

# Viewpoints

# The Sphagnome Project: enabling ecological and evolutionary insights through a genus-level sequencing project

#### **Summary**

Considerable progress has been made in ecological and evolutionary genetics with studies demonstrating how genes underlying plant and microbial traits can influence adaptation and even 'extend' to influence community structure and ecosystem level processes. Progress in this area is limited to model systems with deep genetic and genomic resources that often have negligible ecological impact or interest. Thus, important linkages between genetic adaptations and their consequences at organismal and ecological scales are often lacking. Here we introduce the Sphagnome Project, which incorporates genomics into a long-running history of Sphagnum research that has documented unparalleled contributions to peatland ecology, carbon sequestration, biogeochemistry, microbiome research, niche construction, and ecosystem engineering. The Sphagnome Project encompasses a genus-level sequencing effort that represents a new type of model system driven not only by genetic tractability, but by ecologically relevant questions and hypotheses.

#### Introduction

The discovery, characterization, and prediction of genes associated with traits, and how those traits influence ecosystem function, are key challenges, especially in the face of changing climatic conditions (Whitham et al., 2006). Climate-driven alteration of biological processes occurs across all levels of organization, and is expected to impact a wide range of ecosystem goods and services including biodiversity, nutrient cycling, climate feedback regulation, and productivity (Rockström et al., 2009). However, our ability to associate genes with traits of ecological interest is generally restricted to plant model systems primarily developed for crop and bioenergy feedstocks, and further limited by the sheer complexity of applying genetic and genomic approaches to multiple species or communities. Yet the need to apply system genetic approaches in complex communities is paramount as evolution takes place within a complex web of genetic interactions among species (Whitham et al., 2006).

Here we argue that the genus *Sphagnum* (peat moss) represents an unparalleled model system for ecological and evolutionary genomics, empowered by its contribution to global carbon cycling and emerging genomic resources. Sphagnum species play a major role in peatland formation, a prime example of ecosystem engineering, whereby the organism manipulates its surrounding habitat. Sphagnum primary production influences carbon and nutrient cycling, such as methane production and soil carbon storage, in many boreal forests and peatlands (Turetsky et al., 2012). Sphagnum ecosystem engineering involves the accumulation of peat that facilitates its own growth while making the surrounding environment hostile for vascular plants (van Breemen, 1995). Ultimately these multi-level processes lead to the formation of peatlands that occupy nearly 3% of the land surface and store 25% of the world's soil carbon as recalcitrant peat (Yu et al., 2010). The latter point has led to the assertion that *Sphagnum* has a greater impact on global carbon fluxes, and therefore climate, than any other single genus of plants (Clymo & Hayward, 1982; van Breemen, 1995).

The Sphagnum sequencing project provides a novel nonfood crop or nonbioenergy feedstock example for a plant-based genome sequencing project aimed specifically at carbon cycling. The project is developing resources for within-species genetic associations with ecologically relevant functional traits, and the extension of those gene-to-trait relationships to additional species within the Sphagnum genus. We refer to this effort collectively as the Sphagnome Project. In the following sections, we provide a brief introduction to the ecology and evolution of this unique plant genus. We then outline a research roadmap that highlights scientific questions relevant to the disclosure and use of a genus-wide genomic resource for Sphagnum in two major areas of distinct but overlapping research: (1) carbon sequestration and global biogeochemistry; and (2) niche construction, ecosystem engineering, and microbial associations. We demonstrate that the Sphagnome Project is an example of a novel model system aimed at addressing ecologically relevant questions and hypotheses across levels of organizations.

### Sphagnum ecology and evolution

#### Functional traits and ecosystem function

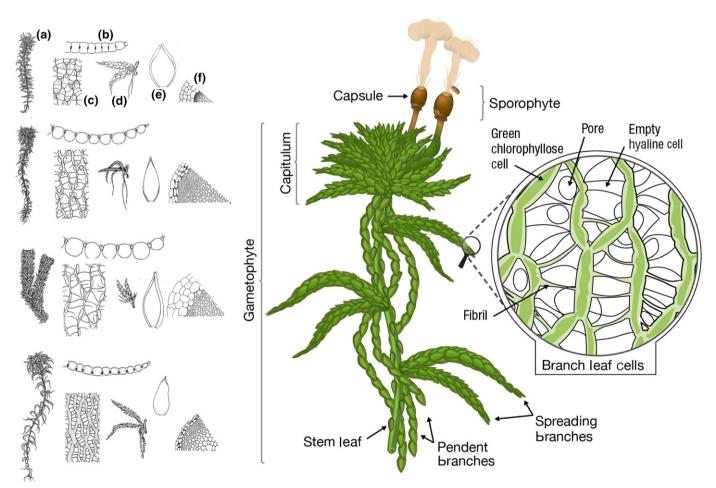
Sphagnum has a remarkable ability to create and then uniquely thrive in nutrient-poor, acidic, and waterlogged conditions. The suite of morphological, physiological, and life history traits that affect *Sphagnum* fitness, herein termed functional traits, enable this 'ecosystem engineer' (Jones et al., 1994) to gain a competitive advantage over other co-occurring species and therefore flourish under relatively harsh environmental conditions. For example, the ability of Sphagnum to store and transport water is controlled

largely by three distinct morphological adaptations – branching architecture, leaf size and arrangement on branches, and hyaline cells (Fig. 1a,b; Rydin & Jeglum, 2013). These traits differ considerably among species, and are associated with highly partitioned microhabitat preferences where Sphagnum species coexist within a peatland. Hummock-forming species, growing c > 30 cm above the water table, have small close-set leaves forming numerous interconnected small capillary spaces (Fig. 1). Spreading branches allow lateral movement of water through the capillary continuum, while numerous close-set pendant branches appressed to the stem form an efficient vertical water-transport system. Consequently, Sphagnum species growing on hummocks can wick moisture and maintain metabolic activity even during drought (Rice & Giles, 1996). In all species, dead hyaline cells in the leaves and the outer cortex of the stems and branches act as water-storage structures.

The capitula at the top of the stems are alive, but a few (c. 5) centimeters down 99% of the light has been absorbed and most of

the Sphagnum cells die (Hayward & Clymo, 1983). From there down to the water table the carpet structure is permeable to water and gases (particularly oxygen) and the damp plant substrates begin to decay in this oxic zone, termed the acrotelm (Ingram, 1978; Clymo & Hayward, 1982). The consequent loss of stem strength and increasing weight eventually result in collapse of the plant structure. This reduces the pore size so water can no longer flow easily through it, and from this point downwards the peat is permanently waterlogged and this is what determines the depth of the water table. In this waterlogged zone, oxygen is consumed by aerobic respiration more rapidly than it can be replenished by diffusion (which is 10 000 times slower in water than it is in air), creating the anoxic catotelm (Clymo, 1983). Hence, through distinct traits, Sphagnum generates environmental conditions that are suitable for its own growth but hostile for the vast majority of other plants (e.g. van Breemen, 1995; Rydin & Jeglum, 2013).

The mechanisms by which *Sphagnum* inhibits fungal and microbial decomposition — and hence promotes peat accumulation — are not



**Fig. 1** Morphological traits of *Sphagnum*. Left panel, four representative species (modified from Crum, 1984). (a) Plant habits showing differences in branch density. (b) Branch leaf cross-sections showing arrangements of larger hyaline cells. As in most mosses, *Sphagnum* leaves consist of a single layer of cells, but unlike in other mosses, the leaf cells are dimorphic, comprising large hyaline cells, dead and empty at maturity, alternating with narrow photosynthetic chlorophyllose cells. In some species (e.g. top), those chlorophyllose cells are not exposed at the leaf surface and in other species they are exposed at the inner or outer surface. (c) Surface view of branch leaf cells, showing variously arranged pores on hyaline cells. The chlorophyllose cells are very narrow, forming a network around each hyaline cell. (d) Branch fascicles, each including so-called spreading and pendent branches. (e) Branch leaf. (f) Stem cross-section showing variously developed, sometimes enlarged outer cortex cells. Right panel, one (haploid) gametophyte plant with stalked capsules releasing spores (modified from Weston *et al.*, 2015). Inset, detail of branch leaf cells showing differentiation of chlorophyllose and hyaline cells.

fully understood, but involve both the external environment engineered by the species, as well as the internal biochemistry of its plant tissue, particularly the low nitrogen: carbon (N:C) ratio (a reflection of the unusually efficient use of nitrogen in producing new biomass) (Bragazza et al., 2006). A passive mechanism for intrinsic decay resistance in the oxic acrotelm layer is suggested by the correlation of microbial decomposition of Sphagnum litter with the relative amounts of structural vs metabolic carbohydrates (Turetsky et al., 2008). Active mechanisms of antimicrobial activity are also implicated, mainly through acid hydrolysis of cell-wall polysaccharides, fragments of which are released into the soil water as 'sphagnan' (Hájek et al., 2011). The precise mechanisms for the antimicrobial activity of sphagnan are still under investigation, but may involve lowering soil pH, reducing availability of nitrogen and carbon, or interfering with extracellular enzymes by immobilizing them in a polyelectrolyte complex (Hájek et al., 2011). Soluble phenolic compounds, either leached directly from Sphagnum tissue or produced during its breakdown, may play a more minor role in tissue preservation, physically protecting polysaccharides through the formation of humic substances (Hájek et al., 2011). While environmental factors such as soil oxygen profiles serve as important regulators of peat decomposition (cf. Freeman et al., 2001) it is clear that a variety of mechanisms contribute to slow decomposition of *Sphagnum* tissue, thereby retarding the turnover of organic biomass in peatlands and sequestering carbon in the form of peat for centuries.

### Phylogeny and evolution

Like all mosses, the haploid gametophyte is the dominant life cycle stage for Sphagnum (Fig. 1). Haploid spores germinate into a filamentous protonema, quickly followed by a thalloid protonemal phase, before transitioning into mature haploid gametophytes. A single spore can result in a large clonal biomass through vegetative growth. Furthermore, the ability to propagate clonally is ubiquitous in Sphagnum and typical clone sizes vary among species (Cronberg, 1996). In S. austinii, one clone occurs throughout North America and the same dominates in Europe (Kyrkjeeide et al., 2016). A single clone of S. subnitens extends from Oregon to the westernmost Aleutian Islands (Karlin et al., 2011). Reproductive seasons are species-specific and sperm require water to access the egg cell in the archegonial venter to form the zygote. The formation of the zygote marks the beginning of the brief diploid stage of development and at maturity meiosis occurs within the capsule, producing haploid spores.

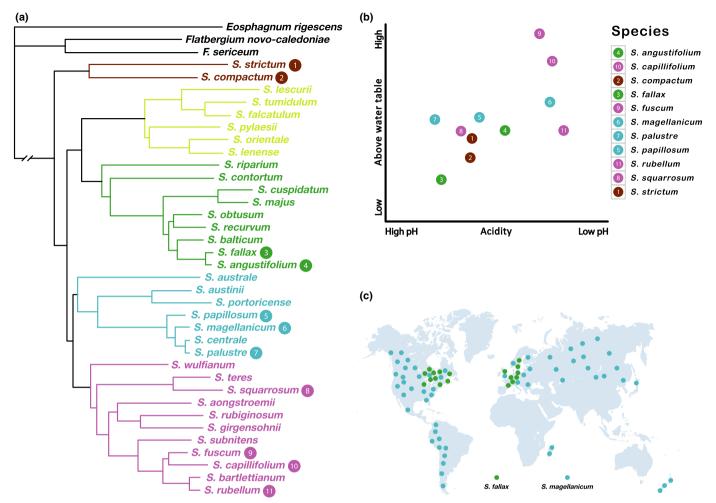
Sphagnum is one of four genera in the class Sphagnopsida (phylum Bryophyta: mosses), an ancient lineage of land plants. Molecular phylogenies suggest the Sphagnopsida diverged from other mosses > 250–350 million years ago (mya) (Shaw et al., 2010), and fossils of peat moss-like fragments, which are the oldest known land plant macrofossils to date, have been found in the Ordovician rocks (c. 500 mya, Cardona-Correa et al., 2016). Fossil Sphagnum and close relatives are recognized by the unique cell pattern in leaves. Three of the genera in the Sphagnopsida contain just one or two species each, and none of them form extensive peats nor do they dominate wetlands as do species of Sphagnum. With

200–300 species, *Sphagnum* is by far the largest genus in the Sphagnopsida and the most important for peatlands. *Sphagnum* species share a common ancestor in the late Tertiary, a surprisingly recent radiation considering the great antiquity of Sphagnopsida (Shaw *et al.*, 2010). This recent radiation, which may have occurred following the mid-Miocene climatic optimum, coincides with the rise of boreal peatlands in the Northern Hemisphere (Greb *et al.*, 2006).

Today, Sphagnum occurs on all continents aside from Antarctica (Crum, 1984). The genus dominates wetland habitats throughout the boreal zone of the Northern Hemisphere but is also diverse at tropical latitudes, especially in South America (as well as in tropical Africa and Asia). At tropical latitudes, Sphagnum sometimes occurs in high altitude peatlands, but in lower altitude tropical regions they typically grow on wet soil banks, along streams, and on dripping rocks, and do not accumulate substantial amounts of peat. Sphagnum comprises five major subgenera (Fig. 2a; Shaw et al., 2016a). The small subgenus Rigida (c. 2-4 species), sister to the four other subgenera, sometimes occur in peatlands, but its species are never dominant and are not major peat-formers. Most Sphagnum species belong to the remaining two clades, both of which include important peat-forming species. The species in one clade (subgenera Cuspidata + Subsecunda) generally occupy hollows close to or at the water table, whereas those in the other clade (subgenera Sphagnum + Acutifolia) generally create lawns and raised hummocks more distant from the water table (Fig. 2b). For decades, peatland ecologists have noted that individual Sphagnum species have narrow realized niches along this hydrological gradient - from low hollow to high hummock (Vitt & Slack, 1984). Sphagnum species also exhibit narrow preferences along a chemical gradient, with some species preferring acidic ombrotrophic bogs and other species preferring fens with more neutral pH. Unlike preferences along the hydrological gradient, species preferences along the chemical gradient do not exhibit a strong phylogenetic signal (Johnson et al., 2015). During the rapid radiation of modern Sphagnum, microhabitat preferences along the chemical gradient plausibly evolved simultaneously in unrelated groups, creating natural experiments with which the genetic basis of microhabitat preferences can be disentangled from phylogenetic history.

# Developing resources for a tractable *Sphagnum* model system with evolutionary and ecological relevance

Genomic resources for *Sphagnum* are rapidly expanding (https://phytozome.jgi.doe.gov). The Sphagnome Project will provide two high quality reference genomes (*S. magellanicum* and *S. fallax*), sequences for 15 additional species across the *Sphagnum* phylogeny (Fig. 2), and shallow sequencing of *c.* 200 individual members from a haploid-sib pedigree. A draft genome for *S. fallax* is now available on https://phytozome.jgi.doe.gov. These Sphagnome Project resources are motivated by two overarching aims: (1) identifying genetic associations with ecologically relevant functional traits within species; and (2) extending those gene-to-trait relationships to additional species within genus.



**Fig. 2** Distribution, phylogeny and habitat preference of species within the Sphagnome Project. (a) A recent phylogeny based on Shaw *et al.* (2016a) with colored branches representing subgenus designations (brown, *Rigida*; yellow, *Subsecunda*; green, *Cuspidata*; blue, *Sphagnum*; purple, *Acutifolia*) and colored circles next to species being sequenced with the Sphagnome Project; (b) generalized habitat preferences for *Sphagnum* species typical of boreal peatlands, in relation to pore water pH and height above water table; (c) global distribution of *Sphagnum fallax* (green) and *Sphagnum magellanicum* (blue). Note that *Sphagnum affine* (*Sphagnum*), *Sphagnum cribrosum* (*Subsecunda*), *Sphagnum fimbriatum* (*Acutifolia*) and *Sphagnum molle* (*Acutifolia*) are not in the figure because they are not boreal peatland species, but have been sequenced as part of the Sphagnome Project.

#### Sphagnum pedigree sequencing and gene-to-trait mapping

The Sphagnome Project is producing high-quality reference genomes for S. magellanicum Brid. and S. fallax H. Klinggr (Shaw et al., 2016b). These two peat-forming species are in different subgenera, occupy very different microhabitats in boreal peatlands, and will provide strong contrasts for investigating phylogenetic and ecological differences (Fig. 2; Johnson et al., 2015). To fulfill the first aim focusing on within-species variation, the Sphagnome Project will conduct resequencing of c. 200 individuals from a S. fallax pedigree to generate a high quality genetic linkage map that will facilitate gene-to-trait experimental approaches (Fig. 3) and genome assembly. The pedigree was developed from single stem descent propagation using sporelings germinated from a single field collected sporophyte; all individuals are haploid sibs. Because S. fallax has separate gametophytic sexes, pedigree individuals can be maintained in clonal culture without risk of intra-gametophytic selfing. Preliminary data show vast phenotypic variation among haploid siblings in response to laboratory growth conditions, temperature and pH (Shaw et al., 2016b). Sphagnum is haploid in its dominant life cycle stage, which eliminates the confounding heterozygosity that can mask allele expression. Therefore, the F<sub>1</sub> (gametophytic) generation can be used in trait mapping, which is not possible for genetic studies in diploid nonbryophyte organisms where, at a minimum, a segregating F<sub>2</sub> pedigree is required. Furthermore, the paternal genotype can be reconstructed by subtracting the progeny genetic markers from the maternal markers. This latter point is especially important, as controlled crosses are currently difficult to perform in Sphagnum. As recently shown in the Sphagnum moss-relative Physcomitrella patens (Stevenson et al., 2016), the simplified genetics of mosses coupled with linkage-analysis can provide a powerful means of predicting phenotypes from DNA markers and their underlying causal alleles (Fig. 3).

Recent advances in maintaining *Sphagnum* tissue cultures (Beike *et al.*, 2015) have improved the reliability of producing axenic

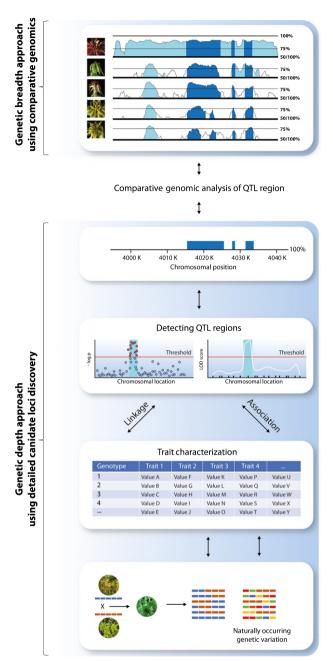


Fig. 3 Schematic of the proposed depth and breadth genetic approaches (Fig. 4 legend). In gene-to-trait studies, linkage-based and association mapping are main approaches used to discover (or map) the genetic basis of quantitative phenotypic variation. Both assume that there is variation for the traits of interest within the population being studied. The linkage-based method relies on individuals with known relationships to each other and DNA variants (termed genetic markers) that segregate through the population. The genetic marker is 'linked' through proximity to the causal loci and they therefore segregate together. Association mapping does not require known relationships among individuals within the population, but instead relies on historical recombination from many generations of random mating. Together these methods constitute the 'genetic depth' approach discussed in the text aimed at identifying candidate genes (lower panel) that are then included in phylogenomic and comparative genome analyses (upper panel). These analyses are simplified by the fact that Sphagnum gametophytes are typically haploid. Two allopolyploid species (S. palustre, S. papillosum) are included to address subsidiary issues related to the evolution of polyploid genomes.

cultures that produce Sphagnum plants that are morphologically similar to field-collected specimens. The Sphagnome Project encompasses a developing germplasm collection that includes culture material for all species being sequenced and a S. fallax haploid-sib pedigree. The low stature of Sphagnum and ease of establishing populations in trans-well culture plates that have relatively small 'bench top' space requirements enable rapid phenotyping that is necessary for gene-to-trait studies (Fig. 3). Further, this germplasm collection can be used to test responses of Sphagnum genotypes to different environmental conditions. Because the complete genomes of these genotypes will already be known as a result of resequencing, genetic associations can be made as soon as phenotypic data are collected. Due to the small size of Sphagnum and other mosses, imaging-based phenotyping will be especially useful in this effort. Single images can capture data on hundreds of individuals, entire populations, and mixed communities, simultaneously aiding the linkage of genes to traits. The broader collection of gene-to-trait associations can be integrated in network models to form a systems biology view of the trait combinations and their correlations underlying phenotype expression and adaptation (Chitwood & Topp, 2015).

#### A genus-wide approach

Extending gene-to-trait relationships beyond a single species is necessary for understanding the evolution of ecosystem function in Sphagnum-dominated peatlands. Traits important for ecosystem function differ among species, including productivity and resource acquisition, resource allocation such as production of secondary compounds, and decomposition rates (Bengtsson et al., 2016; Limpens et al., 2017). Therefore, in addition to the intensive within-species resequencing approach described earlier, the Sphagnome Project includes the sequencing of 31 individuals across 15 species representing the five major clades within *Sphagnum* (Fig. 2). This information, combined with ongoing and existing transcriptome resources (Devos et al., 2016), will provide the basis for genuslevel phylogenomics and comparative genomic analyses in Sphagnum (Fig. 3). This approach is especially useful for the majority of traits in *Sphagnum* where interspecific variation seems to be greater than intraspecific variation (e.g. Bengtsson et al., 2016). Genetic associations will be tested using models that incorporate phylogenetic comparative methods (e.g. Blomberg & Garland, 2002; Revell, 2009) to account for phylogenetic distance when identifying gene-to-trait relationships.

Through this sequencing effort, gene-to-trait relationships of multiple species will be placed within a broader phylogenomic landscape thereby identifying evolutionary patterns associated with microhabitat preferences and functional traits (Figs 2b, 4). While a few recent studies have taken a genus-wide approach to genetic associations (e.g. Haudry et al., 2013; Novikova et al., 2016; Pease et al., 2016) the Sphagnome Project encompasses species that co-occupy and engineer the same ecosystem. We anticipate that these genus-wide sequences, phenotype data, and comparative gene-to-trait relationships will enable the detection of genes under purifying or positive selection as well as gene family evolution associated with major ecological and biogeographic shifts.

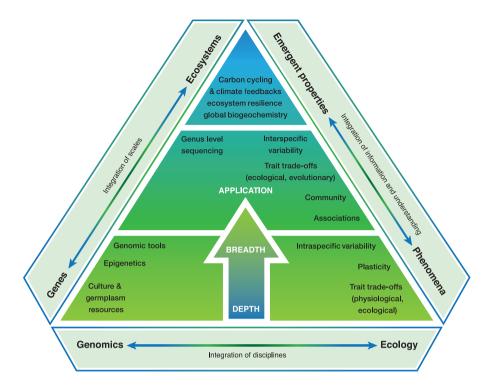


Fig. 4 An integrated approach for *Sphagnum* as a model system linking genetic information on genes underlying functional traits (depth) with phylogenomic analyses (breadth) to large-scale, emergent properties at the level of the ecosystem. Increases in the availability of genomic resources and recent developments of germplasm resources can facilitate collaborative research across multiple disciplines. Understanding the genetic basis of integrated traits will facilitate our understanding of trait-trade-offs, fitness and selection, and response to environmental change.

# Facilitating new ecological and evolutionary understanding

What is the biological basis of unique *Sphagnum* traits or combinations of traits, and how do these trait combinations extend beyond the organism?

Tissue chemistry is a noted functional trait for Sphagnum (Clymo & Hayward, 1982). Polyuronic acids (cell-wall polysaccharides that form a pectin-like polymer) comprise 10-30% of Sphagnum dry mass. They have a high cation exchange capacity (CEC) initially satisfied with H+, which is rapidly exchanged for cations in rainwater, thus making the water around the plants acidic (Clymo & Hayward, 1982) and make cation nutrients unavailable to microbes and other plants (Stalheim et al., 2009). However, the question of a possible link between unique organic compounds and niche engineering by Sphagnum remains a matter of active research (Hájek, 2009; Limpens et al., 2017). It has long been speculated that living Sphagnum benefits from peat formed over time through the accumulation of dead Sphagnum biomass (van Breemen, 1995). Should this be viewed as one type of extended phenotype, where the phenotype of vertically accumulating peat (dead Sphagnum material) changes the function of living Sphagnum at the surface? Sphagnum plants clearly modify their environment in several important ways, but how this influences selection on future offspring and other recipient organisms is unknown. We believe that the Sphagnum genomic resource offers one of the best opportunities to explore these questions and ultimately identify the genetic basis for the traits responsible for ecosystem engineering in Sphagnum. For example, what is the genetic basis of tissue chemistry traits, and do these traits impart a fitness

advantage from a nutrient competition perspective? Furthermore, how do these traits extend beyond the organism? For example, do hummock formation traits covary with tissue chemistry and decomposition rates, and how will these currently adapted trait combinations influence fitness to changing environmental conditions? In regard to niche engineering, is there evidence for an extended phenotype in Sphagnum, and if so, what is the unit of selection, and at which level does selection occur (Whitham et al., 2003)? Do neighborhood effects, such as the genetic effect of an individual on trait values of neighboring individuals influence how Sphagnum traits interact with the environment? How important is clonality to the extended *Sphagnum* phenotype? These important questions extend into much broader spheres of the Sphagnome Project (Fig. 4) and general ecological and evolutionary theory.

# Did adaptation to spatially or temporally varying climate variation spark *Sphagnum* species radiations?

Genus-wide phylogenetic analyses of geographic ranges support the view that the two major peat-forming, crown clades within *Sphagnum* (*Acutifolia+ Sphagnum*; *Cuspidata+ Subsecunda*) (Fig. 2a,b) originated and first diversified in the Northern Hemisphere (J. Shaw *et al.*, unpublished). By contrast, phylogenetic analyses of large seed plant clades that span tropical and Northern Hemisphere ranges usually reveal tropical origins and rare expansions into cold northern climates (Jansson *et al.*, 2013). *Sphagnum* represents one of a small minority of groups that appear to have initially diversified at northern latitudes and subsequently extended their ranges into the tropics. Phylogenetic patterns indicate that southward range expansions were followed by

evolutionary radiations that gave rise to groups of tropical species nested within larger boreal clades.

Moreover, nonboreal radiations occurred in each of the four large subgenera of Sphagnum, providing phylogenetic patterns that can be used as replicated natural experiments to account for shared ancestry when investigating the genetic basis of adaptation and the evolution of functional traits associated with range expansions. In addition to these radiations, a few individual boreal Sphagnum species have extended their ranges into tropical habitats, presumably more recently. Interspecific and intraspecific comparative analyses can be harnessed to address several questions. What genes, gene families, and genomic regions underwent changes associated with range expansions from boreal to tropical climate zones? Are the same genomic features associated with intraspecific and with interspecific range changes across climate zones? Are the same or similar genomic changes associated with climate adaptation in different Sphagnum subgenera, associated with independent range changes? Clarifying functional trait and genomic changes associated with migrations into warmer climates can provide informative analogies to how Sphagnum mosses and, perhaps, other plants may respond to current climate warming.

## What are the factors that limit or facilitate local-scale adaptive evolution?

There has been much interest regarding the importance of phenotypic plasticity relative to local adaptation in response to environmental heterogeneity, and how such responses can ultimately extend to influence ecosystem function (Miner et al., 2005). The sequenced haploid-sib pedigree coupled with phenotype screening will provide the resources necessary for quantitative genetics to determine the extent to which a phenotypic change has a quantitative genetic basis (see the 'Developing resources for a tractable Sphagnum model system with evolutionary and ecological relevance' section). Plasticity is inferred as the proportion of phenotypic variance not explained by genetics (Merilä & Hendry, 2014). The use of common gardens, especially when established among multiple environments with appropriate replication and controls, provides a powerful approach to disentangle genetic from plastic contributions to phenotype. The sequenced Sphagnum haploid-sib pedigree and emerging research community surrounding the Sphagnome Project make the establishment of common gardens with characterized genotypes a reality. Finally, the demonstration that allele frequency shifts occur confirms that evolution has occurred, with the challenge being the need to determine if changes in specific allele frequencies are relevant to the traits and phenomena being investigated. The sequencing of 15 Sphagnum species and nearly 200 progeny individuals provides an ideal system to determine shared and species-specific components of the collective genome and relationships that co-occur with phylogenetic signals. For example, does a gene family expansion coincide with the lineage diversification to novel environments? Together with common garden experiments we will begin to address questions centering on the relative importance of local adaptation vs phenotypic plasticity in Sphagnum responses to environmental heterogeneity.

What is the role of *Sphagnum* and its interacting microbiome in ecosystem carbon and nitrogen cycling?

Hyaline cells not only play a vital function as water storage organs, but also create a novel and safe habitat for a diverse microflora spanning all domains of life (Fig. 1; Bragina et al., 2012; Kostka et al., 2016). The Sphagnum-associated microbiome seems to be divided into two broad categories: those that are host species specific, with specificity maintained across both the sporophyte and gametophyte generations (Bragina et al., 2012), and those that are host species agnostic with environmental factors such as pH and nutrient availability explaining much of the community structure (Larmola et al., 2014). With a raised pH, hyaline cells may serve as 'oases' for microbes in acidic peatland pore waters. The ecological function of Sphagnum symbionts is just beginning to be explored, with evidence pointing to strong linkages with the cycling of both carbon (i.e. methane oxidation) and nitrogen (i.e. nitrogen fixation). For example, diazotrophic cyanobacteria were shown to contribute up to 35% of cellular nitrogen to the Sphagnum host (Berg et al., 2013; Lindo et al., 2013) while methanotrophic bacteria can provide 5-20% of Sphagnum's CO2 demand through methane oxidation (Raghoebarsing et al., 2005; Kip et al., 2010). Together, methanotrophy and N2 fixation are tightly linked and was estimated to provide over one-third of the new nitrogen input in a coastal peatland (Larmola et al., 2014), although see Ho & Bodelier (2015). Therefore, a number of critical questions concerning the Sphagnum microbiome remain, for example what are the signaling and communication pathways between Sphagnum and its microbiome, and do these interactions represent true beneficial symbioses. How do protists and miroeukayotes influence peatland carbon and nitrogen cycles (Jassey et al., 2015)? More questions than answers remain, and achieving a comprehensive understanding of the Sphagnum microbiome will benefit greatly from the application of comparative and functional genomics to evaluate microbial community profiles across Sphagnum lineages and environments, and metatranscriptomics to evaluate symbiotic pathways and metabolism.

### How do we model *Sphagnum* genotype-by-environment interactions?

The understanding of Sphagnum trait characteristics and the population genetics underlying trait distributions may have important implications for modeling biogeochemistry and vegetation dynamics, both within an ecosystem and across regions up to a global scale. However, the Sphagnum trait characterization needed to inform these models is lacking for many high-latitude process-based models (Turetsky et al., 2012). Many ecosystem and regional models have adopted the concept of plant functional types (PFTs), where PFTs are defined as groupings of plant species that share similar characteristics and roles in ecosystem function. However, recent work suggests that parameterization of PFTs with current trait values may not be valid under future environmental conditions because trait values and trait-trait relationships may change under future environmental conditions (van Bodegom

et al., 2012; Scheiter et al., 2013). In this regard, we will benefit from population genomics programs – like the Sphagnome Project - where population genetics, genomics and phenotype analysis can be used to statistically model genome features (such as single nucleotide polymorphism (SNP) distributions) to trait value predictions. The 'trait values' are then entered as parameter values in physiological models. An elegant example of this approach was presented by Reuning et al. (2015), where quantitative trait locus (QTL) analysis was used to genetically parameterize a physiological model to predict transpiration of specific Arabidopsis genotypes. An intriguing question is whether such 'genome informed' ecophysiological models can be used to decipher the mechanisms of local adaptation, which provides deeper insights into heritable variation and trait covariances (and trade-offs) responsible for evolutionary dynamics (Weinig et al., 2014).

#### **Conclusions**

The Sphagnome Project seeks to resolve important and general issues in ecology and evolution including: (1) the niche differentiation and co-occurrence of many closely related Sphagnum species within the same wetland habitat; (2) the genetic regulation of the unique chemical traits that define the central role of Sphagnum species in engineering those habitats; (3) the importance of Sphagnum in determining biodiversity patterns of other organisms, including microbes; and (4) the role of Sphagnum genetics and physiology on biogeochemistry and hydrology at ecosystem to global scales. With new genomic resources already available and growing rapidly, we are poised to utilize the Sphagnum system for linking genomes and phenotypic traits to community assembly, ecosystem function, and evolutionary processes. Moreover, the Sphagnum system can provide unique insights into the phylogenetic history of genome and trait evolution, and allow predictions about how these organismal features are likely to respond to future environmental change.

## **Acknowledgements**

The authors thank Drs Stan Wullschleger, Paul Hanson and two anonymous reviewers for comments on the manuscript. Work related to sequencing efforts are supported by the US Department of Energy (DOE) Joint Genome Institute by the Office of Science under contract no. DE-AC02-05CH11231; and germplasm establishment and maintenance is supported by the US DOE, Office of Science, Office of Biological and Environmental Research, Early Career Research Program. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the US DOE under contract no. DE-AC05-00OR22725. The authors thank the National Evolutionary Synthesis Center (NESCent), NSF #EF-0905606 and the New Phytologist Trust for sponsoring workshops on the Sphagnome Project.

#### **Author contributions**

D.J.W., A.J.S. and M.R.T. conceived the Sphagnome project and solicited community input; D.J.W., A.J.S., M.R.T., M.G.J. and G.G. wrote the paper; Z.L., L.R.B., S.K.R., D.T.H., K.A.M.E., E.D., R.J.N., J.E.K., J.B.G., H.R., J.L., E-S.T., A.C., B.W.B., E.S.E., T.A.O., M.B.N., E.A.L. and R.S.C. conceived of and contributed to the ecological, physiology and modeling section; H.K.S., P.S., M.J. and B.T.P. developed the evolutionary genetic sections; J.S., W.M., K.K.U., J-G.C., P.R. and D.J. contributed to the bioinformatics and quantitative genetics.

David J. Weston<sup>1,2</sup>\*, Merritt R. Turetsky<sup>3</sup>, Matthew G. Johnson<sup>4</sup>, Gustaf Granath<sup>5</sup>, Zoë Lindo<sup>6</sup>, Lisa R. Belyea<sup>7</sup>, Steven K. Rice<sup>8</sup>, David T. Hanson<sup>9</sup>, Katharina A. M. Engelhardt<sup>10</sup>, Jeremy Schmutz<sup>11,12</sup>, Ellen Dorrepaal<sup>13</sup>, Eugénie S. Euskirchen<sup>14</sup>, Hans K. Stenøien<sup>15</sup>, Péter Szövényi<sup>16</sup>, Michelle Jackson<sup>17</sup>, Bryan T. Piatkowski<sup>17</sup> Wellington Muchero<sup>1</sup>, Richard J. Norby<sup>2,18</sup>, Joel E. Kostka<sup>19</sup> Jennifer B. Glass<sup>19</sup>, Håkan Rydin<sup>20</sup>, Juul Limpens<sup>21</sup> Eeva-Stiina Tuittila<sup>22</sup>, Kristian K. Ullrich<sup>23</sup>, Alyssa Carrell<sup>1</sup> Brian W. Benscoter<sup>24</sup>, Jin-Gui Chen<sup>1</sup>, Tobi A. Oke<sup>3</sup>, Mats B. Nilsson<sup>25</sup>, Priya Ranjan<sup>26</sup>, Daniel Jacobson<sup>1</sup> Erik A. Lilleskov<sup>27</sup>, R. S. Clymo<sup>28</sup> and A. Jonathan Shaw<sup>17</sup>

<sup>1</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA; <sup>2</sup>Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA; <sup>3</sup>Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada; <sup>4</sup>Department of Biological Sciences, Texas Tech University, Lubbock, TX 79414, USA; <sup>5</sup>Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, SE-750 07, Uppsala, Sweden; <sup>6</sup>Department of Biology, The University of Western Ontario, London, ON N6A 5B7, Canada; <sup>7</sup>School of Geography, Queen Mary University of London, London, E1 4NS, UK; <sup>8</sup>Department of Biological Sciences, Union College, Schenectady, NY 12308, USA; <sup>9</sup>Department of Biology, University of New Mexico, Albuquerque, <sup>10</sup>Appalachian Lab, University of Maryland Center of Environ-

NM 87131, USA;

mental Science, Frostburg, MD 21532, USA; <sup>11</sup>HudsonAlpha Institute of Biotechnology, Huntsville, AL 35806, USA;

<sup>12</sup>Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA;

<sup>13</sup>Climate Impacts Research Center, Department of Ecology and Environmental Science, Umeå University,

98107, Abisko, Sweden;

<sup>14</sup>Institute of Arctic Biology, University of Alaska, Fairbanks, AK 99775, USA;

<sup>15</sup>NTNU University Museum, Norwegian University of Science and Technology, NO-7491, Trondheim, Norway; <sup>16</sup>Department of Systematic and Evolutionary Botany, University of Zurich, 8008, Zurich, Switzerland;

> <sup>17</sup>Department of Biology, Duke University, Durham, NC 27708, USA;

<sup>18</sup>Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA; <sup>19</sup>Schools of Biology and Earth & Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA; <sup>20</sup>Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, SE-75236, Uppsala, Sweden; <sup>21</sup>Plant Ecology and Nature Conservation Group, Department of Environmental Sciences, Wageningen University, Droevendaalse steeg 3a, NL-6708 PD, Wageningen, the Netherlands; <sup>22</sup>Peatland and Soil Ecology Group, School of Forest Sciences, University of Eastern Finland, Joensuu, Finland; <sup>23</sup>Max-Planck Institute for Evolutionary Biology, 24306, Plön, Germany; <sup>24</sup>Department of Biological Sciences, Florida Atlantic University, Davie, FL 33314, USA;

<sup>25</sup>Department of Forest Ecology and Management, Swedish

University of Agricultural Sciences, Skogsmarksgränd, SE-901 83, Umeå, Sweden;

<sup>26</sup>Department of Plant Sciences, University of Tennessee, 2431 Joe Johnson Drive, Knoxville, TN 37996-4561, USA; <sup>27</sup>US Forest Service, Northern Research Station, 410 MacInnes Dr., Houghton, MI 49931, USA; <sup>28</sup>School of Biological & Chemical Sciences, Queen Mary University of London, London, E1 4NS, UK (\*Author for correspondence: tel +1 865 241 8323;

email westondj@ornl.gov)

#### References

- Beike AK, Spagnuolo V, Luth V, Steinhart F, Ramos-Gomez J, Krebs M, Adamo P, Rey-Asensio AI, Fernandez JA, Giordano S et al. 2015. Clonal in vitro propagation of peat mosses (Sphagnum L.) as novel green resources for basic and applied research. Plant Cell Tissue and Organ Culture 120: 1037-1049.
- Bengtsson F, Granath G, Rydin H. 2016. Photosynthesis, growth, and decay traits in Sphagnum – a multispecies comparison. Ecology and Evolution 6: 3325–3341.
- Berg A, Danielsson A, Svensson BH. 2013. Transfer of fixed-N from N2-fixing cyanobacteria associated with the moss Sphagnum riparium results in enhanced growth of the moss. Plant and Soil 362: 271-278.
- Blomberg SP, Garland T. 2002. Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. Journal of Evolutionary Biology 15:
- van Bodegom PM, Douma JC, Witte JPM, Ordonez JC, Bartholomeus RP, Aerts R. 2012. Going beyond limitations of plant functional types when predicting global ecosystem-atmosphere fluxes: exploring the merits of traits-based approaches. Global Ecology and Biogeography 21: 625-636.
- Bragazza L, Freeman C, Jones T, Rydin H, Limpens J, Fenner N, Ellis T, Gerdol R, Hajek M, Hajek T et al. 2006. Atmospheric nitrogen deposition promotes carbon loss from peat bogs. Proceedings of the National Academy of Sciences, USA 103: 19386-19389.
- Bragina A, Berg C, Cardinale M, Shcherbakov A, Chebotar V, Berg G. 2012. Sphagnum mosses harbour highly specific bacterial diversity during their whole lifecycle. ISME Journal 6: 802-813.
- van Breemen N. 1995. How Sphagnum bogs down other plants. Trends in Ecology & Evolution 10: 270-275.
- Cardona-Correa C, Piotrowski MJ, Knack JJ, Kodner RE, Geary DH, Graham LE. 2016. Peat moss-like vegetative remains from Ordovician carbonates. International Journal of Plant Sciences 177: 523-538.
- Chitwood DH, Topp CN. 2015. Revealing plant cryptotypes: defining meaningful phenotypes among infinite traits. Current Opinion in Plant Biology 24: 54-60.
- Clymo RS. 1983. Peat. In: Gore AJP, ed. Ecosystems of the World, vol. 4A. Mires: swamp, bog, fen and moor. New York, NY, USA: Elsevier, 159-224.

- Clymo RS, Hayward PM. 1982. The ecology of Sphagnum. New York, NY, USA: Chapman and Hall.
- Cronberg N. 1996. Clonal structure and fertility in a sympatric population of the peat mosses, Sphagnum rubellum and S. capillifolium. Canadian Journal of Botany 74: 1375-1385.
- Crum HA. 1984. North American flora, series II, part 11, Sphagnopsida, Sphagnaceae. New York, NY, USA: New York Botanical Garden.
- Devos N, Szovenyi P, Weston DJ, Rothfels CJ, Johnson MG, Shaw AJ. 2016. Analyses of transcriptome sequences reveal multiple ancient large-scale duplication events in the ancestor of Sphagnopsida (Bryophyta). New Phytologist 211: 300-318.
- Freeman C, Ostle N, Kang H. 2001. An enzymic 'latch' on a global carbon store. Nature 409: 149.
- Greb SF, DiMichele WA, Gastaldo RA. 2006. Evolution and importance of wetlands in Earth history. In: Greb SF, DiMichele WA, eds. Wetlands through time. McLean, VA, USA: Geological Society of America Special Paper 399, 1-40.
- Hájek T. 2009. Habitat and species controls on Sphagnum production and decomposition in a mountain raised bog. Boreal Environment Research 14: 947-
- Hájek T, Ballance S, Limpens J, Zijlstra M, Verhoeven JTA. 2011. Cell-wall polysaccharides play an important role in decay resistance of Sphagnum and actively depressed decomposition in vitro. Biogeochemistry 103: 45-57.
- Haudry A, Platts AE, Vello E, Hoen DR, Leclercq M, Williamson RJ, Forczek E, Joly-Lopez Z, Steffen JG, Hazzouri KM et al. 2013. An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. Nature Genetics 45: 891-U228.
- Hayward PM, Clymo RS. 1983. The growth of Sphagnum: experiments on, and simulation of, some effects of light flux and water-table depth. Journal of Ecology 71: 845-863.
- Ho A, Bodelier PLE. 2015. Diazotrophic methanotrophs in peatlands: the missing link? Plant and Soil 389: 419-423.
- Ingram HAP. 1978. Soil layers in mires: function and terminology. Journal of Soil Science 29: 224-227.
- Jansson R, Rodriguez-Castaneda G, Harding LE. 2013. What can multiple phylogenies say about the latitudinal diversity gradient? A new look at the tropical conservatism, out of the tropics, and diversification rate hypotheses. Evolution 67:
- Jassey VE, Signarbieux C, Hättenschwiler S, Bragazza L, Buttler A, Delarue F, Fournier B, Gilbert D, Laggoun-Défarge F, Lara E et al. 2015. An unexpected role for mixotrophs in the response of peatland carbon cycling to climate warming. Scientific Reports 5: 16931.
- Johnson MG, Granath G, Tahvanainen T, Pouliot R, Stenoien HK, Rochefort L, Rydin H, Shaw AJ. 2015. Evolution of niche preference in Sphagnum peat mosses. Evolution 69: 90-103.
- Jones CG, John HL, Moshe S. 1994. Organisms as ecosystem engineers. In: Samson FB, Knopf FL, eds. *Ecosystem management*. New York, NY, USA: Springer, 130–147.
- Karlin EF, Andrus RE, Boles SB, Shaw AJ. 2011. One haploid parent contributes 100% of the gene pool for a widespread species in northwest North America. Molecular Ecology 20: 753-767.
- Kip N, van Winden JF, Pan Y, Bodrossy L, Reichart GJ, Smolders AJP, Jetten MSM, Damste JSS, Op den Camp HJM. 2010. Global prevalence of methane oxidation by symbiotic bacteria in peat-moss ecosystems. Nature Geoscience 3:
- Kostka JE, Weston DJ, Glass JB, Lilleskov EA, Shaw AJ, Turetsky MR. 2016. The Sphagnum microbiome: new insights from an ancient plant lineage. New Phytologist 211: 57-64.
- Kyrkjeeide MO, Hassel K, Flatberg KI, Shaw AJ, Brochmann C, Stenoien HK. 2016. Long-distance dispersal and barriers shape genetic structure of peatmosses (Sphagnum) across the Northern Hemisphere. Journal of Biogeography 43: 1215-1226.
- Larmola T, Leppänen SM, Tuittila ES, Aarva M, Merilä P, Fritze H, Tiirola M. 2014. Methanotrophy induces nitrogen fixation during peatland development. Proceedings of the National Academy of Sciences, USA 111: 734–739.
- Limpens J, Bohlin E, Nilsson MB. 2017. Phylogenetic or environmental control on the elemental and organo-chemical composition of Sphagnum mosses? Plant and Soil 417: 69-85.

- Lindo Z, Nilsson MC, Gundale MJ. 2013. Bryophyte–cyanobacteria associations as regulators of the northern latitude carbon balance in response to global change. *Global Change Biology* 19: 2022–2035.
- Merilä J, Hendry AP. 2014. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications* 7: 1–14.
- Miner BG, Sultan SE, Morgan SG, Padilla DK, Relyea RA. 2005. Ecological consequences of phenotypic plasticity. *Trends in Ecology & Evolution* 20: 685–692.
- Novikova PY, Hohmann N, Nizhynska V, Tsuchimatsu T, Ali J, Muir G, Guggisberg A, Paape T, Schmid K, Fedorenko OM et al. 2016. Sequencing of the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. *Nature Genetics* 48: 1077–1082.
- Pease JB, Haak DC, Hahn MW, Moyle LC. 2016. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. *PLoS Biology* 14: e1002379.
- Raghoebarsing AA, Smolders AJP, Schmid MC, Rijpstra WIC, Wolters-Arts M, Derksen J, Jetten MSM, Schouten S, Damste JSS, Lamers LPM et al. 2005. Methanotrophic symbionts provide carbon for photosynthesis in peat bogs. Nature 436: 1153–1156.
- Reuning GA, Bauerle WL, Mullen JL, McKay JK. 2015. Combining quantitative trait loci analysis with physiological models to predict genotype-specific transpiration rates. *Plant, Cell & Environment* 38: 710–717.
- Revell LJ. 2009. Size-correction and principal components for interspecific comparative studies. *Evolution* 63: 3258–3268.
- Rice SK, Giles L. 1996. The influence of water content and leaf anatomy on carbon isotope discrimination and photosynthesis in *Sphagnum. Plant, Cell & Environment* 19: 118–124.
- Rockstrom J, Steffen W, Noone K, Persson A, Chapin FS, Lambin EF, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ et al. 2009. A safe operating space for humanity. Nature 461: 472–475.
- Rydin H, Jeglum JK. 2013. The biology of peatlands, 2<sup>nd</sup> edn. Oxford, UK: Oxford University Press.
- Scheiter S, Langan L, Higgins SI. 2013. Next-generation dynamic global vegetation models: learning from community ecology. New Phytologist 198: 957–969.
- Shaw AJ, Devos N, Cox CJ, Boles SB, Shaw B, Buchanan AM, Cave L, Seppelt R. 2010. Peatmoss (Sphagnum) diversification associated with Miocene Northern Hemisphere climatic cooling? Molecular Phylogenetics and Evolution 55: 1139–1145.
- Shaw AJ, Devos N, Liu Y, Cox CJ, Goffinet B, Flatberg KI, Shaw B. 2016a.
  Organellar phylogenomics of an emerging model system: *Sphagnum* (peatmoss).
  Annals of Botany 118: 185–196.

- Shaw AJ, Schmutz J, Devos N, Shu S, Carrell AA, Weston DJ. 2016b. The Sphagnum Genome Project: a new model for ecological and evolutionary genomics. Genomes and evolution of charophytes, bryophytes, lycophytes and ferns. In: Rensing SA, ed. *Advances in Botanical Research* 78: 167–187.
- Stalheim T, Ballance S, Christensen BE, Granum PE. 2009. Sphagnan—a pectin-like polymer isolated from Sphagnum moss can inhibit the growth of some typical food spoilage and food poisoning bacteria by lowering the pH. Journal of Applied Microbiology 106: 967–976.
- Stevenson SR, Kamisugi Y, Trinh CH, Schmutz J, Jenkins JW, Grimwood J, Wellington M, Tuskan GA, Rensing SA, Lang D et al. 2016. Genetic analysis of Physcomitrella patens identifies ABSCISIC ACID NON-RESPONSIVE, a regulator of ABA responses unique to basal land plants and required for desiccation tolerance. Plant Cell 28: 1310–1327.
- Turetsky MR, Bond-Lamberty B, Euskirchen E, Talbot J, Frolking S, McGuire AD, Tuittila ES. 2012. The resilience and functional role of moss in boreal and arctic ecosystems. *New Phytologist* 196: 49–67.
- Turetsky MR, Crow SE, Evans RJ, Vitt DH, Wieder RK. 2008. Trade-offs in resource allocation among moss species control decomposition in boreal peatlands. *Journal of Ecology* 96: 1297–1305.
- Vitt DH, Slack NG. 1984. Niche diversification of Sphagnum relative to environmental factors in northern Minnesota peatlands. Canadian Journal of Botany–Revue Canadienne De Botanique 62: 1409–1430.
- Weinig C, Ewers BE, Welch SM. 2014. Ecological genomics and process modeling of local adaptation to climate. *Current Opinion in Plant Biology* 18: 66–72.
- Weston DJ, Timm CM, Walker AP, Gu LH, Muchero W, Schmutz J, Shaw AJ, Tuskan GA, Warren JM, Wullschleger SD. 2015. *Sphagnum* physiology in the context of changing climate: emergent influences of genomics, modelling and host–microbiome interactions on understanding ecosystem function. *Plant, Cell & Environment* 38: 1737–1751.
- Whitham TG, Bailey JK, Schweitzer JA, Shuster SM, Bangert RK, Leroy CJ, Lonsdorf EV, Allan GJ, DiFazio SP, Potts BM et al. 2006. A framework for community and ecosystem genetics: from genes to ecosystems. Nature Reviews Genetics 7: 510–523.
- Whitham TG, Young WP, Martinsen GD, Gehring CA, Schweitzer JA, Shuster SM, Wimp GM, Fischer DG, Bailey JK, Lindroth RL et al. 2003. Community and ecosystem genetics: a consequence of the extended phenotype. Ecology 84: 559–573.
- Yu ZC, Loisel J, Brosseau DP, Beilman DW, Hunt SJ. 2010. Global peatland dynamics since the Last Glacial Maximum. Geophysical Research Letters 37: L13402.

**Key words:** ecological genomics, ecosystem engineering, evolutionary genetics, genome sequencing, niche construction, peatlands, Sphagnome, *Sphagnum*.