

# 1     **The Sphagnome Project: enabling ecological and evolutionary** 2                    **insights through a genus-level sequencing project**

## 3 4     **Authors:**

5     <sup>1,16</sup>David J. Weston, westondj@ornl.gov, orcid.org/0000-0002-4794-9913

6     <sup>2</sup>Merritt R. Turetsky, mrt@uoguelph.ca

7     <sup>3</sup>Matthew G. Johnson, matt.johnson@ttu.edu, orcid.org/0000-0002-1958-6334

8     <sup>4</sup>Gustaf Granath, Gustaf.Granath@gmail.com, orcid.org/0000-0002-3632-9102

9     <sup>5</sup>Zoë Lindo, zlindo@uwo.ca, orcid.org/0000-0001-9942-7204

10    <sup>6</sup>Lisa R. Belyea, l.belyea@qmul.ac.uk, orcid.org/0000-0002-9964-2913

11    <sup>7</sup>Steven K. Rice, rices@union.edu

12    <sup>8</sup>David T. Hanson, dthanson@unm.edu

13    <sup>9</sup>Katharina A. M. Engelhardt, kengelhardt@al.umces.edu, orcid.org/0000-0002-9185-4292

14    <sup>10</sup>Jeremy Schmutz, jschmutz@hudsonalpha.org

15    <sup>11</sup>Ellen Dorrepaal, ellen.dorrepaal@umu.se

16    <sup>12</sup>Eugénie S. Euskirchen, seeuskirchen@alaska.edu

17    <sup>13</sup>Hans K. Stenøien, stenoiens@ntnu.no, orcid.org/0000-0003-2191-332X

18    <sup>14</sup>Péter Szövényi, peter.szovenyi@systbot.uzh.ch

19    <sup>15</sup>Michelle Jackson, michelle.jackson@duke.edu

20    <sup>15</sup>Bryan T. Piatkowski, bryan.piatkowski@duke.edu, orcid.org/0000-0002-1334-8431

21    <sup>1</sup>Wellington, Muchero, mucherow@ornl.gov

22    <sup>16,17</sup>Richard J. Norby, norbyrj@ornl.gov, orcid.org/0000-0002-0238-9828

23    <sup>18</sup>Joel E. Kostka, joel.kostka@biology.gatech.edu

24    <sup>18</sup>Jennifer B. Glass, jennifer.glass@eas.gatech.edu, orcid.org/0000-0003-0775-2486

25    <sup>19</sup>Håkan Rydin, hakan.rydin@ebc.uu.se, orcid.org/0000-0002-7582-3998

26    <sup>20</sup>Juul Limpens, juul.limpens@wur.nl, orcid.org/0000-0001-5779-0304

27    <sup>21</sup>Eeva-Stiina Tuittila, eeva-stiina.tuittila@uef.fi

28    <sup>22</sup>Kristian K. Ullrich, ullrich@evolbio.mpg.de, orcid.org/0000-0003-4308-9626

29    <sup>1</sup>Alyssa Carrell, alyssacarrell@gmail.com

30    <sup>23</sup>Brian W. Benschoter, brian.benschoter@fau.edu, orcid.org/0000-0002-2706-4667

31    <sup>1</sup>Jin-Gui Chen, chenj@ornl.gov orcid.org/0000-0002-1752-4201

32    <sup>2</sup>Tobi A. Oke, ooke@uoguelph.ca, orcid.org/0000-0002-7290-2754

33    <sup>24</sup>Mats B. Nilsson, mats.b.nilsson@slu.se, orcid.org/0000-0003-3765-6399

34 <sup>25</sup>Priya Ranjan, pranjan@utk.edu, orcid.org/0000-0002-0357-1939

35 <sup>1</sup>Daniel Jacobson, jacobsonda@ornl.gov, orcid.org/0000-0002-9822-8251

36 <sup>26</sup>Erik A. Lilleskov, elilleskov@fs.fed.us, orcid/0000-0002-9208-1631

37 <sup>27</sup>R. S. Clymo, R.Clymo@QMUL.ac.uk

38 <sup>15</sup>A. Jonathan Shaw, shaw@duke.edu

39

#### 40 **Author Affiliations:**

41 <sup>1</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA

42 <sup>2</sup>Department of Integrative Biology, University of Guelph, Guelph, ON, Canada

43 <sup>3</sup>Department of Biological Sciences, Texas Tech University, Lubbock, TX 79414 USA

44 <sup>4</sup>Department of Ecology, Swedish University of Agricultural Sciences Box 7044, SE-750 07  
45 Uppsala, Sweden.

46 <sup>5</sup>The University of Western Ontario, Department of Biology, London, Ontario, N6A 5B7, Canada

47 <sup>6</sup>School of Geography, Queen Mary University of London, London, UK

48 <sup>7</sup>Department of Biological Sciences, Union College, Schenectady, NY 12308, USA

49 <sup>8</sup>Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA

50 <sup>9</sup>University of Maryland Center of Environmental Science, Appalachian Lab, Frostburg, MD  
51 21532 USA

52 <sup>10</sup>HudsonAlpha Institute of Biotechnology, Huntsville, AL 35806, USA

53 <sup>11</sup>Climate Impacts Research Centre, Department of Ecology and Environmental Science, Umeå  
54 University, 98107 Abisko, Sweden

55 <sup>12</sup>Institute of Arctic Biology, University of Alaska, Fairbanks, AK 99775, USA

56 <sup>13</sup> NTNU University Museum, Norwegian University of Science and Technology, NO-7491  
57 Trondheim, Norway

58 <sup>14</sup> Department of Systematic and Evolutionary Botany, University of Zurich, Switzerland

59 <sup>15</sup>Department of Biology, Duke University, Durham, NC, USA

60 <sup>16</sup>Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, 37831,  
61 USA

62 <sup>17</sup>Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831,  
63 USA

64 <sup>18</sup>Schools of Biology and Earth & Atmospheric Sciences, Georgia Institute of Technology,  
65 Atlanta, GA 30332, USA

66 <sup>19</sup>Department of Ecology and Genetics, Uppsala University, Norbyvägen 18D, SE-75236  
67 Uppsala, Sweden

68 <sup>20</sup> Plant Ecology and Nature conservation group, Department of Environmental Sciences,  
69 Wageningen University, Droevendaalse steeg 3a, NL-6708 PD, Wageningen, the Netherlands  
70 <sup>21</sup>Peatland and Soil Ecology Group, School of Forest Sciences, University of Eastern Finland,  
71 Joensuu Campus, Finland  
72 <sup>22</sup>Max-Planck Institute for Evolutionary Biology, Plön, Germany  
73 <sup>23</sup>Department of Biological Sciences, Florida Atlantic University, Davie, FL 33314, USA  
74 <sup>24</sup>Department of Forest Ecology and Management, Swedish University of Agricultural Sciences,  
75 Skogsmarksgränd, SE-901 83 Umeå, Sweden  
76 <sup>25</sup>Department of Plant Sciences, University of Tennessee, Knoxville, TN, USA  
77 <sup>26</sup>U.S. Forest Service, Northern Research Station, 410 MacInnes Dr., Houghton, MI 49931  
78 <sup>27</sup>School of Biological & Chemical Sciences, Queen Mary University of London, London E1  
79 4NS, UK

80  
81 This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-  
82 00OR22725 with the U.S. Department of Energy. The United States Government retains and the  
83 publisher, by accepting the article for publication, acknowledges that the United States  
84 Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or  
85 reproduce the published form of this manuscript, or allow others to do so, for United States  
86 Government purposes. The Department of Energy will provide public access to these results of  
87 federally sponsored research in accordance with the DOE Public Access Plan  
88 (<http://energy.gov/downloads/doe-public-access-plan>).

89  
90 Corresponding author:  
91 David J. Weston  
92 Climate Change Science Institute & Biosciences Division  
93 Oak Ridge National Laboratory  
94 Oak Ridge, TN 37831-6301  
95 Phone: (865) 241-8323  
96 Email: westondj@ornl.gov

97  
98 Word count – Summary 137; Introduction 467; Main Text 3,711; Conclusion 149;  
99 Acknowledgement 106; Total excluding references 3,570  
100 Number of references - 64  
101 Colour Figures – 4; Tables - 0

## 102 **Summary**

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

116

117 **Keywords:** ecological genomics, ecosystem engineering, evolutionary genetics,  
118 genome sequencing, genomics, niche construction, peatlands, Sphagnome, *Sphagnum*

119

## 120 **I. Introduction**

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

133           Here we argue that the genus *Sphagnum* (peat moss) represents an unparalleled -  
134 model system for ecological and evolutionary genomics, empowered by its contribution  
135 to global carbon cycling and emerging genomic resources. *Sphagnum* species play a  
136 major role in peatland formation, a prime example of ecosystem engineering, whereby  
137 the organism manipulates its surrounding habitat. *Sphagnum* primary production  
138 influences carbon and nutrient cycling, such as methane production and soil carbon  
139 storage, in many boreal forests and peatlands (Turetsky *et al.*, 2012). *Sphagnum*  
140 ecosystem engineering involves the accumulation of peat that facilitates its own growth  
141 while making the surrounding environment hostile for vascular plants (van Breemen,  
142 1995). Ultimately these multi-level processes lead to peatland formation that occupy  
143 nearly 3% of the land surface and store 25% of the world’s soil carbon as recalcitrant peat  
144 (Yu *et al.*, 2010). The latter point has led to the assertion that *Sphagnum* has a greater  
145 impact on global carbon fluxes, and therefore climate, than any other single genus of  
146 plants (Clymo & Hayward, 1982; van Breemen, 1995).

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

160

## 161 **II *Sphagnum* ecology and evolution**

162

### 163 **1. Functional traits and ecosystem function**

164

165 *Sphagnum* has a remarkable ability to create and then uniquely thrive in nutrient-poor,  
166 acidic, and waterlogged conditions. The suite of morphological, physiological, and life  
167 history traits that affect *Sphagnum* fitness, herein termed functional traits, enable this  
168 ‘ecosystem engineer’ (Jones *et al.*, 1994) to gain a competitive advantage over other co-  
169 occurring species and therefore flourish under relatively harsh environmental conditions.  
170 For example, the ability of *Sphagnum* to store and transport water is controlled largely by  
171 three distinct morphological adaptations – branching architecture, leaf size and  
172 arrangement on branches, and hyaline cells (Fig. 1a,b; Rydin & Jeglum, 2013). These  
173 traits differ considerably among species, and are associated with highly partitioned  
174 microhabitat preferences where *Sphagnum* species coexist within a peatland. Hummock-  
175 forming species, growing ca. > 30 cm above the water table, have small close-set leaves  
176 forming numerous interconnected small capillary spaces (Fig. 1). Spreading branches  
177 allow lateral movement of water through the capillary continuum, while numerous close-  
178 set pendant branches appressed to the stem form an efficient vertical water-transport  
179 system. Consequently, *Sphagnum* species growing on hummocks can wick moisture and  
180 maintain metabolic activity even during drought (Rice & Giles, 1996). In all species,  
181 dead hyaline cells in the leaves and the outer cortex of the stems and branches act as  
182 water-storage structures.

183         The capitula at the top of the stem are alive, but a few (~5) cm down 99 % of the  
184 light has been absorbed and most of the *Sphagnum* cells die (Hayward & Clymo, 1983).  
185 From there down to the water table the carpet structure is permeable to water and gases  
186 (particularly O<sub>2</sub>) and the damp plant substrates begin to decay in this oxic zone, termed  
187 the acrotelm (Ingram, 1978; Clymo & Hayward 1982). The consequent loss of stem  
188 strength and increasing weight eventually result in collapse of the plant structure. This  
189 reduces the pore size so water can no longer flow easily through it, and from this point  
190 downwards the peat is permanently waterlogged and this is what determines the depth of  
191 the water table. In this waterlogged zone, oxygen is consumed by aerobic respiration  
192 more rapidly than it can be replenished by diffusion (which is 10,000 times slower in  
193 water than it is in air), creating the anoxic catotelm (Clymo, 1983). Hence, through  
194 distinct traits, *Sphagnum* generates environmental conditions that are suitable for its own

195 growth but hostile for the vast majority of other plants (e.g., van Breemen, 1995; Rydin  
196 & Jeglum, 2013).

197 The mechanisms by which *Sphagnum* inhibits fungal and microbial  
198 decomposition -- and hence promotes peat accumulation -- are not fully understood, but  
199 involve both the external environment engineered by the species, as well as the internal  
200 biochemistry of its plant tissue, particularly the low N:C ratio (a reflection of the  
201 unusually efficient use of N in producing new biomass) (Bragazza *et al.*, 2006). A  
202 passive mechanism for intrinsic decay resistance in the oxic acrotelm layer is suggested  
203 by the correlation of microbial decomposition of *Sphagnum* litter with the relative  
204 amounts of structural versus metabolic carbohydrates (Turetsky *et al.*, 2008). Active  
205 mechanisms of antimicrobial activity are also implicated, mainly through acid hydrolysis  
206 of cell-wall polysaccharides, fragments of which are released into the soil water as  
207 ‘sphagnum’ (Hájek *et al.*, 2011). The precise mechanisms for the antimicrobial activity of  
208 sphagnum are still under investigation, but may involve lowering soil pH, reducing  
209 availability of nitrogen and carbon, or interfering with extracellular enzymes by  
210 immobilizing them in a polyelectrolyte complex (Hájek *et al.*, 2011). Soluble phenolic  
211 compounds, either leached directly from *Sphagnum* tissue or produced during its  
212 breakdown, may play a more minor role in tissue preservation, physically protecting  
213 polysaccharides through the formation of humic substances (Hájek *et al.*, 2011). While  
214 environmental factors such as soil oxygen profiles serve as important regulators of peat  
215 decomposition (cf Freeman *et al.*, 2001) it is clear that a variety of mechanisms  
216 contribute to slow decomposition of *Sphagnum* tissue, thereby retarding the turnover of  
217 organic biomass in peatlands and sequestering carbon in the form of peat for centuries.

218

## 219 **2. Phylogeny and evolution**

220

221 Like all mosses, the haploid gametophyte is the dominant life cycle stage for *Sphagnum*  
222 (Fig. 1). Haploid spores germinate into a filamentous protonema, quickly followed by a  
223 thalloid protonemal phase, before transitioning into mature haploid gametophytes. A  
224 single spore can result in a large clonal biomass through vegetative growth. Furthermore,  
225 the ability to propagate clonally is ubiquitous in *Sphagnum* and typical clone sizes vary

226 among species (Cronberg, 1991). In *S. austinii*, one clone occurs throughout North  
227 America and the same dominates in Europe (Kyrkjeeide *et al.*, 2016). A single clone of *S.*  
228 *subnitens* extends from Oregon to the westernmost Aleutian Islands (Karlin *et al.*, 2011).  
229 Reproductive seasons are species-specific and sperm require water to access the egg cell  
230 in the archegonial venter to form the zygote. The formation of the zygote marks the  
231 beginning of the brief diploid stage of development and at maturity meiosis occurs within  
232 the capsule, producing haploid spores.

233 *Sphagnum* is one of four genera in the class Sphagnopsida (phylum Bryophyta:  
234 mosses), an ancient lineage of land plants. Molecular phylogenies suggest the  
235 Sphagnopsida diverged from other mosses more than 250-350 mya (Shaw *et al.*, 2010),  
236 and fossils of peat moss-like fragment, which are the oldest known land plant  
237 macrofossils to date, have been found in the Ordovician rocks (~500 mya, Cardona-  
238 Correa *et al.*, 2016). Fossil *Sphagnum* and close relatives are recognized by the unique  
239 cell pattern in leaves. Three of the genera in the Sphagnopsida contain just one or two  
240 species each, and none of them form extensive peats nor do they dominate wetlands as do  
241 species of *Sphagnum*. With 200-300 species, *Sphagnum* is by far the largest genus in the  
242 Sphagnopsida and the most important for peatlands. *Sphagnum* species share a common  
243 ancestor in the late Tertiary, a surprisingly recent radiation considering the great antiquity  
244 of Sphagnopsida (Shaw *et al.*, 2010). This recent radiation, which may have occurred  
245 following the mid-Miocene climatic optimum, coincides with the rise of boreal peatlands  
246 in the northern hemisphere (Greb *et al.*, 2006).

247 Today, *Sphagnum* occurs on all continents aside from Antarctica (Crum, 1984).  
248 The genus dominates wetland habitats throughout the boreal zone of the Northern  
249 Hemisphere but is also diverse at tropical latitudes, especially in South America (as well  
250 as in tropical Africa and Asia). At tropical latitudes, *Sphagnum* sometimes occurs in high  
251 altitude peatlands, but in lower altitude tropical regions they typically grow on wet soil  
252 banks, along streams, and on dripping rocks, and do not accumulate substantial amounts  
253 of peat. *Sphagnum* comprises five major subgenera (Fig. 2a; Shaw *et al.*, 2016a). The  
254 small subgenus *Rigida* (ca. 2-4 species), sister to the four other subgenera, sometimes  
255 occur in peatlands, but its species are never dominant and are not major peat-formers.  
256 Most *Sphagnum* species belong to the remaining two clades, both of which include



257 important peat-forming species. The species in one clade (subgenera *Cuspidata* +  
258 *Subsecunda*) generally occupy hollows close to or at the water table, whereas those in the  
259 other clade (subgenera *Sphagnum* + *Acutifolia*) generally create lawns and raised  
260 hummocks more distant from the water table (Fig. 2b). For decades, peatland ecologists  
261 have noted that individual *Sphagnum* species have narrow realized niches along this  
262 hydrological gradient—from low hollow to high hummock (Vitt & Slack, 1984).  
263 *Sphagnum* species also exhibit narrow preferences along a chemical gradient, with some  
264 species preferring acidic ombrotrophic bogs and other species preferring fens with more  
265 neutral pH. Unlike preferences along the hydrological gradient, species preferences along  
266 the chemical gradient do not exhibit a strong phylogenetic signal (Johnson *et al.*, 2015).  
267 During the rapid radiation of modern *Sphagnum*, microhabitat preferences along the  
268 chemical gradient plausibly evolved simultaneously in unrelated groups, creating natural  
269 experiments with which the genetic basis of microhabitat preferences can be disentangled  
270 from phylogenetic history.

271

### 272 **III. Developing resources for a tractable *Sphagnum* model system with** 273 **evolutionary and ecological relevance**

274

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

284

#### 285 **1. *Sphagnum* pedigree sequencing and gene to trait mapping**

286

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

311           Recent advances in maintaining *Sphagnum* tissue cultures (Beike *et al.*, 2015)  
312 have improved the reliability of producing axenic cultures that produce *Sphagnum* plants  
313 that are morphologically similar to field-collected specimens. The Sphagnome Project  
314 encompasses a developing germplasm collection that includes culture material for all  
315 species being sequenced and a *S. fallax* haploid-sib pedigree. The low stature of  
316 *Sphagnum* and ease of establishing populations in trans-well culture plates that have

317 relatively small ‘bench top’ space requirements enable rapid phenotyping that is  
318 necessary for gene-to-trait studies (Fig. 3). Further, this germplasm collection can be used  
319 to test responses of *Sphagnum* genotypes to different environmental conditions. Because  
320 the complete genomes of these genotypes will already be known as a result of re-  
321 sequencing, genetic associations can be made as soon as phenotypic data are collected.  
322 Due to the small size of *Sphagnum* and other mosses, imaging-based phenotyping will be  
323 especially useful in this effort. Single images can capture data on hundreds of individuals,  
324 entire populations, and mixed communities, simultaneously aiding the linkage of genes to  
325 traits. The broader collection of gene to trait associations can be integrated in network  
326 models to form a systems biology view of the trait combinations and their correlations  
327 underlying phenotype expression and adaptation (Chitwood & Topp, 2015).

328

## 329 **2. A genus-wide approach**

330

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

346 Through this sequencing effort, gene-to-trait relationships of multiple species will  
347 be placed within a broader phylogenomic landscape thereby identifying evolutionary

348 patterns associated with microhabitat preferences and functional traits (Fig. 2b & Fig. 4).  
349 While a few recent studies have taken a genus-wide approach to genetic associations  
350 (e.g., Haudry *et al.*, 2013; Pease *et al.* 2016; Novikova *et al.*, 2016) the Sphagnum  
351 Project encompasses species that co-occupy and engineer the same ecosystem. We  
352 anticipate that these genus-wide sequences, phenotype data, and comparative gene-to-  
353 trait relationships will enable the detection of genes under purifying or positive selection  
354 as well as gene family evolution associated with major ecological and biogeographic  
355 shifts.

356

#### 357 **IV. Facilitating new ecological and evolutionary understanding**

358

##### 359 **1. What is the biological basis of unique *Sphagnum* traits or combinations of traits,** 360 **and how do these trait combinations extend beyond the organism?**

361

362 Tissue chemistry is a noted functional trait for *Sphagnum* (Clymo & Hayward, 1982).  
363 Polyuronic acids (cell-wall polysaccharides that form a pectin-like polymer) comprise 10-  
364 30% of *Sphagnum* dry mass. They have a high cation exchange capacity (CEC) initially  
365 satisfied with H<sup>+</sup>, which is rapidly exchanged for cations in rainwater, thus making the  
366 water around the plants acidic (Clymo & Hayward, 1982) and make cation nutrients  
367 unavailable to microbes and other plants (Stalheim *et al.*, 2009). However, the question  
368 of a possible link between unique organic compounds and niche engineering by  
369 *Sphagnum* remains a matter of active research (Hájek, 2009; Limpens *et al.*, 2017). It has  
370 long been speculated that living *Sphagnum* benefits from peat formed over time through  
371 the accumulation of dead *Sphagnum* biomass (van Breemen, 1995). Should this be  
372 viewed as one type of extended phenotype, where the phenotype of vertically  
373 accumulating peat (dead *Sphagnum* material) changes the function of living *Sphagnum* at  
374 the surface? *Sphagnum* plants clearly modify their environment in several important  
375 ways, but how this influences selection on future offspring and other recipient organisms  
376 is unknown. We believe that the *Sphagnum* genomic resource offers one of the best  
377 opportunities to explore these questions and ultimately identify the genetic basis for the  
378 traits responsible for ecosystem engineering in *Sphagnum*. For example, what is the

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

391

## 392 **2. Did adaptation to spatially or temporally varying climate variation spark**

### 393 *Sphagnum* species radiations?

394

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

405 Moreover, non-boreal radiations occurred in each of the four large subgenera of  
406 *Sphagnum*, providing phylogenetic patterns that can be used as replicated natural  
407 experiments to account for shared ancestry when investigating the genetic basis of  
408 adaptation and the evolution of functional traits associated with range expansions. In  
409 addition to these radiations, a few individual boreal *Sphagnum* species have extended

410 their ranges into tropical habitats, presumably more recently. Inter- and intraspecific  
411 comparative analyses can be harnessed to address several questions. What genes, gene  
412 families, and genomic regions underwent changes associated with range expansions from  
413 boreal to tropical climate zones? Are the same genomic features associated with intra-  
414 and with interspecific range changes across climate zones? Are the same or similar  
415 genomic changes associated with climate adaptation in different *Sphagnum* subgenera,  
416 associated with independent range changes? Clarifying functional trait and genomic  
417 changes associated with migrations into warmer climates can provide informative  
418 analogies to how *Sphagnum* mosses and, perhaps, other plants may respond to current  
419 climate warming.

420

### 421 **3. What are the factors that limit or facilitate local-scale adaptive evolution?**

422

423 There has been much interest regarding the importance of phenotypic plasticity relative to  
424 local adaptation in response to environmental heterogeneity, and how such responses can  
425 ultimately extend to influence ecosystem function (Miner *et al.*, 2005). The sequenced  
426 haploid-sib pedigree, coupled with phenotype screening will provide the resources  
427 necessary for quantitative genetics to determine the extent to which a phenotypic change  
428 has a quantitative genetic basis (Section III). Plasticity is inferred as the proportion of  
429 phenotypic variance not explained by genetics (Merilä *et al.*, 2014). The use of common  
430 gardens, especially when established among multiple environments with appropriate  
431 replication and controls, provides a powerful approach to disentangle genetic from plastic  
432 contributions to phenotype. The sequenced *Sphagnum* haploid-sib pedigree and emerging  
433 research community surrounding the Sphagnome Project make the establishment of  
434 common gardens with characterized genotypes a reality. Finally, the demonstration that  
435 allele frequency shifts occur confirms that evolution has occurred, with the challenge  
436 being the need to determine if changes in specific allele frequencies are relevant to the  
437 traits and phenomena being investigated. The sequencing of 15 *Sphagnum* species and  
438 nearly 200 progeny individuals provides an ideal system to determine shared and species-  
439 specific components of the collective genome and relationships that co-occur with  
440 phylogenetic signals. For example, does a gene family expansion coincide with the

441 lineage diversification to novel environments? Together with common garden  
442 experiments we will begin to address questions centering on the relative importance of  
443 local adaptation versus phenotypic plasticity in *Sphagnum* responses to environmental  
444 heterogeneity.

445

#### 446 **4. What is the role of *Sphagnum* and its interacting microbiome in ecosystem carbon** 447 **and nitrogen cycling?**

448

449 Hyaline cells not only play a vital function as water storage organs, but also create  
450 a novel and safe habitat for a diverse microflora spanning all domains of life (Fig. 1b;  
451 Bragina *et al.*, 2012; 2014; Kostka *et al.*, 2016). The *Sphagnum*-associated microbiome  
452 seems to be divided into two broad categories. Those that are host species specific, with  
453 specificity maintained across both the sporophyte and gametophyte generations (Bragina  
454 *et al.*, 2012), and those that are host species agnostic with environmental factors such as  
455 pH and nutrient availability explaining much of the community structure (Larmola *et al.*,  
456 2014). With a raised pH, hyaline cells may serve as ‘oases’ for microbes in acidic  
457 peatland pore waters. The ecological function of *Sphagnum* symbionts is just beginning  
458 to be explored, with evidence pointing to strong linkages with the cycling of both carbon  
459 (i.e., methane oxidation) and nitrogen (i.e., nitrogen fixation). For example, diazotrophic  
460 cyanobacteria were shown to contribute up to 35% of cellular N to the *Sphagnum* host  
461 (Berg *et al.*, 2013; Lindo *et al.*, 2013) while methanotrophic bacteria can provide 5–20%  
462 of *Sphagnum*’s CO<sub>2</sub> demand through CH<sub>4</sub> oxidation (Raghoebarsing *et al.*, 2005; Kip *et*  
463 *al.*, 2010). Together, methanotrophy and N<sub>2</sub> fixation are tightly linked and was estimated  
464 to provide over one-third of the new N input in a coastal peatland (Larmola *et al.* 2014),  
465 although see Ho & Bodelier (2015). Therefore, a number of critical questions concerning  
466 the *Sphagnum* microbiome remain, for example what are the signaling and  
467 communication pathways between *Sphagnum* and its microbiome, and do these  
468 interactions represent true beneficial symbioses. How do protists and microeukaryotes  
469 influence peatland C and N cycles (Jassey *et al.*, 2015)? More questions than answers  
470 remain, and achieving a comprehensive understanding of the *Sphagnum* microbiome will  
471 benefit greatly from the application of comparative and functional genomics to evaluate

472 microbial community profiles across *Sphagnum* lineages and environments, and meta-  
473 transcriptomics to evaluate symbiotic pathways and metabolism.

474

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

476

477 The understanding of *Sphagnum* trait characteristics and the population genetics  
478 underlying trait distributions may have important implications for modeling  
479 biogeochemistry and vegetation dynamics, both within an ecosystem and across regions  
480 up to a global scale. However, the *Sphagnum* trait characterization needed to inform these  
481 models is lacking for many high-latitude process-based models (Turetsky *et al.*, 2012).  
482 Many ecosystem and regional models have adopted the concept of plant functional types  
483 (PFTs), where PFTs are defined as groupings of plant species that share similar  
484 characteristics and roles in ecosystem function. However, recent work suggests that  
485 parameterization of PFTs with current trait values may not be valid under future  
486 environmental conditions because trait values and trait-trait relationships may change  
487 under future environmental conditions (Scheiter *et al.*, 2013, van Bodegom *et al.*, 2012).  
488 In this regard, we will benefit from population genomics programs – like the Sphagnome  
489 Project – where population genetics, genomics and phenotype analysis can be used to  
490 statistically model genome features (such as single nucleotide polymorphism (SNP)  
491 distributions) to trait value predictions. The ‘trait values’ are then entered as parameter  
492 values in physiological models. An elegant example of this approach was presented by  
493 Reuning *et al.* (2014), where QTL analysis was used to genetically parameterize a  
494 physiological model to predict transpiration of specific *Arabidopsis* genotypes. An  
495 intriguing question is whether such ‘genome informed’ ecophysiological models can be  
496 used to decipher the mechanisms of local adaptation, which provides deeper insights into  
497 heritable variation and trait covariances (and trade-offs) responsible for evolutionary  
498 dynamics (Weinig *et al.*, 2014).

499

## 500 **V. Conclusions**

501 The Sphagnome Project seeks to resolve important and general issues in ecology and  
502 evolution including (1) the niche differentiation and co-occurrence of many closely



503 related *Sphagnum* species within the same wetland habitat, (2) the genetic regulation of  
504 the unique chemical traits that define the central role of *Sphagnum* species in engineering  
505 those habitats, 3) the importance of *Sphagnum* in determining biodiversity patterns of  
506 other organisms, including microbes, and (4) The role of *Sphagnum* genetics and  
507 physiology on biogeochemistry and hydrology at ecosystem to global scales. With new  
508 genomic resources already available, and growing rapidly, we are poised to utilize the  
509 *Sphagnum* system for linking genomes and phenotypic traits to community assembly,  
510 ecosystem function, and evolutionary processes. Moreover, the *Sphagnum* system can  
511 provide unique insights into the phylogenetic history of genome and trait evolution, and  
512 allow predictions about how these organismal features are likely to respond to future  
513 environmental change.

514

#### 515 **Author Contributions:**

516 DJW, AJS, MRT conceived the Sphagnome project and solicited community input; DJW,  
517 AJS, MRT, MGJ, and GG wrote the paper; ZL, LRB, SKR, DTH, KAME, ED, ESK,  
518 RJN, JEK, JBG, HR, JL, EST, AC, BWB, TAO, MBN, EAL and RSC conceived of and  
519 contributed to the ecological, physiology and modeling section; HKS, PS, MJ, BTP,  
520 developed the evolutionary genetic sections; JS, WM KKU, JGC, PR, DJ, contributed to  
521 the bioinformatics and quantitative genetics.

522

#### 523 **References**

- 524 **Beike AK, Spagnuolo V, Luth V, Steinhart F, Ramos-Gomez J, Krebs M, Adamo P,**  
525 **Rey-Asensio AI, Fernandez JA, Giordano S, et al. 2015.** Clonal in vitro  
526 propagation of peat mosses (*Sphagnum* L.) as novel green resources for basic and  
527 applied research. *Plant Cell Tissue and Organ Culture* **120**: 1037-1049.
- 528 **Belyea LR, Clymo RS. 2001.** Feedback control of the rate of peat formation.  
529 *Proceedings of the Royal Society B-Biological Sciences* **268**: 1315-1321.
- 530 **Belyea LR. 1996.** Separating the effects of litter quality and microenvironment on  
531 decomposition rates in a patterned peatland. *Oikos* **77**: 529-539.
- 532 **Bengtsson F, Granath G, Rydin H. 2016.** Photosynthesis, growth, and decay traits in  
533 *Sphagnum* - a multispecies comparison. *Ecology and Evolution* **6**: 3325-3341.
- 534 **Berg A, Danielsson A, Svensson BH. 2013.** Transfer of fixed-N from N<sub>2</sub>-fixing  
535 cyanobacteria associated with the moss *Sphagnum riparium* results in enhanced  
536 growth of the moss. *Plant and Soil* **362**: 271-278.
- 537 **Blomberg SP, Garland T. 2002.** Tempo and mode in evolution: phylogenetic inertia,  
538 adaptation and comparative methods. *Journal of Evolutionary Biology* **15**: 899-

539 910.

540 **Bragazza L, Freeman C, Jones T, Rydin H, Limpens J, Fenner N, Ellis T, Gerdol R,**  
541 **Hajek M, Hajek T, et al. 2006.** Atmospheric nitrogen deposition promotes  
542 carbon loss from peat bogs. *Proceedings of the National Academy of Sciences of*  
543 *the United States of America* **103**: 19386-19389.

544 **Bragina A, Berg C, Cardinale M, Shcherbakov A, Chebotar V, Berg G. 2012.**  
545 Sphagnum mosses harbour highly specific bacterial diversity during their whole  
546 lifecycle. *ISME Journal* **6**: 802-813.

547 **Cardona-Correa C, Piotrowski MJ, Knack JJ, Kodner RE, Geary DH, Graham LE.**  
548 **2016.** Peat moss-like vegetative remains from Ordovician carbonates.  
549 *International Journal of Plant Sciences* **177**: 523-538.

550 **Chitwood DH, Topp CN. 2015.** Revealing plant cryptotypes: defining meaningful  
551 phenotypes among infinite traits. *Current Opinion in Plant Biology* **24**: 54-60.

552 **Clymo RS. 1963.** Ion exchange in Sphagnum and its relation to bog ecology. *Annals of*  
553 *Botany* **27**: 309-&.

554 **Clymo RS. 1983.** Peat. In: Gore AJP, ed. *Ecosystems of the World, Vol. 4A, Mires:*  
555 *Swamp, Bog, Fen and Moor.* New York: Elsevier, 159-224.

556 **Clymo RS, Hayward PM. 1982.** *The ecology of Sphagnum.* New York: Chapman and  
557 Hall.

558 **Cronberg N. 1996.** Clonal structure and fertility in a sympatric population of the peat  
559 mosses, *Sphagnum rubellum* and *S. capillifolium*. *Canad. J. Bot.* **74**: 1375-1385.

560 **Crum HA. 1984.** *North American flora, series II, part 11, Sphagnopsida, Sphagnaceae.*  
561 New York: New York Botanical Garden.

562 **Devos N, Szovenyi P, Weston DJ, Rothfels CJ, Johnson MG, Shaw AJ. 2016.**  
563 Analyses of transcriptome sequences reveal multiple ancient large-scale  
564 duplication events in the ancestor of Sphagnopsida (Bryophyta). *New Phytologist*  
565 **211**: 300-318.

566 **Freeman C, Ostle N, Kang H. 2001.** An enzymic 'latch' on a global carbon store. *Nature*  
567 **409**: 149.

568 **Greb SF, DiMichele WA, Gastaldo RA. 2006.** Evolution and importance of wetlands in  
569 Earth history. In Greb SF, DiMichele WA, eds. *Wetlands through time.*  
570 Geological Society of America Special Paper 399,. 1-40.

571 **Hájek T, Ballance S, Limpens J, Zijlstra M, Verhoeven JTA. 2011.** Cell-wall  
572 polysaccharides play an important role in decay resistance of Sphagnum and  
573 actively depressed decomposition in vitro. *Biogeochemistry* **103**: 45-57.

574 **Hájek T. 2009.** Habitat and species controls on Sphagnum production and decomposition  
575 in a mountain raised bog. *Boreal Environment Research* **14**: 947-958.

576 **Haudry A, Platts AE, Vello E, Hoen DR, Leclercq M, Williamson RJ, Forczek E,**  
577 **Joly-Lopez Z, Steffen JG, Hazzouri KM, et al. 2013.** An atlas of over 90,000  
578 conserved noncoding sequences provides insight into crucifer regulatory regions.  
579 *Nature Genetics* **45**: 891-U228.

580 **Hayward PM, Clymo RS. 1983.** The growth of *Sphagnum*: experiments on, and  
581 simulation of, some effects of light flux and water-table depth. *Journal of Ecology*  
582 **71**: 845-863.

583 **Ho A, Bodelier PLE. 2015.** Diazotrophic methanotrophs in peatlands: the missing link?  
584 *Plant and Soil* **389**: 419-423.

585 **Ingram HAP. 1978.** Soil layers in mires: function and terminology. *Journal of Soil*  
586 *Science* **29**: 224-227.

587 **Jansson R, Rodriguez-Castaneda G, Harding LE. 2013.** What can multiple  
588 phylogenies say about the latitudinal diversity gradient? A new look at the  
589 tropical conservatism, out of the tropics, and diversification rate hypotheses.  
590 *Evolution* **67**: 1741-1755.

591 **Jassey VE, Signarbieux C, Hättenschwiler S, Bragazza L, Buttler A, Delarue F,**  
592 **Fournier B, Gilbert D, Laggoun-Défarge F, Lara E, Mills RT. 2015.** An  
593 unexpected role for mixotrophs in the response of peatland carbon cycling to  
594 climate warming. *Scientific reports* **5**

595 **Johnson MG, Granath G, Tahvanainen T, Pouliot R, Stenoien HK, Rochefort L,**  
596 **Rydin H, Shaw AJ. 2015.** Evolution of niche preference in Sphagnum peat  
597 mosses. *Evolution* **69**: 90-103.

598 **Jones CG, John HL, Moshe S. 1994. Organisms as ecosystem engineers. In:**  
599 ***Ecosystem management*. New York: Springer, 130-147.**

600 **Karlin EF, Andrus RE, Boles SB, Shaw AJ. 2011.** One haploid parent contributes  
601 100% of the gene pool for a widespread species in northwest North America.  
602 *Molecular Ecology* **20**: 753-767.

603 **Kip N, van Winden JF, Pan Y, Bodrossy L, Reichart GJ, Smolders AJP, Jetten**  
604 **MSM, Damste JSS, Op den Camp HJM. 2010.** Global prevalence of methane  
605 oxidation by symbiotic bacteria in peat-moss ecosystems. *Nature Geoscience* **3**:  
606 617-621.

607 **Kostka JE, Weston DJ, Glass JB, Lilleskov EA, Shaw AJ, Turetsky MR. 2016.** The  
608 Sphagnum microbiome: new insights from an ancient plant lineage. *New*  
609 *Phytologist* **211**: 57-64.

610 **Kyrkjeeide MO, Hassel K, Flatberg KI, Shaw AJ, Brochmann C, Stenoien HK.**  
611 **2016.** Long-distance dispersal and barriers shape genetic structure of peatmosses  
612 (Sphagnum) across the Northern Hemisphere. *Journal of Biogeography* **43**: 1215-  
613 1226.

614 **Larmola T, Leppänen SM, Tuittila ES, Aarva M, Merilä P, Fritze H, Tirola M.**  
615 **2014.** Methanotrophy induces nitrogen fixation during peatland development.  
616 *Proceedings of the National Academy of Sciences* **111**: 734-739.

617 **Limpens J, Bohlin E, Nilsson MB. 2017.** Phylogenetic or environmental control on the  
618 elemental and organo-chemical composition of Sphagnum mosses? *Plant and Soil*  
619 **417**: 69-85.

620 **Lindo Z, Nilsson MC, Gundale MJ. 2013.** Bryophyte-cyanobacteria associations as  
621 regulators of the northern latitude carbon balance in response to global change.  
622 *Global Change Biology* **19**: 2022-2035.

623 **Merilä J, Hendry AP. 2014.** Climate change, adaptation, and phenotypic plasticity: the  
624 problem and the evidence. *Evolutionary Applications* **7**: 1-14.

625 **Miner BG, Sultan SE, Morgan SG, Padilla DK, Relyea RA. 2005.** Ecological  
626 consequences of phenotypic plasticity. *Trends in Ecology & Evolution* **20**: 685-  
627 692.

628 **Novikova PY, Hohmann N, Nizhynska V, Tsuchimatsu T, Ali J, Muir G,**  
629 **Guggisberg A, Paape T, Schmid K, Fedorenko OM, et al. 2016.** Sequencing of  
630 the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation

631 and abundant trans-specific polymorphism. *Nature Genetics* **48**: 1077-+.

632 **Pease JB, Haak DC, Hahn MW, Moyle LC. 2016.** Phylogenomics reveals three sources  
633 of adaptive variation during a rapid radiation. *Plos Biology* **14**.

634 **Raghoebarsing AA, Smolders AJP, Schmid MC, Rijpstra WIC, Wolters-Arts M,**  
635 **Derksen J, Jetten MSM, Schouten S, Damste JSS, Lamers LPM, et al. 2005.**  
636 Methanotrophic symbionts provide carbon for photosynthesis in peat bogs. *Nature*  
637 **436**: 1153-1156.

638 **Reuning GA, Bauerle WL, Mullen JL, McKay JK. 2015.** Combining quantitative trait  
639 loci analysis with physiological models to predict genotype-specific transpiration  
640 rates. *Plant Cell and Environment* **38**: 710-717.

641 **Revell LJ. 2009.** Size-correction and principal components for interspecific comparative  
642 studies. *Evolution* **63**: 3258-3268.

643 **Rice SK, Giles L. 1996.** The influence of water content and leaf anatomy on carbon  
644 isotope discrimination and photosynthesis in Sphagnum. *Plant Cell and*  
645 *Environment* **19**: 118-124.

646 **Rydin H, Jeglum JK. 2013.** *The Biology of Peatlands*, 2<sup>nd</sup> ed. Oxford, UK: Oxford  
647 University Press.

648 **Rockstrom J, Steffen W, Noone K, Persson A, Chapin FS, Lambin EF, Lenton TM,**  
649 **Scheffer M, Folke C, Schellnhuber HJ, et al. 2009.** A safe operating space for  
650 humanity. *Nature* **461**: 472-475.

651 **Scheiter S, Langan L, Higgins SI. 2013.** Next-generation dynamic global vegetation  
652 models: learning from community ecology. *New Phytologist* **198**: 957-969.

653 **Shaw AJ, Devos N, Cox CJ, Boles SB, Shaw B, Buchanan AM, Cave L, Seppelt R.**  
654 **2010.** Peatmoss (Sphagnum) diversification associated with Miocene Northern  
655 Hemisphere climatic cooling? *Molecular Phylogenetics and Evolution* **55**: 1139-  
656 1145.

657 **Shaw AJ, Devos N, Liu Y, Cox CJ, Goffinet B, Flatberg KI, Shaw B. 2016a.**  
658 Organellar phylogenomics of an emerging model system: Sphagnum (peatmoss).  
659 *Annals of Botany* **118**: 185-196.

660 **Shaw AJ, Schmutz J, Devos N, Shu S, Carrell AA, Weston DJ 2016b.** The Sphagnum  
661 Genome Project: A new model for ecological and evolutionary genomics. In:  
662 Rensing SA, ed. *Genomes and evolution of charophytes, bryophytes, lycophytes*  
663 *and ferns. Advances in Botanical Research* **78**: 167-187.

664 **Stalheim T, Ballance S, Christensen BE, Granum PE. 2009.** Sphagnan - a pectin-like  
665 polymer isolated from Sphagnum moss can inhibit the growth of some typical  
666 food spoilage and food poisoning bacteria by lowering the pH. *Journal of Applied*  
667 *Microbiology* **106**: 967-976.

668 **Stevenson SR, Kamisugi Y, Trinh CH, Schmutz J, Jenkins JW, Grimwood J,**  
669 **Wellington M, Tuskan GA, Rensing SA, Lang D, et al. 2016.** Genetic analysis  
670 of *Physcomitrella patens* identifies ABCISIC ACID NON-RESPONSIVE  
671 (ANR), a regulator of ABA responses unique to basal land plants and required for  
672 desiccation tolerance. *The Plant Cell* tpc-00091.

673 **Turetsky MR, Bond-Lamberty B, Euskirchen E, Talbot J, Frohling S, McGuire AD,**  
674 **Tuittila ES. 2012.** The resilience and functional role of moss in boreal and arctic  
675 ecosystems. *New Phytologist* **196**: 49-67.

676 **Turetsky MR, Crow SE, Evans RJ, Vitt DH, Wieder RK. 2008.** Trade-offs in resource

677 allocation among moss species control decomposition in boreal peatlands. *Journal*  
678 *of Ecology* **96**: 1297-1305.

679 **van Bodegom PM, Douma JC, Witte JPM, Ordonez JC, Bartholomeus RP, Aerts R.**  
680 **2012.** Going beyond limitations of plant functional types when predicting global  
681 ecosystem-atmosphere fluxes: exploring the merits of traits-based approaches.  
682 *Global Ecology and Biogeography* **21**: 625-636.

683 **van Breemen N. 1995.** How Sphagnum bogs down other plants. *Trends in Ecology &*  
684 *Evolution* **10**: 270-275.

685 **Vile MA, Wieder RK, Zivkovic T, Scott KD, Vitt DH, Hartsock JA, Iosue CL,**  
686 **Quinn JC, Petix M, Fillingim HM, et al. 2014.** N<sub>2</sub>-fixation by methanotrophs  
687 sustains carbon and nitrogen accumulation in pristine peatlands. *Biogeochemistry*  
688 **121**: 317-328.

689 **Vitt DH, Slack NG. 1984.** Niche diversification of Sphagnum relative to environmental  
690 factors in northern Minnesota peatlands. *Canadian Journal of Botany-Revue*  
691 *Canadienne De Botanique* **62**: 1409-1430.

692 **Waddington JM, Morris PJ, Kettridge N, Granath G, Thompson DK, Moore PA.**  
693 **2015.** Hydrological feedbacks in northern peatlands. *Ecohydrology* **8**: 113-127.

694 **Weinig C, Ewers BE, Welch SM. 2014.** Ecological genomics and process modeling of  
695 local adaptation to climate. *Current Opinion in Plant Biology* **18**: 66-72.

696 **Weston DJ, Timm CM, Walker AP, Gu LH, Muchero W, Schmutz J, Shaw AJ,**  
697 **Tuskan GA, Warren JM, Wullschleger SD. 2015.** Sphagnum physiology in the  
698 context of changing climate: emergent influences of genomics, modelling and  
699 host-microbiome interactions on understanding ecosystem function. *Plant Cell*  
700 *and Environment* **38**: 1737-1751.

701 **Whitham TG, Bailey JK, Schweitzer JA, Shuster SM, Bangert RK, Leroy CJ,**  
702 **Lonsdorf EV, Allan GJ, DiFazio SP, Potts BM, et al. 2006.** A framework for  
703 community and ecosystem genetics: from genes to ecosystems. *Nature Reviews*  
704 *Genetics* **7**: 510-523.

705 **Whitham TG, Young WP, Martinsen GD, Gehring CA, Schweitzer JA, Shuster SM,**  
706 **Wimp GM, Fischer DG, Bailey JK, Lindroth RL, et al. 2003.** Community and  
707 ecosystem genetics: A consequence of the extended phenotype. *Ecology* **84**: 559-  
708 573.

709 **Yu ZC, Loisel J, Brosseau DP, Beilman DW, Hunt SJ. 2010.** Global peatland  
710 dynamics since the Last Glacial Maximum. *Geophysical Research Letters* **37**:  
711 L13402.

## 713 **Figures Legends**

714

715 **Figure 1:** Morphological traits of *Sphagnum*. Left, four representative species (modified  
716 from Crum, 1984), A. Plant habit showing differences in branch density. B. Branch leaf  
717 cross sections showing arrangements of larger hyaline cells. As in most mosses,  
718 *Sphagnum* leaves consist of a single layer of cells, but unlike in other mosses, the leaf

719 cells are dimorphic, comprising large hyaline cells, dead and empty at maturity,  
720 alternating with narrow photosynthetic chlorophyllose cells. In some species (e.g., top),  
721 those chlorophyllose cells are not exposed at the leaf surface and in other species they are  
722 exposed at the inner or outer surface. C. Surface view of branch leaf cells, showing  
723 variously arranged pores on hyaline cells. The chlorophyllose cells are very narrow,  
724 forming a network around each hyaline cell. D. Branch fascicles, each including so-called  
725 spreading and pendent branches. E. Branch leaf. F. Stem cross section showing  
726 variously developed, sometimes enlarged outer cortex cells. Right, one (haploid)  
727 gametophyte plant with stalked capsules releasing spores (modified from Weston *et al.*,  
728 2015). Far right, detail of branch leaf cells showing differentiation of chlorophyllose and  
729 hyaline cells.

730

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

741

742 **Figure 3:** Schematic of the proposed depth and breadth genetic approaches. In gene-to-  
743 trait studies, linkage-based and association mapping are main approaches used to  
744 discover (or map) the genetic basis of quantitative phenotypic variation. Both assume that  
745 there is variation for the traits of interest within the population being studied. The  
746 linkage-based method relies on individuals with known relationships to each other and  
747 DNA variants (termed genetic markers) that segregate through the population. The  
748 genetic marker is ‘linked’ through proximity to the causal loci and they therefore  
749 segregate together. Association mapping does not require known relationships among

750 individuals within the population, but instead relies on historical recombination from  
751 many generations of random mating. Together these methods constitute the ‘genetic  
752 depth’ approach discussed in text aimed at identifying candidate genes (bottom) that are  
753 then included in phylogenomic and comparative genome analyses (top). These analyses  
754 are simplified by the fact that *Sphagnum* gametophytes are typically haploid. Two  
755 allopolyploid species (*S. palustre*, *S. papillosum*) are included to address subsidiary  
756 issues related to the evolution of polyploid genomes.

757

758

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

766

767 **Acknowledgements:** We thank Drs. Stan Wullschleger, Paul Hanson and two  
768 anonymous reviewers for comments on the manuscript. Work related to sequencing  
769 efforts are supported by the U.S. Department of Energy (DOE) Joint Genome Institute by  
770 the Office of Science under Contract No. DE-AC02-05CH11231; and germplasm  
771 establishment and maintenance is supported by the U.S. DOE, Office of Science, Office  
772 of Biological and Environmental Research, Early Career Research Program. Oak Ridge  
773 National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy  
774 under contract DE-AC05-00OR22725. We thank the National Evolutionary Synthesis  
775 Center (NESCent), NSF #EF-0905606 and the New Phytologist Trust for sponsoring  
776 workshops on the Sphagnum Project.

777

778

779

780