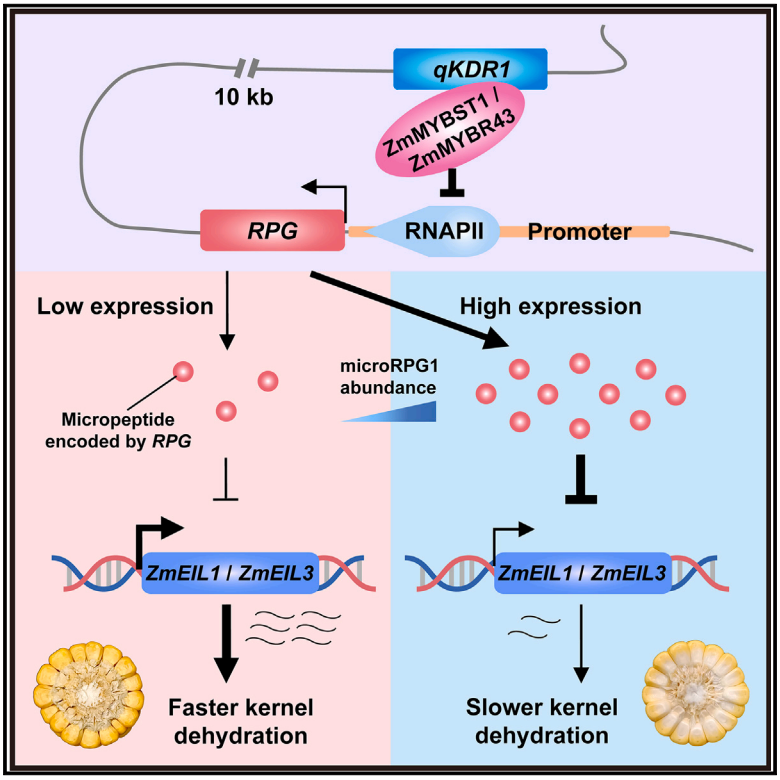
Cell

A *Zea* genus-specific micropeptide controls kernel dehydration in maize

Graphical abstract



Highlights

- *qKDR1* is a major QTL controlling kernel dehydration, which is a key production trait
- *qKDR1* regulates *RPG*, which encodes a 31 amino acid *Zea* genus-specific micropeptide
- *RPG* originated *de novo* from a non-coding sequence
- microRPG1 controls kernel dehydration by precisely regulating ethylene gene expression

Authors

Yanhui Yu, Wenqiang Li, Yuanfang Liu, ..., Alisdair R. Fernie, David Jackson, Jianbing Yan

Correspondence

yjianbing@mail.hzau.edu.cn

In brief

Kernel dehydration rate is a key determinant of moisture content that affects mechanized harvesting and kernel quality in maize. The microRPG1, a *Zea* genus-specific micropeptide that originated *de novo* from a non-coding sequence, controls kernel dehydration rate by precisely regulating the expression of genes in the ethylene signaling pathway in the kernels after filling.

Yu et al., 2025, Cell *188*, 1–16
January 9, 2025
https://doi.org/10.1016/j.cell.2024.10.030

CellPress

Cell

CellPress
OPEN ACCESS

Article

A *Zea* genus-specific micropeptide controls kernel dehydration in maize

Yanhui Yu,[1,2,8] Wenqiang Li,[1,2,8] Yuanfang Liu,[1,2,8] Yanjun Liu,[1,2] Qinzhi Zhang,[1] Yidan Ouyang,[1] Wenya Ding,[1] Yu Xue,[3] Yilin Zou,[1] Junjun Yan,[1] Anqiang Jia,[4] Jiali Yan,[1] Xinfei Hao,[1] Yujie Gou,[3] Zhaowei Zhai,[1] Longyu Liu,[1,2] Yang Zheng,[1,2] Bao Zhang,[1,2] Jieting Xu,[5] Ning Yang,[1,2] Yingjie Xiao,[1,2] Lin Zhuo,[1,2] Zhibing Lai,[1,2] Ping Yin,[1,2] Hai-Jun Liu,[4] Alisdair R. Fernie,[6] David Jackson,[1,7] and Jianbing Yan[1,2,4,9,*]

[1]National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, Hubei 430070, China
[2]Hubei Hongshan Laboratory, Wuhan, Hubei 430070, China
[3]Key Laboratory of Molecular Biophysics of Ministry of Education, Hubei Bioinformatics and Molecular Imaging Key Laboratory, Center for Artificial Intelligence Biology, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China
[4]Yazhouwan National Laboratory, Sanya 572024, China
[5]WIMI Biotechnology Co., Ltd., Changzhou, Jiangsu 213000, China
[6]Department of Root Biology and Symbiosis, Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm 14476, Germany
[7]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
[8]These authors contributed equally
[9]Lead contact
*Correspondence: yjianbing@mail.hzau.edu.cn
https://doi.org/10.1016/j.cell.2024.10.030

SUMMARY

Kernel dehydration rate (KDR) is a crucial production trait that affects mechanized harvesting and kernel quality in maize; however, the underlying mechanisms remain unclear. Here, we identified a quantitative trait locus (QTL), *qKDR1*, as a non-coding sequence that regulates the expression of *qKDR1 REGULATED PEPTIDE GENE* (*RPG*). *RPG* encodes a 31 amino acid micropeptide, microRPG1, which controls KDR by precisely modulating the expression of two genes, *ZmETHYLENE-INSENSITIVE3-like 1* and *3*, in the ethylene signaling pathway in the kernels after filling. microRPG1 is a *Zea* genus-specific micropeptide and originated *de novo* from a non-coding sequence. Knockouts of microRPG1 result in faster KDR in maize. By contrast, overexpression or exogenous application of the micropeptide shows the opposite effect both in maize and *Arabidopsis*. Our findings reveal the molecular mechanism of microRPG1 in kernel dehydration and provide an important tool for future crop breeding.

INTRODUCTION

Maize is one of the most important crops in the world, with an annual global production of 1.21 billion tons in 2021 (https://www.fao.org/). Mechanized harvesting of maize kernels is an effective way to solve the continuous increase of labor costs and improve production efficiency. However, mechanized harvesting has not yet been achieved in China and many other countries, and the mechanized harvesting area of maize kernel is less than 5% in China, partly due to a lack of suitable maize varieties.[1–3] Kernel dehydration rate (KDR) is a key determinant of moisture content at harvest that affects mechanized harvesting and kernel quality.[4–6] A high moisture content at harvest seriously limits harvesting efficiency and increases the cost of drying and storage.[7,8] This represents a considerable problem as most maize varieties in China have high grain water content at harvest, ranging from 30% to 40%, but the moisture content suitable for mechanized harvesting is from 15% to 25%, seriously limiting the widespread application of mechanical harvesting.[1,9–12] Farmers usually delay harvest to reduce moisture content, which leads to a delay in the planting time of the next crop and greatly affects the annual yield of the crops, especially in the Huang-Huai-Hai region of China with a Wheat-Maize Rotation System.[1,4,7] High moisture content can also lead to kernel breakage during mechanical harvesting,[6,13,14] and can cause kernel mold, ear rot, and ear sprouting under warm and humid environmental conditions.[10,15–17] Thus, increasing KDR and reducing kernel moisture content at harvest is essential and has become a major goal of modern maize breeding.[18,19]

Many studies have identified quantitative trait loci (QTLs) controlling KDR in maize over the past two decades.[1,4,6,10,13,18,20–23] However, very few QTLs have been cloned,[21] rendering the molecular mechanism of maize KDR unsolved. Furthermore, the genes controlling KDR are also rarely reported in plants,[24] and the underlying mechanisms are largely unknown.

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

Recent advances have shown that some so-called long non-coding RNAs (lncRNAs) have small open reading frames (sORFs) less than 100 amino acids in length.[25–27] These sORFs can be translated into small proteins, often referred to as micropeptides.[28–30] Many functional peptides translated from lncRNAs have been studied in humans, leading to molecular insights into their functions.[30–34] While a few functional micropeptides have been identified in plants,[35–40] relatively little is known about their roles.

Furthermore, the origin of new genes can generate and maintain remarkable phenotypic diversity.[41–43] These new genes can originate *de novo* from non-coding sequences and affect adaptation, morphological diversity, and speciation of organisms.[43–47] However, new genes in plants that are both novel and phenotypically impactful remain largely uncharacterized.[48]

In this study, we fine-mapped and cloned a QTL for KDR in maize, *Kernel Dehydration Rate1* (*qKDR1*), to a non-coding intergenic DNA sequence. *qKDR1* regulates the expression of a 31 aa micropeptide encoding gene, *qKDR1 Regulated Peptide Gene* (*RPG*), located ~10 kilobases upstream of *qKDR1*. Two MYB-related transcription factors, *ZmMYBST1* and *ZmMYBR43*, bind to the *qKDR1* region to repress *RPG* expression. microRPG1 (micropeptide of *RPG* ORF1) is derived from a sORF generated by a single nucleotide mutation at the base of the genus *Zea* that created a start codon. The microRPG1 micropeptide controls KDR by regulating the expression of genes in the ethylene signaling pathway *ZmETHYLENE-INSENSITIVE3-like 1* and *3* (*ZmEIL1* and *3*). Modulation of microRPG1 expression or exogenous application of the peptide in the distantly related dicot model *Arabidopsis* impacts KDR. Our findings shed light on the molecular mechanism underlying KDR and provide useful insights for future crop breeding, underscore the significant role of micropeptides, and present an excellent model for *de novo* gene evolution.

RESULTS

qKDR1 is the causal locus for KDR

We mapped four QTLs for KDR in a maize recombinant inbred line (RIL) population developed from a cross between the maize inbred lines K22 and DAN340[49,50] (Figure 1A). The QTL with the largest effect, *qKDR1* on the long arm of chromosome 1 and explaining 9.8% of the phenotypic effect, was selected for positional cloning. To examine the allelic effects, we developed a pair of *qKDR1* near isogenic lines (NILs) (NIL^{DAN340} and NIL^{K22}) (Figure 1B) from a heterogeneous inbred family library.[54] Compared with NIL^{DAN340}, NIL^{K22} had a more rapid KDR with up to 1.8% reduction in moisture content at harvest, but no obvious difference in other agronomic traits (Figure 1B; Table S1).

The QTL was further narrowed down to a 1,417 base pair (bp) intergenic non-coding region according to the maize B73 reference sequence version 4.0 (Figure 1C). NIL^{DAN340} had a 6,181-bp insertion in this region compared with NIL^{K22} (Figure 1C; Table S2). To confirm the function of *qKDR1*, we deleted the target sequence using CRISPR-Cas9 in two different maize inbred lines, Zheng58 (similar to NIL^{K22}) and B104 (similar to NIL^{DAN340}) (Figure 1D). Both knockout lines had slower KDR compared with wild-type siblings (Figures 1E and 1F), with few effects on other agronomic traits in two different environments (Table S3). We next surveyed the natural variation of the 6,181-bp QTL region insertion in a diverse maize inbred line population,[55] and identified two additional insertions or deletions (indels) in this region by polymerase chain reaction (PCR). Finally, five different haplotypes were identified (Figure 1G), and four of them, Hap2–Hap5 (with indel), had a lower KDR compared with Hap1 (Figure 1H). The collective data all show statistically significant differences based on different methods (Figure 1I). These results indicate that the 1,417-bp sequence of *qKDR1* is a key determinant for KDR variation, as knocking out this sequence resulted in a reduced KDR in maize.

qKDR1 acts as a silencer

An intriguing question is, as a non-coding sequence, how does the 1,417-bp sequence function. To answer this question, we performed transient transcriptional activity assays in maize protoplasts. The *qKDR1* sequences from the two NILs, $qKDR1^{2D}$ (a *qKDR1* fragment of NIL^{DAN340}) and $qKDR1^{2K}$ (a *qKDR1* fragment of NIL^{K22}), were independently fused downstream of the β-glucuronidase (GUS) gene driven by a minimal *Ubiquitin* promoter (mpUbi) (named pUbi2D and pUbi2K) and using the firefly luciferase (LUC) gene as an internal control (Figures 2A and 2B). Compared with GUS activity of the mpUbi vector (pUbiVector), the activities of pUbi2D and pUbi2K were both significantly reduced. Moreover, the GUS activity of pUbi2K was lower than pUbi2D, demonstrating that GUS expression was reduced by the presence of both *qKDR1* alleles, and the inhibitory effect of $qKDR1^{2K}$ was greater than that of $qKDR1^{2D}$ (Figures 2A–2C). This suggests that the different *qKDR1* alleles act to repress target gene expression. We next fused a series of *qKDR1* fragments with different lengths downstream of the GUS gene (named pUbi3D and pUbi3K-pUbi5K) (Figures 2A and 2B). Transient transcriptional activity assays indicated that all of them displayed a similar repressive activity to pUbi2K (Figures 2A–2C), but pUbi2D led to higher GUS activity. These results suggest that a 369-bp sequence of *qKDR1* is the core silencer sequence and that the 6,181-bp insertion in NIL^{DAN340} reduces the silencing activity.

RPG is the *qKDR1* target gene

To identify the gene(s) regulated by *qKDR1*, we performed RNA sequencing (RNA-seq) on four NILs harboring different *qKDR1* genotypes (Figure 2D). We rationalized that genes that were differentially expressed according to the genotypes of *qKDR1* would be candidate target genes (Figures 2E and 2F). We identified 17 genes, of which three were upregulated and 14 were downregulated (Figures 2E and 2F; Table S4). Among these genes, only one, which we named *RPG*, was close to *qKDR1*, being located ~10-kb upstream (Figure 1C). Consistent with the KDR phenotype, *RPG* was highly expressed in maize kernels after 30 days after pollination (DAP) (Figures S1D, S4A, and S4B). Real-time quantitative PCRs (qPCRs) revealed that *RPG* expression was lower in NIL^{K22} than in NIL^{DAN340} in the late stage of maize kernel maturation (Figure 2G). In the diverse inbred line population, we also found that the *RPG* expression in Hap2 and Hap5 (with indels in *qKDR1*) was significantly higher than

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

CellPress
OPEN ACCESS

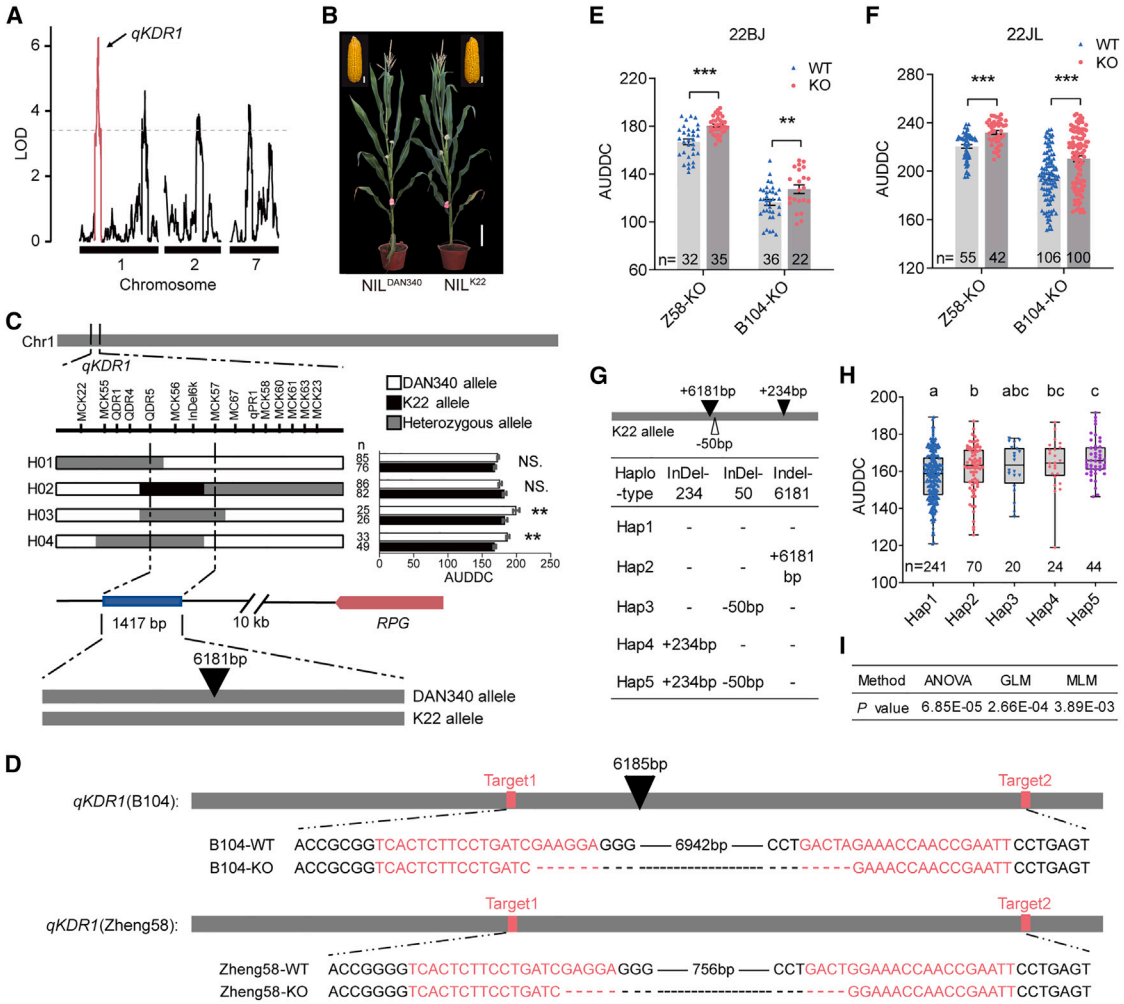


Figure 1. *qKDR1* is the causal locus for kernel dehydration rate

(A) QTL mapping for KDR in maize in the DAN340/K22 RIL population. LOD, logarithm of odds. *qKDR1* is the largest KDR QTL. The dashed gray line at LOD of 3.3 indicates the threshold LOD value for putative QTLs. See Zenodo: https://doi.org/10.5281/zenodo.13924714 for experimental data.

(B) Gross morphologies of NIL^{DAN340} and NIL^{K22}. Scale bars, 20 cm for the plant, 2 cm for the ear.

(C) *qKDR1* was mapped on chromosome 1 and was fine-mapped to a 1,417-bp non-coding region (corresponding to B73 reference sequence version 4.0, Chr1:20007756–20009147) flanked by markers QDR5 and MCK57. The gray vertical dashed lines show the boundaries of the 1,417-bp region of *qKDR1*. The graphical genotypes of the four critical recombinant types are showed on the left, and the bar graphs on the right show the phenotypes of their self-pollinated homozygous progenies. White, black, and gray segments represent alleles homozygous for DNA340, alleles homozygous for K22, and heterozygous alleles, respectively. The pink box represents the 2,013-bp full-length mRNA of *RPG*. The 1,417-bp region of *qKDR1* (blue box) is located ~10-kb downstream of the transcription termination site (TTS) of *RPG*. The black triangle represents the 6,181-bp insertion in NIL^{DAN340} compared with NIL^{K22}. See Zenodo: https://doi.org/10.5281/zenodo.13924714 for experimental data.

(D) Sequences of CRISPR-Cas9 knockout lines with deletions in the target sites (Z58-KO and B104-KO). The target sites are highlighted in pink. The dashed lines indicate deletions.

(E and F) Knockout of *qKDR1* decelerated KDR in maize in field trials performed in Beijing in 2022 (22BJ, E) and Jilin in 2022 (22JL, F), respectively.

(G) Diverse maize inbred line population were divided into five haplotypes by different *qKDR1* genotypes. Hap1, without indel in the *qKDR1* region, has similar genotypes to NIL^{K22}. Hap2, with 6,181-bp insertion, has similar genotypes to NIL^{DAN340}.

(H) The five haplotypes with different *qKDR1* genotypes in the diverse maize inbred line population showed significant phenotypic differences by haplotype analysis. Different letters at top of columns indicate significant differences at $p < 0.05$ (one-way ANOVA followed by Fisher's LSD [least significant difference] multiple-comparison test).

(I) The five haplotypes with different *qKDR1* genotypes in the diverse maize inbred line population showed significant phenotypic differences by using GLM (with population structure) and MLM (with population structure and kinship). Association studies were performed with general linear model (GLM) or mixed linear model (MLM) implemented by TASSEL (trait analysis by association, evolution, and linkage).[51–53] ANOVA, one-way ANOVA in (H).

Data are represented as mean ± SEM. $^{**}p < 0.01$, $^{***}p < 0.001$ (Student's t test).

n is the sample size, a sample represents the phenotype from a plant (C, E, and F) or from a plot (average of multiple plants of a row) (H).

See also Tables S1–S3.

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

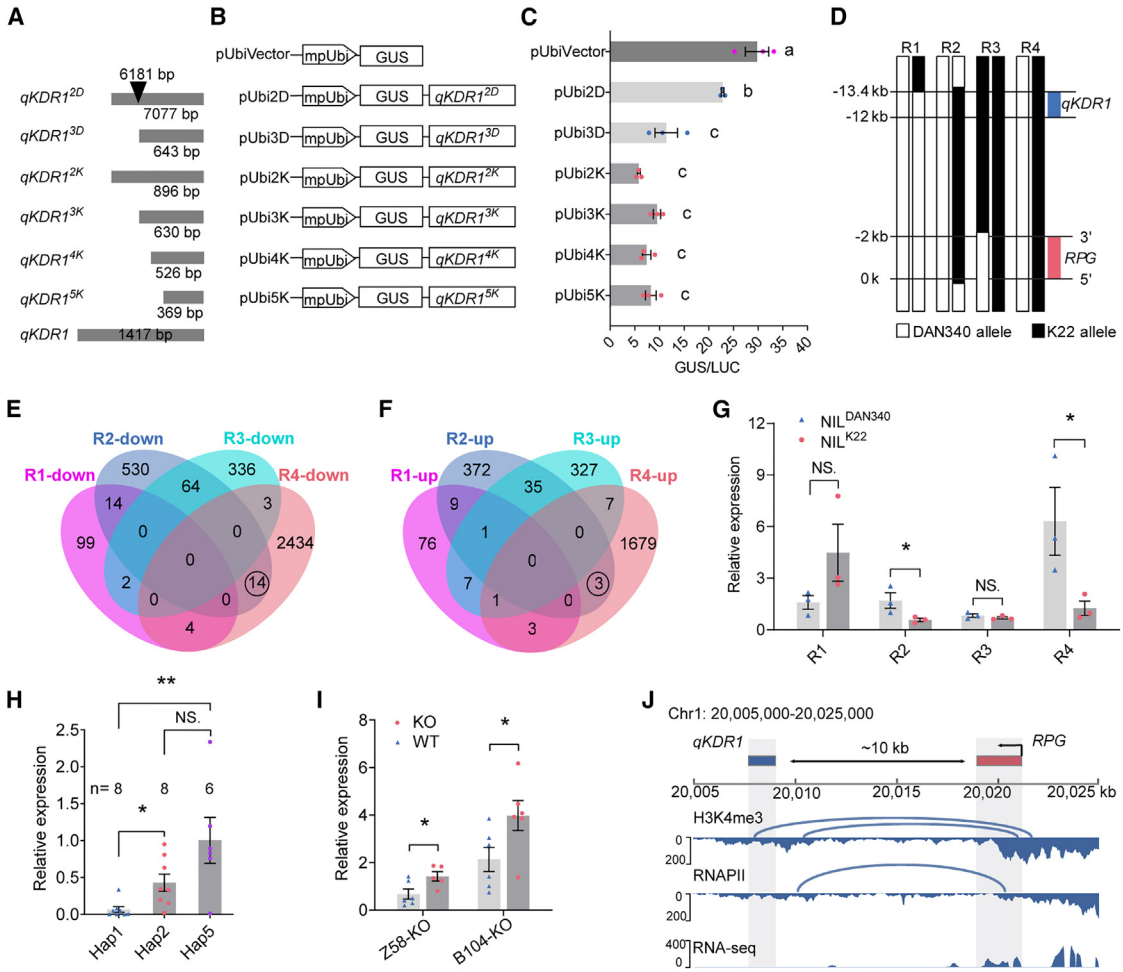


Figure 2. *qKDR1* acts as a silencer and regulates kernel dehydration rate through *RPG*

(A) *qKDR1* fragments used to generate constructs.

(B) Schematic diagram of constructs used for the transient transcriptional activity assays.

(C) Ratio of GUS/LUC activity in transient transcriptional activity assays performed in maize protoplasts. Data are represented as mean ± SEM ($n = 3$). Different letters at top of columns indicate significant differences at $p < 0.05$ (one-way ANOVA followed by Fisher's LSD multiple-comparison test).

(D) Schematic diagram of four NILs with different *qKDR1* genotypes. The numbers on the left represent the position relative to the transcription start site (TSS) of *RPG*.

(E and F) Differential expression analysis of four NILs (RNA-seq). Downregulated genes are shown in (E), and upregulated genes are shown in (F). The black circles represent candidate genes.

(G) *RPG* expression in 40 days after pollination (DAP) kernels of four NILs with different *qKDR1* genotypes. The expression levels of *RPG* were quantified using qPCR and normalized to maize *ACTIN* ($n = 3$).

(H) *RPG* expression in 45 DAP kernels of three haplotypes with different *qKDR1* genotypes. The expression levels of *RPG* were quantified using qPCR and normalized to maize *ACTIN* ($n \geq 6$).

(I) *RPG* expression in 45 DAP kernels of two *qKDR1* knockout lines. The expression levels of *RPG* were quantified using qPCR and normalized to maize *ACTIN* ($n \geq 5$).

(J) Chromatin interaction analysis by paired-end-tag sequencing (ChIA-PET) between *qKDR1* and *RPG*. Chromatin interactions inferred from H3K4me3 and RNA polymerase II (RNAPII) occupancy were detected between the *RPG* and *qKDR1* loci. The bottom panel shows gene expression (RNA-seq). Light gray columns represent the locations of *qKDR1* and *RPG*. Data from Peng et al.[56]

Data are represented as mean ± SEM. $^*p < 0.05$, $^{**}p < 0.01$, NS, not significant (Student's t test).

n is the biological replicate size, a biological replicate is an independent maize protoplast transformation experiment (C) or is from a plant sample (maize kernels at the late maturation stage) (G–I).

See also Figure S1 and Tables S4 and S5.

that in Hap1 (without indel in *qKDR1*) (Figure 2H). As expected, *RPG* expression was also higher in both of the *qKDR1* knockout lines (Figure 2I). Furthermore, chromatin interactions inferred from H3K4me3 and RNA polymerase II (RNAPII) occupancy[56] were also detected between *RPG* and the *qKDR1* locus (Figure 2J).

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

To further confirm that *RPG* is regulated by *qKDR1*, we fused the 883-bp *RPG* promoter upstream of the firefly LUC gene (named pRPGVector), and the Renilla luciferase (REN) gene was used as an internal control. The *qKDR1* fragments of different lengths from NIL^{DAN340} or NIL^{K22} were also fused downstream of the LUC gene (named pRPG3D and pRPG3K-pRPG5K) (Figures S1A and S1B). Compared with the *RPG* promoter-driven LUC activity, all of them displayed lower LUC activity (Figures S1A–S1C). This result is consistent with transient transcriptional activity assays of mpUbi-driven GUS activity (Figures 2A–2C). These results thus collectively suggest that *qKDR1* functions as a repressor of *RPG* expression, and the 369-bp region of *qKDR1* is sufficient for silencing activity.

Published chromatin immunoprecipitation sequencing (ChIP-seq) data[57] showed that 43 transcription factors can bind to the *qKDR1* region, two of which are annotated as transcriptional repressors (Table S5, *Zm00001d000184*, MYB-related-transcription factor 31, named *ZmMYBST1* and *Zm00001d029875*, MYB-related-transcription factor 43, named *ZmMYBR43*), consistent with the role of *qKDR1* acting as a silencer. *ZmMYBST1* and *ZmMYBR43* were expressed in a similar pattern as *RPG* with sequence similarity of 96% (Figure S1D),[58,59] suggesting that they may affect *RPG* expression. mRNA *in situ* hybridization was performed to confirm the expression patterns of *ZmMYBST1*, *ZmMYBR43*, and *RPG*, and the results showed that they were all expressed in the embryo, aleurone layer, and placento-chalazal region (Figure S1E). To evaluate the effects of *ZmMYBST1* and *ZmMYBR43* on the *qKDR1*-regulated expression of *RPG*, we performed transient transcriptional activity assays, in which the coding sequence of *ZmMYBST1* and *ZmMYBR43* driven by the *CaMV 35S* promoter were used as effectors, and pRPG3D and pRPG3K were used as the reporters, respectively (Figure S1F). Compared with the control, co-expression with *ZmMYBST1* or *ZmMYBR43* resulted in strongly repressed LUC activities of the two reporters (Figure S1G). Additionally, double mutants of *ZmMYBST1* and *ZmMYBR43* generated by CRISPR-Cas9 significantly decelerated KDR compared with wild-type siblings (Figure S1H). These results suggest that ZmMYBST1 and ZmMYBR43 target the *qKDR1* region to repress *RPG* expression *in vivo*.

RPG encodes a functional micropeptide

RPG was not annotated as a gene in maize (B73 version 4.0). However, high-depth RNA-seq reads mapped to and were presumably transcribed by the *RPG* region (Figure 3A). We thus performed rapid-amplification of cDNA ends (RACE), including both 5′-RACE and 3′-RACE, and found two full-length *RPG* transcripts, named T01 and T02, with lengths of 2,013 and 1,723 bp in NIL^{K22} (Figure 3B; Table S2). Next, to identify possible protein(s) encoded by *RPG*, we screened putative ORFs of the two transcripts using ORF finder (https://www.ncbi.nlm.nih.gov/orffinder) and identified twelve putative ORFs (Figure 3C), with the largest encoding a micropeptide of 58 amino acids. Small RNA-seq (sRNA-seq) revealed that *RPG* likely does not function by producing sRNAs (Figure 3D). However, ribosome profiling sequencing (Ribo-seq) revealed that *RPG* mRNA was ribosome bound in the ORF1, ORF2, and ORF3 regions, suggesting that it encodes one or more micropeptides (Figure 3E).

To assess whether *RPG* functions by encoding micropeptides rather than a lncRNA, we overexpressed the full-length *RPG* or a mutant *RPG* with all 12 ATG start codons mutated to CTG (Figure S2E). Overexpression of wild-type *RPG* resulted in a slower KDR (Figures S2A–S2D), whereas the non-translatable mutant *RPG* did not alter the KDR (Figures S2E–S2H). These results imply that *RPG* functions by encoding micropeptides rather than as a lncRNA.

We found two natural deletions in the *RPG* region in a diverse maize inbred line population.[55] These removed most of the ORF2 and ORF3 sequences, and the different haplotypes had no obvious KDR differences (Figures S2I–S2L). We also knocked out ORF2 and ORF3 by a 40-bp deletion induced by CRISPR-Cas9 (Figure S2K), and these lines had similar KDR compared with wild-type siblings (Figures S2M and S2N). By contrast, three frameshift mutations in ORF1 with 1-, 2-, or 4-bp deletions (Figures 3F and 3G) all led to a faster KDR (Figures 3H–3K). These findings indicate that ORF1 encodes the functional *RPG* micropeptide. We next evaluated the effect of ORF1 on the kernel moisture content across different environments. At harvest, the kernel moisture content of ORF1 knockout lines decreased between 1.8% and 17.0% (average of 7.1%) in different environments compared with wild type (Figures 3L–3O). Moreover, the microRPG1 knockout lines had no obvious effects on other agronomic traits in three different environments (Table S6). These results suggest that editing of *RPG* ORF1 has potential for breeding maize with higher KDRs.

ORF1 encodes a micropeptide

To further investigate the influence of ORF1 on the KDR phenotype, we expressed ORF1 and ORF1m (in which the ORF1 start codon was mutated to CTG) fused with FLAG tags under the control of the *Ubiquitin* promoter, in maize inbred line B104 (Figure 4A). The wild-type ORF1 construct, which expresses the ORF1 micropeptide, significantly decelerated KDR (Figures 4B, 4C, S3A, and S3B). By contrast, expression of the ORF1m construct did not alter the KDR (Figures 4D, 4E, S3C, and S3D). These results provide additional evidence that the ORF1 micropeptide product of the *RPG* gene affects maize KDR.

To further confirm that the ORF1 was translated in maize, we fused a GFPm ORF, lacking a start codon to the C terminus of ORF1, and driven by the 883-bp *RPG* promoter in protoplasts (Figure 4F). Indeed, we observed fluorescence of the ORF1-GFPm fusion protein in transformed cells (Figure 4G). However, mutation of the ORF1 start codon from ATG to ATT abolished the expression of the ORF1-GFPm fusion protein (Figure 4G). We used anti-GFP antibodies in western blotting to confirm expression of the ORF1-GFPm fusion protein, and expression was absent when the ORF1 ATG was mutated (Figure 4H). Similar results were obtained when we used the *CaMV 35S* promoter in place of the *RPG* promoter (Figures S3E and S3F).

We next asked if we could detect the endogenously produced *RPG* ORF1 micropeptide. We probed western blots of total plant protein and different concentrations of the synthetic ORF1 micropeptide (ORF1p) using an anti-ORF1 antibody. Indeed, we detected the ORF1 micropeptide of the expected size in protein

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

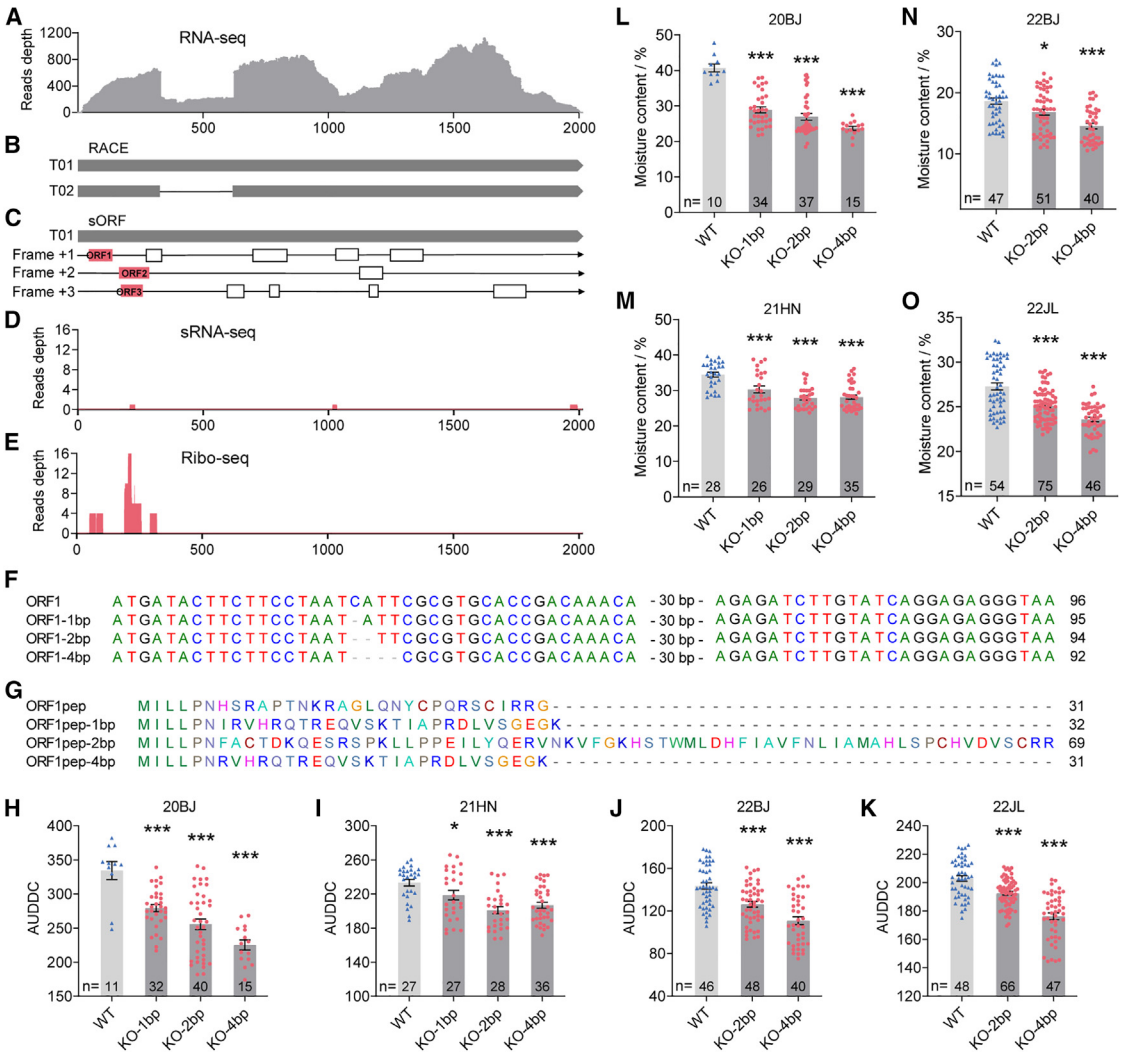


Figure 3. *RPG* encodes a functional micropeptide

(A) RNA-seq reads depth in *RPG* gene region.

(B) 5′-RACE and 3′-RACE reveal the full-length *RPG* cDNA, which contains two transcripts (T01 and T02).

(C) Twelve putative ORFs were identified by ORF finder in the T01 transcript. Pink and white rectangular boxes represent ORFs in *RPG*.

(D) sRNA-seq read depth in the *RPG* gene region.

(E) Ribo-seq read depth in the *RPG* gene region. *RPG* RNA is bound by the ribosome in ORF1, ORF2, and ORF3 regions, and their regions are demarcated in pink (C).

(F) Nucleotide sequences of three ORF1 (corresponding to B73 reference sequence version 4.0, Chr1:20020966–20021061) frameshift mutations (1-, 2-, and 4-bp deletion), −30 bp represents the AGAGAGCAGGTCTCCAAAACTATTGCCCCC sequence.

(G) Amino acid sequences of three ORF1 frameshift mutations (1-, 2-, and 4-bp deletion).

(H–K) Frameshift mutations in ORF1 had a faster KDR in 20BJ (H), in 21HN (I), in 22BJ (J), and in 22JL (K).

(L–O) The effect of ORF1 knockout on moisture content at harvest in different environments. ORF1 knockouts had lower moisture content in 20BJ (L), in 21HN (M), in 22BJ (N), and in 22JL (O).

20BJ, 21HN, 22BJ, and 22JL represent the field trials performed in Beijing in 2020, Hainan in 2021, Beijing in 2022, and Jilin in 2022, respectively (H–O).

The numbers on the right represent the nucleotide and amino acid positions in the full-length sequences (F and G).

Data are represented as mean ± SEM. *$p < 0.05$, ***$p < 0.001$ (Student's t test).

n is the sample size, a sample represents the phenotype from a plant.

See also Figure S2 and Tables S2 and S6.

extracts from maize kernels (Figure 4I). The endogenous ORF1 micropeptide was also verified by immunoprecipitation (IP) (Figure S3G) and identification by mass spectrometry (MS) (Figures 4J and S3H–S3J). These results indicate that the ORF1 micropeptide is produced endogenously in maize kernels. Thus, we named this 31 amino acid residue micropeptide as microRPG1 (micropeptide of *RPG* ORF1) (Figure 4J). microRPG1 appears not related to any known peptides in public databases.

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

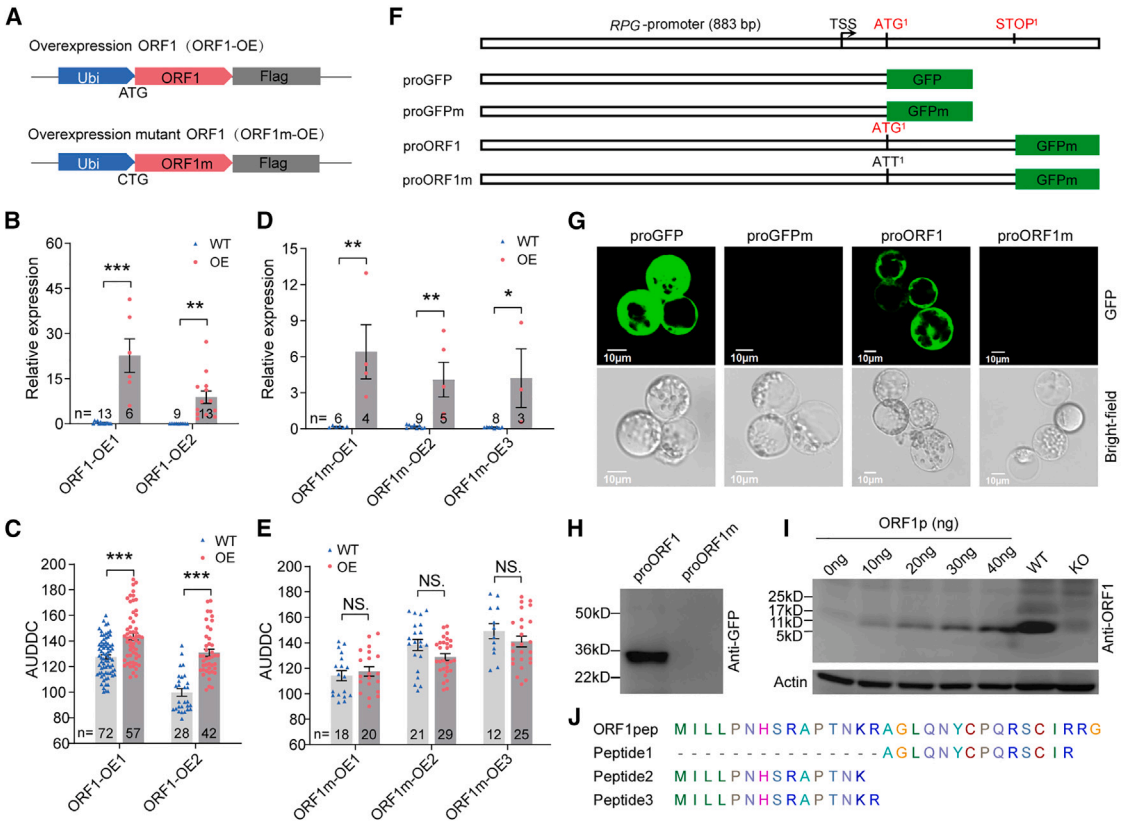


Figure 4. ORF1 encodes a micropeptide (microRPG1)

(A) Constructs used for overexpression of ORF1 and ORF1m (ORF1 start codon ATG mutated to CTG). Both ORF1 and ORF1m were fused with FLAG tags and driven by the *ubiquitin* promoter.

(B) ORF1 expression in two ORF1 overexpression lines (ORF1-OE). The expression levels of ORF1 were quantified using qPCR and normalized to maize *ACTIN* ($n \geq 6$).

(C) ORF1 overexpression slowed down KDR in two different events in field trials performed in Beijing in 2022 (22BJ).

(D) ORF1m expression in three different transgenic overexpression lines (ORF1m-OE1-3). The expression levels of ORF1m were quantified using qPCR and normalized to maize *ACTIN* ($n \geq 3$).

(E) ORF1m overexpression did not change KDR in field trials performed in Beijing in 2022 (22BJ).

(F) Schematic diagram of *RPG* promoter constructs used for GFP assays. proGFP, proGFPm (GFP start codon ATG mutated to CTT), proORF1, and proORF1m (ORF1 start codon ATG mutated to ATT) were analyzed in maize protoplasts. TSS, transcription start site.

(G) Representative images of GFP expression in maize protoplasts. Top panel GFP fluorescence (green) in maize protoplast cells was collected using a laser confocal microscope. Bottom panel shows bright-field images. Scale bars, 10 μm.

(H) Western blotting analysis of proORF1 and proORF1m transformed cells in (G) using anti-GFP antibody.

(I) Western blot analysis showing the presence of ORF1 micropeptide of the expected size in wild-type (WT) maize kernel extracts but not in knockout (KO) extracts. Different concentrations of synthetic ORF1 micropeptide (0∼40 ng) were used in the western blot analysis, and actin was used as the loading control.

(J) Endogenous ORF1 micropeptide was identified by mass spectrometry. Top, predicted ORF1 micropeptide of 31 amino acid residues. Bottom three rows, three unique peptides identified by mass spectrometry.

Data are represented as mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, NS, not significant (Student's t test).

n is the biological replicate size, a biological replicate is from a plant sample (maize leaves) (B and D). *n* is the sample size, a sample represents the phenotype from a plant (C and E).

See also Figure S3.

microRPG1 may control kernel dehydration through regulation of ethylene signaling

To ask how microRPG1 functions, we performed RNA-seq using microRPG1 knockout and overexpression lines compared with wild type as well as *Arabidopsis* treated with microRPG1. 384 genes were significantly downregulated in the microRPG1 knockout line and upregulated in microRPG1 overexpression line and exogenous micropeptide application, and 121 genes showed an opposite trend (Figures 5A and 5B). *RPG* is highly expressed in the late stage of maize kernel development (after 30 DAP) when the dry matter has been completely filled (Figures S1D, S4A, and S4B), implying that microRPG1 may function only during specific developmental periods. Since *RPG* is specifically expressed in the late stage of kernel maturation, we screened for differentially expressed genes with similar expression patterns to *RPG* by *Z* score (≥2). As a

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

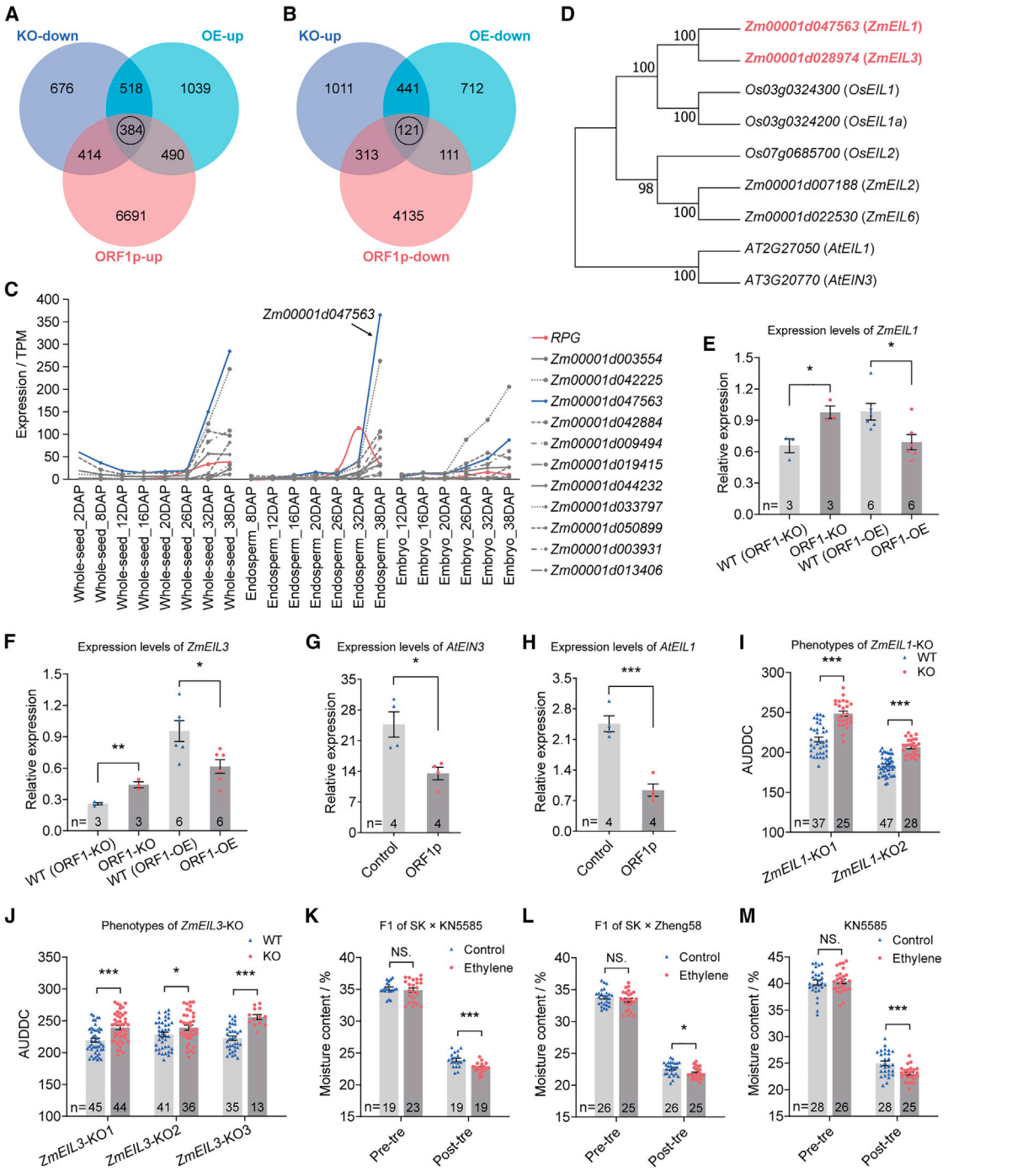


Figure 5. microRPG1 may control kernel dehydration through regulation of ethylene signaling

(A and B) Differential expression analysis of microRPG1 knockout, overexpression, and exogenous application by RNA-seq. The black circles represent overlapped genes.

(C) 11 genes specifically expressed in the late stages of seed (kernel) and endosperm displayed similar expression pattern with *RPG* in seed, endosperm, and embryo. Among them, *Zm00001d047563* showed the highest expression level. The expression levels of 11 genes in various tissues from B73 by RNA-seq.[58,59]

(D) Phylogenetic tree of EIN3/EIL in maize, rice, and *Arabidopsis*.

(E–H) The expression levels of *EIL* genes are regulated by microRPG1. The expression level of *ZmEIL1* (E) and *ZmEIL3* (F) in microRPG1 knockout (ORF1-KO, 37 DAP kernels) and overexpression (ORF1-OE, 40 DAP kernels) lines. The expression level of *AtEIN3* (G) and *AtEIL1* (H) in silique (*Arabidopsis* ripe silique, 72 days after sowing) treated with water (control) or 2.0 μM of ORF1p. The expression levels of *EIL* genes were quantified using qPCR and normalized to maize or *Arabidopsis ACTIN* ($n \geq 3$).

(legend continued on next page)

result, 11 genes (Figure S4C) had similar expression pattern to *RPG* in seed, endosperm, and embryo (Figure 5C), and these genes are more likely to be regulated by microRPG1 and play a similar role in the late stage of maize kernel development. Among them, *Zm00001d047563* an ortholog of *Ethylene-Insensitive3/Ethylene-Insensitive3-like 1* (*EIN3/EIL1*), a key component in ethylene signaling in *Arabidopsis* and rice (Figure 5D),[60–65] had the highest expression (Figure 5C). *Zm00001d047563* has four paralogs in maize; however, only *Zm00001d047563* (named *ZmEIL1*) and *Zm00001d028974* (named *ZmEIL3*) were expressed in the kernel and had similar expression pattern to *RPG* (Figures S1E and S4D). To confirm the expression pattern of microRPG1 micropeptide, we performed immunohistochemistry. The results showed that it was expressed in the embryo, aleurone layer, and placento-chalazal region (Figure S4E), consistent with the mRNA expression of *RPG* (Figure S1E). qPCR confirmed that *ZmEIL1* and *ZmEIL3* were upregulated in the microRPG1 knockout and downregulated in the microRPG1 overexpression lines (Figures 5E and 5F). This suggests that microRPG1 expression in the late stage of maize kernel development inhibits ethylene signaling, resulting in slower dehydration. Similarly, the *Arabidopsis* paralogs *EIN3* and *EIL1* were downregulated following exogenous application of microRPG1 (Figures 5G and 5H). These results indicate that microRPG1 represses the expression of *EIN3*-related genes in both maize and *Arabidopsis*.

To investigate the influence of *ZmEIL1* and *ZmEIL3* on the KDR phenotype, we next generated two knockout lines for *ZmEIL1* and three for *ZmEIL3* by CRISPR-Cas9 (Figures S4F and S4G). All of the knockout lines significantly decelerated KDR compared with wild-type sibling plants (Figures 5I and 5J). We also investigated the effects of *ZmEIL1* or *ZmEIL3* knockout on kernel quality traits and seed germination rate and found that knockout of *EIL* genes affected germination rate significantly and kernel quality slightly (Figures S5A–S5L). These results suggest that *EIL* genes play important roles in plant development and that the fine-tuning of their expression by microRPG1 can be beneficial for agricultural production.

Considering that *EIN3/EIL1* are key components in ethylene signaling, we also applied exogenous ethylene to different maize lines and found that this could accelerate KDR (Figures 5K–5M). Furthermore, we tested five marker genes for ethylene signaling[62,66–72] in microRPG1 knockout and overexpression lines and following exogenous micropeptide application. We found that these marker genes were downregulated when microRPG1 was overexpressed or exogenously applied and upregulated when microRPG1 was knocked out, supporting the idea that microRPG1 regulates ethylene signaling (Figures S5M–S5O). Our results indicate that ethylene signaling can affect KDR, suggesting a possible downstream mechanism for microRPG1 in controlling this trait. This finding is consistent with the well-known role of ethylene in stimulating fruit ripening.[71–73]

Origin of microRPG1 micropeptide

To investigate the origins of the microRPG1 micropeptide, we analyzed its syntenic regions in the Poaceae. We found homologous sequences in the genus *Zea* and *Tripsacum* but absent from other members of the Poaceae. Sequence alignments indicated that the 96-bp microRPG1 sequence was highly conserved in all maize and teosinte genomes (Table S2). Although similar sequences were present in *Tripsacum dactyloides*, they lacked a start codon, suggesting they are not translated. Phylogenetic analysis showed that a nucleotide (ACG to ATG) mutation led to the evolution of the ORF1 micropeptide (Figures 6, S6A, and S6B). These observations, together with the above finding that no known peptide associations exist in public databases, suggest that the microRPG1 micropeptide originated *de novo* from a non-coding sequence after the recent split between the genus *Zea* and *Tripsacum* ~0.65 million years ago (mya) (Figure 6).[45,74–78]

microRPG1 functions similarly in *Arabidopsis*

Given our demonstration that the microRPG1 micropeptide is able to control the ancient seed dehydration pathway in maize, we wondered whether it could exert similar effects in distantly related plant species. We therefore investigated whether exogenous application of synthetic ORF1p would function in *Arabidopsis* (Col-0). We applied different ORF1p concentrations ranging from 0.01 to 2 μM. Remarkably, silique ripening was significantly delayed at peptide concentrations of 1 and 2 μM (Figures S7A and S7B). A 2 μM concentration of ORF1p was selected for additional studies that again delayed silique ripening compared with plants treated with a control peptide (scORF1p), with the same amino acid composition but a scrambled sequence (Figures 7A–7E). Application of ORF1p significantly increased the moisture content of *Arabidopsis* seeds, but there was no significant difference in flowering time (Figures 7F and 7G). This result is similar to microRPG1 overexpression in maize, which leads to a slower KDR. To confirm this result, we generated three microRPG1 overexpression lines in *Arabidopsis*, and all of them significantly delayed silique ripening (Figures S7C and S7D). Our results indicate that microRPG1 micropeptide functions similarly in the distantly related *Arabidopsis* species as in maize.

The peptide treatment experiment suggests that ORF1p can enter cells, presumably by crossing the plasma membrane, echoing the behavior of some reported peptides.[79,80] To test this idea, we labeled ORF1p with fluorescein (FAM, carboxyfluorescein) (FAM-ORF1p) and incubated it with *Arabidopsis* seedlings. Interestingly, the fluorescein-labeled ORF1p rapidly

(I and J) *ZmEIL1* (I) and *ZmEIL3* (J) knockouts had slower KDR in field trials performed in Jilin in 2022 (22JL).
(K–M) Ethylene treatment resulted in a faster KDR in F1 of SK × KN5585 (K), in F1 of SK × Zheng58 (L), and in KN5585 (M). SK, KN5585, and Zheng58 are maize inbred lines. Pre-tre and post-tre represent pre-treatment and post-treatment, respectively.
Data are represented as mean ± SEM. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, NS, not significant (Student's t test).
n is the biological replicate size, a biological replicate is from a plant sample (E–H). *n* is the sample size, a sample represents the phenotype from a plant (I–M).
See also Figures S4 and S5.

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

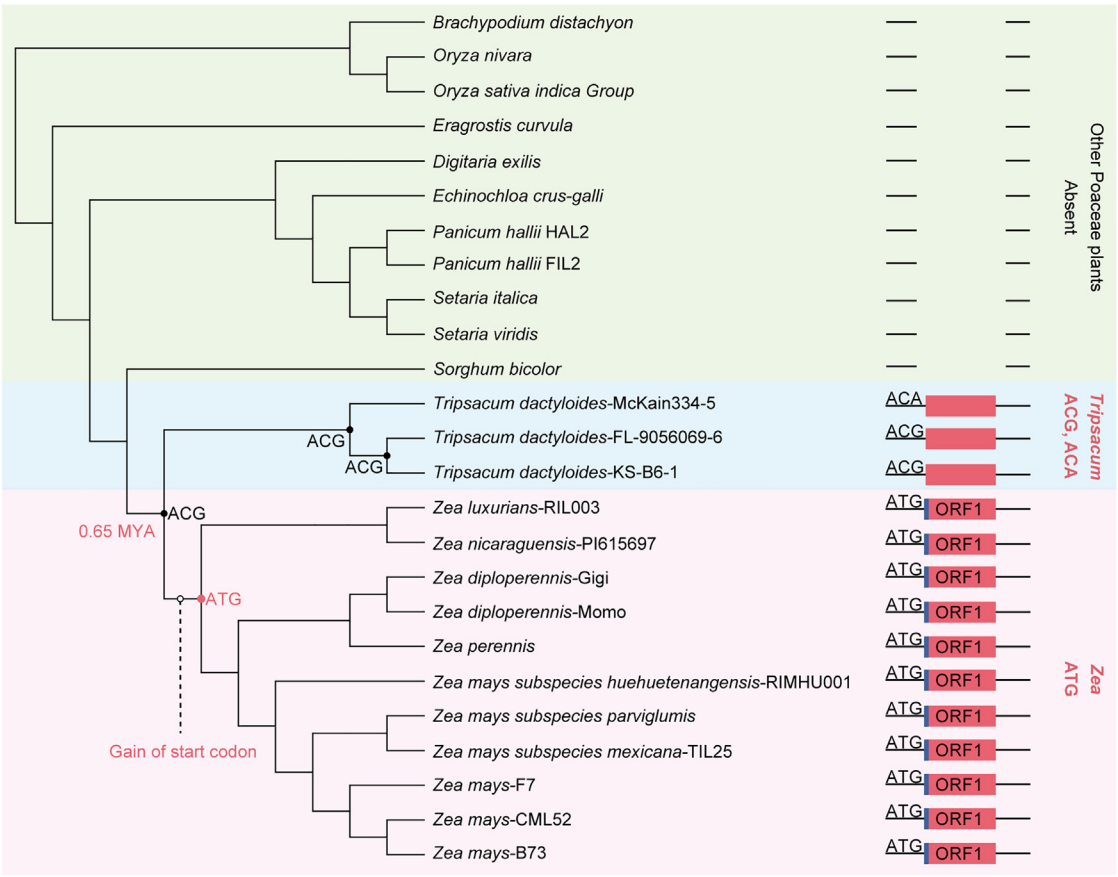


Figure 6. Origin of microRPG1 micropeptide

The phylogenetic tree shows the *de novo* origin of microRPG1. microRPG1 homologous sequences were detected in the genus *Zea* and *Tripsacum* but absent in other Poaceae plants. A single nucleotide (ACG to ATG) mutation at the start codon led to ORF1 that can be translated into microRPG1 protein. Pink boxes represent the microRPG1 homologous sequences. Blue bars represent the start codon of microRPG1. ACG and ACA represent the non-start codon sequences in the genus *Tripsacum*. The number in the phylogenetic tree is an assumed divergence time (million years ago).
See also Figure S6 and Table S2.

penetrated into the root (~0.5 h) but took longer to penetrate other parts of the seedling (Figures S7E and S7F). At 24 h after application, FAM-ORF1p was detected in most parts of the root and began to appear in the cotyledons (Figures 7H and S7E). FAM-ORF1p was detected in leaf veins (Figures 7H and S7G), suggesting that it is transported to the shoot tissues. To address whether ORF1p acts intracellularly in *Arabidopsis* rather than through cell surface receptors, we observed the subcellular localization of FAM-ORF1p. We found that FAM-ORF1p could enter root cells and localized in the nucleus, cytoplasm, and plasma membrane (Figures S7H and S7I), suggesting it may function through different mechanisms. These results were confirmed by incubation of FAM-ORF1p with maize protoplasts (Figure S7J). We also fused ORF1 with GFP and transformed it into maize protoplasts, and observed fluorescence in the nucleus, cytoplasm, and plasma membrane (Figure S7K). Taken together, these data suggest that the ORF1p micropeptide can enter cells and plays a similar role in *Arabidopsis* and maize, implying conserved function across plant species.

DISCUSSION

In this study, we identified and cloned *qKDR1*, a QTL for KDR in maize, and showed that it corresponds to a non-coding, intergenic DNA sequence. Two MYB-related-transcription factors bind to the *qKDR1* region to regulate the expression of its putative target-regulated gene, *RPG*. *RPG* encodes a functional micropeptide of 31 amino acids, and this micropeptide alters KDR by regulating the expression of two key ethylene signaling genes (Figure 7I). We thus identified a *Zea* genus-specific micropeptide and constructed a possible regulatory pathway for maize kernel dehydration, providing useful insights and research directions for our in-depth understanding and applications of seed dehydration in plants.

Non-coding DNA sequences located in intergenic regions can function either as enhancers or silencers and may be bound by transcription factors to regulate the expression of genes located up to megabases away.[81,82] Several such regulatory elements have been functionally characterized in plants, including an enhancer of *teosinte branched1* (*tb1*),[83] *Vegetative*

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

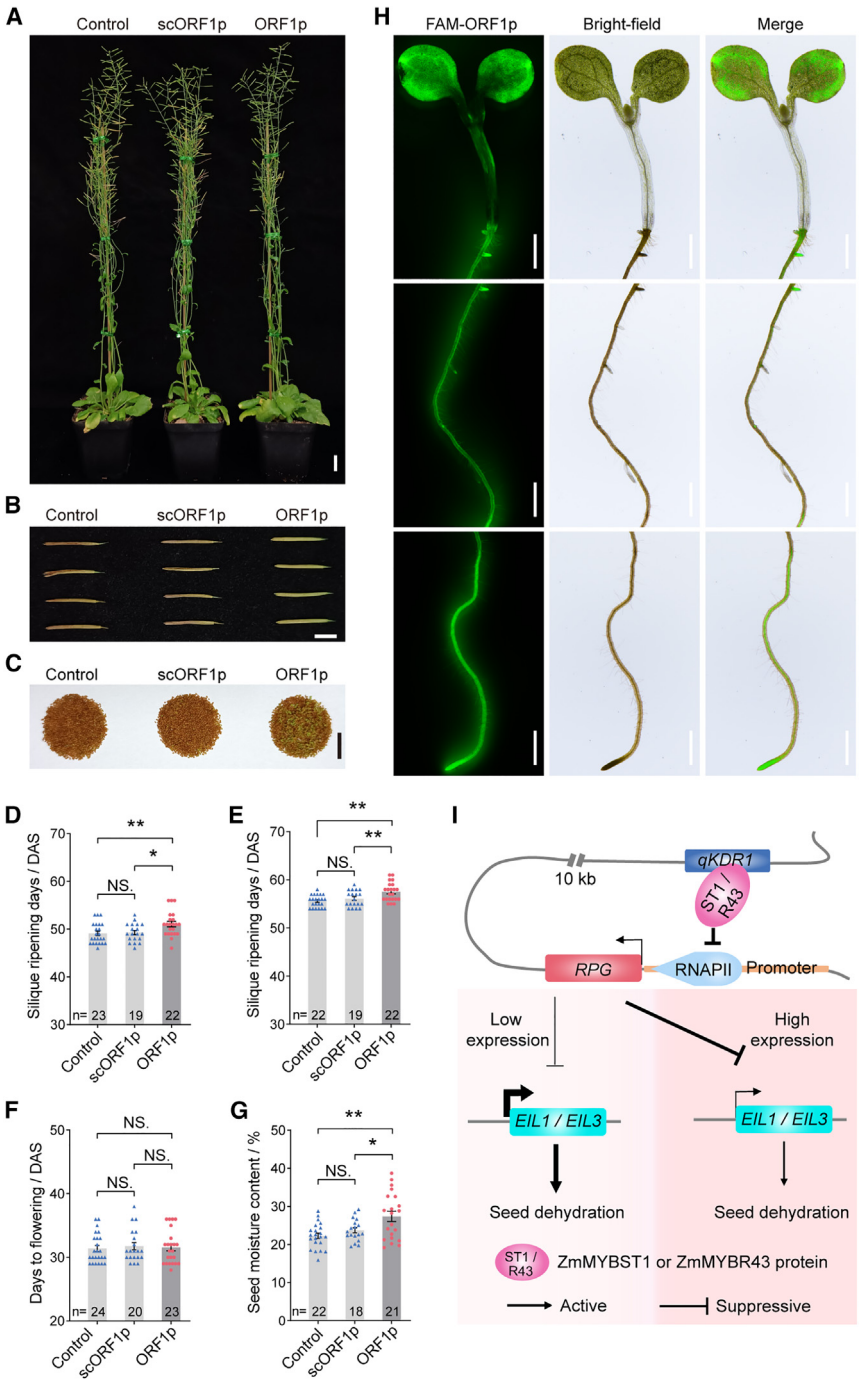


Figure 7. microRPG1 could enter cells and function similarly in *Arabidopsis*, and a model of microRPG1 regulates seed dehydration

(A) Representative image of plants treated with water (control) or 2.0 μM of scORF1p or ORF1p. Scale bar, 2 cm.

(B and C) Representative images of silique (B) and seeds (C) treated with water (control) or 2.0 μM of scORF1p or ORF1p. Scale bar, 0.5 cm.

(D and E) Days required for silique ripening for plants treated with water (control) or 2.0 μM of scORF1p or ORF1p. Silique ripening days were recorded when the siliques began to ripen (D) and when 50% of the siliques were ripe (E). DAS, days after sowing.

(F) Days to flowering of plants treated with water (control) or 2.0 μM of scORF1p or ORF1p. DAS, days after sowing.

(G) Seed moisture content of plants treated with water (control) or 2.0 μM of scORF1p or ORF1p.

(H) FAM-ORF1p is absorbed by seedlings. The figure shows green fluorescence, bright-field, and merged microscopic images of Col-0 seedlings in the presence of 10 μM FAM-ORF1p after incubation for 24 h. Scale bars, 1 mm.

(I) A model of how microRPG1 regulates seed dehydration. *qKDR1* represses *RPG* transcription by recruiting ZmMYBST1 and ZmMYBR43. The 6,181-bp insertion in NIL^{DAN340} reduces the silencing activity of *qKDR1*. *RPG* encodes a microRPG1 micropeptide of 31 amino acid residues, which slows down seed dehydration by regulating the expression of *EIL* genes.

scORF1p, the same amino acid composition as the ORF1p but a scrambled sequence.

n is the sample size, a sample represents the phenotype from a plant (D–F) or from all the seeds of a plant (G).

FAM-ORF1p, 5-FAM fluorescein-labeled micro RPG1.

Data are represented as mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, NS, not significant (Student's t test).

See also Figure S7.

to generative transition1 (*Vgt1*),[84,85] *Kernel Row Number4* (*KRN4*),[86] and *Upright Plant Architecture2* (*UPA2*)[87] in maize, as well as enhancer of *Grain Width5* (*GW5*)[88] and silencer of *Frizzy Panicle* (*FZP*)[89] in rice. Here, we cloned a non-coding DNA sequence, *qKDR1*, that is located ~10 kilobases downstream of *RPG* and can repress *RPG* expression. However, we have not yet revealed the molecular mechanism by which the *qKDR1* region interacts with the *RPG* promoter, and this is a priority for future research.

Here, we also identified a *Zea* genus-specific micropeptide encoded by the *RPG* locus. Although the core DNA sequences are also present in *Tripsacum*, a single nucleotide mutation (C to T) occurred exclusively in the genus *Zea*, creating a start codon that allowed translation of the functional microRPG1. There are various mechanisms for the origin of new genes,[41–43] but it is rare to report that a non-functional sequence becomes a functional gene through a single-base mutation, making this an excellent example to understand the origin of new genes.

The maize kernel moisture content suitable for mechanized harvesting ranges from 15% to 25%, but the moisture content of most maize varieties in China and many other global regions typically ranges from 30% to 40% at harvest.[1,10–12] We found that microRPG1 knockouts decreased the moisture content

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

at harvest by an average of 7.1% and up to 17.0%, depending on the environment and harvest time. Our findings suggest that manipulating microRPG1 to alter the KDR in maize may have great potential in future breeding for mechanized harvesting, since almost all maize lines containing the micropeptide. The microRPG1 knockout lines had no obvious effects on other agronomic traits in three different environments (Table S6). Interestingly, we found that microRPG1 alters KDR by regulating the expression of *ZmEIL1* and *ZmEIL3*, and similar KDR phenotypes were observed in knockouts of *ZmEIL1* or *ZmEIL3*, or following exogenous ethylene application to maize kernels. This hints that the ancient ethylene signaling pathway is involved in maize kernel dehydration. The knockout of *EIL* genes showed some adverse phenotypes, but microRPG1 knockouts did not. This suggests that fine-tuning of *ZmEIL1* and *ZmEIL3* expression by microRPG1 has obvious application value (Figures S5A–S5L; Table S6). By precisely regulating the spatiotemporal expression of additional genes in this pathway, we might find other ways to control maize kernel dehydration and eventually pave the way for greater understanding of the underlying mechanisms. The finding that exogenous application of microRPG1 invokes similar responses in *Arabidopsis* suggests that its utility may reach far beyond maize. Further exploring whether this micropeptide plays a role in other species and how it works will be our research focus in the coming years.

Limitations of the study

qKDR1 is a non-coding regulatory sequence located in an intergenic region downstream of *RPG*; however, how the *qKDR1* region interacts with the *RPG* promoter remains to be fully explored. The molecular mechanisms of microRPG1 have not been fully elucidated. Micropeptides may act intracellularly, or function through cell surface receptors. We found that microRPG1 can be taken up by cells and localizes to the nucleus, cytoplasm, and plasma membrane, suggesting that it may function in multiple ways. In addition to experiments mentioned above, the discovery of intracellular targets or cell surface receptors of microRPG1 remains to be achieved.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jianbing Yan (yjianbing@mail.hzau.edu.cn).

Materials availability

Plasmids, transgenic plants, and unique/stable reagents generated in this study will be made available on request for scientific research, but we may require a payment and a completed materials transfer agreement for commercial purpose.

Data and code availability

- All data are available in the main text and the supplemental information or at public databases. Sequence data were deposited in Genome Sequence Archive (GSA)[90] of the National Genomics Data Center[91] (https://ngdc.cncb.ac.cn/gsa), with accession numbers GSA: CRA013670, CRA013671, CRA013672, CRA013673. Experimental data of genetic mapping, fine mapping, or agronomic traits were deposited in Zenodo: https://doi.org/10.5281/zenodo.13924714. These accession numbers are also listed in the key resources table.
- This study does not report original code.
- Any additional information required to reanalyze the data reported in this study is available from the lead contact upon request.

ACKNOWLEDGMENTS

This research was supported by the National Key Research and Development Program of China (2022YFD1201502), the National Natural Science Foundation of China (32430076, 32321005, and 32201851), the Science and Technology major program of Hubei Province (2021ABA011), and the 111 Project Crop Genomics and Molecular Breeding (B20051) and by NSF IOS 2129189 to D.J.

AUTHOR CONTRIBUTIONS

Jianbing Yan designed this study. Y.Y., Yuanfang Liu, Q.Z., W.D., Y. Zou, Junjun Yan, Jiali Yan, Y.G., Z.Z., L.L., and B.Z. performed experiments. Y.Y. and Yuanfang Liu analyzed data. Yanjun Liu, Y.O., Y. Xue, A.J., N.Y., Y. Xiao, L.Z., Z.L., and P.Y. contributed valuable suggestions in this study. Y.Y., W.L., Yuanfang Liu, Q.Z., X.H., Y. Zheng, and J.X. collected phenotypic data. Y.Y., Yuanfang Liu, H.-J.L., A.R.F., D.J., and Jianbing Yan wrote and edited the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

Y.Y., W.L., Yuanfang Liu, J.X., L.Z., and Jianbing Yan are authors on patent applications related to this work. J.X. is employed by WIMI Biotechnology Co., Ltd.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Maize and growth conditions
  - *Arabidopsis* and growth conditions
- METHOD DETAILS
  - Plant materials
  - Phenotypic data collection and analysis
  - QTL mapping
  - Fine mapping of *qKDR1*
  - RNA extraction and expression analysis
  - Field trials of agronomic, flowering time and yield-related traits
  - Sequence and haplotype analysis of *qKDR1*
  - Rapid-amplification of cDNA ends
  - RNA and small RNA sequencing
  - Ribosome profiling sequencing
  - GFP assay and subcellular localization
  - Western blotting
  - Co-immunoprecipitation (co-IP)
  - Transient transcriptional activity assays
  - Micropeptide synthesis
  - Identification of mature phenotype in *Arabidopsis*
  - Micropeptide absorption in *Arabidopsis*
  - Synteny analysis
  - Phylogenetic tree
  - Identification of differentially expressed genes with similar expression patterns to *RPG*
  - mRNA *in situ* hybridization
  - Immunohistochemistry
  - Determination of kernel oil content
  - Determination of kernel starch content
  - Seed germination rate assays
  - Exogenous ethylene treatment

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2024.10.030.

Received: January 14, 2024
Revised: June 9, 2024
Accepted: October 16, 2024
Published: November 12, 2024

REFERENCES

1. Liu, J., Yu, H., Liu, Y., Deng, S., Liu, Q., Liu, B., and Xu, M. (2020). Genetic dissection of grain water content and dehydration rate related to mechanical harvest in maize. BMC Plant Biol. *20*, 118. https://doi.org/10.1186/s12870-020-2302-0.
2. Li, S.K., Wang, K.R., Xie, R.Z., and Ming, B. (2018). Grain mechanical harvesting technology promotes the transformation of maize production mode. Sci. Agric. Sin. *51*, 1842–1844. https://doi.org/10.3864/j.issn.0578-1752.2018.10.003.
3. Wang, C., Shu, Z., Cheng, B., Jiang, H., and Li, X. (2018). Advances and perspectives in maize mechanized harvesting in China. J. Anhui Agric. Univ. *45*, 551–555. https://doi.org/10.13610/j.cnki.1672-352x.20180620.005.
4. Jia, T., Wang, L., Li, J., Ma, J., Cao, Y., Lübberstedt, T., and Li, H. (2020). Integrating a genome-wide association study with transcriptomic analysis to detect genes controlling grain drying rate in maize (*Zea may*, L.). Theor. Appl. Genet. *133*, 623–634. https://doi.org/10.1007/s00122-019-03492-0.
5. Crane, P.L., Miles, S.R., and Newman, J.E. (1959). Factors associated with varietal diferences in rate of feld drying in corn. Agron. J. *51*, 318–320. https://doi.org/10.2134/agronj1959.00021962005100060003x.
6. Yin, S., Liu, J., Yang, T., Li, P., Xu, Y., Fang, H., Xu, S., Wei, J., Xue, L., Hao, D., et al. (2020). Genetic analysis of the seed dehydration process in maize based on a logistic model. Crop J. *8*, 182–193. https://doi.org/10.1016/j.cj.2019.06.011.
7. Wang, W., Ren, Z., Li, L., Du, Y., Zhou, Y., Zhang, M., Li, Z., Yi, F., and Duan, L. (2022). Meta-QTL analysis explores the key genes, especially hormone related genes, involved in the regulation of grain water content and grain dehydration rate in maize. BMC Plant Biol. *22*, 346. https://doi.org/10.1186/s12870-022-03738-y.
8. de Jager, B., Roux, C.Z., and Kühn, H.C. (2004). An evaluation of two collections of South African maize (*Zea mays* L.) germ plasm: 2. The genetic basis of dry-down rate. S. Afr. J. Plant Soil *21*, 120–122. https://doi.org/10.1080/02571862.2004.10635035.
9. Zhou, G., Hao, D., Xue, L., Chen, G., Lu, H., Zhang, Z., Shi, M., Huang, X., and Mao, Y. (2018). Genome-wide association study of kernel moisture content at harvest stage in maize. Breed. Sci. *68*, 622–628. https://doi.org/10.1270/jsbbs.18102.
10. Dai, L., Wu, L., Dong, Q., Zhang, Z., Wu, N., Song, Y., Lu, S., and Wang, P. (2017). Genome-wide association study of field grain drying rate after physiological maturity based on a resequencing approach in elite maize germplasm. Euphytica *213*, 182. https://doi.org/10.1007/s10681-017-1970-9.
11. Plett, S. (1994). Corn kernel breakage as a function of grain moisture at harvest in a prairie environment. Can. J. Plant Sci. *74*, 543–544. https://doi.org/10.4141/cjps94-097.
12. Li, L.L., Xue, J., Xie, R.Z., Wang, K.R., Ming, B., Hou, P., Gao, S., and Li, S.K. (2018). Effects of grain moisture content on mechanical grain harvesting quality of summer maize. Acta Agron. Sin. *44*, 1747–1754. https://doi.org/10.3724/sp.J.1006.2018.01747.
13. Zhang, J., Zhang, F., Tang, B., Ding, Y., Xia, L., Qi, J., Mu, X., Gu, L., Lu, D., and Chen, Y. (2020). Molecular mapping of quantitative trait loci for grain moisture at harvest and field grain drying rate in maize (*Zea mays* L.). Physiol. Plant. *169*, 64–72. https://doi.org/10.1111/ppl.13048.
14. Li, S., Zhang, C., Lu, M., Yang, D., Qian, Y., Yue, Y., Zhang, Z., Jin, F., Wang, M., Liu, X., et al. (2020). QTL mapping and GWAS for field kernel water content and kernel dehydration rate before physiological maturity in maize. Sci. Rep. *10*, 13114. https://doi.org/10.1038/s41598-020-69890-3.
15. Xiang, K., Reid, L.M., Zhang, Z.M., Zhu, X.Y., and Pan, G.T. (2012). Characterization of correlation between grain moisture and ear rot resistance in maize by QTL meta-analysis. Euphytica *183*, 185–195. https://doi.org/10.1007/s10681-011-0440-z.
16. Kebebe, A.Z., Reid, L.M., Zhu, X., Wu, J., Woldemariam, T., Voloaca, C., and Xiang, K. (2015). Relationship between kernel drydown rate and resistance to Gibberella ear rot in maize. Euphytica *201*, 79–88. https://doi.org/10.1007/s10681-014-1185-2.
17. Capelle, V., Remoué, C., Moreau, L., Reyss, A., Mahé, A., Massonneau, A., Falque, M., Charcosset, A., Thévenot, C., Rogowsky, P., et al. (2010). QTLs and candidate genes for desiccation and abscisic acid content in maize kernels. BMC Plant Biol. *10*, 2. https://doi.org/10.1186/1471-2229-10-2.
18. Sala, R.G., Andrade, F.H., Camadro, E.L., and Cerono, J.C. (2006). Quantitative trait loci for grain moisture at harvest and field grain drying rate in maize (*Zea mays*, L.). Theor. Appl. Genet. *112*, 462–471. https://doi.org/10.1007/s00122-005-0146-5.
19. Qu, J., Zhong, Y., Ding, L., Liu, X., Xu, S., Guo, D., Blennow, A., and Xue, J. (2022). Biosynthesis, structure and functionality of starch granules in maize inbred lines with different kernel dehydration rate. Food Chem. *368*, 130796. https://doi.org/10.1016/j.foodchem.2021.130796.
20. Liu, X.J., Wang, Z.H., Wang, X., Li, T.F., and Zhang, L. (2010). Primary mapping of QTL for dehydration rate of maize kernel after physiological maturing. Acta Agron. Sin. *36*, 47–52. https://doi.org/10.3724/sp.J.1006.2010.00047.
21. Li, W., Yu, Y., Wang, L., Luo, Y., Peng, Y., Xu, Y., Liu, X., Wu, S., Jian, L., Xu, J., et al. (2021). The genetic architecture of the dynamic changes in grain moisture in maize. Plant Biotechnol. J. *19*, 1195–1205. https://doi.org/10.1111/pbi.13541.
22. Wang, Z., Wang, X., Zhang, L., Liu, X., Di, H., Li, T., and Jin, X. (2012). QTL underlying field grain drying rate after physiological maturity in maize (*Zea mays* L.). Euphytica *185*, 521–528. https://doi.org/10.1007/s10681-012-0676-2.
23. Li, Y., Dong, Y., Yang, M., Wang, Q., Shi, Q., Zhou, Q., Deng, F., Ma, Z., Qiao, D., and Xu, H. (2014). QTL detection for grain water relations and genetic correlations with grain matter accumulation at four stages after pollination in maize. J. Plant Biochem. Physiol. *2*, 121. https://doi.org/10.4172/2329-9029.1000121.
24. Manfre, A.J., LaHatte, G.A., Climer, C.R., and Marcotte, W.R. (2009). Seed dehydration and the establishment of desiccation tolerance during seed maturation is altered in the *Arabidopsis thaliana* mutant *atem6-1*. Plant Cell Physiol. *50*, 243–253. https://doi.org/10.1093/pcp/pcn185.
25. Andrews, S.J., and Rothnagel, J.A. (2014). Emerging evidence for functional peptides encoded by short open reading frames. Nat. Rev. Genet. *15*, 193–204. https://doi.org/10.1038/nrg3520.
26. Orr, M.W., Mao, Y., Storz, G., and Qian, S.B. (2020). Alternative ORFs and small ORFs: shedding light on the dark proteome. Nucleic Acids Res. *48*, 1029–1042. https://doi.org/10.1093/nar/gkz734.
27. Takahashi, F., Hanada, K., Kondo, T., and Shinozaki, K. (2019). Hormone-like peptides and small coding genes in plant stress signaling and development. Curr. Opin. Plant Biol. *51*, 88–95. https://doi.org/10.1016/j.pbi.2019.05.011.
28. Makarewich, C.A., and Olson, E.N. (2017). Mining for micropeptides. Trends Cell Biol. *27*, 685–696. https://doi.org/10.1016/j.tcb.2017.04.006.

29. Bazzini, A.A., Johnstone, T.G., Christiano, R., Mackowiak, S.D., Obermayer, B., Fleming, E.S., Vejnar, C.E., Lee, M.T., Rajewsky, N., Walther, T.C., et al. (2014). Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. EMBO J. *33*, 981–993. https://doi.org/10.1002/embj.201488411.

30. Anderson, D.M., Anderson, K.M., Chang, C.L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R., et al. (2015). A micropeptide encoded by a putative long non-coding RNA regulates muscle performance. Cell *160*, 595–606. https://doi.org/10.1016/j.cell.2015.01.009.

31. Huang, J.Z., Chen, M., Chen, D., Gao, X.C., Zhu, S., Huang, H., Hu, M., Zhu, H., and Yan, G.R. (2017). A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. Mol. Cell *68*, 171–184.e6. https://doi.org/10.1016/j.molcel.2017.09.015.

32. Matsumoto, A., Pasut, A., Matsumoto, M., Yamashita, R., Fung, J., Monteleone, E., Saghatelian, A., Nakayama, K.I., Clohessy, J.G., and Pandolfi, P.P. (2017). mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. Nature *541*, 228–232. https://doi.org/10.1038/nature21034.

33. Sun, L., Wang, W., Han, C., Huang, W., Sun, Y., Fang, K., Zeng, Z., Yang, Q., Pan, Q., Chen, T., et al. (2021). The oncomicropeptide APPLE promotes hematopoietic malignancy by enhancing translation initiation. Mol. Cell *81*, 4493–4508.e9. https://doi.org/10.1016/j.molcel.2021.08.033.

34. Zhang, C., Zhou, B., Gu, F., Liu, H., Wu, H., Yao, F., Zheng, H., Fu, H., Chong, W., Cai, S., et al. (2022). Micropeptide PACMP inhibition elicits synthetic lethal effects by decreasing CtIP and poly(ADP-ribosyl)ation. Mol. Cell *82*, 1297–1312.e8. https://doi.org/10.1016/j.molcel.2022.01.020.

35. Sousa, C., Johansson, C., Charon, C., Manyani, H., Sautter, C., Kondorosi, A., and Crespi, M. (2001). Translational and structural requirements of the early nodulin gene *enod40*, a short-open reading frame-containing RNA, for elicitation of a cell-specific growth response in the alfalfa root cortex. Mol. Cell. Biol. *21*, 354–366. https://doi.org/10.1128/mcb.21.1.354-366.2001.

36. Röhrig, H., Schmidt, J., Miklashevichs, E., Schell, J., and John, M. (2002). Soybean *ENOD40* encodes two peptides that bind to sucrose synthase. Proc. Natl. Acad. Sci. USA *99*, 1915–1920. https://doi.org/10.1073/pnas.022664799.

37. Casson, S.A., Chilley, P.M., Topping, J.F., Evans, I.M., Souter, M.A., and Lindsey, K. (2002). The *POLARIS* gene of *Arabidopsis* encodes a predicted peptide required for correct root growth and leaf vascular patterning. Plant Cell *14*, 1705–1721. https://doi.org/10.1105/tpc.002618.

38. Narita, N.N., Moore, S., Horiguchi, G., Kubo, M., Demura, T., Fukuda, H., Goodrich, J., and Tsukaya, H. (2004). Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in *Arabidopsis thaliana*. Plant J. *38*, 699–713. https://doi.org/10.1111/j.1365-313X.2004.02078.x.

39. Blanvillain, R., Young, B., Cai, Y.M., Hecht, V., Varoquaux, F., Delorme, V., Lancelin, J.M., Delseny, M., and Gallois, P. (2011). The *Arabidopsis* peptide kiss of death is an inducer of programmed cell death. EMBO J. *30*, 1173–1183. https://doi.org/10.1038/emboj.2011.14.

40. De Coninck, B., Carron, D., Tavormina, P., Willem, L., Craik, D.J., Vos, C., Thevissen, K., Mathys, J., and Cammue, B.P.A. (2013). Mining the genome of *Arabidopsis thaliana* as a basis for the identification of novel bioactive peptides involved in oxidative stress tolerance. J. Exp. Bot. *64*, 5297–5307. https://doi.org/10.1093/jxb/ert295.

41. Long, M., Betrán, E., Thornton, K., and Wang, W. (2003). The origin of new genes: glimpses from the young and old. Nat. Rev. Genet. *4*, 865–875. https://doi.org/10.1038/nrg1204.

42. Kaessmann, H. (2010). Origins, evolution, and phenotypic impact of new genes. Genome Res. *20*, 1313–1326. https://doi.org/10.1101/gr.101386.109.

43. Chen, S., Krinsky, B.H., and Long, M. (2013). New genes as drivers of phenotypic evolution. Nat. Rev. Genet. *14*, 645–660. https://doi.org/10.1038/nrg3521.

44. Zhou, Y., Zhang, C., Zhang, L., Ye, Q., Liu, N., Wang, M., Long, G., Fan, W., Long, M., and Wing, R.A. (2022). Gene fusion as an important mechanism to generate new genes in the genus *Oryza*. Genome Biol. *23*, 130. https://doi.org/10.1186/s13059-022-02696-w.

45. Zhang, L., Ren, Y., Yang, T., Li, G., Chen, J., Gschwend, A.R., Yu, Y., Hou, G., Zi, J., Zhou, R., et al. (2019). Rapid evolution of protein diversity by de novo origination in *Oryza*. Nat. Ecol. Evol. *3*, 679–690. https://doi.org/10.1038/s41559-019-0822-5.

46. McLysaght, A., and Hurst, L.D. (2016). Open questions in the study of de novo genes: what, how and why. Nat. Rev. Genet. *17*, 567–578. https://doi.org/10.1038/nrg.2016.78.

47. Liang, M., Chen, W., LaFountain, A.M., Liu, Y., Peng, F., Xia, R., Bradshaw, H.D., and Yuan, Y.W. (2023). Taxon-specific, phased siRNAs underlie a speciation locus in monkeyflowers. Science *379*, 576–582. https://doi.org/10.1126/science.adf1323.

48. Huang, Y., Chen, J., Dong, C., Sosa, D., Xia, S., Ouyang, Y., Fan, C., Li, D., Mortola, E., Long, M., et al. (2022). Species-specific partial gene duplication in *Arabidopsis thaliana* evolved novel phenotypic effects on morphological traits under strong positive selection. Plant Cell *34*, 802–817. https://doi.org/10.1093/plcell/koab291.

49. Pan, Q., Li, L., Yang, X., Tong, H., Xu, S., Li, Z., Li, W., Muehlbauer, G.J., Li, J., and Yan, J. (2016). Genome-wide recombination dynamics are associated with phenotypic variation in maize. New Phytol. *210*, 1083–1094. https://doi.org/10.1111/nph.13810.

50. Xiao, Y., Tong, H., Yang, X., Xu, S., Pan, Q., Qiao, F., Raihan, M.S., Luo, Y., Liu, H., Zhang, X., et al. (2016). Genome-wide dissection of the maize ear genetic architecture using multiple populations. New Phytol. *210*, 1095–1106. https://doi.org/10.1111/nph.13814.

51. Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., McMullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. *38*, 203–208. https://doi.org/10.1038/ng1702.

52. Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y., and Buckler, E.S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics *23*, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308.

53. Zhang, Z., Ersoz, E., Lai, C.Q., Todhunter, R.J., Tiwari, H.K., Gore, M.A., Bradbury, P.J., Yu, J., Arnett, D.K., Ordovas, J.M., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. Nat. Genet. *42*, 355–360. https://doi.org/10.1038/ng.546.

54. Liu, N., Liu, J., Li, W., Pan, Q., Liu, J., Yang, X., Yan, J., and Xiao, Y. (2018). Intraspecific variation of residual heterozygosity and its utility for quantitative genetic studies in maize. BMC Plant Biol. *18*, 66. https://doi.org/10.1186/s12870-018-1287-4.

55. Yang, X., Gao, S., Xu, S., Zhang, Z., Prasanna, B.M., Li, L., Li, J., and Yan, J. (2011). Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. Mol. Breed. *28*, 511–526. https://doi.org/10.1007/s11032-010-9500-7.

56. Peng, Y., Xiong, D., Zhao, L., Ouyang, W., Wang, S., Sun, J., Zhang, Q., Guan, P., Xie, L., Li, W., et al. (2019). Chromatin interaction maps reveal genetic regulation for quantitative traits in maize. Nat. Commun. *10*, 2632. https://doi.org/10.1038/s41467-019-10602-5.

57. Tu, X., Mejía-Guerra, M.K., Valdes Franco, J.A., Tzeng, D., Chu, P.Y., Shen, W., Wei, Y., Dai, X., Li, P., Buckler, E.S., et al. (2020). Reconstructing the maize leaf regulatory network using ChIP-seq data of 104 transcription factors. Nat. Commun. *11*, 5089. https://doi.org/10.1038/s41467-020-18832-8.

58. Stelpflug, S.C., Sekhon, R.S., Vaillancourt, B., Hirsch, C.N., Buell, C.R., de Leon, N., and Kaeppler, S.M. (2016). An expanded maize gene

expression atlas based on RNA sequencing and its use to explore root development. Plant Genome *9*. https://doi.org/10.3835/plantgenome2015.04.0025.

59. Chen, J., Zeng, B., Zhang, M., Xie, S., Wang, G., Hauck, A., and Lai, J. (2014). Dynamic transcriptome landscape of maize embryo and endosperm development. Plant Physiol. *166*, 252–264. https://doi.org/10.1104/pp.114.240689.

60. Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W., and Ecker, J.R. (1997). Activation of the ethylene gas response pathway in *Arabidopsis* by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. Cell *89*, 1133–1144. https://doi.org/10.1016/s0092-8674(00)80300-1.

61. Alonso, J.M., Stepanova, A.N., Solano, R., Wisman, E., Ferrari, S., Ausubel, F.M., and Ecker, J.R. (2003). Five components of the ethylene-response pathway identified in a screen for *weak ethylene-insensitive* mutants in *Arabidopsis*. Proc. Natl. Acad. Sci. USA *100*, 2992–2997. https://doi.org/10.1073/pnas.0438070100.

62. Dolgikh, V.A., Pukhovaya, E.M., and Zemlyanskaya, E.V. (2019). Shaping ethylene response: the role of EIN3/EIL1 transcription factors. Front. Plant Sci. *10*, 1030. https://doi.org/10.3389/fpls.2019.01030.

63. Mao, C., Wang, S., Jia, Q., and Wu, P. (2006). *OsEIL1*, a rice homolog of the *Arabidopsis EIN3* regulates the ethylene response as a positive component. Plant Mol. Biol. *61*, 141–152. https://doi.org/10.1007/s11103-005-6184-1.

64. Yang, C., Lu, X., Ma, B., Chen, S.Y., and Zhang, J.S. (2015). Ethylene signaling in rice and *Arabidopsis*: conserved and diverged aspects. Mol. Plant *8*, 495–505. https://doi.org/10.1016/j.molp.2015.01.003.

65. Zhao, H., Yin, C.C., Ma, B., Chen, S.Y., and Zhang, J.S. (2021). Ethylene signaling in rice and *Arabidopsis*: new regulators and mechanisms. J. Integr. Plant Biol. *63*, 102–125. https://doi.org/10.1111/jipb.13028.

66. Zhang, F., Qi, B., Wang, L., Zhao, B., Rode, S., Riggan, N.D., Ecker, J.R., and Qiao, H. (2016). EIN2-dependent regulation of acetylation of histone H3K14 and non-canonical histone H3K23 in ethylene signalling. Nat. Commun. *7*, 13018. https://doi.org/10.1038/ncomms13018.

67. Zhang, F., Wang, L., Qi, B., Zhao, B., Ko, E.E., Riggan, N.D., Chin, K., and Qiao, H. (2017). EIN2 mediates direct regulation of histone acetylation in the ethylene response. Proc. Natl. Acad. Sci. USA *114*, 10274–10279. https://doi.org/10.1073/pnas.1707937114.

68. Chang, K.N., Zhong, S., Weirauch, M.T., Hon, G., Pelizzola, M., Li, H., Huang, S.S.C., Schmitz, R.J., Urich, M.A., Kuo, D., et al. (2013). Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in Arabidopsis. eLife *2*, e00675. https://doi.org/10.7554/eLife.00675.

69. Zhong, S., Shi, H., Xue, C., Wang, L., Xi, Y., Li, J., Quail, P.H., Deng, X.W., and Guo, H. (2012). A molecular framework of light-controlled phytohormone action in Arabidopsis. Curr. Biol. *22*, 1530–1535. https://doi.org/10.1016/j.cub.2012.06.039.

70. Nemhauser, J.L., Hong, F., and Chory, J. (2006). Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. Cell *126*, 467–475. https://doi.org/10.1016/j.cell.2006.05.050.

71. Liu, M., Pirrello, J., Chervin, C., Roustan, J.P., and Bouzayen, M. (2015). Ethylene control of fruit ripening: revisiting the complex network of transcriptional regulation. Plant Physiol. *169*, 2380–2390. https://doi.org/10.1104/pp.15.01361.

72. Huang, W., Hu, N., Xiao, Z., Qiu, Y., Yang, Y., Yang, J., Mao, X., Wang, Y., Li, Z., and Guo, H. (2022). A molecular framework of ethylene-mediated fruit growth and ripening processes in tomato. Plant Cell *34*, 3280–3300. https://doi.org/10.1093/plcell/koac146.

73. Grierson, D. (2013). Ethylene and the control of fruit ripening. Mol. Biol. Biochem. Fruit Ripening, 43–73. https://doi.org/10.1002/9781118593714.ch3.

74. Chen, L., Luo, J., Jin, M., Yang, N., Liu, X., Peng, Y., Li, W., Phillips, A., Cameron, B., Bernal, J.S., et al. (2022). Genome sequencing reveals evidence of adaptive variation in the genus *Zea*. Nat. Genet. *54*, 1736–1745. https://doi.org/10.1038/s41588-022-01184-y.

75. Peng, Y., Yan, H., Guo, L., Deng, C., Wang, C., Wang, Y., Kang, L., Zhou, P., Yu, K., Dong, X., et al. (2022). Reference genome assemblies reveal the origin and evolution of allohexaploid oat. Nat. Genet. *54*, 1248–1258. https://doi.org/10.1038/s41588-022-01127-7.

76. Guo, L., Qiu, J., Ye, C., Jin, G., Mao, L., Zhang, H., Yang, X., Peng, Q., Wang, Y., Jia, L., et al. (2017). *Echinochloa crus-galli* genome analysis provides insight into its adaptation and invasiveness as a weed. Nat. Commun. *8*, 1031. https://doi.org/10.1038/s41467-017-01067-5.

77. Gaut, B.S. (2002). Evolutionary dynamics of grass genomes. New Phytol. *154*, 15–28. https://doi.org/10.1046/j.1469-8137.2002.00352.x.

78. Paterson, A.H., Bowers, J.E., and Chapman, B.A. (2004). Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc. Natl. Acad. Sci. USA *101*, 9903–9908. https://doi.org/10.1073/pnas.0307901101.

79. Dupont, E., Prochiantz, A., and Joliot, A. (2011). Penetratin story: an overview. Methods Mol. Biol. *683*, 21–29. https://doi.org/10.1007/978-1-60761-919-2_2.

80. Habault, J., and Poyet, J.L. (2019). Recent advances in cell penetrating peptide-based anticancer therapies. Molecules *24*, 927. https://doi.org/10.3390/molecules24050927.

81. Weber, B., Zicola, J., Oka, R., and Stam, M. (2016). Plant enhancers: a call for discovery. Trends Plant Sci. *21*, 974–987. https://doi.org/10.1016/j.tplants.2016.07.013.

82. Denker, A., and de Laat, W. (2015). A long-distance chromatin affair. Cell *162*, 942–943. https://doi.org/10.1016/j.cell.2015.08.022.

83. Clark, R.M., Wagler, T.N., Quijada, P., and Doebley, J. (2006). A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. Nat. Genet. *38*, 594–597. https://doi.org/10.1038/ng1784.

84. Castelletti, S., Tuberosa, R., Pindo, M., and Salvi, S. (2014). A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL *Vgt1*. G3 (Bethesda) *4*, 805–812. https://doi.org/10.1534/g3.114.010686.

85. Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K.A., Meeley, R., Ananiev, E.V., Svitashev, S., Bruggemann, E., et al. (2007). Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. Proc. Natl. Acad. Sci. USA *104*, 11376–11381. https://doi.org/10.1073/pnas.0704145104.

86. Du, Y., Liu, L., Peng, Y., Li, M., Li, Y., Liu, D., Li, X., and Zhang, Z. (2020). *UNBRANCHED3* expression and inflorescence development is mediated by *UNBRANCHED2* and the distal enhancer, *KRN4*, in maize. PLoS Genet. *16*, e1008764. https://doi.org/10.1371/journal.pgen.1008764.

87. Tian, J., Wang, C., Xia, J., Wu, L., Xu, G., Wu, W., Li, D., Qin, W., Han, X., Chen, Q., et al. (2019). Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. Science *365*, 658–664. https://doi.org/10.1126/science.aax5482.

88. Liu, J., Chen, J., Zheng, X., Wu, F., Lin, Q., Heng, Y., Tian, P., Cheng, Z., Yu, X., Zhou, K., et al. (2017). *GW5* acts in the brassinosteroid signalling pathway to regulate grain width and weight in rice. Nat. Plants *3*, 17043. https://doi.org/10.1038/nplants.2017.43.

89. Bai, X., Huang, Y., Hu, Y., Liu, H., Zhang, B., Smaczniak, C., Hu, G., Han, Z., and Xing, Y. (2017). Duplication of an upstream silencer of *FZP* increases grain yield in rice. Nat. Plants *3*, 885–893. https://doi.org/10.1038/s41477-017-0042-4.

90. Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., Dong, L., Zhang, Z., Yu, C., Sun, Y., et al. (2021). The Genome Sequence Archive family: toward explosive data growth and diverse data types. Genomics Proteomics Bioinformatics *19*, 578–583. https://doi.org/10.1016/j.gpb.2021.08.001.

91. Xue, Y., Bao, Y., Zhang, Z., Zhao, W., Xiao, J., He, S., Zhang, G., Li, Y., Zhao, G., and Chen, R. (2022). Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2022. Nucleic Acids Res. *50*, D27–D38. https://doi.org/10.1093/nar/gkab951.

92. Yang, J., Carena, M.J., and Uphaus, J. (2010). Area under the dry down curve (AUDDC): a method to evaluate rate of dry down in maize. Crop Sci. *50*, 2347–2354. https://doi.org/10.2135/cropsci2010.02.0098.

93. Zeng, Z.B. (1994). Precision mapping of quantitative trait loci. Genetics *136*, 1457–1468. https://doi.org/10.1093/genetics/136.4.1457.

94. Zeng, Z.B., Kao, C.H., and Basten, C.J. (1999). Estimating the genetic architecture of quantitative traits. Genet. Res. *74*, 279–289. https://doi.org/10.1017/s0016672399004255.

95. Yang, N., Liu, J., Gao, Q., Gui, S., Chen, L., Yang, L., Huang, J., Deng, T., Luo, J., He, L., et al. (2019). Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement. Nat. Genet. *51*, 1052–1059. https://doi.org/10.1038/s41588-019-0427-6.

96. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

97. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

98. Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M.G., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics *29*, 1035–1043. https://doi.org/10.1093/bioinformatics/btt087.

99. Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A., and Dewey, C.N. (2010). RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics *26*, 493–500. https://doi.org/10.1093/bioinformatics/btp692.

100. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

101. Axtell, M.J. (2013). ShortStack: comprehensive annotation and quantification of small RNA genes. RNA *19*, 740–751. https://doi.org/10.1261/rna.035279.112.

102. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25.

103. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36. https://doi.org/10.1186/gb-2013-14-4-r36.

104. Wang, H., Hou, J., Ye, P., Hu, L., Huang, J., Dai, Z., Zhang, B., Dai, S., Que, J., Min, H., et al. (2021). A teosinte-derived allele of a MYB transcription repressor confers multiple disease resistance in maize. Mol. Plant *14*, 1846–1863. https://doi.org/10.1016/j.molp.2021.07.008.

105. Song, S., Huang, H., Gao, H., Wang, J., Wu, D., Liu, X., Yang, S., Zhai, Q., Li, C., Qi, T., et al. (2014). Interaction between MYC2 and ETHYLENE INSENSITIVE3 modulates antagonism between jasmonate and ethylene signaling in *Arabidopsis*. Plant Cell *26*, 263–279. https://doi.org/10.1105/tpc.113.120394.

106. Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., and Xia, R. (2020). TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol. Plant *13*, 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009.

107. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. *20*, 238. https://doi.org/10.1186/s13059-019-1832-y.

108. Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. *26*, 1641–1650. https://doi.org/10.1093/molbev/msp077.

109. Jackson, D., Veit, B., and Hake, S. (1994). Expression of maize *KNOTTED1* related homeobox genes in the shoot apical meristem predicts patterns of morphogenesis in the vegetative shoot. Development *120*, 405–413. https://doi.org/10.1242/dev.120.2.405.

110. Liu, H.J., Jian, L., Xu, J., Zhang, Q., Zhang, M., Jin, M., Peng, Y., Yan, J., Han, B., Liu, J., et al. (2020). High-throughput CRISPR/Cas9 mutagenesis streamlines trait gene identification in maize. Plant Cell *32*, 1397–1413. https://doi.org/10.1105/tpc.19.00934.

111. Liu, Y., Cao, Y., Zhang, Q., Li, X., and Wang, S. (2018). A cytosolic triosephosphate isomerase is a key component in XA3/XA26-mediated resistance. Plant Physiol. *178*, 923–935. https://doi.org/10.1104/pp.18.00348.

112. Fu, X., Liu, C., Li, Y., Liao, S., Cheng, H., Tu, Y., Zhu, X., Chen, K., He, Y., and Wang, G. (2021). The coordination of OsbZIP72 and OsMYBS2 with reverse roles regulates the transcription of *OsPsbS1* in rice. New Phytol. *229*, 370–387. https://doi.org/10.1111/nph.16877.

113. Chen, T., Zhang, H., Liu, Y., Liu, Y.X., and Huang, L. (2021). EVenn: easy to create repeatable and editable Venn diagrams and Venn networks online. J. Genet. Genomics *48*, 863–866. https://doi.org/10.1016/j.jgg.2021.07.007.

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Polyclonal anti-ORF1	ABclonal	This paper
Rabbit Polyclonal anti-GFP	Proteintech	Cat # 50430-2-AP; RRID: AB_11042881
Mouse Monoclonal anti-actin	ABclonal	Cat # AC009; RRID: AB_2771701
Rabbit control IgG (IgG)	ABclonal	Cat # AC005; RRID: AB_2771930
Mouse Monoclonal anti-Rabbit IgG Antibody (M205) HRP	GenScript	Cat # A01827
HRP Goat Anti-Mouse IgG (H+L)	ABclonal	Cat # AS003; RRID: AB_2769851
HRP Goat Anti-Rabbit IgG	Servicebio	Cat # GB23303; RRID: AB_2811189
Anti-Digoxigenin-AP, Fab fragments	Roche	Cat # 11093274910; RRID: AB_514497
Bacterial and virus strains		
*Trans*1-T1 Phage Resistant Chemically Competent Cell	TransGen Biotech	Cat # CD501
Chemicals, peptides, and recombinant proteins		
Protease Inhibitor Cocktail	MCE	Cat # HY-K0010
RNase-free DNase I	Huayueyang Biotech	Cat # ZH0146
MUG	Sangon Biotech	Cat # A602251
Ethrel	Coolaber	Cat # CE5121
FM4-64	Thermo Fisher Scientific	Cat # T13320
DAPI	Thermo Fisher Scientific	Cat # 62248
ORF1p	This paper	N/A
scORF1p	This paper	N/A
FAM-ORF1p	This paper	N/A
Critical commercial assays		
Huayueyang plant RNA extraction kit	Huayueyang Biotech	Cat # ZH120
EasyScript one-step gDNA removal and cDNA synthesis supermix	TransGen Biotech	Cat # AE311
AceQ qPCR SYBR green master mix	Vazyme	Cat # Q111
SMARTer® RACE 5′/3′ kit	Takara Bio USA, Inc	Cat # 634858
BCA protein assay kit	Sangon Biotech	Cat # C503021
Tricine-SDS-PAGE gel preparation kit	Sangon Biotech	Cat # C641100
Pierce™ classic magnetic IP/Co-IP kit	Thermo Fisher Scientific	Cat # 88804
Dual-Luciferase® reporter assay system	Promega	Cat # E1980
SP6/T7 transcription kit	Roche	Cat # 10999644001
Deposited data		
sRNA-seq and Ribo-seq data	This paper	GSA: CRA013670, CRA013671
RNA-seq data	This paper	GSA: CRA013672, CRA013673
Experimental data of genetic mapping, fine mapping or agronomic traits	This paper	Zenodo: https://doi.org/10.5281/zenodo.13924714
Experimental models: Organisms/strains		
Arabidopsis thaliana Col-0	Lab stock	N/A
Arabidopsis thaliana Col-0 *p35S::ORF1-3×HA*/WT	This paper	N/A
Maize inbred line B73 wild-type	This paper	N/A
Maize inbred line B104 wild-type	This paper	N/A
Maize inbred line B104 *qkdr1*	This paper	N/A
Maize inbred line Zheng58 wild-type	This paper	N/A

(*Continued on next page*)

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Maize inbred line Zheng58 *qkdr1*	This paper	N/A
Maize inbred line B104 *pUbi*::RPG/WT	This paper	N/A
Maize inbred line B104 *pUbi*::*RPGm*/WT	This paper	N/A
Maize inbred line B104 *pUbi*::*ORF1-Flag*/WT	This paper	N/A
Maize inbred line B104 *pUbi*::*ORF1m-Flag*/WT	This paper	N/A
Maize inbred line B104 *microrpg1*	This paper	N/A
Maize inbred line B104 *orf2 & orf3*	This paper	N/A
Maize inbred line B104 *mybst1 & mybr43*	This paper	N/A
Maize inbred line KN5585 wild-type	This paper	N/A
Maize inbred line KN5585 *eil1*	This paper	N/A
Maize inbred line KN5585 *eil3*	This paper	N/A
Maize inbred line DAN340 wild-type	This paper	N/A
Maize inbred line K22 wild-type	This paper	N/A
Oligonucleotides		
Primers for fine mapping of *qKDR1*, see Table S7	This paper	N/A
Primers for qPCR, see Table S7	This paper	N/A
Primers for transgenic functional validation, see Table S7	This paper	N/A
Primers for sequence and haplotype analysis of *qKDR1*, see Table S7	This paper	N/A
Primers for 5'-RACE and 3'-RACE, see Table S7	This paper	N/A
Primers for vector construction, see Table S7	This paper	N/A
Primers for mRNA *in situ* hybridization, see Table S7	This paper	N/A
Recombinant DNA		
Plasmid: pUbiVector	Wang et al.[104]	N/A
Plasmid: pUbi2D	This paper	N/A
Plasmid: pUbi3D	This paper	N/A
Plasmid: pUbi2K	This paper	N/A
Plasmid: pUbi3K	This paper	N/A
Plasmid: pUbi4K	This paper	N/A
Plasmid: pUbi5K	This paper	N/A
Plasmid: pRPGVector	This paper	N/A
Plasmid: pRPG3D	This paper	N/A
Plasmid: pRPG3K	This paper	N/A
Plasmid: pRPG4K	This paper	N/A
Plasmid: pRPG5K	This paper	N/A
Plasmid: Effector-ZmMYBST1	This paper	N/A
Plasmid: Effector-ZmMYBR43	This paper	N/A
Plasmid: proGFP	This paper	N/A
Plasmid: proGFPm	This paper	N/A
Plasmid: proORF1	This paper	N/A
Plasmid: proORF1m	This paper	N/A
Plasmid: ORF1-GFPm	This paper	N/A
Plasmid: ORF1m-GFPm	This paper	N/A
Plasmid: ORF1m-GFP	This paper	N/A
Plasmid: sg-*qKDR1*	This paper	N/A
Plasmid: sg-*ZmMYBST1* & *ZmMYBR43*	This paper	N/A
Plasmid: sg-ORF1	This paper	N/A
Plasmid: sg-ORF2 &ORF3	This paper	N/A
Plasmid: sg-*ZmEIL1*	This paper	N/A

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Plasmid: sg-*ZmEIL3*	This paper	N/A
Plasmid: ORF1-OE	This paper	N/A
Plasmid: ORF1m-OE	This paper	N/A
Plasmid: RPG-OE	This paper	N/A
Plasmid: RPGm-OE	This paper	N/A
Plasmid: at-OE	This paper	N/A
Plasmid: SCAMP1-RFP	Liu et al.[111]	N/A
Plasmid: GHD7-mCherry	Fu et al.[112]	N/A
Software and algorithms		
E-Venn	Chen et al.[113]	http://www.ehbio.com/test/venn/#/
ImageJ	NIH	https://imagej.nih.gov/ij/
GraphPad Prism 9	GraphPad	https://www.graphpad.com/scientific-software/prism/
Bioedit	Bioedit	https://bioedit.software.informer.com/
MEGA X	MEGA X	https://www.megasoftware.net/
TBtools	Chen et al.[106]	https://github.com/CJ-Chen/TBtools/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Maize and growth conditions

Maize (*Zea mays* L.) seeds of *qKDR1*, microRPG1, *ZmEIL1* and *ZmEIL3* knockout lines and corresponding wild types were germinated on wet paper in 13 cm × 13 cm Petri dishes for seed germination rate assays. Seeds were grown in plant growth chamber at 26°C and 70% relative humidity, with a photoperiod of 16-h-light (10000 lx)/8-h-dark. Etiolated seedlings of maize inbred line B73 were sown in soil nutrition bowls and grown in plant growth chamber at 26°C, 70% relative humidity, and in the dark. Field-grown plants were grown and cultivated in the experimental field in Hainan (Sanya; 109.19°E, 18.38°N), Hubei (Wuhan; 114.32°E, 30.58°N), Henan (Xinxiang; 113.81°E, 35.20°N), Liaoning (Shenyang; 123.47°E, 41.68°N), Jilin (Gongzhuling; 124.83°E, 43.51°N) or Beijing (39.9°N, 116.3°E), China.

Arabidopsis and growth conditions

Arabidopsis thaliana Col-0 seeds were surface sterilized with 75% ethanol (1 min) and 95% ethanol (5 min), then plated on half-strength Murashige & Skoog (½MS) medium supplemented with sucrose and agar (Coolaber, Cat # PM10621-307), stratified for at least 3 days at 4°C, and then vertically grown in plant growth chamber at 22°C and 70% relative humidity, with a photoperiod of 16-h-light (10000 lx)/8-h-dark. After growing to 2∼4 true leaves, plants were transplanted into soil nutrition bowls to grow.

METHOD DETAILS

Plant materials

A recombinant inbred line (RIL) population of 201 lines derived from the cross between the maize inbred lines DAN340 and K22[49,50] were used to detect quantitative trait loci (QTLs) for kernel dehydration rate (KDR). A pair of near isogenic lines (NILs) were developed from heterogeneous inbred family (HIF) that was heterozygous at *qKDR1* region, NILs homozygous for DAN340 and K22 across the target region were designated NIL^{DAN340} and NIL^{K22}, respectively.

A total of 497 diverse maize inbred lines[55] were used for haplotype and genotype frequencies analyses. Out of these plant materials, a set of 339 maize inbred lines was used for ORF2 and ORF3 haplotype analyses; a set of 399 maize inbred lines was used for *qKDR1* haplotype analyses.

Phenotypic data collection and analysis

In the field, we used a GE BLD5604 moisture meter to measure the moisture content. In order to eliminate the influence of developmental period on KDR, we controlled that every plot was pollinated on the same day and the pollination date of each line was recorded. The moisture content was measured on the same interval days after pollination for each line and its comparison group. For QTL mapping, 17 plants with the same genotype were planted in one row in the field, five randomly chosen plants were measured in each row, and the average value represented the value of this genotype. For fine mapping and transgenic functional validation, all lines and their corresponding comparison groups were planted in neighboring rows in the field. The transient moisture content was measured in two successive stages before harvest, and the kernels were harvested after the second moisture content measurement, so the second moisture content was the moisture content at harvest (Figure S2O). The kernel moisture content of each plant at each

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

stage was measured twice, and the average value was taken as the phenotype of the kernel moisture content for this plant at that stage (*n* = sample size, a sample represents the phenotype from a plant). Kernel moisture content was obtained by measuring the kernels from the middle of the ear on the plant. Then, the area under the dry-down curve (AUDDC) was calculated to quantify the kernel dehydration rate.[21,92] A smaller AUDDC value represents a faster dehydration rate, while a larger AUDDC value represents a slower dehydration rate. To make a fair evaluation, we compared our method with the absolute moisture content obtained by the oven drying method which is typically performed in the lab. The absolute moisture content was calculated as (fresh weight - dry weight) / fresh weight. To evaluate the reliability of our method, all mature kernels from the middle of the ear were measured by oven drying method and meter reading (our method, using a GE BLD5604 moisture meter) in two NILs and found that the correlation coefficient (R^2) was 0.84, indicating that our method is reliable.

QTL mapping

The DAN340/K22 RIL population was planted in a randomized block design across five geographical locations in China: Hainan (Sanya; 109.19°E, 18.38°N) in 2013; Hubei (Wuhan; 114.32°E, 30.58°N), Henan (Xinxiang; 113.81°E, 35.20°N), Liaoning (Shenyang; 123.47°E, 41.68°N) and Jilin (Gongzhuling; 124.83°E, 43.51°N) in 2014. Moisture content of maize kernels was measured at four successive stages: 40, 46, 52 and 58 days after pollination (DAP), the area under the dry-down curve (AUDDC) is calculated to quantify the kernel dehydration rate. The best linear unbiased predictor (BLUP) value for each line was calculated using the mixed liner model to eliminate the influence of environmental effects. The BLUP values for each line were used as the phenotype to perform QTL mapping.

QTL mapping was conducted using composite interval mapping implemented in Windows QTL Cartographer version 2.5 (https://brcwebportal.cos.ncsu.edu/qtlcart/WQTLCart.htm).[93,94] Zmap (model 6) was used to detect QTL throughout the genome with a walking speed of 0.5-cM, and the default values for background controls of 5 for control markers, 10 for window size (cM) and forward regression method for regression method. A threshold for significant QTLs was determined by 1,000 permutations at a significance level of $P < 0.05$. The confidence interval for each QTL position was calculated using the 2-LOD drop of the peak. For *qKDR1*, the QTL confidence interval was 55.61 cM ~63.11 cM on Chromosome 1, corresponding to 16.75 Mb ~20.03 Mb.

Fine mapping of *qKDR1*

To fine-map *qKDR1*, heterozygous individuals derived from HIFs (Heterogeneous Inbred Families) were planted to screen new recombinant events.[54] The heterozygous region of HIFs for *qKDR1* was 16.81 Mb ~23.62 Mb on Chromosome 1, and the HIFs did not detect QTLs in other heterozygous intervals except for the *qKDR1* interval. We detected HIF parents in heterozygous regions except the *qKDR1* interval, and then self-pollinated and selected the offspring that were homozygous in other regions for fine mapping. Identification of new recombinants using flanking markers in the *qKDR1* region, and new molecular markers were developed to determine the breakpoints of identified recombinants. For each new recombinant, progeny tests were performed by comparing the KDR of NIL^{DAN340} and NIL^{K22} homozygous individuals from F3 families.[95] The significant difference of KDR between NIL^{DAN340} and NIL^{K22} homozygous individuals were compared using Student's *t* test. If NIL^{DAN340} and NIL^{K22} homozygous individuals show significant difference in KDR, the parental recombinant was assumed to be heterozygous for the target QTL; otherwise, the parental recombinant was homozygous. The QTL was narrowed down to a 1417-bp non-coding region (corresponding to B73 reference sequence version 4.0, Chr1:20007756-20009147) by integrating the QTL mapping information of all recombinants. The sequences of the primers used in fine mapping of *qKDR1* in this study are listed in Table S7.

RNA extraction and expression analysis

Total RNA was extracted from different maize tissues, including maize kernels (37DAP for microRPG1 knockout line, 40DAP for microRPG1 overexpression line and four NILs with different *qKDR1* genotypes, 45DAP for three haplotypes with different *qKDR1* genotypes from a diverse maize inbred line population and *qKDR1* knockout line, and 5DAP to 45DAP for *RPG* expression pattern of NIL^{DAN340} and NIL^{K22}) and maize leaves (seedlings), or *Arabidopsis* ripe siliques (72 days after sowing) using the Huayueyang plant RNA extraction kit (Huayueyang Biotech, Cat # ZH120) and treated with RNase-free DNase I (Huayueyang Biotech, Cat # ZH0146) to remove genomic DNA. The first strand cDNA was synthesized by *EasyScript* one-step gDNA removal and cDNA synthesis supermix (TransGen Biotech, Cat # AE311). qPCR was performed on the real-time PCR (CFX96 Real-Time System, Bio-Rad) by using AceQ qPCR SYBR green master mix (Vazyme, Cat # Q111). The expression levels of *RPG*, ORF1, *ZmEIL1*, *ZmEIL3*, *ZmPIF4*, *ZmEBF1*, *ZmLZF1*, *ZmERF1*, *ZmACO12*, *AtEIN3*, *AtEIL1*, *AtPIF3*, *AtEBF2*, *AtLZF1*, *AtERF1* and *AtACO12* were quantified and normalized to that of maize or *Arabidopsis ACTIN*, respectively. Each group or genotype contained at least two independent biological replicates, where one biological replicate (*n* = biological replicate size) is derived from one plant sample (maize kernels, maize leaves or *Arabidopsis* ripe siliques), and each replicate was measured by at least two technical replicates with similar results. Primers for quantifying the expression levels of *RPG*, ORF1, *ZmEIL1*, *ZmEIL3*, *ZmPIF4*, *ZmEBF1*, *ZmLZF1*, *ZmERF1*, *ZmACO12*, *AtEIN3*, *AtEIL1*, *AtPIF3*, *AtEBF2*, *AtLZF1*, *AtERF1* and *AtACO12* are listed in Table S7.

Field trials of agronomic, flowering time and yield-related traits

Field trials were used to measure agronomic and yield-related traits of two NILs and the gene-edited lines. All NILs and knockout lines and their corresponding wild-types were planted in neighboring rows in the field. For NILs, plants were grown in Hainan

(Sanya; 109.19°E, 18.38°N), China, in 2018. All plants were self-pollinated, flowering-related traits were recorded in flowering period, two agronomic traits were investigated after pollination, and five yield-related traits were measured after harvest (Figure 1B; Table S1). For gene-edited plants of *qKDR1* and its wild-type plants, plants were grown in Beijing (39.9°N, 116.3°E) and Jilin (Gongzhuling; 124.83°E, 43.51°N), China, in 2022. All plants were self-pollinated, flowering-related traits were recorded in flowering period, two agronomic traits were investigated after pollination, and five yield-related traits were measured after harvest (Table S3). For gene-edited plants of microRPG1 and its wild-type plants, plants were grown in Hainan (Sanya; 109.19°E, 18.38°N) in 2021, Beijing (39.9°N, 116.3°E) in 2022, and Jilin (Gongzhuling; 124.83°E, 43.51°N) in 2022, China, respectively. All plants were self-pollinated, flowering-related traits were recorded in flowering period, two agronomic traits were investigated after pollination, and five yield-related traits were measured after harvest (Table S6). N = sample size, a sample represents the phenotype derived from a plant. The large sample size differences between flowering time or agronomic traits and yield-related traits are due to self-pollination failure or poor pollination resulting in fewer plants being harvested in some cases.

Sequence and haplotype analysis of *qKDR1*

qKDR1 sequences of NIL^{DAN340} and NIL^{K22} were cloned through PCR using primers MCK57_3F and MCK57_R. Then, the sequences of NIL^{DAN340} and NIL^{K22} were determined using a series of primers (Table S7). Sequence comparisons and analyses were performed in MEGA X and BioEdit. Haplotype analysis of *qKDR1* was performed by three pairs of primers (Table S7), of which MCK57_3F and MCK57_R were used to identify the distribution of a 6181-bp insertion (InDel6181), QDR_F and QDR_R were used to identify the distribution of a 50-bp deletion (InDel50), and MCK57_7F and MCK57_R were used to identify the distribution of a 234-bp insertion (InDel234) in the diverse maize inbred line population.

Rapid-amplification of cDNA ends

The full-length sequences of *RPG* in NIL^{K22} and NIL^{DAN340} in the late stage of maize kernels (45DAP) were identified by the rapid-amplification of cDNA ends (RACE). RACE experiment was performed using SMARTer® RACE 5'/3' kit (Takara Bio USA, Inc, Cat # 634858) following the manufacturer's protocol. High quality total RNAs of NIL^{K22} and NIL^{DAN340} were extracted and used for the 5'-RACE and 3'-RACE with gene specific primers. The final products of *RPG* were verified by Sanger sequencing and compared with DNA reference sequence of maize (B73 version 4.0). The gene specific primers used in RACE were listed in Table S7.

RNA and small RNA sequencing

Total RNA was extracted from different maize tissues (maize kernels at the late maturation stage at 37DAP for microRPG1 knockout line, 40DAP for microRPG1 overexpression line and NILs) or *Arabidopsis* tissue (ripe siliques, 72 days after sowing) using the Huayueyang plant RNA extraction kit (Huayueyang Biotech, Cat # ZH120) and treated with RNase-free DNase I (Huayueyang Biotech, Cat # ZH0146) to remove genomic DNA. The total RNAs were used to perform RNA sequencing (RNA-seq) or small RNA sequencing (sRNA-seq). For RNA-seq, the Illumina NovaSeq 6000 and DNBSEQ-T7 platforms were used for sequencing with paired-end reads generation. Cleandata of RNA-seq was obtained and compared to the maize reference genome (version 4.0, http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-48/plants/fasta/zea_mays/dna) by RSEM v1.3.3[96] combined with Bowtie v2.4.1,[97] and the expression levels were calculated with maize annotation information (version 4.48, http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-48/plants/gtf/zea_mays). Then, the differentially expressed genes were calculated by EBSeq,[98,99] and the differentially expressed genes were selected with significances of $P < 0.05$ and $\geq$ 1.2 fold change for NILs, $P < 0.05$ and $\geq$ 1.5 fold change for microRPG1 knockout and overexpression, and $P < 0.05$ and $\geq$ 3.0 fold change for microRPG1 exogenous application. For sRNA-seq, the Illumina NovaSeq 6000 platform was used for sequencing with single-end reads generation. Cleandata of sRNA-seq was obtained by using Cutadapt,[100] and mapped to the maize reference genome (version 4.0, http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-48/plants/fasta/zea_mays/dna) by ShortStack[101] combined with Bowtie.[102]

Ribosome profiling sequencing

In order to investigate whether *RPG* functions by encoding protein(s), ribosome profiling sequencing (Ribo-seq) was performed in a late stage of maize kernel maturation (40DAP). Library construction and sequencing were performed by Novogene (Novogene Co., Ltd.), the Illumina NovaSeq 6000 platform was used for sequencing with single-end reads generation. Cleandata of Ribo-seq was obtained and mapped to the maize reference genome (version 4.0, http://ftp.ebi.ac.uk/ensemblgenomes/pub/release-48/plants/fasta/zea_mays/dna) by using TopHat v2.1.1.[103]

GFP assay and subcellular localization

GFP assay and subcellular localization were performed in maize protoplasts. For GFP assay, the protoplasts were cultured at 23°C in the dark for at least 16 h, and GFP fluorescence was imaged using a laser confocal microscope (Olympus, FV12000MPE). For subcellular localization of FAM-ORF1p in *Arabidopsis* and maize seedlings, images were collected using a laser confocal microscope (Olympus, FV12000MPE) in the presence of 10 μM FAM-ORF1p after incubation for 24 h. For subcellular localization of FAM-ORF1p in maize protoplasts, images were collected using a laser confocal microscope (Olympus, FV12000MPE) in the presence of 10 μM FAM-ORF1p after incubation for 24 h. For subcellular localization of ORF1-GFP in maize protoplasts, images were collected

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

using a laser confocal microscope (Zeiss, LSM980). The primers used for GFP assay and subcellular localization were listed in Table S7.

Western blotting

The plant tissues were frozen in liquid N_2 and ground into powder. The powder or cells were suspended in IP lysis buffer (Thermo Fisher Scientific, Cat # 87788) or SDS lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 2% SDS and 5% glycerol) with 1 mM PMSF and 1% protease inhibitor cocktail (MCE, Cat # HY-K0010), and then sonication was performed for 30 cycles, with 5 s on and 10 s off for each cycle (Diagenode, Bioruptor Plus). The lysates were then centrifuged at 12000 rpm for 10 min at 4°C. The protein concentrations were determined by BCA protein assay kit (Sangon Biotech, Cat # C503021). The total protein for detection of endogenously produced ORF1 micropeptide was extracted by PTM BioLab Co. Inc. For detection of endogenously produced ORF1 micropeptide or ORF1-GFP fusion protein, the proteins were denatured at 96°C for 5 min and separated using 4-20% Tris-Glycine-SDS-PAGE (Sangon Biotech, Cat # C651105) with ColorMixed protein marker (Solarbio, Cat # PR1930) or SeeBlue® Plus2 Pre-Stained standard (Thermo Fisher Scientific, Cat # LC5925). For detection of endogenously produced ORF1 micropeptide following immunoprecipitation (IP), the tissue lysates were denatured at 96°C for 5 min and separated using Tricine-SDS-PAGE gel preparation kit (Sangon Biotech, Cat # C641100) with Spectra Multicolor low range protein ladder (Thermo Fisher Scientific, Cat # 26628). The proteins were separated by SDS-PAGE, and then transferred onto a polyvinylidene fluoride (PVDF) membrane using a wet transfer apparatus (Bio-Rad Laboratories). Western blotting was performed using anti-GFP (1:3000) or anti-ORF1 (1:3000) antibodies. For the detection of actin in maize, western blotting was performed using anti-actin (1:5000) antibody. The antibodies used in western blotting were anti-GFP (Proteintech, Cat # 50430-2-AP), anti-actin antibody (ABclonal, Cat # AC009) and rabbit control IgG (IgG) (ABclonal, Cat # AC005). The polyclonal antibodies were generated in rabbit against a peptide (HSRAPTNKRAGLQNYC) from ORF1 micropeptide (ABclonal).

Co-immunoprecipitation (co-IP)

The ORF1 complexes in transgenic overexpression lines of full-length *RPG* (OEWT), transgenic overexpression line of ORF1 (OE) and wild-type plants (WT) were co-immunoprecipitated (co-IPed) using anti-ORF1 with IgG as a control at 4°C for 16 h. Co-IP experiments were performed using a Pierce™ classic magnetic IP/Co-IP kit (Thermo Fisher Scientific, Cat # 88804) following the manufacturer's protocol. Then, the complexes were separated using Tricine-SDS-PAGE for detection of endogenously produced ORF1 micropeptide via western blotting, and used for protein identification using mass spectrometry.

Transient transcriptional activity assays

The transient transcriptional activity assays were performed in maize protoplasts. For the GUS-LUC transient transcriptional activity assays, a series of fragments of *qKDR1* was generated and fused into the downstream of the β-glucuronidase (GUS) gene driven by a minimal *Ubiquitin* promoter (mpUbi), respectively. The firefly luciferase (LUC) gene was an internal control. For the dual-luciferase (LUC-REN) transient transcriptional activity assays, the 883-bp *RPG* promoter was fused into upstream of the firefly luciferase (LUC) gene in pGreenII 0800-LUC vector, and then named pVector. The Renilla luciferase (REN) gene was used as an internal control. The *qKDR1* fragments from NIL^{DAN340} and NIL^{K22} were fused into the downstream of the LUC gene in pVector, respectively. To evaluate the effects of *ZmMYBST1* and *ZmMYBR43* on the expression of *RPG* regulated by *qKDR1*, the coding sequence (CDS) of *ZmMYBST1* or *ZmMYBR43* was fused into the downstream of the *CaMV 35S* promoter in pGreenII 62-SK vector, and used as effectors. The empty pGreenII 62-SK vector was used as a control. The maize protoplasts were collected from the leaves of 10∼14-day-old etiolated seedlings of inbred line B73. The GUS signal was collected using 4-Methylumbelliferyl-β-D-glucuronic acid (MUG, Sangon Biotech) as a substrate, and was performed as described previously.[104,105] After transformation and incubation, the protein samples were divided into 3 × 20 μL replicates. GUS reaction buffer (0.03 M Na_2HPO_4, 0.02 M NaH_2PO4, 0.01 M EDTA, 0.1% Triton, 0.07% β-mercaptoethanol, 0.1% MUG) was pre-warmed at 37°C, and 150 μL was added to all replicates at the same time. Then, each mixture was divided into 3 × 50 μL replicates. After incubation at 37°C for 1∼5 h, 200 μL 0.2 M Na_2CO_3 was added to terminate the reaction. The fluorescence signal of the final mixture was detected at 465 nm under excitation at 355 nm. The LUC and REN signals were collected using Dual-Luciferase® reporter assay system (Promega, Cat # E1980) following the manufacturer's protocol. Relative GUS activity was calculated by normalizing GUS activity to LUC activity, and relative LUC activity was calculated by normalizing LUC activity to REN activity. The primers used in the transient transcriptional activity assays were listed in Table S7.

Micropeptide synthesis

The synthetic micropeptides (purity > 95%) were synthesized by Genscript (https://www.genscript.com.cn). The micropeptides were dissolved in water, and stored at −80°C until use. ORF1p: MILLPNHSRAPTNKRAGLQNYCPQRSCIRRG, scORF1p: NTRQCCGANL PQYRRIHLARLRGSMPNPKIS, 5-FAM fluorescein-labeled microRPG1 (FAM-ORF1p).

Identification of mature phenotype in *Arabidopsis*

Arabidopsis seeds were surface sterilized with 75% ethanol and 95% ethanol, then plated on half-strength Murashige & Skoog (½MS) medium supplemented with sucrose and agar (Coolaber, Cat # PM10621-307), stratified for at least 3 days at 4°C, and then vertically grown in plant growth chamber at 22°C and 70% relative humidity, with a photoperiod of 16-h-light/8-h-dark.

Treatment groups adopted half-strength MS medium containing different concentrations (0.01 to 2.0 μM) of synthetic ORF1 micropeptide (ORF1p) or 2.0 μM scORF1p (with the same amino acid composition as the ORF1p but a scrambled sequence), and the control group replaced the micropeptides with water. After growing to 2∼4 true leaves, plants were transplanted into soil nutrition bowls to grow.

Treatment groups of *Arabidopsis* were exogenously watered and sprayed with ORF1p or scORF1p until the siliques ripened, while the control group used water. The date of flowering, and the date when siliques began to ripen and the date when 50% siliques ripening were recorded. After harvesting the seeds, the fresh weight and dry weight after drying at 105°C were measured, and then the moisture content was calculated.

Micropeptide absorption in *Arabidopsis*

2.0 μM and 10.0 μM FAM-ORF1p was prepared and incubated with *Arabidopsis* seedlings for 0.5 h∼24 h. After the treatment, seedlings were rinsed with ddH_2O for 5∼6 times. The fluorescence of FAM in plants was observed by fluorescence microscope (Nikon, SMZ25). FAM fluorescence was excited with a 492 nm laser and recorded in the 518 nm emission range. ImageJ was used to analyze the fluorescence images.

Synteny analysis

Synteny analysis of the region encompassing ∼20 flanking genes on each side of microRPG1 was performed in Poaceae plants (the genomes were from Ensembl Plants (http://plants.ensembl.org/index.html)) except the genus *Zea* and *Tripsacum* by using TBtools.[106] Synteny analysis of the region encompassing two flanking genes of microRPG1 was performed in the genus *Zea* and *Tripsacum* (the genomes were from maize-pangenome (https://maize-pangenome-ensembl.gramene.org) and MaizeGDB (https://www.maizegdb.org)). The 96-bp sequence of microRPG1 with 20-bp of 5' and 3' flanking sequence or with 300-bp of 5' flanking sequence and 200-bp of 3' flanking sequence was compared with the syntenic sequences between the two flanking genes by using TBtools ($P < 0.01$). Bioedit was used to compare the microRPG1 and its homologous sequences in the genus *Zea* and *Tripsacum*. The divergence time of species refers to several published studies.[74–78]

Phylogenetic tree

The protein files for Poaceae plants in the tree were obtained from Ensembl Plants (http://plants.ensembl.org/index.html) and MaizeGDB (https://www.maizegdb.org). Subsequently, the identification of single-copy orthologous proteins was performed using OrthoFinder v2.2.6,[107] after selecting the longest transcript for each gene. The construction of the species phylogenetic tree was carried out using FastTree v2.1.10.[108]

For EIN3/EIL in maize, rice and *Arabidopsis*, blastp was used to align ZmEIL1 homologous sequences in National Center for Biotechnology Information (NCBI, https://blast.ncbi.nlm.nih.gov) with reference proteins (refseq_protein). The homologous sequences of ZmEIL1 in maize, rice and *Arabidopsis* were aligned by MUSCLE in MEGA X. The phylogenetic tree was constructed using the maximum likelihood method in MEGA X.

Identification of differentially expressed genes with similar expression patterns to *RPG*

The overlapping differentially expressed genes form microRPG1 knockout, overexpression and exogenous application were used to calculate Z-scores and the genes were considered as specific expression in the late stages of seed (kernel) and endosperm if they had a Z-score of 2 or higher. The Z-scores of the genes expression in 38DAP seed (kernel) and 38DAP endosperm were calculated, and the expression of genes in all tissues (except for the seed and endosperm of 26DAP and 32DAP) of seed, endosperm and embryo from B73[58,59] was used to calculate Z-scores.

mRNA *in situ* hybridization

The 30DAP kernels of maize inbred line B104 were fixed in a 4% polyformaldehyde solution (4g paraformaldehyde dissolved in 100 mL 1 × PBS, pH 6.5∼7.0) and used for mRNA *in situ* hybridization. The kernels were dehydrated in a series of ethanol concentrations and embedded, then sectioned to a thickness of 5 μm. To generate probes for *ZmMYBST1*, *ZmMYBR43*, *RPG*, *ZmEIL1* and *ZmEIL3*, T7 promoter sequence (CATTAATACGACTCACTATAGGG) was added into reverse primers and probe fragments were amplified by PCR using gene specific primers listed in Table S7. Digoxigenin-labeled antisense probes were then transcribed using an *in vitro* transcription kit (Roche, Cat # 10999644001) following the manufacturer's protocol. RNA hybridization and immunologic detection were performed as described previously.[109] The hybridized signals were observed by microscope (Nikon, SMZ25).

Immunohistochemistry

The 30DAP kernels of maize inbred line B104 were dehydrated in a series of ethanol concentrations and embedded, then sectioned to a thickness of 5 μm. The sections were dewaxed and rehydrated, followed by antigen repair. Next, the tissues were incubated in 3% hydrogen peroxide solution for 25 min at room temperature to block endogenous peroxidase, then blocked in 3% BSA for 30 min at room temperature. The tissues were incubated with anti-ORF1 antibody (1:300) overnight at 4°C and subsequently incubated with the corresponding HRP conjugated secondary IgG antibody for 50 min at room temperature. Cellular nuclei were stained with hematoxylin. Immunohistochemistry was performed by Wuhan Servicebio Technology Co., Ltd.

Please cite this article in press as: Yu et al., A *Zea* genus-specific micropeptide controls kernel dehydration in maize, Cell (2025), https://doi.org/10.1016/j.cell.2024.10.030

Determination of kernel oil content

Maize kernels were dried thoroughly and ground into powder. About 0.2g was weighed into a test tube and the weight recorded. Subsequently, 4.5 mL of sulfuric acid methanol solution was added to each test tube, and the mixture was shaken and mixed well, and then 100 μL internal standard was added. The solution was incubated in a water bath at 90°C for 2.5 h∼3 h. The sample was subsequently cooled to room temperature, 2 mL ultrapure water was added and the mixture was shaken well, then 2 mL n-hexane was added and the mixture was shaken well. Following this, the mixture was centrifuged at 3000 rpm for 10 min at room temperature and the supernatant was taken. The oil contents were subsequently determined by liquid chromatography.

Determination of kernel starch content

Maize kernels were dried thoroughly and ground into powder. About 0.02g was weighed into a test tube and the weight recorded. Subsequently, 1 mL anhydrous ether was added and the mixture was shaken well, then centrifuged at 8000 *g* for 10 min at room temperature and the supernatant was discarded. Subsequently, 1 mL 80% ethanol was added to the precipitate and shaken well. The mixture was incubated in a water bath at 80°C for 30 min, prior to centrifugation at 8000 *g* for 10 min at room temperature, after which the supernatant was discarded. Then 0.5 mL of ultrapure water was added to the precipitate and the mixture boiled in a water bath for 15 min to fully gelatinize the sample. After cooling, 0.5 mL of 6 mol/L hydrochloric acid was added and the mixture was heated at 95°C for 30 min, prior to being centrifuged at 8000 *g* for 10 min at room temperature, with the supernatant taken for measurement. After diluting with ultrapure water, 50 μL of supernatant was taken into a 1.5 mL test tube. 250 μL of anthrone reagent was added and mixed well prior to incubation in a water bath at 95°C for 10 min. After cooling to room temperature, the absorption at 620 nm was detected. The determination of kernel starch content was performed by Wuhan ProNets Biotechnology Co., Ltd.

Seed germination rate assays

Maize seeds of each genotype were immersed in water for about 8 h, and then transferred to wet paper in a 13 cm × 13 cm Petri dish for germination. Seeds were grown in a plant growth chamber at 26°C and 70% relative humidity, with a photoperiod of 16-h-light/8-h-dark. Seed germination was considered complete when the root length was > 2 mm. The number of germinated seeds was counted 88 h after sowing. Each genotype contained 24∼36 seeds per biological replicate, and three biological replicates were performed.

Exogenous ethylene treatment

The moisture content of maize kernels was investigated before exogenous ethylene treatment (pre-treatment), and then 277 mM ethrel (Coolaber, Cat # CE5121) was prepared and used to treat maize kernels in the late stage of kernel maturation in F1 of SK × KN5585, F1 of SK × Zheng58, and KN5585 in the field. In detail, the ear husks were peeled back, then 10 mL was sprayed onto the kernels on each ear, followed by covering the ear with a paper bag, and this procedure was repeated after the solution evaporated (∼12 hours later). Four days post-treatment, the moisture content of maize kernels was investigated again to evaluate the effects of ethylene treatment.

Transgenic functional validation

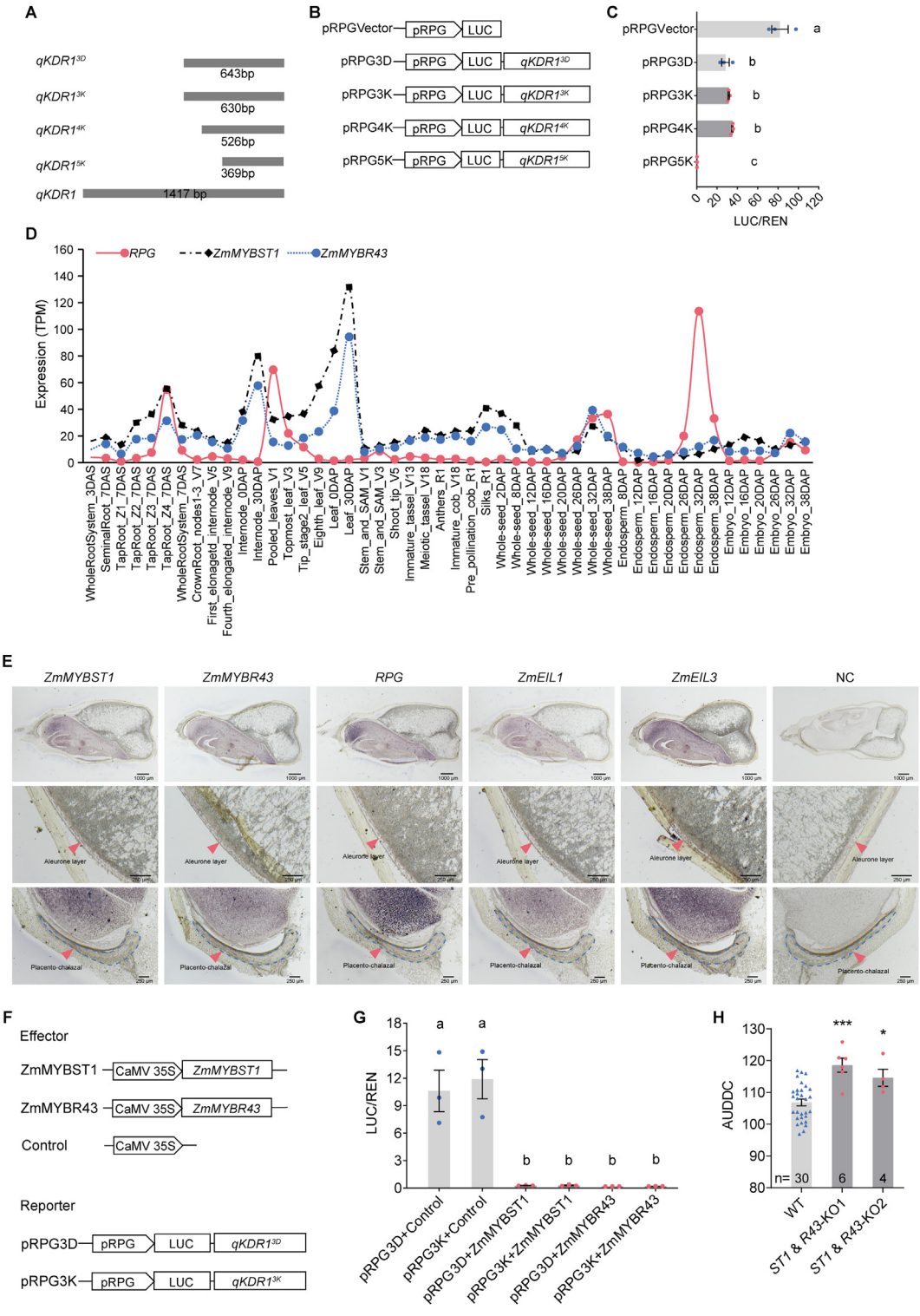
To generate overexpression lines, the related sequences driven by the *Ubiquitin* promoter were transformed into maize inbred line B104 at Wimi Biotechnology (Changzhou, China). The transgenic-positive and transgenic-negative plants were identified in each generation, and qPCR was performed to determine the expression levels of transgenes in the overexpression lines. The CRISPR-Cas9 technique was used to generate knockout lines. The target sites were designed and cloned into pCXB053 vector, and then transformed into maize inbred line B104 or KN5585 at Wimi Biotechnology.[110] The knockout lines were genotyped by PCR and sequencing to identify the mutant plants. The knockout lines with homozygous mutations were used for phenotypic analysis and field trials. The primers used in transgenic functional validation were listed in Table S7. All the overexpression and knockout lines and their corresponding wild-type were planted in neighboring rows in field and the moisture content of maize kernels was measured.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data for quantification analyses are represented as mean ± SEM. Significance of difference was examined by Student's *t* test (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$) or one-way ANOVA followed by Fisher's LSD multiple-comparison test. Different letters at top of columns indicate significant differences at $P < 0.05$. The statistical analyses for all experiments were performed with the GraphPad Prism 9 software (https://www.graphpad.com/scientific-software/prism/). The exact value of *n*, what *n* represents and the statistical details of each experiment are described in the figure legends or tables.

CellPress
OPEN ACCESS

Supplemental figures



(legend on next page)

Figure S1. ZmMYBST1 and ZmMYBR43 proteins target the *qKDR1* region to decrease *RPG* expression, related to Figure 2

(A) *qKDR1* fragments used to generate constructs.

(B) Schematic diagram of constructs used for the transient transcriptional activity assays.

(C) LUC and REN activity were detected via transient transcriptional activity assays performed in maize protoplasts (n = 3).

(D) *ZmMYBST1* and *ZmMYBR43* displayed similar expression pattern with *RPG*. *RPG*, *ZmMYBST1*, and *ZmMYBR43* expression in various tissues from B73 by RNA-seq.[58,59] Pink dots, black diamonds, and blue dots represent *RPG*, *ZmMYBST1*, and *ZmMYBR43* expression, respectively.

(E) mRNA *in situ* hybridization pattern of *ZmMYBST1*, *ZmMYBR43*, *RPG*, *ZmEIL1*, and *ZmEIL3* in 30 days after pollination (DAP) kernels of B104. The top, middle, and bottom rows show representative section images of the kernel, aleurone layer, and placento-chalazal region, respectively. Pink arrowheads point to the aleurone layer or placento-chalazal region. Blue dashed lines mark the placento-chalazal regions. NC, negative control, the probe was replaced with ultrapure water. The hybridization signals appear in purple. Scale bars, 1,000 μm (the top row) and 250 μm (the middle and bottom rows).

(F) Schematic diagram of the effectors and reporters used for the transient transcriptional activity assays.

(G) ZmMYBST1 and ZmMYBR43 decrease *RPG* expression. LUC and REN activity were detected via transient transcriptional activity assays and performed in maize protoplasts (n = 3).

(H) *ZmMYBST1* and *ZmMYBR43* knockouts had slower KDR in field trials performed in Jilin in 2023 (23JL). *ST1* and *R43* represents a double mutation of *ZmMYBST1* and *ZmMYBR43*.

Data are represented as mean ± SEM. Different letters at top of columns indicate significant differences at $p < 0.05$ (one-way ANOVA followed by Fisher's LSD multiple-comparison test). *$p < 0.05$, ***$p < 0.001$ (Student's t test).

n = biological replicate size, a biological replicate is an independent maize protoplast transformation experiment (C and G). n = sample size, a sample represents the phenotype from a plant (H).

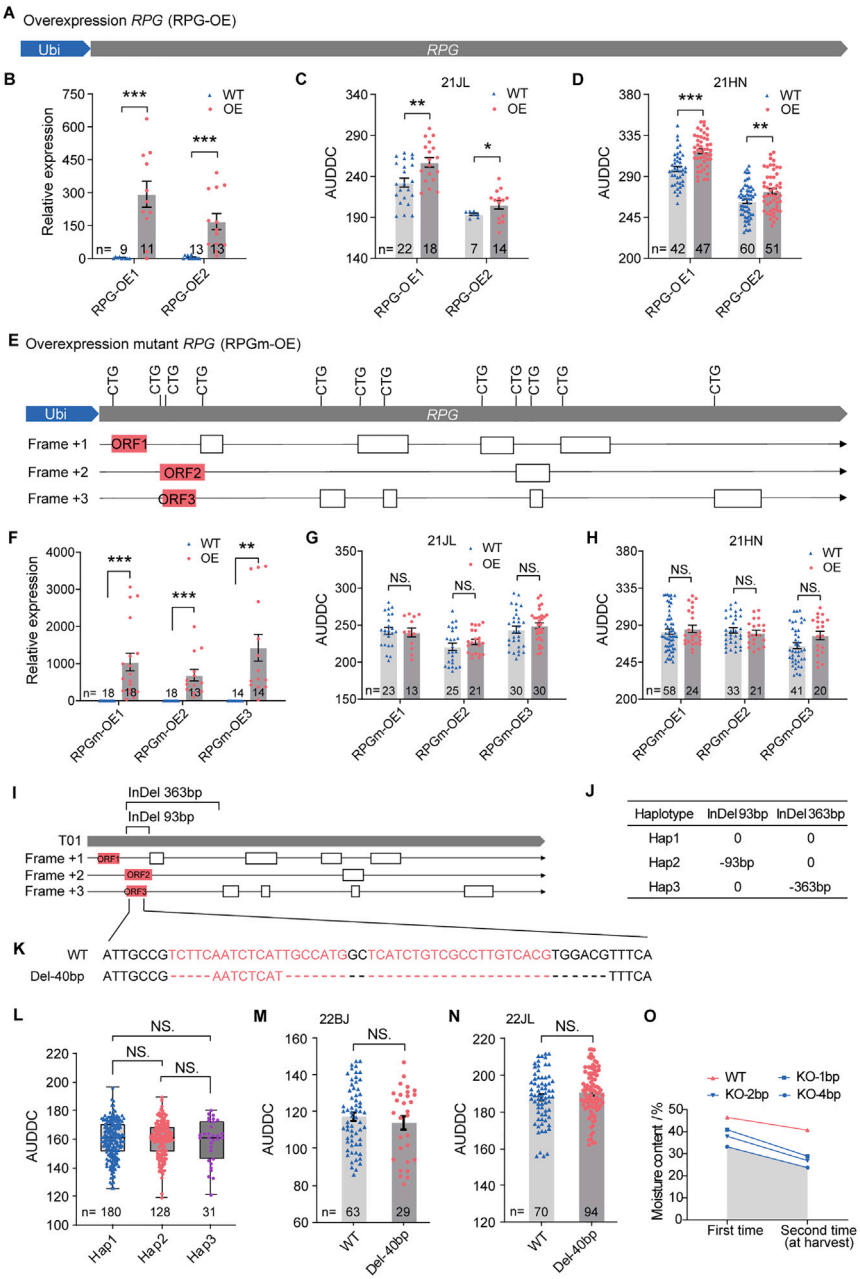


Figure S2. *RPG* does not function as a lncRNA, and ORF2 and ORF3 are not functional, related to Figure 3

(A) Construct used for *RPG* overexpression. Expression of the 2013-bp full-length *RPG* is driven by the ubiquitin promoter.

(B) *RPG* expression in two *RPG* overexpression lines (RPG-OE). The expression levels of *RPG* were quantified using qPCR and normalized to maize *ACTIN* ($n \geq 9$).

(C and D) *RPG* overexpression slowed down KDR in maize in field trials performed in Jilin in 2021 (21JL, C) and Hainan in 2021 (21HN, D), respectively.

(E) Construct used for mutant *RPG* overexpression. All 12 ORF start codons on *RPG* are mutated (the start codon ATGs are mutated to CTGs). Expression of the 2,013-bp full length of mutant *RPG* is driven by the ubiquitin promoter.

(F) Mutant *RPG* expression in three transgenic overexpression lines (RPGm-OE). The expression levels of mutant *RPG* were quantified using qPCR and normalized to maize *ACTIN* ($n \geq 13$).

(G and H) Mutant *RPG* overexpression did not change KDR in maize in field trials performed in Jilin in 2021 (21JL, G) and Hainan in 2021 (21HN, H), respectively.

(I) Two deletions in the ORF2 and ORF3 regions (indel 93 bp and indel 363 bp).

(J) 339 maize inbred lines were divided into three haplotypes.

(K) Sequences of knockout line generated by CRISPR-Cas9 technique (Del-40 bp). The target sites are highlighted with pink. The dashed lines indicate deletions.

(L) No phenotypic difference among the three haplotypes in the diverse maize inbred line population.

(M and N) Knockout of ORF2 and ORF3 do not change KDR in maize in field trials performed in Beijing in 2022 (22BJ, M) and Jilin in 2022 (22JL, N), respectively.

(legend continued on next page)

(O) The dry-down curves corresponding to Figure 3H. The transient moisture content was measured in two successive stages before harvest, and the kernels were harvested after the second moisture content measurement was completed, so the second moisture content was the moisture content at harvest. The shaded area represents the AUDDC of 4-bp deletion mutation in ORF1 (KO-4 bp).

Data are represented as mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, NS, not significant (Student's t test).

n = biological replicate size, a biological replicate is from a plant sample (maize leaves) (B and F). *n* = sample size, a sample represents the phenotype from a plant (C, D, G, H, M, and N) or from a plot (average of multiple plants of a row) (L).

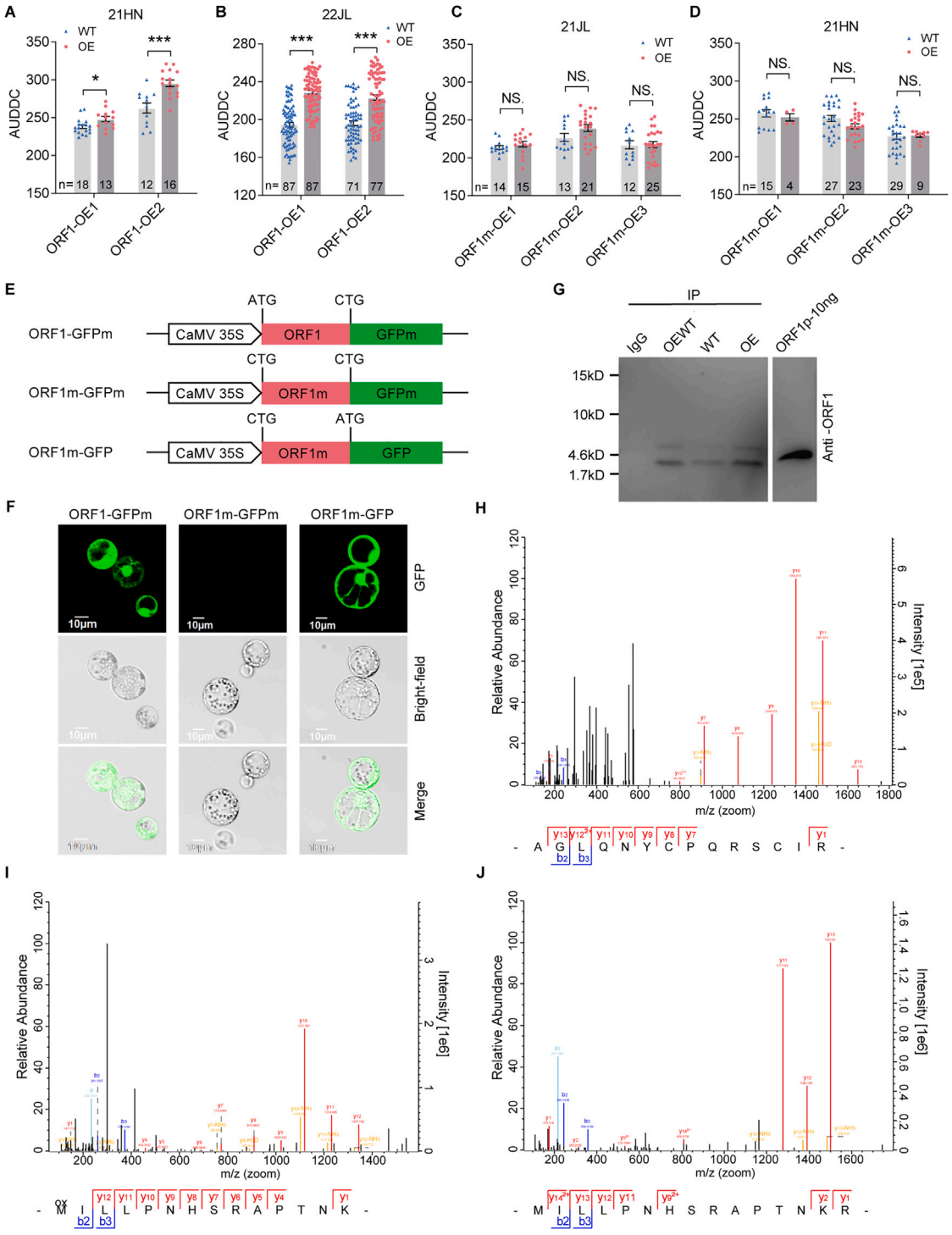


Figure S3. ORF1 encodes a microRPG1 micropeptide, related to Figure 4

(A and B) ORF1 overexpression slowed down KDR in maize in field trials performed in Hainan in 2021 (21HN, A) and Jilin in 2022 (22JL, B), respectively.

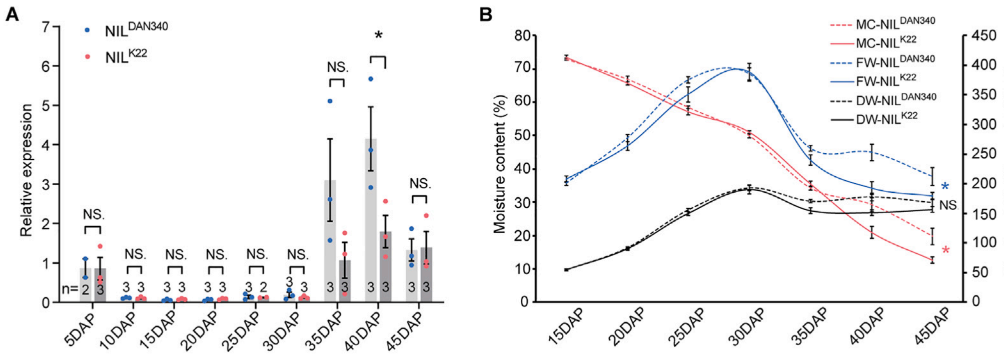
(C and D) ORF1m overexpression did not change KDR in maize in field trials performed in Jilin in 2021 (21JL, C) and Hainan in 2021 (21HN, D), respectively.

(E) Schematic diagram of different constructs used for the GFP assay (*CaMV 35S* promoter). ORF1-GFPm, ORF1m-GFPm, and ORF1m-GFP were analyzed in maize protoplasts. ORF1-GFPm, ORF1m-GFPm, and ORF1m-GFP represent constructs of 35S::ORF1-GFPm, 35S::ORF1m-GFPm, and 35S::ORF1m-GFP. 35S, *CaMV 35S* promoter. GFPm, the GFP start codon ATG is mutated to CTG. ORF1m, the ORF1 start codon ATG is mutated to CTG.

(F) Representative images of GFP array analysis in maize protoplasts (*CaMV 35S* promoter). GFP fluorescence (green), bright-field, and merged images were collected using a laser confocal microscope. Scale bars, 10 μm.

(legend continued on next page)

(G) The natural, endogenous ORF1 micropeptide is verified by immunoprecipitation. The ORF1 micropeptide was immunoprecipitated from maize kernel extracts by anti-ORF1 antibody and was detected by western blotting. IgG, rabbit control IgG. OEWT, overexpression of full-length *RPG*. OE, overexpression of ORF1. ORF1-10 ng, 10 ng of synthetic ORF1 micropeptide. The blank space in the middle eliminated irrelevant lanes.
(H–J) The endogenous ORF1 micropeptide in maize kernel was identified by mass spectrometry (MS) following immunoprecipitation (IP). Three unique peptides identified by mass spectrometry are shown.
Data are represented as mean ± SEM. $*p < 0.05$, $***p < 0.001$, NS, not significant (Student's t test).
n = sample size, a sample represents the phenotype from a plant.



C

Gene	Function	WT (ORF1-KO)	ORF1-KO	WT (ORF1-OE)	ORF1-OE	Z-score (seed38)	Z-score (endosperm38)
Zm00001d003554	22.0 kDa heat shock protein	238.64	1009.59	4405.04	1200.56	2.04	3.13
Zm00001d042225	Heavy metal transport/detoxification superfamily protein	132.20	259.78	501.47	275.81	2.01	2.21
Zm00001d047563	EIL1; Ethylene-Insensitive3-like 1	57.63	99.35	140.81	74.86	2.18	2.96
Zm00001d042884	Acyl-coenzyme A oxidase	24.48	48.52	83.19	48.48	2.12	2.42
Zm00001d009494	Putative MATE efflux family protein	19.49	113.15	144.34	54.08	2.54	2.11
Zm00001d019415	Cytochrome P450 71B10	19.51	71.18	111.12	57.03	2.58	2.43
Zm00001d044232	Protein EARLY FLOWERING 3	16.50	32.35	49.12	24.54	2.52	2.13
Zm00001d033797	alpha/beta-Hydrolases superfamily protein	16.10	37.03	114.76	49.57	2.02	2.80
Zm00001d050899	Cell number regulator 2	6.72	38.47	116.16	42.63	2.85	2.26
Zm00001d003931	ACR4	7.33	29.18	61.37	26.47	2.80	2.03
Zm00001d013406	alpha/beta-Hydrolases superfamily protein	7.13	21.11	65.84	25.70	2.16	2.92

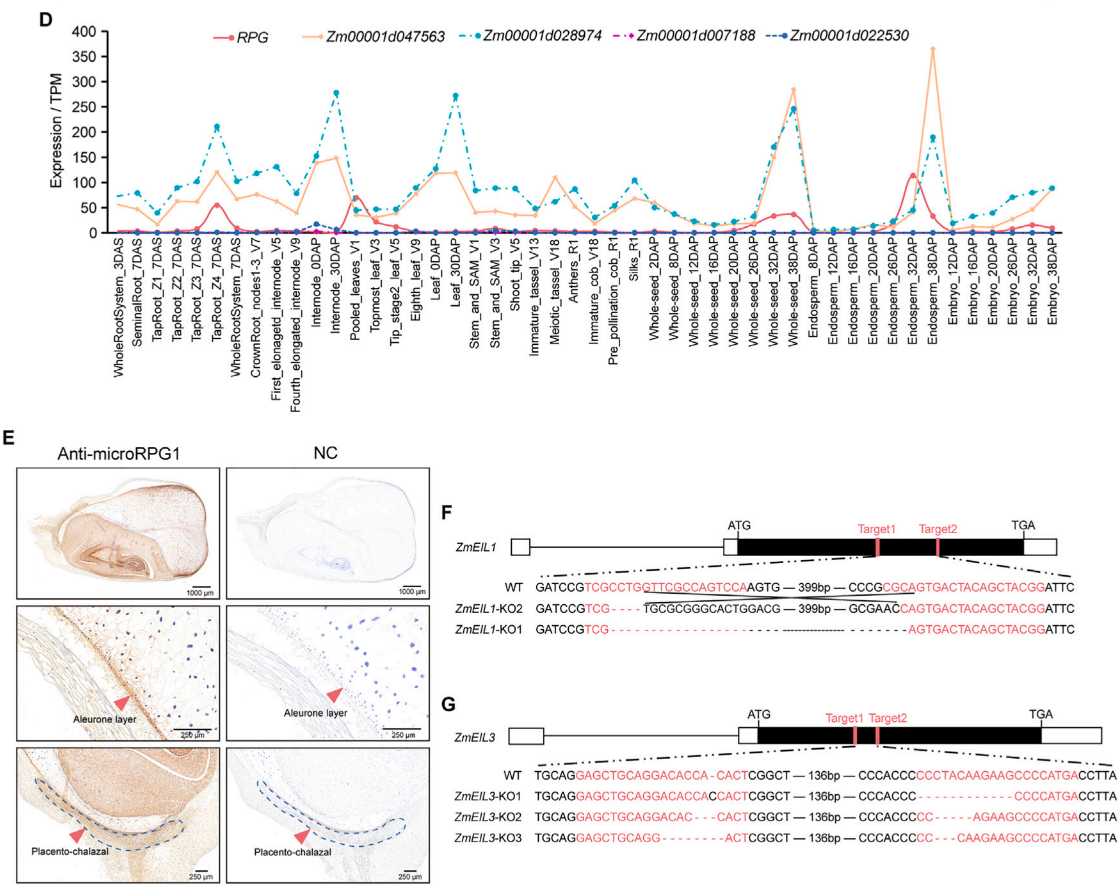


Figure S4. microRPG1 may control kernel dehydration through regulation of ethylene signaling, related to Figure 5

(A) *RPG* expression pattern of NIL^{DAN340} and NIL^{K22} in maize kernels. The expression levels of *RPG* were quantified using qPCR and normalized to maize *ACTIN*. DAP, days after pollination.

(B) Dynamic curves of absolute moisture content, fresh weight, and dry weight of NIL^{DAN340} and NIL^{K22} in maize kernels (n = 6). DAP, days after pollination. MC-NIL^{DAN340}, MC-NIL^{K22}, FW-NIL^{DAN340}, FW-NIL^{K22}, DW-NIL^{DAN340}, and DW-NIL^{K22} represent absolute moisture content of NIL^{DAN340}, absolute moisture

(legend continued on next page)

content of NIL^{K22}, fresh weight of NIL^{DAN340}, fresh weight of NIL^{K22}, dry weight of NIL^{DAN340}, dry weight of NIL^{K22}, respectively. Asterisks and NS, represent the significance of NIL^{DAN340} and NIL^{K22} at harvest (45 DAP). Absolute moisture content = (fresh weight − dry weight)/fresh weight.

(C) The differential expression genes with similar expression pattern to *RPG* in seed (kernel) and endosperm were screened by *Z* score. Show results with *Z* score $\geq$ 2.

(D) Expression patterns of *ZmEIL1* (*Zm00001d047563*), *ZmEIL2* (*Zm00001d007188*), *ZmEIL3* (*Zm00001d028974*), and *ZmEIL6* (*Zm00001d022530*). The expression levels of *EIL* genes in various tissues from B73 by RNA-seq.[58,59]

(E) Representative immunohistochemistry images of microRPG1 micropeptide expression in 30 DAP kernels of B104. The top, middle, and bottom rows show representative section images of the kernel, aleurone layer, and placento-chalazal region, respectively. Pink arrowheads point to the aleurone layer or placento-chalazal region. Blue dashed lines mark the placento-chalazal regions. NC, negative control, the primary antibody was replaced with PBS. The immunohistochemical signals appear in brown. Scale bars, 1,000 (the top row) and 250 μm (the middle and bottom rows).

(F and G) Sequences of *ZmEIL1* (F) and *ZmEIL3* (G) knockout lines generated by CRISPR-Cas9 technique with deletions, insertion, and inversion in the target sites. The target sites are highlighted with pink. The dashed lines indicate deletions. There is an inversion in *ZmEIL1*-KO2.

Data are represented as mean ± SEM. *$p < 0.05$, NS, not significant (one-tailed Student's t test).

n = biological replicate size, a biological replicate is from a plant sample (maize kernels) (A). *n* = sample size, a sample represents the phenotype from a plant (B).

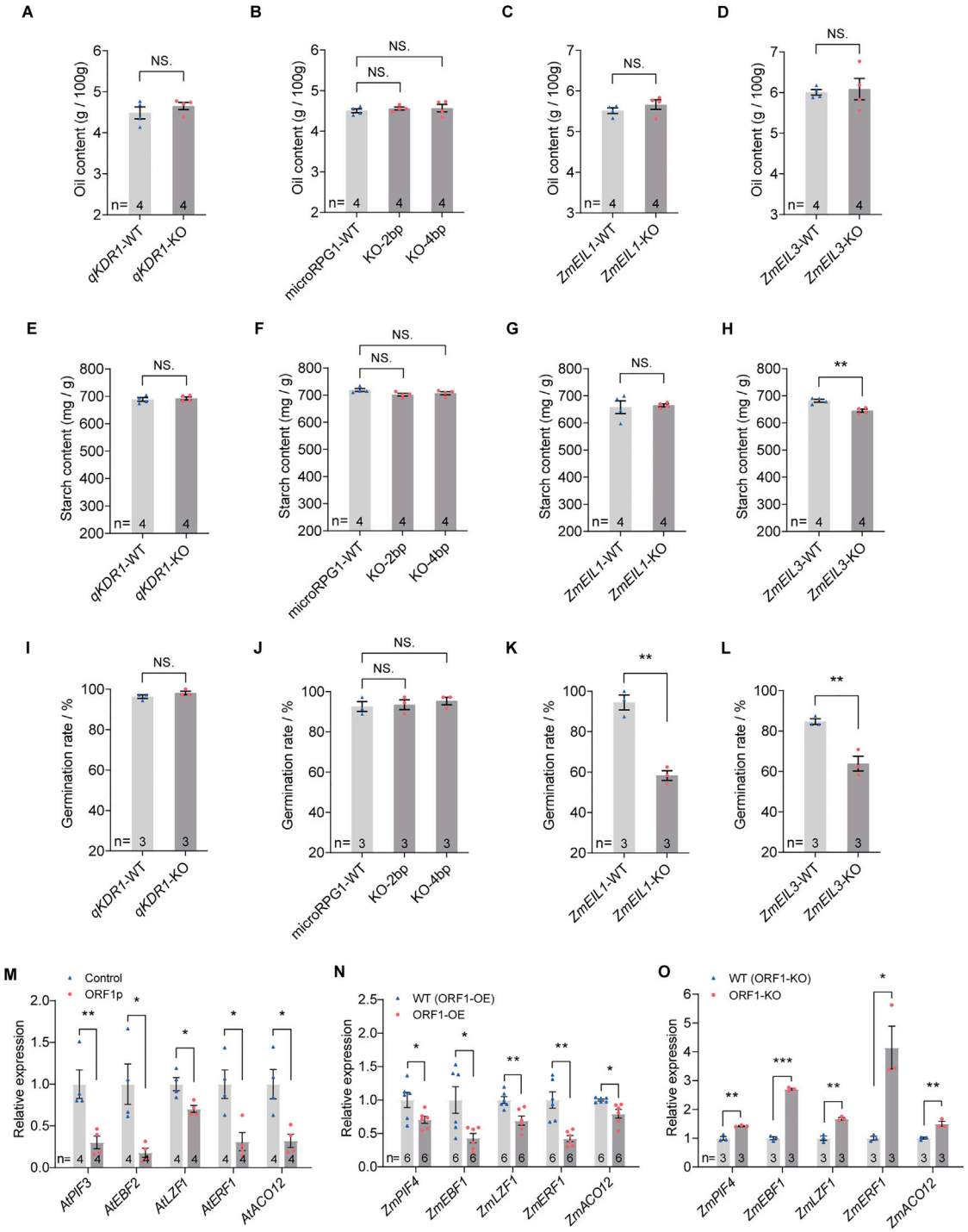


Figure S5. The effects of *ZmEIL1* or *ZmEIL3* knockout on kernel quality traits and seed germination rate, and microRPG1 regulates ethylene signaling, related to Figure 5

(A–D) Kernel oil contents of *qKDR1* knockout and corresponding wild type (A), microRPG1 knockout and corresponding wild type (B), *ZmEIL1* knockout and corresponding wild type (C), and *ZmEIL3* knockout and corresponding wild type (D) (n = 4).

(E–H) Kernel starch contents of *qKDR1* knockout and corresponding wild type (E), microRPG1 knockout and corresponding wild type (F), *ZmEIL1* knockout and corresponding wild type (G), and *ZmEIL3* knockout and corresponding wild type (H) (n = 4).

(I–L) Seed germination rates of *qKDR1* knockout and corresponding wild type (I), microRPG1 knockout and corresponding wild type (J), *ZmEIL1* knockout and corresponding wild type (K), and *ZmEIL3* knockout and corresponding wild type (L) (n = 3).

(legend continued on next page)

(M–O) Relative expression levels of ethylene signaling marker genes in following exogenous micropeptide application and corresponding control plants (treated with water (control) or 2.0 μM of ORF1p, ripe siliques of 72 days after sowing) in *Arabidopsis* (n = 4) (M), microRPG1 overexpression line (ORF1-OE, 40 DAP kernels) and corresponding wild type in maize (n = 6) (N) and microRPG1 knockout line (ORF1-KO, 37 DAP kernels) and corresponding wild type in maize (n = 3) (O). The expression levels of ethylene signaling marker genes were quantified using qPCR and normalized to maize or *Arabidopsis ACTIN*.
Data are represented as mean ± SEM. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, NS, not significant (Student's t test).
n = biological replicate size, a biological replicate is from a plant sample (a mixture of multiple maize kernels in A–H). n = biological replicate size, a biological replicate contained 36 seeds (I and J) or 24 seeds (K and L). n = biological replicate size, a biological replicate is from a plant sample (*Arabidopsis* ripe siliques in M, maize kernels at the late maturation stage in N and O).

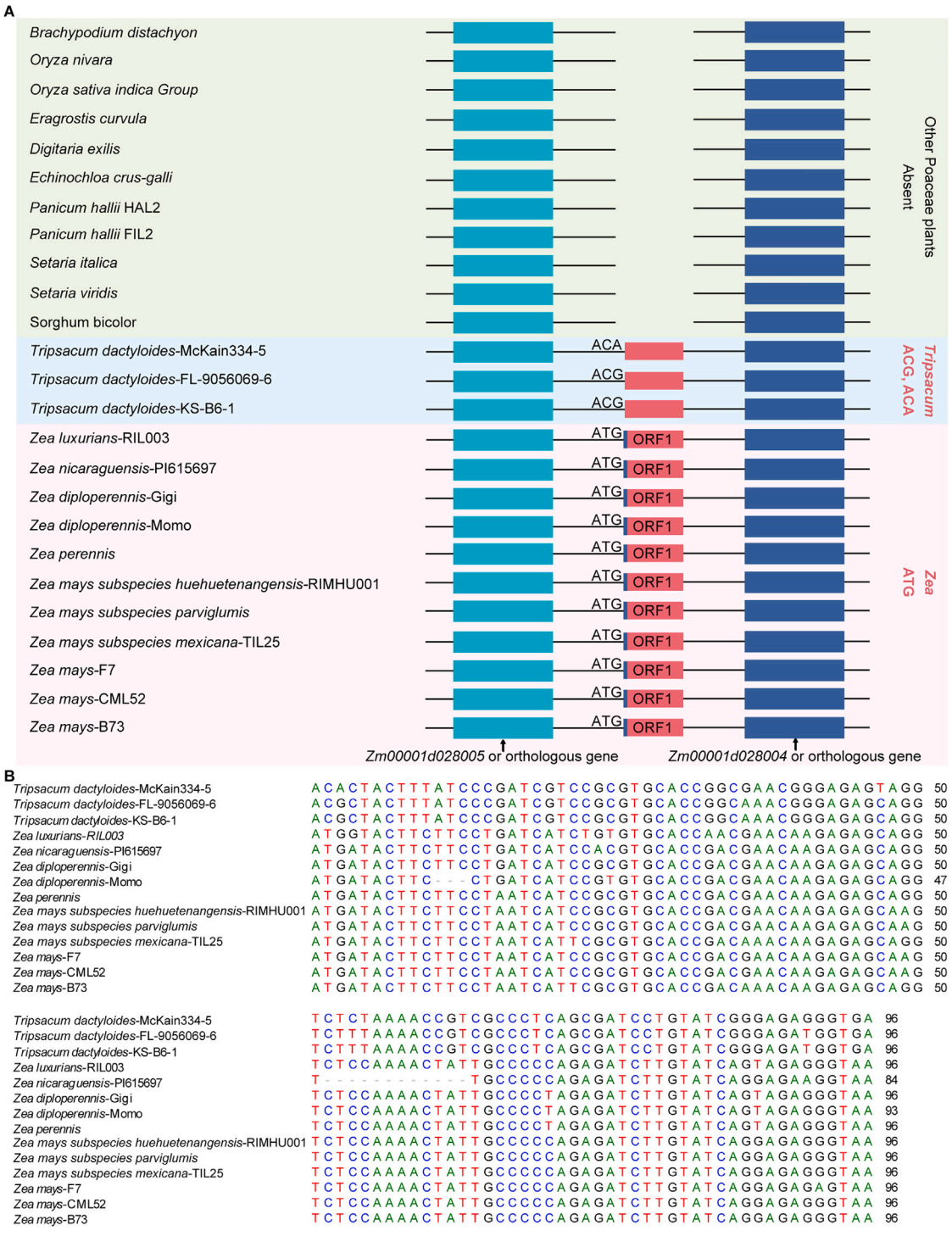
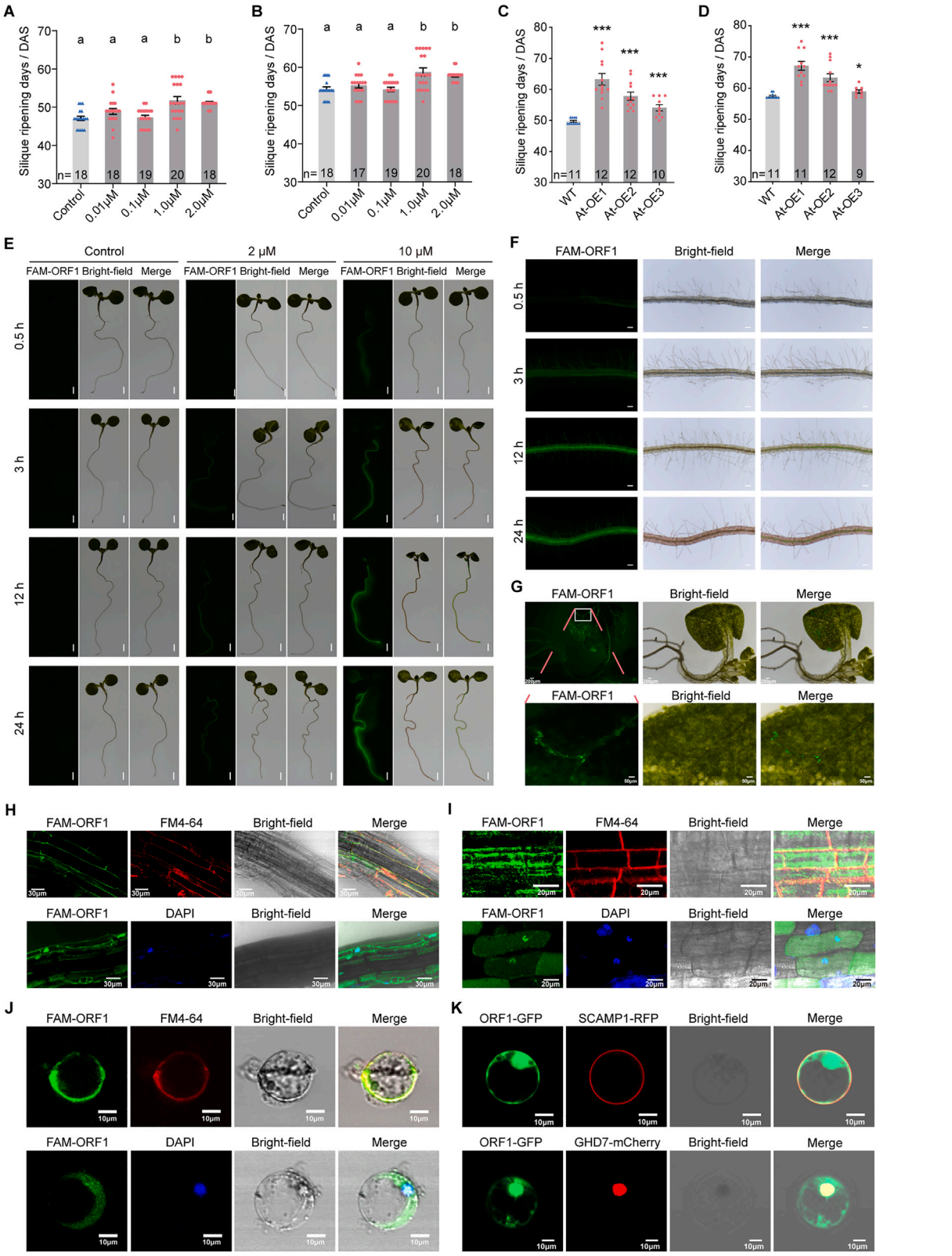


Figure S6. Origin of microRPG1 micropeptide, related to Figure 6

(A) The *de novo* origin of microRPG1. The homologous sequences of microRPG1 were detected in the genus *Zea* and *Tripsacum* but absent in other Poaceae plants. Pink boxes represent the homologous sequences of microRPG1. Blue bars represent the start codon of microRPG1. Cyan and blue boxes represent the 5′ and 3′ flanking genes of microRPG1 or their corresponding orthologous genes in the syntenic regions, respectively. ATG represents the start codon of microRPG1. ACG and ACA represent the non-start codon sequences in the genus *Tripsacum*.

(B) Sequence alignment of microRPG1 and its homologous sequences in the genus *Tripsacum*. The numbers on the right represent the nucleotide positions in the full-length sequences. Several species (*Tripsacum dactyloides*-McKain334-5, *Tripsacum dactyloides*-KS-B6-1, *Zea diploperennis*-Gigi and *Zea perennis*) have more than one syntenic sequence, but they are highly similar and have the same start codon, and the one most similar to microRPG1 is selected for display.



(legend on next page)

Figure S7. microRPG1 delays silique ripening, FAM-ORF1p can be absorbed by root cells and detected in leaf veins, and the subcellular localization patterns of microRPG1, related to Figure 7

(A and B) Applying synthetic ORF1p with different concentrations in *Arabidopsis* (Col-0). Different concentrations of ORF1p (0.01–2 μM) were applied in Col-0. The concentration of 1 and 2 μM of ORF1p significantly delayed silique ripening compared with the control. Silique ripening days were recorded when the siliques began to ripen (A), and when 50% of the siliques were ripe (B). DAS, days after sowing.

(C and D) microRPG1 overexpression delayed silique ripening in *Arabidopsis*. Silique ripening days were recorded when the siliques began to ripen (C), and when 50% of the siliques were ripe (D). DAS, days after sowing.

(E) FAM-ORF1p can be absorbed by seedlings. *Arabidopsis* seedlings were treated with 2 μM or 10 μM of FAM-ORF1p. The figure shows representative image of green fluorescence, bright-field, and merged microscopic images of Col-0 seedlings in the presence of 2 or 10 μM FAM-ORF1p after incubation for 0.5∼24 h. Scale bars, 1 mm.

(F) FAM-ORF1p can be absorbed by root cells of seedlings. *Arabidopsis* seedlings were treated with 10 μM of FAM-ORF1p. The figure shows representative image of green fluorescence, bright-field, and merged microscopic images of Col-0 roots in the presence of 10 μM FAM-ORF1p after incubation for 0.5∼24 h. Scale bars, 100 μm.

(G) FAM-ORF1p can be seen in leaf veins of *Arabidopsis* seedlings. The figure shows representative image of green fluorescence, bright-field, and merged microscopic images of Col-0 seedlings in the presence of 10 μM FAM-ORF1p after incubation for 24 h. Scale bars, 200 (the top row) and 50 μm (the bottom row).

(H) FAM-ORF1 is localized in the nucleus, cytoplasm, and plasma membrane in root cells of *Arabidopsis* seedlings in the presence of 10 μM FAM-ORF1p.

(I) FAM-ORF1 is localized in the nucleus, cytoplasm, and plasma membrane in root cells of maize seedlings in the presence of 10 μM FAM-ORF1p.

(J) FAM-ORF1 is localized in the nucleus, cytoplasm, and plasma membrane in maize protoplast cells in the presence of 10 μM FAM-ORF1p.

(K) ORF1-GFP is localized in the nucleus, cytoplasm, and plasma membrane in maize protoplast cells. SCAMP1-RFP and GHD7-mCherry were used as markers for the plasma membrane and nucleus, respectively.

Scale bars, 30 (H), 20 (I), and 10 μm (J and K).

Data are represented as mean ± SEM. Different letters at top of columns indicate significant differences at $p < 0.05$ (one-way ANOVA followed by Fisher's LSD multiple-comparison test). *$p < 0.05$, ***$p < 0.001$ (Student's t test).

n = sample size, a sample represents the phenotype from a plant.