Pure culture and metagenomic approaches to investigate cellulose and xylan degradation©

Sita Ng

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Master of Science

in

The Faculty of Science

Applied Bioscience

University of Ontario Institute of Technology

April 2010

Certificate of Approval

Abstract

Lignocellulose is composed of lignin, hemicellulose, and cellulose. Lignocellulose waste is a sustainable and renewable resource available for use in biotechnological applications. Efficient enzyme production and enzymes with high catalytic activity are needed for the use of lignocellulose. The study of cellulases and xylanases that degrade cellulose and xylan into constituent monosaccharides is required to advance industrial application of these enzymes. The use of a traditional pure culture approach to discover and characterize cellulases and xylanases from novel actinomycete isolates and the use of metagenomics to uncover previously unidentified cellulase genes was undertaken. Actinomycetes were cultivated from soil samples and the isolate with the best cellulase and xylanase activity was subjected to strain improvement through protoplast fusion. Enhanced enzymatic activity was found in one fusant. Differential release of sugars from xylan was observed through gas chromatographic analysis between the parental and fusant cultures. Genome shuffling was observed in 16S rRNA genes after protoplast fusion. Finally, one putative endo-β-1,4-glucanase was discovered in a metagenomic library created from cellulose-enriched potting soil.

Keywords: lignocellulose, cellulase, xylanase, metagenomics, actinomycete

Acknowledgements

I would like to thank Dr. Janice Strap for her support, encouragement and guidance. Also, I would like to thank my committee members Dr. Dario Bonetta, Dr. Julia Green-Johnson and Dr. Ayush Kumar for their helpful insight, constructive criticism and advice.

My gratitude also extends to present and past members of Dr. Strap's lab and have in some way made a contribution to this thesis: Kelsey Dermott, Aakash Modi, Andrew Latos, Sabrina Keates, Li He, Kathy Tang, Leena Chellakudam, Vanessa Serapio, Lisa Wagar, Meghan Perry, Jennifer Lavis, Ted Paulsen, Karen Lilwall, Evelyn Cardinez, Sandy Clarke and Mateusz Wojtas.

I would like to acknowledge my fellow colleagues for their support and understanding throughout the process. My deepest gratitude goes to Laura Benninger, Mathumai Ganeshakumar, Salma Siddiqui and David Brown. Without their continual support and encouragement, I would have never been able to complete this project.

Finally, I would like to thank my two families for believing in me and for their unconditional support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

(v/v)	volume per volume
(w/v)	weight per volume
Amp	Ampicillin
APS	Ammonium persulfate
BLAST	Basic Local Alignment Search Tool
bp	Base pair(s)
C-CRA	Cellulose Congo Red Agar
CFU	Colony Forming Unit
CMC	Carboxymethyl Cellulose
CMC-CRA	Carboxymethyl Cellulose Congo Red Agar
CTAB	Cetyltrimethylammonium bromide
DCM	Dichloromethane
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotide triphosphates
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium Bromide
GC	Gas chromatography
gDNA	Genomic deoxyribonucleic acid
IPTG	Isopropyl β-D-thiogalactopyranoside
kb	Kilobase
LB	Luria Bertani
Mb	Megabases
mL	Millilitre
MW	Molecular weight
PBF	Bulk rhizosphere on PDA
P Buffer	Protoplast buffer
PCB	Conservation bulk on PDA
PCR	Polymerase Chain Reaction
PDA	Potato Dextrose Agar
PEG	Polyethylene glycol
PRF	Forest rhizosphere on PDA
PSC	Potting Soil Cellulose
PSP	Potting Soil on PDA
PSY	Potting Soil on YDA
R2YE	Sucrose Yeast Extract Medium
rDNA	Ribosomal deoxyribonucleic acid
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
rRNA	Ribosomal Ribonucleic acid
sp.	Single species
spp.	Multiple species
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA Buffer
TEMED	N,N,N',N'-tetramethylethylenediamine
TES	N-tris(hydroxymethyl) methyl-2-aminoethane-sulfonic acid

Tris	Tris(hydroxymethyl)aminomethane
WBF	Bulk rhizosphere on WYE
WCB	Conservation bulk on WYE
WCR	Conservation rhizosphere on WYE
WRF	Forest rhizosphere on WYE
WYE	Water Yeast Extract
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
YDA	Yeast Dextrose Agar
YEME	Yeast Extract-Malt Extract Medium
YRF	Forest rhizosphere on YDA
µg/µL	Microgram per microlitre
µL	Microlitre

I. INTRODUCTION

Bioethanol production from plant matter can be an alternative sustainable energy. Lignocellulose is the plant's main structural component composed of three major polymers: lignin, hemicellulose and cellulose. Cellulose and xylan, the backbone of hemicellulose, make up the plant cell wall (Talmadge et al., 1973; Keegstra et al., 1973) and thus, are bountiful renewable resources. The degradation of the complex polymer, cellulose and xylan of lignocellulose, has an application in bioethanol production through fermentation of sugars such as xylose, or glucose formed as by-products through the hydrolysis of xylan and cellulose, respectively. One limitation of bioethanol production through this means is the expense involved in the production of enzymes such as xylanases and cellulases (Lynd et al., 1991). For bioethanol production to be commercially viable, the discovery of cellulases and xylanases with greater catalytic activity to degrade xylan and cellulose is necessary (Hill et al., 2006). Soil microorganisms produce enzymes that degrade lignocellulose for use as an energy source for survival. This thesis research focused on the identification of soil actinomycetes that produced potent cellulases and xylanases that degrade cellulose and xylan, respectively. To accomplish this and facilitate the discovery of these enzymes, the project was divided into two related research directions:

1. Identification of actinomycete species cultivated from different soil types that degrade cellulose and xylan
2. Mine the soil metagenome for cellulase genes.

I.1. Lignocellulose

Lignocellulose is composed of three components: lignin, hemicellulose and cellulose. Lignin is a highly-branched random phenylpropanoid polymer formed by free-radical condensation of aromatic alcohols (Bouxin et al., 2010; Brown, 1969; Gang et al., 1999; Higuchi, 1990). Hemicellulose is covalently connected to lignin (Jung, 1989). The primary heteropolymers of hemicelluloses are xylan, mannan, galactan, and arabinan (Li et al., 2000). The primary sugar monomers of hemicelluloses that form these polymers are D-xylose, D-mannose, D-galactose, and L-arabinose (Li et al., 2000). Hemicellulose is connected to cellulose microfibrils through hydrogen bonding (Bauer et al., 1973). The interactions between recalcitrant lignin and the highly-ordered crystalline structure of cellulose and hemicellulose create a barrier protecting lignocellulose from degradation. The complexity of lignocellulose requires enzymes with various substrate specificities to completely degrade the recalcitrant lignocellulose to its monomers of which some are sugars that can be fermented to ethanol.

I.2. Biofuels

Alternative sustainable energy is sought after due to the high demand for energy and the limited amount of current energy sources. A continual increase in consumption of fossil fuels while resources are depleting has driven the demand for alternative fuel sources such as ethanol (Goldemberg, 2007). An advantage of using ethanol is a decrease in carbon dioxide emissions (Hill et al., 2006). Lignocellulose biomass can be used as a resource for the production of biofuels, such as bioethanol. Glucose from the degradation of cellulose can be fermented to produce ethanol (Delgenes et al., 1996). Likewise, pentoses from the degradation of xylan can also be fermented to ethanol (Hahn-Hägerdal et al., 1994).

Lignocellulose, the major structural component of plant material, can be a cheap, abundant and renewable energy source obtained from agricultural waste (Hill et al., 2006; Belkacemi et al., 2002), municipal solid waste (Chester & Martin, 2009; Li et al., 2007), as well as waste from forestry and pulp and paper industries (Lynd et al., 1991; Goldemberg, 2007). In Brazil and the United States of America (U.S.A.), fuel ethanol is already being produced from corn and sugar cane (Goldemberg, 2007). Presently, cornstarch has been used for the production of ethanol but this requires a large amount of agricultural land that is normally used for food production (Hill et al., 2006). The use of lignocellulose wastes is a more practical solution than corn crops, as it does not require the use of valuable land resources (Hill et al., 2006). Furthermore, lignocellulose residues resulting from crop harvesting are disposed of through burning (Levine, 2000; Crutzen & Andreae, 1990). However the burning causes air pollution and will damage the soil by reducing nitrogen and water retention (Levine, 2000; Crutzen & Andreae, 1990). To avoid this, the lignocellulose residues can be used for bioethanol production instead (Farrell et al., 2006).

Ethanol production costs are high compared to fossil fuels and therefore it has not replaced fossil fuel usage (Goldemberg, 2007). The use of lignocellulosic biomass is impaired in bioethanol production because the enzymes and chemicals needed for this conversion are expensive (Lynd et al., 1991). Also, due to the recalcitrant nature of lignocellulose, xylanases that degrade xylan efficiently need to be discovered to allow cellulases access to cellulose so that it may be degraded into constituent sugars to increase the efficient use of lignocellulose for bioethanol production. Otherwise, ethanol will not be a cost effective sustainable alternative. Bioethanol production from abundant

and renewable lignocellulose can be a good alternative to fossil fuels with potentially lowered production costs.

I.3. Hemicellulose

Xylan is a plant cell wall polysaccharide that is a major component of hemicellulose. Hemicellulose is a highly branched heteropolymer containing sugar residues such as hexoses (D-galactose, L-galactose, D-mannose, L-rhamnose, L-fucose), pentoses (D-xylose, L-arabinose), and uronic acids (D-glucuronic acid) (Li et al., 2000). Xylan is found between lignin and cellulose and is thought to be important in fibre cohesion and for the integrity of the cell wall (Keegstra et al., 1973; Talmadge et al., 1973; Iwamoto et al., 2008). Xylan helps to protect cellulose from degradation through its covalent interactions with lignin and non-covalent interactions with cellulose (Bauer et al., 1973). Xylan is formed by a xylose backbone linked by ß-1,4-xylosidic bonds with the constituents of arabinosyl, glucoronosyl and acetyl residues (Bauer et al., 1973; Li et al., 2000; Keegstra et al., 1973; Talmadge et al., 1973). The variability of xylan is due to the diversity of the number of neutral or uronic monosaccharide subunits or short oligosaccharide chains that are linked to a ß-1,4-linked D-xylopyranosyl backbone (Bauer et al., 1973). The composition of hemicellulose can vary depending on the plant resulting in complexity in the structure of hemicellulose.

I.4. Hemicellulases

Various xylanases with different substrate specificity are required to completely degrade xylan depending on the structure of hemicellulose. Hemicellulases can either be glycoside hydrolases or carbohydrate esterases, which hydrolyze the glycosidic bonds, and the ester linkages of acetate and ferulic acid respectively (Coutinho & Henrissat, 1999; Henrissat & Bairoch, 1996). Hemicellulases that are glycoside hydrolases include:

β-1,4-endoxylanase (Biely et al., 1997), β-xylosidase (Deshpande et al., 1986), α-L-arabinofuranosidase (Margolles & Reyes-Gavilán, 2003), α-D-glucuronidase (de Vries et al., 1998) and β-mannanases (Stoll et al., 1999). Endo-1,4-β-xylanases randomly cleave the ß-1,4 backbone (Biely et al., 1997). β-xylosidase hydrolyzes xylobiose (Deshpande et al., 1986). Enzymes, such as α-L-arabinofuranosidase (Margolles & Reyes-Gavilán, 2003; Matte & Forsberg, 1992), α-glucuronidase (de Vries et al., 1998), and β-mannanases (Stoll et al., 1999) cleave the side groups attached to the backbone of xylan to allow β-1,4-endoxylanases access to the backbone of xylan to completely degrade xylan (Biely et al., 1997). The α-L-arabinofuranosyl groups of arabinans, arabinoxylans, and arabinogalactan present in xylan are terminally cleaved at the non-reducing end by α-L-arabinofuranosidase (Margolles & Reyes-Gavilán, 2003; Matte & Forsberg, 1992). α-glucuronidase cleaves the α-1,2-glycosidic linkages between xylose and D-glucuronic acid or the 4-*O*-methyl ether linkage (de Vries et al., 1998). β-mannanases release short β-1,4-manno-oligomers from mannan (Stoll et al., 1999).

Hemicellulases that are carbohydrate esterases include: acetyl xylan esterase (Biely et al., 1986) and feruloyl esterases (Blum et al., 2000). Acetyl xylan esterase is required to completely liberate the acetic acids of glucuronoxylans (Biely et al., 1986). The ester bond between arabinose and ferulic acid is hydrolyzed by feruloyl esterases (Blum et al., 2000). To access the backbone of xylan, the side groups need to be removed to allow β-1,4-endoxylanase to degrade xylan to release xylose for fermentation to ethanol.

I.5. Cellulose

Cellulose is a structural polysaccharide found in plant cell walls and is composed of cellulose chains that form intramolecular and intermolecular hydrogen bonds to create

inflexible, insoluble microfibrils (Chanzy et al., 1979; Nishiyama et al., 2002). Cellobiose is two glucose subunits linked together through ß-1,4 bonds to form the repeating unit comprising cellulose chains (Kolpak & Blackwell, 1976) (Figure I.6.1.). The cellulose chains form organized microfibrils and run parallel to each other (Kolpak & Blackwell, 1976). This linear polysaccharide is formed from repeating glucose units linked by ß-1,4-glucosidic bonds containing crystalline regions and amorphous regions, where the microfibril chains are disordered compared to crystalline regions (Nieduszynski & Preston, 1970). The structural integrity of cellulose is created by the organized crystalline regions because they are more resistant to hydrolysis than the amorphous regions (Kolpak & Blackwell, 1976). Therefore, the more crystalline regions cellulose has, the more difficult it is to degrade. Microorganisms that produce cellulases to effectively degrade cellulose completely to glucose are required. The glucose can then be fermented to ethanol.

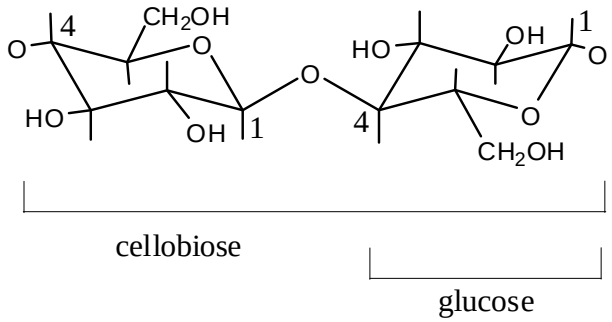


Figure I.6.1. Cellobiose and glucose (Adapted from Béguin and Aubert (1994)).

I.6. Cellulases

Cellulases exhibit three types of activity. The three known modes of action of cellulases include: endoglucanase, exoglucanase, and β-glucosidase (Davies & Henrissat, 1995). Endoglucanases or endo-1,4-ß-glucanases hydrolytically cleave the intramolecular ß-1,4-glucosidic bonds of cellulose (Davies & Henrissat, 1995; Liu et al., 2009; Okazaki

& Moo-Young, 1978) (Figure I.7.1.). New chain ends are formed from this random cleavage and thus, allowing further degradation by subsequent cellulases (Okazaki & Moo-Young, 1978; Liu et al., 2009). Either cellobiose or glucose is formed from the cleavage of cellulose chains at the ends (reducing and non-reducing) by exoglucanases (Davies & Henrissat, 1995; Divne et al., 1998) (Figure I.7.1.). Lastly, β-glucosidases cleave cellobiose (Figure I.7.1.) into glucose residues and this prevents the accumulation of cellobiose (Davies & Henrissat, 1995; Parry et al., 2001; Umezurike, 1979). Accumulation of cellobiose leads to the inhibition of exoglucanases (Okazaki & Moo-Young, 1978). Hydrolysis of cellulose is dependent on the rate-limiting step of the depolymerisation of cellulose by endoglucanases and exoglucanases. To completely degrade cellulose to obtain glucose for fermentation into ethanol, a concerted attack of cellulose by all three modes of action is required.

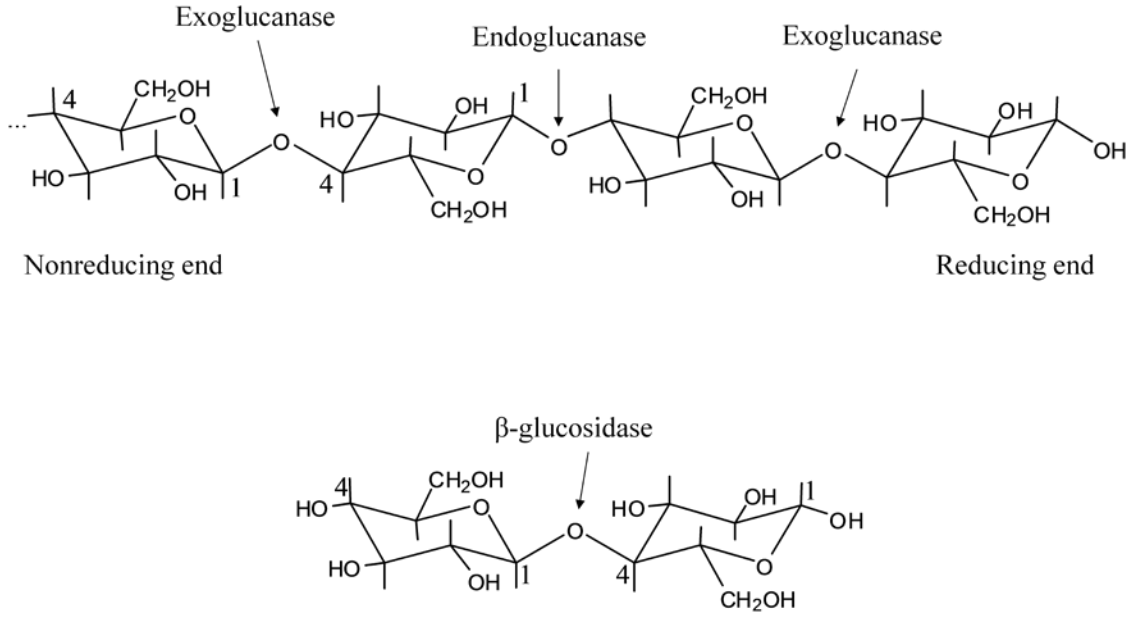


Figure I.7.1. Cleavage sites of cellulose by cellulases: exoglucanse, endoglucanase, and β-glucosidase (Davies & Henrissat, 1995).

Part 1: Pure culture analysis of cellulose and xylan degradation

II. ABSTRACT

For bioethanol production from lignocellulose to be economically viable, a decrease in enzyme cost is required (Lynd et al., 1991). Novel enzymes can be discovered from environmental isolates that are known to produce a diversity of enzymes. Actinomycete isolates were cultivated from potting soil, UOIT forest rhizosphere and bulk soil, and UOIT conservation area rhizosphere and bulk soils. Soil was used because soil contains lignocellulose-degrading actinomycetes. Isolates were screened for cellulase and xylanases activities and a list of isolates producing broad spectrum cellulase and xylanase activity was compiled. The top cultivated strain with the highest cellulase and xylanase activities was subjected to strain improvement through protoplast fusion. Greater than 30% of the fusants produced from three fusion events had improved cellulase and xylanase activities as determined by agar diffusion method. Xylan degradation products between a set of fusants and parents were analyzed by gas chromatography. Fusants and parents digested xylan differently resulting in 10-55 nmol/mg of xylose present in the inoculated cultures over different time points. Less than 3.4 nmol/mg of fucose was liberated from xylan by three fusants, FA1-14, FA9-13 and FA9-14. Trace amounts of glucose was observed in the inoculated samples and uninoculated controls between 48-192 and 432-576 hours. Rhamnose was only observed at 576 hours in the uninoculated control but was not observed at any other time point. Less than 3.5 nmol/mg of arabinose was observed in uninoculated controls, and inoculated cultures of FA1-14, FA2-4 and parental WBF90B. The 16S rRNA gene of one fusant was sequenced and appears to be a novel strain related to other *Streptomyces* spp. The sequenced fusant had a similar 16S rDNA restriction fragment length polymorphism (RFLP) banding pattern to one of its parents. Finally, the 16S rRNA genes of fusants

were compared to their parents by RFLP analysis to determine if any genome shuffling occurred in the conserved 16S rRNA gene. Genome shuffling was observed to occur in the 16S rRNA gene of one fusant.

III. INTRODUCTION

III.1. Actinomycetes

Actinomycetes are filamentous, Gram-positive bacteria of the order *Actinomycetales* that produce an array of primary and secondary metabolites (Embley & Stackebrandt, 1994; Fiedler et al., 2005; Fiedler, 1993). Extracellular hydrolytic enzymes are produced by these organisms, allowing them to compete for nutrients within their natural habitat (Ball & McCarthy, 1988; McCarthy et al., 1985). Actinomycetes can use lignocellulose to support growth as it is the most abundant soil carbon source (Trigo & Ball, 1994). Biomass degrading microorganisms, such as actinomycetes, can work synergistically with other organisms to produce different enzymes for a concerted enzymatic attack to degrade lignocellulose present in soil plant litter. Actinomycetes that produce enzymes to degrade lignocellulose can be cultivated from the soil and used to degrade xylan and cellulose into substituent sugars for bioethanol production as an alternative sustainable energy source.

The genus *Streptomyces* are a group of soil-inhabiting, saprophytic and sporulating bacteria with a high G+C DNA (70%) content (Hopwood, 2006). They are able to colonize plant litter in the soil because of their hyphal growth and release extracellular enzymes to degrade lignocellulose for survival. Cultivated *Streptomyces spp*. have been shown to produce cellulases and xylanases; for instance, *Streptomyces lividans* (Hurtubise et al., 1995), *Streptomyces viridosporus* (Ramachandra et al., 1987) and *Streptomyces thermoviolaceus* produce xylanase (Garg et al., 1996). Secreted extracellular enzymes can be obtained from culture supernatants of cultured novel *Streptomyces* for industrial degradation of lignocellulose for bioethanol production. The degradative enzymes that degrade lignocellulose well can be further improved through

strain improvement techniques (Pettey & Crawford, 1984; Gokhale & Deobagkar, 1989) to meet the needs of the industrial process such as pH and temperature stability.

Streptomyces are ideal candidates for the discovery of cellulases and xylanases and for genetic enhancement of these enzymes to increase their applicability in industrial processes such as bioethanol production. Streptomycetes are an attractive group of organisms for strain improvement and genetic manipulation, including cloning and mutagenesis, (Kieser et al., 2000) as the full genome has been sequenced for *Streptomyces avermitilis* MA-4680 (Ikeda et al., 2003) and *S. coelicolor* A3(2) (Bentley et al., 2002). Therefore there is vast knowledge on the genetics of Streptomycetes. A DNA cloning system for *S. lividans* has been established (Bibb et al., 1980; Kieser et al., 1982) which makes *S. lividans* a suitable heterologous host (Katz et al., 1983; Martinez et al., 2004; Wang et al., 2000) for producing large amounts of extracellular proteins (Strickler et al., 1992). It is possible to clone and overproduce cellulase and xylanase genes from novel *Streptomyces spp.* in cultures of *S. lividans*.

III.2. Strain Improvement: Protoplast fusion

Microbial enzymes are biocatalysts used in industrial applications, for example in bioethanol production xylanases and cellulases are used. For industrial use, enzymes may have to be optimized to be suitable for use in industrial applications such as substrate specificity, temperature stability and catalytic activity to make these enzymes more effective in the process (May et al., 2000; Miyazaki et al., 2006; Song & Rhee, 2000; Hermes et al., 1990; Gokhale & Deobagkar, 1989). Strain improvement involves the modification of genes encoding a property of interest, such as enzymes, through mutagenesis or genome shuffling, and screening for improved properties; the process can be repeated with improved progeny for another round of strain improvement until the

desired outcome is generated for the industrial process (Hermes et al., 1990; Schaeffer et al., 1976; Zhang et al., 2002).

Chemical mutagenesis, a classical strain improvement technique, involves introducing random point mutations with a mutagen. This strategy is tedious because multiple rounds of mutagenesis and selection are required until a desired phenotype is generated (Zhang et al., 2002). Through chemical mutagenesis, introduction of point mutations into different genes encoding lignocellulose degrading enzymes can occur but mutations may be silent, nonsense or point mutations where no improvement occurs. Also, chemical mutagenesis results in deleterious mutations and many mutants need to be screened before a mutant with improvement is found (Zhang et al., 2002).

An alternative strain improvement technique is genome shuffling, or protoplast fusion. In this approach, protoplasts are created by the removal of the cell wall of bacteria then fused together for random homologous recombination of the genome to occur (Hopwood & Wright, 1978). Homologous recombination for multiple genes or operons can yield improved activity in several genes such as those encoding different cellulases or xylanases. Mutants created in this fashion are screened for improved phenotypic traits compared to original strains (Hopwood et al., 1977).

Protoplast fusion has been successfully used to enhance degradation of lignocellulose. For example, cellulase and xylanase activity was improved compared to parental strains in studies where inter-strain genome shuffling between *Cellulomonas* spp. and *Bacillus subtilis* (Gokhale & Deobagkar, 1989), *Penicillium echinulatum* and *Trichoderma harzianum* (Dillon et al., 2008), and intra-strain improvement as in *Trichoderma reesei* (Prabavathy et al., 2006) and *Trichoderma reesei* QM 9414 fusions

(Sandhu & Bawa, 1992). Furthermore, endoglucanase activity was enhanced in the hybrid created either from protoplast fusant of *Streptomyces viridosporus* T7A to itself or *S. setonii* 75Vi2 (Ramachandra et al., 1987). Besides cellulase activity, lignin degradation was enhanced by protoplast fusion of two *Streptomyces* spp. (Pettey & Crawford, 1984). Genome shuffling is therefore effective in enhancing lignocellulose degrading enzymatic activity in various microorganisms.

Besides strain improvement for degradative enzymes, genome shuffling improved thermotolerance, ethanol tolerance and ethanol productivity of *Saccharomyces cerevisiae* (Shi et al., 2009). Genome shuffling is useful in strain improvement for multiple factors and recursive recombination can be completed until a strain is created with the desired properties. A fusant able to degrade cellulose and ferment glucose into ethanol was created from the fusion of *Trichoderma reesei* and *Saccharomyces cerevisiae* (Srinivas et al., 1995). It is conceivable that a microorganism such as an actinomycete that degrades lignocellulose could be fused with a microorganism that can produce ethanol, like *Saccharomyces cerevisiae*, to make a designer strain capable of lignocellulose degradation and fermentation of constituent sugars to produce bioethanol.

III.3. 16S rRNA Genes for Taxonomy

The 16S ribosomal RNA (rRNA) creates the small subunit of the ribosome necessary for protein synthesis in cells. Due to their important role in protein synthesis, 16S rRNA is universal, conserved and existed early on in evolution (Fox et al., 1977).

16S rRNA genes are used as a molecular marker to taxonomically identify prokaryotes (Woese, 1987). 16S rRNA genes evolve at different rates compared to other genes forming conserved nucleotide regions and variable regions (Sogin et al., 1971; Woese et al., 1975). The conserved regions evolve slowly while the variable regions can

evolve much more rapidly. The small subunit (SSU) 16S rRNA gene is present in all cells and most gene variants can be amplified by degenerate primers because the 16S rRNA gene is large enough to contain significant differences in sequence to be used as a marker (Woese, 1987). The variable regions can be used to phylogenetically characterize microorganisms because the differences in the variable regions distinguish one bacterium from another, while the conserved regions are the same (Woese, 1987; Weisburg et al., 1991). A ribosomal database has been created for phylogenetic studies (Olsen et al., 1991; Larsen et al., 1993; Maidak et al., 1996; Cole et al., 2009) allowing characterization of novel microorganisms. Actinomycetes, like many other bacteria, have been characterized by their 16S rRNA gene sequences (Rheims et al., 1996; Niner et al., 1996).

III.4. Part 1 Overview

There is a need to discover lignocellulolytic enzymes from novel isolates that will completely degrade lignocellulose into its constituents for bioethanol production. Cellulases and xylanases that completely degrade cellulose and xylan, respectively, without the need to increase concentrations of expensive enzymes, will make bioethanol production more feasible as an alternative energy source.

Soil actinomycetes produce lignocellulose degrading enzymes for their survival in the environment and are therefore a potential source of industrially useful enzymes. In this thesis work, novel soil actinomycetes that degraded cellulose and xylan well were cultivated and subjected to strain improvement (Figure III.4.1.).

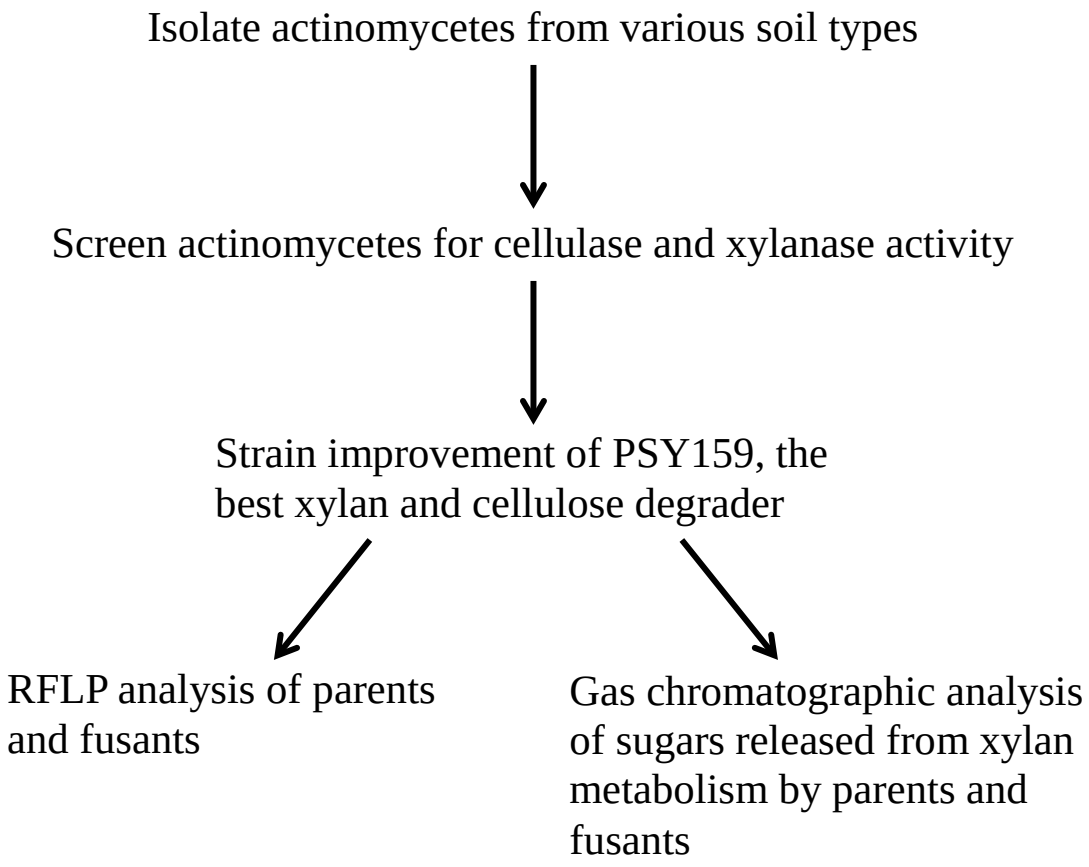


Figure III.4.1. Overview of pure culture analysis of actinomycetes that degrade cellulose and xylan

IV. MATERIALS AND METHODS

IV.1. Materials

All materials, chemicals and antibiotics were purchased from Bioshop, Burlington, ON or Fisher Scientific, Fair Lawn, NJ unless otherwise stated. Garden Club All Purpose Potting Soil was purchased from Canadian Tire (purchased 2007). Potato dextrose agar (PDA) was purchased from Bioshop, Burlington, ON. Agar agar powder food additive was purchased from C.L.T. Intertrade Co. Limited, Bangkok, Thailand. Restriction enzymes were purchased from Fermentas, Burlington, ON and used according to manufacturer's instructions. Congo red, cellulose, carboxymethyl cellulose and birchwood xylan were purchased from Sigma, St. Louis, MO. BeechNut oatmeal was purchased from BeechNut Nutrition Corp. Canajoharie, NY. GoTaq® Hot Start Polymerase and pGEM®-T Easy Vector Systems were purchased from Promega Corporation, Madison, WI. EZ-10 Spin Column Plasmid DNA Minipreps Kit, EZ-10 Spin Column DNA Cleanup Kit and EZ-10 Spin Column PCR Products Purification Kit were purchased from Bio Basic Inc., Markham, ON. For gas chromatography, 2 mL Screw Top 9 mm Wide Opening Vials with 200 µL glass insert with polymer spring and FactorFour™ Column VF-23ms were purchased from Varian Associates Inc., Palo Alto, CA.

IV.2. Bacterial strains

Escherichia coli strain JM109 (*rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17 (rK–,mK+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F´, *tra*D36, *pro*AB, *lac*$I^qZ\Delta M15$]) was used for this study (Messing et al., 1981). Designation for environmental isolates cultivated from various soil samples and on various media is listed in Table IV.2.1.

Table IV.2.1. Soil isolate designations and isolation media

Isolate Designation	**Isolation Medium**
PSY	Potting soil on Yeast Dextrose Agar (YDA)
PSP	Potting soil on Potato Dextrose Agar (PDA)
WRF	Forest rhizosphere on Water Yeast Extract (WYE)
WBF	Bulk rhizosphere on WYE
WCB	Conservation bulk on WYE
WCR	Conservation rhizosphere on WYE
PRF	Forest rhizosphere on PDA
PBF	Bulk rhizosphere on PDA
PCB	Conservation bulk on PDA
PCR	Conservation rhizosphere on PDA
YRF	Forest rhizosphere on YDA

IV.2.1. Maintenance of *Escherichia coli*

Escherichia coli strain JM109 was provided as frozen glycerol stocks (20% v/v) and maintained as such at -20°C. Recombinant *E. coli* were maintained as frozen glycerol stocks (20% v/v) with antibiotic selection at -20°C.

IV.2.2. Maintenance of isolates

All environmental isolates were maintained as frozen glycerol spore stocks (20% v/v) at -20°C. Fusants were maintained on YDA plates at 4°C.

IV.3. Isolation of bacterial strains from potting soil

Bacteria were isolated from soil as follows: 1.02 g of All Purpose Potting Soil was weighed out and placed in a sterile tube containing 9 mL of 0.85% saline solution. The tube was vortexed for approximately 7 minutes. Serial ten-fold dilutions were performed and 100 µL from each serial tenfold dilution was plated to obtain final dilution of 10^{-4}, 10^{-5}, and 10^{-6} onto triplicate onto three plates of Potato Dextrose Agar (PDA) (3.9% (w/v) of PDA) and Yeast Dextrose Agar (YDA) (0.1% (w/v) yeast extract, 0.3% (w/v) dextrose, 1.5% (w/v) agar). Plates were incubated at 30°C for 8 days while observations were taken daily to identify the type of microorganisms present. Actinomycete colonies were identified between 7-10 days and purified. Morphologically unique colonies characteristic of actinomycetes were observed as leathery colonies or with sporulated aerial mycelia but not mucoid like or fungi like (colonies over taking plate with large masses of mycelia and not leathery colonies). Pure cultures were maintained on PDA or YDA. The actinomycete colonies were phenotypically described for cultural and morphological characteristics such as size, elevation, margin, shape, surface appearance of colonies – colony shape, diffusible pigment production and colour,

and spore colour if applicable. Glycerol spore stocks were prepared for strains that sporulated.

IV.4. Spore stock preparation

Into microfuge tubes, 100 µL of 0.85% (w/v) saline solution was dispensed. A colony was excised and placed in the tube and a sterile pestle was used to homogenize the colony. An additional 100 µL of 0.85% saline solution was added and mixed thoroughly. To enhance sporulation, oatmeal agar (6% (w/v) oatmeal, 1.5% (w/v) agar) was used. On two oatmeal agar plates, 100 µL of the homogenate was plated. Plates were incubated at 29°C until isolates were well-sporulated.

Into sterile 15 mL screw-capped tubes, 5 mL of sterile water (18 MΩ) was dispensed aseptically. A sterile spatula was used to scrape spores from a sporulated lawn and was transferred into 15 mL screw-capped tubes with sterile deionized water. Spores were freed from mycelial fragments by vortexing vigorously for 5 minutes. The vortexed spore suspension was carefully poured into a filter unit (composed of cotton in a sterile 15 mL screw-capped tube with a hole punctured at the tip) and allowed to filter by gravity. Sterile, deionized water (18 MΩ) was added to the filter unit to wash through spores. Filtered spores were centrifuged at 1801 *g* for 15 minutes at 4˚C in a Legend RT centrifuge (Sorvall). Supernatant was decanted and discarded. The spore pellet was resuspended in the remaining liquid. An equal volume of 40% (v/v) glycerol was added, vortexed well to mix, and stored at -20°C. These spore stocks were used as inoculum to screen isolates for cellulase and xylanase activity.

IV.5. Degradation assays with actinomycete isolates

To assess the potential of environmentally isolated actinomycetes to degrade cellulose and xylan, *in vitro* degradation assays were performed by culturing

actinomycete isolates on mineral salts agar containing, 0.05% (w/v) cellulose or 0.05% (w/v) carboxymethyl cellulose (CMC) and 0.005% (w/v) congo red (CR) dye. The concentration of xylan used was based on published values (Yang et al., 1995). 5 µL of spores of isolates were spotted onto cellulose-CR or CMC-CR agar plates. Congo red is a dye that binds to ß-1-4-glycosidic linkages of cellulose (Teather & Wood, 1982). As cellulose or CMC is degraded, a clear zone becomes visible around the growth of the isolate due to the release of the congo red dye. CMC congo red agar (0.05% (w/v) of carboxymethyl cellulose, 0.57 mM K_2HPO_4, 1.35 mM KCl, 0.20 mM $MgSO_4$, 0.005% (w/v) yeast extract, 1.5% (w/v) agar) plates and cellulose congo red agar (0.05% (w/v) of cellulose, 0.57 mM K_2HPO_4, 1.35 mM KCl, 0.20 mM $MgSO_4$, 0.005% (w/v) yeast extract, 1.5 % (w/v) agar) plates were used for screening actinomycetes for cellulase activity.

Insoluble xylan from 1% birchwood was used to make defined xylan medium (5.7 mM K_2HPO_4, 17.1 mM NaCl, 15.1 mM $(NH_4)_2SO_4$, 1% (w/v) xylan from birchwood, 20.0 mM $CaCO_3$, 1.8% (w/v) agar) (Yang et al., 1995) to screen for xylanase activity. For this assay, spores were also spotted on xylan agar. Clearance zones indicating degradation were clearly visible on xylan containing media, therefore staining with a dye was unnecessary. Zones of degradation were measured to determine the amount of enzyme activity. Clearance zones were measured by subtracting the zone diameter from the diameter of isolate. Inoculated plates were incubated at 30°C for approximately 72 hours and photos were taken every 24 hours using a FluorChem SP (Alpha Innotech) gel documentation system.

IV.6. Strain improvement through protoplast fusion

Actinomycete isolates with the best cellulase and xylanase activity were cultivated and identified. The isolate with the largest clearance zones on cellulose, CMC and xylan agar was used for strain improvement with 3 other cellulase or xylanase producing isolates. Strain improvement was completed by protoplast fusion (genome shuffling) (Kieser et al., 2000) of two parental strains. The resulting progeny (fusants) were screened for improvement compared to parents for cellulase and xylanase activity by agar diffusion method.

Spores were inoculated into 5 mL of YEME (Yeast Extract - Malt Extract Medium: 0.3% (w/v) yeast extract, 0.5% (w/v) Bacto-peptone, 0.3% (w/v) malt extract, 1% (w/v) glucose, 34% (w/v) sucrose) supplemented with 2 mL/L $MgCl_2$ (2.5 M), 25 mL/L glycine (20%), 7.5 mL/L tiger's milk (1% (w/v) L-arginine, 0.75% (w/v) L-cystine, 0.75% (w/v) L-histidine, 0.75% (w/v) L-histidine, 0.75% (w/v) DL-homoserine, 0.75% (w/v) L-leucine, 0.75% (w/v) L-phenylalanine, 0.75% (w/v) L-proline, 0.15% (w/v) adenine, 0.15% (w/v) uracil, 0.01% (w/v) nicotinamide) to make Super YEME and incubated at 30°C at 150 rpm in a C25 Incubated Floor Model Shaker (New Brunswick Scientific) for 3 days. Liquid culture was transferred to a 50 mL screw-capped tube and centrifuged at 1801 *g* for 15 minutes at 4°C to harvest the cells. Then, the supernatant was discarded and the pellet was resuspended with 15 mL of 10.3% (w/v) chilled sucrose and centrifuged again at 1801 *g* for 15 minutes. The supernatant was decanted and the mycelial pellet was resuspended in 1 mL of aqueous lysozyme solution (1 mg/mL). The suspension was incubated at 30°C for 15-60 minutes (strain dependent) and was monitored for protoplast formation via phase contrast microscopy. The mixture was triturated three times and further incubated for 15 minutes. Then, 5 mL of Protoplast (P)

buffer (10.3% (w/v) sucrose, 1.79 mM K_2SO_4, 26.52 mM $MgCl_2$) was added. Prior to use the following components were added in order: 1.0 mL KH_2PO_4 (0.5%), 10 mL $CaCl_2$ (3.68%), 10 mL N-tris(hydroxymethyl) methyl-2-aminoethane-sulfonic acid (TES) buffer (5.73%, pH 7.2), 200 µL trace element solution (0.29 mM $ZnCl_2$, 0.74 mM $FeCl_3$, 0.06 mM $CuCl_2$, 0.05 mM $MnCl_2$, 0.03 mM $Na_2B_4O_7$, 0.01 mM $(NH_4)_6Mo_7O_{24}$)) was added and triturated thrice. Sterile cotton wool in 15 mL screw-capped tubes was used to filter protoplasts. The filtrate was centrifuged at 1801 *g* for 7 minutes at 4°C. Supernatant was discarded and the protoplasts were gently resuspended in 1.5 mL of P buffer. The turbidity of the parental strains was adjusted with P-buffer visually. For protoplast fusion to occur between the parental strains, 100 µL of each parent was transferred into a microfuge tube. The protoplasts were pelleted by centrifugation at 1801 *g* for 7 minutes at 4°C. Supernatant was discarded and the pellet was resuspended by gentle tapping and the addition of 0.8 mL of 50% (w/v) polyethylene glycol (PEG) 8000 molecular weight in P Buffer immediately and gently triturating once. The reaction tubes were stored at room temperature for 5 minutes. Protoplasts were gently plated onto R2YE for regeneration of cell walls (R2 medium: 12.88% (w/v) sucrose, 1.79 mM K_2SO_4, 132.86 mM $MgCl_2$, 1.25% (w/v) glucose, 0.0125% (w/v) casamino acids, 2.75% (w/v) agar. R2YE: R2 plus the following (after R2 was autoclaved): 10 mL KH_2PO_4 (0.5%), 80 mL $CaCl_2$ (3.68%), 15 mL L-Proline (20%), 100 mL TES buffer (5.73%, pH 7.2), 2 mL trace element solution, 5 mL NaOH (1N), 40 mL yeast extract (10%), 6 mL tiger milk) with the following dilutions of 10^0, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4} with P-buffer as diluent. The plates were incubated at 30°C for 5-7 days and the fusants were transferred to YDA when they were large enough to subculture.

IV.7. Degradation assays for fusants

Fusants resulting from protoplast fusion events were patched and screened on mineral salts agar containing, 0.05% (w/v) cellulose (same as above) or 0.05% (w/v) CMC (same as above) and 0.005% (w/v) congo red (CR) dye. Insoluble xylan from 1% (w/v) birchwood (same as above) was used to screen for xylanase activity. Inoculated plates were incubated at 30°C for approximately 27-40 hours. For the xylan assay, the plates were flooded with iodine after 27-40 hours. Iodine binds to xylan (Williams, 1983). Clearance zones appear clear on a dark background. Photos of the plates were taken using FluorChem SP (Alpha Innotech) gel documentation system after 26-40 hours.

IV.8. Comparison of sugars released from birchwood xylan by parents and putative fusants by gas chromatography

Fusants with greater cellulose and xylan degradation zones when compared to the parental strains, as determined by agar diffusion method, were investigated for sugar release by the putative fusants and the parents during growth in liquid culture containing 1% (w/v) birchwood xylan by the gas chromatography (GC) of alditol acetate derivatives (Blakeney et al., 1983).

Mineral salts solution (0.02 M Na_2HPO_4, 0.01 M KH_2PO_4, 0.54 mM $MgSO_4$, 2.28 mM NaCl, 0.3 mM $CaCl_2$) without the xylan was prepared and autoclaved. Then, 1% xylan medium was created after the 15 g of xylan was washed and added to the mineral salts solution. Since the birchwood xylan contained free xylose, five washes of xylan with water were performed to minimize free xylose present. Birchwood xylan (7.5 g) was added to two 50 mL screw-capped tubes for a total of 15 g of xylan. The xylan was autoclaved and then washed with sterile water, vortexed for 5 minutes and inverted 25 times to mix. The xylan was pelleted by centrifugation for 15 minutes at 4°C at 1801 *g*

and the supernatant was discarded. This wash was repeated four times. In subsequent washes, the xylan pellet was compact; therefore a sterile spatula was used to resuspend the pellet.

The washed xylan was transferred to mineral salts medium by obtaining two aliquots of 50 mL mineral salts solution in 50 mL screw-capped tubes from the 1.5 L mineral salts solution previously autoclaved for quantitative transfer. One aliquot of mineral salts was added to the washed xylan pellet and this pellet was resuspended by dislodging pellet with sterile spatula, vortexed for 1 minute and the tube was inverted 25 times. The resuspended xylan was poured into the media bottle containing mineral salts solution. The screw-capped tube with washed xylan was rinsed with mineral salt solution from the other aliquot. The rinsate was transferred into the media bottle containing mineral salts solution. The tubes were rinsed once more to ensure all the xylan was transferred and poured into media bottle containing mineral salts solution. This was repeated for the transfer of xylan from the other screw-capped tube containing the other aliquot of approximately 7.5 g of xylan.

Spore stock concentration of parents and fusants were quantified using a Petroff-Hauser counting chamber to determine the volume of spores to inoculate into the birchwood xylan in mineral salts liquid cultures. Duplicate 100 mL of 1% birchwood xylan in 500 mL erlenmeyer flasks were inoculated with 10^6 spores/mL of parents, PSY159 and WBF90B, and fusants, FA1-14, FA2-4, FA9-13, FA9-14. Two 100 mL aliquots of uninoculated 1% washed birchwood xylan in mineral salts was used as control. Cultures were shaken at 150 rpm at 30°C.

Gas chromatographic analysis of sugars from the degradation of xylan by parents and fusants was analyzed based on the analysis method by Blakeney et al. (1983) for an alditol acetate derivitization of neutral sugars released during xylan degradation. Samples were obtained every 48 hours by removing a 200 µL aliquot aseptically from each culture into a microfuge tube. This was then centrifuged at 17 000 *g* for 15 minutes at room temperature in a microcentrifuge, accuSpin Micro 17 (Fisher Scientific). The supernatant was transferred to a new microfuge tube and the pellet was discarded.

At 528 hours, the pellets obtained from uninoculated aliquots were saved and the pellet was hydrolyzed and derivatized to determine the sugars present in birchwood xylan. 100µL of sterile water was added to the uninoculated culture pellets and the sugars present in the pellet were analyzed by the same analysis performed with the supernatants.

The pH was measured using pH strips (Whatman). The standard mix contained 2 mg/mL of the sugars generally found in the plant cell wall; rhamnose, fucose, ribose, inositol, xylose, arabinose, glucose, galactose and mannose, all of which are monosaccharides (Blakeney et al., 1983). 8 µg of the standard mix, containing the aforementioned sugars, was derivatized for each sample set.

Glycosidic linkage between sugars of the polysaccharides present in the supernatant were hydrolyzed by the addition of 1M H_2SO_4, vortexed for a few seconds and incubated at 100°C for 1 hour in a heating block. The tubes were then cooled to room temperature. After cooling, 25 µL was aliquoted into three separate microfuge tubes. To each aliquot, the internal standard inositol 2 mg/mL was added to a final concentration of 4 µg. Then, 10 µL of 9 M NH_3 was added and vortexed for 3 seconds. 100 µL of 2%

(w/v) $NaBH_4$ in dimethyl sulfoxide (DMSO) was added. The samples were vortexed for 3 seconds. Samples were then incubated at 40°C for 90 minutes. 25 µL of acetic acid was added and vortexed for 3 seconds. 25 µL of 1-methylimidazole and 400 µL of acetic anhydride was added then vortexed. This was then incubated at room temperature for 10 minutes. 800 µL of water (18 MΩ) was added and mixed by inversion until precipitate dissolved. Then, 200 µL dichloromethane (DCM) was added and mixed by inversion 30 times. The tubes were placed at 4°C for 2 hours to overnight. The DCM phase (bottom layer) was collected into 2 mL Screw Top 9 mm Wide Opening Vials with 200 µL glass insert with polymer spring and used for gas chromatography analysis.

Separation of the sugars was completed by gas chromatography-mass spectrometry (Varian 3900 Gas Chromatographer equipped with a flame ionization detector), with helium as a carrier gas (flow rate of 10.0 mL/minute) on a FactorFour™ Column VF-23ms (30 meters, 0.250 mm ID, 0.25 µm film thickness) capillary column. Injection (1.0 µL) was completed in split mode with a ratio of 20:13, utilizing a Varian CompiPal autosampler. The temperatures for the injector and detector were 240°C and 300°C respectively. The oven temperature was ramped from 140°C to 180°C at 10°C per minute, and held for five minutes followed by an increase of 3°C per minute until a final temperature of 250°C, where the oven was held for 31.33 minutes. Identification was completed by retention time comparison between sample and standard. Galaxie workstation software (Varian) was used to identify and quantify the peaks of alditol acetates.

In order to quantify the sugars present in the samples, an internal standard of inositol was incorporated into the samples. A response factor was calculated for the

standard for the sugars in the standard mix, in relation to the internal standard, inositol and the amount of sugar (nmol/mg of dry weight of xylan) was calculated:

Response factor for each sugar in standard mix (Q) = area of the peak for standard sugar/internal standard

Amount of sugar (nmol/mg of dry weight of xylan) = ([(peak area of sugar/Q_{sugar})/Area $peak_{internal\ standard}$*amount of internal standard in sample]/molecular $weight_{sugar}$)/mg of dry weight of xylan

The amount of sugar present in the uninoculated and inoculated samples for the six replicates was averaged and the standard error was calculated at each time point sampled. It was assumed there was 1% (w/v) of xylan present but this was inaccurate because birchwood xylan is not soluble.

IV.9. Genomic DNA extraction

Genomic DNA (gDNA) of actinomycete isolates was extracted for 16S rRNA gene amplification by PCR for taxonomic identification and to determine phylogenetic relatedness of cultivated actinomycete isolates. Due to difficulties in obtaining genomic DNA from the actinomycete isolates, several procedures were attempted as described below.

IV.9.1. Salt-extraction of genomic DNA

Genomic DNA extraction was performed as described in Aljanabi and Martinez (1997). For growth of actinomycetes in liquid culture, 5 µL of glycerol spore stock was inoculated into 5 mL of yeast dextrose broth (0.1% (w/v) yeast extract, 0.3% (w/v) dextrose), shaken at 150 rpm at 30°C for 1-3 days. Mycelia were harvested by centrifugation at 1801 *g* for 15 minutes. The supernatant was decanted and discarded. The mycelial pellet was washed in chilled 10.3% (w/v) sucrose by centrifugation. The

supernatant was discarded and the pellet was stored at -20°C or used immediately for DNA extraction.

Salt homogenizing buffer (400 µL) (0.4 M NaCl, 10 mM Tris–HCl pH 8.0 and 2 mM EDTA pH 8.0) was added to 750 µL of mycelia and resuspended by triturating. Then 40 µL of 20% SDS and 8 µL of 20 mg/mL proteinase K were added, mixed well, and incubated at 55-65°C for 1 hour. Then, 300 µL of 6 M NaCl was added, vortexed for 30 seconds and centrifuged for 30 minutes at 10 000 *g*. The supernatant was transferred to a fresh tube. An equal volume of isopropanol was added and incubated at -20°C for 1 hour. Samples were centrifuged for 20 minutes at 10 000 *g* at room temperature. The DNA pellet was washed with 70% ethanol, dried and finally resuspended in 300-500 µL sterile water.

IV.9.2. Actinomycete colony Polymerase Chain Reaction (PCR)

Actinomycete colony PCR was attempted for PCR amplification of 16S rDNA because salt extraction method did not yield any genomic DNA (gDNA) for WCB26. Single colonies were aseptically transferred from plates and placed in a microfuge tube. To the colonies, 200 µL of TE buffer pH 8.0 (10 mM Tris-Cl, pH 7.5, 1 mM EDTA) with 1% triton X-100 and 2 mg/mL lysozyme was added and the colony was homogenized using a sterile pestle. The homogenate was incubated at 37°C for 30 minutes then incubated at 99°C for 15 minutes and cooled on ice for 2 minutes. The tube was centrifuged at 16 200 *g* for 5 minutes to remove cell debris. The resulting supernatant containing the DNA was used as template for 16S rRNA gene amplification by PCR (see section IV.10.) and stored at -20°C.

IV.9.3. DNAzol extraction of genomic DNA

A 500 µL aliquot was obtained from the liquid culture of WCB26 grown in yeast dextrose broth (see above) for four days. To the aliquot, 1 mL of DNAzol (Invitrogen) was added, mixed and centrifuged at 10 000 *g* for 10 minutes at room temperature. The supernatant was transferred to a fresh microfuge tube and 100% cold ethanol was added. This was then mixed by inversion for five to eight times and stored at room temperature for 3 minutes; a cloudy precipitate of DNA formed. The DNA was collected by centrifugation at 10 000 *g* for 30 minutes at room temperature. The supernatant was decanted and 800 µL of ice cold 70% ethanol was added. The tube was inverted several times to mix and centrifuged at 17 000 *g* for 2 minutes at room temperature. The supernatant was decanted and the DNA pellet was washed once more. The DNA was air dried and resuspended in 300 µL of sterile water.

IV.10. 16S rRNA gene amplification by Polymerase Chain Replication (PCR)

To accomplish the taxonomic identification of the isolates with cellulase and xylanase activity, universal bacterial primers (Lane, 1991) were used to amplify the rRNA gene from genomic DNA of each isolate. The 16s rRNA primers that were used are: 27F (5' AGA GTT TGA TCM TGG CTC AG 3'), where M=C:A and 1492R (5' TAC GGY TAC CTT GTT ACG ACT T 3'), Y=C:T (Lane, 1991). The total reaction volume was 50 µL and the reaction conditions were: up to 250 ng DNA template, 1X Green GoTaq® Flexi Buffer (Promega), 0.2 mM dNTPs, 0.4 µM each of forward and reverse primers, 2 mM $MgCl_2$ and 1.25 U of GoTaq® Hot Start Polymerase (Promega). Reaction conditions were denaturation at 95°C for 5 minutes, followed by 30 cycles at 94°C for 30 seconds 55°C for 30 seconds, 72°C for 1 minute and finally 72°C for 5 minutes (Thermo Scientific Px2 Thermal Cycler).

To remove oligonucleotides and unincorporated dNTPs, the EZ-10 Spin Column DNA Cleanup Kit or EZ-10 Spin Column PCR Products Purification Kit was used and as per manufacturer's protocol (Bio Basic Inc.). Purified amplicons were stored at -20°C.

IV.11. Cloning of Polymerase Chain Reaction (PCR) amplified 16S rRNA genes into pGEM®-T Easy Vector and transformation into *E. coli*

PCR amplified 16S rRNA genes from select actinomycete isolates were cloned into pGEM®-T Easy Vector according to manufacturer's protocol (Promega). The pGEM®-T Easy Vector system was used for the ligation of PCR amplicons to the vector and subsequent transformation into *E. coli* JM109.

Cloning of 16S rRNA gene was performed through ligation of the PCR amplicons to pGEM®-T Easy Vector and then transformation into *E. coli* JM109. First, ligation of the PCR amplicons to pGEM®-T Easy Vector was performed as per manufacturer's instructions. Frozen *E. coli* JM109 High Efficiency Competent Cells (Promega) were thawed on ice and a 25 μL aliquot of cells was added to a microfuge tube containing 2 μL of the ligation reaction. The transformation reactions were incubated on ice for 20 minutes. The cells were heat-shocked for 45–50 seconds in a water bath at 42°C and then incubated on ice for 2 minutes. A 950 μL aliquot of Super Optimal broth – Catabolite repression (SOC) medium (2% (w/v) tryptone, 0.5 % (w/v) yeast extract, 0.01 M NaCl, 0.0025 M KCl, 0.02 M Mg^{2+} stock, filter sterilized, 0.02 M glucose, filter-sterilized) was added. The cells were incubated for 1.5 hours at 37°C with shaking at approximately 150 rpm in an orbital shaker. For each transformation culture, 100 μL was plated onto duplicate Luria Bertani (LB) (1.0% (w/v) tryptone, 0.5% (w/v) yeast extract, 85.5 mM NaCl, adjust pH to 7.0 with NaOH) containing 100 μg/mL ampicillin, 0.5 mM IPTG

(Isopropyl p-D-thiogalactopyranoside), 80 µg/mL X-Gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) agar plates and incubated overnight at 37°C.

IV.12. Plasmid Purification

The alkaline lysis method (Sambrook & Russell, 2001) was used to determine if vectors containing inserts (PCR amplified 16S rRNA gene) were obtained. A single clone was inoculated into 3 mL of LB broth containing 0.1 mg/mL ampicillin and incubated overnight at 37°C and shaken at 250 rpm. Culture was harvested in 1.5 mL microfuge tube by centrifugation for 30 seconds at 17 000 *g* at room temperature. The supernatant was completely removed by aspiration to ensure the pellet was dry. The cell pellet was resuspended in 100 µL of ice-cold alkaline lysis solution I (50 mM glucose, 25 mM Tris-Cl (pH 8.0), 10 mM EDTA (pH 8.0)) with vigorous vortexing. Then, 200 µL of alkaline lysis solution II (0.2 N NaOH (freshly diluted from a 10 N stock), 1% SDS) was added and quickly mixed by inversion five times and tubes were placed on ice for a minute. Next, 150 µL of alkaline lysis solution III (60 mL of 5 M potassium acetate, 11.5 mL of 17 M glacial acetic acid, 28.5 mL of H_2O) was added and mixed by inversion several times. The tubes were incubated on ice for 5 minutes. Tubes were centrifuged at maximum speed (17 000 *g*) for 5 minutes and supernatant was transferred to a fresh tube. An equal volume of phenol:chloroform:isoamyl alcohol pH 8.0 (25:24:1, v/v) was added and the mixture was vortexed. This suspension was centrifuged at room temperature at 17 000 *g* for 2 minutes. The aqueous phase was transferred to a new tube and 2 volumes of room temperature absolute ethanol were added, mixed and the DNA was allowed to precipitate at room temperature for 2 minutes. The plasmid DNA was collected by centrifugation at 17 000 *g* for 5 minutes at room temperature. The supernatant was removed and the DNA was washed by addition of 1 mL of cold 70% ethanol and

centrifugation at 17 000 *g* for 2 minutes at room temperature. The supernatant was removed and the DNA was air-dried at room temperature, 5-10 minutes. The DNA was resuspended in 50 µL of TE pH 8.0 (10 mM Tris-Cl, pH 7.5, 1 mM EDTA) containing DNAse-free pancreatic RNAse (20 µg/mL). Plasmid DNA was stored at -20°C.

For sequencing, plasmids were purified using EZ-10 Spin Column Plasmid DNA Minipreps Kit (Bio Basic Inc.) according to the manufacturer's instructions. Plasmid DNA was stored at -20°C.

IV.13. Cycle sequencing of fusant FA1-14

The DNA sequence of the amplified 16S rRNA gene of the fusant FA1-14 was determined by cycle sequencing using 3130 Genetic Analyzer (Applied Biosystems). Cloned 16S rDNA was purified by alkaline lysis and was used as the template for sequencing. The concentration of DNA was determined to be 2.78 µg/µL. 4 µL of 3.1 sequencing mix (Applied Biosystems), 2 µL of 5X sequencing buffer, 0.16 pmol/µL of 27F and 1492R primers, 300 ng of DNA template to a final volume 20 µL were combined. The conditions for the sequencing reaction were as follows: denaturation at 96°C for 3 minutes, then 25 cycles of: 96°C for 10 sec, 50°C for 5 sec, 60°C for 4 minutes.

Unincorporated dye terminators were removed by transferring reactions to a 1.5 mL microfuge tube and adding: 20 µL of sterile distilled water and 60 µL of 100% isopropanol. This was vortexed and left to incubate in the dark for 15 minutes to allow precipitation of the extension products. This was then centrifuged at 17 000 *g* for 20 minutes and the supernatant was discarded. To the pellet, 250 µL of 75% isopropanol was added then centrifuged at 17 000 *g* for 5 minutes. The supernatant was discarded and the samples were placed in a 90°C heating block for 1 minute to dry the samples. To the

dried samples, 15 µL of Hi-Di formamide (Applied Biosystems) was added and vortexed to mix. The tubes were then incubated at 95°C for 2 minutes and then chilled on ice for another 2 minutes. This was vortexed and briefly centrifuged for 1 minute at 17 000 *g*. The solution was kept on ice until it was to be loaded into the sequencer. The samples were loaded onto 3130 Genetic Analyzer (Applied Biosystems) for capillary electrophoresis. Sequences were viewed using ContigExpress software from (Vector NTI v10). Nearest relatives of the FA1-14 (1105 bp) were identified by comparison to submitted GenBank sequences through nucleotide-nucleotide Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990). Sequences with 98% identity from BLAST and FA1-14 sequence were aligned using ClustalW (Thompson et al., 1994). Phylogenetic analysis was performed using ClustalTree from San Diego Supercomputer Biology Workbench and neighbour joining method (Saitou & Nei, 1987). Bootstrap analysis (1000 replicates) with resampling two times was used to estimate node confidences of the phylogenetic tree which was viewed using njplot (http://pbil.univ-lyon1.fr/software/njplot.html) and TreeView (Page, 1996).

IV.14. Restriction Fragment Length Polymorphism analysis

PCR amplified 16S rRNA genes of parents used for protoplast fusion and fusants were analyzed by Restriction Fragment Length Polymorphism (RFLP) to ascertain if there were any differences in banding pattern corresponding to genome shuffling in 16S rRNA genes during protoplast fusion. First, the 16S rDNA amplicons were digested with FastDigest® *Hae*III or *Hha*I (Fermentas) by incubation at 37°C for 1 hour and 20 minutes. The final reaction volume was 30 µL and 2 µL 10X FastDigest® buffer, 1 µL of enzyme, 10 µL (0.2 µg) of DNA were added.

The digested fragments were then separated by electrophoresis on a 3% (w/v) Tris-acetate-EDTA (TAE) agarose gel. 6X DNA loading dye (16% (w/v) sucrose, 100 mM EDTA (pH 8), 0.000 625% (w/v) Bromophenol blue) was added to the digests for a final concentration of one times. The digests were electrophoresed on a 3% (w/v) TAE agarose gel at 100V for 1 hour. Molecular weight markers, 0.5 – 10 kb DNA ladder (Bioshop) or 100 bp ladder – 1 kb (New England Biolabs) were electrophoresed on the gel with the samples to determine which marker would allow better analysis. Agarose gels were stained in ethidium bromide (0.5 µg/mL) for 10 minutes and visualized on UV transilluminator and photographed using FluorChem SP gel documentation system (Alpha Innotech).

IV.14.1. Polyacrylamide gels

Polyacrylamide gels were used to obtain better resolution of digested 16S rDNA for RFLP analysis. DNA loading dye (1X) was added to the digested DNA prior to loading. Digested DNA were electrophoresed on 8% (5.9 mL water, 2 mL of 5X Tris-borate-EDTA buffer (TBE), 2 mL 40% (19:1) acrylamide, 100 µL 10% Ammonium persulfate (APS), 10 µL N,N,N',N'-tetramethylethylenediamine (TEMED), total volume 10 mL), 4% (6.9 mL water, 2 mL of 5X TBE, 1 mL 40% (19:1) acrylamide, 100 µL 10% APS, 10 µL TEMED, total volume 10 mL) and 5% (6.65 mL water, 2 mL of 5X TBE, 1.25 mL 40% (19:1) acrylamide, 100 µL 10% APS, 10 µL TEMED, total volume 10 mL). 0.5X TBE was used as the running buffer and the gel ran for 1 hour at 100V. Molecular weight markers 0.5 – 10 kb DNA ladder or 100 bp ladder – 1 kb were loaded for a reference. Agarose gels were stained in ethidium bromide (0.5 µg/mL) for 10 minutes and visualized on UV transilluminator and photographed using FluorChem SP gel documentation system (Alpha Innotech).

V. RESULTS

V.1. Actinomycetes isolation from various soil samples

Over 800 actinomycete isolates have been previously isolated in the Strap lab from University of Ontario Institute of Technology (UOIT) forest rhizosphere soil, UOIT forest bulk soil, conservation area rhizosphere soil and bulk soil and potting soil types to comprise a laboratory culture collection of actinomycetes. As part of this investigation, 186 isolates were isolated in pure culture from potting soil out of an original 280 cultivated from this source. Out of these isolates, spore stocks were created for 143 isolates. The remaining 41 were screened for cellulase activity on carboxymethyl cellulose (CMC) – congo red agar (CRA) (0.05%) but did not produce any cellulase or were not amenable to spore stock preparation; therefore no spore stocks were made for the 41 isolates and these cultures were discarded.

V.2. Screen of pure culture isolates for cellulase and xylanase activity

Actinomycete isolates cultivated from potting soil, UOIT forest bulk soil, UOIT forest rhizosphere soil, conservation area bulk soil and conservation area rhizosphere soil were screened for the production of extracellular cellulase and xylanase activities. Two forms of cellulose, insoluble cellulose and a soluble form of cellulose, carboxymethyl cellulose, were used to screen the isolates for cellulase activity. Degradation assays were performed by growing actinomycete isolates on mineral salts agar containing, cellulose (0.05%) or CMC (0.05%) and congo red (CR) dye (Figure V.2.1.A-B.). Congo red is a dye that binds to cellulose. Spores of each isolate were spotted onto cellulose-CR or CMC-CR agar plates. As cellulose or CMC is degraded, a clear zone becomes visible around the isolate due to the release of the congo red dye. Insoluble xylan from birchwood (1%) was used to screen for xylanase activity. After 3 days of growth, plates

were stained with iodine to visualize the degradation (clearance) zones which appear clear against a dark background (Figure V.2.1.C). Clearance zones were measured to determine the amount of enzyme activity.

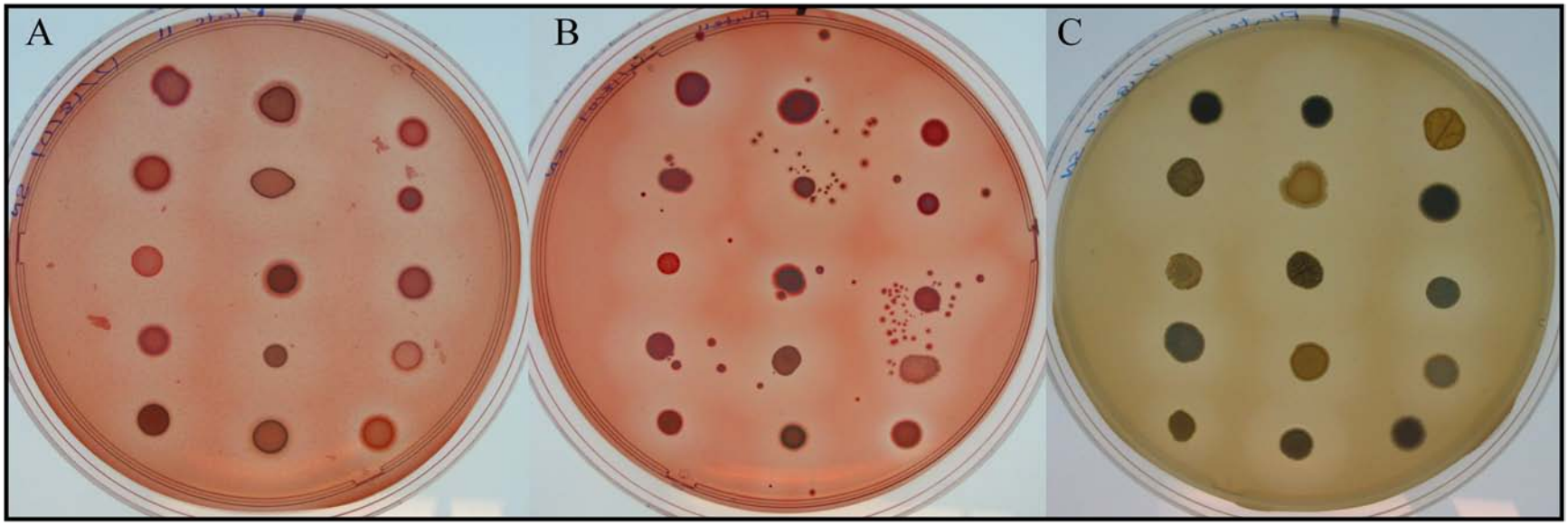


Figure V.2.1. Example screen of isolates by agar diffusion method on A) cellulose congo red agar; B) CMC- congo red agar; and C) xylan stained with iodine. From top to bottom, left to right: PSP25, PSP31, PSP32i, PSP32ii, PSP46, PSP91a, PSY53b, PSY64, PSY66, PSY71, PSY72, PSY83, PSY93, WBF15aiii, WBF90B.

A total of 391 sporulating isolates were screened for cellulase and xylanase activity. The number of isolates screened from bulk soil was 138, 110 from rhizosphere, and 143 from potting soil. Actinomycetes isolated from potting soil had greater clearance zones in all the assays compared to rhizosphere or bulk isolates (Figure V.2.2-3.). For the cellulose assay, rhizosphere and bulk isolates were comparable with the isolates that degraded cellulose (Figure V.2.2.A.). For the CMC assay, rhizosphere isolates had larger clearance zones than bulk isolates overall (Figure V.2.2.B.). Finally, in the xylan assay, rhizosphere isolates had slightly greater zones on xylan agar than the bulk isolates (Figure V.2.2.C.). A total of 283 broad spectrum degrading strains were identified. Of these, 70% produced both xylanases and cellulases (Figure V.2.3.). Most of the potting soil isolates screened (>80%) produced cellulases that degraded cellulose and CMC and xylanases to degrade xylan, while greater than 60% of the rhizosphere isolates and bulk isolates displayed these activities (Figure V.2.3.).

From the 391 total isolates assayed, 283 were found to have broad spectrum cellulase and xylanase activity. The isolates that had the largest clearance zones for all three assays are listed in Table V.2.1. Most of these isolates that exhibited multiple enzyme activities were cultivated from potting soil (PS). Four isolates from bulk soil, PBF57, PBF86, WBF3, and WCB26, and an isolate from rhizosphere forest soil, PRF42, were observed to have cellulase and xylanase activity as they had the largest clearance zones on cellulose, CMC and xylan agar (Table V.2.1.). PSY159 was the best isolate that produced the greatest clearance zones on cellulose, CMC and xylan (Table V.2.1.).

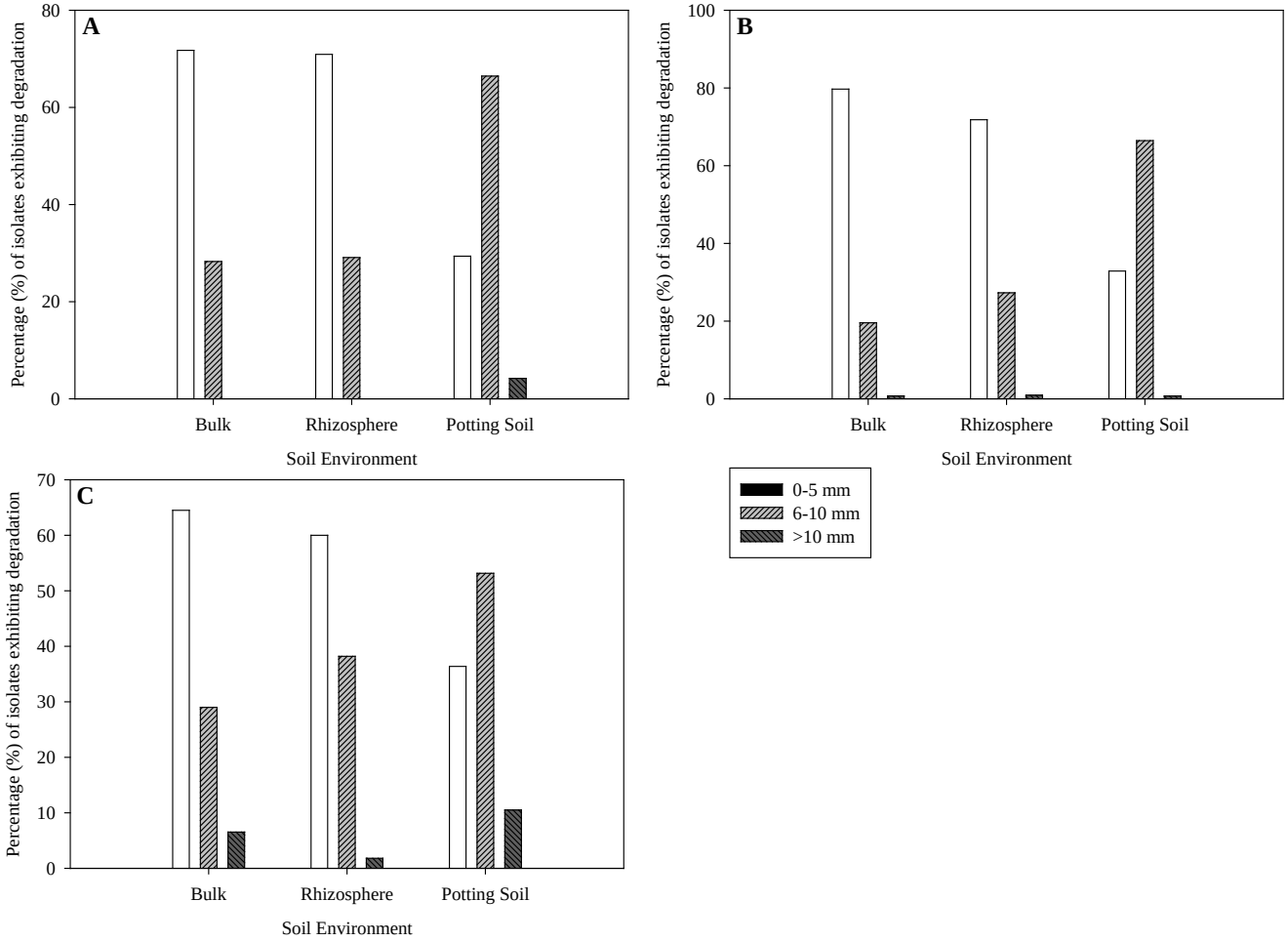


Figure V.2.2. Percentage of isolates from bulk, rhizosphere and potting soil with various clearance zone radii on A) cellulose, B) CMC, and C) xylan agar. n = 391: bulk – n = 138; rhizosphere – n = 110; potting soil – n = 143. Clearance zones were determined by taking into account the size of the spot. Clearance zone radius = (diameter of zone – diameter of spot)/2

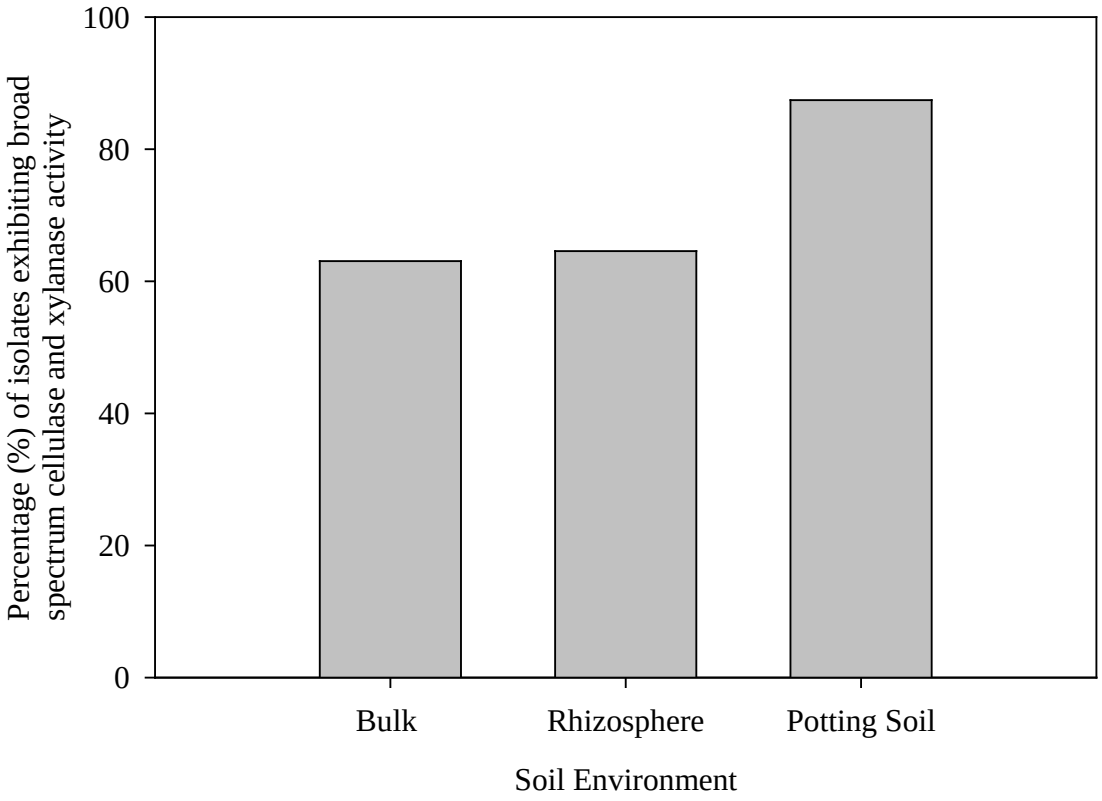


Figure V.2.3. Comparison of the percentage of isolates screened from different soil environments that produced broad spectrum cellulase activity by degrading cellulose and CMC and xylanase activity from degrading xylan. n = 391: bulk , n = 138; rhizosphere, n = 110; potting soil, n = 143.

Table V.2.1. Best strains from 391 sporulating actinomycete isolates screened that exhibit broad spectrum cellulase and xylanase activity as observed by clearance zones at ~72 h incubated at 30°C on Cellulose-Congo Red Agar (CCRA), CMC-Congo Red Agar (CMC-CRA), and Xylan Agar plates

Isolate	**CCRA (mm)***	**CMC-CRA (mm)***	**Xylan agar (mm)***
PSY159	10	11	13
PSY79	11	9	11
PSY57	11	9	10
PSP57	9	9	12
PSY147	8	9	13
PSP46	8	9	13
PSY102	9	9	11
PSY144	8	10	11
PSY38	11	7	9
PSY16	10	8	9
PRF42	8	12	9
PBF57	8	9	12
PBF86	8	8	13
WBF3	8	8	13
WCB26	8	7	13

*Clearance zones were determined by taking into account the size of the colony growth of the isolate. Clearance zone radius = (diameter of zone – diameter of spot)/2

V.3. Strain improvement of isolates

Once screening of the actinomycete isolates for cellulase and xylanase activity was complete, isolate PSY159 was discovered to exhibit the best overall cellulase and xylanase activity. Strain improvement through protoplast fusion was performed to produce strains with even greater cellulase and xylanase activity than the best producer PSY159. Three fusion experiments were performed between four strains: A) PSY159 and WBF90B, B) PSY159 and PSP55, C) PSY159 and WCB26.

These isolates were chosen as the parents for protoplast fusion for the following reasons: PSY159 had the greatest broad spectrum activity cellulase and xylanase activity (Figure V.3.1.). WBF90B had very clear cellulase clearance zones on CMC and cellulose but did not have very large diffusible zones (Figure V.3.1.). PSP55 was one of the best strains with broad spectrum cellulase and xylanase activity in the initial screening on cellulose, CMC and xylan agar (Figure V.3.1.) and it grew quickly in Super YEME liquid culture. WCB26 was one of the isolates with the best broad spectrum activity (Table V.2.1. and Figure V.3.1.) and was from conservation bulk soil compared to potting soil.

From the three protoplast fusion events, a total of 501 fusants were screened for cellulase and xylanase activity (Figure V.3.2.). For protoplast fusion between PSY159 and WBF90B (A), 162 fusants were screened, for fusion between PSY159 and PSP55 (B), 141 fusants were screened, for fusion between PSY159 and WCB26 (C), 198 fusants were screened. Of the total number of fusants screened, there were six that showed overall improvement of broad spectrum cellulase and xylanase activity. Cellulase and xylanase activity was determined by measuring the clearance zone from the edge of the colony to the edge of the clearance zone and this was compared to the zones produced by the two parents from each fusion experiment (Table V.3.1.).

From each fusion experiment, the fusants that showed improvement in cellulase and xylanase activity were selected for further experimentation. The fusion between PSY159 and WBF90B had four fusants, FA1-14, FA2-4, FA9-13, and FA9-14, that demonstrated improved cellulase and xylanase activity compared to the parents. In the fusion between PSY159 and PSP55 and between PSY159 and WCB26B, there were not many fusants with improved cellulase and xylanase activity when compared to their respective parents. Subsequently, the fusant or fusants with improved cellulose and xylanase activity, when compared to the parents, were chosen from each of the fusion experiments, PSY159 & PSP55 and PSY159 & WCB26B. FB11-14 created from fusion PSY159 & PSP55, and FC17-4 created from PSY159 & WCB26 were the fusants that were selected for further study.

The clearance zones on cellulose, CMC and xylan agar for each of the fusion experiments were reanalyzed to account for the growth of the strain because the number of cells and hence, colony size, will affect the amount of enzyme produced. Growth was determined by measuring the clearance zone from one edge of the zone to the opposite edge of the zone with the colony included. This re-analysis resulted in a different outcome compared to the data obtained for clearance zone measurements from the edge of strain to the edge of the clearance zone. Only two, FA9-13 and FA9-14, out of the six original best strains, FA1-14, FA2-4, FA9-13, FA9-14, FB11-14 and FC17-4, had larger clearance zones on cellulose, CMC and xylan when compared to both parents for the respective fusion experiments that generated these fusants (Table V.3.1.). Furthermore, only one isolate, FA9-15, out of all three protoplast fusion events (Table V.3.1.) exhibited clearance zones greater than 2 mm when compared to both parents on cellulose, CMC

and xylan. Also, the colour of the spores of the fusants were more similar to the spores of PSY159 (Figure V.3.2.).

Protoplast fusion between the parents PSY159 and WBF90B had the largest number of fusants with improved cellulase or xylanase activity compared to when PSY159 was fused with PSP55 or WCB26 when improvement of fusants was compared to both parents (Figure V.3.2.A.). Fusion between PSY159 and WCB26 had the largest number of fusants with a decrease in cellulase and xylanase activity (Figure V.3.2.B.). Protoplast fusion between PSY159 and WBF90B was determined to be the best for strain improvement.

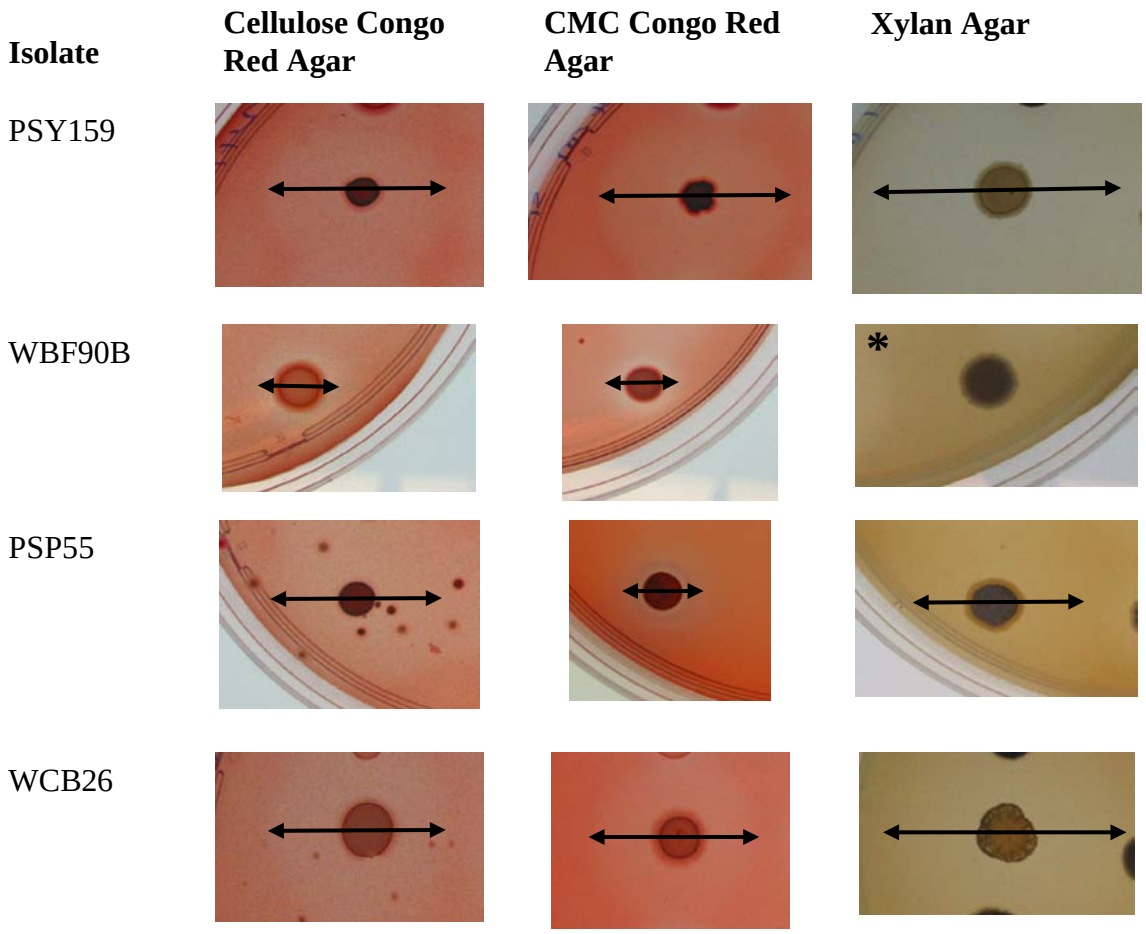


Figure V.3.1. Isolates used as parents in protoplast fusion experiments screened on cellulose congo red agar, CMC congo red agar and xylan agar displaying clearance zones indicating cellulase and xylanase activity, respectively. Arrows indicate clearance zones from one edge to the opposite edge. *No clearance zone observed.

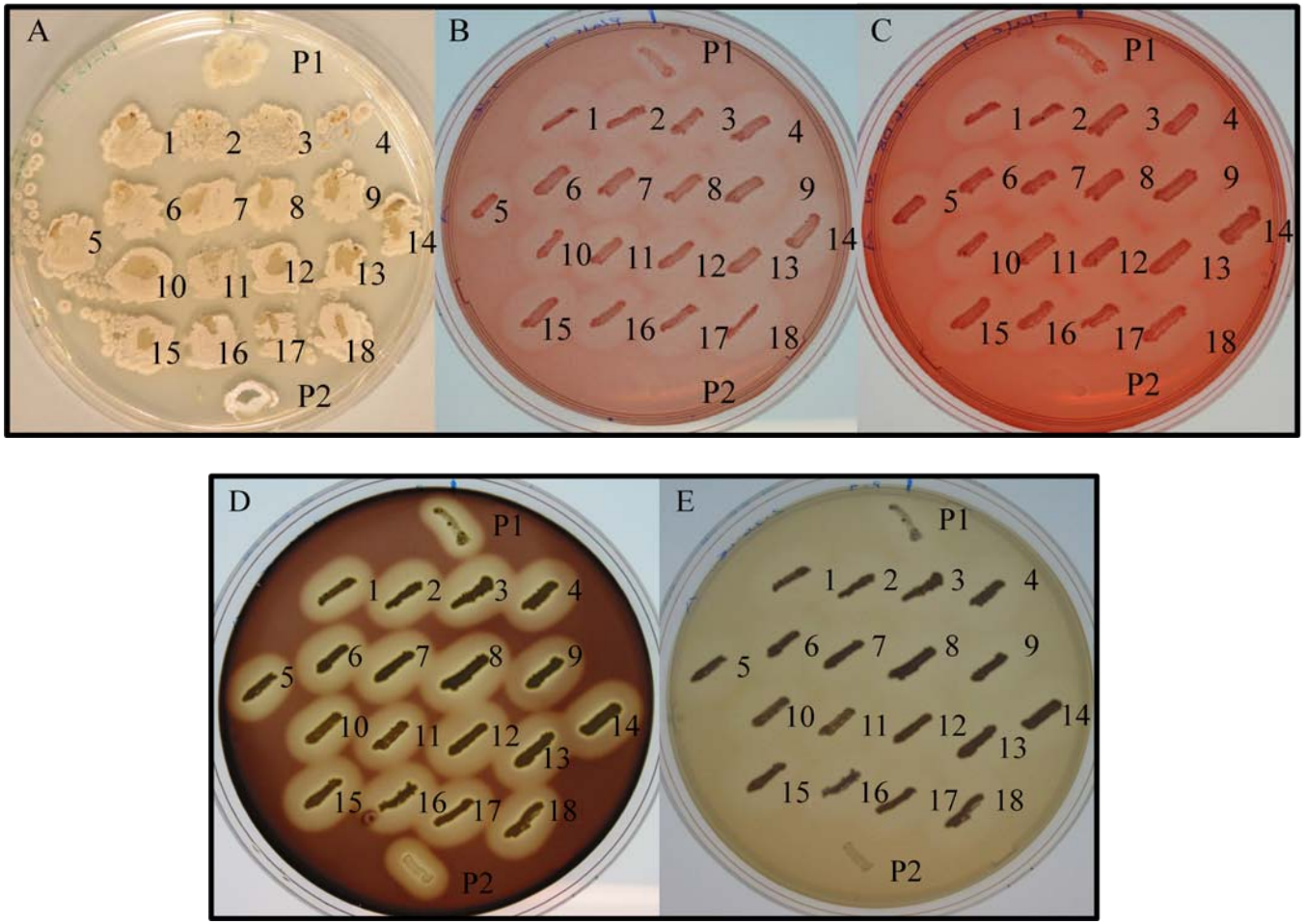


Figure V.3.2. Example of screening A) fusants for cellulase and xylanase activity from fusion between parents PSY159 and WBF90B on B) C-CRA, C) CMC-CRA, D) xylan agar stained with iodine and E) xylan agar before staining with iodine. C-CRA, CMC-CRA and xylan agar plates were incubated for 40 hours at 30°C. A-E) Parents P1 – PSY159 and P2 – WBF90B are designated.

Table V.3.1. Fusants that showed improved broad spectrum cellulase and xylanase activity compared to parental strains as observed by clearance zones at ~27-40 hours incubated at 30°C on Cellulose-Congo Red Agar (C-CRA), CMC-Congo Red Agar (CMC-CRA), and xylan agar plates from each fusion experiment between parents: PSY159 & WBF90B, n=162; PSY159 & PSP55, n=141; and PSY159 & WCB26, n=198. Reanalysis of data to account for growth[b] showed only one fusant, FA9-15, with improved activity broad spectrum activity.

			Clearance zone difference(mm)[a]			Clearance zone difference(mm)[b]		
	Fusant ID	**Parent**	**C-CRA**	**CMC-CRA**	**Xylan Agar**	**C-CRA**	**CMC-CRA**	**Xylan Agar**
1	FA1-14	PSY159	1	1	1	3	-1[c]	-1
		WBF90B	4	5	5	4	3	3
2	FA2-4	PSY159	1	3	1	-2	1	-1
		WBF90B	6	4	4	3	3	3
3	FA9-13	PSY159	3	4	3	4	3	2
		WBF90B	5	6	3	6	5	2
4	FA9-14	PSY159	4	3	3	4	3	1
		WBF90B	6	5	3	6	5	1
5	FB11-14	PSY159	3	3	1	3	2	-1
		PSP55	3	5	3	3	4	1
6	FC17-4	PSY159	3	0	0	1	1	-3
		WCB26	5	3	1	3	2	-3
7	FA9-15	PSY159	3	3	3	4	3	3
		WBF90B	5	5	3	6	5	3

a. Clearance zone difference measured from edge of strain to edge of clearance zone.

b. Clearance zone difference measured from edge of strain to edge of clearance zone with growth accounted for.

c. Clearance zone difference which is smaller for fusant compared to parent.

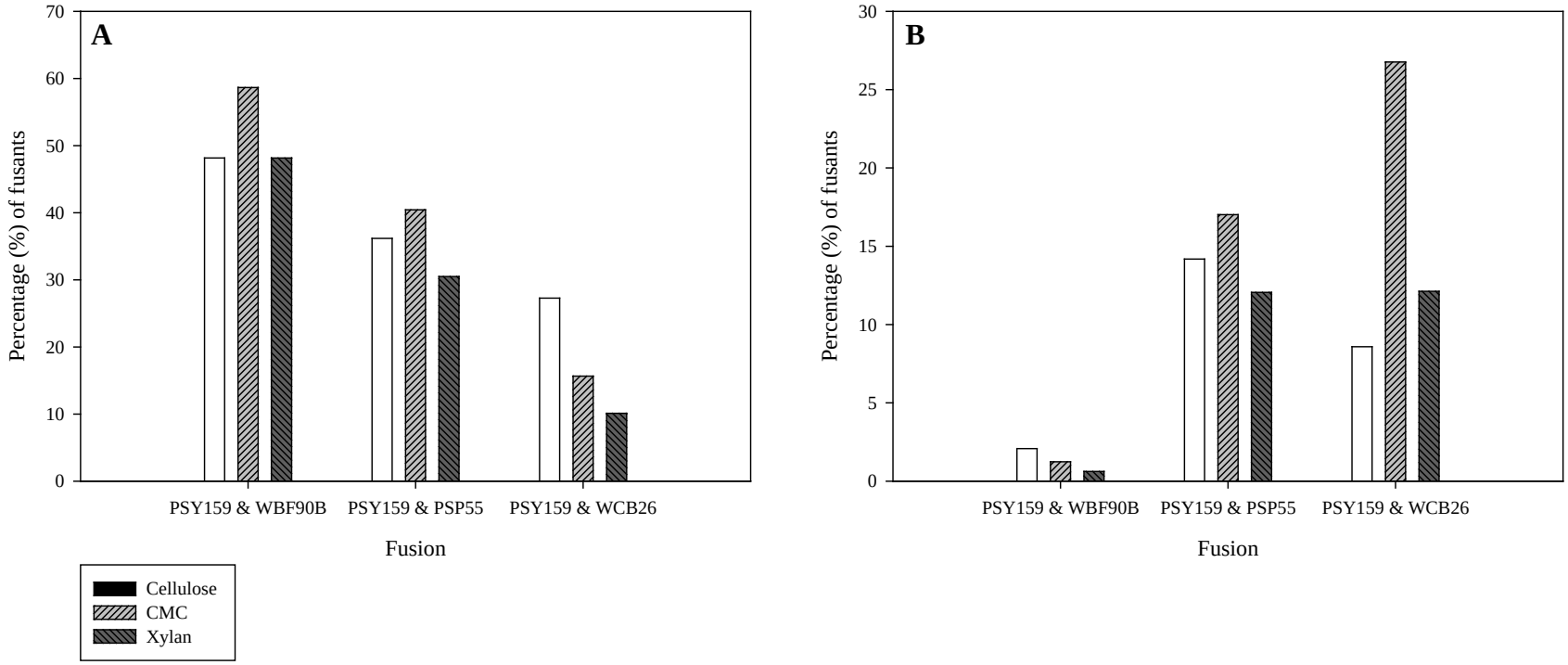


Figure V.3.3. Percentage of fusants that exhibit cellulase activity on Cellulose-Congo Red Agar (C-CRA), CMC-Congo Red Agar (CMC-CRA), and xylanase activity on xylan agar plates from each fusion experiment: between parents PSY159 & WBF90B, n=162; PSY159 & PSP55, n=141; PSY159 & WCB26, n=198 which showed A) improvement compared to both parents and showed B) decreased enzymatic activity compared to both parents.

V.4. Gas chromatographic analysis of sugars released from birchwood xylan by select fusants and their parents, PSY159 and WBF90B

The degradation products generated by fusants with improved cellulase and xylanase activity were compared to their parents to confirm the improvement of cellulase and xylanase activity. The sugars released from xylan degradation between the four fusants, FA1-14, FA2-4, FA9-13, FA9-14, and their two parents, PSY159 and WBF90B, from protoplast fusion experiment A, were examined by gas chromatography to observe any difference in release of sugars from xylan between fusants and between fusants and parents. Fusants, FA1-14, FA2-4, FA9-13, FA9-14, showed improvement in cellulase and xylanase activity compared to both parents and so were chosen for further characterization. Neutral sugars released during xylan degradation by fusants and parents were analyzed to determine if fusants and parents degraded xylan in a similar manner. Each fusant and parent was grown in 1% liquid birchwood xylan and the supernatant of samples was sampled every 48 hours for a total of 576 hours for analysis of sugars present in the plants (Blakeney et al., 1983), rhamnose, fucose, ribose, arabinose, xylose, mannose, galactose and glucose, present. Supernatant from an uninoculated control was analyzed over the course of the experiment as well. Growth of microorganisms was apparent on the side of the flasks while there was some xylan residue adhering to the side of the flask as well, the growth of the microorganisms was a distinct colour compared to the xylan. Data was not obtained between 192-432 hours due to errors with the instrument. There were column problems where the column got clogged from adding some of the aqueous phase, injection septum needed to be changed and other gas tanks had to be changed. Samples were not re-run because most of the sample evaporated by the time the problems were resolved. The pH of each culture was measured at each time

point using pH strips and was found to be stable at pH 8 throughout the entire experiment.

Sugars analyzed in this experiment were: rhamnose, fucose, ribose, arabinose, xylose, mannose, galactose and glucose because these are monosaccharides present in the plant (Blakeney et al., 1983). Trace amounts of rhamnose, galactose and glucose and large amounts of arabinose and xylose were present in the xylan starting material (Figure V.4.1.). Ribose, galactose and mannose were not observed at any time point in any of the supernatants even though rhamnose and galactose were observed in birchwood xylan. No data was obtained at 192 hours for PSY159 because the sample evaporated. Trace amounts of glucose were observed at most of the time points of the inoculated sample supernatants and were comparable to the amount in the uninoculated supernatants (Figure V.4.2.-V.4.7.). Less than 3.4 nmol/mg of fucose per milligram of xylan was present at various times points and only with fusants, FA1-14, FA9-13 and FA9-14 (Figure V.4.2.-V.4.7.). In the culture inoculated with FA1-14 at 96 hours, there was 3.3 nmol/mg of fucose present (Figure V.4.4.). At 96 and 144 hours in FA9-13, 1.58 nmol/mg and 0.65 nmol/mg of fucose was present, respectively (Figure V.4.6.). At 144 hours in FA9-14, 0.61 nmol/mg of fucose was present (Figure V.4.7.). 1.21 nmol/mg of rhamnose was observed in uninoculated control at 576 hours and was not present at any other time point for uninoculated control or inoculated cultures (Figure V.4.2.-V.4.7.). At 96 hours, trace amounts of arabinose was observed: 3.14 nmol/mg in uninoculated control (Figure V.4.2.-V.4.7.), 1.51 nmol/mg in FA1-14 (Figure V.4.4.) and 1.15 nmol/mg in FA2-4 (Figure V.4.5.). At 432 hour, WBF90B showed trace amounts, 0.09 nmol/mg, of arabinose (Figure V.4.3.).

For all time points in all of the inoculated and uninoculated controls, xylose was present (Figure V.4.2.-V.4.7.). For most time points, in all of the inoculated samples, the amount of xylose was less than the amount of xylose present in the uninoculated control (Figure V.4.2.-V.4.7.) indicating the possibility that the inoculated microorganism was consuming the xylose present. When the amount of uninoculated xylose was less than the amount in inoculated samples, xylanases were produced by the inoculated microorganism to degrade xylan. Therefore, an increase in the amount of xylose present was observed (Figure V.4.2.-V.4.7.).

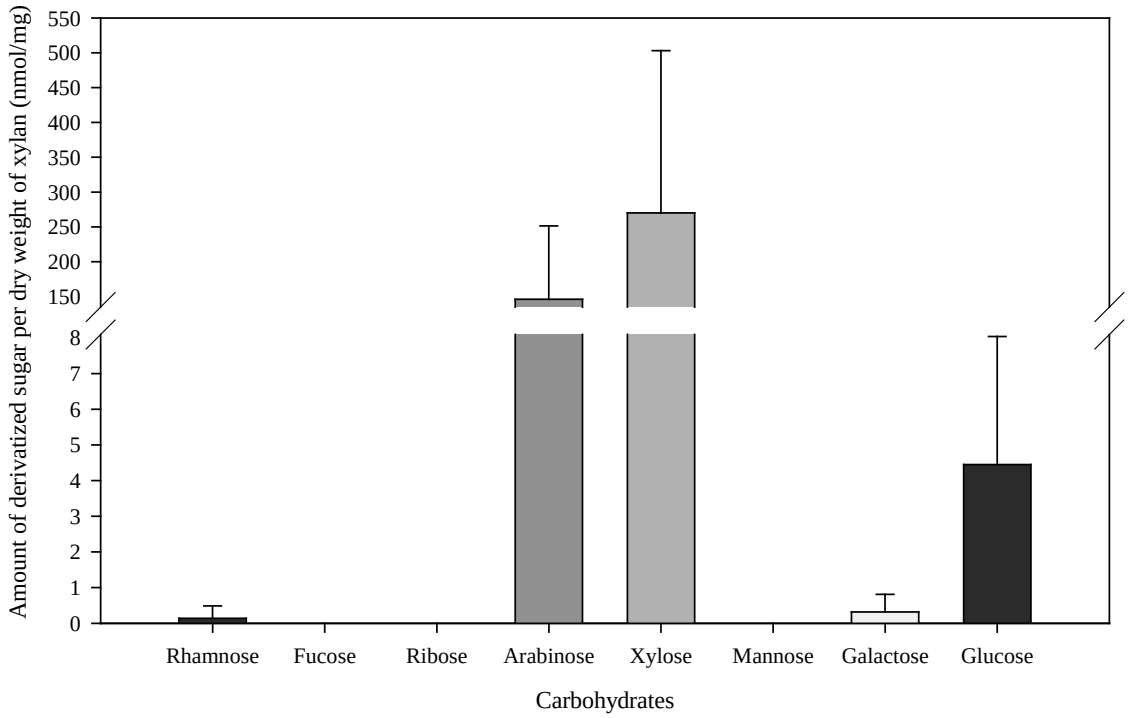
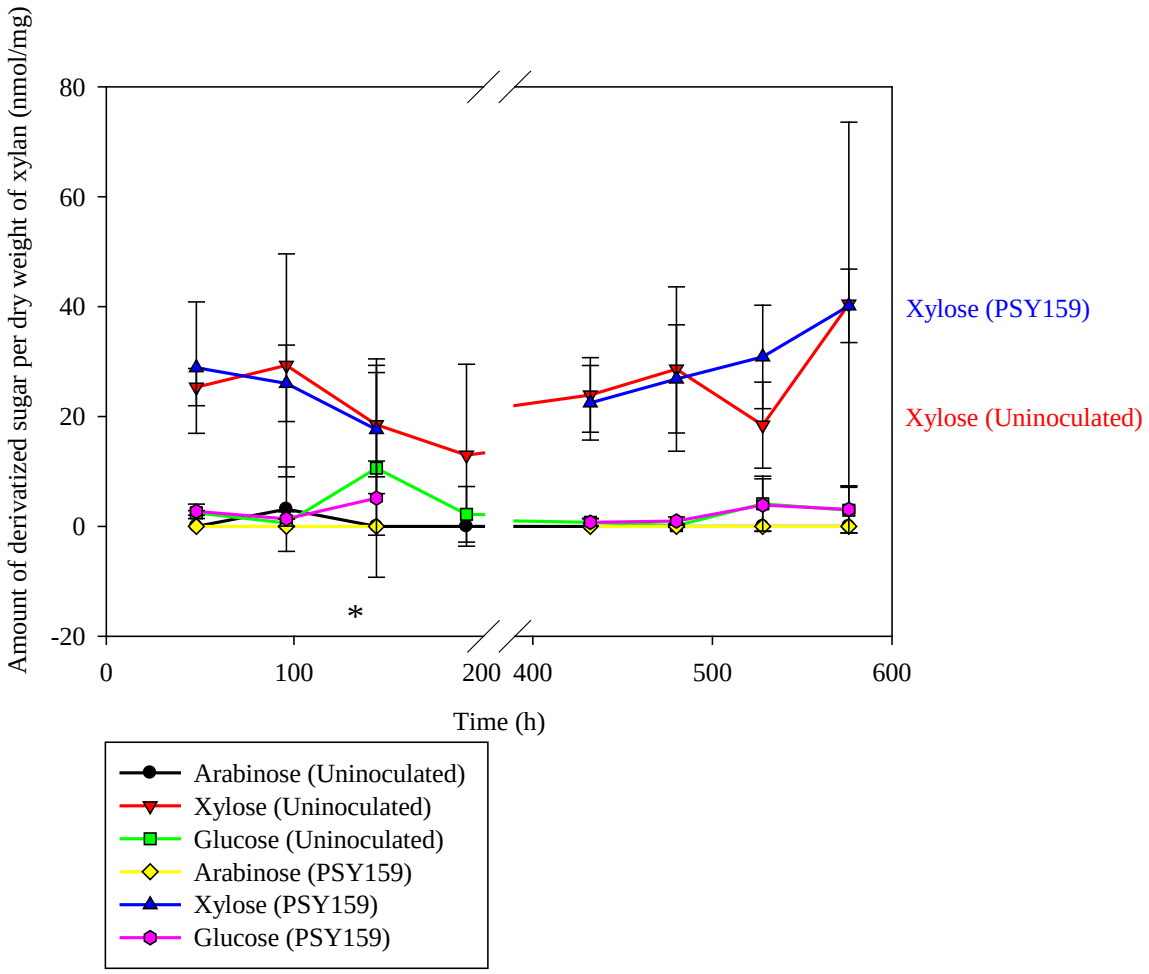


Figure V.4.1. Alditol acetate derivitization of sugars present in birchwood xylan pellet from uninoculated control determined by gas chromatography. Error bars indicate standard error of six replicates.



*No data was obtained at 192 hours for PSY159.

Figure V.4.2. Alditol acetate derivitization of sugars, arabinose, xylose and glucose present in culture supernatant of uninoculated control and PSY159 respectively, grown in birchwood xylan determined by gas chromatography. Six replicates for each 48 hour time point from 48-192, 432-576 hours were performed. Error bars indicate standard error of six replicates.

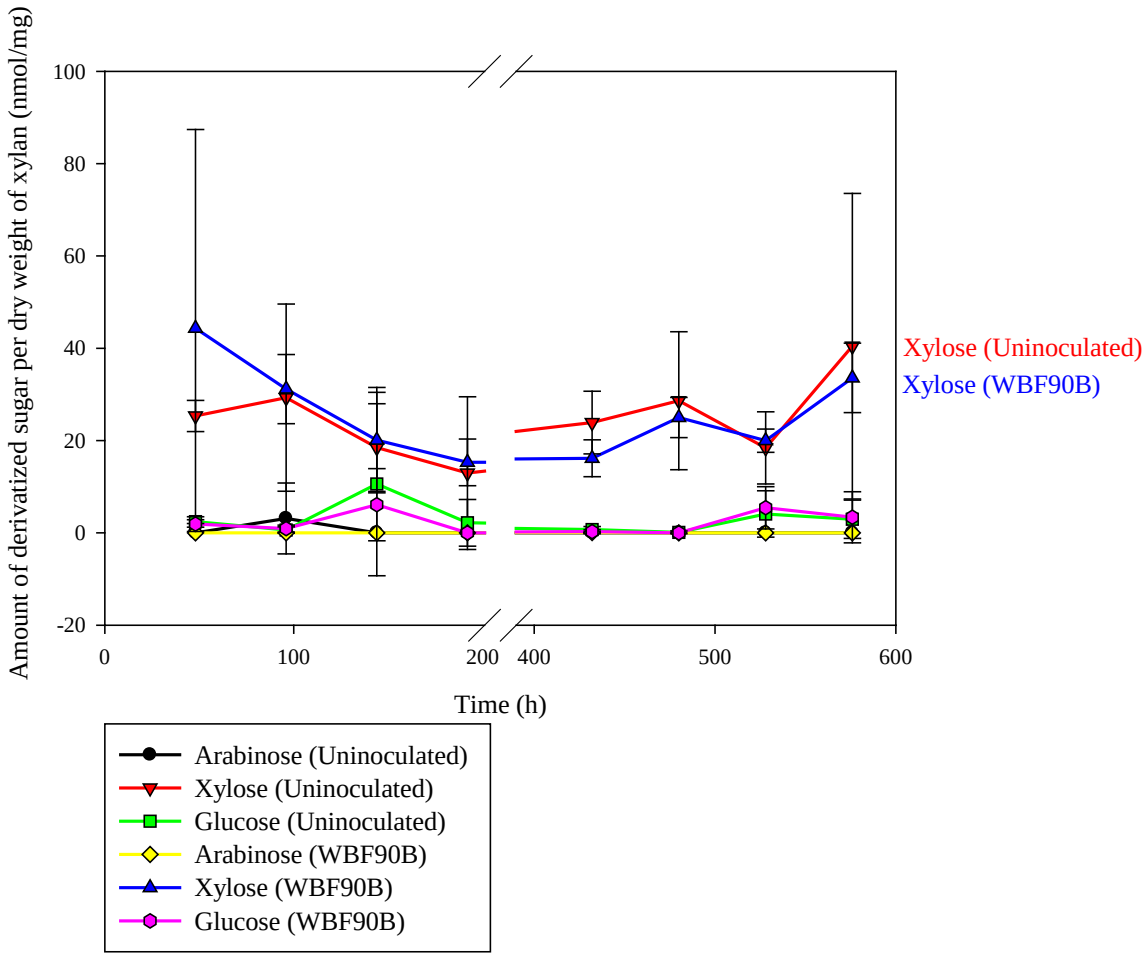


Figure V.4.3. Alditol acetate derivitization of sugars, arabinose, xylose and glucose present in culture supernatant of uninoculated control and WBF90B respectively, grown in birchwood xylan determined by gas chromatography. Six replicates for each 48 hour time point from 48-192, 432-576 hours were performed. Error bars indicate standard error of six replicates.

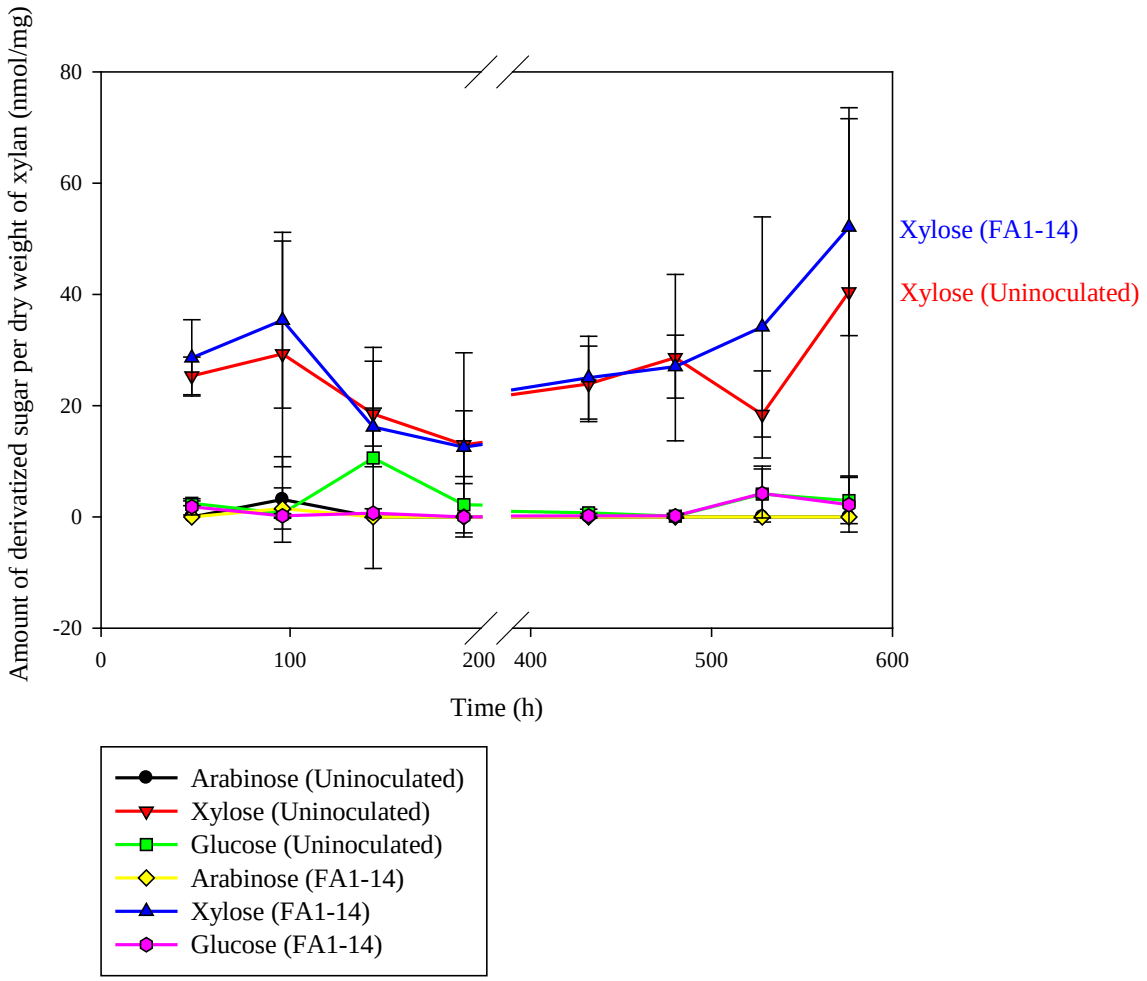


Figure V.4.4. Alditol acetate derivitization of sugars, arabinose, xylose and glucose present in culture supernatant of uninoculated control and FA1-14 respectively, grown in birchwood xylan determined by gas chromatography. Six replicates for each 48 hour time point from 48-192, 432-576 hours were performed. Error bars indicate standard error of six replicates.

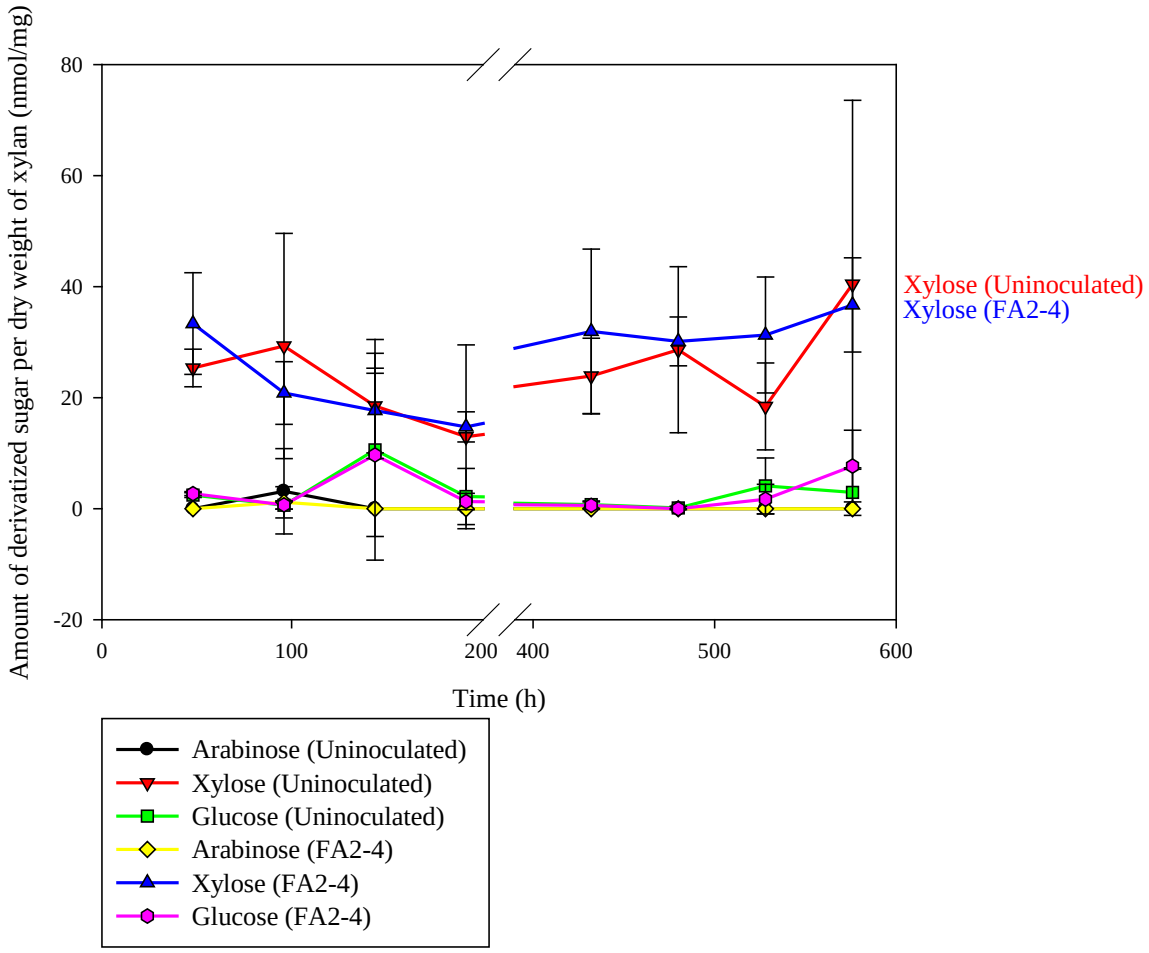


Figure V.4.5. Alditol acetate derivitization of sugars, arabinose, xylose and glucose present in culture supernatant of uninoculated control and FA2-4 respectively, grown in birchwood xylan determined by gas chromatography. Six replicates for each 48 hour time point from 48-192, 432-576 hours were performed. Error bars indicate standard error of six replicates.

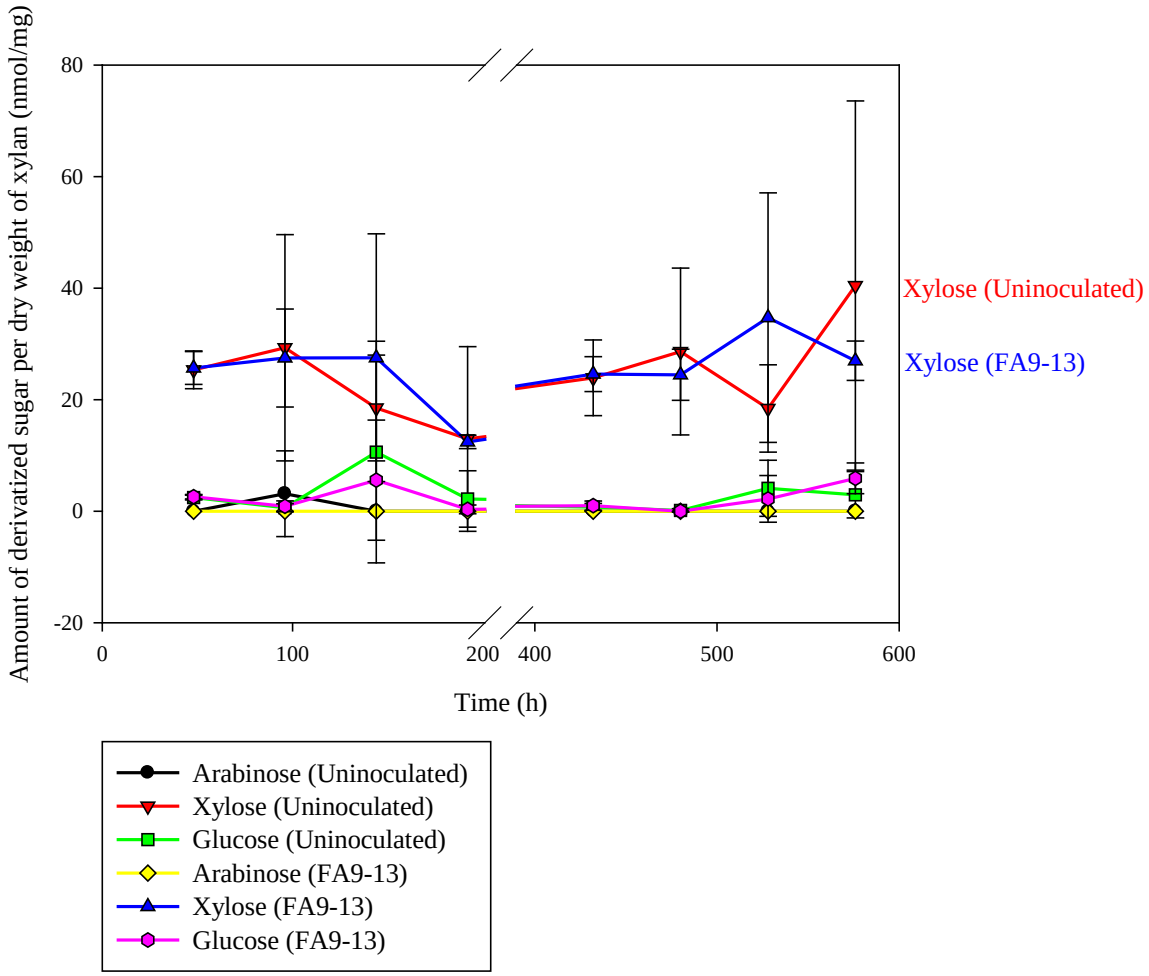


Figure V.4.6. Alditol acetate derivitization of sugars, arabinose, xylose and glucose present in culture supernatant of uninoculated control and FA9-13 respectively, grown in birchwood xylan determined by gas chromatography. Six replicates for each 48 hour time point from 48-192, 432-576 hours were performed. Error bars indicate standard error of six replicates.

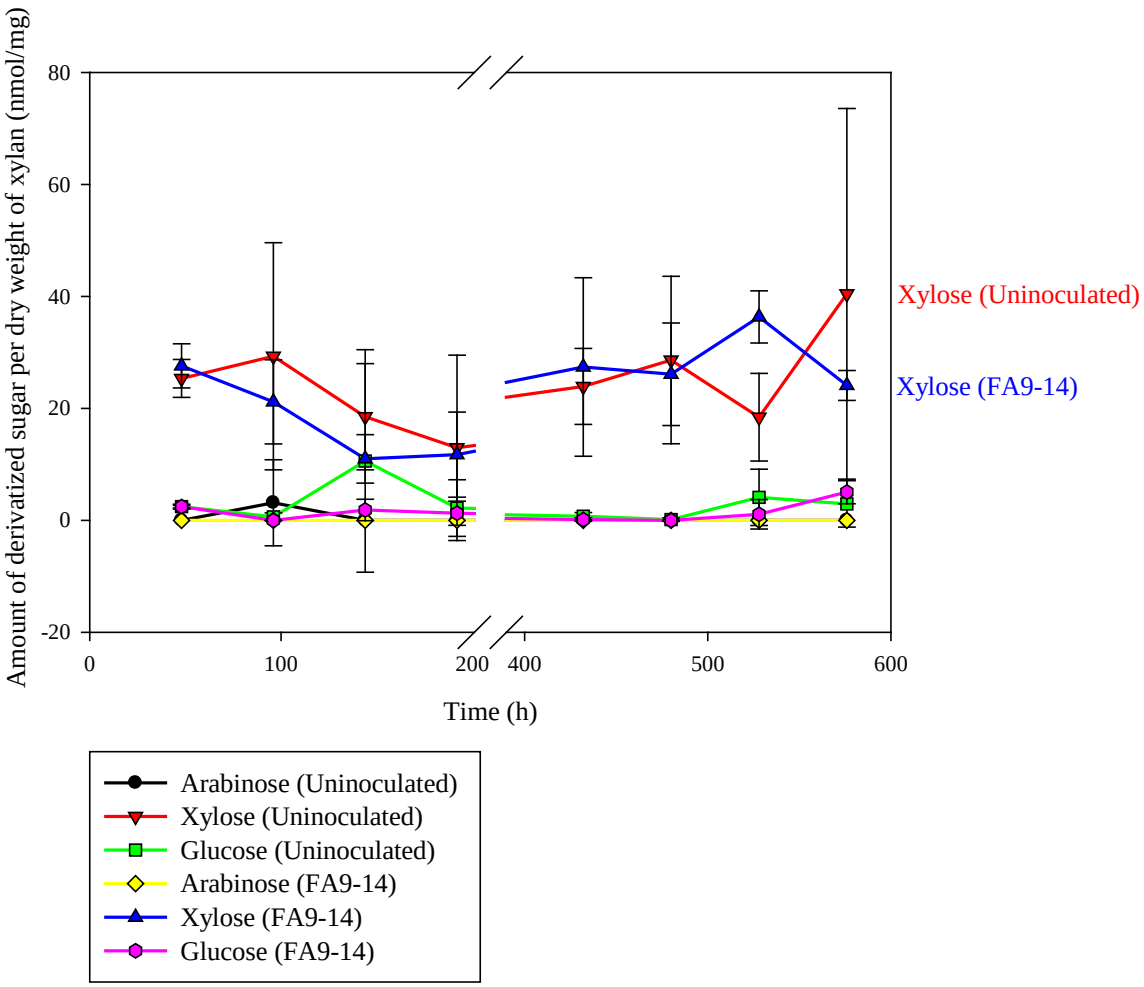


Figure V.4.7. Alditol acetate derivitization of sugars, arabinose, xylose and glucose present in culture supernatant of uninoculated control and FA9-14 respectively, grown in birchwood xylan determined by gas chromatography. Six replicates for each 48 hour time point from 48-192, 432-576 hours were performed. Error bars indicate standard error of six replicates.

V.5. 16S rRNA gene sequence analysis of fusant FA1-14

Fusant FA1-14 showed improvement in cellulase and xylanase activity compared to parents PSY159 and WBF90B. The 16S rRNA gene of FA1-14 was amplified by PCR, cloned, sequenced and compared to previously described sequences by Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) to determine the nearest relatives and if fusant FA1-14 is a novel strain compared to strains already described. The fusant FA1-14 was randomly selected for sequencing and the parental strains were not sequenced in this study.

The 1105 bp PCR amplicon of the 16S rRNA gene sequence of FA1-14 had 98% identity to 16S ribosomal RNA genes of various *Streptomyces* spp. (Table V.5.1.). Fusant FA1-14 16S rRNA gene was also compared to 16S rRNA gene sequences of: *Bacillus subtilis* subsp. subtilis 168, *Streptomyces avermitilis* MA-4680, *Streptomyces coelicolor* A3(2), and *Streptomyces viridosporus* strain NRRL 2414T (Table V.5.2.). *Bacillus subtilis* subsp. subtilis 168 is a Gram positive bacterium with low G+C content that is not a streptomycete. *Bacillus subtilis* (Kunst et al., 1997) was used to root the tree because all other strains were *Streptomyces* sp. The complete genome sequence is known for *Streptomyces avermitilis* MA-4680 (Ikeda et al., 2003) and *Streptomyces coelicolor* A3(2) (Bentley et al., 2002), and were therefore used for comparison. *Streptomyces viridosporus* strain NRRL 2414^{T} is a known lignocellulose degrader (Crawford et al., 1983; Ramachandra et al., 1988) and was also selected for comparison. Phylogenetic analysis was performed using the neighbour joining method (Saitou & Nei, 1987) (Figure V.5.1.). Fusant FA1-14 is a novel strain compared to previously identified strains because it is present in its own clade (Figure V.5.1.).

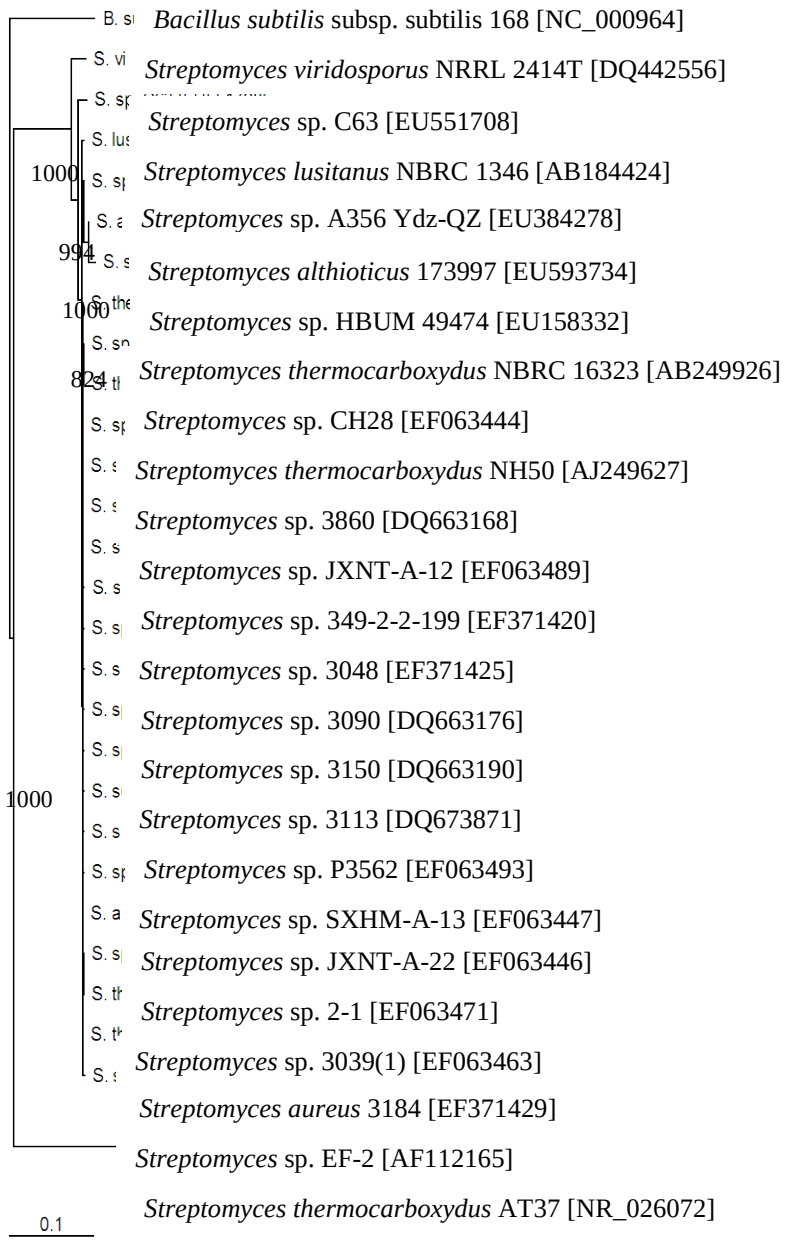


Figure V.5.1. Neighbour-joining tree based on partial 16S rRNA gene sequences of fusant FA1-14 with its nearest phylogenetic relatives. The numbers at the nodes specify the level of bootstrap support based on 1000 re-sampled datasets. The scale bar represents 0.1 nucleotide substitutions per site. The outgroup, *Bacillus subtilis* was used to root the tree. GenBank accession numbers are in the square brackets.

Table V.5.1. Comparison of fusant FA1-14 to GenBank matches with 98% identity

Match	GenBank accession number	Identity* (%)
Streptomyces sp. A356 Ydz-QZ	EU384278	1088/1105 (98%)
Streptomyces aureus 3184	EF371429	1088/1105 (98%)
Streptomyces sp. 3048	EF371425	1088/1105 (98%)
Streptomyces sp. 349-2-2-199	EF371420	1088/1105 (98%)
Streptomyces sp. JXNT-A-12	EF063489	1088/1105 (98%)
Streptomyces sp. 3860	DQ663168	1088/1105 (98%)
Streptomyces thermocarboxydus NBRC 16323	AB249926	1088/1105 (98%)
Streptomyces thermocarboxydus	AB098079	1088/1105 (98%)
Streptomyces thermocarboxydus NH50	AJ249627	1088/1105 (98%)
Streptomyces sp. P3562	EF063493	1087/1105 (98%)
Streptomyces sp. 2-1	EF063471	1087/1105 (98%)
Streptomyces sp. 3039(1)	EF063463	1087/1105 (98%)
Streptomyces sp. SXHM-A-13	EF063447	1087/1105 (98%)
Streptomyces sp. JXNT-A-22	EF063446	1087/1105 (98%)
Streptomyces sp. CH28	EF063444	1088/1106 (98%)
Streptomyces sp. 3113	DQ673871	1087/1105 (98%)
Streptomyces sp. 3150	DQ663190	1087/1105 (98%)
Streptomyces sp. 3090	DQ663176	1087/1105 (98%)
Streptomyces sp. EF-2	AF112165	1087/1105 (98%)
Streptomyces thermocarboxydus AT37	NR_026072	1087/1105 (98%)
Streptomyces althioticus 173997	EU593734	1088/1107 (98%)
Streptomyces lusitanus NBRC 13464	AB184424	1086/1105 (98%)
Streptomyces sp. HBUM 49474	EU158332	1086/1105 (98%)
Streptomyces sp. 1A01649	EF012124	1085/1105 (98%)
Streptomyces sp. C63	EU551708	1068/1086 (98%)

*Nucleotides identical to the nearest GenBank relative/query

Table V.5.2. Comparison of fusant FA1-14 to other organisms

Reference	GenBank accession number	Identity* (%)
Bacillus subtilis subsp. subtilis 168	NC_000964	640/787 (81%)
Streptomyces avermitilis MA-4680	NC_003155	1025/1074 (95%)
Streptomyces coelicolor A3(2)	NC_003888	1047/1080 (96%)
Streptomyces viridosporus NRRL 2414T	DQ442556	1034/1079 (95%)

*Nucleotides identical to the nearest GenBank relative/query

V.6. Comparison of 16S rRNA genes between parents and fusants from protoplast fusion events

Restriction fragment length polymorphism (RFLP) analysis was performed to compare the 16S rRNA gene of each fusant to its respective parents from the protoplast fusion events between PSY159 & WBF90B, PSY159 & PSP55, and PSY159 & WCB26. To our knowledge, no one has evaluated 16S rRNA genes after genome shuffling. Clones containing 1.5 kb PCR amplified 16S rRNA genes from select fusants, FA1-14, FA2-4, FA9-13, FA9-14, FB11-14 and FC17-4 and their respective parents were digested by *Hae*III to compare the banding patterns. The digested DNA was electrophoresed on a 3% agarose gel and showed most fusants were more like parent PSY159 (Figure V.6.1.). FA1-14, FA2-4, FA9-13, and FA9-14 showed a 200 bp band absent in parent WBF90B but was present in PSY159. Fusion B fusant, FB11-14, showed a similar banding pattern as both parents, PSY159 and PSP55. WCB26, one of the parents for fusion C was missing a 200 bp band but was present in the other parent PSY159 and the fusant FC17-4. Furthermore, the banding pattern of WCB26 was different than the others because there was a 312 bp band present. Therefore, FA1-14, FA2-4, FA9-13, FA9-14 and FC17-4 had the same banding pattern as PSY159 while FB11-14 had the same banding pattern as both parents, PSY159 and PSP55. Since the resolution was poor on 3% agarose gel and *Hae*III cuts into the vector multiple times, it was difficult to ascertain differences in 16S rRNA gene sequences. Therefore, to obtain better resolution, polyacrylamide gels were used and only the PCR amplicons of 16S rRNA genes were digested to distinguish any differences present between parents and fusants.

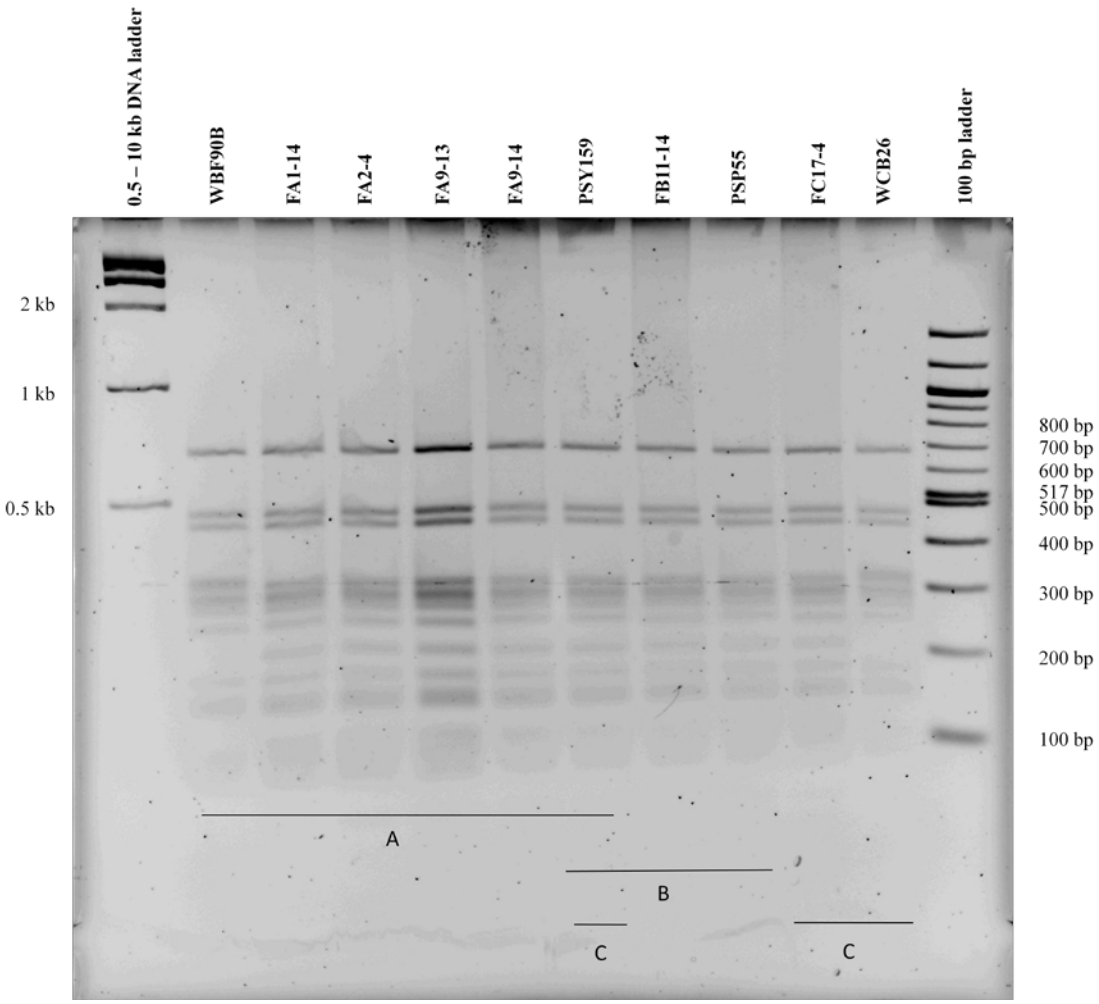


Figure V.6.1. Cloned 16S rRNA genes of fusants and parents from fusion experiments between parents PSY159 & WBF90B, PSY159 & PSP5 and PSY159 & WCB26 digested with *Hae*III and electrophoresed on a 3% agarose gel.

V.6.1. Comparison of 16S rRNA genes between parents and fusants using polyacrylamide gels

The differences in restriction fragment length polymorphisms of 16S rRNA gene sequences of parents and fusants after genome shuffling was compared to determine if genome shuffling occurred within the 16S rRNA gene. Polyacrylamide gel concentrations of 8%, 4% and 5% were used to resolve *Hha*I digested 16S rRNA gene amplicons. An 8% polyacrylamide gel was used to obtain resolution between 6-400 bp. *Hha*I digested 16S rRNA gene amplicons are greater than 400 bp, therefore resolution was not obtained (Figure V.6.1.A-B.). *Hha*I digested 16S rRNA gene amplicons were separated with a 4% polyacrylamide gel (Figure V.6.1.C-D.). To obtain resolution between 80-500 bp, 5% polyacrylamide gel electrophoresis of 16S rRNA gene amplicons was performed (Figure V.6.1.E-F).

The banding pattern of 16S rRNA digested amplicons of FA1-14, FA2-4, FA9-13, and FA9-14 were compared to parents PSY159 and WBF90B. The molecular weight of the bands could not be determined for the 4% polyacrylamide gel because there was unpolymerized polyacrylamide in the wells for the 4% polyacrylamide gel (Figure V.6.1.C.), thus only the banding patterns were considered. A band approximately 300 bp was missing in WBF90B but it was present in FA1-14, FA2-4, FA9-13, FA9-14 and PSY159 (Figure V.6.1.C.). The bands add up to approximately 1.5 kb for PSY159 and FA1-14, meaning the 16S rRNA gene of FA1-14 was the same as 16S rRNA gene of PSY159. The bands do not add up to 1.5 kb, the expected molecular weight for 16S rRNA amplified gene, for FA2-4, FA9-13, FA9-14 and WBF90B (Figure V.6.1.C.). Approximately 100 bp was missing for FA2-4, FA9-13, FA9-14 while WBF90B was missing 500 bp. The concentration of 16S rRNA gene amplicons for FA2-4, FA9-13,

FA9-14 was less concentrated than 16S rRNA gene amplicons of PSY159 or FA1-14. A faint 100 bp band was observed for more concentrated 16S rRNA gene amplicons of PSY159 and FA1-14. The 100 bp could be missing for FA2-4, FA9-13 and FA9-14 as a result of the concentration difference, resulting in very faint 100 bp bands or 100 bp had been digested into smaller fragments. Therefore, FA2-4, FA9-13, FA9-14 has the potential to be similar to PSY159.

FB11-14 was the fusant created between protoplast fusion of PSY159 and PSP55. FB11-14 banding pattern was similar to both parents as observed on the 4% polyacrylamide gel (Figure V.6.1.D.). The 5% polyacrylamide gel showed a faint band less than 100 bp present in both parents but not the fusant (Figure V.6.1.F.). Also, fusant FB11-14 was missing a band that was approximately 100 bp when other bands for FB11-14 were added but it did equal the theoretical size of the 1.5 kb amplified 16s rRNA gene (Figure V.6.1.F.). Likewise, through observation of the 8% (Figure V.6.1.B.) and 4% (Figure V.6.1.D.) polyacrylamide gels, the smallest band of the parents, PSY159 and PSP55, was present but this band was absent for FB11-14. The concentration of amplified 16S rRNA gene for FB11-14 was less than both parents. The 100 bp band could be present but it was too faint to be observed or it was digested into smaller fragments. The banding pattern between both parents PSY159 and PSP55 was the same, but FB11-14 did not have the 100 bp fragment, meaning FB11-14 was different than both parents.

Fusant FC17-4 was derived by protoplast fusion between PSY159 and WCB26. The banding pattern between WCB26 and fusant FC17-4 was different and visualized more clearly on the 5% polyacrylamide gel (Figure V.6.1.F.). Fusant FC17-4 and parent

PSY159 contained a ~300 bp band which was not present in parent WCB26 (Figure V.6.1.D.). All the bands add up to 1.5 kb for PSY159 (Figure V.6.1.D.). WCB26 16S rRNA gene does not add up to 1.5 kb because about 400 bp was missing (Figure V.6.1.D). FC17-4 was missing about 250 bp of DNA in the 4% polyacrylamide gel (Figure V.6.1.D.). In Figure 6.1.1.F. there appeared to be a 150 bp present. Therefore, only 100 bp was missing for FC17-4. However, this missing 100 bp could be too faint to see because the concentration of DNA for FC17-4 was less than the parents. Thus, the banding pattern of FC17-4 was more like PSY159 if this unseen faint 100 bp was taken into account.

In conclusion, the differences seen in the polyacrylamide gels were similar to the differences seen in the agarose gel wherein the 16S rRNA gene of most of the fusants, FA1-14, FA2-4, FA9-13, FA9-14 and FC17-4, most resembled PSY159. The 16S rRNA gene sequence of FB11-14 was not the same as both parents because it did not have the same banding pattern. Thus, 16S rRNA genes are not protected from genome wide shuffling.

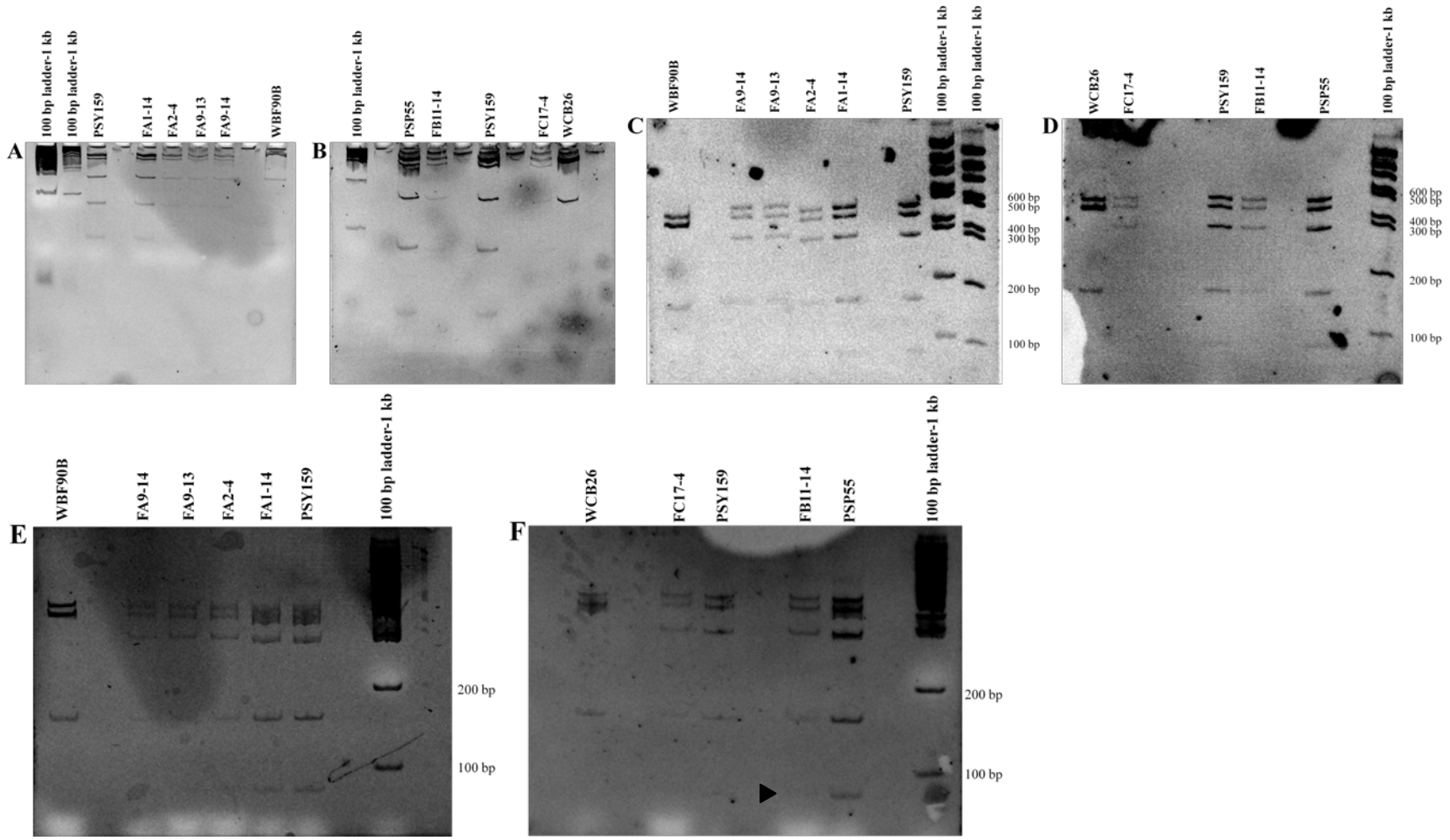


Figure V.6.1.1. 16S rRNA genes digested with *Hha*I and electrophoresed on A, B) 8% polyacrylamide gel, C, D) 4% polyacrylamide gel, and E, F) 5% polyacrylamide gel. Digested 16S rRNA genes were amplified for fusants and parents from fusions between parents A, C, E) PSY159 & WBF90B, B, D, F) PSY159 & PSP55, and PSY159 & WCB26. ▶ Indicates missing band.

VI. DISCUSSION

VI.1. Screening of actinomycetes isolated from soil for cellulase and xylanase activity

Several soil types were used to isolate environmental actinomycetes and these isolates were subsequently assayed for cellulase activity utilizing *in vitro* cellulose, CMC and xylan degradation assays. The soils used in this study were: University of Ontario Institute of Technology (UOIT) forest rhizosphere soil, UOIT forest bulk soil, conservation area rhizosphere soil, conservation area bulk soil and potting soil.

Rhizosphere refers to the soil close to the plant root system (Marschner et al., 2001). In contrast to rhizosphere soil, bulk soil is obtained away from the plant root system (Basil et al., 2004). The rhizosphere isolates had slightly larger clearance zones on CMC and xylan assays than bulk soil isolates (Figure V.2.2.B. & Figure V.2.2.C.). However, there was virtually no difference between the percentage of rhizosphere isolates that exhibited cellulase and xylanase (65%) activities compared to bulk isolates (63%) (Figure V.2.3.). A greater observed percentage of rhizosphere isolates able to degrade cellulose, CMC and xylan may be due to the greater diversity of actinomycetes generally found in rhizosphere soil compared to bulk soil (Basil et al., 2004). The observed difference in diversity is influenced by the root exudates released by the plants which affects the soil of the rhizosphere attracting microorganisms to the area (Smalla et al., 2001) which are able to degrade polymeric organic matter like lignocellulose (Lynch, 1990). There would be less plant litter in the bulk soil because it is farther away from the plant. The less lignocellulose present, the less likely the inhabitants of the soil would need to produce the enzymes to degrade lignocellulose. Although, this may be true, the

enzymes produced by specific microorganisms found in bulk soil can be effective at degrading the scarce lignocellulose for use as a food source.

Potting soil was chosen for cultivation of actinomycetes that may produce cellulases and xylanases because it is enriched with organic nutrients and is likely to house an abundant population of actinomycetes. Isolates from potting soil had a greater percentage of isolates with larger clearance zones screened on cellulose, CMC and xylan agar (Figure V.2.2.) compared to other rhizosphere or bulk soil isolates. A higher percentage (87%) of isolates with broad spectrum activity was isolated from potting soil compared to other cultivated isolates (Figure V.2.3.; Table V.2.1.). Actinomycetes isolated from potting soil dominated the list of isolates that produced large clearance zones on cellulose and xylan containing agar (Table V.2.1.). One possible explanation for this is potting soil was enriched with humus by the manufacturer. Humus is created from decaying biomatter such as plants and other dead organisms (Waksman, 1925; MacCarthy, 2001). Actinomycetes produce extracellular hydrolytic enzymes to degrade humus, enabling them to compete for nutrients and use it as a carbon source to grow (Dari et al., 1995).

Environmental isolates were screened on cellulose, CMC and xylan agar plates by spotting 5 µL of spores onto agar plates containing each substrate. The concentrations of spores could have varied when isolates were screened because the concentrations of spores in the stock solutions were not quantified. This was a preliminary screen and therefore it would be difficult to determine the concentration of every spore stock for mass screening with equal inoculum. To ensure the best isolates are indeed the ones with the best cellulase and xylanase activity, the stock concentration for the best isolates could

be counted and used to re-screen each substrate with equal concentration of spores as inoculum. This re-screening process would allow verification that the isolates do indeed display the best cellulase and xylanase activities compared to other isolates.

Another limitation to screening the strains for cellulase and xylanase activities is that the growth rate can differ between strains on any given substrate. As the growth rate and utilization of substrate for the previously uncharacterized environmental isolates is unknown, an equal inoculum would standardize the assays accounting for the possibility that each isolate produces enzymes with different degradative catalytic activities. Further investigation of the cellulase and xylanase activities of the environmental isolates screened in this study will determine which isolates are candidates for future applications.

In conclusion, actinomycete isolates from potting soil isolates were able to produce cellulases and xylanases that degraded cellulose, CMC and xylan producing some of the largest clearance zones observed in the assays.

VI.2. Strain improvement through protoplast fusion

Strains isolated from potting soil, rhizosphere and bulk soil showed large clearance zones on cellulose, CMC and xylan. Strain PSY159 was the isolate best able to degrade cellulose, CMC and xylan and therefore, was chosen for strain improvement through protoplast fusion, a genome shuffling technique (Hopwood et al., 1977), in an attempt to enhance cellulase and xylanase activity.

Fusants were produced from the protoplast fusion of PSY159 (parent) with strains: WBF90B, an isolate that degraded cellulose and CMC very thoroughly; and PSP55 and WCB26 which were shown to have cellulase and xylanase activities. Parents and fusants were patched onto cellulose congo red, CMC congo red and xylan agar to screen for cellulase and xylanase activities, respectively. The amount of activity was

determined by measuring the size of clearance zones created from the degradation of the substrate, from the edge of the colony to the edge of the clearance zone. Four fusants, FA1-14, FA2-4, FA9-13, and FA9-14 from the fusion between PSY159 and WBF90B had greater clearance zones on the assays compared to both parents (Table V.3.1.). For fusion between PSY159 and PSP55, only one fusant, FB11-14 had greater clearance zones on the assays compared to both parents (Table V.3.1.). The fusion between PSY159 and WCB26 yieled only one fusant, FC17-4 which was found to have improved cellulase and xylanase activities compared to both parents (Table V.3.1.).

Two fusants, FA9-13 and FA9-14, with the largest clearance zones based on measuring from the edge of the growth streak to the edge of the clearance zone, when the growth was not considered, were incubated for 40 hours compared to others that were incubated for at least 26 hours to maximum of 40 hours. Longer incubation time could have contributed to larger clearance zones for FA9-13 and FA9-14 since these two fusants had greater clearance zones than most of the other fusants (Table V.3.1.).

Six fusants, out of 501 screened, with the best broad spectrum activity, cellulase and xylanase activity, were identified out of the three fusion events (Table V.3.1.). These six fusants were identified by measuring the clearance zones from the fusant to the edge of the zone. The inoculum concentration was not controlled during the agar assay screening. Parent strains were used as an internal control to ensure that the fusants could be compared to parents on the agar assay plates. However, initially, the growth of the strain was not accounted for through this analysis method. The data was subsequently re-analyzed to account for growth of the parents and the fusants by normalizing the data through subtraction of the growth of the organism from the clearance zones that they

produced. The normalization allowed for a better comparison of improvements in cellulose or xylan degradation between parents because the greater the streak of the strain on the agar plate, the more enzymes produced. Therefore, the colony growth should be accounted for, to remove error associated with large clearance zones arising from more biomass.

In each of the agar plates, the sole carbon source was cellulose or xylan. If the strain could not utilize the available carbon source then it would not grow. Certain strains may be able to grow at a faster rate if they were more capable of utilizing cellulose or xylan. An equal concentration of spores patched onto the assay plates would normalize the data to account for growth.

Only one fusant, FA9-15, had a clearance zones greater than two millimetres larger compared to both parents on all three assays when the growth was also accounted for (Table V.3.1.). Also, only two, FA9-13 and FA9-14, of the original six fusants with the best cellulase and xylanase determined when the growth was not measured, had improved cellulase and xylanase activity when compared to both parents on all three substrates when the growth of the strain was analyzed (Table V.3.1.). Therefore, the growth has an effect on the clearance zones produced because the other four original strains, FA1-14, FA2-4, FB11-14 and FC17-4, did not have larger clearance zones than the parents when the growth was accounted for (Table V.3.1.).

Protoplast fusion between PSY159 and WBF90B had the greatest percentage of fusants, 162 screened, that showed improvement over their parental strains and the lowest percentage of fusants with decreased activity in all three assays compared to both parents (Figure V.3.3.). PSY159 had large clearance zones on all three assays, cellulose, CMC

and xylan. WBF90B produced small zones on cellulose and CMC, but degraded the cellulose completely resulting in a clear but less diffusible zone observed and had no xylanase activity. The protoplast fusion event that produced the greatest percentage of isolates with improved cellulase and xylanase activity was between an isolate, PSY159, that produced a large clearance zone due to diffusible enzymatic activity and an isolate, WBF90B, that produced cellulase that degraded the cellulose completely, but the cellulase was either less diffusible in the agar, or secreted at low levels as evidenced by a smaller clearance zone.

A greater percentage of improved cellulase and xylanase activity and a lower percentage of decreased enzymatic activity compared to both parents were observed from fusants between two potting soil isolates, PSY159 and PSP55, compared to fusants from PSY159 and WCB26, an isolate cultivated from conservation bulk soil (Figure V.3.3.). When WCB26 was initially screened on cellulose, CMC and xylan agar, WCB26 had greater cellulase and xylanase activity than PSP55. Genome shuffling is random (Hopwood & Wright, 1978; Sankoff & Goldstein, 1989) and therefore, the reason why one strain had a greater percentage of improvement than another strain cannot be controlled.

The agar diffusion method was used as a preliminary screen to determine which fusants had improved cellulase and xylanase activities compared to parent strains. However, verification of the improvement by doing replicate assays on cellulose, CMC and xylan agar would validate these initial screens and allow for statistical analysis. Screening a large number of fusants on replicate assays would be time intensive because many agar plates would be required. Also, when streaking fusants on the assay plates,

spacing the streaks further apart would allow a more accurate measurement of the clearance zones because some of the clearance zones overlapped. For some plates, the clearance zones overlapped because the incubation time was longer than other plates. The plates where the clearance zones overlapped were the first batch of plates screened and the incubation time had not yet been determined. Subsequently, the incubation time was then decreased for other plates to prevent overlapping of clearance zones. When rescreening, incubating all the plates the same amount of time would allow clearance zones between fusants to be compared more accurately. Since an internal control, where parents were streaked onto the same plate as the fusants, was used, the difference in incubation time was not considered to affect the comparison between improvements of fusants and parents.

Agar diffusion data presents a good method of screening but does not give any indication of the utilization of degradation products. Therefore, fusants, FA1-14, FA2-4, FA9-13, FA9-14, and their parents, PSY159 and WBF90B, were used in an analysis to compare the release of sugars from the degradation of 1% (w/v) birchwood xylan. PSY159 and WBF90B and their and fusants were used because these fusants had better cellulase activity than fusants from protoplast fusion between PSY159 and PSP55 or PSY159 and WCB26. The neutral sugars released from the degradation of xylan were analyzed by gas chromatography. In order to do so, the neutral sugars were derivatized into alditol acetates (Blakeney et al., 1983).

The samples were derivatized and set up to be analyzed by the instrument every 48 hours. Due to several technical problems with the gas chromatographer, no data was

collected between time points 192 hours and 432 hours as the samples prematurely evaporated.

Xylan was washed several times before incorporation into the media, but free xylose remained in the liquid media. Therefore, different xylan sources other than birchwood xylan should be investigated. If birchwood xylan is used, xylan must be washed several more times to remove the free xylose and other free sugars that are present. An anthrone assay could be used to monitor the amount of free xylose released at each wash step. The uninoculated xylan-containing medium should be derivatized as a control to determine the amount of sugars present. The large amount of xylose and glucose present in the birchwood xylan, even after washing, skewed the results because in the uninoculated flasks a large amount of xylose was present (Figure V.4.2-7.). The increase in the concentration of free monosaccharides in the inoculated flasks would not be due to degradation of the xylan if monosaccharide continually leeches from the xylan.

Each of the six strains was inoculated into duplicate flasks; samples were taken from each flask and aliquoted into three replicate tubes for derivatization. Therefore, for each sample there were six replicates. The uninoculated control flasks showed large variation between time points due to sugars leaching from the xylan over the course of the experiment. The use of more than two biological replicates and thoroughly washed xylan substrate would have given more representative data.

In order to better quantify the level of strain improvement achieved in select fusants, growth of each parent and fusant should have been followed during the time course experiment. In doing so, the growth of the microorganism could have been related to the amount of degradation observed. Since these microorganisms grow on the substrate

by producing mycelia which penetrate the substrate and release degradative enzymes, a sample could be taken from the inoculated cultures and the cells pelleted from the supernatant by centrifugation. The microbial biomass could then be determined by subtracting the dry weight of the inoculated pellet from the dry weight of the pellet of the uninoculated control. The difference in the dry weight of the pellets would indicate the amount of biomass in the inoculated cultures. A growth curve of each microorganism could then be compared to the amount of xylose and other sugars released over time. Xylanase activity assays (Chen et al., 1997) could also have been performed for each time point so that the amount of xylose in the culture could be correlated to the growth of each microorganism to demonstrate xylanases were responsible for the monosaccharides present in the cultures.

To further confirm degradation of xylan by extracellular xylanases produced by the microorganisms, the culture supernatant could be examined for the presence of xylanases released by each strain through zymogram analysis (Nakamura et al., 1993) using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). This would permit further characterization of the improved strains compared to the parental strains.

The culture flasks used for the analysis of xylan degradation by gas chromatography did not contain exactly 1% birchwood xylan because birchwood xylan was insoluble and when the medium was dispensed into each flask, there was a possibility that each flask did not contain an equal amount of birchwood xylan. Therefore, due to the variability of xylan present in each culture flask, the amount of free sugars released during substrate degradation could not be accurately determined. To

alleviate this problem, washed xylan could be dried to constant weight and a known amount could be added into each flask prior to the addition of the liquid medium.

The amount of sugars observed for each strain was different at different time points. Xylanases were produced by the microorganisms to degrade the xylan into xylose which was observed when the amount of xylose in the inoculated cultures was greater than the amount of xylose in the uninoculated controls. There was no fusant culture with a very large amount of xylose compared to the parents. Determination of the timing and quantity of xylose released in the culture would allow for the extraction of xylose for application of this sugar for fermentation into bioethanol. Therefore, utilization of xylan degradation products by a microorganism will impact its use in industrial processes. In conclusion, the xylan degradation profile was different for all strains analyzed.

VI.3. Limitation to congo red assays to examine cellulose degradation

In order to assess cellulose degradation, congo red dye was incorporated into cellulose agar. Congo red is a dye that binds to ß-1-4-glycosidic linkages of cellulose (Teather & Wood, 1982). A clearance zone is formed which is visible around the growth of the isolate due to the release of the congo red dye as cellulose or CMC is degraded. The possibility that the dye was decolourized by the microorganisms assessed in this assay can be ruled out because the clearance zones were opaque to translucent, indicating the ß-1-4-glycosidic linkages were broken and congo red could not bind to it.

Instead of incorporating congo red into the media, the plates could be stained with congo red, destained with NaCl and stained with HCl using Teather and Wood's (1982) method to view clearance zones. Staining with HCl gives a better contrast of zones because congo red is a pH dependent dye and turns blue under acidic conditions (Mera &

Davies, 1984). If the clearance zone is present after staining with HCl, then the zone was not created due to dye decolourization or a pH effect.

VI.4. 16S rRNA gene sequence analysis of FA1-14

In attempts to generate strains improved for degradation ability, protoplast fusion was performed. FA1-14 was created from the protoplast fusion between a strain cultivated from potting soil, PSY159 and a strain cultivated from UOIT forest bulk soil, WBF90B. PCR amplification of the 16S rRNA gene of fusant FA1-14 was performed and sequenced to compare previously identified 16S rRNA genes for taxonomic identification of FA1-14. FA1-14 was chosen to be sequenced because cellulase and xylanase activity was improved in comparison to both parental strains. Phylogenetic analysis of FA1-14 placed it in its own clade (Figure V.5.1.) indicating it is a novel isolate compared to previously described strains. The parental strains were not sequenced and FA1-14, may not be novel in comparison its parents. In fact, RFLP analysis showed that FA1-14 has the same gene banding pattern as PSY159 using the 16S rRNA gene.

VI.5. Comparison of 16S rRNA gene sequences between parental strains and fusants through restriction fragment length polymorphism (RFLP)

Since the 16S rRNA gene is often used as a molecular clock (Woese, 1987), it was of interest to determine if protoplast fusion affected this gene because this has not been studied to our knowledge and gene shuffling involving conserved genes is rare (Conant & Wagner, 2005). The 16S rRNA genes of select fusants were compared to the 16S rRNA gene of their parents through restriction fragment length polymorphism (RFLP).

The 16S rRNA genes cloned into pGEM T-Easy vector were digested with *Hae*III and electrophoresed on an agarose gel. The banding pattern of fusants from the fusion

between PSY159 and WBF90B and PSY159 and WCB26 seemed to be similar to PSY159 (Figure V.6.1.). Fusant FB11-14 had the same banding pattern as both parents. Analysis was difficult because *Hae*III cuts pGEM T-Easy multiple times resulting in many bands present on the gel.

In an attempt to increase the resolution of RFLP analysis, polyacrylamide gels were used to resolve the fragments generated by *Hha*I digested PCR-amplified 16S rRNA genes. Similar results were obtained between the agarose gel and the polyacrylamide gels: fusants from the two protoplast fusions between parents, PSY159 and WBF90B, and PSY159 and WCB26 had similar banding pattern as PSY159.

FB11-14, a fusant from the fusion of PSY159 and PSP55, had the same banding pattern as both parents based on the agarose gel analysis (Figure V.6.1.). While the RFLP pattern of parental strains PSY159 and PSP55 were identical by both gel types, the RFLP pattern of fusant, FB11-14 differed when examined by polyacrylamide gel electrophoresis. Analysis of the 5% polyacrylamide gel showed FB11-14 was missing a 100 bp band that was observed in the banding pattern of both parents (Figure V.6.1.F.). However, the addition of the band sizes for FB11-14 did not equal to 1.5 kb, the size of the PCR amplified 16S rRNA gene, and 100 bp was missing from the 1.5 kb (Figure V.6.1.F.). Similarly, a small band was absent for FB11-14 on 8% (Figure V.6.1.B.) and 4% (Figure V.6.1.D.) polyacrylamide gels whereas the small band was present for both parents. Therefore, FB11-14 had a different banding pattern than both parents meaning the 16S rRNA gene sequence of the fusant was different than both parents. In conclusion, genome shuffling can occur in conserved 16S rRNA genes.

The restriction enzymes *Hae*III and *Hha*I are tetracutters that recognize GC rich recognition sites (Fermentas, 2010). Actinomycetes are Gram-positive bacteria with a high GC content. *Hae*III and *Hha*I have recognition sites in the 16S rRNA gene of Gram-positive bacteria (Pukall et al., 1998). *Hae*III has been used to differentiate actinomycetes such as *Streptomyces* (Steingrube et al., 1997). The use of restriction enzymes that recognize GC rich DNA sequences allowed for the analysis of different fragment patterns between parents and fusants by RFLP.

The smaller bands on polyacrylamide gels were not clearly observed because the DNA concentration loaded onto the gel was higher for the parents than the DNA concentration of the fusant, FB11-14. Further analysis of fusant, FB11-14 and other fusants compared to parents with an equal concentration loaded on polyacrylamide gels will be necessary to verify genome shuffling in 16S rRNA genes.

VI.6. Extraction of genomic DNA from strain WCB26

One of the problems experienced during this project was in extraction of the genomic DNA using the method of Aljanabi and Martinez (1997) for environmental isolate WCB26, a parent used in protoplast fusion C along with PSY159. Several methods including colony PCR (Ishikawa et al., 2000), DNAzol extraction (Klein et al., 1997), and a method to extract DNA from soils (Zhou et al., 1996) were used to extract the genomic DNA from this isolate. Colony PCR and DNAzol extraction failed. Ultimately, using the SDS-based DNA extraction method of Zhou et al. (1996), the WCB26 culture lysed and genomic DNA was successfully extracted for 16S rRNA gene amplification.

VI.7. Conclusion

Soil isolates with cellulase and xylanase activity were identified and subjected to strain improvement through protoplast fusion which resulted in strains with better cellulase and xylanase activity than the cultivated soil isolates.

Genome shuffling was shown to occur in the 16S rRNA gene of one of the fusants compared to parental strains through RFLP analysis. Further examination of the 16S rRNA gene of other fusants will verify if genome shuffling occurs in conserved 16S rRNA gene and implications in the taxonomic identification of microorganisms.

VI.8. Future directions

Restriction fragment length polymorphism analysis of a fusant generated during strain improvement by protoplast fusion indicated that 16S rRNA genes may not be protected from genome shuffling. It will be of interest to pursue this by generating multiple sets of fusants with different parental strains to replicate the phenomenon.

The cellulase and xylanase enzymes produced by the isolates described in this thesis could be purified and characterized. Protein characterization could include the sequencing of the purified protein by mass spectrometry. The gene corresponding to the protein of interest could then be cloned using a reverse genetics approach and characterized. The biochemical and physical properties of each purified enzyme such as optimum temperature, pH stability and enzyme kinetics, K_m , V_{max}, and k_{cat} by measuring the activity of purified enzyme could be determined (Chen et al., 1997; Lee et al., 2006). The applicability of degradative enzymes produced by the environmental isolates to industrial applications such as bioethanol production could then be determined with the knowledge of the parameters for the best optimal activity of the enzymes.

Part 2: Mining the soil metagenome for cellulases

VII. ABSTRACT

A metagenomic library provides the opportunity to exploit microorganisms for their extensive range of metabolites without culturing them. This method can be used to discover novel biocatalysts from previously unstudied microorganisms.

In this study, enrichment cultures with potting soil, UOIT forest rhizosphere and UOIT forest bulk soil inoculated in 1% cellulose and 1% CMC media were grown for 10 months. Cells from potting soil enriched with 1% cellulose were used to create a functional metagenomic library which was used to screen for cellulase genes on carboxymethyl cellulose containing medium. One cellulase producing clone was identified out of 1,920 clones screened.

VIII. INTRODUCTION

Metagenomics is a technique that can be used to discover novel cellulases. Metagenomics, also known as environmental genomics or community genomics, is the study of the total complement of DNA from an environmental source (Handelsman et al., 1998).

Presently, only a small portion of the estimated diversity in the microbial world has been uncovered. Many microorganisms are unculturable as pure cultures under standard laboratory conditions (Handelsman et al., 1998). It has been estimated that 99% of microorganisms have not been discovered due to these limitations (Amann et al., 1995). In one gram of soil, it has been estimated that 3 000 to 10 000 unique genomes are present and this number is likely to be an underestimation because rare species were not taken into account (Ovreas & Torsvik, 1998; Torsvik et al., 1990). Metagenomics allows for the study of the genomes of these not-yet-cultured microorganisms.

Microorganisms produce a variety of valuable primary and secondary metabolites with a wide range of activity which can be exploited to improve biotechnological processes and production. Mining the metagenomes of such microorganisms allows for the discovery of metabolites such as therapeutic compounds (MacNeil et al., 2001; Courtois et al., 2003; Seow et al., 1997), enzymes (Ferrer et al., 2005; Entcheva et al., 2001; Elend et al., 2006; Knietsch et al., 2003; Rondon et al., 2000; Voget et al., 2003) but more importantly to this thesis study, cellulases (Healy et al., 1995; Voget et al., 2006; Pang et al., 2009) and xylanases (Brennan et al., 2004; Lee et al., 2006; Hu et al., 2008). It is likely that novel catalytic activities of cellulases and xylanases can be identified from uncultured microorganisms through metagenomics that suit different

industrial applications and may provide an efficient solution to the economical barrier for cellulose utilization.

Metagenomics has already aided in the discovery of lignocellulolytic enzymes from various sources. For example, an enzyme with laccase-like activity was discovered, through functional-screening, in a metagenomic library of DNA obtained from bovine rumen microflora (Beloqui et al., 2006). The novel laccase was characterized and exhibited higher efficiency than previously studied laccases, displaying its potential in biotechnological applications. Functional screening allowed for the discovery of a novel xylanase gene which was then cloned, expressed and characterized biochemically from a metagenomic library created from soil (Hu et al., 2008). The product, XynH, exhibited different properties than other described xylanases, making it a better candidate for industrial application such as in bioethanol production (Hu et al., 2008). Moreover, cellulases and xylanases were shown to be present through metagenomic sequence analysis of the hindgut of termites by comparing catalytic domains homologous to glycoside-hydrolases (Warnecke et al., 2007). A cold active xylanase was cloned and characterized from manure wastewater metagenomic library (Lee et al., 2006) which has beneficial uses for lower temperature applications compared to enzymes that have maximum activity at higher temperatures. Xylanases with different substrate specificities which were phylogenetically distant compared to previously described xylanases, were discovered from a metagenomic library created from insect intestinal tracts of termites and moths (Brennan et al., 2004). Novel enzymes that evolved independently could be useful in combination with already described xylanases to optimize the degradation of xylan. Potential novel cellulases and xylanases with greater catalytic activity in different

reaction conditions than previously described enzymes can be discovered through the power of function-driven metagenomics. The cost-efficient enzymes can be used for feedstock processing in the production of bioethanol.

VIII.1. Mining the metagenome

Two approaches are used to analyze a metagenomic library: a sequence-driven approach where the library is screened for specific sequences or motifs of interest; or a function-driven approach, where expressed traits of interest are detected by screening. Using a function-driven approach, clones expressing a fully functional gene product can be identified for a number of specific functions (Seow et al., 1997; Courtois et al., 2003). Another advantage of function-driven approach is that it does not require prior sequence knowledge and novel genes not previously described are detected (Brennan et al., 2004). Drawbacks of the function-driven method include the dependence on expression of the genes in a foreign host and proper folding to yield the production of a functional gene product (Gabor et al., 2004). A function-driven approach allowed for the discovery of a putative novel cellulase in this thesis work.

Enrichment strategies have been used by others to successfully construct metagenomic libraries. Libraries constructed from cellulose-enriched samples and screened for cellulase activity had a greater number of positive clones, compared to studies without an enrichment step (Feng et al., 2007; Grant et al., 2004; Kim et al., 2008; Pang et al., 2009; Rees et al., 2003) (Table VIII.1.1.). Furthermore, a major limitation of metagenomics is the recovery of high quality or high molecular weight DNA (Zhou et al., 1996). Enrichment cultures can be used in the lab to selectively enhance the isolation of genomic DNA with desired activities within environmental soil samples (Borneman, 1999). Therefore, enrichment of soil samples with cellulose before construction of the

metagenomic library was used to decrease the number of clones that would need to be screened before a cellulase-positive clone was detected, in this thesis work (Figure VIII.2.1.).

Table VIII.1.1. Number of positive clones compared to number of clones screened for metagenomic libraries constructed with an enrichment step compared to metagenomic libraries constructed without an enrichment step

No enrichment step vs enrichment step	Type of metagenomic library	# positive clones/# clones screened	Gene of interest	Reference
No enrichment	Soil	1/70 000	cellulase	Kim et al., 2008
	Compost soil	1/25 000	cellulase	Pang et al., 2009
	Lake Nakuru water	1/60 000	cellulase	Rees et al., 2003
	Soil	1/ 105 000	oxidoreductase	Knietsch et al., 2003b
With an enrichment step	Soil with glycerol and 1,2-propanediol for polyol-consuming microorganisms	1/60 000	oxidoreductase	Knietsch et al., 2003a
	Lake Nakuru water and CMC	1/15 000	cellulase	Rees et al., 2003
	Soil enriched for agarolytic activity	1/213	cellulase	Voget et al., 2006
	Thermophilic anaerobic digesters fed with dried Napiergrass and dried Bermudagrass	1/1250	cellulase	Healy et al., 1995
	Rabbit ceca fed with only grass	1/8125	cellulase	Feng et al., 2007
	Wadi el Natrun lake sediments enriched with cellulose	1/8750	cellulase	Grant et al., 2004
	Wadi el Natrun soda soil enriched with cellulose	1/3083	cellulase	Grant et al., 2004

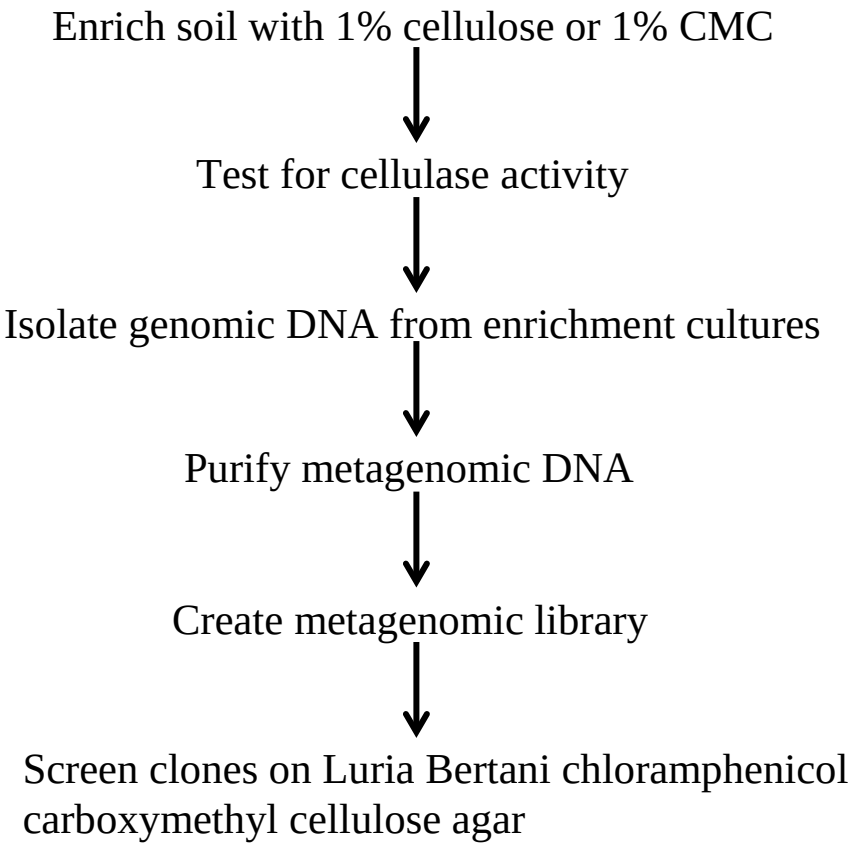


Figure VIII.2.1. Overview of mining the soil metagenome for cellulases

IX. MATERIALS AND METHODS

IX.1. Materials

All materials, chemicals and antibiotics were purchased from Bioshop, Burlington, ON or Fisher Scientific, Fair Lawn, NJ unless otherwise stated. Agarase was purchased from Fermentas, Burlington, ON. All Purpose Potting Soil was purchased from Canadian Tire, in 2007. The CopyControl™ Fosmid Library Production Kit was purchased from EPICENTRE Biotechnologies, Madison, WI.

IX.2. Bacterial strain

EPI300™-T1^R Phage T1-resistant *Escherichia coli* plating strain (F- mcrA Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80d*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 D(*ara*, *leu*)7697 *gal*U *gal*K λ- *rps*L *nup*G *ton*A) (Wild et al., 2002) was used for this study.

IX.2.1. Maintenance of *Escherichia coli*

EPI300™-T1^R Phage T1-resistant *E. coli* plating strain was provided as frozen glycerol stocks and maintained as such. *E. coli* containing fosmids were maintained in Luria Bertani (LB)-Hogness (1X) (For 10X Hogness: 40% (w/v) glycerol, 0.036 M K_2HPO_4, 0.013 M KH_2PO_4, 0.02 M tri-sodium citrate, 0.01 M $MgSO_4$ in deionized water)-Chloramphenicol (12.5 µg/mL) at -20°C and the archive of clones was stored at -80°C.

IX.3. Soil enrichments

Enrichment cultures containing soil and 1% (w/v) cellulose or 1% (w/v) caboxymethyl cellulose (CMC) were used to select for microorganisms that produced cellulases. Enrichment cultures of UOIT forest rhizosphere soil, UOIT forest bulk soil and potting soil contained one gram of soil and either insoluble 1% (w/v) cellulose or 1% (w/v) CMC, a soluble form of cellulose in mineral salts medium (37.33 mM Na_2HPO_4,

14.55 mM KH_2PO_4, 0.8 mM $MgSO_4$, 3.42 mM NaCl, 0.45 mM $CaCl_2$). The growth conditions of these enrichments were 30°C at 150 rpm on a C25 Incubated Floor Model Shaker (New Brunswick Scientific). The enrichment cultures were grown for 10 months at room temperature, either stationary or aerated at 150 rpm at 30°C.

IX.4. Cellulase and xylanase activity screens of soil enrichments

Activity screens of supernatants from soil enrichments were performed to determine if there was any cellulase or xylanase activity present. Activity screens were performed on 0.05% (w/v) cellulose (0.05% (w/v) of cellulose, 0.57 mM K_2HPO_4, 1.35 mM KCl, 0.20 mM $MgSO_4$, 0.005% (w/v) yeast extract, 1.5 % (w/v) agar) or 0.05% (w/v) CMC agar (0.05% (w/v) of carboxymethyl cellulose, 0.57 mM K_2HPO_4, 1.35 mM KCl, 0.20 mM $MgSO_4$, 0.005% (w/v) yeast extract, 1.5% (w/v) agar) and 5 mL of 1% (w/v) congo red (CR) dye and 1% birchwood xylan agar plates (5.7 mM K_2HPO_4, 17.1 mM NaCl, 15.1 mM $(NH_4)_2SO_4$, 1% (w/v) xylan from birchwood, 20.0 mM $CaCO_3$, 1.8% (w/v) agar) (Yang et al., 1995). The concentration of 0.05% cellulose allowed sufficient degradation to be seen in a shorter period of time than 0.5% of cellulose. The concentration of xylan used was based on published values (Yang et al., 1995). A leather punch was used to create holes in the agar and the resulting agar plug was aseptically removed. The supernatant from the enrichment was aseptically removed and 20 µL of supernatant was dispensed in the holes. 20 µL of water was added to one agar hole as a control. Plates were incubated at 30°C and photographed using a FluorChem SP (Alpha Innotech) gel documentation system after 72 hours. For the xylan degradation assay, the plates were flooded with iodine after 27-40 hours. Activity was determined by observing clearance zones around the holes in the agar.

IX.5. Separation of culture supernatant from cells

Cells were harvested from enrichments for later metagenomic DNA extraction as follows. Enrichment cultures were quantitatively transferred into 50 mL screw-capped tubes. The tubes were centrifuged at 1801 *g* for 15 minutes at 4°C to pellet the cells. Supernatant was decanted into 50 mL screw-capped tubes. The cell pellets were stored at 4°C until cells were used for metagenomic DNA extraction.

IX.6. Metagenomic DNA Extraction

Metagenomic DNA extraction was performed as described in Zhou et al. (1996). Metagenomic DNA was extracted from cells harvested from enrichments. 5 g of enriched soil samples were mixed with 13.5 mL of DNA extraction buffer (100 mM Tris-HCl [pH 8.0], 100 mM sodium EDTA [pH 8.0], 100 mM sodium phosphate [pH 8.0], 1.5 M NaCl, 1% (w/v) CTAB) and 100 μL of proteinase K (10 mg/mL) and horizontally shaken at 225 rpm for 30 minutes at 37°C. Then 1.5 mL of 20% SDS was added to the sample and incubated at 65°C in a water bath for 2 hours with gentle end-over-end inversions every 15 to 20 minutes. After incubation, the samples were centrifuged at 6000 *g* for 10 minutes at room temperature. The supernatant was transferred to 50 mL Oak Ridge tubes. Samples were extracted twice more with 4.5 mL of the extraction buffer and 0.5 mL of 20% SDS, vortexed for 10 seconds, incubated at 65°C for 10 minutes and centrifuged as before. All the supernatants were combined and an equal volume of chloroform:isoamyl alcohol (24:1, v/v) was added. The sample was centrifuged at 6000 *g* for 10 minutes and the aqueous phase was transferred to a new Oak Ridge tube. DNA was precipitated by adding 0.6 volume of room temperature 100% isopropanol and incubated at room temperature for 1 hour. Next, centrifugation at 16 000 *g* for 20 minutes at room temperature was done to pellet the DNA. The DNA was washed with cold 70% ethanol,

air-dried and resuspended in sterile distilled water for a final volume of 2 mL. DNA was stored at 4°C for a month. DNA was electrophoresed on a 0.85% (w/v) Tris-acetate-EDTA (TAE) agarose gel to confirm quality and approximate quantity compared to mlolecular weight marker, λ*Hind*III (Bioshop).

IX.7. Purification of Metagenomic DNA from humic acids

The extracted metagenomic DNA was brown in colour indicating the presence of humic acids. Humic acids interfere with cloning (Miller et al., 1999; Tebbe & Vahjen, 1993). Therefore, removal of humic acids from the metagenomic DNA before cloning was required. Several methods were used to purify metagenomic DNA from humic acids: trough method (Harnpicharnchai et al., 2007); agarase (Fermentas) digestion of low melting point agarose; and Cetyltrimethylammonium Bromide (CTAB) (Zhou et al., 1996).

IX.7.1. Trough method

A trough method (Harnpicharnchai et al., 2007) was used to separate genomic DNA from humic acids which co-purified with metagenomic DNA. A 0.85% (w/v) TAE agarose gel was prepared and the gel was loaded to separate genomic DNA from humic acids. Humic acids electrophorese through the gel faster than high molecular weight genomic DNA (Harnpicharnchai et al., 2007). The gel was electrophoresed for 30 minutes at 100 V and stained with ethidium bromide (0.5 µg/mL) for 10 minutes. Since UV light nicks (Witkin, 1976) the DNA, a Darkreader Transilluminator (Clare Chemical Research) was used to visualize DNA on the gel. A small trough was cut below the metagenomic DNA band and trough buffer (15% (w/v) Polyethylene glycol (PEG) 8000, 50% (v/v) 2X TAE was added to the trough. TAE buffer (40 mM Tris acetate and 1 mM Ethylenediaminetetra-acetic acid (EDTA)) was removed so the volume of the TAE buffer

only came up to half of the gel on both sides. The gel was electrophoresed until the metagenomic DNA ran into the trough. The trough buffer containing the DNA was then transferred to a screw-capped tube. The gel was viewed again on the Darkreader Transilluminator (Clare Chemical Research) to confirm all the metagenomic DNA had been collected. An equal volume of chloroform:isoamyl alcohol (24:1, v/v) was added to the screw-capped the containing the DNA and trough buffer. This was centrifuged at 17 000 *g* for 10 minutes. The aqueous phase was transferred to a new microfuge tube and 1/10 volume of 3 M sodium acetate and 0.6 volume of 100% room temperature isopropanol was added. This was allowed to stand at room temperature for 1 hour to precipitate the DNA. The DNA was pelleted by centrifugation at 3716 *g* for 20 minutes at room temperature and washed with cold 70% ethanol. DNA was resuspended with TE buffer pH 8.0 (10 mM Tris-Cl, pH 7.5, 1 mM EDTA). Genomic DNA was stored at 4°C.

IX.7.2. Purification of DNA from Low Melting Point agarose

Since the trough method gave a low yield of DNA, purification of DNA from humic acids was performed by extracting the DNA from low melting point agarose. A thin 1% (w/v) low melting point agarose gel was prepared. The gel was loaded and electrophoresed at 80V-90V for 30-45 minutes. The agarose gel was stained with 0.5 µg/mL ethidium bromide and viewed on the Darkreader Transilluminator (Clare Chemical Research) to prevent nicking of DNA. Metagenomic DNA was cut from the gel and 200 mg pieces were transferred into microfuge tubes. The tubes were incubated for 10 minutes at 70°C to ensure all of the gel melted. The tubes were then transferred to 42°C to equilibrate for 5 minutes before adding 1 U of agarase per 100 mg of agarose. This was gently mixed and incubated for 30-60 minutes at 42°C. The enzyme was then inactivated at 70°C for 10 minutes then the tubes were chilled on ice for 5 minutes. The

undigested oligosaccharides were pelleted by centrifugation at room temperature at 17 000 *g* for 10 minutes. The supernatant was transferred using cut tips to prevent shearing of DNA. DNA was precipitated by the addition of 1/10 volume of 3 M sodium acetate and 1 volume of room temperature isopropanol. This was mixed gently and incubated at -20°C overnight. The tubes were centrifuged at 17 000 *g* for 20 minutes at room temperature and the supernatants were decanted. The pellet was washed twice with cold 70% ethanol and air dried for 10-15 minutes. DNA was resuspended in TE buffer pH 8.0. Purified metagenomic DNA was electrophoresed on a 0.85% TAE agarose gel to check for quality, stained with 0.5 µg/mL ethidium bromide for 10 minutes, visualized on a UV transilluminator and photographed using FluorChem SP (Alpha Innotech) gel documentation system.

IX.7.3. Removal of humic acids with Cetyltrimethylammonium Bromide

Removal of humic acids using Cetyltrimethylammonium Bromide (CTAB) was used (Zhou et al., 1996). An equal volume of phenol:chloroform:isoamyl alcohol pH 8.0 (25:24:1, v/v) was added to potting soil metagenomic DNA and mixed by inversion. This was centrifuged at 17 000 *g* for 10 minutes at room temperature and the aqueous phase was transferred to a fresh microfuge tube. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1, v/v) pH 8 was added, mixed and centrifuged as previously described. The aqueous phase was transferred to a fresh microfuge tube and CTAB was added to a final concentration of 2.5% and mixed by inversion. This was centrifuged 17 000 *g* for 10 minutes at room temperature. Three layers were obtained, the pellet, a loose brown layer and the supernatant. The supernatant from previous centrifugation was transferred and an equal volume of chloroform:isoamyl alcohol (24:1, v/v) was added. This was centrifuged at 17 000 *g* for 5 minutes at room

temperature and 1/10 volume of 3M sodium acetate and an equal volume of 100% isopropanol at room temperature was added. To the loose brown layer, TE buffer pH 8.0 was added and transferred to a new tube. Then, this loose brown layer was processed the same way as the supernatant. The DNA was precipitated by incubation on ice for a few minutes then transferred to -80°C for 10-15 minutes. The DNA was pelleted by centrifugation at 17 000 *g* for 20 minutes and washed with cold 70% ethanol and resuspended in TE buffer.

Various CTAB concentrations were used to determine the optimal concentration necessary to purify DNA away from the humic acids. 500 µL aliquots of metagenomic DNA extracted from potting soil was transferred to a new microfuge tube. CTAB at final concentrations of 0%, 0.5%, 1%, 1.5%, 2%, 2.5% were added. This was then mixed by inversion 10 times and centrifuged at 17 000 *g* for 10 minutes at room temperature. The supernatant was transferred to a new tube and an equal volume of chloroform:isoamyl alcohol (24:1, v/v) was added and mixed. This was centrifuged for 5 minutes at room temperature and 1/10 volume of 3M sodium acetate and an equal volume of 100% isopropanol at room temperature was added. This was incubated overnight at -20°C. The DNA was pelleted by centrifugation at 17 000 *g* for 20 minutes and washed with 70% ethanol and resuspended in TE buffer.

IX.8. Construction of metagenomic library using CopyControl™ HTP Fosmid Library Production Kit

Purified DNA was used to create a fosmid library using the CopyControl™ HTP Fosmid Library Production Kit (EPICENTRE) as per the manufacturer's protocol. Purified metagenomic DNA was triturated 45 times to shear the DNA. The sheared DNA was end repaired to generate blunt-ended, 5'-phosphorylated DNA. The end repaired

DNA was size selected by electrophoresis on a 1% low melting point TAE agarose gel overnight at 30-35V (constant voltage) at room temperature with 100 ng of Fosmid Control DNA loaded as the size marker. The gel was stained with 0.5 µg/mL ethidium bromide and viewed on a Darkreader Transilluminator (Clare Chemical Research). The DNA that migrated with the 36 kb Fosmid Control DNA marker was excised. A known quantity of the gel slice was placed in microfuge tubes. One milligram of solidified agarose was assumed to equal 1 µL of molten agarose after melting. GELase 50X Buffer was warmed to 45°C and low melting point (LMP) agarose was melted by incubating the tube at 70°C for 10-15 minutes then transferring the tube to 45°C. Warmed GELase 50X Buffer was added to 1X final concentration and for 100 µL of melted agarose, 1 U of GELase Enzyme Preparation was added. The solution was gently mixed and incubated for 1 hour to overnight at 45°C. The enzyme was inactivated at 70°C for 10 minutes. The tubes were chilled for 5 minutes and centrifuged at 17 000 *g* for 20 minutes to pellet insoluble oligosaccharides. The supernatant was transferred to a new microfuge tube and the DNA was precipitated by the addition of 1/10 volume of 3 M sodium acetate (pH 7.0) and 0.6 volume of 100% isopropanol at room temperature and incubated at -20°C for one hour. The precipitated DNA was pelleted by centrifugation for 20 minutes at 17 000 *g*. The supernatant was removed and discarded. The DNA pellet was washed twice with ice cold 70% ethanol. The DNA pellet was air dried for 5-10 minutes with the tube inverted. The DNA pellet was resupended in TE buffer pH 8.0. The DNA concentration was quantified by electrophoresis of an aliquot of DNA with known concentrations (50 ng/µL, 25 ng/µL, 12.5 ng/µL, 10 ng/µL, 5 ng/µL) of diluted Fosmid Control DNA on a 0.85% (w/v) agarose gel for 20 minutes at 100V.

The following components were combined for the ligation of 40 kb metagenomic DNA with the vector in 10 µL total reaction volume: 1X Fast-Link Ligation Buffer, 1 mM ATP, 0.5 µg/µL CopyControl pCC2FOS Vector, 0.25 µg of 40 kb concentrated insert DNA, 2U Fast-Link DNA Ligase. The ligation reaction was incubated at room temperature for 2 hours and the reaction was transferred to 70°C for 10 minutes to inactivate the Fast-Link DNA Ligase.

The titer of packaged phage particles was determined with the equation:

(# of colonies) (dilution factor) (1000 µL/mL)/(volume of phage plated [µL]

The appropriate dilution (10^0) that gave 20-200 colonies was used to infect more cells. Infected cells were plated onto LB amended with 12.5 µg/mL chloramphenicol and grown overnight at 37°C. Clones were inoculated into 96 well microtiter plates containing 200 µL LB supplemented with 1X Hogness (For 10X Hogness: 40% (v/v) glycerol, 0.036 M K_2HPO_4, 0.013 M KH_2PO_4, 0.02 M tri-sodium citrate, 0.01 M $MgSO_4$ in deionized water) and 12.5 µg/mL of chloramphenicol and were grown overnight at 37°C at 250 rpm. A replicate set of plates was inoculated to serve as an archive. The working set of plates was stored at -20°C while the archive set of plates was stored at -80°C.

IX.9. Screening of the soil metagenomic library for cellulase activity

The soil metagenomic library was screened for cellulase activity on LB supplemented with 0.1% CMC and 12.5 µg/mL of chloramphenicol at 26-37°C for 2-4 days. Plates were then stained with 0.1% congo red for 15-20 minutes and destained with 1 M NaCl for 15 minutes. Photos were taken of the plates using a digital camera. The plates were then flooded with 1 N HCl for 5 minutes to increase contrast of any clearance

zones. Congo Red is a pH dependent dye and is blue under acidic conditions (Mera & Davies, 1984).

IX.10. Fosmid purification

To confirm that the library consisted of appropriate-sized inserts, 16 clones were randomly chosen. Clones were inoculated into 5 mL of LB Hogness (1X) 12.5 µg/mL chloramphenicol and incubated overnight at 37°C at 250 rpm. A glycerol stock of each clone was created with 800 µL of culture and 200 µL of LB Hogness (1X) 12.5 µg/mL chloramphenicol. Fosmids were purifed by the alkaline lysis method (Sambrook & Russell, 2001). The fosmid DNA was resuspended in 50 µL of TE pH 8.0 containing 20 µg/mL DNAse-free pancreatic RNAse. Fosmid DNA was stored at -20°C.

X. RESULTS

X.1. Enrichment cultures using various soil types

A functional soil metagenomic library was constructed from cellulose-enrichment cultures to screen for cellulases. Enrichment cultures containing potting soil, UOIT forest rhizosphere soil and UOIT forest bulk soil in 1% cellulose or 1% CMC liquid media were created. These enrichment cultures were inoculated for the selection of cellulase producing microorganisms. To increase the number of cells for extraction of DNA, subculturing of the enrichment cultures was performed.

X.2. Cellulase and xylanase activity screens of soil enrichments

Soil enrichments were initiated with the purpose of constructing a metagenomic library to discover cellulase genes and xylanase genes. Therefore, screening of cellulase and xylanase activity of enrichment cultures prior to library construction was performed. The enrichment cultures were screened for cellulase and xylanase activity present in the culture supernatant every few months for 10 months. Initially, activity screens of enrichment cultures supernatants were performed on 0.05% (w/v) CMC (Figure X.2.1.B.) and 1% (w/v) congo red (CR) dye. Then, assays with 0.05% (w/v) cellulose (Figure X.2.1.A.) and 1% (w/v) congo red (CR) dye, and 1% birchwood xylan agar plates (Figure X.2.1.C.) stained with iodine were used to screen for cellulase and xylanases as well. Cellulase and xylanase activities were observed in the enrichment cultures.

Overall, potting soil enrichment supernatants showed greater xylanase activity than rhizosphere soil enrichment supernatants. Potting soil enrichment supernatants inoculated in cellulose containing medium had larger xylan clearance zones than when inoculated in CMC containing medium. On cellulose and CMC agar, CMC enriched rhizosphere forest soil supernatants had better cellulase and xylanase activity than potting

soil enrichment supernatants assayed on CMC. Supernatants of potting soil enrichment cultures containing cellulose had better cellulase and xylanase activity than supernatants of rhizosphere forest soil enriched with cellulose. Therefore, potting soil inoculated with cellulose and rhizosphere soil inoculated with CMC had better cellulase and xylanase activities than other enrichments. Furthermore, bulk soil supernatants had better cellulase activity on cellulose than on CMC. Supernatants from enrichment of bulk soil containing cellulose medium was better than rhizosphere soil cellulose enrichment supernatants. Better xylanase activity was observed in bulk soil supernatants than rhizosphere soil supernatants. Enrichment cultures without aeration had similar cellulase and xylanase activities as the aerated enrichment cultures. Therefore, enrichment cultures containing potting soil and cellulose or enrichment cultures of rhizosphere soil amended with CMC were optimal enrichments for cellulase activity selection.

The results from the activity screens showed that potting soil enriched with 1% cellulose had greater cellulase activity compared to the other enrichments. Also, when screening actinomycete isolates purified from potting soil for cellulase activity, potting soil isolates were better at degrading cellulose compared to bulk or rhizosphere soil isolates. Therefore, cells from cellulose enrichments of potting soil were used for the construction of the metagenomic library.

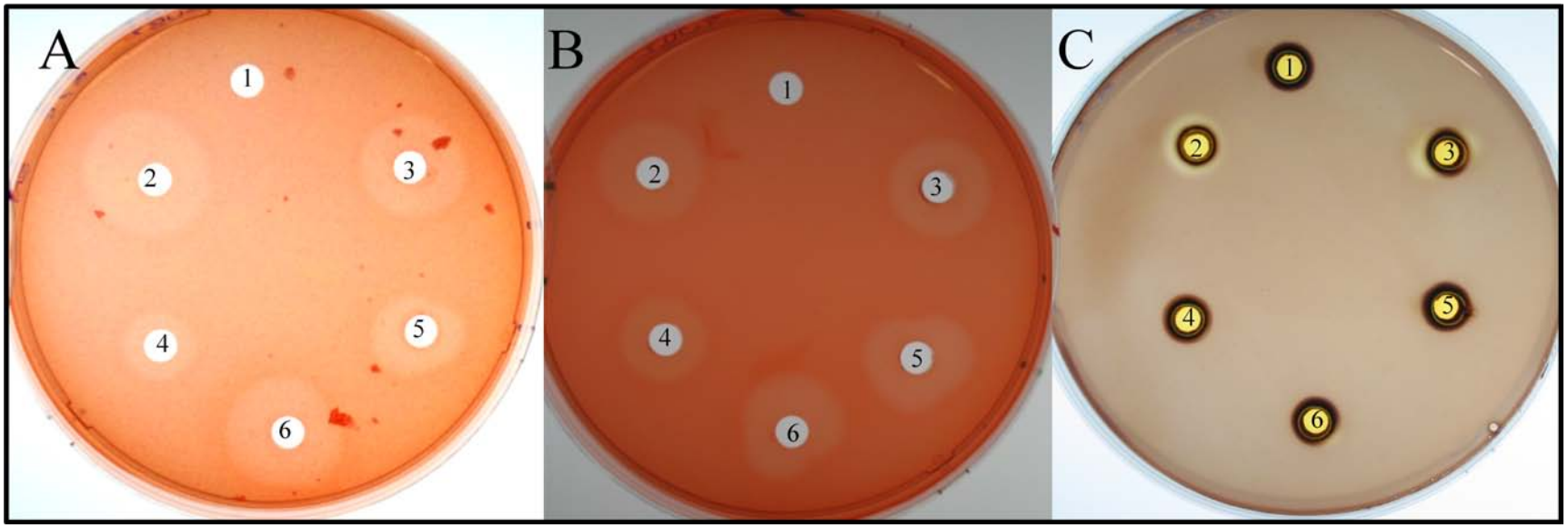


Figure X.2.1. Cellulase activity exhibited by enrichment subcultures on A) C-CRA, B) CMC-CRA and C) xylan agar stained with iodine. 1. dH_2O (control); 2-5: Supernatants from cultures incubated on shaker at 150 rpm at 30°C for 9 months; 2. 1% cellulose + potting soil; 3. 1% CMC + potting soil; 4. 1% cellulose + forest rhizosphere soil; 5. 1% CMC + forest rhizosphere soil; 6. 1% CMC + UOIT forest rhizosphere soil was incubated at room temperature and stationary on the bench. Clearance zones were measured after 72 hours of incubating plates at 30°C, the size of the well was subtracted from the zone then the radii was calculated.

X.3. Screening of the metagenomic library

The metagenomic library, designated PSC (Potting Soil Cellulose), was constructed from the enrichment cultures of potting soil and 1% cellulose. A total of 1,920 clones were screened on CMC containing medium. Only one clone exhibited cellulase activity. This clone, PSC21-G6, produced a putative extracellular endoglucanase that could diffuse through the agar to create a faint but large clearance zone on CMC agar (Figure X.3.1.).

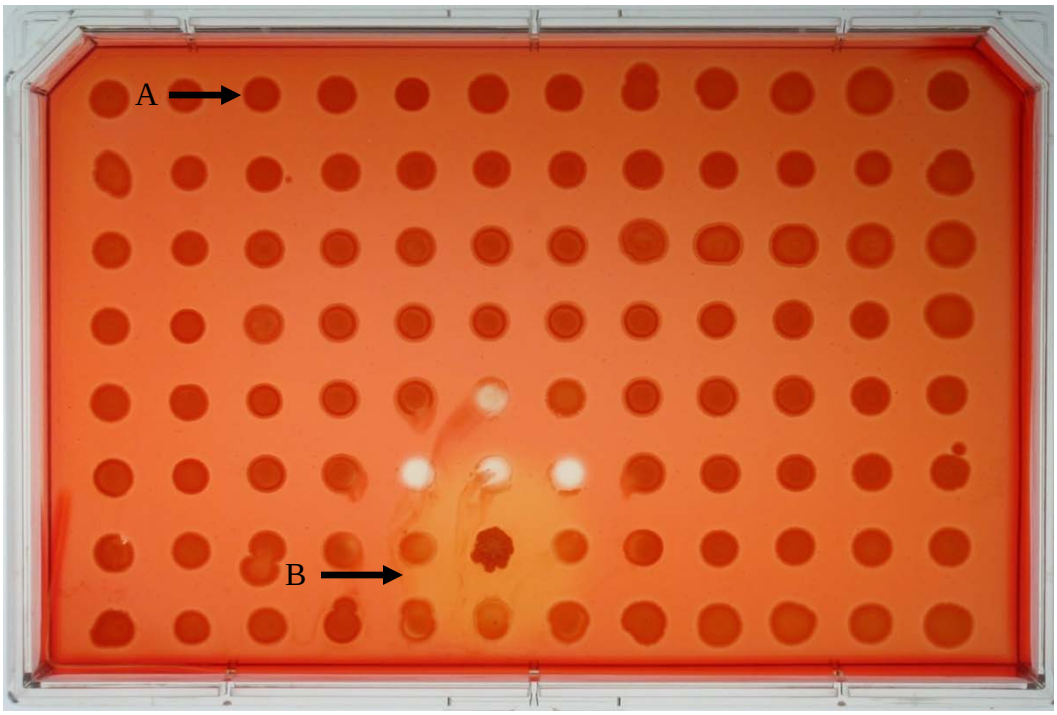


Figure X.3.1. PSC metagenomic library clones plated on LB agar supplemented with 0.1% CMC and 12.5 µg/mL of chloramphenicol stained with congo red and destained with 1 M NaCl to examine cellulase production. A) clone B) a large clearance zone produced by clone PSC21-G6.

X.4. Purification of fosmids for analysis

Sixteen clones were randomly selected for confirmation of insert. It was determined all 16 fosmid isolates contained inserts that were at least 40 kb (Figure X.4.1.).

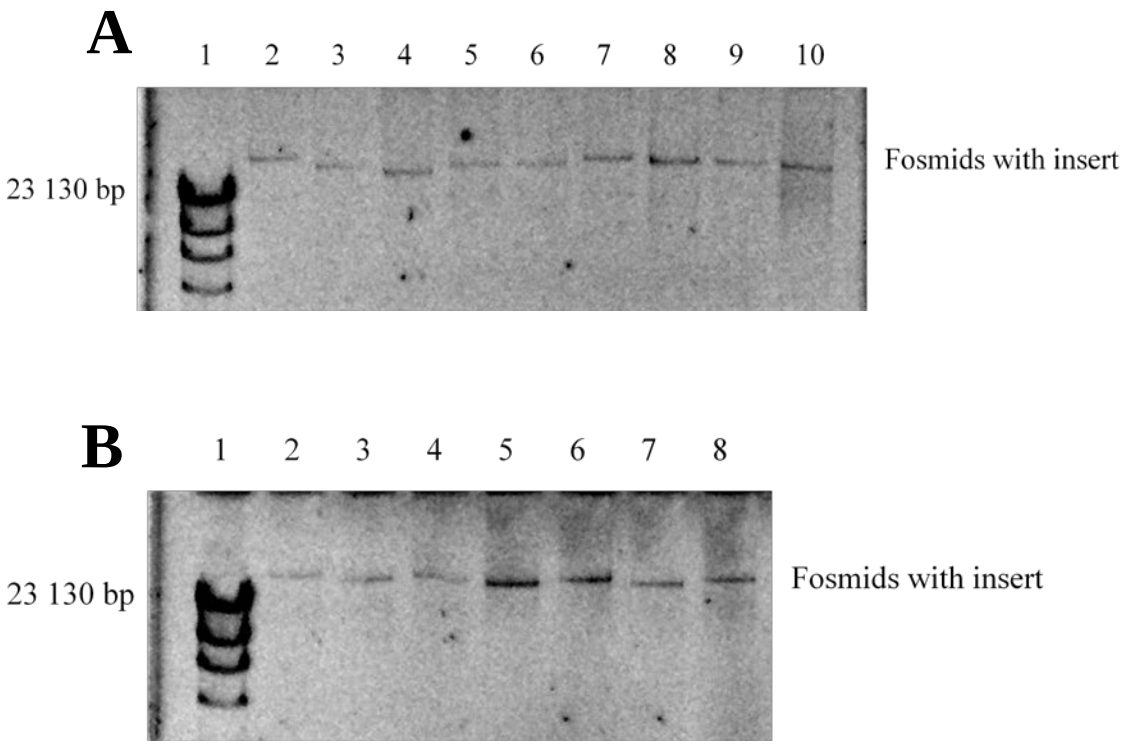


Figure X.4.1. Purified, undigested fosmids from randomly selected clones electrophoresed on a 0.85% agarose gel. The fosmid vector, pCC2FOS, is 8181 bp. Lane 1A and 2B) λ*Hind*III; A) 2-10 purified fosmids and B) 2-8 purified fosmids.

XI. DISCUSSION

Enrichment cultures were created from potting soil, rhizosphere soil and bulk soil types in 1% carboxymethyl cellulose (CMC) and 1% cellulose liquid media and grown for 10 months. Initially, enrichment cultures were screened for cellulase activity on 0.05% CMC agar. Later, 0.05% cellulose and 1% birchwood xylan agar were incorporated to screen for cellulase and xylanase activity, respectively. Cellulose and xylan were used to screen for cellulase and xylanase activity because these media types were used in the screening of actinomycete isolates cultured from soil.

The potting soil cellulose enrichment culture was observed to have cellulase activity throughout the 10 months of enrichment. In order for the organisms to grow on the primary carbon source in the cellulose enrichment cultures, production of extracellular cellulases would be necessary for organisms to degrade cellulose into glucose. Therefore, a functional metagenomic library was constructed from DNA extracted from the enrichment culture of potting soil inoculated into 1% cellulose medium to increase the likelihood of obtaining cellulase positive clones.

Xylanase activity was also observed in the cellulose enrichment cultures. This was probably because cellulase and xylanase activity have been shown to be linked in several microorganisms. For example, in the soil bacterium *Thermobifida fusca*, cellobiose not only induces cellulase expression but it induces xylanase production (Chen & Wilson, 2007). Moreover, cellulase has been shown to be able to degrade xylan to a certain extent because cellulose is associated with xylan in nature (Gilkes et al., 1988; Hall et al., 1988). In addition, cellulase and xylanase genes have been shown to be clustered in the fungal genus, *Piromyces* (Ali et al., 1995). Furthermore, two endoglucanases and three xylanases (endoxylanases B and C and β-xylosidase) are

regulated by a single transcriptional activator, XlnR in *Aspergillus niger* (van Peij et al., 1998) indicating these enzymatic activities can be co-regulated.

XI.1. Screening of the metagenomic library for cellulase activity

The metagenomic library of 1,920 clones constructed from organisms grown in enrichments of potting soil with 1% cellulose for 10 months resulted in one putative cellulase positive clone when screened on CMC agar stained with congo red dye. Congo red is a dye that binds to β-1-4-glucosidic linkages of cellulose (Teather & Wood, 1982). Endoglucanases will cleave the intramolecular β-1-4-glucosidic linkages randomly and the dye will be released. CMC is used for the screening of endoglucanases (Ghose, 1987). Therefore, this putative cellulase activity is likely to be an endo-β-1,4-glucanase because CMC was used to screen for cellulase activity and degradation of CMC was observed. Smaller clearance zones were observed around some of the clones in Figure X.3.1. This could indicate some cell associated cellulase activity because the cellulase did not diffuse out to degrade the cellulose. The cell-associated activity was not further investigated in this study.

XI.2. Comparison of metagenomic libraries created with and without an enrichment step

Metagenomic libraries from a variety of sources have been constructed by enrichment and non enrichment methods. It has been observed that libraries which were enriched and screened for a gene activity of interest had a greater number of positive clones, compared to studies which did not employ enrichment methods (Grant et al., 2004; Feng et al., 2007; Kim et al., 2008; Pang et al., 2009; Rees et al., 2003) (Table XI.2.1.). Since cellulose enrichments were used to create the metagenomic library

described in this thesis, microorganisms that produce cellulase to degrade natural cellulose would have been selected for.

Table XI.2.1. Number of positive clones compared to number of clones screened for metagenomic libraries constructed with an enrichment step compared to metagenomic libraries constructed without an enrichment step

No enrichment step vs enrichment step	Type of metagenomic library	# positive clones/# clones screened	Gene of interest	Reference
No enrichment	Soil	1/70 000	cellulase	Kim et al., 2008
	Compost soil	1/25 000	cellulase	Pang et al., 2009
	Lake Nakuru water	1/60 000	cellulase	Rees et al., 2003
	Soil	1/ 105 000	oxidoreductase	Knietsch et al., 2003b
With an enrichment step	Soil with glycerol and 1,2-propanediol for polyol-consuming microorganisms	1/60 000	oxidoreductase	Knietsch et al., 2003a
	Lake Nakuru water and CMC	1/15 000	cellulase	Rees et al., 2003
	Soil enriched for agarolytic activity	1/213	cellulase	Voget et al., 2006
	Thermophilic anaerobic digesters fed with dried Napiergrass and dried Bermudagrass	1/1250	cellulase	Healy et al., 1995
	Rabbit ceca fed with only grass	1/8125	cellulase	Feng et al., 2007
	Wadi el Natrun lake sediments enriched with cellulose	1/8750	cellulase	Grant et al., 2004
	Wadi el Natrun soda soil enriched with cellulose	1/3083	cellulase	Grant et al., 2004

XI.3. Assumptions and limitations to metagenomics

There are assumptions and limitations to the metagenomic library that was constructed in this study. Enrichments do limit the diversity of the consortia because only organisms that can use the substrate, cellulose or the degradative products such as glucose, can survive. This however, does not limit the diversity of microorganisms within the enrichment that produce the specific activity of interest. A decrease in the biodiversity due to culture enrichment was not important in this study because biodiversity was not the aim of this study.

In this study, only endo-β-1,4-glucanase activity was screened for using carboxymethyl cellulose (Ghose 1987) resulting in assay bias. Other assays can be used to screen for other cellulase activities. For example, β-glucosidase activity can be measured through a cellobiose assay (Ghose 1987). Exo-β-1,4-glucanase activity can be determined by either the Avicel assay (Ng et al., 1977) or filter paper assay (Mandels & Reese, 1965).

The metagenome is composed of the DNA from organisms present in a given environment. The average genome size in an environment is difficult to determine because there are many uncultured microorganisms that have not been studied. The number of fosmid clones in a metagenomic library required to give an adequate representation of the metagenome can be determined by the equation $N = \ln (1-P) / \ln (1-f)$, where N is number of fosmid clones required; P is the desired probability; and f is the proportion of the genome contained in a single clone (Sambrook & Russell, 2001). If there are approximately 6,000 genomes present (Torsvik et al., 1990; Ovreas & Torsvik, 1998) and the average genome size is assumed as 4 Mb (Raes et al., 2007), then that would equal 2.4×10^{10} bases. The number of clones needed to guarantee about 99% of

the DNA sequences to be contained within a fosmid library made with 40 kb insert size is: $N = \ln (1\text{-}0.99)/\ln (1 - [4 \times 10^4 \text{ bases}/2.4 \times 10^{10} \text{ bases}]) = 2.7 \times 10^6$ clones. To ensure a good representation of organisms used as a source in the construction of the metagenomic library, more than 2.7×10^6 fosmid clones are necessary. In this study, only 1,920 clones were selected and screened. A more accurate representation of the genetic composition in potting soil would require more fosmid clones be screened.

In addition to the assumption of the average genome size, during the construction of the library, only 25-40 kb fragments were used as inserts. The PSC library would therefore only represent about 48 Mb, with 25 kb insert, to maximum of 76.8 Mb, with a 40 kb insert. A greater number of fosmid clones would be required to ensure a greater representation of the metagenome due to the insert size restriction. If 10^7 prokaryotic cells (Gans et al., 2005), with a range of 3,000 to 10,000 different prokaryotic genomes (Torsvik et al., 1990; Ovreas & Torsvik, 1998) and genome size ranging from 1.5 Mb – 8 Mb (Raes et al., 2007) is estimated to be present in one gram of soil, a large amount of genetic information is present in one gram of soil. It has been estimated that 10^6 Bacterial Artificial Chromosomes (BAC) clones, with an insert size of 100 kb, must be screened for coverage of the all the distinct prokaryotic species in one gram of soil (Handelsman et al., 1998).

Escherichia coli is a Gram-negative bacterium and was used as the host for the construction of the metagenomic library. A drawback of using *E. coli* or any foreign host is that expression of the gene and the gene product, cellulases in this case, is limited and dependent on the host because the heterologous host may not contain the cellular components required to express or secrete functional cellulases (Gabor et al., 2004). In

general, problems involved with the selection of the host that must be considered when undertaking metagenomic investigations include: poor transcription, translation, and excretion of product (Gabor et al., 2004). As well, improper protein folding may be problematic because the proper chaperones may not be present in the host to produce functional proteins (Ferrer et al., 2003; Ferrer et al., 2004; Gabor et al., 2004). Furthermore, improper production or incorporation of cofactors necessary for the function of the protein where expression of protein is present but the protein cannot function without required cofactors (Gabor et al., 2004). Lastly, codon usage can be different depending on the organisms (Sharp & Li, 1987). Codon bias is where organisms will preferentially use certain codons to code for an amino acid instead of using other synonymous codons (Sharp et al., 2005). This has the potential to contribute to low protein expression and therefore low observed activity because the protein cannot be expressed by the organism due to codon bias (Grote et al., 2005).

Typically, the host that is used for propagating metagenomic libraries is *E. coli* (Handelsman et al., 1998). The reasons for this are that batch production, separation, and downstream processing methods used in the production of valuable products are already well-studied for *E. coli* (Daniel, 2004). *Streptomyces* and *Pseudomonas* strains have been used as a host to express soil prokaryotic genes (Ono et al., 2007; Martinez et al., 2004; Courtois et al., 2003). *Pseudomonas* and *E. coli* are both Gram-negative bacteria but *Pseudomonas* is naturally found in soil (Cho & Tiedje, 2000) whereas *E. coli* is found in fecal matter (Parveen et al., 1999). *Streptomyces* are Gram-positive bacteria. Advantages of using *Streptomyces* or other actinomycetes as the heterologous host are that they possess a greater number of complex promoters (Strohl, 1992), they can post-

transcriptionally modify products that *E. coli* cannot (Gabor et al., 2004), they can express high G+C DNA content genes (Muto & Osawa, 1987) and actinomycetes are known to produce an array of metabolites so there is a greater chance that the biosynthetic machinery is present to express and produce these products (Wilkinson et al., 2002).

Low copy number vectors are an advantage since certain sequences are subjected to modifications such as deletions and therefore are not clonable in *Escherichia coli*. The fosmid vector used in this study was a single copy vector to avoid the problems associated with over expression of certain gene products (Wang & Kushner, 1991; Renault et al., 1996), which has the potential to destroy the host. Fosmids are low copy number cosmids, plasmids with a *cos* site to allow the packaging of DNA into λ phages (Collins & Hohn, 1978), are more stable than multicopy cosmids (Kim et al., 1992).

The organisms present in the enrichment culture of potting soil inoculated in 1% cellulose liquid medium was used as the source of DNA extraction to construct the functional metagenomic library. A direct lysis approach was used to extract DNA from the organisms without the separation of cells from the soil.

First, not all cells lyse the same way during the DNA extraction process (Kauffmann et al., 2004). A limitation of direct lysis of organisms from soil is that the DNA recovered can contain contaminants, such as humic acids, which interfere with enzymatic reactions including ligases for cloning, restriction endonucleases for digesting DNA, and transformation of the DNA (Tebbe & Vahjen, 1993; Miller et al., 1999). In this study, direct lysis resulted in co-purification of DNA with humic acids. Therefore, the DNA had to be further purified from the humic acids (Figure XI.3.1.). Since humic

acids migrate faster than genomic DNA during electrophoresis (Harnpicharnchai et al., 2007), gel purification can be used. Two gel electrophoresis methods were used to purify DNA from humic acids in this study. In both methods, it was difficult to visualize where the DNA migrated in the gel because DNA concentration was too dilute and a thick gel was used. Therefore to overcome these problems, concentrating the DNA and pouring a thin gel allowed visualization of where DNA migrated in the gel.

To overcome the low yields obtained by gel electrophoresis methods, cetyltrimethylammonium bromide (CTAB) was used to purify humic acids from DNA because CTAB has been reported to reduce humic acid contamination by complexing with humic acids (Zhou et al., 1996) and precipitate anionic nucleic acids from solution (Sibatani, 1970; Jones, 1953). Despite the use of various concentrations of CTAB, humic acids co-purified with the DNA.

Figure XI.3.1. Metagenomic DNA contaminated with humic acids

XI.4. Conclusion

Although the PSC metagenomic library was screened for cellulase activity, it can also be used to screen for other activities of interest such as xylanases because a large amount of DNA is contained in the library. The PSC library was successful in that one cellulase positive clone was discovered in 1,920 clones.

XI.5. Future directions

Several experiments should be completed to evaluate the metagenomic library. First, restriction fragment analysis of cloned insert DNA would ensure that cloning was random and that the metagenome was represented in the library. In addition, this would allow estimation of the average insert size in the library. Also, digestion of the cellulase positive clone, subsequent subcloning and re-screening would identify the gene responsible for the cellulase activity. There could be 7-15 genes in an insert of 40 kb because bacterial genes are about 1-2 kb in size (Xu et al., 2006).

To determine the nucleotide sequence of cellulase genes present, *in vitro* transposon mutagenesis (Voget et al., 2006) can be used and sequencing from the ends of the transposon could be conducted to determine the sequence of the cellulase gene (Voget et al., 2006). The clone that exhibits cellulase activity could be sequenced and compared to known cellulase sequences present in the NCBI BLAST database (Altschul et al., 1990). Using blastx and tblastx, the cellulase sequence from the clone can be compared to any possible sequences in the database so phylogenetic analysis of the translated protein sequence can be completed to compare the relationship of the protein to other identified proteins. The nucleotide sequence would be translated in all six reading frames and compared against blastx, a protein sequence database. Then, finally the six-frame translations of the nucleotide sequence would be compared to six-frame translations of a nucleotide sequence database, tblastx, to ensure the experimental protein is compared to any possible translated nucleotide sequences.

The extracellular cellulase expressed by the positive clone could be purified for proteomic analysis. The biochemical and physical properties of the enzyme such as

optimum temperature, pH stability and enzyme kinetics (Lee et al., 2006; Hurtubise et al., 1995; Feng et al., 2007) can be determined.

With the construction of the metagenomic library, activity screens for other novel enzymes can be completed. The metagenomic library can be used for screening cellulases that degrade insoluble cellulose, and xylanases and peroxidases, both of which take part in lignocellulose degradation (Biely et al., 1986; Ramachandra et al., 1988).

XII. REFERENCES

Ali, B. R., Zhou, L., Graves, F. M., Freedman, R. B., Black, G. W., Gilbert, H. J. & Hazelwood, G. P. (1995). Cellulases and hemicellulases of the anaerobic fungus *Piromyces* constitute a multiprotein cellulose-binding complex and are encoded by multigene families. *FEMS Microbiology Letters* **125**, 15-21.

Aljanabi, S. M. & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research* **25**, 4692–4693.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410.

Amann, R. I., Ludwig, W. & Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews* **59**, 143-169.

Ball, A. S. & McCarthy, A. J. (1988). Saccharification of straw by actinomycete enzymes. *Microbiology* **134**, 2139-2147.

Basil, A. J., Strap, J. L., Knotek-Smith, H. M. & Crawford, D. L. (2004). Studies on the microbial populations of the rhizosphere of big sagebrush (*Artemisia tridentata*). *Journal of Industrial Microbiology & Biotechnology* **31**, 278-288.

Bauer, W. D., Talmadge, K. W., Keegstra, K. & Albersheim, P. (1973). The structure of plant cell walls: II. The hemicellulose of the walls of suspension-cultured sycamore cells. *Plant Physiology* **51**, 174–187.

Belkacemi, K., Turcotte, G. & Savoie, P. (2002). Aqueous/steam-fractionated agricultural residues as substrates for ethanol production. *Industrial & Engineering Chemistry Research* **41**, 173-179.

Beloqui, A., Pita, M., Polaina, J., Martínez-Arias, A., Golyshina, O. V., Zumárraga, M., Yakimov, M. M., García-Arellano, H., Alcalde, M., Fernández, V. M., Elborough, K., Andreu, J. M., Ballesteros, A., Plou, F. J., Timmis, K. N., Ferrer, M. & Golyshin, P. N. (2006). Novel polyphenol oxidase mined from a metagenome expression library of bovine rumen: biochemical properties, structural analysis, and phylogenetic relationships. *The Journal of Biological Chemistry* **281**, 22933-22942.

Bentley, S. D., Chater, K. F., Cerdeño-Tárraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J.

& Hopwood, D. A. (2002). Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147.

Bibb, M., Schottel, J. L. & Cohen, S. N. (1980). A DNA cloning system for interspecies gene transfer in antibiotic-producing *Streptomyces*. *Nature* **284**, 526-531.

Biely, P., MacKenzie, C. R., Puls, J. & Schneider, H. (1986). Cooperativity of esterases and xylanases in the enzymatic degradation of acetyl xylan. *Nature Biotechnology* **4**, 731–733.

Biely, P., Vrsanská, M., Tenkanen, M. & Kluepfel, D. (1997). Endo-beta-1,4-xylanase families: differences in catalytic properties. *Journal of Biotechnology* **57**, 151-166.

Blakeney, A. B., Harris, P. J., Henry, R. J. & Stone, B. A. (1983). A simple and rapid preparation of alditol acetates for monosaccharide analysis. *Carbohydrate Research* **113**, 291-299.

Blum, D. L., Kataeva, I. A., Li, X. L. & Ljungdahl, L. G. (2000). Feruloyl esterase activity of the *Clostridium thermocellum* cellulosome can be attributed to previously unknown domains of XynY and XynZ. *Journal of Bacteriology* **182**, 1346–1351.

Borneman, J. (1999). Culture-independent identification of microorganisms that respond to specified stimuli. *Applied and Environmental Microbiology* **65**, 3398–3400.

Bouxin, F., Baumberger, S., Pollet, B., Haudrechy, A. & Dole, P. (2010). Acidolysis of a lignin model: investigation of heterogeneous catalysis using Montmorillonite clay. *Bioresource Technology* **101**, 736-744.

Brennan, Y., Callen, W. N., Christoffersen, L., Dupree, P., Goubet, F., Healey, S., Hernández, M., Keller, M., Li, K., Palackal, N., Sittenfeld, A., Tamayo, G., Wells, S., Hazlewood, G. P., Mathur, E. J., Short, J. M., Robertson, D. E. & Steer, B. A. (2004). Unusual microbial xylanases from insect guts. *Applied and Environmental Microbiology* **70**, 3609-3617.

Brown, S. (1969). Biochemistry of lignin formation. *BioScience* **19**, 115-121.

Béguin, P. & Aubert, J. P. (1994). The biological degradation of cellulose. *FEMS Microbiology Reviews* **13**, 25-58.

Chanzy, H., Imada, K., Mollard, A., Vuong, R. & Barnoud, F. (1979). Crystallographic aspects of sub-elementary cellulose fibrils occurring in the wall of rose cells cultured in vitro. *Protoplasma* **100**, 303-316.

Chen, C., Chen, J. L. & Lin, T. Y. (1997). Purification and characterization of a xylanase from *Trichoderma longibrachiatum* for xylooligosaccharide production. *Enzyme and Microbial Technology* **21**, 91-96.

Chen, S. & Wilson, D. B. (2007). Proteomic and transcriptomic analysis of extracellular proteins and mRNA levels in *Thermobifida fusca* grown on cellobiose and glucose. *Journal of Bacteriology* **189**, 6260-6265.

Chester, M. & Martin, E. (2009). Cellulosic ethanol from municipal solid waste: a case study of the economic, energy, and greenhouse gas impacts in California. *Environmental Science & Technology* **43**, 5183-9.

Cho, J. C. & Tiedje, J. M. (2000). Biogeography and degree of endemicity of fluorescent *Pseudomonas* strains in soil. *Applied and Environmental Microbiology* **66**, 5448-5456.

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, A. S., McGarrell, D. M., Marsh, T., Garrity, G. M. & Tiedje, J. M. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research* **37**, D141-145.

Collins, J. & Hohn, B. (1978). Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 4242-4246.

Conant, G. & Wagner, A. (2005). The rarity of gene shuffling in conserved genes. *Genome Biology* **6**, R50.

Courtois, S., Cappellano, C. M., August, P. R., Martinez, A., Kolvek, S. J., Hopke, J., Osburne, M. S., Nalin, R., Guérineau, M., Jeannin, P., Simonet, P., Pernodet, J. L., Ball, M., Francou, F. X., Normand, P. & Helynck, G. (2003). Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Applied and Environmental Microbiology* **69**, 49-55.

Coutinho, P. M. & Henrissat, B. (1999). Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering*vol. 246, Gilbert, H., Davies, G., Henrissat, H. & Svensson, B. , 3-14. Cambridge: The Royal Society of Chemistry.

Crawford, D. L., Pometto, A. L. & Crawford, R. L. (1983). Lignin degradation by *Streptomyces viridosporus*: isolation and characterization of a new polymeric lignin degradation intermediate. *Applied and Environmental Microbiology* **45**, 898-904.

Crutzen, P. J. & Andreae, M. O. (1990). Biomass burning in the tropics: impact on atmospheric chemistry and biogeochemical cycles. *Science* **250**, 1669-1678.

Daniel, R. (2004). The soil metagenome--a rich resource for the discovery of novel natural products. *Current Opinion in Biotechnology* **15**, 199-204.

Dari, K., Béchet, M. & Blondeau, R. (1995). Isolation of soil *Streptomyces* strains capable of degrading humic acids and analysis of their peroxidase activity. *FEMS Microbiology Ecology* **16**, 115-122.

Davies, G. & Henrissat, B. (1995). Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853-859.

Delgenes, J. P., Moletta, R. & Navarro, J. M. (1996). Effects of lignocellulose degradation products on ethanol fermentations of glucose and xylose by *Saccharomyces cerevisiae*, *Zymomonas mobilis*, *Pichia stipitis*, and *Candida shehatae*. *Enzyme and Microbial Technology* **19**, 220-225.

Deshpande, V., Lachke, A., Mishra, C., Keskar, S. & Rao, M. (1986). Mode of action and properties of xylanase and beta-xylosidase from *Neurospora crassa*. *Biotechnology and Bioengineering* **28**, 1832-1837.

Dillon, A. J., Camassola, M., Henriques, J. A., Fungaro, M. H., Azevedo, A. C., Velho, T. A. & Laguna, S. E. (2008). Generation of recombinants strains to cellulases production by protoplast fusion between *Penicillium echinulatum* and *Trichoderma harzianum*. *Enzyme and Microbial Technology* **43**, 403-409.

Divne, C., Ståhlberg, J., Teeri, T. T. & Jones, T. A. (1998). High-resolution crystal structures reveal how a cellulose chain is bound in the 50 A long tunnel of cellobiohydrolase I from *Trichoderma reesei*. *Journal of Molecular Biology* **275**, 309-325.

Elend, C., Schmeisser, C., Leggewie, C., Babiak, P., Carballeira, J. D., Steele, H. L., Reymond, J. L., Jaeger, K. E. & Streit, W. R. (2006). Isolation and biochemical characterization of two novel metagenome-derived esterases. *Applied and Environmental Microbiology* **72**, 3637-45.

Embley, T. M. & Stackebrandt, E. (1994). The molecular phylogeny and systematics of the actinomycetes. *Annual Review of Microbiology* **48**, 257-89.

Entcheva, P., Liebl, W., Johann, A., Hartsch, T. & Streit, W. R. (2001). Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia. *Applied and Environmental Microbiology* **67**, 89-99.

Farrell, A. E., Plevin, R. J., Turner, B. T., Jones, A. D., O'Hare, M. & Kammen, D. M. (2006). Ethanol can contribute to energy and environmental goals. *Science* **311**, 506-508.

Feng, J. X., Feng, Y., Duan, C. J., Pang, H., Mo, X. C., Wu, C. F., Yu, Y., Hu, Y. L., Wei, J. & Tang, J. L. (2007). Cloning and identification of novel cellulase genes from uncultured microorganisms in rabbit cecum and characterization of the expressed cellulases. *Applied Microbiology and Biotechnology* **75**, 319-328.

Fermentas (2010). Products - All - Conventional Restriction Enzymes. http://www.fermentas.com/en/products/all/conventional-restriction-enzymes.

Ferrer, M., Chernikova, T. N., Timmis, K. N. & Golyshin, P. N. (2004). Expression of a temperature-sensitive esterase in a novel chaperone-based *Escherichia coli* strain. *Applied and Environmental Microbiology* **70**, 4499-504.

Ferrer, M., Chernikova, T. N., Yakimov, M. M., Golyshin, P. N. & Timmis, K. N. (2003). Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nature Biotechnology* **21**, 1266-1267.

Ferrer, M., Golyshina, O. V., Chernikova, T. N., Khachane, A. N., Martins Dos Santos, V. A., Yakimov, M. M., Timmis, K. N. & Golyshin, P. N. (2005). Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin. *Chemistry & Biology* **12**, 895-904.

Fiedler, H. P. (1993). Biosynthetic capacities of actinomycetes. 1 Screening for secondary metabolites by HPLC and UV-visible absorbance spectral libraries. *Natural Product Research* **2**, 119-128.

Fiedler, H. P., Bruntner, C., Bull, A. T., Ward, A. C., Goodfellow, M., Potterat, O., Puder, C. & Mihm, G. (2005). Marine actinomycetes as a source of novel secondary metabolites. *Antonie van Leeuwenhoek* **87**, 37-42.

Fox, G. E., Pechman, K. R. & Woese, C. R. (1977). Comparative cataloging of 16S ribosomal ribonucleic acid: Molecular approach to procaryotic systematics. *International Journal of Systematic and Evolutionary Microbiology* **27**, 44-57.

Gabor, E., Alkema, W. & Janssen, D. (2004). Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environmental Microbiology* **6**, 879–886.

Gang, D. R., Costa, M. A., Fujita, M., Dinkova-Kostova, A. T., Burlat, V., Martin, W., Sarkanen, S., Davin, L. B. & Lewis, N. G. (1999). Regiochemical control of monolignol radical coupling: A new paradigm for lignin and lignan biosynthesis. *Chemistry & Biology* **6**, 143-151.

Gans, J., Wolinsky, M. & Dunbar, J. (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* **309**, 1387-1390.

Garg, A. P., McCarthy, A. J. & Roberts, J. C. (1996). Biobleaching effect of *Streptomyces thermoviolaceus* xylanase preparations on birchwood kraft pulp. *Enzyme and Microbial Technology* **18**, 261-267.

Ghose, T. K. (1987). Measurement of cellulase activities. *Pure and Applied Chemistry* **59**, 257–268.

Gilkes, N. R., Warren, R. A., Miller, R. C. & Kilburn, D. G. (1988). Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease and the effect on catalysis. *The Journal of Biological Chemistry* **263**, 10401-10407.

Gokhale, D. V. & Deobagkar, D. N. (1989). Differential expression of xylanases and endoglucanases in the hybrid derived from intergeneric protoplast fusion between a *Cellulomonas sp.* and *Bacillus subtilis*. *Applied and Environmental Microbiology* **55**, 2675-2680.

Goldemberg, J. (2007). Ethanol for a sustainable energy future. *Science* **315**, 808-810.

Grant, S., Sorokin, D. Y., Grant, W. D., Jones, B. E. & Heaphy, S. (2004). A phylogenetic analysis of Wadi el Natrun soda lake cellulase enrichment cultures and identification of cellulase genes from these cultures. *Extremophiles* **8**, 421-429.

Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C. & Jahn, D. (2005). JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Research* **33**, W526-531.

Hahn-Hägerdal, B., Jeppsson, H., Olsson, L. & Mohagheghi, A. (1994). An interlaboratory comparison of the performance of ethanol-producing micro-organisms in a xylose-rich acid hydroysate. *Applied Microbiology and Biotechnology* **41**, 62-72.

Hall, J., Hazlewood, G. P., Barker, P. J. & Gilbert, H. J. (1988). Conserved reiterated domains in *Clostridium thermocellum* endoglucanases are not essential for catalytic activity. *Gene* **69**, 29-38.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. & Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology* **5**, R245-249.

Harnpicharnchai, P., Thongaram, T., Sriprang, R., Champreda, V., Tanapongpipat, S. & Eurwilaichitr, L. (2007). An efficient purification and fractionation of genomic DNA from soil by modified troughing method. *Letters in Applied Microbiology* **45**, 387-391.

Henrissat, B. & Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *The Biochemical Journal* **316**, 695-696.

Hermes, J. D., Blacklow, S. C. & Knowles, J. R. (1990). Searching sequence space by definably random mutagenesis: improving the catalytic potency of an enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 696-700.

Higuchi, T. (1990). Lignin biochemistry: Biosynthesis and biodegradation. *Wood Science and Technology* **24**, 23-63.

Hill, J., Nelson, E., Tilman, D., Polasky, S. & Tiffany, D. (2006). Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11206-11210.

Hopwood, D. A. & Wright, H. M. (1978). Bacterial protoplast fusion: recombination in fused protoplasts of *Streptomyces coelicolor*. *Molecular & General Genetics* **162**, 307-317.

Hopwood, D. A. (2006). Soil to genomics: the *Streptomyces* chromosome. *Annual Review of Genetics* **40**, 1-23.

Hopwood, D. A., Wright, H. M., Bibb, M. J. & Cohen, S. N. (1977). Genetic recombination through protoplast fusion in *Streptomyces*. *Nature* **268**, 171-174.

Hu, Y., Zhang, G., Li, A., Chen, J. & Ma, L. (2008). Cloning and enzymatic characterization of a xylanase gene from a soil-derived metagenomic library with an efficient approach. *Applied Microbiology and Biotechnology* **80**, 823-830.

Hurtubise, Y., Shareck, F., Kluepfel, D. & Morosoli, R. (1995). A cellulase/xylanase-negative mutant of *Streptomyces lividans* 1326 defective in cellobiose and xylobiose uptake is mutated in a gene encoding a protein homologous to ATP-binding proteins. *Molecular Microbiology* **17**, 367-377.

Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M. & Omura, S. (2003). Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nature Biotechnology* **21**, 526-531.

Ishikawa, J., Tsuchizaki, N., Yoshida, M., Ishiyama, D. & Hotta, K. (2000). Colony PCR for detection of specific DNA sequences in actinomycetes. *Actinomycetologica* **14**, 1-5.

Iwamoto, S., Abe, K. & Yano, H. (2008). The effect of hemicelluloses on wood pulp nanofibrillation and nanofiber network characteristics. *Biomacromolecules* **9**, 1022-1026.

Jones, A. S. (1953). The isolation of bacterial nucleic acids using cetyltrimethylammonium bromide (cetavlon). *Biochimica et biophysica acta* **10**, 607-612.

Jung, H. G. (1989). Forage lignins and their effects on fiber digestibility. *Agronomy Journal* **81**, 33-38.

Katz, E., Thompson, C. J. & Hopwood, D. A. (1983). Cloning and expression of the tyrosinase gene from *Streptomyces antibioticus* in *Streptomyces lividans*. *Journal of General Microbiology* **129**, 2703-2714.

Kauffmann, I. M., Schmid, R. D. & Schmitt, J. (2004). DNA isolation from soil samples for cloning in different hosts. *Applied Microbiology and Biotechnology* **64**, 665–670.

Keegstra, K., Talmadge, K. W., Bauer, W. D. & Albersheim, P. (1973). The structure of plant cell walls: III. A model of the walls of suspension-cultured sycamore cells based on the interconnections of the macromolecular components. *Plant Physiology* **51**, 188-197.

Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. (2000). *Practical Streptomyces Genetics*. Norwich, United Kingdom: John Innes Foundation.

Kieser, T., Hopwood, D. A., Wright, H. M. & Thompson, C. J. (1982). pIJ101, a multi-copy broad host-range *Streptomyces* plasmid: functional analysis and development of DNA cloning vectors. *Molecular & General Genetics* **185**, 223-228.

Kim, S. J., Lee, C. M., Han, B. R., Kim, M. Y., Yeo, Y. S., Yoon, S. H., Koo, B. S. & Jun, H. K. (2008). Characterization of a gene encoding cellulase from uncultured soil bacteria. *FEMS Microbiology Letters* **282**, 44-51.

Kim, U. J., Shizuya, H., de Jong, P. J., Birren, B. & Simon, M. I. (1992). Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Research* **20**, 1083-1085.

Klein, A., Barsuk, R., Dagan, S., Nusbaum, O., Shouval, D. & Galun, E. (1997). Comparison of methods for extraction of nucleic acid from hemolytic serum for PCR amplification of hepatitis B virus DNA sequences. *Journal of Clinical Microbiology* **35**, 1897-1899.

Knietsch, A., Bowien, S., Whited, G., Gottschalk, G. & Daniel, R. (2003). Identification and characterization of coenzyme B12-dependent glycerol dehydratase-and diol

dehydratase-encoding genes from metagenomic DNA libraries derived from enrichment cultures. *Applied and Environmental Microbiology* **69**, 3048-3060.

Kolpak, F. J. & Blackwell, J. (1976). Determination of the structure of cellulose II. *Macromolecules* **9**, 273-278.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Düsterhöft, A., Ehrlich, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S. Y., Glaser, P., Goffeau, A., Golightly, E. J., Grandi, G., Guiseppi, G., Guy, B. J., Haga, K., Haiech, J., Harwood, C. R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M. F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaerrblanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S. M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R. P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., Oreilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S. H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P., Purnelle, B., et al. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249-256.

Lane, D. J. (1991). 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics*, Stackebrandt, E. & Goodfellow, M., 115–175. Chichester, UK: Wiley and Sons Ltd.

Larsen, N., Olsen, G. J., Maidak, B. L., McCaughey, M. J., Overbeek, R., Macke, T. J., Marsh, T. L. & Woese, C. R. (1993). The ribosomal database project. *Nucleic Acids Research* **21**, 3021-3023.

Lee, C. C., Kibblewhite-Accinelli, R. E., Wagschal, K., Robertson, G. H. & Wong, D. W. (2006). Cloning and characterization of a cold-active xylanase enzyme from an environmental DNA library. *Extremophiles* **10**, 295-300.

Levine, J. (2000). Global biomass burning: a case study of the gaseous and particulate emissions released to the atmosphere during the 1997 fires in Kalimantan and Sumatra, Indonesia. In *Biomass Burning and its Inter-Relationships with the Climate System*, Innes, J. L., Beniston, M. & Verstraete, M. M. , 15-31. Dordrecht: Kluwer Academic Publishers.

Li, A., Antizar-Ladislao, B. & Khraisheh, M. (2007). Bioconversion of municipal solid waste to glucose for bio-ethanol production. *Bioprocess and Biosystems Engineering* **30**, 189-196.

Li, K., Azadi, P., Collins, R., Tolan, J., Kim, J. S. & Eriksson, K. L. (2000). Relationships between activities of xylanases and xylan structures. *Enzyme and Microbial Technology* **27**, 89-94.

Liu, H., Fu, S., Zhu, J., Li, H. & Zhan, H. (2009). Visualization of enzymatic hydrolysis of cellulose using AFM phase imaging. *Enzyme and Microbial Technology* **45**, 274-281.

Lynch, J. M. (1990). Beneficial interactions between micro-organisms and roots. *Biotechnology Advances* **8**, 335-346.

Lynd, L. R., Cushman, J. H., Nichols, R. J. & Wyman, C. E. (1991). Fuel ethanol from cellulosic biomass. *Science* **251**, 1318-1323.

MacCarthy, P. (2001). The principles of humic substances. *Soil Science* **166**, 738-751.

MacNeil, I. A., Tiong, C. L., Minor, C., August, P. R., Grossman, T. H., Loiacono, K. A., Lynch, B. A., Phillips, T., Narula, S., Sundaramoorthi, R., Tyler, A., Aldredge, T., Long, H., Gilman, M., Holt, D. & Osburne, M. S. (2001). Expression and isolation of antimicrobial small molecules from soil DNA libraries. *Journal of Molecular Microbiology and Biotechnology* **3**, 301-308.

Maidak, B. L., Olsen, G. J., Larsen, N., Overbeek, R., McCaughey, M. J. & Woese, C. R. (1996). The Ribosomal Database Project (RDP). *Nucleic Acids Research* **24**, 82-85.

Mandels, M. & Reese, E. T. (1965). Inhibition of cellulases. *Annual Review of Phytopathology* **3**, 85-102.

Margolles, A. & Reyes-Gavilán, C. G. (2003). Purification and functional characterization of a novel α-L-Arabinofuranosidase from *Bifidobacterium longum* B667. *Applied and Environmental Microbiology* **69**, 5096–5103.

Marschner, P., Yang, C. H., Lieberei, R. & Crowley, D. E. (2001). Soil and plant specific effects on bacterial community composition in the rhizosphere. *Soil Biology and Biochemistry* **33**, 1437-1445.

Martinez, A., Kolvek, S. J., Yip, C. L., Hopke, J., Brown, K. A., MacNeil, I. A. & Osburne, M. S. (2004). Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Applied and Environmental Microbiology* **70**, 2452-2463.

Matte, A. & Forsberg, C. W. (1992). Purification, characterization, and mode of action of endoxylanases 1 and 2 from *Fibrobacter succinogenes* S85. *Applied and Environmental Microbiology* **58**, 157–168.

May, O., Nguyen, P. T. & Arnold, F. H. (2000). Inverting enantioselectivity by directed evolution of hydantoinase for improved production of L-methionine. *Nature Biotechnology* **18**, 317-320.

McCarthy, A. J., Peace, E. & Broda, P. (1985). Studies on the extracellular xylanase activity of some thermophilic actinomycetes. *Applied Microbiology and Biotechnology* **21**, 238-244.

Mera, S. L. & Davies, J. D. (1984). Differential congo red staining: the effects of pH, non-aqueous solvents and the substrate. *The Histochemical Journal* **16**, 195-210.

Messing, J., Crea, R. & Seeburg, P. H. (1981). A system for shotgun DNA sequencing. *Nucleic Acids Research* **9**, 309.

Miller, D. N., Bryant, J. E., Madsen, E. L. & Ghiorse, W. C. (1999). Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Applied and Environmental Microbiology* **65**, 4715-4724.

Miyazaki, K., Takenouchi, M., Kondo, H., Noro, N., Suzuki, M. & Tsuda, S. (2006). Thermal stabilization of *Bacillus subtilis* family-11 xylanase by directed evolution. *The Journal of Biological Chemistry* **281**, 10236-10242.

Muto, A. & Osawa, S. (1987). The guanine and cytosine content of genomic DNA and bacterial evolution. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 166-169.

Nakamura, S., Wakabayashi, K., Nakai, R., Aono, R. & Horikoshi, K. (1993). Purification and some properties of an alkaline xylanase from alkaliphilic *Bacillus* sp. strain 41M-1. *Applied and Environmental Microbiology* **59**, 2311–2316.

Ng, T. K., Weimer, T. K. & Zeikus, J. G. (1977). Cellulolytic and physiological properties of *Clostridium thermocellum*. *Archives of Microbiology* **114**, 1-7.

Nieduszynski, I. & Preston, R. D. (1970). Crystallite size in natural cellulose. *Nature* **225**, 273-274.

Niner, B. M., Brandt, J. P., Villegas, M., Marshall, C. R., Hirsch, A. M. & Valdés, M. (1996). Analysis of partial sequences of genes coding for 16S rRNA of actinomycetes isolated from *Casuarina equisetifolia* nodules in Mexico. *Applied and Environmental Microbiology* **62**, 3034-3036.

Nishiyama, Y., Langan, P. & Chanzys, H. (2002). Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron x-ray and neutron fiber diffraction. *Journal of the American Chemical Society* **124**, 9074–9082.

Okazaki, M. & Moo-Young, M. (1978). Kinetics of enzymatic hydrolysis of cellulose: analytical description of a mechanistic model. *Biotechnology and Bioengineering* **20**, 637-663.

Olsen, G. J., Larsen, N. & Woese, C. R. (1991). The ribosomal RNA database project. *Nucleic Acids Research* **19**, 2017-2021.

Ono, A., Miyazaki, R., Sota, M., Ohtsubo, Y., Nagata, Y. & Tsuda, M. (2007). Isolation and characterization of naphthalene-catabolic genes and plasmids from oil-contaminated soil by using two cultivation-independent approaches. *Applied Microbiology and Biotechnology* **74**, 501-510.

Ovreas, L. & Torsvik, V. V. (1998). Microbial diversity and community structure in two different agricultural soil communities. *Microbial Ecology* **36**, 303-315.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences: CABIOS* **12**, 357-358.

Pang, H., Zhang, P., Duan, C. J., Mo, X. C., Tang, J. L. & Feng, J. X. (2009). Identification of cellulase genes from the metagenomes of compost soils and functional characterization of one novel endoglucanase. *Current Microbiology* **58**, 404-408.

Parry, N. J., Beever, D. E., Owen, E., Vandenberghe, I., Van Beeumen, J. & Bhat, M. K. (2001). Biochemical characterization and mechanism of action of a thermostable beta-glucosidase purified from *Thermoascus aurantiacus*. *The Biochemical Journal* **353**, 117-127.

Parveen, S., Portier, K. M., Robinson, K., Edmiston, L. & Tamplin, M. L. (1999). Discriminant analysis of ribotype profiles of *Escherichia coli* for differentiating human and nonhuman sources of fecal pollution. *Applied and Environmental Microbiology* **65**, 3142-3147.

Pettey, T. M. & Crawford, D. L. (1984). Enhancement of lignin degradation in *Streptomyces* spp. by protoplast fusion. *Applied and Environmental Microbiology* **47**, 439-440.

Prabavathy, V. R., Mathivanan, N., Sagadevan, E., Murugesan, K. & Lalithakumari, D. (2006). Intra-strain protoplast fusion enhances carboxymethyl cellulase activity in *Trichoderma reesei*. *Enzyme and Microbial Technology* **38**, 719-723.

Pukall, R., Brambilla, E. & Stackebrandt, E. (1998). Automated fragment length analysis of fluorescently-labeled 16S rDNA after digestion with 4-base cutting restriction enzymes. *Journal of Microbiological Methods* **32**, 55-63.

Raes, J., Korbel, J. O., Lercher, M. J., von Mering, C. & Bork, P. (2007). Prediction of effective genome size in metagenomic samples. *Genome Biology* **8**, R10.

Ramachandra, M., Crawford, D. & Pometto, A. (1987). Extracellular enzyme activities during lignocellulose degradation by *Streptomyces* spp.: a comparative study of wild-type and genetically manipulated strains. *Applied and Environmental Microbiology* **53**, 2754-2760.

Ramachandra, M., Crawford, D. L. & Hertel, G. (1988). Characterization of an extracellular lignin peroxidase of the lignocellulolytic actinomycete *Streptomyces viridosporus*. *Applied and Environmental Microbiology* **54**, 3057-63.

Rees, H. C., Grant, S., Jones, B., Grant, W. D. & Heaphy, S. (2003). Detecting cellulase and esterase enzyme activities encoded by novel genes present in environmental DNA libraries. *Extremophiles* **7**, 415-421.

Renault, P., Corthier, G., Goupil, N., Delorme, C. & Ehrlich, S. D. (1996). Plasmid vectors for gram-positive bacteria switching from high to low copy number. *Gene* **183**, 175-182.

Rheims, H., Spröer, C., Rainey, F. A. & Stackebrandt, E. (1996). Molecular biological evidence for the occurrence of uncultured members of the actinomycete line of descent in different environments and geographical locations. *Microbiology* **142**, 2863-2870.

Rondon, M. R., August, P. R., Bettermann, A. D., Brady, S. F., Grossman, T. H., Liles, M. R., Loiacono, K. A., Lynch, B. A., MacNeil, I. A., Minor, C., Tiong, C. L., Gilman, M., Osburne, M. S., Clardy, J., Handelsman, J. & Goodman, R. M. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and Environmental Microbiology* **66**, 2541-2547.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Sambrook, J. & Russell, D. W. (2001). *Molecular cloning: a laboratory manual*. New York: CSH Laboratory Press.

Sandhu, D. K. & Bawa, S. (1992). Improvement of cellulase activity in *Trichoderma*. *Applied Biochemistry and Biotechnology* **34-35**, 175-183.

Sankoff, D. & Goldstein, M. (1989). Probabilistic models of genome shuffling. *Bulletin of Mathematical Biology* **51**, 117-124.

Schaeffer, P., Cami, B. & Hotchkiss, R. D. (1976). Fusion of bacterial protoplasts. *Proceedings of the National Academy of Sciences of the United States of America* **73**, 2151-2155.

Seow, K. T., Meurer, G., Gerlitz, M., Wendt-Pienkowski, E., Hutchinson, C. R. & Davies, J. (1997). A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *Journal of Bacteriology* **179**, 7360-8.

Sharp, P. M. & Li, W. H. (1987). The codon adaptation index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Research* **15**, 1281-1295.

Sharp, P. M., Bailes, E., Grocock, R. J., Peden, J. F. & Sockett, R. E. (2005). Variation in the strength of selected codon usage bias among bacteria. *Nucleic Acids Research* **33**, 1141-1153.

Shi, D. J., Wang, C. L. & Wang, K. M. (2009). Genome shuffling to improve thermotolerance, ethanol tolerance and ethanol productivity of *Saccharomyces cerevisiae*. *Journal of Industrial Microbiology & Biotechnology* **36**, 139-147.

Sibatani, A. (1970). Precipitation and counting of minute quantities of labeled nucleic acids as cetyltrimethylammonium salt. *Analytical Biochemistry* **33**, 279-285.

Smalla, K., Wieland, G., Buchner, A., Zock, A., Parzy, J., Kaiser, S., Roskot, N., Heuer, H. & Berg, G. (2001). Bulk and rhizosphere soil bacterial communities studied by denaturing gradient gel electrophoresis: plant-dependent enrichment and seasonal shifts revealed. *Applied and Environmental Microbiology* **67**, 4742-4751.

Sogin, S. J., Sogin, M. L. & Woese, C. R. (1971). Phylogenetic measurement in procaryotes by primary structural characterization. *Journal of Molecular Evolution* **1**, 173-184.

Song, J. K. & Rhee, J. S. (2000). Simultaneous enhancement of thermostability and catalytic activity of phospholipase A(1) by evolutionary molecular engineering. *Applied and Environmental Microbiology* **66**, 890-894.

Srinivas, D., Jagannadha Rao, K., Theodore, K. & Panda, T. (1995). Direct conversion of cellulosic material to ethanol by the intergeneric fusant *Trichoderma reesei* QM 9414/*Saccharomyces cerevisiae* NCIM 3288. *Enzyme and Microbial Technology* **17**, 418-423.

Steingrube, V. A., Wilson, R. W., Brown, B. A., Jost, K. C., Blacklock, Z., Gibson, J. L. & Wallace, R. J. (1997). Rapid identification of clinically significant species and taxa of aerobic actinomycetes, including *Actinomadura, Gordona, Nocardia,*

Rhodococcus, Streptomyces, and Tsukamurella isolates, by DNA amplification and restriction endonuclease analysis. *Journal of Clinical Microbiology* **35**, 817-822.

Stoll, D., Stålbrand, H. & Warren, R. A. (1999). Mannan-degrading enzymes from *Cellulomonas fimi*. *Applied and Environmental Microbiology* **65**, 2598–2605.

Strickler, J. E., Berka, T. R., Gorniak, J., Fornwald, J., Keys, R., Rowland, J. J., Rosenberg, M. & Taylor, D. P. (1992). Two novel *Streptomyces* protein protease inhibitors. Purification, activity, cloning, and expression. *The Journal of Biological Chemistry* **267**, 3236-3241.

Strohl, W. R. (1992). Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Research* **20**, 961-974.

Talmadge, K. W., Keegstra, K., Bauer, W. D. & Albersheim, P. (1973). The structure of plant cell walls: I. The macromolecular components of the walls of suspension-cultured sycamore cells with a detailed analysis of the pectic polysaccharides. *Plant Physiology* **51**, 158-173.

Teather, R. M. & Wood, P. J. (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Applied and Environmental Microbiology* **43**, 777-780.

Tebbe, C. C. & Vahjen, W. (1993). Interference of humic acids and DNA extracted directly from soil in detection and transformation of recombinant DNA from bacteria and a yeast. *Applied and Environmental Microbiology* **59**, 2657-2665.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680.

Torsvik, V., Goksoyr, J. & Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Applied and Environmental Microbiology* **56**, 782-787.

Trigo, C. & Ball, A. S. (1994). Is the solubilized product from the degradation of lignocellulose by actinomycetes a precursor of humic substances? *Microbiology* **140**, 3145-3152.

Umezurike, G. M. (1979). The cellulolytic enzymes of *Botryodiplodia theobromae* Pat. Separation and characterization of cellulases and beta-glucosidases. *The Biochemical Journal* **177**, 9-19.

Voget, S., Streit, W. R., Leggewie, C., Uesbeck, A., Raasch, C. & Jaeger, K. E. (2003). Prospecting for novel biocatalysts in a soil metagenome. *Applied and Environmental Microbiology* **69**, 6235-6242.

Waksman, S. A. (1925). What is humus? *Proceedings of the National Academy of Sciences of the United States of America* **11**, 463-468.

Wang, G. Y., Graziani, E., Waters, B., Pan, W., Li, X., McDermott, J., Meurer, G., Saxena, G., Andersen, R. J. & Davies, J. (2000). Novel natural products from soil DNA libraries in a streptomycete host. *Organic Letters* **2**, 2401-2404.

Wang, R. F. & Kushner, S. R. (1991). Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100**, 195-199.

Warnecke, F., Luginbuhl, P., Ivanova, N., Ghassemian, M., Richardson, T. H., Stege, J. T., Cayouette, M., McHardy, A. C., Djordjevic, G., Aboushadi, N., Sorek, R., Tringe, S. G., Podar, M., Martin, H. G., Kunin, V., Dalevi, D., Madejska, J., Kirton, E., Platt, D., Szeto, E., Salamov, A., Barry, K., Mikhailova, N., Kyrpides, N. C., Matson, E. G., Ottesen, E. A., Zhang, X., Hernandez, M., Murillo, C., Acosta, L. G., Rigoutsos, I., Tamayo, G., Green, B. D., Chang, C., Rubin, E. M., Mathur, E. J., Robertson, D. E., Hugenholtz, P. & Leadbetter, J. R. (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**, 560-565.

Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *Journal of Bacteriology* **173**, 697-703.

Wild, J., Hradecna, Z. & Szybalski, W. (2002). Conditionally amplifiable BACs: switching from single-copy to high-copy vectors and genomic clones. *Genome Research* **12**, 1434-1444.

Wilkinson, C. J., Hughes-Thomas, Z. A., Martin, C. J., Böhm, I., Mironenko, T., Deacon, M., Wheatcroft, M., Wirtz, G., Staunton, J. & Leadlay, P. F. (2002). Increasing the efficiency of heterologous promoters in actinomycetes. *Journal of Molecular Microbiology and Biotechnology* **4**, 417-426.

Williams, A. G. (1983). Staining reactions for the detection of hemicellulose-degrading bacteria. *FEMS Microbiology Letters* **20**, 253-258.

Witkin, E. M. (1976). Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriological Reviews* **40**, 869–907.

Woese, C. R. (1987). Bacterial evolution. *Microbiology and Molecular Biology Reviews* **51**, 221-271.

Woese, C. R., Fox, G. E., Zablen, L., Uchida, T., Bonen, L., Pechman, K., Lewis, B. J. & Stahl, D. (1975). Conservation of primary structure in 16S ribosomal RNA. *Nature* **254**, 83-86.

Xu, L., Chen, H., Hu, X., Zhang, R., Zhang, Z. & Luo, Z. W. (2006). Average gene length is highly conserved in prokaryotes and eukaryotes and diverges only between the two kingdoms. *Molecular Biology and Evolution* **23**, 1107-1108.

Yang, V. W., Zhuang, Z., Elegir, G. & Jeffries, T. W. (1995). Alkaline-active xylanase produced by an alkaliphilic *Bacillus* sp. isolated from kraft pulp. *Journal of Industrial Microbiology and Biotechnology* **15**, 434-441.

Zhang, Y. X., Perry, K., Vinci, V. A., Powell, K., Stemmer, W. P. & del Cardayre, S. B. (2002). Genome shuffling leads to rapid phenotypic improvement in bacteria. *Nature* **415**, 644-646.

Zhou, J., Bruns, M. A. & Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology* **62**, 316-322.

de Vries, R. P., Poulsen, C. H., Madrid, S. & Visser, J. (1998). aguA, the gene encoding an extracellular alpha-glucuronidase from *Aspergillus tubingensis*, is specifically induced on xylose and not on glucuronic acid. *Journal of Bacteriology* **180**, 243-249.

van Peij, N. N., Gielkens, M. M., de Vries, R. P., Visser, J. & de Graaff, L. H. (1998). The transcriptional activator XlnR regulates both xylanolytic and endoglucanase gene expression in *Aspergillus niger*. *Applied and Environmental Microbiology* **64**, 3615-3619.