Turning the aquatic weed *Azolla* into a sustainable crop

Layout by Kimberley Piek
Cover photo | design: Ronald Leito | Kimberley Piek
Printed by ProefschriftMaken || www.proefschriftmaken.nl
ISBN: 978-90-393-6788-9
LPP contribution series: 48

The reserach described in this thesis was funded by the University of Utrecht and the Laboratory of Paleobotany & Palynology (LPP) Foundation. Additional financial support was provided by Climate KIC.

Turning the aquatic weed *Azolla* into a sustainable crop

De transformatie van het aquatische onkruid *Azolla*
tot een duurzaam gewas

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op vrijdag 7 juli 2017 des avonds te 6.00 uur

door

Paul Brouwer

geboren op 28 april 1988 te Utrecht

Promotoren:	Prof. dr. J.C.M. Smeekens Prof. dr. G-J. Reichart
Copromotoren:	Dr. H. Schluepmann Dr. K.G.J. Nierop

Table of contents

Chapter 1

General introduction

Demand driven by a growing and wealthier world population

Over the last decades the human population on Earth has expanded exponentially from approximately one billion people in 1800 to 7.3 billion in 2015 (United Nations, 2015) and is expected to reach more than 9 billion by 2050 (United Nations, 2015; Lukovenkov et al., 2011). The rate of population growth is decelerating and a major driver of this deceleration is increased wealth (Godfray et al., 2010). Producing sufficient food for the additional 1.7 billion people in the next 33 years will become a major challenge, whereas today still 12.9% of the global population is undernourished (FAO et al., 2015). Higher purchasing power additionally results in a higher consumption and a greater demand for meat, dairy, and fish (Godfray et al., 2010; Speedy, 2003). The production of meat and dairy requires a 2.3 up to 13 fold input of grains as feed and thereby puts even more pressure on global food supply (Sabaté and Soret, 2014), whereas global fish stocks are declining due to overexploitation (FAO, 2016). Not only does increased wealth results in higher demand for food and feed, also the consumption of energy (Ferguson et al., 2000) and materials (Wiedmann et al., 2015) have a positive correlation to wealth. Hence a growing and wealthier world population will boost future demand for food, energy and materials, increasing pressure on Earth's ecosystem.

Negative environmental impacts of current production systems

The negative environmental impacts of our production systems are becoming increasingly well understood and more tangible. Since the start of the industrial revolution in 1760, mankind has used vast quantities of fossil fuels to power the production of food and commodities, as well as to provide for transportation, heat and, more recently, electricity. Before the industrial revolution, CO_2 drawdown by photosynthesis, weathering and ocean sedimentation, and CO_2 emission from respiration, freshwater outgassing and volcanism were in balance (Ciais et al., 2013). The amount of carbon stored in the atmosphere is only small compared to the carbon locked in soils, fossil fuel reserves, and especially oceans (Ciais et al., 2013). However, due to a slow net exchange between the atmosphere and other reservoirs, it may take decades to thousands of years until atmospheric CO_2 is relocated into terrestrial and ocean carbon resevoirs (Ciais et al., 2013). The burning of fossil resources within a limited time frame has therefore led CO_2 piling up in the atmosphere. In 2015 the average CO_2 concentration measured at the Mauna Lao observatory station exceeded 400ppm, versus 317 ppm in 1960 and 200ppm during pre-industrial times (Keeling, 2008). The emission of CO_2 and other greenhouse gasses (GHGs) by humans has caused

global temperature to rise 0.8 °C and resulted in more extreme weather events such as heat waves, droughts, floods, cyclones, and wildfires, melting of glaciers and sea ice and changing migration patterns for wildlife (IPCC, 2014). Consequently, these climatic changes have negative socio-economic impacts, such as loss of crop yields, material damage and decreased availability of drinking water (IPCC, 2014). If emissions of greenhouse gasses are not reduced, global temperature may rise more than 4°C compared to 1900, along with an exponential increase in associated impacts on environment and societies (IPCC, 2014).

Between 1960 to 2000 global grain production increased from 1 billion to 2 billion tons per year, a period often referred to as the 'green revolution' (Khush, 2001). A major breakthrough enabling this 'green revolution' was the development of crop varieties with improved response to nitrogen fertilizer (Khush, 2001). Effectively, the use of artificial fertilizer has increased 7 fold over this period (Tilman, 1998). Concomitantly, the production of nitrogen fertilizer by the Haber-Bosch process has increased terrestrial N input by 71% and global N input by 34% (Figure 1) (Gruber and Galloway, 2008). The nitrogen fertilizer is only partly incorporated in the plant biomass, as 16% of the nitrogen input into croplands may leach into the freshwater hydrological cycle and another 20% is emitted to the atmosphere in gaseous form (Wang et al., 2017). Additionally, a large share of the nitrogen in the harvested biomass, which serves human consumption, is excreted and ends up in freshwater bodies (Gierlinger, 2014; Cease et al., 2015).

The increased use of nitrogen fertilizer from 1960 onwards was accompanied by the use of phosphorous fertilizer. Unlike nitrogen, phosphorous is a much less mobile element. Volatilization is negligible and it is hardly present in a soluble form as it quickly precipitates (Smil, 2000). Run-off of phosphorous from land to oceans proceeds therefore mainly due to erosion and as particulate P (Smil, 2000). Once in the ocean, phosphorous is buried in marine sediments at rates of 25-35 M ton $year^{-1}$ (Smil, 2000). The slow process of tectonic uplift is required to bring this P back to land surfaces. Effectively, natural ecosystems depend on very efficient recycling of the available phosphorous. By mining ancient phosphorous sediments, mankind adds 15 Mt $year^{-1}$ of phosphorous fertilizer to the global phosphorous cycle, representing 52-58% of the total phosphate input of croplands (Smil, 2000).

Intensive use of artificial fertilizers has more than doubled nitrogen and phosphorous run-off into freshwater, and eventually marine ecosystems (Figure 1) (Gruber and Galloway, 2008; Smil, 2000). The run-off of phosphorous and nitrogen into fresh-water ecosystems promotes the growth of few species adapted to a high nutrient environment and inhibits growth of species that are limited by other factors, thereby decreasing biodiversity (Erisman et al., 2013). The same process occurs in coastal areas where high nutrient runoff entering the oceans causes algal blooms. By fueling microbial respiration, algal blooms can furthermore lead to coastal zones

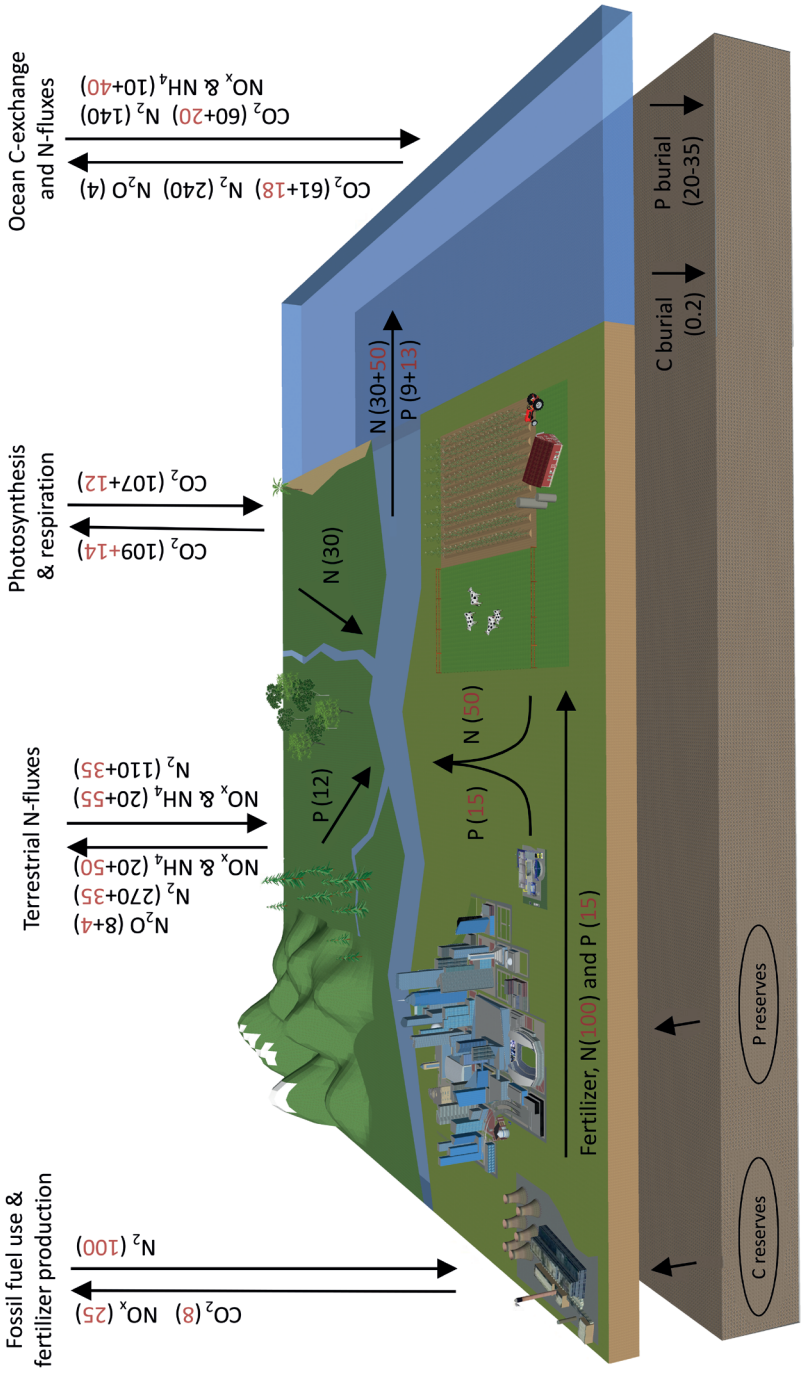


Figure 1. Influence of human activities on global carbon (C), nitrogen (N) and phosphorous (P) cycles. Numbers in red indicate fluxes due to human activity, whilst black numbers indicate natural fluxes. Data for carbon fluxes are taken from Ciais et al. (2013) and provided as billion ton carbon (Ciais et al., 2013). Nitrogen fluxes are adapted from Gruber & Galloway (2008) and provided in million tons of nitrogen (Gruber and Galloway, 2008). Phosphorous fluxes are adapted from Smil (2000) and provided in million tons of phosphorous (Smil, 2000).

becoming hypoxic or even anoxic, resulting in mass mortality among marine species (Diaz and Rosenberg, 2008). Such 'dead zones' occur in coastal areas on all continents and their frequency has doubled since 1960 (Diaz and Rosenberg, 2008).

Volatilization of ammonia is a major source of gaseous nitrogen emission (Wang et al., 2017). Additionally, ammonium salts can be oxidized (nitrified) into nitrate under aerobic conditions. In low oxygen soils or waterways nitrate can serve as an oxygen donor for respiration of denitrifying bacteria, which produce the inert N_2 gas accompanied by lower levels of gaseous NO_x and N_2O intermediates. In wastewater treatment facilities nitrification and denitrification are actively coupled to remove nitrogen from sewage and these facilities are therefore also a source of NO_x and N_2O (Kampschreur et al., 2009). Global ammonia and NO_x emissions have more than tripled due to human activities (Figure 1; Gruber and Galloway, 2008). Ammonia and NO_x are generally removed from the atmosphere within 1-5 days by precipitation or dry deposition downwind of the source (Holland et al., 2005). Deposited ammonium and nitrous oxides can thereby impact the biodiversity of waterways, grasslands and forests via acidification and eutrophication (Bobbink et al., 1998). The emitted N_2O on the other hand, remains in the atmosphere for an average lifetime of 191 years where it acts as a greenhouse gas with nearly 300 times the global warming potential of CO_2 (Hartmann et al., 2013). Human activities have increased N_2O emissions with 33%, resulting in a 20% increase of atmospheric N_2O concentrations up to 324 ppb in 2011, currently making it the third most important GHG contributing to global warming (Gruber and Galloway, 2008; Hartmann et al., 2013).

The carbon cycle and the nutrient cycles are closely interconnected. The Haber-Bosch process for the production of ammonium from atmospheric N_2 is estimated to account for 1.2% of the global energy consumption (Fischedick et al., 2014). Burning of fossil fuels further directly emits 25 M ton of NO_x into the atmosphere and thereby greatly contributes to nitrogen deposition and subsequent eutrophication (Gruber and Galloway, 2008). The uptake of CO_2 by terrestrial and marine photosynthesis on the other hand is improved by increased availability of nitrogen and phosphorous (Reed et al., 2015).

Apart from impacting global carbon and nutrient cycles, current production systems rely heavily on finite fossil resources, which makes them unsustainable by definition. Food production relies on the mining of phosphate rock. Recent reserve estimates indicate phosphate rock reserves represent 300-400 years of current production (Dawson and Hilton, 2011). The production of transportation fuels and materials relies on oil reserves, which account for 40-150 years of current production, depending on whether unconventional sources such as oil shales and tar sands are included (McGlade and Ekins, 2015; Sorrell et al., 2010; IEA, 2016b; Zou et al., 2015). Electricity and heat generation is mostly dependent on natural gas and coal, for which total recoverable resources equal 240 – 430 years of current production

(McGlade and Ekins, 2015; IEA, 2016c; IEA, 2016a).

Scarcity of these finite resources may occur before resources are depleted, when a peak in production is followed by a decline due to asymmetry in resource discovery and decreased production efficiency because resources are depleted and/or lower quality resources are exploited (Bardi, 2009). When applying this concept of peak production, scarcity is predicted to occur within this century for oil, gas, coal, as well as phosphorous fertilizer (Maggio and Cacciola, 2012; Rhodes, 2013; Cordell et al., 2009). When resource scarcity drives the exploitation of lower quality resources, this further affects energy requirements and associated environmental impacts. For example, life cycle GHG emissions of the recently expanding shale oil production are estimated between 97 – 271 kg CO_2-eq GJ^{-1} versus 71-88 kg CO_2-eq GJ^{-1} for conventional oil (Nduagu and Gates, 2015).

Hence, not only are we faced with a growing world population with a higher demand for food, fuels and commodities, the negative effects of our production system need to be reduced. By disrupting global carbon and nutrient cycles, current production systems contribute to global warming and biodiversity loss. The reliance on finite resources further threatens the security of production systems. Policies to decrease (over)consumption and human population growth are obvious solutions, but are generally not initiated due to a lack of politically feasibility (Stephenson et al., 2010; Grasso, 2016; Bryant et al., 2009). Therefore we also need to develop technological solutions to improve productivity while decreasing impact on climate, biodiversity and reliance on finite resources.

Roles of plants in future production systems

Using photosynthesis and the subsequent catalytic conversions by numerous enzymes, plants convert water, CO_2 and nutrients into a vast set of complex organic molecules. Protein, carbohydrates, lipids and polyphenols represent major groups of plant compounds and differ from eachother by their fundamental building blocks. Protein are comprised of amino acid building blocks, carbohydrates include monosaccharides (sugars) and polysaccharides, lipids are composed of either long alkyl-chains or cyclic compounds (terpenoids incl. sterols) with few functional groups and polyphenols consist of aromatic rings with often multiple hydroxyl groups and esters thereof (Figure. 2). Plant compounds are not limited to these major groups and numerous variations exist in their specific chemical structure.

Food

Plants are the basis of global food production by providing the carbohydrates, lipids and proteins and many essential vitamins required for a healthy diet. Although

some food ingredients can be chemically synthetized, the chemical complexity and variety of the nutrients synthetized by plants infers that plants will stay at the basis of food production in the foreseeable future. Increasing the variety of food textures and nutritive ingredients is now accomplished by conversion of plant biomass into animal or fungi biomass. Vitamin B12 is among the few compounds required for the human diet, which is not synthetized by plants. Although vitamin B12 can also be produced by algae and fungi (Watanabe et al., 2014; Nakos et al., 2017), its main source in the human diet is meat. Meat and dairy can be a more concentrated source of protein, fatty acids and iron than plants (McAfee et al., 2010). However, consumption of red meat and particularly processed meat in Western countries exceeds recommended values by 2-4 fold, negatively affecting health (McAfee et al., 2010; Rohrmann et al., 2013). Given that the production of meat and dairy requires a 2.3 up to 13 fold input of plant biomass, a shift towards more direct consumption of plants is likely needed to sustain future food production (Sabaté and Soret, 2014; Aiking, 2014). In parallel, alternative feed sources may be sought that reduce the environmental impact of the dairy and meat production life cycles.

Energy and fuel

As plants capture and reduce CO_2 during their growth they can, in theory, be used as a carbon-neutral energy source. Since the energy harnessed by plants is stored as high density chemical energy, plants are especially interesting for the production of transportation fuels. Combustion of biomass to produce heat and electricity only partly uses the storability and high energy density of plant biomass. Additionally, there are many other renewable alternatives to generate electricity and heat, i.e. solar energy, wind energy, geothermal energy and hydropower, that have lower environmental impact in terms of GHG emissions (Ericsson et al., 2004; Cherubini and Ulgiati, 2010). Although electric vehicles and electrified public transport may become major modes of transport in densely populated areas, long distance transport over air and water and land transport in remote areas will remain dependent on high energy-dense fuels. In these cases biofuels may be tme most suitable replacements for fossil oil-derived gasoline, diesel or kerosene (GG-MOVE, 2011; de Jong et al., 2017).

Two dominant ways to produce biofuel are currently used commercially: the fermentation of sugars into bio-ethanol by yeasts and the conversion of fats into fatty acid methyl esters, or biodiesel, using alkaline catalysts. More recently efforts have been directed towards lignocellulose: common plant cell wall matrices composed of lignin and polysaccharides, including cellulose, hemicellulose and pectin. To produce bio-ethanol from lignocellulosic feedstocks, pre-treatment is required to break down the strong lignocellulosic matrix, followed by chemical or enzymatic hydrolysis

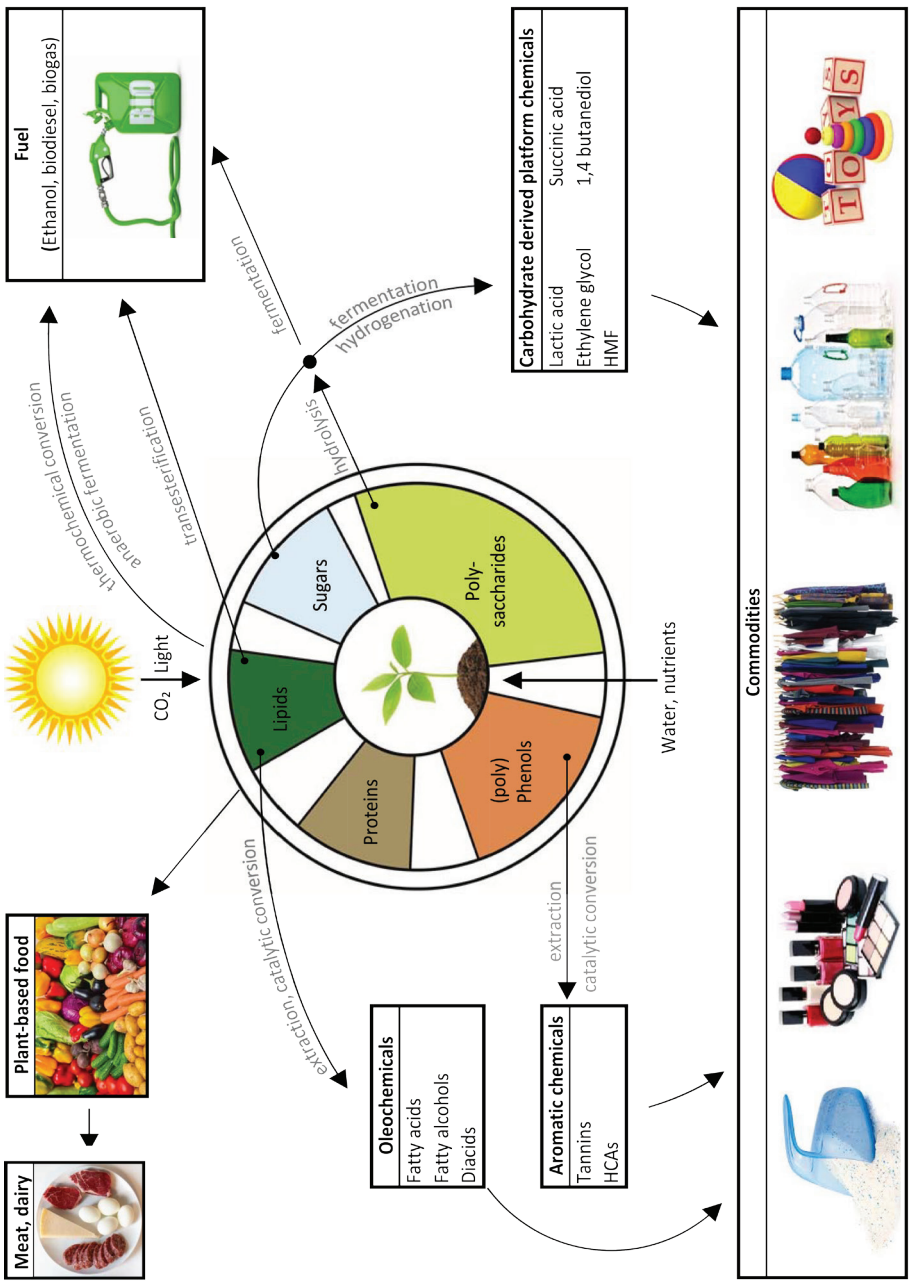


Figure 2. Routes for production of food, fuel and commodities using plants as feedstock. Abbreviations are HMF: hydroxymethylfurfural and HCAs: hydroxycinnamic acids.

into monosaccharides (Sarkar et al., 2012). Depending on their composition, these monosaccharides can then feed into the traditional bio-ethanol production process, whereas residual lignin is combusted to provide heat and electricity for the processing plant. Lignocellulosic biomass, as well as organic waste, can also be used for the production of biogas by anaerobic fermentation. By upgrading, i.e. removing CO_2 and trace compounds such as H_2S, bio-gas can be used in LPG vehicles (Rotunno et al., 2017). Although biogas has similar or greater environmental benefits compared to bio-ethanol and bio-diesel, it is much less widely adapted, likely due to higher costs and minimal infrastructure existing for gas-fuel vehicles (Börjesson and Mattiasson, 2008). Alternatively, direct thermochemical conversion processes, including pyrolysis, gasification and liquefaction, use high temperatures and/or pressures, to generate intermediates such as bio-oils by pyrolysis and syngas by gasification (Alonso et al., 2010). Intermediates are then upgraded chemically to produce drop-in hydrocarbons with characteristics similar to gasoline and diesel fuels (Alonso et al., 2010).

Chemicals and polymers

The petrochemical industry uses oil to produce various chemicals and polymers utilized in everyday commodities, including polyesters and nylon processed in clothing, soft plastic used as packaging material, soaps used in cosmetics, hard plastics applied in housing and synthetic rubbers used in the automotive industry (Roddy, 2013). In 2014, 13% of the global oil production was utilized for non-energy uses (IEA, 2016b), mainly serving the petrochemical industry. To produce these chemicals and polymers from the non-functionalized oil hydrocarbons requires oxygenation steps and other chemical modifications to add functional groups that provide the desired function or enable polymerization (Zakzeski et al., 2010). A major advantage of plant biomass is that it already contains many functionalized molecules, which can in theory directly feed into production processes, but this requires the development of novel ‘green’ chemistry. In recent years many ‘green chemistry’ approaches have been developed using carbohydrates as the starting point. By microbial fermentation and catalytic conversions, plant carbohydrates can be converted into platform chemicals such as lactic acid, ethylene glycol, hydroxymethylfurfural (HMF), succinic acid and 1,4-butadiol (Madhavan Nampoothiri et al., 2010; van Uytvanck et al., 2014; Rosatella et al., 2011; Lee et al., 2011; Burgard et al., 2016). These platform chemicals subsequently serve as the basic ingredients for the production of (partly) bio-based polymers, including polylactic acid (PLA), polyethylene terephthalate (PET), polyethylene furandicarboxylate (PEF), polyamides (PA) and poly(butylene succinate (PBS) (van Uytvanck et al., 2014; Datta and Henry, 2006; Eerhart et al., 2012; Bechthold et al., 2008). These bio-based polymers currently represent a 4.16 million ton market,

replacing oil-based polymers in packaging, textiles, automotive, consumer electronics, toys, etc. (European Bioplastics, 2016).

Lipids are traditionally used in commodities such as soaps and candles, but can also be converted into bio-chemicals, also referred to as oleochemicals. Fatty acids can be hydrogenated into fatty alcohols, which are widely used as surfactants in washing and cleaning formulas (Behr et al., 2008; Biermann et al., 2011). Fatty acids with additional functional groups can be produced by catalytic conversion of unsaturated fatty acids, or extracted from specific plants that produce them naturally (Behr et al., 2008). For example, ricinoleic acid (12-hydroxy-9-octadenenoic acid) extracted from castor oil is used for the production of polyurethanes (PU), which represented 41.6% of the total bio-based polymer market in 2016 (European Bioplastics, 2016). Glycerol is an abundant byproduct of the industrial conversion of plant triglycerides into fatty acids and biodiesel, and is used in cosmetics, paint, automotive, food, tobacco, pharmaceutical, pulp and paper, leather and textile industries (Behr et al., 2008; da Silva et al., 2009). Using glycerol as the carbon source for industrial microbiology, further allows for bioconversion into platform chemicals, such as 1,3-propanediol, dihydroxyacetone, ethanol and succinate (da Silva et al., 2009).

Plant polyphenols contain one or multiple aromatic rings, making them obvious replacements for fossil-oil derived benzene, toluene, and xylenes (BTX), which are important building blocks for the petrochemical industry (Bruijnincx and Weckhuysen, 2013). Lignin is a complex insoluble polyphenol that is highly abundant in plants and therefore has been widely investigated for the production of aromatic chemicals by means of catalysis, fungal, enzymatic and thermochemical conversions (Zakzeski et al., 2010; Gupta et al., 2016). Soluble plant polyphenols are generally less abundant in plants, but are more easily extractable compared to lignin and are already widely utilized in industry. Soluble tannins are used in leather production and novel applications include insulating foams and adhesives (Pizzi, 2006; Kemppainen et al., 2014). Soluble hydroxycinnamic acids and esters thereof find applications as active ingredient in medicines and as formulation stabilizers in food, cosmetic and pharmaceutical industry (Silva et al., 2014). p-Coumaric and caffeic acid have further been proposed for the synthesis of bioplastics resistant to high temperature (Chauzar et al., 2012).

The search for novel, sustainable, plant feedstocks

Plants will remain at the core of global food production and plant based production of fuels and chemicals is increasing due to concerns over global warming and oil scarcity. The oil crisis in 1973 stimulated several governments to formulate biofuel policies. As a result Brazil is now the largest producer of bio-ethanol from sugarcane, supplying 40% of its domestic transportation fuel and exporting another 100 million gallons annually (Solomon et al., 2007). The USA is the second largest producer of

bio-ethanol, mainly from corn, whereas the European Union is the largest producer of biodiesel, mostly from rapeseed (Solomon et al., 2007; Demirbas, 2009).

The sustainability of producing fuels and chemicals from conventional crops, however, has been questioned. With a notable exception for sugarcane ethanol, life cycle assessment (LCA) of these biofuel production systems indicated lower GHG savings than expected due to direct and indirect use of fossil fuels, land-use change and fertilizer input (Cherubini and Ulgiati, 2010). Heavy reliance on fertilizers further impacts global nitrogen and phosphorous cycles and increases dependence on finite phosphate resources. Between 2005 and 2008 a spike in world prices of food commodities, including rice (+25%), wheat (+70%) and corn (+80%) further led to widespread, although not entirely substantiated, concern on the impacts of biofuel production on hunger and deprivation in poor regions (Timilsina, 2012; Sims et al., 2010; Ivanic and Martin, 2008; Mitchell, 2008).

Effectively, these developments in the biofuel sector have been underlying the search of novel plant feedstocks that do not compete with conventional crops for arable land and are efficient in terms of land-use, fossil fuel input and fertilizer inputs. These 'second generation' feedstocks include wastes from agriculture and forestry, such as rice and wheat straw, and wood chips (Sarkar et al., 2012). These wastes predominantly comprise of lignocellulose and can therefore be used to generate carbohydrates for bio-ethanol production and the production of bio-chemicals. Vegetable oil wastes in the form of used cooking oil can be recycled to produce biodiesel (Mandolesi de Araújo et al., 2013). Due to its low costs, used cooking oil has become the second largest biodiesel feedstock in the EU (Flach et al., 2015)

Second generation feedstocks further include novel crops that can be cultivated on marginal lands, unsuitable for conventional agriculture. Perennial grasses such as *Miscanthus* (*Miscanthus giganteus* as well as other species of the genus) and switchgrass (*Panicum virgatum*) have been widely studied as lignocellulosic feedstocks for bioethanol production. *Miscanthus* can obtain high yields of 20-44 t dry weight (dw) ha^{-1} $year^{-1}$, whilst requiring minimal fertilization by efficiently reallocating nutrients to its root system during the winter (Lewandowski et al., 2016; Tubeileh et al., 2016; Lewandowski et al., 2000). For switchgrass yields between 10-27 t dw ha^{-1} $year^{-1}$ have been reported, requiring 50 kg N ha^{-1} $year^{-1}$ for optimal growth (Tubeileh et al., 2016; Carriquiry et al., 2011). Additionally non-edible oil crops, such as *Jatropha*, have been suggested to produce biodiesel on marginal lands (Atabani et al., 2012). *Jatropha* growing on semi-arid wasteland yields 0.5-1.7 t dw ha^{-1} $year^{-1}$ of dry seeds, which generally contain 50% of oil (Wani et al., 2016; Yong et al., 2010). Higher yields up to 5 t dw ha^{-1} $year^{-1}$ have been reported, but rely on active irrigation and fertilization, which increases its competition with food crops (Wani et al., 2016; Yong et al., 2010).

For lipid and biodiesel production, microalgae have been the most researched-

novel feedstock. Being aquatic organisms, algae can be produced independently of arable land. Cultivation of marine algae can further be expanded to the world's oceans. For both freshwater and marine microalgae productivities of over 100 t dw ha^{-1} $year^{-1}$ have been reported under ideal growth conditions (Mata et al., 2010). Most common algae have lipid contents between 20-50%, although lipid content is generally inversely related to productivity (Mata et al., 2010; Yang et al., 2011). Still, oil yields per hectare are estimated at 10 to 20 fold that of conventional oil-crops (Mata et al., 2014). Due to the combination of high productivity and land-independent production algae are often referred to as a third generation feedstock (Yang et al., 2011). Microalgae do require substantial amounts of nutrients to grow, which contributes to the disruption in global nitrogen and phosphorous cycles when supplied as synthetic fertilizer. However, by cultivating algae with nutrient-rich runoff from conventional agriculture or wastewater treatment facilities, a positive effect on both nutrient cycles can be obtained (Woertz et al., 2009). Macroalgae, or seaweeds, can also reach yields above 100 t ha^{-1} $year^{-1}$ (Gao and McKinley, 1994). Lipid contents of macroalgae are generally in the order of 1-6% of the biomass, making seaweeds less attractive for biodiesel production than microalgae (Suutari et al., 2015; Chen et al., 2015). However, seaweeds contain 20-30% carbohydrates that may serve as feedstock for bioethanol or the production of various bio-chemicals via carbohydrate conversions (Suutari et al., 2015; Chen et al., 2015; Bikker et al., 2016).

Although the search for alternative crops has been mainly driven by the biofuel sector, many of the crops were later investigated for food or feed production. Rapeseed was traditionally used for the production of fuel oil and commodities such as soaps and candles and only later used for feed and food when (canola) varieties were bred with reduced amounts of anti-nutritional erucic acid and glucosinolate (Loganes et al., 2016). Similarly researchers are now breeding *Jatropha* to produce edible oilseeds (King et al., 2009). Some seaweeds are interesting alternatives for conventional protein crops, due to their high protein content, i.e. 26.6% 18.9% and 22.6% in *Porphyra*, *Undaria* and *Ulva*, respectively, as well as a high content of essential amino acids (Bikker et al., 2016; Dawczynski et al., 2007). A bio-refinery concept, i.e. the generation of multiple value-adding products from a biological feedstock, may be used to produce a combination of food/feed, fuels and chemicals from these feedstocks (Bikker et al., 2016; Bruins and Sanders, 2012)

Free-floating aquatic plants

Although limited by freshwater availability, free-floating aquatic plants share many of the beneficial properties attributed to algae, including land-independent production, rapid growth and potential for nutrient removal from waste streams.

Free floating aquatic plants are land plants that adapted to an aquatic environ-

ment with a very high nutrient availability. By assimilating nutrients directly from the water, they do not require to be rooted in the soil. Their position on top of the water, allows them to outcompete submerged macrophytes and phytoplankton for light (McCann, 2016; van Gerven et al., 2015; Janse, 1998). This specific adaptation has occurred multiple times during evolution since free floating aquatic plants can be found among several land plant lineages (Chambers et al., 2008). *Azolla* and *Salvinia* are aquatic ferns, believed to have evolved 90 million years ago (Metzgar et al., 2007). Duckweeds are small monocotelydon seed plants belonging to the Lemneaceae family, further subdivided into *Spirodela*, *Lemna*, *Wolffiella* and *Wolffia* species (Les et al., 2002; Tippery et al., 2015). Water lettuce, or *Pistia stratioles*, is closely related to the *Lemna* species as both belong to the order of Araceae (Les et al., 2002). Water hyacinth, or *Eichhornia crassipes*, is also a monocotelydon seed plant, but part of the order of Commelinales (Ness et al., 2011).

Free-floating aquatic plants are renowned for their rapid exponential growth. In general, *Azolla* and *Salvinia* fern species were reported to double their biomass in 2 to 4 days, although for *Azolla* species doubling times lower than 2 days have been reported (Cary and Weerts, 1992; Sah et al., 1989; Cheng et al., 2010; Jampeetong and Brix, 2009; Maejima et al., 2001). Doubling times below 2 days, have been repeatedly observed for duckweeds, including *Lemna gibba* (1.4-1.7 days), *Wolffiella* hyaline (1.3 days) and *Lemna minor* (1.6 days) (Ziegler et al., 2015; Gale et al., 1989; Yilmaz, 2007). Productivity per unit of surface area was reported to range between 30-50 t dw ha^{-1} $year^{-1}$ for *Azolla*, *Salvinia, Lemna* and *Pistia* (Becerra et al., 1990; Henry-Silva et al., 2008), while *Eichhornia crassipes* stands out with productivities of over 80 t dw ha^{-1} $year^{-1}$ being reported (Tucker and Debusk, 1981; Reddy and D'Angelo, 1990).

Several free floating aquatic plants are invasive weeds because of their rapid colonization of water surfaces (Pieterse and Murphy, 1993). Dense floating mats can lead to decreased oxygen exchange, while enhancing oxygen consumption by microbial decomposition of the dead sinking biomass in the dark water body underneath, leading to anoxic waters and associated loss of aerobic life (Janse, 1998). Also the physical obstruction of waterways can negatively impact local communities (Opande et al., 2004). The rapid spread of species originating from South America upon introduction in Africa and Australia is a testimony of their invasive character, taking *Salvinia molesta* and *Eichhornia crassipes* as prime examples (Coetzee and Hill, 2008; Kriticos and Brunel, 2016; Sands and Kassulke, 1986; van Thielen et al., 1994). As a result, several free floating plant species are listed as noxious invasive weeds in Europe, China, US and South America (Gassmann et al., 2006; Wang et al., 2016; Lozano and Brundu, 2016; Truernit et al., 2008)

However, the rapid growth and efficient uptake of nutrients from eutrophic waterways also make aquatic plants ideal candidates to mitigate human effects on global

nitrogen and phosphorous cycles. Free floating aquatic plants can recover nutrients close to emission sources, unlike marine algae, while generating large amounts of biomass per unit of surface area. Various studies indicated successful removal of nitrogen and phosphorous from manure waste, domestic and industrial wastewaters, agricultural runoff and eutrophic natural freshwater ecosystems (Muradov et al., 2014; Zhao et al., 2014; Olguín et al., 2003; Ozengin and Elmaci, 2007; Wang et al., 2012). However, to exploit the full potential of floating aquatic plants safely, control over their invasive traits will need to be gained.

Azolla, a special free-floating aquatic plant

The aquatic fern *Azolla* is a special case among free-floating aquatic plants because it has evolved a symbiosis with the cyanobacterium *Nostoc azollae*, which fixes atmospheric nitrogen (N_2). The N_2-fixing cyanobacteria are located in a special pocket in the upper leaf lobe of *Azolla*, where they develop heterocysts at high frequency over vegetative cells (Figure 3) (Becking, 1987). The nitrogenase enzyme complex, which is inactivated by O_2, is well enough protected in heterocysts from O_2 produced by photosynthesis of the fern leaves. Gas exchange is ensured via a pore in the leaf pocket, visible at the back side of the upper leaf (Becking, 1987). Fixed N_2 is likely provided to the plant in the form of ammonia (Peters and Meeks, 1989; Meeks et al., 1987) and research suggests that in return the plant provides sugars to the cyanobacteria (Rozen et al., 1986; Ran et al., 2010). Although such an exchange between microbes and plants is common, the *Azolla-Nostoc* symbiosis is special due to the high interdependency between symbiont and host. Cyanobacteria are vertically transmitted to new generations by nesting into the female sporocarps, which are the disseminated reproductive organs of *Azolla* ferns (Figure 3) (Becking, 1987; Ran et al., 2010; Zheng et al., 2009). The cyanobacteria within the female reproductive organs differentiate into akinete resting forms (Figure 3) (Ran et al., 2010; Zheng et al., 2009). Upon germination, the cyanobacteria akinetes develop into a colony of vegetative cells that associates with the fern shoot meristem (Becking, 1987). As the plant grows, cyanobacteria are dispensed in the leaves where they start to form heterocyst cells and perform N_2 fixation (Becking, 1987). Long-lasting co-evolution has led the genome of the cyanobacteria to degrade and lose functions that are crucial for free living *Nostoc* relatives, such as an operational glycolysis and phosphate storage pathways (Ran et al., 2010). *Nostoc azollae* therefore likely lost its ability to live outside of the plant.

This unique symbiosis allows *Azolla* to thrive in both N-depleted waters as well as N-rich waters, broadening its habitat compared to other free-floating plants. *Azolla* ferns occur across many continents and across various climates. Currently seven species are recognized: *A. filiculoides*, *A. rubra*, *A. caroliniana*, *A. microphylla*,

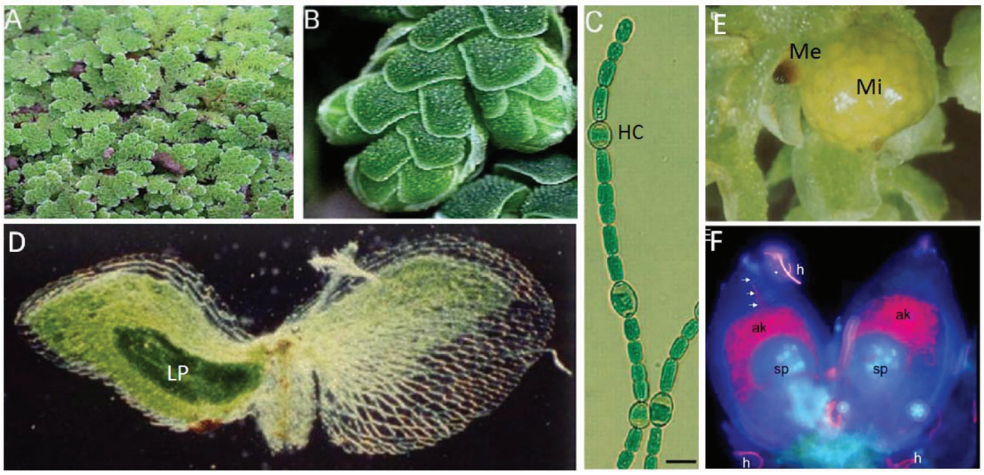


Figure 3. Characteristic features of *Azolla filiculoides*. (A) Sporpohyte photographed from the top (Ran et al., 2010). (B) Close-up of sporophyte (Ran et al., 2010). (C) Vegetative filaments of *Nostoc azollae* with bigger heterocysts (HC) cells (Ran et al., 2010). (D) Section of an *Azolla* leaf showing the upper leaf lobe containing the leaf pocket (LP) and the supporting lower leaf lobe (Lumpkin and Plucknett, 1980). (E) Megasporocarp (Me) and microsporocarp (Mi) containing the female and male spores, located on the bottom of the sporophyte (Becking, 1987). (F) *Nostoc azollae* akinete (ak) resting located in the tip of the megasporocarp, above the megaspore (sp), documented by fluorescence microscopy (Ran et al., 2010).

A. mexicana, *A. pinnata* and *A. nilotica* (Metzgar et al., 2007; Pereira et al., 2011). *A. filiculoides* is widely distributed, occurring in temperate climates in North America and Europe as well as (sub)tropical climates in South America, Africa and Asia (Watanabe et al., 1992). *A. caroliniana*, *A. microphylla* and *A. mexicana* predominantly occur in South America and southern Northern America, although *A. caroliniana* has also been observed in temperate climates in USA and the Netherlands. (Watanabe et al., 1992). *A. pinnata* is distributed in tropical regions in Asia, Africa and Australia (Watanabe et al., 1992). *A. rubra*, is endemic to Australia and New Zeeland and *N. nilotica* is confined to mid-eastern Africa from Sudan to Tanzania (Watanabe et al., 1992; Wagner, 1997).

Fossil records of well-preserved spores indicate that *Azolla* species have been widespread on planet Earth for millions of years. The oldest fossils of *Azolla* ancestors date back to the Campanian and Maastrichtian (66-84 million years ago) and have been found in current Argentina, Bolivia, USA and India (Vajda and McLoughlin, 2005; Cúneo et al., 2014; Hall, 1969; Hall and Swanson, 1968; Nambudiri and Chitaley, 1991). Younger fossils dating from the middle Eocene (49 million years ago) have been found in massive amounts in sediments of the Arctic Ocean and surrounding Nordic seas (Brinkhuis et al., 2006; Barke et al., 2012). At that time the Arctic was enclosed by land masses forming a gigantic lake, with a top layer of

fresh water to support *Azolla* growth (Brinkhuis et al., 2006). The high abundance of *Azolla* fossils and their correlation with high levels of total organic carbon in the sediments, suggest *Azolla* blooms episodically covered this gigantic artic lake during a period of 0.8-1.2 million years (Brinkhuis et al., 2006; Speelman et al., 2009a). This period is referred to as the *Azolla* event (Brinkhuis et al., 2006). Interestingly, the timing of this period correlates with a reduction in global atmospheric CO_2 that initiated a cooldown of the Earth to its current climate (Speelman et al., 2009a). It was estimated that the *Azolla* event would have contributed a 55- 470 ppm reduction of global CO_2 concentration (Speelman et al., 2009a). In laboratory studies, yields of exponentially growing *Azolla* responded positively to elevated CO_2 concentrations as biomass yields increased 32-79% when CO_2 concentrations were increased from ambient levels to 680-1000 ppm (Cheng et al., 2010; Speelman et al., 2009a; van Kempen et al., 2016). Increasing CO_2 concentrations further to 1600-1920 ppm, resulted in a 200-250% yield increase over ambient CO_2 (Speelman et al., 2009a; van Kempen et al., 2016). *Azolla* may therefore be very effective at capturing CO_2 from concentrated sources.

Its unique N_2-fixing capability further makes *Azolla* particularly suitable for the recovery of excessive phosphorous from wastewater, since nitrogen is often limiting the efficiency of biological removal of phosphorous from waste streams. In a study on diluted swine manure, *Azolla* removed 100% of the ammonia and 100% of the phosphate, whereas *Lemna* could remove all ammonium, but only 60.1% of the phosphate (Muradov et al., 2014). By efficiently capturing excess CO_2 and nutrients, *Azolla* cultivation may contribute to prevent further disruption of global carbon, nitrogen and phosphorous cycles and associated negative impacts, depending on how the produced biomass is utilized.

Azolla biomass has been applied as a bio-fertilizer in Rice fields in south-east Asia for over multiple centuries (Wagner, 1997). *Azolla* plants are generally intercropped with Rice and after an initial period of growth plants are incorporated into the soil and upon their decomposition release fixed nitrogen to the rice plants (Wagner, 1997). Using *Azolla* biomass as a bio fertilizer to replace artificial nitrogen fertilizer reduces pollution by nitrogen losses of rice cultivation and provides added value to farmers (Mishra and Dash, 2014; Mian and Stewart, 1985; Watanabe et al., 1989). Additionally, recent investigations have indicated that dual cropping of rice with *Azolla* reduces CH_4 emissions of paddy's (Liu et al., 2017), although this may strongly depending on local conditions as increases of CH_4 emissions have also been reported for waterways dominated by free-floating plants (Kosten et al., 2016). Although the application of *Azolla* as biofertilizer has clear advantages, direct use of *Azolla* biomass as feedstock for food, energy and materials may result in greater environmental impact as well as higher added value to farmers.

Azolla species are reported to contain between 16.4% and 28.9% crude protein

(Becerra et al., 1990; Costa et al., 1999; Datta, 2011; Basak et al., 2002; Alalade and Iyayi, 2006; Fasakin and Balogun, 2001). Several researchers investigated the use of *Azolla* as an alternative protein feed supplement for various animals, including pigs, chicken, ducks, fish and sheep (Becerra et al., 1990; Datta, 2011; Basak et al., 2002; Alalade and Iyayi, 2006; Cherryl et al., 2014; Leterme et al., 2010; Naghshi et al., 2014; Abdel-Tawwab, 2008; Ahmed et al., 2016; Acharya et al., 2015). The rate at which *Azolla* could be included into the diet of these animals without causing adverse effects on their body weight, or digestibility, ranged from 6% in sheep to 25% in fish (Table 1). What limits the inclusion rate of *Azolla* has been speculated upon, i.e. suboptimal amino acid composition, high fiber content and high lignin content have been suggested as possible causes, but not elucidated further as of yet.

Table 1. Summary of feeding trials conducted with *Azolla* biomass.

Azolla species	Fresh / dried	Animal	Inclusion rate[1]	Source
A. pinnata	Sun-dried	Chicken	10%	Alalade & Iyayi (2006)
A. pinnata	Sun-dried	Chicken	5%	Basak et al. (2002)
A. pinnata	Dried	Chicken	5%	Naghshi et al. (2014)
A. pinnata / A. microphylla (1:1)	Dried	Fish, Rohu	25%	Datta et al. (2011)
A. pinnata	Sun-dried	Fish, Tilapia	25%	Abdel-Tawwab (2008)
A. pinnata	Fresh	Perkin Ducks	10%	Acharya et al. (2015)
A. pinnata	Sun-dried	Pigs	20%	Cherryl et al. (2014)
A. filiculoides	Dried	Pigs	15%	Leterme et al. (2010)
A. filiculoides	Drained overnight	Pigs	11%	Becerra et al. (1990)
A. (cristiata)	Sun-dried	Sheep	6%	Ahmed et al. (2016)

[1]Inclusion rate refers to the percentage of *Azolla* in the diet (on a dry weight basis) that could be fed to the animals without negatively affecting weight gain or digestibility with respect to a reference diet.

Apart from feed, *Azolla* biomass can serve as a feedstock for the production of bio-energy and biomaterials. Anaerobic fermentation of whole *Azolla* yielded 188.5 L of biogas per kg dry biomass (Jain et al., 1992). Depending on the growth season, *A. caroliniana* biomass contained 12.7-16.4% of lipids which may be used for the production of biodiesel or oleochemicals (Paoletti et al., 1987). Special wax lipids with mid-chain hydroxyl groups have been reported in *Azolla* in the context of biomarkers in geochemical studies (Speelman et al., 2009b), but may also be of interest for the chemical industry. *A. caroliniana* was reported not to contain lignin, but tannins instead (Nierop et al., 2011). Depending on their composition and concentration, tannins as well as other soluble (poly)phenols identified in *Azolla*, could be extracted

and purified to replace oil-derived aromatic compounds used by various industries (Teixeira et al., 2001; Ishikura, 1982). Rough chemical analyses have further indicated that *Azolla* biomass contains 16.0% of cellulose and 14.3% hemicellulose, which may be used for the production of bioethanol fuel or platform chemicals (Costa et al., 1999; Alalade and Iyayi, 2006; Ahmed et al., 2016).

The combination of land-independent production, high growth rate, N_2-fixing capability and valuable chemical composition makes *Azolla* a potential novel crop for the production of food, fuel and/or materials. However, to utilize this potential in a production system several challenges need to be addressed. These challenges relate to three crucial elements of a future production chain: generation of starting material, cultivation and processing. Challenges in each part of the production chain need to be addressed in parallel as the strength of the entire production chain will depend on the weakest link. Hence the aim of this thesis is to research these major technological challenges and fill knowledge gaps to enable development of an *Azolla* production chain.

Challenges for developing an *Azolla* production chain

Providing reliable starting material

At the basis of the *Azolla* production chain is the starting material. For a plant based production chain to function, sufficient starting material needs to be provided to the cultivation system. Firstly to commence cultivation and secondly to restart cultivation in case the cultivation system fails intermittently. Furthermore, the starting material needs to be reliable in generating reproducible yields and respond reproducibly to the biotic and abiotic environment. A way to physically store varieties suited for production needs is therefore required. Currently, vegetative propagation is used to maintain *Azolla* accessions (Wagner, 1997; Watanabe et al., 1992). Maintaining accessions by vegetative reproduction is both laborious and prone to mixing errors, and therefore an unreliable way to store varieties suited for production needs or wild relatives to conserve biodiversity. Furthermore, vegetative reproduction does not allow for breeding and selection of superior varieties. The latter has been paramount in the yield improvement of major crops. For major crops, such as wheat, the domestication process started 10,000 years ago when humans developed methods to control the sexual reproduction of seed plants and selected varieties with favorable traits (Zeder, 2015; Wang et al., 1999; Diamond, 2002; Allard, 1999). In the last decades breeding has been accelerated by the development of new techniques, starting with mutagenesis to rapidly expand the pool of genetic variation to select

from, followed by targeted approaches involving gene insertion and recently gene modification by CRISPR/CAS (Hilscher et al., 2017; Schiml et al., 2014; Oladosu et al., 2016). The development of new breeding techniques runs in parallel with a rapid development of next generation sequencing techniques, that allows to target breeding approaches (Mir et al., 2012; Scheben et al., 2016; Varshney et al., 2013; Jacob et al., 2016; Salgotra et al., 2014).

In the case of *Azolla*, as well as most ferns, the sexual life cycle is not yet fully understood and methods for controlling progression through the life cycle are still lacking. Ferns produce spores instead of seeds, which may lack key characteristics that facilitated domestication of seed plants, such as ease of collection and long-term storability of reproductive organs (Li and Olsen, 2016; Purugganan and Fuller, 2009). Controlling the reproduction of ferns requires a reliably timed reproduction, or sporulation, and the capability to harvest, store, disseminate and germinate spores. Hence, to provide reliable starting material for *Azolla* cultivation and further improve this starting material by domesticating *Azolla*, we need to develop methods to control its sexual reproduction. In parallel, a better understanding of fern molecular biology will be needed, firstly to help develop these methods, and secondly to accelerate the subsequent breeding process.

Growth and nitrogen fixation in continuous production systems

The cultivation of *Azolla* requires a different approach than conventional land-based cropping systems. Conventional cropping systems operate as a batch process, where plants transgress through an exponential growth phase, followed by a linear growth phase and are harvested in the last asymptotic growth phase. In the exponential growth phase plants are not competing for light since the standing crop per unit of surface area is low. As plants grow bigger, competition for light and/or nutrients kicks in and growth becomes linear. Although in the exponential growth phase, the relative growth rate (RGR) is highest, in the linear growth phase the absolute growth rate (AGR) per surface area is at its maximum. *Azolla* is well studied for its phenomenal RGRs during exponential growth (Cary and Weerts, 1992; Sah et al., 1989; Cheng et al., 2010; Maejima et al., 2001; Speelman et al., 2009a; van Kempen et al., 2016; van Kempen et al., 2013; Arora and Singh, 2003). However, to maximize yields per unit of surface area, continuous harvesting and nutrient supply is needed to keep plants in the linear growth phase. Becerra et al. (1990) are among the few that analyzed *Azolla* growth in a continuous production system, reaching a productivity of 39 t dw biomass ha^{-1} $year^{-1}$ in an open pond system in Colombia (Becerra et al., 1990). Methodologies on how to set up such a continuous production system are lacking, but are required to advance *Azolla* cultivation to a next level.

Similarly, effects of nitrogen fertilizer and CO_2 addition have been studied in

the context of batch-wise exponential growth, but not in the context of continuous production systems (Cary and Weerts, 1992; Cheng et al., 2010; Speelman et al., 2009a; van Kempen et al., 2016). The effects of nitrogen fertilizer and CO_2 addition may be very different in a continuous production system wherein plants are competing for light (energy limitation for CO_2 fixation at high CO_2 concentrations) and nutrients. Batch experiments have indicated that N_2 fixation in *Azolla* is providing all the nitrogen required during exponential growth, i.e. exogenous nitrogen fertilizer do not greatly improve RGRs. Whether N_2 fixation by *N. azollae* also sustains high growth rates in a continuous production system awaits experimental confirmation.

From chemical composition to potential products

The specific chemical composition of *Azolla* biomass will determine which products may be produced from *Azolla* biomass and the overall feasibility of an *Azolla* production chain versus current production systems and other alternatives. Thus far only rough chemical compositions of *Azolla* have been reported. Crude protein contents frequently reported for *Azolla* are based on multiplying *Azolla*'s nitrogen content with a standard conversion factor of 6.25, but a complete analysis of the amino acid composition is lacking. Lipids have been investigated by Paoletti et al. (1987), but quantitative contributions of other lipid groups, such as sterols or special functionalized wax lipids have not been investigated (Paoletti et al., 1987; Speelman et al., 2009b). Concentrations of cellulose, hemicellulose and lignin reported for *Azolla* are based on gravimetrical differences between neutral detergent fiber and acid detergent fiber and sulfuric acid hydrolysis (Costa et al., 1999; Alalade and Iyayi, 2006; Ahmed et al., 2016). Lignin contents of *Azolla* based on these methods range from 5.7-41.0% of the dry weight, which would be even higher than lignin-rich tissues such as wood, whilst analysis by pyrolysis coupled to gas chromatography and mass spectrometry (GC-MS) revealed no true lignin in *A. caroliniana* (Nierop et al., 2011). Instead *A. caroliniana* contained tannins, but its quantitative contribution to *Azolla* biomass remains to be determined. Several other soluble polyphenolic compounds have been independently detected in *Azolla* biomass, but a comprehensive overview is lacking (Teixeira et al., 2001; Ishikura, 1982). Therefore, more precise methods may be employed to further elucidate the chemical composition of *Azolla*. With a better insight into the exact chemical composition we can start to develop processes to produce sustainable and value-adding products from *Azolla* biomass.

Outline of this thesis

Chapter 2 aims to provide more insight into the reproductive biology of *A. filiculoides*. Methods to harvest, store, fertilize and germinate *A. filiculoides* spores are developed. By comparing profiles of RNA sequences obtained from sporophytes and sporocarps, molecular pathways are identified that are similar to those controlling the transition to sexual reproduction in seed plants.

In Chapter 3 focusses on the N_2-fixation in *Azolla*. It is investigated to what extent N_2-fixation by the cyanobacteria can sustain the high productivity of *A. filiculoides* and how the plant-cyanobacteria symbiosis responds to nitrogen fertilizer. The diurnal nitrogen supply by *N. azollae* is evaluated and RNA sequencing is used to gain insight into the molecular basis of metabolic and structural adaptations of the fern to facilitate N_2-fixation by *N. azollae*.

In Chapter 4 a methodology to grow *Azolla* continuously at high productivity is developed This methodology is then used to compare the growth potential of the temperate species *A. filiculoides* with the tropical species *A. pinnata*, and determine the effect of CO_2 concentration on their growth. Growth analysis is combined with analysis of chemical composition to obtain more realistic insight into the potential applications of *Azolla* as a novel crop. In particular, the amino acid composition of *Azolla* is investigated in the context of using *Azolla* as a protein feed to replace soybean meal. Additionally, first insights into lipid and polyphenol content of rapid growing *Azolla* are presented.

Chapter 5 elaborates on the lipid composition of *A. filiculoides* grown under the highly productive conditions introduced in Chapter 4. The fatty acid content of *A. filiculoides* is investigated to determine the potential of *Azolla* lipids to serve as feedstock for the production of biodiesel. Functionalized wax lipids, previously reported as *Azolla* biomarkers, are investigated in the context of their potential as oleochemicals.

Chapter 6 elaborates on the polyphenol composition of *Azolla*. In Chapter 4, (poly)phenols were found to be remarkably abundant in *Azolla* species, which may affect biomass digestibility depending on the type of (poly)phenols present. Hence, this chapter provides a comprehensive overview of soluble and insoluble phenolic compounds in *Azolla*. Genome sequence annotation of *A. filiculoides* is combined with RNA sequencing data from Chapters 2 and 3, to identify biosynthetic pathways leading to the production of (poly)phenolic compounds.

In Chapter 7 extraction methods are developed to produce protein concentrates from *Azolla*. Alkaline protein extraction followed by acid precipitation of protein serves as a starting point due to its low complexity and compatibility with high water content feedstocks; preferred for local on-farm processing. Findings from Chapter 4 and Chapter 6 strongly point to (poly)phenols as limiting factors for the digestibility of *Azolla* feed. Therefore extraction methods that maximize protein yields and minimize polyphenol content are developed.

Lastly, I discuss limitations and implications of the work presented in this thesis. I evaluate how the content of this thesis may contribute to overcoming challenges in the development of an *Azolla* production chain and which key challenges remain to be tackled in future *Azolla* research.

Chapter 2

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, Institute of Environmental Biology, Utrecht University, the Netherlands
[2] Institute for Plant Biochemistry, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University Düsseldorf, Germany;
[3] Department of Earth Sciences, Faculty of Geosciences, Utrecht University, the Netherlands
[4] Agrosystems research, Wageningen Science & Research, the Netherlands

Adapted from Brouwer, P., Bräutigam, A., Külahoglu, C., Tazelaar, A.O.E., Kurz, S., Nierop, K.G.J., van der Werf, A., Weber, A.P.M., Schluepmann, H. (2014). Azolla domestication towards a biobased economy? New Phytologist, 202 (3), pp. 1069-1082.

Abstract

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 bio-chemicals 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 and 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 and 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 and Donze, 1981; Becking, 1987; Peters and Meeks, 1989; Wagner, 1997; Zheng et al., 2009; Carrapico, 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 sexu-

al 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 sporo*carps

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 x10 or x5 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 x16 or x40 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 spore*s 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 for documenting megaspore germination section.

Table 1. Cryopreservation protocols tested on *Azolla filiculoides* spores. (a) Cryopreservation pre-treatment definitions1 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	2M DMSO, 3.2M Glycerol	Tree fern, *Dicksonia Sellowiana*, Rogge et al. (2000)
DMSO + EG + PVP	2M DMSO, 4.8M EG, 2% PVP	Zebrafish embryo, Riesco et al. (2012)
Sucrose	0.4M Sucrose	Fern, *Pteris adscensionis*, Barnicoat et al. (2011)
Trehalose	0.4M Trehalose	Fern, *Pteris adscensionis*, Barnicoat et al. (2011)
DMSO + EG + Glucose + Trehalose	2M DMSO, 2.4M EG, 3.2M Glucose, 0.4M 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 megasp- orocarps tested

In order to optimize cryopreservation involving drying as pre-treatment, 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, 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 Fast-QC 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, Usadel et al., 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 (Af-APTF: TAGAGATGCATGTGGGTGCAGT; AfAPTR: AAAAGCGGTTTACCACCCAGTT) (Salmi and 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: TTTGCATCTTTCGCTCTCTCA; AfCALR: CCAAGCTGCACAATGTAAGGA). 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 343740 (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 (Figure 1a) in pairs, first as microsporocarp pairs, then as the sporophytes reached maturity as micro and megasporocarp pairs (Figure 1b). The megasporocarp (Figure 1c), containing the female megaspore and cyanobacteria akinetes, quickly sunk. Detached microsporocarps generally burst open, releasing the microsporangia (Figure 1d); once these burst, massulae were released and entangled in filamentous appendages of the epispore wall of the megaspore (Figure 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 (Figure 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

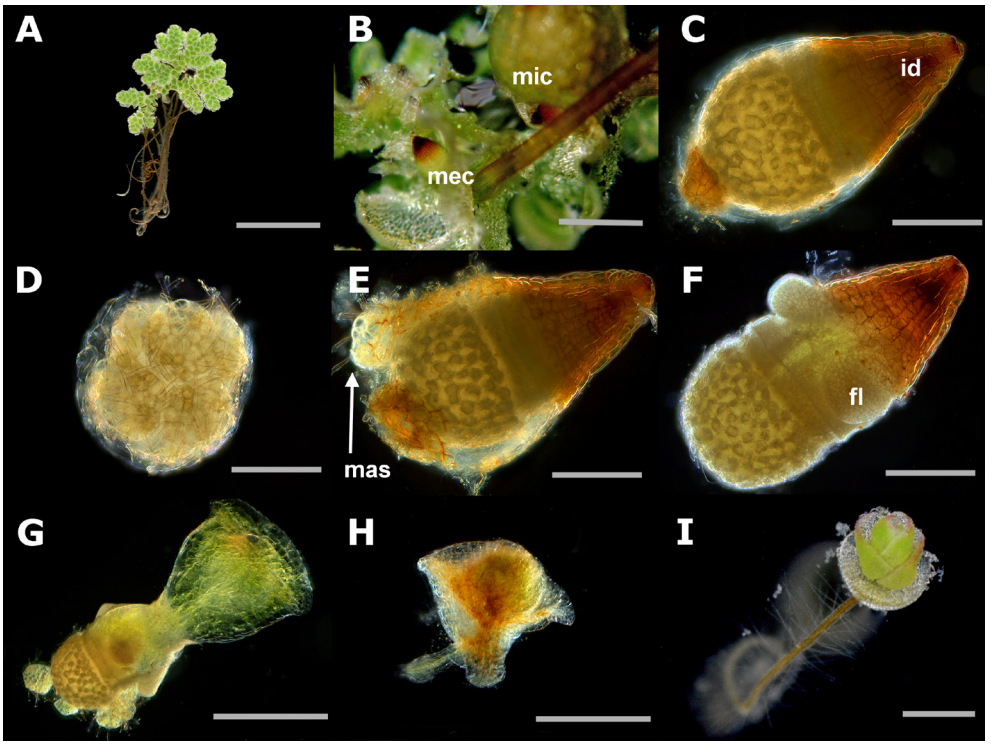


Figure 1. Key stages in the sexual reproduction of Azolla filiculoides. (a) Sporophyte; bar, 10 mm. (b) Megasporocarps (mec) and microsporocarp (mic) at the underside of a sporulating plant; bar, 1 mm. (c) Detached megasporocarp with indusium cap (id); bar, 0.2 mm. (d) Microsporangium containing four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

viable megaspores developed further into sporelings with a cotyledon (Figure 1g). Root development was visible shortly after (Figure 1h). Sporelings became sufficiently buoyant to float at the surface when first leaves appeared (Figure 1i). Germination rates up to 30% were achieved after in vitro fertilization, of which 95% occurred within 26 d (Figure 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 m*egasprores 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% (Figure 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.

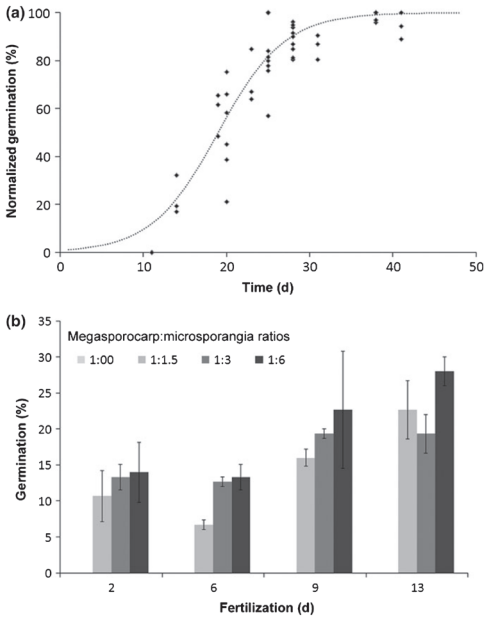


Figure 2. *Azolla filiculoides* germination and in vitro fertilization. (a) Germination frequencies over a 5-wk period: germination was scored when cotyledons had emerged from megasporocarps, data from several experiments, frequencies were normalized by dividing the number of germinated megasporocarps at given time points by the number of germinated megasporocarps after 40 d for each experiment. The trend-line is the best-fit sigmoid function with b = 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.

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. 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 (Figure 3a) and exhibited heterocysts at equal frequency to untreated controls (Figure 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

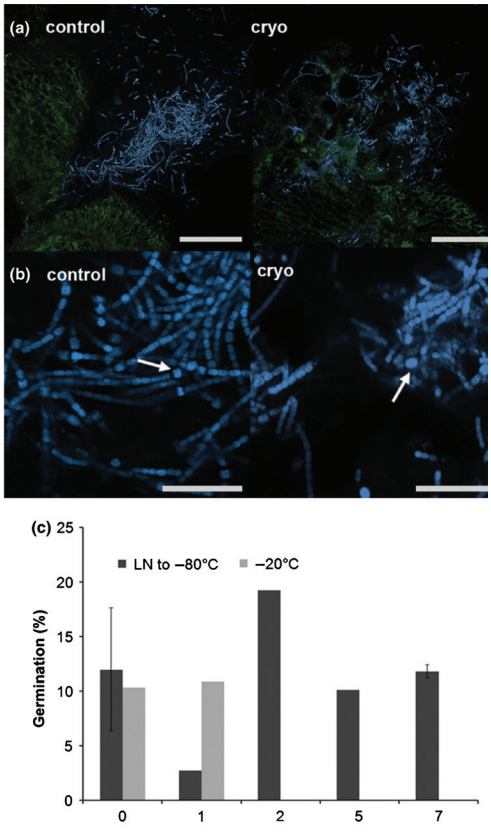


Figure 3. Survival of the symbiont *Nostoc azollae* and long-term cryopreservation of the *Azolla* symbiosis. (a) *N. azollae* filaments in 7 wk *Azolla* fronds developing from a fresh megasporocarp (control) and a megasporocarp cryopreserved for 2 months at -80°C (cryo); bar,0.15 mm. Confocal fluorescence 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.

and stored at -80°C (or frozen and stored at -20°C) for up to 7 months (Figure 3c). 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 and 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.

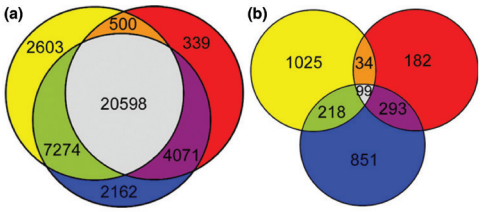


Figure 4. Distribution of unigenes over megasporocarps, microsporocarps and sporophytes. mRNA was extracted from megasporocarps, microsporocarps and nonsporulating sporophytes, sequenced, then sequences assembled in a unigene database as described in the Materials and Methods section. 77%, 76% and 74% 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.

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 (Figure 4a). mRNA extracted from the differing tissues were thus reproducibly from the fern *A. filiculoides*.

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 (Figure 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

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)	micro (1144)	sporo. (1025)[1]	mega (1%)	micro (3%)[2]	sporo. (3%)[2]	mega	micro	sporo
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.

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 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* (Figure 5a). The *Azolla* FT-like protein contains the motifs D-P-D-x-P-S-P-S (Figure 5a, Box I) and G-x-H-R (Figure 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 (Figure 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 (Figure 5b) along with *SOC1* and *LFY* (Figure 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* (Figure 6). Nonetheless *Azolla* proteins like FT, SPL1 and CAL1 are

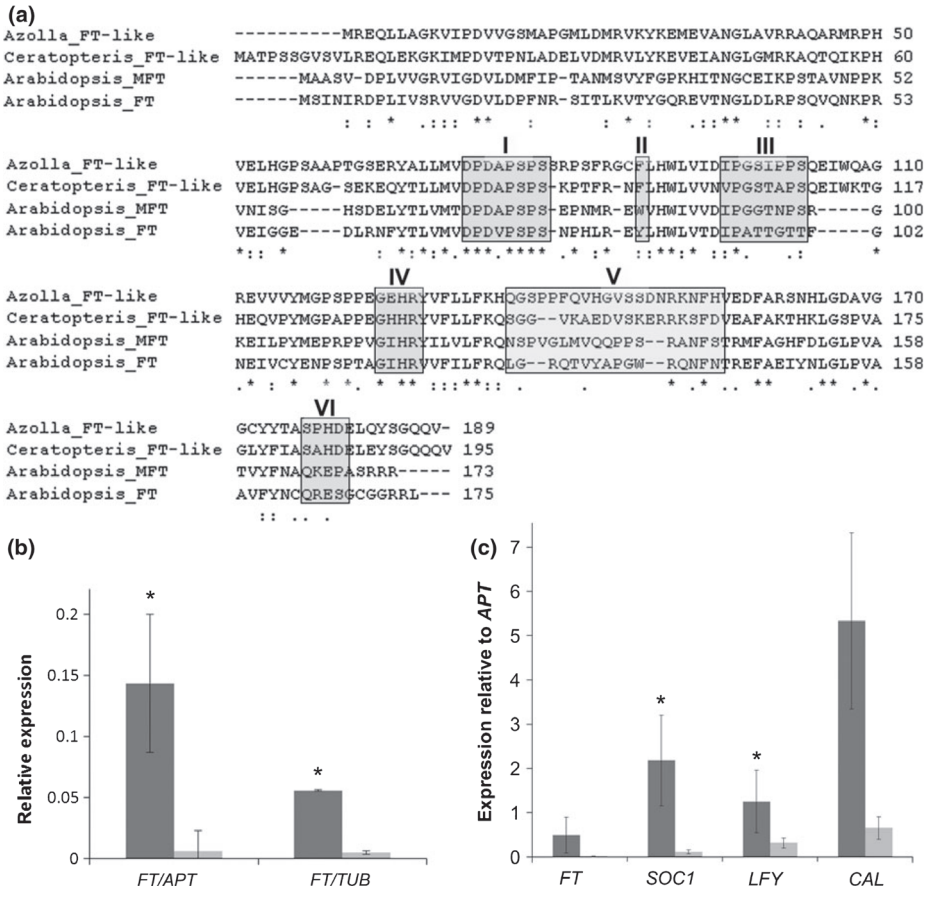


Figure 5. *Azolla* FT-like. (a) Alignment of fern FT-like proteins with FT and MFT from *Arabidopsis*. Read IDs were extracted from the *A. filiculoides* database matching a theoretical DNA encoding the FT from *Ceratopteris* using BLAST then assembled using CAP3. The longest resulting contig covers 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.*, $P < 0.05$.

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 (Figure 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 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 and 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 efficient cyopreservation of varieties of *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 angio-

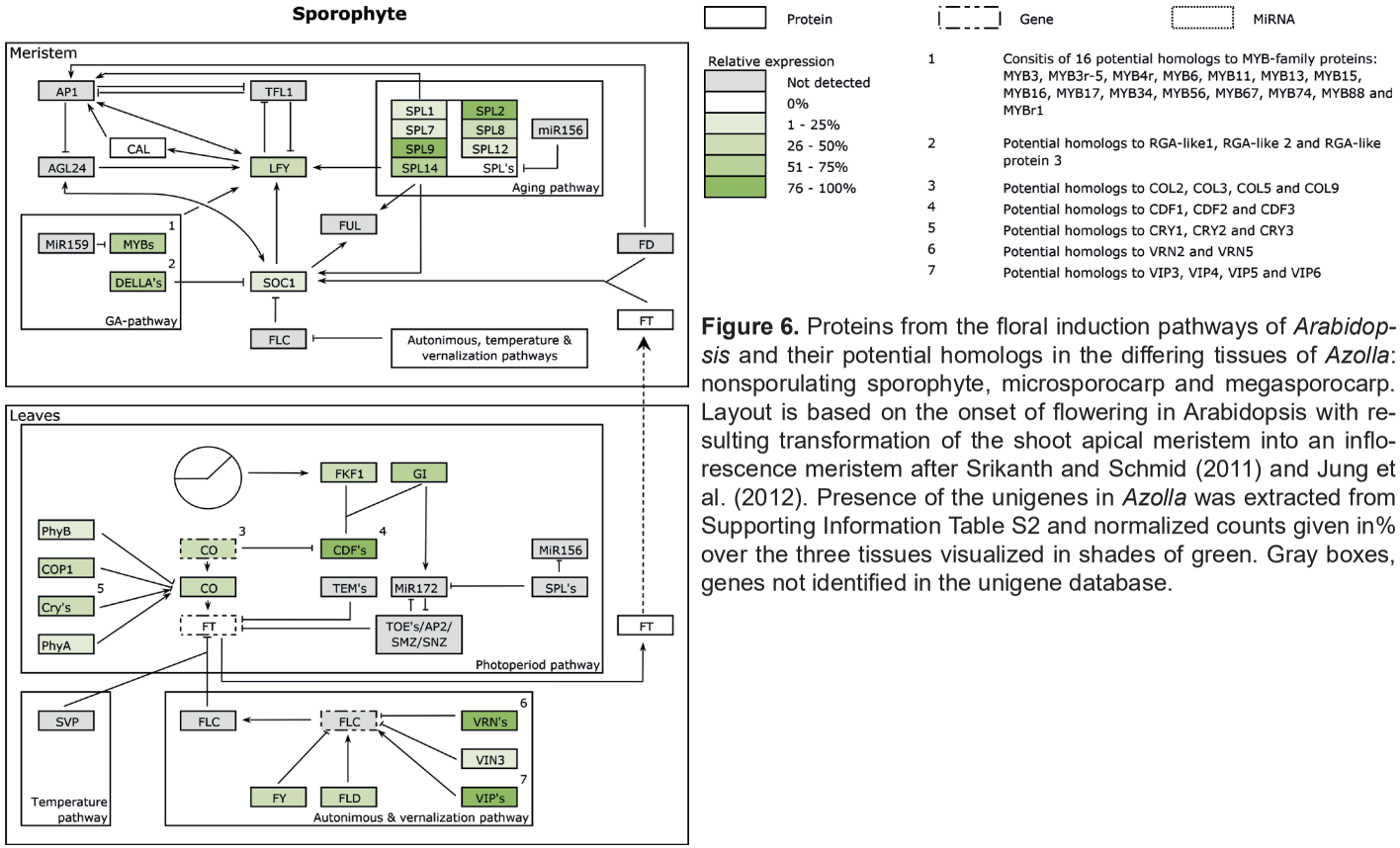
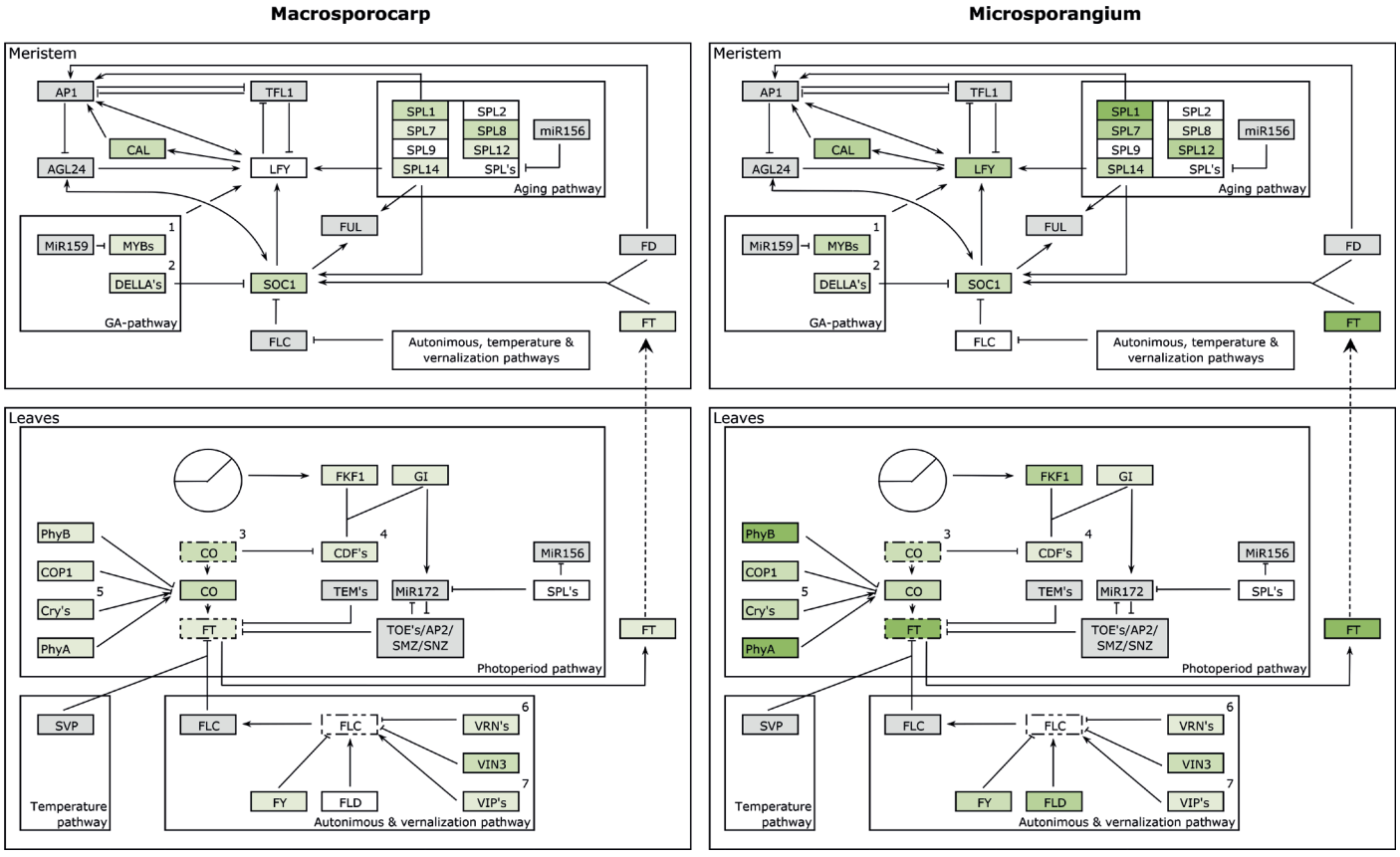


Figure 6. Proteins from the floral induction pathways of *Arabidopsis* and their potential homologs in the differing tissues of *Azolla*: nonsporulating sporophyte, microsporocarp and megasporocarp. Layout is based on the onset of flowering in Arabidopsis with resulting transformation of the shoot apical meristem into an inflorescence meristem after Srikanth and 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.


Macrosporocarp
Microsporangium
Meristem
AP1
TFL1
CAL
AGL24
LFY
SPL1
SPL2
SPL7
SPL8
SPL9
SPL12
SPL14
SPL's
miR156
Aging pathway
MiR159
MYBs
DELLA's
GA-pathway
SOC1
FUL
FD
FT
FLC
Autonimous, temperature & vernalization pathways
Leaves
FKF1
GI
PhyB
COP1
Cry's
PhyA
CO
CDF's
TEM's
MiR172
SPL's
TOE's/AP2/ SMZ/SNZ
Photoperiod pathway
SVP
Temperature pathway
VRN's
VIN3
VIP's
FY
FLD
Autonimous & vernalization pathway

sperms 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 and 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 mapping of sequenced reads onto the corresponding genome 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 and Duckett, 1980). Microsporocarps are still in the process of maturing when attached to the plant and meiotic development has not been completed (Lucas and 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} $year^{-1}$ dry weight yield (Becerra et al., 1990), 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.

Supporting information

Additional supporting information (Tables & Notes) may be found in the online version of this article (open access): http://onlinelibrary.wiley.com/doi/10.1111/nph.12708/abstract

Chapter 3

Profiles of diel molecular responses to N_2 fixation by *Nostoc azollae* reveal that metabolic, structural and vascular cooperation sustain the high productivity of *Azolla* ferns without nitrogen fertilizer

Paul Brouwer[1], Andrea Bräutigam[2], Valerie A. Buijs[1], Anne O. E. Tazelaar[1], Adrie van der Werf[3], Urte Schlüter[2], Gert-Jan Reichart[4], Anthony Bolger[5], Björn Usadel[5], Andreas P. M. Weber[2] and Henriette Schluepmann[1]

[1] Molecular Plant Physiology, Institute of Environmental Biology, Utrecht University, the Netherlands;
[2] Institute for Plant Biochemistry, Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich Heine University Düsseldorf, Germany;
[3] Agrosystems research, Wageningen University & Research, the Netherlands;
[4] Department of Earth Sciences, Faculty of Geosciences, Utrecht University, the Netherlands;
[5] Institute for Botany and Molecular Genetics, BioSC, RWTH Aachen University, Germany.

Adapted from: Brouwer, P., Bräutigam, A., Buijs, V.A., Tazelaar, A.O.E., van der Werf, A., Schlüter, U., Reichart, G-J., Bolger, A., Usadel, B., Weber, A.P.M., Schluepmann, H. (2017) Metabolic adaptation, a specialized leaf organ structure and vascular responses to diurnal N_2 fixation by Nostoc azollae sustain the astonishing productivity of azolla ferns without nitrogen fertilizer. Frontiers in Plant Science, 8, art. no. 442

Abstract

Sustainable agriculture demands reduced input of man-made nitrogen (N) fertilizer, yet N_2 fixation limits the productivity of crops with heterotrophic diazotrophic bacterial symbionts. We investigated floating ferns from the genus *Azolla* that host phototrophic diazotrophic *Nostoc azollae* in leaf pockets and belong to the fastest growing plants. Experimental production reported here verified N-fertilizer independent production of nitrogen-rich biomass that surpassed productivity of canonical crops reaching annual yields per ha of 1200 kg^{-1} N fixed and 35 t dry biomass. Stable ^{15}N isotope labelling showed that N_2 fixation peaked at noon, reaching 0.4 mg N g^{-1} dry weight h^{-1}, indicating that nitrogenase of the heterocystic bacteria thriving in the peculiar leaf pocket organs is entirely protected from O_2 released by photosynthesis. *Azolla* ferns therefore merit consideration as protein crops in spite of the fact that little is known about the fern's physiology to enable domestication. The integration of fern specific transcript profiling and microscopic observations reported here show that the ferns adapted to the phototrophic N_2-fixing symbionts *Nostoc azollae* by 1) adjusting metabolically to nightly absence of N supply using responses ancestral to ferns and seed plants; 2) developing a specialized xylem-rich vasculature surrounding the leaf pocket organ; 3) responding to N-supply by controlling transcripts of genes mediating nutrient transport, allocation and vasculature development. Unlike other non-seed plants, the *Azolla* fern clock is shown to contain both the morning and evening loops; the evening loop is known to control rhythmic gene expression in the vasculature of seed plants and therefore may have evolved along with the vasculature in the ancestor of ferns and seed plants.

Introduction

Large yield increases have been achieved by selecting crop varieties particularly responsive to fertilizer applications (Evenson and Gollin, 2003). Yet nitrogen use efficiency of crops when inorganic nitrogen salts are applied to soil is low: as much as 50% of the applied fertilizer is not utilized resulting in often unwanted eutrophication of surrounding ecosystems. Chemical synthesis of nitrogen (N) fertilizer is presently consuming 1-2% of the yearly fossil fuel derived energy and thus is a major contributor to CO_2 pollution (Erisman et al., 2008). In addition, denitrification of the wasted N fertilizer leads to release of N_2O, with greenhouse warming effects up to 300 fold those of CO_2 (Galloway et al., 2004). Taken together, although spectacular yield increases obtained during the past century with most crops are founded on N fertilizers, their extensive use is unsustainable, and alternatives must be found to mitigate environmental impact of agriculture.

Reducing N-fertilizer applications without yield penalty may be possible with improved nitrogen use efficiency of crops combined with a more precise application of N-fertilizer (Xu et al., 2012) or by using crops fixing N_2 from the atmosphere. N_2 fixation is generally carried out by bacteria associated with the plants, as in root nodules of leguminous crops and in extracellular spaces of sugar cane, or by bacteria freely living in the soil and waterways (dos Santos et al., 2012; Pankievicz et al., 2015). The most established N_2-fixing crop symbiosis is the soybean (*Glycine max*) reaching 230 million metric tonnes annual production in 2008, ranking 4th in production area after wheat, rice, and maize (Hartman et al., 2011). Soybean harbours rhizobia which, after a complex communication between plant and bacteria is triggered, live in nodules of plant roots in facultative symbiosis (Freiberg et al., 1997). Rhizobia in the nodules are fed by plant sugars and secrete ammonium in return (Udvardi and Poole, 2013). Soybean plants reject colonization by rhizobia if nitrogen fertilizer is available. Modern high yield cultivars, are generally more responsive to N-fertilizer which contributes some 40% to protein in the seed (Wilson et al., 2014). Similar trends are found with the high yielding N_2-fixing forage alfalfa (*Medicago sativa*), that yields 21 t ha^{-1} $year^{-1}$ dry weight upon addition of 150 kg N ha^{-1} $year^{-1}$ (Anglade et al., 2015). To further emancipate plant protein production from N-fertilizer, fixation rates of symbionts in established crops will need to be improved (Oldroyd and Dixon, 2014). Alternatively, plant/bacteria symbioses that fix N_2 more efficiently will need to be domesticated. The latter will counter threats associated with increasingly reduced biodiversity on farm land (Stamp et al., 2012). Development of new or orphan crops is timely due to recent advances in genetics that tremendously speed up breeding.

N_2-fixing ferns from the genus *Azolla* have been used as bio-fertilizer in taro and rice cultivation in China, Vietnam and Senegal (Shi and Hall, 1988; Wagner, 1997); *Azolla* species distribution is global from temperate to tropical regions. N_2 fixation

of these ferns is generally thought to be carried out by the phototrophic bacterium *Nostoc azollae*, which was reported to have an eroded genome as a consequence of its symbiotic lifestyle (Ran et al., 2010). The notion that *N. azollae* are the only symbionts in *Azolla* has been challenged based on electron microscopic observations and this raises questions concerning their role in N trafficking within the symbiosis (Carrapiço, 1991; Zheng et al., 2009). *Azolla-Nostoc* symbioses have probably evolved some 90 M years ago (Metzgar et al., 2007), before the legume-rhizobia symbioses that evolved 64 M years ago (Herendeen et al., 1999; Kistner and Parniske, 2002). *N. azollae* are transmitted vertically during the fern life cycle and thrive inside specialized leaf pockets of the fern that close fully upon maturation after motile hormogonia have infected young leaves (Kaplan and Peters, 1981; Rai et al., 2000; Zheng et al., 2009). Leaves of *A. filiculoides* are arranged in two rows, each leaf has two lobes, one lobe rests on the water surface holding the other upper lobe in an aerial position exposed to light (Supplementary Figure 1). The upper leaf lobe exhibits a single leaf pocket with a single adaxially located pore (Supplementary Figure 1) presumed to mediate exchanges of gasses (Veys et al., 1999).

It is hypothesized that the fern contributes sugars while *N. azollae* releases 40% of the fixed nitrogen into the leaf pocket as ammonia (Peters and Meeks, 1989), which is then fixed by the glutamate synthase and glutamine synthetase (GS/GOGAT) cycle of the fern (Meeks et al., 1987). Molecular mechanisms underpinning the fern responses to N supplied by *N. azollae* are entirely lacking, yet a gene database was recently established from transcriptome sequencing (Brouwer et al., 2014). *Azolla* ferns were reported to reach doubling times as low as 2 days in their exponential growth phase without reduced nitrogen source, yielding an N-content of 4-5% of the dry weight (Peters et al., 1980). The ferns therefore ought to be considered as high protein crops in their own right. Data on productivities without N fertilizer is lacking for ferns grown in the linear growth phase in closed-canopy cultures, the latter being crucial to supress algae and outgrow fungal and insect pests.

Here we set out to examine the productivity of *A. filiculoides* under continuous production conditions over 138 days. The effect of various inorganic N-fertilizers on the productivity in terms of biomass, the N-content of the biomass and expression of key fern and bacterial marker genes were analysed simultaneously. Confocal laser scanning microscopy was used to reveal features of the leaf pocket important for metabolic connectivity of the symbiosis. *A. filiculoides* clones with and without *N. azollae* were used to verify the contribution of *N. azollae* to the fern growth rates. $^{15}N_2$-fixation rates were measured during the diel cycle and after acclimation to 2 mM NH_4NO_3. Diel transcriptome analysis was used to probe the fern diel physiological response to different nitrogen sources. RNA sequence assemblies were further used to examine ammonium transporters and markers of vasculature development.

Materials and methods

Plant materials

Azolla filiculoides Lam. was collected in Utrecht as described by Brouwer et al. 2014. *A. pinnata* R. Br. was from the International Rice Research Institute Biofertilizer Collection (code PI 0535, originally from Siguriya in Sri Lanka provided by S. Kulasooriya in 1984; (Watanabe et al., 1992)).

Growth conditions

In continuous production conditions, the Standard *Azolla* Medium (SAM) was liquid IRRI medium (Watanabe et al., 1992) with the following modifications: 0.32 µM, $CuSO_4$, 0.835 µM $ZnSO_4$ and 17.9 µM Fe-EDTA. Light was 300 µmol m^{-2} s^{-1} for 12 h and day time temperature was 22 °C whilst night-time temperature was 15 °C with the exception in the $^{15}N_2$-fixation experiment where temperature was kept a constant 21 °C. Mat density before the continuous harvest commenced reached 2-3 kg m^{-2} fresh weight (FW) biomass.

To obtain sterile cultures of the symbiosis, frond pieces (<1mm^3) of *A. filiculoides* were surface sterilized using bleach at 1% available chlorine for 40 s and four consecutive rinses in sterile water before culture on agar solidified SAM medium. SAM was solidified with 0.6% agarose (Duchefa, Netherlands). Fern clones without *N. azollae* were selected on agar solidified SAM with 60 µg ml^{-1} Erythromycin as previously described (Forni et al., 1991); absence of *N. azollae* was routinely verified by confocal microscopy using the characteristic bacterial fluorescence as a marker (Brouwer et al., 2014).

Sterile cultures of surface sterilized plants (1 L liquid medium with and without 2 mM NH_4NO_3) were grown in enclosed glass containers with a stream of air (78 L h^{-1}) pumped through 0.45 µm filters using aquarium pumps (SuperFish Air flow mini); these sterile cultures were used in all of the DNA and RNA sequencing experiments. Sterile cultures were also used for the $^{15}N_2$-fixation experiments.

Confocal laser scanning microscopy

Propidium iodide/periodic acid staining and mounting of the tissues were essentially as described by Truernit et al. (2008). A Leica SP2 laser scanning confocal fluorescence microscope, equipped with either × 10 or × 40 objectives and a laser with excitation wavelength of 405 nm, was used to visualize propidium iodide at 530-560 nm and auto-fluorescence at 515-530 nm.

RNA extractions and quantitative RT-PCR

Plant material was dry blotted, 50 mg FW snap frozen in liquid N_2, then ground frozen with two glass beads using the Tissue Lyzer II (Qiagen) set for 2x1.5 min at maximum speed. RNA was extracted with the Spectrum Plant Total RNA kit applying protocol B (Sigma-Aldrich). RNA was then treated with DNase (5 units for 3 µg RNA) 30 min at 37 °C and the reaction stopped by 10 min at 65 °C in the presence of EDTA (2 mM).

Primers (polyT and random hexamers at 0.030 and 0.074 µg $µl^{-1}$ final concentrations respectively in the reverse transcription reaction) and 1 µg DNase treated RNA were denatured 5 min at 72 °C before reverse transcription with MLV reverse transcriptase (5 Units, Fermentas) for 10 min at 37 °C then 40 min at 42 °C. Primers used for q-RT PCR are listed in Supplementary Table 1. Quantitative RT-PCR was performed as described in Brouwer et al. (2014).

RNA sequencing and bioinformatic analyses

Surface sterilized *A. filiculoides* were grown in excess sterile SAM with and without 2 mM NH_4NO_3 for 7 days in triplicate replicate cultures. Growth was in a chamber with light set to start at 06:00 for 12 h. On day 8 samples were collected, snap frozen at 2, 8, 14 and 20 h of the diel cycle, then total RNA extracted and DNAse treated as above. RNA integrity, sequencing library quality, and fragment size were verified on a 2100 Bioanalyzer (Agilent). Libraries were prepared using the TruSeq RNA Sample Prep Kit v2 (Illumina), and library quantification was performed with a Qubit 2.0 (Invitrogen). All libraries were sequenced on the HISEQ2000 Illumina platform in paired-end mode.

Paired-end reads were mapped against the transcriptome database (Brouwer et al., 2014) using CLC genomics workbench® (CLC bio, Denmark) using default parameters (mismatch cost 2, insertion cost 3, deletion cost 3, length fraction 0.8, similarity fraction 0.8, auto-detect paired distance, strand specific both and maximum number of hits for a read 30). Only paired matches were counted. Read counts were used for statistical analysis with edgeR in classic mode (Robinson et al., 2010) using the Bioconductor package. Results were corrected for multiple hypothesis testing using Benjamini Hochberg FDR correction (Benjamini and Hochberg, 1995) and considered significant if $q<0.01$. Read counts were normalized to total mapped reads for each sample and expressed as reads per million reads (rpm; Supplementary Tables 2 and 4). Principle component analysis and clustering were conducted on all transcripts which being summed exceeded 10 rpm using the MultiExperiment-Viewer (http://www.tm4.org/).

Clock genes were annotated as the closest homologue to the *Arabidopsis* clock

genes (Nakamichi, 2011). The diel transcripts were identified by isolating transcripts which display differential expression between at least one pair of two adjacent time points then loaded into Mapman for visualization and enrichment tests (Thimm et al., 2004). Functional annotations were imported using the Mercator Pipeline (Lohse et al., 2014) based on the *Azolla* transcriptome sequences (Brouwer et al., 2014). K-means clusters were built using the MultiExperimentViewer with Pearson distance metric.

N-content determinations and $^{15}N_2$ fixation rates

Ferns (100 mg FW) were grown in pots with 43 ± 4 ml of sterile SAM and a remaining air space of 262 ± 4 ml. To the air space 25 ml of $^{15}N_2$ was added using air-tight syringes; overpressure was slowly brought to ambient pressure using a release needle. Plants were subsequently grown in these conditions for 2 h at the following time-points: 8, 14, 20, 2 and 8 h in the diel cycle where the 12 h day started at 7 h. Ferns devoid of cyanobacteria at the 14 h time-point, in the light, were used as the no fixation control. In addition, the effect of 7 day acclimation to 2 mM NH_4NO_3 on ferns with cyanobacteria was tested at the 14 h time-point. Samples were frozen at -20°C, homogenized and freeze-dried before analysis of the dry weights and isotope abundance determinations. All data points represent the average of triplicate biological samples.

Total N content and stable nitrogen isotope ratios ($\delta^{15}N$) were analyzed on a ThermoScience Delta Plus isotope ratio mass spectrometer connected online to a Carlo Erba Instruments Flash 1112 elemental analyzer. The $\delta^{15}N$ of each sample was expressed relative to atmospheric dinitrogen as (‰) and calculated as the following ratio:

$$\frac{([^{15}N_{sample}] - [^{15}N_{air}]) * 10^3}{[^{15}N_{air}]}$$

with $[15N_{sample}]$ the concentration of ^{15}N isotope in the sample, $[15N_{air}]$ the concentration of ^{15}N in standard air. Instrumental precision was better than 0.2‰, based on analytical standards with known isotope composition. N_2 fixation rates were determined using the following calculation:

$$\frac{[^{15}N_{t2}] * [N_{biomass\ t2}] - [^{15}N_{t0}] * [N_{biomass\ t0}]}{2 * [^{15}N_{Airspace}]}$$

with $[15N_{ti}]$ the concentration of ^{15}N before ($i=0$) or after 2 h incubation ($i=2$) with

$^{15}N_2$ enriched air in the bottle, [$N_{biomass\ ti}$] the N content in the biomass before ($i=0$) and after 2 h incubation ($i=2$) and [$15N_{airspace}$] the ratio of ^{15}N over total N in the air space. We assumed no isotope discrimination during the fixation process and therefore rates of fixation calculated may be somewhat underestimated.

Results

Azolla filiculoides yielded 35 t dry weight ha^{-1} year^{-1} in an experimental continuous harvest system

To evaluate the potential of species from the genus *Azolla* as high protein biomass crops, *A. filiculoides* was grown under continuous production conditions for 138 days (Figure 1A). When harvesting 33% of the surface twice per week, accumulated biomass increased linearly (Figure 1B). From the slope of dry biomass accumulation, productivity was calculated at 97.2 kg ha^{-1} day^{-1}, corresponding to

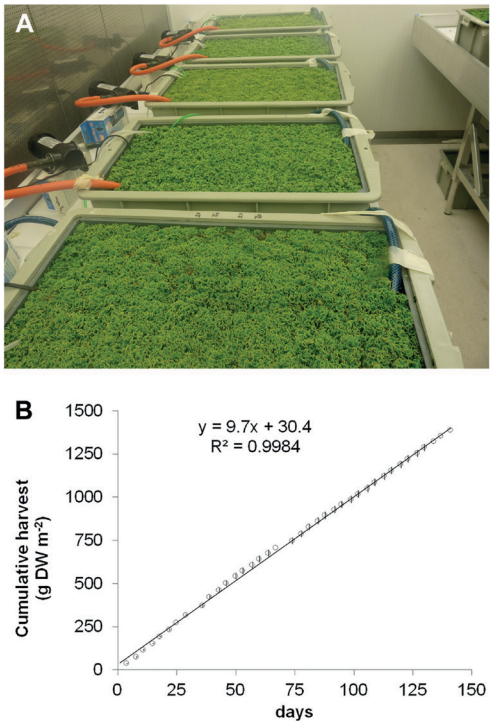


Figure 1. *Azolla filiculoides* yield potential under continuous production conditions. Ferns were grown with 300 µmol s^{-1} cm^{-2} photosynthetic active radiation at 25°C for the 12 h day and at 22 °C during the 12 h night. (A) Closed-canopy cultures with plants in the linear growth phase for continuous harvest. (B) Cumulative dry biomass harvested. Once standing crop formed a closed mat at 160 g dry weight m^{-2}, continuous harvest was at a rate of 33% of the biomass twice per week. Linear regression equation is shown along with its fit; n = 6, standard deviations were smaller than the labels.

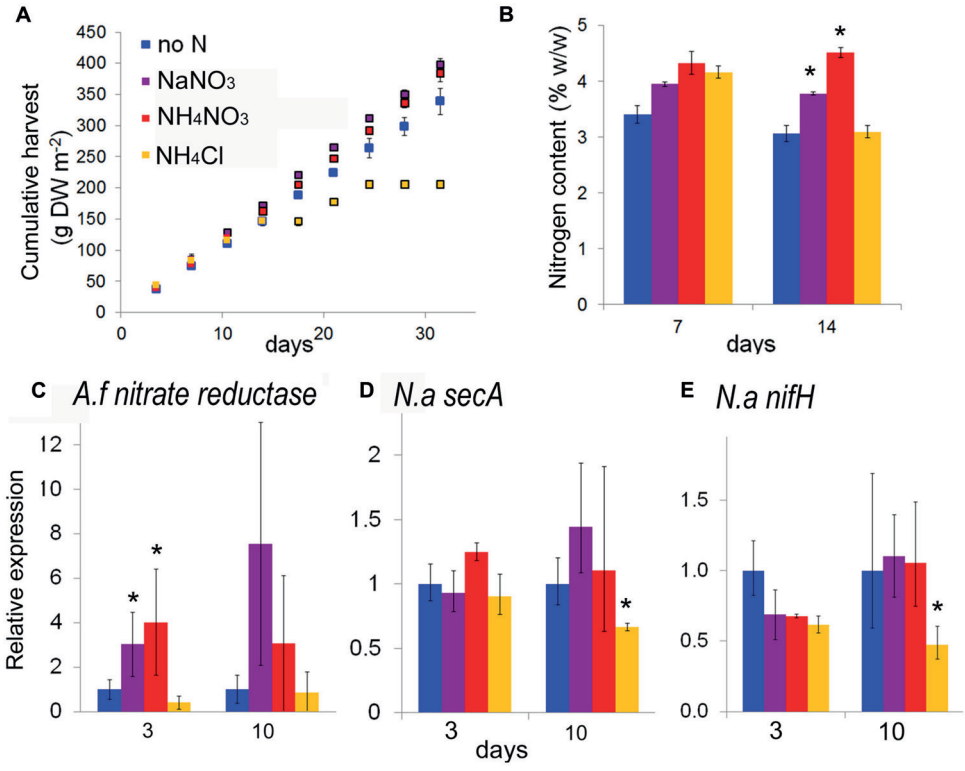


Figure 2. *Azolla filiculoides* response to N-fertilizer under continuous production conditions. Ferns were grown as in Figure 1 but this time in medium without (no N) or with 4 mM $NaNO_3$, or 2 mM NH_4NO_3, or 4 mM NH_4Cl. Biomass harvested was split to determine cumulative harvest (A) nitrogen content (B) and to extract mRNA. Reverse transcription of the RNA was with an excess random hexamers to analyze levels of *A. filiculoides* and *Nostoc azollae* transcripts by quantitative PCR: *A. filiculoides* nitrate reductase Afcontig_35782 (C), *N. azollae secA* (D) and *nifH* (E). Standard deviation is for n = 3, T-test significance with reference to no N for $P < 0.05$ is marked with *

an annual productivity of 35.5 t ha^{-1} biomass dry weight (DW) for *A. filiculoides*. Very similar results were obtained with *A. pinnata* (data not shown). To test whether yield performance of these ferns was influenced by N-fertilizers often encountered in runoff from agricultural land, we examined the effect of several inorganic nitrogen sources in the media used for continuous production. Growth of *A. filiculoides* was in standard *Azolla* medium (SAM) without nitrogen or with 4 mM $NaNO_3$, NH_4Cl or 2 mM NH_4NO_3. We harvested 33% of the culture area twice weekly over a period of 4 weeks, whilst replenishing nutrients weekly. Cumulative harvest over 30 d indicated slight benefits with N fertilizer (up to 15%, Figure 2A). N fertilizers moreover increased N accumulation in the biomass, from 3.41 up to maximally 4.36 % w/w N

with $NaNO_3$ (Figure 2B) and to induced expression of nitrate reductase from the fern (Figure 2C). $NaNO_3$ as well as NH_4NO_3 did not affect transcript abundance of the bacterial housekeeping gene *N. azollae secA* suggesting that bacterial growth was not affected by NO_3-containing media (Figure 2D). In contrast, NH_4Cl reduced *A. filiculoides* productivity and led to a reduction of *N. azollae secA* and *nifH* transcript detection (Figure 2D,E). *NifH* encodes the catalytic subunit of the nitrogenase enzyme and is intact in *N. azollae* (Ran et al., 2010). The above results demonstrate that *A. filiculoides* does not require N fertilizer to sustain high productivity but that it reacts to and benefits from additional N sources in the medium.

Growth of *A. filiculoides* without *N. azollae* entirely relied on exogenous N fertilizer

To evaluate the contribution from *N. azollae* to fern productivity we raised surface sterilized *A. filiculoides*, then for every fern, fronds were grown without and with erythromycin so as to obtain clonal material with and without cyanobacteria. Complete absence of *N. azollae* was verified using confocal microscopy, quantitative PCR on, and sequencing of the extracted plant DNA (data not shown). *A. filiculoides* without *N. azollae* failed to grow without a source of N (Figure 3A). Growth in all other conditions was exponential. Optimum supplementation of ferns without *N. azollae* was obtained using 4 or 8 mM $NaNO_3$, or 2 mM NH_4NO_3, and yielded growth rates approaching those of ferns with *N. azollae* (Figure 2A, Figure 3A). Given the possibility of amino acids as exchange substrates between fern and symbiont in addition to ammonia (Kaplan and Peters, 1981; Meeks et al., 1987), amino acids were tested for their ability to restore growth keeping the total N supplied in the amino acids at 2 mM in the medium (Figure 3B). Whilst arginine and proline supported highest average growth rates that were not significantly different from ferns with *N. azollae*, growth rates on amino acids mostly correlated with those from the seed plant *Arabidopsis thaliana* (Forsum et al., 2008) and, therefore, unlikely reflected trafficking inside leaf pockets. Close-up inspection confirmed that the leaf pockets were well removed from the medium and enclosed entities except for the leaf pocket pore open to air (Figure 3C and Supplementary Figure 1). Confocal microscopy further revealed a conspicuous vasculature curving around the pockets that was closely connected to the pocket fluid (Figure 3C). This vasculature differed from that in leaf lobes without leaf pocket: it was made of several layers of tracheid cells with reinforced cell walls containing pectin brightly stained by the propidium iodide/periodic acid stain (Supplementary Figure 3).

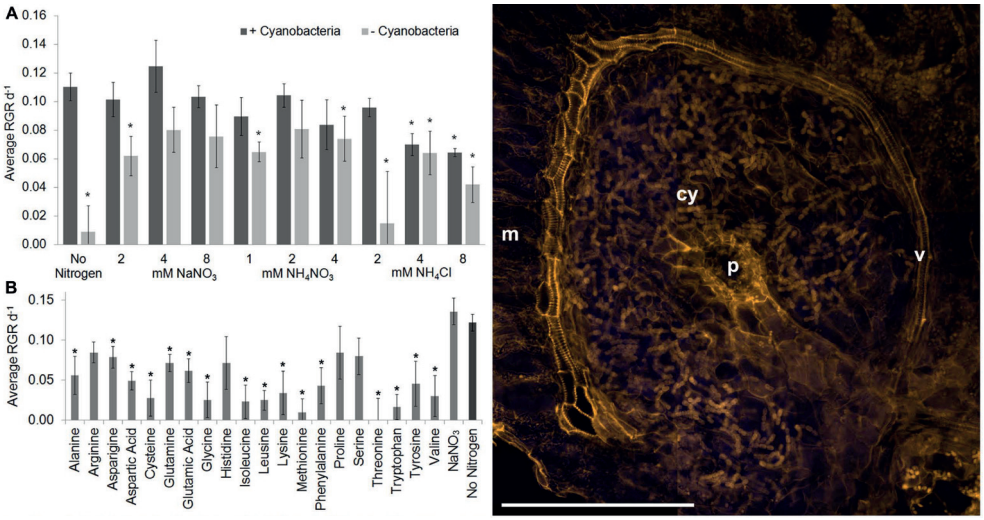


Figure 3. Role of *N. azollae* in sustaining fern growth and specific structures of the leaf pocket. Clonal ferns with (+cyano) and without N. azollae (-cyano) were grown on agar solidified Standard *Azolla* Medium (SAM). (A) Growth in inorganic N-fertilizer ranging from 2 to 8 mM. (B) Growth with 2 mM nitrogen in differing amino acids, or with 4 mM $NaNO_3$ (NO_3), and +cyano on SAM for comparison. Growth was measured as leaf area. Ferns without *N. azollae* failed to grow on SAM but in all other conditions growth fitted an exponential curve. Standard deviations with n = 3, T-test significance with reference to ferns +cyano on SAM for $P < 0.05$ is marked with *. (C) Structures revealed by propidium iodide staining using confocal microscopy 405 nm excitation, 560 nm emission (orange); auto-fluorescence at 515–530 nm emission (blue). v, vasculature; p, pore; cy, *N. azollae* bacteria; m, mesophyll cells; scale bar: 200 µm.

N_2-fixation rates reached 0.4 mg N g^{-1} dw h^{-1} at noon and dropped 7.5 times in the dark

N. azollae are phototrophic bacteria, therefore, we set out to measure diel $^{15}N_2$ fixation in *Azolla*. We first established that N_2-fixation rates were constant over at least 12 h when keeping light, temperature and CO_2 concentrations constant (data not shown). Ferns were exposed to $^{15}N_2$-enriched air for 2 h starting at 8, 14, 18, 20 and 8 h in the diel cycle of a growth chamber with 12 h light starting at 7 h. Ferns without *N. azollae* constituted the control non-fixing ferns at the 14 h time point. The relative enrichment of ^{15}N over air (‰^{15}N) in *Azolla* biomass at 20 and 2 h in the dark was up to 7.5 times lower than that measured at noon (Figure 4A). *N. azollae*, therefore, fixed the bulk amounts of N_2 during daytime.

We further assessed $^{15}N_2$-fixation rates in ferns acclimated for a week to +N (2 mM NH_4NO_3) or –N (SAM) during peak N_2 fixation at 14 h (Figure 4B). Whilst the

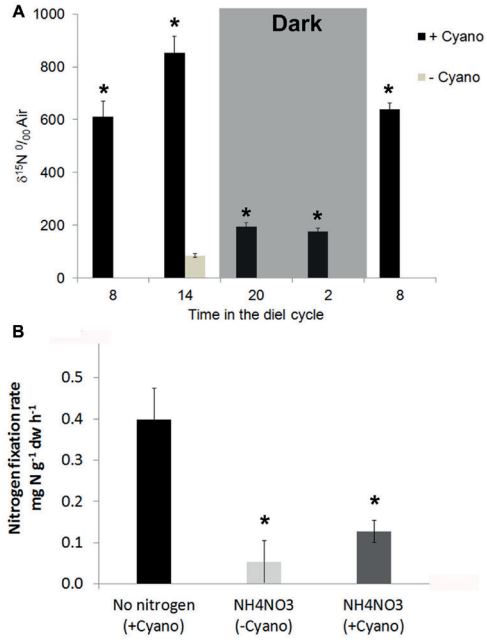


Figure 4. $^{15}N_2$ fixation during the diel cycle and after acclimation to 2 mM NH_4NO_3. (A) Raw ^{15}N enrichment after 2 h exposure to $^{15}N2$ averaged for the time points 8, 14, 20, and 2h during the diel cycle. Ferns were on SAM without nitrogen, the 12 h day started at 6 h. Temperature was kept a constant 21°C. Ferns with *N. azollae* (+Cyano, black) and without (–Cyano, light gray). (B) $^{15}N_2$-fixation rates averaged over 2 h at the timepoint 14 h in the diel cycle after a week acclimation on SAM without [No nitrogen (+Cyano), black and with 2 mM NH_4NO_3 (NH_4NO_3 (+Cyano), dark gray]; control ferns without *N. azollae* [NH_4NO_3 (–Cyano), light gray]. Standard deviations with n = 3. T-test significance marked with * was P < 0.05 for +Cyano data points with reference to ferns –Cyano on SAM for (A); and with reference to +Cyano ferns on no N in (B).

N_2-fixation rate in ferns grown without nitrogen reached 0.394 mg N g^{-1} DW h^{-1}, N-fertilizer reduced $^{15}N_2$-fixation rates four fold. This large decrease in fixation rates further indicated that responses studied in ferns +N resulted from N-uptake rather than the combination of fixation and uptake.

Transcriptional patterns of the diel cycle dominate over the response to N-fertilizer

The fern's ability to react to N-fertilizer prompted an analysis of the transcriptional mechanisms underpinning these responses; the specificity with which cDNA libraries were generated afforded stringent spatial resolution thus distinguishing for the first time processes in the fern from those in the cyanobacteria. Given the diel behavior of nitrogen assimilation genes in the angiosperm *A. thaliana* (Cheng et al., 1991; Gutierrez et al., 2008), we chose to compare the diel transcript profiles of ferns acclimated for a week to SAM without nitrogen (-N) or with 2 mM NH_4NO_3 (+N). Ferns were sampled from triplicate cultures at 2, 8, 14 and 20 h within the 24 h cycle of a growth cabinet with 12 h light starting at 6 h in the morning. RNA-seq resulted in 17.3 to 37.8 million read pairs per sample of which between 47 to 55% mapped as read pairs and 63 to 65% mapped as single end reads to the transcriptome database ((Brouwer et al., 2014); Supplementary Table 2_mapping_stats). Transcriptional investments in functional categories were similar in *A. filiculoides* and *Arabidopsis*

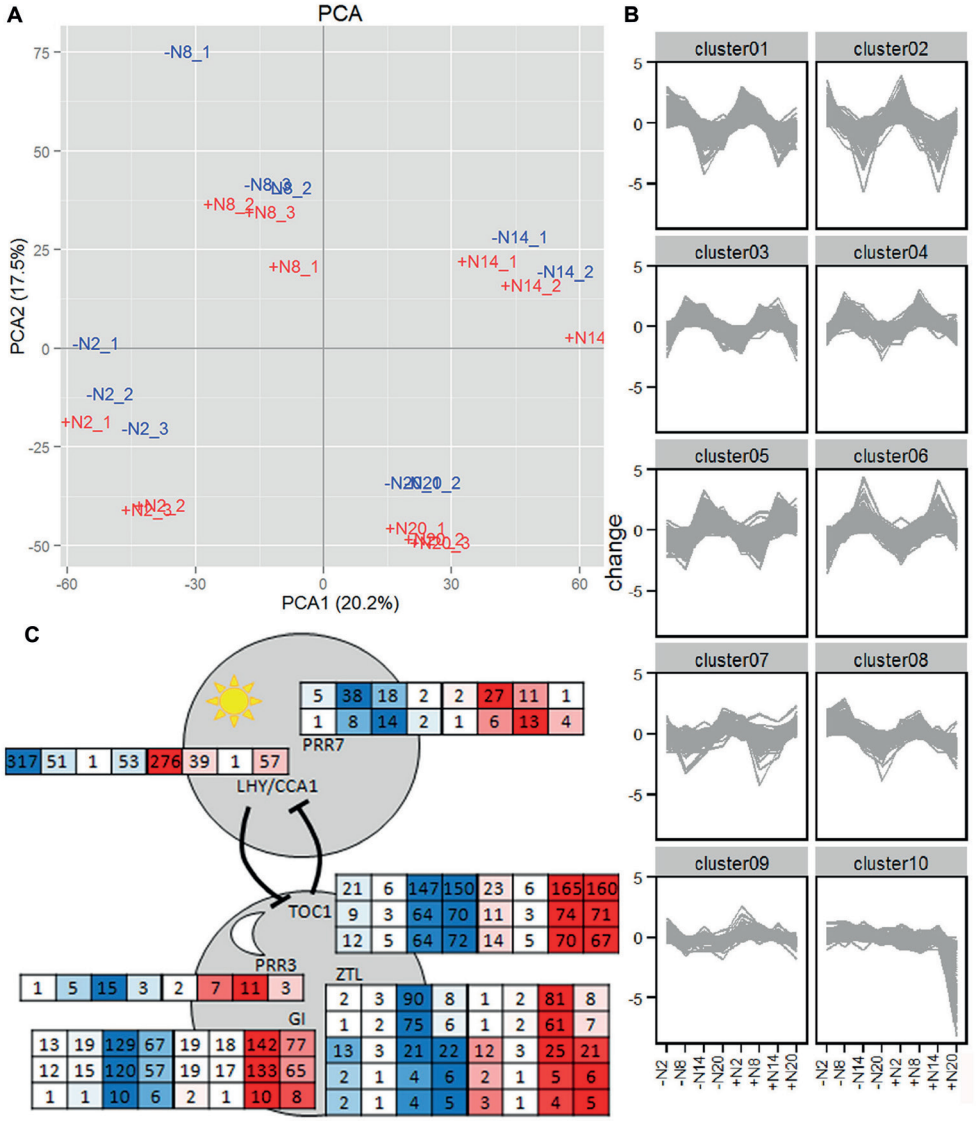


Figure 5. Diel transcript profiles in *A. filiculoides* grown with and without 2 mM NH_4NO_3. Ferns were acclimated to SAM without (–N, in blue) and with 2 mM NH_4NO_3 (+N, in red) for a week then harvested from triplicate cultures at 2, 8, 14, and 20 h with 12 h day-light starting at 6 h. RNA was extracted, sequenced and sequences analyzed as described in the section "Materials and Methods." (A), Principal Component Analysis (PCA) and (B), K-means clustering included all genes with a read count > 10 rpm for all time points and conditions. K-means clusters were built with Pearson distance metric. (C) Clock components from the morning (above) and evening (below) loops with their respective read counts in rpm, on medium without N (in blue) and with N (in red) at 2, 8, 14, and 20 h in each box, respectively.

(Supplementary Figure 3). Principle component analysis indicated four groups that corresponded to the different time points during the day which separated in both the first and second dimension and explained 38% of the variation in the dataset (Figure 5A). The first two dimensions also resolved the N-fertilized samples: +N samples projected apart from the –N samples in the same direction at all time points.

Diel transcripts were defined as transcripts with read counts significantly changed in at least two adjacent time points in either the +N or –N series. Of all transcripts detected, 4872 or 13% were diel (Supplementary Table 2_diel_transcripts). K-means clustering of diel genes identified clusters (Figure 5B; Supplementary Table 2_clusters) with genes peaking at 2 h (8h into the dark, cluster 1 and 2), genes peaking at 8 h (2h into the light, cluster 3 and 4), genes peaking at 14 h (8h into the light, cluster 5 and 6) and genes peaking at 20 h (2h into the dark, cluster 7) with no difference between the –N and +N samples, indicating that diel rhythms of transcription were unresponsive to N-fertilizer, with the exception of genes in cluster 10 (369 contigs) that sharply decreased 2 h into dark if N was supplemented (Figure 5B; Supplementary Table 2_cluster_10). Cluster 10 genes included genes from the Calvin-Benson cycle, PS1 and PS2 light harvesting complexes, and glycine decarboxylase P-protein 2 (Af_7035) from photorespiration; in addition it included the amino acid sensor ACT domain repeat 3-like protein (Af_11070) and three sugar transporters (Af_3233, Af_17210 and Af_14354).

Clock-gene expression patterns corresponded to those in *Arabidopsis* with the morning loop peaking at 2 h and 8 h and the evening loop peaking at 14 h and 20 h although the number of isoforms did not correspond to the numbers in *Arabidopsis*; clock-gene transcripts did not respond to N fertilizer (Figure 5C; Supplementary Table 2_clock_genes). Diel transcripts were enriched in starch and sucrose metabolism, pathways of photosynthesis, nutrient assimilation and selected pathways in secondary, lipid and amino acid metabolism (Supplementary Table 3). Diel rhythms of gene expression and clock components in *A. filiculoides* ferns were hence similar to those in angiosperms. In the fern, moreover, diel rhythms of gene expression dominated over the response to N fertilizer.

Transcriptional profiles of ferns grown without N fertilizer reflected early morning recovery from night-time N deficiency

Acclimation to +N (2 mM NH_4NO_3) compared to –N changed the amounts of 526 transcripts significantly when all contigs were counted that were changed in at least one time point (Supplementary Table 2_N_responsive). The changes were, however, not uniform throughout the day: most changes occurred during darkness at 20 and 2 h with 250 and 240 changed contigs respectively (Figure 6A). In contrast only 28 contigs were changed at noon (14 h; Figure 6A).

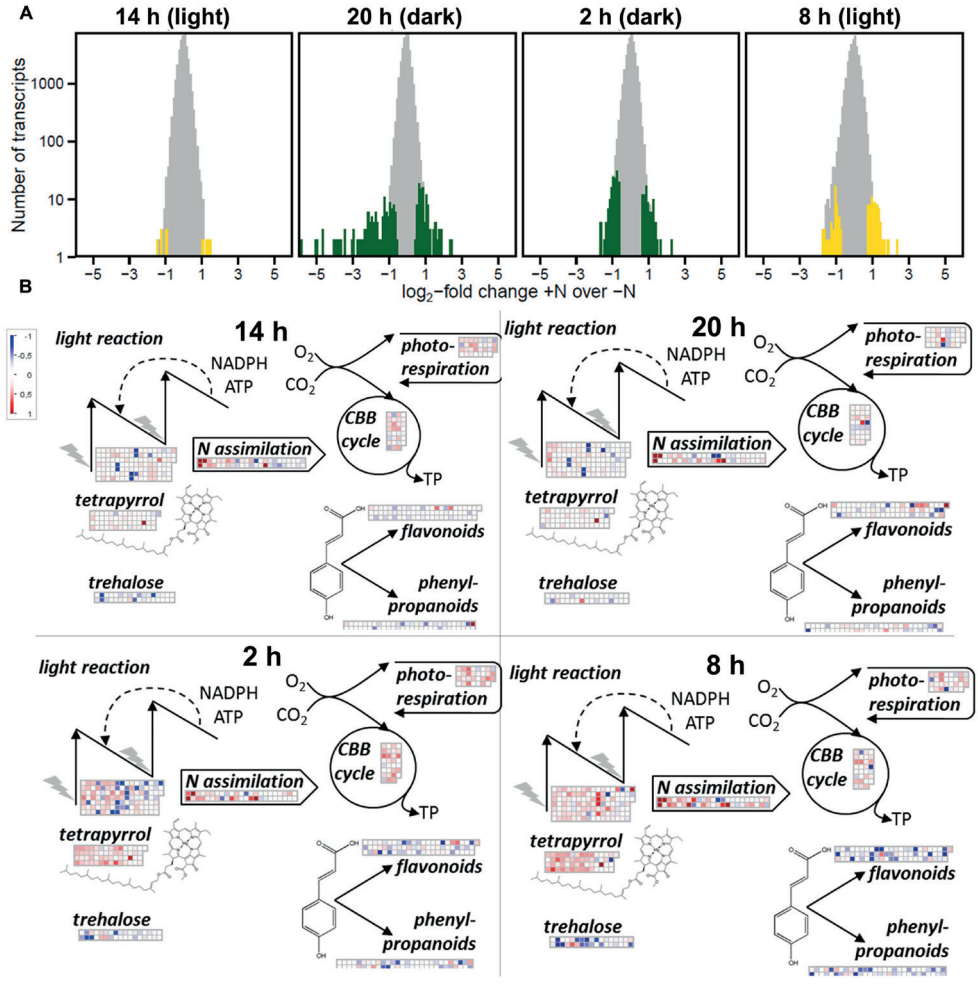


Figure 6. Metabolic responses of *A. filiculoides* to diurnal supply of reduced N from *N. azollae*. (A) Number of genes differentially transcribed comparing log 2-fold change of the read-count ratios in ferns with 2 mM NH_4NO_3 (+N) over without (–N). Yellow, time points in the light; green, time points in the night. (B) Mapman overview of transcriptional changes in metabolism comparing transcription +N over –N. Noon (14 h), evening (20 h; 2 h into dark), night (2 h), and morning (8 h; 2 h into the 12 h day). Upregulation in –N is depicted in red, downregulation is depicted in blue, maximal coloration is set at twofold changes.

To visualize N-dependent transcriptional investments in central metabolic pathways, transcript read counts were submitted to the Mapman software (Figure 6B; Thimm et al., 2004). In the middle of the night (2 h) and in the morning (8 h), +N conditions led to accumulation of transcripts involved in tetrapyrrol synthesis ($q<10^{-9}$),

the Calvin-Benson cycle ($q<10^{-20}$), and in the morning transcripts of the light reaction ($q<10^{-9}$). In contrast, transcripts from flavonoid ($q<10^{-3}$), phenylpropanoid synthesis ($q<10^{-3}$) and trehalose metabolism ($q<10^{-2}$) accumulated in –N conditions, most strongly in the morning at 8 h (Figure 6B; Supplementary Table 4). This pattern essentially reflected the pattern reported in N-starved *Arabidopsis* upon nitrate addition (Scheible et al., 2004) and therefore suggested N deficit at night in ferns grown in the absence of N.

Pathways for the assimilation of N into amino acids remained unchanged but ureic acid metabolism changed with N supply

Transcriptional investment in N-assimilation pathways into amino acids, remained unchanged when comparing +N with –N with very few exceptions: transcripts accumulated for two contigs of glutamine dehydrogenase (af2747 and af3257) in ferns on +N but were much less abundant than transcripts of GOGAT and GS enzymes (Table 1; Supplementary Figure 4), and therefore may represent enzymes in specific tissues exposed to +N from the medium.

At night in plants grown –N, stored ureides may have been remobilized by al-

Table 1. Transcripts associated with N metabolism that responded to +/−N conditions at the time points 2, 8, 14, or 20 h of the diel cycle.

	Average read abundance in rpm at each time point of the diel cycle							
	-N				+N			
Protein	2	8	14	20	N2	N8	N14	N20
Allantoate amidohydrolase	475	712	101	57	166	170	73	48
Allantoate amidohydrolase	85	123	20	11	35	33	13	7
Ammonium transporter 2	4	3	4	4	0	1	0	1
Urea transporter	69	56	96	69	20	25	50	34
Aspartate aminotransferase 5	16	17	19	17	15	10	16	0
Ureide permease 2	36	43	30	24	22	26	27	21
Glutamine synthase clone R1	97	128	110	107	113	88	106	1
NADH-GOGAT	147	136	153	172	167	153	163	239
Aspartate aminotransferase 1	18	19	16	21	30	28	17	21
Glutamate dehydrogenase 1	20	21	9	16	67	46	25	47
Glutamate dehydrogenase 1	3	4	2	2	14	10	4	10
Nitrate reductase 1	24	6	3	8	38	35	22	48
Nitrite reductase 1	20	16	18	24	85	141	90	238
Nitrite reductase 1	27	22	24	31	133	199	126	289

* q-value was corrected for multiple hypothesis testing by Benjamini Hochberg, significant changes are bold.

lantoate amidohydrolase as reads from this enzyme were abundant. Reads of this enzyme were low at night in ferns grown +N (Table 1). Consistently in ferns grown +N at night, N-remobilization likely was decreased since transcripts of a transporter of urea and ureide permease were fewer whilst urea biosynthesis likely was maintained since transcripts of the aspartate aminotransferase 1 and NADH-dependent GOGAT accumulated (Table 1).

Genes affected by +N throughout the day included the nitrite and nitrate reductase genes (Table 1) and confirm that fern metabolism responded to nitrate in the medium. Similarly increased transcript of the HPP family nitrite transporter (af1275; Supplementary Table 2_N_responsive) in ferns +N suggested increased nitrite transport to chloroplasts in fern cells. This nitrate transporter is believed to be cyanobacterial origin and was reported absent in mosses (Maeda et al., 2014) Increased transcript abundance in ferns –N of an AMT2 type ammonium transporter (AMT; Table 1), two phosphate transporters homologous to AtPHT1 and AtPHT3 (af380, af1508; Supplementary Table 2_N_responsive), and a Nodule INception (NIN) transcription factor with homology to AtNLP7 (af7964; Supplementary Table 2_N_responsive) suggest that ferns with active N_2 fixation regulate P and N transport coordinately as known from the angiosperm symbioses with arbuscular mycorrhiza (Breuillin-Sessoms et al., 2015). Manual assembly of AMT transporter sequences allowed construction

Continuation of Table 1.

Fold change +/- N*				Q-value +N versus –N*			
2	8	14	20	2vsN2	8vsN8	14vsN14	20vsN20
-1.51	**-2.06**	-0.46	-0.25	**2.0E-04**	**3.4E-03**	1	1
-1.27	**-1.88**	-0.56	-0.56	**9.2E-03**	**5.2E-03**	1	3.7E-01
-1.65	**-1.25**	**-1.93**	**-1.25**	**1.5E-06**	**1.2E-03**	**2.8E-07**	**3.0E-03**
-1.73	-1.10	-0.92	**-1.00**	**3.6E-16**	4.1E-02	7.0E-01	**3.9E-03**
-0.11	-0.75	-0.24	**-4.03**	1.0E+00	7.6E-01	1	**4.9E-29**
-0.66	-0.72	-0.16	-0.13	**3.9E-03**	1.1E-01	1	1
0.21	-0.54	-0.05	**-5.87**	1	9.9E-01	1	**8.1E-92**
0.19	0.16	0.09	**0.47**	8.6E-01	8.1E-01	1	**7.2E-03**
0.72	0.55	0.10	0.04	**2.0E-03**	1.1E-01	1	1
1.69	1.12	**1.42**	**1.50**	**3.0E-11**	1.1E-02	**8.3E-07**	**2.2E-08**
1.96	1.23	0.98	1.78	**2.0E-09**	2.4E-02	1.6E-01	7.9E-15
0.64	**2.36**	**2.63**	**2.49**	7.7E-02	**2.8E-11**	**2.2E-14**	**6.3E-45**
2.01	**3.03**	**2.22**	**3.27**	**1.5E-14**	**1.5E-15**	**6.6E-11**	**1.4E-33**
2.25	**3.11**	**2.36**	**3.16**	**3.5E-14**	**3.5E-23**	**3.1E-17**	**2.8E-38**

of a phylogenetic tree showing that the fern AMT transporters are both of the MEB (AMT2 type in *Arabidopsis* which is electroneutral) and AMT (AMT1 type in *Arabidopsis*) clades (Figure 7; McDonald and Ward, 2016; Neuhäuser et al., 2009); AMT2 had increased read counts in ferns with actively N_2-fixing cyanobacteria (Table 1). *N. azollae* AMT clustered with the bacterial AmtB clade, yet the *N. azollae* AmtB that is functional was not in the genome location of closely related strains which was inactivated by a mobile element (Supplementary Figure 5).

Key genes of the vasculature were upregulated under –N growth conditions

Xylem was a prominent feature of the vasculature curving around the leaf pocket (Figure 3C). This specialized vasculature may have as much a structural function as a nutrient and water transport function. In ferns –N compared to +N, transcripts of fern homologues of the xylem cystein peptidase accumulated (XCP3, af1829) and of the CYTOKININ OXIDASE decreased (Azfi34233), suggesting active xylem formation (Supplementary Table 2_N_responsive). Further investigation of xylem specific genes revealed homologues of VND6 and LBD15 as well as IRX3 in the *A. filiculoides* sequences. Phloem cells lined the xylem cells for most of the vasculature curving around the leaf pocket (Figure 3C and Supplementary Figure 2B,C). To test a possible role of vasculature in the response to differential N supply, full length gene-homologues characteristic for the vasculature were manually assembled then reads mapped to quantify differential accumulation of the mRNA with more accuracy. Transcripts of a fern homologue to AtGL22 were amongst the most differentially accumulating transcripts (13 x averaged over all time points) under conditions -N compared to +N, GL22 is known from *Arabidopsis* to regulate allocation of phloem nutrient contents.

Discussion

Light-driven N_2 fixation by *N. azollae* entirely supported the high yields of *Azolla* biomass

Clones of *A. filiculoides* without *N. azollae* required N fertilizer for growth and inorganic nitrogen sufficed to reach growth rates nearing those of clones with symbionts (Figure 3A). *N. azollae* supplied sufficient N to support high biomass yields with little gain from N-fertilizer (Figure 2A). N_2-fixation rates for *Azolla* on medium without N-fertilizer peaked at 0.4 mg N g^{-1} DW h^{-1} during midday but averaged

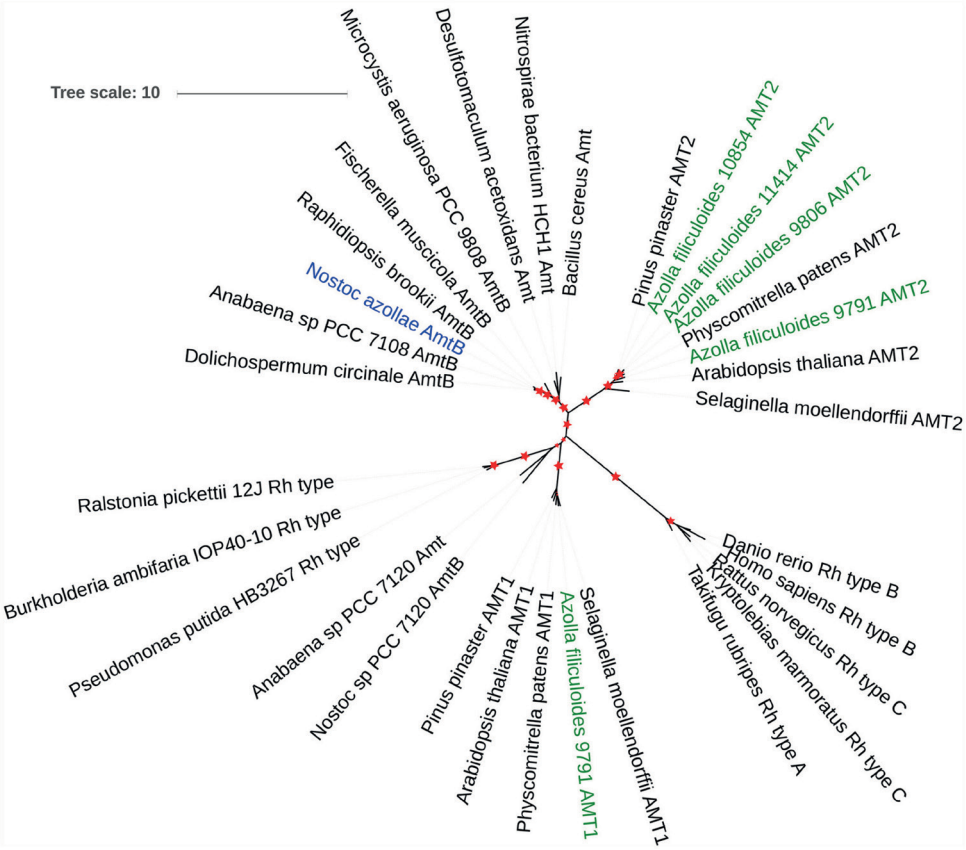


Figure 7. The origin of ammonium transporters from *A. filiculoides* and *N. azollae*. Sequences were manually assembled (Supplementary List of AMT Protein Sequences) then aligned with MUSCLE, the phylogenetic tree calculated using PhyML with 100 bootstraps (Dereeper et al., 2008) and then visualized using iTol (Letunic and Bork, 2016), bootstrap values ranging from 70 to 100% are represented by the increasing red star sizes.

0.17 mg g^{-1} DW h^{-1} over the diel cycle; this corresponded to the N increase in the biomass observed in our in growth experiments (Figure 2), reaching 0.33 g m^{-2} d^{-1} or 0.15 mg g^{-1} DW h^{-1}. To compare, N_2-fixation rates in soybean averaged 0.08 mg g^{-1} DW h^{-1} (Hung et al., 2013). Diel variation of N_2 fixation in clover (*Medicago truncatula-Sinorhizobium meliloti*) was small and entirely dependent upon temperature (Cabeza et al., 2015). The diel variation reported here for *A. filiculoides* was measured under constant temperature and therefore the high N_2-fixation was diurnal, and depended on light.

N. azollae is a filamentous cyanobacterium. Typically N_2-fixation is sequestered to heterocysts in filamentous cyanobacteria and occurs mostly during the day when, firstly, cells of the filament supply carbon to fuel heterocyst glucose-6-phosphate

dehydrogenase (G6PDH) providing reductant NADPH and, secondly, light energy captured by heterocyst photosystem I supplies ATP; G6PDH was required for N_2-fixation and growth of *Nostoc sp.* in the dark (Gallon, 1981; Summers et al., 1995). *N. azollae* lack glycolysis pathway enzymes and therefore may channel all carbon through G6PDH to generate NADPH required to reduce N_2; the *N. azollae* genome retained PS1 for ATP production in light (Ran et al., 2010). In contrast, heterotrophic *Rhizobium meliloti* fixes N_2 in the darkness of root nodules and therefore requires carbohydrates to synthesize NADPH as well as ATP: malate provided by the plant provides NADPH via malic enzyme, and is further respired via oxidative phosphorylation specialized for micro-aerobic environments. The malic enzyme gene is required for N_2 fixation by *R. meliloti* (Driscoll and Finan, 1993). These differences may underlie the higher N_2-fixation rates observed here with phototrophic diazotrophic symbionts in *Azolla* compared to heterotrophic diazotrophic symbionts in legumes.

Azolla yields without N-fertilizer input compete favorably with established plant protein crops

Nitrate fertilizer at 4 mM increased productivity of *Azolla* by merely 15% (Figure 1C) compared to 175% and 132% increase reported when 5 mM NO_3^- was administered to elite breed soybeans 40 and 60 days after planting (Yashima et al., 2003). RNAseq and quantitative RT-PCR revealed that the fern responded to exogenous NO_3^- with the accumulation of transcripts from nitrite and nitrate reductase (Figure 2C; Table 1), and a nitrite transporter (Supplementary Table 2) consistent with previous data of *A. pinnata* showing that labelled NO_3^- is incorporated in roots and then is seen to move slowly to shoots (Ito and Watanabe, 1983). Exogenous NH_4NO_3 at 2 mM resulted in a 78.4% decrease in $^{15}N_2$-fixation by the symbiont (Figure 4B). Similarly, N_2 fixation decreased by 20% to 70% in 99 different strains of *Azolla* on 2.86 mM $(NH_4)_2SO_4$ (Okoronkwo et al., 1989). This may be due to diffusion of external NH_4 into the leaf pocket: complete inhibition of N_2 fixation took place with 1 mM NH_4^+ but not NO_3^- in the related free-living *Anabaena cylindrica* (Ohmori and Hattori, 1974). Alternatively, the fern may contribute to actively regulate N_2-fixation in the symbiont by controlling gas exchange via the leaf-pocket pore (Veys et al., 1999) and O_2 release. Increased O_2 exposure as evidenced by increased PS2 and hemoglobin expression in ferns +N further suggest this (Supplementary Table 2); NO_3^- or urea from the fern unlikely affect N_2 fixation in the symbiont directly since genes for import and assimilation of NO_3^- and urea were reported missing in *N. azollae* (Ran et al., 2010). N_2 fixation by rhizobia in the nodules of legumes is also inhibited by NH_4^+: NH_4SO_4 at 5 mM rapidly decreased N_2 fixation in alfalfa roots exposed over a period of 5 days (Cabeza et al., 2015). In plant symbioses with either cyanobacteria or rhizobia, therefore, the N-substrate exchanged when present in the surrounding

medium inhibits N_2-fixation of the bacteria.

In contrast to the large majority of agricultural crops, *Azolla* species grow in fresh shallow surface water. Once the surface of the water was covered by exponential growth, our experimental farming system produced biomass linearly and stably over periods exceeding 130 days (Figure 1B). Given a measured N-content of 3.5% w/w biomass DW and 35 t ha^{-1} annual productivity without any N-input (Figure 1), *A. filiculoides* fixed over 1200 kg N ha^{-1} $year^{-1}$ and therefore has potential as plant protein crop in temperate regions. Established forage crops that still require some 150 kg ha^{-1} $year^{-1}$ N-fertilizer to reach 25 t ha^{-1} $year^{-1}$ fix some 600 kg N ha^{-1} $year^{-1}$ in the harvested material (Anglade et al., 2015; de Visser et al., 2014). Modern soybean varieties yielding 5 t beans ha^{-1} with no added fertilizer to the soil fixed up to 300 kg N ha^{-1} per crop (Gelfand and Robertson, 2015). Our experimental system aimed at industrialization and thus was closed and entirely controlled, but promising yields have been reported in an open system of 468 m^2 in the tropics (Colombia), where *A. filiculoides* yielded 39 t ha^{-1} $year^{-1}$, albeit in this study nitrogen-containing chicken manure was supplied to the fern (Becerra et al., 1990). Feeding *Azolla* biomass to pigs as protein-rich admixtures to soymeal supported growth as well as control soymeal diets (Becerra et al., 1990; Leterme et al., 2009; Leterme et al., 2010). Continuous production is particularly suited for on-farm processing as it requires processing units of smaller capacity. *Azolla* ferns efficiently use run-off water from fields and remove both nitrogen and phosphate (Shilton et al., 2012), thereby preventing eutrophication if harvested regularly and contributing to closing the nitrogen and phosphate cycles. If grown as dense mats the ferns may help re-solve iron oxides, while releases adsorbed phosphate and therefore be an alternative to mining increasingly rare phosphate fertilizer from sediments of shallow freshwater (Cordell et al., 2011). The findings reported here warrant a detailed assessment of potential (agro-) ecosystem benefits and services of *Azolla* farming, including potential threats to protected wetland ecosystems.

How do the ferns deal with the diurnal cyanobacterial N_2 fixation?

If hosting cyanobacteria is such an efficient solution to fixing large amounts of N_2, we wondered why taxa hosting cyanobacterial N_2 fixation are comparatively rare and what adaptations would be required by the host plants.

Since N_2 fixation was mostly light driven, the fern host alternated between high N availability during the day and low N availability during the night; low N availability may be restricted to the later part of the night due to buffering by the high NH_4^+ concentrations in the *N. azollae* and leaf pocket fluid (Kaplan and Peters, 1981; Meeks et al., 1987). Consequently, the transcriptional patterns differed in *A. filiculoides*

supplemented with NH_4NO_3 and without. The morning time point without NH_4NO_3 (Figure 5B) displayed transcripts highly reminiscent of those in *A. thaliana* when supplied with nitrate after N-starvation (Scheible et al., 2004). The responses to low N availability were therefore conserved between ferns and seed plants. The transcriptional response to low nitrogen availability disappeared during the day with the noon time point having the fewest changes between +N and –N conditions and the highest *N. azollae* N_2-fixation rates. The magnitude of the fern transcriptional response to N fertilizer was therefore much lower than that reported for spermatophytes entirely dependent on externally available N (Wang et al., 2003; Vidal et al., 2013). As N supply was short during the night, it is likely that *Azolla* possess a storage system for N which balances availability during the day and the night, this could have been reflected by the prominence of urea transporters in the list of genes accumulating in ferns without N-fertilizer. Night-time accumulation of transcripts encoding enzymes of allantoin metabolism in ferns without N fertilizer points to allantoin as an intermediate in N storage. To host phototrophic diazotrophic cyanobacteria, therefore, ferns adapted existing responses required to cope with irregular supply of N-fertilizer from the environment that evolved in ancestors common to ferns and seed plants.

What features ensure metabolic connectivity between host and cyanobacteria?

Streptophyte algae and bryophytes were found to regularly associate with N_2-fixing cyanobacteria as well as rhizobia, suggesting that evolution of land plants, including the fern *Azolla*, happened after mechanisms important for both these interactions evolved (Knack et al., 2015). Recognition of specific cyanobacteria evolved with chloroplast endosymbiosis, but recognition competency may have been lost and subsequently acquired again with a novel mechanism by the common ancestor of vascular plants. Mechanisms employed by the seed plants *Gunnera* and *Blasia* to control development of *N. punctiforme* hormogonia may therefore also occur in *Azolla* ferns (Liaimer et al., 2015). Specific regulators and transporters for the exchange of nutrients with intracellular organelles such as the chloroplasts may not have been recruited for extracellular symbiosis, but instead have evolved from extracellular interactions such as those with arbuscular mycorrhiza, rhizobia or, *Nostoc* species. This would allow differential regulation of the transport processes. The AMT2 and PHT1 family transporters and the NIN transcription factor transcripts seen to accumulate in *Azolla* actively fixing N_2 could therefore play a role in the interaction with *N. azollae*. Active mobile elements inside *N. azollae* likely contributed to novel regulation of the ammonium transporter expression required for the symbiosis (Supplementary Figure 5) in addition to the metabolic adaptations such as the lacking glycolysis pathway reported earlier (Ran et al., 2010).

The most striking adaptation of the fern to *N. azollae* is the enclosed leaf-pocket organ with its specific pore, leaf-pocket hair cells, and extensive curving xylem-rich vasculature (Figure 3C; Supplementary Figure 2B,C). The pore structure is likely important for gas exchange; it is lined with specific teat cells possibly having a role in defense against invading organisms and water repulsion (Veys et al., 2002; Veys et al., 1999). The hair cells have been proposed to resemble transfer cells based on EM-cytology detection of dense cytoplasm adjacent to wall ingrowths and abundant ER with ribosomes and mitochondria, and therefore play a role in nutrient exchange (Duckett et al., 1975) consistent with their close contact with the phloem (Supplementary Figure 2B). Finally, the prominent and specific development of xylem cells surrounding the leaf pockets was consistent with a water conducing function; stark reinforcement of the tracheids furthermore revealed a structural role of the xylem cells not generally described in ferns (Figure 3C). Presence of VND6 in *Azolla* suggested that these NAC factors control secondary wall thickening in ferns as in seed plants (Lucas et al., 2013). Accumulation of transcripts related to vasculature, and nutrient transport under conditions requiring N_2 fixation (manually assembled homologs to AtGL22, 13 x, and AtXCP1, 9 x) suggested that vasculature functions are essential for nutrient supply to cyanobacteria, given that the leaf pockets are enclosed structures that do not communicate with surrounding medium. Hosting cyanobacteria in a leaf therefore may require exquisite control over water and nutrient supply as well as structural reinforcements and a specialized pore structure so as to hold the heavy pool with bacteria in the air whilst at the same time ensuring light and gas supply.

The evening loop of the clock is ancestral to ferns and seed plants

Differential diel transcript accumulation was found for 13 % of the fern genes. The types of genes which changed and the magnitude of changes observed here were comparable to those changing in diel cycles in spermatophytes (Blasing et al., 2005) and occurred in all diel patterns (Figure 6), hence the combined output of clock and diel signals such as temperature and light results in similar outputs among vascular plants. Among the transcription factors, the CONSTANS-like group of C2C2 transcription factors was enriched in both *A. filiculoides* and *A. thaliana* diurnal transcripts confirming their ancient role in diurnal gene expression. The best blast match to the C2C2 CONSTANS type transcription factor of *Chlamydomonas reinhardtii* Cre06.g278159, is also diurnal ((Zones et al., 2015) in Supplemental Table 16). Nitrogen fertilizer does not influence the diurnal gene expression in *A. thaliana* (Blasing et al., 2005) and did so only for a very limited number of transcripts in *A. filiculoides* (Figure 5B), which may have been a result of diel variability in N-supply in –N ferns rather than

less N availability per se (Figure 4). The similarity also points to an evolutionary origin of the diel patterns before the split of fern and seed plants. While no diel transcriptome analyses from mosses were available for comparison, the analyses of the single celled *C. reinhardtii* show both different and overlapping patterns. Gene expression in the alga was largely diel with ~80% of genes in diel rhythms (Zones et al., 2015) compared to only 13% in this study and 10-30% depending on criteria in *A. thaliana* (Blasing et al., 2005); the difference may reflect asymmetric coupling of tissue specific clocks in multicellular organisms (Endo et al., 2014).

A. filiculoides harboured the classical clock of seed plants (Figure 5C). Mosses have been shown to only contain a single loop of the clock comprising the morning but not the evening loop (Holm et al., 2010), while some algae contain various clock components including those of the evening loop (Noordally and Millar, 2014). In *Arabidopsis* vasculature transcripts were mostly controlled by the evening loop, whilst the mesophyll cell transcripts were mostly controlled by the morning loop; also the vasculature clock was shown to override that of mesophyll cells (Endo et al., 2014). In the ferns, whilst the evening loop was complete with TOC1, ZTL, GI and PRR3, some detected with multiple transcripts in the original assembly (Brouwer et al., 2014), the morning loop was less complex compared to *Arabidopsis* (Figure 5C). Possibly, therefore, the ancestor of vascular plants evolved an evening loop with a function in the vasculature.

Acknowledgements

We thank Arnold van Dijk for expert technical support at the Organic Geochemistry Department, Utrecht University. We further thank Dr. U Ganeteg at the Umeå Plant Science Centre, Department or Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden, for help with the interpretation of *Azolla* growth rates on amino acids. We acknowledge funding from the European Institute of Technology for the Climate KIC Pathfinder project AzoFaST and from the Deutsche Forschungsgemeinschaft (grants EXC 1028 and WE 2231/9-2).

Supplementary material

The supplementary information can be found in the online version of this article (open access): http://journal.frontiersin.org/article/10.3389/fpls.2017.00442/full

Chapter 4

Maintaining productive cultures of *Azolla*: methodology and the effect of CO_2 concentrations and species on biomass yield, chemical composition and suitability as feed

Paul Brouwer[1,2], Henriette Schluepmann[1], Klaas G.J. Nierop[2], Janneke Elderson[3], Willem de Visser[3], Ingrid van der Meer[4], Peter K. Bijl[2,5],Gert-Jan Reichart[2,6], Sjef Smeekens[1], Adrie van der Werf[3]

[1] Molecular Plant Physiology, Institute of Environmental Biology. Utrecht University, the Netherlands;
[2] Department of Earth Sciences, Faculty of Geosciences, Utrecht University, the Netherlands;
[3] Agrosystems research, Wageningen University & Research, the Netherlands;
[4] Bioscience, Wageningen University & Research; the Netherlands;
[5] Laboratory of Paleobotany & Palynology (LPP) Foundation, the Netherlands;
[6] Royal Netherlands Institute for Sea Research (NIOZ), the Netherlands.

Manuscript submitted

Abstract

Since available arable land is limited and nitrogen fertilizers pollute the environment, future production systems may be developed with crops that do not rely on them. Using growth modeling, we established a continuous harvest system with *Azolla* species yielding 90.0-97.2 kg dry weight (DW) ha^{-1} d^{-1}. Under ambient CO_2, N_2-fixation by the fern's cyanobacterial symbionts accounted for all nitrogen in the biomass. Proteins made up 17.6-20.8 %DW (4.9 x total nitrogen), depending on species and CO_2 treatment, and contained more essential amino acids than protein from soybean. Elevating the atmospheric CO_2 concentration (800ppm) boosted biomass production by 36-47%, without decreasing protein content. Choice of species and CO_2 concentrations further affected the biomass content of lipids (7.9-10.0 %DW) and (poly)phenols (2.1-6.9 %DW). Further research is needed to investigate the role of the high (poly)phenol content in limiting in the inclusion rate of *Azolla* in animal diets, whereas breeding and bio-refining strategies may be used to further improve biomass value.

Introduction

Growing worldwide demands for food, feed, energy and industrial products threatens natural ecosystems and global climate. Emissions by industry and agriculture increase concentrations of atmospheric greenhouse gasses (GHGs) and pollute surface waters (IPCC, 2013). Alternatives are sought to reduce climate impact and dependency on fossil resources, for example by fostering a sustainable bio-based economy. In a bio-based economy biomass not only provides food and feed, but also a major part of our materials, chemicals, and energy (Scott et al., 2007; Dornburg et al., 2010; Bos and Sanders, 2013; Sheldon, 2014). The main challenge for a bio-based economy is to produce sufficient biomass without encroaching on the arable land required for food and feed crops and without leaching nutrients into the environment. Novel agro-systems are therefore considered using alternative highly productive crops that use nutrients efficiently, contain valuable compounds and do not compete for arable lands. These include land based crops, such as *Miscanthus*, grown on marginal lands with minimal fertilizer input for the production of e.g. ethanol fuel or chemical building blocks (Brosse et al., 2012). Also algae are widely considered as novel bio-based agro-systems due to their high production potential, land-independent production and chemical composition suited for the production of fuel, chemicals and food (Brennan and Owende, 2010; Neveux et al., 2015). Many macrophytes share the advantageous traits of algae and demand less energy to harvest, but investigations have mainly focused on utilizing their fast growth for wastewater treatment by removing nutrients and/or heavy metals (Mohedano et al., 2012; Wagner, 1997; Dhote and Dixit, 2009). Although highly relevant, their potential for the industrialized production of food, fuels and/or chemicals has not yet been explored.

Here we consider the floating fresh-water fern *Azolla* for industrialized feedstock production. While species diversity was higher in episodes in the geologic past (Collinson et al., 2013), the genus *Azolla* today consists of 6 species worldwide (Wagner, 1997) that thrive in tropical to temperate regions of the world (Lumpkin and Plucknett, 1980; van Hove, 1989; van Hove and Lejeune, 2002). Growth of *Azolla* is sustained by nitrogen (N_2)-fixing cyanobacteria, *Nostoc azollae*, that reside in cavities of the fern leaves and supply ample nitrogen for both the endophyte and its host (Ran et al., 2010; Peters and Meeks, 1989). N_2 fixed is likely released as NH_4^+ in the leaf cavities and taken up by the ferns (Meeks et al., 1985). N_2 fixation rates vary greatly and values of 1-3.6 kg nitrogen (N) ha^{-1} d^{-1} have been reported with rates up to 1100 kg N ha^{-1} $year^{-1}$ (Watanabe, 1982; Hall et al., 1995). *Azolla* ferns may therefore be farmed commercially without nitrogen fertilizer, thereby reducing the production costs of the *Azolla* biomass and mitigating GHG emissions associated with the production and application of nitrogen fertilizer.

Azolla ferns exhibit extremely high growth rates when grown individually in an open canopy. Relative growth rates (RGR) of over 0.5 d^{-1}, or doubling times of less than 2 d have been reported, making *Azolla* one of the fastest growing plants on Earth (Maejima et al., 2001; cf. van der Werf et al., 1998). In batch experiments, elevated atmospheric CO_2 concentrations were reported to increase the speed of exponential growth of *A. filiculoides* (Speelman et al., 2009; Cheng et al., 2010; van Kempen et al., 2016). RGRs, however, cannot be extrapolated to productivity of a farming system since they represent exponential growth in a situation where surface area is not limiting and no inter- and intra-species competition for light and nutrients occurs. In a production system surface area is limiting, making that productivity per surface area needs to be maximized and not relative growth. Since considerable competition between individual *Azolla* plants occurs, environmental conditions may affect plant growth and chemical composition differently in a production system compared to exponential growth conditions.

Azolla biomass crude protein content was reported at 20-40%DW (% for the remainder of the chapter), making it useful as a protein feed (Buckingham et al., 1978; Reddy and D'Angelo, 1990; Cary and Weerts, 1992; Alalade and Iyayi, 2006; Leterme et al., 2010; Datta, 2011). *Azolla* biomass was further reported to contain 7.9-16 % lipids (Paoletti et al., 1987; Abou et al., 2007; Brouwer et al., 2015) and up to 5% (poly)phenolic compounds (Dewanji, 1993; Sánchez-Viveros et al., 2011; Forni et al., 2012; Deval et al., 2012). Lipids determine the caloric value of whole *Azolla* feed, whereas (poly)phenolic compounds, in particular condensed tannins, may decrease digestibility of whole *Azolla* feed (Nierop et al., 2011; Skrabanja et al., 2000; Bravo, 1998). However, the previous studies use *Azolla* grown exponentially in batch culture or obtained from its natural environment as subject for analysis. Analysis of protein, lipids and (poly)phenol content of biomass grown under productive conditions is needed to assess the true potential value of *Azolla* for the bio-based economy.

Here we aim to provide the methodologies and data required for developing *Azolla* farming systems. We investigate two *Azolla* species that differ in habitat and genetic makeup, i.e. *A. filiculoides* growing in temperate climates in Europe and *A. pinnata* growing in tropical climates in Asia. We first characterize growth parameters for these species and use modelling to devise protocols for continuous harvesting. We then use these protocols to estimate yield potential for each species and investigate the effect of elevating CO_2 (to 800ppm) on yield potential. Lastly, protein, amino acid, crude lipid and total (poly)phenol contents were determined for both species grown under ambient and elevated CO_2 and are discussed in relation to the application of biomass as feed.

Methods

Growth conditions

Azolla filiculoides and *A. pinnata* were obtained from the International Rice Research Institute under accession numbers 1052 and 534, respectively. Plants were grown in a growth room with the following conditions: average photosynthetic flux density 300 µmol m^{-2} s^{-1} for a 16 h light period. Light was provided by fluorescent tubes and incandescent lamps. Unless mentioned otherwise, day temperature was set at 23 °C and night temperature at 20 °C resulting in an average day-time water and air temperature (1 cm above the canopy) of 24.8 and 24.6 °C, respectively. Average night-time temperature of water and air were 23.3 and 21.1 °C, respectively.

Growth analysis and modelling

Plants were grown in aerated 30 L containers with a surface area of 1505 cm^2 containing a non-limiting nutrient solution, including nitrate: KNO_3 2 mM, $Ca(NO_3)_2.4H_20$ 0.34 mM, KH_2PO_4 0.42 mM, $MgSO_4.7H_2O$ 0.21 mM, $MnSO_4.H_2O$ 3.6 µM, H_3Bo_3 13.4 µM, $ZnSO_4.5H_2O$ 0.53 µM, $CuSO_4.5H_20$ 0.10 µM, Fe-EDTA 3.6 µM, $Na_2Mo_4{\cdot}2H_2O$ 0.16 µM. The solution was replaced every two weeks; pH varied between 6 and 7, and when necessary pH was adjusted using H_2SO_4. CO_2 concentration in the growth room was maintained at ambient concentrations (400-450 ppm). Each container was divided into 15 compartments using a Plexiglas construction. In each compartment two plants were inserted. Twice a week one randomly chosen compartment, per species and per container (n=6), was destructively harvested. Plants were oven-dried for 3 days at 30-35 °C and subsequently their dry weight was determined. In the analysis two growth phases were distinguished: exponential growth and linear growth. Plant growth was modelled by:

$$for\ t < t_{e\to l}: y_t = y_0 e^{RGR\cdot t}$$

$$for\ t_{e\to l} < t: y_t = AGR \cdot t + b$$

Where t = time (d), $t_{e\to l}$ = time point when the growth phase shifts from exponential to linear phase (d), y_t = standing crop at time t (kg ha^{-1}), y_0 = standing crop at the start of the experiment (kg ha^{-1}), RGR = relative growth rate (d^{-1}) ,b = standing crop at the start of the linear growth phase (kg ha^{-1}) and AGR = absolute growth rate (kg ha^{-1} d^{-1}). Standing crop under continuous harvesting was modelled by:

$$y_n = (y_{n-1} + AGR * p) * (1 - A_h)$$

Where y_n = standing crop after harvest n (kg ha^{-1}) p = harvest period (d) and A_h = harvest area (%). When the culture is in equilibrium, standing crop is a function of growth, harvest period and harvest area, i.e:

$$y_{n,eq} = y_{n-1} = AGR\left(\frac{p}{A_h} - p\right)$$

Harvest experiments

A. filiculoides and *A. pinnata* were grown indoors using 30 L containers similar to growth analysis experiment (n=5 for each species). The net surface area of each container was 1495 cm^2. Depending on the harvest scheme, 14% or 33% of the surface area was harvested daily or respectively twice a week. Medium replenishment and nutrient composition were optimized by calculating nutrient depletion:

$$N_t = N_0 - \frac{AGR \cdot t \cdot A}{V} + int(t/p_n) \cdot \frac{N_r}{V}$$

Where N_t = Nutrient concentration at time t (mg L^{-1}), N_0 = Nutrient concentration of the initial solution (mg L^{-1}), AGR = average growth rate (g DW m^{-2} d^{-1}), t = time (d), A = culture area (m^2), V = medium volume (m^3), p_n = period after which nutrients are added (d). N_r = Nutrients added (mg L^{-1}). The AGR was obtained from our growth analyses and biomass nutrient composition was based on literature data (Alalade and Iyayi, 2006). The growth medium contained 0.7 mM KNO_3, 0.1 mM$Ca(NO_3)_2.4H_2O$, 0.13 mM KH_2PO_4 and 0.1 mM $MgSO_4 \cdot 7H_2O$ of macronutrients and 4.7 µM Fe-EDTA, 2.2 µM $MnSO_4 \cdot H_2O$, 0.1 µM $Na_2Mo_4 \cdot 2H_2O$, 8.1 µM H_3B_3, 0.06 µM $CuSO_4 \cdot 5H_2O$ and 3.1 µM $ZnSO_4 \cdot 5H_2O$ of trace elements. The complete 30 L solution was replaced every two weeks, in between these periods, medium concentrate was supplied to 1x concentration every 3.5 d. CO_2 concentration in the growth room was maintained at ambient level (400-450 ppm). Harvested biomass was dried, weighed and stored until further analyses.

In the outdoor experiment, *A. filiculoides* was grown from June 23 until august 4 2014 in Wageningen, the Netherlands, in 6 m^2 outdoor ponds with each a volume of 3500 L (n=6). Every week 40% of the total surface area was harvested. Composition of the growth medium and nutrient replenishment were identical to the indoor harvest experiments.

Nitrogen and phosphorus uptake

Nitrogen and phosphorus concentration in the growth medium were analyzed during the harvest experiments at ambient CO_2, using a DR5000 Analyzer (Hach-Lange). Shortly after the solution was replaced aliquots of the solution were taken from each container and analyzed for nitrate-N and ortho-P concentrations. The

same procedure was followed three days later shortly before harvest of 33% of the surface area. After harvesting, water samples were taken and fresh nutrients were added as described above. This provided data for nutrient depletion over the course of three harvest intervals.

Protein and amino acid analysis

Plant material was ground after freezing with liquid nitrogen and subsequently freeze-dried. Total biomass nitrogen was analyzed in all growth replicates using the Dumas method (Silliker, Merieux Nutriscience, Netherlands). Total amino acid analyses were performed in triplicate and using three different methods: 1) acid hydrolysis for most of the amino acids, 2) oxidation followed by acid hydrolysis for cysteine and methionine, 3) alkaline hydrolysis for tryptophan. Amino acids in acid hydrolysates were determined using ion-exchange liquid chromatography with post-column Ninhydrin derivatisation and photometric detection (BiochromAAA) and Li-Citrate and Na-Citrate buffer system elution. Tryptophan was quantified by reversed phase HPLC. Amino acids were expressed as mg g^{-1} total amino acids (AA). Protein concentrations were estimated by correcting total AA for the gain of a water molecule during hydrolysis.

Crude lipid and total phenol analysis

Plant material of each batch was dried for three days at 30 °C, then ground (Retsch ZM200 grinder equipped with a 1 mm sieve). One gram of material was used for dry weight determination after further oven drying at 105°C for 24h followed by 1h cooling in an exsiccator with activated silica. The remainder of the material was Soxhlet extracted with a 9:1 dichloromethane (DCM) to methanol (MeOH) solution for 24 hours with a water-bath temperature of 80°C. The collected extracts were dried using a rotary evaporator and weighed to determine the amount of 'crude lipids'.

After Soxhlet extraction plant residues were collected and dried. Total phenol was determined in triplicate after extraction of 100 mg biomass with 10 mL of aqueous acetone (30:70, v/v) for 24 h at 20 °C using the Folin-Ciocalteu assay (Waterman and Mole, 1994) in two technical replicates. Tannic acid was used as the calibration standard. Literature data, used for comparison, presented in gallic acid equivalent was converted to tannic acid equivalent according to the relative reactivity with the Folin-Ciocalteu reagent (Everette et al., 2010).

Results and discussion

Growth analyses for modelling of a continuous harvest system

To model a continuous production system for *Azolla* species, we determined growth rates experimentally in a temperate species, *A. filiculoides*, and a subtropical species, *A. pinnata*. Ferns were cultured in excess and well aerated medium and their growth analyzed at 300 µmol m^2 s^{-1} in a 16h, 23°C, day-cycle reducing the temperature to 20°C at night (Figure 1a, b). Growth was exponential for the first 8 d (R^2=0.999 and R^2=1.000 respectively), then shifted to linear 8-12 d after the start of the experiment, when cultures reached a standing crop i.e. the total dry weight biomass present on the surface area, between 300-600 kg ha^{-1}. Growth remained linear at least until a standing crop of 1800 kg ha^{-1}.

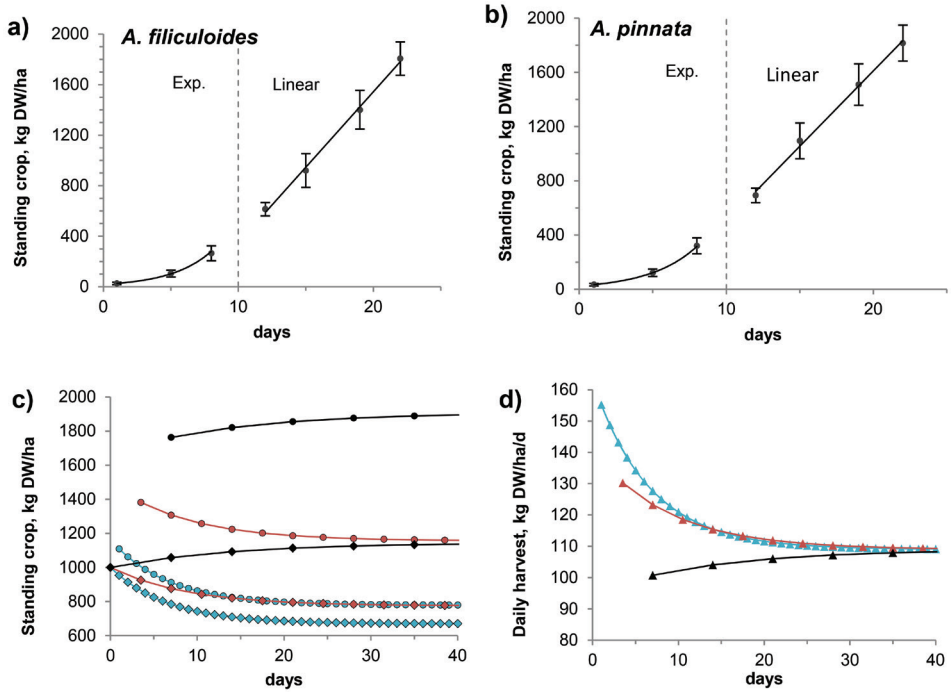


Figure 1. Growth analyses and derived model predictions. Standing crop as dry weight (DW) over time is given for growth analysis of *A. filicuoides* (a) and *A. pinnata* (b) cultures (n=6). Model predictions are given for standing crop before harvest (circle), standing cropafter harvest (diamond) (c) and daily biomass harvest (d) for three harvest schemes: 40% once a week (black) , 33% twice a week (red) and 14% daily (blue).

Relative growth rates (RGR) calculated over the first three time points of exponential growth were 0.337±0.078 d^{-1} for *A. filiculoides* and 0.317±0.034 d^{-1} for A.*pinnata* and were within the range 0.12 to 0.5 d^{-1} previously reported (van Kempen et al., 2013; Speelman et al., 2009; Sah et al., 1989; Cary and Weerts, 1992; Cheng et al., 2010; Debusk and Reddy, 1987; Maejima et al., 2001). RGRs obtained compared with those from the macrophytes *Salvinia natans* (0.28 d^{-1}) (Jampeetong and Brix, 2009) and *Eichhornia crassipes* (0.23 d^{-1}) (Xie et al., 2004), but were somewhat lower than those reported for duckweeds *Wolffiella hyaline* (0.52 d^{-1}) and *Lemna gibba* (0.50 d^{-1}) (Ziegler et al., 2015). The high RGRs reflect the ability of *Azolla* plants, and other floating macrophytes, to expand horizontally on the surface avoiding competition as long as the surface is not fully covered, opposed to the vertical growth of land plants for which competition for light starts shortly after emergence (Goudriaan and Monteith, 1990).

However, by definition, the highest growth per unit of surface area is not obtained in the exponential growth phase but in the linear growth phase, as the slope of standing crop versus time reaches is maximum (Figure 1). Hence when considering *Azolla* species, and other macrophytes, for their potential biomass yields, absolute growth rates (AGR) relative to production surface needed to be considered instead of RGRs. AGRs were at a maximum during the linear growth phase when increases of 119±14 kg dry weight (DW) ha^{-1} d^{-1} and 111±7 kg DW ha^{-1} d^{-1} were determined for *A. filiculoides* and *A. pinnata* respectively (Figure 1a, 1b). Growth rates of the two species did not differ significantly ($p<0.05$) from each other.

Using the absolute growth rates derived experimentally, production in a continuous harvest system was modelled varying input parameters: initial standing crop, harvest period and harvest area (Material and Methods). The model predicted that equilibrium conditions would be independent of initial standing crop. That is, although a high initial standing crop results in short term high yields, the long term cumulative harvest is unaltered. Our model indicated that, once in equilibrium, standing crop is a direct function of harvesting period and harvest area. Assuming an initial standing crop of 1000 kg ha^{-1} and an absolute growth of 109 kg DW ha^{-1} d^{-1}, the model was used to predict outcome in terms of standing crop before and after harvest in three harvesting schemes: 14% daily, 33% twice a week and 40% once a week (Figure 1c). When harvesting 14% daily, standing crop after harvest was close to the lower limit of the linear growth phase, whereas when harvesting 40% once a week the standing crop before harvest was around 1800 kg ha^{-1}. The daily harvest, however, was equal for all harvesting schemes once the standing crop had reached its equilibrium phase (Figure 1d). Therefore the model predicted that harvesting period and harvest area may be chosen freely without loss in yield as long as standing crop remains in the linear growth phase before and after harvest.

Harvest experiments to verify the growth model and determine productivity

Preliminary experiments showed that standing crop evolves over time as predicted by the model (Supplemental data A). However, productivity could not be determined since cultures were close to equilibrium only at the end of the experiments. Determining productivity would either require experiments to be performed over longer time periods such that initial perturbations become negligible, or require starting the experiment with an initial standing crop close to the expected equilibrium. The former approach was taken by maintaining cultures of *A. filiculoides* and *A. pinnata* for over 100 d, while harvesting 33% of the surface area twice a week. *A. filiculoides* cultures equilibrated at around 1000 kg ha^{-1} before harvest and 670 kg ha^{-1} after harvest (Figure 2a). *A. pinnata* cultures reached a lower equilibrium, i.e. 900 kg ha^{-1} before harvest and 600 kg ha^{-1} after harvest (Figure 2b). For both species,

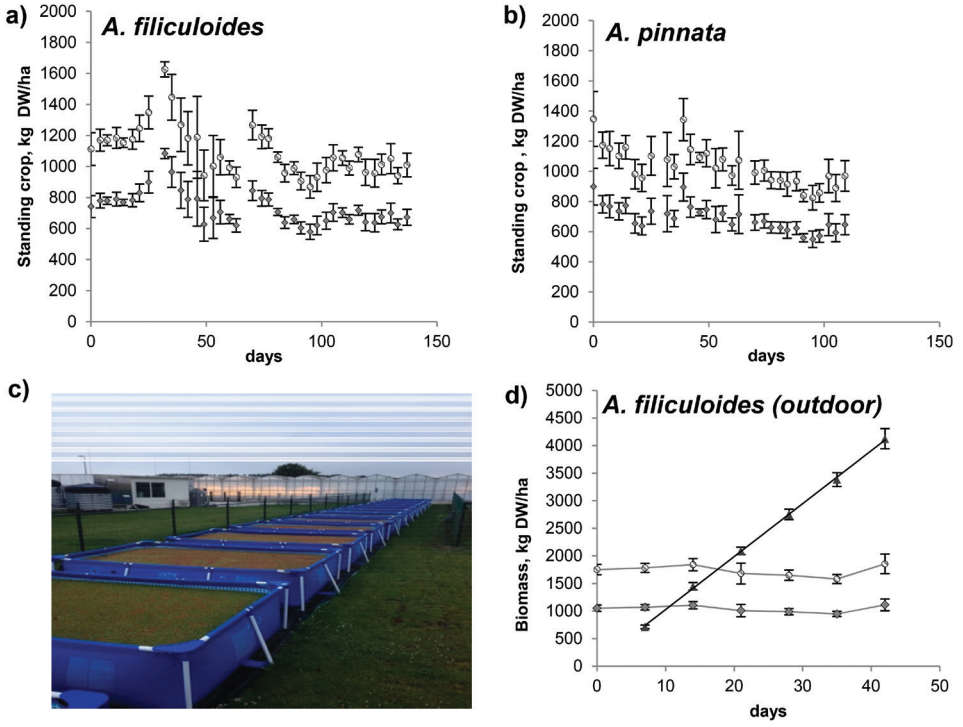


Figure 2. Continuous harvest experiments in controlled and outdoor conditions. Standing crop DW is given before (circle) and after (diamond) harvest of *A. filicuoides* (a) and *A. pinnata* (b) cultures (n=5).(c) The outdoor growth setup. (d) Outdoor standing crop before harvest, after harvest and cumulative harvest (triangle) (n=6).

equilibrium was reached later than expected due to perturbations in the standing crop caused by delaying two harvests at days 27 and 67. However, at all time points, standing crop remained within the linear phase defined in the growth analysis.

To test whether this approach was applicable to a larger scale and an outdoor climate, *A. filiculoides* collected from a local ditch was grown in 6 m^2 open ponds in June 2014 in the Netherlands (Figure 2c). Each week 40% of the pond's total surface area was harvested. To determine productivity, initial standing crop was chosen such that it was close to the expected equilibrium. Standing crop remained stable during the experiment, varying around 1740 kg ha^{-1} before harvest and 1040 kg ha^{-1} after harvest. Cumulative harvested biomass increased linearly over a period of 42 d (Figure 2d), resulting in a productivity of 97.1±4.1 kg DW ha^{-1} d^{-1} (n=6).

We therefore conclude that a continuous harvest system can be used to maintain cultures in the linear growth phase and to get predictable biomass yields, both under controlled conditions and outdoors in the Dutch temperate climate, at least for the period and light intensities tested.

Nutrient input during indoor growth experiments

Monitoring nitrogen and phosphate nutrients in the medium during indoor growth experiments revealed that phosphate was, as expected, depleted during growth and replenished when adding nutrients. The nitrate in the medium, however, was not depleted and as a consequence increased stepwise reaching 3.4 mM after 2 weeks (Figure 3b,d). The N_2-fixing symbiont *N. azollae* apparently provided all nitrogen incorporated in both species, and nitrate concentrations up to 3.4 mM did not inhibit N_2-fixation. This is consistent with 80-97 % of the N incorporated in *A. pinnata* biomass deriving from N_2 fixation at 1 mM NO_3 (Sah et al., 1989). At concentrations of nitrate and, especially, ammonium above 5 mM, however, N_2-fixation was inhibited in *Azolla* ferns (Ito and Watanabe, 1983). We conclude that *Azolla* ferns have particular potential for production systems without N-fertilizer, consistent with their long-established use as bio-fertilizers in rice paddy fields (Shi and Hall, 1988; Wagner, 1997).

Biomass yields in ambient and elevated CO_2

The cumulative harvested biomass from the indoor experiments is depicted in Figure 3a,c. Harvested biomass accumulated linearly over time, yielding 97.2±4.0 kg DW ha^{-1} d^{-1} and 90.0±6.2 kg DW^{-1} ha^{-1} d^{-1} for *A. filiculoides* and *A. pinnata* respectively. These yields were somewhat lower than expected from the growth analysis, which could be attributed to frequent handling of the plants and susceptibility to pathogens. Productivity of *A. filiculoides* grown outdoors in the Dutch summer climate, i.e. 97.1±4.1 kg DW ha^{-1} d^{-1} over 42 d, was close to the productivity predicted by

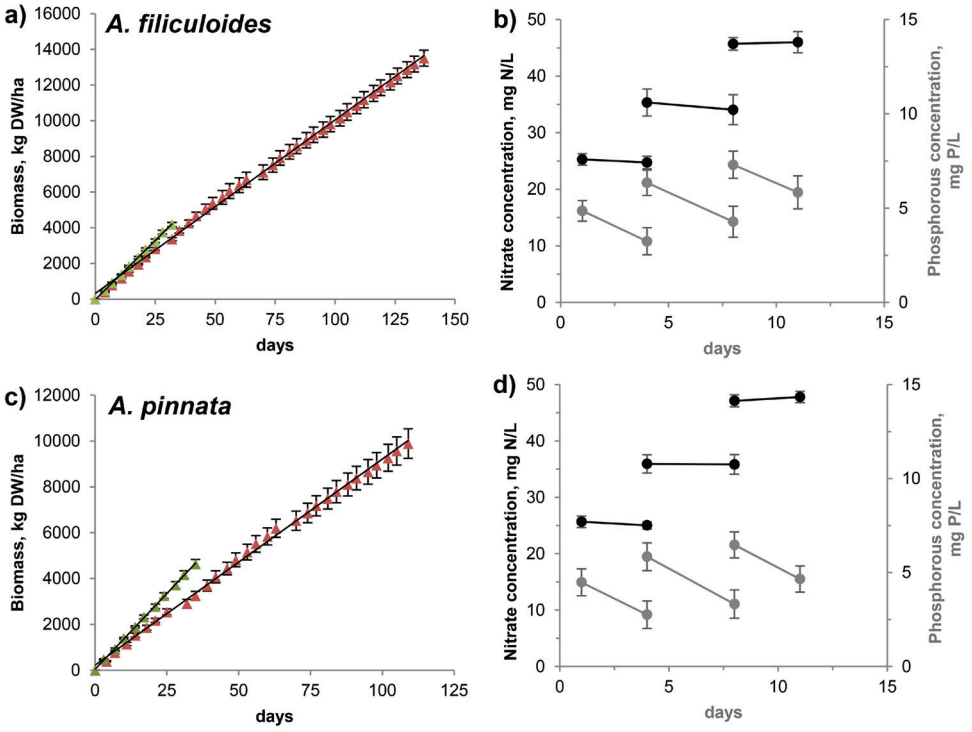


Figure 3. Biomass yields and nutrient use of *A. filiculoides* and *A. pinnata* in productive conditions. Cumulative harvest in dry weight (DW) for ambient (red) and elevated (800ppm) CO_2 conditions (green) (n=5) (a,b). Nitrate (black lines) and phosphate (grey lines) concentrations in the medium over a monitoring period of 11 d during the ambient CO_2 experiment (n=3) (c,d).

indoor experiments at ambient CO_2 concentrations.

The only previously reported continuous harvest pilot using *Azolla*, carried out over 154 d in outdoor ponds recycling chicken farm waste in Colombia, reached similar productivities (107 kg DW ha^{-1} d^{-1}) (Becerra et al., 1990). For two other floating macrophytes: *Salvinia molesta* and *Pistia stratioles*, average productivities reached 94 and 117 kg DW ha^{-1} d^{-1} over a period of 122 d in an outdoor facility in Saõ Paolo, Brazil (Henry-Silva et al., 2008); *Eichhornia crassipes* (water hyacinth) was yet more productive at 245 kg DW ha^{-1} d^{-1} when harvested over a one-year period in an outdoor system in Florida, USA (Reddy and D'Angelo, 1990). Compared to arable crops, the biomass productivity of *Azolla* was not exceptionally high: maximum biomass increases for *Zea mays* and *Brassica napus* of over 200 kg DW ha^{-1} d^{-1} have been observed (Rajcan and Tollenaar, 1999; Diepenbrock, 2000). In these crops, however, high growth rates are restricted to vegetative growth after seedling establishment. In contrast, *Azolla* ferns exhibit continuously fast vegetative growth as long

as they are regularly harvested and given sufficient nutrients.

The possibility that productivity could be increased further by adding CO_2 was tested by elevating CO_2 concentrations in the growth chamber to 800 ppm, while keeping all other environmental factors constant. Again, harvested biomass accumulated linearly over the 40 d tested; the productivities, however, increased to 132.4±4.0 kg DW ha^{-1} d^{-1} for *A. filiculoides* and 132.3±5.4 kg DW ha^{-1} d^{-1} for *A. pinnata* (Figure 3 a,b), corresponding to an increase of 36% and 47% for *A. filiculoides* and *A. pinnata* respectively. This yield increase was only slightly lower than observed for ferns growing exponentially in batch experiments (50%) (Speelman et al., 2009), hence adding CO_2 can also considerably boost yields of continious production systems.

Nitrogen & protein content in *Azolla* biomass

Given the need for sustainable protein crops, we measured total nitrogen and amino acid (AA) contents in the biomass of *A. filiculoides* and *A. pinnata*, obtained from the indoor harvest experiments. Nitrogen concentrations in *Azolla* biomass varied between 3.7-4.1% (Table 1). The sum of amino acids (AA) ranged from 20.8 to 24.4% and the AAs contained 82-88% of the total nitrogen in the biomass. Protein concentration was estimated at 17.6-20.8% (Table 1). The ratio of protein concentration over nitrogen concentration in AA ranged from 5.6 to 5.7, revealing that *Azolla* proteins contain more nitrogen than the reference protein bovine albumin serum (BSA), which has a ratio of 6.25. Since a considerable fraction of the nitrogen was not incorporated in amino acids, the overall ratio between protein and total nitrogen was calculated at 4.9-5.0. The simplified estimation of crude protein by multiplying

Figure 4. Concentration of nitrogen (n=6), total amino acids (AA) and protein (n=3) (as % of the dry weight) and derived ratio's for *A. filiculoides* and *A. pinnata* grown at ambient and elevated CO_2 (800 ppm).

	Ambient CO_2				Elevated CO_2			
	A. filiculoides		*A. pinnata*		*A. filiculoides*		*A. pinnata*	
Total nitrogen	4.0 ± 0.3	a	3.7 ± 0.3	b	4.1 ± 0.2	a	3.7 ± 0.1	b
Total AA	22.7 ± 2.1	a	20.5 ± 1.9	b	24.3 ± 1.5	a	21.2 ± 0.3	b
AA nitrogen	3.5 ± 0.5	a	3.1 ± 0.3	b	3.6 ± 0.2	a	3.2 ± 0.0	b
Non-AA nitrogen	0.4 ± 0.1	a	0.5 ± 0.1	a	0.6 ± 0.1	a	0.6 ± 0.1	a
Protein estimate	19.5 ± 1.8	a	17.6 ± 1.6	b	20.8 ± 1.3	a	18.2 ± 0.3	b
Ratio protein : AA Nitrogen	5.6 ± 0.9	a	5.7 ± 0.7	a	5.8 ± 0.5	a	5.7 ± 0.1	a
Ratio protein : total nitrogen	4.9 ± 0.1	a	4.9 ± 0.1	a	5.0 ± 0.1	a	4.9 ± 0.1	a

Superscript letters indicate possible significant differences($p<0.05$), based on a two-way ANOVA.

total nitrogen with 6.25 is therefore unsuitable for *Azolla* biomass, which has also been reported for leaves of Nigerian crops (3.24-5.17), duckweed (4.8), water hyacinth (4.6), and *Ulva lactuca* (4.6) (Magomya et al., 2014; Zhao et al., 2014; Bikker et al., 2016). In *Azolla*, the additional nitrogen may be present as inorganic nitrogen, such as ammonium in the leaf pockets (Canini et al., 1990), or as nitrogen-containing organic compounds other than amino acids.

Statistical analyses indicated that *A. filiculoides* contained significantly more nitrogen and protein ($P<0.05$) than *A. pinnata*. This significant difference also held for the total AA and AA-nitrogen, whilst the amount of non-AA nitrogen was similar across species and treatments. Increasing CO_2, although increasing growth rates, did not lead to a decrease in nitrogen, nor in protein content. Although not significant ($P>0.05$) the ferns grown under elevated CO_2 seem to have had a slightly increased protein and nitrogen content, in addition to their increased biomass productivity observed in Figure 3a,c. Since more biomass was produced under elevated CO_2 net nitrogen-assimilation increased as well. Increased nitrogen fixation under elevated CO_2 has been reported in the N_2-fixing legumes *Galactia elliottii* (Hungate et al., 1999) and *Medicago sativa* (Luscher et al., 2000), which was attributed to increased transport of photosynthates to the nodules (Soussana and Hartwig, 1996). Alternatively, increased CO_2 may have directly increased N_2-fixation by cyanobacteria in the fern leaf cavities. This is in line with previous studies showing nitrogenase activity being increased 2.6 fold when the cyanobacterium *Anabaena fertilissima* was grown at 6% CO_2 compared to ambient CO_2 (Chinnasamy et al., 2009) and 900 ppm CO_2 increasing nitrogen fixation rates 1.5-3 fold compared to ambient CO_2 in the marine cyanobacterium *Trichodesmium* (Levitan et al., 2007). Still, experiments carried out here did not rule out enhanced nitrate-uptake at elevated CO_2. Irrespective, the addition of CO_2 from industry, electricity or heat generation to an *Azolla* production system is expected to boost biomass productivity with no reduction of protein content.

Based on the productivities obtained in this study, annual yields of 32.8-35.5 t DW ha^{-1} $year^{-1}$ and 48.3 t DW ha^{-1} $year^{-1}$ can be expected from a fully controlled indoor production system using ambient CO_2 and 800 ppm CO_2 concentrations, respectively. Depending on the species, the annually harvested biomass contains 6.8-7.9 t protein ha^{-1} $year^{-1}$ for a controlled system at ambient CO_2 and 10.3-11.8 t protein ha^{-1} $year^{-1}$ at elevated CO_2. Given the productivity of *A. filiculoides* in the outdoor system, the stable monthly photosynthetically active radiation during the growth season in the Netherlands from May to September and the lack of response of *A. filiculoides* productivity to lower temperatures (Supplemental data A), minimum productivity outdoors of 15 t DW ha^{-1} $year^{-1}$ appears feasible in a temperate climate, corresponding to 3.6 t protein ha^{-1} $year^{-1}$. Protein yields obtained in tropical regions of up to 9 t protein ha^{-1} $year^{-1}$ (Becerra et al., 1990) were considerably higher than those predicted in a temperate climate and close to yields of 6.8-7.9 t protein ha^{-1} $year^{-1}$ obtained in

controlled conditions. Also thriving in the (sub)tropics, water hyacinth exhibited faster growth rates (54.93 t DW ha^{-1} $year^{-1}$), but contained only 12.7% protein, thereby reaching production rates of 7.0 t protein ha^{-1} $year^{-1}$, in a field test in China (Zhao et al., 2014). In the same experiment duckweed grew slower (26.5 t DW ha^{-1} $year^{-1}$), but had a high protein content of 25.8 %, thus yielding 6.8 t protein ha^{-1} $year^{-1}$ (Zhao et al., 2014). Depending on growth rates, marine algae, such as *Derbesia tenuissima*, have produced even higher protein yields i.e. 9.3 – 15.8 t protein ha^{-1} $year^{-1}$ (Neveux et al., 2015). The most established protein crop, soybean (*Glycine max*) can yield 2 t protein ha^{-1} $year^{-1}$ in subtropical regions, but only fixes 50% of the organic nitrogen exported from the field as seed (Salvagiotti et al., 2008) and legumes used for protein feed in Europe whilst yielding 15-21 t DW ha^{-1} $year^{-1}$ or 2.2 t N ha^{-1} $year^{-1}$ still require at least 150 kg N fertilizer ha^{-1} $year^{-1}$ (Anglade et al., 2015; Pugesgaard et al., 2015). The major advantage of *Azolla* over other floating plants and green algae, we conclude, is that it does not require N-fertilizer input. Its advantage over established crops derives from the combination of no N-input, no requirement for arable land and high yields.

Azolla biomass amino acid profile suits feed applications

In feeding trials *Azolla* could be included in diets at rates of 10% for poultry, 15% for sows and 25% for tilapia fish, but higher inclusion rates negatively affected animal weight and overall digestibility, meaning that commercial soybean meal or fishfeed could not be fully replaced by *Azolla* (Leterme et al., 2010; Abdel-Tawwab, 2008; Becerra et al., 1990; Alalade and Iyayi, 2006). The amino acid composition of *Azolla* was previously suggested to limit the inclusion rate (Alalade and Iyayi, 2006; Sanginga and van Hove, 1989). To determine whether the *Azolla* biomass from the productive system was suitable as a source of feed we analyzed individual amino acids for all species and treatments. Although concentrations of several individual amino acids differed between species and CO_2 treatment, differences were small (<10%) (Table 2). Noteworthy exceptions include concentrations of tryptophan and alanine, which increased by 18% and 17% when *A. filiculoides* was grown in elevated versus ambient CO_2 concentrations. Concentrations of cysteine and histidine in *A. filiculoides* decreased by respectively 14% and 11% in ferns grown at elevated versus ambient CO_2 (Table 2).

Compared to soybean meal, *Azolla* contained a higher proportion of all essential amino acids except for histidine (Figure 4a). Histidine is, however, not the limiting amino acid in many feeds (Figure 4b). For Nile tilapia and chicken the sulfur containing AA (SAA) were limiting (Figure 5). Levels of methionine were 68-72% of the level recommended for Nile tilapia and 91-97% of those for laying hens. Levels of combined methionine and cysteine equaled 92-95% of the recommended concen-

Table 1. Amino acid concentrations in biomass of *A. filiculoides* and *A. pinnata* grown at ambient and elevated (800 ppm) CO_2 concentrations (n=3).

	Amino acid concentration, mg g^{-1} total amino acids			
	Ambient CO_2		Elevated CO_2	
Amino acid	*A. filiculoides*	*A. pinnata*	*A. filiculoides*	*A. pinnata*
Arginine (Arg)	69.5 ± 5.2 ^{b}	66.9 ± 4.7 ^{b}	71.4 ± 7.3 ^{a}	69.7 ± 1.4 ^{b}
Histidine (His)	21.4 ± 1.9 ^{b}	18.9 ± 1.1 ^{b}	21.7 ± 1.7 ^{a}	21.6 ± 0.3 ^{b}
Isoleucine (Ile)	50 ± 4.2 ^{a}	51.8 ± 3.4 ^{a}	49.7 ± 5.6 ^{b}	50.2 ± 0.9 ^{a}
Leucine (Leu)	88.6 ± 8.1 ^{ab}	88.1 ± 5.5 ^{b}	90.3 ± 9.5 ^{a}	92.2 ± 1.5 ^{c}
Lysine (Lys)	57.6 ± 6.0 ^{b}	54.2 ± 3.3 ^{a}	55.2 ± 6.9 ^{a}	53.6 ± 0.5 ^{a}
Methionine (Met)	18.1 ± 0.6 ^{a}	19.2 ± 1.5 ^{a}	19.1 ± 1.2 ^{a}	19.5 ± 0.5 ^{a}
Phenylalanine (Phe)	59.2 ± 5.4 ^{b}	54.2 ± 2.8 ^{b}	59.4 ± 5.8 ^{a}	58.9 ± 0.8 ^{b}
Threonine (Thr)	50.8 ± 4.9 ^{b}	53.2 ± 3.3 ^{a}	49.1 ± 4.3 ^{c}	50.7 ± 0.8 ^{b}
Tryptophan (Trp)	15.5 ± 1.0 ^{a}	18.3 ± 0.9 ^{bc}	17.2 ± 1.0 ^{c}	16.6 ± 0.5 ^{ab}
Valine (Val)	60.3 ± 5.2 ^{a}	60.4 ± 3.5 ^{a}	60.3 ± 6.3 ^{a}	60.3 ± 0.8 ^{a}
Alanine (Ala)	63.7 ± 5.9 ^{a}	74 ± 5.5 ^{b}	69.3 ± 6.0 ^{d}	71.6 ± 1.4 ^{c}
Aspartic acid (Asp) + Asparagine (Asn)	104 ± 8.2 ^{b}	104 ± 6.5 ^{a}	99.6 ± 8.4 ^{b}	101 ± 1.8 ^{a}
Cysteine (Cys)	11.6 ± 0.3 ^{b}	9.98 ± 0.5 ^{b}	11.8 ± 0.6 ^{a}	11.2 ± 0.1 ^{b}
Glutamicacid (Glu) + Glutamine (Gln)	144 ± 18 ^{b}	137 ± 7.4 ^{a}	137 ± 9.8 ^{ab}	131 ± 2.0 ^{a}
Glycine (Gly)	54.9 ± 4.6 ^{a}	55.2 ± 3.1 ^{a}	55.2 ± 5.3 ^{a}	55.5 ± 0.6 ^{b}
Proline (Pro)	42.9 ± 4.2 ^{a}	42.6 ± 2.4 ^{b}	45.6 ± 5.3 ^{a}	45 ± 0.3 ^{b}
Serine (Ser)	49 ± 4.6 ^{a}	51.6 ± 3.4 ^{b}	50.3 ± 4.3 ^{c}	51.8 ± 0.8 ^{c}
Tyrosine (Tyr)	38.9 ± 3.9 ^{ab}	40.8 ± 3.0 ^{a}	38.3 ± 4.0 ^{c}	39.7 ± 1.1 ^{bc}

Superscript letters indicate possible significant differences ($p<0.05$), based on two-way ANOVA and a fischer LSD post hoc test.

tration for Nile Tilapia, 77-79% of that for laying hens and 82-85% of that for broilers. For dairy cows, lysine (79-85 %), methionine (91-97 %) and threonine (93-102%) were slightly limiting (Figure 4b). But when compared to soybean meal *Azolla* was superior because of its higher relative concentrations of methionine and threonine (Figure 5a). As Becerra et al. (1990) reported, an improved composition of essential amino acids, including methionine and cysteine, was consistently seen in pig diets of which 15-30% of soy protein was replaced by *Azolla* (Becerra et al., 1990). *Azolla* essential amino acid composition, furthermore, met all requirements for growing pigs (50-88 kg) and adult humans (Figure 5b) (NRC, 1998; FAO, 2013). We conclude that slight deficiencies in specific essential amino acids with respect to recommended levels for chicken, tilapia and dairy cows cannot explain the limited inclusion rate

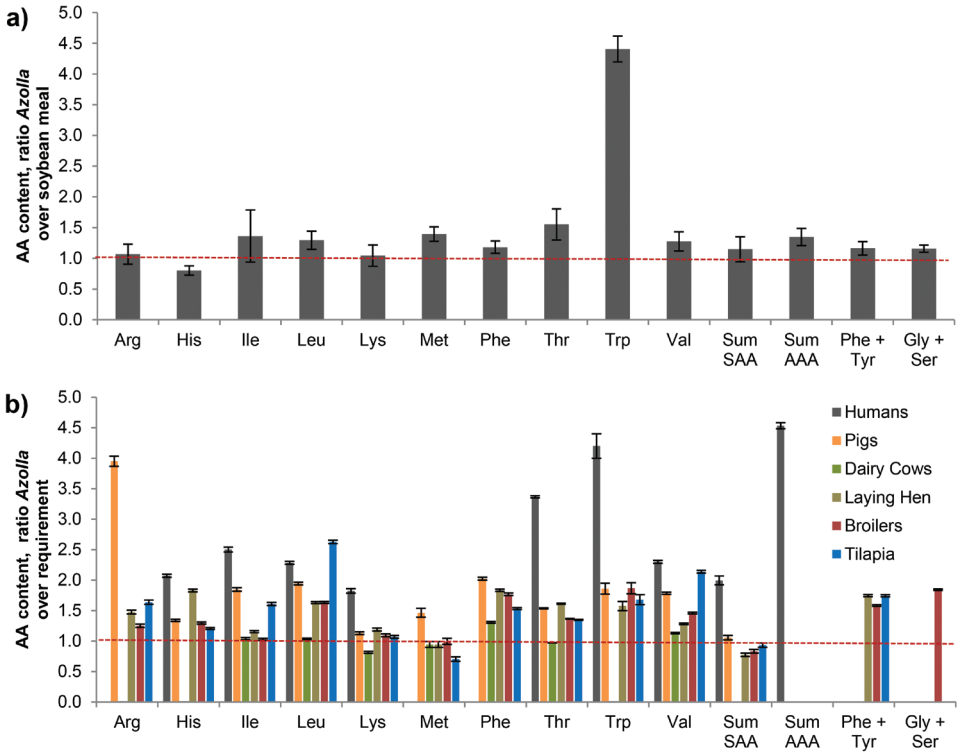


Figure 5. Essential amino acid (AA) concentrations in protein from *Azolla* compared to soybean (*Glycine max*) meal and their suitability for feed. (a) Depicts the ratios of the average AA concentration in *Azolla* species over the average of soybean meal, over three independent studies (Ravindran et al., 2014; Eweedah et al., 1997; Karr-Lilienthal et al., 2005). The sum of sulfur containing amino acids (Sum SAA), the sum of the aromatic amino acids (Sum AAA) and the combinations of Phe+Tyr and Glyc+Ser are also compared. In (b) ratios are shown of amino acid concentrations in *Azolla* over concentration-minima required in food for humans and feed for pigs, dairy cows, laying hens, broilers and tilapia (FAO, 2013; NRC, 1998; NRC, 2001; NRC, 1994; Santiago and Lovell, 1988) Standard deviations indicate variation between *Azolla* biomass of both species and treatments.

of *Azolla* biomass in feeding trials. To domesticate *Azolla* as a protein crop, therefore, other components affecting biomass digestibility need to be characterized, then either bred or extracted out.

Selecting species and growth conditions on (poly)phenol- and lipid-content to improve biomass value

Suggested causes for lowering digestibility of *Azolla* biomass also included high lignin and low energy content (Becerra et al., 1990; Leterme et al., 2010; Abdel-Tawwab, 2008; Alalade and Iyayi, 2006). Lignin is, however, not present in *Azolla*, which instead contains (poly)phenolic tannins (Nierop et al., 2011). These tannins likely decrease digestibility when binding to proteins upon digestion of the biomass (Skrabanja et al., 2000; Bravo, 1998). Total soluble (poly)phenols and crude lipids were therefore measured in biomass harvested from *A. filiculoides* and *A. pinnata* grown at ambient and elevated CO_2 (Figure 5). Soluble (poly)phenol levels, assayed as tannic acid (TA) equivalents per unit of dry weight, were high (20.9-69.1 mg g^{-1}; Figure 5a) compared to foodstuff, such as fruits (1.47-6.22 mg g^{-1}), beans (1.26-6.67 mg g^{-1}) and nuts (0.56-9.2 mg g^{-1}) (Nicoletto et al., 2013; Martínez et al., 2012; Wolosiak et al., 2011; Malencic et al., 2007; Kornsteiner et al., 2006; Everette et al., 2010). Soluble (poly)phenols were 3.3 times more abundant in *A. pinnata* reaching 6.91 % (TA-eq.) at ambient CO_2 (Figure 5a). Crude lipid yields represented roughly 10 % of the dry weight when ferns of either species were grown at ambient CO_2 (Figure 5b). Hence, besides higher yield and protein content, *A. filiculoides* contained less (poly)phenols than *A. pinnata* when grown at ambient CO_2, making it a more suitable protein feed.

In *A. pinnata*, elevated CO_2 altered neither lipids nor soluble (poly)phenols significantly. In *A. filiculoides*, however, elevated CO_2 decreased (0.78 fold) crude lipid

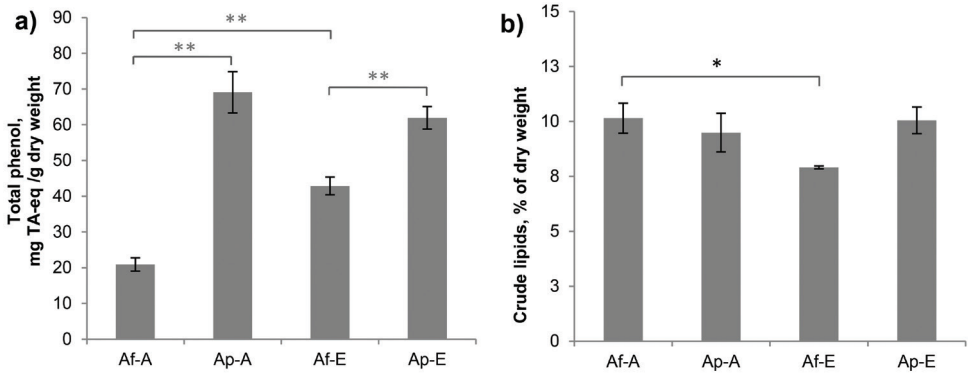


Figure 6. The effect of *Azolla* species and CO_2 concentration on total phenol and crude lipid content. Total phenol is expressed as mg Tannic Acid (TA) equivalent (a) and Lipids are measured by weight (b) in dry biomass of *A. filiculoides* (Af) and *A. pinnata* (Ap) grown under ambient (Ap-A, Af-A) and elevated (800 ppm) CO_2 (Ap-E, Af-E) (n=3).

and increased (2.1 fold) total soluble (poly)phenols (Figure 5a,b). Although adding CO_2 increased protein yield in *A. filiculoides*, it reduced cysteine and lipid content and increased total (poly)phenol content, which may reduce caloric value and digestibility of the *Azolla* biomass. More research is needed to confirm the effects of polyphenol and lipid content on biomass digestibility. This knowledge may then be used to increase biomass value further, by using breeding and selection to produce *Azolla* varieties with more favorable (chemical) properties, which is now possible due to recent advances in cryo-preservation and in vitro fertilization of *Azolla* spores (Brouwer et al., 2014). Alternatively, or additionally, multiple products may be refined from *Azolla* biomass thereby increasing the value of the production chain. Protein extraction would ideally yield a more concentrated product free of any anti-nutritional factors. Besides extracting proteins, lipids may be isolated and converted into high quality biodiesel (Brouwer et al., 2015). Various (poly)phenolic compounds are of interest as food additives due to their health-promoting effects (Bravo, 1998; Yokohira et al., 2008) or are useful as chemical building blocks for various aromatic compounds, which are currently synthetized from the benzene, toluene and xylene fraction of fossil oil (Bruijnincx and Weckhuysen, 2013). In particular, caffeic acid from *Azolla* (Ishikura, 1982) may be used for synthesis of bioplastics resistant to high temperatures (Chauzar et al., 2012). Efficient and economical extraction processes, however, need to be developed for *Azolla* to enable such bio-refining in the future.

Conclusions

Production systems using continuous harvesting allow to obtain high biomass yields for *A. filiculoides* and *A. pinnata*, without using nitrogen fertilizer. The high levels of protein and favorable amino acid content with respect to soybean suit application of the biomass as feed. However the high content of (poly)phenols (up to 7% of the dry weight) is likely responsible for the reported limits in the inclusion rate of *Azolla* in the diet of animals. In terms of biomass yield, protein content and (poly) phenol content, *A. filiculoides* is a more suitable species for the use as animal feed than *A. pinnata*. Our results show a clear case for combining *Azolla* cultivation with CO_2 waste streams, as elevating CO_2 concentrations to 800ppm boosted productivity to 48.3 t DW ha^{-1} $year^{-1}$ without loss in protein content. Increasing CO_2 concentrations did increase (poly)phenol content and decreased lipid content. The effects of (poly)phenols on applicability of *Azolla* as protein feed need further investigation. Breeding and selection for more favorable *Azolla* strains as well as bio-refining to produce multiple value-adding products may be used to further enhance the value of an *Azolla* production chain.

Acknowledgement

We thank Niño Paul Meynard Banayo, Stephan Haefele and Inez Hortense Slamet-Loedin from the International Rice Research Institute (IRRI, Philippines) for providing us with IRRI *Azolla* accessions. We further acknowledge funding from Climate KIC for the Pathfinder project AzoFaSt and financial support by the LPP foundation.

Supplemental data

The supplemental data is available on request at p-brouwer@live.nl and will be made available online upon publication.

Chapter 5

Lipid yield and composition of *Azolla filiculoides* and the implications for biodiesel production

Paul Brouwer[1,2], Adrie van der Werf[3], Henriette Schluepmann[1], Gert-Jan Reichart[2] and Klaas G. J. Nierop[2].

[1] Molecular Plant Physiology, Institute of Environmental Biology, Utrecht University,the Netherlands;
[2] Department of Earth Sciences, Faculty of Geosciences, Utrecht University,the Netherlands;
[3] Agrosystems research, Wageningen University & Research, the Netherlands

Adapted from: Brouwer, P., van der Werf, A., Schluepmann, H., Reichart, G-J., Nierop, K.G.J. (2016). Lipid Yield and Composition of Azolla filiculoides and the Implications for Biodiesel Production. Bioenergy Research, 9 (1), pp. 369-377.

Abstract

The aquatic fern *Azolla* is one of the fastest-growing nitrogen-fixing plants on Earth and therefore considered as a potential source of biomass for bioenergy production. The lipid fraction from *Azolla filiculoides* was analyzed to investigate whether it suited biodiesel production. Since the productivity of *Azolla* is further increased at higher CO_2 concentrations, *A. filiculoides* biomass was produced at 800 ppm CO_2 mimicking a cultivation system utilizing CO_2 waste from industry. The harvested biomass contained 7.92±0.14 % dry weight (dw) crude lipids. Drying conditions did not significantly affect lipid composition or yields, indicating that drying conditions may be energetically optimized without the risk of product loss. Total lipid extracts contained 4.2±0.38 % free fatty acids. Of the crude lipid fraction, 41±13 % consisted of fatty acids that were converted into fatty acid methyl esters upon saponification in methanol. Unique mid-chain (di)hydroxy compounds constituted 7.2±2.8 % of the crude lipids. Based on the fatty acid profile, it was estimated that *Azolla* biodiesel meets requirements set by the EN14214 standard on fuel density, cetane number, and iodine value. The cold filter plugging point (CFPP), however, is expected to be too high due to relatively high concentrations of lignoceric acid and the presence of the mid-chain (di)hydroxy compounds. To produce high-quality biodiesel from *Azolla* lipids a fractionation step will be required removing these compounds. As an advantage, the long-chain alcohols and (di)hydroxy fatty acids obtained after fractionation may provide a valuable secondary product stream with applications to chemical industry and nutrition.

Introduction

Since fossil resources are finite and their large-scale use as fuels has a negative impact on global climate, alternative resources for energy and materials are needed. It is expected that, with the depletion of fossil resources, plants will need to provide an important proportion of our requirements for fuels and chemicals (EPSO, 2005).

However, not all plant-based alternatives are equally sustainable. Current biodiesel feedstocks, i.e., soybean, rapeseed, palm, and sunflower oil, suffer from low land use efficiency, high input requirement (de Vries et al., 2010), and competition for arable land with food production. Hence, to produce fuels and chemicals in a more sustainable way, novel crops are sought that are lowinput, land-efficient, and highly productive. The floating water fern *Azolla* is one such potential novel crop. It can be cultivated in closed systems on non-arable land or in natural occurring freshwater basins. It is known for its high growth rates: In a 154-day outdoor growth experiment, Becerra et al. (1990) estimated the annual productivity of *Azolla filiculoides* to be 39 t dry weight per hectare. Moreover, *Azolla* reaches high growth rates without inorganic nitrogen in its growth medium (Sah et al., 1989), due to its symbiotic relationship with nitrogen-fixing cyanobacteria (*Nostoc azollae*). Active CO_2 supply may increase the productivity of *Azolla* even further by 53% at 760 ppm CO_2 compared to that at 340 ppm (Speelman et al., 2009a).

Until now, *Azolla* has been applied as biofertilizer in rice paddy fields, investigated as a biosorbent of heavy metals (Wagner, 1997), and evaluated as raw animal feed (Abdel-Tawwab, 2008; Alalade and Iyayi, 2006; Leterme et al., 2010) or protein feed (Fasakin, 1999). For biofuel or bioenergy production, so far, anaerobic fermentation to produce biogas (Abassi et al., 1990; Jain et al., 1992) and pyrolysis of whole *Azolla* biomass have been suggested (Muradov et al., 2014).

In the present study, we focus on the lipid fraction of *Azolla* as a source of biodiesel. The lipid fraction contains ester-bound fatty acids (FAs) which may be converted into biodiesel using standard alkaline trans-methylation. Biodiesel quality can be predicted from the FA profile of the feedstock (Moser and Vaughn, 2012). Therefore, various researchers have developed, and used, numerical models to estimate biodiesel quality indicators including density, cetane number, iodine value, and cold filter plugging point (CFPP), using FA profiles (Ramos et al., 2009; Lapuerta et al., 2009; Lapuerta et al., 2010; Arias-Peñaranda et al., 2013; Yadav et al., 2012; Hussain et al., 2014). To assess the quality of biodiesel produced from *Azolla* using these numerical models, quantitative information on its lipid composition is required.

Existing reports on *Azolla* lipids either were incomplete (Paoletti et al., 1987) or did not focus on compound quantification (Speelman et al., 2009b). Therefore, we aimed to analyze all major lipids of *A. filiculoides*. To mimic the situation in a production system utilizing CO_2 waste, biomass was obtained from productive cultures

grown at two times the atmospheric CO_2 concentration, i.e., 800 ppm CO_2, comparable to Speelman et al. (2009a).

Quantification of compounds derived from dried biomass may be affected by the drying condition (Deinum and Maassen, 1994). Since *Azolla* is an aquatic crop with a high water content, drying will be a likely step in any biomass processing system and needs to be performed with the least energy costs, without affecting extractability and composition of lipids. Therefore, this study first investigates the lipid yields and composition of *A. filiculoides* in relation to drying conditions. Second, we evaluate the implications for biodiesel production, using numerical models to estimate biodiesel quality.

Materials and methods

Biomass production and pre-treatment

A. filiculoides was obtained from the International Rice Research Institute (IRRI) under accession number 1052. Details on the original site of collection are provided in the IRRI germplasm collection catalogue (Watanabe et al., 1992). Plants were cultivated in a growth chamber at Wageningen University. Growth in continuous culture was performed, in five separate 30L containers, at a CO_2 concentration of 800 ppm, a day/night temperature of 23/20 °C, a 16-h day length, and a photosynthetic flux density of 325 µmol m^{-2} s^{-1}. Fresh nutrients were supplied every 3 days with the following concentrations: 0.7 mM KNO_3, 0.1 mM $Ca(NO_3)_2$, 0.13 mM KH_2PO_4, and 0.1 mM $MgSO_4 \cdot 7H_2O$ of macronutrients and 4.7 µM Fe-EDTA, 2.2 µM $MnSO_4 \cdot H_2O$, 0.1 µM $Na_2Mo_4 \cdot 2H_2O$, 8.1 µM H_3B_3, 0.06 µM $CuSO_4 \cdot 5H_2O$, and 3.1 µM $ZnSO_4 \cdot 5H_2O$ of trace elements.

The entire medium was refreshed every 2 weeks. Every 3 or 4 days, one third of the surface area of each container was harvested. From all five containers, a batch of fresh *Azolla* biomass was harvested for the analysis. Each batch was divided into three aliquots which were dried for 3 days at 30 °C, dried for 1 day at 65 °C, or freeze-dried, respectively. After drying, the biomass was ground using a Retsch ZM200 grinder equipped with a 1-mm sieve.

To determine the dry weight of each aliquot after drying and grinding, 1 g of biomass was pre-weighed in alumina cups, which were placed into an oven set at 105 °C. They were left to dry for 24 h after which they were transferred into an exsiccator. After cooling for 1 h inside the exsiccator, the alumina cups were taken out and immediately weighed using a Bosch SAE 200 scale with a sensitivity of 0.1 mg.

Total lipid extraction and saponification

The freeze-dried and ground *Azolla* biomass was Soxhlet-extracted with a 7.5:1 dichloromethane (DCM) to methanol (MeOH) solution for 24 h. The collected extracts were dried using a rotary evaporator and weighed to determine the amount of crude lipids. Subsequently, these total lipid extracts (TLEs) were re-dissolved in 9:1 DCM/MeOH and ran over a Na_2SO_4 column to remove any traces of water. Afterward, the TLEs were dried under continuous nitrogen flow.

For TLE analysis, an aliquot was methylated using diazomethane at room temperature, which was subsequently purified over a silica (60-Å pore size) column. Next, the extract was silylated by bis(trimethylsilyl)trifluoroacetamide (BSTFA) in pyridine at 60 °C for 20 min. Five to ten micrograms of squalane was added as an internal standard.

A selection of lipid extracts was saponified in 2 N KOH in MeOH (96 %) at 70 °C for 2 h. After this, the solution was acidified to pH<5.5 using 2 N HCl. One milliliter of ultrapure water (from Milli-Q water purifier) was added, and the solution was three times extracted by 1 ml of DCM. The DCM fractions were combined and dried under continuous nitrogen flow. From these saponified extracts (SEs), traces of water were removed using a Na_2SO_4 column, after which the same derivation steps were conducted as for the TLEs.

Lastly, TLEs and SEs were run on a Hewlett Packard Gas Chromatograph–Flame Ionization Detector (GC-FID) and Thermo-Finnigan TraceGC ultra-Trace DSQ Gas Chromatograph–Mass Spectrometer (GC-MS). Both machines were equipped with a Varian CP-Sil5CB column (30 m, 0.32 mm i.d., and film thickness of 0.10 µm). The GC was operated at constant pressure of 100 kPa, whereas the GC-MS was operated at constant flow of 1.6 ml min^{-1}. Temperature profiles were identical for each analysis, i.e., 70 °C during on-column injection, increasing to 130 °C at a rate of 20 °C min^{-1}, then increasing to 320 °C at a rate of 4 °C min^{-1}, and finally, an isothermal hold for 20 min. The MS operated within a scanning range of m/z 50–800.

Identification of the compounds was carried out by their mass spectra using a NIST library or by interpretation of the spectra, by their retention times, and/or by comparison with literature data. Quantification was conducted by integration of peak areas of each individual peak/compound relative to that of the internal standard. In the case of co-eluting peaks, peak separation was achieved by comparing relative intensities of specific fragment ions and using their ratio to quantify each of the compounds identified. Dotriacontane was added as a tracer before Soxhlet extraction to analyze losses during the extraction and subsequent workup procedure and allow to correct compound concentrations obtained from GC and GC-MS analysis. In total, 11 TLEs were analyzed, 4 of which were extracted from biomass batches dried at 30 °C, 4 extracted from batches dried at 65 °C, and 3 extracted from freeze-dried

batches. Nine SEs were analyzed, three for each drying condition.

Analysis of variance

To determine whether drying conditions affect lipid yields and composition, a one-way analysis of variance (ANOVA) was performed using SPSS 20 statistics software. As we were interested in the specific effect of predefined drying conditions, i.e., freeze-dried, 3 days at 30 °C, and 1 day at 65 °C, we used a straightforward fixed effects model:

$$\gamma_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

where y_{ij} = lipid or compound concentration (%), μ = common value, α_i = treatment effect, and ε_{ij} = residual error due to variation between biomass batches and variations introduced during analysis. For determining the effect of drying conditions on crude lipid yield, all five replicates were used (j=5), whereas for determining the effect of drying conditions on individual compounds, a random subset of at least three replicates were used (j=3). For individual compounds, computing was done on the basis of normalized concentrations to eliminate the effect of the variance in total compound detection. In cases that the F test provided a value with $p<0.05$, Tukey's test was used to determine which drying condition showed significant deviation.

Estimation of biodiesel properties

Fatty acid methyl ester (FAME) composition was used to estimate biodiesel properties, i.e., density, cetane number, iodine value, and cold filter plugging point (CFPP) of *Azolla* biodiesel. To estimate these properties, numerical models are available in the literature. The density (ρ) was calculated by Eq. (1) by assuming a perfect mixture of FAMEs. The densities of pure FAMEs were estimated using Eq. (2) (Lapuerta et al., 2010):

$$\rho_B = \sum X_i \rho_i \tag{1}$$

$$\rho_i = 851.471 + \frac{250.718db + 280.899 - 92.180(m-1)}{1.214 + n} \tag{2}$$

where X_i = relative content of the FAME (%), db = number of double bonds in the FAME, m = number of carbon atoms in the alcohol used for esterification, and n = number of carbon atoms in the original fatty acid. Similarly, the cetane number (CN) is calculated in Eq. (3) by assuming a perfect mixture of FAMEs, and cetane numbers of pure FAMEs are estimated using the formula in Eq. (4) (Lapuerta et al., 2009):

$$CN_{\mathrm{B}} = \sum X_{\mathrm{i}} CN_{\mathrm{i}} \tag{3}$$

$$CN_{\mathrm{i}} = -21.157 + \left(7.965 - 1.785db + 0.235db^2\right)n - 0.099n^2 \tag{4}$$

where CN_B = cetane number of the blend and CN_i = cetane number of pure FAME. Iodine value and cold filter plugging point were estimated using formulas derived by Ramos et al. (2009). Equation (5) provides the iodine value as a function of the degree of unsaturation (DU), whereas DU is given by Eq. (6) (Ramos et al., 2009):

$$Iodine\ value = 87.396 \cdot DU + 1.6691 \tag{5}$$

$$DU = (monosaturated\ Cn:1,\ wt.\%) + 2 \cdot (polysaturated\ Cn{:}{>}1,\ wt.\%) \tag{6}$$

The $CFPP$ was estimated with Eq. (7) using the long chain saturation factor $LCSF(A)$, as defined in Eq. (8) (Ramos et al., 2009):

$$CFPP = 8.9243 \cdot LCSF(A) - 19.325 \tag{7}$$

$$LCSF(A) = MP_{C18} \cdot C18(wt.\%) + MP_{C20} \cdot C20(wt.\%) + MP_{C22} \cdot C22(wt.\%) + MP_{C24} \cdot C24(wt.\%) \tag{8}$$

where MP_i = melting point of the saturated FAME (in degrees Celsius).

For comparison with other biodiesel feedstocks, cetane number, iodine value, and CFPP of rapeseed, soybean, and palm were obtained from Ramos et al. (2009), whereas reported FAME distributions were used to calculate biodiesel density. For the green algae *Scenedesmus incrassatulus*, all biodiesel parameters were calculated using the FAME distribution reported by Arias-Peñaranda et al. (2013). European standards for biodiesel properties were drawn from EN 14214 (CEN, 2008).

Results and discussion

Lipid yield, general composition, and effects of drying conditions

Cultures of *A. filiculoides* grown at 800 ppm CO_2 had an average productivity of 13.24 g m^{-2} d^{-1} over a period of 40 days. The harvested biomass was either dried for 3 days at 30 °C, dried for 1 day at 65 °C, or freeze-dried, before lipid extraction. Table 1 provides the crude lipid yields from biomass dried at each of these conditions. Small differences were observed between drying conditions, but differences between groups were not found to be significant by ANOVA as shown in Table 2. The average crude lipid yield of all extractions equals 7.92±0.14% of the dry weight (dw).

Online Resource 1 provides the average compound concentrations of the TLEs and SEs. Based on the SEs, 4.6±1.4% dw of lipids could be accounted for using quantification on the GC/GC-MS, whereas only 0.84±0.40% dw of lipids are detected in TLEs. Differences between the TLEs and the SEs are summarized in Figure 1 for all major (groups of) compounds in *Azolla*. Saponification in methanol converted the ester-bound fatty acids into fatty acid methyl esters (FAMEs). Similarly, phytol was derived from chlorophyll upon saponification. The amount of sterols did not change significantly. Saponification further resulted in the detection of higher amounts of mid-chain (di)hydroxy compounds, due to the hydrolysis of esters of C26-C36 (di)hydroxy fatty acids ((di)OH FAs).

Table 1. Crude lipid yield of *Azolla* batches exposed to different drying conditions and results of analysis of variance

Drying condition	N	Crude lipid yield (% dw)
30 °C, 3 days	5	7.75±0.49
65 °C, 1 day	5	7.94±0.22
freeze drying, 3 days	5	8.16±0.30
Average	5	7.92±0.14

N denotes the number of replicates. Crude lipid yield is given as average ± standard deviation

Variations in the concentrations of individual compounds can mainly be explained by the variation in total compound detection, as indicated by the much lower variations in normalized concentrations. The ANOVA test only showed a significant ($p<0.05$) difference between drying conditions for 9,10-dihydroxynonacosane (C29 20,21(ω9,ω10)diol) in TLEs and phytol in SEs, as indicated with an asterisk in Online Resource 1. All results of the ANOVA test on TLEs and SEs are provided in Online Resource 2. For SEs, most phytol was detected in batches dried at 65 °C and least in freeze-dried batches. This may indicate that extraction of chlorophyll and derivatives is less efficient from freeze-dried material. Overall, the effect of drying conditions on the lipid yield and composition appears to be of minor importance within the scope of this study. Therefore, it was decided not to distinguish between drying conditions for the further analysis.

Table 2. Results of analysis of variance (ANOVA) in crude lipid yield

	Sum of squares	Degrees of freedom	Mean square	F	p-value
Between groups	$4.53 \cdot 10^{-5}$	2	$2.27 \cdot 10^{-5}$	1.936	·187
Within groups	$1.40 \cdot 10^{-4}$	12	$1.17 \cdot 10^{-5}$		
Total	$1.86 \cdot 10^{-4}$	14			

F provides the result of the Fisher test. The p value gives the likelihood of the found F value for the given degrees of freedom

Fatty acid distribution

The FAs were estimated to constitute 3.2±1.0% dw of *Azolla* biomass and 41±13% of the crude lipid fraction. The average FA profiles of TLEs and SEs of *Azolla* are given in Figure 2. Palmitic acid (C16:0 FA), linoleic acid (C18:2 FA), and linolenic acid (C18:3 FA) are by far the dominating FAs. Furthermore, the concentration of lignoceric acid (C24:0 FA) in *Azolla* is relatively high compared to that in other plant species, such as soy, palm, and rapeseed (Ramos et al., 2009).

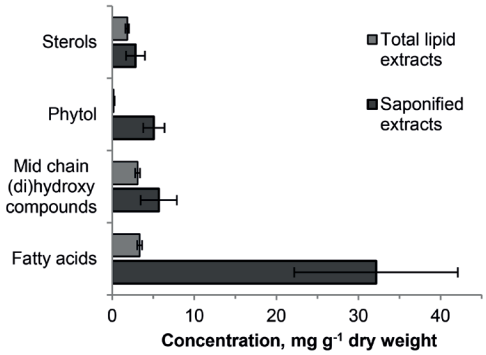


Figure 1. Concentrations of major (groups of) compounds. Error bars indicate standard deviation (N=5, N=4 for TLEs and SEs, respectively).

In TLEs, free fatty acids (FFAs) amounted to 4.2±0.38% of the crude lipid fraction. Many of the FAs identified in SEs could not be detected in TLEs. These include pentadecylic acid (C15:0 FA), palmitoleic acid (C16:1 FA), hexadecatrienoic acid (C16:3), arachidic acid (C20:0 FA), eicosatrienoic acid (C20:3 FA), nervonic acid (C24:1 FA), and cerotic acid (C26:0 FA).

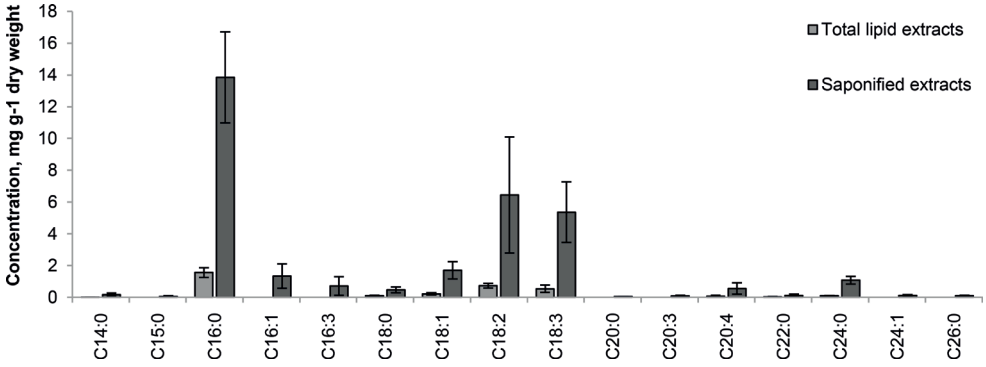


Figure 2. Fatty acid concentrations in TLEs and SEs. Error bars indicate the standard deviation, N=5 for TLEs and N=4 for SEs.

Mid-chain (di)hydroxy compounds

The mid-chain (di)hydroxy compounds in *Azolla* are characterized by their long chain, i.e., 26 to 52 carbon atoms, and the presence of at least one mid-chain hy-

droxy group located at either the ω20 or ω9 position on the carbon chain. This allows the methyl and trimethylsilyl (TMS) derivatives to be identified using typical fragment ions at m/z 369 and m/z 215, respectively (Speelman et al., 2009b). As illustrated in Figure 3, they can be further subdivided in (di)OH alkanes and (di)OH alcohols, and (di)OH FAs and (di)OH wax esters. For each of these classes, examples of structural formulas are given in Figure 3a–h. The (di)OH alkanes consist of alkanols and diols with solely mid-chain hydroxy groups and thus have an alkane backbone. In contrast, the (di)OH alcohols are diols and triols, both have a hydroxy group at the 1-position as mid-chain hydroxy group(s) and are therefore termed (di)OH alcohols. Similarly, the (di)OH fatty acids possess a carboxylic acid group at the 1 position and mid-chain hydroxy group(s). All (di)OH wax esters identified are palmitate esters of (di)OH alcohols. Four of the mid-chain (di)hydroxy compounds identified were not previously reported (Speelman et al., 2009b). In Online Resource 3, the mass spectra of the methyl ester and TMS ethers of these compounds are displayed. Some compounds identified by Speelman et al. (Speelman et al., 2009b) were not detected in the TLEs or in SEs, including 7-hexacosanol (C26 7(ω20) alkanol), 9-octacosanol

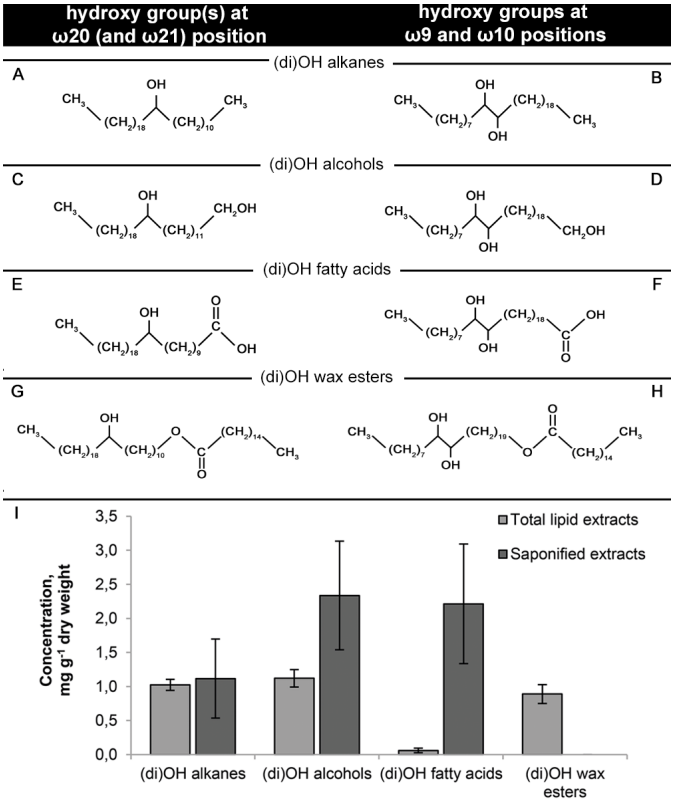


Figure 3. a–h Examples of structural formulas of each compound type and class: a 12- hentriacontanol (C31 12(ω20) alkanol), b 9,10- dihydroxynonacosane (C29 20,21(ω9,ω10) diol), c 1,13-dotriacontanediol (C32 1,13 (1,ω20) diol), d 1,20,21-nonacosanetriol (C29 1,20,21 (1,ω9,ω10) triol), e 11- hydroxytriacontanoic acid (C30 11(ω20) OH FA), f 20,21-dihydroxynonacosanoic acid (C29 20,21(ω9,ω10) diOH FA), g 11-hydroxytriacontane palmitate ester (C46 (ω20) OH wax ester), h 20,21- dihydroxynonacosane palmitate ester (C45 (ω9,ω10) diOH wax ester). i Concentrations of major classes of mid-chain (di)hydroxy compounds in TLEs and SEs. Error bars indicate the standard deviation, N=5 for TLEs and N= 4 for SEs.

(C28 (ω20) alkanol), and 17-hydroxyhexadotriacontane palmitate ester (C52 (ω20) OH wax ester). Small quantities of odd-numbered OH FAs were detected but provided too little signal to be quantified.

In Figure 3i, the concentrations are summarized for all types of mid-chain (di) hydroxy compounds. In TLEs, 0.31±0.028% dw of mid-chain (di)hydroxy compounds were detected, consisting of even amounts of (di)OH alkanes and (di)OH alcohols and a slightly lower amount of (di)OH wax esters. The most abundant compounds in TLEs are the 20,21-dihydroxynonacosane palmitate ester (C45 (ω9,ω10) diOH wax-esters) and the C29 20,21(ω9,ω10) diol. In SEs, the mid-chain (di)hydroxy compounds amounted to 0.57±0.22% dw of the biomass and 7.2±2.8% of the crude lipid fraction, respectively. Upon saponification, the (di)OH wax esters were efficiently hydrolyzed, which is apparent from their absence from SEs and the increase in the concentration of their breakdown products: C16:0 FA and (di)OH alcohols, i.e., C30-C34 1,ω20 diols, C30-C32 1,ω20,ω21 triols, and in particular, 1,20,21-nonacosanetriol (C29 1,20, 21 (1,ω9,ω10) triol). Surprisingly, (di)OH FAs turned out to be the dominant compounds in SEs, whereas in the TLEs, only C29 20,21 (ω9,ω10) diOH FA was faintly detected. The increase in (di)OH FAs after saponification suggests that they are part of lipid esters that are too large to be GC amendable and thereby undetected in TLEs. The estimated mass of all (di)OH FAs was approximately 0.22±0.092% dw. Overall, the compounds with a hydroxy group at position ω20 occur in more diverse chain lengths compared to the compounds with hydroxy groups at positions ω9 and ω10, whereas both types occur in equal amounts.

Implications for biodiesel production

The ester-bound FAs are by far the dominant lipids in *A. filiculoides*, i.e., 3.1±1.0% dw. Other *Azolla* species might have a higher lipid content, such as *A. caroliniana*, for which a crude lipid fraction of 12.7–16.4% dw was reported (Paoletti et al., 1987). However, this difference may as well be caused by adjustment of the biomass composition to the growth rate and/or a shift in carbon allocation under high CO_2 concentrations, as observed in higher plants (Poorter et al., 1997). Due to its continuous productivity, considerable amount of oil can be produced from an *Azolla* production system on an annual basis. In a controlled production system with active CO_2 supply, we achieved an oil production equivalent to 1500 kg ha^{-1} $year^{-1}$. In an outdoor system, Becerra et al. (1990) obtained a biomass productivity of 39,000 kg ha^{-1} $year^{-1}$, corresponding to an oil production of approximately 1200 kg ha^{-1} $year^{-1}$, assuming a similar lipid content. Common biodiesel feedstocks such as soybean, rapeseed, and palm have annual oil productivity of 406, 1307, and 5462 kg ha^{-1} $year^{-1}$, respectively (Karmakar et al., 2010). As an example of other novel biodiesel sources, the annual oil productivity of the microalgae *S. incrassatulus* can be estimated between

2300 and 3900 kg ha^{-1} $year^{-1}$, mainly depending on annual productivity (Petkov et al., 2012). Hence, with oil as the sole product, *Azolla* cultivation is likely outcompeted by palm oil and microalgae. Utilizing whole *Azolla* biomass can yield higher amounts of bioenergy. Muradov et al. (2014) obtained 33 % of bio-oil after pyrolysis of whole *Azolla* biomass that can be directly used as diesel fuel supplement. In their approach, all biomass components, including protein, polyphenols, and carbohydrates, are converted into pyrolysis products. The advantage of using only the lipid extract for biodiesel production is that other biomass components can be extracted separately and commercialized as higher-value products. This holds in particular for the protein fraction which is a major (20-25%) component of *Azolla* biomass and potentially valuable as feed (Alalade and Iyayi, 2006; Leterme et al., 2010; Fasakin, 1999). If we compare *Azolla* to current protein crop soy, the amount of biodiesel that can be produced from the *Azolla* lipid fraction is nearly three times higher.

Production of biodiesel from *Azolla* lipids will require a conversion process that tolerates a high FFA content. As with many other biodiesel feedstocks, the FFA content of *Azolla*, i.e., 4.2±0.38%, is higher than the limit, i.e., 2.5%, for base-catalyzed conversion (Leung et al., 2010). The FFA content can be reduced by using acid-catalyzed methylation of the FFAs prior to alkaline hydrolysis (Berchmans and Hirata, 2008). Alternatively, high FFA tolerating methylation methods can be employed, such as the use of sulphonated synthetic carbon catalysts (Toda et al., 2005).

Biodiesel quality indicators: density, cetane number, iodine value, and cold filter plugging point (CFPP) estimated using FA profile and available numerical models are provided in Figure 4, along with (estimated) values for a selection of other biodiesel sources and the limits set by the EN 14214 standard. Values were calculated from the average fatty acid profile in SEs (n=4). The biodiesel has an estimated density of 880±2.9 kg m^{-3}, a cetane number of 63±4.0, and an iodine value of 80±15. The values are all well within limits set by the EN 14214 standard. Although no European-wide limits are provided for the CFPP, it can be seen that the estimated CFPP for *Azolla* biodiesel, i.e. 7.5 ± 1.4 °C, is too high for general operating conditions. This is mainly the result of the relatively high content of lignoceric (C24:0) in *Azolla* biomass, i.e., 3.4±0.32% of the FAME fraction.

As the numerical models employed are limited to the FA profile, they do not take into account the effects of other lipid constituents. The mid-chain (di)hydroxy compounds amount to 7.2±2.8% of the crude lipid fraction. Although exact temperature characteristics are unknown, the high chain length, lack of double bonds, and GC behavior indicate that these compounds have high melting points and can therefore be a serious source of filter plugging and wax settling issues. Hence, fractionation into a FAME fraction and a fraction containing the high melting point mid-chain (di) hydroxy compounds will be necessary to ensure biodiesel quality in terms of cold temperature characteristics. To enhance the CFPP, this fractionation step may be

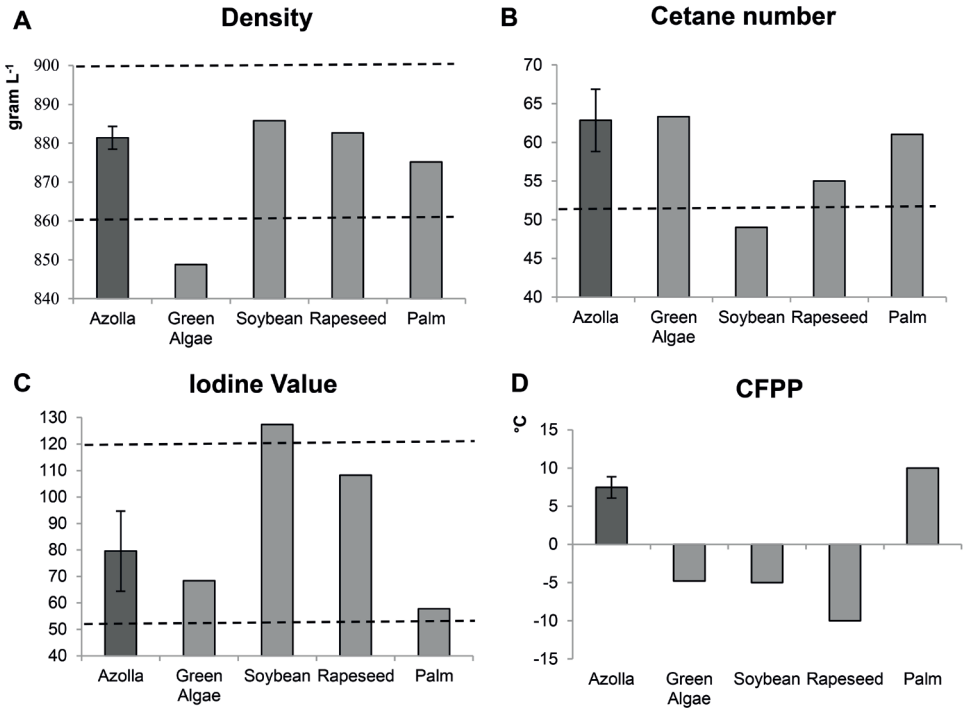


Figure 4. Estimated density, cetane number, iodine value, and cold filter plugging point (CFPP) for biodiesel derived from *Azolla* compared to other feedstock sources. Dotted lines indicate limits set by the EN 14214 European standard. Parameters for rapeseed, soybean, and palm were taken from Ramos et al. (2009) and parameters for green algae (*Scenedesmus incrassatulus*) were calculated from the fatty acid distribution obtained from Arias-Peñaranda et al. (2013).

extended to remove the C24:0 FA. Removal of C24:0 by crystallization fractionation and solvent fractionation was demonstrated by Pérez et al. (2010) in the case of peanut biodiesel. In the case of *Azolla*, full removal up to C24:0 FA would lead to a drastic decrease in CFPP to -12±1.83 °C, while density, cetane number, and iodine value would change only slightly and remain within EU limits at 880±5.2 kg m^{-3}, 61.0±3.9, and 82±16, respectively. Hence, with added processing, high-quality biodiesel can be produced from the *Azolla* lipid fraction. Although an additional fractionation step infers additional costs, the second fraction rich in mid-chain (di)hydroxy compounds may also provide a second product stream, when the unique long-chain fatty alcohols, i.e., (di)OH alcohols, and(di)OH FAs find applications as biochemicals. Surfactants, often applied as detergents in washing and cleaning formulas, and various specialty chemicals are manufactured on the basis of fatty alcohols (Biermann et al., 2011; Behr et al., 2008). However, the alcohols in *Azolla* have multiple hydroxy groups and a typical chain length of C26–C36, which is much longer than

that of the long-chain alcohols produced from FAs or petroleum (C14– C24). In food industry, fatty alcohols with a more similar chain length, such as octacosanol, are currently sold as food additive, due to their cholesterol-lowering effect in animals and humans (Hargrove et al., 2004). If the *Azolla* fatty alcohols exhibit a similar function, such an application could be of commercial interest.

OH FAs are, among others, used as surfactants, lubricants, and synthetic precursors in the polymer industry (Mutlu and Meier, 2010; Cao and Zhang, 2013). They are difficult to synthesize via chemical routes due to inertness of the fatty acid chain (Cao and Zhang, 2013). Currently, the main commercially produced OH FA is ricinoleic acid (12-hydroxy-9-octadenenoic acid), which is extracted from castor oil and used for, among others, the synthesis of polyurethanes (PUs) and diacids (Mutlu and Meier, 2010).

Hence, various possible applications of these compounds exist. Whether, in addition to biodiesel production, further separation of the fatty alcohols and (di)OH FAs is worthwhile depends on the value of the biochemical versus its processing costs. The fact that they have longer chain length and more hydroxy groups compared to currently commercialized compounds may result in a high product value but also makes it difficult to assess the commercialization potential of these compounds beforehand and therefore requires further research.

Conclusions

The lipid fraction of *Azolla* was investigated in the context of biodiesel production. Biomass drying conditions do not affect lipid yield or composition, indicating that drying conditions may be freely chosen so that they are energetically optimal. After saponification in methanol 3.2±1.0% dw of FAMEs are obtained. From the lipid composition, it is predicted that high-quality biodiesel can be produced from the *Azolla* lipid fraction but requires an additional fractionation step to decrease the CFPP. The unique long-chain (di)hydroxy fatty acids and fatty alcohols that are separated in this fractionation step may provide a valuable secondary product stream when purified into biochemicals, with possible applications to chemical industry and nutrition.

Acknowledgments

We acknowledge the financial support by Shell and the LPP foundation. We also like to thank Climate-KIC for support via the AzoFaSt pathfinder project.

Online Resources

The online resources can be found in the online version of this article (open access): https://link.springer.com/article/10.1007/s12155-015-9665-3

Chapter 6

Phenolic compounds and their biosynthesis pathways in the aquatic fern *Azolla*

Paul Brouwer[1,2], Klaas G.J. Nierop[2], Ric C.H. de Vos[3], Isabela G. Perry[1], Gert-Jan Reichart[2], Adrie van der Werf[4], Andrea Bräutigam[5], Ingrid van der Meer[3], and Henriette Schluepmann[1].

[1] Molecular Plant Physiology, Institute of Environmental Biology, Utrecht University, The Netherlands;
[2] Department of Earth Sciences, Faculty of Geosciences, Utrecht University, The Netherlands;
[3] Bioscience, Wageningen University & Research, The Netherlands;
[4] Agrosystems Research, Wageningen University & Research, The Netherlands;
[5] Leibniz Institute of Plant Genetics and Crop Plant Research, Germany

manuscript in preparation

Abstract

The aquatic fern *Azolla* is a potential source for the production of feed, fuels and chemicals, due to its high growth rates, symbiosis with the nitrogen-fixing *Nostoc azollae* and production being independent of arable land. Here we investigated (poly)phenolic compounds in *Azolla* relevant to applications of *Azolla* biomass as feed or feedstock for the production of bio-chemicals. Condensed tannins (CTs) were the most abundant soluble (poly)phenols in *Azolla* species, accounting for 4.0% and 5.8% of the biomass of A.*filiculoides* and *A. pinnata*, respectively. *Azolla* biomass further contained high amounts of dicaffeoylquinic acids, with minor contributions of chlorogenic acid and tricaffeoylquinic acid. Caffeoylquinic acids and, particularly, tricaffeoylquinic acids were more abundant in *A. pinnata*, whilst *A. filiculoides* contained more insoluble phenolic compounds. Insoluble phenolic compounds in *A. filiculoides* consisted of coumaric and ferulic acid building blocks, with a small contribution of guiacyl (G) type lignin. Fluorescence typical of insoluble phenolic compounds was observed in the vasculature, likely deposited in tracheid cells of the xylem.

Analysis of the *A. filiculoides* genome and gene-expression data from sporophytes revealed *Azolla* homologues of genes encoding core enzymes in the phenylpropanoid and flavonoid pathways. Conserved lignin biosynthesis genes were present in *A. filiculoides*, but lowly expressed with respect to downstream *AzfiHCT/HQT*-like genes that may have a role in the production of dicaffeoylquinic acids. A Leuco-Anthocyanidin Reductase (*AzfiLAR*-like) gene was the most abundantly expressed gene in secondary metabolism and likely responsible for the high content of condensed tannins in *Azolla* biomass. *AzfiLAR*-like represents a key invention in the common ancestor of ferns and seed plants. Manipulating (poly)phenol biosynthesis in this remarkable fern will permit further development of *Azolla* as a sustainable novel crop.

Introduction

The aquatic fern *Azolla* is a potential source for the production of feed, fuels and chemicals, due to its high growth rates in the absence of nitrogen fertilizer and independence of arable land. Growth in the absence of nitrogen fertilizer is supported by *Nostoc azollae* cyanobacteria, that reside in specialized leaf pockets where they fix atmospheric nitrogen and supply the fixed nitrogen to the plants (Becking, 1987; Meeks et al., 1986). Since proteins are the main component of *Azolla* biomass, application as feed was explored as a local substitute to conventional protein sources such as soybean meal. While *Azolla* was fed to a number of domestic animals, diet inclusion rates were lower than of commercial feeds, such as soybean meal (Leterme et al., 2010; Abdel-Tawwab, 2008; Becerra et al., 1990; Basak et al., 2002; Alalade and Iyayi, 2006). Recent investigations have pointed towards phenolic compounds in *Azolla* as a possible cause for this limited inclusion rate. *A. caroliana* was shown to contain condensed tannins (CTs) (Nierop et al., 2011) which are known to limit protein digestibility via polyphenol-protein interactions (Bravo, 1998). In addition, *A. filiculoides* and *A. pinnata* contain more phenolic compounds than generally encountered in foods (Chapter 4).

Phenolic compounds are also known to have important functions within the plant, i.e. protecting it from herbivores, UV-light and oxygen radicals and are therefore beneficial to *Azolla* cultivation by reducing the impact of abiotic and biotic stress (Barbehenn and Peter Constabel, 2011; Coley, 1988; Stafford, 1991; Pollastri and Tattini, 2011; Demkura et al., 2010). Specific phenolic compounds, such as were found to be health-enhancing and may be extracted as high value products. (Boudet, 2007; Clifford, 2000). Additionally, phenolic compounds may find use as chemical building blocks for the production of aromatic compounds now synthetized from oil (Bruijnincx and Weckhuysen, 2013).

Hence the development of an *Azolla* production chain as a novel source of feed and/or chemicals requires more knowledge of the phenolic compounds present in *Azolla* biomass. Additionally, understanding of the biosynthesis pathways leading to the production of these compounds may allow targeted breeding and selection to increase the value of the biomass.

Several phenolic compounds have been identified in *Azolla* in separate studies, including deoxyanthocyanidins (luteolinidin-5-glucoside), coumarins (aesculetin, 6-(3'-glucosylcaffeoyl)-aesculetin), hydroxycinnamic acids (caffeic acid esters, chlorogenic acid, p-coumaric acid), flavonols (rutin, quercetin), flavones (homoorientin, saponarin) and proanthocyanidins comprised of (epi)catechin subunits (Nierop et al., 2011; Ishikura, 1982; Teixeira et al., 2001). However, a comprehensive inventory of phenolic compounds in *Azolla* has not yet been performed.

Liquid Chromatography – Mass Spectrometry (LC-MS) has been widely used to

analyze soluble phenolic compounds (Flamini, 2013; Parejo et al., 2004; Moco et al., 2007; Bravo et al., 2006). Specifically, high resolution accurate mass MS enables the detection of large numbers of compounds in a single extract and provides valuable information on the chemical composition and thereby the identity of the metabolites (Moco et al., 2006). However, the detection of large (poly)phenols such as CTs by LC-MS is limited by poor column separation and low ionization efficiency of larger complexes (Kalili et al., 2013). An alternative for determining the basic composition of CTs is thermal degradation via flash pyrolysis followed by separation of the small pyrolysis products by Gas Chromatography (GC) and detection by MS. By Thermally assisted Hydrolysis and Methylation (THM), i.e. pyrolysis in the presence of tetramethylammonium hydroxide (TMAH), carboxylic acid and hydroxyl groups are transformed in situ to their corresponding methyl esters and methyl ethers, allowing higher compound detection and resolution of phenolic compounds versus conventional pyrolysis. THM has therefore been widely used to study the basic composition of plant CTs (Nierop et al., 2005; Galletti et al., 1995; Garnier et al., 2003). Since pyrolysis and THM can be performed on both extracts and intact biomass, it has further been used to study insoluble (poly)phenols such as lignins (Rencoret et al., 2016; Filley et al., 1999; Kuroda et al., 2002; Clifford et al., 1995). Although conventional pyrolysis provides less information for the identification of (poly)phenols than THM, it also breaks down polysaccharides and protein which can be used as references to derive semi-quantitative information on (poly)phenol abundance in plant biomass.

Many genes encoding enzymes in the (poly)phenolic biosynthesis pathways have been described in seed plants (Fraser and Chapple, 2011; Saito et al., 2013) and have been used as a basis for comparative genomic studies to predict pathway changes during land plant evolution (Weng and Chapple, 2010; Tohge et al., 2013). However, little is known on pathways in ferns because a fern genome sequence has until now not been available. Ferns are a sister lineage to seed plants and likely have many secondary metabolism pathways in common to all vascular plants. Genes at the start of phenylpropanoid pathway, responsible for the production of the essential substrate: p-coumaroyl-CoA, evolved before the emergence of land plants, presumably from enzymes of primary metabolism (Tohge et al., 2013; Nelson, 2006; Muradov et al., 2014). Additionally many genes required for the synthesis of hydroxycinnamic acids and lignin have been identified in mosses (Tohge et al., 2013). Hence we may expect that many components of the phenylpropanoid pathway are also present in the fern *Azolla*.

Here we set out to characterize both soluble and insoluble (poly)phenolic compounds present in *A. filiculoides* and *A. pinnata* by combining high mass resolution LC-OrbitrapFTMS, flash pyrolysis-GC-MS and THM-GC-MS analyses. Using the recently assembled *A. filiculoides* genome we investigate conserved phenolic biosynthesis pathways using homology searches, phylogenetic computations and analyses

of previously published expression profiling experiments (Chapter 3) (Brouwer et al., 2014). The obtained data is used to identify candidate genes in the *A. filiculoides* genome responsible for the production of the major phenolic compounds in *Azolla*.

Materials and methods

Materials

Phenolic compounds used as references, i.e. caffeic acid, chlorogenic acid, catechin, naringin, tannic acid and quercetin were purchased from Sigma Aldrich. Anhydrous acetone (90%) was obtained from Merck. Ethyl acetate, dichloromethane (DCM) and methanol (MeOH), graded for trace element analysis, were obtained from VWR international. Black spruce tannins were provided by Caroline M. Preston (Lorenz and Preston, 2002) and characterized by Nierop et al. (2005).

Plant material and growth conditions

A. filiculoides was obtained from a pond near Galgenwaard, Utrecht, the Netherlands (Brouwer et al., 2014). *A. pinnata* was obtained from the International Rice Research Institute (IRRI), under accession number 535. Plants used for chemical analysis were grown in a growth chamber providing 16 h light at an intensity of 100 µmol m^{-2} s^{-1}, a stable temperature of 21°C and 70% humidity. The growth medium for *Azolla* species consisted of 0.65 mM $NaH_2PO_4 \cdot H_2O$, 1.02 mM K_2SO_4, 1 mM $CaCl_2 \cdot 2H_2O$, 1.65 mM $MgSO_4 \cdot 7H_2O$, 17.9 µM Fe-EDTA, 9.1 µM $MnCl_2 \cdot 4H_2O$, 1.6 µM $Na_2MoO_4 \cdot 2H_2O$, 18.4 µM H_3BO_3, 0.8 µM $ZnSO_4 \cdot 7H_2O$, 0.3 µM $CuSO_4 \cdot 5H_2O$, 0.2 µM $CoCl_2 \cdot 6H_2O$. For each species, two cultures were maintained and harvested weekly to remain in the linear growth phase over a period of five weeks. Harvested biomass was freeze-dried and the consecutive harvests were pooled into one batch for each culture. These two batches were used for further analysis.

Extraction and separation of phenolic compounds

Freeze-dried plant material was first extracted with DCM:MeOH 9:1 in a Soxhlet device to remove lipids. After lipid extraction, residual *Azolla* was dried at room temperature, weighed and aliquots were taken for dry weight determination and chemical analysis. The residual biomass was extracted in aqueous acetone (70 v/v%) at room temperature in a rotary shaker set at 200 rpm. After 24 h the extract was removed by centrifugation (5000 g) and extracted again with aqueous acetone. A

small aliquot of the extract was taken for total phenol analysis. The extraction was repeated three times, after which the extracts were pooled. The acetone was removed from the pooled extracts using a rotary evaporator. The residual aqueous fraction was firstly extracted with ethyl acetate yielding the ethyl acetate soluble fraction (Figure 1). Subsequently the aqueous fraction was combined with an equal volume of methanol and loaded on a Sephadex LH20 column. The column was eluted with methanol:water (1:1) until the eluens turned colorless to yield the methanol-water fraction. Lastly a tannin fraction was obtained by eluting with aqueous acetone (70 v/v%) (Figure 1). The ethyl acetate fraction was dried under continuous nitrogen flow, whereas the organic phases of the methanol-water and tannin fractions were removed by rotary evaporation, after which the water phases were freeze-dried. All three fractions were split into 6 aliquots used for C/N, Py-GC-MS and THM, total phenol, acid butanol and dw analysis.

Chemical analyses

To determine total carbon and total nitrogen content (C and N), 1-2.5 mg of material was weighed in tin cups and C/N content determined using an elemental analyzer (Fison NA 1500 CNS), connected to a mass spectrometer (Finnigan Delta Plus). For dry weight analysis samples were placed in an oven at 105°C for 24 h, cooled in an exsiccator for 1 h and weighed.

Total phenol was determined by extracting samples with aqueous acetone (70 v/v%) for 24 h at 21 °C, followed by reacting the extract with the Folin-Ciocalteu reagent (Waterman and Mole, 1994). Tanninc acid (TA) was used as the reference standard. To detect condensed tannins, the acid butanol assay was performed on the dried biomass, residues and (poly)phenolic fractions. This was conducted by adding 1.2 ml of n-butanol-HCl solution (5 v/v%) and 20 µl of 2% ferric ammonium sulfate in 2N HCl to the sample, followed by vortexing and heating at 95 °C for 50 minutes. Afterwards, samples were allowed to cool down and absorbance at 550 nm was determined using a spectrophotometer. Black spruce tannins were used as reference for the acid-butanol assay.

A modified vanillin assay was used to determine the degree of polymerization of *Azolla* tannins (Butler et al., 1982). In short, purified tannins were dissolved in either methanol or glacial acetic acid, containing 4% concentrated HCl and 0.5% vanillin. Samples were allowed to react for 20 minutes at 30°C, after which absorbance was read at 500 nm using a spectrophotometer. The degree of polymerization was calculated from the ratio of absorbance in methanol over glacial acetic acid.

Analyses of (poly)phenols using LC-MS was essentially as described previously (van der Hooft et al., 2012). Biomass was extracted with aqueous methanol (75 v/v%) acidified with formic acid (0.1 v/v%). Separation was on a C18 column (150 x 2.0

mm; 3 µm particles) and detection was by both photodiode array (PDA; 200-700 nm; Waters) and LTQ-Orbitrap FTMS (Thermo) with negative electrospray ionization, using a scan range of m/z 100-1.300 at a resolution of 60,000. Compounds identification was by matching their observed accurate masses, retention times and UV/Vis-spectra, if available, with in-house generated LCMS-based phenolic compound libraries from different plant species such as tomato fruit (Moco et al., 2006; Ballester et al., 2016), broccoli (Lopez-Sanchez et al., 2015) and strawberry (Carbone et al., 2009).

Conventional flash pyrolysis was carried out using a Horizon Instruments Curie-Point pyrolyser. Samples were heated for 5 s at 600 °C. The pyrolysis unit was connected to a Carlo Erba GC8060 gas chromatograph and the products were separated with a fused silica column (Var-ian, 25 m, 0.32 mm i.d.) coated with CP-Sil5 (film thickness 0.40 µm). Helium was used as carrier gas. The oven temperature programme was: 40 °C (1 min) to 180°C at 7 °C min^{-1} followed by a ramp at 20 °C min^{-1} to 320 °C (held 15 min) The column was coupled to a Fisons MD800 mass spectrometer (range m/z 45–650, ionization energy 70 eV, cycle time 0.7 s).

For THM analysis samples were pressed onto Curie-point wires, a droplet of a 25% solution of TMAH was added and samples were dried under a 100W halogen lamp. Analysis of the THM products by way of GC–MS was performed as described for conventional pyrolysis. Compound identification was carried out by way of mass spectral comparison using a NIST library interpretation of the spectra, retention

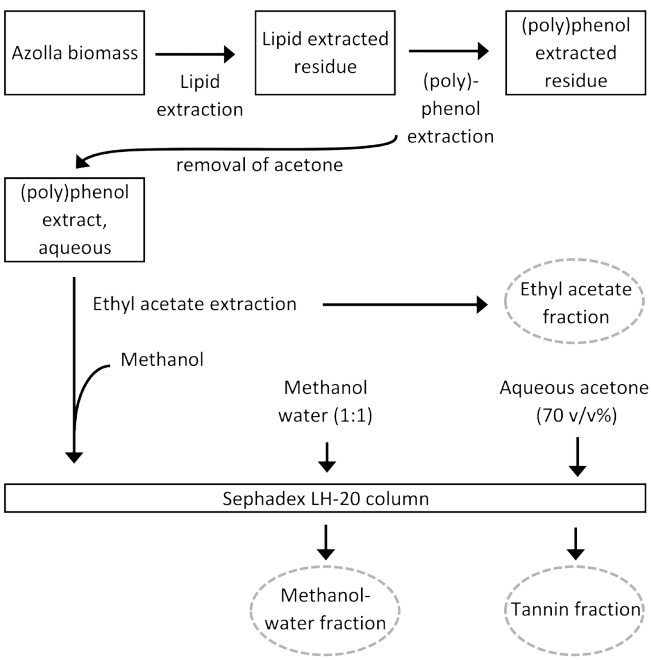


Figure 1. Schematic representation of the (poly) phenol extraction and separation procedure. Lipid extraction was performed with DCM:MeOH 9:1, and (poly) phenol extraction was performed with aqueous acetone (70 v/v%).

times and/or comparison with literature data.

Bioinformatics analysis

Sequence assembly data and annotation data from the *A. filiculoides* accession Galgenwaard was made available to this study by the *Azolla* Genome Sequencing Consortium. To identify candidate enzymes from the biosynthesis pathways of phenolic compounds we submitted the *Azolla* protein candidates from the *Azolla* genome annotation (version 1) to KEGG and Mercator for automatic annotation. For KEGG a reference set representative for Eukaryotes was used. In case automatic annotation yielded no candidate enzymes, we used BLAST searches to identify protein-candidates highly similar to functionally annotated proteins. The candidate enzymes of the phenylpropanoid and flavonoid biosynthesis pathways were aligned with known enzymes from other species using MAFFT. Protein candidates with highly conserved amino acid sequences to other plant species were considered as possible homologs. Subsequently, polygenetic trees were obtained by checking the alignment with Gblocks and construction of phylogenetic trees using Fyml (Dereeper et al., 2008). Proteins that that grouped with functionally annotated genes from other plants species, were considered as homologs.

RNA-sequencing data from *A. filiculodies* accession Galgenwaard was available from previous studies (Brouwer et al., 2014) (Chapter 3). Raw reads of RNA sequencing were screened for quality using fastQC, adapters were removed by fastx and reads were trimmed using Trimmomatic (Bolger et al., 2014). Subsequently, trimmed reads were aligned to coding sequences derived from gene models from to the annotation using bowtie2 in ‘single end’ mode allowing for 1 mismatch for a seed sequence length of 20. The amount of reads per million of reads (RPM) was calculated over the total mapped reads for each sample.

Results

Combining LC-MS and pyrolysis techniques to characterize (poly)phenols in *Azolla*

Complementary detection of hydroxycinnamic acid esters and flavonoids by LC-MS and pyrolysis techniques

A series of phenolic compounds detected by LC-MS in acidic methanol extracts of *A. filiculoides* and *A. pinnata* are provided in Table 1 and a chromatogram indicating major peaks is depicted in Figure 2. Hydroxycinnamic acids were detected at high intensity in both *A. filiculoides* and *A. pinnata* extracts and predominantly comprised of dicaffeoylquinic acids accompanied by lower intensity signals of caffeoylquinic acids (including chlorogenic acid) and tricaffeoylquinic acid. Additionally, both species contained coumaroylquinic acids, whereas feruloylquinic acid was only detected in *A. filiculoides* (Table 1). Procyanidins consisting of three and two (epi)catechin subunits were detected in both species, as well as the epicatechin monomer (Table 1). In addition, *A. pinnata* contained a number of flavonol-glucosides: quercetin-3-O-rutinoside (rutin), quercetin-3-O-glucoside and kaempferol-3-O-rutinoside, while only traces of quercetin-3-O-glucoside were detected in *A. filiculoides* (Table 1, Figure 2).

The ratio between the intensity of $[M-H]^-$ ions of each phenolic compound in *A. filiculoides* and *A. pinnata* extracts, provided in Table 1, indicates that *A. pinnata* generally contains higher levels of phenolic compounds. In particular, *A. pinnata* is more rich in tricaffeoylquinic acids and has broader distribution of dicaffeoylquinic acid isomers than *A. filiculoides*.

Major THM products of *Azolla* biomass are provided in Figure 3 and a full list of identified peak and their relative abundance is provided in Supplemental Table A. In line with the LC-MS data indicating a series of caffeoylquinic acids, for both *A. filiculoides* and *A. pinnata* methylated caffeic acid was a major THM product (Ca) next to 1,4 dimethoxybenzene (Q), which is derived from quinic acid (Supplemental data A) (Wang et al., 2013). Additionally, THM products typical of flavonoids: 1,3,5-trimethoxybenzene (A), 1-methyl-2,4,6-trimethoxybenzene (A) and 3,4-dimethoxybenzoic acid methyl ester (B2), were highly abundant in both *A. filiculoides* and *A. pinnata* (Supplemental Table A) (Nierop et al., 2011; Garnier et al., 2003).

Conventional pyrolysis of whole *Azolla* ferns (Supplemental Table B) yielded catechol in a ratio of 2.0 to (poly)saccharide pyrolysis products, versus 2.8 for *A. pinnata*. Additionally, 4-vinylphenol, indicative of coumaric acid, was detected at a ratio of 0.46 to (poly)saccharide pyrolysis products in *A. filiculoides* versus 0.22 for *A. pinnata*. 4-vinylguaiacol, typical of ferulic acid, was obtained at a ratio of 0.15 to

(poly)saccharide pyrolysis products in *A. filiculoides* versus 0.03 in *A. pinnata* (Supplemental Table B).

Fractionation of soluble (poly)phenols reveals CTs as most abundant compounds

To test whether the THM and pyrolysis products obtained from *Azolla* biomass, typical of flavonoïds and hydroxycinnamic acids, originate from soluble compounds, we performed an extraction of phenolic compounds followed by separation into three soluble fractions i.e. an ethyl acetate fraction, a tannin fraction and a methanol-water fraction. In Table 2 we provide the dry weight and the content of total phenols, CTs, carbon and nitrogen in the three fractions compared to the initial biomass (Table 2).

Biomass of *A. pinnata* contained nearly 1.5 times as much total phenol compared to *A. filiculoides*. Although the ethyl acetate and tannin fraction weighed less than the methanol-water fraction they were more enriched in phenolic compounds as indicated by the total phenol assay. For both species, most of the soluble phenolic compounds were retained in the tannin fractions.

Results of the acid butanol assay indicated that the soluble tannin fraction consisted purely out of CTs. Additionally, some CTs were dissolved by ethyl acetate and consequently ended up in the ethyl acetate fraction. In total, CTs represented 5.3%

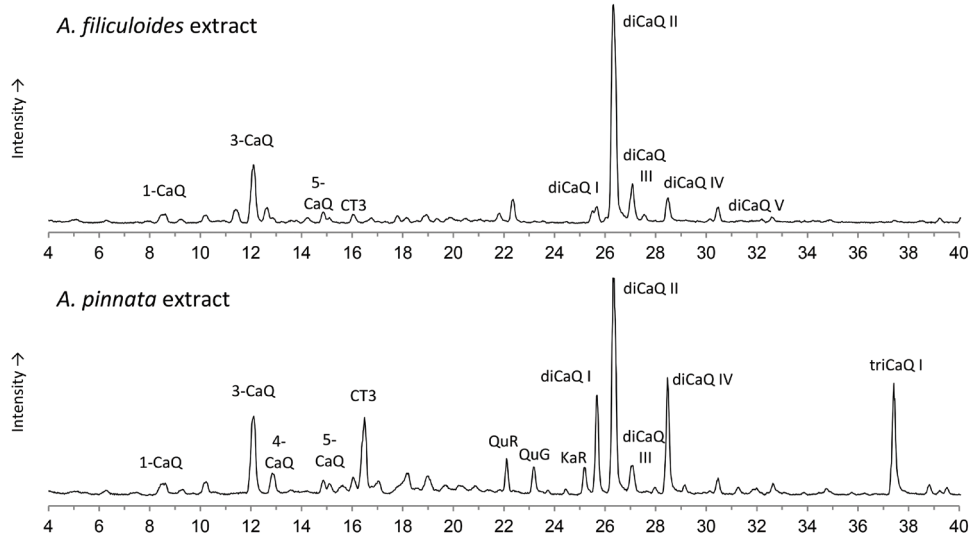


Figure 2. LC-MS chromatogram of methanolic extracts of *A. filiculoides* and *A. pinnata*. Abbreviations refer to: Caffeoylquinic acid (CaQ), (Epi)catechin trimer (CT3), Quercetin-3-O-rutinoside (rutin) (QuR), Quercetin-3-O-glucoside (QuG), Kaempferol-3-O-rutinoside (KaR), dicaffeoylquinic acid (diCaQ) and tricaffeoylquinic acid (triCaQ)

Table 1. List of (poly)phenolic compounds identified by LC-MS

Compound	RT	[M-H]-	Fragment ions[2]	*A. fil* : *A. pin*[3]
1-caffeoylquinic acid	8.55	353.0874	179;191	1 : 1.2
5-caffeoylquinic acid	12.09	353.0875	191	1 : 1.1
3-caffeoylquinic acid	12.86	353.0875	173;179;191	1 : 3.2
Procyanidin, (epi)catechin dimer	13.64	577.1346	-	1 : 0.8
4-caffeoylquinic acid	15.11	353.0875	191	1 : 1.8
Epicatechin	15.28	289.0715	-	1 : 6.9
4-coumaroylquinic acid	16.04	337.0927	191	1 : 1.5
Procyanidin, (epi)catechin trimer	16.64	865.1876	-	1 : 2.3
Feruloylquinic acid	18.17	367.1031	191	1 : 0
5-Coumaroylquinic acid	18.97	337.0928	191	1 : 1.2
Quercetin-3-O- rutinoside (Rutin)	22.12	609.1455	-	1 : 0.1
Quercetin-3-O- glucoside	23.18	463.0877	-	0 : 1
Kaempferol-3-O-rutinoside	25.20	593.1504	-	1 : 22
Dicaffeoylquinic acid I	25.67	515.1192	353	0 : 1
Dicaffeoylquinic acid II	26.32	515.1192	353	1 : 4.8
Dicaffeoylquinic acid III	27.07	515.1192	353	1 : 0.8
Dicaffeoylquinic acid IV	28.48	515.1192	353	1 : 0.7
Dicaffeoylquinic acid V	32.64	515.1192	-	1 : 2.7
Tricaffeoylquinic acid I	37.41	677.1503	515	1 : 4.5
Tricaffeoylquinic acid II	38.82	677.1503	-	1 : 63
Tricaffeoylquinic acid III	39.46	677.1503	-	1 : 99

[1] RT = retention time
[2] Fragments resulting from unintended in-source fragmentation
[3] based on the relative intensity of the [M-H]- ions, corrected for the input weight; extracts were treated identically.

of the dry weight of *A. filiculoides*, of which 4.3% could be extracted by aqueous acetone. In *A. pinnata*, CTs represented 8.0% of the dry weight, of which 6.9% was extractable. The vanillin assay was used to determine the average degree of polymerization of the tannin fraction, which was 7.2±1.2 and 6.7±0.7 for *A. filiculoides* and *A. pinnata* respectively.

The carbon content of the tannin fractions of *A. filiculoides* and *A. pinnata* was 55.7% and 60.5%, respectively, close to the carbon content of pure CTs: 59-66%. Some nitrogen was measured in the tannin fractions suggesting minor contamination with nitrogen-containing compounds. The low concentration of organic matter of the methanol-water fraction (1% of nitrogen and 23% of carbon) indicates a high content of soluble ash.

To link the data from Table 2 to specific (poly)phenolic compounds, each fraction

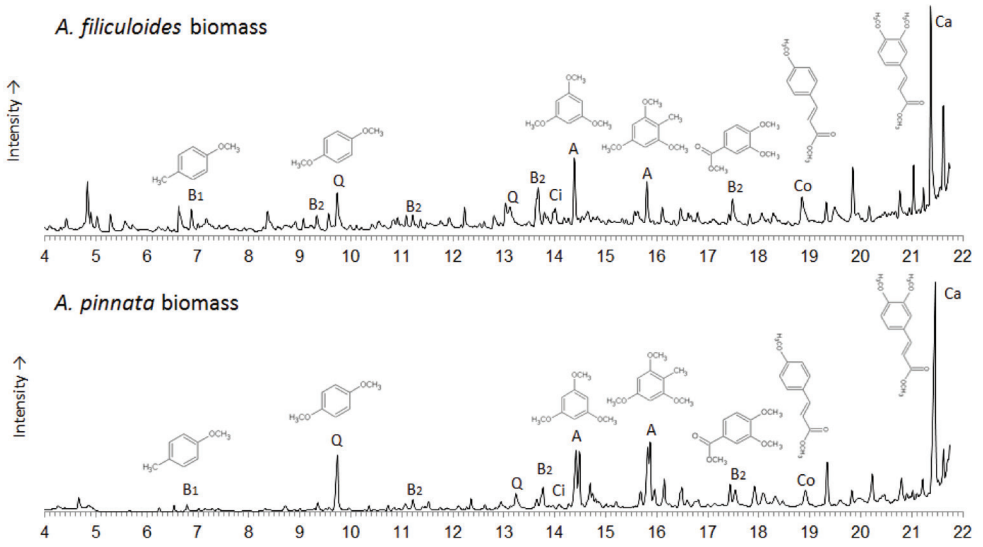


Figure 3. Partial chromatrograms of THM products of *A. filiculoides* and *A. pinnata* biomass. Labels indicate origin of THM products: flavanoïd A-ring (A), flavanoïd B-ring with n hydroxyl groups (Bn), Quinic acid (Q), Caffeic acid (Ca), p-Coumaric acid (Co) and Cinnamic acid (Ci).

was analyzed by LC-MS, THM and conventional pyrolysis GC-MS. In Figure 4, we provide a selection of chromatograms for each of the fractions of *A. filiculoides*. A full set of chromatograms for both *A. filiculoides* and *A. pinnata* is provided in Supplemental Figure S1-3.

The ethyl acetate fraction of *A. filiculoides* contained high intensity signals of dicaffeoylquinic acids, with lower intensity signals of caffeoylquinic acids (Figure 4). The ethyl acetate fraction of *A. pinnata* contained higher levels of tricaffeoylquinic acids than *A. filiculoides* and further contained a procyanidin trimer and flavonol-glucosides. Results of THM supported the predominance of caffeic acid esters in these fractions, as the relative abundance of methylated caffeic acid and the quinic acid derived 1,4-dimethoxybenzene accounted for 66-73% of all (poly)phenolic compounds. The dominance of dicaffeoylquinc acids versus monocaffeoylquinic acids in the ethyl acetate fraction suggested by LC-MS, was also apparent from THM results: the ratio of methylated caffeic acid to 1,4-dimethoxybenzene was twice as high as observed for pure chlorogenic acid (Supplemental Table A). In line with the results from the acid-butanol assay, THM products representing the A-ring of condensed tannins were observed in the ethyl acetate fraction, especially in the case of *A. pinnata* (Supplemental Figure S2, Table A). Conventional pyrolysis of the ethyl acetate fraction yielded predominantly catechol and phenol, derived from caffeic acid and quinic acid, respectively (Supplemental Table B).

Table 2. Chemical composition of the whole *Azolla* biomass and soluble (poly) phenolic fractions.

A. filiculoides	Mass (%)	Total phenol (% TA-eq)	Condensed tannins (% CT-eq)	Nitrogen (%)	Carbon (%)
biomass		3.4 ± 0.4	5.3 ± 0.2	3.7 ± 0.1	42.8 ± 0.4
Ethyl acetate	1.3 ± 0.4	49 ± 3	10 ± 3	n/a	n/a
Tannin	4.0 ± 0.8	63 ± 11	101 ± 23	0.4 ± 0.1	55.7 ± 3.6
Methanol-water	13.0 ± 1.1	10 ± 7	1.6 ± 0.7	1.0 ± 0.0	19.3 ± 0.5
A. pinnata					
biomass		5.0 ± 1.2	8.0 ± 0.1	2.9 ± 0.0	41.6 ± 0.7
Ethyl acetate	2.7 ± 0.1	53 ± 2	38 ± 15	n/a	n/a
Tannin	5.8 ± 2.6	60 ± 3	97 ± 12	0.3 ± 0.1	60.5 ± 2.2
Methanol-water	9.3 ± 1.1	9.1 ± 5.3	2.3 ± 2.0	1.0 ± 0.1	23.0 ± 2.6

LC-MS analysis of the tannin fraction yielded (epi)catechin monomers, dimers, trimers and tetramers, but not higher molecular weight CTs due to a mass detection limit of m/z 1300. A minor contribution of tricaffeoylquinic acids to the tannin fraction was observed for both species. In the tannin fraction of *A. pinnata* flavonol-glucosides were also detected. THM of the tannin fractions yielded the same typical flavonoid products as for the whole biomass. THM products from the flavonoid A-ring, i.e. 1,3,5-trimethoxybenzene and 1-methyl-2,4,6-trimethoxybenzene were most abundant, followed by THM products from a flavonoid B-ring containing two hydroxyl groups, i.e. 3,4-dimethoxybenzoic acid methyl ester and 1,2-dimethoxybenzene. A THM product indicative for a flavonoid B-ring containing one hydroxyl groups, i.e. 4-methoxybenzoic acid methyl ester, was also detected upon THM of the tannin fraction, although roughly 3-4 times less abundant than 3,4-dimethoxybenzoic acid. THM products indicative of a B-ring with three hydroxyl groups were only found in trace amounts. As expected, catechol was also the dominant pyrolysis product from the tannin fraction. The *Azolla* tannin fraction therefore predominantly consisted CTs, or pro-anthocyanidins, comprised of (epi)catechin units, with a lesser contribution of (epi)afzelechin (containing one hydroxyl group at the B-ring).

The methanol-water fractions mainly contained hydroxycinnamic acids, similar to the ethyl acetate fractions. Although dicaffeoylquinic acids remained the major compounds, the fractions were enriched with caffeoylquinic acids and coumarylquinic acids, but depleted of tricaffeoylquinic acids, when compared to the biomass (Figure 4). The methanol-water fraction of *A. filiculoides* further contained most of the feruloylquinic acid. (Epi)catechin dimers and trimers were absent from the methanol-water fractions, but monomeric epicatechin as well as flavonol-glucosides were

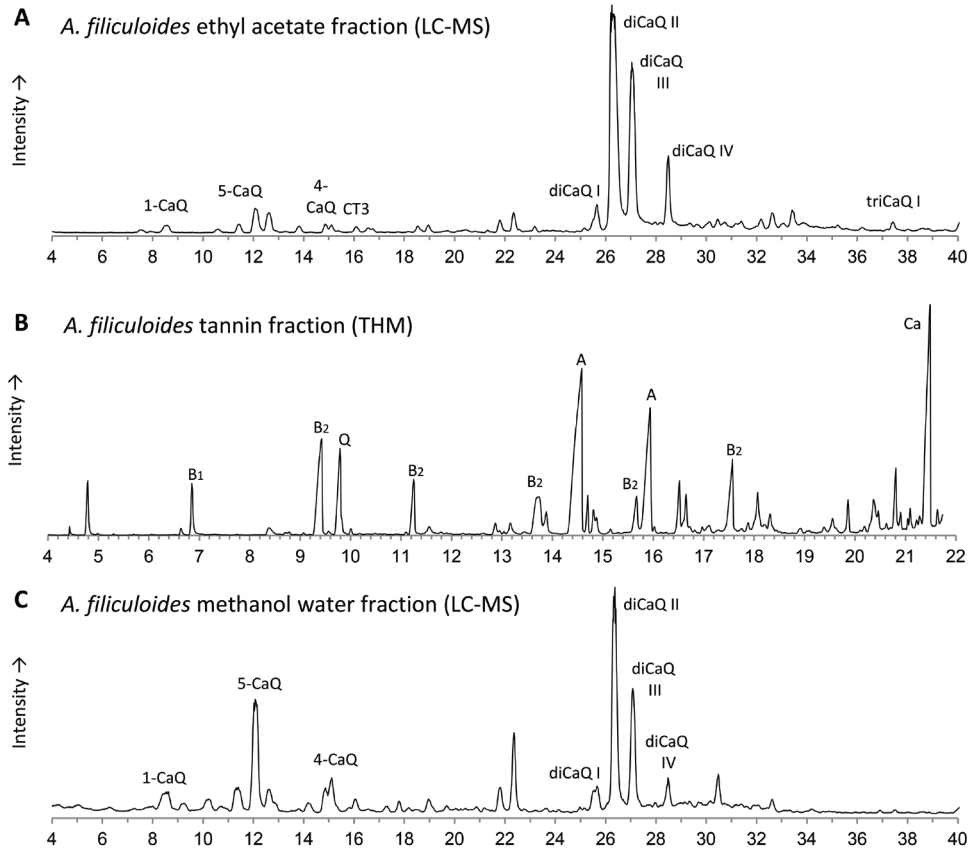


Figure 4. Partial chromatograms of fractions of soluble (poly)phenols of *A. filiculoides* analyzed by LC-MS (A,C) and TMH (B). Labels as in Figure 2 and 3.

detected in *A. pinnata*. In line with LC-MS data, THM of the methanol-water fractions mainly yielded caffeic acid and quinic acid derived THM products (Supplemental Table A). THM of the methanol-water fraction of *A. pinnata* further yielded p-coumaric acid at an abundance of 1.7% in *A. pinnata* and flavonoid A-ring products at 1.29% relative to all detected phenolics compounds. These THM products were less abundant for the methanol-water fraction of *A. filiculoides*, i.e. 0.6% and 0.8% respectively. (Supplemental Table A). Consistent with the low level of total phenols in the methanol-water fraction (Table 2), conventional pyrolysis yielded products from (poly)saccharides at similar abundances compared to products from (poly)phenols as well as amino acid related products: indole and 3-methylindole. The relative abundances of pyrolysis products of ferulic acid, i.e. 4-vinylguaiacol, and coumaric acid, i.e. 4-vinylphenol, were higher in the methanol-water fraction compared to the other fractions, but low compared to the initial biomass.

We conclude that CTs consisting of (epi)catechin, and to a lesser extent, (epi) afzelechin monomers are the most abundant soluble phenolic compounds in *Azolla*, followed by dicaffeoylquinic acids.

Detection of small amounts of lignin in *Azolla*

To identify insoluble (poly)phenolic compounds in *Azolla* we analyzed the residues after lipid extraction and subsequent (poly)phenol extraction for their mass, C/N content, condensed tannin content, pyrolysis and THM products, and compared this with the original biomass.

Lipid extraction removed 13.4% and 13.1% of the initial biomass, corresponding to 9.4-6.9% of the carbon content, for *A. filiculoides* and *A. pinnata*, respectively, but did not decrease the nitrogen content (Table 3). Subsequent (poly)phenol extraction solubilized 22.6% and 21.9% of the biomass for *A. filiculoides* and *A. pinnata* respectively. Apart from 5.8% and 7.3% carbon, (poly)phenol extraction also removed 0.4% and 0.3% of nitrogen for *A. filiculoides* and *A. pinnata*, respectively (Table 3). Although tannin concentrations by the acid-butanol assay showed some variation between the biomass and residue after lipid extraction, lipid extraction did not decrease tannin content significantly. After (poly)phenol extraction 15% and 9% of the tannins remained in the residue of *A. filiculoides* and *A. pinnata*, respectively.

Conventional pyrolysis was used to track the effect of the subsequent extractions on the relative abundance of proteins, lipids and (poly)phenols. Figure 5 depicts the yield of the pyrolysis products in the extracted residues versus the start material and a complete list is provided in Supplemental Table B. The abundance of protein pyrolysis products was similar for the biomass and the lipid extracted residue, in line with the nitrogen content (Figure 5, Table 3). Additionally after (poly)phenol extraction there is little difference in the yield of protein pyrolysis products, suggesting the nitrogen extracted is mainly non amino acid nitrogen. Lipid pyrolysis products were, as expected, much less abundant for the lipid extracted residue and remain similarly abundant after subsequent (poly)phenol extraction. The abundance of catechol is not affected by lipid extraction, but is reduced by (poly)phenol extraction to 33% for *A. filiculoides* and 22% for *A. pinnata*. THM analysis of the (poly)phenol extracted residue revealed that catechol may be obtained from both caffeic acid esters and (epi)catechin type CTs (Supplemental Table A).

The relative abundance of coumaric acid derived 4-vinylphenol is not influenced by lipid extraction for both species, and also not by (poly)phenol extraction in the case of *A. filiculoides*. The latter extraction, however, reduced the relative abundance of 4-vinylphenol to 45% for *A. pinnata*. Unexpectedly, the abundance of feruloyl derived 4-vinylguaiacol was reduced by lipid extraction to 56% and 69% in A.*filiculoides* and *A. pinnata* respectively (Figure 5). Subsequent (poly)phenol extraction did not decrease its abundance further for *A. filiculoides* and only slightly for *A. pinnata*.

The occurrence of 4-vinylguaiacol is also linked to lignin. Although it was previously reported that *A. caroliniana* did not contain lignin (Nierop et al., 2011), specific lignin peaks were detected in residues of *A. filiculoides* biomass subjected to lipid extraction followed by sodium hydroxide extraction at high temperature. The pyrolysis chromatogram of this residue was devoid of signals from soluble (poly)phenols and protein, but contained clear signals from (poly)saccharides as well as G-type lignin specific compounds: 4-(1-propenyl)guaiacol, *cis*-4-(2-propenyl)guaiacol and *trans*-4-(2-propenyl)guaiacol. (Figure 6A) (Kuroda et al., 1995; Kuroda et al., 1990). THM of the same residue further revealed the presence of two stereoisomers of 1-(3,4-dimethoxyphenyl)-1,2,3-trimethoxypropane, which are believed to be indicative of intact β-O-4 linkages, the main bond type in G-type lignin (Figure 6B) (Filley et al., 1999; Kuroda et al., 2002). In untreated biomass, pyrolysis products of the abovementioned compounds were often below the detection limit. For those samples for which pyrolysis did provide signals of three 4-propenylguaiacols, we computed the ratio between them and (poly)saccharide pyrolysis products (Figure 6C). Typical lignin pyrolysis products were only detected in two samples of *A. pinnata* biomass, although at low and varying levels. In *A. filiculoides* biomass we repeatedly detected lignin pyrolysis products at around 1.5% the intensity of (poly)saccharide pyrolysis products (Figure 6C). For comparison, in the DCM and NaOH extracted residue, lignin peaks were detected at 1.6% of the intensity of (poly)saccharide pyrolysis products. In *A. caroliniana* lignin pyrolysis and THM products did not reach the detection limit (Nierop et al., 2011) even after renewed analysis of the same sample. Lignin was also not detected in *A. nilotica* and *A. rubra*.

Copper oxide (CuO) of the biomass of *A. filiculoides* was performed to confirm the presence of G-type lignin. Upon CuO oxidation of *Azolla* biomass, the vanillyl phenols vanillin, vanillic acid and acetovanillone were produced, typical of G-type lignin. (Supplemental Table C). Small amounts of syringaldehyde and acetosyrin-

Table 3. Chemical composition of the whole *Azolla* biomass and the residues after lipid extraction and consecutive (poly)phenol extraction.

A.filiculoides	Mass (%)	Nitrogen[1] (%)	Carbon[1] (%)	Condensed tannins[1] (% CT-eq)
Biomass	-	3.7 ± 0.1	42.8 ± 0.4	5,3 ± 0,2
Lipid extracted residue	86 ± 0	3.7 ± 0.0	33.4 ± 0.4	4.7 ± 0.0
(Poly)phenol extracted residue	62 ± 5	3.3 ± 0.5	27.6 ± 1.7	0.8 ± 0.2
A.pinnata				
Biomass	-	2.9 ± 0.0	41.6 ± 0.4	8.0 ± 0.1
Lipid extracted residue	87 ± 0	2.9 ± 0.0	34.7 ± 0.3	8.3 ± 0.8
(Poly)phenol extracted residue	65 ± 0	2.6 ± 0.2	27.1 ± 0.1	0.7 ± 0.3

[1] Concentrations refer to the percentage of the initial biomass

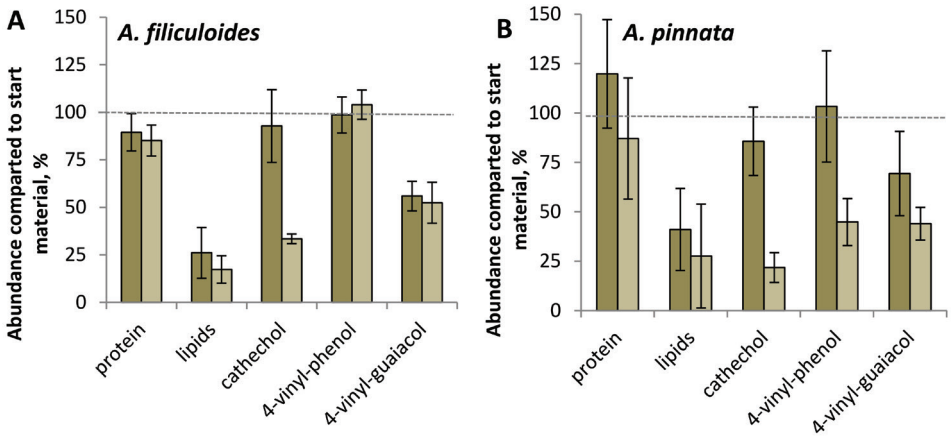


Figure 5. The effect of lipid and (poly)phenol extraction on the abundance of pyrolysis products indicative of proteins, lipids and (poly)phenols in residues of *A. filiculoides* (A) and *A. pinnata* (B). The abundance of pyrolysis products in lipid extracted (dark) and (poly)phenol extracted (light) residues is relative to the initial biomass (set at 100%).

gone were detected in a few samples at a ratio of 0.017±0.021 to vannilyl phenols (Supplemental Table C).

Confocal imaging of *A. filiculoides* biomass stained with propidium periodide shiff stain and destained for soluble phenolic compounds (Truernit et al., 2008), revealed the occurrence of insoluble (poly)phenols around the vascular tissue. The signal was especially located in the stem and less in the vascular loop in the leaf (Figure 6D).

Biosynthesis pathways in *Azolla filiculoides*

Transcriptional investment in (poly)phenol biosynthesis is high compared to *Arabidopsis*

In the *Azolla* genome we identified seven *PAL*-like genes, two *C4H*-like genes and seven *4CL*-like genes, that make up the starting point for (poly)phenol biosynthesis via the phenylpropanoid and flavonoid pathway (Supplemental Figure S4, Supplemental Table D). All AzfiPAL-like proteins had a high identity with known PAL from plants, i.e. 61 to 69%. The two proteins resembling the cytochrome P450 enzyme C4H were near-identical. The AzfiC4H-like protein were 68-83% identical to known C4H in plants and distinctly different from other *Azolla* CYP-like protein. Azfi4CL-like proteins showed more variation, with Azfi4CL-like1-4 being 37-41% identical and Azfi4CL-like 5-8 being 55-63% identical to 4CL from vascular plants.

Expression of *PAL*-, *C4H*- and *4CL*-like genes in *Azolla* sporophytes was high

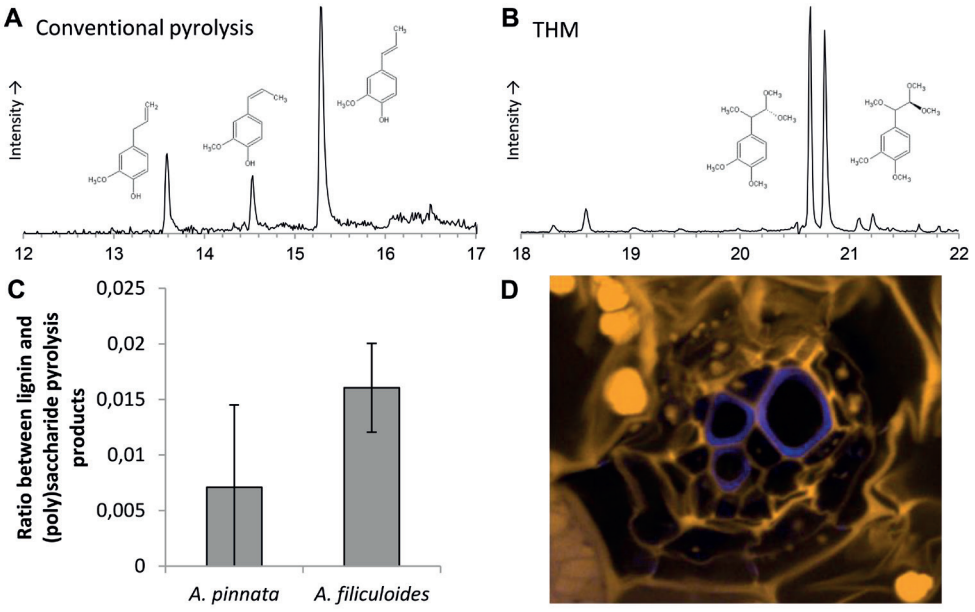


Figure 6. Identification of lignin in *Azolla* biomass (A,B,C) and localization of insoluble (poly) phenols (D). (A) Mass chromatogram (m/z 164) showing the three 4(propenyl)guiacol isomers characteristic of lignin obtained by conventional pyrolysis of *A. filiculoides* biomass subjected to DCM-extraction followed by sodium-hydroxide extraction (pH 12.5) at 95°C. (B) Mass chromatogram (m/z 181) showing *cis*- and *trans*-1-(3,4-dimethoxyphenyl)-1,2,3-trimethoxypropane characteristic of lignin obtained by THM of the same material. (C) Ratios between lignin and (poly)saccharide pyrolysis products of *Azolla* biomass. (D) Autofluoresce in the xylem cells of *A. filiculoides*. The image was taken using confocal scaning fluorescence microscopy by excitation at 405 nm and capturing emission at 505-515 nm (blue autofluorescence) and >560 nm (orange, propidium periodide shiff stain as in Truernit, et al. (2008).

when compared to 20 d old *Arabidopsis* plants (Coolen et al., 2016). The high expression of *PAL* genes in *Azolla* was mainly due to *AzfiPAL*-like 1, that reached an average expression level of 1185 reads per million of reads (RPM), making it the 56th most expressed gene in the *Azolla* diel RNA-seq experiment (Chapter 3). In comparison, *Arabidopsis AtPAL1* was the 2057th most expressed gene in similar diel RNA-seq experiment (Coolen et al., 2016).

Several isoenzymes were differentially expressed in the tissues (Supplemental Table D). For example, *AzfiPAL*-like 2 was predominantly expressed in the sporophyte, *AzfiPAL*-like 5,6 and 7 were mainly expressed in the megasporocarp and *AzfiPAL*-like 4 was predominantly expressed in the microsporocarp. Expression of *Azfi4CL*-like 8 was completely restricted to the spores. The diel expression patterns of *AzfiPAL*-like 1,2, *AzfiC4H*-like 1 and *Azfi4CL*-like 6,7 peaked at 2 h in the night, when plants were grown without nitrogen fertilizer. In contrast, expression of

Azfi4CL-like1,2 peaked during mid-day, whereas *Azfi4Cl*-like 3,4,5 were expressed evenly during the day (Supplemental Table D).

Identification of a highly expressed *AzfiLAR*-like gene

Annotation of flavonoid biosynthesis genes from *A. filiculoides* yielded four *CHS*-like genes, one *CHI*-like gene, one *FNSII*-like gene, two *F3H*-like genes, eight *F3'H*-like genes, two *DFR*-like genes, and one *LAR*-like gene (Figure 7, Supplemental data D). Additionally we identified one *FLS/ANS*-like gene with similarity to genes encoding FLS and ANS proteins in seed plants; phylogenetic analysis of 2OGD genes placed this *Azolla* gene within the branch of *ANS* and *FLS*, but before their diversification (Supplemental Figure S5). The placement of *AzfiF3'H*-like genes by phylogenetic analysis was not well-supported, in contrast to the conserved CYP *AzfiC4H*-like and *AzfiC3H*-like genes (Supplemental Figure S6). Annotation of the eight *AzfiF3'H*-like genes was therefore mainly based on their higher similarity to *F3'H* genes than to the *F3'5'H* and *F5H* genes. Furthermore the phylogenetic analysis suggested a divergence of F3'5'H from F3'H in seed plants, which was well supported by bootstrapping (Supplemental Figure S6).

Analysis of gene expression in the sporophyte revealed a high average expression of *AzfiCHS*-like genes compared to *Arabidopsis*, mostly by *AzfiCHS*-like 2 (1010 RPM) and *AzfiCHS*-like 3 (575 RMP). Although *AzfiCHI*-like is not highly expressed (100 RPM), its expression was high compared to *Arabidopsis*. Additionally, expression of *AzfiF3H*-like genes (260 RPM) and *AzfiF3'H*-like genes (230 RPM) suggest that in *Azolla* substrate is dominantly directed towards the production of dihydroflavonols with two hydroxyl groups on the B-ring. In *Arabidopsis*, expression levels of *AtFNSI* and *AtFLS* were higher than *AzfiFNSII*-like and *AzfiALS/FLS*-like, suggesting more flux towards the production of flavones and flavonols. In both *Azolla* and *Arabidopsis* the expression of *DFR*-like genes was high, promoting conversion of dihydroflavonols into leucoanthocyanidins. The *AzfiFLS/ANS*-like in *Azolla* was expressed as much lower levels than *AtANS* in *Arabidopsis*, in line with the absence of anthocyanidins in the extracts of *Azolla*. Instead *Azolla* highly expressed *AzfiLAR*-like at an average of 1362 RPM making it the most highly expressed secondary metabolism gene and the 44th most expressed gene in the *Azolla* diel RNA-seq. experiment. The expression pattern of *AzfiLAR*-like matched that of other highly expressed genes in the flavonoid biosynthesis pathway, characterised by a peak in gene expression at 8 h in the morning when plants were grown without nitrogen and an even expression during the day when grown with nitrogen constantly available in the medium (Supplemental Table D).

Phylogenetic analysis of the *AzfiLAR*-like gene suggest close relatedness functionally annotated *LAR* genes from gymnosperms and angiosperms, while clearly distinct from *DFR*-like genes and the fungal *LAR* outgroup (Figure 8).

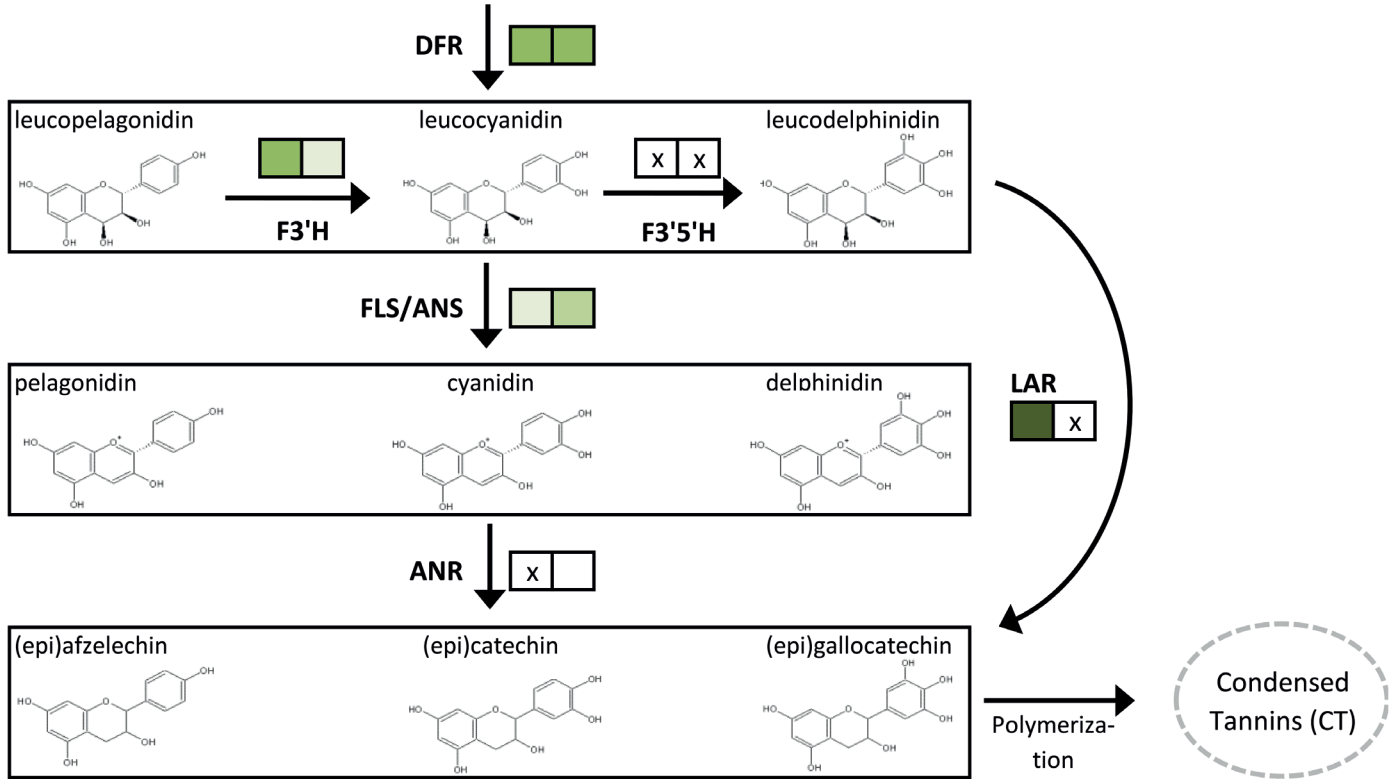
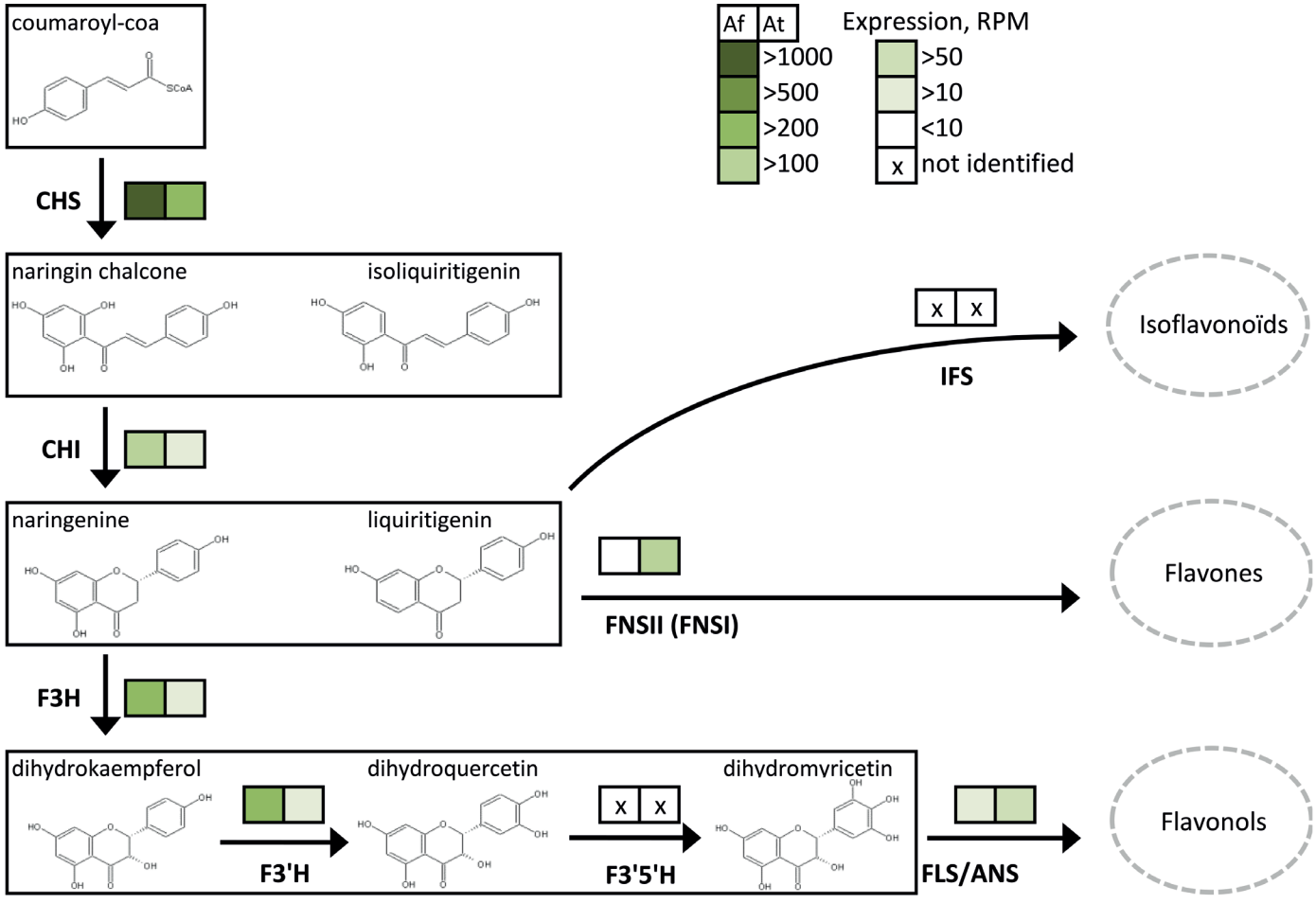


Figure 7. Occurrence and gene expression of enzymes from the flavanoid biosynthesis in *Azolla filiculoides* (Af) compared to *Arabidopsis thaliana* (At). *A. thaliana* seedling data was from Coolen et al, 2016 (Coolen, et al., 2016). Gene expression refers to the steady state accumulation of mRNA quantified by the number of reads matching it, the number of reads is normalized for each sample per million reads (RPM). The colour scheme is the sum of expression over the diel cycle of all identified isoenzymes.


coumaroyl-coa
CHS
naringin chalcone
isoliquiritigenin
CHI
naringenine
liquiritigenin
F3H
dihydrokaempferol
dihydroquercetin
dihydromyricetin
F3'H
F3'5'H
FLS/ANS
IFS
FNSII (FNSI)
Isoflavonoïds
Flavones
Flavonols
Af
At
>1000
>500
>200
>100
Expression, RPM
>50
>10
<10
x
not identified

Expression of the lignin biosynthesis pathway is halted after caffeoyl and feruloyl intermediates

Annotation of the phenylpropanoid pathway (leading up to lignin biosynthesis) in *A. filiculoides* yielded 26 identified genes encoding proteins with similarity to known enzymes for the synthesis of lignin and lignin intermediates (Figure 9). These included three *HQT/HCT*-like genes, one *C3H*-like gene, seven *COMT*-like genes, four *CCoAMT*-like genes, four *CRR*-like genes and four *CAD*-like genes. Phylogenetic analysis indicated a possible homology of AzfiF3'H-like protein to seed plant F5H, but due to the low confidence the presence of an *Azolla* F5H-like protein was left unresolved.

Given the high concentrations of dicaffeoylquinic acids in *Azolla* we investigated

Additional information Figure 7
In seed plants, coumaryl-Coa is converted to naringin or isoliquiritigenin chalcone via chalcone synthase (CHS), followed by conversion into naringenine or liquiritigenin by chalcone isomerase (CHI). Naringenine and liquiritigenin can then be converted by a set of hydroxylating enzymes into isoflavonoids, flavones and/or dihydroflavonols. The substrate for isoflavonoid biosynthesis, 2-hydroxy-isoflavanone, is produced by the cytochrome P450 (P450) enzyme isoflavonoid synthase (IFS). Flavones are also generally produced by a P450 enzyme, i.e. flavanone synthase II (FSII), although in some angiosperms, including Arabidopsis, the hydroxylation is done by the 2-oxoglutarate-dependent (2OGD) enzyme flavanone synthase I (FSI) (Kawai, et al., 2014). Dihydroflavonols are produced by the . enzyme flavanone 3-hydroxylase (F3H). The hydroxylation pattern of the dihydroflavonol B-ring depends on the action of two P450 enzymes: flavonoid 3-hydroxylase (F3'H) and flavonoid 3,5-hydroxylase (F3'5'H), which produce dihydroflavonols with two and three hydroxyl groups on the B-ring, respectively. The conversion of these dihydroflavonols into flavonols is done by the . enzyme flavonol synthase (FLS). To produce anthocyanidins dihydroflavonols are first reduced to leucoanthocyanidins by dihydroflavanonol reductase (DFR) and hydroxylated by the . enzyme anthocyanidin synthase (ANS). Lastly the flavan-3-ol monomers of CTs can be produced by two ways, either by reduction of anthocyanidins by anthocyanidin reductase (ANR) or by direct reduction of leucoanthocyanidins by leucoanthocyanidin reductase (LAR).Naringenine and liquiritigenin can then be converted by a set of hydroxylating enzymes into isoflavonoids, flavones and/or dihydroflavonols. The substrate for isoflavonoid biosynthesis, 2-hydroxy-isoflavanone, is produced by the cytochrome P450 (P450) enzyme isoflavonoid synthase (IFS). Flavones are also generally produce d by a P450 enzyme, i.e. flavanone synthase II (FSII), although in some angiosperms, including Arabidopsis, the hydroxylation is done by the 2-oxoglutarate-dependent (2OGD) enzyme flavanone synthase I (FSI) (Kawai et al., 2014). Dihydroflavonols are produced by the enzyme flavanone 3-hydroxylase (F3H). The hydroxylation pattern of the dihydroflavonol B-ring depends on the action of two P450 enzymes: flavonoid 3-hydroxylase (F3'H) and flavonoid 3,5-hydroxylase (F3'5'H), which produce dihydroflavonols with two and three hydroxyl groups on the B-ring, respectively. The conversion of these dihydroflavonols into flavonols is done by the enzyme flavonol synthase (FLS). To produce anthocyanidins dihydroflavonols are first reduced to leucoanthocyanidins by dihydroflavanonol reductase (DFR) and hydroxylated by the . enzyme anthocyanidin synthase (ANS). Lastly the flavan-3-ol monomers of CTs can be produced by two ways, either by reduction of anthocyanidins by anthocyanidin reductase (ANR) or by direct reduction of leucoanthocyanidins by leucoanthocyanidin reductase (LAR).

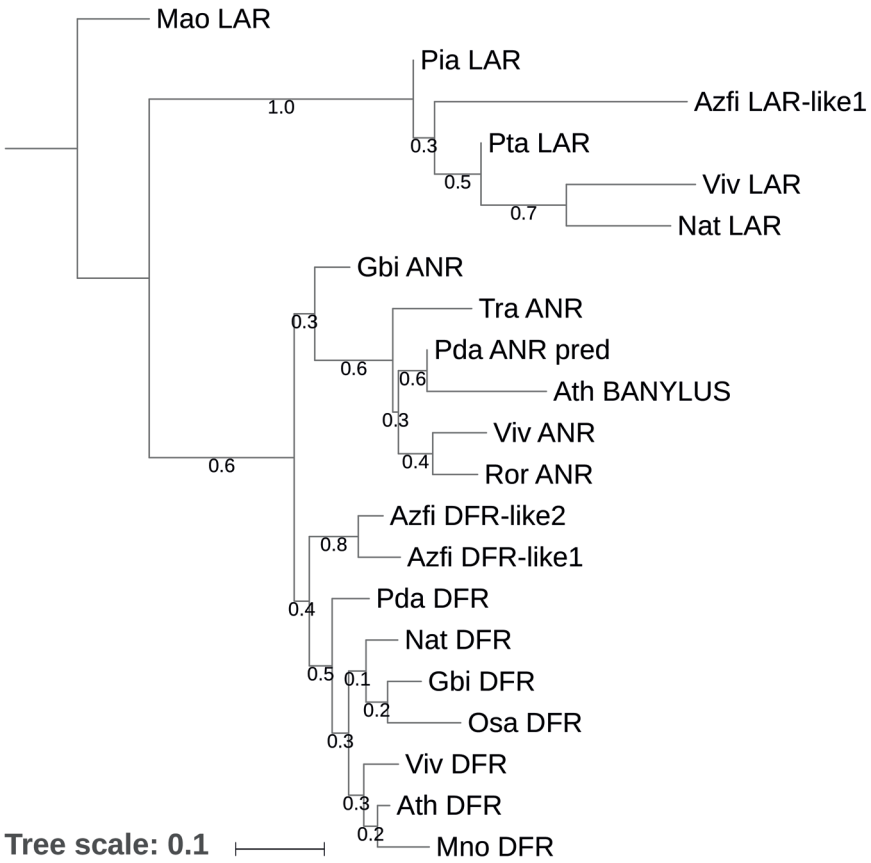


Figure 8. Phylogenetic tree of oxioreductases (DFR and LAR) identified in *Azolla* (Azfi). Alignment was performed at the protein level using MAFFT. The consensus tree was constructed by Fyml using 100 times bootstrapping; bootstrap values are indicated for each node. Abbreviations of reference plant species are: *Picea abies* (Pia), *Pinus taeda* (Pta), *Vitis vinifera* (Viv), *Narcissus tazetta* (Nat), *Ginkgo biloba* (Gbi), *Triticum urartu* (Tra), *Phoenix dactylifera* (Pda), *Arabidopsis thaliana* (Ath), *Rosa rugosa* (Ror), *Oryza sativa* (Osa) and *Morus notabilis* (Mno). A LAR gene of the fungi *Magnaporthe oryzae* (Mao) was used as the outgroup. Sequence accession numbers are provided in Supplemental Table E.

the amino acid sequence of AzfiHCT/HQT-like protein by alignment to HCT/HQT enzymes known to synthesize dicaffeoylquinic acids (Figure 10) (Lallemand et al., 2012; Moglia et al., 2014). All AzfiHQT/HCT-like proteins had the Trp-733 and Gly-238 residues important for hydroxycinnamoyl binding and conserved in many BAHD members, as well as the Thr-660 residue, important for binding of the acyl acceptor. The catalytic His-233 was present in all AzfiHCT/HQT-like proteins. In addition catalytic His-59 was present in AzfiHCT/HQT-like 2 and 3 (Figure 10).

The preference for shikimic acid by HCTs in angiosperms was attributed to Leu-400 and Phe-402 residues and a Met-231 residue was deemed specific for HCT, ensuring less efficient acceptance of feruloyl moieties (Lallemand et al., 2012). AzfiHQT/HCT-like 3 contained all of these residues and therefore resembles HCT In contrast, AzfiHQT/HCT-like1,2 had the HCT-specific Leu-765, but also the HQT-specific Tyr-767, which was also observed in HCT-like genes from *Pinus*, *Physcomitrella* and *Selaginella* (Figure 10).

AzfiHCT/HQT-like protein contained a standard His-234 residue, instead of an Asn-234 residue that boosting the production of dicaffeoylquinic acid by a Coffea HCT mutant (Lallemand et al., 2012). AzfiHQT/HCT-like 2,3 further had a standard Tyr-633 residue, instead of a His-633 residue shown to be necessary for dicaffeoylquinic acid production in Solanaceous species (Moglia et al., 2014). AzfiHQT/HCT-like 1 had an Isoleucine at this location (Figure 10).

AzfiHCT/HQT-like and *CCoAMT*-like genes were moderately expressed in *Azolla* sporophytes. Genes involved in the consecutive enzymatic steps to produce lignin, i.e. *AzfiCRR*-like, *AzfiCAD*-like and *AzfiCOMT*-like, were expressed at much lower rates (Figure 9). This differed in *Arabidopsis* leaves where expression levels of the upstream *AtHCT*, *AtC3H* and *AtCCoAMT* were, firstly, lower compared to *Azolla* and secondly, comparable to expression levels of downstream *AtCRR*, *AtCAD* and *AtCOMT.* Hence in *Azolla* the phenylpopanoid pathway is not geared to the production of lignin, but rather to produce caffeoyl, and feruloyl, intermediates.

The mostly highly expressed isoenzymes in the lignin pathway: *AzfiCRR*-like 2, *AzfiCAD*-like 1 and *AzfiCOMT*-like 4 are all expressed evenly during the day and their expression pattern is not affected when nitrogen is supplied via the growth medium. This diel expression pattern is shared by *AzfiC3H*-like and *AzfiHQT/HCT*-like 3. The expression pattern of *AzfiHQT/HCT*-like 2 and *AzfiCCoAMT*-like 4 slightly peaked at late night/early-morning, similar to key enzymes in the flavonoid pathway. In contrast, the expression of *AzfiHQT/HCT*-like 1 peaked sharply during the late night, mimicking the expression pattern of the highly expressed *AzfiPAL*-like 1.

Discussion

Abundant soluble CTs are likely synthetized by a highly expressed *AzfiLAR*-like gene

The presence of CTs in *A. filiculoides* has been previously determined using histochemical methods (Pereira and Carrapiço, 2007) and Fasakin et al. (2001) reported 4.26% CTs in *A. africana* using the acid butanol assay (Fasakin and Balogun, 2001). Extraction and characterization of CTs from *Azolla* has not yet been per-

formed. This study indicates that CTs or proanthocyanidins, were the most abundant soluble (poly)phenol in both A.*filiculoides* and *A. pinnata*: tannin extracts accounted for 4.0 and 5.8% of the dry weight respectively. Total phenol measured in the extracted tannin fraction equaled 60-63% TA-eq. However, the acid-butanol assay suggests tannin fractions are of high purity (97-101% CTs), which is supported by carbon and nitrogen content and pyrolysis data. Hence the reactivity of extracted *Azolla* CTs to the Folin Ciocalteu reagent was likely below of what has been reported for purified tannins, i.e. 67-104% (Kraus et al., 2003)

Soluble CTs represent on average 68% of the CTs in forage plants, whereas the remaining fraction is either protein-bound or fiber-bound (Terrill et al., 1992). Sample pretreatment and the solvent used are known to influence extraction efficiency of CTs (Palmer et al., 2000). With the procedures used in this study, 81-86% of the CTs in the biomass were solubilized and 9-15% of protein and/or fiber-bound CTs were measured in the residue after extraction. The small underestimation (4-5%) may be due to the formation of protein-(poly)phenol complexes and, generally, the fiber-bound fraction may be underestimated by 10-25% by the acid butanol assay (Adamczyk et al., 2014; Makkar et al., 1999). Nonetheless we show here that the majority of CTs in *Azolla* are soluble and therefore that their dominant role is not structural (Nierop et al., 2011).

THM indicated the structure of *Azolla* CTs consists predominantly of (epi)catechin units, which is supported by identification of the (epi)catechin monomer, dimer and trimer by LC-MS. A smaller contribution of (epi)afzelechin units was detected at 0.26-0.28 of the relative abundance of catechin THM products, whilst contributions of (epi) gallocatechin were negligible: <0.01 of the relative abundance of (epi)catechin THM products. The distribution of CT derived THM products observed in *A. filiculoides* and *A. pinnata* strongly resembles those reported for *A. caroliniana* (Nierop et al., 2011) .

CTs were by far the dominant product of the flavanoid biosynthesis pathway. Glucosylated flavonols were minor soluble (poly)phenolic compounds in *A. pinnata*. In contrast, *A. filiculoides* only contained trace amounts of kaempferol-glucosides. Previously reported flavones: homoorientin and saponarin (Teixeira et al., 2001), were not identified in *A. filiculoides* nor in *A. pinnata*. Similar to CTs, flavonols detected in *A. pinnata* preferentially contained two hydroxyl groups at the B-ring (quercetin) over one hydroxyl group (kaempferol), whilst no flavonol containing three hydroxyl groups (myricetin), was not detected. Since both CTs and flavonols are derived from dihydroflavonols, this shared preference infers the existence of an active F3'H enzyme or enzyme with F3'H function, whilst suggesting a lack, or low activity, of a F3'5'H enzyme.

In light of *Azolla*'s high protein content and favorable amino acid profile a major role of CTs in preventing herbivory is highly plausible; CTs protect various plant

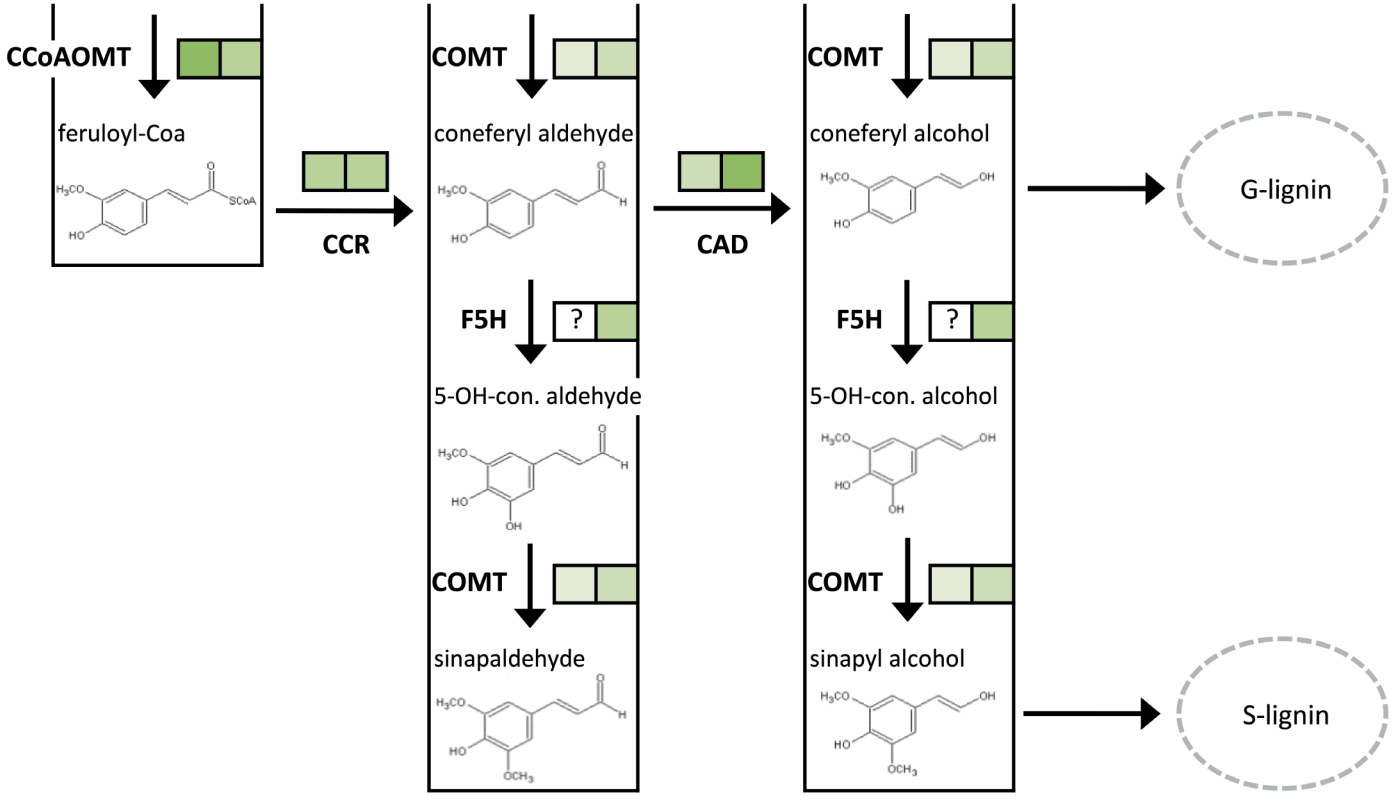
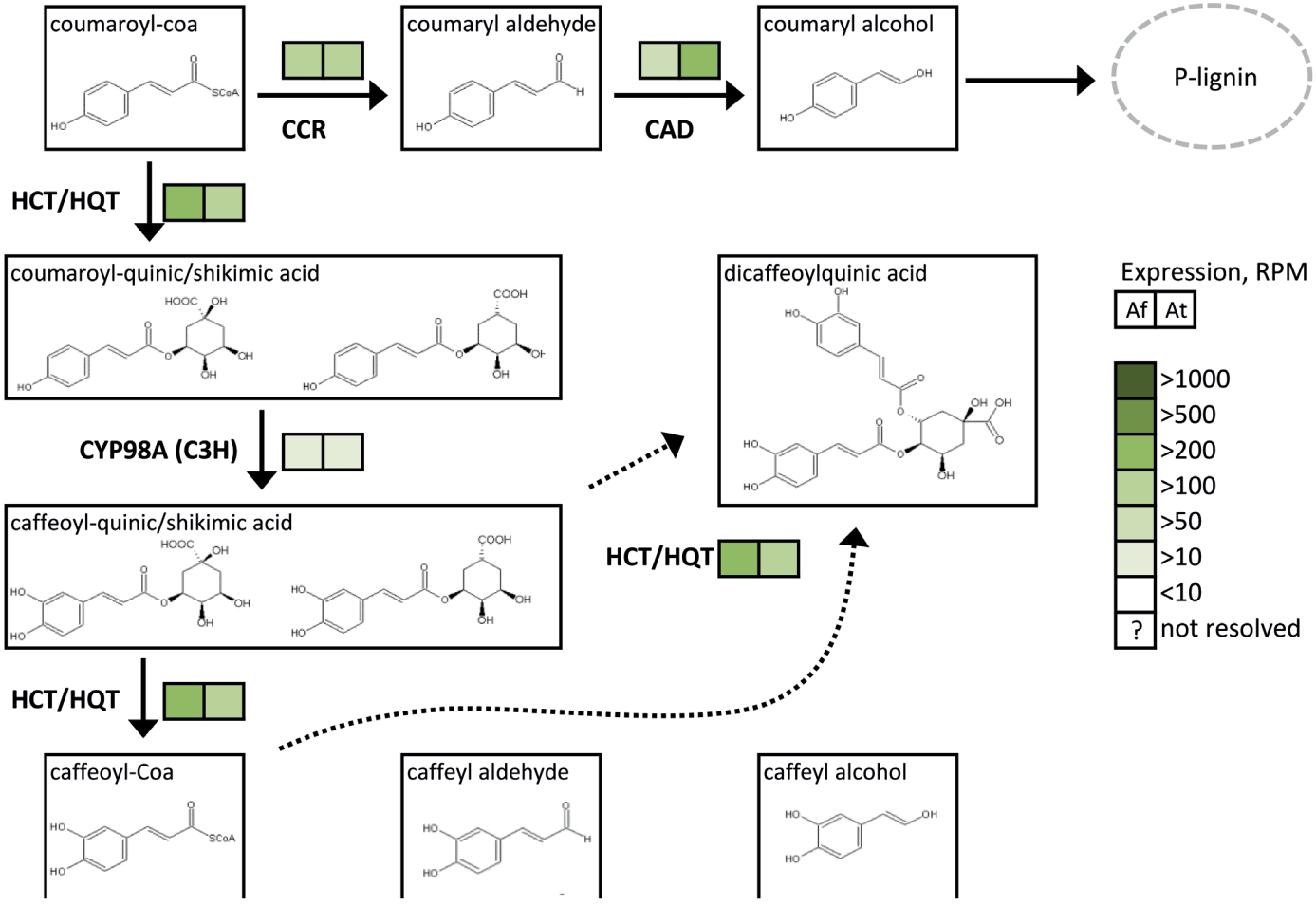


Figure 9. Occurrence and gene expression of enzymes from the flavanoid biosynthesis in *Azolla filiculoides* (Af) compared to *Arabidopsis thaliana* (At). Gene expression was quantified as in Figure 7.


coumaroyl-coa
CCR
coumaryl aldehyde
CAD
coumaryl alcohol
P-lignin
HCT/HQT
coumaroyl-quinic/shikimic acid
CYP98A (C3H)
caffeoyl-quinic/shikimic acid
dicaffeoylquinic acid
HCT/HQT
HCT/HQT
caffeoyl-Coa
caffeyl aldehyde
caffeyl alcohol
Expression, RPM
Af
At
>1000
>500
>200
>100
>50
>10
<10
?
not resolved

Additional information figure 9
Coumaryl-Coa can be directly converted to the P-lignin monomer coumaryl alcohol via cinnamoyl-Coa reductase (CRR) and cinnamyl-alcohol dehydrogenase (CAD). Alternatively, a quinic acid or shikimic acid group can be added by quinate O-hydroxycinnamoyltransferase (HQT) and shikimate O-hydroxycinnamoyltransferase (HCT). The coumarylquinic acid or coumarylshikimic acids are then substrate for the P450 enzyme p-coumarate 3-hydroxylase (C3H), yielding caffeoylquinic acid (chlorogenic acid) or caffeoylshikimic acid. The latter can be converted again by HQT and HCT, this time in a reverse reaction, replacing the quinic acid or shikimic acid by Coenzyme A, yielding caffeoyl-CoA. Caffeoyl-CoA can then be converted to feruloyl-CoA by caffeoyl-CoA O-methyltransferase (CCoAMT). Feruloyl-Coa is then converted to coneferyl aldehyde by CRR and to the G-lignin monomer coneferyl alcohol by CAD. In angiosperms the P450 enzyme ferulate 5-hydroxylase (F5H) can convert the coneferyl group into a 5-hydroxy-coneferyl group. Methylation of the 5-hydroxy-coneferyl group by caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase (COMT) yields a sinapyl-groups, that can be converted into S-lignin. The dashed line indicates possible routes for dicaffeoylquinic acids by HCT/HQT enzymes.

species, in particularly woody species with long-lived leaves, from herbivory by vertebrates and insects (Chapter 4) (Barbehenn and Peter Constabel, 2011; Coley, 1988; Stafford, 1991). However, since anti-herbivory is not universally observed for different combinations of CT structures and herbivore species, other or additional functions may explain the high investment of carbon (Barbehenn and Peter Constabel, 2011; Ayres et al., 1997). In the case of *Azolla*, tannins have been shown to accumulate in simple hairs located within the leaf cavity harboring the nitrogen-fixing *Nostoc Azolla*e (Pereira and Carrapiço, 2007). There, CTs may have a role in regulating the growth of the cyanobacteria symbiont, as was observed for *Frankia* symbionts in *Casuarina glauca* (Laplaze et al., 1999). Also CTs from trees have been shown to effectively reduce growth of cyanobacteria *Microcystis aeruginosa* (Shimada et al., 2010).

The less abundant flavonols in *A. pinnata* are known anti-oxidants, generally believed to play a role in protecting the plant from reactive oxygen species (ROS) in relation to stress factors, such as UV light (Pollastri and Tattini, 2011). More work is needed to understand the roles of flavonoid monomers in *Azolla* and whether their occurrence is specific for *A. pinnata*, or can be invoked in *A. filiculoides*. by environmental cues. For example, the flavones homoorientin and saponarin have not been detected here, but were detected in *A. filiculoides*, grown outdoor in Portugal (Teixeira et al., 2001).

High abundant CTs in *Azolla* were likely produced via the AzfiLAR-like enzyme. AzfiLAR-like was homologous to functionally annotated LAR protein from gymnosperms and angiosperms, while it was evolutionary distant from ANR and DFR reductases. Its diel gene expression pattern matched the gene expression patterns of other genes that were identified as part of the flavonoid biosynthesis pathway. In an alternative pathway, plants convert leucoanthocyanidins into anthocyanidins via ANS and then reduce anthocyanidins to epiflavan-3-ols by ANR (Peng et al., 2012).

Figure 10. Protein sequence alignment of conserved regions of HCT/HQT-like protein in plants. A fungi BAHD acyltransferase was included for comparison. The alignment was performed with MAFFT; consensus characters are coloured. Boxed regions refer to catalytic His residues (1,3), Met residue specific for HCT (2), crucial residue for diCaQ production in Coffee and Tomato (4,6), Gly and Trp residues important for hydroxycinnamoyl binding (5,8), Thr residue, involved in binding of the acyl acceptor (7) and residues important for specificity to shikimate/quinate (9,10) (Lallemand, et al., 2012). Sequence accession numbers are provided in Supplemental Table E.

In this study, we detected no anthocyanidins in *Azolla* biomass. However, *Azolla* is known to produce anthocyanins during biotic and abiotic stress, inferring the presence of an ANS enzyme (Ismail and Mohamed, 2010; Kösesakal, 2014). An *AzfiANS/FLS*-like gene was identified, but was expressed at very low levels, in line with the absence of anthocyanidins in *A. filiculoides*. ANR homologues were not identified using bioinformatics tools, but the epicatechin monomer was detected using LC-MS. However, enzyme extracts of the fern *Dryopteris pycnopteroides* exhibited ANR activity and leaf extracts contained both catechin and epicatechin monomers (Peng et al., 2012). Hence it cannot be ruled out that *Azolla* is capable of producing proanthocyanidins via the ANS/ANR pathway when subjected to different growth conditions or when introducing mutations in the *AzfiLAR*-like gene. In our growth conditions CT production is most likely via *AzfiLAR*-like, which was the most highly expressed secondary metabolism gene of

A. filiculoides; only genes from photosynthesis, calvin cycle and CSA cycle were more highly expressed, underscoring the importance of CT biosynthesis in *Azolla*.

Dicaffeoylquinic acids are abundant in *Azolla*

Dicaffeoylquinic acids were the second major soluble phenolic compound in *Azolla*, accompanied by lesser contributions of caffeoylquinic acid (chlorogenic acid) and tricaffeoylquinic acid. Chlorogenic acid was identified earlier (Ishikura, 1982; Teixeira et al., 2001), but dicaffeoylquinic acid was not yet recognized as a major constituent. The presence of esters of caffeic acid and sugars in *Azolla* have been proposed previously, including caffeic acid 3,4-diglucose, 6-(3'-glucosylcaffeoyl)-aesculetin and the glucose 1,6-diesters of caffeic- and chlorogenic acid, but were not observed in this study coumarylquinic acids were detected in both *A. filucloides* and *A. pinnata*. *A. filiculoides* further contained feruloylquinic acid. However, the abundance of these compounds was minor compared to the caffeoylquinic acids. Hence *Azolla* has a high preference of incorporating caffeic acids into soluble phenolic compounds, with dicaffeoylquinic acid as the main product.

Dicaffeoylquinic acids are synthesized by most plants in quantities <0.1 w/w% (Karaköse et al., 2011; Padda and Picha, 2008). However, in specific tissues of selected plants dicaffeoylquinic acids were shown to accumulate, for example to 0.8-1.3% in green coffee beans (Farah et al., 2005), 2.0% in the immature flowers of artichoke plants (Lattanzio et al., 2009) and 3.1-3.4% in leaves of *Ilex paraguariensis* used for the preparation of mate tea (Bravo et al., 2007). By combining data from LC- MS measurements with the mass and the total phenol content of extracts (Table 2) we can estimate a minimum dicaffeoylquinic acid content of 0.7% and 1.1% in the biomass of *A. filiculoides* and *A. pinnata*, respectively, of which 50-60% was retrieved in the ethyl acetate fraction. However, this estimation depends heavily on the assumption that the extracted dicaffeoylquinic acids have a similar molar reactivity to the folin ciocalteu reagent as previously reported for pure chlorogenic acid (Everette et al., 2010). Hence, future research is needed to provide an exact quantification of dicaffeoylquinic acids in *Azolla*. Nonetheless, results indicate that dicaffeoylquinic acids accumulate in *Azolla* to high levels scarcely encountered in the plant kingdom.

Protection against damaging UV-light has been considered the main function caffeoylquinic acids and these compounds have been shown to accumulate upon UV-B exposure in seed plants (Demkura et al., 2010; Moglia et al., 2014; Izaguirre et al., 2007). Given that *Azolla* plants analyzed in this study were grown without UV-B light, dicaffeoylquinic acids may accumulate to even higher levels in plants grown outdoors.

Remarkable for *Azolla* is the dominance of dicaffeoylquinic acids over monocaffeoylquinic acids. In green coffee beans, artichoke and *Ilex* leaves a high content

of dicaffeoylquinic acid was accompanied by an equal or higher amount of mono-caffeoylquinic acids (Farah et al., 2005; Lattanzio et al., 2009; Bravo et al., 2007). In Coffee and tomato an HCT and HQT enzyme have been shown to produce di-caffeoylquinic acids as a side product to chlorogenic acid. HCT/HQT-like proteins are versatile BAHD acyltransferases that accept a variety of substrates and catalyse both forward, i.e. esterification of a hydroxycinnamoyl CoA with a shikimic/quinic acid and reverse reactions, i.e. production of a hydroxycinnamoyl CoA from an shikimate/quinate ester. While a HCT enzyme in coffee produced dicaffeoylquinic acids from caffeoyl-Coa and chlorogenic acid (Lallemand et al., 2012), a HQT enzyme from tomato produced dicaffeoylquinic acids from two chlorogenic acids (Moglia et al., 2014).

The low ratio between mono- and dicaffeoylquinic acids in *Azolla* suggests the presence of an HCT/HQT-like enzyme with high preference for producing di-caffeoylquinic acid. Of the HCT/HQT-like enzymes identified in *Azolla*, *AzfiHQT/HCT*-like 1 was highly expressed and had a very distinct expression profile from genes identified in upstream lignin biosynthesis, suggesting it may not produce precursors that directly feed into lignin biosynthesis. *AzfiHQT/HCT*-like 1 was one of the few (poly)phenol biosynthesis genes with a similar expression profile as the highly expressed *AzfiPAL*-like 1 that ensures metaboplic flux through the phenylpropanoid and flavanoid pathways. Typical amino acid residues boosting dicaffoylquinic acid synthesis in an Coffee HCT (Asn-234) and Tomato HQT (His-633) were not present in AzfiHQT/HCT-like 1, but an Ile-633 residue was present instead of the standard Tyr-296 residue. Biochemical investigation is needed to determine if AzfiHCT/HQT-like 1, or another enzyme, is responsible for the high dicaffeoylquinic acid production in *Azolla*. Given that coffee HCT and tomato HQT genes have independently evolved to produce dicaffeoylquinic acid via different mechanisms, it is plausible that *Azolla* HCT/HQT-like genes have also, independently, evolved this capability.

Azolla is capable of lignin biosynthesis via a conserved pathway

Analysis of *Azolla* biomass and residues after (poly)phenol extraction revealed the presence of insoluble coumaric and ferulic acid esters in *Azolla*. *A. filiculoides* contained both more coumaric (2 fold) as ferulic (5 fold) acid compared to *A. pinnata* and these compounds were mostly insoluble. Hence whereas in *A. pinnata* biosynthesis is geared towards production of more soluble CTs and hydroxycinnamic acids, *A. filiculoides* produces more insoluble phenolic compounds.

Coumaric acid biosynthesis is the gateway to (poly)phenol biosynthesis, comprising of PAL, C4H and 4CL enzymes and likely pre-dates land plant evolution (Muradov et al., 2014). Several isoenzymes of PAL and 4Cl are already found in the moss *Physcomitrella*, suggesting early differentiation of functions (Tohge et al., 2013). In

Azolla we identified seven *AzfiPAL*-like and eight *Azfi4Cl*-like genes, with diel and tissue specific expression patterns. Two AzfiC4H-like protein had highly conserved amino acid sequences with C4H enzymes from seed plants, in accordance with the low diversification of the C4H enzyme during land plant evolution (Tohge et al., 2013).

The presence of ferulic acids in *A. filiculoides* is in line with moderate expression of *AzfiCCoaOMT*-like genes, which may produce feruloyl-CoA from caffeoyl-CoA. Incorporation of ferulic acid into feruloylquinic acid was observed for *A. filiculoides*, but its extraction did not alter ferulic acid abudance as detected by pyrolysis. Instead more ferulic acids were removed by lipid extraction, which could be explained by esterification of feruloyl-CoA into soluble alkyl ferulates, as in the barks of various trees and during wound healing in potato (Graça, 2010). Insoluble feruloyl compounds may derive from ω-feruloyloxypalmitic acid, which is part of the insoluble suberin polymer (Graça, 2010; Lotfy et al., 1994).

In contrast to previous reports on *A. caroliniana* (Nierop et al., 2011) we repeatedly observed a low amounts of G lignin in *A. filiculoides*. The amount of vanillin produced from *A. filiculoides* by CuO oxidation, i.e. 0.64 mg g^{-1}, was lower than previously reported for leaves of seed plants (1.00-2.17 mg g^{-1}), but comparable to some mosses, including *Polytrichum commune* (0.68 mg g^{-1}), *Selaginella sp.* (0.52 mg g^{-1}) and *Equisetum fluviatile* (0.40 mg g^{-1}) (Logan and Thomas, 1985). All genes required for G-lignin biosynthesis were identified in the *A. filiculoides* genome and were lowly expressed in the sporophyte. Many of these enzymes were already present in mosses and lycophytes, including HCT, C3H, CCoAMT, CRR and COMT-like genes (Tohge et al., 2013). Whilst biosynthesis of G-lignin has been reported in lycophytes, biosynthesis of S-lignin was until recently considered specific for angiosperms, where it relies on the CYP enzyme F5H (Weng and Chapple, 2010; Nelson, 2006; Espiñeira et al., 2011). Here we obtained miniscule amounts of syring aldehyde (0.023 mg g^{-1}) and acetonesyringone (0.005 mg g^{-1}) by CuO oxidation of *A. filiculoides* biomass, whereas the concentration of syring aldehyde is 50-250 fold higher in angiosperm leaves and barks (Logan and Thomas, 1985). Small amounts of S-lignin products have also been previously observed in the ferns *Ceratopteris cornuta* and *Dennstaedtia bipinnata*, and in gymnosperms, whilst higher concentrations have been detected in the spikemoss *Selaginella* (Logan and Thomas, 1985; Espiñeira et al., 2011). In *Selaginella*, it has been shown that S-lignin is synthesized via an independently evolved enzyme (Weng and Chapple, 2010; Espiñeira et al., 2011). Phylogenetic analysis of CYP genes performed here suggested early diversification of F5H from F3'H, but provided insufficient confidence to assign a F5H-like gene in *Azolla*. Additionally it can be questioned to what extend the observed syringyl-phenols can have another origin than S-lignin. Therefore further research is needed to confirm S-lignin biosynthesis in *Azolla*.

It can be concluded that *A. filiculoides* contains a conserved pathway for G-lignin production. However, *Azolla* incorporates only a minimal amount of lignin in its biomass. This may be due to the aquatic habitat of *Azolla* and not be representative of ferns: Espiñeira et al. (2011) remarked on low lignin-like content of plants thriving in aquatic habitats, including *Isoetes fluitans*, *Isoetes histrix*, *Ceratopteris thalictroides*, *Ceratopteris cornuta* and *Posidonia oceanica* (Espiñeira et al., 2011). Since *Azolla* grows horizontally and is completely supported by the water surface it does not need the strong fiber crosslinks required for vertical growth. A low lignin content would further permit the remarkably rapid abscission of branches, which is believed to facilitate its vegetative propagation (Cohen et al., 2014; Uheda and Nakamura, 2000).

However, to what extend other (poly)phenolic compounds comprising of ferulic, and especially, coumaric acids take over the structural function of lignin, or perform different functions, is yet unknown. Fluoresence confocal microscopy suggested localization of insoluble polyphenols in the xylem, consistent with G-lignin preferential deposition in tracheary elements of seed plants (Chapple et al., 1992; Nakashima et al., 2008).

Inventions in flavonoid biosynthesis in the common ancestor of ferns and seed plants?

In contrast to lignin biosynthesis, many key inventions in the flavonoid pathway seem to have occurred after lycophytes and ferns diverged. A CHS-like enzyme is still reported in *P. patens* and lycophyte *S. molendorfi*, but CHI was not detected (Tohge et al., 2013). In *Azolla* we found one *AzfiCHI*-like gene as in monocots (Tohge et al., 2013). The core flavonoid pathway further comprises of three enzyme families, i.e. the 2-oxoglutarate-dependent dioxygenases (2OGD), the Cytochrome P450 hydroxylases (CYP) and oxioreductases (Nelson, 2006; Kawai et al., 2014). The expansion of the 2OGD and CYP enzyme families is believed to be an important driver for the vast diversity of secondary metabolites observed in land plants (Nelson, 2006; Kawai et al., 2014). In *Azolla* we encountered two types of 2OGD enzymes: F3H-like enzymes that likely convert naringenine into gene dihydroflavnonols and an FLS/ANS-like enzyme with similarity to both FLS and ANS. *F3H*-like genes were not found in the moss *P. patens* nor the lycopod *S. molendorfii* (Tohge et al., 2013; Kawai et al., 2014). Several ANS enzymes have been shown to possess FLS activity and it was suggested that the enzymes diverged before the evolution of angiosperms and gymnosperms (Kawai et al., 2014). Polygenetic placement of AzfiFLS/ANS-like protein suggests divergence of FLS and ANS after the evolution of the Salviniales.

Of the hydroxylating CYP enzymes, we could identify one *FNSII*-like (*CYP93B*) gene and eight *F3'H*-like genes (*CYP75A*), but no *IFS*-like gene (*CYP93C*) responsible for isoflavanoid production. The lack of isoflavonoids in *Azolla* is not surprising,

as these are nearly exclusively found in legumes (Nelson, 2006). Flavones, on the other hand have been identified previously in *A. filiculoides* inferring the presence of a FNSII-like gene or a gene with similar function (Teixeira et al., 2001). Although the placement of *AzfiF3'H*-like genes by phylogenetic analysis was generally not well supported, results clearly suggest the diversification of F3'H and F3'5'H after the evolution of the Salviniales. Furthermore, the presence of F3'H-like and not F3'5'H-like genes in *Azolla* is strongly supported by the dominance of flavonoids with two hydroxyl groups at the B-ring and the lack of flavonoids with three hydroxyl groups at the B-ring.

Although flavanoid 2ODG and CYP homologues have not been found in *S. molendorfii* and *P. patens*, monomeric flavonoids are detected in some other mosses and algae (Stafford, 1991; Goiris et al., 2014). Given the disparate distribution of flavonoid groups within the Plantae Kingdom, it is very likely that crucial enzymes have evolved more than once, or that numerous losses have occurred (Stafford, 1991). For example, in an Arctic moss both F3'H and F3'5'H were identified, but were suggested to have evolved independently from the F3'H in ferns, gymnosperms and angiosperms (Liu et al., 2014).

Opposed to monomeric flavonoids, CTs have only been identified in ferns, but not in mosses, lycophytes or algae (Stafford, 1991; Goiris et al., 2014; Hwang et al., 1990; Jones and Firn, 1979). We identified a highly expressed *AzfiLAR*-like gene homologous to functionally annotated *LAR* genes from gymnosperms and angiosperms, while clearly distinct from *DFR*-like genes. Further characterization is needed both of the enzymes identified by orthology searches in this study and of the 30% of proteins that lack annotation in the genome of *A. filiculoides* to fully understand flavonoid biosynthesis. Diversification of oxidoreductases into LAR, however, was a crucial invention during fern evolution, likely affording efficient protection from herbivores.

Implications for an *Azolla* production chain

To develop *Azolla* as a high quality feed, reducing the high levels of CTs may be necessary. This may be achieved by processing the harvested biomass and/or by breeding and selecting for strains with reduced CT content. The LAR enzyme is a very promising breeding target in this regard, as it is the likely route for CT production in *Azolla* and the only *AzfiLAR*-like gene in the *Azolla* genome, yet it cannot be ruled out that its knockdown may redirect the metabolic flux to FLS/ANS and ANR resulting in the production of epicatechin based CTs. Reduced synthesis of CTs will very likely also render *Azolla* more vulnerable to herbivores and possibly deter growth by inhibiting yet unknown functions of these compounds. Hence possible advantages in product quality need to be weighed against disadvantages during cultivation.

Whereas the high CT content of *Azolla* is a possible caveat in the development of a sustainable production chain, the relatively high content of dicaffeoylquinic acids may provide additional value. Dicaffeoylquinic acid derivatives were researched as a potential natural drug for the treatment of various diseases, including the reduction of diabetes type 1 in mice (Tong et al., 2015), inhibition of the integration of the HIV virus in human cell lines (Robinson et al., 1996) and the prevention liver damage in rats (Choi et al., 2005). Artischoke extracts, rich in dicaffeoylquinic and chlorogenic acids, are commercialized as drugs, mainly for the treatment of liver diseases (Lattanzio et al., 2009). Hydrolysis of dicaffeoylquinic acids can alternatively yield caffeic acids used as formulation stabilizers in food, cosmetic and pharmaceutical industry and the synthesis of thermoresistant bioplastics (Silva et al., 2014; Chauzar et al., 2012).

The total phenol content of *Azolla* was shown to depend on growth conditions and species (Chapter 4). *Azolla* plants analyzed here grew relatively slowly, i.e. 3.4±0.3 and 3.3±0.0 g m^{-2} d^{-1}, around one-third of their growth potential, due to the lower light intensity used, i.e. 100 µmol m^{-2} s^{-1} versus 300 µmol m^{-2} s^{-1} (Chapter 4). Consequently these plants had a different chemical composition, including a lower nitrogen and higher lipid content (Supplemental Figure S7). The total phenol content determined in this study was 1.7 fold higher and 1.3 fold lower than reported for rapid-growing *A. filiculoides* and *A. pinnata*, respectively. This indicates a differential response of *A. filiculoides* and *A. pinnata* to elevated light intensity, which was also observed in the case of elevated CO_2 levels (Chapter 3). Given the many specific roles of secondary metabolites, it is to be expected that, similar to seed plants, not only the amount, but also the type of (poly)phenolic compounds being synthetized is highly dependent on the environmental conditions, including temperature, light conditions, presence of abiotic or biotic stress, etc. (Demkura et al., 2010; Moglia et al., 2014; Rivero et al., 2001; Izaguirre et al., 2007). This may also explain why neither flavones nor anthocyanidins were detected in this study versus previous work (Ishikura, 1982; Teixeira et al., 2001). Anthocyanidins provide *Azolla* its red-purple color during cold winters or phosphorous stress (Ismail and Mohamed, 2010; Kösesakal, 2014). Although anthocyanidins were not synthesized in the green *Azolla* ferns grown under favorable conditions, this clearly does not rule out their importance for the plant's functioning. Similarly, compounds that we identify as major compounds, may not be so abundant in ferns grown in other environments. Hence, depending on the climatic conditions and when *Azolla* is harvested, the amounts and types of phenolic compounds will differ, which will impact the value of the harvested biomass. Optimization of growth conditions and development of breeding lines will be required to develop sustainable production of protein and chemicals from *Azolla*.

Supplemental data

The supplemental tables and figures are available on request at: p-brouwer@live.nl and will be made available online upon publication.

Chapter 7

Extracting protein from tannin-rich *Azolla*

Paul Brouwer[1,2], Klaas G.J. Nierop[2],
Wouter J. Huijgen[3], Henriette Schluepmann[1]

[1] Molecular Plant Physiology, Institute of Environmental Biology, Utrecht University, The Netherlands;
[2] Department of Earth Sciences, Faculty of Geosciences, The Netherlands;
[3] Biomass & Energy Efficiency, Energy research Centre of the Netherlands (ECN), The Netherlands.

manuscript in preparation

Abstract

The fast-growing, nitrogen fixing, aquatic fern *Azolla* is a promising novel crop for the production of feed and food, due to its high protein content and favorable amino acid profile compared to soybean meal. However, digestibility of whole biomass is likely limited by the high content of phenolic compounds, particularly, condensed tannins. Here we investigated protein extraction methods for protein yield and (poly)phenol content of extracts obtained from *Azolla*, and a number of other (aquatic) plants. Protein precipitates obtained by extractions at pH 8 contained 16-26% of original biomass nitrogen, depending on extraction temperature. Increasing the extraction pH to 10.5 improved nitrogen recovery in the protein precipitate to maximally 54%. This substantial improvement in yield at pH 10.5 versus pH 8 was also observed in other plants belonging to the Salviniaceae family and the green seaweed *Ulva*, but not for *Lemna* and soybean. In *Azolla*, the improvement in yield is due to solubilization of protein-tannin complexes that are insoluble at pH 8. Extractions at pH 10.5 and pH 12.5 invoked strong interactions between tannins and proteins which prevented their subsequent separation. Two approaches were tested that allow protein extraction from *Azolla* while reducing the content of condensed tannins. Pre-extraction with aqueous acetone (70 v/v%) extracted 76-85% of the condensed tannins and consecutive extraction of the protein at pH 12.5 and 95°C for three hours produced a protein precipitate containing 38% of total biomass nitrogen. Protein extraction at pH 8 and 45°C in the presence of 1.5% PEG 6000 also produced a protein precipitate containing 38% of total biomass nitrogen, while reducing (poly)phenol content with 67%. Digestibility studies need to be performed to confirm effects of (poly)phenol removal on protein digestibility. The choice of methods further depends on their potential to be integrated into a bio-refinery, with (poly)phenols and carbohydrates as potential co-products.

Introduction

The world population is expected to grow to more than 9 billion people in 2030 (United Nations, 2015), which will pose a challenge for global food supply. Especially the amount of digestible protein is limiting in the diets of many population groups (Millward and Jackson, 2004). Currently the major protein crop is soy, which is predominantly produced in (sub)tropical regions of Brazil, USA and Argentina and imported in large quantities in temperate regions, such as western Europe (WWF, 2014). Some of the imported soy is used for direct human consumption, but most of it is utilized for feeding cattle, providing meat protein (WWF, 2014). Given the growing population and the geographically concentrated production of plant protein, novel sources of plant protein are sought, especially for temperate regions.

Azolla is a family of fast-growing, nitrogen-fixing aquatic ferns that thrive in tropical to temperate regions of the world (Lumpkin and Plucknett, 1980; van Hove and Lejeune, 2002; van Hove, 1989) *Azolla* plants are particularly rich in protein and have a favorable amino acid profile compared to soy (Chapter 4) (Becerra et al., 1990). Therefore *Azolla* has been proposed as a novel feed source to replace soy in the diets of pigs and poultry. In feeding trials whole *Azolla* could be included in diets at rates of 5-15%% for poultry, 15% for sows and 25% for tilapia fish. However *Azolla* could not fully replace commercial soybean meal, since at higher inclusion rates negative effects on animal weight and overall digestibility were observed (Becerra et al., 1990; Leterme et al., 2010; Abdel-Tawwab, 2008; Basak et al., 2002; Naghshi et al., 2014; Alalade and Iyayi, 2006).

Recent studies have pointed towards phenolic compounds as a possible cause for the limited inclusion rate. *A. filiculoides* and *A. pinnata* contained more phenolic compounds than what is generally encountered in food (Chapter 4). Dominant phenolic compounds in *Azolla filiculoides* and *Azolla pinnata* include (di)caffeoylquinic acids and proanthocyanidins, or condensed tannins (CT) (Chapter 5). CTs have also been detected in *A. caroliniana* (Nierop et al., 2011) and *A. africana* (Fasakin and Balogun, 2001).

CTs traditionally have been considered anti-nutrients because intake of tannin-rich feed by animals usually leads to a reduced digestibility of protein and a subsequent increase in fecal nitrogen (Bravo, 1998). The digestibility of protein may be reduced by CTs interacting with dietary protein or digestive enzymes, although many mammals are known to produce proline-rich salivary proteins (RPRs) that strongly bind CTs before entering the digestive track (Shimada, 2006). In the case of poultry, production of RPRs has not been reported. Including tannin-rich hulls in the diet of chicken caused reduced digestion of amino acids, starch and lipid compared with the control diet, mainly due to inactivation of digestive enzymes by CTs (Longstaff and Afrc, 1991). In pigs small amounts of PRPs have been reported but exhibited limit-

ed affinity to tannins (Mole et al., 1990). The reduced protein digestibility observed when feeding tannin rich faba bean hulls to pigs, was attributed to the interaction of CTs with both dietary and endogenous proteins in the digestive tract (Jansman et al., 1995). The limited inclusion rate observed in feeding trials with poultry and pigs with *Azolla* may therefore be explained by the 4-8% of CTs in different *Azolla* species (Chapter 5) (Fasakin and Balogun, 2001).

Small phenolic compounds including dicaffeoyquinic acid and flavonol-glucosides are incapable of precipitating protein (Bravo, 1998), but may form covalent bonds with protein by enzymatic and non-enzymatic oxidation as well as non-covalent hydrogen bonds (Bravo, 1998; Budryn and Rachwal-Rosiak, 2013; Le Bourvellec and Renard, 2012). Depending on the type of protein these bonds may lead to decreased protein solubility and/or digestibility (Budryn and Rachwal-Rosiak, 2013) .

A reduced amount of (poly)phenols and especially CTs will therefore very likely increase the digestibility and thereby commercial value of *Azolla* as feed. One approach to enhance the inclusion rate of *Azolla* in animal diets is to feed extracted protein instead of whole *Azolla* as feed.

Alkaline protein extraction has been successfully used to extract protein from seeds, microalgae, macroalgae, grasses and green tea leaves (Sari et al., 2015). Alkaline protein extraction was further used to produce protein extracts (6-7.2% nitrogen) from duckweeds *Lemna*, *Spirodela* and *Wolffia*, although extraction yields were not reported (Rusoff et al., 1980). Alkaline protein extraction has not been previously performed for *Azolla*. Fasakin (1999) used leaf pressing followed by heat congelation on *A. africana* to produce a high purity protein concentrate (11.4% nitrogen) containing 62.1% of the initial pulp nitrogen. Although providing high yields, heat congelation denatures protein whereas alkaline extraction can yield soluble protein, which broadens the scope for product applications.

Therefore, in this study we investigate alkaline protein extraction for the production of protein feed from biomass of *Azolla*, with a focus on the fate of phenolic compounds during this extraction. First we study protein yield in relation to extraction pH and temperature and compare these results with other (novel) feedstocks. Secondly we use pyrolysis and THM to track (poly)phenols during the extraction. Finally, we use the knowledge obtained from these experiments to test approaches for recovering protein from *Azolla* with a lower content of (poly)phenols.

Materials and methods

Plant material and growth conditions

A. filiculoides was obtained from a location at Galgenwaard, Utrecht, the Netherlands. *A. pinnata* was obtained from the IRRI accession, under accession number 535. *S. molesta* was obtained from the collection of the Utrecht University botanical garden. *A. filiculoides*, *A. pinnata* and *S. molesta* were grown in a growth chamber providing 16h light at an intensity of 100 µmol m^{-2} s^{-1} , a stable temperature of 21°C and 70% humidity. The growth medium for *Azolla* species consisted of 0.65 mM $NaH_2PO_4 \cdot H_2O$, 1.02 mM K_2SO_4, 1 mM $CaCl_2 \cdot 2H_2O$, 1.65 mM $MgSO_4 \cdot 7H_2O$, 17.9 µM Fe-EDTA, 9.1 µM $MnCl_2 \cdot 4H_2O$, 1.6 µM $Na_2MoO_4 \cdot 2H_2O$, 18.4 µM H_3BO_3, 0.8 µM $ZnSO_4 \cdot 7H_2O$, 0.3 µM $CuSO_4 \cdot 5H_2O$, 0.2 µM $CoCl_2 \cdot 6H_2O$. Growth medium of *S. molesta* was supplemented with 2 mM NH_4NO_3. Medium pH was adjusted to 5.5 using KOH. Every two weeks the growth medium was replaced completely. One third of the biomass was harvested weekly and freeze-dried. Batches of freeze-dried biomass were pooled and used for laboratory scale protein extractions. Freeze-dried biomass of *Lemna minor* was provided by Wageningen Research & University, the green seaweed *Ulva lactuca* was obtained from OceanHarvest (Ireland) and soybeans where acquired commercially.

For larger scale protein extractions, *A. filiculoides* was grown in a greenhouse in the botanical gardens of Utrecht University from August 2016 to November 2016, without additional light or heating. Plants were grown in 12L containers with a surface area of 459 cm^2. The growth medium was as described above and was replaced each two weeks. Harvested material was directly used for protein extraction.

Materials

Sodium hydroxide, polyethylene glycol (3350 & 6000 Da) and sodium-ascorbate (≥98%) were purchased from Sigma Aldrich. Hydrochloric acid, (38%) and anhydrous acetone (99%) were obtained from Merck. Tannic acid, chlorogenic acid, catechin and Bovine Serum Albumin (BSA), used for reference, were purchased from Sigma Aldrich. Black spruce tannins were provided by Caroline M. Preston (Lorenz and Preston, 2002) and characterized by Klaas G.J. Nierop (Nierop et al., 2005).

Protein extraction

For laboratory scale protein extraction 20 mg of freeze-dried *A. filiculoides* bio-

mass was ground using a tissue-lyser. After grinding 1 ml of solvent was added and tubes were vortexed. The solvent consisted of water with different concentrations of NaOH and possible additives, including polyethylene glycol (PEG, MW 3350 and 6000 Da) and vitamin C in the form of sodium ascorbate. The pH was adjusted during the extraction to remain around 8.0, 10.5 and 12.5. The extraction was performed for 4 h at either room temperature (RT), 45 °C or 95 °C. Solubilized compounds were separated from the insoluble fraction by 10 minutes of centrifugation at 6000x g. The protein precipitate was obtained by adjusting a known amount of the soluble fraction to pH 3.5±0.25 using HCl followed by 10 minutes of centrifugation at 6000x g. The fractions, i.e. insoluble, soluble, protein precipitate and supernatant, were collected in pre-weighed tubes, freeze-dried and weighed. The extraction was performed in nine technical replicates. Of these replicates, three were used for carbon and nitrogen (C and N) determination, two were used for pyrolysis and THM analysis and three were used for total phenol analysis. One replica was stored as a back-up.

For larger scale extraction fresh *Azolla* biomass was surface dried on a filter paper, weighed and a aliquot was taken to determine the dry weight. Surface dried biomass and solvent were mixed in a blender at a 1:50 fresh weight: solvent ratio. After 2x1 minute of mixing the solution was transferred to 50 ml tubes, after which the procedure was followed as described above. Extractions where performed in technical triplicates.

Chemical analysis

For total carbon and total nitrogen content (C and N), 1-2.5 mg of the freeze-dried material was analyzed using an elemental analyzer (Fisons NA 1500 CNS) connected to a mass spectrometer (Finnigan Delta Plus). Total phenol was determined by extracting samples with aqueous acetone (70%) for 24 h at 20 °C followed by reacting the samples with Folin-Ciocalteu reagent (Waterman and Mole, 1994). Insoluble fractions and soluble protein fractions were adjusted to a neutral to acidic pH with HCl prior to the extraction. Tannic acid was used as the calibration standard. For detection of condensed tannins, the acid butanol assay was performed by adding to the freeze-dried sample 1.2 ml of n-butanol-HCl solution (5 v/v%) and 20 µl of 2% ferric ammonium sulfate in 2N HCl. Reagents were vortexed and heated at 95 °C for 50 minutes before measuring absorbance at 550 nm(Makkar et al., 1999). Black spruce tannins were used as the reference standard.

For characterization of soluble and insoluble carbohydrates in *Azolla* biomass, first two successive Soxhlet extractions using water and ethanol were performed to remove non-structural components (Grisel et al., 2014; Wildschut et al., 2013). The extracted *Azolla* feedstock and aqueous extract were hydrolyzed in two steps: 12 M H_2SO_4 at 30 °C for 1 h, followed by 1.2 M H_2SO_4 at 100°C for 3 h. The hydro-

lysate was analyzed using High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD, ICS3000, Dionex, Sunnyvale, CA) with a CarboPac PA1 column. Analysis was performed using a eluents sequence for lignocellulosic carbohydrates (Grisel et al., 2014). Carbohydrate determination in the starting material and insoluble residue of protein extraction was performed without pre-extractions and using a eluents sequence optimized for seaweed carbohydrates (i.e., including uronic acids). The ash content of *Azolla* feedstock was determined by combustion in duplicate at 550 °C.

Pyrolysis was carried out using a Horizon Instruments Curie-Point pyrolyser. Samples were heated for 5 s at 600 °C. The pyrolysis unit was connected to a Carlo Erba GC8060 gas chromatograph and the products were separated with a fused silica column (Varian, 25 m, 0.32 mm i.d.) coated with CP-Sil5 (film thickness 0.40 µm). Helium was used as carrier gas. The oven temperature program was: 40 °C (1 min) to 180 °C at 7 °C min^{-1} and subsequently to 320 °C (held 15 min) at 20 °C min^{-1}. The column was coupled to a Fisons MD800 mass spectrometer (range m/z 45–650, ionization energy 70 eV, cycle time 0.7 s).

Thermally assisted Hydrolysis and Methylation (THM), i.e. pyrolysis in the presence of tetramethylammonium hydroxide (TMAH), was also employed to provide more insight into the polyphenolic composition. Prior to THM, samples were pressed onto Curie-point wires, a droplet of TMAH (25% in water) was added and samples were dried under a 100 W halogen lamp. Analysis of the THM products by way of GC–MS was identical to conventional pyrolysis. Compound identification was carried out by way of mass spectral comparison using a National Institute of Standards and Technology (NIST) library, interpretation of the spectra, retention times and/or comparison with literature data.

Calculations

The mass balance of protein extractions was made by firstly correcting the mass of each freeze-dried fraction (insoluble, soluble, protein precipitate and supernatant) for the content of Na+ , Cl-, sodium ascorbate and/or PEG in the remaining solvent. The mass of the soluble fraction, protein precipitate and supernatant was further corrected for the volume used for analysis, leading to the corrected mass (M_{fc}) as following:

$$M_{fc} = \frac{(V_t/V_a) \cdot (M_f - V_{rf} \cdot C_s)}{M_{start}}$$

Where V_t = the total volume of the soluble fraction after centrifugation, V_a is the volume used for analysis, M_f = mass of the fraction after freeze-drying, V_{rf} = the residual volume of solvent left in the fraction before analysis, C_s = the mass concen-

tration of the solvent, M_{start} = mass of the starting material after freeze-drying. In the cases that the soluble fraction and the supernatant were not analyzed, $V_t = V_a$. The corrected nitrogen balance (N_{fc}) was calculated by:

$$N_{fc} = \frac{(V_t/V_a) \cdot (M_f \cdot N_f)}{(M_{start} \cdot N_{start})}$$

Where N_f = the nitrogen content in the fraction and N_{start} = the nitrogen content in the starting material. Protein yield was estimated by:

$$N_{fc} \cdot \frac{N_{tot}}{C_{w,start}} \cdot K_b$$

Where $C_{w,\,start}$ = the residual water content of the freeze-dried starting material and K_b stands for the empirically determined ratio between amino acids (AAs) and total biomass nitrogen (N_{tot}), i.e. 4.9 for *Azolla* species (Chapter 4), 4.6 for *U. lactuca*, 4.8 (Bikker et al., 2016) for *L. minor* (Zhao et al., 2014) and 5.7 for Soybean (Petzke et al., 1997).

Lastly, protein purity was estimated by multiplying the nitrogen content (N_f) with K_{AA}, which is the ratio between AAs and amino acid nitrogen (N_{AA}), i.e. 5.7 for *Azolla* species (Chapter 4) and *U. lactuca* (Bikker et al., 2016) and 5.8 for *L. minor* (Zhao et al., 2014) and soybeans (Petzke et al., 1997). Both K_b and K_{AA} have not previously been determined for *S. molesta* and were assumed equal to the ratios for *Azolla*.

Quantification of pyrolysis and THM products was performed by peak integration of mass chromatograms using characteristic fragment ions of the compounds of interest. A list of retention times and characteristic fragment ions used for quantification is provided in Table S1, Supporting information.

Results & Discussion

Growth and chemical composition of *Azolla filiculoides*

Azolla filiculoides cultures grew under constant climatic conditions and were harvested weekly to ensure biomass productivity, which averaged 3.4 g m^{-2} d^{-1} over a period of 28 days. Using nitrogen concentrations we estimated protein content at 18.4% of the dry weight (% for the remainder of the article). The biomass further contained 19.8% polysaccharides, predominantly glucan (67% of the polysaccharide weight), followed by galactose (12.6%), xylose (7.6%) and mannose (6.5%) (Table 1). Lipid extraction yielded 12.8% of crude lipids, whilst extraction with aqueous acetone (70 v/v%) yielded 5.3% tannic acid equivalent (TA-eq) of (poly)phenols, including 4% condensed tannins (CTs). Hot water extraction also yielded 2.1% of sol-

uble sugars, mostly fructose and glucose. The biomass further contained 14.3% of ash, of which 11.6% was solubilized by hot water (Table 1).

The ash content is within the range (8.7-18.5%) previously reported for *A. filiculoides* (Costa et al., 1999; Datta, 2011). The protein content determined here differs from values reported previously (Becerra et al., 1990; Costa et al., 1999; Datta, 2011), since we used a conversion ratio (Kb) specific for *Azolla* (4.9), instead of the standard 6.25 (Chapter 4). Hemicellulose content (10.8-18.1%) and cellulose content (11.0-22.7%) were previously determined by the Van Soest method (Costa et al., 1999), which can overestimate polysaccharide content due to protein remaining in the Neutral Detergent Fiber residue (Godin et al., 2014). Our work indicates a lower content of polysaccharides in *A. filiculoides*, and in particular less hemicellulose than was previously estimated by the Van Soest method.

Table 1. Growth and chemical composition of *A. filiculoides*.

Growth rate (g m^{-2} d^{-1})		3.4
Protein[1]		20.6
Lipids		12.8
Total phenol[2]		5.3
Soluble condensed tannins		4.0
Soluble sugars		2.1
	Fructose	1.2
	Galactose	0.04
	Glucose	0.7
	Xylose	0.15
Polysaccharides		19.8
	Glucan	13.3
	Xylan	1.5
	Galactan	2.5
	Arabinan	0.6
	Mannan	1.3
	Rhamnan	0.6
Ash		14.3

[1] Based on a ratio of protein / total N of 4.9
[2] In tannic acid equivalent

Alkaline extraction at pH 10.5 effectively solubilizes *Azolla* protein

The harvested biomass used to test a series of extractions at pH 8, 10.5 and 12.5 carried out either at RT, 45°C or 95°C. As illustrated in Figure 1A,B, pH had a pronounced effect on the mass and nitrogen (N) balances of the extraction. At the mildest conditions of pH 8 and RT only 31% of the biomass dry weight was solubilized and 10% was precipitated. Of the biomass N, 31% was solubilized and 18% was recovered in the protein precipitate (Figure 1C, D). Increasing the pH to 10.5 yielded more solubilized compounds, which were mostly recovered in the protein precipitate (23%). The increase in pH preferentially solubilized N-rich compounds and acid precipitation recovered 54% of the initial biomass N in the protein precipitate. Increasing the pH from pH 10.5 to 12.5 at RT increased solubilized compounds, but did not increase the yield of the protein precipitate (21%) (Figure 1A,B Figure S1,

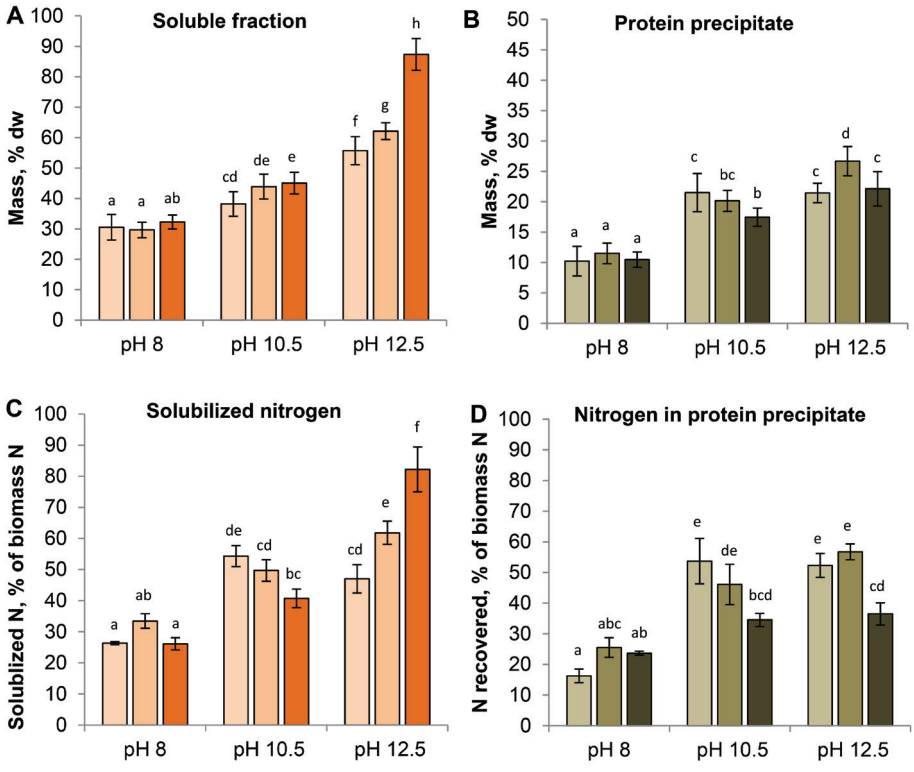


Figure 1. Percentage of initial mass (n=9) and nitrogen (n=3) in the soluble fraction (A,C) and protein precipitate (B,D) obtained by alkaline extraction of *A. filiculoides*. Extractions were performed at pH 8, pH 10.5 and 12.5 for 4 h. Lighter to darker bars indicate extraction temperature: RT, 45°C and 95°C. Statistical analysis was by one-way ANOVA and a tukey post-hoc test; letters indicate significantly differing groups (P<0.05).

Supporting Information). Moreover, both the amount of solubilized N and the amount of N recovered in the protein precipitate remained similar to pH 10.5 (Figure 1C,D).

The effect of temperature was specific for each extraction pH. At pH 8 an increase in temperature from RT to 45°C slightly increased (non-significant) the biomass N recovered in the protein precipitate, but no changes occurred when increasing temperature to 95°C. At pH 10.5 elevated temperature had a slightly negative effect on both solubilized N and, especially, on the N recovered in the protein precipitate. For pH 12.5 extractions, temperature had a much greater effect. At a temperature of 45°C the soluble fraction contained 62% of the initial mass and 67% of the biomass N. The protein precipitate represented 27% of the initial mass and contained 57%

of the biomass N. At 95°C the soluble fraction increased further to 82%, containing 84% of the biomass N. The solubilized compounds, however, were not preferably precipitated with acid but instead remained in the supernatant, resulting in a reduction in mass and N recovered in the protein precipitate with respect to 45°C. (Figure 1, Figure S1, Supporting Information).

Alkaline extraction at pH 12.5 and 95°C was previously reported to provide high protein yields from green tea leaves (Zhang et al., 2014). The improved solubilization of N is likely due to the degradation of membrane-bound protein that would otherwise not become soluble. Considerable degradation of *Azolla* protein was clearly observed after 2h of extraction at pH 12.5 and 95°C (Figure S2, Supporting Information). However, the degraded protein may no longer be susceptible to acid precipitation, explaining the small (41%) fraction of solubilized N being precipitated. Hence extraction at pH 12.5 and 95°C does not improve protein yield compared to extraction performed at pH 10.5 and RT, at least not for the extraction time tested and when used in combination with acid precipitation.

Comparing alkaline extraction of *Azolla* with other feedstocks

A second set of protein extractions was performed to compare protein yields to alternative aquatic plants and algae and to see whether the large increase in extraction efficiency at pH 10.5 versus pH 8 could also be observed for these feedstocks.

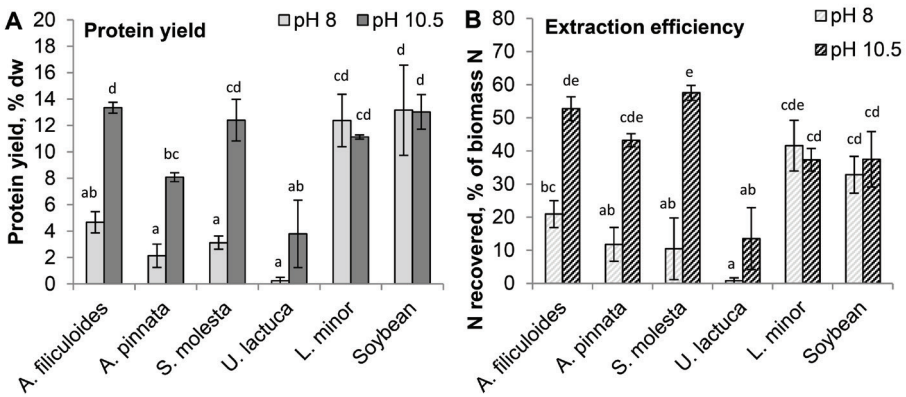


Figure 2. The effect of extraction pH on protein yield (A) and nitrogen recovered in the protein precipitate (B) for protein extractions at RT of *A. filiculoides*, *A. pinnata*, *S. molesta*, *U. lactuca*, *L. minor*, and soybeans (n=3). Statistical analysis was by one-way ANOVA and a tukey post-hoc test; letters indicate significantly differing groups (P<0.05).

Soybean was used as a control. Protein contents of the aquatic plants and soybeans are given in Table 2.

For *A. filiculoides*, absolute protein yield was maximally 13.3% at pH 10.5 and RT, similar to *S. molesta* (12.4%) *L. minor* (12.2%) and *G. max* seeds (12.9%) (Figure 2A).Maximum protein yield from *A. pinnata* was significantly less (8.1%) compared to *A. filiculoides*, mostly due to lower protein content in the biomass (16% versus 21%) (Figure 2A; Table 2). The maximum protein yield from *U. lactuca* was significantly lower compared to other feedstocks except *A. pinnata*, due to much lower extraction efficiency (Figure 2B). This is likely due to the tough cell wall of algae that are generally difficult to disrupt (Sari et al., 2015; Polikovsky et al., 2016). The mechanical destruction of the soybeans also proved more difficult compared to the leaf tissues and the extraction efficiency of 33-37% obtained here is at the low end of values reported in literature (33-71%) (Sari et al., 2015).

The extraction efficiency for *A. filiculoides* at pH 10.5 was similar to the previous experiment, recovering 53% of the biomass N in the protein precipitate (Figure 2B). For *Lemna* the highest amount of N recovered in the protein precipitate was 42% at pH 8, and 58% for *S. molesta* at pH 10.5 (Figure 2B). Given these aquatic plants consist mostly of leaves, these values likely reflect the water soluble protein fraction, which is generally found to lie between 40 and 50% of biomass N in leaves (Takashima et al., 2004; Funk et al., 2013; Fiorentini and Galoppini, 1983). Of the soluble protein RuBisCo generally represents half (Takashima et al., 2004; Fiorentini and Galoppini, 1983). Membrane-associated proteins such as thylakoid components, representing the largest part of the insoluble fraction of leaves (Evans and Seeman, 1989) are unlikely solubilized during extraction at pH 8 and pH 10.5 at RT.

Increasing the extraction pH from 8 to 10.5 had a significantly positive effect on the extraction efficiency for *A. filiculoides* (2.5 fold), *A. pinnata* (3.7 fold) and *S. molesta* (5.5 fold) and a non-significant effect on the extraction efficiency for *U. lactuca* (16.6 fold). However, increasing pH did not change the extraction efficiency

Table 2. Protein content of biomass and mass and purity of protein precipitates

	Start material	pH 8, protein precipitate		pH 10.5, protein precipitate	
Feedstock	Protein content [% dw]	Mass [% dw]	Protein purity [%]	Mass [% dw]	Protein purity [%]
A. filiculoides	21 ± 2	11 ± 1	39 ± 2	26 ± 2	52 ± 2
A. pinnata	16 ± 0	8 ± 3	22 ± 3	25 ± 1	29 ± 1
S. molesta	18 ± 1	9 ± 0	41 ± 2	26 ± 3	46 ± 5
U. lactuca	22 ± 1	0 ± 1	36 ± 13	7 ± 3	51 ± 7
L. minor	25 ± 0	17 ± 2	67 ± 1	17 ± 2	68 ± 1
Soybean	37 ± 3	20 ± 3	65 ± 3	20 ± 4	58 ± 2

for *L. minor* or soybeans, which were statistically identical at both pH levels (Figure 2B). Additionally, the purity of protein precipitates of *Azolla* species, *S. molesta* and *U. lactuca* increased to roughly 50% by increasing the extraction pH from 8 to 10.5 (Table 2). The purity of protein precipitates of *L. minor* and Soybean was generally higher (58-68%) compared to the other feedstocks, and did not improve by increasing pH from 8 to 10.5 (Table 2).

The limited yield at pH 8 is thus shared by species in the Salviniaceae family and is also observed in macroalgae, although the properties that cause these low extraction efficiencies may be different in each species.

Elevating pH enhances yield by preventing protein precipitation by tannins

To determine the fate of the phenolic compounds during protein extraction, we firstly investigated the insoluble fraction for residual (poly)phenols using (conventional) pyrolysis. All pyrolysis products, their likely origin and relative abundance are given in Table S1, Supporting Information. Catechol is the main pyrolysis product from the soluble (poly)phenols in *Azolla* (Chapter 5). Since saccharides are predominantly present as insoluble polysaccharides (Table 1), the yield of pyrolysis products from saccharides is not expected to change due to protein extraction. Therefore a ratio between catechol and saccharide related products is used to track (poly)phenol

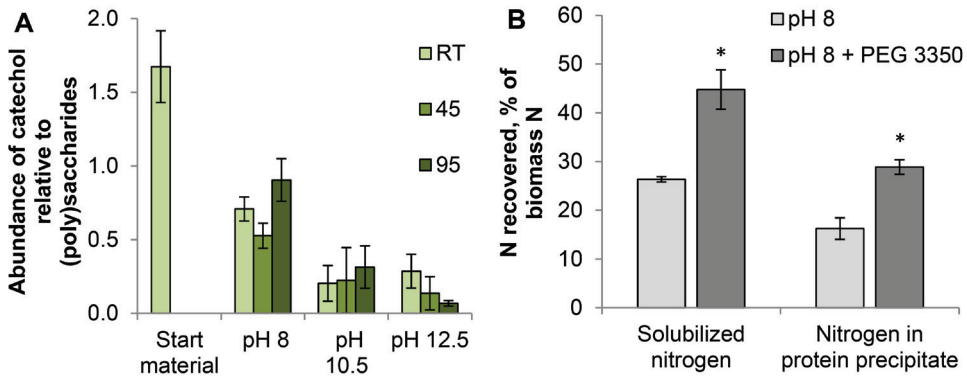


Figure 3. The presence of (poly)phenols in the insoluble fraction (A) and the effect of the addition of PEG to nitrogen solubilization and recovery at pH 8 (B). (A) The presence of polyphenols is indicated by the ratio between (poly)phenol derived catechol and polysccharide derived pyrolysis products (n=2) (Table S1, Supporting information). (B) Nitrogen content of the soluble fraction and protein precipitate were determined for pH 8 extractions at RT with and without addition of PEG 3350 (n=3). The asterix denotes a significant ($p<0.05$) difference as determined by an independent sample T-test.

content of the insoluble fraction (Figure 3A). For pH 8 extractions, the ratio between catechol and polysaccharide-derived pyrolysis products was 0.53-0.90, depending on extraction temperatures (Figure 3A), compared to 1.67 for the starting material. This indicates that a large fraction of the (poly)phenolic compounds does not become soluble during protein extraction at pH 8. However, for pH 10.5 and 12.5 extractions, the ratio between catechol and (poly)saccharide products was much lower (i.e 0.07-0.31), indicating (poly)phenol solubilization.

THM was conducted on the insoluble fraction obtained from pH 8 extractions to determine the type of (poly)phenols that remain insoluble during extraction. In the insoluble fractions obtained from pH 8 extractions at RT and 45°C, THM products typical of flavonoids were enriched compared to the other phenolic compounds (Table S1, Supporting Information). THM products of flavonoid A-rings, i.e. 1,3,5-trimethoxybenzene, 1-methyl-2,4,6-trimethoxybenzene and 1-ethyl-2,4,6-trimethoxybenzene were slightly enriched at both temperatures. THM products from the flavonoid B-ring containing two hydroxyl groups (catechin type) were highly enriched, whereas THM products from a flavonoid B-ring with one hydroxyl group (azfelechin type) were less abundant. However, enrichment of these flavonoid THM products was not observed in the insoluble fraction of extractions performed at pH 8 and 95°C (Table S1, Supporting Information).

Results from pyrolysis contradicted results obtained by the colorimetric Folin Ciocalteu assay which indicated few (poly)phenols in the insoluble fraction of the pH 8 extractions (Figure S3A, Supporting information). Since the total phenol assay involves an aqueous acetone (70 v/v%) extraction we tested the extractability of a series of model (poly)phenols when in the presence of the Bovine Serum Albumin (BSA) reference protein. Whereas chlorogenic acid and catechin were completely extracted, only 50% of black spruce CTs were released from the BSA protein in acetone (70 v/v%) (Figure S4, Supporting information). This suggests CTs form strong interactions with protein at pH 8, that can only partly be broken by aqueous acetone (70 v/v%) extraction, which explains the discrepancy between the colorimetric and thermolytic (pyrolysis and THM) approaches (Nierop and Verstraten, 2006).

Interactions between protein and CTs have been well studied and depend on various parameters such as pH, temperature, protein sequence, and functional groups and the degree of polymerization of the CTs (Saminathan et al., 2014; Zeller et al., 2015). Although most studies have been performed with BSA, Zeller et al. (2015) showed that CTs precipitate Alfalfa leaf protein more strongly than BSA (Zeller et al., 2015). Purified Rubisco was also strongly precipitated by CTs (Mcallister et al., 2005; McNabb et al., 1998).

Increasing pH decreases the precipitation of protein by CTs (Hagerman and Butler, 1981). The large difference in extraction efficiency between pH 8 and pH 10.5 observed for *A. filiculoides*, *A. pinnata* and *S. molesta*, was accompanied by a high

CT content, whereas soybeans and *L. minor* contained 0% and 0.8% of CTs (Figure S5, Supporting Information). As an exception, increasing pH also enhanced protein extraction efficiency from the seaweed *U. lactose*, which contains no CTs. To confirm whether the increase in pH improved protein extraction efficiency by means of decreasing the precipitation of protein by CTs, we performed the protein extraction in the presence of polyethylene glycol (PEG). PEG is known to prevent the interaction between tannins and protein (McNabb et al., 1998). Here, the addition of 1% PEG 3350 during protein extraction from *Azolla* at pH 8 resulted in a significant increase in solubilized N and N recovered in the protein precipitate (Figure 3B).

Hence, although it is known that increasing pH can lead to higher protein extraction efficiency by enhancing protein solubility and aiding in cell wall lysis (Sari et al., 2015), we show that in *Azolla* increasing pH enhances protein extraction by preventing protein precipitation by CTs.

Protein and (poly)phenols strongly interact at high pH

As (poly)phenols seem to be co-extracted from the insoluble fraction upon protein extraction at pH 10.5 and pH 12.5, we investigated the protein precipitates for (poly) phenol content by pyrolysis. Again catechol was considered indicative of (poly)phenolic compounds and was compared to the pyrolysis products typical of protein. The ratio of catechol over protein related products was 0.84-1.2 for protein precipitates

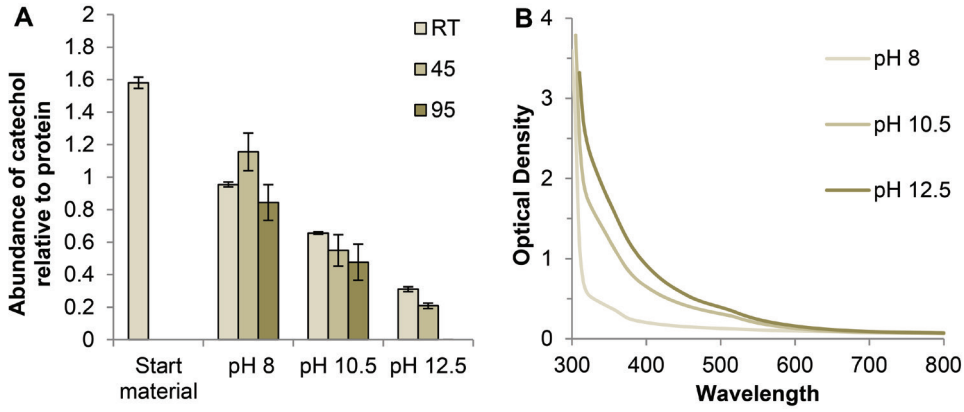


Figure 4. The presence of (poly)phenols in the protein precipitate (A) and coloration of the protein extracts (B). (A) The presence of (poly)phenols is indicated by the ratio between (poly) phenol derived catechol and protein derived pyrolysis products (n=2) (Table S1, Supporting information). (B) Coloration of the protein extract was determined by photospectrometer in diluted (1:10) protein extracts obtained by pH 8, pH 10.5 and pH 12.5 extractions carried out at RT for 4 h.

from pH 8 extractions compared to 1.58 for the starting material, depending on extraction temperature. For protein precipitates from extractions at higher pH, this ratio decreased with increasing pH and temperature, up to no detection of catechol at pH 12.5 and 95°C (Figure 4A).

We also observed that extraction at pH 10.5 and 12.5 resulted in fast red-brown coloration of the protein extract, as opposed to extractions performed at pH 8 (Figure 4B). A general shift of the UV/VIS absorbance to the visual spectrum is typical of oxidative polymerization of (poly)phenols and the rate of oxidation increases exponentially with pH (Jha and Halada, 2011) (Figure S6; Figure S7, Supporting Information). Oxidation of (poly)phenols is believed to result in the formation of highly reactive quinones (le Bourvellec and Renard, 2012; Ozdal et al., 2013). When in the vicinity of proteins, quinones irreversibly react with the sulfhydryl and amino groups of proteins and thereby form protein cross-links (le Bourvellec and Renard, 2012; Damodaran, 1996; Prodpran et al., 2012). These quinone–amino group reactions are known to decrease the digestibility and bioavailability of protein-bound lysine and cysteine (Damodaran, 1996). Additionally, quinones can undergo condensation reactions, resulting in the formation of high molecular weight pigments (Ozdal et al., 2013). The higher degree of polymerization results in a bathochromic shift in the UV/VIS absorption spectra, i.e. a shift towards the visible range noticeable as a red-brown color (Jha and Halada, 2011).

Apart from leading to a very likely decrease in protein product quality, oxidation of (poly)phenolic compounds may also have hindered adequate analysis of (poly) phenol content in the protein extracts. To test whether this affected our results we performed a series of experiments by dissolving pure (poly)phenols in either water or a 0.1M NH_3 solution for 16h. When in the presence of NH_3 (poly)phenols showed a similar bathochromic shift in the UV/VIS spectra as observed in the protein extracts. Additionally, the peak area of polyphenol products detected upon pyrolysis and THM were reduced with 70-99% in (poly)phenol solutions containing 0.1M NH_3, with respect to solutions in water (Figure S8, Supporting Information). Oxidation also decreased the response of the colorimetric total phenol assay to 72% for tannins and 44% for chlorogenic acid. Furthermore, the response of the acid-butanol assay to tannins was reduced to 70% after only two hours exposure to oxidizing conditions (Figure S9, Supporting Information).

Hence, although pyrolysis and THM perform well to track (poly)phenols during pH 8 extractions, at higher pH the oxidation of (poly)phenols leads to a high underestimation of (poly)phenols. The derived ratio's in Figure 4A are therefore underestimating the (poly)phenol content of protein precipitates obtained from extractions at pH 10.5 and 12.5. The coloration of the protein precipitate suggests than considerable amounts of polyphenols are present and have likely formed covalent bonds with protein, obstructing any attempts of subsequent separation.(Khoddami et al., 2013)

To reduce oxidation we tested the reduction of extraction time in combination with the addition of sodium ascorbate (sodium salt of vitamin C). Vitamin C is used to prevent (poly)phenol oxidation during analysis (Khoddami et al., 2013; Häkkinen et al., 1998) and also in food processing (Kolniak-Ostek et al., 2013; Kaack and Austed, 1998; Barril et al., 2012). Reducing the extraction time to 1 h resulted in minimal yield loss and adding sodium ascorbate at 5 mg ml^{-1} proved optimal in reducing the coloration of the protein precipitate (Figure S10, Supporting information). However, even when vitamin C was added to prevent oxidation, separation of protein and CTs by size separation or resin binding proved unsuccessful under alkaline conditions (Table S2, Supporting information). Although this may be explained by an inadequate ability of vitamin C to prevent the formation of covalent bonds non-covalent interactions may have also contributed. Hydrogen bonding between proteins and (poly)phenols is favored in hydrophobic solvents while hydrophobic bonding occurs in hydrophilic or polar solvents, such as water (Oh et al., 1980). In alkaline solution hydrogen bonds are unlikely formed between solutes due to deprotonation of the hydroxide groups and the high concentration of the hydroxide ion as a strongly competing hydrogen bond acceptor. Hydrophobic interactions, however, may still be formed between protein and tannins and could cause the inability of the protein and tannins to be separated by size exclusion when in alkaline solution.

Approaches to separate proteins and (poly)phenols

As separation of polyphenols and protein from alkaline solution proved unsuccessful, we evaluated two alternative approaches: (1) pre-extraction of polyphenols followed by protein extraction and (2) protein extraction in the presence (poly)phenol-binding additives. Here we focused on CTs as they are the major (poly)phenol constituent in *Azolla* biomass and show the highest binding affinity for protein. To improve the relevance of the data produced, we increased the scale of the protein extractions to 2 g dw per extraction and used a blender for mechanical cell-lysis as a more scalable method.

Pre-extraction of (poly)phenolic compounds by aqueous acetone (70%) decreased the content of CTs by 76-85%, in line with previous results (Chapter 5) (Figure 5A). However, the pre-extraction with acetone denatured the protein, making them insoluble at pH 10.5. A subsequent protein extraction at pH 12.5 and 95°C did allow solubilization of protein over time, recovering 38% of biomass N in the protein precipitate after 3 hours (Figure 5B). The partial degradation of protein at pH 12.5 and 95°C is utilized to bring denatured protein back in solution. In contrast, without aqueous acetone pre-extraction, proteins are directly soluble and a maximum of 54% of the biomass N is recovered after 1 hour, whereas further exposure to pH 12.5 and 95°C negatively affects recovery.

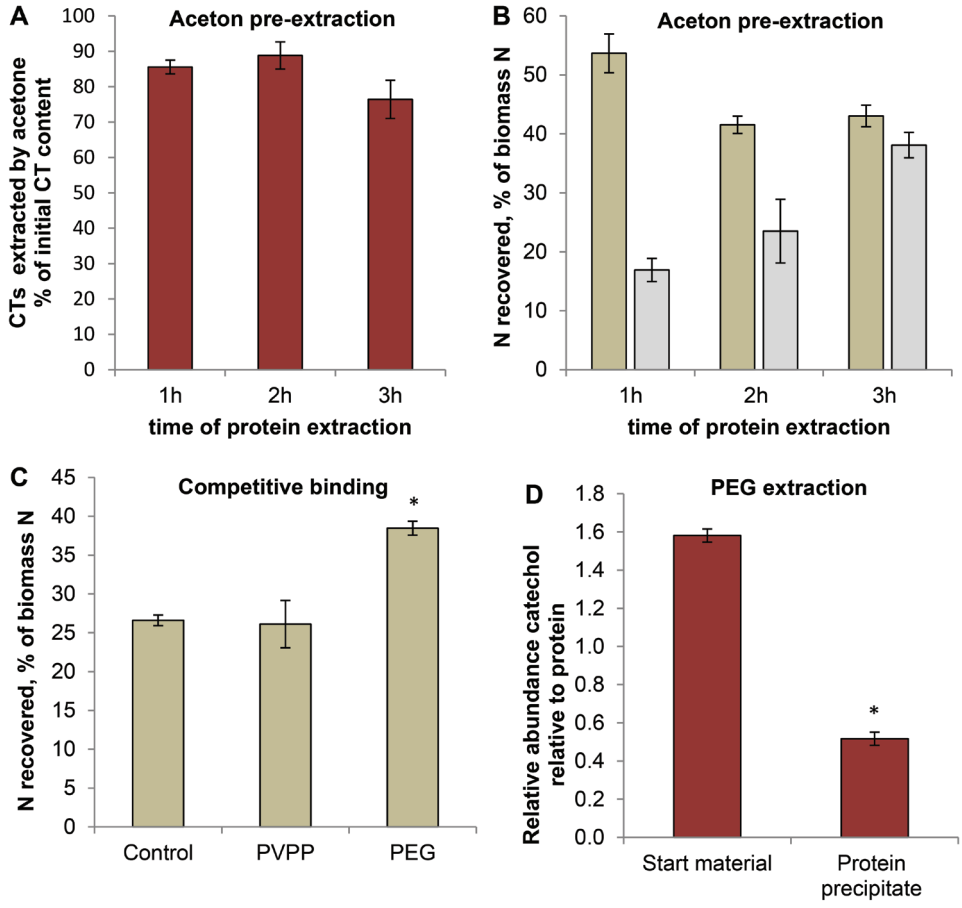


Figure 5. Protein yield and removal of (poly)phenols by aqueous acetone pre-extraction (A,B) and competitive binding (C,D). (A) Amount of CTs extracted by after acetone pre-extraction as determined by the acid-butanol assay (n=3). (B) The amount of biomass N recovered in the protein precipitate obtained by extraction at pH 12.5 and 95°C, without (brown) and with (gray) acetone pre-extraction (n=3). (C) Yield for extractions at pH 8 and 45°C performed with no additive (control), PVPP and 1.5% PEG (n=3). (D) (Poly)phenol content of protein precipitate obtained by extraction with PEG, indicated by the ratio between catechol and protein derived pyrolysis products (n=2) (Table S1, Supporting information). The asterix denotes a significant ($p<0.05$) difference to the control as determined by an independent sample T-test.

The harsh conditions needed to extract protein after aqueous acetone pre-extraction may evoke various chemical reactions between protein and other solubilized compounds that negatively influence protein quality. The coloration of the protein precipitate indicates that the (poly)phenols that remain in the biomass after aque-

ous acetone pre-extraction are fully oxidized. Furthermore, it has been reported that alkali treatment can form non-utilizable amino acids and oxidation can reduce the bioavailability of cysteine, which is often a limiting amino acid in feeds (Sari et al., 2015; Damodaran, 1996).

Using a lower extraction pH in combination with a CT binding additive may provide a better quality feed product. Similar to the small scale extraction described earlier (Figure 3B) we added PEG while extracting protein at pH 8. However, to enhance protein solubility we performed the extraction at 45 °C and used PEG with a molecular weight of 6000 Da, which was reported to bind CTs better than the smaller PEGs (Makkar et al., 1995). Apart from PEG, also polyvinylpolypyrrolidone (PVPP) is known to bind CTs, although slightly less efficient as PEG (Makkar et al., 1995). However, PVPP is used as an insoluble form, which allows easier recovery and recycling compared to the soluble PEG.

The extraction with PEG yielded a protein precipitate containing 38% of the initial N, which was significantly higher than the control for which 26% of the biomass N was recovered in the protein precipitate (Figure 5C). PEG's exceptionally strong binding affinity towards CTs has been attributed to the presence of weak hydrogen bond acceptors in the form of ether bonds combined with a hydrophobic ethyl chain (Oh et al., 1980). Additionally the polymeric nature of PEG allows it to connect to the polymeric tannins at multiple sites. The extraction with insoluble PVPP gave near identical results to the control and therefore did not improve protein solubilization at pH 8 even though it can bind CTs more strongly than protein (Hagerman and Butler, 1981). Since PVPP is insoluble it does not come directly into contact with the CTs when cell walls are mechanically destroyed, as opposed to the soluble PEG, which may (partly) explain the lack of improvement in protein yield at pH 8.

As shown in Figure 5D, the ratio between catechol and protein-derived pyrolysis products, was reduced by 67% in PEG-extracted protein versus the starting material. Binding of CTs by PEG during protein extraction likely further enhances the digestibility of the protein precipitate, as analogously, feeding animals PEG next to a tannin-rich diet, elevates the negative effects of any remaining tannins (Yisehak et al., 2014; Bhatta et al., 2002). The protein precipitate could therefore be directly used as high quality feed. However, this would infer high costs for the protein product as the costs for adding 1-1.5% PEG to the solvent are substantial. To make this process economically feasible, recovery of PEG is required. Possible approaches include using a resin with PEG as the functional group, although the lack of yield improvement with PVPP suggest that rapid contact between the functional molecule and the CTs may be essential. An alternative solution could the modification of the soluble PEG with an end group that specifically binds to a resin. The subsequent separation of PEG from CTs will provide a further challenge due to PEG's stronger binding capacity than protein.

In any case the separation of CTs and protein will infer additional costs to the *Azolla* production chain and it needs to be assessed whether these costs outweigh the increase in value of the final product. Currently we work under the assumption that reduced CT content will increase the digestibility of *Azolla* protein. Including CT separation during alkaline extraction would yield a product with higher protein content without fiber, as well as higher protein digestibility. Although this assumption is highly supported by existing literature, other anti-nutritional factors may be involved and digestibility studies comparing the protein extracts and starting material are needed to verify our assumption.

Protein extraction at the start of an *Azolla* bio-refinery

The choice of processing methods may not solely depend on the value of the protein product. The separation and subsequent purification of (poly)phenols may provide a valuable second product stream. An aqueous acetone (70%) pre-extraction generates a (poly)phenol extract rich in (di)caffeoylquinic acids. Artichoke extracts, rich in dicaffeoylquinic and chlorogenic acids, are commercialized as drugs, mainly for the treatment of liver diseases (Lattanzio et al., 2009). Alternatively, hydrolysis of the *Azolla* (poly)phenol extract would yield caffeic acids that can be used to synthesize derivatives which find application as an active ingredient in medicines and as formulation stabilizers in food, cosmetic and pharmaceutical industry (Silva et al., 2014). Caffeic acid has further been proposed for the synthesis of bioplastics resistant to high temperature (Chauzar et al., 2012).

Although, the fate of the caffeoylquinic acids during PEG extraction is unclear, CTs bound to PEG may be purified and commercialized. It is estimated that annually 180,000 t of CTs are produced, mostly by hot-water extraction from tree barks (Kemppainen et al., 2014). Although most CTs are used in leather production, novel applications are sought, such as replacing fossil-based phenol in insulating foams and adhesives (Kemppainen et al., 2014; Pizzi, 2006).

Additionally, alkaline protein extraction produces a carbohydrate-rich residue. Hydrolysis of the fiber-rich residue obtained by pH 10.5 extractions at RT, yielded 31.7% monosaccharides (Table S3, Supporting Information). This hydrolysate may be ideally suited for conversion into bio-ethanol. *Azolla* contains very small amounts of true lignin, which is favorable since lignin removal via pre-treatment is a major cost for producing ethanol from lignocellulosic biomass (Chapter 5) (Fu et al., 2011; Yang and Wyman, 2008). Enzymatic hydrolysis followed by fermentation by *Saccharomyes cerevisiae* previously yielded 0.09 g ethanol per g of *Azolla* biomass (Miranda et al., 2016), which may be significantly higher for the carbohydrate-rich residue due to the enrichment in polysaccharides as well as the removal of (poly)phenols.

Hence an *Azolla* bio-refinery may provide a range of products, depending on the

choice of extraction methods. The value of each of these products and the effect of subsequent extraction processes on the quality of these products, determines the order of extraction. For the production of caffeic acid derivatives, aqueous acetone pre-extraction is preferred as it yields a high concentration of these compounds, prevents interaction with protein and limits exposure to oxidative conditions. For the production of high quality protein PEG-extraction is more suitable, while potentially yielding a tannin by-product.

Acknowledgement

We would like to thank Arnold van Dijk for performing the numerous nitrogen analyses that form the basis of this paper. We would further like to acknowledge Mark Smit, Wouter Hoogland and Dally Shaffar for their help with the experimental work and the LPP foundation for their financial support.

Supplemental data

The supplemental tables and figures are available on request at: p-brouwer@live.nl and will be made available online upon publication.

General discussion & outlook

Control over the reproductive cycle of *Azolla*

The basis for any cropping system is the starting material. The starting material should be available and also provide reliable yields to farmers. While propagating *Azolla* vegetatively we have encountered frequent losses of accessions, due to improper handling of vegetatively reproducing plants. Also mixing of strains is easy as they can be difficult to differentiate. We therefore envision the sexual reproduction as a way to provide reliable starting material for *Azolla* cultivation. In Chapter 2 we presented our current knowledge of the *Azolla* life cycle and provide basic protocols for the fertilization and subsequent germination of *Azolla* spores. We found that *Azolla* spores cannot be preserved simply by drying, unlike seeds from most seed plants. Instead we adapted a cryopreservation protocol to preserve *Azolla* for longer time periods. Recently we added to this toolkit a protocol for inducing sporulation in *A. filiculoides* by changing light quality, density and nutrient availability. With this protocol we are now able to control each stage of the sexual reproduction of *Azolla filiculoides*. However, we have not yet applied these protocols at the scale required for *Azolla* farming or to breed improved varieties. Additionally, current protocols have only been successfully tested on *A. filiculoides*, whereas, for example, they did not prove successful with *A. pinnata*. We have started exploring the molecular mechanisms underlying sexual reproduction in *A. filiculoides* with the aim to develop methods to induce sporulation that are universally applicable to all *Azolla* species. Future research in this direction is important as cryopreservation of spores can provide *Azolla* researchers and users with a physical databank of genetically diverse *Azolla* species and varieties, that can be used as a resource to further improve *Azolla* based production systems.

Understanding molecular processes in *Azolla*'s: from sequencing to gene function

The study of fern biology is greatly lagging behind that of seed plants. The increased performance and decrease in costs of sequencing technology allowed us to construct a first database of transcribed genes in *Azolla filiculoides* and the third

transcriptome of a fern. This data gave insight in possible biological pathways leading to sexual reproduction in this fern and its response to nitrogen fixation by the *Nostoc* symbionts. However, we also recognize the limitations of assembling RNA sequences from short-read technologies such as illumina or ion torrent without any reference genome: the gene inventory obtained often yields contigs that do not represent full-length genes. The short contigs more often represent lowly expressed genes, which introduces a bias when computing differential expression: genes with low abundance RNA recruit so few reads that their differential expression is highly underestimated. Additionally, gene inventories obtained by assembly of transcript sequencing are inherently limited as only genes expressed under the chosen environmental conditions or plant developmental stages are obtained. For these reasons we also initiated sequencing of the *Azolla* genome. The cyanobacteria-free *Azolla* strains presented in Chapter 3 have been used as the starting material for genome sequencing efforts, which yielded a first genome assembly of a fern, of which the results are soon to be published. A consortium of researchers and businesses are now adding valuable genomic information, including short-read sequencing of six out of seven *Azolla* species and the genome assembly of the closely related *Salvinia cuculata* (Sessa et al., 2014; Li and Pryer, 2014). Bioinformatics tools allow investigation of sequence identity and homology of *Azolla* genes to genes with known functions in seed plants. In this way, genome sequencing data was used to increase the understanding of *Azolla*'s biology, such as polyphenol biosynthesis in Chapter 6 in which we report the discovery of a *LAR*-like gene in *Azolla*, which is likely crucial to the biosynthesis of condensed tannins. However, prediction of the functions of plant genes based on bioinformatics tools is limited and needs to be complemented with functional studies. The fast rate at which descriptive sequencing data is generated seems to outpace knowledge generation on gene function, leading to circular referencing of gene annotation based on gene homology, but relying only on a few studies of gene function. Additionally, novel genes and mechanisms will not be revealed by sequencing efforts. The work presented in this thesis inferring possible gene functions needs to be interpreted in this light. Although we can be quite confident of possible functions of highly conserved genes, such as key biosynthetic enzymes, lesser conserved enzymes or transcription factors may have different substrates, products or targets than is suggested based on gene similarity and/or homology. With greatly increased genome information available, the study of molecular biology in *Azolla* is ready to move towards the study of gene function. Genes encoding key enzymes can now be cloned into recombinant *Escherichia coli* to investigate substrate specify, reaction end products and reaction kinetics of these enzymes. Protocols for spore fertilization, germination and cryopreservation presented in Chapter 2, allow to undertake mutagenesis of *Azolla* spores to generate mutant lines. We have begun attempts to stably transform via callus induction. Callus could be induced by growing

Azolla plants on 1 mg ml^{-1} of 2,4-Dichlorophenoxyacetic acid (2,4-D) or a combination with 6-Benzyl Amino Purine (BAP), as was also published for *A. rubra* (Sini et al., 2014). However subsequent efforts to insert DNA via particle bombardment or Agrobacteria have not yet resulted in successful stable transformation. Stable transformation has only been achieved in ferns belonging to the *Ceratopteris* family, via *Agrobacterium tumefaciens* transformation of spores and particle bombardment of callus (Plackett et al., 2014; Muthukumar et al., 2013). A protocol to transform *Azolla* would be a great addition to available sequencing data and boost future research on *Azolla* by permitting validation of gene function. Transformation may also prove to be a tool directly applicable to generate new *Azolla* varieties with favorable characteristics.

Growth and nitrogen fixation: from controlled conditions to the real-world

The cultivation of *Azolla* requires an approach that differs from that of conventional cropping systems. Conventional cropping systems operate as a batch process, whereas *Azolla* cultivation requires continuous harvesting to obtain high yields. In Chapter 4 we have described methodologies to grow *Azolla* in highly productive systems. We have subsequently used these methodologies to investigate growth potential of two *Azolla* species, i.e. *A. filiculoides* and *A. pinnata*, that differ in genotype and natural habitat. Under constant climatic conditions we obtained stable productivities equaling 35.5 and 32.8 t dry weight (dw) ha^{-1} $year^{-1}$ for *A. filiculoides* and *A. pinnata* respectively. We obtained similar growth rates for *A. filiculoides* in an outdoor growth experiment in the Netherlands, during a month in the summer of 2014. Growth rates of 39 t dw ha^{-1} $year^{-1}$ have also been reported for an outdoor experiment in Colombia (Becerra et al., 1990). Growth experiments presented in Chapter 3 were conducted using a richer medium and productivities equaling 37-44 t ha^{-1} $year^{-1}$ were observed over the course of three weeks. The productivities presented in Chapter 4 are thus conservative estimates of the growth potential under constant climatic conditions, which may be improved by further optimizing growth conditions.

In Chapter 3 we have further shown that the nitrogen fixation by *Nostoc azollae* can fully sustain these high growth rates. Nitrogen-fixation rates reached the equivalent of 1200 kg Nitrogen ha^{-1} $year^{-1}$. Why *Nostoc azollae* fixes nitrogen so efficiently is not yet fully understood. We have shown that atmospheric nitrogen is fixed during the day, indicating *N. azollae* uses its own photosynthesis apparatus for nitrogen fixation similar to free living filamentous cyanobacteria (Peterson et al., 1981). Interestingly the light absorption of *N. azollae* has been shown to be complementary to that of *Azolla*, which may partly explain why the *Azolla-Nostoc* symbiosis can so effectively fix both carbon and nitrogen (Ray et al., 1979). Nitrogen fixation rates were

such that they sustained *Azolla* at the peak productivities obtained experimentally and addition of external nitrogen fertilizer did not meaningfully boost the productivity: hence adding nitrogen fertilizer makes little sense for *Azolla* cultivation.

If exogenous nitrogen is present in the environment, when for example, growing *Azolla* on wastewater, it depends on the form of nitrogen and concentrations thereof, how the plant-cyanobacteria symbiosis will respond. Ammonium was rapidly taken up by *A. filiculoides* during growth experiments presented in Chapter 3, in line with previous results on *A. pinnata* that showed that ammonium uptake is much more rapid than nitrate uptake (Ito and Watanabe, 1983). Previous results further suggest that ammonia is responsible for the inhibition of nitrogen fixation by 2 mM NH_4NO_3 observed in Chapter 3 (Ito and Watanabe, 1983; Sah et al., 1989). We, however, observed ammonia toxicity when supplied as 4 mM NH_4Cl to productive cultures, whereas other authors where able to grow *Azolla* on 5-20 mM NH_4 in batch experiments (Sah et al., 1989; Maejima et al., 2001; Maejima et al., 2002). It was previously shown that the tolerance to ammonium was strain specific and an *A. filiculoides* strain was among the least tolerant ones (Maejima et al., 2001). In our productive cultures we observed quick medium acidification due to ammonium uptake. The initial acidification of the medium, or the high amounts of potassium hydroxide used to stabilize the pH, may have contributed to the toxicity observed and can be countered by an increasing medium buffer capacity. Nonetheless ammonium toxicity is an important factor to take into account when growing *Azolla* on ammonia-rich wastewaters.

In the case of nitrate, we observed zero uptake during growth experiments presented in Chapter 4. In Chapter 3, however, we did observe nitrate removal from the growth medium and a beneficial effect of nitrate on plant growth and nitrogen content. A major difference is that in experiments presented in Chapter 4 the growth medium was aerated, which was not the case in the Chapter 3 experiments. Recent experiments in our group indicated high levels of denitrification when *Azolla* is grown on medium containing nitrate and depleted of oxygen. Hence, the nitrate added during Chapter 3 growth experiments, may have been removed from the medium by denitrifying bacteria instead of by plant uptake. In this case the beneficial effect of nitrate addition to plant growth is not due to nitrate uptake. We observed during the Chapter 3 growth experiments that nitrate addition to the growth medium stabilized medium pH and required almost no pH correction during the course of the experiment. Whether this pH stabilizing effect is the main mechanism by which growth rate is increased when nitrate is added needs further investigation. In any case, uncontrolled denitrification during *Azolla* cultivation is unwanted as it produces the potent greenhouse gas N_2O. If *Azolla* is grown on nitrate-rich wastewaters aeration is likely required to prevent N_2O production and ensure a sustainable cultivation system. Given that the beneficial effect of nitrate is only small and likely unrelated

to nitrate uptake, nitrate-rich wastewaters can better be utilized to grow other plant species than *Azolla*. Phosphorous-rich waste streams are ideal for growing *Azolla*, since there its nitrogen-fixing potential is fully benefitted from and the cultivation system would contribute to restoring the phosphorous cycle and reduce dependency on finite phosphorous rock. Waste streams rich in ammonium and phosphorous can also be utilized, as long as ammonium concentrations are kept at non-toxic levels.

The experiments presented in this thesis are mainly performed under controlled conditions, which are required to understand the effect of single parameters on growth and nitrogen fixation. We now know what productivities *Azolla* is capable of and that this productivity can be sustained by nitrogen fixation. With the methodologies and knowledge at hand, future research should focus on introducing the complexities of real-world cultivation systems, including varying climates and the use of waste nutrients and investigate how these affect growth and nitrogen fixation.

Using chemical composition to evaluate biomass applications: a dynamic property

In this thesis we have tried to get more detailed insight in the chemical composition of *Azolla* biomass in the context of using the biomass for the production of sustainable feed, fuel and chemicals. In Chapter 4 we have shown that productive *Azolla* has a very favorable amino acid profile for use as feed, as compared to soybean meal. We have further provided a more realistic conversion factor to estimate protein content from total nitrogen content, i.e. 4.9 opposed to the standard factor of 6.25 generally used in *Azolla* literature. This conversion factor is similar for many novel feed alternatives including duckweed (4.8) and seaweeds (4.6) (Zhao et al., 2014; Bikker et al., 2016). Even though the true protein content of *Azolla* is lower than generally reported, the amount of protein that can be produced annually per hectare using *Azolla* is still 3.5-4.5 times higher than when using soybean. Surprisingly, adding CO_2 during growth further increased the protein content of *Azolla*, which would boost protein yields to 11.8 t dw ha^{-1} $year^{-1}$, nearly 6 fold that of soybean. Using *Azolla*, instead of soybean for animal feed could thereby help to improve land-use efficiency, mitigate effects on nitrogen and phosphate cycles and decrease CO_2 emissions associated with soybean transport when performed locally.

Feeding trails indicate that dried *Azolla* can be included in the diet of chicken (5-10%), sheep (6%), pigs (11-20%) and fish (25%), without negative effects on animal growth or digestibility (Becerra et al., 1990; Alalade and Iyayi, 2006; Basak et al., 2002; Naghshi et al., 2014; Datta, 2011; Abdel-Tawwab, 2008; Acharya et al., 2015; Cherryl et al., 2014; Leterme et al., 2010; Ahmed et al., 2016). These studies thus already show a substantial potential for introducing *Azolla* in diets of animals raised for meat or dairy production, but this is not yet performed at large scale. To

enhance the use of *Azolla* in feed, the inclusion rate can be increased to levels used for commercial protein feeds by eliminating anti-nutritional factors. The amino acid profile and lignin content, have among others, been put forward as possible causes for the limited inclusion rate, but the results presented in this thesis suggest neither amino acid profile, nor lignin content, to be responsible. Rather the high amounts of soluble polyphenols, and especially tannins, are obvious candidates for limiting the digestibility of *Azolla*. We have shown condensed tannins to amount to 5 and 8% of the dry weight of *A. filiculoides* and *A. pinnata*, respectively (Chapter 6). Condensed tannins are well studied anti-nutritional factors, especially when present in high concentrations, such as in *Azolla*. Both polyphenol and protein content therefore determine the suitability of *Azolla* as feed. In Chapter 4 we showed that not only the choice of species, but also the growth conditions strongly affect polyphenol content of *Azolla*. Whereas adding CO_2 increases protein yield of *A. filiculoides*, it also doubles polyphenol content.

Our work further indicates an inverse relation between lipid content and growth speed, reaching 12.8%, 10.55% and 7.9% for *A. filiculoides* growing at a rate of 3.4, 9.7 and 13.2 g dw m^{-2} day^{-1} (Chapter 6 and Chapter 4). Similarly, high lipid contents of 16.4% were measured in *A. caroliniana* growing in winter versus 14.3% in autumn and 12.7% in summer in Florence, Italy (Paoletti et al., 1987). Of fast growing *A. filiculoides*, 41% of the lipids consisted of fatty acids that could be converted to biodiesel (Chapter 5). Even though biodiesel from *Azolla* can meet EU quality standards, the inverse relation between growth and lipid content makes that the amount of biodiesel that can be produced per hectare is much lower compared to microalgae. The fact that lipids need to be extracted using (expensive) solvents, suggests that growing *Azolla* solely for biodiesel production will likely result in less added value and environmental benefits compared to many other biodiesel feedstocks. However, as a side product of a bio-refinery biodiesel production may be feasible.

Although we have tried to better link productivity and chemical composition, we have limited our experiments to constant climatic conditions. Although such conditions may be obtainable in some tropical regions or climate controlled greenhouses, in many low-cost cultivation systems growth conditions will change over time. Effectively the suitability of *Azolla* biomass for a certain application, may also change over time. If for example, biomass becomes rich in polyphenols and lipids during autumn, the production of chemicals and biodiesel may become more valuable than feed production. However, recent greenhouse experiments conducted in our lab have also indicated that at moderate temperatures (15°C) and sufficient light, protein content can increase to 25%, suiting feed production. Further research is needed to better understand the effect of real-world, dynamic, growth conditions, on the chemical composition of *Azolla*. Especially nitrogen, tannin and lipid contents can be easily assessed over many time intervals and directly relate to the suitability of the biomass

for a given applications.

Processing *Azolla*: from methods to processes

The step from biomass to product is essential for *Azolla* cultivation to have impact and added value. The least complex application is to directly utilize wet *Azolla* biomass, as feed for example. Processing may be used to further increase the added value of the product, but also infers further costs and environmental impact. The high moisture content of *Azolla* makes that conventional centralized processing facilities are unlikely economically feasible due to the high costs of transporting the wet biomass to these facilities. In looking for ways to process *Azolla*, we therefore should aim at local solutions. In Chapter 7 we show that 40-50% of the protein in *Azolla* is extractable by simple alkaline protein extraction. However, the presence of polyphenols, and especially condensed tannins, interferes with protein extraction and the quality of the end product. To produce added value feed products polyphenols need to be separated as well. This is a drawback compared to other feedstocks, such as duckweeds, but may also provide an opportunity when commercializing the obtained polyphenols. Condensed tannins have traditional use in the leather industry and are investigated as adhesives and in Chapter 6 we also show that *Azolla* contain relatively high levels of dicaffeoylquinic and chlorogenic acids, which have possible applications in medicine, cosmetics, foods and chemicals. To separate protein and polyphenols, we propose two extraction methods, or bio-refinery strategies, for *Azolla* in Chapter 7. Aqueous acetone pre-extraction removes 80% of the polyphenols, but requires harsh alkaline extraction to obtain the protein from the residue. The large amounts of organic solvents needed may prove a difficulty for local processing. By prioritizing polyphenol extraction, this approach relies on a high value of the extracted polyphenols, while creating a lower quality protein feed product. The second approach uses poly-ethylene glycol (PEG) during protein extraction to bind condensed tannins. This approach provides a high quality protein product, but it may prove difficult to reutilize condensed tannins. A method to extract protein and polyphenols separately from biomass would not only benefit the *Azolla* production chain, but also the utilization of many other feeds that are rich tannins (Makkar, 2003).

The bio-refinery approaches presented in Chapter 7 provide a possible direction for accomplishing the separation of protein and polyphenols, but much further research is needed to develop economical feasible and scalable processes. In the case of acetone extraction, further purification of the lipid-polyphenol mixture is needed to obtain marketable products. Such a purification approach may be very similar to the analytical procedure used in Chapter 6. The quality of the feed product may be improved by using enzymatic protein degradation instead of thermochemical deg-

radation. In the case of PEG aided protein extraction, the fate of the PEG-polyphenol complexes during extraction needs to be understood and controlled. The latter can be accomplished by selective separation or filtration of PEG-tannin complexes over protein, or by using PEG bound to solid support structures, easy recoverable end-groups or magnetic beads. Subsequently reutilization of the PEG will likely be required to make the process economically feasible and to generate the polyphenol by-product. The latter will be dominated by condensed tannins, but may also contain smaller polyphenols such as dicaffeoylquinic acids as their interaction with PEG is yet unknown.

The proposed bio-refinery concepts are both based on wet processing of *Azolla* to separate protein and polyphenols, which is a logical starting-point as it starts with utilizing the highest-value components and any prior manipulations are likely to invoke reactions between the two, lowering their value. Further research may subsequently look at commercialization of residual processing streams. The polysaccharide-rich insoluble fraction is particularly interesting in this regards, due to its high C_6 sugar content, making it suitable for fermentation into chemical building blocks for the petrochemical industry or bioethanol as transportation fuel.

Additionally, further research is needed to assess the quality of the products generated. The protein concentrates obtained from experiments in Chapter 7 have a lower relative polyphenol content compared to whole *Azolla*, but this does not directly imply better digestibility, as the presence of other anti-nutritional factors cannot be ruled out. Digestibility studies are thus needed to confirm the enhanced value of *Azolla* protein concentrates.

Finally, linking the knowledge on *Azolla*'s molecular biology to the quality of end-products allows to improve the whole production chain at an accelerated pace. The discovery of a Leucoanthocyanidin reductase-like gene in *Azolla* (AzfiLAR-like) responsible for the high production of condensed tannins, allows us to use targeted breeding approached to obtain *Azolla* without condensed tannins. Although effects on plant growth and susceptibility to herbivores are likely to occur, biomass of an *Azolla* lar-mutant could likely be included in the diet of animals at higher rates and be easier processed into protein concentrates, improving its value for feed and food production.

Azolla with respect to other (novel) sustainable feedstocks

An *Azolla* production chain is one of many novel plant feedstocks that may aid in providing sufficient food, fuel and materials in the future, while reducing the environmental impact with respect to current production systems. Each of these plant feedstocks has its own limitations in terms of where they can be cultivated

and their preferable application, which is mainly determined by their chemical composition. It is unlikely that one alternative will become dominant. Rather a range of alternatives will be used, each addressing a specific niche in terms of cultivation area and/or application. Since *Azolla* is a free-floating aquatic plant, its cultivation is restricted to areas with sufficient fresh-water availability. Closed production systems may be used to extend cultivation to areas where water is less available. Using *Azolla* to remove excess nutrients and/or pollutants from waterways may actually benefit drinking water availability in some of these areas. Its nitrogen fixing capability makes it ideally suited for cultivation on waste streams rich in phosphorous that would otherwise not be utilized, which would lead to loss of phosphorous as run-off or retained in the soil as e.g. iron phosphates. Subsiding agricultural areas is one example, where cultivation of *Azolla* can aid in stopping subsidence by submerging the soil and at the same time recover phosphorous from the soil. In cases where nitrogen is the dominant waste nutrient other free-floating plants may be equally or better alternatives, depending on the preferred application of the biomass.

The chemical composition of *Azolla* favors its use for feed production. The high protein content, favorable amino acid profile and high yield per hectare are major advantages over the production of plant protein using soybean. However, to fully replace soybean meal the tannin content of the produced feed, and possibly other anti-nutritional factors, need to be reduced by processing or breeding to obtain a product with similar or higher quality as soybean meal. Macroalgae and duckweeds are other promising feed alternatives with similarly high protein contents as *Azolla* (Chapter 7). As for *Azolla*, the inclusion rates proposed for macroalgea and duckweeds are generally below those of commercial feeds (Bikker et al., 2016; Fasakin et al., 1999; Anderson et al., 2011; Haustein et al., 1992; Gardiner et al., 2008; Valente et al., 2006). For macroalgae the limited inclusion rate has been ascribed to the high mineral content, whereas anti-nutritional factors in duckweeds are yet unknown (Bikker et al., 2016; Gardiner et al., 2008). Hence these alternatives may also require processing or breeding efforts to enhance product quality. However, processing of *Azolla* biomass is more complex compared to duckweeds, due to the yield limiting effect of protein-tannin interactions at low pH (Chapter 7). Protein extraction of macroalgae, requires more extensive processing, likely due to the strong cell wall that prevents release of protein from the cells (Chapter 7) (Sari et al., 2015).

Azolla may also prove to be a very suitable candidate for the production of aromatic compounds. Although lignocellulosic crops such as trees and grasses are rich in lignin, the production of homogenous product chemicals from lignin is still challenging (Rinaldi et al., 2016; Bugg and Rahmanpour, 2015; Wong, 2009) . The soluble, less complex, polyphenols in *Azolla* are easier to extract and process, while still representing a significant annual production volume. In contrast, for the production of fuels and chemicals from polysaccharides, macroalgae, agricultural residues,

forestry residues and grasses may be better alternatives than *Azolla* as they generally contain higher amounts of polysaccharides (Suutari et al., 2015; Lee and Kuan, 2015; Frankó et al., 2016; Raj et al., 2015). Except for macroalgae, water content is also much lower, favoring centralized processing. Additionally, for the production of biodiesel microalgae are better alternatives, due to their faster growth rate, higher lipid content and the ability to use freshwater or marine cultivation systems (Mata et al., 2010). In contrast to biodiesel, the production of long-chain mid-chain hydroxy-lipids may represent an application unique to *Azolla* biomass. However, due to the low concentrations of these compounds their value needs to be high to make extraction and purification economically feasible.

Even though more obvious feedstocks exists for the production of fuels and polysaccharide-derived chemicals, it may become feasible to generate these products from *Azolla* when integrated in a bio-refinery. In a bio-refinery, first the highest value compounds are retrieved, after which it may become economical and sustainable to also extract lower-value products. Hence, in an *Azolla* bio-refinery first protein feed and aromatic compounds may be produced and subsequently polysaccharide-derived chemicals and biofuels.

The limitations of cultivation and chemical composition of *Azolla* together dictate its possible application. A clear opportunity exists for cultivating *Azolla* on phosphorous-rich waste streams to produce firstly feed, secondly chemicals and lastly fuel. The content of this thesis has contributed knowledge and methods required to make such a production chain reality. Theoretically, such a production chain would lead to less environmental impact when replacing fossil fuels and preventing phosphorous and/or nitrogen to leach into the environment. To get a more factual benchmark of such a production chain against conventional production systems and other alternatives, life cycle analysis needs to be performed in parallel to the technological-driven research proposed above.

References

Abbasi, S.A., Nipaney, P.C., Schaumberg, G.D. (1990). Bioenergy poten- tial of eight common aquatic weeds. Biol Wastes 34:359–66.

Abdel-Tawwab, M. (2008). The preference of the omnivorous-macrophagous, *tilapia zillii* (Gervais), to consume a natural free-floating Fern, *Azolla pinnata*, J. World Aquacult. Soc. 39, 104-112.

Abou, Y., Fiogb, E.D. Micha, J. (2007). A preliminary assessment of growth and production of Nile tilapia, *Oreochromis niloticus* L., Fed *Azolla*-based-diets in earthen ponds, J. Appl. Aquac. 19, 55-69.

Acharya, P., Mohanty, G.P., Pradhan, C.R., Mishra, S.K., Beura, N.C., Moharana, B. (2015). Exploring the effects of inclusion of dietary fresh *Azolla* on the performance of White Pekin broiler ducks, Vet. World. 8 1293-1299.

Adamczyk, B., Kiikkilä, O., Kitunen, V., Smolander, A. (2014). Can we measure condensed tannins from tannin-protein complexes? - A case study with acid-butanol assay in boreal forest soil organic layer, Eur. J. Soil Biol. 64, 40-45.

Ahmed, H.A., Ganai, A.M., Beigh, Y.A., Sheikh, G.G., Reshi, P.A. (2016). Performance of growing sheep on *Azolla* based diets, Ind. J. Ani. Res. 50, 721-724.

Aiking, H. (2014). Protein production: Planet, profit, plus people? Am. J. Clin. Nutr. 100.

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

Allard, R.W. (1999). History of plant population genetics, Annu. Rev. Genet. 33, 1-27.

Alonso, D.M., Bond, J.Q., Dumesic, J.A. (2010). Catalytic conversion of biomass to biofuels, Green Chem. 12 1493-1513.

Anderson, K.E., Lowman, Z., Stomp, A.-., Chang, J. (2011). Duckweed as a feed ingredient in laying hen diets and its effect on egg production and composition, Int. J. Poult. Sci. 10, 4-7.

Anglade, J., Billen, G., Garnier, J. (2015). Relationships for estimating N_2 fixation in legumes: incidence for N balance of legume-based cropping systems in Europe. Ecosphere 6: art 37.

Antunes, A.P.M., Watkins, G.M., Duncan, J.R., (2001). Batch studies on the removal of gold(III) from aqueous solution by *Azolla filiculoides*. Biotechnology Letters 23: 249–251.

Arias-Peñaranda M.T., Cristiani-Urbina, E., Montes-Horcasitas, C., Esparza-Garciia, F., Torzillo, G., Cañizares-Villanueva, R.O. (2013). *Scenedesmus incrassatulus* CLHE-Si01: a potential source of re- newable lipid for high quality biodiesel production. Bioresour Technol 140:158–164.

Arora, A., Singh, P.K. (2003). Comparison of biomass productivity and nitrogen fixing potential of *Azolla* SPP, Biomass Bioenergy. 24, 175-178.

Atabani, A.E., Silitonga, A.S., Badruddin, I.A., Mahlia, T.M.I., Masjuki, H.H., Mekhilef, S. (2012). A comprehensive review on biodiesel as an alternative energy resource and its characteristics, Renewable Sustainable Energy Rev. 16 2070-2093.

Ayres, M.P., Clausen, T.P., MacLean Jr., S.F., Redman, A.M., Relchardt, P.B. (1997). Diversity of structure and antiherbivore activity in condensed tannins, Ecology. 78 1696-1712.

Ballester, A.-R., Tikunov, Y., Molthoff, J., Grandillo, S., Viquez-Zamora, M., de Vos, R., de Maagd, R.A., van Heusden, S., Bovy, A.G. (2016). Identification of loci affecting accumulation of secondary metabolites in tomato fruit of a *Solanum lycopersicum* × *Solanum chmielewskii* introgression line population, Frontiers in Plant Science. 7, 1428.

Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., de Pamphilis, C., Albert, V.A., Aono, N., Aoyama, T., Ambrose, B.A. et al. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. Science 332: 960–963.

Barbehenn, R.V., Peter Constabel, C. (2011). Tannins in plant-herbivore interactions, Phytochemistry. 72 1551-1565.

Bardi, U. (2009). Peak oil: The four stages of a new idea, Energy. 34, 323-326.

Barril, C., Clark, A.C., Scollary, G.R. (2012). Chemistry of ascorbic acid and sulfur dioxide as an antioxidant system relevant to white wine, Anal. Chim. Acta. 732, 186-193.

Barke, J., van der Burgh, J., van Konijnenburg-van Cittert, J.H.A., Collinson, M.E., Pearce, M.A., Bujak, J., Heilmann-Clausen, C., Speelman, E.N., van Kempen, M.M.L., Reichart, G-J., Lotter, A.F., Brinkhuis, H. (2012). Coeval Eocene blooms of the freshwater fern *Azolla* in and around Arctic and Nordic seas, Palaeogeogr. Palaeoclimatol. Palaeoecol. 337-338, 108-119.

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.

Basak, B., Pramanik, A.H., Rahman, M.S., Tarafdar, S.U., Roy, B.C. (2002). *Azolla* (*Azolla pinnata*) as a feed ingredient in broiler ration, International Journal of Poultry Science 1, 29-30-34.

Becerra, M., Murgueitio, E., Reyes, G., Preston, T.R. (1990). *Azolla filiculoides* as partial replacement for traditional protein supplements in diets for growing fattening pigs based on sugar cane juice, Livestock Research for Rural Development. 2, 15-22.

Bechthold, I., Bretz, K., Kabasci, S., Kopitzky, R., Springer, A. (2008). Succinic acid: A new platform chemical for biobased polymers from renewable resources, Chem. Eng. Technol. 31, 647-654.

Becking, J.H. (1987). Endophyte transmission and activity in the *Anabaena-Azolla* association, Plant Soil. 100, 183-212.

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

Behr, A., Westfechtel, A., Pérez Gomes, J. (2008). Catalytic processes for the technical use of natural fats and oils, Chemical Engineering and Technology. 31, 700-714.

Benjamini, Y., Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series 57: 289–300.

Berchmans, H.J., Hirata, S. (2008). Biodiesel production from crude *Jatropha* curcas *L. seed* oil with a high content of free fatty acids. Bioresour Technol 99:1716–1721.

Bhatta, R., Shinde, A.K., Vaithiyanathan, S., Sankhyan, S.K., Verma, D.L. (2002). Effect of polyethylene glycol-6000 on nutrient intake, digestion and growth of kids browsing *Prosopis cineraria*, Anim. Feed Sci. Technol. 101, 45-54.

Biermann, U., Bornscheuer, U., Meier, M.A.R., Metzger, J.O., Schäfer, H.J. (2011). Oils and fats as renewable raw materials in chemistry, Angewandte Chemie - International Edition. 50 3854-3871.

Bikker, P., van Krimpen, M.M., van Wikselaar, P., Houweling-Tan, B., Scaccia, N., van Hal, J.W., Huijgen, W.J.J., Cone, J.W., López-Contreras, A.M. (2016). Biorefinery of the green seaweed *Ulva lactuca* to produce animal feed, chemicals and biofuels, J. Appl. Phycol. 1-15.

Bläsing, O.E., Gibon, Y., Günther, M., Höhne, M., Morcuende, R., Osuna, D., Thimm, O., Usadel, B., Scheible, W-R., Stitt, M. (2005). Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in *Arabidopsis*. Plant Cell 17: 3257–3281.

Bobbink, R., Hornung, M., Roelofs, J.G.M. (1998). The effects of air-borne nitrogen pollutants on species diversity in natural and semi-natural European vegetation, J. Ecol. 86, 717-738.

Bolger, A.M., Lohse, M., Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data, Bioinformatics. 30 2114-2120.

Bos, H.L., Sanders, J.P.M. (2013). Raw material demand and sourcing options for the development of a bio-based chemical industry in Europe.: Part 1: Estimation of maximum demand, Biofuel Bioprod. Biorefining. 7, 246-259.

Börjesson, P., Mattiasson, B. (2008). Biogas as a resource-efficient vehicle fuel, Trends Biotechnol. 26, 7-13.

Boudet, A. (2007). Evolution and current status of research in phenolic compounds, Phytochemistry. 68 2722-2735.

le Bourvellec, C., Renard, C.M.G.C. (2012). Interactions between polyphenols and macromolecules: Quantification methods and mechanisms, Crit. Rev. Food Sci. Nutr. 52, 213-248.

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, K.L., Carr, K.M., Gowik, U., Mass, J., Lercher, M.J. et al. (2011a). An mRNA blueprint for C_4 photosynthesis derived from comparative transcriptomics of closely related C_3 and C_4 species. Plant Physiology 155: 142–156.

Brautigam, A., Mullick, T., Schliesky, S., Weber, A.P. (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.

Bravo, L., Goya, L., Lecumberri, E. (2007). LC/MS characterization of phenolic constituents of mate (Ilex paraguariensis, St. Hil.) and its antioxidant activity compared to commonly consumed beverages, Food Res. Int. 40, 393-405.

Bravo, M.N., Silva, S., Coelho, A.V., Boas, L.V., Bronze, M.R. (2006). Analysis of phenolic compounds in Muscatel wines produced in Portugal, Anal. Chim. Acta. 563, 84-92.

Bravo, L. (1998). Polyphenols: Chemistry, dietary sources, metabolism, and nutritional significance, Nutr. Rev. 56, 317-333.

Brennan, L., Owende, P. (2010). Biofuels from microalgae—A review of technologies for production, processing, and extractions of biofuels and co-products, Renewable and Sustainable Energy Reviews. 14, 557-577.

Breuillin-Sessoms, F., Floss, D.S., Gomez, S.K., Pumplin, N., Ding, Y., Levesque-Tremblay, V., Noar, R.D., Daniels, D.A., Bravo, A., Eaglesham, J.B., Benedito, V.A., Udvardi, M.K., Harrison, M.J. (2015). Suppression of arbuscule degeneration in *Medicago truncatula* phosphate transporter4 mutants is dependent on the ammonium transporter 2 family protein AMT2; 3. Plant Cell 27: 1352-1366.

Brinkhuis, H., Schouten, S., Collinson, M.E., Sluijs, A., Damsté, J.S.S., Dickens, G.R., Huber, M., Cronin, T.M., Onodera, J., Takahashi, K., Bujak, J.P., Stein, R., Van Der Burgh, J., Eldrett, J.S., Harding, I.C., Lotter, A.F., Sangiorgi, F., Cittert, H.V.K.-., De Leeuw, J.W., Matthiessen, J., Backman, J., Moran, K., Clemens, S., Eynaud, F., Gattacceca, J., Jakobsson, M., Jordan, R., Kaminski, M., King, J., Koc, N., Martinez, N.C., Mcinroy, D., Moore Jr, T.C., O'Regan, M., Pâlike, H., Rea, B., Rio, D., Sakamoto, T., Smith, D.C., St John, K.E.K., Suto, I., Suzuki, N., Watanabe, M., Yamamoto, M. (2006). Episodic fresh surface waters in the Eocene Arctic Ocean, Nature. 441, 606-609.

Brosse, N., Dufour, A., Meng, X., Sun, Q., Ragauskas, A. (2012).*Miscanthus*: A fast-growing crop for biofuels and chemicals production, Biofuel Bioprod. Biorefining. 6, 580-598.

Brouwer, P., Bräutigam A., Külahoglu C., Tazelaar, A.O.E., Kurz, S., Nierop, K.G.J., van der Werf, A., Weber, A.P.M., Schluepmann, H. (2014). *Azolla* domestication towards a biobased economy? New Phytologist 202: 1069–1082.

Brouwer, P., van der Werf, A., Schluepmann, H., Reichart, G.J., Nierop, K.G.J. (2015). Lipid Yield and Composition of *Azolla filiculoides*and the Implications for Biodiesel Production, Bioenergy Research, 9 (1), pp. 369-377.

Brown, N.J., Newell, C.A., Stanley, S., Chen, J.E., Perrin, A.J., Kajala, K., Hibberd, J.M. (2011). Independent and parallel recruitment of preexisting mechanisms underlying C_4 photosynthesis. Science 331: 1436–1439.

Bruijnincx, P.C.A., Weckhuysen, B.M. (2013). Shale Gas Revolution: An Opportunity for the Production of Biobased Chemicals? Angewandte Chemie International Edition. 52 11980-11987.

Bruins, M.E., Sanders, J.P. (2012). Small-scale processing of biomass for biorefinery, Biofuel Bioprod. Biorefining. 6, 135-145.

Bryant, L., Carver, L., Butler, C.D., Anage, A. (2009). Climate change and family planning: Least-developed countries define the agenda, Bull. WHO. 87, 852-857.

Buckingham, K.W., Ela, S.W., Morris, J.G., Goldman, C.R. (1978). Nutritive value of the nitrogen-fixing aquatic fern *Azolla filiculoides*, J. Agric. Food Chem. 26 1230-1234.

Budryn, G., Rachwal-Rosiak, D. (2013). Interactions of Hydroxycinnamic Acids with Proteins and Their Technological and Nutritional Implications, Food Rev. Int. 29, 217-230.

Bugg, T.D.H., Rahmanpour, R. (2015). Enzymatic conversion of lignin into renewable chemicals, Curr. Opin. Chem. Biol. 29, 10-17.

Burgard, A., Burk, M.J., Osterhout, R., van Dien, S., Yim, H. (2016). Development of a commercial scale process for production of 1,4-butanediol from sugar, Curr. Opin. Biotechnol. 42, 118-125.

Bushart, T.J., Cannon, A.E., ul Haque, A., San Miguel, P., Mostajeran, K., Clark, G.B., Marshall Porterfield, D., Roux, S.J. (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.

Butler, L.G., Price, M.L., Brotherton, J.E. (1982). Vanillin assay for proanthocyanidins (condensed tannins): Modification of the solvent for estimation of the degree of polymerization, J. Agric. Food Chem. 30 1087-1089.

Cabeza, R.A., Liese, R., Fischinger, S.A., Sulieman, S., Avenhaus, U., Lingner, A., Hein, H., Koester, B., Baumgarten, V., Dittert, K., Schulze, J. (2015). Long-term non-invasive and continuous measurements of legume nodule activity. The Plant journal : for cell and molecular biology 81: 637–48.

Canini, A., GrilliCaiola, M., Mascini, M. (1990). Ammonium content, nitrogenase activity and heterocyst frequency within the leaf cavities of *Azolla filiculoides* Lam, FEMS Microbiol. Lett. 71, 205-210.

Cao, Y., Zhang, X. (2013).Production of long-chain hydroxy fatty acids by microbial conversion. Appl Microbiol Biotechnol 97:3323–31.

Carbone, F., Preuss, A., de Vos, R.C.H., D'Amico, E., Perrotta, G., Bovy, A.G., Martens, S., Rosati, C. (2009). Developmental, genetic and environmental factors affect the expression of flavonoid genes, enzymes and metabolites in strawberry fruits, Plant, Cell and Environment. 32 (8), pp. 1117-1131.

Carrapico, 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.

Carrapiço, F. (1991). Are bacteria the third partner of the *Azolla-Anabaena* symbiosis? Plant and Soil 137:157-60.

Carriquiry, M.A., Du, X., Timilsina, G.R. (2011). Second generation biofuels: Economics and policies, Energy Policy. 39 4222-4234.

Cary, P.R., Weerts, P.G.J. (1992). Growth and nutrient composition of *Azolla pinnata* R. Brown and *Azolla filiculoides* Lamarck as affected by water temperature, nitrogen and phosphorus supply, light intensity and pH, Aquat. Bot. 43, 163-180.

Cease, A.J., Capps, K.A., Gates, K.K., Mccrackin, M.L., Nidzgorski, D.A. (2015). Consumer-driven nutrient dynamics in urban environments: The stoichiometry of human diets and waste management, Oikos. 124, 931-948.

CEN (2008). EN 14214 Automotive fuels - Fatty acid methyl esters (FAME) for diesel engines - Requirements and test methods, European committee for Standardization, Brussels.

Chambers, P.A., Lacoul, P., Murphy, K.J., Thomaz, S.M. (2008). Global diversity of aquatic macrophytes in freshwater, Hydrobiologia. 595, 9-26.

Chapple, C.C.S., Vogt, T., Ellis, B.E., Somerville, C.R. (1992). An *Arabidopsis* mutant defective in the general phenylpropanoid pathway, Plant Cell. 4 1413-1424.

Chauzar, M., Tateyama, S., Ishikura, T., Matsumoto, K., Kaneko, D., Ebitani, K., Kaneko, T. (2012). Hydrotalcites catalyze the acidolysis polymerization of phenolic acid to create highly heat-resistant bioplastics, Adv. Funct. Mater. 22 3438-3444.

Chen, H., Zhou, D., Luo, G., Zhang, S., Chen, J. (2015). Macroalgae for biofuels production: Progress and perspectives, Renewable Sustainable Energy Rev. 47, 427-437.

Cheng, W., Sakai, H., Matsushima, M., Yagi, K., Hasegawa, T. (2010). Response of the floating aquatic fern *Azolla filiculoides* to elevated CO_2, temperature, and phosphorus levels, Hydrobiologia. 656, 5-14.

Cheng, C.L., Acedo, G.N., Dewdney, J., Goodman, H.M., Conkling, M. (1991). Differential expression of the two *Arabidopsis* nitrate reductase genes. Plant Physiology 96: 275–279.

Cherryl, D.M., Prasad, R.M.V., Jagadeeswara Rao, S., Jayalaxmi, P., Eswar Rao, B. (2014). Effect of inclusion of *Azolla pinnata* on the haematological and carcass characteristics of crossbred large white yorkshire pigs, Veterinary World. 7, 78-82.

Cherubini, F., Ulgiati, S. (2010). Crop residues as raw materials for biorefinery systems - A LCA case study, Appl. Energy. 87, 47-57.

Chinnasamy, S., Ramakrishnan, B., Bhatnagar, A., Goyal, S.K., Das, K.C. (2009).Carbon and nitrogen fixation by *Anabaena fertilissima* under elevated CO_2 and temperature, J. Freshw. Ecol. 24, 587-596.

Choi, J., Park, J.-., Lee, K.-., Park, K.-., Kim, W.-., Lee, J.-., Jung, H.-., Park, H.-. (2005). In vivo anti-hepatotoxic effects of *Ligularia fischeri var. spiciformis* and the identification of the active component, 3,4-dicaffeoylquinic acid, J. Med. Food. 8, 348-352.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R.B., Piao, S., Thornton, P. (2013). Carbon and Other Biogeochemical Cycles. , Stocker, T.F., Qin, D., Plattner, G., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom; New York, USA.

Clifford, D.J., Carson, D.M., McKinney, D.E., Bortiatynski, J.M., Hatcher, P.G. (1995). A new rapid technique for the characterization of lignin in vascular plants: thermochemolysis with tetramethylammonium hydroxide (TMAH), Org. Geochem. 23, 169-175.

Clifford, M.N. (2000). Chlorogenic acids and other cinnamates - Nature, occurrence, dietary burden, absorption and metabolism, J. Sci. Food Agric. 80, 1033-1043.

Coetzee, J.A., Hill, M.P. (2008). Biological control of water hyacinth - The South African experience, EPPO Bull. 38, 458-463.

Cohen, M.F., Gurung, S., Fukuto, J.M., Yamasaki, H. (2014). Controlled free radical attack in the apoplast: A hypothesis for roles of O, N and S species in regulatory and polysaccharide cleavage events during rapid abscission by *Azolla*, Plant Sci. 217-218, 120-126.

Coley, P.D. (1988). Effects of plant growth rate and leaf lifetime on the amount and type of anti-herbivore defense, Oecologia. 74, 531-536.

Collinson, M.E., Smith, S.Y., van Konijnenburg-van Cittert, J.H.A., Batten, D.J., Refidaff, C., Barke, J., Marone, F. (2013). New observations and synthesis of paleogeneheterosporous water ferns, Int. J. Plant Sci. 174, 350-363.

Coolen, S., Proietti, S., Hickman, R., Davila Olivas, N.H., Huang, P.-., Van Verk, M.C., Van Pelt, J.A., Wittenberg, A.H.J., De Vos, M., Prins, M., Van Loon, J.J.A., Aarts, M.G.M., Dicke, M., Pieterse, C.M.J., Van Wees, S.C.M. (2016). Transcriptome dynamics of *Arabidopsis* during sequential biotic and abiotic stresses, Plant J. 86, 249-267.

Cordell, D., Rosemarin, A., Schröder J.J., Smit, A.L. (2011). Towards global phosphorus security: A systems framework for phosphorus recovery and reuse options. Chemosphere 84: 747–758.

Cordell, D., Drangert, J.-., White, S. (2009). The story of phosphorus: Global food security and food for thought, Global Environ. Change. 19, 292-305.

Costa, M.L., Santos, M.C., Carrapiço, F. (1999). Biomass characterization of *Azolla filiculoides* grown in natural ecosystems and wastewater, Hydrobiologia. 415, 323-327.

Cúneo, N.R., Gandolfo, M.A., Zamaloa, M.C., Hermsen, E. (2014). Late Cretaceous aquatic plant world in Patagonia, Argentina, PLoS ONE. 9.

Damodaran, S. (1996). Amino acids, peptides, and proteins. In: Fennema, O.R. (Ed.), Food chemistry . Marcel Dekker, Inc., pp. 321-429.

Datta, R., Henry, M. (2006). Lactic acid: Recent advances in products, processes and technologies - A review, J. Chem. Technol. Biotechnol. 81, 1119-1129.

Datta, S.N. (2011). Culture of *Azolla* and its efficacy in diet of *Labeo rohita*, Aquaculture. 310, 376-379.

Dawczynski, C., Schubert, R., Jahreis, G. (2007). Amino acids, fatty acids, and dietary fibre in edible seaweed products, Food Chem. 103, 891-899.

Dawson, C.J., Hilton, J. (2011). Fertiliser availability in a resource-limited world: Production and recycling of nitrogen and phosphorus, Food Policy. 36.

Debusk, W.F., Reddy, K.R. (1987). Growth and nutrient uptake potential of *Azollacaroliniana*willd.andSalviniarotundifoliawilld. as a function of temperature, Environ. Exp. Bot. 27, 215-221.

Demirbas, A. (2009). Political, economic and environmental impacts of biofuels: A review, Appl. Energy. 86.

Demkura, P.V., Abdala, G., Baldwin, I.T., Ballaré, C.L. (2010). Jasmonate-dependent and -independent pathways mediate specific effects of solar ultraviolet B radiation on leaf phenolics and antiherbivore defense, Plant Physiol. 152, 1084-1095.

Deinum, B., Maassen, A. (1994). Effects of drying temperature on chemical composition and in vitro digestibility of forages. Anim Feed Sci Technol 46:75–86.

Der, J.P., Barker, M.S., Wickett, N.J., de Pamphilis, C.W., Wolf, P.G. (2011). De novo characterization of the gametophyte transcriptome in bracken fern. Pteridium aquilinum. BMC Genomics 12: 99.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., Gascuel, O. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research 1: W465-9.

Deval, C.G., Mane, A.V., Joshi, N.P., Saratale, G.D. (2012). Phytoremediation potential of aquatic macrophyte*Azollacaroliniana* with references to zinc plating effluent, Emirates J. Food Agric. 24, 208-223.

Dewanji, A. (1993). Amino acid composition of leaf proteins extracted from some aquatic weeds, Journal of Agricultural and Food Chemistry. 41, 1232-1236.

Dhote, S., Dixit, S. (2009).Water quality improvement through macrophytes - A review, Environ.Monit. Assess. 152, 149-153.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication, Nature. 418, 700-707.

Diaz, R.J., Rosenberg, R. (2008). Spreading dead zones and consequences for marine ecosystems, Science. 321, 926-929.

Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L.): A review, Field Crops Res. 67, 35-49.

Dornburg, V., van Vuuren, D., van de Ven, G., Langeveld, H., Meeusen, M., Banse, M., van Oorschot, M., Ros, J., van den Born, J.G., Aiking, H., Londo, M., Mozaffarian, H., Verweij, P., Lysen, E., Faaij, A. (2010). Bioenergy revisited: Key factors in global potentials of bioenergy, Energy Environ. Sci. 3, 258-267.

Driscoll, B.T., Finan, T.M. (1993). NAD+-dependent malic enzyme of *Rhizobium meliloti* is required for symbiotic nitrogen fixation. Molecular Microbiology 7: 865–873.

Duckett, J.G., Toth, R., Soni, S.L. (1975). An ultrastructural study of the *Azolla*, *Anabaena azollae* relationship. New Phytologist 75: 111-118.

Eerhart, A.J.J.E., Faaij, A.P.C., Patel, M.K. (2012). Replacing fossil based PET with biobased PEF; Process analysis, energy and GHG balance, Energy Environ. Sci. 5, 6407-6422.

Endo, M., Shimizu, H., Nohales, M.A., Araki, T., Kay, S.A. (2014).. Tissue-specific clocks in *Arabidopsis* show asymmetric coupling. Nature 515: 419-422.

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

Ericsson, K., Huttunen, S., Nilsson, L.J., Svenningsson, P. (2004). Bioenergy policy and market development in Finland and Sweden, Energy Policy. 32, 1707-1721.

Erisman, J.W., Galloway, J.N., Seitzinger, S., Bleeker, A., Dise, N.B., Petrescu, A.M., Leach, A.M., de Vries, W. (2013). Consequences of human modification of the global nitrogen cycle. Philos. Trans. R. Soc. Lond. , B, Biol. Sci. 368.

Erisman, J.W., Sutton, M.A., Galloway, J., Klimont, Z., Winiwarter, W. (2008). How a century of ammonia synthesis changed the world. Nature Geoscience 1: 636–639.

Espiñeira, J.M., Novo Uzal, E., Gómez Ros, L.V., Carrión, J.S., Merino, F., Ros Barceló, A., Pomar, F. (2011). Distribution of lignin monomers and the evolution of lignification among lower plants, Plant Biol. 13, 59-68.

European Bioplastics (2016). Bioplastics market data (2016).

Evans, J.R., Seeman, J.R. (1989). The allocation of protein nitrogen in the photosynthetic apparatus: costs,consequences and control, Briggs, W.R. (Ed.), Photosynthesis. Alan R. Liss, New York, pp. 183-184-205.

Evenson R.E., Gollin D. (2003). Assessing the impact of the green revolution (1960 to 2000. Science 300: 758–762.

Everette, J.D., Bryant, Q.M., Green, A.M., Abbey, Y.A., Wangila, G.W., Walker, R.B. (2010).Thorough study of reactivity of various compound classes toward the folin-Ciocalteu reagent, J. Agric. Food Chem. 58, 8139-8144.

Eweedah, N., Gundel, J., Mátrai, T. (1997).Protein degradability, amino acid composition, trypsin inhibitor and urease activity of raw and heat-treated fullfat soybean. Arch. Tierernahr. 50, 361-367.

FAO (2013).Dietary protein quality evaluation in human nutrition, FAO, Rome.

FAO (2016). The State of World Fisheries and Aquaculture 2016. Contributing to food security and nutrition for all, FAO, Rome.

FAO, IFAD, WFP (2015). The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress, FAO, Rome.

Farah, A., de Paulis, T., Trugo, L.C., Martin, P.R. (2005). Effect of roasting on the formation of chlorogenic acid lactones in coffee, J. Agric. Food Chem. 53, 1505-1513.

Fasakin, E.A., Balogun, A.M. (2001). Nutritional and anti-nutritional analyses of *Azolla africana* Desv. and *Spirodela polyrrhiza* L. Schleiden as feedstuffs for fish production. In: 14th Annual Conference of the Fisheries Society of Nigeria (FISON) , 19-23 January 1998 ,Ibadan, Nigeria, pp. 31-39.

Fasakin, E.A. (1999). Nutrient quality of leaf protein concentrates produced from water fern (*Azolla africana* Desv) and duckweed (*Spirodela polyrrhiza* L. Schleiden), Bioresour. Technol. 69, 185-187.

Ferguson, R., Wilkinson, W., Hill, R. (2000). Electricity use and economic development, Energy Policy. 28, 923-934.

Filley, T.R., Minard, R.D., Hatcher, P.G. (1999). Tetramethylammonium hydroxide (TMAH) thermochemolysis: Proposed mechanisms based upon the application of ^{13}C-labeled TMAH to a synthetic model lignin dimer, Org. Geochem. 30, 607-621.

Fiorentini, R., Galoppini, C. (1983). The proteins from leaves, Plant Food Hum Nutr. 32, 335-350.

Fischedick, M., R., J., Abdel-Aziz, A., Acquaye, A., Allwood, J.M., Ceron, J.P., Geng, Y., Kheshgi, H., Lanza, A., Perczyk, D., Price, L., Santalla, E., Sheinbaum, C., Tanaka, K. (2014). Industry. , Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Farahani, E., Kadner, S., Seyboth, K., Adler, A., Baum, I., Brunner, S., Eickemeier, P., Kriemann, B., Savolainen, J., Schlömer, S., von Stechow, C., Zwickel, T., Minx, J.C. (Eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Fisher, R.A. (1922). On the interpretation of v2 from contingency tables, and the calculation of P. Journal of the Royal Statistical Society 85: 87–94.

Flach, B., Lieberz, S., Rondon, M., Williams, B., Teiken, C. (2015). EU Biofuels Annual 2015, NL5028,.

Flamini, R. (2013). Recent Applications of Mass Spectrometry in the Study of Grape and Wine Polyphenols, ISRN Spectroscopy. 2013, 45.

Floyd, K.S., Bowman, J.L. (2007). The ancestral developmental tool kit of land plants. International Journal of Plant Sciences 168: 1–35.

Fornara, F., Panigrahi, K.C.S., Gissot, L., Sauerbrunn, N., Rühl, M., Jarillo, J.A., Coupland, G. (2009). *Arabidopsis* DOF Transcription Factors Act Redundantly to Reduce CONSTANS Expression and Are Essential for a Photoperiodic Flowering Response, Developmental Cell. 17, 75-86.

Forni, C., Braglia, R., Harren, F.J.M., Cristescu, S.M. (2012). Stress responses of duckweed (*Lemna minor* L.) and water velvet (*Azolla filiculoides* Lam.) to anionic surfactant sodium-dodecyl-sulphate (SDS), Aquat. Toxicol. 110-111, 107-113.

Forni, C.I., Tel-Or, E., Bar, E., Grilli-Caiola, M. (1991). Effects of antibiotic treatments on *Azolla-Anabaena* and *Arthrobacter*. In Nitrogen Fixation, Springer Netherlands p151–156.

Forsum, O., Svennerstam, H., Ganeteg, U., Näsholm T. (2008). Capacities and constraints of amino acid utilization in *Arabidopsis*. New Phytologist 179: 1058–69.

Frankó, B., Galbe, M., Wallberg, O. (2016). Bioethanol production from forestry residues: A comparative techno-economic analysis, Appl. Energy. 184, 727-736.

Fraser, C.M., Chapple, C. (2011). The Phenylpropanoid Pathway in *Arabidopsis*, American Society of Plant Biologists (Ed.), The *Arabidopsis* Book.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W.J., Rosenthal, A., Perret, X. (1997). Molecular basis of symbiosis between *Rhizobium* and legumes. Nature 387: 394–401.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez Jr., M., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A., Wang, Z.-. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass, Proc. Natl. Acad. Sci. U. S. A. 108, 3803-3808.

Funk, J.L., Glenwinkel, L.A., Sack, L. (2013). Differential Allocation to Photosynthetic and Non-Photosynthetic Nitrogen Fractions among Native and Invasive Species, PLoS ONE. 8.

Gale, J., Smernoff, D.T., Macler, B.A., MacElroy, R.D. (1989). Carbon balance and productivity of *Lemna gibba*, a candidate plant for CELSS, Adv. Space Res. 9, 43-52.

Galletti, G.C., Modafferi, V., Poiana, M., Bocchini, P. (1995). Analytical pyrolysis and thermally assisted hydrolysis-methylation of wine tannin, J. Agric. Food Chem. 43, 1859-1863.

Gallon, J.R. (1981). The oxygen sensitivity of nitrogenase: a problem for biochemists and micro-organisms. Trends in Biochemical Sciences 6: 19–23.

Galloway, J.N., Dentener, F.J., Capone, D.G., Boyer, E.W., Howarth, R.W., Seitzinger, S.P., Asner, G.P., Cleveland, C.C., Green, P.A., Holland, E.A., Karl, D.M., Michaels, A.F., Porter, J.H., Townsend, A.R., Vöosmarty C.J. (2004). Nitrogen Cycles: Past, Present, and Future. Biogeochemistry 70: 153–226

Gao, K., McKinley, K.R. (1994). Use of macroalgae for marine biomass production and CO_2 remediation: a review, J. Appl. Phycol. 6, 45-60.

Gardiner, G.E., Campbell, A.J., O'Doherty, J.V., Pierce, E., Lynch, P.B., Leonard, F.C., Stanton, C., Ross, R.P., Lawlor, P.G. (2008). Effect of *Ascophyllum nodosum* extract on growth performance, digestibility, carcass characteristics and selected intestinal microflora populations of grower-finisher pigs, Anim. Feed Sci. Technol. 141, 259-273.

Garnier, N., Richardin, P., Cheynier, V., Regert, M. (2003). Characterization of thermally assisted hydrolysis and methylation products of polyphenols from modern and archaeological vine derivatives using gas chromatography-mass spectrometry, Anal. Chim. Acta. 493, 137-157.

Gassmann, A., Cock, M.J.W., Shaw, R., Evans, H.C. (2006). The potential for biological control of invasive alien aquatic weeds in Europe: A review, Hydrobiologia. 570, 217-222.

Gelfand, I.R.G.P. (2015). A reassessment of the contribution of soybean biological nitrogen fixation to reactive N in the environment. Biogeochemistry 123: 175–184.

van Gerven, L.P.A., de Klein, J.J.M., Gerla, D.J., Kooi, B.W., Kuiper, J.J., Mooij, W.M. (2015). Competition for light and nutrients in layered communities of aquatic plants, Am. Nat. 186, 72-83.

GG-MOVE (2011). Future transport fuels: Report of the European experts group on future transport fuels. European Comission, Brussels.

Gierlinger, S. (2014). Food and feed supply and waste disposal in the industrializing city of Vienna (1830–1913): a special focus on urban nitrogen flows, Reg. Environ. Change. 15, 317-327.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people, Science. 327, 812-818.

Godin, B., Agneessens, R., Gerin, P., Delcarte, J. (2014). Structural carbohydrates in a plant biomass: Correlations between the detergent fiber and dietary fiber methods, J. Agric. Food Chem. 62, 5609-5616.

Goiris, K., Muylaert, K., Voorspoels, S., Noten, B., De Paepe, D., E Baart, G.J., De Cooman, L. (2014). Detection of flavonoids in microalgae from different evolutionary lineages, J. Phycol. 50, 483-492.

Goudriaan, J., Monteith, J.L. (1990). A Mathematical Function for Crop Growth Based on Light Interception and Leaf Area Expansion, Annals of Botany. 66, 695-701.

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

Graça, J. (2010). Hydroxycinnamates in suberin formation, Phytochem. Rev. 9, 85-91.

Grasso, M. (2016). The Political Feasibility of Consumption-Based Carbon Accounting, New Polit. Econ. 21, 401-413.

Grisel, R.J.H., van der Waal, J.C., de Jong, E., Huijgen, W.J.J. (2014). Acid catalysed alcoholysis of wheat straw: Towards second generation furan-derivatives, Catal Today. 223, 3-10.

Gruber, N., Galloway, J.N. (2008). An Earth-system perspective of the global nitrogen cycle, Nature. 451, 293-296.

Gupta, V.K., Kubicek, C.P., Berrin, J-G., Wilson, D.W., Couturier, M., Berlin, A., Filho, E.X.F., Ezeji, T. (2016). Fungal Enzymes for Bio-Products from Sustainable and Waste Biomass, Trends Biochem. Sci. 41, 633-645.

Gutiérrez R., Stokes, T.L., Thum, K., Xu, X., Obertello, M., Katari, M.S., Tanurdzic, M., Dean, A., Nero, D.C., McClung, C.R., Coruzzi, G.M. (2008). Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. Proceedings

of the National Academy of Sciences of the United States of America 105: 4939–4944.
Hagerman, A.E., Butler, L.G. (1981). The specificity of proanthocyanidin-protein interactions. J. Biol. Chem. 256, 4494-4497.
Häkkinen, S.H., Kärenlampi, S.O., Heinonen, I.M., Mykkänen, H.M., Törrönen, A.R. (1998). HPLC method for screening of flavonoids and phenolic acids in berries, J. Sci. Food Agric. 77, 543-551.
Hall, D.O., Markov, S.A., Watanabe, Y., Krishna Rao, K. (1995).The potential applications of cyanobacterial photosynthesis for clean technologies, Photosynth Res. 46, 159-167.
Hall, J.W. (1969). Studies on Fossil *Azolla*: Primitive Types of Megaspores and Massulae from the Cretaceous, Am. J. Bot. 56, 1173-1180.
Hall, J.W., Swanson, N.P. (1968). Studies on Fossil *Azolla*: *Azolla montana*, a Cretaceous Megaspore with Many Small Floats, Am. J. Bot. 55, 1055-1061.
Hamisch, D., Randewig, D., Schliesky, S., Brautigam, A., Weber, A.P.M., Geffers, R., Herschbach, C., Rennenberg, H., Mendel, R.R., 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.
Hargrove, J.L., Greenspan, P., Hartle, D.K. (2004). Nutritional signifi- cance and metabolism of very long chain fatty alcohols and acids from dietary waxes. Exp Biol Med 229:215–26.
Hartmann, D.L., Klein Tank, A.M.G., Rusticucci, M., Alexander, L.V., Brönnimann, S., Charabi, Y., Dentener, F.J., Dlugokencky, E.J., Easterling, D.R., Kaplan, A., Soden, B.J., Thorne, P.W., Wild, M., Zhai, P.M. (2013). Observations: Atmosphere and Surface. In: Stocker, T.F., Qin, D., Plattner, G., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
Hartman, G.L., West, E.D., Herman, T.K. (2011). Crops that feed the World 2. Soybean—worldwide production, use, and constraints caused by pathogens and pests. Food Security 3: 5–17.
Haustein, A.T., Gilman, J.B., Gastañaduy, A. (1992). Compensatory growth in broiler chicks fed on *Lemna gibba*, Br. J. Nutr. 68, 329-335.
Hedman, H., Källman, T., Lagercrantz, U. (2009). Early evolution of the MFT-like gene family in plants. Plant Molecular Biology 70: 359–369.
Henry-Silva, G.G., Camargo, A.F.M., Pezzato, M.M. (2008). Growth of free-floating aquatic macrophytes in different concentrations of nutrients, Hydrobiologia. 610, 153-160.
Herd, Y.R., Cutter, E.G., 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, Y.R., Cutter, E.G., 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.
Herendeen, P. (1999). A Preliminary Conspectus of the Allon Flora from the Late Cretaceous (Late Santonian) of Central Georgia, U.S.*A. on* JSTOR. Annals of the Missouri Botanical Garden 86:407–471.
Holm, K., Källman T., Gyllenstrand, N., Hedman, H., Lagercrantz, U. (2010). Does the core circadian clock in the moss *Physcomitrella patens* (Bryophyta) comprise a single loop? BMC Plant Biology 10: 109.
Hilscher, J., Bürstmayr, H., Stoger, E. (2017). Targeted modification of plant genomes for precision crop breeding, Biotechnol. J. 12.
Holland, E.A., Braswell, B.H., Sulzman, J., Lamarque, J-F. (2005). Nitrogen deposition onto the United States and Western Europe: Synthesis of observations and models, Ecol. Appl. 15, 38-57.
van der Hooft, J.J.J., Vervoort, J., Bino, R.J., Beekwilder,J., de Vos, R.C.H. (2012). Spectral trees as a robust annotation tool in LC–MS based metabolomics, Metabolomics. 8 (4), 691-703.
van Hove, C. (1989). *Azolla*: and its multiple uses with emphasis on Africa. FAO, Rome, Italy.
van Hove, C., Lejeune, A. (2002). The*Azolla-Anabaena* symbiosis, Biol. Environ. 102, 23-26.
Hung, N.V.P., Watanabe, S., Ishikawa, S., Ohtake, N., Sueyoshi, K., Sato, T., Ishii, S., Fujimaki, S., Ohyama, T. (2013). Quantitative analysis of the initial transport of fixed nitrogen in nodulated soybean plants using ^{15}N as a tracer. Soil Science and Plant Nutrition 59: 888-895.
Hungate, B.A., Dijkstra, P., Johnson, D.W., Hinkle, C.R., Drake, B.G. (1999). Elevated CO_2 increases

nitrogen fixation and decreases soil nitrogen mineralization in Florida scrub oak, Global Change Biol. 5, 781-789.

Hussain, J., Ruan, Z., Nascimento, I.A., Liu, Y., Liao, W. (2014). Lipid profiling and corresponding biodiesel quality of Mortierella isabellina using different drying and extraction methods. Bioresour Technol 169:768–72.

Hwang, T.-., Kashiwada, Y., Nonaka, G.-., Nishioka, I. (1990). 4-Carboxymethyl flavan-3-ols and procyanidins from Davallia divaricata, Phytochemistry. 29, 279-282.

IEA (2016a). Key Coal Trends. Excerpt from: Coal Information. International Energy Agency, Paris.

IEA (2016b). Key oil trends. Excerpt from: Oil information, International Energy Agency, Paris.

IEA (2016c). World Energy Statistics 2016, International Energy Agency, Paris.

IPCC (2013).Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA (1535 pp).

IPCC (2014). Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, USA.

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 University Press, Cambridge, United Kingdom and New York, USA.

Ishikura, N. (1982). 3-Desoxyanthocyanin and other phenolics in the water fern *Azolla*, Bot Mag Tokyo. 95, 303-308.

Ismail, G.S.M., Mohamed, H.E. (2010). Alteration in growth and thylakoid membrane lipid composition of *Azolla caroliniana* under phosphate deficiency, Biol. Plant. 54, 671-676.

Ito, O., Watanabe, I. (1983). The relationship between combined nitrogen uptakes and nitrogen fixation in *Azolla-Anabaena* symbiosis. New Phytologist 95: 647–654.

Ivanic, M., Martin, W. (2008). Implications of higher global food prices for poverty in low-income countries, Agric. Econ. 39, 405-416.

Izaguirre, M.M., Mazza, C.A., Svatoš, A., Baldwin, I.T., Ballaré, C.L. (2007). Solar ultraviolet-B radiation and insect herbivory trigger partially overlapping phenolic responses in *Nicotiana attenuata* and *Nicotiana longiflora*, Ann. Bot. 99, 103-109.

Jacob, C., Carrasco, B., Schwember, A.R. (2016). Advances in breeding and biotechnology of legume crops, Plant Cell Tissue Organ Cult. 127, 561-584.

Jain, S.K., Gujral, G.S., Jha, N.K., Vasudevan, P. (1992). Production of biogas from *Azolla pinnata* R.Br and *Lemna minor* L.: effect of heavy metal contamination. Bioresour Technol 41:273–7

Jampeetong, A., Brix, H. (2009). Effects of NH_4 + concentration on growth, morphology and NH_4 + uptake kinetics of *Salvinia natans*, Ecol. Eng. 35, 695-702.

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

Janse, J.H. (1998). A model of ditch vegetation in relation to eutrophication, Water Sci. Technol. 37, 139-149.

Jansman, A.J., Verstegen, M.W., Huisman, J., van den Berg, J.W. (1995). Effects of hulls of faba beans (*Vicia faba* L.) with a low or high content of condensed tannins on the apparent ileal and fecal digestibility of nutrients and the excretion of endogenous protein in ileal digesta and feces of pigs. J. Anim. Sci. 73, 118-127.

Jensen, E.S., Peoples, M.B., Boddey, R.M., Gresshoff, P.M., Hauggaard-Nielsen, H., Alves, B.J.R., Morrison, M.J. (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.

Jha, P.K., Halada, G.P. (2011). The catalytic role of uranyl in formation of polycatechol complexes, Chem. Cent. J. 5.

Jones, C.G., Firn, R.D. (1979). Some allelochemicals of *Pteridium aquilinum* and their involvement in

resistance to *Pieris brassicae*, Biochem. Syst. Ecol. 7, 187-192.

de Jong, S., Hoefnagels, R., van Stralen, J., Londo, M., Slade, R., Faaij, A., Junginger, M. (2017). Renewable Jet Fuel in the European Union – Scenarios and Preconditions for Renewable Jet Fuel Deployment towards 2030.

Jung, J.H., Ju, Y., Seo, P.J., Lee, J.H., Park, C.M. (2012). The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in *Arabidopsis*. Plant Journal 69: 577–588.

Kaack, K., Austed, T. (1998). Interaction of vitamin C and flavonoids in elderberry (*Sambucus nigra* L.) during juice processing, Plant Foods Hum. Nutr. 52, 187-198.

Kajala, K., Brown, N.J., Williams, B.P., Borrill, P., Taylor, L.E., Hibberd, J.M. (2012) Multiple *Arabidopsis* genes primed for recruitment into C_4 photosynthesis. Plant Journal 69: 47–56.

Kalili, K.M., Vestner, J., Stander, M.A., De Villiers, A. (2013). Toward unraveling grape tannin composition: Application of online hydrophilic interaction chromatography × reversed-phase liquid chromatography-time-of-flight mass spectrometry for grape seed analysis, Anal. Chem. 85, 9107-9115.

Kampschreur, M.J., Temmink, H., Kleerebezem, R., Jetten, M.S.M., van Loosdrecht, M.C.M. (2009). Nitrous oxide emission during wastewater treatment, Water Res. 43, 4093-4103.

Kaplan, D., Peters, G.A. (1981). The *Azolla-Anabaena azollae* relationship. X. $^{15}N_2$ fixation and transport in main stem axes. New Phytologist 89: 337–346.

Kar, P.P., Mishra, S., Singh, D.P. (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.

Karaköse, H., Jaiswal, R., Kuhnert, N. (2011). Characterization and quantification of hydroxycinnamate derivatives in *Stevia rebaudiana* leaves by LC-MS n, J. Agric. Food Chem. 59, 10143-10150.

Karmakar, A., Karmakar, S., Mukherjee, S. (2010). Properties of various plants and animals feedstocks for biodiesel production. Bioresour Technol 101:7201–10.

Karr-Lilienthal, L.K., Grieshop, C.M., Spears, J.K., Fahey Jr., G.C. (2005). Amino acid, carbohydrate, and fat composition of soybean meals prepared at 55 commercial US soybean processing plants, J. Agric. Food Chem. 53, 2146-2150.

Kawai, Y., Ono, E., Mizutani, M. (2014). Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants, Plant J. 78, 328-343.

Keeling, R.F. (2008). Atmospheric science: Recording earth's vital signs, Science. 319, 1771-1772.

van Kempen, M.M.L., Smolders, A.J.P., Bögemann, G.M., Lamers, L.P.M., Roelofs, J.G.M. (2016). Interacting effects of atmospheric CO_2 enrichment and solar radiation on growth of the aquatic fern *Azolla filiculoides*, Freshw. Biol. 61, 596-606.

van Kempen, M.M.L., Smolders, A.J.P., Bögemann, G.M., Lamers, L.L.M., Visser, E.J.W., Roelofs, J.G.M. (2013). Responses of the *Azolla filiculoides* Stras. - *Anabaena azollae* Lam. association to elevated sodium chloride concentrations: Amino acids as indicators for salt stress and tipping point, Aquatic Bot. 106, 20-28.

Kemppainen, K., Siika-aho, M., Pattathil, S., Giovando, S., Kruus, K. (2014). Spruce bark as an industrial source of condensed tannins and non-cellulosic sugars, Ind. Crops Prod. 52, 158-168.

Kent, W.J. (2002). BLAT – The BLAST-like alignment tool. Genome Research 12: 656–664.

Khoddami, A., Wilkes, M.A., Roberts, T.H. (2013). Techniques for analysis of plant phenolic compounds, Molecules. 18, 2328-2375.

Khush, G.S. (2001). Green revolution: The way forward, Nat. Rev. Gen. 2, 815-822.

King, A.J., He, W., Cuevas, J.A., Freudenberger, M., Ramiaramanana, D., Graham, I.A. (2009). Potential of *Jatropha* curcas as a source of renewable oil and animal feed, J. Exp. Bot. 60, 2897-2905.

Kolniak-Ostek, J., Oszmianski, J., Wojdylo, A. (2013). Effect of l-ascorbic acid addition on quality, polyphenolic compounds and antioxidant capacity of cloudy apple juices, Eur. Food Res. Technol. 236, 777-798.

Kösesakal, T. (2014). Effects of seasonal changes on pigment composition of *Azolla filiculoides* lam, Am. Fern J. 104, 58-66.

Kosten, S., Piñeiro, M., de Goede, E., de Klein, J., Lamers, L.P.M., Ettwig, K. (2016). Fate of methane in aquatic systems dominated by free-floating plants, Water Research. 104, pp. 200-207.

Kistner, C., Parniske, M. (2002). Evolution of signal transduction in intracellular symbiosis. Trends in Plant Science 7: 511–518.

Knack, J.J., Wilcox, L.W., Delaux, P.M., Ané J.M., Piotrowski, M.J., Cook, M.E., Graham, J.M., Graham, L.E. (2015). Microbiomes of streptophyte algae and bryophytes suggest that a functional suite of microbiota fostered plant colonization of land. International Journal of Plant Sciences 176: 405-20.

Kornsteiner, M., Wagner, K.-.,Elmadfa, I. (2006). Tocopherols and total phenolics in 10 different nut types, Food Chem. 98, 381-387.

Kraus, T.E.C., Yu, Z., Preston, C.M., Dahlgren, R.A., Zasoski, R.J. (2003). Linking chemical reactivity and protein precipitation to structural characteristics of foliar tannins, J. Chem. Ecol. 29, 703-730.

Kriticos, D.J., Brunel, S. (2016). Assessing and managing the current and future pest risk from water hyacinth, (Eichhornia *crassipes*), an invasive aquatic plant threatening the environment and water security, PLoS ONE. 11.

Kuroda, K.-., Nakagawa-Izumi, A., Dimmel, D.R. (2002). Pyrolysis of lignin in the presence of tetramethylammonium hydroxide (TMAH): Products stemming from ß-5 substructures, J. Agric. Food Chem. 50, 3396-3400.

Kuroda, K.-., Suzuki, A., Kato, M., Imai, K. (1995). Analysis of rice (*Oryza sativa* L.) lignin by pyrolysis-gas chromatography, J. Anal. Appl. Pyrolysis. 34, 1-12.

Kuroda, K.-., Inoue, Y., Sakai, K. (1990). Analysis of lignin by pyrolysis-gas chromatography. I. Effect of inorganic substances on guaiacol-derivative yield from softwoods and their lignins, J. Anal. Appl. Pyrolysis. 18, 59-69.

Lallemand, L.A., Zubieta, C., Lee, S.G., Wang, Y., Acajjaoui, S., Timmins, J., McSweeney, S., Jez, J.M., McCarthy, J.G., McCarthy, A.A. (2012). A structural basis for the biosynthesis of the major chlorogenic acids found in coffee, Plant Physiol. 160, 249-260.

Laplaze, L., Gherbi, H., Frutz, T., Pawlowski, K., Franche, C., Macheix, J.-., Auguy, F., Bogusz, D., Duhoux, E. (1999). Flavan-containing cells delimit Frankia-infected compartments in *Casuarina glauca* nodules, Plant Physiol. 121, 113-122.

Lapuerta, M., Rodríguez-Fernández, J., de Mora, E.F. (2009) Correlation for the estimation of the cetane number of biodiesel fuels and implications on the iodine number. Energy Policy 37: 4337–44

Lapuerta, M., Rodríguez-Fernández, J., Armas, O. (2010) Correlation for the estimation of the density of fatty acid esters fuels and its implications. A proposed biodiesel cetane index. Chem Phys Lipids 163:720–7.

Lattanzio, V., Kroon, P.A., Linsalata, V., Cardinali, A. (2009). Globe artichoke: A functional food and source of nutraceutical ingredients, J. Funct. Foods. 1, 131-144.

Lee, J.W., Kim, H.U., Choi, S., Yi, J., Lee, S.Y. (2011). Microbial production of building block chemicals and polymers, Curr. Opin. Biotechnol. 22, 758-767.

Lee, W.-., Kuan, W.-. (2015). *Miscanthus* as cellulosic biomass for bioethanol production, Biotechnol. J. 10, 840-854.

Les, D.H., Crawford, D.J., Landolt, E., Gabel, J.D., Kimball, R.T. (2002). Phylogeny and systematics of Lemnaceae, the duckweed family, Syst. Bot. 27, 221-240.

Leterme, P., Londoño, A.M., Ordoñez, D.C., Rosales, A., Estrada, F., Bindelle, J., Buldgen, A. (2010). Nutritional value and intake of aquatic ferns (*Azolla filiculoides* Lam. and *Salvinia molesta* Mitchell.) in sows, Anim. Feed Sci. Technol. 155, 55-64.

Letunic, I., Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Research gkw290.

Levitan, O., Rosenberg, G., Setlik, I., Setlikova, E., Grigel, J., Klepetar, J., Prasil, O., Berman-Frank, I. (2007). Elevated CO_2 enhances nitrogen fixation and growth in the marine cyanobacterium Trichodesmium, Global Change Biol. 13, 531-538.

Leung, D.Y.C., Wu, X., Leung, M.K.H. (2010). A review on biodiesel production using catalyzed transesterification. Appl Energy 87: 1083–95.

Lewandowski, I., Clifton-Brown, J., Trindade, L.M., Van Der Linden, G.C., Schwarz, K.-., Müller-Sämann, K., Anisimov, A., Chen, C.-., Dolstra, O., Donnison, I.S., Farrar, K., Fonteyne, S., Harding, G., Hastings, A., Huxley, L.M., Iqbal, Y., Khokhlov, N., Kiesel, A., Lootens, P., Meyer, H., Mos, M., Muylle, H., Nunn, C., Özgüven, M., Roldán-Ruiz, I., Schüle, H., Tarakanov, I., Der Weijde, T., Wagner, M., Xi, Q., Kalinina, O. (2016). Progress on optimizing miscanthus biomass production for the european bioeconomy: Results of the EU FP_7 project OPTIMISC, Front. Plant Sci. 7.

Lewandowski, I., Clifton-Brown, J.C., Scurlock, J.M.O., Huisman, W. (2000). *Miscanthus*: European experience with a novel energy crop, Biomass Bioenergy. 19, 209-227.

Li, F.-., Pryer, K.M. (2014). Crowdfunding the *Azolla* fern genome project: A grassroots approach, GigaScience. 3,.

Li, L., Olsen, K.M. (2016). To Have and to Hold: Selection for Seed and Fruit Retention During Crop Domestication, Curr. Top. Dev. Biol. 119, 63-109.

Liaimer, A., Helfrich, E.J., Hinrichs, K., Guljamow, A., Ishida, K., Hertweck, C., Dittmann, E. (2015). Nostopeptolide plays a governing role during cellular differentiation of the symbiotic cyanobacterium *Nostoc* punctiforme. Proceedings of the National Academy of Sciences 112: 1862-7.

Liu, J., Xu, H., Jiang, Y., Zhang, K., Hu, Y., Zeng, Z. (2017). Methane emissions and microbial communities as influenced by dual cropping of *Azolla* along with early Rice, Nature Scientific Reports. 7, art. no. 40635.

Liu, S., Ju, J., Xia, G. (2014). Identification of the flavonoid 3'-hydroxylase and flavonoid 3',5'-hydroxylase genes from Antarctic moss and their regulation during abiotic stress, Gene. 543, 145-152.

Logan, K.J., Thomas, B.A. (1985). Distribution of lignin derivatives in plants, New Phytol. 99, 571-585.

Loganes, C., Ballali, S., Minto, C. (2016). Main properties of canola oil components: A descriptive review of current knowledge, Open Agric. J. 10, 69-74.

Lohse, M., Nagel, A., Herter, T., May, P., Schroda, M., Zrenner, R., Tohge, T., Fernie, A.R., Stitt, M., Usadel, B. (2014). Mercator: A fast and simple web server for genome scale functional annotation of plant sequence data. Plant, Cell and Environment 37: 1250–1258.

Longstaff, M., Afrc, J.M.M. (1991). The inhibitory effects of hull polysaccharides and tannins of field beans (*Vicia faba* L.) on the digestion of amino acids, starch and lipid and on digestive enzyme activities in young chicks, Br. J. Nutr. 65, 199-216.

Lopez-Sanchez, P., de Vos, R.C.H., Jonker, H.H., Mumm, R., Hall, R.D., Bialek, L., Leenman, R., Strassburg, K., Vreeken, R., Hankemeier, T., Schumm, S., van Duynhoven, J. (2015). Comprehensive metabolomics to evaluate the impact of industrial processing on the phytochemical composition of vegetable purees, Food Chemistry. 168, 348-355.

Lorenz, K., Preston, C.M. (2002). Characterization of high-tannin fractions from humus by carbon-13 cross-polarization and magic-angle spinning nuclear magnetic resonance, J. Environ. Qual. 31, 431-436.

Lotfy, S., Negrel, J., Javelle, F. (1994). Formation of ω-feruloyloxypalmitic acid by an enzyme from wound-healing potato tuber discs, Phytochemistry. 35, 1419-1424.

Lozano, V., Brundu, G. (2016). Prioritisation of aquatic invasive alien plants in South America with the US Aquatic Weed Risk Assessment, Hydrobiologia. 1-16.

Lucas, R.C., Duckett, J.G. (1980). A cytological study of the male and female sporocarps of the heterosporous fern *Azolla filiculoides* Lam. New Phytologist 85: 409–418.

Lucas, W.J., Groover, A., Lichtenberger, R., Furuta, K., Yadav, S.R., Helariutta, Y., He, X-Q., Fukuda, H., Kang, J., Brady, S.M., Patrick, J.W., Sperry, J., Yoshida, A., Lopez-Millan, A-F., Grusak, M.A., Kachroo, P. (2013). The plant vascular system: evolution, development and Functions. Journal of Integrative Plant Biology 55: 294-388.

Lukovenkov, A.V., Belokurov, V.V., Varfolomeev, S.D. (2011). The stability and limits of human population growth, Moscow Univ. Chem. Bull. 66, 318-321.

Lumpkin, T.A., Plucknett, D.L. (1980). *Azolla*: Botany, physiology, and use as a green manure, Econ. Bot. 34, 111-153.

Luscher, A., Hartwig, U.A., Suter, D., Nosberger, J. (2000). Direct evidence that symbiotic N_2 fixation in fertile grassland is an important trait for a strong response of plants to elevated atmospheric CO_2, Global Change Biol. 6, 655-662.

Madhavan Nampoothiri, K., Nair, N.R., John, R.P. (2010). An overview of the recent developments in polylactide (PLA) research, Bioresour. Technol. 101, 8493-8501.

Maeda, S.I., Konishi, M., Yanagisawa, S., Omata, T. (2014). Nitrite transport activity of a novel HPP family protein conserved in cyanobacteria and chloroplasts. Plant Cell Physiology 55: 1311–1324

Maejima, K., Uheda, E., Kitoh, S., Shiomi, N. (2002). Differences in growth rate, nitrogen fixation and numbers of cyanobionts and heterocysts among three *Azolla pinnata* var. *pinnata* strains, Environ.

Exp. Bot. 47, 143-147.
Maejima, K., Kitoh, S., Uheda, E., Shiomi, N. (2001). Response of 19 *Azolla* strains to a high concentration of ammonium ions, Plant Soil. 234, 247-252.
Maggio, G., Cacciola, G. (2012). When will oil, natural gas, and coal peak? Fuel. 98, 111-123.
Magomya, A.M., Kubmarawa, D., Ndahi, J.A., Yebpella, G.G. (2014). Determination Of Plant Proteins Via The Kjeldahl Method And Amino Acid Analysis: A Comparative Study, International Journal of scientific & Technology Research. 3 (4) 68-72.
Mahlberg, P.G., 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, M.A., 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.
Makkar, H.P.S. (2003). Effects and fate of tannins in ruminant animals, adaptation to tannins, and strategies to overcome detrimental effects of feeding tannin-rich feeds, Small Ruminant Res. 49, 241-256.
Makkar, H.P.S., Gamble, G., Becker, K. (1999). Limitation of the butanol-hydrochloric acid-iron assay for bound condensed tannins, Food Chem. 66, 129-133.
Makkar, H.P.S., Blümmel, M., Becker, K. (1995). Formation of complexes between polyvinyl pyrrolidones or polyethylene glycols and tannins, and their implication in gas production and true digestibility in in vitro techniques, Br. J. Nutr. 73, 897-913.
Malencic, D., Popovic, M., Miladinovic, J. (2007).Phenolic content and antioxidant properties of soybean (Glycine max (L.)Merr.) seeds, Molecules. 12, 576-581.
Mandolesi de Araújo, C.D., de Andrade, C.C., de Souza, E., Silva, E., Dupas, F.A. (2013). Biodiesel production from used cooking oil: A review, Renewable Sustainable Energy Rev. 27, 445-452.
Martínez, R., Torres, P., Meneses, M.A., Figueroa, J.G., Pérez-Álvarez, J.A., Viuda-Martos, M. (2012). Chemical, technological and in vitro antioxidant properties of mango, guava, pineapple and passion fruit dietary fibre concentrate, Food Chem. 135, 1520-1526.
Mata, T.M., Martins, A.A., Caetano, N.S. (2010). Microalgae for biodiesel production and other applications: A review, Renewable Sustainable Energy Rev. 14, 217-232.
Mata, T.M., Mendes, A.M., Caetano, N.S., Martins, A.A. (2014). Sustainability and economic evaluation of microalgae grown in brewery wastewater, Bioresour. Technol. 168, 151-158.
McAfee, A.J., McSorley, E.M., Cuskelly, G.J., Moss, B.W., Wallace, J.M.W., Bonham, M.P., Fearon, A.M. (2010). Red meat consumption: An overview of the risks and benefits, Meat Sci. 84, 1-13.
Mcallister, T.A., Martinez, T., Hee, D.B., Muir, A.D., Yanke, L.J., Jones, G.A. (2005). Characterization of condensed tannins purified from legume forages: Chromophore production, protein precipitation, and inhibitory effects on cellulose digestion, J. Chem. Ecol. 31, 2049-2068.
McCann, M.J. (2016). Evidence of alternative states in freshwater lakes: A spatially-explicit model of submerged and floating plants, Ecol. Model. 337, 298-309.
McDonald, T.R., Ward J.M. (2016). Evolution of electrogenic ammonium transporters (AMTs). Frontiers in plant science 7: 352.
McGlade, C., Ekins, P. (2015). The geographical distribution of fossil fuels unused when limiting global warming to 2°C, Nature. 517, 187-190.
McNabb, W.C., Peters, J.S., Foo, L.Y., Waghorn, G.C., Jackson, F.S. (1998). Effect of condensed tannins prepared from several forages on the in vitro precipitation of ribulose-1,5-bisphosphate carboxylase (Rubisco) protein and its digestion by trypsin (EC 2.4.21.4) and chymotrypsin (EC 2.4.21.1), J. Sci. Food Agric. 77, 201-212.
Meeks, J.C., Steinberg, N.A., Enderlin, C.S., Joseph, C.M., Peters, G.A. (1987). *Azolla-Anabaena* Relationship : XIII. Fixation of [13N]N_2, Plant Physiology. 84, 883-886.
Meeks, J.C., Steinberg, N., Joseph, C.M., Enderlin, C.S., Jorgensen, P.A., Peters, G.A. (1985). Assimilation of exogenous and dinitrogen-derived $^{13}NH_4^+$ by *Anabaena azollae* separated from *Azolla caroliniana* Willd, Arch. Microbiol. 142, 229-233.
Metzgar, J.S., Schneider, H., Pryer, K.M. (2007). Phylogeny and divergence time estimates for the fern genus *Azolla* (Salviniaceae), Int. J. Plant Sci. 168, 1045-1053.
Meyer, R.S., DuVal, A.E., Jensen, H.R. (2012). Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. New Phytologist 196: 29–48.

Mian, M.H., Stewart, W.D.P. (1985). A^{15}N tracer study to compare nitrogen supply by *Azolla* and ammonium sulphate to IR8 rice plants grown under flooded conditions, Plant Soil. 83, 371-379.

Millward, D.J., Jackson, A.A. (2004). Protein/energy ratios of current diets in developed and developing countries compared with a safe protein/energy ratio: Implications for recommended protein and amino acid intakes, Public Health Nutr. 7, 387-405.

Mir, R.R., Zaman-Allah, M., Sreenivasulu, N., Trethowan, R., Varshney, R.K. (2012). Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops, Theor. Appl. Genet. 125, 625-645.

Miranda, A.F., Biswas, B., Ramkumar, N., Singh, R., Kumar, J., James, A., Roddick, F., Lal, B., Subudhi, S., Bhaskar, T., Mouradov, A. (2016). Aquatic plant *Azolla* as the universal feedstock for biofuel production, Biotechnol. Biofuels. 9,.

Mishra, P., Dash, D. (2014). Rejuvenation of Biofertilizer for Sustainable Agriculture and Economic Development, Consilience: The Journal of Sustainable Development. 11, 41-61.

Mitchell, D. (2008). A note on rising food prices, Policy Research Working Paper No. 4682. World Bank, Washington, DC.

Moco, S., Capanoglu, E., Tikunov, Y., Bino, R.J., Boyacioglu, D., Hall, R.D., Vervoort, J., de Vos, R.C.H. (2007). Tissue specialization at the metabolite level is perceived during the development of tomato fruit, J. Exp. Bot. 58, 4131-4146.

Moco, S., Bino, R.J., Vorst, O., Verhoeven, H.A., de Groot, J., van Beek, T.A., Vervoort, J., Ric de Vos, C.H. (2006). A liquid chromatography-mass spectrometry-based metabolome database for tomato, Plant Physiol. 141, 1205-1218.

Mohedano, R.A., Costa, R.H.R., Tavares, F.A., Belli Filho, P. (2012). High nutrient removal rate from swine wastes and protein biomass production by full-scale duckweed ponds, Bioresour.Technol. 112, 98-104.

Moglia, A., Lanteri, S., Comino, C., Hill, L., Knevitt, D., Cagliero, C., Rubiolo, P., Bornemann, S., Martin, C. (2014). Dual catalytic activity of hydroxycinnamoyl-coenzyme a quinate transferase from tomato allows it to moonlight in the synthesis of both mono- and dicaffeoylquinic acids, Plant Physiol. 166, 1777-1787.

Mole, S., Butler, L.G., Iason, G. (1990). Defense against dietary tannin in herbivores: A survey for proline rich salivary proteins in mammals, Biochem. Syst. Ecol. 18, 287-293.

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

Moser, B.R., Vaughn, S.F. (2012). Efficacy of fatty acid profile as a tool for screening feedstocks for biodiesel production. Biomass Bioenergy 37:31–41.

Muradov, N., Taha, M., Miranda, A.F., Kadali, K., Gujar, A., Rochfort, S., Stevenson, T., Ball, A.S., Mouradov, A. (2014). Dual application of duckweed and azolla plants for wastewater treatment and renewable fuels and petrochemicals production, Biotechnology for Biofuels. 7.

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.

Mutlu, H., Meier, M.A.R. (2010). Castor oil as a renewable resource for the chemical industry. Eur J Lipid Sci Technol 112:10–30.

Naghshi, H., Khojasteh, S., Jafari, M. (2014). Investigation of the effect of different levels of *Azolla* (*Azolla pinnata*) on performance and characteristics of Cobb Broiler Chicks, International Journal of farming and allied sciences. 3, 45-49.

Nakamichi, N. (2011). Molecular mechanisms underlying the *Arabidopsis* circadian clock. Plant Cell Physiology 52: 1709–18.

Nakashima, J., Chen, F., Jackson, L., Shadle, G., Dixon, R.A. (2008). Multi-site genetic modification of monolignol biosynthesis in alfalfa (*Medicago sativa*): Effects on lignin composition in specific cell types, New Phytol. 179, 738-750.

Nakos, M., Pepelanova, I., Beutel, S., Krings, U., Berger, R.G., Scheper, T. (2017). Isolation and analysis of vitamin B12 from plant samples, Food Chem. 216, 301-308.

Nambudiri, E.M.V., Chitaley, S. (1991). Fossil *Salvinia* and *Azolla* from the Deccan Intertrappean beds

of India, Rev. Palaeobot. Palynol. 69, 325-336.
Nduagu, E.I., Gates, I.D. (2015). Unconventional Heavy Oil Growth and Global Greenhouse Gas Emissions, Environ. Sci. Technol. 49, 8824-8832.
Nelson, D.R. (2006). Plant cytochrome P450s from moss to poplar, Phytochem. Rev. 5, 193-204.
Ness, R.W., Graham, S.W., Barrett, S.C.H. (2011). Reconciling gene and genome duplication events: Using multiple nuclear gene families to infer the phylogeny of the aquatic plant family Pontederiaceae, Mol. Biol. Evol. 28, 3009-3018.
Neuhäuser B., Dynowski, M., Ludewig, U. (2009). Channel-like NH_3 flux by ammonium transporter AtAMT2. FEBS Letters 583: 2833–2838
Neveux, N., Magnusson, M., Maschmeyer, T., de Nys, R., Paul, N.A. (2015). Comparing the potential production and value of high-energy liquid fuels and protein from marine and freshwater macroalgae, GCB Bioenergy. 7, 673-689.
Nicoletto, C., Tosini, F., Sambo, P. (2013).Effect of grafting and ripening conditions on some qualitative traits of 'Cuore di bue' tomato fruits, J. Sci. Food Agric. 93, 1397-1403.
Nierop, K.G.J., Speelman, E.N., de Leeuw, J.W., Reichart, G-J. (2011). The omnipresent water fern *Azolla caroliniana* does not contain lignin, Org. Geochem. 42, 846-850.
Nierop, K.G.J., Verstraten, J.M. (2006). Fate of tannins in Corsican pine litter, J. Chem. Ecol. 32, 2709-2719.
Nierop, K.G.J., Preston, C.M., Kaal, J. (2005). Thermally assisted hydrolysis and methylation of purified tannins from plants, Anal. Chem. 77, 5604-5614.
Noordally, Z.B., Millar, A.J. (2014). Clocks in algae. Biochemistry 54: 171-183
NRC (1994).Nutrient requirements of poultry, 9th ed. National Academy Press, Washington, D.C.
NRC (1998).Nutrient requirements of swine, 10th ed. National Academy Press, Washington, D.C.
NRC (2001).Nutrient requirements of dairy cattle, 7th ed. National Academy Press, Washington, D.C.
Oh, H.I., Hoff, J.E., Armstrong, G.S., Haff, L.A. (1980). Hydrophobic interaction in tannin-protein complexes, J. Agric. Food Chem. 28, 394-398.
Ohmori, M., Hattori, A. (1974). Effect of ammonia on nitrogen fixation by the blue-green alga *Anabaena cylindrica*. Plant Cell Physiology 15: 131–142.
Okoronkwo, N., van Hove, C., Eskew, D.L. (1989). Evaluation of nitrogen-fixation by different strains of the *Azolla-Anabaena* symbiosis in the presence of a high-level of ammonium. Biology and Fertility of Soils 7: 275–278.
Oladosu, Y., Rafii, M.Y., Abdullah, N., Hussin, G., Ramli, A., Rahim, H.A., Miah, G., Usman, M. (2016). Principle and application of plant mutagenesis in crop improvement: A review, Biotechnol. Biotechnol. Equip. 30, 1-16.
Oldroyd, G.E.D., Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. Current Opinion in Biotechnology 26: 19–24.
Olguín, E.J., Rodríguez, D., Sánchez, G., Hernández, E., Ramírez, M.E. (2003). Productivity, protein content and nutrient removal from anaerobic effluents of coffee wastewater in *Salvinia minima* ponds, under subtropical conditions, Acta Biotechnol. 23, 259-270.
Opande, G.O., Onyango, J.C., Wagai, S.O. (2004). Lake Victoria: The water hyacinth (*Eichhornia crassipes* [MART.] SOLMS), its socio-economic effects, control measures and resurgence in the Winam gulf, Limnologica. 34, 105-109.
Ozdal, T., Capanoglu, E., Altay, F. (2013). A review on protein-phenolic interactions and associated changes, Food Res. Int. 51, 954-970.
Ozengin, N., Elmaci, A. (2007). Performance of Duckweed (*Lemna minor* L.) on different types of wastewater treatment, J. Environ. Biol. 28, 307-314.
Padda, M.S., Picha, D.H. (2008). Quantification of phenolic acids and antioxidant activity in sweetpotato genotypes, Sci. Hortic. 119, 17-20.
Palmer, B., Jones, R.J., Wina, E., Tangendjaja, B. (2000). The effect of sample drying conditions on estimates of condensed tannin and fibre content, dry matter digestibility, nitrogen digestibility and PEG binding of Calliandra calothyrsus, Anim. Feed Sci. Technol. 87, 29-40.
Pankievicz, V.C.S., do Amaral, F.P., Santos, K.F.D.N., Agtuca, B., Xu, Y., Schueller, M.J., Arisi, A.C.M., Steffens, M.B.R., de Souza, E.M., Pedrosa, F.O., Stacey, G., Ferrier, R.A. (2015). Robust

biological nitrogen fixation in a model grass-bacterial association. Plant Journal for Cell and Molecular Biology 81: 907–19.

Paoletti, C., Bocci, F., Lercker, G., Capella, P., Materassi, R. (1987). Lipid composition of *Azolla caroliniana* biomass and its seasonal variation, Phytochemistry. 26, 1045-1047.

Parejo, I., Jauregui, O., Sánchez-Rabaneda, F., Viladomat, F., Bastida, J., Codina, C. (2004). Separation and characterization of phenolic compounds in fennel (*Foeniculum vulgare*) using liquid chromatography-negative electrospray ionization tandem mass spectrometry, J. Agric. Food Chem. 52, 3679-3687.

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

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

Peng, Q.-., Zhu, Y., Liu, Z., Du, C., Li, K.-., Xie, D.-. (2012). An integrated approach to demonstrating the ANR pathway of proanthocyanidin biosynthesis in plants, Planta. 236, 901-918.

Pereira, A.L., Martins, M., Oliveira, M.M., Carrapiço, F. (2011). Morphological and genetic diversity of the family *Azolla*ceae inferred from vegetative characters and RAPD markers, Plant Syst. Evol. 297, 213-226.

Pereira, A.L., Carrapiço, F. (2007). Histochemistry of simple hairs from the foliar cavities of *Azolla filiculoides*, Plant Biosyst. 141, 323-328.

Pérez A., Casas A., Fernández C.M., Ramos M.J., Rodríguez L. (2010). Winterization of peanut biodiesel to improve the cold flow proper- ties. Bioresour Technol 101:7375–81.

Peters, G.A., Meeks, J.C. (1989). The *Azolla-Anabaena* Symbiosis: Basic Biology, Annu. Rev. Plant Physiol. Plant Mol. Biol. 40, 193-210.

Peters, G.A., Toia, Jr. R.E., Evans, W.R., Crist, D.K., Mayne, B.C., Poole, R.E. (1980). Characterization and comparison of five N_2-fixing *Azolla-Anabaena* associations, I. Optimization of growth conditions for biomass increase and N content in a controlled environment. Plant, Cell and Environment 3: 261–269.

Peters, G.A., Perkins, S.K. (1993). The *Azolla-Anabaena* symbiosis: endophyte continuity in the *Azolla* life-cycle is facilitated by epidermal trichomes. New Phytologist 123: 65–75.'

Peterson, R.B., Dolan, E., Calvert, H.E., Ke, B. (1981). Energy transfer from phycobiliproteins to Photosystem I in vegetative cells and heterocysts of *Anabaena Variabilis*, Biochim. Biophys. Acta Bioenerg. 634, 237-248.

Petkov, G., Ivanova, A., Iliev, I., Vaseva, I. (2012). A critical look at the microalgae biodiesel. Eur J Lipid Sci Technol 114:103–11.

Petzke, K.J., Ezeagu, I.E., Proll, J., Akinsoyinu, A.O., Metges, C.C. (1997). Amino acid composition, available lysine content and in vitro protein digestibility of selected tropical crop seeds, Plant Foods Hum. Nutr. 50, 151-162.

Pieterse, A.H., Murphy, K.J. (1993). Aquatic Weeds: The Ecology and Management of Nuisance Aquatic Vegetation, Second edition ed. Oxford University Press.

Pizzi, A. (2006). Recent developments in eco-efficient bio-based adhesives for wood bonding: Opportunities and issues, J. Adhes. Sci. Technol. 20, 829-846.

Plackett, A.R.G., Huang, L., Sanders, H.L., Langdale, J.A. (2014). High-efficiency stable transformation of the model fern species *Ceratopteris richardii* via microparticle bombardment, Plant Physiol. 165, 3-14.

Polikovsky, M., Fernand, F., Sack, M., Frey, W., Müller, G., Golberg, A. (2016). Towards marine biorefineries: Selective proteins extractions from marine macroalgae *Ulva* with pulsed electric fields, Innovative Food Sci. Emerg. Technol. 37, 194-200.

Pollastri, S., Tattini, M. (2011). Flavonols: Old compounds for old roles, Ann. Bot. 108, 1225-1233.

Poorter, H., van Berkel, Y., Baxter, R., den Hertog, J., Dijkstra, P., Gifford, R.M. et al. (1997). The effect of elevated CO_2 on the chemical composition and construction costs of leaves of 27 C_3 species. Plant, Cell Environ 20:472–82.

Prodpran, T., Benjakul, S., Phatcharat, S. (2012). Effect of phenolic compounds on protein cross-linking and properties of film from fish myofibrillar protein, Int. J. Biol. Macromol. 51, 774-782.

Pugesgaard, S., Schelde, K., Larsen, S.U., Lærke, P.E., Jørgensen, U. (2015). Comparing annual and perennial crops for bioenergy production - influence on nitrate leaching and energy balance, GCB Bioenergy. 7, 1136-1149.

Purugganan, M.D., Fuller, D.Q. (2009). The nature of selection during plant domestication, Nature. 457, 843-848.

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 Azolla Use Fuzhou, Fuijan, China. International Rice Research Institute, Los Baños, Philippines: 33–38.

Rai, A.N., Soderback, E., Bergman, B. (2000). Cyanobacterium-plant symbioses. New Phytologist 147: 449–481.

Raj, T., Kapoor, M., Gaur, R., Christopher, J., Lamba, B., Tuli, D.K., Kumar, R. (2015). Physical and chemical characterization of various indian agriculture residues for biofuels production, Energy Fuels. 29, 3111-3118.

Rajcan, I., Tollenaar, M. (1999). Source: Sink ratio and leaf senescence in maize: I. Dry matter accumulation and partitioning during grain filling, Field Crops Res. 60, 245-253.

Ramos, M.J., Fernández, C.M., Casas, A., Rodríguez, L., Pérez, A. (2009). Influence of fatty acid composition of raw materials on biodiesel properties. Bioresour Technol 100:261–8

Ran, L., Larsson, J., Vigil-Stenman, T., Nylander, J.A.A., Ininbergs, K., Zheng, W.-, Lapidus, A., Lowry, S., Haselkorn, R., Bergman, B. (2010). Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium, PLoS ONE. 5.

Ravindran, V., Abdollahi, M.R., Bootwalla, S.M. (2014).Nutrient analysis, metabolizable energy, and digestible amino acids of soybean meals of different origins for broilers, Poult.Sci. 93, 2567-2577.

Ray, T.B., Berger, C., Mayne, R.E., Toia, J., and Peters, G.A. (1979). *Azolla-Anabaena* Relationship: VIII. Photosynthetic Characterization of the Association and Individual Partners , Plant Physiology. 64, 791-795.

Reddy, K.R., D'Angelo, E.M. (1990). Biomass yield and nutrient removal by water hyacinth (*Eichhornia crassipes*) as influenced by harvesting frequency, Biomass. 21, 27-42.

Reed, S.C., Yang, X., Thornton, P.E. (2015). Incorporating phosphorus cycling into global modeling efforts: A worthwhile, tractable endeavor, New Phytol. 208, 324-329.

Rencoret, J., del Río, J.C., Nierop, K.G.J., Gutiérrez, A., Ralph, J. (2016). Rapid Py-GC/MS assessment of the structural alterations of lignins in genetically modified plants, J. Anal. Appl. Pyrolysis. 121, 155-164.

Rhodes, C.J. (2013). Peak phosphorus - Peak food? The need to close the phosphorus cycle, Sci. Prog. 96, 109-152.

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

Rigbi, M., Rosinski, J., Siegelman, H.W., Sutherland, J.C. (1980). Cyanobacterial phycobilisomes: selective dissociation monitored by fluorescence and circular dichroism. Proceedings of the National Academy of Sciences, USA 77: 1961– 1965.

Rinaldi, R., Jastrzebski, R., Clough, M.T., Ralph, J., Kennema, M., Bruijnincx, P.C.A., Weckhuysen, B.M. (2016). Paving the Way for Lignin Valorisation: Recent Advances in Bioengineering, Biorefining and Catalysis, Angew. Chem. Int. Ed. 55, 8164-8215.

Rivero, R.M., Ruiz, J.M., García, P.C., López-Lefebre, L.R., Sánchez, E., Romero, L. (2001). Resistance to cold and heat stress: Accumulation of phenolic compounds in tomato and watermelon plants, Plant Sci. 160, 315-321.

Robinson, M.D., McCarthy, D.J., Smyth, G.K. (2010). EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–40

Robinson Jr., W.E., Reinecke, M.G., Abdel-Malek, S., Jia, Q., Chow, S.A. (1996). Inhibitors of HIV^{-1} replication that inhibit HIV integrase, Proc. Natl. Acad. Sci. U. S. A. 93, 6326-6331.

Roddy, D.J. (2013). Biomass in a petrochemical world, Interface Focus. 3.

Rogge, G.D., Viana, A.M., Randi, A.M. (2000). Cryopreservation of spores of Dicksonia sellowiana: an

endangered tree fern indigenous to South and Central America. Cryo-Letters 21: 223–230.

Rohrmann, S., Overvad, K., Bueno-de-Mesquita, H.B., Jakobsen, M.U., Egeberg, R., Tjønneland, A., Nailler, L., Boutron-Ruault, M.-., Clavel-Chapelon, F., Krogh, V., Palli, D., Panico, S., Tumino, R., Ricceri, F., Bergmann, M.M., Boeing, H., Li, K., Kaaks, R., Khaw, K.-., Wareham, N.J., Crowe, F.L., Key, T.J., Naska, A., Trichopoulou, A., Trichopoulos, D., Leenders, M., Peeters, P.H.M., Engeset, D., Parr, C.L., Skeie, G., Jakszyn, P., Sánchez, M.-., Huerta, J.M., Redondo, M.L., Barricarte, A., Amiano, P., Drake, I., Sonestedt, E., Hallmans, G., Johansson, I., Fedirko, V., Romieux, I., Ferrari, P., Norat, T., Vergnaud, A.C., Riboli, E., Linseisen, A.J. (2013). Meat consumption and mortality - results from the European Prospective Investigation into Cancer and Nutrition, BMC Med. 11.

Rosatella, A.A., Simeonov, S.P., Frade, R.F.M., Afonso, C.A.M. (2011). 5-Hydroxymethylfurfural (HMF) as a building block platform: Biological properties, synthesis and synthetic applications, Green Chem. 13, 754-793.

Rotunno, P., Lanzini, A., Leone, P. (2017). Energy and economic analysis of a water scrubbing based biogas upgrading process for biomethane injection into the gas grid or use as transportation fuel, Renew. Energy. 102, 417-432.

Rozen, A., Arad, H., Schönfeld, M., Tel-Or, E. (1986). Fructose supports glycogen accumulation, heterocysts differentiation, N_2 fixation and growth of the isolated cyanobiont *Anabaena azollae*, Arch. Microbiol. 145, 187-190.

Rusoff, L.L., Blakeney Jr., E.W., Culley Jr., D.D. (1980). Duckweeds (Lemnaceae family): A potential source of protein and amino acids, J. Agric. Food Chem. 28, 848-850.

Sabaté, J., Soret, S. (2014). Sustainability of plant-based diets: Back to the future, Am. J. Clin. Nutr. 100.

Sah, R.N., Goyal, S.S., Rains, D.W. (1989). Interactive effects of exogenous combined nitrogen and phosphorus on growth and nitrogen fixation by *Azolla*, Plant Soil. 117, 1-8.

Saito, K., Yonekura-Sakakibara, K., Nakabayashi, R., Higashi, Y., Yamazaki, M., Tohge, T., Fernie, A.R. (2013). The flavonoid biosynthetic pathway in *Arabidopsis*: Structural and genetic diversity, Plant Physiol. Biochem. 72, 21-34.

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

Salgotra, R.K., Gupta, B.B., Stewart Jr., C.N. (2014). From genomics to functional markers in the era of next-generation sequencing, Biotechnol. Lett. 36, 417-426.

Salmi, M.L., Roux, S.J. (2008). Gene expression changes induces by spaceflight in single-cells of the fern *Ceratopteris* richardii. Planta 229: 151–159.

Saminathan, M., Tan, H.Y., Sieo, C.C., Abdullah, N., Wong, C.M.V.L., Abdulmalek, E., Ho, Y.W. (2014). Polymerization degrees, molecular weights and protein-binding affinities of condensed tannin fractions from a *Leucaena leucocephala* hybrid, Molecules. 19, 7990-8010.

Sánchez-Viveros, G., Ferrera-Cerrato, R., Alarcón, A. (2011). Short-term effects of arsenate-induced toxicity on growth, chlorophyll and carotenoid contents, and total content of phenolic compounds of *Azolla filiculoides*, Water Air Soil Pollut. 217, 455-462.

Sands, D.P.A., Kassulke, R.C. (1986). Assessment of *Paulinia acuminata* [Orthoptera: Acrididae] for the biological control of *Salvinia molesta* in Australia, Entomophaga. 31, 11-17.

Sanginga, N., van Hove, C. (1989). Amino acid composition of azolla as affected by strains and population density, Plant Soil. 117, 263-267.

Santiago, C.B., Lovell, R.T. (1988). Amino acid requirements for growth of Nile tilapia, J. Nutr. 118, 1540-1546.

dos Santos, P.C., Fang, Z., Mason, S.W., Setubal, J.C., Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. BMC Genomics 13: 162

Sari, Y.W., Mulder, W.J., Sanders, J.P.M., Bruins, M.E. (2015). Towards plant protein refinery: Review on protein extraction using alkali and potential enzymatic assistance, Biotechnol. J. 10, 1138-1157.

Sari, Y.W., Bruins, M.E., Sanders, J.P.M. (2013). Enzyme assisted protein extraction from rapeseed, soybean, and microalgae meals, Ind. Crops Prod. 43, 78-83.

Sarkar, N., Ghosh, S.K., Bannerjee, S., Aikat, K. (2012). Bioethanol production from agricultural wastes:

An overview, Renew. Energy. 37, 19-27.

Scheben, A., Yuan, Y., Edwards, D. (2016). Advances in genomics for adapting crops to climate change, Curr. Plant Biol. 6, 2-10.

Scheible, W-R., Morcuende, R., Czechowski, T., Fritz, C., Osuna, D., Palacios-Rojas, N., Schindelasch, D., Thimm, O., Udvardi, M.K., Stitt, M. (2004). Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen. Plant Physiology 136: 2483–2499

Schiml, S., Fauser, F., Puchta, H. (2014). The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in *Arabidopsis* resulting in heritable progeny, Plant J. 80, 1139-1150.

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

Scott, E., Peter, F., Sanders, J. (2007).Biomass in the manufacture of industrial products-the use of proteins and amino acids, Appl. Microbiol.Biotechnol. 75, 751-762.

Sessa, E.B., Banks, J.A., Barker, M.S., Der, J.P., Duffy, A.M., Graham, S.W., Hasebe, M., Langdale, J., Li, F.-., Marchant, D.B., Pryer, K.M., Rothfels, C.J., Roux, S.J., Salmi, M.L., Sigel, E.M., Soltis, D.E., Soltis, P.S., Stevenson, D.W., Wolf, P.G. (2014). Between Two Fern Genomes, GigaScience. 3.

Sheldon, R.A. (2014). Green and sustainable manufacture of chemicals from biomass: State of the art, Green Chem. 16, 950-963.

Shepherd, J., Bachis, E. (2014).Changing supply and demand for fish oil, Aquac.Econ. Manage. 18, 395-416.

Shi, D.J., Hall, D.O. (1988). The *Azolla-Anabaena* association: Historical perspective, symbiosis and energy metabolism. The Botanical Review 54: 353–386

Shimada, K., Kitamura, N., Yoshida, Y., Matsushima, H., Asada, Y. (2010). Effects of water soluble extract from the dead leaves of deciduous trees on the growth control of toxic cyanobacteria like Microcystis aeruginosa, WIT Trans. Ecol. Environ. 135, 71-82.

Shimada, T. (2006). Salivary proteins as a defense against dietary tannins, J. Chem. Ecol. 32, 1149-1163.

Shilton, A.N., Powell, N., Guieysse, B. (2012). Plant based phosphorus recovery from wastewater via algae and macrophytes. Current Opinion in Biotechnology 23: 884–9

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 *Azolla* Use Fuzhou, Fuijan, China. Los Ban~os, Philippines: International Rice Research Institute, 27–32.

Sini, S., Smitha, S.S., Madhusoodanan, P.V. (2014). Induction of sporocarp development in vitro in the mosquito fern, *Azolla rubra* R. Br. Annals of Plant Sciences. 4, 994-1002.

Skrabanja, V., Lærke H.N., Kreft I. (2000).Protein-polyphenol interactions and in vivo digestibility of buckwheat groat proteins, Pflugers Archiv Eur. J. Physiol. 440, R129-R131.

da Silva, G.P., Mack, M., Contiero, J. (2009). Glycerol: A promising and abundant carbon source for industrial microbiology, Biotechnol. Adv. 27, 30-39.

Silva, T., Oliveira, C., Borges, F. (2014). Caffeic acid derivatives, analogs and applications: A patent review (2009-2013), Expert Opin. Ther. Pat. 24, 1257-1270.

Sims, R.E.H., Mabee, W., Saddler, J.N., Taylor, M. (2010). An overview of second generation biofuel technologies, Bioresour. Technol. 101, 1570-1580.

Smil, V. (2000). Phosphorus in the environment: Natural flows and human interferences, Annu. Rev. Energy Environ. 25, 53-88.

Smith, K.A., McTaggart, I.P., 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.

Solomon, B.D., Barnes, J.R., Halvorsen, K.E. (2007). Grain and cellulosic ethanol: History, economics, and energy policy, Biomass Bioenergy. 31, 416-425.

Sommer, M., Brautigam, A., Weber, A.P. (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.

Sorrell, S., Speirs, J., Bentley, R., Brandt, A., Miller, R. (2010). Global oil depletion: A review of the evidence, Energy Policy. 38, 5290-5295.

Soussana, J.F., Hartwig, U.A. (1996). The effects of elevated CO_2 on symbiotic N_2 fixation: A link between the carbon and nitrogen cycles in grassland ecosystems, Plant Soil. 187, 321-332.

Speedy, A.W. (2003). Global Production and Consumption of Animal Source Foods, J. Nutr. 133.

Speelman, E.N., van Kempen, M.M.L., Barke, J., Brinkhuis, H., Reichart, G.J., Smolders, A.J.P., Roelofs, J.G.M., Sangiorgi, F., De Leeuw, J.W., Lotter, A.F., Sinninghe Damsté, J.S. (2009a. The Eocene Arctic *Azolla* bloom: Environmental conditions, productivity and carbon drawdown, Geobiology. 7, 155-170.

Speelman, E.N., Reichart, G-J., de Leeuw, J.W., Rijpstra, W.I.C., Sinninghe Damsté, J.S. (2009b. Biomarker lipids of the freshwater fern *Azolla* and its fossil counterpart from the Eocene Arctic Ocean, Org. Geochem. 40, 628-637.

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

Stafford, H.A. (1991). Flavonoid evolution: An enzymic approach, Plant Physiol. 96, 680-685.

Stamp, P., Messmer, R., Walter, A. (2012). Competitive underutilized crops will depend on the state funding of breeding programmes: an opinion on the example of Europe. Plant Breeding 131: 461–464

Stephenson, J., Newman, K., Mayhew, S. (2010). Population dynamics and climate change: What are the links? J. Public Health (United Kingdom). 32, 150-156.

Summers, M.L., Wallis, J.G., Campbell, E.L., Meeks, J.C. (1995). Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium *Nostoc* sp. strain ATCC 29133. Journal of Bacteriology 177: 6184–6194

Suutari, M., Leskinen, E., Fagerstedt, K., Kuparinen, J., Kuuppo, P., Blomster, J. (2015). Macroalgae in biofuel production, Phycol. Res. 63, 1-18.

Takashima, T., Hikosaka, K., Hirose, T. (2004). Photosynthesis or persistence: Nitrogen allocation in leaves of evergreen and deciduous *Quercus* species, Plant Cell Environ. 27, 1047-1054.

Teixeira, G., Carrapiço, F., Gomes, E.T. (2001). C-glycosylflavones in the genus *Azolla*, Plant Biosyst. 135, 233-237.

Terrill, T.H., Rowan, A.M., Douglas, G.B., Barry, T.N. (1992). Determination of extractable and bound condensed tannin concentrations in forage plants, protein concentrate meals and cereal grains, J. Sci. Food Agric. 58, 321-329.

van Thielen, R., Ajuonu, O., Schade, V., Neuenschwander, P., Adité, A., Lomer, C.J. (1994). Importation, releases, and establishment of Neochetina spp. (Col.: Curculionidae) for the biological control of water hyacinth, *Eichhornia crassipes* (Lil.: Pontederiaceae), in Benin, West Africa, Entomophaga. 39, 179-188. Tilman, D. (1998). The greening of the green revolution, Nature. 396, 211-212.

Timilsina, G.R. (2012). Biofuels: The food versus fuel debate, CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour. 7,.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y., Stitt, M. (2004). MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant Journal 37: 914–939.

Tippery, N.P., Les, D.H., Crawford, D.J. (2015). Evaluation of phylogenetic relationships in Lemnaceae using nuclear ribosomal data, Plant Biol. 17, 50-58.

Toda, M., Takagaki, A., Okamura, M., Kondo, J.N., Hayashi, S., Domen, K. et al. (2005). Green chemistry: biodiesel made with sugar catalyst. Nature 438:178.

Tohge, T., Watanabe, M., Hoefgen, R., Fernie, A.R. (2013). The evolution of phenylpropanoid metabolism in the green lineage, Crit. Rev. Biochem. Mol. Biol. 48, 123-152.

Toia, R.E. Jr, Buzby, K.M., Peters, G.A. (1987). Sporocarps of *Azolla mexicana* Presl. 1. isolation and purification of megasporocarps and microsporangia. New Phytologist 106: 271–279.

Tong, J., Ma, B., Ge, L., Mo, Q., Zhou, G., He, J., Wang, Y. (2015). Dicaffeoylquinic Acid-Enriched Fraction of *Cichorium glandulosum* Seeds Attenuates Experimental Type 1 Diabetes via Multipathway Protection, J. Agric. Food Chem. 63, 10791-10802.

Truernit, E., Bauby, H., Dubreucq, B., Grandjean, O., Runions, J., Barthélémy, J., Palauqui, J.-. (2008). High-resolution whole-mount imaging of three-dimensional tissue organization and gene

expression enables the study of phloem development and structure in *Arabidopsis*, Plant Cell. 20, 1494-1503.

Tubeileh, A., Rennie, T.J., Goss, M.J. (2016). A review on biomass production from C_4 grasses: Yield and quality for end-use, Curr. Opin. Plant Biol. 31, 172-180.

Tucker, C.S., Debusk, T.A. (1981). Seasonal growth of *Eichhornia crassipes* (Mart.) solms: Relationship to protein, fiber, and available carbohydrate content, Aquatic Bot. 11, 137-141.

Udvardi, M., Poole, P.S. (2013). Transport and metabolism in legume-rhizobia symbioses. Annual Review of Plant Biology 64: 781–805.

Uheda, E., Nakamura, S. (2000). Abscission of *Azolla* branches induced by ethylene and sodium azide, Plant Cell Physiol. 41, 1365-1372.

United Nations (2015). World Population Prospects: The 2015 Revision, Key Findings and Advance Tables. Working Paper No. ESA/P/WP.241.

Usadel, B., Nagel, A., Steinhauser, D., Gibon, Y., Blasing, O.E., Redestig, H., Sreenivasulu, N., Krall, L., Hannah, M.A., 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.

van Uytvanck, P.P., Hallmark, B., Haire, G., Marshall, P.J., Dennis, J.S. (2014). Impact of biomass on industry: Using ethylene derived from bioethanol within the polyester value chain, ACS Sustainable Chem. Eng. 2, 1098-1105.

Vajda, V., McLoughlin, S. (2005). A new Maastrichtian-Paleocene *Azolla* species from of Bolivia, with a comparison of the global record of coeval *Azolla* microfossils, Alcheringa. 29, 305-329.

Valente, L.M.P., Gouveia, A., Rema, P., Matos, J., Gomes, E.F., Pinto, I.S. (2006). Evaluation of three seaweeds *Gracilaria bursa-pastoris*, *Ulva rigida* and *Gracilaria cornea* as dietary ingredients in European sea bass (Dicentrarchus labrax) juveniles, Aquaculture. 252, 85-91.

Varshney, R.K., Mohan, S.M., Gaur, P.M., Gangarao, N.V.P.R., Pandey, M.K., Bohra, A., Sawargaonkar, S.L., Chitikineni, A., Kimurto, P.K., Janila, P., Saxena, K.B., Fikre, A., Sharma, M., Rathore, A., Pratap, A., Tripathi, S., Datta, S., Chaturvedi, S.K., Mallikarjuna, N., Anuradha, G., Babbar, A., Choudhary, A.K., Mhase, M.B., Bharadwaj, C., Mannur, D.M., Harer, P.N., Guo, B., Liang, X., Nadarajan, N., Gowda, C.L.L. (2013). Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics, Biotechnol. Adv. 31, 1120-1134.

Veys, P., Waterkeyn, L., Lejeune, A., van Hove, C. (1999). The pore of the leaf cavity of *Azolla*: morphology, cytochemistry and possible functions. Symbiosis 27: 33-57.

Veys, P., Lejeune, A., van Hove, C. (2002). The pore of the leaf cavity of *Azolla* species: teat cell differentiation and cell wall projections. Protoplasma 219: 31-42.

Vidal, E.A., Moyano, T.C., Krouk, G., Katari, M.S., Tanurdzic, M., McCombie, W.R., Coruzzi, G.M., Gutiérrez R.A. (2013). Integrated RNA-seq and sRNA-seq analysis identifies novel nitrate-responsive genes in *Arabidopsis thaliana* roots. BMC Genomics 14: 701.

de Visser, C.L.M., Schreuder, R., Stoddard, F. (2014). The EU's dependency on soya bean import for the animal feed industry and potential for EU produced alternatives. OCL 21: D407.

de Vries, S.C., van de Ven, G.W.J., van Ittersum, M.K., Giller, K.E. (2010). Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion tech- niques. Biomass Bioenergy 34:588–601.

Wagner, G.M. (1997). *Azolla*: A Review of Its Biology and Utilization, Bot. Rev. 63, 1-26.

Wang, C., Houlton, B.Z., Dai, W., Bai, E. (2017). Growth in the global N_2 sink attributed to N fertilizer inputs over 1860 to 2000, Sci. Total Environ. 574, 1044-1053.

Wang, H., Wang, Q., Bowler, P.A., Xiong, W. (2016). Invasive aquatic plants in China, Aquatic Invasions. 11, 1-9.

Wang, Z., Li, X., Zhen, S., Li, X., Wang, C., Wang, Y. (2013). The important role of quinic acid in the formation of phenolic compounds from pyrolysis of chlorogenic acid, J Therm Anal Calor. 114, 1231-1238.

Wang, R., Okamoto, M., Xing, X., Crawford, N.M. (2003). Microarray analysis of the nitrate response

in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiology 132: 556–67

Wang, R.-., Stec, A., Hey, J., Lukens, L., Doebley, J. (1999). The limits of selection during maize domestication, Nature. 398, 236-239.

Wang, Z., Zhang, Z., Zhang, J., Zhang, Y., Liu, H., Yan, S. (2012). Large-scale utilization of water hyacinth for nutrient removal in Lake Dianchi in China: The effects on the water quality, macrozoobenthos and zooplankton, Chemosphere. 89, 1255-1261.

Wani, S.P., Garg, K.K., Chander, G. (2016). Water needs and productivity of *Jatropha curcas* in India: Myths and facts, Biofuel Bioprod. Biorefining. 10, 240-254.

Watanabe, F., Yabuta, Y., Bito, T., Teng, F. (2014). Vitamin B12-containing plant food sources for vegetarians, Nutr. 6, 1861-1873.

Watanabe, I., Roger, P.A., Ladha, J.K., van Hove, C. (1992). Biofertilizer germplasm collections at IRRI. International Rice Research Institute, Los Baños, Philippines.

Watanabe, I., Ventura, W., Mascariña, G., Eskew, D.L. (1989). Fate of *Azolla* spp. and urea nitrogen applied to wetland rice (*Oryza sativa* L.), Biol Fert Soils. 8, 102-110.

Watanabe, I. (1982). *Azolla-Anabaena* symbiosis – its physiology and use tropical agriculture, Dommerques, Y.R., Diem, H.G. (Eds.), Microbiology of Trpoical Soils and Plant Productivity. MartinusNijhoff/ Dr. W. Junk, The Hague, the Netherlands, pp. 169-185.

Waterman, P.G., Mole, S. (1994).Analysis of Phenolic Plant Metabolites.Blackwell Scientific Publications, Oxford.

Weng, J.-., Chapple, C. (2010). The origin and evolution of lignin biosynthesis, New Phytol. 187, 273-285.

van der Werf, A., Geerts, R.H.E.M., Jacobs, F.H.H., Korevaar, H., Oomes, M.J.M., de Visser, W. (1998). The importance of relative growth rate and associated traits for competition between species during vegetation succession. , H. Lambers,, H. Poorter& M.M.I. Van Vuuren (Ed.), Inherent Variation in Plant Growth. Physiological mechanisms and Ecological Consequences.Backhuys Publishers, Leiden, pp. 489-502., pp. 489-502.

Wiedmann, T.O., Schandl, H., Lenzen, M., Moran, D., Suh, S., West, J., Kanemoto, K. (2015). The material footprint of nations, Proc. Natl. Acad. Sci. U. S. A. 112, 6271-6276.

Wildschut, J., Smit, A.T., Reith, J.H., Huijgen, W.J.J. (2013). Ethanol-based organosolv fractionation of wheat straw for the production of lignin and enzymatically digestible cellulose, Bioresour. Technol. 135, 58-66.

Willemse, M.T.M. (2009). Evolution of plant reproduction: from fusion and dispersal to interaction and communication. Chinese Science Bulletin 54: 2390–2403.

Wilson, E.W., Rowntree, S.C., Suhre, J.J., Weidenbenner, N.H., Conley, S.P., Davis, V.M., Diers, B.W., Esker, P.D., Naeve, S.L., Specht, J.E., et al. (2014). Genetic Gain × Management Interactions in Soybean: II. Nitrogen Utilization. Crop Science 54: 340.

Woertz, I., Feffer, A., Lundquist, T., Nelson, Y. (2009). Algae grown on dairy and municipal wastewater for simultaneous nutrient removal and lipid production for biofuel feedstock, J. Environ. Eng. 135, 1115-1122.

Wolosiak, R., Druzynska, B., Piecyk, M., Worobiej, E., Majewska, E., Lewicki, P.P. (2011). Influence of industrial sterilisation, freezing and steam cooking on antioxidant properties of green peas and string beans, Int. J. Food Sci. Technol. 46, 93-100.

Wong, D.W.S. (2009). Structure and action mechanism of ligninolytic enzymes, Appl. Biochem. Biotechnol. 157, 174-209.

WWF (2014). The Growth of Soy: Impacts and Solutions, World Wildlife fund, Washington D.C., USA.

Xie, Y., Wen, M., Yu, D., Li, Y. (2004). Growth and resource allocation of water hyacinth as affected by gradually increasing nutrient concentrations, Aquatic Bot. 79, 257-266.

Xu, G., Fan, X., Miller, A.J. (2012). Plant Nitrogen Assimilation and Use Efficiency. Annual Review of Plant Biology 63: 153–182.

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

Yadav, S., Kumar, S., Hussain, Z., Suneja, P., Yadav, S.K., Nizar, M.A. et al. (2012). *Guizotia abyssinica*

(L.f.) cass.: an untapped oilseed re- source for the future. Biomass Bioenergy 43:72–78.

Yang, B., Wyman, C.E. (2008). Potential for low cost pretreatment of cellulosic biomass for biological production of fuels and chemicals, TAPPI Int. Bioenergy Bioproducts Conf.

Yang, J., Xu, M., Zhang, X., Hu, Q., Sommerfeld, M., Chen, Y. (2011). Life-cycle analysis on biodiesel production from microalgae: Water footprint and nutrients balance, Bioresour. Technol. 102, 159-165.

Yashima, H., Fujikake, H., Sato, T., Ohtake, N., Sueyoshi, K., Ohyama, T. (2003). Systemic and local effects of long-term application of nitrate on nodule growth and N_2 fixation in soybean (*Glycine max* [L.] Merr.). Soil Science and Plant Nutrition 49: 825–834.

Yilmaz, D.D. (2007). Effects of salinity on growth and nickel accumulation capacity of *Lemna gibba* (Lemnaceae), J. Hazard. Mater. 147, 74-77.

Yisehak, K., Biruk, K., Abegaze, B., Janssens, G.P.J. (2014). Growth of sheep fed tannin-rich *Albizia gummifera* with or without polyethylene glycol, Trop. Anim. Health Prod. 46, 1113-1118.

Yokohira, M., Yamakawa, K., Saoo, K., Matsuda, Y., Hosokawa, K., Hashimoto, N., Kuno, T., Imaida, K. (2008). Antioxidant effects of flavonoids used as food additives (purple corn color, enzymatically modified isoquercitrin, and isoquercitrin) on liver carcinogenesis in a rat medium-term bioassay, J. Food Sci. 73, C561-C568.

Yong, J.W.H., Ng, Y.F., Tan, S.N., Chew, A.Y.L. (2010). Effect of fertilizer application on photosynthesis and oil yield of *Jatropha curcas* L. Photosynthetica. 48, 208-218.

Zakzeski, J., Bruijnincx, P.C.A., Jongerius, A.L., Weckhuysen, B.M. (2010). The catalytic valorization of lignin for the production of renewable chemicals, Chem. Rev. 110, 3552-3599.

Zeder, M.A. (2015). Core questions in domestication research, Proc. Natl. Acad. Sci. U. S. A. 112, 3191-3198.

Zeller, W.E., Sullivan, M.L., Mueller-Harvey, I., Grabber, J.H., Ramsay, A., Drake, C., Brown, R.H. (2015). Protein precipitation behavior of condensed tannins from Lotus pedunculatus and *Trifolium repens* with different mean degrees of polymerization, J. Agric. Food Chem. 63, 1160-1168.

Zhang, C., Sanders, J.P.M., Bruins, M.E. (2014). Critical parameters in cost-effective alkaline extraction for high protein yield from leaves, Biomass Bioenergy. 67, 466-472.

Zhao, Y., Fang, Y., Jin, Y., Huang, J., Bao, S., Fu, T., He, Z., Wang, F., Zhao, H. (2014). Potential of duckweed in the conversion of wastewater nutrients to valuable biomass: A pilot-scale comparison with water hyacinth, Bioresour. Technol. 163, 82-91.

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 Phytol. 181, 53-61.

Ziegler, P., Adelmann, K., Zimmer, S., Schmidt, C., Appenroth, K.-. (2015). Relative in vitro growth rates of duckweeds (Lemnaceae) - the most rapidly growing higher plants, Plant Biol. 17, 33-41.

Zones, J.M., Blaby, I.K., Merchant, S.S., Umen, J.G. (2015). High-resolution profiling of a synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. Plant Cell 27, 2743-2769.

Zou, C., Zhai, G., Zhang, G., Wang, H., Zhang, G., Li, J., Wang, Z., Wen, Z., Ma, F., Liang, Y., Yang, Z., Li, X., Liang, K. (2015). Formation, distribution, potential and prediction of global conventional and unconventional hydrocarbon resources, Shiyou Kantan Yu Kaifa. 42, 13-25.

Summary

With a growing and wealthier world population, demand for food, fuels and materials is increasing. However, the way we are currently producing our food, fuels and materials has severe negative effects on Earth's ecosystems, including rapid global warming and a loss of biodiversity. To mitigate these negative effects we have to search for new production systems with less environmental impact. Plants will play an important role in these future production systems as they are crucial for our food production and can also be used to produce energy dense transportation fuels and complex chemicals that are now derived from finite oil resource. However, many conventional plant-based production systems are not sustainable, due to intense land-use and excessive input of fossil fuel and fertilizer. Therefore, novel plant feedstocks are sought that provide high amounts of biomass per surface area, whilst requiring minimal input in terms of fossil fuels and fertilizers.

The fern *Azolla* is a free-floating aquatic plant that is capable of rapid growth, but does not rely on the availability of nitrogen fertilizer, due to a special symbiosis with a nitrogen-fixing cyanobacteria. These properties make it an invasive weed in freshwater bodies, such as ditches, lakes or even rivers. However these same properties also make it an promising novel feedstock. The aim of this thesis is to provide knowledge and methods necessary to turn the aquatic weed *Azolla* into a sustainable crop. Using interdisciplinary research we address multiple facets of a potential production chain: generating starting material, cultivation and processing into value adding products.

Being a fern, *Azolla* reproduces by way of spores, instead of seeds. In Chapter 2 we propose to use the sexual reproduction of *Azolla* via these spores to provide reliable starting material for cultivation. We developed methods to collect, store, fertilize and germinate *Azolla* spores. Unlike seed plants, *Azolla* spores could not be stored in a dried form. Drying for 7 days followed by freezing in liquid nitrogen (cryopreservation) allowed to preserve *Azolla* spores at -80°C for up to 7 months, as well as the cyanobacteria symbiont residing inside the spores. The formation of spores (sporulation) could not yet be induced. To better understand the molecular mechanism leading to sporulation, a first transcriptome database of *Azolla* was constructed. The *Azolla* transcriptome database contained many genes alike those regulating sexual reproduction in seed plants including Flowering-locus T (FT). These genes were upregulated in sporulating plants versus non-sporulating plants, suggesting a possible role in regulating sporulation in ferns.

The cultivation of *Azolla* requires a different approach than conventional cropping systems. Whereas most conventional cropping systems operate as a batch

process, *Azolla* cultivation requires continuous harvesting to obtain high yields. In Chapter 4 we have described methodologies to grow *Azolla* in highly productive systems. Growth potentials of *A. filiculoides* and *A. pinnata* were determined at 35.5 and 32.8 t dry weight (dw) ha^{-1} $year^{-1}$, respectively. Addition of CO_2 up to 800 ppm boosted biomass productivity further to 48.3 t dw ha^{-1} $year^{-1}$.

Under ambient CO_2, N_2-fixation by the fern's cyanobacterial symbionts accounted for all nitrogen in the biomass. Detailed investigation of the nitrogen-fixation in Chapter 3 confirmed the capability of the *Azolla-Nostoc* symbiosis to grow rapidly without N-fertilizer, fixing 1200 kg N ha^{-1} $year^{-1}$. Experiments with ^{15}N-labelled nitrogen gas revealed that fixation predominantly occurred during the day, suggesting *Nostoc azollae* uses light to fix nitrogen gas (phototropic), similar to other nitrogen-fixing cyanobacteria. Analysis of diel transcript profiles and microscopic investigations show that the ferns adapted to the phototrophic N_2-fixing symbionts *N. azollae* by 1) adjusting metabolically to nightly absence of N supply using responses ancestral to ferns and seed plants; 2) developing a specialized xylem-rich vasculature surrounding the leaf pocket organ; 3) responding to N-supply by controlling transcripts of genes mediating nutrient transport, allocation and vasculature development.

The chemical composition of *Azolla* biomass was studied in detail to determine which products may be produced from *Azolla* biomass. In Chapter 4 we show that rapidly growing *Azolla* contains between 17.6-20.8 % protein, depending on species and CO_2 treatment. *Azolla* proteins are valuable as feed due to their high content of essential amino acids compared to soybean meal. Since *Azolla*'s protein content was not negatively influenced by CO_2 concentration, the production of *Azolla* protein can be greatly enhanced by utilizing CO_2 waste streams.

Rapidly growing *Azolla* species further contained 7.9-10.0 % lipids and 2.1-6.9 % (tannic acid equivalent) of (poly)phenols, depending on the CO_2 concentration. In Chapter 5 the lipids of *A. filiculoides* grown at elevated CO_2 were analyzed using gas chromatography and mass spectrometry to assess suitability for biodiesel production. The crude lipid extracts consisted of 41±13 % of fatty acids that were converted into fatty acid methyl esters (biodiesel) upon transesterification in methanol. Based on the fatty acid profile it was estimated that *Azolla* biodiesel meets requirements set by the EN14214 standard on fuel density, cetane number and iodine value. However, a fractionation step will be required to decrease the high Cold Filter Plugging Point (CFPP). Through such a fractionation step, small amounts of unique long-chain alcohols and (di)hydroxy fatty acids may be recovered that have possible applications in chemical industry and nutrition.

The relatively high content of (poly)phenols in *Azolla*, compared to foods, may play a role in limiting biomass digestibility of *Azolla* feed. Therefore we performed a detailed characterization of (poly)phenolic compounds in *Azolla* in Chapter 6 by com-

bining colorimetric and mass spectrometry techniques. Condensed tannins, known to limit protein digestibility, were the most abundant soluble phenolic compounds in *Azolla*, accounting for 4.0% and 5.8% of the biomass of A.*filiculoides* and *A. pinnata*, respectively. We further identified dicaffeoylquinic acids as the second most abundant soluble phenolic compound in both *Azolla* species and detected (very) small amounts of G-lignin in *A. filiculoides*.

To understand how these phenolic compounds are synthetized by *Azolla* we analyzed the *A. filiculoides* genome and gene-expression data from sporophytes. We identified various genes with homology to genes responsible for (poly)phenol biosynthesis in seed plants. A Leuco-Anthocyanidin Reductase (*AzfiLAR*-like) gene was the most abundantly expressed gene in the secondary metabolism and likely responsible for the production of the condensed tannins that occur in high amounts in *Azolla*. Manipulating (poly)phenol biosynthesis by knocking out the *AzfiLAR*-like gene could be one approach to decrease the content of condensed tannins in *Azolla* which may enhance biomass digestibility.

In Chapter 7 we investigated the extraction of protein from *Azolla* as a way to generate a valuable protein feed, whilst minimizing the content of anti-nutritional (poly)phenols. We found that the yield of alkaline protein extraction at pH 8 is limited by formation of insoluble protein-tannin complexes. By performing alkaline extraction at pH 10.5 up to 54% of the total biomass nitrogen was recovered in the protein extract, but the increase in pH resulted in strong interactions between (poly)phenols and proteins that limited their subsequent separation. To produce protein extracts with reduced (poly)phenol content we tested two alternative methods: pre-extraction using aqueous acetone and competitive binding of tannins to polyethylene glycol during extraction. Both methods produced a protein precipitate containing 38% of total biomass nitrogen, while considerably reducing (<-67%) (poly)phenol content. Which method is preferable depends on how it can be integrated into a bio-refinery and the value of potential co-products, such as (poly)phenols and carbohydrates.

Altogether, the content of this thesis provides practical tools and novel insights to help researchers and businesses to further develop *Azolla* into a sustainable crop.

Nederlandse samenvatting

Met een groeiende en steeds welvarendere wereldbevolking groeit de vraag naar voedsel, brandstoffen en materialen. De manier waarop we op dit moment deze goederen produceren heeft grote negatieve gevolgen, waaronder de opwarming van de aarde en een verlies aan biodiversiteit. Om deze negatieve gevolgen te beperken moeten we op zoek naar nieuwe productiesystemen die minder impact hebben op het milieu. Planten zullen een belangrijke rol spelen in toekomstige productiesystemen gezien zij cruciaal zijn voor onze voedselvoorziening, maar ook gebruikt kunnen worden voor het produceren van energierijke brandstoffen en complexe chemicaliën die nu worden geproduceerd uit olie, hetgeen uiteindelijk op zal raken. Echter, productieketens die conventionele gewassen gebruiken zijn niet altijd duurzaam, omdat er veel land, kunstmest en fossiele brandstof nodig is. Daarom wordt er gezocht naar alternatieve gewassen waarmee veel biomassa geproduceerd kan worden per oppervlak met minimale input van kunstmest en fossiele brandstoffen.

De varen *Azolla* is een drijvende waterplant die snel kan groeien, maar niet afhankelijk is van stikstofmest, door een speciale symbiose met een stikstof-fixerende cyanobacterie. Door deze eigenschappen is het een invasieve waterplant die sloten, meren en zelfs rivieren kan bedekken. Echter, deze eigenschappen maken het ook een veelbelovend kandidaat-gewas. Het doel van deze thesis is om de kennis en methoden te leveren die nodig zijn om van het aquatische onkruid *Azolla* een duurzaam gewas te maken. Met multidisciplinair onderzoek worden verschillende facetten van een potentiele productieketen belicht: het genereren van startmateriaal, de teelt en het verwerken van biomassa tot producten met toegevoegde waarde.

Als varen, plant *Azolla* zich voort door middel van sporen in plaats van zaaden. In Hoofdstuk 2 stellen wij voor om de seksuele voortplanting via de sporen te gebruiken om betrouwbaar startmateriaal te genereren voor de teelt. We hebben daarvoor methodes ontwikkeld om sporen te kunnen verzamelen, bewaren, bevruchten en te ontkiemen. In tegenstelling tot zaden, konden de sporen van *Azolla* niet bewaard worden door deze te drogen. Door sporen kort te drogen en vervolgens te bevriezen in vloeibare stikstof, ook wel cryopreservatie genoemd, konden deze bij een temperatuur van -80°C voor minstens 7 maanden bewaard worden. Ook de cyanobacterie die zich in de megasporen nestelt, was hierna nog levensvatbaar. De formatie van sporen (sporulatie) kon nog niet teweeg gebracht worden. Om de moleculaire mechanismen die sporulatie teweegbrengen beter te begrijpen werd een eerste transcriptoom

database van *Azolla* geconstrueerd door zijn RNA te sequencen. De transcriptoom database bevatte veel genen overeenkomstig met genen die seksuele voortplanting in zaadplanten reguleren, waaronder Flowering-locus T (FT). Deze genen kwamen hoger tot expressie in planten die aan het sporuleren waren, wat suggereert dat deze ook een rol hebben in het reguleren van sporulatie in varens.

De teelt van *Azolla* vereist een andere aanpak dan de teelt van conventionele gewassen. Waar de meeste conventionele gewassen eenmalig geoogst worden aan het einde van het groeiseizoen, moeten *Azolla* teelten continue geoogst worden om een hoge opbrengst te behalen. In Hoofdstuk 4 beschrijven wij methodologieën om *Azolla* zo te groeien dat een hoge productiviteit behaald wordt. Het groeipotentieel van *Azolla filiculoides* en *Azolla pinnata* werd vastgesteld op respectievelijk 35.5 en 32.8 ton droge stof (ds) hectare^{-1} jaar^{-1}. Het toevoegen van CO_2, tot een concentratie van 800 ppm, verhoogde de productiviteit verder tot 48.3 ton ds hectare^{-1} jaar^{-1}.

Bij atmosferische CO_2 concentraties, werd alle stikstof in de *Azolla* biomassa geleverd door stikstof-fixatie door de symbiotische cyanobacteriën. Gedetailleerd onderzoek naar de stikstof-fixatie in Hoofdstuk 3 bevestigde dat de *Azolla-Nostoc* symbiose in staat is om zeer snel te groeien zonder enige toevoeging van stikstofbemesting, waarbij het 1200 kg stikstof hectare^{-1} jaar^{-1} kan vastleggen. Experimenten met ^{15}N-verreikt stikstofgas toonden aan dat de stikstoffixatie voornamelijk gedurende de dag plaatvindt, wat suggereert dat *Nostoc azollae* net als andere cyanobacteriën energie uit licht gebruikt om stikstof vast te leggen. Microscopie en analyse van genexpressie gedurende de dag toonden aan dat de plant zich op verschillende manieren heeft aangepast aan de *N. azollae* symbiont: 1) aanpassing van zijn stofwisseling aan de afwezigheid van stikstof gedurende de nacht op vergelijkbare wijze als in zaadplanten, 2) het ontwikkelen van een gespecialiseerd xyleem-rijk vatenstelsel die de bladholte omringt waarin de cyanobacteriën leven, 3) het reageren op de stikstoftoevoer door het reguleren van de transcriptie van genen actief in het transport en allocatie van nutriënten en de ontwikkeling van het vatenstelsel.

De chemische samenstelling van *Azolla* biomassa is in detail onderzocht om te bepalen welke producten uit *Azolla* biomassa gemaakt kunnen worden. In Hoofdstuk 4 laten we zien dat snelgroeiende *Azolla* tussen de 17.6-20.8 % eiwit bevat, afhankelijk van de soort en de CO_2 concentraties waaronder zij gegroeid zijn. *Azolla* eiwitten zijn waardevol als veevoer doordat zij een hoge concentratie van essentiële aminozuren bevatten in vergelijking met sojameel. Omdat de eiwitconcentratie niet negatief beïnvloed wordt door de CO_2 concentratie kan de productie van *Azolla* eiwitten sterk verhoogd worden door CO_2 afvalstromen te gebruiken in de teelt.

Snelgroeiende *Azolla* bevat verder 7.9-10 % lipiden en 2.1-6.9% (equivalent van tanninezuur) (poly)fenolen, afhankelijk van de CO_2 concentratie. Om te bepalen of

Azolla lipiden geschikt zijn voor de productie van biodiesel, zijn in Hoofdstuk 5 de lipiden van *A. filiculoides* (gegroeid onder 800 ppm CO_2) in verder detail geanalyseerd met behulp van gas chromatografie en massaspectrometrie. Het ruwe lipidenextract bestond voor 41±13 % uit vetzuren die konden worden omgezet in biodiesel. Op basis van de vetzuur-compositie kon worden ingeschat dat *Azolla* biodiesel voldoet aan de eisen voor energiedichtheid, cetaangetal en joodadditiegetal die gesteld zijn in de Europese EN14214 standaard. Echter een extra fractioneringsstap zal nodig zijn om de filtreerbaarheid bij lage temperaturen te verbeteren. Daarbij zou een kleine hoeveelheid unieke langketenige alcoholen en (di)hydroxyvetzuren verkregen kunnen worden met mogelijke toepassing in de chemische industrie en voedingsmiddelenindustrie.

De relatief hoge concentratie (poly)fenolen in *Azolla* in vergelijking met andere voedselgewassen zou een rol kunnen spelen in het limiteren van de verteerbaarheid van *Azolla* als veevoer. Daarom onderzochten we in Hoofdstuk 6 welke specifieke (poly)fenolen er in *Azolla* voorkomen met behulp van colorimetrie en massa spectrometrie. Gecondenseerde tannines, bekend voor hun negatieve invloed op de verteerbaarheid van eiwitten, bleken de meest veelvoorkomende oplosbare (poly)fenolen in *Azolla* te zijn, goed voor 4.0% en 5.0% van de biomassa van respectievelijk *A. filiculoides* en *A. pinnata*. Verder identificeerden wij dicaffeoylquinic acid als de tweede meest veelvoorkomende (poly)fenol in beide *Azolla* soorten en detecteerden wij kleine hoeveelheden G-lignine in *A. filiculoides*. Om te begrijpen hoe deze (poly) fenolen door de plant worden aangemaakt analyseerden wij het genoom van *A. filiculoides* en bijbehorende genexpressie data. Wij identificeerden verschillende genen die homoloog zijn aan genen die verantwoordelijk zijn voor de biosynthese van (poly)fenolen in zaadplanten. Een leuco-anthocyanidin reductase (*AzfiLAR*-like) gen kwam het meest tot expressie van alle genen in het secundaire metabolisme van *Azolla* en is waarschijnlijk verantwoordelijk voor de productie van de gecondenseerde tannines die in hoge mate in *Azolla* biomassa aanwezig zijn. Het manipuleren van de (poly)fenol biosynthese door het uitschakelen van het *AzfiLAR*-like gen kan een manier zijn om de concentratie van gecondenseerde tannines in *Azolla* te verlagen om zo de verteerbaarheid van de biomassa te verbeteren.

In hoofdstuk 7 onderzochten wij de extractie van eiwitten uit *Azolla* om zo hoogwaardig eiwitrijk voer te produceren met een minimale concentratie van antinutritionele (poly)fenolen. We vonden dat de opbrengst van basische eiwit-extractie bij pH 8 werd beperkt door de vorming van onoplosbare eiwit-tannine complexen. Wanneer de extractie werd uitgevoerd bij pH 10.5 kwamen veel meer eiwitten in oplossing en werd tot 54% van de stikstof in de initiële biomassa in het eiwit extract teruggewonnen. Echter de verhoging in pH zorgde voor sterke interacties tussen eiwitten en (poly)fenolen die de scheiding van beide componenten belemmerde.

Om eiwit-extracten met een verminderd (poly)fenol gehalte te verkrijgen werden twee alternatieve methoden ontwikkeld: een voor-extractie met een waterige oplossing van aceton en competitieve binding van tannines aan polyethylene glycol gedurende de extractie. Met beide methoden kon 38% van de stikstof in de initiële biomassa in het eiwit-extract worden teruggewonnen, terwijl de concentratie van (poly)fenolen aanzienlijk (<-67%) werd verminderd. Welke methode de voorkeur heeft hangt sterk af van hoe deze geïntegreerd kan worden in een bio-raffinaderij en de waarde van potentiele co-producten, zoals (poly)fenolen en vezels.

Alles samengenomen biedt de inhoud van deze thesis praktische methoden en nieuwe inzichten om onderzoekers en ondernemers te helpen *Azolla* verder te ontwikkelen als duurzaam gewas.

Dankwoord | Acknowledgement

Zoals zoveel mooie dingen in het leven is dit proefschrift het resultaat van de input van velen. Graag zou ik dan ook een aantal mensen bedanken voor hun bijdrage te beginnen met mijn promotoren en copromotoren: Sjef, Gert-Jan, Henriette en Klaas.

Sjef, wat betreft de inhoud zag jij meestal al snel dat het wel goed zat. “Een kwestie van uitprinten en een nietje erdoorheen” zei je op het einde, waar ik dan met grote voorzichtigheid op antwoorde. Je gooide mij ook regelmatig voor de leeuwen als er weer ergens een presentatie moest worden gegeven, waar ik redelijke PR vaardigheden aan over heb gehouden. Toch zullen onze gesprekken aan de koffietafel mij het meest bijblijven, over de grote uitdagingen van de 21ste eeuw, de onvoorstelbaar mooie technologieën die er zijn die om de wereld een stukje duurzamer te maken en de frustratie over de onwetendheid van mensen die deze vooruitgang belemmeren. Bedankt dat je het *Azolla* project zo gesteund hebt en proost op nog vele jaren mooi onderzoek en goed golf natuurlijk.

Gert-Jan, hoewel we elkaar na jouw ‘vertrek’ naar NIOZ niet zo vaak gezien hebben, waren gesprekken met jou altijd erg prettig. Waar ik mezelf van tevoren al zat druk te maken over of m’n experimenten en de manuscripten was jouw eerste opmerking meestal: “Ow maar dat gaat heel goed. Niks meer aan doen toch?”. Na onze gesprekken was ik dan meestal ook weer ontspannen en vol goede moed. Bedankt dat je, ondanks dat er geen foraminifeer aan te pas is gekomen, mijn promotor wilde zijn.

Henriette, you have been a great example of thinking outside of the box and not going with the norm. Whereas most people rejected me when I was looking for an internship on biomass and biofuels, you were crazy enough to accept me: a master-student, trained mostly in physics, to start an internship on plant biology and chemistry just because he seemed motivated. Well I think that worked out well :) . Your push to publish my papers yesterday, although sometimes frustrated me, in the end allowed me to finish relatively rapidly. Thank you for teaching me about the marvelous world of biology, which gave me a new way to look at the living world, although respecting its complexity because you were always capable of coming up with yet another incredibly complex molecular mechanism that may explain our results! I hope the *Azolla* project will provide a lot of fun in the years to come, which I don’t doubt with Laura at your side, and that you may have many more great sailing adventures with Ed!

Klaas, jij was altijd de eerste die reageerde op manuscripten, degene die als rots in de branding mijn thesis planning in de gaten hield en toen na twee jaar mijn contract afliep in vogelvlucht geld wist te vinden om mij dat broodnodige halve jaar extra te geven. Dit alles terwijl je zelf niet eens meer bij de UU werkzaam was. Hiervoor ben ik jou enorm veel dank verschuldigd. Eveneens bedankt dat je mij met enthousiasme hebt geïntroduceerd in de wereld van de analytische chemie, jouw geduld met mijn onbenullige vragen die ik tot laat in mijn promotietraject ben blijven stellen, én het herstellen van mijn consistente typfouten. Daarnaast was het ook altijd gezellig, bijpraten met de lunch (ik was altijd op de hoogte hoe AZ ervoor stond) en met als hoogtepunt het 'plant wax' congres in Ascona en de mini-road trip die we daarheen gemaakt hebben. Lange dagen en korte nachtjes (mede omdat je door mijn oordoppen heen wist te snurken), maar met elke avond driegangendiners en Kid Creole door de speakers, hielden we de sfeer er goed in. Bedankt voor alles en op nog vele jaren aan ontdekkingen, lekkere biertjes en goed voetbal in Alkmaar natuurlijk.

I want to extend my gratitude to the members of the reading committee: Pita Verweij, Henk Brinkhuis, Ronald de Vries, Badraldin Sayed Tabatabaei, Andrea Bräutigam and Adrie van der Werf, for their time and effort they invested in reading my thesis and discussing its content. Andrea thank you for all your bioinformatics work on this project and the in-between bioinformatics lessons. Adrie het was een plezier om samen te werken, jij hield mij altijd scherp om de relevantie voor de maatschappij en vooral ook BV Nederland in de gaten te houden ("allemaal leuk en aardig bedacht Paul, maar daar betaald dus niemand voor") en ook vooral om plezier te hebben in 't werk. Badraldin, where the lab-life of a PhD can be lonely it was an absolute pleasure to have you come and visit and work together on our projects, make a lot of jokes and discuss all aspects of life. You have been exceptionally kind and hospital: always bringing gifts, including delicious Gaz!, and incessantly inviting me to come to Iran, which I most definitely will! Also thank you for teaching me the power of observation, which I believe will be of a lot of use to me, also outside of academia (sorry I had to leave).

Tevens wil ik mijn paranimfen, Lennard en Ruud, bedanken. Lennard, wat een eer dat 'de man van 250k' mijn paranimf wil zijn. Jij bent bovendien de aanstichter van dit alles, toen je mij als huisgenoot doorverwees naar Henriette. Het was erg fijn om een goede vriend op de afdeling te hebben om mee te hangen in de gang, begripvolle blikken over de koffietafel te wisselen, wilde zij-projecten te verzinnen en de frustraties of juist het enthousiasme over de resultaten te delen. Succes met de laatste loodjes van jou eigen PhD, begin met schrijven zou ik zeggen en dan vanaf 2 mei volgende jaar eindeloos playstationen en avocado's enten! Ruud, in vergelijking met jou voel ik mij toch een beetje een hobby-wetenschapper, dus des te mooier dat jij

uit Genève wil afdalen om mijn paranimf te zijn. Wat hebben we een hoop mooie discussies gehad, wilde plannen verzonnen en lol gehad op de Otterstraat én daarna. Ik ga nu de wetenschap verlaten om het aan de echte kandidaat-nobelprijs-winnaars zoals jij over te laten, al heb ik nog wel wat ideeën over braincontrol waar ik nog wel input over wil geven hoor, of in Utrecht bij de vuurkorf of in Genève aan de Bains de Pâquis.

Dank aan alle anderen die betrokken waren bij het *Azolla* project. Peter Bijl, jij speelde een belangrijke rol in het gaande houden van het *Azolla* project via de LPP stichting, door mijn PhD project te financieren, contacten met geïnteresseerden te organiseren en het produceren van de '*Azolla* movie'. Dan Brinkhuis, bedankt voor het maken van deze mooie film over ons project, welke ons én de *Azolla* gemeenschap enorm heeft geholpen (we zitten al over de 30.000 views!, ondanks mijn wilde haar dat ik toch even had moeten laten knippen van tevoren ;). Jan de Leeuw, bedankt voor jouw input op de *Azolla* massa-spectra en de presentatie in Ascona.

Fred, Gerard, Roel en alle andere medewerkers en vrijwilligers van de Botanische Tuinen. Bedankt dat jullie altijd behulpzaam waren wanneer team *Azolla* weer eens kasruimte nodig had, er om 17u 's middags een collectie *Salvinia*'s uit Duitsland moest worden veiliggesteld of we opzoek waren naar dat ene specifieke plantje om te analyseren. Voor mijn project zijn jullie van onschatbare waarde gebleken en ik hoop dat jullie nog lang van onschatbare waarde kunnen zijn voor de Universiteit. Ronald Leito en Frits Kindt bedankt voor alle uitleg over de microscopen, foto opstellingen, z-stacks én alle prachtige foto's die jullie gemaakt hebben die zo belangrijk waren om ons werk te kunnen communiceren.

Thanks to my many current and past colleagues and visitors in the Molecular Plant Physiology group: Martijn, Marcel, Evelien, Jolanda, Laura, Savani, Sjors, Nikita, Dorrit, Lazaro, Monika, Wouter, Johannes, Jeba, Bing, Magdalena, Jeroen, Alessia, and Nicole, for helping this layman understand more about plant molecular biology. Marcel bedankt voor jou uitvoerige en zeer enthousiaste uitleg over ontwikkeling en bloei. Evelien bedankt voor jouw geduld mij qPCR en cloning uit te leggen, ook als ik het spoor weer eens helemaal bijster was. Jolanda als echte "madre de la familia" hield je in de gaten hoe het met de mensen ging, wist je te vertellen wat je als PhD nou eigenlijk moest regelen en stak me een hart onder de riem als ik in de stress zat. Veel dank hiervoor en natuurlijk alle hulp in het lab! Lazaro, thank you for all the Spanish lessons, Espero vamos reencontrarnos en Cuba una dia! Esa la cosa! Savani, succes met de laatste loodjes en mocht je toch ooit nog een carrière ambiëren als crimineel meesterbrein ben ik van de partij. Sjors, wat een eer dat ik met een echte van der Horst heb mogen werken, veel succes met jouw PhD en tot weerziens in de boulderhal. Laura, wat een geweldige opvolgster ben jij. Geduren-

de je master was je mij al bioinformatica aan het leren, je wist het *Azolla* project bij de UU toekomst te geven door zelf geld binnen te halen en je bent bovenal super sociaal. Ik heb dan ook geen enkele twijfel dat het goed gaat komen met je PhD.

Many thanks to all the current and past colleagues and visitors in the Earth Sciences department: Francien, Loes, Frédérique, Tjerk, Marieke, Niels, Lottie, Jeifei, Arnold, Linda, Dominika, Anita and Coen. Francien bedankt dat jij mij als vreemde planten-man als onderdeel van het lab hebt opgenomen. Loes, eigenlijk vanaf dag één dat wij samen startten aan de PhD was jij diegene die zorgde dat ik op de hoogte was en betrokken werd bij alle activiteiten van de groep. Bedankt ook voor alle gezelligheid tijdens de het plant-wax congres en vooral ook de road-trip erheen en terug met Herr Willy en Kid Creole. Succes met het afronden van je PhD, af op 1 mei hè maak ons PhD's trots! ;). Arnold, met een totaal van 795 koolstof en stikstof-metingen mag ik wel zeggen dat jij cruciaal geweest bent voor het proefschrift van deze Bierbrouwer. Dat je nog een hoop meetplezier mag hebben en vooral ook een hoop fietsplezier.

Thanks to all the students that helped me with my research in the past three years: Mark, Wijdan, Anne, Valerie, Dally en Wouter. Mark jij was een uitzonderlijk gemotiveerde en enthousiaste student, dat bleek ook wel toen je later nog een keer terug kwam voor een 3-weken project eiwitten extraheren en in je stage op ECN mij in contact bracht met Wouter Huijgen om de suikerinhoud van *Azolla* in kaart te brengen. Het verbaast mij niets dat je door bent gegaan voor je master in Wageningen. Bedankt voor al 't werk dat je verrichte, veel succes met jouw masterstudie en wij komen elkaar ongetwijfeld nog tegen. Dally and Wouter, your work has been very important for finishing the last chapter of this thesis. Thank you for all the effort you put into this project and I wish you both a very bright career in science and/or business.

Naast de vele mensen die mij in mijn werk hebben gesteund zijn er minstens evenveel naast mijn werk die hebben bijgedragen aan het afronden van dit proefschrift. Pap en Mam, dit boekje is er dankzij het analytisch vermogen en de doorzettingskracht die ik van jullie heb gekregen én jullie onvoorwaardelijke steun voor wat het dan ook is dat ik in mijn hoofd haal om te willen doen. Het was heerlijk af en toe weer een week thuis te bivakkeren, waar ik in alle rust (én onder het genot van michelen-ster-waardig eten uit de moestuin) aan mijn proefschrift kon schrijven.

Maartje en Reinhoud, bedankt dat dat ik altijd op jullie kan rekenen, van het regelen van meubels en spulletjes voor ons nieuwe huis tot ontspannen-met-katten-therapie wanneer ik er doorheen zat. Maartje, wellicht ben je als grote zus meer een voor-

beeld geweest dan ik ooit heb toegegeven: gaan studeren in Utrecht (waar ik inmiddels door de bezoekjes aan het IBB een mooi beeld van had gekregen), een master cum laude halen moest ik natuurlijk ook wel evenaren en 't promoveren (ondanks jouw meervoudige waarschuwingen). Maar ik denk dat we het er nu over eens kunnen zijn dat we die academische wereld allebei wel zo'n beetje uitgespeeld hebben. Reinhoud, jij moet nog even de eindbaas verslaan, wat je gezien je arbeidsethos en enthousiasme over de experimenten zeker moet lukken, maar misschien ben je wel de meest verstandige en begin je een Zeeuwse algenkwekerij?

Dank aan al mijn vrienden voor de mooie avonden, weekendjes weg en vakanties die zorgden voor mooie discussies, nieuwe inzichten, maar ook vooral de broodnodige ontspanning en training van de lachspieren.

Annemarie, er zijn veel mensen die hebben bijgedragen aan de afronding van dit boekje, maar er is er maar één zonder wie het nooit was afgekomen. Jij bent gedurende deze hele tijd de rem geweest die ik nodig had om niet door te schieten in m'n werk, maar ook van 't leven te genieten en wat kunnen wij dat goed samen zeg! Sorry dat ik jouw plannen voor nieuwe avonturen soms in de weg zat, want tja die PhD moest af. Maar dat is nu officieel voorbij en dit is het begin van 'het jaar van het risico' en al onze avonturen die daarop mogen volgen!

Curriculum Vitae

Paul Brouwer was born on April 28th 1988 in Utrecht, the Netherlands. In 2006 he graduated from high school (VWO, Etty Hillesum Lyceum in Deventer). After a very brief period studying Architecture in Delft, he decided to start a bachelor in Science and Innovation Management at the University of Utrecht. During this bachelor study he developed an interest in sustainable energy and materials, which he pursued further in the master Energy Science at the University of Utrecht. As part of his master studies he perfomed an internship on solar energy, researching the optical and thermochemical performance of a novel solar concentrator module developed at the start-up Suncycle in Eindhoven. For a second internship he choose to focus on biomass production for sustainable energy and materials and came across the *Azolla* project at the University of Utrecht. After a succesfull internship on *Azolla* domestication and biodiesel production he obtained his master of Science degree (*cum laude*) in 2013 and was able to continue his work on *Azolla* as a junoir researcher. During this period he participated in the Climate KIC 'pioneers into practice' exchange program and worked for one month at the University of València, Spain, investigating the removal of heavy metals from wastewater using plant biomass. In 2014 he continued his work on *Azolla* as a PhD candidate, the results of which are presented in this thesis. Currently he is continuing with the development of aquatic farming systems as an independent entrepreneur.