Title:

A Silk-Expressed Pectin Methylesterase Confers Cross-Incompatibility Between Wild and Domesticated Strains of *Zea mays*

Authors:

Yongxian Lu[1], Samuel A. Hokin[1], Jerry L. Kermicle[2], Mathew M. S. Evans[1*]

Affiliations:

[1]Department of Plant Biology, Carnegie Institution for Science, Stanford, CA 94305

[2]Laboratory of Genetics, University of Wisconsin, Madison, WI 53706

*Correspondence to: mevans@carnegiescience.edu

Abstract:

Despite being members of the same species, some strains of wild teosinte maintain themselves as a distinct breeding population by blocking fertilization by pollen from neighboring maize plants. These teosinte strains may be in the process of evolving into a separate species, since reproductive barriers that block gene flow are critical components in speciation. This trait is conferred by the *Teosinte crossing barrier1-s* (*Tcb1-s*) haplotype, making *Tcb1* a speciation gene candidate. *Tcb1-s* contains a female gene that blocks non-self-type pollen and a male function that enables self-type pollen to overcome that block. The *Tcb1-female* gene encodes a Pectin Methylesterase, implying that modification of the pollen cell wall by the pistil is a key mechanism by which these teosinte females reject foreign (but closely related) pollen.

One sentence summary:

The *Tcb1-female* gene encodes a Pectin Methylesterase that in teosinte silks prevents fertilization by maize pollen.

Short title:

Reproductive isolation and mate rejection in *Zea*

Main text:

Maize (*Zea mays* ssp *mays*) was domesticated from annual teosinte (*Zea mays* ssp *parviglumis*) in the Balsas River valley of Mexico (*1*). In some locations, sympatric populations of domesticated maize and annual teosinte grow in intimate associate and flower synchronously, but rarely produce hybrids (*2, 3*). In sexually reproducing plants, reproductive barriers exist at different stages, including pre-pollination, post-pollination, and post-fertilization. Post-pollination barriers depend on interaction between the pollen grain and the female reproductive organs (stigma, style, and ovule). In *Zea mays*, haplotypes at three loci, *Gametophyte factor1-s (Ga1-s)*, *Gametophyte factor2-s (Ga2-s)*, and *Teosinte crossing barrier1-s (Tcb1-s)*, confer Unilateral Cross-Incompatibility. While *Ga1-s* and *Ga2-s* are widespread in domesticated maize, *Tcb1-s* is almost exclusively found in wild teosinte populations. The *Tcb1-s* haplotype, like *Ga1-s* and *Ga2-s*, confers unilateral cross-incompatibility against varieties carrying the *tcb1* (or *ga1* or *ga2*) haplotype. Viewed otherwise, *Tcb1-s* provides a pollen function that overcomes the crossing barrier. The latter view is preferred since pollen containing both *Tcb1*-s and *tcb1* haplotypes fertilizes *Tcb1-s* plants, indicating that *Tcb1-s* compatibility is not overcome by the *Tcb1-s* : *tcb1* mismatch, as is also the case for the *Ga1* and *Ga2* systems (*4, 5*). *Tcb1-s* was first described in

44 teosinte subspecies *mexicana* Collection 48703 from the central and southern Mexico; this strain
45 also contained the male-only haplotype, *Ga1-m*, of the *Ga1* locus which together with male and
46 female functions of *Tcb1*, make up the Teosinte Incompatibility Complex (TIC) (*2, 3*).

47 Collections of teosinte of both *mexicana* and *parviglumis* subspecies from the central
48 Mexican plateau carry *Tcb1-s* (*6*). *Tcb1-s* confers to females the ability to block fertilization by
49 maize (*tcb1* type) pollen by restricting pollen tube growth (*7*). In the reciprocal cross, teosinte
50 pollen is able to fertilize maize, although poorly when in competition with maize pollen (*3*). *Tcb1*
51 was proposed to be a candidate speciation gene contributing to isolation of diverging maize and
52 teosinte populations, as wild teosinte populations respond to the pressure of cultivated, closely
53 related varieties of domesticated maize (*6*).

54 The male and female functions of *Tcb1-s* are tightly linked but separable by recombination
55 (*7*). Thus, there are four functional classes at this locus (Table S1 for gene content and origin):
56 *Tcb1-s* has both functional male and female genes, *Tcb1-m* has only the functional male gene (*6,*
57 *7*), *Tcb1-f* has only the functional female gene, and the *tcb1* haplotype found in almost all maize
58 lines has neither of the two functional genes. In teosinte, *Tcb1-f* activity in the silks prevents
59 fertilization by maize (*tcb1*) pollen, while *Tcb1-m* activity in pollen enables fertilization of *Tcb1-*
60 *f* females (*7*).

61 To clone the *Tcb1* genes, fine mapping of *Tcb1-s::Col48703* haplotype was performed
62 based on a *tcb1* backcross population with a population of approximately 15,000 chromosomes.
63 Using maize B73 genome as a reference (*8*), the *Tcb1* locus was delimited to a region spanning
64 480 kb on the short arm of chromosome 4. Within this region, there are eleven annotated genes.
65 However, all of these were ruled out as candidates for *Tcb1* functions because they either had
66 identical sequence with identical expression levels between *tcb1* and *Tcb1-s* haplotypes or no

67 expression in the silk or pollen in *Tcb1-s* or *tcb1* (mapping markers included in Table S2). The
68 *Tcb1* genes, therefore, are likely absent from the maize genome. This is not surprising considering
69 the widespread structural variations in genomes between maize lines and between teosinte
70 populations (*9*).

71 To identify *Tcb1-f* knockout mutants, maize lines homozygous for the *Tcb1-s::Col48703*
72 haplotype and carrying active *Mutator* transposons were crossed to maize inbred A195 *su1*. The
73 progeny are expected to be heterozygous for *Tcb1-s* with *su1* approximately 6 cM away and in
74 repulsion (*3*). Due to the rejection of the *tcb1* pollen (which is predominantly *su1*), about 3% of
75 the kernels in every ear with functional *Tcb1-f* were expected to be *sugary* in this open-pollinated
76 population, while any ears without a crossing barrier were predicted to segregate *su1* at 25%. Out
77 of a population of approximately 6,000 individuals, two exceptional ears were found. One ear
78 segregated for 25.6% sugary. This allele is termed *tcb1-f(KO1)*. The second isolate contained a
79 sector of about 45 kernels within which the segregation was one-fourth sugary despite sugary
80 segregating at ~3% over the rest of the ear. This allele is termed *tcb1-f(KO2)*. Mixed pollination
81 tests with the progeny of both individuals show that the loss of function is heritable, and both
82 variants fertilized a *Tcb1-s/tcb1* strain normally, indicating the retention of the male function of
83 *Tcb1-s* (*Tcb1-f* mutated, but *Tcb1-m* intact) (Fig.1). In the case of *tcb1-f(KO2)*, progeny of seeds
84 within the loss-of-function side of the ear inherited the knock-out, while those on the other side of
85 the ear inherited fully functional *Tcb1-s*.

86 RNA from silks of four genotypes were subjected to short read RNA-seq. Transcript
87 models were assembled *de novo* from the RNA-seq reads, and expression levels of genes were
88 compared between these two knockout mutants, a standard maize inbred line W22 (genotype *tcb1*),
89 and a functional *Tcb1-s* line (a W22 subline to which the *Tcb1-s::Col48703* haplotype had been

introduced by backcrossing). One gene, named here *Pertunda* (Roman fertility goddess who enables penetration, parallel to the control of pollen tube penetration in pistils), encoding a maize Pectin Methylesterase38 (PME38) homolog, was identified as a candidate for the *Tcb-f* gene. *Pertunda* is highly expressed in *Tcb1-s* silks (with a peak read depth of ~100,000) compared to the standard maize *tcb1* W22 silks, *tcb1-f(KO2)* silks (maximum read depths of ~100) and *tcb1-f(KO1)* silks (maximum read depth of ~10,000 for the 5' end and ~100 for the 3' end of the transcript model) (Fig.2 a). Based on the mRNA sequence, PCR primers were designed to isolate a BAC (Bacteria Artificial Chromosome) clone from a library we constructed from a maize line to which the *Tcb1-s::Col48703* haplotype had been introduced by backcrossing. By comparing mRNA and gene sequences, a 99-base intron was identified in *Pertunda*, which explains the gap between the two signal peaks in *Tcb1-s*. The intron region showed the same level of expression as that from the whole gene region in *tcb1-f(KO2)*and in the W22 maize line and in the downstream gene region in *tcb1-f(KO1)*. qRT-PCR confirmed this expression difference (Fig.2 b). *Pertunda* is not present in the maize B73 reference genome, which is consistent with the mapping data, and the closest homologs of *Pertunda* are located at the *ga1* locus (*10*).

In addition to the two knockout mutants from the active *Mutator* transposon population, several additional lines derived from the *Tcb1-s::Col48703* accession have lost female barrier function. One was recovered during early backcrossing of the *Tcb1-s::Col48703* haplotype into maize (*2*). Mixed pollination confirmed this is a *Tcb1-male* only plant (Fig.S1). Additionally, two independent *Tcb1-s* lines were isolated in which the barrier gradually lost the strength during ten generations of backcrossing to maize and selection for *Tcb1-male* function (*7*). These two lines were named as *tcb1-f::silent lineage1* (*tcb1-f::sl1*) and *tcb1-f::silent lineage2* (*tcb1-f::sl2*), based on the "progressive" manner of the barrier weakening. *Pertunda* has much lower expression in

these three additional *tcb1* lines, as shown with qRT-PCR (Fig. 2 b). Expression of *Pertunda* was also tested on the two *Tcb1-m* recombinants from the mapping population, which have lost the *Tcb1-female* gene by recombination. Again, *Pertunda* expression was much lower than in *Tcb1-s* lines (Fig. 2 b).

Using a PCR-based dCAPS (Derived Cleaved Amplified Polymorphic Sequence) marker designed for the *Pertunda* gene, it was shown that *Pertunda* maps to the *tcb1* locus (Fig. S2). This marker was then tested on the fifteen closest recombinants from the mapping population of ~15,000 individuals (including four recombinants between the *Tcb1-f* and *Tcb1-m* genes) (*7*). Of the fifteen plants, six carried *Tcb1-f* and blocked maize pollen, and nine lacked the barrier. Results showed that all the six recombinants that carry the barrier had the *Pertunda* gene, while in all nine recombinants that are receptive to maize pollen, *Pertunda* was absent. This shows a perfect physical linkage between *Pertunda* and the *Tcb1-f* barrier function. RNA-seq data suggest that the mutation in *Pertunda* occurred somewhere in the first exon in the *tcb1-f(KO1)* mutant (Fig. 2a). PCR data confirmed there was a disruption within the coding region of *Pertunda in KO1* (Fig. S3). Quite differently, in *tcb1-f(KO2)* mutant silk RNA-seq reads had the same low level of expression as *tcb1* silks along the whole *Pertunda* transcript. Whole genome resequencing of both mutants identified a *Hopscotch* retrotransposon insertion in the first exon in *tcb1-f(KO1)*, close to the site where *Pertunda* expression drops sharply. PCR spanning both ends of the insertion confirmed the insertion event and the border sequences (Fig. S4). In contrast, in *tcb1-f(KO2)*, *Pertunda* was fully assembled, consistent with the PCR data that the coding region is present (Fig. S3). The *tcb1-f(KO2)* allele then could either be mutated in a regulatory region, potentially hundred kilobases away from the coding region, or could be an epi-allele. Similarly, no mutations were found in the coding region of *Pertunda* in the *Tcb1-m* line or the *tcb1-f::sl1* or *tcb1-f::sl2* lines described above.

The *tcb1-f(KO2)*, *tcb1-f::sl1*, and *tcb1-f::sl2* lines were tested for reversion to *Tcb1-s* in double mutants with *mediator of paramutation1* (*mop1*) mutation. *MOP1* encodes a RNA-dependent RNA polymerase and is a key component of RNA-directed DNA Methylation (*11*). *mop1* mutations reactivate silenced genes and affect broad developmental programs (*12*). Re-activation of the *Tcb1-f* function was rare; in only ~14-22% of the *mop1* females tested, did the loss-of-function plants show some recovery of *Tcb1-f* function. Pollen competition experiments were performed for full strength *Tcb1-s* females, *tcb1* females, and the *tcb1-f* loss of function lines without sequence changes (primarily *tcb1-f(KO2)*, *tcb1-f::sl1*, and *tcb1-f::sl2*) (Fig. 3). All of the *Tcb1-s* ears tested showed strong preference for *Tcb1-s* pollen (0-7% kernels from *tcb1* pollen regardless of the ratio of the two pollen types in the mix as indicated by the neutral ear) with the kernel ratio on the test ear and control ear being different from each other at $p<0.0001$ (Fisher exact test) (Fig 3b). Of the 36 *mop1; tcb1-f loss of function* females tested only one had as strong of a pollen preference as full strength *Tcb1-s* females, but five had a difference between the test and control ears at $p<0.0001$ and an additional three females could be included if the stringency was relaxed to $p<0.01$ (Fig. 3b). These partial revertants included plants of lines *tcb1-f(KO2)*, *tcb1-f::sl1*, and *tcb1-f::sl2*. Of the twelve loss of function plants tested that were heterozygous wild-type for *mop1*, none of the plants passed the more stringent $p<0.0001$ threshold and one passed the less stringent $p<0.01$ threshold. It may be that maintaining the silenced derivatives of *tcb1-f* with *mop1* for multiple generations would increase the revertant frequency.

A subset of homozygous *mop1 tcb1-f::sl2* plants were tested at random for *Pertunda* expression in silks prior to pollination. Among the seven tested plants, one plant, yx57-13, showed about four hundred fold higher expression compared to that of the standard W22 maize and eight times higher than *tcb1-f::sl2* plants (Fig. 3C). This plant was the only one of those tested for

159 *Pertunda* expression that recovered the ability to reject *tcb1* pollen, although not as efficiently as
160 full strength *Tcb1-s* plants, which have still higher expression of *Pertunda* than this revertant. This
161 indicates a correlation between *Pertunda* expression level and the female barrier strength, and
162 further supports *Pertunda* as the *Tcb1-f* gene.

163 In addition to the *Tcb1-s::Col48703* strain descried above, three other teosinte-derived
164 *Tcb1-s* lines, two from ssp. *mexicana* and one ssp. *parviglumis* (*6*), were tested for *Pertunda*
165 expression in silk tissue. In all three lines, *Pertunda* expression levels are extremely high and
166 comparable to that of the original central plateau *TIC* haplotype *Tcb1-s::Col48703* (Fig. 4).
167 Interestingly, even though none of the modern north American maize lines tested to date carry the
168 *Tcb1-s* haplotype, an ancient Maiz Dulce variety, Jalisco78 line 1222-2, grown at intermediate
169 altitudes in southwestern Mexico carries *Tcb1-s* (*13*). This is a specialty line that may have
170 undergone selection for cross-incompatibility factors similarly to *Ga1-s* in maize popcorn lines.
171 Whether this Maiz Dulce line acquired *Tcb1-s* from nearby teosinte populations during its origin
172 is unknown, but maize lines from this region have been shown to have substantial introgression
173 from *mexicana* teosintes (*14*). Predicted *Pertunda* coding sequences are identical in all five *Tcb1-*
174 *s* lines: three *mexicana* accessions, one *parviglumis* accession, and the Maiz Dulce line. One
175 Single Nucleotide Polymorphism (SNP) in the intron separates these lines into two groups: one
176 group including the *parviglumis* line (Col104-4a) and one *mexicana* line (Col109-4a), and the
177 other group including two *mexicana* lines (Col48703 and Col207-5d) and the Maiz Dulce line (Fig.
178 S5).

179 The most similar gene to *Pertunda* is a candidate PME gene for *Ga1-female* function. This
180 gene, termed *ZmPME3*, was found to be expressed in the silks of *Ga1-s*, but not in *ga1* silks, and
181 *ZmPME3* was located to the *Ga1* mapping region(*10*). Alignment of the *ZmPME3* and *Pertunda*

show that the two PMEs differ in nine amino acids (Fig. S6). The number of polymorphisms (15 of 1296 nucleotides) between *Pertunda* and *ZmPME3* suggests that these two genes diverged approximately 175,000 years ago, well before the split between the *mexicana* and *parviglumis* subspecies of teosinte and just before the split between *Zea mays parviglumis* and *Zea luxurians*, using calculated nucleotide substitution rates for maize (*15*) and a calculated time since the split between *mexicana* and *parviglumis* of ~60,000 years and *parviglumis* and *luxurians* of ~140,000 years (*16*). It will be interesting to test whether the *Tcb1-male* and *Ga1-male* genes diverged at a similar time, suggesting they were already adjacent before divergence.

The *Tcb1* and *Ga1* barriers may share a similar mechanism, but because they are mostly cross-incompatible with one another they likely differ in their interacting partners. However, *Tcb1-s* and *Ga1-s* are not fully cross-incompatible. In situations where pollen rejection is not absolute, *Tcb1-s* pollen has a competitive advantage over *tcb1* pollen on *Ga1-s* or *Ga2-s* silks. This is true for all combinations of interactions between crossing barrier loci (*3, 5*) and is consistent with them encoding related proteins, although the behavior of pollen tubes during rejection by each system is slightly different (*7*). *Pertunda* encodes a group 1 type of PME without an N-terminal Pectin Methylesterase Inhibitor (PMEI) domain (*17*), and contains a predicted signal peptide, so it has the potential to be secreted and interact directly with the pollen tube to remove methyl-esters from the pectin wall of the pollen tube. Esterified pectins are typically associated with the tip of the growing pollen tube, while de-esterified pectins are enriched distally, and there is a correlation between pectin de-esterification and increased cell wall stiffness (*18*). Pollen tubes have a "soft tip-hard shell" structure, in that the tip region of the tube cell wall has a single pectin layer that is strong enough to withstand turgor pressure, but plastic enough to allow cell expansion and growth (*19*). Inside pollen tubes, pectin is synthesized and esterified in Golgi compartments

before delivery to the tip cell wall via vesicle trafficking (*20*), where it can be de-esterified by PMEs (*21*). Pollen cells finely tune the stiffness of the tip cell wall to sustain pollen tube elongation. Either under- or over-supply of PME activity can result in disturbed pollen tube growth and compromised male fertility (*22-25*). The PERTUNDA (and ZMPME3) protein falls into the Plant 1a clade of mature PME enzymes (*26*) (Fig. S7).

In summary, genetic and genomic data identify *Pertunda* as the *Tcb1-female* barrier gene. Teosinte lines carrying *Pertunda* block maize pollen that lacks the male function provided by *Tcb1-m*. That the *Tcb1-f* gene encodes a cell wall modifying enzyme is consistent with the model that incompatibility with *tcb1* occurs via incongruity rather than active targeting of a *Tcb1-m* encoded protein (*4*). It will be interesting to test how universal this barrier mechanism is among sexual reproducing plants. Surprisingly, it was shown that another PME family member is encoded by the *Ga1-male* gene (*27*) (in a very distinct clade, Plant X2, of PME enzymes), raising the possibility that the biochemical barrier to pollen and the ability of pollen to overcome that barrier are conferred by different classes of PME proteins.

The grass family is known to have widely distributed self-incompatibility (SI) among species, however, the molecular nature of the SI genes and how it is related to interspecific cross-incompatibility are not known(*28, 29*). The grasses also have an unusually high species diversity for a family with abiotic pollinators (*30*). Identification of the *Tcb1-female* gene may facilitate research into the mechanisms of speciation in the grasses. Agriculturally, this work may help managing specialty crop populations by preventing pollen contamination. It may also facilitate development of breeding tools to enrich crop genetic pools by backcrossing crops to their ancestors for the purposes of yield increase or enhanced stress resistance.

References and Notes:

1. Y. Matsuoka *et al.*, A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci U S A* **99**, 6080-6084 (2002).
2. J. L. Kermicle, J. O. Allen, Cross-Incompatibility between Maize and Teosinte. *Maydica* **35**, 399-408 (1990).
3. M. M. S. Evans, J. L. Kermicle, Teosinte crossing barrier1, a locus governing hybridization of teosinte with maize. *Theoretical and Applied Genetics* **103**, 259-265 (2001).
4. J. L. Kermicle, M. M. S. Evans, Pollen-pistil barriers to crossing in maize and teosinte result from incongruity rather than active rejection. *Sex Plant Reprod* **18**, 187-194 (2005).
5. J. L. Kermicle, M. M. S. Evans, The Zea mays sexual compatibility gene ga2: naturally occurring alleles, their distribution, and role in reproductive isolation. *J Hered* **101**, 737-749 (2010).
6. J. L. Kermicle, A selfish gene governing pollen-pistil compatibility confers reproductive isolation between maize relatives. *Genetics* **172**, 499-506 (2006).
7. Y. Lu, J. L. Kermicle, M. M. Evans, Genetic and cellular analysis of cross-incompatibility in Zea mays. *Plant reproduction* **27**, 19-29 (2014).
8. Y. Jiao *et al.*, Improved maize reference genome with single-molecule technologies. *Nature* **546**, 524-527 (2017).
9. R. A. Swanson-Wagner *et al.*, Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. *Genome Res* **20**, 1689-1699 (2010).
10. A. N. Moran Lauter, M. G. Muszynski, R. D. Huffman, M. P. Scott, A Pectin Methylesterase ZmPme3 Is Expressed in Gametophyte factor1-s (Ga1-s) Silks and Maps to that Locus in Maize (Zea mays L.). *Front Plant Sci* **8**, 1926 (2017).
11. M. Alleman *et al.*, An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442**, 295-298 (2006).
12. J. E. Dorweiler *et al.*, mediator of paramutation1 is required for establishment and maintenance of paramutation at multiple maize loci. *Plant Cell* **12**, 2101-2118 (2000).
13. Z. G. Jones, M. M. Goodman, M. D. Krakowsky, Identification of maize-derived dominant gametophyte factors. *Euphytica* **209**, 63-69 (2016).
14. J. van Heerwaarden *et al.*, Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc Natl Acad Sci U S A* **108**, 1088-1092 (2011).
15. R. M. Clark, S. Tavare, J. Doebley, Estimating a nucleotide substitution rate for maize from polymorphism at a major domestication locus. *Mol Biol Evol* **22**, 2304-2312 (2005).
16. J. Ross-Ibarra, M. Tenaillon, B. S. Gaut, Historical divergence and gene flow in the genus Zea. *Genetics* **181**, 1399-1413 (2009).
17. J. Pelloux, C. Rusterucci, E. J. Mellerowicz, New insights into pectin methylesterase structure and function. *Trends Plant Sci* **12**, 267-277 (2007).

18. E. Parre, A. Geitmann, Pectin and the role of the physical properties of the cell wall in pollen tube growth of Solanum chacoense. *Planta* **220**, 582-592 (2005).
19. M. W. Steer, J. M. Steer, Tansley Review No 16 Pollen-Tube Tip Growth. *New Phytologist* **111**, 323-358 (1989).
20. A. Y. Cheung, H. M. Wu, Structural and signaling networks for the polar cell growth machinery in pollen tubes. *Annu Rev Plant Biol* **59**, 547-572 (2008).
21. F. Micheli, Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *Trends Plant Sci* **6**, 414-419 (2001).
22. M. Bosch, A. Y. Cheung, P. K. Hepler, Pectin methylesterase, a regulator of pollen tube growth. *Plant Physiol* **138**, 1334-1346 (2005).
23. G. W. Tian, M. H. Chen, A. Zaltsman, V. Citovsky, Pollen-specific pectin methylesterase involved in pollen tube growth. *Dev Biol* **294**, 83-91 (2006).
24. N. Rockel, S. Wolf, B. Kost, T. Rausch, S. Greiner, Elaborate spatial patterning of cell-wall PME and PMEI at the pollen tube tip involves PMEI endocytosis, and reflects the distribution of esterified and de-esterified pectins. *Plant J* **53**, 133-143 (2008).
25. A. Sanati Nezhad, M. Packirisamy, A. Geitmann, Dynamic, high precision targeting of growth modulating agents is able to trigger pollen tube growth reorientation. *Plant J* **80**, 185-195 (2014).
26. O. Markovic, S. Janecek, Pectin methylesterases: sequence-structural features and phylogenetic relationships. *Carbohydr Res* **339**, 2281-2295 (2004).
27. Z. Zhang *et al.*, A PECTIN METHYLESTERASE gene at the maize Ga1 locus confers male function in unilateral cross-incompatibility. *Nat Commun* **9**, 3678 (2018).
28. J. Heslop-Harrison, Pollen-stigma interaction and cross-incompatibility in the grasses. *Science* **215**, 1358-1364 (1982).
29. B. Yang, D. Thorogood, I. Armstead, S. Barth, How far are we from unravelling self-incompatibility in grasses? *New Phytol* **178**, 740-753 (2008).
30. M. E. Dodd, J. Silvertown, M. W. Chase, Phylogenetic Analysis of Trait Evolution and Species Diversity Variation among Angiosperm Families. *Evolution* **53**, 732-744 (1999).
31. H. Wilkes, Teosinte: the closest relative of maize. *Bussey Institution of Harvard University, Cambridge, Massachusetts*, (1967).
32. M. M. S. Evans, J. L. Kermicle, Teosinte crossing barrier1, a locus governing hybridization of teosinte with maize. *Theoretical and Applied Genetics* **103**, 259-265 (2001).
33. M. G. Grabherr *et al.*, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**, 644-652 (2011).
34. A. Bankevich *et al.*, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455-477 (2012).
35. T. D. Wu, S. Nacu, Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26**, 873-881 (2010).
36. J. P. Huelsenbeck, F. Ronquist, MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755 (2001).

Acknowledgements:

The authors would like to thank Kathy Barton for the support that kept this project alive, encouragement, and insight and members of the Evans and Barton labs past and present for helpful discussions. We would also like to thank Beverly Oashgar for help with the screen for loss-of-function mutants and David Heller, Lance Cabalona, Clayton Coker, Amber Glowacki, and Hannah Vahldick for help growing plants and making crosses. We would also like to thank Jeffrey Ross-Ibarra for help in calculating the divergence time between *ZmPME3* and *Pertunda*. The authors declare that they have no competing interests in this work. This work was supported by National Science Foundation Award number IOS-0951259 and by United States Department of Agriculture-National Research Initiative Competitive Grants Program Award number 35301-13314.

Supplementary Materials:

Materials and Methods

Figures S1-S7

Tables S1-S2

References (31-36)

Fig. 1. Mixed pollination test of the *Tcb1-f* mutants. (A) scheme of the experiment: Two pollen donor lines and three pollen receiver lines were used. Pollen from a *tcb1; R1-self color* maize line (purple circles), produces purple kernels, while pollen from test plants (yellow circles) produces

white or yellow kernels. Pollen from the two donors was mixed and put on three receiving ears: (1) *Tcb1-s* ears to verify the *Tcb1-male* function from the *KO* line; (2) *tcb1-f(KO)* ears to test the presence/absence of the female barrier in the knockout mutant individuals; and (3) a maize (*tcb1; r1*) neutral ear to identify the ratio of viable pollen grains from the two donors in the mixture. (B) Ears from the three pollen recipients for the *tcb1-f(KO1)*) test. (C) Ears from the three pollen recipients for the *tcb1-f(KO2)* test. In both tests, pollen from the *KO* plants successfully fertilized the *Tcb1-s* ear (left ear in B and C), while the ears from *KO* plants showed no barrier to *tcb1* maize pollen with a similar frequency of purple kernels on mutant ears and the neutral test ears (middle and right ears in B and C).

Fig. 2. Gene expression profiling identifies *Pertunda* as a *Tcb1-f* candidate gene. RNA samples collected from silks of different genotypes was analyzed by RNA-Seq and RT-PCR. (A) *Pertunda* gene structure is shown above the graph (solid line indicates single intron), and RNA-Seq read depth is shown for the *Pertunda* gene. (B) *Pertunda* gene expression, compared to *tubulin* levels as a control, as measured by qRT-PCR in a variety of loss-of-function *Tcb1-s* lines: *Tcb1-s*, full strength *Tcb1-s* barrier line; *tcb1-f(KO1)* and *tcb1-f(KO2)*, two loss-of-function alleles from a *Mutator* transposon mutagenesis; *Tcb1-m*, a spontaneous *Tcb1-male* only line; *tcb1-f::silent lineage1* and *tcb1-f::silent lineage2*; and *K452-13* and *J456-13*, two *Tcb1-male* only lines that lost *Tcb1-f* by recombination.

Fig. 3. Reversion of *tcb1-f* loss-of-function. (A) Revertant ear (marked by asterisk in B and C) pollinated by a mix of *tcb1; R1-sc* and *Tcb1-m*; *r1* pollen showing higher frequency of yellow kernels on the test ear than the *tcb1* control ear. (B) Results of mixed pollination tests on four genotypes: *tcb1*, *Tcb1-s*, *tcb1-f(loss-of-function(lof))* alleles, and *tcb1-f(lof) mop1* double mutants.

355 The percentage of kernels from *tcb1* pollen on the test ear are plotted on the x-axis and the
356 percentage of kernels from *tcb1* pollen on the control *tcb1* ear on the y-axis. Equal percentages in
357 the two ears indicates no barrier (line with slope=1). Red circles indicate ears with significantly
358 fewer *tcb1* kernels in the test ear vs. the control ear at $p<0.0001$ and blue circles at $0.0001<p<0.01$.
359 The loss-of-function line is indicated for each revertant. (C) *Pertunda* gene expression level and
360 barrier strength. Barrier strength is expressed as the ratio of kernels from *Tcb1-s* vs *tcb1* pollen on
361 the test ear vs. the control *tcb1* ear (Columns), with a fraction of 1.0 indicating no barrier and
362 significantly higher values a functional barrier. *Pertunda* RNA levels are expressed as relative to
363 the *tubulin* control gene (Orange Line).

364 **Fig.4.** Expression of *Pertunda* in different *Tcb1-s* lines collected in different locations in Mexico
365 (*6*). W22 is a standard maize *tcb1* line, *Tcb1-s::Col48703*, *Tcb1-s::*109-4a, and *Tcb1-s::*207-5d
366 are three independent collections of *Zea mays ssp. mexicana* tesointe lines; *Tcb1-s::*104-4a, a
367 *Zea mays ssp. parviglumis* line; and DGF1222-2 from Maiz Dulce, an ancient Mexican maize
368 sweet corn variety (*13*).

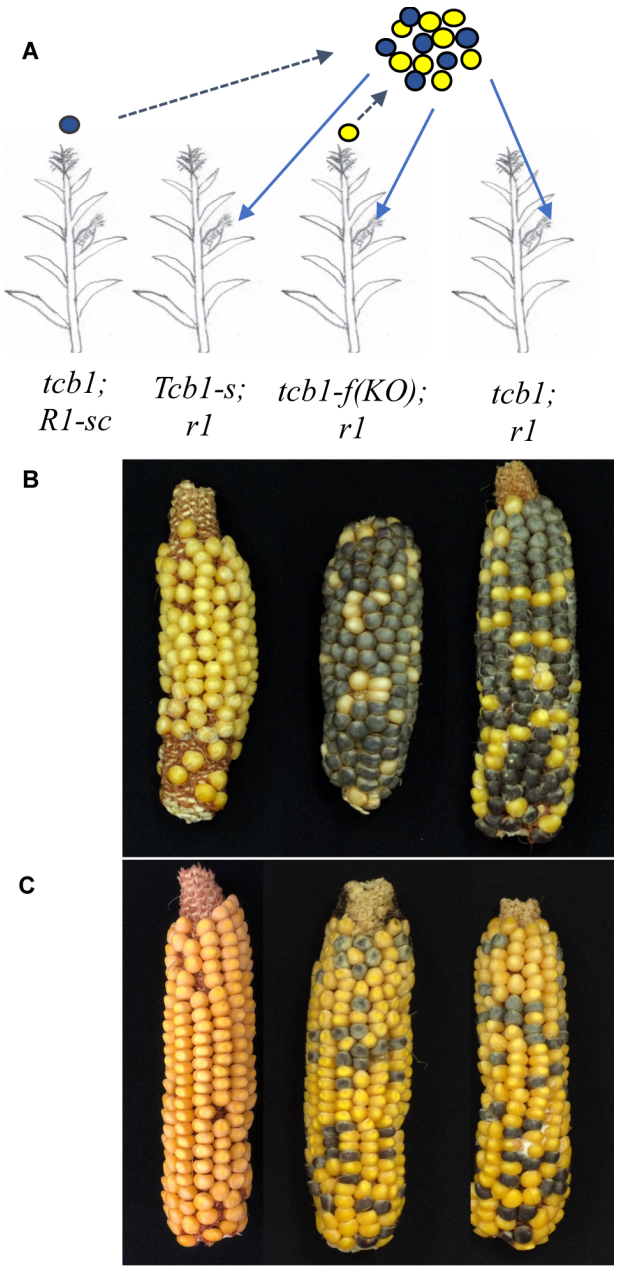
369

370

371

372 **Figures and Legends**

373



374

375

376

377 **Fig. 1.** Mixed pollination test of the *Tcb1-f* mutants. (A) scheme of the experiment: Two pollen
378 donor lines and three pollen receiver lines were used. Pollen from a *tcb1; R1-self color* maize line
379 (purple circles), produces purple kernels, while pollen from test plants (yellow circles) produces
380 white or yellow kernels. Pollen from the two donors was mixed and put on three receiving ears:
381 (1) *Tcb1-s* ears to verify the *Tcb1-male* function from the *KO* line; (2) *tcb1-f(KO)* ears to test the
382 presence/absence of the female barrier in the knockout mutant individuals; and (3) a maize (*tcb1;*
383 *r1*) neutral ear to identify the ratio of viable pollen grains from the two donors in the mixture. (B)
384 Ears from the three pollen recipients for the *tcb1-f(KO1)*) test. (C) Ears from the three pollen
385 recipients for the *tcb1-f(KO2)* test. In both tests, pollen from the *KO* plants successfully fertilized
386 the *Tcb1-s* ear (left ear in B and C), while the ears from *KO* plants showed no barrier to *tcb1* maize
387 pollen with a similar frequency of purple kernels on mutant ears and the neutral test ears (middle
388 and right ears in B and C).
389
390
391

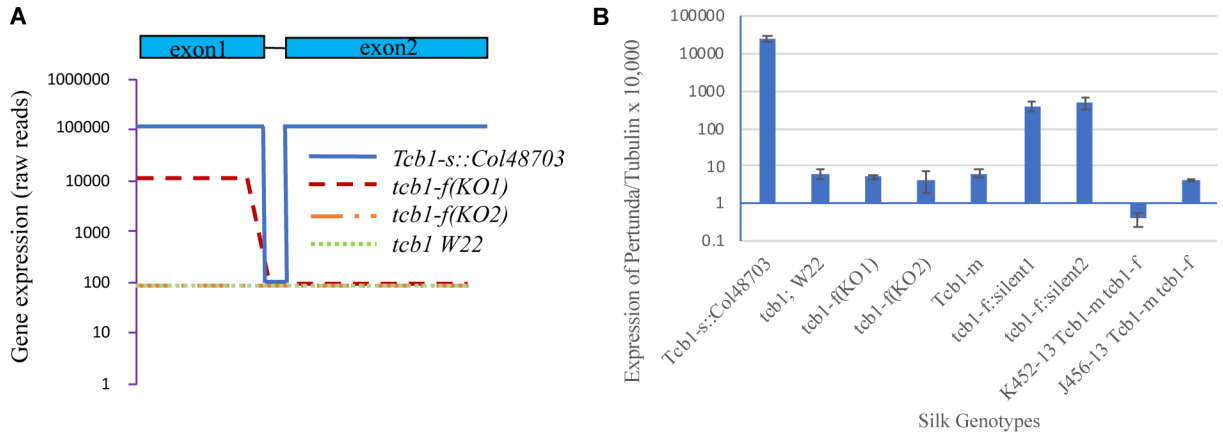
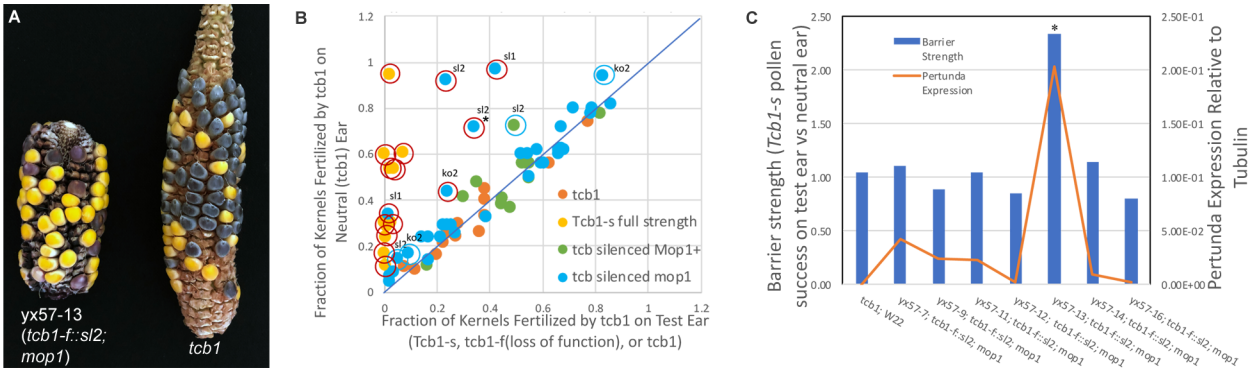


Fig. 2. Gene expression profiling identifies *Pertunda* as a *Tcb1-f* candidate gene. RNA samples collected from silks of different genotypes was analyzed by RNA-Seq and RT-PCR. (A) *Pertunda* gene structure is shown above the graph (solid line indicates single intron), and RNA-Seq read depth is shown for the *Pertunda* gene. (B) *Pertunda* gene expression, compared to *tubulin* levels as a control, as measured by qRT-PCR in a variety of loss-of-function *Tcb1-s* lines: *Tcb1-s*, full strength *Tcb1-s* barrier line; *tcb1-f(KO1)* and *tcb1-f(KO2)*, two loss-of-function alleles from a *Mutator* transposon mutagenesis; *Tcb1-m*, a spontaneous *Tcb1-male* only line; *tcb1-f::silent lineage1* and *tcb1-f::silent lineage2*; and *K452-13* and *J456-13*, two *Tcb1-male* only lines that lost *Tcb1-f* by recombination.

408



409

410

411 **Fig. 3.** Reversion of *tcb1-f* loss-of-function. (A) Revertant ear (marked by asterisk in B and C)
412 pollinated by a mix of *tcb1; R1-sc* and *Tcb1-m*; *r1* pollen showing higher frequency of yellow
413 kernels on the test ear than the *tcb1* control ear. (B) Results of mixed pollination tests on four
414 genotypes: *tcb1*, *Tcb1-s*, *tcb1-f*(*loss-of-function(lof)*) alleles, and *tcb1-f(lof) mop1* double mutants.
415 The percentage of kernels from *tcb1* pollen on the test ear are plotted on the x-axis and the
416 percentage of kernels from *tcb1* pollen on the control *tcb1* ear on the y-axis. Equal percentages in
417 the two ears indicates no barrier (line with slope=1). Red circles indicate ears with significantly
418 fewer *tcb1* kernels in the test ear vs. the control ear at $p<0.0001$ and blue circles at $0.0001<p<0.01$.
419 The loss-of-function line is indicated for each revertant. (C) *Pertunda* gene expression level and
420 barrier strength. Barrier strength is expressed as the ratio of kernels from *Tcb1-s* vs *tcb1* pollen on
421 the test ear vs. the control *tcb1* ear (Columns), with a fraction of 1.0 indicating no barrier and
422 significantly higher values a functional barrier. *Pertunda* RNA levels are expressed as relative to
423 the *tubulin* control gene (Orange Line).

424

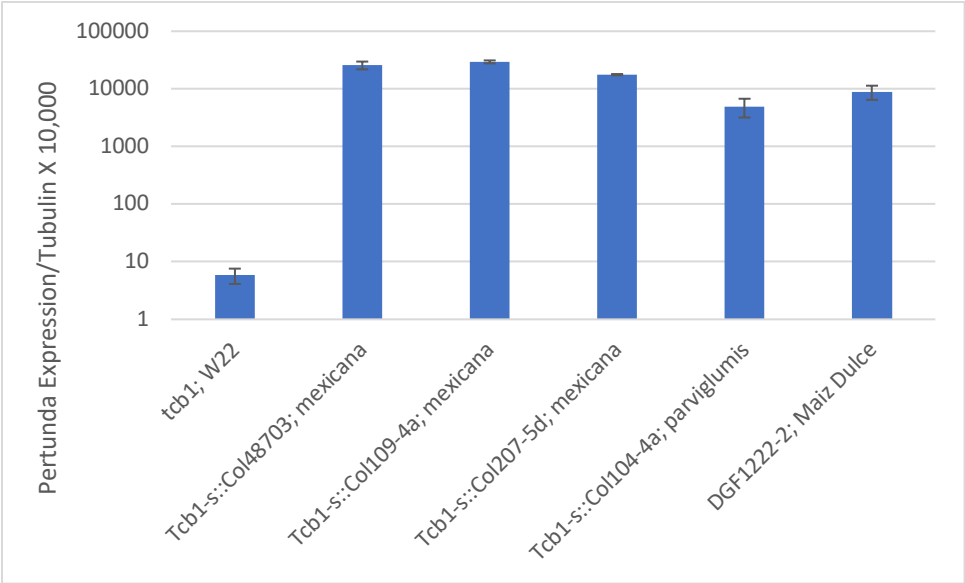


Fig.4. Expression of *Pertunda* in different *Tcb1-s* lines collected in different locations in Mexico (*6*). W22 is a standard maize *tcb1* line, *Tcb1-s::Col48703*, *Tcb1-s::*109-4a, and *Tcb1-s::*207-5d are three independent collections of *Zea mays ssp. mexicana* tesointe lines; *Tcb1-s::*104-4a, a *Zea mays ssp. parviglumis* line; and DGF1222-2 from Maiz Dulce, an ancient Mexican maize sweet corn variety (*13*).

Supplementary Materials for

A Silk-Expressed Pectin Methylesterase Confers Cross-Incompatibility Between Strains of *Zea mays*

Yongxian Lu[1], Samuel A. Hokin[1], Jerry L. Kermicle[2], Mathew M. S. Evans[1]

Correspondence to: mevans@carnegiescience.edu

This PDF file includes:

Materials and Methods

Figs. S1 to S7

Tables S1 to S2

Materials and Methods

Maize and teosinte lines and growth conditions

All maize and teosinte lines used in this study have been described previously (*3, 6, 7, 13*). Plants were grown under field conditions at either Stanford, California or Madison, Wisconsin.

Tcb1-s mapping

As described before (*3*), a Central Plateau teosinte collection 48703 (*31*) carrying the *Tcb1-s* barrier was backcrossed to the Mid-western US dent inbred W22 to incorporate the *Tcb1-s* locus into a maize background. This *Tcb1-s* strain was crossed to a chromosome 4 maize tester line *virescent17* (*v17*) *brown midrib3* (*bm3*) *sugary1* (*su1*), and the F1 was then backcrossed to the same tester line. Recombinants carrying crossovers between the four visual markers were tested for the *Tcb1-s* male and female functions in reciprocal crosses with *Tcb1-s/su1* F1 plants. PCR mapping markers were developed to refine the location of crossovers in these recombinants

Tcb1-f knockout mutant screen

To identify loss-of-function mutants of *Tcb1-female*, a *Ga1-m Tcb1-s* active *Mutator* strain was crossed to maize inbred A195 *su1* (*tcb1*), and then the progeny were grown as an open-pollinated block. Most of the progeny are expected to be heterozygous for *Tcb1-s* and *su1* in repulsion with *su1* approximately 6 cM away from the *tcb1* locus (*32*). Due to the rejection of the *tcb1* pollen (which is predominantly *su1*), about 3% of the kernels in every ear with functional *Tcb1-f* are expected to be *sugary* in this open-pollinated population, while those without a crossing barrier were predicted to segregate *su1* at 25%.

474 Mixed pollination experiments

475 For the mixed pollination testing of the two *tcb1-f* knockout mutants, two pollen donor lines and
476 three pollen receiver lines were used. Pollen from a maize line (*tcb1*) that does not have the *Tcb1*
477 barrier genes but carries the endosperm color marker *R1-self color* (*R1-sc*) will produce purple
478 kernel after fertilization of the lines used, while pollen from the knockout plants and the *Tcb1-m*
479 plants carry *r1-r* and produce anthocyaninless kernels that are white or yellow. After being
480 collected from the two donors and mixed, pollen was put on the three receiving ears: (1) a *Tcb1-s*
481 tester ear was used to verify the presence of the *Tcb1* male function from the *tcb1-f::KO* pollen;
482 (2) the *tcb1-f (KO)* ear was used to test the presence/absence of the female barrier function in the
483 knockout mutant; and (3) a maize (*tcb1*) neutral ear was used to assay the percentage of viable
484 pollen grains from the two donors in the mixture. The same protocol was used on the spontaneous
485 *Tcb1-m* plant, except the *Tcb1-m* plant being tested was substituted for the *KO* plant. For mixed
486 pollinations of the *tcb1-f::silent lineage mop1* double mutant plants, pollen from the same *R1-sc*
487 *tcb1* line and a *r1-r Tcb1-s* tester line was collected, mixed, and applied to the individual silent
488 line ears and the neutral maize ears.

489 Silk tissue collection, RNA isolation and cDNA synthesis

490 Plants for RNA isolation were grown in summer field conditions in Stanford, CA. Silk tissues were
491 collected around 11 am, immediately put into liquid nitrogen in the field, and stored at -80°C. Total
492 RNA was isolated from silks with Trizol reagent (Invitrogen), DNase-treated, and either subjected
493 to Illumina short read paired end RNA-seq, or used to synthesize the 1st strand cDNA with the
494 Superscript IV RT kit (Invitrogen).

495 Quantitative RT-PCR

Each line/genotype had three biological replicates, and each in turn had three technical repeats. Tubulin (Zm00001d033850) was used as a reference gene. In each line, relative expression level of *Pertunda* was obtained by comparing *Pertunda* to tubulin.

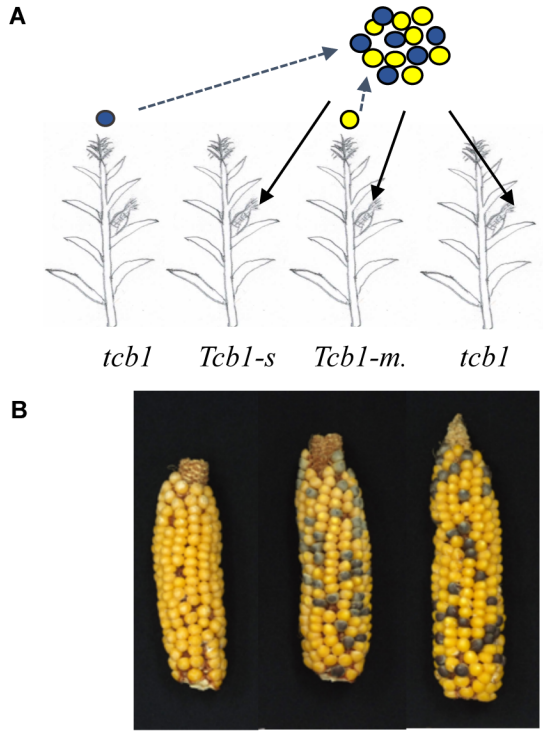
Sequencing, assembly and analysis

All the RNA and DNA sequencing works were done with Illumina Paired-end sequencing by Novogene (CA, USA). RNA-seq reads from all samples were combined and *de novo* assembled with Trinity v2.4.0.(*33*) The gene in contig DN33598_c7_g3_i1 was identified as the *Tcb1-f* candidate gene due to its extremely high expression in the functional *Tcb1-s* line and the almost no expression in the *KO* mutants and a standard W22 maize line. PCR primers were designed based on the DN33598_c7_g3_i1 sequence, and one BAC clone was fished out from library made from maize line into which the *Tcb1-s::Col48703* haplotype had been introgressed. The BAC sequencing reads were assembled with SPAdes v3.11.1.(*34*) NODE_62, a contig that is 13656 bp with coverage of 4029, was identified as having the *Tcb1-f* candidate gene. Whole genome sequencing reads from the two *KO* mutants were individually assembled with SPAdes v3.11.1 and BLASTed against NODE_62. Also, the mutant sequencing reads were mapped against NODE_62 using GSNAP (*35*). Combining both approaches identified the *hopscotch* retrotransposon insertion in the *tcb1-f(KO1)* mutant allele.

For phylogenetic analyses, alignments were made using the ClustalW algorithm in MegAlign (DNASTAR). The predicted mature PME enzymes and the Arabidopsis PME family members were taken from Markovic and Janecek, as were the subfamily designations (*26*). Phylogenies were produced from these alignments using MrBayes v3.2.0 using default settings for amino acid analysis (*36*). The MrBayes analysis was performed for 4,100,000 generations at which point the standard deviation of the split frequencies was below 0.004.

519 **Supplementary Figures**

520



521

522 **Fig. S1**. Mix pollination testing the spontaneous *Tcb1-m* plant. a. scheme of the experiment: Details
523 in "Method" section. b. ears from the three pollen receivers for *Tcb1-m* plant test. Pollen from the
524 *Tcb1-m* plant successfully fertilized the *Tcb1-s* ear and produced yellow kernels (left ear in b),
525 while the ears from *Tcb1-m* plant indeed had lost the barrier to block maize pollen as shown by the
526 purple kernels produced by *tcb1* pollen on the ears (middle ear in b).

527

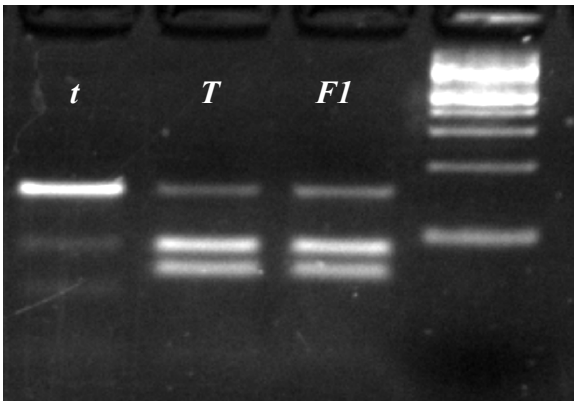


Fig. S2. A dCAPS marker to test presence/absence of *Pertunda* in recombinants from the mapping population. The Marker was designed in the way that only the PCR amplicon from the *Tcb1* genomic DNA (*T*), but not the unspecific PCR product amplified from the maize gnomic DNA (*t*) would be cut by the enzyme Hae III. after PCR and enzyme digestion.

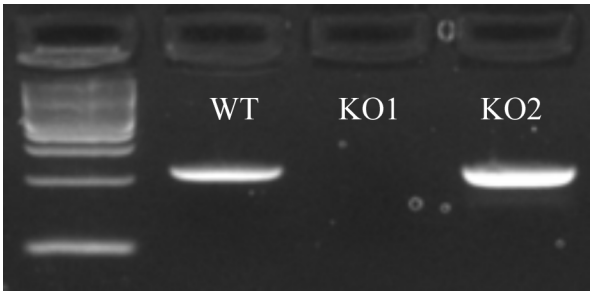


Fig. S3. PCR detection of *Pertunda* in the two *tcb1-f(KO)* mutants. PCR primers spanning the possible mutation site in the *tcb1-f(KO1)* was designed and tested on the *Tcb1-s* (WT), *tcb1-f(KO1)* and *tcb1-f(KO2)*. Using *tcb1-f(KO1)* genomic DNA as template failed to produce amplicon, while *Pertunda* can be detected in the *tcb1-f(KO2)* mutant.

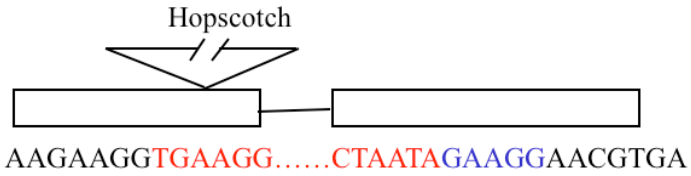


Fig. S4. Identification of the mutation in the *tcb1-f(KO1)* mutant. Whole genome resequencing showed a Hopscotch retrotransposon insertion in the first exon. PCR primers based on the gene and the retrotransposon were used to confirm the insertion and the border sequences. Black bases, *Pertunda* gene sequence; Red bases, retrotransposon sequence; Blue bases, bases duplicated from the left border leading the retrotransposon sequence.

```
Col104-4a parviglumis   ACAGTGAGTGAGTAACTCTGCCGATCGATATTTCACTGCAGTTGAG…
Col109-4a mexicana      ACAGTGAGTGAGTAACTCTGCCGATCGATATTTCACTGCAGTTGAG…
Col48703  mexicana      ACAGTGAGTGAGTAACTGTGCCGATCGATATTTCACTGCAGTTGAG…
Col207-5d mexicana      ACAGTGAGTGAGTAACTGTGCCGATCGATATTTCACTGCAGTTGAG…
Maiz Dulce              ACAGTGAGTGAGTAACTGTGCCGATCGATATTTCACTGCAGTTGAG…
                        ********** ****************************
```

Fig. S5. Partial DNA alignment of the *Pertunda* gene intron between the different *Tcb1-s* lines. Collections 109-4a, 48703 and 207-5d, ssp. *mexicana* teosintes; Collection104-4a, ssp. *parviglumis* teosinte; DGF1222, a traditional maize Dulce sweetcorn variety from Mexico (*13*).

```
PERTUNDA  MVGGVRRCGLGLAMAVALLLAALVVVASGGAETRQKLPAGSGNDDDHAAVLSRLSNVIDP
ZmPME3    MVGGVRRCGLGLAMAVALLLAALVVVASGGAEMRQKLPAGSGNDDDHAAVLSRLSNVIDP
          ******************************** ************************

PERTUNDA  PGSWPPRADAVVAKRCRGVAAPPPCYTSIQAAVDHAPAPQEAEEVEDKYVVHVLAGVYDE
ZmPME3    PGSWPPRADAVVAKRCGGVAAPPPCYTSIQAALKAASAPQEAEEVEDKYVVHVLAGVYDE
          **************** ***************:. * ***********************

PERTUNDA  TVNITRRNVMLIGDGVGATVITGNKSNATGVHMDMTATVNALGHGFIAQNLTIRNTAGPE
ZmPME3    TVNITRRNVMLIGDGVGATVITGNKSNATGVHMDMTATVNALGHGFIAQNLTIRNTAGPD
          ************************************************************:

PERTUNDA  GRQAVALRSNSNKSVVYWCSIEGHEDTLYVENGIQFYLQTSIWGTVDFVFGNAQAMFQSC
ZmPME3    GRQAVALRSNSNKSVVYCCSIEGHEDTLYVENGIQFYLQTSIWGTVDFVFGNAQAMFQSC
          ***************** ****************************************

PERTUNDA  ALLVRRPPKGKHNVLTAQGCNNASRESGFSFHMCTVEAAPGVDLDGVETYLGRPYRNFSH
ZmPME3    ALLVRRPPKGKHNVLTAQGCNNASRESGFSFHMCTVEAAPGVDLDGVETYLGRPYRNFSH
          ************************************************************

PERTUNDA  VAFIKSYLSRVVSPNGWVAWNKNKVVEDTTRTILYLEYGNDGAGADTAGRVKWPGFRVLN
ZmPME3    VAFIKSYLSRVVSPNGWVAWNKNKVVDDTTRTILYLEYGNDGAGADTAGRVKWPGFRVLN
          ************************:***********************************

PERTUNDA  TDDEAIAYTADTFINASKWVPEPIQYVHTLGTAPPPRA
ZmPME3    TDDEAIAYTADTFINASKWVPEPIQYVHTLGTAPPPRA
          ************************************
```

Fig. S6. Alignment of Pertunda and ZmPME3 proteins with Clustal Omega.

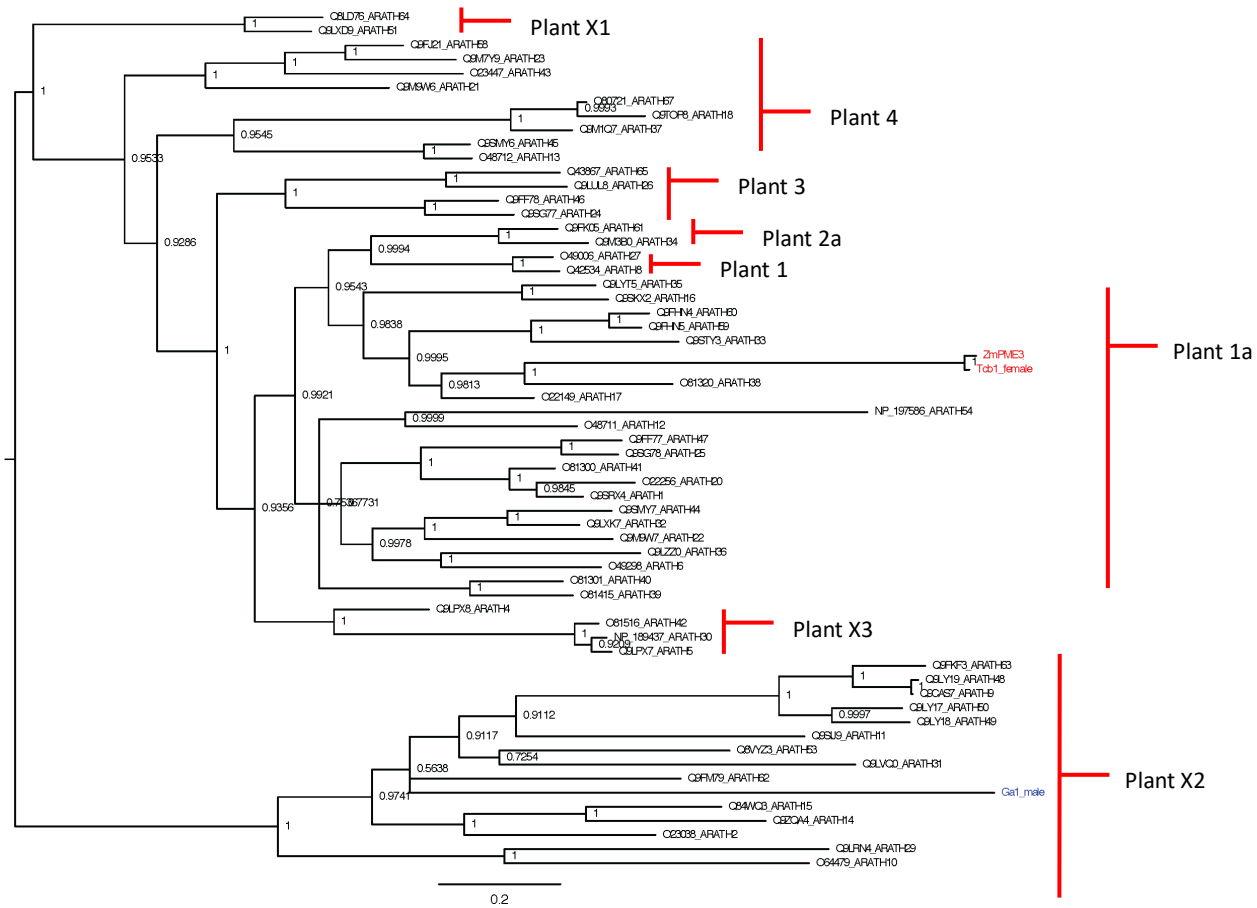


Figure S7. Phylogenic tree of mature PME enzymes (predicted pre and pro domains removed) of Arabidopsis PME proteins and predicted PME proteins encoded by cross-incompatibility loci of *Zea mays*.

631

Table S1. Haplotypes of the *tcb1* locus

Haplotype	Gene Content	Female Barrier	Pollen Penetration	Principal source
Tcb1-s	*Tcb1-f (Pertunda) Tcb1-m*	+	+	Teosinte
Tcb1-f	*Tcb1-f (Pertunda)*	+	-	Recombination
Tcb1-m	*Tcb1-m*	-	+	Teosinte
tcb1	-	-	-	Maize

632

633

634

635

Table S2. PCR primers for mapping *Tcb1-s* relative to the maize B73 reference genome

Marker	Primer sequences	Marker types	Note
053600	FP: AGCAGGTGCCGCCCG T RP: AAGCGAGGGTTTGCTCGAT	Indel	*Tcb1-s* left border
444073 up8	FP: GAGCAGGGTCAGCTGGAAGGAAC RP: GGCATCTCTTGCTTGGCGGGC	dCAPS	*Tcb1-f* left border
Gene 0.1 up	FP: GATAAGTTTGCAAGTGGCCCATCA RP: AGAGATGTTCTGCAATGGGCCTA	Indel	*Tcb1-m* right border
106164	FP: GGCTGAAGCAGCAGAGCCATCCTAA RP: CTGTGTGTGGGATCGGATCCCTA	Indel	*Tcb1-s* right border

636

637

638

639

640

641