Title:

An evolutionarily conserved plant RKD factor controls germ cell differentiation

Authors:

Satoshi Koi[1,2], Tetsuya Hisanaga[1], Katsutoshi Sato[1], Masaki Shimamura[3], Katsuyuki T. Yamato[4], Kimitsune Ishizaki[5,6], Takayuki Kohchi[6] and Keiji Nakajima[1,7,*]

Affiliation:

[1] Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan

[2] Faculty of Science, Osaka City University, 2000 Kisaichi, Katano, Osaka 576-0004, Japan

[3] Department of Biological Science, Graduate School of Science, Hiroshima University, 1-3-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan

[4] Faculty of Biology-Oriented Science and Technology, Kinki University, 930 Nishimitani, Kinokawa, Wakayama 649-6493, Japan

[5] Graduate School of Science, Kobe University, 1-1 Rokkodai, Nada-ku, Kobe, Hyogo 657-8501, Japan

[6] Graduate School of Biostudies, Kyoto University, Kitashirakawa-Oiwake, Sakyo-ku, Kyoto 606-8502, Japan

[7] PRESTO, Japan Science and Technology Agency, 4-1-8, Honcho, Kawaguchi, Saitama 332-0012, Japan

Contact information:

[*]Keiji Nakajima: email, k-nakaji@bs.naist.jp: tel, 81-743-72-5560: fax, 81-743-72-5569

Summary

In contrast to animals, where the germ cell lineage is established during embryogenesis, plant germ cells are generated in reproductive organs via reprogramming of somatic cells. The factors that control germ cell differentiation and reprogramming in plants are poorly understood. Members of the RKD subfamily of plant-specific RWP-RK transcription factors have been implicated in egg cell formation in Arabidopsis based on their expression patterns and ability to cause an egg-like transcriptome upon ectopic expression [1]; however, genetic evidence of their involvement is lacking, due to possible genetic redundancy, haploid lethality, and the technical difficulty of analyzing egg cell differentiation in angiosperms. Here, we analyzed the factors that govern germ cell formation in the liverwort *Marchantia polymorpha*. This recently revived model bryophyte has several characteristics that make it ideal for studies of germ cell formation, such as low levels of genetic redundancy, readily accessible germ cells, and the ability to propagate asexually via gemma formation [2, 3]. Our analyses revealed that MpRKD, a single RWP-RK factor closely related to angiosperm RKDs, is preferentially expressed in developing eggs and sperm precursors in *M. polymorpha*. Targeted disruption of Mp*RKD* had no effect on the gross morphology of the vegetative and reproductive organs, but led to striking defects in egg and sperm cell differentiation, demonstrating that Mp*RKD* is an essential regulator of germ cell differentiation. Together

with previous findings [1, 4-6], our results suggest that *RKD* factors are evolutionarily conserved regulators of germ cell differentiation in land plants.

Results

RKD genes are associated with land plant reproduction

We previously reported that RKD4, a member of the RWP-RK family of putative transcription factors, regulates pattern formation in Arabidopsis embryogenesis [6]. To investigate the range of expression of *RKD* genes in plants, we searched for *RKD* homologs in the expressed sequence tag (EST) databases at PlantGDB (http://www.plantgdb.org/), using a sequence spanning the conserved RWP-RK and C-terminal domains of Arabidopsis RKD4 as query. Most ESTs thus identified were found to originate from reproductive cells, such as *Nicotiana tabacum* (tobacco) zygotes, *Glycine max* (soybean) immature seeds, *Triticum aestivum* (wheat) egg cells, *Brachypodium distachyon* flowers, *Malus domestica* (apple) fruit, and *Marchantia polymorpha* (liverwort) sexual organs [1, 7-10], suggesting that RKD factors have evolutionarily conserved roles in land plant reproduction.

We performed phylogenetic analyses to analyze the origin of RKDs and other RWP-RK proteins in the course of land plant evolution. This attempt was also to examine how broadly the *RKD* genes known to be expressed in the reproductive cells occur in the *RKD* phylogeny. To this end, we retrieved entire sets of full-length RWP-RK protein sequences, including those corresponding to the EST clones, from the whole genome sequences of soybean [11] and Arabidopsis [12], as well as from the

draft genome of the liverwort, *M. polymorpha,* which we have sequenced in collaboration with the Joint Genome Institute (JGI). For comparison, we included the entire set of RWP-RK proteins encoded in the genome of another model bryophyte, *Physcomitrella patens* [13], the distantly related algal Minus-Dominance (MID) and NIT2 proteins [14-16], and a lone pair of non-plant RWP-RK proteins from the amoeba *Dictyostelium discoideum* [17] in our analysis. A maximum-likelihood tree constructed from the conserved RWP-RK domain and the carboxy-terminal flanking region indicated that the RKD proteins expressed in the reproductive cells reside in a single phylogenetic clade, and are separated from the MID and NIN-like proteins (NLPs) known to function in sex determination in algae and nodule formation/nitrogen responses in angiosperms, respectively (Figure 1A) [14, 18, 19]. The RKD clade of our tree corresponds to the RKD(A) subfamily previously designated by Chardin et al. [4], and the RKD sequences known to be expressed in reproductive cells were distributed over the entire RKD clade (Figure 1A). Taken together, our in silico analyses suggest that members of the RKD subfamily of plant RWP-RK factors are functionally associated with reproductive development in the land plant lineage, and have a single phylogenetic origin.

The *RKD* gene of liverwort is preferentially expressed in developing egg and sperm precursors

Given that Arabidopsis *RKD4* is necessary for pattern formation in the early stages of the diploid phase

of the lifecycle [5, 6], we were interested in determining why *RKD* homologs exist in bryophytes, as these organisms lack elaborate morphogenesis during their diploid growth phase. We focused our analysis on the liverwort, *M. polymorpha*, because it has several characteristics that facilitate genetic analysis of reproductive development, such as low levels of genetic redundancy, amenability to targeted gene disruption, and asexual reproduction via gemma formation [2, 3]. Of the four RWP-RK proteins encoded in the liverwort genome (Genbank accession numbers, KU987909-KU987912; see table S1 for detail), only one is phylogenetically classified into the RKD clade, and therefore named MpRKD (Figure 1A) [20]. MpRKD shares high levels of sequence similarity with *Arabidopsis* RKDs at the RWP-RK and carboxy-terminal domains, while little homology was found in the amino-terminal region (Figures 1B). The moss *P. patens* also possesses a single RKD member (named PpRKD), and this protein has high levels of sequence similarity with MpRKD along the entire polypeptide (Figure 1B).

M. polymorpha is a dioecious species with multicellular male and female haploid forms that produce morphologically distinct umbrella-like structures named antheridiophores and archegoniophores, respectively (Figures 2A and 2B) [2, 3, 21]. Multiple antheridia form on the upper side of the antheridiophore. A young antheridium is composed of inner spermatogenous cells (SCs) and an outer jacket layer. SCs divide both transversely and vertically, eventually giving rise to numerous sperm mother cells (SMCs). A SMC divides diagonally to form two triangular spermatids, which then

differentiate into flagellated sperm cells (Figure 2A) [21, 22]. By contrast, multiple archegonia are formed below the lobes of each archegoniophore via tightly regulated cell division sequences (Figure 2B). Briefly, a cylindrical structure composed of a single central cell file surrounded by a peripheral cell layer is formed initially. Cells in the peripheral layer eventually form a flask-shaped archegonial vessel, while those in the central cell file divide anticlinally to give rise to a column consisting of an egg, a ventral canal cell (VCC), neck canal cells (NCCs), and cover cells (CVCs) (Figure 2B). During subsequent maturation phases, the egg adopts a morphology that is strikingly distinct from that of the other cells in the column, while the VCC and NCCs degenerate to form a passage for the sperm (Figure 2B, see [3] for a recent review).

RNA in situ hybridization revealed that Mp*RKD* transcript accumulated to high levels in the egg precursor, while negligible signal was detected in the VCC, NCCs, and wall cells (Figure 2C). Expression of Mp*RKD* in the egg cell persists until maturation (Figures 2D). By contrast, no signal was detected after fertilization (Figure 2E). Due to high levels of background signal, we were unable to analyze Mp*RKD* expression in antheridia reliably. Instead, we analyzed Mp*RKD* expression using multiple independent transgenic lines expressing a β-glucuronidase (GUS) reporter under a 6.6-kb Mp*RKD* promoter fragment. Histochemical GUS staining revealed that the Mp*RKD* promoter was active in developing antheridia (Figure 2F), in which GUS staining was confined to the inner SCs, whereas

only faint staining was detected in the outer jacket layer (Figures 2A and 2G). Though weaker than in antheridia, GUS staining was also detected in the meristematic foci of thalli (Figure 2H). Furthermore, we detected Mp*RKD* expression in the edge of gemma cups in lines expressing the red-fluorescent tdTomato reporter [23] driven by the Mp*RKD* promoter (Figure 2I). Based on the strong expression of Mp*RKD* in the egg and SCs, we hypothesize that Mp*RKD* functions in germ cell differentiation. To test this possibility, we performed genetic analyses of Mp*RKD* as described below.

Loss-of-function Mp*rkd* mutants are defective in egg cell formation

To functionally link Mp*RKD* to germ cell formation, we generated loss-of-function Mp*rkd* mutants by homologous recombination (HR) [2]. Our initial attempt to delete a 4.1-kb region encompassing the entire MpRKD-coding region was unsuccessful, likely because this needed to generate a large gap in the genomic DNA. We therefore changed our strategy and deleted an 896-bp region that encompassed the conserved RWP-RK domain (Figure S1A). PCR-based screening of 584 transformants identified three independent Mp*rkd* mutant lines (hereafter referred to as Mp*rkd-1*, Mp*rkd-2*, and Mp*rkd-3*[20]). PCR analysis of the Y and X chromosome-linked markers, *RBM27* and *RHF73* [24], respectively, revealed that the three Mp*rkd* lines were all female (Figure S1B). RT-PCR analysis of thalli indicated that none of the three Mp*rkd* lines expressed Mp*RKD* transcripts covering the conserved RWP-RK domain, while

two of the three lines (Mp*rkd-2* and Mp*rkd-3*) expressed transcripts corresponding to the amino-terminal region alone (Figure S1C). We clonally purified these three mutant lines through two rounds of regeneration from a single gemma before conducting phenotypic analyses [2].

Apart from exhibiting slightly asynchronous thallus growth and lacking outgrowths at the edge of gemma cups (Figures S1D-S1F), the three Mp*rkd* mutants did not exhibit conspicuous morphological defects during vegetative growth. The gemmae developed normally, and hence asexual reproduction proceeded normally in the Mp*rkd* mutants (Figure S1F). After the transition to the reproductive phase, the three Mp*rkd* mutants and wild-type plants all formed archegoniophores, though this process was also asynchronous in Mp*rkd* (Figures S1G-S1J). These observations indicate that MpRKD does not have critical functions in either vegetative growth, the transition to the reproductive phase, or in the formation of reproductive organs *per se*.

In contrast to the weak effect of mutation of Mp*RKD* on the gross morphology of liverworts, egg cell formation was dramatically affected in all Mp*rkd* mutants. In wild-type archegonia, an egg precursor is formed and differentiates into mature eggs with characteristic features, such as large, spherical cell shape, dense cytoplasm, large nuclei and nucleoli, and detachment from the surrounding wall layer (Figures 3A-3D). Although the egg precursor, VCC, and NCCs all formed in Mp*rkd* archegonia as in the wild type (Figures 3E and 3F), the egg precursor further divided in random

orientations (Figure 3G). None of the resulting cells differentiated into mature eggs; instead, they developed large central vacuoles and remained attached to each other and to the surrounding wall cells (Figure 3H). Consistent with the weak but reproducible expression of Mp*RKD* in archegonial cells described below (Figure 2K), the wall cell layer of Mp*rkd* mutants often exhibited a disorganized and double-layered cellular pattern, though the VCCs and NCCs degenerated normally as in the wild type (compare Figures 3C-3D and 3G-3H).

Since no differentiation markers were available for liverwort egg cells, we next performed an ultrastructural comparison of egg cell differentiation in the wild-type and Mp*rkd* plants using transmission electron microscopy (TEM). As previously described, egg cell maturation in wild-type liverwort is accompanied by conspicuous changes in both intra- and intercellular structures [25]. Cell walls between the egg and surrounding cells are typically degraded asymmetrically on the egg side (Figures 3I and 3J). This leads to the formation of a spherical egg protoplast, while the surrounding wall cells retain apparently intact cell walls and remain attached to each other (Figures 3I and 3J). Numerous vacuolar vesicles and highly electron-dense lipid bodies differentiate in wild-type eggs (Figures 3I and 3J) [25]. By contrast, the corresponding cells in the Mp*rkd* archegonia retained largely intact cell walls, and hence only a narrow space was formed between the egg and surrounding wall cells (Figures 3K and 3L). A dense cytoplasm and vesicles were not observed in the mutant, and large central vacuoles

developed in the daughter cells produced by additional divisions of the egg precursor (Figures 3K and 3L). These observations suggest that the egg precursors failed to differentiate into functional egg cells in Mp*rkd* archegonia, and instead divided to form a mass of vacuolated cells similar to other archegonial cells.

Introduction of a transgenic copy of Mp*RKD* rescued the egg defects of Mp*rkd*, indicating that the observed phenotypes were indeed caused by the loss of Mp*RKD* (Figures S2A-S2D). The Mp*rkd* mutant phenotypes were also rescued by expressing MpRKD-Citrine fusion proteins under the Mp*RKD* promoter (Mp*RKD*pro:Mp*RKD*-Citrine). In the rescued plants, strong yellow fluorescence was observed in the nuclei of eggs and their precursor cells (Figures 2K-2M), consistent with the postulated functions of RWP-RK proteins as transcription regulators [17, 19, 26]. Additionally, this experiment revealed that MpRKD proteins were present only at low levels in the nuclei of wall cells, VCCs, and NCCs in immature archegonia (Figures 2K and 2L). Taken together, these results indicate that Mp*RKD* is essential for egg differentiation in liverworts, while it is dispensable for the establishment of basic archegonial architecture.

Mp*RKD* is required for sperm attraction

Since VCCs and NCCs normally degenerate in Mp*rkd* archegonia as in the wild type, resulting in a

passage for sperm entry, we were interested in determining whether Mp*rkd* archegonia retained the ability to attract sperm. To test this, we applied wild-type sperm to wild-type and Mp*rkd* archegonia. After fixation and staining with Hoechst dye, we examined sperm entry by fluorescence microscopy. In the wild type, numerous sperm cells were detected within the archegonial neck region and egg periphery (Figure 3M). By contrast, no sperm was observed in Mp*rkd* mutant archegonia (Figure 3N), and this phenotype was rescued by the Mp*RKD* transgene (Figure S2E). These observations indicate that not only had Mp*rkd* eggs failed to mature, but also that Mp*rkd* archegonia are defective in sperm attraction, though identity and origin of the sperm-attracting signal are to be investigated.

Mp*RKD* is required for sperm formation

The Mp*RKD* promoter showed strong transcriptional activity in SCs, which give rise to sperm cells (Figure 2G), and indeed the Mp*RKD* EST was identified in a male sexual organ cDNA library [9]. We were therefore interested in establishing whether Mp*RKD* is required for sperm formation. Since the three Mp*rkd* lines were all female, we generated male Mp*rkd* mutants by fertilizing the complemented Mp*rkd* archegonia with wild-type sperm, followed by selection of male Mp*rkd* segregants that lacked the complementing transgene. As in Mp*rkd* females, the vegetative growth of Mp*rkd* males was essentially normal, except for the loss of outgrowth at the edge of gemma cups (data not shown). Antheridiophores

developed normally in male Mp*rkd* after transitioning to reproductive growth. However, a comparison of the cellular morphology of Mp*rkd* and wild-type antheridia revealed a striking difference. Whereas all clonal cell packets in wild-type antheridia differentiate into sperm in a highly synchronous manner (Figures 4A and 4B), SC division and sperm differentiation were asynchronous in Mp*rkd* antheridia (Figures 4C and 4D). At maturation, Mp*rkd* antheridia contained vacuolated cells of various sizes, which appeared to result from abortion of SC packets (Figure 4D). Such defects are not completely penetrant, as sperm appeared to mature in some mutant antheridia (Figures 4E and 4F). These antheridial defects were also rescued by the Mp*RKD* transgene, indicating that they too were caused by loss of MpRKD function (Figures S2F and S2G). These observations clearly demonstrate that MpRKD is required for normal differentiation of sperm, while it is dispensable for vegetative growth and the formation of male reproductive organs *per se*.

Discussion

Despite the broad occurrence of *RKD* genes in plants, their loss-of-function phenotypes have been reported only for Arabidopsis *rkd4* mutants, in which pattern formation in early embryos is severely disrupted [5, 6]. Overexpression of *RKD4* caused seedlings to form embryonic callus, and somatic embryos formed on this callus when *RKD4* overexpression was terminated [6]. Similarly, ectopic

expression of *RKD1* and *RKD2* led to overproliferation of root meristematic cells [27], and more importantly, overexpression of *RKD2* conferred an egg-like transcriptome to the seedlings [1]. Consistently, *RKD* transcripts are enriched in the eggs of Arabidopsis and wheat, and reporter analysis showed that Arabidopsis *RKD1* and *RKD2* are specifically expressed in the egg cell [1, 28]. The shared expression patterns of *RKDs* in the eggs of phylogenetically distant angiosperm species suggest that at least some RKD proteins have evolutionarily conserved roles in egg cell differentiation and/or functions. However, experimental evidence for this has been lacking [1]. While several T-DNA insertion alleles are available for Arabidopsis *RKDs*, we and others have so far failed to detect egg defects in *rkd1 rkd2* double mutants, suggesting further functional redundancy among the gene members, incomplete loss of function of the insertion alleles, haploid lethality, or combinations of these phenomena [1, 27]. While Arabidopsis has been used extensively to study plant development, it is not an ideal system to study germ cell differentiation. In Arabidopsis, egg cells differentiate deep inside the ovule, which in turn develops within the carpels enclosed in the flower bud. A detailed characterization of egg defects is therefore difficult to perform in Arabidopsis, even though gene requirement can be inferred from the lack of certain genotypes in the progeny.

In this study, we used *M. polymorpha* to assess the evolutionary origin of RKD functions in land plants. The genome of *M. polymorpha* exhibits strikingly low levels of gene duplication [2], and

this is also the case for RWP-RK proteins; only four RWP-RK genes exist in the *M. polymorpha* genome, one in the RKD clade, one in the NLP clade, and two in the other (Figure 1A). Our present study unambiguously demonstrated that the single liverwort homolog of angiosperm *RKDs* is preferentially expressed in developing egg and sperm precursors, but not in zygotes. This is in contrast to Arabidopsis *RKD4*, which is required exclusively for post-fertilization patterning, even though it is expressed from the egg stage to the globular stage embryo [5]. Therefore, either *RKD4* acquired broader spatiotemporal expression patterns and its protein product gained patterning functions in embryos, or extant liverworts lost the post-fertilization expression and the patterning functions. Considering the lack of elaborate embryo patterning in liverworts, we envisage the former scenario to be true, though characterization of RKDs in intermediary species is needed to confirm this possibility.

M. polymorpha is among the few plant species in which targeted gene disruption is practical, though genes essential for its haploid-dominant lifecycle cannot be knocked out. We successfully isolated loss-of-function Mp*rkd* mutants, and phenotypic characterization revealed that these plants had striking defects in egg and sperm formation. By contrast, only minute defects were found during other developmental stages. The lack of egg maturation in Mp*rkd* archegonia is exactly as predicted for the loss of egg-expressed *RKDs* in Arabidopsis, and suggests that plant RKD factors retain evolutionarily conserved functions in egg differentiation. To corroborate this possibility, we performed mutual

complementation experiments, in which Mp*RKD* was expressed under the *RKD4* promoter in the Arabidopsis *rkd4* mutant, and either *RKD2* or *RKD4* was expressed under the Mp*RKD* promoter in the liverwort Mp*rkd* mutant. However, the mutant defects were not rescued in any cases. These findings suggest that RKD proteins in extant land plants have been already diversified and can no longer control downstream regulatory networks in phylogenetically distant plant species, even if their common ancestors likely functioned in germ cell differentiation.

In conclusion, our present study provides solid evidence that a single *RKD* homolog plays essential roles in germ cell differentiation in liverworts. Similar results are reported by the accompanying paper in which Mp*RKD* expression was repressed by artificial miRNAs [29]. The broad occurrence of *RKD* genes in land plants, together with the findings that at least some *RKD* genes are expressed in egg cells and activate egg-specific genes suggest that RKD proteins have an evolutionarily conserved role in germ cell differentiation [1, 5]. Additionally, this study demonstrates the power of *M. polymorpha* as a genetic tool to investigate germ cell differentiation in plants [2, 30]. As RWP-RK proteins act as transcriptional regulators [17, 26], it is tempting to isolate the target genes of MpRKD in liverwort and to perform functional analyses of these genes by producing their knock-out mutants.

Acknowledgements

We thank Takashi Araki (Kyoto University) for valuable comments on sperm formation in liverworts, John Bowman (Monash University) and staff at the DOE-Joint Genome Institute (JGI) for sequencing the *M. polymorpha* genome, and Kaori Kitayama, Masako Kanda Emi Murata, Kayo Hashimoto (NAIST), Sakiko Ishida (Kyoto University), and Hajime Akashi (Hiroshima University) for technical assistance. This work was supported by a JST-PRESTO grant to K.N., JSPS Grant-in-Aid for Scientific Research (B) (243700229) to K.N., MEXT Grant-in-Aid for Scientific Research on Priority Areas (25113007) to K.N., and a grant from the Kurata Memorial Hitachi Science and Technology Foundation to K.N.

Author Contributions

S.K., T.H., K.S., M.S., and K.N. conducted the experiments; S.K., K.T.Y., K.I., T.K., and K.N. designed the experiments; and S.K., T.H., M.S., and K.N. wrote the paper.

References

1. Koszegi, D., Johnston, A.J., Rutten, T., Czihal, A., Altschmied, L., Kumlehn, J., Wust, S.E., Kirioukhova, O., Gheyselinck, J., Grossniklaus, U., et al. (2011). Members of the RKD transcription factor family induce an egg cell-like gene expression program. Plant J *67*, 280-291.
2. Ishizaki, K., Nishihama, R., Yamato, K.T., and Kohchi, T. (2016). Molecular Genetic Tools and Techniques for Marchantia polymorpha Research. Plant Cell Physiol *57*, 262-270.

3. Shimamura, M. (2016). Marchantia polymorpha; taxonomy, phylogeny, and morphology of a model system. Plant Cell Physiol *57*, 230-256.
4. Chardin, C., Girin, T., Roudier, F., Meyer, C., and Krapp, A. (2014). The plant RWP-RK transcription factors: key regulators of nitrogen responses and of gametophyte development. J Exp Bot *65*, 5577-5587.
5. Jeong, S., Palmer, T.M., and Lukowitz, W. (2011). The RWP-RK factor GROUNDED promotes embryonic polarity by facilitating YODA MAP kinase signaling. Curr Biol *21*, 1268-1276.
6. Waki, T., Hiki, T., Watanabe, R., Hashimoto, T., and Nakajima, K. (2011). The Arabidopsis RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. Curr Biol *21*, 1277-1281.
7. Gasic, K., Gonzalez, D.O., Thimmapuram, J., Liu, L., Malnoy, M., Gong, G., Han, Y., Vodkin, L.O., Aldwinckle, H.S., Carroll, N.J., et al. (2009). Comparative analysis and functional annotation of a large expressed sequence tag collection of apple. Plant Genome *2*, 23-38.
8. Sha, A.H., Li, C., Yan, X.H., Shan, Z.H., Zhou, X.A., Jiang, M.L., Mao, H., Chen, B., Wan, X., and Wei, W.H. (2012). Large-scale sequencing of normalized full-length cDNA library of soybean seed at different developmental stages and analysis of the gene expression profiles based on ESTs. Mol Biol Rep *39*, 2867-2874.
9. Yamato, K.T., Ishizaki, K., Fujisawa, M., Okada, S., Nakayama, S., Fujishita, M., Bando, H., Yodoya, K., Hayashi, K., Bando, T., et al. (2007). Gene organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system. Proc Natl Acad Sci U S A *104*, 6472-6477.
10. Zhao, J., Xin, H., Qu, L., Ning, J., Peng, X., Yan, T., Ma, L., Li, S., and Sun, M.X. (2011). Dynamic changes of transcript profiles after fertilization are associated with de novo transcription and maternal elimination in tobacco zygote, and mark the onset of the maternal-to-zygotic transition. Plant J *65*, 131-145.
11. Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature *463*, 178-183.
12. Arabidopsis-Genome-Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature *408*, 796-815.
13. Rensing, S.A., Lang, D., Zimmer, A.D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P.F., Lindquist, E.A., Kamisugi, Y., et al. (2008). The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science *319*, 64-69.

14. Camargo, A., Llamas, A., Schnell, R.A., Higuera, J.J., Gonzalez-Ballester, D., Lefebvre, P.A., Fernandez, E., and Galvan, A. (2007). Nitrate signaling by the regulatory gene NIT2 in Chlamydomonas. Plant Cell *19*, 3491-3503.
15. Ferris, P.J., and Goodenough, U.W. (1997). Mating type in Chlamydomonas is specified by mid, the minus-dominance gene. Genetics *146*, 859-869.
16. Geng, S., De Hoff, P., and Umen, J.G. (2014). Evolution of sexes from an ancestral mating-type specification pathway. PLoS Biol *12*, e1001904.
17. Schauser, L., Wieloch, W., and Stougaard, J. (2005). Evolution of NIN-like proteins in Arabidopsis, rice, and Lotus japonicus. J Mol Evol *60*, 229-237.
18. Castaings, L., Camargo, A., Pocholle, D., Gaudon, V., Texier, Y., Boutet-Mercey, S., Taconnat, L., Renou, J.P., Daniel-Vedele, F., Fernandez, E., et al. (2009). The nodule inception-like protein 7 modulates nitrate sensing and metabolism in Arabidopsis. Plant J *57*, 426-435.
19. Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. Nature *402*, 191-195.
20. Bowman, J.L., Araki, T., Arteaga-Vazquez, M.A., Berger, F., Dolan, L., Haseloff, J., Ishizaki, K., Kyozuka, J., Lin, S.S., Nagasaki, H., et al. (2016). The naming of names: guidelines for gene nomenclature in Marchantia. Plant Cell Physiol *57*, 257-261.
21. Durand, E.J. (1908). The development of the sexual organs and sprorogonium of Marchantia polymorpha. The Bulletin of the Torrey Botanical Club *35*, 321-335.
22. Higo, A., Niwa, M., Yamato, K.T., Yamada, L., Sawada, H., Sakamoto, T., Kurata, T., Shirakawa, M., Endo, M., Shigenobu, S., et al. (2016). Transcriptional Framework of Male Gametogenesis in the Liverwort Marchantia polymorpha L. Plant Cell Physiol *57*, 325-338.
23. Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol *22*, 1567-1572.
24. Fujisawa, M., Hayashi, K., Nishio, T., Bando, T., Okada, S., Yamato, K.T., Fukuzawa, H., and Ohyama, K. (2001). Isolation of X and Y chromosome-specific DNA markers from a liverwort, Marchantia polymorpha, by representational difference analysis. Genetics *159*, 981-985.
25. Zinsmeister, D.D., and Carothers, Z.B. (1972). The Fine Structure of Oogenesis in Marchantia polymorpha. American Journal of Botany *61*, 499-512.
26. Soyano, T., Shimoda, Y., and Hayashi, M. (2015). NODULE INCEPTION antagonistically regulates gene expression with nitrate in Lotus japonicus. Plant Cell Physiol *56*, 368-376.
27. Waki, T., Miyashima, S., Nakanishi, M., Ikeda, Y., Hashimoto, T., and Nakajima, K. (2013). A

GAL4-based targeted activation tagging system in Arabidopsis thaliana. Plant J *73*, 357-367.

28. Wuest, S.E., Vijverberg, K., Schmidt, A., Weiss, M., Gheyselinck, J., Lohr, M., Wellmer, F., Rahnenfuhrer, J., von Mering, C., and Grossniklaus, U. (2010). Arabidopsis female gametophyte gene expression map reveals similarities between plant and animal gametes. Curr Biol *20*, 506-512.
29. Rövekamp, M., Bowman, J.L., and Grossniklaus, U. (2016). Marchantia MpRKD regulates the gametophyte-sporophyte transition by keeping egg cells quiescent in the absence of fertilization. Curr Biol *in press*.
30. Bowman, J.L. (2016). A Brief History of Marchantia from Greece to Genomics. Plant Cell Physiol *57*, 210-229.

Figure legends

Figure 1. Phylogenetic analysis of RWP-RK proteins and placement of RKD proteins expressed in reproductive cells.

(A) A maximum-likelihood tree of RWP-RK proteins constructed using the conserved RWP-RK and carboxyl-terminal domains of representative plant species. See Table S1 for sequences and accession numbers. Numbers at the braches indicate bootstrap values calculated from 1000 replicates. Only values higher than 50 are shown. Scale bar indicates evolutionary distance expressed as the ratio of amino acid substitutions. Species are indicated by colored circles and abbreviated as follows: At, *Arabidopsis thaliana*; Bd, *Brachypodium distachyon*; Cr, *Chlamydomonas reinhardtii*; Ddi, *Dictyostelium*

discoideum; Gm, *Glycine max*; Md, *Malus domestica*; Mp, *Marchantia polymorpha;* Nt, *Nicotiana tabacum*; Pp, *Physcomitrella patens*; Ta, *Triticum aestivum*; Vc, *Volvox carteri*. Arrows indicate RKD members with known expression in reproductive cells. The RKD clade is shaded in blue, whereas the NLP clade is shaded in yellow.

(B) Comparison of domain organization between RKD proteins from *M. polymorpha*, *P. patens*, and *A. thaliana*. Conserved regions are colored as indicated.

Figure 2. Expression patterns of Mp*RKD* in liverworts.

(A, B) Diagrams illustrating sperm formation in the antheridium (A) and egg formation in the archegonium (B) of *M. polymorpha*. Nuclei are not drawn for cells in the antheridium, except for in the final step of sperm formation.

(C-E) RNA in situ hybridization demonstrating that Mp*RKD* transcripts accumulate to high levels in the developing egg (C, D) and are absent in the zygote (E).

(F-H) Histochemical GUS staining demonstrating Mp*RKD* transcription in the antheridium (F, G) and meristematic foci of the thallus (H). Arrows indicate some antheridia (F) and a meristematic focus (H) expressing the GUS reporter.

(I) Red fluorescence from the tdTomato reporter driven by the Mp*RKD* promoter, indicating Mp*RKD*

transcription at the edge of a gemma cup (arrowhead) and a meristematic focus (arrow).

(J-M) Expression patterns of MpRKD-Citrine fusion proteins under the Mp*RKD* promoter (Mp*RKD*pro:Mp*RKD*-Citrine) in Mp*rkd-1* mutants. Note the strong nuclear fluorescence in the egg and its precursor cells (arrows in K-M), and additional weak expression in the surrounding wall cells, VCC, and NCCs (arrowheads in K and L). Cell shapes are delineated by dashed lines in (L) and (M). Bars, C-E, J-M, 10 µm; F, I, 1 mm; G, 50 µm; H, 1 cm.

Figure 3. MpRKD functions are required for egg differentiation in liverwort.

(A-H) Archegonium development of the wild type (A-D) and Mp*rkd*-1 mutants (E-H).

(I-L) Ultrastructural comparison of egg maturation in the wild type (I, J) and Mp*rkd-2* (K, L).

(J) and (L) Magnified views of the boxed regions in (I) and (K), respectively. Arrows indicate the presence of cell walls in the wall cells of wild-type and Mp*rkd-2* archegonia. The arrowhead in (J) indicates the absence of visible cell walls in the wild-type egg cell. Note the presence of cell walls in the Mp*rkd* mutant egg (yellow arrow in L).

(M, N) Sperm attraction assay. Sperm from wild-type antheridia were applied to wild-type (M) and Mp*rkd-2* (N) archegonia, and their entry into the archegonia was visualized by Hoechst staining. Arrows indicate sperm inside the wild-type archegonium (M).

Bars, A-I, K, M, N, 10 µm; J, L, 2 µm. Cell type labelling is shown in Figure 2B. See also Figure S2.

Figure 4. MpRKD functions are required for sperm formation.

(A-D) Sperm formation in wild-type (A, B) and Mp*rkd*-3 (C-F) antheridia. (B), (D), and (F) are higher magnifications of the boxed regions in (A), (C), and (E), respectively. Note the asynchronous differentiation of spermatogenous cells (SCs) in Mp*rkd*-3 mutants (C). In contrast to wild-type antheridia, which were filled with small cells that synchronously differentiated into sperm (B), Mp*rkd* antheridia contained large vacuolated cells (asterisks in D) and packets of small vacuolated cells (delineated by a dashed line in D). Normal sperm appeared to form in some Mp*rkd* antheridia (E, F).

Bars, A, C, E, 50 µm; B, D, F, 10 µm. See also Figure S2.