

A *cis*-acting bidirectional transcription switch controls sexual dimorphism in the liverwort

Tetsuya Hisanaga, Keitaro Okahashi, Shohei Yamaoka, Tomoaki Kajiwara, Ryuichi Nishihama, Masaki Shimamura, Katsuyuki T. Yamato, John L. Bowman, Takayuki Kohchi and Keiji Nakajima.

Review timeline:

Submission date:	11 th July 2018
Editorial Decision:	13 th August 2018
Revision received:	17 th November 2018
Accepted:	26 th November 2018

Editor: Anne Nielsen

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision

13th August 2018

Thank you for submitting your manuscript for consideration by the EMBO Journal. It has now been seen by three referees whose comments are shown below.

As you will see from the reports, all three referees express interest in the findings reported in your manuscript although they also raise a number of concerns that you will have to address before they can support publication of the manuscript here. Most importantly, refs #2 and #3 ask for additional characterization of the SUF RNA transcript structure, expression pattern and mechanism of action. I understand that a full mechanistic delineation is beyond the scope of the current study but I would strongly suggest you to follow ref #2's recommendation to test for a *cis* vs *trans* effect. Regarding point 4 from this reviewer, I would encourage you to run the computational analysis suggested but establishing evidence for conservation of this regulatory mechanism will not be a requirement for publication here. Ref #3 raises similar concerns as ref #2 and in addition stresses that the RNAseq data will need to be provided (preferably with an accession number available to the referees as also described in our author guidelines) and that the phylogenetic analysis for MpFGMYB should be extended.

Given the referees' positive recommendations, I would like to invite you to submit a revised version of the manuscript, addressing the comments of all three reviewers.

REFeree REPORTS.

Referee #1:

This manuscript describes the function of a *Marchantia polymorpha* locus that controls sexual development. The MpFGMYB transcription factor promotes female development, and an antisense transcript expressed, MpSUF, in males represses the expression of the MYB transcription factor.

The most important data in the manuscript are:

MpFGMYB is expressed in developing archaegoniophores (female, egg-producing structures) and not detectable in antheridiophores (male sperm producing structures). Genetically female *Mpfgmyb* loss of function mutants develop as males and produce sperm. However, these sperms are morphologically defective. This suggests that genes on the Y chromosome are required for sperm development.

MpSUF is expressed in males. MpSUF is an antisense MpFGMYB transcript and promotes maleness in male plants. Genetically male *Mpsuf* mutants are feminized; they produced archaegoniophores.

The arabidopsis genes that are most similar to MpFGMYB are also involved in female gametophyte (megagametophyte) development. This suggests that the same molecular mechanism controls the development of female gametophyte structures in these diverged groups of plants. This suggests that the FGMYB functioned in the last common ancestor of Arabidopsis and Marchantia.

This is a complete study and all relevant data are presented. I have no additions to suggest.

Referee #2:

This manuscript by Hisanaga et al describes the identification of a conserved MYB transcription factor (FGMYB) in the liverwort *Marchantia* that controls female gametophyte development. Loss of this gene results in female-to-male sex conversion. The gene is evolutionarily conserved also in higher plants. The authors go on to show that expression of FGMYB is regulated by a long-non-coding antisense RNA called SUF, which is upregulated in male gametophytes. Loss of SUF function results in a male-to-female sex conversion. The manuscript is very clearly written and the data support well the conclusions. Functional genomic analysis of developmental pathways in lower plants such as liverworts is a very interesting area that will provide important insights into the evolutionary dynamics of plant development. This study provides significant advances to our understanding of female gametophyte determination. To further strengthen the manuscript, I have a few suggestions for improvement.

1) The mutually exclusive transcription of FGMYB and SUF provides a tentative mechanism that is very appealing. However, besides the transcriptional analysis and the SUF deletion the manuscript lacks mechanistic insight into how this is happening. It would strengthen the manuscript considerably to dig a little deeper into this. With the current evidence, it is possible that either SUF RNA inhibits FGMYB transcription (action in trans) or that transcription of the SUF locus per se inhibits FGMYB transcription (action in cis). As a strain with a deletion of SUF is already available, it should be relatively easy to test whether supplying SUF RNA in trans (from a transgene) can complement the *suf* deletion phenotype. This would provide important evidence whether SUF acts in cis or in trans.

2) It would also be helpful to characterize the SUF transcript a bit more to confirm and complement the RNA-seq data. Where does it start, where does it end, does it have introns etc.? The RT-PCR confirmation provided for SUF (Fig. 5b and Fig. S6) is confusing as one Figure shows a double band and the other does not. Could the authors explain this double band and provide RT-PCR evidence for the full-length SUF transcript?

3) The authors have previously described RKD as a crucial regulator of egg cell maturation in *Marchantia* (Koi, 2016). As RKD appears as a possible target of FGMYB, the authors should check expression of RKD in *fgmyb* and *suf* mutant strains.

4) Given that FGMYB is conserved in higher plants such as *A. thaliana*, it would be interesting to see whether the regulation through an antisense RNA is also likely conserved. Is there any evidence that any of the Arabidopsis orthologues also have a SUF-like antisense RNA (not necessarily related in sequence)? Addressing this (even if only by carefully analysing available and published transcriptome data) would strengthen the role of SUF in the regulation of FGMYB.

Small comments:

- p. 9 "shorter nuclei" - unclear what is meant by this. Please clarify.

- Fig. 4a-c: Please clarify what can be seen here by indicating the nucleus etc.
- Fig. 5a: The scales of the reads of sense and antisense are vastly different (20 vs. 8), yet the font of the scale is so small, that it is hard to appreciate this. While it may not be practical to show everything at the same scale, a simple solution would be to make the y axis labelling larger so that this difference could be seen more clearly.

Referee #3:

The Authors have identified a sex determining factor in *Marchantia*, the new model to study evolution of land plants. Comparing transcriptomes between vegetative haploid tissues and the female reproductive organs (archegonia), a MYB TF is identified, which is named FEMALE GAMETOHYTIC SPECIFIC MYB (MpFGMYB). This TF belongs to a small clade of MYB genes. The Arabidopsis genes from this clade are also expressed in the female gametophyte of *At*. KO mutant of this gene in female plants results in a dramatic conversion of sexual organ development into male organs - antheridiophores-. This conversion is marked by the change of morphology and the striking development of male gametes instead of female gametes on female mutant KO plants. The phenotype is complemented by expression of WT as well as the GFP- fusion protein used as reporter. However, sperm showed defect in flagella development and motility that could be the result of the lack of genes on the Y chromosome that is obviously absent from these mutant female plants. The study proceeds with demonstration that MpFGMYB expression is suppressed by an antisense SUPPRESSOR OF FEMINIZATION (SUF) that is expressed only in male gametophytes after induction of sexual reproduction by FR. The KO of SUF in male results in production of female looking reproductive organs that differentiate archegoniophores. Deletion of the entire locus (absence of both sense MpFGMYB and SUF) resulted in WT-looking male reproductive organs with fertile sperm. Overall experiments are conducted correctly and the results obtained are unambiguous. These first striking findings open the path to our understanding of the origin of sexual differentiation in land plants. In addition the study provides a valuable model in addition to FLC in angiosperms to study the bases of transcriptional regulation by antisense lnc RNAs. As such the results of this work are of interest to a broad readership.

Major points

1. study is based on the analysis of transcriptomes of archegonia and thalli which is not provided by the Authors. It is essential and conditional to publication that this dataset is shared in a public database and Table of the results provided in supplementary materials.
2. The specific expression of MpFGMYB and SUF are the cornerstone of this study. So the Authors provide data more extensive and convincing than a simple RT-PCR. For example a Quantitative RTPCR survey of expression of MpFGMYB and SUF across different stages of vegetative and reproductive development of male and female gametophytes and in the sporophyte. This should also include different stages of antheridiophore and archegoniophore development after induction by FR light, including times when the organogenesis has not started. Given the function of the gene, one expects an expression that pre-dates the morphological development of the archegoniophore.
3. The authors highlight a conserved domain present in MpFGMYB and its putative orthologs in the moss *Physcomitrella*. Could the Authors provide a more extensive search for potential orthologs amongst bryophytes to ascertain the nature of this motif. At least the Authors should provide a basal study of expression of the potential orthologs of MpFGMYB in *physcomitrella* that does not show reproductive organ differentiation but only production of antheridia and archegonia on the same plant. Could the Authors tell us at least if an ortholog of SUF exists in *Physcomitrella*?

Minor points

1. According to the phylogenetic analysis there are four other MYB genes in *At* in the same clade as MpFGMYB. Could the authors comment on their expression and if they are not expressed in the female gametophyte, what could differentiate these genes?
2. In the discussion the Authors point out that RKD might be regulated by MpFGMYB. Again, either the Authors attempt to obtain data related to this question and discuss the result. This would have been quite easy. Or they can remove this point of discussion.

3. In the discussion there are statement regarding potential parallels in the role of putative orthologs of MpFGMYBs in At. These should be toned down very much or rather removed because there is no data besides the expression of these putative orthologs in the embryo sac, to propose a potential orthology. The At protein does not show the same N-terminal domain conserved amongst bryophyte. If the Authors wish to discuss this point they should attempt cross-complementation of Mp mutant with the At gene. Or at least an engineered fusion between the At MYB domain replacing the Mp MYB domain of MpFGMYB.
4. In the Introduction the concept of definition of sexual organ identity is confused with the definition of gamete differentiation at several points. This should not be the case and the Authors should rigorously compare what is truly comparable between angiosperms and bryophytes.
5. In the abstract the statement that bidirectional transcription is the key mechanism should be toned down because it is not shown in this study whether the Antisense SUF plays a role directly in the repression of MpFGMYB.
6. The last sentence of the Abstract is not substantiated by findings presented here and should be modified.

1st Revision - authors' response

17th November 2018

Response to the referees' comments

Referee #1:

This manuscript describes the function of a *Marchantia polymorpha* locus that controls sexual development. The MpFGMYB transcription factor promotes female development, and an antisense transcript expressed, MpSUF, in males represses the expression of the MYB transcription factor.

The most important data in the manuscript are:

MpFGMYB is expressed in developing archaegoniophores (female, egg-producing structures) and not detectable in antheridiophores (male sperm producing structures). Genetically female *Mpfgmyb* loss of function mutants develop as males and produce sperm. However, these sperms are morphologically defective. This suggests that genes on the Y chromosome are required for sperm development.

MpSUF is expressed in males. MpSUF is an antisense MpFGMYB transcript and promotes maleness in male plants. Genetically male *Mpsuf* mutants are feminized; they produced archaegoniophores.

The arabidopsis genes that are most similar to MpFGMYB are also involved in female gametophyte (megagametophyte) development. This suggests that the same molecular mechanism controls the development of female gametophyte structures in these diverged groups of plants. This suggests that the FGMYB functioned in the last common ancestor of Arabidopsis and Marchantia.

This is a complete study and all relevant data are presented. I have no additions to suggest.

[We thank the referee for his/her highly favorable comments on our study.](#)

Referee #2:

This manuscript by Hisanaga et al describes the identification of a conserved MYB transcription factor (FGMYB) in the liverwort *Marchantia* that controls female gametophyte development. Loss of this gene results in female-to-male sex conversion. The gene is evolutionarily conserved also in higher plants. The authors go on to show that expression of FGMYB is regulated by a long-non-coding antisense RNA called SUF,

which is upregulated in male gametophytes. Loss of *SUF* function results in a male-to-female sex conversion. The manuscript is very clearly written and the data support well the conclusions. Functional genomic analysis of developmental pathways in lower plants such as liverworts is a very interesting area that will provide important insights into the evolutionary dynamics of plant development. This study provides significant advances to our understanding of female gametophyte determination. To further strengthen the manuscript, I have a few suggestions for improvement.

We thank the referee for his/her constructive comments on our paper. Our answers to the suggestions are as follows.

1) The mutually exclusive transcription of *FGMYB* and *SUF* provides a tentative mechanism that is very appealing. However, besides the transcriptional analysis and the *SUF* deletion the manuscript lacks mechanistic insight into how this is happening. It would strengthen the manuscript considerably. With the current evidence, it is possible that either *SUF* RNA inhibits *FGMYB* transcription (action in trans) or that transcription of the *SUF* locus per se inhibits *FGMYB* transcription (action in cis). As a strain with a deletion of *SUF* is already available, it should be relatively easy to test whether supplying *SUF* RNA in trans (from a transgene) can complement the *suf* deletion phenotype. This would provide important evidence whether *SUF* acts in cis or in trans.

While the first version of this manuscript was under review, we obtained new data indicating a *cis*-acting mode of *SUF* on *MpFGMYB* suppression. We originally intended to use these data for separate publication in future, but now decided to incorporate them to the present paper. As suggested by the referee, we introduced *SUF* overexpression constructs into *suf* [Y] mutants and found no complementation despite high levels of *SUF* transcripts accumulation. Additionally, we introduced genomic fragment of the *MpFGMYB/SUF* locus with and without the putative *SUF* promoter and TSS (*gMpFGMYB-L* and *gMpFGMYB-S*, respectively) into WT [Y]. While *gMpFGMYB-S* with no predicted *SUF* expression could confer partial sex conversion to females, *gMpFGMYB-L* with predicted *SUF* expression could not. These data indicate the inability of transgenic *SUF* copies to suppress endogenous *MpFGMYB*, as well as of endogenous *SUF* to suppress transgenic *MpFGMYB* copies. Based on these observations, we concluded that *SUF* acts *in cis* to suppress *MpFGMYB* expression in males. We added these new findings as Fig 6 and described them in the last part of the Results section and also in the Discussion. The title for the paper has been changed to feature this important new finding.

2) It would also be helpful to characterize the *SUF* transcript a bit more to confirm and complement the RNA-seq data. Where does it start, where does it end, does it have introns etc.? The RT-PCR confirmation provided for *SUF* (Fig. 5b and Fig. S6) is confusing as one Figure shows a double band and the other does not. Could the authors explain this double band and provide RT-PCR evidence for the full-length *SUF* transcript?

We admit that characterization of *SUF* transcript structures is important, and thus performed 3' and 5' RACE PCRs for *SUF*. The data indicate that *SUF* transcripts have a well-defined structure with an invariable 5' end point and polyadenylation site, indicating that *SUF* represents a strictly defined transcription unit. We added these data in Appendix Fig S3. As for the double band, we apologize for the confusing figures. For RT-PCR in Fig 5b, we used a primer pair flanking the *SUF* intron, but the primers used in Fig S6a (currently Fig EV4B) do not. Since total RNA samples were treated with DNase I and the bands appeared specifically in wild-type males, we think the doubled band in Fig 5B to be derived from spliced (lower band) and unspliced (upper band) variants of *SUF* transcripts. We explained this in the legend to Fig 5. In vernalization-dependent suppression of *FLC*, unspliced *COOLAIR* lncRNA molecules are shown to form dense clouds at the locus from which they are produced leading to local change in epigenetic state (Rosa et al., Nat

Commun. 2016, 7: 13031). While it is hypothesized that unspliced *SUF* transcripts act similarly to unspliced *COOLAIR* lncRNA, it is still highly speculative and hence is not mentioned in the revised manuscript.

3) The authors have previously described RKD as a crucial regulator of egg cell maturation in *Marchantia* (Koi, 2016). As RKD appears as a possible target of GAMYB, the authors should check expression of RKD in *fgmyb* and *suf* mutant strains.

While we are interested in checking whether RKDs are transcriptional targets of FGMYPs, it would be difficult to test this possibility by simply examining expression of *MpRKD* in *Mpfgmyb* and *suf* mutants, because *MpRKD* expression is activated not only in female archegonia but also in male antheridia (Koi et al., 2016 Curr. Biol. 26: 1775-1781) likely by other regulators than *MpFGMYB*. In response to the referee #3's minor comment 2, we removed the part describing possible relationship between *MpFGMYB* and *MpRKD* from the Discussion.

4) Given that FGMYP is conserved in higher plants such as *A. thaliana*, it would be interesting to see whether the regulation through an antisense RNA is also likely conserved. Is there any evidence that any of the *Arabidopsis* orthologues also have a *SUF*-like antisense RNA (not necessarily related in sequence)? Addressing this (even if only by carefully analyzing available and published transcriptome data) would strengthen the role of *SUF* in the regulation of GAMYB.

We searched for publically available RNA-seq datasets from *Arabidopsis* pollen (Qin et al., 2018, F1000Res. 7: 93, doi:[10.12688/f1000research.13311.1]) and various tissues of *Physcomitrella patens* (Perroud et al., 2018, Plant J. 95: 168-182; Koshimizu et al, 2018, Nat Plants 4: 36–45). In the *Arabidopsis* pollen dataset, we did not detect antisense transcripts in none of the three *FGMYB* loci (*MYB64*, *MYB119* and *MYB98*), consistent with the lack of lncRNA-dependent regulation in *A. thaliana* male gametophytes discussed in the previous version (currently Fig 7C). We found that *FGMYB* homologues in *P. patens* are preferentially expressed in gametophores which include archegonia. We also detected a certain numbers of reads of the *FGMYBs* in antheridia. Although distribution patterns of reads along the loci were somewhat different between gametophores and antheridia, potentially suggesting accumulation of antisense lncRNAs, it is so far difficult to comment on possible occurrence of antisense transcription for *PpFGMYBs*, as these RNA-seq data are not from stranded libraries.

Small comments:

- p. 9 "shorter nuclei" - unclear what is meant by this. Please clarify.

We changed "shorter nuclei" to "incompletely condensed nuclei" (currently in p. 10).

- Fig. 4a-c: Please clarify what can be seen here by indicating the nucleus etc.

We now annotated the images with arrows and arrowheads which indicate flagella and nuclei, respectively.

- Fig. 5a: The scales of the reads of sense and antisense are vastly different (20 vs. 8), yet the font of the scale is so small, that it is hard to appreciate this. While it may not be practical to show everything at the same scale, a simple solution would be to make the y axis labelling larger so that this difference could be seen more clearly.

We now labeled the Y axis with larger fonts to present the data more clearly.

Referee #3:

The Authors have identified a sex determining factor in *Marchantia*, the new model to study evolution of land plants. Comparing transcriptomes between vegetative haploid tissues and the female reproductive organs (archegonia), a MYB TF is identified, which is named FEMALE GAMETOPHYTIC SPECIFIC MYB (MpFGMYB). This TF belongs to a small clade of MYB genes. The Arabidopsis genes from this clade are also expressed in the female gametophyte of *At*. KO mutant of this gene in female plants results in a dramatic conversion of sexual organ development into male organs - antheridiophores. This conversion is marked by the change of morphology and the striking development of male gametes instead of female gametes on female mutant KO plants. The phenotype is complemented by expression of WT as well as the GFP- fusion protein used as reporter. However, sperm showed defect in flagella development and motility that could be the result of the lack of genes on the Y chromosome that is obviously absent from these mutant female plants. The study proceeds with demonstration that MpFGMYB expression is suppressed by an antisense SUPPRESSOR OF FEMINIZATION (SUF) that is expressed only in male gametophytes after induction of sexual reproduction by FR. The KO of SUF in male results in production of female looking reproductive organs that differentiate archegoniophores. Deletion of the entire locus (absence of both sense MpFGMYB and SUF) resulted in WT-looking male reproductive organs with fertile sperm. Overall experiments are conducted correctly and the results obtained are unambiguous. These first striking findings open the path to our understanding of the origin of sexual differentiation in land plants. In addition the study provides a valuable model in addition to FLC in angiosperms to study the bases of transcriptional regulation by antisense lnc RNAs. As such the results of this work are of interest to a broad readership.

We thank the referee for his/her constructive comments on our paper. Our answers to the suggestions are as follows.

Major points

1. study is based on the analysis of transcriptomes of archegonia and thalli which is not provided by the Authors. It is essential and conditional to publication that this dataset is shared in a public database and provided in supplementary materials.

We thank the referee for pointing this out. We now deposited our transcriptome data to the DDBJ sequence Read Archive with an accession number of DRA006846. We added this information to the Materials and Methods section with a subheading Data availability. We also included the list of the 23 genes enriched in female gametophytes of both *M. polymorpha* and *A. thaliana* as Table EV1. When we were converting the transcript ID of the original dataset (CUFF.xxx) to the *Marchantia* gene ID (Mapolyxxxxsxxxx), we found that two of the selected transcripts are from an identical locus. Therefore, the number of selected genes has been reduced to 23.

2. The specific expression of MpFGMYB and SUF are the cornerstone of this study. So the Authors provide data more extensive and convincing than a simple RT-PCR. For example a Quantitative RT PRC survey of expression of MpFGMYB and SUF across different stages of vegetative and reproductive development of male and female gametophytes and in the sporophyte. This should also include different stages of antheridiophore and archegoniophore development after induction by FR light, including times when the organogenesis has not started. Given the function of the gene, one expects an expression that pre-dates the morphological development of the archegoniophore.

We appreciate the referee's constructive comments and carried out detailed expression analyses of MpFGMYB as shown in Fig 2D and Fig EV1 in the revised version. After confirming with real-time RT-PCR a steep rise of MpFGMYB expression levels upon induction of reproductive growth by FR irradiation (Fig. 2D), we used the

gMpFGMYBresist-Citrine (Fig EV1) to determine when and where the expression of *MpFGMYB* is initiated. Consistent with the predicted role of *MpFGMYB* in establishing female sexual morphologies, the expression of *MpFGMYB* was found to initiate around 10 days after the induction at the ventral side of the apical notch region (Fig EV1C) where archegoniophores later develop (Fig EV1D). As for *SUF* expression, our previous RT-PCR analyses (now shown in Fig EV4B) detected *SUF* transcripts both in vegetative thalli and antheridiophores. This was confirmed by real-time RT-PCR (Fig. EV4A). This analysis also revealed that antheridiophores exhibited considerably higher accumulation of *SUF* transcripts than thalli. Somewhat unexpectedly, *MpFGMYB* and *SUF* are coexpressed in the sporophytes. While our tentative interpretation for this is spatial or even temporal separation of *MpFGMYB* and *SUF* expression domains, this can not be resolved by currently available RT-PCR data obtained from bulked plant samples. We added the real-time RT-PCR data and corresponding discussion to the revised manuscript. As additional expression was detected in the sporophytes, name of the gene has been changed from "*FEMALE GAMETOPHYTE-SPECIFIC MYB*" to "*FEMALE GAMETOPHYTE MYB*".

3. The authors highlight a conserved domain present in *MpFGMYB* and its putative orthologs in the moss *Physcomitrella*. Could the Authors provide a more extensive search for potential orthologs amongst bryophytes to ascertain the nature of this motif. that does not show reproductive organ differentiation but only production of antheridia and archegonia on the same plant. Could the Authors tell us at least if an ortholog of *SUF* exists in *Physcomitrella*?

As we described in our answer to the referee #2's comment 4) above, we found that the *P. patens* *FGMYB* homologues are preferentially expressed in gametophores that include archegonia. We also detected a certain number of reads of *FGMYBs* in antheridia, and the distribution pattern of reads along the loci appeared different from those in gametophores. Although these differences suggest occurrence of antisense transcription at the *P. patens* *FGMYB* loci in antheridia, we think it still premature to mention this in the text, as the RNA-seq data are not derived from stranded libraries. As *P. patens* is monoicous and thus produces archegonia and antheridia in a close proximity in a single individual, more sophisticated experiments are required to unambiguously examine whether similar bidirectional transcription operates for spatiotemporal regulation of sexual morphologies in this bryophyte species. This will be done in future as a collaborative study with moss experts.

Minor points

1. According to the phylogenetic analysis there are four other MYB genes in *At* in the same clade as *MpFGMYB*. Could the authors comment on their expression and if they are not expressed in the female gametophyte, what could differentiate these genes?

We thank the referee for pointing out this important question. Among the four MYB genes, *MYB115* and *MYB118* are reported to be expressed in the endosperm and activate fatty acid synthesis, whereas *MYB22* and *MYB100* expression is generally quite low in various tissue types (Troncoso-Ponce et al., 2016 Plant Cell 28: 2666-2682). Thus at least some extant *FGMYB* members are expressed outside the female gametophyte, but rather function after fertilization. We speculate this to be a result of evolutionary cooption of gametophytic regulators to sporophytic functions, as has been proven the case for many developmental genes (see for example; Jang et al., 2011, Development 138: 2273-2281 and Kubota et al., 2014, Nat. Commun. 5: 3668), and also suggested for regulators of sexual reproduction by us (Koi et. al., 2016, Curr. Biol. 26: 1775-1781). We discussed this in the revised version.

2. In the discussion the Authors point out that RKD might be regulated by *MpFGMYB*. Again, either the Authors attempt to obtain data related to this question and discuss the result. This would have been quite easy. Or they can remove this point of discussion.

In response to this comment, we removed the corresponding paragraph from the Discussion.

3. In the discussion there are statement regarding potential parallels in the role of putative orthologs of MpFGMYBs in At. These should be toned down very much or rather removed because there is no data besides the expression of these putative orthologs in the embryo sac, to propose a potential orthology. The At protein does not show the same N-terminal domain conserved amongst bryophyte. If the Authors wish to discuss this point they should attempt cross-complementation of Mp mutant with the At gene. Or at least an engineered fusion between the At MYB domain replacing the Mp MYB domain of MpFGMYB.

We agree with the referee; we think regulators for female sexual differentiation in gametophytic phase have common origin between Mp and At, but function of these regulators including their targets could be diverse. To make this point clear, we added a text into the Discussion as follows "To what extent FGMYBs regulate conserved sets of target genes across the land plant lineage, however, remains an open question, as female gametophytes have been considerably diversified in morphology, while retaining their central function, egg cell production, in the course of land plant evolution." A known role of some homologs in sporophytes is also mentioned in the following part in response to the minor point #1.

4. In the Introduction the concept of definition of sexual organ identity is confused with the definition of gamete differentiation at several points. This should not be the case and the Authors should rigorously compare what is truly comparable between angiosperms and bryophytes.

We apologize for the confusion between sexual organ identity and gamete differentiation in the Introduction. What we report here is regulation of gametophytic sexual differentiation processes in the liverwort and its comparison with the angiosperm counterpart, in particular developmental regulation of the female gametophytes, the embryo sacs. In the revised manuscript, sexual organ identity is mentioned only in the second paragraph where sex determination of floral organs is described.

5. In the abstract the statement that bidirectional transcription is the key mechanism should be toned down because it is not shown in this study whether the Antisense *SUF* plays a role directly in the repression of MpFGMYB.

As described in response to referee #2, results from additional experiments indicate that *SUF* acts *in cis*. This finding is included in the Abstract, and also discussed in the main text. We also changed the title of the paper featuring this important notion, rather than emphasizing conserved roles of the MYB subfamily in the previous one. This is also to respond to the minor point 3. However, as the referee pointed out, it is still difficult to conclude that the repression is truly at the transcription step. Therefore, in the Abstract, the sentence has been changed to "Thus the bidirectional transcription module at the MpFGMYB/*SUF* locus acts as a toggle between female and male sexual differentiation in *M. polymorpha* gametophytes". Similar expressions in the main text have been also revised accordingly.

6. The last sentence of the Abstract is not substantiated by findings presented here and should be modified.

In response to this comment, we changed the last sentence of the Abstract from "Thus, conserved regulators are inherently used in the female gametophytes development across

land plants but with diverse upstream regulation" to "Thus, phylogenetically related MYB transcription factors regulate female gametophyte development across land plants".

Accepted

26th November 2018

Thank you for submitting a revised version of your manuscript. It has now been seen by two of the original referees whose comments are shown below.

As you will see they both find that all criticisms have been sufficiently addressed and I am therefore pleased to inform you that your manuscript has now been accepted for publication in The EMBO Journal. Before we can send your files off to production I would ask you to address the following points

REFERE REPORTS

Referee #2:

In the revised version, Hisanaga et al. have made a very thorough attempt to address the concerns that were raised previously. The mode of action of SUF has been clarified and the additional data strengthen the conclusions considerably. I recommend acceptance of the revised manuscript for publication.

Minor comment: It would be helpful to add the number of clones that have been analyzed for the 5' and 3' RACE experiments, and the number of clones that support each position.

Referee #3:

The authors have addressed my comments adequately and the work is ready for publication

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Keiji Nakajima

Journal Submitted to: EMBO Journal

Manuscript Number: EMBOJ-2018-100240R

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if $n < 5$, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
 - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ^2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
 - are tests one-sided or two-sided?
 - are there adjustments for multiple comparisons?
 - exact statistical test results, e.g., P values = x but not P values < x;
 - definition of 'center values' as median or average;
 - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself.

Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	The way sample size was chosen is not described in the text. However, it has been maximized within practically possible ranges and to match the standard of the field. In each graph, data points are plotted along with the bar and S.D.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes. Statistical tests are described in the legend to each figure. Raw measurement data and statistical values are submitted together with the manuscripts.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Gene expression levels among biological replicates are assumed to match Gaussian distribution, and thus parametric statistical analysis, such as t-test and Tukey-Kramer method were used.
Is there an estimate of variation within each group of data?	Yes. Mean $\pm$ S.D. values are presented. P values are reported in the Source data file.
Is the variance similar between the groups that are being statistically compared?	Yes

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

<http://www.antibodypedia.com>

<http://1degreebio.org>

<http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo>

<http://grants.nih.gov/grants/olaw/olaw.htm>

<http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm>

<http://ClinicalTrials.gov>

<http://www.consort-statement.org>

<http://www.consort-statement.org/checklists/view/32-consort/66-title>

<http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun>

<http://datadryad.org>

<http://figshare.com>

<http://www.ncbi.nlm.nih.gov/gap>

<http://www.ebi.ac.uk/ega>

<http://biomodels.net/>

<http://biomodels.net/miriam/>

<http://jij.biochem.sun.ac.za>

http://oba.od.nih.gov/biosecurity/biosecurity_documents.html

<http://www.selectagents.gov/>

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18. Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	Yes, RNA-seq data has been deposited at the DDBJ database, and it is mentioned in the Data availability section.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right)).	Yes, RNA-seq data has been deposited at a public database. Source data (RT-qPCR data and uncropped gel images) are submitted with the manuscript.
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	No
---	----