Supplementary Materials

1. Materials and Methods:

Plant material and growth conditions

The maize cultivar A188 was used in all experiments. Plants were grown at 28°C day/ 20°C night in a 16 hr light/8 hr dark cycle with a light intensity of 230 µE m-2 s-1. Seeds were germinated in three-inch diameter pots containing peat-based soil and grown for three weeks before transferring to eleven-inch or 40 L pots to flowering. Anthers were manually dissected at v12 stage and meiotic stage was determined by acetocarmine staining. For heat stress experiments, maize plants grown in eleven-inch pots were transferred to growth chambers for 3 days (16 h light at 35°C/ 8 h dark at 25°C, light intensity 230 µE m-2 s-1, humidity 75%). Tobacco plants (*Nicotiana benthamiana*) were grown on M2 soil (Levington Advance, UK) at 22°C in a 16 h light/ 8 h dark cycle with light intensity of 100 µE m-2 s-1.

Identification of transposon insertion mutants

To identify transposon insertion lines for Zm00001d007786 (GRMZM2G05903) and Zm00001d013063 (GRMZM2G123063) we screened a Mutator insertion mutant population generated by Biogemma, an Ac/Ds mutant population (*1*) and an UniformMu mutant population (*2*). Insertion lines were confirmed by PCR (Supplementary Table S6) and backcrossed to A188 inbred for four generations before analysis.

Vector construction and generation of transgenic plants

We generated a MAGO1/2-RNAi vector by subcloning a portion corresponding to position 2151-2400 of Zm00001d007786 (GRMZM2G05903) and a fragment corresponding to position 1001-1250 of Zm00001d013063 (GRMZM2G123063) into a pCsVMV::intOsActin-intStLS1-terSbHSP vector using Golden Gate cloning.

To generate a chemically inducible helper component-proteinase (HC-Pro) construct we synthetized a 2,211 bp fragment from the Wheat Streak Mosaic Virus WSMV containing Gateway Recombination flanking sequences (Invitrogen, USA). The synthetic fragment was subcloned into a two-component dexamethasone-inducible expression binary vector named pZZ-TOP, which is derived from pTF101.1 and carries a LhG4:GR synthetic gene fusion and a bi-

33 directional pOp6 promoter (*3*). This vector enables the co-expression of a Beta-Glucuronidase

34 (GUS) reporter and HC-Pro after exposure to 20mM Dexamethasone (DEX).

35 To enable the epidermal expression of HC-Pro, we generated a pZmHDZIV6-LhG4 vector using

36 MultiSite Gateway Recombination (Invitrogen, USA). The native 3164 bp promoter and the 945

37 bp native terminator of Zm00001d002234 (GRMZM2G001289) was cloned into pDONR221 P1-

38 P4 and pDONR221 P3-P2, respectively. The maize codon-optimized LhG4 was cloned into

39 pDONR221 P4r-P3r. The resulting three entry vectors were then recombined into the binary

40 vector pAL010, a derivative of pTF101.1 carrying a Gateway Recombination Cassette and a bi-

41 directional pOp6 promoter enabling the co-expression of a Beta-Glucuronidase (GUS) and NLS-

42 tdTomato reporters. All constructs were fully sequenced before transformation in maize using

43 *Agrobacterium tumefaciens* strains LBA4404 (MAGO1/2-RNAi and pZmHDZIV6-LhG4) or

44 EHA101 (HC-Pro).

45 To determine the in vivo activity of MAGO proteins we generated a firefly Luciferase silencing

46 reporter system. First, we generated a construct containing a nopaline synthase (NOS) promoter

47 (pNOS), a firefly luciferase (FLUC) fused to four miR2118-target sequences (PHAS), a

48 truncated-GFP (ΔGFP) and a NOS terminator (tNOS). This synthetic fragment was cloned in

49 pBINPLUS using *Hin*dIII and *Bam*HI restriction enzyme digestion. The miR2118 target region

50 was designed to have optimal hybridization energies with the RNA target. The resulting

51 construct was digested with *Sma*I and *Eco*RI to enable the insertion of a p35S::GUS:tNOS

52 fragment derived from the pSLJ4J8 vector. We also generated a fragment containing the

53 octopine synthase (OCS) promoter, an artificial miRNA based on maize miR2118c (amiR2118)

54 and a nopaline synthase (NOS) terminator. This fragment was cloned into pBINPLUS using *Asc*I

55 restriction enzyme digestion. To simultaneously express MAGO and amiR2118, we generated

56 codon-optimized MAGO1 and 2, containing a FLAG tag in the carboxy-terminal end, which

57 were cloned into pCsVMV::intOsActin-terSbHSP using SapI resulting in the binary vectors

58 pBIOS11743 and pBIOS11746, respectively. For the phosphorylation study, we generated

59 codon-optimized MAGO2 phosphomimetic (S>E) and phosphoresistant (S>A) forms, containing

60 a FLAG tag at the carboxy end, which were cloned into pCsVMV::intOsActin-terSbHSP using

61 SapI resulting in the binary vectors pBIOS11747 and pBIOS11748, respectively. MAGO2

62 mutants for sRNA binding (Y676E) and cleavage (D835E) were generated using Q5® Site-

63 Directed Mutagenesis Kit (NEB, UK). The pOCS::amiR2118:tNOS was sub-cloned into

64   pBIOS11743 and pBIOS11746 using AscI restriction enzyme digestion. Generated constructs

65   were fully sequenced and transformed in *Nicotiana benthamiana* using *Agrobacterium*

66   *tumefaciens* GV3101.

67   To study the in vivo localization of MAGO proteins, codon-optimised coding regions for each

68   gene were subcloned in pGWB441 (*4*) to generate c-terminal EYFP protein fusion after Gateway

69   Recombination (Invitrogen, USA). All constructs generated in this study were fully sequenced

70   and transformed in *Nicotiana benthamiana* using *Agrobacterium tumefaciens* GV3101.

71

Collection of maize meiocytes

73   The meiotic stage of anthers was determined using acetocarmine staining. Meiocytes were

74   isolated by manual micromanipulation in RNase-free PBS and collected using microglass

75   pipettes controlled by UMP3 UltraMicroPump (WPI). Collected meiocytes were directly frozen

76   in liquid nitrogen and stored at -80°C.

77

Antisera preparation and immunopurification

79   Polyclonal antisera were raised in rabbit against synthetic peptides for MAGO1

80   (VETEHQQGKRSIYRI) or MAGO2 (CVAAREGPVEVRQLPK) (Eurogentec, Liege, BE). To

81   generate HC-Pro antisera, a partial DNA fragment was chemically synthesized (Integrated DNA

82   Technology, UK) and cloned in pET29a (Novagen, Merck, Darmstadt, Germany). A soluble

83   fraction of HC-Pro was isolated and purified by metal affinity chromatography and polyclonal

84   antiserum was raised in rabbit (Eurogentec, Liege, BE). All antisera were affinity purified using

85   a Sulpholink coupling gel system (Pierce, Rockford, IL).

86

MAGO immunoprecipitation

88   For immunoprecipitation, maize anthers were isolated by micromanipulation and ground in

89   extraction buffer (20 mM Tris-Cl pH 7.5, 300 mM NaCl, 5 mM MgCl2, 5 mM DTT, 1% (v/v)

90   cOmplete EDTA-free protease inhibitor cocktail (Merck, Darmstadt, Germany). The lysates

91   were pre-cleared with protein-A agarose beads (Sigma-Aldrich, Poole, UK) and incubated with

92   anti-MAGO1 (1:200), anti-MAGO2 (1:200) or anti-HC-Pro (1:100) antibodies for 2h at 4°C.

93   Protein-A agarose beads were added to the sample and incubated for a further 2h. Subsequently,

94   the beads were washed 4-5 times using extraction buffer supplemented with 0.5% NP-40 (Sigma,

95   St. Louis, MO). For western blot analysis, beads were suspended in 1x SDS loading buffer and

96   heated at 95°C. For isolation of small RNA, the bead slurry was digested with proteinase K (100

97   μg·mL-1) and incubated for 1 h at 37°C prior to RNA extraction using TRIZOL® Reagent

98   (Invitrogen, UK).

99

100 **RNA extraction**

101 Total RNAs were extracted from immature tassels, anthers, meiocytes or leaves with TRIzol®

102 Reagent (Invitrogen, UK) as per manufacturer's instruction. RNAs for sequencing were

103 extracted using Direct-zol RNA miniprep kit (ZYMO Research, Cambridge). The extracted

104 RNAs were quantified using Nanodrop (ThermoFisher, UK) and the quality checked with a

105 Bioanalyzer 2100 (Agilent, UK).

106

107 **Preparation of RNA libraries and sequencing analysis**

108 To isolate small RNAs (sRNAs), total RNA was fractionated on a 15% polyacrylamide TBE-

109 Urea gel (Novex, UK). The gel was stained with ethidium bromide for 5 mins at room

110 temperature and visualised with a UV illuminator. Gel pieces were macerated by incubation with

111 300 mM NaCl overnight followed by RNA precipitation. For sRNA sequencing, fractionated

112 sRNAs were used to construct libraries using the TruSeq Small RNA Sample Preparation Kit

113 (Illumina, UK) and sequenced in single-end 50 base mode an Illumina HiSeq platform

114 (University of Delaware). For total RNA sequencing, total RNAs with RNA integrity number

115 (RIN) >8.0 were used for library construction using a TruSeq RNA Sample Preparation Kit

116 (Illumina, UK) and sequenced in single-end 150 base mode an Illumina NexSeq platform

117 (University of Warwick). NanoPARE libraries were prepared as described previously (Schon et

118 al. 2018). Briefly, cDNA libraries were generated from 5 ng of total RNA using the original

119 Smart-seq2 protocol (Picelli et al. 2013). Five nanograms of cDNA was tagmented using Nextera

120 DNA Flex library preparation kit as described in the manufacture instructions. Tagmented cDNA

121 was purified using Zymo DNA Clean and Concentrator kit and eluted with 20ul nuclease-free

122 water. This purified tagmented product was split into halves and used as a substrate for final

123 enrichment PCR with either Tn5.1/TSO or Tn5.2/TSO oligonucleotide primer sets (Schon et al.

124 2018). PCR reaction products from Tn5.1/TSO and Tn5.2/TSO oligonucleotide primer sets were

125 pooled together and purified using Beckman Coulter AMPureXP DNA beads. The nanoPARE

126 libraries were sequenced in single-end 50 base mode on an Illumina Hi-Seq 2500 instrument.

127 Library details and number of reads are provided in Table S7.

128

129 **Induction of HC-Pro in developing anthers**

130 Maize HC-Pro transgenic plants were grown to adult stage and whorls were cut open to reveal

131 the developing male inflorescence (tassels). Each floret was filled with a DEX solution (50 µM

132 DEX, 0.1% Silwet-77) and allowed to grow to maturity under normal growth conditions. To

133 determine the efficiency of DEX-induction, anthers were submerged in GUS solution (50 mM

134 sodium phosphate, 1 mM potassium ferricyanide, 1 mM potassium ferrocyanide, 0.1% Triton X-

135 100, 10 mM EDTA, 1.0 mg·mL-1 5-Bromo-4-chloro-3-indolyl-β-d-glucuronic acid), incubated

136 at 37°C for 16-24 h until blue precipitate was observed. DEX--treated and mock-treated anthers

137 were sampled to determine pollen viability. To check induction of HC-Pro, total protein was

138 extracted from anthers using QB buffer (100 mM KPO4 (pH 7.8), 1 mM EDTA, 1% Triton X-

139 100, 10% glycerol, 1 mM DTT, Protease inhibitor cocktail (Roche, UK). The extracted proteins

140 were quantified by Quick Strat Bradford Protein Assay (Bio-Rad, UK). Western blot detection

141 was carried out with immunopurified anti-HC-Pro antisera and images recorded using an

142 ImageQuant gel documentation instrument (GE Healthcare Life Sciences, UK).

143

144 **Confocal microscopy analysis**

145 To determine transcriptional activation of pZmHDZIV6::LhG4 in anther epidermis, transgenic

146 plants were grown to reproductive stage and tassels were fixed in SR2200 solution (4% PFA in

147 PBS (pH 7.4), 0.1% SR2000 (Renaissance Chemicals)), vacuum infiltrated, washed with PBS

148 and submerged in ClearSee solution (10% xylitol (w/v), 15% sodium deoxycholate (w/v), 25%

149 urea (w/v) (Kurihara et al. 2015). The samples were vacuum infiltrated, incubated until tassels

150 were cleared, washed and stored in PBS. The cleared tissue was embedded in 4% low melting

151 agarose in PBS and the embedded tissue was mounted onto vibratome blocks, 150   m sections

152 were cut by Lancer Vibratome Series 1000 (TPI, USA). Tissue slices were placed onto glass

153 slides, covered with a coverslip and imaged with a LSM710 confocal microscope (Zeiss, Jena,

154 GE).

155

156 **cDNA synthesis and RT-PCR**

157 Total RNA was treated with Ambion® TURBO DNase kit (Life technologies, USA). DNase-
158 treated RNAs were used for cDNA synthesis using Superscript® reverse transcriptase II
159 (Invitrogen, UK). Semi-quantitative RT-PCR was performed using templates as the synthesized
160 cDNAs. GAPDH was used for data normalization. For Quantitative real-time RT-PCR (qRT-
161 PCR) the optimal number of cycles was determined for each gene. PCR cycling conditions
162 included denaturing at 95ºC for 15 s, annealing at 57ºC for 30 s and extension at 72ºC for 45,
163 using a Bio-rad qRT-PCR machine (BioRad, UK). Changes in expression levels were calculated
164 via the $\Delta\Delta$Ct method. To ensure primer specificity, qRT-PCR was done when the melting curve
165 showed a single peak. To quantify small RNAs we used stem loop qRT-PCR following
166 previously reported methods (Yang et al. 2014; Varkonyi-Gasic, 2017) with minor modifications.
167 Reverse transcription was performed using the RevertAid First Strand cDNA Synthesis Kit
168 (Thermo Scientific, USA) as per the manufacturer's instructions. For stem loop qRT-PCR, we
169 used a 10 µL of RT reaction mixture containing 1 µL of RNA, 1 µL of RT primer (5 µM) and 1
170 µL of U6 RT primer (5 µM), 1 µL of 10 mM dNTP Mix, 2 µL of reaction buffer, 0.5 µL of
171 Ribolock RNase inhibitor (20 U/µL), 0.5 µL revertAid M-MuLV Reverse Transcriptase (200
172 U/µL). The mixture was incubated at 25℃ for 5 min, and then incubation was continued at 42℃
173 for 60 min. The reaction was inactivated by heating at 70℃ for 5 min and to which 1 µL of RT
174 product, 5 µL of SYBR Green real-time PCR Master Mix, and 1 µL of primer (forward and
175 reverse, 1 µM each) was added. Reactions were incubated in a PCR cycler at 95℃ for 3 min,
176 followed by 40 cycles of 95℃ for 5 s, 62℃ for 35 s. Primer sequences are listed in
177 Supplementary Table S6.
178

179 **Display and high-throughput sequencing of retrotransposon insertions**
180 For the detection of retrotransposon insertions, we used Splinkerette PCR (*5*). Genomic DNA
181 was isolated from leaves using a urea gDNA extraction method (*6*). Genomic DNA was digested
182 by *Bst*1 (NEB) overnight and cleaned with a MinElute DNA clean-up kit (Qiagen). End repair
183 was carried out by incubating overnight with T4 DNA polymerase (NEB) followed by A-tailing
184 before clean-up with a MinElute DNA cleanup kit (Qiagen). To generate the Splinkerette adapter,
185 Long-strand adaptor and Short-strand adaptor oligos were synthesised and annealed by heating
186 for 10 mins at 72ºC and allow them to cool at room temperature. Fragmented DNA was ligated
187 to a Splinkerette adapter by T4 ligase (NEB) followed by clean-up with MinElute DNA cleanup

188   kit (Qiagen). Ligated genomic DNA was then used for two rounds of nested PCR with Phusion
189   High-Fidelity Polymerase (NEB). Splink1 primer and retrotransposon-specific round-1 primer
190   was used for the first round of PCR. This PCR product was used as template for the second PCR
191   with Splink2 primer and retrotransposon-specific round 2 primers. Each PCR product was
192   resolved on a 6% Acrylamide gel and stained with ethidium bromide before imaging.
193   For the high-throughput sequencing to identify and map retroelement insertion sites we followed
194   the method developed by Dooner, Wang, Huang, Li, He, Xiong and Du (*7*) with minor
195   modifications. An equal amount of young leaf-tissue was harvested from 10 plants and DNA was
196   extracted by the Urea method. A modified Splinkerette-PCR was used to isolate the retroelement
197   insertion sites. DNA was sheared using a Bioruptor sonication system (Diagenode, Belgium) to a
198   mean size of ~1.7-kb and size-selected by 0.8xAgencourt AMPure XP beads (Beckman Coulter,
199   Brea, CA). The protocol of KAPA library preparation Kits (Kapa Biosystems Inc., Wilmington,
200   MA) was followed in subsequent end-repairing, A-tailing and adaptor-ligation procedures. PCR
201   amplifications followed the protocol of Phusion High-Fidelity Polymerase (NEB, Ipswich, MA).
202   Biotin-Splink1 primers were used for 1st round PCR using Phusion High-Fidelity Polymerase
203   (NEB, Ipswich, MA) and the PCR product was purified using Dynabeads® M-280 Streptavidin
204   (Thermo Fisher Scientific, Carlsbad, CA). These purified products were used as template for
205   second PCR using primers for different retroelements that were barcoded to allow sample
206   multiplexing. The oligonucleotides required for Splinkerette-PCR are listed in Supplementary
207   Table S6. We constructed sequencing libraries, after amplicons were end-repaired, A-tailed, and
208   ligated to Illumina TruSeq Single Index Barcoded adaptors from the Illumina TruSeq LT DNA
209   kit (Illumina, Kapa Biosciences). Adaptor-ligated DNA was amplified in a PCR reaction with 1X
210   Kapa HF PCR Master Mix (Kapa Biosciences), and 1X TruSeq PCR Primer Cocktail (Illumina).
211   Libraries were sequenced on an Illumina NextSeq platform (University of Warwick). Primer
212   sequences are listed in Supplementary Table S6. Library details and number of reads are
213   provided in Table S7.
214

215   **Sequencing data and statistics analyses**
216   Analyses of small RNA sequencing data were carried out using previously described methods (*8*).
217   Mapping of small RNAs to AGPv4 reference genome was performed using Bowtie. Any read
218   with more than 50 perfect matches ("hits") to the genome was excluded from further analysis.

219   Abundances of small RNAs in each library were normalized to "TP10M" (transcripts per 10

220   million) based on the total count of genome-matched reads in that library. For the analysis of

221   RNA-seq data, reads were trimmed and mapped to AGPv4 reference genome using TopHat2 (*9*)

222   and the expression annotated genes and transposons was quantified using TEtranscripts (*10*). For

223   the analysis of NanoPARE sequencing data, we used a described analysis pipeline and

224   determined candidates for sRNA-mediated cleavage of retrotransposons using EndCut (*11*).

225

226   **Phosphoproteomic analysis**

227   Protein from anthers were extracted by adding three times the volume of extraction buffer (50

228   mM HEPES, 150 mM NaCl, 1 mM EDTA, 20 mM NaF, 1 mM Na2MoO4, 1% (v/v) SDS, 1 mM

229   PMSF, 2 µM Calyculin A, 1 mM NaVO4, 1 mM DTT, Protease inhibitor cocktail (Roche)) to

230   0.5 g of tissue. After 30 min the samples were spun for 15 min at 4,000 g (4°C) to remove debris.

231   The supernatant was transferred to a new tube, centrifuged for 30 min at 16,000 g (4°C),

232   transferred to a new tube and treated using the FASP protocol (*12*). Samples were loaded on

233   Amicon® Ultra-2 mL Centrifugal Filters with a cutoff of 3 kDa and diluted with 1 ml 8 M urea

234   until 1 ml of buffer was passed through the column. Reduction and alkylation of the cysteine

235   residues was carried out by adding a combination of 5 mM tris (2-carboxyethyl) phosphine

236   (TCEP) and 10 mM iodoacetamide (IAA) for 30 min at room temperature in the dark, followed

237   by six washes with 25 mM Hepes (pH 7.5). The protein was digested with trypsin (Promega

238   Trypsin Gold, mass spectrometry grade) overnight at 37 °C at an enzyme-to-substrate ratio of

239   1:100 (w:w). After digestion the peptides were suspended in 80% acetonitrile (AcN), 5%

240   trifluoroacetic acid (TFA) and the insoluble matter was spun down at 4000 g for 10 min. The

241   supernatant was used for the enrichment of phosphopeptides as previously described with minor

242   modifications (*13*). The peptide concentration was measured with a Qubit™ fluorometer

243   (Invitrogen) and 1 µg total peptides were used for each sample. The Titansphere TiO2 10 µm

244   beads (GL Sciences Inc.) were equilibrated in a buffer containing 20 mg/mL 2,5-

245   dihydroxybenzoic acid (DHB), 80% ACN and 5% TFA in a ratio of 10 µl DHBeq per 1 mg

246   beads for 10 min with gentle shaking at 600 rpm. TiO2 beads were used in a ratio of 1:2 peptide-

247   bead ration (w:w). The TiO2 solution was added to each sample and incubated for 60 min at

248   room temperature. This step was repeated one more time. The samples were then spun down at

249   3000 g for 2 min and resuspended in 100 µL Wash buffer I (10% AcN, 5% TFA). The

250   resuspended beads were added to self-made C8-columns. C8-colums were made of 200 µL

251   pipette tips with 2 mm Empore™Octyl C8 (Supelco) discs. The columns were spun down at

252   2600 g for 2 min, washed with 100 µL Wash buffer II (40% AcN, 5% TFA) and 100 µL Wash

253   buffer III (40% AcN, 5% TFA). The peptides were eluted from the TiO_2 beads with 20 µL 5%

254   ammonium hydroxide and subsequently with 20 µL 20% ammonium hydroxide in 25% AcN.

255   Both eluates were pooled, the volume was reduced to 5 µL in a centrifugal evaporator (20–30

256   min) and acidified with 100 µL of buffer A (2% AcN, 1% TFA). Samples were desalted with a

257   self-made C18 column (Empore™Octadecyl C18). C18 were made in the same way as the C8-

258   columns. Before adding the samples, the C18-columns were activated with 50 µL methanol and

259   washed with 50 µL AcN and 50 µL buffer A* (2% AcN, 0.1% TFA). Samples were loaded on

260   the C18-column and spun at 2000 g for 7 min. The columns were washed with 50 µL ethyl

261   acetate and 50 µL buffer A* and then eluted consecutively with 20 µL 40% AcN and 20 µL 60%

262   AcN. Samples were then vacuum-dried and prior to MS analysis resuspended in 50 µL buffer A*.

263

264   **Mass spectrometry**

265   Reversed phase chromatography was used to separate tryptic peptides prior to mass

266   spectrometric analysis. Two columns were utilised, an Acclaim PepMap µ-precolumn cartridge

267   300 µm i.d. x 5 mm 5 µm 100 Å and an Acclaim PepMap RSLC 75 µm x 25 cm 2 µm 100 Å

268   (Thermo Scientific). The columns were installed on an Ultimate 3000 RSLCnano system

269   (Dionex). Mobile phase buffer A was composed of 0.1% formic acid in water and mobile phase

270   B 0.1 % formic acid in acetonitrile. Samples were loaded onto the µ-precolumn equilibrated in 2%

271   aqueous acetonitrile containing 0.1% trifluoroacetic acid for 8 min at 10 µL min-1 after which

272   peptides were eluted onto the analytical column at 300 nL min-1 by increasing the mobile phase

273   B concentration from 3% B to 35% over 40 min and then to 90% B over 4 min, followed by a 15

274   min re-equilibration at 3% B.

275   Eluting peptides were converted to gas-phase ions by means of electrospray ionization and

276   analysed on a Thermo Orbitrap Fusion (Q-OT-qIT, Thermo Scientific). Survey scans of peptide

277   precursors from 350 to 1500 m/z were performed at 120K resolution (at 200 m/z) with a 4×10^5

278   ion count target. Tandem MS was performed by isolation at 1.6 Th using the quadrupole, HCD

279   fragmentation with normalized collision energy of 35, and rapid scan MS analysis in the ion trap.

280   The MS2 ion count target was set to 1×10^4 and the max injection time was 200 ms. Precursors

9

281 with charge state 2–7 were selected and sampled for MS2. The dynamic exclusion duration was

282 set to 45 s with a 10ppm tolerance around the selected precursor and its isotopes. Monoisotopic

283 precursor selection was turned on. The instrument was run in top speed mode with 2 s cycles.

284

285 **Mass spectrometry data analysis**

286 A label-free peptide relative quantification analysis was performed in Progenesis QI for

287 Proteomics (Nonlinear Dynamics, Durham). To identify peptides, peak lists were created by

288 using Progenesis QI. The raw data was searched against maize B73 RefGen_4 Working Gene set.

289 Peptides were generated from a tryptic digestion with up to two missed cleavages,

290 carbamidomethylation of cysteines as fixed modifications, oxidation of methionine and

291 phosphorylation of serine, threonine and tyrosine as variable modifications. Precursor mass

292 tolerance was 5 ppm and product ions were searched at 0.8 Da tolerances. Scaffold (TM, version

293 4.4.5, Proteome Software Inc.) was used to validate MS/MS based peptide and protein

294 identifications. Peptide identifications were accepted if they could be established at greater than

295 95.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if

296 they could be established at greater than 99.0% probability and contained at least one identified

297 peptide. Proteins that contained similar peptides and could not be differentiated based on

298 MS/MS analysis alone were grouped to satisfy the principles of parsimony. List of differentially

299 accumulated phosphopeptides are listed in Table S4.

300

301 **Protein structure modelling**

302 To model the conserved catalytic domain of MAGO2, we used MODELLER (*14*). First, we

303 scanned the PDB database (http://www.rcsb.org/) to identify proteins with known structure

304 whose sequences could be best aligned with that of MAGO2. This search identified the *K.*

305 *polysporus* Argonaute (PDB ID 4F1N) (*15*) as a good match for roughly the last two thirds of the

306 MAGO2 sequence. Moreover, three other proteins, namely the human Argonaute1-3, were found

307 to be viable templates covering the rest of the MAGO2 sequence, with varying degree of overlap

308 with KpAGO. Of these, we chose the Argonaute 2, because of the high resolution and

309 crystallographic quality of one of its available structures (PDB ID 4Z4D) (*16*). To use both

310 structure as templates for our modelling, we first aligned them to each other, and then aligned the

311 fitted structures to the sequence of MAGO2. We then used this multiple-template alignment to

312 produce 64 base models of MAGO2, refining the loop regions in each of them twice. Each

313 refinement was repeated 16 independent times, resulting in a total of 1,024 different models. To

314 find the best one, we assessed each model using a high-resolution version of the DOPE (Discrete

315 Optimized Protein Energy) method (*17*), and picked the model with the best score, checking it by

316 hand to ensure it contained no knotted loops or other unphysical structures.

317

318 Having produced an initial model, we refined it using molecular dynamics (MD) simulations, to

319 obtain a realistic final structure. All MD simulations were carried out using Amber18 (*18*). To

320 prepare the parameters for the simulations, we first added hydrogens to the pdb file of the model

321 using the pdb4amber and reduce programs (*19*). To ensure the model was properly folded, we

322 decided to run an initial relaxation simulation using implicit solvation, to exploit the speed-up in

323 conformational sampling that this method provides. More specifically, we used the Generalized

324 Born model (*20*) with a set of optimized atomic parameters for proteins (*21*). Thus, to create the

325 topology parameters, we used the ff14SBonlysc force field, which uses the same parameters as

326 the ff99SB force field (*22*) for the backbone, but full quantum-mechanics ones for the side-

327 chains (*23*), and which is known to work best in this setup.

328 We then minimized the structure with 16000 steps of steepest descent, before heating the system

329 gradually over 0.5 ns from 0 K to 295.15 K, using a Langevin thermostat with collision

330 frequency of 2.0 ps-1. For this and subsequent implicit-solvation steps we constrained the length

331 of the bonds with hydrogens using SHAKE (*24*), imposed a cutoff for nonbonded pair and

332 effective Born radii calculations of 24 Å, and used an integration step of 2 fs. Also, forces

333 involving the derivatives with respect to the effective Born radii were computed every 2

334 integration steps. After heating the system, it was allowed to relax at constant temperature, and

335 computed the total potential energy and its individual contributions (bond energy, dihedral angle

336 energy, van der Waals 1-4 interaction energy, electrostatic 1-4 energy, total van der Waals

337 interaction energy and total electrostatic energy). We stopped the relaxation when we observed at

338 least 20 ns of stability in each of the components, as well as in the total potential energy, which

339 we took as an indicator that no further conformational changes were likely to occur (*25*).

340 To obtain a more realistic model we then performed one more relaxation in explicit solvent. We

341 prepared the starting topology from the last simulation frame of the previous step using the same

342 procedure described above. However, this time we solvated the protein using the TIP3P water

343   model (*26*) in a truncated octahedral box imposing a minimum distance between the edges of the

344   box and the atoms of the protein of 8 Å. Also, for this simulation step we used the full ff14SB

345   force field (*23*), as the current gold-standard for simulations with explicit water molecules, and

346   neutralized the charge of the protein adding 16 Cl- ions, treated via the parameters by Joung and

347   Cheatham (*27, 28*). We then found the optimal distribution of water molecules by constraining

348   the protein atoms via a harmonic potential with a coupling constant of 500 kcal/(mol Å2) and

349   minimizing the potential energy of the system. Subsequently, we removed the constrains and

350   minimized the whole system again, allowing every atom to move. For both minimization steps

351   we used steepest descent and stopped the minimization process when the root-mean-square of

352   the components of the potential energy gradient became smaller than 0.05 kcal/(mol Å2). For

353   this and all other explicit-solvation steps, we used a nonbonded interaction cutoff of 8 Å,

354   constrained the hydrogen-involving bonds using SHAKE, used an integration step of 2 fs, and

355   evaluated slowly-varying terms in the force field at every step.

356   Having obtained a minimized structure, we heated it using the same protocol described above.

357   Then, we equilibrated the system at a constant temperature and constant pressure of 1 bar for 4.5

358   ns using a Monte Carlo barostat with pressure relaxation time of 1.0 ps and attempting a volume-

359   change move every 100 integration steps. Finally, we relaxed the system with the same protocol

360   as above, before annealing it to 0 K by decreasing the temperature gradually over 5 ns and

361   minimizing the resulting structure.

362   To produce a model for the phosphorylated MAGO2 protein, we started from the final non-

363   phosphorylated structure we obtained. We then mutated the relevant residues to their

364   phosphorylated versions, and produced a topology for MD simulations. Since we started this step

365   from an already realistic model of the protein, we had no need to perform an implicit-solvation

366   step before passing to an explicit-solvent simulation. Thus, the protocol we used is the same as

367   the one described above for the explicit-solvation case, with the key changes that we used the

368   phosaa10 force field for the phosphorylated residues (*29, 30*), and the ff99SB for the rest of the

369   protein. This last choice was due to the fact that phosaa10 uses the same assumptions as ff99SB.

370   Thus, its use would not be compatible with force fields of the ff14SB family. Also note that due

371   to the extra negative charges of the phosphoryl groups, only 8 chlorine counter-ions were needed

372   to neutralize the total charge of the system. We then produced a relaxed structure of the system

373   using the exact same steps described previously. To compute the electrostatic potential surfaces,

374   we used PBSA to solve a linearized version of the Poisson-Boltzmann equation, using a level-

375   set-function implementation of the dielectric interface, imposing a smooth molecular surface via

376   density function calculation (*31*), and estimating the nonpolar free energy of solvation as the sum

377   of a cavity term and a dispersion term. For this calculation, we considered an ionic strength of

378   150 mM, a solvent probe radius of 1.4 Å and a solvent-accessible arc resolution of 0.25 Å.

379

380   **2. Figures:**

381
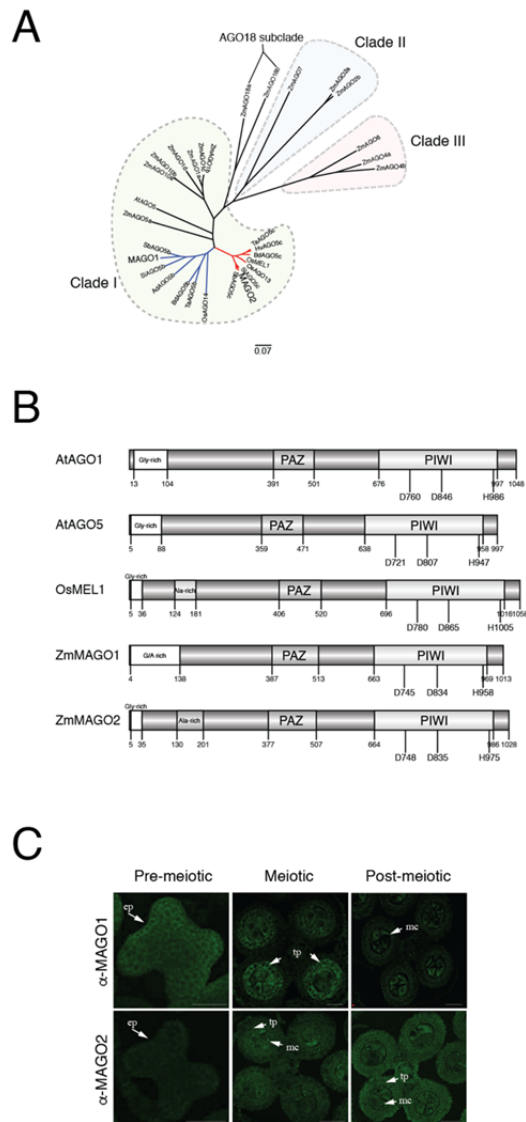
382

383 **Fig. S1. Identification of two Male-Associated Argonaute-like (MAGOs) in maize.**

384 (A) Distance-based phylogeny tree between MAGOs and other monocotyledonous Argonaute-

385 like proteins constructed using the Neighbour-joining method.

13

386   (B) Schematic diagram showing the conserved domains present in MAGO and related
387   Argonaute-like proteins.   PAZ, Piwi Argonaute and Zwille domain; PIWI, PIWI domain.
388   Catalytic amino acid residues (DDH) are indicated.
389   (C) Immunodetection of MAGO1 and MAGO2 in developing anthers using specific antisera.
390   White arrow, accumulation of MAGO protein; ep, epidermis; tp, tapetum; mc, meiocyte. Scale
391   bars are 50 μm.
392

393
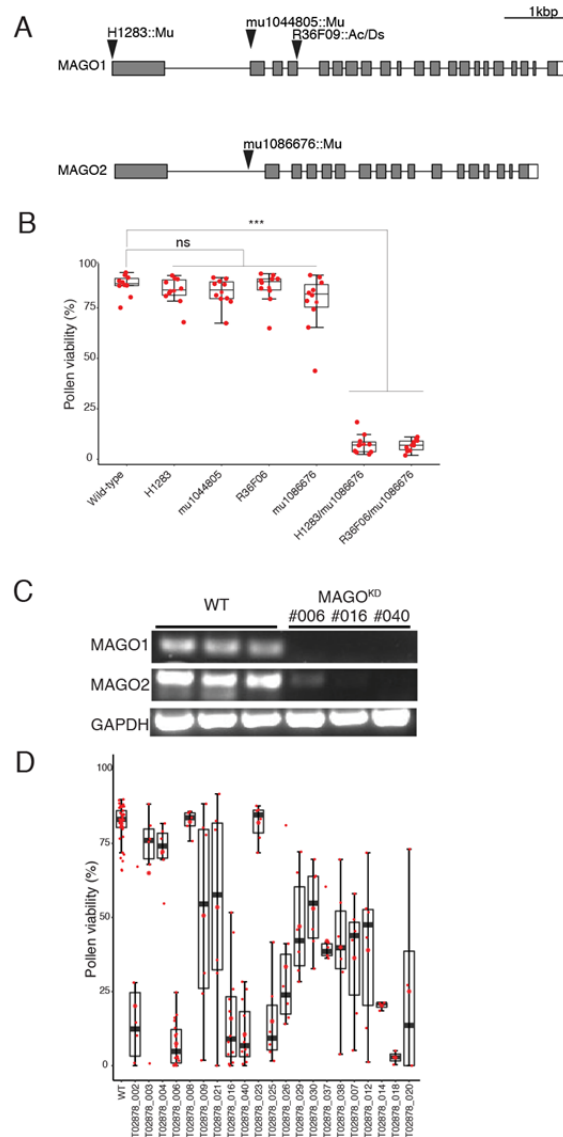
394
Fig. S2. Identification and characterization of transposon insertions for MAGO1 and MAGO2 and down-regulation RNAi lines (MAGO^{KD}).
(A) Schematic diagram showing four independent transposon insertion mutant alleles identified for MAGO1 and MAGO2. Black arrowhead, transposon insertion; Grey box, exon; White box, untranslated region.
(B) Pollen viability in wild-type and homozygous transposon insertion plants grown under field conditions. n≥ 10 independent plants analysed per genotype. Differences between groups were determined by Tukey HSD, ***p < 0.001; n.s. no-significant.
(C) Accumulation of MAGO 1 and 2 transcripts in pre-meiotic anthers determined by RT-PCR. GAPDH was used as a constitutive control.
(D) Pollen viability in field-grown wild-type and 21 independent MAGO^{KD} lines. n≥ 6 hemizygous T2 plants analysed per genotype; more than 10 anthers analysed per plant. Black line, median; Red star, mean.

408
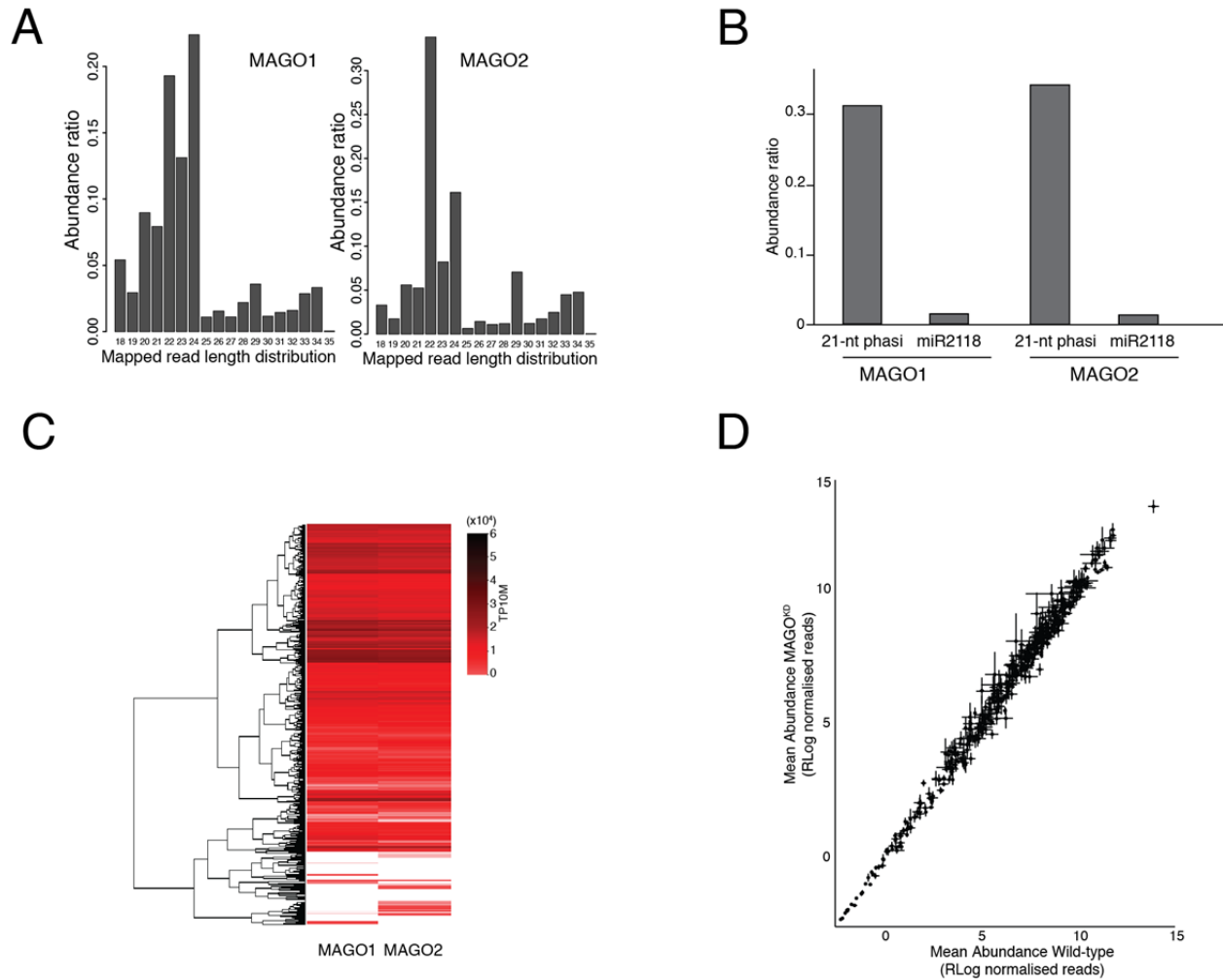
409
410
Fig. S3. Identification of small RNAs associated with MAGO1 and 2.
(A) Abundance of different sRNA classes in immunoprecipitated MAGO1 and 2 protein fractions determined by sRNA sequencing. n= 2 independent biological replicates. TP10M, transcript per ten million reads.
(B) Abundance of 21-nt phasiRNA and miR2118 trigger in MAGO1 and MAGO2 immunoprecipitated fractions.
(C) Abundance of different 21-nt phasiRNA classes bound to MAGO1 and 2.
(D) Mean abundance plot of 21-nt phasiRNA in pre-meiotic anthers from wild-type and MAGO^{KD} plants. Mean from 3 independent biological replicates; Black line, Standard Deviation.
420
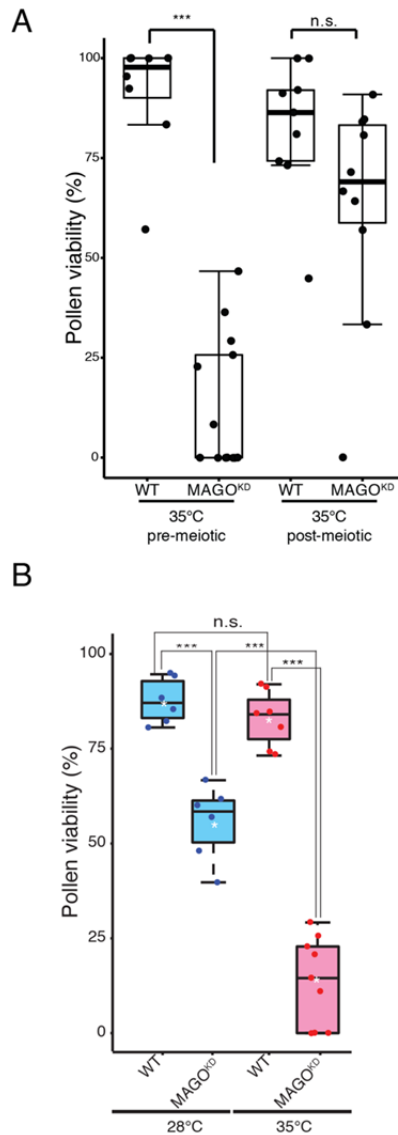
Fig. S4. MAGO1 and 2 are required before meiosis to sustain male fertility under heat stress. (A) Pollen viability in wild-type and MAGO[KD] plants grown under normal conditions (28ºC) and subjected to heat stress (72h/35ºC) before or after meiosis. n≥ 7 plants, more than 6 anthers per plant analysed. Differences between groups were determined by one-way ANOVA, ***p < 0.001; n.s. no-significant.
(B) Pollen viability in wild-type and MAGO[KD] plants under normal conditions (28ºC) and subjected to heat stress (35ºC) before meiosis (n= 10 plants; 10 anthers each). Differences between groups were determined by one-way ANOVA, ***p < 0.001; n.s. no-significant. Black line, median; White star, mean.
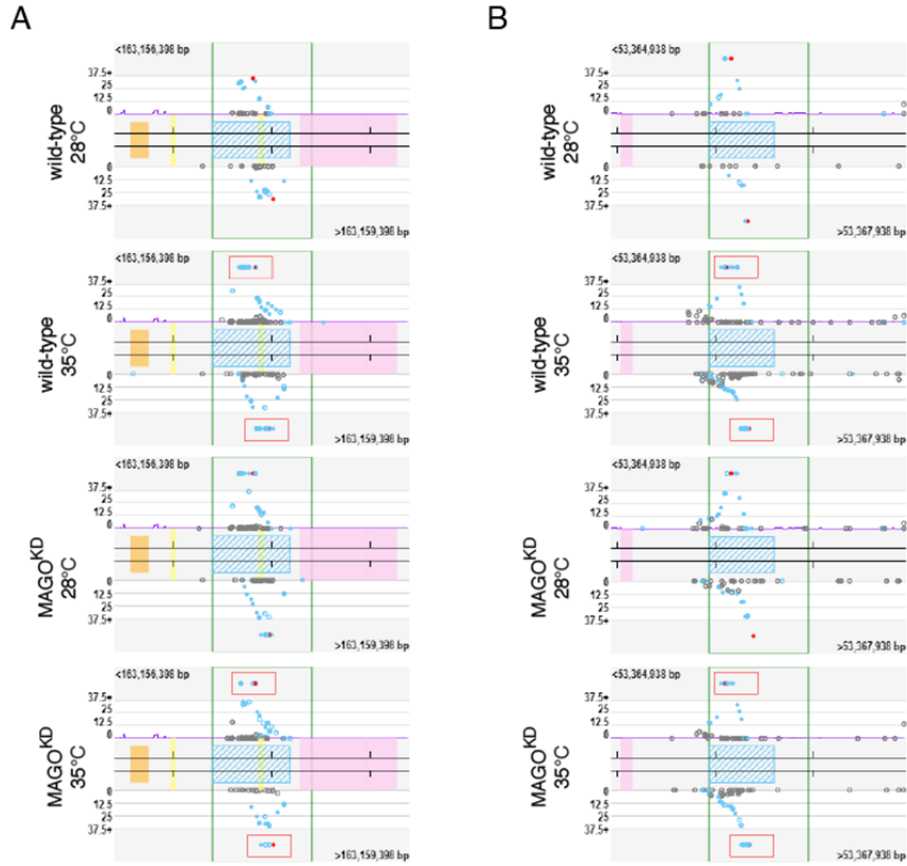
433
434
Fig. S5. Abundance of heat-induced phasiRNAs (Hphasi) from four PHAS loci in pre-meiotic anthers from wild-type and MAGO[KD] plants grown under normal conditions (28ºC) and subjected to heat stress (72h/35ºC) before or after meiosis.
(A) Abundance of 21-nt sRNAs in Hphasi_22 locus. (B) Abundance of 21-nt sRNAs in Hphasi_123 locus. n = 3 independent biological replicates. Red box, Hphasi generating region; Blue dots, common phasiRNAs; Red dot, unique phasiRNAs.
441

18
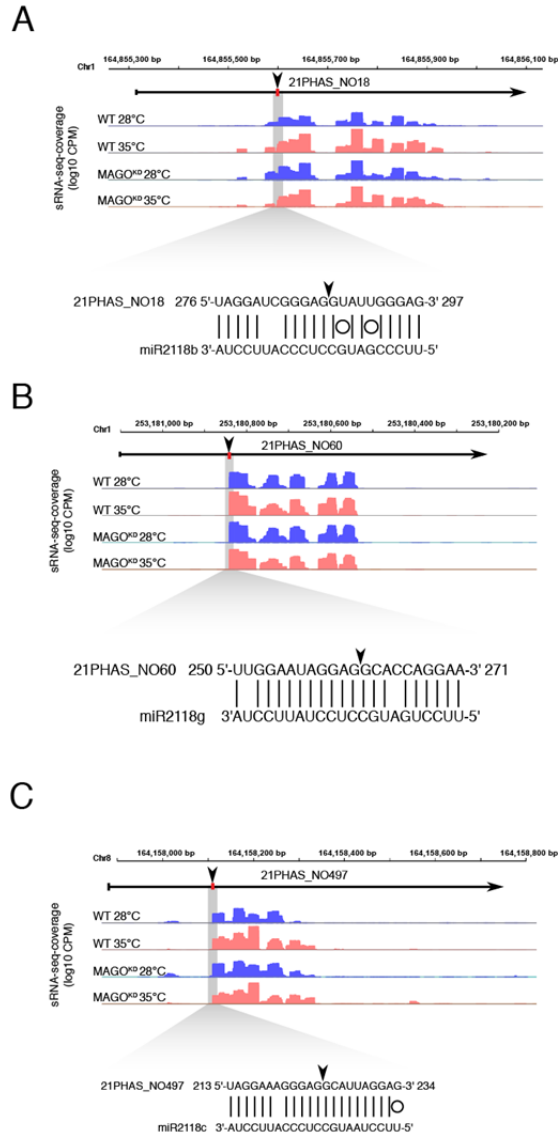
Fig. S6. Small-RNA-seq coverage of three 21-nt Hphasi generated from miR2118-slicing of different PHAS precursors in wild-type and MAGO[KD] plants grown under normal conditions (28ºC) and subjected to heat stress (72h/35ºC) before or after meiosis. n ≥ 3 independent biological replicates. Black arrowhead and red box indicate the location of predicted sites for miRNA-directed slicing remnants.
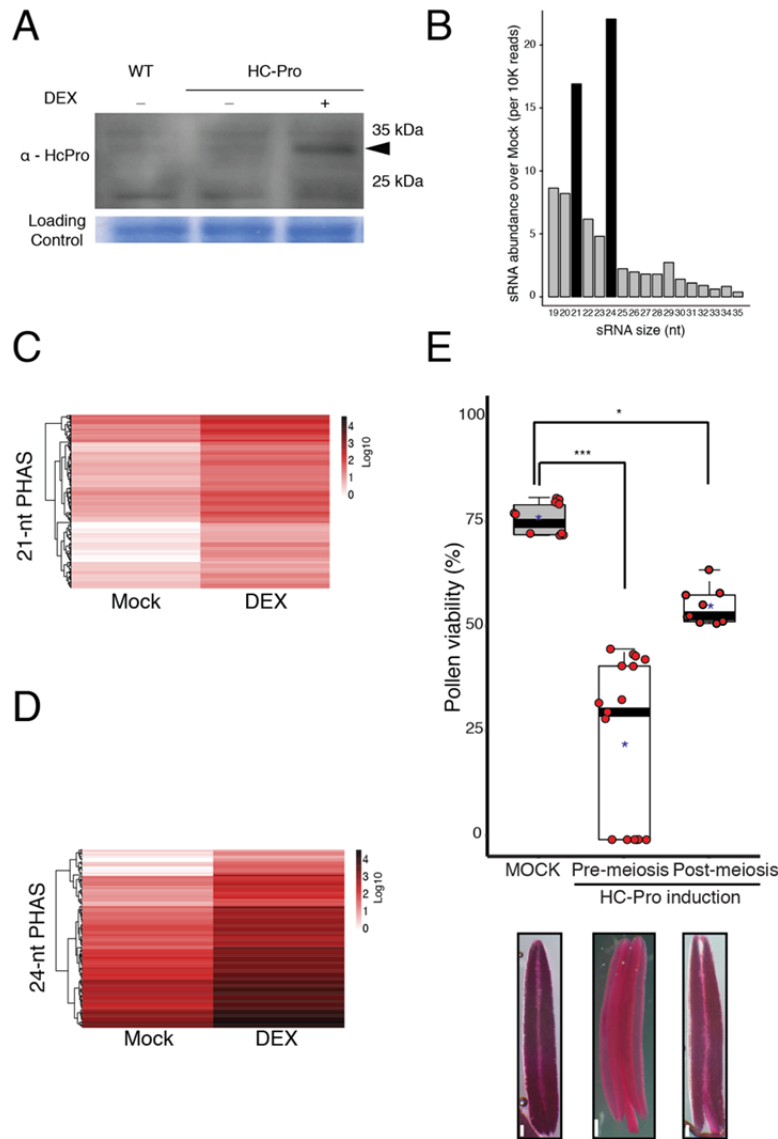
**Fig. S7. Controlled expression of HC-Pro enables the sequestration of small RNAs in maize anthers.**

(A) Western blot detection of HC-Pro accumulation in anthers of wild-type and HC-Pro plants treated with mock and 20 uM DEX.

(B) Relative abundance of sRNAs bound to HC-Pro in pre-meiotic anthers determined by immunoprecipitations coupled to sRNA sequencing. n= 2 independent biological replicates.

(C) Relative abundance of 21-nt phasiRNAs bound to HC-Pro in anthers and determined by immunoprecipitations coupled to sRNA sequencing. n= 2 independent biological replicates.

(D) Relative abundance of 24-nt phasiRNAs bound to HC-Pro in anthers and determined by immunoprecipitations coupled to sRNA sequencing. n= 2 independent biological replicates.

(H) Pollen density in anthers from two independent HDZIV6>>HC-Pro lines (n≥50 anthers; 10 plants each genotype). Differences between groups were determined by one-way ANOVA, ***p < 0.001. Black line, median; Red star, mean. Below, representative anthers after Alexander's Staining. Scale bars 100 μm.
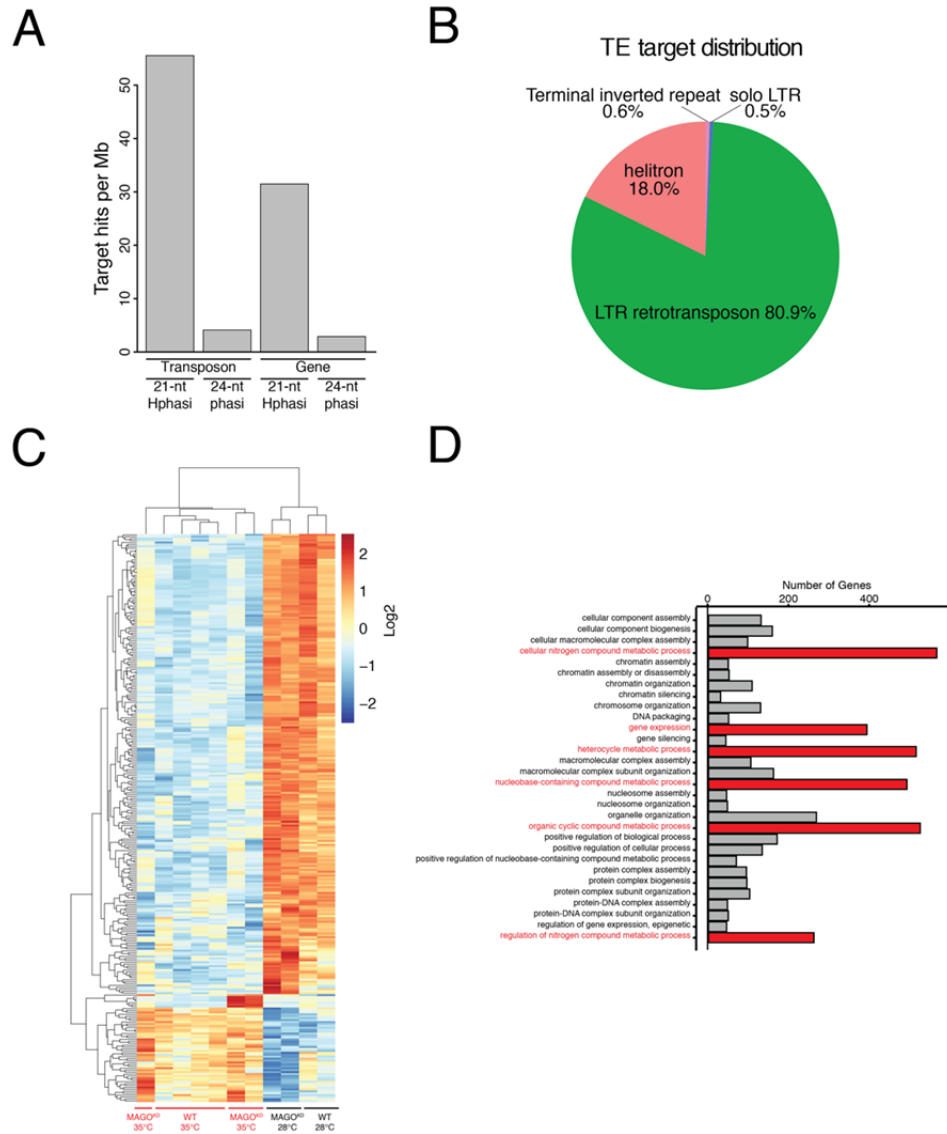
467
468
Fig. S8. Predicted targets of 21-nt Hphasi in the maize genome and impact of heat stress in gene expression.
(A) Distribution of 21-nt HphasiRNA and 24-nt phasiRNA targets according to their genomic location.
(B) Distribution of 21-nt HphasiRNA targets against annotated transposons.
(C) Heatmap showing transcriptional changes in pre-meiotic anthers caused by heat stress (72h/35ºC). n≥ 2 independent biological replicates.
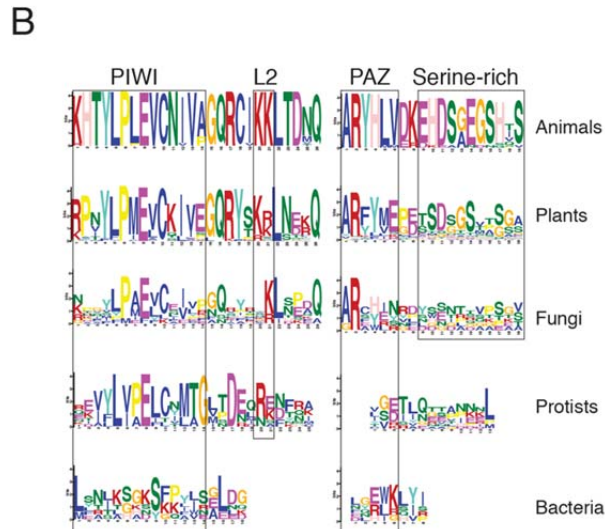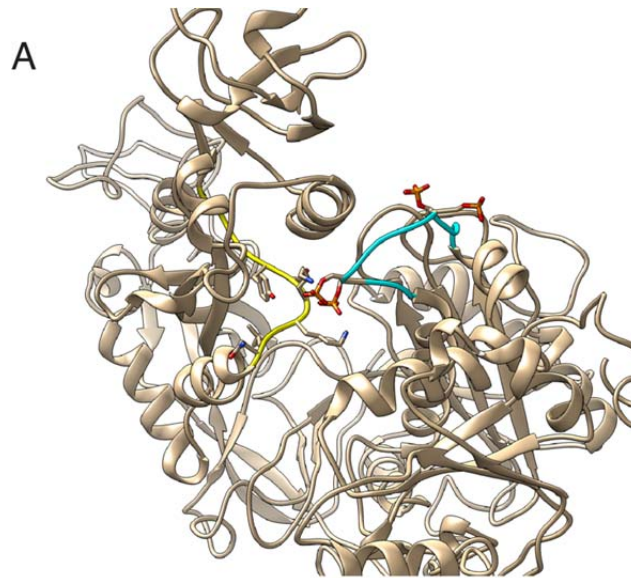(D) Gene Ontology (GO) analysis showing genes sets enriched within the differentially expressed categories.
478
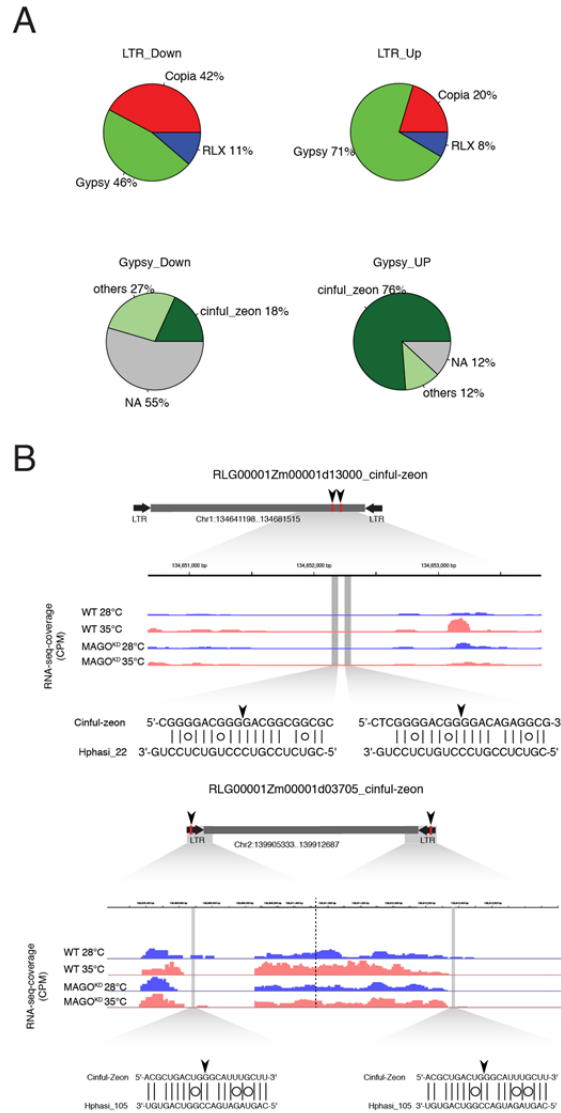
479
480
Fig. S9. Deregulation of retrotransposon in MAGO[KD] pre-meiotic anthers after exposure to heat stress.

(A) Frequency of Gypsy, Copia and RLX retrotransposons de-regulated in pre-meiotic anther of MAGO[KD] plants after exposure to a heat stress (35°C) and frequency of Cinful-zeon and other retrotransposons of the Gypsy-class de-regulated in pre-meiotic anther of MAGO[KD] plants after exposure to a heat stress (35°C).

(B) Coverage of RNA-seq of two different Gypsy-class retrotransposons targeted by Hphasi on anthers of wild-type and MAGO[KD] plants grown under normal conditions (28°C) and subjected to heat stress (72h/35°C) before meiosis. n≥ 3 independent biological replicates. Black arrowhead and red box indicate the location of predicted sites for miRNA-directed slicing remnants.

491

492
493
494 **Fig. S10. MAGO2 phosphopeptides identified by Mass Spectrometry analysis.**
495 (A) Schematic representation of MAGO2 with annotated serine and threonine residues found to
496 be hypo-phosphorylated in response to heat stress.
497 (B) Representative spectra of some MAGO2 phosphopeptides.
498

Fig. S11. Phosphorylation of conserved serine residues in MAGO2 and other catalytically active argonautes.
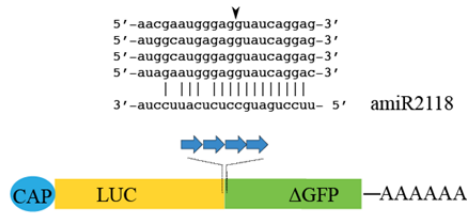
(A) Ribbon structure of the PIWI loop and L2 loop of MAGO2 showing the location of dynamically phosphorylated serine (S989, S990, S994 and S998) residues (red) in the PIWI loop (blue).

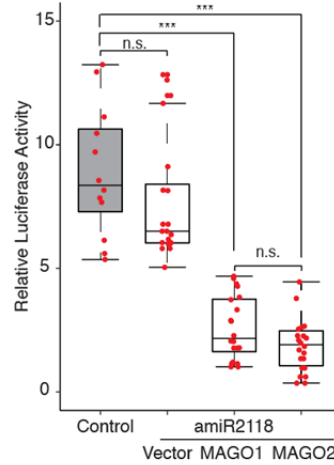(B) Amino acid residues conserved in four protein domains of catalytic argonautes.

A

```
5'-aacgaaugggaggguaucaggag-3'
5'-auggcaugagaggguaucaggag-3'
5'-auggcaugggaggguaucaggag-3'
5'-auagaaugggaggguaucaggac-3'
     |  ||| |||||||||||||
3'-auccuuacucuccguaguccuu- 5'     amiR2118
```

CAP LUC ΔGFP —AAAAAA

B

Relative Luciferase Activity

Control — amiR2118 — Vector MAGO1 MAGO2

C

Control   Vector

MAGO1   MAGO2
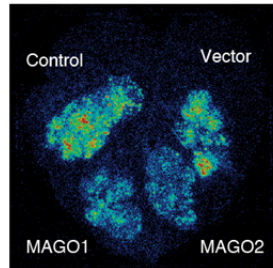
508
509
510 **Fig. S12. *In vivo* silencing activity of MAGO1/2 assessed in *Nicotiana benthamiana* using a**
511 **Luciferase-amiRNA reporter.**
512 (A) Schematic diagram of the MAGO sensor construct created by inserting four copies of a
513 miR2118 binding sites in the luciferase (LUC) gene. The predicted cleavage of the target
514 sequences by amiR2118 is indicated by an arrowhead.
515 (B) Relative luciferase activity in leaves transformed with a control LUC reporter and in
516 combination with amiR2118 and MAGO1 and 2. n= 10 independent biological replicates.
517 Differences between groups were determined by one-way ANOVA, ***$p < 0.001$; n.s. no-
518 significant. Black line, median; Red star, mean.
519 (C) Bioluminescent imaging of a representative leaf transformed with the LUC reporter used to
520 quantify MAGO1 and 2 silencing activity.
521

522

523

3. Tables:

525

Table S1. List of differentially expressed genes in pre-meiotic anthers from wild-type and MAGO^{KD} plants.

Table S2. List of predicted LTRs targeted by Hphasi.

Table S3. List of new transposon insertions determined by LTR-sequencing.

Table S4. List of differentially accumulated phosphopeptides from pre-meiotic anthers of wild-type plants exposed to heat stress.

Table S5. Conservation of phosphorylated serine and threonine residues in different Argonaute-like proteins.

Table S6. List of oligonucleotides and synthetic DNA constructs.

Table S7. Next-Generation-Sequencing library details.

536

4. Multimedia Files:

538

Supplementary Movie. Electrostatic potential distribution on the molecular surface of the central cleft of native and phosphorylated MAGO2. Charge: negative (red), positive (blue) and hydrophobic (grey) residues.

542

5. References:

1. Y. Li, G. Segal, Q. Wang, H. K. Dooner, Gene tagging with engineered Ds elements in maize. *Methods in Molecular Biology*, (2013).
2. D. R. McCarty *et al.*, Steady-state transposon mutagenesis in inbred maize. *Plant Journal*, (2005).
3. M. Samalova, B. Brzobohaty, I. Moore, pOp6/LhGR: A stringently regulated and highly responsive dexamethasone-inducible gene expression system for tobacco. *Plant Journal*, (2005).
4. T. Nakagawa *et al.*, Improved gateway binary vectors: High-performance vectors for creation of fusion constructs in transgenic analysis of plants. *Bioscience, Biotechnology and Biochemistry*, (2007).
5. A. G. Uren *et al.*, A high-throughput splinkerette-pcr method for the isolation and sequencing of retroviral insertion sites. *Nature Protocols*, (2009).
6. J. Chen, S. Dellaporta. (1994).
7. H. K. Dooner *et al.*, Spontaneous mutations in maize pollen are frequent in some lines and arise mainly from retrotranspositions and deletions. *Proc Natl Acad Sci U S A* **116**, 10734-10743 (2019).
8. J. Zhai *et al.*, Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proceedings of the National Academy of Sciences of the United States of America*, (2015).
9. D. Kim *et al.*, TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology*, (2013).
10. Y. Jin, O. H. Tam, E. Paniagua, M. Hammell, TEtranscripts: A package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics*, (2015).
11. M. A. Schon, M. J. Kellner, A. Plotnikova, F. Hofmann, M. D. Nodine, NanoPARE: Parallel analysis of RNA 5′ ends from low-input RNA. *Genome Research*, (2018).
12. J. R. Wiśniewski, A. Zougman, N. Nagaraj, M. Mann, Universal sample preparation method for proteome analysis. *Nature Methods*, (2009).
13. T. E. Thingholm, T. J. D. Jørgensen, O. N. Jensen, M. R. Larsen, Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nature Protocols*, (2006).
14. A. Šali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, (1993).
15. K. Nakanishi, D. E. Weinberg, D. P. Bartel, D. J. Patel, Structure of yeast Argonaute with guide RNA. *Nature*, (2012).
16. N. T. Schirle, J. Sheu-Gruttadauria, S. D. Chandradoss, C. Joo, I. J. MacRae, Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. *eLife*, (2015).
17. M.-y. Shen, A. Sali, Statistical potential for assessment and prediction of protein structures. *Protein Science*, (2006).
18. D. A. Case, Amber 18. *University of California, San Francisco*, (2018).

27

586   19.   J. M. Word, S. C. Lovell, J. S. Richardson, D. C. Richardson, Asparagine and glutamine:
587   Using hydrogen atom contacts in the choice of side-chain amide orientation. *Journal of*
588   *Molecular Biology*, (1999).
589   20.   J. Mongan, C. Simmerling, J. A. McCammon, D. A. Case, A. Onufriev, Generalized born
590   model with a simple, robust molecular volume correction. *Journal of Chemical Theory*
591   *and Computation*, (2007).
592   21.   H. Nguyen, D. R. Roe, C. Simmerling, Improved generalized born solvent model
593   parameters for protein simulations. *Journal of Chemical Theory and Computation*,
594   (2013).
595   22.   V. Hornak *et al.* (2006).
596   23.   J. A. Maier *et al.*, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone
597   Parameters from ff99SB. *Journal of Chemical Theory and Computation*, (2015).
598   24.   J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Numerical integration of the cartesian
599   equations of motion of a system with constraints: molecular dynamics of n-alkanes.
600   *Journal of Computational Physics*, (1977).
601   25.   K. Retzer *et al.*, Evolutionary conserved cysteines function as cis-acting regulators of
602   arabidopsis PIN-FORMED 2 distribution. *International Journal of Molecular Sciences*,
603   (2017).
604   26.   W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of
605   simple potential functions for simulating liquid water. *The Journal of Chemical Physics*,
606   (1983).
607   27.   I. S. Joung, T. E. Cheatham, Determination of alkali and halide monovalent ion
608   parameters for use in explicitly solvated biomolecular simulations. *Journal of Physical*
609   *Chemistry B*, (2008).
610   28.   S. Joung, T. E. Cheatham, Molecular dynamics simulations of the dynamic and energetic
611   properties of alkali and halide ions using water-model-specific ion parameters. *Journal*
612   *of Physical Chemistry B*, (2009).
613   29.   N. Homeyer, A. H. C. Horn, H. Lanig, H. Sticht, AMBER force-field parameters for
614   phosphorylated amino acids in different protonation states: Phosphoserine,
615   phosphothreonine, phosphotyrosine, and phosphohistidine. *Journal of Molecular*
616   *Modeling*, (2006).
617   30.   T. Steinbrecher, J. Latzer, D. A. Case, Revised AMBER parameters for bioorganic
618   phosphates. *Journal of Chemical Theory and Computation*, (2012).
619   31.   X. Ye, J. Wang, R. Luo, A revised density function for molecular surface calculation in
620   continuum solvent models. *Journal of Chemical Theory and Computation*, (2010).
621