Identification of the sex-determining factor in the liverwort *Marchantia polymorpha* reveals unique evolution of sex chromosomes in a haploid system

**Highlights**

- *Feminizer (BPCU)* linked to the liverwort female sex chromosome was identified

- *BPCU* regulates the autosomal sex-determining locus, *FGMYB/SUF*

- *BPCU* additionally shares a function for reproductive induction with its gametolog

- The liverwort sex chromosomes are the oldest among any other known systems

**Authors**

Miyuki Iwasaki, Tomoaki Kajiwara, Yukiko Yasui, ..., Katsuyuki T. Yamato, John L. Bowman, Takayuki Kohchi

**Correspondence**

john.bowman@monash.edu (J.L.B.), tkohchi@lif.kyoto-u.ac.jp (T.K.)

**In brief**

Iwasaki et al. identify a female sex determinant, *Feminizer*, on the female sex chromosome of a haploid plant, *Marchantia polymorpha*. Unexpectedly, *Feminizer* also plays a role in reproductive induction, which is shared by its male gametolog. The liverwort sex chromosomes diverged around 430 mya, older than any known sex chromosome systems.

Iwasaki et al., 2021, Current Biology 31, 5522–5532

December 20, 2021 © 2021 The Authors. Published by Elsevier Inc.

https://doi.org/10.1016/j.cub.2021.10.023
Identification of the sex-determining factor in the liverwort *Marchantia polymorpha* reveals unique evolution of sex chromosomes in a haploid system

Miyuki Iwasaki,1 Tomoaki Kajiwara,1 Yukiko Yasui,1 Yoshihiro Yoshitake,1 Motoki Miyazaki,1 Shogo Kawamura,1 Noriyuki Suetsugu,1 Ryuich Nishihama,1,2 Shohei Yamaoka,1 Dierk Wanke,3 Kenji Hashimoto,2 Kazuyuki Kuchitsu,2 Sean A. Montgomery,4 Shilpi Singh,5 Yasuhiro Tanizawa,6 Masaru Yagura,6 Takako Mochizuki,6 Mika Sakamoto,6 Yasukazu Nakamura,6 Chang Liu,7 Frédéric Berger,1 Katsuyuki T. Yamato,8 John L. Bowman,5,* and Takayuki Kohchi1,9,*

1Graduate School of Biostudies, Kyoto University, Kyoto 606-8502, Japan
2Faculty of Science and Technology, Tokyo University of Science, Chiba 278-8510, Japan
3Department Biologie I, Ludwig-Maximilians-University (LMU), München 80638, Germany
4Gregor Mendel Institute (GMI), Austrian Academy of Sciences, Vienna Biocenter (VBC), 1030 Vienna, Austria
5School of Biological Sciences, Monash University, Melbourne, VIC 3800, Australia
6National Institute of Genetics, Research Organization of Information and Systems, Mishima, Shizuoka 411-8540, Japan
7Institute of Biology, University of Hohenheim, Stuttgart 70599, Germany
8Faculty of Biology-Oriented Science and Technology (BOST), Kindai University, Kinokawa, Wakayama 649-6493, Japan
9Lead contact

*Correspondence: john.bowman@monash.edu (J.L.B.), tkohchi@lif.kyoto-u.ac.jp (T.K.)
https://doi.org/10.1016/j.cub.2021.10.023

SUMMARY

Sex determination is a central process for sexual reproduction and is often regulated by a sex determinant encoded on a sex chromosome. Rules that govern the evolution of sex chromosomes via specialization and degeneration following the evolution of a sex determinant have been well studied in diploid organisms. However, distinct predictions apply to sex chromosomes in organisms where sex is determined in the haploid phase of the life cycle: both sex chromosomes, female U and male V, are expected to maintain their gene functions, even though both are non-recombining. This is in contrast to the X-Y (or Z-W) asymmetry and Y (W) chromosome degeneration in XY (ZW) systems of diploids. Here, we provide evidence that sex chromosomes diverged early during the evolution of haploid liverworts and identify the sex determinant on the *Marchantia polymorpha* U chromosome. This gene, *Feminizer*, encodes a member of the plant-specific BASIC PENTACYSTEINE transcription factor family. It triggers female differentiation via regulation of the autosomal sex-determining locus of *FEMALE GAMETOPHYTE MYB* and *SUPPRESSOR OF FEMINIZATION*. Phylogenetic analyses of *Feminizer* and other sex chromosome genes indicate dimorphic sex chromosomes had already been established 430 mya in the ancestral liverwort. *Feminizer* also plays a role in reproductive induction that is shared with its gametolog on the V chromosome, suggesting an ancestral function, distinct from sex determination, was retained by the gametologs. This implies ancestral functions can be preserved after the acquisition of a sex determination mechanism during the evolution of a dominant haploid sex chromosome system.

INTRODUCTION

Sex-specific gene expression leading to sexual differentiation is generally triggered by a sex-determining gene, often carried by a sex chromosome. Since the early 20th century, different sex determination systems by sex chromosomes, including XX/XY, ZW/ZZ, XX/XX, and ZO/ZZ, have been described in animals.1 Similarly, some flowering plants possess XX/XY (or ZW/ZZ) systems directing sex-specific differentiation, leading to flower development during the diploid life phase.2–4 Sex chromosomes in some flowering plants, including persimmon2 and poplar,5 have a single master gene that directs sex determination, as in some animal XX/XY systems. In other flowering plants, such as kiwifruit6 and asparagus,7 however, sex determination involves allelic differences in two linked genes on the heterogametic sex chromosome, which act to promote male development and suppress female development.8 It is also notable that sex chromosomes in a number of dioecious angiosperm species are homomorphic or show minor structural divergence.9 In diploid systems, suppression of recombination between the sex chromosomes in heterozygotes restricts structural and functional differentiation to the sex chromosome specific to either males (Y) or females (W). The other sex chromosome (X or Z) remains similar to the ancestral state, carrying functional copies of ancestrally present genes, whose presence allows genetic degeneration of the Y- or W-linked regions, sometimes including loss of most ancestral genes. Dosage compensation may evolve to compensate for males’ low X-linked gene expression in XY
systems (or females’ low Z expression in ZW systems). Understanding of sex determination mechanisms and sex chromosome evolution has largely come from studying diploid organisms.

In contrast to animals and flowering plants, sexual differentiation takes place in the dominant haploid life phase of bryophytes (e.g., liverworts and mosses) and most algae, with species having male and female individuals termed dioecious. In these groups, males and females are defined by their gamete size, with females producing large immotile egg cells and males producing small motile sperm. Sex chromosomes in haploid systems show fundamentally distinct genetic and evolutionary behavior from those in diploid systems in which one sex is hemizygous (XY or ZW) and one homozygous (XX or ZZ). In haploid systems, each individual carries only a single sex chromosome. The male and female sex chromosomes are termed U and V, respectively, to emphasize the differences from X and Y (or Z and W) chromosomes. Fertilization results in diploid UV sporophytes, a brief non-sexually differentiated phase of the life cycle, another difference from diploid systems. Due to their similar inheritance, the U and V chromosomes are predicted to evolve similarly. Specifically, both chromosomes are predicted to undergo little or no degeneration, and any degeneration is predicted to affect both the U and V chromosomes similarly, in contrast to the diploid systems outlined above.

Many liverwort species are dioecious and often have heteromorphic sex chromosomes, and dioecy is the likely ancestral condition for liverworts. In Marchantia polymorpha, plants possess either a U or a V sex chromosome and develop into females or males, respectively. Genetic analyses of phylogenetically diverse liverworts indicate that the U chromosome carries a feminizing locus and that the V chromosome carries multiple loci required for sperm motility. Further, haploid plants spontaneously containing both U and V chromosomes develop as females, supporting the presence of a dominant U-linked feminizing locus, in whose absence plants develop as males. Genomic analyses of M. polymorpha have shown that both U and V chromosomes have lower gene densities, and higher repetitive content, than the autosomes, similar to the symmetrical changes in other non-recombining U/V systems, in the mosses Ceratodon purpureus and Syntrichia caninervis and the brown alga Ectocarpus. However, the sex-determining genes have yet to be identified in liverworts or any other complex multicellular haploid organism, and only recently has it become possible to study the evolution of sex chromosomes in haploid systems and test their predicted differences from diploid systems.

Here, we identify a U-linked transcription factor gene, BASIC PENTACYSTEINE ON THE U CHROMOSOME (BPCU), as the master sex determinant, Feminizer, in M. polymorpha. BPCU has a homolog (gametolog) on the V chromosome, BASIC PENTACYSTEINE ON THE V CHROMOSOME (BPCV), and both share a conserved function in reproductive induction, suggesting retention of an ancestral function despite the BPCU gametolog evolving an additional function in female sex determination. Phylogenetic analyses of BPCU and BPCV, and their orthologs in other liverworts suggest that these gametologs, and thus the U and V chromosomes, had been already established 430 mya in the ancestral liverwort.

RESULTS

Chromosome-level assembly of the U chromosome

As a basis for the search for Feminizer, we constructed a chromosome-level assembly of the U chromosome using Pac-Bio long-read sequencing and Hi-C technologies, as was done previously for the V chromosome. The U chromosome carries rDNA repeats at each of its ends and, between these, includes a 4.5-Mb region containing genes. Previous analyses identified 74 and 154 genes on the U and V chromosomes, respectively, but the chromosome-level assembly of the U chromosome and manually refined annotation of both chromosomes (excluding transposable element genes and merging some split genes) result in smaller numbers of 47 and 93, respectively (Data S1A). As was noted previously, most U chromosome genes that do not have a V gametolog appear to be fragmentary or have highly similar autosomal paralogs, suggesting that they may be recent immigrants to the U chromosome (Data S1A). The eight M. polymorpha autosomes carry much larger numbers of genes, between 1,900 and 3,000.

BPCU on the U chromosome is necessary and sufficient for female sex determination

We confirmed the aforementioned predictions of a dominant U-linked sex-determining Feminizer locus in M. polymorpha, because plants carrying both U and V sex chromosomes that arose spontaneously via meiotic non-disjunction (wild type [WT] [LV], the sex chromosome background is given in brackets for clarity) developed female sex organs (Figures 1A–1C). To identify the U-linked Feminizer gene, we screened, via genome editing of U-chromosome candidate genes, for mutations that impact sex organ morphology (Data S1A). Two independently obtained mutants of BPCU formed male sex organs (Figures 1D–1F), suggesting that BPCU is necessary for development of female sex organs. Surprisingly, both of these bpcU mutants possessed not only a U chromosome harboring the mutated bpcU gene but also a V chromosome (Figure 1G). This observation and analysis by flow cytometry (Figure S1) suggest that the two bpcU mutants were derived from spores that had inherited both sex chromosomes, presumably via non-disjunction. To further test BPCU’s dominant feminizing activity, we introduced a genomic fragment of BPCU spanning 2 kb upstream to 3 kb downstream of the BPCU transcribed region into the genome of WT male plants that carry only the V chromosome. The genetically male plants transformed with BPCU (gBPCU [V]) displayed a male-to-female sex conversion phenotype, although no egg cells developed within their sex organs (Figure 1H), a phenotype previously reported for an autosomal sex conversion mutant. The loss of egg cell differentiation in these genotypes is likely due to the lack of U-linked egg-cell differentiation gene(s). We concluded that BPCU is necessary and sufficient to induce the development program of female sex organs.

BPCU acts as Feminizer by regulating the autosomal sex-determining locus, FGMYB/SUF

We next analyzed the expression profiles of sex-specific genes in the sex-converted plants and detected gene
expression patterns reflecting the morphological sex phenotype (Figure 2A). Critically, expression of the autosomal locus FEMALE GAMETOPHYTE MYB (FGMYB), which promotes female sex differentiation,28 was correlated with the morphological sex observed in different genotypes. In males, the expression of FGMYB is suppressed by expression of an antisense long non-coding RNA called SUPPRESSOR OF FEMINIZATION (SUF), which is transcribed from a promoter located 3′ of the FGMYB locus.29 Consistent with the sex conversion phenotype, the mutants (bpcU-101 and bpcU-102) had decreased FGMYB and increased SUF expression compared with WT females (Figure 2B, left). Furthermore, the BPCU-carrying genetically male plants (gBPCU) [V] had decreased SUF and increased FGMYB expression compared with WT males (Figure 2B, right). BPCU belongs to an alanine-zipper-containing group of the BARLEY B RECOMBINANT/BASIC PENTACYSTEINE (BBR/BPC) proteins that bind to GAGA elements31,32 and repress gene expression via epigenetic regulation.33 We demonstrated that, in vitro, BPCU binds to GAGA motifs enriched in the promoter of SUF (Figure S2). Thus, we hypothesized that, in females, BPCU directly suppresses SUF expression, consequently relieving FGMYB from suppression by SUF. In support of this hypothesis, fgmyb mutations in the gBPCU [V] background suppressed the feminization induced by gBPCU (Figures 2C and S3). We also observed differences in chromatin status with the repressive H3K27me3 peaks expanded into the gene region of SUF in WT [U], but not in WT [V] and bpcU [U] (Figure 2D), which is consistent with a model that BPCU directly represses SUF through chromatin modification. These observations indicate that BPCU acts as the U-linked Feminizer and that it directly regulates the autosomal sex-determining locus complex, FGMYB/SUF, in M. polymorpha.

BPCU and its gametolog, BPCV, are required for induction of sexual reproduction, but only BPCU confers the feminization

In contrast with other sex-determining genes identified in diploid organisms and also in the haploid alga Volvox carteri,33 the Feminizer, BPCU, has a paired homolog, BPCV, on the male sex chromosome (Figure S4A). Such gene pairs are called “gametologs” and originate from a common ancestral autosomal gene. Gametologs generally conserve essential functions that are maintained during sex chromosome evolution. Indeed, the highly conserved C-terminal domain of BPCV, which is known as a DNA-binding domain of BPCs, also bound the same target sequences as BPCU in vitro (Figure S2A). BPCU and BPCV are both expressed primarily in the female and male reproductive organs, respectively, although they are expressed in all tissues examined (Figure S4B). Because the bpcU mutants (bpcU-101 and bpcU-102) obtained by the initial screening were aneuploid with the V chromosomes, we generated bpcU as well as bpcV mutants in a haploid background (Figures 3A and S5). Both bpcU and bpcV haploid mutants failed to induce sex organ formation under growth conditions that induce reproductive development in WT (Figures 3B and 3E). The introduction of a genomic fragment of BPCU or BPCV complemented the reproductive induction phenotype of bpcU (Figures 3C-3E). However, bpcU mutants complemented with BPCV developed male sex organs instead of female sex organs (Figures 3C and 3D). These observations indicate that BPCU and BPCV share a function required to induce sex organ development.

The sex chromosomes in liverworts were established around 430 mya, which is the oldest among the known sex chromosomes

We first examined the antiquity of the BPCU and BPCV gametologs within liverworts and noted that both gametologs are
present in species representing the three extant lineages of liverworts, indicating that the divergence of these gametologs began in the ancestral liverwort (Figure 4). Further, the ancestral land plant possessed a single BPC transcription factor gene, a condition likely retained in common ancestors of each of the bryophyte lineages. The identification of BPCU as *Feminizer* in *M. polymorpha* sheds light on liverwort U/V sex chromosome evolution. Sex chromosomes have their evolutionary origin when an ancestral gene on an autosomal pair mutates or when a gene duplicates onto one member of this chromosome pair and functions as a sex determiner. During the evolution of a U/V sex chromosome system, the sex-determining gene might have arisen in a non-recombining region or a non-recombining region might have evolved subsequently. Thereafter, successive chromosomal rearrangements could create evolutionary strata of U-V sequence divergence and potential genetic degeneration, which, as outlined above, should affect both members of the pair. However, essential genes required in the haploid generation could be retained on both the U and V and over time diverge into pairs of gametologs that differ in sequence but retain a conserved function. In contrast, genes necessary in only one sex should be retained.

Figure 2. **BPCU acts as *Feminizer* by regulating the autosomal sex-determining locus, FGMYB/SUF**

(A) RT-PCR analysis of sex-specific gene expression in *bpcU* [UV] and g*BPCU* [V]. Total RNAs from WT [U] (Tak-2), WT [V] (Tak-1), UV plant, *bpcU* (*bpcU*-101 [UV] and *bpcU*-102 [UV]), and g*BPCU* [V] (lines no. 3 and no. 4) were used for detecting expression of the genes indicated above. Mp3g14390 and Mp6g01560 are genes expressed specifically in males and Mp5g13880 and Mp7g02820 in females. Mp3g23400 (*EF1*) is ubiquitously expressed gene.

(B) Expression of FGMYB and SUF. qRT-PCR was performed for RNAs from sex organs. Experiments shown in the left and right panels were performed separately for the loss-of-function lines (*bpcU*-101 [UV]; *bpcU*-102 [UV]) and the gain-of-function lines (g*BPCU* [V] no. 3 and no. 4), respectively, with WT [U] (Tak-2) and WT [V] (Tak-1) as controls. Bars represent mean ± SE. Symbols above the bars indicate grouping by p < 0.05 in a Tukey-Kramer test (n = 3).

(C) Genetic suppression of BPCU-dependent feminization by *fgmyb* mutation. Sex organs 1 month after reproductive induction are shown for two independently obtained *fgmyb* mutants. Left and right panels: receptacles and gamete-containing organs are shown, respectively. Scale bars, 1 mm (left panels); 50 μm (right panels).

(D) H3K27me3 status at the FGMYB/SUF locus in the vegetative stage in WT [U], WT [V], and *bpcU*-1 [U]. See Figure 3 for the genotype and the phenotype of *bpcU*-1 [U]. IGV browser screenshot is shown. H3K27me3 and RNA-seq tracks show bigwig files normalized to 1 × genomic coverage (reads per genome coverage [RPGCs]) and bins per million mapped reads (BPMs), respectively (window size = 10 bp). Peak calling was done using SICER2 with the default parameters (window size = 200 bp; false discovery rate [FDR] = 0.01). H3 was used as a control for peak calling. See also Figures S2 and S3.
on the respective sex chromosome and could be lost from the other.\textsuperscript{17}

Because the sex-determining gene or genes should be located in the oldest stratum,\textsuperscript{10} the \textit{M. polymorpha BPCU} probably identifies the oldest non-recombining region of the \textit{U} chromosome. In this plant, divergence of \textit{U-V} genes could not be estimated using synonymous site substitution values (\textit{pS}), because these values are at or near saturation for almost half of the 19 gametolog pairs (Data S1 \textit{B}). Instead, we therefore used phylogenetic analyses based on all alignable coding sequences of each of the gametologs (Figure 4; Data S2). The timing of the incorporation of each pair of \textit{M. polymorpha} gametologs into non-recombining regions of the \textit{M. polymorpha} sex chromosomes can be inferred by the tree topology. If the \textit{M. polymorpha} \textit{U} and \textit{V} gametologs are less diverged from \textit{Haplomitriopsida} sequences than from each other, the divergence of the \textit{M. polymorpha} gametologs was inferred to have occurred before the divergence of the \textit{Haplomitriopsida} and the \textit{Marchantia}.

Conversely, if the \textit{M. polymorpha} \textit{U} and \textit{V} gametologs are less diverged from one another than they are to some other liverwort sequences, then their divergence time was inferred based on the relative positions in the tree topology of the

Figure 3. \textit{BPCU} and its gametolog \textit{BPCV} are required for induction of sexual reproduction, but only \textit{BPCU} confers the feminization

(A) Gene organizations of \textit{BPCU} and \textit{BPCV} loci and schematic illustrations of \textit{bpcU} and \textit{bpcV} mutations. The coding regions are colored. Gray boxes indicate 5' and 3' untranslated regions. Arrowheads indicate the positions of gRNAs. The deleted region in \textit{BPCU} is shown by a bent broken line.

(B) No sex organ formation in \textit{bpcU} [\textit{U}] and \textit{bpcV} [\textit{V}]. Female WT [\textit{U}] (Tak-2; top left), male WT [\textit{V}] (Tak-1; bottom left), \textit{bpcU-1} [\textit{U}] (top right), and \textit{bpcV-1} [\textit{V}] (bottom right) plants are shown. Images 1 month after reproductive induction are shown. Arrowheads indicate sex organs. Scale bar, 10 mm.

(C) Genetic complementation of the reproductive induction phenotype of \textit{bpcu} with \textit{BPCU} or \textit{BPCV}. Genomic \textit{BPCU}-transformed \textit{bpcU} [\textit{U}] (g\textit{BPCU}\#2 \textit{bpcU-1} [\textit{U}]; top) and genomic \textit{BPCV}-transformed \textit{bpcU} [\textit{U}] (g\textit{BPCV}\#1 \textit{bpcU-1} [\textit{U}]; bottom) plants are shown. Arrowheads indicate sex organs. Scale bar, 10 mm.

(D) Magnified images of sex organs in g\textit{BPCU}\#2 \textit{bpcU-1} [\textit{U}] (top) and g\textit{BPCV}\#1 \textit{bpcU-1} [\textit{U}] (bottom). Sex organs 1 month after reproductive induction are shown. Left and right panels: receptacles and gamete-containing organs are shown, respectively. Scale bars, 1 mm (left panels); 50 \textmu m (right panels).

(E) Loss of \textit{BPCU} or \textit{BPCV} impairs reproductive induction. Gemmae were grown under white light for 10 days and transferred to the reproductive induction condition (STAR Methods). Days at which a visible receptacle was first formed after reproductive induction are shown. No sex organs were formed in \textit{bpcU} (\textit{bpcU-1}, \textit{bpcU-2}, and \textit{bpcU-3}) and \textit{bpcV} (\textit{bpcV-1}, \textit{bpcV-2}, and \textit{bpcV-3}) over 60 days. Bars represent mean ± SD (\textit{n} = 3).

See also Figures S4 and S5.
Figure 4. Phylogenetic tree of BPCU and BPCV

The phylogram was reconstructed using nucleotide alignments as described in the STAR Methods section. The liverwort sequences are shown in purple, and the clades containing BPCU and BPCV are highlighted in yellow and turquoise, respectively. The phylogenetic positions of the BPCU and BPCV of *M. polymorpha* are demarcated by arrows with the pS value between the two *M. polymorpha* gametologs indicated (0.77). Taxa are color coded as follows: liverworts, purple; Haplomitriopsida, dark purple; Marchantiopsida, bright purple; Jungermanniopsida, violet; hornworts, dark green; mosses, light green; lycophytes, orange; and ferns, brown. Numbers at branches indicate posterior probability values. See also Data S1B and S2.
DISCUSSION

The chromosome level assemblies of the *M. polymorpha* U and V chromosomes allow a test of the theoretical predictions that UV sex chromosomes should evolve symmetrically\(^{19}\) and that both should show minimal degeneration.\(^{23}\) While the *M. polymorpha* U and V chromosomes have evolved symmetrically, the low gene numbers (<5% of each autosome) suggest that they have undergone substantial degeneration.\(^{19–21}\) The low gene density is consistent with this, as repetitive sequence accumulation can occur if there are few genes.\(^{19–21}\) No obvious synteny is detectable between the U and V chromosomes, and the evolutionary strata inferred are not contiguous, implying that each sex chromosome has undergone independent chromosomal rearrangements.

Our results demonstrate that a U-linked transcription factor gene, *BPCU*, is the sex determinant *Feminizer* in *M. polymorpha*. Unexpectedly, however, *BPCU* is also essential for reproductive induction in females, a function which, in males, is carried out by its gametolog on the V chromosome, *BPCV*. The conserved function in the reproductive transition of both *BPCU* and *BPCV* suggests that this presumably ancestral function was retained despite only *BPCU* evolved a sex-determining (feminizing) function promoting female organ development (Figure 6). That *BPCU* and *BPCV* have apparently equivalent, and required, roles in the reproductive transition explains why the originally isolated *bpcU* mutants were aneuploid, with *BPCV* enabling the transition to reproductive development in the *bpcU* [UV] background.

Including *BPCU* and *BPCV*, seven ancient gametolog pairs are present on the *M. polymorpha* sex chromosomes, implying that large non-recombining regions were already established in the ancestral liverwort, which is estimated to have existed in the Mid-Silurian, ca. 430 mya. This age precedes any other known sex chromosome system. The oldest XY and ZW angiosperm sex chromosomes evolved in the Cenozoic (30–40 mya),\(^9\) the U/V sex chromosome systems in green\(^{39}\) and brown algae\(^{40}\) in Data S1C. See Data S2 and Figure 4 for phylogenetic trees of each of the gametologs. For the Marchantopsida and Jungermanniopsida, we included only dioicous species where the transcriptome (*) or genome (\(\square\)) was apparently derived from a single sex (i.e., where only single orthologs were detected for all the genes that we infer to be in the oldest evolutionary stratum [shaded purple]). Predicted nodes of the origins (color coded) of the non-recombining regions of the *M. polymorpha* U and V chromosomes are indicated on the chromosome schematic and the presently accepted liverwort phylogenetic tree (above the matrix).\(^{37,38}\) Geologic ages of numbered nodes as previously estimated\(^{37,38}\) are listed at left of the tree. See also Data S1A, S1C, and S2.
functions in the initiation of reproductive development, like its ancestral function in the reproductive transition but did not acquire a role in specifying female sex organ development, perhaps via its regulation of FGMYB. This allele evolved into BPCU, while the alternative allele (now the gametolog), BPCV, retained the ancestral function in the reproductive transition but did not acquire a sex-determining function. This can potentially explain the fact that, in addition to its role in sex determination, BPCU functions in the initiation of reproductive development, like its gametolog BPCV. This scenario is distinct from that described for other sex-determining systems. For example, in therian mammals, the male sex determinant SRY does not share the non-sex determination function of its X-linked gametolog, SOX3, which was presumably retained from the common ancestral gene. Likewise, the sex determinant MID in the chlorophyte alga Volvox carteri evolved directly from an already male-specific gene within the ancestral mating-type locus. Our findings provide a distinct trajectory for the evolution of a sex determinant (the Feminizer) in another haploid system and further implies that a pair of gametologs that initially appear functionally equivalent may have been differentiated to serve females and males unequally.

**STAR METHODS**

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

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Plant materials and culture
- **METHOD DETAILS**
  - Plant transformation
  - U chromosome assembly
  - Genetic nomenclature
  - Microscopy
  - RNA extraction and reverse-transcription
  - Plasmid construction for expressing BPCU and BPCV
  - Genome editing
  - Genotyping and sex diagnosis
  - Generation of transgenic lines expressing genomic BPCU and genomic BPCV
  - RT-PCR
  - Quantitative real-time RT-PCR
  - DPI-ELISA with BPCU and selected DNA-probes
  - Flow cytometry analysis
  - Chromatin profiling analysis
  - Analysis of sex chromosome evolution
  - Data analysis of CUT&RUN
  - Transcriptome Mapping
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - General statistical analyses
  - Gene expression analysis of BPCU and BPCV

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2021.10.023.

**ACKNOWLEDGMENTS**

Authors thank Qi Zhou for critical reading of the manuscript, Aino Komatsu for isolation of the segregated plants with both U and V chromosomes, and Keitaro Okahashi for initial contribution to select candidate genes. Computations were performed on the NIG supercomputer at ROIS National Institute of Genetics, the supercomputer of ACCMS at Kyoto University, and the supercomputer of the HPC Computing Center at Kindai University. We thank Takefumi

Figure 6. Genetic model of feminization by BPCU and reproductive induction by BPCU and BPCV

The female and male plants carry BPCU on the U chromosome and BPCV on the V chromosome, respectively. BPCU and BPCV share a function involved in inducing sexual reproduction. Following the transition to reproductive development, the female thallus initiates expression of the transcription factor FGMYB that promotes development of female reproductive organs. In gametophytes that do not carry BPCU, expression of the antisense long non-coding (Inc) RNA SUF prevents FGMYB expression, leading to male reproductive development. The U-linked Feminizer, BPCU functions to repress antisense SUF expression in females, allowing FGMYB expression and leading to female reproductive development.
Kondo and Yukari Sando (NGS core facility of Graduate School of Biostudies, Kyoto University) for supporting the RNA sequencing (RNA-seq) analysis. Research was funded by JSPS/MEXT KAKENHI, grant numbers: JP17H07424 and JP19H05675 to T. Kohchi, 20510185 and 24510272 to K.T.Y., JP20H05780 to S.Y., JP20K15783 to Y.T., JP16H06279 (PAGS) to Y.N., JP19K16166 to Y. Yoshitake, and JP19K16167 to Y. Yasui. T. Kohchi was also supported by JSPS Bilateral Joint Research Projects JPJSBP120192003. D.W. was supported by the German Research Foundation (DFG; WA2814/4-1). C.L. was supported by European Research Council under the European Union's Horizon 2020 research and innovation programme 757600. F.B. was supported by the FWF grant DK1238 chromosome dynamics. J.L.B. was supported by Australian Research Council DP200100225.

AUTHOR CONTRIBUTIONS

M.I. and T. Kajiwara performed a majority of the experiments by the help and supervision of Y. Yasui, Y. Yoshitake, M.M., N.S., R.N., S.Y., and T. Kohchi. Y. Yoshitake, T. Kajiwara, S.K., and S.A.M. performed bioinformatic analysis on sex chromosome evolution. J.L.B. performed bioinformatic analysis on sex chromosome evolution. D.W., K.H., and K.K. performed RNA binding experiments. T. Kohchi, F.B., J.L.B., and K.T.Y. performed genome annotation and constructed the genome database under the supervision of Y.N. S.S. and J.L.B. performed bioinformatic analysis on sex chromosome evolution. D.W., K.H., and K.K. performed RNA binding experiments. T. Kohchi, F.B., J.L.B., and K.T.Y. designed the project and contributed to conceptualize the work. T. Kohchi, K.T.Y., F.B., and J.L.B. wrote the manuscript with input from all authors. All authors jointly interpreted the data and thoroughly checked the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 7, 2021
Revised: August 2, 2021
Accepted: October 8, 2021
Published: November 3, 2021

REFERENCES

1. Bachtrog, D., Kirkpatrick, M., Mank, J.E., McDaniel, S.F., Pires, J.C., Rice, W., and Valenzuela, N. (2011). Are all sex chromosomes created equal? Trends Genet. 27, 350–357.
2. Akagi, T., and Charlesworth, D. (2019). Pleiotropic effects of sex-determining genes in the evolution of dioecy in two plant species. Proc. R. Soc. B: Biol. Sci. 286, 20191805.
3. Akagi, T., Henry, I.M., Tao, R., and Comai, L. (2014). Plant genetics. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. Science 346, 646–650.
4. Boualem, A., Troade, C., Camps, C., Lemhendi, A., Morin, H., Sari, M.A., Fraenkel-Zagouri, R., Kovalski, I., Dogimont, C., Perl-Treves, R., and Bendahmane, A. (2014). A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. Science 350, 688–691.
5. Müller, N.A., Kersten, B., Leite Montalvo, A.P., Mank, J.E., Pires, J.C., Rice, W., and Valenzuela, N. (2011). Are all sex chromosomes created equal? Trends Genet. 27, 350–357.
6. Akagi, T., Henry, I.M., Tao, R., and Comai, L. (2014). Plant genetics. A Y-chromosome-encoded small RNA acts as a sex determinant in persimmons. Science 346, 646–650.
7. Müller, N.A., Kersten, B., Leite Montalvo, A.P., Mank, J.E., Pires, J.C., Rice, W., and Valenzuela, N. (2011). Are all sex chromosomes created equal? Trends Genet. 27, 350–357.
8. Charlesworth, B., and Charlesworth, D. (1978). Model for evolution of dioecy and gynodioecy. Am. Nat. 112, 975–997.
9. Renner, S.S., and Müller, N.A. (2021). Plant sex chromosomes defy evolutionary model of expanding recombination suppression and genetic degeneration. Nat. Plants 7, 392–402.
10. Coelho, S.M., Guenon, J., Lipinska, A.P., Cock, J.M., and Umen, J.G. (2018). UV chromosomes and haploid sexual systems. Trends Plant Sci. 23, 794–807.
11. Coelho, S.M., Mignerot, L., and Cock, J.M. (2019). Origin and evolution of sex-determination systems in the brown algae. New Phytol. 222, 1751–1756.
12. Higo, A., Kawashima, T., Borg, M., Zhao, M., López-Vidriero, I., Sakayama, H., Montgomery, S.A., Sekimoto, H., Hackenberg, D., Shimamura, M., et al. (2018). Transcription factor DUO1 generated by neo-functionalization is associated with evolution of sperm differentiation in plants. Nat. Commun. 9, 5283.
13. Liu, X., Bogaert, K., Engelen, A.H., Leliæft, F., Roleda, M.Y., and De Clerck, O. (2017). Seaweed reproductive biology: environmental and genetic controls. Bot. Mar. 60, 89–108.
14. Bull, J.J. (1978). Sex chromosomes in haploid dioecy: a unique contrast to Muller's theory for diploid dioecy. Am. Nat. 112, 245–250.
15. Lewis, K.R. (1961). The genetics of bryophytes. T. Brit. Bryol. Soc. 4, 111–130.
16. Bachtrog, D., Mank, J.E., Peichel, C.L., Kirkpatrick, M., Otto, S.P., Ashman, T.L., Hahn, M.W., Kitano, J., Mayrose, I., Ming, R., et al.; Tree of Sex Consortium (2014). Sex determination: why so many ways of doing it? PLoS Biol. 12, e1001899.
17. Bull, J.J. (1983). Evolution of Sex Determining Mechanisms (The Benjamin/ Cummings Publishing Company).
18. Bowman, J.L. (2016). A brief history of Marchantia from Greece to genomics. Plant Cell Physiol. 57, 210–229.
19. Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., et al. (2017). Insights into land plant evolution garnered from the Marchantia polymorpha genome. Cell 171, 287–304.e15.
20. Montgomery, S.A., Tanizawa, Y., Galik, B., Wang, N., Ito, T., Mochizuki, K., Akincheva, S., Bowman, J.L., Cognat, V., Marechal-Drouard, L., et al. (2020). Chromatin organization in early land plants reveals an ancestral association between H3K27me3, transposons, and constitutive heterochromatin. Curr. Biol. 30, 573–588.e7.
21. Yamato, K.T., Ishizaki, K., Fujisawa, M., Okada, S., Nakayama, S., Fujishita, M., Bando, H., Yodoya, K., Hayashi, K., Bando, T., et al. (2007). Gene organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system. Proc. Natl. Acad. Sci. USA 104, 6472–6477.
22. Haupt, G. (1932). Beiträge zur Zylogenie der Gattung Marchantia (L.). I. Zeitschr. Ind. Abst. Vererbungs. 62, 367–428.
23. Carey, S.B., Jenkins, J., Lovell, J.T., Maumus, F., Payton, A.C., Shu, S., Tiley, G.P., Fernandez-Pozo, N., Healey, A., et al. (2021). To dry perchance to live: insights from the genome of the desiccation-tolerant lichen Pertusaria phaeophyta. Nature 583, 1339–1356.
24. Silva, A.T., Gao, B., Fisher, K.M., Mishler, B.D., Ekwealar, J.T.B., Stark, L.R., Li, X., Zhang, D., Brinda, J.C., et al. (2021). To dry perchance to live: insights from the genome of the desiccation-tolerant bryophyte Syntrichia caninervis. Plant J. 105, 1339–1356.
25. Ahmed, S., Cock, J.M., Pessia, E., Luthringer, R., Cornier, A., Robuchon, M., Sterck, L., Peters, A.F., Dittami, S.M., Corre, E., et al. (2014). A haploid model of sex determination in the brown alga Ectocarpus sp. Curr. Biol. 24, 1945–1957.
26. Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Armit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293.
27. Nakayama, S., Fujishita, M., Sone, T., and Ohyama, K. (2001). Additional locus of rDNA sequence specific to the X chromosome of the liverwort, Marchantia polymorpha. Chromosome Res. 9, 469–473.

28. Hisanaga, T., Okahashi, K., Yamaoka, S., Kajiwara, T., Nishihama, R., Shimamura, M., Yamato, K.T., Bowman, J.L., Kohchi, T., and Nakajima, K. (2019). A cis-acting bidirectional transcription switch controls sexual dimorphism in the liverwort. EMBO J. 38, 663–669.

29. Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14, 178–192.

30. Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-seq data. Bioinformatics 25, 1952–1958.

31. Sangwan, I., and O’Brien, M.R. (2002). Identification of a soybean protein that interacts with GAGA element dinucleotide repeat DNA. Plant Physiol. 129, 1788–1794.

32. Theune, M.L., Bloss, U., Brand, L.H., Ladwig, F., and Wanke, D. (2019). The timescale of early land plant evolution. Proc. Natl. Acad. Sci. USA 116, 2107–2112.

33. Geng, S., De Hoff, P., and Umen, J.G. (2014). Evolution of sexes from an ancestral mating-type specification pathway. PLoS Biol. 12, e1001904.

34. Li, F.W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M.S., Bennett, T., Blizquez, M.A., et al. (2020). Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat. Plants 6, 259–272.

35. Forrest, L.L., Davis, E.C., Long, D.G., Crandall-Stotler, B.J., Clark, A., and Hollingsworth, M.L. (2006). Unraveling the evolutionary history of the liverworts (Marchantiophyta): multiple taxa, genomes and analyses. Bryologist 109, 1–31.

36. He-Nygrén, X., Juslén, A., Ahonen, I., Glenny, D., and Pippoo, S. (2006). Illuminating the evolutionary history of liverworts (Marchantiophyta) - towards a natural classification. Cladistics 22, 1–31.

37. Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, A., Coelho, S.M. (2017). Multiple gene movements into and out of haploid dimorphism in the liverwort. EMBO J.

38. Yamamoto, K., Hamaji, T., Kawai-Toyooka, H., Matsuzaki, R., Takahashi, T., Shirakawa, M., Endo, M., Shigenobu, S., et al. (2016). Transcriptional framework of male gametogenesis in the liverwort Marchantia polymorpha. Plant Cell Physiol. 57, 325–338.

39. Thorne, L., Ooka, K., Nishihama, R., Yoshitake, Y., Ishida, S., Inoue, K., Saito, M., Okahashi, K., Bao, H., Nishida, H., Yamaguchi, K., et al. (2018). Generative cell specification requires transcription factors evolutionarily conserved in land plants. Curr. Biol. 28, 479–486.e5.

40. Katsura, Y., Kondo, H.X., Ryan, J., Harley, V., and Satta, Y. (2018). The evolutionary process of mammalian sex determination genes focusing on marsupial SYR1. BMC Evol. Biol. 18, 3.

41. Deblaere, R., Bytuber, B., De Greve, H., Deboeck, F., Schell, J., Van Montagu, M., and Leemans, J. (1985). Efficient octopine Ti plasmid-derived vectors for Agrobacterium-mediated gene transfer to plants. Nucleic Acids Res. 13, 4777–4788.

42. Gamborg, O.L., Miller, R.A., and Ojima, K. (1968). Nutrient requirements of suspension cultures of soybean root cells. Exp. Cell Res. 50, 151–158.

43. Pluthero, F.G. (1993). Rapid purification of high-activity Taq DNA polymerase. Nucleic Acids Res. 21, 4850–4851.

44. Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclelease strategy for high-resolution mapping of DNA binding sites. eLife 6, e21856.

45. Higo, A., Niwa, M., Yamato, K.T., Yamada, L., Sawada, H., Sakamoto, T., Kurata, T., Shirakawa, M., Endo, M., Shigenobu, S., et al. (2016). Transcriptional framework of male gametogenesis in the liverwort Marchantia polymorpha. Plant Cell Physiol. 57, 325–338.

46. Sugano, S., Nishihama, R., Shira, T., Matsuda, Y., Hisanaga, T., Shimada, T., Hara-Nishimura, I., Asakabe, K., and Kohchi, T. (2018). Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in Marchantia polymorpha. PLoS ONE 13, e0205117.

47. Koele, E., Suetanag, N., Imano, W., Moto, S., Nomura, Y., Stoloz, S.C., Nakagami, H., Kohchi, T., and Nishihama, R. (2020). Regulation of photosynthetic carbohydrate metabolism by a Raf-like kinase in the liverwort Marchantia polymorpha. Plant Cell Physiol. 61, 631–643.

48. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890.

49. Patho, R., Duggal, G., Love, M.I., irzary, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14, 417–419.

50. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

51. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.

52. Danecsek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFTools. GigaScience 10, gbi008.

53. Ramírez, F., Ryan, D.P., Grünig, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dünard, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44 (W1), W160–W165.

54. Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 4, 1521.

55. Ruan, J., and Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. Nat. Methods 17, 155–158.

56. Wu, T.D., and Watanabe, K.C. (2005). GMAP: a genomic mapping and alignment web server for deep-sequencing data analysis. Nucleic Acids Res. 33, W112–W116.
Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., and Bollback, J.P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294, 2310–2314.

Korber, B. (2000). HIV signature and sequence variation analysis. In Computational Analysis of HIV Molecular Sequences, A.G. Rodrigo, and G.H. Learn, Jr., eds. (Kluwer Academic), pp. 55–72.

Chiyoda, S., Linley, P.J., Yamato, K.T., Fukuzawa, H., Yokota, A., and Kohchi, T. (2007). Simple and efficient plastid transformation system for the liverwort Marchantia polymorpha L. suspension-culture cells. Transgenic Res. 16, 41–49.

Ishizaki, K., Nishihama, R., Yamato, K.T., and Kohchi, T. (2016). Molecular genetic tools and techniques for Marchantia polymorpha research. Plant Cell Physiol. 57, 262–270.

Kubota, A., Kita, S., Ishizaki, K., Nishihama, R., Yamato, K.T., and Kohchi, T. (2014). Co-option of a photoperiodic growth-phase transition system during land plant evolution. Nat. Commun. 5, 3668.

Ishizaki, K., Chiyoda, S., Yamato, K.T., and Kohchi, T. (2008). Agrobacterium-mediated transformation of the haploid liverwort Marchantia polymorpha L., an emerging model for plant biology. Plant Cell Physiol. 49, 1084–1091.

Kubota, A., Ishizaki, K., Hosaka, M., and Kohchi, T. (2013). Efficient Agrobacterium-mediated transformation of the liverwort Marchantia polymorpha using regenerating thalli. Biosci. Biotechnol. Biochem. 77, 167–172.

Bowman, J.L., Araki, T., Arteaga-Vazquez, M.A., Berger, F., Dolan, L., Haseloff, J., Ishizaki, K., Kyozuka, J., Lin, S.S., Nagasaki, H., et al. (2016). The naming of names: guidelines for gene nomenclature in Marchantia. Plant Cell Physiol. 57, 257–261.

Ishizaki, K., Mizutani, M., Shimamura, M., Masuda, A., Nishihama, R., and Kohchi, T. (2013). Essential role of the E3 ubiquitin ligase nopperabo1 in schizogenous intercellular space formation in the liverwort Marchantia polymorpha. Plant Cell 25, 4075–4084.

Fujisawa, M., Hayashi, K., Nishio, T., Bando, T., Okada, S., Yamato, K.T., Fukuzawa, H., and Ohyama, K. (2001). Isolation of X and Y chromosome-specific DNA markers from a liverwort, Marchantia polymorpha, by representation difference analysis. Genetics 159, 981–985.

Inoue, K., Nishihama, R., Kataoka, H., Hosaka, M., Manabe, R., Nomoto, M., Tada, Y., Ishizaki, K., and Kohchi, T. (2016). Phytochrome signaling is mediated by PHYTOCHROME INTERACTING FACTOR in the liverwort Marchantia polymorpha. Plant Cell 28, 1406–1421.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 25, 402–408.

Brand, L.H., Kirchner, T., Hummel, S., Chaban, C., and Wanke, D. (2010). DPI-ELISA: a fast and versatile method to specify the binding of plant transcription factors to DNA in vitro. Plant Methods 6, 25.

Fischer, S.M., Böser, A., Hirsch, J.P., and Wanke, D. (2016). Quantitative analysis of protein-DNA interaction by qDPI-ELISA. Methods Mol. Biol. 1482, 49–66.

Nishihama, R., Ishizaki, K., Hosaka, M., Matsuda, Y., Kubota, A., and Kohchi, T. (2015). Phytochrome-mediated regulation of cell division and growth during regeneration and sporulation development in the liverwort Marchantia polymorpha. J. Plant Res. 128, 407–421.

Dong, S., Zhao, C., Zhang, S., Wu, H., Mu, W., Wei, T., Li, N., Wan, T., Liu, H., Cui, J., et al. (2019). The amount of RNA editing sites in liverwort organellar genes is correlated with GC content and nuclear PPR protein diversity. Genome Biol. Evol. 11, 3233–3239.

Leebens-Mack, J.H., Barker, M.S., Carpenter, E.J., Deyholos, M.K., Gitzendanner, M.A., Graham, S.W., Grosse, I., Li, Z., Melkonian, M., Mirarab, S., et al.; One Thousand Plant Transcriptomes Initiative (2019). One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574, 679–685.

Linde, A.M., Sawangproh, W., Cronberg, N., Szövényi, P., and Lagercrantz, U. (2020). Evolutionary history of the Marchantia polymorpha complex. Front. Plant Sci. 11, 829.

RadhaKrishnan, G.V., Keller, J., Rich, M.K., Vernié, T., Mbainga Mbainga, D.L., Vigneron, N., Coutret, L., Clemente, H.S., Libourel, C., Cheema, J., et al. (2020). An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. Nat. Plants 6, 280–289.

Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matacic, N., Ayvampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A., et al. (2014). Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc. Natl. Acad. Sci. USA 111, E4859–E4868.

Neil, M., and Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–426.
# STAR METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Antibodies** | | |
| Rabbit polyclonal H3 | Abcam | Cat# ab1791; RRID: AB_302613 |
| Rabbit polyclonal H3K27me3 | Millipore | Cat# 07-449; RRID: AB_310624 |
| HRP Anti-6X His tag antibody | Abcam | Cat# ab1187; RRID: AB_298652 |
| **Bacterial and virus strains** | | |
| *Escherichia coli* DH5α | Widely distributed | N/A |
| *Agrobacterium tumefaciens* GV2260 | Deblaere et al. | N/A |
| *Escherichia coli* BL21-CodonPlus-RIL | Agilent | Cat# 230240 |
| **Biological samples** | | |
| HEK293 DNA | Danhua Jiang, Beijing, China | N/A |
| **Chemicals, peptides, and recombinant proteins** | | |
| Gamborg’s B5 salts | Gamborg et al. | N/A |
| hygromycin B | Nacalai Tesque | Cat# 07296-24 |
| cefotaxime (CLAFORAN) | Sanofi | Cat# 6132409D1050 |
| TRIzol reagent | ThermoFisher Scientific | Cat# 15596018 |
| ReverTra Ace | Toyobo Life Science | Cat# TR-T-101 |
| RNase A | ThermoFisher Scientific | Cat# EN0531 |
| RNase A | Sigma-Aldrich | Cat# R5125 |
| RQ1 RNase-Free DNase | Promega | Cat# M610A |
| Proteinase K | Thermo Fisher Scientific | Cat# EC0491 |
| Taq DNA polymerase | Plutero | N/A |
| Bio-Mag Plus Concanavalin A coated beads | polysciences | Cat# 86057 |
| cOmplete Protease Inhibitor Cocktail | Roche | Cat# 11697498001 |
| pA-MNase | Skene and Henikoff | N/A |
| propidium iodide | Nacalai Tesque | Cat# 29037-76 |
| ortho-phenylenediamine (OPD) | Merck-Sigma/Aldrich | Cat# P5412 |
| sulfuric acid (2 N H₂SO₄) | Roth | Cat# 2609.1 |
| SYBR Green Nucleic Acid Gel Stain | Lonza | Cat# 50513 |
| **Critical commercial assays** | | |
| KOD FX Neo DNA polymerase | Toyobo Life Science | Cat# KFX-201 |
| KOD One PCR Master Mix | Toyobo Life Science | Cat# KMM-101 |
| pENTR/D-TOPO Cloning kit | Thermo Fisher Scientific | Cat# K240020 |
| Gateway LR clonase II Enzyme mix | Thermo Fisher Scientific | Cat# 11791020 |
| NucleoSpin Gel and PCR Clean-up Kit | Macherey & Nagel | Cat# 740609.50 |
| Pre-blocked clear flat-bottom Streptavidin Coated Plates (96-wells) | Pierce Thermo Fisher Scientific | Cat# 15124 |
| Penta-His HRP Conjugate Kit | QIAGEN | Cat# 34460 |
| **Deposited data** | | |
| *Marchantia polymorpha* genome v3.1 | Bowman et al.¹⁹ | https://marchantia.info; SRA: SRR1800537 |
| *Marchantia polymorpha* genome v5.1 | Montgomery et al.²⁰ | https://marchantia.info; SRA: PRJNA553138 |
| *Marchantia polymorpha* genome v6.1 | This paper | https://marchantia.info; SRA: PRJDB11173 |
| Iso-seq and RNA-seq for gene annotations | Montgomery et al.²⁰ | SRA: PRJD88530 and PRJNA251267 |
| Tak-1 thallus CUT&RUN | Montgomery et al.²⁰ | SRA: PRJNA553138 |

(Continued on next page)
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Tak-2 and bpcU thallus CUT&RUN | This paper | SRA: PRJNA757234 |
| Tak-1, Tak-2 and bpcU thallus RNA-seq | This paper | SRA: PRJNA757569 |
| RNA-seq for expression analyses in various tissues | Bowman et al., Higo et al., and Yamaoka et al. | SRA: DRR050351-53; DRR050346-48; DRR118943-45; DRR118949-51; SRR4450260-62; SRR4450254-56 |

### Experimental models: Organisms/strains

- **Marchantia polymorpha Tak-1** Bowman et al. 19 N/A
- **Marchantia polymorpha Tak-2** Bowman et al. 19 N/A
- **Marchantia polymorpha BC3-38** Yamaoka et al. 50 N/A
- **Marchantia polymorpha UV plant** This paper N/A
- **Marchantia polymorpha bpcU-101** This paper N/A
- **Marchantia polymorpha bpcU-102** This paper N/A
- **Marchantia polymorpha gBPCU** This paper N/A
- **Marchantia polymorpha gBPCU fgmyb-31** This paper N/A
- **Marchantia polymorpha gBPCU fgmyb-41** This paper N/A
- **Marchantia polymorpha bpcU-1** This paper N/A
- **Marchantia polymorpha bpcU-2** This paper N/A
- **Marchantia polymorpha bpcU-3** This paper N/A
- **Marchantia polymorpha bpcV-1** This paper N/A
- **Marchantia polymorpha bpcV-2** This paper N/A
- **Marchantia polymorpha bpcV-3** This paper N/A
- **Marchantia polymorpha gBPCU bpcU-1** This paper N/A
- **Marchantia polymorpha gBPCU bpcU-1** This paper N/A
- **Marchantia polymorpha gBPCU bpcU-1** This paper N/A
- **Marchantia polymorpha gBPCU bpcU-1** This paper N/A
- **Marchantia polymorpha gBPCV bpcU-1** This paper N/A
- **Marchantia polymorpha gBPCV bpcU-1** This paper N/A
- **Marchantia polymorpha gBPCV bpcU-1** This paper N/A
- **Marchantia polymorpha gBPCV bpcU-1** This paper N/A

### Oligonucleotides

See Table S1 This paper N/A

### Recombinant DNA

- **pENTR D-TOPO** Thermo Fisher Scientific Cat# 45-0218
- **BCPU gene pENTRdTOPO** This paper N/A
- **BCPV gene pENTRdTOPO** This paper N/A
- **pMpGWB101** Ishizaki et al. 51 GenBank: LC057443
- **pMpGWB301** Ishizaki et al. 51 GenBank: LC057517
- **BCPU gene pMpGWB101** This paper N/A
- **BCPU gene pMpGWB301** This paper N/A
- **BCPV gene pMpGWB301** This paper N/A
- **pMpGE_En03** Sugano et al. 52 GenBank: LC090755
- **pMpGE_En04** Koide et al. 53 N/A
- **pBC-GE12** Koide et al. 53 N/A
- **pBC-GE23** Koide et al. 53 N/A
- **pBC-GE34** Koide et al. 53 N/A
- **BCPU.1 gRNA pMpGE_En03** This paper N/A
- **BCPU gRNALDs pMpGE_En04** This paper N/A
- **BCPV.1 gRNA pMpGE_En03** This paper N/A
- **BCPV.2 gRNA pMpGE_En03** This paper N/A

(Continued on next page)
Continued

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| FGMYB gRNA3 pMpGE_En03 | This paper | N/A |
| FGMYB gRNA4 pMpGE_En03 | This paper | N/A |
| pMpGE010 | Sugano et al. | GenBank: LC090756 |
| pMpGE011 | Sugano et al. | GenBank: LC090757 |
| pMpGE017 | Koide et al. | N/A |
| BPCU.1 gRNA pMpGE010 | This paper | N/A |
| BPCU gRNA Lds pMpGE017 | This paper | N/A |
| BPCV.1 gRNA pMpGE010 | This paper | N/A |
| BPCV.2 gRNA pMpGE010 | This paper | N/A |
| FGMYB gRNA3 pMpGE011 | This paper | N/A |
| FGMYB gRNA4 pMpGE011 | This paper | N/A |
| BPCU cds pENTRdTOPO | This paper | N/A |
| BPCV cds pENTRdTOPO | This paper | N/A |
| pET-DEST42 | Thermo Fisher Scientific | Cat# 12276010 |
| BPCU cds pET-DEST42 | This paper | N/A |
| BPCV cds pET-DEST42 | This paper | N/A |

Software and algorithms

| Software | Source | URL |
|----------|--------|-----|
| fastp v0.20.1 | Chen et al. | https://github.com/OpenGene/fastp |
| Salmon v0.14.1 | Patro et al. | https://github.com/COMBINE-lab/salmon |
| bowtie2 v2.4.1 | Langmead and Salzberg | https://github.com/BenLangmead/bowtie2 |
| STAR v2.7.3a | Dobin et al. | https://github.com/alexdobin/STAR |
| samtools v1.9.0 | Danecek et al. | https://github.com/samtools/samtools |
| deepTools v3.5.0 | Ramirez et al. | https://github.com/deeptools/deepTools |
| faCount | UCSC Genome Browser | https://hgdownload.cse.ucsc.edu/admin/exe/faCount |
| SICER2 | Zang et al. | https://zanglab.github.io/SICER2/ |
| tximport | Soneson et al. | https://bioconductor.org/packages/release/bioc/html/tximport.html |
| IGV v2.8.0 | Thorvaldsdóttir et al. | https://software.broadinstitute.org/software/igv/ |
| Redbean v2.5 | Ruan and Le | https://github.com/ruanjue/wtdbg2 |
| NextPolish v1.3.1 | Hu et al. | https://github.com/Nextomics/NextPolish |
| Pilon v1.23 | Walker et al. | https://github.com/broadinstitute/pilon |
| NextDenovo v2.2-beta.0 | GrandOmicns | https://github.com/Nextomics/NextDenovo |
| 3D-DNA pipeline | Dudchenko et al. | https://github.com/aidenlab/3d-dna |
| GMAP ver. 2019.09.12 | Wu and Watanabe | http://research-pub.gene.com/gmap/ |
| Se-Al v2.0a11 | Institute of Evolutionary Biology University of Edinburgh | http://tree.bio.ed.ac.uk/software/seal/ |
| MrBayes 3.2.1 | Huelsenbeck and Ronquist | https://nbisweden.github.io/MrBayes/index.html |
| FigTree v1.4.0 | Huelsenbeck et al. | http://tree.bio.ed.ac.uk/software/figtree/ |
| SNAP v2.1.1 | Korber | http://hcv.lanl.gov/content/sequence/SNAP/SNAP.html |

Other

| Equipment | Source | URL |
|-----------|--------|-----|
| BD FACSARIA III | BD Biosciences | https://www.bdbiosciences.com/en-us |
| BD Accuri C6 Flow Cytometer | BD Biosciences | https://www.bdbiosciences.com/en-us |
| CFX96 real-time PCR detection system | Bio-Rad Laboratories | https://www.bio-rad.com |
| NanoDrop 2000 spectrophotometer | Thermo Scientific | https://www.thermofisher.com/us/en/home.html |
RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Takayuki Kohchi (tkohchi@lif.kyoto-u.ac.jp).

Materials availability
Plasmids and transgenic lines generated during this study are available from the Lead Contact without restriction.

Data and code availability
Data have been deposited at DDBJ/ENA/GenBank databases and are publicly available as of the date of publications. Accession numbers are listed in the Key resources table. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plant materials and culture
The female accession, Takaragaike-2 (Tak-2), and the male accession, Takaragaike-1 (Tak-1), and their backcrossed female line BC3-38 of Marchantia polymorpha ssp. ruderalis were used in this study. The UV plant (Figure 1C) was isolated among progeny sporelings derived from a cross between BC3-38 and Tak-1. These plants were grown primarily on half-strength Gamborg’s B5 medium containing 1% agar, or on vermiculite supplemented by 2,000-fold-diluted HYPONeX (Hyponex Japan, Osaka, Japan) under continuous white light (50-60 μmol photons m⁻² s⁻¹; CCFL OPT-40C-N-L, Optorom) at 22°C. To induce reproductive growth, 10-day-old gemmalings or older thalli were transferred to white light conditions supplemented with far-red light (30 μmol photons m⁻² s⁻¹, VBL-TFL600-IR730*, Valore, Kyoto, Japan).

METHOD DETAILS

Plant transformation
Agrobacterium-mediated transformation of M. polymorpha was performed using standard protocols for sporelings for the initial screening or thalli for all the other lines. Transformants were selected on the media containing 100 μg ml⁻¹ cefotaxime and either 10 μg ml⁻¹ hygromycin or 0.5 μM chlorsulfuron according to the selection marker genes in the binary vectors. After co-cultivation with Agrobacterium, plants were washed with sterile water and plated on the selective media. Regenerating antibiotic-resistant transformants (T1) were transferred to new selective media and grown until gemmae (G1) were produced.

U chromosome assembly
Draft genome assembly was performed from PacBio reads using Redbean ver. 2.5 with default settings. The draft contigs were polished by NextPolish ver 1.3.1 using PacBio long reads and Illumina short reads, followed by a further round of polishing by Pilon ver. 1.23 with Illumina short reads. Separately, a draft genome was also reconstructed using NextDenovo v2.2-beta.0 (https://github.com/Nextomics/NextDenovo), NextPolish, and Pilon, which was then merged into the draft genome obtained from Redbean to compensate missing regions based on comparisons with the reference genomes obtained from MarpolBase (Mptak1v5.1 and JGi3.1, https://marchantia.info). The resultant genome sequences were subjected to Hi-C scaffolding with the 3D-DNA pipeline. The in situ Hi-C library preparation of 2-week-old Tak-2 thalli, sequencing, and data processing were performed as described previously. Finally, the order and orientation of the contigs within scaffolds were manually corrected by comparing them to the v3.1 genome sex-chromosomes. Gene annotation was lifted-over from the reference genomes by aligning transcript sequences using GMAP ver. 2019.09.12 s. Full-length transcriptome sequences (Iso-seq) and RNA-seq reads were obtained from INSDC (BioProject IDs: PRJDB8530 and PRJNA251267), mapped to the newly obtained Tak-2 genome, and used as evidence for gene annotations. All U chromosome-linked genes were manually curated and spurious gene models that did not show any expression level or share homology to known functional domains were eliminated. New gene models were also identified based on evidence from Iso-seq and RNA-seq. The updated gene IDs for the sex chromosome-linked genes are listed in Data S1A.

Genetic nomenclature
BPCU and BPCV correspond to the locus codes, MpUg00370 (this study) and MpVg00350, respectively. Genetic nomenclature in M. polymorpha is as outlined previously. Prefix “Mp” in gene symbols was omitted for simplicity since all genes used in this study were derived from M. polymorpha. The gene symbols, BPCU and BPCV, were renamed from BPC1 and BPC2, respectively, from the reported genome analysis to clearly designate their location on specific sex chromosomes. Locus codes and gene symbols are also available through MarpolBase (https://marchantia.info).
Microscopy
The stalk of the developing sex organs (archegoniophores or antheridiophores) was cut off, placed on agar medium, and the entire receptacles were photographed with a stereomicroscope SZX16 (OLYMPUS, Tokyo, Japan). The inner gamete-containing organs were removed using a needle on a glass slide and photographed with a microscope Axiophot (ZEISS, Oberkochen, Germany).

RNA extraction and reverse-transcription
Total RNA was extracted from stages 1–2 sex organs or 10-day-old thalli with TRIzol reagent (Thermo Fisher Scientific) according to the manufacture’s instruction. DNA in the solution was digested using RQ1 RNase-Free DNase (Promega, Madison, USA) by incubation at 37 °C for 15 min. The quantity and quality of extracted RNA were assessed using a NanoDrop 2000 spectrophotometer (Thermo Scientific). 500 ng of total RNA was reverse-transcribed to synthesize the first-strand cDNA using ReverTra Ace (Toyobo Life Science) with the above constructs.

Plasmid construction for expressing BPCU and BPCV
Plasmids used in this study were constructed using Gateway cloning system (Thermo Fisher Scientific, Waltham, MA) for M. polymorpha, as described below. Primers used for PCR amplification are listed in Table S1. The constructs were verified by Sanger sequencing.

To create the vectors for expressing BPCU and BPCV in plants, the gene sequences of BPCU (2-kb upstream to 3-kb downstream of the transcribed region) and BPCV (1.5-kb upstream to 2-kb downstream of the transcribed region) were amplified from Tak-2 and Tak-1 genomic DNAs, respectively, using KOD FX Neo DNA polymerase (Toyobo Life Science, Osaka, Japan), and cloned into pENTR/D-TOPO (Thermo Fisher Scientific). The resulting entry vectors (BPCU-gene-pENTRdTOPO and BPCV-gene-pENTRdTOPO) were recombined with pMpGW101 and pMpGW301 using Gateway LR clonase II Enzyme mix (Thermo Fisher Scientific) to create BPCU-gene-pET-DEST42 and BPCV-gene-pET-DEST42. These vectors were introduced into BL21-RIL cells for the protein expression described below.

Genome editing
To create mutants for BPCU (bpcU-101, bpcU-102, and bpcU-3), BPCV (bpcV-1, bpcV-2 and bpcV-3) and FGMYB (fgmyb-31 and fgmyb-41), oligonucleotides listed in Table S1 were used as guide RNA (gRNA) sequences for CRISPR/Cas9-mediated genome editing with pMpGE_En03 (Addgene 71353), and pMpGE010 (Addgene 71536) or pMpGE011 (Addgene 71537). To generate large deletion mutants for BPCU (bpcU-1 and bpcU-2), four pairs of oligonucleotides (Table S1) were cloned as gRNA sequences into pMpGE_En04, pBc-GE12, pBc-GE23, and pBc-GE34, and further combined to generate an entry clone having four gRNAs. The resultant gRNA sequences were subcloned into pMpGE017 by the Gateway LR reaction as described previously. For the initial Feminizer screening, the gene-edited mutants were generated via transformation of F1 spores from a cross between Tak-2 and Tak-1 with the above constructs. bpcU mutants (bpcU-1 and bpcU-2) and bpcV mutants (bpcV-1, bpcV-2, and bpcV-3) were generated in Tak-2 and Tak-1, respectively, and bpcU-3 in the BC3-38 background, via thallus transformation.

Genotyping and sex diagnosis
Genomic DNA was extracted from a small piece of thallus (3 mm x 3 mm) crushed with micropestle in 100 μL of genotyping buffer containing 100 mM Tris-HCl, pH 9.5, 1M KCl, and 100 mM EDTA as described previously, and used for PCR using KOD FX Neo DNA polymerase. For gene-edited mutant identification, gRNA-targeted regions were amplified from the genomic DNA of transformants, and sequenced using the primer pairs listed in Table S1. Sex diagnosis was performed using the primer sets for rhf73 and rbm27 that are specific to sex chromosomes as described previously.

Generation of transgenic lines expressing genomic BPCU and genomic BPCV
To create the transgenic plants carrying the BPCU and BPCV transgenes, thallus transformation was used to introduce BPCU gene pMpGW101 and BPCU gene pMpGW301 into Tak-1 and bpcU-1, respectively, and also BPCV gene pMpGW301 into bpcU-1. To confirm genomic integration of the transgenes, G1 gametophytes were used for PCR amplification using KOD One PCR Master Mix and primers listed in Table S1, and progeny of G1 plants were used for the experiments. More than 10 lines were independently isolated for each construct, and the data from multiple representative lines were shown.

RT-PCR
RT-PCR was performed in 10 μL reaction solution, containing KOD One PCR Master Mix, 1 μL diluted cDNA and 0.3 μL each of 10 μM forward and reverse primers listed in Table S1 PRM (Mp3g14390) and LC7 (Mp5g01560) were amplified under the following cycle conditions: initial incubation at 94 °C for 2 min, followed by 37 cycles of 98 °C for 10 s, 58 °C for 5 s and then 68 °C for 1 s. Mp5g13880 was amplified under the following cycle conditions: initial incubation at 94 °C for 2 min, followed by 35 cycles of 98 °C.
for 10 s, 60 °C for 5 s and then 68 °C for 1 s. Mp7g02820 was amplified under the following cycle conditions: initial incubation at 94°C for 2 min, followed by 35 cycles of 98°C for 10 s, 58°C for 5 s and then 68°C for 1 s. EF1 (Mp3g23400) was amplified under the following cycle conditions: initial incubation at 94°C for 2 min, followed by 31 cycles of 98°C for 10 s, 58°C for 5 s and then 68°C for 1 s.

### Quantitative real-time RT-PCR

Quantitative real-time RT-PCR (qRT-PCR) was performed in 25 μL reaction solution, containing homemade Taq DNA polymerase, SYBR Green I Nucleic Acid Gel Stain (Lonza, Basel, Switzerland), 5 μL diluted cDNA, and 1 μL each of 10 μM forward and reverse primers listed in Table S1. Reactions were carried out with a CFX96 real-time PCR detection system (Bio–Rad Laboratories, Hercules, CA), under the following cycle conditions: initial incubation at 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, and 60°C for 30 s. Three technical and three biological replicates were measured for each reaction. EF1 (Mp3g23400) was used as internal control. Relative expression levels were calculated by the 2−ΔΔCT method.

### DPI-ELISA with BPCU and selected DNA-probes

E. coli BL21-RIL cells harboring BPCU-cds-pET-DEST42 or BPCV-cds-pET-DEST42 were used for the protein expression. Crude protein extraction was performed with HEPES buffer (4 mM HEPES pH 7.5, 100 mM KCl, 8% (v/v) glycerol, 0.2% biotin-free BSA, 5 mM dithiothreitol (DTT), supplemented with 1x Complete Protease Inhibitor without EDTA (Roche, Basel, Switzerland). Protein extracts from non-transformed E. coli BL21-RIL cells served as controls. Semiquantitative DPI-ELISA experiments were performed with 5 μg of crude protein extracts (in 30 μl buffered solution) and biotinylated double-stranded (ds) DNA-probes. Promoter probes were PCR amplified by using biotinylated (sense) and non-biotinylated (antisense) primer combinations. Residual biotinylated primers were removed with a NucleoSpin Gel and PCR Clean-up Kit (Macherey & Nagel, Duren, Germany). DPI-ELISA with purified promoter probes was performed according to the general protocol. The use of streptavidin-coated ELISA plates (5 pmol/well) allows for the semiquantitative and comparable readout within the same protein extract. Pre-blocked clear flat-bottom Streptavidin Coated Plates (96-wells) (Pierce Thermo Fisher Scientific) were used with Saffire II multiplate reader (Tecan, Mannedorf, Switzerland) for photometric measurements. Detection of 6xHis-epitope tagged BPCU was carried out with either Penta-HRP (OPD) as a substrate. The HRP reaction was stopped in all wells by addition of sulfuric acid (2 N H2SO4) as soon as a visible coloration appeared. Average signal intensities and standard deviation were computed from two technical replicates on the same plate and from two independent protein extracts. Relative signal intensities were calculated relative to the highest photometric absorbance of a reference protein-DNA combination.

### Flow cytometry analysis

Flow cytometry was performed according as previously described with slight modifications. One hundred mg of 9-day-old M. polymorpha thalli and 37-day-old A. thaliana (Col-0) rosette leaves were chopped in 1.5 mL of Gaibrath Buffer (45 mM MgCl2, 30 mM Trisodium citrate dihydrate, 20 mM MOPS, pH 7.0) supplemented with 0.1% v/v Triton X-100 using a single edge blade (AccuTec Blades, Verona, USA) for 5 min on ice, and consecutively passed through 40 μm and 10 μm cell strainers (SYSMEX, Kobe, Japan). RNaseA (Sigma-Aldrich, St. Louis, USA) was added to a final concentration of 50 μg ml−1 and incubated on ice for 15 min, then propidium iodide (Nacalai Tesque, Kyoto, Japan) was added to the same final concentration and incubated on ice for 30 min. The prepared samples were analyzed by a BD Accuri C6 Flow Cytometer (BD Biosciences, Franklin Lakes, USA) for 300,000 events each.

### Chromatin profiling analysis

The CUT&RUN (Cleavage Under Targets & Release Using Nuclease) assays were performed as previously described with slight modifications. DAPI stained 40,000 nuclei were sorted by BD FACSARIA III (BD Biosciences), 10 μL of BioMagPlus Concanavalin A bead slurry (Polysciences, #86057) was washed twice in 700 μL of binding buffer (20 mM HEPES pH 7.9, 10 mM KCl, 1 mM CaCl2 and 1 mM MnCl2) and resuspended in 10 μL of binding buffer, and added to the nuclear suspension. The mixture was thoroughly mixed on a rotator for 10 min at room temperature to allow binding of nuclei to the beads. The nuclei were pelleted with a magnet stand and resuspended with 50 μL of antibody buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM spermidine, 1 x protease inhibitor cocktail EDTA free (Roche) and 2 mM EDTA pH8.0). 0.5 μg of antibody was thoroughly mixed and incubated overnight at 4°C. Nuclei were washed twice on a magnet stand with 1 mL of wash buffer and incubated with 700 μg ml−1 of Protein A–micrococcal nuclease fusion protein (pA-MN) in 50 μL of wash buffer at room temperature. After 10 minutes, the mixture was washed twice with 1 mL wash buffer to remove unbound pA-MN. Nuclei were resuspended in 150 μL of wash buffer and chilled on ice at 0°C. 3 μL of 100 mM CaCl2 was added to activate pA-MN with incubation at 0°C for 120 min. The reaction was stopped by the addition of 100 μL of 2 x STOP buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 50 μg ml−1 RNase A, 50 μg ml−1 glycogen and MNase-digested spike-in DNA (HEK293 cells)) and incubated in thermo-mixier at 37°C for 10 minutes at 500 rpm. The protein–DNA complex was released by centrifugation and then digested by 0.05 μg proteinase K and 0.1% SDS at 70°C for 10 min. DNA was extracted by Phenol: Chloroform: Isoamyl-alcohol (25:24:1) with phase lock gel tubes. Library preparation of CUT&RUN was performed using a NEB Ultra II library preparation kit. 150-500 bp of DNA was selected by SPRI beads. DNA amount and fragment size were analyzed.
by Nanodrop 3300 and Fragment analyzer respectively. Libraries were pooled at similar molar amounts and sequenced using the Nextseq 500 platform. Paired-end sequencing was performed (read length, 42 bp \( \times 2 \); index, 6 bp).

**Analysis of sex chromosome evolution**

Predicted orthologs of *M. polymorpha* gametolog were collected from available liverwort genome and transcriptome sequence data.\(^8^2\)–\(^8^6\) Since the available liverwort transcriptomes were not generated to investigate sex chromosome evolution, we initially applied two filters. First, sequences derived from monoocious species were removed as monoicy is associated with a loss of sex chromosomes. Second, for analyses of gametologs other than BPC, we included only dioicous liverwort species for which only a single sex appeared to be sampled, because the transcriptome data appeared to be derived from a single sex (based on the presence of a single gametolog for each gene that we inferred to be in the oldest evolutionary stratum). An exception was the Haplocomitriopsida species, where we included all identified sequences since these sequences were critical for identifying the oldest evolutionary stratum. For the phylogenetic analysis of BPC orthologs, all identified sequences from dioicous liverwort species were included. Moss and hornwort orthologs were also included to ascertain whether any non-recombining regions pre-date the origin of liverworts, and vascular plant (lycophyte and fern) orthologs were used to root the trees.\(^8^3\),\(^8^6\)

For phylogenetic analyses of the gametologs (Figure 4; Data S2), all site types in coding regions (including non-synonymous as well as synonymous sites) were used. 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/). Poorly alignable sequences were removed, and alignments of nucleotides were employed in subsequent Bayesian analysis using MrBayes 3.2.1.\(^6^6\),\(^6^7\) The analysis was run for 1,000,000 generations (except for BPCU/MpUg00370 which was run for 10,000,000 generations), which was sufficient for convergence of the two simultaneous runs (split frequencies < 0.05). To allow a burn-in phase, the first 50% of the trees generated 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.

The proportions of synonymous substitutions per synonymous site (Ks values) were calculated with SNAP v2.1.1 (https://hcv.lanl.gov/content/sequence/SNAP/SNAP.html) as described previously.\(^6^8\),\(^8^7\) Ks values for the 19 pairs of *M. polymorpha* gametologs were previously reported.\(^1^9\) To relate these values to the liverwort phylogeny, we aligned sequences of eleven autosomal orthologs of liverwort genes across the same phylogenetic groups as used for studying gametologs, and estimated Ks values separately for each gene, with the standard deviation calculated as a measure of variation in values of different genes at each phylogenetic node (Data S1B).

**Data analysis of CUT&RUN**

CUT&RUN reads were preprocessed to filter out low-quality reads with Fastp v0.20.1,\(^5^4\) and mapped to the MpTakv6 genome presented in this paper using Bowtie2 v2.4.1\(^1^6\) with the “–very –sensitive” flag. Reads with MapQ less than 10 were removed with Samtools v1.9.0,\(^5^8\) and duplicates or inserts less than 140 bp reads were removed with alignmentSieve in Deeptools v3.5.0.\(^5^9\) Deduplicated reads from biological replicates were merged. We called peaks for chromatin marks by SICER2 with default parameters.\(^3^0\) Bigwig files were generated using “bamcoverage” in Deeptools\(^5^9\) with “–extendReads –exactScaling –binSize 10 –normalizeUsing RPGC” flags. The effective genome size was calculated from MpTakv6 with faCount.

**Transcriptome Mapping**

RNA-seq reads were preprocessed to filter out low-quality reads with Fastp v0.20.1,\(^5^4\) and mapped to the MpTakv6 genome presented in this paper using STAR v2.7.3a\(^5^7\) with default parameters. Bigwig files were generated using “bamcoverage" in Deeptools\(^5^9\) with “–binSize 10 –normalizeUsing BPM” flags.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

**General statistical analyses**

All statistical analyses were conducted in R (version 3.6.0). The specific tests, sample, size, and p value are displayed in the corresponding legends.

**Gene expression analysis of BPCU and BPCV**

Gene expression data from female sex organs: DRR050351, DRR050352, DRR050353,\(^1^9\) male sex organs: DRR050346, DRR050347, DRR050348,\(^4^9\) female thalli: DRR118943, DRR118944, DRR118945,\(^5^0\) male thalli: DRR118949, DRR118950, DRR118951,\(^5^0\) spores: SRR4450260, SRR4450261, SRR4450262,\(^1^9\) and sporelings: SRR4450254, SRR4450255, SRR4450256\(^1^9\) were downloaded from Sequence Read Archive. Transcript abundances were quantified by Salmon package v1.40 with validateMappings option.\(^5^0\) Transcripts Per Million values of transcripts were integrated into values corresponding to genome annotations by tximport package.\(^6^0\)
Supplemental Information

Identification of the sex-determining factor in the liverwort *Marchantia polymorpha* reveals unique evolution of sex chromosomes in a haploid system

Miyuki Iwasaki, Tomoaki Kajiwara, Yukiko Yasui, Yoshihiro Yoshitake, Motoki Miyazaki, Shogo Kawamura, Noriyuki Suetsugu, Ryuichi Nishihama, Shohei Yamaoka, Dierk Wanke, Kenji Hashimoto, Kazuyuki Kuchitsu, Sean A. Montgomery, Shilpi Singh, Yasuhiro Tanizawa, Masaru Yagura, Takako Mochizuki, Mika Sakamoto, Yasukazu Nakamura, Chang Liu, Frédéric Berger, Katsuyuki T. Yamato, John L. Bowman, and Takayuki Kohchi
Figure S1. Flow cytometry analysis of WT [U], WT [V], UV plant, \textit{bpcU-101} and \textit{bpcU-102}, Related to Figure 1.

Flow cytometry was performed for 9-day-old \textit{M. polymorpha} thalli and 37-day-old \textit{A. thaliana} (Col-0) rosette leaf. Histograms of fluorescence signals from a propidium iodide-stained nuclei are shown.
**Figure S2. GAGA-binding of BPCU and BPCV in vitro, Related to Figure 2.**

BPCU binds to a single GAGA motif *in vitro* based on different DNA-Protein Interaction (DPI)-ELISA studies with 6xHis-epitope tagged BPCU protein extracts and selected dsDNA-probes. (A) Binding of recombinant BPCU and BPCV to selected oligonucleotide probes *in vitro*. DNA-sequences of the double stranded oligonucleotide probes are displayed on the right. Only the biotinylated sense strands are given in 5’ to 3’-orientation. GAGA/TCTC tetranucleotides are colored red. Values represent relative signal intensities compared to the strongest signal [BPCU with K4-probe] (100%). Empty BL21-RIL cell extracts serve as a negative background control in all experiments. Light gray background shading indicates level of confidence for significant binding based on all BL21 control signals (*t*-test *p* < 0.05). (B) Schematic overview of the FGMYB/SUF locus, indicating the positions and sizes for biotinylated dsSUF promoter region probes. Gene organization of FGMYB and SUF (Top): gray and pink boxes indicate UTR and the coding region of FGMYB, respectively. Blue boxes indicate transcribed region of SUF. Distribution of cumulative 10-mer counts of GAGA-motifs (bottom). (C) BPCU binds to biotinylated SUF promoter region probes *in vitro*. DPI-ELISA experiment with indicated immobilized SUF promoter probes as well as respective mutated probes (SUFmut), in which most of the GAGA-motifs were replaced by GTGT. Values represent relative signal intensities compared to the strongest signal [BPCU with Prom-SUF^{2198}-probe] (100%).
Figure S3. Genetic characterization of gBPCU [V] plants harboring a fgmyb mutation, Related to Figure 2.

(A) Gene organization of FGMYB and SUF locus and schematic illustration of fgmyb mutations. The coding region of FGMYB is colored pink. Gray boxes indicate 5’- and 3’-untranslated regions. The SUF transcript is colored blue. Arrowheads indicate the positions of gRNAs used for generating fgmyb-31 and fgmyb-41. (B) Mutations in fgmyb. The nucleotide sequences of the wild-type (top) and the mutants (bottom) are aligned. The sequences for gRNAs are underlined. The PAM sequences for CRISPR/Cas9 are shown in bold. Deletions and insertions are shown in red. (C) Diagnosis of genetic sex using U chromosome-linked rhf73 and V chromosome-linked rbm27 markers. (D) Genotyping with BPCU-specific primers.
Figure S4. Comparison of the deduced amino acid sequences and expression analysis of *BPCU* and *BPCV*, Related to Figure 3.

(A) The deduced amino acid sequences of *BPCU* and *BPCV* are aligned. Conserved amino acids are highlighted. (B) Expression of *BPCU* and *BPCV* in various developmental stages. Expression profiles of *BPCU* and *BPCV* in thalli, sexual organs (gametangiophores containing gametangia), spores, and sporelings were quantified from public RNAseq data (see Methods). Y axis indicates Transcripts Per Million (TPM). Bars represent mean ± s. d. (n = 3).
Figure S5. Mutations in *bpcU* and *bpcV*, related to Figure 3.

The nucleotide sequences of the wild-type and the mutants are aligned. The sequences of the gRNA are underlined. The PAM sequences for CRISPR/Cas9 are shown in bold. Deletions and insertions are shown in red. *bpcU* (A) and *bpcV* (B).
| Primer name | Sequences 5’>3’ | References | Notes |
|-------------|-----------------|------------|-------|
| **For genome editing** (bpcU) in Figure 1, Figure 3, and Figure S4 |
| X18-11.1 gRNA F | CTCGGAGCTCTCGTCCTAG | this study | bpcU-101, bpcU-102, bpcU-3 |
| X18-11.1 gRNA R | AAAATCGCTCTAGAGAACAGA | this study | bpcU-101, bpcU-102, bpcU-3 |
| X18-11.1-Ds gRNA F | CTCGGAGCTCTCGTCCTAG | this study | bpcU-101, bpcU-102, bpcU-3 |
| X18-11.1-Ds gRNA R | AAAATCGCTCTAGAGAACAGA | this study | bpcU-101, bpcU-102, bpcU-3 |
| **For genotyping sequencing** (bpcU-1) and (bpcU-2) |
| YB-17.1 gRNA F | CTCGGAGCTCTCGTCCTAG | this study | bpcU-1, bpcU-2 |
| YB-17.1 gRNA R | AAAATCGCTCTAGAGAACAGA | this study | bpcU-1, bpcU-2 |
| YB-17.2 gRNA F | CTCGGAGCTCTCGTCCTAG | this study | bpcU-1, bpcU-2 |
| YB-17.2 gRNA R | AAAATCGCTCTAGAGAACAGA | this study | bpcU-1, bpcU-2 |
| **For genome editing (mpfgmyb)** in Figure 1 and Figure S4 |
| MpFGMYB gRNA 3F | CTCGGAGCTCTCGTCCTAG | this study | mpfgmyb-3 |
| MpFGMYB gRNA 3R | AAAGGACGCTCTAGAGAACAGA | this study | mpfgmyb-3 |
| MpFGMYB gRNA 4F | CTCGGAGCTCTCGTCCTAG | this study | mpfgmyb-4 |
| MpFGMYB gRNA 4R | AAAGGACGCTCTAGAGAACAGA | this study | mpfgmyb-4 |
| **For sex diagnosis PCR in Figure 1G and Figure S3** |
| ge02-Fw | ATCTGGGAGCTCTACCTG | this study | ge02-Fw |
| ge02-Rv | TTCTGGAGCTCTACCTG | this study | ge02-Rv |
| **For RT-PCR in Figure 2A** |
| Mpfg101fW | CTCGGAGCTCTACCTG | this study | EF1 |
| Mpfg101rW | CTAGTCCGGAGCTCTACCTG | this study | EF1 |
| Mpfg101rW | CTAGTCCGGAGCTCTACCTG | this study | EF1 |
| **For qRT-PCR of** FGMYB | MpFGMYB-qPCR-R11 | this study | FGMYB |
| **For genomic construct in Figure 1H and Figure 3** |
| bpcU-1, bpcU-2 | | this study | bpcU-1, bpcU-2 |
| bpcU-1, bpcU-2 | | this study | bpcU-1, bpcU-2 |
| bpcU-1, bpcU-2 | | this study | bpcU-1, bpcU-2 |
| bpcU-1, bpcU-2 | | this study | bpcU-1, bpcU-2 |
| bpcU-1, bpcU-2 | | this study | bpcU-1, bpcU-2 |
| **For qRT-PCR of** FGMYB | MpFGMYB-qPCR-R11 | this study | FGMYB |
| **For genomic construct in Figure 3** |
| YB-17 genehead+2 F | CTCGGAGCTCTACCTG | this study | gPCV |
| YB-17 genehead+2 R | CTCGGAGCTCTACCTG | this study | gPCV |
| **For expression constructs** of BPCU and BPCV in Figure 2 and transgene genotyping of BPCU and BPCV in Figure 1, Figure 3, and Figure S3 |
| X18-11 cdS F2 | CTCGGAGCTCTACCTG | this study | gPCV |
| X18-11 cdS R2 | CTCGGAGCTCTACCTG | this study | gPCV |
| **For BPCU genomic construct in Figure 1H and Figure 3** |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| **For BPCV genomic construct in Figure 3** |
| BPCV-1, BPCV-2 | | this study | BPCV |
| BPCV-1, BPCV-2 | | this study | BPCV |
| BPCV-1, BPCV-2 | | this study | BPCV |
| BPCV-1, BPCV-2 | | this study | BPCV |
| **For BPCU genomic construct in Figure 1H and Figure 3** |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| **For expression constructs** of BPCU and BPCV in Figure 2 and transgene genotyping of BPCU and BPCV in Figure 1, Figure 3, and Figure S3 |
| X18-11 cdS F2 | CTCGGAGCTCTACCTG | this study | gPCV |
| X18-11 cdS R2 | CTCGGAGCTCTACCTG | this study | gPCV |
| **For BPCU genomic construct in Figure 3** |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| **For BPCU genomic construct in Figure 3** |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
| BPCU-1, BPCU-2 | | this study | BPCU |
Sufficient-S4 GTGGCCTGTATTGGGCTTCAAGTGC this study Promoter SUF/SUF mut e.g for DPI-ELISA

| Genome Editing | gRNA F | gRNA R | Study Details |
|---------------|--------|--------|--------------|
| X17-09.1      | CTCGGACTCCGGTCGACAAACC | this study | Mapoly0017s0009, MpUg00080 |
| X17-09.1      | AAACGGTTGCTGACCGGGAGGC | this study | Mapoly0017s0009, MpUg00080 |
| X18-01.1      | CTCGAGCTGTGATTGGGCTTCAAGTGC | this study | Mapoly0018s0001, MpUg000290 |
| X18-01.1      | AAACGTTGCTGACCAATGAGC | this study | Mapoly0018s0001, MpUg000290 |
| X18-11.1      | CTCGAGGAGCTCCGGTGACT | this study | BPCU, Mapoly0018s0001, MpUg000370 |
| X18-11.1      | AAACCTGAGCAAGGTGCTGACT | this study | BPCU, Mapoly0018s0001, MpUg000370 |
| X18-04.1      | CTCGAGGAGCTCTCTGAGCT | this study | Mapoly0018s0004, MpUg000330 |
| X18-04.1      | AAACATGGTACGAGGAACCTG | this study | Mapoly0018s0004, MpUg000330 |
| X17-03.1      | CTCGAGCCTAATGTGACCAAGC | this study | Mapoly0017s0003, MpUg00020 |
| X17-03.1      | AAACGTTGACAGTGGCTGACT | this study | Mapoly0017s0003, MpUg00020 |
| X17-20.1      | CTCGCTGTGCTTGAGCTGACT | this study | Mapoly0017s0002, MpUg00100 |
| X18-21.1      | AAACACTAAGAGCTTACCAGA | this study | Mapoly0018s0021, MpUg000440 |
| X18-21.1      | CTCGAATTCCCTCCTAGCAT | this study | Mapoly0018s0021, MpUg000440 |

Table S1. List of oligonucleotides used in this study, related to STAR Methods.
Supplemental References:

S1. Brand, L.H., Kirchler, T., Hummel, S., Chaban, C., and Wanke, D. (2010). DPI-ELISA: a fast and versatile method to specify the binding of plant transcription factors to DNA in vitro. Plant Methods 6, 25.

S2. Theune, M.L., Bloss, U., Brand, L.H., Ladwig, F., and Wanke, D. (2019). Phylogenetic analyses and GAGA-motif binding studies of BBR/BPC proteins lend to clues in GAGA-motif recognition and a regulatory role in brassinosteroid signaling. Front. Plant Sci. 10, 466.

S3. Hisanaga, T., Okahashi, K., Yamaoka, S., Kajiwara, T., Nishihama, R., Shimamura, M., Yamato, K.T., Bowman, J.L., Kohchi, T., and Nakajima, K. (2019). A cis-acting bidirectional transcription switch controls sexual dimorphism in the liverwort. EMBO J. 38, 663-669.

S4. Fujisawa, M., Hayashi, K., Nishio, T., Bando, T., Okada, S., Yamato, K.T., Fukuzawa, H., and Ohyama, K. (2001). Isolation of X and Y chromosome-specific DNA markers from a liverwort, Marchantia polymorpha, by representational difference analysis. Genetics 159, 981-985.