Metagenomic Mining of Feruloyl Esterases from Termite Enteric Flora

Konanani Rashamuse[1§*], Tina Ronneburg[1,2§], Walter Sanyika[1,2], Kgama Mathiba[1], Edwin Mmutlane[1], and Dean Brady[1,2]

[1] CSIR Biosciences, Meiring Naudé Road; Brummeria; Pretoria; 0001, South Africa

[2] Departments of Biotechnology and Food Technology, Tshwane University of Technology, Pretoria, 0001, South Africa

* Corresponding Authors: Mailing address: Dr Konanani Rashamuse
CSIR Bioscience
Building 18
PO Box 395
Pretoria, 0001
South Africa

Tel.: (+2712) 841 3682

Fax: (+2712) 841 3388

E-mail: KRashamuse@csir.co.za

[§]The authors contributed equally to this work

Key words: Termite, Feruloyl esterases, Metagenomics, Biorefinery

Abstract

A metagenome expression library was created from *Trinervitermes trinervoides* termite hindgut symbionts and subsequently screened for feruloyl esterase (FAE) activities resulting in seven recombinant fosmids conferring feruloyl esterase phenotypes. The amino acid sequence lengths of the seven FAE encoding open reading frames (ORFs) ranged from 260 to 274 aa and encoded polypeptides of between 28.9 to 31.4 kDa. The highest sequence identity scores for the seven ORFs against GenBank database were between 45 and 59% to a number of carboxyl ester hydrolyses. The seven FAE primary structures contained sequence motifs that corresponding well with a classical penta-peptide (G-x-S-x-G) serine hydrolyse signature motif which harbours the catalytic serine residue in other FAE families. Six of the seven *fae* genes were successfully expressed heterologously in *Escherichia coli* and the purified enzymes exhibited temperature optima range of 40-70^{o}C and the pH optima of between 6.5 and 8.0. The k_{cat}/K_{M} ratios for the six characterised FAEs showed the following order of substrate preference: methyl sinapate > methyl ferulate > ethyl ferulate. All six FAEs showed poor conversion rates against methyl *p*-coumarate and methyl caffeate both of which lacked the methoxy (O-CH_3) group substituent on the aromatic ring of the ester substrates, emphasising the requirement for at least one methoxy group on the aromatic ring of the hydroxycinnamic acid ester substrate for optimal FAE activity.

Introduction

Plant cell walls are the most abundant source of organic carbon on earth (Gilbert 2010). The structural composition of different plant cell walls have been shown to be predominantly made up of polysaccharides including celluloses, hemicelluloses and pectins (Gilbert 2010). These polysaccharides are ester-linked to hydroxycinnamic acid derivates (such as coumaric, sinapic and ferulic,) which undergo *in vivo* oxidative coupling reactions to either polymerize into lignin macromolecules or form dehydrodimers of the acids (Ishii 1997; Fry 1986).

The ferulated polysaccharides are critical entities in directing cell wall cross-linkages, which also serve as a defensive mechanism against the invasion by plant pathogens (Mohnen et al. 2008). Thus, any successful bio-based technology for the conversion of plant biomass would to a large extent depend on the depolymerization of the covalent cross-linkages between cell wall polysaccharides and hydroxycinnamic acids (Faulds 2010; Wong 2006). Besides ester linkages, ether bridges are also known besides dimerization and oligomerization of the hydroxycinnamic acids (Ishii 1997). The recent developments in the fields of enzymatic bioconversion of plant cell walls for biofuel and biochemical feedstocks productions have highlighted the requirement for the synergistic interaction between plant cell wall accessory enzymes (such as feruloyl esterases, acetyl xylan esterases and α-L-arabinofuranosidases) and lignocellulases (cellulases, xylanases and pectinases) which directly hydrolyse the plant biopolymers into sugar monomers (Biely 2003; Collins et al. 2005; Gilbert 2010; Faulds 2010).

Feruloyl esterases (FAEs, EC 3.1.1.73), also often referred to as phenolic acid, sinapic acid or cinnamoyl esterases, represent a subclass of carboxylester hydrolases (EC 3.1.1.-) that catalyze the release of hydroxycinnamic acids that are generally found esterified to

polysaccharides such arabinoxylans and pectins (Crepin et al. 2004). The removal of the hydroxycinnamic acids by FAEs allows the depolymerisation of cellulose and hemicellulose polymers from each other, making the plant cell wall more accessible to cellulases and hemicellulases (Benoit et al. 2008; Scharf and Tartar 2008). In addition to catalyzing the hydrolytic release of hydroxycinnamic acid derivatives, FAEs catalyzing trans-esterification reactions of ferulolyated oligosaccharides have also been reported (Vafiada et al. 2009; Hatzakis et al. 2003).

Feruloyl esterases are have been classified into four types (A-D), based on their amino acid sequences, their preference towards mono- and di-ferulates, and their specificity for particular substitutions on the hydroxycinnamic acid ring (Crepin et al. 2004). Recently a new classification scheme has also been developed based on the sequence-derived descriptors and pharmacophoric properties (Udath et al. 2011). The applications of these enzymes in the industry cover a broad spectrum; including pulp and paper processing (Record et al. 2003) and as animal feed additives to facilitate nutrient assimilation (Krueger et al. 2008; Nsereko et al. 2008). Recently, there has been renewed interest in these enzymes not only because of their important role in the lignocellulosic biomass conversion for renewable energy production, but also for the release of phenolic acids which are valuable by-products for various downstream bioprocesses (Laszlo et al. 2003; Faulds 2010). For example, hydroxycinnamic acid derivatives (such as *p*-coumaric, ferulic, sinapic and caffeic acid) have widespread potential applications due to their antimicrobial, photoprotectant, anti-tumour and antioxidant properties (Graf 1992), as well as their use as flavour precursors (Shiyi and Kin-Chor 2004).

Termite species are very efficient in digesting lignocellulosic based materials with the support of microbial symbionts associated with the hindgut (Wheeler et al. 2009; Warnecke et al. 2007; Watanabe et al. 1998). The P3 section of the termites' hindgut in particular serves as a metabolic engine involved in the breakdown of a wide variety of compounds with

the aid of enzymes derived from the host termites as well as the hindgut associated obligate symbionts (Arima and Woodley 2008). A recent report investigating microbial diversity of the P3 section of the hindgut (of a higher termite species, *Trinervitermes trinervoides*) using culture-independent metagenomic approach revealed a high level of prokaryotic symbionts diversity associated with the hindgut (Sanyika et al. 2012), making the hindgut particularly the P3 section an ideal source for screening of novel enzymes. Esterases have long been implicated to be part of termites' physiology and metabolism, mainly involved in the metabolic pathways that provide carbon and energy sources (Davis et al. 1995; Ruvolo-Takasusuki and Collet 2000). Esterolytic activities associated with either the host termites or associated mutualistic symbionts is well documented in the literature (Wheeler et al. 2009; Rashamuse et al. 2012) but specific esterolytic activities relating to the lignocellulosic deconstruction in the termite hindgut has not yet fully exploited. In this report we describe the construction of a shotgun metagenomic library using a community DNA derived from the hindgut prokaryotic symbionts of the *Trinervitermes trinervoides* termite species and subsequent functional screening and biochemical characterisation of recombinant feruloyl esterases associated with lignocellulosic biomass conversion.

Materials and Methods

Strains and plasmids

E. coli EPI100-T1^{R} (F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ-rpsL nupG trfA dhfr) was used as the host for cloned environmental DNA (Epicentre Biotechnologies, USA). *E. coli* DH5α was used as the host for routine sub-cloning, while *E. coli* BL21 (DE3) was used as the expression host. The cloning vector, pJET1.2 was used for sub-cloning of the PCR products and enzyme-restricted DNA fragments. The vector pET20b (+) (Merck) was used for protein expression. All the restriction and modification enzymes were used as recommended by manufacturers. Primers were synthesized and purchased from Inqaba Biotech (Pretoria, South Africa). The standard molecular biology protocols were followed as described by Sambrook and Russell (2001).

Termite sampling and DNA extraction

Sugarcane feeding *Trinervitermise trinervoides* termite species were collected from Komatipoort, Mpumalanga, South Africa (25° 23' 7.8" S 31° 52' 53.1" E). Approximately 500 worker termites were dissected and the aqueous content of the P3 hindgut collected as previously reported (Warnecke et al. 2007). A total high molecular weight community DNA from the hindgut was isolated using a method previously described by Gilliespie et al. (2005) with some modifications. The modifications of the protocol included a 2 minute glass bead (0.5 mm diameter) beating step using a Genie disruptor (Scientific industries Inc., USA) and the addition of a column purification step using the 40 kb cut-off spin column (Zymo Research, USA).

Library construction and screening

The purified high molecular weight DNA (>30 kb) was subjected to metagenome library creation using the EpiFOS™ Fosmid Library Production Kit (Epicentre Biotechnologies, USA). Primary screening of recombinant fosmid clones in *E. coli* EPI300-T1^{R} was performed on LB agar plates supplemented with isopropyl-β-D-thiogalactoside (IPTG, 0.1 mM), chloramphenicol (12.5 µg/ml), tributyrin 1 % (v/v) and gum arabic 0.1 % (w/v), followed by incubation at 37 °C. The identified esterase positive clones on tributyrin agar were further screened for FAE activities on 0.4 % ethyl ferulate agar plates (Donaghy et al.1998).

Nucleotide sequencing and sequence analysis

High throughput 454-nucleotide sequencing (Roche) was performed by Inqaba Biotech (South Africa). Sequence analysis and manipulations were performed using CLC Combine Workbench (CLCBIO, Denmark) and BioEdit (Hall 1990) software with the aid of the Basic Local Alignment Search tool for proteins (BLASTP) search engine (Altschul et al. 1997). The signal peptide predictions were conducted using SignalP 3.0 server located at http://www.cbs.dtu.dk/services/SignalP/ (Bendten et al. 2004). Theoretical M_W and pI were predicted using the EXPASY server (http://www.expasy.ch) (Appel et al. 1994). Polymerase chain reactions (PCR) amplifications were achieved by using the HiFi™ DNA polymerase (KAPA Biosystems, South Africa).

Phylogenetic analysis

Non-redundant database sequences from the NCBI database (http://blast.ncbi.nlm.nih.gov/Blast) were used as reference controls for the inference of phylogenetic position of the FAE sequences identified in this study. Based on the sequence identities, the closely related protein sequences were identified through homology searches using the BLASTP (Altschul et al. 1997) against the NCBI non-redundant database. Sequence alignments and editing were done using the BioEdit software (Hall 1990). The alignments were manually inspected and trimmed to equal length in BioEdit. The *MEGA 4*

software (Tamura et al. 2007) was used for phylogenetic inference using the Neighbour-Joining method (Saitou and Nei 1987). The acetyl xylan esterase (Axe) sequences were used as an out-group to root the phylogenetic tree. A 1000 bootstrap replicates were used with pair-wise deletion of gaps.

Recombinant expression of FAE encoding genes

The primer pairs corresponding to the respective full length gene sequences (*fae*1-7), were designed to introduce *Nde*I and *Xho*I restriction sites at the 5´ and 3´ end of the genes respectively, for directional cloning of the seven *fae* genes into the pET20b (+) expression vector. The reverse primers lacked the stop codon to allow the recombinant proteins to be expressed in-frame with the 6x-His tag sequence at the C-terminus of the proteins. The amplified PCR products were digested with *Nde*I/*Xho*I DNA restriction enzymes, followed by ligation into the pET20b (+) plasmid linearised with the same enzymes. The resulting expression constructs (p20FAE1-7) were placed under the control of the T7 promoter, inducible with IPTG.

Protein expression and purification

Overnight *E. coli* BL21 cultures (5 ml) harbouring p20FAE1-7 expression constructs were used to inoculate 50 ml LB broth/ ampicillin (100 µg ml^{-1}). The cultures were grown at 37 ºC to an OD_{600nm} of between 0.6 - 0.8. Induction was initiated by the addition of IPTG (1 mM final concentration) and the cultures were further incubated at 25 ºC for at least 6 hours, followed by centrifugation (8 000 *x g*, 10 min, 4 ºC). Harvested cells were re-suspended in 10 ml bacterial protein extraction reagent (B-PER, Pierce, USA) and incubated at room temperature for 10 min. The suspension was centrifuged (13 000 *x g*, 10 min, 4 ºC) and the supernatant fractions were used as a source for intracellular soluble proteins.

Solid ammonium sulphate was added with rapid stirring to 15 % saturation, followed by incubation on ice for at least 1 h at 4 ^{o}C and precipitated proteins were removed by

centrifugation (20 000 *x g*, 30 min, 4 ^{o}C). The supernatant was loaded onto a HiLoad^{TM} 16/10 Phenyl Sepharose high performance hydrophobic interaction column (GE healthcare) pre-equilibrated with 50 mM Tris-HCl (pH 7.5) containing 0.3 M ammonium sulphate. Proteins bound to the hydrophobic matrix were eluted over 10 column volumes at a flow rate of 2 ml min^{-1} with a linear gradient of decreasing ammonium sulphate concentration (0.3 - 0 M) in Tris-HCl buffer (50 mM, pH 7.5). Fractions containing esterase activity were pooled and concentrated using 10 kDa cut-off VIVAspin column before loading onto a HiPrep^{TM} 26/10 Sephacryl size exclusion column (GE Healthcare), pre-equilibrated with 50 mM Tris-HCl (pH 7.5) containing 0.3 M NaCl. Fractions containing esterase activity were pooled and the protein concentrations of the purified samples were estimated by the method of Bradford (1976), using bovine serum albumin (BSA) as a standard.

Determination of Protein molecular Masses

The molecular mass subunits of the purified proteins were determined by sodium dodecyl sulphate-polyacrylamide gel eletrophoresis (SDS-PAGE) based on the Laemmli (1970) protocol. Proteins were visualized following Coomassie brilliant blue G20 staining. The molecular mass of the native proteins were determined using analytical Superdex 200 GL 10/300 gel filtration column, pre-equilibrated with 50 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl. The following proteins were used as reference standards: Thyroglobulin (670 000 Da), γ-Globulin (160 000 Da), Ovalbumin (44 000 Da), Myoglobin (17 000 Da) and Vitamin B12 (135 Da).

Feruloyl esterase assay

Unless otherwise stated, routine feruloyl esterase activities were performed in triplicate. One unit of enzyme activity was defined as the amount of enzyme that releases 1 µmol of ferulic acid per minute from ethyl ferulate and the specific activity was given in units per milligram of protein. The reaction mixture (1 ml) was carried out in MOPS buffer (50 mM, pH 6.5) that contained 1 mM methyl ferulate (dissolved in acetonitrile/ K_2HPO_4 (90:10)). The reactions

were started upon the addition of 10 µl of the appropriately diluted enzyme solutions. Reactions were incubated at 40 °C for 5 min and were stopped by addition of 200 µl HCl (36 % (v/v)). The identical reaction mixture excluding the enzyme was included as a control to correct for autohydrolysis of the substrate.

HPLC Analysis

The released hydroxycinnamic acid derivatives (ferulic acid, sinapate, *p*-coumarate and caffeate) were quantified using Waters 2690 HPLC system equipped with the diode array detector 996. Separation was achieved on a XTerra® MS C18 3.5 µm (3.0 x 50 mm) column with a guard column. The detector wavelength was set at 320 nm. The mobile phases were at a flow rate of 0.5 ml min^{-1} and involved isocratic triflouroacetic acid: acetonitrile (70:30 (v/v)) for ethyl ferulate (EFA), methyl ferulate (MFA) and methyl *p*-coumarate (MpCA) substrates, and a linear gradient of triflouroacetic acid: acetonitrile (90:10 (v/v)) was used for chlorogenic acid and methyl caffeate containing substrate reactions.

Biochemical characterisation

All biochemical characteristics for the six FAEs were determined by using EFA (1 mM) as a substrate. The temperature optima of the purified enzymes were determined between 30 – 80 ^{o}C. The standard reaction mixture was incubated at the desired temperatures and the reactions were started by adding appropriately dilute enzyme volume.

The thermostability profiles of the purified enzymes were determined by incubating the enzyme in MOPS buffer (100 mM, pH 6.5) at temperatures ranging 10 ^{o}C above and below the temperature optimum of each individual enzyme. The residual activities were determined at timed intervals by measuring the release of ferulic acid. The effect of pH on the purified FAEs was determined over a pH range of 3 - 11 in 50 mM universal buffer (Britton and Robinson, 1931).

Substrate specificity preferences for the purified FAEs were determined using the standard assay in the presence of 1 mM of the following esters: Ethyl ferulate (EFA), methyl ferulate (MFA), methyl *p*-coumarate (M*p*CA), methyl sinapate (MSA), methyl caffeate (MCA) and chlorogenic acid (ChA). Kinetic parameters for the purified FAEs were determined for EFA, MFA and MSA as starting substrates. Initial substrate concentrations were varied between 0.25 - 3 mM. The experimental data of initial velocities *versus* substrate concentrations were used to determine the kinetic constants (V_{max}, $K_{M,}$ k_{cat}, k_{cat}/K_M) of the respective purified enzymes.

General esterase activities were determined using 1 mM of *p*-nitrophenyl esters as described by Rashamuse et al. (2009). Substrate preferences were determined using 1 mM of the specified *p*-nitrophenyl esters of various chain lengths: *p*-nitrophenyl acetate (C2), *p*-nitrophenyl butyrate (C4), *p*-nitrophenyl valerate (C5), *p*-nitrophenyl octanoate (C8), *p*-nitrophenyl dedecanoate (C12) and *p*-nitrophenyl palmitate (C16). Experimental initial velocities data vs substrate concentrations (with coefficients of variation of (<5%) were fitted to the Michaelis–Menten equation by weighted non-linear regression (hyperbola) analysis.

Accession Numbers

The nucleotide sequences of the seven *fae* genes were submitted to GenBank, under the following accession numbers: FAE1 (KC493563), FAE2 (KC493564), FAE3 (KC493565), FAE4 (KC493566), FAE5 (KC493567), FAE6 (KC493568) and FAE7 (KC493569).

Results

Metagenomic library construction and screening

A fosmid library (4 x 10^5 cfu ml^{-1}) with an average insert size of 30 kb was constructed using the CopyControl Fosmid Library Production Kit which represented 1.3 Gp of cloned DNA from the termite hindgut symbionts. Only 10% (~40 000 colony forming units, CFU) of the library was screened and a total of sixty eight (68) esterase positive fosmids were identified on tributyrin agar, representing a hit rate of 1:588. The 68 esterase positive fosmids identified were further screened for feruloyl esterase (FAE) activity on ethyl ferulate agar plates (Donaghy et al.1998). A total of seven fosmid clones (pFosFae 1-7) tested positive for FAE activity as indicated by zones of clearance around wells containing recombinant fosmid cultures. The restriction fragment length polymorphisms (RFLP) of the seven positive clones using *Bam*HI and *Not*I restriction enzymes showed non-redundant patterns and the estimated insert sizes ranging between 31 and 37 kb.

Nucleotide Sequencing and analysis of Deduced Primary structures

In order to locate the genes encoding FAE activities within the seven fosmids, a random 454-shotgun sequencing of the DNA inserts was performed. Translational analysis of the nucleotide sequences of the seven fosmid clones resulted in seven open reading frames (ORFs) encoding putative FAE activities being identified. The nucleotide sequence lengths of the FAE encoding ORFs ranged from 780 to 822 bp, with the GC content of between 36 and 54% (Table 1). Translational analysis of the seven ORFs revealed polypeptides ranging between 260 to 274 aa, and the corresponding predicted subunit molecular masses ranged of between 28.9 and 31.4 kDa. A common esterase G-x-S-x-G corresponding motif was identified in all seven FAEs primary structures (Supplementary Data: Fig. S1). The deduced amino acid sequences of FAEs exhibited typical α/β-hydrolase fold features based on the InterProScan tertiary structure prediction program (http://www.expasy.ch). All seven ORFs

lacked the N-terminal leader peptide sequences in their primary structures (Bendten et al. 2004), suggesting that the encoded proteins are either secreted intracellularly or exported to the preplasmic or extracellular environment through other secretion pathways other than the two-step Xcp-dependent secretion pathways that are mediated by signal peptide sequences (Bendten et al. 2004).

Homology Searches and Phylogenetic classification

The deduced amino acid sequences were employed to search for identical protein sequences in the GenBank database from a global amino acid alignment. The search report revealed highest sequence identities of between 45 - 59 % to a number of prokaryotic carboxyl ester hydrolases (Table 1). A pair-wise sequence matrix showing the identities of FAE amino acid sequences amongst each other revealed that the primary structures of these FAEs were significantly different, with the percentage differences ranging from 36 - 84% (Supplementary Data: Table S1).

To gain a further understanding regarding the evolutionary relationships of the identified FAEs, their amino acid sequences were compared with the related sequences from previously characterised feruloyl esterases (and acetyl xylan esterases as an out-group) using neighbouring-joining distance analyses, yielding a phylogenetic tree shown in Fig. 1. The FAE sequences from this study clustered together independently of the four known FAE families (type A-D) reported by Crepin et al (2004). The independent clustering of the FAE sequences observed suggested that the FAE sequences identified in this study form a novel family of feruloyl acid esterases. The application of a recently developed FAEs classification scheme which is based sequence-derived descriptors and pharmacophoric properties (Udatha et al. 2011) classified the seven FAE sequences identified into feruloyl esterase family 1B (FEF 1B). None of family 1B homologues have been biochemical characterised to date (Udatha et al. 2011) and as a result biochemical characterisation of the seven FAEs will provide for the first time biochemical data for this family.

Recombinant expression and Purification

In order to heterologously express *fae1-7* genes, the full length sequences were directionally ligated into pET20b (+) to allow the expression of the genes under a strong T7 promoter and in-frame with the His-tag sequence. With the exception of FAE3 (which was largely expressed as insoluble biological inactive inclusion bodies), the rest of FAEs were recombinantly produced in the soluble cytoplasmic fraction of corresponding *E. coli* cells (Supplementary Data: Fig. S2).

Initial attempts to purify soluble recombinant FAEs using one-step immobilized metal ion affinity chromatography (IMAC) purification procedure failed, with more than 95% of the recombinant enzymes recovered in the flow-through fractions. The western blot analysis using *anti*-His antibody confirmed the presence of the His-tags sequence in all FAEs (Data not shown). In addition, sequencing of the expression constructs confirmed that all the genes were in-frame with the His-tag.

A classical protein purification procedure involving three steps (ammonium sulphate fractionation, hydrophobic interaction and size exclusion chromatography) was then followed. The purification data table for the purified FAEs is shown in Table S2 (supplementary Data). The final yields of between 3.0 - 23% were achieved while purifications folds increased to between 2.4 and 19.8. The purified FAEs showed respective protein bands using SDS PAGE, which were in agreement with the theoretical molecular mass subunit predicted from the amino acid sequences (Supplementary Data: Fig. S3). The native molecular mass determination using analytical size exclusion column revealed that FAE1, FAE2, FAE4, FAE7 had dimeric native structures, while FAE5 and FAE6 possibly had trimeric native structures (supplementary Data: Table S3). It was presumed that the multimeric forms of these enzymes resulted in the His-tag folding inside the proteins, making

the His-tags inaccessible to resin binding and hence the unsuccessful attempt to purify these enzymes based on the one-step IMAC purification procedure.

Biochemical characterisation

The purified FAEs revealed a narrow pH range optima of between 6.5 and 8.0. FAE4 revealed a pH optimum of 6.5. The pH optima for FAE5 and FAE6 was at 7.0, while that of FAE1 and FAE2 was at 8.0. Both FAE1 and FAE4 showed broader pH stability compared to the other enzymes, retaining more that 60% activity between pH 6.0 – 8.0 (Table 2). FAE2, FAE5 and FAE7 showed narrow pH stability with 60% activity retention between pH 7.0 - 8.0. The same narrow pH stability was observed with FAE6 between pH 6.0 and 7.0.

The purified FAEs showed different temperature optima: 50 ^{o}C for FAE1, 60 ^{o}C for FAE2, 70 ^{o}C for FAE4, 40 ^{o}C for FAE5, FAE6 and FAE7 (Fig S5: Supplementary Data). FAE1 showed moderate thermostability with a half life of 1.5 h at 50 ^{o}C. FAE2 showed good thermostability with half life >8 h at 60 ^{o}C. FAE4 showed excellent thermostability ($t_{1/2}$ = 2.5 h) at 80 ^{o}C and $t_{1/2}$ > 8 h at 70 °C. FAE4, showed no significant activity changes between 30-60 ^{o}C over an 8 h period. Although the temperature optima for FAE5, FAE6 and FAE7 were determined to be 40 ^{o}C, these enzymes were significantly thermostable ($t_{1/2}$ > 8 h) at 50 ^{o}C (Table 2 and Fig S4: supplementary Data).

The initial rates of hydrolysis for the six FAEs were determined for different hydroxycinnamate esters (Table 3). Maximal hydrolytic rates were obtained with methyl ferulate (for FAE1, FAE4 and FAE6) and methyl sinapate (for FAE2, FAE5 and FAE7). In addition of showing highest preference to methyl ferulate, FAE1 also showed significant activity (10% conversion) against chlorogenic acid, which could not be hydrolysed by any of the other five FAEs (Data not shown). All FAEs showed low conversion rate (less than 50%) against methyl caffeate and methyl *p*-coumarate.

A summary of the kinetic constants of the six FAEs against three most readily hydrolyzed substrates (EFA, MFA and MSA) is shown in Table 4. The K_M values indicated highest affinity for MSA, followed by MFA and EFA. All six FAEs displayed high catalytic turnover (k_{cat}) against EFA. The catalytic efficiencies (as measured by k_{cat}/K_M ratios) suggested that the three substrates (EFA, MFA and MSA) were being converted at different efficiencies under the conditions investigated, with MSA being the most favoured substrate by all six enzymes.

The substrate specificities of the purified FAEs against different fatty acyl substrates were determined using *p*-nitrophenyl esters of C2 - C16 (Table S4: Supplemetary Data). FAE2 selectivity was highest (100%) against *p*-NP-C2, while FAE4, FAE5 and FAE7 showed maximum activity (100%) against *p*-NP-C4. FAE1 and FAE6 showed highest preference to *p*-NP-C5. With an exception of FAE6 which retained 80% against *p*-NP-C8, the hydrolytic activities of the other five FAEs decreased dramatically against medium chain length esters (C8 - C12). All FAEs were inactive against the longest chain length tested suggesting that they are true carboxyl ester hydrolyses and not lipases (Bornscheuer, 2002).

Discussion

Activity-based screening of metagenomic libraries has been shown to be very useful in discovering new classes of enzymes with novel and useful functions (Elend et al. 2006). In this study the application of the culture-independent metagenomic approach, coupled with functional screening was demonstrated and led to the discovery of seven feruloyl esterase encoding genes and their subsequent purification and biochemical characterisation.

All known esterase are serine hydrolases, with the putative catalytic serine residue typically located within the classical pentapeptide (G-X-S-X-G) signature sequence motifs (where X represents any amino acids (Bornscheuer 2002; Jaeger et al. 1999). The three-dimensional structures of known esterases have all revealed that the topological location of the serine residue within the G-X-S-X-G is conserved (Bornscheuer 2002). The serine residue has been found located at the top of a tight bend (also referred to as a nucleophilic elbow) found in all esterase structures reported to date (Bornscheuer 2002). This nucleophilic elbow has been shown to form only when the amino acids at the -2 and +2 positions relative to the catalytic serine have small side chains, explaining the predominance of glycine residues at this location (Villeneuve et al. 2000). A multiple sequence alignment was used to analyse the primary structures of the seven feruloyl esterases (Supplementary Data: Fig S1). Analyses of the amino acid sequences revealed the presence of the following sequences: G-L-S-M-G sequences in FAE1, FAE2, FAE3, FAE4, FAE5 and FAE7, and G-I-S-M-G in FAE6, which correspond well with the classical (G-x-S-x-G) pentapeptide sequence. The G-x-S-x-G sequence motif has been shown to harbour the catalytic serine in other feruloyl esterases and many other carboxylesterases (Bornscheuer 2002; Jaeger et al. 1999; Arpigny and Jaeger 1999).

Furthermore, the alignment showed the presence of the corresponding region (-HG-, amino acid positions 54-55) within the FAE primary structures. The oxyanion region is

characterised by a conserved -HG- dipeptide and a short tetra-peptide hydrophobic stretch upstream the conserved HG dipeptide (Bornscheuer 2002). The oxyanion region has been proven to stabilize the tetrahedral intermediates of feruloyl esterase hydrolytic reactions (Hermoso et al. 2004). The conserved aspartic and histidine residues (amino acid positions 233 and 251, respectively), which form a hydrogen network with the catalytic serine (amino acid position 131) were also observed from the alignment. Based on the multiple sequence alignment, the following catalytic triad residues were deduced: FAE1 (Ser131, Asp222, and His251), FAE2 (Ser126, Asp215, and His244), FAE3 (Ser128, Asp215, and His244), FAE4 (Ser125, Asp216, and His245), FAE5 (Ser124 Asp215 and His244), FAE6 (Ser123 Asp212 and His241), and FAE7 (Ser121 Asp215, and His244).

Moreover, the multiple sequence alignment also revealed highly conserved tryptophan residues at position 63 (Supplementary Data: Figure S1), with an exception of FAE6 in which the Tyr63 residue was replaced by less polar Phe63. Based on the highly conserved nature of Tyr63, we suggest that Tyr63 and Phe63 could play a role analogous to that proposed for Tyr80 from the x-ray crystal structures of the *Aspergillus niger* FAEA (Hermoso et al. 2004). In AnFaeA structure, ferulic acid interacts through the OH group at C4 and the OCH_3 group at the C3 with Tyr80-OH in the AnFAEA-substrate complex (Hermoso et al. 2004). Previous mutational studies have also shown that the replacement of Tyr80 by a non aromatic residue caused a significant loss in FAE activity (Tarbouriech et al. 2005).

Recently, there have been reports on FAEs from termite gut environment (Chandrasekharaiah et al. 2011, Wheeler et al. 2009). However, a comparison of the FAE amino acid sequences identified in this study against the four esterases (RfEst1, RfEst2, RfEst3 and RfEst4) from the gut of the termite *Reticulitermes flavipes* (Wheeler *et al.,* 2009) showed poor sequence conservation across the alignment (Data not shown). Moreover, no significant sequence identity (20%) could be established between FAE sequences from this study and the FAE from the symbionts of termite gut of *Coptotermes formosanus*

(Chandrasekharaiah *et al.,* 2011). Moreover, no significant sequence identity could also be established between FAE sequences from this study and those FAE sequences from other metagenome studies (Table S1: Supplementary Data).

A considerable number of recombinant FAEs have been over expressed and purified from a number of genera (Crepin et al. 2004). Amongst the previously characterised FAEs, the majority are monomeric, although some dimeric and tetrameric forms have also been reported (Crepin et al. 2004; Topakas et al. 2005; Wong 2006). In general, the subunit molecular masses of the known FAEs range from 29 to 58 kDa, with an exception of FAE_Xyn10A and 10B from *Clostridium thermocellum* and FAE1 from *Aureobasidium pullulans* that have subunit molecular masses of 90, 116 and 201 kDa, respectively (Rumbold et al. 2003). The native molecular mass of the six characterised FAEs were mostly dimeric, with an exception of FAE5 and FAE6 which possibly had a trimeric native structure (Table S3: Supplementary Data).

Initial rates of hydrolysis of the six FAEs showed low conversion rate against M*p*CA and MCA both of which lacked the methoxy group substitute in their aromatic ring. The observations support the previous structural studies that indicated that at least one methoxy group substituent on the aromatic ring of hydroxycinnamic acid derivative is a requirement for catalytic activity (Andersen et al. 2002). Substrate specificities of FAEs against various cinnamic acid methyl esters differs widely in literature. For example, FAEs from *A. niger* (Kroon et al. 1996), *Penicillium expansum* (Donaghy and McKay 1997) and *Penicillium funiculosum* (Kroon et al. 2000) showed poor preference for MSA, whereas *A. niger* FAE III (Faulds and Williamson 1994) shows no activity towards MCA.

The kinetic constants of the six FAEs showed high affinity and substrate preference for MSA, which has two methoxy group substitutes on the aromatic ring. These observations further support homology modelling studies, which showed that the increase in the hydroxy

substitution decreases the FAE affinity to the substrate (Andersen et al. 2002) while increasing the number of methoxy groups increases the FAE affinity to the substrate (Andersen et al. 2002). The k_{cat}/K_M ratios in this study showed the following order of substrate preference: MSA>MFA>EFA, which was in contrast to a number of fungal FAEs (Rumbold et al. 2003; Faulds and Williamson 1994; Castanares et al. 1992) which showed the following order of preference: MpCA>MCA >MFA>MSA.

The thermostability profile of FAE4 deserves special mention. FAE4 has a temperature optimum of 70^{o}C and a half life of 2.5h at 80^{o}C. Literature search reveals that majority of known FAEs are optimally active between 37 to 60^{o}C (Wong et al. 2006), with an exception of feruloyl esterase Tx-Est1 from *Thermobacilllus xylanilyticus* (Rakotoarivonina et al. 2011) and TfFAE from *Thermoanaerobacter tengcongensis* (Abokitse et al. 2010) which represent two of the most thermostable FAEs reported to date. Tx-Est1 had a half life of 30 min at 65^{o}C, while TfFAE had a half life of 50 min at 80 ^{o}C. Thus, FAE4 characterised in this study represents one of the most thermostable FAE known to date and could be useful in the development of a biorefinery process at high temperatures (Turner et al. 2007).

The FAEs characterised in this study showed no lipase activity on olive oil/Rhodamine B lipase specific assay (Kouker and Jaeger 1987), suggesting that the enzymes were “true” carboxyl ester hydrolases (Bornscheuer 2002). The observation was further confirmed by substrate specific profiling (using *p*-nitrophenyl esters of different chain lengths) which revealed that FAEs preferred short chain lengths, a typical characteristic of true carboxylesterases (Jaeger et al. 1999). Due to the lack of the following substrates; FAX (2-*O*[5-*O*-*trans*-feruloyl)-β-L-arabinofuranosyl]-D-xylopyranose) and FAXX (*O*-[5-*O*-(*trans*feruloyl)-α-L-arabinofuranosyl]-(1->3)-*O*-β-D-xylopyranosyl-(1->4)-D-xylopyranose), we could not classify the identified FAEs using Crepin et al. (2004) feruloyl esterase classification scheme.

In conclusion, we have characterised six FAEs from *Trinervitermes trinervoides* hindgut symbionts with different physiochemical properties and substrate preferences that may be useful in developing an enzyme cocktail for the breakdown of different ferulated polysaccharides. We are currently evaluating the synergistic interactions of the six FAEs together with xylanases and cellulases for the release of sugar monomers from sugar baggase.

Acknowledgements

The work was supported by the Technology Innovation Agency of South Africa and the CSIR-ZA. The authors would also like to thank Mr Harris Tshwane Manchidi, Johan Kemp, and Dr Fritha Hennessey for the help with sample collection. Special thanks to Cornel Du Toit, UP Department of Entomology, University of Pretoria for the help with termite collection and dissection and Dr Daniel Visser for his input during the screening programme. TR and KR contributed equally to this work.

References

Abokitse K, Wu M, Bergeron H, Grosse S, Lau P (2010) Thermostable feruloyl esterase for the bioproduction of ferulic acid from triticale bran. Appl Microbial Biotechnol 87:195-203

Altschul SF, Madden TS, Schäffer, AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25:3389-3402

Andersen A, Svendsen A, Vind J, Lassen SF, Hjort,C, Borch K, Patkar SA (2002) Studies on ferulic acid esterase activity in fungal lipases and cutinases. Colloids Sur Biointer 26:47-55

Appel RD, Bairoch A, Hochstrasser DF (1994) A new generation of information retrieval tools for biologists: the example of the EXPASY WWW server. Trends Bioche Sci 19: 258-260

Arima K and Woodley J (2008) Computational methods for understanding bacterial and archeal genomes; In Xu Y, Gogatern P (eds): Metagenomics. CH 14, Imperial College Press, pp 345-350

Arpigny KL and Jaeger KE (1999) Bacterial lipolytic enzymes: classification and properties. J Biochem 343:177-183

Bendten JD, Nielsen H, von Heijnie G, Brunak S (2004). Improved prediction of signal peptide: SignalP 3.0. J Mol Biol 340:783-795

Benoit I, Danchin EJ, Bleichrodt RJ, de Vries RP (2008). Biotechnological applications and potential of fungal feruloyl esterases based on prevalence, classification and biochemical diversity. Biotechnol Lett 30:387–396.

Biely P (2003) Xylanolytic Enzymes. In: J.R. Whitaker, A. Voragen and D. Wong, Editors, Handbook of Food Enzymology, Marcel Dekker Inc., NY, pp. 879–916

Bornscheuer UT (2002) Microbial carboxylesterases: classification, properties and application in biocatalysis. FEMS Microbiol Rev 26:73-81

Bradford MM (1976) A rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248-254

Britton HT K and Robinson RA (1931) Universal buffer solutions and the dissociation constant of veronal. *J Chem Soc* 1456–1462

Castanares A, McCrae SI, Wood TM (1992) Purification and properties of a feruloyl/p-coumaroyl esterase from the fungus *Penicillium pinophilum*. Enzyme Microbial Technol 14:875–884

Chandrasekharaiah M, Thulasi A, Bagath M, Kumar DP, Sunil Singh Santosh, SS ,Palanivel C, Vazhakkala LJ, Sampath, KT (2011) Molecular cloning, expression and characterization of a novel feruloyl esterase enzyme from the symbionts of termite (*Coptotermes formosanus*) gut. BMB Reports 44:52-53

Collins T, Gerday C, Feller G (2005) Xylanases, xylanase families and extremophilic xylanases. FEMS Microbiol Rev 29:3–23

Crepin VF, Faulds CB, Connerton IF (2004) Functional classification of the microbial feruloyl esterases. Appl Microbiol Biotechnol 63:647–652

Davis RW, Kamble ST, Prabhakaran SK (1995): Characterization of general esterases in workers of the Eastern subterranean termite. J Econ Entomol 88:574–578

Donaghy J, and McKay AM (1997) Purification and characterization of a feruloyl esterase from the fungus *Penicillium expansum*. J Appl Microbiol 83:718–726

Donaghy J, Kelly PF, McKay AM (1998) Detection of ferulic acid esterase production by *Bacillus sp*. and *Lactobacilli*. Appl Microbiol Biotechnol 50:257–260

Elend C, Scheisser C, Leggewie C, Babiak P, Carballeira D, Steele L, Reymond L, Jaeger K, Streit W (2006) Isolation and biochemical characterization of two novel metagenome-derived esterases. Appl Environ Microbiol 72:3637-2645

Faulds CB (2010) What can feruloyl esterases do for us. Phytochem Rev 9:121–132

Faulds CB and Williamson G (1994) Purification and characterization of a ferulic acid esterase (FAE-III) from *Aspergillus niger* specificity of the phenolic moiety and binding to microcrystalline cellulose. Microbiology 140:779–787

Fry SC (1986) Cross-linking of matrix polysaccharides in the growing cell walls of angiosperms. Annu Rev Plant Physiol 37:165-186

Gilbert HJ (2010) The biochemistry and structural biology of plant cell wall deconstruction. Plant Physio 153:444-455

Gilliespie DE, Rondon MR, Goodman RM, Handelsman J, Williamson LL (2005). Metagenomic library from uncultured microorganisms In: Osborn, A.M and Smith, C.J. Molecular microbial ecology, Taylor and Francis group, New York Ch1, pp 261-279.

Graf E (1992) Antioxidant potential of ferulic acid. Free Radic Biol Med 13:435–448

Hall TA (1990) BioEdit: A user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl acids Symp. Ser. 41:95-98

Hatzakis NS, Daphnomili D, Smonou I (2003). Ferulic acid esterase from *Humicola insolens* catalyzes enentioselective transesterificstion of secondary alcohols. J Mol Catal 21: 309-311

Hermoso JA, Aparicio SJ, Molina R, Juge N, Gonzalez R, Faulds CB (2004) The crystal structure of feruloyl esterase A from *Aspergillus niger* suggest evolutive functional convergence in feruloyl esterase family. J Mol Biol 338:495-506

Ishii T (1997) Structure and functions of feruloylated polysaccharides. Plant science 127:111-127

Jaeger KE, Dijkstra BW, Reetz. MT (1999) Bacterial biocatalysis: molecular biology, three-Dimensional structures and biotechnological applications of lipases. Ann Rev Microbiol 53:315-351

Kouker G, and Jaeger KE (1987) Specific and sensitive plate assay for bacterial lipases. Appl Environ Microbiol 53:211-213

Kroon PA, Faulds CB, and Williamson G (1996) Purification and characterization of a novel esterase induced by growth of *Aspergillus niger* on sugar-beet pulp. Biotechnol Appl Biochem 23:255–262

Kroon P, Williamson G, Fish NM, Archer DB, Belshaw NJ (2000) A modular esterase from *Penicillium funiculosum* which releases ferulic acid from plant cell walls and binds crystalline cellulose contains a carbohydrate binding module. European Journal of Biochem 267:6740-6752

Krueger NA, Adesogan AT, Staples CR, Krueger WK, Dean DB, Littell RC (2008) The potential to increase digestibility of tropical grasses with a fungal, ferulic acid esterase enzyme preparation. Animal Feed Sci Technol 145:95–108

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.

Laszlo JA, Compton DL, Eller FJ, Taylor SL, Isbell TA (2003) Packed-bed bioreactor synthesis of feruloylated monoacyl- and diacylglycerols: clean production of a ‘‘green’’ sunscreen. Green Chem 5:382–386

MohneN D, Bar-Peled M, Somerville C (2008) Biosynthesis of plant cell wall. In ME Himmerl, Ed, Biomass recalcitrance. Blackwell Publishing, Oxford, pp 266-277

Nsereko VL, Smiley BK, Rutherford WM, Spielbauer A, Forrester KJ, Hettinger, GH, Harman EK, Harman BR (2008) Influence of inoculating forage with lactic acid bacterial strains that produce ferulate esterase on ensilage and ruminal degradation of fiber. Animal Feed Sci Technol 145:122–135

Rakotoarivonina H, Hermant B, Chabbert B, Touzel JP, Remond C (2011) A thermostable feruloyl esterase from the hemicellulolytic bacterium *Thermobacillus xylanilyticus* relases phenolic acids from non pretreated plant cell walls. Appl Microbial Biotechnol 90:541-552

Rashamuse K, Mabizela-Mokoena N, Sanyika, W, Mabvakure B, Brady D (2012) Accessing diversity from termite hindgut symbionts through metagenomics. J Mol Microbiol Biotechnol. 222:277-86

Rashamuse KJ Magomani V, Ronneburg T, Brady D (2009) A novel family VIII carboxylesterase derived from a leachate metagenome library exhibits promiscuous beta-lactamase activity on nitrocefin. Appl Microbiol Biotechnol 83:491-500

Record E, Asther M, Sigoillot C, Pages S, Delattre M, Haon M, van den Hondel CA, Sigoillot JC, Lesage-Meessen L, Asther M (2003) Overproduction of the *Aspergillus niger* feruloyl esterase for pulp bleaching applications. Appl Microbiol Biotechnol 62:349–355

Rumbold K, Biely P, Mastihubova M, Marinka Gudelj M, Gubitz M, Robra KH, Prior BA (2003) Purification and Properties of a Feruloyl Esterase Involved in Lignocellulose Degradation by *Aureobasidium pullulans*. 69:5622–5626

Ruvolo-Takasusuki MCR, Collet T (2000) Characterization of *Nasutitermes globiceps* esterases. Biochem Genet 38:367–375

Saitou N, Nei, M (1987) The neighbour-joining method: A new method for reconstructing phylogenetic trees. J Mol Bio Evol 4:406-425

Sambrook, J, Russell DW (2001) Molecular cloning: a laboratory manual. 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanyika TW, Rashamuse KJ, Hennessy F, Brady D (2012) Luminal hindgut bacterial diversities of the grass and sugarcane feeding termite *Trinervitermes trinervoides*. African J Micro Res 6:2639-2648

Scharf ME and Tartar A. (2008) Termite digestomes as sources for novel lignocellulases. Biofuels Bioprod Biorefin 2:540–552

Shiyi O and Kin-Chor K (2004) Ferulic acid: pharmaceutical functions, preparation and applications in foods. J Sci of Food and Agric 84:1261-1269

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 10:1093/molbev/msm1092.

Tarbouriech N, Prates JA, Fontes CM, Davies GJ (2005) Molecular determinants of substrate specificity in the feruloyl esterase module of xylanase 10B from *Clostridium thermocellum*. Acta Crystallogr D 61:194-197

Topakas E, Christakopolous P, Faulds C (2005) Comparison of mesophilic and thermophilic feruloyl esterases: Characterization of their substrate specificity for methyl phenylalkanoates. J Biotechnol 115:355–336

Turner P, Mamo G, Gand Karlsson EN (2007) Potential and utilisation of thermophiles and thermostable enzymes in biorefining. Microb Cell Fact 6:1-23

Udatha DB, Kouskoumvekaki I, Olsson L, Panagiotou G (2011). The interplay of descriptor-based computational analysis with pharmacophore modeling builds the basis for a novel classification scheme for feruloyl esterases. Biotechnol Adv 29:94-110

Vafiadi C, Nahmias VR, Faulds CB,Christakopoulos P (2009) Feruloyl esterase-catalysed synthesis of glycerol sinapate using ionic liquids mixtures. J Biotechnol 139:124-129

Villeneuve P, Muderhwa JM, Graille J, Haas MJ (2000) Customizing lipases for biocatalysis: a survey of chemical, physical and molecular biological approaches. J Mol Catal 9:113-148.

Warnecke F, Luginbu¨hl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT, Cayouette M,McHardy AC, Djordjevic G, Aboushadi N, Sorek R, Tringe SG, Podar M, Martin HG, Kunin V, Dalevi D, Madejska J, Kirton E, Platt D, Szeto E, Salamov A, Barry K, Mikhailova N, Kyrpides NC, Matson EG, Ottesen EA, Zhang X, Herna´ndez M, Murillo C, Acosta LG, Rigoutsos I, Tamayo G, Green BD, Chang C, Rubin EM, Mathur EJ, Robertson DE, Hugenholtz P, Leadbetter JR (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature 450:560–565

Watanabe H, Noda H, Lo N (1998) A cellulase gene of termite origin. Nature 394:330–331

Wheeler MM, Tarver MR, Coy MR, Scharf EM (2009) Characterization of four esterase genes and esterase activity from the gut of the termite *Reticulitermes flavipes*. Archive of Insect Biochem and Physiol 73:30–48

Wong, D.W.S, (2006) Ferolyl esterase: A key enzymes in Biomass degradation. Appl. Biochem and Biotech 133:87-109

List of figures

Figure 1: Phylogenetic relationships between FAEs identified in the study: FAE1 (KC493563), FAE2 (KC493564), FAE3 (KC493565), FAE4 (KC493566), FAE5 (KC493567), FAE6 (KC493568) and FAE7 (KC493569) and GenBank obtained FAEs: NcFaeD-3.660 from *Neurospora crassa*, PfFaeA from *Penicillium funiculosum* (AJ312296), PfXYLD from *Pseudomonas fluorescens* XYLD (X58956), NcFaeD-3.544 from *N. Crassa*, PeESTA from *Piromyces equi* ESTA (AF164516), NcFae-1 from *N. crassa* (AJ293029), PfFaeB from *Pen. funiculosum* (AJ291496), AaAXE from *Aspergillus awamori* AXE (D87681), AnAXE from *Asp. niger* AXE (A22880), AfAXE from *Asp. ficuum* AXE (AF331757), PpAXEI from *Pen. purpurogenum* AXE I (AAM93261), OspFaeA from *Orpinomyces sp.* PC-2 FaeA (AF164351), CtXynZ from *Clostridium thermocellum* XynZ (M22624), RspXyn1 from *Ruminococcus sp*. Xyn1 (S58235), RfXynE from *R. flavefaciens* XynE (AJ272430), CtXynY from *C. thermocellum* XynY (X83269), RaXynB from *R. albus* XynB (AB057588), RfAXE from *R. flavefaciens* AXE (AJ238716), OspAXE from *Orpinomyces sp.* PC-2 AXE A (AAC14690), AnFaeB from *Asp. niger* FaeB (AJ309807), TsFaeC from *Talaromyces stipitatus* FaeC (AJ505939), AoEST from *Asp. oryzae* selective esterase (CAD28402), AspChloro from *Acinetobacter sp*. chlorogenate esterase (AAL54855), AoTAN from *Asp. oryzae* tannase (JC5087), RfXynB from *R. flavefaciens* (Z35226), AnFaeA from *Asp. niger* (AF361950), AaFaeA from *Asp. awamori* (Q9P979), AtFaeA from *Asp. turbingensis* (Y09331), TlLipase from *Thermomyces lanuginosus* (O59952), RmLipase from *Rhizomucor miehei* (P19515), PpAXE from *P. purporogenum* xylan esterase II (AF015285), TrAXE from *T. reesei* acetyl xylan esterase I (S71334), BfCINI from *Butyrivibrio fibrisolvens* E14 cinnamoyl ester hydrolase I (T44624), BfCINII from *B. fibrisolvens* E14 cinnamoyl ester hydrolase II (AAB57776), NpAXE from *Neocallimastix patriciarum* AXE (U66253), ApESTA from *Asp. parasiticus* ESTA esterase (AF417002), AspAEST from *Acinetobacter sp*. AEST esterase (BAB68337), BpAXE from *Bacillus pumilus* AXE (AJ249957), SlAXE from *Streptomyces lividans* AXE (AAC06115)

List of Tables:

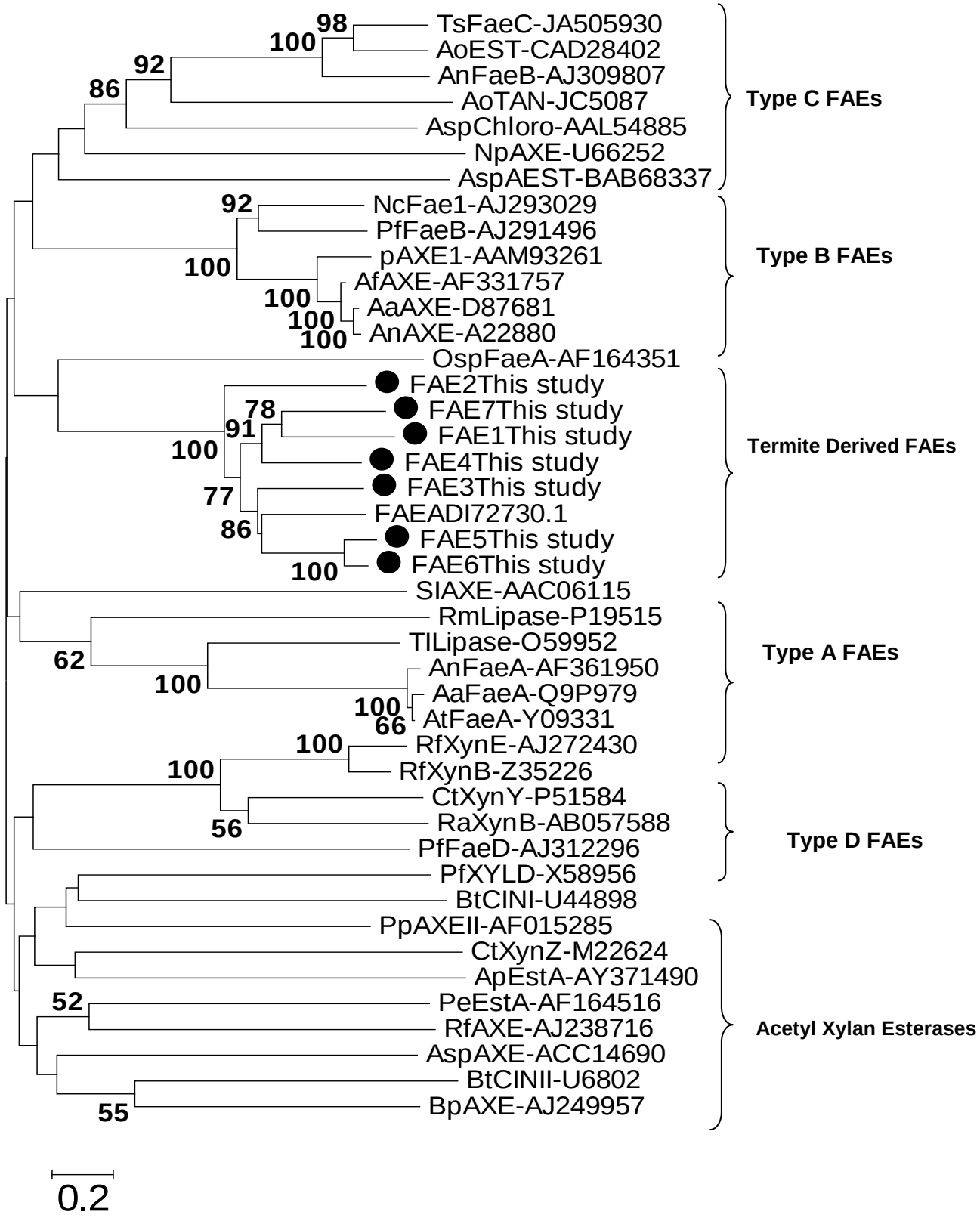
98
100
92
86
TsFaeC-JA505930
AoEST-CAD28402
AnFaeB-AJ309807
AoTAN-JC5087
AspChloro-AAL54885
NpAXE-U66252
AspAEST-BAB68337
Type C FAEs
92
100
100
100
100
NcFae1-AJ293029
PfFaeB-AJ291496
pAXE1-AAM93261
AfAXE-AF331757
AaAXE-D87681
AnAXE-A22880
Type B FAEs
OspFaeA-AF164351
78
91
100
77
86
100
FAE2This study
FAE7This study
FAE1This study
FAE4This study
FAE3This study
FAEADI72730.1
FAE5This study
FAE6This study
Termite Derived FAEs
SIAXE-AAC06115
62
100
100
66
RmLipase-P19515
TILipase-O59952
AnFaeA-AF361950
AaFaeA-Q9P979
AtFaeA-Y09331
Type A FAEs
100
100
RfXynE-AJ272430
RfXynB-Z35226
56
CtXynY-P51584
RaXynB-AB057588
PfFaeD-AJ312296
Type D FAEs
PfXYLD-X58956
BtCINI-U44898
PpAXEII-AF015285
CtXynZ-M22624
ApEstA-AY371490
52
PeEstA-AF164516
RfAXE-AJ238716
AspAXE-ACC14690
BtCINII-U6802
55
BpAXE-AJ249957
Acetyl Xylan Esterases
0.2

Figure 1:

Table 2:

Genes name	nt. length (bp)	GC content (%)	a.a. length	Mw, kDa (pI)	Identity (Similarity) %	Best Hit (Accession number)	Organism name
fae1	822	50.7	274	31.4 (5.65)	51 (67)	putative esterase (ZP_05495774.1)	*Clostridium papyrosolvens*
fae2	780	51.7	263	29.8 (4.93)	49 (65)	acetylesterase (ZP_04857794.1)	*Ruminococcus sp.*
fae3	789	53.9	260	28.9 (5.32)	56 (68)	putative esterase (ZP_06883368.1)	*Clostridium lentocellum*
fae4	804	42.3	268	31.1 (5.71)	59 (72)	putative esterase (YP_002572111.1)	*Anaerocellum thermophilum*
fae5	801	46.7	267	30.1 (5.16)	45 (63)	putative esterase (YP_002572111.1)	*Caldicelulosiruptor becscii*
fae6	783	36.3	261	30.1 (5.54)	59 (72)	Predicted esterase (CBK96609.1)	*Eubacterium siraeum*
fae7	789	38.2	261	29.6 (5.67)	58 (72)	Predicted esterase (CBK96609.1)	*Clostridium lentocellum*

a.a. = amino acids, bp = base pairs, nt = nucleotide.

Table 3:

Parameters		Units	FAE1	FAE2	FAE4	FAE5	FAE6	FAE7
pH optimum		pH	7.0	7.0	7.0	7.0	7.0	7.0
pH range (>60%)		pH	6.0-8.0	7.0-8.0	6.0-8.0	7.0-8.0	6.0-7.0	7.0-8.0
Temperature optimum		^{o}C	50	60	70	40	40	40
Temperature range (>60%)		^{o}C	40-60	40-60	50-90	40-60	30-60	20-50
Temperature stability ($T_{1/2}$)	30^{o}C	hr	>8	>8	>8	>8	>8	>8
	40^{o}C	hr	>8	>8	>8	>8	7	>8
	50^{o}C	hr	1.5	>8	>8	>8	<1	>8
	60^{o}C	hr	ND	>8	>8	6	ND	6
	70^{o}C	hr	ND	<1	>8	ND	ND	ND
	80^{o}C	hr	ND	ND	2.5	ND	ND	ND

Table 4:

Substrates	FAE1	FAE 2	FAE 4	FAE 5	FAE 6	FAE 7
	Relative rates of hydrolysis (%)					
Ethyl ferulate (EFA)	70.0 (±9.5)	67.7 (±7.0)	81.1 (±6.8)	75.0 (±7.2)	76.3 (±0.7)	77.9 (±3.6)
Methyl ferulate (MFA)	100.0 (±6.8)	73.1 (±5.3)	100.0 (±3.9)	82.0 (±2.7)	100.0 (±2.7)	87.8 (±4.7)
Methyl sinapate (MSA)	54.2 (±4.0)	100.0 (±2.7)	85.7 (±5.8)	100.0 (±5.6)	54.3 (±4.7)	100.0 (±5.0)
Methyl *p*-coumarate (MpCA)	10.9 (±8.1)	24.2 (±1.8)	29.1 (±8.8)	13.8 (±7.0)	19.7 (±1.4)	18.4 (±0.0)
Methyl caffeate (MCA)	10.5 (±4.2)	38.8 (±8.7)	8.5 (±7.2)	22.2 (±3.3)	8.3 (±12.0)	2.7 (±9.4)
Chlorogenic Acid (ChA)	16.1 (±0.2)	0.0	0.0	0.0	0.0	0.0

All experiments were performed in triplicates and the values represent the means ± standard error from the triplicate values measured

Table 4:

Substrate	FAE1				FAE2				FAE4			
	K_M (mM)	V_{max} (U mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	K_M (mM)	V_{max} (U mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	K_M (mM)	V_{max} (U mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
EFA	2.8	66.3	530.2	192.8	0.7	15.15	142.4	203.4	0.7	22.38	183.3	261.8
	(±0.28)	(±5.4)	(±14.3)	(±15.5)	(±0.1)	(±1.06)	(±10.2)	(±14.6)	(±0.18)	(±2.07)	(±16.9)	(±13.2)
MFA	0.4	21.6	172.9	465.1	0.50	15.85	148.9	297.8	0.8	31.05	254.4	318.6
	(±0.08)	(±2.6)	(±12.1)	(±26.8)	(±0.1)	(±0.68)	(±6.3)	(±12.6)	(±0.16)	(±2.10)	(±17.2)	(±20.4)
MSA	0.1	67.2	539.9	5390.	0.3	9.03	84.8	282.6	0.4	17.98	147.3	368.5
	(±0.06)	(±0.7)	(±15.8)	(±45.5)	(±0.1)	(±0.38)	(±5.4)	(±12.9)	(±0.07)	(±0.08)	(±6.8)	(±17.65)
Substrate	**FAE5**				**FAE6**				**FAE7**			
	K_M (mM)	V_{max} (U mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	K_M (mM)	V_{max} (U mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})	K_M (mM)	V_{max} (U mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
EFA	2.1	53.2	335.7	159.8	0.8	39.94	326.5	408.1	2.5	61.88	332.4	132.9
	(±0.6)	(±8.5)	(±15.4)	(±12)	(±0.18)	(±3.50)	(±28.7)	(±35.3)	(±0.7)	(±10.07)	(±54.1)	(±16.8)
MFA	1.5	47.9	298.4	198.9	0.6	48.11	462.8	771.3	1.7	51.90	278.8	164.7
	(±0.3)	(±3.34)	(±21.7)	(±13.7)	(±0.14)	(±3.47)	(±33.3)	(±58.3)	(±0.3)	(±4.07)	(±21.8)	(±13.9)
MSA	0.9	31.7	197.9	219.8	0.11	12.61	121.3	1102.7	1	33.97	182.5	182.5
	(±0.1)	(±1.46)	(±9.1)	(±9.8)	(±0.01)	(±0.42)	(±4.42)	(±37.4)	(±0.01)	(±1.54)	(±8.7)	(±8.8)

Enzyme concentration in the assay = 5 µg ml^{-1}.

k_{cat} was calculated assuming a single site per monomeric protein.