Pectinmethylesterase from the rice weevil, Sitophilus oryzae: cDNA isolation and sequencing, genetic origin, and expression of the recombinant enzyme

Authors: Shen, Zhicheng, Pappan, Kirk, Mutti, Navdeep S., He, Qi-Jiong, Denton, Michael, et al.

Pectinmethylesterase from the rice weevil, *Sitophilus oryzae*: cDNA isolation and sequencing, genetic origin, and expression of the recombinant enzyme

Zhicheng Shen[1], Kirk Pappan[3]*, Navdeep S. Mutti[2], Qi-Jiong He[3], Michael Denton[3]**, Yu Zhang[2], Michael R. Kanost[3], John C. Reese[2], and Gerald R. Reeck[3]

[1] Institute of Applied Entomology, Zhijiang University, Hangzhou, China
[2] Department of Entomology, Kansas State University, Manhattan, Kansas 66506
[3] Department of Biochemistry, Kansas State University, Manhattan, Kansas 66506
* Current address: Department of Pathology, Washington University School of Medicine, St. Louis, MO 63110
**Current address: Northrop Grumman Information Technology, 8111-18th Street, Building 803, Brooks AFB, San Antonio, TX 78235
reeck@ksu.edu

Abstract

A cDNA clone encoding pectinmethylesterase of the rice weevil, *Sitophilus oryzae* (L.) has been isolated and sequenced. The cDNA clone was expressed in cultured insect cells and active pectinmethylesterase was purified from the culture medium, thus confirming that the cDNA encodes pectinmethylesterase. *In situ* hybridization indicated that the enzyme's transcript was present in the midgut. Weevils treated with tetracycline so that they lack genes of known symbiotic organisms still contained the pectinmethylesterase gene, indicating that the gene is encoded by the rice weevil genome. The rice weevil enzyme is most similar in sequence to bacterial pectinmethylesterases. Given this and the enzyme's apparently rather general absence from animal species, we suggest the possibility that this gene was transferred horizontally to an ancient weevil, possibly from a bacterial symbiont, and exists in *Sitophilus* species now as a result of that ancestral horizontal transfer.

Keywords: pectinase, pectin esterase, polygalacturonase, gut, digestive enzyme, pectin, symbiont, *Wolbachia*, SOPE, gene transfer

Introduction

Polygalacturonase and pectinmethylesterase from the rice weevil, *Sitophilus oryzae*, have been purified to near homogeneity (Shen et al., 1996; Shen et al., 1999), marking, in each case, the first purification and characterization of the enzyme from an animal source. A cDNA clone for the rice weevil polygalacturonase was isolated (Shen et al., 2003). Studies based on that cDNA clone strongly indicated that polygalacturonase is encoded by the rice weevil's genome, rather than a genome of a symbiotic microorganism, suggesting that there may have been a horizontal transfer of the gene from a fungus to an ancestral weevil.

However, the previous work on pectinmethylesterase left open the question the genetic origin of the enzyme. The gene for this enzyme could reside either in the genome of the insect itself or in that of a symbiotic microorganism. The latter possibility was raised by Campbell (1989) for *Sitophilus* pectinases. This was a plausible suggestion, given the scarcity of polygalacturonases and pectinmethylesterases in the animal kingdom, the widespread occurrence of the enzymes in bacteria and fungi, and the harboring of endosymbionts by *Sitophilus* species (Heddi et al., 1998). The

nonredundant database at the National Center for Biotechnology Information now contains over 500 sequences of pectinmethylesterases from bacterial, fungal and plant sources that comprise a single pectinesterase family (Pfam, 2004), but no pectinmethylesterase of an animal origin. However, Eigenheer et al. (2003) have reported expressed sequence tags encoding a putative pectinmethylesterase from the midguts of the weevil *Ips pini*.

We here report the isolation and sequencing of a cDNA clone that encodes rice weevil pectinmethylesterase, the expression of the enzyme in cultured insect cells, and an experiment with drug-treated insects that strongly indicates that the encoding gene resides in the insect's genome. Thus, rice weevil pectinmethylesterase, like its counterpart, polygalacturonase, appears to be a gut digestive enzyme encoded by the weevil's genome. It seems likely to have transferred there from a bacterium, based on sequence similarities to extant pectinmethylesterases.

Materials and Methods

Insects and chemicals

Rice weevils were reared on wheat kernels and maintained

Shen Z, Pappan K, Mutti NS, He Q-J, Denton M, Zhang Y, Kanost MR, Reese JC, Reeck GR. 2005. Pectinmethylesterase from the rice weevil, *Sitophilus oryzae*: cDNA isolation and sequencing, genetic origin, and expression of the recombinant enzyme. 9pp. *Journal of Insect Science*, 5:21, Available online: insectscience.org/5.21

2

at room temperature. Pectin (P 2157) was obtained from Sigma (St. Louis, MO). All other chemicals were analytical grade.

N-terminal amino acid sequencing

From a sample of purified pectinmethylesterase from the rice weevil (Shen et al., 1999) the N-terminal sequence ?QTAPGTAS was obtained by automated Edman degradation using the Applied Biosystems 473 Protein Sequencing System at the Kansas State University Biotechnology Core Facility. As indicated by the question mark, the N–terminal residue was ambiguous, as is commonly the case for the first cycle in automated Edman degradation.

Polymerase Chain Reaction

Genomic DNA was isolated from 2 g of adult rice weevils according to Campbell et al. (1992) and was used as template in PCR. Three degenerate primers (PE1, PE2, and PE4) were synthesized based on the N-terminal sequence of the purified rice weevil pectinmethylesterase sequence (PE1 and PE2) or on a highly conserved stretch of amino acid residues, GTQDTL (PE4), found in bacterial, fungal, and plant pectinmethylesterases. The sequences of the primers were:

PE1 - 5' AA(A/G)CA(A/G)CA(A/G)ACIGCICCNGG 3'
PE2 - 5' AA(A/G)CA(A/G)ACIGCICCIGGIACNGC 3'
PE4 - 5' AG(C/G)GTGTC(C/T)TG(G/A)TANCC 3'

where I and N denote deoxyinosine and total degeneracy, respectively. PCR was carried out as previously described (Shen et al., 2003) using 200 pmol of either PE1 or PE2, and 200 pmol of PE4.

Cloning of PCR products and isolation of full-length clone from cDNA library

DNA was purified from the single intense gel band from PCR and cloned in pGEM-T as previously described (Shen et al., 2003). Ligated plasmids were transformed into competent *Escherichia coli* strain XL1 blue cells and the bacteria spread on LB agar plates containing 100 μg/ml ampicillin, 0.5 mM IPTG, and 80 μg/ml X-gal. Plasmid DNA isolated from several white colonies was purified and digested with *Pst I* and *Apa I*. The digested DNA products were separated by agarose gel electrophoresis and inserts were sequenced. One clone (PE24), of length 609 bp, showed strong sequence similarity to known pectinmethylesterase sequences and was selected to construct a probe for screening of a whole-body rice weevil cDNA library (see Shen et al., 2003). From that screening, a clone, PE6, with a 1.6 kb insert, was selected. Sequencing of the insert revealed an open reading frame that appeared to encode the complete rice weevil pectinmethylesterase. This clone (accession number AY841894) was used in the experiments described here.

Construction of phylogenetic tree

The amino acid sequence encoded by the longest open reading frame of clone PE6 (i.e., the rice weevil pectinmethylesterase sequence) and 18 other pectinmethylesterase amino acid sequences were aligned with ClustalW version 1.82 (www.ebi.ac.uk/clustalw). This alignment was used for tree building, using the parsimony analysis in PAUP 4.0 b10 (Swafford, 2000), with unweighted characters and including gaps, and bootstrapping (1000 "fast"

replications).

Genomic DNA isolation, Southern blotting, analysis of several genes in tetracycline-treated rice weevils, and in situ hybridization

These procedures were conducted as described in Shen et al. (2003). The pectinmethylesterase PCR primers in the experiment on tetracycline-treated weevils were: forward, 5' TCA AAT CCA CAA TGC CAG AC 3'; and, reverse, 5' CGC TTC TAC GAT CAG CTT TTA C 3'. The expected (and observed) product size was 185 bp.

Expression of recombinant pectinmethylesterase in cultured insect cells

A DNA fragment encoding the entire amino acid sequence of rice weevil pectinmethylesterase was generated by PCR using the PE6 insert, in pBLUESCRIPT, as a template and two primers, PMESFN and PMESFC. PMESFN contained a Bam H1 restriction site and had the following sequence:

5' CATGGATCCAAACAATGAAAAATAATCG 3'.

PMESFC contained an EcoRI restriction site and had the following sequence:

5' TTGGAATTCTAAAATTAGGCACC 3'.

The PCR product was recovered from a low melt agarose gel using a Wizard Miniprep Kit (Promega, www.promega.com), digested with BamH1 and EcoRi, then ligated to Bam H1- and EcoRI-digested pVL1393 (Pharmingen, www.pharmingen.com). The ligated construct, pVL1393-PMEASE, was transformed into *E. coli* strain XL1-Blue. Clones containing the pectinmethylesterase insert were confirmed by restriction digests and DNA sequencing.

Three micrograms of pVL1391-PMEASE plasmid were co-transfected with 0.5 μg of BaculoGold DNA into 2 x 10⁶ *Spodoptera frudiperda* SF21 cells (Invitrogen, www.invitrogen.com) to produce a recombinant baculovirus for pectinmethylesterase expression, following the instructions for the BaculoGold Baculovirus Expression System (Pharmingen). For expression, Sf21 cells (8.5 x 10⁶) were cultured in 15 ml of SF-900 II serum-free medium (Invitrogen) and infected with the recombinant baculovirus at a multiplicity of infection of 5. After 5 days of incubation at 27° C the cell medium was collected and centrifuged at 10,000 x g for 20 min.

From the supernatant, pectinmethylesterase was purified by dialysis against 20 mM Tris-Cl (pH 8.7), and then successive chromatography on Q-Sepharose at pH 8.7, where the vast majority of the activity was in the flow through, and S-Sepharose at pH 4.6, where the pectinmethylesterase activity corresponded with the single peak of bound, and subsequently eluted, material (He, 1998). Purified pectinmethylesterase (14μg/ml) was obtained from the culture medium. The purified protein was subjected to N-terminal sequencing as described above.

Results

Sequence of the pectinmethylesterase cDNA clone and recombinant expression of the enzyme

The nucleotide sequence of cDNA clone PE6 (Fig. 1), was identified on the basis of sequence similarity to pectinmethylesterases from other organisms. The N-terminal sequence of the protein was predicted by PSORT to be a signal peptide for an extracellular protein
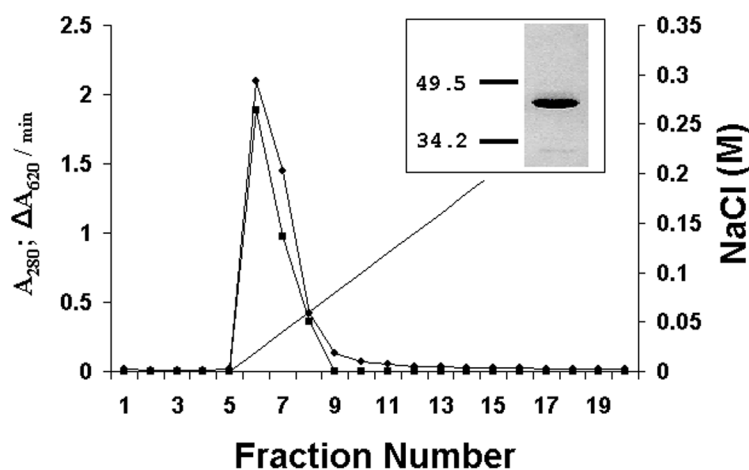
```
                                    CGGCACGAGAGAGATATTAAATCATAGATATAAACA   34
          ATGAAAATAATCGTATTATTACTACTAGCCGTAGTATTAGCTTCGGCGGATCAAACAGCA   94
           M  K  I  I  V  L  L  L  L  A  V  V  L  A  S  A ↑ D  Q  T  A   20
          CCTGGAACAGCATCTCGCCCAATCTTGACAGCATCAGAATCGAATTATTTCACTACCGCA  154
            P  G  T  A  S  R  P  I  L  T  A  S  E  S  N  Y  F  T  T  A   40
          ACTTATCTTCAAGGATGGTCGCCACCTTCAATTTCAACGAGCAAAGCTGATTATACTGTT  214
            T  Y  L  Q  G  W  S  P  P  S  I  S  T  K  A  D  Y  T  V      60
          GGTAATGGCTACAATACAATACAAGCCGCTGTAAATGCTGCTATTAATGCAGGAGGCACC  274
            G  N  G  Y  N  T  I  Q  A  A  V  N  A  A  I  N  A  G  G  T   80
          ACAAGAAAATATATTAAAATTAATGCTGGAACATATCAAGAAGTTGTTTATATTCCCAAC  334
            T  R  K  Y  I  K  I  N  A  G  T  Y  Q  E  V  V  Y  I  P  N  100
          ACAAAGGTGCCTCTTACGATTTATGGTGGAGGCAGTAGCCCATCAGACACACTGATAACA  394
            T  K  V  P  L  T  I  Y  G  G  G  S  S  P  S  D  T  L  I  T  120
          TTGAACATGCCCGCTCAAACTACTCCATCAGCTTATAAAAGTCTTGTTGGGTCACTTTTT  454
            L  N  M  P  A  Q  T  T  P  S  A  Y  K  S  L  V  G  S  L  F  140
          AATTCAGCGGACCCTGCCTATAGTATGTATAATAGCTGGCGATCTAAAAGTGGCGCTATT  514
            N  S  A  D  P  A  Y  S  M  Y  N  S  W  R  S  K  S  G  A  I  160
          GGAACTAGTTGCTCTACAGTGTTTTGGGGTAAAGCACCAGCTGTTCAAATCGTAAATCTC  574
            G  T  S  C  S  T  V  F  W  G  K  A  P  A  V  Q  I  V  N  L  180
          TCTATTGAAAATTCGGCTAAAAATACCGGAGATCAGCAAGCCGTCGCTCTACAAACTAAC  634
            S  I  E  N  S  A  K  N  T  G  D  Q  Q  A  V  A  L  Q  T  N  200
          AGTGACCAAATTCAAATCCACAATGCCAGACTGTTGGGTTATCAAGACACACTGTATGCT  694
            S  D  Q  I  Q  I  H  N  A  R  L  L  G  Y  Q  D  T  L  Y  A  220
          GGGTCAGGAAGTTCCAGCGTCGAAAGGAGTTACTATACCAATACCTATATTGAAGGAGAT  754
            G  S  G  S  S  S  V  E  R  S  Y  Y  T  N  T  Y  I  E  G  D  240
          ATCGATTTTGTTTTCGGTGGTGGATCTGCTATATTTGAAAGTTGCACTTTCTATGTAAAA  814
            I  D  F  V  F  G  G  G  S  A  I  F  E  S  C  T  F  Y  V  K  260
          GCTGATCGTAGAAGCGACACTTCCGTAGTTTTTGCACCTGATACCGATCCTCATAAGATG  874
            A  D  R  R  S  D  T  S  V  V  F  A  P  D  T  D  P  H  K  M  280
          TACGGATATTTCGTTTATAAATCTACAATTACAGGAGATAGTGCCTGGTCATCCTCAAAG  934
            Y  G  Y  F  V  Y  K  S  T  I  T  G  D  S  A  W  S  S  S  K  300
          AAAGCATACCTAGGACGTGCTTGGGATTCTGCCGTATCCTCGTCAAGTGCATACGTTCCA  994
            K  A  Y  L  G  R  A  W  D  S  A  V  S  S  S  S  A  Y  V  P  320
          GGAACTTCACCTAATGGTCAGCTAATTATAAAGGAGAGCACCATTGATGGAATAATTAAT 1054
            G  T  S  P  N  G  Q  L  I  I  K  E  S  T  I  D  G  I  I  N  340
          ACTTCTGGGCCTTGGACAACAGCAACTAGCGGCAGAACATATTCTGGAAATAATGCCAAC 1114
            T  S  G  P  W  T  T  A  T  S  G  R  T  Y  S  G  N  N  A  N  360
          AGTAGAGACCTTAATAATGACAACTACAATAGATTTTGGGAATATAACAATTCTGGAAAT 1174
            S  R  D  L  N  N  D  N  Y  N  R  F  W  E  Y  N  N  S  G  N  380
          GGTGCCTAATTTTATAATGACAATTTATAGTAATAAATATTGAACCTAANAAAAAAAAA  1234
            G  A  ★     382
          AA            1236
```

Figure 1. Nucleotide sequence of pectinmethylesterase cDNA clone PE6 from *Sitophilus oryzae*, with the translated sequence of the protein. The arrow indicates a signal-peptide cleavage site predicted by PSORT (psort.nibb.ac.jp; von Heijne 1986). The underlined portion of the underlined amino acid sequence corresponds to sequence determined on the enzyme purified from rice weevils (Shen et al. 1996; Methods of this paper).

(http://psort.nibb.ac.jp), with cleavage predicted between residues 16 and 17. The underlined stretch of amino acid residues corresponds to sequence obtained by N-terminal sequencing of the purified pectinmethylesterase itself. The mass of the predicted mature protein (39.2 kD) corresponded well with the mass of the purified enzyme, as determined by SDS gel electrophoresis, 38 kDa (Shen et al. 1999). Thus, it appears that a cDNA clone that encodes the entirety of rice weevil pectinmethylesterase was obtained. The poly(A) tail in the cDNA is a feature that has relevance to the genetic origin of the enzyme, as discussed below.

Shen Z, Pappan K, Mutti NS, He Q-J, Denton M, Zhang Y, Kanost MR, Reese JC, Reeck GR. 2005. Pectinmethylesterase from the rice weevil, 4
Sitophilus oryzae: cDNA isolation and sequencing, genetic origin, and expression of the recombinant enzyme. 9pp. *Journal of Insect Science*, 5:21,
Available online: insectscience.org/5.21

Figure 2. Chromatography of recombinant pectinmethylesterase from *Sitophilus oryzae* on S-Sepharose at pH 4.6. The sample was the flow-through from a Q-Sepharose chromatography (see Methods). A_{280} is indicated by circles and ΔA_{620}/min (a measure of pectinmethylesterase activity) by squares. The insert shows SDS polyacrylamide gel electrophoresis of fractions 6 – 8, combined, with the masses (in kDa) of two standards, carbonic anhydrase and ovalbumin. The estimated mass of this band was 40 kDa.

Expression of the recombinant enzyme in cultured insect cells was used to directly confirm that clone PE6 encodes pectinmethylesterase. Figure 2 shows elution of the enzyme from S-Sepharose and SDS gel electrophoresis of the purified protein. Sequencing of the purified protein revealed an N-terminal sequence of DQTAPGT, which corresponded to the predicted N-terminal sequence of the protein inferred from the nucleotide sequence of clone PE6 (Fig. 1).

Similarity in sequence to other pectinmethylesterases

The sequence of rice weevil pectinmethylesterase was aligned with pectinmethylesterase sequences from bacteria, plants and fungi (Figure 3). The highest level of sequence similarity is to bacterial pectinmethylesterases, a fact reflected in the phylogram in Figure 4. Among bacterial sequences, the highest level of similarity was toward the enzymes from enterobacteria — the *Erwinia* species, *E. coli* and *Salmonella typhimurium.*

In situ hybridization

The rice weevil gut was probed with digoxygenin-labeled pectinmethylesterase RNA and gave a strong signal throughout the midgut, including the ventriculus (Figure 5A). This is consistent with the notion that the pectinmethylesterase, like polygalacturonase (Shen et al. 2003), is a midgut digestive enzyme in the rice weevil. Of course, the current data do not exclude the possibility that the pectinmethylesterase gene is expressed in other tissues.

Heddi et al. (1998) reported that SOPE (the principal symbiont of *S. oryzae*) is absent from the rice weevil intestine and is instead present only in bacteriomes in specialized organs. A second rice weevil endosymbiont, *Wolbachia*, is present in the gut, but in the only *Wolbachia* genome completely sequenced, a symbiont of *Drosophila melanogaster* (accession number NC_002978,

www.ncbi.nlm.nih.gov), there is no gene with sequence similarity to rice weevil pectinmethylesterase. In conjunction with these observations, finding the rice weevil pectinmethylesterase transcript in the weevil midgut suggests that this enzyme is encoded by the insect's genome and not that of either rice weevil endosymbiont.

Studies on antibiotic-treated insects

Adult rice weevils were reared for several weeks on wheat flour to which we added tetracycline (0.5 mg/g of flour). This procedure is known to rid the rice weevil of its two bacterial symbionts, SOPE (the principal endosymbiont) and *Wolbachia* (Heddi et al. 1998). At several times after exposure to the antibiotic, we used PCR to evaluate the presence of genes for the rice weevil (18S-rRNA gene), for *Wolbachia* and SOPE (16S-rRNA genes) and for the pectinmethylesterase gene. The symbiont genes disappeared during the tetracycline treatment (first *Wobachia* and, later, SOPE) whereas the 18S-rRNA gene of the rice weevil and the pectinmethylesterase gene were unaffected (Figure 6).

Northern and Southern blot analyses

Northern blot analysis of total rice weevil RNA using a pectinmethylesterase probe revealed a single band of 1600 bases (result not shown). In Southern analysis we observed simple banding patterns (Figure 7). The cDNA PE6 contains sites only for PuvII (position 554) and EcoRV (position 754) among the enzymes we used. Although definitive conclusions cannot be drawn, the results suggest a single gene (based especially on the HindIII and PvuII digestions), possibly with two alleles (based on the XbaI and EcoRV digestions) or possibly with XbaI, EcoRV and EcoRI sites in introns.

Discussion

Pectinmethylesterase, which has been extensively studied in bacterial, fungal and plant systems, is one part of a coordinated enzymatic attack on pectin. Removal of methyl groups from pectin by pectinmethylesterase allows polygalacturonase to hydrolyze glycosidic linkages in the de-esterified substrate, pectate (Bateman and Bashman 1976).

At this point, the function of these pectinases in the rice weevil is not known. Their occurrence in the gut suggests that their role is digestive. This could either be for direct uptake of nutrients (galacturonides) from the digestion of pectin, or it could be aimed at the disruption of pieces of plant tissues, as pectin "cements" plant cell walls (Bateman and Bashman 1976), thus allowing access of other digestive enzymes to the contents of the plant cells.

Campbell (1989) reported rather high pectinase activities in *Sitophilus* species which allowed the purification of the enzymes, although this required roughly 3000-fold purification in each case (Shen et al. 1996; Shen et al. 1999). Campbell (1989) suggested that the enzymes might be encoded by a symbiotic microorganism, namely by the bacterial endosymbiont of *Sitophilus*. This was a reasonable suggestion, since pectinases are found in many bacteria but, among animals, only in a few insect species (Adams and McAllan, 1956; Laurema and Nuorteva, 1961; Campbell and Dryer 1985). But our studies on a cDNA that encodes rice weevil pectinmethylesterase point to its source as the insect genome itself,

```
Arabidopsis thaliana      KYKTVQEAVDSAPDT-NRTVKFVIRIREGVYEETVRVPFEKKNVVFIGDG
Lycopersicon esculentum   KYRTLAEAVAAAPDK-SKT-RYVIYVKRGTYKENVEVSSRKMNLMIIGDG
Aspergillus oryzae        DYDTLSAAVN-ALST-TETATQTIFIEEGTYDEQVYIPSLAGKLIIYGQT
Cochliobolus carbonum     KYKTVQSAVN-ALDK-STSKAQCIFIAKGTYTEQVYIQQLNSPLTIYGET
Erwinia carotovora        TYANIQQAVNAALQAQQGRAHIDIKVLPGSYIGTVYLPADSPPLTLFGAG
Escherichia coli          THTTIQAAVDAAIIKRTNKRQY-IAVMPGEYQGTVYVPAAPGGITLYGTG
Rice weevil               GYNTIQAAVNAAINAGGTTRKY-IKINAGTYQEVVYIPNTKVPLTIYGGG
                           : .:   **  *          * :  *  *   * :     : : *


Arabidopsis thaliana      MGKTVITGSLNVGQP-------------------------------------
Lycopersicon esculentum   MYATIITGSLNVVD-------------------------------------
Aspergillus oryzae        EDDTTYTGNLVNITHNI-----------------------------KLAD
Cochliobolus carbonum     KDTRSYSGNTVTITQRK-----------------------------SQND
Erwinia carotovora        EQPDNVVIQLALDSMFSPATYQKTVNPHGEYQPGDPAWYMYDLCASKQNA
Escherichia coli          EKPIDVKIGLSLDGGMSPADWRHDVNPRGKYMPGKPAWYMYDSCQSKRSD
Rice weevil               SSPSDTLITLNMPAQTTPSAYKSLVG--SLFNSADPAYSMYNSWRSK-SG


Arabidopsis thaliana      GMTTFESATVGVLGDGFMARDLTIEN----TAGADAHQAVAFRSDSDFSV
Lycopersicon esculentum   GSTTFHSATLAAVGKGFILQDICIQN----TAGPAKHQAVALRVGADKSV
Aspergillus oryzae        VANDDETATVRNHSPNSSIYNLNIIN----TCGQVCHQALAVSPYGNGQG
Cochliobolus carbonum     TPSNDETATLRAHTPNLKVYNINLVN----TRGQGS-QALAVSAWADKQG
Erwinia carotovora        TIDTLCAAVVWSQSDNFQMKNLTVVNSLLDSVDSRAHQAVALRTDGDKVQ
Escherichia coli          SIGVLCSAVFWSQNNGLQLQNLTIENTLGDSVDAGNHPAVALRTDGDQVQ
Rice weevil               AIGTSCSTVFWGKAPAVQIVNLSIENSAKN---TGDQQAVALQTNSDQIQ
                           ::..          :: : *            *:*.  .:


Arabidopsis thaliana      LENCEFLGNQDTLYAHSL------------RQFYKQCRIQGNVDFIFGN-
Lycopersicon esculentum   INRCRIDAYQDTLYAHSQ------------RQFYRDSYVTGTIDFIFGN-
Aspergillus oryzae        YYGCQFTGYQDTLLAQSG------------NQVYAHNLIEGAVDFIFGQH
Cochliobolus carbonum     YYGCSFKGYQDTVLAQLG------------NQVYAKSYIEGATDFIFGQQ
Erwinia carotovora        LERVRLIGRQDTFFVNTSNLHNEYVTDRYSRVYIKDSYIEGDVDYVFGR-
Escherichia coli          INNVNILGRQNTFFVTNSGVQNRLETNRQPRTLVTNSYIEGDVDIVSGR-
Rice weevil               IHNARLLGYQDTLYAGSG-------SSSVERSYYTNTYIEGDIDFVFGG-
                           : . *:*.  .           .   . : *  * : *


Arabidopsis thaliana      SAAVFQDCDILIASKHSKLEQGGANNAITAHGRIDASQSTGFVFLNCSIN
Lycopersicon esculentum   AAVVFQKCQ-LVARKPGKYQQ----NMVTAQGRTDPNQATGTSIQFCDII
Aspergillus_oryzae        ARAWFQDCDIRVLKGPS-------SGYITANGRSSETDTSYYVIHKSSVA
Cochliobolus carbonum     ARAWFDGCDIGVL-GAQ-------VGWITASGRAADSNPSYYVFNKSKVS
Erwinia carotovora        ATAVFNRVHFHTVSSRGA------KDIHVFAPDSMPWVQYGFLAVKCRFT
Escherichia coli          GAVVFDNTEFRVVNSRTQ------QEAYVFAPATLSNIYYGFLAVNSRFN
Rice weevil               GSAIFESCTFYVKADRRS------DTSVVFAPDTDPHKMYGYFVYKSTIT
                           . . *:             .           . .


Arabidopsis thaliana      GTEEYMKEFQANPEGHKNFLGRPWKEFSRTVFVNCNLESLISPDGWMPWN
Lycopersicon esculentum   ASPDLKPVVKEFP----TYLGRPWKKYSRTVVMESYLGGLIDPSGWAEWH
Aspergillus oryzae        AADGNDVPSGTYY------LGRPWSQYARVLYQETSMTDVINSAGWSVWS
Cochliobolus carbonum     PAVGATVSSGVYY------LGRPWGNYARVVFQHTDLSNVINPQGWAAWN
Erwinia carotovora        GDEGFSRERKAKL-------GRAWDQGARQT----GYQPGLTANGQLVIR
Escherichia coli          ---AFG-DGVAQL-------GRSLDVDAN-------------TNGQVVIR
Rice weevil               GDSAWSSSKKAYL-------GRAWDSAVSSSS---AYVPGTSPNGQLIIK
                                            **.                      . *


Arabidopsis thaliana      G-DFALKTLYYGEYKNTGPGSVRSSRVPWS--SEIPEKHVDV-YSVANFI
Lycopersicon esculentum   G-DFALKTLYYGEFMNNGPGAGTSKRVKWPGYHVITDPAEAMSFTVAKLI
Aspergillus oryzae        TTQANTENVTFAEYGNTGAGAE-GTRASFS-------EKLSEPVAISTIL
Cochliobolus carbonum     TGDARTDHVLFGEYQNTGAGAS-GKR-SFG-------QTLKAALSITDIL
Erwinia carotovora        DSTIDASYDREQPWGAAATTARPFVGNTDSA-------RNLDDMHFNRLW
Escherichia coli          DSAINEGFNTAKPWADAVISNRPFAGNTGSVDDNDEIQRNLNDTNYNRMW
Rice weevil               ESTIDGIINTSGPWTTATSG-RTYSGNNANS-------RDLNNDNYNRFW
                           :                                              :


Arabidopsis thaliana      QADE
Lycopersicon esculentum   QGGS
Aspergillus oryzae        GSD-
Cochliobolus carbonum     GSN-
Erwinia carotovora        EYNN
Escherichia coli          EYNN
Rice weevil               EYNN
```

Figure 3. Alignment of pectinmethylesterase amino acid sequence from *Sitophilus oryzae* with representative pectinmethylesterase sequences from bacteria, fungi and plants. Sequences (and accession numbers) are: rice weevil, this paper; *Erwinia carotovora* YP_048235; *Escherichia coli* NP_415293; *Cochliobolus carbonum* ADD43340; *Aspergillus oryzae* BAA75474; *Lycopersicon esculentum* CAA52704; *Arabidopsis thaliana* AAM91523. The sequences were aligned using ClustalW version 1.82 (www.ebi.ac.uk/clustalw). Each aligned sequence was trimmed at the N-terminus and C-terminus, to show the region of highest similarity among the 7 sequences. The rice weevil sequence shown, for instance, begins at residue 47 of the mature protein and ends 5 residues short of the C-terminal residue. Asterisks indicate the presence of identical residues in all sequences; colons represent an intermediate level of similarity at a given position; and periods a lower but still marked level of similarity. (See the ClustalW website for definitions: http://www.ebi.ac.uk/clustalw.)
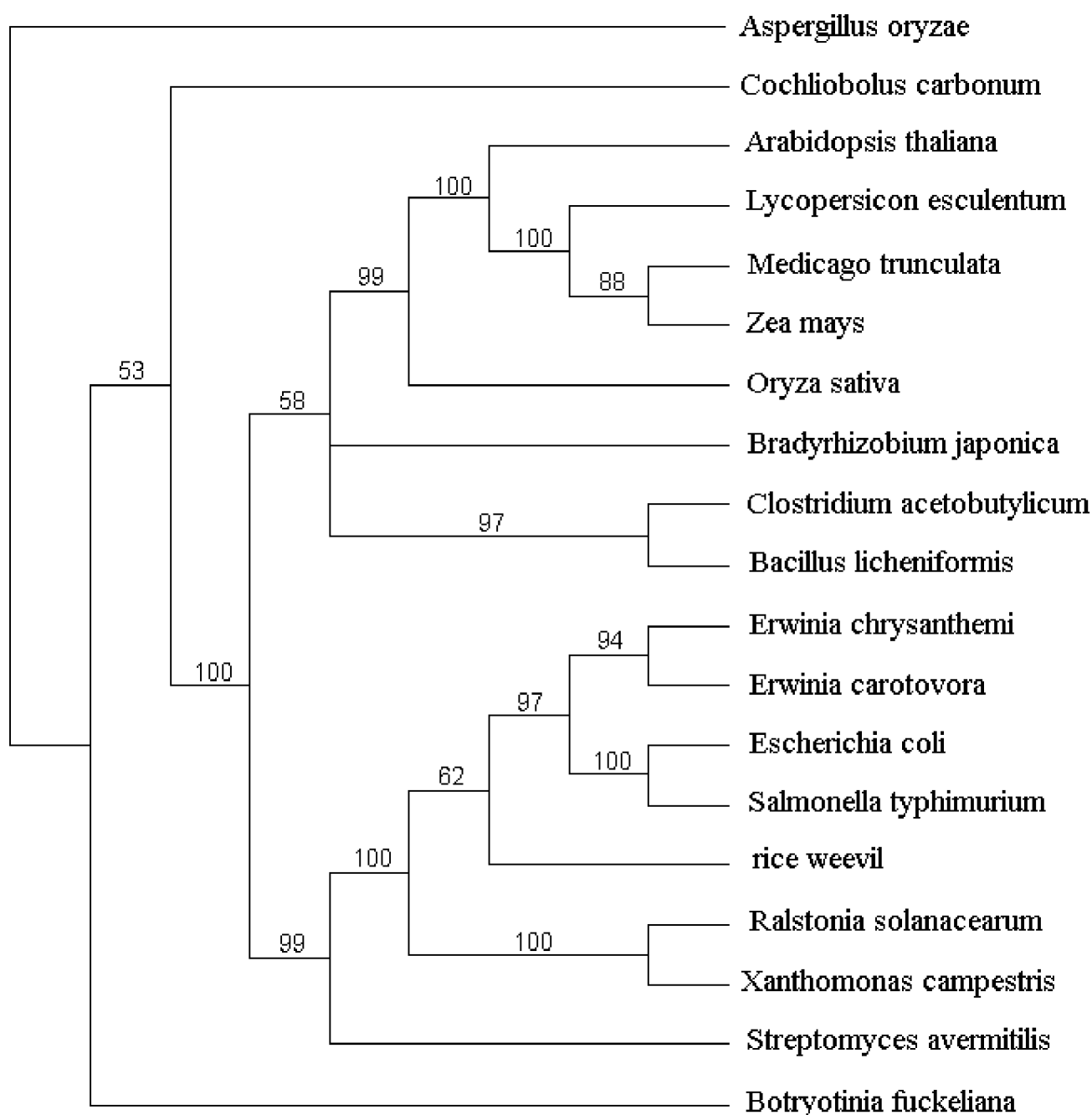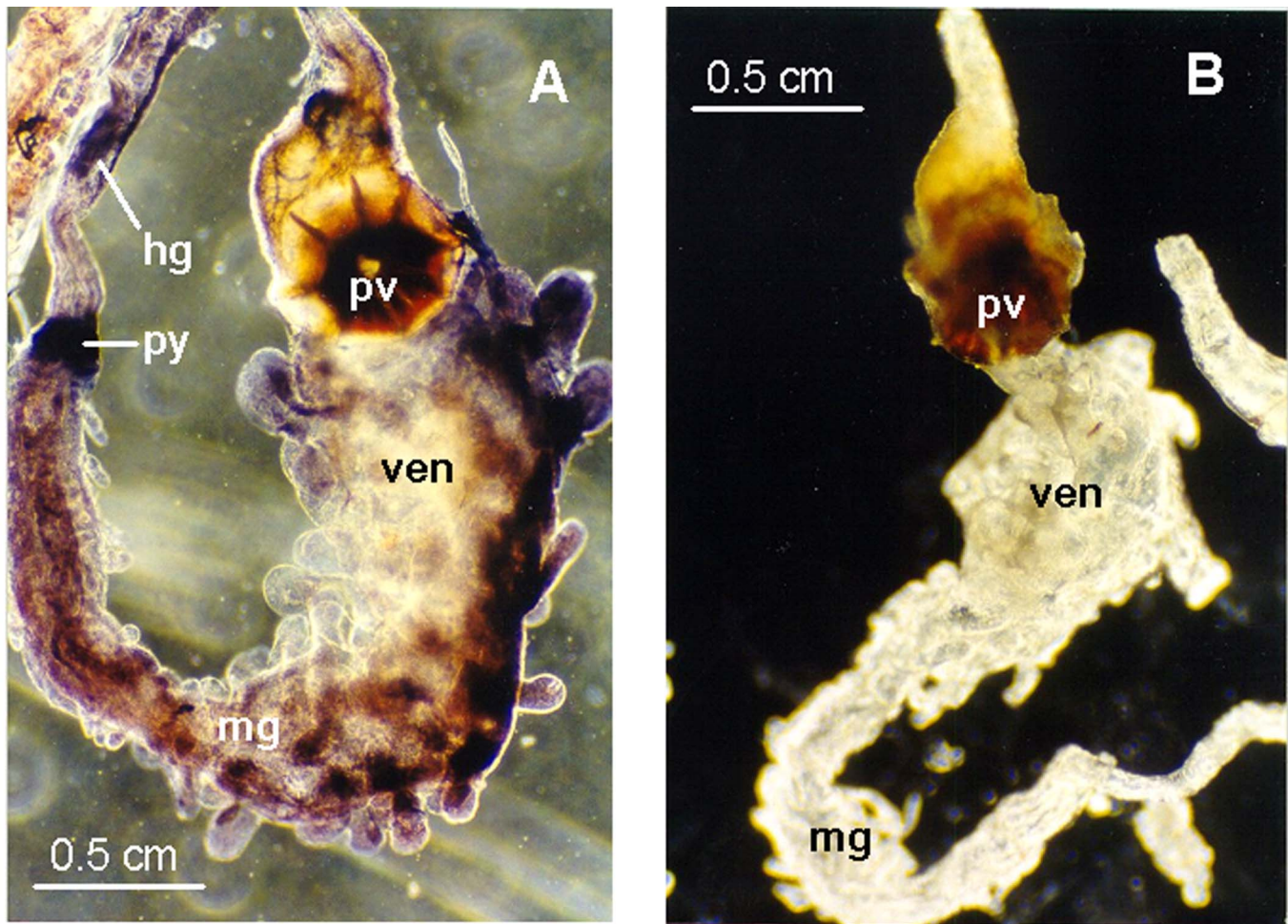
Figure 4. Phylogram of pectinmethylesterase sequences. The phylogram was created starting with a multiple alignment created with ClustalW version 1.82. The alignment included several sequences in addition to those shown in Figure 3. The additional sequences are: *Erwinia chrysanthemi* CAA59151; *Salmonella typhimurium* NP_459765; *Ralstonia solanacearum* P24791; *Xanthomonas campestris* NP_637620; *Streptomyces avermitilis* NP_827553; *Bradyrhizobium japonica* NP_768634; *Bacillus licheniformis* YP_093018; *Clostridium acetobutylicum* NP_349964; *Oryza sativa* NP_916048 ; *Medicago trunculata* CAB65290; *Zea mays* CAA73733; *Botryotinia fuckeliana* CAC29255. In creating the alignment on which this phylogram is based, we used the entire amino acid sequences of the proteins. The phylogram was generated as described in Methods. The consensus tree that is shown is rooted with the *Aspergillus* sequence. A number in the phylogram indicates the percentage of trees that included the nearest branch point (in 1000 trees constructed).

and not that of a symbiotic organism. The pieces of evidence in support of this conclusion are several, but perhaps the most convincing is shown in Figure 6, namely, the retention of the pectinmethylesterase gene during tetracycline treatment that rids *S. oryzae* of its two endysymbionts. Our studies on rice weevil polygalacturonase (Shen et al. 2003) and pectinmethylesterase (this

paper) reveal similarities between these enzymes, but there is an interesting difference in that rice weevil pectinmethylesterase is most similar in sequence to bacterial enzymes, whereas the rice weevil polygalacturonase sequence is most similar to fungal enzymes (Shen et al. 2003).

Pectinase activities have been reported in only a few insects

Shen Z, Pappan K, Mutti NS, He Q-J, Denton M, Zhang Y, Kanost MR, Reese JC, Reeck GR. 2005. Pectinmethylesterase from the rice weevil,
Sitophilus oryzae: cDNA isolation and sequencing, genetic origin, and expression of the recombinant enzyme. 9pp. *Journal of Insect Science*, 5:21,
Available online: insectscience.org/5.21

7

Figure 5. Detection of pectinmethylesterase mRNA in adult *Sitophilus oryzae* digestive tracts by *in situ* hybridization. Methods are given in Shen et al. (2003). The control (B) lacked exposure to the digoxygenin-labeled probe. A control lacking the alkaline-phosphatase-coupled antibody was also devoid of signal (not shown). The full complement of reagents (A) gave strong staining stained throughout the midgut (including the ventriculus, the enlarged anterior portion of the midgut) and weak staining in the anterior portion of the hindgut. Abbreviations: hg, hindgut; mg, midgut; pv, proventriculus; py, pylorus; ven, ventriculus.

(Adams and McAllan, 1956; Laurema and Nuorteva, 1961; Campbell and Dryer 1985; Ma et al. 1990; Shen et al. 1996, 1999; Doostdar et al. 1997). Genes that encode pectinmethylesterase that are recognizably similar in sequence to known pectinmethylesterases do not occur in the several insect species for which the entire genome has been sequenced, or in *C. elegans*, the *Wolbachia* symbiont of *D. melanogaster* or *Buchnera aphidocola*, an endosymbiont of aphids known to have pectinmethylesterase activity (Campbell and Dryer 1985; Ma et al. 1990). Thus, the most plausible explanation, whether in *Sitophilus* or in aphids, is that these pectinase genes reside in the insects' genomes as a result of an ancestral horizontal gene transfer. In the case of the rice weevil pectinmethylesterase, this transfer would appear to have been from a bacterium.

The study of pectinmethylesterases in weevils or other coleopterans is in its early stages. This enzyme and polygalacturonase is known to be present in comparable amounts in three *Sitophilus* species (Campbell, 1989; Reeck et al. unpublished results). We have been unable to find any trace of pectinmethylesterase activity or an encoding gene in either *Tribolium* or *Tenebrio*, two other phytophagous coleopteran species. Eigenheer et al. (2003) have reported ESTs encoding a putative pectinmethylesterase from the midgets of the weevil *Ips pini*. The longest of those ESTs encodes an amino acid sequence that aligns with residues 87 - 304 of mature rice weevil pectinmethylesterase, with identical amino acid residues at 63% of the positions and no gaps in the alignment. This strongly suggests that the occurrence of pectinmethylesterase extends beyond *Sitophilus* species among the weevils. It may be that the occurrence of both pectinases is largely limited to weevils, among coleopterans, but a better understanding of the distributions and functions of these enzymes awaits further investigation.

Acknowledgements
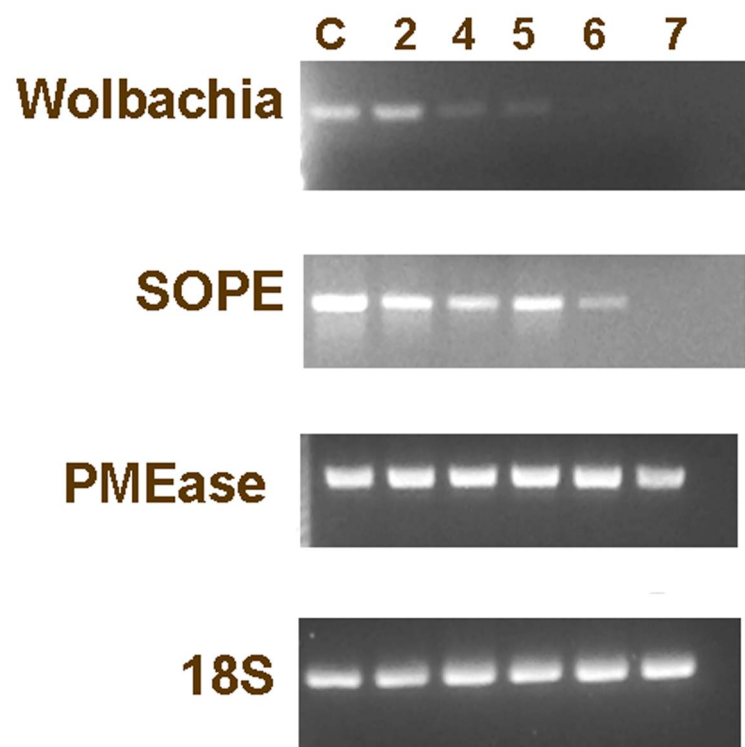
Figure 6. PCR analysis of several genes in tetracycline-treated *Sitophilus oryzae*. DNA prepared from rice weevils was analyzed after 2, 4, 5, 6 and 7 weeks of exposure to tetracycline (0.5 µg/g flour). The control (C) is analysis of DNA from the normal rearing of the weevils for 7 weeks in the absence of the drug. The number of cycles was constant for the PCR analysis of DNA prepared after the various times of rearing: 35 cycles for analysis of the *Wolbachia* 16S-RNA gene and 30 cycles for the other genes (*S. oryzae* 18S-RNA gene, the SOPE 16S-RNA gene, and the rice weevil pectinmethylesterase gene). Experimental details are given in Shen et al. (2003).
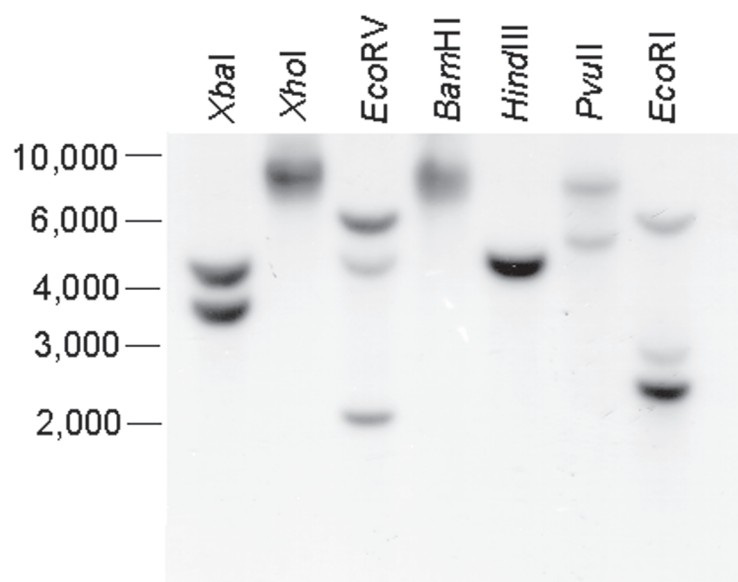
Figure 7. Southern-blot analysis of rice weevil pectinmethylesterase. DNA isolated from adult rice weevils was digested with various restriction endonucleases, as indicated, and probed with radioactively labeled pectinmethylesterase cDNA as described in Shen et al. (2003). The numbers on the left indicate positions of markers, in kb pairs.

State University Museum of Entomological and Prairie Arthropod Research, Manhattan, Kansas.

References

Adams JB, McAllan JV. 1956. Pectinase in the saliva of *Myzus persicae* (Sulz.) (Homoptera: Aphididae) *Canadia Journal of Zoology* 34: 541-543.

Bateman DF, Bashman HG. 1976. Degradation of plant cell walls and membranes by microbial enzymes. In *Encyclopedia of Plant Physiology New Series*, *Physiological Plant Pathology* (R. Heiteffus & P. H. Williams, eds) pp. 316 355. Springer Verlag, Berlin.

Campbell BC, Dreyer DL. 1985. Host-plant resistance of sorghum: differential hydrolysis of sorghum pectic substances by polysaccharases of greenbug biotypes (*Schizaphis graminum*, Homoptera: Aphididae). *Archives of Insect Biochemistry and Physiology* 2: 203-215.

Campbell BC. 1989. On the role of microbial symbiotes in herbivorous insects. In: *Insect-Plant Interactions* Bernays EA, editor. 1-44. Boca Raton: CRC Press.

Campbell BC, Bragg TS, Turner CE. 1992. Phylogeny of symbiotic bacteria of four weevil species (Coleoptera: Curculionedae) based on analysis of 16S ribosomal DNA. *Insect Biochemistry and Molecular Biology* 22: 415-421.

Doostdar H, McCollum TG, Mayer RT. 1997. Purification and characterization of an endo-polygalacturonase from the gut of West Indies sugarcane rootstalk borer weevil (*Diaprepes abbreviatus* L.) larvae. *Comparative Biochemistry and Physiology* 118B: 861-867.

Eigenheer AL, Keeling CI, Young S, Tittiger C. 2003. Comparison of gene representation in midguts from two phytophagous insects, *Bombyx mori* and *Ips pini*, using expressed sequence tags. *Gene*. 316: 127-36.

He QJ.1998. *Expression and purification of recombinant rice weevil,* Sitophilus oryzae *(L.) pectin methylesterase*. 91 pp. Master's thesis, Kansas State University, Manhattan, KS.

Heddi A, Grenier AM., Khatchadourian C, Charles H, Nardon P. 1998. Four intracellular genomes direct weevil biology: Nuclear, mitochondrial, principal endosymbiont, and *Wolbachia*. *Proceedings of the National Academy of Sciences (USA)* 96: 6814-6819.

Laurema S, Nuorteva P. 1961. On the occurrence of pectin polygalacturonase in the salivary glands of Heteroptera and Homoptera Auchenorrhyncha. *Annales Entomologici Fennici.* 27: 89-93.

Ma R, Reese JC, Black WC, Bramel-Cox P. 1990. Detection of pectinesterase and polygalacturonase from salivary secretions of living greenbugs*, Schizaphis graminum* (Homoptera: Aphididae*). Journal of Insect Physiology* 36: 507-512.

Pfam (Protein families database). 2004. Pectinesterase. http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF01095.

Shen Z, Denton M, Mutti N, Pappan K, Kanost MR, Reese JC, Reeck GR. 2003. Polygalacturonase from *Sitophilus oryzae*: Possible horizontal transfer of a pectinase gene from fungi to weevils. 9pp. *Journal of Insect Science*, 3:24, Available

Shen Z, Pappan K, Mutti NS, He Q-J, Denton M, Zhang Y, Kanost MR, Reese JC, Reeck GR. 2005. Pectinmethylesterase from the rice weevil, *Sitophilus oryzae*: cDNA isolation and sequencing, genetic origin, and expression of the recombinant enzyme. 9pp. *Journal of Insect Science*, 5:21, Available online: insectscience.org/5.21

online: insectscience.org/3.24

Shen ZC, Manning G, Reese JC, Reeck GR. 1999. Pectin methylesterase from the rice weevil, *Sitophilus oryzae* (L.) (Coleoptera: Curculionidae): Purification and characterization. *Insect Biochemistry and Molecular Biology* 29: 209-214.

Shen ZC, Reese JC, Reeck GR. 1996. Purification and characterization of polygalacturonase from the rice weevil, *Sitophilus oryzae* (Coleoptera: Curculionidae). *Insect Biochemistry and Molecular Biology* 26: 427-433.

Swafford DL. 2000. *Phylogenetic analysis with parsimony* (and other methods) Version 4.0 b10. Sinauer Associates, Sunderland, Massachusetts.

von Heijne G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research.* 14: 4683-4690.