TAS3 miR390-dependent loci in non-vascular land plants: towards a comprehensive reconstruction of the gene evolutionary history

Sergey Y. Morozov [Corresp., 1], **Irina A. Milyutina** [1], **Tatiana N. Erokhina** [2], **Liudmila V. Ozerova** [3], **Alexey V. Troitsky** [1], **Andrey G. Solovyev** [1, 4]

[1] Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, Russia

[2] Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow, Russia

[3] Tsitsin Main Botanical Garden, Russian Academy of Science, Moscow, Russia

[4] Institute of Molecular Medicine, Sechenov First Moscow State Medical University, Moscow, Russia

Corresponding Author: Sergey Y. Morozov
Email address: morozov@genebee.msu.su

Trans-acting small interfering RNAs (ta-siRNAs) are transcribed from protein non-coding genomic TAS loci and belong to a plant-specific class of endogenous small RNAs. These siRNAs have been found to regulate gene expression in most taxa including seed plants, gymnosperms, ferns and mosses. In this study, bioinformatic and experimental PCR-based approaches were used as tools to analyze TAS3 and TAS6 loci in transcriptomes and genomic DNAs from representatives of evolutionary distant non-vascular plant taxa such as Bryophyta, Marchantiophyta and Anthocerotophyta. We revealed previously undiscovered TAS3 loci in plant classes Sphagnopsida and Anthocerotopsida, as well as TAS6 loci in Bryophyta classes Tetraphidiopsida, Polytrichopsida, Andreaeopsida and Takakiopsida. These data further unveil the evolutionary pathway of the miR390-dependent TAS3 loci in land plants. We also identified charophyte alga sequences coding for SUPPRESSOR OF GENE SILENCING 3 (SGS3), which is required for generation of ta-siRNAs in plants, and hypothesized that the appearance of TAS3-related sequences could take place at a very early step in evolutionary transition from charophyte algae to an earliest common ancestor of land plants.

TAS3 miR390-dependent loci in non-vascular land plants: towards a comprehensive reconstruction of the gene evolutionary history

Sergey Y. Morozov[1], Irina A. Milyutina[1], Tatiana N. Erokhina2, Liudmila V. Ozerova[3], Alexey V. Troitsky[1], Andrey G. Solovyev[1,4]

[1]Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russian Federation

[2]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow, Russian Federation

[3]Tsitsin Main Botanical Garden, Russian Academy of Science, Moscow, Russian Federation

[4]Institute of Molecular Medicine, Sechenov First Moscow State Medical University, Moscow, Russian Federation

Corresponding author: Sergey Y. Morozov, morozov@genebee.msu.su

Abbreviations: dsRNA – double-stranded RNA; miRNA – microRNA; siRNA – small interfering RNA; ssRNA – single-stranded RNA; tasiARF - trans-acting siRNA specific for ARF gene; ta-siRNA - trans-acting siRNA

ABSTRACT

Trans-acting small interfering RNAs (ta-siRNAs) are transcribed from protein non-coding genomic TAS loci and belong to a plant-specific class of endogenous small RNAs. These siRNAs have been found to regulate gene expression in most taxa including seed plants, gymnosperms, ferns and mosses. In this study, bioinformatic and experimental PCR-based approaches were used as tools to analyze TAS3 and TAS6 loci in transcriptomes and genomic DNAs from representatives of evolutionary distant non-vascular plant taxa such as Bryophyta, Marchantiophyta and Anthocerotophyta. We revealed previously undiscovered TAS3 loci in plant classes Sphagnopsida and Anthocerotopsida, as well as TAS6 loci in Bryophyta classes Tetraphidiopsida, Polytrichopsida, Andreaeopsida and Takakiopsida. These data further unveil the evolutionary pathway of the miR390-dependent TAS3 loci in land plants. We also identified charophyte alga sequences coding for SUPPRESSOR OF GENE SILENCING 3 (SGS3), which is required for generation of ta-siRNAs in plants, and hypothesized that the appearance of TAS3-related sequences could take place at a very early step in evolutionary transition from charophyte algae to an earliest common ancestor of land plants.

INTRODUCTION

Plant chromosomal loci of trans-acting small interfering RNAs (ta-siRNAs) and microRNAs (miRNAs) encode non-protein-coding and protein-coding precursor transcripts, which are synthesized by RNA polymerase II and include cap-structures and poly-(A) tails. In plants, primary miRNA transcripts forming internal imperfect hairpins are processed by a protein complex including Dicer-like protein 1 (DCL1), HYL1 and SERRATE to give RNA duplexes with 2-nucleotide 3’-overhangs, which are then terminally methylated by specific RNA methylase HEN1. One strand of such duplexes, being typically of 21 nucleotides in length and representing a mature miRNA, is selectively recruited by Argonaut (AGO) family protein to an effector complex targeting a specific RNA for AGO-mediated endonucleolytic cleavage or

translational repression (Rogers and Chen, 2013; Axtell, 2013; Bologna and Voinnet, 2014; Borges and Martienssen, 2015; Chorostecki et al., 2017).

Some specific microRNAs are able to initiate production of ta-siRNAs (and other secondary phased RNAs - phasiRNAs) by an step-by-step processing of long double-stranded RNA by DCL4 from a start point defined by miRNA-directed cleavage of a single-stranded RNA precursor in a "phased" pattern. These PHAS loci include non-coding TAS genes and genes encoding penta-tricopeptide repeat-containing proteins (PPRs), nucleotide-binding and leucine-rich repeat-containing proteins (NB-LRRs), or MYB transcription factors (Allen and Howell, 2010; Zhai et al., 2011; Xia et al., 2013; Fei et al., 2013; Axtell, 2013; Yoshikawa, 2013; Zheng et al., 2015; Komiya, 2017; Liu et al., 2018; Deng et al., 2018). Biogenesis of ta-siRNAs includes initial AGO-dependent miRNA binding at single or dual sites of the precursor transcripts and their subsequent cleavage. The further process is dependent on plant RNA-dependent RNA polymerase 6 (RDR6) and SGS3 proteins participating in the formation of dsRNA, which is then cleaved in a sequential and phased manner by DCL4 with assistance of DRB4 (dsRNA binding protein). The resulting ta-siRNAs (mostly of 21 bp in length), similar to miRNAs, are methylated by HEN1 protein (Allen and Howell, 2010; Axtell, 2013; Fei et al., 2013; Yoshikawa, 2013; Bologna and Voinnet, 2014; Komiya, 2017; Deng et al., 2018).

Arabidopsis TAS3a transcript, first identified by Allen et al. (2005), gives rise to two near-identical 21-nucleotide tasiARFs targeting the mRNAs of some auxin-responsive transcription factors (ARF2, ARF3/ETT and ARF4). Most angiosperm TAS3 primary transcripts are recognized by miR390 and cleaved by AGO7 at the 3′ target site, whereas the 5′ miRNA target site is non-cleavable. However, the number of miR390 cleavage sites, organization of tasiARF sequence blocks and phasing registers may vary among different TAS3 genes of vascular plants (Allen and Howell, 2010; Axtell, 2013; Fei et al., 2013; Zheng et al., 2015; Xia et al., 2013; 2017; de Felippes et al., 2017; Komiya, 2017; Deng et al., 2018). Moreover, miR390 may additionally target and inhibit protein-coding gene transcripts, such as StCDPK1 related to auxin-responsive pathway (Santin et al., 2017).

Previously, we described a new method for identification of plant TAS3 loci based on PCR with a pair of oligodeoxyribonucleotide primers mimicking miR390. The method was found to be efficient for dicotyledonous plants, cycads, conifers, and mosses (Krasnikova et al., 2009; 2011; 2013; Ozerova et al., 2013). Importantly, at that time the structural and functional

information on bryophyte TAS3 loci was available only for the model plant *Physcomitrella patens* (Arif et al., 2013), and we used our PCR-based approach as a phylogenetic profiling tool to identify relatives of *P. patens* TAS3 loci in 26 additional moss species of class Bryopsida and several mosses of classes Polytrichopsida, Tetraphidopsida and Andreaeopsida. Moreover, we found a putative pre-miR390 genomic sequence for an additional moss class, Oedipodipsida (Krasnikova et al., 2013). Our studies revealed that a representative of Marchantiophyta (liverwort *Marchantia polymorpha*, class Marchantiopsida) could also encode a candidate miR390 gene and a potential TAS3-like locus (Krasnikova et al., 2013). This finding extended the known evolutionary history of TAS3 loci to the proposed most basal land plant lineage (Ruhfel et al., 2014; Bowman et al., 2017). In addition, we sequenced putative pre-miR390 genomic locus for *Harpanthus flotovianus* (Marchantiophyta, class Jungermanniopsida**)** (Krasnikova et al., 2013)**.** Later, our findings of TAS3-like and miR390 loci were experimentally confirmed in the studies of the transcriptomes of Marchantiophyta plants *M. polymorpha* (Lin et al., 2016; Tsuzuki et al., 2016) and *Pellia endiviifolia* (class Jungermanniopsida**)** (Alaba et al., 2015).

New genomic and transcriptomic sequence data for basal Viridiplantae appeared in NCBI (http://ncbi.nlm.nih.gov/sra) and Phytozome (http://www.phytozome.net) databases prompted us to perform new experimental and *in silico* analyses of TAS3 loci in basal taxons of Viridiplantae. In this paper, we identified previously unrecognized TAS3 loci in classes Sphagnopsida and Anthocerotopsida, as well as composite TAS6/TAS3 loci in classes Tetraphidiopsida, Polytrichopsida, Andreaeopsida and Takakiopsida. Additionally, we revealed SGS3-coding sequences in charophytes and analyzed their evolutionary links.

MATERIALS AND METODS

Dried material for *Sphagnum angustifolium and S. girgensohnii* were taken from herbarium at Department of Biology, Moscow State University. Total DNA was extracted from dry plants using the Nucleospin Plant Extraction Kit (Macherey-Nagel, Germany) according to the protocol of the manufacturer. For PCR amplification, the following degenerate primers were used: a forward primer Spha-TASP (5’-GGCGRTAWCCYTACTGAGCTA-3’) and reverse primer Spha-TASM (5’- TAGCTCAGGAGRGATAMMBMRA-3’). For PCR, 30 cycles were used with a melting temperature of 94°C – 3', and the next steps are as follows: an annealing temperature

94°C – 20", 65°C –20", 58°C –30", and an extending temperature of 72◦C followed by a final extension at 72◦C for 5'. PCR products were separated by electrophoresis of samples in a 1.5% agarose gel and purified using the Gel Extraction Kit (Qiagen, Germany). For cloning, the PCR-amplified DNA bands isolated from gel were ligated into pGEM-T (Promega). The resulting clones were screened by length in 1,5% agarose gel. The plasmids were used as templates in sequencing reactions with an automated sequencer (Applied Biosystems) 3730 DNA Analyzer with facilities of "Genom" (Moscow, Russia).

Sequences for comparative analysis were retrieved from NCBI (http://www.ncbi.nlm.nih.gov/), Phytozome (http://www.phytozome.net) and 1000 Plant Transcriptome Project ("1KP") (http://1kp-project.com/blast.html). Sequence similarities were analysed by NCBI Blast at http://blast.ncbi.nlm.nih.gov/BlastAlign.cgi. The presence of open reading frames within retrieved sequences was analysed at http://web.expasy.org/translate/. The nucleic acid sequences and deduced amino acid sequences were analyzed and assembled using the NCBI. Conserved domains in the amino acid sequences SGS3 were identified using the CD-Search of the NCBI.

The sequences of SGS3 protein and TAS3 nucleotide sequences were aligned by MAFFT version 7 software (Katoh et al., 2014). The phylogenetic tree was constructed by the Neighbor-Joining method with 1000 bootstrap replications in MEGA7 (Kumar et al., 2016). The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1).

RESULTS

TAS3 loci in Bryophyta (classes Sphagnopsida and Takakiopsida)

It is commonly accepted that mosses of classes Sphagnopsida and Takakiopsida represent most basal lineages in Bryophyta (Shaw et al., 2010; 2011; Rosato et al., 2016). Previously, using primers, which have allowed us to detect pre-miR390 and TAS3 loci in Bryopsida and some other moss classes, we failed to identify pre-miR390 and TAS3 genes in genus Sphagnum (Krasnikova et al., 2013). However, a predicted sequence of pri-miR390 from *Sphagnum fallax* was recently reported (Xia et al., 2017). This finding prompted us to re-evaluate the occurrence

of TAS3-like loci in Sphagnopsida. To this end, we designed a new pair of degenerated PCR primers Spha-TASP and Spha-TASM, which differed from those used previously (Krasnikova et al., 2011; 2013). As a positive control, we used plasmid DNA carrying cloned TAS3 gene of *Andreaea rupestris,* a representative of basal Bryophyta (Krasnikova et al., 2013). Like the positive control, two total DNA probes from *Sphagnum angustifolium and S. girgensohnii* gave a single main PCR product of the expected size (Fig. 1). Cloning and sequencing of these PCR fragments revealed two TAS3-like primary structures having 285 (*S. angustifolium*) and 292 (*S. girgensohnii*) bases in length and exhibiting 96% identity (e-value = 2e-131). We named these loci as Sphan-285 and Sphgi-292, (Fig. 2, Fig. S1 and Table 1). Despite that the degenerate miR390-mimicking primers were used for amplification of these loci, miR390 recognition sites in TAS3 species well corresponded to genomic loci of TAS3 extracted from genomic and transcriptomic data of other plants from genus *Sphagnum* (Fig. 2).

Peatmosses *S. angustifolium and S. girgensohnii* belong to subgenera *Cuspidata* and *Acutifolia*, respectively (Shaw et al., 2010, 2016). To extend search for TAS3-like loci inside genus *Sphagnum* we performed bioinformatics analysis of the nucleotide sequences in databases available at NCBI (Sequence Read Archive) and Phytozome (version 12.1). Phytozome has recently released genome assembly of bog moss *S. fallax* (version 0.5). Bog moss belongs to subgenus *Cuspidata* and represents the most closely related moss to *S. angustifolium* (Shaw et al., 2016). BLASTN search at Phytozome allowed us to reveal a TAS3-like locus (supercontig super_37), which has 100% identity to the TAS3 locus of *S. angustifolium* sequenced in this study (Fig. S1 and Table 1). Unexpectedly, we found an additional TAS3-like locus in *S. fallax* (transcript Sphfalx0293s0011, supercontig super_293). This TAS3 locus in bog moss has 277 nucleotides in length and showed only a distant relation to the *S. angustifolium* TAS3 (Fig. 2, Fig. S1 and Table 1).

To further analyze Sphagnopsida TAS3-related loci, we used BLAST analysis of Sequence Read Archive (SRA), which is the NCBI database collecting sequence data obtained by the use of next generation sequence (NGS) technology. Assembly of sequence reads of *S. recurvum* (subgenus *Cuspidata*) retrieved by BLAST search using *S.fallax* sequences as queries revealed two TAS3 loci (Table 1). The first locus (Sphre-283) is 283 nucleotides in length and has 98% identity to Sphan-285. The second locus (Sphre-277) shows 98% identity to

Sphfalx0293s0011 (Table 1, Fig. S1). These findings indicate that two distant TAS3 loci in species of a particular subgenus of genus *Sphagnum* are extremely similar.

We also analyzed the SRA database of subgenus *Sphagnum* (Shaw et al., 2010, 2016). It was found that *S. magellanicum* belonging to this subgenus also encode two TAS3 loci called Sphma-285 (285 nt size) and Sphma-286 (286 nt size) (Fig. S1 and Table 1). Unlike *S. fallax* and *S. recurvum*, in *S. magellanicum* TAS3 loci are more similar, showing 86% identity (Fig. 2). Both Sphma-285 and Sphma-286 had 85% identity to Sphan-285 (Fig. 2). It was found that TAS3-like locus (Sphpa) from one more representative of subgenus *Sphagnum* (*S. palustre*) exhibited 98% identity to Sphma-285 (Fig. S1 and Table 1). The SRA database also contained sequence reads of two representatives from subgenus *Subsecunda* (Shaw et al., 2010, 2016). Our BLAST analysis and subsequent assembly of retrieved reads revealed a single TAS3 locus in *S. cribrosum* (Sphcri, 291 nt size) showing 95% identity to Sphan-285 and 81% identity to Sphma-286 (Fig. 2, Fig. S1 and Table 1) and a partial TAS3-like sequence in *S. lescurii* (Fig. S1 and Table 1).

Analysis of the SRA database of *Takakia lepidozioides* (class Takakiopsida) allowed us to reveal only one TAS3-like sequence (Takle-207) (Fig. S1 and Table 1). The same sequence was revealed in a longer assembly which was found recently upon search of 1KP database (Xia et al., 2017).

Since Takakiopsida and Sphagnopsida are most basal sister lines to all other Bryophyta (Shaw et al., 2010, 2011; Rosato et al., 2016; Puttick et al., 2018), it was very interesting to compare the structural organization of Takakiopsida and Sphagnopsida TAS3 loci with other classes of Bryophyta. Our previous detailed analysis of approximately 40 TAS3 loci in Bryophyta (Krasnikova et al., 2011; 2013) showed that the general structure of moss TAS3 is similar in all taxa and fits the structural organization of *Physcomitrella patens* genes, comprising dual miR390 target sites on the 5' and 3' borders and internal monomeric tasiAP2 sequence followed by tasiARF sequence positioned in 20-30 bases. We revealed that phylogenetic tree of TAS3-like loci in Bryophyta showed clear subdivision of their sequences into two main clades (see Fig. 5 in Krasnikova et al., 2013). The first group was formed by a cluster of sequences close to *P. patens* TAS3 species PpTAS3a, PpTAS3d, and PpTAS3f, and the second one – by those close to PpTAS3b, PpTAS3c, and PpTAS3e. The recent paper on the structure of TAS3 loci in lower land plants (Xia et al., 2017) has shown the structure-functional basis for this

phylogenetic subdivision. TAS3 species of the first group (PpTAS3a/PpTAS3d/PpTAS3f cluster) were shown to form class III of TAS3-like loci and contain, in addition to the previously reported tasiAP2 and tasiARF-a2 sequences, newly discovered tasiARF-a3 sequence positioned 5' according to tasiAP2 (Fig. 3). Among TAS3 species of basal Bryophyta, *Andreaea rupestris* locus 13-Aru (Krasnikova et al., 2013) belongs to class III (Fig. 3).

BLAST comparison of *T. lepidozioides* TAS3 with known Bryopsida loci showed that Takle-207 belongs to class II of TAS3 with typical positioning of tasiAP2 and tasiARF-a2 sequences (Fig. 2 and Fig. S1). On the other hand, none of Sphagnopsida TAS3-like sequences (Table 1) showed conventional internal structural organization of the most moss TAS3 species. The only recognizable conserved site, except miR390-targeting regions, was identified as tasiARF-a2 sequence, which was found to be conserved between two very distant TAS3 loci in *S. fallax* and *S. recurvum* (Fig. 2). The mentioned above tasiARF sequences, tasiARF-a2 and tasiARF-a3, showed no sequence similarity suggesting their independent origins. These tasiRNAs were found to be formed from different strands of the TAS3 dsRNA intermediate and target different regions of ARF genes (Xia et al., 2017). Inhibition of production of both tasiARF RNAs in *P. patens* resulted in obvious developmental defects exhibited, in particular, as alterations in gametophore initiation, protonemal branch determinacy and caulonemal differentiation (Plavskin et al., 2016).

Comparison of nucleotide sequences between TAS3 species of several moss classes revealed in many plants obvious similarity of nucleotide sequence blocks including tasiAP2 site and immediate upstream 21 bp block occurring in the same 21-bp-phase (Fig. 4A). We hypothesized that this sequence block may correspond to novel previously unrecognized ta-siRNA in many moss species. This putative siRNA in its single-stranded form really presents in *P. patens* transcriptome as minus-sense 21-nucleotide ta-siRNA (see NCBI SRA accessions SRX903096-SRX903105) like tasiARF RNA (Arif et al., 2012). Thus, we speculated that novel hypothetical ta-siRNA might be produced from TAS3, and its minus-strand is complementary to uncharacterized well-conserved, protein-coding moss mRNAs which have homologs also in conifers and angiosperms (Fig. 4B; Fig. S2).

TAS3 loci in Anthocerotophyta

Taking into account the finding of TAS3-like loci in classes Sphagnopsida and Takakiopsida and previously published data (Krasnikova et al., 2013; Xia et al., 2017), one can conclude that the only remaining blind-spot in land plants with respect to TAS3 is represented by phylum Anthocerotophyta. Relationships between liverworts, mosses and hornworts are still obscure. Moreover, the question remains which bryophyte phylum is a sister line to all other land plants (Qiu, 2008; Shaw et al., 2011; Harrison, 2017; Puttick et al., 2018). Recent molecular phylogenetic analysis, in which three bryophyte lineages were resolved, revealed that a clade with mosses and liverworts could form a sister group to the tracheophytes, whereas the hornworts is sister line to all other land plants (Wickett et al., 2014). However, analyses of the plastid genome sequences suggested another branching order of the phylogenetic tree, with hornworts rather than moss/liverwort clade being a sister group to tracheophytes (Lewis et al., 1997; Samigullin et al., 2002; Ruhfel et al., 2014; Lemieux et al., 2016). Moreover, some very recent nuclear gene comparisons also suggested that hornworts could be a sister clade to tracheophytes, and liverworts plus mosses might be closer to a common ancestor of land plants (Rosato et al., 2016; Bowman et al., 2017). However, this ancestor could have more tracheophyte-like characteristics than Setaphyta (liverworts/mosses) due to secondary simplification (Puttick et al., 2018).

Analysis of the SRA database of Anthocerotophyta revealed a putative TAS3-like sequence in *Folioceros fuciformis* (family *Anthocerotaceae*). Unexpectedly, the discovered TAS3-like sequence (Folfu) was found to be 244 nucleotides in length and obviously similar to Bryophyta class III TAS3 species (Fig. 3, Fig. S3 and Table 2). The identity of Folfu to some moss TAS3 sequences exceeds 80% being therefore even higher than between some related Bryopsida species (Fig. 3). Thus these data clearly indicate a close relation of TAS3 in Anthocerotophyta to Bryophyta TAS3 (excepting Sphagnopsida).

TAS3 loci in Marchantiophyta

Some of the recent molecular phylogenetic reconstructions suggested that Marchantiophyta species could represent a sister clade to all other land plants (see above). Therefore, finding and comparative analyses of TAS3 loci in this taxon represented a significant interest for understanding early events in TAS3 evolution. In contrast to class Marchantiopsida, where putative TAS3 and pre-miR390 loci were previously identified (Krasnikova et al., 2013;

Lin et al., 2016; Tsuzuki et al., 2016), for class Jungermanniopsida only potential pre-miR390 loci were found in *Pellia endiviifolia* and *Harpanthus flotovianus* (Krasnikova et al., 2013; Alaba et al., 2015). Assuming that miR390 was found to be among eight most conserved miRNA species in land plants (Xia et al., 2013; You et al., 2017; Liu et al., 2018), Jungermanniopsida could be expected to encode TAS3 loci.

To detect new potential TAS3 loci, we performed BLAST analysis of the SRA database for species of class Jungermanniopsida using *Marchantia polymorpha* TAS3 sequence (1-Mpo) as a query. Using this approach we revealed a set of reads and assembled a single TAS3-like locus (Pelen-192) for *Pellia endiviifolia* (192 nt size). In addition, TAS3 locus of 226 nucleotides in length was found in *Metzgeria crassipilis* (Metcr-226) (Fig. 5, Table 2, Fig. S3). The latter locus was also recently revealed in a search of 1KP database (Xia et al., 2017).

TAS3 1-Mpo sequence was further used for BLAST analysis of other Marchantiopsida sequences available at the NCBI SRA database. As a result, we retrieved sequence reads and assembled five full-length TAS3-like sequences in *Plagiochasma appendiculatum* (Plaap-247), *Dumortiera hirsuta* (Dumhi-243), *Marchantia emarginata* (Marem-262), *Ricciocarpos natans* (Ricna-235) and *Conocephalum japonicum* (Conja-252) (Fig. 5, Table 2, Fig. S3). Recent bioinformatics analysis of 1KP database revealed three additional full-length TAS3-like sequences in *Conocephalum conicum*, *Lunularia cruciata* and *Marchantia paleaceae* (Xia et al., 2017) (Table 2). Thus, totally 11 TAS3-like loci have been found in Marchantiophyta.

TAS6 loci in Bryophyta

Previous studies of *P. patens* revealed three novel non-coding PHAS loci (TAS6) which were located in rather close genomic proximity to PpTAS3 loci (PpTAS3a, PpTAS3d, and PpTAS3f) and expressed as common RNA precursors with these TAS3 species (Cho et al., 2012; Arif et al., 2012, 2013). Moreover, miR529 and miR156 were suggested to influence accumulation of ta-siRNAs specific not only for TAS6, but also for PpTAS3a (Cho et al., 2012). We have found that localization of TAS6 loci close to TAS3 genes in common transcripts was not unique for *P. patens* (subclass Funariidae), since these loci were also found to be encoded by three other mosses of subclasses Bryidae and Dicranidae (Krasnikova et al., 2013).

For further search of the combined TAS6/TAS3 loci, we performed bioinformatics analysis of 1KP database. Although nucleotide sequences of miR156 and related miR529, as well as their recognition sites in RNA transcripts, are highly conserved among land plants (Morea et al., 2016; Axtell & Meyers, 2018), the internal sequences between dual miR156/miR529 recognition sites show little or no similarity even between different TAS6 loci of *P. patens* (Arif et al., 2012). So we used, as queries for BLAST search, the individual full-length TAS6/TAS3 loci including most characterized locus encoding PpTAS3a (Fig. 6), as well as those for PpTAS3d and PpTAS3f. First, it was found that in addition to four previously found Bryopsida species, encoding TAS6/TAS3 loci, these loci could be revealed in basal subclasses Timmiidae (*Timmia austriaca)* and Diphysciidae (*Diphyscium foliosum*) (Shaw et al., 2011) (Table 3, Fig. S4). List of TAS6/TAS3 loci in other moss subclasses was also significantly extended: we found 18 new loci in Bryidae, seven loci in Dicranidae and four loci in Funariidae (Table 3, Fig. S4). These novel loci showed recognizable but varying sequence similarities to the PpTAS3a-containing locus (Fig. 6). Second, most importantly, putative TAS6/TAS3 loci were revealed in 4 basal classes of Bryophyta, namely, Tetraphidiopsida, Polytrichopsida, Andreaeopsida and Takakiopsida (Table 3, Fig. S4). These novel loci had a similar organization to Bryopsida TAS6/TAS3 species (Fig. 6). However, no TAS6-specific sequence signatures were found in the vicinity of genomic *S. fallax* and *M. polymorpha* TAS3 loci upon analysis of the corresponding Phytozome genome contigs.

Phylogeny of SGS3 as a characteristic molecular component of TAS3 pathway

It was shown that some species of green algae could encode ancient types of dicer-like proteins, RDRs, and AGOs. On the other hand, no encoded SGS3 proteins were revealed for these algae (Zheng et al., 2015). Since SGS3 was found to be essential for production of tasiARF RNAs in moss *P. patens* (Plavskin et al., 2016), we performed sequence analyses to identify possible SGS3 genes in charophytes. For identification of SGS3 protein orthologs among land nonvascular plants and charophytes, we used as a query the most conserved region of *P. patens* SGS3 including short zinc binding zf-XS domain and RNA recognition XS domain (Bateman, 2002; Zhang & Trudeau, 2008). Importantly, the short N-terminal zf-XS domain is characteristic for functional SGS3 proteins, since the XS domain-containing protein of *Selaginella moellendofii* lacking TAS-generating machinery (Banks et al., 2011) possesses no zf-XS domain

upstream of XS domain and instead contains the C-terminal RING zf region (see NCBI accession XP_002979112). However, it should be noted that the lack of TAS3 pathway and SGS3 is not universal for lycophytes (Xia et al., 2017).

In addition to class Bryopsida, SGS3 protein sequences were revealed for members of classes Marchantiopsida, Jungermanniopsida, Anthocerotopsida, Takakiopsida and Sphagnopsida (Fig. 7 and Fig. S5). Most importantly, search for the SGS3 coding sequences in transcriptomes of four charophyte classes (Zygnemophyceae, Coleochaetophyceae, Charophyceae, and Klebsormidiophyceae) also revealed the SGS3-like proteins in representatives of all these taxa (Fig. 7, Fig. S5, Fig. S6). This observation was in agreement with the fact that SGS3-like coding sequence was found in the fully sequenced and annotated genome of *Klebsormidium nitens* (NCBI accession GAQ92898) (Hori et al., 2014). Moreover, the characteristic motifs of land plant SGS3 proteins (Bateman, 2002) were revealed in the protein sequences from charophyte algae (Fig. S5, Fig. S6).

Importantly, in the dendrogram based on comparisons of 27 aligned SGS3 protein sequences, the position of charophytes (Fig. 7) nearly corresponded to the commonly accepted Viridiplantae phylogenetic tree (Shaw et al., 2011; Delwiche & Cooper, 2015; Harrision, 2017), where class Zygnemophyceae (*Spirogyra pratensis*) was a sister group for all land plants. It has become clear that evolving the SGS3-like genes was not directly connected to the appearance of TAS loci in Viridiplantae, since Chlorophyta species, lacking SGS3, encode not only critical enzyme machinery including DCLs, RDRs, and AGOs (You et al., 2017), but also PHAS loci (Zheng et al., 2015). Despite our extensive searches, no SGS3 genes could be identified also in brown and red algae, and this is in agreement with previously published data on green algae (Zheng et al., 2015).

DISCUSSION

Our current analyses revealed previously undiscovered TAS3 loci in bryophytes from classes Takakiopsida, Sphagnopsida and Anthocerotopsida. In *Folioceros fuciformis* (family *Anthocerotaceae*) we found a TAS3-like sequence which is obviously similar to Bryopsida class III TAS3 species (Fig. 3), whereas Takakiopsida TAS3 locus showed relatedness to class II of TAS3 with typical positioning of tasiAP2 and tasiARF-a2 sequences and no tasiARF-a3 sites

(Xia et al., 2017). Unexpectedly, all predicted Sphagnopsida TAS3 loci showed no conventional internal structural organization of the moss class II and class III TAS3 species. Excepting miR390-targeting regions, the only recognizable conserved site was tasiARF-a2 sequence (Fig. 2).

It was shown that structural organization of Marchantiopsida and Jungermanniopsida TAS3 loci were quite similar, whereas Marchantiophyta TAS3 species were obviously different from those of Bryophyta. These loci were proposed to belong to TAS3 class I with species containing two conserved sequence blocks which presumably represent functional ta-siRNAs (Xia et al., 2017). One of these blocks was found in the vicinity of the 3'-terminal miR390 binding site and corresponded to Bryopsida tasi-AP2 sequence (Krasnikova et al. 2013), whereas another one (tasiARF-a1), unique among lower land plants, was located closer to the 5'-terminal miR390 binding site in Marchantiopsida and Jungermanniopsida TAS3 (Tsuzuki et al., 2016; Xia et al., 2017) (Fig. 5, Fig. S3).

Assuming occurrence of tasiARFs as potential products of TAS3 in all main lineages of land plants (Marchantiophyta, Anthocerotophyta and Bryophyta), recent paper proposed that the earliest function of TAS3 could contribute to the production of ta-siRNAs targeting ARF genes, and, since green algae encode no ARF genes, TAS3 likely appeared first in land plants (Xia et al., 2017). However, extensive comparative sequence analysis showed that charophyte algae representing the sister group to all land plants (colonized terrestrial environments approximately 480 million years ago, see for references Fischer (2018) and Puttick et al. (2018)) could also encode ARF-like proteins including all sequence domains typical for bryophyte and angiosperm ARFs (Mutte et al., 2017). Moreover, our current data showed that TAS3-like loci are encoded by the representatives of all main taxa among non-vascular plants. These observations suggest that the TAS3 evolution started in a common ancestor of land plants, likely belonging to a still unknown lineage of charophytes. Identification of the canonical motifs of land plant SGS3 in charophyte proteins (see above) indirectly supports this speculation. However, it should be kept in mind that evolving the SGS3-like genes could not be connected solely to the appearance of PHAS loci in Viridiplantae, since green algae and brown algae species were found to encode not only essential silencing machinery enzymes including DCLs, RDRs and AGOs, but also PHAS loci (Billoud et al., 2014; Zheng et al., 2015; Singh et al., 2015; Zhang et al., 2016; Dueck et al., 2016; You et al., 2017; Cock et al., 2017). Finally, it can be proposed that the failure to identify

charophyte TAS3 loci may be related to (i) the incompleteness of the available sequence data; (ii) evolving by charophytes the one-hit TAS3 genes (de Felippes et al., 2017); or (iii) the use of miRNA species with sequences other than land plant miR390 for TAS precursor processing.

CONCLUSIONS

The current data on the structural organization of TAS3-like loci in all main classes of land non-vascular plants reveal three types of TAS3 loci, namely, (i) Bryopsida-like TAS3 (classes II and III, Xia et al., 2017) found in Bryophyta plants (excepting Sphagnopsida) and Anthocerotopsida, (ii) Marchantiophyta-like TAS3 (class I, Xia et al., 2017) and (iii) Sphagnopsida-like TAS3 (this paper). Clearly recognizable common ta-siRNAs is represented in these TAS3 types by tasiARF sequences. Occurrence of primitive SGS3 and ARF genes in charophytes (Mutte et al., 2017 and this paper) supports the idea that TAS3-like genes first appeared in the hypothetical common precursor of land plants (Xia et al., 2017) to regulate auxin signaling (Leiser, 2018).

ACKNOWLEDGEMENTS

We thank researchers who contributed samples used in this study to the 1KP initiative.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences: The new sequences generated for this study are available as a nexus file in the Supplemental Material. All sequences used in this study are available on GeneBank (new sequences accession numbers MF682529 and MF682530).

REFERENCES

Alaba S, Piszczalka P, Pietrykowska H, Pacak AM, Sierocka I, Nuc PW, Singh K, Plewka P, Sulkowska A, Jarmolowski A, Karlowski WM, Szweykowska-Kulinska Z. **2015.** The liverwort *Pellia endiviifolia* shares microtranscriptomic traits that are common to green algae and land plants. *New Phytology* **206**:352–367. DOI: 10.1111/nph.13220.

Allen E, Xie Z, Gustafson A, Carrington J. 2005. MicroRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* **121**:207–221.

Allen E, Howell M. 2010. miRNAs in the biogenesis of trans-acting siRNAs in higher plants. *Seminars in Cell and Developmental Biology* **21**:798–804. DOI 10.1016/j.semcdb.2010.03.008.

Arif MA, Fattash I, Ma Z, Cho SH, Beike AK, Reski R, Axtell MJ, Frank W. 2012. DICER-LIKE3 activity in Physcomitrella patens DICER-LIKE4 mutants causes severe developmental dysfunction and sterility. *Mol. Plant* **5**:1281-1294. doi: 10.1093/mp/sss036.

Arif MA, Frank W, Khraiwesh B. 2013. Role of RNA interference (RNAi) in the Moss Physcomitrella patens. *Int. J. Mol. Sci*. **14**:1516-1540. doi: 10.3390/ijms14011516.

Axtell MJ. 2013. Classification and comparison of small RNAs from plants. *Annu. Rev. Plant Biol.* **64**:137–159. DOI 10.1146/annurev-arplant-050312-120043.

Axtell MJ & Meyers BC. 2018. Revisiting criteria for plant miRNA annotation in the era of big data. Plant Cell doi: 10.1105/tpc.17.00851.

Bateman, A. 2002. The SGS3 protein involved in PTGS finds a family. *BMC Bioinformatics* **3**: 21. doi: 10.1186/1471-2105-3-21.

Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, dePamphilis C, Albert VA, Aono N, Aoyama T, Ambrose BA, Ashton NW, Axtell MJ, Barker E, Barker MS, Bennetzen JL, Bonawitz ND, Chapple C, Cheng C, Correa LG, Dacre M, DeBarry J, Dreyer I, Elias M, Engstrom EM, Estelle M, Feng L, Finet C, Floyd SK, Frommer WB, Fujita T, Gramzow L, Gutensohn M, Harholt J, Hattori M, Heyl A, Hirai T, Hiwatashi Y, Ishikawa M, Iwata M, Karol KG, Koehler B, Kolukisaoglu U, Kubo M, Kurata T, Lalonde S, Li K, Li Y, Litt A, Lyons E, Manning G, Maruyama T, Michael TP, Mikami K, Miyazaki S, Morinaga S, Murata T, Mueller-Roeber B, Nelson DR, Obara M, Oguri Y, Olmstead RG, Onodera N, Petersen BL, Pils B, Prigge M, Rensing SA, Riaño-Pachón DM, Roberts AW, Sato Y, Scheller HV, Schulz B, Schulz C, Shakirov EV, Shibagaki N, Shinohara N, Shippen DE, Sørensen I, Sotooka R, Sugimoto N, Sugita M, Sumikawa N, Tanurdzic M, Theissen G, Ulvskov P, Wakazuki S, Weng JK, Willats WW, Wipf D, Wolf PG, Yang L, Zimmer AD, Zhu Q, Mitros T, Hellsten U, Loqué D, Otillar R, Salamov A, Schmutz J, Shapiro H, Lindquist E, Lucas S, Rokhsar D, Grigoriev IV. 2011. The selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**:960-963.

Billoud B, Nehr Z, Le Bail A, Charrier B. 2014. Computational prediction and experimental validation of microRNAs in the brown alga Ectocarpus siliculosus. *Nucleic Acids Res*. **42**:417-429. doi: 10.1093/nar/gkt856.

Bologna N, Voinnet O. 2014. The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annu Rev Plant Biol.* **65**:473-503. DOI 10.1146/annurev-arplant-050213-035728.

Borges F & Martienssen RA. 2015. The expanding world of small RNAs in plants. *Nat. Rev. Mol. Cell. Biol.* **16**:727-741. doi: 10.1038/nrm4085.

Bowman JL, Kohchi T, Yamato KT, Jenkins J, Shu S, Ishizaki K, Yamaoka S, Nishihama R, Nakamura Y, Berger F, Adam C, Aki SS, Althoff F, Araki T, Arteaga-Vazquez MA, Balasubrmanian S, Barry K, Bauer D, Boehm CR, Briginshaw L, Caballero-Perez J, Catarino B, Chen F, Chiyoda S, Chovatia M, Davies KM, Delmans M, Demura T, Dierschke T, Dolan L, Dorantes-Acosta AE, Eklund DM, Florent SN, Flores-Sandoval E, Fujiyama A, Fukuzawa H, Galik B, Grimanelli D, Grimwood J, Grossniklaus U, Hamada T, Haseloff J, Hetherington AJ, Higo A, Hirakawa Y, Hundley HN, Ikeda Y, Inoue K, Inoue SI, Ishida S, Jia Q, Kakita M, Kanazawa T, Kawai Y, Kawashima T, Kennedy M, Kinose K, Kinoshita T, Kohara Y, Koide E, Komatsu K, Kopischke S, Kubo M, Kyozuka J, Lagercrantz U, Lin SS, Lindquist E, Lipzen AM, Lu CW, De Luna E, Martienssen RA, Minamino N, Mizutani M, Mizutani M, Mochizuki N, Monte I, Mosher R, Nagasaki H, Nakagami H, Naramoto S, Nishitani K, Ohtani M, Okamoto T, Okumura M, Phillips J, Pollak B, Reinders A, Rövekamp M, Sano R, Sawa S, Schmid MW, Shirakawa M, Solano R, Spunde A, Suetsugu N, Sugano S, Sugiyama A, Sun R, Suzuki Y, Takenaka M, Takezawa D, Tomogane H, Tsuzuki M, Ueda T, Umeda M, Ward JM, Watanabe Y, Yazaki K, Yokoyama R, Yoshitake Y, Yotsui I, Zachgo S, Schmutz J. 2017. Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. *Cell* **171**:287-304.e15. doi: 10.1016/j.cell.2017.09.030.

Cho SH, Coruh C, Axtell MJ. 2012. miR156 and miR390 regulate tasiRNA accumulation and developmental timing in *Physcomitrella patens*. *Plant Cell* **24**:4837–4849. doi: 10.1105/tpc.112.103176.

Chorostecki U, Moro B, Rojas AML, Debernardi JM, Schapire AL, Notredame C, Palatnik JF. 2017. Evolutionary Footprints Reveal Insights into Plant MicroRNA Biogenesis. *Plant Cell* **29**:1248-1261. doi: 10.1105/tpc.17.00272.

Cock JM, Liu F, Duan D, Bourdareau S, Lipinska AP, Coelho SM, Tarver JE. 2017. Rapid Evolution of microRNA Loci in the Brown Algae. *Genome Biol. Evol*. **9**:740-749. doi: 10.1093/gbe/evx038.

Delwiche CF & Cooper ED. 2015. The evolutionary origin of terrestrial life. *Curr. Biol*. **25**: R899–R910

Deng P, Muhammad S, Cao M, Wu L. 2018. Biogenesis and regulatory hierarchy of phased small interfering RNAs in plants. *Plant Biotechnology Journal.* doi: 10.1111/pbi.12882.

Dueck A, Evers M, Henz SR, Unger K, Eichner N, Merkl R, Berezikov E, Engelmann JC, Weigel D, Wenzl S, Meister G. 2016. Gene silencing pathways found in the green alga Volvox carteri reveal insights into evolution and origins of small RNA systems in plants. *BMC Genomics* **17**:853.

Fei Q, Xia R, Meyers B. 2013. Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* **25**:2400-2415.

de Felippes FF, Marchais A, Sarazin A, Oberlin S, Voinnet O. 2017. A single miR390 targeting event is sufficient for triggering TAS3-tasiRNA biogenesis in Arabidopsis. *Nucleic Acids Res*. **45**:5539-5554. doi: 10.1093/nar/gkx119.

Fischer WW. 2018. Early plants and the rise of mud. *Science* **359**:994-995.

Harrison CJ. 2017. Development and genetics in the evolution of land plant body plans. *Philosophical Transactions of the Royal Society B: Biological Sciences* **372** DOI: 10.1098/rstb.2015.0490.

Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N. 2014. Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. *Nat. Commun*. **5**:3978. doi: 10.1038/ncomms4978.

Katoh K & Standley DM. 2014. MAFFT Multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780. doi:10.1093/molbev/mst010.

Krasnikova M, Milyutina I, Bobrova V, Troitsky A, Solovyev A, Morozov S. 2009. Novel miR390-dependent transacting siRNA precursors in plants revealed by a PCR-based experimental approach and database analysis. *Journal of Biomedicine and Biotechnology*. Article ID 952304. doi: 10.1155/2009/952304.

Krasnikova M, Milyutina I, Bobrova V, Ozerova L, Troitsky A, Solovyev A, Morozov S. 2011. Molecular diversity of mir390-guided trans-acting siRNA precursor genes in lower land plants: experimental approach and bioinformatics analysis. *Sequencing*. Article ID 703683. doi: 10.1155/2011/703683.

Krasnikova M, Goryunov D, Troitsky A, Solovyev A, Ozerova L, Morozov S. 2013. Peculiar evolutionary history of miR390-guided TAS3-like genes in land plants. *Scientific World Journal.* Article ID 924153. doi: 10.1155/2013/924153.

Komiya R. 2017. Biogenesis of diverse plant phasiRNAs involves an miRNA-trigger and Dicer-processing. *J Plant Res.* **130**:17-23. DOI 10.1007/s10265-016-0878-0.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**:1870-1874.

Leyser O. 2018. Auxin signaling. *Plant Physiology* **176:**465-479. doi: 10.1104/pp.17.00765.

Lemieux C, Otis C, Turmel M. 2016. Comparative Chloroplast Genome Analyses of Streptophyte Green Algae Uncover Major Structural Alterations in the Klebsormidiophyceae,

Coleochaetophyceae and Zygnematophyceae. *Front. Plant Sci*. **7**:697. doi: 10.3389/fpls.2016.00697.

Lewis LA., Mishler BD, Vilgalys R. 1997. Phylogenetic relationships of the liverworts (Hepaticae), a basal embryophyte lineage, inferred from nucleotide sequence data of the chloroplast gene *rbcL*, *Mol. Phylogenet. Evol*., **7**: 377-393.

Lin PC, Lu CW, Shen BN, Lee GZ, Bowman JL, Arteaga-Vazquez MA, Liu LY, Hong SF, Lo CF, Su GM, Kohchi T, Ishizaki K, Zachgo S, Althoff F, Takenaka M, Yamato KT, Lin SS. 2016. Identification of miRNAs and Their Targets in the Liverwort Marchantia polymorpha by Integrating RNA-Seq and Degradome Analyses. *Plant Cell Physiology* **57**:339-358. doi: 10.1093/pcp/pcw020.

Liu H, Yu H, Tang G, Huang T. 2018. Small but powerful: function of microRNAs in plant development. *Plant Cell Reports* doi: 10.1007/s00299-017-2246-5.

Morea EG, da Silva EM, e Silva GF, Valente GT, Barrera Rojas CH, Vincentz M, Nogueira FT. 2016. Functional and evolutionary analyses of the miR156 and miR529 families in land plants. *BMC Plant Biology* **16**:40. doi: 10.1186/s12870-016-0716-5.

Mutte S, Kato H, Rothfels C, Melkonian M, Wong G K-S, Weijers D. 2017. Origin and evolution of the nuclear auxin response system. *bioRxiv* 220731; doi: https://doi.org/10.1101/220731

Ozerova L, Krasnikova M, Troitsky A, Solovyev A, Morozov S. 2013. TAS3 genes for small ta-siARF RNAs in plants belonging to subtribe Senecioninae: occurrence of prematurely terminated RNA precursors. *Mol Gen Mikrobiol Virusol* (Moscow) **28**:79-84.

Plavskin Y, Nagashima A, Perroud PF, Hasebe M, Quatrano RS, Atwal GS, Timmermans MC. 2016. Ancient trans-Acting siRNAs Confer Robustness and Sensitivity onto the **Auxin** Response. *Dev. Cell* **36**:276-289. doi: 10.1016/j.devcel.2016.01.010.

Puttick MN, Morris JL, Williams TA, Cox CJ, Edwards D, Kenrick P, Pressel S, Wellman CH, Schneider H, Pisani D, Donoghue PCJ. 2018. The Interrelationships of Land Plants and the Nature of the Ancestral Embryophyte. *Curr Biol.* **28**:733-745. doi: 10.1016/j.cub.2018.01.063.

Samigullin TK, Yacentyuk SP, Degtyaryeva GV, Valieho-Roman KM, Bobrova VK, Capesius I, Martin WF, Troitsky AV, Filin VR, Antonov AS. 2002. Paraphyly of bryophytes and close relationship of hornworts and vascular plants inferred from chloroplast rDNA spacers sequence analysis, *Arctoa* **11**: 31-43.

Qiu YL. 2008. Phylogeny and evolution of charophytic algae and land plants. *J. Syst. Evol*. **46**, 287–306. doi:10.3724/SP.J.1002.2008.08035.

Rogers K, Chen X. 2013. Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell*, **25**:2383-2399. DOI 10.1105/tpc.113.113159.

Rosato M, Kovařík A, Garilleti R, Rosselló JA. 2016. Conserved Organisation of 45S rDNA Sites and rDNA Gene Copy Number among Major Clades of Early Land Plants. *PLoS One* **11**:e0162544. doi: 10.1371/journal.pone.0162544.

Ruhfel BR, Gitzendanner MA, Soltis PS, Soltis DE, Burleigh JG. 2014. From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evol. Biol*. **14**:23. doi: 10.1186/1471-2148-14-23.

Santin F, Bhogale S, Fantino E, Grandellis C, Banerjee AK, Ulloa RM. 2017. Solanum tuberosum StCDPK1 is regulated by miR390 at the posttranscriptional level and phosphorylates the auxin efflux carrier StPIN4 in vitro, a potential downstream target in potato development. *Physiol. Plant* **159**:244-261. doi: 10.1111/ppl.12517.

Shaw AJ, Cox CJ, Buck WR, Devos N, Buchanan AM, Cave L, Seppelt R, Shaw B, Larraín J, Andrus R, Greilhuber J, Temsch EM. 2010. Newly resolved relationships in an early land plant lineage: Bryophyta class Sphagnopsida (peat mosses). *American Journal of Botany* **97**:1511–1531. doi: 10.3732/ajb.1000055.

Shaw AJ, Szövényi P, Shaw B. 2011. Bryophyte diversity and evolution: windows into the early evolution of land plants. *American Journal of Botany* **98**:352–369. doi: 10.3732/ajb.1000316.

Shaw AJ, Devos N, Liu Y, Cox CJ, Goffinet B, Flatberg KI, Shaw B. 2016. Organellar phylogenomics of an emerging model system: Sphagnum (peatmoss). *Ann. Bot*. **118**:185-196. doi: 10.1093/aob/mcw086.

Singh RK, Gase K, Baldwin IT, Pandey SP. 2015. Molecular evolution and diversification of the Argonaute family of proteins in plants. *BMC Plant Biology* **15**:23. doi: 10.1186/s12870-014-0364-6.

Tsuzuki M, Nishihama R, Ishizaki K, Kurihara Y, Matsui M, Bowman JL, 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 Cell Physiology* **57**: 359–372. doi: 10.1093/pcp/pcv182.

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS. 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc. Natl. Acad. Sci. U S A* **111**:E4859-68. doi: 10.1073/pnas.1323926111.

Xia R, Meyers BC, Liu Z, Beers EP, Ye S, Liu Z. 2013. MicroRNA superfamilies descended from miR390 and their roles in secondary small interfering RNA Biogenesis in Eudicots. *Plant Cell* **25**:1555–1572. doi: 10.1105/tpc.113.110957.

Xia R, Xu J, Meyers BC. 2017. The emergence, evolution, and diversification of the miR390 TAS3 ARF pathway in land plants. *Plant Cell* **29**:1232-1247. doi: 10.1105/tpc.17.00185.

You C, Cui J, Wang H, Qi X, Kuo L-Y, Ma H, Gao L, Mo B, Chen X. 2017. Conservation and divergence of smallRNA pathways and microRNAs in plants. *Genome Biology* **18**:158 doi: 10.1186/s13059-017-1291-2.

Yoshikawa M. 2013. Biogenesis of *trans*-acting siRNAs, endogenous secondary siRNAs in plants. *Genes Genet Syst*. 88:77–84.

Zhai J, Jeong D, De Paoli E, Park S, Rosen B, Li Y, González A, Yan Z, Kitto S, Grusak M, Jackson S, Stacey G, Cook D, Green P, Sherrier D, Meyers M. 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev*. **25**:2540-2553. DOI 10.1101/gad.177527.111.

Zhang D, Trudeau VL. 2008. The XS domain of a plant specific SGS3 protein adopts a unique RNA recognition motif (RRM) fold. *Cell Cycle* **7**:2268-2270.

Zheng Y, Wang Y, Wu J, Ding B, Fei Z. 2015. A dynamic evolutionary and functional landscape of plant phased small interfering RNAs. *BMC Biol*. **13**:32. doi: 10.1186/s12915-015-0142-4.

669 **FIGURE LEGENDS**
670
671 **Figure 1:** Analysis of PCR products in 1.5% agarose gel. Amplification of genomic DNA
672 sequences flanked by miR390 and miR390* sites. PCR products were obtained on genomic
673 DNAs with degenerate primers. *Sphagnum angustifolium* (1), *Sphagnum girgensohnii* (2),
674 *Andreaea rupestris* (3). (M), DNA size markers including bands ranging from 100 bp to 1000 bp
675 with 100 bp step (Sibenzyme).
676
677 **Figure 2:** Analysis of TAS3 loci in genus *Sphagnum.*
678 **(A)** Multiple sequence alignment of available nucleotide sequences of TAS3-like loci from
679 mosses of genus *Sphagnum* along with TAS3 loci of *Takakia lepidozioides* and *Folioceros*
680 *fuciformis*. Alignment was generated at MAFFT6 program. The miR390 target sites are in
681 yellow, and putative tasiARF-a2 site is in green; tasiAP2 site is in blue. **(B)** The minimal
682 evolution phylogenetic tree based on analysis of the aligned TAS3 genes from mosses of genus
683 *Sphagnum*. This tree was generated according MAFFT6 program. For full plant names and
684 accession numbers see Table 1.
685
686 **Figure 3:** Multiple sequence alignment of selected available nucleotide sequences of class III
687 TAS3-like loci from Bryopsida mosses along with TAS3 locus of *Folioceros fuciformis*.
688 Alignment was generated at MAFFT6 program. The miR390 target sites are in yellow; putative
689 tasiARF-a2 site is in green; tasiAP2 is in blue, and tasiARF-a3 is shaded. Note that all sequences
690 are cut in the 5'-terminal TAS3 area to exclude non-aligned regions.
691
692 **Figure 4:** Novel potential ta-siRNA in Bryopsida plants.
693 **(A)** Multiple sequence alignments of nucleotide sequence blocks including tasiAP2 site and
694 preceding 21 bp site of putative ta-siRNA of *Andreaea rupestris* TAS3-like locus WOGB_
695 2010369 with the corresponding transcript sequences of moss TAS3 loci. BLASTN was used at
696 1KP blast site. For the complete TAS3 transcript sequences see Xia et al., 2017. The putative
697 tasiAP2 site is in blue, and preceding putative ta-siRNA site is in violet. *Andreaea*1 - *Andreaea*
698 *rupestris* WOGB_ 2010369; *Andreaea2* - *Andreaea_rupestris* WOGB_ 2002765; *Tetraphis1* -
699 *Tetraphis_pellucida* HVBQ_2019753; *Tetraphis2* - *Tetraphis_pellucida* HVBQ_2011866;
700 *Tetraphis3* - *Tetraphis_pellucida* HVBQ _2005644; *Plagiomnium* - *Plagiomnium_insigne*

BGXB_2010105; *Leucobryum* - *Leucobryum_glaucum* RGKI_2062694; *Racomitrium* - *Racomitrium_varium* RDOO_2117129; *Philonotis Philonotis fontana* ORKS_2058791; *Dicranum* - *Dicranum_scoparium* NGTD_2078536; *Encalypta – Encalypta streptocarpa* KEFD_2058811; *Ceratodon* - *Ceratodon_purpureus* FFPD_2044193; *Niphotrichum* - *Niphotrichum_elongatum* ABCD_2000143; *Funaria* - *Funaria* sp. XWHK _2042016; *Schwetschkeop – Schwetschkeopsis fabronia* IGUH_2166854; *Aulacomnium* - *Aulacomnium heterostichum* WNGH_2088134; *Syntrichia* - *Syntrichia_princeps* GRKU_2074985; *Diphyscium1* - *Diphyscium_foliosum* AWOI_2069791; *Diphyscium2* - *Diphyscium_foliosum* AWOI_2006305; *Pohlia* - *Pohlia nutans* GACA01023180; *Bryum argenteum* Unigene33538 GCZP01053768; PHYSCO TAS3A – *Physcomitrella patens* TAS3a BK005825.

(B) The example target transcript sequence (lower case letters) from *Leucobryum albidum* (VMXJ_2127900) is presented alongside with the predicted novel ta-siRNA shown in blue and above the transcript. Lower case letters in ta-siRNA indicate non-Watson-Crick pairing positions. Complementary mRNA sequences are in violet; conserved amino acid sequence signatures are in yellow. Numbers indicate codon positions of the target gene.

Figure 5: Analysis of TAS3 loci in Marchantiophyta plants.

(A) Multiple sequence alignment of available nucleotide sequences of TAS3-like loci from Marchantiophyta plants along with TAS3 loci of *Takakia lepidozioides* and *Folioceros fuciformis*. Alignment was generated at MAFFT6 program. The miR390 target sites are in yellow, and putative tasiARF-a2 site is in green; tasiAP2 site is in blue. **(B)** The minimal evolution phylogenetic tree based on analysis of the aligned TAS3 genes from Marchantiophyta plants. This tree was generated according MAFFT6 program. For full plant names and accession numbers see Table 2.

Figure 6: Pairwise sequence comparisons of selected nucleotide sequences of TAS6/TAS3-like loci from mosses with TAS6/TAS3 of *Physcomitrella patens* precursor RNA (accession JN674513). BLASTN was used at 1KP blast site. The miR390 target sites are in yellow; putative miR156/miR529 sites are underlined; tasiAP2 is in blue; putative tasiARF-a2 site is in green; tasiARF-a3 is shaded.

732

733 **Figure 7:** The phylogenetic tree constructed from conserved regions of SGS3 protein sequences
734 from 27 selected streptophytes by the Neighbor-Joining method with 1000 bootstrap replications.
735 There were a total of 419 positions in the final dataset. Evolutionary analyses were conducted in
736 MEGA7 (Kumar et al., 2016). Bootstrap support values ≥50% are shown. The evolutionary
737 distances were computed using the JTT matrix-based method and are in the units of the number
738 of amino acid substitutions per site. The rate variation among sites was modeled with a gamma
739 distribution (shape parameter = 1). The tree was rooted at two Klebsormidiaceae charopytes.
740 Accession numbers from NCBI ot PHYTOZOME data banks see in Fig. S5.

Figure 1

Analysis of PCR products in 1.5% agarose gel.

Amplification of genomic DNA sequences flanked by miR390 and miR390* sites. PCR products were obtained on genomic DNAs with degenerate primers. *Sphagnum angustifolium* (1), *Sphagnum girgensohnii* (2), *Andreaea rupestris* (3). (M), DNA size markers including bands ranging from 100 bp to 1000 bp with 100 bp step (Sibenzyme).

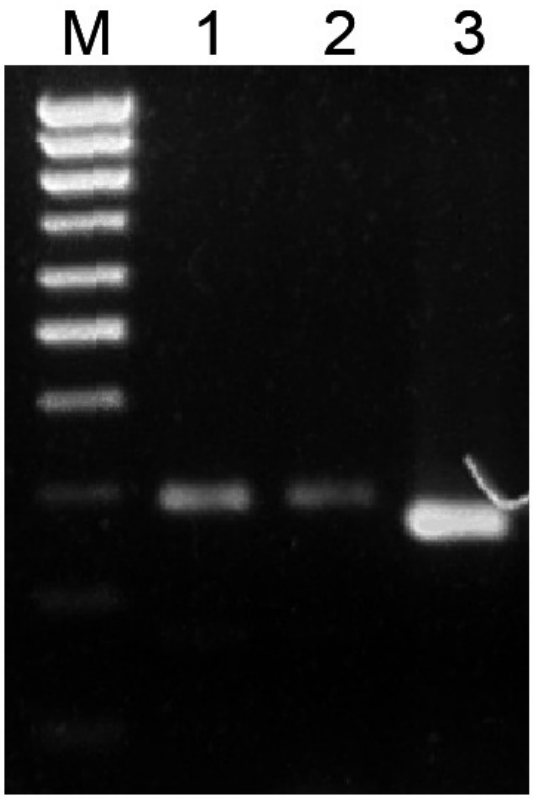

M
1
2
3

Figure 2(on next page)

Analysis of TAS3 loci in genus *Sphagnum.*

(A) Multiple sequence alignment of available nucleotide sequences of TAS3-like loci from mosses of genus *Sphagnum* along with TAS3 loci of *Takakia lepidozioides* and *Folioceros fuciformis*. Alignment was generated at MAFFT6 program. The miR390 target sites are in yellow, and putative tasiARF-a2 site is in green; tasiAP2 site is in blue. **(B)** The minimal evolution phylogenetic tree based on analysis of the aligned TAS3 genes from mosses of genus *Sphagnum*. This tree was generated according MAFFT6 program. For full plant names and accession numbers see Table 1.

(A)

```
Sphan-285          GGCGGTAACCCTTCTGAGC-TAAGTTTAAAC-GGATAGGG-TTTGTGTTTTGCAAGTAGATT------TGTGTGTTTTTTAATGTCTTTT
S.fallax super_37  GGCGGTAACCCTTCTGAGC-TAAGTTTAAAC-GGATAGGG-TTTGTGTTTTGCAAGTAGATT------TGTGTGTTTTTTAATGTCTTTT
Sphfalx0293s001    GGCGGTAACCCTTCTGAGC-TAAGTTTAAAC-GGATAGGG-TTTGTGTTTTGCAAGTAGATT------TGTGTGTTTTTTAATGTCTTTT
Sphma-285          GGCGGTAACCCTTCTGAGC-TAAGTTTGAGC-GGATAGGG-TTTGTGTTTTGCAAGTAGATT------TGTGTGTTTTTTAATGTCTTTT
Sphre-283          GGCGGTAACCCTTCTGAGC-TAAGTTTAAGC-GGATAGGG-TTTGTGTTTTGCAAGTAGATT--------TGTGTTTTTTAATGTCTTTT
Sphpa              GGCGGTAACCCTTCTGAGC-TAAGTTTGAGC----TAGGG-TTTGAGCCTTGCAAGTAGAATTGGATTTGTGTGTTTTTTAATGTCTTTT
Sphgi-292          GGCGGTAACCCTTCTGAGCGTAAGTTTAAGCAAGATAGGG-TTTGTGTTTTGCAAGTAGATTTG--TGTATGTGTTTTTTAATGTCTTTC
Sphcri             GGCGGTAACCCTTCTGAGC-TAAGTTTAAACAAGATAGGGTTTTGTGTTTTGCAAGTAGATTTG--TGTGTGTG-TTTTTAATGTCTTTC
Sphma-286          GGCGGTAACCCTTCTGAGC-TAAATTTTGAGGGAATAGGG-TTTGAGCCTTGCAAGTAGAAT------TGTGAAATTTTTAATGTATTTC
Sphre-277          GGCGATATCCTTTCTGAGC-TAA--TTTGCGGATTTAAGG-TTTGATTTGCAACATTAA------------------TGTGATGTTTTTG
Takakia            GGCGCTAACCTTCCTGAGC-TAA----------------------------GCCAGTAG------------------------------
Folioceros         GGCGTTATCCTTCCTGAGC-TGA----------------------------GAAAGAAG------------------------------
                   **** ** **.*.****** *.*                            .  *  *.

Sphan-285          AGTAAGGAAGGAAGCTGAATGTTAGGGTTAACATAATTA--------TTATGTTTTTAG-TATAAGCCCTTGT---TTCAGATATGAATT
S.fallax super_37  AGTAAGGAAGGAAGCTGAATGTTAGGGTTAACATAATTA--------TTATGTTTTTAG-TATAAGCCCTTGT---TTCAGATATGAATT
Sphfalx0293s001    AGTAAGGAAGGAAGCTGAATGTTAGGGTTAACATAATTA--------TTATGTTTTTAG-TATAAGCCCTTGT---TTCAGATATGAATT
Sphma-285          AGTAAGGAAGGAAGCTGAATGTTAGGGTTAACATAATTA--------TTATGTTTTTAG-TATAAGCCCTTGT---TTCAGATATGAATT
Sphre-283          AGTAAGGAAAGAAGCTGAATGTTAGGGTTAACATAATTA--------TTATGTTTTTAG-TATAAACCCTTGT---TTCAGATATGAATT
Sphpa              AGTAAGGAAGGAAACTGAATGTTAGGGTTAACATAATTA---------TTATG-TTTTAG-TATAAGCCCTTGT---TTCAGATATGAATT
Sphgi-292          AGAAAGGAAGGAAGCTGAATGTTAGGGTTAACATAATTAT-------TTATGTTTTTAG-TATAAGCCCTTGT---TTCAGATTTGAATT
Sphcri             AGAAAGGAAGGAAGCTGAATGTTAGGGTTAACATAATTA--------TTATGTTTTGAGAGATAAACCCTTGT---TTCAGATTTGAATT
Sphma-286          AGTAAGGAAGGAACATGAAAGTTACGGTTATCATAATTA--------TTATGTTTTTAG-TAAATGCCCTTGT---TTGAAATATGACTT
Sphre-277          AGTTAGTATGGGTTATGAATTTAGTGATTCTTTTGTTTT---------TAAATCAACAAT-TATTGATCATTGC---AAGAACAATAATGT
Takakia            ----------AGGGTGGGTTGAGGGG-----------------------GGGCACTAGG--ACACTTCCCGGC-----CTTGTGCCGGAT
Folioceros         -GCAAGGGTGGGGGTGGCGTGGCGGGCGGCGCCTTGTTAACGGGGTGTTAAGCACCAAC--GGACGCCCTGGCAGCCTCAGACGCCACCC
                             ..    *      . *                          .     . .   .    .* . *.           . .  .

Sphan-285          CTATAGCTTGAAGACATGACAAAC----ATGTTGTTCGTCATCTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAATGTTTGCACAT
S.fallax super_37  CTATAGCTTGAAGACATGACAAAC----ATGTTGTTCGTCATCTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAATGTTTGCACAT
Sphfalx0293s001    CTATAGCTTGAAGACATGACAAAC----ATGTTGTTCGTCATCTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAATGTTTGCACAT
Sphma-285          CTATAGCTTGAAGACATGATAAAC----ATGTTATTCCTCATTTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAATGTTTGCTCAT
Sphre-283          CTATAGCTTGAAGACATGACAAAC----ATGTTGTTCGTCATCTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAGTGTTTGCACAT
Sphpa              CTATAGCTTGAAGACATGATAAAC----ATGTTATTCCTCATCTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAATGTTTGCACAT
Sphgi-292          CTATAGCTTGAAGACATGACAAAC----ATGTTGTTCGTCATCTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAGTGTTTGCACAT
Sphcri             CTATAGCTTGAAGACATGACAAAC----ATGTTGTTCCTCATCTCATGATCA-CCTGCAGACCTACCCTTGAGACAAAATGTTTGCACAT
Sphma-286          ATATAGCTTGAAGACATAATAAAA----AAGAAATTCATCATTTCATGACCT-CCTGCACAACCTCCTTCGAGATAAAATGTTTGCACAT
Sphre-277          TGATCGAATTCAAAAGTCATTTACTATGTTTGTACCCATCACCTCATTCTCATCTTGCAGACCTACCCTTGCGACAAAGTGTGTGCAGAT
Takakia            ATGGTGGCCTAGGGTGTGATGAGT----GCTTTA-CCAGCACCTCACATTGGCCCAGCCGTCCTACCCTTGGTACAAGGGGACTGCAACT
Folioceros         ACGGCTCCGTAGGGTGTGATGAGT----GCTTTACCTAGCGCTCAGCCCCTGGCGAGCCCACCTACCCTTGTGACACGGGCCTGGCAGAT
                     .        ... .* *.  .          . ..  *.... ..  .   *  **    *. **.*.*  *.* ..     **   *

Sphan-285          TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
S.fallax super_37  TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
Sphfalx0293s001    TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
Sphma-285          TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
Sphre-283          TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
Sphpa              TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
Sphgi-292          TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
Sphcri             TATTGCAACAT-CTTGTCAATTTAGTTATCACTCCTGAGCTA
Sphma-286          TATTGAAACAT-CTCGTCAATTTAGTTATCACTCCTGAGCTA
Sphre-277          TATTGCAACAC-CTTGTCAATTTCGATATCACTCCTGAGCTA
Takakia            CTTTGCGCCATCCTTGTAAATTTGTTTATCACTCCTGAGCTA
Folioceros         CCCTGCACGGCCCCTGTCGG-TTACGTATCACTCCTGAGCTA
                   . .** .  .. *..** .. **   ****************
```

(B)

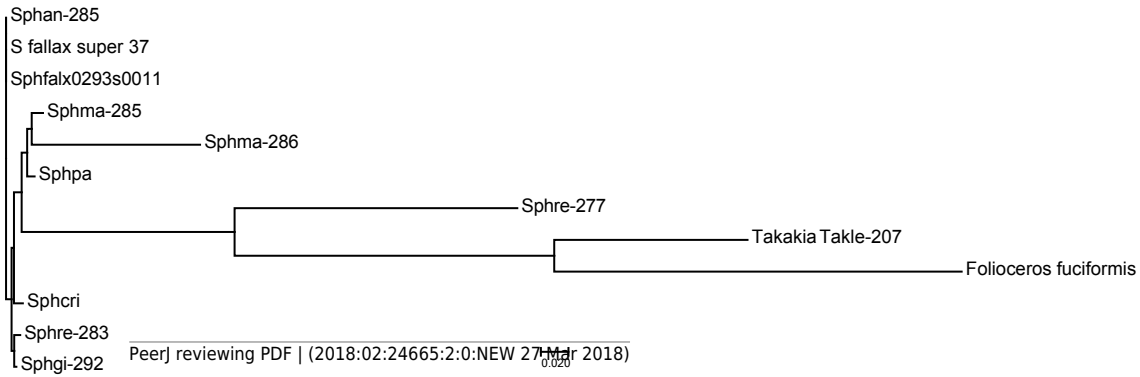


Figure 3(on next page)

Multiple sequence alignment of selected available nucleotide sequences of class III TAS3-like loci from Bryopsida mosses along with TAS3 locus of *Folioceros fuciformis*.

Alignment was generated at MAFFT6 program. The miR390 target sites are in yellow; putative tasiARF-a2 site is in green; tasiAP2 is in blue, and tasiARF-a3 is shaded. Note that all sequences are cut in the 5'-terminal TAS3 area to exclude non-aligned regions.

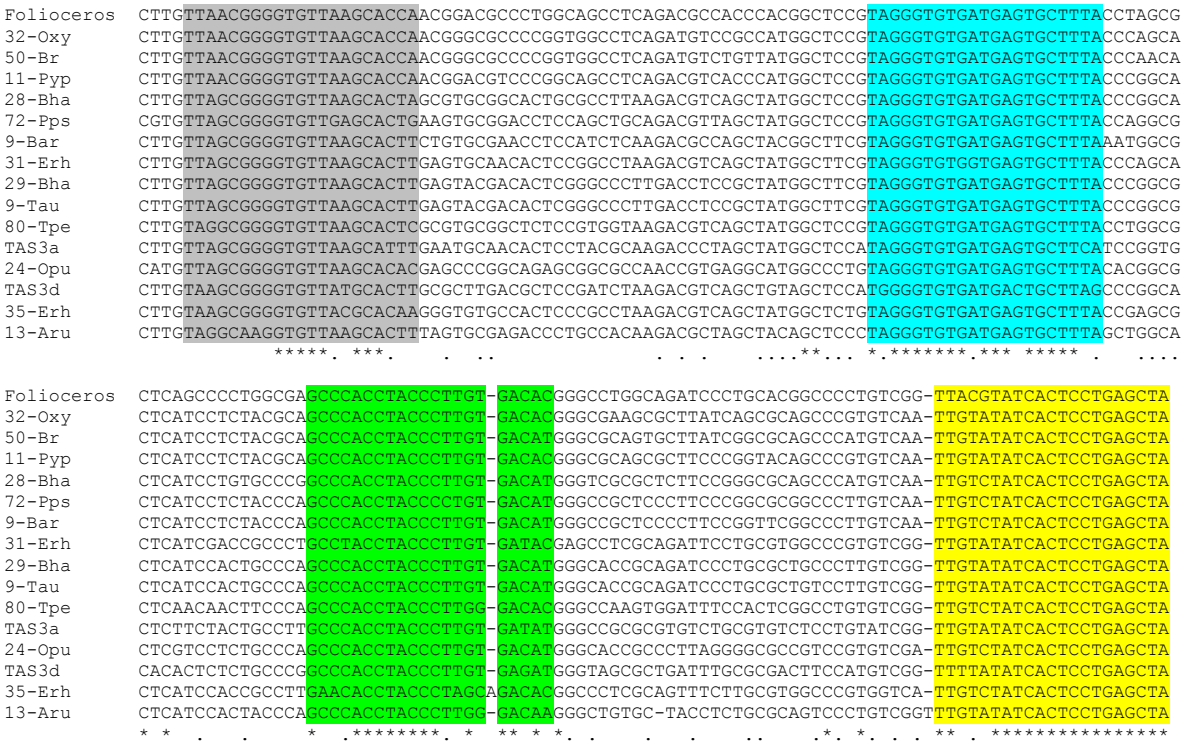

Folioceros CTTGTTAACGGGGTGTTAAGCACCAACGGACGCCCTGGCAGCCTCAGACGCCACCCACGGCTCCGTAGGGTGTGATGAGTGCTTTACCTAGCG
32-Oxy CTTGTTAACGGGGTGTTAAGCACCAACGGGCGCCCCGGTGGCCTCAGATGTCCGCCATGGCTCCGTAGGGTGTGATGAGTGCTTTACCCAGCA
50-Br CTTGTTAACGGGGTGTTAAGCACCAACGGGCGCCCCGGTGGCCTCAGATGTCTGTTATGGCTCCGTAGGGTGTGATGAGTGCTTTACCCAACA
11-Pyp CTTGTTAACGGGGTGTTAAGCACCAACGGACGTCCCGGCAGCCTCAGACGTCACCCATGGCTCCGTAGGGTGTGATGAGTGCTTTACCCGGCA
28-Bha CTTGTTAGCGGGGTGTTAAGCACTAGCGTGCGGCACTGCGCCTTAAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTACCCGGCA
72-Pps CGTGTTAGCGGGGTGTTGAGCACTGAAGTGCGGACCTCCAGCTGCAGACGTTAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTACCAGGCG
9-Bar CTTGTTAGCGGGGTGTTAAGCACTTCTGTGCGAACCTCCATCTCAAGACGCCAGCTACGGCTTCGTAGGGTGTGATGAGTGCTTTAAATGGCG
31-Erh CTTGTTAGCGGGGTGTTAAGCACTTGAGTGCAACACTCCGGCCTAAGACGTCAGCTATGGCTTCGTAGGGTGTGGTGAGTGCTTTACCCAGCA
29-Bha CTTGTTAGCGGGGTGTTAAGCACTTGAGTACGACACTCGGGCCCTTGACCTCCGCTATGGCTTCGTAGGGTGTGATGAGTGCTTTACCCGGCG
9-Tau CTTGTTAGCGGGGTGTTAAGCACTTGAGTACGACACTCGGGCCCTTGACCTCCGCTATGGCTTCGTAGGGTGTGATGAGTGCTTTACCCGGCG
80-Tpe CTTGTAGGCGGGGTGTTAAGCACTCGCGTGCGGCTCTCCGTGGTAAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTACCTGGCG
TAS3a CTTGTTAGCGGGGTGTTAAGCATTTGAATGCAACACTCCTACGCAAGACCCTAGCTATGGCTCCATAGGGTGTGATGAGTGCTTCATCCGGTG
24-Opu CATGTTAGCGGGGTGTTAAGCACACGAGCCCGGCAGAGCGGCGCCAACCGTGAGGCATGGCCCTGTAGGGTGTGATGAGTGCTTTACACGGCG
TAS3d CTTGTAAGCGGGGTGTTATGCACTTGCGCTTGACGCTCCGATCTAAGACGTCAGCTGTAGCTCCATGGGGTGTGATGACTGCTTAGCCCGGCA
35-Erh CTTGTAAGCGGGGTGTTACGCACAAGGGTGTGCCACTCCCGCCTAAGACGTCAGCTATGGCTCTGTAGGGTGTGATGAGTGCTTTACCGAGCG
13-Aru CTTGTAGGCAAGGTGTTAAGCACTTTAGTGCGAGACCCTGCCACAAGACGCTAGCTACAGCTCCCTAGGGTGTGATGAGTGCTTTAGCTGGCA
*****. ***. . .. .**... *.*******.*** ***** . ....
Folioceros CTCAGCCCCTGGCGAGCCCACCTACCCTTGT-GACACGGGCCTGGCAGATCCCTGCACGGCCCCTGTCGG-TTACGTATCACTCCTGAGCTA
32-Oxy CTCATCCTCTACGCAGCCCACCTACCCTTGT-GACACGGGCGAAGCGCTTATCAGCGCAGCCCGTGTCAA-TTGTATATCACTCCTGAGCTA
50-Br CTCATCCTCTACGCAGCCCACCTACCCTTGT-GACATGGGCGCAGTGCTTATCGGCGCAGCCCATGTCAA-TTGTATATCACTCCTGAGCTA
11-Pyp CTCATCCTCTACGCAGCCCACCTACCCTTGT-GACACGGGCGCAGCGCTTCCCGGTACAGCCCGTGTCAA-TTGTATATCACTCCTGAGCTA
28-Bha CTCATCCTGTGCCCGGCCCACCTACCCTTGT-GACATGGGTCGCGCTCTTCCGGGCGCAGCCCATGTCAA-TTGTCTATCACTCCTGAGCTA
72-Pps CTCATCCTCTACCCAGCCCACCTACCCCTGT-GACATGGGCCGCTCCCTTCCCGGCGCGGCCCTTGTCAA-TTGTCTATCACTCCTGAGCTA
9-Bar CTCATCCTCTACCCAGCCCACCTACCCTTGT-GACATGGGCCGCTCCCCTTCCGGTTCGGCCCTTGTCAA-TTGTCTATCACTCCTGAGCTA
31-Erh CTCATCGACCGCCCTGCCTACCTACCCTTGT-GATACGAGCCTCGCAGATTCCTGCGTGGCCCGTGTCGG-TTGTATATCACTCCTGAGCTA
29-Bha CTCATCCACTGCCCAGCCCACCTACCCTTGT-GACATGGGCACCGCAGATCCCTGCGCTGCCCTTGTCGG-TTGTATATCACTCCTGAGCTA
9-Tau CTCATCCACTGCCCAGCCCACCTACCCTTGT-GACATGGGCACCGCAGATCCCTGCGCTGTCCTTGTCGG-TTGTATATCACTCCTGAGCTA
80-Tpe CTCAACAACTTCCCAGCCCACCTACCCTTGG-GACACGGGCCAAGTGGATTTCCACTCGGCCTGTGTCGG-TTGTCTATCACTCCTGAGCTA
TAS3a CTCTTCTACTGCCTTGCCCACCTACCCTTGT-GATATGGGCCGCGCGTGTCTGCGTGTCTCCTGTATCGG-TTGTATATCACTCCTGAGCTA
24-Opu CTCGTCCTCTGCCCAGCCCACCTACCCTTGT-GACATGGGCACCGCCCTTAGGGGCGCCGTCCGTGTCGA-TTGTCTATCACTCCTGAGCTA
TAS3d CACACTCTCTGCCCGGCCCACCTACCCTTGT-GAGATGGGTAGCGCTGATTTGCGCGACTTCCATGTCGG-TTTTATATCACTCCTGAGCTA
35-Erh CTCATCCACCGCCTTGAACACCTACCCTAGCAGACACGGCCCTCGCAGTTTCTTGCGTGGCCCGTGGTCA-TTGTCTATCACTCCTGAGCTA
13-Aru CTCATCCACTACCCAGCCCACCTACCCTTGG-GACAAGGGCTGTGC-TACCTCTGCGCAGTCCCTGTCGGTTTGTATATCACTCCTGAGCTA
* * . . * .********. * ** * *. . . . .. .*. *. . . ** . ***************

Figure 4(on next page)

Novel potential ta-siRNA in Bryopsida plants.

(A) Multiple sequence alignments of nucleotide sequence blocks including tasiAP2 site and preceding 21 bp site of putative ta-siRNA of *Andreaea rupestris* TAS3-like locus WOGB_ 2010369 with the corresponding transcript sequences of moss TAS3 loci. BLASTN was used at 1KP blast site. For the complete TAS3 transcript sequences see Xia et al., 2017. The putative tasiAP2 site is in blue, and preceding putative ta-siRNA site is in violet. *Andreaea1 - Andreaea rupestris* WOGB_ 2010369; *Andreaea2 - Andreaea_rupestris* WOGB_ 2002765; *Tetraphis1 - Tetraphis_pellucida* HVBQ_2019753; *Tetraphis2 - Tetraphis_pellucida* HVBQ_2011866; *Tetraphis3 - Tetraphis_pellucida* HVBQ_2005644; *Plagiomniu*m - *Plagiomnium_insigne* BGXB_2010105; *Leucobryum - Leucobryum_glaucum* RGKI_2062694; *Racomitrium - Racomitrium_varium* RDOO_2117129; *Philonotis - Philonotis fontana* ORKS_2058791; *Dicranum - Dicranum_scoparium* NGTD_2078536; *Encalypta – Encalypta streptocarpa* KEFD_2058811; *Ceratodon - Ceratodon_purpureus* FFPD_2044193; *Niphotrichum - Niphotrichum_elongatum* ABCD_2000143; *Funaria - Funaria* sp. XWHK_2042016; *Schwetschkeop – Schwetschkeopsis fabronia* IGUH_2166854; *Aulacomnium - Aulacomnium heterostichum* WNGH_2088134; *Syntrichia - Syntrichia_princeps* GRKU_2074985; *Diphyscium1 - Diphyscium foliosum* AWOI_2069791; *Diphyscium2 - Diphyscium foliosum* AWOI_2006305; *Pohlia - Pohlia nutans* GACA01023180; *Bryum argenteum* - Unigene33538 GCZP01053768; PHYSCO TAS3A – *Physcomitrella patens* TAS3a BK005825.

(B) The example target transcript sequence (lower case letters) from *Leucobryum albidum* (VMXJ_2127900) is presented alongside with the predicted novel ta-siRNA shown in blue and above the transcript. Lower case letters in ta-siRNA indicate non-Watson-Crick pairing positions. Complementary mRNA sequences are in violet; conserved amino acid sequence signatures are in yellow. Numbers indicate codon positions of the target gene.

(A)

```
Andreaea1       AAGACGTCAGCTATGGCTCCCTAGGGTGTGATGAGTGCTTTA   43
Andreaea2       AAGACGCTAGCTACAGCTCCCTAGGGTGTGATGAGTGCTTTA  190
Tetraphis1      AAGACGTCAGCTATGGCTCCCTAGGGTGTGATGAGTGCTTTA  137
Tetraphis2      AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  182
Tetraphis3      AAGACGCCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  440
Plagiomnium     AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  315
Leucobryum      AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA 1006
Racomitrium     AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  304
Philonotis      AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  314
Dicranum        AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA   58
Encalypta       AAGACGTCAGCTATGGCTCCATAGGGTGTGATGAGTGCTTTA 1009
Ceratodon       AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  192
Niphotrichum    AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  455
Funaria         AAGACGTCAGCTATGGCTCTGTAGGGTGTGATGAGTGCTTTA  250
Schwetschkeop   AAGACGTCAGCTATGGCTCCGTAGGGTGTGATGAGTGCTTTA 1632
Aulacomnium     AAGACGTCAGTTATGGCTCCGTAGGGTGTGATGAGTGCTTTA  278
Syntrichia      AGGACGTCAGCTATGGTTCCGTAGGGTGTGATGAGTGCTTTA  154
Diphyscium1     AAGACGTCAGCTATGTCTTCATAGGGTGTGATGAGTGCTTTA  047
Diphyscium2     AAGACGTCACTTATGGCTTCCTAGGGTGTGATGAGTGCTTAA  486
Pohlia          AAGACGCCAGCTATGGCTTCGTAGGGTGTGATGAGTGCTTTA   38
Bryum           AAGACGCCAGCTACGGCTTCGTAGGGTGTGATGAGTGCTTTA  240
PHYSCO TAS3A    AAGACCCTAGCTATGGCTCCATAGGGTGTGATGAGTGCTTCA  158
```

(B)

```
                                                3'-UUCUGCaGUCGAUaCCGAGGc
caagacagttacggatcatctggacgtaccggaggtggtctagggcaagatgatagctacggctcat
 Q  D  S  Y  G  S  S  G  R  T  G  G  G  L  G  Q  D  D  S  Y  G  S  - 28

                                  3'-UUCUGCaGUCGAUACCGAGGc
tcatctggtcgcactggaggtggcctcggccaagacgacagctatggttcctctggccgc
 S  S  G  R  T  G  G  G  L  G  Q  D  D  S  Y  G  S  S  G  R  - 56

                        3'-UUCUGCaGUCGAUaCCGAGGc
ggtggcggcctcggtaaagatgatagctacggctcctcaggacgtaccgggggaaataca
 G  G  G  L  G  K  D  D  S  Y  G  S  S  G  R  T  G  G  N  T  - 96

    3'-UUCUGCaGuCGAUACCgAGGc
cttggacaagatgattcgtatggatcctctggacaaactggcggtaacagcggatacggc
 L  G  Q  D  D  S  Y  G  S  S  G  Q  T  G  G  N  S  G  Y  G  - 136
```

Figure 5(on next page)

Analysis of TAS3 loci in Marchantiophyta plants.

(A) Multiple sequence alignment of available nucleotide sequences of TAS3-like loci from Marchantiophyta plants along with TAS3 loci of *Takakia lepidozioides* and *Folioceros fuciformis*. Alignment was generated at MAFFT6 program. The miR390 target sites are in yellow, and putative tasiARF-a2 site is in green; tasiAP2 site is in blue.

(B) The minimal evolution phylogenetic tree based on analysis of the aligned TAS3 genes from Marchantiophyta plants. This tree was generated according MAFFT6 program. For full plant names and accession numbers see Table 2.

(A)

```
Takakia          GGCGCTAACCTTCCTGAGCTAAGCCAGTAGAG--------------------------------GGTGGGTTGAGGGGGGGCAC-----
Folioceros       GGCGTTATCCTTCCTGAGCTGAGAAAGAAGGCAAG------------------------GGTGGGGGTGGCGTGGCGGGCGGCGCCTTGT
1-Mpo            GACGGTATCCTTCTTGAGCTAAAAAG----------ATGTAGC-----------TTCCTGCTACATCTCACACGACACATCTCATTTGAA
Marchantia       GACGGTATCCTTCTTGAGCTAGGAAGAAGGAG----ATGTAGC-----------TTCCTGCTACATCTCACACGACATGTCTCGTTTGCA
Conocephal. Jap  GACGGTATCCTTCTTGAGCTAGGAGGAAAGAG----ATGTAGCGAGCGGGACCTTACCTGCTACATCTCACACGACGTGTCTCGTTTGAA
Conocephal. Con  GACGGTATCCTTCTTGAGCTAGGAGGAAAGAG----ATGTATCAAGCGGTGCCTTACCTGCTACATCTCACACGACATGTCTCGTTTGAA
Dumortiera       GACGGTATCTTTCTTGAGCTAGGAGGAACGAG----ATGTAGC---------TGTTCCTGCTACATCTCACACGACATGTCTCATTTGTA
Plagiochasma     GACGGTATCCTTCTTGAGCTAGGAGGAAAGAG----ATGTAGCGTT-----TAGTGCTAGCTACATCTCACACGACATGTCTCGTTTGTG
Ricciocarpos     GACGGTATCCTTCCTGAGCTAAGGGGACAGGA----AAAT-----------------CTGCTACATCTCACACGATACGTTTCTTTGGTA
Lunularia        GACGTTATCCTTCTTGAGCTAGGAAGAAGGAG----ATGTAGC-----------TTCCTGCTACATCTCACACGACATGTTTCACTAGAA
Metzgeria        GACGTTATCCTTCTTGAGCTAGACAGACGTGACAGGATGTATC---------TGGCTACTGCTGCATCTCACACTATGGGAGGGCTGGCGA
Pellia           GACGTTATCCGTCTTGAGCTAGATAGATGTGG-----------------------TTCTGCTGCATCTCACACTACATGCGGCCCT----
                 *.** ** *. *..******...  .                                  * .  . . . . .    .

Takakia          -----------------------------TAGGACACTTCCCGGC-------------CTTGTGCCGGATATGGTGGC----------C
Folioceros       TAACGGGGTGTTAAGCAC------------CAACGGACGCCCTGGC--------AGCCTCAGACGCCACCCACGGCTCC----------G
1-Mpo            TGTTCAAATCTTTAGTGACTGAATCGAATACTAAAGTTAATTTGACTTCAATAGAGACTAGTTTGCGGGAGAAACTGTGCCAGTTAGCAG
Marchantia       TGTTCAAATCTCTCGGATGTAAGTCACATACAGAAGTTAATTTGACGGCAAGCGAGACACATGTGCGGGACGGACACCCCTGGTTAGCAT
Conocephalum     TGTGCATGT----------------------CACAATTTAATTGGATGATCAACGAGATAAATGTGTTGGATGGACTCTCCTGGCTAGCAT
Conocephalum     TGTGCAAGT----------------------CACAATTTAATTCGATGATCAACGAGATAAATGTGTTGGATGGACTCTCCTGGCTAGCAT
Dumortiera       TGTTCAAAT----------------------CAAAAATTAATTTGACAACACACGAGACAGACGTGAGGGACAGAGACCTTTTGCTAGCAT
Plagiochasma     TGTTCAAAT----------------------CAGAAGTCAATTTGATGACAAAAGAGACAGACGTGTGAGACGGACGCCTTTGGCCAGCAT
Ricciocarpos     TCTTCAAAT----------------------CAGAATATGATTCGAAGATACACGAGATAAATTTGTGATACGGACACCTCTGCCTAAAAT
Lunularia        AATCTAG------------------------------------------TAGTGAAACAATTGTGTGGGACAGGCCTGCCTGGTAAGCAC
Metzgeria        GCTGGAGCTGTT-------------------------------------------CTCAGCCTTTATGGTGAGATACCCTTGGTCAGCTT
Pellia           ------------------------------------------------------CTCTTGCTGGCAATGACAACTAGGTGAGCAT
                                                                          .   .   . .

Takakia          TAGGGTGTGATGAGTGCTTTACCAG-CACCTCACATTGGCCCAGCCGTCCTACCCTTGGTACAAGGGGACTGCAACTCTTTGCGCCATCC
Folioceros       TAGGGTGTGATGAGTGCTTTACCTAGCGCTCAGCCCCTGGCGAGCCCACCTACCCTTGTGACACGGGCCTGGCAGATCCCTGCACGGCCC
1-Mpo            GAGGGTGTGATGAGTGCTTTACCT--------------GGTCCAGGATCCCCACCCCCTCCT----CCACTGCCTATTTCTAGGCTCGCG
Marchantia       GAGGGTGTGATCAGTGCTTTACCC-------------GGCTCGGGTTCCCGTCCCCTTCCC----CCACTGCCTATGTCTAGGCTCGCC
Conocephalum     GAGGGTGTGATGAGTGCTTTACCA-------------TACAGGGGTTCCCGTCCTTTCTTC----CCGTCGCATATGTCTAGGTTCGCC
Conocephalum     TAGGGTGTGATGAGTGCTTTACCA-------------TACAGGGGTTCCCGTCCTTTCTTC----CCGTCGCATATGTCTAGGTTCGCC
Dumortiera       GAGGGTGTGATGAGTGCTTTACCA-------------GGCAAGGGTTCACGTCCTTTTCTC----CCATTGCCTATGTCTAGGCTCGCC
Plagiochasma     GAGGGTGTGATGAGTGCTTTACCA-------------GGCCGAGGTTCCCGTCCTTTTCTC----CCATTGCCTATGTCTAGGCTCGCC
Ricciocarpos     GAGGGTGTGATGAGTGCTTTACTA-------------GGCAGGGGTTCACGTCCATTTCTC----CCACTGCATATGTCTAGGTTCGCC
Lunularia        TAGAGTGTGATAGGTGCTTTACCA-------------GACGGGGGCTCCCGTCCCCTTCCC----CCACTGCCTATGTCTAGGCTCGCC
Metzgeria        TAGGGTGTGATATGTGCTTGACGT-------------GGCAGTCTCACCCGCGCCCCTGGG----CAAATGCCTGCATCTTCCTATCTG
Pellia           GGAGGTGTGATATGTGCTTGACGT-------------GACTTTCCCTCCTGCCCCTGATTT----GTTTTGCCTGTTGCTAGGCTAGTC
                  ...*******  ****** **                .       * .  * .            **   .  .*   .   .

Takakia          TTGTAAATTTGTTTATCACTCCTGAGCTA
Folioceros       CTGTCGG-TTACGTATCACTCCTGAGCTA
1-Mpo            -TTAC---CTGCCTATCCCTCTTGAGCTA
Marchantia       -TGAC---CTGCCTATCCCTCTTGAGCTA
Conocephalum     -TGAC---CTGCCTATCCCTCTTGAGCTA
Conocephalum     -TGAC---CTGCCTATCCCTCTTGAGCTA
Dumortiera       -TGAC---CTGCCTATCCCTCTTGAGCTA
Plagiochasma     -TGAC---CTGCCTATCCCTCTTGAGCTA
Ricciocarpos     -TGAC---CTGCCTATCCCTCTTGAGCTA
Lunularia        -TGAC---CTGCCTATCCCTCTTGAGCTA
Metzgeria        -TTAA---TTTCCTATCCCTTCTGAGCTA
Pellia           TTGTA---TCTCCTATCCCTCCTGAGCTA
                  *      .. . **** **..*******
```

(B)

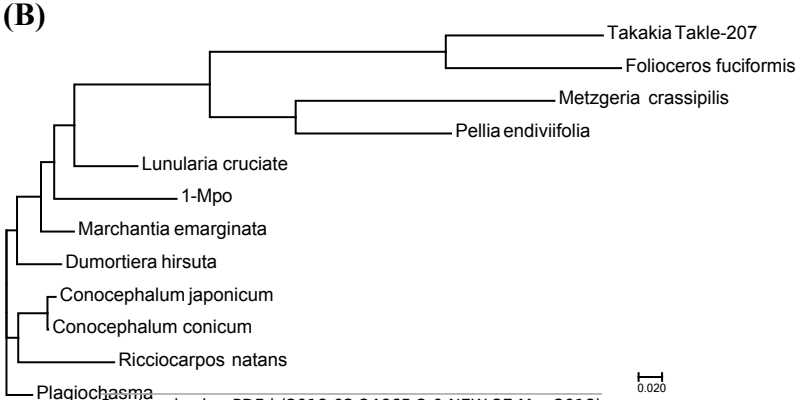
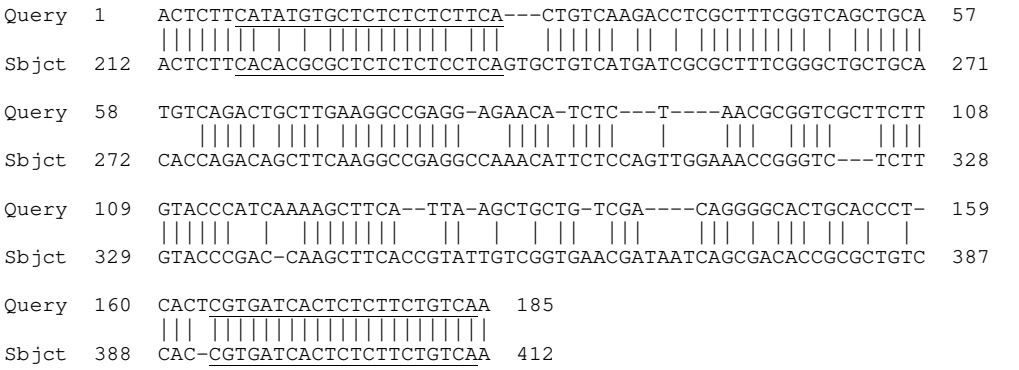


Figure 6(on next page)

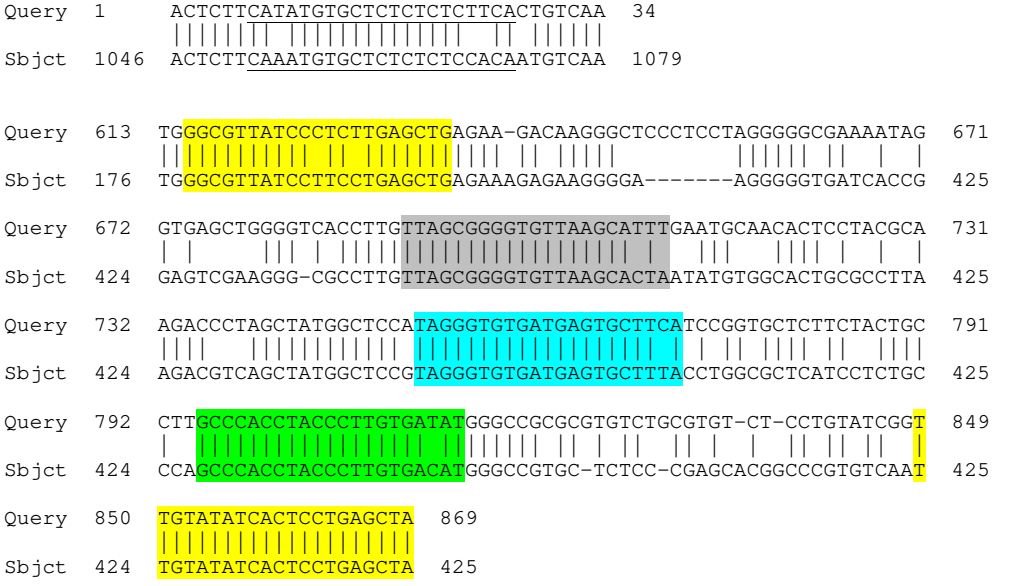
Pairwise sequence comparisons of selected nucleotide sequences of TAS6/TAS3-like loci from mosses with TAS6/TAS3 of *Physcomitrella patens* precursor RNA (accession JN674513).

BLASTN was used at 1KP blast site. The miR390 target sites are in yellow; putative miR156/miR529 sites are underlined; tasiAP2 is in blue; putative tasiARF-a2 site is in green; tasiARF-a3 is shaded.

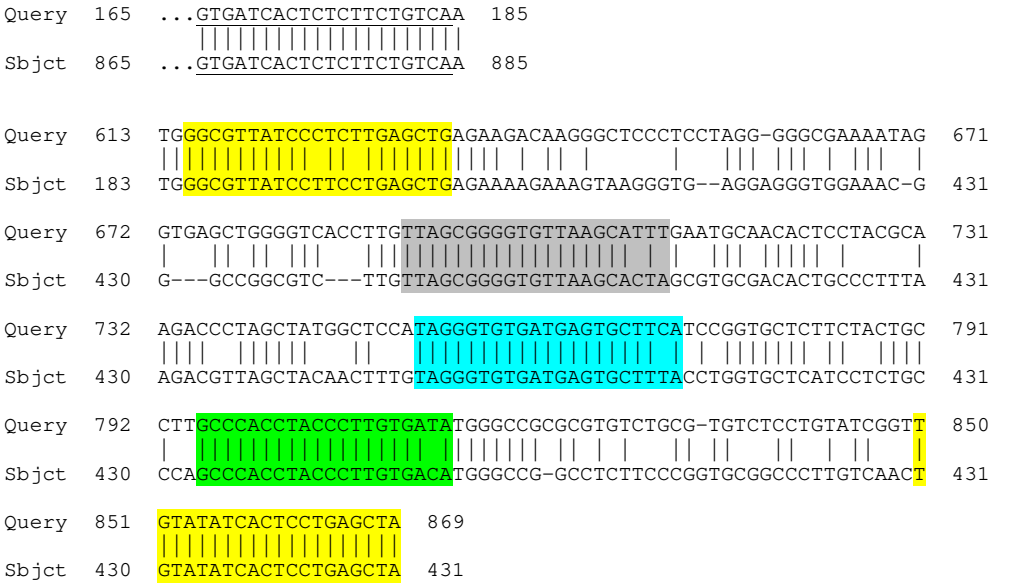
2058811 *Encalypta streptocarpa* (accession KEFD_2058811)



2058791 *Philonotis fontana* (accession ORKS_2058791)



2050742 *Hedwigia ciliata* (accession YWNF _2050742)



2069791 *Diphyscium foliosum* (accession AWOI_2069791)

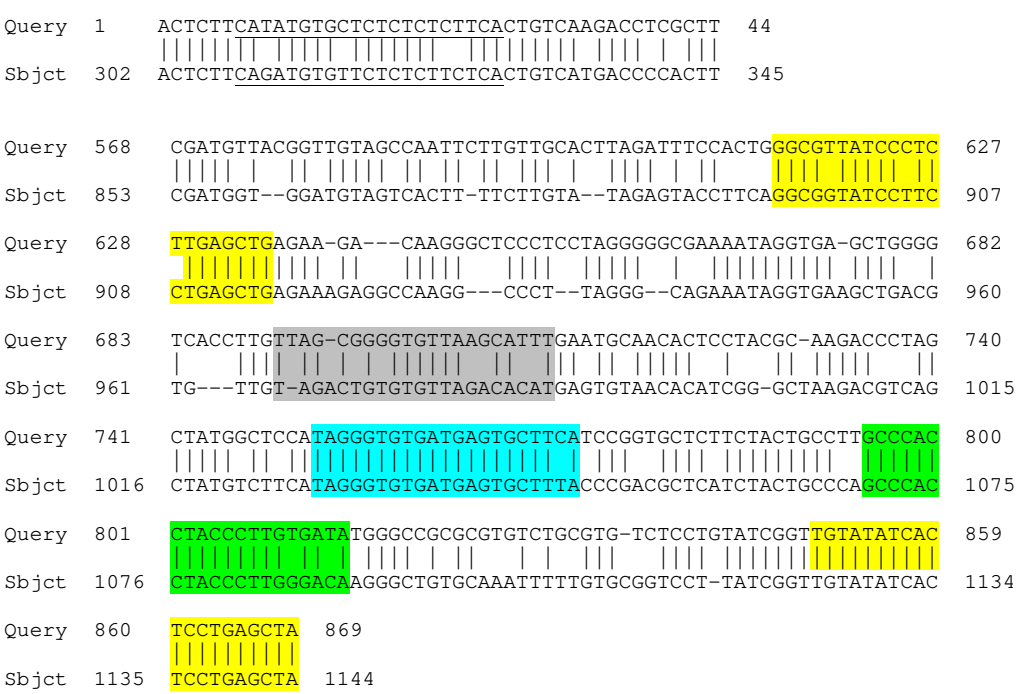


Figure 7 (on next page)

The phylogenetic tree constructed from conserved regions of SGS3 protein sequences from 27 selected streptophytes by the Neighbor-Joining method with 1000 bootstrap replications.

There were a total of 419 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar et al., 2016). Bootstrap support values ≥50% are shown. The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). The tree was rooted at two Klebsormidiaceae charopytes. Accession numbers from NCBI ot PHYTOZOME data banks see in Fig. S5.

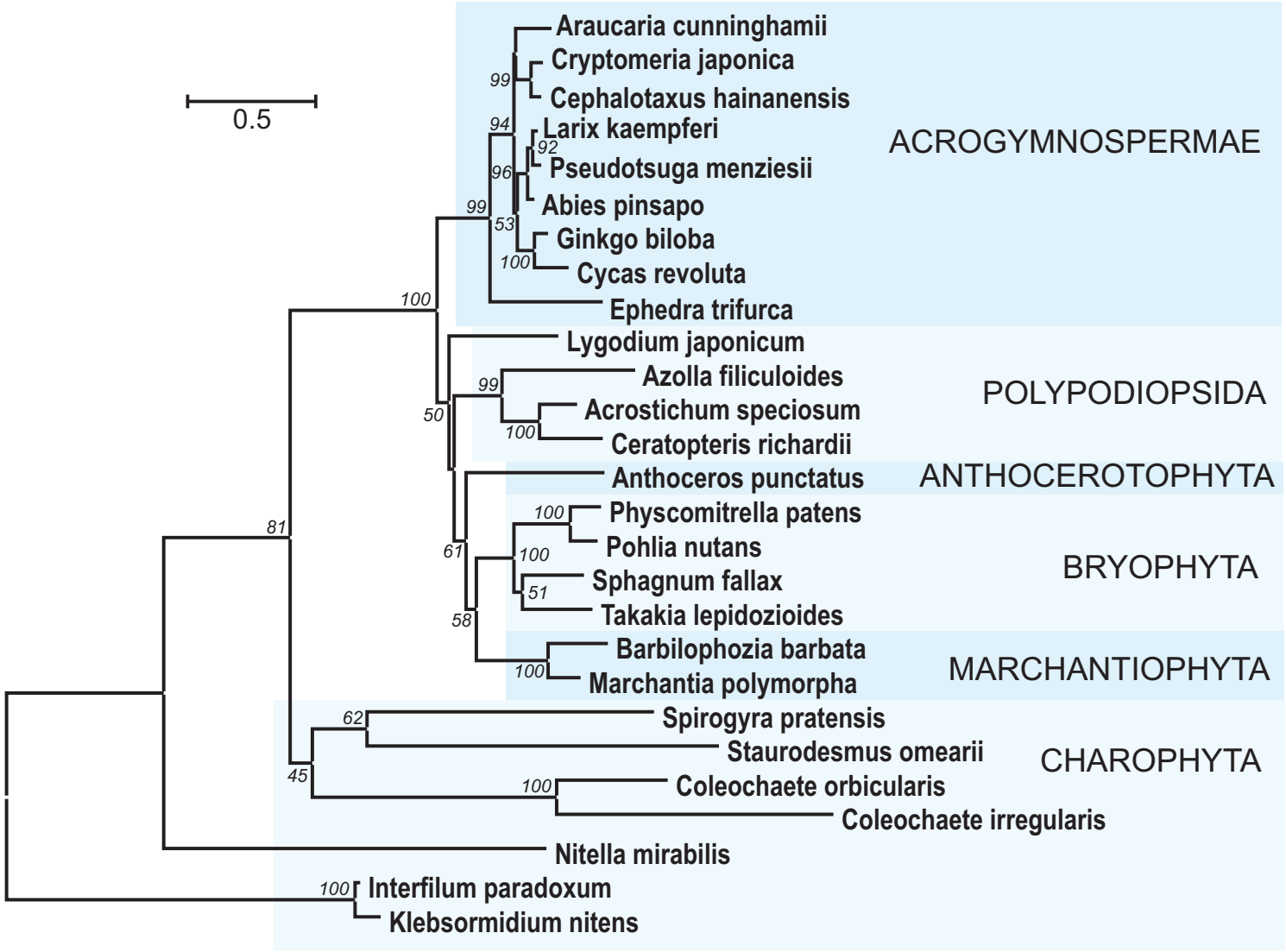

0.5
Araucaria cunninghamii
Cryptomeria japonica
Cephalotaxus hainanensis
Larix kaempferi
Pseudotsuga menziesii
Abies pinsapo
Ginkgo biloba
Cycas revoluta
Ephedra trifurca
ACROGYMNOSPERMAE
Lygodium japonicum
Azolla filiculoides
Acrostichum speciosum
Ceratopteris richardii
POLYPODIOPSIDA
Anthoceros punctatus
ANTHOCEROTOPHYTA
Physcomitrella patens
Pohlia nutans
Sphagnum fallax
Takakia lepidozioides
BRYOPHYTA
Barbilophozia barbata
Marchantia polymorpha
MARCHANTIOPHYTA
Spirogyra pratensis
Staurodesmus omearii
Coleochaete orbicularis
Coleochaete irregularis
CHAROPHYTA
Nitella mirabilis
Interfilum paradoxum
Klebsormidium nitens
99
94
92
96
99
53
100
100
50
99
100
81
61
100
100
51
58
100
62
45
100
100

Table 1(on next page)

List of the putative TAS3 loci in Sphagnopsida and Takakiopsida

Table 1.
List of the putative TAS3 loci in Sphagnopsida and Takakiopsida

Plant species	Locus name	Subgenus	Length	Sequence source
Sphagnum angustifolium	Sphan-285	*Cuspidata*	285 nts	MF682529
S. girgensohnii	Sphgi-292	*Acutifolia*	292 nts	MF682530
S. fallax	contig super_37	*Cuspidata*	285 nts	SRX2120232
S. fallax	Sphfalx0293s0011	*Cuspidata*	277 nts	Sphfalx0293s0011*
S. recurvum	Sphre-283	*Cuspidata*	283 nts	SRX1513231
S. recurvum	Sphre-277	*Cuspidata*	277 nts	SRX1513231
S. magellanicum	Sphma-285	*Sphagnum*	285 nts	SRX2330962
S. magellanicum	Sphma-286	*Sphagnum*	286 nts	SRX2330962
S. palustre	Sphpa	*Sphagnum*	partial	SRX1516347
S. cribrosum	Sphcri	*Subsecunda*	291 nts	ERX443237
S. lescurii	Sphle	*Subsecunda*	partial	ERX337183
Takakia lepidozioides	Takle-207	Not applicable	207 nts	ERX2100030 SKQD-2076588**

* - PHYTOZOME accession; ** - 1KP accession (Xia et al., 2017). Different sphagnum subgenera are colored specifically.

Table 2(on next page)

List of the putative TAS3 loci in Anthocerotophyta and Marchantiophyta

Table 2.
List of the putative TAS3 loci in Anthocerotophyta and Marchantiophyta

Plant species	Class/subclass	Order	Length	Sequence source
Folioceros fuciformis	Anthocerotopsida/Anthocerotidae	Anthocerotales	244 nts	SRS2162762
Marchantia polymorpha 1-Mpo	Marchantiopsida/Marchantiidae	Marchantiales	256 nts	KC812742
Marchantia emarginata	Marchantiopsida/Marchantiidae	Marchantiales	262 nts	SRX1952816
Conocephalum japonicum	Marchantiopsida/Marchantiidae	Marchantiales	252 nts	SRX1952810
Ricciocarpos natans	Marchantiopsida/Marchantiidae	Marchantiales	235 nts	ERX337127
Dumortiera hirsuta	Marchantiopsida/Marchantiidae	Marchantiales	243 nts	SRX1126014
Plagiochasma appendiculatum	Marchantiopsida/Marchantiidae	Marchantiales	247 nts	SRX1741567
Conocephalum conicum	Marchantiopsida/Marchantiidae	Marchantiales	248 nts	ILBQ_2006554*
Lunularia cruciata	Marchantiopsida/Marchantiidae	Lunulariales	220 nts	TXVB_2071521*
Marchantia paleaceae	Marchantiopsida/Marchantiidae	Marchantiales	257 nts	HMHL_2051051*
Metzgeria crassipilis	Jungermanniopsida/Metzgeriidae	Metzgeriales	226 nts	ERX337128
Pellia endiviifolia	Jungermanniopsida/Pelliidae	Pelliales	192 nts	SRX726500

* - 1KP accession (Xia et al., 2017).

Table 3(on next page)

List of the putative TAS6/TAS3 loci of Bryophyta in transcribed sequences found in 1KP database

Table 3.
List of the putative TAS6/TAS3 loci of Bryophyta in transcribed sequences found in 1KP database

Plant species	Class/subclass	Order	Length* and type	Sequence source
Timmia austriaca	Bryopsida/Timmiidae	Timmiales	TAS6/TAS3 (874nts)	ZQRI-2061439 ZQRI-2063082
Thuidium delicatulum	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (837nts)	EEMJ-2003175
Hypnum subimponens	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (823nts)	LNSF-2068452
Pseudotaxiphyllum elegans	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (1590nts)	QKQO-2009669
Anomodon attenuatus	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (843nts)	QMWB-2059873
Anomodon rostratus	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (829nts)	VBMM-2003482
Schwetschkeopsis fabronia	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (854nts)	IGUH-2166854
Leucodon sciuroides	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (852nts)	ZACW-2016434
Fontinalis antipyretica	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (1410nts)	DHWX-2007057
Rhytidiadelphus loreus	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (830nts)	WSPM-2009782
Rhynchostegium serrulatum	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (853nts)	JADL-2047695
Climacium dendroides	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (809nts)	MIRS-2012325
Calliergon cordifolium	Bryopsida/Bryidae	Hypnales	TAS6 (95nts)	TAVP-2006322
Neckera douglasii	Bryopsida/Bryidae	Hypnales	TAS6/TAS3 (839nts)	TMAJ-2023603
Plagiomnium insigne	Bryopsida/Bryidae	Bryales	TAS6/TAS3 (914nts)	BGXB-2010105
Orthotrichum lyellii	Bryopsida/Bryidae	Orthotrichales	TAS6 (192nts)	CMEQ-2080784
Hedwigia ciliata	Bryopsida/Bryidae	Hedwigiales	TAS6/TAS3 (877nts)	YWNF-2050742
Philonotis fontana	Bryopsida/Bryidae	Bartramiales	TAS6/TAS3 (893nts)	ORKS-2058791
Aulacomnium heterostichum	Bryopsida/Bryidae	Rhizogoniales	TAS6/TAS3 (863nts)	WNGH-2088134
Scouleria aquatic	Bryopsida/Dicranidae	Scouleriales	TAS6/TAS3 (partial)	BPSG-2088977
Syntrichia princeps	Bryopsida/Dicranidae	Pottiales	TAS6/TAS3 (partial)	GRKU-2074985
Leucobryum glaucum	Bryopsida/Dicranidae	Dicranales	TAS6/TAS3 (763nts)	RGKI-2062694
Leucobryum albidum	Bryopsida/Dicranidae	Dicranales	TAS6/TAS3 (763nts)	VMXJ-2128109
Dicranum scoparium	Bryopsida/Dicranidae	Dicranales	TAS6 (105nts)	NGTD-2092412

Ceratodon purpureus	Bryopsida/Dicranidae	Pseudoditrichales	TAS6/TAS3 (1121nts)	FFPD-2005850 SRX2065999
Racomitrium varium	Bryopsida/Dicranidae	Grimmiales	TAS6/TAS3 (724nts)	RDOO-2117129
Physcomitrium_sp.	Bryopsida/Funariidae	Funariales	TAS6 (partial)	YEPO-2071108
Physcomitrium_sp.	Bryopsida/Funariidae	Funariales	TAS6 (178nts)	YEPO-2000016
Physcomitrium_sp.	Bryopsida/Funariidae	Funariales	TAS6/TAS3 (821nts)	YEPO-2016361
Encalypta streptocarpa	Bryopsida/Funariidae	Encalyptales	TAS6/TAS3 (883nts)	KEFD-2058811
Diphyscium foliosum	Bryopsida/Diphysciidae	Diphyscales	TAS6/TAS3 (832nts)	AWOI-2069791
Tetraphis pellucida	Tetraphidopsida	Tetraphidales	TAS6 (partial)	HVBQ-2112923
Atrichum angustatum	Polytrichopsida	Polytrichales	TAS6/TAS3 (810nts)	ZTHV-2082998
Andreaea rupestris	Andreaeopsida	Andreaeales	TAS6/TAS3 (869nts)	WOGB-2010369
Takakia lepidozioides	Takakiopsida	Takakiales	TAS6/TAS3 (1040nts)	SKQD-2076588

* - The length indicates total size of TAS6-TAS3 complex element (from the 5' miR529 target site in TAS6 to 3' miR390 target site in TAS3) or isolated TAS6 (between miR529 and miR156 target sites).