Electronic Supplementary Information

EXPERIMENTAL DETAILS: MATERIALS AND METHODS

Plant Materials and Enriched Tissue Fractions

Two maize genotypes available at the International Maize and Wheat Improvement Center (CIMMYT) were used in the current study. P84 c2 (population derived from Caribbean accessions) was used as resistant check to the maize weevil, *Sitophilus zeamais*, while CML290 (tropical maize line) was considered as susceptible check to the MW based on our previous reports.[6-7] Insect bioassay was performed according with García-Lara et al.[6] For both genotypes, grain tissue samples were prepared from three separate replicates generated from 20 kernel samples for each genotype. Studies were conducted using individual tissues (embryo, endosperm and pericarp). Each tissue was removed manually after imbibing seeds in distilled water for 20 min at 4°C. The samples were dried and ground to a fine powder using a cyclone mill fitted with a 1-mm screen and stored at –20°C until extraction.

Peroxidase Extraction and In-Solution Assay.

The peroxidase extraction procedure was modified from Bestwick et al.[6] Ground samples of grain tissue (0.10 g) were hydrated overnight (4°C) with 0.8 mL of extraction buffer containing 100 mM sodium phosphate, pH 6.8. The sample was homogenized for 1 min and then centrifuged at 12,000 g for 30 min at 4°C. Protein concentrations were determined by the Bradford method.[15] Peroxidase activity was measured with guaiacol (2-methoxyphenol, Sigma Aldrich) at 470 nm using a UV–visible light spectrophotometer (BioTek Instruments, Inc., model Synergy™ 2, Vermont, USA). The assay mixture

contained 880 μL of 10 mM guaiacol in 50 mM potassium phosphate buffer, pH 6.6; 110-120 μL of extracted supernatant; 80 to 90 μL of sterile water; and 20 μL of 3% H_2O_2 to initiate the assay. Specific activity was expressed in units per milligram (U/mg) of protein content. At least three biological replicates were tested in this study.

Native Gel and One-Dimensional Polyacrylamide Gel Electrophoresis (1D-PAGE)

Native gel and SDS-PAGE for soluble peroxidases was performed according to the protocol described previously.[16] However, in order to avoid denatured of active peroxidases, the SDS-PAGE was slightly modified: 2x sample buffer was prepared without adding reducing agent, and the sample was not boiled before running the gel. Other studies had demonstrated the preservation of catalytic activity for hemoproteins under similar SDS-PAGE conditions.[8] For POD isoform separation SDS-PAGE was performed in 0.75 mm thick 9% acrylamide gels in a Mini-Protean II electrophoresis cell (Bio-Rad, Hercules, CA). The loaded sample (10 μL) was adjusted for the same amount of protein (20 μg). The molecular weight of the peroxidase was estimated by using protein standards (Bio Rad Lab, Inc. Cat.161-0374EDU).

Sequential Gel Staining and Image Analysis

Native and polyacrylamide gel was incubated for 15 min at 25°C in a potassium phosphate buffer (50 mM, pH 6.6) containing 10 mM guaiacol and 3% H_2O_2 immediately after electrophoresis. An orange-brown staining band appeared at the gel sites where POD activity was present. The image was captured using an ImageScanner III (GE) with the software LabScan 6.0. Subsequently, for SDS-PAGE, the gel was stained for protein with Coomassie Blue R250 (0.1% Coomassie Blue R250 in 40% methanol/10% acetic acid for 2 h and destained until clear with 40% methanol/10% acetic acid) and scanned again.

Protein Digestion for MS Analysis

Gel bands of interest were manually excised and digested with a protocol previously described.[17-18] Briefly, gel pieces were washed and destained with water, NH_4HCO_3 and acetonitril. Following, the gel pieces were reduced with DL-dithiothreitol and alkylated with iodoacetamide. The proteins were digested overnight at 37°C with sequencing grade trypsin (Promega. Madison, WI). The resulting peptides were extracted 1h with a 1:1 mixture of acetonitril/0.1% trifluoroacetic acid and dried in a vacuum centrifuge. Prior to LC-MS/MS the peptides were re-dissolved in 0.1% formic acid and transferred to FAMOS vials.

NanoLC-MS/MS Analysis

For the nanoLC-MS/MS analysis of the tryptic digests an Agilent 1100 series LC system with a nano pump for gradient formation and a capillary pump for sample loading was employed. The chromatography was carried out with an Agilent HPLC-ChipCube, allowing for direct infusion of the column outlet into the mass spectrometer electrospray ionization source. The mass spectrometry device was an Agilent LC/MSD Trap XCT Ultra. An Agilent chip G4240-62001 with a 43mmx75µm separation column and a 4mm 40nL enrichment column, both filled with Zorbax 300SB-C18 5µm resin was used for separation of the peptides. Gradient buffer A was 0.1% formic acid, gradient buffer B 98% acetonitril/0.1% formic acid. The capillary loading pump was washing with 60% methanol. The enrichment and analysis column were equilibrated with 3% buffer B, and then 8 µL of tryptic digest was loaded on the enrichment column and washed 5min with 3% B. After switching to the analysis column, B was increased to 45% during the next 27min. After this elution step, the column was cleaned for 3min with 95% B and re-equilibrated with 3% B

for 6.5min before the next sample. Flow rate was 0.4µL/min. For data accumulation a smart target ion count of 200,000 was set. The scan mode was extended, positive, *m/z* range 100-1400. Up to 5 peptides per mass spectrum were automatically chosen for fragmentation, with preference for > doubly charged precursors and exclusion of singly charged ones.

Generation of Custom Protein Sequence Database

A custom protein sequence database in FASTA format was constructed to ensure a high coverage of relevant *Zea mays*/POD sequences for the X!Tandem searches. The first component was SwissProt protein database (www.ebi.ac.uk/uniprot), containing proteins of many different species. Secondly, the protein sequences associated with the taxa *Zea mays* were exported from the NCBI database (www.ncbi.nlm.nih.gov) and added to the custom database. Finally, the PeroxiBase (www.peroxibase.isb-sib.ch), a specialized peroxidase protein database, maintained at the University of Geneva, was included.[12,19] To allow the estimation of a false discovery rate, a concatenated target-decoy data base was built according to the suggestions of Elias and Gygi[25]. This was done by reversing all sequences in the data base (FastaTools 0.9) and merging those to the original entries in the data base using a Perl script (mergeFasta.pl).

Protein Identification from nanoLC-MS/MS Data

Protein identification was carried out on a local installation of the Trans-Proteomic Pipeline, developed at the Seattle Proteome Center (www.tools.proteomeventer.org). [20,21] The initial assignment of LC-MS/MS mzXML spectra to peptides was based on a shared peak count approach (22) and performed by the X! Tandem algorithm from the Global Proteome Machine Organization(23, 24). An asymmetric monoisotopic mass error of -2 to 4 m/z was allowed for the precursor, a tolerance of 0.4 Da for the fragments. Atomic or

molecular masses were calculated manually with the molecular weight calculator adopted from Wieser.[25] X! Tandem search results were validated by PeptideProphet[26] and ProteinProphet.[27] Protein identifications with a ProteinProphet probability score > 0.9, the relevant LC-MS/MS spectra and meta data were submitted to the PRoteomics IDEntifications (PRIDE) database[29] (www.ebi.ac.uk/pride). Conversion of raw and identification data to PRIDE XML was done using the PRIDE converter[30], v2.3.2 1. For safe identification, additionally two unique peptides are demanded. No decoy hit meets this requirement.

LC-MS/MS raw data and identification results can be accessed through the PRIDE data base (www.ebi.ac.uk/pride), with a reviewer login:

Username: review55220

Password: d4.+2Y+Z

After acceptance of the manuscript, the data will be made publicly available with the accession numbers ***12910-14***. The current name of the project is POD_MAIZE

Bioinformatics

Multiple sequence alignments were performed by T-COFFEE[31] Version_5.05, hosted at the EBI server (www.ebi.ac.uk). The molecular weight and isoelectric point of an amino acid sequence was calculated by the ExPASy ProtParam[32] tool (www.expasy.org). InterProScan[33] (EBI, www.ebi.ac.uk) is a meta-search engine for different repositories and programs and allows a thorough scan for protein domain families and functional sites. For analyzing the sequences of the identified putative peroxidases, InterProScan was run with all available options.

ADDITIONAL REFERENCES

15 M.M. Bradford, *Anal. Biochem.*, 1976, **72,** 248.

16 J. Sambrook, D. Russell, and J. Sambrook, J. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press: Woodbury, NY, 2001.

17 A. Shevchenko, M. Wilm, O. Vorm, and M. Mann, *Anal. Chem.*, 1996, **68,** 850.

18 A. Shevchenko, H. Tomas, J. Havlis, J. Olsen and M. Mann, *Nat Protoc.*, 2006, **1**, 2856.

19 Peroxibase - the peroxidase database. http://peroxibase.isb-sib.ch/.

20 Spc proteomics tools. http://tools.proteomecenter.org/software.php.

21 A. Keller, J. Eng, N. Zhang, X. Li, and R. Aebersold, *Mol. Syst. Biol.*, 2005, **1**, msb4100024-E1.

22 F. Forner, L. Foster and S. Toppo, *Current Bioinformatics,* 2007, **2,** 63.

23 The x! tandem project. http://www.thegpm.org/TANDEM/index.html.

24 R. Craig and R. Beavis, *Bioinformatics*, 2004, **20**, 1466.

25 M. Wieser, *Pure Appl. Chem.*, 2006, **78**, 2051.

26 A. Keller, A. Nesvizhskii, E. Kolker and R. Aebersold, *Anal. Chem.*, 2002, **74**, 5383.

27 A. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, *Anal. Chem.*, 2003, **75,** 4646.

28 J.E. Elias and S.P. Gygi. *Nature Methods,* 2007, **4,** 207

29 J. Vizcaino, R. Côté, F. Reisinger, J. Foster, M. Mueller, J. Rameseder, et al. Proteomics 2009, **9,** 4276

30 H. Barsnes, J. Vizcaino, I. Eidhammer, and L. Martens, *Nat Biotech* 2009, **27,** 598

31 C. Notredame, D.G. Higgins, J. Heringa, *JMB* 2000, **302,** 205

32 E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel and A. Bairoch, (In) John M. Walker (ed): *The Proteomics Protocols Handbook*, Humana Press, 2005, **571**

33 E.M. Zdobnov and R. Apweiler, *Bioinformatics* 2001, **17**, 847

SUPPORTING FIGURES

Figure S1. Amino acid sequence of ZmPrx139 (PeroxiBase) in weevil- susceptible maize. Red: Peptides found in embryo sample SM1/2, Green: Peptide found in endosperm, blue: found in both fractions. Regions shaded in yellow indicate POD motifs assigned by InterProScan (accession PR00461).

```
ZmPrx139 (PeroxiBase)

        10         20         30         40         50         60
MAHAATLAAA AAVVVLIVVC CCSSSEAASH SANARQPPLA PGLSLDFYKR SCPRAESVVR
        70         80         90        100        110        120
RFVQDAVRRD AGLAAGLLRL HFHDCFVQGC DASVLLDGSA TGPGEQQAPP NLTLRPTAFK
       130        140        150        160        170        180
AINDIHDRLQ KECGGAVVSC SDVLALAARD SVVVSGGPSY KVPLGRRDSA SFATREDVLS
       190        200        210        220        230        240
GLPPPTAAVP ALLAVLSKIN LDATDLVALS GGHTIGLGHC ASFESRLFPR QDPTLNATFA
       250        260        270        280        290        300
GHLRRTCPAK GTDRRTPLDV RTPDAFDNKY YVNLVNREGL FTSDQDLFSN PGTRPLVEKF
       310        320        330        340        350
ARSQRDFFDQ FAFSVVKMGQ IKVLTGAQGQ IRRNCSARNA GMTWSVVDEA AQQSFVF
```

Figure S2. Amino acid sequence of B4FFK9_MAIZE in weevil-resistant maize. Red: Peptides found in embryo sample RM, Green: Peptide found in endosperm sample RE. Regions shaded in yellow indicate potential plant POD motifs as found by sequence similarity (see figure 2).

```
B4FFK9_MAIZE (Uniprot)

        10         20         30         40         50         60
MASSDGKPLP TPASVGGGGG SSTAPPGQPT TVASKVLDMG AAAMQSLRPV KQAKQHMCTF
        70         80         90        100        110        120
ALYAHDPKRQ VETHHYVSRL NQDFLQCAVY DSDKADARLI GVEYIVSRKI FDSLPAEEQR

       130        140        150        160        170        180
LWHSHAHEIK SGLWTSPHVA GLLEKAELDH MAATFGKFWC TWQVDRGDRL PLGAPALMVS
       190        200        210        220        230        240
PQADPAAAVR PDLVRKRDDR YGLSTEELRA ARADVEAPAE EHPGQADYWL RHRKGFAVDV
       250
VPHEMKRHAP FP
```