

Gamete expression of TALE class HD genes activates the diploid sporophyte program in *Marchantia polymorpha*

Tom Dierschke^{1,2}, Eduardo Flores-Sandoval¹, Madlen I Rast-Somssich¹, Felix Althoff², Sabine Zachgo², John L Bowman^{1*}

¹School of Biological Sciences, Monash University, Melbourne, Australia; ²Botany Department, University of Osnabrück, Osnabrück, Germany

Abstract Eukaryotic life cycles alternate between haploid and diploid phases and in phylogenetically diverse unicellular eukaryotes, expression of paralogous homeodomain genes in gametes primes the haploid-to-diploid transition. In the unicellular chlorophyte alga *Chlamydomonas*, KNOX and BELL TALE-homeodomain genes mediate this transition. We demonstrate that in the liverwort *Marchantia polymorpha*, paternal (sperm) expression of three of five phylogenetically diverse BELL genes, MpBELL234, and maternal (egg) expression of both MpKNOX1 and MpBELL34 mediate the haploid-to-diploid transition. Loss-of-function alleles of MpKNOX1 result in zygotic arrest, whereas a loss of either maternal or paternal MpBELL234 results in variable zygotic and early embryonic arrest. Expression of MpKNOX1 and MpBELL34 during diploid sporophyte development is consistent with a later role for these genes in patterning the sporophyte. These results indicate that the ancestral mechanism to activate diploid gene expression was retained in early diverging land plants and subsequently co-opted during evolution of the diploid sporophyte body.

*For correspondence:

John.Bowman@monash.edu

Competing interest: The authors declare that no competing interests exist.

Funding: See page 20

Received: 20 March 2020

Preprinted: 07 April 2020

Accepted: 15 September 2021

Published: 17 September 2021

Reviewing Editor: Sheila McCormick, University of California, Berkeley, United States

© Copyright Dierschke et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

Life cycles of eukaryotes alternate between haploid and diploid phases, initiated by meiosis and gamete fusion, respectively. Expression of paralogous homeodomain (HD) genes in the two gametes and the subsequent heterodimerization of the respective proteins in the zygote direct the haploid-to-diploid transition in gene expression in phylogenetically diverse eukaryotes, including the ascomycete fungus *Saccharomyces cerevisiae* (Goutte and Johnson, 1988; Herskowitz, 1989), the basidiomycete fungi *Coprinopsis cinerea* and *Ustilago maydis* (Gillissen et al., 1992; Hull et al., 2005; Kues et al., 1992; Spit et al., 1998; Urban, 1996), the Amoebozoa *Dictyostelium discoideum* (Hedgethorne et al., 2017), the brown alga *Ectocarpus* (Arun et al., 2019), the red alga *Pyropia yezoensis* (Mikami et al., 2019), and the unicellular chlorophyte alga *Chlamydomonas reinhardtii* (Ferris and Goodenough, 1987; Lee et al., 2008; Nishimura et al., 2012; Zhao et al., 2001). This broad phylogenetic distribution suggests this was an ancestral function of HD genes (reviewed by Bowman et al., 2016b). In Viridiplantae, the paralogs are two subclasses, KNOX and BELL, of TALE class HD genes. In *Chlamydomonas*, the minus (-) gamete expresses a KNOX protein (GSM1) and the plus (+) gamete expresses a BELL protein (GSP1), and upon gamete fusion, the two proteins heterodimerize and translocate to the nucleus, activating zygotic gene expression (Lee et al., 2008). GSM1 and GSP1 are necessary for diploid gene expression and, when ectopically expressed together in vegetative haploid cells, are sufficient to induce the diploid genetic program (Ferris and Goodenough, 1987; Lee et al., 2008; Nishimura et al., 2012; Zhao et al., 2001). Biologically, the expression of a unique paralog in each type of gamete, coupled with the requirement for heterodimerization for functionality, is a

mechanism to preload gametes such that immediately following gamete fusion/fertilization a distinct diploid genetic program is initiated.

Heterodimerization of TALE-HD paralogs is mediated by subclass-specific protein domains N-terminal of the homeodomain conserved in phylogenetically diverse eukaryotes (Bellaoui et al., 2001; Bürglin, 1997). For example, Viridiplantae KNOX proteins and metazoan MEIS proteins share a homologous heterodimerization domain, thus defining a TALE-HD subclass, named MEINOX, present in the ancestral eukaryote (Bürglin, 1997). In contrast, heterodimerization domains of MEINOX partners are not as well conserved (Joo et al., 2018). Some MEINOX partners in animals have a characteristic PBC-C heterodimerization domain (Bürglin, 1997), and PBC-related domains are found in some algal BELL-related TALE-HD proteins, suggesting its presence in the ancestral eukaryote (Joo et al., 2018). In contrast, previously characterized land plant BELL proteins possess conserved SKY and BELL domains, collectively termed POX (pre-homeobox), that mediate heterodimerization with KNOX partners (Bellaoui et al., 2001; Hackbusch et al., 2005; Smith et al., 2002). POX domains are either highly divergent from, or unrelated to, PBC domains. Finally, in other algal BELL-related TALE-HD proteins, neither heterodimerization domain is evident (Joo et al., 2018). Regardless, when tested for the dimerization potential, Archaeplastida BELL-related TALE-HD proteins interact only with MEINOX TALE-HD partners and not with other BELL-related TALE-HD proteins (Joo et al., 2018).

Land plants are characterized by an alternation of generations whereby complex multicellular bodies develop in both haploid (gametophyte) and diploid (sporophyte) phases of the life cycle (Hofmeister, 1862), and elements of the KNOX/BELL system have been implicated in regulating the land plant haploid-to-diploid transition. In the moss *Physcomitrium patens*, one subclass of KNOX genes, KNOX1, is required for proper proliferation and differentiation in the diploid body (Sakakibara et al., 2008; Singer and Ashton, 2007), while another subclass, KNOX2, acts to suppress the haploid genetic program during diploid development (Sakakibara et al., 2013). In addition to being expressed during sporophyte development, both classes of KNOX genes are expressed in the egg cell during gamete formation (Sakakibara et al., 2013; Sakakibara et al., 2008). The expression of KNOX genes in egg cells suggests that this gamete may correspond to the (-) gamete in *Chlamydomonas*, which also expresses a KNOX protein, implying that the male gamete in land plants might express BELL proteins, as does the (+) gamete in *Chlamydomonas*. Ectopic expression of a *Physcomitrium* BELL paralog, PpBELL1, has been noted to induce the diploid genetic program in specific cell types of the gametophyte, but PpBELL1 expression was detected in the egg cell (Horst et al., 2016), rather than the male gamete, as might have been expected if it is functionally similar to *Chlamydomonas* BELL. However, expression of PpBELL1 has also been reported to be induced in the sperm by activation of glutamate channels, with PpBELL1 subsequently functionally active in the sporophyte (Ortiz-Ramírez et al., 2017). These observations prompted our investigation of the homologous genetic programs in the liverwort *Marchantia polymorpha*.

Results

***M. polymorpha* possesses phylogenetically diverse TALE homeodomain proteins**

The *M. polymorpha* genome encodes nine TALE-HD-related family members: four KNOX genes and five BELL genes (Bowman et al., 2017; Figure 1), all of which are expressed in either the haploid sexual organs or the diploid sporophyte or both, with minimal or no expression detected by reverse transcription-polymerase chain reaction (RT-PCR) in the haploid vegetative thallus (Figure 1—figure supplement 1). Of the four KNOX genes, three are of KNOX1 subclass, but with only one encoding a HD, and one of KNOX2 subclass (Figure 1—figure supplement 1). The two subclasses arose via gene duplication in an ancestral charophycean alga (Sakakibara, 2016; Joo et al., 2018; Frangedakis et al., 2017). Expression of MpKNOX1 (Mp5g01600), which encodes a HD, is predominantly detected in archegoniophores and young sporophytes. In contrast, the two other KNOX1 genes (MpKNOX1A, Mp4g12450; MpKNOX1B, Mp2g11140) lacking a HD are predominantly expressed in antheridiophores. Phylogenetic analysis indicates that these HD-less paralogs are liverwort specific and are likely evolved in the ancestral liverwort (Figure 1—figure supplement 1). In contrast, expression of MpKNOX2 (Mp7g05320) is not detected in unfertilized reproductive organs and is expressed primarily during sporophyte development.

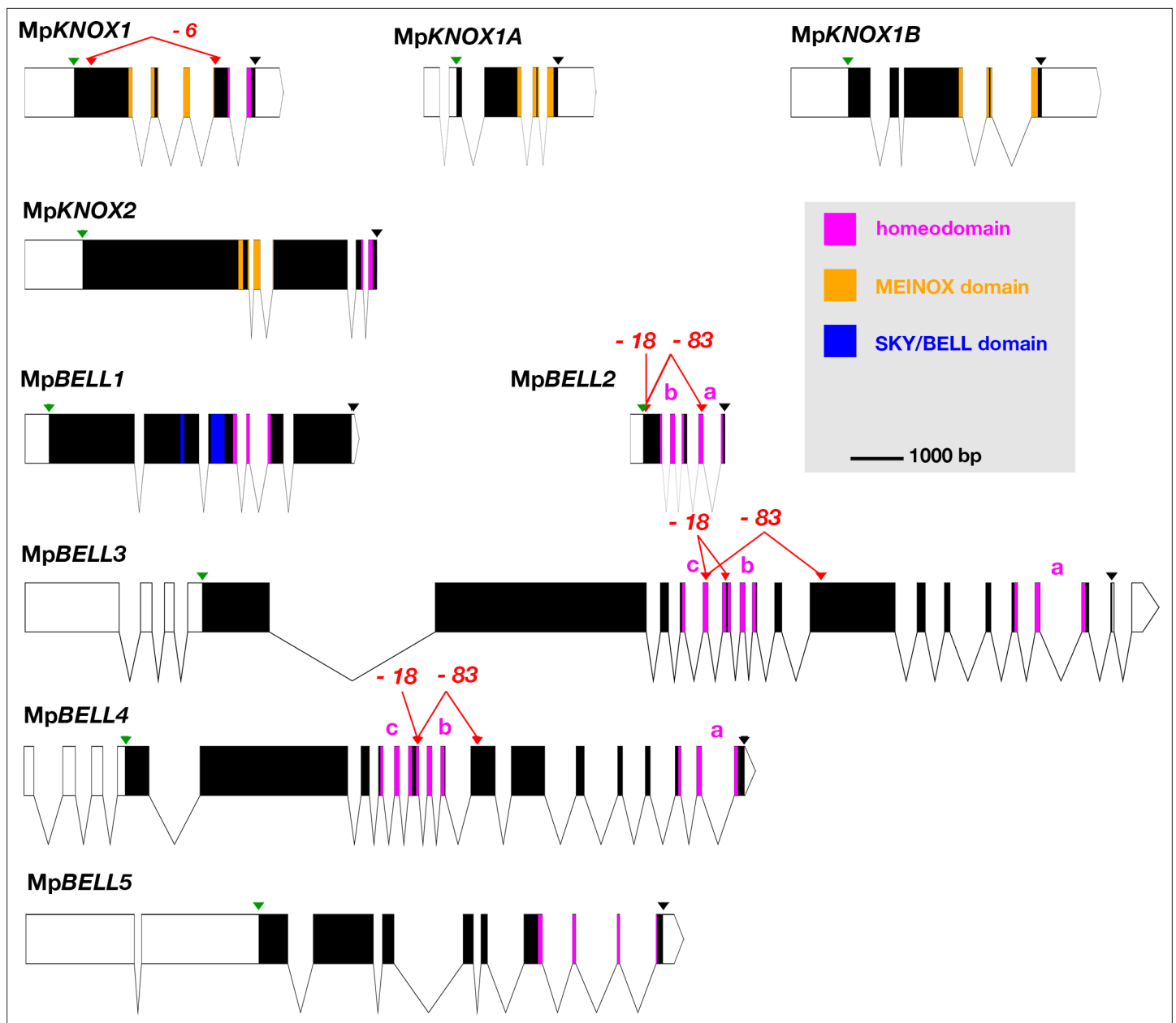


Figure 1. Schematic representations of the nine TALE-HD loci of *Marchantia polymorpha*. White, 5' and 3' untranslated regions (UTRs); black, coding exons; thin lines, introns; green triangles, start codon; black triangles, stop codon; red triangles, guide RNA-targeted positions. The mutant alleles generated via CRISPR-Cas9 are described in more detail in table 1; molecular lesions of some alleles can be found in **Figure 1—source data 1**. All protein annotation models are based on the *Marchantia* genome assembly of v5.1 except for MpKNOX2. The MpKNOX2 model is based on sequences derived from reverse transcription-polymerase chain reaction (RT-PCR). In genes with multiple homeodomains, they are denoted a, b, and c. Gene models were assembled using wormweb (<http://wormweb.org/exonintron>).

The online version of this article includes the following figure supplement(s) for figure 1:

Source data 1. Molecular lesions of alleles listed in **Figure 1** and Table 1.

Figure supplement 1. Expression of *Marchantia polymorpha* TALE-HD genes.

Figure supplement 2. Unrooted Bayesian phylogram of Viridiplantae KNOX-TALE class homeodomain genes.

The *M. polymorpha* BELL genes reside in three phylogenetically distinct clades, each with its own characteristic HD (**Figure 2; Figure 2—figure supplement 1**). Like some algal BELL-related proteins (**Joo et al., 2018**), the presence of a PBC domain is not easily discernible in any of the predicted *M. polymorpha* BELL protein sequences. Only one *M. polymorpha* BELL gene (MpBELL1, Mp8g18310)

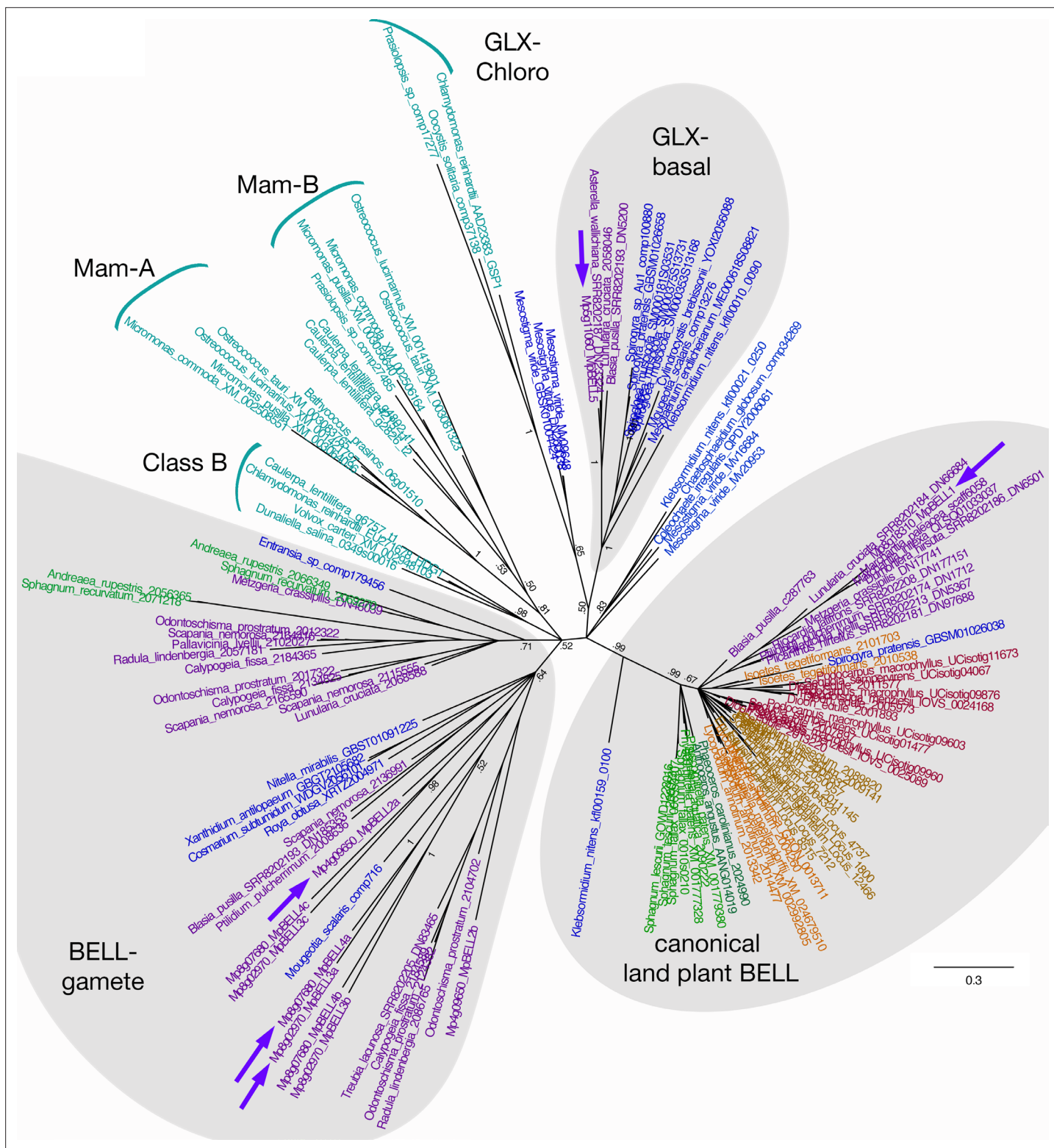


Figure 2. Unrooted Bayesian phylogram of Viridiplantae BELL-TALE class homeodomain genes. Tree was constructed using a nucleotide alignment of the homeodomain. The five *Marchantia polymorpha* BELL genes are highlighted (purple arrows). A clade of genes representing previously identified ‘canonical’ land plant BELL genes is recovered with high support. This clade includes MpBELL1 and charophycean algal sequences dating as far back as prior to the divergence of *Klebsormidium*. These canonical land plant BELL genes also share significant sequence similarity outside the homeodomain, including the SKY and BELL domains. The remaining BELL genes, including all chlorophyte sequences, sequences from charophycean algae, and the other four *M. polymorpha* sequences, reside in either other well-supported clades or in polytomies. Several subclades (labeled) previously identified

Figure 2 continued on next page

Figure 2 continued

are evident (Joo et al., 2018). MpBELL2/3/4 resides in a polytomy composed of liverwort and charophyte sequences, with other liverwort, moss, and charophyte sequences residing in a second clade; these are labeled 'BELL-gamete'. The multiple homeodomains of MpBELL2, MpBELL3 and MpBELL4 are designated a, b, and c (see Figure 1). MpBELL5 resides, with other liverwort and charophyte sequences, in the GLX-basal clade. Chlorophyte, light blue; charophyte, dark blue; liverwort, purple; moss, green; hornwort, dark green; lycophyte, brown; fern, orange; seed plant, red. Numbers at branches indicate posterior probability values >50%; branches explicitly shown have probability values >50%, whereas polytomies represent nodes with probability values <50%; values within subclades are omitted for clarity.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Comparison of Viridiplantae TALE homeodomains.

is phylogenetically related to previously described, 'canonical', land plant BELL genes. MpBELL1 harbors a conserved canonical land plant BELL homeodomain sequence and a discernible, albeit divergent, BELL domain characteristic of other land plant BELL genes. Similar to MpKNOX2, MpBELL1 expression is not detected in reproductive organs prior to fertilization and is expressed predominantly during sporophyte development (Figure 1—figure supplement 1). The other four *M. polymorpha* BELL genes fall into two distinct clades and are more closely related to algal TALE-HD genes than to the canonical land plant BELL genes (Figure 2). Three of the algal-like BELL genes, MpBELL2 (Mp4g09650), MpBELL3 (Mp8g02970), and MpBELL4 (Mp8g07680), each encode two or three homeodomains (Figure 1). The carboxyl HD is most conserved with those upstream becoming progressively less conserved, consistent with their origins being via successive intragenic duplications. Notably, MpBELL3 and MpBELL4 both encode large proteins, while MpBELL2 encodes a much smaller protein consisting of only two HDs. These three genes, named here BELL-gamete, arose via gene duplications within liverworts, and related genes are found throughout liverworts and in some moss lineages that are sister groups to the vast majority of moss diversity (Figure 2). The fourth algal-like sequence, MpBELL5 (Mp5g11060), is phylogenetically distinct from the other four *M. polymorpha* BELL genes, and resides in the previously defined 'GLX-basal' clade (Joo et al., 2018). Orthologs of MpBELL5 exist in other liverworts (Figure 2). The most parsimonious interpretation is that a diversity of BELL paralogs arose in a charophyte algal ancestor, with the ancestral land plant likely possessing three paralogs representing the canonical land plant, GLX-basal and BELL-gamete lineages (Figure 2). This diversity has persisted in *M. polymorpha* and other liverworts and, to a lesser extent, in mosses.

Maternal MpKNOX1 is required for post-zygotic embryo development

During gametophyte generation, MpKNOX1 is expressed specifically in the egg cell but is not detected at the stage prior, when the venter canal cell is present (Figure 3A–B). During differentiation of the egg cell, MpKNOX1 appears to be expressed in a pulse early, with the signal diminishing in older archegonia (Figure 3B). Following fertilization, MpKNOX1 is expressed throughout the developing sporophyte up until the time when future sporogenous cells become distinct (Figure 3C–E), with expression thereafter becoming undetectable during meiotic stages (Figure 3F). A female harboring a loss-of-function allele consisting of a 2.3 kb deletion spanning the MpKNOX1 locus (Figure 1; Table 1) was crossed with wild-type males. In contrast to control crosses using wild-type males and females where sporophyte production was observed (Figure 3G–I), mature sporophytes never developed on gametangiophores from >100 crosses between wild-type males and Mpknx1-6^{9e} females (Figure 3J–L). However, on closer inspection, on rare occasions (3/129 fertilization events) did we encounter arrested multicellular embryos within senescing archegoniophores (Figure 3—figure supplement 1). A closer examination of female Mpknx1-6^{9e} mutants revealed a developmental arrest of the zygote following fertilization, with the zygote failing to undergo cytokinesis (Figure 3K). However, two gametophytic tissues, the calyptra and the pseudoperianth, whose growth is dependent upon fertilization (Hofmeister, 1862), commence development normally. In Mpknx1-6^{9e} mutants, as in wild type, the calyptra undergoes periclinal divisions indicating successful fertilization and implying an intergenerational (sporophyte to gametophyte) signal inducing its development (Figure 3H–I and K–L). Likewise, the pseudoperianth, a ring of tissue surrounding the archegonium that only develops post-fertilization, initially develops normally in Mpknx1-6^{9e} mutants (Figure 3H–I and K–L). Subsequently, both the calyptra and pseudoperianth are developmentally arrested between 1 and 2 weeks post fertilization (wpf) in Mpknx1-6^{9e} mutants. In addition, Mpknx1-6^{9e} mutants exhibit a senescent archegoniophore phenotype. Unfertilized wild-type archegoniophores remain green for multiple

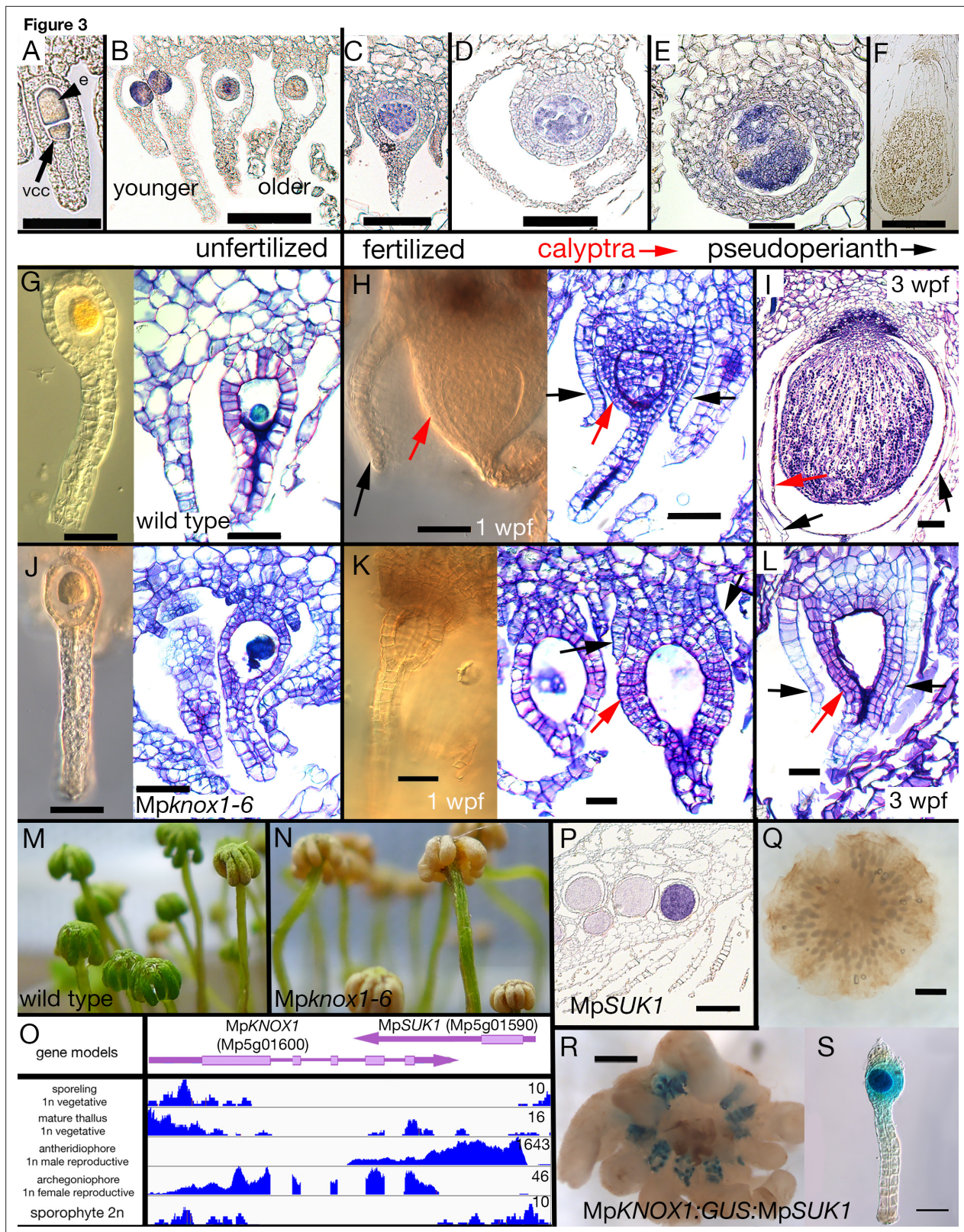


Figure 3. Maternal *MpKNOX1* is required for post-zygotic embryo development. (A–F) *MpKNOX1* expression pattern. *MpKNOX1* transcripts could be detected via *in situ* hybridization in the egg (e) cells of unfertilized archegonia (B), but not in the venter canal cell (A; vcc). After fertilization, expression is detected in young developing sporophytes (approximately 1 week post fertilization (wpf); C, D), with expression continuing until foot and seta begin to differentiate (E). *MpKNOX1* is not expressed in older sporophytes where sporogenous tissue has differentiated (F). (G–L) Comparison of wild-type

Figure 3 continued on next page

Figure 3 continued

and *Mpknox1* development; unfertilized (**G, J**), 1 wpf (**H, K**), and 3 wpf (**I, L**). *Mpknox1-6^{9e}* egg cells appear wild-type-like (unfertilized; **G, J**). In contrast to wild-type (**H, I**), the embryo of *Mpknox1-6^{9e}* mutants does not develop post fertilization (**K, L**). The zygote fails to undergo cytokinesis, with the nucleus disappearing after 3 weeks, leaving an empty space (approximately 3 wpf; **L**). The initial outgrowth of the pseudoperianth (black arrows) and calyptra (red arrows) post-fertilization is not affected in *Mpknox1-6^{9e}* mutants, but their development is arrested as well approximately 1 wpf (**K, L**). (**M, N**) Archegoniophores of *Mpknox1-6^{9e}* mutants begin to senesce approximately 2 weeks after maturation (**N**), while wild-type archegoniophores of the same age remain green (**M**). (**O**) Based on RNA-sequencing (RNA-seq) data and associated gene models, the 3' end of *MpKNOX1* overlaps with the 3' untranslated region (UTR) of *MpSUK1*, which is transcribed from the opposite strand. Predicted full-length *MpKNOX1* transcripts, that is, those including exons 2 and 3, are present primarily in the archegoniophore, consistent with the *in situ* data and semiquantitative reverse transcription-polymerase chain reaction (sqRT-PCR) data in **Figure 1—figure supplement 1**. FPKM (Fragments Per Kilobase of exon model per Million mapped fragments) scales are set to a standard within each tissue to allow potential weakly expressed transcripts to be visualized. (**P**) Signal of *MpSUK1* *in situ* hybridization is detected in antheridia of the antheridiophore. (**Q, R**) Expression of *proMpKNOX1:GUS:SUK1* in an archegoniophore (**R**), an archegonium (**S**), and an antheridiophore (**Q**). Scale bar = 50 μ m (**A, E, G, J, K, L, S**); 100 μ m (**B, C, D, H, I**); 200 μ m (**P**); 500 μ m (**F, R**); 1000 μ m (**Q**).

The online version of this article includes the following source data and figure supplement(s) for figure 3:

Figure supplement 1. *Mpknox1-6^{9e}* mutants rarely form embryos on senescing archegoniophores.

Figure supplement 1—source data 1. Developmental staging data of crosses between *Mpknox1-6* females and wild-type males.

Figure supplement 2. Expression profile of *proMpKNOX1* transcriptional β -glucuronidase reporter gene fusions unravels regulation by antisense transcript *MpSUK1*.

Figure supplement 3. *MpKNOX1A* and *MpKNOX1B* lack the homeodomain and are both primarily expressed in the antheridiophores.

weeks before they senesce (**Figure 3M**). In contrast, *Mpknox1-6^{9e}* archegoniophores begin to senesce around 2 weeks post maturation (**Figure 3N**), suggesting that an *MpKNOX1*-dependent signal is required for archegoniophore maintenance.

At the *MpKNOX1* locus, a convergently transcribed gene, *SUPPRESSOR OF KNOX1* (*MpSUK1*, Mp5g01590), is expressed predominantly in antheridia (**Figure 3O–P**), suggesting that *MpKNOX1* may be regulated by an antisense transcript as has been described for *MpFGMYB* (Mp1g17210) (**Hisanaga, 2019**). Consistent with this hypothesis, a transcriptional fusion of 4.6 kb 5' of *MpKNOX1* coding sequence with a β -glucuronidase (*GUS*) reporter coding sequence and 3' terminator results in expression in both egg cells and antheridia (**Figure 3—figure supplement 2**). In contrast, incorporation of 2.7 kb of the *MpSUK1* antisense transcriptional unit immediately downstream of the *GUS* coding sequence (*proMpKNOX1:GUS:SUK1*) results in female-specific expression (**Figure 3; Figure 3—figure supplement 2**). However, if a 3' terminator sequence is inserted between the *GUS* coding sequence and the *MpSUK1* sequences, the antisense repression is lost, suggesting a requirement of transcriptional overlap but in a non-sequence-specific mechanism (**Figure 3—figure supplement 2**). *MpSUK1* also encodes a 225 aa protein, including a 55 aa domain conserved across most liverworts but not found in other land plant lineages, that does not overlap with the *MpKNOX1* coding sequence (**Figure 3O**).

In addition to the full-length *MpKNOX1* gene, two additional *KNOX1*-related genes, *MpKNOX1A* and *MpKNOX1B*, are encoded in the *M. polymorpha* genome. Both genes encode a conserved *KNOX1* MEINOX domain (**Figure 3—figure supplement 3A**) but lack a HD and are predominantly expressed in the antheridia and sperm cells (**Figure 3—figure supplement 3B–E**).

MpBELL234 are expressed in antheridia, archegonia, and young embryos

The acropetal development of antheridia within antheridiophores results in the youngest being peripheral. *MpBELL2*, *MpBELL3*, and *MpBELL4* expression was detected in stage 2–4 antheridia (**Higo et al., 2016**) within antheridiophores as assayed by *in situ* hybridization (**Figure 4A, C and E**). Likewise, translational *MpBELL2* and transcriptional *MpBELL3* and *MpBELL4* *GUS* reporter gene fusion lines using 3.5 kb (*MpBELL2*), 5.6 kb (*MpBELL3*), and 5.8 kb (*MpBELL4*) of sequence 5' to a transcriptional start site exhibit signals in antheridia of developing antheridiophores (**Figure 4B, D, F and G**). In these reporter lines, no signal was detected in unfertilized archegoniophores or developing sporophytes, nor was any expression observed in the vegetative gametophyte. However, in contrast to the above reporter lines, strand-specific transcriptomic analysis suggests that shorter transcripts, likely produced from an alternative promoter, are generated at the *MpBELL3* and *MpBELL4* loci in archegoniophores and sporophytes (**Figure 4—figure supplement 1A, B**). Indeed, expression of both

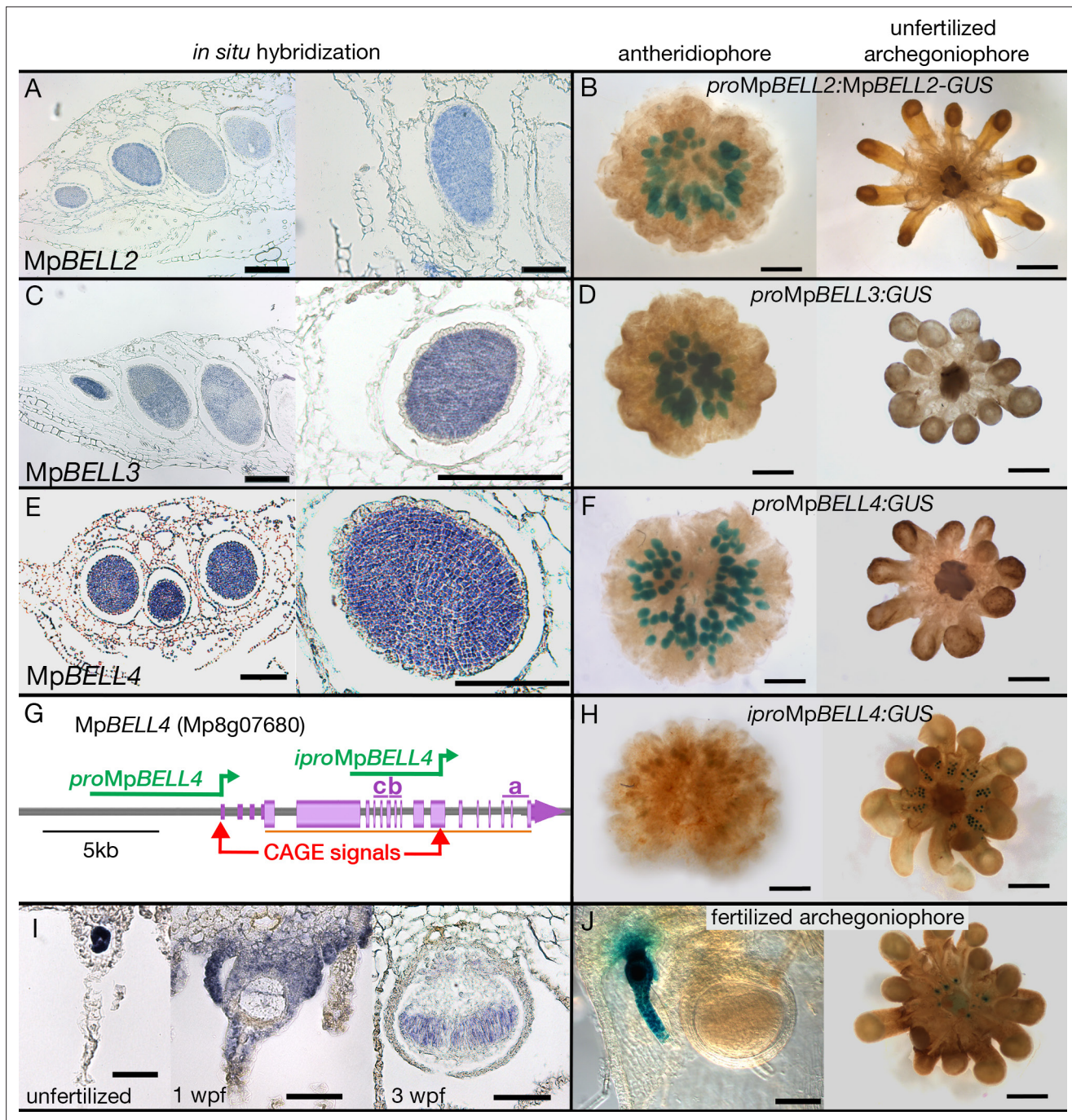


Figure 4. Expression patterns of MpBELL2, MpBELL3, and MpBELL4. *In situ* localization of all three MpBELL mRNAs in antheridia as seen in antheridiophore cross-sections; younger antheridia are toward the left (**A**, **C**, **E**). Antheridiophores with translational (**B**) or transcriptional (**D**, **F**) β -glucuronidase reporter gene fusions of MpBELL2 (**B**), MpBELL3 (**D**), and MpBELL4 (**F**). These reporter genes all harbor sequences 5' of the longest predicted transcript at each of the loci; for example, *proMpBELL4* in (**G**). Signal appears strongest in young- to medium-aged antheridia (stage 3 and stage 4; *Higo et al., 2016*), with the signal being lost in older antheridia toward the center of the antheridiophore (**B**, **F**), possibly due to draining of spermatogenous tissue. Staining was not observed in unfertilized archegonia (**B**, **D**, **F**). (**I**) *In situ* hybridization using the full-length MpBELL4 coding sequence exhibits signals in egg cells of unfertilized archegonia, gametophytic tissues surrounding fertilized archegonia at 1 week post fertilization (wpf), as well as sporophytes up to at least 3 wpf. (**H**, **J**) Reporter genes constructed using an alternative promoter internal of the MpBELL4 locus (*iproMpBELL4*) marked with a Cap Analysis of Gene Expression (CAGE) signal (**G**) exhibit a signal in unfertilized archegonia, but not in antheridia (**H**). Signal remains visible in developing unfertilized archegonia within the center of the archegoniophore, but no signal was observed in sporophytes at 2 wpf (**J**). Scale bar = 200 μ m for left panels in **A**, **C**, **E** and the right panel in **C**, panels **B**, **D**, **F**, **H**, and the right panel in **J**; 100 μ m for right panels in **A**, **E**, **I** and the left panel in **J**; 50 μ m for the left and middle panel in **I**.

Figure 4 continued on next page

Figure 4 continued

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Expression profiles of MpBELL loci.

Figure supplement 2. MpBELL3 locus is a source of multiple transcripts in antheridia versus archegonia/sporophytes.

Figure supplement 3. MpBELL4 transcripts in archegonia/sporophytes lack homeodomains.

genes was detected in egg cells of unfertilized archegonia, gametophytic tissues surrounding fertilized archegonia at 1 wpf, as well as sporophytes up to at least 3 wpf via *in situ* hybridization (**Figure 4I; Figure 4—figure supplement 1C**). An MpBELL4 reporter construct using a 4.2 kb sequence upstream of a CAGE (Cap Analysis of Gene Expression) signal (**Montgomery et al., 2020**) corresponding to the presumed alternative transcriptional start site of the archegonial/sporophytic transcript results in an egg-cell and post-fertilization gametophytic signal; however, no expression was observed in the older sporophytes (**Figure 4G–H and J**). Directed RT-PCR experiments also confirmed the existence of a shorter transcript encoding homeodomain sequences produced from MpBELL3 in archegoniophores and during sporophyte development (**Figure 4—figure supplement 2**).

Paternal and maternal MpBELL are required for proper embryo development

To examine whether MpBELL2/3/4 could provide the male counterpart to the female MpKNOX1, we created loss-of-function alleles for each gene (**Table 1**). When used as male parents in crosses with wild-type females, *Mpbell2*, *Mpbell3*, and *Mpbell4* single mutants did not exhibit an aberrant phenotype; however, when *Mpbell3 Mpbell4* (*Mpbell34*) double mutant or *Mpbell2 Mpbell3 Mpbell4* (*Mpbell234*) triple mutant males were used in crosses with wild-type females, mature sporophytes were formed at a reduced frequency compared to crosses between wild-type parents (**Figure 5A–B**). To quantify the reduction in fertility, we first examined crosses between wild-type parents in greater detail. A majority of fertilization events in crosses between wild-type males and wild-type females produced mature sporophytes (**Figure 5A–B**). However, nearly 40 % of wild-type crosses were either aborted at the zygote stage or arrested at a globular multicellular stage, characteristic of about 1 wpf. As we did not observe sporophytes arrested at later stages, sporophytes that surpass the 1 wpf stage must generally progress to maturity. In contrast, most fertilization events derived from crosses between wild-type females and *Mpbell(2)34* males resulted in sporophyte development that only progressed to a globular stage characteristic of the first week of wild-type development, followed by sporophyte arrest. A smaller number of fertilization events were aborted at zygote formation, resembling the phenotype observed in female *Mpknox1-6^{9e}* mutants (**Figure 5A–B**). However, mature sporophytes with viable spores were produced in about 17 % of fertilization events. In individual crosses in which mature sporophytes form, the corresponding archegoniophores remain viable until after the sporophytes have matured, while unfertilized archegoniophores undergo senescence. The archegoniophores with only arrested sporophytes senesced in a manner similar to unfertilized archegoniophores.

Since both MpBELL3 and MpBELL4 expression was detected in the egg cell, we performed reciprocal crosses between *Mpbell(2)34* females and wild-type males and we observed a similar distribution of sporophyte phenotypes compared to wild-type females crossed with *Mpbell234* males: aborted zygotes, aborted multicellular embryos, and phenotypically normal sporophytes (**Figure 5A–B**). However, in crosses where both the male and female harbored mutations in both MpBELL3 and MpBELL4, no mature sporophytes were formed, with a majority of embryos arrested at the zygote stage (**Figure 5A–B**), a phenotype similar to, albeit less severe than, that observed for *Mpknox1* mutants. Similar sporophyte phenotype ratios in reciprocal crosses between *Mpknox1* and *Mpbell34* have been described independently (**Hisanaga, 2021**).

We next examined whether MpKNOX1 and MpBELL proteins could interact. We chose antheridial-expressed full-length MpBELL4 for analysis due to the truncated nature of MpBELL2 and the extreme length of MpBELL3. In a split Yellow Fluorescent Protein (YFP) BiFC assay in *Nicotiana benthamia* leaf epidermal cells, MpKNOX1 on its own was cytoplasmically localized, but when co-expressed with nuclear localized MpBELL4, MpKNOX1 signal became nuclear (**Figure 5C–F**). A similar interaction was observed with MpKNOX2 and MpBELL1 (**Figure 5—figure supplement 1**). These interactions are selective, since neither interaction between MpKNOX1 and MpBELL1 nor between MpKNOX2

Table 1. Alleles used in this study. Nomenclature is as outlined previously (Bowman et al., 2016c; Montgomery et al., 2020).

Name	Gene number(s)	Type	Details	Gender	Putative effect
Mpknox1-6 ^{9e}	Mp5g01600	CRISPR 2 gRNAs	2360 nt del	Female	Null
Mpbell2-1 ^{9e}	Mp4g09650	CRISPR 1 gRNAs	1 nt in	Female	Hypomorph
Mpbell2-2 ^{9e}	Mp4g09650	CRISPR 1 gRNAs	10 nt del	Male	Hypomorph
Mpbell2-11 ^{9e}	Mp4g09650	CRISPR 1 gRNAs	1 nt del	Male	Hypomorph
Mpbell3-4 ^{9e}	Mp8g02970	CRISPR 2 gRNAs	2217 nt del	Male	Null
Mpbell3-10 ^{9e}	Mp8g02970	CRISPR 2 gRNAs	2218 nt del	Male	Null
Mpbell4-5 ^{9e}	Mp8g07680	CRISPR 2 gRNAs	1135 nt del	Male	Null
Mpbell4-24 ^{9e}	Mp8g07680	CRISPR 2 gRNAs	1135 nt del	Male	Null
Mpbell34-4 ^{9e} (Mpbell3-1 ^{9e} Mpbell4-1 ^{9e})		CRISPR 4 gRNAs		Male	
MpBELL2	Mp4g09650		wt		Wild type
Mpbell3-1 ^{9e}	Mp8g02970		6 nt in/1 nt del; 1 nt in		Hypomorph
Mpbell4-1 ^{9e}	Mp8g07680		1 nt del		Hypomorph
Mpbell234-15 ^{9e} (Mpbell2-15 ^{9e} Mpbell3-15 ^{9e} Mpbell4-15 ^{9e})		CRISPR 4 gRNAs		Male	
Mpbell2-15 ^{9e}	Mp4g09650		23 nt in/10 nt del		Hypomorph
Mpbell3-15 ^{9e}	Mp8g02970		49 nt in/1 nt del; 1 nt in		Hypomorph
Mpbell4-15 ^{9e}	Mp8g07680		1 nt del		Hypomorph
Mpbell234-83 ^{9e} (Mpbell2-83 ^{9e} Mpbell3-83 ^{9e} Mpbell4-83 ^{9e})		CRISPR 6 gRNAs		Male	
Mpbell2-83 ^{9e}	Mp4g09650		110 nt in/1401 nt del		Null?
Mpbell3-83 ^{9e}	Mp8g02970		2217 nt del		Null?
Mpbell4-83 ^{9e}	Mp8g07680		7 nt in/1150 nt del		Null?
Mpbell234-7 ^{9e} (Mpbell2-7 ^{9e} Mpbell3-7 ^{9e} Mpbell4-7 ^{9e})		CRISPR 4 gRNAs		Female	

Table 1 continued on next page

Table 1 continued

Name	Gene number(s)	Type	Details	Gender	Putative effect
Mpbell2-7 ^{9e}	Mp4g09650		16 nt in/8 nt del		Hypomorph
Mpbell3-7 ^{9e}	Mp8g02970		1 nt.in/30 nt del; wt		Hypomorph
Mpbell4-7 ^{9e}	Mp8g07680		44 nt in		Hypomorph
Mpbell234-18 ^{9e} (Mpbell2-18 ^{9e} Mpbell3-18 ^{9e} Mpbell4-18 ^{9e})		CRISPR 4 gRNAs		Female	
Mpbell2-18 ^{9e}	Mp4g09650		13 nt in/7 nt del		Hypomorph
Mpbell3-18 ^{9e}	Mp8g02970		97 nt in/304 nt del; 1 nt del		Hypomorph
Mpbell4-18 ^{9e}	Mp8g07680		1 nt del		Hypomorph
<small>pro</small> EF1:XVE>> amiR-MpE(z) ^{15kmIR166}	Mp5g18040	Transgene	Knock-down	Both	Hypomorph
<small>pro</small> EF1:MpKNOX1	Mp5g01600	Transgene	Ectopic overexpression	Both	Hypermorph
<small>pro</small> EF1:XVE>> MpBELL3	Mp8g02970	Transgene	Inducible overexpression	Both	Hypermorph
<small>pro</small> EF1:MpKNOX1; <small>pro</small> EF1:XVE>> MpBELL3	Mp5g01600	Transgene	Ectopic overexpression	Both	Hypermorph
<small>pro</small> MpBELL2:MpBELL2-GUS	Mp4g09650	Transgene	Translational β -glucuronidase gene fusion	Both	
<small>pro</small> MpBELL3:GUS	Mp8g02970	Transgene	Transcriptional β -glucuronidase gene fusion	Both	
<small>pro</small> MpBELL4:GUS	Mp8g07680	Transgene	Transcriptional β -glucuronidase gene fusion	Both	
<small>internal_pro</small> MpBELL4:GUS	Mp8g07680	Transgene	Transcriptional β -glucuronidase gene fusion	Both	
<small>pro</small> MpKNOX1:GUS	Mp5g01600	Transgene	Transcriptional β -glucuronidase gene fusion	Both	
<small>pro</small> MpKNOX1:GUSterm:SUK1	Mp5g01600 Mp5g01590	Transgene	Transcriptional β -glucuronidase gene fusion	Both	
<small>pro</small> MpKNOX1:GUS:SUK1term	Mp5g01600 Mp5g01590	Transgene	Transcriptional β -glucuronidase gene fusion	Both	
<small>pro</small> MpKNOX1A:MpKNOX1A-GUS	Mp4g12450	Transgene	Translational β -glucuronidase gene fusion	Both	
<small>pro</small> MpKNOX1B:MpKNOX1B-GUS	Mp2g11140	Transgene	Translational β -glucuronidase gene fusion	Both	

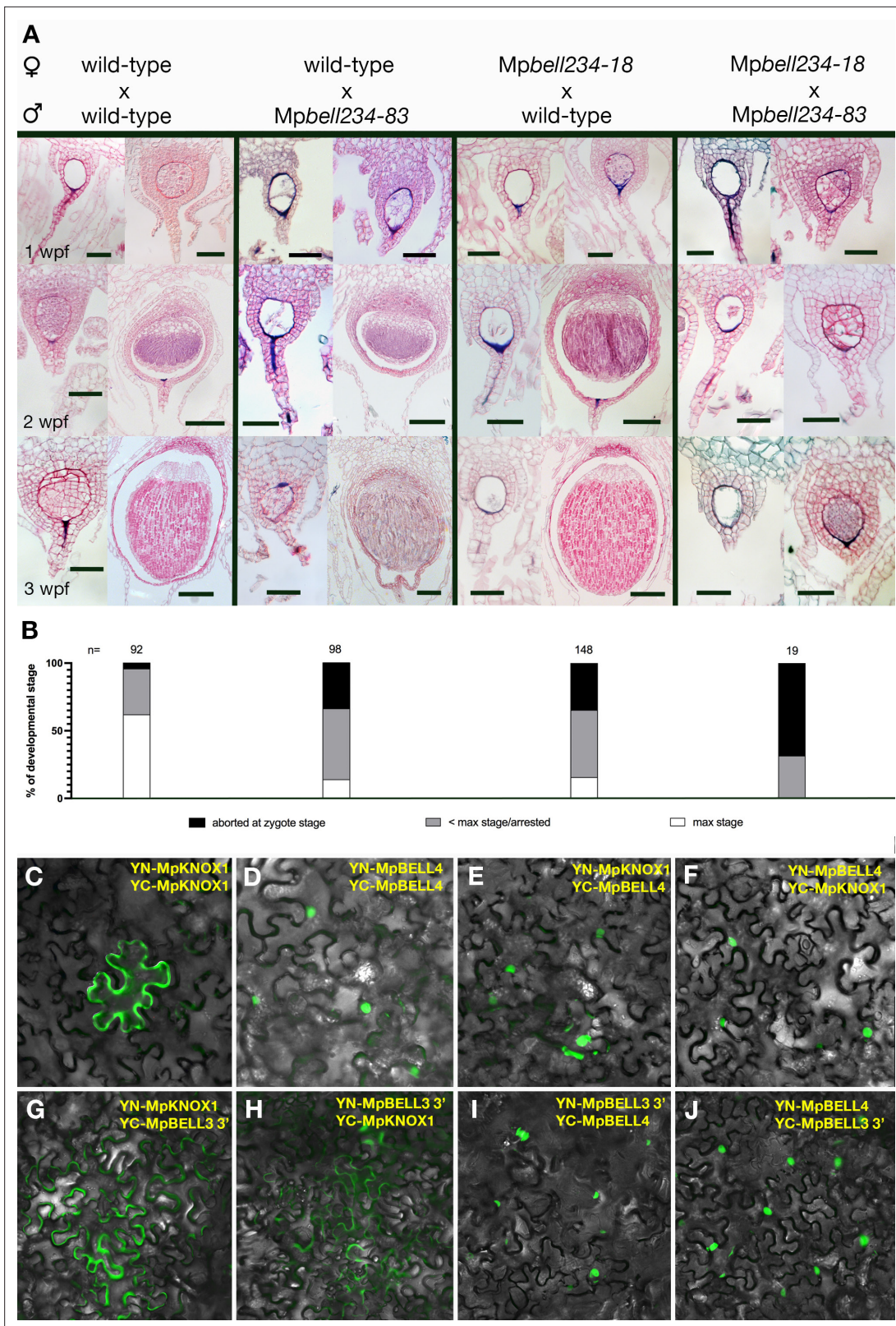


Figure 5. Sporophyte development entails BELL activity in both male and female gametes. (A) Wild-type females crossed with wild-type males (left column), reciprocal crosses between wild-type and *Mpbell* mutant alleles (middle two columns), and crosses between *Mpbell234* homozygotes (right column) observed after 1, 2, and 3 weeks post fertilization (wpf). All crosses produced aborted embryos or arrested sporophytes (left pictures in each panel). However, crosses between *Mpbell234-18*⁹⁰ and *Mpbell234-83*⁹⁰ only produced embryos aborted at the zygote stage or arrested at approximately
Figure 5 continued on next page

Figure 5 continued

the 1 wpf stage. See Table 1 for details on mutant alleles. (B) Percentages of developmental stages observed for each of the crosses. Plants were crossed once and then examined after 1, 2, or 3 wpf. Observed embryos were grouped into the following developmental stages: aborted zygotes, sporophytes that were arrested at a stage younger than the maximum stage possible, and sporophytes that had reached the maximum stage expected. N = the total number of observed fertilization events. See **Figure 5—source data 1** for details. (C–F) BiFC (Bimolecular fluorescence complementation) assay of protein-protein interaction of MpKNOX1 and MpBELL4 in *Nicotiana benthamiana* leaves. Homodimers of MpKNOX1 show cytoplasmic localization (C), while MpBELL4 homodimers localize to the nucleus (D). Co-expression of both proteins resulted in nuclear localization of heterodimers (E, F). (G–J) BiFC assays of the protein derived from the shorter 3' transcript of MpBELL3 (**Figure 4—figure supplement 2**) with MpKNOX1 show weak cytoplasmic interaction (G, H). However, MpBELL4 interacts strongly with _{short}MpBELL3 and results in nuclear localization (I, J). Scale bar = 50 μm, for all left pictures in (A) panels, all double mutant pictures, and all 1 wpf pictures; 100 μm for right pictures of the 2 wpf panel; 200 μm for all right pictures of the 3 wpf panel.

The online version of this article includes the following figure supplement(s) for figure 5:

Source data 1. Developmental staging data of crosses between wild-type and *Mpbell234* mutants.

Figure supplement 1. MpKNOX1/MpBELL4 and MpKNOX2/MpBELL1 exhibit specificity in their interactions.

Figure supplement 2. Co-expression strategy to express MpKNOX1 and MpBELL3 simultaneously in the vegetative gametophyte.

and MpBELL4 was observed (**Figure 5—figure supplement 1**). In contrast to full-length MpBELL4, we did not observe any efficient interaction between the shorter sporophyte-expressed MpBELL3 and MpKNOX1 in this heterologous system (**Figure 5G–H**). Surprisingly, however, in this system, full-length MpBELL4 and sporophyte-expressed MpBELL3 interacted (**Figure 5I–J**).

Expression of the fifth *M. polymorpha* MpBELL gene, MpBELL5, may be restricted to the archeogoniophore among tissues examined (**Figure 4—figure supplement 1A, B**). As with MpKNOX1 (**Figure 3**) and MpFGMYB (*Hisanaga, 2019*), the MpBELL5 transcript may also be opposed by a convergently transcribed locus in antheridia (**Figure 4—figure supplement 1A, B**).

Activation of sporophyte gene expression in the vegetative gametophyte

The expression patterns and mutant phenotypes of MpBELL2/3/4 and MpKNOX1 genes are consistent with a role in activating diploid gene expression, and thus we examined whether ectopic (co-) expression of these genes in vegetative gametophyte is sufficient to activate diploid gene expression (**Figure 5—figure supplement 2A**). Ectopic expression of MpKNOX1 alone does not alter expression of MpKNOX2, MpBELL1, or MpBELL3 in the vegetative gametophyte (**Figure 5—figure supplement 2B**). However, ectopic expression of either MpBELL3 or co-expression of MpKNOX1 and MpBELL3 in the vegetative gametophyte for 72 hr is sufficient to activate both MpKNOX2 and MpBELL1, whose expression is normally limited to sporophyte development, or, in the case of MpBELL1, also after continuous far-red light induction (*Inoue et al., 2019*; **Figure 5—figure supplement 2B**). Thus, in this context, MpBELL3 alone can activate MpKNOX1 expression, a scenario perhaps reminiscent of post-zygotic activation of MpKNOX1 after fertilization (**Figure 3C, Figure 3—figure supplement 2D**).

Discussion

An ancestral function for TALE-HD genes in Viridiplantae

The eukaryotic life cycle alternates between haploid and diploid phases, initiated by meiosis and gamete fusion, respectively. Organisms spanning the phylogenetic diversity of eukaryotes, including ascomycete and basidiomycete fungi, Amoebozoa, brown algae, and chlorophyte algae, utilize paralogous homeodomain proteins that heterodimerize following gamete fusion to initiate the diploid genetic program, lending support for the idea that this may have been an ancestral function of homeodomain proteins. Our observations in *M. polymorpha* indicate that TALE-HD proteins, specifically MpKNOX1 and MpBELL2/3/4, are initially supplied in gametes, and that this gametic expression is required for diploid sporophyte development (**Figure 6**). MpKNOX1 is absolutely required in the egg cell, as is MpBELL3/4 in either the sperm or egg. This scenario is reminiscent of that characterized in *Chlamydomonas*, wherein KNOX is supplied by the (-) gamete and BELL via the (+) gamete, and that once together in the zygote, the diploid genetic program can be activated. Thus, the basic tenets of the genetic regulation of the haploid-to-diploid transition are conserved in two widely divergent

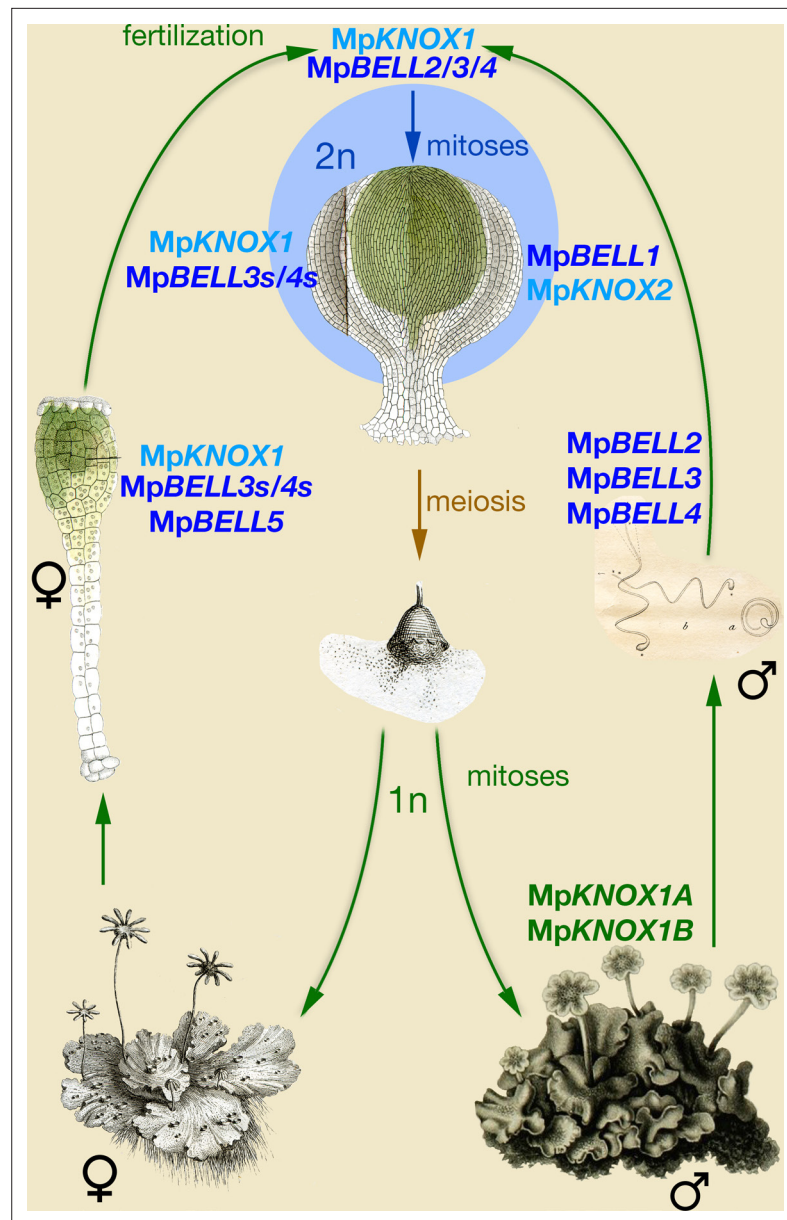


Figure 6. Model of KNOX and BELL function in *Marchantia polymorpha*. Three MpBELL genes (MpBELL2/3/4) are expressed in antheridia in the cells that will differentiate into sperm cells. MpKNOX1 is expressed in the egg cell of the archegonia, along with three MpBELL genes (MpBELL5 and MpBELL3/4). In contrast to the antheridia, where full-length transcripts are found, alternative short transcripts of MpBELL3/4 (MpBELL3s and MpBELL4s; see **Figure 4—figure supplements 2 and 3**) are produced in the egg cell. After fertilization, MpBELL3/4 derived from both sperm and egg cells and MpKNOX1 derived from the egg cell act together to activate the zygotic program. MpBELL1 and MpKNOX2 form a distinct heterodimer that acts at later stages of sporophyte development. The homeodomain lacking MpKNOX1A and MpKNOX1B genes expressed in the antheridia along with MpBELL2/3/4 potentially prevents functionality by sequestering MpBELL2/3/4 proteins. The function of MpBELL5 is unknown. Images are obtained from **Marchant, 1713; Mirbel, 1835; Thuret, 1851; and Unger, 1837.**

The online version of this article includes the following figure supplement(s) for figure 6:

Figure supplement 1. Phylogenetic perspective on HD gene-mediated haploid-to-diploid transitions in eukaryotes.

Viridiplantae lineages, consistent with the notion that such a system was present in the common ancestor of extant Viridiplantae.

The obvious major difference between the roles of KNOX and BELL in *Marchantia* when compared to *Chlamydomonas* is the activity of MpBELL34 in both the egg and sperm, rather than being confined to the sperm. Thus, questions on how the system operates mechanistically and why it might have evolved are paramount. The functions of maternal MpKNOX1 and paternal MpBELL34 could be presumed to represent the ancestral condition, with the role of maternal MpBELL34 being a derived condition. The absolute requirement for maternal MpKNOX1 (Figure 3) and its ability to heterodimerize with full-length MpBELL3/4 (Figure 5), which is expressed only paternally (Figure 4), suggest that this ancestral system functions in *Marchantia*, albeit with a reduced efficacy. In *Marchantia*, although neither maternal nor paternal MpBELL is absolutely required, both are necessary for efficient sporophyte development. One plausible hypothesis is that maternally supplied MpBELL3/4 evolved as a backup system to ensure diploid development following fertilization, concomitant with the evolution of sperm with reduced cytoplasmic content and highly condensed nuclei, anatomical features that could reduce the efficiency of MpBELL protein delivery by sperm. That both MpKNOX1 and ^{short}MpBELL3 are present in the egg cell prior to fertilization and that they do not efficiently interact, at least in a split YFP BiFC assay in *Nicotiana* leaf cells (Figure 5), suggests that the act of fertilization may trigger a biochemical change required for their interaction. Although ^{short}MpBELL3 can interact with full-length MpBELL4 (Figure 5), it is likely not critical in planta, given the absolute requirement of MpKNOX1 for sporophyte development. The requirement of both maternal and paternal MpBELL for efficient sporophyte production may also provide a mechanism to ensure activation of sporophyte development only occurs following fertilization. Finally, it is of note that a similar system, wherein two TALE-HD genes are both expressed in male and female gametes, has been described for the brown alga *Ectocarpus* (Arun et al., 2019). As *Ectocarpus* also possesses multicellular haploid and diploid generations and is anisogamous, perhaps similar evolutionary pressures contributed to the independent evolution of similar systems.

Given the phylogenetic affinity of mosses as a sister group to liverworts (Renzaglia et al., 2018), we might expect a similar KNOX/BELL program to regulate the haploid-to-diploid transition in *Physcomitrium*. The phenotype of Mpknx1 alleles is more extreme than triple-null loss-of-function alleles of all PpKNOX1 loci where a sporophyte with viable spores forms (Sakakibara et al., 2008), suggesting that PpKNOX2 may also play a role in the transition in this species (Sakakibara et al., 2013). In *Physcomitrium*, PpBELL1 expression has been observed in sperm (Ortiz-Ramírez et al., 2017) via transcriptome analysis and in the egg via reporter gene (Horst et al., 2016). While the latter study reported a lack of male expression, this could be due to a loss of expression in this *Physcomitrium* accession (Meyberg et al., 2020), or alternatively, if the *Physcomitrium* PpBELL1 genomic architecture is similar to MpBELL3/4, the reporter gene insertion may have excluded regulatory sequences required for alternative transcription start sites. Hence, aspects of the ancestral system regulating the haploid-to-diploid transition could be present in multiple early diverging land plant lineages.

In contrast to the situation in bryophytes, in angiosperms, paternally supplied pluripotency factors related to BABYBOOM, an AP2/ERF transcription factor, can activate the zygotic genetic program (Khanday et al., 2019; Conner et al., 2017; Conner et al., 2015). Furthermore, despite the extensive literature on KNOX/BELL function in angiosperms, the only potential known vestige of this ancestral system existing in angiosperms is the presence of KNOX2 in the *Arabidopsis* female gametophyte, possibly in the egg cell (Pagnussat et al., 2007). Thus, it seems that the ancestral function has been lost entirely in some derived land plants, as it has presumably in the metazoan lineage (Bowman et al., 2016b).

Since liverworts are placental organisms, with the diploid sporophyte nourished by the maternal haploid gametophyte, there is opportunity for extensive intergenerational communication. For example, that the maternal gametophytic calyptra and pseudoperianth initiate their development in Mpknx1 mutants implicates a MpKNOX1-independent non-cell autonomous zygote-derived signal activated post-fertilization. In contrast, the premature senescence of Mpknx1 archegoniophores suggests a MpKNOX1-dependent intra-gametophytic signal for the maintenance of archegonial viability. Further, the premature senescence of wild-type archegoniophores bearing only aborted embryos suggests a signal emanating from older sporophytes to maintain the viability of the maternal archegoniophore tissues until sporophyte development is completed. Finally, that aborted embryos

are found even in crosses involving only wild-type parents suggests possible maternal control over allocation of resources across multiple embryos, a phenomenon similar to maternal control over fruit set in angiosperms (Stephenson, 1981).

Roles of MpKNOX1 in *Marchantia* sporophyte

A key land plant innovation was the evolution of a multicellular diploid generation, the embryo, via mitoses, interpolated between gamete fusion and meiosis (Bower, 1908). In seed plants, ferns, and mosses, KNOX1 activity is associated with continued sporophyte cell proliferation, including sporophyte apical meristem activity (Hay and Tsiantis, 2010; Sakakibara et al., 2008; Sano et al., 2005). While *M. polymorpha* and other liverworts lack apical meristems during sporophyte development (Kienitz-gerloff, 1874), MpKNOX1 expression is detected throughout developing sporophytes until the inception of sporogenous cell differentiation (Figure 3), consistent with MpKNOX1 having a role in cell proliferation during *M. polymorpha* sporophyte development. Thus, while MpKNOX1 retains the ancestral function in the haploid-to-diploid transition, the lack of zygotic cell division in Mpknx1 mutants and the MpKNOX1 sporophytic expression pattern suggest that neofunctionalization of KNOX1 contributed to the evolution of the embryo via stimulation of cell proliferation. That MpBELL3/4 are also expressed during these stages of sporophyte development (Figure 4; Figure 4—figure supplements 1 and 2, 3) suggests these genes encode the heterodimeric partners of MpKNOX1 during sporophyte development. This is supported by the subfunctionalization of MpBELL genes with respect to heterodimerization partners—MpBELL3/4 specifically heterodimerize with MpKNOX1 while MpBELL1 interacts only with MpKNOX2 (Figure 5—figure supplement 2). This subfunctionalization resembles that observed for KNOX/BELL interactions in angiosperms (Furumizu et al., 2015) but represents an independent evolutionary event since all angiosperm BELL genes are MpBELL1 orthologs.

In eukaryotic lineages in which multicellularity evolved in the diploid phase of the life cycle, genes involved in the haploid-to-diploid transition, or paralogs thereof, have been repeatedly co-opted into roles directing development in the diploid generation. Most conspicuously, despite the loss of their role in zygotic gene activation (see below), both non-TALE- (eg, Hox) and TALE-HD genes act to pattern the metazoan body (Merabet and Mann, 2016; Pearson et al., 2005; Lewis, 1978). Likewise, in the basidiomycete fungus *Coprinopsis*, the same HD heterodimer that initiates the diploid genetic program also directs early developmental stages in the multicellular diploid (Kamada, 2002). In the syncytial chlorophyte alga *Caulerpa lentillifera*, differential TALE-HD expression in the diploid body was speculated to influence differentiation of fronds and stolons, but functional data are lacking (Arimoto, 2019). Finally, in land plants, as the multicellular sporophyte evolved increasing complexity, KNOX/BELL genetic modules were co-opted to direct development of novel organs and tissues via regulation of cell proliferation (KNOX1) or differentiation (KNOX2) in conjunction with other gene regulatory networks (Furumizu et al., 2015; Hay and Tsiantis, 2010).

The remaining TALE-HD genes of *M. polymorpha*

The two KNOX1-related sequences lacking homeodomains (MpKNOX1A and MpKNOX1B) are reminiscent of similar, albeit independently evolved, proteins in angiosperms that act as inhibitors of KNOX/BELL function by forming inactive heterodimers primarily with BELL partners (Magnani and Hake, 2008; Kimura et al., 2008). The antheridial expression of MpKNOX1A and MpKNOX1B prompts the hypothesis that the encoded proteins may act as a safeguard against the inappropriate activity of MpBELL in the male gametophyte. This is particularly pertinent since ectopic expression of MpBELL3 alone in the vegetative gametophyte can activate MpKNOX1, MpKNOX2, and MpBELL1 (Figure 5—figure supplement 2). The lack of an aberrant phenotype observed in lines constitutively expressing MpKNOX1 in male gametophytes would be consistent with this hypothesis.

In contrast to the other MpBELL genes, MpBELL5 transcripts are only detected in archegoniophores. Given the strong heterodimerization affinity preferences of BELL proteins for KNOX proteins (Bellaoui et al., 2001; Joo et al., 2018), the only conspicuous potential partner in *M. polymorpha* would be MpKNOX1. Hence, MpBELL5 might play an as-yet-undefined role in the activation of zygotic gene expression. At the MpBELL5 locus, a convergently transcribed locus is expressed predominantly in antheridia (Figure 4—figure supplement 1), suggesting MpBELL5 may be repressed by an anti-sense transcript in a manner similar to MpKNOX1 (Figure 3) and MpFGMYB (Hisanaga, 2019). Such

male-expressed antisense transcript-mediated repression apparently provides a general mechanism for female-specific expression of autosomal genes, with sex-specific regulation ultimately being linked to the feminizing locus on the female sex chromosome (Bowman, 2016a; Knapp, 1935; Lorbeer, 1936).

Marchantia BELL gene diversity resembles that of charophyte algae

The phylogenetic diversity of BELL genes in *M. polymorpha*, and liverworts more broadly, more closely resembles the diversity observed in charophyte algae than that of other land plants (Joo et al., 2018; Lee et al., 2008). All previously described land plant BELL genes form a single clade that evolved within the charophycean algae (Figure 2). The single *M. polymorpha* gene in this clade, MpBELL1, is the only MpBELL gene predominantly expressed during sporophyte development. The other MpBELL genes reside in two other phylogenetically distinct clades, GLX-basal and BELL-gamete, both of which include charophyte sequences (Figure 2). This phylogenetic pattern suggests an ancient proliferation of BELL paralogs within the charophycean algal ancestor, with *M. polymorpha* (and other liverworts and some mosses) retaining this diversity that was subsequently lost in most other land plant lineages. The structural diversity (two or three homeodomains) of the MpBELL2/3/4 paralogs and the extensive sequence diversity of BELL-gamete clade genes are consistent with rapid evolution of genes involved in reproductive processes (Swanson and Vacquier, 2002). Finally, the expression pattern of MpBELL5, a member of the GLX-basal clade that evolved early in the charophytes, also suggests an as-yet-unresolved function in reproduction. Similarly, the *Chlamydomonas* genome encodes a second BELL paralog, HDG1, that is expressed in both (+) and (-) gametes, suggesting a role in reproduction, but whose function is unknown.

An ancestral function for TALE-HD genes in eukaryotes

In Viridiplantae (*Chlamydomonas* and *Marchantia*), the haploid-to-diploid transition is mediated by two TALE-HD genes. In the red alga *Pyropia yezoensis*, KNOX gene expression is detected in the diploid conchosporangium, but not in haploid thalli (Mikami et al., 2019). The phylogenetic distribution of KNOX and BELL subfamilies and their heterodimerization affinities (Joo et al., 2018) suggest that the Archaeplastida common ancestor utilized KNOX/BELL TALE-HD genes to mediate the haploid-to-diploid transition. Likewise, in the brown alga *Ectocarpus*, two TALE-HD proteins, OUROBOROS and SAMSARA, mediate the transition (Arun et al., 2019). In contrast, in both ascomycete and basidiomycete fungi, the haploid-to-diploid transition is mediated by heterodimerization of a TALE-HD and a non-TALE-HD protein (Gillissen et al., 1992; Goutte and Johnson, 1988; Herskowitz, 1989; Hull et al., 2005; Kues et al., 1992; Spit et al., 1998; Urban, 1996). In the Amoebozoa *Dictyostelium*, the two homeodomain-like proteins controlling the haploid-to-diploid transition are highly divergent, rendering the phylogenetic affinities enigmatic (Hedgethorpe et al., 2017). As these taxa span eukaryotic phylogenetic diversity, one role of homeodomain genes in the ancestral eukaryote was to regulate the haploid-to-diploid transition—the evolution of the homeodomain in the ancestral eukaryote was associated with evolution of a novel life cycle (Figure 6—figure supplement 1). In arguably the most intensively studied taxon, the Metazoa, zygotic gene activation has been replaced by maternally derived pluripotency factors (Schulz and Harrison, 2019), but homeodomain genes have been retained.

Given the ancestral eukaryote possessed a minimum of two HD genes, including one TALE-HD and one non-TALE-HD (Bharathan et al., 1997; Bürglin, 1997; Derelle et al., 2007), it is an intriguing question as to whether, ancestrally, both TALE and non-TALE paralogs acted in the diploid-to-haploid transition. Either possible ancestral scenario (TALE + non-TALE or TALE + TALE) requires either a co-option of a pre-existing paralog (eg, non-TALE), or alternatively, a TALE gene duplication and transference of function to the new paralog, respectively. To resolve the ancestral condition, broader phylogenetic sampling and functional analyses across additional unicellular Bikont lineages, particularly in the Excavata and the paraphyletic sister groups of the Metazoa (choanoflagellates, Filasterea, and Ichthyosporea), might be informative (Figure 6—figure supplement 1). Resolution of the ancestral condition could inform whether the primary ancestral function of both TALE-HD and non-TALE-HD was regulating the haploid-to-diploid transition or whether non-TALE-HD genes had another fundamental function in ancestral eukaryotes.

Materials and methods

Ecotypes, plant growth, transformation, and induction

M. polymorpha ssp *ruderalis*, ecotype BoGa, obtained from the Botanical Garden of Osnabrueck, Germany, was used in SZ lab. Plant cultivation, transformation, and induction were performed according to *Ishizaki et al., 2008* and *Althoff, 2014*. *M. polymorpha* ssp *ruderalis*, ecotype MEL, was used in JLB lab and grown, transformed, and induced according to *Ishizaki et al., 2008* and *Flores-Sandoval et al., 2015*. In case of double transformations, sporelings were co-transformed using two constructs featuring different selectable markers. Spores were drained and plated on selection media for approximately 2 weeks and then subjected to a second round of selection before being transferred to ½ B5 plates. For induction of reproductive organs, plants were transferred to white light supplemented with far red light (735 nm; 45 μmol/m²/s) on ½ B5 media supplemented with 1 % glucose.

Genotyping, RNA extraction, cDNA synthesis, and semiquantitative reverse transcription-polymerase chain reaction

DNA was extracted using a modified protocol from *Edwards et al., 1991*. Instead of vacuum drying, the pelleted DNA was air-dried. Amplicons were directly sequenced. RNA was extracted from ~100 mg wet weight tissue using Rneasy kit from Qiagen following the manufacturer's instructions, including an on-column Dnase I digestion. The same amount of total RNA was used for complementary DNA (cDNA) synthesis using either Superscript II (Invitrogen) (SZ lab) or Bioscript (Bioline) (JLB lab) reverse transcriptase according to the manufacturer's instructions with oligo-dT₁₅ primer. 5' and 3' RACE-ready cDNA synthesis was performed using SMARTScribe reverse transcriptase (Takara) according to the manufacturer's instructions but with the addition of random primer mix (NEB) to allow reverse transcription of long mRNA templates. See **Supplementary file 1, Table 1a** for primer sequences.

In situ hybridization, sectioning, staining, and microscopy

M. polymorpha ssp *ruderalis*, ecotype BoGa, tissue fixation, embedding, sectioning, and hybridization with digoxigenin (DIG)-labeled antisense RNA probes were performed according to *Zachgo, 2002*. Sections were either stained in toluidine blue O (TBO) alone or counterstained with ruthenium red (RR) according to *Retamales and Scharaschkin, 2014*. Staining time for RR and TBO was 90 s and 60 s, respectively. Microscopic slides were observed using a Leica MZ16 FA microscope and pictures were taken with a Leica DFC490 camera. Plants were observed using either a Lumar dissecting microscope (Zeiss) and photographed with AxioCam HRc and AxioVision software (both Zeiss) or a Leica stereomicroscope (Leica M165 FC) and photographed with an integrated Leica DFC490 camera. See **Supplementary file 1, Table 1b** for probe sequences.

Bimolecular fluorescence complementation

Coding sequences of KNOX and BELL genes were seamlessly cloned into the pGREEN derivatives (*Hellens et al., 2000*) pAMON/pSUR for N-terminal split VENUS (I152L) fusion and into pURIL/pDOX for C-terminal split VENUS (I152L) fusion (*Lampugnani et al., 2016*). All constructs feature the 35 S promoter driving the N- (pAMON/pURIL) or C-terminal (pSUR/pDOX) part of the split VENUS (I152L) fluorophore for translational fusion of the gene-of-interest in frame with the split VENUS (I152L). All constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 for transfection of plant leaves (*Bracha-Drori et al., 2004*). For fluorescence microscopy, usually three, approximately 50 × 50 mm, abaxial leaf tissue fragments were examined under an Axioskop two mot plus (Zeiss) microscope and photographed using AxioCam HRc and AxioVision software. For YFP filter set 46, excitation BP 500/20; beam splitter FT 515; emission BP 535/30 was used. See **Supplementary file 1, Table 1c** for primer sequences.

Overexpression and transcriptional and translational fusion constructs

Primers used for overexpression with the endogenous, constitutive MpEF1α promoter (*Althoff, 2014*), inducible overexpression, and transcriptional and translational fusion constructs are shown in **Supplementary file 1, Table 1c**. The complete 1.8 kb MpKNOX1 coding sequence and a 4.8 kb fragment of the MpBELL3 coding sequence were amplified from cDNA and seamlessly cloned (NEBuilder, NEB) into pENTR-D (Invitrogen) via NotI/Ascl sites. The MpKNOX1 coding sequence was subsequently

recombined into pMpGWB403 (Addgene entry #68668) for constitutive expression, whereas the MpBELL3 coding sequence was recombined into the estrogen-inducible binary pHART XVE (Flores-Sandoval et al., 2016). Regulatory sequences upstream of the transcriptional start site of MpKNOX1 (4.6 kb), MpBELL3 (5.6 kb), and MpBELL4 (5.8 kb) and internal of the MpBELL4 locus (4.1 kb) were amplified and cloned into pENTR-D and subsequently recombined using LR-clonase II (Invitrogen) into pMpGWB104 (Addgene entry #68558) featuring the GUS reporter gene. To elucidate the transcriptional regulation at the MpKNOX1/MpSUK1 locus, the 4.6 kb regulatory region of MpKNOX1 was seamlessly cloned (NEBuilder, NEB) into the HindIII/XbaI site 5' of the Gateway cassette of pMpGWB401 (Addgene entry #68666). The reversely transcribed MpSUK1 locus (2.7 kb) was seamlessly cloned into either the SacI site or AscI site (5' or 3' of the NOS terminator). Subsequently, the GUS reporter gene was recombined using pENTR-GUS (Invitrogen). For translational fusions of MpBELL2, MpKNOX1A, and MpKNOX1B, the regulatory sequence 5' upstream of the transcriptional start site including the genomic gene locus and excluding the stop codon was amplified and cloned into pRITA, a shuttle vector featuring the GUS reporter gene and NOS terminator. MpBELL2 and MpKNOX1A were cloned into the KpnI or Sall/KpnI site, respectively. MpKNOX1B was seamlessly cloned using Infusion cloning (Clontech). All constructs were subsequently transferred to the binary vector pHART using NotI.

CRISPR/Cas9-mediated gene knock-out

Potential guide RNAs (gRNAs) were obtained and screened for off targets in the *M. polymorpha* genome using CRISPOR website (<http://crispor.tefor.net/>). Quadruple gRNAs were each cloned into pMpGE_En04, pBC-14, pBC-23, or pBC-34 using BsaI sites. Double gRNAs were cloned into pMpGE_En04 and pBC-14 (Hisanaga, 2019). These guides were subsequently assembled into the pMpGE_En04 backbone using BglI sites. pMpGE_En04 were gateway cloned into pMpGE010 (Sugano et al., 2018; Sugano et al., 2014; Hisanaga, 2019; Addgene entry #71,536), featuring Cas9 nuclease. In case of six gRNAs, the remaining two were cloned into pMpGWB401 (Ishizaki et al., 2015; Addgene entry #68,666) and plants doubly transformed (Ishizaki et al., 2008). See **Supplementary file 1, Table 1d** for gRNA sequences.

GUS staining

Generally, detection of GUS activity in a minimum of three independent lines was performed overnight at 37 °C in GUS staining solution (0.05 M NaPO₄, pH 7.2, 2 mM K₃[Fe(CN)₆], 2 mM K₄[Fe(CN)₆], 0.3 % (v/v) Triton X-100) supplemented with 0.6 mg/ml b-D-glucopyranosiduronic acid (X-gluc) dissolved in N,N-dimethylformamide (DMF). Only detection of *pro*MpKNOX1:GUS activity in archegoniophores was performed for 4 hr in staining solution containing 5 mM K₃[Fe(CN)₆] and 5 mM K₄[Fe(CN)₆]. Antheridiophores were cleared in 70 % (w/v) chloral hydrate in 10 % (v/v) glycerol solution.

Phylogenetics

Predicted TALE-HD sequences were assembled from land plants, charophytes, and chlorophytes, via Genbank or additional sources as detailed below. Gymnosperm sequences were obtained from Congenie (congenie.org), fern sequences from *Equisetum* and *Azolla* transcriptomes (de Vries et al., 2016; Vanneste et al., 2015), bryophyte sequences, other than *M. polymorpha* (Phytosome) and *P. patens* (Genbank), from available transcriptomes (Wickett et al., 2014; Dong et al., 2019; Radhakrishnan et al., 2020), *Klebsormidium* and *Caulerpa* sequences derived from genome sequences (Hori et al., 2014; Arimoto, 2019), and other charophyte and chlorophyte sequences from Genbank and additional transcriptomes (Cooper and Delwiche, 2016; Ju et al., 2015).

Complete or partial coding nucleotide sequences were manually aligned as amino acid translations using Se-AL v2.0a11 for Macintosh (<http://tree.bio.ed.ac.uk/software/seal/>). We excluded ambiguously aligned sequences to produce an alignment of 231 nucleotides (77 amino acids) for 132 BELL sequences. Alignments of KNOX genes included the homeodomain, MEINOX (KNOX) and ELK domains (Joo et al., 2018), comprising 630 nucleotides (210 amino acids) for 124 KNOX sequences. Alignments of nucleotides and amino acids were employed in the subsequent Bayesian analysis. Bayesian phylogenetic analysis was performed using Mr Bayes 3.2.1 (Huelsenbeck and Ronquist, 2001; Huelsenbeck et al., 2001). The Bayesian analyses for the nucleotide data sets were run for 15,000,000 (BELL) or 5,000,000 (KNOX) generations, which was sufficient for convergence of the two

simultaneous runs (BELL, 0.0357; KNOX, 0.0423). In both cases, to allow for the burn-in phase, 50 % of the total number of saved trees were discarded. The graphic representation of the trees was generated using the FigTree (version 1.4.0) software (<http://tree.bio.ed.ac.uk/software/figtree/>). Sequence alignments and command files used to run the Bayesian phylogenetic analyses can be provided upon request.

Acknowledgements

We thank our colleagues at the Joint Genome Institute (JGI) for their work producing the *Marchantia* genome sequence; the work conducted by the US Department of Energy (DOE) JGI, a DOE Office of Science User Facility, is supported by the Office of Science of the US DOE under contract DE-AC02-05CH11231. This work was supported by Monash University and received funding from the Australian Research Council (FF0551326, DP130100177, DP170100049, DP210101423) to JLB. SZ acknowledges support from the German Research Foundation (SFB944, P13). We thank Keisuke Inoue for providing pMpGE_En04, pBC-12, pBC-14, pBC-23, and pBC-34, and Edwin Lampugnani for BiFC vectors. We also thank Claudia Gieshoidt for *Marchantia* tissue culture support.

Additional information

Funding

Funder	Grant reference number	Author
Australian Research Council	FF0551326	Tom Dierschke Eduardo Flores-Sandoval John L Bowman
Australian Research Council	DP130100177	Tom Dierschke Eduardo Flores-Sandoval John L Bowman
Australian Research Council	DP170100049	Tom Dierschke Eduardo Flores-Sandoval Madlen I Rast-Somssich John L Bowman
Deutsche Forschungsgemeinschaft	SFB944	Tom Dierschke Felix Althoff Sabine Zachgo
Australian Research Council	DP210101423	Tom Dierschke John L Bowman
Deutsche Forschungsgemeinschaft	P13	Tom Dierschke Felix Althoff Sabine Zachgo

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Tom Dierschke, Conceptualization, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review and editing; Eduardo Flores-Sandoval, Investigation, Methodology; Madlen I Rast-Somssich, Felix Althoff, Investigation; Sabine Zachgo, Methodology; John L Bowman, Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing - original draft, Writing - review and editing

Author ORCIDs

Tom Dierschke  <http://orcid.org/0000-0002-3547-5765>

John L Bowman  <http://orcid.org/0000-0001-7347-3691>

Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.57088.sa1>

Author response <https://doi.org/10.7554/eLife.57088.sa2>

Additional files

Supplementary files

- Supplementary file 1. Primers and guide RNAs used in this study. **Table 1a**. Primer name, gene number, sequence, and the corresponding figure used for sqRT-PCR and RACE. **Table 1b**. Primer name, gene name, and sequence used for *in situ* probe preparation. **Table 1c**. Primer name, gene name, and sequence used for overexpression, inducible overexpression, and transcriptional and translational β -glucuronidase reporter constructs. **Table 1d**. Guide RNA sequences used to generate mutant alleles.
- Transparent reporting form

Data availability

All data generated or analysed during this study are included in the manuscript and supporting files.

References

- Adl SM**, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V, Heiss A, Hoppenrath M, Lara E, Le Gall L, Lynn DH, McManus H, Mitchell EAD, Mozley-Stanridge SE, Parfrey LW, Pawlowski J, et al. 2012. The revised classification of eukaryotes. *The Journal of Eukaryotic Microbiology* **59**: 429–493. DOI: <https://doi.org/10.1111/j.1550-7408.2012.00644.x>, PMID: 23020233
- Althoff F**. 2014. Comparison of the mpef1 α and CAMV35 promoters for application in Marchantia polymorpha overexpression studies. *Transgenic Research* **23**: 234–244. DOI: <https://doi.org/10.1007/s11248-013-9746-z>
- Arimoto A**. 2019. A siphonous macroalgal genome suggests convergent functions of homeobox genes in algae and land plants. *DNA Research* **26**: dsz002. DOI: <https://doi.org/10.1093/dnares/dsz002>
- Arun A**, Coelho SM, Peters AF, Bourdareau S, Pérès L, Scornet D, Strittmatter M, Lipinska AP, Yao H, Godfroy O, Montecinos GJ, Avia K, Macaisne N, Troadec C, Bendahmane A, Cock JM. 2019. Convergent recruitment of tale homeodomain life cycle regulators to direct sporophyte development in land plants and brown algae. *eLife* **8**: e43101. DOI: <https://doi.org/10.7554/eLife.43101>, PMID: 30644818
- Bellaoui M**, Pidkowich MS, Samach A, Kushalappa K, Kohalmi SE, Modrusan Z, Crosby WL, Haughn GW. 2001. The arabidopsis BELL1 and KNOX TALE Homeodomain proteins interact through a domain conserved between plants and animals. *The Plant Cell* **13**: 2455–2470. DOI: <https://doi.org/10.1105/tpc.010161>, PMID: 11701881
- Berger MF**, Badis G, Gehrke AR, Talukder S, Philippakis AA, Peña-Castillo L, Alleyne TM, Mnaimneh S, Botvinnik OB, Chan ET, Khalid F, Zhang W, Newburger D, Jaeger SA, Morris QD, Bulyk ML, Hughes TR. 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133**: 1266–1276. DOI: <https://doi.org/10.1016/j.cell.2008.05.024>, PMID: 18585359
- Bharathan G**, Janssen BJ, Kellogg EA, Sinha N. 1997. Did homeodomain proteins duplicate before the origin of angiosperms, fungi, and metazoa? *PNAS* **94**: 13749–13753. DOI: <https://doi.org/10.1073/pnas.94.25.13749>, PMID: 9391098
- Bower FO**. 1908. Origin of a Land Flora: A Theory Based on the Facts of Alternation. London: MacMillan.
- Bowman JL**. 2016a. A brief history of Marchantia from Greece to Genomics. *Plant & Cell Physiology* **57**: 210–229. DOI: <https://doi.org/10.1093/pcp/pcv044>, PMID: 25766905
- Bowman JL**, Sakakibara K, Furumizu C, Dierschke T. 2016b. Evolution in the cycles of life. *Annual Review of Genetics* **50**: 133–154. DOI: <https://doi.org/10.1146/annurev-genet-120215-035227>, PMID: 27617970
- Bowman JL**, Araki T, Arteaga-Vazquez MA, Berger F, Dolan L, Haseloff J, Ishizaki K, Kyoizuka J, Lin S-S, Nagasaki H, Nakagami H, Nakajima K, Nakamura Y, Ohashi-Ito K, Sawa S, Shimamura M, Solano R, Tsukaya H, Ueda T, Watanabe Y, et al. 2016c. The naming of names: Guidelines for gene nomenclature in Marchantia. *Plant & Cell Physiology* **57**: 257–261. DOI: <https://doi.org/10.1093/pcp/pcv193>, PMID: 26644462
- Bowman JL**, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, Yamaoka S, Nishihama R, Nakamura Y, Berger F, Adam C, Aki SS, Althoff F, Araki T, Arteaga-Vazquez MA, Balasubramanian S, Barry K, Bauer D, Boehm CR, Briginshaw L, et al. 2017. Insights into land plant evolution garnered from the MARCHANTIA polymorpha genome. *Cell* **171**: 287–304. DOI: <https://doi.org/10.1016/j.cell.2017.09.030>, PMID: 28985561
- Bracha-Drori K**, Shichrur K, Katz A, Oliva M, Angelovici R, Yalovsky S, Ohad N. 2004. Detection of protein-protein interactions in plants using bimolecular fluorescence complementation. *The Plant Journal* **40**: 419–427. DOI: <https://doi.org/10.1111/j.1365-313X.2004.02206.x>, PMID: 15469499
- Bürglin TR**. 1997. Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Research* **25**: 4173–4180. DOI: <https://doi.org/10.1093/nar/25.21.4173>, PMID: 9336443
- Conner JA**, Mookkan M, Huo H, Chae K, Ozias-Akins P. 2015. A parthenogenesis gene of apomict origin elicits embryo formation from unfertilized eggs in a sexual plant. *PNAS* **112**: 11205–11210. DOI: <https://doi.org/10.1073/pnas.1505856112>, PMID: 26305939
- Conner JA**, Podio M, Ozias-Akins P. 2017. Haploid embryo production in rice and maize induced by PSASGR-BBML transgenes. *Plant Reproduction* **30**: 41–52. DOI: <https://doi.org/10.1007/s00497-017-0298-x>, PMID: 28238020
- Cooper E**, Delwiche C. 2016. Green algal transcriptomes for phylogenetics and comparative genomics. Figshare. <https://doi.org/10.6084/m9.figshare>

- Crooks GE**, Hon G, Chandonia J-M, Brenner SE. 2004. WEBLOGO: A sequence logo generator. *Genome Research* **14**: 1188–1190. DOI: <https://doi.org/10.1101/gr.849004>, PMID: 15173120
- de Vries J**, Fischer AM, Roettger M, Rommel S, Schluepmann H, Bräutigam A, Carlsbecker A, Gould SB. 2016. Cytokinin-induced promotion of root meristem size in the fern *azolla* supports a shoot-like origin of euphyllophyte roots. *The New Phytologist* **209**: 705–720. DOI: <https://doi.org/10.1111/nph.13630>, PMID: 26358624
- Derelle R**, Lopez P, Guyader HL, Manuel M. 2007. Homeodomain proteins belong to the ancestral molecular toolkit of eukaryotes. *Evolution & Development* **9**: 212–219. DOI: <https://doi.org/10.1111/j.1525-142X.2007.00153.x>
- Dong S**, Zhao C, Zhang S, Wu H, Mu W, Wei T, Li N, Wan T, Liu H, Cui J, Zhu R, Goffinet B, Liu Y. 2019. The Amount of RNA Editing Sites in Liverwort Organellar Genes Is Correlated with GC Content and Nuclear PPR Protein Diversity. *Genome Biology and Evolution* **11**: 3233–3239. DOI: <https://doi.org/10.1093/gbe/evz232>, PMID: 31651960
- Edwards K**, Johnstone C, Thompson C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19**: 1349–1350. DOI: <https://doi.org/10.1093/nar/19.6.1349>, PMID: 2030957
- Ferris PJ**, Goodenough UW. 1987. Transcription of novel genes, including a gene linked to the mating-type locus, induced by *chlamydomonas* fertilization. *Molecular and Cellular Biology* **7**: 2360–2366. DOI: <https://doi.org/10.1128/mcb.7.7.2360-2366.1987>, PMID: 3614194
- Flores-Sandoval E**, Eklund DM, Bowman JL. 2015. A simple auxin transcriptional response system regulates multiple morphogenetic processes in the liverwort *Marchantia polymorpha*. *PLOS Genetics* **11**: e1005207. DOI: <https://doi.org/10.1371/journal.pgen.1005207>, PMID: 26020649
- Flores-Sandoval E**, Dierschke T, Fisher TJ, Bowman JL. 2016. Efficient and inducible use of artificial microRNAs in *Marchantia polymorpha*. *Plant & Cell Physiology* **57**: 281–290. DOI: <https://doi.org/10.1093/pcp/pcv068>, PMID: 25971256
- Frangedakis E**, Saint-Marcoux D, Moody LA, Rabinowitsch E, Langdale JA. 2017. Nonreciprocal complementation of KNOX gene function in land plants. *The New Phytologist* **216**: 591–604. DOI: <https://doi.org/10.1111/nph.14318>, PMID: 27886385
- Furumizu C**, Alvarez JP, Sakakibara K, Bowman JL. 2015. Antagonistic roles for *knox1* and *knox2* genes in patterning the land plant body plan following an ancient gene duplication. *PLOS Genetics* **11**: e1004980. DOI: <https://doi.org/10.1371/journal.pgen.1004980>, PMID: 25671434
- Gillissen B**, Bergemann J, Sandmann C, Schroeder B, Bölker M, Kahmann R. 1992. A two-component regulatory system for self/non-self recognition in *Ustilago maydis*. *Cell* **68**: 647–657. DOI: [https://doi.org/10.1016/0092-8674\(92\)90141-x](https://doi.org/10.1016/0092-8674(92)90141-x), PMID: 1739973
- Goutte C**, Johnson AD. 1988. $\alpha 1$ protein alters the DNA binding specificity of $\alpha 2$ repressor. *Cell* **52**: 875–882. DOI: [https://doi.org/10.1016/0092-8674\(88\)90429-1](https://doi.org/10.1016/0092-8674(88)90429-1), PMID: 3127056
- Hackbusch J**, Richter K, Müller J, Salamini F, Uhrig JF. 2005. A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins. *PNAS* **102**: 4908–4912. DOI: <https://doi.org/10.1073/pnas.0501181102>, PMID: 15781858
- Hay A**, Tsiantis M. 2010. KNOX genes: Versatile regulators of plant development and diversity. *Development* **137**: 3153–3165. DOI: <https://doi.org/10.1242/dev.030049>, PMID: 20823061
- Hedgethorpe K**, Eustermann S, Yang J-C, Ogden TEH, Neuhaus D, Bloomfield G. 2017. Homeodomain-like DNA binding proteins control the haploid-to-diploid transition in *Dictyostelium*. *Science Advances* **3**: e1602937. DOI: <https://doi.org/10.1126/sciadv.1602937>, PMID: 28879231
- Hellens RP**, Edwards EA, Leyland NR, Bean S, Mullineaux PM. 2000. Pgreen: A versatile and flexible binary Ti vector for agrobacterium-mediated plant transformation. *Plant Molecular Biology* **42**: 819–832. DOI: <https://doi.org/10.1023/a:1006496308160>, PMID: 10890530
- Herskowitz I**. 1989. A regulatory hierarchy for cell specialization in yeast. *Nature* **342**: 749–757. DOI: <https://doi.org/10.1038/342749a0>, PMID: 2513489
- Higo A**, Niwa M, Yamato KT, Yamada L, Sawada H, Sakamoto T, Kurata T, Shirakawa M, Endo M, Shigenobu S, Yamaguchi K, Ishizaki K, Nishihama R, Kohchi T, Araki T. 2016. Transcriptional framework of male gametogenesis in the liverwort *Marchantia polymorpha* l. *Plant & Cell Physiology* **57**: 325–338. DOI: <https://doi.org/10.1093/pcp/pcw005>, PMID: 26858289
- Hisanaga T**. 2019. A cis-acting bidirectional transcription switch controls sexual dimorphism in the liverwort. *The EMBO Journal* **10**: e100240. DOI: <https://doi.org/10.15252/embj.2018100240>
- Hisanaga T**. 2021. Deep Evolutionary Origin of Gamete-Directed Zygote Activation by KNOX/BELL Transcription Factors in Green Plants. *bioRxiv*. DOI: <https://doi.org/10.1101/2020.04.08.031930>
- Hofmeister WFB**. 1862. On the Germination, Development, and Fructification of the Higher Cryptogamia, and on the Fructification of the Coniferae. Ray Society. DOI: <https://doi.org/10.5962/bhl.title.60930>
- Hori K**, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, Moriyama T, Ikeuchi M, Watanabe M, Wada H, Kobayashi K, Saito M, Masuda T, Sasaki-Sekimoto Y, Mashiguchi K, Awai K, et al. 2014. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nature Communications* **5**: 3978. DOI: <https://doi.org/10.1038/ncomms4978>, PMID: 24865297
- Horst NA**, Katz A, Pereman I, Decker EL, Ohad N, Reski R. 2016. A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction. *Nature Plants* **2**: 15209. DOI: <https://doi.org/10.1038/nplants.2015.209>, PMID: 27250874

- Huelsenbeck JP**, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755. DOI: <https://doi.org/10.1093/bioinformatics/17.8.754>, PMID: 11524383
- Huelsenbeck JP**, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**: 2310–2314. DOI: <https://doi.org/10.1126/science.1065889>, PMID: 11743192
- Hull CM**, Boily M-J, Heitman J. 2005. Sex-specific homeodomain proteins Sxi1alpha and Sxi2a coordinately regulate sexual development in *Cryptococcus neoformans*. *Eukaryotic Cell* **4**: 526–535. DOI: <https://doi.org/10.1128/EC.4.3.526-535.2005>, PMID: 15755915
- Inoue K**, Nishihama R, Araki T, Kohchi T. 2019. Reproductive induction is a far-red high irradiance response that is mediated by phytochrome and phytochrome interacting factor in *Marchantia polymorpha*. *Plant & Cell Physiology* **60**: 1136–1145. DOI: <https://doi.org/10.1093/pcp/pcz029>, PMID: 30816950
- Ishizaki K**, Chiyoda S, Yamato KT, Kohchi T. 2008. Agrobacterium-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant & Cell Physiology* **49**: 1084–1091. DOI: <https://doi.org/10.1093/pcp/pcn085>, PMID: 18535011
- Ishizaki K**, Nishihama R, Ueda M, Inoue K, Ishida S, Nishimura Y, Shikanai T, Kohchi T. 2015. Development of Gateway binary vector series with four different selection markers for the liverwort *Marchantia polymorpha*. *PLOS ONE* **10**: e0138876. DOI: <https://doi.org/10.1371/journal.pone.0138876>, PMID: 26406247
- Joo S**, Wang MH, Lui G, Lee J, Barnas A, Kim E, Sudek S, Worden AZ, Lee J-H. 2018. Common ancestry of heterodimerizing tale homeobox transcription factors across metazoa and archaeplastida. *BMC Biology* **16**: 136. DOI: <https://doi.org/10.1186/s12915-018-0605-5>, PMID: 30396330
- Ju C**, Van de Poel B, Cooper ED, Thierer JH, Gibbons TR, Delwiche CF, Chang C. 2015. Conservation of ethylene as a plant hormone over 450 million years of evolution. *Nature Plants* **1**: 14004. DOI: <https://doi.org/10.1038/nplants.2014.4>, PMID: 27246051
- Kamada T**. 2002. Molecular genetics of sexual development in the mushroom *Coprinus cinereus*. *BioEssays* **24**: 449–459. DOI: <https://doi.org/10.1002/bies.10083>, PMID: 12001268
- Keeling PJ**. 2014. The marine microbial eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLOS Biology* **6**: e1001889. DOI: <https://doi.org/10.1371/journal.pbio.1001889>
- Khanday I**, Skinner D, Yang B, Mercier R, Sundaresan V. 2019. A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature* **565**: 91–95. DOI: <https://doi.org/10.1038/s41586-018-0785-8>, PMID: 30542157
- Kienitz-gerloff F**. 1874. vergleichende untersuchungen über die entwicklungsgeschichte des lebermoos - sporogoniums. *Botanische Zeitung* **32**: 161–172.
- Kimura S**, Koenig D, Kang J, Yoong FY, Sinha N. 2008. Natural variation in leaf morphology results from mutation of a novel KNOX gene. *Current Biology* **18**: 672–677. DOI: <https://doi.org/10.1016/j.cub.2008.04.008>, PMID: 18424140
- Knapp E**. 1935. Untersuchungen über die wirkung von Röntgenstrahlen an Dem Lebermoos sphaerocarpus, Mit Hilfe der tetraden-analyse. I. *Zeitschrift Für Induktive Abstammungs- Und Vererbungslehre* **70**: 309–349. DOI: <https://doi.org/10.1007/BF00328659>
- Kues U**, Richardson WV, Tymon AM, Mutasa ES, Gottgens B, Gaubatz S, Gregoriades A, Casselton LA. 1992. The combination of dissimilar alleles of the A alpha and A beta gene complexes, whose proteins contain homeo domain motifs, determines sexual development in the mushroom *Coprinus cinereus*. *Genes & Development* **6**: 568–577. DOI: <https://doi.org/10.1101/gad.6.4.568>
- Lampugnani ER**, Ho YY, Moller IE, Koh P-L, Golz JF, Bacic A, Newbigin E. 2016. A Glycosyltransferase from *Nicotiana glauca* Pollen Mediates Synthesis of a Linear (1,5)- α -L-Arabinan When Expressed in *Arabidopsis*. *Plant Physiology* **170**: 1962–1974. DOI: <https://doi.org/10.1104/pp.15.02005>, PMID: 26850276
- Lee J-H**, Lin H, Joo S, Goodenough U. 2008. Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. *Cell* **133**: 829–840. DOI: <https://doi.org/10.1016/j.cell.2008.04.028>, PMID: 18510927
- Lewis EB**. 1978. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**: 565–570. DOI: <https://doi.org/10.1038/276565a0>, PMID: 103000
- Lorbeer G**. 1936. DIE Umwandlung EINES HAPLOIDEN, GENOTYPISCH WEIBLICHEN Gametophyten von *Sphaerocarpus donnellii* in Einen Männlichen mit HILFE von röntgenstrahlen. *Planta* **25**: 70–83. DOI: <https://doi.org/10.1007/BF01909306>
- Magnani E**, Hake S. 2008. KNOX lost the ox: The *Arabidopsis* KNATM gene defines a novel class of KNOX transcriptional regulators missing the homeodomain. *The Plant Cell* **20**: 875–887. DOI: <https://doi.org/10.1105/tpc.108.058495>, PMID: 18398054
- Marchant J**. 1713. Nouvelle Découverte Des Fleurs et des Graines D'Une Plante Rangée Par Les Botanistes Fous Le Genre Du Lichen. Histoire de l'Académie Royale des Sciences, Avec les Mémoires de Mathématique et de Physique, pour La Meme Cette Académie. *Année MDCCXIII*: 229–234.
- Merabet S**, Mann RS. 2016. To be specific or not: The critical relationship between Hox and TALE proteins. *Trends in Genetics* **32**: 334–347. DOI: <https://doi.org/10.1016/j.tig.2016.03.004>, PMID: 27066866
- Meyberg R**, Perroud PF, Haas FB, Schneider L, Heimerl T, Renzaglia KS, Rensing SA. 2020. Characterisation of evolutionarily conserved key players affecting eukaryotic flagellar motility and fertility using a moss model. *The New Phytologist* **227**: 440–454. DOI: <https://doi.org/10.1111/nph.16486>, PMID: 32064607

- Mikami K**, Li C, Irie R, Hama Y. 2019. A unique life cycle transition in the red seaweed *pyropia yezoensis* depends on apospory. *Communications Biology* **2**: 299. DOI: <https://doi.org/10.1038/s42003-019-0549-5>, PMID: 31396579
- Mirbel CF**. 1835. Recherches anatomiques et physiologiques sur le *Marchantia polymorpha*. (Anatomical and physiological researches on *Marchantia polymorpha*. *Mém. Acad. Roy. Sc. Inst. France* **13**: 337–436.
- Montgomery SA**, Tanizawa Y, Galik B, Wang N, Ito T, Mochizuki T, Akimcheva S, Bowman JL, Cognat V, Maréchal-Drouard L, Ekker H, Hong SF, Kohchi T, Lin SS, Liu LYD, Nakamura Y, Valeeva LR, Shakirov EV, Shippen DE, Wei WL, et al. 2020. Chromatin organization in early land plants reveals an ancestral association between h3k27me3, transposons, and constitutive heterochromatin. *Current Biology* **30**: 573–588. DOI: <https://doi.org/10.1016/j.cub.2019.12.015>, PMID: 32004456
- Nishimura Y**, Shikanai T, Nakamura S, Kawai-Yamada M, Uchimiya H. 2012. Gsp1 triggers the sexual developmental program including inheritance of chloroplast dna and mitochondrial dna in *Chlamydomonas Reinhardtii*. *The Plant Cell* **24**: 2401–2414. DOI: <https://doi.org/10.1105/tpc.112.097865>, PMID: 22715041
- Noyes MB**, Christensen RG, Wakabayashi A, Stormo GD, Brodsky MH, Wolfe SA. 2008. Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* **133**: 1277–1289. DOI: <https://doi.org/10.1016/j.cell.2008.05.023>, PMID: 18585360
- Ortiz-Ramírez C**, Michard E, Simon AA, Damineli DSC, Hernández-Coronado M, Becker JD, Feijó JA. 2017. Glutamate receptor-like channels are essential for chemotaxis and reproduction in mosses. *Nature* **549**: 91–95. DOI: <https://doi.org/10.1038/nature23478>, PMID: 28737761
- Pagnussat GC**, Yu HJ, Sundaresan V. 2007. Cell-fate switch of synergid to egg cell in arabidopsis eostre mutant embryo sacs arises from misexpression of the bel1-like homeodomain gene blh1. *The Plant Cell* **19**: 3578–3592. DOI: <https://doi.org/10.1105/tpc.107.054890>, PMID: 18055603
- Pearson JC**, Lemons D, McGinnis W. 2005. Modulating HOX gene functions during animal body patterning. *Nature Reviews. Genetics* **6**: 893–904. DOI: <https://doi.org/10.1038/nrg1726>, PMID: 16341070
- Radhakrishnan GV**, Keller J, Rich MK, Vernié T, Mbadinga Mbadinga DL, Vigneron N, Cottret L, Clemente HS, Libourel C, Cheema J, Linde AM, Eklund DM, Cheng S, Wong GKS, Lagercrantz U, Li FW, Oldroyd GED, Delaux PM. 2020. An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. *Nature Plants* **6**: 280–289. DOI: <https://doi.org/10.1038/s41477-020-0613-7>, PMID: 32123350
- Renzaglia KS**, Villarreal Aguilar JC, Garbary DJ. 2018. Morphology supports the setaphyte hypothesis: Mosses plus liverworts form a natural group. *Bryophyte Diversity and Evolution* **40**: 11. DOI: <https://doi.org/10.11646/bde.40.2.1>
- Retamales HA**, Scharaschkin T. 2014. A staining protocol for identifying secondary compounds in myrtaceae. *Applications in Plant Sciences* **2**: apps.1400063. DOI: <https://doi.org/10.3732/apps.1400063>, PMID: 25309840
- Sakakibara K**, Nishiyama T, Deguchi H, Hasebe M. 2008. Class 1 KNOX genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development. *Evolution & Development* **10**: 555–566. DOI: <https://doi.org/10.1111/j.1525-142X.2008.00271.x>, PMID: 18803774
- Sakakibara K**, Ando S, Yip HK, Tamada Y, Hiwatashi Y, Murata T, Deguchi H, Hasebe M, Bowman JL. 2013. KNOX2 genes regulate the haploid-to-diploid morphological transition in land plants. *Science* **339**: 1067–1070. DOI: <https://doi.org/10.1126/science.1230082>, PMID: 23449590
- Sakakibara K**. 2016. Technological Innovations Give Rise to a New Era of Plant Evolutionary Developmental Biology. Elsevier.
- Sano R**, Juarez CM, Hass B, Sakakibara K, Ito M, Banks JA, Hasebe M. 2005. KNOX homeobox genes potentially have similar function in both diploid unicellular and multicellular meristems, but not in haploid meristems. *Evolution and Development* **7**: 69–78. DOI: <https://doi.org/10.1111/j.1525-142X.2005.05008.x>
- Schulz KN**, Harrison MM. 2019. Mechanisms regulating zygotic genome activation. *Nature Reviews. Genetics* **20**: 221–234. DOI: <https://doi.org/10.1038/s41576-018-0087-x>, PMID: 30573849
- Singer SD**, Ashton NW. 2007. Revelation of ancestral roles of KNOX genes by a functional analysis of *Physcomitrella* homologues. *Plant Cell Reports* **26**: 2039–2054. DOI: <https://doi.org/10.1007/s00299-007-0409-5>, PMID: 17724598
- Smith HMS**, Boschke I, Hake S. 2002. Selective interaction of plant homeodomain proteins mediates high dna-binding affinity. *PNAS* **99**: 9579–9584. DOI: <https://doi.org/10.1073/pnas.092271599>, PMID: 12093897
- Spit A**, Hyland RH, Mellor EJ, Casselton LA. 1998. A role for heterodimerization in nuclear localization of a homeodomain protein. *PNAS* **95**: 6228–6233. DOI: <https://doi.org/10.1073/pnas.95.11.6228>, PMID: 9600947
- Stephenson AG**. 1981. Flower and fruit abortion: Proximate causes and ultimate functions. *Annual Review of Ecology and Systematics* **12**: 253–279. DOI: <https://doi.org/10.1146/annurev.es.12.110181.001345>
- Sugano SS**, Shirakawa M, Takagi J, Matsuda Y, Shimada T, Hara-Nishimura I, Kohchi T. 2014. Crispr/cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L. *Plant & Cell Physiology* **55**: 475–481. DOI: <https://doi.org/10.1093/pcp/pcu014>, PMID: 24443494
- Sugano SS**, Nishihama R, Shirakawa M, Takagi J, Matsuda Y, Ishida S, Shimada T, Hara-Nishimura I, Osakabe K, Kohchi T. 2018. Efficient crispr/cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLOS ONE* **13**: e0205117. DOI: <https://doi.org/10.1371/journal.pone.0205117>, PMID: 30379827
- Swanson WJ**, Vacquier VD. 2002. The rapid evolution of reproductive proteins. *Nature Reviews. Genetics* **3**: 137–144. DOI: <https://doi.org/10.1038/nrg733>, PMID: 11836507
- Thuret G**. 1851. Recherches Sur les zoospores des Algues et les anthéridies des Cryptogames. *Annales de Sciences Naturelles Botanique* **3**: 5–39.

- Unger F.** 1837. Weitere Beobachtungen über die Samenthiere der Pflanzen. *Nova Acta Physico Medica A C L* **18**: 786–796. DOI: <https://doi.org/10.1002/andp.18983000104>
- Urban M.** 1996. Identification of the pheromone response element in *Ustilago maydis*. *Molecular & General Genetics* **251**: 31–37. DOI: <https://doi.org/10.1007/BF02174341>
- Vanneste K, Sterck L, Myburg AA, Van de Peer Y, Mizrahi E.** 2015. Horsetails are ancient polyploids: Evidence from *equisetum giganteum*. *The Plant Cell* **27**: 1567–1578. DOI: <https://doi.org/10.1105/tpc.15.00157>, PMID: [26002871](https://pubmed.ncbi.nlm.nih.gov/26002871/)
- Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, Ruhfel BR, Wafula E, Der JP, Graham SW, Mathews S, Melkonian M, Soltis DE, Soltis PS, Miles NW, Rothfels CJ, et al.** 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *PNAS* **111**: E4859–E4868. DOI: <https://doi.org/10.1073/pnas.1323926111>, PMID: [25355905](https://pubmed.ncbi.nlm.nih.gov/25355905/)
- Zachgo S.** 2002. In Situ Hybridization. Oxford University Press.
- Zhao H, Lu M, Singh R, Snell WJ.** 2001. Ectopic expression of a *Chlamydomonas mt+*-specific homeodomain protein in *mt-* gametes initiates zygote development without gamete fusion. *Genes & Development* **15**: 2767–2777. DOI: <https://doi.org/10.1101/gad.919501>, PMID: [11641281](https://pubmed.ncbi.nlm.nih.gov/11641281/)