

Uniwersytet im. Adama Mickiewicza w Poznaniu

Szkoła Doktorska Nauk Przyrodniczych

Wydział Biologii



Rozprawa doktorska

**Charakterystyka funkcjonalna miRNA specyficznych dla
wątrobowców w rozwoju narządów płciowych
*Marchantia polymorpha***

Bharti Aggarwal

Praca napisana pod kierunkiem

Prof. dr. hab. Zofii Szweykowskiej-Kulińskiej

oraz dr Haliny Pietrykowskiej

Poznań, 2025

Adam Mickiewicz University, Poznan, Poland

Doctoral School of Natural Sciences

Faculty of Biology



Doctoral dissertation

**Functional characterization of liverwort-specific miRNAs in sexual
organ development of *Marchantia polymorpha***

Bharti Aggarwal

Thesis written under the supervision of
Prof. dr. hab. Zofia Szweykowska-Kulińska
and dr Halina Pietrykowska

Poznań, 2025

FUNDING

This work was supported by the Polish National Science Center NCN (OPUS grant: UMO/2020/39/B/NZ3/00539) and IDUB (The Initiative of Excellence Research University 05/IDUB/2019/94) at Adam Mickiewicz University, Poznan.

PUBLICATIONS

Some of the findings described in this dissertation have been published in the following research paper:

Aggarwal, B., Karlowski, M.W., Nuc, P., Jarmolowski, A., Pietrykowska, H., Szweykowska-Kulinska, Z. (2024). MicroRNAs differentially expressed in vegetative and reproductive organs of *Marchantia polymorpha*- insights into their expression pattern, gene structures and function. (RNA Biology. doi: 10.1080/15476286.2024.2303555. Impact Factor: 4.76)

Review article:

1. Pietrykowska, H., Alisha, A., Aggarwal, B., Watanabe, Y., Ohtani, M., Jarmolowski, A., Sierocka, I., Szweykowska-Kulinska, Z. (2023). Conserved and non-conserved RNA-target modules in plants: Lessons for better understanding Marchantia development. (Plant Molecular Biology. DOI: 10.1007/s11103-023-01392-y. Impact Factor: 4.33)

Other publications:

1. Review article: Kandhol, N., Aggarwal, B., Bansal, R., Parveen, N., Singh, V.P., Chauhan, D.K., Sonah, H., Sahi, S., Grillo, R., Videa, J.P., Deshmukh, R., Tripathi, D.K. (2022) Nanoparticles as a potential protective agent for arsenic toxicity alleviation in plants. (Environment Pollution. DOI: 10.1016/j.envpol.2022.118887. Impact Factor: 8.9)
2. Research article: *Kumawat, S., *Aggarwal, B., Rana, N., Mandlik, R., Mehra, A., Shivaraj, S.M., Sonah, H., Deshmukh, R., (2021). Identification of Aquaporins and deciphering their role under salinity stress in pomegranate

- (*Punica granatum*). (Journal of Plant Biochemistry and Biotechnology. DOI: 10.1007/s13562-021-00738-1. Impact Factor: 1.6) *Equally contributed
3. Book chapter: Kaur, S., Vas, S., Aggarwal, B., Kaur, J., Bhardwaj, A. (2024). Multiplex genome editing approaches for plant genetic manipulations. In Alok, A., Kumar, J., Shekhawat, M. (Eds.) Basics of CRISPR/Cas mediated Plant genome editing. <https://doi.org/10.1201/9781003247685>.
 4. Book chapter: Aggarwal, B., Rajora, N., Raturi, G., Dhar, H., Kadam, S. B., Mundada, P. S., ... & Sonah, H. (2023). Biotechnology and urban agriculture: A partnership for the future sustainability. (Plant Science. doi: 10.1016/j.plantsci.2023.111903)
 5. Book chapter: Aggarwal, B., Vats, S., Kaushal, L., Singh, A., Padalkar, G., Yadav, H., Kumar, V., Sinha, S., Umate, S.M. (2023). Biofortification of Maize (*Zea mays*). In Deshmukh, R., Nadaf, A., Ansari, W.A., Singh, K., Sonah, H. (Eds.) Biofortification in Cereals. Springer Nature, Singapore. (https://doi.org/10.1007/978-981-19-4308-9_8)
 6. Book chapter: Rahim, M.S. Parveen, A., Aggarwal, B., Madhavan, A., Kumar, P., Kumar, V., Rana, N., Bansal, R., Deshmukh, R., Roy, J. (2021) Computational tools and approaches for Aquaporin (AQP) research. In Tripathi D, Roy Choudhury A, Deshmukh R (Eds.) Metal and nutrient Transporters in abiotic stress. 1-32. <https://doi.org/10.1016/B978-0-12-817955-0.00001-6>

Supervisor

Prof. dr. hab. Zofia Szweykowska-Kulińska

Department of Gene Expression, Faculty of Biology, Adam Mickiewicz University
Poznan, Poland

Co-Supervisor

dr Halina Pietrykowska

Department of Gene Expression, Faculty of Biology, Adam Mickiewicz University
Poznan, Poland

ਮੈਂ ਸ਼ੁਕਰਗੁਜ਼ਾਰ ਹਾਂ

ਮੇਰੇ ਮਾਤਾ ਜੀ ਦਾ, ਜਿਨ੍ਹਾਂ ਨੇ ਜ਼ਿੰਦਗੀ ਤੋਂ ਰੁਖਸਤ ਲੈ ਲਿਆ, ਪਰ ਉਹਨਾਂ ਦਾ ਪਿਆਰ ਹਰ ਪਲ ਮੇਰੇ ਨਾਲ ਰਿਹਾ।

ਮੇਰੇ ਪਿਤਾ ਜੀ ਦਾ, ਜਿਨ੍ਹਾਂ ਨੇ ਹਰ ਮੋੜ 'ਤੇ ਵਿਸ਼ਵਾਸ ਕਰਦੇ ਹੋਏ ਮੇਰੀ ਹੱਸਲਾ ਅਫ਼ਜ਼ਾਈ ਕੀਤੀ।

ACKNOWLEDGEMENTS

My most heartfelt thanks to Prof. dr. hab. Zofia Szweykowska-Kulińska for scientific advices and constant encouragement, without which none of this work would have been possible. I also thank my co-supervisor dr Halina Pietrykowska for giving me the majority of my hands-on-training in the lab and sharing more than protocol knowledge with me. I appreciate all the bioinformatic help from Prof. dr. hab. Wojciech Karlowski and willingness to always answer my questions. I also thank Prof. dr. hab. Artur Jarmołowski for his scientific encouragement.

The opportunity to work and learn in the lab was unique and thrilling. Thanks to all the members of ZEG - Dawid, Kasia, Ola-Świda, Andrzej, Mateusz, Iza, Łukasz, Kuba for helping me from the early days of my PhD and for stimulating discussions that helped me to make progress in my research. Thanks to all fellow PhD camaraderie – Marta, Monika, Zuza and Daria for keeping up my spirit and helping me in times of need. Thanks to all the Master's students - Ksenia, Oscar and Gabriella for a fun-filled environment.

Thank you Prof. UAM dr. hab. Bożena Sikora, Małgorzata Klimorowska, and Magdalena Walczak for paving the way for me to make it to the finish line.

Thank you Ms. Arleta Kucz for always being patient with me and helping me in all the documentation so that I can attend the insightful conferences.

I am indebted to my best friends – Romica, Renuka, Navpreet, Rimpay, Aman, Sandeep & Pankaj for their unwavering support despite-of the time differences and long distances. Thank you Anand Maurya for solving all my technical problems and giving me useful advices. Thank you Gulsamal for being my oasis in the sometimes never ending desert of our shared lab building. Thank you Chandni and Vipul for being the most gracious hosts as we bonded over festival celebrations. Thanks to all of my amazing friends for shepherding me through!

A deep sense of gratitude for my family, especially my siblings. Words cannot express how much I missed you all and home during these years. Thank you for being my strength and supporting me even from distance. Thank you Divit and Sia for always cheering me up.

A special note of thanks to my human diary, Christoph, for being my anchor and for always uplifting my confidence. This accomplishment is as much yours as it is mine. I feel truly blessed to receive love from my Kentrath Parivar. You gave me a haven much closer to home where I could always return to rest my ship when the storms grew too strong.

Thank you all for playing your part in this journey!

Table of Contents

| | |
|--|-----------|
| ABBREVIATIONS..... | 12 |
| STRESZCZENIE | 13 |
| ABSTRACT | 15 |
| Chapter 1: Introduction | 17 |
| 1.1 <i>M. polymorpha</i> as a model organism..... | 20 |
| 1.2 Marchantia as a model for forward and reverse genetic analyses | 20 |
| 1.3 The life cycle of Marchantia | 22 |
| 1.4 The anatomy of Marchantia | 24 |
| 1.4.1 The anatomy of vegetative thallus | 24 |
| 1.4.2 Asexual reproduction..... | 26 |
| 1.4.3 Sexual reproduction..... | 28 |
| 1.4.4 Development of archegoniophores and archegonia..... | 28 |
| 1.4.5 Development of antheridiophores and antheridia..... | 31 |
| 1.4.6 Fertilization and sporophyte development | 33 |
| 1.5 Regulators of gene expression throughout plant development: miRNAs | 35 |
| 1.5.1 miRNAs are generated via a complex biogenesis pathway in plants..... | 36 |
| 1.5.2 Mode of action of miRNAs | 38 |
| 1.5.3 Vignette of miRNA/target mRNA modules steering Marchantia's development. | 39 |
| Chapter 2: Aim of the thesis | 47 |
| Chapter 3: Materials..... | 48 |
| 3.1 Plant material, propagation and maintenance..... | 48 |
| 3.2 Bacterial strains: Culture and growth parameters | 48 |
| 3.3 Culture media and reagents | 49 |
| 3.3.1 Bacterial culture media | 49 |
| 3.3.2 Antibiotics | 49 |
| 3.3.3 Plant culture media..... | 50 |
| 3.3.4 Nucleic acid molecular weight markers..... | 52 |
| 3.4 Solutions and Buffers | 52 |
| 3.4.1 Electrophoresis solution..... | 52 |
| 3.4.2 Agarose gel | 53 |
| 3.4.3 DNA extraction buffer for genotyping..... | 54 |
| 3.4.4 RNA isolation solutions..... | 54 |

| | | |
|-------------------------------|---|-----------|
| 3.4.5 | GUS staining solutions | 55 |
| 3.4.6 | Northern Blot Hybridization solutions | 57 |
| 3.5 | Kits | 59 |
| 3.6 | Enzymes | 60 |
| 3.7 | Vectors | 61 |
| 3.8 | Oligonucleotides used | 61 |
| 3.9 | Softwares used..... | 62 |
| Chapter 4. Methodology | | 64 |
| 4.1 | <i>Escherichia coli</i> (DH-5α) transformation | 64 |
| 4.2 | <i>Agrobacterium tumefaciens</i> transformation | 64 |
| 4.3 | Polymerase Chain Reaction (PCR) | 65 |
| 4.4 | Agarose gel electrophoresis | 65 |
| 4.5 | Genomic DNA isolation | 65 |
| 4.6 | DNA isolation for genotyping | 66 |
| 4.7 | RNA extraction | 67 |
| 4.8 | DNase treatment | 68 |
| 4.9 | cDNA synthesis | 69 |
| 4.10 | RT-qPCR..... | 70 |
| 4.11 | CRISPR/Cas9 based genome editing..... | 71 |
| 4.12 | Plasmid DNA isolation | 72 |
| 4.13 | DNA sequencing (Sanger method) | 72 |
| 4.14 | Crossing of <i>Marchantia</i> plants | 72 |
| 4.15 | <i>Agrobacterium</i> -mediated transformation of <i>Marchantia</i> sporelings | 72 |
| 4.16 | Phenotypic analysis..... | 74 |
| 4.17 | Protocol for the transfer of plants from soil to <i>in vitro</i> culture..... | 74 |
| 4.18 | Construction of promoter reporter plants | 74 |
| 4.19 | GUS staining of <i>Marchantia</i> vegetative and reproductive tissues | 75 |
| 4.20 | Northern hybridization | 75 |
| 4.20.1 | RNA electrophoresis in dPAG..... | 75 |
| 4.20.2 | Transfer of RNA from gel to membrane | 76 |
| 4.20.3 | EDC-cross linking | 76 |
| 4.20.4 | Labeling of oligonucleotide probes..... | 76 |
| 4.20.5 | Hybridization and washing | 77 |

| | | |
|---------------------------|---|-----------|
| 4.20.5 | Hybridization signal detection | 77 |
| 4.21 | RNA-seq library preparation for Illumina sequencing | 78 |
| 4.21.1 | RNA fragmentation and priming | 78 |
| 4.21.2 | First-strand cDNA synthesis | 78 |
| 4.21.3 | Second-strand cDNA synthesis | 79 |
| 4.21.4 | Purification of double-stranded cDNA using beads | 79 |
| 4.21.5 | End prep of cDNA library | 79 |
| 4.21.6 | Adaptor ligation..... | 80 |
| 4.21.7 | PCR enrichment of adaptor-ligated DNA | 80 |
| 4.21.8 | Assess library quality | 81 |
| Chapter 5: Results | | 82 |
| 5.1 | Liverwort-specific miRNAs differentially expressed in Marchantia vegetative and generative organs | 82 |
| 5.1.1 | MpmiR11737a/b | 83 |
| 5.1.1.1 | Pri-MpmiR11737a/b and mature MpmiR11737a/b show different expression and accumulation patterns in Marchantia..... | 83 |
| 5.1.1.2 | Identification of full-length transcript of Mp <i>MIR11737a</i> gene | 84 |
| 5.1.1.3 | Mp <i>MIR11737a</i> gene structure identification | 86 |
| 5.1.1.4 | Identification of putative target of MpmiR11737a from degradome data..... | 87 |
| 5.1.2 | MpmiR11865/11865* | 89 |
| 5.1.2.1 | MpmiR11865* and MpmiR11865 exhibit opposite accumulation patterns in Marchantia..... | 89 |
| 5.1.2.2 | Pri-MpmiR11865 transcript is mainly expressed in gametangiophores of Marchantia..... | 91 |
| 5.1.2.3 | Identification of full-length transcript of Mp <i>MIR11865</i> gene | 92 |
| 5.1.2.4 | Mp <i>MIR11865</i> gene structure identification | 93 |
| 5.1.2.5 | Identification of putative targets of MpmiR11865 and MpmiR11865* from degradome data | 93 |
| 5.2 | Functional characterization of liverwort-specific MpmiR11887 | 96 |
| 5.2.1 | Pri-MpmiR11887 and MpmiR11887 presence exclusively in antheridiophores . | 96 |
| 5.2.2 | Identification of full-length transcript of Mp <i>MIR11887</i> gene..... | 97 |
| 5.2.3 | Mp <i>MIR11887</i> gene structure identification..... | 98 |
| 5.2.4 | Characterization of Mp <i>MIR11887</i> gene promoter <i>in planta</i> activity | 99 |
| 5.2.5 | Analysis of CRISPR/Cas9 KO plants for Mp <i>MIR11887</i> | 101 |

| | | |
|---|---|------------|
| 5.2.5.1 | Genotyping and validation of Mp <i>MIR11887</i> CRISPR/Cas9 knockouts | 101 |
| 5.2.5.2 | Phenotypic analysis of Δ <i>mpmir11887^{ge}</i> mutant plants..... | 104 |
| 5.2.6 | Identification of putative target of MpmiR11887 from transcriptomic and degradome data analyses..... | 108 |
| 5.3 | Functional characterization of liverwort-specific MpmiR11796..... | 112 |
| 5.3.1 | Pri-MpmiR11796 and MpmiR11796 are upregulated in archegoniophores | 112 |
| 5.3.2 | Identification of full-length transcript of Mp <i>MIR11796</i> gene..... | 113 |
| 5.3.3 | Mp <i>MIR11796</i> gene structure identification..... | 114 |
| 5.3.4 | Characterization of Mp <i>MIR11796</i> gene promoter <i>in planta</i> activity | 114 |
| 5.3.5 | Generation of CRISPR/Cas9 KO plants for Mp <i>MIR11796</i> | 117 |
| 5.3.6 | Phenotypic analysis of Δ <i>mpmir11796^{ko}</i> mutant plants..... | 120 |
| 5.3.7 | Identification of putative mRNA target of MpmiR11796 from transcriptomic and degradome data analyses..... | 133 |
| 5.3.8 | GO analyses..... | 134 |
| 5.4 | Identification of novel miRNAs expressed specifically or abundantly in archegoniophores | 141 |
| 5.4.1 | FAN-1 | 141 |
| 5.4.2 | FAN-3..... | 143 |
| 5.4.3 | FAN-4..... | 144 |
| Chapter 6. Discussion | | 146 |
| Chapter 7: Conclusions and Future perspectives | | 161 |
| Literature cited..... | | 163 |
| Supplementary data..... | | 175 |

ABBREVIATIONS

| | |
|-------------------|--|
| AA | amino acid |
| ACT | actin |
| AGO1 | ARGONAUTE1 |
| AR | archegonial receptacle |
| BLAST | basic local alignment search tool |
| bp | base pair |
| Cas9 | CRISPR associated protein 9 |
| cDNA | complementary DNA |
| CDS | coding sequence |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| DNA | deoxyribonucleic acid |
| dPAG | denaturing polyacrylamide gel |
| dsDNA | double-stranded DNA |
| EC | egg cell |
| Fa | archegoniophore |
| FR | far-red light |
| Fv | female vegetative thallus |
| G1/G2 | Generation1/Generation2 |
| GUS | β-glucuronidase |
| GSP | gene specific primer |
| IP | immunoprecipitation |
| kb | kilo basepairs |
| KO | knock out |
| Ma | antheridiophore |
| MIR | microRNA gene |
| miRNA | microRNA |
| miRNA* | passenger microRNA strand |
| Mpo/Mp | <i>Marchantia polymorpha</i> |
| mRNA | messenger RNA |
| Mv | male vegetative thallus |
| ncRNA | non-coding RNA |
| NGS | next-generation sequencing |
| nt | nucleotide |
| PCR | polymerase chain reaction |
| Pen | <i>Pellia endiviifolia</i> |
| pre-miRNA | precursor-miRNA |
| pri-miRNA | primary-miRNA |
| PRM | protamine-like |
| R:FR ratio | red: far-red ratio |
| RACE | rapid amplification of cDNA ends |
| RISC | RNA induced silencing complex |
| RLM-RACE | RNA ligase mediated RACE |
| RNA | ribonucleic acid |
| RPM | reads per million |
| RT-PCR | reverse transcription PCR |
| RT-qPCR | reverse transcription-quantitative PCR |
| SC | spermatogenous cells |
| sRNA | small RNA |
| sgRNA | single guide RNA |
| snRNA | small nuclear RNA |
| t-plot | target-plot |
| TAK-1 | TAKARAGAIKE-1 (<i>M. polymorpha</i> accession name; male) |
| TAK-2 | TAKARAGAIKE-2 (<i>M. polymorpha</i> accession name; female) |
| tasiRNA | trans-acting small-interfering RNA |
| TF | transcription factor |
| TSS | transcription start site |
| UTR | untranslated region |
| WT | wild-type |

STRESZCZENIE

Prawidłowe formowanie organów generatywnych oraz sukces reprodukcyjny wymagają precyzyjnej kontroli procesów rozwojowych na wszystkich etapach cyklu życiowego rośliny. Procesy te opierają się na złożonych sieciach zależności genetycznych, które są w bardzo małym stopniu poznane u roślin niższych. U wątrobowców zidentyfikowano dziewięć konserwatywnych miRNA, obecnych u wszystkich roślin lądowych, oraz zestaw miRNA specyficznych dla tej grupy roślin. Rola wątrobowcowo-specyficznych miRNA w *Marchantia polymorpha* pozostaje w dużej mierze niezbadana. Niektóre z nich wykazują zróżnicowane wzorce ekspresji w organach generatywnych *Marchantia*. W celu zbadania ich roli w rozmnażaniu płciowym przeprowadziłam analizy eksperymentalne wzorców ekspresji różnych miRNA specyficznych dla wątrobowców i wybrałam dwa miRNA do dalszych badań.

Jeden z badanych miRNA, MpmiR11887, wykazywał wyjątkowo wysoką ekspresję w męskich gametangioforach – anteridioforach. Analizy baz danych genomowych oraz badania eksperymentalne wykazały, że gen *MpMIR11887* stanowi niezależną jednostkę transkrypcyjną. Co więcej, pre-MpmiR11887 znajduje się tuż za krótkim ORF-em kodującym potencjalne białko o długości 65 aminokwasów, którego funkcja pozostaje nieznana, choć może ono reprezentować miPeP. Barwienie histochemiczne GUS ujawniło, że promotor *MpMIR11887* jest aktywny głównie w męskich organach generatywnych, tj. młodych anteridiach oraz komórkach spermatogenicznych anteridioforów, a słabiej w archegoniach. Rośliny te również charakteryzowały się wcześniejszym rozwojem anteridioforów z większymi dyskami anteridialnymi oraz większymi dojrzałymi anteridiami. Wyniki te wskazują, że MpmiR11887 bierze udział w regulacji rozwoju anteridialnych dysków i anteridiów.

Drugi z tych miRNA, MpmiR11796, wykazywał zróżnicowaną ekspresję w organach wegetatywnych i generatywnych *Marchantia*, z wyraźnie wysoką akumulacją w żeńskich gametangioforach – archegonioforach. Analizy baz danych genomowych oraz eksperymenty 5'-RLM i 3'-RACE wykazały, że gen *MpMIR11796* stanowi niezależną jednostkę transkrypcyjną pozbawioną intronów, częściowo nakładającą się z genem kodującym białko. Za pomocą histochemicznego barwienia GUS

zaobserwowano silną aktywność promotora genu *MpMIR11796* w ryzoidach języczkowych, znajdujących się w ramionach i nóżkach rodniostanów osobników żeńskich. Ponadto męskie osobniki z nokautem genu *MpMIR11796* ($\Delta mpmir11796ko$), uzyskane metodą CRISPR/Cas9, wytwarzały mniej ryzoidów podczas rozwoju gemm, natomiast rośliny żeńskie $\Delta mpmir11796ko$ wykazywały zmniejszone rozmiary plechy i archegonioforów, skrócenie nóżki, zmniejszoną liczbę ryzoidów języczkowych oraz zaburzenia w rozwoju komórki jajowej. Doświadczenia nad krzyżowaniem zmutowanych osobników wykazały także poważne zaburzenia w wytwarzaniu gametangioforów oraz w rozwoju gamet w kolejnym pokoleniu. Obserwacje te wskazują, że *MpmiR11796* odgrywa ważną rolę w ogólnym rozwoju *Marchantii*, a w szczególności w rozwoju struktur i funkcjonowaniu organów rozrodczych.

Analizy danych transkryptomicznych i degradomowych pozwoliły na identyfikację potencjalnych docelowych mRNA dla *MpmiR11796* i *MpmiR11887*.

ABSTRACT

The developmental processes must be meticulously controlled at all stages of a plant's life cycle to ensure proper sexual organ development and reproductive success. These processes rely on complex genetic networks, which remain unknown in non-vascular plants. In liverworts, nine conserved miRNAs common to all land plants and a set of non-conserved (liverwort-specific) miRNAs were identified. However, the functional roles of the non-conserved miRNAs in *Marchantia polymorpha* remain largely unexplored. Notably, some of them show differential expression patterns across the generative organs of *Marchantia*. To investigate their roles in sexual reproduction, I conducted experimental analyses on various liverwort-specific miRNAs and selected two candidates for in-depth studies.

One of these miRNAs, MpmiR11887, had an exceptionally high expression in male gametangiophores - antheridiophores. Genomic database and experimental analyses revealed that Mp*MIR11887* gene represents an independent transcriptional unit. Moreover, the pre-MpmiR11887 is located just downstream of short ORF encoding a putative 65-AA protein of unknown function, which may represent a miPeP. Histochemical GUS staining revealed that the Mp*MIR11887* promoter is active predominantly in the male generative organs i.e. young antheridia and spermatogenous cells of antheridiophores and weakly in the archegonia of archegoniophores. Additionally, $\Delta\text{mpmir11887}^{\text{ko}}$ plants showed early development of antheridiophores with larger antheridial discs and larger size of mature antheridia. These results show that MpmiR11887 is involved in at least fine-tuning the antheridial discs and antheridia development.

Another miRNA, MpmiR11796, exhibited differential expression across *Marchantia*'s vegetative and reproductive organs, with notably high accumulation in female gametangiophores – archegoniophores. Genomic database analyses along-with 5'-RLM and 3'-RACE experiments revealed that Mp*MIR11796* gene represents an independent intron-less transcription unit, although sharing sequence overlap with a protein-coding gene. A strong Mp*MIR11796* gene promoter activity was observed in the pegged rhizoids located within digitate rays and stalk of archegoniophores using

histochemical GUS staining. Furthermore, CRISPR/Cas9 male $\Delta\text{mpmir}11796^{\text{ko}}$ plants produced less rhizoids during gemmae development, while female $\Delta\text{mpmir}11796^{\text{ko}}$ plants showed reduction in thallus and archegonial receptacle size, reduced stalk length, reduced number of pegged rhizoids and abnormalities in egg cell development. Furthermore, crossing experiments revealed severe impairment in gametangiophores production and gamete development in the next generation. These observations show that MpmiR11796 plays an important role in the overall development of *Marchantia*, but predominantly in sexual reproductive organ formation and function.

Finally, transcriptome and degradome data analyses identified potential targets for MpmiR11887 and MpmiR11796.

Chapter 1: Introduction

The transition of plants to terrestrial life occurred by mid-Ordovician (approximately 450 million years ago) – an event that was one of the most transformative events in Earth's history (Servais et al., 2019; Moody, 2020; Strother & Foster, 2021). This transition profoundly influenced the terrestrial environment and laid the foundation for the emergence of complex life forms on land. In order to survive and thrive in this new habitat, the early land plants had to overcome numerous challenges and the need for new structural and reproductive adaptations arose. Such key adaptations include sensing environmental signals (light and gravity), evolution of cuticle, emergence of extracellular structural components (sporopollenin / lignin / pectic acid), resistance to prolonged dehydration, development of intercellular communication networks (plasmodesmata, plant hormones, receptor, their ligands), evolution of multicellular meristems, three-dimensional growth, presence of air pores / stomata, development of rooting system, evolution of embryo (sporophyte, 2n), alternation of generations (Oliver et al., 2005; Bowman et al., 2007; Pires & Dolan, 2012; Yeats & Rose, 2013; Bowman et al., 2016; Bowman et al., 2017; Rensing, 2018; Moody, 2020; Kohchi et al., 2021). These innovations were accompanied by several changes at the molecular level leading to emergence of sophisticated gene regulatory pathways. These regulatory pathways very often are controlled by small non-coding RNA (ncRNAs), particularly microRNAs (miRNAs), which fine-tune gene expression and shape the developmental destiny of an organism. It is especially fascinating to explore how such regulatory mechanisms function in *Marchantia polymorpha*, a representative of one of the three bryophyte lineages. The study of miRNA-mediated mechanisms in *Marchantia* could offer valuable insights into ancient molecular strategies that supported plant life on land and provide a comprehensive understanding of evolutionary developmental biology.

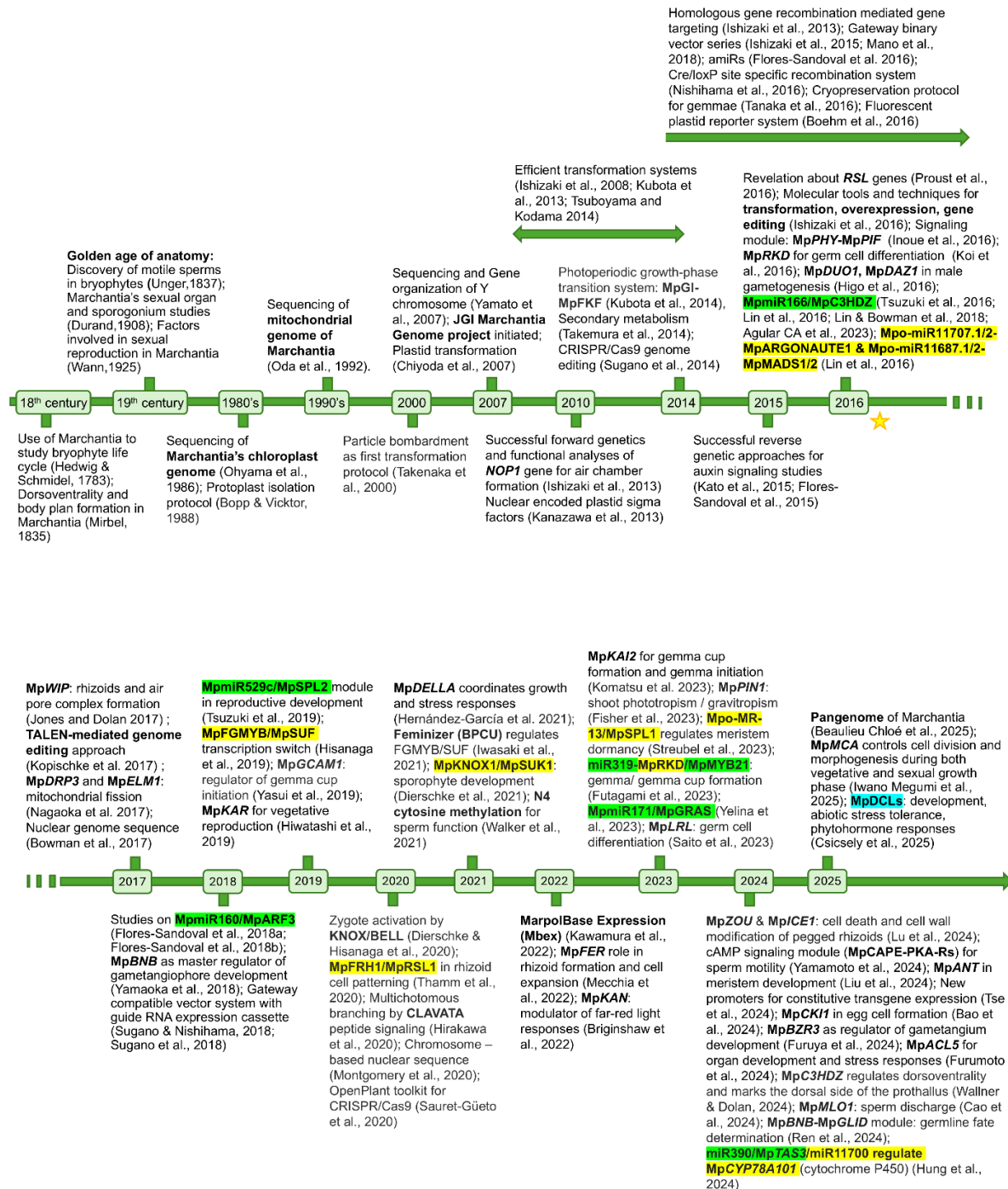


Figure 1. The timeline of major discoveries in Marchantia research. The figure illustrates the chronological progression of discoveries in Marchantia, including development in the genetic molecular tool-kit, identification of factors involved in Marchantia asexual and sexual reproduction, spanning from the 18th century to present year, 2025. The conserved miRNA- target mRNA modules has been highlighted in

green, whereas, the unique and novel miRNA/target mRNA modules have been highlighted in yellow. MpDCLs, the regulatory kingpin, are highlighted in turquoise.

Abbreviations: **NOP1**, NOPPERABO1; **GI**, GIGANTEA; **FKF**, FLAVIN-BINDING KELCH REPEAT F-BOX; **ARF**, AUXIN RESPONSE FACTOR; **CAPE**, COMBINED ADENYLYL CYCLASE WITH PHOSPHODIESTERASE; **PKA-R**, cAMP-DEPENDENT PROTEIN KINASE; **FGMYB**, FEMALE GAMETOPHYTE MYB; **KNOX**, KNOTTED1-LIKE HOMEODOMAIN; **MIR**, microRNA; **PHY**, PHYTOCHROME; **PIF**, PHYTOCHROME-INTERACTING FACTOR; **SPL**, SQUAMOSA PROMOTER-BINDING PROTEIN-LIKE; **SUF**, SUPPRESSOR OF FEMINIZATION; **SUK**, SUPPRESSOR OF KNOX; **Cre/LoxP**, CYCLIZATION RECOMBINASE/LOCUS OF CROSSING OVER; **RSL**, ROOT HAIR DEFECTIVE SIX-LIKE; **RKD**, RWP-RK DOMAIN-CONTAINING PROTEIN; **DUO1**, DUO POLLEN1; **DAZ1**, DUO1-ACTIVATED ZINC FINGER; **C3HDZ**, CLASS III HOMEODOMAIN LEUCINE ZIPPER1; **WIP**, WOUND INDUCED POLYPEPTIDE; **TALEN**, TRANSCRIPTION ACTIVATOR-LIKE EFFECTOR NUCLEASE; **DRP3**, DYNAMIN-RELATED PROTEIN3; **ELM1**, ELONGATED MITOCHONDRIA1; **GCAM1**, GEMMA CUP-ASSOCIATED MYB; **KAR**, KARAPPO; **BPCU**, BASIC PENTACYSTEINE ON THE U CHROMOSOME; **FRH1**, FEW RHIZOIDS1; **FER**, FERONIA; **GRAS**, GIBBERELLIN INSENSITIVE, REPRESSOR OF GA INSENSITIVE, SCARECROW-LIKE; **DELLA**, ASPARTIC ACID-GLUTAMIC ACID-LEUCINE-LEUCINE-ALANINE; **MCA**, MID1-COMPLEMENTING ACTIVITY2; **DCL**, DICER-LIKE PROTEIN; **ZOU**, ZHOUP1; **ICE1**, INDUCER OF CBF EXPRESSION1; **ANT**, AINTEGUMENTA; **CKI1**, CYTOKININ-INDEPENDENT1; **BZR3**, BRASSINAZOLE-RESISTANT1; **ACL5**, ACAULIS5; **KAI2**, KARRIKIN INSENSITIVE2; **PIN1**, PIN-FORMED1; **MLO1**, MILDEW RESISTANCE LOCUS O; **BNB**, BONOBO; **GLID**, GERMLINE IDENTITY DETERMINANT; **LRL**, LOTUS JAPONICUS Roothairless LIKE1; **KAN**, KANADI; **PIN1**, PIN-FORMED1.

1.1 *M. polymorpha* as a model organism

The fossil records and currently accepted phylogenies suggest that liverworts represent one of the ancient lineages of land plants (*Embryophyta*) (Morris et al., 2018; Delaux et al., 2019). *M. polymorpha*, a complex thalloid liverwort (also called umbrella liverwort), is found in multiple geographic regions across different continents and can colonize man-disturbed habitats (Duckett et al., 2024). Long before its use in modern biological research, *Marchantia* was documented in the ancient Greek medical texts as a plant applied on open wounds to prevent infection and reduce inflammation (Bowman, 2016; Bowman et al., 2016). Since 18th century, it has been used as a model for biological studies. A timeline of major milestones in *Marchantia* knowledgebase are presented in Fig. 1 and selected miRNA/ target mRNA modules involved in *Marchantia* development are discussed in detail in the section 1.6 of this chapter.

1.2 *Marchantia* as a model for forward and reverse genetic analyses

Land plants evolved from an ancestral charophycean algae (also known as streptophyte algae). Liverworts, hornworts and mosses represent three early diverging plants and a sister group to all other vascular land plants (Bowman et al., 2007; Morris et al., 2018; Puttick et al., 2018). *Marchantia*, a representative liverwort, occupies a pivotal position in the land plant phylogeny (Bowman, 2022). The last common ancestor of *Marchantia* and model plant for angiosperms i.e. *Arabidopsis*, was also the last common ancestor of all land plants (Ligrone et al., 2012). Therefore, discoveries made in *Marchantia* can provide valuable insights into the earliest stages of land plant evolution. Given its phylogenetic position, simple body plan, and expanding molecular toolkit, *Marchantia* serves an attractive model system to investigate the gene regulatory mechanisms and infer ancient evolutionary events (Bowman et al., 2022). Bryophytes exhibit unique morphologies in comparison to seed plants like a lack of vascular system, absence of lignified wall, presence of motile sperms and dominance of haploid gametophyte generation (Shaw & Renzaglia, 2004; Pressel & Duckett, 2019). This reflects their evolutionary divergence, but many core genetic regulators, including major transcription factor families, have been conserved across these lineages (Floyd & Bowman, 2007; Rensing et al., 2008; Banks et al., 2011).

Forward genetics offers an unbiased method to identify genes based on the mutant phenotypes. The haploid gametophyte dominant life cycle of *Marchantia* makes it a powerful system for forward genetic screens to discover genes. For example, one of the first mutants identified through a forward genetic screen was *bonobo*, which constitutively produced gametangiophores (Yamaoka et al., 2004). Because of haploid dominant phase, the phenotypic effects of mutations are directly visible in the transformed generation, simplifying the functional gene analyses. *Marchantia* is easy to propagate as it reproduces both sexually (producing spores) and asexually (producing gemmae). Isogenic populations of the same age can be rapidly grown via gemmae - which are produced in large numbers on mature thalli (Bowman, 2016). Being a dioecious species with sex determined by the presence of sex chromosomes (U (♀) or V (♂) chromosomes), *Marchantia* produces reproductive structures on the vegetative thallus during long day conditions in nature and far-red light induction *in vitro* (Bowman, 2016). This makes sexual crosses relatively simple, i.e., the sperms collected from the antheridial disc can be directly applied to the female archegonial receptacle (Shimamura, 2016). Moreover, fertile sperms could be cryopreserved, allowing crosses to be performed whenever reproductive female plants are available (Togawa et al., 2018).

The genome of *Marchantia* is relatively small, about 280 Mb, and mostly consists of single-copy regulatory genes (Bowman et al., 2017). Notably, the absence of whole genome duplications facilitates reverse genetic approaches, providing a significant advantage when compared to other plant species that have undergone genome duplications (Bowman et al., 2017; Lang et al., 2018). Therefore, the core gene regulatory networks are retained in a much simpler form. For instance, sexual reproduction processes in *Marchantia*, including germline cell specification and differentiation, vegetative to reproductive phase transition etc. are regulated by transcription factors, many of which have orthologs in angiosperms (Yamaoka et al., 2018; Saito et al., 2023; Bao et al., 2024). This conservation of core molecular mechanisms makes *Marchantia* an ideal system to study the developmental innovations and biological phenomena from an evolutionary point of view.

The genetic transformation of *Marchantia* is straightforward and highly efficient. A variety of genetic and genomic resources have now been available, including gateway vectors for different applications such as promoter activity assays (Ishizaki et al., 2015; Mano et al., 2018), CRISPR/Cas9-mediated genome editing (Kubota et al., 2013; Sugano et al., 2014; Sugano & Nishihama, 2018; Sugano et al., 2018; Iwakawa et al., 2021), artificial miRNA approach for the suppression of target transcripts (Flores-Sandoval et al., 2016) etc. The well-established and efficient *Agrobacterium*-mediated transformation of spores enables the generation of a large number of transformants, allowing targeted mutagenesis with relative ease (Ishizaki et al., 2008; Kubota et al., 2013; Tsuboyama & Kodama, 2018). With this expanding molecular tool-kit, *Marchantia* is becoming an increasingly feasible model for addressing diverse biological questions. In conclusion, *Marchantia* possesses several features that makes it especially well-suited for answering the fundamental questions in plant biology.

1.3 The life cycle of *Marchantia*

As a typical bryophyte, *Marchantia* has a haploid gametophyte dominant life cycle (see Fig. 2), the anatomical details of which have been described in the next chapter. It reproduces:

- Vegetatively / asexually via multicellular propagules known as gemmae, and
- Sexually through the development of highly modified thalloid structures called gametangiophores (antheridiophores producing antheridia and archegoniophores producing egg bearing archegonia) and subsequently upon fertilization, the diploid sporophyte undergoes meiosis to produce haploid spores for dispersal. These haploid spores initiate the gametophyte stage of the life cycle (Bowman, 2016; Shimamura, 2016; Kohchi et al., 2021; Bowman et al., 2022).

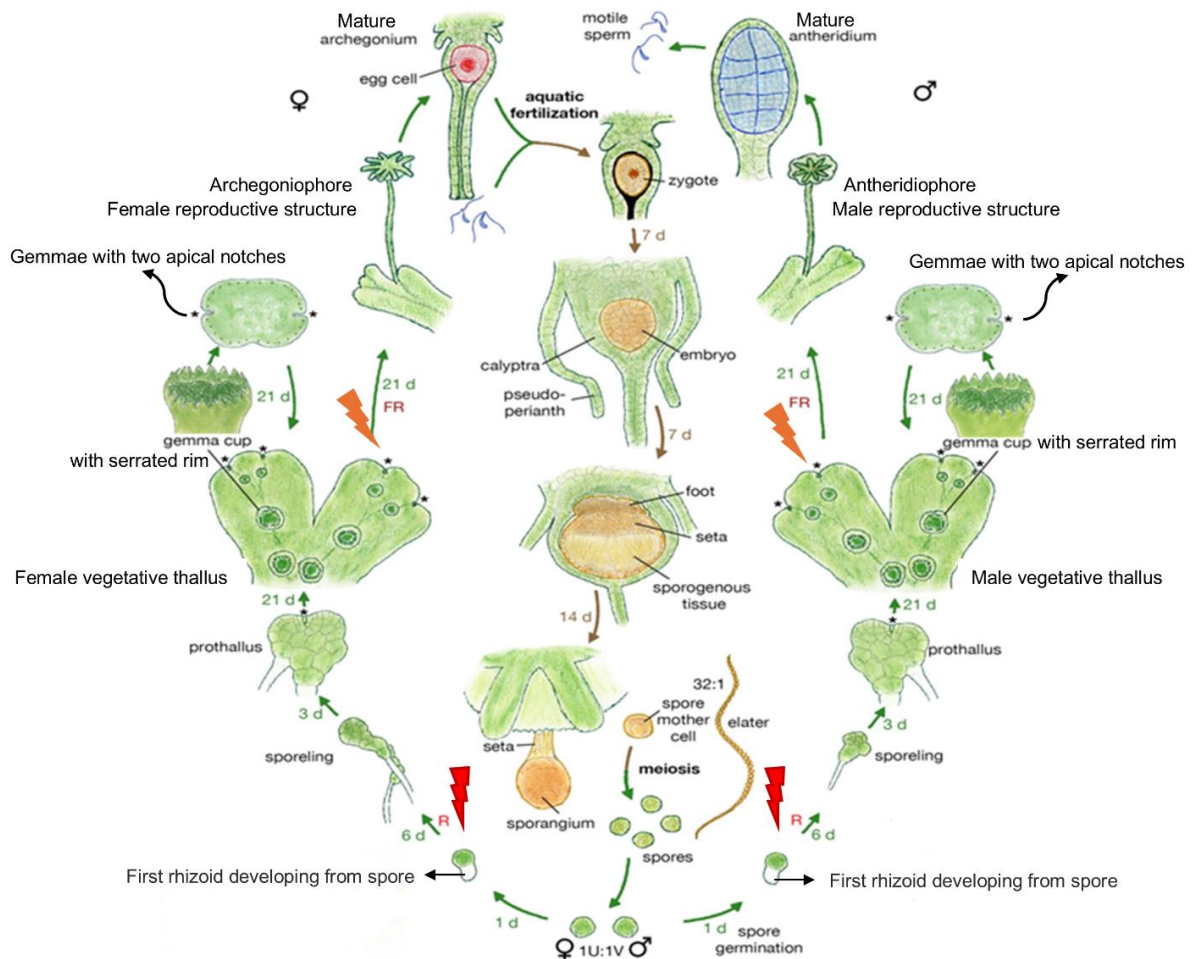


Figure 2. Life cycle of *Marchantia* [adapted from (Bowman et al., 2022) with modifications].

1.4 The anatomy of Marchantia

In the dominant gametophyte phase of its life cycle, *Marchantia* forms a flat, thalloid plant body, from which specialized structures develop.

1.4.1 The anatomy of vegetative thallus

The first stage of the gametophytic phase in *Marchantia*'s life cycle commences with the development of sporelings to male or female vegetative thallus under appropriate conditions (long day / red light under experimental conditions) (Inoue et al., 2019; Bowman et al., 2022). The initial asymmetric cell division in the spore produces one small cell (rhizoid initial cell) and a large cell which continues to undergo subsequent cell divisions. This results in the formation of rhizoids on one end and an irregular mass of photosynthetic tissue on the other end of sporeling (protonema). The cells of protonema then divide irregularly developing into a spherical prothallus. Following this, prothallus undergoes a transition from two-dimensional to three-dimensional growth enabling the production of specialized differentiated tissues (Nishihama et al., 2015; Shimamura, 2016; Bowman et al., 2022).

The mature thallus consists of upper assimilatory region, producing air chambers, the middle storage region and the ventral epidermis with ventral scales and rhizoids. The air chambers on the dorsal surface are filled with photosynthetic filaments, which are connected with the atmosphere outside through a tiny air pore (Fig. 3) (Shimamura, 2016; Kohchi et al., 2021; Bowman et al., 2022). The mature thallus is thick at the midrib region (runs down the center of the thallus from the apical notch region) and becomes thinner towards the margin and clearly shows dorsoventrality (Fig. 4). On the ventral surface of the thallus, three rows of ventral scales (called median, intermediate, marginal) are present (Fig. 4). The median ventral scales are attached to the thallus and include the bundle of pegged rhizoids (Bowman, 2016; Shimamura, 2016).

Rhizoids are unicellular and originate from ventral scales or ventral epidermal cells (Cao et al., 2014; Shimamura, 2016). Also, rhizoids usually do not grow right next to each other, instead, each rhizoid cell is surrounded by non-rhizoid cells which help the rhizoids to space out and rather not cluster together (Cao et al., 2014). Finally, no rhizoids develop on the dorsal epidermis.

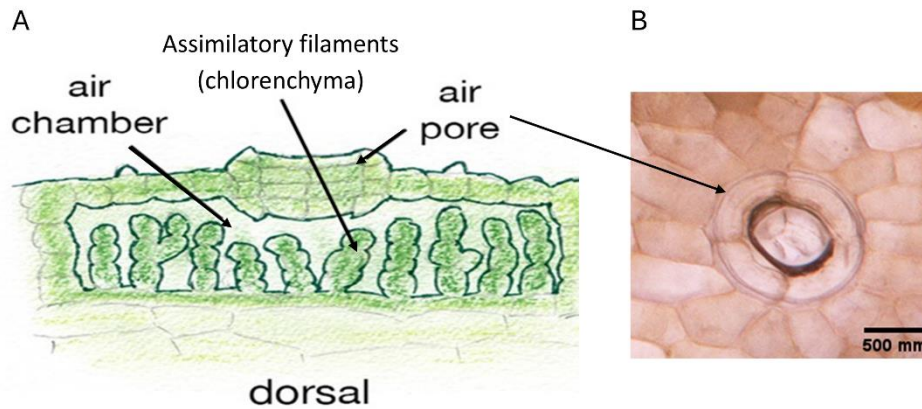


Figure 3. (A) Schematic illustration of air chamber present on the dorsal side of thallus (adapted from Bowman et al., 2022 with modifications); (B) Microscopic image of an air pore on the dorsal surface of thallus.

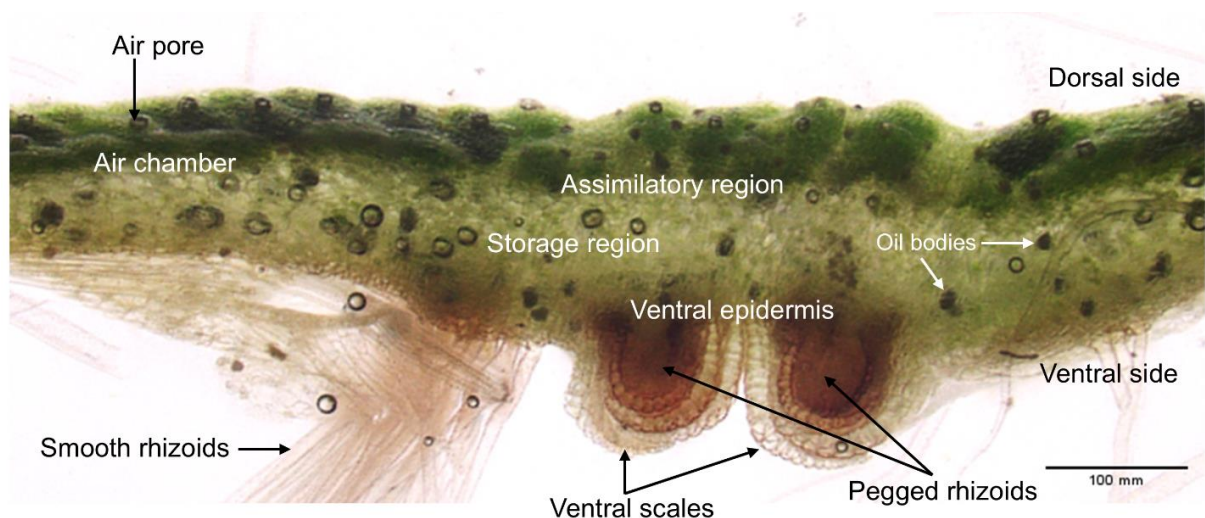


Figure 4. Microscopic image of transverse section of vegetative thallus.

There are two types of rhizoids on the ventral surface of the thallus. Smooth rhizoids grow in a perpendicular direction from the thallus into the soil/medium for anchorage and nutrition. The fungal infection usually takes place via smooth rhizoids (Bowman et al., 2022). Pegged (also called tuberculate) rhizoids grow in bundles along the midrib and are covered by ventral scales. Since there are no internal water-conducting tissues in *Marchantia*, these pegged rhizoids take on that role. They are long, reinforced with peg-like structures and become dead at maturity – just like water conducting cells

(WCCs) in vascular plants (Lu et al., 2024). Being basically hollow, they can move water efficiently through capillary action. Moreover, pegged rhizoids have elastic, cavitation-resistant walls which help them against desiccation (Duckett et al., 2014). The stalk of gametangiophores contains pegged rhizoids which may function as pathways not only for water but also for sperm cells (Cao et al., 2013; Shimamura, 2016). The role of pegged rhizoids as a fertilization route was experimentally verified through dye experiments using 1% Methylene Blue or Toluidine Blue. The movement of dye in archegonial grooves was observed and notably, within 30-60 min, the dye reached the archegonia (Duckett et al., 2014; Pressel & Duckett, 2019). Additionally, the inside parts of digitate rays of archegonial receptacles also consist of pegged rhizoids. They are distributed in such a way that they surround archegonia on all sides and distance between them and archegonia is even less than 1 mm (Shimamura, 2016).

1.4.2 Asexual reproduction

Asexual reproduction is achieved by the formation of gemmae continuously produced from the base of gemma cups (Fig. 5A) on the dorsal thallus surface (both in male and female). The central part of gemmae is several cells thick (Fig. 5C). Gemmae remain dormant inside the gemma cup, and this dormancy is promoted by auxin (Eklund et al., 2015). When removed from the cup or when water droplets fall into gemma cups, gemmae are splashed out (Fig. 5B). After dispersal, gemmae begin to develop rhizoids and possess two meristematic regions (two apical cells on opposite side of the disc; Fig. 5C, 5D) which undergo active cell divisions and ultimately develop into thalli. Initially, rhizoids develop from both dorsal and ventral surface of gemmae, but after 2-3 days of growth, dorsoventrality is established under the influence of light and gravity in an auxin dependent manner, and subsequently rhizoids develop only from the ventral surface (Bowman, 2016; Shimamura, 2016; Kohchi et al., 2021).

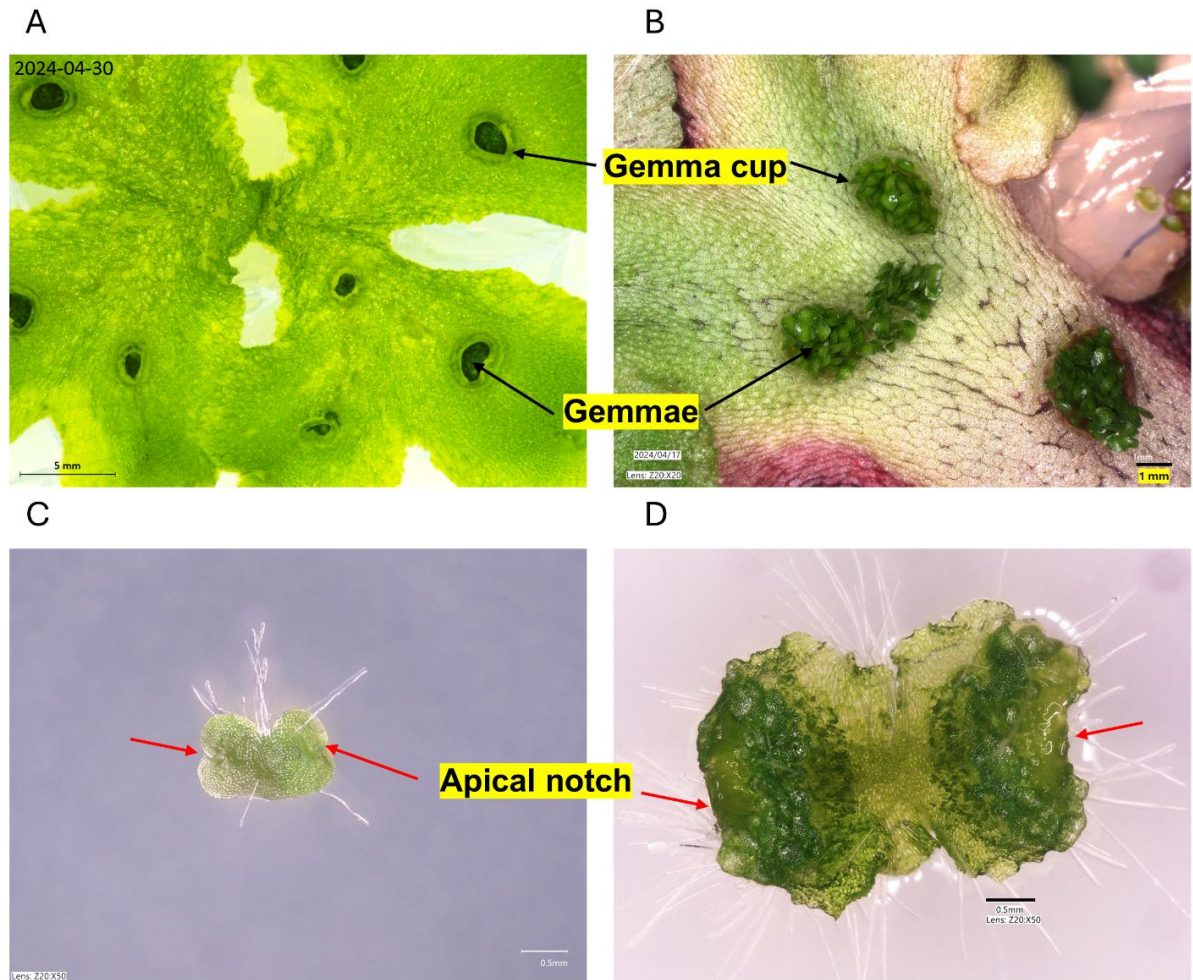


Figure 5. (A) Light microscopic image of 15-day old thallus showing developing gemmae at the base of gemma cup. (B) Light microscopic image of 1-month old thallus showing gemmae dispersal outside of gemma cup. (C) Light microscopic image of 2-day old gemma. (D) Light microscopic image of 7-day old gemma with two apical notches marked by red arrows.

1.4.3 Sexual reproduction

Under long day conditions (and exposition to far-red light in laboratory conditions), *Marchantia* transits from the vegetative to the reproductive developmental phase (Inoue et al., 2019). The induction of the sexual reproductive program initiates visible gametangiophore formation from the apical notch of one of the two thallus branches from bifurcation. The aerial reproductive structures emerging from the mature thalli are basically modified thallus branches known as antheridiophores in male plants and archegoniophores in female plants. Both structures are composed of a stalk which elevates the antheridial disc/archegonial receptacle several centimeters off the ground so that the spores from elevated mature sporophyte is efficiently dispersed (Shimamura, 2016).

1.4.4 Development of archegoniophores and archegonia

The archegoniophores have a highly lobed structure called archegonial receptacle, having 9-11 finger-like rays (called digitate rays; Fig. 6A) and has similar anatomy as thallus in having air chambers, ventral scales and rhizoids (Fig. 6B, 6C, 6D) (Shimamura, 2016). In fact, the digitate rays develop from the inward rolling of the highly branched thallus suitable for retaining water droplets. Inspired by this, a pipette like tool capable of grabbing, transporting and releasing water in a controlled manner was developed (Nakamura et al., 2018). The archegoniophore stalk also has dorsoventrality similar to that of thallus. One could recognize two grooves with bundles of pegged rhizoids and air chambers in the transverse section (Fig. 6D). These pegged rhizoids function as efficient conduits, transporting water and nutrients to the archegonial receptacle. The archegoniophore stalk always elongates in *Marchantia* regardless of whether fertilization occurs (Shimamura, 2016).

Female sexual organs, archegonia (Fig. 6F), develop at the base of individual digitate rays (except between the first-two formed digitate rays) of archegonial receptacle (Fig. 6C, 6E). In a young archegoniophore, the neck of archegonia protrudes from the margin of archegonial receptacle towards the periphery. According to the rule, the neck cells of archegonia are formed prior to the division of SCC (secondary central cell having the potential to divide into egg cell and ventral canal cell) (Durand, 1908;

Shimamura, 2016). In the mature archegoniophore, the archegonia are arranged in the order of developmental stage with the youngest archegonia near the stalk and the older archegonia along the periphery (Fig. 6E). The archegonia mature in succession from the peripheral ones, ensuring that they are capable of fertilization by the time stalk elongation begins (Shimamura, 2016).

A protrusion of superficial cells, known as the archegonial initial (AI) extends outward and undergoes transverse division to form outer archegonial initial (OA) (or upper hemispherical cell) and basal archegonial initial (BA) (or primary stalk cell). Through three asymmetrical divisions, OA is divided into primary axial cell (PA) and three primary peripheral cells (PPs) (or jacket initial cells) (Fig. 7A). PA divides to form the primary cover cell (PCV) and primary central cell (PCC), whereas, PPs divide to form six rows of cells consisting of upper neck initials (NIs) and lower ventral initials (VIs). NIs continue to divide and form archegonial neck and VIs undergo divisions to form a cavity for the egg known as the venter of archegonia (Fig. 7B). PCC undergoes further asymmetrical divisions to generate a secondary central cell (SCC) and a primary neck canal (PNC) cell. The SCC enlarges and divides asymmetrically to produce 2 daughter cells, a smaller ventral canal cell (VCC) and a large cell that ultimately matures into an egg cell (EC). The VCC gradually disorganizes into mucilaginous strands and degenerates along-with neck canal cells to make a path for sperm cells to reach the egg. While the EC is maintained in the cytoplasm-rich form in the archegonial venter, several times VCC remains in the venter in a cytoplasm-rich form similar to the egg (Fig. 7C). The names of each cell of archegonia follow description published by (Shimamura, 2016; Ren et al., 2024).

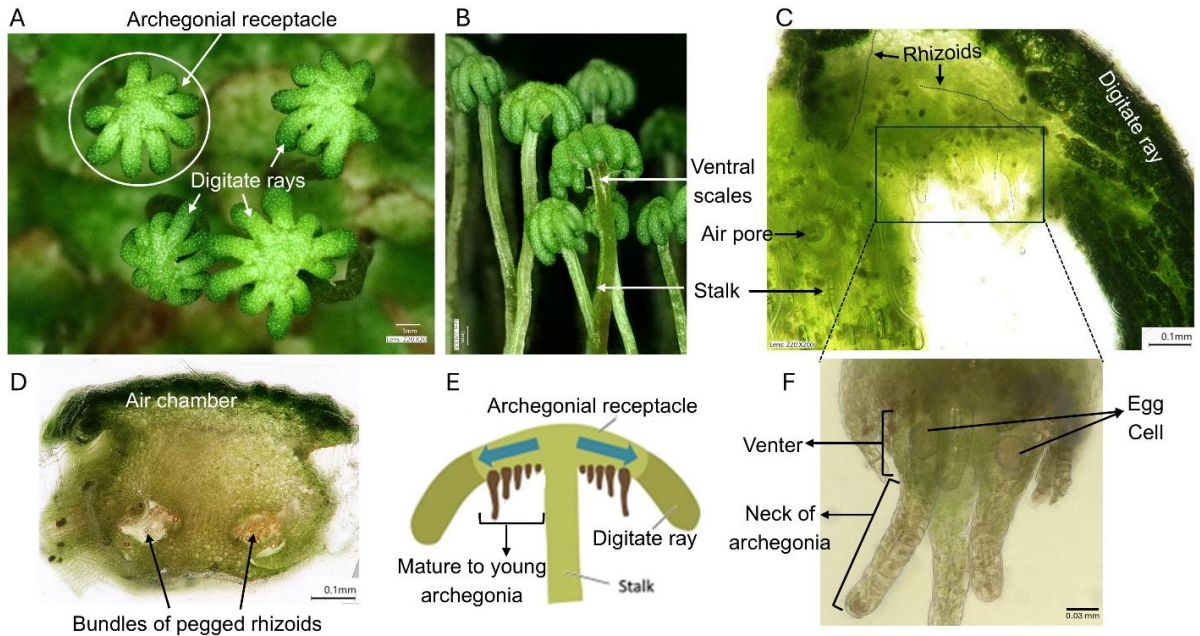


Figure 6. Female gametangiophore - Archegoniophore. (A) Microscopic image of the top-view of mature archegoniophores. (B) Microscopic image of the side-view of mature archegoniophores. (C) Longitudinal section of an archegonial receptacle. (D) Transverse section of the stalk of the archegoniophore. (E) Schematic drawing of an archegoniophore showing the order of developmental stage of archegonia (adapted from Cui et al., 2023, with several modifications). (F) Microscopic image of mature archegonia containing a single egg cell developed in the venter of archegonium.

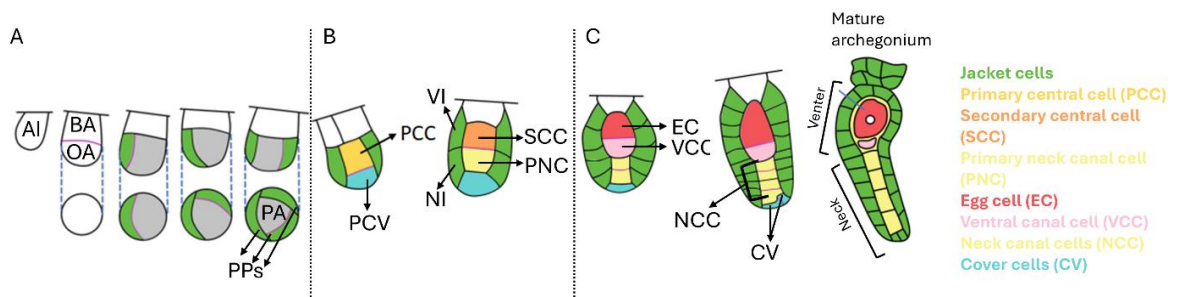


Figure 7. Schematic illustration of successive stages in developing archegonium (adapted from (Bao et al., 2024) with several modifications). Abbreviations: AI, archegonial initial; OA, outer archegonial initial; BA, basal archegonial initial; PA, primary axial cell; PPs, primary peripheral cells; PCC, primary central cell; PCV,

primary cover cell; NI, neck initials; VI, ventral initials; SCC, secondary central cell; PNC, primary neck canal; EC, egg cell; VCC, ventral canal cell; NCC, neck canal cells; CV, cover cells.

1.4.5 Development of antheridiophores and antheridia

The antheridiophores emerge from the apical meristem of the male vegetative thallus under long day conditions or when supplemented with far-red light in a laboratory environment (Fig. 8A, 8B). The mature antheridial disc (Fig. 8C) has similar anatomy as thallus in having air chambers, ventral scales and rhizoids. There are two different types of chambers on the dorsal side of the antheridial discs. Air chambers with barrel shaped pores similar to vegetative thallus and antheridial chambers, which contain a single antheridium (Fig. 8E). The stalk of antheridiophore also contains two grooves of rhizoids, but, unlike the archegoniophore stalk, there are no air chambers on the antheridiophore stalk (Fig. 8F) (Shimamura, 2016).

The male sexual organ, antheridium (plural - antheridia) develop in the cavity on the upper side of the antheridial disc. They are arranged in the order of developmental stage inside the antheridial disc, with the youngest one near the margin and the older ones near the center of the antheridial disc (Fig. 8D). The development of antheridia also begins from antheridial initial cell (AIC) which undergoes division to form two cells: outer antheridial cell (OAC) which will develop into antheridium and basal antheridial cell (BAC) which will form the stalk later. OAC undergoes divisions to form a filament of 2-3 cells, which divide further by transverse divisions to form a young antheridium consisting of 2-3 tiers of four cells each (Fig. 9A). This is followed by periclinal divisions to form inner spermatogenous cells (SCs) and an outer layer of jacket cells. Male gametogenesis is initiated when SCs are specified. The SCs undergo nearly-synchronous multiple cell divisions and differentiate into cuboidal-shaped spermatid mother cells (SMCs) resulting in the increased volume of antheridia (Fig. 9B). Each SMC divides diagonally to produce immature spermatids that eventually differentiate into biflagellate motile sperms (Fig. 9C) (Durand, 1908; Shimamura, 2016).

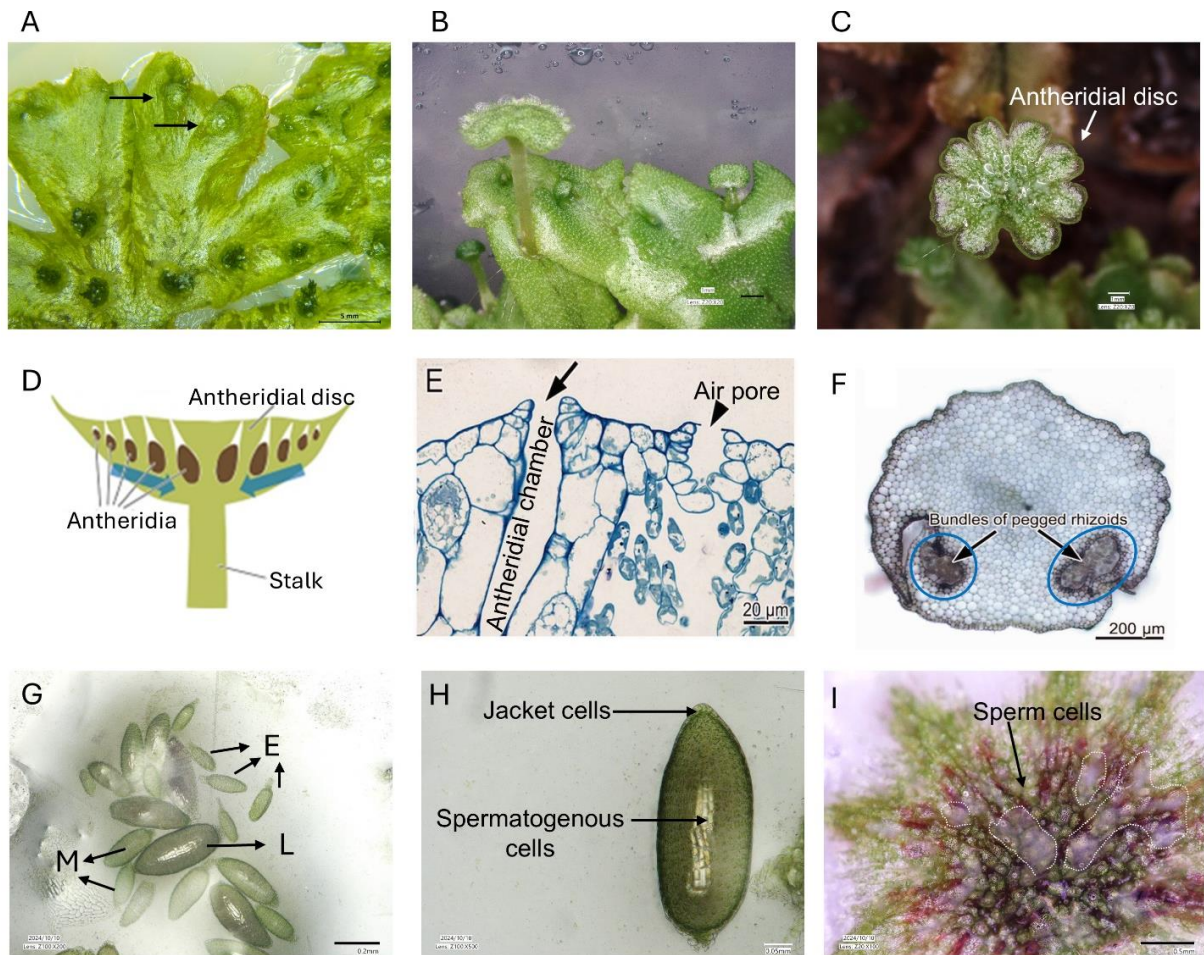


Figure 8. Male gametangiophore - Antheridiophore. (A) Microscopic image of the developing antheridiophores. (B) Microscopic image of the side-view of developing antheridiophores from apical notches of male thallus. (C) Microscopic image of top view of antheridial disc. (D) Schematic drawing of an antheridiophore showing the order of developmental stage of antheridia. (E) Vertical section of antheridial pore and air pore (adapted from Shimamura, 2016, with modifications). (F) Transverse section of the stalk of the antheridiophore [adapted from (Shimamura, 2016) with modifications]. (G) Antheridia manually dissected from antheridial disc. E – early, M – middle, L – late stage of antheridia development. (H) Microscopic image of an antheridium. (I) Sperm discharge on the antheridial disc. The dotted lines show the mucilaginous mass containing sperm cells released on the surface of antheridial disc through antheridial cavity.

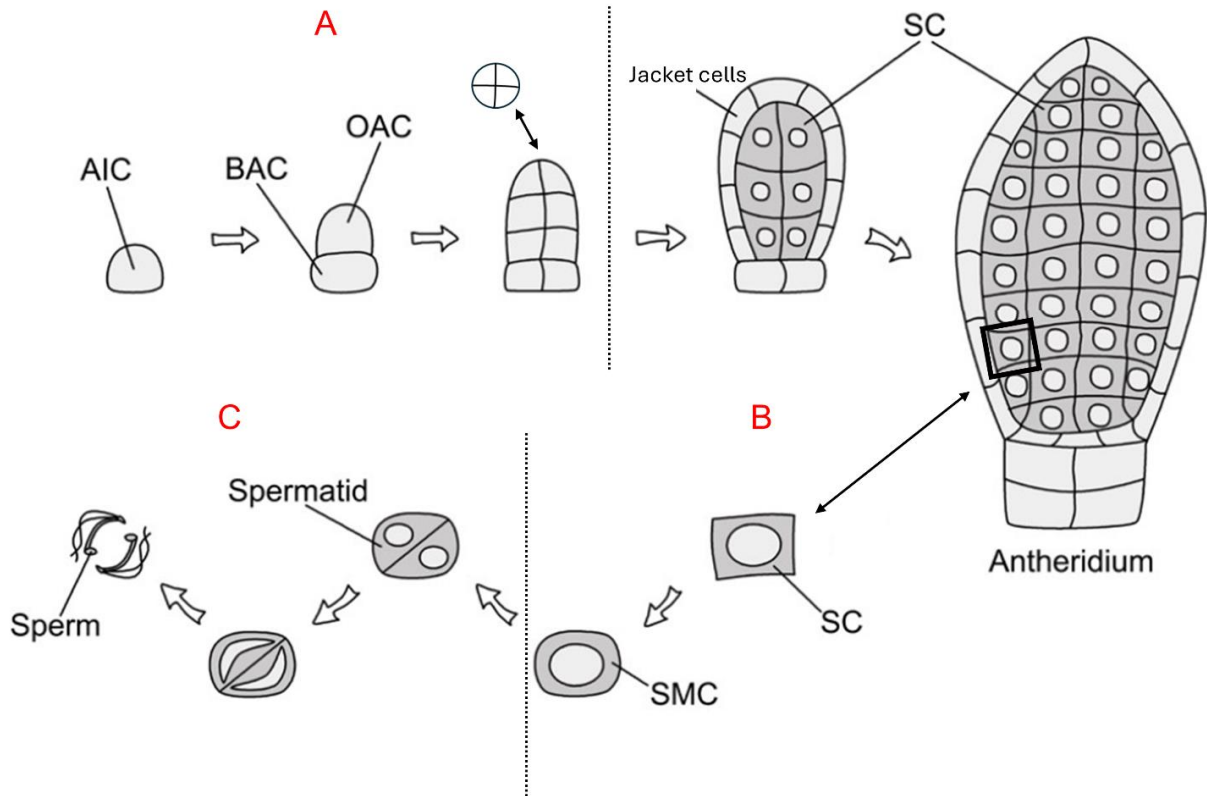


Figure 9. Schematic illustration of antheridium development. [The image has been taken from (Ren et al., 2024) with several modifications]. Abbreviations: AIC, antheridial initial cell; OAC, outer antheridial cell; BAC, basal antheridial cell; SC, spermatogenous cells; SMC, spermatid mother cells.

1.4.6 Fertilization and sporophyte development

The cytosolic calcium levels in the tip cells of antheridia induce programmed cell death (PCD). The sperms burst out through these degenerated tip cells into water and flow onto the antheridial disc surface (Fig. 10A, see Fig. 8I) (Cao et al., 2024). At the time of release, sperms are immotile as they are encapsulated in the cell wall remnants. Soon after the exposure to external water, the flagella starts to beat making the sperm rotate inside its capsule. With time, the flagellar beating becomes vigorous and sperms finally break out of their capsules and hatch (Shimamura, 2016). The motile sperms reach the archegonial receptacle, where they fertilize the egg-cells within archegonia. The tip of the mature archegonial neck opens upon contact with water and extrudes a mucilaginous mass making a way for sperm cells to enter the archegonium

(Shimamura, 2016). The sperms swim towards the egg cell by detecting an attractant released from the egg cell itself or its surrounding cells (Fig. 10B). The chemical identity of these attractants has not been elucidated yet. Although, the mucilaginous mass released from archegonia is considered to contain material for sperm attraction. The mechanism of polyspermy block remains unknown (Shimamura, 2016).

After fertilization, the diploid zygote develops into a multicellular embryo, which produces three different tissues: foot, seta, and capsule after 2-3 weeks of fertilization (Fig. 10C). The capsule (hanging upside down under archegonial receptacle) contains the sporogenous tissue from which two cell types differentiate: spore mother cells and elaters. The spore mother cells undergo meiosis to produce thousands of haploid spores. The elaters undergo PCD and help in the spore dispersal by hygroscopic movements (Fig. 10D). Furthermore, these spores germinate under favorable conditions and develop into mature thalli (Bowman, 2016; Shimamura, 2016; Bowman et al., 2022).

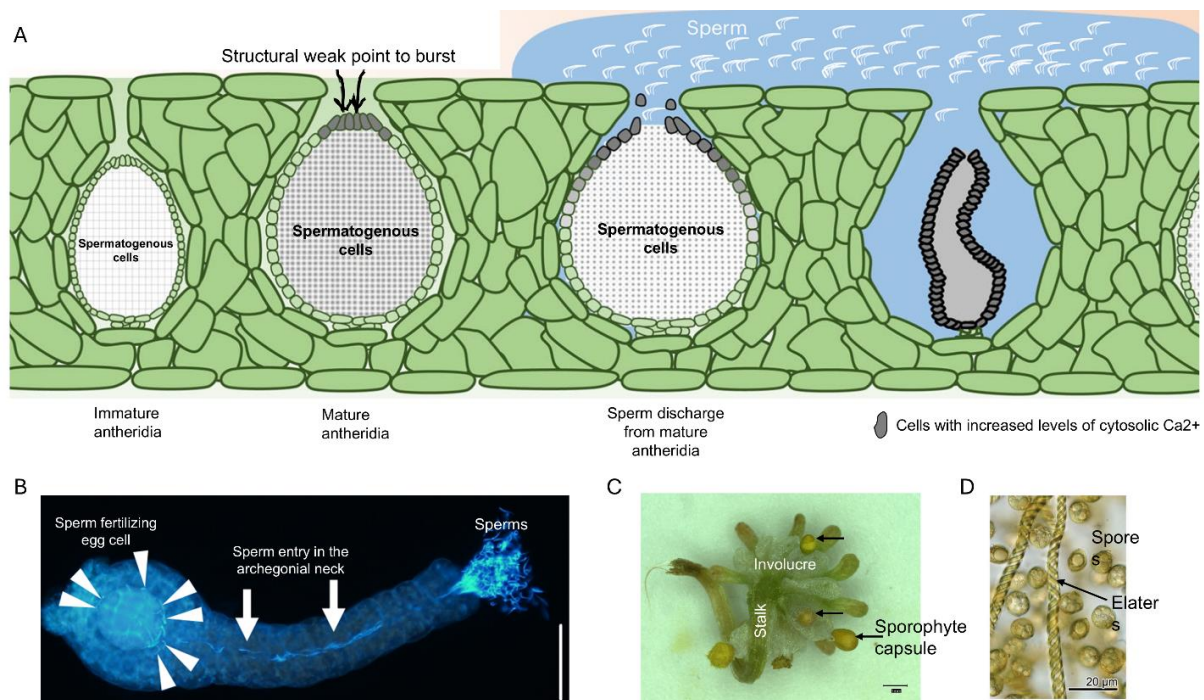


Figure 10. Release of sperm cells, fertilization and sporophyte formation. (A) Schematic illustration of sperm discharge from antheridial tip cells (adapted from Cao et al., 2024, with modifications). (B) Microscopic image of sperm entry in the archegonia [adapted from (Yamamoto et al., 2024) with modifications]. (C) Microscopic

image of archegoniophore post-fertilization with yellow capsules exposed to the outside by seta elongation. (D) Microscopic image of spores and elaters (adapted from Shimamura, 2016 with modifications).

1.5 Regulators of gene expression throughout plant development: miRNAs

Proper sexual organ development and successful reproduction requires precise regulation of gene expression throughout a plant's life cycle. This regulation is controlled by complex genetic networks that integrate the external environmental signals with the endogenous cues to guide development. These networks operate through multiple, interconnected layers of gene regulation, collectively shaping the overall growth and developmental trajectory of the plants. One of the key regulatory mechanisms by which transcription factors and a variety of other genes are regulated, is through the activity of small non-coding RNAs (ncRNAs) (Bartel, 2004; Voinnet, 2009). These ncRNAs even contributed to key innovations that enabled plants to survive and thrive in the terrestrial environments 450-470 million years ago (Fattash et al., 2007; Arif et al., 2013).

Among other classes of ncRNAs, miRNAs have emerged as fundamental players in regulating gene expression and overall plant developmental mechanisms (Bajczyk et al., 2023). miRNAs (18-24 nucleotides in length) are conserved post-transcriptional regulators of gene expression not only in plants but also in metazoans (Bartel, 2018). In fact, they were first identified in the nematode, *Caenorhabditis elegans* and the paradigm for miRNA function comes from the study of lin-4 and let-7 miRNAs which controlled larval developmental transitions (Lee et al., 1993; Olsen & Ambros, 1999; Reinhart et al., 2000). They control gene expression at the post-transcriptional level either through mRNA cleavage or by translational inhibition and this function of miRNAs is based on the sequence specificity and complementarity to their target mRNAs (Brodersen et al., 2008; Rogers & Chen, 2013).

1.5.1 miRNAs are generated via a complex biogenesis pathway in plants

Plants and animals share numerous similarities in the mechanisms of miRNA action, but there are substantial differences in their biogenesis pathways. The biogenesis of canonical plant miRNAs is a multistage process happening entirely in the cell nucleus (Bajczyk et al., 2023). Most miRNA genes (*MIR* genes) in plants, represent independent transcriptional units (Szarzynska et al., 2009; Pietrykowska et al., 2022) whereas some of them can be found encoded within the introns of protein coding or non-coding genes (Knop et al., 2017; Stepień et al., 2017).

MIR genes are mainly transcribed by RNA polymerase II (RNA Pol II) (Xie et al., 2005; Kim et al., 2011) generating a long primary-miRNA (pri-miRNA) transcript which holds the miRNA/miRNA* (miRNA strand/passenger miRNA strand) duplex. Pri-miRNAs are 5'-capped and 3'-polyadenylated (Jones-Rhoades & Bartel, 2004; Xie et al., 2005) and fold into a hairpin structure which is recognized by a multiprotein complex: the microprocessor. While assisted by numerous proteins, most importantly, the microprocessor core complex (DCL1, SE, and HYL1) cleaves the pri-miRNAs (Fig. 11) (Park et al., 2002; Han et al., 2004; Yang et al., 2006). DCL1 (DICER-LIKE1) protein initially cleaves the pri-miRNAs resulting in the release of precursor-miRNAs (pre-miRNAs) which also have a characteristic stem-loop structure (Kurihara & Watanabe, 2004). Since DCL proteins have dominant function for cleaving pri-miRNAs and producing small RNAs (sRNAs), their complete knock-out (KO) has been found to be lethal in *Arabidopsis* (Schauer et al., 2002; Nodine & Bartel, 2010). SE (SERRATE), a zinc finger protein whose KO mutation is also lethal and HYL1 (HYPONASTIC LEAVES1), a double stranded RNA-binding protein, are mainly needed for efficient and accurate processing of pri-miRNAs (Han et al., 2004; Yang et al., 2006; Bielewicz et al., 2023; Jozwiak et al., 2023). Eventually, the pre-miRNA is cleaved into miRNA/miRNA* duplex, which consists of a guide miRNA strand that is complementary to the target mRNA, and the passenger miRNA* strand, which is degraded in majority of the cases (Bologna et al., 2009). The miRNA/miRNA* duplex is 2'-O-methylated at 3'-ends by HEN1 (HUA ENHANCER1 METHYLASE) for protection against

degradation (Plotnikova et al., 2013; Yu et al., 2005). The mature and methylated miRNA is loaded into AGO1 (ARGONAUTE1) in the nucleus and exported to the cytosol as AGO1-miRNA complex (also known as RNA-induced silencing complex) via CRM1/EXPO1 exportin (CHROMOSOMAL MAINTENANCE1/ EXPORTIN1) (Bologna et al., 2018).

Marchantia also possesses this miRNA biogenesis machinery:

- Microprocessor core proteins [MpDCLs (MpDCL1a, MpDCL1b, MpDCL3), MpHYL1, MpSE]
- 44 Microprocessor auxiliary proteins (including MpHEN1, MpHEN2, MpDDL, MpCHR2/BRM)
- 19 Microprocessor regulatory proteins (including EXPORTIN1A/XPO1A, CRM1B/XPO1B, MpHST1)
- miRISC formation proteins (MpAGO1, MpAGO4, MpAGO10)

(Lin et al., 2016; Tsuzuki et al., 2016; Bowman et al., 2017; Lin & Bowman, 2018; Pietrykowska et al., 2022).

Most of the microprocessor core and auxiliary proteins reported in Arabidopsis were also identified in Marchantia (Pietrykowska et al., 2022). Although, these proteins have been well characterized in Arabidopsis, similar studies remain unexplored in Marchantia. However, functional studies of the MpDCL^{ge} mutants were reported recently. The authors observed defects in the 3D growth and reproductive organ formation in MpDCL^{ge} mutants reflecting disruptions in the regulatory sRNA dependent pathways (Csicsely et al., 2025). Also, another interesting study revealed that instead of miR168 regulating AtAGO1 in Arabidopsis, the two liverwort-specific miRNAs: Mpo-miR11707.1 and Mpo-miR11707.2 regulate MpAGO1 (Lin et al., 2016). Recently, Hong et al., 2024 studied the link between MpAGO1 and high-temperature stress. Under these conditions, Mpo-miR11707 incorporates into MpAGO1, regulating MpAGO1 mRNA levels and, therefore facilitate precise gene regulation for plant resilience to environmental stress (Hong et al., 2024).

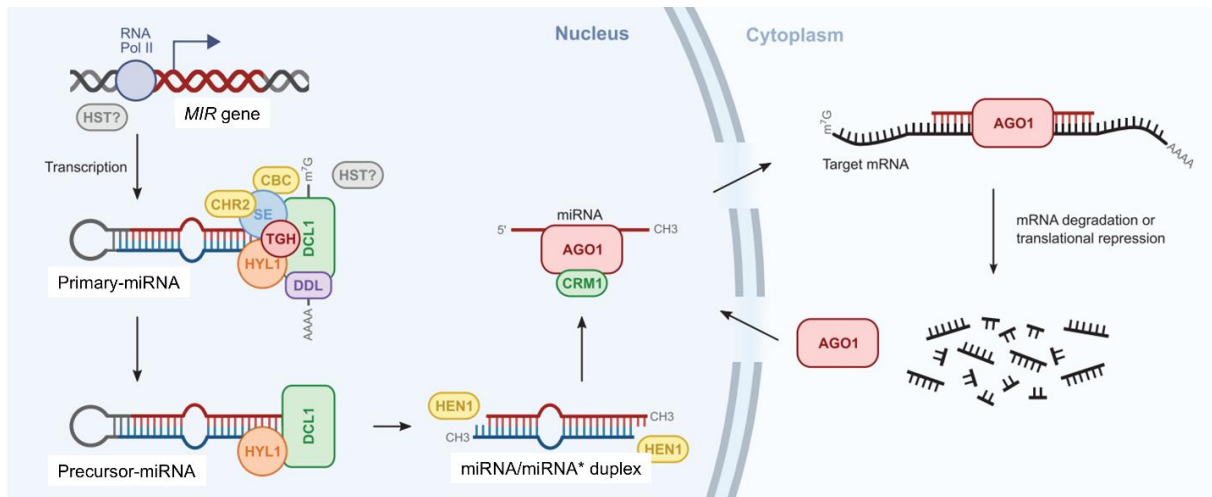


Figure 11. A diagram of miRNA biogenesis in plants. [adapted from Pietrykowska et al., (2022) with modifications]. Abbreviations: RNA POL II, RNA POLYMERASE II; MIR GENE, MICRORNA GENE; DCL1, DICER-LIKE1; HYL1, HYPONASTIC LEAVES1; SE, SERRATE; TGH, TOUGH; DDL, DAWDLE; CBC, CAP-BINDING PROTEIN COMPLEX, CHR2, CHROMATIN REMODELLING2; HST, HASTY; HEN1, HUA-ENHANCER1; CRM1, CHROMOSOMAL MAINTENANCE1; AGO1, ARGONAUTE1.

1.5.2 Mode of action of miRNAs

Plant miRNAs bind to their target mRNAs with near-perfect complementarity (Rhoades et al., 2002; Bartel & Bartel, 2003). For instance, miR165/166 directs the cleavage of PHB (PHABULOSA), PHV (PHAVOLUTA), and REV (REVOLUTA) mRNA target site (Rhoades et al., 2002) and it was observed that the cleavage rate is reduced 14-fold when an additional mismatch is incorporated into the target (Tang et al., 2003). This led to the hypothesis that the high degree of complementarity is required for the effective slicing of the mRNA target by AGO proteins (Bartel & Bartel, 2003; Tang et al., 2003). However, for instance, miR171 has perfect antisense complementarity to three mRNAs encoding SCARECROW-like (SCL) transcription factors (Llave, Xie, et al., 2002; Reinhart et al., 2002). In several cases, miRNAs identify and guide the cleavage of their target mRNAs through imperfect base pairing with the slice site located within the pair complementarity in the coding sequence or UTR regions. There are various examples identified where the number of miRNA-target mismatches are tolerated and even maintained through the evolution (Rhoades et al., 2002). Hence,

during computational target prediction, the regulatory targets are also identified by allowing a 1-3-nt gap in the search algorithm (Llave, Kasschau, et al., 2002; Rhoades et al., 2002; Bartel & Bartel, 2003; Tang et al., 2003). Simultaneously, there are several examples of targets which are regulated at protein level in the absence of noticeable changes in the mRNA level suggesting plant miRNAs interference via translational repression (Iwakawa & Tomari, 2013). Similar to angiosperms, miRNAs in *Marchantia* have cleavage sites located mainly in the coding sequence of target mRNA, with few examples of cleavage sites located in the 3'-UTR region, for instance, MpmiR11667.6/Mp1g06610 mRNA target, MpmiR11670.1/Mp8g00880 mRNA target (Lin et al., 2016; Aggarwal et al., 2024). At least conserved miRNAs show almost perfect complementarity with their conserved target mRNAs in *Marchantia* (Tsuzuki et al., 2016).

1.5.3 Vignette of miRNA/target mRNA modules steering *Marchantia*'s development

Although a vast body of research has investigated the roles of genes, transcription factors, and miRNA-mediated gene expression regulation in *Marchantia* development, this chapter specifically focuses on the current knowledge of miRNA/target mRNA modules that have been characterized as regulators in *Marchantia*'s developmental processes.

Extensive research across angiosperms has established the critical role of miRNAs in regulating a wide range of plant developmental and physiological processes. These include leaf morphogenesis, shoot apical meristem development, floral development, vegetative to reproductive phase transition, embryo development, stress responses, maintenance of nutrient homeostasis, hormone regulation, and signal transduction (Aukerman & Sakai, 2003; Palatnik et al., 2003; Achard et al., 2004; Baker et al., 2005; Mallory et al., 2005; Williams et al., 2005; Alonso-Peral et al., 2010; Liu & Vance, 2010; Huijser & Schmid, 2011). Gaining a deeper understanding of miRNA-mediated regulatory mechanisms in liverworts is essential for gaining insights into ancestral mechanisms of plant development and for understanding how early land plants adapted to terrestrial environment.

Almost a decade ago, sRNA sequencing provided insights into the miRNA repertoire of *Pellia endiviifolia* and *Marchantia polymorpha* (Alaba et al., 2015; Lin et al., 2016; Tsuzuki et al., 2016). In *Marchantia*, a total of 265 miRNA genes were identified (Lin et al., 2016; Tsuzuki et al., 2016). These *MIR* genes were found to be located within various genomic regions (5'-UTRs, coding regions, 3'-UTRs, introns, intergenic regions) (Lin et al., 2016; Tsuzuki et al., 2016; Bowman et al., 2017; Pietrykowska et al., 2022). However, deeper experimental studies provided evidences that many of the *MIR* genes represent independent transcriptional units (Aggarwal et al., 2024).

Since 2016, functional studies in *Marchantia* have primarily focused on conserved miRNAs, with very few examples of non-conserved (liverwort-specific) miRNAs explored to date (described below). While the conserved miRNAs tend to control the central developmental processes of ancestral origin (phase transition, dorsoventral patterning, cell cycle progression, cell specification, differentiation and proliferation), the liverwort-specific miRNAs could provide insights into lineage specific innovations and evolutionary processes (Bowman et al., 2022). Therefore, detailed functional characterization of these liverwort-specific miRNAs are required for understanding their precise functions in *Marchantia* development.

Nine conserved miRNA families (miR160, miR166, miR171, miR319, miR390, miR529/miR156, miR408, miR530/1030 and miR536), which are present in all other land plants, were found in *Marchantia*. These conserved miRNAs often have conserved targets (Lin et al., 2016; Tsuzuki et al., 2016; Lin & Bowman et al., 2018; For review (Pietrykowska et al., 2023). Here, I will explain the studies performed on conserved miRNAs and their targets in *Marchantia* (Fig. 12):

- I. A single genomic locus of Mp*MIR160* encodes a 246 nt stem-loop structure (Flores-Sandoval et al., 2016). miR160 regulates the expression of Mp*ARF3*, a class C ARF in *Marchantia* (Lin et al., 2016; Tsuzuki et al., 2016). Mp*ARF3* regulates a plethora of developmental processes including cell differentiation, proliferation and phase transition from asexual to sexual reproduction (Flores-Sandoval et al., 2018a; Flores-Sandoval et al., 2018b; Tsuzuki et al., 2016). The over-expression of Mpmir160 caused major aberrations in thallus surface area, apical notches, number of gemmae, air pores, rhizoids, ectopic antheridia on the dorsal surface of

antheridial disc, fused archegoniophores, and auxin insensitivity. Similar phenotypes were observed in *Mparf3* mutant plants (Flores-Sandoval et al., 2018a; Flores-Sandoval et al., 2018b). Moreover, *MpmiR160* was found to be upregulated in reproductive organs of *Marchantia*, suggesting their important role in sexual reproduction (Aggarwal et al., 2024). Hence, **MpMIR160-MpARF3** module influences multiple developmental processes in *Marchantia* and shows conserved activity across land plants (Flores-Sandoval et al., 2018a; Flores-Sandoval et al., 2018b).

- II. In the case of *miR165/miR166* family, two **MpMIR166** loci were identified in *Marchantia*: *miR166a* and *miR166b*. Most of the reads were obtained from *MIR166a* (Tsuzuki et al., 2016). *MpmiR166* expression was upregulated in both male and female vegetative and generative tissues in *Marchantia* (Aggarwal et al., 2024), suggesting its broad functional role in *Marchantia* development. The overexpression of *MpMIR166a* caused curling of vegetative thallus towards ventral side and improper development of gemma cups and downregulated the expression levels of its conserved target, *MpC3HDZ* (Floyd & Bowman, 2004; Tsuzuki et al., 2016). Recently, *MpC3HDZ* was reported as a marker of prothallus, regulating the dorsoventrality development and *de-novo* meristem formation (Wallner & Dolan, 2024).
- III. *SCARECROW-LIKE GRAS* transcription factors influence accumulation of chlorophyll. *MpGRAS10/MpSCL* has been shown to be a target of *miR171* in *Marchantia* (Lin et al., 2016; Lin & Bowman, 2018; Yelina et al., 2023). KO mutants of **MpMIR171** and *Mpscl* were indistinguishable from WT plants in chlorophyll content and chloroplast area, hence no impact on chloroplast development was observed in *Marchantia*. However, thallus morphology in *Mpscl* mutants was observed to be stunted with inward-curling edges suggesting the role of this module in vegetative development (Yelina et al., 2023).
- IV. *MpmiR319* is encoded within two loci in *Marchantia*: *MpmiR319a*, *MpmiR319b* (Tsuzuki et al., 2016) and they were shown to contribute differently to the formation of gemmae and gemma cup. Surprisingly, double KO of both loci did not result in enhanced developmental defects and RNA-seq analysis showed no significant

changes in the abundance of target transcripts between WT and double KO mutant. This may be due to the miR319 cleavage site being located within 3'-UTR of *MpRKD* mRNA, allowing 5'-fragment to persist in the cell and detected as intact mRNA. Functional evidence supporting the role of **MpmiR319/MpRKD** module in asexual reproduction came from the study miR319-resistant *MpRKD*, which showed defects in gemmae/gemma cup formation. Unexpectedly, miR319-resistant *MpMYB21* did not show such defects, despite confirmed cleavage of *MpMYB21* mRNA by miR319 in RACE assay (Tsuzuki et al., 2016; Futagami et al., 2025). Hence, the authors suggest that while miR319 family is conserved across land plants, its regulatory role is diverse.

- V. Another conserved miRNA, miR390 targets a single *MpTAS3* gene in *Marchantia*. *MpTAS3* contains 5' and 3' miR390 target sites that allows to produce complementary phased siRNAs similar to other land plants (Tsuzuki et al., 2016). *TAS3* is a highly conserved ncRNA locus which produces mature tasiRNAs: tasiR-ARFs and tasiR-AP2 in *P. patens* (Axtell et al., 2006), but, it was reported that *Marchantia*'s *TAS3* locus expresses only tasiR-AP2 (Krasnikova et al., 2013; Tsuzuki et al., 2016). miR390/*MpTAS3* module was further characterized and a new regulatory role of **miR390/MpTAS3/miR11700/MpCYP78A101** was discovered (Hung et al., 2024). First, the authors identified tasiARF target site in *MpARF2*, with the evidence of cleavage from the degradome data. Second, immunoprecipitation (IP) experiments with *MpAGO1* verified the loading of miRNAs and tasiRNAs as miR390 was enriched and low value for tasiARF was observed. Third, since *AtDCL4* is majorly involved in tasiRNA processing, similar results were proved for *MpDCL4* as *TAS3* tasiRNAs were present in Tak-1 WT plants, but not in CRISPR/Cas9 *dcl4^{ge}* mutants. Fourth, a novel tasiRNA, tasi78A was found to be significantly enriched in *MpAGO1*-IP profile. Interestingly, tasi78A had reverse complementary sequence to tasiAP2, but originating from another phasing position. Moreover, tasi78A showed high number of degradome reads at the target site in *MpCYP78A101* transcript. In addition to this, the authors also found another miRNA target site on *MpCYP78A101* mRNA for *MpmiR11700*, as supported by evidence from degradome and enrichment in *MpAGO1*-IP profile. The promoter assays

suggested overlapping expression patterns for miR390/MpTAS3 module, MpMIR11700, and MpCYP78A101 in the apical notch area of gemma. Furthermore, they observed upregulated levels of MpCYP78A101 mRNA in the MIR390^{ge} and downregulated levels of MpCYP78A101 in overexpression plants of tasi78A and miR11700, proving that miR390, tasi78A, and miR11700 coordinately regulate target expression (Hung et al., 2024).

- VI. The miRNA closely related to miR156 in angiosperms was found in Marchantia i.e. MpmiR529 (overlap sequence of 16-18 nt with miR156). miR529ab was predicted to target KANADI transcription factor and miR529c was proven to target MpSPL2 mRNA (Tsuzuki et al., 2016). Furthermore, functional analysis also showed that MpmiR529c represses MpSPL2 gene expression and regulates sexual reproductive development. Since the reproductive transition in Marchantia requires FR light, the Mpmir529c null mutant plants and plants expressing miR529c-resistant MpSPL2 developed gametangiophores in the absence of FR light stimulus, suggesting role of **MpmiR529c/MpSPL2** module in reproductive transition (Tsuzuki et al., 2019).
- VII. **MpmiR408** has two loci in Marchantia: miR408a and miR408b, which were predicted to target laccase and uclacyanin mRNAs, as the pairings between these mRNAs and miR408 are conserved in Marchantia (Tsuzuki et al., 2016). However, RACE analysis did not detect cleavage at the predicted sites, suggesting that this conserved miRNA might regulate a different target gene or that the silencing activity occurs at levels too low for detection (Tsuzuki et al., 2016). Moreover, 22 targets with significant reads in degradome were predicted for miR408a, which has two isoforms (Lin et al., 2016), but to date, no functional studies have been conducted with regards to its role in Marchantia development. The strong upregulation of MpmiR408ab in archegoniophores was reported (Aggarwal et al., 2024).
- VIII. Compared with *P. patens* having 10 loci for miR1030, **MpmiR1030** was found to have a single locus in Marchantia and it was predicted to target NB-LRR family protein in Marchantia (Tsuzuki et al., 2016). MpmiR1030 exhibited strong upregulation in archegoniophores in Marchantia (Aggarwal et al., 2024), but so far,

there is no experimental evidence demonstrating its functional role in archegoniophore or overall development of *Marchantia*.

- IX. **MpmiR536** is encoded at a single locus in *Marchantia*, in contrast to the 6 loci found in *P. patens*. It was predicted to target ABA insensitive 3 mRNA in *Marchantia* (Tsuzuki et al., 2016), however, no functional characterization has been reported to date.

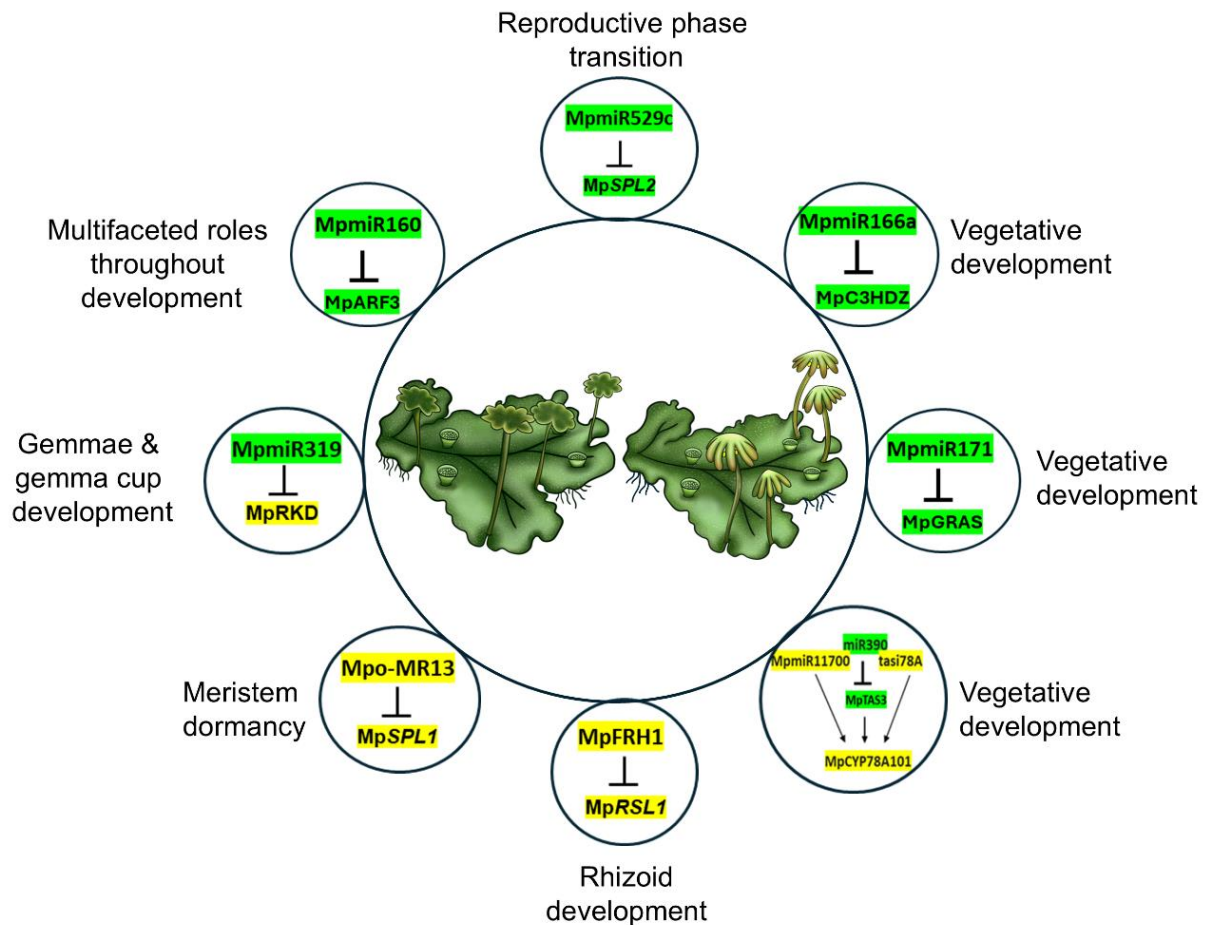


Figure 12. Experimentally validated miRNA/target mRNA modules involved in *Marchantia* development. The conserved miRNA/target mRNA modules have been highlighted in bright green, the unique miRNA/target mRNA modules have been highlighted in yellow. miR319 being conserved miRNA is highlighted in green but the target being unique is highlighted in yellow.

Studies on non-conserved miRNAs include (Fig. 12):

- I. *AtSPL8*, a homolog of *MpSPL1*, is not a target of any miRNA, but *Mpo-MR13* (*MpmiR11671*) was discovered as a unique miRNA targeting *MpSPL1* gene in *Marchantia* (Lin et al., 2016; Tsuzuki et al., 2016; Streubel et al., 2023). The dichotomous branching of *Marchantia* thallus is determined by the number of active and dormant meristems at the apical notch region. The meristem dormancy is promoted by transcription factor *MpSPL1*, as *Mpspl1* loss-of-function plants produced no dormant meristems. This activity of *MpSPL1* was further inhibited by *Mpo-MR13*, as *Mpspl* gain-of-function and *Mpo-mir13* loss-of-function plants produced more dormant meristems. Hence, **Mpo-MR13/MpSPL1** module was proven to be important for meristem dormancy. Additionally, *Mpo-MR13* is regulated by *MpPIF* mediated phytochrome signaling. In light conditions (high R:FR ratio), *MpPIF* is inactive, allowing the expression of the *Mpo-MR-13* miRNA, which further represses its target *MpSPL1*, preventing dormancy in daughter meristems. As a result, both daughter meristems grow into branches. In shade (low R:FR ratio), *MpPIF* becomes active and represses *Mpo-MR-13*, leading to the accumulation of *MpSPL1*. This causes one of the two daughter meristems to become dormant (Streubel et al., 2023).
- II. A liverwort-specific miRNA, *MpFRH1* (*FEW RHIZOIDS1*; also *MpMIR11681*), was discovered through a mutagenesis screen for genes involved in rhizoid development (Honkanen et al., 2018). *RSL* (*ROOT HAIR DEFECTIVE SIX-LIKE*) transcription factors are ancestral master regulators for the development of filamentous rooting structures (Pires et al., 2013). In *Marchantia*, *Mprsl1* loss-of-function plants and *Mpfrh1* gain-of-function plants produced very few or no rhizoids, empty gemma cups or rarely developed gemmae (Proust et al., 2016; Honkanen et al., 2018). Additionally, *Mpfrh1* loss-of-function plants developed more rhizoid cell clusters similar to *Mprsl1* gain-of-function transgenic plants. Moreover, higher levels of *MpRSL1* mRNA were found in *Mpfrh1* loss-of-function mutants (Thamm et al., 2020). Hence, **MpFRH1/MpRSL1** module was established in which *MpFRH1* miRNA regulates *MpRSL1* mRNA by mediating its cleavage and thus, acts as a repressor of rhizoid cell development in *Marchantia* (Honkanen et al., 2018).

Thamm et al., 2020 also reported a unique mechanism of lateral inhibition specific to liverworts mediated by MpFRH1 during epidermal cell patterning in Marchantia. Since higher expression levels of Mp*RSL1* favor the specification of an epidermal cell to be specified as a rhizoid cell, Mp*RSL1* induces the expression of MpFRH1 which further acts as brake, preventing the neighboring cells from becoming rhizoid cells and hence, these neighboring cells develop as flat epidermal cells (Thamm et al., 2020).

Taken together, reports from several studies have shed light on the importance of conserved as well as non-conserved miRNA/target mRNA modules in Marchantia's developmental processes.

Chapter 2: Aim of the thesis

miRNAs perform pleiotropic functions in various processes of plant development. Liverwort-specific miRNAs revealed differential expression in *Marchantia* organs. Some of them showed enrichment / exclusive accumulation in the *Marchantia* sexual organs. In the research presented in this thesis, I hypothesized that these liverwort-specific miRNAs are involved in the reproductive development of *Marchantia*.

Following tasks were performed to verify my hypothesis:

1. To identify selected liverwort-specific *MIR* gene structures, expression patterns of their pri-miRNAs and their mRNA target prediction.
2. To select the most promising liverwort-specific miRNA-mRNA target modules for in-depth analysis.
3. To generate GUS expressing plants under selected liverwort-specific *MIR* gene promoters.
4. To perform CRISPR/Cas9 gene editing for the generation of knock-out mutants of selected *MIR* genes.
5. To perform the phenotypic characterization of obtained CRISPR/Cas9 KO mutants.
6. To define the functional roles of selected miRNAs in vegetative and reproductive development.

Chapter 3: Materials

3.1 Plant material, propagation and maintenance

M. polymorpha ssp. *ruderalis* male accession Takaragaike-1 (Tak-1) and female accession Takaragaike-2 (Tak-2) were used for performing all experiments. The Tak-1 and Tak-2 samples were brought from Prof. Takayuki Kohchi laboratory (Graduate School of Biostudies, Kyoto University, Japan). Gemma or apical part of the thallus was used for the propagation of plant cultures. For *in-vitro* culture: *M. polymorpha* was grown in axenic cultures in petri dishes (Sarstedt, Germany) and polypropylene boxes (Duchefa Biochemie, Netherlands) containing solid half-strength Gamborg medium (Gamborg et al., 1976) (Duchefa Biochemie, Netherlands) supplemented with 1% (w/v) sucrose (Merck, USA), and with 0.8% (w/v) agar medium (BioShop, Canada) and pH adjusted to 5.5 – 5.6 with 0.1 M KOH. The cultures were placed at 22°C under continuous white light. To induce gametangiophore development, thalli were transferred to high-power blue irradiation (470 nm) as described in (Aggarwal B et al., 2024). For soil culture: Thalli were propagated on Jiffy-7® (Jiffy International AS, Netherlands) 42 mm peat pellets and were kept under similar conditions as *in-vitro* cultures.

3.2 Bacterial strains: Culture and growth parameters

Escherichia coli DH5α competent cells were used for cloning and plasmid amplification. In the case of destination vector for gateway cloning, *E. coli* DB3.1 cells were used for plasmid amplification. Both *E. coli* DH5α and DB3.1 cells were cultured on LB (Luria-Bertani) medium supplemented with 1.5 % agar at 37 °C.

Agrobacterium tumefaciens GV3101 competent cells were used for transformation of *M. polymorpha* sporelings. *A. tumefaciens* was cultured for 2-3 days at 28 °C in LB medium containing rifampicin, gentamicin (This *A. tumefaciens* strain is resistant to these antibiotics) and spectinomycin (plasmid propagation).

The competent cells for *E. coli* and *A. tumefaciens* were always ready to use in the Department of Gene Expression.

3.3 Culture media and reagents

3.3.1 Bacterial culture media:

LB Broth media composition

| Component | Final concentration | Amount |
|---|---------------------|--------------|
| Bacto-Tryptone (BioShop, Canada) | 1 % | 5 g |
| Yeast extract (Carl Roth GmbH, Germany) | 0.5 % | 2.5 g |
| NaCl (Merck, USA) | 1 % | 5 g |
| H ₂ O MilliQ | - | Up to 500 ml |

The medium was autoclaved at 121 °C for 20 min. For LB agar medium, 7.5 g agar (BioShop, Canada) was added and autoclaved. After cooling to ~50 °C, appropriate antibiotics were added, medium was poured into sterile petri dishes and kept at 4 °C until use.

3.3.2 Antibiotics:

| Antibiotics | Manufacturer | Stock concentration | Working concentration |
|---------------|-------------------------------|---------------------|-----------------------|
| Ampicillin | BioShop, Canada | 100 mg/ml | 100 µg/ml |
| Gentamicin | Merck, USA | 20 mg/ml | 20 µg/ml |
| Kanamycin | BioShop, Canada | 50 mg/ml | 50 µg/ml |
| Rifampicin | BioShop, Canada | 50 mg/ml | 100 µg/ml |
| Spectinomycin | BioShop, Canada | 50 mg/ml | 50 µg/ml |
| Cefotaxime | Merck, USA | 100 mg/ml | 50 µg/ml |
| Hygromycin B | Thermo Fisher Scientific, USA | 50 mg/ml | 10 µg/ml |

Others:

X-Gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside) solution (BioShop, Canada) [20 mg/ml, dissolved in N, N-dimethylformamide]. The solution was then stored at 4°C in a dark bottle.

100 mM acetosyringone (Merck, USA) was dissolved in Dimethyl sulfoxide (DMSO) (Merck, USA) and stored at -20°C until further use.

3.3.3 Plant culture media:

Preparation of OM51C medium

A) 1000X Gamborg's B5 micro-elements stock:

| Components | Final concentration | Amount |
|---|---------------------|--------------|
| Na ₂ MoO ₄ X 2H ₂ O (Merck, USA) | 1 mM | 25 mg |
| CuSO ₄ X 5H ₂ O (Merck, USA) | 0.01 mM | 2.5 mg |
| CoCl ₂ X 6H ₂ O (Merck, USA) | 0.01 mM | 2.5 mg |
| ZnSO ₄ X 7H ₂ O (Merck, USA) | 6 mM | 200 mg |
| MnSO ₄ X H ₂ O (Merck, USA) | 50 mM | 1 g |
| H ₃ BO ₃ (Merck, USA) | 40 mM | 300 mg |
| H ₂ O MilliQ | - | Up to 100 ml |

To completely dissolve all of the reagents, the components were added in the exact order stated and autoclaved at 121 °C for 20 minutes.

B) Gamborg's B5 vitamin mix:

| Components | Final concentration | Amount |
|-----------------------------|---------------------|---------------|
| myo-Inositol (Merck, USA) | 554.94 μ M | 100 mg |
| Nicotinic Acid (Merck, USA) | 8.12 μ M | 1 mg |
| Pyridoxine HCl (Merck, USA) | 4.86 μ M | 1 mg |
| Thiamine HCl (Merck, USA) | 29.65 μ M | 10 mg |
| H ₂ O MilliQ | - | Up to 1000 ml |

The components were thoroughly mixed, following which the solution was subjected to autoclaving at a temperature of 121 °C for a duration of 20 minutes.

C) 10X OM51C Medium Stock:

| Components | Final Concentration | Amount |
|--|---------------------|---------------|
| CaCl ₂ X 2H ₂ O (Merck, USA) | 20 mM | 3 g |
| EDTA (Thermo Fisher Scientific, USA) | 1 mM | 0.4 g |
| KH ₂ PO ₄ (Merck, USA) | 20 mM | 2.75 g |
| KNO ₃ (Merck, USA) | 200 mM | 20 g |
| MgSO ₄ X 7H ₂ O (Merck, USA) | 15 mM | 3.7 g |
| NH ₄ NO ₃ (Merck, USA) | 40 mM | 4 g |
| Gamborg's B5 micro-elements | - | 10 ml |
| Gamborg's B5 Vitamin mix | - | 10 ml |
| 0.075% KI solution (Merck, USA) | 0.00075 % | 10 ml |
| H ₂ O MilliQ | | Up to 1000 ml |

It was ascertained that the components had been adequately mixed. Aliquots of 50 ml were stored at -20 °C until further use.

OM51C media components:

| Components | Final concentration | Amount |
|---------------------------------|---------------------|--------------|
| 10X OM51C medium stock | 1X | 50 ml |
| Sucrose (Merck, USA) | 110 mM | 20 g |
| Casein hydrolysate (Merck, USA) | 0.2 % | 1 g |
| L-glutamine (Merck, USA) | 4 mM | 0.3 g |
| H ₂ O MilliQ | - | Up to 500 ml |

The pH was adjusted to 5.5 using 1 N KOH. For the purpose of preparing agar plates, 5 g of agar was added. The media was then autoclaved at 120 °C for 20 minutes.

3.3.4 Nucleic acid molecular weight markers

- A) GeneRuler™ 100 bp Plus DNA Ladder (Thermo Fisher Scientific, USA)
- B) GeneRuler™ 1 kb Plus DNA Ladder (Thermo Fisher Scientific, USA)

3.4 Solutions and Buffers

3.4.1 Electrophoresis solution:

10X TAE Buffer

| Components | Final concentration | Amount |
|--------------------------------------|---------------------|---------------|
| Tris-Base (Merck, USA) | 400 mM | 48.5 g |
| Glacial Acetic Acid (Merck, USA) | 200 mM | 11.4 ml |
| EDTA (Thermo Fisher Scientific, USA) | 10 mM | 20 ml |
| H ₂ O MilliQ | - | Up to 1000 ml |

The buffer was diluted to a concentration of 1X for further use.

3.4.2 Agarose gel

| Components | Final concentration | Amount |
|-----------------------------------|---------------------|---------|
| Agarose (Prona Agarose, Spain) | 1-1.5 % | 1-1.5 g |
| 1X TAE | 1X | 100 ml |

The concentration of agarose in the gel was determined by the size of DNA fragments to be separated. The appropriate amount of agarose and TAE buffer was added to the 250 ml flask, which was then boiled in a microwave until completely dissolved. The solution was then cooled down to ~50 °C before the addition of 1 % Ethidium Bromide solution (3 µl / 50 ml) (Merck, USA). For transparent samples, DNA was mixed with a 6x orange loading buffer.

6X orange loading dye was prepared as follows:

| Components | Final concentration | Initial concentration | Amount per 10 ml |
|--|---------------------|-----------------------|------------------|
| Tris HCl (pH=7.6) (Merck, USA) | 10 mM | 1 M | 0.1 ml |
| Orange G (Avantor Performance Materials, Poland) | 0.33 % | 1 % | 3 ml |
| Xylene cyanol | 0.03 % | 1 % | 0.3 ml |
| Glycerol | 60 % | 99 % | 6 ml |
| EDTA (Thermo Fisher Scientific, USA) | 60 mM | 1 M | 0.6 ml |

The component mixture was autoclaved and transferred to 0.5 ml tubes.

3.4.3 DNA extraction buffer for genotyping

A) 1 M Tris (pH = 9.5) buffer:

| Components | Final concentration | Amount |
|-------------------------|---------------------|-------------|
| Tris base (Merck, USA) | 1 M | 6.05 g |
| H ₂ O MilliQ | - | Up to 50 ml |

The pH was adjusted to 9.5 using 5 M HCl.

B) DNA Extraction buffer:

| Components | Final concentration | Amount |
|---|---------------------|--------------|
| 1 M Tris; pH 9.5 | 0.1 M | 10 ml |
| 0.5 M EDTA; pH 8.0 (Thermo Fisher Scientific, USA) | 10 mM | 2 ml |
| 3 M KCl (Merck, USA) | 1 M | 33.2 ml |
| H ₂ O MilliQ | - | Up to 100 ml |

All components were mixed and subsequently filtered using MILLEX-HP filters (Ø 0.45 µm) (MilliporeSigma, USA).

3.4.4 RNA isolation solutions

DEPC-treated RNase free water: 1 ml of Diethyl Pyrocarbonate (DEPC) (Merck, USA) was added to 1 L of MilliQ H₂O, agitated vigorously, and then incubated for the entire night under a fume hood. The DEPC-treated water was autoclaved twice the following day for 20 minutes at 121 °C.

Trizol-like reagent (Chomczynski & Sacchi, 1987) ready to use was prepared in the Department of Gene Expression.

0.5 M EDTA: The appropriate amount of EDTA (14.51 g for 100 ml) (Merck, USA) was weighed and added to about 80 ml of MilliQ water. The pH was adjusted to 8.0 with

NaOH pellets to increase the solubility. The volume was adjusted to 100 MilliQ water and stored at room temperature or 4 °C for longer storage.

10 % (w/v) N-Lauroylsarcosine Sodium salt solution (sarkosyl) (Merck, USA) for sRNA enrichment: 5 g of sarkosyl was weighed and added to 50 ml of MilliQ water in a 50 ml falcon. The solution was stored at 4 °C for further use.

2X RNA loading dye was prepared as follows:

| Components | Final concentration | Initial concentration | Amount per 10 ml |
|--------------------------------------|---------------------|-----------------------|------------------|
| Tris HCl (pH = 7.6) (Merck, USA) | 10 mM | 1 M | 0.1 ml |
| Bromophenol Blue | 0.025 % | 1 % | 250 µl |
| Xylene cyanol | 0.025 % | 1 % | 250 µl |
| Formamide | 94 % | 99 % | 9.4 ml |
| EDTA (Thermo Fisher Scientific, USA) | 0.5 mM | 1 M | 5 µl |

The component mixture was autoclaved and stored in -20 °C. Finally, 2X RNA loading dye was mixed in equal amounts with RNA samples before running on Agarose gel electrophoresis and northern blot gel.

3.4.5 GUS staining solutions

A) Phosphate Buffer (pH = 7.2) components:

| Components | Final concentration | Amount |
|---|---------------------|--------------|
| 1 M NaH ₂ PO ₄ (Merck, USA) | 0.29 M | 29 ml |
| 1 M Na ₂ HPO ₄ (Merck, USA) | 0.72 M | 72 ml |
| H ₂ O MilliQ | - | Up to 100 ml |

The component mix was filtered using MILLEX-HP filters (Ø 0.45 µm).

B) GUS premix solution components:

| Components | Final concentration | Amount |
|---|---------------------|--------------|
| 0.5 M NaPO ₄ (pH = 7.2) | 50 mM | 50 ml |
| 0.5 M EDTA (pH = 8.0) (Thermo Fisher Scientific, USA) | 100 mM | 10 ml |
| 10 % TritonX-100 (Merck, USA) | 0.1 % | 0.5 ml |
| H ₂ O MilliQ | - | Up to 500 ml |

The mixture was filtered using MILLEX-HP filters (Ø 0.45 µm) and stored in a dark bottle in 4 °C.

C) GUS staining solution components:

| Components | Final concentration | Amount |
|---|---------------------|-------------|
| GUS Premix solution | - | 9.8 ml |
| 100 mM K ₄ [Fe(CN) ₆] (Merck, USA) | 0.5 mM | 50 µl |
| 100 mM K ₃ [Fe(CN) ₆] (Merck, USA) | 0.5 mM | 50 µl |
| 100 mM X-Gluc (Thermo Fisher Scientific, USA) | 1 mM | 100 µl |
| H ₂ O MilliQ | - | Up to 10 ml |

GUS staining solution was prepared right before use. Potassium-ferrocyanide and -ferricyanide stocks were kept at 4 °C. X-Gluc was dissolved in N,N-Dimethylformamide (DMF) (Merck, USA) and stored at -20 °C.

3.4.6 Northern Blot Hybridization solutions

A) Acrylamide stock (40 %) for RNA gels

| Components | Final concentration | Amount |
|-------------------------------|---------------------|--------------|
| Acrylamide (BioShop, Canada) | 38 % (w/v) | 190 g |
| Bis-acrylamide (Merck, USA) | 2 % (w/v) | 10 g |
| DEPC-treated H ₂ O | - | Up to 500 ml |

The components were filtered with a Ø 0.45-µm filter and store at 4 °C in a dark bottle.

B) 20X MOPS-NaOH components:

| Components | Final concentration | Amount |
|--|---------------------|--------------|
| MOPS (Merck, USA) | 20X | 41.85 g |
| 1 M Sodium Acetate (Merck, USA) | 40 mM | 20 ml |
| 0.5 M EDTA (pH = 8.0) (Thermo Fisher Scientific, USA) | 20 mM | 20 ml |
| DEPC-treated H ₂ O | - | Up to 500 ml |

MOPS was dissolved in DEPC-treated water and the pH was adjusted to 7.0 with 5 N NaOH. Then, sodium acetate and EDTA were added. The buffer was stored at room temperature at a place that is not exposed to light.

C) dPAG (denaturing Polyacrylamide gel) components:

| Components | Final concentration | Amount |
|-------------------------------|---------------------|-------------|
| 40 % acrylamide (19:1) | 15 % | 22.5 ml |
| Urea (Merck, USA) | 15 % | 25.2 g |
| 20X MOPS-NaOH | 1X | 3 ml |
| DEPC-treated H ₂ O | - | Up to 60 ml |

The components were dissolved by heating at 75-100 °C until the solution becomes clear. The component mix was then filtered using MILLEX-HP filters (Ø 0.45 µm). 700

μl of Ammonium persulfate (APS) (BioShop, Canada) and 20 μl of Tetramethylethylenediamine (TEMED) (Merck, USA) were added to the gel and immediately poured between the glass plates. The gel was ready for use after 30 minutes.

D) EDC cross-linking reagent:

| Components | Final concentration | Amount |
|---------------------------------------|---------------------|-------------|
| 12.5 M 1-methylimidazole (Merck, USA) | 0.13 M | 245 μl |
| EDC (Merck, USA) | 0.16 M | 0.753 g |
| DEPC-treated H ₂ O | - | Up to 24 ml |

245 μl of 12.5 M 1-methylimidazole was added to 9 ml of DEPC-treated water. Then, 0.753 g of 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) was added right before usage and the volume was adjusted to 24 ml with DEPC-treated water.

E) Hybridization buffer:

| Components | Final concentration | Amount |
|---|---------------------|--------------|
| 1 M Na ₂ HPO ₄ (Merck, USA) | 375 mM | 37.5 ml |
| 1 M NaH ₂ PO ₄ (Merck, USA) | 125 mM | 12.5 ml |
| 10 % SDS (BioShop, Canada) | 3.5 % | 35 ml |
| DEPC-treated H ₂ O | - | Up to 100 ml |

The buffer was filtered using MILLEX-HP filters (Ø 0.45 μm) and stored at room temperature and heated to 37 °C before use.

F) 20X SSC buffer:

| Components | Final concentration | Amount |
|--------------------------------------|---------------------|---------------|
| NaCl (Merck, USA) | 17.5 % | 175.3 g |
| Sodium citrate tribasic (Merck, USA) | 8.8 % | 88.2 g |
| DEPC-treated H ₂ O | - | Up to 1000 ml |

The buffer was filtered using MILLEX-HP filters (Ø 0.45 µm) and stored at room temperature.

G) Washing buffer:

| Components | Final concentration | Amount |
|-------------------------------|---------------------|---------------|
| 20X SSC buffer | 2X | 100 ml |
| 10 % SDS | 1 % | 10 ml |
| DEPC-treated H ₂ O | - | Up to 1000 ml |

The buffer was filtered using MILLEX-HP filters (Ø 0.45 µm) and stored at room temperature.

3.5 Kits

| Kit | Manufacturer | Objective |
|--|-------------------------------|---|
| GeneJET Plasmid Miniprep kit | Thermo Fisher Scientific, USA | Plasmid isolation |
| TURBO DNA-free™ Kit | Thermo Fisher Scientific, USA | DNA digestion for RNA purification |
| GeneRacer™ Core Kit | Thermo Fisher Scientific, USA | cDNA preparation for 5'-RLM, 5'-RACE and 3'-RACE |
| Direct-zol™ RNA Mini prep | Zymo Research, USA | RNA isolation |
| Ribominus™ Plant Kit for RNA seq | Thermo Fisher Scientific, USA | Depletion of rRNA for RNA seq library preparation |
| High sensitivity DNA reagents | Agilent Technologies, USA | Quality control analysis of NGS libraries |
| Qubit™ 1X dsDNA HS Assay tube | Thermo Fisher Scientific, USA | Quantification of NGS libraries |
| NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760, E7765) | New England BioLabs, USA | NGS library preparation |

| | | |
|--|-------------------------------|---|
| GeneJET Plant Genomic DNA purification Mini kit | Thermo Fisher Scientific, USA | Genomic DNA isolation |
| GeneJET Gel Extraction and DNA Cleanup Micro Kit | Thermo Fisher Scientific, USA | DNA extraction from agarose gel and cleanup of PCR products |

3.6 Enzymes

| Enzyme | Manufacturer | Objective |
|---|-------------------------------|---|
| TURBO™ DNase (2 U/μl) | Thermo Fisher Scientific, USA | DNase treatment |
| CloneAmp HiFi PCR Premix | Takara, Japan | Gene amplification |
| 50X Advantage® 2 polymerase mix | Takara Japan | 5'-RLM, 5'-RACE, 3'-RACE reactions |
| DreamTaq DNA polymerase (5 U/μl) | Thermo Fisher Scientific, USA | Gene amplification and colony PCR reactions |
| SuperScript™ III Reverse Transcriptase (200 U/μl) | Thermo Fisher Scientific, USA | cDNA synthesis |
| Gateway™ LR Clonase™ II enzyme mix | Thermo Fisher Scientific, USA | Gateway cloning reactions |
| T4 DNA ligase (5 U/μl) | Thermo Fisher Scientific, USA | Ligation of insert and vector |
| RNasin® Plus Ribonuclease inhibitor (40 U/μl) | Promega, USA | Ribonuclease inhibitor during cDNA synthesis |
| T4 Polynucleotide Kinase (10000 U/ml) | New England BioLabs, USA | Radioactive labelling of oligonucleotide probes |
| KAPA 3G Plant DNA Polymerase (2.5 U/μl) | KAPA Biosystems, USA | Genotyping PCR reaction |
| Shrimp Alkaline Phosphatase (rSAP) (1000 U/ml) | New England BioLabs, USA | Exo-SAP reaction for direct sequencing of genotyping PCR products |
| Exonuclease I (ExoI) | New England | Exo-SAP reaction for direct |

| | | |
|---------------------------|--------------|---------------------------------------|
| (20,000 U/ml) | BioLabs, USA | sequencing of genotyping PCR products |
| GoTaq® qPCR Master Mix 2X | Promega, USA | RT-qPCR |

3.7 Vectors

| Vector | Manufacturer | Objective |
|----------------|---|--|
| pENTR™/D-TOPO® | Thermo Fisher Scientific, USA | Entry vector for gateway® based cloning |
| pGEM® T Easy | Promega, USA | Cloning 5'-RLM, 5'-RACE, 3'-RACE products |
| pMpGE_En03 | Kindly provided by Prof. Takayuki Kohchi, Kyoto University, Japan | Entry vector for CRISPR/Cas9 gRNA expression cassette |
| pMpGE010 | | Destination vector for CRISPR/Cas9 genome editing |
| pMpGWB104 | | Destination vector with no promoter and GUS fusion protein |

3.8 Oligonucleotides used

Oligo(dT) solution

Oligo(dT)₁₈ primer (Thermo Scientific) is a single stranded 18-mer oligonucleotide with 5' and 3' hydroxyl ends. The primer is supplied as 20X concentrated aqueous solution ready to be used for cDNA synthesis.

Other primers

Designed primers were purchased either as a lyophilizate form or in TE buffer from Merck Life Science Sp. z o. o., an affiliate of Merck KGaA, Darmstadt, Germany (Supplementary Table 8).

3.9 Softwares used

The following databases and softwares were utilized in the design and analysis of the experiments:

MarpolBase (<https://marchantia.info/>) (Bowman et al., 2017; Montgomery et al., 2020)
Genome database for Marchantia

CRISPRdirect (<https://crispr.dbcls.jp/>) (Naito et al., 2015)
Software for the design of guideRNAs for CRISPR/Cas9

MBEX (<https://mbex.marchantia.info/>) (Kawamura et al., 2022)
MarpolBase Expression database

eFP (Marchantia Atlas eFP Browser at bar.toronto.ca) (Tan et al., 2023)
Database used to visualize gene expression in Marchantia tissues

ImageJ (Fiji) (Schindelin et al., 2012)
Applications for this software are numerous and include phenotypic analysis (e.g. measures of the size, shape, etc.)

MAFFT (version 7) (<https://mafft.cbrc.jp/alignment/server/>) (Rozewicki et al., 2019)
Multiple sequence alignment program for amino acid or nucleotide sequences

MultiGauge (ver. 2.2) (Fuji Film Inc.)
Software for analyses of northern blot gels

Primer3web (version 4.1.0) (<https://primer3.ut.ee/>) (Untergasser et al., 2012)
Software used for primer designing

RNAfold (Folder ver. 1.11) (Hofacker, 2003)
Software used for the prediction of miRNA stem-loop secondary structures

SDS (version 2.4) (Thermo Fisher Scientific, USA)
Software used to design and analyze RT-qPCR experiments

GeneSnap (SynGene, UK)

Software for the visualization of agarose gels

VHX-7000 software (Keyence, Japan)

Software for capturing microscopic images

CLC Genomics Workbench (Qiagen, Germany)

Software for aligning RACE sequences with the genome/gene/plasmid

SnapGene (GSL Biotech LLC, USA)

Software for everyday molecular biology

Chapter 4. Methodology

4.1 *Escherichia coli* (DH-5α) transformation

The *E. coli* DH5α chemically competent cells (100 µl in 1.5 ml microcentrifuge tube) were thawed on ice for 15 minutes. The desired plasmid (~80 – 100 ng/µl) or ligation reaction mix (~5 µl) was mixed with the thawed competent cells by pipetting, and incubated on ice for 20-30 minutes. Then, the samples were subjected to heat shock at 42 °C for 2 minutes, and immediately chilled on ice for 2 - 5 minutes. Finally, 1 ml of the LB medium without antibiotics was added to each tube. After that, the mixture was incubated for 1 - 1.5 hours at 37 °C while being shaken at 350 rpm (Thermomixer Comfort, Eppendorf, Germany) in order to recover. 200-250 µl of the recovery mixture was spread on LB agar plates supplemented with X-Gal for blue-white screening and the appropriate antibiotics for each cloning vector. The plates were grown at 37 °C overnight.

Antibiotic resistance for different plasmids:

| Plasmid | Antibiotic resistance | Reference |
|----------------|--------------------------|--------------------------|
| pGEM-T easy | Ampicillin (100 µg/ml) | Addgene |
| pENTR™/D-TOPO® | Kanamycin (50 µg/ml) | Thermo Fisher Scientific |
| pMpGWB102 | Spectinomycin (50 µg/ml) | (Ishizaki et al., 2015) |
| pMpGWB103 | Spectinomycin (50 µg/ml) | |
| pMpGWB104 | Spectinomycin (50 µg/ml) | |

4.2 *Agrobacterium tumefaciens* transformation

The electro-competent cells were first thawed on ice for around 15 - 20 minutes. 500 - 1000 ng of the desired plasmid construct was added to the tube and mixed by gentle pipetting. The mixture was then transferred to the pre-chilled electroporation cuvette (Carl Roth GmbH+Co. KG, Germany). This cuvette was then placed into Gene Pulser X cell™ system's ShockPod cuvette chamber (Bio-Rad, USA). A pulse of 2.5 kV was applied and 1 ml of liquid LB media was added immediately and mixed by pipetting.

This bacterial culture was transferred to 1.5 ml microcentrifuge tube and placed in the incubator for 4 hours at 28 °C with shaking at 120 rpm. After 4 hours, the bacterial culture was centrifuged at 865 g for 5 minutes and plated onto LB agar plates with desired antibiotics. The transformed colonies could be observed after 2 - 3 days.

4.3 Polymerase Chain Reaction (PCR)

For standard PCR reactions, the fragments were amplified using 0.25 U DreamTaq DNA polymerase together with 10X DreamTaq DNA Green Buffer (includes 20 mM MgCl₂), 0.2 mM each dNTP, 10 µM forward primer and 10 µM reverse primer in a 10 µl reaction. Genomic DNA or cDNA or bacterial colony were used as a template. Negative controls contained water as a template. In order to amplify the promoter or coding regions, that were to be used for gateway cloning, the reactions were performed using CloneAmp HiFi PCR Premix. 12.5 µl of CloneAmp together with 2.5 µl each of 10 µM forward and reverse primer and 1 µl of template were used in a 25 µl reaction.

4.4 Agarose gel electrophoresis

For running agarose-gel electrophoresis, a small-gel horizontal system was used. The prepared samples were loaded on 1 - 1.5 % agarose gel and subjected to electrophoresis in 1X TAE buffer at a constant current of 50 mA (gel dimensions: 7X10 cm). Before loading on the gel, the RNA samples were denatured at 85 °C for 3 minutes. After the completion of electrophoresis run, the gel was visualized with Gene Snap software in G-Box Chemii (Syngene, UK).

4.5 Genomic DNA isolation

~100 mg of finely grounded plant material was used to isolate genomic DNA using GeneJET Plant Genomic DNA purification Mini kit according to the manufacturer's instructions. The concentration of isolated DNA was measured by DS-11 Denovix spectrophotometer (Denovix, USA) and integrity of DNA was assessed using 0.8 % agarose gel electrophoresis.

4.6 DNA isolation for genotyping

A small piece (approximately 1 - 2 mm²) of thallus tissue from desired plant was collected and put to BioTube™ (Carl Roth Laboratories, Germany). 100 µl of DNA extraction buffer was added to the microtubes together with one 5 mm or two 3 mm glass beads (Merck, USA). The samples were crushed for 3 minutes in the TissueLyser II (Qiagen, USA). After crushing, 400 µl of autoclaved water was added to each Biotube and mixed properly. 1 µl of this crude extract was used as a template for genomic PCR reaction.

PCR reaction:

| | |
|---------------------------------|-----------------|
| <u>2X KAPA Buffer</u> | <u>6.25 µl</u> |
| <u>KAPA polymerase (0.25U)</u> | <u>0.1 µl</u> |
| <u>Forward primer (10 mM)</u> | <u>1 µl</u> |
| <u>Reverse primer (10 mM)</u> | <u>1 µl</u> |
| <u>Template (crude extract)</u> | <u>1 µl</u> |
| <u>Water MilliQ</u> | <u>3.5 µl</u> |
| <u>Total</u> | <u>12.85 µl</u> |

PCR cycle:

94 °C, 2 min → (98 °C, 68 °C - 10 sec, X min) x 35 cycles → 68 °C, 1 min → 12 °C, ∞.
X= estimated length of product (kb)

The PCR products were loaded on 1 % agarose gel and DNA was separated by gel electrophoresis and visualized under UV light with Transilluminator (Thistle Scientific Ltd., UK) and G-BOX Chemii (Syngene, UK). For direct sequencing of PCR products, Exo-SAP reaction was performed.

Exo-SAP master mix:

| | |
|----------------------------------|-----------------|
| <u>10X Exo I Buffer</u> | <u>0.4 µl</u> |
| <u>10X CutSmart</u> | <u>0.4 µl</u> |
| <u>Exonuclease I (0.5U)</u> | <u>0.025 µl</u> |
| <u>rSAP (0.25U)</u> | <u>0.25 µl</u> |
| <u>Autoclaved H₂O</u> | <u>1.25 µl</u> |
| <u>Total</u> | <u>2.3 µl</u> |

2.3 µl of Exo-SAP mix was added to each PCR reaction tube containing 12.85 µl of genomic PCR reaction. The reaction mixture was incubated at 37°C for 30 minutes

and then at 95 °C for 10 minutes. Then, 10 µl water was added to the reaction mixture. For sequencing, 2 µl of PCR product, 1.5 µl forward primer and 1.5 µl of water were added to the sequencing PCR tube.

For sex identification, plants were genotyped using the same crude extract and primers which bind to specific regions on the sex chromosomes. For the X-chromosome, the rhf73 marker was amplified and for the Y chromosome, the rbm27 marker was amplified (Fujisawa et al., 2001).

4.7 RNA extraction

RNA isolation from antheridia or sperm cells: Antheridia were dissected manually in water while the sperm cells are released out from the antheridial discs into water. Both antheridia or sperm cells were collected in a 2 ml microcentrifuge tube and centrifuged at 18,670 g to settle antheridia at the bottom of the tube or to obtain a pellet of sperm cells, respectively. After removing the supernatant completely, the microcentrifuge tube containing antheridia / sperm cell pellet was dropped in the liquid nitrogen and stored at -80 °C. 200 µl of Trizol (for antheridia) and 500 µl of Trizol (for sperm cell pellet) was added to the microcentrifuge tubes. Next, one 3 mm / 5 mm glass beads were added and samples were lysed using TissueLyser II. Next, all the samples were centrifuged at 18,670 g for 10 minutes at room temperature. The supernatant was transferred to a new microcentrifuge tube and equal concentration of ethanol was added and mixed properly. From this point, RNA isolation was proceeded using the Direct-zol RNA Miniprep kit according to the manufacturer's instructions. RNA was eluted in 25-30 µl DEPC-treated water.

RNA isolation from whole tissues: ~200 mg mature thallus or antheridiophores or archegoniophores were frozen and ground to powder in liquid nitrogen. To isolate RNA from whole powdered tissue, 700 µl - 1 ml Trizol (mixed with 0.5 M EDTA and 10% w/v N-Laurylsarcosine Sodium salt solution for sRNA enrichment) was added to the tube. The samples were vortexed vigorously, incubated at room temperature for 5 minutes. After this, the samples were centrifuged at 18,670 g until the particulate debris from homogenized tissue is removed. From this point, RNA isolation was proceeded using

the Direct-zol RNA Miniprep kit according to the manufacturer's instructions. RNA was eluted in 50 µl DEPC-treated water.

RNA concentration was measured by DS-11 Denovix spectrophotometer (Denovix, USA). The efficiency of RNA per mg of tissue were as follows:

| Tissue Type | Efficiency of RNA in mg of tissue |
|---------------------------|-----------------------------------|
| Male vegetative Thallus | ~1.5 µg / 20 mg |
| Antheridiophores | ~1 µg / 20 mg |
| Antheridia | ~0.1 µg / 100-150 antheridia |
| Sperm cells | ~0.04 µg / thick pellet |
| Female vegetative Thallus | ~1.5 µg / 20 mg |
| Archegoniophores | ~0.4 µg / 10 mg |

4.8 DNase treatment

The concentration of isolated RNA in each sample was equalized, and contaminating genomic DNA was digested using the Turbo™ DNA-free kit. To 6 - 10 µg RNA in DEPC treated water, 5 µl of 10X Turbo™ DNase buffer, 1 µl of 10X Turbo™ DNase were added and incubated at 37 °C for 45 minutes. Then, 150 µl of DEPC-treated water and 200 µl of Acid phenol:chloroform:IAA (125:24:1) (Thermo Fisher Scientific, USA) was added, vortexed for 10 seconds and centrifuged at 18,620 g for 5 minutes at room temperature.

The upper aqueous phase (approximately 200 µl) was transferred to new microcentrifuge tube. To this, 2 µl of GlycoBlue co-precipitant (15 mg/ml) (Thermo Fisher Scientific, USA), 0.1 times volume of 3 M Sodium Acetate (Thermo Fisher Scientific, USA) and 2.5 times volume of 100 % Ethanol (Merck, USA) were added and precipitated at -20 °C overnight or -80 °C for at least 30 minutes. Next, the samples were centrifuged at 18,620 g for 20 minutes at 4 °C. The supernatant was removed carefully and the blue colored RNA pellet was washed two times with 70 % Ethanol. The pellet was air dried for 5 minutes and then dissolved in DEPC-treated water. RNA concentration was measured using DS-11 Denovix spectrophotometer and samples

were subjected to agarose gel electrophoresis to check RNA integrity. The isolation efficiency of RNA after DNase treatment were as follows:

| Tissue Type | Efficiency of RNA |
|---------------------------|-------------------|
| Male vegetative Thallus | 90 % |
| Antheridiophores | 90 % |
| Antheridia | 80 % |
| Sperm cells | 60 % |
| Female vegetative Thallus | 80 % |
| Archegoniophores | 50 % |

4.9 cDNA synthesis

For routine gene expression analysis: 2 - 5 µg of DNaseI-treated RNA together with 1 µl of Oligo(dT)₁₈ primer and 1 µl 10 mM dNTPs were incubated at 65 °C for 5 minutes to dissolve any secondary RNA structures. Next, it was incubated on ice for 2 minutes and a mixture containing 4 µl of 5X First strand buffer (Thermo Fisher Scientific, USA), 1 µl of 0.1 M DTT (Thermo Fisher Scientific, USA), 1 µl of RNasin® Plus Ribonuclease inhibitor and 1 µl of SuperScript™ III Reverse Transcriptase were added and reverse transcription was performed at 50 °C for 90 minutes, then 55 °C for 10 minutes and finally 75 °C for 15 minutes. Finally, cDNA was diluted 1:1 with nuclease-free water. To test the successful synthesis of cDNA, PCR was performed using MpACT7 primers (in Supplementary Table 8). cDNA quality was assessed as good when clear and uniformly strong bands were visible on 1% agarose gel.

For RNA-ligase mediated rapid amplification of cDNA ends (5'-RLM and 3'-RACE): To obtain the full length sequence of primary-miRNA transcripts, 5'-RLM and 3'-RACE experiments were performed. Gene specific primers (reverse primers for 5'-RLM RACE and forward primers for 3'-RACE) were designed within the precursor miRNA region (both 'standard' and 'nested' for the first and second round of PCR, respectively). cDNA template was synthesized from 2 - 5 µg of the total DNaseI -

treated RNA using GeneRacer™ kit according to the manufacturer's instructions. All the templates were checked with MpACT7 primers. 5'-RLM and 3'-RACE (round 1 and 2) reactions were set up according to manufacturer's instructions. The PCR products were then separated on 1 -1.5 % agarose gel. The products of desired length were cut from the gel and purified using GeneJET Gel Extraction and DNA Cleanup Micro kit according to the manufacturer's protocol. The purified products were cloned into pGEM T-easy vector using T4 DNA ligase. In general, insert to vector ratio of 5:1 was used in each reaction. The reaction was further incubated at 4 °C overnight and later used for transformation into *E. coli* DH5α. After transformation, the positive, white clones were confirmed by colony PCR. Plasmids were isolated from the confirmed clones and ~100 ng plasmid samples were sent for sequencing.

To determine the slice site of miRNA targets, 5'-RACE experiments were performed. The experiment set-up was similar to 5'-RLM with a single exception. During the library preparation, the removal of mRNA cap structure from DNaseI - treated RNA step was omitted, with the intention to analyze only truncated mRNAs in the 5'-RACE experiments.

4.10 RT-qPCR

RT-qPCR was used for the measurement of steady-state mRNA levels. The reaction contained 1 µl of cDNA template together with the 5 µl of GoTaq® qPCR Master Mix 2X, 2 µl each for both forward and reverse primers (0.5 mM). The total reaction volume was 10 µl and qPCR was conducted using the 7900 HT Fast-Real time PCR system (Thermo Fisher Scientific, US). Three biological and two technical replicates were used for each sample. MpACT7 and MpAPT were chosen as reference genes on the basis of analysis of *M. polymorpha* reference genes suitable for RT-qPCR (Saint-Marcoux et al., 2015). The following cycling conditions were used:

| Temperature | Duration | |
|-------------|----------|-----------|
| 95 °C | 10 min | |
| 95 °C | 15 sec | 40 cycles |
| 60 °C | 1 min | |
| 95 °C | 15 sec | |
| 60 °C | 15 sec | |
| 95 °C | 15 sec | |

At the end of the cycles, a dissociation curve analysis was performed using SDS 2.4 software to confirm that only a single amplicon had been formed. From three Ct values for each biological replicate, average was calculated and normalized to that of housekeeping gene to determine delta Ct values. The relative expression levels were calculated using the formula $2^{-\Delta\Delta C_t}$. The statistical significance of the results was estimated using a student's t-test at three significance levels: *p < 0.05, **p < 0.01, and ***p < 0.001.

4.11 CRISPR/Cas9 based genome editing

For generation of CRISPR/Cas9 genome edited plants by single guide RNA approach, the protocol by Sugano et al., (2014) was followed with some modifications. CRISPRdirect tool was explored to design the nucleotide sequences to be used as sgRNA and the PAM (= NGG) motif was screened within the genomic loci. The sequence with the highest specificity for the target and minimum off-targets hits were chosen. The PAM (= NGG) motif sequence was removed for ordering the oligo nucleotide sequences. sgRNA oligonucleotides used in this research are listed in (Supplementary Table 8).

sgRNAs DNA oligonucleotides were first annealed in TE at 95 °C for 5 minutes. Next, the gRNAs cassettes were ligated into pMpGE_En03 entry vector cleaved using BsaI restriction enzyme. The gRNA cassettes was next transferred to the destination vector pMpGE010 using LR cloning (for the cloning of inserts (flanked by *attL* sites) from gateway entry vectors into the destination binary vectors (containing *attR* sites).

Gateway™ LR Clonase™ II enzyme mix was used with ~30 ng of the entry vector and ~50 ng of the destination vector. The resultant destination vectors were further used for *Agrobacterium*-mediated transformation of *M. polymorpha* sporelings.

4.12 Plasmid DNA isolation

Plasmid DNA was isolated from 3 ml of liquid LB bacterial cultures (with appropriate antibiotics) using a GeneJET Plasmid Miniprep kit according to the manufacturer's protocol. DNA concentration was measured using DS-11 Denovix spectrophotometer.

4.13 DNA sequencing (Sanger method)

The routine DNA sequencing was performed using the Sanger method in the Laboratory of Molecular Biology Techniques at the Faculty of Biology (Adam Mickiewicz University) in Poznan, Poland.

4.14 Crossing of *Marchantia* plants

The gemmae's from Tak-1 and Tak-2 accessions were set up in sterilized soil under the same conditions mentioned in (Section 3.1). The plants were monitored daily for the emergence of sexual structures, antheridiophores and archegoniophores, respectively. Sterile water was added to both male and female reproductive structures to create a moist environment conducive for sperm mobility. Then, sperm transfer was manually assisted using a pipette. The sterile water containing sperm was collected from the top of antheridiophores and deposited onto the archegoniophores digitate rays and left undisturbed for at least one hour. This process was repeated two to three times to facilitate the production of sporophyte. Upon sporophyte development, spores were collected from the archegoniophores using forceps in a 1.5 ml microcentrifuge tubes and stored at -80 °C.

4.15 *Agrobacterium*-mediated transformation of *Marchantia* sporelings

The method of transformation was followed after Ishizaki et al. (2008). Intact WT spores obtained from crosses between Tak-1 and Tak-2 plants were collected in a 1.5

ml microcentrifuge tube. 500 µl of water was added to this tube of dried spores and the spore clumps were gently pressed against the side of the tube with the pipette tip to make a uniform yellowish colored suspension.

Next, 500 µl of NaDCC (sodium dichloroisocyanurate) (Merck, USA) was added to the tube for surface sterilization of spores. It was vortexed and incubated at room temperature for 3 minutes, then centrifuged at 18,620 g for 1 minute. The supernatant was removed and spores were rinsed 3 times with 1 ml of sterile water. Then, 500 µl of sterilized spore suspension was added to 50 ml of liquid OM51C medium containing flask followed by incubating for 7 days under continuous white light at 22 °C with 120 rpm shaking.

On the fifth day, a single colony of *A. tumefaciens* strain GV3101 carrying the vector of interest was inoculated into 5 ml of liquid LB media containing 20 µg/ml of gentamicin, 100 µg/ml of rifampicin, and 50 µg/ml of spectinomycin, and incubated with 120 rpm shaking at 28 °C for 1 - 2 days.

On the seventh day, the bacterial culture was centrifuged at 2000 g for 15 minutes and resuspended in 10 ml of fresh OM51C liquid medium containing 100 µM acetosyringone which induces the virulence genes of *Agrobacterium*. The mixture was incubated for 5-6 hours with 120 rpm shaking at 28 °C.

In the subsequent experiment, 7-day old cultures of *Marchantia* sporelings were supplemented with 100 µM of acetosyringone and 1 ml of induced *Agrobacterium* culture. The sporelings and bacteria were then co-cultivated for 2 days on a 120 rpm shaker under continuous white light conditions. Next, the transformed sporelings were collected on a sieve (40 µm BD sterile cell strainer) (Becton Dickinson, USA) and rinsed thoroughly with ~250 ml of autoclaved water.

The washed sporelings were spread on OM51C agar plates supplemented with 100 µg/ml cefotaxime to inhibit *A. tumefaciens* growth and 10 µg/ml hygromycin B to select transformed sporelings. The transformants were visible after 2 - 3 weeks. G2 generation of the transformants were genotyped and taken as the starting point for further phenotypic analysis.

4.16 Phenotypic analysis

All phenotypic analyses of vegetative and reproductive structure cross-sections of *Marchantia* were performed using the VHX-7000 digital microscope (Keyence, Japan) and all images were captured with the VHX-7000 software.

4.17 Protocol for the transfer of plants from soil to *in vitro* culture

The apical part of the thalli were cut with a scalpel under the binocular and placed in sterile water in a 50 ml falcon. The thalli fragments were washed 3 times in the sterile water by vigorous shaking. A drop of Tween20 was added and washing with sterile water was repeated three times again. Next, the plants were washed in 70% ethanol for 15 seconds and in 5 % ACE (chlorine bleach) for about 5 minutes. Finally, the thalli fragments were washed with sterile water and dried on sterile blotting paper and placed on half strength Gamborg medium. The thalli fragments become transparent except the apical cell. In about 2 weeks, the apical cell can be observed to grow into new thallus.

4.18 Construction of promoter reporter plants

To investigate *in vivo* expression levels of *MIR* genes, their promoters sequences corresponding to ~2.5 - 3 kb upstream of the pre-miRNA region were amplified using genomic DNA as a template for PCR. Gene-specific primers were designed with the addition of CACC sequence to the forward primer to clone in the entry vector pENTR™/D-TOPO®. The amplified promoter fragments were then ligated into the binary vector pMpGWB-104 using Gateway™ LR Clonase™ II enzyme mix.

The recombinant plasmids were transformed into *E. coli* DH5α competent cells, and positive clones were selected on LB agar plates containing appropriate antibiotics. Colony PCR was used to confirm the presence of the promoter fragment in the vector. Sanger sequencing was performed to verify the sequence and orientation of the insert. The verified plasmids were introduced into the *Marchantia* genome via agrobacterium-mediated transformation. The histochemical GUS staining for the respective gene promoters were observed in various tissues of *Marchantia*.

4.19 GUS staining of *Marchantia* vegetative and reproductive tissues

To obtain GUS stained *M. polymorpha* plants, the protocol described in Ishizaki et al., (2012) was used with several modifications. Firstly, the vegetative tissues (e.g., thalli, gemmae) and reproductive tissues (e.g., antheridiophores, antheridia, archegoniophores, and archegonia) of desired plants were harvested and put to a 24-well or 6-well plate (Merck, USA). 1 ml of GUS staining solution (Section 4.5) was added to each well to submerge the tissues in GUS solution. The plants were subjected to vacuum infiltration for 3 times (15 minutes each) inside a Vacuotherm™ oven (Thermo Fisher Scientific, USA). The plant tissue samples were then incubated at 37 °C in the dark in the incubator (New Brunswick; Eppendorf, Germany) overnight. Next day, the staining solution was removed and the plants were washed with ethanol with an increasing gradient (70 %, 80 %, 95 %). Tissues were incubated in each ethanol concentration for 10 –15 minutes until clear. Cleared tissues were mounted on glass slides for microscopic observation. GUS staining was visualized using VHX-7000 Keyence digital microscope. The tissues from WT plants were used as negative controls to confirm the specificity of the staining. Tissues with strong GUS activity showed distinct blue coloration, while negative controls showed no staining.

4.20 Northern hybridization

The glass plates (dimensions 15x15 cm), comb and sides were cleaned properly with ethanol (70 %). The glass plates were fixed on all three sides first with a tape and second with dPAG gel to prevent any leakage. dPAG was prepared as mentioned in (Chapter 3, section 4.6).

4.20.1 RNA electrophoresis in dPAG

RNA-PAGE was performed using an Owl™ P10 DS dual gel system (Thermo Fisher Scientific, USA). A short term pre-electrophoresis was performed for 10 min at 300 Volt in 1X MOPS buffer before loading the samples. The sample wells were cleaned using a syringe before and after pre-electrophoresis. Then, ~10 - 15 µg of RNA samples were mixed with an equal amount of 2X RNA loading dye and denatured at 90 °C for

5 min along-with the [³²P] ATP-labeled Decade RNA marker (Thermo Fisher Scientific, USA) for sizing bands. Finally, electrophoresis of RNA samples on 15 % dPAG was performed at 300 V for 2.5 - 3 hours.

4.20.2 Transfer of RNA from gel to membrane

For transferring the RNA from the gel to the membrane, 6 sheets of 3 MM Whatman paper and 1 sheet of Amersham Hybond-NX Nylon membrane (Capitol Scientific, USA) were cut according to the size of the gel (~15 x 15 cm). First, three sheets of Whatman paper were soaked in DEPC-treated water and placed on the anode platform of the Trans-Blot® SD Semi-Dry Electrophoretic Transfer Cell (Bio-Rad Laboratories, USA). Then, the nylon membrane was soaked and placed on top of these 3 sheets. Now the gel was placed on top of membrane and covered with 3 more sheets of Whatman paper dipped in DEPC water. A glass rod was carefully rolled over the surface of thick paper to remove all the air bubbles. Then, the cathode platform was assembled and the transfer was carried out for 1 hour at 20 V.

4.20.3 EDC-cross linking

After transfer, the nylon membrane with RNA was placed with RNA side face up on one sheet of 3 MM Whatman Paper and saturated with ~24 ml of freshly prepared cross-linking EDC reagent (as described in Chapter 3 section 4.6). Then, it was wrapped in Saran foil and incubated at 55 °C for 2 hours in the oven.

4.20.4 Labeling of oligonucleotide probes

The DNA oligonucleotide designed for mature miRNA detection was radiolabeled at its 5' end with T4 Polynucleotide Kinase (10 U/μl) and [γ-P³²] ATP, 6000 Ci/mmol, 10 mCi/ml (Hartmann Analytic, Germany).

T4-PNK reaction

| Component | Final concentration | Amount |
|-------------------------------|---------------------|------------------|
| 10X T4-PNK Buffer | 1X | 5 μ l |
| T4-PNK | 20U | 2 μ l |
| Oligonucleotide probe | 10 μ M | 2 μ l |
| γ -ATP (6000 Ci/mmol) | 50 pmol | 5 μ l |
| DEPC-treated H ₂ O | - | Up to 50 μ l |

The reaction mixture was incubated at 37 °C for 45 minutes. To remove the unincorporated [γ^{32} P]-ATP from the samples, CentriPure MINI Spin Columns Desalt Z-25 (EMP Biotech GmbH, Germany) were used, according to the manufacturer's protocol.

4.20.5 Hybridization and washing

After cross-linking, the membrane was rolled with RNA face side up and put in the glass hybridization bottle. 10 ml of pre-warmed hybridization buffer (as described in Chapter 3 section 4.6) was added to the bottle and incubated at 37 °C for 30 - 45 minutes. Then, the radiolabeled probes were denatured at 85 °C for 2 minutes and along-with the radiolabeled U6 probe. The hybridization buffer was replaced with a fresh portion and the desired radiolabeled probe (50 μ l) and U6 probe (2 -10 μ l) were also added to the hybridization bottle and left for overnight incubation at 37 °C. The next day, the membrane was washed with a washing buffer (as described in section 4.6 of Chapter 3).

4.20.6 Hybridization signal detection

The exposition lasted from 2 hours to 14 days depending on the accumulation level of the miRNA under study. The hybridization signal was scanned using Phosphor-imager-Fluorescent Image Analyzer FLA-5100 (Fuji Film Inc.) and the results were quantified using Multi Gauge software Ver. 2.2 (Fuji Film Inc.).

4.21 RNA-seq library preparation for Illumina sequencing

Total RNA was extracted from archegoniophores and antheridia of WT and mutant plants (as described in section 7 of Chapter 4). 3 - 6 µg of RNA was subjected to DNaseI treatment (as described in section 8 of Chapter 4). 3 - 5 µg of DNase-treated RNA samples were further processed to deplete ribosomal RNA by using Ribominus™ Plant kit for RNA seq following instructions from the provider. The efficiency of RNA after depletion of rRNA was 80 % for both archegoniophores and antheridia. ~100 ng of rRNA depleted RNA samples were prepared in 5 µl of nuclease-free water and following steps were followed.

4.21.1 RNA fragmentation and priming

The fragmentation and priming reaction was assembled on ice and the following components were mixed thoroughly by pipetting up and down several times.

| Fragmentation and Priming mix | Volume |
|--|--------|
| rRNA depleted RNA | 5 µl |
| NEBNext First Strand Synthesis Reaction Buffer | 4 µl |
| Random Primers | 1 µl |
| Total Volume | 10 µl |

The PCR reaction tubes were then incubated at 94 °C for 8 minutes with the heated lid set to 105 °C.

4.21.2 First-strand cDNA synthesis

| First-strand cDNA synthesis reaction | Volume |
|---|--------|
| Fragmented and primed RNA | 10 µl |
| NEBNext Strand Specificity Reagent | 8 µl |
| NEBNext First Strand Synthesis Enzyme Mix | 2 µl |
| Total Volume | 20 µl |

The above reaction was assembled on ice and mixed thoroughly by pipetting and then incubated in a preheated thermal cycler with heated lid set at $\geq 80\text{ }^{\circ}\text{C}$ as follows:

Step 1: $25\text{ }^{\circ}\text{C}$ – 10 minutes

Step 2: $42\text{ }^{\circ}\text{C}$ – 15 minutes

Step 3: $70\text{ }^{\circ}\text{C}$ – 15 minutes

Step 4: $4\text{ }^{\circ}\text{C}$ – ∞

4.21.3 Second-strand cDNA synthesis

| Second-strand cDNA synthesis reaction | Volume |
|---|------------------|
| First-Strand synthesis product | 20 μl |
| NEBNext Second Strand Synthesis Reaction Buffer with dUTP mix | 8 μl |
| NEBNext Second Strand Synthesis Enzyme Mix | 4 μl |
| Nuclease free water | 48 μl |
| Total Volume | 80 μl |

The above reaction was assembled on ice and mixed thoroughly by pipetting and then incubated in a thermal cycler for 1 hour at $16\text{ }^{\circ}\text{C}$ with heated lid off. In the meantime, transfer the purification beads from $-20\text{ }^{\circ}\text{C}$ to the workbench.

4.21.4 Purification of double-stranded cDNA using beads

The reaction was purified by AMPure beads following the instructions in the manual and finally 50 μl of purified reaction was transferred to new PCR tube.

4.21.5 End prep of cDNA library

| End prep reaction | Volume |
|---|------------------|
| Second-Strand synthesis product | 50 μl |
| NEBNext Ultra II End prep Reaction Buffer | 7 μl |
| NEBNext Ultra II End prep Enzyme Mix | 3 μl |
| Total Volume | 60 μl |

The above reaction was assembled on ice and mixed well by pipetting. The samples were incubated in a thermal cycler with the heated lid set at $\geq 75^{\circ}\text{C}$ as follows:

Step 1: 20°C – 30 minutes

Step 2: 65°C – 30 minutes

Step 3: 4°C – ∞

4.21.6 Adaptor ligation

The NEBNext Adaptor was diluted 5-fold in ice-cold Adaptor Dilution Buffer prior to setting up the ligation reaction.

| Ligation reaction | Volume |
|--------------------------------------|--------------------|
| End prepped DNA | 60 μl |
| Diluted Adaptor | 2.5 μl |
| NEBNext Ligation Enhancer | 1 μl |
| NEBNext Ultra II Ligation Master Mix | 30 μl |
| Total Volume | 93.5 μl |

The above reaction was assembled on ice and mixed well by pipetting. The samples were incubated for 15 minutes at 20°C in a thermal cycler with the heated lid off. Next, 3 μl of USER Enzyme was added to the ligation reaction, resulting in a total volume of 96.5 μl . After mixing properly, the reaction was incubated at 37°C for 15 minutes with the heated lid $\geq 45^{\circ}\text{C}$. Then, ligation reaction was purified using AMPure beads as per manufacturer's instructions. Finally, 15 μl of supernatant was transferred to new tube.

4.21.7 PCR enrichment of adaptor-ligated DNA

| Components | Volume |
|--------------------------------|------------------|
| Adaptor Ligated DNA | 15 μl |
| Index primer | 5 μl |
| Universal PCR Primer | 5 μl |
| NEBNext Ultra II Q5 Master Mix | 25 μl |
| Total Volume | 50 μl |

The above reaction was mixed well by pipetting and quickly spin in a microcentrifuge. The samples were then placed in a thermal cycler with the heated lid set to 105 °C and PCR amplification was performed with the following cycling conditions:

| Cycle step | Temperature | Time | Cycles |
|----------------------|-------------|------------|--------|
| Initial denaturation | 98 °C | 30 seconds | 1 |
| Denaturation | 98 °C | 10 seconds | 8 |
| Annealing/extension | 65 °C | 75 seconds | |
| Final extension | 65 °C | 5 minutes | 1 |
| Hold | 4 °C | ∞ | |

At this point, the libraries were diluted 5-fold and 1 µl of this diluted library was run on Agilent Bioanalyzer using Agilent High sensitivity DNA kit (Agilent, United States). If the library yield is too low to quantify, the samples were subjected to more rounds of PCR enrichment and checked again on the Agilent Bioanalyzer. After this, the PCR reaction was purified using AMPure beads and finally 20 µl of the supernatant was transferred to a clean PCR tube and stored at -20 °C.

4.21.8 Assess library quality

The final library quality was again checked on Agilent Bioanalyzer DNA chip and quantified on Qubit Fluorometer using Qubit dsDNA HS assay kit.

Chapter 5: Results

5.1 Liverwort-specific miRNAs differentially expressed in *Marchantia* vegetative and generative organs

Alaba et al., (2015) identified 42 novel miRNA families in *Pellia endiviifolia*. Tsuzuki et al., (2016), after comparison of *P. endiviifolia* and *M. polymorpha* micro-transcriptomes, found only one common miRNA, PenmiR8163/MpmiR11737ab, shared between *Pellia endiviifolia* (a simple thalloid liverwort) and *M. polymorpha* (a complex thalloid liverwort). Pietrykowska et al., (2022) identified two more miRNAs shared between *P. endiviifolia* and *M. polymorpha* (PenmiR8170/MpmiR11865* and PenmiR8185/MpmiR11889; hereafter referred to as liverwort-specific miRNAs). Thus, despite the diversity in morphology and reproductive strategies between these two liverwort species, they share a few miRNAs, while generally maintaining distinct miRNA species.

sRNA NGS and northern blot analysis revealed the differential expression profiles for liverwort-specific PenmiR8163/MpmiR11737a/b and PenmiR8170/MpmiR11865* in *Marchantia*'s vegetative and generative organs (Pietrykowska et al., 2022; Aggarwal et al., 2024). Additionally, it was shown by Pietrykowska, (2020) that PenmiR8185/MpmiR11889 and some other liverwort-specific miRNAs, namely, MpmiR11887 and MpmiR11796 (not present in *Pellia* but present in other liverwort species) exhibit intriguing expression patterns across *Marchantia*'s organs.

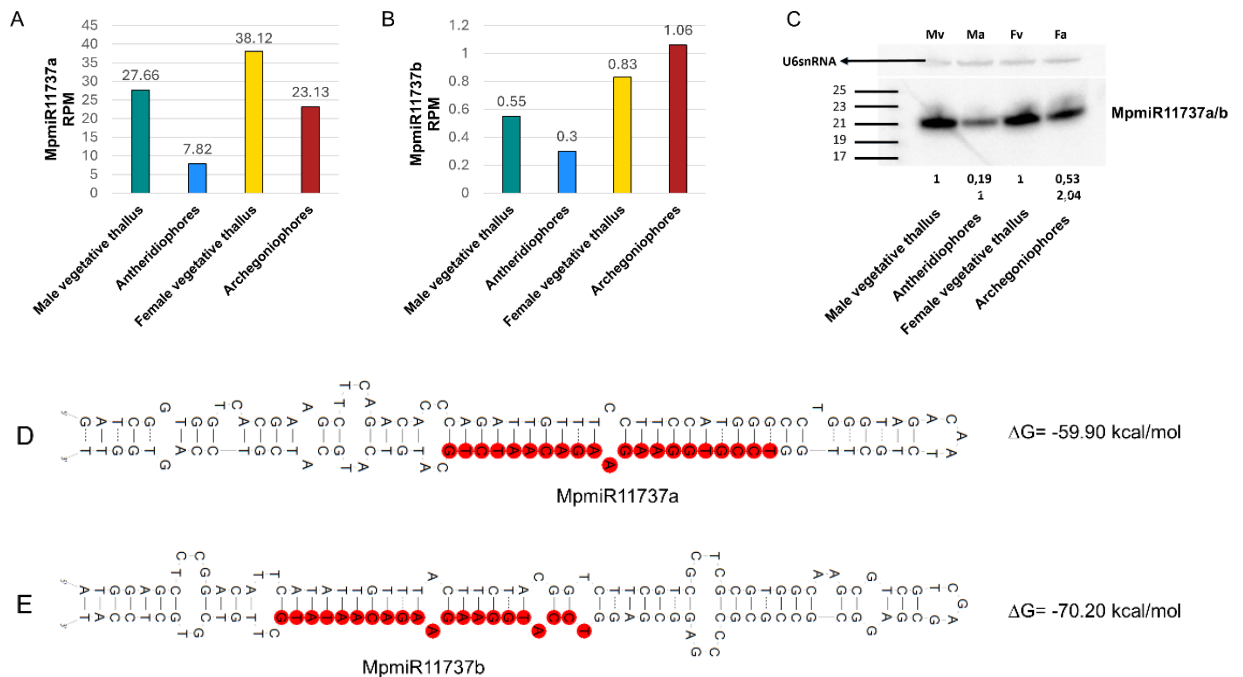
In this thesis, I provide for the above miRNAs: *MIR* gene structures, pri-miRNA transcript levels, *MIR* gene promoter activities, and the identification of putative targets for these four liverwort-specific miRNAs. Furthermore, I also present the functional characterization of two of them, MpmiR11887 and MpmiR11796, shedding light on their roles in *Marchantia*'s reproductive development.

5.1.1 MpmiR11737a/b

The MpmiR11737 family in *Marchantia* comprises two members, MpmiR11737a and MpmiR11737b. The miRNAs differ by 2nt substitutions and according to MarpolBase, they are encoded within two different protein-coding genes i.e. Mp5g12920 and Mp8g07030 genes, respectively (Bowman et al., 2017; Montgomery et al., 2020).

5.1.1.1 Pri-MpmiR11737a/b and mature MpmiR11737a/b show different expression and accumulation patterns in *Marchantia*

sRNA NGS data and northern blot hybridization results from Pietrykowska, (2020) revealed that MpmiR11737a accumulates at higher levels in male and female vegetative thalli compared to *Marchantia* reproductive organs (Fig. 13A, 13C). On the other hand, the level of MpmiR11737b is extremely low, as seen in sRNA NGS data (Fig. 13B). The pre-MpmiR11737a and pre-MpmiR11737b form classical stem loop structures (Fig. 13D, 13E). RT-qPCR analysis showed that the expression pattern of pri-MpmiR11737a aligns with that of mature MpmiR11737a (Fig. 13F). In the case of pri-miR11737b, RT-qPCR analysis showed higher transcript levels in reproductive organs (Fig. 13G). These findings indicate that MpmiR11737a and MpmiR11737b exhibit distinct accumulation patterns in *Marchantia*.



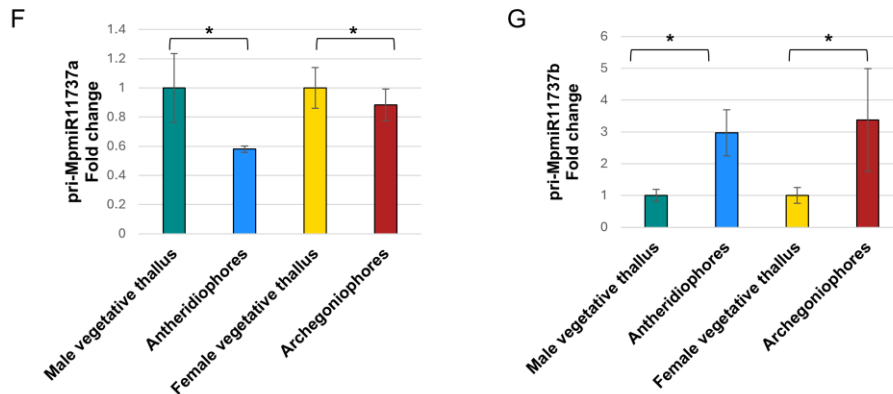


Figure 13. MpmiR11737a/b accumulation and pri-MpmiR11737a/b expression patterns. (A) & (B) sRNA NGS sequencing results with normalized read counts presented above each bar; RPM- Reads Per Million. (C) Northern hybridization; the numbers presented below northern blot show the relative strengths of signal of miRNA bands; control signals were arbitrarily established as 1; signal intensity differences were separately obtained for male vegetative thalli control/antheridiophores, female vegetative thalli control/archegoniophores (numbers in first row) and antheridiophores control/archegoniophores (numbers in second row); Left side of the blot - RNA marker depicting 17-25 nucleotide long RNAs; U6 snRNA hybridization-RNA loading control. (D) & (E) Structures of pre-MpmiR11737a and pre-MpmiR11737b as predicted by RNA Fold (Folder ver. 1.11); miRNA sequences are highlighted in red and minimum free energy (ΔG) of predicted structure is written on the right side. (F) & (G) RT-qPCR expression levels of pri-MpmiR11737a and pri-MpmiR11737b transcript; *p-value < 0.05. All images have been derived from (Aggarwal et al., 2024) with some modifications.

5.1.1.2 Identification of full-length transcript of MpMIR11737a gene

The analysis genomic and transcriptomic database of Marchantia (MarpolBase) revealed that MpmiR11737a is encoded within the second long intron of the Mp5g12920 gene, which encodes a chloroplast PsbP-like protein1. To determine the full-length of pri-MpmiR11737a, 5'-RLM and 3'-RACE experiments were performed using cDNA prepared (as described in section 4.9). In 5'-RLM RACE approach, all

mRNAs deprived of the cap structure are discarded, and thus in this technique, identified 5'-end represent transcription start site (TSS).

The nucleotide sequence of pre-miR11737a was retrieved from MarpolBase. Primers were designed within the precursor region and used to amplify the most distal 5' and 3' ends. The products obtained from 5'-RLM and 3'-RACE reactions were eluted from the gel, cloned and sequenced. The longest products were approximately 784 bp in the 5'-RLM RACE reaction and ~827 bp in the 3'-RACE reaction (Fig. 14A, 14B). To confirm that the longest 5' and 3' ends belonged to the same precursor molecule, RT-PCR was performed using primers designed to the furthest 5' and 3' ends of the RACE products. The full-length transcript of Mp*MIR11737a* gene, measuring 1131 bp, was obtained via RT-PCR and confirmed by Sanger DNA sequencing (Fig. 14C).

The second member of MpmiR11737 family i.e. MpmiR11737b was found to be located within the 5'-UTR of Mp8g07030 gene, which encodes a protein of unknown function. Despite several attempts, amplification of the precursor region by 5'-RLM and 3'-RACE was unsuccessful (Fig. 15B). This is likely due to very low expression levels of the pri-MpmiR11737b, as inferred from the low levels of mature MpmiRNA11737b observed in sRNA NGS data across all tissues (Fig. 13B).

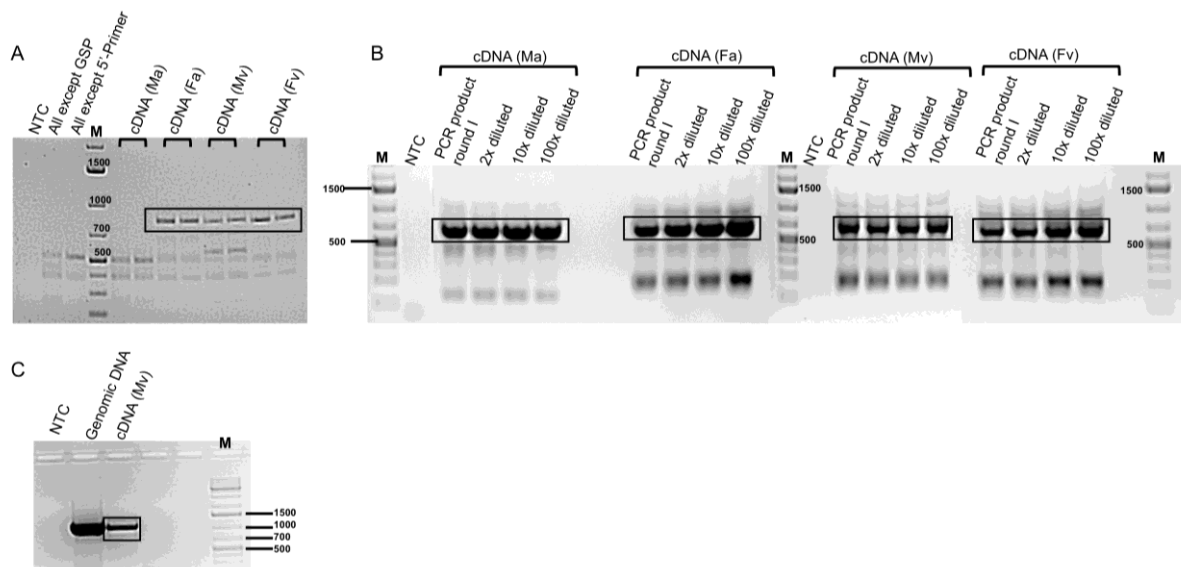


Figure 14. Agarose gel electrophoresis of products obtained with 5'-RLM (A), 3'-RACE (B) and RT-PCR (C) for the full-length of pri-MpmiR11737a. The specific amplified products have been boxed. For 5'-RLM, three negative controls were used to identify

specific RACE products: All components except template (NTC), all components except Gene specific primer (GSP) and all components except GeneRacer™ 5' or 3' primer; For 3'-RACE, a negative control reaction without template (NTC) was used and in addition to this, Genomic DNA was used as a positive control for RT-PCR; cDNA templates from four different tissues were used: Ma= antheridiophores, Fa= archegoniophores, Mv= male vegetative thallus and Fv= female vegetative thallus. M – GeneRuler™ 1 kb Plus DNA Ladder.

5.1.1.3 **MpMIR11737a gene structure identification**

The sequences obtained from 5'-RLM and 3'-RACE were aligned with the genomic and transcriptomic databases using CLC Genomics Workbench, allowing for the identification of gene structures. The MpMIR11737a gene sequence was found to overlap with the sequence of protein coding gene from which it was derived (Mp5g12920). The TSS of MpMIR11737a gene was identified within the long intron, while the termination site was located in the third exon of its host gene (Mp5g12920). Therefore, it can be concluded that MpMIR11737a functions as an intron-less, independent transcriptional unit while sharing sequence overlap with the protein coding Mp5g12920 gene (Fig. 15A).

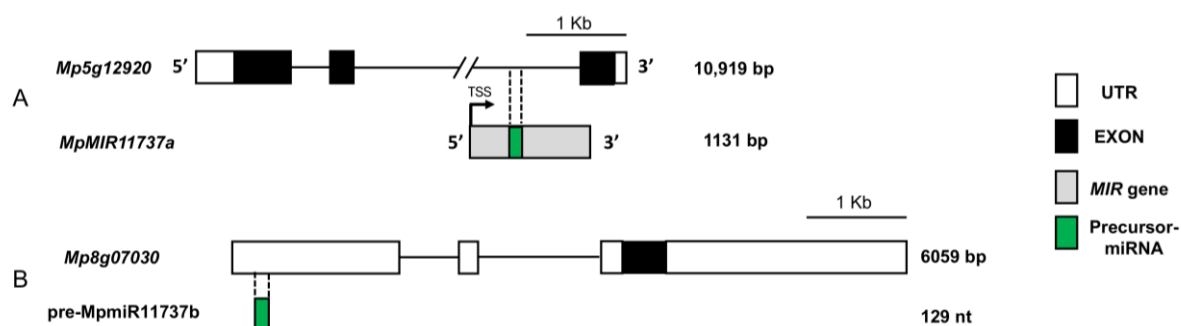


Figure 15. Diagram showing MpMIR11737a gene structure. Mp5g12920, the protein coding gene (upper row) overlapping with MIR11737a gene (lower row) and Mp8g07030 overlapping with pre-MpmiR11737b sequence has been shown. The length of protein-coding gene and MIR gene is written on the right side. The scale bar corresponds to 1 kb. The diagrams have been derived from (Aggarwal et al., 2024) with some modifications.

5.1.1.4 Identification of putative target of MpmiR11737a from degradome data

The degradome sequencing data identified a putative mRNA target for MpmiR11737a i.e. *Mp1g15010* gene, which encodes an uncharacterized protein. The putative miRNA slice site is located within the 3'-UTR region (Fig. 16A). The t-plot (target-plot) analysis showed a significant peak at position 877 within the 3'-UTR, indicating the cleavage site (Fig. 16B). The base-pairing between miRNA and the putative target mRNA sequence exhibits near-perfect complementarity (Fig. 16B). Furthermore, RT-qPCR analysis revealed a reverse correlation between the expression level of MpmiR11737a and its putative target *Mp1g15010* transcript (Fig. 16C).

Table 1. Putative mRNA target for MpmiR11737a revealed by degradome sequencing:

| miRNA ID | Target gene ID | mRNA length (nt) | mRNA:miRNA pairing region (nt) | Cleavage (nt of mRNA)+ | Cleavage (nt of mRNA)* |
|-------------|------------------|------------------|--------------------------------|------------------------|------------------------|
| MpmiR11737a | <i>Mp1g15010</i> | 1695 | 863:885 | - | c:877(5.39 2 118)* |

Cleavage occurs between the 9th/10th nucleotide (+) or 10th/11th nucleotide (*) of the miRNA sequence.

c:877 (5.39|2|118)*

c refers to the cleavage site and the values in bracket correspond to degradome sequencing metrics.

877 means nucleotide position of cleavage in mRNA.

5.39 means normalized degradome read abundance at this cleavage site.

2 is the rank of the signal.

118 means the number of raw degradome reads mapped to this cleavage site (Addo-Quaye et al., 2009; German et al., 2009).

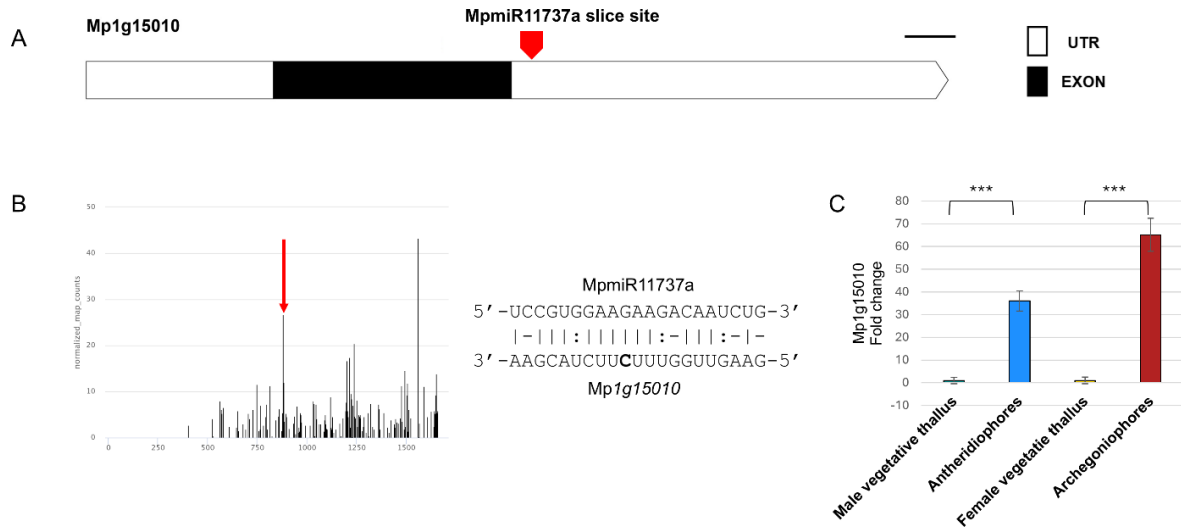


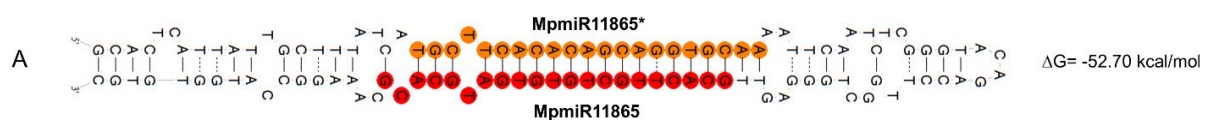
Figure 16. MpmiR11737a putatively targets mRNA of Mp1g15010 gene. (A) Structure of Mp1g15010 gene; the putative slice site for miRNA is shown by downward pointing arrow; (B) Based on degradome data, t-plot for corresponding putative target mRNA; the predicted cleavage site is pointed with red arrow; Alongside the t-plot, complementary base pairing between miRNA and target mRNA is presented; cleavage site is indicated by nucleotide marked in bold. (C) RT-qPCR analysis of expression level of Mp1g15010 transcript; ***p-value < 0.001. The images, (B) and (C) have been derived from (Aggarwal et al., 2024) with some modifications.

5.1.2 MpmiR11865/11865*

As previously mentioned, liverwort-specific miRNA Pen-miR8170/MpmiR11865*, was also identified in the sRNA NGS data of *M. polymorpha* (Pietrykowska et al., 2022; Aggarwal et al., 2024). Pietrykowska thesis, (2020) demonstrated that MpmiR11865* represents a passenger miRNA (miRNA*) within a known precursor of MpmiR11865 (Fig. 17A). Surprisingly, no miRNA species corresponding to MpmiR11865 were detected in sRNA NGS data of *P. endiviifolia*.

5.1.2.1 MpmiR11865* and MpmiR11865 exhibit opposite accumulation patterns in Marchantia

Pietrykowska, (2020) and Aggarwal et al., (2024) showed sRNA NGS data revealing that MpmiR11865* is more abundant in female vegetative thalli and archegoniophores (Fig. 17B). Northern blot analysis further confirmed that MpmiR11865* accumulates ~3 fold higher in archegoniophores as compared to female vegetative thalli. However, unexpectedly, northern blot analysis revealed that MpmiR11865* primarily accumulates in the antheridiophores (Fig. 17C). This is not in agreement with NGS data, however, northern hybridization was repeated three times with three different biological replicas. In contrast, our studies found that mature MpmiR11865 is strongly upregulated in archegoniophores, as shown by both NGS and northern hybridization (Fig. 17D, 17E). Based on northern blot results, we can conclude that MpmiR11865* exhibits an opposite expression profile to that of MpmiR11865, suggesting that these are two independent miRNA species originating from the same precursor.



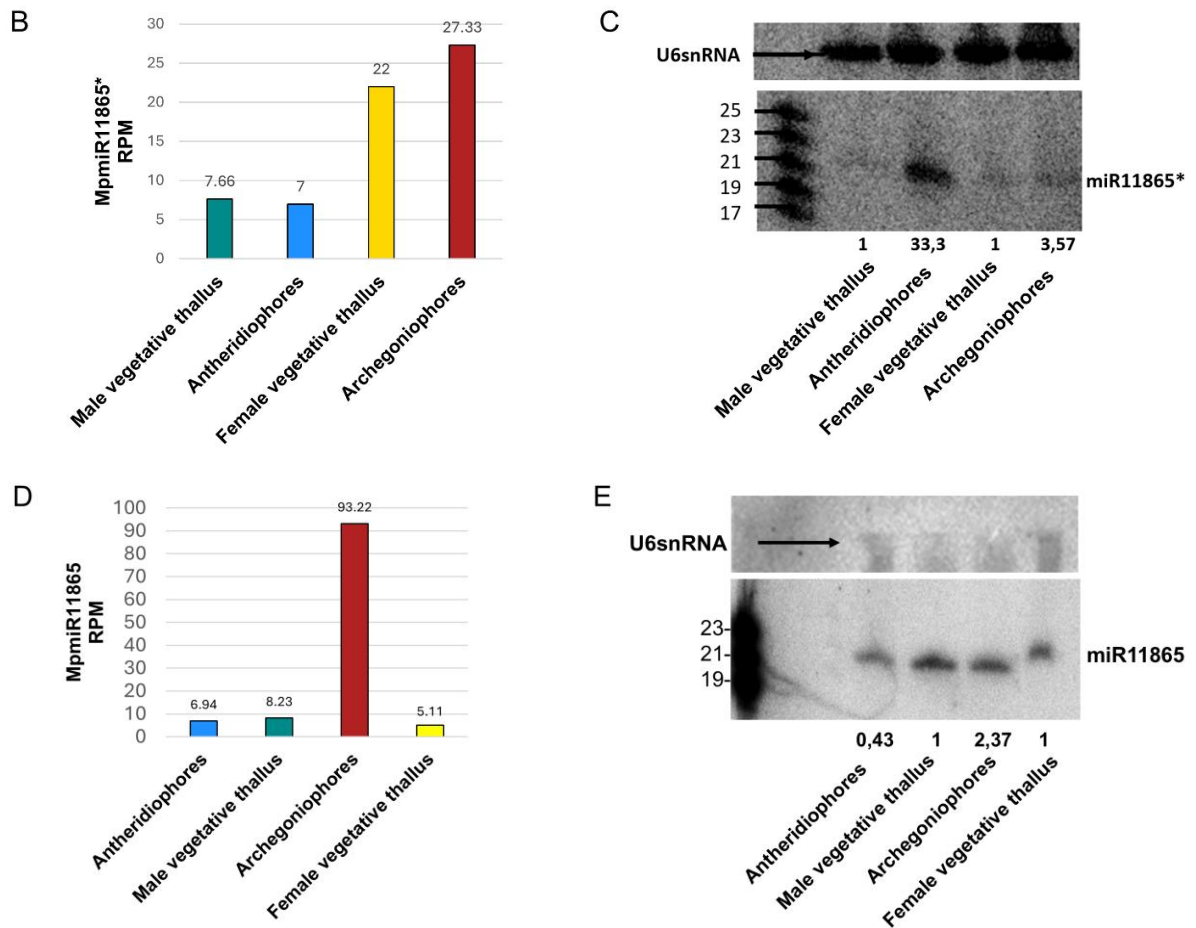


Figure 17. MpmiR11865*/11865 accumulation patterns. (A) Structure of pre-MpmiR11865 as predicted by RNA Fold (Folder ver. 1.11); miRNA sequence is highlighted in red and miRNA* sequence is highlighted in orange; the minimum free energy (ΔG) of predicted structure is written on the right side. (B) & (D) sRNA NGS sequencing results for MpmiR11865* and MpmiR11865, respectively, with normalized read counts presented above each bar; RPM- Reads Per Million. (C) & (E) Northern hybridization; the numbers presented below northern blot show the relative strengths of signal of miRNA bands; control signals were arbitrarily established as 1; signal intensity differences were separately obtained for male vegetative thalli control/antheridiophores, female vegetative thalli control/archegoniophores; Left side of the blot - RNA marker depicting 17-25 nucleotide long RNAs; U6 snRNA hybridization-RNA loading control. All the images have been derived from (Aggarwal et al., 2024) with some modifications.

5.1.2.2 Pri-MpmiR11865 transcript is mainly expressed in gametangioophores of Marchantia

The level of pri-MpmiR11865 transcript is more abundant in both antheridiophores and archegoniophores, which correlates with the northern blot results showing strong accumulation of MpmiRNA11865* in antheridiophores, and strong accumulation of MpmiR11865 in archegoniophores (Fig. 18). However, the differential accumulation pattern of MpmiR11865* and MpmiR11865 in the respective gametangioophores suggests that functional miRNAs derived from the miRNA/miRNA* duplex are differently selected by AGO1 in antheridiophores and archegoniophores. This also suggests that MpmiR11865* and MpmiR11865 might have separate regulatory functions in antheridiophores and archegoniophores, respectively. This also highlights the complexity of these miRNAs regulation in Marchantia.

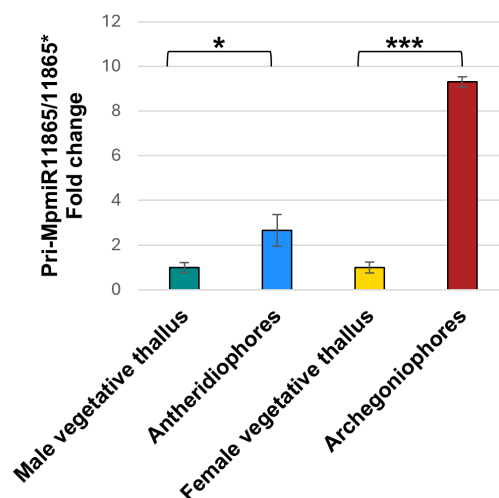


Figure 18. Pri-MpmiR11865 expression pattern in generative organs of Marchantia. RT-qPCR expression level of pri-MpmiR11865*/11865 transcript; *p-value < 0.05; ***p-value < 0.001. The image has been derived from Aggarwal et al., (2024) with some modifications.

5.1.2.3 Identification of full-length transcript of MpMIR11865 gene

MarpolBase analysis revealed that pre-MpmiR11865 sequence aligns with the genomic sequence on chromosome 5. Primers were designed within this region and extended towards the most distal ends in the 5' and 3' directions. 5'-RLM and 3'-RACE experiments yielded the longest products, measuring ~937 bp and ~517 bp, respectively (Fig. 19A, 19B). Using RT-PCR, full-length transcript of MpMIR11865 gene was determined (~1588 bp), which was confirmed by Sanger DNA sequencing as well (Fig. 19C).

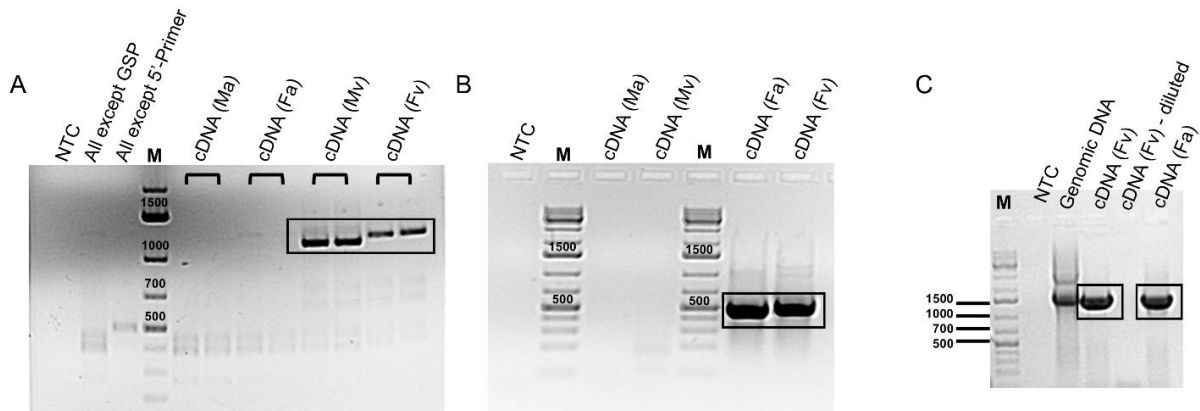


Figure 19. Agarose gel electrophoresis of products obtained with 5'-RLM (A), 3'-RACE (B) and RT-PCR (C) for the full-length of pri-MpmiR11865. The specific amplified products have been boxed. For 5'-RLM, three negative controls were used to identify specific RACE products: All components except template (NTC), all components except Gene specific primer (GSP) and all components except GeneRacer™ 5' or 3' primer; For 3'-RACE, a negative control reaction without template (NTC) was used and in addition to this, Genomic DNA was used as a positive control for RT-PCR; cDNA templates from four different tissues were used: Ma= antheridiophores, Fa= archegoniophores, Mv= male vegetative thallus and Fv= female vegetative thallus. M – GeneRuler™ 1 kb Plus DNA Ladder.

5.1.2.4 MpMIR11865 gene structure identification

The alignment of 5'-RLM and 3'-RACE products with the genomic and transcriptomic database using CLC Genomics Workbench enabled the identification of gene structure. The MpMIR11865 gene sequence overlapped with a region on chromosome 5 (Fig. 20). Moreover, the TSS and termination site did not overlap with any protein-coding genes. This shows that MpMIR11865 gene is transcribed independently without being part of a polycistronic transcript encoding proteins. Therefore, MpMIR11865 gene represents an intron-less independent transcriptional unit.

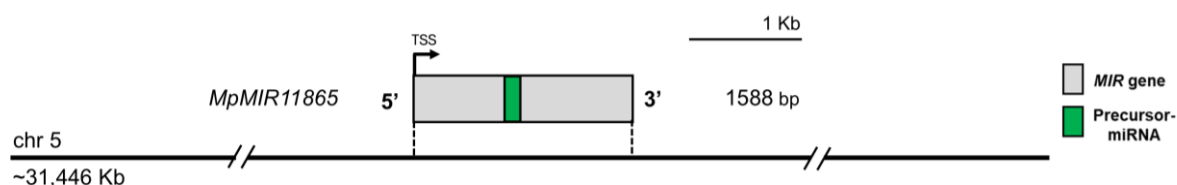


Figure 20. Schematic representation of MpMIR11865 gene structure. Chromosome 5 is represented as black horizontal line. The length of MIR gene is written on the right side and scale bar corresponds to 1 kb. The image has been derived from (Aggarwal et al., 2024) with some modifications.

5.1.2.5 Identification of putative targets of MpmiR11865 and MpmiR11865* from degradome data

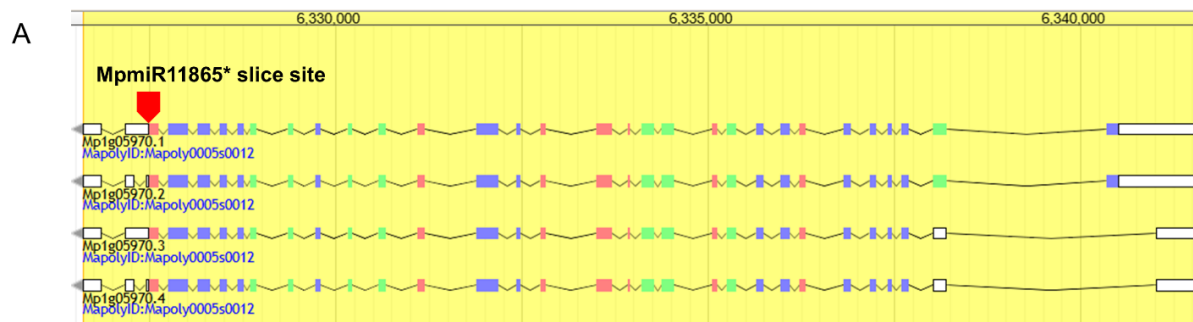
Mp1g05970 gene, which encodes a tRNA ligase1 protein was identified as the potential target of MpmiR11865* from our degradome sequencing data. This gene produces four mRNA isoforms that can be potentially cleaved. The putative cleavage site is always the same and lies within the 3'-UTR region (Fig. 21A). However, its nucleotide cleavage numeration differs depending on the isoform tested. In the t-plot analysis, a sharp peak at position 4523 (for isoform Mp1g05970.1, Mp1g05970.2) and 3863 (for isoform Mp1g05970.3, Mp1g05970.4) was observed as the dominant cleavage site. Moreover, the putative target shows high sequence complementarity to MpmiR11865* sequence (Fig. 21B). Additionally, RT-qPCR analysis revealed low expression levels of the identified target mRNA in antheridiophores, showing an inverse correlation with MpmiR11865* levels (Fig. 21C).

Similarly, for MpmiR11865, we identified a potential new target encoding a nuclear AAA+ ATPase (Valosin-containing protein subfamily) protein encoded by *Mp6g13460* gene. This gene produces two mRNA isoforms. The putative cleavage site is always the same and is located within the 3'-UTR region (Fig. 21D). However, its nucleotide cleavage numeration differs depending on the isoform tested. The t-plot analysis revealed a prominent peak at predicted position for isoform *Mp6g13460.2* and shows strong complementarity to MpmiR11865 sequence (Fig. 21E). However, for isoform *Mp6g13460.1*, the putative cleavage site gained the lowest score i.e. 4. That is why it is not shown in the presented t-plot (Fig. 21E). The degradome data was supported by RT-qPCR analysis, which revealed downregulation of *Mp6g13460* expression in archegoniophores, demonstrating a reverse correlation between expression level of target mRNA and accumulation level of MpmiR11865 (Fig. 21F).

Table 2. Putative mRNA targets for MpmiR11865* and MpmiR11865 revealed by degradome sequencing:

| miRNA ID | Target gene ID | mRNA length (nt) | mRNA:miRNA pairing region (nt) | Cleavage (nt of mRNA)+ | Cleavage (nt of mRNA)* |
|-------------|--------------------|------------------|--------------------------------|------------------------|------------------------|
| MpmiR11865* | <i>Mp1g05970.1</i> | 5093 | 4509:4533 | - | c:4523(9.34 1 48)* |
| | <i>Mp1g05970.2</i> | 4947 | 4509:4533 | - | c:4523(9.46 1 48)* |
| | <i>Mp1g05970.3</i> | 4431 | 3849:3873 | - | c:3863(9.74 1 48)* |
| | <i>Mp1g05970.4</i> | 4287 | 3850:3873 | - | c:3863(9.93 1 48)* |
| MpmiR11865 | <i>Mp6g13460.1</i> | 5468 | 5029:5047 | - | c:5037(2.91 4 33)* |
| | <i>Mp6g13460.2</i> | 4745 | 4621:4639 | - | c:4629(3.58 1 33)* |

Cleavage occurs between the 9th/10th nucleotide (+) or 10th/11th nucleotide (*) of the miRNA sequence.



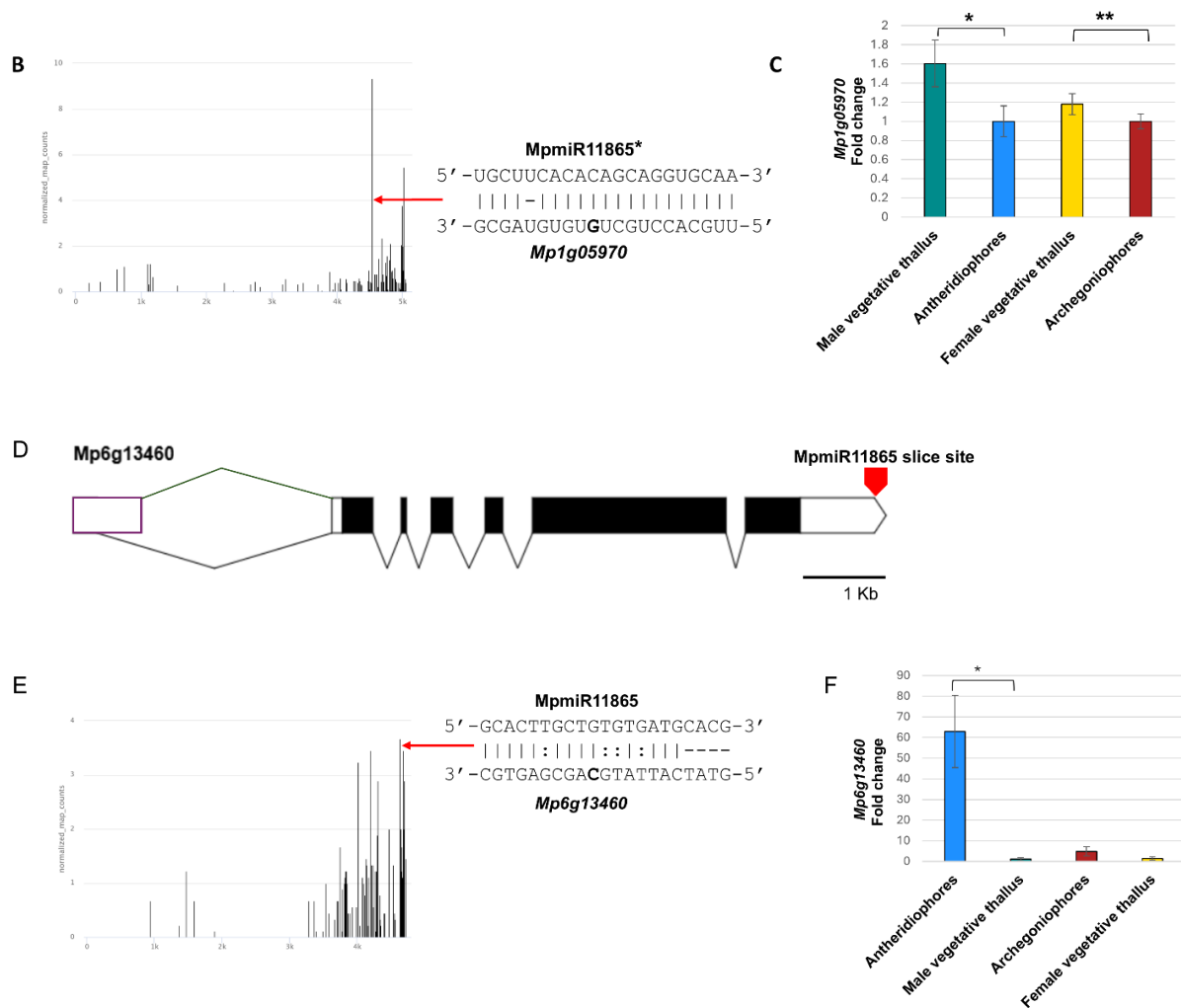


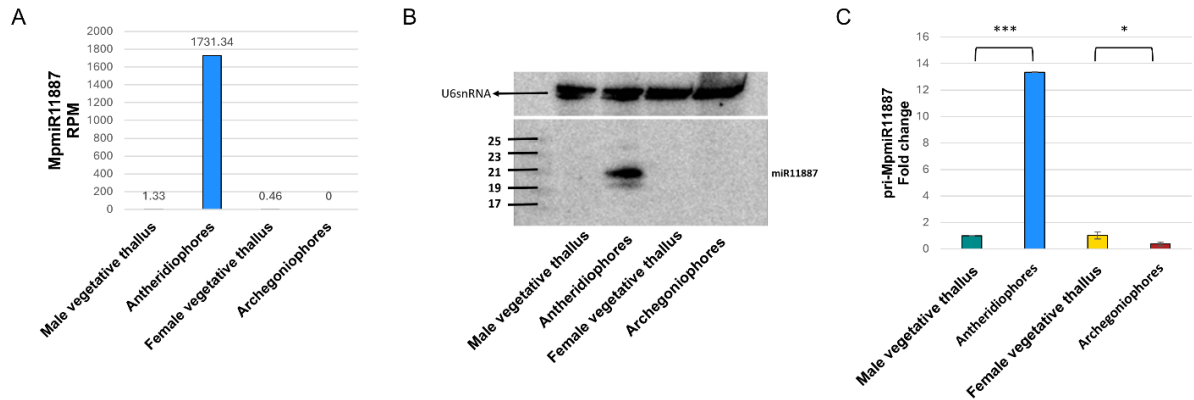
Figure 21. Putative mRNA targets of MpmiR11865* and MpmiR11865. (A) & (D) Structures of *Mp1g05970* and *Mp6g13460* gene; the putative slice site of respective miRNA is shown by downward pointing arrow. (B) & (E) Based on degradome data, the t-plot of corresponding putative target mRNAs: *Mp1g05970* and *Mp6g13460*; the predicted cleavage site is pointed with red arrow; Alongside the t-plot, complementary base pairing between miRNA and target mRNA is presented; cleavage site is indicated by nucleotide marked in bold. (C) & (F) RT-qPCR expression level of *Mp1g05970* and *Mp6g13460* transcripts; *p-value < 0.05; *p-value < 0.05. The images: (B), (C), (E) and (F) have been derived from (Aggarwal et al., 2024) with some modifications.

5.2 Functional characterization of liverwort-specific MpmiR11887

After analyzing the accumulation profile of liverwort-specific miRNAs in sexual organs of *Marchantia*, we decided to perform detailed studies on two examples: MpMIR11887 which is exclusively expressed in antheridiophores (male sexual organs) and MpMIR11796 which shows dominant expression in archegoniophores (female sexual organs). BLAST analysis in onekp database version 5 (The 1000 Plants Project, <https://db.cngb.org/blast/>) revealed that MpmiR11887 is present only in *M. polymorpha*, while MpmiR11796 is present in two liverwort species: *Lunularia cruciata* and *M. polymorpha*.

5.2.1 Pri-MpmiR11887 and MpmiR11887 presence exclusively in antheridiophores

sRNA NGS data and northern hybridization results were congruent, showing that MpmiR11887 accumulates exclusively in antheridiophores (Fig. 22A, 22B) (Pietrykowska, 2020; Aggarwal et al., 2024). The pre-MpmiR11887 is relatively long in this case but still maintains a classical stem loop structure (Fig. 22D). RT-qPCR analysis further detected almost exclusively high expression of pri-MpmiR11887 in antheridiophores (Fig. 22C). Both pri-MpmiR11887 and mature MpmiR11887 are exclusively present in antheridiophores, which suggests that the level of MpmiR11887 is transcriptionally regulated in male reproductive organs. The exclusive accumulation of mature MpmiR11887 in antheridiophores indicates its potential significance in the development of male reproductive structures in *Marchantia*.



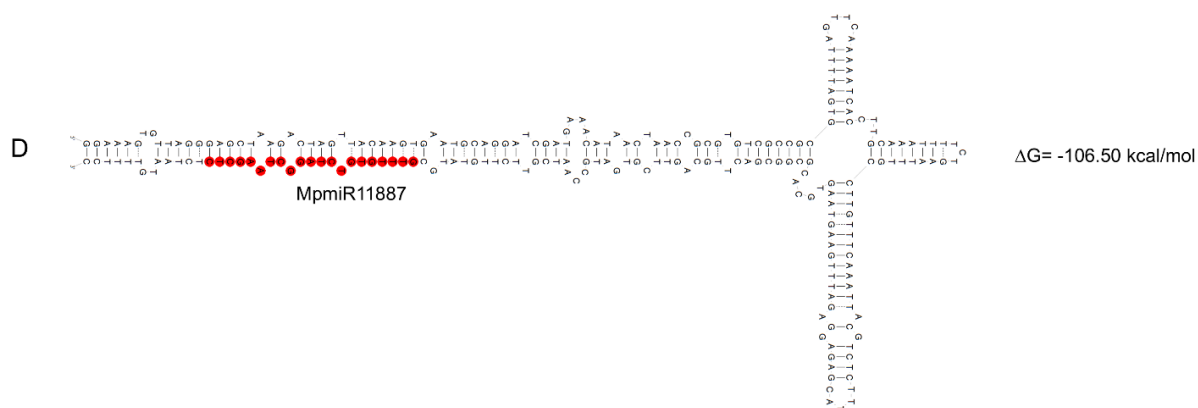


Figure 22. MpmiR11887 accumulation and pri-MpmiR11887 expression pattern. (A) sRNA NGS sequencing results with normalized read counts presented above each bar; RPM- Reads Per Million. (B) Northern hybridization; Left side of the blot - RNA marker depicting 17-25 nucleotide long RNAs; U6 snRNA hybridization - RNA loading control. (C) RT-qPCR analysis for expression pattern of pri-MpmiR11887; *p-value < 0.05; ***p-value < 0.001. (D) Structure of pre-MpmiR11887 as predicted by RNA Fold (Folder ver. 1.11); miRNA sequence is highlighted in red and minimum free energy (ΔG) of predicted structure written on the right side. The images have been derived from (Aggarwal et al., 2024) with some modifications.

5.2.2 Identification of full-length transcript of *MpMIR11887* gene

The analysis of Marchantia genomic database revealed that MpmiR11887 sequence overlaps with the 3'-UTR of *Mp6g01830* gene, which encodes a 65-amino acid putative protein of unknown function. The pre-miRNA sequence was retrieved from the database, and primers were designed within this region. 5'-RLM and 3'-RACE reactions were performed, which yielded the longest products of approximately 575 bp and ~560 bp, respectively (Fig. 23A, 23B). RT-PCR confirmed that the full-length of *MpMIR11887* gene is ~1211 bp (Fig. 23C).

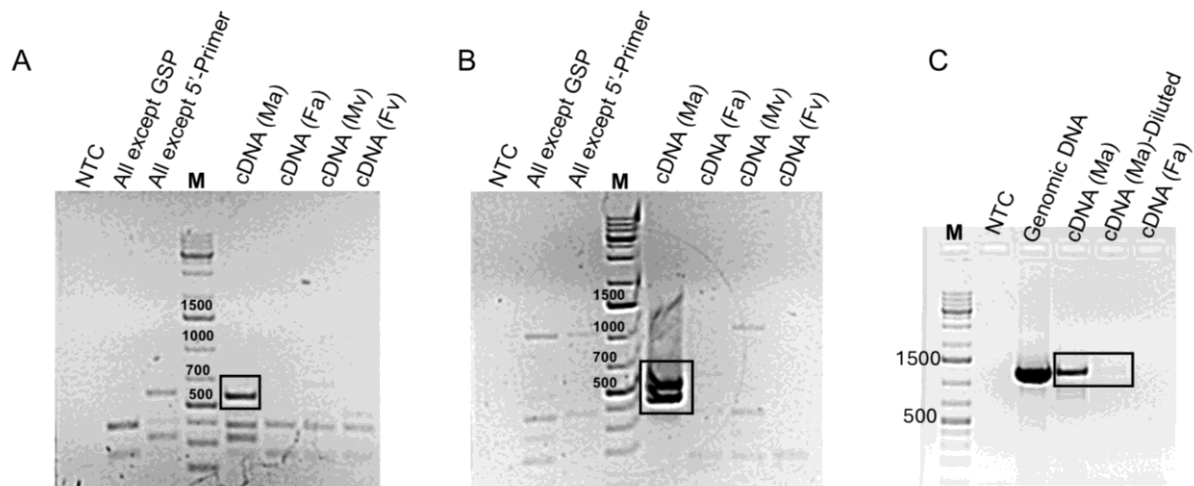


Figure 23. Agarose gel electrophoresis of products obtained with 5'-RLM (A), 3'-RACE (B) and RT-PCR (C) for the full-length of pri-MpmiR11887. The specific amplified products have been boxed. For 5'-RLM and 3'-RACE, three negative controls were used to identify specific RACE products: All components except template (NTC), all components except Gene specific primer (GSP) and all components except GeneRacer™ 5' or 3' primer; For RT-PCR, a negative control reaction without template (NTC) and genomic DNA was used as a positive control; cDNA templates from four different tissues were used: Ma= antheridiophores, Fa= archegoniophores, Mv= male vegetative thallus and Fv= female vegetative thallus. M – GeneRuler™ 1 kb Plus DNA Ladder.

5.2.3 MpMIR11887 gene structure identification

The sequences obtained from 5'-RLM and 3'-RACE were aligned with the genomic and transcriptomic database using CLC Genomics Workbench, enabling the identification of gene structures. The MpMIR11887 gene sequence was found to overlap with the sequence of a putative protein coding gene (Mp6g01830). The TSS of MpMIR11887 was observed to be the same as its putative host gene, while the termination site extended beyond the host gene's termination site (Fig. 24). This finding strongly suggests that MpMIR11887 gene represents independent transcriptional unit while putative 65 AA long upstream sequence of pre-MpmiR11887 may represent miPeP (Lauressergues et al., 2015; Lauressergues et al., 2022). Since two termination

sites could be observed in 3'-RACE reaction, it could imply that there are some regulatory elements controlling the termination of the gene transcription.

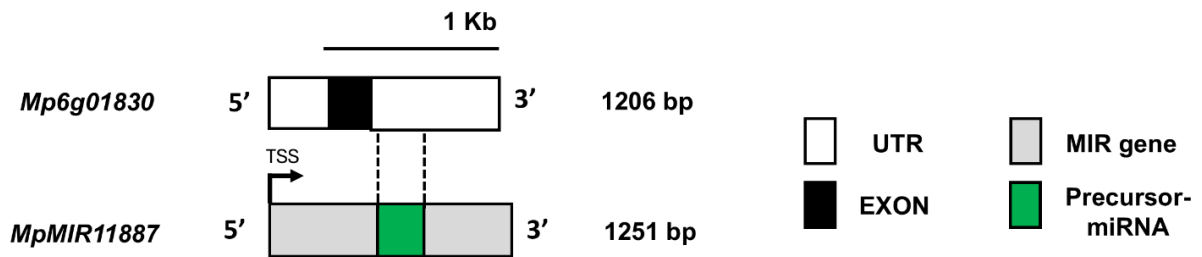


Figure 24. Schematic representation of Mp*MIR11887* gene structure. Mp6g01830, the protein coding gene, (upper row) overlapping with Mp*MIR11887* gene (lower row). The length of protein-coding gene and *MIR* gene is written on the right side. The scale bar corresponds to 1 kb. The diagram has been derived from (Aggarwal et al., 2024) with some modifications.

5.2.4 Characterization of Mp*MIR11887* gene promoter *in planta* activity

To investigate the *in-vivo* expression levels of Mp*MIR11887* gene, the promoter sequence corresponding to ~1.9 kb upstream of the mature MpmiR11887 was fused to the GUS reporter gene (Fig. 25). The obtained transgene was then introduced into the *Marchantia* genome via *Agrobacterium*-mediated transformation. The histochemical GUS staining revealed that the promoter is active in the generative organs of *Marchantia* i.e. in the young antheridia and spermatogenous cells (Fig. 26A, 26B). Additionally, GUS expression was also observed in the apical notch of female vegetative thalli and archegonia (Fig. 26C).

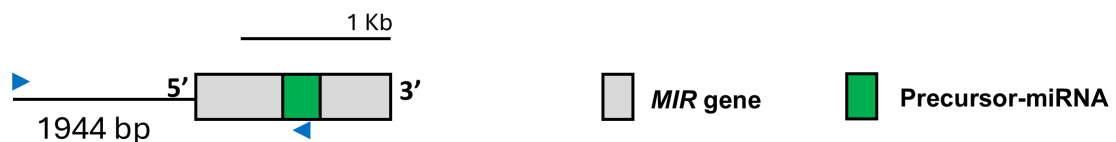


Figure 25. Schematic diagram showing the selection of promoter region used to drive the expression of GUS under Mp*MIR11887* gene promoter.

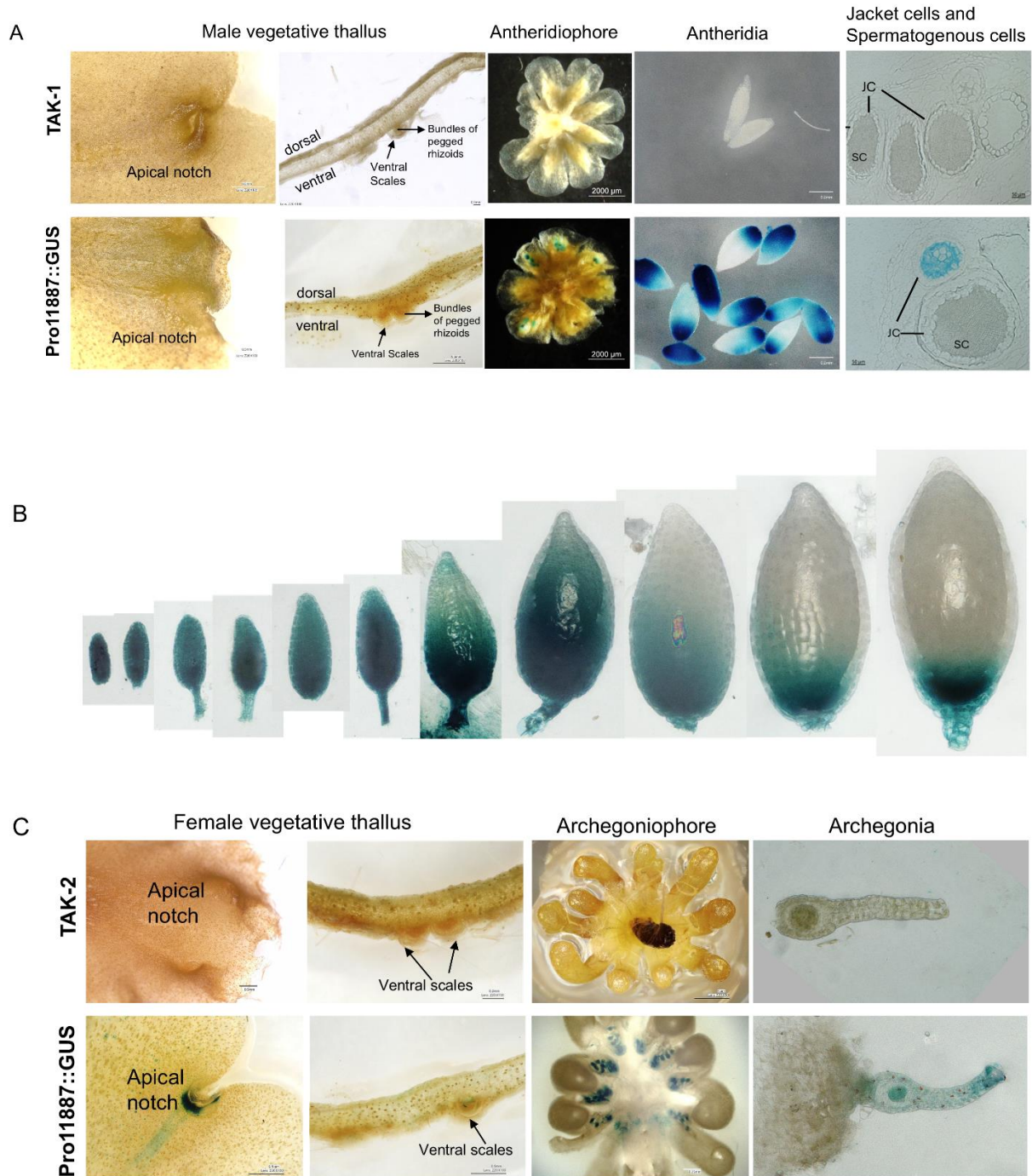


Figure 26. Spatiotemporal profile of *MpMIR1187* gene promoter activity. (A) Left to right: The GUS stained images in male tissue i.e. male vegetative thallus; antheridiophore; antheridia; high magnification images of antheridial disc cross-section showing antheridia with spermatogenous cells surrounded by jacket cells. The image on the extreme right from WT and *pro1187::GUS* were prepared by Prof. Robert

Malinowski from Institute of Plant Genetics of Polish Academy of Sciences, Poznan, Poland. (B) The GUS stained images of antheridia from stage 4 antheridiophore development (Higo et al., 2016) are shown in ascending order based on their size; young to mature stage shown left to right, respectively. (C) Left to right: The GUS stained images in female tissues i.e. female vegetative thallus and archegoniophore, including archegonia.

5.2.5 Analysis of CRISPR/Cas9 KO plants for MpMIR11887

5.2.5.1 Genotyping and validation of MpMIR11887 CRISPR/Cas9 knockouts

To determine the biological functions of MpmiR11887 in *M. polymorpha*, Pietrykowska H. generated knock-out (KO) mutants using CRISPR/Cas9 system. A set of guide RNAs were designed (Fig. 27), sgRNA1 targeting the 5'-end of the pre-MpmiR11887 (including miRNA* region) and sgRNA2 targeting the mature MpmiR11887 region. $\Delta mpmir11887^{#1,8,10,11,13}$ mutants were established using sgRNA1, whereas $\Delta mpmir11887^{#3,5,6,9}$ mutants were established using sgRNA2. I confirmed the presence of mutations in these mutant plants (8 male and 1 female mutant plants) (Fig. 28, Table 3).

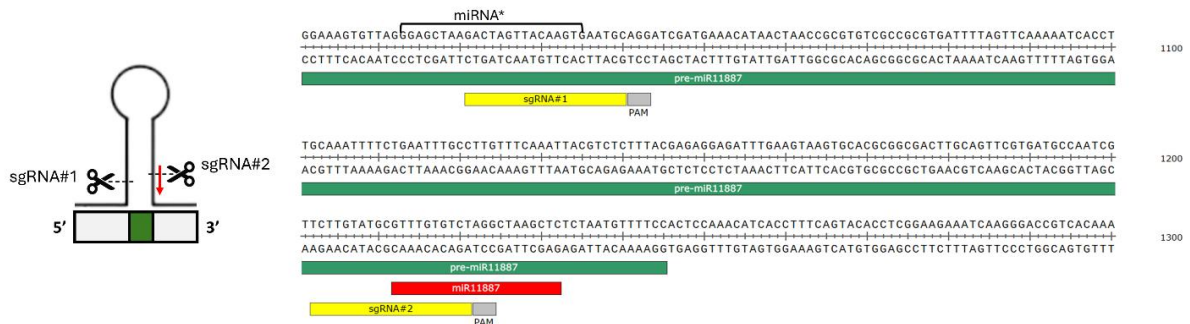


Figure 27. Schematic representation of the positions of sgRNA1 and sgRNA2 (in yellow), PAM sequence (highlighted in grey on the image on right) targeting the pre-MpmiR11887 (in green) and mature MpmiR11887 region (highlighted in red on the image on right), respectively; sgRNA2 overlaps 13 nt out of 21 nt miRNA sequence. Below the stem-loop structure shown on the left side, MpMIR11887 gene structure is shown.

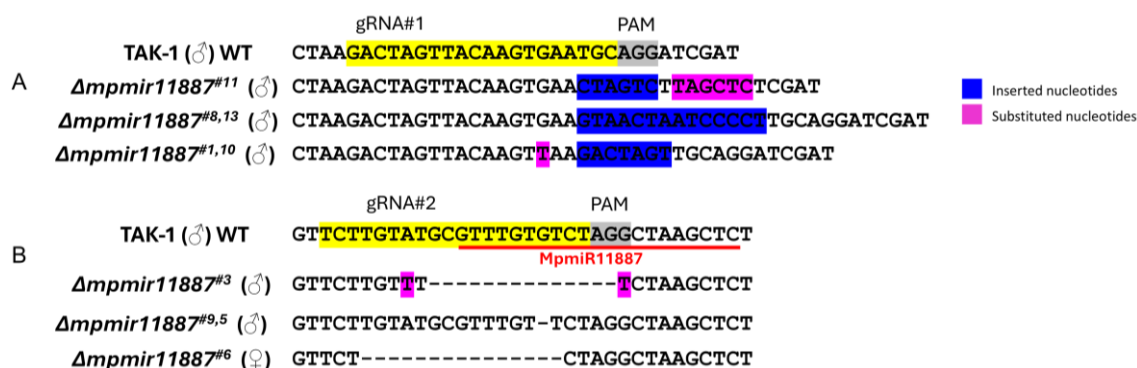


Figure 28. Schematic representation of the sequence editing in the *Δmpmir11887*^{ge} mutant plants that led to mutations within mature miRNA region and downstream of miRNA* region.

Table 3. List of CRISPR/Cas9 genome-edited mutant plants for MpMIR11887 gene with their sex and genotypes.

| Sr. No. | Mutant plant | Sex | Mutation type |
|---------|-----------------------------------|-----|--|
| 1. | <i>Δmpmir11887</i> ^{#1} | M | Insertion of 7 nt, substitution of 1 nt in pre-miRNA region encompassing 1 nt substitution within miRNA* region |
| 2. | <i>Δmpmir11887</i> ^{#3} | M | Deletion of 14 nt encompassing 10 nt deletion from mature miRNA region |
| 3. | <i>Δmpmir11887</i> ^{#5} | M | Deletion of 1 nt in mature miRNA region |
| 4. | <i>Δmpmir11887</i> ^{#6} | F | Insertion of 14 nt in pre-miRNA region encompassing 8 nt deletion from miRNA region |
| 5. | <i>Δmpmir11887</i> ^{#8} | M | Insertion of 14 nt in pre-miRNA region |
| 6. | <i>Δmpmir11887</i> ^{#9} | M | Deletion of 1 nt in mature miRNA region |
| 7. | <i>Δmpmir11887</i> ^{#10} | M | Insertion of 7 nt , substitution of 1 nt in pre-miRNA region encompassing 1 nt substitution within miRNA* region |
| 8. | <i>Δmpmir11887</i> ^{#11} | M | Insertion of 6 nt , substitution of 2 nt in pre-miRNA region |
| 9. | <i>Δmpmir11887</i> ^{#13} | M | Insertion of 14 nt in pre-miRNA region |

Northern blot hybridization was used to confirm the absence of MpmiR11887 in $\Delta mpmir11887^{#3}$, $\Delta mpmir11887^{#9}$, $\Delta mpmir11887^{#11}$ male mutant plants, confirming that the mutations in these plants resulted in the loss of miRNA expression (Fig. 29). In all $\Delta mpmir11887^{ge}$ mutant plants, the structure of pre-MpmiR11887 showed conformational changes except $\Delta mpmir11887^{#9}$ and $\Delta mpmir11887^{#5}$ where a single nucleotide was deleted from mature miRNA region (Fig. 30). Surprisingly, even though no conformational changes were observed in precursor structure in $\Delta mpmir11887^{#9}$, it was a complete KO. Hence, subsequent experiments were conducted using $\Delta mpmir11887^{#3,9,11}$ KO plants.

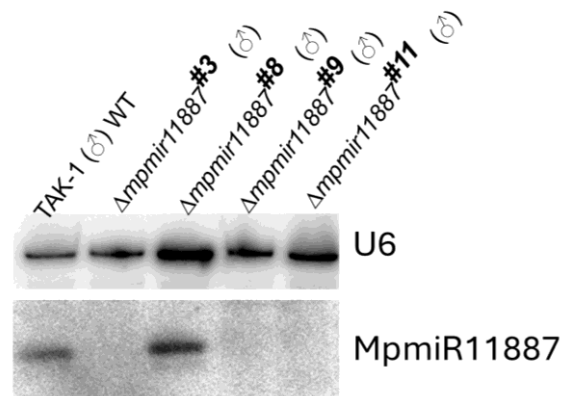


Figure 29. Northern blot analysis of $\Delta mpmir11887^{ge}$ mutant plants. U6 was used as RNA loading control.

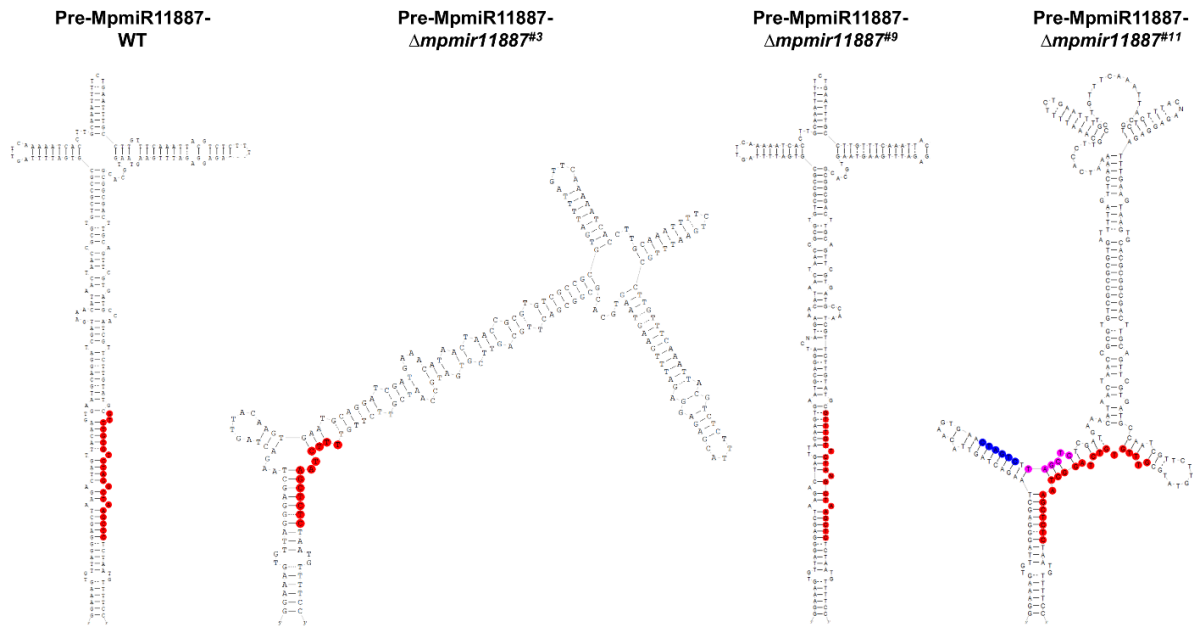


Figure 30. Secondary structures of the pre-MpmiR11887 in WT and $\Delta mpmir11887^{ko}$ ($\Delta mpmir11887^{#3}$, $\Delta mpmir11887^{#9}$, $\Delta mpmir11887^{#11}$) mutant plants as predicted by RNA Fold (Folder ver. 1.11); Both miRNA region in WT and in mutant pre-MpmiR11887, where only a part of miRNA sequence is affected, are highlighted in red. In the case of $\Delta mpmir11887^{#9}$ only 1nt was deleted from the middle part of mature miRNA sequence.

5.2.5.2 Phenotypic analysis of $\Delta mpmir11887^{ge}$ mutant plants

As described earlier, *Marchantia* undergoes vegetative growth via gemmae, which develop into a dichotomous branching plant body called the thallus. Initial observations of the phenotypic changes in developing gemmae at varying time intervals were recorded in all the three KO mutant plants ($\Delta mpmir11887^{#3}$, $\Delta mpmir11887^{#9}$, $\Delta mpmir11887^{#11}$) (Fig. 31). During the early stages of gemmae development, the KO plants exhibited larger thallus area in comparison to WT Tak-1 plants. However, despite the differences in early growth dynamics of thallus, by two weeks, the size of KO plants was equivalent to that of WT plants. Additionally, in 5-week-old KO plants, the antheridiophores had already developed at a mature age, whereas, in WT plants, they were only beginning to emerge (Fig. 32A, 32B). This suggests that MpmiR11887 regulates the developmental timing of male gametophyte production. Consistent with

the increased size of vegetative thalli, the diameter of antheridial discs in KO plants were also significantly wider than those in WT plants (Fig. 32C, 32D). Moreover, the mutant antheridial discs contained larger antheridia more frequently. The size of mature antheridia in KO mutant plants was often noticeably bigger than in WT plants (Figure 32E, 32F). While antheridial discs in WT plants harbored only a few large mature antheridia, the antheridial discs in KO plants featured a greater number of mature antheridia clustered around larger size values (Fig. 32F). Upon application of water to the dorsal surface of the antheridiophores, both WT and KO plants discharged white sperm masses with no noticeable differences. Microscopic examination confirmed that the sperm cells produced by KO plants were motile.

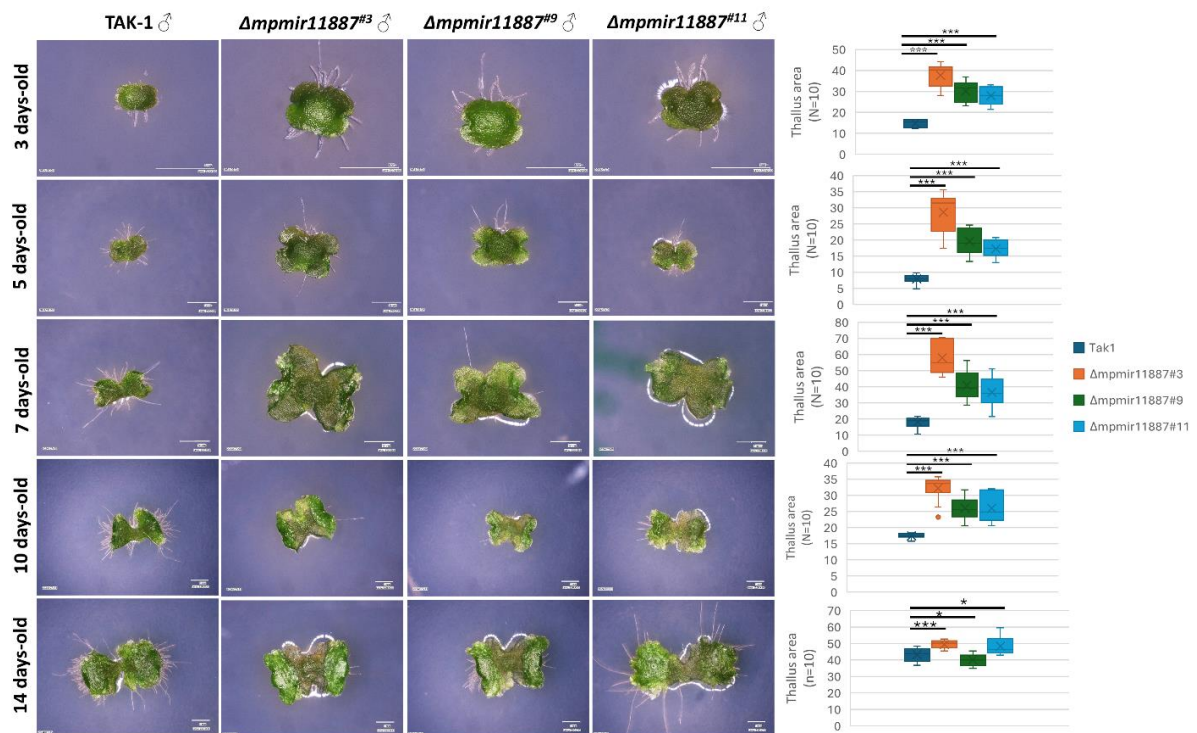


Figure 31. Dorsal view of growing gemma from WT Tak-1 plants and $\Delta mpmir11887^{ko}$ plants ($\Delta mpmir11887^{\#3}$, $\Delta mpmir11887^{\#9}$, $\Delta mpmir11887^{\#11}$). The growth interval points have been mentioned on the left i.e. from 3-14 days old. The thallus area was measured from 10 individual gemmae and the box plots on the right side show the thallus area differences between WT and KO plants. The thallus area was measured using ImageJ software. The statistics were performed using student's t-test (*p-value < 0.05; ***p-value < 0.001).

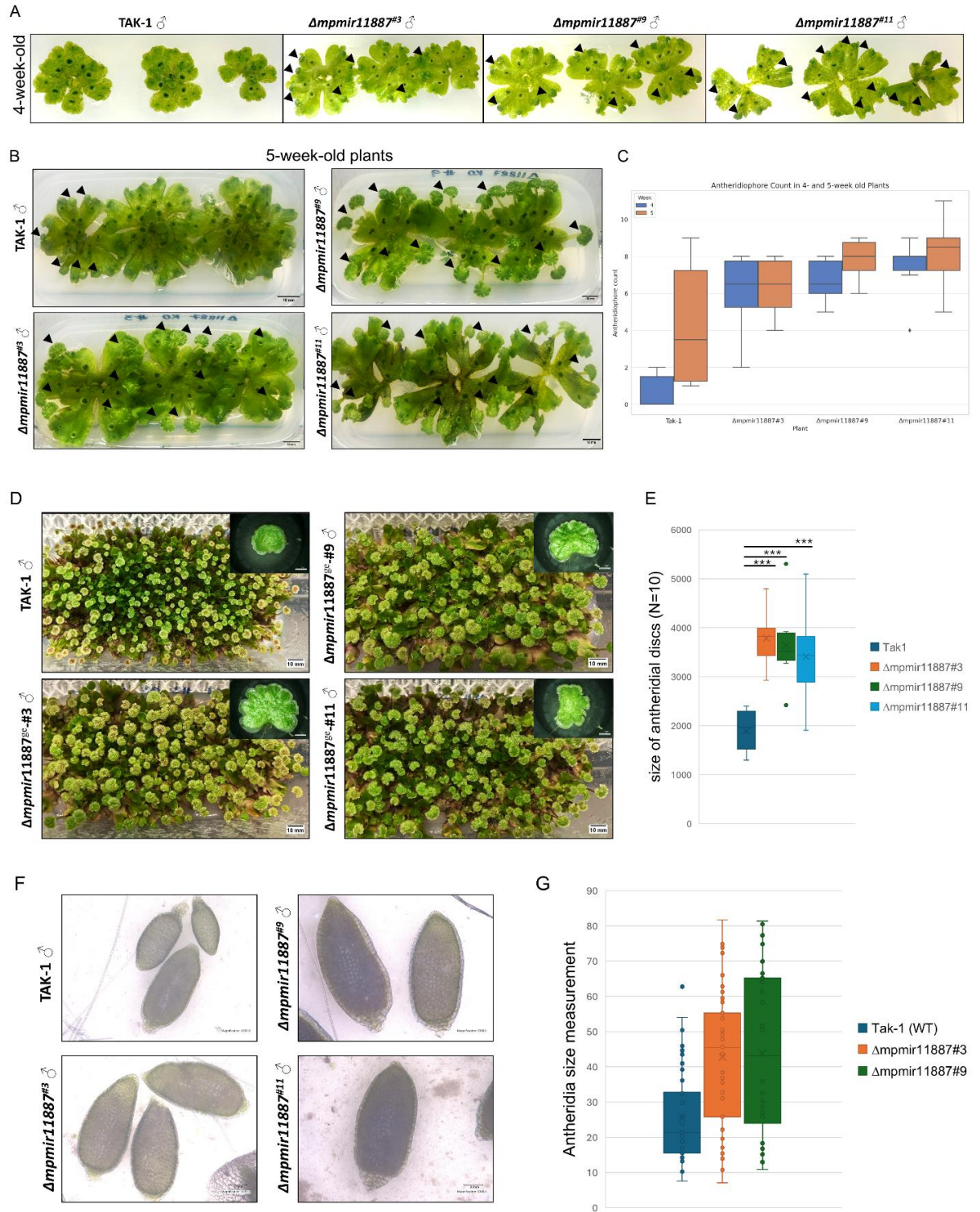


Figure 32. Phenotypic analysis of antheridiophores and antheridia in WT and $\Delta mpmir11887^{ko}$ plants ($\Delta mpmir11887^{#3}$, $\Delta mpmir11887^{#9}$, $\Delta mpmir11887^{#11}$). (A)

Photographs of 4-week-old plants subjected to FR light to induce sexual organ formation in WT Tak-1 and $\Delta mpmir11887^{ko}$ plants. (B) Photographs of 5-week-old WT Tak-1 and $\Delta mpmir11887^{ko}$ plants. (C) Box-plot showing the number of antheridiophores in 4- and 5-week old WT and $\Delta mpmir11887^{ko}$ plants ($n = 6$). (D) Photographs of antheridiophores in 5-week-old WT Tak-1 and $\Delta mpmir11887^{ko}$ plants. (E) Box-plot showing the differences in the size of antheridial discs in WT and $\Delta mpmir11887^{ko}$ plants. The statistics were performed using student's t-test (**p-value < 0.001). (F) Microscopic images of isolated antheridia from antheridial disc of WT and $\Delta mpmir11887^{ko}$ plants. (G) Box-plot showing the differences in the size of antheridia in WT and $\Delta mpmir11887^{ko}$ plants.

Larger antheridia might produce more sperm cells, which might lead to higher fertilization efficiency. To check the reproductive success, I crossed WT Tak-2 (♀) plants with sperm cells from WT Tak-1 (♂) plants, $\Delta mpmir11887^{#3}$ (♂) and $\Delta mpmir11887^{#9}$ (♂) KO plants. Preliminary observations indicated that KO plants were capable of successful fertilization, as yellow-colored mature sporophytes were formed in all the cases (Fig. 33). We can conclude that sperm cells produced by $\Delta mpmir11887^{#3}$ and $\Delta mpmir11887^{#9}$ KO plants are capable of fertilization. However, quantitative analyses are required to determine whether fertilization efficiency is affected when KO sperm cells are used to fertilize WT Tak-2 archegonia and female mutant i.e. $\Delta mpmir11887^{#6}$.

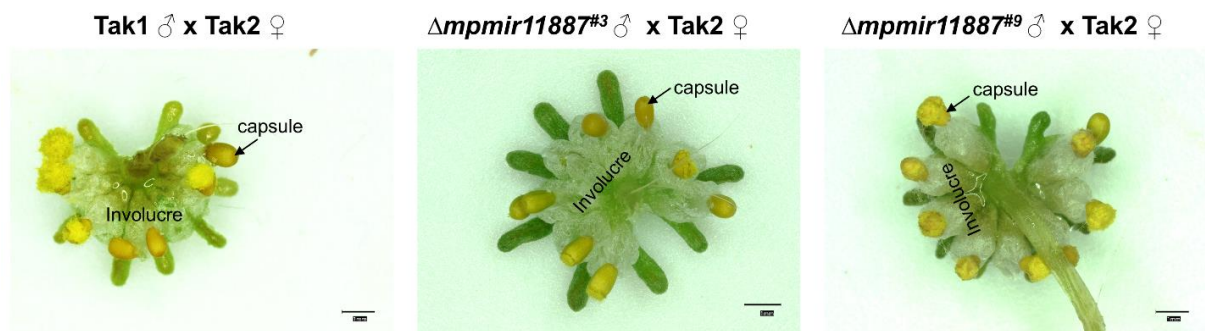


Figure 33. Microscopic images of the sporophytes obtained by crossing WT and KO sperm cells with WT Tak-2 plants. Yellow capsules can be seen as exposed outside the involucre due to seta elongation.

5.2.6 Identification of putative target of MpmiR11887 from transcriptomic and degradome data analyses

Degradome sequencing data indicated that MpmiR11887 targets Mp1g20730 mRNA, which encodes β -tubulin protein. RT-qPCR revealed that Mp1g20730 mRNA is downregulated in antheridiophores, showing reverse correlation with MpmiR11887 level (Aggarwal et al., 2024). However, RACE assays could not detect the cleavage of this target mRNA. RT-qPCR analysis further showed downregulation of Mp1g20730 mRNA (encoding β -tubulin) in antheridiophores and sperm cells of $\Delta mpmir11887^{ko}$ plants (graphs not shown in the thesis). Since we expected upregulation of target mRNA in $\Delta mpmir11887^{ko}$ plants, we concluded that selected mRNA is not the target for MpmiR11887.

Based on degradome data, I selected another gene: Mp7g16780 gene as a potential target for MpmiR11887. This gene encodes a chromatin related protein, MpATX1 and produces two mRNA isoforms that could be cleaved by MpmiR11887. The putative slice site was predicted to be within the coding sequence in this case. The degradome data was supported by the upregulated expression levels of MpATX1 transcript in the antheridiophores of all $\Delta mpmir11887^{ko}$ plants as compared to WT (Fig. 34A). Furthermore, I found that MpATX1 is expressed at extremely low levels in the sperm cells [Fig. 34B; (Tan et al., 2023)], and interestingly, MpATX1 levels were found to be upregulated in the sperm cells of all $\Delta mpmir11887^{ko}$ plants (Fig. 34C). Given that ATX genes encode histone-lysine N-methyltransferases and are involved in histone modifications, and additionally histones are replaced by protamines during spermatogenesis, I decided to examine the level of protamine mRNA – encoded by MpPRM gene (Mp3g14390) in Marchantia. Interestingly, an increased level of this protamine mRNA was observed in sperm cells of all $\Delta mpmir11887^{ko}$ plants as compared to WT sperm cells (Fig. 34D).

Despite repeated attempts, RACE assays were unsuccessful in detecting the cleavage site on MpATX1 mRNA, as all the products obtained were 40 bp upstream of the predicted slice site. Since a strong promoter activity of MpMIR11887 was observed in antheridia, RNA-seq libraries were prepared from antheridia isolated from WT Tak-1 and $\Delta mpmir11887^{#3}$ KO mutant using three biological replicates for each, expecting to

facilitate the mRNA target identification (described in Chapter 4 section 21). Although an increased transcript levels of *Mp1g20730* (β -tubulin), *MpATX1*, and *MpPRM* was observed in the transcriptomic dataset (derived from antheridia and not from whole antheridiophores and sperm cells), the fold change values were low and not statistically significant. Since this part of the study was inconclusive, I looked for other potential mRNA targets. Bioinformatic analysis of transcriptomic data combined with degradome sequencing data was done to identify mRNA targets that are significantly upregulated in *Δ mpmir11887^{#3}* KO and contain sequence complementary with *MpmiR11887* (analysis performed by Prof. W. Karlowski). Surprisingly, the transcriptomic analysis identified a single potential target gene, *Mp4g22015*. Unfortunately, this gene encodes an uncharacterized protein. The putative slice site lies within the 5'-UTR region and the t-plot indicated a significant peak at predicted cleavage site (Fig. 34F).

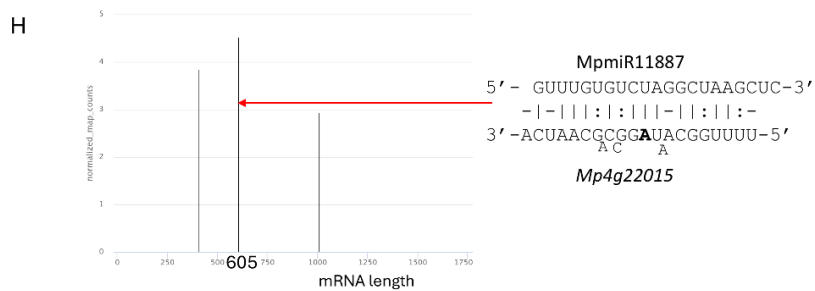
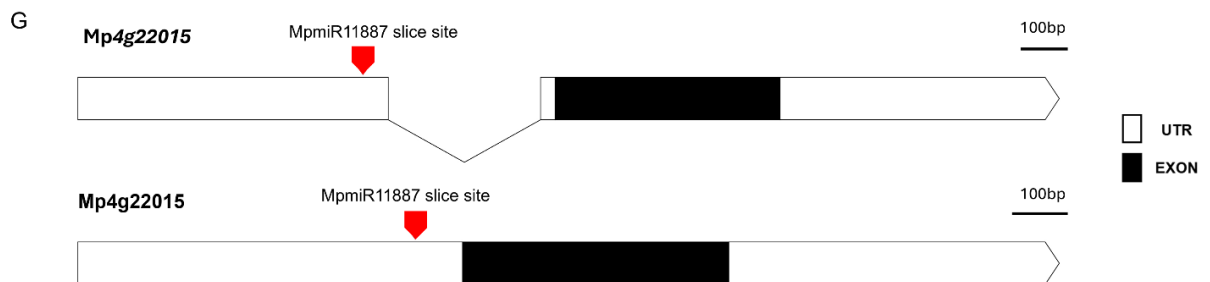
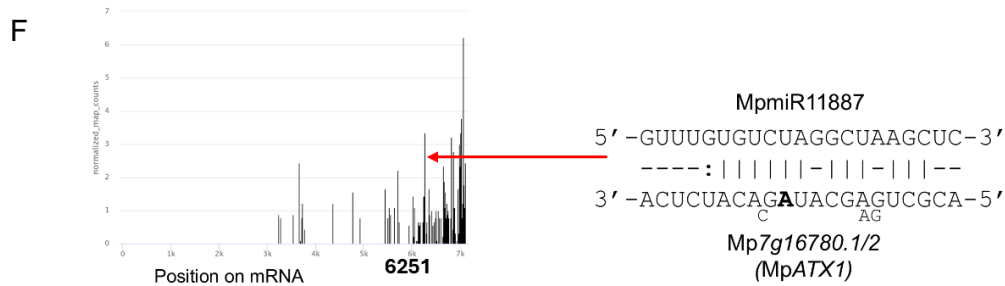
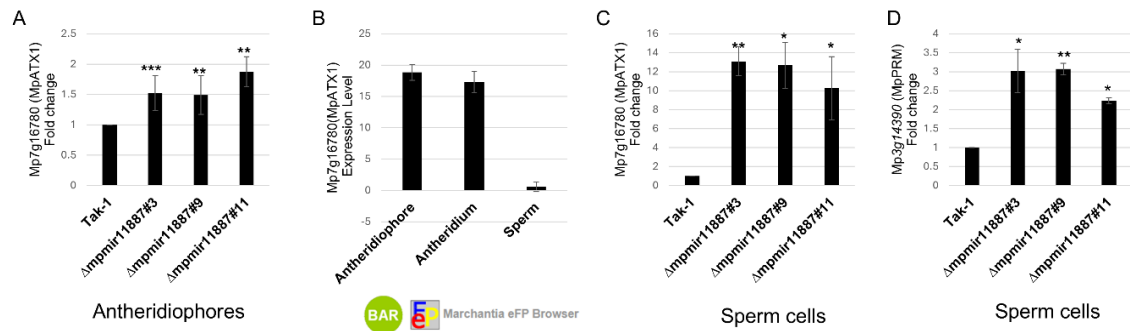
Table 4. Putative mRNA target for *MpmiR11887* as revealed by transcriptomic and degradome data analyses:

| miRNA ID | Target gene ID | mRNA length (nt) | mRNA:miRNA pairing region (nt) | Cleavage (nt of mRNA)+ | Cleavage (nt of mRNA)* |
|-------------------|------------------|------------------|--------------------------------|------------------------|------------------------|
| <i>MpmiR11887</i> | <i>Mp4g22015</i> | 1758 | 595:617 | - | c:605(1.23 1 20)* |

When I performed RACE assay to experimentally validate the predicted miRNA-guided cleavage of *Mp4g22015* mRNA, the resulting RACE product was a chimeric fragment comprising 108 bp fragment aligned to chromosome 4 fused to a portion of *Mp4g22015* mRNA sequence, 50 bp downstream of predicted slice site. There are some technical and scientific explanations for this:

- 1) During the ligation of adaptors at 5'-ends or during reverse transcription step, non-specific events can occur, which may produce chimeric cDNA fragments.
- 2) The genome is wrongly assembled at this particular locus.

Still no firm conclusions can be currently drawn regarding the identification of *MpmiR11887* mRNA target. Further experiments are required and will be performed later.



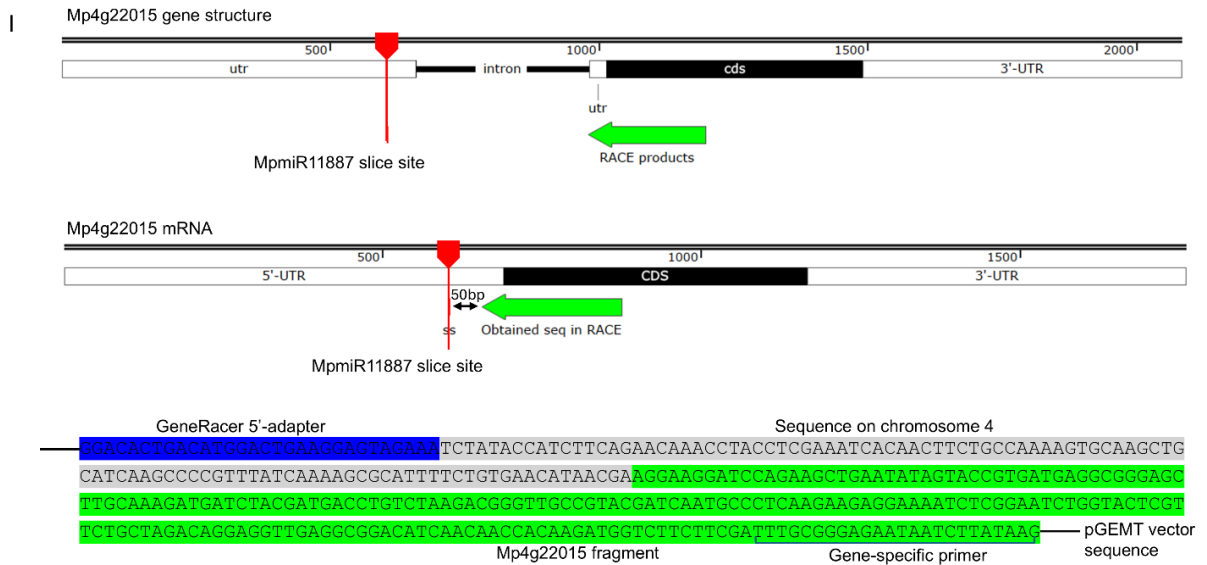


Figure 34. Putative targets of MpmiR11887. (A) RT-qPCR expression level of MpATX1 mRNA transcript in the antheridiophores of WT Tak-1 and $\Delta mpmir11887^{ko}$ plants, normalized by MpACT7 (n = 3); **p-value < 0.01; ***p-value < 0.001. (B) Expression levels of MpATX1 gene retrieved from Marchantia eFP Browser (Tan et al., 2023). (C) & (D) RT-qPCR expression level of MpATX1 and MpPRM mRNA transcripts in the sperm cells of WT Tak-1 and $\Delta mpmir11887^{ko}$ plants, normalized by MpACT7 (n=3); *p-value < 0.05; **p-value < 0.01. (E) Structure of MpATX1 mRNA; the putative slice site for MpmiR11887 is shown by red downward pointing arrow; Scale bar corresponds to 1 kb. (F) Based on degradome data, the t-plot of MpATX1 mRNA; the predicted cleavage site is pointed with red arrow; Alongside the t-plot, complementary base pairing between miRNA and target mRNA is presented; cleavage site is indicated by nucleotide marked in bold. (G) Structure of Mp4g22015 gene and mRNA; the putative slice site for MpmiR11887 is shown by red downward pointing arrow; Scale bar corresponds to 100 bp. (H) Based on degradome data, the t-plot of corresponding putative target mRNA Mp4g22015; the predicted cleavage site is pointed with red arrow; Alongside the t-plot, complementary base pairing between miRNA and target mRNA is presented; cleavage site is indicated by nucleotide marked in bold. (I) RACE results for Mp4g22015.

5.3 Functional characterization of liverwort-specific MpmiR11796

5.3.1 Pri-MpmiR11796 and MpmiR11796 are upregulated in archegoniophores

sRNA NGS and northern hybridization results from (Pietrykowska thesis, 2020) and (Aggarwal et al., 2024) showed high accumulation of MpmiR11796 in the female gametophyte - archegoniophores (Fig. 35A, 35B). The pre-MpmiR11796 also forms a classical stem loop structure (Fig. 35D). RT-qPCR analysis further detected lower levels of pri-MpmiR11796 in archegoniophores as compared to female vegetative thalli (Fig. 35C). The observation of higher accumulation of mature MpmiR11796 in archegoniophores, while lower levels of pri-MpmiR11796 in archegoniophores compared to female vegetative thalli, could be due to the fact that the pre-MpmiR11796 is more rapidly processed into mature MpmiR11796, indicating increased post-transcriptional pri-miRNA processing efficiency. This suggestion is also supported by the fact that no products could be amplified in 5'-RLM RACE in archegoniophore (Fa) tissue (Fig. 36A).

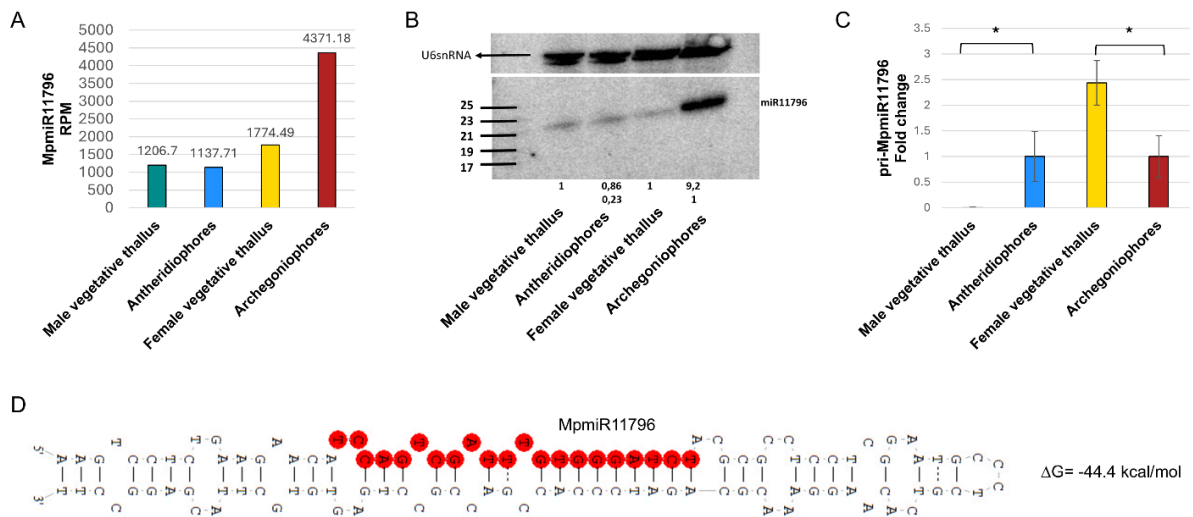


Figure 35. MpmiR11796 accumulation and pri-MpmiR11796 expression pattern. (A) sRNA NGS sequencing results with normalized read counts presented above each bar; RPM- Reads Per Million. (B) Northern hybridization; the numbers presented below the blot show the relative strengths of signal of miRNA bands; control signals were arbitrarily established as 1; signal intensity differences were separately obtained for male vegetative thalli control/antheridiophores, female vegetative thalli control/

archegoniophores (numbers in first row) and antheridiophores control/archegoniophores (numbers in second row); Left side of the blot - RNA marker depicting 17-25 nucleotide long RNAs; U6 snRNA hybridization-RNA loading control. (C) RT-qPCR analysis for expression pattern of pri-MpmiR11796; *p-value < 0.05. (D) Structure of pre-MpmiR11796 as predicted by RNA Fold (Folder ver. 1.11); miRNA sequence is highlighted in red and minimum free energy (ΔG) of predicted structure is written on the right side. The diagrams have been derived from (Aggarwal et al., 2024) with some modifications.

5.3.2 Identification of full-length transcript of MpMIR11796 gene

The analysis of Marchantia genomic database revealed that MpmiR11796 sequence overlaps with the 5'-UTR of Mp4g11670 gene, which encodes a protein of unknown function. The pre-miRNA sequence was retrieved from the database, and primers were designed within this region. 5'-RLM and 3'-RACE reactions were performed, where no products were amplified in archegoniophore (Fa) tissue suggesting lower levels of pri-MpmiR11796. The 5'-RLM and 3'-RACE yielded the longest amplified product of approximately 120 bp and ~418 bp, respectively. RT-PCR further confirmed that the full-length of MpMIR11796 gene is ~505 bp (Fig. 36).

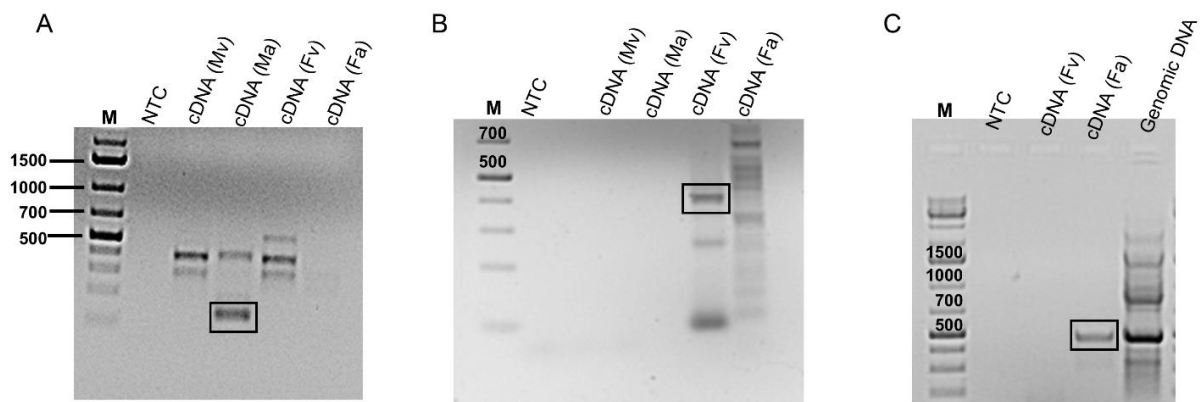


Figure 36. Agarose gel electrophoresis of products obtained with 5'-RLM (A), 3'-RACE (B) and RT-PCR (C) for the full-length of pri-MpmiR11796. The specific amplified products have been boxed. For 5'-RLM and 3'-RACE, a negative control reaction without template (NTC) was used. For RT-PCR, negative control was the reaction without template (NTC) and positive control was the reaction using genomic DNA as

template; cDNA templates from four different tissues were used: Ma= antheridiophores, Fa= archegoniophores, Mv= male vegetative thallus and Fv= female vegetative thallus. M – GeneRuler™ 1 kb Plus DNA Ladder.

5.3.3 MpMIR11796 gene structure identification

Sequences obtained from 5'-RLM and 3'-RACE were aligned with genomic and transcriptomic database using CLC Genomics Workbench, allowing the identification of gene structures. The analysis revealed that MpMIR11796 gene sequence overlapped with a putative protein-coding gene (Mp4g11670). Notably, the TSS of MpMIR11796 differed from that of its host gene, being located within the 5'-UTR region, while the termination site ended within the first intron (Fig. 37). This finding indicates that MpMIR11796 gene represents an independent transcriptional unit while sharing sequence overlap with a putative protein coding gene.

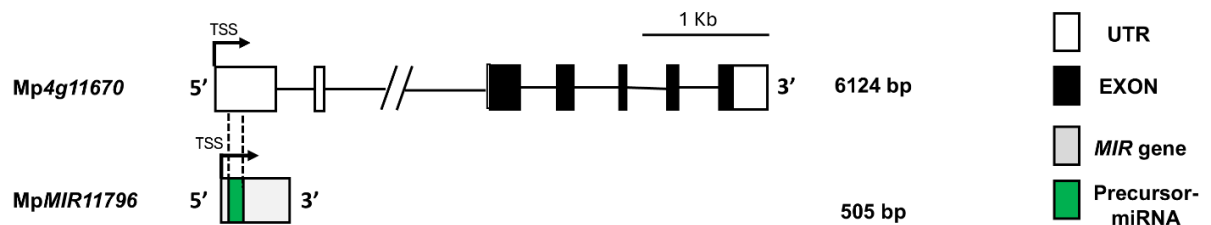


Figure 37. Schematic representation of MpMIR11796 gene structure. Mp4g11670, the protein coding gene (upper row) overlapping with MpMIR11796 gene (lower row). The lengths of protein-coding gene and MIR gene is written on the right side. The scale bar corresponds to 1 kb. The diagram has been derived from (Aggarwal et al., 2024) with some modifications.

5.3.4 Characterization of MpMIR11796 gene promoter *in planta* activity

To investigate the *in vivo* expression levels of MpMIR11796 gene, a ~2.5 kb promoter sequence upstream of the pre-MpmiR11796 was fused to the GUS reporter gene (Fig. 38). The resulting transgene was then introduced into the Marchantia genome via agrobacterium-mediated transformation.

Histochemical GUS staining detected promoter activity in the rhizoids present in digitate rays of archegoniophores (Fig. 39A). I also observed strong activity in the stalk

of archegoniophore, where detailed analysis revealed expression within the pegged rhizoids present inside the stalk. Additionally, the expression could be observed in the neck of some archegonia (indicated by red arrow) (Fig. 39A). I also observed promoter activity in the ventral scales of female vegetative thallus (Fig. 39B), antheridiophores (Fig. 39C) and male vegetative thallus (Fig. 39D).

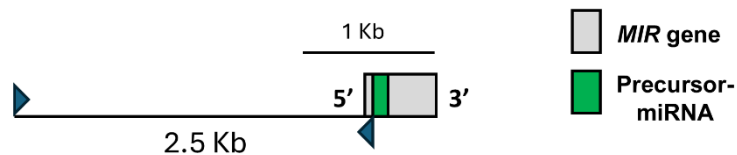
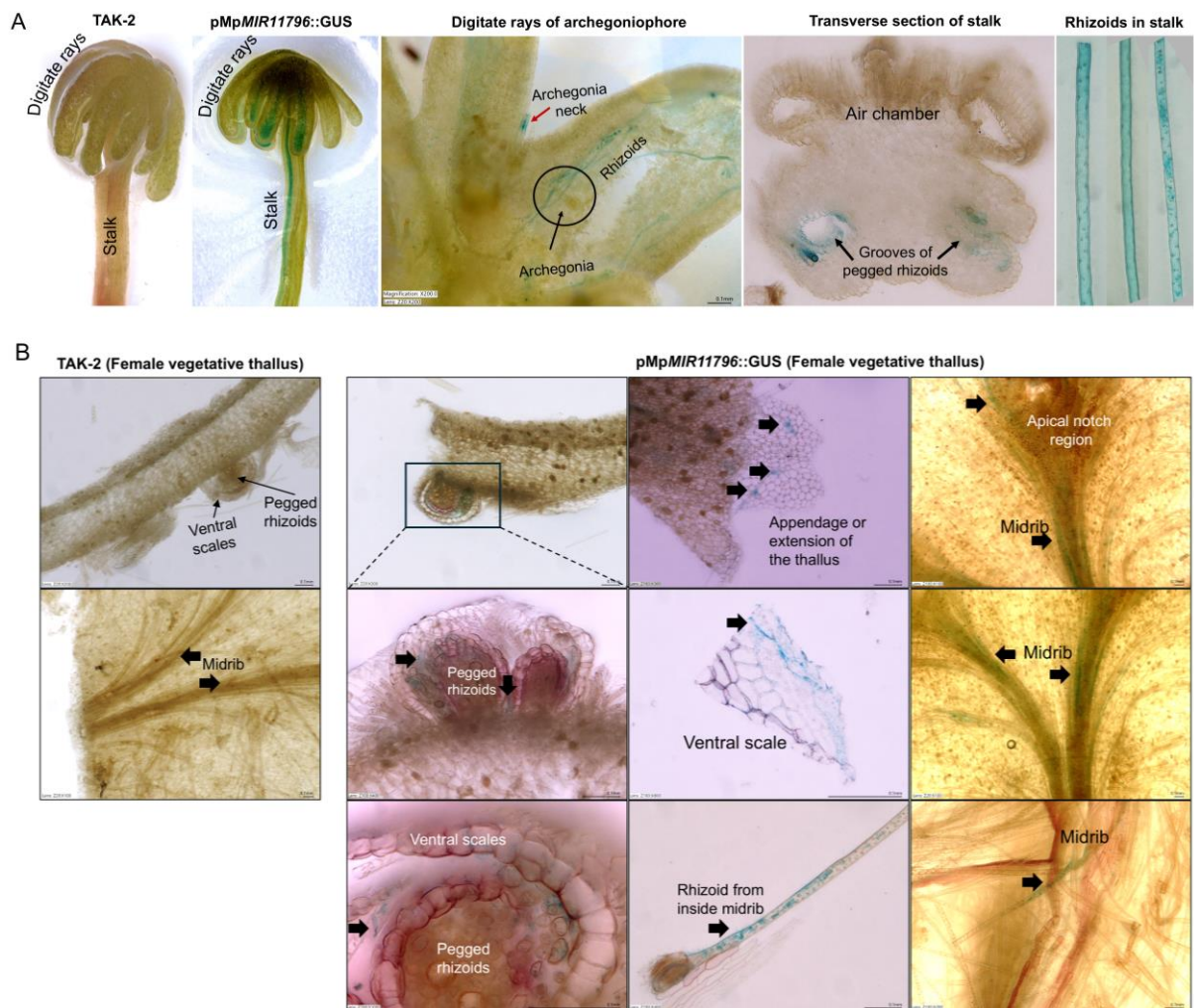


Figure 38. Schematic diagram showing the selection of promoter region used to drive the expression of GUS under Mp*MIR11796* gene promoter.



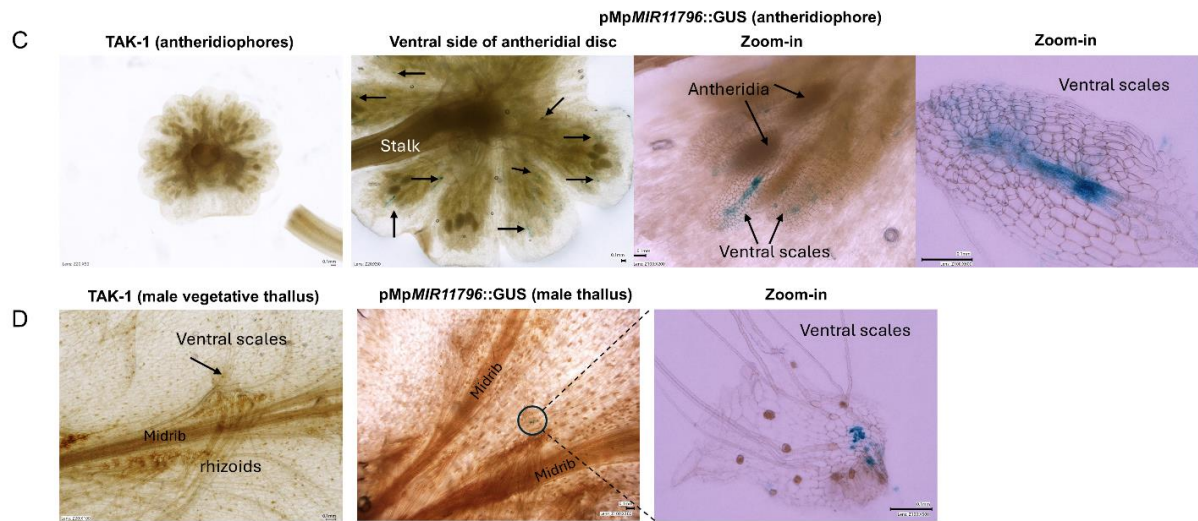


Figure 39. Spatiotemporal profile of *MpMIR11796* gene promoter activity. (A) From left to right: GUS-unstained WT mature archegoniophore; Mature archegoniophore from *pMpMIR11796::GUS* transformants stained with GUS showing the promoter activity in digitate rays and stalk; Digitate rays of archegoniophore showing the network of rhizoids inside, which sandwich archegonia (marked in circle); GUS staining observed in the neck of archegonia (shown with red arrow); Transverse section of the stalk of archegoniophore showing GUS staining in the grooves of pegged rhizoids; Dissected pegged rhizoids from the stalk of archegoniophore. (B) Different cross-sections of female vegetative thallus from GUS-unstained WT and *pMpMIR11796::GUS* transformants. (C) and (D) Microscopic images and zoom-in sections of antheridiophore and male vegetative thallus from GUS-unstained WT and *pMpMIR11796::GUS* transformants, respectively.

5.3.5 Generation of CRISPR/Cas9 KO plants for MpMIR11796

To elucidate the possible biological function(s) of MpmiR11796 in *Marchantia*, I wanted to observe the phenotypic changes after knocking down its production. To achieve this, two sgRNAs (sgRNA1 and sgRNA2) were designed using CRISPRdirect tool targeting the miRNA region (Fig. 40). The vectors for CRISPR/Cas9 genome editing were constructed to introduce in-del mutations at the respective MpMIR11796 locus and introduced into the spores via *Agrobacterium*-mediated transformation protocol (as described in section 15 of Chapter 4). The T1 and G1 plants were selected on hygromycin-containing media. The targeted genomic sequence was amplified in G1 as well as G2 plants and analyzed by sanger sequencing.

I obtained many male and female mutant plants with point mutations within the pre-MpmiRNA11796 region and I selected the plants that acquired significant insertions/deletions/ substitutions affecting the secondary structure of pre-miRNA11796. $\Delta mpmir11887^{#29}$ mutant was established using sgRNA1, whereas $\Delta mpmir11887^{#12,102,25,26,36}$ mutants were established using sgRNA2. I confirmed the location of mutations in these plants (4 male and 2 female mutant plants) (Fig. 41, Table 5).

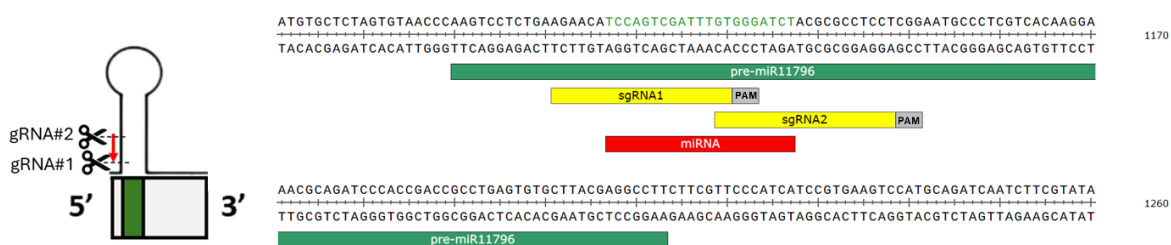


Figure 40. Schematic representation of the positions of sgRNA1 and sgRNA2 (in yellow), PAM sequence (highlighted in grey in the image on the right) targeting the pre-MpmiR11796 (in green) and mature MpmiR11796 region (highlighted in red in the image on the right); sgRNA1 overlaps 17nt out of 21nt miRNA sequence; sgRNA2 overlaps 9nt out of 21nt miRNA sequence. Below the stem-loop structure on the left side, MpMIR11796 gene structure is shown.

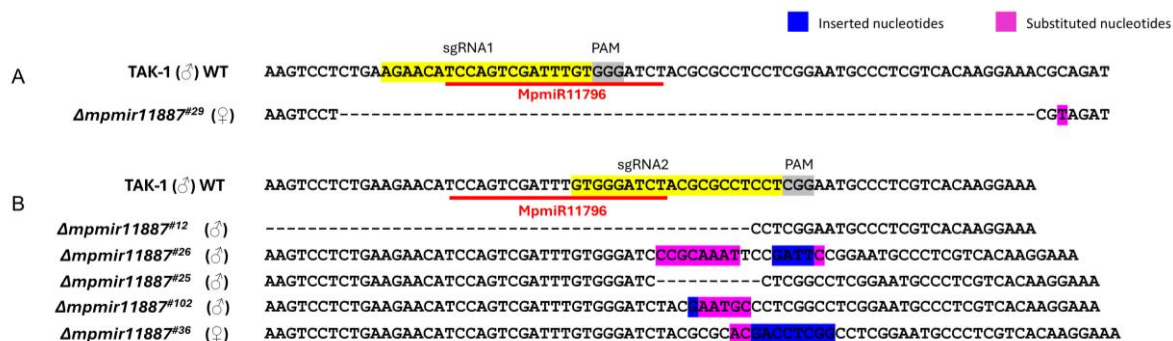


Figure 41. Schematic representation of the sequence editing in the Δ mpmir11796^{ge} mutant plants that led to mutations within mature miRNA region or pre-miRNA region.

Table 5. List of CRISPR/Cas9 genome-edited mutant plants for MpMIR11796 gene with their sex and genotypes.

| Sr. No. | Mutant plant | Sex | Mutation type |
|---------|-------------------------------------|-----|---|
| 1. | Δ mpmir11796 ^{#12} | M | Deletion of 46 nt in pre-miRNA region leading to deletion of entire mature miRNA region |
| 2. | Δ mpmir11796 ^{#102} | M | Insertion of 1 nt, substitution of 5 nt in pre-miRNA region |
| 3. | Δ mpmir11796 ^{#25} | M | Deletion of 10 nt from the pre-miRNA region encompassing 3 nt deletion from mature miRNA region |
| 4. | Δ mpmir11796 ^{#26} | M | Insertion of 4 nt, substitution of 9 nt in pre-miRNA region encompassing 1 nt substitution within mature miRNA region |
| 5. | Δ mpmir11796 ^{#29} | F | Deletion of 66 nt, substitution of 1 nt in pre-miRNA region leading to deletion of entire mature miRNA region |
| 6. | Δ mpmir11796 ^{#36} | F | Insertion of 8 nt, substitution of 2 nt within pre-miRNA region |

I further confirmed the complete loss of MpmiR11796 in $\Delta\text{mpmir11796}^{\#12}$ (♂), $\Delta\text{mpmir11796}^{\#26}$ (♂), $\Delta\text{mpmir11796}^{\#29}$ (♀), and $\Delta\text{mpmir11796}^{\#50}$ (♀) mutant plants using northern blot hybridization (Fig. 42). In all $\Delta\text{mpmir11796}^{ge}$ mutant plants, the structure of pre-MpmiR11796 showed conformational changes (Fig. 43). These mutant plants were successfully established as KO plants and were further used for phenotypic analysis.

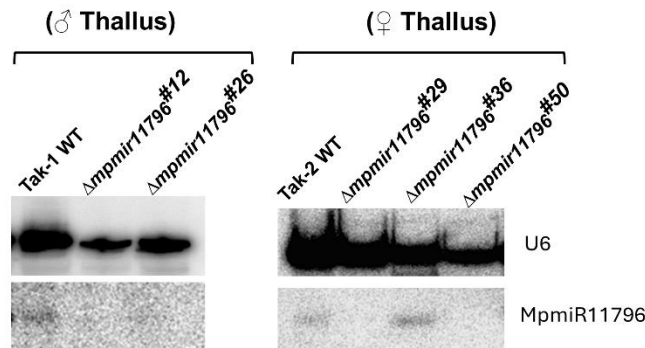


Figure 42. Northern blot analysis of $\Delta\text{mpmir11796}^{ge}$ mutant plants. U6 was used as RNA loading control.

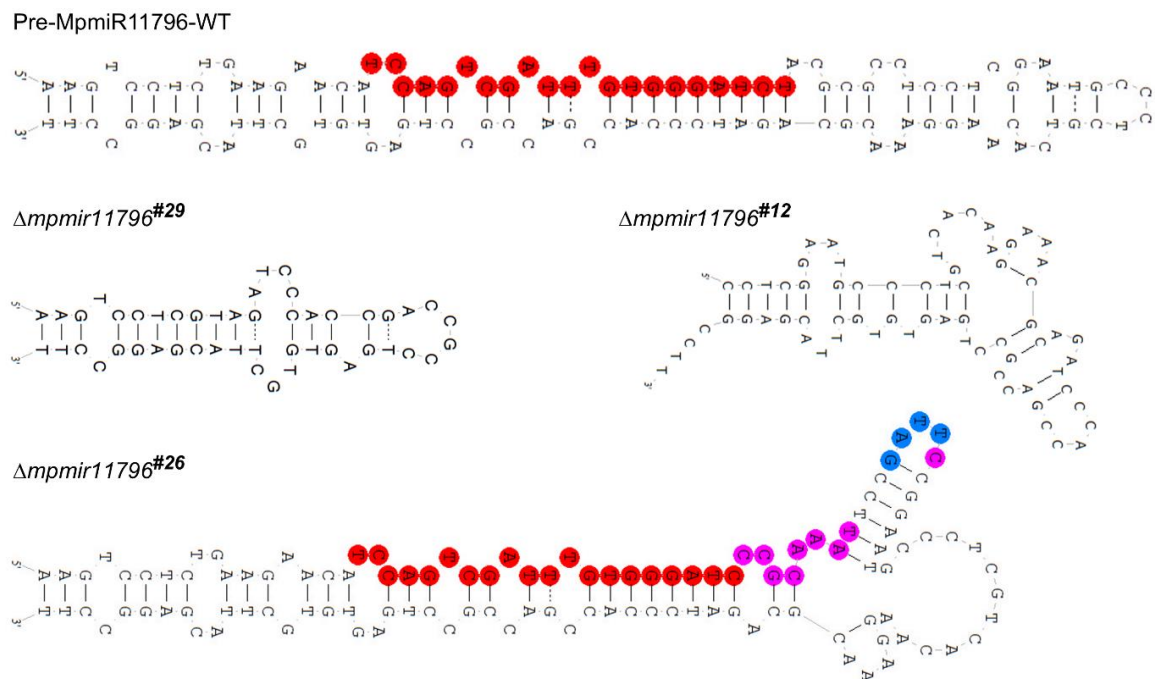


Figure 43. Secondary structures of the pre-MpmiR11796 in WT and $\Delta\text{mpmir11796}^{ge}$ KO plants as predicted by RNA Fold (Folder ver. 1.11); miRNA region in WT precursor

and remaining miRNA region after mutation in $\Delta mpmir11887^{#26}$ is highlighted in red, substitutions are highlighted in pink and insertions are highlighted in blue.

5.3.6 Phenotypic analysis of $\Delta mpmir11796^{ko}$ mutant plants

Under experimental conditions, I observed the vegetative growth of gemmae obtained from WT Tak-2 plants and $\Delta mpmir11796^{#29}$ KO plants. The KO plants showed retarded growth and smaller thallus area was observed as compared to WT (Fig. 44). However, I noticed that the mutants – though able to produce gemma cups – developed significantly reduced number of gemma cups as compared to WT (Fig. 44). Generally, the ventral scales cover the bundles of pegged rhizoids on the ventral side of WT thallus, but the structure and layering of ventral scales is disrupted in $\Delta mpmir11796^{#29}$ KO plants (Fig. 45). These observations suggest that MpmiR11796 is involved in the development of various structures of female vegetative thallus in Marchantia.

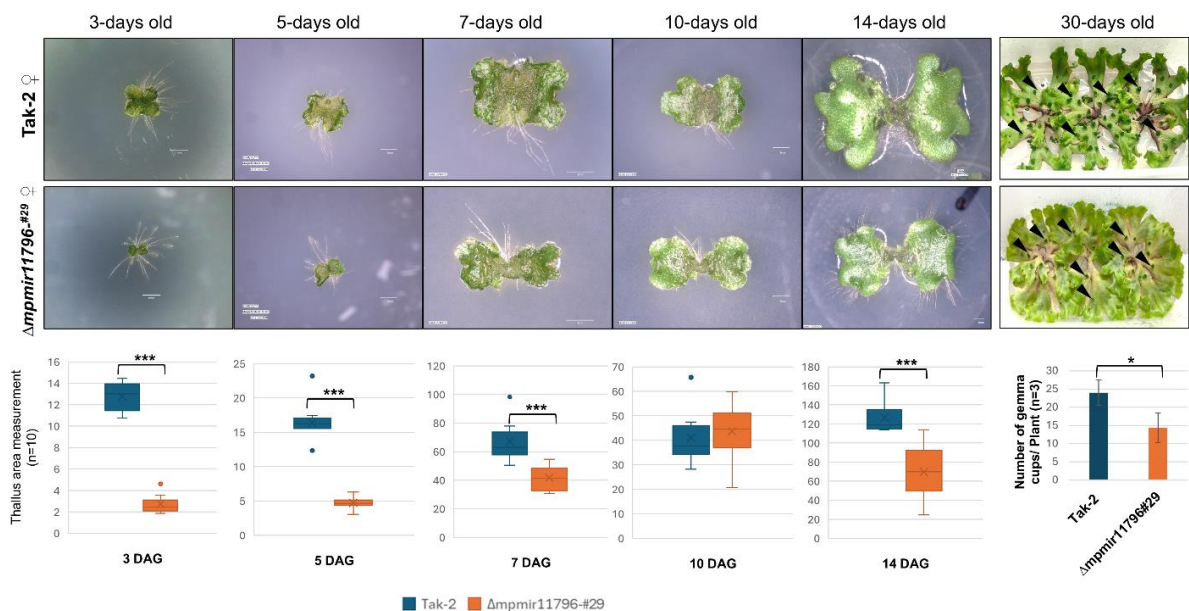


Figure 44. Dorsal view of growing gemma from WT Tak-2 plants and $\Delta mpmir11796^{#29}$ KO plants (upper two panels); growth interval points are mentioned on top i.e. 3-14 days old; 30-day-old thalli from WT and $\Delta mpmir11796^{#29}$ KO plants are shown on the image on extreme right. The arrowheads on the images point towards gemma cups. From left to right (bottom panel): Box plots showing the thallus area differences between WT and $\Delta mpmir11796^{#29}$ KO plants. The thallus area was measured from 10

individual gemmae using ImageJ software and statistics were performed using student's t-test (*p-value < 0.05; ***p-value < 0.001); Bar graph on the right shows the significant differences in the number of gemma cups produced by WT and $\Delta mpmir11796^{#29}$ KO plants. The statistics were performed using student's t-test (*p-value < 0.05).

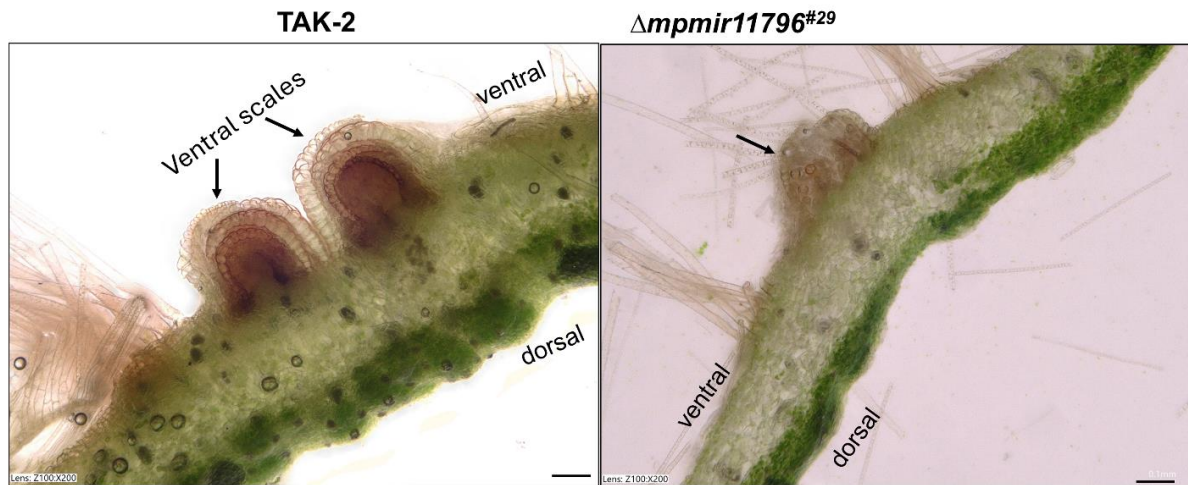


Figure 45. Microscopic images of transverse sections of vegetative thallus in WT and $\Delta mpmir11796^{#29}$ KO plants.

Next, I wanted to check whether MpmiR11796 KO affects archegonia/archegoniophore development. The $\Delta mpmir11796^{#29}$ KO plants developed archegoniophores under FR illumination as early as WT plants. However, upon examining in detail, the archegoniophore development was affected at macroscopic and microscopic levels. The archegonial receptacles (AR; or archegonial disk or carpocephalum) were significantly smaller in $\Delta mpmir11796^{#29}$ KO plants (Fig. 46) and often produce an extra small arm on the AR in approximately 29 % (n = 23/80) of archegoniophores of $\Delta mpmir11796^{#29}$ KO plants (Fig. 47). Moreover, I also observed that $\Delta mpmir11796^{#29}$ KO archegoniophores fail to produce rhizoids and ventral scales unlike WT archegoniophores in *Marchantia* (Fig. 48, 49).

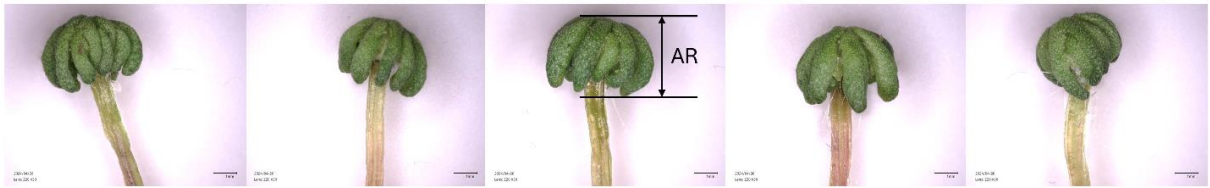
Moreover, the stalks of archegoniophore in $\Delta mpmir11796^{#29}$ KO plants did not undergo elongation even though WT archegoniophores started to mature and developed elongated brown stalks (Fig. 50A, 50B). The stalks of another KO line i.e. $\Delta mpmir11796^{#50}$ also appeared smaller compared to WT plants (Fig. 50B), however,

Δmpmir11796^{#29} KO plants were analyzed more deeply. The stalk length and diameter were significantly decreased in *Δmpmir11796^{#29}* KO plants (Fig. 50C). The stalk of the archegoniophore has two longitudinal grooves containing bundles of pegged rhizoids, however, the archegoniophore stalk in the mutant *Δmpmir11796^{#29}* had significantly lesser area, which also indicates reduction in the size of two grooves of rhizoids and simultaneously lesser bundles of pegged rhizoids comparative to WT (Fig. 51).

Next, I also observed various abnormalities during gemmae development in male *Δmpmir11796^{ko}* plants i.e. *Δmpmir11796^{#12}*, *Δmpmir11796^{#26}*. Both KO plants produced a significantly reduced number of smooth rhizoids during gemmae development (Fig. 52A, 52B). The experiment was replicated twice. In the first repetition, gemmae growth was monitored from day 3 to day 17, during which both KO plants exhibited reduced number of smooth rhizoids as compared to WT plants (Fig. 52). In the second repetition, the gemma growth was observed until day 21, where by this stage, the KO plants had developed rhizoids in numbers comparable to WT (Fig. 53). Notably, the *Δmpmir11796^{ko}* plants produced antheridiophores that showed no distinguishable morphological differences from those of WT antheridiophores (Fig. 52C).

Given that MpMIR11796 gene promoter activity was observed in the rhizoids of digitate rays and the rhizoids inside the archegoniophore stalk, together with phenotypes observed in *Δmpmir11796^{ko}* plants i.e. reduced size of the archegonial heads, shorter archegoniophore stalks, improper ventral scales and almost negligible production of rhizoids in archegoniophores, it is indicated that MpmiR11796 is required for proper archegoniophore growth and development. This also suggests that MpmiR11796 may regulate genes involved in polarized cell growth, cell proliferation, cell elongation, cell differentiation.

TAK-2



$\Delta mpmir11796^{#29}$

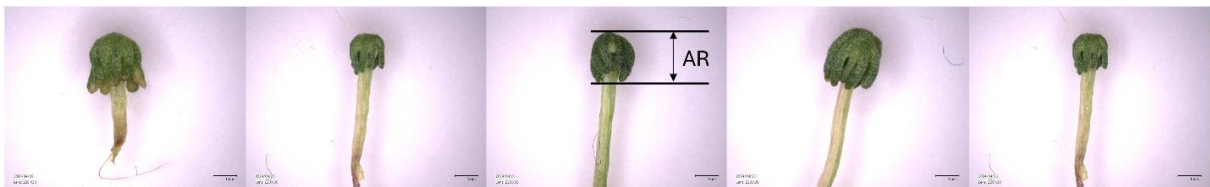


Figure 46. Microscopic images showing the side view of mature archegoniophores from WT and $\Delta mpmir11796^{#29}$ KO plants. AR: Archegonial Receptacle.

TAK-2



$\Delta mpmir11796^{#29}$



Figure 47. Microscopic images of mature archegoniophores from WT and $\Delta mpmir11796^{#29}$ KO plants. The black arrowheads show the extra small arm produced on the archegonial receptacle.



Figure 48. Microscopic images of cross sections of mature archegoniophores from WT and $\Delta mpmir11796^{#29}$ KO plants. The white arrow shows the rhizoids in the WT archegoniophore and Δ shows absence of rhizoids in archegoniophores produced by $\Delta mpmir11796^{#29}$ KO plants.

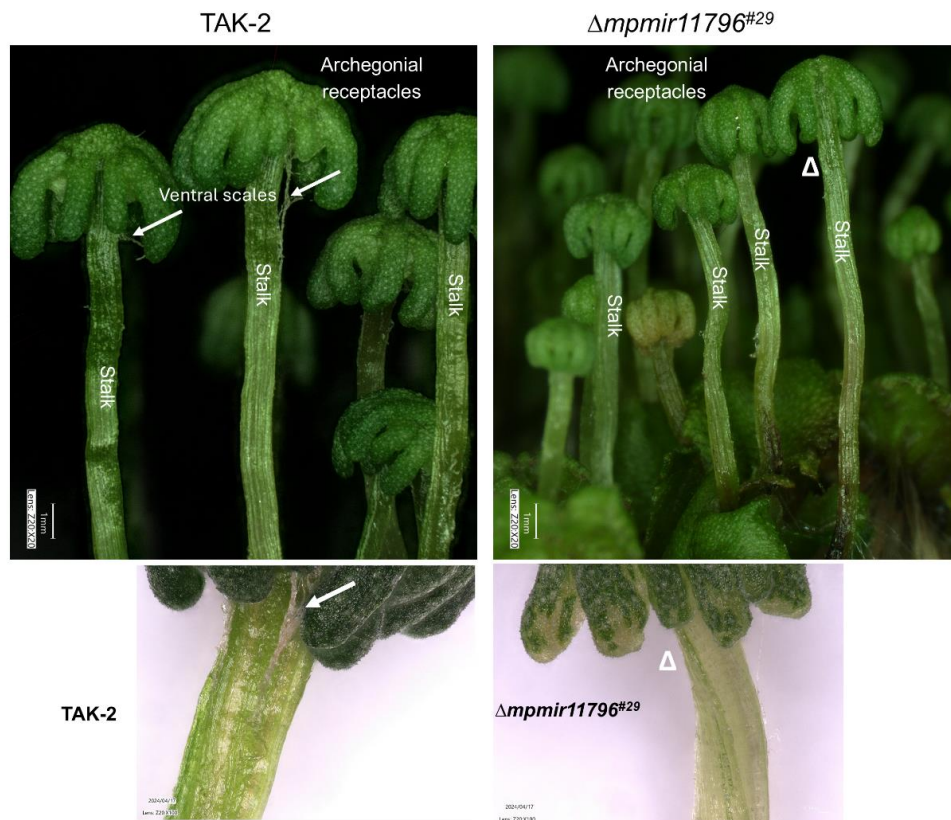


Figure 49. Microscopic images of mature archegoniophores from WT and $\Delta mpmir11796^{#29}$ KO plants. The white arrows point to the ventral scales in the

archegoniophores of WT plants. Δ shows absence of ventral scales in the archegoniophores produced by $\Delta mpmir11796^{#29}$ KO plants.

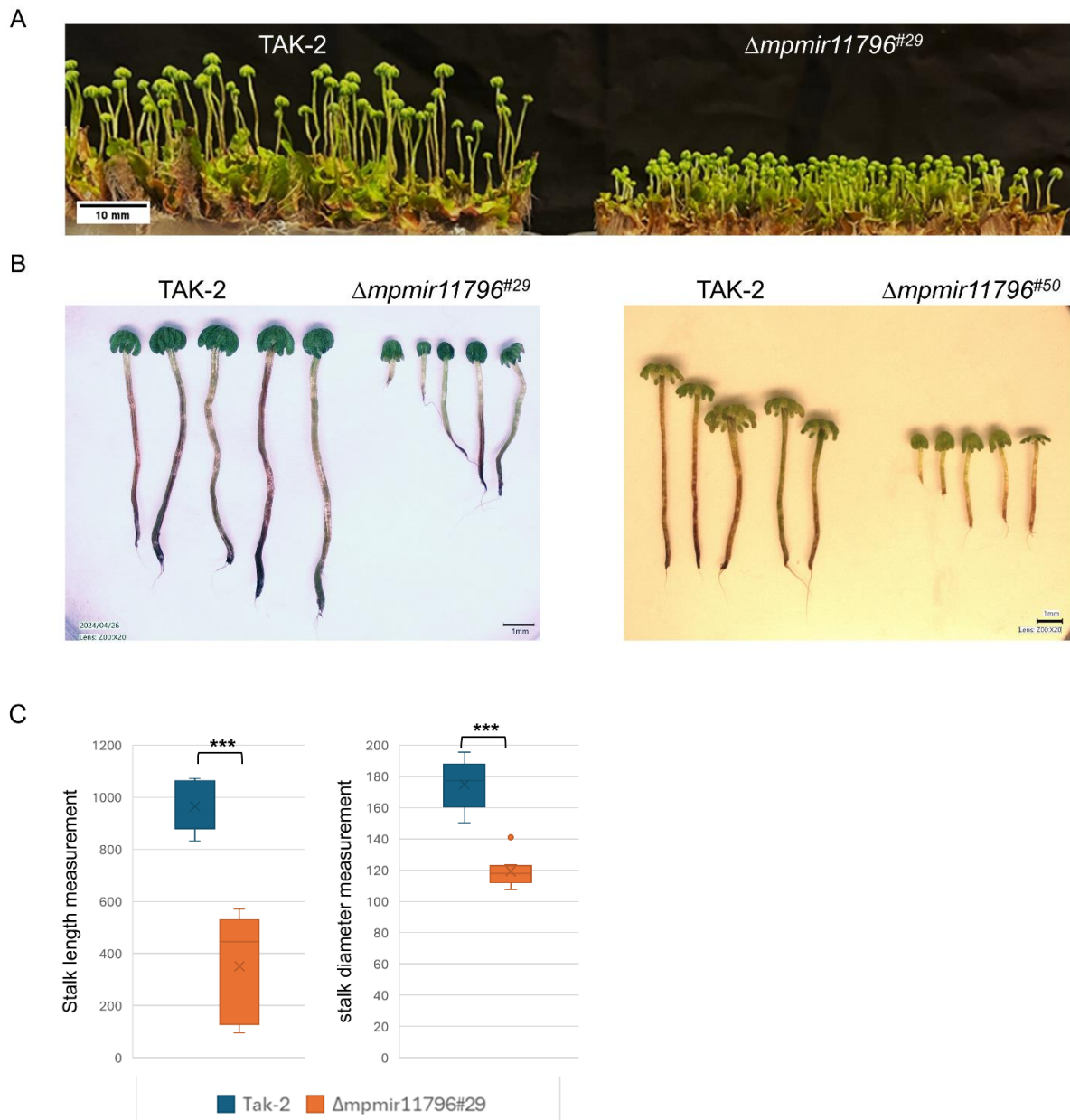
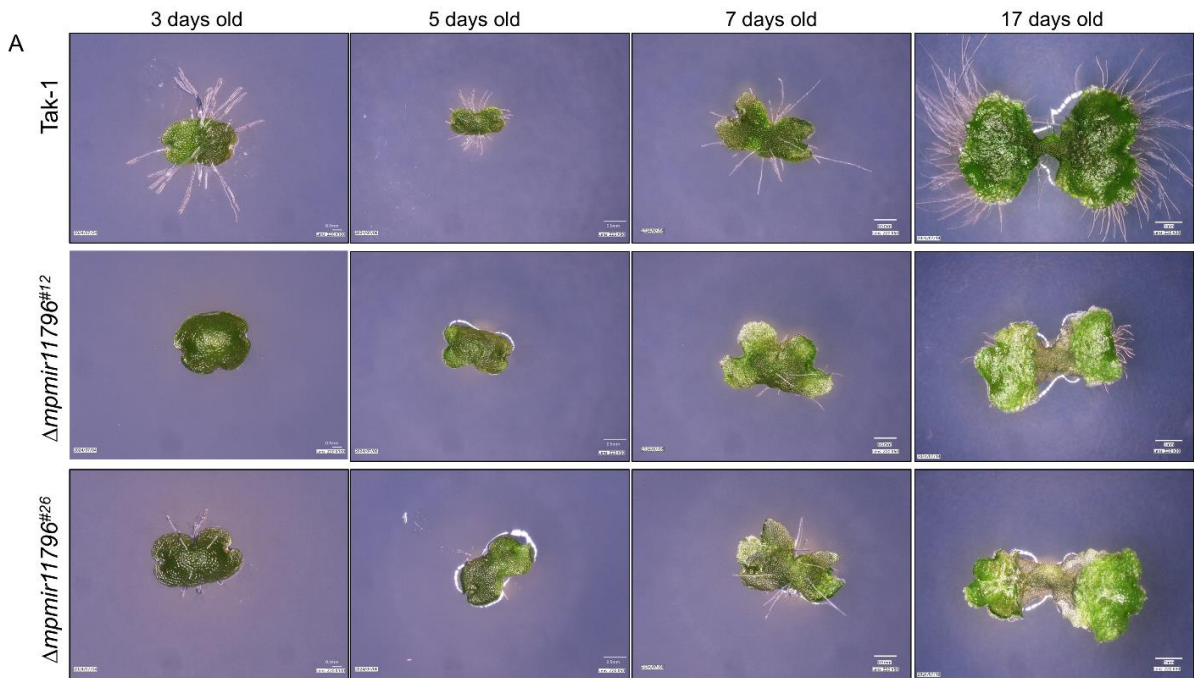


Figure 50. (A) Overall morphology of archegoniophores of 1.5-month-old WT and $\Delta mpmir11796^{#29}$ KO plants. (B) Microscopic images of archegoniophores from WT, $\Delta mpmir11796^{#29}$, and $\Delta mpmir11796^{#50}$ KO plants. The browning of stalk indicates mature archegoniophores. (C) Box-plot showing the quantification of the length and

diameter of the stalk of archegoniophore in WT and $\Delta mpmir11796^{#29}$ KO plants. (n = 10; ***p-value < 0.001).



Figure 51. Transverse section of the stalk of archegoniophore in WT and $\Delta mpmir11796^{#29}$ KO plants. The box plot on the right shows the quantification of area of the stalk of archegoniophore in WT and $\Delta mpmir11796^{#29}$ KO plants (n = 13; ***p-value < 0.001).



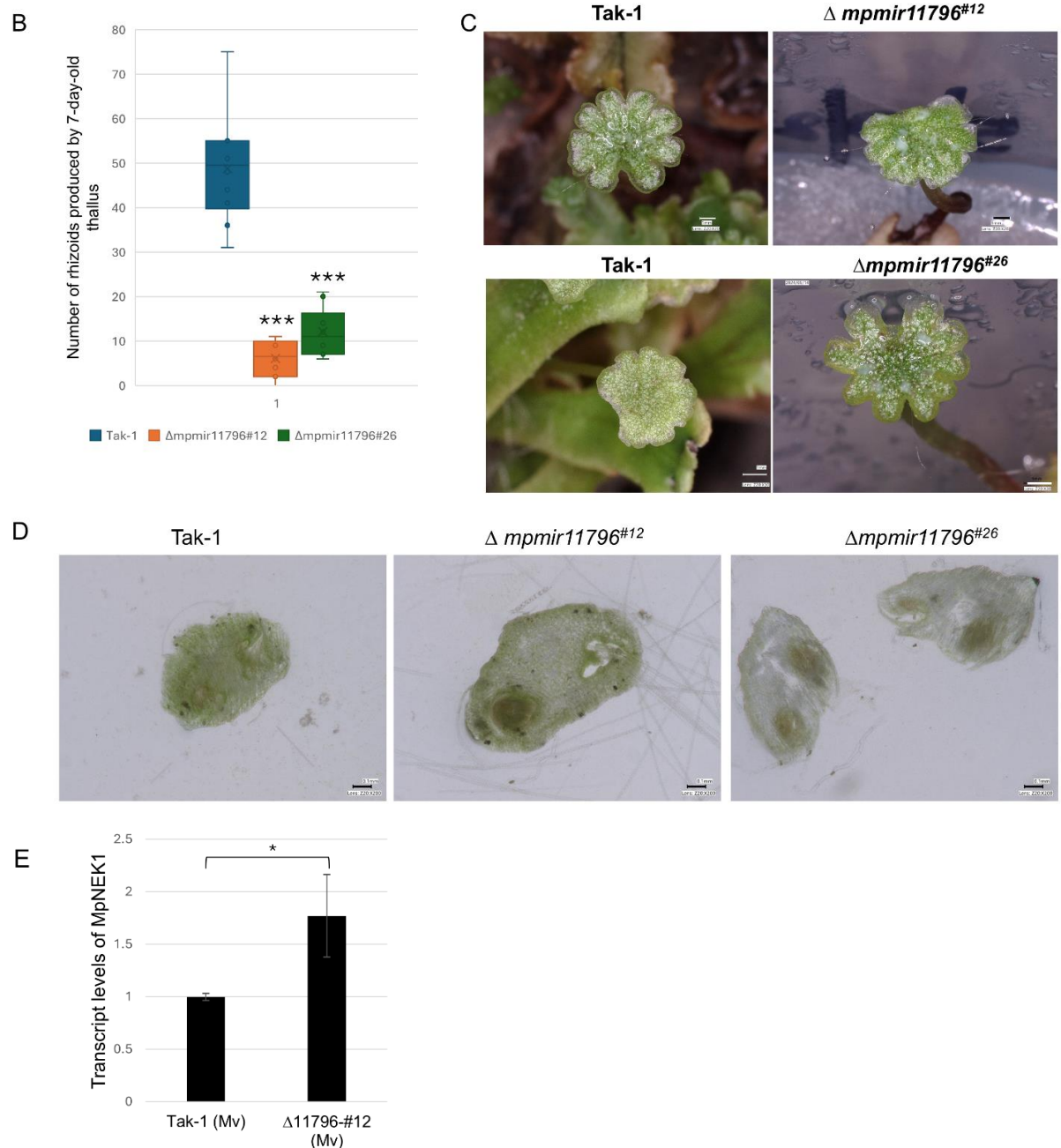


Figure 52. (A) Dorsal view of growing gemmae from WT and $\Delta mpmir11796^{ko}$ male mutant plants. The growth interval points are mentioned on top i.e. 3-17 days old. (B) Box plot showing the quantification of the number of rhizoids produced by 7-day-old WT and $\Delta mpmir11796^{ko}$ male plants ($n = 10$; *** p -value < 0.001). (C) Microscopic images of mature antheridiophores produced by WT and $\Delta mpmir11796^{ko}$ male mutant plants. (D) Transverse section of the stalk of antheridiophore in WT, $\Delta mpmir11796\#12$ and $\Delta mpmir11796\#26$ KO plants. (E) RT-qPCR expression levels of MpNEK1 transcript

in WT and $\Delta mpmir11796^{#12}$ KO plants, normalized by MpACT7 (n = 3); asterisk indicates significant difference (t-test, p-value < 0.05).

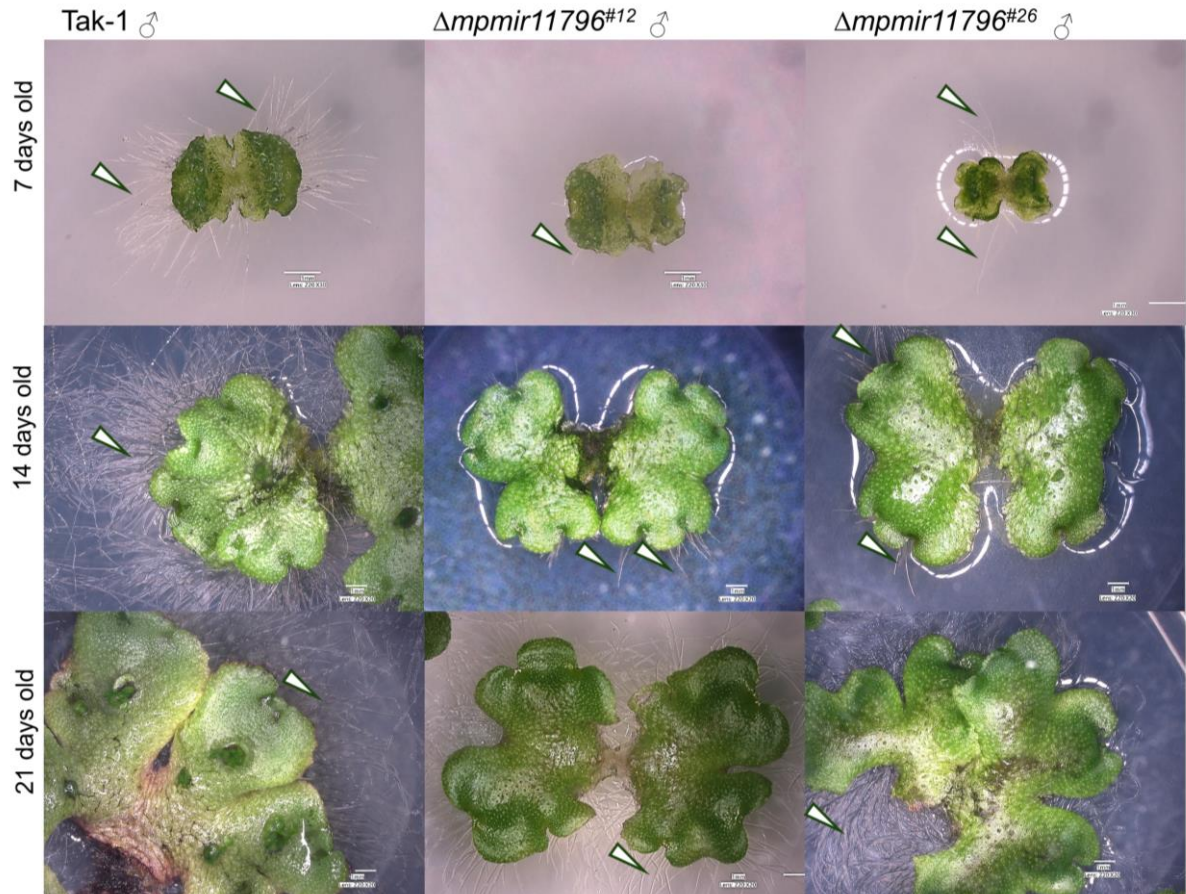


Figure 53. Dorsal view of growing gemmae from WT and $\Delta mpmir11796^{ko}$ male mutant plants. The growth interval points are mentioned on left i.e. 7-21 days old. The arrowheads point to the rhizoids produced by male WT Tak-1 and male $\Delta mpmir11796^{ko}$ plants.

Next, I examined the structure of archegonia in WT and $\Delta mpmir11796^{#29}$ KO plants. During the archegonia maturation in *M. polymorpha*, the young archegonium with a primary central cell (PCC) undergoes asymmetrical divisions to generate a secondary central cell (SCC) and a primary neck canal (PNC) cell. The SCC enlarges and divides asymmetrically to produce 2 daughter cells, a smaller ventral canal cell (VCC) and a large cell that ultimately matures into an egg cell (Fig. 7). I observed several

aberrations in the archegonial cells produced by $\Delta mpmir11796^{#29}$ KO plants. Unlike the round shape of EC in WT archegonia, 8.8 % (n = 62/700) of the archegonia in $\Delta mpmir11796^{#29}$ KO plants had an unusual shape. The ventral canal cell (VCC) in the venter of WT archegonia gradually disorganizes into mucilaginous strands during the development, however, 14.7 % (n = 103/700) of the archegonia in $\Delta mpmir11796^{#29}$ KO plants showed unusual masses possibly due to unusual division between the VCC and EC. Various archegonia with mucilaginous masses or blobs on the EC were also observed. Possibly due to some disturbances in these asymmetrical divisions, I observed 1.4 % (n = 10/700) of the archegonia with double egg cells in the venter while a single egg cell is generally present in each WT archegonia venter (Fig. 54). Collectively, these results demonstrate that MpmiR11796 is involved in the asymmetrical divisions during egg cell development in Marchantia.

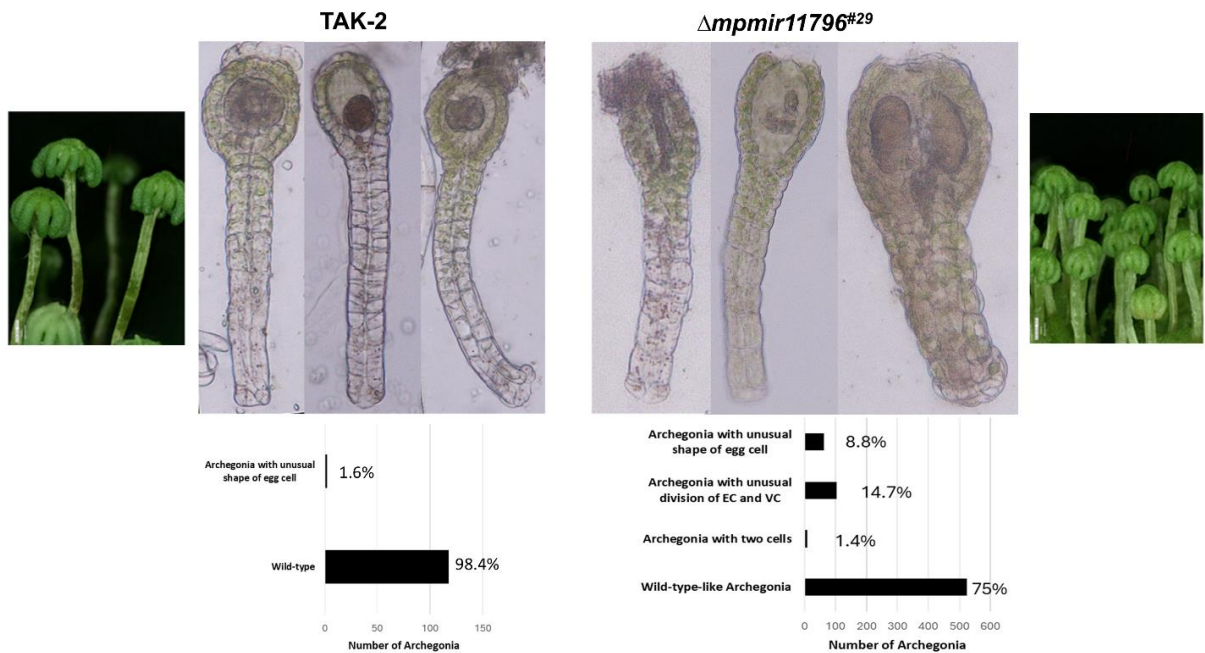


Figure 54. Microscopic images showing morphology of isolated archegonia from female WT Tak-2 plants and $\Delta mpmir11796^{#29}$ KO plants. EC, egg cell; VCC, ventral canal cell. The graphs below the images shows the percentage of archegonia with WT like and/or abnormal features. For WT (n=5) and $\Delta mpmir11796^{#29}$ KO plants (n=18) archegoniophores were dissected.

To further test if the $\Delta mpmir11796^{#29}$ KO plants were able to undergo successive fertilization, I crossed them with WT sperm cells (Case I) and $\Delta mpmir11796^{#12}$ male mutant sperm cells (Case II). The fertilization efficiency was low in both cases as compared to WT cross (Table 6) with case I producing more adverse effects. The obtained sporophytes were cultured to observe the growth and development of next generation (Fig. 55). Delayed production of gametangiophores was observed in both cases but finally case II produced gametangiophores, but with abnormal gametes (Fig. 56). Case I also produced 1-2 antheridiophores and archegoniophores after 4 months. This implies a critical maternal role of MpmiR11796. It has to be stated that the results were obtained from one biological replicate and another biological replicate needs to be performed to draw firm conclusions.

Table 6. Summary of the crossing experiments performed.

| | Tak1 ♂ x Tak2 ♀ | Case I: Tak1 ♂ x $\Delta mpmir11796^{#29}$ ♀ | Case II: $\Delta mpmir11796^{#12}$ ♂ x $\Delta mpmir11796^{#29}$ ♀ |
|---|--|---|---|
| Number of archegoniophores used for fertilization | 18 | 15 | 20 |
| Number of archegoniophores with sporophyte obtained | 17 | 4 | 12 |
| Fertilization efficiency | 94 % | 27 % | 60 % |
| Sporophyte development | 4-5 gametangiophores produced after 1.5 months and continued | 1-2 gametangiophores produced after 4 months | 3-4 gametangiophores produced after 3 months, but malformed |

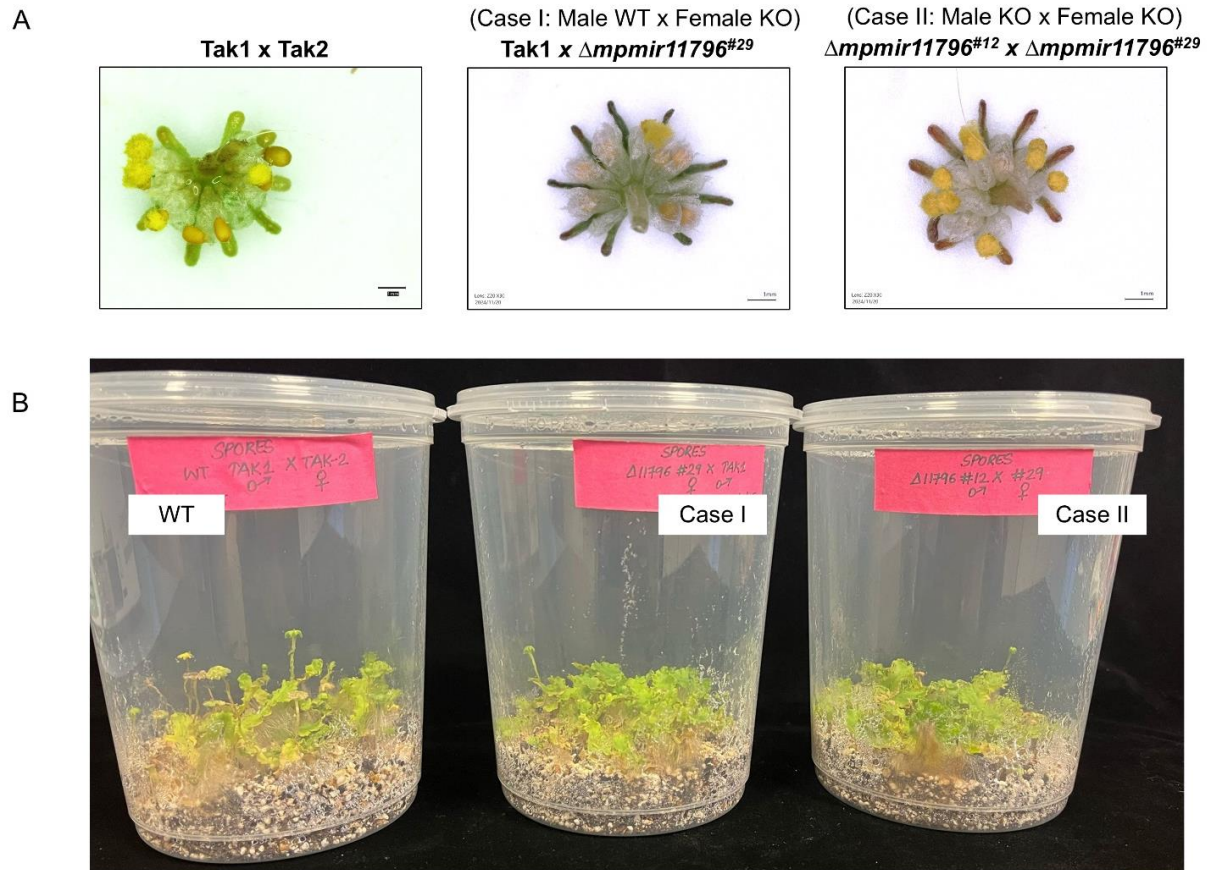


Figure 55. Sporophytes obtained after crossing experiments. (A) Microscopic images of sporophyte obtained after crosses. (B) Photograph of plants after germination of spores obtained from crossing between male WT Tak-1 and female WT Tak-2 plants, Case I (male WT Tak1 sperms used to fertilize female $\Delta mpmir11796^{#29}$ mutant plants) and Case II (male $\Delta mpmir11796^{#12}$ mutant sperms used to cross female $\Delta mpmir11796^{#29}$ mutant plants).

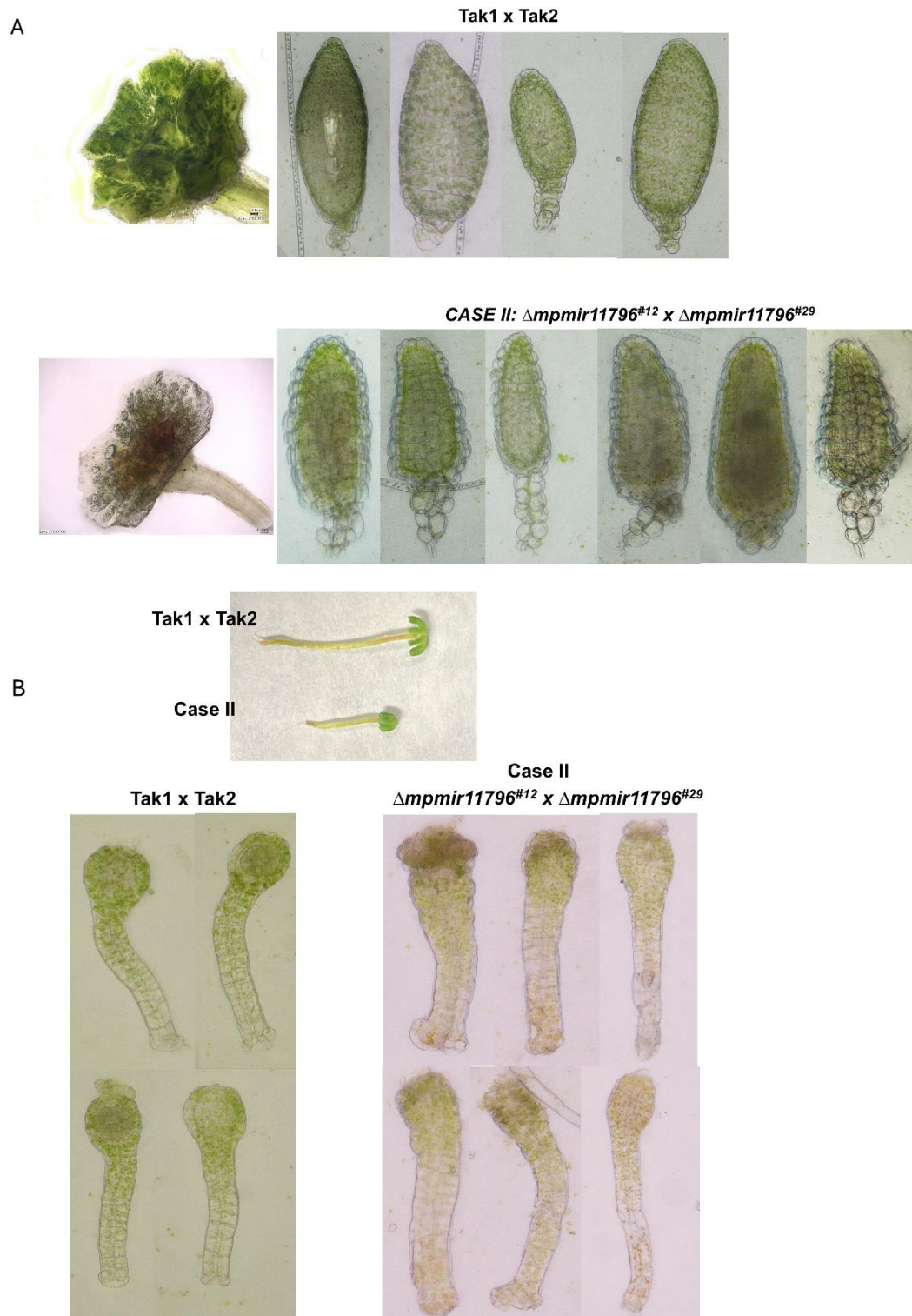


Figure 56. (A) Microscopic images of antheridiophores and antheridia obtained after sporophyte germination from WT and $\Delta mpmir11796^{ko}$ crossings. (B) Microscopic

images of archegoniophores and archegonia obtained after sporophyte germination from WT and $\Delta mpmir11796^{ko}$ crossings.

5.3.7 Identification of putative mRNA target of MpmiR11796 from transcriptomic and degradome data analyses

Degradome sequencing data identified Mp4g20750 mRNA as a potential target for MpmiR11796. Mp4g20750 gene encodes a protein belonging to linker histone H1 and H5 family. RT-qPCR analysis revealed exclusive expression of target mRNA in antheridiophores, showing reverse correlation with MpmiR11796 expression levels (Aggarwal et al., 2024). RACE assays detected a single cleavage product at the predicted slice site. Transcriptome analysis (from archegoniophores) showed an increase in the transcript levels of Mp4g20750 mRNA, but the fold change values were low and not statistically significant. Furthermore, RT-qPCR analysis showed downregulation of Mp4g20750 mRNA in male $\Delta mpmir11796^{ko}$ plants (graphs not shown in the thesis). This suggests that Mp4g20750 is probably not the true target of MpmiR11796.

Subsequently, Mp1g02530 gene was selected as a potential target for MpmiR11796 from degradome data as this gene encodes an AAA+ ATPase and was found to be identical to cell division cycle protein (CDC48A) in Arabidopsis. The putative slice site was identified within the coding sequence (Fig. 57A, 57B). Moreover, RT-qPCR analysis revealed that it is upregulated in $\Delta mpmir11796^{ko}$ plants (Fig. 57C). Despite repeated attempts, RACE assays were unsuccessful in detecting the cleavage site as all the products obtained were 13 bp upstream of predicted slice site. The transcriptomic analysis combined with degradome sequencing data identified some potential targets for MpmiR11796 (Table 7). RACE assays have to be performed to prove the proper target.

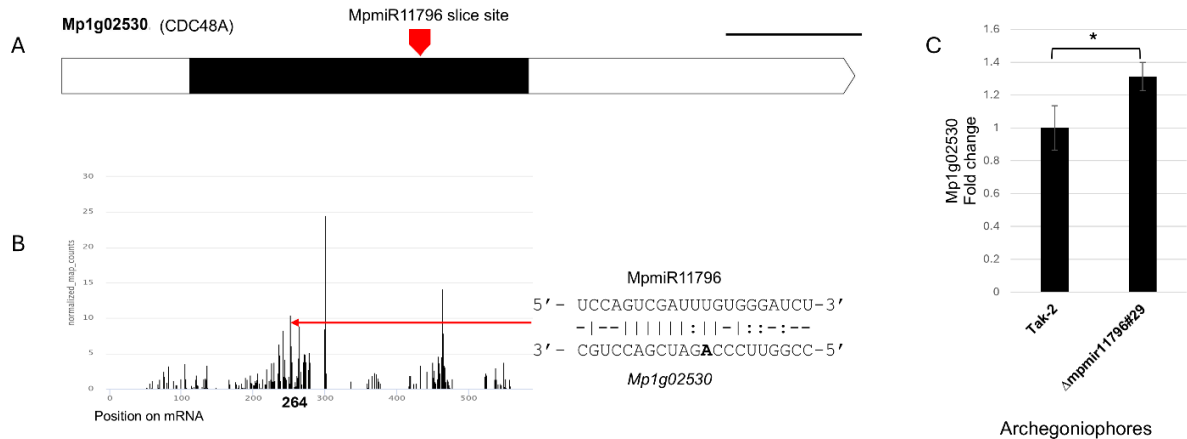
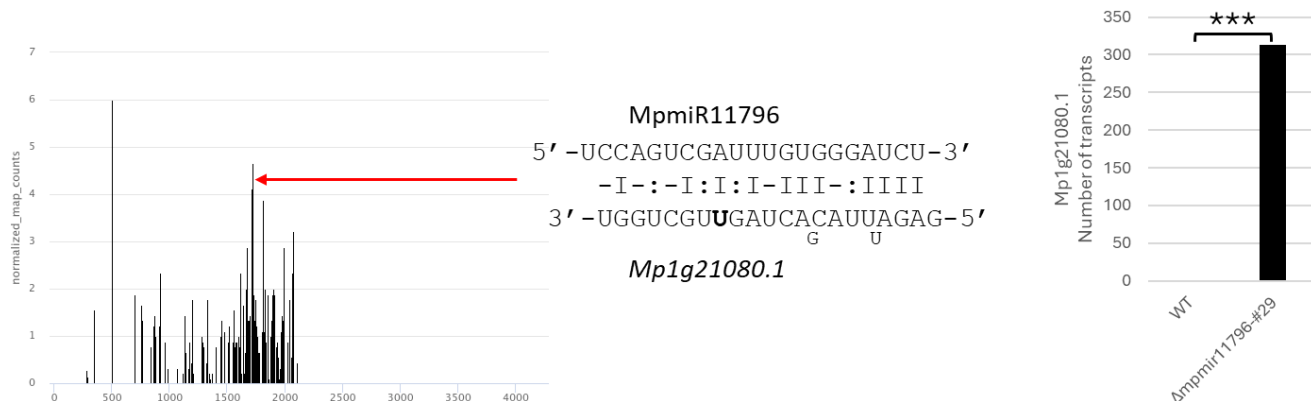


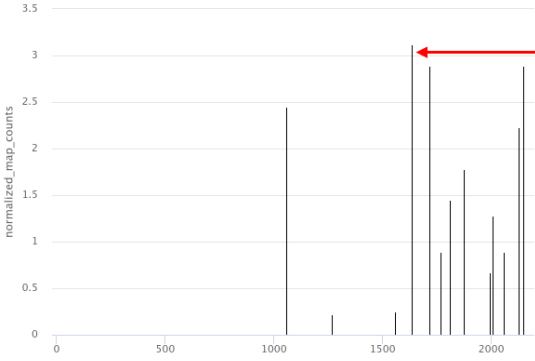
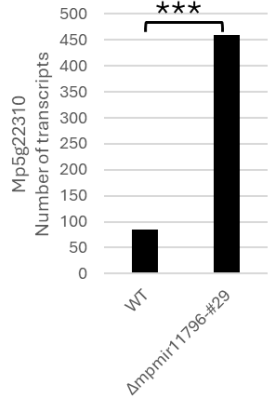
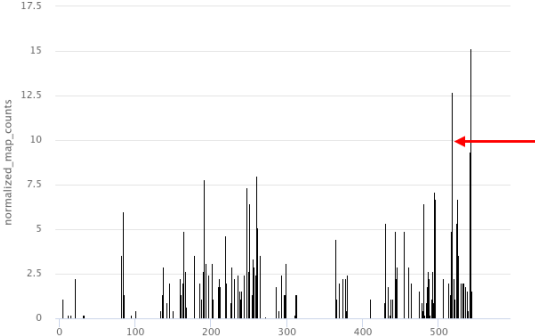
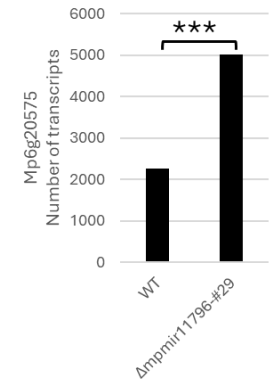
Figure 57. Putative target for MpmiR11796 identified from degradome data. (A) Structure of Mp 1g02530 gene; the putative slice site for miRNA is shown by downward pointing arrow. (B) Based on degradome data, the t-plot of corresponding putative target mRNA Mp 1g02530; the predicted cleavage site is pointed with red arrow; Alongside the t-plot, complementary base pairing between miRNA and target mRNA is presented; cleavage site is indicated by nucleotide marked in bold. (C) RT-qPCR expression level of Mp 1g02530 transcript; *p-value<0.05.


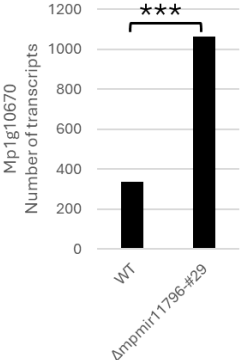
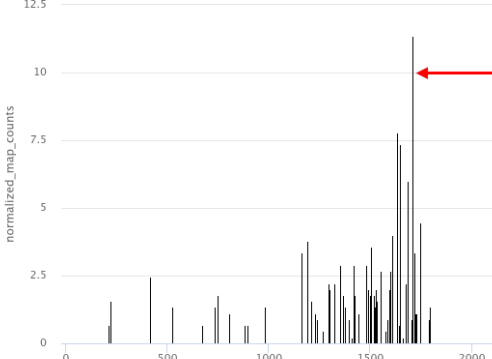
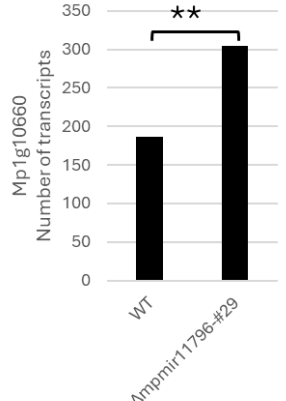
5.3.8 GO analyses

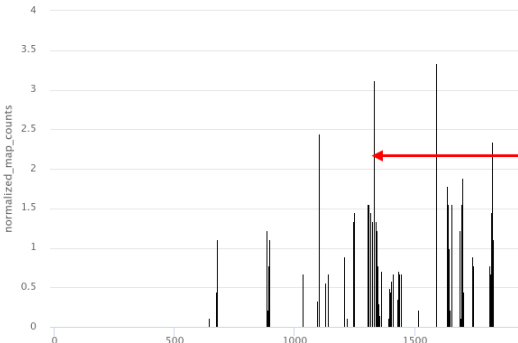
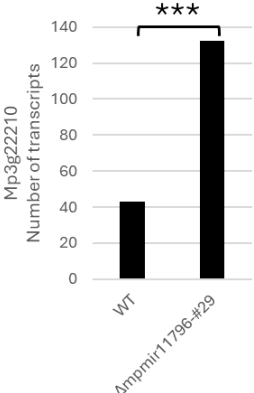
Differentially expressed genes (DEGs) from transcriptome analysis were subjected to Gene Ontology (GO) analysis (Performed by Prof. W. Karlowski). This led to the identification of two main categories of biological processes affected by the knock-out of MpmiR11796. These two categories include: ATP-binding (12 genes) and protein-binding (28 genes). The ATP-binding category included plasma membrane H⁺-transporting ATPases, protein tyrosine and serine/threonine kinases, ABC transporters, proteases belonging to AAA+ ATPase superfamily, and a DNA repair protein. The protein binding category includes ubiquitin-like proteins, apoptotic ATPases, potassium channel, Ankyrin repeat and DHHC-type Zn-finger domain containing proteins, pro-apoptotic proteases, F-box domain containing proteins, bromodomain containing protein, Leucine-rich repeat proteins, LRR receptor-like serine/threonine-protein kinases, WD-domain containing proteins etc. The details have been listed in Supplementary Table 9.

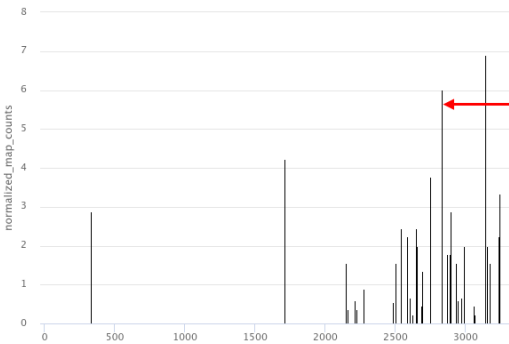
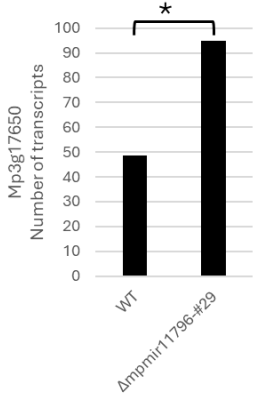
Table 7. Putative mRNA targets selected for MpmiR11796 as revealed by transcriptomic and degradome data analyses:

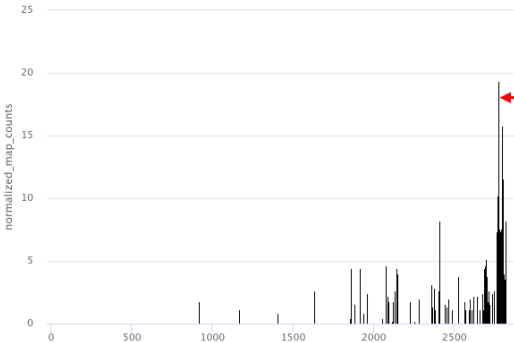
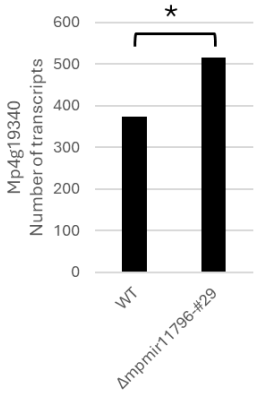
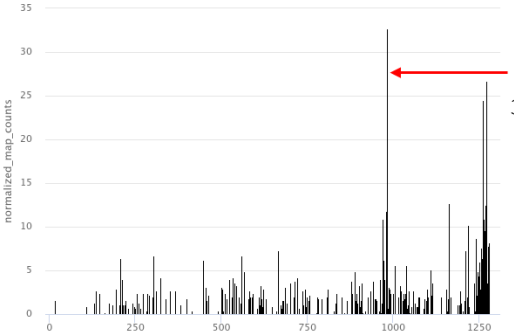
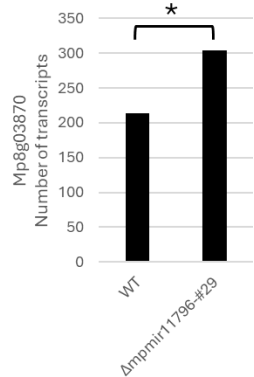
| Sr. No. | Target gene ID | mRNA length (nt) | mRNA:miRNA pairing region (nt) | Cleavage (nt of mRNA)+ | Cleavage (nt of mRNA)* | Location of putative slice site within mRNA |
|---------|--|------------------|--------------------------------|------------------------|------------------------|---|
| 1. | Mp1g21080.1 | 4243 | 1703:1725 | c:1718(4.48 2 42)+ | - | 3' - UTR |
| | <p>>Mp1g21080.1 [PANTHER:PTHR47978]; [SMART:SM00175] rab_sub_5; [SMART:SM00174] rho_sub_3; [ProSiteProfiles:PS51421] small GTPase Ras family profile.; [NCBIfam:TIGR00231] small GTP-binding protein domain; [CDD:cd04123] Rab21; [G3DSA:3.40.50.300]; [G3DSA:3.40.50.300:FF:000550] ras-related protein Rab-21; [SMART:SM00176] ran_sub_2; [MobiDBLite:mobidb-lite] consensus disorder prediction; [PRINTS:PR00449] Transforming protein P21 ras signature; [Pfam:PF00071] Ras family; [ProSiteProfiles:PS51419] small GTPase Rab1 family profile.; [SMART:SM00173] ras_sub_4; [SUPERFAMILY:SSF52540] P-loop containing nucleoside triphosphate hydrolases; [KEGG:K07890] Ras-related protein Rab-21; [KOG:KOG0088] GTPase Rab21, small G protein superfamily; [R]; [MapolyID:Mapoly0001s0443]; [GO:0003924] GTPase activity; [GO:0032482] Rab protein signal transduction; [GO:0005794] Golgi apparatus; [GO:0005525] GTP binding; [GO:0005768] endosome</p> <div>  <p>normalized_map_counts</p> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>-I-:-I:I:I-III-:IIII</p> <p>3' -UGGUCGUUGAUCACAUUAGAG-5'</p> <p>Mp1g21080.1</p> <p>Number of transcripts</p> <p>Mp1g21080.1</p> <p>WT</p> <p>ΔmpmiR11796-#29</p> <p>***</p> </div> | | | | | |
| 2. | Mp5g22310 | 2176 | 1623:1637 | c:1632(2.16 1 14)+ | - | Coding sequence |
| | <p>>Mp5g22310.1 [PANTHER:PTHR31479] ALPHA/BETA-HYDROLASES SUPERFAMILY PROTEIN; [SUPERFAMILY:SSF53474] alpha/beta-Hydrolases; [Coils:Coil] Coil; [G3DSA:3.40.50.1820] alpha/beta hydrolase; [MapolyID:Mapoly0166s0025]</p> | | | | | |

| | | | | | | |
|----|---|------|-----------|-------------------|----------------------|----------|
| | <div>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>-I--III-III:III::-I-I</p> <p>3' -CUGCCCAGGUAAAGCACUUCACAA-5'</p> <p><i>Mp5g22310</i></p> </div>  </div> | | | | | |
| 3. | Mp6g20575 | 588 | 504:527 | c:517(4.99 2 57)+ | - | 3' - UTR |
| | <div> <p>>Mp6g20575.1 [Pfam:PF01439] Metallothionein; [GO:0046872] metal ion binding</p>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>: -I-I--IIIII:I::--</p> <p>3' -GAGGCAACUAAAUAUUUAGAU-5'</p> <p>AU A</p> <p><i>Mp6g20575</i></p> </div>  </div> | | | | | |
| 4. | Mp1g10670 | 3250 | 3083:3105 | - | c:3098(16.50 1 261)* | 3' - UTR |
| | <p>>Mp1g10670.1 [Pfam:PF17808] Fn3-like domain from Purple Acid Phosphatase; [PANTHER:PTHR45778] PURPLE ACID PHOSPHATASE-RELATED; [SUPERFAMILY:SSF49363] Purple acid phosphatase, N-terminal domain; [CDD:cd00839] MPP_PAPs; [G3DSA:2.60.40.380]; [Pfam:PF14008] Iron/zinc purple acid phosphatase-like protein C; [G3DSA:3.60.21.10]; [SUPERFAMILY:SSF56300] Metallo-dependent phosphatases; [Pfam:PF16656] Purple acid Phosphatase, N-terminal domain; [Pfam:PF00149] Calcineurin-like phosphoesterase; [KEGG:K22390] acid phosphatase type 7; [KOG:KOG1378] Purple acid phosphatase; [G];</p> | | | | | |

| | | | | | | |
|----|--|------|-----------|--------------------|---|----------|
| | <p>[MapolyID:Mapoly0014s0160]; [GO:0003993] acid phosphatase activity; [GO:0016787] hydrolase activity; [GO:0046872] metal ion binding</p> <div>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>: -II:IIIII:-:II:-III:</p> <p>3' -UGUGUUAGCUAGCUACUAGGG-5'</p> <p>CA G GU</p> <p>Mp1g10670</p> </div>  </div> | | | | | |
| 5. | Mp1g10660 | 2386 | 1695:1717 | c:1705(5.52 1 51)* | - | 3' - UTR |
| | <p>>Mp1g10660.1 [SMART:SM00855] PGAM_5; [G3DSA:3.40.50.1240]; [ProSitePatterns:PS00175] Phosphoglycerate mutase family phosphohistidine signature.; [Pfam:PF00300] Histidine phosphatase superfamily (branch 1); [SUPERFAMILY:SSF53254] Phosphoglycerate mutase-like; [PANTHER:PTHR48100] BROAD-SPECIFICITY PHOSPHATASE YOR283W-RELATED; [CDD:cd07067] HP_PGM_like; [G3DSA:3.40.50.1240:FF:000018] Phosphoglycerate mutase; [KEGG:K22200] 2-carboxy-D-arabinitol-1-phosphatase [EC:3.1.3.63]; [KOG:KOG0235] Phosphoglycerate mutase; [G]; [MapolyID:Mapoly0014s0161]; [GO:0003824] catalytic activity; [GO:0005737] cytoplasm; [GO:0016791] phosphatase activity</p> <div>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>-III:IIIII:II-::I--II</p> <p>3' -UGGUUAGCUAGACCUUCCCGA-5'</p> <p>C UU C</p> <p>Mp1g10660</p> </div>  </div> | | | | | |

| | | | | | | |
|----|--|------|-----------|---|--------------------|-----------------|
| 6. | Mp3g22210 | 1910 | 1321:1337 | - | c:1329(3.29 2 28)+ | Coding sequence |
| | <p>>Mp3g22210.1 [Coils:Coil] Coil; [SUPERFAMILY:SSF48264] Cytochrome P450; [PRINTS:PR00463] E-class P450 group I signature; [G3DSA:1.10.630.10:FF:000126] Predicted protein; [G3DSA:1.10.630.10] Cytochrome P450; [CDD:cd20618] CYP71_clan; [Pfam:PF00067] Cytochrome P450; [ProSitePatterns:PS00086] Cytochrome P450 cysteine heme-iron ligand signature.; [PRINTS:PR00385] P450 superfamily signature; [PANTHER:PTHR47944] CYTOCHROME P450 98A9; [KOG:KOG0156] Cytochrome P450 CYP2 subfamily; [Q]; [MapolyID:Mapoly1191s0001]; [GO:0005506] iron ion binding; [GO:0016705] oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; [GO:0004497] monooxygenase activity; [GO:0020037] heme binding</p> <div>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>I-I---I:III:I-I---:-I</p> <p>3' -AAAGGGUGUUAAGCCCGAAGCAA-5'</p> <p>G</p> <p>Mp3g22210</p> </div>  </div> | | | | | |
| 7. | Mp3g17650 | 3284 | 2817:2835 | - | c:2829(3.30 2 27)+ | 3' - UTR |
| | <p>>Mp3g17650.1 [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:1.10.510.10:FF:000358] Putative leucine-rich repeat receptor-like serine/threonine-protein kinase; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [Pfam:PF13855] Leucine rich repeat; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [Pfam:PF00560] Leucine Rich Repeat; [G3DSA:3.80.10.10:FF:000041] LRR receptor-like serine/threonine-protein kinase ERECTA; [CDD:cd14066] STKc_IRAK; [Pfam:PF08263] Leucine rich repeat N-terminal domain; [SUPERFAMILY:SSF52058] L domain-like; [G3DSA:3.80.10.10:FF:000400] Nuclear pore complex protein NUP107; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [SMART:SM00369] LRR_typ_2; [Pfam:PF00069] Protein kinase domain; [ProSitePatterns:PS00108] Serine/Threonine protein kinases active-site signature.; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [PANTHER:PTHR27008]</p> | | | | | |

| | | | | | | |
|----|---|------|-----------|--------------------|---|----------|
| | <p>OS04G0122200 PROTEIN; [SMART:SM00220] serkin_6; [ProSiteProfiles:PS51450] Leucine-rich repeat profile.; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG0618] Serine/threonine phosphatase 2C containing leucine-rich repeats, similar to SCN circadian oscillatory protein (SCOP); C-term missing; [T]; [MapolyID:Mapoly0039s0031]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding</p> <div>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>--I-III:-III-I--I-:</p> <p>3' -UUGGCAGUAGAACCCAGGACGACU-5'</p> <p>U</p> <p>Mp3g17650</p> </div>  </div> | | | | | |
| 8. | Mp4g19340 | 2838 | 3759:2777 | c:2771(6.26 1 87)* | - | 3' - UTR |
| | <p>>Mp4g19340.1 [NCBIfam:TIGR00820] ZIP zinc/iron transport family; [PANTHER:PTHR11040] ZINC/IRON TRANSPORTER; [MobiDBLite:mobidb-lite] consensus disorder prediction; [Pfam:PF02535] ZIP Zinc transporter; [KEGG:K14709] solute carrier family 39 (zinc transporter), member 1/2/3; [KOG:KOG1558] Fe²⁺/Zn²⁺ regulated transporter; [P]; [MapolyID:Mapoly0169s0010]; [GO:0055085] transmembrane transport; [GO:0030001] metal ion transport; [GO:0046873] metal ion transmembrane transporter activity; [GO:0005385] zinc ion transmembrane transporter activity; [GO:0071577] zinc ion transmembrane transport; [GO:0005886] plasma membrane; [GO:0016020] membrane; MpZIP3; Transcription factor; Fe homeostasis protein</p> | | | | | |

| | | | | | | |
|----|--|------|---------|---------------------|---|-----------------|
| | <div>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>-II--IIII-III-III:</p> <p>3' -ACAUGUAAGACCAAAGGU-5'</p> <p>Mp4g19340.1</p> </div>  </div> | | | | | |
| 9. | Mp8g03870 | 1298 | 972:987 | c:981(11.78 1 147)* | - | Coding sequence |
| | <p>>Mp8g03870.1 [PANTHER:PTHR23164] EARLY ENDOSOME ANTIGEN 1; [G3DSA:3.30.40.10] Zinc/RING finger domain, C3HC4 (zinc finger); [ProSiteProfiles:PS50848] START domain profile.; [SMART:SM00064] fyve_4; [SUPERFAMILY:SSF57903] FYVE/PHD zinc finger; [G3DSA:3.30.530.20]; [Pfam:PF01852] START domain; [Pfam:PF01363] FYVE zinc finger; [ProSiteProfiles:PS50178] Zinc finger FYVE/FYVE-related type profile.; [SUPERFAMILY:SSF55961] Bet v1-like; [CDD:cd00177] START; [KOG:KOG1729] FYVE finger containing protein; N-term missing; C-term missing; [R]; [MapolyID:Mapoly0012s0177]; [GO:0008289] lipid binding; [GO:0046872] metal ion binding</p> <div>  <div> <p>MpmiR11796</p> <p>5' -UCCAGUCGAUUUGUGGGAUCU-3'</p> <p>----:IIII:II:IIII-:I-</p> <p>3' -CCUACUAGCUGAAUACCCCGUUGUC-5'</p> <p>Mp8g03870.1</p> </div>  </div> | | | | | |

5.4 Identification of novel miRNAs expressed specifically or abundantly in archegoniophores

Our sRNA NGS data identified several potential novel miRNAs specifically expressed in archegoniophores. Out of bioinformatically identified 10 miRNA species, I was able to confirm the accumulation of 3 novel potential miRNAs: FAN-1 (highly accumulated in archegoniophores), FAN-3 (differentially expressed in vegetative and reproductive organs), and FAN-4 (specifically expressed in archegoniophores).

5.4.1 FAN-1

sRNA NGS sequencing results revealed the presence of a 24 nt miRNA, highly enriched in archegoniophores and also detectable in antheridiophores, although to a much lesser extent (Fig. 58A). Interestingly, genome database analysis revealed that this miRNA is encoded within the precursor of another annotated miRNA, Mpo-miR11693b, and likely represents its passenger strand (Tsuzuki et al., 2016). Previous sRNA NGS data by Tsuzuki et al., (2016) as well as our NGS results, revealed high accumulation of Mpo-miR11693b predominantly in archegoniophores (Fig. 58C) (Tsuzuki et al., 2016). Notably, Mpo-miR11693b accumulates at higher levels in archegoniophores as compared to FAN-1, which suggests that FAN-1 is the miRNA* (passenger strand) of Mpo-miR11693b (Fig. 58A, 58C).

Moreover, I noticed certain discrepancies in the annotation of this miRNA across different versions of genome database. In version 3.1, 20 nt out of 24 nt FAN-1 sequence was annotated as Mpo-miR11693b, whereas in versions 5.1 and 7.1, Mpo-miR11693b was annotated on the opposite strand. Despite these differences, our sRNA NGS and northern blot hybridization results consistently confirmed high accumulation of FAN-1 in archegoniophores and, to a lesser extent, in antheridiophores (Fig. 58A, 58B). Interestingly, the miRNA signal was found closer in size to a 21 nt RNA marker rather than 24 nt (Fig. 58B). Furthermore, the sequence of pre-miRNA does not overlap with any protein coding gene and has a classical stem-loop structure (Fig. 58D), indicative of its potential to generate a functional mature miRNA, and possibly the corresponding miRNA*.

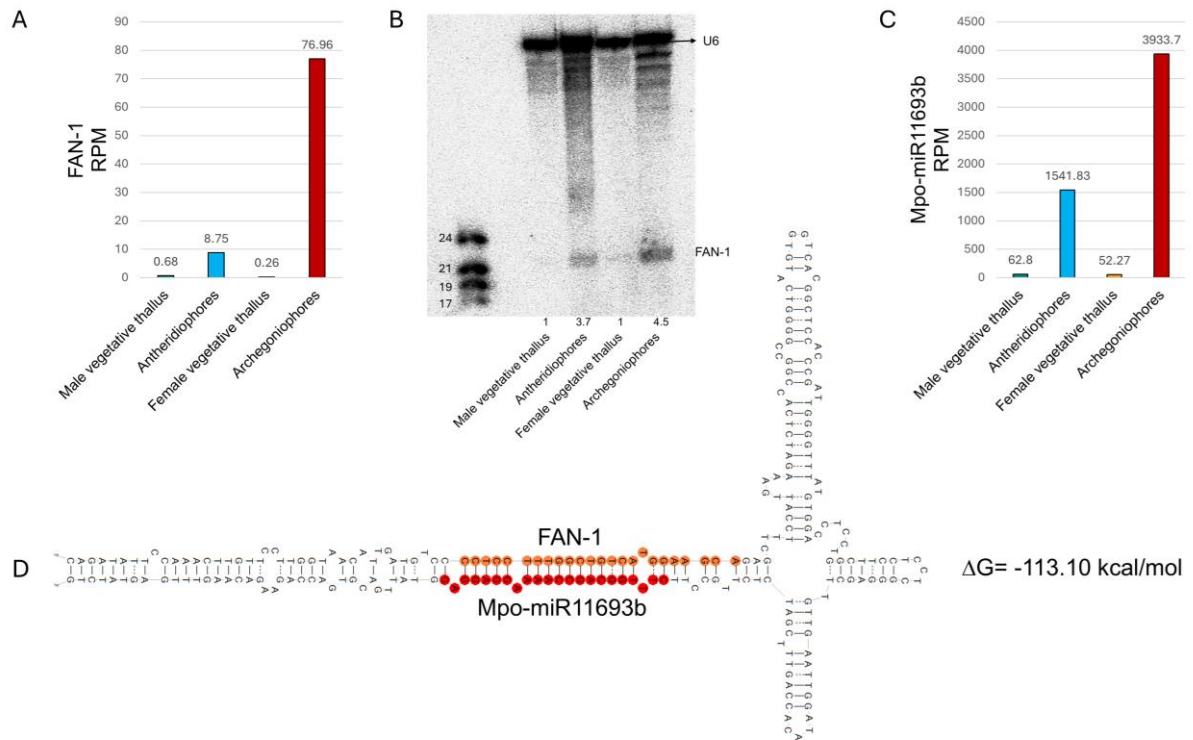


Figure 58. FAN-1 accumulation pattern. (A) and (C) sRNA NGS sequencing results with normalized read counts presented above each bar; RPM- Reads Per Million. (B) Northern hybridization; numbers presented below the blot show relative strengths of signal of miRNA bands, control signals were arbitrarily established as 1, signal intensity differences were separately obtained for male vegetative thalli control/ antheridiophores, female vegetative thalli control/ archegoniophores; Left side of the blot - RNA marker depicting 17-24 nucleotide long RNAs; U6 snRNA hybridization - RNA loading control. (D) Structure of FAN-1 pre-miRNA as predicted by RNA Fold (Folder ver. 1.11); Mpo-miR11693b miRNA sequence found in Tsuzuki et al., (2016) and our sRNA NGS data is highlighted in red and FAN-1 miRNA* sequence is highlighted in orange; minimum free energy (ΔG) of predicted structure is written on the right side.

5.4.2 FAN-3

Another miRNA identified in our sRNA NGS data showed accumulation in the generative organs of *Marchantia*, particularly enriched in archegoniophores (Fig. 59A). However, northern hybridization revealed downregulated levels in antheridiophores and archegoniophores when compared with male and female vegetative thalli (Fig. 59B). This discrepancy could be explained by genome database analyses, which showed that this miRNA is encoded by multiple loci located on both sex chromosomes (both U and V) and autosomal chromosomes. Structural analyses confirmed that its precursor forms a classical stem-loop structure (Fig. 59C), supporting the potential of FAN-3 to function as a mature miRNA. Notably, FAN-3 pre-miRNA sequence found in different loci is identical. Nevertheless, further experimental validation is required to explain these observations.

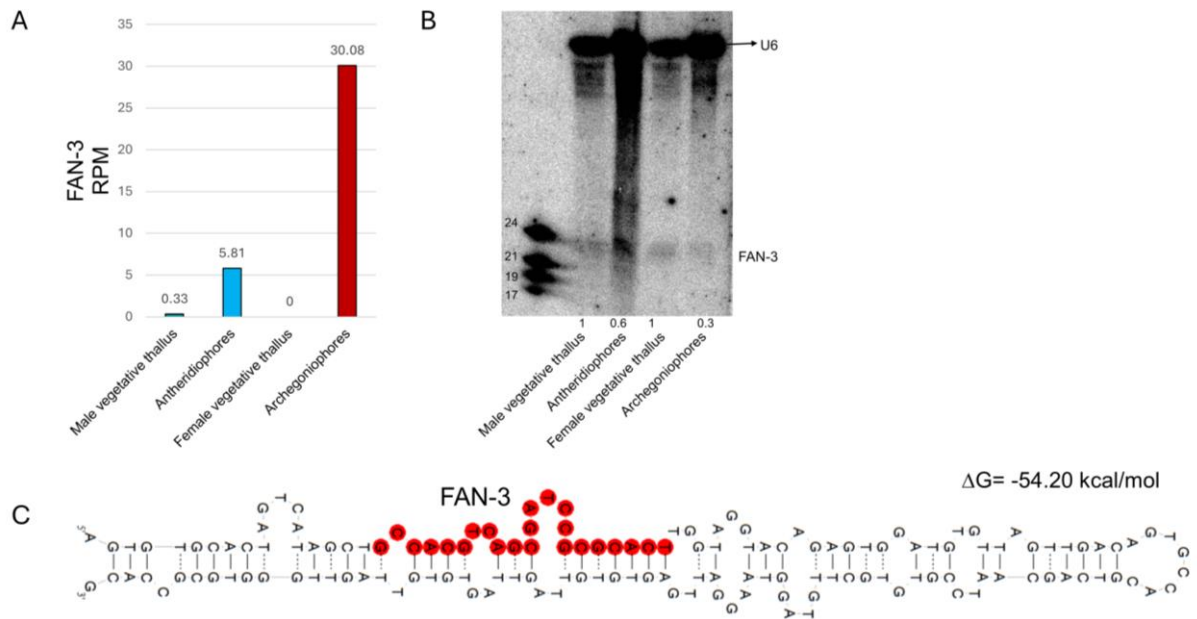


Figure 59. FAN-3 accumulation pattern. (A) sRNA NGS sequencing results with normalized read counts presented above each bar; RPM- Reads Per Million. (B) Northern hybridization; numbers presented below the blot show relative strengths of signal of miRNA bands, control signals were arbitrarily established as 1, signal intensity differences were separately obtained for male vegetative thalli control/ antheridiophores, female vegetative thalli control/ archegoniophores; Left side of the

blot - RNA marker depicting 17-24 nucleotide long RNAs; U6 snRNA hybridization - RNA loading control. (D) Structure of FAN-3 pre-miRNA as predicted by RNA Fold (Folder ver. 1.11); miRNA sequence is highlighted in red and minimum free energy (ΔG) of predicted structure is written on the right side.

5.4.3 FAN-4

sRNA NGS analysis revealed the presence of a miRNA of 22 nt in length: FAN-4, which is enriched in archegoniophores (Fig. 60A). This accumulation was further confirmed by northern blot hybridization (Fig. 60B). Structural analysis demonstrated that its pre-miRNA forms a classical stem-loop structure (Fig. 60D). Interestingly, genome database analysis showed that another annotated miRNA, Mpo-miR11886, originates from the same genomic locus and shares sequence similarity with FAN-4, with 18 nt overlap out of 22 nt. Previous sRNA NGS data by Tsuzuki et al., (2016) and our NGS data showed that Mpo-miR11886 also accumulates predominantly in archegoniophores (Fig. 60C) (Tsuzuki et al., 2016). However, there are differences in the lengths between the annotated Mpo-miR11886 precursor and FAN-4 precursor-miRNA. The predicted pre-miRNA structure for FAN-4 is 201 nt in length (Fig. 60D), while Mpo-miR11886 has annotated precursor-miRNA length of 156 nt (Fig. 60E), and both structures partially overlap. These findings suggest that FAN-4 and Mpo-miR11886 are sequence isomiR variants derived from the same precursor. Despite this, the accumulation of FAN-4 was confirmed exclusively in archegoniophores by northern blot hybridization (Fig. 60B). Additionally, database analysis revealed that a part of predicted as well as annotated precursor (44 bp) overlaps with the 5' - UTR of protein-coding gene of unknown function, encoded from the opposite strand (Fig. 60F).

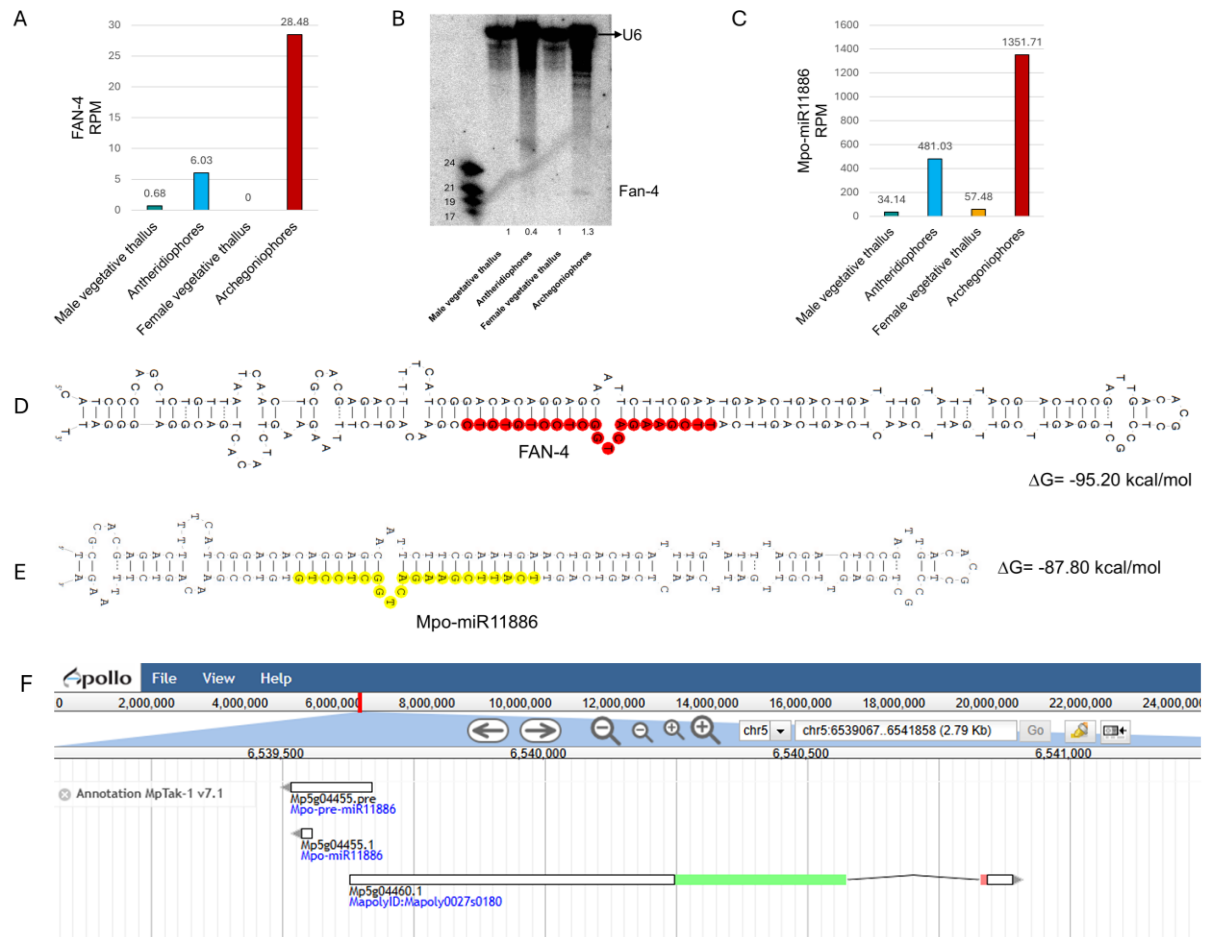


Figure 60. FAN-4 accumulation pattern. (A) and (C) sRNA NGS sequencing results with normalized read counts presented above each bar; RPM- Reads Per Million. (B) Northern hybridization; numbers presented below the blot show relative strengths of signal of miRNA bands, control signals were arbitrarily established as 1, signal intensity differences were separately obtained for male vegetative thalli control/ antheridiophores, female vegetative thalli control/ archegoniophores; Left side of the blot - RNA marker depicting 17-24 nt long RNAs; U6 snRNA hybridization - RNA loading control. (D) & (E) Structures of pre-miRNAs of FAN-4 and annotated Mpo-miR11886 as predicted by RNA Fold (Folder ver. 1.11); FAN-4 sequence is highlighted in red and Mpo-miR11886 is highlighted in yellow; minimum free energy (ΔG) of predicted structures is written on the right side. (F) Screenshot from genome browser showing overlapping region between precursor-miRNA and 5'-UTR of protein-coding gene (Mp5g04460).

Chapter 6. Discussion

In this thesis, I tried to understand the role of liverwort-specific miRNAs in shaping the development of *Marchantia*. The important roles of conserved miRNAs shared across land plants have been well-established, as they predominantly regulate transcription factors involved in fundamental plant development processes and stress tolerance mechanisms (Lin et al., 2016; Tsuzuki et al., 2016; Lin & Bowman, 2018; Flores-Sandoval et al., 2018; Tsuzuki et al., 2019; Yelina et al., 2023; Futagami et al., 2025). These analyses of conserved miRNA functions have provided important insights into the conserved and fundamental regulatory pathways in early-diverging plant lineages. However, the role of liverwort-specific miRNAs – a unique group likely contributing to lineage-specific innovations – remains largely unexplored, with only two examples functionally characterized to date (MpFRH1/MpRSL1 and Mpo-MR-13/MpSPL1) (Honkanen et al., 2018; Thamm et al., 2020; Streubel et al., 2023).

Of the 265 miRNAs identified in *Marchantia*, only nine belong to conserved miRNA families shared among all land plants. The remaining non-conserved miRNAs were either kept (but lost in the lineage of seed plants) or acquired during evolution, after the divergence from the common ancestor of seed and non-seed plants (Axtell et al., 2007; Lin et al., 2016; Tsuzuki et al., 2016). In my thesis, I aimed to bridge this research gap by analyzing the expression patterns of six selected liverwort-specific miRNAs (MpmiR11737a/b, MpmiR11865*/11865, MpmiR11887, and MpmiR11796) and conducting functional studies on two of them (MpmiR11887 and MpmiR11796). The phenotypic characterization of these two miRNA KO plants have provided valuable insights into their biological roles in *Marchantia*'s vegetative and reproductive development.

Liverwort-specific miRNA family: MpmiR11737

The MpmiR11737 family (MpmiR11737a and MpmiR11737b) represents the first example of a miRNA shared with *P. endiviifolia* (Pen-miR8163) (Alaba et al., 2015; Tsuzuki et al., 2016; Pietrykowska et al., 2022). *P. endiviifolia* and *M. polymorpha* shared a common ancestor around 250 million years ago (Feldberg et al., 2014).

Despite their ancient divergence, they still share some common liverwort-specific miRNAs, which strongly suggests their important role in liverworts. A single miRNA sequence was identified for Pen-miR8163, suggesting that there is a single miR8163 family member in *Pellia* genome. Alternatively, there is a possibility that additional family members could be present in *Pellia*, however, they encode the same sequence of mature miRNA.

The analysis of accumulation profiles of both MpmiR11737a and MpmiR11737b revealed distinct expression patterns across vegetative and reproductive organs of *Marchantia*, suggesting the potential for divergent biological roles. The gene structure in the case of MpmiR11737a was identified. Even though it is encoded within a long intron (~8 kb) of a protein-coding gene (*Mp5g12920*), my research proved that it represents an independent transcriptional unit, as TSS lies within this intron and termination occurs within the 3rd exon of its host gene. The overlap between protein-coding genes and *MIR* genes is rarely observed in case of higher plants (Bielewicz et al., 2012; Szarzynska et al., 2009). Nevertheless, there are examples in *Arabidopsis* where *MIR* genes overlap with protein-coding genes: At-miR402 is encoded within a protein-coding gene overlapping with *AT1G77230* encoding a pentatricopeptide protein of unknown function (Knop et al., 2017). Another example: At-miR400 is co-transcribed from *AT1G32583* gene, which encodes tapetum determinant protein, and miRNA is encoded within the first intron of this gene (Yan et al., 2012). The accumulation of these two miRs in *Arabidopsis* is regulated via alternative splicing events in heat stress. We cannot exclude that the accumulation of liverwort-specific miRNAs in *Marchantia*, encoded within introns of protein-coding genes, are also regulated in response to environmental cues.

The second member of MpmiR11737 family, MpmiR11737b, exhibits extremely low abundance in sRNA NGS data. This, in conjunction with unsuccessful 5'-RLM and 3'-RACE attempts to identify the full-length transcript of *MpMIR11737b* gene, suggests that MpmiR11737b could represent a functionally redundant miRNA, with the major regulatory role attributed to MpmiR11737a. Alternatively, it is plausible that this miRNA accumulates upon specific conditions during *Marchantia*'s life cycle.

The degradome sequencing data analysis identified Mp *1g15010*, which encodes an uncharacterized protein, as a putative target of MpmiR11737a. Several aspects of target prediction appears to hold considerable promise, including a significant peak in t-plot analysis, near-perfect complementary base-pairing between the miRNA and target mRNA, and a reverse correlation between mature MpmiR11737a and Mp1g15010 mRNA levels. However, it is important to note that experimental validation is still required to confirm Mp *1g15010* as a *bona fide* target of MpmiR11737a. Although, Mp1g15010 encodes a protein of unknown function, future functional characterization will be necessary to achieve comprehensive understanding of miRNA/mRNA target interaction and the biological role of Mp *1g15010*. Additionally, MpmiR11737b differs from MpmiR11737a by 2nt, one mismatch near 5'-end and the other near 3'-end, which suggests that both miRNAs may possibly have the same target mRNA.

Liverwort-specific miRNAs: MpmiR11865*/MpmiR11865

MpmiR11865* represents another example of miRNA shared with *P. endiviifolia* (*Pen-miR8170*) (Pietrykowska et al., 2022). Interestingly, MpmiR11865* was initially recognized as a passenger strand within a known MpmiR11865 precursor (Tsuzuki et al., 2016; Pietrykowska, 2020). Notably, no corresponding miRNA species was detected in sRNA NGS data of *P. endiviifolia*, suggesting that MpmiR11865 may have gained stability specifically in Marchantia. Northern blot hybridization revealed another intriguing fact: both MpmiR11865* and MpmiR11865 exhibit opposite accumulation profiles, suggesting that both are functionally active miRNAs derived from the same precursor, and potentially involved in different regulatory mechanisms depending on the tissue of expression (Aggarwal et al., 2024). Reports from angiosperms have shown that in some cases, miRNA* can act as a functional molecule and may accumulate to higher levels than their corresponding miRNAs during plant development and in response to stress (Rajagopalan et al., 2006; Bologna & Voinnet, 2014; Swida-Barteczka et al., 2023). Under certain circumstances, the miRNA* retain several features of active miRNAs and serve as a guide strand, ultimately triggering the silencing of a new target (Manavella et al., 2013; Swida-Barteczka et al., 2023). Hence, both miRNA and miRNA* can be loaded onto RISC and can have different

regulatory outcomes. These observations mark an intriguing starting point for investigating the functional relevance of MpmiR11865* and MpmiR11865. Moreover, Mp*MIR11865* may represent the first example of a gene in liverworts, encoding both miRNA and miRNA* as functional molecules.

Furthermore, as in the case of Mp*MIR11737a*, my studies revealed that Mp*MIR11865* gene represents an intron-less independent transcription unit, not overlapping with any protein coding gene. Degradome sequencing identified two different targets for miRNA* and miRNA. Mp *1g05970*, encoding a tRNA ligase 1 protein, was identified as a putative target of MpmiR11865*, while Mp*6g13460*, encoding a nuclear AAA+ ATPase protein, was predicted as a potential target of MpmiR11865 (Aggarwal et al., 2024). In both cases, inverse expression pattern between the miRNAs and their putative targets supported these predictions, although further experimental validations are still necessary.

Liverwort-specific miRNA: MpmiR11887

MpmiR11887 represents another interesting example, being exclusively expressed in antheridiophores (Pietrykowska, 2020; Aggarwal et al., 2024). As in the previously described case of MpmiR11737a, MpmiR11887 is encoded within 3'-UTR of a protein-coding gene (Mp*6g01830*), encoding a short open reading frame (ORF) of unknown function. There is a possibility that this ORF may encode microRNA-encoded peptide (miPeP). Such coding abilities of pri-miRNAs have been reported in *Arabidopsis thaliana*, *Medicago truncatula*, *Glycine max*, *Vitis vinifera*, *Hordeum vulgare* and even *Homo sapiens* (Lauressergues et al., 2015; Couzigou et al., 2016; Fang et al., 2017; Sharma et al., 2020; Chen et al., 2022; Chojnacka et al., 2023). miPEPs have been shown to activate the transcription of their corresponding *MIR* genes, which subsequently increases the accumulation of mature miRNA (Ren et al., 2021). Alternatively, we cannot exclude the possibility that the short ORF is not translated, and such peptide is not present in Marchantia.

I observed an intriguing phenomenon that despite the exclusive accumulation of MpmiR11887 in antheridiophores, GUS expression under the Mp*MIR11887* gene promoter was also observed in female tissues. Why do we see the Mp*MIR11887*

promoter activity in female tissues? The pri-MpmiR11887 can be detected in female tissues (Fig. 22C), although mature MpmiR11887 is not (Fig. 22A, 22B). This could result in the production of a short peptide (above called miPeP) and exclude the production of mature miRNA in female tissues. This may explain the detection of GUS activity in apical notches of female vegetative thallus and archegonia (Fig. 26C). Alternatively, we may observe ectopic non-specific expression of MpMIR11887 promoter in female tissues, for which further studies could be performed. Moreover, the promoters are often not 100% tissue-specific unless key elements (inducers or repressors) are included.

Phenotypic analysis of CRISPR/Cas9 generated male $\Delta mpmir11887^{ko}$ plants revealed that MpmiR11887 is involved in regulating both vegetative and reproductive development in Marchantia. Although the $\Delta mpmir11887^{ko}$ plants exhibited larger thallus area in the initial stages of gemmae development, the size differences between $\Delta mpmir11887^{ko}$ plants and WT were negligible upon two weeks of gemmae growth. Notably, weak expression of pri-MpmiR11887 could be detected in male vegetative thalli (Fig. 22C), despite the exclusive accumulation of mature mpmiR11887 in antheridiophores. Given the absence of mature MpmiR11887 in vegetative tissue, it is possible that a short peptide (miPeP) encoded within the pri-MpmiR11887 may contribute to the observed effects. Additionally, $\Delta mpmir11887^{ko}$ plants carry mutation within the pre-MpmiR11887 region while the short ORF expression (miPeP) is presumably unaffected. Thus, the potential role of this ORF in this case needs further investigation. This intriguing observation requires careful in-depth studies and a detailed dissection of both – the miRNA and peptide within this same gene.

Interestingly, the timing of reproductive development was noticeably accelerated in KO plants. In 5-week-old KO plants, mature antheridiophores were already fully developed, while WT plants were only initiating their formation. Additionally, consistent with the larger thalli, the antheridial discs and antheridia of KO plants were significantly wider than those of WT. Despite these developmental differences, the functional aspects of reproduction appeared unaffected, as KO and WT plants both released motile sperm cells upon water application. Thus, while MpMIR11887 gene function impacts the timing of antheridiophore development, but is not essential for the basic

formation of motile and functional sperm cells ($\Delta mpmir11887^{ko}$ sperm cells were able to fertilize WT female plants producing sporophytes). While phenotypic analyses confirmed normal sperm release in $\Delta mpmir11887^{ko}$ plants, detailed studies on sperm cell morphology, flagellar beating etc. have still to be performed. Therefore, further investigations are necessary to fully assess whether MpmiR11887 influences more aspects of sperm cell development or function that might not be apparent through general observations alone. Several studies gave insights into the molecular basis of sperm function in Marchantia, which I will discuss below. However, none of these described genes were affected significantly in their transcript levels in $\Delta mpmir11887^{ko}$ antheridia transcriptome.

cAMP (ADENOSINE 3',5' CYCLIC MONOPHOSPHATE) is a key universal secondary messenger, synthesized from ATP by ADENYLYL CYCLASE (AC) and hydrolyzed to AMP by cAMP PHOSPHODIESTERASE (PDE). In Marchantia, a dual functioning enzyme, MpCAPE (COMBINED AC WITH PDE) has been shown to play a pivotal role in sperm motility, particularly in promoting linear forward motility and regulating flagellar beating, although it is not essential for spermatogenesis (Yamamoto et al., 2024). Interestingly, while $proMpCAPE:GUS$ signals increased with antheridia maturation, $proMpMIR11887:GUS$ signals decreased with antheridia maturation. This suggests that, at some stage of antheridia development, there may be a transient overlap in the expression patterns of MpCAPE and MpMIR11887 and it is worth to check the motility patterns in MpmiR11887 KO plants.

Marchantia has four *MLO* genes, but only MpMLO1 showed expression in the jacket cells as well as spermatogenous cells of antheridia. Experimental analyses demonstrated that the jacket cells in *Mpmlo1* mutant did not undergo cell shrinkage and sperm discharge was abolished. Hence, MpMLO1 was found to be essential for PCD of antheridial tip cells to trigger the sperm outflow (Cao et al., 2024). The other three *MLO* genes showed expression specifically in the sperm cells, which also overlaps with MpMIR11887 gene expression, suggesting potential roles in spermatogenous cell differentiation, which remains unexplored. Interestingly, the transcript levels of MpMLO2 and MpMLO3 were increased (Log2 fold change ~ -2.26) in our transcriptomic dataset for $\Delta mpmir11887^{ko}$ antheridia, though statistically not

significant. RT-qPCR analysis should be performed to confirm MpMLO2 and MpMLO3 mRNAs essential upregulation in the $\Delta mpmir11887^{ko}$ antheridia. It might be an attractive possibility to test whether MpmiR11887 regulates sperm cell differentiation directly or indirectly via the MLO gene family, as the regulatory pathways could be more complex in liverworts than expected.

MpFER (FERONIA) has also been suggested to be needed for the development of sperm cells. A single CrRLK1L gene (*Catharanthus roseus* Receptor like protein kinase 1-like) present in Marchantia, MpFER, is expressed throughout its life cycle. Hence, it has a broad involvement in both vegetative and reproductive development, including rhizoid formation, cell expansion, male and female gametogenesis and sporophyte development (Mecchia et al., 2022). However, the transcriptional regulation and upstream components in the CrRLK1L pathway in Marchantia remain to be elucidated. Given the role of MpFER in male gametogenesis and the male-specific expression of MpmiR11887, it would be intriguing to investigate whether this miRNA or its target is associated with this network.

The MYB transcription factor, MpDUO1 (*DUO POLLEN 1*) is known to be the master regulator controlling sperm cell differentiation in Marchantia (Higo et al., 2018). Furthermore, it was shown that MpDUO1 regulates several antheridium expressed genes including MpDAZ1 (*DUO1-ACTIVATED ZINC FINGER 1*), MpLC7 (*DYNEIN LIGHT CHAIN 7*), and MpPRM (gene encoding protamine-like sperm nuclear protein) to coordinate sperm differentiation properly (Higo et al., 2016; Higo et al., 2018). MpPRM transcript levels were also upregulated in sperm cells of $\Delta mpmir11887^{ko}$ plants. A link between MpMIR11887 activity and the above stated genes could be studied in more details.

A single MCA (*MID1-COMPLEMENTING ACTIVITY*) gene is present in Marchantia. MpMCA (Mp5g19510) was found to be involved in cell proliferation and development via $[Ca^{2+}]_{cyt}$ regulation at the actively dividing cells (apical notch region, egg cells, antheridium cells) during both vegetative growth as well as sexual reproduction (Iwano Megumi et al., 2025). Detailed analysis of Mpmca mutants revealed distorted shapes of both jacket cells and spermatogenous cells (SC) and moreover, SC were filled with

granular structures, suggesting impaired cell division. Detailed microscopic analysis of SC in $\Delta mpmir11887^{ko}$ mutants should also be carefully and deeply inspected.

It should be highlighted that the sperm development is a complex process where *MpMIR11887* gene might be one piece of the puzzle. We identified various putative targets for *MpmiR11887* in our degradome data, out of which *MpATX1* gene was quite interesting. *MpATX1* encodes a chromatin related protein and is the homolog of histone lysine methyl transferase *ATX2* in *Arabidopsis* whose promoter driven GUS expression was reported to be mainly concentrated in pollen grains (Saleh et al., 2008). The putative slice site lies within the coding sequence and degradome data was supported by the upregulated expression levels of *MpATX1* mRNA in the antheridiophores and sperm cells of all the three $\Delta mpmir11887^{ko}$ plants as compared to WT. It is interesting to further explore this area of research and to look whether overexpression lines of *MpATX1* produce the same phenotype as $\Delta mpmir11887^{ko}$ plants. Additionally, another potential target mRNA has been identified by us in the transcriptomic data: *Mp4g22015* gene with unknown function, which is highly upregulated in $\Delta mpmir11887^{ko}$ plants. Further studies are needed to assess whether *MpmiR11887* targets both mRNAs or one of them.

Liverwort-specific miRNA: *MpmiR11796*

MpmiR11796 revealed high levels of mature miRNA accumulation in archegoniophores (Tsuzuki et al., 2016; Pietrykowska, 2020; Aggarwal et al., 2024). Interestingly, RT-qPCR analysis detected lower levels of pri-*MpmiR11796* in archegoniophores when compared to female vegetative thalli, suggesting that pre-*MpmiR11796* undergoes tissue-specific processing i.e. in this case, pri-*MpmiR11796* is efficiently processed into mature miRNA in archegoniophores, thereby reflecting post-transcriptional regulation affecting the final accumulation of mature miRNA. Moreover, in angiosperms, it is known that level of mature miRNA does not always correlate with the level of pri-miRNA (Barciszewska-Pacak et al., 2015).

Like previously described cases, *MpMIR11796* gene acts as an independent transcriptional unit despite having partially overlapping sequence with a protein-coding gene (Aggarwal et al., 2024). Interestingly, *MpMIR11796* gene promoter activity via

histochemical GUS staining showed strong expression in the rhizoids inside archegoniophore stalk and digitate rays, ventral scales of vegetative thallus and antheridiophores. In this thesis, the GUS expression was checked in the older thallus plants. However, juvenile gemmae (e.g. 3 – 7 days old) should also be checked to see if the promoter is active in rhizoid precursor cells or rhizoids produced by vegetative thallus, to find a clue if MpmiR11796 is involved in polarized tip growth or growth polarity. Although, GUS expression could be observed in specific cells on the extensions of female vegetative thallus, it is not clear if these cells are rhizoid precursor cells (see Fig. 39B).

Analysis of rhizoid growth in *Marchantia* has been very helpful in understanding polarized tip growth as they undergo persistent polarized cell elongation (Shimamura, 2016; Kohchi et al., 2021). Many studies have tried to understand the molecular basis of this phenomenon. A total of 33 genes required for the growth of rhizoids were identified. These genes included cell wall biosynthesis and integrity sensing, vesicle transport and cytoskeleton, and genes with unknown functions (Honkanen et al., 2016). Although, the transcript levels of these genes showed a trend of increase or decrease, the differences were not significant in our transcriptomic dataset for female $\Delta mpmir11796^{ko}$ plants. *Marchantia* has single copy counterparts of *Arabidopsis* genes involved in root hair growth, which indicates that these genes might have additional developmental roles. Additionally, other rhizoid developmental genes remain to be discovered (Honkanen et al., 2016).

The stabilization of microtubule cytoskeleton is necessary for root hair growth and orientation (Bibikova et al., 1999). NEK (Never-in-mitosis (NIMA)-related kinases) is a serine/threonine protein kinase which regulates directional growth in plants (Motosé et al., 2008; Sakai et al., 2008). A single *NEK* gene, *MpNEK1*, was found to be involved in tip growth of rhizoids in *Marchantia*. *MpNEK1* maintains the growth polarity of rhizoids by reorganizing and destabilizing the microtubules, without affecting their elongation. Therefore, *Mpnek1* mutant plants produced curly and spiral rhizoids due to fluctuations in the direction of growth (Otani et al., 2018). Since microtubules promote rhizoid elongation via organelle transport, the overexpression of *MpNEK1* suppresses rhizoid tip growth by affecting the organization of microtubules and/or microtubule-

dependent processes (Mase et al., 2025). Interestingly, RT-qPCR analysis showed that MpNEK1 transcript levels are upregulated in thalli of male $\Delta mpmir11796^{ko}$ mutant plants (Fig. 52E), whereas, transcriptome analysis from archegoniophores of $\Delta mpmir11796^{ko}$ mutant plants showed significant downregulation of MpNEK1 transcript levels. My results also show that the male $\Delta mpmir11796^{ko}$ mutant plants produced defects in rhizoid growth during early stages of gemma development, however, mature thalli produces rhizoids comparable to WT plants. Alternatively, the female $\Delta mpmir11796^{ko}$ mutant plants exhibited no differences in the rhizoid growth. It would be valuable to investigate if MpmiR11796 affects microtubule organization during polarized tip growth of rhizoids.

Moreover, MpNEK1 was also shown to be involved in cell division and proliferation as *Mpnek1* mutants showed suppression in the growth of both gametangiophores. But the lengths of archegoniophore stalks were strongly reduced in female *Mpnek1* mutants, however, the lengths of antheridiophore stalks in male *Mpnek1* mutants were reduced to a smaller magnitude as compared to WT plants (Mase et al., 2025). Interestingly, I observed similar results in the case of female $\Delta mpmir11796^{ko}$ mutant plants, which produced archegoniophores with extreme reduction in the size of stalks whereas male $\Delta mpmir11796^{ko}$ mutant plants produced antheridiophores and stalks with almost no noticeable differences when compared to WT plants. This suggests that MpMIR11796 also might be involved in cell division and proliferation of female reproductive structures. We also found a putative target for MpmiR11796 from transcriptomic and degradome data analyses: Mp3g17650 which also encodes for a serine/threonine protein kinase and has leucine-rich repeats (Table 7). However, further experiments need to be performed to confirm it as a target of MpmiR11796.

Furthermore, a highly conserved AAA-ATPase, MpKTN (KATANIN) was identified as a microtubule-severing enzyme and proved to play an important role in the formation of polar organisers (Attrill & Dolan, 2024). Polar organisers are liverwort-specific centrosome-like MTOCs (microtubule organizing centers), which contain endoplasmic reticulum, vesicles, microtubule ends and high accumulation of γ -tubulin, and are crucial for proper cell division plane alignment (Brown & Lemmon, 2011; Buschmann et al., 2016). AAA-ATPases perform diverse cellular activities, and this class remains

largely unexplored in *Marchantia*. From degradome analysis, I identified a potential target mRNA for MpmiR11796: Mp1g02530. According to MarpolBase, this gene encodes AAA+-ATPase with Vps4 C-terminal oligomerization domain. But further bioinformatic analysis revealed that it is identical to AtCDC48A (cell division cycle48) protein, which is crucial for cell division plane orientation and early embryo patterning in *Arabidopsis* (Gong et al., 2024).

Polarized cell growth further contributes to the change in shape of the cell, meaning establishment of cell polarity is essential as it ultimately guides the cell division: symmetric or asymmetric (Naramoto et al., 2022). Asymmetrical cell divisions were one of the key innovations in sessile land plants, regulating numerous plant physiological and developmental processes. The daughter cells generated from asymmetrical cell divisions acquire different developmental fate depending on the polarity information within the mother cell (De Smet & Beeckman, 2011). A recent study revealed that MpCKI1 (CYTOKININ-INDEPENDENT1) gene is a key regulator for the specification of female germline. Specifically, MpCKI1 establishes the polarity required for proper asymmetric cell divisions during archegonium development in *Marchantia* (Bao et al., 2024). However, the molecular basis underlying these asymmetrical divisions in *Marchantia* remain to be explored further in detail. Similarly, the female $\Delta mpmir11796^{ko}$ mutant plants produced several aberrations in the division of secondary central cell (SCC). 25 % of mature archegonia contained either double egg cells or unusual division between egg cell and ventral canal cell or unusually shaped egg cells. This suggests that MpmiR11796 and its target mRNA have a potential role in establishing the polarity of precursor egg cell and that its KO shows perturbations in the final stages of cell division which gives rise to egg cell. Since 75 % of the archegonia produced by female $\Delta mpmir11796^{ko}$ mutant plants were WT like, it may suggest that MpmiR11796-target mRNA module effect is under the threshold and/or masked by other regulators. It may also indicate that some epi-transcriptomic mechanisms might be involved in the regulation of asymmetrical divisions, in which MpmiR11796 could be involved. These asymmetrical divisions maintain cellular identity and secure the development of egg cells for successful fertilization (Durand,

1908; Shimamura, 2016). Hence, in female $\Delta mpmir11796^{ko}$ mutants, it is not clear whether it is a result of cytokinesis failure or a defect in cell specification or patterning. To recapitulate, among the 33 genes identified as involved in rhizoid growth, some genes were found to encode proteins involved in vesicle transport and cytoskeleton (Honkanen et al., 2016). The tip elongation requires the delivery and deposition of large amount of material (produced in the Golgi apparatus) to the elongating tip (Scheller & Ulvskov, 2010; Driouich et al., 2012). This relies on secretory vesicles as they deliver the required cell wall material, membranes and proteins to facilitate root hair elongation. Generally, much more membrane is delivered than required, therefore, the excess membrane needs to be recycled through endocytosis (Ketelaar et al., 2008). The regulatory checkpoint here is that the vesicles need to be delivered to the desired location, and be marked with their identity and their target, so that they can be trafficked accordingly. Coordinated activity of numerous factors are involved in conferring this identity, which includes Rab GTPases and SNARE protein families (Lipka et al., 2007; Nielsen et al., 2008; Ovečka et al., 2012). The most promising target found in transcriptome analysis combined with degradome analysis was MpRAB21 (Mp1g21050), which is a small GTPase involved in plethora of developmental processes, including membrane trafficking. MpRAB21 has two mRNA isoforms in Marchantia. Both mRNA isoforms encode 205-AA protein and the only difference lies in mRNA isoform 2 containing a shorter 5'-UTR. Surprisingly, transcriptomic analysis revealed that isoform1 is extremely upregulated whereas isoform2 is downregulated in the KO mutant. Literature analysis revealed that RAB21, the ancient molecular machineries for higher order biological functions, are present in basal species such as *Physcomitrium patens*, *Marchantia polymorpha*, *Selaginella moellendorffii*, but they were lost several times during plant evolution (Rensing et al., 2008; Banks et al., 2011; Klöpper et al., 2012; Uemura & Ueda, 2014).

The membrane-trafficking system has been uniquely developed in each organism and it has not been experimentally validated in Marchantia yet. The expression patterns and subcellular localization of all members of RAB GTPases were examined in Marchantia (Minamino et al., 2018). RAB21 has not been identified in the Arabidopsis genome, but it is present in animals (Klopper et al., 2012). In Marchantia, RAB21 is

ubiquitously transcribed in all tissues and citrine-fused MpRAB21 under the control of CaMV 35S promoter showed localization at punctate structures (Minamino et al., 2018). Moreover, RAB21 localizes to the early endosomes in mammalian cells and plays an important role in exocytosis and endocytosis (Simpson et al., 2004; Pellinen et al., 2006). Since MpRAB21 has not been confirmed to be the target of MpmiR11796, no firm conclusions could be drawn. But it will be an interesting project to functionally characterize MpRAB21, as its role in endosomal transport remains unexplored in *Marchantia* and to check if MpmiR11796 has a role in it.

Endosomes are involved in protein recycling, degradation, and participate in receptor-mediated signaling pathway. Proteins in plant plasma membrane (PM) are constitutively recycled between PM and early endosomes. Example of such protein includes PIN1 (PIN-FORMED1), PM H⁺-ATPase, and water channel plasma membrane intrinsic protein2 (PIP2a) (Otegui & Spitzer, 2008). Many receptors, for example receptor tyrosine kinases are activated at the PM. Various studies have shown that endosomes are signaling platforms in animals as they found the accumulation of various receptors in endosomes. In plants, endosomal trafficking plays a role in signaling via PM kinase receptors, for instance, BRI1 (Brassinosteroid insensitive1) which controls cell expansion and division (Geldner et al., 2007). PM proteins tagged with ubiquitin by E3 ligases undergo degradation and are delivered to endosomes. Endosomal protein complexes (ESCRTs) sort them into vesicles, giving rise to multivesicular endosome (MVB) formation inside the endosome (Otegui & Reyes, 2010).

Interestingly, the GO analysis provided some more insights into the biological processes affected by the knock-out of MpmiR11796 (Supplementary Table 9). One of these category of biological processes included ATP-binding, which included 12 genes. Interestingly, out of 12 genes, 3 genes encode PM H⁺-ATPases, 6 encode tyrosine and serine/threonine protein kinases, 1 encodes ABC transporter, 1 encodes CLP protease, and 1 encodes DNA mismatch repair protein. The second category included protein binding, including 28 genes. Out of these 28 genes, 8 genes encode F-box domain containing protein, 4 genes encode protein tyrosine and serine/threonine

kinases, and others have been presented in Table 8. At least 13 genes have potential direct evidence in endosomal trafficking from the literature discussed above.

Hence, polarized growth, growth polarity, coordination of cytoskeleton organization and vesicular trafficking are all connected within the developmental processes essential for plant growth and development. A molecular outline of these events need to be characterized in more details in *Marchantia* as the mechanisms have been understood in rhizoids to some extent.

Novel miRNAs

In my thesis, I confirmed the accumulation of 3 putative novel miRNAs in vegetative and reproductive organs of *Marchantia*. Similar to the previously described case of MpmiR11865*, FAN-1 was identified to be encoded within the precursor of another annotated miRNA, Mpo-miR11693b. Based on sRNA NGS data, lower accumulation of FAN-1 as compared to Mpo-miR11693b and its origin from the same precursor, supports the notion that FAN-1 is the passenger strand (miRNA*). Although miRNA* is degraded in most cases, I already discussed various examples above regarding evidences of miRNA* acting as functional molecule (Manavella et al., 2013; Swida-Barteczka et al., 2023). FAN-1's high accumulation in archegoniophores suggests its functional role and it may target mRNA independently, a phenomenon observed in various plant systems, including Mp*MIR11865* gene example already shown in *Marchantia*. The northern blot signal for FAN-1 was found to be closer to 21nt rather than 24nt seen in NGS. This may suggest the presence of iso-miRs which requires further studies. Moreover, the accumulation of both FAN-1 and MpmiR11693b in archegoniophores suggests that these miRNAs could be involved in female reproductive development.

Northern blot results revealed higher accumulation of FAN-3 in antheridiophores as well as detectable expression in other tissues. Its presence at multiple genomic loci could explain the discrepancy between NGS and northern blot results. Many plant miRNAs are known to originate from transposable elements (TEs) (Li et al., 2011). There is a possibility that FAN-3 might be located within TEs and this would explain uniform pre-miRNA hairpin structure of FAN-3 on multiple chromosomes. Although its

tissue-specific accumulation could be influenced by chromatin accessibility and modifications (as several reports have shown that TE-derived miRNAs have regulatory roles in stress responses), epigenetic silencing and activation depending on the chromatin state may affect final accumulation of FAN-3 (Nosaka et al., 2012; Noshay et al., 2021). This suggests that FAN-3 may contribute to fine-tuning target gene expression at a specific developmental stage or under certain environmental conditions during *Marchantia*'s life-cycle. Also, there is a possibility that the scaffold has not been correctly assembled in *Marchantia* genome. Therefore, further studies are needed to dissect the potential role of FAN-3 in *Marchantia*.

These findings underscores the complex and dynamic nature of miRNAs in *Marchantia* and highlight the need for further validation to clarify their biological roles. Our findings also suggest that although current annotations are extensive, the miRNA repertoire of *Marchantia* is still incomplete and additional functional miRNAs contributing to the developmental complexity and stress response mechanisms remain to be uncovered.

Chapter 7: Conclusions and Future perspectives

Based on my results, I found that:

1. Liverwort-specific MpMIR11737a, MpMIR11887, and MpMIR11796 genes represent independent transcriptional units, despite sharing sequence overlap with protein-coding genes. In the case of MpMIR11887, the mature MpmiR11887 is encoded within the 3'-UTR of a protein-coding gene containing a short ORF of unknown function, which might encode a miPeP. Inactivation of this ORF using CRISPR/Cas9 may clarify the potential role of the short peptide.
2. Liverwort-specific MpmiR11737a and MpmiR11737b exhibit different accumulation patterns across Marchantia vegetative and reproductive organs. With degradome data, I identified Mp1g15010 gene (encoding an uncharacterized protein) as a putative target for MpmiR11737a/b. Both miRNAs may possibly target the same mRNA, as they differ by only 2nt. Further studies are needed to prove or reject this hypothesis. Preparation of KO and over-expression (OE) mutant plants of these miRNAs is required to explain their role in Marchantia life cycle.
3. MpmiR11865* and MpmiR11865 represent the first example in liverworts, where both miRNA and miRNA* are probably active functional molecules, having opposite accumulation patterns and therefore have separate targets. Further studies are required to confirm their target identities as well as the role of both miRNAs.
4. Liverwort-specific MpmiR11887 showed exclusive accumulation in antheridiophores. CRISPR-Cas9 generated $\Delta mpmir11887^{ko}$ plants showed morphological differences in antheridiophore development. Detailed morphological observations are required for analyzing the motility pattern and flagellar beating of sperms from $\Delta mpmir11887^{ko}$ plants. It would also be valuable to perform crosses between both male and female $\Delta mpmir11887^{ko}$ plants to identify any defects in sporophyte development.
5. The putative targets for MpmiR11887 identified from degradome data include MpATX1 and Mp4g22015. Both mRNA targets show upregulated expression levels in $\Delta mpmir11887^{ko}$ plants. Further studies, including KO and OE of putative targets are needed to assess whether MpmiR11887 targets both mRNAs or one of them.

6. Liverwort-specific MpmiR11796 showed miRNA accumulation predominantly in archegoniophores. CRISPR/Cas9 generated female $\Delta mpmir11796^{ko}$ plants showed various abnormalities, of which the most spectacular include reduced ventral scales on female vegetative thallus, reduced size of archegonial receptacles, reduced length of archegoniophore stalks, and severe impairment in gametangiophore production and gamete development after fertilization.
7. Several putative mRNA targets for MpmiR11796 were identified from degradome and transcriptomic data analyses. All mRNA targets show upregulated expression levels in $\Delta mpmir11796^{ko}$ plants. Further studies are needed to establish a comprehensive miRNA-target mRNA module.
8. The following questions remain intriguing to fully understand the biological significance of MpmiR11796:
 - a. Is MpmiR11796 involved in rhizoid development?
 - b. Does MpmiR11796 affect the microtubule cytoskeleton?
 - c. Does the expression of plant cell death effector genes during pegged rhizoid development differ in $\Delta mpmir11796^{ko}$ plants as compared to WT?
 - d. Is MpmiR11796 involved in regulating egg cell expansion and the establishment of growth polarity during asymmetric divisions, thereby enabling proper egg cell development for fertilization?

Literature cited

- Achard, P., Herr, A., Baulcombe, D. C., & Harberd, N. P. (2004). Modulation of floral development by a gibberellin-regulated microRNA. *Development*, 131(14), 3357-3365. <https://doi.org/10.1242/dev.01206>
- Addo-Quaye, C., Snyder, J. A., Park, Y. B., Li, Y. F., Sunkar, R., & Axtell, M. J. (2009). Sliced microRNA targets and precise loop-first processing of MIR319 hairpins revealed by analysis of the *Physcomitrella patens* degradome. *RNA*, 15(12), 2112-2121. <https://doi.org/10.1261/rna.1774909>
- Aggarwal, B., Karłowski, W. M., Nuc, P., Jarmolowski, A., Szweykowska-Kulinska, Z., & Pietrykowska, H. (2024). MiRNAs differentially expressed in vegetative and reproductive organs of *Marchantia polymorpha*—insights into their expression pattern, gene structures and function. *RNA biology*, 21(1), 1-12.
- Alaba, S., Piszczalka, P., Pietrykowska, H., Pacak, A. M., Sierocka, I., Nuc, P. W., Singh, K., Plewka, P., Sulkowska, A., Jarmolowski, A., Karłowski, W. M., & Szweykowska-Kulinska, Z. (2015). The liverwort *Pellia endiviifolia* shares microtranscriptomic traits that are common to green algae and land plants. *New Phytol*, 206(1), 352-367. <https://doi.org/10.1111/nph.13220>
- Alonso-Peral, M. M., Li, J., Li, Y., Allen, R. S., Schnippenkoetter, W., Ohms, S., White, R. G., & Millar, A. A. (2010). The microRNA159-regulated GAMYB-like genes inhibit growth and promote programmed cell death in *Arabidopsis*. *Plant Physiol*, 154(2), 757-771. <https://doi.org/10.1104/pp.110.160630>
- Arif, M. A., Frank, W., & Khraiweh, B. (2013). Role of RNA interference (RNAi) in the Moss *Physcomitrella patens*. *Int J Mol Sci*, 14(1), 1516-1540. <https://doi.org/10.3390/ijms14011516>
- Attrill, S. T., & Dolan, L. (2024). KATANIN-mediated microtubule severing is required for MTOC organisation and function in *Marchantia polymorpha*. *Development*, 151(20).
- Aukerman, M. J., & Sakai, H. (2003). Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2-Like Target Genes. *The Plant Cell*, 15(11), 2730-2741. <https://doi.org/10.1105/tpc.016238>
- Axtell, M. J., Jan, C., Rajagopalan, R., & Bartel, D. P. (2006). A two-hit trigger for siRNA biogenesis in plants. *Cell*, 127(3), 565-577. <https://doi.org/10.1016/j.cell.2006.09.032>
- Axtell, M. J., Snyder, J. A., & Bartel, D. P. (2007). Common functions for diverse small RNAs of land plants. *Plant Cell*, 19(6), 1750-1769. <https://doi.org/10.1105/tpc.107.051706>
- Bajczyk, M., Jarmolowski, A., Jozwiak, M., Pacak, A., Pietrykowska, H., Sierocka, I., Swida-Barteczka, A., Szewc, L., & Szweykowska-Kulinska, Z. (2023). Recent Insights into Plant miRNA Biogenesis: Multiple Layers of miRNA Level Regulation. *Plants (Basel)*, 12(2). <https://doi.org/10.3390/plants12020342>
- Baker, C. C., Sieber, P., Wellmer, F., & Meyerowitz, E. M. (2005). The *early extra petals1* Mutant Uncovers a Role for MicroRNA *miR164c* in Regulating Petal Number in *Arabidopsis*. *Current Biology*, 15(4), 303-315. <https://doi.org/10.1016/j.cub.2005.02.017>
- Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., DePamphilis, C., Albert, V. A., Aono, N., Aoyama, T., & Ambrose, B. A. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science*, 332(6032), 960-963.
- Bao, H., Sun, R., Iwano, M., Yoshitake, Y., Aki, S. S., Umeda, M., Nishihama, R., Yamaoka, S., & Kohchi, T. (2024). Conserved CK1-mediated signaling is required for female

- germline specification in *Marchantia polymorpha*. *Current Biology*, 34(6), 1324-1332. e1326.
- Barciszewska-Pacak, M., Milanowska, K., Knop, K., Bielewicz, D., Nuc, P., Plewka, P., Pacak, A. M., Vazquez, F., Karlowski, W., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2015). Arabidopsis microRNA expression regulation in a wide range of abiotic stress responses. *Frontiers in Plant Science*, 6, 410. <https://doi.org/10.3389/fpls.2015.00410>
- Bartel, B., & Bartel, D. P. (2003). MicroRNAs: at the root of plant development? *Plant physiology*, 132(2), 709-717.
- Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116(2), 281-297. [https://doi.org/10.1016/s0092-8674\(04\)00045-5](https://doi.org/10.1016/s0092-8674(04)00045-5)
- Bartel, D. P. (2018). Metazoan micrnas. *Cell*, 173(1), 20-51.
- Bibikova, T. N., Blancaflor, E. B., & Gilroy, S. (1999). Microtubules regulate tip growth and orientation in root hairs of *Arabidopsis thaliana*. *The Plant Journal*, 17(6), 657-665.
- Bielewicz, D., Dolata, J., Bajczyk, M., Szewc, L., Gulanicz, T., Bhat, S. S., Karlik, A., Jozwiak, M., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2023). Hyponastic Leaves 1 Interacts with RNA Pol II to Ensure Proper Transcription of MicroRNA Genes. *Plant and Cell Physiology*, 64(6), 571-582. <https://doi.org/10.1093/pcp/pcad032>
- Bielewicz, D., Dolata, J., Zielezinski, A., Alaba, S., Szarynska, B., Szczesniak, M. W., Jarmolowski, A., Szweykowska-Kulinska, Z., & Karlowski, W. M. (2012). mirEX: a platform for comparative exploration of plant pri-miRNA expression data. *Nucleic Acids Res*, 40(Database issue), D191-197. <https://doi.org/10.1093/nar/gkr878>
- Bologna, N. G., Iselin, R., Abriata, L. A., Sarazin, A., Pumplin, N., Jay, F., Grentzinger, T., Dal Peraro, M., & Voinnet, O. (2018). Nucleo-cytosolic Shuttling of ARGONAUTE1 Prompts a Revised Model of the Plant MicroRNA Pathway. *Mol Cell*, 69(4), 709-719 e705. <https://doi.org/10.1016/j.molcel.2018.01.007>
- Bologna, N. G., Mateos, J. L., Bresso, E. G., & Palatnik, J. F. (2009). A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. *EMBO J*, 28(23), 3646-3656. <https://doi.org/10.1038/emboj.2009.292>
- Bologna, N. G., & Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in *Arabidopsis*. *Annu Rev Plant Biol*, 65, 473-503. <https://doi.org/10.1146/annurev-arplant-050213-035728>
- Bowman, J. L. (2016). A brief history of *Marchantia* from Greece to genomics. *Plant and Cell Physiology*, 57(2), 210-229.
- Bowman, J. L. (2022). The liverwort *Marchantia polymorpha*, a model for all ages. *Curr Top Dev Biol*, 147, 1-32. <https://doi.org/10.1016/bs.ctdb.2021.12.009>
- Bowman, J. L., Araki, T., & Kohchi, T. (2016). *Marchantia*: past, present and future. In (Vol. 57, pp. 205-209): Oxford University Press.
- Bowman, J. L., Arteaga-Vazquez, M., Berger, F., Briginshaw, L. N., Carella, P., Aguilar-Cruz, A., Davies, K. M., Dierschke, T., Dolan, L., & Dorantes-Acosta, A. E. (2022). The renaissance and enlightenment of *Marchantia* as a model system. *The Plant Cell*, 34(10), 3512-3542.
- Bowman, J. L., Floyd, S. K., & Sakakibara, K. (2007). Green genes—comparative genomics of the green branch of life. *Cell*, 129(2), 229-234.
- Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S. Q., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S. S., Althoff, F., Araki, T., Arteaga-Vazquez, M. A., Balasubramanian, S., Barry, K., Bauer, D., Boehm, C. R., . . . Schmutz, J. (2017). Insights into Land Plant Evolution Garnered from the *Marchantia polymorpha* Genome. *Cell*, 171(2), 287-+. <https://doi.org/10.1016/j.cell.2017.09.030>
- Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., & Voinnet, O. (2008). Widespread Translational Inhibition by Plant

- miRNAs and siRNAs. *Science*, 320(5880), 1185-1190. <https://doi.org/doi:10.1126/science.1159151>
- Brown, R. C., & Lemmon, B. E. (2011). Dividing without centrioles: innovative plant microtubule organizing centres organize mitotic spindles in bryophytes, the earliest extant lineages of land plants. *AoB plants*, 2011, plr028.
- Buschmann, H., Holtmannspotter, M., Borchers, A., O'Donoghue, M. T., & Zachgo, S. (2016). Microtubule dynamics of the centrosome-like polar organizers from the basal land plant *Marchantia polymorpha*. *New Phytol*, 209(3), 999-1013. <https://doi.org/10.1111/nph.13691>
- Cao, M.-X., Li, S.-Z., & Li, H.-J. (2024). MpMLO1 controls sperm discharge in liverwort. *Nature Plants*, 10(6), 1027-1038.
- Chen, Q.-j., Zhang, L.-p., Song, S.-r., Wang, L., Xu, W.-p., Zhang, C.-x., Wang, S.-p., Liu, H.-f., & Ma, C. (2022). vvi-miPEP172b and vvi-miPEP3635b increase cold tolerance of grapevine by regulating the corresponding MIRNA genes. *Plant Science*, 325, 111450.
- Chojnacka, A., Smoczynska, A., Bielewicz, D., Pacak, A., Hensel, G., Kumlehn, J., Maciej Karłowski, W., Grabsztunowicz, M., Sobieszczuk-Nowicka, E., & Jarmolowski, A. (2023). PEP444c encoded within the MIR444c gene regulates microRNA444c accumulation in barley. *Physiologia Plantarum*, 175(5), e14018.
- Chomczynski, P., & Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, 162(1), 156-159. [https://doi.org/https://doi.org/10.1016/0003-2697\(87\)90021-2](https://doi.org/https://doi.org/10.1016/0003-2697(87)90021-2)
- Couzigou, J.-M., André, O., Guillotin, B., Alexandre, M., & Combier, J.-P. (2016). Use of microRNA-encoded peptide miPEP172c to stimulate nodulation in soybean. *New Phytologist*, 211(2), 379-381.
- Csicsely, E., Oberender, A., Georgiadou, A.-S., Alz, J., Kiel, S., Gutsche, N., Zachgo, S., Grünert, J., Klingl, A., Top, O., & Frank, W. (2025). Identification and characterization of DICER-LIKE genes and their roles in *Marchantia polymorpha* development and salt stress response. *The Plant Journal*, 121(3), e17236. <https://doi.org/https://doi.org/10.1111/tpj.17236>
- De Smet, I., & Beeckman, T. (2011). Asymmetric cell division in land plants and algae: the driving force for differentiation. *Nature Reviews Molecular Cell Biology*, 12(3), 177-188.
- Delaux, P.-M., Hetherington, A. J., Coudert, Y., Delwiche, C., Dunand, C., Gould, S., Kenrick, P., Li, F.-W., Philippe, H., & Rensing, S. A. (2019). Reconstructing trait evolution in plant evo–devo studies. *Current Biology*, 29(21), R1110-R1118.
- Driouch, A., Follet-Gueye, M.-L., Bernard, S., Kousar, S., Chevalier, L., Vitré-Gibouin, M., & Lerouxel, O. (2012). Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Frontiers in Plant Science*, 3, 79.
- Duckett, J. G., Ligrone, R., Renzaglia, K. S., & Pressel, S. (2014). Pegged and smooth rhizoids in complex thalloid liverworts (Marchantiopsida): structure, function and evolution. *Botanical Journal of the Linnean Society*, 174(1), 68-92.
- Duckett, J. G., Pressel, S., & Kowal, J. (2024). The biology of *Marchantia polymorpha* subsp. *ruderalis* Bischl. & Boissel. Dub in nature. *Frontiers in Plant Science*, 15, 1339832.
- Durand, E. J. (1908). The development of the sexual organs and sporogonium of *Marchantia polymorpha*. *Bulletin of the Torrey Botanical Club*, 35(7), 321-335.
- Eklund, D. M., Ishizaki, K., Flores-Sandoval, E., Kikuchi, S., Takebayashi, Y., Tsukamoto, S., Hirakawa, Y., Nonomura, M., Kato, H., & Kouno, M. (2015). Auxin produced by the indole-3-pyruvic acid pathway regulates development and gemmae dormancy in the liverwort *Marchantia polymorpha*. *The Plant Cell*, 27(6), 1650-1669.
- Fang, J., Morsalin, S., Rao, V. N., & Reddy, E. S. P. (2017). Decoding of non-coding DNA and non-coding RNA: pri-micro RNA-encoded novel peptides regulate migration of cancer cells. *Journal of Pharmaceutical Sciences and Pharmacology*, 3(1), 23-27.

- Fattash, I., Voß, B., Reski, R., Hess, W. R., & Frank, W. (2007). Evidence for the rapid expansion of microRNA-mediated regulation in early land plant evolution. *BMC Plant Biology*, 7, 1-19.
- Feldberg, K., Schneider, H., Stadler, T., Schäfer-Verwimp, A., Schmidt, A. R., & Heinrichs, J. (2014). Epiphytic leafy liverworts diversified in angiosperm-dominated forests. *Scientific Reports*, 4(1), 5974.
- Flores-Sandoval, E., Dierschke, T., Fisher, T. J., & Bowman, J. L. (2016). Efficient and inducible use of artificial microRNAs in *Marchantia polymorpha*. *Plant and Cell Physiology*, 57(2), 281-290.
- Flores-Sandoval, E., Romani, F., & Bowman, J. L. (2018)a. Co-expression and transcriptome analysis of *Marchantia polymorpha* transcription factors supports class C ARFs as independent actors of an ancient auxin regulatory module. *Frontiers in Plant Science*, 9, 1345.
- Flores-Sandoval, E., Eklund, D. M., Hong, S. F., Alvarez, J. P., Fisher, T. J., Lampugnani, E. R., Golz, J. F., Vázquez-Lobo, A., Dierschke, T., & Lin, S. S. (2018)b. Class C ARF s evolved before the origin of land plants and antagonize differentiation and developmental transitions in *Marchantia polymorpha*. *New Phytologist*, 218(4), 1612-1630.
- Floyd, S. K., & Bowman, J. L. (2004). Gene regulation: ancient microRNA target sequences in plants. *Nature*, 428(6982), 485-486. <https://doi.org/10.1038/428485a>
- Floyd, S. K., & Bowman, J. L. (2007). The ancestral developmental tool kit of land plants. *International journal of plant sciences*, 168(1), 1-35.
- Fujisawa, M., Hayashi, K., Nishio, T., Bando, T., Okada, S., Yamato, K. T., Fukuzawa, H., & Ohyama, K. (2001). Isolation of X and Y chromosome-specific DNA markers from a liverwort, *Marchantia polymorpha*, by representational difference analysis. *Genetics*, 159(3), 981-985. <https://doi.org/10.1093/genetics/159.3.981>
- Futagami, K., Tsuzuki, M., Yoshida, M., & Watanabe, Y. (2025). MpmiR319 promotes gemma/gemma cup formation in the liverwort *Marchantia polymorpha*. *Journal of Experimental Botany*. <https://doi.org/10.1093/jxb/eraf148>
- Gamborg, O. L., Murashige, T., Thorpe, T. A., & Vasil, I. K. (1976). Plant Tissue Culture Media. *In Vitro*, 12(7), 473-478. <http://www.jstor.org/stable/20170323>
- Geldner, N., Hyman, D. L., Wang, X., Schumacher, K., & Chory, J. (2007). Endosomal signaling of plant steroid receptor kinase BRI1. *Genes & Development*, 21(13), 1598-1602.
- German, M. A., Luo, S., Schroth, G., Meyers, B. C., & Green, P. J. (2009). Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat Protoc*, 4(3), 356-362. <https://doi.org/10.1038/nprot.2009.8>
- Gong, W., Bak, D. T., Wendrich, J. R., Weijers, D., & Laux, T. (2024). CDC48A, an interactor of WOX2, is required for embryonic patterning in *Arabidopsis thaliana*. *Plant Cell Reports*, 43(7), 174. <https://doi.org/10.1007/s00299-024-03158-2>
- Han, M.-H., Goud, S., Song, L., & Fedoroff, N. (2004). The *Arabidopsis* double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. *Proceedings of the National Academy of Sciences*, 101(4), 1093-1098. <https://doi.org/doi:10.1073/pnas.0307969100>
- Higo, A., Kawashima, T., Borg, M., Zhao, M., López-Vidriero, I., Sakayama, H., Montgomery, S. A., Sekimoto, H., Hackenberg, D., & Shimamura, M. (2018). Transcription factor DUO1 generated by neo-functionalization is associated with evolution of sperm differentiation in plants. *Nature communications*, 9(1), 5283.
- Higo, A., Niwa, M., Yamato, K. T., Yamada, L., Sawada, H., Sakamoto, T., Kurata, T., Shirakawa, M., Endo, M., & Shigenobu, S. (2016). Transcriptional framework of male

- gametogenesis in the liverwort *Marchantia polymorpha* L. *Plant and Cell Physiology*, 57(2), 325-338.
- Hofacker, I. L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Research*, 31(13), 3429-3431. <https://doi.org/10.1093/nar/gkg599>
- Hong, S.-F., Wei, W.-L., Pan, Z.-J., Yu, J.-Z., Cheng, S., Hung, Y.-L., Tjita, V., Wang, H.-C., Komatsu, A., Nishihama, R., Kohchi, T., Chen, H.-M., Chen, W.-C., Lo, J.-C., Chiu, Y.-H., Yang, H.-C., Lu, M.-Y., Liu, L.-Y. D., & Lin, S.-S. (2024). Molecular Insights into MpAGO1 and Its Regulatory miRNA, miR11707, in the High-Temperature Acclimation of *Marchantia polymorpha*. *Plant and Cell Physiology*, 65(9), 1414-1433. <https://doi.org/10.1093/pcp/pcae080>
- Honkanen, S., Jones, V. A., Morieri, G., Champion, C., Hetherington, A. J., Kelly, S., Proust, H., Saint-Marcoux, D., Prescott, H., & Dolan, L. (2016). The mechanism forming the cell surface of tip-growing rooting cells is conserved among land plants. *Current Biology*, 26(23), 3238-3244.
- Honkanen, S., Thamm, A., Arteaga-Vazquez, M. A., & Dolan, L. (2018). Negative regulation of conserved RSL class I bHLH transcription factors evolved independently among land plants. *Elife*, 7. <https://doi.org/10.7554/eLife.38529>
- Huijser, P., & Schmid, M. (2011). The control of developmental phase transitions in plants. *Development*, 138(19), 4117-4129. <https://doi.org/10.1242/dev.063511>
- Hung, Y.-L., Hong, S.-F., Wei, W.-L., Cheng, S., Yu, J.-Z., Tjita, V., Yong, Q.-Y., Nishihama, R., Kohchi, T., & Bowman, J. L. (2024). Dual regulation of cytochrome P450 gene expression by two distinct small RNAs, a novel tasiRNA and miRNA, in *Marchantia polymorpha*. *Plant and Cell Physiology*, 65(7), 1115-1134.
- Inoue, K., Nishihama, R., Araki, T., & Kohchi, T. (2019). Reproductive Induction is a Far-Red High Irradiance Response that is Mediated by Phytochrome and PHYTOCHROME INTERACTING FACTOR in *Marchantia polymorpha*. *Plant and Cell Physiology*, 60(5), 1136-1145. <https://doi.org/10.1093/pcp/pcz029>
- Ishizaki, K., Chiyoda, S., Yamato, K. T., & Kohchi, T. (2008). Agrobacterium-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant and Cell Physiology*, 49(7), 1084-1091.
- Ishizaki, K., Nishihama, R., Ueda, M., Inoue, K., Ishida, S., Nishimura, Y., Shikanai, T., & Kohchi, T. (2015). Development of gateway binary vector series with four different selection markers for the liverwort *Marchantia polymorpha*. *PLoS One*, 10(9), e0138876.
- Iwakawa, H.-o., & Tomari, Y. (2013). Molecular Insights into microRNA-Mediated Translational Repression in Plants. *Molecular Cell*, 52(4), 591-601. <https://doi.org/10.1016/j.molcel.2013.10.033>
- Iwakawa, H., Melkonian, K., Schlüter, T., Jeon, H.-W., Nishihama, R., Motose, H., & Nakagami, H. (2021). Agrobacterium-mediated transient transformation of *Marchantia* liverworts. *Plant and Cell Physiology*, 62(11), 1718-1727.
- Jones-Rhoades, M. W., & Bartel, D. P. (2004). Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell*, 14(6), 787-799. <https://doi.org/10.1016/j.molcel.2004.05.027>
- Jozwiak, M., Bielewicz, D., Szweykowska-Kulinska, Z., Jarmolowski, A., & Bajczyk, M. (2023). SERRATE: a key factor in coordinated RNA processing in plants. *Trends in Plant Science*, 28(7), 841-853. <https://doi.org/10.1016/j.tplants.2023.03.009>
- Kawamura, S., Romani, F., Yagura, M., Mochizuki, T., Sakamoto, M., Yamaoka, S., Nishihama, R., Nakamura, Y., Yamato, K. T., Bowman, J. L., Kohchi, T., & Tanizawa, Y. (2022). MarpolBase Expression: A Web-Based, Comprehensive Platform for Visualization and Analysis of Transcriptomes in the Liverwort *Marchantia polymorpha*. *Plant and Cell Physiology*, 63(11), 1745-1755. <https://doi.org/10.1093/pcp/pcac129>

- Ketelaar, T., Galway, M., Mulder, B., & Emons, A. (2008). Rates of exocytosis and endocytosis in *Arabidopsis* root hairs and pollen tubes. *Journal of Microscopy*, 231(2), 265-273.
- Kim, Y. J., Zheng, B., Yu, Y., Won, S. Y., Mo, B., & Chen, X. (2011). The role of Mediator in small and long noncoding RNA production in *Arabidopsis thaliana*. *The EMBO journal*, 30(5), 814-822. <https://doi.org/10.1038/emboj.2011.3>
- Klöpfer, T. H., Kienle, N., Fasshauer, D., & Munro, S. (2012). Untangling the evolution of Rab G proteins: implications of a comprehensive genomic analysis. *BMC biology*, 10, 1-17.
- Knop, K., Stepień, A., Barciszewska-Pacak, M., Taube, M., Bielewicz, D., Michalak, M., Borst, J. W., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2017). Active 5' splice sites regulate the biogenesis efficiency of *Arabidopsis* microRNAs derived from intron-containing genes. *Nucleic Acids Res*, 45(5), 2757-2775. <https://doi.org/10.1093/nar/gkw895>
- Kohchi, T., Yamato, K. T., Ishizaki, K., Yamaoka, S., & Nishihama, R. (2021). Development and Molecular Genetics of *Marchantia polymorpha*. *Annu Rev Plant Biol*, 72, 677-702. <https://doi.org/10.1146/annurev-arplant-082520-094256>
- Krasnikova, M. S., Goryunov, D. V., Troitsky, A. V., Solovyev, A. G., Ozerova, L. V., & Morozov, S. Y. (2013). Peculiar evolutionary history of miR390-guided TAS3-like genes in land plants. *ScientificWorldJournal*, 2013, 924153. <https://doi.org/10.1155/2013/924153>
- Kubota, A., Ishizaki, K., Hosaka, M., & Kohchi, T. (2013). Efficient *Agrobacterium*-mediated transformation of the liverwort *Marchantia polymorpha* using regenerating thalli. *Bioscience, biotechnology, and biochemistry*, 77(1), 167-172.
- Kurihara, Y., & Watanabe, Y. (2004). *Arabidopsis* micro-RNA biogenesis through Dicer-like 1 protein functions. *Proc Natl Acad Sci U S A*, 101(34), 12753-12758. <https://doi.org/10.1073/pnas.0403115101>
- Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., Piednoel, M., Gundlach, H., Van Bel, M., & Meyberg, R. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal*, 93(3), 515-533.
- Lauressergues, D., Couzigou, J. M., Clemente, H. S., Martinez, Y., Dunand, C., Becard, G., & Combier, J. P. (2015). Primary transcripts of microRNAs encode regulatory peptides. *Nature*, 520(7545), 90-93. <https://doi.org/10.1038/nature14346>
- Lauressergues, D., Ormancey, M., Guillotin, B., San Clemente, H., Camborde, L., Duboe, C., Tourneur, S., Charpentier, P., Barozet, A., Jauneau, A., Le Ru, A., Thuleau, P., Gervais, V., Plaza, S., & Combier, J. P. (2022). Characterization of plant microRNA-encoded peptides (miPEPs) reveals molecular mechanisms from the translation to activity and specificity. *Cell Rep*, 38(6), 110339. <https://doi.org/10.1016/j.celrep.2022.110339>
- Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, 75(5), 843-854.
- Li, Y., Li, C., Xia, J., & Jin, Y. (2011). Domestication of transposable elements into microRNA genes in plants. *PLoS One*, 6(5), e19212.
- Ligrone, R., Duckett, J. G., & Renzaglia, K. S. (2012). Major transitions in the evolution of early land plants: a bryological perspective. *Annals of botany*, 109(5), 851-871.
- Lin, P. C., Lu, C. W., Shen, B. N., Lee, G. Z., Bowman, J. L., Arteaga-Vazquez, M. A., Liu, L. Y. D., Hong, S. F., Lo, C. F., Su, G. M., Kohchi, T., Ishizaki, K., Zachgo, S., Althoff, F., Takenaka, M., Yamato, K. T., & Lin, S. S. (2016). Identification of miRNAs and Their Targets in the Liverwort *Marchantia polymorpha* by Integrating RNA-Seq and Degradome Analyses. *Plant and Cell Physiology*, 57(2), 339-358. <https://doi.org/10.1093/pcp/pcw020>
- Lin, S. S., & Bowman, J. L. (2018). MicroRNAs in *Marchantia polymorpha*. *New Phytol*, 220(2), 409-416. <https://doi.org/10.1111/nph.15294>

- Lipka, V., Kwon, C., & Panstruga, R. (2007). SNARE-ware: the role of SNARE-domain proteins in plant biology. *Annu. Rev. Cell Dev. Biol.*, 23(1), 147-174.
- Liu, J., & Vance, C. P. (2010). Crucial roles of sucrose and microRNA399 in systemic signaling of P deficiency: a tale of two team players? *Plant Signaling & Behavior*, 5(12), 1556-1560. <https://doi.org/10.4161/psb.5.12.13293>
- Llave, C., Kasschau, K. D., Rector, M. A., & Carrington, J. C. (2002). Endogenous and Silencing-Associated Small RNAs in Plants[W]. *The Plant Cell*, 14(7), 1605-1619. <https://doi.org/10.1105/tpc.003210>
- Llave, C., Xie, Z., Kasschau, K. D., & Carrington, J. C. (2002). Cleavage of <i>Scarecrow-like</i> mRNA Targets Directed by a Class of <i>Arabidopsis</i> miRNA. *Science*, 297(5589), 2053-2056. <https://doi.org/doi:10.1126/science.1076311>
- Lu, Y.-T., Loue-Manifel, J., Bollier, N., Gadiant, P., De Winter, F., Carella, P., Huguin, A., Grey, S., Marnas, H., & Simon, F. (2024). Convergent evolution of water-conducting cells in *Marchantia* recruited the ZHOUP1 gene promoting cell wall reinforcement and programmed cell death. *Current Biology*, 34(4), 793-807. e797.
- Mallory, A. C., Bartel, D. P., & Bartel, B. (2005). MicroRNA-directed regulation of *Arabidopsis* AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell*, 17(5), 1360-1375. <https://doi.org/10.1105/tpc.105.031716>
- Manavella, P. A., Koenig, D., Rubio-Somoza, I., Burbano, H. A., Becker, C., & Weigel, D. (2013). Tissue-specific silencing of *Arabidopsis* SU (VAR) 3-9 HOMOLOG8 by miR171a. *Plant physiology*, 161(2), 805-812.
- Mano, S., Nishihama, R., Ishida, S., Hikino, K., Kondo, M., Nishimura, M., Yamato, K. T., Kohchi, T., & Nakagawa, T. (2018). Novel gateway binary vectors for rapid tripartite DNA assembly and promoter analysis with various reporters and tags in the liverwort *Marchantia polymorpha*. *PLoS One*, 13(10), e0204964.
- Mase, H., Sumiura, A., Yoshitake, Y., Kohchi, T., Takahashi, T., & Motose, H. (2025). Involvement of a NIMA-related kinase in cell division of the liverwort *Marchantia polymorpha*. *Plant and Cell Physiology*, pcaf021.
- Mecchia, M. A., Rövekamp, M., Giraldo-Fonseca, A., Meier, D., Gadiant, P., Vogler, H., Limacher, D., Bowman, J. L., & Grossniklaus, U. (2022). The single *Marchantia polymorpha* FERONIA homolog reveals an ancestral role in regulating cellular expansion and integrity. *Development*, 149(19), dev200580.
- Minamino, N., Kanazawa, T., Era, A., Ebine, K., Nakano, A., & Ueda, T. (2018). RAB GTPases in the basal land plant *Marchantia polymorpha*. *Plant and Cell Physiology*, 59(4), 850-861.
- Montgomery, S. A., Tanizawa, Y., Galik, B., Wang, N., Ito, T., Mochizuki, T., Akimcheva, S., Bowman, J. L., Cognat, V., Marechal-Drouard, L., Ekker, H., Hong, S. F., Kohchi, T., Lin, S. S., Liu, L. D., Nakamura, Y., Valeeva, L. R., Shakirov, E. V., Shippen, D. E., . . . Berger, F. (2020). Chromatin Organization in Early Land Plants Reveals an Ancestral Association between H3K27me3, Transposons, and Constitutive Heterochromatin. *Curr Biol*, 30(4), 573-588 e577. <https://doi.org/10.1016/j.cub.2019.12.015>
- Moody, L. A. (2020). Three-dimensional growth: a developmental innovation that facilitated plant terrestrialization. *Journal of Plant Research*, 133(3), 283-290.
- Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C. H., Yang, Z., Schneider, H., & Donoghue, P. C. J. (2018). The timescale of early land plant evolution. *Proc Natl Acad Sci U S A*, 115(10), E2274-E2283. <https://doi.org/10.1073/pnas.1719588115>
- Motose, H., Tominaga, R., Wada, T., Sugiyama, M., & Watanabe, Y. (2008). A NIMA-related protein kinase suppresses ectopic outgrowth of epidermal cells through its kinase activity and the association with microtubules. *The Plant Journal*, 54(5), 829-844.

- Naito, Y., Hino, K., Bono, H., & Ui-Tei, K. (2015). CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics*, 31(7), 1120-1123. <https://doi.org/10.1093/bioinformatics/btu743>
- Nakamura, K., Hisanaga, T., Fujimoto, K., Nakajima, K., & Wada, H. (2018). Plant-inspired pipettes. *Journal of The Royal Society Interface*, 15(140), 20170868.
- Naramoto, S., Hata, Y., Fujita, T., & Kyojuka, J. (2022). The bryophytes *Physcomitrium patens* and *Marchantia polymorpha* as model systems for studying evolutionary cell and developmental biology in plants. *The Plant Cell*, 34(1), 228-246.
- Nielsen, E., Cheung, A. Y., & Ueda, T. (2008). The regulatory RAB and ARF GTPases for vesicular trafficking. *Plant physiology*, 147(4), 1516-1526.
- Nishihama, R., Ishizaki, K., Hosaka, M., Matsuda, Y., Kubota, A., & Kohchi, T. (2015). Phytochrome-mediated regulation of cell division and growth during regeneration and sporeling development in the liverwort *Marchantia polymorpha*. *Journal of Plant Research*, 128, 407-421.
- Nodine, M. D., & Bartel, D. P. (2010). MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. *Genes Dev*, 24(23), 2678-2692. <https://doi.org/10.1101/gad.1986710>
- Nosaka, M., Itoh, J.-I., Nagato, Y., Ono, A., Ishiwata, A., & Sato, Y. (2012). Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice.
- Noshay, J. M., Marand, A. P., Anderson, S. N., Zhou, P., Mejia Guerra, M. K., Lu, Z., O'Connor, C. H., Crisp, P. A., Hirsch, C. N., & Schmitz, R. J. (2021). Assessing the regulatory potential of transposable elements using chromatin accessibility profiles of maize transposons. *Genetics*, 217(1), 1-13.
- Oliver, M. J., Velten, J., & Mishler, B. D. (2005). Desiccation tolerance in bryophytes: a reflection of the primitive strategy for plant survival in dehydrating habitats? *Integrative and comparative biology*, 45(5), 788-799.
- Olsen, P. H., & Ambros, V. (1999). The *lin-4* regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Developmental Biology*, 216(2), 671-680.
- Otani, K., Ishizaki, K., Nishihama, R., Takatani, S., Kohchi, T., Takahashi, T., & Motose, H. (2018). An evolutionarily conserved NIMA-related kinase directs rhizoid tip growth in the basal land plant *Marchantia polymorpha*. *Development*, 145(5), dev154617.
- Otegui, M. S., & Spitzer, C. (2008). Endosomal functions in plants. *Traffic*, 9(10), 1589-1598.
- Ovečka, M., Illés, P., Lichtscheidl, I., Derksen, J., & Šamaj, J. (2012). Endocytosis and vesicular recycling in root hairs and pollen tubes. *Endocytosis in Plants*, 81-106.
- Palatnik, J. F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J. C., & Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. *Nature*, 425(6955), 257-263. <https://doi.org/10.1038/nature01958>
- Park, W., Li, J., Song, R., Messing, J., & Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Curr Biol*, 12(17), 1484-1495. [https://doi.org/10.1016/s0960-9822\(02\)01017-5](https://doi.org/10.1016/s0960-9822(02)01017-5)
- Pellinen, T., Arjonen, A., Vuoriluoto, K., Kallio, K., Fransen, J. A., & Ivaska, J. (2006). Small GTPase Rab21 regulates cell adhesion and controls endosomal traffic of β 1-integrins. *Journal of Cell Biology*, 173(5), 767-780.
- Pietrykowska, H., Alisha, A., Aggarwal, B., Watanabe, Y., Ohtani, M., Jarmolowski, A., Sierocka, I., & Szweykowska-Kulinska, Z. (2023). Conserved and non-conserved RNA-target modules in plants: lessons for a better understanding of *Marchantia* development. *Plant Molecular Biology*, 113(4), 121-142. <https://doi.org/10.1007/s11103-023-01392-y>

- Pietrykowska, H., Sierocka, I., Zielezinski, A., Alisha, A., Carrasco-Sanchez, J. C., Jarmolowski, A., Karlowski, W. M., & Szweykowska-Kulinska, Z. (2022). Biogenesis, conservation, and function of miRNA in liverworts. *J Exp Bot*, 73(13), 4528-4545. <https://doi.org/10.1093/jxb/erac098>
- Pires, N. D., & Dolan, L. (2012). Morphological evolution in land plants: new designs with old genes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1588), 508-518.
- Pires, N. D., Yi, K., Breuninger, H., Catarino, B., Menand, B., & Dolan, L. (2013). Recruitment and remodeling of an ancient gene regulatory network during land plant evolution. *Proceedings of the National Academy of Sciences*, 110(23), 9571-9576. <https://doi.org/doi:10.1073/pnas.1305457110>
- Plotnikova, A., Baranauskė, S., Osipenko, A., Klimašauskas, S., & Vilkaitis, G. (2013). Mechanistic insights into small RNA recognition and modification by the HEN1 methyltransferase. *Biochemical Journal*, 453(2), 281-290. <https://doi.org/10.1042/bj20121699>
- Pressel, S., & Duckett, J. G. (2019). Do motile spermatozooids limit the effectiveness of sexual reproduction in bryophytes? Not in the liverwort *Marchantia polymorpha*. *Journal of Systematics and Evolution*, 57(4), 371-381.
- Proust, H., Honkanen, S., Jones, V. A., Morieri, G., Prescott, H., Kelly, S., Ishizaki, K., Kohchi, T., & Dolan, L. (2016). RSL Class I Genes Controlled the Development of Epidermal Structures in the Common Ancestor of Land Plants. *Curr Biol*, 26(1), 93-99. <https://doi.org/10.1016/j.cub.2015.11.042>
- Puttick, M. N., Morris, J. L., Williams, T. A., Cox, C. J., Edwards, D., Kenrick, P., Pressel, S., Wellman, C. H., Schneider, H., & Pisani, D. (2018). The interrelationships of land plants and the nature of the ancestral embryophyte. *Current Biology*, 28(5), 733-745. e732.
- Rajagopalan, R., Vaucheret, H., Trejo, J., & Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev*, 20(24), 3407-3425. <https://doi.org/10.1101/gad.1476406>
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 403(6772), 901-906. <https://doi.org/10.1038/35002607>
- Reinhart, B. J., Weinstein, E. G., Rhoades, M. W., Bartel, B., & Bartel, D. P. (2002). MicroRNAs in plants. *Genes Dev*, 16(13), 1616-1626. <https://doi.org/10.1101/gad.1004402>
- Ren, X., Zhang, X., Qi, X., Zhang, T., Wang, H., Twell, D., Gong, Y., Fu, Y., Wang, B., & Kong, H. (2024). The BNB–GLID module regulates germline fate determination in *Marchantia polymorpha*. *The Plant Cell*, 36(9), 3824-3837.
- Ren, Y., Song, Y., Zhang, L., Guo, D., He, J., Wang, L., Song, S., Xu, W., Zhang, C., & Lers, A. (2021). Coding of non-coding RNA: insights into the regulatory functions of Pri-MicroRNA-encoded peptides in plants. *Frontiers in Plant Science*, 12, 641351.
- Rensing, S. A. (2018). Great moments in evolution: the conquest of land by plants. *Current opinion in plant biology*, 42, 49-54.
- Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., Nishiyama, T., Perroud, P. F., Lindquist, E. A., Kamisugi, Y., Tanahashi, T., Sakakibara, K., Fujita, T., Oishi, K., Shin, I. T., Kuroki, Y., Toyoda, A., Suzuki, Y., Hashimoto, S., . . . Boore, J. L. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science*, 319(5859), 64-69. <https://doi.org/10.1126/science.1150646>
- Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., & Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell*, 110(4), 513-520. [https://doi.org/10.1016/s0092-8674\(02\)00863-2](https://doi.org/10.1016/s0092-8674(02)00863-2)

- Rogers, K., & Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell*, 25(7), 2383-2399. <https://doi.org/10.1105/tpc.113.113159>
- Rozewicki, J., Li, S., Amada, K. M., Standley, D. M., & Katoh, K. (2019). MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Research*, 47(W1), W5-W10. <https://doi.org/10.1093/nar/gkz342>
- Saint-Marcoux, D., Proust, H., Dolan, L., & Langdale, J. A. (2015). Identification of reference genes for real-time quantitative PCR experiments in the liverwort *Marchantia polymorpha*. *PLoS One*, 10(3), e0118678. <https://doi.org/10.1371/journal.pone.0118678>
- Saito, M., Momiki, R., Ebine, K., Yoshitake, Y., Nishihama, R., Miyakawa, T., Nakano, T., Mitsuda, N., Araki, T., & Kohchi, T. (2023). A bHLH heterodimer regulates germ cell differentiation in land plant gametophytes. *Current Biology*, 33(22), 4980-4987. e4986.
- Sakai, T., Honing, H. v. d., Nishioka, M., Uehara, Y., Takahashi, M., Fujisawa, N., Saji, K., Seki, M., Shinozaki, K., & Jones, M. A. (2008). Armadillo repeat-containing kinesins and a NIMA-related kinase are required for epidermal-cell morphogenesis in *Arabidopsis*. *The Plant Journal*, 53(1), 157-171.
- Saleh, A., Alvarez-Venegas, R., Yilmaz, M., Le, O., Hou, G., Sadler, M., Al-Abdallat, A., Xia, Y., Lu, G., Ladunga, I., & Avramova, Z. (2008). The Highly Similar *Arabidopsis* Homologs of *Trithorax* ATX1 and ATX2 Encode Proteins with Divergent Biochemical Functions. *The Plant Cell*, 20(3), 568-579. <https://doi.org/10.1105/tpc.107.056614>
- Schauer, S. E., Jacobsen, S. E., Meinke, D. W., & Ray, A. (2002). DICER-LIKE1: blind men and elephants in *Arabidopsis* development. *Trends Plant Sci*, 7(11), 487-491. [https://doi.org/10.1016/s1360-1385\(02\)02355-5](https://doi.org/10.1016/s1360-1385(02)02355-5)
- Scheller, H. V., & Ulvskov, P. (2010). Hemicelluloses. *Annual review of plant biology*, 61(1), 263-289.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676-682. <https://doi.org/10.1038/nmeth.2019>
- Servais, T., Cascales-Miñana, B., Cleal, C. J., Gerrienne, P., Harper, D. A., & Neumann, M. (2019). Revisiting the Great Ordovician Diversification of land plants: Recent data and perspectives. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 534, 109280.
- Sharma, A., Badola, P. K., Bhatia, C., Sharma, D., & Trivedi, P. K. (2020). Primary transcript of miR858 encodes regulatory peptide and controls flavonoid biosynthesis and development in *Arabidopsis*. *Nat Plants*, 6(10), 1262-1274. <https://doi.org/10.1038/s41477-020-00769-x>
- Shaw, J., & Renzaglia, K. (2004). Phylogeny and diversification of bryophytes. *American journal of botany*, 91(10), 1557-1581.
- Shimamura, M. (2016). *Marchantia polymorpha*: Taxonomy, Phylogeny and Morphology of a Model System. *Plant and Cell Physiology*, 57(2), 230-256. <https://doi.org/10.1093/pcp/pcv192>
- Simpson, J. C., Griffiths, G., Wessling-Resnick, M., Fransen, J. A., Bennett, H., & Jones, A. T. (2004). A role for the small GTPase Rab21 in the early endocytic pathway. *Journal of cell science*, 117(26), 6297-6311.
- Stepien, A., Knop, K., Dolata, J., Taube, M., Bajczyk, M., Barciszewska-Pacak, M., Pacak, A., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2017). Posttranscriptional coordination of splicing and miRNA biogenesis in plants. *Wiley Interdiscip Rev RNA*, 8(3). <https://doi.org/10.1002/wrna.1403>
- Streubel, S., Deiber, S., Rotzer, J., Mosiolek, M., Jandrasits, K., & Dolan, L. (2023). Meristem dormancy in *Marchantia polymorpha* is regulated by a liverwort-specific miRNA and a

- clade III SPL gene. *Curr Biol*, 33(4), 660-674 e664. <https://doi.org/10.1016/j.cub.2022.12.062>
- Strother, P. K., & Foster, C. (2021). A fossil record of land plant origins from charophyte algae. *Science*, 373(6556), 792-796.
- Sugano, S. S., & Nishihama, R. (2018). CRISPR/Cas9-based genome editing of transcription factor genes in *Marchantia polymorpha*. *Plant Transcription Factors: Methods and Protocols*, 109-126.
- Sugano, S. S., Nishihama, R., Shirakawa, M., Takagi, J., Matsuda, Y., Ishida, S., Shimada, T., Hara-Nishimura, I., Osakabe, K., & Kohchi, T. (2018). Efficient CRISPR/Cas9-based genome editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS One*, 13(10), e0205117.
- Sugano, S. S., Shirakawa, M., Takagi, J., Matsuda, Y., Shimada, T., Hara-Nishimura, I., & Kohchi, T. (2014). CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia polymorpha* L. *Plant and Cell Physiology*, 55(3), 475-481.
- Swida-Barteczka, A., Pacak, A., Kruszka, K., Nuc, P., Karlowski, W. M., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2023). MicroRNA172b-5p/trehalose-6-phosphate synthase module stimulates trehalose synthesis and microRNA172b-3p/AP2-like module accelerates flowering in barley upon drought stress. *Frontiers in Plant Science*, 14, 1124785. <https://doi.org/10.3389/fpls.2023.1124785>
- Szarzynska, B., Sobkowiak, L., Pant, B. D., Balazadeh, S., Scheible, W. R., Mueller-Roeber, B., Jarmolowski, A., & Szweykowska-Kulinska, Z. (2009). Gene structures and processing of *Arabidopsis thaliana* HYL1-dependent pri-miRNAs. *Nucleic Acids Res*, 37(9), 3083-3093. <https://doi.org/10.1093/nar/gkp189>
- Tan, Q. W., Lim, P. K., Chen, Z., Pasha, A., Provart, N., Arend, M., Nikoloski, Z., & Mutwil, M. (2023). Cross-stress gene expression atlas of *Marchantia polymorpha* reveals the hierarchy and regulatory principles of abiotic stress responses. *Nat Commun*, 14(1), 986. <https://doi.org/10.1038/s41467-023-36517-w>
- Tang, G., Reinhart, B. J., Bartel, D. P., & Zamore, P. D. (2003). A biochemical framework for RNA silencing in plants. *Genes Dev*, 17(1), 49-63. <https://doi.org/10.1101/gad.1048103>
- Thamm, A., Saunders, T. E., & Dolan, L. (2020). MpFEW RHIZOIDS1 miRNA-Mediated Lateral Inhibition Controls Rhizoid Cell Patterning in *Marchantia polymorpha*. *Curr Biol*, 30(10), 1905-1915 e1904. <https://doi.org/10.1016/j.cub.2020.03.032>
- Togawa, T., Adachi, T., Harada, D., Mitani, T., Tanaka, D., Ishizaki, K., Kohchi, T., & Yamato, K. T. (2018). Cryopreservation of *Marchantia polymorpha* spermatozoa. *Journal of Plant Research*, 131, 1047-1054.
- Tsuboyama, S., & Kodama, Y. (2018). AgarTrap protocols on your benchtop: simple methods for *Agrobacterium*-mediated genetic transformation of the liverwort *Marchantia polymorpha*. *Plant Biotechnology*, 35(2), 93-99.
- Tsuzuki, M., Futagami, K., Shimamura, M., Inoue, C., Kunimoto, K., Oogami, T., Tomita, Y., Inoue, K., Kohchi, T., Yamaoka, S., Araki, T., Hamada, T., & Watanabe, Y. (2019). An Early Arising Role of the MicroRNA156/529-SPL Module in Reproductive Development Revealed by the Liverwort *Marchantia polymorpha*. *Curr Biol*, 29(19), 3307-3314 e3305. <https://doi.org/10.1016/j.cub.2019.07.084>
- Tsuzuki, M., Nishihama, R., Ishizaki, K., Kurihara, Y., Matsui, M., Bowman, J. L., Kohchi, T., Hamada, T., & Watanabe, Y. (2016). Profiling and Characterization of Small RNAs in the Liverwort, *Marchantia polymorpha*, Belonging to the First Diverged Land Plants. *Plant and Cell Physiology*, 57(2), 359-372. <https://doi.org/10.1093/pcp/pcv182>
- Uemura, T., & Ueda, T. (2014). Plant vacuolar trafficking driven by RAB and SNARE proteins. *Current opinion in plant biology*, 22, 116-121.

- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research*, 40(15), e115-e115. <https://doi.org/10.1093/nar/gks596>
- Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. *Cell*, 136(4), 669-687. <https://doi.org/10.1016/j.cell.2009.01.046>
- Wallner, E.-S., & Dolan, L. (2024). Reproducibly oriented cell divisions pattern the prothallus to set up dorsoventrality and de novo meristem formation in *Marchantia polymorpha*. *Current Biology*, 34(19), 4357-4367. e4354.
- Williams, L., Grigg, S. P., Xie, M. T., Christensen, S., & Fletcher, J. C. (2005). Regulation of *Arabidopsis* shoot apical meristem and lateral organ formation by microRNA miR166g and its AtHD-ZIP target genes. *Development*, 132(16), 3657-3668. <https://doi.org/10.1242/dev.01942>
- Xie, Z. X., Allen, E., Fahlgren, N., Calamar, A., Givan, S. A., & Carrington, J. C. (2005). Expression of *Arabidopsis* MIRNA genes. *Plant physiology*, 138(4), 2145-2154. <https://doi.org/10.1104/pp.105.062943>
- Yamamoto, C., Takahashi, F., Suetsugu, N., Yamada, K., Yoshikawa, S., Kohchi, T., & Kasahara, M. (2024). The cAMP signaling module regulates sperm motility in the liverwort *Marchantia polymorpha*. *Proceedings of the National Academy of Sciences*, 121(16), e232221121.
- Yamaoka, S., Nishihama, R., Yoshitake, Y., Ishida, S., Inoue, K., Saito, M., Okahashi, K., Bao, H., Nishida, H., Yamaguchi, K., Shigenobu, S., Ishizaki, K., Yamato, K. T., & Kohchi, T. (2018). Generative Cell Specification Requires Transcription Factors Evolutionarily Conserved in Land Plants. *Current Biology*, 28(3), 479-+. <https://doi.org/10.1016/j.cub.2017.12.053>
- Yamaoka, S., Takenaka, M., Hanajiri, T., Shimizu-Ueda, Y., Nishida, H., Yamato, K. T., Fukuzawa, H., & Ohyama, K. (2004). A mutant with constitutive sexual organ development in *Marchantia polymorpha* L. *Sexual Plant Reproduction*, 16, 253-257.
- Yan, K., Liu, P., Wu, C.-A., Yang, G.-D., Xu, R., Guo, Q.-H., Huang, J.-G., & Zheng, C.-C. (2012). Stress-induced alternative splicing provides a mechanism for the regulation of microRNA processing in *Arabidopsis thaliana*. *Molecular Cell*, 48(4), 521-531.
- Yang, L., Liu, Z., Lu, F., Dong, A., & Huang, H. (2006). SERRATE is a novel nuclear regulator in primary microRNA processing in *Arabidopsis*. *The Plant Journal*, 47(6), 841-850. <https://doi.org/https://doi.org/10.1111/j.1365-313X.2006.02835.x>
- Yeats, T. H., & Rose, J. K. (2013). The formation and function of plant cuticles. *Plant physiology*, 163(1), 5-20.
- Yelina, N. E., Frangedakis, E., Schreier, T. B., Rever, J., Tomaselli, M., Haseloff, J., & Hibberd, J. M. (2023). Streamlined regulation of chloroplast development in the liverwort *Marchantia polymorpha*. *bioRxiv*, 2023.2001. 2023.525199.
- Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R. W., Steward, R., & Chen, X. (2005). Methylation as a Crucial Step in Plant microRNA Biogenesis. *Science*, 307(5711), 932-935. <https://doi.org/doi:10.1126/science.1107130>

Supplementary data

Supplementary Table 8. List of primers, guide RNAs, and probes used in this study.

| No. | Primers | Sequence (5'→3') |
|------------------------|----------------|----------------------------|
| 5'-RLM RACE | | |
| 1. | 5RLM_737a_R1 | CAGGCCCATGGAAGGAAACA |
| 2. | 5RLM_737a_R2 | GTGACCACCGATCCATAGATG |
| 3. | 5RLM_11865_R1 | GCAAGTGCAACTCCCAGCCAA |
| 4. | 5RLM_11865_R2 | GCACCTGCTGTGTGAAGCAT |
| 5. | 5RLM_11887_R1 | GTCTTAGCTCCCTAACACTTTCC |
| 6. | 5RLM_11887_R2 | GCGACACGCGGTTAGTTATG |
| 7. | 5RLM_11796_R1 | AAGGCCTCGTAAGCACACTCA |
| 8. | 5RLM_11796_R2 | CGAGGAGGCGCGTAGATCCCA |
| 3'-RACE | | |
| 1. | 3RACE_737a_R1 | CATCTATGGATCGGTGGTCAC |
| 2. | 3RACE_737a_R2 | GTAGACAATCTGCTTGGTCCG |
| 3. | 3RACE_11865_R1 | GGATTCTAGTCAGGGCACTCAT |
| 4. | 3RACE_11865_R2 | CAATCTCGGGTACAGACCTT |
| 5. | 3RACE_11887_R1 | AGTAAGTGACGCGGCGACTT |
| 6. | 3RACE_11887_R2 | GTATGCGTTTGTGTCTAGGC |
| 7. | 3RACE_11796_R1 | GATTTGTGGGATCTACGCGCCT |
| 8. | 3RACE_11796_R2 | AGATCCCACCGACCGCCTGAGT |
| Full length transcript | | |
| 1. | FLT_737a_F1 | GTATACAGTCCTCGCTCTGA |
| 2. | FLT_737a-F2 | CAGTCCTCGCTCTGAAAAGCAC |
| 3. | FLT_737a_R1 | CTGCAAGCGATAATTCTTCCG |
| 4. | FLT_737a_R2 | ATACTATATCAAAAGTGACAGTAGAC |
| 5. | FLT_11865_F1 | ACTGGCCTGCCGCGATCGGA |
| 6. | FLT_11865_F2 | GGAGCGACCTGTTGATCCTTTGC |
| 7. | FLT_11865_R1 | CGAGCCATTCACTTTGTAGAT |
| 8. | FLT_11865_R2 | CACGCGTTCGCATTTCTTAACATACG |

| | | |
|-----|--------------|----------------------------|
| 9. | FLT_11887_F1 | AGTACCCTTTTCGATCGAGGTCA |
| 10. | FLT_11887_R1 | GTGCAGTTCTCCTTCAGTAGGAAGAG |
| 11. | FLT_11887_R2 | CACAGTAACTCGAGGAAGTAC |
| 12. | FLT_11796-F | TAAAGAGCAATGCCACTCTCGGG |
| 13. | FLT_11796-R | GCAATTATGCAATGTTCTGTCTGC |

RT-qPCR

| | | |
|-----|----------------|-------------------------|
| 1. | pri11737a-F | GATCGGTGGTCACGAAGCTT |
| 2. | pri11737a-R | AACCACTGGACGATGCATGA |
| 3. | pri11865-F | GCACTCATTATTGCTTTATC |
| 4. | pri11865-R | GCACACCATGGCCTTTGCGT |
| 5. | pri11796-F | AAGTCCTCTGAAGAACATCC |
| 6. | pri11796-R | AAGGCCTCGTAAGCACACTCA |
| 7. | pri11887-F | GGAAAGTGTTAGGGAGCTAA |
| 8. | pri11887-R | GGAAAACATTAGAGAGCTTAGC |
| 9. | pri11737b-F | ATGGAGCTCCGGACATTCAT |
| 10. | pri11737b-R | ATGGAGCACGACAAGCATAT |
| 11. | MpACT-F | AGGCATCTGGTATCCACGAG |
| 12. | MpACT-R | ACATGGTCGTTCCCTCCAGAC |
| 13. | MpAPT-F | CGAAAGCCCAAGAAGCTACC |
| 14. | MpAPT-R | GTACCCCGGTTGCAATAAG |
| 15. | cdc796target-F | AATCACGAGGATTTGGTTCCG |
| 16. | cdc796target-R | CAATGAATGGAGTTTAGTTAT |
| 17. | 887T_7g16780_F | TTCGCCCTATAATTGCCGACA |
| 18. | 887T_7g16780_R | GGTTGATGAGATGTGCTATGC |
| 19. | tubulinFss | AGCTGACACTACGGTTCTGGT |
| 20. | tubulinRss | GAGCTTCGACACATTTCACTGCC |

5'-RACE

| | | |
|----|-------------------|----------------------------|
| 1. | cdc11796target-R1 | CGGAGACGGTAATGTGGATGAC |
| 2. | cdc11796target-R2 | CACACTACCGTTGACGGTCTT |
| 3. | MpATX1-Race1 | GAACCCCGACAACCAGTACACCCA |
| 4. | MpATX1-Race2 | CAATAGCATGTCAACAGTTCATCCTT |

| | | |
|--|-------------------|--------------------------|
| 5. | RACE-ATX1-R3 | CCGACAACCAGTACACCCACAAT |
| 6. | RACE-ATX1-R4 | TAGACCTCTTCCGATTCAGCCGA |
| 7. | tubulin_RACE_R2 | ACGGCAGCAGCGAACATGATGCA |
| 8. | Tubulin11887-rev5 | CAGTGTCATTATTCGGGGGAG |
| 9. | H1_796-REV5 | CTTGGACAGTCAAACCTACGGT |
| 10. | H1_796-REV4 | GACCTCAATGTTTCGCATACA |
| 11. | 4g22015-R1 | CGACACGGTGGATCCTACATT |
| 12. | 4g22015-R2 | CTTATAAGATTATTCTCCCGCAAA |
| Sequencing primers | | |
| 1. | M13-F | CACGACGTTGTAAAACGAC |
| 2. | M13-R | GGATAACAATTTACACAGG |
| gRNAs for CRISPR-Cas9 | | |
| 1. | gRNA1_11887 | GACTAGTTACAAGTGAATGC |
| 2. | gRNA2_11887 | TCTTGTATGCGTTTGTGTCT |
| 3. | gRNA1_11796 | AGAACATCCAGTCGATTTG |
| 4. | gRNA2_11796 | GTGGGATCTACGCGCCTCCT |
| Promoter amplification and cloning | | |
| 1. | P11796-F | CACCGACTGATTCCCTACTAAATA |
| 2. | P11796-R | CTTGGGTACACTAGAGCACA |
| 3. | P11887-F | CACCCGGCCGACATGTTTAGGTAG |
| 4. | P11887-R | CGAAAGGGTACTGTTTTCTAGC |
| Probes for Northern blot hybridization | | |
| 1. | MpmiR11737ab | GAGATTGTTTTCTTCCACGGGG |
| 2. | MpmiR11865* | ATGCATCTCCTCTGTGAAGCA |
| 3. | MpmiR11887 | GAGCTTAGCCTAGACACAAAC |
| 4. | MpmiR11796 | AGATCCCACAAATCGACTGGA |
| 5. | U6 | TCATCCTTGCGCAGGGGCCA |
| 6. | FAN-1 | TGCTTCCATGACAGCCAAAGGAGG |
| 7. | FAN-3 | AGTGCGCGGATCGCTGACGTGGC |
| 8. | FAN-4 | GACACAGGAGCCAGTCTTCGAA |

Primers for genotyping

| | | |
|----|--------------|------------------------|
| 1. | Gtyp_11887_F | GCAAAGTAGGCTCCTCTGTCAA |
| 2. | Gtyp_11887_R | CCTGCGCAGATAAACTGAGA |
| 3. | Gtyp_11796_F | CTGTAATGAGTGCGCGTTCA |
| 4. | Gtyp_11796_R | GATCTTGAGCCATGCTCGGA |

Supplementary Table 9. GO categories and list of genes in each category.

| GO.ID | Description | P-value | FDR |
|------------|-------------|---------|---------|
| GO:0005524 | ATP binding | 2.8E-05 | 2.8E-05 |

| Sr. No. | Gene ID | Annotation in marchantia.info |
|---------|-------------|---|
| 1. | Mp3g07700.1 | >Mp3g07700.1 [ProSitePatterns:PS00154] E1-E2 ATPases phosphorylation site.; [NCBIfam:TIGR01494] HAD-IC family P-type ATPase; [G3DSA:3.40.50.1000]; [NCBIfam:TIGR01647] plasma-membrane proton-efflux P-type ATPase; [SUPERFAMILY:SSF56784] HAD-like; [SFLD:SFLDF00027] p-type atpase; [G3DSA:2.70.150.10]; [PRINTS:PR00119] P-type cation-transporting ATPase superfamily signature; [Pfam:PF00702] haloacid dehalogenase-like hydrolase; [PANTHER:PTHR42861] CALCIUM-TRANSPORTING ATPASE; [SMART:SM00831] Cation_ATPase_N_a_2; [G3DSA:3.40.1110.10:FF:000004] Plasma membrane ATPase; [Pfam:PF00690] Cation transporter/ATPase, N-terminus; [G3DSA:3.40.50.1000:FF:000211] Plasma membrane ATPase; [Pfam:PF00122] E1-E2 ATPase; [G3DSA:2.70.150.10:FF:000004] Plasma membrane ATPase; [SFLD:SFLDS00003] Haloacid Dehalogenase; [SUPERFAMILY:SSF81665] Calcium ATPase, transmembrane domain M; [G3DSA:6.10.140.890]; [PRINTS:PR00120] H ⁺ -transporting ATPase (proton pump) signature; [CDD:cd02076] P-type_ATPase_H; [G3DSA:1.20.1110.10:FF:000045] ATPase 4 plasma membrane-type; [G3DSA:3.40.1110.10]; [G3DSA:1.20.1110.10]; [SUPERFAMILY:SSF81653] Calcium ATPase, transduction domain A; [KEGG:K01535] H ⁺ -transporting ATPase [EC:7.1.2.1]; [KOG:KOG0205] Plasma membrane H ⁺ -transporting ATPase; [P]; [MapolyID:Mapoly0006s0246]; [GO:0008553] P-type proton-exporting transporter activity; [GO:0005524] ATP binding; [GO:0051453] regulation of intracellular pH; [GO:1902600] proton transmembrane transport; [GO:0000166] nucleotide binding; [GO:0120029] proton export across plasma membrane; [GO:0005215] transporter activity; [GO:0016887] ATP hydrolysis activity; [GO:0005886] plasma membrane; [GO:0015662] P-type ion transporter activity; [GO:0016021] membrane; [GO:0034220] monoatomic ion transmembrane transport; [GO:0016020] membrane |
| 2. | Mp3g07690.1 | >Mp3g07690.1 [ProSitePatterns:PS00154] E1-E2 ATPases phosphorylation site.; [SFLD:SFLDF00027] p-type atpase; [Pfam:PF00690] Cation transporter/ATPase, N-terminus; [G3DSA:1.20.1110.10]; [SFLD:SFLDS00003] Haloacid Dehalogenase; [SUPERFAMILY:SSF81665] Calcium ATPase, transmembrane domain M; [G3DSA:3.40.1110.10:FF:000004] Plasma membrane ATPase; [Pfam:PF00702] haloacid dehalogenase-like hydrolase; [CDD:cd02076] P-type_ATPase_H; [NCBIfam:TIGR01494] HAD-IC family P-type ATPase; [G3DSA:2.70.150.10]; |

| | | |
|----|-------------|---|
| | | [G3DSA:3.40.50.1000:FF:000211] Plasma membrane ATPase; [PRINTS:PR00120] H ⁺ -transporting ATPase (proton pump) signature; [SUPERFAMILY:SSF81653] Calcium ATPase, transduction domain A; [PANTHER:PTHR42861] CALCIUM-TRANSPORTING ATPASE; [SUPERFAMILY:SSF56784] HAD-like; [G3DSA:3.40.1110.10]; [Pfam:PF00122] E1-E2 ATPase; [SMART:SM00831] Cation_ATPase_N_a_2; [NCBIfam:TIGR01647] plasma-membrane proton-efflux P-type ATPase; [G3DSA:1.20.1110.10:FF:000045] ATPase 4 plasma membrane-type; [G3DSA:6.10.140.890]; [G3DSA:2.70.150.10:FF:000004] Plasma membrane ATPase; [G3DSA:3.40.50.1000]; [PRINTS:PR00119] P-type cation-transporting ATPase superfamily signature; [KEGG:K01535] H ⁺ -transporting ATPase [EC:7.1.2.1]; [KOG:KOG0205] Plasma membrane H ⁺ -transporting ATPase; [P]; [MapolyID:Mapoly0006s0245]; [GO:0008553] P-type proton-exporting transporter activity; [GO:0005524] ATP binding; [GO:0051453] regulation of intracellular pH; [GO:1902600] proton transmembrane transport; [GO:0000166] nucleotide binding; [GO:0120029] proton export across plasma membrane; [GO:0005215] transporter activity; [GO:0016887] ATP hydrolysis activity; [GO:0005886] plasma membrane; [GO:0015662] P-type ion transporter activity; [GO:0016021] membrane; [GO:0034220] monoatomic ion transmembrane transport; [GO:0016020] membrane |
| 3. | Mp7g16990.4 | >Mp7g16990.4 [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [ProSitePatterns:PS00107] Protein kinases ATP-binding region signature.; [PANTHER:PTHR47209] OS06G0639500 PROTEIN; [Pfam:PF07714] Protein tyrosine and serine/threonine kinase; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [KOG:KOG0198] MEKK and related serine/threonine protein kinases; [T]; [MapolyID:Mapoly0051s0037]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005524] ATP binding |
| 4. | Mp3g12410.1 | >Mp3g12410.1 [ProSitePatterns:PS00154] E1-E2 ATPases phosphorylation site.; [G3DSA:6.10.140.890]; [G3DSA:3.40.1110.10]; [G3DSA:3.40.50.1000:FF:000211] Plasma membrane ATPase; [Pfam:PF00122] E1-E2 ATPase; [SUPERFAMILY:SSF81653] Calcium ATPase, transduction domain A; [G3DSA:1.20.1110.10]; [SFLD:SFLDS00003] Haloacid Dehalogenase; [NCBIfam:TIGR01494] HAD-IC family P-type ATPase; [SUPERFAMILY:SSF81665] Calcium ATPase, transmembrane domain M; [PRINTS:PR00119] P-type cation-transporting ATPase superfamily signature; [G3DSA:1.20.1110.10:FF:000045] ATPase 4 plasma membrane-type; [G3DSA:3.40.1110.10:FF:000004] Plasma membrane ATPase; [G3DSA:3.40.50.1000]; [PRINTS:PR00120] H ⁺ -transporting ATPase (proton pump) signature; [G3DSA:2.70.150.10:FF:000004] Plasma membrane ATPase; [SUPERFAMILY:SSF56784] HAD-like; [Pfam:PF00702] haloacid dehalogenase-like hydrolase; [PANTHER:PTHR42861] CALCIUM-TRANSPORTING ATPASE; [Pfam:PF00690] Cation transporter/ATPase, N-terminus; [CDD:cd02076] P-type_ATPase_H; [G3DSA:2.70.150.10]; |

| | | |
|----|-------------|--|
| | | [SFLD:SFLDF00027] p-type atpase; [SMART:SM00831] Cation_ATPase_N_a_2; [NCBIfam:TIGR01647] plasma-membrane proton-efflux P-type ATPase; [KEGG:K01535] H ⁺ -transporting ATPase [EC:7.1.2.1]; [KOG:KOG0205] Plasma membrane H ⁺ -transporting ATPase; [P]; [MapolyID:Mapoly0050s0045]; [GO:0008553] P-type proton-exporting transporter activity; [GO:0005524] ATP binding; [GO:0051453] regulation of intracellular pH; [GO:1902600] proton transmembrane transport; [GO:0000166] nucleotide binding; [GO:0120029] proton export across plasma membrane; [GO:0005215] transporter activity; [GO:0016887] ATP hydrolysis activity; [GO:0005886] plasma membrane; [GO:0015662] P-type ion transporter activity; [GO:0016021] membrane; [GO:0034220] monoatomic ion transmembrane transport; [GO:0016020] membrane |
| 5. | Mp1g23430.2 | >Mp1g23430.2 [CDD:cd03250] ABCC_MRP_domain1; [Pfam:PF00664] ABC transporter transmembrane region; [SMART:SM00382] AAA_5; [G3DSA:3.40.50.300]; [CDD:cd03244] ABCC_MRP_domain2; [ProSiteProfiles:PS50893] ATP-binding cassette, ABC transporter-type domain profile.; [G3DSA:1.20.1560.10] ABC transporter type 1, transmembrane domain; [SUPERFAMILY:SSF52540] P-loop containing nucleoside triphosphate hydrolases; [ProSiteProfiles:PS50929] ABC transporter integral membrane type-1 fused domain profile.; [SUPERFAMILY:SSF90123] ABC transporter transmembrane region; [Pfam:PF00005] ABC transporter; [G3DSA:1.20.1560.10:FF:000013] ABC transporter C family member 2; [CDD:cd18579] ABC_6TM_ABCC_D1; [PANTHER:PTHR24223] ATP-BINDING CASSETTE SUB-FAMILY C; [MobiDBLite:mobidb-lite] consensus disorder prediction; [G3DSA:3.40.50.300:FF:000997] Multidrug resistance-associated protein 1; [CDD:cd18580] ABC_6TM_ABCC_D2; [G3DSA:3.40.50.300:FF:000163] Multidrug resistance-associated protein member 4; [G3DSA:1.20.1560.10:FF:000003] ABC transporter C family member 10; [ProSitePatterns:PS00211] ABC transporters family signature.; [KOG:KOG0054] Multidrug resistance-associated protein/mitoxantrone resistance protein, ABC superfamily; [Q]; [MapolyID:Mapoly0065s0035]; [GO:0055085] transmembrane transport; [GO:0042626] ATPase-coupled transmembrane transporter activity; [GO:0140359] ABC-type transporter activity; [GO:0005524] ATP binding; [GO:0016887] ATP hydrolysis activity; [GO:0016020] membrane |
| 6. | Mp2g04150.3 | >Mp2g04150.3 [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [MobiDBLite:mobidb-lite] consensus disorder prediction; [G3DSA:3.80.10.10:FF:000129] Leucine-rich repeat receptor-like kinase; [ProSiteProfiles:PS50011] Protein kinase domain profile.; [Pfam:PF13855] Leucine rich repeat; [Pfam:PF07714] Protein tyrosine and serine/threonine kinase; [PANTHER:PTHR45631] OS07G0107800 PROTEIN-RELATED; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [SUPERFAMILY:SSF52058] L domain-like; [Pfam:PF12819] Malectin-like domain; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG0444] Cytoskeletal regulator Flightless-I (contains |

| | | |
|----|-------------|---|
| | | leucine-rich and gelsolin repeats); N-term missing; C-term missing; [Z]; [MapolyID:Mapoly0031s0071]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |
| 7. | Mp3g17650.1 | >Mp3g17650.1 [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:1.10.510.10:FF:000358] Putative leucine-rich repeat receptor-like serine/threonine-protein kinase; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [Pfam:PF13855] Leucine rich repeat; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [Pfam:PF00560] Leucine Rich Repeat; [G3DSA:3.80.10.10:FF:000041] LRR receptor-like serine/threonine-protein kinase ERECTA; [CDD:cd14066] STKc_IRAK; [Pfam:PF08263] Leucine rich repeat N-terminal domain; [SUPERFAMILY:SSF52058] L domain-like; [G3DSA:3.80.10.10:FF:000400] Nuclear pore complex protein NUP107; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [SMART:SM00369] LRR_typ_2; [Pfam:PF00069] Protein kinase domain; [ProSitePatterns:PS00108] Serine/Threonine protein kinases active-site signature.; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [PANTHER:PTHR27008] OS04G0122200 PROTEIN; [SMART:SM00220] serkin_6; [ProSiteProfiles:PS51450] Leucine-rich repeat profile.; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG0618] Serine/threonine phosphatase 2C containing leucine-rich repeats, similar to SCN circadian oscillatory protein (SCOP); C-term missing; [T]; [MapolyID:Mapoly0039s0031]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |
| 8. | Mp1g25800.2 | >Mp1g25800.2 [G3DSA:1.10.8.60:FF:000002] ATP-dependent Clp protease ATP-binding subunit ClpX; [MobiDBLite:mobidb-lite] consensus disorder prediction; [Pfam:PF07724] AAA domain (Cdc48 subfamily); [SUPERFAMILY:SSF52540] P-loop containing nucleoside triphosphate hydrolases; [G3DSA:3.40.50.300]; [G3DSA:3.40.50.300:FF:000560] CLP protease regulatory subunit CLPX3 mitochondrial; [G3DSA:1.10.8.60]; [SMART:SM01086] ClpB_D2_small_2; [NCBIfam:TIGR00382] ATP-dependent protease ATP-binding subunit ClpX; [CDD:cd19497] RecA-like_ClpX; [SMART:SM00382] AAA_5; [PANTHER:PTHR48102] ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX-LIKE, MITOCHONDRIAL-RELATED; [Pfam:PF10431] C-terminal, D2-small domain, of ClpB protein; [KEGG:K03544] ATP-dependent Clp protease ATP-binding subunit ClpX; [KOG:KOG0745] Putative ATP-dependent Clp-type protease (AAA+ ATPase superfamily); N-term missing; [O]; [MapolyID:Mapoly0002s0296]; [GO:0005524] ATP binding; [GO:0051082] unfolded protein binding; [GO:0005759] mitochondrial matrix; [GO:0016887] ATP hydrolysis activity; [GO:0140662] ATP-dependent protein folding chaperone; [GO:0051603] proteolysis involved in protein catabolic process; [GO:0006457] protein folding |

| | | |
|-----|-------------|---|
| 9. | Mp5g09530.1 | >Mp5g09530.1 [SMART:SM00220] serkin_6; [CDD:cd14066] STKc_IRAK; [ProSiteProfiles:PS50011] Protein kinase domain profile.; [ProSitePatterns:PS00107] Protein kinases ATP-binding region signature.; [G3DSA:3.30.200.20:FF:000162] Adenine nucleotide alpha hydrolase-like domain kinase; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [ProSitePatterns:PS00108] Serine/Threonine protein kinases active-site signature.; [PANTHER:PTHR46008] LEAF RUST 10 DISEASE-RESISTANCE LOCUS RECEPTOR-LIKE PROTEIN KINASE-LIKE 1.4; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [Pfam:PF00069] Protein kinase domain; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [MapolyID:Mapoly0095s0007]; [GO:0005524] ATP binding; [GO:0016310] phosphorylation; [GO:0004672] protein kinase activity; [GO:0006468] protein phosphorylation; [GO:0016301] kinase activity |
| 10. | Mp2g17270.1 | >Mp2g17270.1 [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [G3DSA:1.10.510.10:FF:000095] protein STRUBBELIG-RECEPTOR FAMILY 8; [Pfam:PF00560] Leucine Rich Repeat; [CDD:cd14066] STKc_IRAK; [Pfam:PF13855] Leucine rich repeat; [G3DSA:3.30.200.20:FF:000394] Leucine-rich repeat receptor-like protein kinase; [PANTHER:PTHR45631] OS07G0107800 PROTEIN-RELATED; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [Pfam:PF07714] Protein tyrosine and serine/threonine kinase; [ProSitePatterns:PS00107] Protein kinases ATP-binding region signature.; [SMART:SM00220] serkin_6; [SUPERFAMILY:SSF52058] L domain-like; [SMART:SM00369] LRR_typ_2; [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:3.80.10.10:FF:000383] Leucine-rich repeat receptor protein kinase EMS1; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [ProSitePatterns:PS00108] Serine/Threonine protein kinases active-site signature.; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG0444] Cytoskeletal regulator Flightless-I (contains leucine-rich and gelsolin repeats); C-term missing; [Z]; [MapolyID:Mapoly0652s0001]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |
| 11. | Mp4g10940.1 | >Mp4g10940.1 [G3DSA:2.60.120.430]; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [G3DSA:3.80.10.10:FF:000041] LRR receptor-like serine/threonine-protein kinase ERECTA; [SUPERFAMILY:SSF52058] L domain-like; [Pfam:PF07714] Protein tyrosine and serine/threonine kinase; [Pfam:PF12819] Malectin-like domain; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [ProSiteProfiles:PS51450] Leucine-rich repeat profile.; [PANTHER:PTHR45631] OS07G0107800 PROTEIN-RELATED; [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:1.10.510.10:FF:000146] LRR receptor-like serine/threonine-protein kinase IOS1; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG4237] Extracellular matrix protein slit, contains leucine-rich and EGF-like repeats; N-term |

| | | |
|-----|-------------|--|
| | | missing; C-term missing; [WT]; [MapolyID:Mapoly0011s0079]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |
| 12. | Mp4g08420.1 | >Mp4g08420.1 [ProSitePatterns:PS00486] DNA mismatch repair proteins mutS family signature.; [SMART:SM00533] DNAend; [G3DSA:1.10.1420.10]; [G3DSA:3.40.1170.10] DNA repair protein MutS, domain I; [Pfam:PF05188] MutS domain II; [Pfam:PF05190] MutS family domain IV; [G3DSA:1.10.1420.10:FF:000003] DNA mismatch repair protein; [Coils:Coil] Coil; [G3DSA:3.30.420.110] MutS, connector domain; [SUPERFAMILY:SSF48334] DNA repair protein MutS, domain III; [G3DSA:3.30.420.110:FF:000002] DNA mismatch repair protein; [SMART:SM00534] mutATP5; [Pfam:PF00488] MutS domain V; [SUPERFAMILY:SSF52540] P-loop containing nucleoside triphosphate hydrolases; [PANTHER:PTHR11361] DNA MISMATCH REPAIR PROTEIN MUTS FAMILY MEMBER; [SUPERFAMILY:SSF53150] DNA repair protein MutS, domain II; [CDD:cd03285] ABC_MSH2_euk; [G3DSA:3.40.1170.10:FF:000003] DNA mismatch repair protein; [G3DSA:3.40.50.300]; [G3DSA:3.40.50.300:FF:000925] DNA mismatch repair protein MSH2; [Pfam:PF01624] MutS domain I; [PIRSF:PIRSF005813] MSH2; [Pfam:PF05192] MutS domain III; [KEGG:K08735] DNA mismatch repair protein MSH2; [KOG:KOG0219] Mismatch repair ATPase MSH2 (MutS family); [L]; [MapolyID:Mapoly0120s0004]; [GO:0005634] nucleus; [GO:0030983] mismatched DNA binding; [GO:0005524] ATP binding; [GO:0003690] double-stranded DNA binding; [GO:0006312] mitotic recombination; [GO:0140664] ATP-dependent DNA damage sensor activity; [GO:0006298] mismatch repair; [GO:0032301] MutSalph complex |

| GO.ID | Description | P-value | FDR |
|------------|-----------------|---------|--------|
| GO:0005524 | Protein binding | 0.0003 | 0.0003 |

| Sr. No. | Gene ID | Annotation in marchantia.info |
|---------|--------------|---|
| 1. | Mp4g02730.1, | >Mp4g02730.1 [SUPERFAMILY:SSF54236] Ubiquitin-like; [G3DSA:3.10.20.90:FF:000016] Polyubiquitin 3; [CDD:cd01803] Ubl_ubiquitin; [ProSitePatterns:PS00299] Ubiquitin domain signature.; [G3DSA:3.10.20.90]; [PRINTS:PR00348] Ubiquitin signature; [ProSiteProfiles:PS50053] Ubiquitin domain profile.; [PANTHER:PTHR10666] UBIQUITIN; [Pfam:PF00240] Ubiquitin family; [SMART:SM00213] ubq_7; [KEGG:K08770] ubiquitin C; [KEGG:K04551] ubiquitin B; [KOG:KOG0004] Ubiquitin/40S ribosomal protein S27a fusion; C-term missing; [J]; [MapolyID:Mapoly0080s0026]; [GO:0031386] protein tag activity; [GO:0005634] nucleus; [GO:0016567] protein ubiquitination; [GO:0031625] ubiquitin protein ligase binding; [GO:0005515] protein binding; [GO:0019941] modification-dependent protein catabolic process; [GO:0005737] cytoplasm |
| 2. | Mp4g01140.1, | >Mp4g01140.1 [SUPERFAMILY:SSF117281] Kelch motif; [CDD:cd22157] F-box_AtFBW1-like; [ProSiteProfiles:PS50181] F-box domain profile.; [SMART:SM00256] fbox_2; [G3DSA:1.20.1280.50]; [SUPERFAMILY:SSF81383] F-box domain; [Pfam:PF00646] F-box domain; [G3DSA:2.120.10.80]; [PANTHER:PTHR31672] BNACNNG10540D PROTEIN; [KOG:KOG2120] SCF ubiquitin ligase, Skp2 component; N-term missing; C-term missing; [O]; [MapolyID:Mapoly0066s0028]; [GO:0005515] protein binding |
| 3. | Mp3g09210.1, | >Mp3g09210.1 [Coils:Coil] Coil; [SUPERFAMILY:SSF52058] L domain-like; [MobiDBLite:mobidb-lite] consensus disorder prediction; [G3DSA:3.40.50.300]; [G3DSA:1.10.8.430]; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [SUPERFAMILY:SSF52540] P-loop containing nucleoside triphosphate hydrolases; [G3DSA:1.20.930.20]; [G3DSA:1.10.10.10]; [SMART:SM00369] LRR_typ_2; [Pfam:PF00931] NB-ARC domain; [PANTHER:PTHR36766] PLANT BROAD-SPECTRUM MILDEW RESISTANCE PROTEIN RPW8; [Pfam:PF13855] Leucine rich repeat; [G3DSA:1.10.8.430:FF:000003] Probable disease resistance protein At5g66910; [PRINTS:PR00364] Disease resistance protein signature; [KOG:KOG4658] Apoptotic ATPase; [T]; [GO:0050832] defense response to fungus; [GO:0043531] ADP binding; [GO:0005515] protein binding; [GO:0007166] cell surface receptor signaling pathway |
| 4. | Mp6g19250.1, | >Mp6g19250.1 [SUPERFAMILY:SSF54695] POZ domain; [G3DSA:3.30.710.10] Potassium Channel Kv1.1; Chain A; [ProSiteProfiles:PS50097] BTB domain profile.; [CDD:cd18186] |

| | | |
|----|--------------|---|
| | | BTB_POZ_ZBTB_KLHL-like; [Pfam:PF00651] BTB/POZ domain; [MapolyID:Mapoly0045s0138]; [GO:0005515] protein binding |
| 5. | Mp7g15040.1, | >Mp7g15040.1 [ProSiteProfiles:PS50297] Ankyrin repeat region circular profile.; [Pfam:PF00023] Ankyrin repeat; [SUPERFAMILY:SSF48403] Ankyrin repeat; [G3DSA:1.25.40.20]; [MobiDBLite:mobidb-lite] consensus disorder prediction; [ProSiteProfiles:PS50088] Ankyrin repeat profile.; [PANTHER:PTHR24173] ANKYRIN REPEAT CONTAINING; [SMART:SM00248] ANK_2a; [Pfam:PF12796] Ankyrin repeats (3 copies); [KOG:KOG0509] Ankyrin repeat and DHHC-type Zn-finger domain containing proteins; C-term missing; [R]; [KOG:KOG0510] Ankyrin repeat protein; N-term missing; C-term missing; [R]; [KOG:KOG0512] Fetal globin-inducing factor (contains ankyrin repeats); N-term missing; C-term missing; [K]; [MapolyID:Mapoly0009s0188]; [GO:0005515] protein binding |
| 6. | Mp4g19460.1, | >Mp4g19460.1 [G3DSA:1.20.1280.50]; [SUPERFAMILY:SSF81383] F-box domain; [PANTHER:PTHR31672] BNACNNG10540D PROTEIN; [MapolyID:Mapoly0304s0003]; [GO:0005515] protein binding |
| 7. | Mp7g12640.4, | >Mp7g12640.4 [SUPERFAMILY:SSF50494] Trypsin-like serine proteases; [SUPERFAMILY:SSF50156] PDZ domain-like; [SMART:SM00228] pdz_new; [Pfam:PF17820] PDZ domain; [Pfam:PF13365] Trypsin-like peptidase domain; [G3DSA:2.40.10.10]; [MobiDBLite:mobidb-lite] consensus disorder prediction; [Pfam:PF12812] PDZ-like domain; [CDD:cd00987] PDZ_serine_protease; [ProSiteProfiles:PS50106] PDZ domain profile.; [PRINTS:PR00834] HtrA/DegQ protease family signature; [PANTHER:PTHR46366] PRO-APOPTOTIC SERINE PROTEASE NMA111; [G3DSA:2.30.42.10]; [KEGG:K22686] pro-apoptotic serine protease NMA111 [EC:3.4.21.-]; [KOG:KOG1421] Predicted signaling-associated protein (contains a PDZ domain); C-term missing; [R]; [MapolyID:Mapoly0003s0272]; [GO:0006508] proteolysis; [GO:0005515] protein binding; [GO:0004252] serine-type endopeptidase activity |
| 8. | Mp2g07520.1, | >Mp2g07520.1 [SUPERFAMILY:SSF81383] F-box domain; [SMART:SM00256] fbox_2; [G3DSA:1.20.1280.50]; [ProSiteProfiles:PS50181] F-box domain profile.; [Pfam:PF00646] F-box domain; [PANTHER:PTHR31672] BNACNNG10540D PROTEIN; [MapolyID:Mapoly0015s0038]; [GO:0005515] protein binding |
| 9. | Mp4g19400.1, | >Mp4g19400.1 [G3DSA:1.20.1280.50]; [G3DSA:2.120.10.80]; [SUPERFAMILY:SSF117281] Kelch motif; [SUPERFAMILY:SSF81383] F-box domain; [PANTHER:PTHR31672] BNACNNG10540D PROTEIN; [KOG:KOG0379] Kelch repeat-containing proteins; N-term missing; C-term missing; [R]; [MapolyID:Mapoly0169s0004]; [GO:0005515] protein binding |

| | | |
|-----|--------------|---|
| | | |
| 10. | Mp1g08170.1, | >Mp1g08170.1 [SMART:SM00256] fbox_2; [Pfam:PF14299] Phloem protein 2; [Pfam:PF00646] F-box domain; [CDD:cd22162] F-box_AtSKIP3-like; [PANTHER:PTHR31960] F-BOX PROTEIN PP2-A15; [SUPERFAMILY:SSF81383] F-box domain; [MapolyID:Mapoly0036s0061]; [GO:0005515] protein binding |
| 11. | Mp6g00190.1, | >Mp6g00190.1 [SUPERFAMILY:SSF81383] F-box domain; [G3DSA:1.20.1280.50]; [SUPERFAMILY:SSF50965] Galactose oxidase, central domain; [G3DSA:2.120.10.80]; [SMART:SM00256] fbox_2; [PANTHER:PTHR31672] BNACNNG10540D PROTEIN; [MapolyID:Mapoly0163s0003]; [GO:0005515] protein binding |
| 12. | Mp5g09790.1, | >Mp5g09790.1 [G3DSA:1.25.40.10] Tetratricopeptide repeat domain; [Pfam:PF00856] SET domain; [MobiDBLite:mobidb-lite] consensus disorder prediction; [SMART:SM00317] set_7; [CDD:cd20071] SET_SMYD; [SUPERFAMILY:SSF82199] SET domain; [G3DSA:2.170.270.10] SET domain; [ProSiteProfiles:PS50280] SET domain profile.; [SMART:SM00028] tpr_5; [ProSiteProfiles:PS50005] TPR repeat profile.; [Coils:Coil] Coil; [SUPERFAMILY:SSF48452] TPR-like; [PANTHER:PTHR47643] TPR DOMAIN PROTEIN (AFU_ORTHOLOGUE AFUA_5G12710); [KOG:KOG0553] TPR repeat-containing protein; N-term missing; C-term missing; [R]; [KOG:KOG2084] Predicted histone tail methylase containing SET domain; C-term missing; [B]; [MapolyID:Mapoly0048s0091]; [GO:0005515] protein binding |
| 13. | Mp5g24260.1, | >Mp5g24260.1 [MobiDBLite:mobidb-lite] consensus disorder prediction; [Pfam:PF17035] Bromodomain extra-terminal - transcription regulation; [Coils:Coil] Coil; [PRINTS:PR00503] Bromodomain signature; [ProSiteProfiles:PS50014] Bromodomain profile.; [PANTHER:PTHR45926] OSJNBA0053K19.4 PROTEIN; [ProSiteProfiles:PS51525] NET domain profile.; [SMART:SM00297] bromo_6; [Pfam:PF00439] Bromodomain; [G3DSA:1.20.920.10]; [SUPERFAMILY:SSF47370] Bromodomain; [G3DSA:1.20.1270.220]; [KOG:KOG1474] Transcription initiation factor TFIID, subunit BDF1 and related bromodomain proteins; C-term missing; [K]; [MapolyID:Mapoly0010s0030]; [GO:0005515] protein binding |
| 14. | Mp2g08850.1, | >Mp2g08850.1 [G3DSA:3.30.60.30]; [Pfam:PF07648] Kazal-type serine protease inhibitor domain; [SUPERFAMILY:SSF100895] Kazal-type serine protease inhibitors; [PANTHER:PTHR34376] SERINE PROTEASE INHIBITOR, KAZAL-TYPE FAMILY PROTEIN; [KOG:KOG4578] Uncharacterized conserved protein, contains KAZAL and TY domains; C-term missing; [R]; [KOG:KOG3626] Organic anion transporter; N-term missing; C-term missing; [Q]; [MapolyID:Mapoly0015s0170]; [GO:0005515] protein binding |

| | | |
|-----|--------------|--|
| 15. | Mp2g05930.1, | >Mp2g05930.1 [G3DSA:1.25.40.20]; [ProSiteProfiles:PS50088] Ankyrin repeat profile.; [Pfam:PF12796] Ankyrin repeats (3 copies); [PANTHER:PTHR24178] MOLTING PROTEIN MLT-4; [ProSiteProfiles:PS50297] Ankyrin repeat region circular profile.; [SMART:SM00248] ANK_2a; [SUPERFAMILY:SSF48403] Ankyrin repeat; [KOG:KOG4412] 26S proteasome regulatory complex, subunit PSMD10; [O]; [MapolyID:Mapoly0021s0049]; [GO:0005515] protein binding |
| 16. | Mp7g07850.1, | >Mp7g07850.1 [MobiDBLite:mobidb-lite] consensus disorder prediction; [Pfam:PF12796] Ankyrin repeats (3 copies); [PANTHER:PTHR24189] MYOTROPHIN; [ProSiteProfiles:PS50297] Ankyrin repeat region circular profile.; [SUPERFAMILY:SSF48403] Ankyrin repeat; [SMART:SM00248] ANK_2a; [ProSiteProfiles:PS50088] Ankyrin repeat profile.; [G3DSA:1.25.40.20]; [KOG:KOG0512] Fetal globin-inducing factor (contains ankyrin repeats); N-term missing; C-term missing; [K]; [KOG:KOG0521] Putative GTPase activating proteins (GAPs); N-term missing; [T]; [MapolyID:Mapoly0076s0009]; [GO:0005634] nucleus; [GO:0005737] cytoplasm; [GO:0005515] protein binding |
| 17. | Mp1g07690.1, | >Mp1g07690.1 [PANTHER:PTHR38926] F-BOX DOMAIN CONTAINING PROTEIN, EXPRESSED; [SMART:SM00367] LRR_CC_2; [MobiDBLite:mobidb-lite] consensus disorder prediction; [Pfam:PF00646] F-box domain; [CDD:cd09917] F-box_SF; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [SUPERFAMILY:SSF52047] RNI-like; [SUPERFAMILY:SSF81383] F-box domain; [ProSiteProfiles:PS50181] F-box domain profile.; [G3DSA:1.20.1280.50]; [SMART:SM00256] fbox_2; [KOG:KOG1947] Leucine rich repeat proteins, some proteins contain F-box; N-term missing; [R]; [KOG:KOG4341] F-box protein containing LRR; N-term missing; C-term missing; [R]; [MapolyID:Mapoly0036s0015]; [GO:0005515] protein binding |
| 18. | Mp2g04150.3, | >Mp2g04150.3 [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [MobiDBLite:mobidb-lite] consensus disorder prediction; [G3DSA:3.80.10.10:FF:000129] Leucine-rich repeat receptor-like kinase; [ProSiteProfiles:PS50011] Protein kinase domain profile.; [Pfam:PF13855] Leucine rich repeat; [Pfam:PF07714] Protein tyrosine and serine/threonine kinase; [PANTHER:PTHR45631] OS07G0107800 PROTEIN-RELATED; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [SUPERFAMILY:SSF52058] L domain-like; [Pfam:PF12819] Malectin-like domain; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG0444] Cytoskeletal regulator Flightless-I (contains leucine-rich and gelsolin repeats); N-term missing; C-term missing; [Z]; [MapolyID:Mapoly0031s0071]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |

| | | |
|-----|--------------|---|
| 19. | Mp3g20810.1, | >Mp3g20810.1 [Pfam:PF00240] Ubiquitin family; [SUPERFAMILY:SSF54236] Ubiquitin-like; [CDD:cd01791] Ubl_UBL5; [ProSiteProfiles:PS50053] Ubiquitin domain profile.; [G3DSA:3.10.20.90]; [PANTHER:PTHR13042] UBIQUITIN-LIKE PROTEIN 5; [G3DSA:3.10.20.90:FF:000054] Ubiquitin-like protein 5; [KEGG:K13113] ubiquitin-like protein 5; [KOG:KOG3493] Ubiquitin-like protein; [O]; [MapolyID:Mapoly0159s0011]; [GO:0031386] protein tag activity; [GO:0005634] nucleus; [GO:0005515] protein binding; [GO:0000398] mRNA splicing, via spliceosome; [GO:0036211] protein modification process; [GO:0006464] protein modification process; [GO:0005737] cytoplasm |
| 20. | Mp3g22950.1, | >Mp3g22950.1 [G3DSA:3.10.20.90:FF:000016] Polyubiquitin 3; [ProSiteProfiles:PS50053] Ubiquitin domain profile.; [CDD:cd01803] Ubl_ubiquitin; [ProSitePatterns:PS00299] Ubiquitin domain signature.; [SUPERFAMILY:SSF54236] Ubiquitin-like; [PANTHER:PTHR10666] UBIQUITIN; [Pfam:PF00240] Ubiquitin family; [SMART:SM00213] ubq_7; [G3DSA:3.10.20.90]; [PRINTS:PR00348] Ubiquitin signature; [KEGG:K08770] ubiquitin C; [KEGG:K04551] ubiquitin B; [KOG:KOG0004] Ubiquitin/40S ribosomal protein S27a fusion; C-term missing; [J]; [MapolyID:Mapoly0024s0072]; [GO:0031386] protein tag activity; [GO:0005634] nucleus; [GO:0016567] protein ubiquitination; [GO:0031625] ubiquitin protein ligase binding; [GO:0005515] protein binding; [GO:0019941] modification-dependent protein catabolic process; [GO:0005737] cytoplasm |
| 21. | Mp3g17650.1, | >Mp3g17650.1 [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:1.10.510.10:FF:000358] Putative leucine-rich repeat receptor-like serine/threonine-protein kinase; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [Pfam:PF13855] Leucine rich repeat; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [Pfam:PF00560] Leucine Rich Repeat; [G3DSA:3.80.10.10:FF:000041] LRR receptor-like serine/threonine-protein kinase ERECTA; [CDD:cd14066] STKc_IRAK; [Pfam:PF08263] Leucine rich repeat N-terminal domain; [SUPERFAMILY:SSF52058] L domain-like; [G3DSA:3.80.10.10:FF:000400] Nuclear pore complex protein NUP107; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [SMART:SM00369] LRR_typ_2; [Pfam:PF00069] Protein kinase domain; [ProSitePatterns:PS00108] Serine/Threonine protein kinases active-site signature.; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [PANTHER:PTHR27008] OS04G0122200 PROTEIN; [SMART:SM00220] serkin_6; [ProSiteProfiles:PS51450] Leucine-rich repeat profile.; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG0618] Serine/threonine phosphatase 2C containing leucine-rich repeats, similar to SCN circadian oscillatory protein (SCOP); C-term missing; [T]; [MapolyID:Mapoly0039s0031]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |

| | | |
|-----|--------------|---|
| 22. | Mp2g16930.1, | >Mp2g16930.1 [PANTHER:PTHR19877] EUKARYOTIC TRANSLATION INITIATION FACTOR 3 SUBUNIT I; [ProSiteProfiles:PS50082] Trp-Asp (WD) repeats profile.; [SMART:SM00320] WD40_4; [ProSitePatterns:PS00678] Trp-Asp (WD) repeats signature.; [Pfam:PF00400] WD domain, G-beta repeat; [G3DSA:2.130.10.10]; [ProSiteProfiles:PS50294] Trp-Asp (WD) repeats circular profile.; [PRINTS:PR00320] G protein beta WD-40 repeat signature; [Hamap:MF_03008] Eukaryotic translation initiation factor 3 subunit I [EIF3I].; [SUPERFAMILY:SSF50978] WD40 repeat-like; [CDD:cd00200] WD40; [KEGG:K03246] translation initiation factor 3 subunit I; [KOG:KOG0643] Translation initiation factor 3, subunit i (eIF-3i)/TGF-beta receptor-interacting protein (TRIP-1); [JT]; [MapolyID:Mapoly0109s0034]; [GO:0003723] RNA binding; [GO:0002183] cytoplasmic translational initiation; [GO:0005515] protein binding; [GO:0071541] eukaryotic translation initiation factor 3 complex, eIF3m; [GO:0003743] translation initiation factor activity; [GO:0005852] eukaryotic translation initiation factor 3 complex; [GO:0005737] cytoplasm |
| 23. | Mp2g17270.1, | >Mp2g17270.1 [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [G3DSA:1.10.510.10:FF:000095] protein STRUBBELIG-RECEPTOR FAMILY 8; [Pfam:PF00560] Leucine Rich Repeat; [CDD:cd14066] STKc_IRAK; [Pfam:PF13855] Leucine rich repeat; [G3DSA:3.30.200.20:FF:000394] Leucine-rich repeat receptor-like protein kinase; [PANTHER:PTHR45631] OS07G0107800 PROTEIN-RELATED; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [Pfam:PF07714] Protein tyrosine and serine/threonine kinase; [ProSitePatterns:PS00107] Protein kinases ATP-binding region signature.; [SMART:SM00220] serkin_6; [SUPERFAMILY:SSF52058] L domain-like; [SMART:SM00369] LRR_typ_2; [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:3.80.10.10:FF:000383] Leucine-rich repeat receptor protein kinase EMS1; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [ProSitePatterns:PS00108] Serine/Threonine protein kinases active-site signature.; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG0444] Cytoskeletal regulator Flightless-I (contains leucine-rich and gelsolin repeats); C-term missing; [Z]; [MapolyID:Mapoly0652s0001]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |
| 24. | Mp1g17450.1, | >Mp1g17450.1 [MobiDBLite:mobidb-lite] consensus disorder prediction; [PRINTS:PR00320] G protein beta WD-40 repeat signature; [ProSiteProfiles:PS50082] Trp-Asp (WD) repeats profile.; [SUPERFAMILY:SSF50978] WD40 repeat-like; [ProSitePatterns:PS00678] Trp-Asp (WD) repeats signature.; [PANTHER:PTHR43979] PRE-mRNA-PROCESSING FACTOR 17; [CDD:cd00200] WD40; [ProSiteProfiles:PS50294] Trp-Asp (WD) repeats circular profile.; [Pfam:PF00400] WD domain, G-beta repeat; [G3DSA:2.130.10.10]; [G3DSA:2.130.10.10:FF:000034] Pre-mRNA-processing factor 17, putative; [SMART:SM00320] WD40_4; [KEGG:K12816] pre-mRNA-processing factor 17; |

| | | |
|-----|--------------|--|
| | | [KOG:KOG0282] mRNA splicing factor; [S]; [MapolyID:Mapoly0001s0085]; [GO:0071013] catalytic step 2 spliceosome; [GO:0003729] mRNA binding; [GO:0005515] protein binding; [GO:0000398] mRNA splicing, via spliceosome |
| 25. | Mp2g15730.1, | >Mp2g15730.1 [ProSiteProfiles:PS50082] Trp-Asp (WD) repeats profile.; [ProSiteProfiles:PS50294] Trp-Asp (WD) repeats circular profile.; [SUPERFAMILY:SSF50978] WD40 repeat-like; [G3DSA:2.130.10.10]; [Coils:Coil] Coil; [SMART:SM00320] WD40_4; [ProSitePatterns:PS00678] Trp-Asp (WD) repeats signature.; [PANTHER:PTHR19846] WD40 REPEAT PROTEIN; [Pfam:PF00400] WD domain, G-beta repeat; [PRINTS:PR00320] G protein beta WD-40 repeat signature; [KOG:KOG0282] mRNA splicing factor; N-term missing; [S]; [KOG:KOG0642] Cell-cycle nuclear protein, contains WD-40 repeats; N-term missing; C-term missing; [D]; [MapolyID:Mapoly0082s0069, Mapoly0312s0002]; [GO:0046540] U4/U6 x U5 tri-snRNP complex; [GO:0030621] U4 snRNA binding; [GO:0017070] U6 snRNA binding; [GO:0005515] protein binding; [GO:0000398] mRNA splicing, via spliceosome |
| 26. | Mp4g10940.1, | >Mp4g10940.1 [G3DSA:2.60.120.430]; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [G3DSA:3.80.10.10:FF:000041] LRR receptor-like serine/threonine-protein kinase ERECTA; [SUPERFAMILY:SSF52058] L domain-like; [Pfam:PF07714] Protein tyrosine and serine/threonine kinase; [Pfam:PF12819] Malectin-like domain; [SUPERFAMILY:SSF56112] Protein kinase-like (PK-like); [ProSiteProfiles:PS51450] Leucine-rich repeat profile.; [PANTHER:PTHR45631] OS07G0107800 PROTEIN-RELATED; [ProSiteProfiles:PS50011] Protein kinase domain profile.; [G3DSA:1.10.510.10:FF:000146] LRR receptor-like serine/threonine-protein kinase IOS1; [G3DSA:3.30.200.20] Phosphorylase Kinase; domain 1; [G3DSA:1.10.510.10] Transferase(Phosphotransferase) domain 1; [KOG:KOG1187] Serine/threonine protein kinase; [T]; [KOG:KOG4237] Extracellular matrix protein slit, contains leucine-rich and EGF-like repeats; N-term missing; C-term missing; [WT]; [MapolyID:Mapoly0011s0079]; [GO:0006468] protein phosphorylation; [GO:0004672] protein kinase activity; [GO:0005515] protein binding; [GO:0005524] ATP binding |
| 27. | Mp3g02750.1, | >Mp3g02750.1 [PANTHER:PTHR48051]; [SMART:SM00364] LRR_bac_2; [SUPERFAMILY:SSF52058] L domain-like; [ProSiteProfiles:PS51450] Leucine-rich repeat profile.; [Pfam:PF13855] Leucine rich repeat; [G3DSA:3.80.10.10] Ribonuclease Inhibitor; [SMART:SM00369] LRR_typ_2; [Coils:Coil] Coil; [KOG:KOG0472] Leucine-rich repeat protein; C-term missing; [S]; [KOG:KOG1259] Nischarin, modulator of integrin alpha5 subunit action; N-term missing; [TZ]; [MapolyID:Mapoly0007s0263]; [GO:0005515] protein binding |

| | | |
|-----|-------------|--|
| 28. | Mp4g20550.2 | >Mp4g20550.2 [SMART:SM00256] fbox_2; [Pfam:PF00646] F-box domain; [SMART:SM00612] kelc_smart; [PANTHER:PTHR46344] OS02G0202900 PROTEIN; [G3DSA:2.120.10.80]; [SUPERFAMILY:SSF117281] Kelch motif; [CDD:cd22152] F-box_AtAFR-like; [SUPERFAMILY:SSF81383] F-box domain; [Pfam:PF01344] Kelch motif; [KOG:KOG4441] Proteins containing BTB/POZ and Kelch domains, involved in regulatory/signal transduction processes; N-term missing; [TR]; [MapolyID:Mapoly0101s0001]; [GO:0005515] protein binding |
|-----|-------------|--|