Methods

Azolla domestication towards a biobased economy?

Paul Brouwer[1], Andrea Bräutigam[2], Canan Külahoglu[2], Anne O. E. Tazelaar[1], Samantha Kurz[2], Klaas G. J. Nierop[3], Adrie van der Werf[4], Andreas P. M. Weber[2] and Henriette Schluepmann[1]

[1]Molecular Plant Physiology Department, Utrecht University, Padualaan 8, 3584 CH, Utrecht, the Netherlands; [2]Institute for Plant Biochemistry, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; [3]Organic Geochemistry, Utrecht University, Budapestlaan 4, 3584 CD, Utrecht, the Netherlands; [4]Plant Research International, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands

Author for correspondence:
Henriette Schluepmann
Tel: +31 30 253 3289
Email: h.schlupmann@uu.nl

Received: *31 October 2013*
Accepted: *22 December 2013*

New Phytologist (2014) **202:** 1069–1082
doi: 10.1111/nph.12708

Key words: *Azolla filiculoides*, biobased economy, cryopreservation, domestication, fern, RNAseq, sexual reproduction.

Summary

- Due to its phenomenal growth requiring neither nitrogen fertilizer nor arable land and its biomass composition, the mosquito fern *Azolla* is a candidate crop to yield food, fuels and chemicals sustainably. To advance *Azolla* domestication, we research its dissemination, storage and transcriptome.
- Methods for dissemination, cross-fertilization and cryopreservation of the symbiosis *Azolla filiculoides–Nostoc azollae* are tested based on the fern spores. To study molecular processes in *Azolla* including spore induction, a database of 37 649 unigenes from RNAseq of microsporocarps, megasporocarps and sporophytes was assembled, then validated.
- Spores obtained year-round germinated *in vitro* within 26 d. *In vitro* fertilization rates reached 25%. Cryopreservation permitted storage for at least 7 months. The unigene database entirely covered central metabolism and to a large degree covered cellular processes and regulatory networks. Analysis of genes engaged in transition to sexual reproduction revealed a FLOWERING LOCUS T-like protein in ferns with special features induced in sporulating *Azolla* fronds.
- Although domestication of a fern–cyanobacteria symbiosis may seem a daunting task, we conclude that the time is ripe and that results generated will serve to more widely access biochemicals in fern biomass for a biobased economy.

Introduction

In the coming decades we anticipate a rapid increase in world population that will greatly increase global demand for food, although this is constrained by the limited availability of arable land. With the depletion of fossil resources, plants will need to provide an increasingly large proportion of our requirements for energy and chemicals in addition to food (EPSO, 2005). Intensive agriculture using conventional crops is often associated with high inputs and negative climate impacts (Jensen *et al.*, 2012). For example, periodic application of excess nitrogen fertilizer leads to high nitrous oxide (N_2O) emissions (Smith *et al.*, 1997). N_2O has a global warming potential 310-fold higher than CO_2 and is, in terms of impact, the third most important greenhouse gas, after CO_2 and CH_4 (IPCC, 2007). To meet the challenges of lowering N_2O emissions and increasing production, novel crops that require less or no nitrogen fertilizer, use nonarable land with high biomass yields, and feed both food and chemical industries, are much sought after.

Azolla is one such potential crop: it is an aquatic fern that may be cultivated in closed systems on nonarable land as well as in freshwater basins. It is known for its high growth rates, doubling biomass in 2 d under favorable conditions (Wagner, 1997). *Azolla* thrives without the addition of nitrogen fertilizer to sustain its growth because it harbors symbiotic nitrogen-fixing cyanobacteria; this has led to its use as a nitrogen fertilizer in paddy fields of South-east Asia (Wagner, 1997). It is an accumulator of heavy metals and its use in waste water treatment has been demonstrated on a small scale (Costa *et al.*, 1999; Antunes *et al.*, 2001). *Azolla* is a high-protein animal feed, but its limited digestibility (Alalade & Iyayi, 2006; Abdel-Tawwab, 2008) may be the result of it containing tannins not lignin (Nierop *et al.*, 2011), together with other polyphenols (Fasakin, 1999). Compared to algae and free-living diazotrophic cyanobacteria, *Azolla* requires no mixing of the water body and is easier to harvest. To fully harness the potential of *Azolla*, however, its domestication is a prerequisite. Domestication requires protocols for the collection, storage and dissemination of reproductive structures (Willemse, 2009). In addition, breeding varieties adapted to specific uses and permitting containment of this otherwise invasive weed requires control over sexual reproduction. To accelerate the breeding process, genetic and sequence information is required. As with most ferns, *Azolla* is currently neither domesticated nor bred (Meyer *et al.*, 2012). Sequence information is lacking except for the genome

sequence of its endosymbiotic cyanobacteria *Nostoc azollae* (Ran *et al.*, 2010).

Azolla is a member of the *Salviniaceae* and heterosporous: megaspores and microspores each form gametophytes bearing the gametes that mate to form the sporophyte, the dominant diploid phase of *Azolla*. A very small gametophyte yielding a single megaspore develops inside the megasporocarp (Peters & Perkins, 1993). By contrast many microsporangia with microspores packaged in several pseudocellular massulae develop in the microsporocarp (Herd *et al.*, 1985). *Azolla* spores inside the sporocarps generally exhibit strong resistance to external stresses, such as drought (Becking, 1987) and subzero temperatures (Janes, 1998). Using sporocarps to preserve biodiversity may seem a logical approach; however, sporocarps have not been generally available for *Azolla* species and the preservation methods will need to be improved to reach long-term and high viability. Long-term storage by preservation of whole sporophytes or only small parts of the meristems was not reported for *Azolla*. Instead, *Azolla* varieties are currently maintained by *in vitro* subculture in biodiversity collections such as the Biofertilizer Collection of the International Rice Research Institute (IRRI; Watanabe *et al.*, 1992). Continuous subculture, however, is laborious and thus prone to human error, and may promote somaclonal variation and adaptation of the specimens to the artificial environment under which they have been cultured during the past 25 yr. A more reliable method to preserve varieties needs to be developed that allows selection and breeding efforts.

Control over *Azolla* sexual reproduction will be of paramount importance for disseminating existing varieties or breed new varieties. Controlling the production of spores and fertilization will be most critical. A number of authors have described sexual reproduction in *Azolla* species (Becking & Donze, 1981; Becking, 1987; Peters & Meeks, 1989; Wagner, 1997; Zheng *et al.*, 2009; Carrapiço, 2010). These studies mainly focused on the vertical transfer of the cyanobacterial symbiont during the reproduction process. Methods were described that used sporocarps to raise new sporelings field plots, relying on the natural processes of fertilization on the floor bed (Quin-Yuan *et al.*, 1987; Shuying, 1987). None of the publications on sexual reproduction in *Azolla* described fertilization in a controlled laboratory environment, and none described the induction of spore formation *in vitro*. Knowledge of the environmental cues and molecular mechanisms controlling sporulation in *Azolla* is very scarce. Studies have reported outdoor conditions under which sporulation has occurred in different species, such as high population density (Becking, 1987; Janes, 1998), shorter days and colder nights (Kar *et al.*, 2002), high light intensity and high temperature (Becking, 1987). Herd *et al.* (1989) showed the effect of temperature regime on the sporulation of a large variety of *Azolla* species and strains, but could not establish a clear link with sporulation frequency. Also different growth-regulating substances could not induce sporulation in *A. pinnata* and *A. filiculoides* (Herd *et al.*, 1989). Kar *et al.* (2002) later showed that a combination of hormones could increase sporulation frequency and promote megasporocarp formation, but only in cultures that were already sporulating. As *Azolla* species adapted to differing environments over time, they likely evolved differential environmental cues triggering sporulation. To develop a reliable protocol to induce sporulation in several species, a strategy focusing on more downstream, molecular, processes controlling the transition to sexual reproduction will likely be more effective than studying environmental clues alone.

In flowering plants, the transition to sexual reproduction is controlled by multiple input pathways which measure day length, temperature, nutritional status and age of the plant. FLOWERING LOCUS T (FT) signals environmental cues perceived in leaves as it moves via the vasculature to the shoot meristems; there, FT activates LEAFY (LFY) and with it the transition to sexual reproduction. FT was not found in the genomes of the nonvascular lower plants *Physcomitrella* and *Selaginella* (Banks *et al.*, 2011) and has not previously been described in ferns, which were the first plants to evolve highly developed vasculature.

In lower plants LFY is thought to promote vegetative development of the gametophyte, and Floyd & Bowman (2007) proposed that LFY repression after fertilization would be required for development of the extended vegetative growth of the sporophytes in higher plants. LFY from the fern *Ceratopteris* was capable of partially suppressing the phenotype of the Arabidopsis *lfy* mutant but LFY from the moss *Physcomitrella* did not (Maizel *et al.*, 2005). Some targets of LFY are therefore conserved in ferns and higher plants. LFY activates the ABC genes in angiosperms and gymnosperms, thus promoting the transition to sexual reproduction. What induces the transition to sexual reproduction in ferns and other lower plants is mostly unknown.

Azolla belongs to an under-sampled group with regard to transcriptome or genome sequence resources. Studying molecular components that may control the transition to reproductive development in *Azolla* is therefore difficult. RNA sequencing of species without a sequenced genome provides a valuable resource. While the assemblies remain far from perfect (Schliesky *et al.*, 2012), both unigene databases (Brown *et al.*, 2011; Kajala *et al.*, 2012; Sommer *et al.*, 2012) and quantitative gene expression data (Brautigam *et al.*, 2011a; Gowik *et al.*, 2011) have successfully been used to explore the physiology and gene regulation in species without prior sequence resources.

In order to provide the basis needed for domestication of *Azolla*, we begin by describing a method to collect large amounts of *A. filiculoides* spores all year round. We define and illustrate key stages in the germination process then demonstrate *in vitro* fertilization and germination of *Azolla* spores. We further show that cryopreservation of fertilized megaspores using a drying pretreatment is effective for preserving *A. filiculoides* while also preserving the *N. azollae* symbiont, opening the way to genomic characterization of the cryopreserved variety. From sequencing reads of RNA from megasporocarps, microsporocarps and sporophytes we assemble a database of 37 649 unigenes which we annotate so as to provide a resource to molecular research. We then describe genes in *Azolla* possibly involved in inducing spore formation based on what is known from induction of the reproductive phase in flowering plants.

Materials and Methods

Collecting *Azolla* sporocarps

Sporulating *A. filiculoides* Lam. was collected in mid-October 2012 from a ditch in Utrecht, the Netherlands (52°04′24″N; 5°08′53″E) and kept in demineralized water in a glasshouse at 5–15°C and 14 h days with a light intensity of at least 70 μmol m^{-2} s^{-1} Photosynthetic Photon Flux Density (PPFD).

Unfertilized megasporocarps were collected according to Toia *et al.* (1987) with modifications: sporulating plants were placed on top of a stack of sieves with 1000-, 500- and 200-μm mesh sizes and megasporocarps were detached from the stems using a strong water jet. Residue recovered on the 200-μm mesh size sieve was then layered on a 3 M sorbitol: water step gradient and centrifuged at 400 ***g*** for 10 min, resulting in a clear layer of pure megasporocarps. Megasporocarps were washed three times by centrifugation with 50 ml water. To collect mostly fertilized megasporocarps, sediment accumulating at the bottom of the containers in which sporulating cultures were kept was used in the above procedure, except that the sorbitol: water step gradient was layered on top of the residue of the 200-μm mesh size sieve to obtain a clear layer of fertilized megasporocarps. Mature microsporocarps were plucked manually from the plants.

Documenting megaspore germination

A mixture of megasporocarps and microsporangia was germinated in 400 μl of either *Azolla* growth medium at pH 5.5 (Watanabe *et al.*, 1992) or demineralized water in a growth cabinet set at 25°C day : 15°C nights with 12 h light (40–70 μmol s^{-1} m^{-2} PPFD). Spores were scored for germination over a period of 6 wk. Morphological changes to the megasporocarps were tracked under a binocular Leica Axioskop light microscope, using either × 10 or × 5 objectives, and key events documented by digital imaging in dark field using a Leica PFC 420C Camera (Carl Zeiss BV, Sliedrecht, Netherlands). Images of sporulating *Azolla* plants and *Azolla* sporelings were made using a Nikon DXM12000 camera (Nikon Instruments Europe BV, Amsterdam, Netherlands) on either a Zeiss Axiovert 35M reversed microscope or a Zeiss Stemi SV11 stereo microscope (Carl Zeiss). Mature *Azolla* plants were photographed using a Nikon D300S DSLR camera with a 60 mm macro objective.

Documenting *Nostoc azollae*

Nostoc azollae phycobilisomes differ in their spectral emissions from plant pigments (Rigbi *et al.*, 1980), therefore, a Leica SP2 confocal fluorescence microscope, equipped with either × 16 or × 40 objectives and a helium–neon laser with excitation wavelength of 543 nm, was used to visualize *N. azollae* fluorescence in the range 630–670 nm. Plant tissue fluorescence was visualized in the range 560–630 nm and/or 680–750 nm. During capture, images were frame-averaged over 16 frames. To determine the presence of cyanobacteria in sporelings, two glass plates were pressed against each other, thereby squeezing the cyanobacteria cells out of the leaf pockets.

Fertilizing *A. filiculoides* spores *in vitro*

Fifty megasporocarp batches were mixed during 20 s with microsporangia at various ratios in 1 ml water, then incubated in darkness and room temperature (RT) for 2–13 d (fertilization periods), before megasporocarps (without free microsporangia and massulae) were transferred to *Azolla* growth medium (8 ml) and left to develop in 25°C day : 15°C nights with 12 h light (40–70 μmol m^{-2} s^{-1} PPFD). After 2 wk development germination was scored at regular intervals. All conditions were tested in triplicate.

Testing and optimizing cryopreservation protocols for long-term storage of *A. filiculoides*

The cryopreservation protocols summarized in Table 1(a) were tested on batches of 20 megasporocarps. Each condition was evaluated in duplicate. For cryopreservation involving cryoprotectant pretreatment, 1 ml of the cryoprotective solution was added to the megasporocarps immediately before snap-freezing in liquid nitrogen (LN). Batches were kept for 5 min in LN, then thawed on a heating plate set at 30°C for 1.5 min and washed thrice in 1 ml medium. To test cryopreservation using a drying pretreatment, megasporocarps were dried for 1, 4 and 8 d in the fume hood at RT and then snap frozen in LN without added fluid. Batches were kept for 5 min in LN then thawed for 1.5 min at 30°C before adding 1 ml medium. For each cryopreservation protocol, a treatment control was included that was not frozen but still exposed to the cryopreservation pre-treatment; additionally two controls were included that received no treatment. Germination was as in the Documenting megaspore germination section.

In order to optimize cryopreservation involving drying as pretreatment, megasporocarps with attached massulae were collected from sediment and the number of megasporocarps in each batch was increased to 250–600 megasporocarps. Drying conditions tested included 1, 4, 7, 16 and 32 d drying at RT in the fume hood as above. Additional batches were dried for 1, 4 and 7 d at constant temperature (CT) of 26°C. For each cryopreservation pre-treatment a control batch was included that was exposed to the cryopreservation pre-treatment only. Two batches were included that were not subjected to any cryopreservation treatment. These batches served as controls and allowed comparison of germination rates between megaspores collected from plant material and spores collected from growth container sediments. The freezing, thawing and germination of the batches was as above.

In order to test long-term storage, batches of spores collected from sediment and dried for 7 d at RT were frozen in LN, then transferred to −80°C and stored for 1 d, 1, 2, 5 and 7 months. Another set of dried batches was frozen at −20°C and then stored for only 1 and 2 months. A final set of batches was not subjected to a drying treatment, but directly frozen in the −20°C freezer,

Table 1 Cryopreservation protocols tested on Azolla filiculoides spores. (a) Cryopreservation pre-treatment definitions[1] and (b) Germination rate and viability for unfrozen and frozen megasporocarps after cryopreservation pre-treatment[2]

Cryopreservation pre-treatments	Concentration/treatment time	Previously reported use
(a)		
DMSO + Glycerol	2 M DMSO, 3.2 M Glycerol	Tree fern, *Dicksonia Sellowiana,* Rogge *et al.* (2000)
DMSO + EG + PVP	2 M DMSO, 4.8 M EG, 2% PVP	Zebrafish embryo, Riesco *et al.* (2012)
Sucrose	0.4 M Sucrose	Fern, *Pteris adscensionis*, Barnicoat *et al.* (2011)
Trehalose	0.4 M Trehalose	Fern, *Pteris adscensionis*, Barnicoat *et al.* (2011)
DMSO + EG + Glucose + Trehalose	2 M DMSO, 2.4 M EG, 3.2 M Glucose, 0.4 M Trehalose	Duckweed, *Lemna Minor*, Parsons & Wingate (2012)
Drying	1 d, 4 d, 8 d	Tree fern, *Dicksonia Sellowiana*, Rogge *et al.* (2000)

	No freezing			Freezing		
Pre-treatment	*N*	Viability (%)	Germination (%)	*N*	Viability (%)	Germination (%)
(b)						
None (control)	32	67	23	42	0	0
DMSO + Glycerol	29	93	7	34	0	0
DMSO + EG + PVP	36	79	23	41	0	0
Sucrose	37	92	29	49	0	0
Trehalose	50	82	57	70	0	0
DMSO + EG + Glucose + Trehalose	35	71	48	71	0	0
1 d drying	31	51	9.40	36	19	0
4 d drying	24	24	21	27	30	0
8 d drying	31	38	10	36	25	6

[1]Dimethyl sulfoxide (DMSO), ethylene glycol (EG) and Polyvinyl pyrrolidone (PVP).
[2]Viability is the percentage of megasporocarps with floats. Freezing was in liquid nitrogen. Each condition was tested in duplicate. *N*, number of megasporocarps tested.

but these repeatedly had zero germination. Zero and 7 months storage were tested in duplicate and triplicate, respectively.

RNA sequencing and quantitative RT-PCR (qRT-PCR)

Ferns, in general, and heterosporous ferns like *Azolla*, in particular, represent a particularly under-sampled group with regard to sequence information. Tissues from three different developmental phases of the complex lifecycle were therefore chosen for RNA sequencing (RNA-seq) to capture the transcriptome of *Azolla*. Microsporocarps and megasporocarps were plucked from sporulating *Azolla* grown on demineralized water in the glasshouse. Sporophytes were sterile, nonsporulating, grown in medium with and without nitrogen, and collected at 6 h intervals over 24 h starting 1 h before dawn. Total RNA was extracted (Spectrum Plant Total RNA Kit; Sigma-Aldrich) from megasporocarps, microsporocarps and sporophytes, then digested with DNase I. Two replicate extractions were pooled for each of the reproductive tissues and eight replicate extractions were pooled for the sporophyte tissue (corresponding to four time points and growth with/without nitrogen). Poly(A+) RNA was enriched using oligo (dT) Dynabeads (Ambion). To enrich capped transcripts cDNA was synthesized using the Clontech SMARTer kit (Takara Bio Europe SAS, Saint-Germain-en-Laye, France). Libraries were made from 100 ng template independently for each of the three tissues using the Ion Plus Fragment Library Kit with Ion Xpress™ Barcode Adapters (BC12, BC13 and BC01; Life Technologies). PCR amplifications and emulsions were generated using the Ion PGM™ Template OT2 400 Kit and Ion OneTouch™ 2 System (Life Technologies) with emulsions at 8 pM. Sequencing was completed with the Ion 316™ Chip v2 on a Ion PGM™ sequencer.

The resulting sequencing reads were inspected with the FastQC tools (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Using fastx tools (http://hannonlab.cshl.edu/fastx_toolkit/), reads were trimmed by removing bases with a Phred score < 20; reads which were pruned to < 50% of their original length by this step were discarded. The remaining reads were filtered for those which had < 90% of bases with a Phred score above 20 and reads shorter than 50 bases were discarded. The trimmed and filtered reads were assembled using CLC Genomics Workbench software (CLC Bio, Aarhus, Denmark) with default parameters (Brautigam *et al.*, 2011b). Quality was assessed based on read length distribution and unigene annotation (Schliesky *et al.*, 2012). The unigene database was annotated using BLAT (Kent, 2002) against proteins from *Selaginella moellendorffii*, *Arabidopsis thaliana* and *Nostoc azollae* to identify the proportion of genes similar to those in plant and symbiont genomes (Supporting Information Notes S1).

The unigene database was uploaded to the KEGG Automated Annotation Server (KAAS, http://www.genome.jp/tools/kaas/) to test the coverage of common pathways (Moriya *et al.*, 2007). The resulting maps were exported from the server, curated for pathways not present in plants, and sorted according to content (proteins present in the unigene database in green colour; see Notes S2–S4).

All reads were then mapped onto the unigene database with default parameters using CLC Genomics Workbench (Table S1). The relative read counts for each unigene were normalized by the total read counts in each tissue (reads per million mapped reads; rpm). To find signature genes for the microsporocarp, megasporocarp and sporophyte, a small group of unigenes (1%, 3%, and 3%, respectively) with read coverage above 100 rpm were identified for each tissue. Annotation of the highly read unigenes against *A. thaliana* served to assign ontology terms in the ontology list derived from MapMan (Usadel *et al.*, 2006, 2009) in Excel. Fisher's Exact test was then applied to evaluate whether the highly read unigenes were enriched in any one ontology term of pathways (Fisher, 1922); *P*-values were corrected for multiple hypothesis testing by the Bonferroni method.

In order to validate the unigene sequences and quantify gene-expression, RNA was extracted from *A. filiculoides* nonsporulating sporophytes grown in the cabinet, sporulating sporophytes collected from the wild (September 2013) and microsporocarps, and then cDNA synthesized. The primers for qRT-PCR were for the references *AfTUBULIN* (AfTUBF: CCTCCGAAAACT CTCCTTCC; AfTUBR: GGGGGTGATCTAGCCAAAGT) and *AfADENINE PHOSPHORIBOSYLTRANSFERASE* (AfAPTF: TAGAGATGCATGTGGGTGCAGT; AfAPTR: AAAA GCGGTTTACCACCCAGTT) (Salmi & Roux, 2008). Further qRT-PCR primers were for *AfFT* (AfFTF: AAGAGATTTG GCAAGCTGGA; AfFTR: TAGCAACCACCAACAGCATC), *AfSOC1* (AfSOCF: ATGGGATCGTAAGGCTTCAAAA; AfSOCR: AGCAGAGCACACAGGTCTCAAC), *AfLFY* (AfLFYF: GCGGCAAGAGGAAGAGATAGA; AfLFYR: AGT GGATGTGCTCTTGCTGAA) and *AfCAL* (AfCALF: TTTG CATCTTTCGCTCTCTCA; AfCALR: CCAAGCTGCACAA TGTAAGGA). Data was from three biological replicates, significance was assessed by *t*-test with $P < 0.05$.

Results

Azolla filiculoides spores can be collected year round

A difficulty for researching sexual reproduction in *Azolla* is access to sporulating populations. Storage of sporulating *A. filiculoides* in rainwater at temperatures varying from 5 to 15°C and 14 h light with minimum intensity of 40 μmol m^{-2} s^{-1} PPFD d^{-1} resulted in a constantly sporulating population. Megasporocarps were collected by subsequent sieving and purification. Yields from plants varied from 2890 (December 2012) to 55 190 (June 2013) megasporocarps purified from 1 m^2 of *Azolla* mat (standing crop density 2.1–2.8 kg FW m^{-2}). By contrast, yields from the sediment were much higher, ranging from 157 600 (December 2012) to 343 740 (August 2013) purified from 1 m^2 of *Azolla* mat. Sediment FW increased as organic material accumulated, reaching an average of 14 kg m^{-2} in August 2013. Year-round access to large amounts of spores allowed testing of *in vitro* germination, *in vitro* fertilization and preservation methods, and allowed extraction of RNA from sporocarps.

Cotyledon emergence, not float emergence characterizes megaspore fertilization

Key stages in the sexual reproduction of *Azolla* were documented by three dimensional reconstruction of digital images from dark-field microscopy. Sporocarps developed under the sporophyte (Fig. 1a) in pairs, first as microsporocarp pairs, then as the sporophytes reached maturity as micro and megasporocarp pairs (Fig. 1b). The megasporocarp (Fig. 1c), containing the female megaspore and cyanobacteria akinetes, quickly sunk. Detached microsporocarps generally burst open, releasing the microsporangia (Fig. 1d); once these burst, massulae were released and

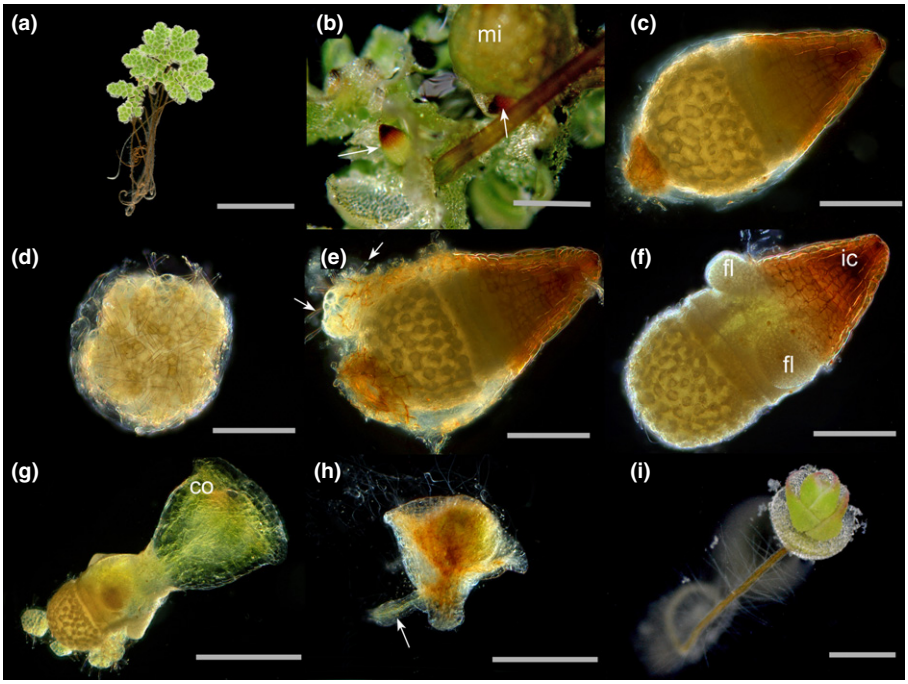

Fig. 1 Key stages in the sexual reproduction of *Azolla filiculoides*. (a) Sporophyte; bar, 10 mm. (b) Megasporocarps (arrows) and microsporocarp (mi) at the underside of a sporulating plant; bar, 1 mm. (c) Detached megasporocarp; bar, 0.2 mm. (d) Microsporangium containing four massulae; bar, 0.2 mm. (e) The massulae's glochidia (arrows) allow it to attach to the megasporocarp; bar, 0.2 mm. (f) Megasporocarp floats (fl) emerge from beneath the indusium cap (ic); bar, 0.2 mm. (g) Cotyledons (co) push away indusium cap and emerge from the megasporocarp; bar, 0.5 mm. (h) Sporeling with root (arrow) detaches from the megasporocarp; bar, 0.5 mm. (i) *Azolla* sporeling floating; bar, 1 mm.

entangled in filamentous appendages of the epispore wall of the megaspore (Fig. 1e). To study spore fertilization and preservation, morphological changes denoting successful fertilization and germination were characterized.

The first morphological change to the megasporocarp was elongation of the megasporocarp and emergence of floats from underneath the indusium cap (Fig. 1f). Nonfertilized megaspores also developed floats, especially when nutrients were present in the medium. Float development could be the result of independent development of the female gametophytes, or rapid proliferation of the cyanobacteria pushing the floats outward. Hence, megasporocarps with floats did not automatically imply that fertilization had taken place but instead revealed whether or not the megaspores were viable. Fertilized viable megaspores developed further into sporelings with a cotyledon (Fig. 1g). Root development was visible shortly after (Fig. 1h). Sporelings became sufficiently buoyant to float at the surface when first leaves appeared (Fig. 1i).

Germination rates up to 30% were achieved after *in vitro* fertilization, of which 95% occurred within 26 d (Fig. 2a). Nutrient availability promoted germination, but was not essential: when a 1 : 1.5 mixture of megasporocarps and microsporangia was placed in medium or demineralized water, germination rates reached 30.1% and 8.1%, respectively. In both cases no nitrogen was present, indicating that sufficient nitrogen reserves were present to allow germination.

To conclude, observations on spore germination revealed that the emergence of cotyledons but not of floats was a reliable measure of megaspore fertilization.

Azolla megasprores can be fertilized *in vitro*

Controlled fertilization is a prerequisite for breeding. We therefore set out to test whether *Azolla* spores could be fertilized *in vitro* and how much time would be required for this process. Megasporocarps collected from sporophytes were incubated for various durations with microsporangia, from burst microsporocarps, then transferred to fresh medium in the absence of massulae and left to germinate. Germination frequencies obtained varied from 7% to 27% (Fig. 2b). Megasporocarps without added microsporangia did not develop sporelings, confirming that megaspores were not fertilized when attached to the sporophytes nor during collection. Incubation for fertilization beyond 6 d from 9 to 13 d increased germination frequencies above 20%. Useful *A. filiculoides* spore fertilization rates were thus obtained *in vitro*, within a practicable time span.

Drying rather than cryoprotectant pre-treatment permits spore survival to cryopreservation

In order to preserve the diversity of natural varieties or varieties developed for breeding and further dissemination, various stages of *Azolla* will need to be preserved over long periods of time without loss of viability or genetic alterations. Since spores are the natural dissemination stages of *Azolla*, we tested whether these could be simply preserved by drying, but spores in dried macrosporocarps were never viable when extending storage to 4 wk.

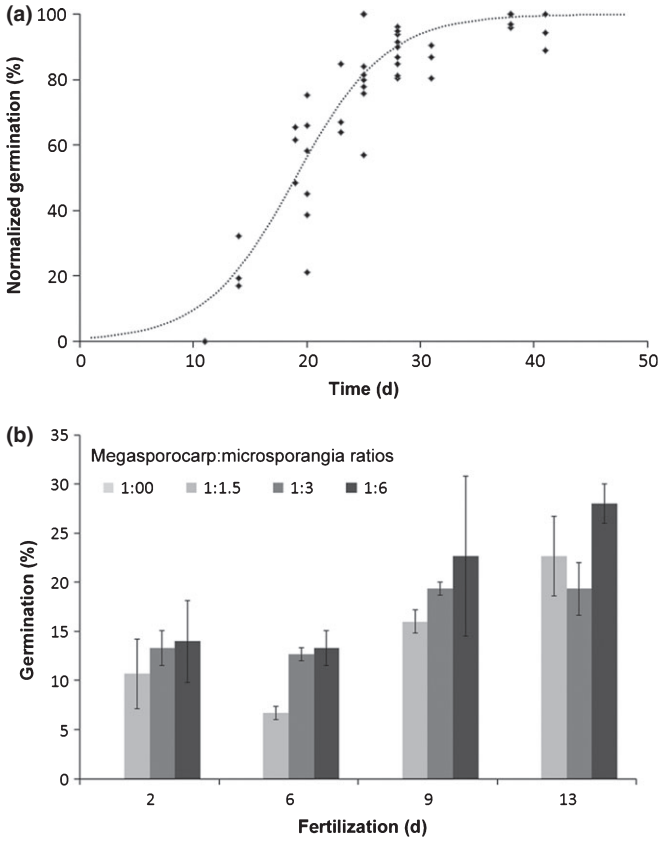

Fig. 2 *Azolla filiculoides* germination and *in vitro* fertilization. (a) Germination frequencies over a 5-wk period: germination was scored when cotyledons had emerged from megasporocarps, data from several experiments, frequencies were normalized by dividing the number of germinated megasporocarps at given time points by the number of germinated megasporocarps after 40 d for each experiment. The trendline is the best-fit sigmoid function with $\beta = 0.4$ and $T0.5 = 18$ d. (b) Germination rate as a function of fertilization period, for different megasporocarp: microsporangia ratios. Spores were mixed, incubated in darkness at room temperature (RT) for 2, 6, 10 and 13 d for fertilization, then megasporocarps only transferred to growth medium for germination. Germination was scored after 40 d. Error bars, ± SE.

Alternatively spores were frozen directly in −20°C or liquid nitrogen (LN), either in medium or without, but neither gave viable spores.

We then tested cryopreservation protocols (Table 1a) employing cryoprotectants. Cryoprotectant mixtures did not generally affect the viability of megaspores or germination frequencies (Table 1b, *No freezing*). Cryoprotectant mixtures, however, did not permit survival to the freeze/thaw cycle (Table 1b, *Freezing*). Sporelings were solely recovered when spores had been dried in a fume hood for 8 d at RT before freezing. Drying pre-treatments moreover allowed 10–30% of the megasporocarps to develop floats, an indication for viability, whilst all other pre-treatments did not (Table 1b, *No freezing*- viability). The low germination of 23% in the untreated control megasporocarps revealed that only about a quarter of the megaspores used for this screening experiment were fertilized. We concluded that drying pre-treatment changed the physiology of fertilized *A. filiculoides* megaspores so as to resist freeze/thaw cycles.

Table 2 Survival of *Azolla filiculoides* spores to the freeze/thaw cycle after various drying pre-treatments

Pre-treatment	Freezing	*N*	Germination (%)
None (control)	No	281	50.96
RT 1 d	LN	391	0.00
RT 4 d	LN	382	0.79
RT 7 d	LN	247	5.66
RT 16 d	LN	283	1.06
RT 32 d	LN	305	2.62
CT 26°C 1 d	LN	371	4.86
CT 26°C 4 d	LN	340	26.50
CT 26°C 7 d	LN	279	50.62

Megasporocarps collected from sediment were dried 1–32 d at either fluctuating room temperature (RT) of constant temperature (CT) of 26°C. *N*, number of megasporocarp tested. Megasporocarps were either not frozen (No) or frozen in liquid nitrogen (LN), thawed and then germinated as described in the Materials and Methods section.

We extended the drying pre-treatment at RT to 32 d and tested drying at higher temperature of 26°C, simulating natural drought conditions. To improve the percentage of fertilized megaspores, megasporocarps from sediment were used which also allowed using greater numbers of megasporocarps per condition to increase the sensitivity of our test. A germination frequency of 51% in the untreated control revealed that half of the megasporocarps from sediment were fertilized and viable (Table 2). The optimal duration of drying at RT was 7 d, after which the frequency of survival decreased. Constant temperature (CT) at 26°C during the drying improved survival over the lower and fluctuating RT tremendously: the batch dried at CT 26°C for 7 d before the freeze/thaw cycle reached 50% germination, which was nearly equal to that of the untreated control (51%) and almost 10 times higher compared to 7 d drying at RT (6%). Spores dried for 7 d at RT and 26°C had a water content of 18.0% and 13.1%, respectively. Hence, drying at 26°C was more efficient and improved survival rates of cryopreserved spores up to a level nearly equal to that of untreated spores.

Nostoc azollae symbionts survive and cryopreservation of the symbiosis is possible without loss of viability for at least 7 months

Survival of the *N. azollae* symbionts residing within the megasporocarp cone tip is especially important as the symbionts fix nitrogen, which is of agronomic importance. Filamentous cyanobacteria were present in *Azolla* fronds recovered from cryopreserved megasporocarps (Fig. 3a) and exhibited heterocysts at equal frequency to untreated controls (Fig. 3b). In addition, all cryopreserved batches from drying optimization experiments yielded sporelings that grew on nitrogen-free medium except those batches subjected to < 4 d drying pre-treatment at RT.

In order to test whether spore viability persisted when frozen over longer periods, batches of fertilized megasporocarps were dried for 7 d at RT, then snap frozen in LN and stored at −80°C (or frozen and stored at −20°C) for up to 7 months (Fig. 3c).

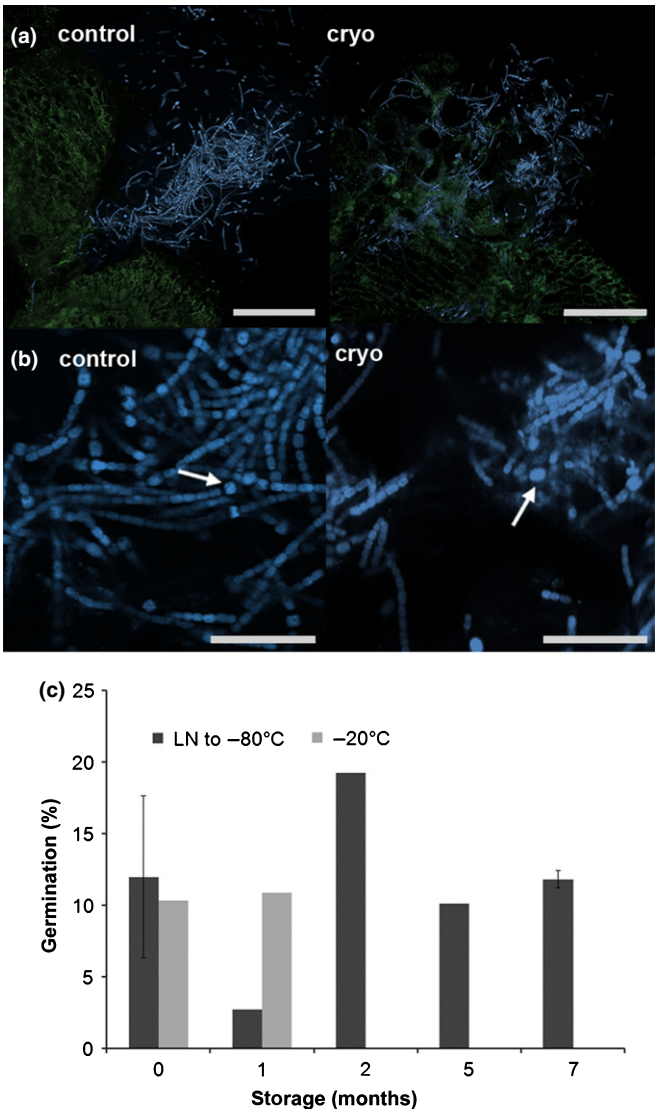

Fig. 3 Survival of the symbiont *Nostoc azollae* and long-term cryopreservation of the *Azolla* symbiosis. (a) *N. azollae* filaments in 7 wk *Azolla* fronds developing from a fresh megasporocarp (control) and a megasporocarp cryopreserved for 2 months at −80°C (cryo); bar, 0.15 mm. Confocal fluorescence microscopy with excitation 534 nm and fluorescence recorded at 680–750 nm (green) and 630–670 nm (blue). (b) Close-up of the filaments as in (a) depicting heterocysts (arrows); bar, 37.5 μm. (c) Germination of cryopreserved spores stored up to 7 months. Megasporocarps with massulae attached were dried for 7 d at room temperature (RT), batches of *c*. 200 megasporocarps were frozen in liquid nitrogen then stored at −80°C or frozen and stored at −20°C. Thawed megasporocarps were scored for germination after 5 wk. Batches stored at −20°C were not viable beyond the first month storage with zero germination after 2 and 5 months. Error bars, ± SD, $n = 3$ batches for 7 months storage.

Germination rates varied between 2.72% and 19.20% and after 7 months averaged 13%. Although a large amount of variation was observed there was no loss in viability related to the storage period over at least 7 months.

In our hands cryopreservation of neither sporophytes nor small explants was successful and therefore if the cryopreservation of *Azolla* varieties were to be contingent on the availability of spores,

a method to reliably induce sporulation will need to be developed. Given the absence of reports on successful induction of sporulation, we chose to investigate molecular pathways that may control transition to sexual reproduction.

A 37 649 unigene database of *A. filiculoides* covers metabolism, cellular processes and regulatory networks extensively

A. filiculoides belongs to a particularly neglected phylogenetic group of heterosporous ferns, the *Salviniaceae*, for which no sequence resource exists. To provide for phylogenetic and molecular studies, sequencing reads were obtained independently from three differing stages of the complex lifecycle of *Azolla*: microsporocarps, sporophytes and megasporocarps. Microsporocarps did not contain contaminating *N. azollae* but were not aseptic. Sporophytes were aseptically grown but contained contaminating *N. azollae*. Megasporocarps, central for *Azolla* reproductive biology, contained contaminating *N. azollae* and were not aseptic. Sequencing reads were cleaned and assembled into a unigene database. Because the assembly could not be benchmarked against closely related species, reads were aggressively cleaned to preclude erroneous assemblies removing two-thirds of all reads (Table S1, Read cleaning). The resulting reads were then assembled into a database comparable with previous assemblies (Notes S1; Brautigam & Gowik, 2010). Annotation against *A. thaliana*, *S. moellendorffii* and *N. azollae* matched two-thirds of the unigene database against the plants but less than one-fifth against the symbiont (Table S1). All but 41 unigenes matched by the symbiont also matched the plants, in most cases with better Expectation values (e-values; Table S2). mRNA purification and cDNA library synthesis thus efficiently discriminated against bacterial transcripts.

Without a close relative with a sequenced genome, the unigene database was benchmarked against core plant pathways as represented in KEGG. Unigenes in the database covered all nuclear encoded genes of both the light and dark reaction of photosynthesis, all but two genes required to synthesize all 20 amino acids, all genes required to synthesize both purine and pyrimidine nucleotides, and the genes for sulfur and ammonia assimilation (Notes S2). Starch and sucrose synthesis, the TCA cycle and glycolysis were completely covered. The synthesis and modification of fatty acids and lipids were largely represented; phenylpropanoid metabolism including lignin precursors, terpenoid synthesis including carotenoids, porphyrin and chlorophyll synthesis were completely covered. Gene groups involved in cellular maintenance such as peroxisomes, the proteasome, the ER including ER trafficking, DNA replication and repair, RNA synthesis and processing were well represented (Notes S3). The regulatory pathways of circadian rhythm, hormone perception and pathogen perception were essentially complete (Notes S4). In summary, the unigene database of *A. filiculoides* containing 37k unigenes covers central metabolism entirely, and cellular processes and regulatory networks to a large degree.

All reads were then mapped on the unigene database: 77%, 76% and 74% of reads from for megasporocarps, microsporocarps and sporophytes, respectively, matched unigenes of the database. Unigenes were labeled 'expressed' if at least one read could be detected. The majority of unigenes – 20 598 out of 37 649 – were expressed in all tissues, while only 7% at most were specific to sporophyte tissue (Fig. 4a). mRNA extracted from the differing tissues were thus reproducibly from the fern *A. filiculoides*.

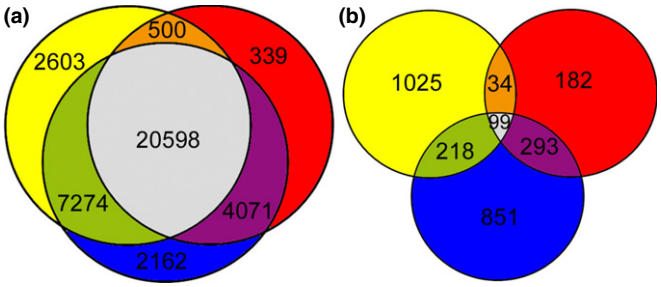

Fig. 4 Distribution of unigenes over megasporocarps, microsporocarps and sporophytes. mRNA was extracted from megasporocarps, microsporocarps and nonsporulating sporophytes, sequenced, then sequences assembled in a unigene database as described in the Materials and Methods section. 77%, 76% and 74% of reads matched unigenes and were used to assign unigene distribution over the differing tissues: microsporocarp (blue), megasporocarp (red), sporophyte (yellow). (a) Distribution taking all unigenes into account. (b) Distribution of highly expressed unigenes counting at least 100 reads per million reads.

The reproducibly high proportion of mapped reads in all tissues allowed us to test whether highly abundant reads from each tissue reflect tissue biology: when only unigenes with rpm > 100 were considered, the different tissues revealed little intersection and thus expression was characteristic (Fig. 4b). Sporophytes exhibited highly read unigenes enriched for Calvin cycle, core nitrogen metabolism, photorespiration, photosystem I and chlorophyll synthesis over other pathways (Table 3); a read signature typical of leaf tissues (Brautigam *et al.*, 2011a,b; Gowik *et al.*, 2011). By contrast, highly read megasporocarp unigenes were enriched in storage protein synthesis, in mitochondrial electron transfer, and ATPase, and in the syntheses of sterols and derivatives (Table 3). Highly read unigenes from microsporocarps were similarly enriched in mitochondrial electron transfer including ATPase and synthesis of sterol and sterol derivatives, and in addition, in proteasome, in cytosolic ribosomes as well as lipid synthesis and transfer proteins. Signatures of highly read unigenes therefore confirmed that both reproductive tissues engaged in storage compound synthesis fueled by catabolic metabolism, which was consistent with collection of the reproductive organ when still on the sporophyte to minimize microbial contamination.

Pathways leading to the onset of reproductive organ formation in Arabidopsis are present in *Azolla*

Genes controlling the onset of flowering in Arabidopsis may possibly also be involved in the transition to the reproductive phase in ferns: 165 unigenes from the *A. filiculoides* database were most similar to Arabidopsis genes involved in flowering and correspond to 64 different Arabidopsis genes (Table S2). Table S3

Table 3 Pathways enriched in *Azolla* megasporocarps (mega), microsporocarps (micro) and sporophytes

Pathway	Unigenes present in all	Highly expressed in			Percentage highly expressed in			Fishers Exact test for enrichment in[3]		
		mega (475)[1]	micro (1144)[1]	sporophyte (1025)[1]	mega (1%)[2]	micro (3%)[2]	sporophyte (3%)[2]	mega	micro	sporophyte
Calvin cycle	37	0	2	7	0	5	19	1	3.13E-01	**5.65E-05**
Nitrogen – core assimilation	27	0	0	13	0	0	48	1	1	**6.30E-14**
Photorespiration	46	2	3	13	4	7	28	1.15E-01	1.65E-01	**1.97E-10**
Photosynthesis PS I	27	0	1	8	0	4	30	1	5.67E-01	**4.27E-07**
Chlorophyll	53	0	1	7	0	2	13	1	1	**5.79E-04**
Storage protein	26	7	12	1	27	46	4	**2.69E-08**	**4.05E-12**	5.14E-01
Lipids – sterol and derivatives	70	8	12	1	11	17	1	**2.99E-06**	**1.29E-06**	1
Mitochondrial electron transfer/ATPase	132	9	16	3	7	12	2	**5.03E-05**	**2.95E-06**	1
Lipids – general	43	2	9	0	5	21	0	1.03E-01	**4.96E-06**	6.33E-01
LTP	21	3	6	2	14	29	10	2.27E-03	**2.93E-05**	1.12E-01
Proteasome	87	2	15	5	2	17	6	3.02E-01	**5.78E-08**	9.04E-02
Ribosome cytosol	197	2	34	6	1	17	3	1	**2.78E-16**	6.63E-01

[1]Number of highly expressed unigenes in this tissue.
[2]Expected % if distribution was even.
[3]Statistically significant *P*-values are indicated in bold text. *P*-values were corrected for multiple hypothesis testing by the Bonferroni method.

scores the presence of proteins associated with control of flowering in genomes of higher and lower plants and in the transcriptomes of the ferns *Ceratopteris richardii* and *A. filiculoides*. All major proteins associated with circadian rhythm and the photoperiod flowering pathways had homologs in *Azolla* tissues. Furthermore a large proportion of proteins associated with the vernalization or autonomous pathway in Arabidopsis had homologs in *Azolla*. LFY was identified in *Azolla*, and a number of SQUAMOSA BINDING PROTEIN-like (SPL) proteins were identified in both *Azolla* and *Ceratopteris*, including SPL1, SPL2 and SPL9-like in *Azolla*. A total of nine MADS-box like proteins in *Azolla* include potential homologs of SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1) and CAULIFLOWER (CAL1).

FT was not found in lower plant genomes and possibly, the integrating role of FT might have evolved along with (lignified) vasculature in ferns. An FT-like protein is present in both *Azolla* and *Ceratopteris* (Fig. 5a). The *Azolla* FT-like protein contains the motifs D-P-D-x-P-S-P-S (Fig. 5a, Box I) and G-x-H-R (Fig. 5a, Box IV) conserved for PEBP-like proteins (Hedman *et al.*, 2009). The amino acid associated with FT/TFL switching differs in the fern FT-like proteins: fern FT-like proteins have F, while Y and H are characteristic for FT and TFL1, respectively (Fig. 5a, Box II). Furthermore MFT has a characteristic P at box VI, whereas fern proteins have a corresponding D instead. In the conserved region indicated by box I both *Azolla* and *Ceratopteris* have an A at the fourth position, similar to MFT and BROTHER OF FT (BFT), instead of a V found in FT and TFL. The *Azolla* FT-like protein was repeatedly detected in microsporocarps and megasporocarps but not in the sporophyte. qRT-PCR further confirmed high expression of *Azolla FT*-like in microsporocarps compared to sporophytes (not shown) and revealed induced expression in sporulating sporophytes compared to nonsporulating sporophytes (Fig. 5b) along with *SOC1* and *LFY* (Fig. 5c). We conclude that an FT-like protein, neither characteristically FT nor MFT, is induced as sporophytes undergo reproductive development.

Mapping the *Azolla* proteins to Arabidopsis flowering pathways reveals that whole Arabidopsis pathways towards flowering are not obviously induced in reproductive tissues of *Azolla* (Fig. 6). Nonetheless *Azolla* proteins like FT, SPL1 and CAL1 are commonly detected in the reproductive organs whilst they were not detected in the sporophyte. By contrast, *Azolla* proteins like SPL9 and SPL2 seem to be restricted to the sporophyte (Fig. 6) as is the *Azolla* protein like Class II KNOX reported to repress gametophytic development in *Physcomitrella* (Table S2; Sakakibara *et al.*, 2013).

Discussion

Methods that may be invaluable for the domestication and breeding of *Azolla* are presented using the species *A. filiculoides* as a starting point. Collecting large amounts of clean fertilized megaspores will be critical for dissemination. Fertilizing spores *in vitro* will be important for breeding. Cryopreserving *Azolla* will be essential to preserve biodiversity and store varieties particularly suited for production. Importantly, cryopreservation of an *A. filiculoides* variety now opens the way to genomic investigations with the safety of cryopreserving the plant genotype sequenced. Our first RNAseq experiment showed that *Azolla* mRNA was mostly devoid of contaminating *N. azollae* RNA and generated a 37 649 unigene database that extensively covered plant metabolism, cellular processes and regulatory networks. Networks controlling the transition to reproductive development

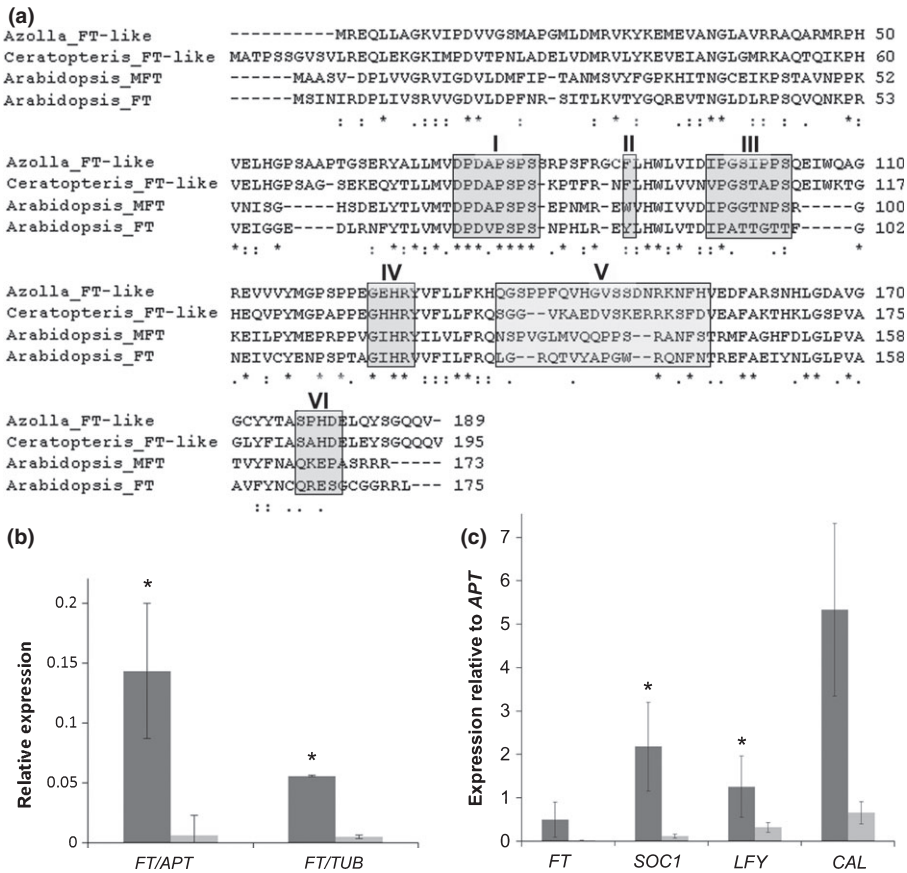

Fig. 5 *Azolla* FT-like. (a) Alignment of fern FT-like proteins with FT and MFT from Arabidopsis. Read IDs were extracted from the *A. filiculoides* database matching a theoretical DNA encoding the FT from *Ceratopteris* using BLAST then assembled using CAP3. The longest resulting contig covers the query FT for 94% of the sequence. Features are boxed: I, PEBP conserved D-P-D-x-P-S-P-S motiv; II, His/Tyr residues involved in FT/TFL switching; III, conserved MFT/FT region; IV, PEBP conserved G-x-H-R motif; V, B-region; VI, box with P characteristic for MFT (Hedman *et al.*, 2009). (b) *AfFT* expression relative to *TUBULIN* or *APT* in sporophytes sporulating (dark gray) from the wild or nonsporulating (light gray) from the growth cabinet. qRT-PCR was as described in the Materials and Methods section, $n = 3$, ± SD. (c) Expression of *AfFT*, *AfSOC1*, *AfLFY* and *AfCAL* relative to *APT* in sporophytes as in (b), $n = 3$, ± SD. *, $P < 0.05$.

in higher plants were present in *Azolla* and included an FT-like protein induced in sporulating *Azolla*.

In order to establish the above methods, we supplemented existing documentation on *Azolla* germination from Becking (1987), Peters & Meeks (1989), Wagner (1997) and Carrapiço (2010). Our conditions for *in vitro* fertilization and germination of the *A. filiculoides* megaspores resembled those described for *Marsilea vestita*, another heterosporous fern from the *Marsileaceae*, a family related to the *Salviniaceae* (Mahlberg & Yarus, 1977), although the time span required for *Azolla* germination is much greater.

Cryopreservation in ferns was achieved previously and is generally based on the desiccation and frost tolerance of the spores (Pence, 2008). Janes (1998) reported successful storage of *A. filiculoides* spores by freezing, in tap water, at −10°C, for at least 19 d. We were unable to reproduce this with material collected in the Netherlands, but this may be due to our material from plants being grown in the relatively protected environment of the glasshouse, where they were not exposed to stressful conditions. A different method for preserving *Azolla* varieties, not relying on the spores, consists of keeping a stem-tip culture of an *Azolla* frond under sterile conditions at low temperature (0–10°C), allowing preservation up to 12 months (Xu *et al.*, 2011). Preserving stem-tip cultures is laborious and precludes scaling up of the procedure. Cryopreservation is potentially more reliable and our results for spores subjected to a controlled drying pre-treatment open the way to long-term high efficiency preservation of varieties from *A. filiculoides*. Whether cryopreservation of spores from other species of *Azolla* will be similarly successful will require further investigation.

Because spores are the natural dissemination form of ferns, the control of spore induction for reliable and mass scale production of dissemination stages of any fern variety is a prerequisite to its domestication. However, very little knowledge on fern spore induction was available as research on ferns lags far behind that of angiosperms and even lycophytes and bryophytes (Muthukumar *et al.*, 2013). There is as yet no single fern genome sequenced and transcriptomes have been published from only two ferns, *Ceratopteris richardii* (Bushart *et al.*, 2013) and *Pteridium aquilinum* (Der *et al.*, 2011). Insight from these was limited, however, due to the evolutionary distance and specific ecological niche of the floating fern family of *Salviniaceae*.

RNA-seq from three different tissues of *Azolla*, the sporophyte and both micro- and megasporocarps, yielded a unigene database. Analysis of the database showed that *Azolla* shares core metabolism and regulation with genomic model plants (Notes S2). The Unigene database completely covers primary metabolism and mostly covers cellular processes and regulatory pathways. The unigene length distribution and annotation was comparable to that produced in other sequencing efforts (Brautigam & Gowik, 2010). To test whether the unigene database is in principle suitable for future quantitative RNA-seq, the reads from the three different tissues were mapped onto the assembly. Two thirds of reads could be mapped. Considering that during

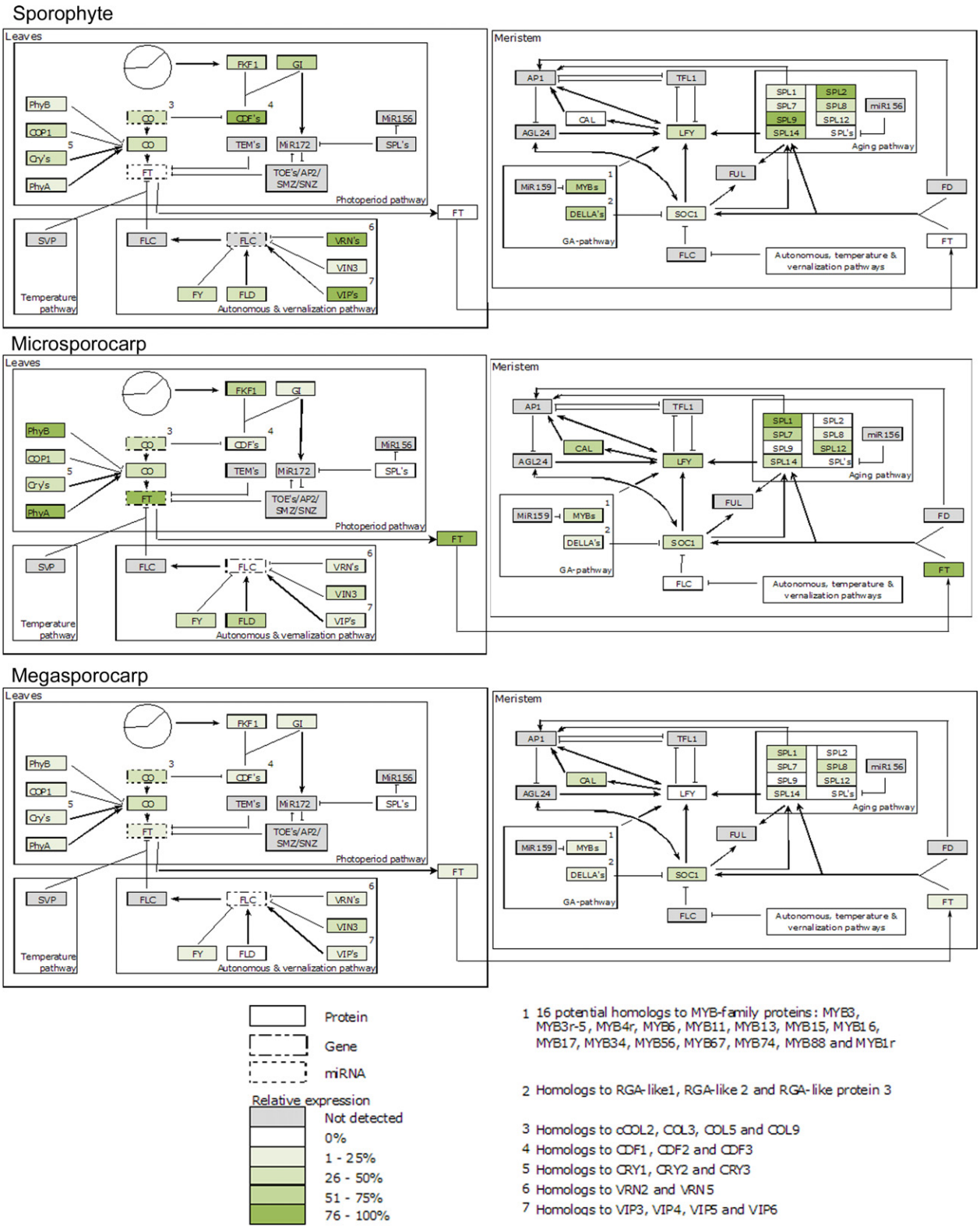

Fig. 6 Proteins from the floral induction pathways of Arabidopsis and their potential homologs in the differing tissues of Azolla: nonsporulating sporophyte, microsporocarp and megasporocarp. Layout is based on the onset of flowering in Arabidopsis with resulting transformation of the shoot apical meristem into an inflorescence meristem after Srikanth & Schmid (2011) and Jung *et al.* (2012). Presence of the unigenes in *Azolla* was extracted from Supporting Information Table S2 and normalized counts given in % over the three tissues visualized in shades of green. Gray boxes, genes not identified in the unigene database.

mapping of sequenced reads onto the corresponding genome *c.* 90% of the reads map (Hamisch *et al.*, 2012) and that during cross-species mapping between 60% (Gowik *et al.*, 2011) and 80% (Brautigam *et al.*, 2011b) of reads map on the reference sequence database, the mapping was efficient. The long read technology of Ion Torrent sequencing employed in this study thus proved to be a cost-effective way to produce a unigene database for a species without sequence resources. The low annotation frequency of the unigene database of *Azolla* (Notes S1) compared to the annotation frequencies of flowering plant RNA-seq efforts (Brautigam *et al.*, 2011b) demonstrates the evolutionary distance of *Azolla* from genomic models and indicates a large potential for new gene discovery, particularly in the light of *Azolla's* unusual secondary metabolism (Nierop *et al.*, 2011).

Assigning ontology terms of the differing pathways to the highly read unigenes indicated that read counts for the pathways reflected the biology of the tissues. Both reproductive tissues are catabolic in nature: the energy-consuming mitochondrial electron transfer chain and its ATPase were overrepresented among the highly read unigenes. The microsporocarp is very rich in lipids and in protein as judged by microscopic stains; the megasporocarp is also very rich in protein but only moderately rich in lipids (Lucas & Duckett, 1980). Microsporocarps are still in the process of maturing when attached to the plant and meiotic development has not been completed (Lucas & Duckett, 1980). Hence, the prevalence of proteasome and cytosolic ribosome components among the highly read unigenes in microsporocarps may reflect the maturation process. Meiosis as a category was not tested but selected genes involved in meiotic processes could be detected in microsporocarp mRNA, but not in the other two (Table S2).

Similar to other RNA-seq efforts, the *Azolla* transcriptome database will serve to elucidate *Azolla* biology, for example the transition to sexual reproduction and sporocarp formation. The *Azolla* unigene database contained candidate homologs of many genes controlling sexual transition in Arabidopsis. Whether these actually have a function in regulating sporulation in *Azolla*, however, is uncertain as many of the flowering-related genes have been associated with multiple developmental functions in seed plants. Several flowering-related genes were only identified because multiple tissues were included in the initial sequencing: *Azolla* proteins like FT, CAL and SPL1 were reproducibly read in the *Azolla* reproductive tissues but not in the sporophytes. By contrast, *Azolla* proteins like SPL9 and SPL2 were reproduciby read in the sporophyte whilst absent in the reproductive organs. *Azolla* FT-like expression was confirmed by qRT-PCR to be induced in sporulating as opposed to nonsporulating sporophytes. A crucial future step will be to attempt transformation of *Azolla* to test the function of genes identified in the unigene database. Only very recently, Muthukumar *et al.* (2013) reported successful transformation of the ferns *P. vittata* and *C. thalictroides* using spores as the transformation targets.

In conclusion, we present for the first time methods for storage and dissemination, as well as an annotated database of genes that may contribute to the domestication of *Azolla*, a candidate crop that is highly productive, reaching 40 tons ha^{-1} yr^{-1} dry weight yield (AzoFaSt Performance Report, 2013), requiring no nitrogen fertilizer and growing in areas not previously used as arable land. The domestication of a fern/cyanobacteria symbiosis may seem like a daunting task, but we feel that the time is ripe. In addition, results generated will serve more widely to access the wealth of biochemicals in biomass hidden within the pteridophytes for the bio-economy of the future.

Acknowledgements

We thank Ronald Leito and Frits Kindt of the Biology Imaging Center, Utrecht University, for help with microscopy. We acknowledge expert technical support by BMFZ, Heinrich Heine University, Düsseldorf. Funding was from the AzoFaSt Pathfinder, Climate-KIC, European Institute of Technology, and the Deutsche Forschungsgemeinschaft (EXC 1028). We further thank anonymous reviewers for helpful comments on the manuscript.

References

Abdel-Tawwab M. 2008. The preference of the omnivorous-macrophagous, *Tilapia zillii* (Gervais), to consume a natural free-floating Fern, *Azolla pinnata*. *Journal of the World Aquaculture Society* **39**: 104–112.

Alalade OA, Iyayi EA. 2006. Chemical composition and the feeding value of Azolla (*Azolla pinnata*) meal for egg-type chicks. *International Journal of Poultry Science* **5**: 137–141.

Antunes APM, Watkins GM, Duncan JR. 2001. Batch studies on the removal of gold(III) from aqueous solution by *Azolla filiculoides*. *Biotechnology Letters* **23**: 249–251.

AzoFaSt Performance Report. 2013. Van der Werf A, ed. *Project performance report pathfinder project "AzoFaSt"*. [WWW document] URL http://www.uu.nl/faculty/science/EN/contact/Researchinstitutes/IOEB/research/groups/mpp/research/Pages/Trehalose.aspx [accessed 1 September 2013].

Banks JA, Nishiyama T, Hasebe M, Bowman JL, Gribskov M, dePamphilis C, Albert VA, Aono N, Aoyama T, Ambrose BA *et al.* 2011. The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332**: 960–963.

Barnicoat H, Cripps R, Kendon J, Sarasan V. 2011. Conservation *in vitro* of rare and threatened ferns - case studies of biodiversity hotspot and island species. *In Vitro Cellular and Developmental Biology – Plant* **47**: 37–45.

Becking JH. 1987. Endophyte transmission and activity in the Anabaena-Azolla association. *Plant and Soil* **100**: 183–212.

Becking JH, Donze M. 1981. Pigment distribution and nitrogen fixation in *Anabaena azollae*. *Plant and Soil* **61**: 203–226.

Brautigam A, Gowik U. 2010. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology (Stuttgart, Germany)* **12**: 831–841.

Brautigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ *et al.* 2011a. An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology* **155**: 142–156.

Brautigam A, Mullick T, Schliesky S, Weber AP. 2011b. Critical assessment of assembly strategies for non-model species mRNA-Seq data and application of next-generation sequencing to the comparison of C(3) and C(4) species. *Journal of Experimental Botany* **62**: 3093–3102.

Brown NJ, Newell CA, Stanley S, Chen JE, Perrin AJ, Kajala K, Hibberd JM. 2011. Independent and parallel recruitment of preexisting mechanisms underlying C(4) photosynthesis. *Science* **331**: 1436–1439.

Bushart TJ, Cannon AE, ul Haque A, San Miguel P, Mostajeran K, Clark GB, Marshall Porterfield D, Roux SJ. 2013. RNA-Seq analysis identifies potential modulators of gravity response in spores of *Ceratopteris* (*Parkeriaceae*): evidence

for modulation by calcium pumps and apyrase activity. *American Journal of Botany* **100**: 161–174.

Carrapiço F. 2010. *Azolla* as a superorganism. Its implications in symbiotic studies. In: Seckbach J, Grube M, eds. *Symbioses and stress: joint ventures in biology. Cellular origin, life in extreme habitats and astrobiology*, **17**. Dordrecht, the Netherlands: Springer, 227–241.

Costa ML, Santos MC, Carrapiço F. 1999. Biomass characterization of *Azolla filiculoides* grown in natural ecosystems and wastewater. *Hydrobiologia* **415**: 323–327.

Der JP, Barker MS, Wickett NJ, dePamphilis CW, Wolf PG. 2011. *De novo* characterization of the gametophyte transcriptome in bracken fern. Pteridium aquilinum. *BMC Genomics* **12**: 99.

EPSO. 2005. European plant science: a field of opportunities. *Journal of Experimental Botany* **56**: 1699–1709.

Fasakin EA. 1999. Nutrient quality of leaf protein concentrates produced from water fern (*Azolla africana* Desv) and duckweed (*Spirodela polyrrhiza* L. Schleiden). *Bioresource Technology* **69**: 185–187.

Fisher RA. 1922. On the interpretation of χ^2 from contingency tables, and the calculation of *P*. *Journal of the Royal Statistical Society* **85**: 87–94.

Floyd KS, Bowman JL. 2007. The ancestral developmental tool kit of land plants. *International Journal of Plant Sciences* **168**: 1–35.

Gowik U, Brautigam A, Weber KL, Weber AP, Westhoff P. 2011. Evolution of C_4 photosynthesis in the genus *Flaveria*: how many and which genes does it take to make C4? *Plant Cell* **23**: 2087–2105.

Hamisch D, Randewig D, Schliesky S, Brautigam A, Weber APM, Geffers R, Herschbach C, Rennenberg H, Mendel RR, Hansch R. 2012. Impact of SO_2 on *Arabidopsis thaliana* transcriptome in wildtype and sulfite oxidase knockout plants analyzed by RNA deep sequencing. *New Phytologist* **196**: 1074–1085.

Hedman H, Källman T, Lagercrantz U. 2009. Early evolution of the *MFT*-like gene family in plants. *Plant Molecular Biology* **70**: 359–369.

Herd YR, Cutter EG, Watanabe I. 1985. A light and electron microscopic study of microsporogenesis in *Azolla microphylla*. *Proceedings of the Royal Society of Edinburgh* **86B**: 53–58.

Herd YR, Cutter EG, Watanabe I. 1989. The effects of temperature and selected growth regulating substances on sporulation in the aquatic fern *Azolla*. *American Fern Journal* **79**: 136–142.

IPCC. 2007. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL, eds. *Climate change 2007: the physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, UK/New York, NY, USA: Cambridge University Press.

Janes R. 1998. Growth and survival of *Azolla filiculoides* in Britain: II. Sexual reproduction. *New Phytologist* **138**: 377–384.

Jensen ES, Peoples MB, Boddey RM, Gresshoff PM, Hauggaard-Nielsen H, Alves BJR, Morrison MJ. 2012. Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries – a review. *Agronomy for Sustainable Development* **32**: 329–364.

Jung JH, Ju Y, Seo PJ, Lee JH, Park CM. 2012. The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in Arabidopsis. *Plant Journal* **69**: 577–588.

Kajala K, Brown NJ, Williams BP, Borrill P, Taylor LE, Hibberd JM. 2012. Multiple Arabidopsis genes primed for recruitment into C(4) photosynthesis. *Plant Journal* **69**: 47–56.

Kar PP, Mishra S, Singh DP. 2002. Azolla sporulation in response to application of some selected auxins and their combination with gibberellic acid. *Biology and Fertility of Soils* **35**: 314–319.

Kent WJ. 2002. BLAT – The BLAST-like alignment tool. *Genome Research* **12**: 656–664.

Lucas RC, Duckett JG. 1980. A cytological study of the male and female sporocarps of the heterosporous fern *Azolla filiculoides* Lam. *New Phytologist* **85**: 409–418.

Mahlberg PG, Yarus S. 1977. Effects of light, pH, temperature, crowding on megaspore germination and sporophyte formation in *Marsilea*. *Journal of Experimental Botany* **28**: 1137–1146.

Maizel A, Busch MA, Tanahashi T, Perkovic J, Kato M, Hasebe M, Weigel D. 2005. The floral regulator LEAFY evolves by substitutions in the DNA binding domain. *Science* **308**: 260–263.

Meyer RS, DuVal AE, Jensen HR. 2012. Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytologist* **196**: 29–48.

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research* **35**: 182–185.

Muthukumar B, Joyce B, Elless M, Stewart N. 2013. Stable transformation of ferns using spores as targets: *Pteris vittata* (Chinese brake fern) and *Ceratopteris thalictroides* (C-fern 'Express'). *Plant Physiology* **163**: 648–658.

Nierop KGJ, Speelman EN, de Leeuw JW, Reichart G-J. 2011. The omnipresent water fern *Azolla caroliniana* does not contain lignin. *Organic Geochemistry* **42**: 846–850.

Parsons JL, Wingate V. 2012. Methods and compositions for the cryopreservation of duckweed. Patent Application Publication US 2012/0190004 A1.

Pence VC. 2008. Cryopreservation of Bryophytes and Ferns. In: Reed BM, ed. *Plant cryopreservation: a practical guide*. New York, NY, USA: Springer, 117–140.

Peters GA, Meeks JC. 1989. The Azolla-Anabaena symbiosis: basic biology. *Annual Review of Plant Physiology and Plant Molecular Biology* **40**: 193–210.

Peters GA, Perkins SK. 1993. The Azolla-Anabaena symbiosis: endophyte continuity in the Azolla life-cycle is facilitated by epidermal trichomes. *New Phytologist* **123**: 65–75.

Quin-Yuan X, Yan-Ru S, Guang-Li Y, Ke-Lin P. 1987. Germination of *Azolla filiculoides* Lam. sporocarps and factors affecting their growth. In: Smith WH, Cervantes EP, eds. *Azolla utilization*, Proceedings of the Workshop on AzoIla Use Fuzhou, Fuijan, China. Los Baños, Philippines: International Rice Research Institute, 33–38.

Ran L, Larsson J, Vigil-Stenman T, Nylander JAA, Ininbergs K, Zheng WW, Lapidus A, Lowry S, Haselkorn R, Bergman B. 2010. Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. *PLoS ONE* **5**: 329–493.

Riesco MF, Martínez-Pastor F, Chereguini O, Robles V. 2012. Evaluation of zebrafish (*Danio rerio*) PGCs viability and DNA damage using different cryopreservation protocols. *Theriogenology* **77**: 122–130.

Rigbi M, Rosinski J, Siegelman HW, Sutherland JC. 1980. Cyanobacterial phycobilisomes: selective dissociation monitored by fluorescence and circular dichroism. *Proceedings of the National Academy of Sciences, USA* **77**: 1961–1965.

Rogge GD, Viana AM, Randi AM. 2000. Cryopreservation of spores of *Dicksonia sellowiana*: an endangered tree fern indigenous to South and Central America. *Cryo-Letters* **21**: 223–230.

Sakakibara K, Ando S, Yip HK, Tamada Y, Hiwatashi Y, Murata T, Deguchi H, Hasebe M, Bowman JL. 2013. *KNOX2* genes regulate the haploid-to-diploid morphological transition in land plants. *Science* **339**: 1067–1070.

Salmi ML, Roux SJ. 2008. Gene expression changes induces by spaceflight in single-cells of the fern *Ceratopteris richardii*. *Planta* **229**: 151–159.

Schliesky S, Gowik U, Weber AP, Brautigam A. 2012. RNA-Seq assembly – are we there yet? *Frontiers in Plant Science* **3**: 220.

Shuying L. 1987. Method for using *Azolla filiculoides* sporocarps to culture sporophytes in the field. In: Smith WH, Cervantes EP, eds. *Azolla utilization*. Proceedings of the Workshop on AzoIla Use Fuzhou, Fuijan, China. Los Baños, Philippines: International Rice Research Institute, 27–32.

Smith KA, McTaggart IP, Tsuruta H. 1997. Emissions of N_2O and NO associated with nitrogen fertilization in intensive agriculture, and the potential for mitigation. *Soil Use and Management* **13**: 296–304.

Sommer M, Brautigam A, Weber AP. 2012. The dicotyledonous NAD malic enzyme C_4 plant *Cleome gynandra* displays age-dependent plasticity of C_4 decarboxylation biochemistry. *Plant Biology (Stuttgart, Germany)* **14**: 621–629.

Srikanth A, Schmid M. 2011. Regulation of flowering time: all roads lead to Rome. *Cellular and Molecular Life Sciences* **68**: 2013–2037.

Toia RE Jr, Buzby KM, Peters GA. 1987. Sporocarps of *Azolla mexicana* Presl. I. isolation and purification of megasporocarps and microsporangia. *New Phytologist* **106**: 271–279.

Usadel B, Nagel A, Steinhauser D, Gibon Y, Blasing OE, Redestig H, Sreenivasulu N, Krall L, Hannah MA, Poree F *et al.* 2006. PageMan: an

interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. *BMC Bioinformatics* 7: 8.

Usadel B, Poree F, Nagel A, Lohse M, Czedik-Eysenberg A, Stitt M. 2009. A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize. *Plant, Cell & Environment* **32**: 1211–1229.

Wagner GM. 1997. *Azolla*: a review of its biology and utilization. *Botanical Review* **63**: 1–26.

Watanabe I, Roger PA, Ladha JK, van Hove C. 1992. *Biofertilizer germplasm collections at IRRI*. Los Baños, Philippines: International Rice Research Institute.

Willemse MTM. 2009. Evolution of plant reproduction: from fusion and dispersal to interaction and communication. *Chinese Science Bulletin* **54**: 2390–2403.

Xu G, Zheng X, Ye H, Wang J, Lin Y. 2011. Method for preserving *Azolla* variety. Patent CN102119659 A.

Zheng W, Bergman B, Chen B, Zheng S, Xiang G, Rasmussen U. 2009. Cellular responses in the cyanobacterial symbiont during its vertical transfer between plant generations in the *Azolla microphylla* symbiosis. *New Phytologist* **181**: 53–61.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 Summaries RNAseq and assembly

Table S2 *Azolla filiculoides* unigene database annotation

Table S3 Genes associated with reproductive phase transition in genomes of higher and lower plants and in transcriptomes of ferns

Notes S1 *A. filiculoides* unigene database.

Notes S2 *A. filiculoides* unigene database proteins in metabolism.

Notes S3 *A. filiculoides* unigene database proteins in cellular processes.

Notes S4 *A. filiculoides* unigene database proteins in regulatory pathways.

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.