COVERAGE IMPACTS BIOMASS COMPOSITION, CONVERSION TO ETHANOL YIELDS AND MICROBIAL COMMUNITIES DURING STORAGE

by

ANNE R. RIGDON

B.S., South Dakota State University, 2006
M.S., University of Nebraska, 2009

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Grain Science and Industry
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2013

Abstract

Increased mandates for the production of transportation fuels from renewable resources have thrust the conversion of lignocellulosic biomass, e.g., energy crops and agricultural residues, to ethanol into commercial production. The conversion of biomass to ethanol has been implemented; transportation and storage logistics are still obstacles to overcome by industry. Limited harvest windows throughout the year necessitate extended periods of biomass storage to maintain a consistent, year-round supply to the biorefinery. Sorghum biomass stored with no coverage (NN), covered with tarp (NT), wrapped in plastic (PN) and covered with a tarp and wrapped in plastic (PT) for six months was analyzed for changes in biomass components—cellulose, hemicellulose and lignin, cellulose and hemicellulose degrading enzymes, and conversion to ethanol yields. Treatment NN had increased enzyme activity, and reduced cellulose content and ethanol yields; while biomass covered maintained enzyme activity, cellulose content and ethanol yields. Sequencing of the Large SubUnit (LSU) region and the internal transcribed spacer (ITS) regions of ribosomal RNA gene gave consistent results of fungal community dynamics in biomass stored as previously described. Fungal community richness and diversity increased, while evenness decreased in uncovered biomass during storage. Covered and uncovered storage treatments and over time were found to exhibit distinctly different fungal communities. In contrast, bacterial communities were found to be unresponsive to storage treatments and durations. *Cladosporium*, *Alternaria* and *Cryptococcus* were found to be the most abundant in the stored biomass. Covering of biomass strongly limits the arrival and establishment of new fungal propagules in stored biomass, reducing biomass degradation by these often pathogenic, saprobic or endophytic communities. Overall, covering of biomass during storage is essential for optimal substrate retention for downstream processing into ethanol. In addition, storage and transportation logistics of three real-world scenarios were evaluated for the conversion of wheat straw, corn stover and sorghum stalks residues to ethanol at a biorefinery located in Southwest Kansas. Economic evaluation revealed that transport and storage of residues at satellite storage facilities was most economical for farmers and would create opportunity for the operation of profitable facilities that would supply the local biorefinery on demand throughout the year.

COVERAGE IMPACTS BIOMASS COMPOSITION, CONVERSION TO ETHANOL YIELDS AND MICROBIAL COMMUNITIES DURING STORAGE

by

ANNE R. RIGDON

B.S., South Dakota State University, 2006
M.S., University of Nebraska, 2009

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Grain Science and Industry
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2013

Approved by:

Major Professor
Dr. Dirk E. Maier

Copyright

Abstract

Increased mandates for the production of transportation fuels from renewable resources have thrust the conversion of lignocellulosic biomass, e.g., energy crops and agricultural residues, to ethanol into commercial production. The conversion of biomass to ethanol has been implemented; transportation and storage logistics are still obstacles to overcome by industry. Limited harvest windows throughout the year necessitate extended periods of biomass storage to maintain a consistent, year-round supply to the biorefinery. Sorghum biomass stored with no coverage (NN), covered with tarp (NT), wrapped in plastic (PN) and covered with a tarp and wrapped in plastic (PT) for six months was analyzed for changes in biomass components—cellulose, hemicellulose and lignin, cellulose and hemicellulose degrading enzymes, and conversion to ethanol yields. Treatment NN had increased enzyme activity, and reduced cellulose content and ethanol yields; while biomass covered maintained enzyme activity, cellulose content and ethanol yields. Sequencing of the Large SubUnit (LSU) region and the internal transcribed spacer (ITS) regions of ribosomal RNA gene gave consistent results of fungal community dynamics in biomass stored as previously described. Fungal community richness and diversity increased, while evenness decreased in uncovered biomass during storage. Covered and uncovered storage treatments and over time were found to exhibit distinctly different fungal communities. In contrast, bacterial communities were found to be unresponsive to storage treatments and durations. *Cladosporium*, *Alternaria* and *Cryptococcus* were found to be the most abundant in the stored biomass. Covering of biomass strongly limits the arrival and establishment of new fungal propagules in stored biomass, reducing biomass degradation by these often pathogenic, saprobic or endophytic communities. Overall, covering of biomass during storage is essential for optimal substrate retention for downstream processing into ethanol. In addition, storage and transportation logistics of three real-world scenarios were evaluated for the conversion of wheat straw, corn stover and sorghum stalks residues to ethanol at a biorefinery located in Southwest Kansas. Economic evaluation revealed that transport and storage of residues at satellite storage facilities was most economical for farmers and would create opportunity for the operation of profitable facilities that would supply the local biorefinery on demand throughout the year.

Table of Contents

List of Figures

List of Tables

Acknowledgements

This material is based upon work supported by National Science Foundation Grant: From Crops to Commuting: Integrating the Social, Technological, and Agricultural Aspects of Renewable and Sustainable Biorefining (I-STAR); NSF Award No.: DGE-0903701 and funding from the Kansas Alliance for BioRefining and BioEnergy (KABB).

I would first like to thank Dr. Dirk Maier for his guidance, support and positive encouragement throughout my doctoral work. He allowed me to work independently and collaboratively within multiple research groups throughout the university. He also encouraged me to step outside my comfort zone by presenting at an international conference and traveling internationally to Europe and Brazil. Without his nomination for the IGERT Fellowship, I may not have had so many opportunities for collaborative research, international travel and professional development.

I would also like to thank my committee, including Dr. Ari Jumpponen, Dr. Jason Bergtold, Dr. Praveen Vadlani and Dr. Scott Staggenborg, who also deserve full acknowledgements for their support and guidance; in addition, to the support and guidance from my outside chair, Dr. Walter Dodds. Each committee member personally assisted with my research and helped me develop personally and professionally. A special thanks to Ari for being a constant positive and motivator during some of the most overwhelming and trying times I've endured.

My deepest thanks to all who have worked with me, including past and present labmates in the Stored Product Protection and Technology Research Group in Grain Science and in the Fungal Ecology Lab Group in the Division of Biology, my friends in the Grain Science and Industry Department and various other people I have met at K-State, throughout Kansas and during my international travels. My gratitude is with my friends and colleagues, with special gratitude to Vernon Schaffer and the staff and students of the Agronomy North Farm, Brian Bartels, Ben Smart, Dave Trumble, Shawn Brown, Alina Akhunova and Hanquan Liang of the KSU Integrated Genomics Facility, Dr. Ron Madl, Jonathan Huss, Dr. Mary Rezac and Keith Rutlin. Personally, I would like to thank my family and friends for their never ending love and understanding, as I never would have made it this far without their encouragement and guidance.

Dedication

I dedicate this to my high school guidance counselor, who once told me ‘college probably isn’t for you’; those words echoing through my mind have given me the motivation to set goals high and the determination to reach them.

Preface

All chapters were written for a specified journal. Therefore, the required journal format was followed for each manuscript, with references following the discussion and figures and tables coming after references in Chapters 3-5. Chapters 1 and 6 were not submitted for publication but followed the same format as the other chapters for consistency.

Chapter 1 - Introduction

The Energy Independence and Security Act (EISA) of 2007 initiated the Renewable Fuel Standards (RFS) program, with the goal of limiting our reliance on petroleum-based transportation fuels from foreign nations and reducing the environmental impacts of a petroleum-based society. The RFS program outlined goals for the production and utilization of transportation fuels, primarily ethanol, from renewable resources, with the goal of shifting current production from first generation feedstocks (maize) to second generation feedstocks (lignocellulosic biomass). Lignocellulosic materials considered for ethanol production include agricultural residues (corn stover, sorghum stalks, wheat straw), short rotation forestry crops and residues, perennial grasses and dedicated energy crops. Many of these crops can only be harvested annually or biannually, requiring varying lengths of storage prior to the conversion to ethanol. Until recently, research related to the storage of lignocellulosic materials was limited to feedstocks primarily preserved for use by livestock, mostly cattle. Due to this, most of the research focused on preserving or increasing the digestibility of the stored feedstock, not including the production of ethanol. In addition, little research has been done to characterize the microbial populations present in stored feedstock, mostly due to limitations in microbiological methodologies. With the advent of high-throughput next-generation sequencing technology, characterizing microbial populations from a mixed sample has become routine. Next-generation sequencing allows for the exploration of microbial populations and population shifts in biomass stored for ethanol production and how these populations may impact ethanol production. Evaluation of the microbial populations present under various storage conditions could help in the identification of ideal storage parameters for maximum ethanol yields.

Initially, transportation fuels-ethanol and biodiesel-produced from renewable resources in the US have followed the first generation pathway, shown in Figure 1.1. Biodiesel production requires the transesterification of oil extracted from common oil crops (soybeans, rapeseed or sunflowers) to produce fatty acid methyl esters or biodiesel. Ethanol produced from grain crops, primarily corn, requires starch to be enzymatically hydrolyzed to glucose for fermentation to ethanol by *Saccharomyces cerevisiae* (Naik et al., 2010). The production of ethanol from grain crops increased over the last decade due to the dramatic increase in petroleum-based transportation fuels. This spike in petroleum also made the biochemical conversion of starch to

ethanol, as a fuel additive, cost-competitive with petroleum-based fuel (Naik et al., 2010). With the explosion of the 'grain alcohol' market, the new debate of food-versus-fuel was ignited, indicating that increased production of ethanol from grains resulted in rising food prices. This debate, in conjunction with production mandates outlined in the RFS, has fueled the research, development and implementation of ethanol production from second generation feedstocks or lignocellulosic biomass.

Second generation biofuels can be produced via two methods, thermochemically and biochemically, as shown in figure 1.2, utilizing plant biomass or lignocellulosic material. Both methods allow for the utilization of the same feedstocks but only the biochemical approach using lignocellulosic materials will be fully detailed. Plant biomass has been identified as a promising source for the production of liquid biofuels due to its raw abundance and renewability. Estimates from Perlack et al.'s (2005) billion-ton study revealed 194 million dry tons of available biomass could be harvested every year. Short rotation forestry crops (poplar, willow and eucalyptus), perennial grasses (miscanthus, switch grass and reed canary grass) and residues from the wood industry, forestry and agriculture (wheat straw, sorghum stalks, and corn stover) have been identified as potential feedstocks for the production of second generation biofuels (Naik et al., 2010; Gomez et al., 2008). Shifting the production of ethanol from first generation grains to second generation lignocellulosic feedstocks would allow more sustainable production practices, reduced CO_2 and greenhouse gas (GHG) emissions and less stress on commodity prices (Stevens and Verhe, 2004).

The biochemical conversion of plant biomass (lignocellulosic material) is a three-step process consisting of pretreatment, enzymatic hydrolysis and fermentation. Along with physical conversion of lignocellulosic biomass to ethanol, many other factors must be taken under consideration for efficient operation of the biorefinery. These factors include feedstock availability, feedstock harvest parameters, feedstock transportation to a biorefinery, and storage requirements to maintain feedstock quality. Logistics of feedstock harvest, storage, handling and transport are the greatest challenges to making lignocellulosic ethanol production cost-competitive to grain ethanol (Hess et al., 2007). Current infrastructure for the transport, storage and handling of grains will not be suitable for biomass feedstock variability, handling properties and bulk density. The commercial biofuels industry will need to develop a new infrastructure or modify the current grain-based one to make biomass storage and transport economical.

A significant amount of lignocellulosic biomass will need to be stored to ensure enough material is available for continuous production operation of the biorefinery, since most biomass can be harvested annually or biannually. The main goal of biomass storage is to prevent losses and to maintain biomass quality for maximum ethanol yield. Storage of biomass in a dried bale form is easier than storage at high moisture contents (Sokhansanj and Hess, 2009). Loss of biomass can occur during or between each unit operation and can be physical or chemical losses. Physical losses can be attributed to machine-induced losses due to moisture content at time of harvest, yield, design features of the machine used, contours of the field and prevailing weather conditions. Breakdown of structural and nonstructural carbohydrates due to abiotic (oxidative reactions or pyrolysis) or biotic (microbial metabolism or respiration) reactions result in chemical losses (Sokhansanj and, Hess 2009). Studies by Shinners and Binversie (2004) found dry matter losses of corn stover bales stored outdoors ranged from 7.0-38.5% after nine months. However, few studies have fully detailed the resulting ethanol yields from biomass obtained from perennial grasses, dedicated energy crops, and agricultural residues when stored under dry conditions over an extended period of time (Emery and Mosier, 2012). Furthermore, abiotic and biotic factors influencing degradation in stored biomass and the resulting impact on ethanol yields have also not been fully evaluated and reported.

Many fungal taxa are able to degrade plant-derived carbon compounds, like cellulose and lignin, making stored biomass a nutrient-rich substrate ideal for colonization by these diverse fungi. Along with fungi, multiple bacterial strains are capable of pathogenicity in plants (Schaad et al., 2001). The impacts of fungal and bacterial colonization of biomass during storage and the resulting impacts on downstream processing, like conversion to ethanol, have not been fully evaluated. Characterization of these fungal and bacterial communities using traditional microbiological methodologies relies on culturing techniques done in the laboratory. However, studies have revealed that more than 99% of the microorganisms (bacteria, archaea and fungi) from environmental samples remain ‘unculturable’ in the laboratory. Due to this, determining the roles and functions these organisms influence in their natural settings is nearly impossible (Sharma et al., 2005). Our inability to cultivate these organisms in the laboratory setting has led to the development of culture-independent techniques, based on DNA sequencing technologies or next-generation sequencing technologies (NGS). Advances in sequencing technologies have grown in leaps and bounds in the last two decades, with pyrosequencing (Roche 454) becoming

one of the most powerful tools available for exploring microorganisms from various sampling environments. NGS offers an enormous volume of data cheaply, ranging from one million to one billion short reads per instrument run (Metzker, 2010).

With the advent of NGS technologies, nuclear ribosomal DNA (rDNA) markers are widely used in monitoring fungal and bacterial community shifts in a system. The choice of sequencing locus is extremely important to fully survey the fungal and bacterial diversity present in a system. In fungi, rDNA includes the small subunit (SSU, 18S), internal transcribed spacer (ITS, ITS1+5.8S+ITS2), and large subunit (LSU, 25-28S) regions (Porter and Golding, 2012). Traditionally, ITS has been targeted in amplicon-based environmental sequencing and has been used extensively in molecular ecology studies (Kruger et al., 2012). The amount of ITS sequence data available in public databases is rapidly expanding, allowing for better annotation and ecological insight. ITS has been the locus of choice for analyzing phylogenetic relationships in fungal communities, primarily due to the ease of designing both broad and selective Polymerase Chain Reaction (PCR) primers for it (see Appendix A). In addition, the wide use of ITS for taxonomy and phylogenetic relationships has also been due to the ease of amplification from small quantities of DNA (due to high copy number of rRNA genes) and the high degree of variation even between closely related species (Kruger et al., 2012). The major caveats of ITS-sequence data analysis are the inability of the sequence data to be aligned and the variability in length of the gene region. Development of primers for LSU-targeted sequencing of the D1/D2 region have been utilized for taxonomic identification in environmental samples (Porter and Golding, 2012). Furthermore, LSU-targeted sequencing results in sequence data that can be aligned to a reference alignment to determine phylogenetic relationships. Public databases have also been expanding rapidly in the amount of LSU sequence data available.

Bacterial taxa can be identified based on DNA encoding the 16S rRNA gene, since regions of the gene can be amplified using PCR primers that bind to conserved sites in most or all species (Liu et al., 2007). Again, large public databases are available for determining bacterial phylogenies based on 16S rRNA sequences. The full-length 16S rRNA gene is roughly 1500 bp and can be used for accurate taxonomic identification and sequence diversity among a bacterial community (Vilo and Dong, 2012). However, NGS technologies, like Roche 454, only have sequence reads of 100-500 bp. Therefore, fragments of the larger gene must be obtained for sequencing, preferable containing variable regions. Due to the extensive nature of the 16S rRNA

gene sequence database, primer selection for amplicon-sequencing plays a less important role, as seen with fungal community assessment.

With the tremendous amount of data generated from sequencing, a multitude of sequence analysis tools are available, primarily through open-source. Analysis of sequencing data is dependent on bioinformatics, which utilizes various statistical and computer algorithms techniques to make biological inferences about the sequences obtained. Multiple factors must be considered when determining which sequence analysis tool to use. The open-source, platform-independent, community-supported program mothur was identified as an ideal pipeline for the analysis of both fungal and bacterial sequence data. The mothur pipeline has the multiple calculators for quantifying key ecological parameters (α and β diversity); visualization tools (Venn diagrams, heat maps, dendrograms); functions for screening sequence collections based on quality; sequence alignment; pairwise sequence alignment and distance calculator; and the ability to call individual commands, either from within mothur or directly from the command line, allowing greater flexibility in analysis (Schloss et al., 2009).

Understanding the abiotic and biotic changes occurring in stored biomass destined for lignocellulosic ethanol production is essential for maximizing ethanol yields and maintaining producer and processer profitability. Maintaining substrate quality for conversion to ethanol during storage can be done through simple storage practices, like coverage with a tarp or wrapping in plastic. Characterization of fungal and bacterial communities present in stored biomass could help decide management practices during long-term storage of biomass destined for lignocellulosic ethanol production.

Economic Considerations for Lignocellulosic Ethanol Production

Multiple unit operations must be undertaken to release the energy trapped in plants into a useable transportation fuel. This is only possible when the benefits of production and utilization outweigh the cost of production. Economic evaluation of field to fuel costs, allows for the feasibility and benefits of lignocellulosic ethanol production to be weighed. Economic evaluation of lignocellulosic ethanol also allows for pin-pointing areas of production that need further optimization for reducing processing costs.

A multitude of factors must be considered when determining exactly how much it will cost a farmer to produce biomass for a lignocellulosic biorefinery. The farmer must consider

direct and fixed expenses incurred during cultivation. Direct costs to consider include fertilizer, herbicides and insecticides, seed, operator and manual labor, fuel for equipment and equipment repair and maintenance, while fixed expenses include cost of implements, tractors and land (Linton et al., 2011). These expenses can vary dramatically depending on which biomass feedstock is used and what harvest, collection and storage methods the farmer chooses to utilize. The costs incurred by the farmer for production must be less than what the biorefinery is willing to pay for the biomass.

Cost considerations for the processor are vastly different from those of the producer. The biorefinery must be able to run pretreatment, enzymatic hydrolysis and fermentation operations, in the case of ethanol production, at optimal levels to ensure the production of ethanol that can be sold competitively with petroleum-based fuels currently on the market. However, both the producer and processor must take into consideration the impact storage has on biomass quality for conversion to ethanol, as a biorefinery requires a continuous supply of consistent quality biomass to operate efficiently and economically.

References

Emery, I.R., Mosier, N.S., 2012. The impact of dry matter loss during herbaceous biomass storage on net greenhouse gas emissions from biofuels production. Biomass Bioenerg. 39, 237-246.

Gomez, L.D., Steele-King, C.G., McQueen-Mason, S.J., 2008. Sustainable liquid biofuels from biomass: the writing's on the walls. New Phytol., 178, 473-485.

Hess, J.R., Wright, C.T., Kenney, K.L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel Biop. Bior., 1, 181-190.

Kruger, D., Kapturska, D., Fischer, C., Daniel, R., Wubet, T., 2012. Diversity measures in environmental sequences are highly dependent on alignment quality—data from ITS and new LSU primers targeting basidiomycetes. PLoS ONE 7:e32139-.

Linton, J.A., Miller, J.C., Little, R.D., Petrolia, D.R., Coble, K.H., 2011. Economic feasibility of producing sweet sorghum as an ethanol feedstock in the southeastern United States. Biomass Bioenerg. 35, 3050-3057.

Liu, Z., Lozupone, C., Hamady, M., Bushman, F.D., Knight, R., 2007. Short pyrosequencing reads suffice for accurate microbial community analysis. Nucleic Acids Res. 35:e120-.

Metzker, M.L., 2010. Sequencing technologies- the next generation. Nat. Rev. Genet., 11, 31-46.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: a comprehensive review. Renew. Sust. Energ. Rev. 14, 578-597.

Perlack, R.D., Wright, L.L., Turhollow, A.F., Graham, R.L., Stokes, B.J., Erback, D.C., 2005. Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. US Department of Energy DE-AC05-000R22725.

Porter, T.M., Golding, G.B., 2012. Factors that affect large subunit ribosomal DNA amplicon sequencing studies of fungal communities: classification method, primer choice, and error. PLoS ONE 7:e35749-.

Schaad, N.W., Jones, J.B., Chun, W., 2001. Laboratory guide for identification of plant pathogenic bacteria, third edition. APS Press.

Schloss, P.D., et al. 2009. Introduction mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. App. Environ. Microb. 75, 7537-7541.

Sharma, R., Ranjan, R., Kapardar, R.K., Grover, A., 2005. 'Unculturable' bacterial diversity: an untapped resource. Curr. Sci. India, 89, 72-77.

Shinners, K.J., Binversie, J.B.N., 2004. Harvest and storage of wet corn stover biomass. American Society of Agricultural Engineers Annual Meeting 2004. ASAE041159. ASAE, St. Joseph, Mich.

Sokhansanj, S., Hess, J.R., 2009. Biomass supply logisitics and infrastructure. In: Mielenz, J.R. (Ed.), *Biofuels: methods and protocols, methods in molecular biology*. 581, 1-25.

Stevens, C.V., Verhe, R., 2004. Renewable bioresources scope and modification for non-food applications. England: John Wiley and Sons Ltd.

Vilo, C., Dong, Q., 2012. Evaluation of RDP classifier accuracy using 16S rRNA gene variable regions. Metagenomics 1: doi:10.4303/mg/235551.

Figures and Tables

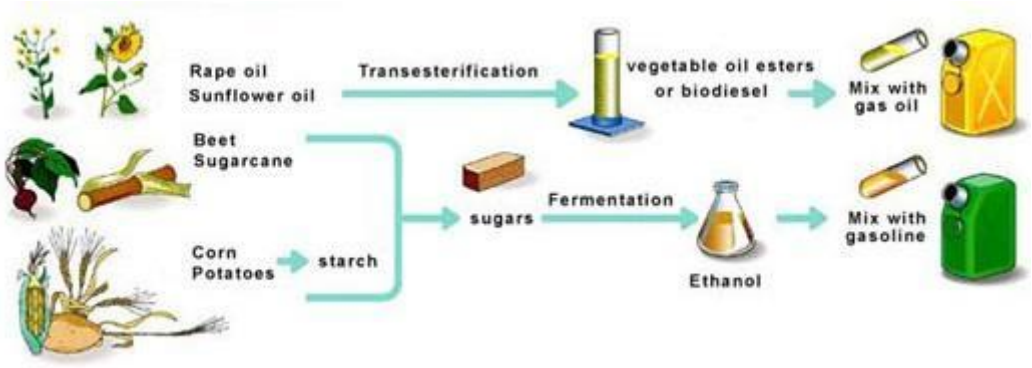


Figure 1.1 Generalized production schemes of first generation biofuels—biodiesel and bioethanol. (http://refuelingthefuture.yolasite.com/first-generation-biofuels.php)

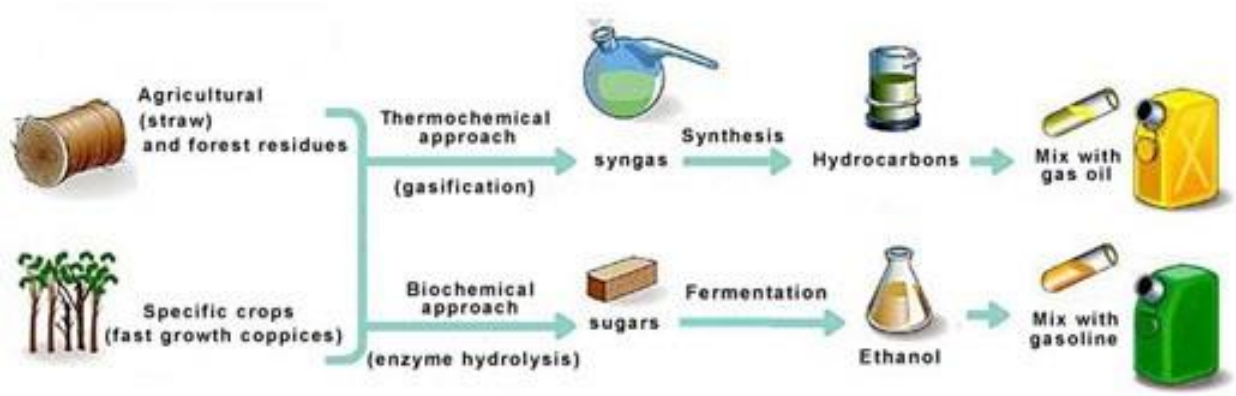


Figure 1.2 Generalized production scheme of second generation biofuels utilizing thermochemical and biochemical methods. (http://refuelingthefuture.yolasite.com/second-generation-biofuels.php)

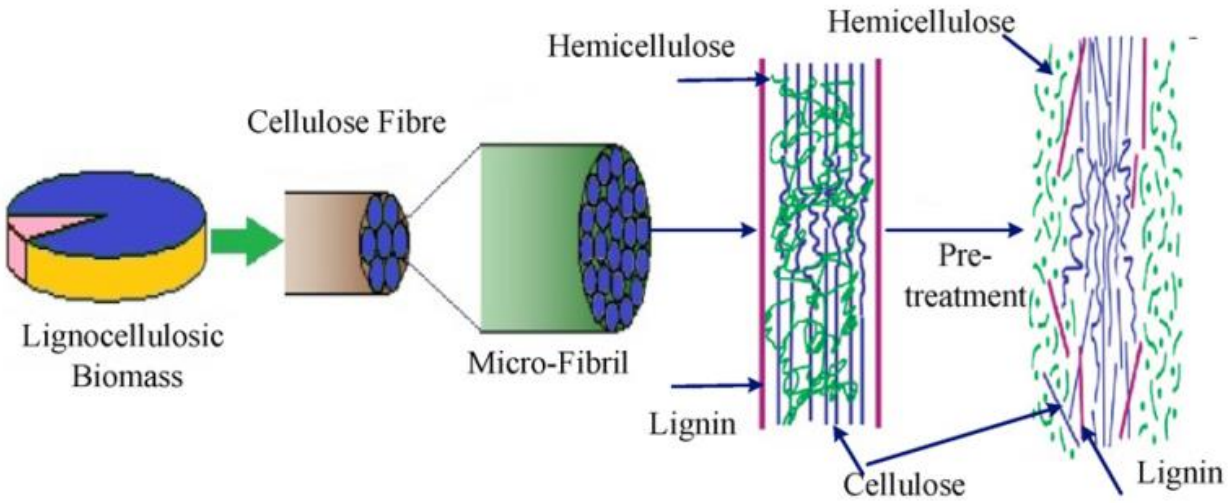


Figure 1.3 Schematic representation of the cellulose-hemicellulose-lignin matrix and the impact of pretreatment. (Chandra et al., 2012)

Chapter 2 - Research Objectives

The overall research objective of this study was to evaluate the changes in stored sorghum biomass destined for lignocellulosic ethanol production. A storage study was conducted using baled photoperiod-sensitive sorghum with four different storage treatments for a duration of six-months, with sampling every two months. Baled sorghum biomass was stored with no coverage (NN), covered with a tarp (NT), wrapped in plastic (PN) or wrapped in plastic and covered with a tarp (PT) with sampling at 0, 2, 4 or 6 months. Specific research objectives with corresponding chapter are outlined below:

- Physical characterization of the sorghum biomass, including biomass components—dry matter, cellulose, hemicellulose and lignin; enzymatic activity—enzymes associated with cellulose and hemicellulose degradation; conversion to ethanol—glucose and ethanol yields, were analyzed for each storage treatment at each sampling point. (Chapter 3)
- Fungal community dynamics were monitored using direct 454-pyrosequencing of an amplicon library produced from the Large SubUnit (LSU) region of the ribosomal RNA gene. (Chapter 4)
- Monitoring of fungal and bacterial communities using direct 454-pyrosequencing of amplicon libraries produced from the internal transcribed spacer (ITS) between the small subunit (5.8S) and large subunit ribosomal RNA genes and the VI-VII region of the 16S ribosomal RNA gene, respectively. (Chapter 5)
- Economic evaluation of storage and transport of baled biomass to a biorefinery using three different delivery scenarios. (Chapter 6)

Chapter 3 - Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop[1]

Abstract

With increased mandates for biofuel production in the US, ethanol production from lignocellulosic substrates is burgeoning, highlighting the need for thorough examination of the biofuel production supply chain. This research focused on the impact storage has on biomass, particularly photoperiod-sensitive sorghum biomass. Biomass quality parameters were monitored and included biomass components, cellulose, hemicellulose and lignin, along with extra-cellular enzymatic activity (EEA) responsible for cellulose and hemicellulose degradation and conversion to ethanol yields. Analyses revealed dramatic decreases in uncovered treatments, specifically reduced dry matter content from 88% to 59.9%, cellulose content from 35.3% to 25%, hemicellulose content from 23.7% to 16.0% and ethanol production of 0.20 g L^{-1} to 0.02 g L^{-1} after 6 months storage along with almost double EEA activities. In contrast, biomass components, EEA and ethanol yields remained relatively stable in covered treatments, indicating covering of biomass during storage is essential for optimal substrate retention and ethanol yields.

Key Words: lignocellulosic ethanol production; ethanol; biomass; storage; extra-cellular enzyme activity; biofuels; composition.

[1] Published: Rigdon, A.R., Jumpponen, A., Vadlani, P.V., Maier, D.E., 2013. Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop. Bioresour. Techno. 132, 269-275.

Introduction

In 2007, the United States government signed into action the Energy Independence and Security Act (EISA), which expanded the Renewable Fuel Standards (RFS) program and outlined government mandated goals for the utilization and production of renewable fuels in the United States by 2022. In 2010 the revised RFS2 mandates outlined that 136 billion liters of renewable fuels be used in the US and 60.5 of 136 billion liters be produced using the cellulosic platform or second generation technologies, along with goals for reduced greenhouse gas (GHG) emissions (Gao et al., 2011). Second generation biofuels are those produced from lignocellulosic materials, including agricultural and forestry residues, solid waste, perennial woody and herbaceous energy crops (Gibbons and Hughes, 2009). The driving factors of the economic competitiveness of cellulosic ethanol production are feedstock cost, availability, storage and transportation expenses (Hess et al., 2007). Estimates from Hess et al. (2007) value feedstock costs to be 35-50% of total ethanol production costs, with the logistics associated with moving the biomass to the biorefinery comprising 50-75% of the feedstock costs. These high logistical costs reduce the potential profit margins for biomass producers and biorefinery operators and are therefore a serious concern (Hess et al., 2007).

A majority of the potential lignocellulosic feedstocks are harvested annually or bi-annually, resulting in extended storage to provide a continuous supply to a biorefinery. As a result, a large proportion of the biomass must be stored and preserved to provide a consistent, year-round feedstock supply to the biorefineries. Identifying the most cost effective method to minimize dry matter losses is essential. Minimizing dry matter loss during storage could play a key role in reducing the total biomass costs for the biomass producer and biorefinery operator and be crucial for the overall profitability of these operations. Biomass can be stored under wet or dry methods, both having advantages and disadvantages. Zheng et al. (2012) found that ensilage (wet storage) of sugar beet pulp (SBP) was an effective combined storage and pretreatment method that resulted in increased ethanol yields (0.2 g ethanol/g SBP) via fermentation with *E. coli* KO11. However, few studies have fully detailed the resulting ethanol yields from biomass stored under dry conditions over an extended period of time or monitored changes occurring during that storage period (Emery and Mosier, 2012).

Extra-cellular enzymes are the main mediators of soil biological processes, including organic matter degradation, mineralization and nutrient cycling (Marx et al., 2001). Substrate

degradation and availability to microbial or plant uptake are believed to be controlled by hydrolytic enzymes (Marx et al., 2001). Monitoring enzymatic activity in stored biomass, specifically related to organic matter degradation, can be used as an indicator of biomass quality. Monitoring the enzymatic degradation of cellulose and hemicellose is of greatest interest for conversion of plant biomass to ethanol. A wide variety of methods have been developed for estimating enzyme activities in soil samples. These methods vary in substrate choices, assay conditions, incubation time and detection methods (e.g., colorimetric, fluorimetric, radiolabelled) (Marx et al., 2001). The fluorescent compound 4-methylumbelliferone (MUB) has great advantages over other enzyme assays and the assay principles have been previously described (Darrah and Harris 1996; Freeman et al., 1995; Hoppe 1983; Marx et al., 2001; Somville 1984). These advantages include high fluorescence of the MUB conjugate allowing minimal quantities to be detected, no reported inhibition or facilitation of the enzymatic activity, and measured enzymatic activity using MUB substrates is similar to the natural processes (Marx et al., 2001). The sensitivity of MUB-linked substrates also allows for the utilization of a 96-well plate format, which conveniently allows rapid measurement of activity for a range of enzymes and a large number of samples.

The goal of this study was to investigate the potential impact of four dry storage methods of sorghum biomass on the conversion to ethanol. We monitored the changes in biomass components, extra-cellular enzyme activity, and conversion to ethanol yields during each storage condition and duration. These studies critically evaluated the effect storage has on biomass components, including dry matter, cellulose, hemicellulose and lignin and enzymes specifically associated with cellulose and hemicellulose degradation.

Materials and Methods

All analyses were carried out as described below. After bale grinding, sub-samples for biomass component or conversion to ethanol and enzymatic activity analyses were immediately stored at -20°C or -80°C, respectively, until analysis.

Feedstock

Photo-period sensitive sorghum cultivar PS1990, a forage hybrid, from Sorghum Partners (Sorghum Partners, LLC, New Deal, TX) was grown on the North Agronomy Farm, Department of Agronomy, Kansas State University in Manhattan, KS. The sorghum was planted on June 3,

2010. Nitrogen was applied 15 to 20 days after planting at a rate of 100 pounds per acre using urea (0-46-0) and herbicide Bicep II Magnum applied at 2.4 quarts per acre. Sorghum was cut and windrowed on October 7, 2010 and allowed to field dry. Cut and windrowed sorghum was baled into small square bales (average size 0.36 by 0.46 by 0.91 m) on October 15, 2010 using a Massey Ferguson (AGCO, Duluth, GA) square baler.

Storage Treatments

The bales were randomly assigned to one of four treatment groups: no plastic/no tarp (NN), no plastic/tarp (NT), plastic/no tarp (PN) and plastic/tarp (PT). All bales were assigned a random number from one to 96 and weighed. Those bales requiring plastic wrap were first wrapped two to three times with Tytan Wrap Premium Silage Film, 750 mm by 1500 m by 25.4 µm (Tytan International, LLC, Lenexa, KS). The PN and PT bales were then placed in two extra-large, industrial strength black plastic trash bags (Husky Brand, 45 gallon with 1mm thickness, Home Depot). The bales were randomly assigned to a sampling time point of zero, two, four or six months with six bales per treatment per sampling time. Bales were placed on wooden pallets in a single layer, grouped by sampling time point and covered with a tarp as needed (treatments NT and PT). Bales from each treatment and sampling time were arranged in the same configuration as well.

Bale Sampling

At sampling each whole bale was individually ground through a FitzMill (FitzMill Comminutor, Fitzpatrick, Elmhurst, IL) with a screen size of 4.76 mm. The shredded biomass was collected and mixed thoroughly for two minutes using a twin shaft paddle mixer (Hayes & Stolz, Fort Worth, TX). From the homogenized mixture, sub-samples were collected for compositional analysis (about one kg) and for further particle size reduction through a Bliss Hammer Mill (Bliss Industries, Ponca City, OK) with screen size of 0.397 mm attached to a Craftsman ShopVac (Sears Holdings Corp., Hoffman Estates, IL). Samples ground through the Bliss Mill were collected, mixed and further sub-sampled for conversion to ethanol, moisture content determination and measurement of extra-cellular enzymatic activity.

Biomass Component Determination

Biomass components were determined by the Ruminant Nutrition Lab of the Department of Animal Sciences at Kansas State University, Manhattan, KS. Sub-samples from three bales from each treatment group at each sampling time were collected after bale grinding and immediately used for biomass component analyses. Analyses included dry matter determination using AOAC Method 930.15, neutral detergent fiber (NDF) using ANKOM Method 6, acid detergent fiber (ADF) using ANKOM Method 8 and acid detergent lignin (ADL) using ANKOM Method 8 (ANKOM Technology, USA). ADF, NDF and ADL values were used to obtain cellulose, hemicellulose and lignin content of the baled biomass sub-samples. In addition, moisture content of each bale at the sampling time was determined using AACC International Method 44-19.01

Extra-Cellular Enzyme Activity Assay

Cellulase and hemicellulase activities were determined using a fluorimetric, 96-well plate assay based on the method outlined by Marx et al. (2001), listed in table 3.1 (all substrate and buffers from Sigma-Aldrich, St. Louis, MO). In brief, a one gram sample of each ground bale was mixed with 100 ml of 50 mM sodium acetate pH 7.0 and sonicated for two minutes to release any enzymes from the ground samples. From the sonicated sample, 20 µl was loaded into designated substrate and standard curve wells of the 96-well plate. Plates were loaded with four substrates linked to 4-methylumbelliferone for determination of enzymatic activity based on fluorescence. The linked substrates were loaded into the plate at final substrate concentrations of 200 and 400 µM, both in duplicate. The standard curve was constructed using 4-methylumbelliferone at a final concentration of 0, 2.5, 5.0, 10.0, 20.0, 30.0, 40.0 and 50 µM. 50 mM sodium acetate pH 7.0 was loaded into the plate first, followed by 4-MUB for the standard curves, 4-MUB linked substrates and then samples. Plates were incubated at 30°C for two hours. After incubation, 10 µl of 1 M sodium hydroxide was added to each well to increase the pH beyond 10 to maximize MUB fluorescence. Fluorescence was measured by a computerized microplate fluorimeter (Victor 3, PerkinElmer, Waltham, MA) with 355 nm excitation and 460 nm emission filters. The enzymatic activity was estimated by regression based on the standard curves.

Conversion to Ethanol

As with the biomass component determination, aliquots from three bales from each treatment and sampling time were used for conversion to ethanol. The conversion of sorghum biomass to ethanol followed methods outlined by Yoo et al. (2011) with some modifications. In brief, substrate (20% w/w) was added to 2% (w/w) sodium hydroxide solution (Sigma Aldrich, St. Louis, MO) and autoclaved (SS-325E; Tomy Tech, USA, Inc., Fremont, CA) at a temperature of 121°C for 30 minutes. The remaining solids were washed with deionized water to neutralize, followed by drying in an air oven at 60°C for 24 h. Enzymatic hydrolysis of the dried pretreated substrate was carried out using the proprietary blend of cellulase enzymes, Cellic® CTec2, provided by Novozymes (Novozymes A/S, Denmark). Enzymatic hydrolysis was carried out with 3.0% enzyme per total solids (g enzyme/g biomass) in 0.1 M citric acid buffer pH 5.0 at 50°C for 48 h and shaking at 120 RPM. The hydrolyzed substrate was inoculated at 2% (v/v) with actively growing *Saccharomyces cerevisiae* Ferm Pro™ (Danville, KY) and incubated at 30°C for 24 hours with shaking at 100 RPM. In addition, 3% (v/v) of a 10% yeast extract solution (BD Biosciences, San Jose, CA) was added at inoculation of the substrate. Actively growing *S. cerevisiae* was obtained through propagation of -80°C glycerol stock cultures in yeast mold broth (BD Biosciences, San Jose, CA) for 24 h at 30°C. 1 mL samples were taken after enzymatic hydrolysis and fermentation for quantification of glucose and ethanol using a binary HPLC system (Shimadzu Corporation, Japan) as described by Oberoi et al. (2011) and a Phenomenex Rezex RPM monosaccharide column (300 x 7.8 mm; Phenomenex, CA). In brief, degassed deionized water was used as a mobile phase at a flow rate of 0.6 mL min^{-1}. The column oven and refractive index detector (RID-10A) were maintained at 80°C and 65°C, respectively. Samples were centrifuged and filtered through Phenomenex 0.45 micron RC membranes prior to injection. Peaks were detected by the refractive index detector and quantified on the basis of area and retention time of the standards (glucoseand ethanol).

Data Analysis

The data were analyzed using PROC GLIMMIX in SAS 9.2 Software (Cary, North Carolina, USA). The treatment and time main effects and treatment*time interactions were tested, followed by post-hoc analysis of treatment. The effects of storage time were evaluated against treatment NN at time 0 as the control. The treatment*time interactions were further

decayed with pairwise comparisons of slice effects. Differences were considered significant at alpha = 0.05. Statistical differences between treatments at a sampling point are indicated by different letters, while * indicates a statistical difference from the control (treatment NN at time 0).

Results and discussion

This study provides detailed information on biomass components, extra-cellular enzyme activity and ethanol yields in dry, baled biomass stored under the four storage conditions over a 6 month period. All dry weights are based on the post-storage moisture content of each bale. At sampling time 0, the average moisture content was determined to be 10-12%, which was also assumed to be the initial moisture content of all the bales. The bales in the uncovered treatment were visibly deteriorated (dark colored, moldy), whereas those from the covered treatments were similar to their initial condition after baling, with little to no noticeable deterioration, even after 6 month storage.

Biomass Component Determination

Sorghum biomass dry matter, cellulose and hemicellulose contents were generally similar in treatment NT, PN and PT, but differed from treatment NN after storage (Figure 3.1), while lignin content was similar between all treatments and storage times, including NN. We focus on the differences between the three covered treatments (NT, PN and PT) and the uncovered, control treatment (NN) at each sampling time. As expected, biomass components did not differ among the treatments at time zero. Dry matter did not change after 6 months of storage in the covered treatments, shown in figure 3.1a. In contrast, dry matter content declined in treatment NN, which also differed from treatments NT, PN and PT after 2, 4 and 6 month storage. Our dry matter results concur with Khanchi et al. (2009), who baled forage sorghum into large square or round bales and stored them either without cover or covered with a tarp. Khanchi et al. (2009) reported that the dry matter losses in uncovered bales stored outdoors ranged from 5.73 to 6.04% after 6 months of storage, whereas dry matter losses in the tarp-covered bales ranged from 5.73 to 6.34%. Shah et al. (2011) reported dry matter losses of 11% in corn stover bales covered with a tarp and 17% dry matter losses in bales wrapped in plastic after 9-month storage. Dry matter losses were more dramatic in this study: uncovered small square bales lost on average 33.2% of dry matter during the 6-month storage. These differences among the studies

are likely attributable to bale geometry, bale size, bale density, biomass feedstock, and sampling methods.

Cellulose contents in the covered treatments NT, PN and PT remained stable compared to the uncovered treatment NN after 2, 4 and 6 month storage (figure 3.1b). In contrast to dry matter, cellulose contents in all treatments after 4-month storage were lower than in the control, with treatment NN having the lowest content of 22.4%. This continued decline in cellulose content for treatments NT, PN and PT was not seen after 6-months storage. While cellulose content in treatments NT and PT did not differ from the control after 6 month storage, it was lower in treatment PN than in the control (figure 3.1b). The reduced cellulose content in treatments NT, PN and PT after 4 months storage could be attributed to environmental conditions, as sampling occurred in mid-February when temperatures and precipitation were at a low (figure 3.2). Additional environmental factors that could have led to biomass degradation are temperature cycling and extremes, moisture (precipitation) and solar degradation. However, at the 6 months storage time, reduced cellulose content was not seen in treatments NT, PN and PT compared to time 0. This could indicate that reduced temperatures and precipitation (seen after 4 months of storage) could greatly influence cellulose degradation, even when the biomass is covered.

The change in hemicellulose content was found to follow the same pattern as dry matter. The uncovered treatment NN was found to have reduced hemicellulose content after 2, 4 and 6 month storage, while the content in the covered treatments (NT, PN and PT) remained stable (figure 3.1c.). Lignin contents did not differ among the treatments at sampling point (figure 3.1d). However, treatment NN differed from control after 2, 4 and 6 month storage. As seen with cellulose and hemicellulose contents, a decrease in lignin content was observed after 4 months storage for all treatments, which could be due to environmental factors at sampling as previously described. Our observations are consistent with those of Shah et al. (2011), who also reported stable cellulose, hemicellulose and lignin contents after 9 months storage of large square corn stover bales covered with a tarp or wrapped in plastic.

Extra-cellular Enzymatic Activity (EEA) Assay

The MUB-linked substrates allowed for the assessment of cellulose and hemicellulose degrading extra-cellular enzyme activities in plant biomass used for ethanol production and

stored for 0, 2, 4 and 6 months under four conditions. Such fluorometrically linked substrates have been primarily used for estimation of extra-cellular enzyme activities in a wide variety of soil and litter samples (DeForest, 2009; Saiya-Cork et al., 2002; Selmants et al., 2005), but not specifically for plant biomass samples. Enzymatic activities reported here are similar or slightly higher than those in soil and litter samples with comparable moisture contents (DeForest, 2009; Selmants et al., 2005), as might be expected based on the abundance of (hemi-)cellulose in our biomass. Cellulose (4-MUB- β-D-cellobioside and 4-MUB- β-D-glucoside) and hemicellulose (4-MUB- β-D-glucuronide hydrate and 4-MUB-β-D-xyloside) degrading enzyme activities are shown in figure 3.3. As with the biomass components, treatment NN at time 0 was considered as a control and none of the EEAs differed among the treatments at time 0. Cellulose degrading activities in the covered treatments NT, PN and PT did not differ across storage times and did not differ from the control after any length of storage (Figures 3.3a and 3.3b). In contrast, these activities were higher in the uncovered treatment than in the control after 2, 4 and 6 month storage. Although the two substrates used to estimate the activities of cellulose-degrading enzymes were largely consistent, cellulase activity as measured by 4-MUB-β-D-glucoside was higher in treatment PT than in the control and did not differ from treatment NN after 6-month storage. Hemicellulase activity as measured by 4-MUB- β-D-glucuronide hydrate (figure 3.3c) did not differ between treatments or when compared to the control after any length of storage, except for treatment PT where estimated activity was higher than in the control after 6-month storage. Hemicellulase activity as measured by 4-MUB-β-D-xyloside did not differ between treatments at time zero as expected. The estimated activity in treatment NN was greater than in the control after 2, 4 and 6 months of storage. Furthermore, this activity was higher in treatment NN than in treatments NT, PN and PT after 2, 4 and 6 month storage. Treatments NT, PN and PT did not differ at 2 and 4 month storage, but differed between treatments after 6 months. Overall, the cellulose and hemicellulose degrading extra-cellular enzyme activities were consistently highest in the uncovered bales. This corroborates the reduced cellulose and hemicellulose contents in these bales. In contrast, the enzyme activities remained low in the covered bales and the cellulose and hemicellulose contents were stable after 6 months of storage.

Conversion to Ethanol

Prior to enzymatic hydrolysis, we optimized sample pretreatment using a dilute acid (2% sulfuric acid) or dilute alkali solutions and autoclaving (results not shown) as described in Brijwani et al. (2010) and Oberoi et al. (2011). The dilute alkali pretreatment (protocol described above) resulted in the highest glucose yield from the biomass and was therefore chosen for further evaluation for the production of ethanol. We also evaluated optimal parameters for enzymatic hydrolysis and fermentation to ethanol by varying enzyme dosage levels and incubation times (results not shown). Figures 3.4a and 3.4b display glucose and ethanol yield per gram of biomass on dry weight basis (prior to pretreatment) displayed. As with the other parameters (biomass components and extra-cellular enzymatic activity) we determined, the glucose and ethanol yields did not differ among the treatments at time 0. We observed no differences in glucose yields between the treatments or between treatments and the control after up to six months of storage, except for an increase in the uncovered treatment NN after 2-month storage compared to the control. Our glucose yields were higher than those reported by Brijwani et al. (2010) and Yoo et al. (2011) for enzymatic hydrolysis of soybean hulls and soybean hulls supplemented with wheat bran, respectively. These differences can be attributed to differences in feedstock component quantities and in the chosen enzyme hydrolysis system. Similarly to glucose yields, ethanol yields did not differ between the treatments at time zero. None of the treatments differed from each other or from the control after 2 or 4 months storage in regards to ethanol yields. However, after 6 months, ethanol yields from biomass in treatment NN were lower than those from treatments NT, PN and PT and from the control. Ethanol yields in treatments NT, PN and PT did not differ from each other or from the control after six months storage. Ethanol yields after 0, 2 and 4 months of storage are similar to those reported by Balat (2011) for conversion of cornstalk to ethanol using dilute alkali pretreatment and enzymatic hydrolysis. During 6-month storage, ethanol yield from covered bales remained stable. In contrast, uncovered bales had a significant decrease in ethanol yields after 6 months of storage. Although the uncovered biomass contained less cellulose, we did not observe congruent decreases in the glucose yields. This may indicate that biomass storage either uncovered or covered does not compromise the ability of the commercial enzymes to hydrolyze cellulose to glucose during enzymatic hydrolysis. Furthermore, the reduced ethanol yields from biomass stored uncovered for 6 months, observed in this study, could indicate the introduction of

compounds inhibitory to ethanol fermentation by *S. cerevisiae*. In addition, it is important to note the large loss of dry matter from the biomass stored uncovered - these losses will ultimately reduce the ethanol volume produced from biomass stored uncovered.

Substrate losses, including dry matter, cellulose and hemicellulose, were consistently greatest in uncovered stored plant biomass left susceptible to environmental elements. The results of this study strongly indicate plant biomass substrate used for the production of lignocellulosic ethanol can be preserved during storage if the biomass is covered.

Conclusions

From biomass harvest to conversion to ethanol, the storage method (covered or uncovered) for preserving dry plant biomass used for lignocellulosic ethanol production was found to play a large role in substrate quality and subsequent ethanol yields. This comprehensive study found congruency in three datasets, specifically reduced cellulose content, increased extra-cellular enzymatic activity associated with cellulose degradation and ultimately reduced conversion to ethanol yields in biomass left uncovered over six-month storage. Congruency was also seen in covered biomass bales, with stability in cellulose content, extra-cellular enzymatic activity and ethanol yields.

Acknowledgements

The authors wish to thank the Center for Sustainable Energy, and Department of Grain Science and Industry and Kansas State University for funding this project. Authors greatly acknowledge Dr. Scott Staggenborg, Department of Agronomy and the staff of the Agronomy North Farm, Kansas State University for expertise related to sorghum production and harvest, along with Dr. Leigh Murray from the Department of Statistics, Kansas State University for assistance with experimental design and data analyses. This material is based upon work supported by National Science Foundation Grant #0903701: "Integrating the Socioeconomic, Technical, and Agricultural Aspects of Renewable and Sustainable Biorefining Program", awarded to Kansas State University. This article is Contribution No. 13-048-J from the Kansas Agricultural Experiment Station, Manhattan, KS 66506, USA.

References

Balat, M., 2011. Production of bioethanol from lignocellulosic materials via the biochemical pathway: a review. Energ. Convers. Manage. 52, 858-875.

Brijwani, K., Oberoi, H.S., Vadlani, P.V., 2010. Production of cellulolytic enzyme system in mixed-culture solid-state fermentation of soybean hulls supplemented with wheat bran. Process Biochem. 45, 120-128.

Darrah, P.R., Harris, P.J., 1986. A fluorimetric method for measuring the activity of soil enzymes. Plant Soil. 92, 81-88.

DeForest, J.L., 2009. The influence of time, storage temperature, and substrate age on potential soil enzyme activity in acidic forest soils using MUB-linked substrates and L-DOPA. Soil Biol. Biochem. 41, 1180-1186.

Emery, I.R., Mosier, N.S., 2012. The impact of dry matter loss during herbaceous biomass storage on net greenhouse gas emissions from biofuels production. Biomass Bioenerg. 39, 237-246.

Freeman, C., Liska, G., Ostle, N.J., Jones, S.E., Lock, M.A., 1995. The use of fluorogenic substrates for measuring enzyme activity in peatlands. Plant Soil. 175, 147-152.

Gao, J., Qian, L., Thelen, K. D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103 (3), 844-850.

Gibbons W.R., Hughes, S.R., 2009. Integrated biorefineries with engineered microbes and high-value co-products for profitable biofuels production. In Vitro Cell. Dev.-Pl. 45, 218-228.

Hess J.R., Wright, C.T., Kenney, K. L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel. Bioprod. Bior. 1, 181-190.

Hoppe, H-G., 1983. Significance of exoenzymatic activities in the ecology of brackish water: measurement by means of methylumbelliferyl-substrates. Mar. Ecol.-Prog. Ser., 11, 299-308.

Khanchi, A., Jones, C.L., Sharma, B., 2009. Characteristics and compositional variation in round and square sorghum bales under different storage conditions. ASABE Meeting Presentation. Paper Number:096672.

Marx, M.C., Wood, M., Jarvis, S.C., 2001. A microplate fluorimetric assay for the study of enzyme diversity in soils. Soil Biol. Biochem. 33, 1633-1640.

Oberoi, H.S., Vadlani, P.V., Nanjudaswamy, A., Bansal, S., Singh, S., Kaur, S., Babbar, N., 2011. Enhanced ethanol production from Kinnow mandarin (*Citrus reticulate*) waste *via* a statistically optimized simultaneous saccharification and fermentation process. Bioresour. Technol. 102, 1593-1601.

Saiya-Cork, K.R., Sinsabaugh, R.L., Zak, D.R., 2002. The effects of long term nitrogen deposition on extracellular enzyme activity in an *Acer saccharum* forest soil. Soil Biol. Biochem. 34, 1309-1315.

Selmants, P.C., Hart, S.C., Boyle, S.I, Stark, J.M., 2005. Red alder (*Alnus rubra*) alters community-level soil microbial function in conifer forests of the Pacific Northwest, USA. Soil Biol. Biochem. 37, 1860-1868.

Shah, A., Darr, M.J., Webster, K., Hoffman, C., 2011. Outdoor storage characteristics of single-pass large square corn stover bales in Iowa. Energies 4, 1687-1695.

Somville, M., 1984. Measurement and study of substrate specificity of exoglucosidase activity in eutropic water. Appl. Environ. Microb., 48, 1181-1185.

Yoo, J., Alavi, S., Vadlani, P., Amanor-Boadu, V., 2011. Thermo-mechanical extrusion pretreatment for conversion of soybean hulls to fermentable sugars. Bioresour. Technol. 102, 7583-7590.

Zheng, Y., Yu, C., Cheng, Y., Lee, C., Simmons, C.W., Dooley, T.M., Zhang, R., Jenkins, B.M., VanderGheynst, J.S., 2012. Integrating sugar beet pulp storage, hydrolysis and fermentation for fuel ethanol production. Appl Energ. 93, 168-175.

Tables and Figures

Table 3.1 Extra-cellular enzymes assayed in biomass samples collected from each treatment at each sampling point, their commission number (EC) and corresponding substrate. (4-MUB=4-methylumbelliferone)

Enzyme	EC	Substrate
Cellulose Degrading		
Cellobiohydrolase	3.2.1.91	4-MUB-β-D-cellobioside
β-1,4-Glucosidase	3.2.1.21	4-MUB-β-D-glucoside
Hemicellulose Degrading		
β-glucuronidase	3.2.1.31	4-MUB-β-D-glucuronide hydrate
β-1,4-Xylosidase	3.2.1.37	4-MUB-β-D-xyloside

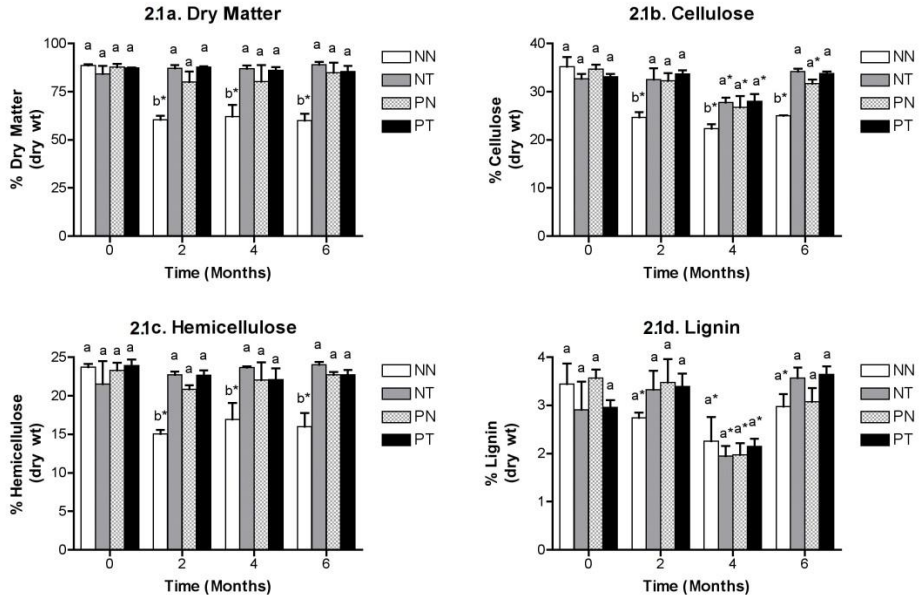


Figure 3.1 Changes in biomass components, including dry matter, cellulose, hemicellulose and lignin as percentage on dry basis. Different letters indicate significant differences within storage period, while * indicates significant difference from control (treatment NN at time 0).

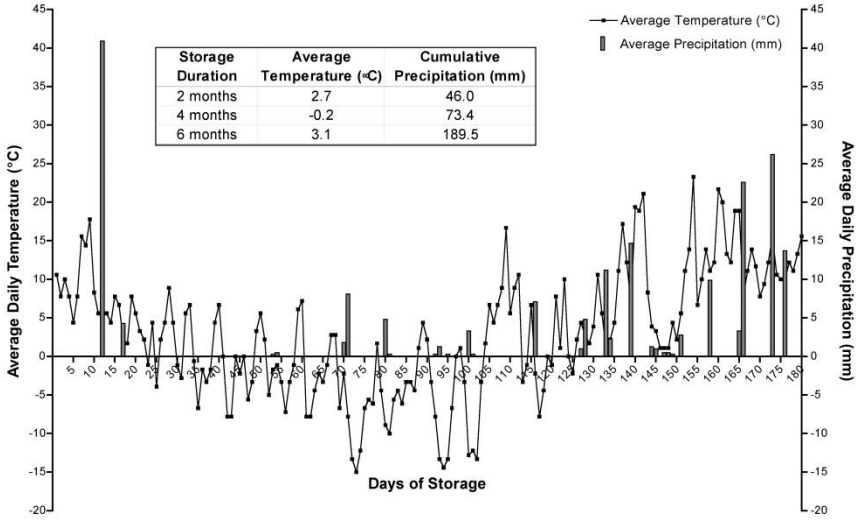


Figure 3.2 Daily temperature and precipitation averages during plant biomass stored under 4 different conditions (NN, NT, PN and PT) and 4 storage durations (0, 2, 4 and 6 months) along with average temperatures and cumulative precipitation for each storage time. Two-months sampling occurred after 60 days of storage on December 13th, 4 months sampling after 120 days of storage on February 21st and 6 months sampling after 180 days on April 18th.

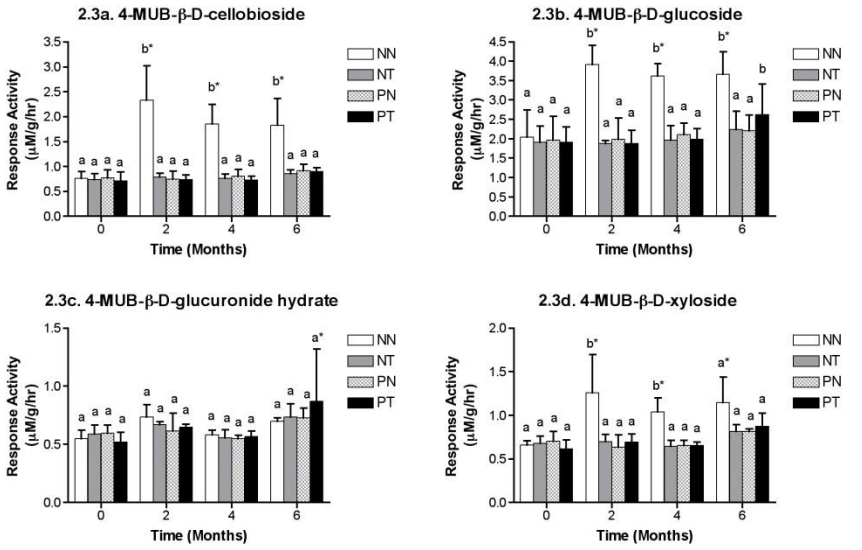


Figure 3.3 Cellulose degrading enzyme activity as measured by 4-MUB-β-D-cellobioside (3.3a) and 4-MUB-β-D-glucoside (3.3b) and hemicellulose degrading enzyme activity as measured by 4-MUB-β-D-glucuronide hydrate (3.3c) and 4-MUB-β-D-xyloside (3.3d) for all treatments over six months of storage. Activities given in µM of activity per g of biomass

per hour on dry weight basis. Different letters indicate significant differences within a sampling time, while * indicates significant difference from the control (treatment NN at time 9).

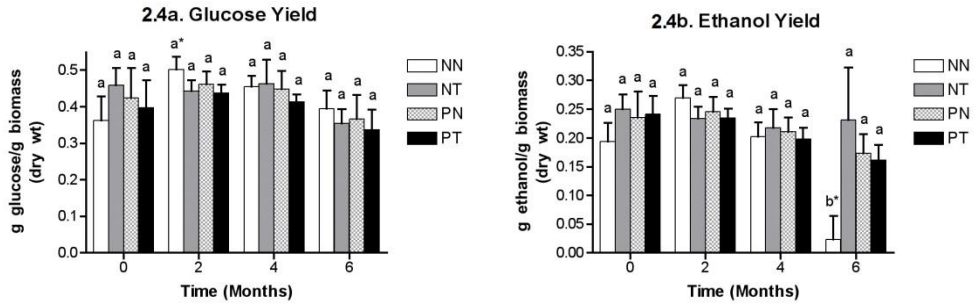


Figure 3.4 Glucose (3.4a) and ethanol (3.4b) yields across treatments over time. Values given on dry weight basis. Different letters indicate significant differences within sampling time, while *indicates significant difference from the control (treatment NN at time 0).

Chapter 4 - Fungal community dynamics in stored biomass based on Large SubUnit ribosomal DNA amplicon sequencing[2]

Abstract

The American biofuels industry has been thrust into high gear due to the passing of the Energy Independence and Security Act (EISA) of 2007 and the subsequent roll out of the Renewable Fuel Standards (RFS) program, outlining production mandates of transportation fuels, specifically ethanol, from renewable resources. The RFS mandates included increasing production from non-grain feedstocks, using second generation feedstocks from lignocellulosic materials, including agricultural (corn stover or wheat straw) and forestry residues, perennial grasses (*Miscanthus* or switchgrass) and short rotation forestry crops (poplar, willow or eucalyptus). Current production, harvest, transport and storage infrastructure and practices are designed for handling grain crops with uniform size and bulk density; however, lignocellulosic materials vary greatly in size and bulk density. For the economic production of ethanol from lignocellulosic materials a biorefinery will need a continuous supply of biomass throughout the year. Unfortunately, many lignocellulosic biomass feedstocks may only be harvested annually or biannually, requiring long term storage. Lignocellulosic ethanol is produced from the breakdown of the cellulose to simple sugar in the biomass, thus maintaining the cellulose quality is essential for profitability. Many common phytopathogenic or saprobic fungi can break down cellulose and monitoring the fungal communities and their dynamics in stored biomass is thus essential. Here, we evaluated the fungal community composition and dynamics by direct 454-pyrosequencing in *Sorghum* biomass during six-month storage either wrapped in plastic (PN), covered with a tarp (NT), wrapped in plastic and covered with a tarp (PT) or stored with no cover (NN). Fungal communities increased in richness (S_{obs}) and declined in evenness during storage in uncovered biomass, while covered treatments remained stable. In contrast to richness, diversity estimators (Simpson's diversity index) did not change during storage in any of the four treatments. Non-metric multidimensional scaling (NMDS) and Analysis of Molecular Variance (AMOVA) indicated that while the fungal communities were indistinguishable among the covered treatments, fungal communities in the uncovered treatments were distinct from the covered

[2] To be submitted to Applied and Environmental Microbiology

treatments. In total, the fungal communities were comprised of a total of 210 Operational Taxonomic Units (OTUs), strongly dominated by the Phylum Ascomycota (162 OTUs) and a few (48 OTUs) assigned to Phylum Basidiomycota. We conclude that covering the stored biomass likely prevents the introduction of allochthonous propagules, thus better preserving the biomass for downstream applications.

Introduction

In 2007, the United States government signed into action the Energy Independence and Security Act (EISA), which expanded the Renewable Fuel Standards (RFS) program and outlined government mandated goals for the utilization and production of renewable fuels in the United States by 2022. In 2010 the revised RFS2 mandates outlined that 136 billion liters of renewable fuels be used in the US and 60.5 liters of that be produced using the lignocellulosic platform or second generation technologies, along with reduction of greenhouse gas (GHG) emissions (Gao *et al.*, 2011). Currently, ethanol in the US is primarily produced using first generation technologies that utilize grain, primarily corn, but also more limited quantities of sorghum and wheat. Second generation biofuels are those produced via biochemical or thermochemical pathways from lignocellulosic materials such as a feedstock, including agricultural and forestry residues, solid waste, perennial woody and herbaceous energy crops (Gibbons and Hughes, 2009). The biochemical pathway requires the enzymatic degradation of cellulose into glucose for fermentation to ethanol; while the thermochemical pathway consists of burning the feedstocks in the absence or presence of oxygen for the production of syngas to be further synthesized and mixed with gasoline. Ethanol is being produced at pilot-scale quantities through the biochemical pathway using lignocellulosic feedstocks; however, multiple commercial-scale plants are projected to come online by the end of 2013.

Feedstock costs, availability, storage and transportation expenses control the economic competitiveness of lignocellulosic ethanol production (Hess *et al.*, 2007). Hess *et al*. (2007) estimate feedstock costs to be 35-50% of total ethanol production costs, with the biomass transportation logistics comprising of 50-75% of the production costs. The proportion of the transportation costs depends on biomass species, yield, location, climate and local economy as well as harvest, collection, preprocessing, transportation and material handling technologies. Reduced biomass producer and biorefinery operator profit margins are a concern when logistical costs exceed 25% of the total biomass costs (Hess et al., 2007). The biomass supply is further complicated by production: feedstock may be harvested only annually or bi-annually, necessitating biomass storage for a consistent, year-round supply operation of the biorefineries. Duration of storage depends on the feedstock availability and can range from six to twelve months. Identifying the most cost effective means to minimize the dry matter loss is thus essential. Reducing dry matter losses during storage could play a key role in cutting the incurring

biomass costs for producers and biorefinery operators. Selection of the optimal storage method plays an important role: dry storage methods reduce dry matter losses compared to wet storage methods (Emery and Mosier, 2012). Thus far, few studies have fully detailed the resulting ethanol yields from biomass stored under dry and wet storage conditions over an extended period of time, i.e., storage after harvest to the next harvest.

Currently, to the best of our knowledge, this is the first report to characterize fungal communities present in lignocellulosic biomass stored over an extended period of time and destined for biofuel production. Furthermore, the impacts of the feedstock-residing fungal communities on biomass quality and ethanol yields have not been evaluated. In the studies reported here, we analyzed the fungal community composition and its dynamics using direct 454-pyrosequencing of the large subunit (LSU) of the ribosomal RNA (rRNA) gene. Our overall goals were to evaluate the effects that different storage conditions and the storage duration have on fungal communities: their richness and diversity as well as composition and dynamics during a six-month storage period. To do this we assigned baled Sorghum biomass to four storage treatments and followed the fungal community dynamics up to six months at two-month intervals. We hypothesized that 1) the richness and diversity of the uncovered treatment (NN) would increase during the incubation period; 2) richness and diversity in the covered treatments (NT, PN and PT) would remain unchanged during the storage period as fungal community would be dominated by a few species due to decreased fluctuations in moisture because of the coverage material; 3) the dominant taxa in the uncovered and covered treatments would differ; and, 4) the fungal community constituents would change over time differently across the different storage conditions. Our results suggest that covering of the stored biomass is adequate to stagnate and slow down the fungal community dynamics implying that simple and cost-effective measures to protect the stored biomass from fluctuations in the environmental conditions may be adequate to protect the biomass integrity for biofuel processing.

Materials and Methods

Biomass Storage Experimental Design

Photo-period sensitive sorghum cultivar PS1990 from Sorghum Partners (Chromatin, New Deal, TX) was grown on the North Agronomy Farm, Department of Agronomy, Kansas State University, Manhattan, KS. The sorghum was planted June 3, 2010 and fertilized 15 to 20

days after planting at a rate of about 112 kg per ha using urea (N-P-K; 46-0-0) and herbicide-treated with Bicep II Magnum (Syngenta, Basel, Switzerland) at 5.6 l per ha. Sorghum was cut and windrowed October 7, 2010 and allowed to dry on site. On October 15, 2010, sorghum was baled into small bales (average size 0.36 by 0.46 by 0.91 m) using a Massey Ferguson (AGCO, Duluth, GA) square baler. A total of 96 bales were weighed and randomly assigned to one of four treatments (no plastic/no tarp [NN], no plastic/tarp [NT], plastic/no tarp [PN] and plastic/tarp [PT]) and to one of four sampling times (0-, 2-, 4- or 6- months after baling) for a total of six replicates per treatment and sampling time. The PN and PT bales were wrapped two to three times with Tytan Wrap Premium Silage Film (750 mm by 1500 m by 25.4 µm; Tytan International, LLC, Lenexa, KS) and placed in two extra-large, industrial strength black plastic trash bags (Husky Brand, 45 gallon with 1mm thickness, (Husky Plastics, Thornhill, Ontario, Canada). The bales assigned to 2-, 4-, and 6-month incubation were grouped by sampling time point and stored in a single layer on wooden pallets at the North Agronomy Farm (Manhattan, KS). Individual tarps (silver heavy duty from Erickson Manufacturing LTD., Marine City, MI) were placed over treatments NT and PT for each sampling time point with the tarp covering the bale tops and sides completely.

DNA Extraction and Amplicon Production

Each bale was ground using a FitzMill (FitzMill Comminutor, Fitzpatrick, Elmhurst, IL) with a screen size of 4.76 mm. The shredded biomass was collected and mixed thoroughly for two minutes using a twin shaft paddle mixer (Hayes & Stolz, Fort Worth, TX). From the homogenized mixture, a 1 kg sub-sample was collected for further particle size reduction through a Bliss Hammer Mill (Bliss Industries, Ponca City, OK) with screen size of 0.397 mm attached to a Craftsman ShopVac (Sears Holdings Corp., Hoffman Estates, IL). Samples ground through the Bliss Mill were collected, mixed and further sub-sampled (about 5 g) for DNA extraction and stored at -80°C.

From the previously prepared sub-samples stored at -80°C, a 1 gram aliquot was ground in liquid N_2 with a mortar and pestle until the N_2 evaporated. The grinding was repeated for a total of three times. The ground sample was transferred into a 50 mL DNA extraction tube (Qiagen DNeasy Plant Maxi Kit, Qiagen Inc, Valencia, CA, USA) and DNA was extracted according to the manufacturer's protocol, with an additional 5-min centrifugation added to

minimize the ethanol carry-over after the addition of Buffer AW. The final elution was carried out in two steps, as recommended, with a 750 µl volume at each step. Extracted DNA was visualized on a 1% TBE agarose gel and quantified using a ND1000 spectrometer (NanoDrop Technologies, Wilmington, DE). Template DNA for each sample was aliquoted into 96-well plates at a working concentration of 10 ng μL^{-1} and stored at -20°C. Remaining DNA was archived at -80°C.

Fungal amplicons for direct 454-pyrosequencing of the large subunit (LSU) of the ribosomal RNA (rRNA) gene were generated using a two-step PCR protocol to reduce primer biases (Berry et al. 2011). In brief, triplicate primary and secondary PCR reactions were carried out in 50 µL reactions each containing 100 ng of template, 200 µM of each forward and reverse primer, 200 µM of each dioxynucleotide, 2.5 mM $MgCl_2$, 10 µL 5x Green GoTaq Flexi Buffer (Promega, Madison, WI), 9.6 µL DNase/RNase free water and 2 U of GoTaq Hot Start Polymerase (Promega, Madison, WI). Primary PCR reactions contained the forward primer LROR (5'-CCGCTGAACTTAAGCATATCAATA-3') (Amend *et al*., 2010) and the reverse primer LR3 (5'-CCGTGTTTCAAGACGGG-3') (Vilgalys and Hester, 1990; see Appendix B, Table B.1). Cycle parameters for the primary PCR consisted of a 94°C initial denaturing for 4 min, followed by 20 cycles of denaturing at 94°C for 1 min, annealing at 53°C for 45 sec and extension at 72° for 2 min and a final extension of 72°C for 8 min in a MasterCycler (Eppendorf, Hamburg, Germany). All PCR products were visualized on a 1.5% TBE agarose gel to ensure the presence of PCR products of the appropriate fragment size. Secondary PCR reactions added sample-specific DNA barcode tags and the 454-linkage adaptors, the modified LR0R-A primer included the 454-adaptor (A-primer) and 10 bp barcode tags unique to each sample with the modified LR3-B primer containing the 454-adaptor (B-primer) (Margulies et al., 2005; see Appendix B, Table B.1). To remove excess PCR primers, the primary PCR products were purified using Diffinity RapidTip (Diffinity Genomics, Inc, West Henrietta, NY) according to the manufacturer's specifications and a 10 µL aliquot used for secondary PCR reactions. The secondary PCR reaction parameters were identical to the primary PCR but included only 5 cycles. Positive controls of *Saccharomyces cerevisiae* genomic DNA and negative controls without template DNA were included; negative controls remained free of PCR products throughout.

The secondary PCR products were purified and normalized using SequalPrep Normalization 96-well Plate (Life Technologies, Grand Island, NY) according to the manufacturer's specifications. The purified products were pooled equimolarly and purified again using Agencourt AMPure XP (Beckman Coulter Inc, Brea, CA) as recommended, with the modification of 1:1 PCR-product to bead loading ratio and the purified amplicon library was 454-pyrosequenced (GS FLX Titanium, Roche Applied Science, Indianapolis, IN) at the Integrated Genomics Facility at Kansas State University (Manhattan, KS).

Data Analysis

The sequence data were analyzed using the mothur software package (v. 1.27.0; Schloss *et al*., 2009) following a modified standard operating procedure outlined by Schloss *et al*. (2011; www.mothur.org/wiki/Schloss_SOP, accessed October 15, 2012). We completed two runs of sequencing and merged resultant files after denoising (Schloss et al., 2011; Quince et al., 2009; Quince et al., 2011) and prior to further analyses. Unique representative sequences of the trimmed data were pre-clustered (pre.cluster) based on 2 bp differences and aligned against a modified LSU training dataset by James et al. (2006), which included the removal of all non-fungal sequences and re-alignment using the MAFFT algorithm. Taxon information obtained from the Ribosomal Database Project (RDP) naïve Bayesian classifier (Wang *et al*., 2007). A total of 4,442 chimeric sequences were removed using the UChime algorithm (Edgar et al., 2011) as implemented through the mothur program. Sequences were further clustered into Operational Taxonomic Units (OTUs) at 97% similarity using the average-neighbor algorithm. The OTUs were assigned to taxa using the naïve Bayesian classifier (Wang *et al*., 2007) against the Ribosomal Database Project's training dataset at a 60% bootstrap support cutoff (Liu *et al*., 2012).

Global singletons were removed from the data set prior to further analysis (Tedersoo *et al.,* 2010). We subsampled 652 sequences from each experimental unit to eliminate bias due to unequal sampling effort (Gihring *et al*., 2011). Sub-sampling resulted in the elimination of bales 64 (NN), 73 (NT), 70 (PN), 82 (PN) and 85 (PT) of 6-month incubation.

Community coverage, richness, diversity and evenness were determined based on collected OTU information. Adequacy of sampling was determined based on Good's coverage (formula: $C = 1 - \frac{n_1}{N}$; where n_1 = the number of OTUs that have been sampled once and N = the

total number of individuals in the sample) and construction of rarefaction curves after each incubation period (Rarefaction formula = $S_n = S_t - \frac{\sum_{i=1}^{S_t} \frac{N-N_i}{n}}{\frac{N}{n}}$; where S_n = average number of OTUs observed after drawing n individuals and S_t = total number of OTUs in sample of N total individuals) . Richness was estimated by looking at the observed number of species (S_{obs}) and estimated through Chao1 ($S_{chao1} = S_{obs} + \frac{n_1(n_1-1)}{2(n_2+2)}$; where S_{chao1} = the estimated richness, S_{obs} = the observed number of species, n_1 = the number of OTUs with only one sequence (i.e. "singletons"), n_2=the number of OTUs with only two sequences (i.e., "doubletons") (Magauran, 1988). The Complement of Simpson's Index of Diversity (1-D; where $D = \frac{\sum n(n-1)}{N(N-1)}$; where n = the total number of organisms of a particular species, N = the total number of organisms of all species) was used to determine diversity, with a higher diversity indicated by a higher value. Evenness (E_D) ($E_D = \frac{\frac{1}{D}}{S}$; where S is the OTU richness in each sample and D is Simpson's Diversity Index shown above) was also determined based on Simpson's Diversity Index. Beta diversity was visualized using non-metric multidimensional scaling (NMDS) based on Yue and Clayton's dissimilarity matrix (θ_{YC}) created from the subsampled sequences. The θ_{YC} distances were used to determine the optimal number of dimensions to represent the data based on treatment and time. Analyses revealed a 2-dimensional output with an overall stress of 0.136 and an R^2 value of 96.1%. Community differences were tested using Analysis of MOlecular VAriance (AMOVA, PERMANOVA; Anderson 2001) and homogeneity of the populations present in each treatment and incubation time was also evaluated using HOmogeneity of MOlecular VAriance (HOMOVA) in mothur. Evaluation of associations of OTUs to treatments was done by correlation analyses (Spearman's Rank Correlation) in mothur as well. Those strongly and significantly correlated with axes separating the treatments were considered to be associated with those treatment conditions. OTUs with strong and significant negative correlations with axis 1 and axis 2 scores are correlated with the uncovered treatment, whereas those positively correlated with the two axes are correlated with the covered treatments.

Statistical analysis included treatment, incubation time or interactive differences evaluation using two-way ANOVA for Good's coverage, S_{obs} and Simpson's Diversity estimates. Further evaluation of temporal changes in community diversity metrics over the 6-month incubation period was visualized using linear regression. Both ANOVA and linear

regression analyses were performed using JMP 7.0.2 (SAS Institute Inc., Cary, NC). Of the 100 most abundant OTUs, the individual response of each OTU was tested for treatment, time and treatment by time interactions using ANOVA with a false discovery rate (FDR) of 0.05.

Results

We obtained a total of 547,356 LSU reads with an average length of 421 bp. After denoising, quality control, and removal of putative chimeras, we retained a total of 124,481 sequences. These reads were clustered into 567 OTUs at 97% similarity. Of the 567 OTUs, 357 were singletons, which were removed prior to further analysis, leaving a total of 210 OTUs. Sequences (.sff) are archived in Short Read Archive at NCBI (SRA XXXXXX-still need to archive). The fungal communities in our stored biomass were strongly dominated by the Phylum Ascomycota (162 OTUs; 77.2% of the sequences), followed by the Phylum Basidiomycota (48 OTUs; 23.8%). The 20 most abundant OTUs (based on sequence counts) comprised 97.7% (121,580 of the 124,481) of the total sequences represented. The OTU, number of sequences and taxonomic information of the top 20 OTUs are shown in Table 4.1.

The sequences represented a total of 210 OTUs, with the most abundant sequence assigned to genera *Cladosporium* (OTU 554 with 59,382 reads). An additional 42 OTUs were assigned to *Cladosporium* representing a total of 61,914 reads in our dataset and variety of Cladosporia species present. The second most abundant sequence represented by *Alternaria* (OTU 542 with 36,447 reads) was found to highly correlate to the uncovered treatment (NN). A total of 9 OTUs were found to be *Alternaria* and represented 36,517 sequence reads. In addition, *Cryptococcus*, *Eurotium* and *Gibberella* rounded out the top five most abundant sequence reads with 5,142, 3,369 and 2,729, respectively.

Good's coverage (98.6% ± 0.004) indicated that our study captured the resident diversity well in our experiment. The coverage estimators neither differed between the treatments ($F_{3,87}$ = 0.4109; P = 0.7456) nor across the incubation length ($F_{3,87}$ = 1.9907; P = 0.1620) but had a treatment by time interaction ($F_{3,87}$ = 3.3673; P = 0.0224) attributable to a decrease in coverage in treatment NN over time (Table 4.2). Although Good's coverage estimator indicated adequate sampling of the system, rarefaction curves (Figure 4.1) suggested that sampling of greater number of sequences (> 652 sequences) would have yielded additional OTU richness, which is

typical of studies using environmental DNA. Rarefaction also indicates further diversity and richness would have been captured if the sequencing effort would have been greater.

Fungal OTU richness was found to be low (24.6 ± 3.717 per sample); however the low richness was captured well in our sampling as indicated by the Good's coverage, which is difficult because of the inherent complexity of the communities. Fungal OTU richness (S_{obs}, Figure 4.2) differed across treatments ($F_{3,87} = 3.2125$; $P = 0.0271$), over time ($F_{3,87} = 14.9070$; $P = 0.0002$) and there was a significant treatment by time interaction ($F_{3,87} = 3.4213$; $P = 0.0210$). The fungal OTU richness was higher in the uncovered (NN) treatment than in the covered treatments (Tukey's HSD at $\alpha = 0.05$). Multiple linear regression analyses show a slight but significant increase in S_{obs} in treatment NN during the six-month incubation, whereas the covered treatments (NT, PN and PT) remained largely unchanged (Table 4.2; Figure 4.2). In contrast to S_{obs}, the extrapolative Chao1 richness estimates neither differed between treatments or over time nor was there a treatment by time interaction (ANOVA: $F_{7,83} = 1.0932$, $P = 0.3752$; Regression analysis: treatment: $F_{3,87} = 0.0291$, $P = 0.9933$, time: $F_{3,87} = 1.3615$, $P = 0.2466$ and treatment by time $F_{3,87} = 2.0411$, $P = 0.1144$). Diversity (1-D; Figure 4.3) varied over time ($F_{3,87} = 20.6330$, $P = <0.0001$) but not between the treatments ($F_{3,87} = 1.2418$, $P = 0.2999$) or show significant treatment by time interaction ($F_{3,87} = 0.3618$, $P = 0.7808$). Further evaluation of the differences over time using multiple linear regression analyses indicated an increase in diversity over time (Table 4.2), which is consistent with the observed increase in richness. In contrast to 1-D, evenness (E_D) did not differ between treatments or change over time, but there was a significant treatment by time interaction ($F_{3,87} = 4.6298$, $P = 0.0048$) (Figure 4.4). Based on our multiple linear regression analysis this interaction is mainly attributable to the decrease in evenness in treatment NN (Tukey's HSD at $\alpha = 0.05$) but not in the treatments covered during the storage. This decrease in evenness in NN suggests a few taxa becoming dominant over the storage period, while remained stable in the covered treatments. In addition to increased richness, the increase in diversity over time is likely due to the increase in evenness in the covered treatments (NT, PN and PT), though linear regression analyses were not significant, all slopes had a positive estimate, except for the uncovered treatment (NN).

Based on the decline in stress we selected two-dimensional solutions (k = 2; stress = 0.136; $R^2 = 96.1\%$). The two-dimensional NMDS solution captured the variability in our data set well: Axis 1 represented 72.8% and Axis 2 23.1% of the variance, 95.9% in total. The NMDS

ordination allowed for a visual representation of fungal communities in different storage conditions during six-month storage. The differences among treatments or between different storage periods were evaluated through AMOVA (Table 4.3). AMOVA indicated that fungal communities were distinct among the storage conditions (Figure 4.5): the covered treatments (NT, PN and PT) were clustering together and were distinct from the uncovered treatment (NN). Similarly to storage conditions, storage duration had an impact on fungal community composition (Figure 4.6; Table 4.3). The communities in the end of the experiment (T6) differed from those in the beginning of our experiment (T0) or after 2 (T2) and 4 (T4) months of storage. However, the communities did not differ among the three first sampling times (T0-T4). We interpret this to indicate accelerating community dynamics during the last two months of storage. This acceleration is likely attributable to environmental conditions that may have facilitated fungal growth in our biomass, as T0, T2, T4 and T6 sampling occurred in October, December, February and April, respectively. HOMOVA indicated no differences among the variances within the communities in different storage treatments or over storage time.

Correlation with the uncovered and covered treatments was determined using Spearman's Rank Correlation Coefficient. OTUs correlated with the uncovered and covered treatments are shown in Tables 4.4 and 4.5, respectively, with a description of the Family ecology.

Based on the individual responses from the 100 most abundant OTUs (FDR; q = 0.05), the influence of individual OTUs in the treatments could be determined. Ten OTUs (10%) responded to storage treatments (Table 4.6), frequencies of four OTUs varied with storage duration (Table 4.7), and three OTUs showed a significant treatment by time interaction effect (Table 4.8). OTUs 413 (*Valsonectria*), 285 (*Sphaeronaemella*), 548 (*Pseudombrophila*), 515 (*Sphaeronaemella*), and 469 (*Pyxidiophora*) were more abundant in the uncovered (NN) than in the covered treatments, but did not differ among the covered treatments (NT, PN and PT; Table 4.6). Only OTU 545 (*Wallemia*) was more abundant in treatment NT compared to all other treatments (NN, PN and PT; Table 4.6). OTU 559 (*Valsaria*) was more abundant in treatment PT than in NN and NT (Table 4.6). Four OTUs shifted in frequencies during the six-month storage. OTUs 560 (*Cladosporium*), 566 (*Eurotium*) and 559 (*Valsaria*) increased over time during 6-month storage, whereas OTU 542 (*Alternaria*) decreased (Table 4.7). Three OTUs showed significant treatment by time interactions. OTU 548 (*Pseudombrophila*) increased in the

uncovered treatment (NN) during storage compared to treatment PT, while treatments NT and PT decreased (Table 4.8). OTU 559 (*Valsaria*) decreased in treatments NN and NT during storage compared to treatment PT, while treatment PN also decreased over time but to a much lesser extent. OTU 545 (*Wallemia*) rapidly increased over time in treatment NT compared to PT, but tended to decrease in treatments NN and PN during storage (Table 4.8). A further breakdown of the family and genera associated with each significant OTU from FDR analysis is shown in Table 4.9 along with family ecology with reference and response effect.

Discussion

We analyzed fungal communities and their dynamics in biomass stored for up to six months under various conditions and sampled every two months. Results from a previous study indicated increased enzymatic activity associated with cellulose and hemicellulose degradation and reduced cellulose and conversion to ethanol of sorghum biomass stored uncovered over a six-month period. Enzymatic activity, cellulose and ethanol conversion of sorghum biomass covered during storage were found to remain stable (Rigdon et al., 2013). These results motivated the current study to evaluate whether there were corresponding concomitant shifts and differences in the communities most likely responsible for the sorghum biomass degradation during storage. Our study represents an extensive and unique evaluation of fungal community dynamics in stored biomass using high-throughput pyrosequencing. Our results highlight the fungal community dynamics during storage and their responses to management during storage.

The most commonly observed genera in our biomass (*Cladosporium, Alternaria* and *Cryptococcus*) are all common air- and soil-borne fungi, with multiple species of *Alternaria* being recognized plant pathogens. Our data highlight temporal dynamics in the stored biomass even in the coarsest taxonomic levels. In the beginning of the storage, the fungal communities were expectedly similar across the treatments. However, after six-month storage, the proportion of Ascomycetes increased markedly in the uncovered treatment, but to a lesser degree in the covered ones. This abundance of Ascomycota over Basidiomycota over the storage duration in the sorghum biomass are in alignment with the assessment of the successional dynamics of rye residues by Poll et al. (2010). This is in contrast with the terminal, late successional basidiomycetous communities suggested to establish in the latter stages of litter decomposition (see Frankland, 1998). However, the observed patterns may be largely attributable to the short

incubation time driven by our focus on a realistic harvest and storage cycle for biomass destined for bioethanol production.

It is perhaps unsurprising that the sorghum biomass was so strongly dominated by *Cladosporium*. It is one of the most common genera of airborne fungi, but includes also a number of saprobic, phytopathogenic, and endophytic species. Endophytic *Cladosporium* is common in switchgrass (*Panicum virgatum*), that is similar to Sorghum, is a C_4 perennial grass native to North America (Kleczewski et al. 2012). Kleczewski et al. (2012) concluded that early plant biomass decomposition is dominated by endophytic fungi, like *Cladosporium*, likely because of their colonization of plant biomass prior to plant senescence or death. In this case, *Cladosporium* was likely present in the sorghum biomass at harvest and continued to inhabit the biomass after harvest and during storage, with the continued introduction of additional *Cladosporium* genera in biomass left uncovered. The sorghum biomass contained an abundance of cellulose and hemicellulose for the continued colonization and endophytic-to-saprobic activity of *Cladosporium* found in our system. Enzymes related to cellulose and hemicellulose degradation were found to increase in uncovered biomass during storage, though whether these enzymes came specifically from *Cladosporium* was not evaluated (Rigdon et al., 2013). Biomass storage without coverage permits biomass colonization by common airborne fungi with saprobic characteristics, such as *Cladosporium,* with the potential to degrade fermentable biomass components, therefore compromising the commercially viable use of the stored biomass (Rigdon et al., 2013). Biomass storage with coverage produces an environment inhibitory to the continued degradation of the available sorghum biomass, specifically cellulose and hemicellulose, as indicated here and reduced enzyme activity seen by Rigdon et al. (2013).This emphasizes the importance of maintaining stable storage conditions that inhibit and reduce fungal activity during storage, thus maintaining the biomass integrity for intended downstream uses, like lignocellulosic ethanol production.

Our coverage estimators suggested a near complete sampling, whereas the rarefaction analyses suggested that more taxa remained to be detected. Regardless, our estimators suggested that the sampling was equally efficient sampling across the four treatments and the four sample time points. The two methods to estimate sampling effort differ fundamentally and the lack of congruence highlights the different views that these tools provide.

The fungal community responses to storage conditions and duration were particularly clear with OTU richness. OTU richness (S_{obs}) differed among treatments and over time. The significant time by treatment interaction highlighted the dramatic increase in richness in treatment NN with increased incubation time. We hypothesized an increase in richness of treatment NN as a result of the continuous deposition of allochthonous propagules on the exposed bales. However, the decrease in evenness in the uncovered bales suggests a few taxa becoming dominant over time. It is likely that these dominant taxa are increasing in abundance over time resulting in the decrease in evenness and increased richness, which is not attributed to the continued deposition of allochthonous propagules throughout storage. In contrast to the uncovered bales, richness estimators in the covered treatments (NT, PN and PT) remained stable with only a slight increase over the six-month incubation. We attribute this lack of increase in observed species richness in the covered treatments to the prevention of the propagule introduction due to the coverage of the plastic and/or tarp.

Previous research by Muller (2005) found bales wrapped in plastic or bale stretch wrap (both used in this study) had a higher carbon dioxide content during storage, indicating a lack of diffusion of gases (carbon dioxide, oxygen) through the plastic/wrap, thus creating an anaerobic environment. Though we did not monitor oxygen or carbon dioxide levels in the covered bales, it is likely an anaerobic environment was reached. In addition, treatments PN and PT were found to have increased abundance of *Cyrenella*, *Alternaria* and *Valsaria*. The genus *Alternaria* are ubiquitous in the environment and have been shown to continue to grow (hyphal elongation) and sporulate in environments lacking oxygen or slightly elevated levels of carbon dioxide (Lukens and Horsfall, 1972). Both *Cyrenella* and *Valsaria* are yeasts, and likely capable of growth in oxygen limiting environments. In contrast to the plastic and/or tarp wrapped treatments, simple coverage with a tarp (treatment NT) likely created a more aerobic environment Furthermore, coverage with a tarp prevented moisture seepage into the bales, but also allowed for moisture to escape the bales, as indicated by the stable dry matter content observed by Rigdon et al. (2013) Interestingly, treatment NT was found to have the highest increase in abundance in the genus *Wallemia* over time (based on ANOVA with FDR of 0.05). Members of the genus *Wallemia* have long been known to be xerophilic/xerotolerant. Thus, an environment with very little moisture, like that of treatment NT, would be conducive for outward growth of this genus (Zalar et al., 2005). In addition to the importance of preventing the arrival of saprobic propagules,

preventing growth and metabolic activity is also important. Biomass coverage reduces growth and metabolic activity, as evidenced by the lack of enzyme activity reported by Rigdon et al. (2013).

Although OTU richness increased over time, more so in treatment NN compared to the others, diversity (1-D) was found to increase over time in all treatments. This overall increase in diversity (1-D) in all treatments over time is attributable to the increased richness and likely indicates an increase in heterogeneity of community in the stored sorghum biomass. This is congruent with our expectations, as those members of community better suited for growth on the sorghum biomass and environment created by coverage, or lack thereof, would have a competitive advantage and continue to grow throughout the 6-month storage period. In contrast to diversity, evenness (E_D) was found to decrease significantly in only treatment NN, while maintaining an upward trajectory in the covered treatments. This decline in evenness in treatment NN suggests the dominance of a few taxa over the 6-month storage time. Based on FDR analysis OTUs 413, 285, 548, 515 and 469, identified as *Valsonectria*, *Sphaeronaemella*, *Pseudombrophila*, *Sphaeronaemella* and *Pyxidiophora*, respectively, were found to be most abundant in the uncovered treatment (NN). Cosmopolitan distribution, plant pathogenicity, and associations of growth on plant or cellulosic materials (building materials) are family and genera characteristics of these OTUs. Cosmopolitan distribution of these OTUs would explain their abundance in the uncovered biomass bales. Their continued persistence in the cellulose-rich sorghum biomass is attributed to their ability to degrade cellulose used as an energy source. Though little is known about the genus *Valsonectria* (originally described as *Diaporthe*), it has been identified as a serious parasite in chestnuts, with mycelial growth throughout the bark and wood of infected trees (Stevens, 1913). Research by Vakili (1992) identified *Sphaeronaemella* as a parasitic biocontrol agent against fusaria in corn seedlings. The increase in *Pseudombrophila* in the uncovered sorghum biomass is congruent with findings by Hansen et al. (2005), in which the genus was isolated from decaying stems and leaves of plants. *Pyxidiophora* species have been isolated from southern pine beetle-infested loblolly pine trees and other mite and beetle habitats in trees and wood (Blackwell, 1986 and Blackwell et al., 1989). Based on the collective substrata these four genera of fungi have been found on, it is likely they possess enzymes needed for the breakdown of cellulose or are parasitic to other fungi present. This would allow them thrive in the sorghum biomass left uncovered over the 6-month storage period.

Our NMDS and subsequent AMOVA analyses indicated that the fungal communities in *Sorghum* biomass stored uncovered were distinct from those in covered biomass. Furthermore, the communities were relatively homogenous among the covered treatments and varied only little temporally. From homogenous fungal communities in the harvested biomass, the environmental conditions in the uncovered biomass facilitated a fungal community succession in the uncovered biomass that become increasingly distinct from the covered treatments. Notably, the fungal community richness was temporally stable in the covered biomass but increased in the uncovered treatments highlighting the arrival and establishment of new airborne propagules.

Our observations on the fungal richness, diversity and evenness clearly indicate that biomass covering is critical to maintain substrate quality during extended storage. It is likely that the fungal communities would continue shifting towards those likely to possess the extra-cellular enzymes needed to break down cellulose, hemicellulose and lignin if the storage period was further extended beyond 6 months employed in this study. We base this speculation on conclusions from studies focusing on fungal community succession in the latter stages of litter decomposition (see Frankland 1998). While sequence-based annotations must be considered with caution (Nilsson et al., 2011), it is tempting to interpret our data in terms of controlling the plant tissue inhabiting pathogens, saprobes and endophytes for improved preservation of stored biomass. These data strongly support our conclusion above: covering stored biomass strongly limits arrival and establishment of new fungal propagules in the stored biomass.

References

Aime, M.C., Matheny, P.B., Henk, D.A., Frieders, E.M., Nilsson, R.H., Piepenbring, M., McLaughlin, D.J., Szabo, L.J., Begerow, D., Sampaio, J.P., Bauer, R., Weib, M., Oberwinkler, F., Hibbett, D.. 2006. An overview of the higher level classification of Puccinomycotina based on combined analyses of nuclear large and small subunit rDNA sequences. Mycologia 98, 896-905.

Amend, A.S., Seifert, K.A., Bruns, T.D.. 2010. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol. Ecol. 19, 5555-5565.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral. Ecol. 26, 32-46.

Bensch, K., Groenewald, J.Z., Dijksterhuis, J., Starink-Willemse, M., Anderson, B., Summerell, B.A., Shin, H.-D., Dugan, F.M., Schroers, H.-J., Braun, J., Crous, P.W., 2010. Species and ecological diversity within the *Cladosporium cladoporioides* complex (*Davidiellaceae, Capnodiales*). Stud. Mycol. 67, 1-94.

Berry, D., Mahfoudh, K.B., Wagner, M., Loy, A., 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl. Environ. Microbiol. 77, 7846-7849.

Blackwell, M., 1986. A new species of *Pyxidiophora* and its *Thaxteriola* anamorph. Mycologia 78, 605-612.

Blackwell, M., Moser, J.C., Wisniewski, J., 1989. Ascospores of *Pyxidiophora* on mites associated with beetles in trees and wood. Mycol. Res. 92, 397-403.

Cain, R.F., Weresub, L.K., 1957. Studies of coprophilous Ascomycetes V. *Sphaeronaemella fimicola*. Can. J. Botany. 35, 119-131.

Edgar R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011 UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194-2200.

Emery, I.R., Mosier, N.S., 2012. The impact of dry matter loss during herbaceous biomass storage on net greenhouse gas emissions from biofuels production. Biomass Bioenerg. 39, 237-246.

Frankland, J.C., 1998. Fungal succession—unraveling the unpredictable. Mycol. Res. 102, 1-15.

Gao, J., Qian, L., Thelen, K.D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103, 844-850.

Gibbons W.R., Hughes, S.R., 2009. Integrated biorefineries with engineered microbes and high-value co-products for profitable biofuels production. In Vitro Cell. Dev.-Pl. 45, 218-228.

Gihring, T.M., Zhang, G., Brandt, C.C., Brooks, S.C., Campbell, J.H., Carroll, S., Criddle, C.S., Green, S.J., Jardine, P., Kostka, J.E., Lowe, K., Mehlhorn, T.L., Overholt, W., Watson, D.B., Yang, Z., Wu, W., Schadt, C.W., 2011. A limited consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. Appl. Environ. Microb. 77, 5955-5965.

Hansen, K., Perry, B.A., Pfister, D.H., 2005. Phylogenetic origins of two cleistothecial fungi, *Orbicula parietina* and *Lasiobolidium orbiculoides*, within the operculate discomycetes. Mycologia 97, 1023-1033.

Hess J.R., Wright, C.T., Kenney, K.L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel. Bioprod. Bior. 1, 181-190.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Fox, C.J., et al., 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443, 818-822.

Kleczewski, N.M., Bauer, J.T., Bever, J.D, Clay, K., Reynolds, H.L., 2012. A survey of endophytic fungi in switchgrass (*Panicum virgatum*) in the Midwest, and their putative roles in plant growth. Fungal Ecol. 5, 521-529.

Lukens, R.J., Horsfall, J.G., 1972. Processes of sporulation in alternaria solani and their response to metabolic inhibitors. Phytopathology 63, 176-182.

Magauran, A.E., 1988. *Ecological Diversity and its Measurement*. Princeton University, Princeton.

Margulies, M., et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376-380.

Mendez-Mayboca, F., Checa, J., Esqueda, M., Chacon, S., 2010. New records of *Loculoascomycetes* from natural protected areas in Sonora, Mexico. Mycotaxon 111, 19-30.

Muller, C.E., 2005. Fermentation patterns of small-bale silage and haylage produced as a feed for horses. Grass Forage Sci. 60, 109-118.

Nilsson, R.H., Tedersoo, L., Lindahl, B.D., Kjoller, R., Carlsen, T., Quince, C., Abarenkov, K., et al., 2011. Towards standardization of the description and publication of next-generation sequencing datasets of fungal communities. New Phytol. 119, 314-318.

Poll, C., Brune, T., Begerow, D., Kandeler, E., 2010. Small-scale diversity and succession of fungi in the detritusphere of rye residues. Microb. Ecol. 59, 130-140.

Quince C, Lanzén, A., Curtis, T.P., Davenport, R.J., Hall, N., Head, I.M., Read, L.F., Sloan, W.T., 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. Nat. Methods 6, 639.

Quince C, Lanzen, A., Davenport, R.J., Turnbaugh, P.J., 2011. Removing noise from pyrosequenced amplicons. BMC Bioinformatics 12, 38.

Rigdon, A.R., Jumpponen, A., Vadlani, P.V., Maier, D.E., 2013. Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop. Bioresource Technol. 132, 269-275.

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:e27310.

Stevens, F.L., 1913. The Fungi Which Cause Plant Disease-book. The MacMillian Company, New York

Tedersoo, L., Nilsson, R.H., Abarenkov, K., et al., 2010. 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. New Phytol. 188, 291-301.

Vakili, N.G., 1992. Biological seed treatment of corn with mycopathogenic fungi. J. Phytopathology 134, 313-323.

Vilgalys, R., Hester, M., 1990. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. J. Bacteriol. 172, 4238-4246.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microb. 73, 5261-5267.

Zalar, P., de Hoog, G.S., Schroers, H.-J., Frank, J.M., Gunde-Cimerman, N., 2005. Taxonomy and phylogeny of the xerophilic genus *Wallemia* (Wallemiomycetes and Wallemiales, cl. et ord. nov.). Anton. Leeuw. Int. J. G. 87, 311-328.

Tables and Figures

Table 4.1 The overall 20 most abundant OTUs found, including OTU number, number of sequences and taxonomic information.

OTU	# of Seq.	Class	Order	Family	Genus
Otu554	59382	Dothideomycetes	Capnodiales	Davidiellaceae	*Cladosporium*
Otu542	**36447**	**Dothideomycetes**	**Pleosporales**	**Pleosporaceae**	***Alternaria***
Otu538	**5142**	**Tremellomycetes**	**Tremellales**	**Tremellaceae**	***Cryptococcus***
Otu566	3369	Eurotiomycetes	Eurotiales	Trichocomaceae	*Eurotium*
Otu541	2729	Sordariomycetes	Hypocreales	Nectriaceae	*Gibberella*
Otu515	2718	Sordariomycetes	Microascales	Microascales incertae sedis	*Sphaeronaemella*
Otu532	2150	Tremellomycetes	Tremellales	Tremellaceae	*Filobasidiella*
Otu545	1949	Wallemiomycetes	Wallemiales	Wallemiaceae	*Wallemia*
Otu560	1596	Dothideomycetes	Capnodiales	Davidiellaceae	*Cladosporium*
Otu425	**1405**	**Sordariomycetes**	**Sordariomycetes incertae sedis**	**Glomerellaceae**	***Colletotrichum***
Otu553	972	Dothideomycetes	Capnodiales	Davidiellaceae	*Davidiella*
Otu559	633	Sordariomycetes	Diaporthales	Diaporthales incertae sedis	*Valsaria*
Otu526	594	Tremellomycetes	Tremellales	Tremellaceae	*Cryptococcus*
Otu534	**571**	**Sordariomycetes**	**Magnaporthales**	**Magnaporthaceae**	***Magnaporthe***
Otu531	413	Saccharomycetes	Saccharomycetales	Saccharomycodaceae	*Hanseniaspora*
Otu561	380	Sordariomycetes	Hypocreales	Bionectriaceae	*Hydropisphaera*
Otu548	377	Pezizomycetes	Pezizales	Pyronemataceae	*Pseudombrophila*
Otu499	300	Saccharomycetes	Saccharomycetales	Saccharomycetaceae	*Pichia*
Otu492	244	Saccharomycetes	Saccharomycetales	Metschnikowiaceae	*Clavispora*
Otu460	**209**	**Cystobasidiomycetes**	**Cystobasidiomycetes incertae sedis**	**Cystobasidiomycetes incertae sedis**	***Cyrenella***
	121580				

(OTUs highly correlated to treatment NN in bold; OTUs highly correlated to treatments NT, PN and PT underlined)

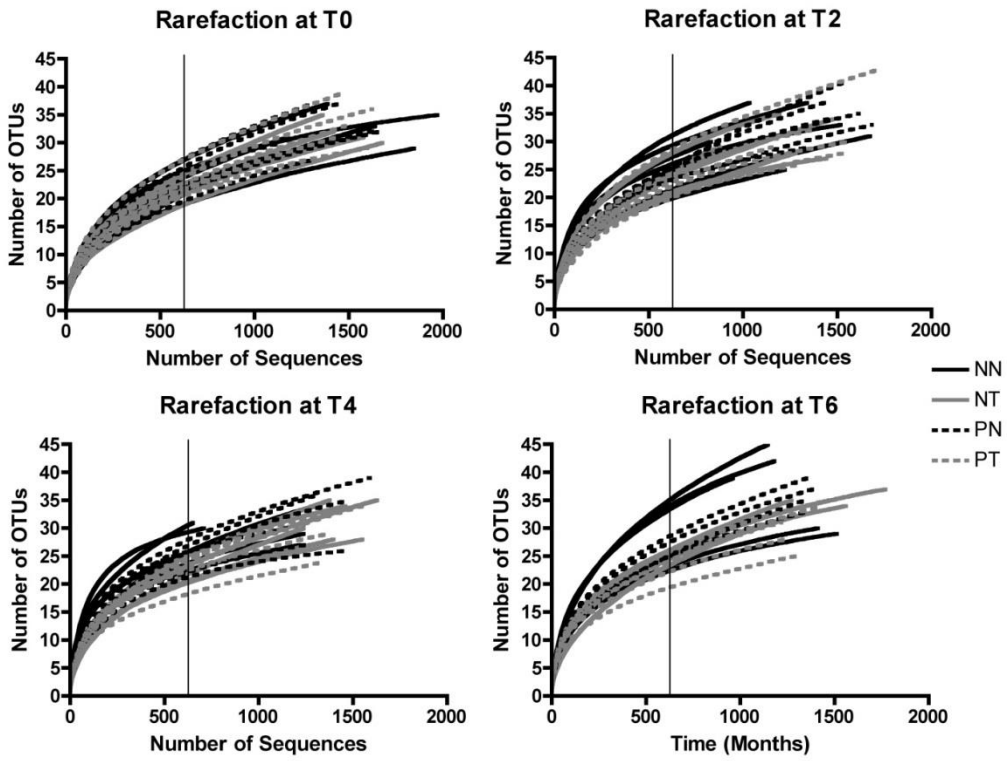


Figure 4.1 Rarefaction curves after 0, 2, 4 and 6-month incubation periods with subsampling cut off after 652 sequences indicated by the vertical line. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

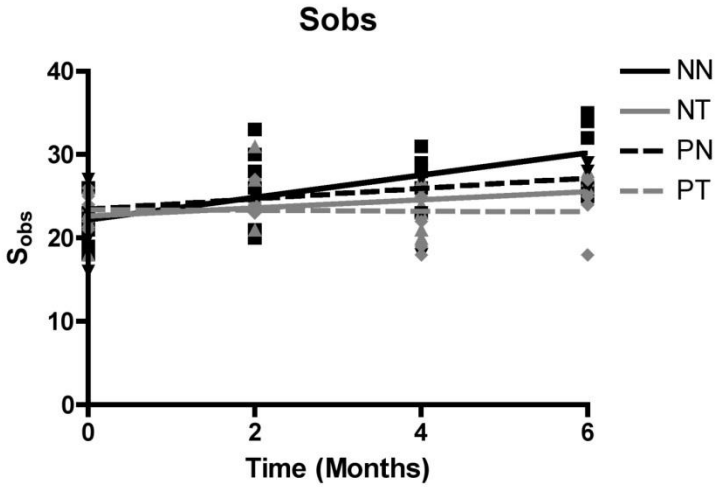


Figure 4.2 Richness across treatments and time based on observed species (S_{obs}) with differences between treatments, over time and treatment by time interactions. (NN

indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

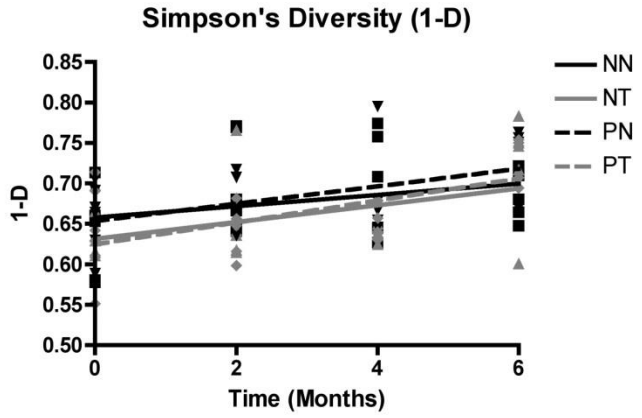


Figure 4.3 Diversity across treatments and time based on the complement of Simpson's Diversity Index (1-D) with diversity in all treatments increasing over time but no significant differences found between treatments or treatment by time interaction. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

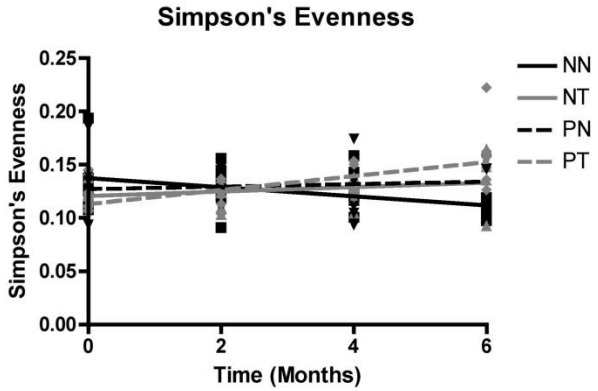


Figure 4.4 Evenness across treatment and time based on Simpson's Index, with uncovered treatment decreasing over six months of storage. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Table 4.2 Multiple Linear Regression Analyses. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Response	Estimate	Estimate STD	t-value	P-value**
Good's Coverage				
Intercept (NN)	0.0002225	0.00072997	0.30474	0.76133
Intercept (NT)	-0.000605	0.00072997	-0.8281	0.409966
Intercept (PN)	-0.000249	0.00074148	-0.3358	0.737849
Intercept (PT)*	***0.9866576***	***0.00069382***	***1422.07***	***7.76E-184***
Slope (NN)	***-0.000894***	***0.00033358***	***-2.6812***	***0.008848***
Slope (NT)	-0.000112	0.00033358	-0.3351	0.738373
Slope (PN)	0.0002101	0.00034465	0.6095	0.543859
Slope (PT)*	-0.000274	0.00019421	-1.4109	0.162007
S_{obs}				
Intercept (NN)	***1.3632064***	***0.59096016***	***2.30677***	***0.02356***
Intercept (NT)	-0.607559	0.59096016	-1.0281	0.306896
Intercept (PN)	0.6157508	0.60027788	1.02578	0.307977
Intercept (PT)*	***22.912838***	***0.5616939***	***40.7924***	***1.09E-56***
Slope (NN)	***0.7412326***	***0.27005301***	***2.74477***	***0.007421***
Slope (NT)	-0.10861	0.27005301	-0.4022	0.688585
Slope (PN)	0.0215853	0.27901572	0.07736	0.938521
Slope (PT)*	***0.6070379***	***0.1572248***	***3.86095***	***0.000223***
1-D				
Intercept (NN)	0.0061813	0.00863378	0.71594	0.476038
Intercept (NT)	-0.010531	0.00863378	-1.2197	0.22603
Intercept (PN)	0.0126051	0.00876991	1.43731	0.154389
Intercept (PT)*	***0.6418194***	***0.00820621***	***78.2115***	***1.59E-79***
Slope (NN)	-0.003495	0.00394541	-0.8859	0.378217
Slope (NT)	-1.22E-05	0.00394541	-0.0031	0.997532
Slope (PN)	0.0003434	0.00407635	0.08424	0.93307
Slope (PT)*	***0.0104339***	***0.00229701***	***4.54236***	***1.87E-05***
Evenness				
Intercept (NN)	-0.0033	0.00392861	-0.8401	0.403283
Intercept (NT)	-0.001897	0.00392861	-0.4827	0.630545
Intercept (PN)	0.0019377	0.00399055	0.48557	0.628554
Intercept (PT)*	***0.1246028***	***0.00373405***	***33.3694***	***7.32E-50***
Slope (NN)	***-0.00565***	***0.00179527***	***-3.1474***	***0.002288***
Slope (NT)	0.0006859	0.00179527	0.38208	0.703376
Slope (PN)	-0.000207	0.00185485	-0.1117	0.911294

Slope (PT)*	0.0013868	0.0010452	1.32678	0.188219

* = Treatment PT was selected as a reference level to emphasize the contrast between the three covered treatments (PT, PN, NT) and the uncovered treatment (NN)

** = P-values test the null hypotheses (H_0: Intercept $_{PN, NT, or NN}$ − Intercept $_{Ref PT}$ = 0; and H_0: Slope $_{PN, NT, or NN}$ − Slope $_{Ref PT}$ = 0). In other words, significant P-values here indicate that the difference between intercept or slope terms for treatments PN, NT, or NN differ from that of PT.

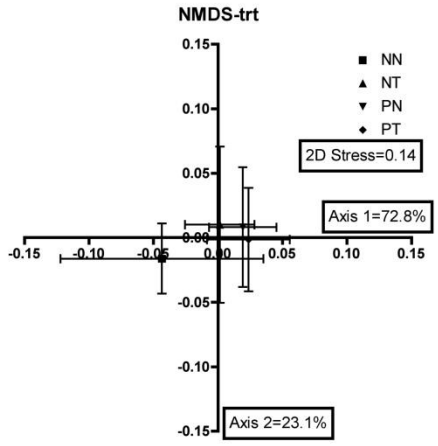


Figure 4.5 Non-metric Multi-dimensional Scaling (NMDS) by treatment. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

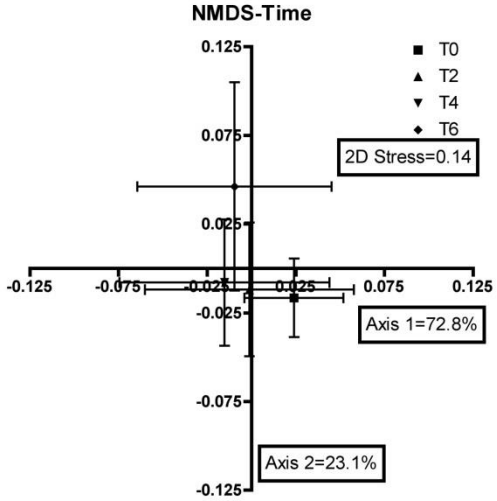


Figure 4.6 Non-metric Multi-dimensional Scaling (NMDS) by treatment. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Table 4.3 Analysis of Molecular Variance (AMOVA) results for NMDS by treatment and by time. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

AMOVA Results for NMDS by treatment		AMOVA Results for NMDS by time	
NN-NT-PN-PT*	***$F_{3,87}$=8.010, p=<0.001***	***T0-T2-T4-T6****	***$F_{3,87}$=7.771, p=<0.001***
NN-NT*	***$F_{1,44}$=8.285, p=0.001***	T0-T2	$F_{1,46}$=3.975, p=0.032
NN-PN*	***$F_{1,43}$=11.464, p=<0.001***	***T0-T4****	***$F_{1,46}$=8.951, p=0.001***
NN-PT*	***$F_{1,44}$=13.604, p=<0.001***	***T0-T6****	***$F_{1,41}$=19.569, p=<0.001***
NT-PN	$F_{1,43}$=3.693, p=0.051	T2-T4	$F_{1,46}$=1.586, p=0.237
NT-PT	$F_{1,44}$=4.289, p=0.032	***T2-T6****	***$F_{1,41}$=9.209, p=0.001***
PN-PT	$F_{1,43}$=1.274, p=0.272	***T4-T6****	***$F_{1,41}$=6.377, p=0.001***

Based on Bonferroni correction for multiple comparison of 0.0083; * indicates significance at $P < 0.0083$).

Table 4.4 OTUs correlated with the uncovered treatment (NN-no coverage) with family and associated family ecology. Correlation based on significantly negative axis scores for axis 1 with negative axis 2 scores.

OTU	Axis 1 Score	Axis 1 P-value	Axis 2 Score	Axis 2 P-value	Family	Associated Family Ecology
542	-0.491048	0.000001	-0.847852	0.0	Pleosporaceae	Necrotrophic pathogens and saprobes, especially with grasses
460	-0.424982	0.000027	-0.201851	0.054903	Cystobasidiomycetes incertae sedis	Many mycoparasites; saprobes; pathogens of plant roots
425	-0.3888948	0.000139	-0.077996	0.45934	Glomerallaceae	Hemibiotrophic-initially endophytic followed by being necrophilic
538	-0.358399	0.000486	-0.429495	0.000022	Tremellaceae	Usually grow on woody substrata, often parasitic on other fungi
534	-0.353792	0.000581	-0.124184	0.238752	Magnaporthaceae	Usually necrotophic on roots

Table 4.5 OTUs correlated with covered treatments (NT-tarp, PN-plastic and PT-plastic and tarp) with family and associated family ecology. Correlation based on significantly positive axis scores for axis 1 with positive axis 2 scores.

OTU	Axis 1 Score	Axis 1 P-value	Axis 2 Score	Axis 2 P-value	Family	Associated Family Ecology
555	0.389697	0.000134	0.411214	0.000051	Wickerhamomycetaceae	Yeast, growth on a wide range of carbon sources
548	0.38197	0.000187	0.275963	0.008103	Pyronemataceae	Saprobic on soil or rotten wood or mycorrhizal
499	0.355812	0.000538	0.326596	0.00158	Saccharomycetaceae	Yeast, cosmopolitan distribution
515	0.342304	0.000896	0.234788	0.025079	Microascales incertae sedis	Saprobic, live in soil, rotting vegetation, some plant pathogenic
492	0.33028	0.001387	0.366257	0.000356	Metschnikowiaceae	Necrotrophic on plant tissue

Table 4.6 Results from FDR by treatment

OTU	F statistic	Treatment Responses
413	$F_{3,86}$=<0.0001	NN different from NT, PN, PT; NN highest mean
285	$F_{3,86}$=<0.0001	NN different from NT, PN, PT; NN highest mean
548	$F_{3,86}$=<0.0001	NN different from NT, PN, PT; NN highest mean
545	$F_{3,86}$=<0.0001	NT different from NN, PN, PT; NT highest mean
515	$F_{3,86}$=<0.0001	NN different from NT, PN, PT; NN highest mean
542	$F_{3,86}$=0.0002	NN different from NT, PN, PT; NN lowest mean
469	$F_{3,86}$=0.0005	NN different from NT, PN, PT; NN highest mean
460	$F_{3,86}$=0.0006	NN different from NT, PN, PT; NN lowest mean
559	$F_{3,86}$=0.0024	PT different from NN, NT (PN, NN, NT not different); PT highest mean
536	$F_{3,86}$=0.0028	NN different from PT, NT (PN, PT, NT not different); NN highest mean

Table 4.7 Results from FDR by time

OTU	Estimate	St. Error	t-ratio	t-test
560	2.7137	0.3316	8.18	<0.0001
566	5.6199	1.0279	5.47	<0.0001
542	-7.2159	1.6311	-4.42	<0.0001
559	1.9318	0.4761	4.06	0.0001

Table 4.8 Results from FDR for treatment by time interaction

OTU	F Statistic	Estimate	St. Error	t-ratio	t-test
548	$F_{3,86}$=<0.0001	NN-time= 3.7854	0.7578	2.86	<0.0001
		NT-time= -1.2618	0.7578	-1.67	0.0997
		PN-time= -1.2618	0.7830	-1.61	0.1109
559	$F_{3,86}$=<0.0001	NN-time= -1.8929	0.8177	-2.29	0.0245
		NT-time= -1.9035	0.8177	-2.33	0.0224
		PN-time= -0.6456	0.8449	-0.76	0.4470
545	$F_{3,86}$=0.0012	NN-time= -1.5955	1.3081	-1.22	0.2260
		NT-time= 5.3942	1.3081	4.12	<0.0001
		PN-time= -1.3368	1.3515	-0.99	0.3255

Table 4.9 OTU, family, genera, associated family ecology with reference and the response effect of those OTUs significant during FDR with bootstrap values in parenthesis analysis.

OTU	Family	Genus	Associated Family Ecology	Reference	Response
413	Bionectriaceae (100)	*Valsonectria* (100)	Cosmopolitan distribution; Pathogenic to crop plants, causing stem and root rots; Associated with plant material	Fungal Families of the World	Treatment effect- NN different from NT, PN and PT
285	Microascales incertae sedis (100)	*Sphaeronaemella* (100)	Species of *Sphaeronaemella* have been associated with plant materials and parasitic to *Eurotium* and *Microascus* species	Cain and Weresub, 1957	Treatment effect- NN different from NT, PN and PT
548	Pyronemataceae (100)	*Pseudombrophila* (100)	Cosmopolitan distribution; frequently found growing on damp plaster etc in buildings; saprobic on soil or rotten wood	Fungal Families of the World	Treatment effect- NN different from NT, PN and PT; Interaction effect- NN increasing over time
545	Wallemiaceae (100)	*Wallemia* (100)	Widely distributed; spoilage organism on dried or desiccated food products; saprobic, capable of growth over wide ranges of water tension	Fungal Families of the World	Treatment effect- NT different from NN, PN and PT; Interaction effect- NT increasing over time
515	Microascales incertae sedis (100)	*Sphaeronaemella* (100)	Genera of *Sphaeronaemella* have been associated with plant materials and parasitic to *Eurotium* and *Microascus* species	Cain and Weresub, 1957	Treatment effect- NN different from NT, PN and PT
542	Pleosporaceae (94)	*Alternaria* (94)	Necrotrophic pathogens and saprobes, especially with grasses	Fungal Families of the World	Treatment effect- NN different from NT, PN and PT; Time effect- decreasing over time
469	Pyxidiophoraceae	*Pyxidiophora*	Primarily found in northern	Fungal Families	Treatment effect- NN

	(100)	(100)	temperate climates; associated with arthropods	of the World	different from NT, PN and PT
460	Cystobasidiomycetes incertae sedis (100)	*Cyrenella* (100)	Little information; most isolated from marine environments	Aime et al., 2006	Treatment effect- NN different from NT, PN and PT
559	Diaporthales incertae sedis (100)	*Valsaria* (100)	Classification to family ambiguous, possibly Diaporthaceae; saprophytic; perennial canker in stone fruit trees	Mendez-Mayboca et al., 2010	Treatment effect- PT different from NN, NT; Interaction effect- NN, NT decreasing over time
536	Cystofilobasidiaceae (100)	*Guehomyces* (100)	Little know; presumably saprobic, more common in colder climates/conditions	Fungal Families of the World	Treatment effect- NN different from NT, PT
560	Davidiellaceae (100)	*Cladosporium* (100)	Genera of Cladosporium cosmopolitan distribution; saprobic; common endophytic or quiescent fungi	Bensch et al., 2010	Time effect- increasing 2.7137 units per month
566	Trichocomaceae (100)	*Eurotium* (100)	Ubiquitous in soil; saprobic; associated with decaying plant material	Fungal Families of the World	Time effect- increasing 5.6199 units per month

Chapter 5 - Fungal communities more responsive than bacterial in biomass stored for lignocellulosic ethanol production[3]

Abstract

Passing of the Energy Independence and Security Act (EISA) of 2007 put into motion the Renewable Fuel Standards (RFS) program and is the major driving force behind the biofuels industry in the US. RFS program mandates for the production of transportation fuels, primarily ethanol, has propelled the biofuels industry into unleashing the potential of lignocellulosic crops (perennial grasses, short rotation forestry crops) and agricultural and forest residues. Current production, harvest, transport and storage infrastructure and practices are designed for handling grain crops with uniform sizes and bulk density; however, lignocellulosic materials vary greatly in these characteristics. In addition, lignocellulosic materials are harvested annually or biannually, requiring extended storage for year-round production of ethanol at the biorefinery. Baled biomass, covered or uncovered outdoors, has been identified as the primary method for storage, thus leaving the biomass susceptible to degradation by environmental conditions and microbial attack. Many common air- and soil-borne fungi and bacteria inhabit stored biomass, are capable of breaking down its components, cellulose, hemicellulose and lignin, and may reduce the amount of material available for conversion to ethanol. Here, we evaluated the fungal and bacterial community composition and dynamics using direct 454-pyrosequencing in Sorghum biomass during six-month storage either wrapped in plastic (PN), covered with a tarp (NT), wrapped in plastic and covered with a tarp (PT) or stored with no cover (NN). Fungal and bacterial communities showed contrasting dynamics during storage and in response to covering treatments. Fungal communities in uncovered biomass shifted in composition diversity and evenness during storage, while in covered treatments these community attributes were comparatively stable. Visualization of the fungal communities using Non-metric Multi-Dimensional Scaling (NMDS) ordination indicated that they were indistinguishable among covered treatments, but distinct from the uncovered treatment. Bacterial communities did not respond to storage coverage and were stable over time in richness, diversity, evenness and composition by treatment and time. These results suggest that understanding fungal community

[3] To be submitted to Applied and Environmental Microbiology

dynamics are crucial in the maintenance of biomass integrity for lignocellulolytic bioethanol conversion, whereas bacterial community dynamics are likely to be of lesser importance.

Introduction

The fermentation of glucose into desired chemicals, like ethanol is perhaps the oldest product obtained through traditional biotechnology (Zaldivar et al., 2001). Ethanol was the original transportation fuel of choice when Henry Ford introduced the original model T car in the 1880s. However, petroleum-derived fuels began flooding the market as a cheaper alternative to ethanol beginning the US's dependence on fossil fuels. The "oil crisis" of the 1970s awakened the awareness of finite supply of fossil fuels, and re-surged the interest in ethanol as a transportation fuel. As prices began to stabilize, interest in ethanol faded, but was renewed in 2007 with the passing of the Energy Independence and Security Act (EISA), which initiated the Renewable Fuel Standards (RFS) program. The RFS program outlined goals for the production and utilization of transportation fuels, primarily ethanol, from renewable resources, with the goal of shifting current production from first generation feedstocks to second generation feedstocks Gao et al., 2011).

First generation feedstocks included beet and sugarcane for the direct fermentation of sugar to ethanol or starch-rich grains, like maize, to be enzymatically hydrolyzed and then fermented to ethanol. Second generation feedstocks include lignocellulosic materials like agricultural residues (wheat straw, corn stover, sorghum stalk), perennial grasses (switchgrass, miscanthus), short rotation forestry crops, municipal solid wastes and residues from the forestry industry. These lignocellulosic materials contain cellulose. Cellulose, the most abundant and renewable polymers found on the earth, is composed of thousands of molecules of glucose linked together via β(1,4)-glycosidic bonds (Zaldivar et al., 2001). Plant-derived cellulose is trapped in the complicated cellulose-hemicellulose-lignin matrix, making direct conversion of cellulose into a more useful chemical very difficult. Releasing the potential power of this untapped resource via microbial bioconversion to useable chemicals, like ethanol, has not fully been revealed. The complex lignocellulose matrix, designed by nature to be resistant to breakdown, must first have the lignin disrupted (pre-treatment) to release cellulose and hemicellulose for enzymatic hydrolysis into glucose to be fermented to ethanol.

While the technical challenges of converting lignocellulosic biomass to ethanol have been resolved, physical harvest, transport and storage logistics have not. The current infrastructure used for transporting agricultural products were designed around the harvest, transport and storage of grains, with uniform sizes and bulk densities. Harvest of lignocellulosic biomass results in material with varying sizes and bulk densities, making handling and transport difficult (Hess et al., 2007). In addition, most lignocellulosic biomass can only be harvested annually or biannually, requiring long-term storage to maintain a consistent supply, sufficient for a biorefinery to operate year-round. Prior research on lignocellulosic biomass storage has focused on maintaining the digestibility of the biomass as a cattle feed, not further downstream processing into ethanol. Baling biomass into large round or square bales has been identified as most cost effective way of collecting the biomass for transport to the biorefinery. In addition, bales can easily be stacked in different configurations for storage, with or without coverage, prior to processing into ethanol. Maintaining substrate quality, primarily cellulose, in the baled biomass is the primary goal during storage to ensure profitability by the biorefinery and maximum ethanol yields. Environmental factors (precipitation, wind, solar radiation) and microbial colonization present the biggest threat to cellulose degradation during storage of biomass. Many common fungal and bacterial strains found in the air and soils are capable of breaking down cellulose, thus making less available for bioconversion to ethanol. Few studies have fully detailed the resulting ethanol yields from biomass stored under various conditions over an extended period of time, i.e., storage after harvest to the next harvest.

The fungal and bacterial communities present in stored lignocellulosic biomass destined for conversion to ethanol have not been fully characterized. Moreover, the impacts of the fungal and bacterial communities residing in biomass during storage on biomass quality and ethanol yields have not been evaluated. In the studies reported here, we 454-pyrosequence analyzed the fungal and bacterial communities in baled sorghum biomass stored under varied conditions and durations. Fungal community analysis consisted of sequencing of amplicons produced from the internal transcribed spacer (ITS) region of the ribosomal RNA (rRNA) gene. The VI-VII region of the 16S rRNA gene was used for amplicon production and sequencing to evaluate the bacterial communities present. We hypothesized that 1) the richness and diversity of the uncovered treatment (NN) would increase in fungal and bacterial communities during the storage duration; 2) richness and diversity in the covered treatments (NT, PN and PT) would remain

unchanged during the storage period as the fungal and bacterial communities would be dominated by a few species due to reduced moisture fluctuations because of coverage; 3) the dominant taxa in the uncovered and covered treatments would differ; 4) the fungal and bacterial community constituents would change over time differently across the different storage conditions. Our results suggest that the fungal community is strongly responsive to storage duration in uncovered biomass, while remaining stable in covered biomass over time. In contrast, the bacterial community remained stable regardless of the storage method or duration. In conjunction with earlier analyses (Rigdon et al., unpublished), these results imply that fungal communities, rather than bacterial, should remain the focus for preservation efforts of stored biomass. Slowing of the fungal community dynamics via coverage is a simple, cost-effective and adequate method to protect stored biomass from fluctuations in the environmental conditions, thus maintaining biomass quality for lignocellulosic ethanol production.

Materials and Methods

Biomass Storage Experimental Design

Photo-period sensitive sorghum cultivar PS1990 from Sorghum Partners (Chromatin, New Deal, TX) was grown on the North Agronomy Farm, Department of Agronomy, Kansas State University, Manhattan, KS. The sorghum was planted June 3, 2010 and fertilized 15 to 20 days after planting at a rate of about 112 kg per ha using urea (N-P-K; 46-0-0) and herbicide-treated with Bicep II Magnum (Syngenta, Basel, Switzerland) at 5.6 l per ha. Sorghum was cut and windrowed October 7, 2010 and allowed to dry on site. On October 15, 2010, sorghum was baled into small bales (average size 0.36 by 0.46 by 0.91 m) using a Massey Ferguson (AGCO, Duluth, GA) square baler. A total of 96 bales were weighed and randomly assigned to one of four treatments (no plastic/no tarp [NN], no plastic/tarp [NT], plastic/no tarp [PN] and plastic/tarp [PT]) and to one of four sampling times (0-, 2-, 4- or 6- months after baling) for a total of six replicates per treatment and sampling time. The PN and PT bales were wrapped two to three times with Tytan Wrap Premium Silage Film (750 mm by 1500 m by 25.4 µm; Tytan International, LLC, Lenexa, KS) and placed in two extra-large, industrial strength black plastic trash bags (Husky Brand, 45 gallon with 1mm thickness, (Husky Plastics, Thornhill, Ontario, Canada). The bales were randomly assigned to 2-, 4-, and 6-month incubation, grouped by sampling time point, and stored in a single layer on wooden pallets at the North Agronomy Farm

(Manhattan, KS). Individual tarps (silver heavy duty from Erickson Manufacturing LTD., Marine City, MI) were placed over treatments NT and PT for each sampling time point with the tarp covering the bale tops and sides completely.

DNA Extraction and Amplicon Production

Each bale was ground using a FitzMill (FitzMill Comminutor, Fitzpatrick, Elmhurst, IL) with a screen size of 4.76 mm. The shredded biomass was collected and mixed thoroughly for two minutes using a twin shaft paddle mixer (Hayes & Stolz, Fort Worth, TX). From the homogenized mixture, a 1 kg sub-sample was collected for further particle size reduction through a Bliss Hammer Mill (Bliss Industries, Ponca City, OK) with screen size of 0.397 mm attached to a Craftsman ShopVac (Sears Holdings Corp., Hoffman Estates, IL). Samples ground through the Bliss Mill were collected, mixed and further sub-sampled (about 5 g) for DNA extraction and stored at -80°C.

From the previously prepared sub-samples stored at -80°C, a 1 gram aliquot was ground in liquid N_2 with a mortar and pestle until the N_2 evaporated. The grinding was repeated for a total of three times. The ground sample was transferred into a 50 mL DNA extraction tube (Qiagen DNeasy Plant Maxi Kit, Qiagen Inc, Valencia, CA, USA) and DNA was extracted according to the manufacturer's protocol, with an additional 5-min centrifugation added to minimize the ethanol carry-over after the addition of Buffer AW. The final elution was carried out in two steps, as recommended, with a 750 µl volume at each step. Extracted DNA was visualized on a 1% TBE agarose gel and quantified using a ND1000 spectrometer (NanoDrop Technologies, Wilmington, DE). Template DNA for each sample was aliquoted into 96-well plates at a working concentration of 10 ng μL^{-1} and stored at -20°C. Remaining DNA was archived at -80°C.

Fungal amplicons for direct 454-pyrosequencing of the internal transcribed spacer (ITS) region of the ribosomal RNA (rRNA) gene were generated using a two-step PCR protocol to reduce primer biases (Berry et al. 2011). In brief, triplicate primary PCR reactions were carried out in 50 µL reactions each containing 100 ng of template, 1 µM of each forward and reverse primer, 200 µM of each dioxynucleotide, 2.5 mM $MgCl_2$, 10 µL 5x Green GoTaq Flexi Buffer (Promega, Madison, WI), 9.6 µL DNase/RNase free water and 2 U of GoTaq Hot Start Polymerase (Promega, Madison, WI). Primary PCR reactions contained the forward primer

ITS1F (5'- CTTGGTCATTTAGAGGAAGTAA-3'; Gardes and Bruns, 1993) and the reverse primer ITS4 (5'- TCCTCCGCTTATTGATATGC-3') (White et al. 1990; see Appendix B, Table B.2). Cycle parameters for the primary PCR consisted of a 94°C initial denaturing for 2 min, then 4 cycles of denaturing at 94°C for 1 min, annealing at 57°C for 1 min with a 1°C decrease with each subsequent cycle and extension at 72° for 2 min, followed by 16 cycles of 94°C for 1 min, 54°C for 1 min and 72°C for 2 min and a final extension of 72°C for 8 min in a MasterCycler (Eppendorf, Hamburg, Germany). All PCR products were visualized on a 2% TBE agarose gel to ensure the presence of PCR products of the appropriate fragment size. Primary PCR products were purified using Diffinity RapidTip (Diffinity Genomics, Inc, West Henrietta, NY) according to the manufacturer's specifications and a 10 µL aliquot used for secondary PCR reactions. Secondary PCR reactions added sample-specific DNA barcode tags and the 454-linkage adaptors, the modified ITS1F-A primer included the 454-adaptor (A-primer) and 10 bp barcode tags unique to each sample with the modified ITS4-B primer containing the 454-adaptor (B-primer) as described in Jumpponen and Jones (2010) (see Appendix B, Table B.2). Secondary PCR reactions were also carried out in 50 µL reactions as described above but with 10 µL of purified primary PCR product for a template. Secondary PCR reaction parameters were identical to the primary PCR but only 5 cycles of denature at 94°C, annealing at 54°C and extension at 72°C were carried out. Positive controls of *Saccharomyces cerevisiae* genomic DNA and negative controls without template DNA were included.

Bacterial amplicons for direct 454-pyrosequencing of the 16S ribosomal RNA (rRNA) gene were also generated using the two-step PCR amplification. The primary PCR included the forward primer 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and the reverse primer 338R (5'-CATGCTGCCTCCCGTAGGAGT-3'), similar to Fierer et al. (2008; see Appendix B). In the secondary PCR reaction, the primer constructs included the forward primer 27F combined with 10 bp barcodes and the 454-adaptor (A-primer) and reverse primer 338R combined with the 454-adaptor (B-primer) for the secondary PCR reactions (see Appendix B, Table B.3). Primary PCR amplification was carried out in 50 µL reactions each containing 100 ng of template, 1 µM of each forward and reverse primer, 200 µM of each dioxynucleotide, 5.0 µL DNase/RNase free water and 25 µL of AmpliTaq Gold 360 Master Mix (Applied Biosystems, Carlsbad, CA). Cycle parameters for the primary PCR were a 94°C initial denaturing for 3 min, then 30 cycles of denaturing at 94°C for 45 sec, annealing at 50°C for 30 sec and extension at 70° for 90 sec with a

final extension of 72°C for 10 min in a MasterCycler (Eppendorf, Hamburg, Germany). Secondary PCR amplification was carried out with the same reaction formulation but with 10 µL of primary PCR product instead of template and only 5 cycles of denature, anneal and extension.

Fungal and bacterial secondary PCR products were purified and normalized with SequalPrep Normalization 96-well Plate (Life Technologies, Grand Island, NY) according to the manufacturer's specifications. The purified products were pooled at equal volumes and purified again using Agencourt AMPure XP (Beckman Coulter Inc, Brea, CA) as recommended, with the modification of 1:1 PCR-product to bead loading ratio. Purified amplicon library was 454-pyrosequenced (GS FLX Titanium, Roche Applied Science, Indianapolis, IN) at the Integrated Genomics Facility at Kansas State University (Manhattan, KS).

Data Analysis-Bioinformatics and Statistics

The fungal and bacterial sequence data were analyzed using mothur (v. 1.27.0; Schloss et al., 2009) following the standard operating procedure outlined by Schloss et al., (2011), with some modifications made for the fungal sequence data. The online SOP (www.mothur.org/wiki/Schloss_SOP was accessed March 4, 2013) was used as a guide during fungal and bacterial sequence analysis. Both datasets were received as sff files, so low quality reads based on flow data were removed using trim.flows with the removal of reads below 450 flows and allowing one mismatch to the barcode and two mismatches to the primer. The datasets were then de-noised using shhh.flows, and the data further trimmed by culling sequences shorter than 250 bp were culled or with homopolymers longer than 8 bp. Primer and barcodes sequences were removed as outlined in the SOP. For the fungal dataset, unique sequences were identified, chimeras and global singletons were identified and removed and each sample was sub-sampled at 1290 sequences, which removed samples 23PTT4, 48NNT4, 73NTT6 and 82PNT6. From the sub-sampled sequence data set, sequences were pairwise aligned using the default Needleman alignment to create a distance matrix. Fasta-formatted sequences of a representative sequence from each OTU (get.oturep) were BLASTn searched against the NCBI non-redundant database and the top hits (lowest e value) were used to assign identities for taxonomic identification.

For the bacterial dataset unique representative sequences were aligned against the SILVA alignment database (Schloss et al., 2011). To ensure all sequences overlapped in the same alignment space, those outside the desired range were removed (screen.seqs) with the sequences

starting at position 1 and the end optimized using mothur. The alignment was filtered to remove columns containing only gaps using filter.seqs and unique sequences were identified again (unique.seqs) and pre-clustered if different by less than 2 bp (pre.cluster). Chimeric sequences were removed from the dataset (Edgar et al., 2011) and the remaining sequences were classified using naïve Bayesian classifier and the RDP training set (Wang et al. 2007). Chloroplast, mitochondrial, and unclassified sequences were removed. A distance matrix was generated (dist.seqs) to identify OTUs and clustered at 97% similarity. Global singletons were removed and the dataset subsampled to an equal 450 sequences per experimental unit to eliminate bias resulting from unequal sampling effort (Gihring et al., 2011). Subsampling lead to the loss of three experimental units from downstream analyses (bale 34 (PN) stored for 4 months; bales 64 (NN) and 82 (PN) stored for 6 months).

We estimated community coverage, richness, diversity and evenness for both the fungi and bacteria. Adequacy of sampling was determined based on Good's coverage (formula $C = 1 - \frac{n_1}{N}$; where n_1 = the number of OTUs that have been sampled once and N = the total number of individuals in the sample) and construction of rarefaction curves after each incubation period (Rarefaction formula = $S_n = S_t - \frac{\sum_{i=1}^{S_t} \frac{N-N_i}{n}}{\frac{N}{n}}$; where S_n = average number of OTUs observed after drawing n individuals and S_t = total number of OTUs in sample of N total individuals). Richness was estimated by observed number of species (S_{obs}) and extrapolated by Chao1 ($S_{chao1} = S_{obs} + \frac{n_1(n_1-1)}{2(n_2+2)}$; where S_{chao1} = the estimated richness, S_{obs} = the observed number of species, n_1=the number of OTUs with only one sequence (i.e., "singletons"), n_2 = the number of OTUs with only two sequences (i.e., "doubletons") (Magauran, 1988). We estimated diversity using the Complement of Simpson's Index of Diversity (1-D; where $D = \frac{\sum n(n-1)}{N(N-1)}$; where n = the total number of organisms of a particular species, N = the total number of organisms of all species) was used to determine diversity, with a higher diversity indicated by a higher value. Evenness (Simpson's Equitability – E_D) ($E_D = \frac{\frac{1}{D}}{S}$; where S is the OTU richness in each sample and D is Simpson's Diversity Index shown above) was also determined based on Simpson's Diversity Index. Compositional differences among the communities were visualized using non-metric multidimensional scaling (NMDS) based on the Bray-Curtis dissimilarity matrix calculated from after subsampling. OTUs correlated with each treatment group were also

determined based on Spearman's Rank Correlation Coefficient. Community differences were tested using Analysis of MOlecular VAriance (AMOVA; Anderson 2001) and homogeneity of the populations present in each treatment and incubation time was also evaluated using HOmogeneity of MOlecular VAriance (HOMOVA) in mothur.

Good's coverage, $S_{obs,}$ and Simpson's diversity data were analyzed using two-way Analysis Of Variance (ANOVA) with a model that included storage treatment, duration of the incubation and their interaction. The temporal dynamics over the 6-month incubation period were further evaluated using multiple linear regression. All analyses were performed using JMP 7.0.2 (SAS Institute Inc., Cary, NC). To evaluate taxon level differences among the treatments, the 100 most abundant OTUs were analyzed using two-way ANOVA and multiple regression. These analyses were False Discovery Rate (FDR; Q = 0.05) corrected to account for multiple comparisons.

Results

Fungal Community Dynamics

Of the 316,828 reads obtained, 119,239 reads remained after quality control, removal of chimeras, trimming to 250 bp, and removal of 356 singletons. The data included 330 OTUs at 97% similarity, 227 of which (66.0% of the sequences) were classified as ascomycotes and 103 basidiomycetes (34.0% of the sequences). The data were strongly dominated by a few common OTUs: 20 most abundant OTUs represented 87.9% of the retained sequences (104,791 of the 119,239). The OTU, number of sequences and taxonomic information of the top 20 OTUs are shown in Table 5.1. The five most abundant OTUs, were assigned to genera *Cryptococcus, Cladosporium, Alternaria, Fusarium* and *Hannella*, which represented 51.8% of the sequences.

Good's coverage (98.1% ± 0.004) indicated that our sampling captured the resident diversity well. The coverage estimators differed neither between the treatments ($F_{3,88}$ = 1.1947; P = 0.3168) or the incubation length ($F_{3,88}$ = 1.3224; P = 0.2534) nor were there any treatment by time interactions ($F_{3,88}$ = 0.1689; P = 0.9191). In contrast to the coverage estimators, rarefaction curves (Figure 5.1) suggested that sampling a greater number of sequences (> 1290 sequences) would have resulted in greater richness, which is typical of studies using environmental DNA. Fungal OTU richness (S_{obs}, Figure 5.2) did not differ between the treatments ($F_{3,88}$ = 0.1321; P = 0.9407), over time ($F_{3,88}$ = 3.3930; P = 0.0690) or treatment by time interaction ($F_{3,88}$ = 0.2973; P

= 0.8273). However, in the multiple linear regression analyses, all treatments had a positive slope, indicating an increase in richness during the six month storage. Similarly to the observed richness, the extrapolative Chao1 richness estimates neither differed between treatments or over time nor was there a treatment by time interaction (ANOVA: $F_{7,84} = 0.4515$, $P = 0.8665$; Regression analysis: treatment: ($F_{3,88} = 0.2751$; $P = 0.8432$), time: ($F_{3,88} = 1.7231$; $P = 0.1929$) and treatment by time: ($F_{3,88} = 0.1751$; $P = 0.9130$). Although there were no significant main effect differences (treatment [$F_{3,88} = 2.2035$; $P = 0.0937$] and time [$F_{3,88} = 2.1422$; $P = 0.1470$]) in diversity (1-D; Figure 5.3), there was a significant treatment by time interaction ($F_{3,88} = 3.5980$, $P = 0.0168$). The multiple linear regression analyses indicated that the fungal diversity decreased in treatment NN ($P = 0.0036$) and increased in treatment NT over time ($P = 0.0390$; Table 5.2). Evenness (E_D) did not to differ between treatments ($F_{3,88} = 1.0155$; $P = 0.3901$), over time ($F_{3,88} = 0.0017$; $P = 0.9673$) but there was a significant treatment by time interaction ($F_{3,88} = 2.8750$; $P = 0.0410$; Figure 5.4). Multiple linear regression analyses indicated that evenness decreased in treatment NN over time ($P = 0.0057$), whereas no similar trend was visible in the covered treatments. The decrease in evenness in NN suggests a few taxa becoming dominant over the storage period.

The community composition data were optimally resolved in three dimensions (k = 3; stress = 0.178; $R^2 = 85.3\%$). Two first axes captured the variability in our data well: Axis 1 represented 41.2% and Axis 2 33.9% of the variance, 75.1% in total. The NMDS ordination allowed for a visual representation of the fungal communities of each storage treatment for the 6-month storage period. Based on our AMOVA, fungal communities were distinct among the storage treatments (Table 5.3; Figure 5.5): the covered treatments (NT, PN and PT) clustered together and away from the uncovered treatment (NN). Although not as dramatic, storage duration also had an impact on the fungal community composition (Figure 5.6; Table 5.3). The early fungal communities in the beginning of the experiment and after two-month storage (T0 and T2) differed from the communities at the end of the experiment (T6). The communities after 4-month storage (T4) neither differed from the early (T0 and T2) nor the final (T6) communities. The lack of differences found between the middle (T4) and late (T6) communities would indicate the changes happening took time with an increase towards the end of storage (T6). The slow shifts in the fungal community over time are likely attributable to the environmental conditions, as T0, T2, T4 and T6 occurred in October, December, February and April, respectively. Low air

temperatures in December and February likely inhibited fungal growth, while warmer temperatures and relatively high air humidity from February to April facilitated fungal growth in our sorghum biomass. HOMOVA indicated no differences in community heterogeneity in the different storage treatments or durations.

The greatest community differences were between the uncovered and covered storage treatments. We examined Spearman's Rank Correlation Coefficients to pinpoint OTUs correlated with the uncovered and covered storage treatments. The OTUs correlated to the covered and uncovered treatments, along with Family ecology are shown in Tables 5.4 and 5.5, respectively. Those OTUs correlated to uncovered treatments had an associated family ecology of being saprobic on plant materials, wood substrata and other fungi. Covered treatments were found to be colonized with ubiquitous soil community members and multiple species from the genera *Cladosporium*.

Analyses of the 100 most abundant OTUs (FDR; q = 0.05) identified 36 OTUs with a significant treatment effect (Table 5.6), 20 with a significant storage time effect (Table 5.7), and 12 a significant treatment by time interaction effect (Table 5.8). Twenty of the 36 OTUs with a significant treatment effect, identified differences between uncovered and covered treatments (NT, PN and PT), underlined in Table 5.6. Of those 20 OTUs, 12 were enriched in the uncovered treatment; only 8 had a lower abundance in the uncovered treatment. The OTUs with a higher abundance in the uncovered treatment were identified as *Heydenia* (OTU052; family: Pyronemataceae), *Gibellulopsis* (OTU006; family: Plectosphaerellaceae), *Plectosphaerella* (OTU022; family: Plectosphaerellaceae), *Acremonium* (OTU081; family: Bionectraceae), *Acremonium* (OTUs 024, 085 and 047; family: Hypocreaceae), while *Sarocladium* (OTU069) had an uncertain Family classification in the Hypocreales order and *Sphaeronaemella* (OTU035) had an uncertain Family classification in the Microascales order. OTUs 053, 051 and 072 belonged to Hypocreales but were unclassified at the Family and Genus levels. All of these are members of the Class Sordariomycetes with the exception of OTU052, which is of the Pezizomycetes class. Twenty OTUs had a significant time effect, with 13 OTUs increasing over time, underlined in Table 5.7, while the remaining seven decreased. The OTUs that increased over time were identified as *Cladosporium* (OTUs 030, 002, 077; family: Davidiellaceae), *Eurotium* (OTU012; family: Trichocomaceae), *Gibellulopsis* (OTU006; family: Plectosphaerellaceae), *Plectosphaerella* (OTU022; family: Plectosphaerellaceae), *Acremonium*

(OTU081; family: Bionectraceae), *Acremonium* (OTUs 024 and 047; family: Hypocreaceae), *Wallemia* (OTU021; family: Wallemiaceae), and *Scarocladium* (OTU069; family: incertae sedis), *Valsaria* (OTU079; family incertae sedis), and OTU053 had uncertain classification to Family but a member of the Hypocreales order. Those OTUs found to decrease over time represented multiple families, including Pleosporaceae, Nectriaceae, Leptosphaeriaceae, Didymellaceae and Ustillaginaceae. Overall, the majority of the communities responsive to treatment and time were from Sordariomycetes. Interestingly among the 12 OTUs with a significant treatment by time interaction, eight (OTUs 033, 024, 053, 081, 022, 047, 069, and 052) increased in the uncovered treatment (NN) over time and remained stable in the covered treatments, shown in Table 5.8. Of those 8 OTUs, 4 (OTUs 024, 081, 047, and 069) were identified as *Acremonium*. *Acremonium* (OTU024) increased the most over time in the uncovered treatment. *Valsaria* (OTU079) and *Cryptococcus* (OTU068) were found to decrease in NN and NT over time, while increasing in PN and PT overtime. Fusarium (OTU023) was found to increase in PN over time and decrease in all other treatments (NN, NT, and PT) over time. Interestingly, *Wallemia* (OTU021) was found to increase the most in the covered treatment NT, while decreasing in the other covered treatments (PN and PT) and uncovered (NN).

Bacterial Community Dynamics

Of the 667,953 raw reads, 190,282 reads remained after quality control, removal of putative chimeras, clustering and alignment to the silva reference, with the removal of sequences classified as mitochondria, chloroplast, Archaea, Eukarya or unknown. A total of 89,755 sequence reads remained. Clustering based on 97% similarity resulted in 2,212 OTUs, with 1,797 singletons being removed. The 20 most abundant OTUs (Table 5.10) represented 63,963 sequences (71.9% of the total sequences). Family Enterobacteriaceae dominated and represented seven of the 20 most abundant OTUs. The remaining families among the 20 most abundant OTUs included Pseudomonadaceae, Sanguibacteriaceae, Nocardiopsaceae, Xanthomonadaceae, Caulobacteraceae, Staphlyococcaceae and Dermabacteraceae. Four of the 20 most abundant OTUs identified at the genus-level to be *Pseudomonas*.

Good's coverage (86.3% ± 0.030) indicated that our study captured the resident diversity well in our experiment, did not differ between the treatments ($F_{3,89} = 0.4779$, $P = 0.6986$), time ($F_{3,89} = 0.1232$, $P = 0.7264$), and had no significant interaction ($F_{3,89} = 2.3063$, $P = 0.0824$). In

contrast to coverage, rarefaction analyses (Figure 5.7) suggested that our sampling was not sufficient, especially with sub-sampling at 450 sequences. Observed richness (S_{obs}; Figure 5.8), extrapolated richness (Chao1; Figure 5.9), and diversity (1-D; Figure 5.10) did not differ between treatments, over storage time, and had no significant treatment by time interaction (Table 5.11). Evenness (E_D; Figure 5.11) did not differ between treatments or over storage time. However, the treatment by time interaction was significant (Table 5.11). Further evaluation of the interaction term revealed the uncovered treatment (NN) increased in evenness over time while the covered treatment (NT) decreased in evenness over time, when compared to treatment PT, shown in Table 5.12. This increase in evenness over storage time in the uncovered treatment indicates that the overall community structure does not become increasingly dominated by a few taxa, which is in dramatic contrast to fungal evenness in the uncovered treatment.

Bacterial community data were optimally resolved in three dimensions (k = 3; stress = 0.174; R^2 = 89.3%). However, the two-dimensional NMDS solution captured the variability in our data set well: Axis 1 represented 57.9% and Axis 2 26.6% of the variance, 84.5% in total. In direct contrast with fungal NMDS by treatment, there was no distinction between fungal communities found in the four treatments (Figure 5.12). This was also the case when evaluating NMDS by time, with no distinction between sampling times in the bacterial communities (Figure 5.13). Our AMOVA indicated no distinction between the bacterial communities in the different storage treatments or storage durations (Table 5.11), in contrast to the fungal communities. The community heterogeneity remained stable as indicated by the non-significant HOMOVA among the communities in the different storage treatments or over storage time (Table 5.11). The lack of distinction between covered and uncovered treatments or over storage period suggests bacterial community dynamics are minimally influenced by storage conditions or duration. Since no distinctions were seen between treatments, Spearman's Rank Correlation Coefficient was not used to determine OTUs associated with each treatment or time, as was done with the fungal OTUs.

The stability and lack of difference in the bacterial community dynamics in different treatments was further highlighted by our OTU level analyses of the 100 most abundant OTUs. None of the 100 most common OTUs responded significantly to storage treatment, storage time or had a significant treatment by time interaction after correction for multiple comparisons as seen with the fungal community (FDR; q = 0.05).

Discussion

We analyzed fungal and bacterial community dynamics in sorghum biomass stored under various storage conditions and durations. Our previous studies indicated higher cellulose- and hemicellulose-modifying enzyme activity congruent with reduced cellulose content and reduced conversion to ethanol yields in sorghum biomass stored uncovered over a six-month period (Rigdon et al. 2013). In contrast to uncovered biomass, biomass stored covered over the six-month period maintained cellulose content and ethanol yields as well as lower enzyme activities (Rigdon et al., 2013). The resulting degradation of sorghum biomass during storage motivated the current study to evaluate concomitant shifts in the fungal and bacterial communities. Our results revealed an abundance of saprobic fungal community members present in uncovered biomass, likely causing the degradation of biomass as seen previously. Our study represents an extensive and unique evaluation of the fungal and bacterial community dynamics in stored biomass using high-throughput pyrosequencing. Our results highlight the impact of management, specifically biomass coverage method, during storage has on fungal community dynamics and to a lesser extent the bacterial community dynamics. Most notably in our results was the responsiveness of the fungal communities to coverage-no coverage, which was not seen in the bacterial communities. More importantly, the results presented here, based on ITS-targeted sequencing, were congruent with LSU-targeted sequencing previously reported by Rigdon et al. (unpublished). Furthermore, the abundance of saprobic fungal community members present in uncovered biomass during storage, with the capability of biomass degradation (i.e., cellulose breakdown) via enzymatic pathways; result in reduced biomass quality for lignocellulosic ethanol production

Ascomycota were found to dominate over Basidiomycota 2:1 (227:103 OTUs) in the sorghum biomass in all storage treatments throughout the entire 6-month storage period (data not shown). These results agree with Poll et al. (2010): Ascomycota dominated rye residues, possibly because their ability to degrade substrates like cellulose. In contrast, terminal, late successional basidiomycetous communities have been suggested to establish in the latter stages of litter decomposition in forest ecosystems, as they are better suited to degrade lignin (Frankland, 1998). The low abundance of basidiomycetes is likely due to the short storage/incubation time driven by our focus on a realistic harvest and storage cycle for biomass destined for lignocellulosic ethanol production. In addition, the sorghum biomass used in this study had a cellulose and

hemicellulose content of roughly 35% and 25% dry weight, respectively; while lignin comprised only 3-4% (dry weight). The high abundance of cellulose and hemicellulose in combination with the low lignin content in the substrate is likely conducive to the maintenance of ascomycete-dominated community.

The most commonly observed fungal genera in our stored biomass (*Cryptococcus, Cladosporium, Alternaria* and *Fusarium*) include common air- and soil-borne fungi as well as potential plant pathogens (*Alternaria* and *Fusarium*). Genus *Cladosporium* is one of the most common of airborne genera, but includes also a number of saprobic, phytopathogenic and endophytic species. Kleczewski et al. (2012) identified *Cladosporium* as a common endophyte in switchgrass (*Panicum virgatum*), a C_4 perennial grass native to North America like sorghum. Kleczewski et al. (2012) reported that *Cladosporium* colonized plant biomass prior to plant senescence or death and continued to dominant plant biomass decomposition. The endophytic-to-saprobic nature of *Cladosporium* was further validated in our system by its increase over time, especially in the covered treatments, as evidenced by OTUs 030, 002 and 077.

Neither bacterial or fungal communities showed a strong change in richness over time or differed across treatments; however, a positive, yet not significant, trajectory was evident, similar to our former LSU study (Rigdon et al., unpublished). Consistency in richness responses (S_{obs} and Chao1) among treatments was not evident in the bacterial communities, indicating a lack of response to coverage, contrary to our hypothesis. We had hypothesized richness in both the fungal and bacterial communities to increase over time in the uncovered treatment, as the lack of coverage would allow the continued deposition of allochthonous propagules.

While fungal diversity treatment and time main effects were not significant, the treatment by time interaction was. Further evaluation of the interaction effects using multiple linear regression indicated a decrease in uncovered (NN) over time, while covered treatments increased over time. In contrast, LSU analysis indicated an increase in diversity in all treatments over time (Rigdon et al., unpublished). In alignment with diversity, the evenness (E_D) of the fungal community also decreased in the uncovered treatment over time. We detected no bacterial OTUs that strongly responded to treatments or shifted in abundance over time. In contrast, many fungal OTUs significantly changed in abundance and differed among treatments.

Our NMDS ordination analyses of the fungal communities combined with AMOVA distinguished the covered (NT, PN and PT) and uncovered treatments (NN) as well as the

communities in the early and late storage. Visual differences seen in ordination were confirmed through AMOVA, further evidence that the fungal community dynamics were greatly influenced by the management method used for biomass storage prior to lignocellulosic ethanol production. Ordination results by treatment and by time in the fungal community were consistent with those reported using LSU-based sequencing (Rigdon et al., unpublished). Unlike the fungal community, no distinction between storage treatments or durations could be seen in the bacterial communities. In addition to the other community metrics (richness, diversity and evenness) observed, the bacterial community dynamics were not influenced by storage method used for lignocellulosic ethanol production.

Many fungal OTUs were found to respond to treatment conditions and storage durations, while bacterial OTUs did not. Consistent with our earlier sequencing of the LSU, the genus *Wallemia* increased most in abundance in the covered treatment NT, while decreasing in all other treatments (based on ANOVA with FDR of 0.05). Though not directly measured, coverage of bales with a tarp prevented the addition of moisture from precipitation events (rain, snow) but still allowed for the evaporation of moisture from fungal respiration. This allowed the bales to maintain low moisture content (Rigdon et al., 2013), likely creating an environment conducive to *Wallemia* – a xerophilic/xerotolerant genus.

While not monitored directly in this study, bales wrapped in plastic or bale stretch wrap (both used in this study), tend to have a higher carbon dioxide concentration during storage, as a result of lack of diffusion of gases through the plastic/wrap and creating a micro-aerobic environment (Muller, 2005). The lack of gas diffusion would also indicate moisture evaporation of the bale would be stopped by the plastic, thus condensing on the interior surface of the plastic and be available for metabolic processes. In addition, treatments PN and PT were found to have increased abundance of *Valsaria, Alternaria, Cryptococcus* and *Bulleromyces*. The genera *Alternaria* are ubiquitous in the environment and have been shown to continue to grow (hyphal elongation) and sporulate in environments low in oxygen or with slightly elevated levels of carbon dioxide (Lukens and Horsfall, 1972). Cryptococcus and Bulleromyces are both yeasts that have been isolated from various plant sources and contain species that have been shown to have antagonistic effects for natural biocontrol against plant pathogens (Roberts 1990; Rodrigues et al., 2009)

The bacterial communities were diverse and had more than 2,000 OTUs, nearly 10 times more diverse than the fungal communities, with just over 200 OTUs. The five most abundant OTUs from the bacterial community were identified as *Pantoea, Pseudomonas, Enterobacter*, unclassified from the Enterobacteraceae and *Sanguibacter*. Multiple species of *Pantoea* have been associated with disease in many economically important agricultural crops and forest tree species worldwide (Coutinho and Venter, 2009). Many of the most common OTUs were assigned to taxa that associate with plant tissues. These are pathogens such as *Pantoea* and *Pseudomonas* and putative endophytes, exemplified by *Sanguibacter*. Specifically, *P. ananatis* has been identified to cause leaf spot symptoms in Sudan grass (*Sorghum sudanense*) and sorghum (*Sorghum bicolor*) (Cota et al., 2010). Many species of *Pseudomonas* have been identified as plant pathogens; however, strain AKM-P6 was found to have growth-promoting properties on the survival and growth of sorghum seedlings at elevated temperatures (Ali et al., 2009). A few species of *Enterobacter* have been identified with nitrogen fixation and associated with the roots of maize, wheat and sorghum (Pederson et al., 1978). Species of *Sanguibacter* have been found to grow endophytically on tobacco (*Nicotiana tabacum*) and have been shown to increase the phytoextraction of cadmium to the plant, improving the overall phytoremediation of metalliferous soils (Mastretta et al., 2009). The collective endophytic and saprobic capabilities of these various genera make it not surprising that they were abundantly found on sorghum biomass stored over a 6-month period. However, the lack of response to storage treatment in the bacterial community further indicates the system is driven by the fungal communities.

Our study highlights the different dynamics of fungal and bacterial communities in lignocellulosic biomass stored in short- and mid-term. While the fungal richness, diversity and evenness clearly indicate that fungal community dynamics are influenced by management practices during storage, bacterial communities remain stable over time and do not differ among the management treatments. This highlights the importance of understanding especially the fungal community behavior over time in stored biomass. Many of the fungi detected in our study likely possess lignocellulose modifying enzyme systems that compromise biomass integrity and its utility in downstream applications. In contrast, bacterial taxa that commonly occupied the biomass in this study represented putative phytopathogens or antagonists highlighting the legacy of bacterial communities that remained in the biomass through storage. Corroborating our earlier studies the present results further suggest that simple biomass management via coverage inhibits

the arrival and establishment of a fungal community, leading to preservation of substrate for lignocellulosic ethanol production.

References

Ali, Sk.Z., Sandhya, V., Grover, M., Kishore, N., Rao, L.V., Venkateswarlu, B., 2009. Pseudomonas sp strain AKM-P6 enhances tolerance of sorghum seedlings to elevated temperatures. Biology and Fertility of Soils 46, 45-55.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral. Ecol. 26, 32-46.

Berry, D., Mahfoudh, K.B., Wagner, M., Loy, A., 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl. Environ. Microbiol. 77, 7846-7849.

Cota, L.V., Costa, R.V., Silva, D.D., Parreira, D.F., Lana, U.G.P., Casela, C.R., 2010. First report of pathogenicity of Pantoea ananatis in sorghum (Sorghum bicolor) in Brazil. Australasian Plant Disease Notes 5, 120-122.

Coutinho, T.A., Venter, S.N., 2009. Pantoea ananatis: an unconventional plant pathogen. Molecular Plant Pathology 10, 325-335.

Edgar R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011 UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194-2200.

Frankland, J.C., 1998. Fungal succession—unraveling the unpredictable. Mycol. Res. 102, 1-15.

Fierer, N., Hamady, M., Lauber, C.L., Knight, R., 2008. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. PNAS. 105, 17994-17999.

Gao, J., Qian, L., Thelen, K.D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103, 844-850.

Gardes, M., Bruns, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes – application to the identification of mycorrhizae and rusts. Mol. Ecol. 2, 113-118.

Gihring, T.M., Zhang, G., Brandt, C.C., Brooks, S.C., Campbell, J.H., Carroll, S., Criddle, C.S., Green, S.J., Jardine, P., Kostka, J.E., Lowe, K., Mehlhorn, T.L., Overholt, W., Watson, D.B., Yang, Z., Wu, W., Schadt, C.W., 2011. A limited consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. Appl. Environ. Microb. 77, 5955-5965.

Hess J.R., Wright, C.T., Kenney, K.L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel. Bioprod. Bior. 1, 181-190.

Jumpponen, A., Jones, K.L., 2010. Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. New Phytol. 186, 496-513.

Kleczewski, N.M., Bauer, J.T., Bever, J.D., Clay, K., Reynolds, H.L., 2012. A survey of endophytic fungi in switchgrass (*Panicum virgatum*) in the Midwest, and their putative roles in plant growth. Fungal Ecol. 5, 521-529.

Lukens, R.J., Horsfall, J.G., 1972. Processes of sporulation in alternaria solani and their response to metabolic inhibitors. Phytopathology 63, 176-182.

Magauran, A.E., 1988. *Ecological Diversity and its Measurement*. Princeton University, Princeton.

Mastretta, C., Taghavi, S., van der Lelie, D., Mengoni, A., Galardi, F., Gonnelli, C., Barac, T., Boulet, J., Weyens, N., Vangronsveld, J., 2009. Endophytic bacteria from seeds of Nicotiana tabacum can reduce cadmium phytotoxicity. International Journal of Phytoremediation 11, 251-267.

Muller, C.E., 2005. Fermentation patterns of small-bale silage and haylage produced as a feed for horses. Grass Forage Sci. 60, 109-118.

Pederson, W.L., Chakrabarty, K., Klucas, R.V., Vidaver, A.K., 1978. Nitrogen fixation (acetylene reduction) associated with roots of winter wheat and sorghum in Nebraska. Applied and Environmental Microbiolgy 35, 129-135.

Poll, C., Brune, T., Begerow, D., Kandeler, E., 2010. Small-scale diversity and succession of fungi in the detritusphere of rye residues. Microb. Ecol. 59, 130-140.

Rigdon, A.R., Jumpponen, A., Vadlani, P.V., Maier, D.E., 2013. Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop. Bioresource Technol. 132, 269-275.

Roberts, R.G., 1990. Postharvest biological control of gray mold of apple by Cryptococcus laurentii. Phytopathology 80, 526-530.

Rodrigues, A., Cable, R.N., Mueller, U.G., Bacci Jr., M., Pagnocca, F.C., 2009. Antagonistic interactions between garden yeasts and microfungal garden pathogens of leaf-cutting ants. Antonie van Leeuwenhoek 96, 331-342.

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:e27310.

Zaldivar, J.J., Nielsen, J.J., Olsson, L.L., 2001. Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. Applied Microbiology and Biotechnology 56, 17-34.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: M.A. Innis, D.H. Glefand, J.J. Sninsky, T.J. White (eds) *PCR Protocols: a Guide to Methods and Applications*, pp. 315-322. Academic Press, New York.

Tables and Figures

Table 5.1 20 most abundant fungal OTUs found, including OTU number, number of sequences represented and taxonomic information.

OTU	# of Seq	Class	Order	Family	Genus
Otu001	14853	Tremellomycetes	Tremellales	Tremellaceae	*Cryptococcus*
Otu002	14577	Dothideomycetes	Capnodiales	Davidiellaceae	*Cladosporium*
Otu003	12945	Dothideomycetes	Pleosporales	Pleosporaceae	*Alternaria*
Otu004	12612	Sordariomycetes	Hypocreales	Nectriaceae	*Fusarium*
Otu005	6800	Tremellomycetes	Tremellales	Tremellaceae	*Hannella*
Otu006	6415	Sordariomycetes	Glomerellales	Plectosphaerellaceae	*Gibellulopsis*
Otu007	4729	Dothideomycetes	Incertae sedis	Incertae sedis	*Epicoccum*
Otu008	4664	Sordariomycetes	Trichosphaeriales	Trichsphaeriaceae	*Nigrospora*
Otu009	4610	Tremellomycetes	Tremellales	Unclassified	Unclassified
Otu010	4399	Sordariomycetes	Hypocreales	Nectriaceae	*Fusarium*
Otu011	3469	Tremellomycetes	Tremellales	Unclassified	*Hannella*
Otu012	2255	Eurotiomycetes	Eurotiales	Trichocomaceae	*Eurotium*
Otu013	2009	Sordariomycetes	Hypocreales	Hypocreaceae	*Acremonium*
Otu014	1798	Tremellomycetes	Tremellales	Tremellaceae	*Bulleromyces*
Otu015	1615	Tremellomycetes	Tremellales	Tremellaceae	*Cryptococcus*
Otu016	1603	Dothideomycetes	Pleosporales	Didymellaceae	*Phoma*
Otu017	1542	Dothideomycetes	Pleosporales	Leptosphaeriaceae	*Ampelomyces*
Otu018	1457	Tremellomycetes	Tremellales	Unclassified	Unclassified
Otu019	1418	Tremellomycetes	Tremellales	Unclassified	*Hannella*
Otu020	1021	Dothideomycetes	Pleosporales	Didymellaceae	*Didymella*

*OTUs correlated to uncovered treatment (NN-no coverage); OTUs correlated to covered treatments (NT-no plastic/trap, PN-plastic/no tarp, PT-plastic/tarp)

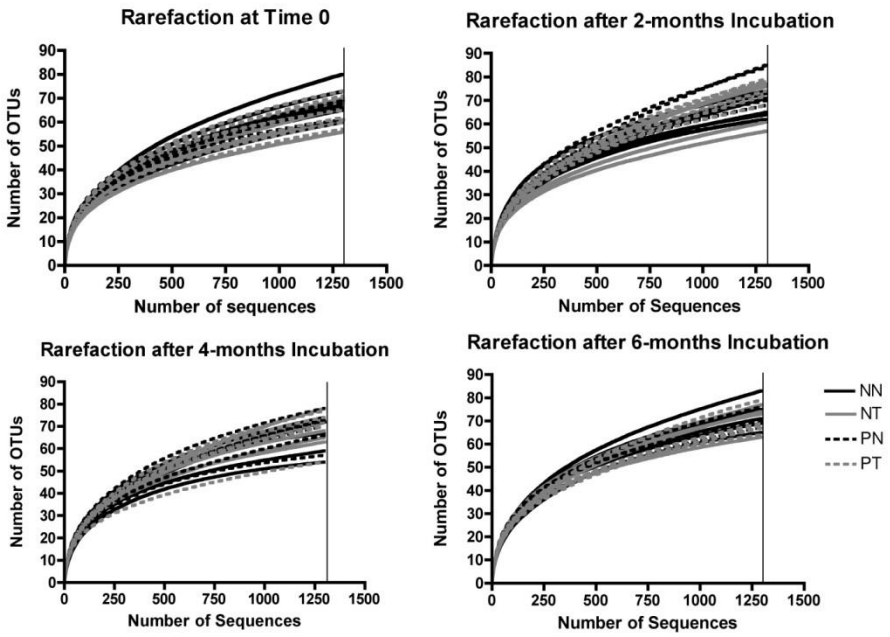


Figure 5.1 Fungal community rarefaction curves after 0, 2, 4 and 6-month incubation periods with subsampling cut off of 1290 sequences indicated by the vertical line. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

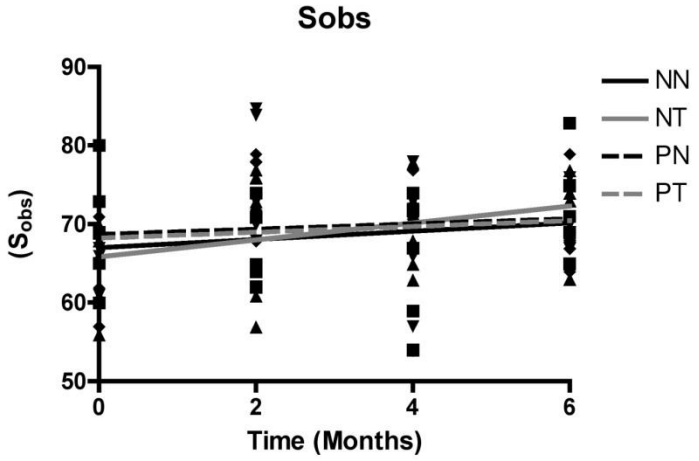


Figure 5.2 Fungal community richness across treatments and time based on observed species (S_{obs}) with differences between treatments, over time and treatment by time interactions. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

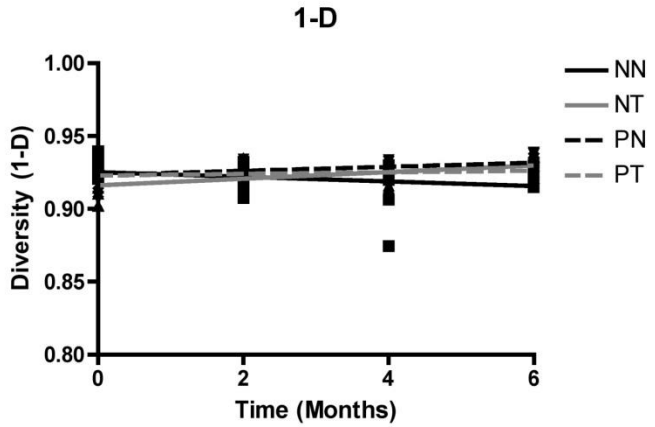


Figure 5.3 Fungal community diversity across treatments and time based on the complement of Simpson's Diversity Index (1-D) with no significant differences found. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

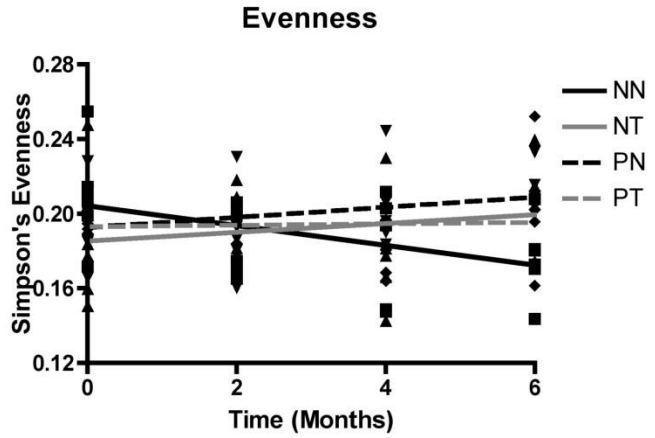


Figure 5.4 Fungal community evenness across treatment and time based on Simpson's Index, with uncovered treatment decreasing after six months of storage. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Table 5.2 Fungal community Multiple Linear Regression Analyses. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Response	Estimate	Estimate STD	t-value	P-value**

Good's Coverage				
Intercept (NN)	0.0012564	0.00078742	1.5955985	0.11433414
Intercept (NT)	-0.0010186	0.00078743	-1.2936334	0.199338
Intercept (PN)	0.00028026	0.00078743	0.35592048	0.72279244
Intercept (PT)*	***0.98066948***	***0.00074703***	***1312.7511***	***6.76E-183***
Slope (NN)	0.00020282	0.00034835	0.58223326	0.56196909
Slope (NT)	5.99E-05	0.00035689	0.16773298	0.86719652
Slope (PN)	-0.0001054	0.00035689	-0.2952798	0.76850877
Slope (PT)*	0.00023413	0.0002036	1.14995309	0.25342663
S_{obs}				
Intercept (NN)	-0.6193325	1.21080796	-0.5115035	0.61034055
Intercept (NT)	-0.1505069	1.21081909	-0.1243017	0.90137335
Intercept (PN)	0.57487056	1.21081909	0.47477824	0.63617712
Intercept (PT)*	***67.4282087***	***1.14870775***	***58.6991851***	***5.74E-70***
Slope (NN)	-0.0545292	0.5356475	-0.1018006	0.91915759
Slope (NT)	0.50415369	0.54879418	0.91865715	0.36090533
Slope (PN)	-0.2451647	0.54879418	-0.4467334	0.6562172
Slope (PT)*	0.576682	0.31307435	1.84199694	0.06900443
1-D				
Intercept (NN)	-0.0032569	0.00169655	-1.9197415	0.0582851
Intercept (NT)	-0.0011147	0.00169657	-0.6570094	0.51297098
Intercept (PN)	***0.0035936***	***0.00169657***	***2.11815936***	***0.03711654***
Intercept (PT)*	***0.92201958***	***0.00160954***	***572.847101***	***1.20E-152***
Slope (NN)	***-0.0022483***	***0.00075054***	***-2.9956117***	***0.003599***
Slope (NT)	***1.61E-03***	***0.00076896***	***2.09719294***	***0.03898299***
Slope (PN)	0.00075265	0.00076896	0.97879825	0.33049029
Slope (PT)*	***0.00064205***	***0.00043867***	***1.46363162***	***1.47E-01***
Evenness				
Intercept (NN)	-0.0051459	0.00425649	-1.2089531	0.23007282
Intercept (NT)	-0.0017459	0.00425653	-0.4101691	0.68272621
Intercept (PN)	0.00663303	0.00425653	1.55831688	0.12291799
Intercept (PT)*	***0.19385707***	***0.00403819***	***48.0059782***	***7.66E-63***
Slope (NN)	***-0.0053423***	***0.00188302***	***-2.8370606***	***0.00570559***
Slope (NT)	0.00233286	0.00192924	1.20921289	0.22997356
Slope (PN)	0.00266016	0.00192924	1.37886453	0.17159729
Slope (PT)*	0.00004524	0.00110059	0.04110483	0.96730987

* = Treatment PT was selected as a reference level to emphasize the contrast between the three covered treatments (PT, PN, NT) and the uncovered treatment (NN)

** = P-values test the null hypotheses (H_0: Intercept $_{PN, NT, or NN}$ − Intercept $_{Ref PT}$ = 0; and H_0: Slope $_{PN, NT, or NN}$ − Slope $_{Ref PT}$ = 0). In other words, significant P-values here indicate that the difference between intercept or slope terms for treatments PN, NT, or NN

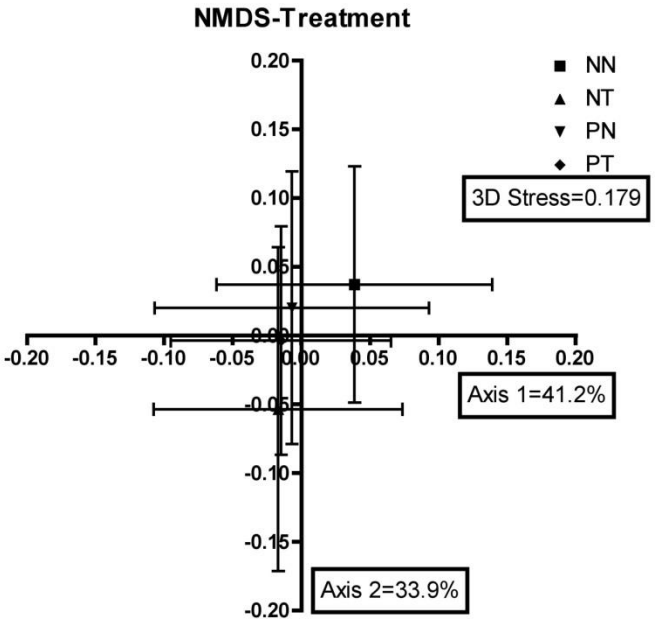


Figure 5.5 Fungal community Non-metric Multi-dimensional Scaling (NMDS) by treatment. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

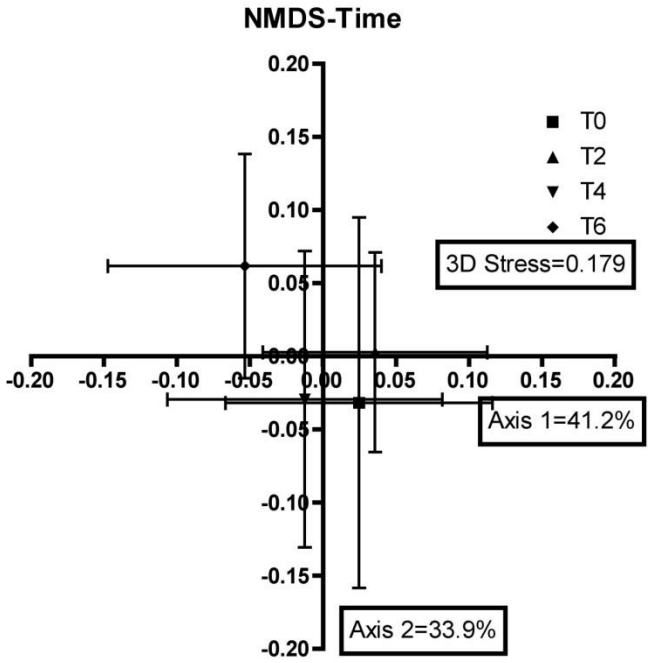


Figure 5.6 Fungal community Non-metric Multi-dimensional Scaling (NMDS) by treatment. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Table 5.3 Fungal community Analysis of Molecular Variance (AMOVA) results for NMDS by treatment and by time. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

AMOVA Results for NMDS by treatment		AMOVA Results for NMDS by time	
NN-NT-PN-PT*	***$F_{3,88}$=6.498, p=<0.001***	***T0-T2-T4-T6****	***$F_{3,88}$=3.056, p=<0.001***
NN-NT*	***$F_{1,44}$=10.732, p=0.001***	T0-T2	$F_{1,46}$=1.821, p=0.074
NN-PN*	***$F_{1,44}$=9.602, p=<0.001***	T0-T4	$F_{1,44}$=2.407, p=0.017
NN-PT*	***$F_{1,44}$=10.387, p=<0.001***	***T0-T6****	***$F_{1,44}$=6.901, p=<0.001***
NT-PN	$F_{1,44}$=2.196, p=0.027	T2-T4	$F_{1,44}$=1.164, p=0.277
NT-PT	$F_{1,44}$=1.711, p=0.071	***T2-T6****	***$F_{1,44}$=4.011, p=0.001***
PN-PT	$F_{1,43}$=0.633, p=0.831	T4-T6	$F_{1,42}$=2.480, p=0.026

Based on Bonferroni correction for multiple comparison of 0.0083; * indicates significance (p-value<0.0083).

Table 5.4 Fungal community OTUs correlated with the uncovered treatment (NN-no coverage) with family and associated family ecology. Correlation based on significantly positive axis scores for axis 1 with positive axis 2 scores.

OTU	Axis 1 Score	Axis 1 P-value	Axis 2 Score	Axis 2 P-value	Family	Genus	Associated Family Ecology
Otu011	0.720	0.0	0.272	0.008	Tremellaceae	*Hannella*	Usually grow on woody substrata, often parasitic on other fungi
Otu006	0.316	0.002	0.619	0.0	Plectosphaerellaceae	*Gibellulopsis*	Saprobic on plant material
Otu043	0.298	0.003	0.368	0.0003	Botryosphaeriaceae	*Macrophoma*	
Otu041	0.249	0.016	0.575	0.0	Hypocreaceae	*Acremonium*	Common in all environmental zones; saprobic on rotting wood and other vegetation
Otu015	0.249	0.016	0.319	0.001	Tremellaceae	*Cryptococcus*	Usually grow on woody substrata, often parasitic on other fungi
Otu004	0.207	0.046	0.315	0.002	Nectriaceae	*Fusarium*	Associated with dead plant material

Table 5.5 Fungal community OTUs correlated with covered treatments (NT-tarp, PN-plastic and PT-plastic and tarp) with family and associated family ecology. Correlation based on significantly negative axis scores for axis 1 with negative axis 2 scores.

OTU	Axis 1 Score	Axis 1 P-value	Axis 2 Score	Axis 2 P-value	Family	Genus	Associated Family Ecology
Otu002	-0.432	0.0	-0.128	0.219	Davidiellaceae	*Cladosporium*	Saprobic, commonly found outdoors
Otu010	-0.398	0.0	-0.283	0.006	Nectriaceae	*Fusarium*	Associated with dead plant material
Otu173	-0.289	0.005	-0.221	0.033	Trichocomaceae	*Aspergillus*	Saprobic, ubiquitious in soil, commonly associated with decaying plant material
Otu232	-0.236	0.023	-0.008	0.939	Davidiellaceae	*Cladosporium*	Saprobic, commonly found outdoors
Otu038	-0.212	0.041	-0.048	0.642	Davidiellaceae	*Cladosporium*	Saprobic, commonly found outdoors
Otu270	-0.210	0.044	-0.019	0.851	Tremellaceae	*Cryptococcus*	Usually grow on woody substrata, often parasitic on other fungi

Table 5.6 Fungal community OTUs with significant treatment responses, including F-statistic, treatment response and genus, underlined indicate uncovered treatment (NN) different from all covered treatments (NT, PN and PT).

OTU	F-Statistic	Treatment Response	Genus
Otu006	$F_{3,86}$=<0.0001	NN different from PN, PT, NT; NN highest mean	*Gibellulopsis* OR *Verticillium*
Otu022	$F_{3,86}$=<0.0001	NN different from PT, PN, NT; NN highest mean	*Plectosphaerella*
Otu053	$F_{3,86}$=<0.0001	NN different from NT, PN, PT; NN highest mean	Unclassified
Otu051	$F_{3,86}$=<0.0001	NN different from PN, NT, PT; NN highest mean	Unclassified
Otu072	$F_{3,86}$=<0.0001	NN different from PT, PN, NT; NN highest mean	Unclassified
Otu003	$F_{3,86}$=<0.0001	NN different from NT, PT, PN; NN lowest mean	*Alternaria*
Otu035	$F_{3,86}$=<0.0001	NN different from PN, NT, PT; NN highest mean	*Sphaeronaemella*
Otu024	$F_{3,86}$=<0.0001	NN different from PN, PT, NT; NN highest mean	*Acremonium*
Otu005	$F_{3,86}$=<0.0001	NN different from PT, PN, NT; NN lowest mean	*Hannaella* OR *Bullera*
Otu069	$F_{3,86}$=<0.0001	NN different from NT, PT, PN; NN highest mean	*Sarocladium* OR *Acremonium*
Otu021	$F_{3,86}$=<0.0001	NT different from PN, NN, PT; NT highest mean	*Wallemia*
Otu027	$F_{3,86}$=<0.0001	NN different from PN, PT, NT; NN lowest mean	*Sporobolomyces*
Otu017	$F_{3,86}$=<0.0001	NN different PN, NT; NN lowest mean	*Ampelomyces*
Otu044	$F_{3,86}$=<0.0001	NN different from PN, PT, NT; NN lowest mean	Unclassified
Otu085	$F_{3,86}$=<0.0001	NN different from NT, PT, PN; NN highest mean	*Acremonium*
Otu007	$F_{3,86}$=<0.0001	NN different from PT, PN, NT; NN lowest mean	*Epicoccum*
Otu047	$F_{3,86}$=<0.0001	NN different from NT, PT, PN; NN highest mean	*Acremonium*
Otu028	$F_{3,86}$=0.0003	NN different from PN, PT, NT; NN lowest mean	*Leptosphaerulina*
Otu031	$F_{3,86}$=0.0005	NN different from PN, PT, NT; NN lowest mean	Unclassified
Otu052	$F_{3,86}$=0.0006	NN different from NT, PT, PN; NN highest mean	*Heydenia*

Otu081	$F_{3,86}$=0.0010	NN different from PN, NT, PT; NN highest mean	*Gliomastix* OR *Acremonium* OR *Periconia*
Otu079	$F_{3,86}$=0.0014	PT different from NN, NT; PT highest mean	*Valsaria*
Otu070	$F_{3,86}$=0.0022	NN different from NT, PT; NN highest mean	*Clavispora*
Otu025	$F_{3,86}$=0.0027	NN different from PN, PT; NN lowest mean	Unclassified
Otu050	$F_{3,86}$=0.0029	NN different from PN, PT; NN lowest mean	*Alternaria*
Otu039	$F_{3,86}$=0.0032	NN different from PN, PT, NT; NN lowest mean	*Sporidiobolus*
Otu026	$F_{3,86}$=0.0043	NN different from PT; NN highest mean	*Metschnikowia*
Otu036	$F_{3,86}$=0.0057	PN different from NT, NN; PN highest mean	*Alternaria*
Otu014	$F_{3,86}$=0.0073	NN different from NT, PN; NN lowest mean	*Bulleromyces*
Otu098	$F_{3,86}$=0.0103	NT different from PT, NN, NT highest mean	*Sporobolomyces*
Otu033	$F_{3,86}$=0.0115	NN different from PT, NN highest mean	*Cryptococcus*
Otu018	$F_{3,86}$=0.0119	NN different from NT, PT; NN lowest mean	Unclassified
Otu048	$F_{3,86}$=0.0131	NN different from PN, PT; NN highest mean	*Hydropisphaera*
Otu056	$F_{3,86}$=0.0139	NN different from PT, NT; NN highest mean	*Guehomyces*
Otu015	$F_{3,86}$=0.0164	PN different from NN; PN highest mean	*Cryptococcus*
Otu001	$F_{3,86}$=0.0165	NT different from PN; NT highest mean	*Cryptococcus* OR *Hannaella*

Table 5.7 Fungal community OTUs with significant time effect, including estimate, standard error, t-ratio, t-test and genus, underlined indicate an increase over time.

OTU	Estimate	St. Error	t-ratio	t-test	Genus
Otu030	1.224705	0.16914	7.25	<0.0001	*Cladosporium*
Otu012	7.416177	1.287207	5.76	<0.0001	*Eurotium*
Otu017	-1.74807	0.310304	-5.63	<0.0001	*Ampelomyces*
Otu007	-3.90841	0.70973	-5.51	<0.0001	*Epicoccum*
Otu079	0.263348	0.051974	5.07	<0.0001	*Valsaria*
Otu024	2.604759	0.679865	3.83	0.0002	*Acremonium*
Otu022	1.542698	0.438965	3.51	0.0007	*Plectosphaerella*
Otu063	-0.14629	0.041942	-3.49	0.0008	*Pseudozyma*

Otu081	0.203468	0.058496	3.48	0.0008	*Gliomastix* OR *Acremonium* OR *Periconia*
Otu047	0.540308	0.156069	3.46	0.0008	*Acremonium*
Otu006	5.064807	1.467912	3.45	0.0009	*Gibellulopsis* OR *Verticillium*
Otu016	-1.05638	0.311996	-3.39	0.0011	*Phoma*
Otu021	2.550663	0.754669	3.38	0.0011	*Wallemia*
Otu010	-2.58523	0.806516	-3.21	0.0019	*Fusarium*
Otu053	0.338747	0.110607	3.06	0.0029	*Unclassified*
Otu005	-2.41759	0.79502	-3.04	0.0031	*Hannaella* OR *Bullera*
Otu069	0.238428	0.083526	2.85	0.0054	*Sarocladium* OR *Acremonium*
Otu002	3.818962	1.338898	2.85	0.0055	*Cladosporium*
Otu077	0.10636	0.03789	2.81	0.0062	*Cladosporium*
Otu003	-3.62304	1.302625	-2.78	0.0067	*Alternaria*

Table 5.8 Fungal community OTUs with significant treatment by time interactions effects, including F-Statistics, estimates, standard errors, t-ratio, t-test and genus, underline indicates an increase in uncovered treatment (NN).

OTU	F-Statistic	Estimate	St. Error	t-ratio	t-test	Genus
Otu033	$F_{3,86}$= <0.0001	NN-time=1.057097 NT-time=-0.37918 PN-time=-0.25182	0.190516 0.195192 0.195192	5.55 -1.94 -1.29	<0.0001 0.0554 0.2005	*Cryptococcus*
Otu024	$F_{3,86}$= <0.0001	NN-time=6.315593 NT-time=-2.50017 PN-time=-1.54184	1.1632 1.191749 1.191749	5.43 -2.10 -1.29	<0.0001 0.0389 0.1993	*Acremonium*
Otu079	$F_{3,86}$= <0.0001	NN-time=-0.26335 NT-time=-0.26335 PN-time=0.132879	0.088924 0.091107 0.091107	-2.96 -2.89 1.46	0.004 0.0049 0.1484	*Valsaria*
Otu053	$F_{3,86}$= <0.0001	NN-time=1.01944 NT-time=-0.26956 PN-time=-0.37019	0.189241 0.193886 0.193386	5.39 -1.39 -1.91	<0.0001 0.1681 0.0596	*Unclassified*
Otu081	$F_{3,86}$= <0.0001	NN-time=0.497555 NT-time=-0.18303 PN-time=-0.13664	0.100082 0.102538 0.102538	4.97 -1.78 -1.33	<0.0001 0.0779 0.1863	*Gloimastix* OR *Acremonium* OR *Periconia*
Otu021	$F_{3,86}$= 0.0001	NN-time=-2.91689 NT-time=6.33613 PN-time=-1.4351	1.291185 1.322875 1.322875	-2.26 4.79 -1.08	0.0265 <0.0001 0.2811	*Wallemia*
Otu022	$F_{3,86}$= 0.0002	NN-time=3.536981 NT-time=-1.02147 PN-time=-1.16141	0.751038 0.769471 0.769471	4.71 -1.33 -1.51	<0.0001 0.1879 0.1350	*Plectosphaerella*
Otu047	$F_{3,86}$= 0.0003	NN-time=1.213347 NT-time=-0.39959 PN-time=-0.44597	0.267023 0.273576 0.273576	4.54 -1.46 -1.63	<0.0001 0.1479 0.1068	*Acremonium*
Otu069	$F_{3,86}$=	NN-time=0.640958	0.142907	4.49	<0.0001	*Sarocladium* OR

	0.0004	NT-time=-0.18969 PN-time=-0.23843	0.146415 0.146415	-1.30 -1.63	0.1987 0.1075	*Acremonium*
Otu052	$F_{3,86}$= 0.0010	NN-time=1.611397 NT-time=-0.50232 PN-time=-0.55892	0.381342 0.390701 0.390701	4.23 -1.29 -1.43	<0.0001 0.2021 0.1563	*Heydenia*
Otu068	$F_{3,86}$= 0.0018	NN-time=-0.13557 NT-time=-0.2399 PN-time=0.273461	0.084108 0.086173 0.086173	-1.61 -2.78 3.17	0.1108 0.0066 0.0021	*Cryptococcus*
Otu023	$F_{3,86}$= 0.0020	NN-time=-0.95939 NT-time=-0.00088 PN-time=1.220822	0.334756 0.342972 0.342972	-2.87 -0.00 3.56	0.0053 0.9980 0.0006	*Fusarium*

Table 5.9 Fungal community OTUs with significant treatment, time or treatment by time interaction effects, including taxonomic information at family and genera level and associated family ecology with refererence.

OTU	Family	Genus	Family Ecology	Reference	Response
Otu006	Plectosphaerellaceae	*Gibellulopsis* OR *Verticillium*			Treatment effect-NN different from PN, PT, NT; Time effect-increasing over time
Otu022	Plectosphaerellaceae	*Plectosphaerella*			Treatment effect-NN different from PT, PN, NT; Time effect-increasing over time; Interaction-NN increasing over time
Otu053	Unclassified	Unclassified			Treatment effect-NN different from NT, PN, PT; Time effect-increasing over time; Interaction-NN increasing over time
Otu051	Unclassified	Unclassified			Treatment effect-NN different from PN, NT, PT;
Otu072	Unclassified	Unclassified			Treatment effect-NN different from PT, PN, NT;
Otu003	Pleosporaceae	*Alternaria*	cosmopolitan; necrotrophic pathogens and sprobes, especially associated with grasses	Fungal Families of the World	Treatment effect-NN different from NT, PT, PN; Time effect-decreasing over time
Otu035	Unclassified	*Sphaeronaemella*			Treatment effect-NN different from PN, NT, PT;

Otu024	Hypocreaceae	*Acremonium*	cosmopolitan; saprobic on rotting wood and other vegetation or parasitic on other fungi	Fungal Families of the World	Treatment effect-NN different from PN, PT, NT; Time effect-increasing over time; Interaction-NN increasing over time, NT decreasing over time
Otu005	Unclassified	*Hannaella* OR *Bullera*			Treatment effect-NN different from PT, PN, NT; Time effect-decreasing over time
Otu069	Incertae sedis	*Sarocladium* OR *Acremonium*			Treatment effect-NN different from NT, PT, PN; Time effect-increasing over time; Interaction-NN increasing over time
Otu021	Wallemiaceae	*Wallemia*	Widely distributed; saprobic, capable of growth over wide ranges of water tension, spoilage on dried and desiccated products	Fungal Families of the World	Treatment effect-NT different from PN, NN, PT; Time effect-increasing over time; Interaction-NN decreasing over time, NT increasing over time
Otu027	Sporidiobolaceae	*Sporobolomyces*	cosmopolitan; saprobic, found in wide variety of habitats	Fungal Families of the World	Treatment effect-NN different from PN, PT, NT;
Otu017	Leptosphaeriaceae	*Ampelomyces*	cosmopolitan, especially prominent in temperate regions; little is known, many crop pathogens	Fungal Families of the World	Treatment effect-NN different PN, NT; Time effect-decreasing over time
Otu044	Unclassified	Unclassified			Treatment effect-NN different from PN, PT,

					NT;
Otu085	Hypocreaceae	*Acremonium*	cosmopolitan; saprobic on rotting wood and other vegetation or parasitic on other fungi	Fungal Families of the World	Treatment effect-NN different from NT, PT, PN;
Otu007	Incertae sedis	*Epicoccum*			Treatment effect-NN different from PT, PN, NT; Time effect-decreasing over time
Otu047	Hypocreaceae	*Acremonium*	cosmopolitan; saprobic on rotting wood and other vegetation or parasitic on other fungi	Fungal Families of the World	Treatment effect-NN different from NT, PT, PN; Time effect-increasing over time; Interaction-NN increasing over time
Otu028	Didymellaceae	*Leptosphaerulina*			Treatment effect-NN different from PN, PT, NT;
Otu031	Unclassified	Unclassified			Treatment effect-NN different from PN, PT, NT;
Otu052	Pyronemataceae	*Heydenia*	cosmopolitan; saprobic on soil or rotten wood, found growing on plaster etc in buildings	Fungal Families of the World	Treatment effect-NN different from NT, PT, PN; Interaction-NN increasing over time
Otu081	Bionectriaceae	*Gliomastix* OR *Acremonium* OR *Periconia*	cosmopolitan; associated with plant material, especially on wood and herbaceous debris	Fungal Families of the World	Treatment effect-NN different from PN, NT, PT; Time effect-increasing over time; Interaction-NN increasing over time
Otu079	Incertae sedis	*Valsaria*			Treatment effect-PT different from NN, NT;

					Time effect-increasing over time; Interaction-NN, NT decreasing over time
Otu070	Metschnikowiaceae	*Clavispora*	widespread; necrotrophic on plant tissue	Fungal Families of the World	Treatment effect-NN different from NT, PT;
Otu025	Unclassified	Unclassified			Treatment effect-NN different from PN, PT;
Otu050	Pleosporaceae	*Alternaria*	cosmopolitan; necrotrophic pathogens and sprobes, especially associated with grasses	Fungal Families of the World	Treatment effect-NN different from PN, PT;
Otu039	Sporidiobolaceae	*Sporidiobolus*	cosmopolitan; saprobic, found in wide variety of habitats	Fungal Families of the World	Treatment effect-NN different from PN, PT, NT;
Otu026	Metschnikowiaceae	*Metschnikowia*	widespread; necrotrophic on plant tissue	Fungal Families of the World	Treatment effect-NN different from PT;
Otu036	Pleosporaceae	*Alternaria*	cosmopolitan; necrotrophic pathogens and sprobes, especially associated with grasses	Fungal Families of the World	Treatment effect-PN different from NT, NN;
Otu014	Tremellaceae	*Bulleromyces*	cosmopolitan; usually growing on woody substrata, often parasitic on other fungi	Fungal Families of the World	Treatment effect-NN different from NT, PN;
Otu098	Sporidiobolaceae	*Sporobolomyces*	cosmopolitan; saprobic, found in wide variety of habitats	Fungal Families of the World	Treatment effect-NT different from PT, NN, NT highest mean
Otu033	Unclassified	*Cryptococcus*			Treatment effect-NN different from PT, NN; Interaction-NN increasing over time
Otu018	Unclassified	Unclassified			Treatment effect-NN different from NT, PT;
Otu048	Bionectriaceae	*Hydropisphaera*	cosmopolitan; associated	Fungal Families	Treatment effect-NN

			with plant material, especially on wood and herbaceous debris	of the World	different from PN, PT;
Otu056	Unclassified	*Guehomyces*			Treatment effect-NN different from PT, NT;
Otu015	Tremellaceae	*Cryptococcus*	cosmopolitan; usually growing on woody substrata, often parasitic on other fungi	Fungal Families of the World	Treatment effect-PN different from NN;
Otu001	Tremellaceae	*Cryptococcus* OR *Hannaella*	cosmopolitan; usually growing on woody substrata, often parasitic on other fungi	Fungal Families of the World	Treatment effect-NT different from PN;
Otu030	Davidiellaceae	*Cladosporium*			Time effect-increasing over time
Otu012	Trichocomaceae	*Eurotium*	cosmopolitan; saprobes with often aggressive colonization strategies, adaptable to extreme environments, ubiquitious in soil communities, extremely common associates of decaying plant material	Fungal Families of the World	Time effect-increasing over time
Otu063	Ustilaginaceae	*Pseudozyma*	cosmopolitan; number of species are important cereal pathogens, Sporisorium on sugarcane and sorghum; biotrophic in living tissues	Fungal Families of the World	Time effect-decreasing over time
Otu006	Plectosphaerellaceae	*Gibellulopsis* OR *Verticillium*			Time effect-increasing over time
Otu016	Didymellaceae	*Phoma*			Time effect-decreasing over time
Otu010	Nectriaceae	*Fusarium*	cosmopolitan; associated with dead plant material or other fungi, often	Fungal Families of the World	Time effect-decreasing over time

			pathogenic, fusarium/gibberella species plant pathogens		
Otu002	Davidiellaceae	*Cladosporium*			Time effect-increasing over time
Otu077	Davidiellaceae	*Cladosporium*			Time effect-increasing over time
Otu068	Incertae sedis	*Sarocladium* OR *Acremonium*			Interaction-NT decreasing over time, PN increasing over time
Otu023	Nectriaceae	*Fusarium*	cosmopolitan; associated with dead plant material or other fungi, often pathogenic, fusarium/gibberella species plant pathogens	Fungal Families of the World	Interaction-NN decreasing over time

Table 5.10 The 20 most abundant bacterial OTUs with taxon information and associated family ecology.

OTU	Size	Phylum	Class	Order	Family	Family Ecology
Otu0001	17580	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0002	9200	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Plant pathogen
Otu0003	8193	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Plant pathogen
Otu0004	7513	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0005	2952	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0006	2658	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0007	2326	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0008	1916	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Plant pathogen
Otu0009	1575	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Plant pathogen
Otu0010	1429	Actinobacteria	Actinobacteria	Actinomycetales	Sanguibacteraceae	Ubiquitous in soil

Otu0011	1360	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0012	1357	Actinobacteria	Actinobacteria	Actinomycetales	Nocardiopsaceae	Halophilic?
Otu0013	1350	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Plant pathogen
Otu0014	1196	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0015	839	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Ubiquitous in soil
Otu0016	580	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Common in soil (nutrient limiting environments)
Otu0017	538	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	
Otu0018	491	unclassified	unclassified	unclassified	unclassified	
Otu0019	467	Actinobacteria	Actinobacteria	Actinomycetales	Dermabacteraceae	
Otu0020	443	Actinobacteria	Actinobacteria	Actinomycetales	unclassified	
	63963					

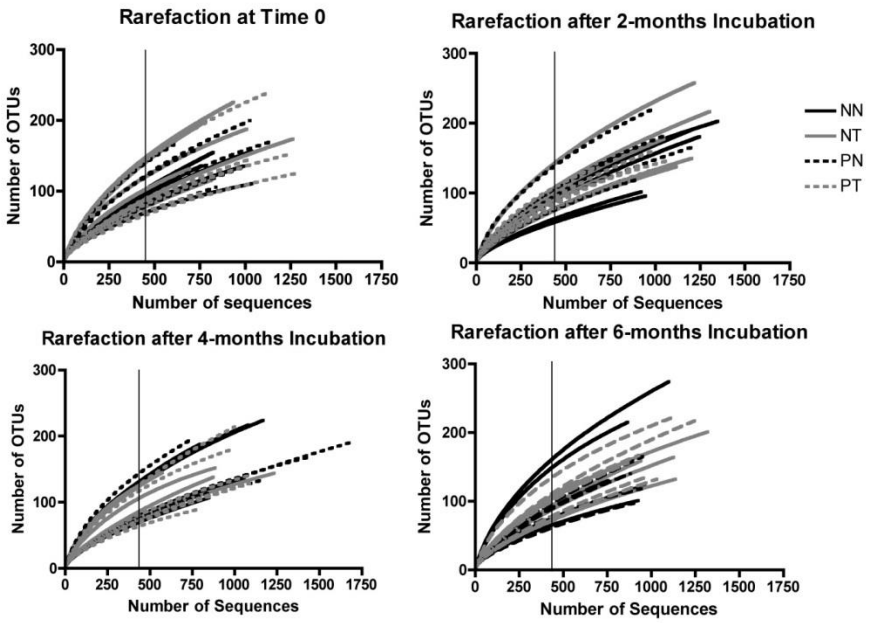


Figure 5.7 Bacterial community rarefaction curves at T0, T2, T4 and T4 months of incubation with subsampling at 450 sequences. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

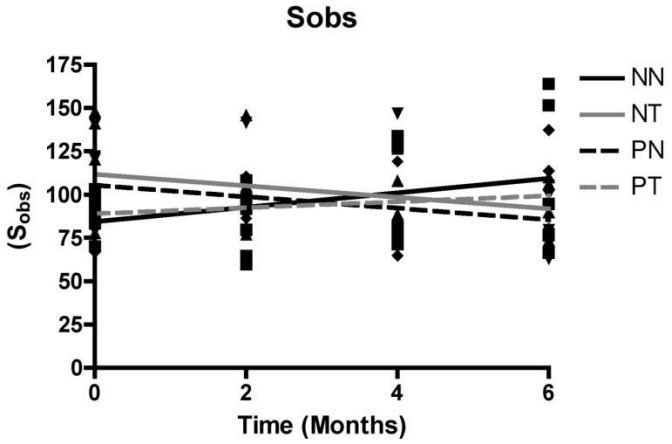


Figure 5.8 Bacterial community richness based on S_{obs} with no differences found between treatments, over time or treatment by time interaction. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

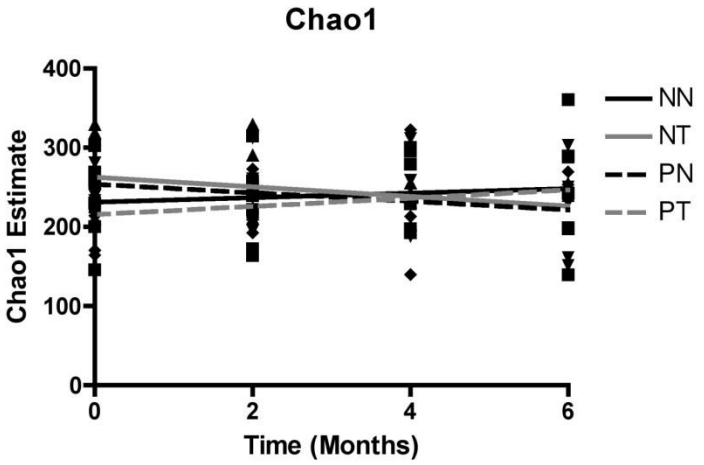


Figure 5.9 Bacterial community estimated richness based on Chao1 with no differences found between treatments, over time or treatment by time interactions. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

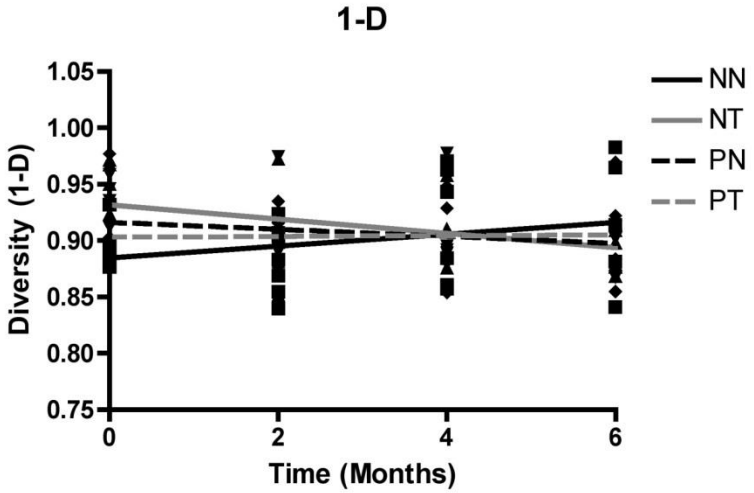


Figure 5.10 Bacterial community diversity (1-D) with no differences between treatments, over time or treatment by time interaction. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

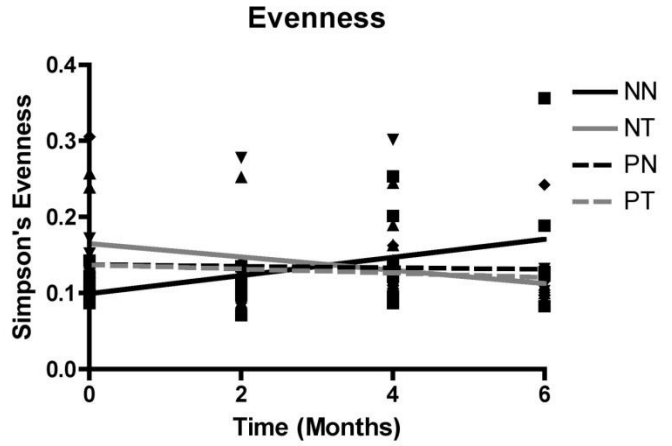


Figure 5.11 Bacterial community evenness with no differences between treatments or over time, treatment by time interaction found significant with evenness increasing in treatment NN and decreasing in treatment NT over time. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Table 5.11 Bacterial community regression, AMOVA and HOMVA analyses from bacterial community assessment metrics. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

Parameter	Treatment	Time	Treatment by Time
S_{obs}	$F_{3,89}$=0.4880, P=0.6916	$F_{3,89}$=0.0281, P=0.8672	$F_{3,89}$=2.7117, P=0.0500
Chao1	$F_{3,89}$=0.3783, P=0.7688	$F_{3,89}$=0.1543, P=0.6954	$F_{3,89}$=1.7297, P=0.1670
Diversity (1-D)	$F_{3,89}$=0.5661, P=0.6389	$F_{3,89}$=0.3535, P=0.5537	$F_{3,89}$=2.2714, P=0.0861
Evenness (E_D)	$F_{3,89}$=0.1434, P=0.9336	$F_{3,89}$=0.0031, P=0.9560	***$F_{3,89}$=2.9586, P=0.0369***
AMOVA	$F_{3,89}$=0.6863, P=0.847	$F_{3,89}$=1.06654, P=0.287	
HOMOVA	$F_{3,89}$=0.9567, P=0.594	$F_{3,89}$=1.5845, P=0.39	

* = Treatment PT was selected as a reference level to emphasize the contrast between the three covered treatments (PT, PN, NT) and the uncovered treatment (NN)
** = P-values test the null hypotheses (H_0: Intercept $_{PN, NT, or NN}$ − Intercept $_{Ref PT}$ = 0; and H_0: Slope $_{PN, NT, or NN}$ − Slope $_{Ref PT}$ = 0). In other words, significant P-values here indicate that the difference between intercept or slope terms for treatments PN, NT, or NN
***AMOVA and HOMOVA- significant at <0.05

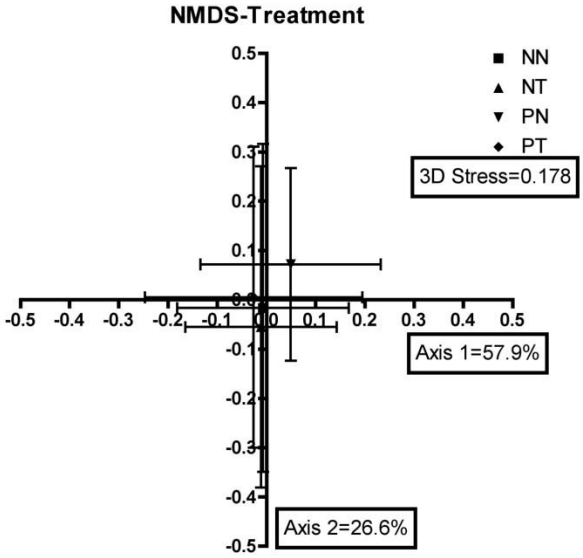


Figure 5.12 Bacterial community Non-metric Multi-dimensional Scaling (NMDS) by treatment based on the bacterial community. (NN indicates uncovered, NT indicates covered with a tarp, PN indicates wrapped with plastic and PT indicates wrapped in plastic and covered with a tarp).

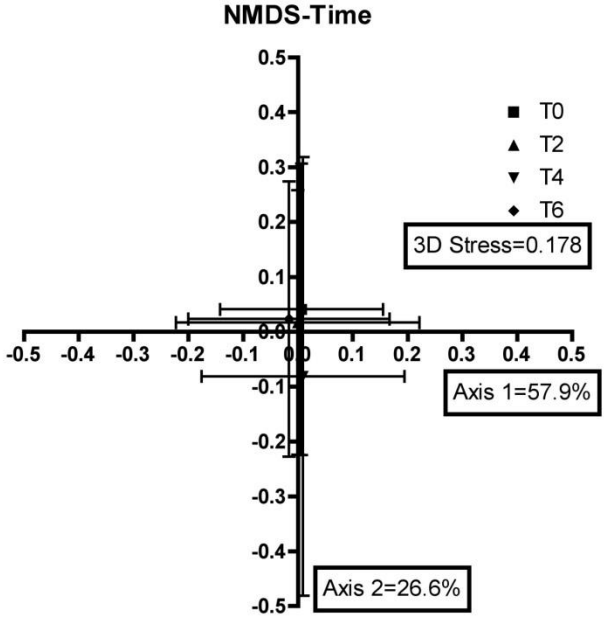


Figure 5.13 Bacterial community Non-metric Multi-dimensional Scaling (NMDS) by time based on the bacterial community. (T0 indicates initial time, T2 indicates after 2-months

incubation, T4 indicates after 4-months incubation and T6 indicates after 6-months incubation).

Chapter 6 - Economic evaluation of baled biomass storage and transport to a Kansas biorefinery

Abstract

Mandates outlined by the Renewable Fuel Standards (RFS) program, established by the Energy Independence and Security Act (EISA) of 2007, have identified lignocellulosic biomass as an ideal renewable resource for the production of the transportation fuel, ethanol. Agricultural residues, like wheat straw, corn stover and sorghum stalks, are abundantly available and are a waste that can be converted to ethanol with relative ease. A challenge in converting agricultural residues to ethanol is in transporting the residues from the field to the biorefinery and storage of the residues to maintain substrate quality for efficient and economical conversion to ethanol throughout the year. The logistics and costs of transportation and storage in three different scenarios were analyzed. Scenario 1 relies on the direct transport of residues to the biorefinery for storage; scenario 2 relies on on-farm/in-field storage after harvest and transport by the farmer to the biorefinery when needed; and scenario 3 introduces satellite storage facilities to which the farmer transports biomass residues for storage and the facility transports to the biorefinery when needed. While the total costs associated with scenario 1 were the lowest, scenario 3 was determined to be an ideal situation as costs were distributed across all supply chain members involved.

Introduction

In 2007, the United States government signed into policy the Energy Independence and Security Act (EISA), which expanded the Renewable Fuel Standards (RFS) program and outlined government mandated goals for the utilization and production of renewable fuels in the United States by 2022. In 2010 the revised RFS2 mandates outlined that 136 billion liters of renewable fuels be used in the US, and 60.5 of the 136 billion liters be produced using cellulosic platform or second generation technologies, along with goals for reduced greenhouse gas (GHG) emissions (Gao et al., 2011). Prior to this bill, the production of biofuels from renewable resources was limited to the production of ethanol from corn and sorghum grain.

With increased interest in the production of liquid transportation fuels from renewable lignocellulosic materials, research has focused on conversion methods of cellulose biomass via

the biochemical pathway. The biochemical conversion of plant biomass (lignocellulosic material) is a three-step process consisting of pretreatment, enzymatic hydrolysis and fermentation. Proposed available biomass has been divided into three categories: wastes, standing forests and energy crops. Biomass in the waste category includes wastes from agricultural production, primarily crop residues like wheat straw, corn stover or sorghum stalks. Standing forests includes the residues from the forestry industry along with short rotation forestry crops like poplar, willow or eucalyptus. Possible energy crops include perennial grasses like switchgrass, miscanthus, big bluestem, and energy sorghum. With the multitude of feedstock options available for lignocellulosic ethanol production, each feedstock type has advantages and disadvantages. In addition, geographic location of a lignocellulosic ethanol biorefinery will limit the feedstocks available locally and seasonally.

With the multitude of feedstocks available for bioconversion to ethanol, determining the available feedstock within a reasonable distance of a biorefinery, the party responsible for storage, the storage method and delivery schedule to maintain a consistent supply to the biorefinery is needed. For this case study we used the agricultural residues abundant in the area surrounding Hugoton, KS. In 2011 Abengoa Bioenergia began the construction of a lignocellulosic ethanol plant that is surrounded by wheat, corn and sorghum fields, which will be used to supply the plant with wheat straw, corn stover and sorghum stalk residues. Martinez and Maier (2011) used GIS-based modeling to quantify the amount of residue from wheat straw, corn stover, and sorghum stalks available to the biorefinery in Hugoton, KS. Based on the available residue around the biorefinery, service areas at 10-mile intervals were determined using the road network in the area. A harvest schedule of the residues available and three different scenarios for delivery and storage were evaluated and the cost to the producer (farmer) and biorefinery were determined.

Assumptions

To fully evaluate the different scenarios proposed, several assumptions were made and held consistent when comparing scenarios. Since composition analysis can be time-consuming and costly, the quality of the residue (substrate) was estimated based on moisture content. Residue with low moisture content (high dry matter content) indicates the cellulose content has likely not degraded during storage. In addition, residue at its driest will be preferable to the

biorefinery, as high moisture residue will interfere with grinding and pre-processing. The delivery price paid to the farmer or satellite facility will be based on dry weight, at a rate of $70 per dry ton, based on USDA markets for the week of August 16, 2013 (USDA, 2013). The price schedule paid to the farmer or satellite facility, based on moisture content, per ton of residue delivered is shown in Table 6.1, with discounts per dry ton starting at biomass over 15% moisture.

Biorefinery Capacity and Requirements

We assumed the biorefinery has a production capacity of 100 million gallons of ethanol per year. Based on the road network around the biorefinery in Hugoton, KS, service areas at 10-mile intervals away from the biorefinery were determined (Figure 6.1). From each service area, the amount of wheat straw, corn stover and sorghum stalk residue that can be conservatively removed, based on soil type and average yield over the past five years was determined (Table 6.2). Based on the available residue in each 10-mile service area, the theoretical amount of ethanol that could be produced from each residue in each service area is shown in Table 6.3. The theoretical amount of ethanol to be produced from the available residue in each service area was calculated using the feedstock-specific theoretical ethanol yields of 128 gallon/dry ton for wheat straw, 130 gallon/dry ton for corn stover, and 113 gallon/dry ton for sorghum stalk (USDOE, 2013). Based on the theoretical ethanol yields, all residues in the 0-10 mile and 10-20 mile service areas and part of the residues in the 20-30 mile service area would need to be collected to supply a 100 million gallon plant production capacity. To reach 100 million gallon production capacity, the plant would need to produce almost 8,400,000 million gallons of ethanol per month. The required biomass residue, in dry tons, needed to supply the 100 million gallon facility would be expected to vary based on feedstock, due to the variability in the conversion factor of each residue to ethanol.

Biomass Production and Harvest Schedule

Costs associated with crop production, harvest and baling of the crops were not taken into consideration, as these costs vary greatly between wheat, corn and sorghum. It was assumed that all residues will be baled into large square bales, due to their efficiency in production and ease of transport and stacking. For all scenarios the same harvesting schedule will be followed with freshly harvested residues utilized first, followed by stored residues. Residues closer to the

biorefinery (service are 0-10 miles) will be utilized first, while those in the outer service areas (10-20 and 20-30 mile) used later. Wheat harvest occurs in June and July, so wheat straw residue harvested in June and July will be utilized at the biorefinery immediately. Left-over wheat straw residue from harvest will be stored and used in August and September. Based on the theoretical ethanol yield of 128 gallons per dry ton, 65,625 dry tons of wheat straw will need to be collected per month during June through September. Corn stover and sorghum stalk residue harvest will occur simultaneously in October and November, and will be utilized first by the biorefinery in those months. Left-over residues will be stored and used by the biorefinery in December and January through May of the following year. Based on the theoretical ethanol yields for corn stover and sorghum stalk, roughly 64,615 and 74,336 dry tons, respectively, will need to be collected to reach a monthly capacity of 8,400,000 gallons at the biorefinery. Sorghum stalk from each service area was considered to be utilized by the biorefinery before corn stover. For all scenarios, residue at harvest will be transported directly to the biorefinery for immediate use without storage coverage; while residue for storage will be transported directly to a storage site (at the biorefinery, satellite facility or on-farm/in-field) to be covered with a tarp or wrapped in plastic; except for scenario 2, in which it is up to the farmer to determine the storage method. Residues at harvest in June, July, October and November will be assumed to maintain 15% moisture content upon delivery to the biorefinery and will not require storage as it will be utilized as it is brought to the facility. Left-over residue bales requiring storage for two months or less will be covered with a tarp, while bales stored longer than two months will be wrapped individually in plastic.

Biomass Storage Conditions and Costs

For all scenarios, moisture content of the residues at harvest and going into storage were assumed to be 15% or less and will remain at 15% if the residue bales are stored under a tarp or wrapped in plastic, regardless of storage duration (Rigdon et al., 2013). Large square bale dimensions are 3 feet wide, 3 feet high and 8 feet long and will be stacked in a configuration containing 2,214 bales, which is 18 feet high, 28 feet wide and 180 feet long (Martinez and Maier, 2011). The area required per bale stack is half an acre, which includes spacing between bale stacks of 1x the width on the shorter side and 1.5x the width for the long side for safety reasons and to leave room to load and unload bales (according to Abengoa Bioenergy). Bale

stacks for storage of two months or less, will be covered with a tarp over the top layer with overhang half-way down the stack. Bales to be stored longer than two months will be individually wrapped in plastic and placed in the same stack configuration. The costs associated with covering residue with a tarp are $3.46 per dry ton and wrapping in plastic costs are $8.23 per dry ton (Darr and Shah, 2012). Coverage of residue bales with a tarp or wrapped in plastic has been shown to have similar dry matter losses of 5-7% and 3-10%, respectively; however, plastic wrap has been found to be a more robust system for long-term storage (Darr and Shah, 2012; Shinners et al., 2007; Shah et al., 2011). In contrast, findings by Rigdon et al. (2013) showed minimal dry matter losses during storage, regardless of coverage type (coverage with tarp or wrapped in plastic). For simplicity, dry matter losses of residue bales covered with a tarp or wrapped with plastic during storage will be considered negligible. Moisture content will also be considered to remain at 15% for residue bales covered with a tarp or wrapped in plastic, while residue bales with no coverage during storage will have a moisture content of 45% at delivery. Discounts for residue will be based on moisture content (Rigdon et al., 2013). Wheat straw residue for use in August and September will be covered with a tarp, in addition to corn stover and sorghum stalk residues for use in December and January. Corn stover and sorghum stalk residues for use in February, March, April and May will be wrapped in plastic for long-term storage. Based on the harvest schedule and the available residues in each service area, the dry tons of each residue feedstock available to the biorefinery or to be stored per service area each month are shown in Table 6.4. A total of 262,500, 363,813 and 87,017 dry tons of wheat straw, corn stover and sorghum stalk, respectively, within 30 miles would be utilized by a 100 MGY biorefinery located near Hugoton, KS. All wheat straw, corn stover and sorghum stalk residue in the 0-10 and 10-20 mile service area will be utilized; while 126,022 and 103,616 dry tons of wheat straw and corn stover, respectively, found in the 20-30 mile service area was needed by the biorefinery (Table 6.4).

Biomass Transport Costs

Transport costs at the biorefinery, at the satellite facility or on the farm/within the field were considered negligible and not included in the transport costs. All residues will be transported by flat-bed semi-truck trailers, with a maximum load weight of 20.9 tons, allowing for a maximum of 28 large square bales at 15% moisture content to be transported per load

(Petrolia, 2008). Transport costs were assumed to be $3.62 per loaded mile (Brechbill et al., 2011). Transport costs were calculated based on the one-way distance traveled and will not include round-trip costs.

Scenarios

For the Hugoton plant we proposed three real-world scenarios for wheat straw, corn stover and sorghum stalk residue transport and storage to maintain a continuous supply of locally available residues to the biorefinery. Each scenario has advantages and disadvantages, both relating to maintaining residue (substrate) quality and economic feasibility. Scenario 1 outlines the storage of all residue bales on-site at the biorefinery; scenario 2 requires all harvested residues to be stored on-farm/in-field by the farmer; and scenario 3 8 utilizes satellite storage facilities located 10 and 20 miles away from the biorefinery.

Results and Discussion

For each scenario the associated costs were calculated and are described in detail below. While some additional assumptions were made in each scenario, general assumptions were mentioned previously. We attempted to account for all possible factors when calculating costs, but for simplicity, assumptions were made and some costs were excluded (i.e., loading/unloading equipment, on-site transport equipment, and land used for storage). All tables show the dry ton requirements of the biorefinery and associated storage and transport cost per month.

Scenario 1-Residue Storage On-site at Biorefinery

For scenario 1, the annual supply of residues will be stored on-site at the biorefinery. Residues baled into large square bales by the custom-harvest company will be transported to the biorefinery for storage, with transport costs paid by the farmer. Residue bales will be used and stored as described in the assumptions. Residue storage will be on the opposite side of the biorefinery, away from ethanol storage as a safety precaution. Advantages of on-site storage include constant inventory to ensure continuous production; control of storage conditions and monitoring capabilities of residue quality, where compromised residue can be processed immediately to reduce losses. The disadvantages of on-site storage include total liability of residue inventory. Large stacks of residue bales have been known to spontaneously combust due to insufficient field drying allowing for internal heating due to microbial metabolism and

chemical reactions (Festensten, 1971; Gregory et al., 1963). Fires within a residue stack could cause explosions in the presence of highly flammable, stored ethanol, and the amount of land area needed for storage is rather large (roughly half an acre per bale stack, as previously mentioned).

Given that all residue bales will be transported immediately after harvest and the moisture content was assumed to be 15%, the farmers will be paid a rate of $70 per dry ton, costing the biorefinery $55,356,070 for a year's supply of feedstock (790,801 dry tons). However, the biorefinery will only be purchasing residue bales during harvest (June, July, October and November), so this cost will be spread over four months. Based on the harvest schedule previously described, the residue of each feedstock available monthly from each service area (Table 6.4), the cost associated with tarp covering or wrapping in plastic was determined. The overall storage related cost for the year is shown in Table 6.5. Residue to be covered with a tarp for use in August, September, December and January, totals 260,480 dry tons, costing a total of $901,621, while residue for use in February, March, April and May (264,262 dry tons) to be wrapped in plastic, will cost $2,174,874 (Table 6.5). The total amount of residue to be stored is 524,742 dry tons, costing at total of $3,076,495 in storage costs. It will be important for the biorefinery to maintain the residue quality for bioconversion because it paid the highest value for the residue. Thus, proper storage management is pivotal.

To defer some costs incurred by the biorefinery, all transportation costs for hauling the residues to the biorefinery are the responsibility of the farmers supplying the residues. The biorefinery's total costs are those associated with purchasing the residues and storage. The total transportation costs to the farmers are variable due to farm/field locations within the service areas. For our analyses, average distances were used for calculating transportation costs, while actual distances will vary by farmer and field location. Fuel prices are also highly volatile, thus increasing variability in transportation costs. The average transportation distance for each service area for the farmers was calculated as the distance from the inner border of the service area plus the average distance traveled within the service area. This was determined to be five miles. So for the 0-10 mile service area the average transport distance is five miles, for the 10-20 mile service area average transport distance is 15 miles, and for the 20-30 mile service area average transport distance is 25 miles. As shown in Table 6.6, transportation costs to the farmers from the 0-10 mile, 10-20 mile and 20-30 mile service areas are $103,537, $1,032,085 and $1,186,455,

respectively, for a grand total of $2,322,077. As previously mentioned, farmers will be paid $70 per dry ton for residue at 15% moisture or less, costing the biorefinery $55,356,070 for 790,801 dry tons. The total storage costs for residue coverage with a tarp or wrapped in plastic is $3,058,046 (Table 6.5) paid by the biorefinery. In total the biorefinery will spend $58,414,116 or $74 per dry ton for feedstock and storage costs for one year operation of a 100 MGY biorefinery.

Scenario 2- Residue Storage On-farm/In-field

Scenario 2 outlines the assumptions, advantages and disadvantages of residue storage on the farm or in the field by the farmer with transport to the biorefinery when needed. Once residues have been harvested by a custom harvest company, large square bales will be stored on the edge of the field from which they were harvested or in a central location designated by the farmer; however, it will be up to the farmer whether they are to be covered with a tarp or wrapped in plastic for on-farm storage. All harvest, storage, and transportation costs will be the responsibility of the farmer. The major disadvantage of on-farm storage is that all liability is on the farmer. Improper storage (residues left uncovered) will result in a major discount to the farmer upon delivery to the biorefinery, and the potential loss of land productivity for storage space on the farm depending whether storage occurs on crop or marginal lands. The cost of residue transport is the responsibility of the farmer, which is a major advantage to the biorefinery, along with reduced need for land area they would have to commit for storage.

For on-farm/in-field storage the farmer will have the option to leave residues uncovered, covered with a tarp or wrapped in plastic. Again, the same configurations for tarp and plastic coverage was assumed and residues left uncovered will be in the same stack configuration as that outlined for the tarp-covered residue. The location of individual farms within the service areas is quite variable, so the storage and transportation costs for the collective farmers within the 30-mile radius of the biorefinery is shown in Table 6.7. The total transportation costs to the farmers, from all service areas is $2,322,077, while storage costs for all service areas totaled $901,233 for tarp coverage and $2,174,876 when bales are wrapped in plastic. For the 0-10 mile service area, the amount paid to the farmers for the residue is $8,368,780 (15% moisture content), which would compensate for the farmers' transportation costs of $103,537 from the field to the biorefinery and result in a return of $6,204,853 above transport and storage costs, which is $51.90 per dry ton. From the 10-20 mile service area the total transportation costs were

$1,032,085 and storage costs of $465,197 and $957,026 for tarp and plastic coverage, respectively, resulting in a return of $25,353,052 or $64 per dry ton after transport and storage costs. From the 20-30 mile service area the transport costs were $1,186,455 and storage costs of $436,036 and $1,217,851 for tarp and plastic coverage, respectively, resulting in a return of $16,339,588 or $59 per dry ton after transport and storage.

Scenario 3- Residue Storage at Satellite Facilities

Scenario 3 outlines the assumptions, advantages and disadvantages of residue storage at satellite facilities. Storage at satellite facilities 10 and 20 miles from the biorefinery were assumed to be run by a third party entity (e.g., grain handling facilities), with four satellite facilities at both 10- and 20-miles from the biorefinery (total of eight satellite storage facilities). Residue will be sold to the satellite facility by the farmer, with the satellite facility subsequently selling the residue to the biorefinery after storage. Payment received by the satellite facility from the biorefinery was assumed to be the same as outlined previously in Table 6.1; however, the maximum amount paid to the farmer by the satellite facility will be $58 per dry ton, with discount delivery price for increased moisture content (Table 6.8). The amount of $58 per dry ton was the maximum amount the satellite facility was assumed to pay the farmer and maintain a reasonable profit margin, based on a 25% overhead for indirect costs (costs not associated with bale storage and transport to the biorefinery, based on satellite locations at 10- and 20-miles and the willingness to assume the risk of preserving the bales during storage). The location of the satellite storage facilities at 10 and 20 miles from the biorefinery are shown in Figure 6.2. Residue collected in the first service area, 0-10 miles, will be transported directly to the biorefinery. Residue collected from the second service area, 10-20 miles, will be transported and stored at the closest satellite facility located at the 10-mile border, while residue collected from the third service area, 20-30 miles, will be transported and stored at the closest satellite facility located at the 20-mile border. Residue brought to the satellite facility will be stored and managed by the satellite facility (including costs for coverage with a tarp or plastic). Residue to be stored two months or less will be covered with a tarp, while residue to be stored for more than 2 months will be wrapped in plastic. Stored residue will subsequently be sold to the biorefinery, with residue stored under a tarp delivered first. The cost of residue transport to the satellite facility was assumed to be the responsibility of the farmer, while the transport costs of the residue from

the satellite facility to the biorefinery was assumed to be the responsibility of the satellite facility. Advantages of this scenario includes splitting transport costs between the farmer and satellite facility, reduced land requirements for residue storage as storage is spread among eight satellite storage facilities, and the hazards of on-site storage at the biorefinery are removed. In addition, liability of preserving the residue during storage is shifted from the farmer and biorefinery to the satellite facilities. The disadvantage of this system is the requirement of multiple storage facilities requiring equipment and personnel to operate and monitor residue during storage.

With the introduction of satellite storage facilities at 10 and 20 miles from the biorefinery, the amount of available residue remains the same for each feedstock, as shown in Table 6.4. However storage and transportation costs are split between the satellite facility and the farmers. Farmers received $58 per dry ton upon direct delivery of all residues to the satellite facilities at harvest with a moisture content of 15%. At the 10-mile satellite facilities a total of 250,743 dry tons were delivered, costing $14,543,094 (paid to the farmer); while the 20-mile satellite facilities received 273,999 dry tons, costing $15,891,942 (paid to the farmer; Table 6.9). Again, the residue will be stored as previously described at the satellite facility, with the amount of residue to be covered with a tarp or wrapped in plastic at satellite facilities located 10 and 20 miles from the biorefinery (Table 6.9). The total storage costs at the 10-mile satellite facilities are $1,422,251 and $1,653,886 for the 20-mile satellite facilities.

As mentioned previously, it was assumed that 20.9 dry tons of residues could be transported on each truck and the cost per truck per loaded mile was $3.62 (one-way). The costs for transporting residues from the satellite facilities to the biorefinery were paid by the satellite facility. Transportation distances from the satellite facilities to the biorefinery were calculated at the furthest point, or 10 miles for transport from the 10-mile satellite facilities and 20 miles for transport from the 20-mile satellite facilities. Table 6.10 shows the 10- and 20-mile satellite facilities cumulative transportation costs for hauling the residue bales from the satellite facilities to the biorefinery. The 10-mile facilities transportation costs totaled $576,139, while the 20-mile facilities transportation costs totaled $949,164. The farmers supplying residues directly to the biorefinery paid transportation costs for transport of residues from the farm to the satellite facilities. The distance travelled by each farmer for transport of residues from the field to the satellite facility is variable, so the average transportation distance within the service areas was assumed to be five miles. The dry tons available, total number of trucks needed for transport and

the associated transportation costs for farmers for residue transport to the biorefinery (0-10 mile service area) and the 10- and 20-mile satellite facilities totaled $103,537, $344,028, and $237,291, respectively (Table 6.11), or $684,856 in total.

From scenario 3, the total residue bale costs, including feedstock cost (at $58 per dry ton; $30,435,036), and transportation ($1,525,303) and storage ($3,076,137) costs paid by the satellite facilities was $35,036,476 (Tables 6.9 and 6.10). Satellite facilities will store residue bales to maintain 15% moisture content, they were assumed to be paid $36,731,940 for 524,742 dry tons of residues, with a return of $1,695,464 above storage, transport and biomass costs.

Comparison of Scenarios

The overall costs of feedstock, transportation and storage, to each supply chain member, i.e., the biorefinery, satellite facilities and farmers, for each scenario is summarized in Table 6.12. Scenario 2 had the highest per dry ton returns after transport for the farmers of $61.17, followed by scenario 1 and scenario 3 with returns after transport and storage of $6.84 and $2.87 per dry ton, respectively. It is important to mention the per dry ton returns after transport and storage in Table 6.12 are based on overall income and expenses over all service areas. In the case of scenario 3, the amount of dry tons collected in service area 20-30 miles is 237,999, which will cost the farmers $1,752,534 to store on-farm/in-field and $26,212,963 to transport to the biorefinery. The total expenses to the farmers in collecting and transporting residue bales directly to the biorefinery is $27,965,497, while only being paid $20,275,220 for the residue bales by the biorefinery, resulting in a reduction of $7,690,277, not including costs associated with residue harvest (Table 6.11). Based on the information presented in Table 6.12, scenario 2 resulted in returns, for both the satellite facility and farmers of $38.72 and $61.17 per dry ton, respectively, after transport and storage. In addition to being profitable for both the satellite facility and farmers, scenario 2 also had the advantage of spreading the land area needed to store the residue bales across a wider area around the biorefinery.

It is also important to note that not all wheat straw and corn stover residues (all sorghum stalk residue was collected, see below) were collected in the 20-30 mile service area, thus allowing for expansion of the capacity of the biorefinery. In our analyses we also assumed that all feedstock residues would be segregated during storage and processing to ethanol, which may not always be the case. It could be more likely that the feedstock residues would be commingled

during storage and processing to ethanol, especially corn stover and sorghum stalk, which are harvested during the same time frame. In the case of corn stover and sorghum stalk, the theoretical ethanol yields from each feedstock are quite different, with sorghum stalk yields being much lower (113 vs 130 gallons per dry ton for sorghum stalk and corn stover, respectively), resulting in a dilution effect on conversion yields. In addition, we assumed all sorghum stalks would be collected and utilized first, followed by corn stover, with the same feedstock price paid to the farmer or satellite facility, regardless of feedstock type. In actuality, the biorefinery or satellite facility may pay a discounted rate or not collect sorghum stalk residue bales due to their reduced yield. However, changing climatic conditions and reduced water availability in Southwest Kansas may cause farmers to shift production from corn to sorghum in future years, although this does not seem to be the trend based on the data collected over the last five years.

Scenario 1 was determined to be the most expensive for the biorefinery, while scenarios 2 and 3 resulted in the only cost to the biorefinery to be that of feedstock procurement. In addition to a lower cost, scenarios 2 and 3 reduce the risk for the biorefinery of preserving residues during storage and allowing the option of paying a reduced price for lower quality feedstocks (higher moisture content) that would likely reduce conversion to ethanol. The latter two scenarios also reduce the amount of land area needed by the biorefinery for residue storage. It is also important to emphasize that during harvest months (i.e., June, July, October and November) residues were directly transported to the biorefinery in each scenario. No residues were covered with a tarp or wrapped in plastic at the biorefinery during harvest months as it was assumed residue was utilized as it arrived at the biorefinery.

Conclusions

Three real-world scenarios were outlined above for the transportation and storage of wheat straw, corn stover and sorghum stalk residues to be harvested for conversion to ethanol by a biorefinery near Hugoton, KS. It was determined, based on theoretical ethanol yields from each feedstock residue, that a 30-mile radius around the biorefinery would be sufficient to supply a 100 MGY biorefinery. A harvest schedule of the three residues was outlined and the storage needs, coverage with a tarp or wrapped in plastic, throughout the year were determined to minimize dry matter losses. Scenario 1 had all harvested residues transported directly to the

biorefinery for storage with farmers responsible for transportation costs and the biorefinery responsible for storage risk and costs for covering residue bales with a tarp or wrapping in plastic. Scenario 2 introduced the utilization of satellite storage facilities 10- and 20-miles, operating similarly to grain elevators, from the biorefinery. Scenario 3 had the responsibility of transportation and storage left to the farmer. Scenario 1 was found to have the lowest total costs and scenario 3 was found to have the highest total costs. In comparison, scenario 2 did not have the lowest costs paid by each party, but it was determined to be the most ideal situation as costs and risks are distributed between supply chain members.

References

Brechbill, S.C., Tyner, W.E., Ileleji, K.E. 2011., The economics of biomass collection and transportation and its supply to Indiana cellulosic and electric utility facilities. Bioenergy Research 4, 141-152.

Darr, M.J., Shah, A., 2012. Biomass storage: an update on industrial solutions for baled biomass feedstocks. Biofuels 3, 321-332.

Festenstein, G.N., 1971. Carbohydrates in hay on self-heating to ignition. J. Sc. Food Agr. 22, 231-234.

Gao, J., Qian, L., Thelen, K. D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103, 844-850.

Gregory, P.H., Lacey, M.E., Festenstein, G.N., Skinner, F.A., 1963. Microbial and biochemical changes during moulding of hay. J. Gen. Microbiol. 33, 147-174.

Martinez, A., Maier, D.E., 2011. Quantifying feedstock availability using a geographical information system. Biological Engineering Transactions. 4, 133-146.

Rigdon, A.R., Jumpponen, A., Vadlani, P.V., Maier, D.E., 2013. Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop. Bioresource Technol. 132, 269-275.

Shah, A., Darr, M.J., Webster, K., Hoffman, C., 2011. Outdoor storage characteristics of single-pass large square corn stover bales in Iowa. Energies 4, 1687-1695.

Shinners, K.J., Adsit, G.S., Binversie, B.N., Digman, M.F., Muck, R.E., Wiemer, P.J., 2007. Single-pass, split-stream harvest of corn grain and stover. Trans. ASABE., 50, 355-363.

USDA. 2013. National Biomass Energy Report. http://www.ams.usda.gov/mnreports/nw_gr310.txt. Accessed August 26, 2013.

USDOE. 2013. Theoretical Ethanol Yield Calculator. National Renewable Energy Laboratory. Washington, D.C.: USDOE National Renewable Energy Laboratory. Available online at http://www1.eere.energy.gov/bioenergy/ethanol_yield_calculator.html. Accessed 09 March 2013.

Tables and Figures

Table 6.1 Delivery price paid by the biorefinery per dry ton residue, based on moisture content. A base price of $70 per dry ton with moisture content of 15% or less was used with discounted delivery price given as moisture content increased.

Moisture Content	Delivery Price, per dry ton
15% and less	$70
15-20%	$65
21-25%	$60
26-30%	$55
31-35%	$50
36-40%	$45
41-45%	$40
46-50%	$35
50% +	$30

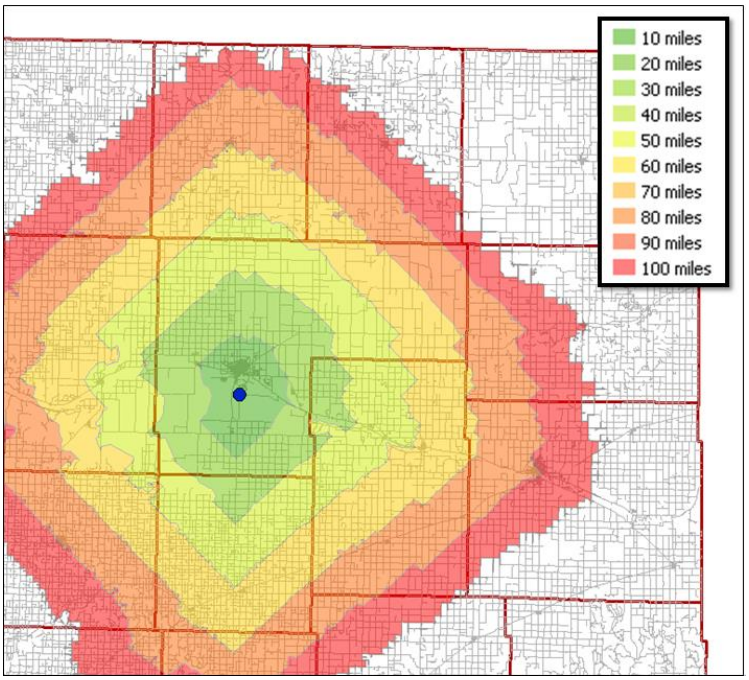


Figure 6.1 GIS-based service areas representing ten 10-mile service areas from the biorefinery (the blue dot) near Hugoton, Kansas. (Martinez and Maier, 2011)

Table 6.2 Estimated annual residue availabile (dry tons) based on five-year average for wheat straw, corn stover and sorghum stalk per 10-mile service area from biorefinery located near Hugoton, KS.

Service Area	0-10 miles		10-20 miles		20-30 miles		30-40 miles		40-50 miles	
Total Dry Tons	119,554		397,248		487,483		821,006		973,507	
	%	Dry Tons	%	Dry Tons	%	Dry Tons	%	Dry Tons	%	Dry Tons
Wheat Straw	27.1%	32,399	26.2%	104,079	42.3%	206,205	45.4%	372,737	39.2%	381,615
Corn Stover	64.8%	77,471	65.5%	260,197	48.6%	236,917	43.8%	359,601	51.2%	498,436
Sorghum Stalk	8.1%	9,684	8.3%	32,972	9.1%	44,361	10.8%	88,669	9.6%	93,457
Total	100.0%	119,554	100.0%	397,248	100.0%	487,483	100.0%	821,006	100.0%	973,507
Cumulative Total		119,554		516,802		1,004,285		1,825,291		2,798,798

Table 6.3 Estimated gallons of ethanol produced annually using theoretical ethanol yields of 128 gal/DT for wheat straw, 130 gal/DT for corn stover, and 113 gal/DT for sorghum stalk per 10-mile service area around the biorefinery location near Hugoton, KS.

Service Area	0-10 miles	10-20 miles	20-30 miles	30-40 miles	40-50 miles
Wheat Straw	4,147,089	13,322,109	26,394,280	48,455,774	48,846,687
Corn Stover	10,071,229	33,825,667	30,799,176	46,748,082	64,796,626
Sorghum Straw	1,094,278	3,725,789	5,012,788	10,019,557	10,560,604
Total	15,312,596	50,873,565	62,206,243	105,223,413	124,203,917
Cumulative Total	15,312,596	66,186,161	128,392,404	233,615,817	357,819,734

Table 6.4 Dry tons available and theoretical ethanol yield of wheat straw (128 gal/DT), corn stover (130 gal/DT) and sorghum stalk (113 gal/DT) in each service area for a biorefinery located near Hugoton, KS.

Service Area		0-10 miles		10-20 miles		20-30 miles		Total Ethanol
Available Biomass of each feedstock residue		Wheat 32,399 Corn 77,471 Sorghum 9,684		Wheat 104,079 Corn 260,197 Sorghum 32,972		Wheat 206,205 Corn 236,917 Sorghum 44,361		
Month	Feedstock	Dry Tons	Ethanol	Dry Tons	Ethanol	Dry Tons	Ethanol	
June	Wheat	32,399	4,147,072	33,226	4,252,928	-	-	8,400,000
July	Wheat	-	-	65,625	8,400,000	-	-	8,400,000
August	Wheat	-	-	5,228	669,184	60,397	7,730,816	8,400,000
September	Wheat	-	-	-	-	65,625	8,400,000	8,400,000
October	Corn	56,198	7,305,708	-	-	-	-	8,400,000
	Sorghum	9,684	1,094,292	-	-	-	-	
November	Corn	21,273	2,765,522	14,682	1,908,642	-	-	8,400,000
	Sorghum	-	-	32,972	3,725,836	-	-	
December	Corn	-	-	64,615	8,400,000	-	-	8,400,000
	Sorghum	-	-	-	-	-	-	
January	Corn	-	-	64,615	8,400,000	-	-	8,400,000
	Sorghum	-	-	-	-	-	-	
February	Corn	-	-	64,615	8,400,000	-	-	8,400,000
	Sorghum	-	-	-	-	-	-	
March	Corn	-	-	51,670	6,717,118	-	-	8,400,000
	Sorghum	-	-	-	-	14,893	1,682,882	
April	Corn	-	-	-	-	39,001	5,070,089	8,400,000
	Sorghum	-	-	-	-	29,468	3,329,911	
May	Corn	-	-	-	-	64,615	8,400,000	8,400,000
	Sorghum	-	-	-	-	-	-	
Total	Wheat	32,399	-	104,079	-	126,022	-	-

Residue Used	**Corn**	77,471	-	260,197	-	103,616	-	-
	Sorghum	9,684	-	32,972	-	44,361	-	-
Left-over Residue	**Wheat**	-	-	-	-	80,183	-	-
	Corn	-	-	-	-	133,301	-	-
	Sorghum	-	-	-	-	-	-	-

Table 6.5 Quantity of wheat straw, corn stover and sorghum stalk residues (dry tons) to be covered with a tarp or wrapped in plastic (and associated costs) throughout the year in each service area for the biorefinery located near Hugoton, KS with a production capacity of 100MGY (total quantity of dry tons required dependent on feedstock and was based on theoretical ethanol yield).

Service Area		0-10 miles	10-20 miles	20-30 miles	Coverage Type			
Available dry tons of each feedstock residue	Wheat	32,399	104,079	206,205	Tarp (dry tons)	Tarp Cost ($3.46/ DT)	Plastic (dry tons)	Plastic Cost ($8.23/ DT)
	Corn	77,471	260,197	236,917				
	Sorghum	9,684	32,972	44,361				
June	Wheat	32,399	33,226	-	-	-	-	-
July	Wheat	-	65,625	-	-	-	-	-
August	Wheat	-	5,228	60,397	65,625	$227,063	-	-
September	Wheat	-	-	65,625	65,625	$227,063	-	-
October	Corn	56,198	-	-	-	-	-	-
	Sorghum	9,684	-	-	-	-	-	-
November	Corn	21,273	14,682	-	-	-	-	-
	Sorghum	-	32,972	-	-	-	-	-
December	Corn	-	64,615	-	64,615	$223,568	-	-
	Sorghum	-	-	-	-	-	-	-
January	Corn	-	64,615	-	64,615	$223,568	-	-
	Sorghum	-	-	-	-	-	-	-
February	Corn	-	64,615	-	-	-	64,615	$531,781
	Sorghum	-	-	-	-	-	-	-
March	Corn	-	51,670	-	-	-	51,670	$425,244
	Sorghum	-	-	14,893	-	-	14,893	$122,567
April	Corn	-	-	39,001	-	-	39,001	$320,976
	Sorghum	-	-	29,468	-	-	29,468	$242,524
May	Corn	-	-	64,615	-	-	64,615	$531,781

	Sorghum	-	-	-	-	-	-	-
Total Dry Tons Covered					260,480	-	264,262	-
Total Cost/Coverage Type					-	$901,621	-	$2,174,874

Table 6.6 Number of trucks needed and average transport costs based on dry tons of wheat straw, corn stover and sorghum stalk residues transported from each service area (in-field) to the biorefinery located near Hugoton, KS.

Service Area		0-10 miles			10-20 miles			20-30 miles		
Available dry tons of each feedstock residue		Wheat		32,399	Wheat		104,079	Wheat		206,205
		Corn		77,471	Corn		260,197	Corn		236,917
		Sorghum		9,684	Sorghum		32,972	Sorghum		44,361
Month	Feedstock	Dry Tons	Trucks Needed	Transport Cost	Dry Tons	Trucks Needed	Transport Cost	Dry Tons	Trucks Needed	Transport Cost
June	Wheat	32,399	1,550	$28,058	33,226	1,590	$86,324	-	-	-
July	Wheat	-	-	-	65,625	3,140	$170,499	-	-	-
August	Wheat	-	-	-	5,228	250	$13,583	60,397	2,890	$261,528
September	Wheat	-	-	-	-	-	-	65,625	3,140	$284,166
October	Corn	56,198	2,689	$48,669	-	-	-	-	-	-
	Sorghum	9,684	463	$8,387	-	-	-	-	-	-
November	Corn	21,273	1,018	$18,423	14,682	702	$38,145	-	-	-
	Sorghum	-	-	-	32,972	1,578	$85,664	-	-	-
December	Corn	-	-	-	64,615	3,092	$167,875	-	-	-
	Sorghum	-	-	-	-	-	-	-	-	-
January	Corn	-	-	-	64,615	3,092	$167,875	-	-	-
	Sorghum	-	-	-	-	-	-	-	-	-
February	Corn	-	-	-	64,615	3,092	$167,875	-	-	-
	Sorghum	-	-	-	-	-	-	-	-	-
March	Corn	-	-	-	51,670	2,472	$134,243	-	-	-
	Sorghum	-	-	-	-	-	-	14,893	713	$64,489
April	Corn	-	-	-	-	-	-	39,001	1,866	$168,880
	Sorghum	-	-	-	-	-	-	29,468	1,410	$127,601
May	Corn	-	-	-	-	-	-	64,615	3,092	$279,792
	Sorghum	-	-	-	-	-	-	-	-	
Total		119,554	5,720	$103,537	397,248	19,007	$1,032,085	273,999	13,110	$1,186,455

Table 6.7 Transportation and storage costs paid by farmers per service areas around the biorefinery near Hugoton, KS and amount paid to farmers for the residue bales delivered.

Service Area	Dry Tons	Amount Paid to Farmer by Biorefinery	Field to Biorefinery Transport Costs to Farmer	Covered with Tarp Costs to Farmer	Wrapped in Plastic Costs to Farmers	Farmer Return after Transport and Storage
0-10 miles	119,554	$8,368,780	$103,537	-	-	$6,204,853
10-20 miles	397,248	$27,807,360	$1,032,085	$465,197	$957,026	$25,353,052
20-30 miles	273,999	$19,179,930	$1,186,455	$436,036	$1,217,851	$16,339,588
Total	790,801	$55,356,070	$2,322,077	$901,233	$2,174,876	$49,957,884

Table 6.8 Amount paid to farmer per dry ton, based on moisture content, by the satellite facilities located at either 10- or 20-miles from a biorefinery.

Moisture Content	Delivery Price, per dry ton
15% and less	$58
15-20%	$53
21-25%	$48
26-30%	$43
31-35%	$38
36-40%	$33
41-45%	$28
46-50%	$23
50% +	$18

Figure 6.2 Satellite storage facilities (black stars) located at a 10- and 20-mile distances from the biorefinery (blue dot) near Hugoton, KS.

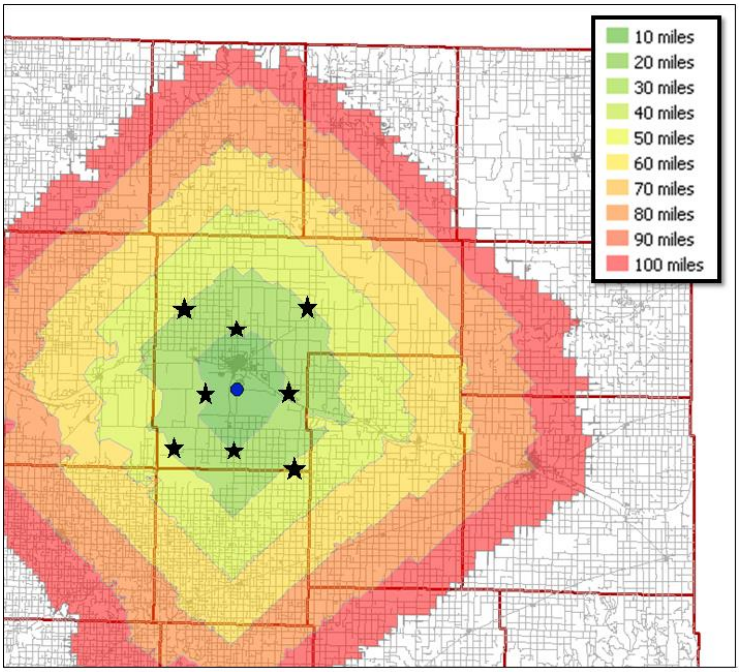


Table 6.9 Quantity and costs for covering bales with a tarp or wrapping in plastic for wheat straw, corn stover and sorghum stalk residues stored at the 10- and 20-mile satellite facilities for a biorefinery near Hugoton, KS.

Storage Costs at 10- and 20-mile Satellite Facilities								
	At 10-mile Facility				At 20-mile Facility			
	Tarp		Plastic		Tarp		Plastic	
	Dry Tons	Cost	Dry Tons	Cost	Dry Tons	Cost	Dry Tons	Cost
June	-	-	-	-	-	-	-	-
July	-	-	-	-	-	-	-	-
August	5,228	$18,089	-	-	60,397	$208,973	-	-
September	-	-	-	-	65,625	$227,063	-	-
October	-	-	-	-	-	-	-	-
November	-	-	-	-	-	-	-	-
December	64,615	$223,568	-	-	-	-	-	-
January	64,615	$223,568	-	-	-	-	-	-
February	-	-	64,615	$531,781	-	-	-	-
March	-	-	51,670	$425,244	-	-	14,893	$122,567
April	-	-	-	-	-	-	68,469	$536,500
May	-	-	-	-	-	-	64,615	$531,781
Total	134,458	$465,225	116,285	$957,026	126,022	$436,036	147,977	$1,217,851
Storage Total	$1,422,251				$1,653,886			
Dry Tons Received	250,743				273,999			
Feedstock Cost ($58/DT)	$14,543,094				$15,891,942			

Table 6.10 Transportation costs for the satellite facility to transport residue bales from the satellite facilities located 10 and 20 miles from the biorefinery near Hugoton, KS (Total storage costs, total paid for residue to the farmer and grand total paid by satellite facilities 10 and 20 miles from the biorefinery).

	From Satellite at 10 miles			From Satellite at 20 miles		
	Dry Tons Available	**Trucks Needed**	**Transport Costs**	**Dry Tons Available**	**Trucks Needed**	**Transport Costs**
June	33,226	1,590	$57,549	-	-	-
July	65,625	3,140	$113,666	-	-	-
August	5,228	250	$9,055	60,397	2,890	$209,222
September	-	-	-	65,625	3,140	$227,333
October	-	-	-	-	-	-
November	47,654	2,280	$82,539	-	-	-
December	64,615	3,092	$111,917	-	-	-
January	64,615	3,092	$111,917	-	-	-
February	51,670	2,472	$89,495	-	-	-
March	-	-	-	14,893	713	$51,591
April	-	-	-	68,469	3,276	$237,184
May	-	-	-	64,615	3,092	$223,834
Total Transport Costs	$576,139			$949,164		

Table 6.11 Transportation costs for the farmers supplying residue bales directly to the biorefinery from the 0-10 mile service area and to satellite facilities located 10 and 20 miles biorefinery near Hugoton, KS.

	0-10 mile Service Area			To 10-mile Satellite			To 20-mile Satellite		
	Dry Tons Available	# of Trucks Needed	Transport Costs	Dry Tons Available	# of Trucks Needed	Transport Costs	Dry Tons Available	# of Trucks Needed	Transport Costs
June	32,399	1,550	$28,058	33,226	1,590	$28,775	-	-	-
July	-	-	-	65,625	3,140	$56,833	-	-	-
August	-	-	-	5,228	250	$4,528	60,397	2,890	$52,306
September	-	-	-	-	-	-	65,625	3,140	$56,833
October	65,882	3,152	$57,056	-	-	-	-	-	-
November	21,273	1,018	$18,423	47,654	2,280	$41,270	-	-	-
December	-	-	-	64,615	3,092	$55,958	-	-	-
January	-	-	-	64,615	3,092	$55,958	-	-	-
February	-	-	-	64,615	3,092	$55,958	-	-	-
March	-	-	-	51,670	2,472	$44,748	14,893	713	$12,898
April	-	-	-	-	-	-	68,469	3,276	$59,296
May	-	-	-	-	-	-	64,615	3,092	$55,958
Total	$103,537			$344,028			$237,291		

Table 6.12 Summary of expenses (feedstock, storage and transport) incurred by the biorefinery, satellite facilities and farmers for each scenario, along with total returns after storage and transport and returns per dry ton (not including overhead) for satellite facility and farmers.

		Scenario 1	Scenario 2	Scenario 3
Biorefinery	**Income**	-	-	-
	Expenses			
	Feedstock	$55,356,070	$55,356,070	$55,356,070
	Storage	$3,076,495	-	-
	Transport	-	-	-
	Total	$58,432,565	$55,356,070	$55,356,070
Farmers	**Income**	$55,356,070	$55,356,070	$51,048,776
	Expenses			
	Feedstock	-	-	-
	Storage	-	$3,053,046	-
	Transport	$2,322,077	$2,322,077	$684,856
	Total	$2,322,077	$5,375,123	$684,856
	Return	$53,033,993	$49,980,947	$50,363,920
	Return/DT	$67	$63	$64
Satellite Facility	**Income**	-	-	$46,987,290
	Expenses			
	Feedstock	-	-	$38,932,326
	Storage	-	-	$3,076,137
	Transport	-	-	$1,637,220
	Total	-	-	$43,645,683
	Return	-	-	$3,341,607
	Return/DT	-	-	$5

Chapter 7 - Overall Conclusions

The following overall conclusions are drawn from this work:

- From biomass harvest to conversion to ethanol, the storage method (covered or uncovered) played a large role in preserving substrate quality and subsequent ethanol yields.
- Uncovered biomass enzymatic activity related to cellulose degradation was found to increase, 0.7 µM/g/hr to 1.9 µM/g/hr based on 4-MUB-β-D-cellobioside and 2.0 µM/g/hr to 3.9 µM/g/hr based on 4-MUB-β-D-glucoside, after 2 months of storage with enzyme activity remaining elevated at 4 and 6 months storage.
- Uncovered biomass cellulose content decreased from 35% to 25% after 2 months of storage and remained at 22% and 25% at 4 and 6 months, respectively.
- Uncovered biomass ethanol yields also decreased from 0.19 g ethanol/g biomass to 0.02 g ethanol/g biomass after 6 months of storage.
- Covered biomass enzymatic activity related to cellulose and hemicellulose degradation, cellulose and hemicellulose content and ethanol yields were found to remain stable during a six-month storage period.
- Enzymatic activity related to cellulose and hemicellulose degradation, biomass components (cellulose, hemicellulose and lignin) and conversion to ethanol yield datasets were all found to be congruent.
- Fungal community richness of uncovered biomass increased over time, while community evenness decreased; community diversity increased over time in covered and uncovered biomass based on LSU-targeted sequencing.
- Fungal communities in covered and uncovered biomass were found to be distinctly different; distinct communities were also seen over time, all based on LSU-targeted sequencing.
- *Cladosporium, Alternaria* and *Cryptococcus* were found to be the most abundant genera present in stored biomass, covered or uncovered; the genera *Wallemia* increased in abundance in the storage treatment NT, likely due to the unique environment created by the tarp, all based on LSU-targeted sequencing.

- ITS- and LSU-targeted sequencing datasets were found to be congruent when interpreting community dynamics in stored biomass.
- Community evenness and diversity in uncovered biomass decreased based on ITS-targeted sequencing.
- Fungal communities in covered and uncovered biomass were found to be distinctly different; distinct communities were also seen over time, all based on LSU-targeted sequencing.
- *Cladosporium, Alternaria* and *Cryptococcus* were found to be the most abundant genera present in stored biomass, covered or uncovered; the genera *Wallemia* increased in abundance in the storage treatment NT, likely due to the unique environment created by the tarp, all based on ITS-targeted sequencing.
- Covering of biomass strongly limits the arrival and establishment of new fungal propagules in stored biomass, reducing biomass degradation by these often pathogenic, saprobic or endophytic communities, based on LSU- and ITS-targeted sequencing.
- Bacterial communities were found to be unresponsive to storage treatment and duration with Enterobacteriaceae representing the most abundant sequences found.
- The quantity of dry tons of wheat straw, corn stover and sorghum stalk residues to supply a 100 MGY biorefinery was available within in a 30-mile radius and was predicted to cost the biorefinery $55,356,070 at $70 per dry ton, assuming no discounts.
- Based on three scenarios outlined, the scenario introducing satellite storage facilities at 10- and 20-miles from the biorefinery was determined to be the ideal situation for residue transport and storage as risks and costs were distributed among the biorefinery, satellite storage facilities and farmers supplying residue.

Chapter 8 - Future Work

As the biofuels industry continues to grow and utilize lignocellulosic feedstocks, research and development is vital for improving yields, reducing environmental impacts and improving the profitability of biorefineries. Based on the results from this work, suggestions for future work can be made.

- Replication of harvest and storage of biomass used for lignocellulosic ethanol production, as only one harvest year and 6-months of storage were included in this work. Extension of storage duration pass 6-months would provide additional data necessary to properly manage storage of biomass.
- Continued evaluation of the microbial communities present in stored biomass. Metagenomic assay to isolate microbial cellulose and hemicellulose degrading enzymes from communities growing in biomass stored for extended durations. Enzymes isolated would likely be capable of degradative activity at ambient temperatures, which would be beneficial to the biorefinery during the conversion of biomass to ethanol.
- Optimization of satellite storage facility locations and logistics to reduce costs to farmers and the biorefinery. Determine ideal location of satellite storage facilities where available biomass has highest density, to reduce transport costs.

Chapter 9 - References

Chapter 1

Emery, I.R., Mosier, N.S., 2012. The impact of dry matter loss during herbaceous biomass storage on net greenhouse gas emissions from biofuels production. Biomass Bioenerg. 39, 237-246.

Gomez, L.D., Steele-King, C.G., McQueen-Mason, S.J., 2008. Sustainable liquid biofuels from biomass: the writing's on the walls. New Phytol., 178, 473-485.

Hess, J.R., Wright, C.T., Kenney, K.L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel Biop. Bior., 1, 181-190.

Kruger, D., Kapturska, D., Fischer, C., Daniel, R., Wubet, T., 2012. Diversity measures in environmental sequences are highly dependent on alignment quality—data from ITS and new LSU primers targeting basidiomycetes. PLoS ONE 7:e32139-.

Linton, J.A., Miller, J.C., Little, R.D., Petrolia, D.R., Coble, K.H., 2011. Economic feasibility of producing sweet sorghum as an ethanol feedstock in the southeastern United States. Biomass Bioenerg. 35, 3050-3057.

Liu, Z., Lozupone, C., Hamady, M., Bushman, F.D., Knight, R., 2007. Short pyrosequencing reads suffice for accurate microbial community analysis. Nucleic Acids Res. 35:e120-.

Metzker, M.L., 2010. Sequencing technologies- the next generation. Nat. Rev. Genet., 11, 31-46.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K., 2010. Production of first and second generation biofuels: a comprehensive review. Renew. Sust. Energ. Rev. 14, 578-597.

Perlack, R.D., Wright, L.L., Turhollow, A.F., Graham, R.L., Stokes, B.J., Erback, D.C., 2005. Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply. US Department of Energy DE-AC05-000R22725.

Porter, T.M., Golding, G.B., 2012. Factors that affect large subunit ribosomal DNA amplicon sequencing studies of fungal communities: classification method, primer choice, and error. PLoS ONE 7:e35749-.

Schaad, N.W., Jones, J.B., Chun, W., 2001. Laboratory guide for identification of plant pathogenic bacteria, third edition. APS Press.

Schloss, P.D., et al. 2009. Introduction mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. App. Environ. Microb. 75, 7537-7541.

Sharma, R., Ranjan, R., Kapardar, R.K., Grover, A., 2005. 'Unculturable' bacterial diversity: an untapped resource. Curr. Sci. India, 89, 72-77.

Shinners, K.J., Binversie, J.B.N., 2004. Harvest and storage of wet corn stover biomass. American Society of Agricultural Engineers Annual Meeting 2004. ASAE041159. ASAE, St. Joseph, Mich.

Sokhansanj, S., Hess, J.R., 2009. Biomass supply logisitics and infrastructure. In: Mielenz, J.R. (Ed.), *Biofuels: methods and protocols, methods in molecular biology*. 581, 1-25.

Stevens, C.V., Verhe, R., 2004. Renewable bioresources scope and modification for non-food applications. England: John Wiley and Sons Ltd.

Vilo, C., Dong, Q., 2012. Evaluation of RDP classifier accuracy using 16S rRNA gene variable regions. Metagenomics 1: doi:10.4303/mg/235551.

Chapter 3

Balat, M., 2011. Production of bioethanol from lignocellulosic materials via the biochemical pathway: a review. Energ. Convers. Manage. 52, 858-875.

Brijwani, K., Oberoi, H.S., Vadlani, P.V., 2010. Production of cellulolytic enzyme system in mixed-culture solid-state fermentation of soybean hulls supplemented with wheat bran. Process Biochem. 45, 120-128.

Darrah, P.R., Harris, P.J., 1986. A fluorimetric method for measuring the activity of soil enzymes. Plant Soil. 92, 81-88.

DeForest, J.L., 2009. The influence of time, storage temperature, and substrate age on potential soil enzyme activity in acidic forest soils using MUB-linked substrates and L-DOPA. Soil Biol. Biochem. 41, 1180-1186.

Emery, I.R., Mosier, N.S., 2012. The impact of dry matter loss during herbaceous biomass storage on net greenhouse gas emissions from biofuels production. Biomass Bioenerg. 39, 237-246.

Freeman, C., Liska, G., Ostle, N.J., Jones, S.E., Lock, M.A., 1995. The use of fluorogenic substrates for measuring enzyme activity in peatlands. Plant Soil. 175, 147-152.

Gao, J., Qian, L., Thelen, K. D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103 (3), 844-850.

Gibbons W.R., Hughes, S.R., 2009. Integrated biorefineries with engineered microbes and high-value co-products for profitable biofuels production. In Vitro Cell. Dev.-Pl. 45, 218-228.

Hess J.R., Wright, C.T., Kenney, K. L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel. Bioprod. Bior. 1, 181-190.

Hoppe, H-G., 1983. Significance of exoenzymatic activities in the ecology of brackish water: measurement by means of methylumbelliferyl-substrates. Mar. Ecol.-Prog. Ser., 11, 299-308.

Khanchi, A., Jones, C.L., Sharma, B., 2009. Characteristics and compositional variation in round and square sorghum bales under different storage conditions. ASABE Meeting Presentation. Paper Number:096672.

Marx, M.C., Wood, M., Jarvis, S.C., 2001. A microplate fluorimetric assay for the study of enzyme diversity in soils. Soil Biol. Biochem. 33, 1633-1640.

Oberoi, H.S., Vadlani, P.V., Nanjudaswamy, A., Bansal, S., Singh, S., Kaur, S., Babbar, N., 2011. Enhanced ethanol production from Kinnow mandarin (*Citrus reticulate*) waste *via* a statistically optimized simultaneous saccharification and fermentation process. Bioresour. Technol. 102, 1593-1601.

Saiya-Cork, K.R., Sinsabaugh, R.L., Zak, D.R., 2002. The effects of long term nitrogen deposition on extracellular enzyme activity in an *Acer saccharum* forest soil. Soil Biol. Biochem. 34, 1309-1315.

Selmants, P.C., Hart, S.C., Boyle, S.I, Stark, J.M., 2005. Red alder (*Alnus rubra*) alters community-level soil microbial function in conifer forests of the Pacific Northwest, USA. Soil Biol. Biochem. 37, 1860-1868.

Shah, A., Darr, M.J., Webster, K., Hoffman, C., 2011. Outdoor storage characteristics of single-pass large square corn stover bales in Iowa. Energies 4, 1687-1695.

Somville, M., 1984. Measurement and study of substrate specificity of exoglucosidase activity in eutropic water. Appl. Environ. Microb., 48, 1181-1185.

Yoo, J., Alavi, S., Vadlani, P., Amanor-Boadu, V., 2011. Thermo-mechanical extrusion pretreatment for conversion of soybean hulls to fermentable sugars. Bioresour. Technol. 102, 7583-7590.

Zheng, Y., Yu, C., Cheng, Y., Lee, C., Simmons, C.W., Dooley, T.M., Zhang, R., Jenkins, B.M., VanderGheynst, J.S., 2012. Integrating sugar beet pulp storage, hydrolysis and fermentation for fuel ethanol production. Appl Energ. 93, 168-175.

Chapter 4

Aime, M.C., Matheny, P.B., Henk, D.A., Frieders, E.M., Nilsson, R.H., Piepenbring, M., McLaughlin, D.J., Szabo, L.J., Begerow, D., Sampaio, J.P., Bauer, R., Weib, M., Oberwinkler, F., Hibbett, D.. 2006. An overview of the higher level classification of Puccinomycotina based on combined analyses of nuclear large and small subunit rDNA sequences. Mycologia 98, 896-905.

Amend, A.S., Seifert, K.A., Bruns, T.D.. 2010. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol. Ecol. 19, 5555-5565.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral. Ecol. 26, 32-46.

Bensch, K., Groenewald, J.Z., Dijksterhuis, J., Starink-Willemse, M., Anderson, B., Summerell, B.A., Shin, H.-D., Dugan, F.M., Schroers, H.-J., Braun, J., Crous, P.W., 2010. Species and ecological diversity within the *Cladosporium cladoporioides* complex (*Davidiellaceae, Capnodiales*). Stud. Mycol. 67, 1-94.

Berry, D., Mahfoudh, K.B., Wagner, M., Loy, A., 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl. Environ. Microbiol. 77, 7846-7849.

Blackwell, M., 1986. A new species of *Pyxidiophora* and its *Thaxteriola* anamorph. Mycologia 78, 605-612.

Blackwell, M., Moser, J.C., Wisniewski, J., 1989. Ascospores of *Pyxidiophora* on mites associated with beetles in trees and wood. Mycol. Res. 92, 397-403.

Cain, R.F., Weresub, L.K., 1957. Studies of coprophilous Ascomycetes V. *Sphaeronaemella fimicola*. Can. J. Botany. 35, 119-131.

Edgar R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011 UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194-2200.

Emery, I.R., Mosier, N.S., 2012. The impact of dry matter loss during herbaceous biomass storage on net greenhouse gas emissions from biofuels production. Biomass Bioenerg. 39, 237-246.

Frankland, J.C., 1998. Fungal succession—unraveling the unpredictable. Mycol. Res. 102, 1-15.

Gao, J., Qian, L., Thelen, K.D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103, 844-850.

Gibbons W.R., Hughes, S.R., 2009. Integrated biorefineries with engineered microbes and high-value co-products for profitable biofuels production. In Vitro Cell. Dev.-Pl. 45, 218-228.

Gihring, T.M., Zhang, G., Brandt, C.C., Brooks, S.C., Campbell, J.H., Carroll, S., Criddle, C.S., Green, S.J., Jardine, P., Kostka, J.E., Lowe, K., Mehlhorn, T.L., Overholt, W., Watson, D.B., Yang, Z., Wu, W., Schadt, C.W., 2011. A limited consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. Appl. Environ. Microb. 77, 5955-5965.

Hansen, K., Perry, B.A., Pfister, D.H., 2005. Phylogenetic origins of two cleistothecial fungi, *Orbicula parietina* and *Lasiobolidium orbiculoides*, within the operculate discomycetes. Mycologia 97, 1023-1033.

Hess J.R., Wright, C.T., Kenney, K.L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel. Bioprod. Bior. 1, 181-190.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Fox, C.J., et al., 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443, 818-822.

Kleczewski, N.M., Bauer, J.T., Bever, J.D, Clay, K., Reynolds, H.L., 2012. A survey of endophytic fungi in switchgrass (*Panicum virgatum*) in the Midwest, and their putative roles in plant growth. Fungal Ecol. 5, 521-529.

Lukens, R.J., Horsfall, J.G., 1972. Processes of sporulation in alternaria solani and their response to metabolic inhibitors. Phytopathology 63, 176-182.

Magauran, A.E., 1988. *Ecological Diversity and its Measurement*. Princeton University, Princeton.

Margulies, M., et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376-380.

Mendez-Mayboca, F., Checa, J., Esqueda, M., Chacon, S., 2010. New records of *Loculoascomycetes* from natural protected areas in Sonora, Mexico. Mycotaxon 111, 19-30.

Muller, C.E., 2005. Fermentation patterns of small-bale silage and haylage produced as a feed for horses. Grass Forage Sci. 60, 109-118.

Nilsson, R.H., Tedersoo, L., Lindahl, B.D., Kjoller, R., Carlsen, T., Quince, C., Abarenkov, K., et al., 2011. Towards standardization of the description and publication of next-generation sequencing datasets of fungal communities. New Phytol. 119, 314-318.

Poll, C., Brune, T., Begerow, D., Kandeler, E., 2010. Small-scale diversity and succession of fungi in the detritusphere of rye residues. Microb. Ecol. 59, 130-140.

Quince C, Lanzén, A., Curtis, T.P., Davenport, R.J., Hall, N., Head, I.M., Read, L.F., Sloan, W.T., 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. Nat. Methods 6, 639.

Quince C, Lanzen, A., Davenport, R.J., Turnbaugh, P.J., 2011. Removing noise from pyrosequenced amplicons. BMC Bioinformatics 12, 38.

Rigdon, A.R., Jumpponen, A., Vadlani, P.V., Maier, D.E., 2013. Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop. Bioresource Technol. 132, 269-275.

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:e27310.

Stevens, F.L., 1913. The Fungi Which Cause Plant Disease-book. The MacMillian Company, New York

Tedersoo, L., Nilsson, R.H., Abarenkov, K., et al., 2010. 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. New Phytol. 188, 291-301.

Vakili, N.G., 1992. Biological seed treatment of corn with mycopathogenic fungi. J. Phytopathology 134, 313-323.

Vilgalys, R., Hester, M., 1990. Rapid genetic identification and mapping of enzymatically amplified ribosomal DNA from several *Cryptococcus* species. J. Bacteriol. 172, 4238-4246.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microb. 73, 5261-5267.

Zalar, P., de Hoog, G.S., Schroers, H.-J., Frank, J.M., Gunde-Cimerman, N., 2005. Taxonomy and phylogeny of the xerophilic genus *Wallemia* (Wallemiomycetes and Wallemiales, cl. et ord. nov.). Anton. Leeuw. Int. J. G. 87, 311-328.

Chapter 5

Ali, Sk.Z., Sandhya, V., Grover, M., Kishore, N., Rao, L.V., Venkateswarlu, B., 2009. Pseudomonas sp strain AKM-P6 enhances tolerance of sorghum seedlings to elevated temperatures. Biology and Fertility of Soils 46, 45-55.

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral. Ecol. 26, 32-46.

Berry, D., Mahfoudh, K.B., Wagner, M., Loy, A., 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. Appl. Environ. Microbiol. 77, 7846-7849.

Cota, L.V., Costa, R.V., Silva, D.D., Parreira, D.F., Lana, U.G.P., Casela, C.R., 2010. First report of pathogenicity of Pantoea ananatis in sorghum (Sorghum bicolor) in Brazil. Australasian Plant Disease Notes 5, 120-122.

Coutinho, T.A., Venter, S.N., 2009. Pantoea ananatis: an unconventional plant pathogen. Molecular Plant Pathology 10, 325-335.

Edgar R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011 UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194-2200.

Frankland, J.C., 1998. Fungal succession—unraveling the unpredictable. Mycol. Res. 102, 1-15.

Fierer, N., Hamady, M., Lauber, C.L., Knight, R., 2008. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. PNAS. 105, 17994-17999.

Gao, J., Qian, L., Thelen, K.D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103, 844-850.

Gardes, M., Bruns, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes – application to the identification of mycorrhizae and rusts. Mol. Ecol. 2, 113-118.

Gihring, T.M., Zhang, G., Brandt, C.C., Brooks, S.C., Campbell, J.H., Carroll, S., Criddle, C.S., Green, S.J., Jardine, P., Kostka, J.E., Lowe, K., Mehlhorn, T.L., Overholt, W., Watson, D.B., Yang, Z., Wu, W., Schadt, C.W., 2011. A limited consortium is responsible for extended bioreduction of uranium in a contaminated aquifer. Appl. Environ. Microb. 77, 5955-5965.

Hess J.R., Wright, C.T., Kenney, K.L., 2007. Cellulosic biomass feedstocks and logistics for ethanol production. Biofuel. Bioprod. Bior. 1, 181-190.

Jumpponen, A., Jones, K.L., 2010. Seasonally dynamic fungal communities in the *Quercus macrocarpa* phyllosphere differ between urban and nonurban environments. New Phytol. 186, 496-513.

Kleczewski, N.M., Bauer, J.T., Bever, J.D., Clay, K., Reynolds, H.L., 2012. A survey of endophytic fungi in switchgrass (*Panicum virgatum*) in the Midwest, and their putative roles in plant growth. Fungal Ecol. 5, 521-529.

Lukens, R.J., Horsfall, J.G., 1972. Processes of sporulation in alternaria solani and their response to metabolic inhibitors. Phytopathology 63, 176-182.

Magauran, A.E., 1988. *Ecological Diversity and its Measurement*. Princeton University, Princeton.

Mastretta, C., Taghavi, S., van der Lelie, D., Mengoni, A., Galardi, F., Gonnelli, C., Barac, T., Boulet, J., Weyens, N., Vangronsveld, J., 2009. Endophytic bacteria from seeds of Nicotiana tabacum can reduce cadmium phytotoxicity. International Journal of Phytoremediation 11, 251-267.

Muller, C.E., 2005. Fermentation patterns of small-bale silage and haylage produced as a feed for horses. Grass Forage Sci. 60, 109-118.

Pederson, W.L., Chakrabarty, K., Klucas, R.V., Vidaver, A.K., 1978. Nitrogen fixation (acetylene reduction) associated with roots of winter wheat and sorghum in Nebraska. Applied and Environmental Microbiolgy 35, 129-135.

Poll, C., Brune, T., Begerow, D., Kandeler, E., 2010. Small-scale diversity and succession of fungi in the detritusphere of rye residues. Microb. Ecol. 59, 130-140.

Rigdon, A.R., Jumpponen, A., Vadlani, P.V., Maier, D.E., 2013. Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop. Bioresource Technol. 132, 269-275.

Roberts, R.G., 1990. Postharvest biological control of gray mold of apple by Cryptococcus laurentii. Phytopathology 80, 526-530.

Rodrigues, A., Cable, R.N., Mueller, U.G., Bacci Jr., M., Pagnocca, F.C., 2009. Antagonistic interactions between garden yeasts and microfungal garden pathogens of leaf-cutting ants. Antonie van Leeuwenhoek 96, 331-342.

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:e27310.

Zaldivar, J.J., Nielsen, J.J., Olsson, L.L., 2001. Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. Applied Microbiology and Biotechnology 56, 17-34.

White, T.J., Bruns, T., Lee, S., Taylor, J., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: M.A. Innis, D.H. Glefand, J.J. Sninsky, T.J. White (eds) *PCR Protocols: a Guide to Methods and Applications*, pp. 315-322. Academic Press, New York.

Chapter 6

Brechbill, S.C., Tyner, W.E., Ileleji, K.E. 2011., The economics of biomass collection and transportation and its supply to Indiana cellulosic and electric utility facilities. Bioenergy Research 4, 141-152.

Darr, M.J., Shah, A., 2012. Biomass storage: an update on industrial solutions for baled biomass feedstocks. Biofuels 3, 321-332.

Festenstein, G.N., 1971. Carbohydrates in hay on self-heating to ignition. J. Sc. Food Agr. 22, 231-234.

Gao, J., Qian, L., Thelen, K. D., Hao, X., da Costa Sousa, L., Lau, M.W., Balan, V., Dale, B.E., 2011. Corn harvest strategies for combined starch and cellulosic bioprocessing to ethanol. Agron. J. 103, 844-850.

Gregory, P.H., Lacey, M.E., Festenstein, G.N., Skinner, F.A., 1963. Microbial and biochemical changes during moulding of hay. J. Gen. Microbiol. 33, 147-174.

Martinez, A., Maier, D.E., 2011. Quantifying feedstock availability using a geographical information system. Biological Engineering Transactions. 4, 133-146.

Rigdon, A.R., Jumpponen, A., Vadlani, P.V., Maier, D.E., 2013. Impact of various storage conditions on enzymatic activity, biomass components and conversion to ethanol yields from sorghum biomass used as a bioenergy crop. Bioresource Technol. 132, 269-275.

Shah, A., Darr, M.J., Webster, K., Hoffman, C., 2011. Outdoor storage characteristics of single-pass large square corn stover bales in Iowa. Energies 4, 1687-1695.

Shinners, K.J., Adsit, G.S., Binversie, B.N., Digman, M.F., Muck, R.E., Wiemer, P.J., 2007. Single-pass, split-stream harvest of corn grain and stover. Trans. ASABE., 50, 355-363.

USDA. 2013. National Biomass Energy Report. http://www.ams.usda.gov/mnreports/nw_gr310.txt. Accessed August 26, 2013.

USDOE. 2013. Theoretical Ethanol Yield Calculator. National Renewable Energy Laboratory. Washington, D.C.: USDOE National Renewable Energy Laboratory. Available online at http://www1.eere.energy.gov/bioenergy/ethanol_yield_calculator.html. Accessed 09 March 2013.

Appendix A - Molecular Microbiological Methods

Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) was developed in the late 1980s and allowed for the amplification of DNA fragments. A single or few strands of DNA could be amplified to several orders of magnitude, generating thousands to millions of copies of a targeted sequence of DNA Bartlett and Stirling, 2003). PCR is commonly used in medical and biological research labs for a variety of applications, including DNA cloning, DNA-based phylogeny or functional gene analysis, diagnosis of hereditary diseases, and the detection and diagnosis of infectious diseases.

The method relies on thermal cycling, repeated cycles of heating and cooling of the reaction mixture to melt DNA and enzymatic replication of the DNA. Primers (short DNA fragments) containing sequences complementary to the targeted region and DNA polymerase are used to amplify DNA, generating targeted fragments. DNA is exponentially amplified (Figure A.1). Repeated cycles include denaturing, annealing and extension. During denaturing temperatures of 98°C and higher are used to denature double-stranded DNA fragments into single strands. In annealing, primers are annealed to single-stranded DNA, and A, C, G, and T nucleotides are added by DNA polymerase (Taq) during extension.

Next-Generation Sequencing

Studies have revealed that more than 99% of the microorganisms (bacteria, archaea and fungi) from environmental samples remain 'unculturable' in the laboratory. Due to this, determining the roles and functions these organisms influence in their natural settings nearly impossible (Sharma et al., 2005). Our inability to cultivate these organisms in the laboratory setting has led to the development of culture-independent techniques, based on DNA sequencing technologies or next-generation sequencing technologies (NGS). Advances in sequencing technologies have grown in leaps and bounds in the last two decades, with pyrosequencing (Roche 454) becoming one of the most powerful tools available for exploring microorganisms from various sampling environments. NGS offers an enormous volume of data cheaply, ranging from one million to one billion short reads per instrument run (Metzker 2010). For the past few years pyrosequencing has been used extensively for obtaining sequence data; however, newer technologies are being developed allowing for even more sequencing reads to be obtained at a

lower cost. Advances in polymerase chain reaction (PCR) based NGS platforms include Roche 454 (pyrosequencing), Illumina, AB SOLiD and Ion Torrent, while single molecule sequencing (SMS) platforms include Helicos HeliScope and Pacific Biosciences system. Each sequencing platform has advantages and disadvantages that must be consider and include read length, maximum number of reads per equipment run, output sequencing data per run, run time and cost (Shorkralla et al., 2012). Table A.1 shows a comparison of current NGS technologies available

With all the data produced from NGS, data analysis is dependent on template preparation, sequence alignment and assembly methods. Robust methods are needed during template preparation to ensure a representative, non-biased source of nucleic acids from the sample. From almost any sample the total DNA can be extracted and targeted gene regions can be amplified and sequenced to determine microbial community structure; or DNA from a single sample can be fragmented and sequenced for genome determination. Depending on the biological applications, sequences or reads obtained from NGS can be aligned to known sequences and used to determine single nucleotide variations in highly related genomes of species of interest. Or sequencing reads obtained of targeted gene regions of the total DNA in a sample allows for clustering based on sequence similarity resulting in taxonomic identification of community members in the sample (Yin et al., 2010). Analysis of sequencing data is dependent on bioinformatics, which utilizes various statistical and computer algorithms techniques to make biological inferences about the sequences obtained. A multitude of programs and open-source pipelines are available to use for sequence analysis, with each having advantages and disadvantages for the biological questions being asked.

As mentioned above, pyrosequencing has become a very powerful tool in obtaining a vast amount of data for a small price. Pyrosequencing is based on the "sequencing by synthesis" principle, in which the complementary strand of single stranded DNA is enzymatically synthesized causing a fluorometric reaction and quantified by a detector. Over 400-600 base pairs can be sequenced in a single pyrosequencing run in 12-14 hours, compared to the "end termination" method used in Sanger sequencing, which is much more time consuming and costly (Metzker 2010). Using emulsion PCR, single strands of DNA are attached to capture beads, creating micro-reactors (Figure A.2). The complementary strands of ssDNA is synthesized when solutions of A, C, G, or T or flushed across the microtiter plate containing a single bead in each well (Figure A.2). As nucleotides are incorporated, a chemical reaction generates light, which is

detected and recorded in a flowgram (Figure A.2). Flowgram data is converted to nucleotide sequences and used for analysis. Pyrosequencing allows for the identification of fungal and bacterial strains to the family and genus level, depending on database used. Changes and shifts in the communities present in different biological samples, like stored biomass for lignocellulosic ethanol production, can be detected through pyrosequencing.

References

Bartlett, J.M.S., Stirling, D., 2003. A short history of the polymerase chain reaction. PCR Protocols. 226, 3-6.

Metzker, M.L., 2010. Sequencing technologies- the next generation. Nat. Rev. Genet. 11, 31-46.

Sharma, R., Ranjan, R., Kapardar, R.K., Grover, A., 2005. 'Unculturable' bacterial diversity: an untapped resource. Curr. Sci. India, 89, 72-77.

Shokralla, S., Spall, J.L., Gibson, J.F., Hajibabaei, M., 2012. Next-generation sequencing technologies for environmental DNA research. Mol. Ecol. 21, 1794-1805.

Yin, C., Jones, K.L., Peterson, D.E., Garrett, K.A., Hulbert, S.H., Paulitz, T.C., 2010. Members of soil bacterial communities sensitive to tillage and crop rotation. Soil Biol. Biochem. 42, 2111-2118.

Figures and Tables

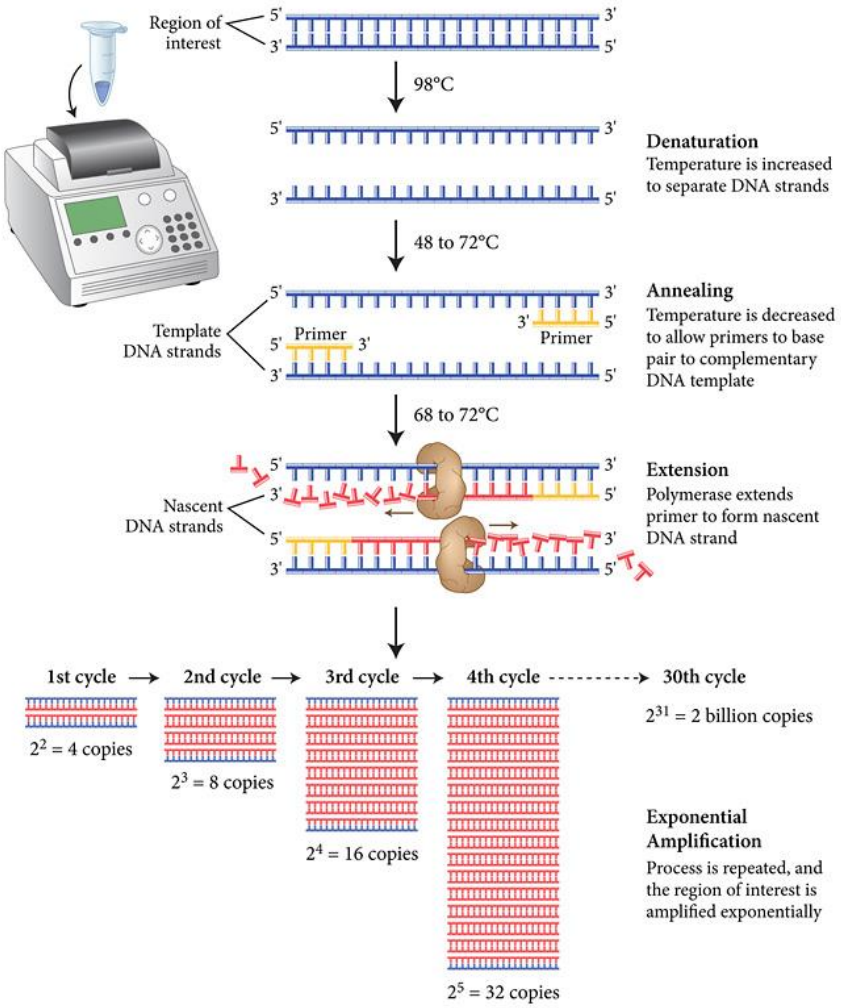
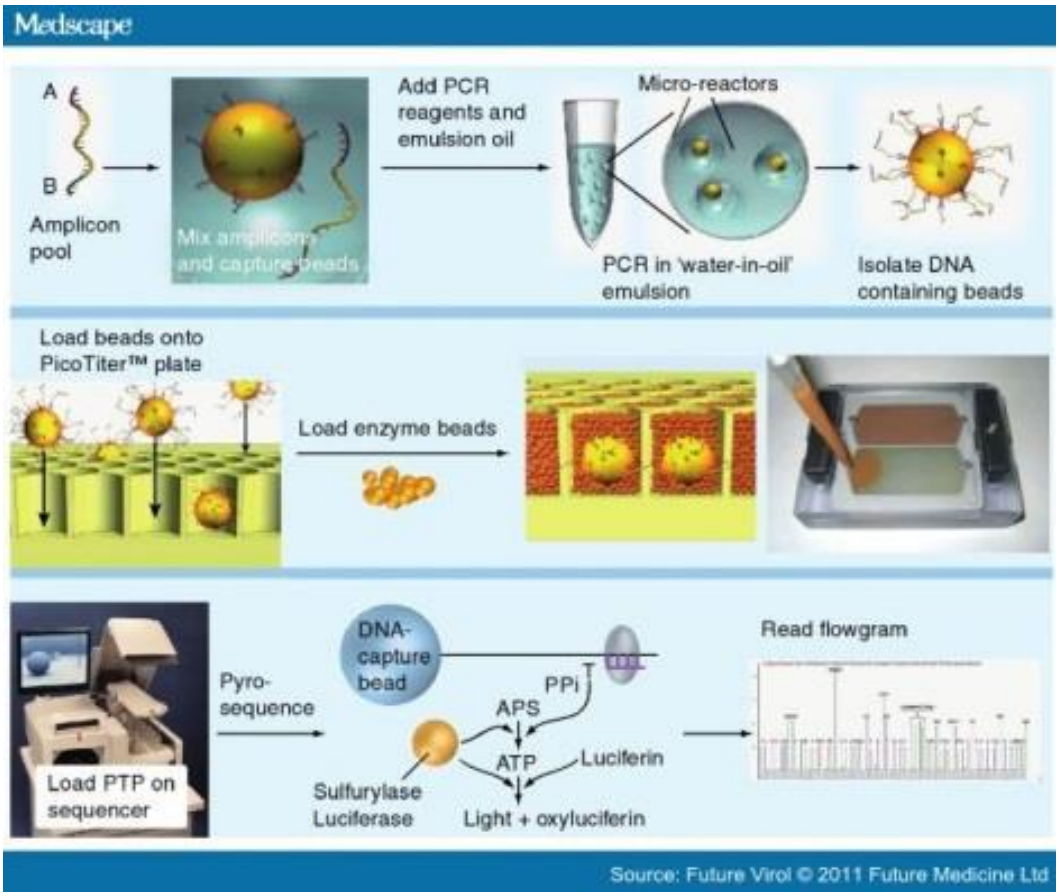


Figure A.1 Polymerase Chain Reaction (PCR) used for the amplification of specific regions within DNA

Table A.1 Comparison of different next-generation sequencing technologies, including read lengths, maximum number of reads per run, sequencing output per run and average run time.

Category	Platform	Read length (bp)	Max. number of reads/run	Sequencing output/run	Run time
PCR-based NGS technologies	Roche 454 GS FLX	400–500	1×10^6	≤500 Mb	10 h
	Roche 454 GS FLX+	600–800	1×10^6	≤700 Mb	23 h
	Roche 454 GS Junior	400–450	1×10^5	~35 Mb	10 h
	Illumina HiSeq 2000	100–200	6×10^9	≤540–600 Gb	11 d
	Illumina HiSeq 1000	100–200	3×10^9	≤270–300 Gb	8.5 d
	Illumina GAIIx	50–75	6.4×10^8	≤95 Gb	7.5–14.5 d
	Illumina MiSeq	100–150	7×10^6	≤1–2 Gb	19–27 h
	AB SOLiD 5500 system	35–75	2.4×10^9	~100 Gb	4 d
	AB SOLiD 5500 xl system	35–75	6×10^9	~250 Gb	7–8 d
	Ion Torrent -314 chip	100–200	1×10^6	≥10 Mb	3.5 h
	Ion Torrent -316 chip	100–200	6×10^6	≥100 Mb	4.7 h
	Ion Torrent -318 chip	100–200	11×10^6	≥1 Gb	5.5 h
SMS technologies	Helicos HeliScope	30–35	1×10^9	~20–28 Gb	≤1 d
	Pacific Biosciences system	≥1500	50×10^3	~60–75 Mb	0.5 h

Figure A.2 Overview of 454-pyrosequencing using emulsion PCR for sequencing by synthesis.



Appendix B - Primers and Barcodes

Below are the forward and reverse primers used for amplicon library production for LSU, ITS and 27F. Each amplicon library utilized barcodes for sample identification and are also included below. For 454-pyrosequencing the A- and B-linkers were used. A-linker: 5' – CCATCTCATCCCTGCGTGTCTCCGACTCAG – 3'; B-linker: 5' – CCTATCCCCTGTGTGCCTTGGCAGTCTCAG – 3'

LSU Primers and Barcodes

LROR (Forward): 5' – CCGCTGAACTTAAGCATATCAATA – 3'; LR3 (Reverse): 5' – CCGTGTTTCAAGACGGG – 3'

Table B.1 LSU Barcodes used in amplicon library production (Chapter 4).

1	ACGAGTGCGT	25	TCGATAGTGA	49	CAGTACTGCG	73	TCGTCGCTCG
2	ACGCTCGACA	26	TCGCTGCGTA	50	CGACAGCGAG	74	CGACGTGACT
3	AGACGCACTC	27	TCTGACGTCA	51	CGATCTGTCG	75	TACAGATCGT
4	AGCACTGTAG	28	TGAGTCAGTA	52	CGCGTGCTAG	76	TCGATCACGT
5	ATCAGACACG	29	TGTAGTGTGA	53	CGCTCGAGTG	77	TACGAGTATG
6	ATATCGCGAG	30	TGTCACACGA	54	CGTGATGACG	78	TGTACTACTC
7	CGTGTCTCTA	31	TGTCGTCGCA	55	CTATGTACAG	79	CATAGTAGTG
8	CTCGCGTGTC	32	ACACATACGC	56	CTCGATATAG	80	TCACGTACTA
9	ACGCGAGTAT	33	CATACTCTAC	57	CTCGCACGCG	81	ACATACGCGT
10	ACTACTATGT	34	CGACACTATC	58	CTGCGTCACG	82	TACACACACT
11	ACTGTACAGT	35	CGAGACGCGC	59	CTGTGCGTCG	83	TACGCTGTCT
12	AGACTATACT	36	CGTATGCGAC	60	TAGCATACTG	84	TCGCACTAGT
13	AGCGTCGTCT	37	AGCTATCGCG	61	TATACATGTG	85	TACTCTCGTG
14	AGTACGCTAT	38	AGTCTGACTG	62	TATCACTCAG	86	ACGACTACAG
15	ATAGAGTACT	39	AGTGAGCTCG	63	TATCTGATAG	87	CGAGAGATAC
16	CACGCTACGT	40	ATAGCTCTCG	64	TCGTGACATG	88	CGTCTAGTAC
17	AGCTCACGTA	41	ATCACGTGCG	65	TCTGATCGAG	89	CAGTAGACGT
18	AGTATACATA	42	ATCGTAGCAG	66	TGACATCTCG	90	TACACGTGAT
19	AGTCGAGAGA	43	ATCGTCTGTG	67	TGAGCTAGAG	91	TAGTGTAGAT
20	AGTGCTACGA	44	ATGTACGATG	68	TGATAGAGCG	92	TCTAGCGACT
21	CGATCGTATA	45	ATGTGTCTAG	69	TGCGTGTGCG	93	TAGAGACGAG
22	CGCAGTACGA	46	CACACGATAG	70	TGCTAGTCAG	94	CGTAGACTAG
23	CGCGTATACA	47	CACTCGCACG	71	TGTATCACAG	95	ATACGACGTA
24	CGTACAGTCA	48	CAGACGTCTG	72	TGTGCGCGTG	96	TCTACGTAGC

ITS Primers and Barcodes

ITS1f (Forward): 5’ – CTTGGTCATTTAGAGGAAGTAA – 3’; ITS4 (Reverse): 5’ – TCCTCCGCTTATTGATATGC – 3’

Table B.2 ITS barcodes used in amplicon library production (Chapter 5).

1	ACGAGTGCGT	25	ACGCGAGTAT	49	AGCTCACGTA	73	ACAGTCGTGC
2	ACGCTCGACA	26	ACTACTATGT	50	AGTATACATA	74	ACATGACGAC
3	AGACGCACTC	27	ACTGTACAGT	51	AGTCGAGAGA	75	ACGACAGCTC
4	AGCACTGTAG	28	AGACTATACT	52	AGTGCTACGA	76	ACGTCTCATC
5	ATCAGACACG	29	AGCGTCGTCT	53	CGATCGTATA	77	ACTCATCTAC
6	ATATCGCGAG	30	AGTACGCTAT	54	CGCAGTACGA	78	ACTCGCGCAC
7	CGTGTCTCTA	31	ATAGAGTACT	55	CGCGTATACA	79	AGAGCGTCAC
8	CTCGCGTGTC	32	CACGCTACGT	56	CGTACAGTCA	80	AGCGACTAGC
9	TCTCTATGCG	33	CAGTAGACGT	57	CGTACTCAGA	81	AGTAGTGATC
10	TGATACGTCT	34	CGACGTGACT	58	CTACGCTCTA	82	AGTGACACAC
11	CATAGTAGTG	35	TACACACACT	59	CTATAGCGTA	83	AGTGTATGTC
12	CGAGAGATAC	36	TACACGTGAT	60	TACGTCATCA	84	ATAGATAGAC
13	ATACGACGTA	37	TACAGATCGT	61	TAGTCGCATA	85	ATATAGTCGC
14	TCACGTACTA	38	TACGCTGTCT	62	TATATATACA	86	ATCTACTGAC
15	CGTCTAGTAC	39	TAGTGTAGAT	63	TATGCTAGTA	87	CACGTAGATC
16	TCTACGTAGC	40	TCGATCACGT	64	TCACGCGAGA	88	CACGTGTCGC
17	TGTACTACTC	41	TCGCACTAGT	65	TCGATAGTGA	89	CATACTCTAC
18	ACGACTACAG	42	TCTAGCGACT	66	TCGCTGCGTA	90	CGACACTATC
19	CGTAGACTAG	43	TCTATACTAT	67	TCTGACGTCA	91	CGAGACGCGC
20	TACGAGTATG	44	TGACGTATGT	68	TGAGTCAGTA	92	CGTATGCGAC
21	TACTCTCGTG	45	TGTGAGTAGT	69	TGTAGTGTGA	93	CGTCGATCTC
22	TAGAGACGAG	46	ACAGTATATA	70	TGTCACACGA	94	CTACGACTGC
23	TCGTCGCTCG	47	ACGCGATCGA	71	TGTCGTCGCA	95	CTAGTCACTC
24	ACATACGCGT	48	ACTAGCAGTA	72	ACACATACGC	96	CTCTACGCTC

27F Primers and Barcodes

27F (Forward): 5’ – AGAGTTTGATCCTGGCTCAG – 3’; 338R (Reverse): 5’ – CATGCTGCCTCCCGTAGGAGT – 3’

Table B.3 27F barcodes used in amplicon library production (Chapter 5).

1	ACGAGTGCGT	25	TCGATAGTGA	49	CAGTACTGCG	73	ACATACGCGT
2	ACGCTCGACA	26	TCGCTGCGTA	50	CGACAGCGAG	74	ACGACTACAG
3	AGACGCACTC	27	TCTGACGTCA	51	CGATCTGTCG	75	ATACGACGTA
4	AGCACTGTAG	28	TGAGTCAGTA	52	CGCGTGCTAG	76	CAGTAGACGT

5	ATCAGACACG	29	TGTAGTGTGA	53	CGCTCGAGTG	77	CATAGTAGTG
6	ATATCGCGAG	30	TGTCACACGA	54	CGTGATGACG	78	CGACGTGACT
7	CGTGTCTCTA	31	TGTCGTCGCA	55	CTATGTACAG	79	CGAGAGATAC
8	CTCGCGTGTC	32	ACACATACGC	56	CTCGATATAG	80	CGTAGACTAG
9	ACGCGAGTAT	33	CATACTCTAC	57	CTCGCACGCG	81	CGTCTAGTAC
10	ACTACTATGT	34	CGACACTATC	58	CTGCGTCACG	82	TACACACACT
11	ACTGTACAGT	35	CGAGACGCGC	59	CTGTGCGTCG	83	TACACGTGAT
12	AGACTATACT	36	CGTATGCGAC	60	TAGCATACTG	84	TACAGATCGT
13	AGCGTCGTCT	37	AGCTATCGCG	61	TATACATGTG	85	TACGAGTATG
14	AGTACGCTAT	38	AGTCTGACTG	62	TATCACTCAG	86	TACGCTGTCT
15	ATAGAGTACT	39	AGTGAGCTCG	63	TATCTGATAG	87	TACTCTCGTG
16	CACGCTACGT	40	ATAGCTCTCG	64	TCGTGACATG	88	TAGAGACGAG
17	AGCTCACGTA	41	ATCACGTGCG	65	TCTGATCGAG	89	TAGTGTAGAT
18	AGTATACATA	42	ATCGTAGCAG	66	TGACATCTCG	90	TCACGTACTA
19	AGTCGAGAGA	43	ATCGTCTGTG	67	TGAGCTAGAG	91	TCGATCACGT
20	AGTGCTACGA	44	ATGTACGATG	68	TGATAGAGCG	92	TCGCACTAGT
21	CGATCGTATA	45	ATGTGTCTAG	69	TGCGTGTGCG	93	TCGTCGCTCG
22	CGCAGTACGA	46	CACACGATAG	70	TGCTAGTCAG	94	TCTACGTAGC
23	CGCGTATACA	47	CACTCGCACG	71	TGTATCACAG	95	TCTAGCGACT
24	CGTACAGTCA	48	CAGACGTCTG	72	TGTGCGCGTG	96	TGTACTACTC