The gut microbiomes of desert *Pachysoma* spp. MacLeay (Coleoptera: Scarabaeidae)

by

Philippa Zena Nel Franzini

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor (Ph.D.) Genetics

In the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

November 2016

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Supervisor: Prof. Don A. Cowan

Co-supervisors: Dr. Jean-Baptise Ramond; Dr. Sandra Ronca

Declaration

I, Philippa Zena Nel Franzini, declare that the thesis/dissertation, which I hereby submit for the degree *Philosophiae Doctor* Genetics (Ph.D.) at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE:

DATE:

Dedication

This thesis is dedicated to my late father. You were my inspiration and my greatest supporter.

You are greatly missed.

Acknowledgments

Foremost, I would like to thank my family for their continued love and support throughout my studies. To my parents who supported me continuously in all my endeavours and who taught me that together we could accomplish anything. To my brother who inspired my love in science and nature as a child. To my husband who has been by my side throughout my Ph.D. and who I couldn't have done this without.

I thank God for all of his blessings and for guiding me through my Ph.D.

I would like to thank the NRF for financial assistance during the Ph.D. study.

I thank the University of Pretoria, Department of Genetics, Centre for Microbial Ecology and Genomics (CMEG) and the Genomics Research Institute (GRI) for allowing me to pursue this study and for supplying financial support and infrastructure.

I owe a huge thank you to my supervisors Prof. Don Cowan, Dr. Jean-Baptise Ramond and Dr. Sandra Ronca for continuously guiding and supporting me. Furthermore, I would like to thank our collaborators Prof Clarke Scholtz and Dr Catherine Sole as well as the members of the Scarab Research Group for their assistance with collections of samples, dissections and support.

I would thank all of my fellow lab-mates at CMEG for your advice, support and friendship throughout this study. In particular I would like to thank Dr, O. Bezuidt, Mr M. van Goethem, Dr A. Valverde and Mr. R. Johnson for your continued assistance.

I especially want to thank Sandra, Storme and Olivier for their encouragement during these past four years.

Surendra Vikram your assistance with bioinformatics analysis was invaluable and you have taught me a great deal during this study.

Abstract

Microbial communities inhabit many environmental niches including the nutrient-rich gut systems of animals, where they are involved in a number of important processes. Insect gut microbiota may assist the host with several functions including synthesis of nutritional components lacking from the host diet and digestion of lignocellulosic materials. It is generally believed that the diet of the host plays an important role in the structure of the gut microbiome. Numerous studies have focused on insects feeding on lignocellulosic diets such as termites, as well as medically and agriculturally important insect species. Few studies have researched the gut microbiota of adult dung beetles. Most scarab beetle species feed on the liquid component of wet dung, whereas *Pachysoma* spp. may feed on lignocellulosic materials within their diet of dry dung, plant detritus or both. This feeding behaviour makes *Pachysoma* an ideal candidate for studying the role that diet has on gut microbiome assembly.

Plant detritus feeding *P. endroedyi* and the dry dung feeding *P. striatum* were collected from Namaqualand, South Africa. The mid- and hindgut of each individual were dissected and mDNA extracted using a phenol-chloroform method. Amplicon sequencing of the bacterial 16S rRNA gene and the fungal ITS region was used to determine inter- and intra-specific differences in microbial community structures. Shotgun sequencing of the entire gut metagenome was carried out on mDNA extracted from whole gut samples. Shotgun sequencing was used for both taxonomic and functional annotation of the *Pachysoma* gut microbiomes.

Both amplicon and shotgun sequencing detected substantial differences in bacterial and fungal diversity between the two *Pachysoma* species. Amplicon sequencing showed the number of bacterial phyla ranged from 6-11 and 4-7 (total 14 phyla) for *P. endroedyi* and *P. striatum,* respectively. Furthermore, a minimal core microbiome was detected with only 2.57% of the bacterial OTUs shared between the two *Pachysoma* species studied. Large intra-specific variations were also noted within both *Pachysoma* species. Fungal communities could not be detected in the gut of *P. endroedyi,* while only two fungal phyla were detected *P.*

striatum gut samples. Metagenome shotgun sequencing detected a greater bacterial diversity (total of 39 phyla) than the 16S rRNA gene amplicon study, although large differences were noted between the two species. Furthermore, shotgun sequencing demonstrated that fungal communities were present in the guts of both *Pachysoma* species. Archaea, viruses and other eukaryotic microorganisms were also present in the gut metagenomes of both *Pachysoma* species.

The functional capacity of the *Pachysoma spp.* gut microbiomes was analysed using shotgun sequencing. Both species had the genetic capacity to degrade cellulose and hemicellulose but not lignin, supporting the suggestion that *P. striatum* feeds on plant material in the dry dung. Furthermore, the functional capacity of the microbiomes of both *Pachysoma* species were comparable, suggesting the ability for both species to feed on either dry dung or plant detritus. The similarity of the functional profiles of the two *Pachysoma* species suggests the existence of a functional rather than phylogenetic core microbiome

This primary study has successfully characterised the phylogenetic and functional profiles of the gut microbiomes of two *Pachysoma* species feeding on different substrates. However, it is still unclear if diet is the primary driver in gut microbiome assembly.

List of Abbreviations

ANOSIM	Analysis of Similarities
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BP	Base pair
CAZy	Carbohydrate-Active enZYme
CCD	Charge coupled device
COG	Clusters of Orthologous Groups
CTAB	Cetyl trimethylammonium bromide
DGGE	Denaturing gradient gel electrophoresis
DictDB	Dictyoptera gut microbiota reference database
DNA	Deoxyribonucleic acid
dNTP	Deoxyribose nucleoside triphosphates
EDTA	Ethylenediaminetetraacetic acid
GH	Glycoside hydrolase
KAAS	KEGG Automatic Annotation System
KEGG	Kyoto Encyclopedia of Genes and Genomes
mDNA	Metagenomic DNA
MEGAN	MEtaGenome ANalyzer
MG-RAST	Metagenomics RAST server
NCBI	National Center for Biotechnology Information
NGS	Next-generation Sequencing
nMDS	Non-Metric Multi-Dimensional Scaling
NR	Non-redundant
ORF	Open reading frame
OTU	Operational taxonomic unit
PBS	Phosphate-buffered saline

PCR	Polymerase chain reaction
pH	Potential of hydrogen
PM	Peritrophic membrane
RDP	Ribosomal Database Project
TRFLP	Terminal Restriction Fragment Length Polymorphism
UV	Ultraviolet

Table of Figures

Figure 1.1: The global distribution of deserts measured against the terrestrial surface temperature (taken from Makhalanyane et al. (2015)). 2

Figure 1.2: Typical vegetation within Namaqualand (Photo courtesy of H. de Klerk). ... 4

Figure 1.3: Basic structure of the larval scarab gut (adapted from Huang et al. (2010)). 14

Figure 2.1: *Pachysoma* samples collected for the present study: a) *P. endroedyi* and b) *P. striatum* (Courtesy of Hennie de Klerk). 48

Figure 2.2: Collection of *Pachysoma* samples: a) Observation of a burrow; b) a typical *Pachysoma* spp. burrow with the food-source (plant detritus) and a beetle inside (Courtesy of Hennie de Klerk). 48

Figure 2.3: Aerial view of the Western coast of South Africa with the Namaqualand Desert collection site indicated by the yellow marker and shown in the insert (Insert map data: Google, AfriGIS (Pty) Ltd, 2016). 49

Figure 2.4: Burrow being examined surrounded by typical vegetation at the collection site (Photo courtesy of H. de Klerk). 49

Figure 2.5: Dissected *P. striatum* showing H: hindgut, M: Midgut and P (square): pyloris (Foregut not shown). 51

Figure 3.1: Cytochrome oxidase I gene Parsimony tree phylogeny of 13 *Pachysoma* spp. Branch colours indicate the diet of the *Pachysoma* spp.: dung (brown), plant detritus (green), polyphagous (blue) and unknown (black). This phylogentic tree was adapted, with permission, from (Sole et al., 2005) and the dietary information taken from both (Harrison et al., 2003) and personal observations by Prof C. Scholtz. Two species of *Scarabaeus*, (*S. proboscideus* and *S. rugosus*) which is the sister-genus to *Pachysoma* and a typical wet-dung-feeder, were used as outgroups. Numbers to the right of the tree indicate the three *Pachysoma* lineages. The two species considered in this study, *P. endroedyi* and *P. striatum*, are indicated with stars (Adapted from Sole et al. (2005)). 64

Figure 3.2: Rarefactions curves showing gut microbial community richness of all *Pachysoma* individuals for bacterial 16S rRNA gene amplicon data of: a) *P. endroedyi*, b) *P. striatum;* and c) fungal ITS region amplicon data of *P. striatum*..... 67

Figure 3.3: nMDS ordination plot based on Bray-Curtis distance matrices of bacterial 16S rRNA gene pyrosequencing data for *P. endroedyi* and *P. striatum* individuals. A stress value of less than 0.1 represents a high quality ordination. *Pachysoma endroedyi* and *P. striatum* are represented by green and inverted brown triangles, respectively. 69

Figure 3.4: Venn diagrams showing distribution of bacterial OTUs between (a) *P. endroedyi* and (b) *P. striatum* individuals based on the 16S rRNA gene pyrosequencing analysis. Shared OTUs are shown in bold. Numerical labels are shown for each individual. 71

Figure 3.5: a) Venn diagram comparing the distribution of fungal OTUs between *P. striatum* individuals based on the ITS region pyrosequencing analysis. Shared OTUs are shown in bold. Numerical labels are given for each individual; b) Relative abundance of fungal phyla in five *P. striatum* individuals based on ITS rRNA gene region pyrosequencing analysis at a 97% identity threshold. 72

Figure 3.6: Comparison of inter-specific differences in relative abundance of bacterial phyla in the gut of two *Pachysoma* spp., *P. endroedyi* and *P. striatum*, based on 16S rRNA gene pyrosequencing analysis. 75

Figure 4.1: Amplification of the a) 18S rRNA and b) 16S rRNA genes from cockroach gut samples subjected to all four host DNA reduction methods. L: molecular weight KAPA Universal Ladder, C: no treatment, 1: removal of PM, 2: filtration, 3: minimal homogenisation and 4: low-speed centrifugation. *Escherichia coli* cells were used as the positive control, therefore, no amplification was expected in gel (a)..... 93

Figure 4.2: Rarefaction curves of gut microbial community richness of a) *P. endroedyi* and b) *P. striatum* gut samples. Eukaryotic contigs derived from Metazoa and Streptophyta were removed before analysis. 96

Figure 4.3: Taxonomic affiliation of shotgun sequencing reads assigned using MG-RAST (a and b), and MEGAN (c and d) for *P. endroedyi* (a and c) and *P. striatum* (b and d) gut metagenomes. 98

Figure 4.4: Diversity and abundance of bacterial phyla in the gut samples of *P. endroedyi* (green) and *P. striatum* (brown) using shotgun (Illumina Hiseq) analysed using MEGAN.. 100

Figure 4.5: Functional assignments of microbial-derived ORFs from both *Pachysoma* species to categories of COGs generated by MEGAN. 102

Figure 4.7: Summary of the plant cell wall degrading enzymes present in the *P. endroedyi* and *P. striatum* gut metagenomes. Only hemicellulases present in at least one of the metagenomes are listed. Enzymes are highlighted according to the metagenome in which they were detected: *P. striatum* (brown) or both *Pachysoma* metagenomes (yellow). 111

Figure 4.8: Adapted KEGG nitrogen metabolism reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes. 113

Figure 4.9: Adapted KEGG peptidoglycan biosynthesis reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes. 115

Figure 4.10: Abundance and diversity of genes encoding energy metabolism pathways (KEGG) for *P. endroedyi* (green) and *P. striatum* (brown) gut microbial metagenomes..... 116

Figure 4.11: Abundance and diversity of genes encoding carbohydrate metabolism pathways (KEGG) for *P. endroedyi* (green) and *P. striatum* (brown) gut microbial metagenomes. 117

Figure 4.12: Adapted KEGG pyridoxine (vitamin B6) metabolism reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes. 118

Figure 4.13: Adapted KEGG Ascorbate (Vitamin C) and Aldarate metabolism reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes. 119

Figure 4.6: Distribution of glycosyl hydrolase (GH) families in the gut metagenomes of *P. endroedyi* and *P. striatum* annotated using hmmer. 104

List of Tables

Table 2.1: List of chemical reagents. 46

Table 2.2: Enzymes used in this study. 46

Table 2.3: Compositions of buffers and solutions used in this study. 47

Table 3.1: Values for sequence reads, OTUs, phyla and diversity indices for bacterial and fungal gut communities of *P. endroedyi* and *P. striatum* individuals 66

Table 3.2: Phylogenetic classification of the most abundant bacterial genera in the gut samples of *P. endroedyi* and *P. striatum*: i.e., representing >2% reads. 77

Table 4.1: Overview of methods tested for efficiency in reducing host DNA contamination during mDNA extraction. 91

Table 4.2: Nanodrop readings of mDNA yields and purity ratios for samples treated with four host DNA reduction methods. 92

Table 4.3: Nanodrop readings of mDNA yields and purity values for replicate *P. endroedyi* and *P. striatum* gut samples sent for shotgun sequencing. 94

Table 4.4: Sequencing metadata and assembly metrics for shotgun sequencing of the *P. endroedyi* and *P. striatum* gut microbiomes using CLC Genomics Workbench, MG-RAST and MEGAN. 95

Table 4.5: Glycosyl hydrolase families detected in the gut metagenomes of *P. endroedyi* and *P. striatum* and the associated KEGG E.C. assignments were applicable. 105

Table of Contents

Declaration II

Dedication III

Acknowledgments IV

Abstract V

List of Abbreviations VII

Table of Figures IX

List of Tables XIII

Chapter 1: Literature Review 1

1.1 Desert Environments 1

1.1.1 The Desert Biome 1

1.1.2 Namaqualand: Study Site 2

1.2 Desert Insect Communities: A Brief Overview 4

1.2.1 Desert Dung Beetles 5

1.2.2 The Genus *Pachysoma* 5

1.3 Diets of Insects 6

1.3.1 The Nutritional Quality of the *Pachysoma* Diet 7

1.3.2 An Overview of Plant Cell Wall Degradation 11

1.4 Structure of the Insect Gut 12

1.5 Insect Gut Microbial Communities 15

1.5.1 Microbial Colonization of the Gut Environment 15

1.5.2 Variations in Gut Microbial Communities 18

1.5.3 Gut Microbial Symbionts and the Relationship between Host Diet and the Gut Microbiome 19

1.5.4 Previous Research on Scarab Beetle Gut Microbiomes 21

1.6 Sequence-Based Methods in Microbial Ecology 22

1.6.1 Amplicon-based Pyrosequencing 23
1.6.2 Environmental Shotgun Sequencing 24
1.7 Aims and Objectives 25
1.7.1 Aims 25
1.7.2 Objectives 26
1.8 References 26
Chapter 2: Materials and Methods 46
2.1 Chemical Reagents 46
2.2 General Methodologies 47
2.2.1 Collection and Storage of *Pachysoma* Samples 47
2.2.2 Insect Identification 50
2.2.3 Gut Dissection 50
2.2.4 Analytical Techniques 51
2.3 454 Amplicon Pyrosequencing Methods – Chapter 3 52
2.3.1 Metagenomic DNA Extraction for Amplicon Pyrosequencing 52
2.3.2 454 Amplicon Pyrosequencing 53
2.4 Hiseq Shotgun Sequencing Methods – Chapter 4 55
2.4.1 Metagenomic DNA Extraction 55
2.4.2 PCR Amplification of 16S rRNA and 18S rRNA Genes 55
2.4.3 Metagenome Sequencing and Data Analysis 56
2.5 References 57
Chapter 3: Amplicon Sequencing of Bacterial and Fungal Gut Communities of *Pachysoma* Species 61
3.1 Introduction 61
3.2 Results and Discussion 63
3.2.1 The Desert Beetle Genus *Pachysoma* 63

3.2.2 Sequencing Outputs and Diversity Indices of the Bacterial 16S rRNA Gene and Fungal ITS Region of the *Pachysoma* Gut Microbiome 65

3.2.3 Inter-specific Variations in Bacterial and Fungal *Pachysoma* Gut Communities 68

3.2.4 Intra-specific Variation of *Pachysoma* Gut Microbial Communities 70

3.2.5 Phylogenetic Diversity of Bacterial and Fungal *Pachysoma* Gut Communities 74

3.3 Conclusion 78

3.4 References 79

Chapter 4: Shotgun Metagenomic Sequencing of the Gut Microbiomes of *Pachysoma endroedyi* and *Pachysoma striatum* 88

4.1 Introduction 88

4.2 Results and Discussion 89

4.2.1 Comparison of Four Previously Published Host DNA Reduction Methods. 89

4.2.2 Shotgun Sequencing of the Gut Microbiomes of *P. endroedyi* and *P. striatum* 94

4.2.3 Community Composition of the *Pachysoma* Metagenome 96

4.2.4 Functional Profiling of the Metagenomes of *P. endroedyi* and *P. striatum*... 101

4.3 Conclusion 120

4.4 References 120

Chapter 5: Concluding Remarks 129

5.1 Studying the *Pachysoma* gut microbiome 129

5.2 Limitations of this study 131

5.3 Future study prospects 133

5.4 References 135

Chapter 1: Literature Review

1.1 Desert Environments

1.1.1 The Desert Biome

Deserts make up approximately a third of the global terrestrial biome (Laity, 2009). Deserts are located globally (Figure 1.1), with 57.2% of the African continent considered arid (Peel et al., 2007), including the world's largest desert (the Sahara Desert (Harris, 2003)), while the only continent absent of any known deserts is Europe (Harris, 2003).

A desert can be defined as an environment where the ratio of precipitation to potential evaporation (P/PET) is less than 1 (Unep, 1992). Using this definition, deserts can further be defined according to four categories of aridity, namely dry sub-arid ($0.5 \leq P/PET < 0.65$), semi-arid ($0.2 \leq P/PET < 0.5$), arid ($0.05 \leq P/PET < 0.2$), and hyperarid ($P/PET < 0.05$). Although most deserts are known as being extremely hot, polar deserts, such as those in Antarctica, are known for their exceptionally cold climates (Cowan and Tow, 2004). For this reason deserts are further classed as hot (>18°C) and cold deserts (<18°C) (Peel et al., 2007).

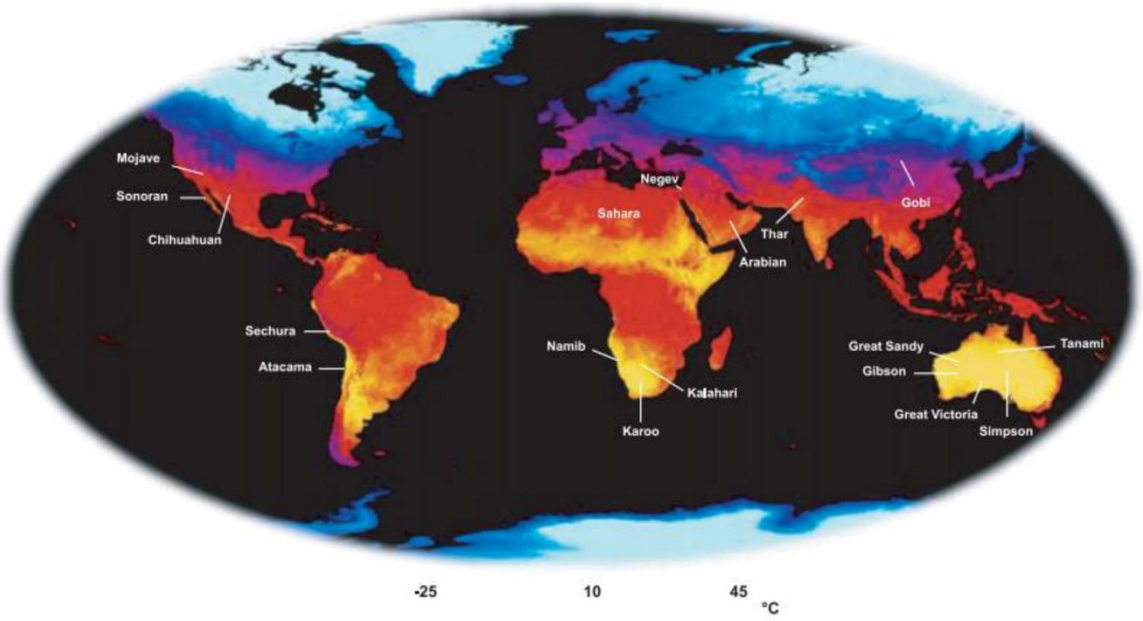


Figure 1.1: The global distribution of deserts measured against the terrestrial surface temperature (taken from Makhalanyane et al. (2015)).

1.1.2 Namaqualand: Study Site

The name Namaqualand is derived from the Khoe-speaking Nama people who originally inhabited the land along with the San people (Wisborg and Rohde, 2005, Benjaminsen et al., 2006). Namaqualand is a semi-arid coastal desert forming part of the Succulent Karoo biome in the Northern Cape Province of South Africa (Cowling et al., 1999, Benjaminsen et al., 2006, Desmet, 2007). Various studies report the approximate size of Namaqualand to vary from 45000-52600km^2 (Cowling et al., 1999, Wisborg and Rohde, 2005, Benjaminsen et al., 2006, Desmet, 2007, May and Lahiff, 2007). The Gariep River separates Namaqualand from Namibia, while the Olifants River and Bokkeveld escarpment form the southern boundary with vast Bushman planes surrounding the eastern side of the desert (Desmet, 2007). The Atlantic Ocean forms the western border of the desert (Benjaminsen et al., 2006, Desmet, 2007).

The climate of Namaqualand varies according to region. Winter rainfall occurs towards the western coastal areas with summer rainfall in the interior regions (Desmet, 2007). The

annual rainfall in the region is relatively low at approximately 50-400 mm (Cowling et al., 1999, Desmet, 2007). Namaqualand is relatively cold due to the south-westerly sea breeze, keeping maximum average summer temperatures below 30°C (Desmet, 2007). In contrast, winter temperatures typically exceed 35°C (Cowling et al., 1999, Desmet, 2007). Due to the coastal location of the Namaqualand, supplemented by high humidity with low temperatures at night, desert fog events and heavy dew are common occurrences, particularly in winter months (Cowling et al., 1999, Desmet and Cowling, 1999, Desmet, 2007) .

Several bioregions (biophysical regions which are classified according to physical environment, climate and flora), are found within this arid ecosystem (Desmet, 2007). Namaqualand covers approximately a quarter of the Succulent Karoo (Desmet, 2007), which is listed as one of the global 200 priority ecoregions (Olson and Dinerstein, 2002) as well as a biodiversity hotspot (Myers et al., 2000). Namaqualand boasts more than 3500 plant species in 135 families (Desmet, 2007). Approximately 25% of the Namaqualand flora is endemic to this region, although this value is estimated to actually be between 40 and 50% (Cowling et al., 1999, Desmet, 2007). The Namaqualand uplands and central Namaqualand coast harbour 1109 and 432 plant species, respectively, 71 and 74 of which are on the Red Data List of endangered species (SKEP, 2008). The most diverse plant families within Namaqualand include Aizoaceae (ice plants), Asteraceae (Aster family) and Iridaceae (Iris family) all consisting of over 200 species (Desmet, 2007). Figure 1.2 shows an example of the typical vegetation which is found in parts of Namaqualand.

Anthropogenic activities impact desert ecosystems, with the most prominent being commercial livestock farming (Benjaminsen et al., 2006, O'Farrell et al., 2011) and mining activities, both of which are prominent in this area (Benjaminsen et al., 2006, May and Lahiff, 2007). The Namaqualand area is considered to be a global conservation priority although only a small portion of the region and the associated vegetation is currently protected (Desmet, 2007, SKEP, 2008).

Figure 1.2: Typical vegetation within Namaqualand (Photo courtesy of H. de Klerk).

1.2 Desert Insect Communities: A Brief Overview

The fauna and flora able to survive in extreme desert environments are typically found in low abundance, although diversity may be high (Harris, 2003). Organisms that manage to survive in desert environments benefit in several ways including exploitation of available resources (e.g. food-sources) and reduced competition (Wharton, 2002). Insects are amongst the most abundant members of the kingdom Animalia found in desert ecosystems due to their ability to adapt to and survive under extreme conditions (Wharton, 2002, Harris, 2003).

The main challenges affecting survival in desert ecosystems are the extreme heat and lack of water (Harris, 2003, Allaby, 2006). Desert fauna are largely dependent on the available plants as a valuable water and nutritional source (Harris, 2003, Allaby, 2006). Therefore, the reduced ability of plants to adapt to desert environments in turn affectively reduces the number of animals that can successfully inhabit these areas (Allaby, 2006).

Numerous insect orders have been reported in arid ecosystems including, but not restricted to Hymenoptera (Marsh, 1986, Struck, 1994, Al-Houty, 1997, Tigar and Osborne, 1999, Forbes et al., 2005, Al-Houty, 2011, Piñero et al., 2011), Isoptera (Al-Houty, 1997, Forbes et al., 2005), Dictyoptera (Al-Houty, 1997, 2011), Odonata (Al-Houty, 1997, Tigar and Osborne, 1999, Suhling et al., 2003, Forbes et al., 2005), Lepidoptera (Struck, 1994, Al-Houty, 1997, Tigar and Osborne, 1999, Forbes et al., 2005, Al-Houty, 2011), Diptera (Struck, 1994, Al-Houty, 1997, Tigar and Osborne, 1999, Forbes et al., 2005, Al-Houty, 2011), Orthoptera (Al-Houty, 1997, Tigar and Osborne, 1999, Forbes et al., 2005, Al-Houty, 2011, Piñero et al., 2011) and Coleoptera (Struck, 1994, Al-Houty, 1997, Tigar and Osborne, 1999, Kazmi and Ramamurthy, 2004, Forbes et al., 2005, Al-Houty, 2011, Piñero et al., 2011), although these differ according to the respective arid region.

1.2.1 Desert Dung Beetles

Coleoptera (commonly known as beetles) is the largest of 30 orders within the class Insecta (Roskov et al., 2016). The subfamily Scarabaeinae (scarab beetles) encompasses more than 225 genera consisting of 5700 valid species of dung beetles (Scholtz et al., 2009), with 670 known species in Southern Africa (Mlambo et al., 2015). Scarabaeinae consists of 11 tribes, namely: Ateuchini, Deltochilini, Eucraniini, Gymnopleurini, Oniticellini, Onitini, Coprini, Onthophagini, Phanaeini, Scarabaeini, and Sisyphini (Bouchard et al., 2011). Contradictory to what the name may suggest, species of dung beetles may also feed on carrion, fruit, fungi and plant detritus (Scholtz et al., 2009, Mlambo et al., 2015). This study will focus on the genus *Pachysoma* MacLeay (1821) which forms part of the tribe Scarabaeini (Bouchard et al., 2011).

1.2.2 The Genus *Pachysoma*

The taxonomic level of *Pachysoma* has been consistently disputed. *Pachysoma* was originally a generic classification (Ferreira, 1953) before being synonymized with its closest relative *Scarabaeus* Linnaeus, 1758 (Mostert and Holm, 1982). Later studies determined

Pachysoma to be a subgenus of *Scarabaeus* (Harrison and Philips, 2003, Forgie et al., 2005). However, *Pachysoma* currently retains full generic status through collective morphometric and genetic evidence (Forgie et al., 2006, Sole et al., 2007, Mlambo et al., 2015).

The genus *Pachysoma* (Scarabaeoidea: Scarabaeidae) consists of 13 species endemic to the south-western African coastal desert (Cape Town, South Africa to Walvis Bay, Namibia) (Harrison et al., 2003). While Scarabaeus are distributed around the world in a variety of habitat types, distributions of individual desert *Pachysoma* species are usually restricted (Harrison et al., 2003).

Pachysoma species differ most significantly from other dung beetles in their diet and foraging behaviour. While most insects of the same genus feed on similar diets, *Pachysoma* species vary in the dietary materials on which they feed: plant detritus, dung or both (Harrison et al., 2003). Furthermore, *Scarabaeus* species typically feed on wet dung from herbivorous mammals. They form balls from the dung and roll them backwards to a suitable area for burial below the surface (Doube, 1990). *Pachysoma* deviate from this behaviour by feeding on dry dung and plant detritus. *Pachysoma* dig a burrow once suitable food has been located, repeatedly foraging and dragging the food forwards to the burrow (Scholtz, 1989). It is hypothesized that *Pachysoma* have probably adapted to feeding on dry dung due to the unavailability of the wet counterpart and the arid environments in which these species are found (Scholtz et al., 2004). In order to exploit this resource, a change in foraging behaviour was necessary as dry dung and detritus cannot be rolled into a ball, thereby modifying the behaviour of *Pachysoma* to dragging small fragments to their burrows (Scholtz et al., 2004).

1.3 Diets of Insects

Insects feed on a wide array of substrates including plant-based diets (e.g. wood, detritus, leaves and roots), nectar, pollen, sap, fungi, fur, feathers, skin, soil, blood, carion and other insects (Brauman, 2000, Douglas, 2009). The chemical composition of each food-source differs, with some dietary components easier to digest than others (Karasov et al., 2011).

Sugars, proteins and lipids are rapidly digested, whereas chitin, lignocelluloses and insoluble starches are harder to digest (Karasov et al., 2011). The composition of refractory and non-refractory chemicals and materials in a variety of dietary substrates is reviewed by Karasov et al. (2011).

Animals feeding on diets high in recalcitrant materials may either assimilate the less recalcitrant compounds in the diet, avoiding indigestible materials or may have adapted to feed on and utilise these recalcitrant materials (Abe and Higashi, 1991). *Pachysoma* species are believed to have adapted to feed on and utilize recalcitrant materials in their diets (Holter and Scholtz, 2011, 2013).

1.3.1 The Nutritional Quality of the *Pachysoma* Diet

All species of *Pachysoma* appear to feed on and utilise refractive materials within their respective diets (dung or plant detritus). Four *Pachysoma* species are known to primarily feed on plant detritus (*P. hippacrates*, *P. glentoni*, *P. schinzi* and *P. endroedyi* (C. Scholtz Pers Comm; Harrison et al. (2003)). Plant detritus (otherwise known as plant litter), refers to decomposing plant matter. The chemical and material composition of plant detritus differs according to plant species (Hättenschwiler et al., 2008). For this reason, a brief overview of the general composition of plant detritus will be given.

Plant detritus consists of primary (plant residues) and secondary (microbial residues and exudates) resources (Kögel-Knabner, 2002). Plant detritus can be composed of two types of plant tissues: namely parenchymatic and woody tissues (Kögel-Knabner, 2002). Polysaccharides (and lignin when relevant) are the most abundant organic compounds of plant detritus (Kögel-Knabner, 2002), forming approximately 90% of the plant cell wall (Doblin et al., 2010). Plant cell wall polysaccharides include cellulose, xylan, mannose, glucomannans, galactan and pectin (Kögel-Knabner, 2002, Hättenschwiler et al., 2008). Proteins account for approximately 10% of the plant cell wall (Doblin et al., 2010). Other components of plant cell walls include polyphenols, lipids and cutin (Kögel-Knabner, 2002, Hättenschwiler et al., 2008). Storage and non-refractory compounds of plant detritus, such as proteins, starches, pigments,

fructan and α-glycans, are easily digested and form an important resource for microorganisms (Kögel-Knabner, 2002, Karasov et al., 2011). Lignified plant material is considered to be nutrient deficient with low concentrations of nitrogen, amino acids, sterols and vitamin B (Dillon and Dillon, 2004, Brune, 2013). However, plant materials with low concentrations of lignin are a protein and nitrogen rich source (Brune, 2013).

The plant litter chemical composition varies, with high carbon concentrations (45.3 and 52.4% (Hättenschwiler et al., 2008)), while nitrogen concentrations (0.68-2.5%) vary according to plant species (Kemp et al., 2003, Hättenschwiler et al., 2008). Along with carbon and nitrogen, phosphorous, iron, aluminium, calcium, potassium, magnesium and manganese concentrations in plant detritus tend to vary over time (Kemp et al., 2003, Goya et al., 2008, Hättenschwiler et al., 2008). However, starch, nitrogen and phosphorous concentrations have been found to be significantly lower in the plant detritus when compared to the living plant (Hättenschwiler et al., 2008). Furthermore, seasonal changes affect the chemical composition of vegetation. Specifically, proteins and carbohydrates are moved to the root systems during the winter seasons, changing the nutritional content of the leafy structures typically fed on by animals (Sinclair, 1975).

Before consumption by insects, detritus undergoes microbial conditioning (Swift et al., 1979), where microorganisms initiate the degradation process and release a portion of the nutrient content. Microbial decomposition over time affects the composition of plant litter (Benner and Hodson, 1985). However, microbial conditioning offers several advantages including the detoxification of allelochemicals, breakdown of the cell wall and immobilization of nitrogen and phosphorous (Nalepa et al., 2001).

Plant detritus and faecal matter are considered to be comparable (Webb, 1977), with dung beetles having been able to routinely move between the two dietary substrates (Cambefort, 1991). An example of this may be *P. benningesi,* which is considered polyphagous, feeding on both dry dung pellets and plant detritus (Harrison et al., 2003) depending on availability of the respective substrate.

In comparison to the known plant detritus feeding *Pachysoma* species, six species are found to feed primarily on dry dung pellets (*P. aesculapius*, *P. garapinum*, *P. denticola*, *P. rotundigenus*, *P. rodriguesi* and *P. striatum*). Dung is excreted organic waste consisting of nutritional materials and chemicals not fully utilised by the animal (Mlambo et al., 2015). Dung can be divided into plant fragments (fibre), which have generally been subjected to mastication by the original animal, and a brown liquid/viscous component (Anderson and Coe, 1974).

In general, dung is composed of digestive juices, ash, albuminous substances, fats, carbohydrates, free living and host associated microorganisms, epithelial cells from the animal, minerals and vitamins (Anderson and Coe, 1974, Lodha, 1974, Müller, 1980, Nalepa et al., 2001, McDowell and Stewart, 2005, Mlambo et al., 2015, Holter, 2016). Furthermore, the chemical and material content, texture and volume of dung varies according to animal species from which the dung is expelled, respective feeding strategies/diets, rainfall and the season, feeding behaviours (I.e., grazing or browsing) as well as whether or not the animal is ruminant (Greenham, 1972, Arman and Hopcraft, 1975, Edwards, 1991, Leeming et al., 1996, Delve et al., 2001, Janecke and Smit, 2015). When larger undigested materials are excreted, the gut mucosa can be damaged resulting in the increase of epithelial cells, mucus and the presence of blood in the stool (Greenham, 1972).

Fibre in dung has been subjected to at least one round of previous digestion, altering the chemical and physical structure of components such as carbohydrates (Lodha, 1974). The concentration of refractive components of plant materials (fibre) in the dung varies across animal species and the ruminant nature of the animal, with cellulose ranging from 15-41%, hemicellulose varying from 7-32% and lignin from 10-44% of the respective faecal organic matter (Holter, 2016).

Nitrogen is a very important chemical component of dung, with greater concentrations in the dung than in the original substrates/vegetation, presumably due to microbial activity in the original animal's intestine (Anderson and Coe, 1974, Lodha, 1974). Concentrations of faecal nitrogen fluctuate according to animal species with nitrogen content higher in dung from

ruminants than non-ruminants (Holter, 2016). Faecal nitrogen is both available from undigested material in the dung or from microbial biomass in the dung and water-soluble nitrogenous compounds (Holter, 2016).

Nine of the ten essential amino acids for insects (Barbehenn et al., 1999, Nation, 2008), as well as some non-essential amino acids are found in dung, albeit sometimes in low concentrations (Müller, 1980, Mason et al., 1989, Rougon et al., 1990). Experimental methods may discriminate against the tenth essential amino acid (Tryptophan) due to strong acids used for protein hydrolysis degrading tryptophan (Edelhoch, 1967). A variety of sterols, including the animal sterol cholesterol, are available from dung of herbivorous and omnivorous animals (Leeming et al., 1996, Tyagi et al., 2008, Derrien et al., 2011), suggesting that the coprophagous insects acquire their essential sterols from their diet.

Dung also contains ash, a mineral residue of soil particles and cell wall silica which constitutes an average of 12 and 16% of the inorganic mass of non-ruminant and ruminant dung, respectively (Holter, 2016). The presence of ash is expected to lower food quality, reducing organic matter concentrations and therefore the availability of nutritional food to the insect (Holter, 2016).

A recent review by Holter (2016) has collected information on the nutritional composition of a large variety of herbivorous mammals to better understand the quality of food afforded to dung beetles. Adult dung beetles feeding on wet dung typically exhibit soft saprophagy, feeding on very small particles in the dung which are generally high in nutritional content (Holter et al., 2002), sifting out large plant fragments (Anderson and Coe, 1974). Holter (2016) hypothesised that it was likely that wet dung feeding beetles feed primarily on microbial biomass and the associated nutrients in the dung. However, coprophagous *Pachysoma* feed on dry dung (Harrison et al., 2003) with studies finding evidence that *Pachysoma* have adapted to feed on the refractive materials in both dung and plant detritus (Holter and Scholtz, 2011, 2013).

Little information is available on dry dung composition compared to that of wet dung (Holter, 2016). As dung ages, concentrations of water-soluble organic matter, hemicellulose and cellulose may decrease significantly, while lignin, ash and water-soluble protein concentrations are likely to remain constant or increase (Lodha, 1974, Holter, 2016). Concentrations of phosphorous also differ between wet and dry dung (McDowell and Stewart, 2005). Nitrogen concentrations are expected to decrease as the dung ages (Holter, 2016). CO_2 production from the dung rapidly increases for the first few days, decreasing as the dung ages and dries out (Anderson and Coe, 1974). CO_2 production is increased when moisture is re-introduced into dry dung (Anderson and Coe, 1974). Some *Pachysoma* species (*P. striatum*) dig burrows below the moisture line, and in so doing may reintroduce moisture into the detritus (Scholtz, 1989). Therefore, it is possible that dry dung has undergone some preconditioning by the involvement of microbial communities, leading to suggestions that *Pachysoma* may feed on the microbial communities and not the refractive materials within the diet (C. Sole Pers. Comm.). However, no evidence of this has been found (Holter et al., 2009). Furthermore, *Pachysoma* have adapted mouthparts and gut systems that suggest they have evolved to feed on cellulosic diets (Holter and Scholtz, 2011, 2013).

1.3.2 An Overview of Plant Cell Wall Degradation

Plant cell walls form a major component of the diet of both detritivorous and coprophagous *Pachysoma* species. Lignin and cellulose are the most abundant natural polymers (Pérez et al., 2002). Cellulose and hemicelluloses are sugar-derived macromolecules, while lignin is an aromatic polymer derived from phenylpropanoids (Pérez et al., 2002). A brief overview of the degradation of lignin, hemicellulose and cellulose is provided, with a focus on microbial degradation.

The first process in plant cell wall degradation is depolarization of lignin (Calderón-Cortés et al., 2012). Lignin is a highly recalcitrant material that is able to withstand the effects of enzymes and chemicals (Ohkuma, 2003). As such, lignin protects other plant cell wall materials including cellulose (Ohkuma, 2003). Few known taxa of organisms are able to

successfully break-down lignin (Ohkuma, 2003). However, there is strong evidence that some insects feeding on healthy wood are able to digest lignin, presumably with the assistance of specific fungal gut symbionts (Geib et al., 2008). Enzymes involved in lignin degradation include laccases, ligninolytic peroxidases (lignin peroxidase, manganese peroxidase and versatile peroxidase), oxidases, dehydrogenases associated with mycelium, aryl-alcohol dehydrogenases (fungal origin) and quinone reductases (Guillén et al., 2005).

Degradation of hemicellulose is the second step in plant cell wall degradation (Calderón-Cortés et al., 2012). Hemicelluloses are non-cellulosic polysaccharides (excluding pectin and starch) forming part of the plant cell wall (Watanabe and Tokuda, 2010). The degradation of cellulose is the final step in plant cell wall degradation (Calderón-Cortés et al., 2012). Enzymes which digest cellulose are collectively termed cellulases and are divided into three classes of hydrolytic enzymes: endoglucanases, exoglucanases (cellobiohydrolases) and β-glucosidases (cellobiases) (Terra and Ferreira, 1994, Lynd et al., 2002, Gilbert, 2010, Watanabe and Tokuda, 2010).

1.4 Structure of the Insect Gut

The insect gut is a long muscular tube (Figure 1.3) extending from the insect's mouth to the anus and is divided into the fore-, mid- and hindgut (Chapman, 1998, Brune, 2013). The foregut, consisting of the buccal cavity, pharynx, oesophagus and crop, begins at the oral cavity and extends to the proventriculus, a simple valve at the entrance of the midgut (Terra and Ferreira, 1994, Chapman, 1998). In several insect species, the proventriculus is a more complex structure, assisting with filtering of food or may have circular plates which grind food at entry into the midgut (Chapman, 1998). Food may be stored in the crop (Terra and Ferreira, 1994, Chapman, 1998) before moving into the midgut. The midgut is the longest gut segment, extending across most of the body cavity (Huang et al., 2010). A characteristic feature of the midgut is the presence of rings of caecae (Huang et al., 2010). Peristaltic movements of circular and longitudinal muscles aid the movement of food through the gut (Vallet-Gely et al.,

2008). Digestions mainly takes place in the midgut where enzymes are secreted and nutrients are absorbed (Terra and Ferreira, 1994, Chapman, 1998). From the midgut, food moves into the hindgut, often via the pyloric sphincter (pylorus) which functions to restrict movement of food back into the midgut (Terra and Ferreira, 1994, Huang et al., 2010). The hindgut is made-up of the ileum and rectum (Terra and Ferreira, 1994, Chapman, 1998). The hindgut is responsible for nutrient and water absorption in some insect groups such as termites (Potrikus and Breznak, 1981, Terra and Ferreira, 1994). The hindgut of different insect species can differ, as some insects retain a straight gut, while others form large dilated compartment(s) commonly termed a fermentation chamber (Terra and Ferreira, 1994, Brune and Friedrich, 2000, Huang et al., 2010).

This fermentation chamber is where a large portion of the gut microbial community is found (Brune and Friedrich, 2000). For example, termites and scarab larvae support sac-like hindgut dilations (Brune and Friedrich, 2000, Huang et al., 2010). Certain insects such as families within Orthoptera do not have fermentation chambers but rather have other specialised hindgut features where the majority of microbes are found (Nation, 1983). Xylophagous species or those with extra-corporeal digestive symbionts generally do not have a fermentation chamber present within the hindgut (Watanabe and Tokuda, 2010).

The fore- and hindgut form from the embryonic ectoderm, resulting in the secretion of a cuticle (intima) (Chapman, 1998) to protect the gut. This lining is shed during ecdysis (i.e. the shedding of the cuticle). In contrast, the midgut is derived from endodermal cells (Chapman, 1998) and in most insect species creates a network of chitin and proteins called the peritrophic membrane (PM) which forms around the food bolus (Terra and Ferreira, 1994, Chapman, 1998, Shao et al., 2001, Terra, 2001, Wang and Granados, 2001). The PM forms a layer between the gut epithelium and the ingested food particles (Chapman, 1998). The PM also acts to concentrate dietary substrates and the necessary digestive enzymes. The formation of the PM results in the division of the midgut into the endo- and ectoperitrophic spaces (Terra and Ferreira, 1994, Chapman, 1998). The PM can be categorised into two types according to

region of the midgut in which it is formed. Type I forms across the entire midgut, typically when specific food is ingested. The Type II PM only forms in a specific area within the anterior region of the midgut, regardless of the presence of food (Lehane, 1997, Terra, 2001). Small pores throughout the PM allow enzymes and small particles to move between the endo- and ectoperitrophic spaces while confining the majority of microoganisms to the endoperitrophic space (Spence and Kawata, 1993, Barbehenn and Martin, 1995, Chapman, 1998, Edwards and Jacobs-Lorena, 2000).

Redox potential, pH, and oxygen and hydrogen concentrations may vary according to the specific gut segment (Brune and Friedrich, 2000, Brune, 2013). Furthermore, gut pH varies according to insect species (Terra and Ferreira, 1994). pH may range from acidic to alkaline varying across insect species and the respective gut segment (Terra and Ferreira, 1994).

The Malpighian tubules are the excretory organs in insects, extending from the hindgut (the pylorus is the site of attachment) into the body cavity (Terra and Ferreira, 1994, Chapman, 1998, Engel and Moran, 2013). The placement and number of Malpighian tubules may differ according to insect species as has been noted in ants (Cook and Davidson, 2006). These organs absorb waste products which are transported to the anterior hindgut, where nitrogenous and food waste accumulates (Engel and Moran, 2013).

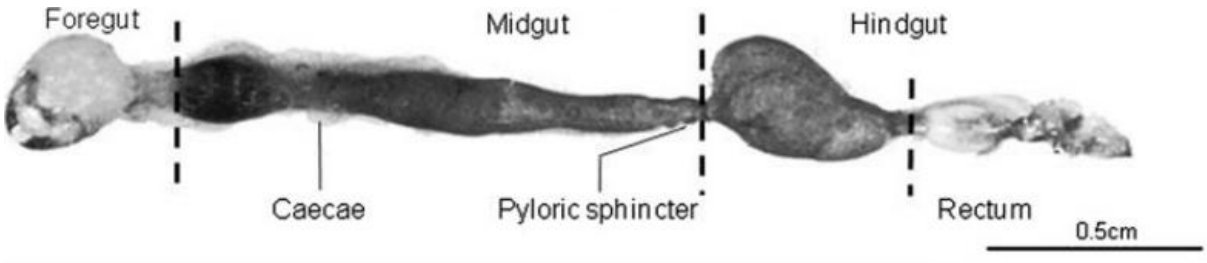


Figure 1.3: Basic structure of the larval scarab gut (adapted from Huang et al. (2010)).

1.5 Insect Gut Microbial Communities

1.5.1 Microbial Colonization of the Gut Environment

The insect gut is an extreme environment, with unstable conditions making colonisation by microbial organisms difficult. Firstly, microorganisms need to persist against harsh physiochemical conditions, digestive processes and host immune defence (Douglas, 2015). Secondly, insects undergo several events in which the gut is shed, affecting the in situ microbial communities (Broderick and Lemaitre, 2012, Engel and Moran, 2013, Douglas, 2015). In insect groups producing a PM, microorganisms are unavoidably expelled during regular shedding of this lining (Engel and Moran, 2013). Likewise, larvae moult several times during development, at which time the lining of the fore- and hindgut is shed, along with it any inhabiting microorganisms (Douglas, 2011, Broderick and Lemaitre, 2012, Engel and Moran, 2013). This results in variable gut microbiomes across the different life stages in holometabolous insects (Engel and Moran, 2013). However, moulting events stop once the insect is an adult, providing a more stable environment for colonisation of the fore- and hindgut (Engel and Moran, 2013).

Microorganisms found in the insect gut environment are typically either resident (naturally occurring), successfully inhabiting the gut environment, or transient, acquired through contact with the environment or ingested with the food (De Vries et al., 2001, Bright and Bulgheresi, 2010). While a small number of transient microorganisms may be harmful (pathogenic) to the host, the majority have little influence on the host, typically being removed from the gut (Douglas, 2009). Some ingested microorganisms are beneficial to the host, as a food-source, assisting with digestion of the ingested food and supplementing the nutritional needs of the host (Douglas, 2009). Transient microorganisms may be prevented from colonising the insect gut by the presence of the resident microflora (Veivers et al., 1982). However a small number of bacteria are capable of persisting in the gut (Vallet-Gely et al., 2008).

Resident microorganisms inherited by the host can either be obligate or facultative. Obligate or primary microorganisms are evolutionarily ancient, living in bacteriocytes (specialised host cells) and are expected to assist the host nutritionally (Dale and Moran, 2006, Moran et al., 2008, Ferrari and Vavre, 2011). However, they may not necessarily be required by the host (Moran et al., 2008). Obligate microorganisms typically have reduced genome size compared to their free-living counterparts (Dale and Moran, 2006, Nakabachi et al., 2006, McCutcheon and Moran, 2012, Bennett and Moran, 2013), retaining genes necessary for interactions with the host (Dale and Moran, 2006, Bennett and Moran, 2013). Facultative microorganisms (secondary microorganisms) are acquired through maternal horizontal transmission and are further categorised as mutualistic or parasitic (Dale and Moran, 2006, Moran et al., 2008, Bright and Bulgheresi, 2010). Facultative microorganisms either invade the tissues and cells of new insect host species (Dale and Moran, 2006) or reside in the hemocoel (Fukatsu et al., 2000). Distribution of facultative mutualists (and the respective effect on the host) is typically influenced by environmental factors, such as location and temperature (Tsuchida et al., 2002, Ferrari and Vavre, 2011)

Gut microorganisms can have both a physiological and evolutionary effect on the host (Douglas, 2011). For example, microbial composition, abundance and activity may impact the physiological welfare of the host (Douglas, 2011). As such, microorganisms may have a negative effect on the host, impacting host growth and metabolism (Broderick and Lemaitre, 2012) and influencing host behaviour (Hosokawa et al., 2008, Sharon et al., 2010, Najarro et al., 2015, Wittman and Fedorka, 2015). The host is able to control microbial (pathogenic) colonisation using two different mechanisms: resistance and tolerance (Schneider and Ayres, 2008). Resistance limits pathogen/microbial colonisation, whereas tolerance reduces the impact that pathogen/microbial colonisation would have on the host (Schneider and Ayres, 2008). Therefore, tolerance mechanisms are proportional to the population size of microorganisms inhabiting the gut of the respective host, with resistance mechanisms preferred in cases of low density populations (Engel and Moran, 2013). Different mechanisms

have been adapted by the host. For example, the host may produce enzymes (e.g. lysozymes) that damage bacterial cell walls, thereby decreasing abundance of microorganisms in either the whole or particular sections of the gut (Daffre et al., 1994, Fujita et al., 2001, Engel and Moran, 2013, Nayduch and Joyner, 2013). The resident microbiota may also assist in reducing colonization of other microorganisms in the gut (Vollaard and Clasener, 1994, Dillon and Dillon, 2004). In contrast, physical features such as the gut linings (e.g. PM) protecting the epithelium restrict access of the microorganisms to the insect itself, allowing the host to tolerate microbial colonisation of the gut (Engel and Moran, 2013).

The types of microorganisms commonly found within the insect gut include, bacteria (Ahn et al., 2012, Boissière et al., 2012, He et al., 2013, Scully et al., 2013, Dietrich et al., 2014, Gauthier et al., 2015), eukaryotes (including fungi)(Suh et al., 2004, Nguyen et al., 2007, Scully et al., 2013, Santana et al., 2015), archaea (Egert et al., 2003, Lemke et al., 2003, Santana et al., 2015) and viruses (Jia et al., 2013, Liu et al., 2013, Chandler et al., 2015).

Bacteria are typically the most abundant microorganisms in the guts of the majority of insect species (Engel and Moran, 2013). The diversity of bacterial taxa within insect guts is far less than mammals with bacterial operational taxonomic units (OTUs) within the majority of insect guts being substantially lower than in mammalian gut systems (Dethlefsen et al., 2008, Boissière et al., 2012, Andongma et al., 2015, Diouf et al., 2015). These differences are likely due to the insect gut environment being affected by moults and other perturbation events as well as the shorter life-span of insects when compared to mammals (Broderick and Lemaitre, 2012). The bacterial phyla Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria are commonly abundant in insect guts (Jones et al., 2013, Yun et al., 2014). Archaea (particularly Thermoplasmatales, Halobacteriales and methanogens) have been recorded in insects such as termites, scarab beetles and cockroaches (Egert et al., 2003, Lemke et al., 2003). Protists, mainly of the phyla Metamonada, Parabasalia and Preaxostyla, appear specific to lower termites and wood-feeding cockroaches (Hongoh, 2010, Douglas, 2015). Fungi are not the only microbial eukaryotes reported in insect guts. Anaerobic ciliates (Clevelandellida) appear

specific to cockroaches and termites (Gijzen et al., 1991, Gijzen and Barugahare, 1992, Gijzen et al., 1994), while trypanosomatids have been reported from Hemiptera, Hymenoptera, and Diptera (Chandler and James, 2013, Maslov et al., 2013).

1.5.2 Variations in Gut Microbial Communities

Variations in gut microbiome assembly are evident within an individual insect and between insects. Within an individual, gut microbial communities may differ axially across the different gut segments. The majority of microorganisms colonising the foregut are found in the crop (when present) (Köhler et al., 2012, Schauer et al., 2012). As the crop functions as a temporary storage space of food, microorganisms are expected to be removed regularly and move into the midgut (Douglas, 2015). In the midgut microbial colonisation is complicated by the secretion of digestive enzymes and a high immune response which may result in inhibited growth or death of the microbial population (Vallet-Gely et al., 2008, Douglas, 2015). Microbial communities successfully colonising the midgut are presumably able to withstand extremes of pH, ionic strength, redox potential and digestive enzymes of this gut segment (Vallet-Gely et al., 2008). Microbial colonisation of the midgut is further complicated when a PM is produced as penetration of this membrane by microorganisms is difficult, resulting in the majority of microbial symbionts being removed by peristalsis with the food (Douglas, 2015). However, insect or microbial chitinases may facilitate the movement of microorganisms across the PM (Tsai et al., 2001, Dostálová and Volf, 2012).

Microbial populations are most dense in the hindgut (particularly the ileum) where digestive enzymes and desiccation stress are negligible (Douglas, 2015). The high nutritional content (in the form of ions and metabolites in the Malpighian waste products) of the hindgut also promotes microbial growth (Douglas, 2015). Microbial community structure may also differ radially across the gut segment, i.e. between the gut wall and gut lumen (Köhler et al., 2012), Gut wall microorganisms persist in the gut, increasing their interactions with the host compared to gut lumen microbiota, which are typically removed with the food (Douglas, 2015)

Variations between insects may include large intra-specific differences between bacterial communities of insects of the same species (Ahn et al., 2012, Boissière et al., 2012, Osei-Poku et al., 2012, Montagna et al., 2015). Such variations may be driven by host genetics (Douglas, 2015) although non-genetic drivers have also been identified (Chandler et al., 2011, Osei-Poku et al., 2012, Wong et al., 2013). Intra-specific microbiome assembly may therefore not be driven only by deterministic processes but may be influenced by stochastic events including ingestion of food-based microorganisms, relationships with other microorganisms in the gut and the gut wall attachment site (Douglas, 2015).

Interactions between different microorganisms successfully colonising the gut can be either negative (antagonistic) or positive in nature. Negative interactions (competition or amensalism) influence the co-occurrence of microbial species and increase host-microbial variation, while positive (commensalism or mutualism) interactions often result in a persistent core microbiome (Wong et al., 2013). A core microbiome can be defined as “Organisms common across microbiomes hypothesized to play a key role in ecosystem function within a habitat.” (Shade and Handelsman, 2012, p. 5). Different models of core microbiomes have been defined for human microbiomes (Hamady and Knight, 2009), ranging from “substantial” (large number of OTUs shared across all individuals), “subpopulation” (OTUs only shared amongst individuals of same subpopulation), “gradient”, “minimal” and “no core”. The persistence of a core microbiome may indicate several scenarios: direct transmission of microbial communities between hosts, selectivity of microbial colonisation displayed by the host or the adaptive quality of microorganisms to successfully inhabit the gut (Engel and Moran, 2013).

1.5.3 Gut Microbial Symbionts and the Relationship between Host Diet and the Gut Microbiome

Symbiosis is any association between different species i.e. microorganisms (mutualist, commensal or parasite) that persistently associate with an insect host (Zook, 1998, Dillon and

Dillon, 2004, Douglas, 2007). It is important to note, however, that microorganisms may relate differently to different hosts and under different conditions, depending on genotype and the physiochemical conditions (i.e., a microorganism may be symbiotic in one situation and pathogenic in another) (Dillon and Charnley, 2002, Ffrench-Constant et al., 2003, Douglas, 2007, Chandler et al., 2008).

Gut microbial symbionts have been studied in an array of insect species and have been reported to assist the host in various ways including production of components of cohesion pheromones (Dillon et al., 2002), protection against parasites/pathogens and other harmful agents (Kaltenpoth et al., 2005, Kaltenpoth, 2009, Douglas, 2011), changing host body colour (Tsuchida et al., 2010), mediating thermal tolerance (Dunbar et al., 2007), contributing to fungus-garden health (Currie et al., 1999, Currie et al., 2003, Kaltenpoth, 2009) and nutrition (Douglas, 2009).

A number of insects are unable to synthesize certain essential nutritional elements including amino acids (Wilson et al., 2010, Douglas, 2013), certain vitamins linked to metabolic enzyme functioning (Douglas, 2013) and sterols (Zdobnov et al., 2002, Behmer and Nes, 2003, Janson et al., 2009). Insects rely on their food-source for acquisition of such nutrients, although some insects feed on diets deficient in these essential components (Douglas, 2013), making it necessary to find alternative methods for acquiring these nutrients.

One such alternative is the acquisition of microbial symbionts that assist the host in exploiting nutrient deficient dietary lifestyles (Akman et al., 2002, Wu et al., 2006). For example, blood-feeding insects such as mosquitoes may be deficient in B vitamins, those feeding on plant sap may lack a source of essential amino acids, and sound-wood feeders are deprived of nitrogen (Douglas, 2009, Douglas, 2011). Microbial symbionts may assist the host in degradation of plant cell walls (Park et al., 2007, Morales-Jiménez et al., 2009, Calderón-Cortés et al., 2012), nutrient, amino acid and vitamin production (Nakabachi and Ishikawa, 1999, Akman et al., 2002, Wu et al., 2006), nitrogen fixation (Nardi et al., 2002, Morales-Jiménez et al., 2009) and production/source of sterols (Janson et al., 2009). Multiple gut

microbial symbionts may collaborate to provide complete nutritional support for the host (McCutcheon and Moran, 2010).

Diet has been shown to have a pronounced effect on gut microbiome assembly. For example, pronounced differences are noted between wild-type insects feeding on natural diets and lab-reared insects fed on artificial diets (Chandler et al., 2011). Furthermore, microbiomes are altered when the major nutritional class of the food (or its concentration) is altered (i.e. the amount of protein, lipid, sugar and/or fibre is altered) (Huang et al., 2013). In the well-studied *Drosophila* model, wild fly microbiota show greater similarity between different *Drosophila* species feeding on the same substrate than closely related *Drosophila* species feeding on different diets (Chandler et al., 2011).

A core microbiome may still persist regardless of the influence of host diet (Huang et al., 2013, Schauer et al., 2014). Physiochemical gut conditions tend to favour particular microbial species and it is therefore expected that gut microbial communities should not be random even when some microbial species are acquired from the environment/diet (Engel and Moran, 2013).

1.5.4 Previous Research on Scarab Beetle Gut Microbiomes

The author is unaware of any previously published studies on the gut microbiomics of dry detritus feeding insects. Very few studies have investigated the gut microbiomes of scarab beetles, and most have concentrated on the larval life stage (Egert et al., 2003, Lemke et al., 2003, Egert et al., 2005, Zhang and Jackson, 2008, Andert et al., 2010, Huang et al., 2010, Huang et al., 2012, Huang and Zhang, 2013). Based on studies of other insect classes, it is expected that the larval gut microbiomes will be different to those of the respective adult scarab beetles (Engel and Moran, 2013).

Previous studies on scarab beetle gut microbiomes are largely restricted to the use of clone libraries and Terminal Restriction Fragment Length Polymorphism (TRFLP)/ Denaturing gradient gel electrophoresis (DGGE) analyses and not newer sequencing technologies (Egert

et al., 2003, Egert et al., 2005, Zhang and Jackson, 2008, Andert et al., 2010, Huang et al., 2012, Huang and Zhang, 2013).

The midgut microbiota of *Pachnoda ephippiata* (Andert et al., 2010), *Pachnoda marginata* (Andert et al., 2010) and *Melolontha melolontha* scarab beetle larvae are less diverse than hindgut communities (Egert et al., 2005). Apart from axial differences, radial differences in gut microbial communities were also noted in *M. melolontha* (Egert et al., 2005). Bacteria showing CMCase activity were found in the hindgut of *Holotrichia parallela* but were typically absent from the midgut (Huang et al., 2012). Huang et al. (2012) also found *Pseudomonas* to be the most abundant taxon of cellulolytic bacteria. Archaea were also detected in the guts of *M. melolontha* (Egert et al., 2005) and *P. ephippiata* (Egert et al., 2003, Lemke et al., 2003).

The effect of environmental factors on scarab gut microbiomes has also been reported with location appearing to significantly impact gut microbiome assembly of *H. parallela* (Huang and Zhang, 2013) and *Costelytra zealandica* (Zhang and Jackson, 2008). Furthermore, Andert et al. (2010) noted that variations in gut microbes between two *Pachnoda* species were greater influenced by host phylogeny than diet. In contrast, Zhang and Jackson (2008) found distinct differences between midgut bacterial communities of *C. zealandica* fed on different diets and starved. However, no such differences were noted in the hindguts of the same individuals. Significant changes in gut bacterial communities have also been noted during larval development in *H. parallela* (Huang and Zhang, 2013) and *Onthophagus taurus* (Estes et al., 2013).

1.6 Sequence-Based Methods in Microbial Ecology

Modern advances in sequencing technology have resulted in the next-generation sequencing (NGS) platforms, which are alternatively named deep or massively parallel sequencing (Metzker, 2010, Chiu and Miller, 2016). These newer technologies result in millions of reads in a single run (Mardis, 2008, Chiu and Miller, 2016). This is accomplished

through preparation of libraries consisting of Deoxyribonucleic acid (DNA) fragments which have not been subjected to cloning or culture-based methods (Mardis, 2008, Morozova and Marra, 2008). The first NGS platform made available was the Roche 454 pyrosequencing instrument (Liu et al., 2012), which was quickly followed by the Illumina (MiSeq/HiSeq/NextSeq), ABI SOLiD, Life Technologies Ion Torrent, and the PacBio RS systems. For the purpose of this discussion, only technologies and platforms used in this study will be described further.

1.6.1 Amplicon-based Pyrosequencing

Pyrosequencing is a bioluminescence sequencing technique that works on measurements of pyrophosphate released during DNA synthesis (Ronaghi, 2001). Pyrosequencing was developed as an alternative to Sanger sequencing at the Royal Institute of Technology (KTH) in Stockholm (Fakruddin et al., 2012). The Roche 454 pyrosequencer uses emulsion polymerase chain reaction (PCR) to amplify single DNA isolates for library construction (Rothberg and Leamon, 2008). Emulsification PCR requires an oil-water interface to produce droplets, whereby each droplet (microreactor) contains a bead which bonds covalently to a single DNA template with attached adaptors (Mardis, 2008, Schlebusch and Illing, 2012). Amplified fragments are produced when PCR is performed across the bead's surface (Schlebusch and Illing, 2012). The Roche 454 pyrosequencer deposits the beads into individual Picotiter plate wells (Hodkinson and Grice, 2015, Goodwin et al., 2016) where sequencing agents (DNA polymerase, Adenosine triphosphate (ATP) sulfurylase, and luciferase) are then added (Hodkinson and Grice, 2015). A single deoxyribose nucleoside triphosphates (dNTP) is added in limited concentrations to extend the primer and discontinue DNA synthesis, which is restarted by the addition of another dNTP (Metzker, 2010). This prompts pyrophosphate release during complementary strand synthesis by nucleotide incorporation (Hodkinson and Grice, 2015). Pyrophosphate release in turn produces a fluorescent signal which is recorded by a charge coupled device (CCD) camera for base calling (Metzker, 2010, Hodkinson and Grice, 2015, Chiu and Miller, 2016).

Fast sequencing time (Liu et al., 2012), longer read lengths resulting in greater identification accuracy and precision (Hodkinson and Grice, 2015), easily automated system (Ronaghi, 2001), parallel processing (Ronaghi, 2001) and overcoming bias introduced through cloning-based methods (Siqueira et al., 2012) are the general advantages of 454 pyrosequencing. Unlike earlier methods, pyrosequencing foregoes the need for gel electrophoresis and labelled reagents such as primers (Ronaghi, 2001). However, as with all sequencing technologies errors do occur, with the most notable being resulting poly-bases (homopolymers) longer than 6 base pairs (bp) (Liu et al., 2012, Hodkinson and Grice, 2015), resulting in measured insertions or deletions being introduced (Metzker, 2010, Schlebusch and Illing, 2012). Further disadvantages include the high reagent cost (Liu et al., 2012). Due to advances in newer sequencing platforms 454 life sciences are discontinuing the platform from 2016 (Hodkinson and Grice, 2015). As with all amplicon-based studies, bias may be introduced through specific primers and PCR amplification cycles (Hodkinson and Grice, 2015).

1.6.2 Environmental Shotgun Sequencing

Shotgun sequencing foregoes targeting a specific genomic locus through shearing all DNA into small fragments that are sequenced independently (Sharpton, 2014). The resultant reads are assembled into larger fragments. This technique results in both taxonomical and functional data being obtained (Eisen, 2007, Sharpton, 2014, Hodkinson and Grice, 2015, Oulas et al., 2015). As no universal marker gene is known for all domains of life, shotgun sequencing has the advantage of deciphering the entire microbial community (Hodkinson and Grice, 2015). By foregoing DNA amplification and associated biases, a clearer picture of exact relative abundances of organisms is given (Hodkinson and Grice, 2015).

Unfortunately, reference genomes are currently limited, decreasing classification efficiency (Petrosino et al., 2009, Hodkinson and Grice, 2015). Furthermore, shotgun sequencing relies on samples with large DNA concentrations (Petrosino et al., 2009, Hodkinson and Grice, 2015). High host DNA concentrations may also hinder the number of

target reads obtained (Petrosino et al., 2009, Sharpton, 2014). The cost of generating whole metagenomes is far greater than amplicon-based studies particularly when the majority of reads obtained are from the host (Sharpton, 2014).

Whole shotgun sequencing was performed using the Illumina Hiseq 2500 platform. Illumina technology uses a sequencing-by-synthesis approach performed in a flow cell (Buermans and Den Dunnen, 2014, Hodkinson and Grice, 2015). Adapter sequences are introduced during library preparation (Liu et al., 2012, Buermans and Den Dunnen, 2014, Hodkinson and Grice, 2015). During loading, double stranded DNA is denatured into single stranded DNA molecules in order to produce a copy of the original template (Liu et al., 2012, Buermans and Den Dunnen, 2014). The original template is removed and replicated (Buermans and Den Dunnen, 2014) using bridge amplification to form clusters consisting of identical fragments (Liu et al., 2012, Hodkinson and Grice, 2015). Labelled dye-terminator nucleotides are washed over and allowed to attach to the DNA fragments (Liu et al., 2012, Buermans and Den Dunnen, 2014, Hodkinson and Grice, 2015). Following imaging, the fluorescent group is removed and the terminator deactivated, allowing for the process to be repeated by the subsequent addition of nucleotides (Buermans and Den Dunnen, 2014, Hodkinson and Grice, 2015). The Illumina Hiseq 2500 platform boasts read lengths of up to 250bp with a throughput of 125-150Gb (Goodwin et al., 2016).

1.7 Aims and Objectives

1.7.1 Aims

Pachysoma species exhibit an unusual feeding behaviour, with species feeding on different diets. Furthermore, most dung beetles feed on wet dung. However, *Pachysoma* spp. utilise the dry alternative (as well as dry plant detritus), removing it from the environment. Unlike wet dung feeders, *Pachysoma* spp. appear unable to remove the plant fragments from the dung before consumption and instead feed on the plant biomass in the dung. The role gut microorganisms may play in the digestion of the two food-sources (dry dung and plant detritus)

as well as how these different diets affect gut microbiome assembly of this ecologically important insect species has not been previously studied. Therefore, this study aims to explore the relationship between the *Pachysoma* host and the associated gut microbiota.

1.7.2 Objectives

1.7.2.1 Diversity Analysis of *Pachysoma* Gut Microbiomes

The technical objectives of this study included:

- The use of amplicon sequencing to characterise the gut bacterial and fungal communities of two desert insects feeding on dry plant detritus and dung pellets.
- A comparison of intra- and inter-specific bacterial and fungal gut communities, to determine any potential influence of host diet and/or host phylogeny.
- The use of shotgun sequencing to determine the composition of the entire gut microbiome of both *Pachysoma* species.

1.7.2.2 Functional Analysis of *Pachysoma* Gut Microbiomes

- The use of shotgun sequencing of the whole *Pachysoma* gut microbial metagenome to determine the functional capacity of the gut community.
- Determination of correlations between the functional genes of the microbial gut communities and the different diets (dung versus plant detritus) of the two *Pachysoma* species.

1.8 References

Abe, T. & Higashi, M., 1991. Cellulose centered perspective on terrestrial community structure. *Oikos,* 60(1), pp. 127-133.

Ahn, J.-H., Hong, I.-P., Bok, J.-I., Kim, B.-Y., Song, J. & Weon, H.-Y., 2012. Pyrosequencing analysis of the bacterial communities in the guts of honey bees *Apis cerana* and *Apis mellifera* in Korea. *Journal of Microbiology,* 50(5), pp. 735-745.

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M. & Aksoy, S., 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nature Genetics,* 32(3), pp. 402-407.

Al-Houty, W., 1997. Checklist of the insect fauna of Kuwait. *Kuwait Journal of Science and Engineering,* 24(1), pp. 145-162.

Al-Houty, W., 2011. Recent insect fauna recorded in Kuwait. *Kuwait Journal of Science and Engineering,* 38(1A), pp. 81-92.

Allaby, M., 2006. *Deserts*. Chelsea House, USA, New York.

Anderson, J.M. & Coe, M.J., 1974. Decomposition of elephant dung in an arid, tropical environment. *Oecologia,* 14(1-2), pp. 111-125.

Andert, J., Marten, A., Brandl, R. & Brune, A., 2010. Inter-and intraspecific comparison of the bacterial assemblages in the hindgut of humivorous scarab beetle larvae (*Pachnoda* spp.). *FEMS Microbiology Ecology,* 74(2), pp. 439-449.

Andongma, A.A., Wan, L., Dong, Y.-C., Desneux, N., White, J.A. & Niu, C.-Y., 2015. Pyrosequencing reveals a shift in symbiotic bacteria populations across life stages of *Bactrocera dorsalis*. *Scientific Reports,* 5.

Arman, P. & Hopcraft, D., 1975. Nutritional studies on East African herbivores. *British Journal of Nutrition,* 33(2), pp. 255-264.

Barbehenn, R.V. & Martin, M.M., 1995. Peritrophic envelope permeability in herbivorous insects. *Journal of Insect Physiology,* 41(4), pp. 303-311.

Barbehenn, R.V., Reese, J.C. & Hagen, K.S., 1999. *The food of insects*. John Wiley & Sons, USA, New York.

Behmer, S.T. & Nes, W.D., 2003. Insect sterol nutrition and physiology: a global overview. *Advances in Insect Physiology,* 31, pp. 1-72.

Benjaminsen, T.A., Rohde, R., Sjaastad, E., Wisborg, P. & Lebert, T., 2006. Land reform, range ecology, and carrying capacities in Namaqualand, South Africa. *Annals of the Association of American Geographers,* 96(3), pp. 524-540.

Benner, R. & Hodson, R.E., 1985. Microbial degradation of the leachable and lignocellulosic components of leaves and wood from *Rhizophora mangle* in a tropical mangrove swamp. *Marine Ecology Progress Series,* 23, pp. 221-230.

Bennett, G.M. & Moran, N.A., 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a phloem-feeding insect. *Genome Biology and Evolution,* 5(9), pp. 1675-1688.

Boissière, A., Tchioffo, M.T., Bachar, D., Abate, L., Marie, A., Nsango, S.E., et al., 2012. Midgut microbiota of the malaria mosquito vector *Anopheles gambiae* and interactions with *Plasmodium falciparum* infection. *PLoS Pathogens,* 8(5), pp. e1002742.

Bouchard, P., Bousquet, Y., Davies, A.E., Alonso-Zarazaga, M.A., Lawrence, J.F., Lyal, C.H.C., et al., 2011. Family-group names in Coleoptera (Insecta). *ZooKeys,* 88, pp. 1-972.

Brauman, A., 2000. Effect of gut transit and mound deposit on soil organic matter transformations in the soil feeding termite: a review. *European Journal of Soil Biology,* 36(3), pp. 117-125.

Bright, M. & Bulgheresi, S., 2010. A complex journey: transmission of microbial symbionts. *Nature Reviews Microbiology,* 8(3), pp. 218-230.

Broderick, N.A. & Lemaitre, B., 2012. Gut-associated microbes of *Drosophila melanogaster*. *Gut Microbes,* 3(4), pp. 307-321.

Brune, A., 2013. Symbiotic associations between termites and prokaryotes. *In:* Rosenberg, E., Delong, E.F., Lory, S., Stackebrandt, E. & Thompson, F. (eds.) *The prokaryotes.* pp. 545-577. Springer Berlin, Heidelberg.

Brune, A. & Friedrich, M., 2000. Microecology of the termite gut: structure and function on a microscale. *Current Opinion in Microbiology,* 3(3), pp. 263-269.

Buermans, H.P.J. & Den Dunnen, J.T., 2014. Next generation sequencing technology: advances and applications. *Biochimica et Biophysica Acta* 1842(10), pp. 1932-1941.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H. & Oyama, K., 2012. Endogenous plant cell wall digestion: a key mechanism in insect evolution. *Annual Review of Ecology, Evolution, and Systematics,* 43, pp. 45-71.

Cambefort, Y., 1991. From saprophagy to coprophagy. *In:* Hanski, I. & Cambefort, Y. (eds.) *Dung beetle ecology.* pp. 22-35. Princeton University Press, USA, New Jersey.

Chandler, J.A. & James, P.M., 2013. Discovery of trypanosomatid parasites in globally distributed *Drosophila* species. *PLoS One,* 8(4), pp. e61937.

Chandler, J.A., Liu, R.M. & Bennett, S.N., 2015. RNA shotgun metagenomic sequencing of northern California (USA) mosquitoes uncovers viruses, bacteria, and fungi. *Frontiers in Microbiology,* 6.

Chandler, J.A., Lang, J.M., Bhatnagar, S., Eisen, J.A. & Kopp, A., 2011. Bacterial communities of diverse *Drosophila* species: ecological context of a host–microbe model system. *PLoS Genetics,* 7(9), pp. e1002272.

Chandler, S.M., Wilkinson, T.L. & Douglas, A.E., 2008. Impact of plant nutrients on the relationship between a herbivorous insect and its symbiotic bacteria. *Proceedings of the Royal Society of London B: Biological Sciences,* 275(1634), pp. 565-570.

Chapman, R.F., 1998. *The insects: structure and function*. Cambridge University Press, UK, Cambridge.

Chiu, C. & Miller, S., 2016. *Next-generation sequencing*. ASM Press, USA, Washington.

Cook, S.C. & Davidson, D.W., 2006. Nutritional and functional biology of exudate-feeding ants. *Entomologia Experimentalis et Applicata,* 118(1), pp. 1-10.

Cowan, D.A. & Tow, L.A., 2004. Endangered antarctic environments. *Annual Review of Microbiology,* 58, pp. 649-690.

Cowling, R.M., Esler, K.J. & Rundel, P.W., 1999. Namaqualand, South Africa–an overview of a unique winter-rainfall desert ecosystem. *Plant Ecology,* 142(1-2), pp. 3-21.

Currie, C.R., Bot, A.N.M. & Boomsma, J.J., 2003. Experimental evidence of a tripartite mutualism: bacteria protect ant fungus gardens from specialized parasites. *Oikos,* 101(1), pp. 91-102.

Currie, C.R., Scott, J.A., Summerbell, R.C. & Malloch, D., 1999. Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. *Nature,* 398(6729), pp. 701-704.

Daffre, S., Kylsten, P., Samakovlis, C. & Hultmark, D., 1994. The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted for expression in the digestive tract. *Molecular and General Genetics* 242(2), pp. 152-162.

Dale, C. & Moran, N.A., 2006. Molecular interactions between bacterial symbionts and their hosts. *Cell,* 126(3), pp. 453-465.

De Vries, E.J., Jacobs, G. & Breeuwer, J.A.J., 2001. Growth and transmission of gut bacteria in the western flower thrips, *Frankliniella occidentalis*. *Journal of Invertebrate Pathology,* 77(2), pp. 129-137.

Delve, R.J., Cadisch, G., Tanner, J.C., Thorpe, W., Thorne, P.J. & Giller, K.E., 2001. Implications of livestock feeding management on soil fertility in the smallholder farming systems of sub-Saharan Africa. *Agriculture, Ecosystems and Environment,* 84(3), pp. 227-243.

Derrien, M., Jardé, E., Gruau, G. & Pierson-Wickmann, A.-C., 2011. Extreme variability of steroid profiles in cow feces and pig slurries at the regional scale: implications for the use of steroids to specify fecal pollution sources in waters. *Journal of Agricultural and Food Chemistry,* 59(13), pp. 7294-7302.

Desmet, P.G., 2007. Namaqualand—a brief overview of the physical and floristic environment. *Journal of Arid Environments,* 70(4), pp. 570-587.

Desmet, P.G. & Cowling, R.M., 1999. Biodiversity, habitat and range-size aspects of a flora from a winter-rainfall desert in north-western Namaqualand, South Africa. *Plant Ecology,* 142, pp. 23-33.

Dethlefsen, L., Huse, S., Sogin, M.L. & Relman, D.A., 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biology,* 6(11), pp. e280.

Dietrich, C., Köhler, T. & Brune, A., 2014. The cockroach origin of the termite gut microbiota: patterns in bacterial community structure reflect major evolutionary events. *Applied and Environmental Microbiology,* 80(7), pp. 2261-2269.

Dillon, R.J. & Charnley, K., 2002. Mutualism between the desert locust *Schistocerca gregaria* and its gut microbiota. *Research in Microbiology,* 153(8), pp. 503-509.

Dillon, R.J. & Dillon, V.M., 2004. The gut bacteria of insects: nonpathogenic interactions. *Annual Reviews in Entomology,* 49(1), pp. 71-92.

Dillon, R.J., Vennard, C.T. & Charnley, A.K., 2002. A note: gut bacteria produce components of a locust cohesion pheromone. *Journal of Applied Microbiology,* 92(4), pp. 759-763.

Diouf, M., Roy, V., Mora, P., Frechault, S., Lefebvre, T., Hervé, V., et al., 2015. Profiling the succession of bacterial communities throughout the life stages of a higher termite *Nasutitermes arborum* (Termitidae, Nasutitermitinae) using 16S rRNA gene pyrosequencing. *PLoS One,* 10(10), pp. e0140014.

Doblin, M.S., Pettolino, F. & Bacic, A., 2010. Evans Review: plant cell walls: the skeleton of the plant world. *Functional Plant Biology,* 37(5), pp. 357-381.

Dostálová, A. & Volf, P., 2012. *Leishmania* development in sand flies: parasite-vector interactions overview. *Parasite and Vectors,* 5, pp. 276.

Doube, B.M., 1990. A functional classification for analysis of the structure of dung beetle assemblages. *Ecological Entomology,* 15(4), pp. 371-383.

Douglas, A.E., 2007. Symbiotic microorganisms: untapped resources for insect pest control. *Trends in Biotechnology,* 25(8), pp. 338-342.

Douglas, A.E., 2009. The microbial dimension in insect nutritional ecology. *Functional Ecology,* 23(1), pp. 38-47.

Douglas, A.E., 2011. Lessons from studying insect symbioses. *Cell Host & Microbe,* 10(4), pp. 359-367.

Douglas, A.E., 2013. Microbial brokers of insect-plant interactions revisited. *Journal of Chemical Ecology,* 39(7), pp. 952-961.

Douglas, A.E., 2015. Multiorganismal insects: diversity and function of resident microorganisms. *Annual Review of Entomology,* 60, pp. 17-34.

Dunbar, H.E., Wilson, A.C.C., Ferguson, N.R. & Moran, N.A., 2007. Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biology,* 5(5), pp. e96.

Edelhoch, H., 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry,* 6(7), pp. 1948-1954.

Edwards, M.J. & Jacobs-Lorena, M., 2000. Permeability and disruption of the peritrophic matrix and caecal membrane from *Aedes aegypti* and *Anopheles gambiae* mosquito larvae. *Journal of Insect Physiology,* 46(9), pp. 1313-1320.

Edwards, P.B., 1991. Seasonal variation in the dung of African grazing mammals, and its consequences for coprophagous insects. *Functional Ecology,* 5(5), pp. 617-628.

Egert, M., Wagner, B., Lemke, T., Brune, A. & Friedrich, M.W., 2003. Microbial community structure in midgut and hindgut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology,* 69(11), pp. 6659-6668.

Egert, M., Stingl, U., Bruun, L.D., Pommerenke, B., Brune, A. & Friedrich, M.W., 2005. Structure and topology of microbial communities in the major gut compartments of *Melolontha melolontha* larvae (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology,* 71(8), pp. 4556-4566.

Eisen, J.A., 2007. Environmental shotgun sequencing: its potential and challenges for studying the hidden world of microbes. *PLoS Biology,* 5(3), pp. e82.

Engel, P. & Moran, N.A., 2013. The gut microbiota of insects–diversity in structure and function. *FEMS Microbiology Reviews,* 37(5), pp. 699-735.

Estes, A.M., Hearn, D.J., Snell-Rood, E.C., Feindler, M., Feeser, K., Abebe, T., et al., 2013. Brood ball-mediated transmission of microbiome members in the dung beetle, *Onthophagus taurus* (Coleoptera: Scarabaeidae). *PLoS One,* 8(11), pp. e79061.

Fakruddin, M., Chowdhury, A., Hossain, M.N., Mannan, K.S. & Mazumda, R.M., 2012. Pyrosequencing-principles and applications. *International Journal of Life Science and Pharma Research,* 2(2), pp. 65-76.

Ferrari, J. & Vavre, F., 2011. Bacterial symbionts in insects or the story of communities affecting communities. *Philosophical Transactions of the Royal Society of London B: Biological Sciences,* 366(1569), pp. 1389-1400.

Ferreira, M., 1953. Monografia dos escarabaeieos da Afirca do Sul. Tribo-Scarabaeini. I Parte Sub-tribo Pachysomides. *Boletim da Sociedade de Estudos da Provincia de Mocambique,* 23, pp. 1-85.

Ffrench-Constant, R., Waterfield, N., Daborn, P., Joyce, S., Bennett, H., Au, C., et al., 2003. *Photorhabdus*: towards a functional genomic analysis of a symbiont and pathogen. *FEMS Microbiology Reviews,* 26(5), pp. 433-456.

Forbes, G.S., Van Zee, J.W., Smith, W. & Whitford, W.G., 2005. Desert grassland canopy arthropod species richness: temporal patterns and effects of intense, short-duration livestock grazing. *Journal of Arid Environments,* 60(4), pp. 627-646.

Forgie, S.A., Philips, T.K. & Scholtz, C.H., 2005. Evolution of the Scarabaeini (Scarabaeidae: Scarabaeinae). *Systematic Entomology,* 30(1), pp. 60-96.

Forgie, S.A., Kryger, U., Bloomer, P. & Scholtz, C.H., 2006. Evolutionary relationships among the Scarabaeini (Coleoptera: Scarabaeidae) based on combined molecular and morphological data. *Molecular Phylogenetics and Evolution,* 40(3), pp. 662-678.

Fujita, A., Shimizu, I. & Abe, T., 2001. Distribution of lysozyme and protease, and amino acid concentration in the guts of a wood-feeding termite, *Reticulitermes speratus* (Kolbe): possible digestion of symbiont bacteria transferred by trophallaxis. *Physiological Entomology,* 26(2), pp. 116-123.

Fukatsu, T., Nikoh, N., Kawai, R. & Koga, R., 2000. The secondary endosymbiotic bacterium of the pea aphid *Acyrthosiphon pisum* (Insecta: Homoptera). *Applied and Environmental Microbiology,* 66(7), pp. 2748-2758.

Gauthier, J.-P., Outreman, Y., Mieuzet, L. & Simon, J.-C., 2015. Bacterial communities associated with host-adapted populations of pea aphids revealed by deep sequencing of 16S ribosomal DNA. *PLoS One,* 10(3), pp. e0120664.

Geib, S.M., Filley, T.R., Hatcher, P.G., Hoover, K., Carlson, J.E., del Mar Jimenez-Gasco, M., et al., 2008. Lignin degradation in wood-feeding insects. *Proceedings of the National Academy of Sciences,* 105(35), pp. 12932-12937.

Gijzen, H.J. & Barugahare, M., 1992. Contribution of anaerobic protozoa and methanogens to hindgut metabolic activities of the American cockroach, *Periplaneta americana*. *Applied and Environmental Microbiology,* 58(8), pp. 2565-2570.

Gijzen, H.J., Broers, C.A., Barughare, M. & Stumm, C.K., 1991. Methanogenic bacteria as endosymbionts of the ciliate *Nyctotherus ovalis* in the cockroach hindgut. *Applied and Environmental Microbiology,* 57(6), pp. 1630-1634.

Gijzen, H.J., van der Drift, C., Barugahare, M. & op den Camp, H.J.M., 1994. Effect of host diet and hindgut microbial composition on cellulolytic activity in the hindgut of the American cockroach, *Periplaneta americana*. *Applied and Environmental Microbiology,* 60(6), pp. 1822-1826.

Gilbert, H.J., 2010. The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiology,* 153(2), pp. 444-455.

Goodwin, S., McPherson, J.D. & McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics,* 17(6), pp. 333-351.

Goya, J.F., Frangi, J.L., Pérez, C.A. & Dalla Tea, F., 2008. Decomposition and nutrient release from leaf litter in *Eucalyptus grandis* plantations on three different soils in Entre Ríos, Argentina. *Bosque,* 29(3), pp. 217-226.

Greenham, P.M., 1972. The effects of the variability of cattle dung on the multiplication of the bushfly (*Musca vetustissima* Walk). *The Journal of Animal Ecology,* 41(1), pp. 153-165.

Guillén, F., Martínez, M.J., Gutiérrez, A. & Del Rio, J.C., 2005. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. *International Microbiology,* 8, pp. 195-204.

Hamady, M. & Knight, R., 2009. Microbial community profiling for human microbiome projects: tools, techniques, and challenges. *Genome Research,* 19(7), pp. 1141-1152.

Harris, N., 2003. *Atlas of the world's deserts*. Taylor & Francis, New York.

Harrison, J.d.G. & Philips, T.K., 2003. Phylogeny of Scarabaeus (*Pachysoma* MacLeay) stat. nov., and related flightless Scarabaeini (Scarabaeidae: Scarabaeinae). *Annals of the Transvaal Museum,* 40, pp. 47-71.

Harrison, J.d.G., Scholtz, C.H. & Chown, S.L., 2003. A revision of the endemic south-western African dung beetle subgenus *Scarabaeus* (*Pachysoma*) MacLeay, including notes on other flightless Scarabaeini (Scarabaeidae: Scarabaeinae). *Journal of Natural History,* 37(3), pp. 305-355.

Hättenschwiler, S., Aeschlimann, B., Coûteaux, M.M., Roy, J. & Bonal, D., 2008. High variation in foliage and leaf litter chemistry among 45 tree species of a neotropical rainforest community. *New Phytologist,* 179(1), pp. 165-175.

He, S., Ivanova, N., Kirton, E., Allgaier, M., Bergin, C., Scheffrahn, R.H., et al., 2013. Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS One,* 8(4), pp. e61126.

Hodkinson, B.P. & Grice, E.A., 2015. Next-generation sequencing: a review of technologies and tools for wound microbiome research. *Advances in Wound Care,* 4(1), pp. 50-58.

Holter, P., 2016. Herbivore dung as food for dung beetles: elementary coprology for entomologists. *Ecological Entomology,* 41(4), pp. 367-377.

Holter, P. & Scholtz, C.H., 2011. Re-establishment of biting mouthparts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on plant litter—old structures reacquired or new ones evolved? *Journal of Morphology,* 272(8), pp. 1007-1016.

Holter, P. & Scholtz, C.H., 2013. Elongated hindguts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on dry dung pellets or plant litter. *Journal of Morphology,* 274(6), pp. 657-662.

Holter, P., Scholtz, C.H. & Wardhaugh, K.G., 2002. Dung feeding in adult scarabaeines (tunnellers and endocoprids): even large dung beetles eat small particles. *Ecological Entomology,* 27(2), pp. 169-176.

Holter, P., Scholtz, C.H. & Stenseng, L., 2009. Desert detritivory: nutritional ecology of a dung beetle (*Pachysoma glentoni*) subsisting on plant litter in arid South African sand dunes. *Journal of Arid Environments,* 73(12), pp. 1090-1094.

Hongoh, Y., 2010. Diversity and genomes of uncultured microbial symbionts in the termite gut. *Bioscience, Biotechnology, and Biochemistry,* 74(6), pp. 1145-1151.

Hosokawa, T., Kikuchi, Y., Shimada, M. & Fukatsu, T., 2008. Symbiont acquisition alters behaviour of stinkbug nymphs. *Biology Letters,* 4(1), pp. 45-48.

Huang, S. & Zhang, H., 2013. The impact of environmental heterogeneity and life stage on the hindgut microbiota of *Holotrichia parallela* larvae (Coleoptera: Scarabaeidae). *PLoS One,* 8(2), pp. e57169.

Huang, S., Sheng, P. & Zhang, H., 2012. Isolation and identification of cellulolytic bacteria from the gut of *Holotrichia parallela* larvae (Coleoptera: Scarabaeidae). *International Journal of Molecular Sciences,* 13(3), pp. 2563-2577.

Huang, S., Zhang, H.Y., Marshall, S. & Jackson, T.A., 2010. The scarab gut: a potential bioreactor for bio-fuel production. *Insect Science,* 17(3), pp. 175-183.

Huang, X.-F., Bakker, M.G., Judd, T.M., Reardon, K.F. & Vivanco, J.M., 2013. Variations in diversity and richness of gut bacterial communities of termites (*Reticulitermes flavipes*) fed with grassy and woody plant substrates. *Microbial Ecology,* 65(3), pp. 531-536.

Janecke, B.B. & Smit, G.N., 2015. Faecal nitrogen of browser and mixed feeder game species during different seasons. *African Journal of Range & Forage Science,* 32(3), pp. 203-212.

Janson, E.M., Grebenok, R.J., Behmer, S.T. & Abbot, P., 2009. Same host-plant, different sterols: variation in sterol metabolism in an insect herbivore community. *Journal of Chemical Ecology,* 35(11), pp. 1309-1319.

Jia, S., Zhang, X., Zhang, G., Yin, A., Zhang, S., Li, F., et al., 2013. Seasonally variable intestinal metagenomes of the red palm weevil (*Rhynchophorus ferrugineus*). *Environmental Microbiology,* 15(11), pp. 3020-3029.

Jones, R.T., Sanchez, L.G. & Fierer, N., 2013. A cross-taxon analysis of insect-associated bacterial diversity. *PLoS One,* 8(4), pp. e61218.

Kaltenpoth, M., 2009. Actinobacteria as mutualists: general healthcare for insects? *Trends in Microbiology,* 17(12), pp. 529-535.

Kaltenpoth, M., Göttler, W., Herzner, G. & Strohm, E., 2005. Symbiotic bacteria protect wasp larvae from fungal infestation. *Current Biology,* 15(5), pp. 475-479.

Karasov, W.H., Martínez del Rio, C. & Caviedes-Vidal, E., 2011. Ecological physiology of diet and digestive systems. *Annual Review of Physiology,* 73, pp. 69-93.

Kazmi, S.I. & Ramamurthy, V., 2004. Coleoptera (Insect) fauna from the Indian Thar Desert, Rajasthan. *Zoos' Print Journal,* 19(4), pp. 1447-1448.

Kemp, P.R., Reynolds, J.F., Virginia, R.A. & Whitford, W.G., 2003. Decomposition of leaf and root litter of Chihuahuan desert shrubs: effects of three years of summer drought. *Journal of Arid Environments,* 53(1), pp. 21-39.

Kögel-Knabner, I., 2002. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biology and Biochemistry,* 34(2), pp. 139-162.

Köhler, T., Dietrich, C., Scheffrahn, R.H. & Brune, A., 2012. High-resolution analysis of gut environment and bacterial microbiota reveals functional compartmentation of the gut in wood-feeding higher termites (*Nasutitermes* spp.). *Applied and Environmental Microbiology,* 78(13), pp. 4691-4701.

Laity, J.J., 2009. *Deserts and desert environments*. Wiley-Blackwell, UK, Chichester.

Leeming, R., Ball, A., Ashbolt, N. & Nichols, P., 1996. Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. *Water Research,* 30(12), pp. 2893-2900.

Lehane, M.J., 1997. Peritrophic matrix structure and function. *Annual Review of Entomology,* 42(1), pp. 525-550.

Lemke, T., Stingl, U., Egert, M., Friedrich, M.W. & Brune, A., 2003. Physicochemical conditions and microbial activities in the highly alkaline gut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology,* 69(11), pp. 6650-6658.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., et al., 2012. Comparison of next-generation sequencing systems. *BioMed Research International,* 2012, pp. 1-11.

Liu, N., Zhang, L., Zhou, H., Zhang, M., Yan, X., Wang, Q., et al., 2013. Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (*Odontotermes yunnanensis*). *PLoS One,* 8(7), pp. e69184.

Lodha, B.C., 1974. Decomposition of digested litter. *In:* Dickinson, C.H. & Pugh, G.J.F. (eds.) *Biology of plant litter decomposition.* pp. 213-241. Academic Press, UK, London.

Lynd, L.R., Weimer, P.J., Van Zyl, W.H. & Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology and Molecular Biology Reviews,* 66(3), pp. 506-577.

MacLeay, W.S., 1821. *Horae entomologicae: or essays on the annulose animals*. S. Bagster, UK, London.

Makhalanyane, T.P., Valverde, A., Gunnigle, E., Frossard, A., Ramond, J.-B. & Cowan, D.A., 2015. Microbial ecology of hot desert edaphic systems. *FEMS Microbiology Reviews,* 39(2), pp. 203-221.

Mardis, E.R., 2008. The impact of next-generation sequencing technology on genetics. *Trends in Genetics,* 24(3), pp. 133-141.

Marsh, A.C., 1986. Checklist, biological notes and distribution of ants in the central Namib Desert. *Madoqua,* 14(4), pp. 333-344.

Maslov, D.A., Votýpka, J., Yurchenko, V. & Lukeš, J., 2013. Diversity and phylogeny of insect trypanosomatids: all that is hidden shall be revealed. *Trends in Parasitology,* 29(1), pp. 43-52.

Mason, V.C., Cook, J.E., Cooper, E.M., Hoadley, C.J. & Cockburn, J.E., 1989. Oven and stack ammoniation of grass hays. 2. Voluntary intake, digestibility and faecal nitrogen excretion in sheep in relation to hay composition. *Animal Feed Science and Technology,* 24(3-4), pp. 313-326.

May, H. & Lahiff, E., 2007. Land reform in Namaqualand, 1994–2005: a review. *Journal of Arid Environments,* 70(4), pp. 782-798.

McCutcheon, J.P. & Moran, N.A., 2010. Functional convergence in reduced genomes of bacterial symbionts spanning 200 My of evolution. *Genome Biology and Evolution,* 2, pp. 708-718.

McCutcheon, J.P. & Moran, N.A., 2012. Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology,* 10(1), pp. 13-26.

McDowell, R.W. & Stewart, I., 2005. Phosphorus in fresh and dry dung of grazing dairy cattle, deer, and sheep. *Journal of Environmental Quality,* 34(2), pp. 598-607.

Metzker, M.L., 2010. Sequencing technologies—the next generation. *Nature Reviews Genetics,* 11(1), pp. 31-46.

Mlambo, S., Sole, C.L. & Scholtz, C.H., 2015. A molecular phylogeny of the African Scarabaeinae (Coleoptera: Scarabaeidae). *Arthropod Systematics & Phylogeny,* 73(2), pp. 303-321.

Montagna, M., Chouaia, B., Mazza, G., Prosdocimi, E.M., Crotti, E., Mereghetti, V., et al., 2015. Effects of the diet on the microbiota of the red palm weevil (Coleoptera: Dryophthoridae). *PLoS One,* 10(1), pp. e0117439.

Morales-Jiménez, J., Zúñiga, G., Villa-Tanaca, L. & Hernández-Rodríguez, C., 2009. Bacterial community and nitrogen fixation in the red turpentine beetle, *Dendroctonus valens* LeConte (Coleoptera: Curculionidae: Scolytinae). *Microbial Ecology,* 58(4), pp. 879-891.

Moran, N.A., McCutcheon, J.P. & Nakabachi, A., 2008. Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics,* 42, pp. 165-190.

Morozova, O. & Marra, M.A., 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics,* 92(5), pp. 255-264.

Mostert, L.E. & Holm, E., 1982. Notes on the flightless Scarabaeina (Coleoptera, Scarabaeidae) with a description of a new species. *Cimbebasia,* 5, pp. 273-284.

Müller, Z.O., 1980. *Feed from animal wastes:state of knowledge*. Food and Agriculture Organization of the United Nation, Italy, Rome.

Myers, N., Mittermeier, R.A., Mittermeier, C.G., Da Fonseca, G.A.B. & Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature,* 403(6772), pp. 853-858.

Najarro, M.A., Sumethasorn, M., Lamoureux, A. & Turner, T.L., 2015. Choosing mates based on the diet of your ancestors: replication of non-genetic assortative mating in *Drosophila melanogaster*. *PeerJ,* 3, pp. e1173.

Nakabachi, A. & Ishikawa, H., 1999. Provision of riboflavin to the host aphid, *Acyrthosiphon pisum*, by endosymbiotic bacteria, *Buchnera*. *Journal of Insect Physiology,* 45(1), pp. 1-6.

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H.E., Moran, N.A. & Hattori, M., 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science,* 314(5797), pp. 267.

Nalepa, C.A., Bignell, D.E. & Bandi, C., 2001. Detritivory, coprophagy, and the evolution of digestive mutualisms in Dictyoptera. *Insectes Sociaux,* 48(3), pp. 194-201.

Nardi, J.B., Mackie, R.I. & Dawson, J.O., 2002. Could microbial symbionts of arthropod guts contribute significantly to nitrogen fixation in terrestrial ecosystems? *Journal of Insect Physiology,* 48(8), pp. 751-763.

Nation, J.L., 1983. Specialization in the alimentary canal of some mole crickets (Orthoptera: Gryllotalpidae). *International Journal of Insect Morphology and Embryology,* 12(4), pp. 201-210.

Nation, J.L., 2008. *Insect physiology and biochemistry*. CRC Press, USA.

Nayduch, D. & Joyner, C., 2013. Expression of lysozyme in the life history of the house Fly (*Musca domestica* L.). *Journal of Medical Entomology,* 50(4), pp. 847-852.

Nguyen, N.H., Suh, S.-O. & Blackwell, M., 2007. Five novel *Candida* species in insect-associated yeast clades isolated from Neuroptera and other insects. *Mycologia,* 99(6), pp. 842-858.

O'Farrell, P.J., De Lange, W.J., Le Maitre, D.C., Reyers, B., Blignaut, J.N., Milton, S.J., et al., 2011. The possibilities and pitfalls presented by a pragmatic approach to ecosystem service valuation in an arid biodiversity hotspot. *Journal of Arid Environments,* 75(6), pp. 612-623.

Ohkuma, M., 2003. Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Applied Microbiology and Biotechnology,* 61(1), pp. 1-9.

Olson, D.M. & Dinerstein, E., 2002. The Global 200: Priority ecoregions for global conservation. *Annals of the Missouri Botanical garden,* 89(2), pp. 199-224.

Osei-Poku, J., Mbogo, C.M., Palmer, W.J. & Jiggins, F.M., 2012. Deep sequencing reveals extensive variation in the gut microbiota of wild mosquitoes from Kenya. *Molecular Ecology,* 21(20), pp. 5138-5150.

Oulas, A., Pavloudi, C., Polymenakou, P., Pavlopoulos, G.A., Papanikolaou, N., Kotoulas, G., et al., 2015. Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinformatics and Biology Insights,* 9, pp. 75-88.

Park, D., Oh, H., Jeong, W., Kim, H., Park, H. & Bae, K.S., 2007. A culture-based study of the bacterial communities within the guts of nine longicorn beetle species and their exo-enzyme producing properties for degrading xylan and pectin. *Journal of Microbiology-Seoul,* 45(5), pp. 394-401.

Peel, M.C., Finlayson, B.L. & McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* 4(2), pp. 439-473.

Pérez, J., Munoz-Dorado, J., de la Rubia, T. & Martinez, J., 2002. Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. *International Microbiology,* 5(2), pp. 53-63.

Petrosino, J.F., Highlander, S., Luna, R.A., Gibbs, R.A. & Versalovic, J., 2009. Metagenomic pyrosequencing and microbial identification. *Clinical Chemistry,* 55(5), pp. 856-866.

Piñero, F.S., Tinaut, A., Aguirre-Segura, A., Miñano, J., Lencina, J.L., Ortiz-Sánchez, F.J. & Pérez-López, F.J., 2011. Terrestrial arthropod fauna of arid areas of SE Spain: diversity, biogeography, and conservation. *Journal of Arid Environments,* 75(12), pp. 1321-1332.

Potrikus, C.J. & Breznak, J.A., 1981. Gut bacteria recycle uric acid nitrogen in termites: a strategy for nutrient conservation. *Proceedings of the National Academy of Sciences,* 78(7), pp. 4601-4605.

Ronaghi, M., 2001. Pyrosequencing sheds light on DNA sequencing. *Genome Research,* 11(1), pp. 3-11.

Roskov, Y., Abucay, L., Orrell, T., Nicolson, D., Bailly, N., Kirk, P., et al. 2016. *Species 2000 & ITIS Catalogue of Life, 31st October 2016* [Online]. the Netherlands: Species 2000: Naturalis, Leiden. Available: Digital resource at www.catalogueoflife.org/annual-checklist/2016 [Accessed 2016].

Rothberg, J.M. & Leamon, J.H., 2008. The development and impact of 454 sequencing. *Nature Biotechnology,* 26(10), pp. 1117-1124.

Rougon, D., Rougon, C., Levieux, J. & Trichet, J., 1990. Variations in the amino-acid content in zebu dung in the Sahel during nesting by dung-beetles (Coleoptera, Scarabaeidae). *Soil Biology and Biochemistry,* 22(2), pp. 217-223.

Santana, R.H., Catão, E.C.P., Lopes, F.A.C., Constantino, R., Barreto, C.C. & Krüger, R.H., 2015. The gut microbiota of workers of the litter-feeding termite *Syntermes wheeleri* (Termitidae: Syntermitinae): archaeal, bacterial, and fungal communities. *Microbial Ecology,* 70, pp. 545-556.

Schauer, C., Thompson, C.L. & Brune, A., 2012. The bacterial community in the gut of the cockroach *Shelfordella lateralis* reflects the close evolutionary relatedness of cockroaches and termites. *Applied and Environmental Microbiology,* 78(8), pp. 2758-2767.

Schauer, C., Thompson, C.L. & Brune, A., 2014. Pyrotag sequencing of the gut microbiota of the cockroach *Shelfordella lateralis* reveals a highly dynamic core but only limited effects of diet on community structure. *PLoS One,* 9(1), pp. e85861.

Schlebusch, S. & Illing, N., 2012. Next generation shotgun sequencing and the challenges of de novo genome assembly. *South African Journal of Science,* 108(11-12), pp. 62-70.

Schneider, D.S. & Ayres, J.S., 2008. Two ways to survive infection: what resistance and tolerance can teach us about treating infectious diseases. *Nature Reviews Immunology,* 8(11), pp. 889-895.

Scholtz, C.H., 1989. Unique foraging behaviour in *Pachysoma*. *Journal of Arid Environments,* 16(3), pp. 305-313.

Scholtz, C.H., Harrison, J. & Grebennikov, V., 2004. Dung beetle (*Scarabaeus* (*Pachysoma*)) biology and immature stages: reversal to ancestral states under desert conditions (Coleoptera: Scarabaeidae)? *Biological Journal of the Linnean Society,* 83(4), pp. 453-460.

Scholtz, C.H., Davis, A.L.V. & Kryger, U., 2009. *Evolutionary biology and conservation of dung beetles*. Pensoft Publishers, Moscow.

Scully, E.D., Geib, S.M., Hoover, K., Tien, M., Tringe, S.G., Barry, K.W., et al., 2013. Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. *PLoS One,* 8(9), pp. e73827.

Shade, A. & Handelsman, J., 2012. Beyond the Venn diagram: the hunt for a core microbiome. *Environmental Microbiology,* 14(1), pp. 4-12.

Shao, L., Devenport, M. & Jacobs-Lorena, M., 2001. The peritrophic matrix of hematophagous insects. *Archives of Insect Biochemistry and Physiology,* 47(2), pp. 119-125.

Sharon, G., Segal, D., Ringo, J.M., Hefetz, A., Zilber-Rosenberg, I. & Rosenberg, E., 2010. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences,* 107(46), pp. 20051-20056.

Sharpton, T.J., 2014. An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science,* 5.

Sinclair, A.R.E., 1975. The resource limitation of trophic levels in tropical grassland ecosystems. *The Journal of Animal Ecology,* 44(2), pp. 497-520.

Siqueira, J.F., Fouad, A.F.J. & Rôças, I.N., 2012. Pyrosequencing as a tool for better understanding of human microbiomes. *Journal of Oral Microbiology,* 4.

SKEP 2008. Succulent Karoo ecosystem programme rapid appraisal report. *A discussion paper in preparation for strategic planning 2009-2014.* Cape Town.

Sole, C.L., Bastos, A.D.S. & Scholtz, C.H., 2007. Do individual and combined data analyses of molecules and morphology reveal the generic status of '*Pachysoma*' MacLeay (Coleoptera: Scarabaeidae)? *Insect Systematics & Evolution,* 38(3), pp. 311-330.

Spence, K.D. & Kawata, M.Y., 1993. Permeability characteristics of the peritrophic membranes of *Manduca sexta* larvae. *Journal of Insect Physiology,* 39(9), pp. 785-790.

Struck, M., 1994. Flowers and their insect visitors in the arid winter rainfall region of southern Africa: observations on permanent plots. Composition of the anthophilous insect fauna. *Journal of Arid Environments,* 28(1), pp. 45-50.

Suh, S.-O., McHugh, J.V. & Blackwell, M., 2004. Expansion of the *Candida tanzawaensis* yeast clade: 16 novel *Candida* species from basidiocarp-feeding beetles. *International Journal of Systematic and Evolutionary Microbiology,* 54(6), pp. 2409-2429.

Suhling, F., Jödicke, R. & Schneider, W., 2003. Odonata of African arid regions–are there desert species? *Cimbebasia,* 18, pp. 207-224.

Swift, M.J., Heal, O.W. & Anderson, J.M., 1979. *Decomposition in terrestrial ecosystems*. University of California Press, USA, California.

Terra, W.R., 2001. The origin and functions of the insect peritrophic membrane and peritrophic gel. *Archives of Insect Biochemistry and Physiology,* 47(2), pp. 47-61.

Terra, W.R. & Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry,* 109B(1), pp. 1-62.

Tigar, B.J. & Osborne, P.E., 1999. Patterns of biomass and diversity of aerial insects in Abu Dhabi's sandy deserts. *Journal of Arid Environments,* 43(2), pp. 159-170.

Tsai, Y.-L., Hayward, R.E., Langer, R.C., Fidock, D.A. & Vinetz, J.M., 2001. Disruption of *Plasmodium falciparum* chitinase markedly impairs parasite invasion of mosquito midgut. *Infection and Immunity,* 69(6), pp. 4048-4054.

Tsuchida, T., Koga, R., Shibao, H., Matsumoto, T. & Fukatsu, T., 2002. Diversity and geographic distribution of secondary endosymbiotic bacteria in natural populations of the pea aphid, *Acyrthosiphon pisum*. *Molecular Ecology,* 11(10), pp. 2123-2135.

Tsuchida, T., Koga, R., Horikawa, M., Tsunoda, T., Maoka, T., Matsumoto, S., et al., 2010. Symbiotic bacterium modifies aphid body color. *Science,* 330(6007), pp. 1102-1104.

Tyagi, P., Edwards, D.R. & Coyne, M.S., 2008. Use of sterol and bile acid biomarkers to identify domesticated animal sources of fecal pollution. *Water, Air, & Soil Pollution,* 187(1-4), pp. 263-274.

Unep, 1992. *World atlas of desertification*. Edward Arnold, UK, London.

Vallet-Gely, I., Lemaitre, B. & Boccard, F., 2008. Bacterial strategies to overcome insect defences. *Nature Reviews Microbiology,* 6(4), pp. 302-313.

Veivers, P.C., O'Brien, R.W. & Slaytor, M., 1982. Role of bacteria in maintaining the redox potential in the hindgut of termites and preventing entry of foreign bacteria. *Journal of Insect Physiology,* 28(11), pp. 947-951.

Vollaard, E.J. & Clasener, H.A., 1994. Colonization resistance. *Antimicrobial Agents and Chemotherapy,* 38(3), pp. 409.

Wang, P. & Granados, R.R., 2001. Molecular structure of the peritrophic membrane (PM): identification of potential PM target sites for insect control. *Archives of Insect Biochemistry and Physiology,* 47(2), pp. 110-118.

Watanabe, H. & Tokuda, G., 2010. Cellulolytic systems in insects. *Annual Review of Entomology,* 55, pp. 609-632.

Webb, D.P., 1977. Regulation of deciduous forest litter decomposition by soil arthropod feces. *In:* Matson, W.J. (ed.) *The role of arthropods in forest ecosystems.* pp. 57-69. Springer Berlin, Heidelberg.

Wharton, D.A., 2002. *Life at the limits: organisms in extreme environments*. Cambridge University Press UK, Cambridge.

Wilson, A.C.C., Ashton, P.D., Calevro, F., Charles, H., Colella, S., Febvay, G., et al., 2010. Genomic insight into the amino acid relations of the pea aphid, *Acyrthosiphon pisum*, with its symbiotic bacterium *Buchnera aphidicola*. *Insect Molecular Biology,* 19(s2), pp. 249-258.

Wisborg, P. & Rohde, R., 2005. Contested land tenure reform in South Africa: experiences from Namaqualand. *Development Southern Africa,* 22(3), pp. 409-428.

Wittman, T. & Fedorka, K.M., 2015. Male mate choice for unparasitized females in *Drosophila melanogaster*. *Journal of Insect Behavior,* 28(1), pp. 37-43.

Wong, A.C.-N., Chaston, J.M. & Douglas, A.E., 2013. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *The ISME Journal,* 7(10), pp. 1922-1932.

Wu, D., Daugherty, S.C., van Aken, S.E., Pai, G.H., Watkins, K.L., Khouri, H., et al., 2006. Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. *PLoS Biology,* 4(6), pp. e188.

Yun, J.-H., Roh, S.W., Whon, T.W., Jung, M.-J., Kim, M.-S., Park, D.-S., et al., 2014. Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied and Environmental Microbiology,* 80(17), pp. 5254-5264.

Zdobnov, E.M., Von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R.R., et al., 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science,* 298(5591), pp. 149-159.

Zhang, H. & Jackson, T.A., 2008. Autochthonous bacterial flora indicated by PCR-DGGE of 16S rRNA gene fragments from the alimentary tract of *Costelytra zealandica* (Coleoptera: Scarabaeidae). *Journal of Applied Microbiology,* 105(5), pp. 1277-1285.

Zook, D., 1998. A new symbiosis language. *ISS Symbiosis News,* 1(3), pp. 1-3.

Chapter 2: Materials and Methods

2.1 Chemical Reagents

Chemical reagents, buffers, enzymes and the respective suppliers are listed in Table 2.1-2.3 below.

Table 2.1: List of chemical reagents.

Chemical	Supplier	Country
Acetic acid	Merck	Germany
Ammonium acetate (NH_4Ac)	Merck-Saarchem	South Africa
Bromophenol blue	Sigma-Aldrich	USA
Chloroform	Merck	Germany
Disodium phosphate (Na_2HPO_4)	Merck-Saarchem	South Africa
Ethanol (EtOH)	Illovo sugar	South Africa
Ethylenediaminetetraacetic acid (EDTA)	Merck	Germany
GelRed™ (10000x)	Biotium	USA
Glycerol	Merck-Saarchem	South Africa
Hydrochloric acid (HCl)	Merck-Saarchem	South Africa
Isoamyl alcohol	Merck	Germany
Monopotassium phosphate (KH_2PO_4)	Merck-Saarchem	South Africa
SeaKem LE agarose	Lonza	USA
Sodium chloride (NaCl)	Sigma-Aldrich	USA
Tris (hydromethyl) aminomethane (Tris base)	Merck	Germany

Table 2.2: Enzymes used in this study.

Enzyme	Supplier	Country
Lysozyme	Sigma-Aldrich	USA
Proteinase K	Sigma-Aldrich	USA
DreamTaq™ DNA polymerase	Fermentas Life Science	Lithuania

Table 2.3: Compositions of buffers and solutions used in this study.

Buffer/Solution	Composition
0.5M EDTA	37.22% [w/v] in dH2O
0.5M Tris-HCl	6% [w/v] Tris-base in dH20
Ringer solution	0.12 g/L $CaCl_2$, 0.105 g/L KCl, 0.05 g/L $NaHCO_3$, 2.25 g/L NaCl (Sigma-Aldrich, USA)
CTAB	0.1M Tris-HCl [pH8.0], 1.4M NaCl, 0.02M EDTA [pH8.0]
Phenol:chloroform:isoamyl	25 [v/v] phenol: 24 [v/v] chloroform: 1 [v/v] isoamyl alcohol (Sigma-Aldrich, USA)
Chloroform:isoamyl	24 [v/v] chloroform: 1 [v/v] isoamyl alcohol
Phosphate buffered saline (PBS)	8 g/liter NaCl, 0.2 g/liter KCl, 1.44 g/liter Na_2HPO_4, 0.24 g/liter KH_2PO_4 [pH 7.4]
Tris-acetate-EDTA (TAE)	40mM Tris (pH 7.6), 20mM acetic acid, 1mM EDTA
6x GelRed™-Loading buffer solution	30x GelRed™, 5ml 6x Loading buffer
6x Loading buffer	0.25% [w/v] Bromophenol blue, 40% [v/v] glycerol

2.2 General Methodologies

2.2.1 Collection and Storage of *Pachysoma* Samples

Adult individuals of *P. endroedyi* (Figure 2.1a) and *P. striatum* (Figure 2.1b) from single breeding populations, were collected (Figures 2.2a and b) by the Scarab Research Group from the University of Pretoria. Samples were collected in September and October 2014 (for the amplicon sequencing study; see chapter 3) and a separate set of samples were collected in August 2015 (for the shotgun sequencing study; see chapter 4) from coastal sandveld near Kommandokraal, Namaqualand, South Africa (S31°29'58.4" E18°12'29.2"; Figures 2.3-2.4) under the Cape Nature permit number 0056-AAA008-00041. Beetles were identified on site. Due to their large body size, 99% ethanol was injected into their abdomens using sterile syringes for gut preservation (Montagna et al., 2015). Insects were then stored in 99% ethanol at -80˚C, until dissection. Furthermore, five cockroach (*Blaptica dubia*) samples were sourced from a pet shop for comparison of host DNA contamination reduction methods (Chapter 4). Gut dissections of cockroach samples were carried out as described for the *Pachysoma* samples (Section 2.2.3).

a)  b) 

Figure 2.1: *Pachysoma* samples collected for the present study: a) *P. endroedyi* and b) *P. striatum* (Courtesy of Hennie de Klerk).

a)  b) 

Figure 2.2: Collection of *Pachysoma* samples: a) Observation of a burrow; b) a typical *Pachysoma* spp. burrow with the food-source (plant detritus) and a beetle inside (Courtesy of Hennie de Klerk).

Figure 2.3: Aerial view of the Western coast of South Africa with the Namaqualand Desert collection site indicated by the yellow marker and shown in the insert (Insert map data: Google, AfriGIS (Pty) Ltd, 2016).

Figure 2.4: Burrow being examined surrounded by typical vegetation at the collection site (Photo courtesy of H. de Klerk).

2.2.2 Insect Identification

Morphological identification of each species was carried out by Prof. C. Scholtz and experts from the Scarab Research Group of the University of Pretoria. Taxonomic keys (Scholtz and Holm, 1985, Harrison et al., 2003) were used to morphologically identify the two *Pachysoma* species (see Chapter 3).

2.2.3 Gut Dissection

Gut dissections were performed under a Zeiss Stemi 2000-C dissection microscope (Zeiss, Oberkochen, Germany) as previously described (Correa et al., 2012) with modifications. All equipment was sterilised before use with 10% bleach and 70% EtOH. The average body length of *P. endreodyi* ranges from 20.7-26.4mm, and the one of *P. striatum* ~19 mm (Harrison et al., 2003). The insects were placed in a wax-lined glass Petri dish with quarter strength autoclaved Ringer solution. The thorax and abdominal integument were removed using scissors before pinning the specimen to the wax layer in the Petri dish. Forceps were used to remove the membranes covering the internal organs. The rectum was pulled downwards, moving the gut gently out of the body cavity. The mid- and hindguts were separated and stored in 1.5ml eppendorf tubes at -20°C until metagenomic DNA (mDNA) extraction for the amplicon study (Chapter 3). For the shotgun sequencing study (Chapter 4), whole-guts were removed and stored at -20°C until mDNA extraction. Photographs were taken of both the mid- and hindgut using an AxioCam ERc5s (Zeiss, Oberkochen, Germany, Figure 2.5).

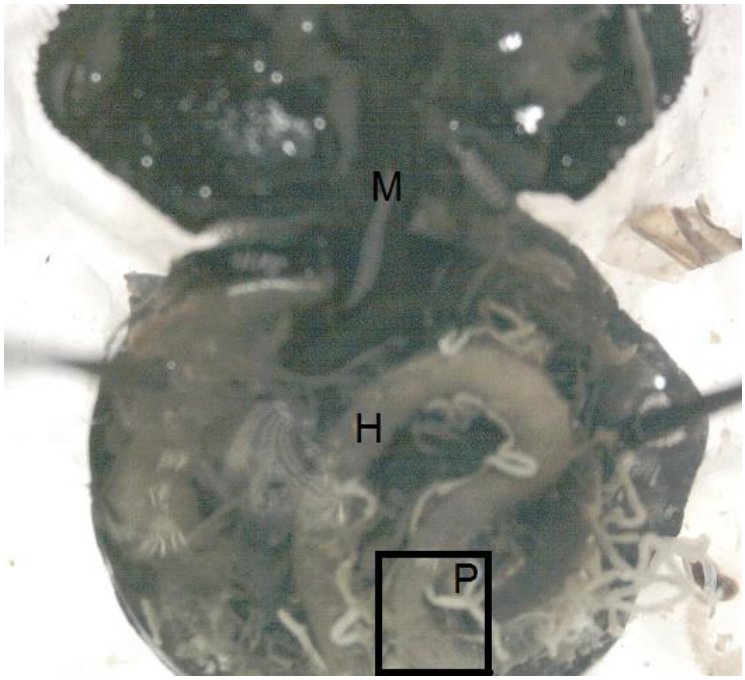


Figure 2.5: Dissected *P. striatum* showing H: hindgut, M: Midgut and P (square): pyloris (Foregut not shown).

2.2.4 Analytical Techniques

2.2.4.1 Gel Electrophoresis

Metagenomic DNA (see Sections 2.3.1 and 2.4.1) and PCR amplicons (see Section 2.3.2) were separated using 1% agarose gels prepared with 1X TAE buffer and SeaKem LE agarose. Samples were loaded with a "1X Gel Red / 3X loading buffer" mixture. After migration (100V, 35-60min), DNA was visualised under Ultraviolet (UV) light (Ultraviolet transilluminator, Spectroline, USA). Amplicon and mDNA sizes were determined by comparing their migration to the KAPA Universal Ladder (KAPA Biosystems, USA). Photographs of agarose gels were captured with the Molecular Imager Gel Doc XR+ (BioRad, USA).

2.2.4.2 Spectrophotometry

The Nanodrop 2000 spectrophotometer (Thermoscientific, Massachusetts, USA) was used to determine DNA concentrations and quality. DNA concentrations were calculated as a

value of [OD260 nm x ng/µl]. The purity of each sample was determined using the $OD_{260/280}$ and $OD_{260/230}$ ratios. A $A_{260/280}$ ratio of between 1.8 and 2 is regarded as pure (Nanodrop, 2007). The $A_{260/230}$ ratio is a secondary measure of purity and a preferred value of between 2.0 and 2.2 generally indicates a pure sample (Nanodrop, 2007).

2.3 454 Amplicon Pyrosequencing Methods – Chapter 3

2.3.1 Metagenomic DNA Extraction for Amplicon Pyrosequencing

Gut mDNA extraction was performed using a modified version of the protocols previously described by Calderón-Cortés et al. (2010) and Shi et al. (2012). Gut sections were weighed and crushed in liquid nitrogen using a sterilised epi-crusher. For 10mg of gut tissue, 100µl of a preheated (60°C) 2% Cetyl trimethylammonium bromide (CTAB) solution was added. The mixtures were incubated for 30min at 60°C before centrifugation for 5min at 10000rpm. The supernatant was transferred to a clean collection tube and enzymatic digestion of the gut samples was carried out with the addition of 2µl lysozyme (5mg/ml) per 100µl CTAB solution for 30min at 37°C under continuous shaking (120 rpm). 0.5µl Proteinase K (20mg/ml) per 100µl CTAB solution was then added (Priya et al., 2012), followed by an overnight incubation at 55°C with continuous shaking. One volume phenol:chloroform:isoamyl alcohol (25:24:1) solution was added. Tubes were inverted and centrifuged at 13000rpm at 4°C for 4min. One volume chloroform:isoamyl alcohol (24:1) solution was added to the top aqueous phase and the mixtures were inverted before centrifugation at 13000rpm at 4°C for 15min. This step was repeated until no protein contamination was observed (Priya et al., 2012). DNA was precipitated with 3M NH_4Ac (Lagisz et al., 2010) and ice cold 99.9% EtOH followed by overnight incubation at -20°C. Mixtures were centrifuged for 60min at 14000rpm at 4°C. The DNA pellet was washed twice with ice cold 70% EtOH and allowed to dry completely for 2 hours. The DNA pellet was resuspended in 50µl filter-sterilized nanopure H_2O overnight at 4°C (Lagisz et al., 2010), and stored at -20°C for downstream analysis.

2.3.2 454 Amplicon Pyrosequencing

2.3.2.1 454 Pyrosequencing of the Bacterial 16S Gene and Fungal ITS Region

The gut mDNA of five individuals from each *Pachysoma* spp. was sent to Molecular Research (MR DNA, Texas, United States of America; http://mrdnalab.com) for 16S rRNA gene and ITS region pyrosequencing using the Roche 454 GS FLX Titanium platform. The primers 27F (AGRGTTTGATCMTGGCTCAG; Lane (1991)) and 338R (AGTGCTGCCTCCCGTAGGAGT; Fierer et al. (2008)) were used to amplify the 16S rRNA gene region. Fungal specific fITS9 (GAACGCAGCRAAIIGYGA; Ihrmark et al. (2012)) and ITS4 (TCCTCCGCTTATTGATATGC; Gardes and Bruns (1993)) primers were used for the amplification of the ITS region.

2.3.2.2 Data Analysis

Raw pyrosequencing reads were filtered and analysed using mothur version 1.35.1 (Accessed May 2015 - January 2016) (Schloss et al., 2009, Schloss et al., 2011). Fasta, quality and flow files were extracted from the sff files using the sff.info command. For the bacterial 16S rRNA gene pyrosequencing reads, filtering of poor quality reads was done using the shhh.flows command allowing for one or two mismatches between barcodes and primers, respectively. Remaining sequences were quality filtered with the trim.seqs command allowing for maximum homopolymers of 8bp and a minimum sequence length of 100bp. Sequences were aligned to the SILVA reference database (http://www.arbsilva.de/download/arb-files/) using the align.seqs command. The screen.seqs and filter.seqs commands were used to retain only overlapping sequences. Chimeras were identified and removed using the chimera.uchime command. Sequences were classified against five databases, namely the Ribosomal Database Project (RDP), SILVA, National Center for Biotechnology Information (NCBI), The Dictyoptera gut microbiota reference Database (DictDb; data shown) and GreenGenes with a confidence threshold of 80%. OTUs were clustered for each individual beetle before removal of singletons using the remove.rare command. Samples were

subsampled to 1718 reads (Table 3.1, Chapter 3); i.e., the lowest number of reads across all samples.

ITS sequence reads were analysed as described above, with minor differences outlined previously (Bell et al., 2014). Filtering of poor quality reads was done using the trim.seqs command, allowing for one or four mismatches between barcodes and primers, respectively. Sequences were trimmed to 200bp using the chop.seqs command to ensure all sequences were the same length. Sequences were classified against the UNITE database (version 6) with a confidence threshold of 50% and subsampled to the lowest number of OTUs across all samples (107; Table 3.1, Chapter 3) for statistical analyses.

Square-root transformation of the data was performed to reduce the effect of highly dominant species and increase the effect of less abundant species (Clarke et al., 2006). To compare the composition of bacterial and fungal communities within the guts of the two *Pachysoma* spp. studied, the Bray-Curtis coefficient (Bray and Curtis, 1957) was used to build a dissimilarity matrix from the pre-treated data. To visualise (dis)similarities between the two *Pachysoma* spp. gut communities, Non-Metric Multi-Dimensional Scaling (nMDS) plots were constructed using Primer 6 software (version 6.1.5.81; Primer E Ltd, Plymouth, UK) and created by calculating distances between communities (Clarke, 1993). Kruskal's stress value was used to evaluate the ordination of sample placement in both two- and three-dimensional nMDS plots. A stress value greater than 0.2 suggests random placement, whereas a value lower than 0.1 suggests good correlation with the sample matrices (Clarke, 1993). Differences between the bacterial gut community structures of each *Pachysoma* spp. were tested for significance using one-way global Analysis of Similarities (ANOSIM) in Primer 6 software (version 6.1.5.81; Primer E Ltd, Plymyth, UK), using 10 000 permutations (Clarke, 1993). According to Clarke (1993) an R value closer to 1 indicates greater dissimilarity between datasets than within a dataset. Venn plots were created using R statistical package 2.15.1 in vegan (www.rproject.org), to determine the intra- and inter-specific distribution of bacterial and fungal OTUs. Diversity indices and rarefaction curves were generated in mothur using the

summary.single and rarefaction.single commands, respectively. Rarefaction curves were generated for each species in Microsoft Excel. Phylogenetic comparisons of both the bacterial and fungal dataset, were done using the relative abundance of all reads in the dataset so as to ensure inclusion of rare taxa. Relative abundances (%) were calculated from the number of reads of the specific microbial taxa divided by the total number of reads for the particular *Pachysoma* individual. Nucleotide sequences for both the bacterial and fungal datasets have been uploaded to NCBI (http://www.ncbi.nlm.nih.gov/) Short Read Archive (SRA) under the accession number SRP071915.

2.4 Hiseq Shotgun Sequencing Methods – Chapter 4

2.4.1 Metagenomic DNA Extraction

DNA was extracted following the method described in Section 2.3.1 with modifications to reduce host DNA contamination (Liu et al., 2011). Briefly, each whole-gut (160-290 mg) was placed in a 1.5ml eppendorf tube with 1ml Phosphate-buffered saline (PBS) solution. Samples were crushed using a sterilized epi-crusher. Samples were vortexed for 15s followed by three rounds of low-speed centrifugation (2900rpm at 4˚C for 10 min) to remove eukaryotic DNA before extraction (Liu et al., 2011). After each round, the supernatant was transferred to a new 15ml tube and the pellet resuspended in 1ml PBS. The supernatants were then pooled and centrifuged at 2400rpm at 4˚C for 10min. The supernatant was removed into a new tube. This step was repeated 4 times and the final supernatant was centrifuged at 9000xg at 4˚C for 15 min. The pellet was resuspended in 1ml PBS and 100µl of a 2% preheated (60˚C) CTAB solution was added for every 10mg of gut sample. DNA extraction then followed the method described in Section 2.3.1

2.4.2 PCR Amplification of 16S rRNA and 18S rRNA Genes

The 18S rRNA gene was amplified using the insect specific primers 18S intfw-st12 (ATCAAGAACGAAAGTTAGAG; Haring and Aspöck (2004)) and 18S rev1

(ATGGGGAACAATTGCAAGC; Haring and Aspöck (2004)) following a previous protocol (Sole et al., 2013) with minor modifications. Each PCR mixture (50µl) contained 0.2mM dNTPs, 0.5mM each primer, 0.01U DreamTaq (Fermentas, Lithuania), 1X Dream Taq Buffer and 1µl DNA template. The thermal cycle included an initial denaturation step at 95°C for 2min, followed by 30 cycles of 10s denaturation at 95°C , 20s annealing at 50°C and 1min extension at 72°C , with a final extension step at 72°C for 5min. The bacteria specific 16S rRNA gene was amplified using the primers 27F (AGAGTTTGATCCTGGCTCAG; Lane (1991)) and 1492R (TACGGYTACCTTGTTACGACTT; Lane (1991)). Each PCR mixture (final volume of 50µl) contained 0.2mM dNTPs, 0.4mg/ml BSA, 0.5mM of each primer, 0.01U DreamTaq, 1X Dream Taq Buffer and 1µl DNA template. The thermal cycle included an initial denaturation step at 94°C for 3min, followed by 25 cycles of denaturation at 94°C for 30s, annealing at 55°C for 1.30min and extension at 72°C for 2.30min with a final extension step at 72°C for 5min. PCR amplicons were visualized on 1% agarose gels.

2.4.3 Metagenome Sequencing and Data Analysis

Metagenomic DNA was extracted from the gut samples of three individuals per *Pachysoma* species, and pooled prior to sequencing using the Illumina Hiseq 2500 platform (Molecular Research (MR DNA), Texas, United States of America; http://mrdnalab.com). Reads were quality trimmed and assembled into contigs using the CLC Genomic Workbench (version 8.5.1; CLC Bio). Reads were trimmed by 15bp on the 5’ end and 1bp on the 3’ end allowing for no ambiguous bases and a minimum length of 75bp. For the assembly, a minimum contig length of 200bp, with a mismatch cost of 2bp and insertion and deletion costs of 3bp was accepted. Length and similarity fractions of 0.5 and 0.8, respectively, were accepted. High quality assemblies (including unmapped reads) were uploaded to the Metagenomics RAST server (MG-RAST)(Meyer et al., 2008) for taxonomic annotation. The assemblies were also annotated with Basic Local Alignment Search Tool (BLASTX) analysis using Diamond (Buchfink et al., 2015) against the non-redundant (NR) protein database with sensitive parameters (e-value 1e-3) and uploaded to the MEtaGenome ANalyzer (MEGAN)(Huson et

al., 2007, Huson et al., 2016). Contigs (or unmapped reads) predicted to originate from Metazoa (animals) and Streptophyta (plants) were extracted and removed from the original datasets, to ensure that only microbial open reading frames (ORFs)/contigs were analysed further. Prodigal was used for protein translation followed by a BLASTX and BLASTP analysis using Diamond with sensitive parameters and an e-value of 1e-3. The microbial datasets were analysed in MEGAN for both taxonomic classification and gene prediction using SEED and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Carbohydrate-Active enZYme (CAZy)(Cantarel et al., 2009) family classifications were based on Pfam (Bateman et al., 2004) domain assignments of the microbial datasets. KEGG E.C assignments were allocated using the KEGG Automatic Annotation System (KAAS)(Moriya et al., 2007). Sequences for both datasets have been uploaded to NCBI (http://www.ncbi.nlm.nih.gov/) Short Read Archive (SRA) under the accession number SRP071915.

2.5 References

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., et al., 2004. The Pfam protein families database. *Nucleic Acids Research,* 32(suppl 1), pp. D138-D141.

Bell, T.H., Hassan, S.E.-D., Lauron-Moreau, A., Al-Otaibi, F., Hijri, M., Yergeau, E. & St-Arnaud, M., 2014. Linkage between bacterial and fungal rhizosphere communities in hydrocarbon-contaminated soils is related to plant phylogeny. *The ISME Journal,* 8(2), pp. 331-343.

Bray, J.R. & Curtis, J.T., 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecological Monographs,* 27(4), pp. 325-349.

Buchfink, B., Xie, C. & Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods,* 12(1), pp. 59-60.

Calderón-Cortés, N., Quesada, M., Cano-Camacho, H. & Zavala-Páramo, G., 2010. A simple and rapid method for DNA isolation from xylophagous insects. *International Journal of Molecular Sciences,* 11(12), pp. 5056-5064.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B., 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research,* 37(suppl 1), pp. D233-D238.

Clarke, K.R., 1993. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology,* 18(1), pp. 117-143.

Clarke, K.R., Somerfield, P.J. & Chapman, M.G., 2006. On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray–Curtis coefficient for denuded assemblages. *Journal of Experimental Marine Biology and Ecology,* 330(1), pp. 55-80.

Correa, V.R., Majerczak, D.R., Ammar, E.-D., Merighi, M., Pratt, R.C., Hogenhout, S.A., et al., 2012. The bacterium *Pantoea stewartii* uses two different type III secretion systems to colonize its plant host and insect vector. *Applied and Environmental Microbiology,* 78(17), pp. 6327-6336.

Fierer, N., Hamady, M., Lauber, C.L. & Knight, R., 2008. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences,* 105(46), pp. 17994-17999.

Gardes, M. & Bruns, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes-application to the identification of mycorrhizae and rusts. *Molecular Ecology,* 2(2), pp. 113-118.

Haring, E. & Aspöck, U., 2004. Phylogeny of the Neuropterida: a first molecular approach. *Systematic Entomology,* 29(3), pp. 415-430.

Harrison, J.d.G., Scholtz, C.H. & Chown, S.L., 2003. A revision of the endemic south-western African dung beetle subgenus *Scarabaeus* (*Pachysoma*) MacLeay, including notes on other flightless Scarabaeini (Scarabaeidae: Scarabaeinae). *Journal of Natural History,* 37(3), pp. 305-355.

Huson, D.H., Auch, A.F., Qi, J. & Schuster, S.C., 2007. MEGAN analysis of metagenomic data. *Genome Research,* 17(3), pp. 377-386.

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al., 2016. MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Computational Biology,* 12(6), pp. e1004957.

Ihrmark, K., Bödeker, I., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., et al., 2012. New primers to amplify the fungal ITS2 region–evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiology Ecology,* 82(3), pp. 666-677.

Lagisz, M., Port, G. & Wolff, K., 2010. A cost-effective, simple and high-throughput method for DNA extraction from insects. *Insect Science,* 17(5), pp. 465-470.

Lane, D.J., 1991. 16S/23S rRNA sequencing. *In:* Stackebrandt, E. & Goodfellow, M. (eds.) *Nucleic acid techniques in bacterial systematics.* pp. 125-175. John Wiley and Sons, UK, Chichester.

Liu, N., Yan, X., Zhang, M., Xie, L., Wang, Q., Huang, Y., et al., 2011. Microbiome of fungus-growing termites: a new reservoir for lignocellulase genes. *Applied and Environmental Microbiology,* 77(1), pp. 48-56.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., et al., 2008. The metagenomics RAST server–a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics,* 9(1), pp. 386.

Montagna, M., Chouaia, B., Mazza, G., Prosdocimi, E.M., Crotti, E., Mereghetti, V., et al., 2015. Effects of the diet on the microbiota of the red palm weevil (Coleoptera: Dryophthoridae). *PLoS One,* 10(1), pp. e0117439.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C. & Kanehisa, M., 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research,* 35(suppl 2), pp. W182-W185.

Nanodrop 2007. 260/280 and 260/230 ratios NanoDrop® ND-1000 and ND-8000 8-sample spectrophotometers. *Technical Support Bulletin T009.* USA, Delaware: NanoDrop Technologies, Inc.

Priya, N.G., Ojha, A., Kajla, M.K., Raj, A. & Rajagopal, R., 2012. Host plant induced variation in gut bacteria of *Helicoverpa armigera*. *PLoS One,* 7(1), pp. e30768.

Schloss, P.D., Gevers, D. & Westcott, S.L., 2011. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One,* 6(12), pp. e27310.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology,* 75(23), pp. 7537-7541.

Scholtz, C.H. & Holm, E., 1985. *Insects of southern Africa*. Butterworths, South Africa, Durban.

Shi, Z., Wang, L. & Zhang, H., 2012. Low diversity bacterial community and the trapping activity of metabolites from cultivable bacteria species in the female reproductive system of the oriental fruit fly, *Bactrocera dorsalis* Hendel (Diptera: Tephritidae). *International Journal of Molecular Sciences,* 13(5), pp. 6266-6278.

Sole, C.L., Scholtz, C.H., Ball, J.B. & Mansell, M.W., 2013. Phylogeny and biogeography of southern African spoon-winged lacewings (Neuroptera: Nemopteridae: Nemopterinae). *Molecular Phylogenetics and Evolution,* 66(1), pp. 360-368.

Chapter 3: Amplicon Sequencing of Bacterial and Fungal Gut Communities of *Pachysoma* Species

3.1 Introduction

The microbial gut communities of a wide range of insect species have been investigated (for reviews see Dillon and Dillon (2004), Douglas (2015), Broderick and Lemaitre (2012), Brune (2013), Huang et al. (2010), Engel and Moran (2013)). The gut environment is considered to be an unstable system, as microorganisms face secretion of digestive enzymes, physical disturbance, habitat shedding during insect moults and other physiochemical conditions that are typically unfavourable for colonisation (Dillon and Dillon, 2004, Engel and Moran, 2013, Douglas, 2015). However, there are significant benefits to gut colonisation, including high nutrient availability and protection from external environmental stressors (Hooper et al., 2002, Douglas, 2015).

The relationships between host and gut microbiota range across the full spectrum of interactions; i.e., from pathogenic to obligate mutualism (Dillon and Dillon, 2004). When beneficial to their host, insect-associated microbial communities may participate in a number of activities including degradation of recalcitrant materials such as lignocellulose (Park et al., 2007, Douglas, 2009, Calderón-Cortés et al., 2012, Douglas, 2013, He et al., 2013), the production of nutrients and vitamins (Douglas, 2009, Douglas, 2013, 2015), the production of components of cohesion pheromones (Dillon et al., 2002), nitrogen fixation and utilisation of nitrogenous waste products (Nardi et al., 2002, Douglas, 2009, Douglas, 2013, 2015), protection against parasites (Koch and Schmid-Hempel, 2011, Douglas, 2015), change in body colouration (Tsuchida et al., 2010) and sterol synthesis (Douglas, 2009, Douglas, 2013).

Insect gut microbiomes are known to differ between insect species, driven by variations in the gut structure, different host lifecycles, host phylogeny and diet (Colman et al., 2012, Engel and Moran, 2013, Douglas, 2015). The gut microbiome is also influenced within the

individual insect or species, varying according to host life-stage (Vasanthakumar et al., 2008, Wang et al., 2011, Arias-Cordero et al., 2012, Andongma et al., 2015, Diouf et al., 2015), and/or diet (Miyata et al., 2007, Wang et al., 2011, Bertino-Grimaldi et al., 2013, Huang et al., 2013, Montagna et al., 2015b). Host diet influences gut microbial communities as they adapt to dietary changes through the induction of enzymes and changes in community structure (Kaufman and Klug, 1991, Santo Domingo et al., 1998, Bertino-Grimaldi et al., 2013). However, a core community may persist through major dietary changes (Huang et al., 2013, Schauer et al., 2014).

Studies on insect-microbial associations have mainly focused on termites (Brune and Friedrich, 2000, Ohkuma, 2003, Warnecke et al., 2007, Husseneder, 2010, Brune, 2013, Poulsen, 2015), but also on agriculturally important species such as honeybees (Hamdi et al., 2011, Moran, 2015), and medically important insects such as mosquitoes (Wang et al., 2011, Boissière et al., 2012, Osei-Poku et al., 2012, Wang et al., 2012, Coon et al., 2014). Little attention has been given to dung beetles, which are common and abundant insects in virtually all terrestrial environments and which facilitate nutrient cycling and bioturbation (Nichols et al., 2008). The desert dung beetle genus *Pachysoma*, from the Scarabaeini tribe, of which the quintessential scarab genus, *Scarabaeus* is also a member, consists of 13 species endemic to the south-west African coast (Harrison et al., 2003, Sole et al., 2005). Members of *Pachysoma* exhibit atypical feeding behaviour. While most adult dung beetles feed, by filtration, on minute particulate fragments in wet dung (Holter and Scholtz, 2011, 2013), adult *Pachysoma* feed on various and varying dry food-sources: plant detritus, dung pellets or both. These substrates are collected on the soil surface and masticated with specially-adapted mouthparts (Figure 3.1) (Scholtz, 1989, Harrison et al., 2003, Sole et al., 2005, Holter and Scholtz, 2011).

Given that insect gut microorganisms are known to be involved in the degradation of recalcitrant materials such as lignocellulosic compounds (Douglas, 2009, Calderón-Cortés et al., 2012, Engel and Moran, 2013, Douglas, 2015), it follows that the gut microbiomes of desert

insects may play a significant role in carbon-turnover in desert ecosystems. By studying the gut microbiome diversity of *Pachysoma* spp. feeding on different plentiful and readily-available substrates, it is possible to consider the effects of host diet and/or host phylogeny on gut microbiome assembly processes. This study was designed to characterise the gut microbial (bacterial and fungal) assemblages of coprophagous (*P. striatum*) (Harrison et al., 2003) and detritivorous (*P. endroedyi*; C. Scholtz Pers. Comm.) members of the same genus from the same location and to potentially determine whether host diet and/or host phylogeny could be deterministic factors in *Pachysoma* gut microbial community assembly.

3.2 Results and Discussion

3.2.1 The Desert Beetle Genus *Pachysoma*

The distribution of the *Pachysoma* species is restricted to the arid coastal regions of south-western Africa, principally because of the flightless nature of the genus (Harrison et al., 2003). The genus *Pachysoma* forms three distinct lineages, supporting six (lineage 1), four (lineage 2) and three (lineage 3) species, respectively (Sole et al., 2005). *Pachysoma endroedyi* is located in lineage 1 and *P. striatum* in lineage 2 (Figure 3.1). The driving forces behind the formation of these three lineages are currently unknown. However, it has been noted that all members of lineage 3 have a uniform diet (Figure 3.1) and originate from desert areas with a consistent aridity index (Harrison et al., 2003, Sole et al., 2005), whereas both the aridity index of the desert locations from which lineage 1 and 2 members originate, and their diets, fluctuate (Figure 3.1).

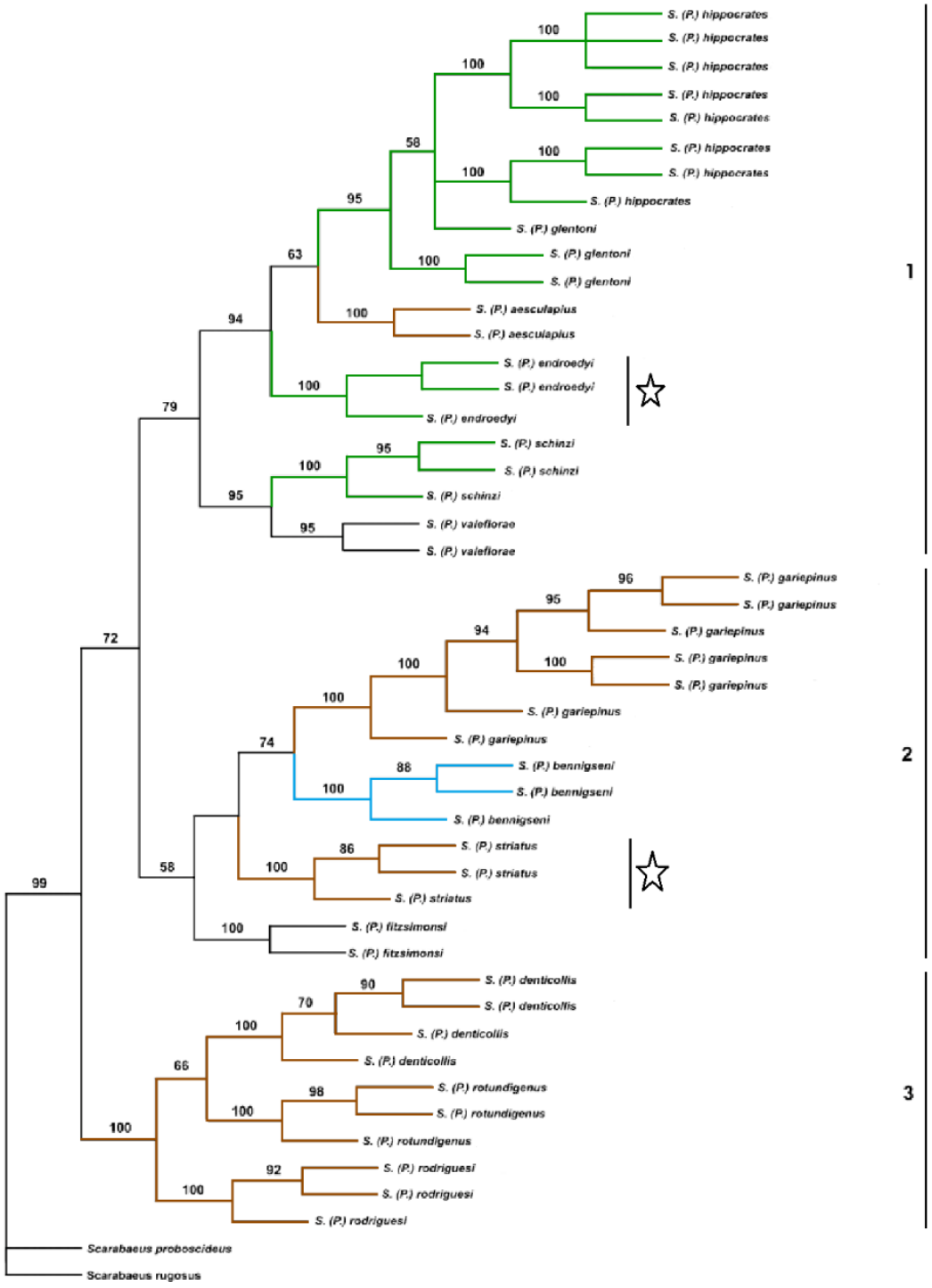


Figure 3.1: Cytochrome oxidase I gene Parsimony tree phylogeny of 13 *Pachysoma* spp. Branch colours indicate the diet of the *Pachysoma* spp.: dung (brown), plant detritus (green), polyphagous (blue) and unknown (black). This phylogentic tree was adapted, with permission, from (Sole et al., 2005) and the dietary information taken from both (Harrison et al., 2003) and personal observations by Prof C. Scholtz. Two species of *Scarabaeus*, (*S. proboscideus* and *S. rugosus*) which is the sister-genus to *Pachysoma* and a typical wet-dung-feeder, were used as outgroups. Numbers to the right of the tree indicate the three *Pachysoma* lineages. The two species considered in this study, *P. endroedyi* and *P. striatum*, are indicated with stars (Adapted from Sole et al. (2005)).

The diet of *P. striatum* consists predominantly of the dry dung pellets (Scholtz, 1989, Harrison et al., 2003) of various small native mammalian herbivores and sheep. Despite observations from a decade ago stating that *P. endroedyi* was a polyphagous feeder (Harrison et al., 2003), numerous and wide-scale recent observations suggest that *P. endroedyi* is a detrivore (Prof C. Scholtz, pers. comm.), the classification adopted in this study. *Pachysoma* species have specialised anatomical and physiological features for mastication and digestion of fragments from plant detritus and dry dung (Holter and Scholtz, 2011, 2013).

The linkage between host and gut microbiome is believed to be bidirectional, in that gut microorganisms can provide nutritional assistance to the insect host (Douglas, 2011, 2013) while the host diet influences the gut microbiome assembly (Miyata et al., 2007, Wang et al., 2011, Bertino-Grimaldi et al., 2013, Huang et al., 2013, Montagna et al., 2015b). However, host phylogeny may also impact gut microbiome composition (Colman et al., 2012, Yun et al., 2014), irrespective of the diet.

3.2.2 Sequencing Outputs and Diversity Indices of the Bacterial 16S rRNA Gene and Fungal ITS Region of the *Pachysoma* Gut Microbiome

The gut microbiomes of five detritivorous *P. endroedyi* and five coprophagous *P. striatum* individuals were determined by 16S rRNA gene amplicon sequencing. After removal of chimeras and singletons, 39050 bacterial and 1492 fungal reads remained, with mean read lengths of 238bp and 100bp, respectively. Only 462 bacterial reads were obtained for *P. endroedyi* individual 3 (Table 3.1), which was therefore removed from further analysis. Considerable variation in the number of bacterial sequence reads was noted between individuals, ranging from 1718 to 2817 and 3911 to 10106 for *P. endroedyi* and *P. striatum,* respectively. However, Good's coverage (>0.97 for all samples), rarefaction and chao1 diversity indices suggested that the coverage of *Pachysoma* bacterial gut communities (Figures 3.2a and b) were sufficient for a valid comparison between individuals. The fungal ITS region could not be amplified in samples from the detritivorous species *P. endroedyi,*

despite repeated attempts. The absence of fungi in the insect gut has previously been noted for individuals of various insect groups including Neuroptera and Coleoptera (using culture-dependent techniques) (Nguyen et al., 2007). In the fungal ITS sequence datasets for *P. striatum,* diversity indices and rarefaction curves showed low coverage for all but *P. striatum* individual 2, suggesting that the fungal diversity was generally underestimated (Table 3.1; Figure 3.2c).

Table 3.1: Values for sequence reads, OTUs, phyla and diversity indices for bacterial and fungal gut communities of *P. endroedyi* and *P. striatum* individuals.

	Individual	Number of reads	Number of OTUs	Phyla	Singletons	Chao	Invsimpson	Shannon	Coverage
Bacterial 16S rRNA gene	*P. endroedyi* 1	2120	213	10	105	244.61	19.01	3.88	0.97
	P. endroedyi 2	1718	258	11	133	282.34	81.00	4.91	0.97
	P. endroedyi 3	462	97	11	42	112.62	21.47	3.82	0.94
	P. endroedyi 4	2175	271	6	177	287.59	62.44	4.80	0.98
	P. endroedyi 5	2817	317	9	193	335.83	60.13	4.83	0.98
	P. striatum 1	10106	157	4	84	174.53	8.77	2.82	1.00
	P. striatum 2	4901	158	4	87	194.96	16.64	3.38	0.99
	P. striatum 3	4208	140	6	51	183.05	8.63	2.88	0.99
	P. striatum 4	3911	119	4	71	125.84	8.12	2.93	1.00
	P. striatum 5	6620	172	5	100	201.29	13.36	3.32	0.99
Fungal ITS region	*P. striatum* 1	136	88	1	156	179.50	96.63	4.29	0.55
	P. striatum 2	939	202	2	602	222.81	51.93	4.56	0.94
	P. striatum 3	199	106	2	223	248.38	88.35	4.40	0.66
	P. striatum 4	107	70	2	153	157.50	65.94	4.04	0.53
	P. striatum 5	111	63	2	146	134.75	52.18	3.88	0.62

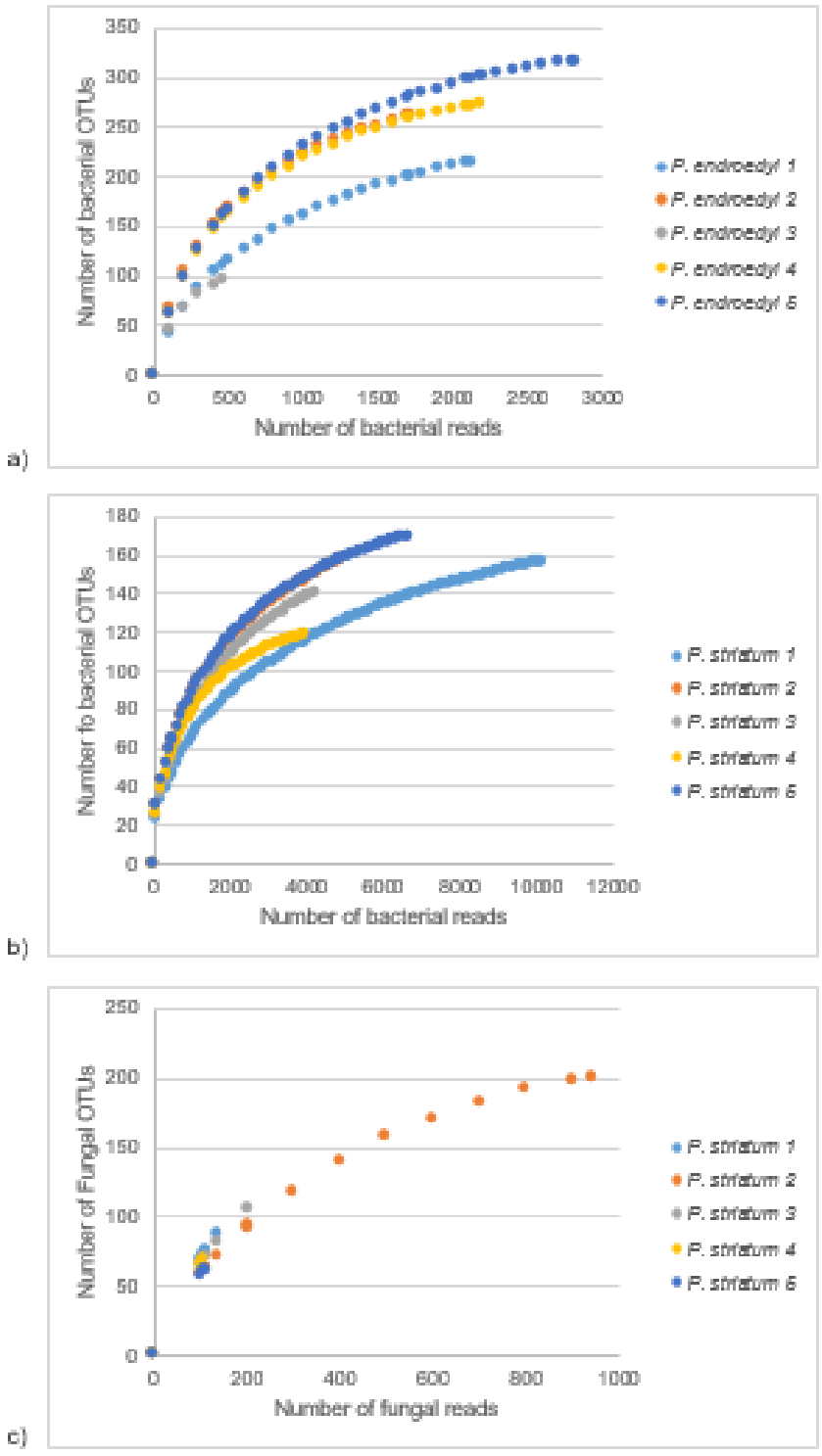


Figure 3.2: Rarefactions curves showing gut microbial community richness of all *Pachysoma* individuals for bacterial 16S rRNA gene amplicon data of: a) *P. endroedyi*, b) *P. striatum;* and c) fungal ITS region amplicon data of *P. striatum*.

A total of 1009 bacterial and 294 fungal OTUs were detected at an identity threshold of 97% (Table 3.1). Numbers ranged from 213 to 317 and 119 to 172 in the *P. endroedyi* and *P. striatum* gut samples, respectively (Table 3.1). These values are comparable with results obtained for termite and cockroach gut microbiomes (Sabree and Moran, 2014). It should be noted that the fungal ITS sequence read lengths were short (only 100bp), which could explain the poor phylogenetic resolution of *P. striatum* fungal gut communities (Lim et al., 2010).

In both *Pachysoma* spp., the number of bacterial 16S rRNA sequence reads was inversely proportional to the number of bacterial OTUs; i.e., *P. striatum* gut samples had a higher average number of bacterial reads (5949 ± 2550) but a lower average number of bacterial OTUs (149 ± 20) when compared to *P. endroedyi* (2208 ± 455 reads and 265 ± 43 OTUs, respectively). Those data suggest that the gut bacterial communities of *P. striatum* are composed of a relatively low number of dominant phylotypes at high abundance (Ahn et al., 2012, Gauthier et al., 2015, Montagna et al., 2015b). Contrastingly, the *P. endroedyi* gut bacterial community may include a higher bacterial diversity (Boucias et al., 2013, Montagna et al., 2015b). This inverse relationship, and the higher Shannon diversity index of the *P. endroedyi* gut bacterial community (4.6 ± 0.5) compared with the *P. striatum* gut community (3.1 ± 0.3; Table 3.1), suggests that competition is greater in the *P. striatum* gut than in *P. endroedyi*. This difference may be a reflection of the different diets, as insects feeding on simple diets (e.g., the coprophagous diet of *P. striatum*) commonly have a lower gut bacterial diversity than those feeding on more complex diets (e.g., the detritivorous diet of *P. endroedyi*) (Colman et al., 2012, Yun et al., 2014).

3.2.3 Inter-specific Variations in Bacterial and Fungal *Pachysoma* Gut Communities

The gut bacterial communities of *P. endroedyi* and *P. striatum* were significantly different, sharing only 3.7% of bacterial OTUs (Figure 3.3; ANOSIM [R=1.00, p<0.008]). Both host phylogeny and host diet could be driving forces for the observed differences (Chandler

et al., 2011, Colman et al., 2012). For example, the Hymenopteran gut microbiome has previously been shown to be influenced by host phylogeny, while the gut microbiomes of detritivorous insects (e.g., certain termites, Coleoptera and Diptera) are dictated by diet (Colman et al., 2012). Gut bacterial communities of *Drosophila* spp. also appear to be impacted by host diet rather than host phylogeny (Chandler et al., 2011). In Coleoptera (the order in which *Pachysoma* is placed), gut bacterial communities are significantly different to those of other insect groups (Colman et al., 2012), indicating that host phylogeny is a significant driving force for gut microbiome assembly. However, within Coleoptera, significant similarities in bacterial assemblages of certain beetles with similar diets (e.g., those feeding on live arboreal tissue) have also been noted (Colman et al., 2012), which suggests that diet may also be a deterministic factor. It should, however, be noted that no coprophagous insects were included in this study (Colman et al., 2012), making a direct comparison with *Pachysoma* speculative.

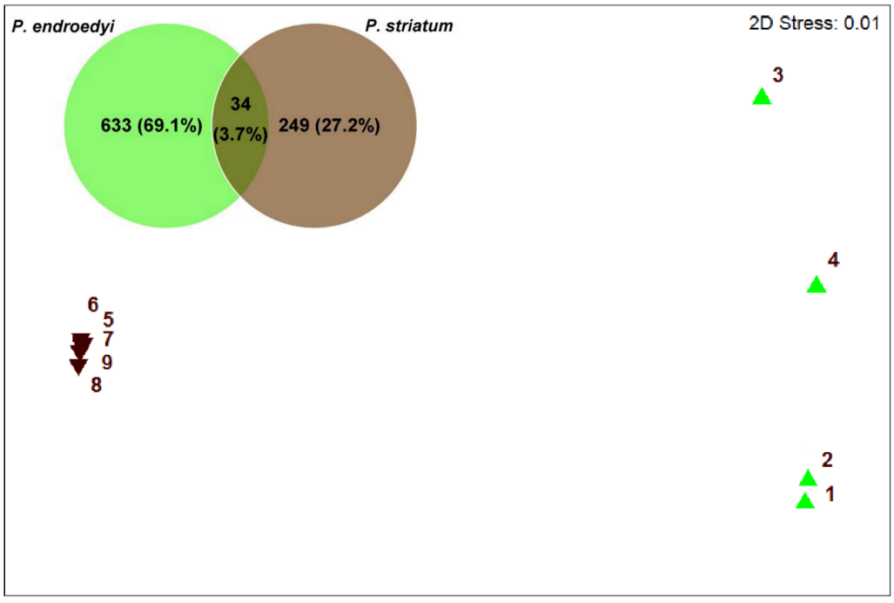


Figure 3.3: nMDS ordination plot based on Bray-Curtis distance matrices of bacterial 16S rRNA gene pyrosequencing data for *P. endroedyi* and *P. striatum* individuals. A stress value of less than 0.1 represents a high quality ordination. *Pachysoma endroedyi* and *P. striatum* are represented by green and inverted brown triangles, respectively.

It is not possible to compare the gut fungal communities of the two insect species studied, given that despite numerous attempts it was not possible to PCR-amplify fungal ITS sequences from the detritivorous *P. endroedyi.* While it is unlikely that fungal species are completely absent from the gut microbiome of this species, this negative result suggests that they may represent a relatively minor fraction of the total gut microbial diversity. To fully confirm this, the sample size should be increased and *P. endroedyi* individuals from multiple breeding populations should be investigated.

It is expected that host diet would be a contributing factor in the presence (or absence) of fungi in the *Pachysoma* gut. For example, true yeasts (Saccharomycetes) are typically observed in the guts of litter-, plant- and wood-feeding insects (Suh et al., 2008, Scully et al., 2013, Santana et al., 2015), but not in those of predacious insects (Nguyen et al., 2007, Shao et al., 2015).

3.2.4 Intra-specific Variation of *Pachysoma* Gut Microbial Communities

Large intra-specific differences in *Pachysoma* gut communities were noted, with the majority of OTUs being unique to each *Pachysoma* individual (Figures 3.4a and b, Figure 3.5a) and only 11 (1.1%) and 17 (3.3%) bacterial OTUs being shared between individuals of *P. endroedyi* and *P. striatum*, respectively. Furthermore, only two non-abundant fungal OTUs (ranging from 1.6-1.7% of the community) were shared among the five *P. striatum* individuals (Figure 3.5a). Such intra-specific differences, relating to the relative abundances and diversity of bacterial members of gut communities, are not uncommon, as has been observed for honeybees (*Apis cerana* and *A. mellifera* (Ahn et al., 2012)), mosquitoes (*Aedes* spp., *Culex* spp., *Anopheles* spp., *Mansonia* spp.; (Boissière et al., 2012, Osei-Poku et al., 2012)) and the red palm weevils *Rhynchophorus ferrugineus* and *R. vulneratus* (Montagna et al., 2015b), among others. A recent study on the gut microbiomes of 218 different insect species from 21 orders (Yun et al., 2014) indicated that 46% of the total number of bacterial OTUs detected (n=9301) were only observed in single individuals. The large intra-specific variation noted in

Pachysoma could be influenced by the stochastic, and transient, process of microorganisms entering the gut with the food-source (Douglas, 2015) and, for *P. striatum*, the different amounts of feeding material contained in the guts of each individual (Dillon and Dillon, 2004). Furthermore, it cannot be excluded that the 'time of feeding' prior to sampling may also have had an influence on intra-specific gut microbiome variability (Dillon and Dillon, 2004).

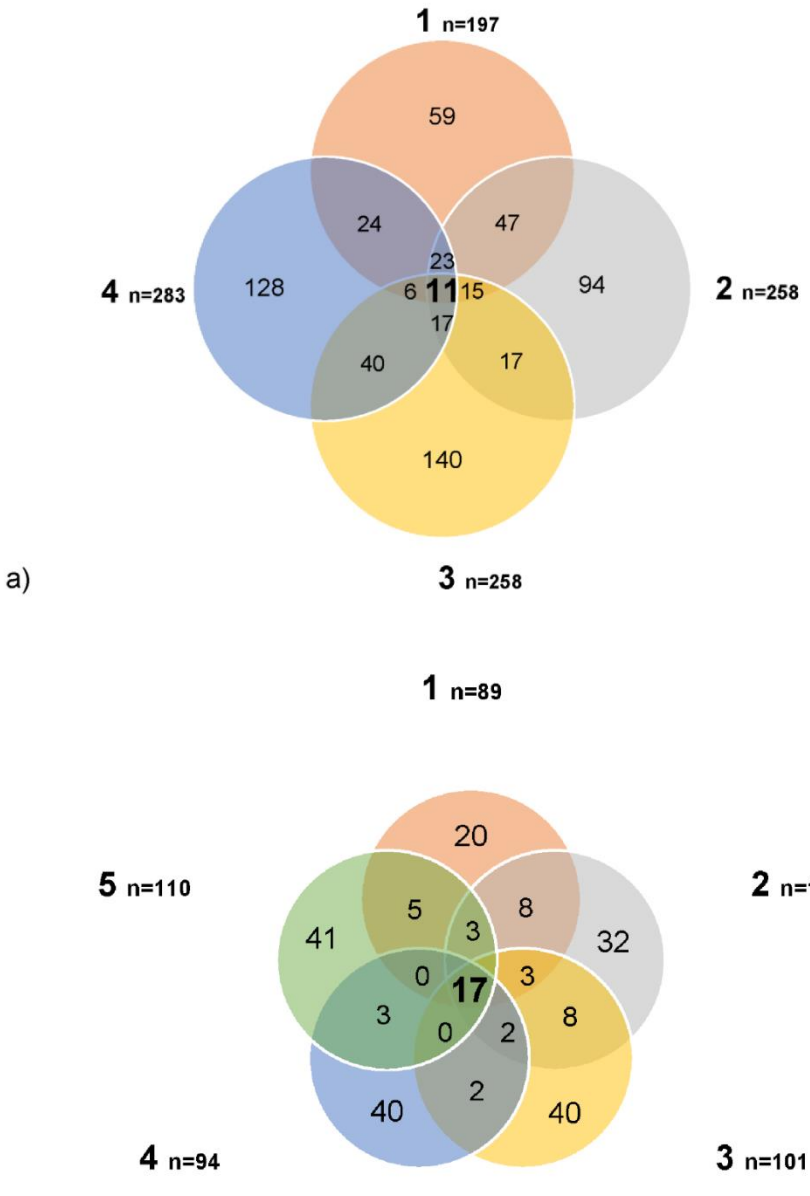
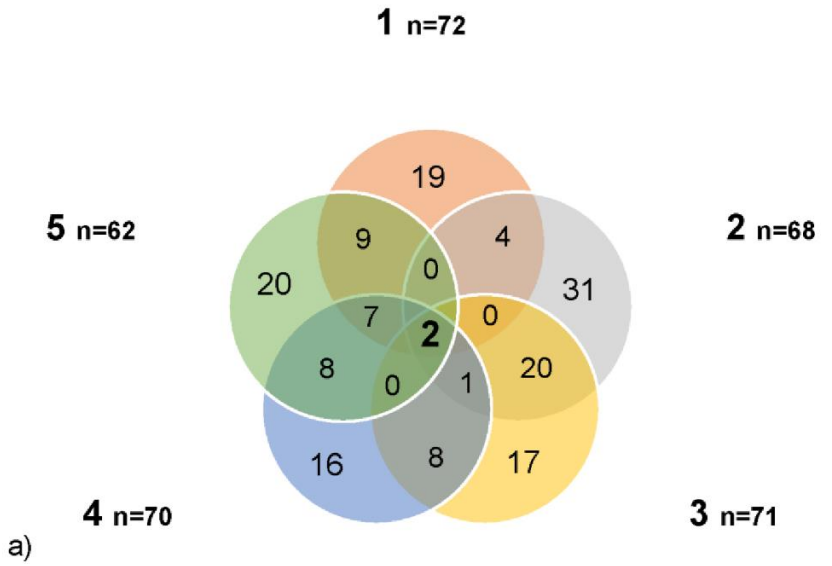


Figure 3.4: Venn diagrams showing distribution of bacterial OTUs between (a) *P. endroedyi* and (b) *P. striatum* individuals based on the 16S rRNA gene pyrosequencing analysis. Shared OTUs are shown in bold. Numerical labels are shown for each individual.



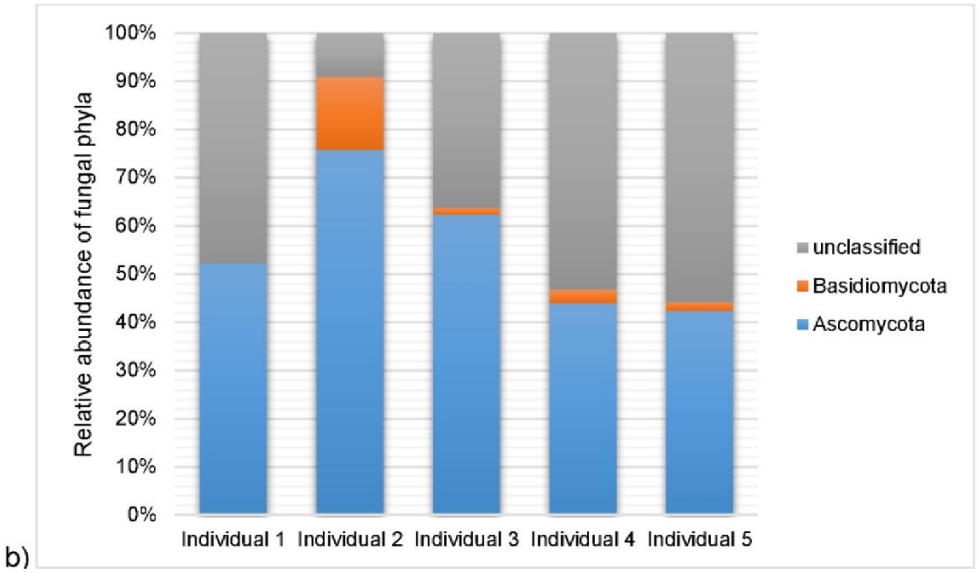


Figure 3.5: a) Venn diagram comparing the distribution of fungal OTUs between *P. striatum* individuals based on the ITS region pyrosequencing analysis. Shared OTUs are shown in bold. Numerical labels are given for each individual; b) Relative abundance of fungal phyla in five *P. striatum* individuals based on ITS rRNA gene region pyrosequencing analysis at a 97% identity threshold.

Of the shared bacterial OTUs, only one (assigned to the phylum Bacteroidetes) and eight (4 Firmicutes, 2 Actinobacteria, 1 Bacteroidetes and 1 Proteobacteria) were abundant (i.e., represented >2% reads) in the *P. endroedyi* and *P. striatum* gut samples, respectively. This distribution is strongly suggestive that the *Pachysoma* gut core community is very small, as has been proposed for the "minimal core" model (Hamady and Knight, 2009). Other studies have noted the presence of consistent core microbial communities within individuals of the same insect species (e.g., the bed bug *Cimex lectularius*; (Meriweather et al., 2013) and bumble bee *Bombus terrestris* (Billiet et al., 2015, Meeus et al., 2015)), or across taxonomic levels (e.g., across the ant tribe Cephalotini; (Anderson et al., 2012)). In the termite *Reticulitermes flavipes*, a substantial core bacterial microbiome (65% shared OTUs) was noted, regardless of the artificial feeding diet, suggesting that host phylogeny may play a more important role than host diet in the assembly of the gut microbiome (Huang et al., 2013). Similar results have been noted in cockroaches (Schauer et al., 2014). However, with a minimal core microbiome in both *Pachysoma* spp., phylogeny appears less important than diet. Furthermore, a minimal core gut microbiome may result from negative interactions between gut microorganisms, such as antagonism or amensalism, or indicate, as for *Drosophila* (Wong et al., 2013), the establishment of 'non-gut-specific' microorganisms.

It has been suggested that a 'functional' rather than a 'phylogenetic' core microbiome may be more informative in determining the assembly of gut microbiomes (Karasov et al., 2011). In studies on humans, which typically follow the minimal core model, functional gene diversity appears to be broadly similar across individuals (Turnbaugh et al., 2009, Karasov et al., 2011). Therefore, there may be a functional core community in each *Pachysoma* sp. studied, displaying shared metabolic capacities (Turnbaugh et al., 2009); i.e., exhibiting functional redundancy. As such, it has been suggested that a comparison of functional properties of hosts feeding on different diets can guide an understanding of the functional roles of different gut microbiomes (Karasov et al., 2011).

3.2.5 Phylogenetic Diversity of Bacterial and Fungal *Pachysoma* Gut Communities

The gut bacterial diversity of *P. endroedyi* was higher (6-11 phyla; Figure 3.6) than that of *P. striatum* (4-7 phyla; Figure 3.6). The *P. endroedyi* gut samples were dominated by Bacteroidetes (18.0-54.8%), Firmicutes (10.0-34.6%), Proteobacteria (8.7-18.1%) and Planctomycetes (2.5-25.7%), while Actinobacteria (0.1-22.5%), Elusimocrobia (0-9.3%) and Synergistetes (0-7.3%) showed highly variable abundances (Figure 3.6). The remaining 7 phyla each represented less than 2% of the community and were often detected in single insects. In *P. striatum,* Bacteroidetes (3.0-57.1%), Firmicutes (18.9-56.2%), Proteobacteria (6.4-32.1%) and Actinobacteria (5.2-21.0%) were also dominant phyla although the relative abundances varied between individuals (Figure 3.6). Three minor phyla (<2% abundance) were only detected in two *P. striatum* individuals, namely Deferribacteres, Planctomycetes and Synergistetes (Figure 3.6). All the identified bacterial phyla have previously been reported in insect microbiomes, with the phyla Firmicutes, Proteobacteria, Bacteroidetes and Actinobacteria commonly abundant in insect gut samples (Jones et al., 2013, Yun et al., 2014).

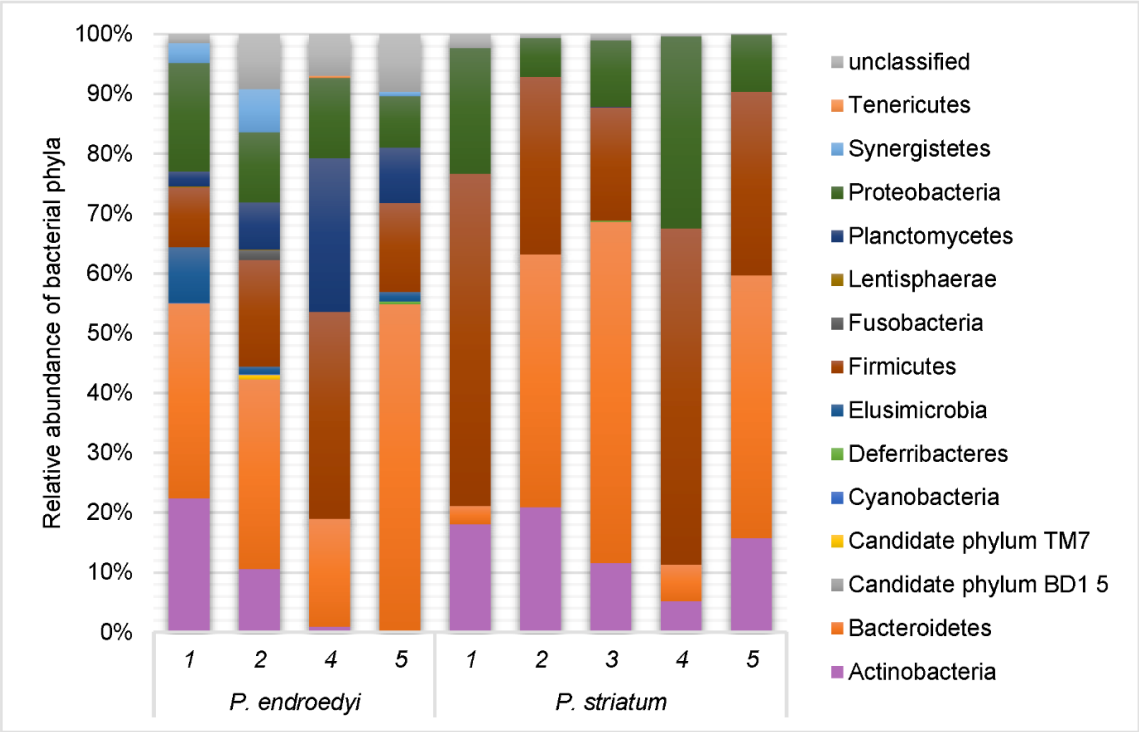


Figure 3.6: Comparison of inter-specific differences in relative abundance of bacterial phyla in the gut of two *Pachysoma* spp., *P. endroedyi* and *P. striatum*, based on 16S rRNA gene pyrosequencing analysis.

The presence of specific bacterial phyla and/or their relative abundances in insect gut samples may be linked to host diet. For example, certain insects with simpler diets (e.g., feeding on pollen and nectar (Ahn et al., 2012), fruit (Wang et al., 2014, Andongma et al., 2015), or sap (Pandey and Rajagopal, 2016)), contain gut bacterial communities which are typically dominated by heterotrophic Proteobacteria and/or Firmicutes. In contrast, Bacteroidetes (along with other phyla) were highly abundant in the gut microbiomes of insects feeding on plant materials such as wood and leaves (Köhler et al., 2012, Boucias et al., 2013, Schauer et al., 2014, Montagna et al., 2015b, Waite et al., 2015). The *P. striatum* gut bacterial communities did not display these patterns, suggesting that coprophagous diets may structure insect gut communities differently.

Fifteen and 11 bacterial genera were abundant (>2% relative abundance of reads) within the guts of *P. endroedyi* and *P. striatum*, respectively (Table 3.2). Only two of these genera

were abundant in both species (*Dysgonomonas* and unclassified Enterobacteriaceae; Table 3.2). *Dysgonomonas* was less abundant in *P. endroedyi* gut samples (2.8% ± 0) than in *P. striatum* (26.3% ± 0.2), in which it was the most abundant genus. *Dysgonomonas* have been reported to be present at high abundance in the gut system of the fungus-growing termite (*Macrotermes annandalei*) and red palm weevil larvae (*Rhynchophorus ferrugineus*) (Tagliavia et al., 2014, Zhang et al., 2014). Two species of *Dysgonomonas* have previously been characterised from the gut of termites (Yang et al., 2014, Pramono et al., 2015). Both species have been found to ferment glucose and xylan as a sole carbon source and to produce acetic acid as the major end-product (Yang et al., 2014, Pramono et al., 2015), suggesting roles in both the lignocellulosic biomass degradation pathway and in providing readily metabolisable substrates for ingestion by the host. The large difference in the abundance of this phylotype in the two *Pachysoma* species suggests a key nutritional role in *P. striatum* but not in *P. endroedyi.*

Table 3.2: Phylogenetic classification of the most abundant bacterial genera in the gut samples of *P. endroedyi* and *P. striatum*: i.e., representing >2% reads.

Phylum	Family	Genus	*P. endroedyi* (%)	*P. striatum* (%)
Actinobacteria	Propionibacteriaceae	*Proponiobacterium 1*	8.2 ±0.1	0.1 ± 0
		Tessaracoccus	0	7.6 ± 0
	unclassified	*unclassified*	0.1 ± 0	5.2 ± 0
Bacteroidetes	Bacteroidaceae	*Bacteroides*	7.6 ± 0	0
	Marinilabiaceae	*Uncultured 1*	2.7 ± 0	0
	Porphyromonadaceae 1	***Dysgonomonas***	2.8 ± 0	**26.3 ± 0.2**
	Porphyromonadaceae 4	*Proteiniphilum*	0.2 ± 0	3.5 ± 0
	Rikenellaceae	*Alisitpes IV*	6.5 ± 0	0
		unclassified	3.2 ± 0	0
	unclassified	*unclassified*	5.3 ± 0	0.1 ± 0
Elusimicrobia	Endomicrobiaceae	*Endomicrobium*	3.0 ± 0	0
Firmicutes	Enterococcaceae	*Vagococcus*	0.3 ± 0	4.9 ± 0
	Family XI Incertae Sedis	*unclassified*	0	10.9 ± 0.1
	Lachnospiraceae	*Uncultured 13*	0.9 ± 0	2.9 ± 0
		unclassified	3.3 ± 0	0.7 ± 0
	Ruminococcaceae	*Termite cockroach cluster*	5.3 ± 0.1	0.1 ± 0
		unclassified	3.3 ± 0	1.4 ± 0
	Veillonellaceae	*Anaeroarcus-Anaeromusa*	0	11.5 ± 0.1
Planctomycetes	**unclassified**	***unclassified***	**11.3 ± 0.1**	0
Proteobacteria	unclassified	*unclassified*	0	3.1 ± 0
	Insect cluster	*unclassified*	2.2 ± 0	0
	Enterobacteriaceae	*unclassified*	4.2 ± 0.1	7.4 ± 0.1
	Enterobacteriaceae 1	*unclassified*	0	2.5 ± 0.1
Synergistetes	Synergistaceae	*Candidatus Tammella*	2.8 ± 0	0

Percentages are the average read relative abundances in each species (*P. endroedyi*: n=4; *P. striatum*: n=5). Colours depict the species in which the bacterial genus is abundant: *P. endroedyi* (green), *P. striatum* (brown) or both (blue). The most abundant genus of each species is shown in bold.

An unclassified Planctomycetes dominated the gut samples of *P. endroedyi* (11.3% ± 0.1 relative abundance of reads). To the best of our knowledge this is the first report of an insect gut microbiome dominated by Planctomycetes. Planctomycetes were only detected in a single *P. striatum* individual at very low abundance (0.01%). Planctomycetes have previously

been detected in the guts of the termites *Syntermes wheeleri* and *Nasutitermes* spp. (Köhler et al., 2012, Santana et al., 2015), the cockroach *Shelfordella lateralis* (Schauer et al., 2014), adult and larval beetles (*Cryptocephalus* spp., *Prionoplus reticularis* and *Pachnoda* spp. (Andert et al., 2010, Reid et al., 2011, Montagna et al., 2015a)), the tree weta *Hemideina thoracica* (Waite et al., 2015) and the mosquito *Aedes albopictus* (Zouache et al., 2012), but only in low abundances (<1-5% relative abundance).

Ascomycota was the most abundant fungal phylum (42.3-75.7%) in all *P. striatum* gut samples, which is typical for insect gut microbiomes (Suh et al., 2004, Suh et al., 2005, Scully et al., 2013, Santana et al., 2015). Basidiomycota were not ubiquitously detected, and were observed only in the gut samples of four of the five *P. striatum* individuals (1.8-15.2%; Figure 3.5b). A substantial proportion of fungal ITS sequence reads could not be classified, even at the phylum level (9.1-55.9%; Figure 3.5b). Unfortunately, relatively little is known about insect gut fungal diversity (compared to bacterial diversity (Gibson and Hunter, 2010)), with the majority of published studies being based on culture-dependent methods which are typically biased when compared with culture-independent methods (Zhang et al., 2003, Gibson and Hunter, 2010).

3.3 Conclusion

This is the first study to investigate the gut microbiomes of any dung beetle feeding on dry food-sources and to compare those of closely related adult dung beetle species with very different diets but from the same locality. *Pachysoma* spp. are ecologically important in arid environments where they undoubtedly participate in nutrient cycling and bioturbation (Nichols et al., 2008). We have demonstrated that, as predicted, the gut microbiomes differed significantly between two species which feed on different substrates. However, both populations showed large intra-specific variations. Thus, to further characterise the gut microbiomes of these *Pachysoma* species, the number of individuals studied should be increased and populations from different sites investigated. Such experiments would make it

possible to evaluate whether inter-specific variation was higher than intra-specific variation within a single *Pachysoma* species.

We are unable to fully assess whether host phylogeny or the host diet is the dominant driver of the *Pachysoma* gut microbiomes. Nevertheless, we provide evidence that diet probably plays a significant role, particularly noting the fact that the gut microbiomes of the detritivorous *P. endroedyi* (feeding on complex food-sources) have higher bacterial diversities than those of the coprophagous species (feeding on relatively simple food-sources) (Colman et al., 2012, Yun et al., 2014). Functional gene analysis of the microbiomes of *P. endroedyi* and *P. striatum* could potentially assist in confirming the role that host diet plays in *Pachysoma* gut microbiome assembly (He et al., 2013).

3.4 References

Ahn, J.-H., Hong, I.-P., Bok, J.-I., Kim, B.-Y., Song, J. & Weon, H.-Y., 2012. Pyrosequencing analysis of the bacterial communities in the guts of honey bees *Apis cerana* and *Apis mellifera* in Korea. *Journal of Microbiology,* 50(5), pp. 735-745.

Anderson, K.E., Russell, J.A., Moreau, C.S., Kautz, S., Sullam, K.E., Hu, Y., et al., 2012. Highly similar microbial communities are shared among related and trophically similar ant species. *Molecular Ecology,* 21(9), pp. 2282-2296.

Andert, J., Marten, A., Brandl, R. & Brune, A., 2010. Inter-and intraspecific comparison of the bacterial assemblages in the hindgut of humivorous scarab beetle larvae (*Pachnoda* spp.). *FEMS Microbiology Ecology,* 74(2), pp. 439-449.

Andongma, A.A., Wan, L., Dong, Y.-C., Desneux, N., White, J.A. & Niu, C.-Y., 2015. Pyrosequencing reveals a shift in symbiotic bacteria populations across life stages of *Bactrocera dorsalis*. *Scientific Reports,* 5.

Arias-Cordero, E., Ping, L., Reichwald, K., Delb, H., Platzer, M. & Boland, W., 2012. Comparative evaluation of the gut microbiota associated with the below-and above-ground life stages (larvae and beetles) of the forest cockchafer, *Melolontha hippocastani*. *PLoS One,* 7(12), pp. e51557.

Bertino-Grimaldi, D., Medeiros, M.N., Vieira, R.P., Cardoso, A.M., Turque, A.S., Silveira, C.B., et al., 2013. Bacterial community composition shifts in the gut of *Periplaneta americana* fed on different lignocellulosic materials. *SpringerPlus,* 2(1), pp. 609.

Billiet, A., Meeus, I., Nieuwerburgh, F., Deforce, D., Wäckers, F. & Smagghe, G., 2015. Colony contact contributes to the diversity of gut bacteria in bumblebees (*Bombus terrestris*). *Insect Science,* Online version.

Boissière, A., Tchioffo, M.T., Bachar, D., Abate, L., Marie, A., Nsango, S.E., et al., 2012. Midgut microbiota of the malaria mosquito vector *Anopheles gambiae* and interactions with *Plasmodium falciparum* infection. *PLoS Pathogens,* 8(5), pp. e1002742.

Boucias, D.G., Cai, Y., Sun, Y., Lietze, V.U., Sen, R., Raychoudhury, R. & Scharf, M.E., 2013. The hindgut lumen prokaryotic microbiota of the termite *Reticulitermes flavipes* and its responses to dietary lignocellulose composition. *Molecular Ecology,* 22(7), pp. 1836-1853.

Broderick, N.A. & Lemaitre, B., 2012. Gut-associated microbes of *Drosophila melanogaster*. *Gut Microbes,* 3(4), pp. 307-321.

Brune, A., 2013. Symbiotic associations between termites and prokaryotes. *In:* Rosenberg, E., Delong, E.F., Lory, S., Stackebrandt, E. & Thompson, F. (eds.) *The prokaryotes.* pp. 545-577. Springer Berlin, Heidelberg.

Brune, A. & Friedrich, M., 2000. Microecology of the termite gut: structure and function on a microscale. *Current Opinion in Microbiology,* 3(3), pp. 263-269.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H. & Oyama, K., 2012. Endogenous plant cell wall digestion: a key mechanism in insect evolution. *Annual Review of Ecology, Evolution, and Systematics,* 43, pp. 45-71.

Chandler, J.A., Lang, J.M., Bhatnagar, S., Eisen, J.A. & Kopp, A., 2011. Bacterial communities of diverse *Drosophila* species: ecological context of a host–microbe model system. *PLoS Genetics,* 7(9), pp. e1002272.

Colman, D.R., Toolson, E.C. & Takacs-Vesbach, C.D., 2012. Do diet and taxonomy influence insect gut bacterial communities? *Molecular Ecology,* 21(20), pp. 5124-5137.

Coon, K.L., Vogel, K.J., Brown, M.R. & Strand, M.R., 2014. Mosquitoes rely on their gut microbiota for development. *Molecular Ecology,* 23(11), pp. 2727-2739.

Dillon, R.J. & Dillon, V.M., 2004. The gut bacteria of insects: nonpathogenic interactions. *Annual Reviews in Entomology,* 49(1), pp. 71-92.

Dillon, R.J., Vennard, C.T. & Charnley, A.K., 2002. A note: gut bacteria produce components of a locust cohesion pheromone. *Journal of Applied Microbiology,* 92(4), pp. 759-763.

Diouf, M., Roy, V., Mora, P., Frechault, S., Lefebvre, T., Hervé, V., et al., 2015. Profiling the succession of bacterial communities throughout the life stages of a higher termite *Nasutitermes arborum* (Termitidae, Nasutitermitinae) using 16S rRNA gene pyrosequencing. *PLoS One,* 10(10), pp. e0140014.

Douglas, A.E., 2009. The microbial dimension in insect nutritional ecology. *Functional Ecology,* 23(1), pp. 38-47.

Douglas, A.E., 2011. Lessons from studying insect symbioses. *Cell Host & Microbe,* 10(4), pp. 359-367.

Douglas, A.E., 2013. Microbial brokers of insect-plant interactions revisited. *Journal of Chemical Ecology,* 39(7), pp. 952-961.

Douglas, A.E., 2015. Multiorganismal insects: diversity and function of resident microorganisms. *Annual Review of Entomology,* 60, pp. 17-34.

Engel, P. & Moran, N.A., 2013. The gut microbiota of insects–diversity in structure and function. *FEMS Microbiology Reviews,* 37(5), pp. 699-735.

Gauthier, J.-P., Outreman, Y., Mieuzet, L. & Simon, J.-C., 2015. Bacterial communities associated with host-adapted populations of pea aphids revealed by deep sequencing of 16S ribosomal DNA. *PLoS One,* 10(3), pp. e0120664.

Gibson, C.M. & Hunter, M.S., 2010. Extraordinarily widespread and fantastically complex: comparative biology of endosymbiotic bacterial and fungal mutualists of insects. *Ecology Letters,* 13(2), pp. 223-234.

Hamady, M. & Knight, R., 2009. Microbial community profiling for human microbiome projects: tools, techniques, and challenges. *Genome Research,* 19(7), pp. 1141-1152.

Hamdi, C., Balloi, A., Essanaa, J., Crotti, E., Gonella, E., Raddadi, N., et al., 2011. Gut microbiome dysbiosis and honeybee health. *Journal of Appplied Entomology,* 135(7), pp. 524-533.

Harrison, J.d.G., Scholtz, C.H. & Chown, S.L., 2003. A revision of the endemic south-western African dung beetle subgenus *Scarabaeus* (*Pachysoma*) MacLeay, including notes on other flightless Scarabaeini (Scarabaeidae: Scarabaeinae). *Journal of Natural History,* 37(3), pp. 305-355.

He, S., Ivanova, N., Kirton, E., Allgaier, M., Bergin, C., Scheffrahn, R.H., et al., 2013. Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS One,* 8(4), pp. e61126.

Holter, P. & Scholtz, C.H., 2011. Re-establishment of biting mouthparts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on plant litter—old structures reacquired or new ones evolved? *Journal of Morphology,* 272(8), pp. 1007-1016.

Holter, P. & Scholtz, C.H., 2013. Elongated hindguts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on dry dung pellets or plant litter. *Journal of Morphology,* 274(6), pp. 657-662.

Hooper, L.V., Midtvedt, T. & Gordon, J.I., 2002. How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annual Review of Nutrition,* 22(1), pp. 283-307.

Huang, S., Zhang, H.Y., Marshall, S. & Jackson, T.A., 2010. The scarab gut: a potential bioreactor for bio-fuel production. *Insect Science,* 17(3), pp. 175-183.

Huang, X.-F., Bakker, M.G., Judd, T.M., Reardon, K.F. & Vivanco, J.M., 2013. Variations in diversity and richness of gut bacterial communities of termites (*Reticulitermes flavipes*) fed with grassy and woody plant substrates. *Microbial Ecology,* 65(3), pp. 531-536.

Husseneder, C., 2010. Symbiosis in subterranean termites: a review of insights from molecular studies. *Environmental Entomology,* 39(2), pp. 378-388.

Jones, R.T., Sanchez, L.G. & Fierer, N., 2013. A cross-taxon analysis of insect-associated bacterial diversity. *PLoS One,* 8(4), pp. e61218.

Karasov, W.H., Martínez del Rio, C. & Caviedes-Vidal, E., 2011. Ecological physiology of diet and digestive systems. *Annual Review of Physiology,* 73, pp. 69-93.

Kaufman, M.G. & Klug, M.J., 1991. The contribution of hindgut bacteria to dietary carbohydrate utilization by crickets (Orthoptera: Gryllidae). *Comparative Biochemistry and Physiology Part A: Physiology,* 98(1), pp. 117-123.

Koch, H. & Schmid-Hempel, P., 2011. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proceedings of the National Academy of Sciences,* 108(48), pp. 19288-19292.

Köhler, T., Dietrich, C., Scheffrahn, R.H. & Brune, A., 2012. High-resolution analysis of gut environment and bacterial microbiota reveals functional compartmentation of the gut in wood-feeding higher termites (*Nasutitermes* spp.). *Applied and Environmental Microbiology,* 78(13), pp. 4691-4701.

Lim, Y.W., Kim, B.K., Kim, C., Jung, H.S., Kim, B.-S., Lee, J.-H. & Chun, J., 2010. Assessment of soil fungal communities using pyrosequencing. *The Journal of Microbiology,* 48(3), pp. 284-289.

Meeus, I., Parmentier, L., Billiet, A., Maebe, K., van Nieuwerburgh, F., Deforce, D., et al., 2015. 16S rRNA amplicon sequencing demonstrates that indoor-reared bumblebees (*Bombus terrestris*) harbor a core subset of bacteria normally associated with the wild host. *PLoS One,* 10(4), pp. e0125152.

Meriweather, M., Matthews, S., Rio, R. & Baucom, R.S., 2013. A 454 survey reveals the community composition and core microbiome of the common bed bug (*Cimex lectularius*) across an urban landscape. *PLoS One,* 8(4), pp. e61465.

Miyata, R., Noda, N., Tamaki, H., Kinjyo, K., Aoyagi, H., Uchiyama, H. & Tanaka, H., 2007. Influence of feed components on symbiotic bacterial community structure in the gut of the wood-feeding higher termite *Nasutitermes takasagoensis*. *Bioscience, Biotechnology, and Biochemistry,* 71(5), pp. 1244-1251.

Montagna, M., Gómez-Zurita, J., Giorgi, A., Epis, S., Lozzia, G. & Bandi, C., 2015a. Metamicrobiomics in herbivore beetles of the genus *Cryptocephalus* (Chrysomelidae): toward the understanding of ecological determinants in insect symbiosis. *Insect Science,* 22(3), pp. 340-352.

Montagna, M., Chouaia, B., Mazza, G., Prosdocimi, E.M., Crotti, E., Mereghetti, V., et al., 2015b. Effects of the diet on the microbiota of the red palm weevil (Coleoptera: Dryophthoridae). *PLoS One,* 10(1), pp. e0117439.

Moran, N.A., 2015. Genomics of the honey bee microbiome. *Current Opinion in Insect Science,* 10, pp. 22-28.

Nardi, J.B., Mackie, R.I. & Dawson, J.O., 2002. Could microbial symbionts of arthropod guts contribute significantly to nitrogen fixation in terrestrial ecosystems? *Journal of Insect Physiology,* 48(8), pp. 751-763.

Nguyen, N.H., Suh, S.-O. & Blackwell, M., 2007. Five novel *Candida* species in insect-associated yeast clades isolated from Neuroptera and other insects. *Mycologia,* 99(6), pp. 842-858.

Nichols, E., Spector, S., Louzada, J., Larsen, T., Amezquita, S., Favila, M. & Network, T.S.R., 2008. Ecological functions and ecosystem services provided by Scarabaeinae dung beetles. *Biological Conservation,* 141(6), pp. 1461-1474.

Ohkuma, M., 2003. Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Applied Microbiology and Biotechnology,* 61(1), pp. 1-9.

Osei-Poku, J., Mbogo, C.M., Palmer, W.J. & Jiggins, F.M., 2012. Deep sequencing reveals extensive variation in the gut microbiota of wild mosquitoes from Kenya. *Molecular Ecology,* 21(20), pp. 5138-5150.

Pandey, N. & Rajagopal, R., 2016. Molecular characterization and diversity analysis of bacterial communities associated with *Dialeurolonga malleswaramensis* (Hemiptera: Aleyrodidae) adults using 16S rDNA amplicon pyrosequencing and FISH. *Insect Science,* 5(23), pp. 704-711.

Park, D., Oh, H., Jeong, W., Kim, H., Park, H. & Bae, K.S., 2007. A culture-based study of the bacterial communities within the guts of nine longicorn beetle species and their exo-enzyme producing properties for degrading xylan and pectin. *Journal of Microbiology-Seoul,* 45(5), pp. 394-401.

Poulsen, M., 2015. Towards an integrated understanding of the consequences of fungus domestication on the fungus-growing termite gut microbiota. *Environmental Microbiology,* 17(8), pp. 2562-2572.

Pramono, A.K., Sakamoto, M., Iino, T., Hongoh, Y. & Ohkuma, M., 2015. *Dysgonomonas termitidis* sp. nov., isolated from the gut of the subterranean termite *Reticulitermes speratus*. *International Journal of Systematic and Evolutionary Microbiology,* 65(2), pp. 681-685.

Reid, N.M., Addison, S.L., Macdonald, L.J. & Lloyd-Jones, G., 2011. Biodiversity of active and inactive bacteria in the gut flora of wood-feeding huhu beetle larvae (*Prionoplus reticularis*). *Applied and Environmental Microbiology,* 77(19), pp. 7000-7006.

Sabree, Z.L. & Moran, N.A., 2014. Host-specific assemblages typify gut microbial communities of related insect species. *SpringerPlus,* 3(1), pp. 138.

Santana, R.H., Catão, E.C.P., Lopes, F.A.C., Constantino, R., Barreto, C.C. & Krüger, R.H., 2015. The gut microbiota of workers of the litter-feeding termite *Syntermes wheeleri* (Termitidae: Syntermitinae): archaeal, bacterial, and fungal communities. *Microbial Ecology,* 70, pp. 545-556.

Santo Domingo, J.W., Kaufman, M.G., Klug, M.J., Holben, W.E., Harris, D. & Tiedje, J.M., 1998. Influence of diet on the structure and function of the bacterial hindgut community of crickets. *Molecular Ecology,* 7(6), pp. 761-767.

Schauer, C., Thompson, C.L. & Brune, A., 2014. Pyrotag sequencing of the gut microbiota of the cockroach *Shelfordella lateralis* reveals a highly dynamic core but only limited effects of diet on community structure. *PLoS One,* 9(1), pp. e85861.

Scholtz, C.H., 1989. Unique foraging behaviour in *Pachysoma*. *Journal of Arid Environments,* 16(3), pp. 305-313.

Scully, E.D., Geib, S.M., Hoover, K., Tien, M., Tringe, S.G., Barry, K.W., et al., 2013. Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. *PLoS One,* 8(9), pp. e73827.

Shao, M.-W., Lu, Y.-H., Miao, S., Zhang, Y., Chen, T.-T. & Zhang, Y.-L., 2015. Diversity, bacterial symbionts and antibacterial potential of gut-associated fungi isolated from the *Pantala flavescens* larvae in China. *PLoS One,* 10(7), pp. e0134542.

Sole, C.L., Scholtz, C.H. & Bastos, A.D.S., 2005. Phylogeography of the Namib Desert dung beetles *Scarabaeus* (*Pachysoma*) MacLeay (Coleoptera: Scarabaeidae). *Journal of Biogeography,* 32(1), pp. 75-84.

Suh, S.-O., McHugh, J.V. & Blackwell, M., 2004. Expansion of the *Candida tanzawaensis* yeast clade: 16 novel *Candida* species from basidiocarp-feeding beetles. *International Journal of Systematic and Evolutionary Microbiology,* 54(6), pp. 2409-2429.

Suh, S.-O., McHugh, J.V., Pollock, D.D. & Blackwell, M., 2005. The beetle gut: a hyperdiverse source of novel yeasts. *Mycological Research,* 109(3), pp. 261-265.

Suh, S.O., Nguyen, N.H. & Blackwell, M., 2008. Yeasts isolated from plant-associated beetles and other insects: seven novel *Candida* species near *Candida albicans*. *FEMS Yeast Research,* 8(1), pp. 88-102.

Tagliavia, M., Messina, E., Manachini, B., Cappello, S. & Quatrini, P., 2014. The gut microbiota of larvae of *Rhynchophorus ferrugineus* Oliver (Coleoptera: Curculionidae). *BMC Microbiology,* 14(1), pp. 136.

Tsuchida, T., Koga, R., Horikawa, M., Tsunoda, T., Maoka, T., Matsumoto, S., et al., 2010. Symbiotic bacterium modifies aphid body color. *Science,* 330(6007), pp. 1102-1104.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., et al., 2009. A core gut microbiome in obese and lean twins. *Nature,* 457(7228), pp. 480-484.

Vasanthakumar, A., Handelsman, J., Schloss, P.D., Bauer, L.S. & Raffa, K.F., 2008. Gut microbiota of an invasive subcortical beetle, *Agrilus planipennis* Fairmaire, across various life stages. *Environmental Entomology,* 37(5), pp. 1344-1353.

Waite, D.W., Dsouza, M., Biswas, K., Ward, D.F., Deines, P. & Taylor, M.W., 2015. Microbial community structure in the gut of the New Zealand insect Auckland tree weta (*Hemideina thoracica*). *Archives of Microbiology,* 197(4), pp. 603-612.

Wang, A., Yao, Z., Zheng, W. & Zhang, H., 2014. Bacterial communities in the gut and reproductive organs of *Bactrocera minax* (Diptera: Tephritidae) based on 454 pyrosequencing. *PLoS One,* 9(9), pp. e106988.

Wang, S., Ghosh, A.K., Bongio, N., Stebbings, K.A., Lampe, D.J. & Jacobs-Lorena, M., 2012. Fighting malaria with engineered symbiotic bacteria from vector mosquitoes. *Proceedings of the National Academy of Sciences,* 109(31), pp. 12734-12739.

Wang, Y., Gilbreath III, T.M., Kukutla, P., Yan, G. & Xu, J., 2011. Dynamic gut microbiome across life history of the malaria mosquito *Anopheles gambiae* in Kenya. *PLoS One,* 6(9), pp. e24767.

Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., et al., 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature,* 450(7169), pp. 560-565.

Wong, A.C.-N., Chaston, J.M. & Douglas, A.E., 2013. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *The ISME Journal,* 7(10), pp. 1922-1932.

Yang, Y.-J., Zhang, N., Ji, S.-Q., Lan, X., Shen, Y.-L., Li, F.-L. & Ni, J.-F., 2014. *Dysgonomonas macrotermitis* sp. nov., isolated from the hindgut of a fungus-growing termite. *International Journal of Systematic and Evolutionary Microbiology,* 64(9), pp. 2956-2961.

Yun, J.-H., Roh, S.W., Whon, T.W., Jung, M.-J., Kim, M.-S., Park, D.-S., et al., 2014. Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied and Environmental Microbiology,* 80(17), pp. 5254-5264.

Zhang, M., Liu, N., Qian, C., Wang, Q., Wang, Q., Long, Y., et al., 2014. Phylogenetic and functional analysis of gut microbiota of a fungus-growing higher termite: Bacteroidetes from higher termites are a rich source of β-glucosidase genes. *Microbial Ecology,* 68(2), pp. 416-425.

Zhang, N., Suh, S.-O. & Blackwell, M., 2003. Microorganisms in the gut of beetles: evidence from molecular cloning. *Journal of Invertebrate Pathology,* 84(3), pp. 226-233.

Zouache, K., Michelland, R.J., Failloux, A.-B., Grundmann, G.L. & Mavingui, P., 2012. Chikungunya virus impacts the diversity of symbiotic bacteria in mosquito vector. *Molecular Ecology,* 21(9), pp. 2297-2309.

Chapter 4: Shotgun Metagenomic Sequencing of the Gut Microbiomes of *Pachysoma endroedyi* and *Pachysoma striatum*

4.1 Introduction

A variety of methods have been used to study the insect gut microbiome including culturing, clone libraries, amplicon sequencing and most recently shotgun sequencing of the entire gut metagenome (Suh et al., 2004a, Ahn et al., 2012, Bertino-Grimaldi et al., 2013, He et al., 2013, Liu et al., 2013, Scully et al., 2013, Santana et al., 2015). Until recently, newer sequencing technologies were priced outside the range acceptable for most studies. However, as costs decrease, shotgun sequencing is becoming a favourable method for sequencing whole microbial communities in a variety of habitats (Venter et al., 2004, Adriaenssens et al., 2015, Mendes et al., 2015, Adriaenssens et al., 2016, Vikram et al., 2016). Shotgun sequencing has also been utilised for studying the gut microbiomes of well-studied taxonomic groups of insects including honeybees (Engel et al., 2012), mosquitoes (Chandler et al., 2015), termites (He et al., 2013), ticks (Carpi et al., 2011) and wood-feeding beetles (Jia et al., 2013, Scully et al., 2013).

Shotgun metagenomic sequencing allows access to the entire gut microbiome i.e. both microbial diversity and metabolic potential (Sankar et al., 2015). As a result, it is possible to study the set of functional genes and pathways that can assist a host in biological processes, with enzymes involved in digestion of recalcitrant carbohydrates (e.g., lignocellulosic compounds) important in industry and ecology alike (He et al., 2013, Scully et al., 2013)). However, a disadvantage of shotgun sequencing the insect gut microbiome is the large majority of resulting sequences being of host origin (Carpi et al., 2011, Hunter et al., 2011, Feehery et al., 2013). To overcome this challenge, the majority of studies use targeted

amplicon sequencing. Even studies which do use shotgun sequencing typically only report functional data, using amplicon sequencing for taxonomic analysis (He et al., 2013, Scully et al., 2013). However, shotgun sequencing overcomes primer bias introduced in amplicon sequencing and offers description of the entire gut microbial composition not just particular groups (e.g. bacteria or fungi) (Hodkinson and Grice, 2015).

This is the first study investigating the gut microbiome, and associated metabolic potentials, of a scarab beetle using this high throughput technology. In this chapter, shotgun sequencing was used to determine the taxonomic composition and functional capacity of the *Pachysoma* gut microbiome of two species feeding on substrates with different chemical and material compositions. Therefore, this chapter aims to answer the following main research questions:

1. What is the taxonomic composition of the *Pachysoma* gut metagenome and how does microbial community structure differ between the two *Pachysoma* spp. in comparison with the amplicon study.
2. Are there notable differences in the functional capacity of the two *Pachysoma* spp. microbiomes and their associated diets?
3. What is the capacity of microbial communities to digest refractive polycarbohydrates and how does it compare between the two different diets (plant detritus versus dry dung)?

4.2 Results and Discussion

4.2.1 Comparison of Four Previously Published Host DNA Reduction Methods.

Shotgun sequencing randomly fragments and sequences “everything” within a DNA sample (Auburn et al., 2011). However, mDNA samples from the insect gut contain large concentrations of DNA from the insect host (Hunter et al., 2011, Feehery et al., 2013). This so-called host DNA contamination is thus problematic when studying gut microbiomes, with the majority of reads being of host origin (Carpi et al., 2011, Hunter et al., 2011, Feehery et

al., 2013). For example, it has been observed when studying insect gut microbiomes that over 80% of the shotgun reads can originate from the insect host (Carpi et al., 2011, Feehery et al., 2013, Jia et al., 2013). Therefore in order to fully utilise shotgun sequencing to its full potential, it is important to implement methods to reduce host DNA contamination.

Four previously published methods to reduce host DNA contamination were tested on cockroach samples (Table 4.1). Ideally, these methods must allow unbiased access to all microorganisms within the sample (Krishnan et al., 2014) while reducing the concentration of host DNA extracted.

Table 4.1: Overview of methods tested for efficiency in reducing host DNA contamination during mDNA extraction.

Method	No treatment (control)	Removal of Peritrophic Membrane	Filtration	Minimal homogenisation	Low-speed centrifugation
Description	Methods used for amplicon sequencing samples	Removal of insect-derived PM during dissection	Extensive homogenisation to effectively release microorganisms from gut wall followed by filtration	Minimal homogenisation to reduce amount of host DNA released	Low-speed centrifugation to separate microbial cells from eukaryotic cells
Host/source from original study	*Pachysoma* spp. (this study)	*Anoplophora glabripennis* gut (wood feeding beetle)	*Ips pini* gut (bark beetle)	Human ileal and colonic biopsies	*Macrotermes annandalei*gut (fungus-growing termites)
Dissection	Whole-gut dissected	PM removed before gut added to tube for homogenisation	Whole-gut dissected (as per control method)	Whole-gut dissected (as per control method)	Whole-gut dissected (as per control method)
Homogenization	Add CTAB and homogenise using a tissue grinder	Add CTAB and homogenise using a tissue grinder (as per control method)	Add CTAB; sonicate guts for 30s; macerate guts further with a epi-crusher and vortex at medium speed for 10s; filter contents through filter paper	Vortex at medium speed for 10s; centrifuge minimal cycle	Add PBS followed by multiple centrifugation cycles at 800xg
mDNA extraction	Typical phenol-chloroform method	Phenol-chloroform method (as per control method)	Phenol-chloroform method (as per control method)	Phenol-chloroform method (as per control method)	Phenol-chloroform method (as per control method)
Shortcomings	High concentrations of host DNA contamination	May be removing important microorganisms embedded in the PM	May remove a portion of fungal symbionts with host DNA	Homogenisation may not be efficient to release entire gut microbial community	May remove a portion of fungal symbionts with host DNA
Reference		Scully et al. (2013)	Delalibera et al. (2007)	Carbonero et al. (2011)	Liu et al. (2011)

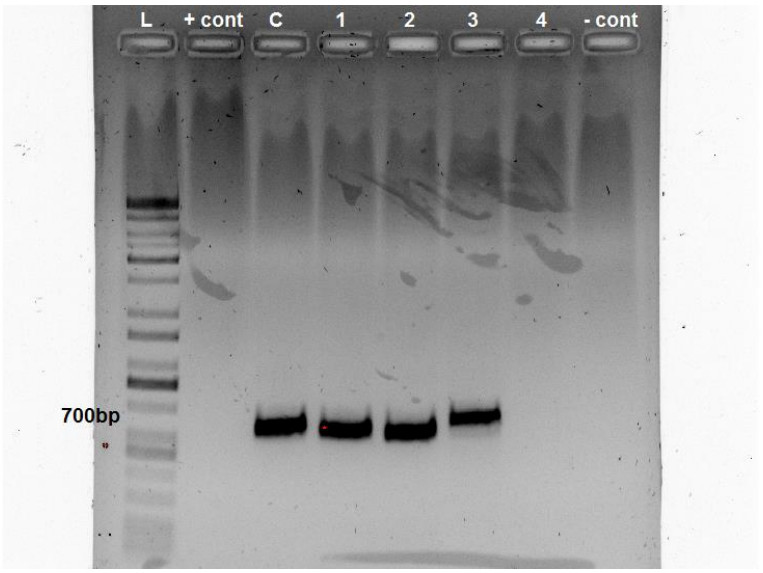
The four tested methods yielded high quality DNA with good purity values compared to that of the no treatment (control) method (Table 4.2). PCR amplification of the insect 18S rRNA

gene was carried out with insect specific primers (detailed in Chapter 2.4.2). Only the low-speed centrifugation method was unsuccessful in amplifying the insect 18S rRNA gene (Figure 4.1a). Therefore, it appears that this method may successfully reduce the concentration of host DNA. Furthermore, the low-speed centrifugation method yielded the lowest concentration of mDNA. The 16S rRNA gene was amplified from all five samples using bacteria specific primers (See Chapter 2.4.2), suggesting that bacterial DNA was still present in the samples (Figure 4.1b).

Table 4.2: Nanodrop readings of mDNA yields and purity ratios for samples treated with four host DNA reduction methods.

Sample ID	DNA concentration (ng/µl)	$A_{260/280}$	$A_{260/230}$
No treatment	3327.9	1.2	1.1
Removal of PM	1211.1	2.0	2.1
Filtration	471.8	1.9	1.7
Minimal homogenisation	60.3	1.8	1.8
Low speed centrifugation	16.3	1.9	2.2

a)



b)

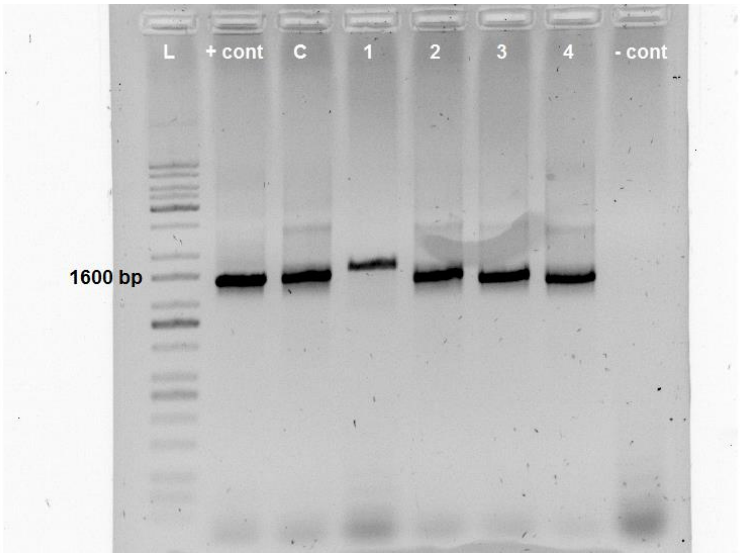


Figure 4.1: Amplification of the a) 18S rRNA and b) 16S rRNA genes from cockroach gut samples subjected to all four host DNA reduction methods. L: molecular weight KAPA Universal Ladder, C: no treatment, 1: removal of PM, 2: filtration, 3: minimal homogenisation and 4: low-speed centrifugation. *Escherichia coli* cells were used as the positive control, therefore, no amplification was expected in gel (a).

4.2.2 Shotgun Sequencing of the Gut Microbiomes of *P. endroedyi* and *P. striatum*

Metagenomic DNA extracted using the low-speed centrifugation method from the gut samples of three individuals of *P. endroedyi* and *P. striatum* (Table 4.3) were pooled in equal concentrations and sent to Molecular Research (MR DNA), for shotgun sequencing using the Illumina Hiseq platform (Section 2.4.3).

Table 4.3: Nanodrop readings of mDNA yields and purity values for replicate *P. endroedyi* and *P. striatum* gut samples sent for shotgun sequencing.

Species	Replicate	DNA concentration (ng/µl)	$A_{260/280}$	$A_{260/230}$
P. endroedyi	1	99.7	1.4	0.9
	2	171.5	1.7	1.2
	3	77.3	1.4	0.7
P. striatum	1	141.7	1.5	0.7
	2	45.9	1.6	1.0
	3	157.1	2.0	1.7

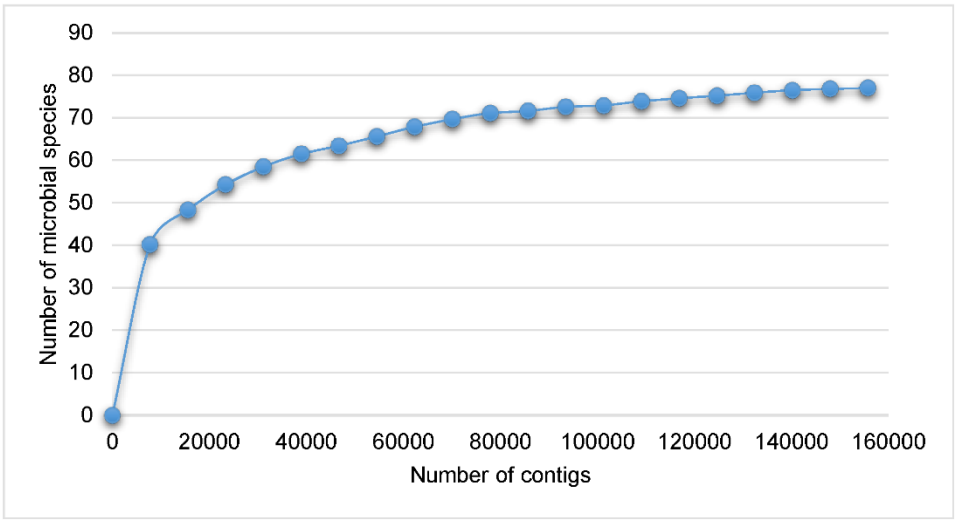
19 590 694 and 26 098 638 paired-end reads were recovered for *P. endroedyi* and *P. striatum*, respectively (Table 4.4). *De novo* assembly of high-quality reads using CLC genomic workbench resulted in 231,054 and 223,703 contigs. Both assemblies had N50 scores (the largest scaffold length where 50% of the genomic assembly is represented by scaffolds greater than this size (Kingsford et al., 2010)) of 1234 and 787 for *P. endroedyi* and *P. striatum*, respectively (Table 4.4). These N50 scores were comparable with those of another shotgun sequencing study on the insect gut microbiome of the Asian longhorned beetle, *Anoplophora glabripennis* (N50=938) (Scully et al., 2013), which along with the other metadata and metrics (Table 4.4) suggest that high quality assemblies were created. Rarefaction curves indicated successful sampling of the *P. endroedyi* gut microbiome (Figure 4.2a), suggesting that

complete microbial richness was analysed. Rarefaction curves for the *P. striatum* gut microbiome did not reach asymptote (Figure 4.2b), suggesting that the complete microbial richness may not be displayed.

Table 4.4: Sequencing metadata and assembly metrics for shotgun sequencing of the *P. endroedyi* and *P. striatum* gut microbiomes using CLC Genomics Workbench, MG-RAST and MEGAN.

Parameters	*P. endroedyi*	*P. striatum*
No. raw reads	19 590 694	26 098 638
No. reads after quality control	17 207 062	24 260 166
Average read length	132.63	131.96
Number of nucleotides	2 282 237 909	3 201 399 203
Number of contigs	231 054	223 703
N25	6 346	14 132
N50	1 234	787
N75	452	405
Number of unassembled reads	3 644 008	2 764 456
Assembled reads	13 563 054	21 495 710
Minimum contig length (bp)	82	76
Maximum contig length (bp)	220 532	256 410
Average contig length	799	707
Ribosomal RNA (%)	0.9	3.9
Annotated protein (%)	65.6	32.6
Unknown protein (%)	31.9	62.7
Number of ORFs	267 509	127 539
Number of ORFs assigned to COG (%)	101 542 (37.9)	44 762 (35.1)

a)



b)

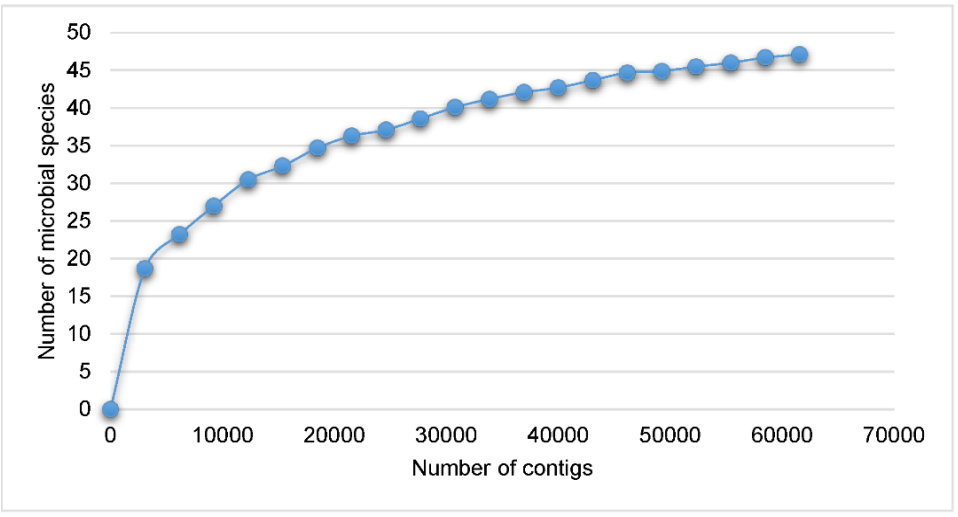


Figure 4.2: Rarefaction curves of gut microbial community richness of a) *P. endroedyi* and b) *P. striatum* gut samples. Eukaryotic contigs derived from Metazoa and Streptophyta were removed before analysis.

4.2.3 Community Composition of the *Pachysoma* Metagenome

Taxonomic annotation of contigs was done using both MG-RAST and MEGAN (Figure 4.3). The results of both analyses were comparable to each other with the exception of eukaryotes in *P. striatum*, where approximately twice the number of eukaryotic phylotypic signals in the gut of *P. striatum* were detected by MEGAN (22.1% contigs) compared to MG-RAST (11.2% contigs). However, in both analyses the majority of eukaryotic contigs were

associated to Metazoa (animals), which will be removed from further analysis. Therefore, the MEGAN results were used further for both taxonomic and functional annotation.

As expected, Bacteria was the most dominant domain in the guts of both *Pachysoma* species (Figure 4.3), which is typical for insect gut systems (Belda et al., 2011, Liu et al., 2013, Shi et al., 2013, Rahman, 2016). This result also illustrates that the low-speed centrifugation method does appear to successfully reduce host DNA contamination, as has been demonstrated previously (Liu et al., 2011). This is further collaborated by the fact that few or no contigs were associated with either *Pachysoma* species using both MG-RAST and MEGAN (MG-RAST: *P. endroedyi* = 0 contigs, *P. striatum* = 0 contigs; MEGAN: *P. endroedyi* = 0 contigs, *P. striatum* = 1 *Pachysoma* sp. contig). Furthermore, Eukaryotes were detected in lower percentages than recorded in studies where host DNA reduction methods were not used (Carpi et al., 2011, Jia et al., 2013). However, the number of Eukaryotic contigs was substantially higher in the coprophagous *P. striatum* (22.1% contigs) than the detritivorous *P. endroedyi* (0.1% contigs). Archaea and viruses were also detected in low abundances (<2% contigs) in both *Pachysoma* species.

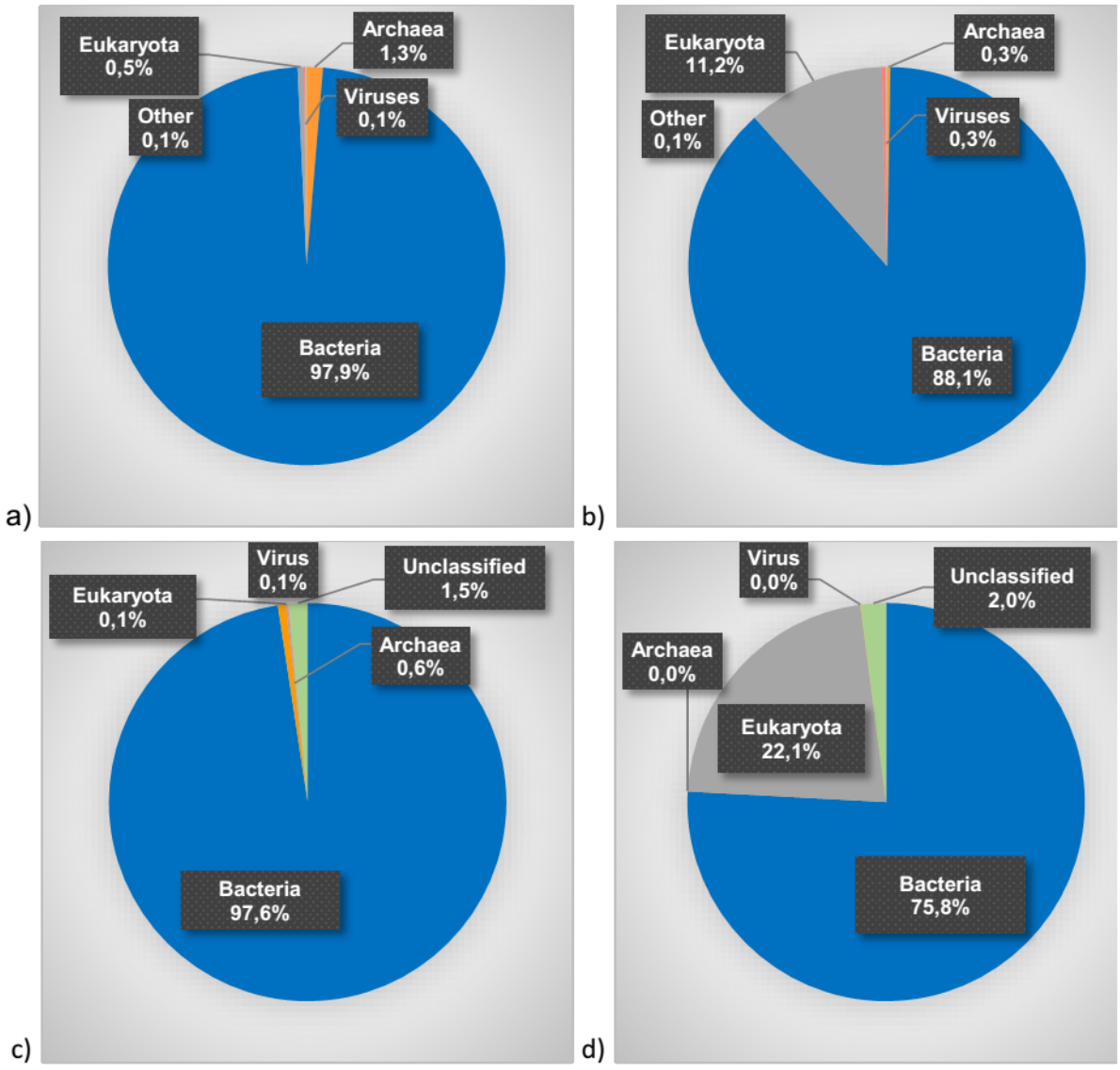


Figure 4.3: Taxonomic affiliation of shotgun sequencing reads assigned using MG-RAST (a and b), and MEGAN (c and d) for *P. endroedyi* (a and c) and *P. striatum* (b and d) gut metagenomes.

Shotgun sequencing revealed a greater bacterial diversity than detected using amplicon sequencing. A total 39 bacterial phyla, all of which were found in the guts of the plant detritus feeding *P. endroedyi*, while only 24 bacterial phyla were detected in the gut of the coprophagous *P. striatum.* Bacteria that could not be classified past domain were also present in both *Pachysoma* gut metagenomes. Although the diversity is greater than recorded in the

amplicon study where a total 14 bacterial phyla were detected (Chapter 3), both studies showed great differences in diversity between the two species which could be due to dietary differences as previously discussed (Chapter 3; Franzini et al. (2016)). Furthermore, in the shotgun sequencing datasets, four phyla were the dominant (>2% contigs) bacterial communities in the gut of both *Pachysoma* species, specifically Bacteroidetes (most dominant [50.6% of bacteria-derived contigs] in *P. endroedyi*), Firmicutes (most dominant [36.4% of bacteria-derived contigs] in *P. striatum*), Proteobacteria and Actinobacteria (Figure 4.4). These abundances may be representative of specific diets as already discussed in the amplicon study (Chapter 3). Bacteriodetes are typically highly abundant in plant feeding insects (Boucias et al., 2013, Schauer et al., 2014, Montagna et al., 2015, Waite et al., 2015), with Proteobacteria and/or Firmicutes most dominant in insects feeding on simpler diets (Ahn et al., 2012, Wang et al., 2014, Andongma et al., 2015, Pandey and Rajagopal, 2016). This corresponds with Bacteriodetes being most abundant in gut samples of the plant detritus feeding *P. endroedyi* and Firmicutes most abundant in *P. striatum*. The remaining phyla represented abundances of less than 2% of the contigs for both *Pachysoma* spp. This result was comparable to *P. striatum* amplicon sequencing results but varied from that for *P. endroedyi*, where seven phyla were considered abundant (>2% 16S rRNA sequence reads). The largest inconsistency was the difference in abundance of the bacterial phyla Planctomycetes and Elusimicrobia between the two sequencing technologies.

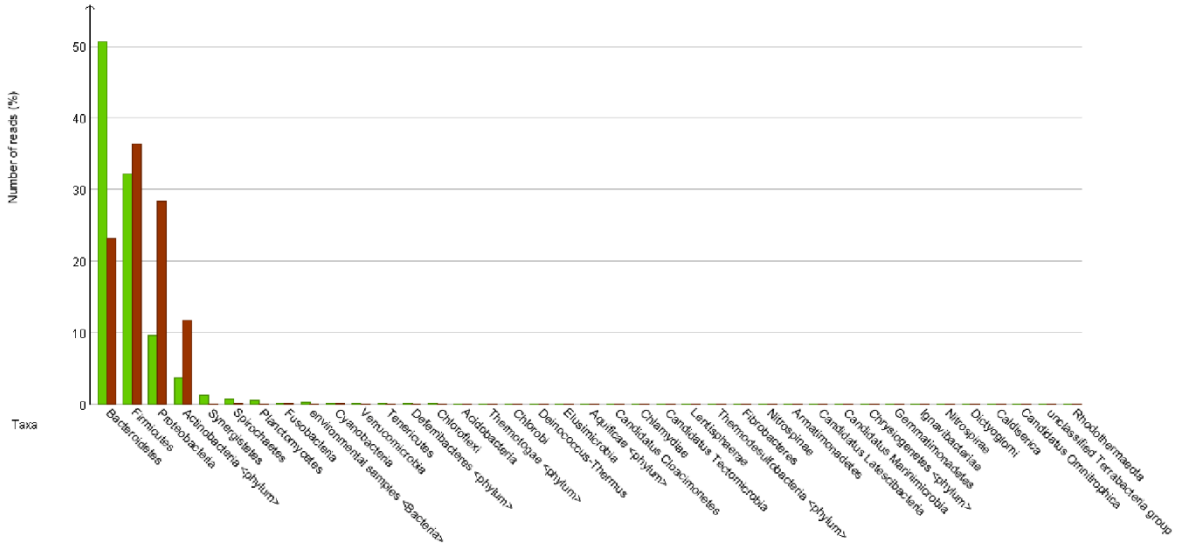


Figure 4.4: Diversity and abundance of bacterial phyla in the gut samples of *P. endroedyi* (green) and *P. striatum* (brown) using shotgun (Illumina Hiseq) analysed using MEGAN.

Metazoa (animals) was the most abundant Eukaryotic kingdom accounting for 58.9% and 99.1% of the eukaryotic-derived contigs for *P. endroedyi* and *P. striatum*, respectively. Possible causes for this high number of Metazoa-derived contigs in the *P. striatum* metagenome include, various sources contaminating the dung including the animal that excreted the faeces or as part of the original diet (Paxinos et al., 1997, Symondson, 2002, Jarman et al., 2004). The majority of Metazoa-derived contigs were from the class Insecta (84.5%) followed by Arachnida (1.0%), Chromadorea (1.0%) and Mammalia (0.9%).

A significant result, when comparing results of the 2 technologies, is the presence of fungal gut communities in both *Pachysoma* species. This is in contrast to the amplicon study where the ITS region could not be amplified for *P. endroedyi* gut samples. Furthermore, discrepancies were noted between the fungal results of the two technologies with regard to the detection of yeasts in the gut samples. True yeasts (Saccharomycetes) were detected using shotgun sequencing in low abundances in both *P. endroedyi* (2 [8.3%] fungal contigs) and *P. striatum* (1 [1.0%] fungal contig) gut samples, respectively but were absent from *P. striatum* samples in the amplicon study. Saccharomycetes are typically detected in insect guts

using culture-dependent methods and clone libraries (Suh and Blackwell, 2004, Suh et al., 2004b, Suh et al., 2004a, Suh et al., 2005, Nguyen et al., 2007).

4.2.4 Functional Profiling of the Metagenomes of *P. endroedyi* and *P. striatum*.

Microbial gut communities are considered to fully or partially assist the host in digestion of different components within the diet (Douglas, 2009). Therefore, the functional capacity of the gut microbial communities of *P. endroedyi* and *P. striatum* were determined after removal of all Metazoa and Streptophyta derived ORFs from each respective dataset.

Only 37.9% and 35.1% microbial-derived ORFs from *P. endroedyi* and *P. striatum* gut, respectively, could be classified into 22 functional categories against the Clusters of Orthologous Groups (COG) database (Figure 4.5). Genes from eight COG categories were overrepresented in both datasets namely; carbohydrate transport and metabolism (G; 10.8-12.4%), replication, recombination and repair (L; 9.1-11.1%), amino acid transport and metabolism (E; 10.3-10.7%), cell wall/membrane/envelope biogenesis (M; 8.7-9.2%), translation, ribosomal translation and biogenesis (J; 8.1-8.2%), energy production and conversion (C; 7.8-8.2%), Inorganic ion transport and metabolism (P; 7.1-8.0%) and transcription (K; 6.6-8.3%). The enrichment of such functional genes were not significantly different between the microbiomes of the two *Pachysoma* species ($p>0.05$). It appears that *Pachysoma* gut communities may be suitably adapted to support their own nutritional needs through enrichment of metabolic and cellular processing genes (Rahman, 2016). Furthermore, the enrichment of such genes may be linked to the content of plant biomass in the respective diets, particularly regarding genes involved in carbohydrate transport and metabolism (G), energy production and conversion (C) and cell wall/membrane/envelope biogenesis (M) potentially indicating microorganisms need to attach to plant cell walls (Suen et al., 2011). Functional categories absent from the annotated *Pachysoma* ORFs were nuclear structure (category Y) and the poorly categorized R (general function prediction only) and S (function unknown), which are commonly abundant in metagenomic datasets from insect gut samples

(Belda et al., 2011, Shi et al., 2013, Rahman, 2016). The absence of such genes could indicate a weakness in the capacity of the software to efficiently bin genes into COG categories. For example, it is noted that a large percentage of genes were not categorised (*P. endroedyi*: 37.9%, *P. striatum*=35.1%). As the COG categories R and S are bins for genes where the function cannot be clearly identified, it is possible that such genes may have been classed as 'unidentified' and would therefore not be included in the accepted COG classifications (Figure 4.5).

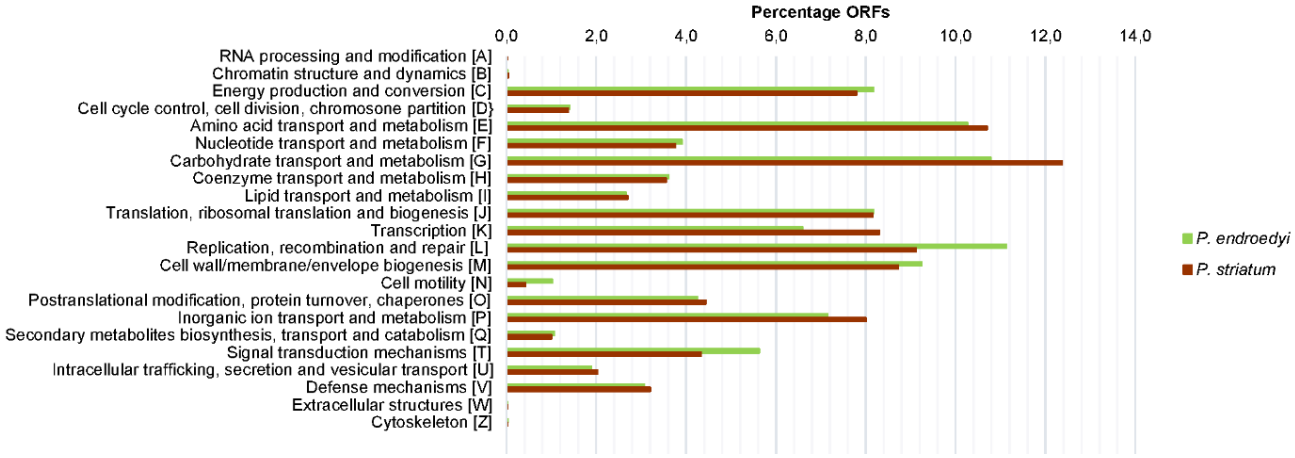


Figure 4.5: Functional assignments of microbial-derived ORFs from both *Pachysoma* species to categories of COGs generated by MEGAN.

4.2.4.1 Candidate Genes for Carbohydrate Transport and Metabolism

The focus of this study is the effect of host diet on microbiome assembly and the associated role of the respective microorganisms in digestion of that particular diet. As plant biomass is present in both diets (Holter and Scholtz, 2011, 2013), genes involved in carbohydrate transport and metabolism are of interest in this study. The diet of *P. endroedyi* consists basically of undigested plant fragments (Harrison et al., 2003) that may have been subjected to previous degradation by environmental microorganisms (microbial conditioning; Swift et al. (1979)). In contrast, the coprophagous diet of *P. striatum* contains smaller plant

fragments that have already been subjected to one round of digestion/mastication (Anderson and Coe, 1974) along with associated microbial conditioning. It could be expected that the composition of plant polymers would differ between the two diets due to the faecal plant matter having been previously digested. However, it is unknown whether or not faecal plant matter has been partially degraded or not. Furthermore, without knowing the species of plant in the two diets it would be speculative to suggest that different functional capacities would be needed to digest the respective plant biomass.

Genes encoding glycoside hydrolases (GHs; enzymes which hydrolyse glycosidic bonds between carbohydrates (http://www.cazy.org/; Lombard et al. (2014)) were classified into 51 and 53 families for *P. endroedyi* and *P. striatum*, respectively, based on combined Pfam domain and KEGG enzyme class (E.C.) assignments (Figure 4.6). The GH profiles were similar between the two species with the hemicellulose and cellulose GH2 (*P. endroedyi*=21.1%; *P. striatum*=15.4%) and GH3 (*P. endroedyi*=9.0%; *P. striatum*=9.8%) CAZyme families being the most abundant for the gut microbiomes of both *Pachysoma* species. Furthermore, abundances were mostly similar between the GH families of the two *Pachysoma* species, with the exception of GH1 (*P. endroedyi* = 9.1% ORFs, *P. striatum* = 1.5% ORfs; Figure 4.6). Genes encoding GHs from family one are typically cellulases and hemicellulases. Therefore, it is interesting that the coprophagous *P. striatum* has a substantially larger, albeit not statistically significant ($p>0.05$), abundance of genes in this family (7.55% ORFs). Plant fragments found in dung have previously undergone digestion/mastication by the primary animal (Anderson and Coe, 1974). Therefore, the large difference in abundance of GH1 enzymes suggests they may be necessary to degrade plant fragments previously subjected to digestion and expelled in the dung. Eight families were reported in only one of the two *Pachysoma* species, in low abundances (<6 ORFs). The associated KEGG E.C. assignments of all GH families are listed in Table 4.5.

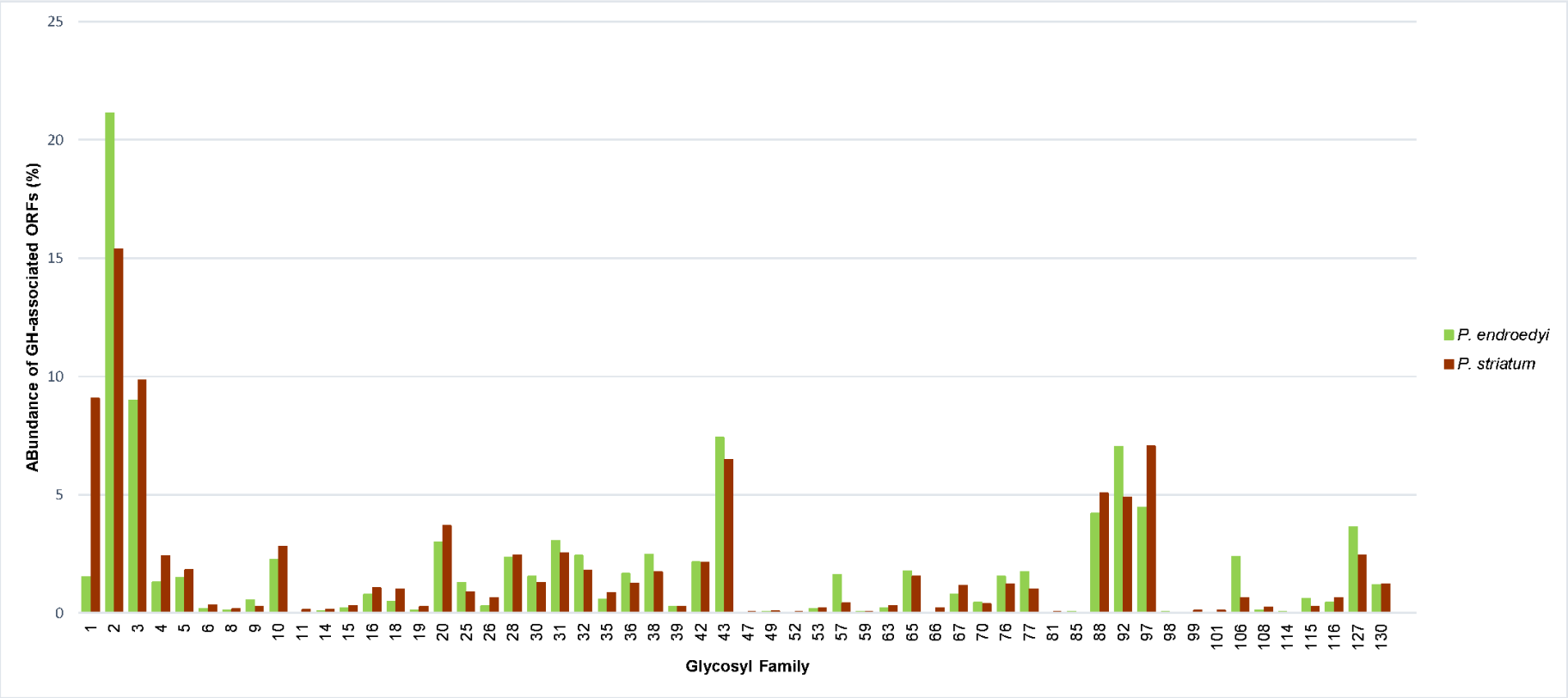


Figure 4.6: Distribution of glycosyl hydrolase (GH) families in the gut metagenomes of *P. endroedyi* and *P. striatum* annotated using hmmer.

Table 4.5: Glycosyl hydrolase families detected in the gut metagenomes of *P. endroedyi* and *P. striatum* and the associated KEGG E.C. assignments were applicable.

Family	*P. endroedyi* ORFs	*P. striatum* ORFS	Pfam HMM Name	Pfam accession	KEGG ID	KEGG functions	KEGG ECs	*P. endroedyi* ORFs	*P. striatum* ORFS
1	74	269	Glyco_hydro_1	PF00232.16	K01223	6-phospho-beta-glucosidase	[EC:3.2.1.86]	56	248
					K05350	beta-glucosidase	[EC:3.2.1.21]	5	0
					K01220	6-phospho-beta-galactosidase	[EC:3.2.1.85]	1	2
2	1034	456	Glyco_hydro_2_C	PF02836.15	K01190	beta-galactosidase	[EC:3.2.1.23]	174	73
					K01195	beta-glucuronidase	[EC:3.2.1.31]	3	5
					K01192	beta-mannosidase	[EC:3.2.1.25]	8	9
					K18577	mannosylglycoprotein endo-beta-mannosidase	[EC:3.2.1.152]	5	2
					K12308	beta-galactosidase	[EC:3.2.1.23]	254	0
					K01195	beta-glucuronidase	[EC:3.2.1.31]	7	6
					K01192	beta-mannosidase	[EC:3.2.1.25]	3	4
					K01190	beta-galactosidase	[EC:3.2.1.23]	0	85
					K18577	mannosylglycoprotein endo-beta-mannosidase	[EC:3.2.1.152]	4	2
			Glyco_hydro_2	PF00703.19	K01190	beta-galactosidase	[EC:3.2.1.23]	188	74
					K01192	beta-mannosidase	[EC:3.2.1.25]	5	2
					K01195	beta-glucuronidase	[EC:3.2.1.31]	5	6
					K18577	mannosylglycoprotein endo-beta-mannosidase	[EC:3.2.1.152]	3	1
3	439	292	Glyco_hydro_3	PF00933.19	K05349	beta-glucosidase	[EC:3.2.1.21]	176	127
					K01207	beta-N-acetylhexosaminidase	[EC:3.2.1.52]	12	11
			Glyco_hydro_3_C	PF01915.20	K05349	beta-glucosidase	[EC:3.2.1.21]	153	118
					K15920	beta-D-xylosidase 5	[EC:3.2.1.37]	2	2
4	63	71	Glyco_hydro_4	PF02056.14	K01222	6-phospho-beta-glucosidase	[EC:3.2.1.86]	6	18
					K07406	alpha-galactosidase	[EC:3.2.1.22]	14	2
					K01232	maltose-6'-phosphate glucosidase	[EC:3.2.1.122]	1	6
			Glyco_hydro_4C	PF11975.6	K01222	6-phospho-beta-glucosidase	[EC:3.2.1.86]	2	18
					K07406	alpha-galactosidase	[EC:3.2.1.22]	16	5
					K01232	maltose-6'-phosphate glucosidase	[EC:3.2.1.122]	2	6

5	73	54	Cellulase	PF00150.16	K01179	endoglucanase	[EC:3.2.1.4]	16	9
					K01210	glucan 1,3-beta-glucosidase	[EC:3.2.1.58]	0	2
5	73	54	Cellulase	PF00150.16	K19355	mannan endo-1,4-beta-mannosidase	[EC:3.2.1.78]	3	1
6	8		GHL6	PF14871.4		No EC evidence			
		10	Glyco_hydro_6	PF01341.15	K01179	endoglucanase	[EC:3.2.1.4]	0	1
8	5	5	Glyco_hydro_8	PF01270.15	K01179	endoglucanase	[EC:3.2.1.4]	0	3
9	27	8	Glyco_hydro_9	PF00759.17	K01179	endoglucanase	[EC:3.2.1.4]	4	2
10	110	83	GHL10	PF02638.13		No EC evidence			
			Glyco_hydro_10	PF00331.18					
11	0	4	Glyco_hydro_11	PF00457.15	K01181	endo-1,4-beta-xylanase	[EC:3.2.1.8]	0	1
14*	4	4	Glyco_hydro_14	PF01373.15		Beta-Amylase	[EC 3.2.1.2]	4	0
15	10	9	Glyco_hydro_15	PF00723.19		No EC evidence			
16	38	31	Glyco_hydro_16	PF00722.19		No EC evidence			
18	23	30	Glyco_hydro_18	PF00704.26	K06306	spore germination protein		3	1
					K01183	chitinase	[EC:3.2.1.14]	5	10
					K17523	chitinase-3-like protein 1/2		0	1
19	5	8	Glyco_hydro_19	PF00182.17	K03791	putative chitinase		2	2
					K01183	chitinase	[EC:3.2.1.14]	1	0
20	146	109	Glyco_hydro_20	PF00728.20	K12373	hexosaminidase	[EC:3.2.1.52]	78	56
			Glyco_hydro_20b	PF02838.13	K12373	hexosaminidase	[EC:3.2.1.52]	0	37
					K01197	hyaluronoglucosaminidase	[EC:3.2.1.35]	0	1
25	62	26	Glyco_hydro_25	PF01183.18	K07273	lysozyme	[EC 3.2.1.17]	25	16
					K06381	spoIID; stage II sporulation protein D		1	0
26	14	19	Glyco_hydro_26	PF02156.13	K01218	mannan endo-1,4-beta-mannosidase	[EC:3.2.1.78]	7	10
					K07214	enterochelin esterase and related enzymes		0	1
28	114	72	Glyco_hydro_28	PF00295.15	K18650	exo-poly-alpha-galacturonosidase	[EC:3.2.1.82]	0	1
30	75	38	Glyco_hydro_30	PF02055.14	K01201	glucosylceramidase	[EC:3.2.1.45]	26	14
					K15924	glucuronoarabinoxylan endo-1,4-beta-xylanase	[EC:3.2.1.136]	0	1
			Glyco_hydr_30_2;	PF14587.4	K01201	glucosylceramidase	[EC:3.2.1.45]	0	5
			Glyco_hydro_30C	PF17189.2	K01201	glucosylceramidase	[EC:3.2.1.45]	10	14
31	149	75	Glyco_hydro_31	PF01055.24	K01811	alpha-D-xyloside xylohydrolase	[EC:3.2.1.177]	66	33
					K01187	alpha-glucosidase	[EC:3.2.1.20]	20	18
					K05546	alpha 1,3-glucosidase	[EC:3.2.1.84]	2	1
					K07407	alpha-galactosidase	[EC:3.2.1.22]	1	2
32	118	53	Glyco_hydro_32C	PF08244.10	K03332	fructan beta-fructosidase	[EC:3.2.1.80]	4	2
					K01193	beta-fructofuranosidase	[EC:3.2.1.26]	1	0

			Glyco_hydro_32N	PF00251.18	K01193	beta-fructofuranosidase	[EC:3.2.1.26]	6	4
32	118	53	Glyco_hydro_32N	PF00251.18	K03332	fructan beta-fructosidase	[EC:3.2.1.80]	16	3
35	28	25	Glyco_hydro_35	PF01301.17	K12308	beta-galactosidase	[EC:3.2.1.23]	8	17
36	81	37	Glyco_hydro_36C	PF16874.3	K07407	alpha-galactosidase	[EC:3.2.1.22]	14	11
			Glyco_hydro_36N	PF16875.3	K07407	alpha-galactosidase	[EC:3.2.1.22]	20	11
38	121	51	Glyco_hydro_38	PF01074.20	K01191	alpha-mannosidase	[EC:3.2.1.24]	32	14
					K15524	mannosylglycerate hydrolase	[EC:3.2.1.170]	3	9
			Glyco_hydro_38C	PF07748.11	K01191	alpha-mannosidase	[EC:3.2.1.24]	33	9
					K15524	mannosylglycerate hydrolase	[EC:3.2.1.170]	0	6
39	13	8	Glyco_hydro_39	PF01229.15		No EC evidence			
42	105	63	Glyco_hydro_42	PF02449.13	K12308	beta-galactosidase	[EC:3.2.1.23]	30	24
			Glyco_hydro_42C	PF08533.8	K12308	beta-galactosidase	[EC:3.2.1.23]	2	2
			Glyco_hydro_42M	PF08532.8	K12308	beta-galactosidase	[EC:3.2.1.23]	9	4
43	363	192	GH43_C	PF16369.3	K06113	arabinan endo-1,5-alpha-L-arabinosidase	[EC:3.2.1.99]	2	5
			Glyco_hydro_43	PF04616.12	K01198	xylan 1,4-beta-xylosidase	[EC:3.2.1.37]	15	20
					K06113	arabinan endo-1,5-alpha-L-arabinosidase	[EC:3.2.1.99]	41	17
					K15921	arabinoxylan arabinofuranohydrolase	[EC:3.2.1.55]	5	2
					K03544	ATP-dependent Clp protease ATP-binding subunit ClpX		1	0
47	1	1	Glyco_hydro_47	PF01532.18	K01230	mannosyl-oligosaccharide alpha-1,2-mannosidase	[EC:3.2.1.113]	1	0
						alpha-mannosidase	[EC 3.2.1.113]	0	1
49	3	2	Glyco_hydro_49	PF03718.11		No EC evidence			
52*	1	1	Glyco_hydro_52	PF03512.11		beta-xylosidase	[EC 3.2.1.37]	1	1
53	8	6	Glyco_hydro_53	PF07745.11	K01224	arabinogalactan endo-1,4-beta-galactosidase	[EC:3.2.1.89]	6	6
57	79	12	Glyco_hydro_57	PF03065.13	K07405	alpha-amylase	[EC:3.2.1.1]	46	10
					K16149	1,4-alpha-glucan branching enzyme	[EC:2.4.1.18]	6	0
59	3	1	Glyco_hydro_59	PF02057.13	K01202	galactosylceramidase	[EC:3.2.1.46]	1	0
63	10	9	Glyco_hydro_63	PF03200.14	K01228	mannosyl-oligosaccharide glucosidase	[EC:3.2.1.106]	1	0
						putative isomerase		0	3
65	86	46	Glyco_hydro_65C	PF03633.13	K00691	maltose phosphorylase	[EC:2.4.1.8]	5	2
			Glyco_hydro_65M	PF03632.13	K00691	maltose phosphorylase	[EC:2.4.1.8]	11	5
			Glyco_hydro_65N	PF03636.13	K00691	maltose phosphorylase	[EC:2.4.1.8]	6	3
66	0	6	Glyco_hydro_66	PF13199.4	K05988	dextranase	[EC:3.2.1.11]	0	6
67	39	34	Glyco_hydro_67C	PF07477.10	K01235	alpha-glucuronidase	[EC:3.2.1.139]	11	9
			Glyco_hydro_67M	PF07488.10	K01235	alpha-glucuronidase	[EC:3.2.1.139]	16	12
			Glyco_hydro_67N	PF03648.12	K01235	alpha-glucuronidase	[EC:3.2.1.139]	3	6

70	21	11	Glyco_hydro_70	PF02324.14		No EC evidence			
76	75	36	Glyco_hydro_76	PF03663.12	K08257	mannan endo-1,6-alpha-mannosidase	[EC:3.2.1.101]	1	0
76	75	36				uncharacterized protein		2	0
			Glyco_hydro_76	PF03663.12		alpha-1,6-mannanase	[EC 3.2.1.101]	0	36
77	85	30	Glyco_hydro_77	PF02446.15	K00705	4-alpha-glucanotransferase	[EC:2.4.1.25]	74	28
81	0	1	Glyco_hydro_81	PF03639.11	K01180	endo-1,3(4)-beta-glucanase	[EC:3.2.1.6]	0	1
85*	2	0	Glyco_hydro_85	PF03644.11		endo-β-N-acetylglucosaminidase	[EC 3.2.1.96]	2	0
88	206	150	Glyco_hydro_88	PF07470.11	K15532	unsaturated rhamnogalacturonyl hydrolase	[EC:3.2.1.172]	46	35
					K18581	unsaturated chondroitin disaccharide hydrolase	[EC:3.2.1.180]	6	8
					K09955	uncharacterized protein		2	0
					K06888	uncharacterized protein		0	1
92	344	145	Glyco_hydro_92	PF07971.10		No EC evidence			
97	218	209	Glyco_hydro_97	PF10566.7	K01187	alpha-glucosidase	[EC:3.2.1.20]	63	47
			GH97_C	PF14509.4	K01187	alpha-glucosidase	[EC:3.2.1.20]	41	37
			GH97_N	PF14508.4	K01187	alpha-glucosidase	[EC:3.2.1.20]	55	43
98	2	0	Glyco_hydro_98C	PF08307.9		No EC evidence			
			Glyco_hydro_98M	PF08306.9					
99	0	3	Glyco_hydro_99	PF16317.3		No EC evidence		0	0
101	0	3	Glyco_hydro_101	PF12905.5		endo-α-N-acetylgalactosaminidase	[EC 3.2.1.97]	0	3
106	116	19	Glyco_hydro_106	PF17132.2	K02856	L-rhamnose-H+ transport protein		1	
						alpha-L-rhamnosidase	[EC 3.2.1.40]	0	19
108*	6	7	Glyco_hydro_108	PF05838.10		N-acetylmuramidase	[EC 3.2.1.17]	6	7
114*	2	0	Glyco_hydro_114	PF03537.11		endo-α-1,4-polygalactosaminidase	[EC 3.2.1.109]	2	0
115	29	8	Glyco_hydro_115	PF15979.3		No EC evidence			
116	22	19	Glyco_hydr_116N	PF12215.6	K17108	non-lysosomal glucosylceramidase	[EC:3.2.1.45]	0	1
			DUF608	PF04685.11	K17108	non-lysosomal glucosylceramidase	[EC:3.2.1.45]	0	1
127	177	72	Glyco_hydro_127	PF07944.10	K09955	uncharacterized protein		97	47
130	58	36	Glyco_hydro_130	PF04041.11	K18785	beta-1,4-mannooligosaccharide/beta-1,4-mannosyl-N-acetylglucosamine phosphorylase	[EC:2.4.1.319/ 2.4.1.320]	14	15
					K16212	4-O-beta-D-mannosyl-D-glucose phosphorylase	[EC:2.4.1.281]	8	8
Glycosyl hydrolase catalytic core	10	4	Glyco_hydro_cc	PF11790.6	K15924	glucuronoarabinoxylan endo-1,4-beta-xylanase	[EC:3.2.1.136]	0	1

Given the large differences in bacterial diversity, it is surprising that functional diversity of GH families was similar between the two *Pachysoma* microbiomes. Of particular interest are enzymes involved in the breakdown of the major components of plant cell walls, namely cellulose, hemicellulose and other plant polysaccharides. Cellulases were classified into six GH families, namely GH1, GH3, GH5, GH6 (only *P. striatum*), GH8 (only *P. striatum*) and GH9. Their corresponding KEGG E.C. assignments (assigned using KAAS) suggest that only two classes of cellulases were present in both *Pachysoma* spp., namely endoglucanases, and β-glucosidases. Endoglucanases cleave β-1-4-glucosidic bonds while β-glucosidases hydrolyses cellobiose to glucose (Terra and Ferreira, 1994, Lynd et al., 2002). No ORFs were assigned to the third class of enzymatic activity, the exoglucanases (E.C. 3.2.1.74; 3.2.1.91), which cleave cellobiose or glucose from cellulose polysaccharide chains (Terra and Ferreira, 1994, Lynd et al., 2002). Although microbial exoglucanases have been recorded from the gut of the Asian longhorned beetle (Geib et al., 2009, Scully et al., 2013), such enzymes have not been associated with higher termites (Do et al., 2014, Rahman, 2016), and appear to be restricted to populations of protozoa from lower termites (Zhou et al., 2007, Do et al., 2014, Rahman, 2016). Furthermore, insects do not appear to produce exogenous exoglucanses (Calderón-Cortés et al., 2012). It has been suggested that the presence of an elongated gut and mastication of the plant materials (features exhibited by *Pachysoma* (Holter and Scholtz, 2011, 2013)) by the insect may account for the absence of such exoglucanases (Calderón-Cortés et al., 2012). However, it should be noted that all GH families with known exoglucanase activity were present in both *Pachysoma* species, and therefore these enzymes may be present in the metagenomes although KEGG E.C. assignment was not possible.

KEGG E.C assignments showed the presence of a number of hemicellulases potentially capable of acting on a variety of substrates (Table 4.5). Xylanases were underrepresented in both *Pachysoma* metagenomes. Only two ORFs from the *P. striatum* metagenome had KEGG E.C. assignments to xylanases from GH11 and GH30, whereas no xylanases were assigned

from the *P. endroedyi* metagenome (Table 4.5). However, GH families known to include xylanases were present within both metagenomes. This result is unexpected as xylan is the most abundant hemicellulose (Kulkarni et al., 1999, Shallom and Shoham, 2003) and an abundant polymer of plant cell walls (Scheller and Ulvskov, 2010) which varies between plant species (Kulkarni et al., 1999, Beg et al., 2001, Collins et al., 2005, Scheller and Ulvskov, 2010). Furthermore, no genes encoding for α-arabinofuranosidase, which are also involved in xylan breakdown (Beg et al., 2001, Shallom and Shoham, 2003), were detected in either *Pachysoma* gut metagenome. Arabinofuranosidases are abundant in metagenomes from grass decomposing systems including insect gut samples (He et al., 2013), suggesting the absence of this plant type from *Pachysoma* diets. However, other enzymes assisting in xylan breakdown (Beg et al., 2001, Shallom and Shoham, 2003) were present, namely β-xylosidase (GH3, GH43 and GH52) and α-glucuronidase (GH67). The majority of enzymes were linked to mannose and galactose metabolism (Table 4.5), which are significant components of hemicellulases in a variety of plants (Scheller and Ulvskov, 2010) Figure 4.7 gives summary of the cellulases and hemicellulases present in the *Pachysoma* gut metagenomes.

Microbial conditioning of the food-source may be responsible for the absence of lignin degrading genes in both metagenomes. Although the lignin content of the *Pachysoma* diet is potentially lower than that of wood-feeding insects, it is not expected to be absent from either the dung or plant detritus substrates. The absence of such genes is recorded in both higher and lower termites feeding on a variety of diets which contain substantial concentrations of lignin, such as wood (Warnecke et al., 2007, He et al., 2013, Rahman, 2016). Microbial conditioning may be responsible for the absence of such genes in fungus-growing termites (Hyodo et al., 2000, Ohkuma et al., 2001, Hyodo et al., 2003). Furthermore, disruption of the physical integrity of the plant biomass may reduce lignin concentrations (Hyodo et al., 1999).

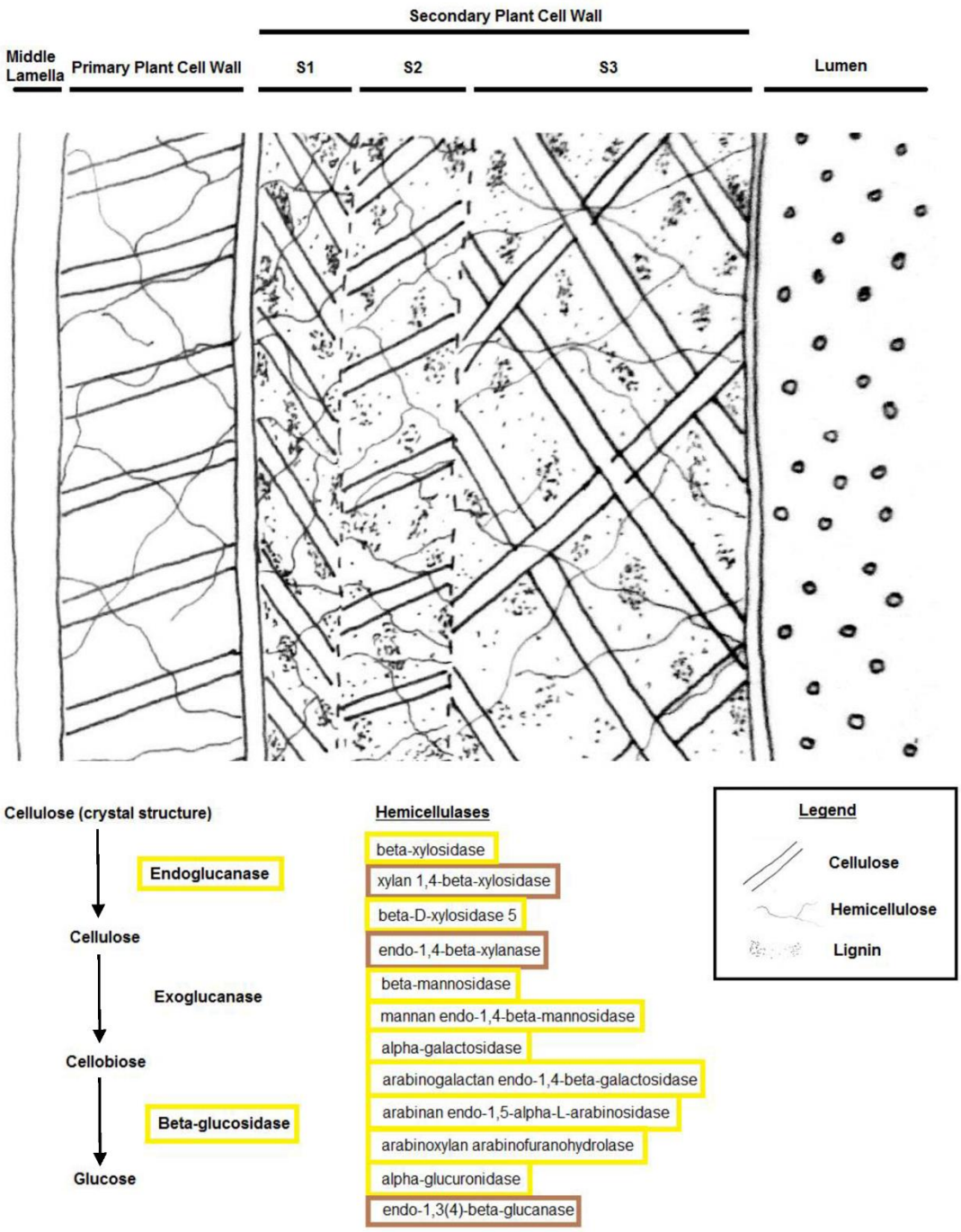


Figure 4.7: Summary of the plant cell wall degrading enzymes present in the *P. endroedyi* and *P. striatum* gut metagenomes. Only hemicellulases present in at least one of the metagenomes are listed. Enzymes are highlighted according to the metagenome in which they were detected: *P. striatum* (brown) or both *Pachysoma* metagenomes (yellow).

KEGG E.C assignments (after removal of possible false positive identifications) could only be made for 43.9% and 55.1% of the ORFs assigned to a GH family for the *P. endroedyi* and *P. striatum* gut microbial metagenomes, respectively. This result, coupled with the low percentage of ORFs that could be classified into COG categories (*P. endroedyi* = 37.9% and *P. striatum* = 35.1%), suggests that novel enzymes may be present within the *Pachysoma* metagenomes.

4.2.4.2 Other Nutritional Pathways in the *Pachysoma* Gut Metagenome

Due to the similarities of plant degrading potential between the two *Pachysoma* gut microbiomes, other nutritional pathways were compared to determine variations in the functional capacity of the two *Pachysoma* microbiomes dependent on diet. Comparisons were made between genes from the abundant COG categories E (amino acid transport and metabolism) in the context of nitrogen metabolism, C (energy production and conversion) and G (carbohydrate transport and metabolism) by assessing the respective KEGG pathways. Furthermore, vitamin and sterol metabolism were also considered, due to the importance of these nutritional components to insects.

4.2.4.2.1 Nitrogen Metabolism

Nitrogen is an essential nutrient to insects with those feeding on nitrogen rich diets possibly meeting their nutritional needs through their diet, whereas those feeding on nutrient poor diets, such as wood, commonly rely on microbial symbionts to fix nitrogen (Douglas, 2009). Nitrogen content of plant detritus varies (Kemp et al., 2003, Hättenschwiler et al., 2008), whereas dung is considered a nitrogen rich source (Liao et al., 2004), although the nitrogenous content of naturally dried dung and plant detritus fed on by *Pachysoma* is not well understood. However, it appears that the nitrogen content of both *Pachysoma* diets is sufficient due to the paucity of nitrogen fixation genes. Furthermore, only a small percentage of ORFs (*P. endroedyi*=2.1%; *P. striatum*=2.5%) were assigned to KEGG nitrogen metabolism pathways. The majority of genes were involved in the assimilation of ammonium and synthesis of the α-

amino acid L-glutamate (Figure 4.8) corresponding with the insects utilizing fixed nitrogen from the diet. Similar pathways were also dominant in a wood feeding beetle (Scully et al., 2013).

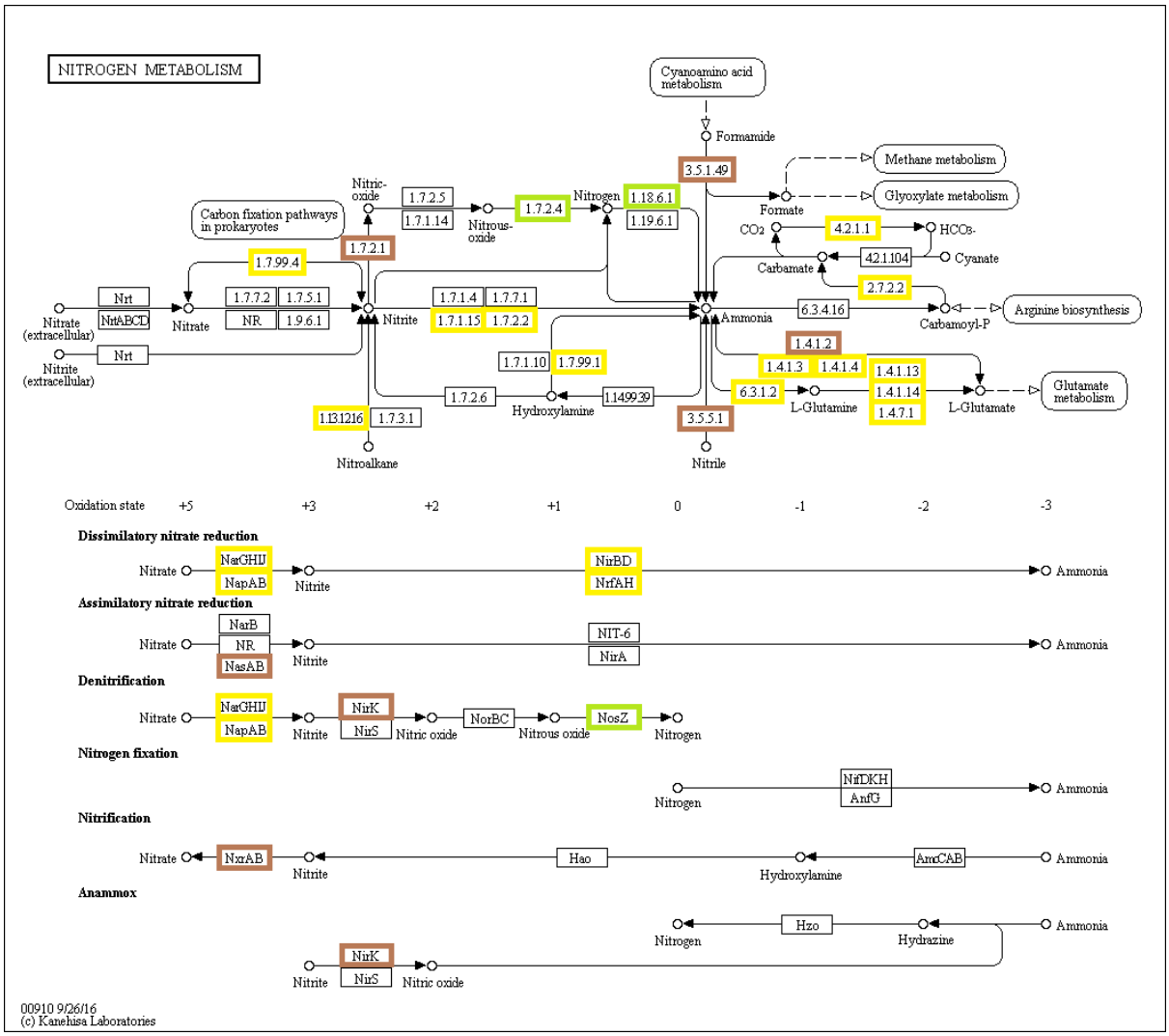


Figure 4.8: Adapted KEGG nitrogen metabolism reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes.

Microbial debris excreted with the dung (Mlambo et al., 2015) or new microbial decomposers may serve as a source of nitrogen in the dung (He et al., 2013). As both the dung and plant detritus are expected to undergo microbial conditioning before consumption by *Pachysoma*, it is possible that microbial cells are a major source of nitrogen for both

Pachysoma species and would explain the lack of nitrogen fixation genes in the *P. striatum* metagenome. This is consistent with the presence of genes encoding full or partial peptidoglycan metabolism pathways (a component of microbial cell walls) in both metagenomes (Figure 4.9). Furthermore, a large percentage of genes encoded to the COG category E amino acid transport and metabolism (10.3-10.7%). Full or partial pathways were present in both species for metabolism of all the essential amino acids. Little difference was noted in the genes encoding amino acid metabolism between the two species despite the availability of amino acids from dung (Barbehenn et al., 1999, Nation, 2008), although the amino acid composition of the exact diet fed on by *P. endroedyi* and *P. striatum* is unknown. This suggests that *P. endroedyi* may be able to metabolise amino acids from dung, enabling them to potentially feed on this food-source when plant detritus is unavailable.

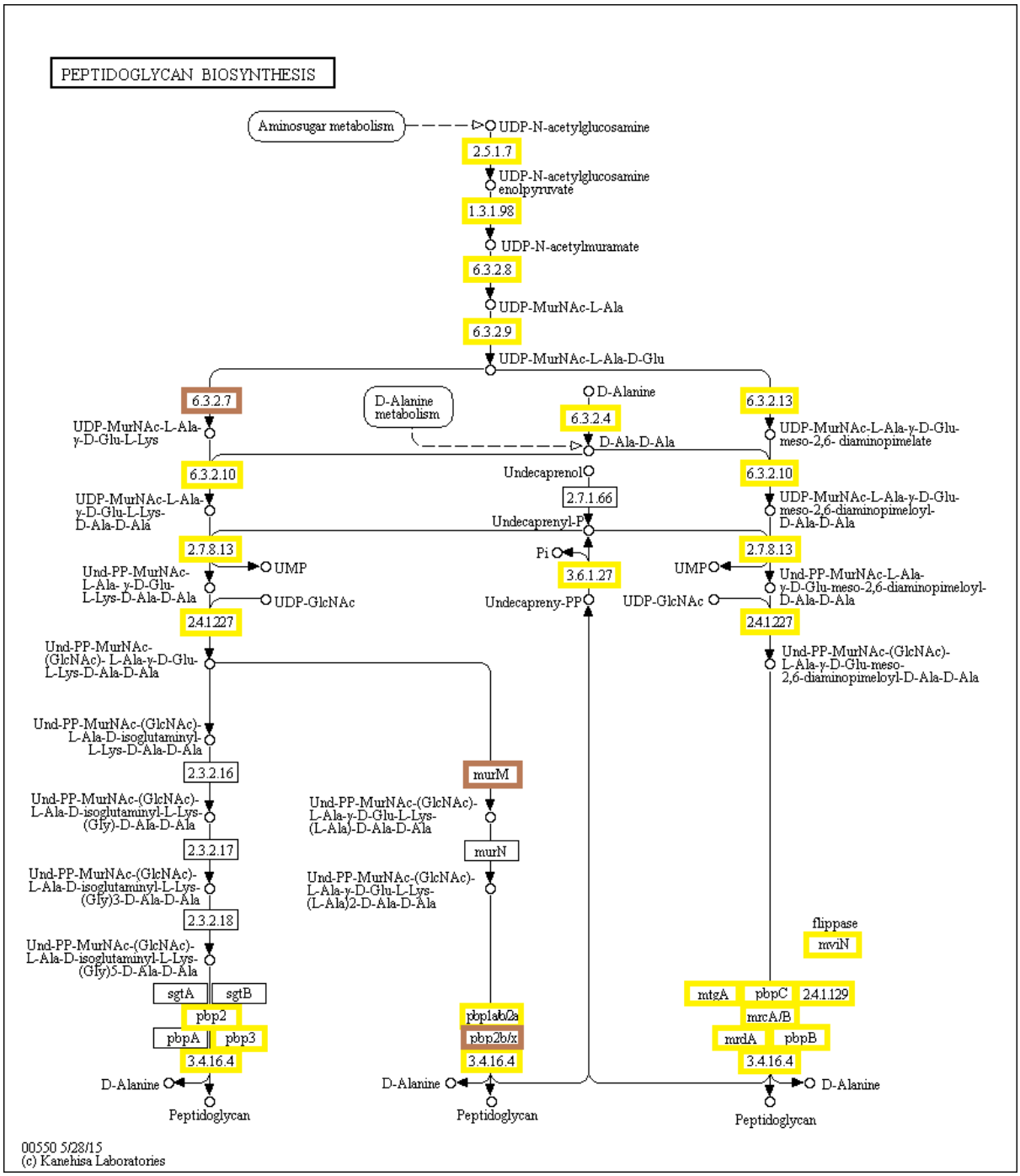


Figure 4.9: Adapted KEGG peptidoglycan biosynthesis reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes.

4.2.4.2.3 Energy Metabolism

Genes encoding oxidative phosphorylation (*P. endroedyi* = 19.2% and *P. striatum* = 25.6% energy metabolism ORFs) and methane metabolism (*P. endroedyi* = 29.1% and *P. striatum* =22.8% energy metabolism ORFs) pathways were abundant within both *Pachysoma* spp. metagenomes (Figure 4.10). The high abundance of genes involved in methane metabolism is not proportional to the abundance of methanogens within the gut. However, upon further inspection few genes were involved in synthesis of methane with the methane molecular marker gene, *mcrA* (K00399) (detected from "Candidatus *Methanomethylophilus alvus*", a known methanogen from the human gut; (Borrel et al., 2012)) only detected in the *P. endroedyi* metagenome. This corresponds with the diets of *Pachysoma,* especially dung, being substantial methane sources (Keppler et al., 2006, Willén, 2011)

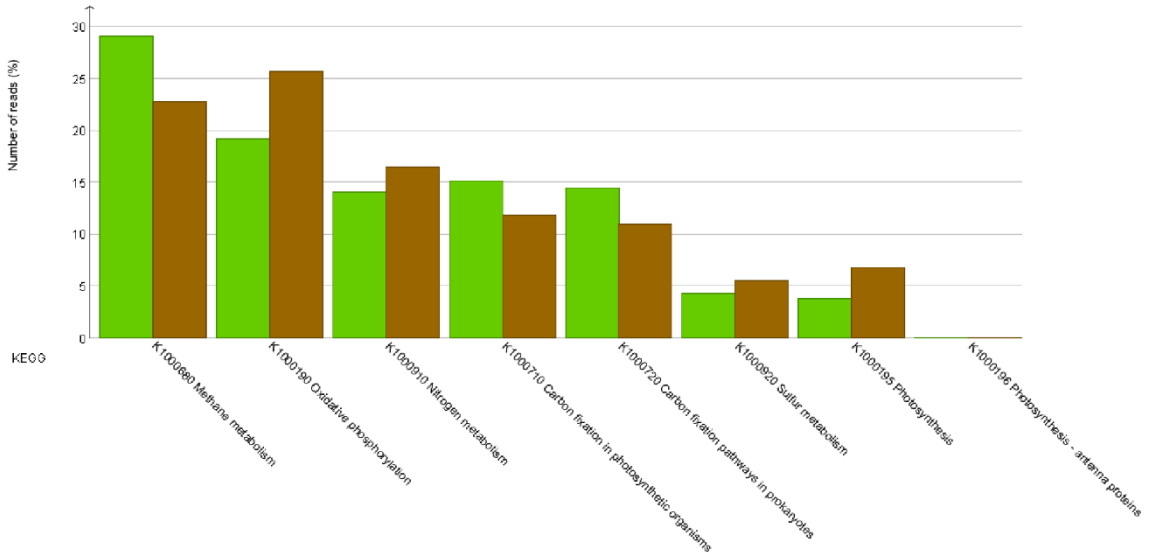


Figure 4.10: Abundance and diversity of genes encoding energy metabolism pathways (KEGG) for *P. endroedyi* (green) and *P. striatum* (brown) gut microbial metagenomes.

4.2.4.2.2 Carbohydrate Metabolism

Both *Pachysoma* metagenomes had an abundance of genes encoding the two major polymer hydrolysis pathways: glycolysis (*P. endroedyi* = 11.0% and *P. striatum* =11.6%

carbohydrate metabolism ORFs) and pentose phosphate (*P. endroedyi* = 8.9% and *P. striatum* =7.5% carbohydrate metabolism ORFs) pathways (Figure 4.11). This is similar to results noted in other plant feeding insects such as termites (Rahman, 2016). Genes encoding proteins for the starch and sucrose metabolism pathway were more abundant within *P. endroedyi* than *P. striatum* which corresponds with the larger concentration of plant materials in detritus than dung. This further suggests that *P. striatum* may not feed exclusively on plant biomass but also on the brown liquid component of the dung.

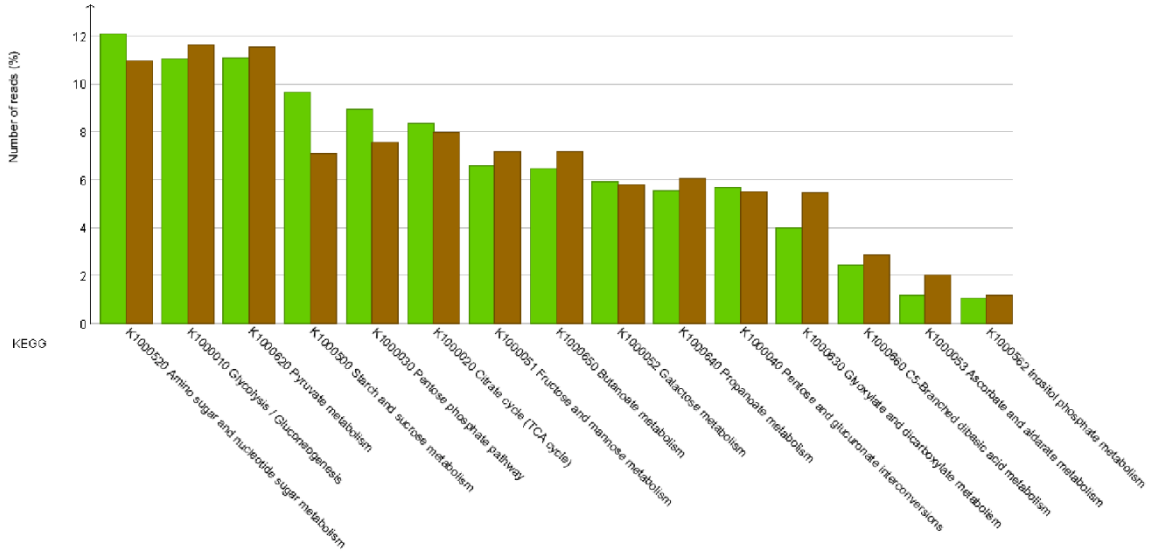


Figure 4.11: Abundance and diversity of genes encoding carbohydrate metabolism pathways (KEGG) for *P. endroedyi* (green) and *P. striatum* (brown) gut microbial metagenomes.

4.2.4.2.4 Vitamin, Sterol and Fatty Acid Synthesis

Water soluble (Vitamin B complex and C) and fat soluble (Vitamin A and E) vitamins are considered essential for all insect species (Barbehenn et al., 1999). Vitamin B acts as co-enzymes with pyridoxine (Vitamin B_6), folate (Vitamin B_9), riboflavin and FAD (Vitamin B_2), biotin (Vitamin B_7), thiamine (Vitamin B_1), nicotinic acid (Vitamin B_3), cobalamin (Vitamin B_{12}) and pantothenate (Vitamin B_5) essential to insects (Barbehenn et al., 1999). As noted with

plant degrading enzymes, profiles of genes involved in vitamin metabolism were similar for both species, although *P. striatum* typically had a higher diversity of genes in low abundances.

A number of genes involved in the metabolism of the essential B vitamins, with the exception of pyridoxine, were detected in the *Pachysoma* gut metagenome of both species. This suggests that microorganisms are, at least in part, responsible for the metabolism of the vitamin B complex. However, only a few genes involved in the pyridoxine metabolism pathway were detected (Figure 4.12), which could be due to the insect being able to synthesize vitamin B6 itself. The absence of such genes is unexpected, with microbial genes for synthesis of pyridoxine detected from the microbiomes of other plant feeding insect species (Scully et al., 2013).

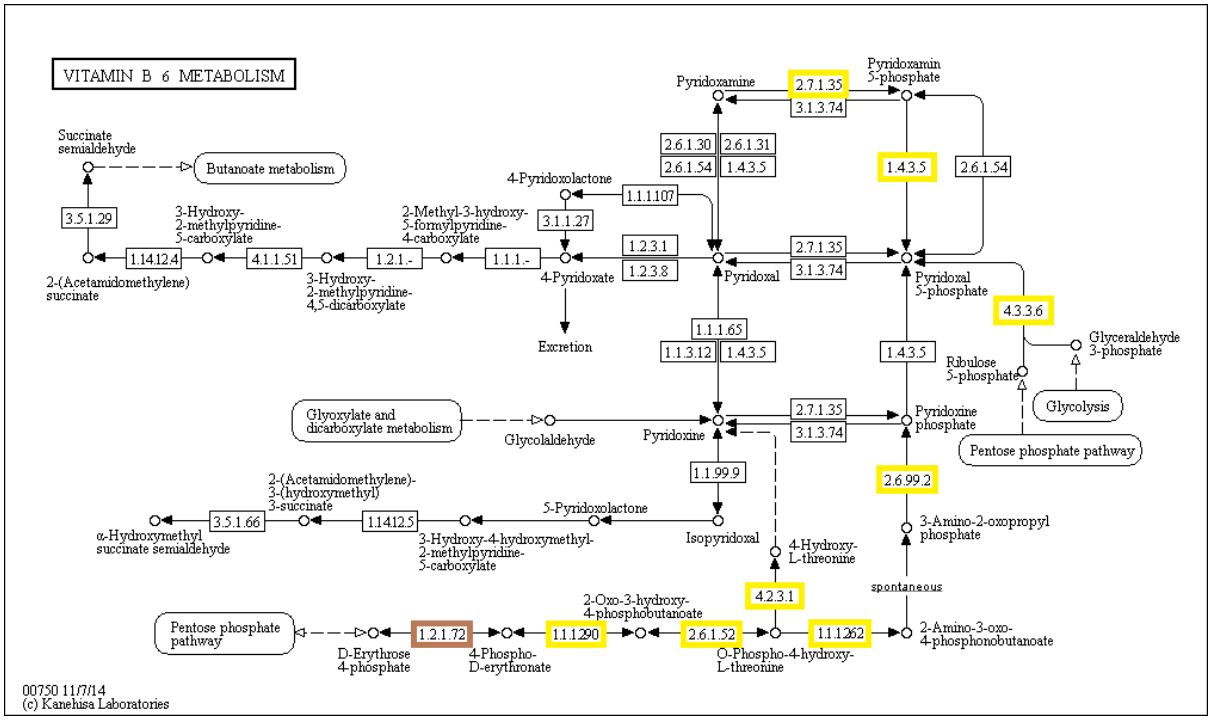


Figure 4.12: Adapted KEGG pyridoxine (vitamin B6) metabolism reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes.

Genes involved in the metabolism of other vitamins were also detected. Vitamin C (L-ascorbic acid) metabolism pathways were not complete in *Pachysoma*. Vitamin C is only required by herbivorous insects (Barbehenn et al., 1999), presumably those feeding on live plants. However, the detection of several genes involved in the synthesis of L-Ascorbate-6P (Figure 4.13) suggests that microbial communities may partly assist in vitamin C synthesis. Vitamin C is readily available in plant substrates and therefore would not be expected to be synthesized by the insect. Similarly, vitamin E is not required by all insects (Barbehenn et al., 1999), which is confirmed by the absence of genes involved in α-tocopherol (vitamin E) synthesis pathways suggesting that if this vitamin is essential to *Pachysoma*, it is derived from the diet or synthesized by the insect. Another suggestion is that vitamin K_1, of which synthesis pathways were present in *Pachysoma*, may be substituted for vitamin E (Barbehenn et al., 1999).

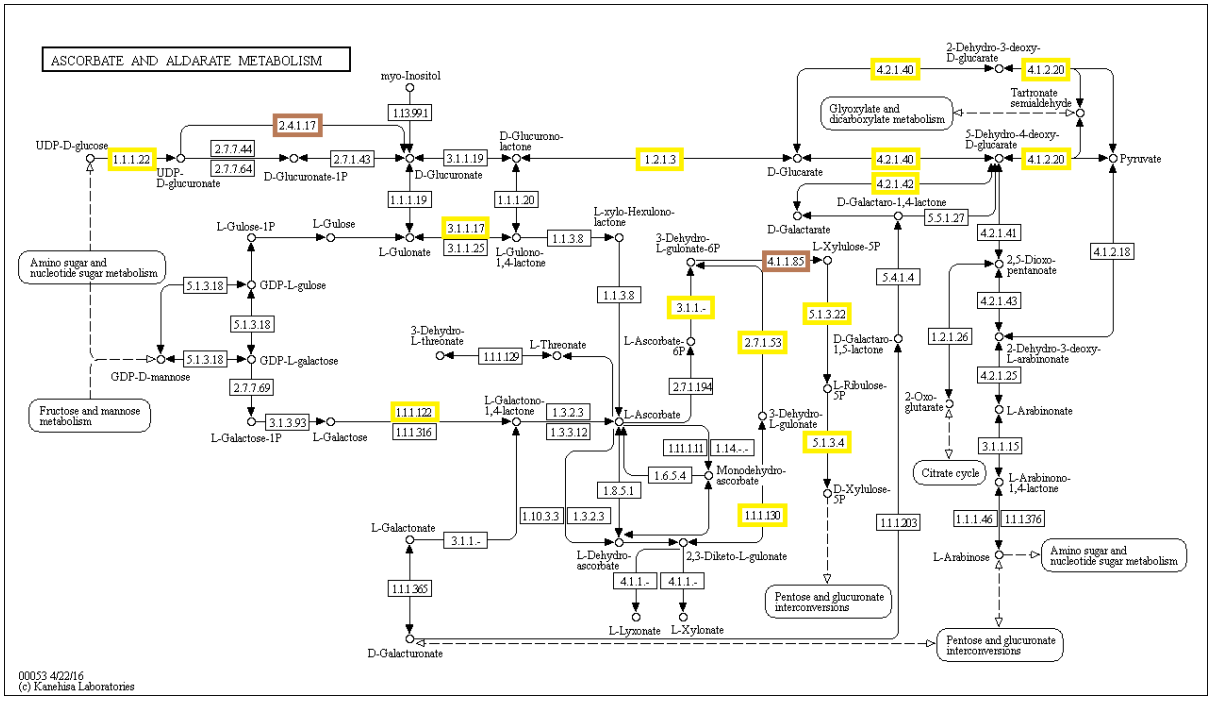


Figure 4.13: Adapted KEGG Ascorbate (Vitamin C) and Aldarate metabolism reference pathway (http://www.genome.jp) depicting which enzymes were present in the *P. endroedyi* (green), *P. striatum* (brown) and both (yellow) metagenomes.

Animals are unable to synthesise sterols, which are essential for insect nutrition. Therefore, the majority of insects derive sterols from their diet. This appears to be the case in *Pachysoma* with no synthesis pathways for the fungal-derived sterol ergesterol or synthesis of the necessary insect sterols 7-dehydrocholesterol and cholesterol (Douglas, 2009). These results further highlight the similarities in the functional capacities of the two *Pachysoma* microbiomes, once again suggesting that individuals of the two species may be able to feed on both diets.

4.3 Conclusion

This is the first shotgun metagenomic study on a dry detritus/dung feeding insect. Here it has been possible to compare the gut microbial community structure, and functional capacity thereof, of two *Pachysoma* species feeding on different diets. The bacterial diversity was shown to be greater than previously expected in *P. endroedyi* gut samples. However, large differences in bacterial diversity were noted, as found for the amplicon study (Chapter 3; Franzini et al. (2016)). In comparison, the functional capacity of both *Pachysoma* microbiomes was similar, with the exception of methane synthesis. Therefore, it would appear that *Pachysoma* species may have a core functional microbiome and that individuals of each *Pachysoma* species may be able to feed on and successfully utilize the other diet (i.e. dung feeding species feed on plant detritus and vice versa) when needed. This corresponds with *P. endroedyi* originally being classified as polyphagous through feeding observations (Harrison et al., 2003). Furthermore, there is evidence that novel GH enzymes, particularly cellulases and hemicellulases, may be present in *Pachysoma* metagenomes and this should be determined further by specialized studies.

4.4 References

Adriaenssens, E.M., van Zyl, L.J., Cowan, D.A. & Trindade, M.I., 2016. Metaviromics of Namib desert salt pans: a novel lineage of haloarchaeal salterproviruses and a rich source of ssDNA viruses. *Viruses,* 8(1), pp. 14-32.

Adriaenssens, E.M., Van Zyl, L., De Maayer, P., Rubagotti, E., Rybicki, E., Tuffin, M. & Cowan, D.A., 2015. Metagenomic analysis of the viral community in Namib Desert hypoliths. *Environmental Microbiology,* 17(2), pp. 480-495.

Ahn, J.-H., Hong, I.-P., Bok, J.-I., Kim, B.-Y., Song, J. & Weon, H.-Y., 2012. Pyrosequencing analysis of the bacterial communities in the guts of honey bees *Apis cerana* and *Apis mellifera* in Korea. *Journal of Microbiology,* 50(5), pp. 735-745.

Anderson, J.M. & Coe, M.J., 1974. Decomposition of elephant dung in an arid, tropical environment. *Oecologia,* 14(1-2), pp. 111-125.

Andongma, A.A., Wan, L., Dong, Y.-C., Desneux, N., White, J.A. & Niu, C.-Y., 2015. Pyrosequencing reveals a shift in symbiotic bacteria populations across life stages of *Bactrocera dorsalis*. *Scientific Reports,* 5.

Auburn, S., Campino, S., Clark, T.G., Djimde, A.A., Zongo, I., Pinches, R., et al., 2011. An effective method to purify *Plasmodium falciparum* DNA directly from clinical blood samples for whole genome high-throughput sequencing. *PLoS One,* 6(7), pp. e22213.

Barbehenn, R.V., Reese, J.C. & Hagen, K.S., 1999. *The food of insects*. John Wiley & Sons, USA, New York.

Beg, Q., Kapoor, M., Mahajan, L. & Hoondal, G.S., 2001. Microbial xylanases and their industrial applications: a review. *Applied Microbiology and Biotechnology,* 56(3-4), pp. 326-338.

Belda, E., Pedrola, L., Peretó, J., Martínez-Blanch, J.F., Montagud, A., Navarro, E., et al., 2011. Microbial diversity in the midguts of field and lab-reared populations of the European corn borer *Ostrinia nubilalis*. *PLoS One,* 6(6), pp. e21751.

Bertino-Grimaldi, D., Medeiros, M.N., Vieira, R.P., Cardoso, A.M., Turque, A.S., Silveira, C.B., et al., 2013. Bacterial community composition shifts in the gut of *Periplaneta americana* fed on different lignocellulosic materials. *SpringerPlus,* 2(1), pp. 609.

Borrel, G., Harris, H.M., Tottey, W., Mihajlovski, A., Parisot, N., Peyretaillade, E., et al., 2012. Genome sequence of “Candidatus *Methanomethylophilus alvus*” Mx1201, a methanogenic archaeon from the human gut belonging to a seventh order of methanogens. *Journal of Bacteriology,* 194(24), pp. 6944-6945.

Boucias, D.G., Cai, Y., Sun, Y., Lietze, V.U., Sen, R., Raychoudhury, R. & Scharf, M.E., 2013. The hindgut lumen prokaryotic microbiota of the termite *Reticulitermes flavipes* and its responses to dietary lignocellulose composition. *Molecular Ecology,* 22(7), pp. 1836-1853.

Calderón-Cortés, N., Quesada, M., Watanabe, H., Cano-Camacho, H. & Oyama, K., 2012. Endogenous plant cell wall digestion: a key mechanism in insect evolution. *Annual Review of Ecology, Evolution, and Systematics,* 43, pp. 45-71.

Carbonero, F., Nava, G.M., Benefiel, A.C., Greenberg, E. & Gaskins, H.R., 2011. Microbial DNA extraction from intestinal biopsies is improved by avoiding mechanical cell disruption. *Journal of Microbiological Methods,* 87(1), pp. 125-127.

Carpi, G., Cagnacci, F., Wittekindt, N.E., Zhao, F., Qi, J., Tomsho, L.P., et al., 2011. Metagenomic profile of the bacterial communities associated with *Ixodes ricinus* ticks. *PLoS One,* 6(10), pp. e25604.

Chandler, J.A., Liu, R.M. & Bennett, S.N., 2015. RNA shotgun metagenomic sequencing of northern California (USA) mosquitoes uncovers viruses, bacteria, and fungi. *Frontiers in Microbiology,* 6.

Collins, T., Gerday, C. & Feller, G., 2005. Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews,* 29(1), pp. 3-23.

Delalibera, I., Vasanthakumar, A., Burwitz, B.J., Schloss, P.D., Klepzig, K.D., Handelsman, J. & Raffa, K.F., 2007. Composition of the bacterial community in the gut of the pine engraver, *Ips pini* (Say)(Coleoptera) colonizing red pine. *Symbiosis,* 43(2), pp. 97-104.

Do, T.H., Nguyen, T.T., Nguyen, T.N., Le, Q.G., Nguyen, C., Kimura, K. & Truong, N.H., 2014. Mining biomass-degrading genes through Illumina-based de novo sequencing and metagenomic analysis of free-living bacteria in the gut of the lower termite *Coptotermes gestroi* harvested in Vietnam. *Journal of Bioscience and Bioengineering,* 118(6), pp. 665-671.

Douglas, A.E., 2009. The microbial dimension in insect nutritional ecology. *Functional Ecology,* 23(1), pp. 38-47.

Engel, P., Martinson, V.G. & Moran, N.A., 2012. Functional diversity within the simple gut microbiota of the honey bee. *Proceedings of the National Academy of Sciences,* 109(27), pp. 11002-11007.

Feehery, G.R., Yigit, E., Oyola, S.O., Langhorst, B.W., Schmidt, V.T., Stewart, F.J., et al., 2013. A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS One,* 8(10), pp. e76096.

Franzini, P.Z.N., Ramond, J.-B., Scholtz, C.H., Sole, C.L., Ronca, S. & Cowan, D.A., 2016. The gut microbiomes of two *Pachysoma* MacLeay desert dung beetle species (Coleoptera: Scarabaeidae: Scarabaeinae) feeding on different diets. *PLoS One,* 11(8), pp. e0161118.

Geib, S.M., del Mar Jimenez-Gasco, M., Carlson, J.E., Tien, M. & Hoover, K., 2009. Effect of host tree species on cellulase activity and bacterial community composition in the gut of larval Asian longhorned beetle. *Environmental Entomology,* 38(3), pp. 686-699.

Harrison, J.d.G., Scholtz, C.H. & Chown, S.L., 2003. A revision of the endemic south-western African dung beetle subgenus *Scarabaeus* (*Pachysoma*) MacLeay, including notes on other flightless Scarabaeini (Scarabaeidae: Scarabaeinae). *Journal of Natural History,* 37(3), pp. 305-355.

Hättenschwiler, S., Aeschlimann, B., Coûteaux, M.M., Roy, J. & Bonal, D., 2008. High variation in foliage and leaf litter chemistry among 45 tree species of a neotropical rainforest community. *New Phytologist,* 179(1), pp. 165-175.

He, S., Ivanova, N., Kirton, E., Allgaier, M., Bergin, C., Scheffrahn, R.H., et al., 2013. Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS One,* 8(4), pp. e61126.

Hodkinson, B.P. & Grice, E.A., 2015. Next-generation sequencing: a review of technologies and tools for wound microbiome research. *Advances in Wound Care,* 4(1), pp. 50-58.

Holter, P. & Scholtz, C.H., 2011. Re-establishment of biting mouthparts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on plant litter—old structures reacquired or new ones evolved? *Journal of Morphology,* 272(8), pp. 1007-1016.

Holter, P. & Scholtz, C.H., 2013. Elongated hindguts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on dry dung pellets or plant litter. *Journal of Morphology,* 274(6), pp. 657-662.

Hunter, S.J., Easton, S., Booth, V., Henderson, B., Wade, W.G. & Ward, J.M., 2011. Selective removal of human DNA from metagenomic DNA samples extracted from dental plaque. *Journal of Basic Microbiology,* 51(4), pp. 442-446.

Hyodo, F., Azuma, J.-I. & Abe, T., 1999. Estimation of effect of passage through the gut of a lower termite, *Coptotermes formosanus* Shiraki, on lignin by Solid-State CP/MAS 13C NMR. *Holzforschung,* 53(3), pp. 244-246.

Hyodo, F., Inoue, T., Azuma, J.-I., Tayasu, I. & Abe, T., 2000. Role of the mutualistic fungus in lignin degradation in the fungus-growing termite *Macrotermes gilvus* (Isoptera; Macrotermitinae). *Soil Biology and Biochemistry,* 32(5), pp. 653-658.

Hyodo, F., Tayasu, I., Inoue, T., Azuma, J.I., Kudo, T. & Abe, T., 2003. Differential role of symbiotic fungi in lignin degradation and food provision for fungus-growing termites (Macrotermitinae: Isoptera). *Functional Ecology,* 17(2), pp. 186-193.

Jarman, S.N., Deagle, B.E. & Gales, N.J., 2004. Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology,* 13(5), pp. 1313-1322.

Jia, S., Zhang, X., Zhang, G., Yin, A., Zhang, S., Li, F., et al., 2013. Seasonally variable intestinal metagenomes of the red palm weevil (*Rhynchophorus ferrugineus*). *Environmental Microbiology,* 15(11), pp. 3020-3029.

Kemp, P.R., Reynolds, J.F., Virginia, R.A. & Whitford, W.G., 2003. Decomposition of leaf and root litter of Chihuahuan desert shrubs: effects of three years of summer drought. *Journal of Arid Environments,* 53(1), pp. 21-39.

Keppler, F., Hamilton, J.T.G., Braß, M. & Röckmann, T., 2006. Methane emissions from terrestrial plants under aerobic conditions. *Nature,* 439(7073), pp. 187-191.

Kingsford, C., Schatz, M.C. & Pop, M., 2010. Assembly complexity of prokaryotic genomes using short reads. *BMC Bioinformatics,* 11, pp. 21.

Krishnan, M., Bharathiraja, C., Pandiarajan, J., Prasanna, V.A., Rajendhran, J. & Gunasekaran, P., 2014. Insect gut microbiome–an unexploited reserve for biotechnological application. *Asian Pacific Journal of Tropical Biomedicine,* 4, pp. S16-S21.

Kulkarni, N., Shendye, A. & Rao, M., 1999. Molecular and biotechnological aspects of xylanases. *FEMS Microbiology Reviews,* 23(4), pp. 411-456.

Liao, W., Liu, Y., Liu, C. & Chen, S., 2004. Optimizing dilute acid hydrolysis of hemicellulose in a nitrogen-rich cellulosic material—dairy manure. *Bioresource Technology,* 94(1), pp. 33-41.

Liu, N., Yan, X., Zhang, M., Xie, L., Wang, Q., Huang, Y., et al., 2011. Microbiome of fungus-growing termites: a new reservoir for lignocellulase genes. *Applied and Environmental Microbiology,* 77(1), pp. 48-56.

Liu, N., Zhang, L., Zhou, H., Zhang, M., Yan, X., Wang, Q., et al., 2013. Metagenomic insights into metabolic capacities of the gut microbiota in a fungus-cultivating termite (*Odontotermes yunnanensis*). *PLoS One,* 8(7), pp. e69184.

Lombard, V., Ramulu, H.G., Drula, E., Coutinho, P.M. & Henrissat, B., 2014. The Carbohydrate-Active Enzymes database (CAZy) in 2013. *Nucleic Acids Research,* 42(D1), pp. D490-D495.

Lynd, L.R., Weimer, P.J., Van Zyl, W.H. & Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology and Molecular Biology Reviews,* 66(3), pp. 506-577.

Mendes, L.W., Tsai, S.M., Navarrete, A.A., De Hollander, M., van Veen, J.A. & Kuramae, E.E., 2015. Soil-borne microbiome: linking diversity to function. *Microbial Ecology,* 70(1), pp. 255-265.

Mlambo, S., Sole, C.L. & Scholtz, C.H., 2015. A molecular phylogeny of the African Scarabaeinae (Coleoptera: Scarabaeidae). *Arthropod Systematics & Phylogeny,* 73(2), pp. 303-321.

Montagna, M., Chouaia, B., Mazza, G., Prosdocimi, E.M., Crotti, E., Mereghetti, V., et al., 2015. Effects of the diet on the microbiota of the red palm weevil (Coleoptera: Dryophthoridae). *PLoS One,* 10(1), pp. e0117439.

Nation, J.L., 2008. *Insect physiology and biochemistry*. CRC Press, USA.

Nguyen, N.H., Suh, S.-O. & Blackwell, M., 2007. Five novel *Candida* species in insect-associated yeast clades isolated from Neuroptera and other insects. *Mycologia,* 99(6), pp. 842-858.

Ohkuma, M., Maeda, Y., Johjima, T. & Kudo, T., 2001. Lignin degradation and roles of white rot fungi: Study on an efficient symbiotic system in fungus-growing termites and its application to bioremediation. *RIKEN Review No 42*, pp. 39-42.

Pandey, N. & Rajagopal, R., 2016. Molecular characterization and diversity analysis of bacterial communities associated with *Dialeurolonga malleswaramensis* (Hemiptera: Aleyrodidae) adults using 16S rDNA amplicon pyrosequencing and FISH. *Insect Science,* 5(23), pp. 704-711.

Paxinos, E., McIntosh, C., Ralls, K. & Fleischer, R., 1997. A noninvasive method for distinguishing among canid species: amplification and enzyme restriction of DNA from dung. *Molecular Ecology,* 6(5), pp. 483-486.

Rahman, N.A., 2016. *Molecular investigation of Australian termites and their gut symbionts.* Doctor of Philosophy, University of Queensland.

Sankar, S.A., Lagier, J.-C., Pontarotti, P., Raoult, D. & Fournier, P.-E., 2015. The human gut microbiome, a taxonomic conundrum. *Systematic and Applied Microbiology,* 38(4), pp. 276-286.

Santana, R.H., Catão, E.C.P., Lopes, F.A.C., Constantino, R., Barreto, C.C. & Krüger, R.H., 2015. The gut microbiota of workers of the litter-feeding termite *Syntermes wheeleri* (Termitidae: Syntermitinae): archaeal, bacterial, and fungal communities. *Microbial Ecology,* 70, pp. 545-556.

Schauer, C., Thompson, C.L. & Brune, A., 2014. Pyrotag sequencing of the gut microbiota of the cockroach *Shelfordella lateralis* reveals a highly dynamic core but only limited effects of diet on community structure. *PLoS One,* 9(1), pp. e85861.

Scheller, H.V. & Ulvskov, P., 2010. Hemicelluloses. *Plant Biology,* 61(1), pp. 263.

Scully, E.D., Geib, S.M., Hoover, K., Tien, M., Tringe, S.G., Barry, K.W., et al., 2013. Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. *PLoS One,* 8(9), pp. e73827.

Shallom, D. & Shoham, Y., 2003. Microbial hemicellulases. *Current Opinion in Microbiology,* 6(3), pp. 219-228.

Shi, W., Xie, S., Chen, X., Sun, S., Zhou, X., Liu, L., et al., 2013. Comparative genomic analysis of the endosymbionts of herbivorous insects reveals eco-environmental adaptations: biotechnology applications. *PLoS Genetics,* 9(1), pp. e1003131.

Suen, G., Weimer, P.J., Stevenson, D.M., Aylward, F.O., Boyum, J., Deneke, J., et al., 2011. The complete genome sequence of *Fibrobacter succinogenes* S85 reveals a cellulolytic and metabolic specialist. *PLoS One,* 6(4), pp. e18814.

Suh, S.-O. & Blackwell, M., 2004. Three new beetle-associated yeast species in the *Pichia guilliermondii* clade. *FEMS Yeast Research,* 5(1), pp. 87-95.

Suh, S.-O., McHugh, J.V. & Blackwell, M., 2004a. Expansion of the *Candida tanzawaensis* yeast clade: 16 novel *Candida* species from basidiocarp-feeding beetles. *International Journal of Systematic and Evolutionary Microbiology,* 54(6), pp. 2409-2429.

Suh, S.-O., Gibson, C.M. & Blackwell, M., 2004b. *Metschnikowia chrysoperlae* sp. nov., *Candida picachoensis* sp. nov. and *Candida pimensis* sp. nov., isolated from the green lacewings *Chrysoperla comanche* and *Chrysoperla carnea* (Neuroptera: Chrysopidae). *International Journal of Systematic and Evolutionary Microbiology,* 54(5), pp. 1883-1890.

Suh, S.-O., McHugh, J.V., Pollock, D.D. & Blackwell, M., 2005. The beetle gut: a hyperdiverse source of novel yeasts. *Mycological Research,* 109(3), pp. 261-265.

Swift, M.J., Heal, O.W. & Anderson, J.M., 1979. *Decomposition in terrestrial ecosystems*. University of California Press, USA, California.

Symondson, W.O.C., 2002. Molecular identification of prey in predator diets. *Molecular Ecology,* 11(4), pp. 627-641.

Terra, W.R. & Ferreira, C., 1994. Insect digestive enzymes: properties, compartmentalization and function. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry,* 109B(1), pp. 1-62.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., et al., 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science,* 304(5667), pp. 66-74.

Vikram, S., Guerrero, L.D., Makhalanyane, T.P., Le, P.T., Seely, M. & Cowan, D.A., 2016. Metagenomic analysis provides insights into functional capacity in a hyperarid desert soil niche community. *Environmental Microbiology,* 18(6), pp. 1875-1888.

Waite, D.W., Dsouza, M., Biswas, K., Ward, D.F., Deines, P. & Taylor, M.W., 2015. Microbial community structure in the gut of the New Zealand insect Auckland tree weta (*Hemideina thoracica*). *Archives of Microbiology,* 197(4), pp. 603-612.

Wang, A., Yao, Z., Zheng, W. & Zhang, H., 2014. Bacterial communities in the gut and reproductive organs of *Bactrocera minax* (Diptera: Tephritidae) based on 454 pyrosequencing. *PLoS One,* 9(9), pp. e106988.

Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., et al., 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature,* 450(7169), pp. 560-565.

Willén, A., 2011. *Methane production from dairy cows: relations between enteric production and production from faeces and urine*. Swedish University of Agricultural Sciences.

Zhou, X., Smith, J.A., Oi, F.M., Koehler, P.G., Bennett, G.W. & Scharf, M.E., 2007. Correlation of cellulase gene expression and cellulolytic activity throughout the gut of the termite *Reticulitermes flavipes*. *Gene,* 395(1), pp. 29-39.

Chapter 5: Concluding Remarks

5.1 Studying the *Pachysoma* gut microbiome

Insect-microbiota relationships have been studied for several years. Such studies have increased knowledge in how microorganisms interact with the host and how these microorganisms may be used for industrial purposes (e.g. biofuel research (Sun and Scharf, 2010)), pest control (Douglas, 2007) and other agricultural activities (e.g. understanding and improving honeybee health (Evans and Schwarz, 2011)) and understanding and treating disease (e.g. malaria (Dong et al., 2009)). However, the majority of gut microbiome studies focus on only a few specific insect taxa including termites (Brauman, 2000, Brune and Friedrich, 2000, Ohkuma, 2003, Brune, 2013), mosquitoes (Wang et al., 2011, Boissière et al., 2012, Osei-Poku et al., 2012, Wang et al., 2012, Coffey et al., 2014), flies (Broderick and Lemaitre, 2012, Wong et al., 2013, Wang et al., 2014, Wong et al., 2015) and honey bees (Ahn et al., 2012, Engel et al., 2012, Sabree et al., 2012). However, few studies have investigated the gut microbiota of the ecologically important adult dung beetles. These beetles are essential for nutrient cycling and removal of decaying dung which may decrease pest and disease prevalence (Nichols et al., 2008).

The lack of studies on the gut microbiota of dung beetles may be attributed to the continued emphasis on insects of medical, economic and agricultural import. A majority of studies also focus on insects feeding primarily on lignocellulosic materials as enzymes produced by lignocellulose degrading microorganisms may be useful in industry. As dung beetles typically feed on wet dung they do not fall into the above categories. Although plant fragments are found in dung (Anderson and Coe, 1974), beetles feeding on wet dung typically remove any plant biomass, exhibiting soft saprophagy (Anderson and Coe, 1974, Holter et al., 2002). In contrast, *Pachysoma* species feed on plant detritus or dry dung (Harrison et al., 2003), where plant biomass cannot easily be removed by the insect before consumption,

making them good candidates for studying microbiomes involved in lignocellulosic degradation. Evidence suggests that *Pachysoma* species feeding on both plant detritus and dung have adapted to feed on plant material (Holter and Scholtz, 2011, 2013), with researchers hypothesising that they are able to successfully digest lignocellulosic diets.

Insect host phylogeny has an important impact on gut microbiome assembly (Colman et al., 2012). Therefore, host phylogeny should always be considered when studying the impact of environmental variables (diet, location, etc.) on gut microbiome assembly. However, host phylogeny may have a greater impact on gut microbiome assembly in some insect groups compared to others (Colman et al., 2012). Host phylogeny is expected to have a less pronounced effect on gut microbiome assembly in insect species from the same genus. Therefore, by studying insect genera with species feeding on different diets, such as *Pachysoma,* we may be able to better understand the effect of host diet on gut microbial communities where the influence of host phylogeny is reduced.

In order to research the relationship between *Pachysoma* species and their respective gut microbiota, this study aimed to determine the taxonomical diversity and functional capacity of the *Pachysoma* gut microbiome. This research acted as the primary study to determine the basic characteristics and potential of the associated microorganisms. This was achieved by comparing the gut microbiomes of plant detritus (*P. endroedyi*) and dry dung (*P. striatum*) feeding insects to answer the following three main research questions:

1. What is the gut microbial diversity of the two *Pachysoma* species and how does it differ between the two species?
2. What is the functional capacity of the *Pachysoma* gut microbiome to assist the host with digestion of the respective food-source?
3. Is host diet or phylogeny the main driver of microbiome assembly?

Next-generation sequencing provided answers to the first two research questions, the implications of which will be discussed further below. However, this study was unable to

conclusively answer the third research question which is discussed further as a limitation (Section 5.2). Suggestions for future studies to better understand the role of host diet and phylogeny in *Pachysoma* gut microbiome assembly will be discussed in Section 5.3.

Both the amplicon and shotgun sequencing studies showed distinct differences in gut bacterial diversity between the two *Pachysoma* speices with the amplicon study indicating a small core microbiome. However, the functional capacities of the two *Pachysoma* gut microbial communities were similar. This is in line with the suggestion that a functional core microbiome may be present in place of a phylogenetic core (Turnbaugh et al., 2009, Karasov et al., 2011). Furthermore, *Pachysoma* species and individuals appear to associate with different microbial communities which assist the host in a similar fashion nutritionally. Furthermore, *P. striatum* does appear to feed on plant biomass found within the dung, unlike dung beetles feeding on wet faeces. All the above would allow the different *Pachysoma* species to feed on both diet types in the situation where the respective specialised diet is not available. For desert dwellers, such an adaption would be invaluable. Although competition for specific food-sources is reduced in desert ecosystems (Wharton, 2002), the specialised diet may be unavailable. Such observations are in line with previous reports of polyphagous feeding behaviour in *P. endroedyi* (Harrison et al., 2003). Since then, unpublished data has suggested a specialised feeding behaviour in *P. endroedyi* (C. Scholtz Pers. Comm.). However, this study may suggest that previous observations were in fact indicating the ability of individuals to feed on both food-sources when necessary.

5.2 Limitations of this study

The insect gut is a complex environment where numerous factors affect microbiome assembly (Engel and Moran, 2013, Douglas, 2015). Although certain variables were controlled during this study (location, season and life-stage), other factors are likely to have an impact on the respective gut microbiota, although the intensity of such is unknown. In particular host diet and phylogeny are both probable role players in microbiome assembly (Colman et al.,

2012, Yun et al., 2014), even though the effect of host phylogeny is expected to have a lesser impact than across higher insect taxa. In this study, these variables are intertwined, with each species feeding on a different diet. Although some evidence was found to suggest that diet is the main role player, further studies would need to confirm these findings (Section 5.3).

A large variety of studies have focused on the insect gut microbiome. However, the majority of studies have been dedicated to particular insect taxa. This specialised focus may be of detriment to research on understudied insect groups. This is of particular note in this study where no literature was found on *Pachysoma* gut microbiomes or specialised dung feeding insects, particularly adult scarab beetles. This means that comparisons need to be made with insects in different families/orders feeding on different substrates to that of *Pachysoma*. Due to the large number of factors causing variations in gut microbiome assembly (Engel and Moran, 2013, Douglas, 2015), comparisons across insect groups is very speculative and therefore was used with caution throughout this study. Furthermore, studies on gut microbiomes of scarab beetles typically focus on larvae (Egert et al., 2003, Lemke et al., 2003, Egert et al., 2005). Although, there is evidence that both *Pachysoma* species feed on plant material within their specific diets, only studies on adult insects were compared with as the microbiomes of larvae have been found to differ considerably from adults in holometabolous insects (Engel and Moran, 2013). This study will hopefully serve as a basis for future studies considering the gut microbiomes of *Pachysoma* species, dry dung feeders, scarab beetles and future diet-microbiome related studies.

A further limiting factor, which affects all insect-microbiome studies, is linked to methodological constraints. The complexity of mDNA extracted from insect guts complicates downstream analyses. Traditional mDNA extraction methods typically yield high concentrations of host DNA (Carpi et al., 2011, Feehery et al., 2013, Jia et al., 2013). In order to overcome this challenge, target-specific primers are commonly used. This has been successful for amplifying the bacterial 16S rRNA gene in insect-microbiome studies as

described in Chapter 3 (Franzini et al., 2016). However, as insects and fungi are closely related (Gibson and Hunter, 2010), primers targeting the fungal ITS region and 18S rRNA gene are typically non-specific, amplifying insect DNA. For this reason the majority of insect-fungal studies utilise culture-dependent methods (Gibson and Hunter, 2010). In this study, primers targeting the fungal ITS region with minimal insect coverage were used (Ihrmark et al., 2012). However, amplification of the ITS region was not successful for the plant detritus feeding *P. endroedyi* (Chapter 3; Franzini et al. (2016)). In comparison, shotgun sequencing of *Pachysoma* gut samples, where host DNA contamination was reduced, detected fungal, archaeal, bacterial and viral communities from both *Pachysoma* species. At current, few studies have utilised shotgun sequencing to study insect gut microbiomes (e.g. He et al. (2013), Jia et al. (2013), Scully et al. (2013), Shi et al. (2013)). When studying the overall microbial community structure in a broad context as in this study, shotgun sequencing is a valuable tool for uncovering phylogenetic and functional information in the absence of bias associated with amplicon sequencing (Hodkinson and Grice, 2015). However, methods which reduce host DNA contamination may skew eukaryotic microbial (i.e. fungi) results (Chapter 4). Therefore, studies focusing on fungal or other eukaryotic microorganisms would do better to utilise alternative methods.

5.3 Future study prospects

Certain suggestions can be made regarding sampling of *Pachysoma* individuals for future studies. Firstly, it is suggested that the individual beetles be collected while foraging or feeding so that the exact food-source can be collected with the respective beetle. This would allow for a parallel analysis of the microorganisms associated with both the exact food-source and the insect gut. Secondly, a large number of samples should be collected and analysed due to the substantial intra-specific variation recorded in this study (Chapter 3; Franzini et al. (2016)).

In order to truly understand the effect diet has on *Pachysoma* gut microbiome assembly it is to necessary know the chemical and material composition of the food on which the individual insects feed. The composition and nutritional quality of wet dung fed on by scarab beetles is fairly well understood (Holter, 2016). However, less information is available on the nutritional composition of dry dung, particularly faeces fed on by adult *Pachysoma* species. In depth detail would need to be known, extending to the plant species which make up the plant detritus and are found in the dung. As the chemical and material composition of plant polymers differs according to plant species (Hättenschwiler et al., 2008), it is important to identify the composition of the respective *Pachysoma* food-source. Some studies have observed the feeding behaviour of the coprophagous *P. striatum*, which appear to feed on rabbit dung (Scholtz, 1989). However, newer observations of several *Pachysoma* species should be made. Furthermore, it is believed that *P. striatum* likely feed on dung from the sheep grazing on the field where the samples are collected (C. Scholtz Pers. Comm.).

Lab-based studies have shown great promise in determining the effect of diet on gut microbiome assembly. However, it is not possible to lab-rear all insect species. *Pachysoma* species are particularly difficult to study in a laboratory environment (C. Scholtz Pers. Comm., (Harrison et al., 2003)). One such possible reason for this is their need to burrow before feeding (Scholtz, 1989, Harrison et al., 2003, Holter et al., 2009). As their burrows are quite deep (Scholtz, 1989), it would be ineffective to keep the insect in cages where variables such as environmental conditions can be controlled. However, attempts to lab-rear *Pachysoma hippocrates* noted the ability of the beetles to forage and burrow successfully although they did not breed (Harrison et al., 2003). However, it may be possible to answer the necessary research questions using specialised field-based studies. The suggested alternative would be to use shotgun sequencing to analyse the gut metagenomes of all 13 *Pachysoma* species. Of the 13 *Pachysoma* species, four are specialised plant detritus feeders, six are specialised dung feeders, one is polyphagous and the dietary specialisation of the remaining two is

unknown (C. Scholtz Pers. Comm., (Harrison et al., 2003)). Furthermore, although all 13 species are found in South Africa or Namibia (Harrison et al., 2003), species are collected from different sample sites (C. Scholtz Pers. Comm.). Therefore, by studying all 13 species, cluster analysis can be used to determine the impact of host diet, phylogeny and location. This should also give insight into the functional capacity of *Pachysoma* and the hypothesis presented in this study that the different species are functionally equipped to feed on both substrates.

Another suggested line of study would be comparing the gut microbiomes of other dry dung feeding scarab beetles, particularly focusing on genera where species feed on different materials, as in *Pachysoma*. One suggested taxa is the South American *Anomiopsoides* genera, which have been found to also include species feeding on dry dung or plant material (Ocampo, 2005). By comparing the gut microbiomes of these species with those of *Pachysoma* species it would be possible to determine differences or similarities between host feeding preferences, phylogeny and/or locality across insect taxa of the same family. However, caution would need to be taken in sampling the food-source along with the beetle and determining the composition of each individual food-source.

This study has succeeded in showing the potential that future studies in *Pachysoma* gut microbiomes would have in further understanding the relationship of host diet and microbiome assembly. There is a lot of research potential in better understanding the ecological impact of the gut microbiomes of these unusual dung beetles.

5.4 References

Ahn, J.-H., Hong, I.-P., Bok, J.-I., Kim, B.-Y., Song, J. & Weon, H.-Y., 2012. Pyrosequencing analysis of the bacterial communities in the guts of honey bees *Apis cerana* and *Apis mellifera* in Korea. *Journal of Microbiology,* 50(5), pp. 735-745.

Anderson, J.M. & Coe, M.J., 1974. Decomposition of elephant dung in an arid, tropical environment. *Oecologia,* 14(1-2), pp. 111-125.

Boissière, A., Tchioffo, M.T., Bachar, D., Abate, L., Marie, A., Nsango, S.E., et al., 2012. Midgut microbiota of the malaria mosquito vector *Anopheles gambiae* and interactions with *Plasmodium falciparum* infection. *PLoS Pathogens,* 8(5), pp. e1002742.

Brauman, A., 2000. Effect of gut transit and mound deposit on soil organic matter transformations in the soil feeding termite: a review. *European Journal of Soil Biology,* 36(3), pp. 117-125.

Broderick, N.A. & Lemaitre, B., 2012. Gut-associated microbes of *Drosophila melanogaster*. *Gut Microbes,* 3(4), pp. 307-321.

Brune, A., 2013. Symbiotic associations between termites and prokaryotes. *In:* Rosenberg, E., Delong, E.F., Lory, S., Stackebrandt, E. & Thompson, F. (eds.) *The prokaryotes.* pp. 545-577. Springer Berlin, Heidelberg.

Brune, A. & Friedrich, M., 2000. Microecology of the termite gut: structure and function on a microscale. *Current Opinion in Microbiology,* 3(3), pp. 263-269.

Carpi, G., Cagnacci, F., Wittekindt, N.E., Zhao, F., Qi, J., Tomsho, L.P., et al., 2011. Metagenomic profile of the bacterial communities associated with *Ixodes ricinus* ticks. *PLoS One,* 6(10), pp. e25604.

Coffey, L.L., Page, B.L., Greninger, A.L., Herring, B.L., Russell, R.C., Doggett, S.L., et al., 2014. Enhanced arbovirus surveillance with deep sequencing: Identification of novel rhabdoviruses and bunyaviruses in Australian mosquitoes. *Virology,* 448, pp. 146-158.

Colman, D.R., Toolson, E.C. & Takacs-Vesbach, C.D., 2012. Do diet and taxonomy influence insect gut bacterial communities? *Molecular Ecology,* 21(20), pp. 5124-5137.

Dong, Y., Manfredini, F. & Dimopoulos, G., 2009. Implication of the mosquito midgut microbiota in the defense against malaria parasites. *PLoS Pathogens,* 5(5), pp. e1000423.

Douglas, A.E., 2007. Symbiotic microorganisms: untapped resources for insect pest control. *Trends in Biotechnology,* 25(8), pp. 338-342.

Douglas, A.E., 2015. Multiorganismal insects: diversity and function of resident microorganisms. *Annual Review of Entomology,* 60, pp. 17-34.

Egert, M., Wagner, B., Lemke, T., Brune, A. & Friedrich, M.W., 2003. Microbial community structure in midgut and hindgut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology,* 69(11), pp. 6659-6668.

Egert, M., Stingl, U., Bruun, L.D., Pommerenke, B., Brune, A. & Friedrich, M.W., 2005. Structure and topology of microbial communities in the major gut compartments of *Melolontha melolontha* larvae (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology,* 71(8), pp. 4556-4566.

Engel, P. & Moran, N.A., 2013. The gut microbiota of insects–diversity in structure and function. *FEMS Microbiology Reviews,* 37(5), pp. 699-735.

Engel, P., Martinson, V.G. & Moran, N.A., 2012. Functional diversity within the simple gut microbiota of the honey bee. *Proceedings of the National Academy of Sciences,* 109(27), pp. 11002-11007.

Evans, J.D. & Schwarz, R.S., 2011. Bees brought to their knees: microbes affecting honey bee health. *Trends in Microbiology,* 19(12), pp. 614-620.

Feehery, G.R., Yigit, E., Oyola, S.O., Langhorst, B.W., Schmidt, V.T., Stewart, F.J., et al., 2013. A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS One,* 8(10), pp. e76096.

Franzini, P.Z.N., Ramond, J.-B., Scholtz, C.H., Sole, C.L., Ronca, S. & Cowan, D.A., 2016. The gut microbiomes of two *Pachysoma* MacLeay desert dung beetle species (Coleoptera: Scarabaeidae: Scarabaeinae) feeding on different diets. *PLoS One,* 11(8), pp. e0161118.

Gibson, C.M. & Hunter, M.S., 2010. Extraordinarily widespread and fantastically complex: comparative biology of endosymbiotic bacterial and fungal mutualists of insects. *Ecology Letters,* 13(2), pp. 223-234.

Harrison, J.d.G., Scholtz, C.H. & Chown, S.L., 2003. A revision of the endemic south-western African dung beetle subgenus *Scarabaeus* (*Pachysoma*) MacLeay, including notes on other flightless Scarabaeini (Scarabaeidae: Scarabaeinae). *Journal of Natural History,* 37(3), pp. 305-355.

Hättenschwiler, S., Aeschlimann, B., Coûteaux, M.M., Roy, J. & Bonal, D., 2008. High variation in foliage and leaf litter chemistry among 45 tree species of a neotropical rainforest community. *New Phytologist,* 179(1), pp. 165-175.

He, S., Ivanova, N., Kirton, E., Allgaier, M., Bergin, C., Scheffrahn, R.H., et al., 2013. Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS One,* 8(4), pp. e61126.

Hodkinson, B.P. & Grice, E.A., 2015. Next-generation sequencing: a review of technologies and tools for wound microbiome research. *Advances in Wound Care,* 4(1), pp. 50-58.

Holter, P., 2016. Herbivore dung as food for dung beetles: elementary coprology for entomologists. *Ecological Entomology,* 41(4), pp. 367-377.

Holter, P. & Scholtz, C.H., 2011. Re-establishment of biting mouthparts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on plant litter—old structures reacquired or new ones evolved? *Journal of Morphology,* 272(8), pp. 1007-1016.

Holter, P. & Scholtz, C.H., 2013. Elongated hindguts in desert-living dung beetles (Scarabaeidae: Scarabaeinae) feeding on dry dung pellets or plant litter. *Journal of Morphology,* 274(6), pp. 657-662.

Holter, P., Scholtz, C.H. & Wardhaugh, K.G., 2002. Dung feeding in adult scarabaeines (tunnellers and endocoprids): even large dung beetles eat small particles. *Ecological Entomology,* 27(2), pp. 169-176.

Holter, P., Scholtz, C.H. & Stenseng, L., 2009. Desert detritivory: nutritional ecology of a dung beetle (*Pachysoma glentoni*) subsisting on plant litter in arid South African sand dunes. *Journal of Arid Environments,* 73(12), pp. 1090-1094.

Ihrmark, K., Bödeker, I., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., et al., 2012. New primers to amplify the fungal ITS2 region–evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiology Ecology,* 82(3), pp. 666-677.

Jia, S., Zhang, X., Zhang, G., Yin, A., Zhang, S., Li, F., et al., 2013. Seasonally variable intestinal metagenomes of the red palm weevil (*Rhynchophorus ferrugineus*). *Environmental Microbiology,* 15(11), pp. 3020-3029.

Karasov, W.H., Martínez del Rio, C. & Caviedes-Vidal, E., 2011. Ecological physiology of diet and digestive systems. *Annual Review of Physiology,* 73, pp. 69-93.

Lemke, T., Stingl, U., Egert, M., Friedrich, M.W. & Brune, A., 2003. Physicochemical conditions and microbial activities in the highly alkaline gut of the humus-feeding larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Applied and Environmental Microbiology,* 69(11), pp. 6650-6658.

Nichols, E., Spector, S., Louzada, J., Larsen, T., Amezquita, S., Favila, M. & Network, T.S.R., 2008. Ecological functions and ecosystem services provided by Scarabaeinae dung beetles. *Biological Conservation,* 141(6), pp. 1461-1474.

Ocampo, F.C., 2005. Revision of the southern South American endemic genus *Anomiopsoides* Blackwelder, 1944 (Coleoptera: Scarabaeidae: Scarabaeinae: Eucraniini) and description of its food relocation behavior. *Journal of Natural History,* 39(27), pp. 2537-2557.

Ohkuma, M., 2003. Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Applied Microbiology and Biotechnology,* 61(1), pp. 1-9.

Osei-Poku, J., Mbogo, C.M., Palmer, W.J. & Jiggins, F.M., 2012. Deep sequencing reveals extensive variation in the gut microbiota of wild mosquitoes from Kenya. *Molecular Ecology,* 21(20), pp. 5138-5150.

Sabree, Z.L., Hansen, A.K. & Moran, N.A., 2012. Independent studies using deep sequencing resolve the same set of core bacterial species dominating gut communities of honey bees. *PLoS One,* 7(7), pp. e41250.

Scholtz, C.H., 1989. Unique foraging behaviour in *Pachysoma*. *Journal of Arid Environments,* 16(3), pp. 305-313.

Scully, E.D., Geib, S.M., Hoover, K., Tien, M., Tringe, S.G., Barry, K.W., et al., 2013. Metagenomic profiling reveals lignocellulose degrading system in a microbial community associated with a wood-feeding beetle. *PLoS One,* 8(9), pp. e73827.

Shi, W., Xie, S., Chen, X., Sun, S., Zhou, X., Liu, L., et al., 2013. Comparative genomic analysis of the endosymbionts of herbivorous insects reveals eco-environmental adaptations: biotechnology applications. *PLoS Genetics,* 9(1), pp. e1003131.

Sun, J.-Z. & Scharf, M.E., 2010. Exploring and integrating cellulolytic systems of insects to advance biofuel technology. *Insect Science,* 17(3), pp. 163.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., et al., 2009. A core gut microbiome in obese and lean twins. *Nature,* 457(7228), pp. 480-484.

Wang, A., Yao, Z., Zheng, W. & Zhang, H., 2014. Bacterial communities in the gut and reproductive organs of *Bactrocera minax* (Diptera: Tephritidae) based on 454 pyrosequencing. *PLoS One,* 9(9), pp. e106988.

Wang, S., Ghosh, A.K., Bongio, N., Stebbings, K.A., Lampe, D.J. & Jacobs-Lorena, M., 2012. Fighting malaria with engineered symbiotic bacteria from vector mosquitoes. *Proceedings of the National Academy of Sciences,* 109(31), pp. 12734-12739.

Wang, Y., Gilbreath III, T.M., Kukutla, P., Yan, G. & Xu, J., 2011. Dynamic gut microbiome across life history of the malaria mosquito *Anopheles gambiae* in Kenya. *PLoS One,* 6(9), pp. e24767.

Wharton, D.A., 2002. *Life at the limits: organisms in extreme environments*. Cambridge University Press UK, Cambridge.

Wong, A.C.-N., Chaston, J.M. & Douglas, A.E., 2013. The inconstant gut microbiota of *Drosophila* species revealed by 16S rRNA gene analysis. *The ISME Journal,* 7(10), pp. 1922-1932.

Wong, A.C.-N., Luo, Y., Jing, X., Franzenburg, S., Bost, A. & Douglas, A.E., 2015. The Host as the Driver of the Microbiota in the Gut and External Environment of *Drosophila melanogaster*. *Applied and Environmental Microbiology,* 81(18), pp. 6232-6240.

Yun, J.-H., Roh, S.W., Whon, T.W., Jung, M.-J., Kim, M.-S., Park, D.-S., et al., 2014. Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied and Environmental Microbiology,* 80(17), pp. 5254-5264.