	E-38	1	214	224	130
Lep_00	Beta-		gi 675378716 gb KFM71618.1 Beta-							
01618	manno		mannosidase, partial [Stegodyphus	Arac	KFM716	4.76		239.		
4-RA	sidase	564	mimosarum]	hnida	18	E-74	75	965	176	132
Lep_00	beta-						62.4			
01618	manno		gi 1279738204 ref XP_022913917.1 bet	Coleo	XP_0229	5.38	113	127.		
5-RA	sidase	546	a-mannosidase [Onthophagus taurus]	ptera	13917	E-32	5	102	141	88
Lep_00	beta-		gi 926610764 ref XP_013790438.1 beta	Arthr			59.1			
01618	manno		-mannosidase-like [Limulus	opod	XP_0137	8.68	836	196.		
3-RA	sidase	1071	polyphemus]	a	90438	E-54	7	438	245	145

Table	3.14.	Continued.
-------	-------	------------

Seque	Seque	Sequ	Hit desc.	Orde	Hit ACC	E-	Simi	Bit	Align	Posi
nce	nce	ence		r		Valu	larit	-	ment	tive
name	desc.	lengt				е	У	Sc	lengt	S
		h						or	h	
								е		
			gi 1227973104 ref XP_021917998.1 bet							
			a-mannosidase [Zootermopsis							
			nevadensis]gi 1227973106 ref XP_0219							
			1/999.1 beta-mannosidase		VD 00101					
					XP_02191					
			nevadensis]gi 122/9/3108 ref XP_0219		7998, VD 02101					
			18000.1 beta-mannosidase		XP_02191					
1 00	. .				7999, VD 00101			20		
Lep_00	beta-		nevadensis]gi 646/1/625 gb KDR20416.		XP_02191	0.00	75.7	28		
05363	manno		1 Beta-mannosidase [Zootermopsis	Blatt	8000,	3.03	709	9.2	005	4 5 0
6-RA	sidase	1464	nevadensis	odea	KDR20416	E-86	3	71	227	172
	alpha-									
Lep_00	manno		gi 1330899240 gb PNF26420.1 hypothet				76.1	77.		
10265	sidase		ical protein B7P43_G16606, partial	Blatt		1.93	904	41		
4-RA	2	381	[Cryptotermes secundus]	odea	PNF26420	E-15	8	1	63	48
			gi 1339064620 ref XP_023715606.1 lyso							
	Lysoso		somal alpha-mannosidase isoform X2							
	mal		[Cryptotermes							
Lep_00	alpha-		secundus]gi 1330931495 gb PNF42229.		XP_02371		63.8	37		
02253	manno		1 Lysosomal alpha-mannosidase	Blatt	5606,	1.1E	429	8.6		
3-RA	sidase	1704	[Cryptotermes secundus]	odea	PNF42229	-117	8	37	484	309

Seque	Sequen	Sequen	Hit desc.	Orde	Hit ACC	E-	Simi	Bit	Align	Posi
nce	ce desc.	ce		r		Valu	larit	-	ment	tive
name		length				е	У	SC or	lengt	S
								01 e	11	
	lysosom							-		
	al		gi 1339064618 ref XP_023715597.1 l							
	alpha-		ysosomal alpha-mannosidase isoform							
	mannos		X1 [Cryptotermes							
Lep_00	idase		secundus]gi 1330931496 gb PNF422		XP_02371		63.3	13		
02253	isoform		30.1 Lysosomal alpha-mannosidase	Blatt	5597,	6.28	093	6.3		
2-RA	X1	588	[Cryptotermes secundus]	odea	PNF42230	E-35	5	46	139	88
	mannos									
	yl-									
	oligosac									
	charide									
	alpha-									
	1,2-									
	mannos		gi 1227996329 ref XP_021929980.1							
Lep_00	idase IA		mannosyl-oligosaccharide alpha-1,2-	_			93.1	53		
06144	isoform		mannosidase IA isoform X2	Blatt	XP_02192		034	4.6		
5-RA	X2	1386	[Zootermopsis nevadensis]	odea	9980	0	5	43	319	297
	lysosom									
	al									
	alpha-									
Lep_00	mannos		gi 985414995 ref XP_015372896.1 P				65.7	26		
00169	idase-		REDICTED: lysosomal alpha-	Hemi	XP_01537	8.4E	894	6.9		
3-RA	like	1128	mannosidase-like [Diuraphis noxia]	ptera	2896	-79	7	29	304	200

Seque	Sequen	Sequen	Hit desc.	Orde	Hit ACC	E-	Simi	Bit	Align	Posi
nce	ce desc.	се		r		Valu	larit	-	ment	tive
name		length				е	У	Sc	lengt	S
								or	h	
								е		
	lysosom									
	al									
	alpha-									
Lep_00	mannos		gi 1233202594 ref XP_022200324.1 l				63.8			
00169	idase-		ysosomal alpha-mannosidase-like	Hemi	XP_02220	1.81	497	17		
4-RA	like	1206	[Nilaparvata lugens]	ptera	0324	E-44	7	1.4	213	136
	lysosom									
	al									
	alpha-									
Lep_00	mannos		gi 662196129 ref XP_008471080.1 P				86.2	20		
00169	idase-		REDICTED: lysosomal alpha-	Hemi	XP_00847	2.07	903	4.1		
5-RA	like	1440	mannosidase-like [Diaphorina citri]	ptera	1080	E-58	2	42	124	107
	lysosom									
	al									
	alpha-									
	mannos									
	idase									
	(manno									
	sidase									
	alpha		gi 668446735 gb KFB36711.1 lysoso							
Lep_00	class 2b		mal alpha-mannosidase				86.9	80.		
00169	membe		(mannosidase alpha class 2b member	Dipte		3.39	565	49		
7-RA	r 1)	438	1) [Anopheles sinensis]	ra	KFB36711	E-16	2	25	46	40

Seque	Sequenc	Sequen	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	e desc.	се				Valu	larit	Scor	ment	tive
name		length				е	У	e	lengt	S
						-		-	h	
	lysosomal									
	alpha-									
	mannosid									
Lep_00	ase		gi 817191866 ref XP_012271344.				60.3			
00169	isoform		1 lysosomal alpha-mannosidase	Hymen	XP_012	1.93E	174	51.2		
2-RA	X1	351	isoform X1 [Orussus abietinus]	optera	271344	-06	6	174	63	38
	lysosomal									
Lep_00	alpha-		gi 1229882973 ref XP_02216163							
00169	mannosid		0.1 lysosomal alpha-mannosidase-	Hemipt	XP_022	3.83E		47.7		
6-RA	ase-like	345	like [Myzus persicae]	era	161630	-05	75	506	40	30
	lysosomal		gi 1048052005 ref XP_01746075							
Lep_00	alpha-		5.1 PREDICTED: lysosomal alpha-				72.1			
06766	mannosid		mannosidase-like, partial		XP_017	3.85E	739	133.		
3-RA	ase-like	591	[Rhagoletis zephyria]	Diptera	460755	-35	1	65	115	83
	lysosomal		gi 1133436203 ref XP_01987431							
Lep_00	alpha-		8.1 PREDICTED: lysosomal alpha-				64.5			
06766	mannosid		mannosidase-like [Aethina	Coleop	XP_019	1.81E	833	48.1		
4-RA	ase-like	318	tumida]	tera	874318	-05	3	358	48	31
	lysosomal		gi 1080066641 ref XP_01857553							
Lep_00	alpha-		4.1 lysosomal alpha-mannosidase							
00152	mannosid		isoform X2 [Anoplophora	Coleop	XP_018	6.07E	73.4	60.8		
8-RA	ase-like	312	glabripennis]	tera	575534	-10	375	474	64	47

Seque	Sequence desc	Seque	Hit desc.	Order	Hit ACC	E- Valu	Simi larit	Bit- Scor	Align ment	Posi tive
name	uesei	length				e	y	e	lengt	s
	mannosyl- oligosacchar ide alpha- 1,2-		gi 1227996327 ref XP_0219299 79.1 mannosyl-oligosaccharide						<u>n</u>	
Lep_00	mannosidas		alpha-1,2-mannosidase IA				76.9			
11751	e IA isoform		isoform X1 [Zootermopsis	Blattod	XP_021	1.9E-	480	374.		
6-RA	X1	1014	nevadensis]	ea	929979	123	5	785	308	237

Table 3.14. Continued.

Seque	Sequence	Sequenc	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	desc.	e length				Valu	larit	Scor	ment	tive
name						е	У	е	lengt	S
							-		h	
			gi 1227998244 ref XP_021930							
			961.1 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		XP_021					
			nevadensis]gi 1227998250 ref		930961,					
			XP_021930964.1 alpha-		XP_021					
			mannosidase 2 [Zootermopsis		930962,					
			nevadensis]gi 1227998252 ref		XP_021					
			XP_021930965.1 alpha-		930963,					
			mannosidase 2 [Zootermopsis		XP_021					
			nevadensis]gi 1227998254 ref		930964,					
			XP_021930966.1 alpha-		XP_021					
			mannosidase 2 [Zootermopsis		930965,					
			nevadensis]gi 646706309 gb		XP_021					
Lep_00	alpha-		KDR13614.1 Alpha-		930966,		58.4			
09200	mannosida		mannosidase 2 [Zootermopsis	Blattod	KDR136	8.41E	112	123.		
0-RA	se 2	735	nevadensis]	ea	14	-30	1	635	214	125

Seque	Sequenc	Sequen	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	e desc.	ce				Valu	larit	Scor	ment	tive
name		length				е	У	е	lengt	S
Lep_00 01493	lysosomal alpha- mannosid ase isoform		gi 1339064620 ref XP_02371560 6.1 lysosomal alpha-mannosidase isoform X2 [Cryptotermes secundus]gi 1330931495 gb PNF 42229.1 Lysosomal alpha- mannosidase [Cryptotermes	Blattod	XP_023 715606, PNF422	3E-	59.7 756	441.	<u>n</u>	
2-RA	X1	2208	secundus]	ea	29	139	4	039	624	373
Lep_00 01493 3-RA	Lysosoma l alpha- mannosid ase	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]	Hymen optera	XP_011 348312, EZA483 46	3.4E- 20	66.9 354 8	92.8 189	124	83
Lep_00 14866 4-RA	mannosyl - oligosacc haride alpha-1,2- mannosid ase IA- like	735	gi 1227996327 ref XP_02192997 9.1 mannosyl-oligosaccharide alpha-1,2-mannosidase IA isoform X1 [Zootermopsis nevadensis]	Blattod ea	XP_021 929979	1.21E -62	70.6 976 7	213. 001	215	152

Table 3.14. Cont	inued.
------------------	--------

Seque	Sequenc	Sequen	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Posi
nce name	e desc.	ce length				Valu e	larit y	Scor e	ment lengt	tive s
							-		h	
	lysosomal		gi 766932221 ref XP_011498322.							
Lep_00	alpha-		1 PREDICTED: lysosomal alpha-				88.7			
05091	mannosid		mannosidase [Ceratosolen solmsi	Hymen	XP_011	9.21E	323	126.		
7-RA	ase	516	marchali]	optera	498322	-32	9	331	71	63
	Lysosoma		gi 1059382461 ref XP_01778605							
Lep_00	l alpha-		2.1 PREDICTED: lysosomal alpha-				74.3			
05091	mannosid		mannosidase isoform X1	Coleop	XP_017	6.15E	243	91.2		
6-RA	ase	666	[Nicrophorus vespilloides]	tera	786052	-19	2	781	74	55
			gi 1339052711 ref XP_02371469							
			1.1 alpha-mannosidase 2							
			[Cryptotermes							
			secundus]gi 1339052713 ref XP_							
			023714692.1 alpha-mannosidase		XP_023					
			2 [Cryptotermes		714691,					
			secundus]gi 1339052715 ref XP_		XP_023					
			023714693.1 alpha-mannosidase		714692,					
			2 [Cryptotermes		XP_023					
Lep_00	alpha-		secundus]gi 1339052717 ref XP_		714693,		73.6			
11611	mannosid		023714694.1 alpha-mannosidase	Blattod	XP_023	1.37E	842	73.9		
3-RA	ase 2	315	2 [Cryptotermes secundus]	ea	714694	-14	1	442	57	42
	lysosomal		gi 1000733841 ref XP_01558936							
Lep_00	alpha-		8.1 PREDICTED: lysosomal alpha-				62.8			
02175	mannosid		mannosidase isoform X1 [Cephus	Hymen	XP_015	3.01E	458	179.		
3-RA	ase	927	cinctus]	optera	589368	-48	5	104	253	159

Table 3.14. Continued	
-----------------------	--

Seque	Sequenc	Sequen	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	e desc.	ce				Valu	larit	Scor	ment	tive
name		length				e	У	e	lengt	S
	lysosomal		gi 985414995 ref XP_015372896.							
Lep_00	alpha-		1 PREDICTED: lysosomal alpha-				57.9			
02175	mannosid		mannosidase-like [Diuraphis	Hemipt	XP_015	1.11E	710	116.		
1-RA	ase	459	noxia]	era	372896	-28	1	701	138	80
			gi 1339057577 ref XP_02371729							
	ER		2.1 ER degradation-enhancing							
	degradati		alpha-mannosidase-like protein 2							
	on-		[Cryptotermes							
	enhancin		secundus]gi 1330893438 gb PNF							
	g alpha-		23714.1 ER degradation-		XP_023					
Lep_00	mannosid		enhancing alpha-mannosidase-		717292,		91.2			
03937	ase-like		like protein 2 [Cryptotermes	Blattod	PNF237		087	814.		
5-RA	protein 2	1389	secundus]	ea	14	0	9	683	455	415
	lysosomal		gi 321449456 gb EFX61897.1 hy							
Lep_00	alpha-		pothetical protein							
03065	mannosid		DAPPUDRAFT_337859, partial	Crusta	EFX618	8.22E	80.9	254.		
7-RA	ase-like	546	[Daphnia pulex]	cea	97	-79	816	603	163	132
	lysosomal									
	alpha-									
	mannosid		gi 751453449 ref XP_011181145.							
Lep_00	ase		1 PREDICTED: lysosomal alpha-				92.3			
03065	isoform		mannosidase isoform X1		XP_011	1.99E	076	79.3		
6-RA	X1	150	[Zeugodacus cucurbitae]	Diptera	181145	-17	9	37	39	36

Seque	Sequenc	Sequen	Hit desc.	Order	Hit ACC	E-	Simi	Bit-	Align	Posi
nce	e desc.	ce				Valu	larit	Scor	ment	tive
name		length				e	У	е	lengt	S
									h	
	lysosomal									
	alpha-									
	mannosid									
Lep_00	ase		gi 1321314822 ref XP_02329964				79.8			
03065	isoform		7.1 lysosomal alpha-mannosidase		XP_023	4.77E	319	168.		
8-RA	X2	585	isoform X2 [Lucilia cuprina]	Diptera	299647	-46	3	318	119	95
	alpha-									
	mannosid									
Lep_00	ase 2-like		gi 1325316915 ref XP_02334018				45.6			
08673	isoform		6.1 alpha-mannosidase 2-like	Phasm	XP_023	1.4E-	521	52.3		
7-RA	X1	720	isoform X1 [Eurytemora affinis]	atodea	340186	05	7	73	138	63
	ER									
	degradati									
	on-									
	enhancin									
	g alpha-		gi 1339066460 ref XP_02372533							
Lep_00	mannosid		3.1 ER degradation-enhancing				72.1			
07609	ase-like		alpha-mannosidase-like protein 3	Blattod	XP_023	1.85E	739	140.		
7-RA	protein 3	813	[Cryptotermes secundus]	ea	725333	-35	1	969	115	83
			gi 1330899240 gb PNF26420.1 h							
Lep_00	alpha-		ypothetical protein				75.4			
07272	mannosid		B7P43_G16606, partial	Blattod	PNF264	1.8E-	098	75.4		
4-RA	ase 2	720	[Cryptotermes secundus]	ea	20	13	4	85	61	46

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Po
nce	desc.	nce			ACC	Valu	larit	Sco	ment	sit
name		lengt				е	У	re	lengt	iv
		h							h	es
			gi 170061340 ref XP_001866193.1 ly							
			sosomal alpha-mannosidase [Culex		XP_001					
			quinquefasciatus]gi 167879594 gb E		86619					
Lep_0	lysosomal		DS42977.1 lysosomal alpha-		3,		60.3	92.		
00047	alpha-		mannosidase [Culex	Dipter	EDS42	2.19	603	048		
98-RA	mannosidase	625	quinquefasciatus]	а	977	E-19	6	5	111	67
			gi 1227972165 ref XP_021917507.1							
			ER degradation-enhancing alpha-							
			mannosidase-like protein 3 isoform							
			X1 [Zootermopsis							
			nevadensis]gi 1227972167 ref XP_02		XP_021					
			1917508.1 ER degradation-enhancing		91750					
	ER		alpha-mannosidase-like protein 3		7,					
	degradation-		isoform X1 [Zootermopsis		XP_021					
	enhancing		nevadensis]gi 646718115 gb KDR20		91750					
Lep_0	alpha-		717.1 ER degradation-enhancing		8,		63.6	123		
00848	mannosidase-		alpha-mannosidase-like 3	Blatto	KDR20	2.05	363	.63		
06-RA	like protein 3	417	[Zootermopsis nevadensis]	dea	717	E-31	6	5	132	84
	Endoplasmic									
	reticulum		gi 1339074955 ref XP_023724784.1							
	mannosyl-		endoplasmic reticulum mannosyl-							
Lep_0	oligosaccharid		oligosaccharide 1,2-alpha-		XP_023					
00297	e 1,2-alpha-		mannosidase isoform X2	Blatto	72478	3.36	86.2	240		12
15-RA	mannosidase	516	[Cryptotermes secundus]	dea	4	E-75	069	.35	145	5
Table 3.14. Continued.										

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Po
nce	desc.	nce			ACC	Valu	arity	Sco	ment	sit
name		lengt				е		re	lengt	iv
		h							h	es
	endoplasmic									
	reticulum		gi 1339074955 ref XP_023724784.							
	mannosyl-		1 endoplasmic reticulum							
Lep_0	oligosaccharid		mannosyl-oligosaccharide 1,2-		XP_023			280		
00297	e 1,2-alpha-		alpha-mannosidase isoform X2	Blattod	72478	1.66E	88.82	.79		15
16-RA	mannosidase	966	[Cryptotermes secundus]	ea	4	-88	353	6	170	1
	mannosyl-									
	oligosaccharid		gi 1227996327 ref XP_021929979.							
Lep_0	e 1,2-alpha-		1 mannosyl-oligosaccharide alpha-		XP_021			431		
00405	mannosidase		1,2-mannosidase IA isoform X1	Blattod	92997	6.9E-	64.87	.40		31
97-RA	IA isoform X1	1698	[Zootermopsis nevadensis]	ea	9	142	603	9	484	4
	lysosomal									
Lep_0	alpha-		gi 675365773 gb KFM58675.1 Lys					206		
00749	mannosidase		osomal alpha-mannosidase, partial	Arachn	KFM58	1.12E	79.54	.06		10
53-RA	isoform X1	549	[Stegodyphus mimosarum]	ida	675	-61	545	8	132	5
Lep_0	Lysosomal		gi 1061483689 gb 0DN00899.1 Ly					332		
00085	alpha-		sosomal alpha-mannosidase	Collem	ODN00	2.2E-	56.91	.79		24
75-RA	mannosidase	1755	[Orchesella cincta]	bola	899	100	244	8	434	7
	lysosomal		gi 321463090 gb EFX74108.1 hyp							
Lep_0	alpha-		othetical protein					271		
00085	mannosidase-		DAPPUDRAFT_252189 [Daphnia	Crusta	EFX74	5.11E	61.59	.16		17
76-RA	like	1730	pulex]	cea	108	-78	42	6	276	0

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Po
nce	desc.	nce			ACC	Valu	larit	Sco	ment	sit
name		lengt				е	У	re	lengt	iv
		h							h	es
	lysosomal									
Lep_0	alpha-		gi 332029389 gb EGI69344.1 Lyso				64.3			
00870	mannosidase		somal alpha-mannosidase	Hymeno	EGI693	1.02	678	229		16
43-RA	isoform X1	822	[Acromyrmex echinatior]	ptera	44	E-67	2	.18	261	8
			gi 1199407503 ref XP_021199415.							
			1 lysosomal alpha-mannosidase-		XP_021					
			like isoform X4 [Helicoverpa		19941					
	lysosomal		armigera]gi 1199407505 ref XP_0		5,					
Lep_0	alpha-		21199416.1 lysosomal alpha-		XP_021			93.		
00772	mannosidase-		mannosidase-like isoform X5	Lepidopt	19941	1.9E	58.8	974		
83-RA	like	339	[Helicoverpa armigera]	era	6	-21	785	5	107	63
			gi 1339064618 ref XP_023715597.							
			1 lysosomal alpha-mannosidase							
			isoform X1 [Cryptotermes		XP_023					
			secundus]gi 1330931496 gb PNF4		71559					
Lep_0	Lysosomal		2230.1 Lysosomal alpha-		7,		66.6	395		
00124	alpha-		mannosidase [Cryptotermes	Blattode	PNF42	1.9E	666	.58		29
77-RA	mannosidase	1635	secundus]	а	230	-124	7	6	435	0
	lysosomal									
Lep_0	alpha-		gi 1351647487 ref XP_024086113.		XP_024		58.6	85.		
00124	mannosidase		1 lysosomal alpha-mannosidase-	Hemipte	08611	1.85	538	500		
78-RA	isoform X2	345	like [Cimex lectularius]	ra	3	E-18	5	1	104	61

Seque	Sequence	Seque	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Po
nce	desc.	nce				Value	arity	Scor	ment	sit
name		lengt						е	lengt	iv
		h							h	es
	lysosomal		gi 1351647487 ref XP_02408							
Lep_0	alpha-		6113.1 lysosomal alpha-							
00124	mannosidase-		mannosidase-like [Cimex	Hemipt	XP_02408	7.12E-	75.40	56.9		
79-RA	like	828	lectularius]	era	6113	07	984	954	61	46
			gi 1227982315 ref XP_02192							
			2781.1 ER degradation-							
			enhancing alpha-							
			mannosidase-like protein 1							
	ER		[Zootermopsis							
	degradation-		nevadensis]gi 646713767 gb							
	enhancing		KDR17988.1 ER degradation-		XP_02192					
Lep_0	alpha-		enhancing alpha-		2781,					
00511	mannosidase-		mannosidase-like 1	Blattod	KDR1798	5.05E-	91.87	291.		14
81-RA	like protein 1	1284	[Zootermopsis nevadensis]	ea	8	91	5	967	160	7
	ER									
	degradation-		gi 751228804 ref XP_011167							
	enhancing		829.1 PREDICTED: ER							
Lep_0	alpha-		degradation-enhancing alpha-							
00627	mannosidase-		mannosidase-like protein 3,	Hymeno	XP_01116	1.73E-	97.67	164.		
28-RA	like protein 3	327	partial [Solenopsis invicta]	ptera	7829	51	442	466	86	84

Seque nce name	Sequenc e desc.	Seq ue nce len gth	Hit desc.	Order	Hit ACC	E- Valu e	Simi larit y	Bit- Scor e	Align ment lengt h	Po sit iv es
Lep_0 00627	ER degradati on- enhancin g alpha- mannosid ase-like protein 3 isoform	200	gi 1227972169 ref XP_021917509.1 E R degradation-enhancing alpha- mannosidase-like protein 3 isoform X2 [Zootermopsis nevadensis]gi 1227972171 ref XP_021 917511.1 ER degradation-enhancing alpha-mannosidase-like protein 3	Plattodaa	XP_021 91750 9, XP_021 91751	2.03 E 22	100	100.5	42	12
Lep_0 00648 16-RA	lysosomal alpha- mannosid ase isoform X2	588	gi 1096257085 gb APA33853.1 semina l fluid protein [Nilaparvata lugens]	Hemipter	APA33 853	3.78 E-12	75	70.47 74	56	42

Seque	Sequence	Sequen	Hit desc.	Order	Hit	E- Val	Sim	Bit	Align	Po
name	uesc.	length			ALL	ue	itv	- Sc	lengt	iv
		0					5	or	h	es
								е		
					XP_023					
			gi 1339052711 ref XP_023714691.1 alp		71469					
			ha-mannosidase 2 [Cryptotermes		1,					
			secundus]gi 1339052713 ref XP_02371		XP_023					
			4692.1 alpha-mannosidase 2		71469					
			[Cryptotermes		2,					
			secundus]gi 1339052715 ref XP_02371		XP_023					
			4693.1 alpha-mannosidase 2		71469					
			[Cryptotermes		3,					
Lep_0	alpha-		secundus]gi 1339052717 ref XP_02371		XP_023	7.8	73.	79.		
00590	mannosida		4694.1 alpha-mannosidase 2	Blattod	71469	1E-	770	33		
16-RA	se 2	252	[Cryptotermes secundus]	ea	4	17	49	7	61	45
	lysosomal									
	alpha-				VD 040	2.4	60	22		
Lep_0	mannosida		gi 817191868 ref XP_012271345.1 lyso		XP_012	2.4	69. 000	23		10
00607	se isoform	000	somal alpha-mannosidase isoform X2	Hymen	27134	는-	029	8.0	260	18
69-RA	XZ	899	[Orussus abietinus]	optera	5	69	85	39	268	5
	lysosomal									
1 0	aipna-		- 11251(47407L - (WD 02400(1121)) -		VD 024	F 2	(0)	1 4		
	mannosida		$gi 135164/48/ ref XP_024086113.1 lys$	Homist	λΡ_024	5.3 (E	60.	14 7 r		11
00396	se isolorm	770		nempt	08011	0E-	320	7.5 17	100	
90-KA	Δ1	//ð	iectularius	era	3	38	52	1/	190	5

Table 3.14. Continue	d.
----------------------	----

Seque	Sequence	Sequen	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Ро
nce	desc.	се			ACC	Value	larit	Sco	ment	sit
name		length					У	re	lengt	iv
									h	es
	lysosomal		gi 1059382461 ref XP_017786052.							
Lep_0	alpha-		1 PREDICTED: lysosomal alpha-		XP_017		81.6	176		
00396	mannosida		mannosidase isoform X1	Coleop	78605	1.29E-	793	.40		10
89-RA	se	867	[Nicrophorus vespilloides]	tera	2	47	9	7	131	7
	lysosomal									
Lep_0	alpha-		gi 662196129 ref XP_008471080.1		XP_008		70.7	253		
00330	mannosida		PREDICTED: lysosomal alpha-	Hemipt	47108	9.03E-	627	.44		16
60-RA	se-like	711	mannosidase-like [Diaphorina citri]	era	0	81	1	7	236	7
	lysosomal		gi 929380069 ref XP_014100402.1							
Lep_0	alpha-		PREDICTED: lysosomal alpha-		XP_014		79.8	189		
00330	mannosida		mannosidase-like, partial		10040	1.74E-	449	.11		10
59-RA	se-like	447	[Bactrocera oleae]	Diptera	2	59	6	9	129	3

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Sim	Bit	Align	Po
nce	desc.	nce			ACC	Val	ilar	-	ment	sit
name		lengt				ue	ity	Sc	lengt	iv
		h						or	h	es
								e		
			gi 1227998244 ref XP_021930961.1 alp		XP_021					
			ha-mannosidase 2 [Zootermopsis		93096					
			nevadensis]gi 1227998246 ref XP_0219		1,					
			30962.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998248 ref XP_0219		2,					
			30963.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998250 ref XP_0219		3,					
			30964.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998252 ref XP_0219		4,					
			30965.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998254 ref XP_0219		5,					
			30966.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
Lep_0	alpha-		nevadensis]gi 646706309 gb KDR13614.		6,	1.2	57.	20		
00776	mannosidase		1 Alpha-mannosidase 2 [Zootermopsis	Blatto	KDR13	2E-	551	3.7		14
42-RA	2	735	nevadensis]	dea	614	57	02	56	245	1
Lep_0	lysosomal		gi 1130264682 ref XP_019770786.1 PRE		XP_019	1.7	54.	93.		
00340	alpha-		DICTED: lysosomal alpha-mannosidase-	Coleo	77078	E-	676	20		
60-RA	mannosidase	465	like [Dendroctonus ponderosae]	ptera	6	20	26	41	139	76

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Sim	Bit	Align	Po
nce	desc.	nce			ACC	Val	ilar	-	ment	sit
name		lengt				ue	ity	Sc	lengt	iv
		h						or	h	es
								е		
			gi 1339064618 ref XP_023715597.1 lyso							
			somal alpha-mannosidase isoform X1		XP_023					
	lysosomal		[Cryptotermes		71559					
Lep_0	alpha-		secundus]gi 1330931496 gb PNF42230.		7,		64.	75		
00118	mannosidase		1 Lysosomal alpha-mannosidase	Blatto	PNF42		504	6.5		52
93-RA	isoform X1	2730	[Cryptotermes secundus]	dea	230	0	28	18	817	7
Lep_0	alpha-		gi 1009600602 ref XP_015928615.1 alp		XP_015	2.2	80.	16		
01241	mannosidase		ha-mannosidase 2 isoform X1	Arach	92861	8E-	188	2.1		
10-RA	2	336	[Parasteatoda tepidariorum]	nida	5	45	68	55	106	85
			gi 1048012429 ref XP_017467120.1 PRE							
			DICTED: lysosomal alpha-mannosidase		XP_017					
			isoform X1 [Rhagoletis		46712					
	lysosomal		zephyria]gi 1048012433 ref XP 017467		0,					
Lep 0	alpha-		123.1 PREDICTED: lysosomal alpha-		XP 017	9.1	60.	31		
00469	mannosidase		mannosidase isoform X1 [Rhagoletis	Dipter	46712	6E-	597	9.3		22
02-RA	isoform X1	1275	zephyria]	a	3	97	83	16	368	3
	ER									
	degradation-									
	enhancing									
	alpha-		gi 1330882322 gb PNF18187.1 ER							1
Lep 0	mannosidase		degradation-enhancing alpha-			6.9	91.	20		1
01157	-like protein		mannosidase-like protein 1, partial	Blatto	PNF18	9E-	596	4.5		10
01-RA	1	399	[Cryptotermes secundus]	dea	187	64	64	27	119	9

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Sim	Bit	Align	Po
nce	desc.	nce lengt			ALL	var	llar itv	- Sc	ment lengt	SIU iv
name		h				uc	ity	or	h	es
								e		•••
			gi 1227982315 ref XP_021922781.1 ER							
	ER		degradation-enhancing alpha-							
	degradation-		mannosidase-like protein 1							
	enhancing		[Zootermopsis		XP_021					
	alpha-		nevadensis]gi 646713767 gb KDR17988.		92278					
Lep_0	mannosidase		1 ER degradation-enhancing alpha-		1,	7.9	81.	13		
00988	-like protein		mannosidase-like 1 [Zootermopsis	Blatto	KDR17	8E-	818	6.7		
56-RA	1	336	nevadensis]	dea	988	37	18	32	88	72
			gi 1339064620 ref XP_023715606.1 lyso							
			somal alpha-mannosidase isoform X2		XP_023					
			[Cryptotermes		71560					
Lep_0	Lysosomal		secundus]gi 1330931495 gb PNF42229.		6,	1.6	62.	66.		
01272	alpha-		1 Lysosomal alpha-mannosidase	Blatto	PNF42	5E-	666	62		
03-RA	mannosidase	231	[Cryptotermes secundus]	dea	229	12	67	54	75	47
	lysosomal									
Lep_0	alpha-		gi 1000733841 ref XP_015589368.1 PRE	Hyme	XP_015	3.3	63.	14		
01368	mannosidase		DICTED: lysosomal alpha-mannosidase	nopte	58936	8E-	953	1.7		11
82-RA	isoform X2	504	isoform X1 [Cephus cinctus]	ra	8	37	49	39	172	0
	lysosomal									
Lep_0	alpha-		gi 929380069 ref XP_014100402.1 PRE		XP_014	2.6	78.	21		
00655	mannosidase		DICTED: lysosomal alpha-mannosidase-	Dipter	10040	6E-	666	4.1		11
33-RA	isoform X1	825	like, partial [Bactrocera oleae]	a	2	67	67	57	150	8

Seque	Sequence	Seque	Hit desc.	Order	Hit	Е-	Sim	Bit	Align	Po
nce	desc.	nce			ACC	Val	ilar	-	ment	sit
name		lengt				ue	ity	Sc	lengt	iv
		h						or	h	es
								е		
Lep_0	alpha-		gi 1009600602 ref XP_015928615.1 alp	_	XP_015	5.7	80.	16		
00828	mannosidase		ha-mannosidase 2 isoform X1	Arach	92861	5E-	952	0.9		
57-RA	2	326	[Parasteatoda tepidariorum]	nida	5	45	38	99	105	85
	lysosomal									
Lep_0	alpha-		gi 1351647487 ref XP_024086113.1 lyso		XP_024	1.1	64.	17		
00992	mannosidase		somal alpha-mannosidase-like [Cimex	Hemip	08611	5E-	356	7.5		13
28-RA	isoform X2	719	lectularius]	tera	3	48	44	63	202	0
Lep_0	lysosomal		gi 985414995 ref XP_015372896.1 PRE		XP_015	3.4	55.	11		
01010	alpha-		DICTED: lysosomal alpha-mannosidase-	Hemip	37289	9E-	704	7.8		
70-RA	mannosidase	462	like [Diuraphis noxia]	tera	6	29	7	57	149	83
					XP_023					
			gi 1339078211 ref XP_023726521.1 alp		72652					
			ha-mannosidase 2 [Cryptotermes		1,					
			secundus]gi 1339078213 ref XP_023726		XP_023					
			522.1 alpha-mannosidase 2		72652					
			[Cryptotermes		2,					
Lep_0	alpha-		secundus]gi 1339078215 ref XP_023726		XP_023		79.	14		
00695	mannosidase		523.1 alpha-mannosidase 2	Blatto	72652		206	44.		83
67-RA	2	3048	[Cryptotermes secundus]	dea	3	0	05	1	1058	8
	lysosomal									
Lep_0	alpha-					2.7	68.	12		
00622	mannosidase		gi 1096257085 gb APA33853.1 seminal	Hemip	APA33	2E-	103	9.4		
35-RA	-like	435	fluid protein [Nilaparvata lugens]	tera	853	33	45	13	116	79

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Po
nce	desc.	nce			ACC	Valu	larit	Sco	ment	sit
name		lengt				e	У	re	lengt	iv
		h							h	es
			gi 1048012429 ref XP_017467120.1 P							
			REDICTED: lysosomal alpha-		XP_017					
			mannosidase isoform X1 [Rhagoletis		46712					
	lysosomal		zephyria]gi 1048012433 ref XP_0174		0,					
Lep_0	alpha-		67123.1 PREDICTED: lysosomal alpha-		XP_017		78.9	350		
00058	mannosidase		mannosidase isoform X1 [Rhagoletis	Dipter	46712	2E-	256	.13		19
70-RA	isoform X1	873	zephyria]	а	3	110	2	2	242	1
			gi 1339064618 ref XP_023715597.1 l							
			ysosomal alpha-mannosidase isoform		XP_023					
			X1 [Cryptotermes		71559					
Lep_0	Lysosomal		secundus]gi 1330931496 gb PNF4223		7,		63.7			
00058	alpha-		0.1 Lysosomal alpha-mannosidase	Blatto	PNF42	1.75	195	269		20
67-RA	mannosidase	1357	[Cryptotermes secundus]	dea	230	E-78	1	.24	328	9
	lysosomal									
Lep_0	alpha-		gi 1247040342 gb PCG75287.1 hypot				73.0			
00058	mannosidase		hetical protein B5V51_11959	Lepid	PCG75	9.48	769	63.		
68-RA	-like	612	[Heliothis virescens]	optera	287	E-10	2	929	52	38

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Sim	Bit	Align	Po
nce	desc.	nce			ACC	Val	ilar	-	ment	sit
name		lengt				ue	ity	Sc	lengt	iv
		h						or	h	es
								е		
					XP_023					
			gi 1339052711 ref XP_023714691.1 alp		71469					
			ha-mannosidase 2 [Cryptotermes		1,					
			secundus]gi 1339052713 ref XP_023714		XP_023					
			692.1 alpha-mannosidase 2		71469					
			[Cryptotermes		2,					
			secundus]gi 1339052715 ref XP_023714		XP_023					
			693.1 alpha-mannosidase 2		71469					
			[Cryptotermes		3,					
Lep_0	alpha-		secundus]gi 1339052717 ref XP_023714		XP_023		78.	79		
00011	mannosidase		694.1 alpha-mannosidase 2	Blatto	71469		509	1.1		45
27-RA	2	2151	[Cryptotermes secundus]	dea	4	0	53	86	577	3

Seque	Sequence	Seque	Hit desc.	Order	Hit	Е-	Sim	Bit	Align	Po
nce	desc.	nce			ACC	Val	ilar	-	ment	sit
name		lengt				ue	ity	Sc	lengt	iv
		h						or	h	es
								е		
			gi 1227998244 ref XP_021930961.1 alp		XP_021					
			ha-mannosidase 2 [Zootermopsis		93096					
			nevadensis]gi 1227998246 ref XP_0219		1,					
			30962.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998248 ref XP_0219		2,					
			30963.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998250 ref XP_0219		3,					
			30964.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998252 ref XP_0219		4,					
			30965.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
			nevadensis]gi 1227998254 ref XP_0219		5,					
			30966.1 alpha-mannosidase 2		XP_021					
			[Zootermopsis		93096					
Lep_0	alpha-		nevadensis]gi 646706309 gb KDR13614.		6,	5.4	66.	23		
00011	mannosidase		1 Alpha-mannosidase 2 [Zootermopsis	Blatto	KDR13	6E-	425	0.7		18
28-RA	2	990	nevadensis]	dea	614	66	99	2	277	4
Lep_0	alpha-		gi 1330899240 gb PNF26420.1 hypothet			5.8	72.	26		
00011	mannosidase		ical protein B7P43_G16606, partial	Blatto	PNF26	8E-	602	4.6		15
29-RA	2	918	[Cryptotermes secundus]	dea	420	82	74	18	219	9

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Sim	Bit	Align	Po
nce	desc.	nce			ACC	Val	ilar	-	ment	sit
name		lengt				ue	ity	Sc	lengt	iv
		h						or	h	es
								е		
			gi 1339064618 ref XP_023715597.1 lyso							
			somal alpha-mannosidase isoform X1		XP_023					
			[Cryptotermes		71559					
Lep_0	Lysosomal		secundus]gi 1330931496 gb PNF42230.		7,		63.	78		
00060	alpha-		1 Lysosomal alpha-mannosidase	Blatto	PNF42		443	3.8		53
52-RA	mannosidase	2921	[Cryptotermes secundus]	dea	230	0	4	67	848	8

Seque	Sequenc	Sequ	Hit desc.	Order	Hit	E- Volu	Simi	Bit-	Align	Pos
name	e uest.	lengt			ALL	e valu	y y	e	lengt	es
		h							h	
	beta-									
	glucuroni									
Lep_00	dase-like		gi 1330907784 gb PNF30346.1 hypoth				78.5			
10114	isoform		etical protein B7P43_G13405	Blatto	PNF3	1.24E	714	116.		
3-RA	X1	414	[Cryptotermes secundus]	dea	0346	-30	3	316	84	66
			gi 241309940 ref XP_002407819.1 bet							
			a-glucuronidase (GusB), putative		XP_00					
			[Ixodes		24078					
Lep_00	beta-		scapularis]gi 215497226 gb EEC06720		19,		70.5			
02223	glucuroni		.1 beta-glucuronidase (GusB), putative	Arach	EEC06	1.97E	128	267.		
8-RA	dase-like	918	[Ixodes scapularis]	nida	720	-87	2	314	234	165
Lep_00	beta-		gi 1344820106 ref XP_023953356.1 be		XP_02		65.2			
02223	glucuroni		ta-glucuronidase-like, partial [Bicyclus	Lepid	39533	3.75E	173	122.		
9-RA	dase	576	anynana]	optera	56	-31	9	865	115	75
			gi 939667583 ref XP_014280874.1 PR		XP_01					
			EDICTED: beta-glucuronidase-like		42808					
			isoform X1 [Halyomorpha		74,					
			halys]gi 939667586 ref XP_014280876		XP_01					
	beta-		.1 PREDICTED: beta-glucuronidase-like		42808					
	glucuroni		isoform X1 [Halyomorpha		76,					
Lep_00	dase		halys]gi 939667589 ref XP_014280877		XP_01		77.0			
00015	isoform		.1 PREDICTED: beta-glucuronidase-like	Hemip	42808	7.49E	588	229.		
6-RA	X3	944	isoform X1 [Halyomorpha halys]	tera	77	-68	2	18	170	131

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

Table 3.15. Continued	
-----------------------	--

Seque	Sequenc	Sequ	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
nce	e desc.	ence			ACC	Valu	larit	Scor	ment	itiv
name		h				e	У	e	h	es
	beta-									
	glucuroni									
Lep_00	dase-like		gi 641675506 ref XP_008187067.1 PR		XP_00		75.9			
00015	isoform		EDICTED: beta-glucuronidase isoform	Hemip	81870	3.27E	656	300.		
8-RA	X1	942	X2 [Acyrthosiphon pisum]	tera	67	-95	7	827	233	177
	beta-									
	glucuroni						0.7.0			
Lep_00	dase-like		gi 1233178675 ref XP_022190762.1 be		XP_02	0.455	87.3	0		
06534	isoform	205	ta-glucuronidase-like [Nilaparvata	Hemip	21907	8.47E	015	97.0	()	
6-RA	XI	285	lugens	tera	62	-23	9	561	63	55
Lep_00	beta-		gi 1316153756 ref XP_023221887.1 be	_	XP_02		67.2			
06534	glucuroni		ta-glucuronidase-like [Centruroides	Arthr	32218	7.47E	131	69.7		
5-RA	dase-like	372	sculpturatus]	opoda	87	-13	1	07	61	41
	beta-									
	glucuroni									
Lep_00	dase-like						77.5			
07599	isoform		gi 1022772830 gb KZS16987.1 Beta-	Crusta	KZS16	8.94E	362	192.		
9-RA	X1	672	glucuronidase [Daphnia magna]	cea	987	-61	3	586	138	107
Lep_00	beta-		gi 1330889853 gb PNF22051.1 hypoth				75.4			
05373	glucuroni		etical protein B7P43_G09739	Blatto	PNF2	3.6E-	966	217.		
8-RA	dase-like	510	[Cryptotermes secundus]	dea	2051	69	9	624	151	114
Lep_00	beta-		gi 1330889853 gb PNF22051.1 hypoth				75.4			
05585	glucuroni		etical protein B7P43_G09739	Blatto	PNF2	2.73E	966	217.		
5-RA	dase-like	468	[Cryptotermes secundus]	dea	2051	-69	9	238	151	114

Table	3.15.	Continued.

Seque	Sequenc	Sequ	Hit desc.	Ord	Hit ACC	E-	Simi	Bit-	Align	Pos
nce	e desc.	ence		er		Val	larit	Scor	ment	itiv
name		lengt				ue	У	е	lengt	es
		h							h	
Lep_00	Beta-		gi 1316153756 ref XP_023221887.1 bet	Crus		6.2	76.2			
11856	glucuroni		a-glucuronidase-like [Centruroides	tace	XP_023	2E-	295	159.		
3-RA	dase	657	sculpturatus]	а	221887	44	1	073	122	93
Lep_00	Beta-		gi 1316153756 ref XP_023221887.1 bet	Crus		4.1	76.2			
11833	glucuroni		a-glucuronidase-like [Centruroides	tace	XP_023	6E-	295	157.		
3-RA	dase	657	sculpturatus]	а	221887	43	1	147	122	93
	beta-									
	glucuroni									
Lep_00	dase-like		gi 1339060577 ref XP_023718876.1 bet			2.8	70.2			
04714	isoform		a-glucuronidase-like isoform X2	Blatt	XP_023	E-	947	480.		
6-RA	X1	1386	[Cryptotermes secundus]	odea	718876	163	8	33	441	310
			gi 1228013990 ref XP_021938962.1 bet							
			a-glucuronidase [Zootermopsis							
			nevadensis]gi 1228013992 ref XP_0219		XP_021					
			38963.1 beta-glucuronidase		938962,					
			[Zootermopsis		XP_021					
Lep_00	beta-		nevadensis]gi 1228013994 ref XP_0219		938963,	3.2	79.0			
03971	glucuroni		38964.1 beta-glucuronidase	Blatt	XP_021	4E-	697	65.8		
3-RA	dase-like	414	[Zootermopsis nevadensis]	odea	938964	11	7	55	43	34
	beta-		gi 241309940 ref XP_002407819.1 beta-							
	glucuroni		glucuronidase (GusB), putative [Ixodes		XP_002					
Lep_00	dase		scapularis]gi 215497226 gb EEC06720.1	Arac	407819,	7.1	79.0			
03971	(GusB),		beta-glucuronidase (GusB), putative	hnid	EEC067	1E-	322	95.9		
4-RA	putative	258	[Ixodes scapularis]	а	20	24	6	005	62	49

Seque	Sequence	Seq	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
nce	desc.	ue			ACC	Valu	larit	Scor	ment	itiv
name		nce				е	У	е	lengt	es
		len							h	
		gth								
	beta-									
Lep_0	glucuronid		gi 646696603 gb KDR08779.1 Beta-				75.3			
00397	ase		glucuronidase [Zootermopsis	Blatto	KDR0	8.4E-	846	90.1		
15-RA	isoform X3	621	nevadensis]	dea	8779	19	2	225	65	49
			gi 1339060575 ref XP_023718875.1 bet							
			a-glucuronidase-like isoform X1		XP_02					
			[Cryptotermes		37188					
Lep_0	beta-		secundus]gi 1330889851 gb PNF22049.		75,		65.7			
00221	glucuronid	220	1 Beta-glucuronidase [Cryptotermes	Blatto	PNF2		575	657.		
44-RA	ase-like	2	secundus]	dea	2049	0	8	136	660	434
			gi 1339060579 ref XP_023718877.1 bet							
			a-glucuronidase-like isoform X3		XP_02					
	beta-		[Cryptotermes		37188					
Lep_0	glucuronid		secundus]gi 1330889852 gb PNF22050.		77,		73.6			
00464	ase-like	116	1 hypothetical protein B7P43_G09739	Blatto	PNF2	1E-	526	377.		
31-RA	isoform X3	1	[Cryptotermes secundus]	dea	2050	124	9	096	334	246
			gi 1339060579 ref XP_023718877.1 bet							
			a-glucuronidase-like isoform X3		XP_02					
			[Cryptotermes		37188					
Lep_0	beta-		secundus]gi 1330889852 gb PNF22050.		77,		69.7			
00286	glucuronid		1 hypothetical protein B7P43_G09739	Blatto	PNF2	4.1E-	777	233.		
14-RA	ase-like	738	[Cryptotermes secundus]	dea	2050	71	8	032	225	157

Table 3.15. Continue	d.
----------------------	----

Seque	Sequence desc	Seque	Hit desc.	Order	Hit ACC	E- Valu	Sim ilar	Bit- Scor	Align ment	Pos itiv
name	ucse.	lengt			nee	e	ity	e	lengt	es
	beta-									
Lep 0	glucuronid		gi 1316206353 ref XP 023213588.1 bet		XP 02		85.			
00286	ase-like		a-glucuronidase-like [Centruroides	Arthr	32135	3.42E	849	162.		
13-RA	isoform X3	750	sculpturatus]	opoda	88	-47	06	925	106	91
	beta-									
Lep_0	glucuronid		gi 1233178675 ref XP_022190762.1 bet		XP_02		87.			
01305	ase-like		a-glucuronidase-like [Nilaparvata	Hemip	21907	3.27E	301	96.6		
36-RA	isoform X1	195	lugens]	tera	62	-23	59	709	63	55
			gi 1339060579 ref XP_023718877.1 bet							
			a-glucuronidase-like isoform X3		XP_02					
	beta-		[Cryptotermes		37188					
Lep_0	glucuronid		secundus]gi 1330889852 gb PNF22050.		77,		67.			
00210	ase-like		1 hypothetical protein B7P43_G09739	Blatto	PNF2	1.89E	582	171.		
05-RA	isoform X1	564	[Cryptotermes secundus]	dea	2050	-48	42	4	182	123
	beta-									
Lep_0	glucuronid		gi 1316206353 ref XP_023213588.1 bet		XP_02					
00210	ase-like		a-glucuronidase-like [Centruroides	Arthr	32135	6.62E		243.		
03-RA	isoform X3	726	sculpturatus]	opoda	88	-79	75	817	192	144
			gi 968019329 ref XP_015011863.1 unc							
			haracterized protein Dere_GG20922,		XP_01					
	beta-		isoform C [Drosophila		50118					
Lep_0	glucuronid		erecta]gi 945202794 gb KQS62122.1 u		63,		72.			
00210	ase		ncharacterized protein Dere_GG20922,	Dipter	KQS6	2.97E	881	71.2		
04-RA	isoform X1	375	isoform C [Drosophila erecta]	а	2122	-13	36	478	59	43

Table 3.15. Continued

Sequ	Sequen	Seque	Hit desc.	Ord	Hit ACC	Е-	Sim	Bit-	Align	Pos
ence	ce desc.	nce		er		Value	ilar	Sco	ment	itiv
nam		lengt					ity	re	lengt	es
е		h							h	
			gi 1227991600 ref XP_021927576.1 beta-							
			glucuronidase-like [Zootermopsis		XP_021					
			nevadensis]gi 1227991602 ref XP_0219275		927576,					
			78.1 beta-glucuronidase-like [Zootermopsis		XP_021					
			nevadensis]gi 1227991604 ref XP_0219275		927578,					
			79.1 beta-glucuronidase-like [Zootermopsis		XP_021					
			nevadensis]gi 1227991606 ref XP_0219275		927579,					
			80.1 beta-glucuronidase-like [Zootermopsis		XP_021					
			nevadensis]gi 1227991608 ref XP_0219275		927580,					
			81.1 beta-glucuronidase-like [Zootermopsis		XP_021					
			nevadensis]gi 1227991610 ref XP_0219275		927581,					
			82.1 beta-glucuronidase-like [Zootermopsis		XP_021					
			nevadensis]gi 1227991612 ref XP_0219275		927582,					
			83.1 beta-glucuronidase-like [Zootermopsis		XP_021					
			nevadensis]gi 1227991614 ref XP_0219275		927583,					
Lep_	beta-		84.1 beta-glucuronidase-like [Zootermopsis		XP_021					
0002	glucuro		nevadensis]gi 646709721 gb KDR15442.1		927584,		64.	314		
2123	nidase-		Beta-glucuronidase [Zootermopsis	Blatt	KDR154	2.8E-	506	.30		
-RA	like	1026	nevadensis]	odea	42	100	17	9	324	209
Lep_										
0002	Beta-		gi 1228362729 ref XP_021967946.1 beta-	Coll			76.	376		
2122	glucuro		glucuronidase-like isoform X1 [Folsomia	emb	XP_021	1.1E-	978	.71		
-RA	nidase	1276	candida]	ola	967946	122	42	1	278	214

Sequ	Sequen	Seque	Hit desc.	Ord	Hit ACC	E-	Sim	Bit-	Align	Pos
ence	ce desc.	nce		er		Value	ilar	Sco	ment	itiv
nam		lengt					ity	re	lengt	es
е		h							h	
	beta-									
Lep_	glucuro									
0002	nidase		gi 1330889853 gb PNF22051.1 hypothetic				69.	204		
2813	isoform		al protein B7P43_G09739 [Cryptotermes	Blatt	PNF220	9.44E-	832	.91		
-RA	X3	609	secundus]	odea	51	64	4	2	179	125
	beta-									
	glucuro									
Lep_	nidase-									
0002	like		gi 641675506 ref XP_008187067.1 PREDIC	Hem			72.			
2814	isoform		TED: beta-glucuronidase isoform X2	ipter	XP_008	4.85E-	602	172		
-RA	X3	612	[Acyrthosiphon pisum]	a	187067	48	74	.17	146	106
Lep_	beta-									
0002	glucuro		gi 1316153756 ref XP_023221887.1 beta-	Arth			71.			
4190	nidase-		glucuronidase-like [Centruroides	ropo	XP_023	1.14E-	985	297		
-RA	like	1134	sculpturatus]	da	221887	94	82	.36	282	203
	beta-									
Lep	glucuro									
0002	nidase						76.	219		
4189	isoform		gi 1321309229 ref XP 023296676.1 beta-	Dipt	XP 023	7.51E-	506	.93		
-RA	X1	621	glucuronidase isoform X2 [Lucilia cuprina]	era	296676	66	02	5	166	127

Seque	Seq	Sequen	Hit desc.	Or	Hit	E-	Sim	Bit-	Align	Pos
nce name	ce	length		aer	ALL	e valu	ity	re	lengt	es
	desc								h	
	•			Coll						
Len ()	Malt			em	ОРМ		79	98		
00578	ase		gil1061480896/gbl0DM982891/Maltase1	bol	9828	4 2 8	166	211		
45-RA	1	1110	[Orchesella cincta]	a	9	E-20	67	7	72	57
			[]	-	XP 0			_		
			gi 1339059590 ref XP_023718348.1 maltase 2-		2371					
Lep_0	Malt		like [Cryptotermes	Blat	8348,		75.	273		
01106	ase		secundus]gi 1330891147 gb PNF22635.1 Maltas	tod	PNF2	2.88	462	.86		
59-RA	1	651	e 2 [Cryptotermes secundus]	ea	2635	E-87	96	3	216	163
					XP_0					
	malt		gi 1339059590 ref XP_023718348.1 maltase 2-		2371					
Lep_0	ase		like [Cryptotermes	Blat	8348,		76.	155		
00910	2-		secundus]gi 1330891147 gb PNF22635.1 Maltas	tod	PNF2	6.51	033	.99		
06-RA	like	522	e 2 [Cryptotermes secundus]	ea	2635	E-43	06	2	121	92
			gi 1339059399 ref XP_023718249.1 maltase 2-		XP_0					
			like isoform X3 [Cryptotermes		2371					
			secundus]gi 1330891182 gb PNF22667.1 hypoth		8249,					
			etical protein B7P43_G07129 [Cryptotermes		PNF2					
Lep_0	Malt		secundus]gi 1330891184 gb PNF22669.1 hypoth	Blat	2667,		78.	192		
00967	ase		etical protein B7P43_G07129 [Cryptotermes	tod	PNF2	1.48	832	.58		
11-RA	1	648	secundus]	ea	2669	E-56	12	6	137	108

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

Seque	Seq	Sequen	Hit desc.	Ord	Hit ACC	E-	Sim	Bit-	Align	Pos
nce	uen	ce		er		Valu	ilar	Sco	ment	itiv
name	се	length				е	ity	re	lengt	es
	desc								h	
	•									
	malt		gi 1339059590 ref XP_023718348.1 maltase		XP_023					
Lep_0	ase		2-like [Cryptotermes		718348,		70.	231		
00772	2-		secundus]gi 1330891147 gb PNF22635.1 Mal	Blatt	PNF226	5.56	935	.87		
59-RA	like	834	tase 2 [Cryptotermes secundus]	odea	35	E-70	96	6	203	144
	sucr									
	ase-									
	isom									
	altas									
	е,									
Lep_0	intes		gi 1067098296 ref XP_018011173.1 PREDICT	Crus			61.			
00354	tinal		ED: maltase-glucoamylase, intestinal-like	tace	XP_018	2.66	370	220		
21-RA	-like	1707	[Hyalella azteca]	а	011173	E-59	72	.32	321	197
Lep_0	Malt						60.	375		
00349	ase		gi 1330891183 gb PNF22668.1 Maltase 1	Blatt	PNF226	1.1E	633	.94		
30-RA	1	1410	[Cryptotermes secundus]	odea	68	-122	48	1	442	268
Lep_0	Malt						64.	107		
01139	ase		gi 1228019306 ref XP_021941640.1 maltase	Blatt	XP_021	5.24	285	.84		
97-RA	1	615	2-like isoform X1 [Zootermopsis nevadensis]	odea	941640	E-25	71	2	98	63
			gi 1339059590 ref XP_023718348.1 maltase		XP_023					
Lep_0	Malt		2-like [Cryptotermes		718348,		75.	280		
00623	ase		secundus]gi 1330891147 gb PNF22635.1 Mal	Blatt	PNF226	4.13	799	.41		
57-RA	1	759	tase 2 [Cryptotermes secundus]	odea	35	E-89	09	1	219	166

Seque	Seq	Sequen	Hit desc.	Ord	Hit ACC	Е-	Sim	Bit-	Align	Pos
nce	uen	ce		er		Valu	ilar	Sco	ment	itiv
name	се	length				е	ity	re	lengt	es
	desc								h	
	•									
	Malt									
	ase-									
	gluc									
	oam									
10	ylas			C			40			
Lep_0	e,		gi[106/100065]ref[XP_01801211/.1]PREDICT	Crus	VD 010	1 4 5	4Z.	71		
00332	intes	570	ED: sucrase-isomaltase, intestinal-like	tace	XP_018	1.45	016	/1.	110	50
14-RA	tinal	570	[Hyalella azteca]	а	01211/	E-12	81	633	119	50
	sucr									
	ase-									
	isom									
	altas									
Law 0	e,			A			71	202		
	intes		gl/6/53/1953/gD/KFM64855.1/Maltase-	Arac	VEMCA	(2)	/1.	202		
	liko	052	giucoamyrase, intestinai, partiai [Stegodyphus	nnia		0.34 E 60	584 7	.21	102	121
IJ-KA	-like	852	mmosarum	a	822	E-00	/	0	165	151
Lep_0	Mait		cil12200E020ElrofIVD 022710247 1 probable	Dlatt	VD 022	1 1 7	/3.	706		
01404 42 DA	ase	102	gi 1339059395 [rel AP_023718247.1 [probable]	Blatt	AP_023	1.1/ E 1(437 F	790	()	47
43-KA	1	193	maitase isoform x1 [Cryptotermes secundus]	ouea	/1824/	E-10	Э	2	04	47
Lop 0	mait		ril1101240010 mof VD 010000021 1 DDEDICT	Hom			07	02		
Lep_0	ase		SILLIUIS49818 [fel]Ar_018900831.1 [PREDICI	inter	VD 010	1 50	0/.	92.		
	2- 1:1-0	270	ED: Maitase A1-like isoform A3 [Bemisia	ipter	AP_018	1.58	03/	048		47
38-KA	пке	3/8	tadacıj	a	900831	E-20	04	5	54	4/

Table 3.16. Co	ntinued.
----------------	----------

Seque	Seq	Sequen	Hit desc.	Ord	Hit	E-	Sim	Bit-	Align	Pos
nce	uen	ce		er	ACC	Val	ilar	Sco	ment	itiv
name	ce	length				ue	ity	re	lengt	es
	desc								h	
	•									
	malt									
	ase-									
	gluc									
	oam									
	ylas									
	e,									
Lep_0	intes			Crus	117040	9.1	60.	143		
00697	tinal	0.00	gi[1022768805]gb]KZS13426.1]Uncharacterize	tace	KZS13	7E-	365	.66	1.64	00
57-RA	-like	969	d protein APZ42_021387 [Daphnia magna]	а	426	36	85	5	164	99
	malt									
	ase									
	1-									
1 0	like			Hym	VD 020	0.0	71	47		
	15010		-: 11152022005 June SVD 020202571 1 June Store 1	eno	XP_020	0.0	/1.	4/.		
	rm	(27	gl[1153833895]ref[XP_020293571.1]maitase 1-	pter	29357	12	428	/50	40	20
35-KA	XZ	627	ike isolorm X2 [Pseudomyrinex graciiis]	a		12	57	0	42	30
	mait		gl 1339059399 ref XP_023718249.1 maitase 2-		XP_023					
	ase		ince isoloi in A5 [Ci ypiolei ines		/1824					
			betical protoin P7D42 C07120 [Comptotormag		ד, סאביי					
Lop 0	isofo		accurdus and 1220901194 and DNE22660 1 humat		FNF22 667	70	71	155		
	15010		betical protoin P7D42 C07120 [Comptotormag	Dlatt	007, DNE22	7.0 0E	/1.	100		
26 DA	1111 V2	207	socundus	odoa	FNF22 660	0E- 11	75	.22	120	01
30-KA	ΛL	30/	secultuusj	ouea	009	44	/3	T	120	71

Table 3.16. Continued

Seque	Sequence	Seque	Hit desc.	Ord	Hit	E-	Sim	Bit-	Align	Pos
nce name	aesc.	nce lengt		er	ALL	vai ue	itv	sco re	lengt	es
		h					5		h	
	maltase-									
	glucoamyla									
Lep_0	se,		gi 1238867836 ref XP_022249217.1 lysos	Arth	XP_022	4.6	70.	126		
00820	intestinal-		omal alpha-glucosidase-like [Limulus	ropo	24921	5E-	707	.33		
89-RA	like	666	polyphemus]	da	7	35	07	1	99	70
	Maltase-		gi 1067072134 ref XP_018019496.1 PRED							
Lep_0	glucoamyla		ICTED: LOW QUALITY PROTEIN: probable	Crus	XP_018	1.1	58.	68.		
00637	se,		maltase-glucoamylase 2, partial [Hyalella	tace	01949	2E-	571	551		
63-RA	intestinal	312	azteca]	а	6	12	43	4	70	41
	maltase-									
	glucoamyla									
Lep_0	se,		gi 1048057520 ref XP_017462059.1 PRED		XP_017	6.3	66.	166		
00637	intestinal-		ICTED: maltase-glucoamylase, intestinal-	Dipt	46205	4E-	447	.00		
64-RA	like	507	like, partial [Rhagoletis zephyria]	era	9	49	37	7	152	101
Lep_0					XP_016	2.5	59.	67.		
00560			gi 1036973924 ref XP_016963459.1 PRED	Dipt	96345	2E-	677	010		
80-RA	maltase 1	279	ICTED: maltase 1 [Drosophila biarmipes]	era	9	12	42	6	62	37
Lep_0				Coll		2.8	78.	100		
00560			gi 1061480896 gb 0DM98289.1 Maltase 1	emb	ODM9	2E-	378	.90		
81-RA	Maltase 1	699	[Orchesella cincta]	ola	8289	22	38	8	74	58
	sucrase-									
Lep_0	isomaltase,		gi 675371953 gb KFM64855.1 Maltase-	Arac		2.4	64.	232		
00868	intestinal-		glucoamylase, intestinal, partial	hnid	KFM64	9E-	682	.64		
62-RA	like	723	[Stegodyphus mimosarum]	а	855	72	54	6	252	163

Seque	Sequence	Seque	Hit desc.	Ord	Hit	E-	Sim	Bit-	Align	Pos
nce	desc.	nce		er	ACC	Val	ilar	Sco	ment	itiv
name		lengt				ue	ity	re	lengt	es
		h							h	
				Hym						
Lep_0				eno	XP_003	6.1	77.			
01247	maltase 2-		gi 1185574477 ref XP_003395914.2 malta	pter	39591	1E-	358	79.		
32-RA	like	234	se 1 [Bombus terrestris]	а	4	17	49	337	53	41
Lep_0			gi 1227980746 ref XP_021921963.1 unch		XP_021	2.5	66.	154		
00051			aracterized protein KIAA0513 isoform X3	Blatt	92196	1E-	839	.83		
24-RA	Maltase 1	678	[Zootermopsis nevadensis]	odea	3	42	38	6	193	129
				Hym						
Lep_0			gi 1317983107 ref XP_012271316.2 unch	eno	XP_012	3.4		170		
00430	maltase 2		aracterized protein LOC105694836	pter	27131	8E-	79.	.24		
03-RA	isoform X1	918	[Orussus abietinus]	a	6	45	2	4	125	99
					XP_018					
					91758					
			gi 1101338755 ref XP_018917587.1 PRED		7,					
			ICTED: maltase 2-like [Bemisia		XP_018					
			tabaci]gi 1101338757 ref XP_018917588.		91758					
			1 PREDICTED: maltase 2-like [Bemisia		8,					
Lep_0			tabaci]gi 1101338759 ref XP_018917589.	Hem	XP_018	1.0	75.	86.		
00430	maltase 2-		1 PREDICTED: maltase 2-like [Bemisia	ipter	91758	8E-	757	270		
02-RA	like	567	tabaci]	a	9	17	58	5	66	50

Seque	Sequence	Seque	Hit desc.	Ord	Hit	E-	Sim	Bit-	Align	Pos
nce	desc.	nce		er	ACC	Val	ilar	Sco	ment	itiv
name		lengt				ue	ity	re	lengt	es
		h							h	
			gi 1339059399 ref XP_023718249.1 malta							
			se 2-like isoform X3 [Cryptotermes		XP_023					
			secundus]gi 1330891182 gb PNF22667.1		71824					
			hypothetical protein B7P43_G07129		9,					
			[Cryptotermes		PNF22					
Lep_0			secundus]gi 1330891184 gb PNF22669.1		667,	2.5	70.			
00060	maltase 2-		hypothetical protein B7P43_G07129	Blatt	PNF22	1E-	848	317		
58-RA	like	2673	[Cryptotermes secundus]	odea	669	96	71	.39	271	192
	Maltase-									
Lep_0	glucoamyla		gi 1067098296 ref XP_018011173.1 PRED	Crus	XP_018	3.4	67.	339		
00688	se,		ICTED: maltase-glucoamylase, intestinal-	tace	01117	E-	320	.34		
67-RA	intestinal	915	like [Hyalella azteca]	а	3	105	26	7	306	206
	sucrase-									
Lep_0	isomaltase,		gi 1067098296 ref XP_018011173.1 PRED	Crus	XP_018		67.			
00114	intestinal-		ICTED: maltase-glucoamylase, intestinal-	tace	01117		084	596		
65-RA	like	1767	like [Hyalella azteca]	а	3	0	08	.66	559	375
	maltase-									
Lep_0	glucoamyla		gi 321476731 gb EFX87691.1 hypothetica	Crus		1.6	57.	410		
00454	se,		l protein DAPPUDRAFT_306567 [Daphnia	tace	EFX87	E-	283	.22		
79-RA	intestinal	1600	pulex]	а	691	133	46	3	508	291
Lep_0						3.5	81.	100		
01081			gi 646696837 gb KDR08865.1 Maltase 1	Blatt	KDR08	4E-	967	.52		
34-RA	Maltase 1	441	[Zootermopsis nevadensis]	odea	865	23	21	3	61	50

Seque nce name	Sequence desc.	Seque nce lengt h	Hit desc.	Ord er	Hit ACC	E- Val ue	Sim ilar ity	Bit- Sco re	Align ment lengt h	Pos itiv es
	sucrase-									
Lep_0	isomaltase,		gi 1067098296 ref XP_018011173.1 PRED	Crus	XP_018	1.5	63.	416		
00362	intestinal-		ICTED: maltase-glucoamylase, intestinal-	tace	01117	E-	636	.77		
31-RA	like	1596	like [Hyalella azteca]	а	3	130	36	2	451	287
Lep_0					XP_021	2.7	84.	97.		
00710			gi 1228013734 ref XP_021938828.1 malta	Blatt	93882	1E-	745	826		
68-RA	Maltase 1	672	se 2-like [Zootermopsis nevadensis]	odea	8	21	76	5	59	50

Seque	Seque	Seque	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
nce	nce	nce		r		Val	larit	Scor	ment	itiv
name	desc.	lengt				ue	У	е	lengt	es
		h							h	
			gi 1339059399 ref XP_023718249.1 malt							
			ase 2-like isoform X3 [Cryptotermes							
			secundus]gi 1330891182 gb PNF22667.1		XP_023					
			hypothetical protein B7P43_G07129		718249,					
			[Cryptotermes		PNF226					
Lep_00	alpha		secundus]gi 1330891184 gb PNF22669.1		67,	8.6	70.5			
08229	amyla		hypothetical protein B7P43_G07129	Blatt	PNF226	1E-	882	218.0		
3-RA	se	573	[Cryptotermes secundus]	odea	69	67	4	09	187	132
	alpha-									
	amyla									
Lep_00	se-like		gi 1000763879 ref XP_015605017.1 PRE	Hym		8.1	80.1			
01835	isofor		DICTED: alpha-amylase 1-like [Cephus	enopt	XP_015	6E-	980	143.6		
7-RA	m X2	423	cinctus]	era	605017	39	2	65	101	81
	alpha-									
Lep_00	amyla					2.7	93.8			
01835	se 2-		gi 540849793 gb AGV15452.1 alpha-	Blatt	AGV154	5E-	461	123.2		
9-RA	like	1356	amylase, partial [Blattella germanica]	odea	52	28	5	5	65	61
	alpha-									
	amyla									
Lep_00	se-like		gi 62955866 gb AAY23288.1 1,4-alpha-D-			2.0	76.7			
08174	isofor		glucan glucanohydrolase precursor	Blatt	AAY232	5E-	441	68.16		
9-RA	m X2	567	[Blattella germanica]	odea	88	11	9	62	43	33

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

Table 3.17. Continued.

Seque	Seque	Sequen	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
nce	nce	се		r		Val	larit	Scor	ment	itiv
name	desc.	length				ue	У	e	lengt	es
									h	
Lep_00	alpha-		gi 998506683 ref XP_015513623.1 PR	Hym		2.8	76.4			
03537	amylas		EDICTED: alpha-amylase 2-like	enopt	XP_015	4E-	705	155.9		
5-RA	e A-like	372	[Neodiprion lecontei]	era	513623	44	9	92	102	78
	alpha-									
Lep_00	amylas					1.3	86.3			
03537	e 4N-		gi 68266131 gb AAY88833.1 alpha-	Dipte	AAY888	9E-	636	76.64		
3-RA	like	561	amylase, partial [Coelopa frigida]	ra	33	14	4	06	44	38
Lep_00	Alpha-					1.1	79.8			
08226	amylas		gi 540849793 gb AGV15452.1 alpha-	Blatt	AGV154	3E-	701	230.3		
9-RA	e 1	567	amylase, partial [Blattella germanica]	odea	52	71	3	35	154	123
Lep_00	alpha-		gi 1000763879 ref XP_015605017.1 P	Hym		1.3	81.4			
13172	amylas		REDICTED: alpha-amylase 1-like	enopt	XP_015	8E-	814	85.50		
4-RA	e 1-like	315	[Cephus cinctus]	era	605017	18	8	01	54	44
	Pancre									
	atic									
Lep_00	alpha-		gi 62955866 gb AAY23288.1 1,4-			8.5	75.6			
07579	amylas		alpha-D-glucan glucanohydrolase	Blatt	AAY232	2E-	097	63.92		
3-RA	е	642	precursor [Blattella germanica]	odea	88	10	6	9	41	31
	Pancre									
	atic									
Lep_00	alpha-			Colle		3.4	69.3			
06260	amylas		gi 1061479081 gb 0DM96628.1 Pancr	mbol	ODM96	8E-	069	111.3		
2-RA	е	663	eatic alpha-amylase [Orchesella cincta]	а	628	26	3	09	101	70

Table 3.17. Continu	led.
---------------------	------

Seque	Seque	Sequen	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
nce	nce	се		r		Val	larit	Scor	ment	itiv
name	desc.	length				ue	У	е	lengt	es
									h	
			gi 602182726 gb AHN91843.1 alpha-							
			amylase [Marsupenaeus		AHN91					
Lep_00	alpha-		japonicus]gi 602182728 gb AHN91844		843,	2.0				
09136	amylas		.1 alpha-amylase [Marsupenaeus	Crust	AHN91	9E-		92.43		
7-RA	e 1-like	252	japonicus]	acea	844	21	86	37	50	43
	alpha-									
	amylas									
Lep_00	e 1-like		gi 1067066079 ref XP_018017054.1 P			1.7	82.4			
08218	isofor		REDICTED: alpha-amylase 1-like	Crust	XP_018	7E-	742	137.5		
3-RA	m X1	702	isoform X1 [Hyalella azteca]	acea	017054	35	3	02	97	80
Lep_00	alpha-					3.7	81.6			
11150	amylas		gi 540849793 gb AGV15452.1 alpha-	Blatt	AGV154	2E-	513	167.1		
9-RA	e 1-like	600	amylase, partial [Blattella germanica]	odea	52	47	8	62	109	89
Lep_00	Alpha-					2.3	77.7			
08330	amylas		gi 85002763 gb ABC68516.1 alpha-	Blatt	ABC685	E-	777	109.3		
3-RA	e 1	687	amylase [Blattella germanica]	odea	16	25	8	83	81	63
Lep_00	alpha-		gi 998506681 ref XP_015513622.1 PR	Hym		6.7	81.5			
09919	amylas		EDICTED: alpha-amylase A-like	enopt	XP_015	1E-	217	147.9		
4-RA	e A-like	519	[Neodiprion lecontei]	era	513622	42	4	02	92	75
			gi 602182726 gb AHN91843.1 alpha-							
			amylase [Marsupenaeus		AHN91					
Lep_00	alpha-		japonicus]gi 602182728 gb AHN91844		843,	2.1				
11447	amylas		.1 alpha-amylase [Marsupenaeus	Crust	AHN91	6E-		91.27		
1-RA	e 1-like	366	japonicus]	acea	844	20	86	81	50	43

Table	3.17.	Continued.
Table	3.17.	continuea.

Seque	Seque	Sequen	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
nce	nce	ce		r		Val	larit	Scor	ment	itiv
name	desc.	length				ue	У	е	lengt	es
									h	
Lep_00	alpha-		gi 1000764249 ref XP_015605210.1 P	Hym		7.6				
10354	amylas		REDICTED: alpha-amylase 1-like	enopt	XP_015	9E-		72.40		
9-RA	e 1-like	354	[Cephus cinctus]	era	605210	14	78	34	50	39
Lep_00	alpha-					3.8	76.1			
11379	amylas		gi 68266167 gb AAY88846.1 alpha-	Dipte	AAY888	8E-	061	149.8		
4-RA	e A-like	378	amylase, partial [Musca domestica]	ra	46	45	9	28	113	86
			gi 1339092808 ref XP_023707620.1 al							
			pha-amylase 1-like [Cryptotermes		XP_023					
Lep_00	alpha-		secundus]gi 1330914763 gb PNF3359		707620,	1.8	78.8			
09364	amylas		7.1 Alpha-amylase 1 [Cryptotermes	Blatt	PNF335	3E-	732	118.6		
7-RA	e 1-like	738	secundus]	odea	97	28	4	27	71	56
Lep_00	alpha-			Zyge		9.8	88.4			
13334	amylas		gi 168830277 gb ACA34383.1 alpha-	ntom	ACA343	6E-	615	96.28		
7-RA	e 1-like	261	amylase, partial [Lepisma saccharina]	а	83	26	4	57	52	46
Lep_00	alpha-		gi 168830254 gb ACA34371.1 alpha-			1.8	73.9			
13139	amylas		amylase, partial [Periplaneta	Blatt	ACA343	7E-	130	154.4		
0-RA	e 1-like	354	americana]	odea	71	47	4	51	115	85
Lep_00	alpha-		gi 1227981869 ref XP_021922548.1 al			1.4	71.3			
03108	amylas		pha-amylase 2-like [Zootermopsis	Blatt	XP_021	2E-	656	276.5		
1-RA	e 1-like	1584	nevadensis]	odea	922548	84	4	59	227	162
Lep_00	alpha-					1.7	78.1			
11268	amylas		gi 139478961 gb AB077639.1 alpha-	Dipte	AB0776	3E-	690	202.6		
7-RA	e A-like	513	amylase, partial [Musca domestica]	ra	39	64	1	01	142	111

Sequenc e name	Sequen	Sequ ence	Hit desc.	Order	Hit ACC	E- Val	Simil arity	Bit- Scor	Align ment	Posit ives
e nume		lengt				ue	unity	e	lengt	1100
		h							h	
			gi 1227958457 ref XP_02192690							
			1.1 alpha-N-							
			acetylgalactosaminidase-like							
			[Zootermopsis							
			nevadensis]gi 1227958459 ref X							
			P_021926910.1 alpha-N-							
			acetylgalactosaminidase-like		XP_0219					
			[Zootermopsis		26901,					
	alpha-		nevadensis]gi 646723001 gb KD		XP_0219					
Lep_000	galactos		R23773.1 Alpha-N-		26910,	3.5				
86810-	idase A-		acetylgalactosaminidase	Blatto	KDR237	4E-		167.		
RA	like	609	[Zootermopsis nevadensis]	dea	73	48	80	933	110	88
	beta-									
	galactos									
	idase-1-									
Lep_001	like		gi 1228405415 ref XP_02196183			1.2				
27369-	protein		1.1 beta-galactosidase-1-like	Colle	XP_0219	2E-	73.97	92.8		
RA	2	324	protein 2 [Folsomia candida]	mbola	61831	23	26	189	73	54
			gi 1339081983 ref XP_02370200							
Lep_000	Beta-		9.1 LOW QUALITY PROTEIN:							
26440-	galactos		beta-galactosidase-like	Blatto	XP_0237		70.03	685.		
RA	idase	2145	[Cryptotermes secundus]	dea	02009	0	106	256	644	451

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

Sequenc	Sequen	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posit
e name	ce desc.	ence				Val	arity	Scor	ment	ives
		lengt				ue	_	е	lengt	
		h							h	
			gi 1000751440 ref XP_01559857							
			4.1 PREDICTED: beta-							
			galactosidase-1-like protein 2							
			isoform X2 [Cephus							
			cinctus]gi 1000751442 ref XP_01							
			5598576.1 PREDICTED: beta-							
			galactosidase-1-like protein 2							
			isoform X2 [Cephus							
			cinctus]gi 1000751444 ref XP_01							
			5598577.1 PREDICTED: beta-							
			galactosidase-1-like protein 2							
			isoform X2 [Cephus							
			cinctus]gi 1000751446 ref XP_01		XP_0155					
			5598578.1 PREDICTED: beta-		98574,					
			galactosidase-1-like protein 2		XP_0155					
			isoform X2 [Cephus		98576,					
			cinctus]gi 1000751448 ref XP_01		XP_0155					
			5598579.1 PREDICTED: beta-		98577,					
	beta-		galactosidase-1-like protein 2		XP_0155					
	galactos		isoform X2 [Cephus		98578,					
	idase-1-		cinctus]gi 1000751450 ref XP_01		XP_0155					
Lep_000	like		5598580.1 PREDICTED: beta-		98579,	2.7				
59101-	protein		galactosidase-1-like protein 2	Hymen	XP_0155	6E-	67.69	230.		
RA	2	1416	isoform X2 [Cephus cinctus]	optera	98580	66	912	335	226	153

Sequenc	Sequenc	Sequenc	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	e length			ALL	valu e	larit v	SCOF e	lengt	es
						č	y	Č	h	0.5
	beta-									
	galactosi									
	dase-1-									
	like									
Lep_001	protein 2		gi 1330893432 gb PNF23708.1 hy				86.8			
23361-	isoform		pothetical protein B7P43_G02480	Blattod	PNF2	1.52	852	108.6		
RA	X2	240	[Cryptotermes secundus]	ea	3708	E-27	5	12	61	53
Lep_000	Beta-		gi 1227962476 ref XP_021942674		XP_0		65.0			
11817-	galactosi		.1 beta-galactosidase isoform X2	Blattod	2194	1.2E	862	445.6		
RA	dase	1926	[Zootermopsis nevadensis]	ea	2674	-146	1	62	464	302
	beta-									
	galactosi									
Lep_000	dase		gi 1227962478 ref XP_021942675		XP_0		84.6			
11816-	isoform		.1 beta-galactosidase isoform X3	Blattod	2194	3.56	153	167.5		
RA	X3	738	[Zootermopsis nevadensis]	ea	2675	E-46	8	48	104	88
	beta-									
	galactosi									
	dase-1-									
	like		gi 1000751438 ref XP_015598573							
Lep_001	protein 2		.1 PREDICTED: beta-galactosidase-		XP_0		86.8			
66804-	isoform		1-like protein 2 isoform X1	Hymeno	1559	5.02	852	107.4		
RA	X2	189	[Cephus cinctus]	ptera	8573	E-27	5	57	61	53
Table 3.18. Continued	l.									
-----------------------	----									
-----------------------	----									

Sequenc e name	Sequenc e desc.	Sequenc e length	Hit desc.	Order	Hit ACC	E- Valu e	Simi larit y	Bit- Scor e	Align ment lengt h	Pos itiv es
	Beta- galactosi									
Lep_000	dase-1-		gi 1215258369 gb 0XA44696.1 Be							
06759-	like		ta-galactosidase-1-like protein 2	Collemb	OXA4	1.77		144.0		
RA	protein 2	814	[Folsomia candida]	ola	4696	E-39	55	5	200	110

Table 3.18. Continued

Sequenc	Sequenc	Sequenc	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	e length			ACC	Valu	larit	Scor	ment	itiv
						е	У	е	lengt	es
									h	
			gi 1000751440 ref XP_015598574							
			.1 PREDICTED: beta-galactosidase-							
			1-like protein 2 isoform X2							
			[Cephus							
			cinctus]gi 1000751442 ref XP_015							
			598576.1 PREDICTED: beta-							
			galactosidase-1-like protein 2		XP_0					
			isoform X2 [Cephus		1559					
			cinctus]gi 1000751444 ref XP_015		8574,					
			598577.1 PREDICTED: beta-		XP_0					
			galactosidase-1-like protein 2		1559					
			isoform X2 [Cephus		8576,					
			cinctus]gi 1000751446 ref XP_015		XP_0					
			598578.1 PREDICTED: beta-		1559					
			galactosidase-1-like protein 2		8577,					
			isoform X2 [Cephus		XP_0					
			cinctus]gi 1000751448 ref XP_015		1559					
			598579.1 PREDICTED: beta-		8578,					
			galactosidase-1-like protein 2		XP_0					
	beta-		isoform X2 [Cephus		1559					
	galactosi		cinctus]gi 1000751450 ref XP_015		8579,					
Lep_000	dase-1-		598580.1 PREDICTED: beta-		XP_0		72.1			
06757-	like		galactosidase-1-like protein 2	Hymeno	1559	1.19	238	229.1		
RA	protein 2	2064	isoform X2 [Cephus cinctus]	ptera	8580	E-63	9	8	226	163

Sequenc	Sequenc	Sequenc	Hit desc.	Order	Hit	E- Valu	Simi	Bit-	Align	Pos
ename	e uesc.	e length			ALL	e valu	y arti	e	lengt	es
									h	
	beta-									
	galactosi									
Lep_000	dase		gi 1330925577 gb PNF39271.1 hy				56.8			
28717-	isoform		pothetical protein B7P43_G16708	Blattod	PNF3	3.01	965	306.6		100
RA	X3	1014	[Cryptotermes secundus]	ea	9271	E-98	5	05	348	198
Lep_000	Beta-			a .		0 = 1	61.0	4.40.0		
28716-	galactosi	720	gi 1022758129 gb KZS04700.1 Be	Crustac	KZS0	3.54	778	140.9	1.68	100
RA	dase	729	ta-galactosidase [Daphnia magna]	ea	4700	E-36	4	69	167	102
Lep_001	beta-		gi 675376294 gb KFM69196.1 Bet		KFM		62.3	1044		
17981-	galactosi		a-galactosidase, partial	Arachni	6919	3.48	931	106.6	445	-
RA	dase-like	1149	[Stegodyphus mimosarum]	da	6	E-24	6	86	117	73
Lep_000	beta-		gi 970886870 ref XP_015108515.		XP_0		82.3			
67905-	galactosi		1 PREDICTED: beta-galactosidase	Hymeno	1510	3.98	529	88.96	_	
RA	dase	699	[Diachasma alloeum]	ptera	8515	E-18	4	69	51	42
	beta-									
	galactosi									
	dase-1-									
	like									
Lep_001	protein 2		gi 1330893432 gb PNF23708.1 hy				83.4			
26791-	isoform		pothetical protein B7P43_G02480	Blattod	PNF2	1.25	645	203.3		
RA	X2	393	[Cryptotermes secundus]	ea	3708	E-62	7	71	127	106
Lep_000	Beta-		gi 1330925577 gb PNF39271.1 hy				73.5			
26029-	galactosi		pothetical protein B7P43_G16708	Blattod	PNF3	2.03	751	222.2		
RA	dase	852	[Cryptotermes secundus]	ea	9271	E-66	3	46	193	142

Table 3.18. Con	ntinued.
-----------------	----------

Sequenc	Sequenc	Sequenc	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	e length			ACC	Valu	larit	Scor	ment	itiv
						e	У	e	lengt	es
									h	
			gi 1227962474 ref XP_021942673							
			.1 beta-galactosidase isoform X1		XP_0					
			[Zootermopsis		2194					
Lep_000	Beta-		nevadensis]gi 646721601 gb KDR		2673,		61.1			
26030-	galactosi		22879.1 Beta-galactosidase	Blattod	KDR2	3.23	940	122.0		
RA	dase	798	[Zootermopsis nevadensis]	ea	2879	E-29	3	94	134	82
	Beta-									
	galactosi		gi 1339057562 ref XP_023717284							
Lep_000	dase-1-		.1 beta-galactosidase-1-like		XP_0		56.3			
08899-	like		protein 2 isoform X2	Blattod	2371		564	579.7		
RA	protein 2	2361	[Cryptotermes secundus]	ea	7284	0	9	11	763	430

Table 3.18. Continued

Sequenc	Sequenc	Sequenc	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	e length			ACC	Valu	larit	Scor	ment	itiv
						е	У	е	lengt	es
									h	
			gi 1000751440 ref XP_015598574							
			.1 PREDICTED: beta-galactosidase-							
			1-like protein 2 isoform X2							
			[Cephus							
			cinctus]gi 1000751442 ref XP_015							
			598576.1 PREDICTED: beta-							
			galactosidase-1-like protein 2		XP_0					
			isoform X2 [Cephus		1559					
			cinctus]gi 1000751444 ref XP_015		8574,					
			598577.1 PREDICTED: beta-		XP_0					
			galactosidase-1-like protein 2		1559					
			isoform X2 [Cephus		8576,					
			cinctus]gi 1000751446 ref XP_015		XP_0					
			598578.1 PREDICTED: beta-		1559					
			galactosidase-1-like protein 2		8577,					
			isoform X2 [Cephus		XP_0					
			cinctus]gi 1000751448 ref XP_015		1559					
			598579.1 PREDICTED: beta-		8578,					
			galactosidase-1-like protein 2		XP_0					
	beta-		isoform X2 [Cephus		1559					
	galactosi		cinctus]gi 1000751450 ref XP_015		8579,					
Lep_000	dase-1-		598580.1 PREDICTED: beta-		XP_0		56.6			
36877-	like		galactosidase-1-like protein 2	Hymeno	1559	8.9E	552	213.3		
RA	protein 2	954	isoform X2 [Cephus cinctus]	ptera	8580	-62	9	86	293	166

Table 3.18. Continued

Sequenc e name	Sequenc e desc.	Sequenc e length	Hit desc.	Order	Hit ACC	E- Valu	Simi larit	Bit- Scor	Align ment	Pos itiv
		8				e	y	e	lengt h	es
	beta-									
Lon 000	galactosi		gi 1000751438 ref XP_015598573				72.2			
26876-	liko		1-like protein 2 isoform X1	Hymono	ΛΡ_U 1550	5 28	72.5	122.0		
RA	nrotein 2	405	[Cenhus cinctus]	ntera	8573	5.20 E-31	3	94	112	81
101		105	gi 1339081983 ref XP 023702009	pteru	0070	1.51	5	<i>у</i> т	112	
Lep 000	Beta-		.1 LOW QUALITY PROTEIN: beta-		XP 0		68.3			
02666-	galactosi		galactosidase-like [Cryptotermes	Blattod	2370	6.8E	301	511.5		
RA	dase	1782	secundus]	ea	2009	-173	3	31	521	356
	beta-									
	galactosi									
Lep_000	dase-like		gi 1227962478 ref XP_021942675		XP_0		84.6			
02668-	isoform		.1 beta-galactosidase isoform X3	Blattod	2194	1.05	153	167.9		
RA	X1	474	[Zootermopsis nevadensis]	ea	2675	E-47	8	33	104	88
	beta-		~110007514201maflyD 015500572							
Lop 000	galactosi		1 DDEDICTED: hota galactoridada		VD O		601			
02498-	liko		1-like protein 2 isoform X1	Hymono	1559		210	659.8		
RA	nrotein 2	2398	[Cenhus cinctus]	ntera	8573	0	5	33	684	468
	heta-	2370		ptera	0373	U	5	55	001	100
	galactosi		gil998516729 ref XP 015519163.							
Lep 000	dase-1-		1 PREDICTED: beta-galactosidase-		XP 0		57.5			
02499-	like		1-like protein 2 [Neodiprion	Hymeno	1551	2.9E	091	407.5		
RA	protein 2	2103	lecontei]	ptera	9163	-131	6	27	546	314

Sequenc e name	Sequenc e desc.	Sequenc e length	Hit desc.	Order	Hit ACC	E- Valu	Simi larit	Bit- Scor	Align ment	Pos itiv
						е	У	e	lengt h	es
	beta-									
	galactosi									
	dase-1-									
	like									
Lep_000	protein 2		gi 1330893432 gb PNF23708.1 hy				84.1			
02500-	isoform		pothetical protein B7P43_G02480	Blattod	PNF2	5.46	269	201.8		
RA	X2	786	[Cryptotermes secundus]	ea	3708	E-60	8	3	126	106
	beta-									
	galactosi									
Lep_001	dase-like		gi 675376294 gb KFM69196.1 Bet		KFM		71.9			
18693-	isoform		a-galactosidase, partial	Arachni	6919	3.34	101	104.7		
RA	X1	601	[Stegodyphus mimosarum]	da	6	E-25	1	6	89	64

Sequenc e name	Seque nce	Seque nce	Hit desc.	Order	Hit ACC	E- Val	Simila rity	Bit- Scor	Align ment	Positi ves
	desc.	lengt b				ue		е	length	
	Endo-	11								
	1,4-									
	beta-		gi 1229883103 ref XP_02216169			1.7				
Lep_0004	xylan		4.1 uncharacterized protein	Hemip	XP_0221	4E-	70.31	60.8		
5817-RA	ase A	231	LOC111027604 [Myzus persicae]	tera	61694	10	25	474	64	45
	Endo-									
	1,4-		gi 391327465 ref XP_003738220.							
	beta-		1 PREDICTED: uncharacterized			1.6				
Lep_0009	xylan		protein LOC100898192	Arach	XP_0037	2E-	58.18	69.7		
7333-RA	ase A	696	[Galendromus occidentalis]	nida	38220	11	182	07	110	64
	Endo-									
	1,4-		gi 391327465 ref XP_003738220.							
	beta-		1 PREDICTED: uncharacterized			5.2				
Lep_0013	xylan		protein LOC100898192	Arach	XP_0037	E-	52.14	72.4		
0240-RA	ase A	546	[Galendromus occidentalis]	nida	38220	13	286	034	140	73

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

Sequenc	Seque	Seque	Hit desc.	Order	Hit ACC	E-	Simila	Bit-	Align	Positi
e name	nce	nce				Val	rity	Scor	ment	ves
	desc.	lengt				ue		е	length	
		h								
			gi 998489397 gb AMK48629.1							
	myrosi		glycoside hydrolase family 1,			7.0				
Lep_0011	nase 1-		partial [Rhynchophorus	Coleopt	AMK486	1E-	70.78	116.		
7435-RA	like	294	ferrugineus]	era	29	33	652	701	89	63
	myrosi		gi 1042851809 gb ANQ90707.			3.1				
Lep_0000	nase 1-		1 beta-glucosidase 2, partial	Lepido	ANQ907	2E-		108.		
4871-RA	like	411	[Manduca quinquemaculata]	ptera	07	29	75	612	84	63
	myrosi		gi 1080062829 ref XP_018573			1.2				
Lep_0000	nase 1-		475.1 myrosinase 1-like	Coleopt	XP_0185	6E-	78.19	185.		
4870-RA	like	537	[Anoplophora glabripennis]	era	73475	54	549	652	133	104
	myrosi		gi 506967929 gb AGM32308.1			7.1				
Lep_0003	nase 1-		beta-glucosidase, partial	Blattod	AGM323	7E-	82.97	72.0		
4695-RA	like	390	[Coptotermes formosanus]	ea	08	14	872	182	47	39
	myrosi		gi 506965870 gb AGM32287.1			4.1				
Lep_0003	nase 1-		beta-glucosidase, partial	Blattod	AGM322	E-	83.90	121.		
4694-RA	like	717	[Coptotermes formosanus]	ea	87	31	805	709	87	73
	myrosi		gi 1069789897 ref XP_018320			2.6				
Lep 0013	nase 1-		188.1 PREDICTED: myrosinase	Coleopt	XP 0183	2E-	88.57	119.		
4073-RA	like	309	1-like [Agrilus planipennis]	era	20188	31	143	398	70	62
	myrosi		gi 1233170929 ref XP_022187			6.9				
Lep_0001	nase 1-		264.1 myrosinase 1-like	Hemipt	XP_0221	1E-	68.08	160.		
9550-RA	like	1608	isoform X2 [Nilaparvata lugens]	era	87264	42	511	999	141	96

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

Sequenc	Seque	Seque	Hit desc.	Order	Hit ACC	E-	Simila	Bit-	Align	Positi
e name	nce	nce				Val	rity	Scor	ment	ves
	desc.	lengt				ue		е	length	
		h								
	myrosi									
	nase 1-									
	like		gi 1229709445 ref XP_022116			1.0				
Lep_0001	isofor		625.1 myrosinase 1-like [Pieris	Lepido	XP_0221	4E-	81.08	59.6		
9549-RA	m X1	279	rapae]	ptera	16625	09	108	918	37	30
	myrosi		gi 506967929 gb AGM32308.1			9.1				
Lep_0002	nase 1-		beta-glucosidase, partial	Blattod	AGM3230	3E-	80.45	193.		
1306-RA	like	564	[Coptotermes formosanus]	ea	8	60	113	356	133	107
			gi 1228411798 ref XP_021965							
			263.1 myrosinase 1-like							
			[Folsomia		XP_0219					
	myrosi		candida]gi 1215254226 gb OX		65263,	7.1				
Lep_0002	nase 1-		A40567.1 Lactase-phlorizin	Collem	OXA4056	9E-	70.17	54.2		
1307-RA	like	288	hydrolase [Folsomia candida]	bola	7	08	544	99	57	40
			gi 1325357344 ref XP_023312			1.5				
Lep_0004	myrosi		869.1 myrosinase 1	Coleopt	XP_0233	E-	73.22	354.		
2970-RA	nase 1	1422	[Anoplophora glabripennis]	era	12869	115	034	369	295	216
	myrosi		gi 636793015 dbj BA085048.1			2.1				
Lep_0004	nase 1-		beta-glucosidase, partial	Blattod	BA08504	7E-	69.14	103.		
8639-RA	like	456	[Salganea esakii]	ea	8	25	894	99	94	65
	myrosi		gi 939630683 ref XP_0142787			1.3				
Lep_0004	nase 1-		11.1 PREDICTED: myrosinase	Hemipt	XP_0142	4E-		79.7		
8640-RA	like	324	1-like [Halyomorpha halys]	era	78711	16	76	221	50	38

Sequenc	Seque	Sequ	Hit desc.	Order	Hit ACC	E-	Simil	Bit-	Align	Posit
e name	nce	ence				Valu	arity	Scor	ment	ives
	desc.	lengt				е		е	lengt	
		h							h	
			gi 1069790020 ref XP_018320243.1							
			PREDICTED: myrosinase 1-like							
			isoform X2 [Agrilus							
			planipennis]gi 1069790022 ref XP_		XP_0183					
Lep_001	myros		018320244.1 PREDICTED:		20243,					
24670-	inase		myrosinase 1-like isoform X3	Coleo	XP_0183	3.14	66.66	73.9		
RA	1-like	261	[Agrilus planipennis]	ptera	20244	E-15	667	442	90	60
			gi 1339073035 ref XP_023723774.1							
			myrosinase 1-like [Cryptotermes		XP_0237					
Lep_000	myros		secundus]gi 1330878957 gb PNF16		23774,					
39278-	inase		841.1 Myrosinase 1 [Cryptotermes	Blatto	PNF168	1E-	59.77	366.		
RA	1-like	1776	secundus]	dea	41	118	011	696	435	260
Lep_001	myros		gi 1080048148 ref XP_018565412.1							
52982-	inase		myrosinase 1 [Anoplophora	Coleo	XP_0185	0.00	62.79	43.8		
RA	1	387	glabripennis]	ptera	65412	094	07	986	43	27
Lep_001	myros		gi 1080062831 ref XP_018573476.1							
14845-	inase		myrosinase 1-like [Anoplophora	Coleo	XP_0185	2.13	68.38	177.		
RA	1-like	537	glabripennis]	ptera	73476	E-51	71	178	155	106
Lep_000	myros		gi 506967929 gb AGM32308.1 beta							
41124-	inase		-glucosidase, partial [Coptotermes	Blatto	AGM323	2.28	85.25	242.		
RA	1-like	1305	formosanus]	dea	08	E-75	641	276	156	133
Lep_001	myros		gi 1133421334 ref XP_019866289.1							
22267-	inase		PREDICTED: myrosinase 1-like	Coleo	XP_0198	1.82	86.41	127.		
RA	1-like	243	[Aethina tumida]	ptera	66289	E-34	975	872	81	70

Table 3.20. Continued.

Sequenc	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	nce			ACC	Value	larit	Scor	ment	itiv
		lengt					У	е	lengt	es
		h							h	
Lep_000	myrosin		gi 506965870 gb AGM32287.1 beta-		AGM		58.8			
65284-	ase 1-		glucosidase, partial [Coptotermes	Blattod	3228	3.18E-	235	202.		
RA	like	735	formosanus]	ea	7	62	3	216	204	120
Lep_000	myrosin		gi 998493172 ref XP_015513162.1		XP_0		63.3			
68929-	ase 1-		PREDICTED: myrosinase 1-like	Hymeno	1551	1.6E-	802	69.7		
RA	like	726	[Neodiprion lecontei]	ptera	3162	11	8	07	71	45
Lep_001			gi 1008445131 ref XP_975665.2 PR		XP_9		83.6			
20227-	myrosin		EDICTED: myrosinase 1 isoform X2	Coleopt	7566	3.3E-	734	51.2		
RA	ase 1	204	[Tribolium castaneum]	era	5	07	7	174	49	41
			gi 1339068038 ref XP_023721112.1		XP_0					
			myrosinase 1-like [Cryptotermes		2372					
			secundus]gi 1339068040 ref XP_02		1112,					
			3721113.1 myrosinase 1-like		XP_0					
			[Cryptotermes		2372					
			secundus]gi 1330885855 gb PNF19		1113,					
			653.1 Myrosinase 1 [Cryptotermes		PNF1					
Lep_000			secundus]gi 1330885856 gb PNF19		9653,		64.0			
28153-	Myrosin		654.1 Myrosinase 1 [Cryptotermes	Blattod	PNF1	2.9E-	540	368.		
RA	ase 1	1590	secundus]	ea	9654	119	5	237	370	237
			gi 1069792344 ref XP_018321256.1							
Lep_001	myrosin		PREDICTED: uncharacterized		XP_0		56.1			
33665-	ase 1-		protein LOC108734274 [Agrilus	Coleopt	1832	3.91E-	904	81.2		
RA	like	321	planipennis]	era	1256	17	8	629	105	59

Table 3.20. Continued

Sequenc	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	nce			ACC	Value	larit	Scor	ment	itiv
		lengt					У	е	lengt	es
		h							h	
Lep_000			gi 1233170929 ref XP_022187264.1		XP_0					
22255-	Myrosin		myrosinase 1-like isoform X2	Hemipt	2218	7.19E-	73.7	271.		
RA	ase 1	735	[Nilaparvata lugens]	era	7264	88	069	552	232	171
Lep_000	myrosin		gi 1042851809 gb ANQ90707.1 beta		ANQ		81.1			
22256-	ase 1-		-glucosidase 2, partial [Manduca	Lepidop	9070	7.93E-	594	107.		
RA	like	396	quinquemaculata]	tera	7	29	2	457	69	56
Lep_000			gi 506967929 gb AGM32308.1 beta-		AGM		82.4			
62260-	Myrosin		glucosidase, partial [Coptotermes	Blattod	3230	1.8E-	175	132.		
RA	ase 1	324	formosanus]	ea	8	37	8	88	91	75
Lep_000			gi 1325352588 ref XP_023311344.1		XP_0					
62259-	myrosin		lactase-phlorizin hydrolase-like	Coleopt	2331	6.3E-		96.2		
RA	ase 1	225	[Anoplophora glabripennis]	era	1344	23	80	857	65	52
Lep_000			gi 1042851809 gb ANQ90707.1 beta		ANQ		83.0			
04129-	myrosin		-glucosidase 2, partial [Manduca	Lepidop	9070	4.49E-	357	156.		
RA	ase 1	435	quinquemaculata]	tera	7	48	1	762	112	93
Lep_000	myrosin		gi 1061476505 gb 0DM94351.1 Lac		ODM		80.8			
04128-	ase 1-		tase-phlorizin hydrolase [Orchesella	Collemb	9435	1.22E-	219	113.		
RA	like	750	cincta]	ola	1	26	2	62	73	59
Lep_000	myrosin		gi 506967929 gb AGM32308.1 beta-		AGM		82.9			
04130-	ase 1-		glucosidase, partial [Coptotermes	Blattod	3230	1.3E-	787	80.1		
RA	like	456	formosanus]	ea	8	16	2	073	47	39

Table 3.20	. Continued.
------------	--------------

Sequenc	Sequenc	Seque	Hit desc.	Order	Hit	E-	Simi	Bit-	Align	Pos
e name	e desc.	nce			ALL	value		SCOL	ment	Itiv
		h					У	е	h	es
Lep_001			gi 506967929 gb AGM32308.1 beta-		AGM		83.5			
06154-	Myrosin		glucosidase, partial [Coptotermes	Blattod	3230	1.69E-	164	132.		
RA	ase 1	324	formosanus]	ea	8	37	8	88	91	76
Lep_000	myrosin		gi 1279707180 ref XP_022915768.1		XP_0		72.0			
22729-	ase 1-		myrosinase 1-like [Onthophagus	Coleopt	2291	9.87E-	164	291.		
RA	like	1096	taurus]	era	5768	93	6	967	243	175
Lep_000	myrosin		gi 1080058576 ref XP_018571131.1		XP_0		73.0			
22730-	ase 1-		myrosinase 1 [Anoplophora	Coleopt	1857	4.89E-	158	94.7		
RA	like	486	glabripennis]	era	1131	21	7	449	63	46
Lep_000	myrosin		gi 636793009 dbj BA085045.1 beta				73.2			
47597-	ase 1-		-glucosidase, partial [Zootermopsis	Blattod	BA08	4.43E-	876	171.		
RA	like	558	nevadensis]	ea	5045	51	7	785	146	107
	myrosin									
	ase 1-									
Lep_000	like		gi 1069782941 ref XP_018335102.1		XP_0		77.0			
47598-	isoform		PREDICTED: myrosinase 1-like	Coleopt	1833	4.66E-	642	121.		
RA	X1	624	[Agrilus planipennis]	era	5102	30	2	709	109	84

Sequen	Seque	Sequ	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
се	nce	ence		r		Value	larit	Scor	ment	itiv
name	desc.	lengt					У	е	lengt	es
		h							h	
			gi 1339068038 ref XP_023721112.1 my							
			rosinase 1-like [Cryptotermes		XP_023					
			secundus]gi 1339068040 ref XP_02372		721112,					
			1113.1 myrosinase 1-like [Cryptotermes		XP_023					
			secundus]gi 1330885855 gb PNF19653.		721113,					
			1 Myrosinase 1 [Cryptotermes		PNF196					
Lep_00			secundus]gi 1330885856 gb PNF19654.		53,		65.0			
015101	Myrosi		1 Myrosinase 1 [Cryptotermes	Blatt	PNF196	1.2E-	526	441.		
-RA	nase 1	1457	secundus]	odea	54	148	3	81	475	309
Lep_00							77.1			
106870	myrosi		gi 1101347123 ref XP_018899386.1 PR	Hemi	XP_018	1.73E-	929	76.2		
-RA	nase 1	303	EDICTED: myrosinase 1 [Bemisia tabaci]	ptera	899386	15	8	554	57	44
Lep_00	myrosi		gi 748995286 gb AJE75665.1 putative				72.0			
111910	nase 1-		glycosyl hydrolase [Chrysomela	Coleo	AJE756	6.21E-	279	180.		
-RA	like	501	lapponica]	ptera	65	53	7	644	143	103
Lep_00	myrosi		gi 218749835 ref NP_001136332.1 glyc	Hym			76.5			
105473	nase 1-		oside hydrolase-like protein [Nasonia	enopt	NP_001	4.12E-	957	72.7		
-RA	like	327	vitripennis]	era	136332	14	4	886	47	36
			gi 1339057501 ref XP_023717250.1 my							
			rosinase 1-like isoform X2							
			[Cryptotermes		XP_023					
Lep_00			secundus]gi 1330893380 gb PNF23656.		717250,		86.6			
096192	Myrosi		1 Myrosinase 1 [Cryptotermes	Blatt	PNF236	1.37E-	666	77.0		
-RA	nase 1	165	secundus]	odea	56	16	7	258	45	39

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
се	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	ce					У	е	lengt	es
		leng							h	
		th								
	myrosi									
	nase 1-									
Lep_00	like		gi 1233170929 ref XP_022187264.1 myr				79.0			
020465	isofor		osinase 1-like isoform X2 [Nilaparvata	Hemi	XP_022	1.09E-	209	196.		
-RA	m X2	501	lugens]	ptera	187264	59	8	052	143	113
Lep_00	myrosi		gi 506967929 gb AGM32308.1 beta-				76.8			
020464	nase 1-		glucosidase, partial [Coptotermes	Blatt	AGM32	1.56E-	115	171.		
-RA	like	885	formosanus]	odea	308	49	9	014	138	106
			gi 1228381282 ref XP_021952145.1 myr							
			osinase 1-like [Folsomia		XP_021					
Lep_00	myrosi		candida]gi 1215269331 gb 0XA55643.1	Colle	952145,		84.8			
020463	nase 1-		Lactase-phlorizin hydrolase [Folsomia	mbol	OXA556	1.64E-	484	62.7		
-RA	like	330	candida]	а	43	10	8	734	33	28
			gi 1339068038 ref XP_023721112.1 myr		XP_023					
			osinase 1-like [Cryptotermes		721112,					
			secundus]gi 1339068040 ref XP_023721		XP_023					
			113.1 myrosinase 1-like [Cryptotermes		721113,					
			secundus]gi 1330885855 gb PNF19653.		PNF196					
Lep_00			1 Myrosinase 1 [Cryptotermes		53,		66.0			
041138	Myrosi		secundus]gi 1330885856 gb PNF19654.	Blatt	PNF196	2.61E-	919	182.		
-RA	nase 1	882	1 Myrosinase 1 [Cryptotermes secundus]	odea	54	51	5	956	174	115

Table 3.20. Continued	
-----------------------	--

Sequen	Seque	Seque	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
се	nce	nce		r		Value	larit	Scor	ment	itiv
name	desc.	lengt					У	е	lengt	es
		h							h	
Lep_00	myrosi		gi 1080062338 ref XP_018573207.1 m				72.7			
041137	nase 1-		yrosinase 1-like isoform X2	Coleo	XP_018	1.27E-	272	92.4		
-RA	like	408	[Anoplophora glabripennis]	ptera	573207	20	7	337	66	48
Lep_00	myrosi		gi 998489397 gb AMK48629.1 glycosi				71.4			
134507	nase 1-		de hydrolase family 1, partial	Coleo	AMK48		285	118.		
-RA	like	294	[Rhynchophorus ferrugineus]	ptera	629	1E-33	7	627	91	65
Lep_00	myrosi		gi 1080062829 ref XP_018573475.1 m				78.9			
027729	nase 1-		yrosinase 1-like [Anoplophora	Coleo	XP_018	1.97E-	115	195.		
-RA	like	744	glabripennis]	ptera	573475	57	6	667	147	116
			gi 1228381282 ref XP_021952145.1 m							
			yrosinase 1-like [Folsomia		XP_021					
Lep_00	myrosi		candida]gi 1215269331 gb 0XA55643.	Colle	952145,					
027731	nase 1-		1 Lactase-phlorizin hydrolase	mbol	OXA556	3.37E-		60.8		
-RA	like	276	[Folsomia candida]	а	43	10	87.5	474	32	28
Lep_00			gi 506967929 gb AGM32308.1 beta-				84.4			
027730	Myrosi		glucosidase, partial [Coptotermes	Blatt	AGM32	6.27E-	444	73.5		
-RA	nase 1	303	formosanus]	odea	308	15	4	59	45	38
Lep_00			gi 1080062823 ref XP_018573472.1 m				73.4			
060394	myrosi		yrosinase 1, partial [Anoplophora	Coleo	XP_018	2.42E-	042	116.		
-RA	nase 1	463	glabripennis]	ptera	573472	29	6	701	94	69

Table 3.20. Continued	
-----------------------	--

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
се	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	се					У	е	lengt	es
		leng							h	
		th								
Lep_00			gi 506967929 gb AGM32308.1 beta-				86.9			
060393	Myrosi		glucosidase, partial [Coptotermes	Blatt	AGM32	1.51E-	565	182.		
-RA	nase 1	423	formosanus]	odea	308	56	2	956	115	100
Lep_00	myrosi		gi 1069789897 ref XP_018320188.1 PRE							
136832	nase 1-		DICTED: myrosinase 1-like [Agrilus	Coleo	XP_018	1.01E-		78.9		
-RA	like	243	planipennis]	ptera	320188	16	84	518	50	42
Lep_00	myrosi		gi 506967929 gb AGM32308.1 beta-				78.6			
101061	nase 1-		glucosidase, partial [Coptotermes	Blatt	AGM32	2.29E-	666	112.		
-RA	like	393	formosanus]	odea	308	29	7	849	75	59
Lep_00			gi 1008445131 ref XP_975665.2 PREDIC				81.9			
074338	myrosi		TED: myrosinase 1 isoform X2 [Tribolium	Coleo	XP_975	1.15E-	148	113.		
-RA	nase 1	330	castaneum]	ptera	665	28	9	62	94	77
Lep_00	myrosi		gi 1080062829 ref XP_018573475.1 myr							
074339	nase 1-		osinase 1-like [Anoplophora	Coleo	XP_018	8.33E-	76.1	93.9		
-RA	like	291	glabripennis]	ptera	573475	22	194	745	67	51
Lep_00							79.2			
073232	Myrosi		gi 1090651209 gb A0Y34571.1 beta-	Blatt	A0Y345	1.77E-	929	266.		
-RA	nase 1	621	glucosidase [Coptotermes formosanus]	odea	71	85	3	929	198	157
Lep_00							91.3			
143227	Myrosi		gi 1233173996 ref XP_022188302.1 myr	Hemi	XP_022	1.33E-	043	68.9		
-RA	nase 1	168	osinase 1-like [Nilaparvata lugens]	ptera	188302	13	5	366	46	42

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
ce	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	ce					У	е	lengt	es
		leng							h	
		th								
Lep_00	myrosi		gi 1080062831 ref XP_018573476.1 myr				78.1			
078589	nase 1-		osinase 1-like [Anoplophora	Coleo	XP_018	2.55E-	954	187.		
-RA	like	690	glabripennis]	ptera	573476	54	9	193	133	104
Lep_00							72.6			
042919	myrosi	119	gi 1325357344 ref XP_023312869.1 myr	Coleo	XP_023	1.1E-	072	354.		
-RA	nase 1	4	osinase 1 [Anoplophora glabripennis]	ptera	312869	116	6	369	303	220
Lep_00	myrosi			Colle			41.1			
045854	nase 1-		gi 1228404666 ref XP_021961442.1 myr	mbol	XP_021	1.2E-	392	77.0		
-RA	like	480	osinase 1-like [Folsomia candida]	а	961442	15	4	258	158	65
Lep_00	myrosi									
060792	nase 1-		gi 668445331 gb KFB35692.1 AGAP006	Dipte	KFB356	1.86E-		77.7		
-RA	like	426	424-PA-like protein [Anopheles sinensis]	ra	92	15	84	962	50	42
Lep_00			gi 1247040849 gb PCG75727.1 hypothet	Lepid						
020108	myrosi		ical protein B5V51_11090 [Heliothis	opter	PCG757	3.61E-	49.0	69.7		
-RA	nase 1	526	virescens]	а	27	12	566	07	159	78
Lep_00	myrosi		gi 939630683 ref XP_014278711.1 PRE				65.0			
020109	nase 1-	110	DICTED: myrosinase 1-like	Hemi	XP_014	5.59E-	793	76.6		
-RA	like	4	[Halyomorpha halys]	ptera	278711	13	7	406	63	41
			gi 1339073035 ref XP_023723774.1 myr		XP_023					
Lep_00	myrosi		osinase 1-like [Cryptotermes		723774,		80.6			
113513	nase 1-		secundus]gi 1330878957 gb PNF16841.	Blatt	PNF168	2.89E-	451	144.		
-RA	like	477	1 Myrosinase 1 [Cryptotermes secundus]	odea	41	39	6	436	93	75

Table 3.20. Continu	.ued
---------------------	------

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
се	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	ce					У	e	lengt	es
		leng							h	
		th								
Lep_00	myrosi		gi 1008445131 ref XP_975665.2 PREDIC				81.9			
058059	nase 1-		TED: myrosinase 1 isoform X2 [Tribolium	Coleo	XP_975	8.9E-	148	113.		
-RA	like	318	castaneum]	ptera	665	29	9	62	94	77
Lep_00	myrosi		gi 506967929 gb AGM32308.1 beta-				73.7			
058058	nase 1-		glucosidase, partial [Coptotermes	Blatt	AGM32	1.35E-	864	125.		
-RA	like	513	formosanus]	odea	308	33	1	561	103	76
Lep_00	myrosi		gi 506967929 gb AGM32308.1 beta-				74.3			
042495	nase 1-		glucosidase, partial [Coptotermes	Blatt	AGM32	5.89E-	016	228.		
-RA	like	708	formosanus]	odea	308	73	8	794	179	133
Lep_00	myrosi		gi 1080062831 ref XP_018573476.1 myr				82.5			
042496	nase 1-		osinase 1-like [Anoplophora	Coleo	XP_018	4.63E-	396	97.4		
-RA	like	480	glabripennis]	ptera	573476	22	8	413	63	52
Lep_00	myrosi		gi 1233170929 ref XP_022187264.1 myr				76.8			
054176	nase 1-	108	osinase 1-like isoform X2 [Nilaparvata	Hemi	XP_022	2.54E-	707	181.		
-RA	like	0	lugens]	ptera	187264	51	5	415	147	113
Lep_00	myrosi						74.3			
093809	nase 1-		gi 636793015 dbj BA085048.1 beta-	Blatt	BA0850	1.26E-	243	93.9		
-RA	like	465	glucosidase, partial [Salganea esakii]	odea	48	21	2	745	74	55
Lep_00										
092612	myrosi		gi 668445331 gb KFB35692.1 AGAP006	Dipte	KFB356	4.36E-		77.0		
-RA	nase 1	717	424-PA-like protein [Anopheles sinensis]	ra	92	14	84	258	50	42

Table 3.20. Continued	
-----------------------	--

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
ce	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	ce					У	е	lengt	es
		leng							h	
		th								
Lep_00	myrosi		gi 1227954946 ref XP_021937620.1 myr				85.4			
042140	nase 1-		osinase 1-like isoform X1 [Zootermopsis	Blatt	XP_021	1.65E-	166	86.6		
-RA	like	234	nevadensis]	odea	937620	19	7	557	48	41
Lep_00			gi 1080062831 ref XP_018573476.1 myr				81.9			
042138	myrosi		osinase 1-like [Anoplophora	Coleo	XP_018	1.68E-	672	93.9		
-RA	nase 1	351	glabripennis]	ptera	573476	21	1	745	61	50
Lep_00	myrosi		gi 506967929 gb AGM32308.1 beta-							
113948	nase 1-		glucosidase, partial [Coptotermes	Blatt	AGM32	9.31E-		116.		
-RA	like	417	formosanus]	odea	308	31	70.4	701	125	88
Lep_00			gi 506965870 gb AGM32287.1 beta-							
082283	Myrosi		glucosidase, partial [Coptotermes	Blatt	AGM32	4.89E-	87.2	117.		
-RA	nase 1	612	formosanus]	odea	287	30	093	857	86	75
Lep_00	myrosi		gi 1114635916 gb APM84101.1 GH1							
150133	nase 1-		beta-glucosidase [Microcerotermes	Blatt	APM84	8.15E-	76.2	115.		
-RA	like	294	annandalei]	odea	101	30	5	931	80	61
Lep_00			gi 1042851809 gb ANQ90707.1 beta-	Lepid			75.8			
106211	myrosi		glucosidase 2, partial [Manduca	opter	ANQ90	4.14E-	064	151.		
-RA	nase 1	438	quinquemaculata]	a	707	46	5	754	124	94
Lep_00			gi 506965870 gb AGM32287.1 beta-				74.2			
089193	Myrosi		glucosidase, partial [Coptotermes	Blatt	AGM32	1.83E-	857	88.5		
-RA	nase 1	780	formosanus]	odea	287	18	1	817	70	52

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	Е-	Simi	Bit-	Align	Pos
се	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	ce					У	е	lengt	es
		leng							h	
		th								
			gi 1339068038 ref XP_023721112.1 myr		XP_023					
			osinase 1-like [Cryptotermes		721112,					
			secundus]gi 1339068040 ref XP_023721		XP_023					
			113.1 myrosinase 1-like [Cryptotermes		721113,					
			secundus]gi 1330885855 gb PNF19653.		PNF196					
Lep_00			1 Myrosinase 1 [Cryptotermes		53,		67.4			
033948	Myrosi	125	secundus]gi 1330885856 gb PNF19654.	Blatt	PNF196	8.2E-	528	406.		
-RA	nase 1	9	1 Myrosinase 1 [Cryptotermes secundus]	odea	54	136	3	371	424	286
Lep_00			gi 1233170926 ref XP_022187263.1 myr				70.4			
024904	Myrosi	141	osinase 1-like isoform X1 [Nilaparvata	Hemi	XP_022	4.3E-	081	434.		
-RA	nase 1	3	lugens]	ptera	187263	147	6	106	392	276
Lep_00	myrosi			Colle			77.5			
024905	nase 1-		gi 1215275228 gb 0XA61537.1 Lactase-	mbol	OXA615	3.85E-	510	65.8		
-RA	like	228	phlorizin hydrolase [Folsomia candida]	а	37	12	2	55	49	38
Lep_00	myrosi						72.4			
013204	nase 1-	131	gi 1279707180 ref XP_022915768.1 myr	Coleo	XP_022	1.55E-	279	306.		
-RA	like	2	osinase 1-like [Onthophagus taurus]	ptera	915768	97	8	99	243	176
Lep_00	myrosi		gi 1080062338 ref XP_018573207.1 myr							
013205	nase 1-		osinase 1-like isoform X2 [Anoplophora	Coleo	XP_018	1.84E-		91.2		
-RA	like	372	glabripennis]	ptera	573207	20	75	781	64	48
Lep_00	myrosi						76.4			
097528	nase 1-		gi 1090651209 gb A0Y34571.1 beta-	Blatt	A0Y345	3.36E-	705	169.		
-RA	like	761	glucosidase [Coptotermes formosanus]	odea	71	47	9	474	119	91

Table 3.20. (Continued.
---------------	------------

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
се	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	ce					У	е	lengt	es
		leng							h	
		th								
Lep_00	myrosi		gi 1325349814 ref XP_023310604.1 myr				74.1			
009457	nase 1-		osinase 1-like [Anoplophora	Coleo	XP_023	6.7E-	935	114.		
-RA	like	831	glabripennis]	ptera	310604	27	5	775	93	69
Lep_00							58.2			
009460	Myrosi		gi 1330893407 gb PNF23683.1 Myrosin	Blatt	PNF236	4.35E-	089	122.		
-RA	nase 1	912	ase 1 [Cryptotermes secundus]	odea	83	29	6	479	134	78
Lep_00	myrosi		gi 506967929 gb AGM32308.1 beta-							
072701	nase 1-		glucosidase, partial [Coptotermes	Blatt	AGM32	7.56E-		125.		
-RA	like	921	formosanus]	odea	308	32	82	561	100	82
	myrosi									
	nase 1-									
Lep_00	like		gi 506967929 gb AGM32308.1 beta-				70.5			
011410	isofor		glucosidase, partial [Coptotermes	Blatt	AGM32	1.49E-	357	133.		
-RA	m X2	339	formosanus]	odea	308	37	1	265	112	79
Lep_00	myrosi						59.1			
011407	nase 1-		gi 1090651209 gb A0Y34571.1 beta-	Blatt	A0Y345	8.8E-	836	98.5		
-RA	like	618	glucosidase [Coptotermes formosanus]	odea	71	22	7	969	98	58
Lep_00	myrosi			Colle			93.3			
011408	nase 1-		gi 1061483168 gb 0DN00405.1 Lactase-	mbol	ODN00	2.78E-	333	58.9		
-RA	like	318	phlorizin hydrolase [Orchesella cincta]	а	405	09	3	214	30	28

Table 3.20. Continued	
-----------------------	--

Sequen	Seque	Seq	Hit desc.	Orde	Hit ACC	E-	Simi	Bit-	Align	Pos
се	nce	uen		r		Value	larit	Scor	ment	itiv
name	desc.	се					У	е	lengt	es
		leng							h	
		th								
Lep_00	myrosi		gi 746857492 ref XP_011059478.1 PRE	Hym			81.5			
011409	nase 1-		DICTED: myrosinase 1 isoform X2	enopt	XP_011	1.08E-	217	150.		
-RA	like	423	[Acromyrmex echinatior]	era	059478	41	4	214	92	75
			gi 1339073035 ref XP_023723774.1 myr		XP_023					
Lep_00	myrosi		osinase 1-like [Cryptotermes		723774,		68.1			
011406	nase 1-		secundus]gi 1330878957 gb PNF16841.	Blatt	PNF168	1.33E-	818	58.9		
-RA	like	474	1 Myrosinase 1 [Cryptotermes secundus]	odea	41	08	2	214	44	30
Lep_00	myrosi		gi 1227984163 ref XP_021923731.1 lact				81.6			
113088	nase 1-		ase-phlorizin hydrolase-like isoform X2	Blatt	XP_021	8.45E-	326	215.		
-RA	like	447	[Zootermopsis nevadensis]	odea	923731	67	5	698	147	120
Lep_00							92.3			
102915	myrosi		gi 1339057505 ref XP_023717252.1 myr	Blatt	XP_023	2.18E-	076	50.8		
-RA	nase 1	318	osinase 1-like [Cryptotermes secundus]	odea	717252	06	9	322	26	24
Lep_00			gi 1080052826 ref XP_018567975.1 lact				62.7			
039189	myrosi		ase-like protein [Anoplophora	Coleo	XP_018	1.85E-	329	135.		
-RA	nase 1	579	glabripennis]	ptera	567975	35	2	576	161	101

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
nce name	desc.	nce lengt			ALL	value	arity	e Scor	lengt	es
	Lutic	11	gil1240050415lomblSUW61272						11	
Lop 00	Lyttt		11 utic polycoccharido							
Lep_00	polysaccilariue		monooyygonago, partial		CIM/61	1 505	75 72	2652		
09/91 E DA	monooxygenas	0/1	Thormobia domostical		272	1.30E- 07	016	205.5	206	156
J-NA	e Lutio	041	cil12400E0417lomblSUV(61272		372	07	010	00	200	150
Lop 00	Lytit		11 utic polycoccharido							
107451	polysaccilariue		monooyygonago, partial		CIM/61	2 OF	02.05	201 7		
0/451 1 D A	monooxygenas	000	Thormobia domostical		272	3.0E- 122	92.05	301.7	214	107
1-KA	e Lutio	000	cil12400E041ElomblSUV61272		373	155	007	19	214	197
Lon 00	Lytic		gi 1349059415 jeiii Dj51w01572							
Lep_00	polysaccharide		.1 Lytic polysaccharide		CILAZC 1	2.275		(7.20		
12/36	monooxygenas	400	monooxygenase, partial		51W61	Z.Z/E-	66.66	67.39		20
3-RA	e	408	[Thermobia domestica]		372	12	667	58	57	38
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
11924	monooxygenas		monooxygenase, partial		SIW61	1.08E-	90.83	221.4		
7-RA	е	383	[Thermobia domestica]		372	72	333	76	120	109
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
13011	monooxygenas		monooxygenase, partial		SIW61	2.9E-	66.23	80.49		
0-RA	е	237	[Thermobia domestica]		373	18	377	25	77	51
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
03953	monooxygenas		monooxygenase, partial		SIW61	1.06E-	70.81	260.7		
7-RA	е	842	[Thermobia domestica]		372	85	34	66	209	148

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

Table 3.21. Continued

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
nce	desc.	nce			ACC	Value	arity	Scor	ment	itiv
name		lengt						е	lengt	es
		h							h	
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
03953	monooxygenas		monooxygenase, partial		SIW61	2.9E-	69.61	213.7		
8-RA	е	834	[Thermobia domestica]		372	67	326	72	181	126
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
06835	monooxygenas		monooxygenase, partial		SIW61	6.59E-	68.87	176.7		
4-RA	е	1281	[Thermobia domestica]		373	51	417	92	151	104
			gi 939644305 ref XP_0142723							
	Lytic		00.1 PREDICTED:							
Lep_00	polysaccharide		uncharacterized protein		XP_01					
15107	monooxygenas		LOC106678350 [Halyomorpha	Hemip	42723	2.78E-	74.50	73.17		
8-RA	e	168	halys]	tera	00	15	98	38	51	38
	Lytic		gi 1009534315 ref XP_015904							
Lep_00	polysaccharide		140.1 uncharacterized protein		XP_01					
09782	monooxygenas		LOC107436837 [Parasteatoda	Arach	59041	2.39E-	66.96	253.8		
6-RA	e	716	tepidariorum]	nida	40	83	429	32	224	150
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
07768	monooxygenas		monooxygenase, partial		SIW61	1.14E-	75.81	289.2		
0-RA	e	876	[Thermobia domestica]		372	96	395	71	215	163

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
nce	desc.	nce			ACC	Value	arity	Scor	ment	itiv
name		h						e	h	es
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
08180	monooxygenas		monooxygenase, partial		SIW61	3.46E-	67.33	201.4		
4-RA	е	889	[Thermobia domestica]		373	62	668	45	199	134
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
09778	monooxygenas		monooxygenase, partial		SIW61	2.08E-	78.35	127.8		
8-RA	е	466	[Thermobia domestica]		373	35	052	72	97	76
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
09909	monooxygenas		monooxygenase, partial		SIW61	4.34E-		236.8		
3-RA	е	684	[Thermobia domestica]		373	77	71	84	200	142
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
11857	monooxygenas		monooxygenase, partial		SIW61	1.64E-	77.84	249.5		
7-RA	е	597	[Thermobia domestica]		373	82	431	95	167	130
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
02939	monooxygenas		monooxygenase, partial		SIW61	1.51E-	80.74	248.4		
6-RA	е	703	[Thermobia domestica]		373	81	534	4	161	130
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
02939	monooxygenas		monooxygenase, partial		SIW61	3.77E-	91.17	61.61		
5-RA	е	228	[Thermobia domestica]		372	11	647	78	34	31

Seque	Sequence	Seque	Hit desc.	Order	Hit	Е-	Simil	Bit-	Align	Pos
nce	desc.	nce			ACC	Value	arity	Scor	ment	itiv
name		lengt h						e	lengt h	es
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
02939	monooxygenas		monooxygenase, partial		SIW61	7.28E-	77.39	177.1		
4-RA	e	648	[Thermobia domestica]		373	54	13	78	115	89
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
06166	monooxygenas		monooxygenase, partial		SIW61	2.06E-	80.74	248.0		
2-RA	е	703	[Thermobia domestica]		373	81	534	54	161	130
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
06166	monooxygenas		monooxygenase, partial		SIW61	5.54E-	90.47	75.48		
3-RA	е	300	[Thermobia domestica]		372	16	619	5	42	38
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
04337	monooxygenas		monooxygenase, partial		SIW61	4.55E-	73.09	246.8		
1-RA	е	867	[Thermobia domestica]		373	80	645	99	197	144
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
04337	monooxygenas		monooxygenase, partial		SIW61	1.07E-	73.05	209.1		
2-RA	е	573	[Thermobia domestica]		373	66	389	49	167	122
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
12201	monooxygenas		monooxygenase, partial		SIW61	1.2E-	75.58	231.4		
7-RA	e	558	[Thermobia domestica]		373	75	14	91	172	130

Table 3.21. Continued

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
nce	desc.	nce			ACC	Value	arity	Scor	ment	itiv
name		lengt						е	lengt	es
		h							h	
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
09384	monooxygenas		monooxygenase, partial		SIW61	2.71E-	79.24	292.7		
5-RA	e	848	[Thermobia domestica]		372	98	528	38	212	168
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
06722	monooxygenas		monooxygenase, partial		SIW61	2.38E-	67.33	202.6		
6-RA	e	938	[Thermobia domestica]		373	62	668	01	199	134
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
09761	monooxygenas		monooxygenase, partial		SIW61	1.25E-	78.35	127.8		
4-RA	e	425	[Thermobia domestica]		373	35	052	72	97	76
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
09840	monooxygenas		monooxygenase, partial		SIW61	6.3E-	93.67	297.3		
8-RA	e	684	[Thermobia domestica]		373	101	089	6	158	148
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
09931	monooxygenas		monooxygenase, partial		SIW61	8.61E-	76.27	286.5		
8-RA	e	660	[Thermobia domestica]		372	97	907	74	215	164

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
nce	desc.	nce			ACC	Value	arity	Scor	ment	itiv
name		lengt						е	lengt	es
		h							h	
			gi 1316145279 ref XP_023217							
			685.1 uncharacterized protein		XP_02					
			LOC111620072 [Centruroides	Arthr	32176					
	Lytic		sculpturatus]gi 1316145289 re	opoda	85,					
Lep_00	polysaccharide		f XP_023217690.1 uncharacteri	-	XP_02					
03894	monooxygenas		zed protein LOC111620076	Scorpi	32176	2.1E-	70.95	242.6		
7-RA	е	876	[Centruroides sculpturatus]	ones	90	78	238	62	210	149
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
03894	monooxygenas		monooxygenase, partial		SIW61	2.36E-	62.24	231.1		
8-RA	е	768	[Thermobia domestica]		373	74	066	06	241	150
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
10667	monooxygenas		monooxygenase, partial		SIW61	7.73E-	77.39	175.2		
8-RA	е	507	[Thermobia domestica]		373	54	13	52	115	89
			gi 1067063705 ref XP_018015							
	Lytic		775.1 PREDICTED:							
Lep_00	polysaccharide		uncharacterized protein		XP_01					
06665	monooxygenas		LOC108672588 [Hyalella	Crusta	80157	9.93E-	72.46	190.6		
0-RA	е	534	azteca]	cean	75	55	377	6	138	100
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
01814	monooxygenas		monooxygenase, partial		SIW61	7.5E-	70.75	273.4		
0-RA	е	824	[Thermobia domestica]		372	91	472	78	212	150

Table	3.21.	Continued.

Seque	Sequence	Seque	Hit desc.	Order	Hit	E-	Simil	Bit-	Align	Pos
nce	desc.	nce			ACC	Value	arity	Scor	ment	itiv
name		lengt						e	lengt	es
-		h							h	
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
13640	monooxygenas		monooxygenase, partial		SIW61	3.45E-	77.39	174.4		
2-RA	е	384	[Thermobia domestica]		373	54	13	81	115	89
	Lytic		gi 1009534315 ref XP_015904							
Lep_00	polysaccharide		140.1 uncharacterized protein		XP_01					
10993	monooxygenas		LOC107436837 [Parasteatoda	Arach	59041	1.08E-	70.55	219.5		
4-RA	е	942	tepidariorum]	nida	40	68	556	5	180	127
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
12188	monooxygenas		monooxygenase, partial		SIW61	1.61E-	74.41	170.2		
5-RA	е	549	[Thermobia domestica]		373	51	86	44	129	96
	Lytic		gi 1349059415 emb SIW61372							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
07232	monooxygenas		monooxygenase, partial		SIW61	1.7E-	78.30	300.0		
2-RA	е	961	[Thermobia domestica]		372	100	189	56	212	166
	Lytic		gi 1349059417 emb SIW61373							
Lep_00	polysaccharide		.1 Lytic polysaccharide							
11965	monooxygenas		monooxygenase, partial		SIW61	6.69E-	81.08	278.4		
6-RA	e	617	[Thermobia domestica]		373	94	108	85	185	150
	Lytic		gi 1009534315 ref XP_015904							
Lep_00	polysaccharide		140.1 uncharacterized protein		XP_01					
07183	monooxygenas		LOC107436837 [Parasteatoda	Arach	59041	1.14E-	73.56	209.9		
8-RA	е	796	tepidariorum]	nida	40	65	322	2	174	128

Table 3.21. Continue	d.
----------------------	----

Seque	Sequence desc.	Sequ	Hit desc.	Ord	Hit	E-	Simil	Bit-	Alig	Pos
nce		ence		er	ACC	Value	arity	Score	nme	itiv
name		lengt							nt	es
		h							leng	
									th	
			gi 1349059417 emb SIW61373.1							
Lep_00	Lytic		Lytic polysaccharide							
08301	polysaccharide		monooxygenase, partial		SIW61	5.03E-	60.39	197.5		
6-RA	monooxygenase	619	[Thermobia domestica]		373	62	604	93	202	122
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
03755	polysaccharide		monooxygenase, partial		SIW61	3.11E-	68.66	169.8		
6-RA	monooxygenase	1029	[Thermobia domestica]		372	49	667	59	150	103
			gi 1349059417 emb SIW61373.1							
Lep_00	Lytic		Lytic polysaccharide							
03755	polysaccharide		monooxygenase, partial		SIW61	1.98E-	78.78	52.37		
7-RA	monooxygenase	537	[Thermobia domestica]		373	06	788	3	33	26
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
08328	polysaccharide		monooxygenase, partial		SIW61	8.76E-	74.41	220.3		
3-RA	monooxygenase	853	[Thermobia domestica]		372	70	86	2	172	128
			gi 1351638095 pdb 5MSZ AChai							
			n A, Lytic Polysaccharide							
Lep_00	Lytic		Monooxygenase AA15 from							
03092	polysaccharide		Thermobia domestica in the Cu(I)		5MSZ_	5.65E-	80.48	60.46		
2-RA	monooxygenase	390	State		А	10	78	22	41	33

Seque	Sequence	Seque	Hit desc.	Ord	Hit	E-	Simil	Bit-	Alig	Pos
nce	desc.	nce		er	ACC	Value	arity	Score	nme	itiv
name		lengt							nt	es
		h							leng	
									th	
	Lytic		gi 1349059417 emb SIW61373.1							
Lep_00	polysaccharide		Lytic polysaccharide							
03092	monooxygenas		monooxygenase, partial		SIW61	8.91E-	79.41	134.0		
4-RA	е	306	[Thermobia domestica]		373	39	176	35	102	81
	Lytic		gi 1349059415 emb SIW61372.1							
Lep_00	polysaccharide		Lytic polysaccharide							
02825	monooxygenas		monooxygenase, partial		SIW61	2.2E-	78.77	300.0		
3-RA	e	987	[Thermobia domestica]		372	100	358	56	212	167
	Lytic		gi 1349059415 emb SIW61372.1							
Lep_00	polysaccharide		Lytic polysaccharide							
02825	monooxygenas		monooxygenase, partial		SIW61	3.75E-	73.51	231.1		
2-RA	e	615	[Thermobia domestica]		372	75	351	06	185	136
	Lytic		gi 1349059417 emb SIW61373.1							
Lep_00	polysaccharide		Lytic polysaccharide							
02825	monooxygenas		monooxygenase, partial		SIW61	1.32E-	93.54	62.77		
1-RA	e	219	[Thermobia domestica]		373	11	839	34	31	29
	Lytic		gi 1349059417 emb SIW61373.1							
Lep_00	polysaccharide		Lytic polysaccharide							
10323	monooxygenas		monooxygenase, partial		SIW61	2.33E-	66.83	218.0		
4-RA	e	753	[Thermobia domestica]		373	69	673	09	196	131

Table 3.21. Continued

Seque	Sequence desc.	Seq	Hit desc.	Ord	Hit	E-	Simil	Bit-	Alig	Pos
nce		uen		er	ACC	Value	arity	Score	nme	itiv
name		ce							nt	es
		leng							leng	
		th							th	
			gi 1349059417 emb SIW61373.1							
Lep_00	Lytic		Lytic polysaccharide							
06615	polysaccharide		monooxygenase, partial		SIW61	3.7E-	87.87	61.61		
4-RA	monooxygenase	228	[Thermobia domestica]		373	11	879	78	33	29
			gi 759036658 ref XP_011345707							
			.1 PREDICTED: uncharacterized							
			protein LOC105284120		XP_01					
			[Ooceraea	Hym	13457					
Lep_00	Lytic		biroi]gi 607367294 gb EZA6144	eno	07,					
04517	polysaccharide		1.1 hypothetical protein	pter	EZA61	3.24E-	71.97	165.2		
1-RA	monooxygenase	835	X777_07774 [Ooceraea biroi]	а	441	48	452	36	157	113
			gi 759036658 ref XP_011345707							
			.1 PREDICTED: uncharacterized							
			protein LOC105284120		XP_01					
			[Ooceraea	Hym	13457					
Lep_00	Lytic		biroi]gi 607367294 gb EZA6144	eno	07,					
04517	polysaccharide		1.1 hypothetical protein	pter	EZA61	4.63E-	66.08	125.5		
2-RA	monooxygenase	357	X777_07774 [Ooceraea biroi]	а	441	35	696	61	115	76
			gi 1009534315 ref XP_01590414							
Lep_00	Lytic		0.1 uncharacterized protein	Arac	XP_01					
09944	polysaccharide		LOC107436837 [Parasteatoda	hnid	59041	1.1E-	73.56	211.4		
1-RA	monooxygenase	724	tepidariorum]	а	40	66	322	6	174	128

Table 3.21. Continued

Seque	Sequence desc.	Seq	Hit desc.	Ord	Hit	Е-	Simil	Bit-	Alig	Pos
nce		uen		er	ACC	Value	arity	Score	nme	itiv
name		ce							nt	es
		leng							leng	
		th							th	
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
08432	polysaccharide		monooxygenase, partial		SIW61	1.61E-	77.84	230.3		
4-RA	monooxygenase	883	[Thermobia domestica]		372	73	091	35	176	137
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
05986	polysaccharide		monooxygenase, partial		SIW61	1.75E-	72.97	234.5		
8-RA	monooxygenase	618	[Thermobia domestica]		372	76	297	72	185	135
			gi 1349059417 emb SIW61373.1							
Lep_00	Lytic		Lytic polysaccharide							
05986	polysaccharide		monooxygenase, partial		SIW61	1.49E-	93.33	60.84		
7-RA	monooxygenase	295	[Thermobia domestica]		373	10	333	74	30	28
			gi 939644305 ref XP_014272300							
Lep_00	Lytic		.1 PREDICTED: uncharacterized	Hem	XP_01					
15668	polysaccharide		protein LOC106678350	ipter	42723	7.34E-	73.58	74.71		
1-RA	monooxygenase	171	[Halyomorpha halys]	а	00	16	491	46	53	39
			gi 1349059417 emb SIW61373.1							
Lep_00	Lytic		Lytic polysaccharide							
07279	polysaccharide		monooxygenase, partial		SIW61	2.21E-	78.26	174.8		
4-RA	monooxygenase	384	[Thermobia domestica]		373	54	087	66	115	90

Table 3.21. Continued

Seque	Sequence desc.	Seq	Hit desc.	Ord	Hit	Е-	Simil	Bit-	Alig	Pos
nce		uen		er	ACC	Value	arity	Score	nme	itiv
name		ce							nt	es
		leng							leng	
		th							th	
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
11844	polysaccharide		monooxygenase, partial		SIW61	1.53E-	91.59	223.4		
7-RA	monooxygenase	380	[Thermobia domestica]		372	73	664	02	119	109
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
07870	polysaccharide		monooxygenase, partial		SIW61	3.4E-	88.04	291.5		
7-RA	monooxygenase	588	[Thermobia domestica]		372	99	348	82	184	162
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
07870	polysaccharide		monooxygenase, partial		SIW61	4.25E-	88.05	245.7		
8-RA	monooxygenase	594	[Thermobia domestica]		372	81	031	43	159	140
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
13784	polysaccharide		monooxygenase, partial		SIW61	1.15E-	75.75	164.0		
4-RA	monooxygenase	471	[Thermobia domestica]		372	49	758	81	132	100
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
06057	polysaccharide	103	monooxygenase, partial		SIW61	4.49E-	73.51	233.8		
3-RA	monooxygenase	8	[Thermobia domestica]		372	74	351	02	185	136
Table 3.21. Continued										

Seque	Sequence desc.	Seq	Hit desc.	Ord	Hit	E-	Simil	Bit-	Alig	Pos
nce		uen		er	ACC	Value	arity	Score	nme	itiv
name		ce							nt	es
		leng							leng	
		th							th	
			gi 1349059415 emb SIW61372.1							
Lep_00	Lytic		Lytic polysaccharide							
00658	polysaccharide		monooxygenase, partial		SIW61	1.05E-	72.85	276.5		
7-RA	monooxygenase	870	[Thermobia domestica]		372	91	714	59	210	153
			gi 1349059417 emb SIW61373.1							
Lep_00	Lytic		Lytic polysaccharide							
00659	polysaccharide		monooxygenase, partial		SIW61	4.21E-	56.52	184.1		
0-RA	monooxygenase	941	[Thermobia domestica]		373	55	174	11	207	117
			gi 1238881740 ref XP_02225383							
			4.1 uncharacterized protein							
			LOC106469570 isoform X1		XP_02					
			[Limulus		22538					
			polyphemus]gi 1238881742 ref		34,					
Lep_00	Lytic		XP_022253835.1 uncharacterize	Arth	XP_02					
00658	polysaccharide	136	d protein LOC106469570 isoform	ropo	22538	6.63E-	55.55	204.9		
8-RA	monooxygenase	3	X1 [Limulus polyphemus]	da	35	60	556	12	288	160
			gi 1009534315 ref XP_01590414							
Lep_00	Lytic		0.1 uncharacterized protein	Arac	XP_01					
00658	polysaccharide		LOC107436837 [Parasteatoda	hnid	59041	1.22E-	65.82	75.48		
9-RA	monooxygenase	335	tepidariorum]	а	40	15	278	5	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	ndoglucanase β-glucosidase		
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		<u> </u>		
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).



Enzyme Code Dist. (Hydrolases) [thermobia_domestica_cds__1_]

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.



Glycoside hydrolases and LPMOs - T. domestica

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*.



Annotation results of coding sequences of *Ctenolepisma longicaudata*

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.



Glycoside hydrolases and LPMOs - C. longicaudata

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 upregulated in the foregut tissue of paper-fed *T. domestica* than other diets, yet LPMO upregulation 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°C.

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°C 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°C. 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*_{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 upregulated 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 329 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 upregulation 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 330 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 upregulation 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 upregulated in the rest of the body sample of *C. longicaudata* fed on CMC compared to protein 332

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 -1,6glucanase, 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 333 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 βglucosidases (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. dom*estica 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 upregulated 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 337

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 βglucosidases, 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 ^a 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 enzymecompatible 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 solventbased 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-13	Paper	0.02
Th_d_00093510-RA	hydrolase family 9	Paper	0.03
	uncharacterized family	Paper	0.005
	31 glucosidase		
Th_d_00048538-RA	KIAA1161-like		
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-13	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	Padj value
Th_d_00005705-RA	hydrolase family 9	Paper	0.001
Th_d_00018098-RA	β-glucuronidase-like	Paper	0.04
	uncharacterized family	Paper	0.02
	31 glucosidase		
Th_d_00048538-RA	KIAA1161-like		
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	СМС	< 0.001
Th_d_00056727-RA	СМС	0.03
Th_d_00064622-RA	СМС	0.02
Th_d_00068335-RA	СМС	0.02
Th_d_00109123-RA	СМС	0.002
Th_d_00110473-RA	СМС	0.04
Th_d_00120034-RA	СМС	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.
10010		Gommada

Diets tested	Up-regulated diet	Padj value
Th_d_00084210-RA	Paper	0.04
Th_d_00120998-RA	Paper	0.03
Switchgrass vs CMC		
Coding sequence		
Th_d_00014400-RA	СМС	< 0.001
Th_d_00064622-RA	CMC	0.001
Th_d_00098297-RA	СМС	0.04
Th_d_00104344-RA	СМС	0.005
Th_d_00109123-RA	СМС	0.003
Th_d_00110473-RA	СМС	0.03
Th_d_00120034-RA	СМС	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-		СМС	0.02
RA	hydrolase family 9		
Th_d_00104189-		СМС	0.02
RA	hydrolase family 9		
Paper vs Protein			
Coding sequence			
Th_d_00071154-		Paper	< 0.001
RA	hydrolase family 9		
Th_d_00029874-		Paper	0.003
RA	hydrolase family 9		
Th_d_00041503-		Paper	0.006
RA	Endoglucanase		
Th_d_00000348-		Paper	0.002
RA	Endoglucanase		
Th_d_00018984-		Paper	0.04
RA	hydrolase family 9		
Th_d_00045439-		Paper	0.008
RA	β-1,4-glucanase		
Th_d_00031500-		Paper	0.003
RA	Endoglucanase		
Th_d_00015659-		Paper	0.04
RA	hydrolase family 9		
Th_d_00007226-		Paper	0.02
RA	β-1,4-endoglucanase 1		
Th_d_00111221-		Paper	0.02
RA	hydrolase family 9		
Th_d_00005705-		Paper	< 0.001
RA	hydrolase family 9		
Th_d_00104189-		Paper	< 0.001
RA	hydrolase family 9		
Th_d_00046715-		Paper	< 0.001
RA	hydrolase family 9		
Th_d_00028499-		Paper	< 0.001
RA	hydrolase family 9		

Table 4.3. Continued

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

Table 4.3. Continued.

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

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 ($P_{adj} < 0.05$).

Diets tested	Up-regulated diet	P _{adj} value
CMC vs Protein		
Coding sequence		
Th_d_00037467-RA	СМС	0.03
Paper vs CMC		
Coding sequence		
Th_d_00056739-RA	СМС	< 0.001
Th_d_00068347-RA	Paper	0.04
Th_d_00083583-RA	СМС	0.04
Th_d_00120998-RA	СМС	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	Padj 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-13	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	СМС	0.04
Lep_00022398-RA	hydrolase family 9	Paper	0.002
	endo- β-1,4-	Paper	0.02
Lep_00034832-RA	glucanase		
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
	endo- β-1,4-	Paper	0.04
Lep_00104311-RA	glucanase		
Lep_00115213-RA	hydrolase family 9	Paper	< 0.001
Lep_00135466-RA	hydrolase family 9	Paper	< 0.001
Switchgrass vs Paper			
Coding sequence			
	endo- β-1,4-	Paper	0.005
Lep_00029521-RA	glucanase		
Lep_00036184-RA	hydrolase family 9	Paper	0.005
	β-1,3-glucan-binding	Paper	< 0.001
Lep_00039163-RA	-like	_	
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 ($P_{adj} < 0.05$).

Diets tested	Enzyme	Up-regulated diet	Padj value
CMC vs Protein			
Coding sequence			
Lep_00004873-RA	Glucosidase	СМС	0.04
Lep_00011407-RA	Glucosidase	Protein	0.04
Lep_00015101-RA	Glucosidase	СМС	0.02
Lep_00020108-RA	Glucosidase	СМС	0.006
Lep_00026156-RA	hydrolase family 9	СМС	0.03
Lep_00036877-RA	β-galactosidase-12	СМС	0.03
Lep_00038648-RA	hydrolase family 9	СМС	0.04
Lep_00052604-RA	Glucosidase	СМС	0.002
Lep_00059101-RA	β-galactosidase-12	СМС	0.006
	uncharacterized	СМС	0.001
	family 31 glucosidase		
Lep_00079073-RA	KIAA1161-like		
	uncharacterized	СМС	< 0.001
	family 31 glucosidase		
Lep_00094746-RA	KIAA1161-like		
Lep_00097528-RA	Glucosidase	СМС	0.02
Lep_00098705-RA	Glucosidase	СМС	0.001
	uncharacterized	СМС	< 0.001
	family 31 glucosidase		
Lep_00102111-RA	KIAA1161-like		
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-12	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	Padj value
	uncharacterized	Protein	0.04
	family 31 glucosidase		
Lep_00102111-RA	KIAA1161-like		
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
	β-1,3-glucan-binding -	Switchgrass	0.04
Lep_00034316-RA	like		
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
	uncharacterized	Protein	0.04
	family 31 glucosidase		
Lep_00077538-RA	KIAA1161-like		
	uncharacterized	Protein	0.03
	family 31 glucosidase		
Lep_00079073-RA	KIAA1161-like		
Lep_00095117-RA	endo- β-1,4-glucanase	Switchgrass	0.005
Lep_00096399-RA	hydrolase family 9	Protein	<0.001
	uncharacterized	Protein	< 0.001
	family 31 glucosidase		
Lep_00102111-RA	KIAA1161-like		
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	Padj 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	СМС	< 0.001
Lep_00016311-RA	endo- β-1,4-glucanase	СМС	< 0.001
Lep_00016312-RA	hydrolase family 9	СМС	< 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
	β-1,3-glucan-binding -	Switchgrass	0.04
Lep_00034316-RA	like		
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	Padj 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-13	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	Padj 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 ($P_{adj} < 0.05$).

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	Padj 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 ($P_{adj} < 0.05$) were blasted against the NCBInr database, and the matching protein and the tissue in *T. domestica* are listed.

Sequence_I	Identified	Organism	Que	Ε	%	Accession	Sam
d	protein		ry	valu	identi		ple
			cov	е	ty		
		Carlatana	er	6	F00/	A A 171 2 2 2 0 1	Dest
1n_a_00000	Cellulase	Coptotermes	90%	66-	59%	AAK12339.1	Rest
340				180			body
Th d 00000	Hypothotic	S	550%	4.0-	1.1.06	DNF26267 1	Forog
351	al	s secundus	5570	61	TT /0	1 11 50507.1	nt
551	nroteinB7P	5 Securitus		01			ut
	43 G18040						
	(β-1,4-						
	glucanase)						
Th_d_00000	Cellulase	Coptotermes	74%	2e-	65%	AAK12339.1	Foreg
352		acinaciformi		45			ut
		S					_
Th_d_00000	β-1,4-	Cryptoterme	77%	1e-	52%	XP_02370492	Rest
353	endoglucan	s secundus		93		9.1	of the
Th d 00000	ase	Diuranhia	2404	20	5704	VD 01527072	Douy
354	2 Elluoglucali	novia	2490	2e- 19	57%	81	of the
554		ΠΟΧΙά		17		0.1	body
Th d 00001	β-	Cephus	98%	0	54%	XP 01559854	Foreg
122	galactosida	cinctus			- , 0	7.1	ut
	se						
Th_d_00005	β-1,4-	Mastoterme	80%	0	66%	CAD54729.1	Both
705	endoglucan	S					
	ase	darwiniensis					_
Th_d_00006	β-	Daphnia	96%	2e-	46%	EFX89607.1	Foreg
652	galactosida	pulex		114			ut
Th d 00006	se ß_	Austrofundu	700/2	60-	1.20%	VD 01396316	Forog
653	p- galactosida	lus limnaeus	7 9 70	48	4270	8 1	nt
033	se	ius inniucus		10		0.1	ut
Th_d_00006	β-	Monodelphis	58%	2e-	70%	XP_00750509	Foreg
654	galactosida	domestica		20		6.1	ut
	se						

Table 4.9. Continued.

Sequence_I d	Identified protein	Organism	Que ry cov	E valu e	% identi ty	Accession	Sam ple
			er		-		
Th_d_00007 226	β-1,4- endoglucan ase	Panesthia cribrata	69%	5e- 172	55%	AAF80584.1	Rest of the body
Th_d_00015 659	β-1,4- endoglucan ase	Zootermopsi s nevadensis	63%	3e- 90	69%	XP_02194132 4.1	Rest of the body
Th_d_00018 098	β- glucuronid ase	Zootermopsi s nevadensis	96%	8e- 146	62%	KDR08779.1	Fore gut
Th_d_00018 632	β- glucosidase	Neotermes koshunensis	98%	1e- 178	61%	3AHZ_A	Rest of the body
Th_d_00018 984	β-1,4- endoglucan ase	Cryptoterme s secundus	42%	3e- 91	66%	PNF24409.1	Rest of the body
Th_d_00028 499	β-1,4- endoglucan ase	Mastoterme s darwiniensis	75%	0	65%	CAD54729.1	Both
Th_d_00029 874	β-1,4- endoglucan ase	Blattella germanica	71%	3e- 76	48%	PSN31180.1	Rest of the body
Th_d_00031 500	β-1,4- endoglucan ase	Cryptoterme s secundus	96%	5e- 125	49%	XP_02370492 9.1	Both
Th_d_00034 779	β-1,4- endoglucan ase	Mastoterme s darwiniensis	88%	0	67%	CAD54729.1	Both
Th_d_00035 294	β-1,4- endoglucan ase	Zootermopsi s nevadensis	87%	8e- 179	59%	XP_02194132 4.1	Both
Th_d_00035 324	β-1,4- endoglucan ase	Coptotermes acinaciformi s	99%	4e- 163	58%	AAK12339.1	Rest of the body
Th_d_00036 630	Endo- β- 1,4- mannanase	Orchesella cincta	63%	1e- 75	44%	ODM91993.1	Fore gut

Table 4.9. Continued.

Sequence_I d	Identified protein	Organism	Que ry	E valu	% identi tv	Accession	Sam ple
			er	C	ey		
Th_d_00038 486	Glycoside hydrolase family 9	Aretaon asperrimus	83%	2e- 171	65%	AMH40360.1	Rest of the body
Th_d_00041 503	β-1,4- endoglucan ase	Nasutiterme s takasagoens is	92%	2e- 141	56%	1KS8_A	Rest of the body
Th_d_00045 439	β-1,4- endoglucan ase	Anoplophor a glabripennis	38%	1e- 21	53%	XP_01856819 4.1	Rest of the body
Th_d_00046 715	β-1,4- endoglucan ase	Zootermopsi s nevadensis	92%	2e- 142	63%	XP_02194132 4.1	Both
Th_d_00048 538	Glucosidase family 31	Periplaneta americana	70%	3e- 174	53%	AIA09350.1	Foreg ut
Th_d_00061 520	β-1,4- endoglucan ase	Zootermopsi s nevadensis	96%	6e- 41	58%	KDR06579.1	Foreg ut
Th_d_00071 154	Cellulase	Haliotis kamtschatk ana	19%	1e- 21	77%	ACS15347.1	Both
Th_d_00093 510	Cellulase	Cryptoterme s secundus	100 %	5e- 38	62%	PNF24409.1	Foreg ut
Th_d_00101 434	Endo- β- 1,4- mannanase	Cherax quadricarin atus	65%	7e- 26	56%	AIN40245.1	Rest of the body
Th_d_00104 189	Endoglucan ase E-4-like	Blattella germanica	56%	2e- 47	74%	PSN33998.1	Rest of the body
Th_d_00107 645	β-1,4- endoglucan ase	Eurytemora affinis	91%	4e- 38	60%	XP_02332980 7.1	Rest of the body
Th_d_00111 221	Endoglucan ase 4-like	Orbicella faveolata	57%	1e- 36	60%	XP_02061486 7.1	Rest of the body

Table 4.9. Continued.

Sequence_I d	Identified protein	Organism	Que ry cov er	E valu e	% identi ty	Accession	Sam ple
Th_d_00118 343	Endoglucan ase 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 ($P_{adj} < 0.05$) were blasted against NCBInr database, and the matching protein and tissue in *C. longicaudata* are listed.

Sequence_I d	Identified protein	Organism	Quer y cove r	E valu e	% identi ty	Accession	Samp le
Lep_000448 91-RA	Uncharacteri zed protein LOC1108405 36 (endo-β- 1,4- glucanase)	Zootermop sis nevadensis	94	0.0	66	XP_0219413 24.1	Both
Lep_000987 05-RA	β- glucosidase	Acanthast er planci	25	5e- 07	55	XP_0220905 46.1	Rest of the body
Lep_000775 38-RA	α- glucosidase family 31	Periplanet a americana	89	0.0	53	AIA09350.1	Rest of the body
Lep_000878 18-RA	Hypothetical protein LOTGIDRAF T 101222 (endo-β-1,6- glucanase)	Lottia gigantea	70	3e- 122	49	XP_0090542 80.1	Rest of the body
Lep_000363 80-RA	Myrosinase 1-like (β- glucosidase)	Cryptoter mes secundus	88	2e- 104	50	XP_0237211 12.1	Rest of the body
Lep_001152 13-RA	Endoglucana se 7-like	Zootermop sis nevadensis	84	1e- 80	72	XP_0219413 22.1	Both
Lep_000048 72-RA	β- glucosidase	Coptoterm es formosanu s	76	2e- 145	56	A0Y34571.1	Rest of the body

Table 4.10. Continued.

Sequenc e_Id	Identified protein	Organism	Query cover	E value	% iden tity	Accession	Sam ple
Lep_0000 4873-RA	Myrosinase 1-like (β- glucosidase)	Cryptotermes secundus	94	8e-50	68	XP_0237237 74.1	Rest of the bod y
Lep_0006 5345-RA	β- glucuronidas e-like	Centruroides sculpturatus	49	2e-11	51	XP_0232218 87.1	Rest of the bod y
Lep_0007 1461-RA	Uncharacteri zed protein LOC1108405 36 (endo-β- 1,4- glucanase)	Zootermopsis nevadensis	94	1e- 105	72	XP_0219413 24.1	Both
Lep_0001 9550-RA	Hypothetical protein B5V51_1109 0 (β- glucosidase)	Heliothis virescens	74	8e-50	34	PCG75727.1	Rest of the bod y
Lep_0002 1305-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 bod y
Lep_0004 8639-RA	β-glucosidase	Salganea esakii	59	7e-24	55	BA085048.1	Rest of the bod y

Table 4.10. Continued.

Sequenc	Identified	Organism	Query	E	%	Accessio	Sam
e_Id	protein	<i>a</i>	cover	value	identity	n	ple
Lep_0005	Hypothetical	Cryptotermes	88	16-67	62		Rest
3730-KA	B7D43 C097	secundus				51.1	01 the
	39 (R-						bod
	glucuronidas						v
	e)						5
Lep_0005	β-	Cephus	74	9e-65	51	XP_0155	Rest
9101-RA	galactosidase	cinctus				98574.1	of
	-1-like						the
	protein 2						bod
Lop 0007	Endoglucanac	Plattolla	77	60.70	60	DCN220	y Poth
1754-RA	e E-4	aermanica	//	00-79	00	98.1	Dotti
Lep 0005	Hypothetical	Cryptotermes	96	9e-68	62	PNF220	Fore
5855-RA	protein	secundus				51.1	gut
	B7P43_G097						-
	39 (β-						
	glucuronidas						
Lap 0001	e) Endeglugenee	Zaatarmaanaia	04	0. 75	71	VD 021	Deat
Lep_0001	Endogiucanas	Zootermopsis	84	8e-75	/1	AP_021 0/1222	Rest
0310-NA	e / like	nevuuensis				1	body
Lep 0001	Cellulase	Coptotermes	71	3e-55	58	AAK123	Rest
6311-RA		acinaciformis				39.1	of the
							body
Lep_0001	Endoglucanas	Blattella	38	8e-44	65	PSN339	Rest
6312-RA	e E-4	germanica				98.1	of the
Lon 0007	Ilumethetical	Cravatatarmaa	20	1.0	47	DNE224	body
Lep_0007	nypolitetical	socundus	69	1 <i>4</i> -7	47	21 1	rest of the
+510-IX	B7P43 G048	secundus		17/		21.1	body
	91						bouy
	(Uncharacteri						
	zed family 31						
	glucosidase)						
Lep_0006	β-glucosidase	Coptotermes	83	1e-60	50	AGM32	Rest
5284-RA		formosanus				287.1	of the
	1						body

Table 4.10. Continued.

Sequenc	Identified	Organism	Query	Ε	%	Accessi	Sam
e_Id	protein		cover	value	identity	on	ple
Lep_0005 2326-RA	Uncharacteri zed protein LOC1108405 36 (endo-β- 1,4- glucanase)	Zootermopsi s nevadensis	83	1e-98	50	XP_021 941324. 1	Both
Lep_0000 6757-RA	β - galactosidase -1-like protein 2	Cephus cinctus	32	4e-62	52	XP_015 598574. 1	Fore gut
Lep_0003 4744-RA	Glycoside hydrolase family 9	Peruphasma schultei	86	1e- 147	50	AMH40 374.1	Fore gut
Lep_0001 5101-RA	Myrosinase 1-like (β- glucosidase)	Cryptoterme s secundus	96	4e- 147	50	XP_023 721112. 1	Rest of the body
Lep_0001 4981-RA	Cellulase	Antipaluria urichi	31	4e-58	57	AOV942 50.1	Rest of the body
Lep_0001 4982-RA	Putative endo-β-1,4- glucanase of EG1	Odontoterm es formosanus	82	5e-77	56	BAD120 08.1	Rest of the body
Lep_0001 4983-RA	Predicted protein (endo-β-1,4- glucanase)	Nematostell a vectensis	31	7e-23	78	XP_001 640311. 1	Rest of the body
Lep_0007 1132-RA	Uncharacteri zed protein LOC1108405 36 (endo-β- 1,4- glucanase)	Zootermopsi s nevadensis	95	3e- 104	72	XP_021 941324. 1	Both
Lep_0005 1683-RA	Glycoside hydrolase family 9	Timema cristinae	90	1e-23	64	AMH40 395.1	Both

Table 4.10. Continued.

Sequenc e_Id	Identified protein	Organism	Query cover	E value	% identit v	Accessio n	Sam ple
Lep_0005 1684-RA	β-1,4- endoglucanas e 1	Panesthia cribrate	94	8e-36	63	AAF8058 4.1	Both
Lep_0007 9073-RA	Hypothetical protein C0J52_14633 (Uncharacteri zed family 31 glucosidase)	Blatella germanica	60	2e-20	65	PSN3130 8.1	Rest of the body
Lep_0010 4311-RA	Hypothetical protein C0J52_21511 (endo-β-1,4- glucanase	Blatella germanica	99	2e-55	57	PSN3118 0.1	Fore gut
Lep_0004 6431-RA	β- glucuronidas e-like isoform X3	Cryptoterme s secundus	86	3e- 123	57	XP_0237 18877.1	Rest of the body
Lep_0002 8614-RA	β- glucuronidas e-like isoform X3	Cryptoterme s secundus	91	1e-69	52	XP_0237 18877.1	Both
Lep_0002 8613-RA	β- glucuronidas e-like	Halyomorph a halys	42	6e-46	72	XP_0242 17164.1	Rest of the body
Lep_0009 5117-RA	Putative endo- β-1,4- glucanase SmEG1	Sinocapriter mes mushae	99	1e-47	53	BAD120 12.1	Rest of the body
Lep_0006 0393-RA	β-glucosidase	Coptotermes formosanus	81	5e-55	70	AGM323 08.1	Rest of the body
Lep_0003 9163-RA	Hypothetical protein D910_11210 (β-1,3- glucanase)	Dendroctonu s ponderosae	68	1e-09	44	ERL9392 4.1	Fore gut

Table 4.10. Continued.

Sequenc e_Id	Identified protein	Organism	Query cover	E value	% identit	Accessio n	Sam ple
Lep_0008 2758-RA	Lactase- phlorizin hydrolase- like (β- glucosidase)	Acanthochro mis polyacanthu s	90	6e-09	30	XP_0220 78461.1	Rest of the body
Lep_0003 8648-RA	Glycoside hydrolase family 9	Ramulus artemis	35	5e-22	65	AMH403 83.1	Rest of the body
Lep_0003 4316-RA	Hypothetical protein DAPPUDRAF T_203138 (β- 1,3- glucanase)	Daphnia pulex	49	3e-36	42	EFX6903 6.1	Rest of the body
Lep_0006 0202-RA	Endoglucanas e E-4	Blattella germanica	73	2e-91	62	PSN3399 8.1	Both
Lep_0002 2398-RA	Glycoside hydrolase family 9	Timema cristinae	88	5e- 176	61	AMH403 92.1	Both
Lep_0002 6029-RA	Hypothetical protein B7P43_G167 08 (β- galactosidase)	Cryptoterme s secundus	66	7e-65	58	PNF3927 1.1	Rest of the body
Lep_0003 6184-RA	Uncharacteri zed protein LOC1108405 36 (endo- β- 1,4- glucanase)	Zootermopsi s nevadensis	84	0.0	66	XP_0219 41324.1	Both
Lep_0001 2530-RA	Endo- β-1,4- mannanase	Daphnia magna	90	3e- 117	54	KZS1075 2.1	Fore gut

Table 4.10. Continued.

Sequenc e_Id	Identified protein	Organism	Query cover	E value	% identit v	Accessio n	Sam ple
Lep_0005 0000-RA	Protein- glucosylgalac tosylhydroxyl ysine glucosidase- like isoform X2	Spodoptera litura	90	3e-74	52	XP_0228 21190.1	Rest of the body
Lep_0000 8899-RA	β- galactosidase -1-like protein 2 isoform X2	Cryptoterme s secundus	95	0.0	43	XP_0237 17284.1	Rest of the body
Lep_0009 4746-RA	α-glucosidase family 31	Periplaneta Americana	83	5e- 161	56	AIA0935 0.1	Rest of the body
Lep_0010 2111-RA	Hypothetical protein L798_13618 (uncharacteri zed family 31 glucosidase)	Zootermopsi s nevadensis	67	3e-47	66	KDR119 65.1	Rest of the body
Lep_0002 0108-RA	Hypothetical protein LOTGIDRAFT _207250 (β- glucosidase)	Lottia gigantea	20	3e-11	83	XP_0090 64666.1	Rest of the body
Lep_0003 4832-RA	β-1,4- endoglucanas e 1	Mastotermes darwiniensis	70	2e-45	64	AAF6372 4.1	Fore gut
Lep_0002 4218-RA	Hypothetical protein LOTGIDRAFT _101222 (β- 1,6- glucanase)	Lottia gigantea	72	1e-19	43	XP_0090 54280.1	Rest of the body

Table 4.10. Continued.

Sequenc e_Id	Identified protein	Organism	Query cover	E value	% identit v	Accessio n	Sampl e
Lep_0003 6877-RA	β- galactosidase -1-like protein 2	Cephus cinctus	88	3e-60	41	XP_0155 98574.1	Rest of the body
Lep_0002 9521-RA	Endo- β-1,4- glucanase	Panesthia angustipenni s	56	4e-52	51	BAG700 31.1	Foreg ut
Lep_0007 8550-RA	Glycoside hydrolase family 9	Peruphasma schultei	92	1e-80	59	AMH403 74.1	Both
Lep_0013 5466-RA	Glycoside hydrolase family 9	Peruphasma schultei	66	6e-24	73	AMH403 74.1	Both
Lep_0005 2604-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_0009 6399-RA	Hypothetical protein B7P43_G096 74 (Endo- β- 1,4- glucanase)	Cryptoterme s secundus	38	1e-47	62	XP_0219 24915.1	Both
Lep_0002 4904-RA	Myrosinase 1-like isoform X1 (β- glucosidase)	Nilaparvata lugens	82	1e- 145	56	XP_0221 87263.1	Rest of the body
Lep_0002 6156-RA	Endoglucanas e E-4	Blattella germanica	35	2e-70	62	PSN3399 8.1	Rest of the body

Table 4.10. Continued.

Sequenc	Identified protein	Organism	Query	E value	%	Accessio	Sampl
e_Id			cover		iden tity	n	е
Lep_0002 6154-RA	Putative endo-β- 1,4-glucanase HsEG4	Hodotermo psis sjostedti	75	8e-43	70	BAD1200 4.1	Rest of the body
Lep_0009 7528-RA	Lactase-phlorizin hydrolase-like (β- glucosidase)	Pomacea canaliculat a	46	8e-47	64	XP_02511 0416.1	Rest of the body
Lep_0000 6775-RA	Uncharacterized protein LOC110840536 (Endo- β-1,4- glucanase)	Zootermop sis nevadensis	93	0.0	60	XP_02194 1324.1	Both
Lep_0001 1407-RA	Lactase-phlorizin hydrolase-like (β- glucosidase)	Pomacea canaliculat a	47	4e-21	47	XP_02511 0416.1	Rest of the body
Lep_0011 3088-RA	Lactase-phlorizin hydrolase-like isoform X2 (β- glucosidase)	Zootermop sis nevadensis	97	3e-65	66	XP_02192 3731.1	Rest of the body
Lep_0004 2852-RA	β-1,4- endoglucanase	Mastoterm es darwiniens is	96	2e-107	61	CAD5473 0.1	Rest of the body
Lep_0004 2851-RA	Cellulase	Antipaluri a urichi	38	2e-10	49	AOV9425 0.1	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*_{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*_{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

385

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.*

387

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

388

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.