The production of bioethanol from non-food agricultural residues represents an alternative energy source to fossil fuels for incorporation into the world's economy. Within the context of bioconversion of plant biomass into renewable energy using improved enzymatic cocktails, Illumina RNA-seq transcriptome profiling was conducted on a strain of *Aspergillus tamarii*, efficient in biomass polysaccharide degradation, in order to identify genes encoding proteins involved in plant biomass saccharification. Enzyme production and gene expression was compared following growth in liquid and semi-solid culture with steam-exploded sugarcane bagasse (SB) (1% *w/v*) and glucose (1% *w/v*) employed as contrasting sole carbon sources. Enzyme production following growth in liquid minimum medium supplemented with SB resulted in 0.626 and 0.711 UI.mL^{-1} xylanases after 24 and 48 h incubation, respectively. Transcriptome profiling revealed expression of over 7120 genes, with groups of genes modulated according to solid or semi-solid culture, as well as according to carbon source. Gene ontology analysis of genes expressed following SB hydrolysis revealed enrichment in xyloglucan metabolic process and xylan, pectin and glucan catabolic process, indicating up-regulation of genes involved in xylanase secretion. According to carbohydrate-active enzyme (CAZy) classification, 209 CAZyme-encoding genes were identified with significant differential expression on liquid or semi-solid SB, in comparison to equivalent growth on glucose as carbon source. Up-regulated CAZyme-encoding genes related to cellulases (CelA, CelB, CelC, CelD) and hemicellulases (XynG1, XynG2, XynF1, XylA, AxeA, arabinofuranosidase) showed up to a 10-fold log2FoldChange in expression levels. Five genes from the AA9 (GH61) family, related to lytic polysaccharide monooxygenase (LPMO), were also identified with significant expression up-regulation. The transcription factor gene XlnR, involved in induction of hemicellulases, showed up-regulation on liquid and semi-solid SB culture. Similarly, the gene ClrA, responsible for regulation of cellulases, showed increased expression on liquid SB culture. Over 150 potential transporter genes were also identified with increased expression on liquid and semi-solid SB culture. This first comprehensive

analysis of the transcriptome of *A. tamarii* contributes to our understanding of genes and regulatory systems involved in cellulose and hemicellulose degradation in this fungus, offering potential for application in improved enzymatic cocktail development for plant biomass degradation in biorefinery applications.

INTRODUCTION

The production of renewable energy is one of the greatest challenges of the twenty-first Century. Whilst dependency upon fossil fuels is associated with depleting oil reserves and greenhouse gas emissions, plant biomass, by contrast, with its' global abundancy, represents a sustainable and environmentally clean energy source (Goldemberg, 2007; Tan et al., 2016).

The production of bioethanol from non-food agricultural residues such as lignocellulosic trash, grasses and woods, is known as second-generation (2G) ethanol (Alvira et al., 2010), and is considered a promising alternative energy source to fossil fuels for incorporation into the world's economy. Brazil is currently one of the principal agricultural producers, as an important supplier of both food and industrial crops. Sugarcane is planted over an area of almost 9 million hectares, with an annual production of over 620 million tons (CONAB, 2018). Whilst around 45% of the crop production is employed for sugar extraction, the majority is used in the bioethanol industry, with estimates of production of 28 billion liters of anhydrous and hydrated ethanol for 2018/2019 (CONAB, 2018). Bioethanol production in Brazil is based almost exclusively on first-generation technologies, whereby the sucrose content of the plant is converted into ethanol. In this process, sugarcane bagasse will accumulate as an agricultural residue (Goldemberg, 2008). Whilst the burning of bagasse currently serves as an energy source in bioethanol mills, as this biomass represents approximately one-third of the energy content of the crop, the conversion of the lignocellulose component of the cell wall into fermentable hexose (glucose) and pentose (e.g., D-xylose and L-arabinose) sugars offers considerable potential for increased 2G ethanol production, potentially by up to 40% (Amorim et al., 2011). Two Brazilian cellulosic ethanol plants came into operation in 2014, with capacities planned for production of up to 1 billion liters of ethanol per year from bagasse (Silva et al., 2017).

For economically viable 2G ethanol production, complete hydrolysis, or saccharification, of plant biomass is required. Such plant material is composed mainly of polysaccharide crystalline microfibers of cellulose (40–50%), followed in abundance by a matrix of various hemicelluloses and pectins (25–35%), in addition to the polyaromatic lignin (15–20%) (Lin and Tanaka, 2006; Ragauskas et al., 2006; Jordan et al., 2012; Guerriero et al., 2016). Efficient biorefinary-based conversion of this material is hampered due to the recalcitrance of lignocelluloses (Chundawat et al., 2011). In the case of sugarcane bagasse, lignocellulose sugars vary in terms of identity and branching, comprising residues of glucose (60%), xylose (13%), arabinose (6%), mannose (3%), galactose (1,5%), and less than 1% fructose and rhamnose (Häkkinen et al., 2012).

As cellulases and hemicellulases remain costly, increasing the costs of 2G bioethanol production, a continued characterization of sources of such enzymes, together with an improved understanding of the mechanisms involved in enzyme secretion and enzyme efficiency are of fundamental importance for the biofuel industry. Hydrolytic enzymes appropriate for fermentation of available sugars in lignocelluloses are known to be secreted by a wide variety of bacteria and filamentous fungi, with the latter often producing not only a diverse array of extracellular lignocellulolitic enzymes, but also displaying efficiency in secretion of such enzymes in high quantities (Phitsuwan et al., 2013). For this reason, fungi are today the principal source of hydrolytic enzymes for this industrial application (Sims et al., 2010; Couturier et al., 2012).

Lignocellulolytic fungi typically produce extensive sets of carbohydrate-active enzymes (CAZymes) that correlate with their geographical origin habitat (Van Den Brink and De Vries, 2011). A number of ascomycete fungi, notably species members of the genera *Trichoderma* and *Aspergillus,* produce a range of cellulases and hemicellulases, which are today applied across numerous relevant industries for production of food, feed, paper, textiles and pharmaceuticals (Archer, 2000; de Souza et al., 2011). Whilst *Aspergillus niger* and *Trichoderma reseei* are currently employed in the production of commercial enzymatic cocktails for lignocellulosic biomass deconstruction (Singhania, 2011; Mohanram et al., 2013), the identification of additional sources of carbohydrate active enzymes will likely increase efficiency in the deconstruction of this biomass. As such, additional species have recently been screened as potential sources of cellulases, hemicellulases and accessory proteins for optimized industrial enzyme production (Brown et al., 2016; Cong et al., 2017; de Gouvêa et al., 2018).

The availability of whole genome sequences for fungi has improved understanding of fungal biodiversity with respect to plant cell wall degradation. *Aspergillus nidulans*, considered the model species of the genus given its' well-elucidated sexual cycle, possesses a genome sequence of 30.06 MB, with 9396 predicted genes (Galagan et al., 2005). Other characterized Aspergillus species of importance for the food, textile, pulp and paper industries, and potentially in 2G ethanol production, include *A. oryzae* and *A. niger*. *A. oryzae* has a total genome size of 37.12 MB, with 12336 predicted genes (Machida et al., 2005). Similarly, *A. niger* possesses a genome of 37.2 MB, with 14600 predicted genes (Pel et al., 2007). Comparison of gene sequences against

the Carbohydrate-Active Enzymes Database (http://www.cazy.org/) (Cantarel et al., 2009) has revealed 186 genes related to polysaccharide hydrolysis in *A. nidulans*, 217 in *A. oryzae* and 171 in *A. niger* (Delmas et al., 2012).

In addition to gene discovery, the annotated genome sequences for these species serve as resources for analysis of the transcriptome in additional *Aspergillus* species without available genome sequences. Such analysis of transcriptional regulation of genes encoding hydrolytic enzymes in *Aspergillus* has been studied in relation to growth on different sugar carbon sources (Andersen et al., 2008; Jørgensen et al., 2009; Salazar et al., 2009). In relation to fermentation of sugarcane bagasse, microarray analysis provided information on gene expression modulation in *A. niger* (Guillemette et al., 2007; de Souza et al., 2011), with cellulases, hemicellulases and transporters identified with increased expression during growth on sugarcane bagasse in comparison to fructose. Subsequent RNAseq analysis of the *A. niger* transcriptome, following growth on wheat straw compared to simple sugars, revealed a CAZy gene representation change from 3% of total mRNA on 1% glucose to 19% on wheat straw, representing numerous enzymes from the classes of Glycoside Hydrolases (GH), Carbohydrate Esterases (CE), and Polysaccharide Lyases (PL) (Delmas et al., 2012). Further RNAseq-based analysis of gene expression in Aspergillus species following growth on sugarcane bagasse as carbon source has also revealed important information regarding regulatory mechanisms and genes encoding plant cell wall degrading enzymes, accessory proteins and transporters (Pullan et al., 2014; Brown et al., 2016; Borin et al., 2017; Cong et al., 2017; de Gouvêa et al., 2018).

The continued characterization of hydrolases, accessory proteins and the regulation of their expression in Aspergillus species that display efficiency in degradation of lignocellulose will further our understanding of their roles in saccharification. Given the importance of *Aspergillus tamarii* as an efficient producer of enzymes such as xylanases (El-Gindy et al., 2015; Monclaro et al., 2016), we utilized an Illumina RNA-seq approach to analyze the transcriptome in this fungus following semi-solid and liquid cultivation on steam-exploded bagasse (SB) compared gene expression following growth on glucose (G). Genes encoding cellulases and hemicellulases, transcription factors and transporters are characterized in relation to their differential expression following fungal growth on each carbon source. Data will benefit the development of improved fungal strains with increased ability to deconstruct lignocellulose and generate value-added bioproducts.

MATERIALS AND METHODS

Strain and Culture Conditions

A stock culture of a strain of *A. tamarii*, code BLU37, was provided by the fungal culture collection at the Enzymology Laboratory, University of Brasilia, Brazil (genetic heritage number 010237/2015-1). The strain was originally isolated into pure culture from natural composting cotton textile waste material in the Vale do Itajaí, Santa Catarina, Brazil (Siqueira et al., 2009) and maintained in the culture collection at −80°C in 50% glycerol.

Species identity reconfirmation was conducted by sequence analysis of the nuclear ribosomal DNA (rDNA) ITS1-5.8S-ITS2 region, together with specific regions of the β-tubulin and calmodulin genes. Genomic DNA was extracted according to Raeder and Broda (1985) from a 3 day old liquid culture in Czapek Yeast Extract medium (CYA) (Pitt and Hocking, 2009) incubated on an orbital shaker at 28°C. Each PCR reaction contained 10 ng genomic DNA, 2,5 $mmol^{-1}$ of each primer, 1 $mmol^{-1}$ dNTPs, 4 $mmol^{-1}$ $MgCl_2$, 1U of Taq Platinum® polymerase (Invitrogen) and 1 x Taq Platinum® polymerase buffer (Invitrogen). Ribosomal DNA ITS regions were amplified using primers ITS5 and ITS4 (White et al., 1990), a β-tubulin gene region with primers Bt_2a and Bt_2b (Glass and Donaldson, 1995), and a calmodulin gene region amplified with primers Cmd5 and Cmd6 (Hong et al., 2006). PCR cycling was performed with the following programs: initial denaturation at 94°C for 4 min, 30 cycles of denaturation at 94°C for 1 min, primer annealing for 1 min, at 50°C for primers ITS5 and ITS4, and at 60°C for primers Bt2a, Bt2b, Cmd5 and Cmd6, extension at 72°C for 1 min, and a final extension period at 72°C for 5 min. PCR products were purified using ExoSAP-IT® (USB, Cleveland, Ohio, USA) and sequenced using Big Dye® Terminator v3.1 Cycle Sequencing chemistry (Applied Biosystems, Foster City, CA, USA) on an ABI 3700 DNA sequencer (Applied Biosystems, Foster City, CA, USA). For molecular identification, sequences were compared against the nucleotide database NCBI using the BLASTn algorithm (Altschul et al., 1990). Ribosomal DNA ITS, β-tubulin and calmodulin gene sequences were deposited in GenBank under accession numbers MH540359, MH544272 and MH544273, respectively.

For analysis of gene expression in BLU37 following exposure to SB or glucose as carbon source, the strain was grown in either liquid or semi-solid minimal medium (KH_2PO_4 7 g; K_2HPO_4 2 g; $MgSO_4$ 0.4 g; $(NH_4)_2SO_4$ 1.6 g, pH 7.0, per liter of distilled water), containing SB (1% *w/v*) or glucose (1% *w/v*) (Sigma Aldrich) as exclusive carbon source. In order to guarantee elimination of reducing sugars, prior to fungal inoculation, SB was repeatedly washed with deionized water until reducing sugars were no longer detectable by the colorimetric dinitrosalicylic acid (DNS) assay (Miller, 1959). Liquid cultures were grown in 100 mL of media in Erlenmeyer flasks, whilst semi-solid cultures were grown on petri plates with media supplemented with agar (15 g L^{-1}). Fungal spores at a concentration of 1×10^8 conidiospores mL^{-1} were used as inocula, with cultures then incubated at 28°C and 150 rpm for 36 and 48 h. Fungal cultures were arranged in a randomized block design, with three replicates for each treatment and time point. Growth treatments were labeled as follows: liquid medium, SB carbon source, 36 h incubation (LB36); liquid medium, glucose carbon source, 36 h incubation (LG36); liquid medium, SB carbon source, 48 h incubation (LB48); liquid medium, glucose carbon source, 48 h incubation (LG48); semi-solid medium SB carbon source, 36 h incubation (SB36); semi-solid medium, glucose carbon source, 36 h incubation (SG36); semi-solid medium, SB carbon source, 48 h incubation